EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN-PH-EP/2004-025
15 June 2004

Observation of the rare decay $K_S \to \pi^0\mu^+\mu^-$

J.R. Batley, G.E. Kalmus[1)], C. Lazzeroni, D.J. Munday, M. Patel, M.W. Slater, S.A. Wotton
Cavendish Laboratory, University of Cambridge, Cambridge, CB3 0HE, U.K.[2)]

R. Arcidiacono, G. Bocquet, A. Ceccucci, D. Cundy[3)], N. Doble[4)], V. Falaleev, L. Gatignon, A. Gonidec, P. Grafström, W. Kubischta, F. Marchetto[5)], I. Mikulec[6)], A. Norton, B. Panzer-Steindel, P. Rubin[7)], H. Wahl[8)]
CERN, CH-1211 Genève 23, Switzerland

E. Monnier[9)], E. Swallow, R. Winston
The Enrico fermi Institute, The University of Chicago, Chicago, Illinois, 60126, U.S.A.

E. Goudzovski, D. Gurev, P. Hristov[10)], V. Kekelidze, V. Kozhuharov, L. Litov, D. Madigozhin, N. Molokanova, Yu. Potrebenikov, S. Stoynev, A. Zinchenko
Joint Institute for Nuclear Research, Dubna, Russian Federation

R. Sacco[11)], A. Walker
Department of Physics and Astronomy, University of Edinburgh, JCMB King's Buildings, Mayfield Road, Edinburgh, EH9 3JZ, U.K.

W. Baldini, P. Dalpiaz, J. Duclos, P.L. Frabetti, A. Gianoli, M. Martini, F. Petrucci, M. Scarpa, M. Savrié
Dipartimento di Fisica dell'Università e Sezione dell'INFN di Ferrara, I-44100 Ferrara, Italy

A. Bizzeti[12)], M. Calvetti, G. Graziani, E. Iacopini, M. Lenti, F. Martelli[13)], G. Ruggiero, M. Veltri[13)]
Dipartimento di Fisica dell'Università e Sezione dell'INFN di Firenze, I-50125 Firenze, Italy

M. Behler, K. Eppard, M. Eppard, A. Hirstius, K. Kleinknecht, U. Koch, L. Masetti, P. Marouelli, U. Moosbrugger, C. Morales Morales, A. Peters, R. Wanke, A. Winhart
Institut für Physik, Universität Mainz, D-55099 Mainz, Germany[14)]

A. Dabrowski, T. Fonseca Martin, M. Szleper, M. Velasco
Department of Physics and Astronomy, Northwestern University, Evanston Illinois 60208-3112, U.S.A.

G. Anzivino, P. Cenci, E. Imbergamo, G. Lamanna, P. Lubrano, A. Michetti, A. Nappi, M. Pepe, M.C. Petrucci, M. Piccini, M. Valdata
Dipartimento di Fisica dell'Università e Sezione dell'INFN di Perugia, I-06100 Perugia, Italy

C. Cerri, G. Collazuol, F. Costantini, R. Fantechi, L. Fiorini, S. Giudici, I. Mannelli, G. Pierazzini, M. Sozzi
Dipartimento di Fisica, Scuola Normale Superiore e Sezione dell'INFN di Pisa, I-56100 Pisa, Italy

C. Cheshkov[10], J.B. Cheze, M. De Beer, P. Debu, G. Gouge, G. Marel, E. Mazzucato, B. Peyaud, B. Vallage
DSM/DAPNIA - CEA Saclay, F-91191 Gif-sur-Yvette, France

M. Holder, A. Maier, M. Ziolkowski
Fachbereich Physik, Universität Siegen, D-57068 Siegen, Germany[15]

C. Biino, N. Cartiglia, M. Clemencic, S. Goy-Lopez, E. Menichetti, N. Pastrone
Dipartimento di Fisica Sperimentale dell'Università e Sezione dell'INFN di Torino, I-10125 Torino, Italy

W. Wislicki
Soltan Institute for Nuclear Studies, Laboratory for High Energy Physics, PL-00-681 Warsaw, Poland[16]

H. Dibon, M. Jeitler, M. Markytan, G. Neuhofer, L. Widhalm
Österreichische Akademie der Wissenschaften, Institut für Hochenergiephysik, A-10560 Wien, Austria[17]

Submitted for publication in Physics Letters B.

---

[1] Present address: Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, UK
[2] Funded by the U.K. Particle Physics and Astronomy Research Council
[3] Present address: Istituto di Cosmogeofisica del CNR di Torino, I-10133 Torino, Italy
[4] Also at Dipartimento di Fisica dell'Università e Sezione dell'INFN di Pisa, I-56100 Pisa, Italy
[5] On leave from Sezione dell'INFN di Torino, I-10125 Torino, Italy
[6] On leave from Österreichische Akademie der Wissenschaften, Institut für Hochenergiephysik, A-1050 Wien, Austria
[7] On leave from University of Richmond, Richmond, VA, 23173, USA; supported in part by the US NSF under award #0140230
[8] Also at Dipartimento di Fisica dell'Università e Sezione dell'INFN di Ferrara, I-44100 Ferrara, Italy
[9] Also at Centre de Physique des Particules de Marseille, IN2P3-CNRS, Université de la Méditerranée, Marseille, France
[10] Present address CERN, CH-1211 Genève 23, Switzerland
[11] Present address Laboratoire de l'Accélerateur Linéaire, IN2P3-CNRS, Université de Paris-Sud, 91898 Orsay, France
[12] Dipartimento di Fisica dell'Università di Modena e Reggio Emilia, via G. Campi 213/A I-41100, Modena, Italy
[13] Istituto di Fisica, Università di Urbino, I-61029 Urbino, Italy
[14] Funded by the German Federal Minister for Research and Technology (BMBF) under contract 7MZ18P(4)-TP2
[15] Funded by the German Federal Minister for Research and Technology (BMBF) under contract 056SI74
[16] Supported by the Committee for Scientific Research grants 5P03B10120, SPUB-M/CERN/P03/DZ210/2000 and SPB/CERN/P03/DZ146/2002
[17] Funded by the Austrian Ministry for Traffic and Research under the contract GZ 616.360/2-IV GZ 616.363/2-VIII, and by the Fonds für Wissenschaft und Forschung FWF Nr. P08929-PHY

Abstract

A search for the decay $K_S \to \pi^0\mu^+\mu^-$ has been made by the NA48/1 Collaboration at the CERN SPS accelerator. The data were collected during 2002 with a high-intensity K_S beam. Six events were found with a background expectation of $0.22^{+0.18}_{-0.11}$ events. Using a vector matrix element and unit form factor, the measured branching ratio is

$$B(K_S \to \pi^0\mu^+\mu^-) = [2.9^{+1.5}_{-1.2}(stat) \pm 0.2(syst)] \times 10^{-9}.$$

1 Introduction

This paper reports the first observation of the decay $K_S \to \pi^0\mu^+\mu^-$ and a measurement of its branching ratio. The analysis was carried out on data taken during 2002 by the NA48/1 experiment at the CERN SPS, which also recently reported the first observation of the decay $K_S \to \pi^0 e^+ e^-$ [1].

The physics interest of the $K_S \to \pi^0\mu^+\mu^-$ decay is that it measures the indirect CP violating contribution of the decay $K_L \to \pi^0\mu^+\mu^-$, thereby allowing the direct CP violating component of the K_L decay to be extracted. This can provide input to the determination of the imaginary part of the element V_{td} of the Cabibbo, Kobayashi, Maskawa (CKM) mixing matrix. In addition, the decay $K_S \to \pi^0\mu^+\mu^-$ can be used to study the structure of the $K \to \pi\gamma^*$ form factor. This is particularly interesting if combined with the $K_S \to \pi^0 e^+ e^-$ results since both decays are expected to be dominated by the exchange of a single virtual photon ($K \to \pi\gamma^* \to \pi\ell^+\ell^-$).

2 Experimental Setup

The beam line and detector built by the NA48 Collaboration to measure the $Re(\epsilon'/\epsilon)$ parameter [2] from a comparison of $K_{S,L} \to \pi^+\pi^-, \pi^0\pi^0$ decays was used, with the modifications described below.

2.1 Beam

The experiment was performed at the CERN SPS accelerator and used a 400 GeV proton beam impinging on a Be target to produce a neutral beam. The spill length was 4.8 s out of a 16.8 s cycle time. The proton intensity was fairly constant during the spill with a mean of 5×10^{10} protons per spill.

The neutral kaon beam line was modified as follows: the K_L beam was blocked and a small, additional sweeping magnet was installed above the K_S collimator. In order to reduce the number of photons in the neutral beam, primarily from π^0 decays, a platinum absorber 24 mm thick was placed in the beam between the target and the main sweeping magnet which deflected charged particles into the collimator. The 5.1 m long K_S collimator, with its axis at an angle of 4.2 mrad to the proton beam direction, selected a beam of neutral long-lived particles (K_S, K_L, Λ^0, Ξ^0, n, γ's, etc.). On average, there were 2×10^5 K_S decays per spill in the fiducial volume downstream of the collimator, with momenta between 60 and 200 GeV (and a mean energy of $\sim$ 110 GeV). An average of 0.017 K_L decays were expected for every K_S decay within the first three K_S lifetimes from the end of the collimator.

2.2 Detector

In order to minimize interactions of the neutral beam, the collimator was immediately followed by a 90 m long evacuated tank. The tank was terminated by a 0.3% radiation length (X_0) thick Kevlar window, except in a region close to the beam which continued in a vacuum pipe through the centre of the downstream detectors.

2.2.1 Tracking

The tracking was performed with a spectrometer housed in a helium gas volume. It consisted of two drift chambers before and two after a dipole magnet with a horizontal transverse momentum kick of 265 MeV. Each chamber had four views, each of which had two sense wire planes. The resulting space points were typically reconstructed with a resolution of 150 μm in each projection. The spectrometer momentum resolution was parameterised as:

$$\sigma_p/p = 0.48\% \oplus 0.015\% \cdot p,$$

where p is in GeV. This gave a resolution of 3 MeV on the reconstructed kaon mass in $K_S \to \pi^+\pi^-$ decays. The track time resolution was 1.4 ns.

2.2.2 Electromagnetic Calorimetry

The detection and measurement of electromagnetic showers were performed with a 27 X_0 deep liquid krypton calorimeter (LKr). The energy resolution was parameterised as [3]:

$$\sigma(E)/E = 3.2\%/\sqrt{E} \oplus 9\%/E \oplus 0.42\%,$$

where E is in GeV. The calorimeter was subdivided into 13,500 cells of transverse dimension 2 cm $\times$ 2 cm, which resulted in a transverse position resolution better than 1.3 mm for single photons with energy above 20 GeV. The π^0 mass resolution was 0.8 MeV, while the time resolution of the calorimeter for a single shower was better than 300 ps.

2.2.3 Scintillator Detectors and Muon Detector

A scintillator hodoscope was located between the spectrometer and the LKr. It consisted of two planes, segmented in horizontal and vertical strips respectively, with each plane arranged in four quadrants. The time resolution for the hodoscope system was 200 ps. Downstream of the LKr calorimeter was an iron-scintillator sandwich hadron calorimeter (HCAL), followed by muon counters (MUC) which consisted of three planes of plastic scintillators, each shielded by an 80 cm thick iron wall. The first two planes, $M1X$ and $M1Y$, consisted of 25 cm wide horizontal ($M1X$) and vertical ($M1Y$) scintillator strips, with a length of 2.7 m. The third plane, $M2X$, consisted of horizontal strips of width 44.6 cm and was mainly used to measure the efficiency of the $M1X$ and $M1Y$ counters. The central strip in each plane was split with a gap of 21 cm to accommodate the beam pipe. The fiducial volume of the experiment was principally determined by the LKr calorimeter acceptance, together with seven rings of scintillation counters which surrounded the decay volume to veto activity outside this region.

2.2.4 Trigger and Readout

The detector was sampled every 25 ns, and samples were recorded in time windows of 200 ns. The extended time window allowed the rate of accidental activity to be investigated from the time sidebands.

The trigger selection for $K_S \to \pi^0\mu^+\mu^-$ candidates consisted of a first-stage hardware trigger followed by a second-stage software trigger. The hardware trigger, for the first 40% of the data taken, selected events satisfying the following conditions:

- at least one hit in each of the horizontal and vertical planes of the hodoscope, within the same quadrant;
- at least three hit wires in at least three views of DCH1 integrated over 200 ns;
- a track vertex located within 90 m of the end of the collimator;
- no hit in the two ring scintillator counters farthest downstream;
- a two-muon signal ($2\mu_{\rm tight}$) from the muon counters.

The $2\mu_{\rm tight}$ signal required at least two hits in each of the first two muon counter planes ($M1X$ and $M1Y$).

For the last 60% of the data taken, a second, parallel, hardware trigger component was added in which the $2\mu_{\rm tight}$ condition was replaced by the following requirements:

- hadron calorimeter energy less than 10 GeV;
- electromagnetic calorimeter energy greater than 15 GeV;
- a two-muon signal ($2\mu_{\rm loose}$) from the muon counters.

The $2\mu_{\rm loose}$ signal was similar to the $2\mu_{\rm tight}$ signal but allowed one of the first two muon planes ($M1X$ or $M1Y$) to contain only a single hit. The addition of the $2\mu_{\rm loose}$ trigger significantly improved the trigger acceptance for the $K_S \to \pi^0\mu^+\mu^-$ signal, as discussed below in Section 3.4.

The software trigger (for the entire data set) required:

- at least two tracks in the drift chambers that were not associated with energetic (>5 GeV) clusters in the LKr;
- less than 10 GeV in the hadronic calorimeter;
- at least one hit in the first two muon counter planes.

Events that satisfied the trigger conditions were recorded and reprocessed with improved calibrations to obtain the final data sample.

3 Data Analysis

The signal channel $K_S \to \pi^0\mu^+\mu^-$ required the identification of two muons of opposite charge accompanied by two additional clusters (photons) in the LKr. The kaon energy was estimated from the sum of the track momenta and cluster energies and was required to be between 60 and 200 GeV.

The invariant mass, $m_{\mu\mu}$, for the two muons from a $K_S \to \pi^0\mu^+\mu^-$ decay is limited by kinematics to:

$$0.211\,\text{GeV} < m_{\mu\mu} < 0.363\,\text{GeV}$$

where the lower limit corresponds to $m_\mu + m_\mu$, while the upper limit is given by $m_K - m_{\pi^0}$.

The known K_S mass was used as a constraint (see Eq. 1 below) to reconstruct the longitudinal vertex position using information from the charged tracks and the clusters. The invariant mass, $m_{\gamma\gamma}$, of the two photons was reconstructed using the measured positions and energies of the two clusters, assuming this vertex position. The total invariant mass, $m_{\mu\mu\pi}$, was calculated using the vertex reconstructed from the charged tracks alone, with the known π^0 mass imposed as a constraint to improve the kaon mass resolution.

3.1 Signal and Control regions

The similarity between the kinematics of the $K_L \to \pi^0\pi^+\pi^-$ and $K_S \to \pi^0\mu^+\mu^-$ channels allowed the former channel to be used to evaluate the resolutions $\sigma_{m_{\gamma\gamma}}$ and $\sigma_{m_{\mu\mu\pi}}$

on $m_{\gamma\gamma}$ and $m_{\mu\mu\pi}$, respectively. These were measured to be $\sigma_{m_{\gamma\gamma}} = 0.8\,\text{MeV}$ and $\sigma_{m_{\mu\mu\pi}} = 3\,\text{MeV}$, in agreement with a detailed Monte Carlo simulation based on GEANT [4].

A signal region and a control region were defined in the $(m_{\gamma\gamma}, m_{\mu\mu\pi})$ plane:

- *SIGNAL REGION:*
  $|m_{\gamma\gamma} - m_{\pi^0}| \leq 2.5\,\sigma_{m_{\gamma\gamma}}$ and $|m_{\mu\mu\pi} - m_K| \leq 2.5\,\sigma_{m_{\mu\mu\pi}}$;
- *CONTROL REGION:*
  $3\,\sigma_{m_{\gamma\gamma}} \leq |m_{\gamma\gamma} - m_{\pi^0}| \leq 6\,\sigma_{m_{\gamma\gamma}}$ and $3\,\sigma_{m_{\mu\mu\pi}} \leq |m_{\mu\mu\pi} - m_K| \leq 6\,\sigma_{m_{\mu\mu\pi}}$.

The signal and control regions were kept masked until the analysis cuts were finalised in order to keep the event selection unbiased. The cuts to reject the background were set using both data and Monte Carlo simulation. An extended control region was used to estimate the contributions from various backgrounds.

3.2 Event selection

In order to identify muons, charged tracks reconstructed in the spectrometer were extrapolated to the MUC planes and associated with MUC hits using cuts on the spatial separation and time difference between the extrapolated track and the MUC hit. Multiple scattering was taken into account before applying the spatial cut, and light propagation along the MUC strips was taken into account before applying the time-difference cut. For muons with momentum greater than 10 GeV, the efficiency of the MUC counters was found to be 0.99 ± 0.01.

Track pairs reconstructed in the spectrometer were classified as muons if the following conditions were satisfied:

- either no LKr cluster was associated to each track, or the energy of any associated cluster was less than 1.5 GeV. A cluster was associated with a track if the separation of the cluster and the track was less than 6 cm and the absolute time difference between the track time and the cluster time was less than 4.5 ns;
- each track had associated hits in the first two planes ($M1X$ and $M1Y$) of the MUC;
- the total energy in the hadron calorimeter was less than 10 GeV;
- for each track, the time measured by either the drift chambers or the trigger hodoscope and the time measured by the muon detector did not differ by more than 4.5 ns. In about 95% of the events the track time was given by the trigger hodoscope, while for the remaining events the time was given by the drift chambers;
- the average of the track times and the average time in the muon detector did not differ by more than 3 ns;
- the tracks had opposite charge;
- the individual track momenta were above 10 GeV in order to have high muon identification efficiency;
- the transverse momentum with respect to the kaon line of flight of each track was below 0.180 GeV;
- the distance between the impact points of the two tracks at the first drift chamber was greater than 10 cm in order to have high trigger vertex reconstruction efficiency;
- the ratio of the track momenta was larger than 1/3 or smaller than 3 in order to reduce the background from Λ and $\bar{\Lambda}$ decays coming either directly from the target or from $\Xi^0(\bar{\Xi}^0) \to \Lambda(\bar{\Lambda})\pi^0$ decays;
- the reconstructed invariant mass assuming $\Lambda \to p\pi^-$ (or $\bar{\Lambda}$) was not compatible with the $\Lambda(\bar{\Lambda})$ mass ($1.1115\,\text{GeV} < m_{p\pi} < 1.1200\,\text{GeV}$);
- $0.21\,\text{GeV} < m_{\mu\mu} < 0.36\,\text{GeV}$.

Clusters reconstructed in the LKr calorimeter were classified as photons if the following conditions were satisfied:

- the energy of the cluster was greater than 3 GeV and less than 100 GeV;
- the distance to the nearest cluster was greater than 10 cm in order to minimize the effect of energy sharing on cluster reconstruction;
- the distance to the nearest extrapolated track was greater than 20 cm.

In addition, a π^0 candidate satisfied the following conditions:

- the total energy of the clusters was greater than 20 GeV;
- the transverse momentum with respect to the kaon line of flight of the π^0 candidate was between 0.05 and 0.25 GeV.

Three quantities related to the decay vertex were computed:

- *total vertex,* z_K.
  To compute the longitudinal vertex position, z_K, the kaon mass was assumed and the kinematical information from all the particles involved in the decay was used. The distance from the vertex to the LKr calorimeter, d_K, was calculated as:

$$d_K = \sqrt{\frac{\sum_{i\neq j} p_i p_j d_{ij}^2}{m_K^2 - 2m_\mu^2 - 2\sum_{i\neq j}(E_i E_j - p_i p_j)}} \tag{1}$$

  where d_{ij} is the separation between the i'th and j'th particles in the LKr plane after linearly extrapolating charged tracks from before the spectrometer magnet, and E_i and p_i are the energy and the momentum of the i'th particle. The vertex is given by $z_K = z$ (LKr position) $- d_K$. This formula follows from energy and momentum conservation applied to the decay $K^0 \to \mu^+\mu^-\gamma\gamma$. The approximation $\cos\alpha_{ij} \sim 1 - \frac{1}{2}(d_{ij}/d_K)^2$ was made, where α_{ij} is the angle defined between the momenta of each pair of particles.
- π^0 *vertex,* z_{π^0} .
  The π^0 vertex position along the beam direction, z_{π^0}, was computed in a similar way to the total vertex, but used only the two photon clusters and imposed the π^0 mass instead of the kaon mass as a constraint.
- *track vertex,* z_{ch}.
  The track vertex was computed by finding the position of the closest distance of approach (CDA) between the two tracks. The CDA was required to be less than 1.5 cm in order to reject muons from pion decay.

Further cuts were applied:

- on the assumption that the observed event was a kaon decay, its proper lifetime was calculated using the z_K, z_{π^0} and z_{ch} vertices, taking the end of the final collimator as the origin. All three calculated proper lifetimes were required to be between 0 and 3 K_S mean lifetimes.
- the energy-weighted centre-of-gravity (COG) of the event was required to be less than 5 cm in order to further reduce the background from accidental activity. The COG was defined as COG$=\sqrt{(\sum_i E_i x_i)^2 + (\sum_i E_i y_i)^2}/\sum_i E_i$, where x_i and y_i are the coordinates of the i'th particle in the LKr plane after linearly extrapolating charged tracks from before the spectrometer magnet, and E_i is the energy of the i'th particle.
- events with extra tracks within the trigger window or extra clusters in the readout window were rejected.

– the time difference Δt between the average track time and the average time of the LKr clusters forming the π^0 was required to be less than 1.5 ns.
– fiducial cuts: tracks were required to be at least 12 cm from the centre of the drift chambers and clusters in the LKr were required to be at least 15 cm from the centre of the LKr, 11 cm from its outside borders, and 2 cm from any defective cell.

After applying all these cuts, the signal and control regions were unmasked. Six events were found in the signal region, and none in the control region (see Fig. 1). The kinematical quantities for these events are summarised in Table 1.

3.3 Background

There are two categories of background: *physical background* and *accidental background*. Physical background was defined as that due to a single kaon decay, for example $K_L \to \pi^+\pi^-\pi^0$, where the two charged pions decayed into muons. Accidental background was caused by the overlap of particles coming from two separate decays that happened to be in time and faked the signal, for example, an overlap between the decays $K_L \to \pi\mu\nu$, where $\pi \to \mu\nu$, and $K_S \to \pi^0\pi^0$, where two photons missed the detector. The accidental background contained two components: (a) overlapping fragments from different proton interactions in the target; and (b) fragments from associated production (KK or ΛK) due to a single proton interaction. The first component could be estimated through a study of out-of-time sidebands.

3.3.1 Physical backgrounds

Only two sources of physical background were found to give a significant contribution:

(1) $K_{L,S} \to \mu^+\mu^-\gamma\gamma$ decays:

The $K_L \to \mu^+\mu^-\gamma\gamma$ branching ratio has been measured by the KTeV Collaboration to be $(1.0^{+0.8}_{-0.6}) \times 10^{-8}$ [5]. Using Monte Carlo simulation, the acceptance for $K_L \to \mu^+\mu^-\gamma\gamma$ decays in NA48 was found to be 5×10^{-3}. This small value originates from a geometrical acceptance of about 20% together with the low probability that the invariant mass of the two photons is consistent with the π^0 mass. The K_S decay into the same final state was taken into account by assuming equal decay rates for K_S and K_L. The total background was estimated to be $0.04^{+0.04}_{-0.03}$ event.

(2) $K_L \to \pi^+\pi^-\pi^0$ decay, where both pions have decayed in flight:

For this background, the reconstructed total mass $m_{\mu\mu\pi}$ lies below the known kaon mass as a consequence of the missing energy of the neutrinos and the use of the muon mass rather than the pion mass in the reconstruction. For the same reasons, the z vertex position (Eq. 1) computed by imposing the K^0 mass constraint is pushed upstream from the actual vertex position of the $\pi^+\pi^-\pi^0$ decay. The shift in the z vertex position leads to a reconstructed two-photon mass $m_{\gamma\gamma}$ lying above the known π^0 mass. These effects are clearly seen in Fig. 2, where the $m_{\gamma\gamma}$ and $m_{\mu\mu\pi}$ distributions are compared for data and Monte Carlo, after applying all cuts and after removing the K_S proper lifetime cuts calculated from z_K. The Monte Carlo gives a good description of the data.

An extended control region was defined in the $(m_{\mu\mu\pi}, m_{\gamma\gamma})$ plane, between 6σ and 12σ around the π^0 and K_S masses. No data event was found in this region. A Monte Carlo sample equivalent to 24 times the 2002 data was used to estimate the background. The cuts on the K_S lifetime, on the transverse momenta of the muons and the pion and on the total energy of the two photons were relaxed, giving 11 Monte Carlo events in the extended control region and no event in the signal region. This led to a background

estimate of $< 2.44/24 = 0.10$ at 90%CL. It was checked that the transverse momentum and photon energy cuts did not significantly affect the shape of the relevant background distributions. After re-applying all cuts except the K_S lifetime cut, 2 of the 11 Monte Carlo events remained in the extended control region. This led to an additional extrapolation factor of $2/11 = 0.18$, giving an overall background estimate for this mode of < 0.018 event at 90%CL.

3.3.2 Accidental backgrounds

Accidental background was studied using data with relaxed timing requirements. As a consequence of the trigger system, an asymmetric timing window was used. Events in the time sidebands ($-115\,\text{ns} \leq \Delta t \leq -3\,\text{ns}$) and ($3\,\text{ns} \leq \Delta t \leq 60\,\text{ns}$) were used to extrapolate the background from the out-of-time to the in-time ($-1.5\,\text{ns} \leq \Delta t \leq +1.5\,\text{ns}$) signal region. Δt was defined as the time difference between the average track time and the average time of the LKr clusters forming the π^0. Given the number of K_S decays in the fiducial volume and the spill length, the probability that two events overlap in time was $\sim 1 \times 10^{-4}$. The out-of-time events are shown in Fig. 3. The readout system introduced a non-uniformity in the event time distribution which changed the effective width of the out-of-time window from (175−6) to (125−6) ns. Due to additional triggering effects, events in this window were recorded only for the $2\mu_{\text{tight}}$ component of the trigger. An extrapolation factor of 1.27 was applied to take into account the relative acceptances of the $2\mu_{\text{tight}}$ and $2\mu_{\text{loose}}$ triggers (see Sections 3.4 and 3.5). Six events were observed in the out-of-time signal region, five of them in the later data-taking period when the $2\mu_{\text{loose}}$ trigger was active, leading to an overall accidental background estimate of $((5\times1.27)+1)\times3/(125-6) = 0.18$ event.

Studies have shown that the majority of the accidental background was due to $K_S \to \pi^+\pi^-$ or $K_L \to \pi^\pm\mu^\mp\nu_\mu$ decays which were in time with two photons from a $K \to \pi^0\pi^0$ decay. The background for $m_{\mu\mu} < 0.3\,\text{GeV}$ was dominated by $K_L \to \pi^\pm\mu^\mp\nu_\mu$ decays, while for $0.30\,\text{GeV} < m_{\mu\mu} < 0.36\,\text{GeV}$, the background was dominated by $K_S \to \pi^+\pi^-$ decays. Background where the two charged tracks came from different decays was examined by looking for events where the two tracks had the same charge. No same-sign event was found in the signal region, consistent with the background estimate above. The accidental background from associated production was found to be negligible.

3.3.3 Background summary

The total background was estimated to be $0.22^{+0.18}_{-0.11}$ event in the signal region. The significant background contributions are summarised in Table 2.

3.4 Trigger efficiency

Hardware trigger:

The trigger efficiency was dominated by the geometrical acceptance of the two-muon signal and by the charged vertex requirement. The efficiencies of the other trigger components were measured, but no correction was necessary since they were also present in the normalisation trigger.

$K_L \to \pi^+\pi^-\pi^0$ events were used to measure the efficiency of the hardware coordinate builders and the microprocessors that reconstructed the tracks and determined the vertex of the decay. The efficiency of this trigger component was found to be 0.910 ± 0.001.

The efficiency of the two-muon signal component of the trigger was estimated from $K_L \to \pi^+\pi^-\pi^0$ decays where both charged pions decayed to muons. The kinematics of

these events is very similar to that of $K_S \to \pi^0\mu^+\mu^-$. The efficiencies of the $2\mu_{\rm tight}$ and $2\mu_{\rm loose}$ signals are shown in Fig. 4 as a function of $m_{\mu\mu}$. A large track separation was effectively imposed by the $2\mu_{\rm tight}$ trigger requirement, resulting in a lower trigger efficiency at low $m_{\mu\mu}$. As shown in Fig. 4, the dependence of the trigger efficiency on $m_{\mu\mu}$ was consistent with that expected from Monte Carlo simulation.

The $2\mu_{\rm tight}$ and $2\mu_{\rm loose}$ trigger efficiencies for $K_S \to \pi^0\mu^+\mu^-$ decays were found to be 0.73 ± 0.03 and 0.93 ± 0.02, respectively. Taking into account the relative amount of data recorded with the $2\mu_{\rm tight}$ and $2\mu_{\rm loose}$ triggers, and including the vertex-finder trigger efficiency, the overall hardware trigger efficiency correction was determined to be 0.77 ± 0.02.

Software trigger:
A small fraction of unbiased events was flagged by the software trigger, but not rejected. Using these events, the efficiency for the software trigger described in Section 2.2.4 was determined to be $1.00^{+0.00}_{-0.02}$.

3.5 Signal Acceptance

The signal acceptance was estimated using Monte Carlo simulation. The amplitude for the decay was taken to be a vector matrix element of the form [6]:

$$A[K_S \to \pi^0\mu^+\mu^-] \propto W(z)(p+p_\pi)^\mu \bar{u}_l(p_-)\gamma_\mu v_l(p_+), \tag{2}$$

where p, p_π, p_- and p_+ are the four-momenta of the kaon, pion, muon and anti-muon, respectively, $z = m^2_{\mu\mu}/m^2_K$, and $W(z)$ is a form factor. As explained in [6], the $W(z)$ dependence on z vanishes to lowest order and for this analysis was represented by the first order polynomial $W(z) = a_S + b_S z$.

The response of the LKr and HCAL detectors to muons is not well simulated by the Monte Carlo and the effect of the LKr and HCAL energy cuts on the muon identification efficiency was therefore studied using data. From a special run with a muon beam, the mean 2μ identification efficiency was determined to be 0.96 ± 0.02. The dependence of the trigger efficiency on $m_{\mu\mu}$ was also determined from data, as described in the previous section.

The acceptance as a function of $m_{\mu\mu}$ is shown in Fig. 5(a), before and after applying the muon identification and trigger efficiency corrections (see Table 4). The reduced acceptance at low $m_{\mu\mu}$ after applying these corrections is due to the $2\mu_{\rm tight}$ component of the trigger.

The expected $m_{\mu\mu}$ distribution, taking into account the muon identification and trigger efficiency corrections, is shown in Fig. 5(b) for several values of b_S/a_S, including a unit form factor, $b_S/a_S = 0$, and $b_S/a_S = 0.4$ as suggested by the Vector Meson Dominance (VMD) model. The reconstructed dilepton mass distribution of the six signal events is also shown in Fig. 5(b) and is consistent with expectation.

The dependence of the overall acceptance on the value of b_S/a_S is shown in Fig. 6. The variation in acceptance seen in the region $-5 \lesssim b_S/a_S \lesssim -2$ arises because the form factor develops a minimum within the kinematically allowed range of $m_{\mu\mu}$ (see Fig. 5(b)). The maximum variation in the acceptance over a wide range of b_S/a_S was used to determine the systematic uncertainty.

Taking into account the dependence of the trigger efficiency on $m_{\mu\mu}$ and the muon identification efficiency, the overall acceptance was found to be

$$0.081 \pm 0.002(stat) \pm 0.004(syst) \tag{3}$$

for kaons in the energy range 60 to 200 GeV decaying in the first three K_S lifetimes after the collimator, where a unit form factor has been assumed. For comparison, the acceptance before applying the trigger and muon identification corrections was 0.109 .

3.6 Normalisation.

The K_S flux was estimated using 132 million $K_S \to \pi^+\pi^-$ events from a minimum bias trigger. Only events that had one vertex, two tracks and no additional clusters outside a modified in-time window were used. Events where the charged pions had decayed were not rejected, and an E/p cut of less than 0.95 was used.

The flux was corrected for $\pi^+\pi^-$ acceptance (0.4579 ± 0.0004), software trigger efficiency (0.9989 ± 0.0004), and the E/p cut (0.9932 ± 0.0001), where the errors are statistical only. No correction was needed for the hardware trigger efficiency, for inefficiencies in the chambers and track reconstruction, or for losses due to multi-vertex and accidental activity since these also applied to the signal channel. To estimate the systematic uncertainty, the cuts used to reject accidental activity and extra clusters were varied. A systematic uncertainty was assigned to take into account the 3% variation observed.

After taking into account the prescaling factor of the trigger, the total K_S flux was found to be $(2.50 \pm 0.08) \times 10^{10}$ in the same fiducial volume as used for the signal channel.

4 Result

As shown in Fig. 1, six events were found in the signal region, with a background estimate of $0.22^{+0.18}_{-0.11}$ event. This is the first observation of the $K_S \to \pi^0\mu^+\mu^-$ decay.

The kinematic properties of the six $K_S \to \pi^0\mu^+\mu^-$ candidates were consistent with those expected based on Monte Carlo simulation of the signal. It was also checked that the number of events observed when the principal analysis cuts were relaxed was consistent with the expected background and that there was no accumulation of events close to any of these cuts.

Using the information summarised in Table 3, the $K_S \to \pi^0\mu^+\mu^-$ branching ratio was measured to be:

$$B(K_S \to \pi^0\mu^+\mu^-) = [2.9^{+1.5}_{-1.2}(stat) \pm 0.2(syst)] \times 10^{-9}, \tag{4}$$

where the statistical error was obtained using the method in Ref. [7] and the systematic uncertainty by combining the individual errors in quadrature.

This result is consistent within error with recent predictions based on Chiral Perturbation Theory [6, 8].

5 Discussion

5.1 Test of Chiral Perturbation Theory

Chiral Perturbation Theory (ChPT) can be used to predict the branching ratio for $K_S \to \pi^0\ell^+\ell^-$ and the corresponding dilepton mass spectrum, $m_{\ell\ell}$. The measurement presented here tests these predictions and constrains the parameters of the model.

The $K_S \to \pi^0\ell^+\ell^-$ branching ratios can be expressed as a function of two parameters, a_S and b_S [6]:

$$B(K_S \to \pi^0 e^+ e^-) = \left[0.01 - 0.76a_S - 0.21b_S + 46.5a_S^2 + 12.9a_S b_S + 1.44b_S^2\right] \times 10^{-10} \tag{5}$$

$$B(K_S \to \pi^0\mu^+\mu^-) = \left[0.07 - 4.52a_S - 1.50b_S + 98.7a_S^2 + 57.7a_S b_S + 8.95b_S^2\right] \times 10^{-11} \tag{6}$$

where the total form factor is $W_S(z) = G_F m_K^2(a_S + b_S z) + W_S^{\pi\pi}(z)$, $z = m_{\ell\ell}^2/m_K^2$, m_K is the kaon mass, $m_{\ell\ell}$ is the invariant mass of the two leptons, and $W_S^{\pi\pi}(z)$ is expected to be small. Assuming VMD, which predicts $b_S = 0.4\, a_S$ [6], the value of the parameter $|a_S|$ can be obtained from the measurement of the individual $K_S \to \pi^0\ell^+\ell^-$ branching ratios via the relations [9]

$$B(K_S \to \pi^0 e^+ e^-) \simeq 5.2 \times 10^{-9}\, a_S^2, \quad (7)$$

$$B(K_S \to \pi^0 \mu^+ \mu^-) \simeq 1.2 \times 10^{-9}\, a_S^2. \quad (8)$$

The VMD model gives an estimate of the ratio of the muon to electron branching ratios of 0.23 [6], while the ratio predicted with a unit form factor is 0.21 [10].

Within the VMD model, the value of the parameter $|a_S|$ can be obtained from the measurement of $B(K_S \to \pi^0\mu^+\mu^-)$ using Eq. 8:

$$|a_S|_{\mu\mu} = 1.54^{+0.40}_{-0.32} \pm 0.06.$$

This value is in agreement with the same quantity extracted from the study of $K_S \to \pi^0 e^+ e^-$ in Ref. [1]:

$$|a_S|_{ee} = 1.06^{+0.26}_{-0.21} \pm 0.07.$$

The ratio $B(K_S \to \pi^0\mu^+\mu^-)/B(K_S \to \pi^0 e^+ e^-)$ is found to be $0.49^{+0.35}_{-0.29} \pm 0.08$, in reasonable agreement with the VMD model prediction of 0.23.

The a_S and b_S parameters can be obtained from a combined analysis of $B(K_S \to \pi^0 e^+ e^-)$ and $B(K_S \to \pi^0\mu^+\mu^-)$ using Eqs. 5 and 6. The 68% confidence level contours in the (a_S, b_S) plane derived from the measured decay rates are shown in Fig. 7(a). The VMD prediction $b_S/a_S = 0.4$ falls within both sets of contours.

The constraints on a_S and b_S from the two branching ratio measurements have been combined using a maximum likelihood method, as shown in Fig. 7(b). Two regions in the (a_S, b_S) plane are preferred, namely

$$a_S = -1.6^{+2.1}_{-1.8}, \quad b_S = 10.8^{+5.4}_{-7.7} \quad (9)$$

$$a_S = 1.9^{+1.6}_{-2.4}, \quad b_S = -11.3^{+8.8}_{-4.5}. \quad (10)$$

The limited statistics do not allow a significant determination of b_S or an assessment of the linear dependence of the form factor on z. Table 4 shows the acceptance as a function of dilepton mass to facilitate the comparison of the two decay channels.

5.2 CPV component of $K_L \to \pi^0\mu^+\mu^-$

The measured branching ratio for the decay $K_S \to \pi^0\mu^+\mu^-$ allows the CPV contribution to the branching ratio of the corresponding K_L decay, $K_L \to \pi^0\mu^+\mu^-$, to be predicted as a function of $\mathrm{Im}(\lambda_t)$ to within a sign ambiguity [11]:

$$B(K_L \to \pi^0\mu^+\mu^-)_{\mathrm{CPV}} \times 10^{12} = C_{\mathrm{MIX}} \pm C_{\mathrm{INT}} \left(\frac{\mathrm{Im}(\lambda_t)}{10^{-4}}\right) + C_{\mathrm{DIR}} \left(\frac{\mathrm{Im}(\lambda_t)}{10^{-4}}\right)^2, \quad (11)$$

where

$$C_{\mathrm{MIX}} = 3.1 \times 10^9\, B(K_S \to \pi^0\mu^+\mu^-), \quad C_{\mathrm{INT}} = 4.6 \times 10^4 \sqrt{B(K_S \to \pi^0\mu^+\mu^-)}, \quad C_{\mathrm{DIR}} = 1.0.$$

$C_{\rm INT}$ is the coefficient for the term due to the interference between the direct ($C_{\rm DIR}$) and indirect ($C_{\rm MIX}$) CPV components, and $\lambda_t = V_{td}V_{ts}^*$.

The predicted dependence of $B(K_L \to \pi^0\mu^+\mu^-)_{\rm CPV}$ on $\mathrm{Im}(\lambda_t)$ is shown in Fig. 8. Taking the central value of the measured branching ratio $B(K_S \to \pi^0\mu^+\mu^-)$ and $\mathrm{Im}(\lambda_t) = 1.36 \times 10^{-4}$ [12] gives:

$$B(K_L \to \pi^0\mu^+\mu^-)_{CPV} \times 10^{12} \approx 8.8_{\rm mixing} \pm 3.3_{\rm interference} + 1.8_{\rm direct}. \qquad (12)$$

The predicted dependence for the $\pi^0 e^+e^-$ channel is also shown in Fig. 8.

The CP conserving (CPC) component of $K_L \to \pi^0\ell^+\ell^-$ decays can be constrained using measurements of the decay $K_L \to \pi^0\gamma\gamma$ [13, 14]. A recent analysis based on ChPT obtained the prediction $(5.2 \pm 1.6) \times 10^{-12}$ [11].

Combining the CPV and the CPC components, the central value for the total $K_L \to \pi^0\mu^+\mu^-$ branching ratio is estimated to be 19×10^{-12} or 13×10^{-12}, depending on the sign of the interference term between the direct and the indirect (mixing) amplitudes. This estimate is consistent with the present experimental upper limit on $B(K_L \to \pi^0\mu^+\mu^-)$ of 3.8×10^{-10}(90% CL) [15]. The corresponding central value for the total $K_L \to \pi^0 e^+e^-$ branching ratio was estimated to be 32×10^{-12} or 13×10^{-12} [1].

Acknowledgments

It is a pleasure to thank the technical staff of the participating laboratories, universities and affiliated computing centres for their efforts in the construction of the NA48 apparatus, in the operation of the experiment, and in the processing of the data. We would also like to thank Gino Isidori and John Ellis for helpful discussions.

References

[1] J. R. Batley *et al.*, Physics Letters B576 (2003) 43.
[2] J. R. Batley *et al.*, Physics Letters B544 (2002) 97.
[3] G. Unal, NA48 Collaboration, in: IX International Conference on Calorimetry, October 2000, Annecy, France, hep-ex/0012011.
[4] GEANT Detector Description and Simulation Tool, CERN Program Library Long Write-up W5013 (1994).
[5] A. Alavi-Harati *et al.*, Physical Review D 62 (2000) 112001.
[6] G. D'Ambrosio, G. Ecker, G. Isidori and J. Portoles, Journal of High Energy Physics 08 (1998) 004.
[7] G.J. Feldman and R.D. Cousins, Phys Rev D57 (1998) 3873.
[8] S. Friot, D. Greynat, and E. de Rafael, "Rare Kaon Decays Revisited," hep-ph/0404136.
[9] G. Buchalla, G. D'Ambrosio and G. Isidori, Nuclear Physics B672 (2003) 387.
[10] G. Ecker, A. Pich, and E. de Rafael, Nuclear Physics B291 (1987) 692.
[11] G. Isidori, C. Smith and R. Unterdorfer, "The rare decay $K_L \to \pi^0\mu\mu$ within the SM," hep-ph/0404127.
[12] M. Battaglia *et al.*, "The CKM matrix and the unitarity triangle," hep-ph/0304132.
[13] A. Lai *et al.*, Physics Letters B536 (2002) 229.
[14] A. Alavi-Harati *et al.*, Physical Review Letters 83 (1999) 917.
[15] A. Alavi-Harati *et al.*, Physical Review Letters 84 (2000) 5279.

Event	K_S momentum (GeV)	τ/τ_S	$m_{\ell\ell}$ (GeV)
$K_S \to \pi^0\mu^+\mu^-$			
1	93.1	0.48	0.241
2	117.9	0.71	0.220
3	147.8	1.74	0.226
4	95.3	2.49	0.241
5	112.7	0.64	0.262
6	101.0	0.96	0.226
$K_S \to \pi^0 e^+ e^-$			
1	84.6	0.74	0.291
2	128.2	0.50	0.267
3	114.1	1.02	0.173
4	83.9	2.09	0.272
5	130.8	1.46	0.303
6	121.2	1.49	0.298
7	94.2	1.64	0.253

Table 1: Kinematics of the six events found in the signal region for $K_S \to \pi^0\mu^+\mu^-$ and of the seven events found for $K_S \to \pi^0 e^+ e^-$ [1].

Background Source	Events
$K_L \to \pi^0\pi^+\pi^-$	$0^{+0.02}_{-0.00}$
$K_L \to \mu^+\mu^-\gamma\gamma$	$0.04^{+0.04}_{-0.03}$
Accidentals	$0.18^{+0.18}_{-0.11}$
Total background	$0.22^{+0.18}_{-0.11}$

Table 2: Summary of the background estimate in the signal region.

Total K_S Flux	$(2.50 \pm 0.08) \times 10^{10}$
Acceptance	$0.081 \pm 0.002 \pm 0.004$
Background	$0.22^{+0.18}_{-0.11}$
Events	6

Table 3: Summary of information used to extract the $K_S \to \pi^0\mu^+\mu^-$ branching ratio.

$K_S \to \pi^0\mu^+\mu^-$		$K_S \to \pi^0 e^+ e^-$	
$m_{\mu\mu}$ (GeV)	Acceptance	m_{ee} (GeV)	Acceptance
0.21-0.23	$0.075^{+0.005}_{-0.008}$	0.000-0.0365	0.0809 ± 0.0015
0.23-0.25	$0.072^{+0.005}_{-0.006}$	0.0365-0.073	0.0877 ± 0.0009
0.25-0.27	$0.088^{+0.002}_{-0.003}$	0.073-0.1095	0.0910 ± 0.0007
0.27-0.29	$0.083^{+0.003}_{-0.003}$	0.1095-0.146	0.0872 ± 0.0007
0.29-0.31	$0.084^{+0.004}_{-0.006}$	0.146-0.1825	0.0785 ± 0.0006
0.31-0.33	$0.097^{+0.003}_{-0.014}$	0.1825-0.219	0.0755 ± 0.0006
0.33-0.35	$0.079^{+0.003}_{-0.012}$	0.219-0.2555	0.0703 ± 0.0007
0.35-0.37	$0.017^{+0.002}_{-0.002}$	0.2555-0.292	0.0664 ± 0.0008
		0.292-0.3285	0.0560 ± 0.0010
		0.3285-0.365	0.0430 ± 0.0020

Table 4: Overall acceptance as a function of the dilepton mass for $K_S \to \pi^0\mu^+\mu^-$ and $K_S \to \pi^0 e^+ e^-$ [1].

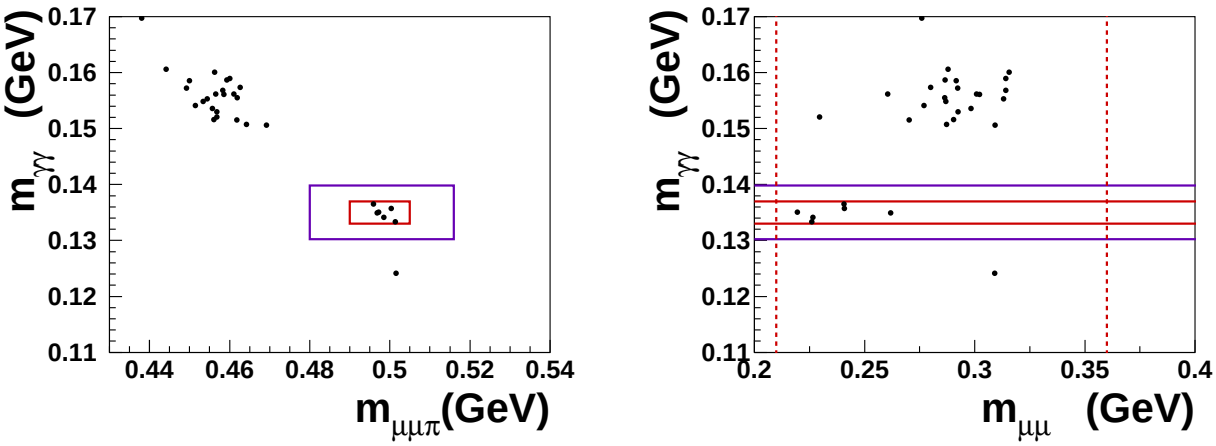

Figure 1: Scatter plot for the events passing all the cuts described in the text: (a) for the $m_{\gamma\gamma}$ versus $m_{\mu\mu\pi}$ plane and (b) for the $m_{\gamma\gamma}$ versus $m_{\mu\mu}$ plane. The 2.5σ and the 6σ signal and control regions and the $m_{\mu\mu}$ kinematic limits are also shown.

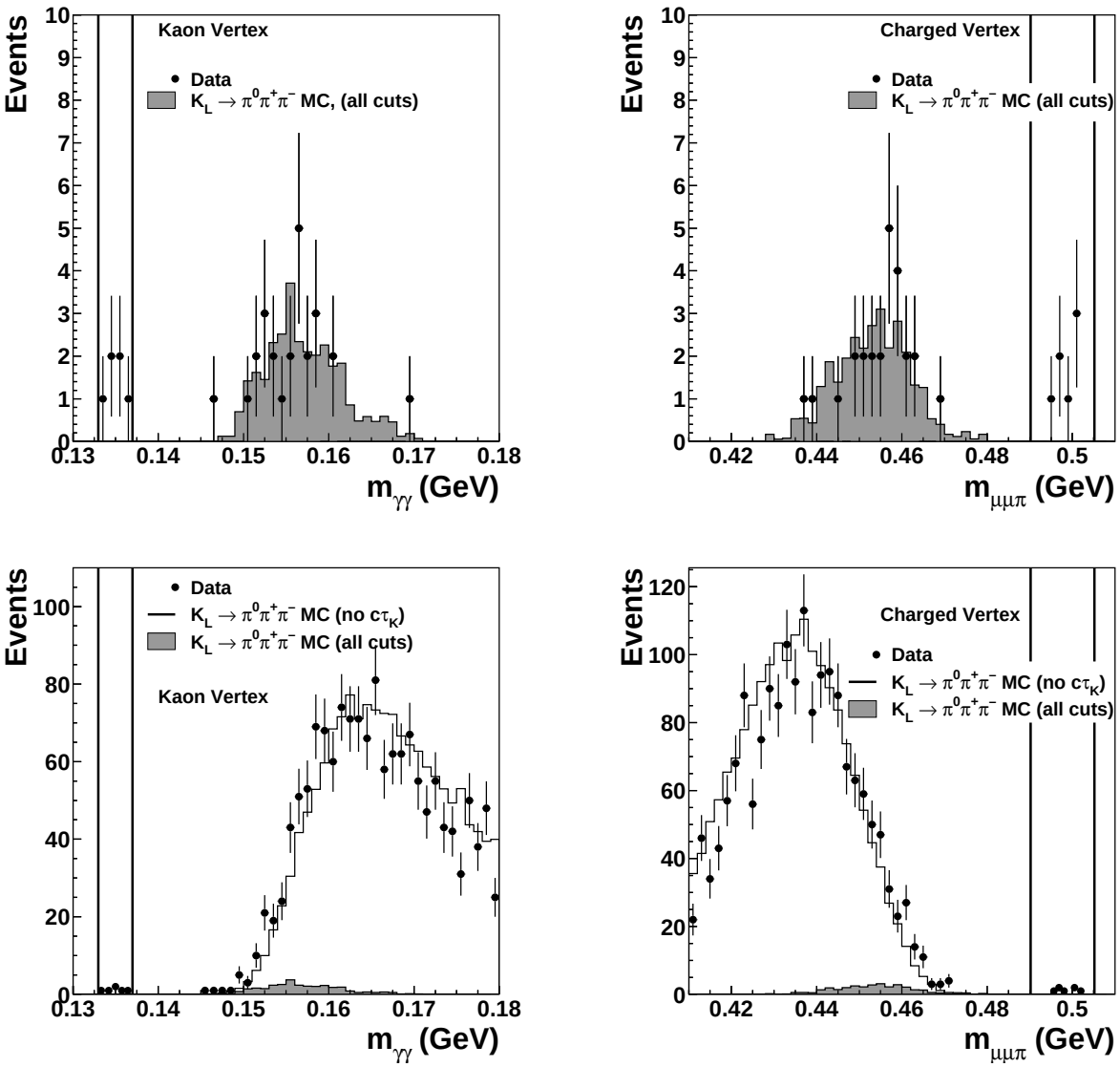

Figure 2: Distributions of $m_{\gamma\gamma}$ and $m_{\mu\mu\pi}$, with the cuts on $m_{\gamma\gamma}$ and $m_{\mu\mu\pi}$ removed, for data (points with error bars) and $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo (histograms). The lower plots do not include the K_S proper lifetime cut calculated from z_K. The vertical lines indicate the $2.5\,\sigma$ signal region.

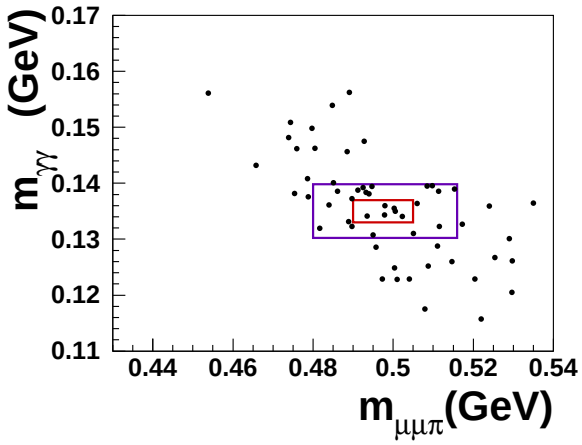

Figure 3: Scatter plot of $m_{\gamma\gamma}$ versus $m_{\mu\mu\pi}$ for events in the out-of-time window. The 2.5σ and the 6σ signal and control regions are also shown.

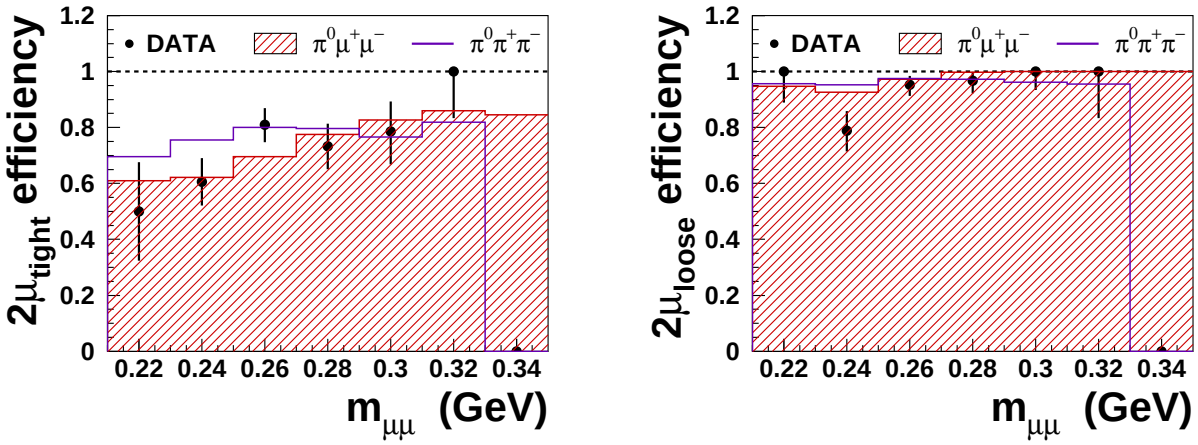

Figure 4: Trigger efficiency measured from data (points with error bars) and estimated from Monte Carlo (histograms) for the two-muon signal component of the trigger, as a function of $m_{\mu\mu}$: (a) for $2\mu_{\rm tight}$, and (b) for $2\mu_{\rm loose}$.

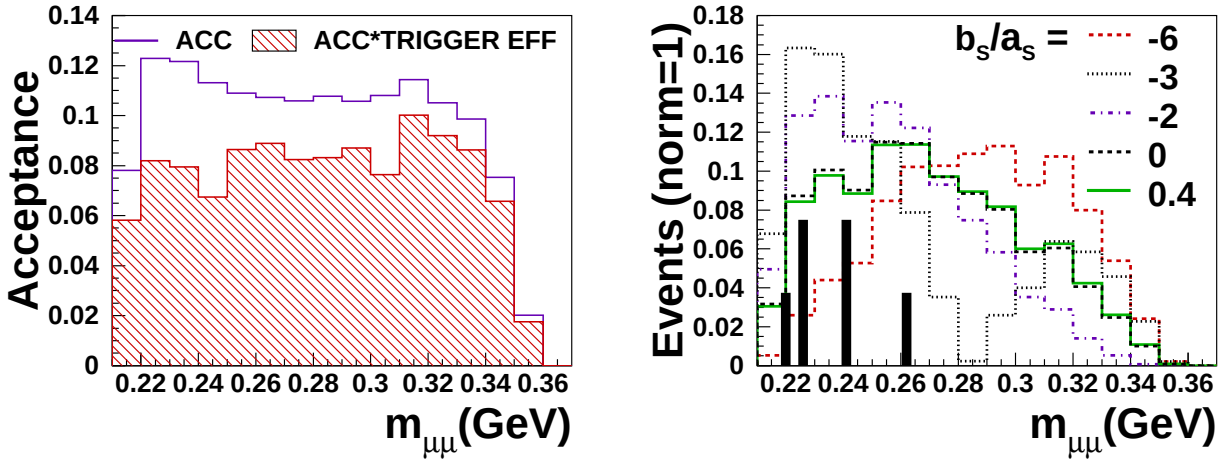

Figure 5: (a) Acceptance as a function of $m_{\mu\mu}$, including (hatched histogram) and not including (open histogram) the trigger efficiency; (b) Expected $m_{\mu\mu}$ distributions of accepted events for several values of b_S/a_S after taking into account the trigger efficiency. The $m_{\mu\mu}$ distribution of the six signal events is also shown.

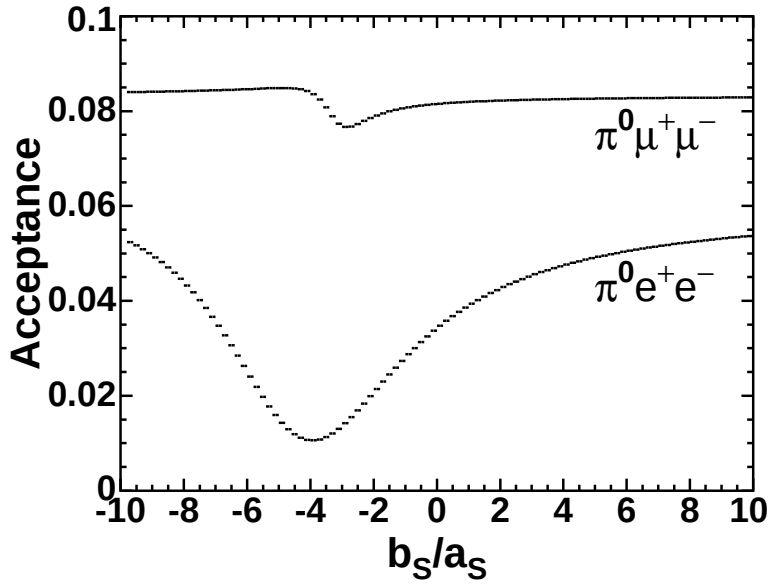

Figure 6: Overall acceptance as a function of b_S/a_S for the $K_S \to \pi^0\mu^+\mu^-$ (upper curve) and $K_S \to \pi^0 e^+ e^-$ (lower curve) channels.

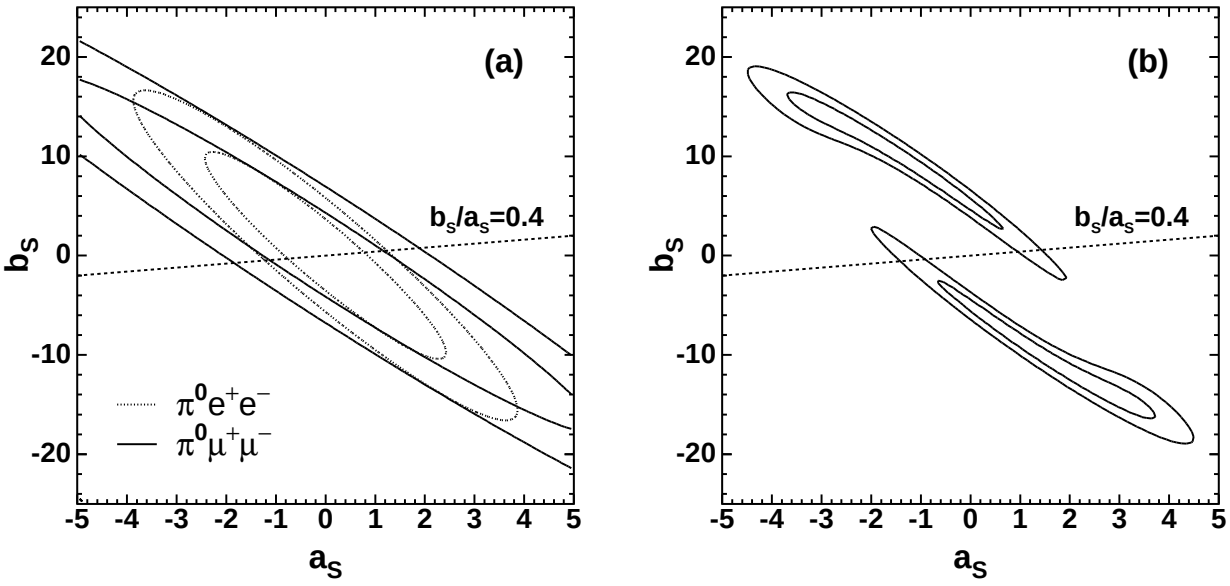

Figure 7: (a) Allowed regions of a_S and b_S determined from the observed number of $K_S \to \pi^0\mu^+\mu^-$ and $K_S \to \pi^0e^+e^-$ events separately. The region between the inner and outer solid (dashed) elliptical contours is the allowed region for $K_S \to \pi^0\mu^+\mu^-$ ($K_S \to \pi^0e^+e^-$) at 68% CL. (b) Allowed regions of a_S and b_S for the $K_S \to \pi^0\mu^+\mu^-$ and $K_S \to \pi^0e^+e^-$ channels combined. The inner (outer) contour of each pair delimits the 1σ (2σ) allowed region from the combined log-likelihood. The dashed straight line in both plots corresponds to $b_S = 0.4a_S$, as predicted by the VMD model.

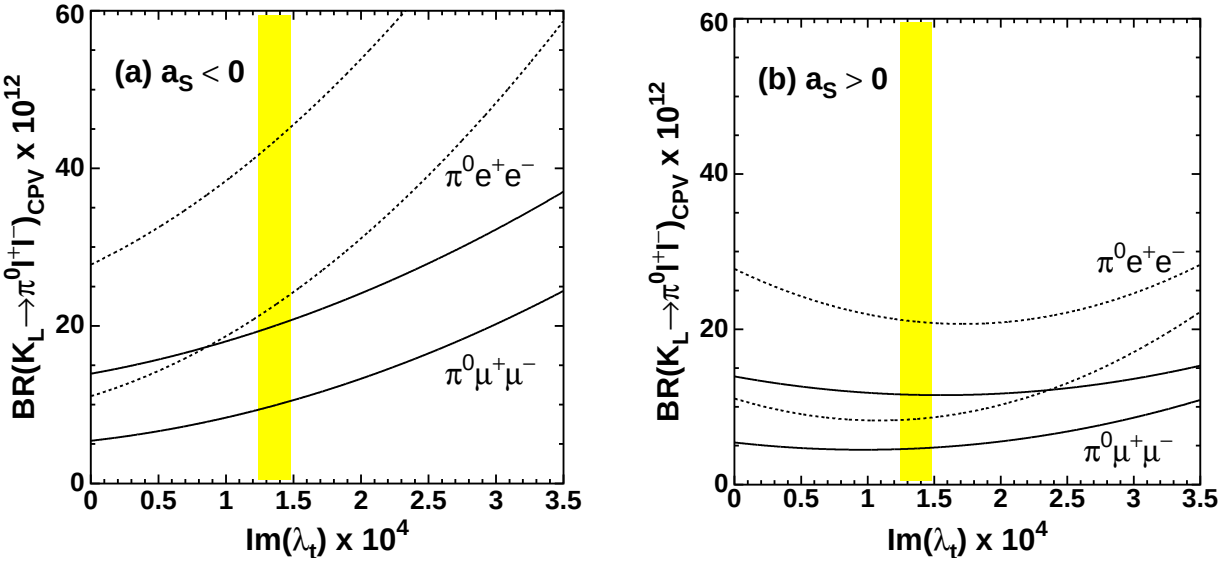

Figure 8: Predicted CPV component of the $K_L \to \pi^0\mu^+\mu^-$ (solid curves) and $K_L \to \pi^0e^+e^-$ (dashed curves) branching ratios as a function of $\mathrm{Im}(\lambda_t)$ assuming (a) $a_S < 0$ and (b) $a_S > 0$. Each pair of curves delimits the allowed range derived from the $\pm 1\sigma$ measured values of $|a_S|$. The vertical shaded band shows the world average value of $\mathrm{Im}(\lambda_t)$ with its uncertainty [12].