First Observation of the Rare Decay Mode $K_L^0 \to e^+ e^-$

D. Ambrose,[1] C. Arroyo,[2] M. Bachman,[3] P. de Cecco,[3] D. Connor,[3] M. Eckhause,[4] K. M. Ecklund,[2] S. Graessle,[1]
A. D. Hancock,[4] K. Hartman,[2] M. Hebert,[2] C. H. Hoff,[4] G. W. Hoffmann,[1] G. M. Irwin,[2] J. R. Kane,[4] N. Kanematsu,[3]
Y. Kuang,[4] K. Lang,[1] R. Lee,[3] R. D. Martin,[4] J. McDonough,[1] A. Milder,[1] W. R. Molzon,[3] M. Pommot-Maia,[2]
P. J. Riley,[1] J. L. Ritchie,[1] P. D. Rubin,[5] V. I. Vassilakopoulos,[1] C. B. Ware,[1] R. E. Welsh,[4] S. G. Wojcicki,[2]
E. Wolin,[4] and S. Worm[1]

(BNL E871 Collaboration)

[1]*University of Texas, Austin, Texas 78712*
[2]*Stanford University, Stanford, California 94305*
[3]*University of California, Irvine, California 92697*
[4]*College of William and Mary, Williamsburg, Virginia 23187*
[5]*University of Richmond, Richmond, Virginia 23173*

In an experiment designed to search for and study very rare two-body decay modes of the K_L^0, we have observed four examples of the decay $K_L^0 \to e^+ e^-$, where the expected background is 0.17 ± 0.10 events. This observation translates into a branching fraction of $8.7^{+5.7}_{-4.1} \times 10^{-12}$, consistent with recent theoretical predictions. This result represents by far the smallest branching fraction yet measured in particle physics. [S0031-9007(98)07665-0]

PACS numbers: 13.20.Eb, 12.15.Mm

We report the first observation of the decay $K_L^0 \to e^+ e^-$, which is strongly suppressed by the GIM mechanism [1] and the helicity structure of the V-A interaction. Thus, it could be sensitive to a new physics process not subject to these effects. The rate of the analogous decay $K_L^0 \to \mu^+ \mu^-$ is dominated by the absorptive contribution which arises from the real two-photon intermediate state, $K_L^0 \to \gamma\gamma \to \mu^+ \mu^-$. The absorptive contribution [2] suggests a $K_L^0 \to e^+ e^-$ branching fraction of 3×10^{-12}. However, recent predictions [3] in the framework of chiral perturbation theory indicate that the two-photon dispersive contribution is larger than the absorptive part in the $K_L^0 \to e^+ e^-$ decay. The predicted branching fraction is approximately 9×10^{-12}. The observation reported here is consistent with these predictions and consequently rules out a significant non-standard-model contribution to this decay.

Data were recorded in 1995 and 1996 by Experiment 871 at the Alternating Gradient Synchrotron (AGS) at Brookhaven National Laboratory (BNL). The experiment built upon experience with BNL E791, which set the best upper limit on $K_L^0 \to e^+ e^-$ [4]. An improved search was made possible by the large increase in AGS intensity with the new Booster and by apparatus modifications to take advantage of higher flux while improving background rejection. Major parts of the spectrometer were more finely segmented to accommodate higher rates. Other important changes were as follows: (a) Redundancy was increased in the most critical areas by making three position measurements in all tracking stations; (b) magnetic fields were set to provide tracks nearly parallel to the beam for two-body decays at the Jacobian peak; and (c) a beam stop [5]

was installed in the first spectrometer magnet to absorb the neutral beam. Taken together, these modifications made possible an improvement of roughly an order of magnitude over the earlier experiment.

The neutral beam was produced by 24 GeV protons on a 1.4 interaction length Pt target at 3.75° with respect to the collimation channel. Typical proton intensity was 1.5×10^{13} per slow spill (1.2–1.6 s duration). This provided about $2 \times 10^8 K_L^0$ per spill ($2 < p_K < 16$ GeV/c) at the entrance of the decay volume, of which 7.5% decayed in the 11 m fiducial length. The experiment is shown in Fig. 1. For brevity, only the detectors most relevant to this analysis will be discussed. These are tracking, electron identification, and trigger hodoscopes.

The tracking section consisted of six chamber stations and two consecutive dipole magnets. The magnets had opposite polarities and provided net transverse momentum kicks of 418 and 216 MeV/c. The topology of a $K_L^0 \to e^+ e^-$ decay is illustrated in Fig. 1. The upstream four tracking stations, where the highest rates occurred, were constructed from 5 mm diameter straw tubes [6] and were operated with a fast gas mixture (50-50 CF_4/C_2H_6). The downstream-most two stations were in a region of lower rates and consisted of drift chambers with 1 cm sense wire spacing operated with a 50-50 Ar/C_2H_6 gas mixture. Each tracking station provided three measurements of position in the horizontal (magnet-bend) x plane and two measurements in the vertical y plane. Since the resolution of the system was limited mainly by multiple Coulomb scattering, the total material was kept to a minimum; including helium bags between the chambers, it amounted to 1.5×10^{-2} radiation lengths. The single
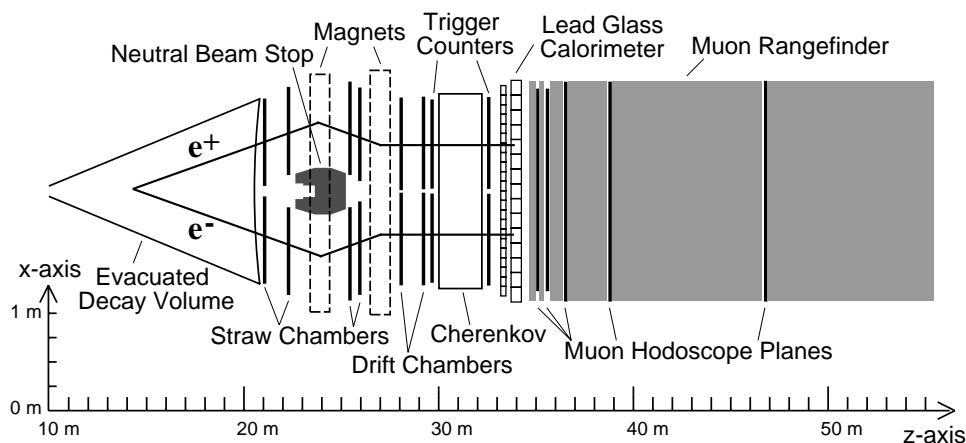
FIG. 1. Plan view of the E871 apparatus. The origin of the z axis is at the target.

wire resolutions were 160 and 120 μm, and the average efficiencies (per wire) were 96% and 98%, for the straw and drift chambers, respectively. The measured mass resolutions for $K_L^0 \to \pi^+\pi^-$ and $K_L^0 \to \mu^+\mu^-$ decay modes were 1.11 and 1.28 MeV/c^2, respectively, compared with the 1.07 and 1.22 MeV/c^2 predicted by Monte Carlo. The $K_L^0 \to e^+e^-$ mass resolution predicted by Monte Carlo is 1.39 MeV/c^2.

Redundant electron identification was achieved by an atmospheric hydrogen threshold Cherenkov counter (CER) and a lead glass array (PBG). The CER had 4×4 arrays of mirror-phototube pairs on each side of the beam. The average photoelectron yield was 5.5 for electrons. The PBG consisted of 216 blocks arranged in two layers, with 3.5 radiation length converter blocks in front of 10.5 radiation length absorber blocks. The measured PBG energy resolution for electrons was $\sigma/E = 0.015 + 0.062/\sqrt{E(\text{GeV})}$. The CER and PBG performance is summarized in Table I.

Two scintillator hodoscopes, one upstream of the CER and one downstream (separated by 2.9 m), were used for triggering. Both hodoscopes had 3.2 cm wide x-measuring slats with phototubes on both ends. The downstream hodoscope had, in addition, 3.0 cm wide y-measuring slats with a phototube on one end. The slats overlapped their nearest neighbors by 3 mm to avoid inefficiency due to cracks.

On-line data selection involved hardware and software triggers. A level 0 (L0) trigger based on the pattern

TABLE I. Performance of particle identification detectors as measured from a clearly identified sample of K_{e3} and $K_{\mu3}$ events. The Cherenkov π and μ rejections are calculated for particles with momenta below their Cherenkov thresholds, 8.3 and 6.3 GeV/c, respectively.

	Cherenkov	Lead glass
e efficiency	0.977 ± 0.001	0.987 ± 0.004
π misidentification	0.0019 ± 0.0002	0.0093 ± 0.0004
μ misidentification	0.0024 ± 0.0002	0.0018 ± 0.0002

of trigger hodoscope hits required two in-time tracks to be parallel within about 30 mrad to, and on opposite sides of, the beam axis. The L0 trigger along with signals from particle identification detectors formed level 1 (L1) triggers for various di-lepton modes. The L1 trigger for e^+e^- required CER hits to be in time and spatially correlated with the trigger hodoscope hits. PBG signals were not required. For events passing a L1 trigger, a software algorithm, run on a farm of eight RISC processors, constituted a level 3 (L3) trigger. The algorithm reconstructed tracks using information from the trigger hodoscopes and all tracking chambers. For e^+e^- events, it required tracks on each side of the spectrometer which formed a vertex in the decay volume. Also, L0 triggers prescaled by a factor of 1000 were recorded for analysis. This minimum bias sample was used to study detector performance and provided data for normalization.

The spectrometer had a geometrical acceptance of 1.89% for $K_L^0 \to e^+e^-$ decays with $9.75 < z < 20.75$ m (see Fig. 1) and kaon momenta between 2 and 16 GeV/c. The trigger requirement of parallelism reduced the acceptance to 1.57%. Analysis criteria further reduced the acceptance to 1.23%.

Potential sources of background include misidentified $K_L^0 \to \pi^\pm e^\mp \nu$ decays, accidental spatial and temporal coincidences of e^+ and e^- from two $K_L^0 \to \pi^\pm e^\mp \nu$ decays, partially measured $K_L^0 \to e^+e^-\gamma$, $K_L^0 \to e^+e^-e^+e^-$, or $K_L^0 \to e^+e^-\gamma\gamma$ decays, and $K_L^0 \to \gamma\gamma$ decays with asymmetric external conversion of both photons in the vacuum window or first straw chamber. The only sources which were not negligible after all analysis criteria were applied are $K_L^0 \to e^+e^-\gamma$ [7] and $K_L^0 \to e^+e^-e^+e^-$ [8] decays. Both of these decays have a low probability of producing an e^+e^- pair with invariant mass near the K_L^0 mass.

Analysis procedures were designed to ensure that selection criteria were not influenced by knowledge of events in or near the signal region. $K_L^0 \to e^+e^-$ events should have measured e^+e^- mass near the K_L^0 mass (497.7 MeV/c^2) and measured transverse momentum (p_T), defined with respect to the parent K_L^0 direction of flight, near zero.

Thus, events with $490 < M_{ee} < 505$ MeV/c^2 and $p_T^2 < 100$ (MeV/c)^2 were excluded from consideration until all selection criteria were finalized.

As a first step in the analysis, events were required to have a good track (with signals from at least two x-measuring wires and one y-measuring wire in each chamber) on each side of the spectrometer. These tracks had to form a good vertex in the decay volume and project to in-time trigger counter hits consistent with the parallelism requirement. Two independent fitting algorithms, which used a full magnetic field map, subsequently determined kinematic quantities for each track. The fitters had different sensitivities to track-finding errors and checked each other's results. Ultimately, all selection criteria were applied to the results of one fitter, with the exception that consistent results for M_{ee} and p_T^2 were required of both.

Events had to satisfy minimal track and vertex quality criteria, be fully contained within the spectrometer, and have $M_{ee} > 475$ MeV/c^2 and $p_T^2 < 900$ (MeV/c)^2 [excluding the region $490 < M_{ee} < 505$ MeV/c^2 and $p_T^2 < 100$ (MeV/c)^2]. In addition, events had to have energy deposition in the PBG characteristic of electrons and good CER pulse height and timing information. The remaining 833 events were studied to determine criteria to further exclude background with minimal acceptance loss. This optimization was based on Monte Carlo simulations and studies of observed $K_L^0 \rightarrow \pi^+ \pi^-$ events. These criteria included better track and vertex quality, tighter timing, a requirement that the momentum asymmetry between the tracks $[|p_{e^+} - p_{e^-}|/(p_{e^+} + p_{e^-})]$ be less than 0.55 (to suppress $K_L^0 \rightarrow \pi^\pm e^\mp \nu$ decays), and a requirement that no additional complete tracks be found in the spectrometer (to suppress events with two K_L^0 decays). Finally, to reduce $K_L^0 \rightarrow e^+ e^- e^+ e^-$ background, criteria were applied to reject events with short partial tracks in the two upstream-most straw chambers pointing to the reconstructed $e^+ e^-$ vertex. The resulting sample comprised 44 events.

After all analysis criteria were finalized, a signal region was defined (elliptical in p_T^2 and M_{ee} and corresponding to about 2.5 sigma in each). As a result of inner bremsstrahlung [9], 23% of $K_L^0 \rightarrow e^+ e^-$ decays fall outside this ellipse. The size of the signal region was chosen to reduce the expected level of background to well below one event. Background estimates were based on Monte Carlo simulations and comparisons to the data. Samples of $K_L^0 \rightarrow e^+ e^- e^+ e^-$ and $K_L^0 \rightarrow e^+ e^- \gamma$ decays were generated [10]. These distributions were absolutely normalized by their measured branching fractions [11] to obtain predictions of 38 ± 8 events from $K_L^0 \rightarrow e^+ e^- \gamma$ and 24 ± 11 events from $K_L^0 \rightarrow e^+ e^- e^+ e^-$ in the region defined by $476 < M_{ee} < 490$ MeV/c^2 and $p_T^2 < 400$ (MeV/c)^2, where 43 events were observed. The uncertainties in the predictions are mainly due to the $K_L^0 \rightarrow e^+ e^- \gamma$ form factor and the fact that the $K_L^0 \rightarrow e^+ e^- e^+ e^-$ efficiency of the partial track cut is not well known. To remove these sources of systematic uncertainty from the estimate of the background in the signal region, we

performed a fit to data to normalize the background distributions, which were then extrapolated into the signal region. The fit in p_T^2 and M_{ee} exploited the fact that in the range $476 < M_{ee} < 490$ MeV/c^2 the p_T^2 distributions for the background decays differ significantly. This procedure yielded estimates of background in the signal region of 0.09 ± 0.11 event from $K_L^0 \rightarrow e^+ e^- e^+ e^-$, 0.08 ± 0.02 event from $K_L^0 \rightarrow e^+ e^- \gamma$, and 0.17 ± 0.10 for the sum of both (to be compared with 0.37 ± 0.14 for the absolutely normalized prediction).

When the full analysis was performed on the excluded region, four $K_L^0 \rightarrow e^+ e^-$ candidates were found inside the signal ellipse, as shown in Fig. 2. These candidates have been carefully scrutinized and exhibit no anomalous features. The probability of observing four background events in the signal region when 0.2 are expected is 6×10^{-5}. A check on the reconstructed mass of $K_L^0 \rightarrow \pi^+ \pi^-$ events collected within minutes of the $K_L^0 \rightarrow e^+ e^-$ candidates rules out transient shifts in the mass scale.

To extract the best estimate of the number of $K_L^0 \rightarrow e^+ e^-$ events, a maximum likelihood fit was performed in the region defined by $476 < M_{ee} < 510$ MeV/c^2 and $p_T^2 < 400$ (MeV/c)^2. In addition to the 50 events inside this region, the inputs to the fit were Monte Carlo distributions in M_{ee} and p_T^2 for $K_L^0 \rightarrow e^+ e^- e^+ e^-$, $K_L^0 \rightarrow e^+ e^- \gamma$, and $K_L^0 \rightarrow e^+ e^-$ decays. The fit estimated the number of events in each distribution, subject to the constraint that the sum equal the observed number of events. Figure 3 shows the fit results for the three distributions versus M_{ee}. The result for the number of $K_L^0 \rightarrow e^+ e^-$ events is $4.20^{+2.69}_{-1.94}$ in the full region of the fit. For comparison, the inset in Fig. 3 shows the M_{ee} distribution of data along with the absolutely normalized Monte Carlo predictions for $K_L^0 \rightarrow e^+ e^- e^+ e^-$ and $K_L^0 \rightarrow e^+ e^- \gamma$.

The $K_L^0 \rightarrow e^+ e^-$ branching fraction B_{ee} was determined from the formula

$$ B_{ee} = N_{ee} \frac{B_{\pi\pi}}{R N_{\pi\pi}} \frac{A_{\pi\pi}}{A_{ee}} \frac{1}{\epsilon_{ee}^{L1}} \frac{1}{\epsilon_{ee}^{L3}} \frac{1}{\epsilon_{ee}^{PID}} f_{\pi\pi} . $$

N_{ee} is the number of $K_L^0 \rightarrow e^+ e^-$ events determined from the likelihood fit. $B_{\pi\pi}$ is the $K_L^0 \rightarrow \pi^+ \pi^-$ branching
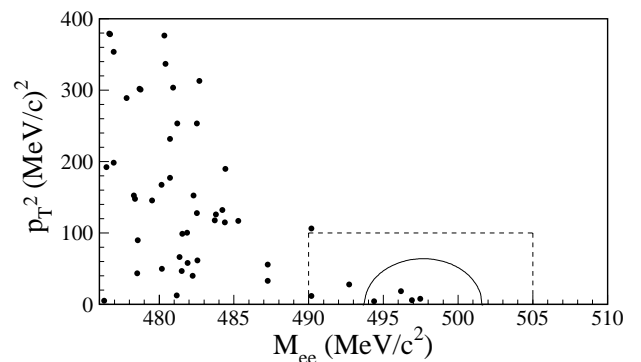

FIG. 2. p_T^2 versus M_{ee} for $K_L^0 \rightarrow e^+ e^-$ candidates. The dashed line shows the exclusion region. The solid curve bounds the signal region.
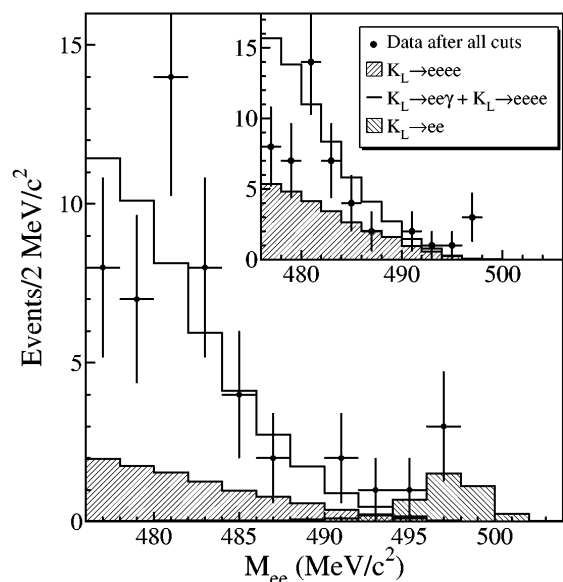
FIG. 3.   Results of the maximum likelihood fit, showing the distributions of $K_L^0 \to e^+e^-$, $K_L^0 \to e^+e^-e^+e^-$, and $K_L^0 \to e^+e^-\gamma$ along with the data versus M_{ee}. The inset shows the absolute Monte Carlo prediction for the background distributions, which have systematic uncertainties described in the text.

fraction [11]. R is the prescale factor for the minimum bias sample used for normalization. It was the product of the hardware prescale (1000) and an additional factor of 20 imposed in software. $N_{\pi\pi}$ is the number of $K_L^0 \to \pi^+\pi^-$ events in this reduced sample (and includes a small correction for the effect of K_S^0 interference). The ratio $A_{\pi\pi}/A_{ee}$, calculated via Monte Carlo, corrects for mode-dependent acceptance differences due to the detector geometry, trigger, and analysis criteria. ϵ_{ee}^{L1} is the L1 ee trigger efficiency with respect to the L0 trigger (which all events had to satisfy). ϵ_{ee}^{L3} is the L3 ee trigger efficiency and was derived from the measured L3 efficiency for $K_L^0 \to \pi^+\pi^-$ by making a small correction to account for the different kinematics of $K_L^0 \to e^+e^-$. ϵ_{ee}^{PID} is the particle identification efficiency and was determined by weighting measured detector efficiencies by the appropriate distributions for $K_L^0 \to e^+e^-$ decays. $f_{\pi\pi}$ accounts for the loss of $K_L^0 \to \pi^+\pi^-$ events due to hadronic interactions in the spectrometer and was based on a GEANT cal-

TABLE II.   Factors entering into the calculation of the $K_L^0 \to e^+e^-$ branching fraction.

Variable	Value
N_{ee}	$4.20^{+2.69}_{-1.94}$
$B_{\pi\pi}$ [11]	$(2.067 \pm 0.035) \times 10^{-3}$
R	$20\,000$
$N_{\pi\pi}$	$83\,531 \pm 381$
$A_{\pi\pi}/A_{ee}$	1.478 ± 0.011
ϵ_{ee}^{L1}	0.977 ± 0.011
ϵ_{ee}^{L3}	0.929 ± 0.003
ϵ_{ee}^{PID}	0.929 ± 0.011
$f_{\pi\pi}$	0.954 ± 0.004

culation. The calculation was checked against special data taken with a single-arm trigger, which provided events in which a pion was lost due to hadronic interaction in the nontrigger arm. The values of these parameters are given in Table II. The calculated branching fraction B_{ee} is $8.7^{+5.7}_{-4.1} \times 10^{-12}$.

In summary, we have made the first observation of the decay $K_L^0 \to e^+e^-$. The measured branching fraction is consistent with recent theoretical predictions [3] and is the smallest measured in particle physics to date.

[1] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970).

[2] L. M. Sehgal, Phys. Rev. **183**, 1511 (1969).

[3] G. Valencia, Nucl. Phys. **B517**, 339 (1998); D. Gomez Dumm and A. Pich, Phys. Rev. Lett. **80**, 4633 (1998).

[4] K. Arisaka et al., Phys. Rev. Lett. **71**, 3910 (1993).

[5] J. Belz et al., e-print 9808037.

[6] S. Graessle et al., Nucl. Instrum. Methods Phys. Res., Sect. A **137**, 138 (1995).

[7] G. D. Barr et al., Phys. Lett. B **240**, 283 (1990); K. E. Ohl et al., Phys. Rev. Lett. **65**, 1407 (1990); W. M. Morse et al., Phys. Rev. D **45**, 36 (1992).

[8] T. Akagi et al., Phys. Rev. D **47**, 2644 (1993); M. R. Vagins et al., Phys. Rev. Lett. **71**, 35 (1993); P. Gu et al., Phys. Rev. Lett. **72**, 3000 (1994); G. D. Barr et al., Z. Phys. C **65**, 361 (1995).

[9] Inner bremsstrahlung is included in the Monte Carlo using the results of D. A. Dicus and W. W. Repko (to be published), which are consistent with L. Bergstrom, Z. Phys. C **20**, 135 (1983) for the e^+e^- mode.

[10] Generation of $K_L^0 \to e^+e^-\gamma$ was based on N. Kroll and W. Wada, Phys. Rev. **98**, 1355 (1955), modified by a form factor from L. Bergstrom, E. Masso, and P. Singer, Phys. Lett. **131B**, 229 (1983), with value $\alpha_{K^*} = -0.28$ (as given by Ref. [11]). Generation of $K_L^0 \to e^+e^-e^+e^-$ was based on T. Miyazaki and E. Takasugi, Phys. Rev. D **8**, 2051 (1973). The introduction of a form factor [according to L. Zhang and J. L. Goity, Phys. Rev. D **57**, 7031 (1998)] had negligible effect on our analysis.

[11] Particle Data Group, Eur. Phys. J. C **3**, 1 (1998).