Measurement of the branching ratio for the rare decay $K_L^0 \to \mu^+\mu^-$

A. P. Heinson,[1] J. Horvath, [2] P. Knibbe,[3] C. Mathiazhagan, W. R. Molzon, and J. Urheim[4]

University of California, Irvine, California 92717

K. Arisaka, R. D. Cousins, T. Kaarsberg,[5] J. Konigsberg,[6] P. Melese,[7] P. Rubin,[8] W. E. Slater, and D. Wagner[9]

University of California, Los Angeles, California 90024

G. W. Hart, W. W. Kinnison, D. M. Lee, R. J. McKee, E. C. Milner,[10] G. H. Sanders, and H. J. Ziock

Los Alamos National Laboratory, Los Alamos, New Mexico 87545

S. Axelrod,[11] K. A. Biery,[12] M. Diwan,[13] G. M. Irwin, K. Lang,[14] J. Margulies,[15] D. A. Ouimette,[16] A. J. Schwartz,[17] and S. G. Wojcicki

Stanford University, Stanford, California 94309

L. B. Auerbach, J. Belz,[18] P. Buchholz,[19] C. Guss,[20] V. L. Highland, S. H. Kettell,[13] W. K. McFarlane,[21] and M. Sivertz[22]

Temple University, Philadelphia, Pennsylvania 19122

G. W. Hoffmann, P. J. Riley, J. L. Ritchie, and A. Yamashita[13]

University of Texas, Austin, Texas 78712

M. D. Chapman,[23] M. Eckhause, J. F. Ginkel,[24] A. D. Hancock, J. R. Kane, C. J. Kenney,[25] Y. Kuang, W. F. Vulcan,[16] R. E. Welsh, R. J. Whyley,[26] R. G. Winter,[27] and M. T. Witkowski[28]

College of William and Mary, Williamsburg, Virginia 23187

(BNL E791 Collaboration)

(Received 29 August 1994)

We have measured the branching ratio for $K_L^0 \to \mu^+\mu^-$ using our full data set obtained during running periods in 1988, 1989, and 1990. The total number of $\mu^+\mu^-$ candidates after a background subtraction is 707, which represents the largest sample to date of this rare decay mode. Our result is $B(K_L^0 \to \mu^+\mu^-) = (6.86 \pm 0.37) \times 10^{-9}$, which is consistent with earlier results and very near the

[1]Present address: University of California, Riverside, CA 92521.
[2]Present address: Purdue University, Lafayette, IN 47907.
[3]Present address: Intermetrics Inc., Warminster, PA 18974.
[4]Present address: California Institute of Technology, Pasadena, CA 91125.
[5]Present address: Sandia National Laboratories, Arlington, VA 22209.
[6]Present address: Harvard University, Cambridge, MA 02138.
[7]Present address: Rockefeller University, New York, NY 10021.
[8]Present address: University of Richmond, Richmond, VA 23173.
[9]Present address: University of Chicago, Chicago, IL 60637.
[10]Present address: College Retirement Equities Fund, New York, NY 10017.
[11]Present address: Measurex Corporation, Cupertino, CA 95014.
[12]Present address: McGill University, Montreal, Canada H3A 2T8.
[13]Present address: BNL, Upton, NY 11973.
[14]Present address: University of Texas, Austin, TX 78712.
[15]Present address: Temple University, Philadelphia, PA 19122.
[16]Present address: CEBAF, Newport News, VA 23606.
[17]Present address: Princeton University, Princeton, NJ 08544.
[18]Present address: Rutgers University, Piscataway, NJ 08855.
[19]Present address: CERN, CH-1211 Geneva 23, Switzerland.
[20]Present address: Cornell University, Ithaca, NY 14853.
[21]Present address: Norfolk State University, Norfolk, VA 23504.
[22]Present address: University of California, San Diego, La Jolla, CA 92093.
[23]Present address: University of New Mexico, Albuquerque, NM 87444.
[24]Present address: University of Colorado, Boulder, CO 80309.
[25]Present address: SLAC, Stanford, CA 94309.
[26]Present address: MCI Communications Corporation, McLean, VA 22102.
[27]Deceased.
[28]Present address: Rensellaer Polytechnic Institute, Troy, NY 12180.

unitarity bound. The upper limit on $B(K_L^0 \to \mu^+\mu^-)$ is used to set an upper limit on the real part of the amplitude $\mathcal{A}(K_L^0 \to \mu^+\mu^-)$; this is subsequently used to constrain the top quark mass m_t and the Wolfenstein parameter ρ of the CKM matrix.

PACS number(s): 13.20.Eb, 12.15.Hh

I. INTRODUCTION

The decay $K_L^0 \to \mu^+\mu^-$ has played a major role in our understanding of weak interactions. The small value of the branching ratio provided strong experimental evidence for the Glashow-Iliopoulos-Maiani (GIM) mechanism of weak decay [1], in which the presence of a second-generation charge-2/3 quark (c) prevents flavor-changing neutral currents from occurring at the tree level, and strongly suppresses them from occurring at the one-loop level (u and c quark contributions cancel). In fact, the GIM mechanism was proposed prior to the discovery of the c quark and allowed its mass to be predicted with surprising accuracy [2].

The small residual rate left after u and c contributions cancel is almost entirely due to the process $K_L^0 \to \gamma\gamma \to \mu^+\mu^-$ where the intermediate photons are real. Using the measured value [3] $B(K_L^0 \to \gamma\gamma) = 5.70 \times 10^{-4}$ and the QED result [4, 5] $\Gamma(K_L^0 \to \gamma\gamma \to \mu^+\mu^-)/\Gamma(K_L^0 \to \gamma\gamma) = 1.20 \times 10^{-5}$, one calculates $B(K_L^0 \to \mu^+\mu^-)_{\gamma\gamma} = 6.83 \times 10^{-9}$. This value, referred to as the "unitarity bound," is the minimum rate at which $K_L^0 \to \mu^+\mu^-$ is expected to occur: other absorptive amplitudes such as those resulting from $\pi\pi$ and 3π intermediate states are small and cannot significantly destructively interfere with $K_L^0 \to \gamma\gamma \to \mu^+\mu^-$ [5, 6].

The rate for $B(K_L^0 \to \mu^+\mu^-)$ also receives contributions from the short-distance electroweak diagrams of Fig. 1. The internal loops within these diagrams are dominated by the heavy t (top) quark. The diagrams together contribute an amplitude proportional to [7, 8]

$$f(x_t)\ |\mathrm{Re}(V_{td}^* V_{ts})|^2\ , \tag{1}$$

where f is a function of $x_t = (m_t/m_W)^2$ and V_{td} and V_{ts} are Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix elements [9]. An accurate measurement of $B(K_L^0 \to \mu^+\mu^-)$ thus constrains the product of CKM matrix elements $(V_{td}^* V_{ts})$ given the top mass m_t, or vice versa [8, 10–14]. We discuss such constraints in the final section of this paper.

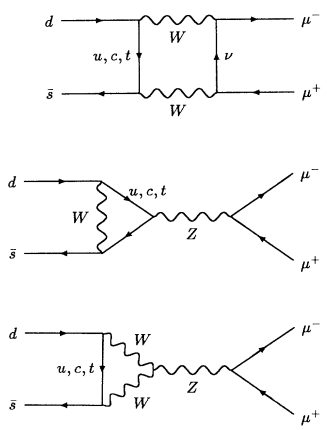

FIG. 1. Short-distance electroweak contributions to the amplitude for $K_L^0 \to \mu^+\mu^-$.

The $K_L^0 \to \mu^+\mu^-$ decay was first searched for with sufficient sensitivity by Clark *et al.* [15] at the Lawrence Berkeley Laboratory Bevatron. This experiment failed to observe any events and published an upper limit of 1.82×10^{-9} at the 90% confidence level, significantly below the unitarity bound. A subsequent pair of experiments at Brookhaven National Laboratory (BNL) (Carithers *et al.* [16]) together having nominally *less* sensitivity than that of Clark *et al.* observed nine events and measured $B(K_L^0 \to \mu^+\mu^-) = 12^{+8}_{-4} \times 10^{-9}$. The BNL experiments ran consecutively and used the same apparatus with slightly different magnetic fields and muon detectors. The disagreement between Carithers and Clark was resolved by a third BNL experiment (Fukushima *et al.* [17]) which observed three events and measured $B = 8.8^{+10.7}_{-5.5} \times 10^{-9}$. This confirmed the existence of $K_L^0 \to \mu^+\mu^-$ but left uncertain whether the branching ratio was above or below the unitarity bound. The situation improved in 1977 when an experiment at Argonne National Laboratory having approximately twice the sensitivity of the previous searches observed 16 $K_L^0 \to \mu^+\mu^-$ events and measured $B = 8.1^{+2.8}_{-1.8} \times 10^{-9}$ (Shochet *et al.* [18]). While this result precluded a branching ratio very much lower than the unitarity bound, the statistical error was too large to rule out nonstandard contributions to the rate. In the mid-1980s two experiments designed to search for $K_L^0 \to \mu^\pm e^\mp$ with high sensitivity also measured $B(K_L^0 \to \mu^+\mu^-)$ with high statistics. The experiments were conducted at KEK (Japan) and BNL, where more intense neutral beams than used previously had become available. The KEK experiment observed 178 $K_L^0 \to \mu^+\mu^-$ decays (Akagi *et al.* [19]), while the BNL experiment observed 707. We report here the final results from the BNL experiment.

The data from this experiment was taken over three years of running —1988, 1989, and 1990. The yields of $K_L^0 \to \mu^+\mu^-$ candidates after a background subtraction are 87, 274, and 346, respectively. The results for the 1988 and 1989 data sets have been previously reported [20–22]. The 1989 data set has now been reanalyzed along with the 1990 data set using a common procedure described here. The 1988 data used a different trigger and off-line pattern-recognition algorithm; since it comprises only 12% of the overall statistics, it was not reanalyzed. Our final sample represents the largest sample to date of $K_L^0 \to \mu^+\mu^-$ events.

The measurement of $B(K_L^0 \to \mu^+\mu^-)$ is challenging because of the branching ratio's small value. To attain a good statistical sample, a very high flux K_L^0 beam is needed. This unfortunately produces high rates in tracking chambers and counters from the dominant decay channels $K_L^0 \to \pi^\pm \ell^\mp \bar{\nu}$ and $K_L^0 \to \pi^+\pi^-\pi^0$. Good

mass resolution is also needed to discriminate against $K_L^0 \to \pi^\pm \mu^\mp \bar{\nu}$ decays followed by $\pi^\pm \to \mu^\pm \bar{\nu}$, which also produce $\mu^+\mu^-$ pairs. To normalize the sensitivity of the experiment, a prescaled sample of $K_L^0 \to \pi^+\pi^-$ decays is recorded simultaneously with the $K_L^0 \to \mu^+\mu^-$ sample and the ratio $\Gamma(K_L^0 \to \mu^+\mu^-)/\Gamma(K_L^0 \to \pi^+\pi^-)$ actually measured. This result is multiplied by the Particle Data Group [3] value for $B(K_L^0 \to \pi^+\pi^-)$ to obtain the branching fraction for $K_L^0 \to \mu^+\mu^-$. Since $K_L^0 \to \pi^+\pi^-$ and $K_L^0 \to \mu^+\mu^-$ decays are topologically similar, several systematic errors cancel in the ratio of partial widths.

II. APPARATUS

A. Detector and beam

The experiment, E791, was performed in the B5 beamline at the Alternating Gradient Synchrotron (AGS) of Brookhaven National Laboratory. In addition to measuring the $K_L^0 \to \mu^+\mu^-$ decay mode, E791 also searched for $K_L^0 \to \mu^\pm e^\mp$ and $K_L^0 \to e^+e^-$, the final results of which have recently been published [23, 24]. Figure 2 shows a plan view of the E791 neutral beamline and detector. The AGS provided $4.5-5.0 \times 10^{12}$ protons per beam spill with a momentum of 24 GeV/c incident on a 1.3 interaction-length Cu target. The cycle time of the machine averaged approximately 3.0 seconds and the duty cycle approximately 40%. The targeting angle was 2.75°, which represented a compromise between maximizing the number of K_L^0's accepted and minimizing the

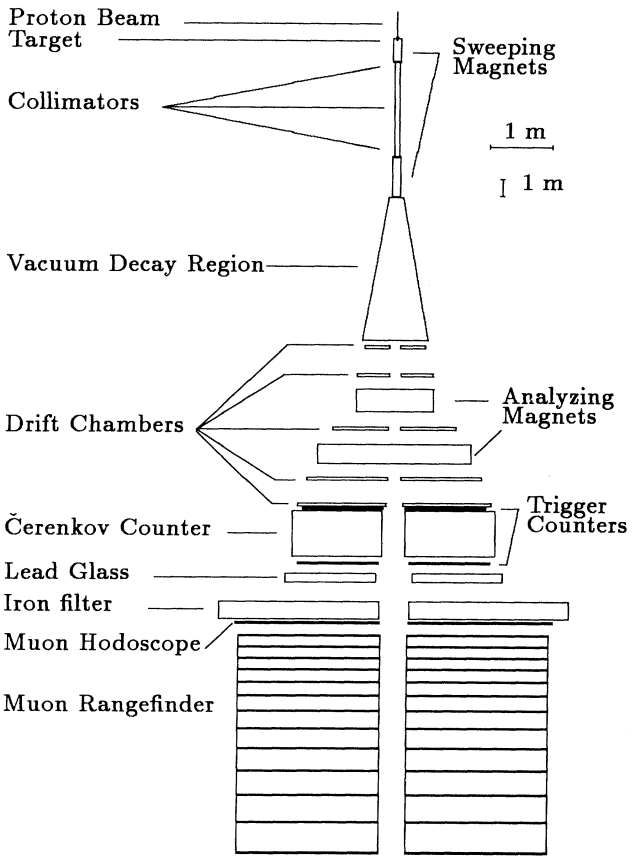

FIG. 2. Plan view of E791 neutral beam line and detector.

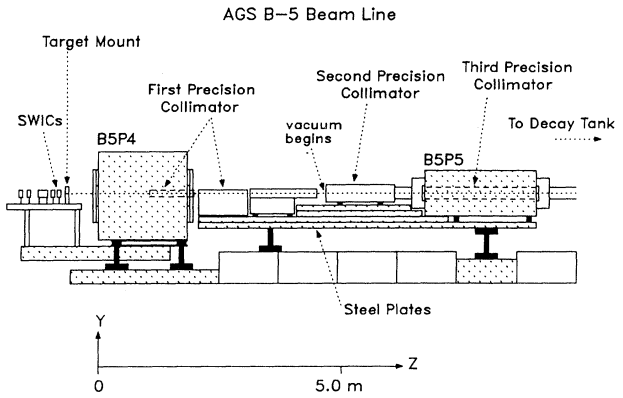

FIG. 3. Elevation view of beam-defining elements in the AGS B5 neutral beam line. The proton beam entered from the left. Lead foils (to convert γ's) inside the gap of the B5P4 magnet are not shown. Also inside this gap was the upstream half of the first precision collimator. The entire region was close packed with concrete, steel, and lead shielding.

n/K_L^0 ratio of the secondary beam. Downstream of the target were a series of collimators to define the beam, lead foils to convert γ's, and two sweeping magnets to clean the beam of charged particles (Fig. 3). The resultant beam consisted almost entirely of neutrons and K_L^0's in the ratio of $\sim$18:1. The collimators defined the beam to subtend a solid angle of 4.1 mrad horizontally $\times$ 15 mrad vertically (FWHM).

At 9.7 m from the target the beam entered an 8.0 m long vacuum tank evacuated to 0.020 torr, which served as the decay volume. The end of the tank was sealed by a flange and Mylar window assembly which differed between the 1989 and 1990 runs. In 1989 the flange contained three rectangular Mylar windows: a small central one through which the neutral beam passed and larger ones on each side through which K_L^0 decay products passed. In 1990 this assembly was replaced by a single round Mylar + Kevlar window.

Immediately downstream of the vacuum tank was the spectrometer, which consisted of two dipole magnets and five drift chamber (DC) stations. There were two stations upstream of the first magnet (a 48D48 with a vertical gap of 37 inches), two stations downstream of the second magnet (a 96D40 with a vertical gap of 44 inches), and one station between the magnets. With this arrangement the magnets made two nearly independent momentum measurements, sharing hits only in the middle drift chamber. The correlation between the measurements was such that a position mismeasurement at DC3 increased one momentum measurement while decreasing the other, leaving the average essentially unchanged. The chambers and all other detectors were arranged in two arms, allowing for passage of the neutral beam down the center through only helium. The coordinate system of the detector was defined as follows: the $+z$ axis pointed along the beam direction, the $+y$ axis pointed vertically upwards, and the $+x$ axis pointed horizontally to the left when viewed from upstream.

Each chamber [25] consisted of two x-measuring and two y-measuring planes with a single-wire resolution of

~ 130 μm. Both sets of planes were staggered by half the cell-size in order to resolve the left-side–right-side ambiguity of a single hit (Fig. 4). The chambers were operated with a 50/50 gas mixture of argon/ethane bubbled through either propanol or a water-ethanol mixture; the temperature of the bubbler was such that the gas saturated when the additive contributed 1–2 % of the total volume. The drift velocity was 51 μm/ns, giving a maximum drift-time for a 5.1 mm radius drift cell of ~ 100 ns. The wire signals were amplified and discriminated, and the discriminator times digitized by time-to-digital converters (TDC's) with a least-count of 2.5 ns [26]. Signals on pairs of adjacent wires were sent to meantimers [27], and the meantimed signals, with a jitter of less than 25 ns, were used in the trigger. Between the 1989 and 1990 running periods, the chambers comprising the most upstream DC pair were replaced by modules 25% larger. The new chambers were positioned 1.5 cm closer to the beam to take advantage of the larger acceptance of the new Mylar + Kevlar vacuum window. The new chambers and window resulted in a 37% increase in the acceptance for $K_L^0 \to \mu^+\mu^-$. Helium bags were placed between all drift chambers to reduce multiple scattering of K_L^0 decay products and also to reduce detector rates from neutron interactions.

The 48D48 and 96D40 dipole fields imposed horizontal impulses of opposite sign having magnitudes 300 MeV/c and 318 MeV/c, respectively. The large B_y (vertical) component of the field was carefully mapped on a two-inch cubic grid consisting of over 100 000 points. The minor B_x (horizontal) and B_z (along beam) components were taken from solutions of the two-dimensional Poisson equation. Spline fits were used to determine B_y between data points; in areas where the magnetic probe could not reach, B_y was calculated by extrapolation using the shape derived from the two-dimensional Poisson model.

Downstream of the spectrometer on each side of the beam were two banks of trigger scintillation counters (TSC's) separated by 3.3 m in z. Each bank had x-measuring and y-measuring slats which were read out to two different types of TDC's: the y-measuring slats

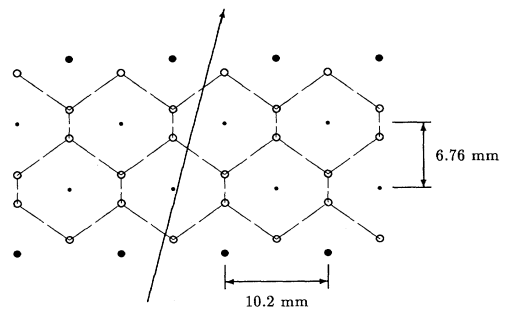

FIG. 4. Two x-measuring or two y-measuring planes of a drift-chamber module. The large open circles are field-shaping wires, the small solid circles are sense wires, and the large solid circles are guard wires. Drift-cell boundaries are denoted by dashed lines. The meantime of the two sense wires hit was used for triggering; the cut on meantime was asymmetric to account for the mean angle of track incidence.

used TDC's with a least-count of 2.5 ns, while the x-measuring slats used TDC's with a much smaller least-count of 0.2 ns [28]. The good timing resulting from the latter helped reject background from accidentals. The x-measuring slats were double ended, i.e., had phototubes on each end. Both x (meantimed) and y signals were used in the trigger.

Between the TSC banks on each side of the beam was a gas threshold Čerenkov counter (CER) used for electron identification. The gas mixture was 40% nitrogen and 60% helium kept at a pressure slightly above atmospheric. The index of refraction was monitored by an interferometer and maintained such that $n - 1 = (140 \pm 5) \times 10^{-6}$; this placed the μ and π thresholds at 6.3 GeV/c and 8.3 GeV/c, respectively. Each Čerenkov counter had eight spherical mirrors arranged in two rows of four, with each mirror reflecting light through a quartz window onto a 5-inch phototube. The tubes were read out to TDC's and analogue to digital converters (ADC's) [29], providing timing and pulse-height information.

Downstream of the Čerenkov counter and TSC banks was a lead-glass array (PbG) which also was used to identify electrons. The array was built in two longitudinal layers to give shower profile information in z. The upstream layer, referred to as the converter blocks, was 3.3 radiation lengths deep and consisted of 52 $10 \times 10 \times 90$ cm^3 blocks oriented with the long axis vertical. The downstream layer, referred to as the absorber blocks, was 10.5 radiation lengths deep and consisted of 216 $15 \times 15 \times 32$ cm^3 blocks arranged with the long axis along the beam direction. The blocks were wrapped in aluminized Mylar or vinyl with the exception of the downstream face of the converter blocks and the upstream face of the absorber blocks: these were left unwrapped so that any radiation damage could be cured *in situ* with UV lamps. Such curing was carried out between the 1989 and 1990 runs. The entire array sat in a light-tight, temperature-controlled enclosure, and the response of the system was monitored using a nitrogen laser and an optical dye. When excited by the laser, the dye emitted light of wavelength similar to that produced by electrons traversing the PbG. The light was fanned-out to each block via an optical fiber, and the response of the block was monitored for calibration purposes.

Downstream of the lead-glass array was an iron wall 91 cm deep to absorb hadrons, and downstream of the wall were two detectors used to identify muon candidates. The first was a "muon hodoscope" (MHO) consisting of 11 vertical x-measuring slats and 14 horizontal y-measuring slats on each side of the beam. The horizontal slats were single ended while the vertical slats were double ended. All signals were read-out to TDC's with a least-count of 0.2 ns, and both x (meantimed) and y signals were used in the trigger.

The most downstream detector was a "muon range finder" (MRG), which consisted of 75 three-inch-deep slabs of marble followed by 25 three-inch-deep slabs of aluminum on each side of the beam [30]. Between consecutive marble or aluminum slabs was a two-inch air gap, thirteen of which were instrumented with x-measuring and y-measuring drift-tube detectors. The most down-

stream drift-tube plane hit determined how many marble or aluminum slabs a muon candidate traversed before stopping. For positive identification of a muon, this range was required to be consistent with that expected from the momentum measurement made in the spectrometer. In addition, the transverse position of hits in the drift-tubes was required to match that expected from projecting the spectrometer tracks. The gaps were instrumented at z-intervals corresponding to 10% increments of range; this gave sufficient discrimination between real muons and pion "punch through."

B. Triggering and data acquisition

The experiment used multiple levels of triggering. The lowest level or L0 trigger required TSC hits in all four x-measuring planes and the two downstream y-measuring planes. The L0 signal was put in coincidence with mean-timer signals from the three most upstream DC's to form a "minimum bias" trigger. The pattern of DC hits required on each side of the beam was as follows: three out of four x-measuring planes and three out of four y-measuring planes in DC's 1 and 2, and one out of two x-measuring planes and one out of two y-measuring planes in DC3. The minimum bias logic thus required at least an eightfold coincidence of DC hits in coincidence with an L0 trigger; this indicated the presence of a track in each spectrometer arm which traversed the entire detector. The L0 trigger rate during the spill was typically 1000 kHz, and the minimum bias rate was typically 70 kHz.

The minimum bias trigger was put in coincidence with a dilepton signal to form the Level 1 (L1) trigger. A dilepton signal consisted of hits in each detector arm from either the Čerenkov counter or the muon hodoscope. For example, if both Čerenkov counters fired an L1 trigger was generated and an "ee" trigger bit would be set. If the Čerenkov counter on one side of the beam fired along with the muon hodoscope on the other side, an L1 "μe" bit would be set. For the muon hodoscope to be considered hit, both an x and a y slat had to have hits and both tubes on the x slat had to have signals. More than one L1 bit could be set for an individual event. The L1 trigger rate was typically 8 kHz.

Events passing the L1 trigger were further reduced a factor of ~ 50 by a filtering algorithm running online in eight 3081E processors [31]. This filtering stage is referred to as the Level 3 (L3) trigger. A Level 2 trigger based in hardware which would have quickly calculated a two-body invariant mass was designed but never implemented. When an L1 trigger was issued, all data from ADC, TDC, and latch crates was digitized and uploaded in parallel into 3081E memory [32]. The L3 algorithm did a fast unpacking of hits in DC's 1, 2, and 3 to reconstruct single-bend tracks. It then used a lookup table of field integrals to calculate track momenta, two-body dilepton masses m_{12}, and a collinearity angle θ_K, where θ_K is the angle between the reconstructed two-body momentum vector and the kaon direction as calculated from the target and vertex positions. Real two-body K_L^0 decays had a two-body momentum pointing back to the target, and hence small θ_K.

The reconstructed tracks were required to satisfy a collinearity cut $\theta_K^2 < 100$ mrad^2, a two-body mass cut $m_{12} > 460$ MeV/c^2, and in 1989 only, a mass cut $m_{12} < 550$ MeV/c^2. The lepton masses chosen for the m_{12} calculation were those corresponding to the L1 trigger bit(s) set. Events passing these criteria were written to magnetic tape. The L3 trigger rate was approximately 180 events/spill and included a sample of minimum bias triggers prescaled by a factor of 2000. These minimum bias triggers were subjected to the L3 algorithm with pion masses used for the m_{12} calculation, but the events were uploaded regardless of whether they passed or not. The result of the L3 calculation was recorded, however, such that the events could be used later to determine the L3 efficiency.

III. OFF-LINE ANALYSIS

A. Pattern recognition and track fitting

All dilepton and minimum bias events written to tape were reconstructed off-line in two "passes." The first pass ran a fast pattern-recognition routine which searched for two tracks, one in each spectrometer arm, which satisfied a loose vertex requirement. The second pass ran a slower, more refined fitting package which used the full magnetic field map and two distinct and complementary fitting algorithms, referred to as QT and FT.

The pattern-recognition routine began in the TSC's by combining hits in upstream and downstream banks to form potential track segments. The segments were projected upstream to find corresponding hits in DC's 5 and 4, and the TSC2-TSC1-DC5-DC4 track segments were projected upstream through the 96D40 analyzing magnet to DC3. Hits were searched for in the x view of DC3 within broad windows on either side of the position to which the track segment projected, to account for track bending. When a hit was found, the track momentum was tentatively determined. Hits in DC1 and DC2 were found by further projecting the track upstream, requiring that the bending angle in the upstream magnet be opposite and approximately equal in magnitude to the bending angle in the downstream magnet. The nonbending y-view segments were required to lie along a straight line. Tracks in the x view and tracks in the y view were subsequently paired to form three-dimensional tracks, and these were required to form a loose vertex inside the decay tank with a distance of closest approach (DOCA) of less than 12 cm. Events passing this criterion had the invariant mass m_{12} and collinearity angle θ_K recalculated. The particle masses used for the m_{12} calculation were the same as those which satisfied the L3 mass requirement. The momentum transverse to the K_L^0 direction as defined by the target and vertex positions (p_T) was also calculated. Events with $m_{12} > 470$ MeV/c^2 and either $\theta_K^2 < 10$ mrad^2 or $p_T^2 < 800$ (MeV/c)^2 were selected for kinematic fitting.

All events to be fit were subjected to both QT and FT algorithms. Both fitters used the full magnetic field map and "swum" tracks using a fourth-order Runge-Kutta al-

gorithm to solve the equations of motion. The fitters resolved the left-side–right-side ambiguity of single hits by minimizing their respective track χ^2's. After track and vertex fitting were completed, m_{12} and θ_K were recalculated a final time using the new track parameters.

The QT fitter did an iterative calculation of track parameters, treating the two magnets as separate spectrometers. For the x (bending) view, the values of hit positions found by the pattern-recognition routine were used to swim tracks along the route DC1 to DC2 to DC3 for the upstream spectrometer and along the route DC5 to DC4 to DC3 for the downstream spectrometer. The differences between the calculated track trajectories at DC3 and the actual hit position were used to correct the particle's momentum separately in each spectrometer. In the y (nonbending) view the procedure was similar, except the differences between the track projections at DC3 and the hit position were used to determine nominal scattering angles at DC2 and DC4. A scattering angle at DC3 was determined from the difference in y-view track angles at DC3.

The procedure was iterated until the track trajectories and hit positions agreed to within 10 μm in the x view; if this condition could not be satisfied, the event failed QT. Events which passed had an x-view χ^2 calculated from the following quantities: (1) δp, the difference between the momenta calculated in upstream and downstream spectrometers; (2) $\delta r \equiv \delta p/\sigma_{\delta p} + q\,\delta\theta_{x3}/\sigma_{\delta\theta_{x3}}$, where q is the electric charge and $\delta\theta_{x3}$ is the angular mismatch in the x view at DC3 between the track fitted upstream and the track fitted downstream. $\delta\theta_{x3}$ is found to be strongly correlated with δp but in a manner opposite for positive and negative tracks; hence the nonintuitive sum above was used in which the correlated parts cancel out and the smaller independent component remains.

The y-view χ^2 was calculated from $\delta\theta_{y2}$, $\delta\theta_{y3}$, and $\delta\theta_{y4}$, the y-view scattering angles at DC2, DC3, and DC4, respectively. The covariances between all quantities entering x- and y-view χ^2's were found from Monte Carlo simulation. As a final step a vertex χ^2 was calculated as the DOCA of the tracks divided by the error in the DOCA due to errors in the track angles:

$$\chi^2_v = \frac{(\mathrm{DOCA})^2}{(z_{\mathrm{DC1}} - z_v)^2\,(\sigma^2_{\theta L} + \sigma^2_{\theta R})}. \tag{2}$$

The mass and collinearity resolutions for QT-fit events show fair agreement with Monte Carlo simulations except for the presence of non-Gaussian tails in the data (see Fig. 5). The vertex χ^2 distributions also agree satisfactorily, but the track χ^2 distributions show significant differences. As shown in Fig. 6, the distribution from the data is broader than that from Monte Carlo simulation and has a longer tail out to larger χ^2 values. These differences are not completely understood. The broadening of the distribution is partly attributed to measurement and interpolation errors in the magnetic field, and the long tail is partly attributed to δ rays.

The FT fitter, in contrast with QT, did a single fit to all hits on a track in the entire spectrometer. The algorithm parametrized a track by the vector $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = (x,\ y,\ dx/dz,\ dy/dz,\ q/p)\ , \tag{3}$$

where the positions and derivatives are those at the vacuum window. The particle was swum through the spectrometer and a track χ^2 calculated based on the differ-

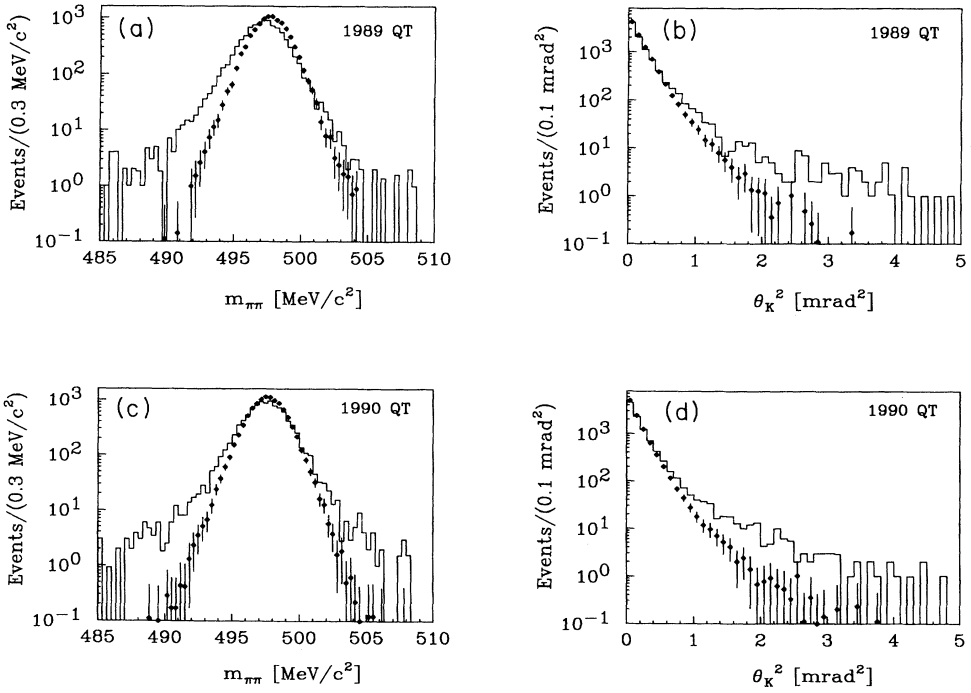

FIG. 5. The $K^0_L \to \pi^+\pi^-$ mass and collinearity resolutions for 1989 (a),(b) and 1990 (c),(d) QT-fit events. The histograms show the data while the circles with error bars show Monte Carlo spectra. The distributions are normalized to the same total number of events. There is fair agreement between Monte Carlo and data except for the presence of non-Gaussian tails in the data and a small shift in the $m_{\pi\pi}$ peak for the 1989 data. This shift is due to a slightly low magnetic field scale factor used in the track fitting; the effect upon the measured value of $B(K^0_L \to \mu^+\mu^-)$ is less than 0.3%.

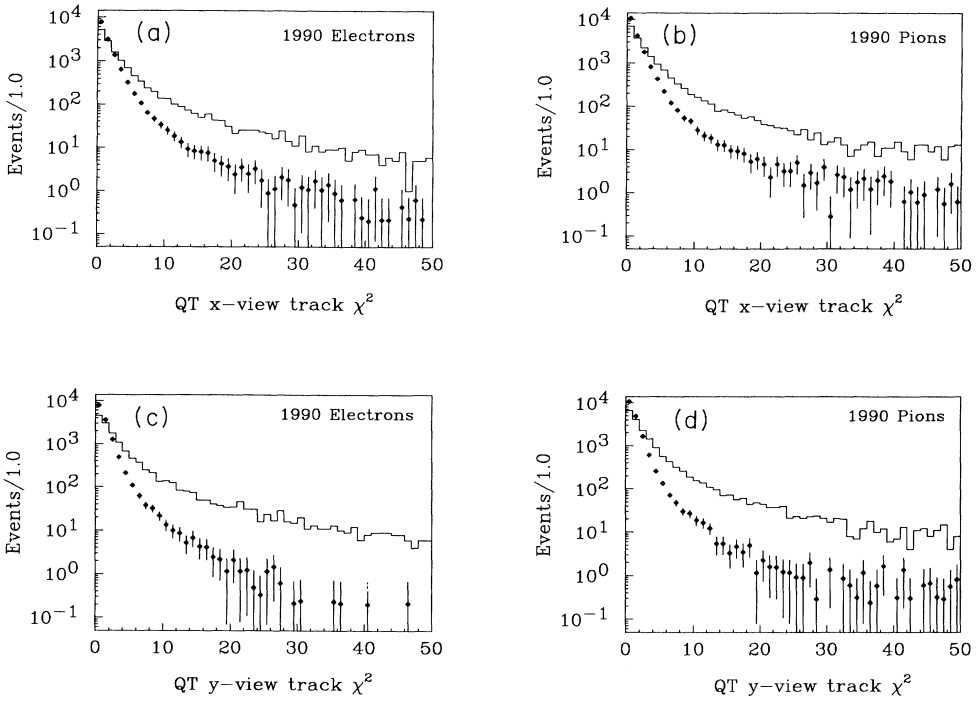

FIG. 6. The QT track χ^2 distribution for electrons (a),(c) and pions (b),(d) from well-identified $K_L^0 \to \pi e \bar{\nu}$ decays. The histograms show the data while the circles with error bars show Monte Carlo spectra. The distributions are normalized to the same total number of events. The electron plots include multiple scattering and brehmsstrahlung effects, while the pion plots include multiple scattering and pion decay. For both samples the distributions from data are broader than those from Monte Carlo and have longer tails out to larger χ^2 values.

ences in position between the calculated track trajectory and the twenty or fewer wires hit:

$$\chi^2 = \sum_i \sum_j (x_i^{\text{data}} - x_i^{\text{swum}})(V^{-1})_{ij}(x_j^{\text{data}} - x_j^{\text{swum}}) , \tag{4}$$

where i and j run over all DC hits. The matrix V_{mn} was momentum dependent and contained covariances $\langle \delta x_{\text{DC}m} \delta x_{\text{DC}n} \rangle$ determined from Monte Carlo simulation. Multiple scattering in the x and y views was assumed to be uncorrelated and V_{mn} consisted of two 10×10 submatrices. Expression (4) was minimized using a matrix inversion technique requiring two to four iterative swims per track.

The vertex fit was similar to the track fit except that instead of fitting for five track parameters, the algorithm fit for nine event parameters:

$$\boldsymbol{\beta} = (x_v,\ y_v,\ z_v,\ dx_1/dz,\ dy_1/dz,\ q_1/p_1,\ dx_2/dz,\ dy_2/dz,\ q_2/p_2). \tag{5}$$

The χ^2 to be minimized was

$$\chi^2 = \sum_i \sum_j (\alpha_i^{\text{data}} - \alpha_i^{\text{theor}})(W^{-1})_{ij}(\alpha_j^{\text{data}} - \alpha_j^{\text{theor}}) , \tag{6}$$

where i and j run over the ten previously fit track parameters, five for each track. The 10×10 covariance matrix W_{mn} was block diagonal with two 5×5 submatrices. Each submatrix was equal to the inverse of the matrix

$$\mathcal{M}_{ij} = \frac{\partial \chi^2(\boldsymbol{\alpha})}{\partial \alpha_i \partial \alpha_j} \qquad (i,j = 1,5), \tag{7}$$

which was determined previously in the track fitting procedure (one small correction was made: two terms were added to account for scattering in the vacuum window). There was $10 - 9 = 1$ degree of freedom in the vertex fit.

The mass and collinearity resolutions resulting from FT were essentially identical to those from QT. We continued the analysis using both fitters independently of each other, with different values of track and vertex χ^2 cuts used for each analysis stream. The resulting $K_L^0 \to \mu^+\mu^-$ branching ratios were then compared to provide a measure of the systematic error in $B(K_L^0 \to \mu^+\mu^-)$ due to biases in fitting and χ^2 cuts. As shown in Sec. VII, the QT and FT branching ratios found are extremely close.

The track and vertex χ^2 cuts were important for rejecting background to the $\mu^+\mu^-$ sample from $K_L^0 \to \pi\mu\bar{\nu}$ decays ($K_{\mu 3}$) and $K_L^0 \to \pi^+\pi^-$ decays in which one or both pions decayed. Figure 7 shows schematically a pion decaying in the horizontal plane of the spectrometer downstream of DC2. For such a decay, the momenta measured in the upstream and downstream halves of the spectrometer do not match, the track angles at DC3 do not match, and the track χ^2 is poor. If the pion had decayed upstream of DC1, the track χ^2 would have been satisfactory but the vertex χ^2 would have been poor (the

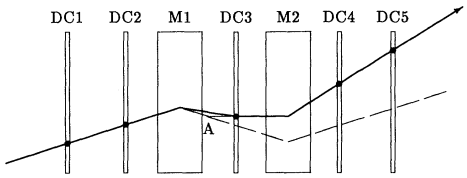

FIG. 7. Schematic diagram of pion decay in the spectrometer. The bending angles are exaggerated for clarity. The pion decayed at point A, with the dashed line showing the trajectory the pion would have followed had it not decayed. The thick solid line denotes the track determined by the QT fitter. The momenta determined for this track in upstream and downstream halves of the spectrometer do not match, nor do the track angles at DC3.

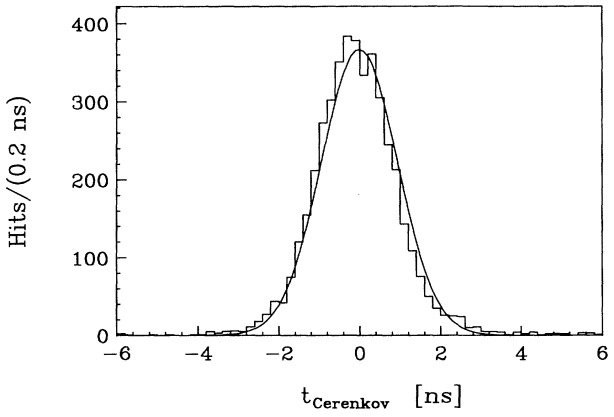

FIG. 8. Distribution of Čerenkov times for electrons identified without using the Čerenkov counter. The standard deviation of a Gaussian fit is 0.9 ns. Tracks identified as electrons were required to have $|t_{\text{CER}}| < 4.0$ ns.

vertex was determined by projecting the DC1–DC2 track segments). Finally, if the pion had decayed between DC1 and DC2, both vertex and track χ^2's would have been poor. Unlike the QT x-view track χ^2, the FT track χ^2 did not include the momentum mismatch δp directly; a cut was thus made explicitly on δp in the FT analysis, where the upstream and downstream momenta were separately determined by fitting to x hit positions in only the three upstream or three downstream drift chambers.

B. Particle identification

After kinematic fitting, particle identification proceeded by projecting tracks from DC5 (downstream of the magnets) to the Čerenkov counter, lead-glass array, muon hodoscope, and muon range finder, and correlating any hits in these detectors with the projected track positions. For example, an electron candidate was identified by a signal in the Čerenkov mirror to which the track projected with a time within 4 ns of the mean event time. The mean event time was defined as the mean time of the four double-ended x-measuring TSC's hit. Figure 8 shows a distribution of Čerenkov times for a sample of electrons identified without using the Čerenkov counter; the standard deviation of a Gaussian fit is 0.9 ns.

The response of the PbG array is illustrated in Fig. 9 for separate samples of electrons, muons, and pions. The signal, called E, was proportional to the total relativistic track length in the blocks. For showering electrons the total track length was proportional to the incident electron energy, and the array was thus calibrated to give $E_{\text{electrons}} \approx p$, where p is the track momentum measured in the spectrometer. Muons did not shower and their signals corresponded to a track length equal to the block depth. This signal was about the same size as that produced by a 400 MeV electron, and since the mean track momentum was ~ 3 GeV/c, $\langle E/p \rangle_{\text{muons}} \approx 1/7$. Pions did not shower electromagnetically but interacted hadronically; since the PbG was 1.2 nuclear interaction lengths deep, 70% of pions interacted somewhere in the glass. Fortunately, most of these interactions gave small E/p. However, if an interaction resulted in one or more π^0's carrying off most of the incoming pion momentum, the resultant E/p could be as large as that for an incident electron. To reject such interactions a cut was made on the quantity E_c/E_T, where E_c is the signal measured in the front converter blocks and E_T the signal measured in the total array (converter + absorber blocks). Since the converters comprised 0.29 interaction lengths, $(e^{-0.29} - e^{-1.2})/(1 - e^{-1.2}) = 64\%$ of pions which did interact in the glass did so after the converters. Over half of pion interactions thus left very little signal in the converters, and interactions with large E/p could be cut by requiring that E_c/E_T be greater than some small value. The loss in electron efficiency was small as the converters comprised 3.3 radiation lengths and 96% of electrons initiated electromagnetic showers there. Thus, the identification criterion for electrons was that they lie above and to the right of the contour indicated in Fig. 9. This contour was determined from a study of $K_L^0 \to \pi e \bar{\nu}$ decays (K_{e3}) and optimized pion rejection while retaining high electron identification efficiency.

The MHO hodoscope was located downstream of 0.91 m of iron, and thus only muons with $p \gtrsim 1.6$ GeV/c, and a small amount of hadronic "punch-through," gave MHO hits. Muon identification in the MHO was based on four criteria: (1) the distance in x between the projected track position and the center of the x-slat hit; (2) the distance in y between the projected track position and the center of the y-slat hit; (3) the difference in time between the x and y hits, where of the two x-slat phototube times, the one closest to that of the y time was used; (4) the difference between the mean time of x and y hits and the mean event time as determined from the x-measuring TSC's. For each of these criteria a confidence level (C.L.) for the track being a muon was found by comparing the time or position match, denoted ξ, with the probability distribution $\mathcal{P}(z)$ expected for real muons:

$$\text{C.L.} \equiv 2 \int_{|\xi|}^{\infty} \mathcal{P}(z)\, dz. \tag{8}$$

The probability distributions $\mathcal{P}(z)$ were symmetric

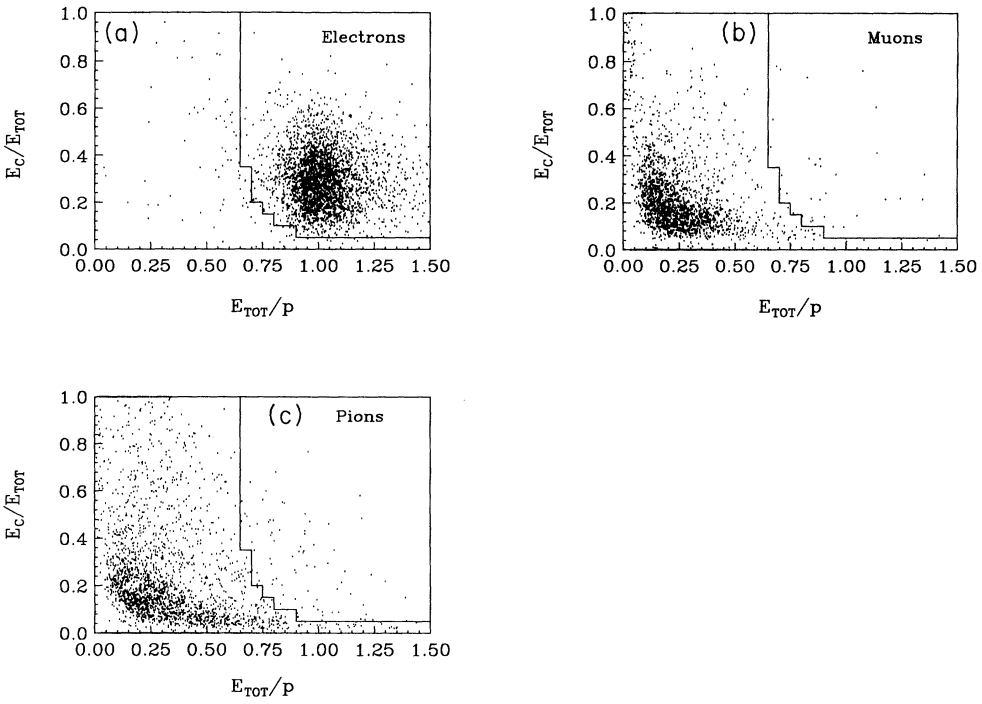

FIG. 9. The response of the lead-glass array to a sample of: (a) electrons, (b) muons, and (c) pions. All particles were identified without using the PbG. E_c is the energy recorded in the converter blocks, and E_T is the energy recorded in the total array (converter + absorber blocks). Tracks identified as electrons were required to lie above and to the right of the contour shown.

around $z = 0$ and normalized to unity; a factor of 2 is included in Eq. (8) so that for $\xi = 0$, an optimal match, C.L. $= 1$.

The probability distributions were found from a separate study of $K_{\mu 3}$ decays and were complicated: the distributions for the x and y position matches were convolutions of a Gaussian (from multiple scattering) with the "square" response function of MHO slats, while the distribution for the time-difference match was the product of a Gaussian and a special error function which accounted for the bias introduced by using the vertical tube time closest to the horizontal tube time. The "widths" or shapes of the distributions had a $1/p$ dependence as expected from multiple scattering effects. This dependence was studied and an analytic form $\sigma(1/p)$ found which modeled it, so that when finding confidence levels for a particular track, probability distributions were used which corresponded to the track's momentum [33].

The four confidence levels resulting from position and time matching were combined into an overall confidence level by mapping each individual confidence level onto an equivalent χ^2 distribution for one degree of freedom, summing the set of four χ^2's, and then mapping the sum back onto a confidence level by integrating the χ^2 distribution for four degrees of freedom. The distribution of the resulting overall confidence level for a sample of muons from $K_{\mu 3}$ decays is shown in Fig. 10. Except for the spike at zero, the distribution is approximately flat as expected. The spike at zero is due to missing or early hits caused by counter inefficiency or accidentals, and a small amount of pion contamination. To select $K_L^0 \to \mu^+\mu^-$ decays, both tracks in an event were required to have overall confidence levels > 0.0005.

The MRG track-finding routine found track-associated hits in each drift-tube plane beginning with the most upstream plane and working towards the most downstream. When two consecutive x planes and two consecutive y planes were found which had no track-associated hits, the last plane with an associated hit was noted. Muon candidates were identified by requiring that the last MRG gap with an associated hit be within three gaps of that expected based on the track's momentum. Figure 11 shows the response of the MRG to a sample of muons from $K_{\mu 3}$ decays and pions from K_{e3} decays (in which the pion is easy to identify). The muon sample shows good matching between the last gap hit and the last gap expected,

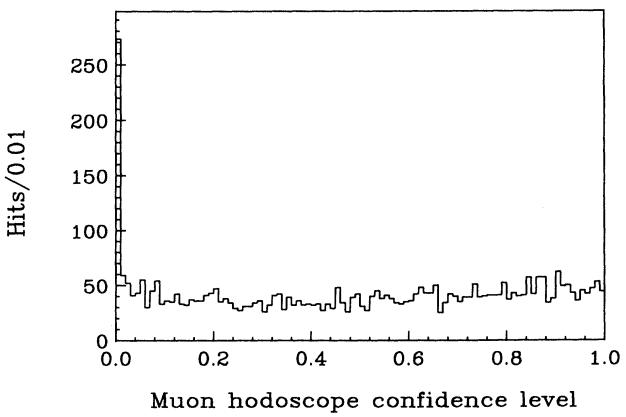

FIG. 10. The distribution of muon hodoscope confidence levels for muons from $K_L^0 \to \pi\mu\bar{\nu}$ decays, where the muons were identified without using the MHO. The distribution is approximately flat as expected, with a spike at zero due to missing or early hits and some pion contamination. Tracks identified as muons were required to have a confidence level > 0.0005.

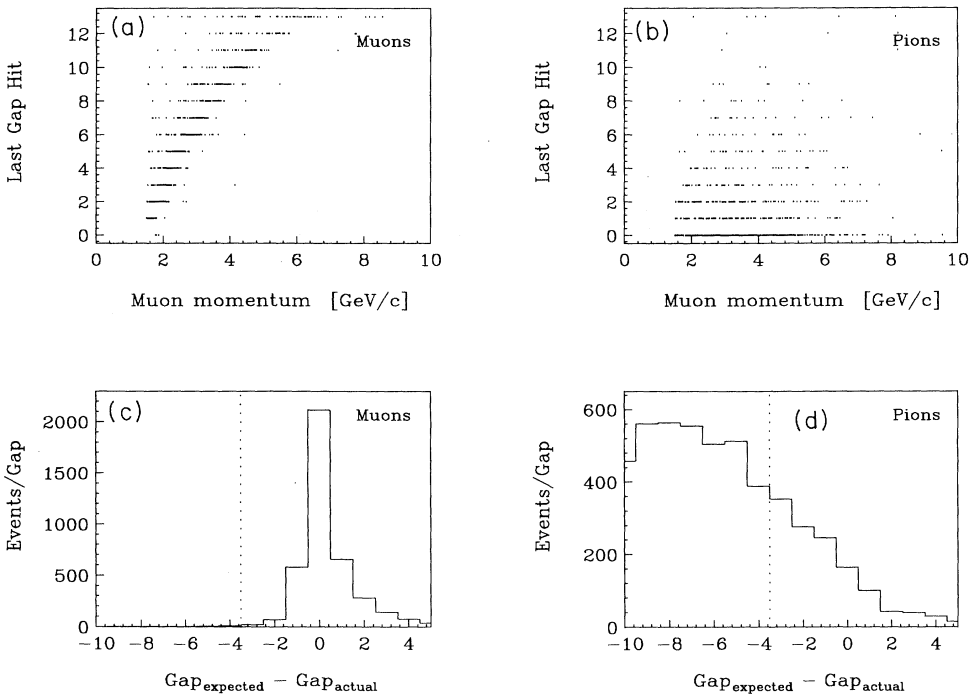

FIG. 11. The response of the muon range finder to a sample of muons from $K_L^0 \to \pi\mu\bar{\nu}$ decays (a),(c) and pions from $K_L^0 \to \pi e\bar{\nu}$ decays (b),(d). In (c) and (d) a gap difference cut of ≥ -3 is indicated by a dotted vertical line. The cut is seen to have high efficiency for muons and to give good rejection against pions.

while the pion sample shows the last gap hit to be significantly upstream of that expected. To select $K_L^0 \to \mu^+\mu^-$ decays, both tracks in an event were required to have a difference between the last gap hit and the last gap expected of ≥ -3.

C. Data reduction and event selection

The dilepton and minimum bias triggers written to tape were required to pass the pattern-recognition algorithm and at least one of the fitters, QT or FT. All successfully fit minimum bias triggers having $m_{\pi\pi} > 450$ MeV/c^2 were separately flagged. These prospective $K_L^0 \to \pi^+\pi^-$ decays were prescaled by a factor of 3 (2) for 1989 (1990), reducing the sample to a more manageable size while preserving a sufficient number of $K_L^0 \to \pi^+\pi^-$ decays to normalize the sensitivity of the experiment. The entire minimum bias sample was prescaled by a factor of 60 (30) for 1989 (1990), which preserved sufficient statistics to measure the efficiency of particle-identification detectors using $K_L^0 \to \pi\ell\bar{\nu}$ decays ($K_{\ell 3}$).

The dilepton events were divided into $\mu^\pm e^\mp$, $\mu^+\mu^-$, and e^+e^- samples, depending on the L1 trigger bit set. Events with more than one bit set were considered in more than one sample. Each sample was required to pass a high mass cut $m_{12} > 480$ MeV/c^2 and a low collinearity or transverse momentum cut: $\theta_K^2 < 10$ mrad^2 or $p_T < 800$ (MeV/c)^2. The samples were also required to have track-associated hits in the relevant detectors: the $\mu^+\mu^-$ sample was required to have hits in left and right MHO counters, the e^+e^- sample hits in left and right CER counters, and the $\mu^\pm e^\mp$ sample hits in the appropriate combination of CER and MHO counter.

The $\mu^+\mu^-$ sample was subsequently required to pass tight "event-quality" cuts:

(1) The vertex z position had to be downstream of 9.75 m, which was safely away from the fringe field of the most downstream sweeper magnet.

(2) $|x/z| < 0.003$ and $|y/z| < 0.010$, where (x, y, z) are the coordinates of the reconstructed vertex. This cut ensured that a decay originated from within the neutral beam, which had horizontal and vertical full-width divergences of 4.1 mrad and 15 mrad, respectively.

(3) $\chi^2_{\text{track}} < 50\,(200)$ and $\chi^2_{\text{vertex}} < 25\,(18)$ for QT (FT), and for FT only: $|p_{\text{front}} - p_{\text{back}}|/(p_{\text{front}} + p_{\text{back}}) < 0.025$. These cuts rejected pion decay, which, when originating from $K_L^0 \to \pi\mu\bar{\nu}$, could contribute background to the $\mu^+\mu^-$ sample. The loss to the $K_L^0 \to \pi^+\pi^-$ sample was accounted for by applying the same cuts to Monte Carlo $K_L^0 \to \pi^+\pi^-$ decays when calculating the acceptance. The momentum matching cut was made only for FT, as this quantity was included in QT's x-view track χ^2.

(4) $(p_+ - p_-)/(p_+ + p_-) < 0.66$. This cut was used to eliminate $\Lambda^0 \to p\pi^-$ decays in the $K_L^0 \to \pi^+\pi^-$ normalization sample; it was also made on $\mu^+\mu^-$ events to reduce potential bias between the $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ samples.

(5) No track trajectory could intersect material in the vacuum window flange, magnet coils, magnet shield plates, or edges of drift chambers. Tracks were required to traverse the PbG array, as the PbG response was used later to estimate K_{e3} background in the $K_L^0 \to \mu^+\mu^-$ sample.

To summarize, the final $\mu^+\mu^-$ sample consisted of pattern-recognized events passing at least one of the fitters, passing the above event-quality cuts, and having

an MHO confidence level > 0.0005 and an MRG gap difference ≥ -3 for each track. The good kinematic consistency between the L3 filter, the pattern-recognition algorithm, and the QT/FT fitters resulted in high selection efficiency, while the redundant muon identification significantly reduced backgrounds. The selection was done identically for both 1989 and 1990 data sets. The change in vacuum windows between the 1989 and 1990 running periods necessitated slightly altering the fiducial cut at the window to account for the different shapes of the flanges.

IV. FINAL EVENT SAMPLES

A. $\mu^+\mu^-$ signal and background

Figure 12 shows θ_K^2 plotted vs $m_{\mu\mu}$ for the $\mu^+\mu^-$ sample selected above. A large clustering of events is visible at low θ_K^2 near $m_K = 497.67$ MeV/c^2. Also visible is some amount of background extending to high $\mu^+\mu^-$ mass. This background is dominated by K_{e3} decays in which the pion decayed in-flight and the electron was misidentified as a muon; the misidentification produces a large two-body invariant mass equal to or exceeding m_K. The background also contains a small number of mismeasured $K_{\mu 3}$ decays, which for perfect momentum measurement have a kinematic end point at 489 MeV/c^2.

To subtract these backgrounds, two regions are defined: a signal region extending ± 6.0 MeV/c^2 around m_K and from 0–2.0 mrad^2 in θ_K^2, and a "background" region covering the same range of $m_{\mu\mu}$ but extending from 2.5–6.5 mrad^2 in θ_K^2. The cuts defining the signal region are chosen to be $\gtrsim 4$ standard deviations in resolution, where the resolutions were determined from a study of $K_L^0 \to \pi^+\pi^-$ events and were found to be 1.5 MeV/c^2 for $m_{\mu\mu}$ and 0.31 mrad for θ_K. For 1990 data (QT) there are 370 events in the signal region and 29 events in the background region. Of this latter group, 16 have one of the muon candidates also satisfying electron criteria and are thus identified as K_{e3} decays. The electron efficiency in this kinematic region was measured in a separate study to be 88.7%, and thus the most likely number of K_{e3} decays in the background region is 18.0. This implies that there are $29 - 18.0 = 11.0$ $K_{\mu 3}$ decays in this region. These backgrounds are subsequently extrapolated into the signal region at lower θ_K^2.

To extrapolate the K_{e3} background, we relax the MHO and MRG cuts in the $\mu^+\mu^-$ sample and instead require that events satisfy CER and PbG electron identification cuts. This results in a large sample of K_{e3} decays. Of these K_{e3}'s, those with $|m_{\mu\mu} - m_K| < 6.0$ MeV/c^2 are histogrammed in θ_K^2 [Fig. 13(a)]. The extrapolation factor needed is then the ratio of the number of events with $0 < \theta_K^2 < 2.0$ mrad^2 to the number with $2.5 < \theta_K^2 < 6.5$ mrad^2. This factor is 1.03, which gives 18.5 expected K_{e3} events in the signal region.

To extrapolate the $K_{\mu 3}$ background, we histogram in θ_K^2 those events from Fig. 12 which have $|m_{\mu\mu} - m_K| < 6.0$ MeV/c^2 but which do *not* have an electron candidate [Fig. 13(b)]. This sample contains very little K_{e3} background and consists of $K_L^0 \to \mu^+\mu^-$ decays at low θ_K^2 and $K_{\mu 3}$ decays at higher θ_K^2. The distribution of $K_{\mu 3}$ decays appears roughly flat in θ_K^2 and we assume this dependence. The level of this background is small enough such that even a 50% error in this assumption introduces negligible error into the branching ratio mea-

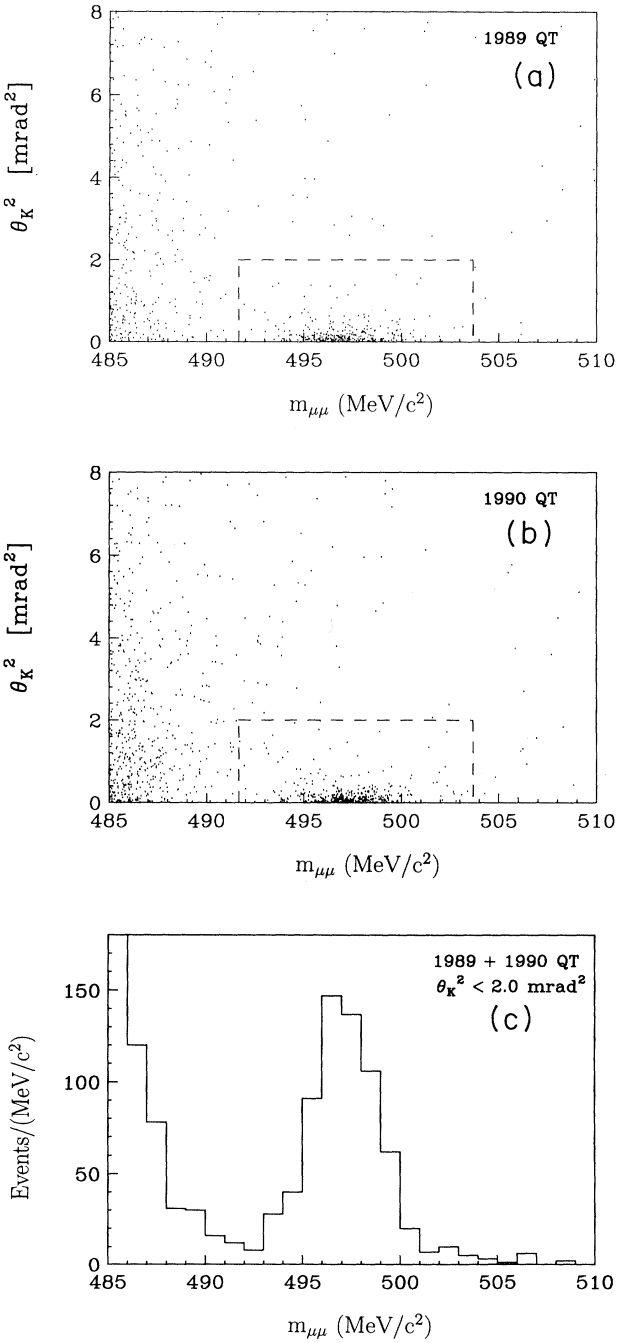

FIG. 12. θ_K^2 vs $m_{\mu\mu}$ for the final $\mu^+\mu^-$ sample: (a) 1989 QT data, and (b) 1990 QT data. The dashed line indicates the signal region defined by $|m_{\mu\mu} - m_K| < 6.0$ MeV/c^2 and $\theta_K^2 < 2.0$ mrad^2. The $m_{\mu\mu}$ distribution for 1989 + 1990 events with $\theta_K^2 < 2.0$ mrad^2 is shown in (c).

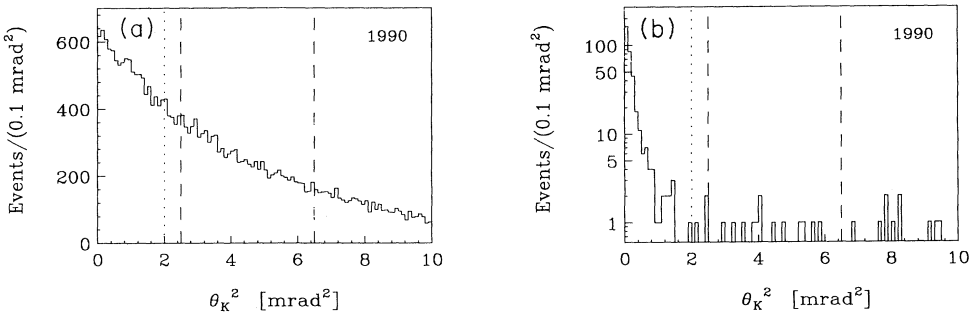

FIG. 13. The θ_K^2 spectrum for: (a) $K_L^0 \to \pi e\bar{\nu}$ events, and (b) $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi\mu\bar{\nu}$ events, where the latter appear at higher θ_K^2. Events plotted must pass all event-quality cuts and satisfy $|m_{\mu\mu} - m_K| < 6.0$ MeV/c^2. The signal region is indicated by a dotted vertical line; the background region is indicated by dashed lines.

surement (the statistical error on signal-region events is much larger). Extrapolating as flat the 11.0 estimated $K_{\mu 3}$'s in the background region gives 5.5 events in the signal region. The final number of $K_L^0 \to \mu^+\mu^-$ candidates is then $370 - 18.5 - 5.5 = 346$. These results are tabulated in Table I along with those for the 1989 data set. We mention that we could have counted K_{e3} events in the signal region directly and not needed to use an extrapolation factor. However, the electron efficiency in this region is more difficult to obtain as $K_L^0 \to \mu^+\mu^-$ decays contaminate the measurement and need to be cut away with an MHO cut.

B. $\pi^+\pi^-$ normalization

The sensitivity of the $K_L^0 \to \mu^+\mu^-$ search is calculated by counting the number of $K_L^0 \to \pi^+\pi^-$ decays reconstructed. These decays were recorded with the minimum bias trigger, required to pass the L3 mass and collinearity cuts off-line, and subjected to the same pattern recognition, fitting, and event-quality cuts as the $\mu^+\mu^-$ sample. The only difference between $\pi^+\pi^-$ and $\mu^+\mu^-$ samples aside from a minimum bias prescale factor is that the former was not subjected to particle-identification cuts. Rather than make such cuts, we prefer to subtract off the well-understood $K_{\mu 3}$ and K_{e3} backgrounds under the $K_L^0 \to \pi^+\pi^-$ peak.

The background subtraction proceeds as follows: we first count the number of events in a signal region identical to that used for the $\mu^+\mu^-$ counting. For 1990 data (QT) there are 40 485 events in this region. To estimate the $K_{\mu 3}$ and K_{e3} background, all events with $\theta_K^2 < 2.0$ mrad^2 are projected onto the $m_{\pi\pi}$ axis [Fig. 14(a)] and all events with $|m_{\pi\pi} - m_K| < 6.0$ MeV/c^2 are projected onto the θ_K^2 axis [Fig. 14(b)]. A large sample of Monte Carlo $K_{\ell 3}$ decays is then generated and the $m_{\pi\pi}$ and θ_K^2 line shapes matched to the data by normalizing to the number of events away from the $K_L^0 \to \pi^+\pi^-$ peak. For the mass spectrum the normalization region is 482–490 MeV/c^2 and 505–520 MeV/c^2, while for the θ_K^2 spectrum the normalization region is 3.0–10.0 mrad^2. The resultant number of $K_{\ell 3}$ decays in the signal region is 8259 ± 118 using the $m_{\pi\pi}$ line shape and 8079 ± 95 using the θ_K^2 line shape, where the errors are statistical and partially correlated. These two estimates differ by more than one standard deviation, and the result that the θ_K^2 estimate is lower than the $m_{\pi\pi}$ estimate is also obtained with 1989 data and with subsets of 1989 and 1990 data selected with tighter kinematic cuts. We thus consider the difference a systematic effect. The Monte Carlo line shapes in both cases closely resemble the data, and we take the central value between estimates as the number of background events and include half the difference as a systematic error. The number of $K_L^0 \to \pi^+\pi^-$ candidates is then $40\,485 - 8169 = 32\,316 \pm 233$ (stat) ± 90 (syst). The background estimates are stable against variation in the specific normalization regions used.

A small correction is made to the $K_L^0 \to \pi^+\pi^-$ yield to account for contributions from $K_S^0 \to \pi^+\pi^-$ decays and interference between $K_S^0 \to \pi^+\pi^-$ and $K_L^0 \to \pi^+\pi^-$ amplitudes. For an initially pure K^0 beam, the time dependence of the decay rate to $\pi^+\pi^-$ is

$$\left|\langle\pi^+\pi^-|H|K^0(t)\rangle\right|^2 \propto e^{-t/\tau_S} + |\eta_{+-}|^2 e^{-t/\tau_L} + 2|\eta_{+-}| e^{-(t/2\tau_S + t/2\tau_L)} \cos(\Delta m_K t/\hbar - \phi_{+-}), \quad (9)$$

TABLE I. The results of the $\mu^+\mu^-$ event counting and background subtraction procedure for QT. The signal region is defined as $|m_{\mu\mu} - m_K| < 6.0$ MeV/c^2 and $\theta_K^2 < 2.0$ mrad^2; the background region is defined as the same mass interval and $2.5 < \theta_K^2 < 6.5$ mrad^2. The 1989 and 1990 data sets were collected over running periods lasting 15 weeks and 12 weeks, respectively.

Year	$\mu^+\mu^-$ in signal region	$\mu^+\mu^-$ in background region with electron	$\mu^+\mu^-$ in background region no electron	Electron efficiency	Electron extrapolation factor	K_{e3} in signal region	$K_{\mu 3}$ in signal region	$K_L^0 \to \mu^+\mu^-$ candidates
1989	292	13	5	0.896	1.152	16.7	1.7	273.5 ± 17.7
1990	370	16	13	0.887	1.027	18.5	5.5	346.0 ± 19.8

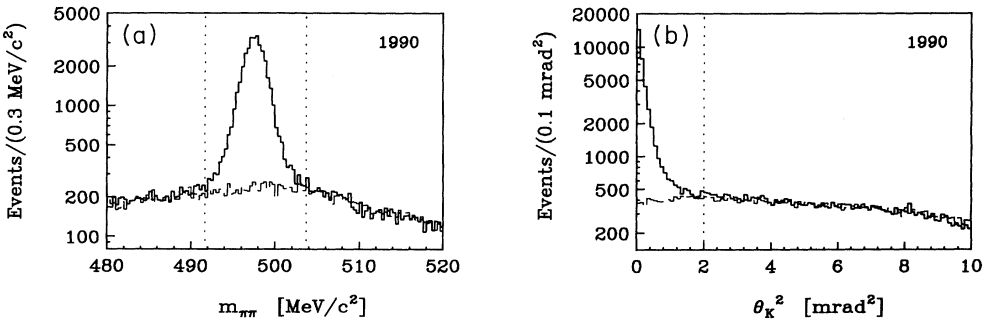

FIG. 14. Matching the Monte Carlo $K_L^0 \to \pi\ell\bar{\nu}$ spectrum (dashed) to the $K_L^0 \to \pi^+\pi^-$ normalization sample selected without particle-identification cuts (solid). (a) Events with $\theta_K^2 < 2.0$ mrad^2, and (b) events with $|m_{\pi\pi} - m_K| < 6.0$ MeV/c^2. The signal region is indicated by dotted vertical lines.

where the first term on the right-hand side corresponds to the K_S^0 contribution to $\pi^+\pi^-$, the second term corresponds to the K_L^0 contribution, and the last term corresponds to the interference between the two and is referred to here as the K_I^0 contribution. The parameters τ_S and τ_L are the lifetimes of the short- and long-lived neutral kaon states, respectively, and Δm_K is their mass difference. The CP-violating parameters $|\eta_{+-}|$ and ϕ_{+-} are the usual magnitude and phase, respectively, of the ratio of decay amplitudes:

$$\frac{\langle\pi^+\pi^-|H|K_L^0\rangle}{\langle\pi^+\pi^-|H|K_S^0\rangle} \equiv \eta_{+-} = |\eta_{+-}|e^{i\phi_{+-}}. \tag{10}$$

To correct for K_S^0 and K_I^0 contributions to the $\pi^+\pi^-$ sample, each event is individually weighted according to its probability of originating from a K_L^0. This probability is the ratio of the second term in Eq. (9) to the sum of all three terms in (9), where the time t of a decay is $(z/p) \times m_K$. The parameters τ_S, τ_L, $\Delta m_K/\hbar$, $|\eta_{+-}|$, and ϕ_{+-} taken from the Particle Data Group [3].

The resultant weights are applied to the data when constructing the mass and collinearity distributions used for the background subtraction procedure. The weighted distributions correspond to those which would have been obtained had the experiment run with a pure K_L^0 beam, with one exception: the semileptonic background in the data is also weighted. This is undesirable because $K_{\ell 3}$ decays receive no significant contribution from K_S^0. To account for this, we apply the K_L^0 weighting to the Monte Carlo $K_{\ell 3}$ background such that the data and Monte Carlo $K_{\ell 3}$ line shapes remain identical.

To use Eq. (9) as prescribed requires making a small correction to the interference term to account for the fact that $\bar{K}^0$'s as well as K^0's were produced in our beam by pA collisions at the target. The rate $\left|\langle\pi^+\pi^-|H|\bar{K}^0(t)\rangle\right|^2$ is identical to that given in Eq. (9) except that the interference term changes sign; this term thus becomes "diluted" when K^0's and $\bar{K}^0$'s are produced together. The dilution factor D has been measured at AGS energies [34] and is parametrized as

$$D(p) \equiv \frac{N_{K^0} - N_{\bar{K}^0}}{N_{K^0} + N_{\bar{K}^0}} = 1 - 1.5\,e^{-0.17p}\,, \tag{11}$$

where p is the kaon momentum in GeV/c. Multiplying the interference term by this factor accounts for $\bar{K}^0$ production.

To summarize, all high-mass $\pi^+\pi^-$ events are individually weighted by the properly normalized K_L^0 term of Eq. (9) before the background subtraction procedure is implemented. The effect on $\pi^+\pi^-$ counting is a reduction in the number of candidate events by 1.1%. All results are listed in Table II for 1989 and 1990 data sets. The errors listed are statistical and systematic, respectively, where the former results from the number of events in the signal region and the latter results from the $K_{\ell 3}$ background subtraction.

V. CALCULATION OF GEOMETRIC ACCEPTANCE

The acceptance of the E791 detector for $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ decays is studied via Monte Carlo simulation. As input to the simulation, we made a preliminary determination from data of the momentum spectrum of K_L^0's exiting our target at 2.75°. The event generator sampled from this distribution as well as from distributions in $|x/z|$ and $|y/z|$ (also determined from data) in order to realistically simulate the divergence of the neutral beam. As a last step the generator sampled from a distribution containing the correlation of p_K with $|y/z|$; because the proton beam was angled downwards with respect to the horizontal, K_L^0's at negative y/z had slightly

TABLE II. The results of the $\pi^+\pi^-$ event counting and background subtraction procedure for QT. The errors listed in the last column are statistical and systematic, respectively. The signal region is the same as that of Table I.

Year	$\pi^+\pi^-$ in signal region	Monte Carlo $K_L^0 \to \pi\ell\bar{\nu}$ decays in signal region			$K_L^0 \to \pi^+\pi^-$ candidates
		$m_{\pi\pi}$ spectrum	θ_K^2 spectrum	Average	
1989	19 060	3895	3736	3816	15 245 ± 157 ± 80
1990	40 485	8259	8079	8169	32 316 ± 233 ± 90

higher momentum than those at positive y/z.

The generated K_L^0's were forced to decay within the vacuum tank, and the decay products were swum through the spectrometer using the magnetic field maps of the analyzing magnets. If the trajectory intersected a drift chamber or downstream TSC counter—the detectors used for pattern recognition and track fitting—a hit was digitized and entered into the event buffer in form identical to that for raw data. All simulated hits included channel inefficiencies and smearing over the measured resolution function of the detector element. The former were taken from periodic measurements of detector efficiencies and included any dead channels. After all hits were packed into the raw event buffer, the analysis of a Monte Carlo event proceeded in a manner identical to that for a real event.

When analyzing Monte Carlo events, two adjustments are made which improve the agreement between data and the Monte Carlo samples. The first adjustment accounts for a difference between the p_K spectrum used to generate events and the final spectrum observed. The adjustment involves weighting the simulated K_L^0 decays according to their momentum p_K such that the two momentum distributions agree. The weighting function $\mathcal{F}(p_K)$ is taken as the ratio of the smoothed momentum spectrum from data to that from the Monte Carlo simulation. For this comparison, fully-reconstructed $K_L^0 \to \pi^+\pi^-$ decays are used. However, taking the ratio is complicated by the fact that the $\pi^+\pi^-$ sample contains $K_{\ell 3}$ background, and the genuine $\pi^+\pi^-$ decays receive K_S^0 and K_I^0 contributions which are not simulated in the Monte Carlo program. Thus for the momentum comparison to be accurate, the Monte Carlo $K_L^0 \to \pi^+\pi^-$ sample is combined with a 22% admixture of Monte Carlo $K_{\ell 3}$ decays (the ratio is that determined from the $\pi^+\pi^-$ background subtraction procedure) which have been "corrected" event-by-event to eliminate K_S^0 and K_I^0 contributions (see Sec. IV B). This modified Monte Carlo sample nominally corresponds to the $\pi^+\pi^-$ sample from data after the latter has been corrected for K_S^0 and K_I^0 contributions.

Since the beam and targeting conditions for the 1989 and 1990 running periods were identical, the K_L^0 momentum spectra for the two years were in principle the same. We thus sum the 1989 and 1990 momentum spectra before determining $\mathcal{F}(p_K)$ in order to improve the statistics. When combining the Monte Carlo samples, their relative normalization is taken to match the ratio of $K_L^0 \to \pi^+\pi^-$ candidates observed for the respective years. The $\mathcal{F}(p_K)$ function obtained is carried through all subsequent Monte Carlo analyses: all distributions and event counting use these weights, which is equivalent to regenerating the Monte Carlo samples using the modified momentum spectrum. To estimate the error in our final acceptances due to uncertainty in the momentum spectrum, we find weighting functions for 1989 data alone and for 1990 data alone; these are used to obtain a measure of how much the acceptance varies as a result of reasonable variation in p_K. We quote this variation, about 1.8% in the ratio of acceptances $A_{\pi\pi}/A_{\mu\mu}$, as an additional systematic error. We note that the momentum spectrum observed is softer than that expected based on the invariant spectrum $Ed^3\sigma/dp^3$ measured by Skubic *et al.* [35] using a 300 GeV/c proton beam (see Fig. 15).

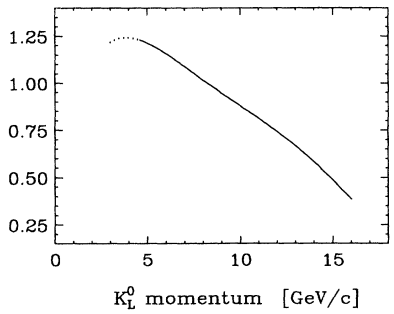

FIG. 15. The ratio of the p_K spectrum from data to that from Monte Carlo simulation, where the Monte Carlo events were generated according to the $Ed^3\sigma/d^3p$ invariant spectrum measured and parametrized by Skubic *et al.* [35]. The dotted part of the curve corresponds to $x < 0.2$, where Skubic *et al.* had no data and their parametrization may not be valid.

The second adjustment made to Monte Carlo events is an event-by-event weighting based on run number which results in the yield of Monte Carlo $K_L^0 \to \pi^+\pi^-$ decays matching the yield from data on a run-by-run basis. This run weighting accounts for variations in acceptance as a function of time (e.g., from variations in channel efficiencies) occurring in combination with changes in beam intensity and targeting efficiency.

The Monte Carlo samples thus have two weights applied: one for momentum and another for run number. The resulting $K_L^0 \to \pi^+\pi^-$ sample looks very similar to data, as shown in Fig. 16. The figure compares the reconstructed K_L^0 momenta, vertex z positions, x and y positions of the pion tracks at the vacuum window, and the fractions of "inbend" and "outbend" events (posi-

TABLE III. Geometric acceptances for $K_L^0 \to \pi^+\pi^-$ and $K_L^0 \to \mu^+\mu^-$ decays (QT), for three different momentum spectra (weightings). The percentages listed are for decays occurring within the vacuum tank. The variation in the ratio $(A_{\pi\pi}/A_{\mu\mu})$ is taken as a systematic error in $B(K_L^0 \to \mu^+\mu^-)$.

	1989 Data			1990 Data		
	$A_{\pi\pi}$ (%)	$A_{\mu\mu}$ (%)	$A_{\pi\pi}/A_{\mu\mu}$	$A_{\pi\pi}$ (%)	$A_{\mu\mu}$ (%)	$A_{\pi\pi}/A_{\mu\mu}$
Momentum spectrum						
Fit to 1989 $\pi^+\pi^-$ spectrum	4.47	3.86	1.157	6.10	5.31	1.149
Fit to 1990 $\pi^+\pi^-$ spectrum	3.47	2.94	1.177	4.72	4.05	1.167
Fit to 1989 + 1990 combined	3.71	3.17	1.173	5.06	4.35	1.163

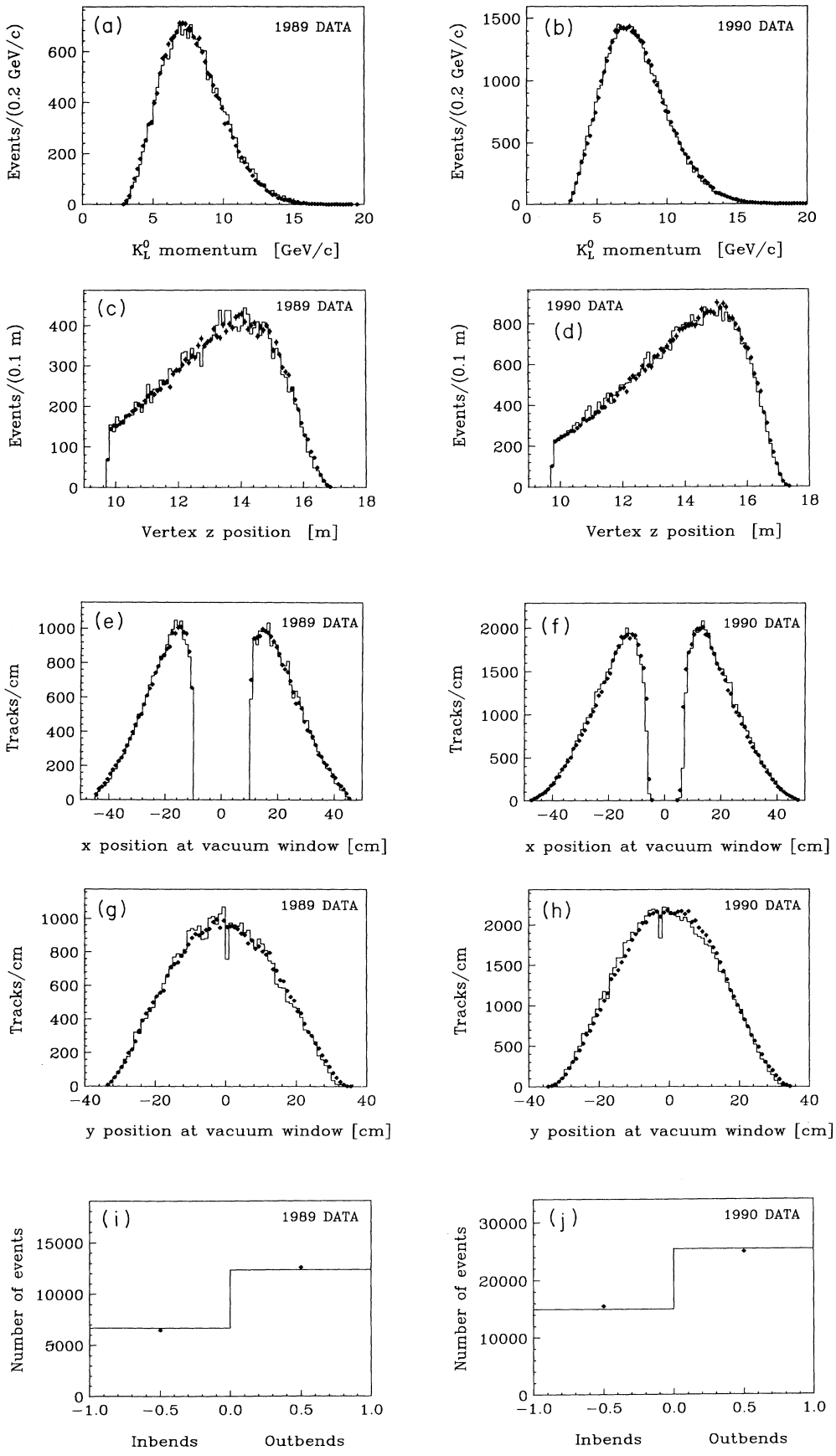

FIG. 16. Comparing the Monte Carlo $K_L^0 \rightarrow \pi^+\pi^-$ sample (circles) with data (histogram). The Monte Carlo events are weighted by momentum and run number (see text), and they include a 22% admixture of $K_{\ell 3}$ decays to reflect the semileptonic background in the data. The momentum weighting was adjusted to give good agreement between data and Monte Carlo. Both data and Monte Carlo $K_{\ell 3}$ decays were "corrected" for K_S^0 and K_I^0 contributions.

tive track on right or left side of detector, respectively). The acceptances for $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ are calculated as the sum of the weights of all Monte Carlo events which pass event-quality cuts and lie inside the signal region, divided by the sum of the weights of all events generated and forced to decay within the vacuum tank. The resultant acceptances and the ratio of acceptances $A_{\pi\pi}/A_{\mu\mu}$ are listed in Table III. The table lists acceptances for all three momentum weightings: that resulting from 1989 data alone, that resulting from 1990 data alone, and that resulting from 1989 + 1990 data combined.

VI. CALCULATION OF DETECTION EFFICIENCIES

A. Level 1 trigger

One source of bias between the $\mu^+\mu^-$ and $\pi^+\pi^-$ samples was the L1 trigger, which required that an MHO x slat (both tubes) and an MHO y slat fired in each detector arm. This requirement is more restrictive than the off-line MHO cuts, which required that only one tube on an x-measuring slat fired; thus the L1 efficiency for events passing all off-line cuts must be separately measured and corrected for. This correction also accounts for a small amount of L1 deadtime.

The L1 efficiency is measured by selecting from minimum bias data a sample of $K_{\mu 3}$ decays in which the pion decayed upstream of the MHO, giving two muons in the event. This sample is subjected to the same particle identification and event-quality cuts as were used to select the $K_L^0 \to \mu^+\mu^-$ candidate sample. Of the $K_{\mu 3}$'s which pass these cuts, the fraction which has the L1 trigger bit set is taken as the L1 trigger efficiency for $K_L^0 \to \mu^+\mu^-$.

There is uncertainty in this calculation due to the fact that $K_L^0 \to \pi\mu\bar{\nu}$ is a three-body decay and the muons populate slightly different regions of the MHO than do muons from $K_L^0 \to \mu^+\mu^-$. To reduce possible systematic error arising from this difference, we use only the highest mass and lowest collinearity $K_{\mu 3}$'s which still provide good statistics. The mass range chosen is 380–480 MeV/c^2 and the collinearity range is $\theta_K^2 < 100$ mrad^2. The resulting L1 efficiencies are 0.984 ± 0.005 for 1989 data and 0.977 ± 0.004 for 1990 data, where the errors given are statistical. To estimate the systematic error we repeat the calculation using $K_{\mu 3}$'s from a lower mass interval; the resulting efficiencies differ from the first result by much less than the statistical errors, implying that the systematic error can be neglected (for the given statistics).

B. Level 3 trigger

All L1 $\mu^+\mu^-$ triggers were required to pass the L3 trigger in order to be written to tape. To measure the L3 efficiency, the algorithm was also performed on minimum bias triggers from which $K_L^0 \to \pi^+\pi^-$ decays were selected. These triggers were written to tape along with the L3 trigger decision regardless of whether they passed or not. Thus, one measure of the L3 efficiency is the fraction of $K_L^0 \to \pi^+\pi^-$ events passing all analysis cuts which have the L3 trigger bit set. This fraction is 0.64 for 1989 data and 0.90 for 1990 data. The improvement in efficiency resulted from a faster and more sophisticated algorithm which looped over more combinations of hits in DC's 1–3 until the L3 mass and collinearity cuts were satisfied.

To reduce bias between the $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ samples, the L3 cut for the latter is made off-line (using the L3 trigger bit) and the *ratio* of separate $\mu^+\mu^-$ and $\pi^+\pi^-$ L3 efficiencies used to correct the ratio of event yields. The separate efficiencies are calculated in two steps. First, the minimum bias $\pi^+\pi^-$ sample is binned in three quantities: vertex z position, K_L^0 momentum, and the opening angle between the two tracks. The fraction of events in each bin which pass the L3 trigger is recorded. There are 20 bins used for each variable, resulting in an overall grid of $20^3 = 8000$ separate efficiencies. Next, Monte Carlo $K_L^0 \to \mu^+\mu^-$ events are generated and used to find the fraction of $K_L^0 \to \mu^+\mu^-$ decays populating each bin. The $K_L^0 \to \mu^+\mu^-$ L3 efficiency is then the sum of the products of bin populations with their respective efficiencies. As a check, we repeat the calculation using the fractional bin populations of Monte Carlo $K_L^0 \to \pi^+\pi^-$ decays and recover essentially the fraction of $K_L^0 \to \pi^+\pi^-$ events passing all analysis cuts which have the L3 trigger bit set. All L3 efficiencies for 1989 and 1990 are listed in Table IV. The ratio of $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ efficiencies is very close to unity, the difference being due to the muons and pions populating slightly different regions of the drift chambers.

The errors on $\pi^+\pi^-$ and $\mu^+\mu^-$ efficiencies listed in the table are the sums in quadrature of statistical and systematic errors, where the statistical errors are propagated from bin efficiencies and the systematic errors are taken as the small difference between the two $\pi^+\pi^-$ efficiency calculations described above. This difference is attributed to uncertainty in the K_L^0 momentum spectrum. The error in the ratio of L3 efficiencies ($\varepsilon_{\pi\pi}/\varepsilon_{\mu\mu}$) is smaller than the error on individual efficiencies because both the statistical errors and systematic errors are strongly correlated (the $\mu^+\mu^-$ and $\pi^+\pi^-$ calculations use the same bin efficiencies and K_L^0 momentum spectrum). The latter correlation was studied and the ratio of L3 efficiencies found to be essentially independent of variations in the p_K spectrum. We thus take the error in the ratio to conservatively be the larger of the statistical errors on $\varepsilon_{\pi\pi}$ and $\varepsilon_{\mu\mu}$.

C. Particle-identification cuts

To calculate the acceptance of MHO and MRG particle-identification cuts (CL > 0.0005 and gap difference ≥ -3), both data and Monte Carlo events are used in a manner similar to that used for the L3 efficiency calculation. First, a sample of $K_{\mu 3}$ decays is selected by requiring that (1) $m_{\pi\pi} > m_K - m_{\pi^0}$ to eliminate $K_L^0 \to \pi^+\pi^-\pi^0$ decays, (2) $m_{\pi\pi} < m_K - 10$ MeV/c^2 to eliminate $K_L^0 \to \pi^+\pi^-$ decays, and (3) one track passes pion identi-

TABLE IV. Level 3 efficiencies for QT events. The Monte Carlo efficiencies are calculated by convoluting L3 efficiencies for minimum bias $K_L^0 \to \pi^+\pi^-$ events binned by K_L^0 momentum, vertex z position, and opening angle, with Monte Carlo-generated distributions of these quantities for $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ decays.

Event sample	1989 efficiency	1990 efficiency
Monte Carlo $K_L^0 \to \pi^+\pi^-$	0.6439 ± 0.0028	0.8975 ± 0.0014
Data $K_L^0 \to \pi^+\pi^-$	0.6434 ± 0.0027	0.8988 ± 0.0014
Monte Carlo $K_L^0 \to \mu^+\mu^-$	0.6404 ± 0.0045	0.8875 ± 0.0026
Monte Carlo ratio $(K_L^0 \to \pi^+\pi^-)/(K_L^0 \to \mu^+\mu^-)$	$1.005(5) \pm 0.0071$	$1.011(3) \pm 0.0030$

fication criteria while the other track passes electron veto criteria. The pion identification criteria is that there is no signal in the Čerenkov counter, no signal in the MHO counter, no signal in the MRG within three gaps of the last gap expected based on track momentum, and E/p in the PbG is less than 0.5. The electron veto criteria is that there is no signal in the CER counter, and E/p is less than 0.5. The CER cut is relaxed if the track momentum is greater than the muon Čerenkov threshold of 6.2 GeV/c. For the MRG efficiency calculation, the muon candidate is also required to have an MHO signal with confidence level >0.0005, i.e., the same requirement as used to select $\mu^+\mu^-$ events. If both tracks satisfy pion identification criteria, the event is rejected; since both MHO and MRG detectors were $\gtrsim 97\%$ efficient, this rejects $\lesssim 3\%$ of "inefficient" muons. All events are subjected to the same event-quality cuts as were used to select $K_L^0 \to \mu^+\mu^-$ candidates.

The muon tracks are binned by momentum and x position, and for each bin the fraction of tracks passing the MHO and MRG particle identification cuts is found. A sample of Monte Carlo $K_L^0 \to \mu^+\mu^-$ events is then generated and used to find the fraction of $K_L^0 \to \mu^+\mu^-$ decays populating each pair of bins, one bin for each track. The pairing of bins accounts for correlation between the daughter particles. The overall $K_L^0 \to \mu^+\mu^-$ efficiency is calculated as the sum of the products of all "pair populations" with their respective efficiencies:

$$\epsilon_{\rm MHO} = \sum_{(i,j=\ \text{MHO bins})} f(i,j)\ \epsilon_i\,\epsilon_j\ ,$$

$$\epsilon_{\rm MRG} = \sum_{(k,\ell=\ \text{MRG bins})} f(k,\ell)\ \varepsilon_k\,\varepsilon_\ell\ ,$$

$$\epsilon_{\rm tot} = \sum_{(i,j=\ \text{MHO bins})}\ \sum_{(k,\ell=\ \text{MRG bins})} f(i,j,k,\ell)\ \epsilon_i\,\epsilon_j\,\varepsilon_k\,\varepsilon_\ell, \qquad (12)$$

TABLE V. Muon hodoscope efficiencies binned by track momentum and x position at the MHO, as found from 1990 $K_L^0 \to \pi\mu\bar{\nu}$ events (QT). The efficiencies found from 1989 data are very similar. The MHO cut is that the confidence level is >0.0005.

	x position at the MHO by left-side slat										
Momentum [GeV/c]	1	2	3	4	5	6	7	8	9	10	11
$p < 1.6$	–	1.0	0.857	0.791	0.803	0.772	0.771	0.792	0.677	0.857	–
$p > 1.6$	0.901	0.942	0.964	0.970	0.975	0.976	0.976	0.965	0.942	0.875	–
	x position at the MHO by right-side slat										
Momentum [GeV/c]	12	13	14	15	16	17	18	19	20	21	22
$p < 1.6$	–	1.0	0.803	0.804	0.800	0.823	0.824	0.721	0.800	0.500	–
$p > 1.6$	0.925	0.941	0.961	0.961	0.968	0.975	0.974	0.967	0.939	0.556	–

TABLE VI. MRG efficiencies binned by track momentum, as found from 1990 $K_L^0 \to \pi\mu\bar{\nu}$ events (QT). The efficiencies found from 1989 data are very similar. The MRG cut is that the (last gap expected − last gap hit) is ≥ -3.

Muon momentum range (GeV/c)										
1.5–1.6	1.6–2.0	2.0–2.5	2.5–3.0	3.0–3.5	3.5–4.0	4.0–4.5	4.5–5.0	5.0–5.5	5.5–6.0	>6.0
1.0000	1.0000	0.9996	0.9982	0.9968	0.9939	0.9944	0.9905	0.9874	0.9762	0.9735

where labels i and k refer to one daughter muon and labels j and ℓ refer to the other. The MHO and MRG bins are not identical: there are 22 MHO bins corresponding to two momentum ranges and 11 x positions, and 11 MRG bins corresponding to momentum only. The two momentum ranges for the MHO are $p < 1.6$ GeV/c and $p > 1.6$ GeV/c, where 1.6 GeV/c is approximately the momentum at which essentially all muons penetrate the 0.9 m deep iron wall just upstream of the MHO. The 11 MHO x positions correspond to the 11 x-measuring slats. All MHO and MRG bin efficiencies are listed in Tables V and VI.

The summations in Eq. (12) are performed as sums over Monte Carlo events:

$$\epsilon_{\rm tot} = \frac{1}{N_{\rm tot}} \sum_{\rm events} \epsilon_i \, \epsilon_j \, \varepsilon_k \, \varepsilon_\ell \ , \tag{13}$$

where $N_{\rm tot}$ is the total number of events generated. The statistical error on $\epsilon_{\rm tot}$ arises from the finite size of the $K_{\mu 3}$ sample used to measure the MHO and MRG bin efficiencies. The systematic error arises from uncertainty in the K_L^0 momentum spectrum and is taken as the variation in $\epsilon_{\rm tot}$ when $\epsilon_{\rm tot}$ is recalculated using different spectra (see Sec. V). The resulting $K_L^0 \to \mu^+\mu^-$ identification efficiencies are listed in Table VII. For comparison, the table also lists MHO efficiencies for several confidence level cuts more restrictive than 0.0005.

D. $\pi^+\pi^-$ interaction correction

In addition to trigger and particle-identification requirements, there was one additional source of bias between $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ decays: pions could interact strongly in detector elements, significantly altering their trajectories or being absorbed completely such that the event fails pattern recognition. To correct for this loss, we calculate the probability of either pion from $K_L^0 \to \pi^+\pi^-$ decay interacting and causing the event to be lost, and adjust the number of $K_L^0 \to \pi^+\pi^-$ events upwards by this amount.

We first measure the fraction of $K_L^0 \to \pi^+\pi^-$ decays in which one of the pions interacts somewhere downstream of the spectrometer but upstream of the second trigger counter (TSC2), causing the event to be lost. The material in this region consists of the last drift chamber DC5, the first trigger counter TSC1, and the Čerenkov counter. To measure this loss, we use a special pattern-recognition algorithm which begins at the front of the spectrometer and works its way downstream [36] rather than the standard algorithm which begins downstream with the TSC's and works its way upstream. The special algorithm in fact does not use TSC information at all. From a sample of K_{e3} and $K_{\mu 3}$ decays which pass the special algorithm, approximately 1.6% of the pion tracks fail the standard algorithm (i.e., the TSC2 slat to which the track projects is not hit); in contrast, less than 0.3% of the lepton tracks fail. The pion loss is measured independently for five ranges or bins of pion momentum and the total loss calculated as the difference between unity and the sum of the products of bin efficiencies × the fractional bin populations for $K_L^0 \to \pi^+\pi^-$ decay. The fractional populations are found from Monte Carlo simulation, and the resultant loss of $K_L^0 \to \pi^+\pi^-$ decays is $3.37 \pm 0.21\%$.

The problem remains of calculating the loss of pions due to interactions upstream of DC5. There is uncertainty arising from the probability of a ~ 3 GeV/c pion interacting in the amount of material traversed. We calculate the probability of an interaction using Particle Data Group values of total nuclear cross sections [3]. Although these were measured with 30–300 GeV/c neutrons [37] rather than low momentum pions, they include the coherent elastic and quasielastic components of strong scattering. Data with pion beams scattering off complex nuclei ($A \gtrsim 4$) do not include such contributions [38], as they contribute predominately in the forward direction and are difficult to disentangle from Coulomb scattering effects. Scattering data does show that pion inelastic cross sections ($\sigma_{\rm tot} - \sigma_{\rm elast} - \sigma_{\rm quasi}$) are systematically 20 − 30% lower than proton or neutron inelastic cross sections [38, 39]; using neutron cross sections from the PDG thus yields an upper bound on

TABLE VII. $K_L^0 \to \mu^+\mu^-$ particle-identification efficiencies (QT). The errors listed are statistical except for the second error in the last column, which results from uncertainty in the K_L^0 momentum spectrum. The MRG cut is that the (last gap expected − last gap hit) is ≥ -3.

Year	C.L.>0.0005	C.L.>0.005	C.L.>0.01	C.L.>0.1	MRG	(C.L.>0.0005) * MRG
1989	0.9309 ± 0.0022	0.9060	0.8892	0.7236	0.9865 ± 0.0013	$0.9184 \pm 0.0026^{+0.0001}_{-0.0002}$
1990	0.9256 ± 0.0016	0.8988	0.8815	0.7091	0.9847 ± 0.0012	$0.9112 \pm 0.0021^{+0.0005}_{-0.0009}$

pion loss upstream [40]. Table VIII lists the number of nuclear collision lengths comprised by the vacuum window, DC1– DC4, and the helium bags; the total is $0.00117 + 0.00232 + 0.00345 = 0.00694$, and the upper bound on $K_L^0 \to \pi^+\pi^-$ loss upstream is 0.0138.

We obtain a lower bound by scaling the 3.37% loss measured in the DC5 – CER region by the ratio of the number of nuclear collision lengths upstream of DC5 to the number of collision lengths in the measurement region. The difference between neutron and pion cross sections should cancel in this ratio. The scaling results in a lower bound because, due to track and vertex χ^2 cuts, pions which interact upstream in the spectrometer are much more likely to be rejected than pions which interact downstream. From Table VIII we find that the number of collision lengths comprised by DC5 + TSC1 + CER is 0.0492, and thus the lower bound is $(0.00694/0.0492) \times 0.0337 = 0.0048$. In the absence of more information we take the fraction of events lost upstream of DC5 to be the central value between upper and lower bounds: $(0.0138 + 0.0048)/2 = 0.0093$. We combine half the difference between the bounds in quadrature with the error on the 3.37% measurement to give the overall systematic error. The final result for $K_L^0 \to \pi^+\pi^-$ loss in the whole detector is $1-(1-0.0337)(1-0.0093) = 0.043 \pm 0.005$.

VII. $\mu^+\mu^-$ BRANCHING RATIO AND SYSTEMATIC ERRORS

To calculate the branching ratio for $K_L^0 \to \mu^+\mu^-$ we combine the event yields and efficiencies found previously as follows:

$$B_{\mu\mu} = \frac{N_{\mu\mu}}{P\,N_{\pi\pi}} \left(\frac{A_{\pi\pi}}{A_{\mu\mu}}\right) \left(\frac{1}{\varepsilon_{\mu\mu}^{L1}}\right) \times \left(\frac{\varepsilon_{\pi\pi}^{L3}}{\varepsilon_{\mu\mu}^{L3}}\right) \left(\frac{1}{\varepsilon_{\mu\mu}^{\mathrm{PID}}}\right) (\delta_{\mathrm{int}})\, B_{\pi\pi}\ , \tag{14}$$

TABLE VIII. Probability of a pion strongly interacting in detector elements. All nuclear collision lengths are from the Particle Data Group Book [3].

Material	Nuclear collision length λ_T (g/cm^2)	Mass traversed (g/cm^2)	Interaction probability
Vacuum window			
1989: Mylar (510 μm)	60.2	0.0707	0.00117
1990: Mylar + Kevlar (130 + 430 μm)	(60.2)	0.050	0.00083
Drift chambers			
Mylar (2 skins, 13 μm)	60.2	0.0035	0.00006
Al (2 skins, 13 μm)	70.6	0.0070	0.00010
Argon	76.4	0.0121	0.00016
Ethane	55.7	0.0100	0.00018
Ethanol	55.7	0.0005	0.00001
Al wires	70.6	0.0041	0.00006
Au plating	113.3	0.0016	0.00001
Total for each DC			0.00058
DC1 – DC4			0.00232
Helium bags			
Mylar (8 skins, 25 μm)	60.2	0.0288	0.00048
Helium	49.9	0.1481	0.00297
Total			0.00345
Trigger counters			
Scintillator	58.4	1.0400	0.01785
Tape	60.2	0.0128	0.00021
Marvelguard (carbon)	60.2	0.2180	0.00362
G10 (fiberglass)	62.6	0.1290	0.00205
Aluminum support	70.6	0.1609	0.00227
Air in support	62.0	0.0073	0.00012
Total for each TSC			0.02622
Čerenkov			
Front Al panel	70.6	0.2150	0.0030
Helium	49.9	0.0330	0.0007
Nitrogen	61.4	0.1520	0.0025
Mirrors	59.2	0.3812	0.0064
Back Al panel	70.6	0.6939	0.0098
Total			0.0224

where $N_{\mu\mu}$ is the number of $\mu^+\mu^-$ candidates observed, $N_{\pi\pi}$ is the number of $\pi^+\pi^-$ candidates, P is the total $\pi^+\pi^-$ prescale factor, $A_{\pi\pi}/A_{\mu\mu}$ is the ratio of $\pi^+\pi^-$ and $\mu^+\mu^-$ acceptances, $\varepsilon_{\mu\mu}^{L1}$ is the $\mu^+\mu^-$ L1 trigger efficiency, $(\varepsilon_{\pi\pi}^{L3})/(\varepsilon_{\mu\mu}^{L3})$ is the ratio of $\pi^+\pi^-$ and $\mu^+\mu^-$ L3 efficiencies, $\varepsilon_{\mu\mu}^{\rm PID}$ is the $\mu^+\mu^-$ particle-identification efficiency, $\delta_{\rm int}$ corrects for the fraction of $K_L^0 \to \pi^+\pi^-$ decays lost due to pion interactions, and $B_{\pi\pi}$ is the branching ratio for $K_L^0 \to \pi^+\pi^-$ taken from the Particle Data Group [3]. All values for the 1989 and 1990 data sets (QT) are listed in Table IX along with their measurement errors. Inserting these values into Eq. (14) yields

$$B(K_L \to \mu^+\mu^-)_{\rm QT} = \begin{cases} (7.58 \pm 0.48 \pm 0.28) \times 10^{-9} & (1989)\,, \\ (6.87 \pm 0.40 \pm 0.22) \times 10^{-9} & (1990)\,, \end{cases} \quad (15)$$

where the first error is statistical and the second systematic. The statistical error results from the number of $\mu^+\mu^-$ events observed, while the systematic error results from propagating the systematic errors on acceptances, efficiencies, and backgrounds. The FT analysis gives

$$B(K_L \to \mu^+\mu^-)_{\rm FT} = \begin{cases} (7.32 \pm 0.48 \pm 0.28) \times 10^{-9} & (1989)\,, \\ (6.70 \pm 0.40 \pm 0.20) \times 10^{-9} & (1990)\,. \end{cases} \quad (16)$$

There is a $\sim 3\%$ difference between the QT and FT results. This difference is due to the number of $\pi^+\pi^-$ candidates observed and indicates a difference in the degree in which pion decays are rejected. Because Monte Carlo events must also pass fitting cuts, the loss of $K_L^0 \to \pi^+\pi^-$ events due to pion decay is nominally accounted for in the $K_L^0 \to \pi^+\pi^-$ acceptance correction. The fact that a small difference remains after the correction indicates possible systematic error in the Monte Carlo simulation. We incorporate this error by taking our measurement to be the central value between QT and FT and adding half the difference in quadrature to the other systematic errors. Our final result is

$$B(K_L \to \mu^+\mu^-) = \begin{cases} (7.45 \pm 0.48 \pm 0.34) \times 10^{-9} & (1989)\,, \\ (6.78 \pm 0.40 \pm 0.23) \times 10^{-9} & (1990)\,. \end{cases} \quad (17)$$

To check for other systematic effects in the data, we calculate the branching ratio for independent subsets of events: "inbend" events (positive track on right), "outbend" events (positive track on left), and different ranges of K_L^0 momentum, vertex z position, and run number (i.e., running time). The results are shown in Fig. 17 along with statistical errors. All variations are consistent within the errors and no obvious systematic trends stand out.

The dependence of the measured branching ratio on the particular choice of cuts is studied by calculating the branching ratio repeatedly as one cut is varied while all others are held fixed. The range over which the cut is varied includes both sides of the value used in the analysis. The resultant change, if any, in $B(K_L^0 \to \mu^+\mu^-)$ is examined for systematic increases or decreases. If a systematic change is found, an additional systematic error is included in the measurement of $B(K_L^0 \to \mu^+\mu^-)$ to account for uncertainty due to the cut. The procedure is repeated for each of 15 cuts used to select the $\mu^+\mu^-$ and $\pi^+\pi^-$ event samples (Table X).

TABLE IX. Values (QT) used for the overall likelihood fit for $B(K_L^0 \to \mu^+\mu^-)$. When two errors are listed, the first is an uncorrelated error while the second is correlated with another error(s), usually on the same quantity but for another year.

Factor	1988	1989	1990
$\mu^+\mu^-$ events in signal region	87	292	370
Background from $K_L^0 \to \pi\ell\bar{\nu}$	0.3	18.5 ± 4.6	24.0 ± 4.7
$\pi^+\pi^-$ prescale	6000	6000	4000
$K_L^0 \to \pi^+\pi^-$ candidates	8226 ± 148	$15245 \pm 157 \pm 80$	$32316 \pm 233 \pm 90$
$A_{\pi\pi}/A_{\mu\mu}$	$1.18 \pm 0.012 \pm 0.055$	$1.173 \pm 0.015^{+0.004}_{-0.016}$	$1.163 \pm 0.013^{+0.004}_{-0.014}$
$\epsilon_{\mu\mu}^{L1}$	0.985 ± 0.015	0.984 ± 0.005	0.977 ± 0.004
$\epsilon_{\pi\pi}^{L3}/\epsilon_{\mu\mu}^{L3}$	1.26 ± 0.03[a]	$1.005(5) \pm 0.0071$	$1.011(3) \pm 0.0030$
$\epsilon_{\mu\mu}^{\rm PID}$	0.928 ± 0.032	$0.9184 \pm 0.0026^{+0.0001}_{-0.0002}$	$0.9112 \pm 0.0021^{+0.0005}_{-0.0009}$
$B(K_L^0 \to \mu^+\mu^-) \times 10^9$	$5.6 \pm 0.6 \pm 0.4$	$7.58 \pm 0.48 \pm 0.28$	$6.87 \pm 0.40 \pm 0.22$

[a]In this analysis the L3 algorithm was not applied to $\pi^+\pi^-$ events.

Figures 18(a) and 18(b) show the change in $B(K_L^0 \to \mu^+\mu^-)$ as the MHO confidence level and MRG gap difference cuts, respectively, are varied. The small changes seen are consistent with statistical fluctuations and do not exhibit systematic trends. Of the 15 cuts studied in this manner, only three exhibit possible trends: the θ_K^2 cut for 1989 data (FT and QT), the mass window cut defining the signal region for 1990 data (QT only), and the track χ^2 cut for 1989 data (FT and QT) and 1990 data (FT only). Figure 18(c) shows the change in the 1989 FT branching ratio as the θ_K^2 cut is varied from a tight cut of 1.5×10^{-6} to a loose cut of 2.5×10^{-6} (the cut used to select $K_L^0 \to \mu^+\mu^-$ and $K_L^0 \to \pi^+\pi^-$ candidates is 2.0×10^{-6}). For the tight cut the branch-

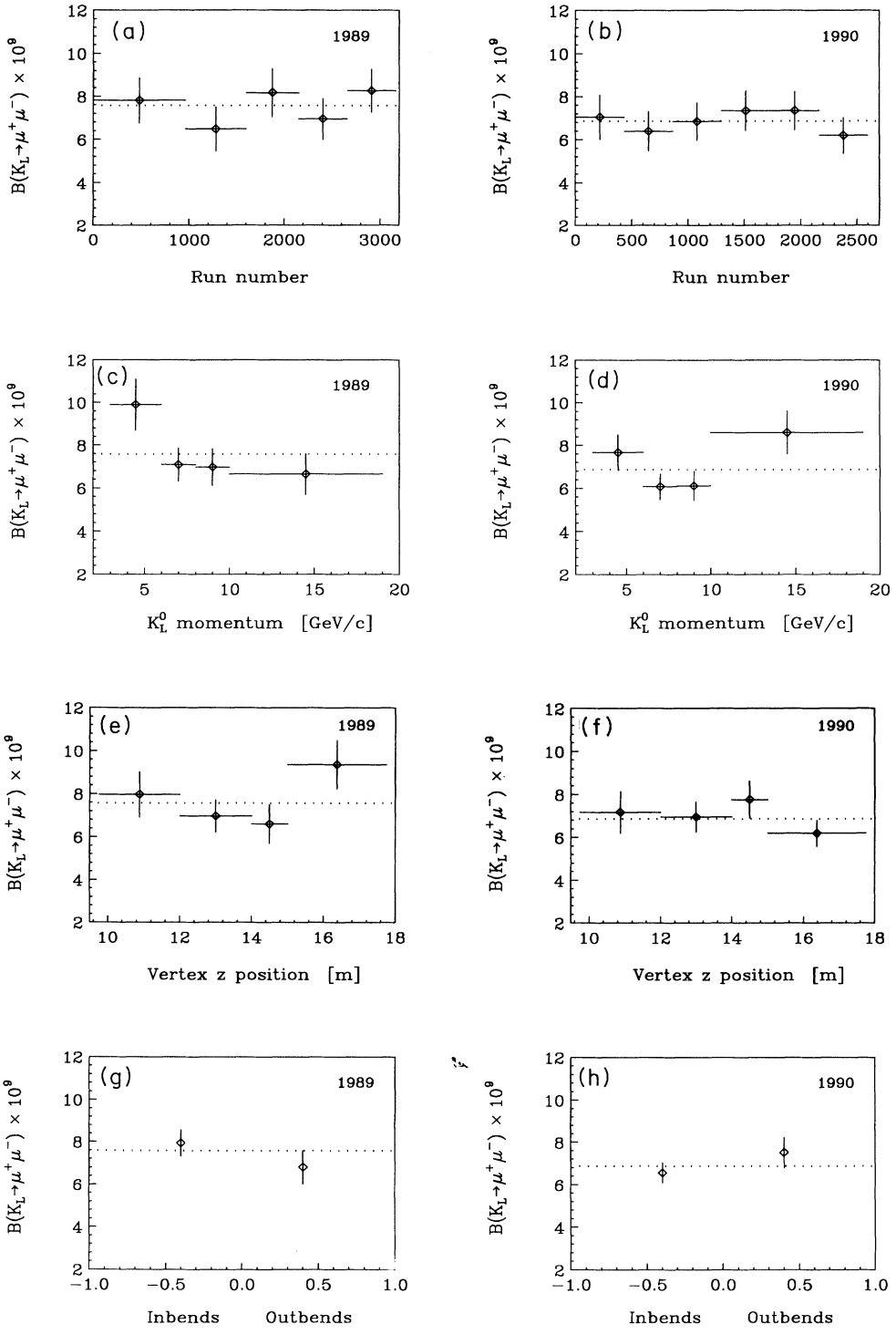

FIG. 17. $B(K_L^0 \to \mu^+\mu^-)$ plotted for independent subsets of data. All measurements are consistent within statistical errors, and no obvious systematic trends stand out. The independent subsets are: (a),(b) different ranges of run numbers, (c),(d) different ranges of K_L^0 momentum, (e),(f) different ranges of vertex z positions, and (g),(h) "inbend" and "outbend" events (positive track on right or left side of detector). The overall measured values $B_{\mu\mu} = 7.58$ (1989 QT) and $B_{\mu\mu} = 6.87$ (1990 QT) are indicated by dotted horizontal lines.

ing ratio rises by 1.8%, while for the loose cut it falls by 1.4%. We choose to include as a systematic error the fractional change which occurs in the direction *away* from the branching ratio of the other fitter; we thus add 1.4% in quadrature to the other systematic errors on $B(K_L^0 \to \mu^+\mu^-)$. Figure 18(d) shows the change in the 1990 QT branching ratio as the mass window cut is varied from a tight cut of 5 MeV/c^2 to a loose cut of 7 MeV/c^2; there is a 1.1% decrease and a 1.9% increase, respectively. In this case we include the 1.1% change as a systematic error as it occurs in the more physically sensible direction of the cut (a 7 MeV/c^2 cut is much looser than needed for high efficiency and includes significantly more $K_{\mu 3}$ background, which is difficult to subtract). Both the θ_K^2 systematic error and the mass window systematic error have been included in Eqs. (15), (16), and (17).

The results for the track χ^2 cuts are shown in Figs. 19(a) (FT) and 19(b) for FT and Figs. 19(c) and 19(d) for QT. The figures show that when the FT track χ^2 cut is tightened from its nominal value, the branching ratio rises towards the higher value of QT. When the QT track χ^2 cut is tightened from its nominal value, the branching ratio (1989) decreases towards the lower value of FT. When FT and QT χ^2 cuts are loosened, no significant change in the branching ratio occurs. Also, when the FT momentum matching cut between upstream and downstream spectrometers is tightened, the branching ratio rises towards QT; when it is loosened, no change occurs. These trends suggest that $B(K_L^0 \to \mu^+\mu^-)$ probably lies between FT and QT values, which is consistent with our taking the central value as our measurement and including half the difference as a systematic error. This error is equivalent to (and substitutes for) including as errors the trends observed in the χ^2 plots.

The final task remaining is to combine 1989 and 1990 results and also to incorporate our previous 1988 result [20] When combining results, care must be taken to correctly account for both correlated and uncorrelated systematic errors: the former include errors such as those arising from the K_L^0 momentum spectrum, while the lat-

TABLE X. Cuts used to select $K_L^0 \to \mu^+\mu^-$ candidate events and the systematic change in $B(K_L^0 \to \mu^+\mu^-)$ observed as cuts were individually varied (see Sec. VII in text). For most cuts no systematic change was observed, only statistical fluctuations.

Cut	Nominal value	Range of variation	Change in $B(K_L^0 \to \mu^+\mu^-)$ 1989	1990
Vertex $\|z\|$	> 9.75 m	9.5 – 10.5	FT: – QT: –	– –
Vertex $\|x/z\|$	< 0.0027	0.0020 – 0.0030	FT: – QT: –	– –
Vertex $\|y/z\|$	< 0.0100	0.0050 – 0.0120	FT: – QT: –	– –
$\|x\|$ at vacuum window (1989 only)	> 0.1069 m	0.0979 – 0.1169 m	FT: – QT: –	n/a n/a
$\|r\|$ at vacuum window (1990 only)	< 0.4794 m	0.4704 – 0.4894 m	FT: n/a QT: n/a	– –
$\|x\|$ at PbG array	> 0.273 m (1989) > 0.303 m (1990)	0.253 – 0.308 m 0.263 – 0.343 m	FT: – QT: –	– –
Track momentum	> 1.5 GeV/c	1.0 – 2.5 GeV/c	FT: – QT: –	– –
$(p_+ - p_-)/(p_+ + p_-)$	< 0.66	0.31 – 0.76	FT: – QT: –	– –
$\|p_{\text{upstrm}} - p_{\text{downstrm}}\|/(p_{\text{upstrm}} + p_{\text{downstrm}})$	< 0.025	0.010 – 0.035	FT: – QT: –	– –
Track χ^2	< 200 (FT) < 50 (QT)	100 – 300 20 – 80	FT: +3.7% QT: −1.6%	+1.9% –
Vertex χ^2	< 18 (FT) < 25 (QT)	10 – 30 11 – 41	FT: – QT: –	– –
$\|m_{\mu\mu} - m_K\|$	< 6 MeV/c^2	5 – 7 MeV/c^2	FT: – QT: –	– $^{+1.9\%}_{-1.1\%}$
θ_K^2	< 2 mrad^2	1.5 – 2.5 mrad^2	FT: $^{+1.8\%}_{-1.4\%}$ QT: $^{+1.3\%}_{-2.0\%}$	– –
MHO confidence level	> 0.0005	0.0005 – 0.1	FT: – QT: –	– –
MRG gap difference	≥ -3	(−4) – (−2)	FT: – QT: –	– –

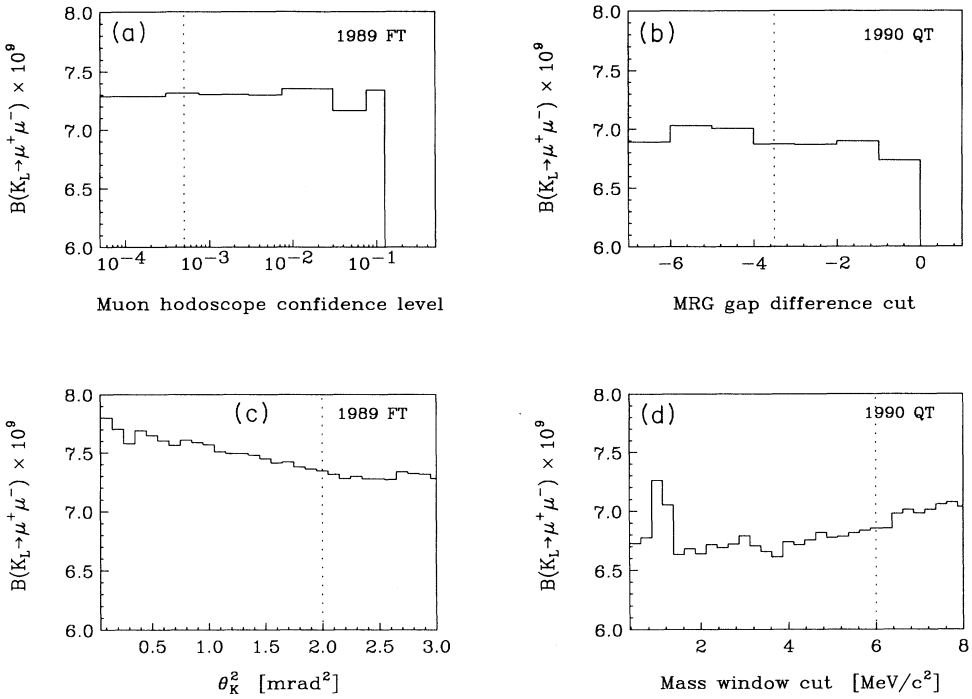

FIG. 18. $B(K_L^0 \to \mu^+\mu^-)$ plotted as a function of: (a) MHO confidence level cut, (b) MRG gap difference cut, (c) θ_K^2 cut, and (d) mass window cut (width of signal region). The values of the cuts used in the analysis are indicated by dotted vertical lines. For (c) and (d) potential trends are seen, and the variations in $B(K_L^0 \to \mu^+\mu^-)$ are included as systematic errors in the measurement (see text).

ter include errors such as those arising from MHO and MRG bin efficiencies. To account for these and all other errors we combine results by doing a single maximum likelihood fit to data from all three years. For one year, the relative likelihood of the branching ratio $B_{\mu\mu}$ is the Poisson probability

$$\mathcal{L}(B_{\mu\mu}) \equiv \frac{\alpha^N e^{-\alpha}}{N!}, \tag{18}$$

where N is the number of $K_L^0 \to \mu^+\mu^-$ candidates observed and α is the number expected:

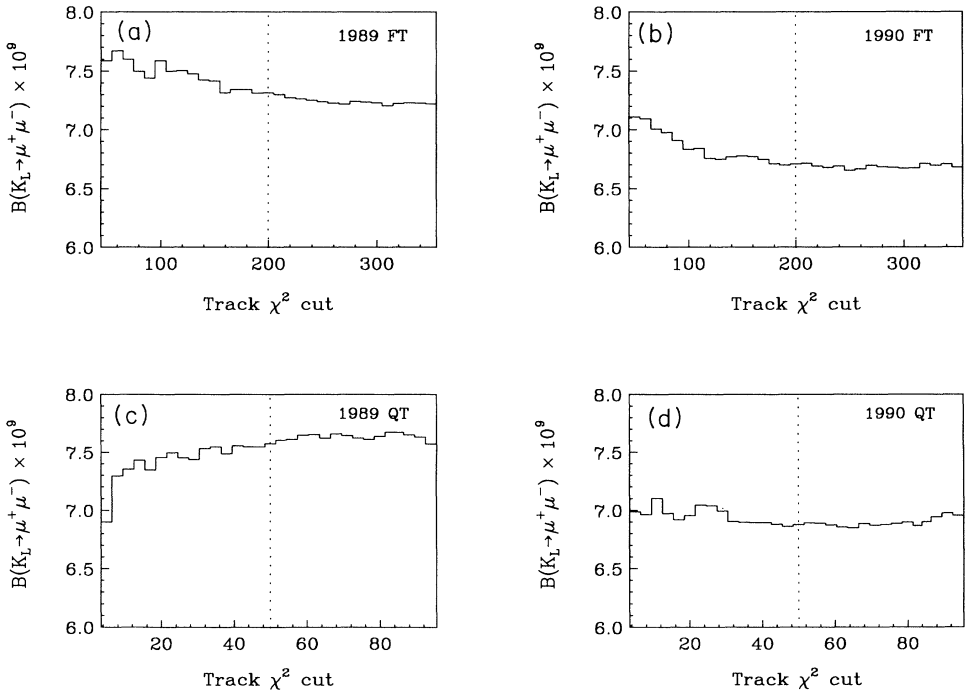

FIG. 19. $B(K_L^0 \to \mu^+\mu^-)$ plotted as a function of FT track χ^2 cut (a),(b) and QT x view track χ^2 cut (c),(d). The values of the cuts used in the analysis are indicated by dotted vertical lines.

$$\alpha = \left(\frac{B_{\mu\mu}}{B_{\pi\pi}}\right)\left(\frac{N_{\pi\pi}}{\delta_{\rm int}}\right)\left(\frac{A_{\mu\mu}}{A_{\pi\pi}}\right)\varepsilon_{\mu\mu}^{L1}\left(\frac{\varepsilon_{\mu\mu}^{L3}}{\varepsilon_{\pi\pi}^{L3}}\right)\varepsilon_{\mu\mu}^{\rm PID}. \tag{19}$$

The parameters $B_{\pi\pi}$, $N_{\pi\pi}$, $\delta_{\rm int}$, $(A_{\mu\mu}/A_{\pi\pi})$, $\varepsilon_{\mu\mu}^{L1}$, $(\varepsilon_{\mu\mu}^{L3}/\varepsilon_{\pi\pi}^{L3})$, and $\varepsilon_{\mu\mu}^{\rm PID}$ are as defined previously in Eq. (14). The uncertainties in these parameters are incorporated into the likelihood function by smearing (i.e., integrating) $\mathcal{L}(B_{\mu\mu})$ over the normalized distribution function for each parameter. The distribution functions are taken to be Gaussian with means equal to the parameters' measured values and standard deviations equal to the measurement errors (see Table IX). When the errors are asymmetric, the distribution functions are taken to be asymmetric. All integrals are evaluated numerically by sampling over the functions, and $\mathcal{L}$ is found to be maximum at the $B_{\mu\mu}$ calculated using Eq. (14), as it should be.

To incorporate data from all three years, we multiply the right-hand-side of Eq. (18) by the Poisson probabilities corresponding to the event yields and efficiencies from the other two years. We then sample over distribution functions corresponding to 21 separate efficiencies, acceptances, and backgrounds. Those errors which are correlated are sampled identically, i.e., using the same random number such that a fluctuation in one corresponds to an identical fluctuation in the other.

The negative of the logarithms of the likelihood functions for the three individual years and for all three years combined are shown in Fig. 20 for QT; the FT functions look similar. The 1989 and 1990 QT and FT functions yield maximum likelihood estimates for $B(K_L^0 \to \mu^+\mu^-)$ equal to the values listed in (15) and (16). The likelihood functions for $1988+1989+1990$ data combined yield maximum likelihood estimates for the branching ratio ($\times 10^9$) of 6.95 (QT) and 6.77 (FT). Defining upper and lower limits for a 68.3% confidence interval as those values where $-\ln\mathcal{L}$ rises by 0.5, we obtain measurements $B_{\mu\mu} = 6.95^{+0.34}_{-0.33}\ 10^{-9}$ (QT) and $B_{\mu\mu} = 6.77^{+0.33}_{-0.32}\ 10^{-9}$ (FT). These confidence intervals include both statistical and systematic errors. We take the central value between QT and FT results as our final result and include half the difference between them as an additional systematic error which we add in quadrature to the other errors:

$$B(K_L^0 \to \mu^+\mu^-) = (6.86 \pm 0.37) \times 10^{-9} \quad (1988+1989+1990). \tag{20}$$

This result is somewhat lower than that of experiment KEK-137, which measured $B(K_L^0 \to \mu^+\mu^-) = (7.9 \pm 0.6 \pm 0.3) \times 10^{-9}$ [19]. If we do not integrate out the parameter $B_{\pi\pi}$ in Eq. (19), we obtain a likelihood function for the ratio of partial widths $\Gamma_{\mu\mu}/\Gamma_{\pi\pi}$. The maximum likelihood estimate for the ratio and the one standard deviation errors are:

$$\frac{\Gamma(K_L^0 \to \mu^+\mu^-)}{\Gamma(K_L^0 \to \pi^+\pi^-)} = (3.38 \pm 0.17) \times 10^{-6} \quad (1988+1989+1990). \tag{21}$$

VIII. THEORETICAL DISCUSSION

The measurement of $B(K_L^0 \to \mu^+\mu^-)$ is important within the framework of the standard model because it and related processes involving internal quark loops are the only way at present to obtain information about V_{td}, the CKM matrix element coupling t and d quarks. Here

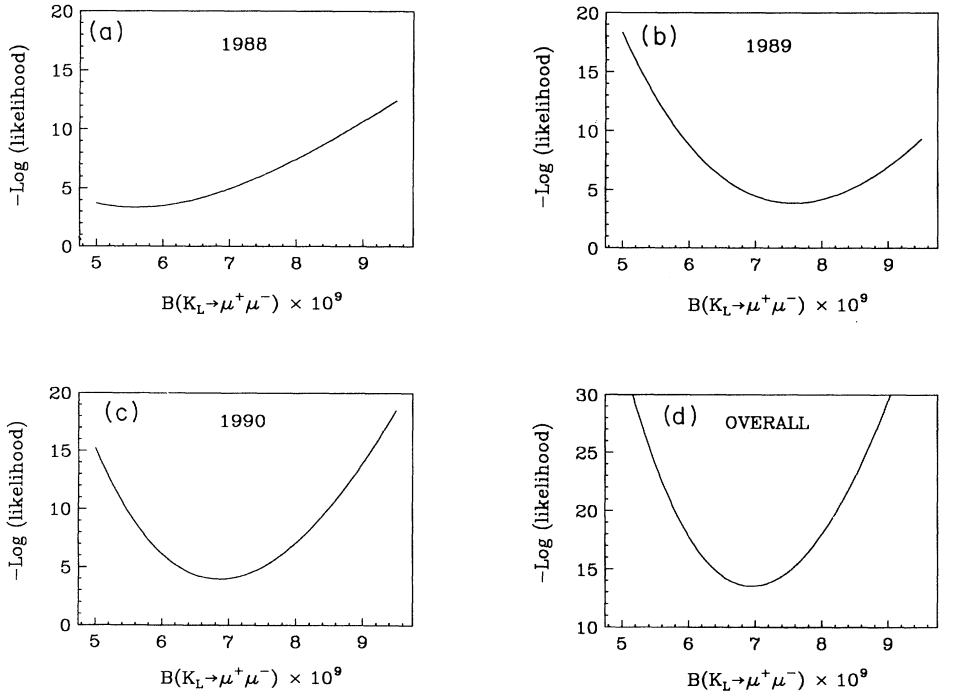

FIG. 20. Negative-log-likelihood functions for $B(K_L^0 \to \mu^+\mu^-)$ corresponding to: (a) previously published 1988 data, (b) 1989 data, (c) 1990 data, and (d) data from all three years combined. The functions shown are for QT data; the FT functions look similar. The maximum likelihood estimates for $B_{\mu\mu}$ are 5.59, 7.58, 6.87, and 6.95, respectively; the maximum likelihood estimate for FT data for all three years combined is 6.77.

we present a calculation showing how such information is extracted from the measured value of the branching ratio; similar calculations can be found in the literature [10–14, 41]. Our calculation differs from those previously published in that we use (1) our overall measurement of $B(K_L^0 \to \mu^+\mu^-)$, (2) recent higher-order calculations of the charm and top quark contributions to the short-distance amplitude, (3) a new measurement of $B(K_L^0 \to \mu^+\mu^-\gamma)$ from which, within one theoretical model, the long-distance amplitude for $K_L^0 \to \mu^+\mu^-$ can be extracted, and (4) the unitarity bound for the ratio of partial widths $\Gamma(K_L^0 \to \mu^+\mu^-)/\Gamma(K_L^0 \to \pi^+\pi^-)$, which has less uncertainty than the unitarity bound for the branching ratio itself.

To extract information from $B(K_L^0 \to \mu^+\mu^-)$ about the t quark requires an understanding of the different contributions to the amplitude $\mathcal{A}(K_L^0 \to \mu^+\mu^-)$. Denoting the real or dispersive part as $\mathrm{Re}\mathcal{A}$ and the imaginary or absorptive part as $\mathrm{Im}\mathcal{A}$ allows the branching ratio to be written

$$B(K_L^0 \to \mu^+\mu^-) \equiv |\mathrm{Re}\mathcal{A}|^2 + |\mathrm{Im}\mathcal{A}|^2 , \tag{22}$$

where $\mathcal{A}$ implicitly includes the K_L^0 lifetime and phase-space factors. The absorptive part $\mathrm{Im}\mathcal{A}$ is almost exclusively due to $K_L^0 \to \gamma\gamma \to \mu^+\mu^-$, which alone gives a branching ratio (or unitarity bound) very close to the measured value. This leaves little room for $|\mathrm{Re}\mathcal{A}|^2$.

$\mathrm{Re}\mathcal{A}$ receives contributions from short-distance electroweak diagrams dominated by the t quark, and also from long-distance electromagnetic processes:

$$\mathrm{Re}\mathcal{A} = \mathcal{A}_{\text{weak}} + \mathcal{A}_{\text{long dist}}. \tag{23}$$

The short-distance weak contribution is calculated within the framework of the standard model [7] and arises from K^0 and $\bar{K}^0$ decay via the loop diagrams of Fig. 1. Because the K_L^0 is a coherent mixture of K^0 and $\bar{K}^0$, the K^0, $\bar{K}^0$ decay amplitudes must be summed in order to obtain the K_L^0 decay amplitude; the result is proportional to the purely real quantity $V_{td}^* V_{ts} + V_{td} V_{ts}^* = 2\mathrm{Re}(V_{td}^* V_{ts})$. This is in contrast to the closely related decay $K^+ \to \pi^+\nu\bar{\nu}$, whose amplitude is proportional to the complex quantity $V_{td}^* V_{ts}$.

The long-distance two-photon contribution in Eq. (23) is nonperturbative and difficult to calculate reliably. This uncertainty in $\mathcal{A}_{\text{long dist}}$ complicates extracting quantitative information about $\mathrm{Re}(V_{td}^* V_{ts})$ and m_t (arising from the internal t propagator) from the measured value of $B(K_L^0 \to \mu^+\mu^-)$. Two different theoretical treatments of $\mathcal{A}_{\text{long dist}}$ have appeared in the literature and are briefly reviewed in Appendix A. The calculation of Bergström *et al.* [11, 12] results in the constraint:

$$-2.9 \times 10^{-5} \;\le\; \mathcal{A}_{\text{long dist}} \;\le\; 0.50 \times 10^{-5} , \tag{24}$$

while the calculation of Ko [14] gives

$$0.27 \times 10^{-5} \;\le\; \mathcal{A}_{\text{long dist}} \;\le\; 4.7 \times 10^{-5}. \tag{25}$$

By subtracting the unitarity bound from the measured value of $B(K_L^0 \to \mu^+\mu^-)$, we can determine $|\mathrm{Re}\mathcal{A}|$. Subtracting $\mathcal{A}_{\text{long dist}}$ from this value then constrains $\mathcal{A}_{\text{weak}}$, and the constraint on $\mathcal{A}_{\text{weak}}$ gives a constraint in the m_t-$|\mathrm{Re}(V_{td}^* V_{ts})|$ plane.

To proceed, we subtract from our result (21) the unitarity bound for the ratio of partial widths

$$\Gamma(K_L^0 \to \mu^+\mu^-)/\Gamma(K_L^0 \to \pi^+\pi^-) .$$

The unitarity bound for the ratio is

$$\frac{\Gamma(K_L^0 \to \gamma\gamma \to \mu^+\mu^-)}{\Gamma(K_L^0 \to \pi^+\pi^-)} = \left[\frac{\Gamma(K_L^0 \to \gamma\gamma \to \mu^+\mu^-)}{\Gamma(K_L^0 \to \gamma\gamma)}\right] \times \left[\frac{\Gamma(K_L^0 \to \gamma\gamma)}{\Gamma(K_L^0 \to \pi^0\pi^0)}\right] \times \left[\frac{\Gamma(K_L^0 \to \pi^0\pi^0)}{\Gamma(K_L^0 \to \pi^+\pi^-)}\right] , \tag{26}$$

$$= \left[1.20 \times 10^{-5}\right] \left[0.632 \pm 0.004 \pm 0.008\right] \left[0.4518 \pm 0.0066\right] , \tag{27}$$

$$= (3.43 \pm 0.07) \times 10^{-6}, \tag{28}$$

where the first term in square brackets results from a QED calculation [4, 5] and the latter two terms are measured [3]. The 2% error in Eq. (28) is less than half the error on the unitarity bound for the branching ratio, which is calculated using the measured value of $B(K_L^0 \to \gamma\gamma)$. Subtracting (28) from (21) gives the result:

$$\frac{|\mathrm{Re}\mathcal{A}(K_L^0 \to \mu^+\mu^-)|^2}{B(K_L^0 \to \pi^+\pi^-)} = (-0.05 \pm 0.18) \times 10^{-6} \tag{29}$$

or, multiplying by

$$B(K_L^0 \to \pi^+\pi^-) = (2.03 \pm 0.04) \times 10^{-3} , \tag{30}$$

yields

$$|\mathrm{Re}\mathcal{A}|^2 = (-1.0 \pm 3.7) \times 10^{-10}. \tag{31}$$

From this result we determine an upper limit for $|\mathrm{Re}\mathcal{A}|^2$. Since the central value lies in a negative or unphysical region, we follow the Bayesian prescription of the Particle Data Group [3] and multiply a Gaussian likelihood function for $|\mathrm{Re}\mathcal{A}|^2$ by a prior probability density for $|\mathrm{Re}\mathcal{A}|^2$ which is constant but set to zero for negative values. The resultant posterior probability density for $|\mathrm{Re}\mathcal{A}|^2$ then vanishes in the unphysical region. We integrate this function upwards to find the value of $|\mathrm{Re}\mathcal{A}|^2$ which bounds 90% of the area; this value is taken as the 90% C.L. upper limit. We obtain

$$|\mathrm{Re}\mathcal{A}|^2 \;<\; 5.6 \times 10^{-10} \quad (90\%\ \text{C.L.}). \tag{32}$$

The magnitude of $\mathcal{A}_{\rm weak}$ can be largest if $\mathcal{A}_{\rm weak}$ and $\mathcal{A}_{\rm long\ dist}$ interfere destructively. In this case, inequalities (24) and (32) imply (for an upper limit only):

$$\left(|\mathcal{A}_{\rm weak}| - 2.9\times10^{-5}\right)^2 < 5.6\times10^{-10} \tag{33}$$

or

$$|\mathcal{A}_{\rm weak}|^2 < 2.8\times10^{-9}. \tag{34}$$

Within the standard model [7],

$$|\mathcal{A}_{\rm weak}|^2 = B(K^+\to\mu^+\nu_\mu)\frac{\tau(K_L^0)}{\tau(K^+)}\left(\frac{\alpha^2}{4\pi^2\sin^4\theta_W}\right)\frac{(1-4m_\mu^2/m_K^2)^{1/2}}{(1-m_\mu^2/m_K^2)^2}\frac{|\mathrm{Re}\sum_i V_{is}^*V_{id}C(x_i)|^2}{|V_{us}|^2}, \tag{35}$$

where the summation i ranges over u, c, and t quarks, the variable $x_i=(m_i/m_W)^2$, and

$$C(x_i)=\frac{x_i}{4}\left[\frac{4-x_i}{1-x_i}+\frac{3x_i\ln x_i}{(1-x_i)^2}\right]. \tag{36}$$

Because of the small value of m_u, only c and t quarks contribute appreciably to the sum. The function $C(x_i)$ undergoes QCD corrections which have recently been calculated to $O(\alpha_s)$ for the t quark [42] and within a next-to-leading-log approximation to all orders for the charm quark [41]. We denote these corrected functions by $C_t(x_t)$ and $C_c(x_c)$, respectively. Expressing CKM matrix elements in terms of the Wolfenstein parameters A, ρ, and $\lambda=0.22$ [43], Eq. (35) is rewritten [41]

$$|\mathcal{A}_{\rm weak}|^2 = (1.71\times10^{-9})A^4|C_t(x_t)|^2 \times\left[1-\rho+\frac{417\,C_c(x_c)}{A^2\,C_t(x_t)}\right]^2, \tag{37}$$

where $\alpha = 1/128$, $\sin^2\theta_W = 0.23$, $\tau_{K^+} = 12.4$ ns, $\tau_{K_L}=51.7$ ns, $B(K^+\to\mu^+\nu)=0.635$, and a factor of 2 has been absorbed into the definitions of C_c and C_t. The third term in square brackets results exclusively from the charm quark and for $m_c=1.4$ GeV/c^2 (renormalization scale $\mu=m_c$) [44] and $m_t=170$ GeV/c^2 ($\mu=m_t$) has magnitude 0.23; the charm contribution is thus non negligible. Combining Eq. (37) with Eq. (34) and taking $m_c=1.4$ GeV/c^2 gives a constraint among A, ρ, and m_t:

$$A^4|C_t(x_t)|^2\left[1-\rho+\frac{0.155}{A^2C_t(x_t)}\right]^2 < 1.62. \tag{38}$$

This constraint corresponds to our 90% C.L. upper limit on $|\mathrm{Re}\mathcal{A}|^2$. The parameter $A=V_{cb}/\lambda^2$ can be determined with relatively small theoretical uncertainty by applying heavy quark effective theory [45] to the measured rate of $\bar{B}\to D^*\ell\bar{\nu}$; a recent average over ALEPH, ARGUS, and CLEO results [46] gives $|V_{cb}| = 0.040\pm0.0025\pm0.002$ or $A = 0.83\pm0.07$. Inserting this into Eq. (38) gives a lower limit on ρ as a function of m_t. This limit is plotted in Fig. 21(a) for $A=0.76$, 0.83, and 0.90 and represents a unique constraint on CKM unitarity. If $\mathcal{A}_{\rm long\ dist}$ is constrained to lie in the range given in Eq. (25), we find constraints upon m_t and ρ as shown in Fig. 21(b). The plot corresponding to $\mathcal{A}_{\rm long\ dist}$ taken from Bergström *et al.* [11, 12] shows that for $A=0.83$ and $m_t=174$ GeV/c^2 [as indicated by data from the Collider Detector at Fermilab (CDF) Collaboration [47]], $\rho > -0.52$. If we repeat the above analysis using the weighted average of our branching ratio and that of experiment KEK-137, we find very similar limits: $|\mathrm{Re}\mathcal{A}|^2 < 6.5\times10^{-10}$ at 90% C.L. and for $m_t=174$ GeV/c^2, $\rho > -0.58$.

IX. CONCLUSIONS

We have measured $B(K_L^0\to\mu^+\mu^-)$ using our full data set collected over three years of running. There are 749 events in our signal region which pass all analysis cuts. Of these, 42.5 events are estimated to be background, leaving 707 $K_L^0\to\mu^+\mu^-$ candidates. Our result for the branching ratio, $(6.86\pm0.37)\times10^{-9}$, has a total uncertainty of 5.4% divided between a 3.7% statistical error and a 4.0% systematic error. The systematic error has the following components: a 2.0% error in the ratio of relative acceptances for $K_L^0\to\pi^+\pi^-$ and $K_L^0\to\mu^+\mu^-$,

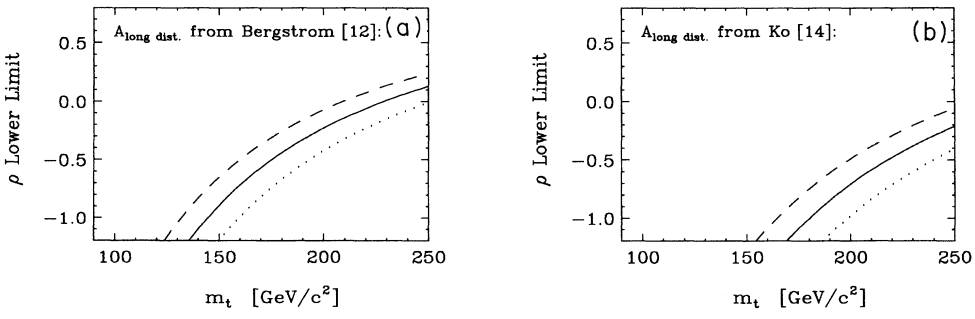

FIG. 21. The lower limit on the Wolfenstein parameter ρ corresponding to our 90% C.L. upper limit on $|\mathrm{Re}\mathcal{A}|^2$, as a function of top mass. The dotted contour corresponds to $A = 0.76$, the solid contour to $A = 0.83$, and the dashed contour to $A = 0.90$. The long-distance dispersive contribution to $\mathcal{A}(K_L^0\to\mu^+\mu^-)$ is taken from the theoretical calculation of: (a) Bergström *et al.* [12], and (b) Ko [14].

a 0.8% error in the $\pi^+\pi^-$ background estimate, a 0.5% error in the pion interaction correction, a 0.4% error in the $\mu^+\mu^-$ L1 trigger efficiency, a 0.3% error in the ratio of $\mu^+\mu^-$ and $\pi^+\pi^-$ L3 efficiencies, a 0.3% error in the $\mu^+\mu^-$ particle-identification efficiency, and a 2.0% error in the branching ratio for $K_L^0 \to \pi^+\pi^-$ as taken from the Particle Data Group [3]. Without the error on $B(K_L^0 \to \pi^+\pi^-)$, the total systematic error is 3.4%.

Our measurement of the ratio of partial widths $\Gamma(K_L^0 \to \mu^+\mu^-)/\Gamma(K_L^0 \to \pi^+\pi^-)$ is $(3.38 \pm 0.17) \times 10^{-6}$. Subtracting off the unitarity bound from our measurement, we obtain an upper limit on the real part of the amplitude $\mathcal{A}(K_L^0 \to \mu^+\mu^-)$: $|\mathrm{Re}\mathcal{A}|^2 < 5.6 \times 10^{-10}$ at 90% C.L. This limit can be combined with phenomenological calculations of the long-distance contribution to the amplitude to obtain a constraint upon the top quark mass m_t and the Wolfenstein parameter ρ of the CKM matrix. The constraint we obtain is plotted in the ρ-m_t plane in Fig. 21 for two different values of the long-distance contribution to $\mathrm{Re}\mathcal{A}$.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the AGS staff, in particular, J. Mills, W. Leonhardt, J. W. Glenn, H. Brown, and R. Brown. We received valuable technical and engineering assistance from G. Daniel, G. Hart, V. Hart, J. Kubic, F. Mansell, M. Roehrig, A. Salminen, A. Tilghman, Q. Trang, and C. Zhang. We thank G. Bonneaud, J. Frank, J. Greenhalgh, P. Guss, M. Ierano, J. Martoff, D. Roberts, W. Wales, and D. Woycheshin for significant contributions at various stages of this work. We also thank the BNL CCD and the Cornell CNSF for assistance with data processing. This work was supported in part by the U.S. Department of Energy, the National Science Foundation, and the Robert A. Welch Foundation.

APPENDIX: LONG-DISTANCE CONTRIBUTION TO THE REAL PART OF $\mathcal{A}(K_L^0 \to \mu^+\mu^-)$

One theoretical treatment of $\mathcal{A}_{\text{long dist}}$ is that of Bergström *et al.* [11, 12], who calculate the $K\gamma^*\gamma^*$ form factor governing $K_L^0 \to \gamma^*\gamma^* \to \mu^+\mu^-$ decay by calculating the $K\gamma^*\gamma$ form factor governing $K_L^0 \to \ell^+\ell^-\gamma$ decay and extrapolating to the case where both photons are off mass shell. The $K\gamma^*\gamma$ form factor accounts for two distinct processes: a pseudoscalar-pseudoscalar $\Delta S = 1$ transition $K_L^0 \to \pi^0, \eta, \eta'$ followed by an electromagnetic transition $\pi^0, \eta, \eta' \to \gamma^*\gamma$; and a $K_L^0 \to K^*\gamma$ bremsstrahlung followed by a vector-vector $\Delta S = 1$ transition $K^* \to \rho, \omega, \phi$, where the vector meson subsequently converts to γ^*. All intermediate hadronic states are necessarily off mass shell. The form factor is written [48]

$$A_{\gamma\gamma^*}(s) = \frac{|A(K_L^0 \to \gamma\gamma)_{\text{expt}}|}{1 - s/m_\rho^2} + \alpha_K \left(\frac{\sqrt{2}\, eG_F f_{K^*K\gamma}}{1 - s/m_{K^*}^2}\right)\left(\frac{m_\rho^2}{f_{K^*} f_\rho^2}\right)\left[\frac{4}{3} - \frac{1}{1 - s/m_\rho^2} - \frac{1}{9(1 - s/m_\omega^2)} - \frac{2}{9(1 - s/m_\phi^2)}\right], \tag{A1}$$

where the first term represents the pseudoscalar-pseudoscalar transition and the second term the vector-vector transition. The variable s is the square of the $\ell^+\ell^-$ invariant mass, the form factor $f_{K^*K\gamma}$ is taken from the measured rate of $K^* \to K\gamma$, and all other masses and decay constants are well measured. The parameter α_K characterizes the strength of the vector-vector transition relative to the pseudoscalar-pseudoscalar transition. Inserting all known values and extrapolating to the case of $K\gamma^*\gamma^*$, Bergström quotes two results for $\mathcal{A}_{\text{long dist}}$ depending on whether in extrapolating from $K\gamma^*\gamma$ to $K\gamma^*\gamma^*$ one photon or both are "saturated" with vector mesons [11]:

$$\mathcal{A}_{\text{long dist}} = \left(\frac{\sqrt{2}}{\pi}\right)\left(\frac{m_\mu}{m_K}\right)\left(1 - 4m_\mu^2/m_K^2\right)^{1/4} |A(K_L^0 \to \gamma\gamma)_{\text{expt}}|\xi\,, \tag{A2}$$

where

$$\xi = (-2.3 - 5.3\alpha_K) \times 10^{-2} \tag{A3}$$

or

$$\xi = (-1.3 - 4.9\alpha_K) \times 10^{-2}. \tag{A4}$$

Taking $|A(K_L^0 \to \gamma\gamma)_{\text{expt}}| = (5.7 \times 10^{-4})^{\frac{1}{2}}$ [3] gives

$$\mathcal{A}_{\text{long dist}} = (-5.0 - 11.5\alpha_K) \times 10^{-5} \tag{A5}$$

or

$$\mathcal{A}_{\text{long dist}} = (-2.8 - 10.6\alpha_K) \times 10^{-5}. \tag{A6}$$

The true $\mathcal{A}_{\text{long dist}}$ is expected to lie between these two extremes.

This model is interesting because the phenomenological parameter α_K is fixed by the measured rate of $K_L^0 \to \mu^+\mu^-\gamma$ or alternatively by the e^+e^- invariant mass spectrum measured in $K_L^0 \to e^+e^-\gamma$ decays. Recent data gives

$$\begin{aligned}
\alpha_K &= -0.280 \pm 0.083\,^{+0.054}_{-0.034} && (K_L^0 \to e^+e^-\gamma,\ \text{BNL E845, 919 events [49]}),\\
\alpha_K &= -0.28 \pm 0.13 && (K_L^0 \to e^+e^-\gamma,\ \text{CERN NA31, 1053 events [50]}),\\
\alpha_K &= -0.15\,^{+0.14}_{-0.12} && (K_L^0 \to \mu^+\mu^-\gamma,\ \text{Fermilab E799, 199 events [51]}),
\end{aligned}$$

where the value quoted for $K_L^0 \to \mu^+\mu^-\gamma$ is based on a preliminary measurement of that branching ratio [51]. Combining the measurements by assuming the errors uncorrelated gives $\alpha_K = -0.25 \pm 0.07$. Inserting this into Eqs. (A5) and (A6) leads to

$$-2.9 \times 10^{-5} \;\le\; \mathcal{A}_{\text{long dist}} \;\le\; 0.50 \times 10^{-5}. \qquad \text{(A7)}$$

An alternative calculation of $\mathcal{A}_{\text{long dist}}$ by Ko [14] uses a hidden symmetry scheme in which vector meson form factors cut off the high momentum behavior of virtual photons [52]. Ko's result is

$$\mathcal{A}_{\text{long dist}} = (1.71 \times 10^{-11})^{\frac{1}{2}} \times (8.8 - 40.7\delta_p)\ , \qquad \text{(A8)}$$

where δ_p is a parameter whose deviation from unity measures the contribution of penguin operators to the amplitude. This parameter is constrained to be in the range $-0.06 \le \delta_p \le 0.20$ by the $\gamma\gamma$ spectrum measured in $K_L^0 \to \pi^0\gamma\gamma$ decay [52]. Thus,

$$0.27 \times 10^{-5} \;\le\; \mathcal{A}_{\text{long dist}} \;\le\; 4.7 \times 10^{-5}. \qquad \text{(A9)}$$

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

[2] M. K. Gaillard and B. W. Lee, Phys. Rev. D **10**, 897 (1974). See also M. K. Gaillard, B. W. Lee, and R. E. Schrock, *ibid.* **13**, 2674 (1976).

[3] Particle Data Group, K. Hikasa *et al.*, Phys. Rev. D **45**, S1 (1992).

[4] C. Quigg and J. D. Jackson, UCRL Report No. 18487, 1965 (unpublished); L. M. Sehgal, Phys. Rev. **183**, 1511 (1969); for a review, see H. Stern and M. Gaillard, Ann. Phys. (N.Y.) **76**, 580 (1973).

[5] B. R. Martin, E. de Rafael, and J. Smith, Phys. Rev. D **2**, 179 (1970).

[6] M. K. Gaillard, Phys. Lett. **35B**, 431 (1971); G. Farrar and S. B. Trieman, Phys. Rev. D **4**, 257 (1971); M. Pratap, J. Smith, and Z. Uy, *ibid.* **5**, 269 (1972); S. Adler, G. Farrar, and S. B. Trieman, *ibid.* **5**, 770 (1972).

[7] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981).

[8] R. E. Shrock and M. B. Voloshin, Phys. Lett. **87B**, 375 (1979).

[9] M. Kobayashi and K. Maskawa, Prog. Theor. Phys. **49**, 282 (1972).

[10] A. J. Buras, Phys. Rev. Lett. **46**, 1354 (1981); F. J. Gilman and J. S. Hagelin, Phys. Lett. **126B**, 111 (1983); N. F. Nasrallah and K. Schilcher, Z. Phys. C **36**, 467 (1987); C. Q. Geng and J. N. Ng, Phys. Rev. D **41**, 2351 (1990); C. S. Kim, J. L. Rosner, and C.-P. Yuan, *ibid.* **42**, 96 (1990).

[11] L. Bergström *et al.*, Phys. Lett. **134B**, 373 (1984).

[12] L. Bergström, E. Massó, and P. Singer, Phys. Lett. B **249**, 141 (1990).

[13] G. Bélanger and C. Q. Geng, Phys. Rev. D **43**, 140 (1991).

[14] P. Ko, Phys. Rev. D **45**, 174 (1992).

[15] A. R. Clark *et al.*, Phys. Rev. Lett. **26**, 1667 (1971).

[16] W. C. Carithers *et al.*, Phys. Rev. Lett. **30**, 1336 (1973); **31**, 1025 (1971).

[17] Y. Fukushima *et al.*, Phys. Rev. Lett. **36**, 348 (1976).

[18] M. J. Shochet *et al.*, Phys. Rev. D **19**, 1965 (1979); Phys. Rev. Lett. **39**, 59 (1977).

[19] T. Akagi *et al.*, Phys. Rev. Lett. **67**, 2618 (1991); T. Inagaki *et al.*, Phys. Rev. D **40**, 1712 (1989); T. Agaki *et al.*, KEK Report No. 94-151, 1994 (unpublished).

[20] C. Mathiazhagan *et al.*, Phys. Rev. Lett. **63**, 2185 (1989).

[21] A. P. Heinson *et al.*, Phys. Rev. D **44**, 1 (1991).

[22] A. J. Schwartz, in *The Fermilab Meeting, Proceedings of the Meeting of the Division of Particles and Fields of the APS*, Batavia, Illinois, 1992, edited by C. H. Albright, P. H. Kasper, R. Raja, and J. Yoh (World Scientific, Singapore, 1993), p. 816; A. Schwartz, in *Intersections Between Particle and Nuclear Physics,* Tucson, Arizona, 1991, edited by Willem T. H. Van Oers (American Institute of Physics, New York, 1992), p. 609; W. R. Molzon, in *Proceedings of the 25th International Conference on High Energy Physics*, edited by K. K. Phua and Y. Yamaguchi (South East Asia Theoretical Physics Association and the Physical Society of Japan, 1991), p. 1063.

[23] K. Arisaka *et al.*, Phys. Rev. Lett. **70**, 1049 (1993).

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

[25] J. Urheim, Ph.D. thesis, University of Pennsylvania, 1990.

[26] R. D. Cousins, C. Friedman, and P. L. Mélèse, IEEE Trans. Nucl. Sci. **36**, 646 (1989).

[27] P. Knibbe, Ph.D. thesis, University of Pennsylvania, 1991.

[28] A. Schwartz and D. A. Ouimette, Report No. KL-243 (unpublished).

[29] K. A. Biery, D. A. Ouimette, and J. L. Ritchie, IEEE Trans. Nucl. Sci. **36**, 650 (1989).

[30] J. Frank *et al.*, IEEE Trans. Nucl. Sci. **36**, 79 (1989); C. J. Kenney *et al.*, *ibid.* **36**, 74 (1989).

[31] P. F. Kunz *et al.*, Report No. SLAC-PUB-3332, 1984 (unpublished).

[32] R. D. Cousins *et al.*, Nucl. Instrum. Methods Phys. Res. **A277**, 517 (1989).

[33] K. Biery, Ph.D. thesis, Stanford University, 1991.

[34] J. H. Christenson *et al.*, Phys. Rev. Lett. **43**, 1212 (1979).

[35] P. Skubic *et al.*, Phys. Rev. D **18**, 3115 (1978).

[36] This algorithm was used for the analysis of 1988 data;

see Ref. [20].
[37] P. V. R. Murphy *et al.*, Nucl. Phys. B**92**, 269 (1975).
[38] J. C. Allaby *et al.*, Yad. Fiz. **12**, 538 (1970) [Sov. J. Nucl. Phys. **13**, 295 (1971)]; S. P. Denisov *et al.*, Nucl. Phys. **B61**, 62 (1973); A. S. Carroll *et al.*, Phys. Lett. **80B**, 319 (1979).
[39] T. J. Roberts *et al.*, Nucl. Phys. **B159**, 56 (1979).
[40] The energy dependence of neutron cross sections is very weak [37], while the energy dependence of pion cross sections is weak for $p_\pi \gtrsim 20$ GeV/c but rises by approximately 15% as p_π decreases from 20 GeV/c to 6.7 GeV/c. This increase is not enough for σ_π to exceed σ_n.
[41] G. Buchalla and A. Buras, Nucl. Phys. **B412**, 106 (1994).
[42] G. Buchalla and A. Buras, Nucl. Phys. **B400**, 225 (1993).
[43] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
[44] S. Narison, Phys. Lett. B **197**, 405 (1987); **216**, 191 (1989). See also: J. Gasser and H. Leutwyler, Phys. Rep. **87**, 77 (1982).
[45] See, for example, M. Neubert *et al.*, in *Heavy Flavors,* edited by A. Buras and M. Lindner (World Scientific, Singapore, 1992).
[46] R. Patterson, in *Proceedings of the XXVII International Conference on High Energy Physics,* Glasgow, Scotland, 1994 edited by I. Knowles and P. Bussey (IOP, Bristol, 1995).
[47] F. Abe *et al.*, Phys. Rev. D **50**, 2966 (1994).
[48] L. Bergström, E. Massó, and P. Singer, Phys. Lett. **131B**, 229 (1983).
[49] K. E. Ohl *et al.*, Phys. Rev. Lett. **65**, 1407 (1990).
[50] G. D. Barr *et al.*, Phys. Lett. B **240**, 283 (1990).
[51] M. B. Spencer, in Proceedings of the 8th Meeting of the American Physical Society Division of Particles and Fields, Albuquerque, New Mexico, 1994 (to be published).
[52] P. Ko, Phys. Rev. D **44**, 139 (1991).