

## Search for rare muon decays with the Crystal Box detector

R. D. Bolton, M. D. Cooper, J. S. Frank,<sup>(a)</sup> A. L. Hallin,<sup>(b)</sup> P. A. Heusi,<sup>(c)</sup> C. M. Hoffman,  
G. E. Hogan, F. G. Mariam, H. S. Matis,<sup>(d)</sup> R. E. Mischke, L. E. Piilonen,<sup>(e)</sup> V. D. Sandberg,  
G. H. Sanders, U. Sennhauser,<sup>(f)</sup> R. Werbeck, and R. A. Williams  
*Los Alamos National Laboratory, Los Alamos, New Mexico 87545*

S. L. Wilson,<sup>(g)</sup> R. Hofstadter, E. B. Hughes, and M. W. Ritter<sup>(h)</sup>  
*Hansen Laboratories and Department of Physics, Stanford University, Stanford, California 94305*

D. Grosnick<sup>(i)</sup> and S. C. Wright  
*University of Chicago, Chicago, Illinois 60637*

V. L. Highland and J. McDonough<sup>(j)</sup>  
*Temple University, Philadelphia, Pennsylvania 19122*  
(Received 16 February 1988)

A search with the Crystal Box detector shows no evidence for the lepton-family-number-nonconserving decays  $\mu \rightarrow e\gamma$ ,  $\mu \rightarrow e\gamma\gamma$ , or  $\mu \rightarrow eee$ . The search provides upper limits for the branching ratios of  $\Gamma(\mu \rightarrow e\gamma)/\Gamma(\mu \rightarrow e\nu\bar{\nu}) < 4.9 \times 10^{-11}$  (90% C.L.),  $\Gamma(\mu \rightarrow e\gamma\gamma)/\Gamma(\mu \rightarrow e\nu\bar{\nu}) < 7.2 \times 10^{-11}$  (90% C.L.), and  $\Gamma(\mu \rightarrow eee)/\Gamma(\mu \rightarrow e\nu\bar{\nu}) < 3.5 \times 10^{-11}$  (90% C.L.). In addition, a limit for the emission of a light scalar or pseudoscalar boson in radiative muon decay is given.

### I. INTRODUCTION

The standard model<sup>1</sup> of electroweak interactions correctly describes the presently available experimental data. In this model, the known quarks and leptons are divided into three generations or families. However, the multiplicity of generations of particles is not understood. This paper describes a new search for neutrinoless transitions from the muon, a second-generation lepton, to the electron of the first generation. The observation of such a lepton-family-number-nonconserving decay would indicate a connection between lepton generations and would be direct evidence for physics beyond the standard model.

The muon's role in the spectrum of elementary particles has been a mystery since its discovery<sup>2</sup> in 1936. For over a decade the muon was thought to be the quantum mediating the strong nuclear force, as predicted by Yukawa.<sup>3</sup> However, in 1947 an experiment<sup>4</sup> using muons absorbed in dense materials clearly showed that the muon does not interact via the strong force. Subsequent high-precision tests of the electromagnetic and weak couplings of muons and electrons have found no significant differences between these two particles except for their masses.

The normal decay of the muon is to an electron and two neutrinos. Neutrinoless decay modes of the muon, such as  $\mu \rightarrow e\gamma$ ,  $\mu \rightarrow eee$ , and  $\mu \rightarrow e\gamma\gamma$ , obey all conservation principles associated with space-time symmetries. However, none of these decays has been observed. To explain this, various lepton-number-conservation laws have been proposed. The first such law was introduced by Konopinski and Mahmoud<sup>5</sup> in 1953, followed by the additive lepton-number-conservation law<sup>6</sup> with separate

lepton-family numbers in 1957, and a third lepton conservation law<sup>7</sup> based on a multiplicative lepton number in 1961. These conservation laws prohibit the neutrinoless decays as well as the process  $\mu^- Z \rightarrow e^- Z$ . The additive lepton-number-conservation law also prohibits muonium-antimuonium conversion and  $\mu^+ \rightarrow e^+ \bar{\nu}_e \nu_\mu$ .

The latter two lepton-number schemes require separate types of neutrinos for the muon and the electron. Pontecorvo and Schwartz<sup>8</sup> independently proposed experiments to reveal their existence. In 1962, an experiment<sup>9</sup> confirmed that there were indeed two different types of neutrinos and supported the hypothesis of separate lepton numbers. With this discovery, interest in the search for neutrinoless decay modes of the muon waned and experimentation essentially ended for about 15 years.

A resurgence of interest in the search for rare muon decays occurred in 1977 when rumors circulated that an experiment at the Swiss Institute for Nuclear Research (SIN) had found a signal for the decay  $\mu \rightarrow e\gamma$ . This report underscored the fact that conservation of muon number is only empirical and without a fundamental basis. The rumors were later refuted, but not until after a burst of both theoretical and experimental activity. Theorists found that the suppression of the branching ratios for lepton-family-number-nonconserving decays to below the experimental upper limits ( $\sim 10^{-8}$ ) was a natural result of the new models. On the experimental side, besides the search at SIN (Ref. 10), two other experiments were quickly assembled to search with improved sensitivity for  $\mu \rightarrow e\gamma$ . One was located at the Tri-University Meson Facility (TRIUMF) (Ref. 11), and the other experiment was at the Los Alamos Meson Physics Facility (LAMPF) (Ref. 12). The results from these experiments showed no evidence for the decay  $\mu \rightarrow e\gamma$  at a

level of  $1.7 \times 10^{-10}$ .

The search for neutrinoless decay modes of the muon has spanned almost 40 years. Figure 1 shows the experimental status of the upper limits for the three neutrinoless decay modes as a function of time. The groupings of points in the plot reflect innovations in muon sources and instrumentation. The first search used muons from cosmic-ray showers and an apparatus consisting of Geiger-Müller tubes and absorbers. The next series of searches, until about 1964, used accelerators and stopped pion beams as sources of muons; the detectors were of various types, such as scintillation counters, water Cherenkov counters, spark chambers, a freon bubble chamber, and energy measurements with NaI(Tl) crystals. The recent searches, from the mid 1970s until the present, utilized high-intensity muon beams from the meson factories and improved spectrometer designs to search for these rare decays with remarkable sensitivities. The present experimental bounds on the branching ratio for the neutrinoless muon decay modes are

$$B_{\mu e\gamma} < 4.9 \times 10^{-11} \text{ (90\% C.L.) , Ref. 13 ,}$$

$$B_{\mu e\gamma\gamma} < 7.2 \times 10^{-11} \text{ (90\% C.L.) , Ref. 14 ,}$$

$$B_{\mu eee} < 2.4 \times 10^{-12} \text{ (90\% C.L.) , Ref. 15 ,}$$

where, for example,

$$B_{\mu e\gamma} \equiv \frac{\Gamma(\mu^+ \rightarrow e^+ \gamma)}{\Gamma(\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu)} . \quad (1)$$

Although the standard model is consistent with all observations, it is unlikely to be the ultimate theory. This is because, for example, it is described by more than one coupling constant, it provides no explanation for the replication of fermion families, and it is unable to predict the fermion masses and various weak-mixing angles.

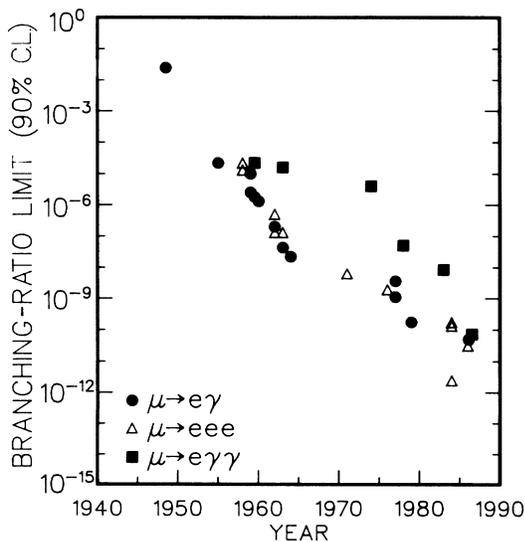


FIG. 1. The experimental upper limits for the branching ratios of the neutrinoless decay modes of the muon vs the year of the measurements.

Finding evidence for the nonconservation of family number would remove an apparently accidental conservation law that is not associated either with a space-time symmetry or with the existence of a massless gauge boson.<sup>16</sup>

Numerous extensions to the standard model have been proposed to overcome its shortcomings. These extensions contain presently undetermined parameters such as mixing angles and particle masses that are an obstacle to their predictive power. However, the relative rates for the various lepton-family-number-nonconserving processes are generally well predicted by each model. Thus, the most sensitive possible searches for a number of these processes are needed to restrict the models.

The simplest addition to the standard model is to assume that the known neutrinos have masses. The  $\tau$  neutrino has the largest experimental upper limit on its mass. Assuming that the  $\tau$  neutrino is the heaviest neutrino and that the other neutrino masses are not important, the branching ratio for  $\mu \rightarrow e\gamma$  is given by<sup>17</sup>

$$B_{\mu e\gamma} = \frac{3\alpha}{32\pi} |U_{\mu\tau}^* U_{e\tau}|^2 \frac{m_{\nu_\tau}^4}{m_W^4} , \quad (2)$$

where  $U_{i\tau}$  describes the mixing between the  $\tau$  neutrino and the other neutrinos. With the upper limit on the  $\tau$ -neutrino mass<sup>18</sup> of 70 MeV and the limit  $|U_{\mu\tau}^* U_{e\tau}|^2 < 2.5 \times 10^{-3}$ , obtained from  $\nu_\mu N$  scattering,<sup>19</sup> the branching ratio is conservatively  $< 10^{-18}$ . The branching ratio for  $\mu \rightarrow e\gamma$  from this simple model is not accessible experimentally. Thus the observation of this decay at a much higher level would be evidence for some other extension to the standard model. Some such extensions, which can result in large branching ratios for neutrinoless muon decays, are described in Sec. VIA together with the impact of this experiment on the parameters of such models. Equation (2) contains an unknown mass to the fourth power. This power dependence is characteristic of many models. Large experimental improvements are needed to significantly improve the limits on unknown masses.

The remainder of this paper discusses the essential information about the experimental search for rare muon decays. The goal of the experiment was a sensitivity in the branching ratios of  $\mu^+ \rightarrow e^+ \gamma$ ,  $\mu^+ \rightarrow e^+ \gamma\gamma$ , and  $\mu^+ \rightarrow e^+ e^- e^+$  of a few parts in  $10^{11}$ . Meeting this goal required a large quantity of muons, precise measurements of the time, energy, and direction of the decay products so that unwanted backgrounds could be suppressed, and a large-solid-angle detector to capture the final-state particles of a three-body decay. Details of the construction, calibration, and operation of the detector are given as well as the methods used to search for rare processes in the presence of backgrounds.

## II. APPARATUS AND EXPERIMENTAL TECHNIQUE

### A. General description

An experiment searching for a rare decay mode must be able to study a large number of particle decays, identi-

fy examples of the desired decay, and reject all background processes. A copious supply of  $\mu^+$ 's was produced at the A2 target at LAMPF and transported to the Crystal Box apparatus by the Stopped Muon Channel (SMC). Because the muon lifetime is long ( $2.2 \mu\text{s}$ ), the decays occurred after the muons came to rest in a thin polystyrene target.

The Crystal Box detector was designed to observe muon decays in which the final-state particles were positrons, electrons, and photons. The signature of these decays at rest is that the sum of the energies of the detected final-state particles,  $E_{\text{tot}}$ , equals the muon mass, the vector sum of the momenta of the detected final-state particles,  $|\mathbf{P}_{\text{tot}}|$ , is zero, and the particles are all produced simultaneously. Background processes can be either coincident (such as  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ ) or random (resulting from the nearly simultaneous decay of several muons each producing one or more detected particle). None of these backgrounds have the above signature. Thus, precise measurements of the energies, momenta, and times of emission of positrons, electrons, and photons enable the desired processes to be identified and unwanted events to be rejected. It should also be noted that muon decays in which all of the final-state particles are detected obey the relation  $E + P \equiv E_{\text{tot}} + |\mathbf{P}_{\text{tot}}| = M_\mu$ . Decays with undetected final-state particles obey  $E + P \leq M_\mu$ , while  $E + P$  is unconstrained for random events.

The Crystal Box detector, shown in Fig. 2, had an array of 396 NaI(Tl) crystals that surrounded the polystyrene target. Positrons, electrons, and photons initiate electromagnetic showers in the NaI(Tl), which produce an amount of scintillation light proportional to the incident energy. A measurement of the charge in the pulse from the photomultipliers (PMT's) coupled to the crystals determined the particle energy. The point of impact of the particle on the front face of the NaI(Tl) array was obtained from the distribution of energies deposited in neighboring crystals. The times of the PMT pulses determined the times of emission of the particles. Charged particles passed through a multilayer drift chamber that determined their trajectories before they reached the

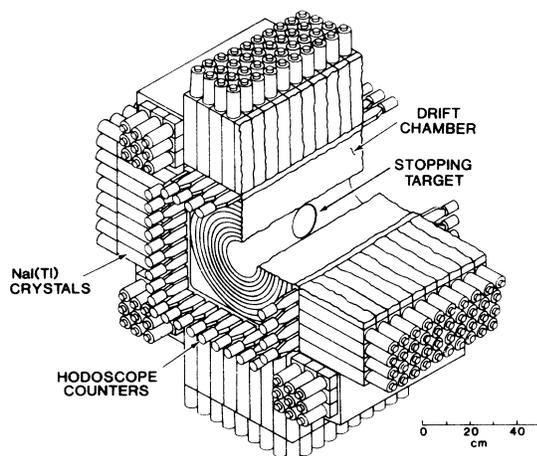


FIG. 2. A schematic cutaway diagram of the Crystal Box detector.

NaI(Tl). There was no applied magnetic field. The intersection of the reconstructed charged-particle trajectory and the target plane determined the decay point or "vertex." Photons were assumed to have started at the vertex and traveled to the point of impact on the NaI(Tl). A layer of plastic scintillation hodoscope counters located between the drift chamber and the crystals was used to differentiate between charged and neutral particles for triggering purposes and to provide precision timing information for charged particles.

## B. Muon beam

The experiment used a surface  $\mu^+$  beam<sup>20</sup> transported from the production target to the apparatus in an evacuated beam channel tuned to  $\sim 26 \text{ MeV}/c$ . The typical average beam intensity was  $4 \times 10^5 \mu^+/\text{s}$  with a duty factor of 6–9%. Because the beam line was  $\sim 30 \text{ m}$  long, the pion contamination at the experiment was extremely small. An electrostatic separator reduced the  $e^+/\mu^+$  ratio at the stopping target from about 10/1 to 1/50. Some of the remaining  $e^+$  contamination came from muon decay in flight after the separator. The ideal beam would have uniformly illuminated the stopping target with no particles stopping on the inner wall of the drift chamber upstream of the target. Because positrons from three different muons decaying at nearly the same time would not emerge from the same point on the target, a uniform beam spot would optimize background rejection in the search for  $\mu^+ \rightarrow e^+ e^+ e^-$ . Muons that stopped on the inner drift-chamber wall were not useful because the measured trajectory intersected the wall twice and it was impossible to determine which interaction was correct.

Lead collimators reduced the beam-spot size so that  $\sim 86\%$  of the beam stopped in the target. The resulting beam spot was still not uniform; the particle stopping density was higher near the center of the target. A thin aluminum sieve, located in the lead collimator, served to flatten out the stopped density. The sieve had fewer holes near the beam centerline than at larger radii so that it intercepted a larger fraction of muons near the center of the target rather than near the outside of the target.

The stopping target was a 0.35-mm-thick elliptical piece of polystyrene mounted at  $45^\circ$  with respect to the beam line so that it presented a circular profile of 6.25-cm radius to the beam. The target was thick enough to stop all muons in the beam that struck it. Polystyrene was chosen as the target material because previous measurements indicated that muons largely depolarize in polystyrene before decaying.<sup>21</sup> Measurements made with a muon-spin-relaxation ( $\mu\text{SR}$ ) apparatus found a residual muon polarization of  $(14.5 \pm 1.4)\%$ ; the measured asymmetry in the angular distribution of single positrons from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$  in the Crystal Box implies a residual muon polarization of  $(19 \pm 8)\%$ , in agreement with the above number.

Beam tuning was performed with a set of movable thin plastic scintillation counters located in the target region. Figure 3 shows a beam scan made with these counters. Final beam parameters were measured by the drift chamber with the apparatus triggered on the detection of

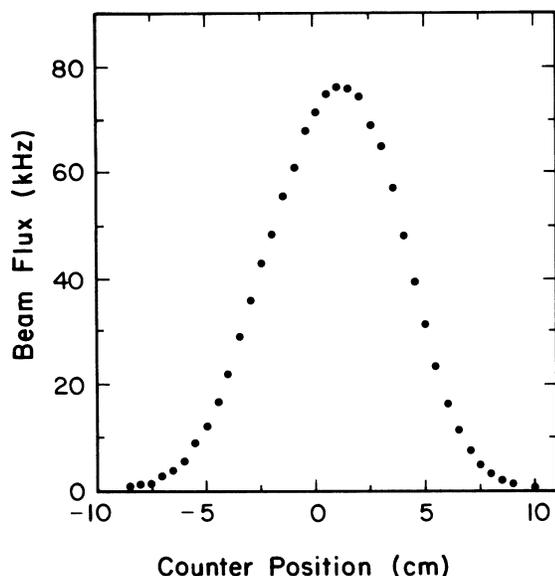


FIG. 3. A one-dimensional beam scan made with movable scintillation counters.

one charged particle ( $1e$  trigger). Figure 4 shows a typical beam spot. The “ring” that appears around the outside of the target is due to particles whose trajectories did not intersect the target; the vertex for these particles was placed at the intersection of the trajectory and the inner-drift-chamber foil closest to the measured points. The vertices for tracks that originated from the drift-chamber foil near the target were placed on the target if their trajectory intersected the target plane.

A hole with a radius of 1 cm was cut in the target, 2 cm above the center of the target. The hole is visible in Fig. 4. This hole allowed a small fraction of the incident beam to pass behind the target and stop in a special counter, the  $I$  counter shown in Fig. 5. The  $I$  counter was located so that it did not interfere with any particle trajectories that originated in the target and intersected the NaI(Tl). Muons passing through the hole in the target stopped in a 2.5-cm-diameter aluminum cup embedded in a 3.8-cm-diameter hollow cylindrical plastic scin-

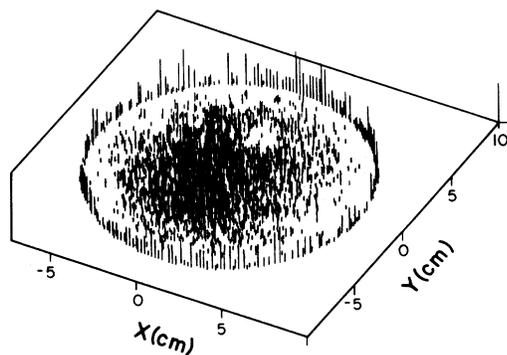


FIG. 4. The distribution of muon stopping locations as determined from the intersection of tracks in the drift chamber with the target. The height of each bar is proportional to the population in each bin.

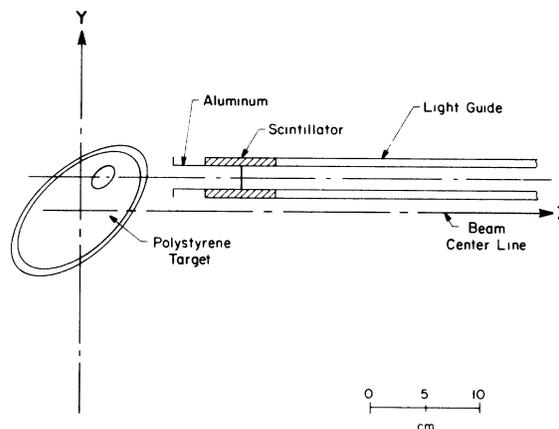


FIG. 5. A diagram showing the  $I$  counter located behind the hole in the target.

tillator located 14.7 cm downstream of the target. The scintillator was coupled to a photomultiplier tube (PMT) with a long cylindrical light pipe. Positrons from muons decaying in the cup traversed the  $I$ -counter scintillator, the drift chamber, the hodoscope counters, and the NaI(Tl). Events with a signal from the  $I$  counter thus originated from a well-defined location at a well-measured time. The  $I$  counter was used for timing calibration and stabilization.

### C. Drift chamber

A cylindrical drift chamber<sup>22</sup> was used as the central electron tracking device in the Crystal Box detector. Requirements for the drift chamber included a large solid angle to optimize the detector acceptance, small size to reduce the amount of NaI(Tl) required, low mass to reduce multiple scattering and positron annihilation in flight, a high track-finding efficiency at instantaneous muon stopping rates up to  $10^7 \text{ s}^{-1}$ , and high resolution in determining the vertex of tracks on the target.

Monte Carlo studies and prototype chamber tests led to the design of a 65-cm-long cylindrical drift chamber with 728 cells arranged in eight concentric layers with radii at the end plates between 105 and 220 mm. The cross section of a cell was  $8 \times 10$  mm at the midpoint of the chamber. The wires of each layer were at alternating stereo angles of  $10^\circ$ – $16^\circ$  from the axis of the drift chamber. Each cell consisted of a  $25\text{-}\mu\text{m}$  gold-plated tungsten sense wire surrounded by eight  $152\text{-}\mu\text{m}$  gold-plated copper-beryllium field wires. Field wire voltage ratios were chosen to keep the field lines as radial as possible, so the drift time was a function only of a track's distance from a sense wire. Materials used in construction of the chamber kept the mass between the target and the NaI(Tl) crystal array down to  $6.73 \times 10^{-3}$  radiation lengths. Uncertainties in the wire positions contributed  $90 \mu\text{m}$  (rms) to the overall position resolution. A gas mixture of 49% argon, 49% ethane, and 2% isopropyl alcohol results in a drift velocity of  $51.4 \mu\text{m/ns}$ .

At low beam rates ( $\leq 2 \times 10^6 \mu^+/\text{s}$ ) the intrinsic resolution was  $130 \mu\text{m}$  rms for each drift-chamber layer. At

high rates ( $\approx 8 \times 10^6 \mu^+/\text{s}$ ) the intrinsic resolution per plane degraded by as much as 30% for events with three electrons ( $3e$  triggers). The efficiency of event reconstruction was  $\geq 97\%$  for a  $1e$  trigger at low beam rates and  $\sim 75\%$  for a  $3e$  trigger at high rates.

#### D. Scintillators

The scintillation-counter hodoscope located between the drift chamber and the NaI(Tl) array was used to define a charged-particle trigger for electrons and a veto signal to reject charged particles in the photon quadrant trigger. The 36-counter hodoscope consisted of 32 rectangular Pilot B scintillators of dimensions 44.4 cm (length)  $\times$  5.3 cm (width)  $\times$  1.3 cm (thickness) and four trapezoidal scintillators with the same length and thickness but a slightly smaller width that were used in the corners. The scintillators were individually wrapped with a single layer of aluminum foil to minimize inert material: each quadrant of the array was then covered with a layer of black paper. The counters were coupled at each end to Amperex 2232B PMT's with adiabatic light pipes approximately 0.5 m long. The length of counters was chosen so they intercepted all charged-particle trajectories that originated at the target and struck the fiducial volume of the NaI(Tl) (the fiducial volume excluded the farthest upstream and farthest downstream columns of crystals).

To tag charged particles that entered the NaI(Tl) array without passing through the plastic scintillators (thus simulating photons), the regions upstream and downstream of the scintillators were covered by an array of 16 guard scintillation counters each measuring 13.3 cm  $\times$  23.8 cm  $\times$  0.3 cm. They overlapped the hodoscope scintillators by about 1 cm.

The average measured time resolution of the individual hodoscope counters was 290 ps [full width at half maximum (FWHM).] The resolution was determined by measuring the timing of positrons relative to the  $I$  counter for normal muon decays from the target.

#### E. NaI(Tl)

The construction, operation, and performance of the array of 396 NaI(Tl) crystal modules have been described in detail elsewhere.<sup>23</sup> There were 360 crystals with dimensions 6.35 cm  $\times$  6.35 cm  $\times$  30.5 cm that were arranged in four quadrants of 10-column  $\times$  9-row arrays of 90 crystals each; columns ran perpendicular to the beam line. In addition, there were 36 crystals, each measuring 6.35 cm  $\times$  6.35 cm  $\times$  76.2 cm, that were arranged in four 3  $\times$  3 crystal arrays, located in the corner spaces between adjacent quadrants.

#### F. Pion beam and liquid-hydrogen target

The primary calibration of the NaI(Tl) utilized photons from the decay of  $\pi^0$ 's produced from  $\pi^-p \rightarrow \pi^0n$  at rest. For these data, the drift chamber was removed from the center of the detector and replaced by a liquid-hydrogen target. The target cup was a 5-cm-long, 3.8-cm-diameter cylinder with hemispherical ends. The beam channel was

tuned to transport 140 MeV/c  $\pi^-$ . The energy of the pions was degraded with graphite to maximize the pion stopping rate. Two thin plastic scintillation counters were located just upstream of the target to define the incident beam. The incident pion intensity was  $10^4 \pi^-/\text{s}$ . It took roughly eight hours to switch from muon to pion data taking or back.

#### G. Trigger electronics

The apparatus was triggered on three processes ( $e^+\gamma$ ,  $e^+\gamma\gamma$ , and  $e^+e^+e^-$ ) independently. To minimize dead time, the trigger electronics had to reduce the rate of candidate decays to a rate of less than 160 Hz instantaneous (10 Hz average).

The first-level  $eee$  trigger required three or more hodoscope counters in time coincidence. Constant-fraction discriminators<sup>24</sup> (CFD's) on the signals from each end of the hodoscope counters removed time slewing. A high-speed meantimer<sup>25</sup> required a coincidence between signals from the two ends of a counter and reduced the position dependence of the signal timing of a hodoscope counter from 3 ns over the length of the scintillator to less than 100 ps. This allowed the use of very tight ( $\sim 5$  ns) on-line timing. A nonadjacency circuit similar to that of Flauger<sup>26</sup> determined the number of separated clusters of scintillators. Demanding three or more clusters imposed a minimum opening angle requirement between electrons. A successful first-level trigger started the data-acquisition system.

The second-level trigger performed energy and geometry cuts based on information latched by the first level. Electrons and positrons were required to have a minimum energy of  $\sim 11$  MeV by demanding that there was a CFD signal from at least one NaI(Tl) crystal in the row behind the struck scintillator or in an adjacent row. The signal threshold on the CFD for each crystal was 6 MeV. There was an additional  $\sim 5$  MeV energy loss in the plastic scintillators and the Al wall of the NaI(Tl) container. A pattern recognition unit then examined the pattern of tagged hodoscope counters for any three scintillators that could have been from a  $\mu^+ \rightarrow e^+e^+e^-$  decay at rest. The patterns used by the recognition logic were derived from the Monte Carlo program (see Sec. III) using a uniform phase-space distribution. If the second-level conditions were not satisfied, the data acquisition was aborted and the data-acquisition modules were reset.

The  $e\gamma$  and  $e\gamma\gamma$  triggers were based on the response of an entire quadrant rather than that of a single scintillator or row of NaI(Tl). Particle differentiation was based on the response of the plastic-scintillator counters. An electron quadrant was defined as a coincidence between one or more NaI(Tl) CFD signals (excluding the farthest upstream and farthest downstream columns of crystals but including the corner crystals nearest the given quadrant) and one or more signals from the meantimers connected to the hodoscope counters in the same quadrant. A photon quadrant was defined as one or more such NaI(Tl) CFD signals in the absence of any signal from the plastic scintillators in front of the NaI(Tl). This plastic-scintillator signal used dead-time-less leading-edge discriminators. A signal from either end of a hodoscope

counter, from a guard counter in that quadrant, or from the upstream PMT of either of the nearest hodoscope counters in the two neighboring quadrants constituted this scintillator signal. These last two counters were included to protect against false photon signals from a shower spreading from one quadrant to the next. A NaI(Tl) quadrant energy signal was generated by linearly adding the analog signals from all of quadrant face crystals plus the signals from the nearest corner crystals and one half of the signals from the adjacent diagonal corner crystals.

The  $e\gamma$  trigger also had two levels. The first level required a coincidence of a photon quadrant signal, the opposite-facing electron quadrant signal, and NaI(Tl) quadrant energy signals (with a threshold of  $\sim 35$  MeV) from each quadrant. The second level had pattern recognition logic designed to assure that the geometry of the event was consistent with the positron and the photon having equal and opposite momentum. The logic used information from the NaI(Tl) CFD's grouped by columns. The logic required at least one pair of columns such that if there was a signal from column  $i$  in the photon quadrant, there was also a signal from a column within the range  $(9-i)$  to  $(13-i)$  in the electron quadrant. Only columns 2 through 9 were used in this logic. A hit in a corner crystal automatically satisfied the column logic requirement.

The  $e\gamma\gamma$  trigger was defined as one and only one electron quadrant, at least two photon quadrants other than the electron quadrant, and  $\geq 70$  MeV deposited in the entire NaI(Tl) array. Events with the second photon in either the electron quadrant or the first photon quadrant were recorded only if they satisfied the  $e\gamma$  logic.

In addition to the data triggers, there were a number of calibration and stabilization triggers. The most important were the single-electron triggers and the pion-calibration triggers. The single-electron triggers were used for timing calibration and energy stabilization. The  $(1e) \cdot I$  trigger was used for timing calibration; it required a signal from one and only one hodoscope counter, at least 10 MeV deposited in the NaI(Tl) quadrant behind that hodoscope, and a signal from the  $I$  counter. The timing of this trigger was determined by the  $I$  counter permitting the measurement of the timing of all the various detector elements relative to a common source. For the other single-electron trigger,  $(1e)$ , a sample of positrons from ordinary muon decay was collected so that drifts in the position of the high-energy edge of the energy spectrum could be measured. This trigger did not include the  $I$  counter because events from the  $I$  counter would have had the spectrum smeared due to the energy loss in the counter. Instead, the trigger simply required a signal from one and only one hodoscope counter.

The pion triggers were used to get the absolute energy calibrations of the NaI(Tl). Two reactions were detected. The trigger for  $\pi^- p \rightarrow \gamma n$  was one and only one photon quadrant with  $\geq 100$  MeV in the NaI(Tl). The trigger for the chain  $\pi^- p \rightarrow \pi^0 n$ ,  $\pi^0 \rightarrow \gamma\gamma$  was two opposite photon quadrants with  $\geq 30$  MeV in each quadrant. The photon energies from  $\pi^0$  decay range from 55.1 to 82.7 MeV. The minimum opening angle between the photons is

156.5°. For both triggers, the timing of the trigger was determined by pion beam counters.

#### H. Data-acquisition electronics

The plastic counter PMT signals were passively split with 20% of the pulse height going to a CAMAC analog-to-digital converter (ADC) for pulse-height analysis, 40% of the pulse height going to the CFD's, and the remaining 40% going to the leading-edge discriminators. Signals from the CFD's for both ends of a hodoscope counter fed the meantimer circuit. The arrival times for particles in the guard counters and the mean-times for hodoscope counters were digitized by TDC's that had a sensitivity of 100 ps/channel.

The processing of the NaI(Tl) signals is described in detail in Ref. 23. Timing and pulse-height information from each NaI(Tl) crystal were digitized for each event. The timing signals were discriminated with CFD's similar to those for the signals from the plastic scintillators but modified to accommodate the slower rise-time characteristics of the NaI(Tl) signals.

The amount of pileup energy in each NaI(Tl) crystal was measured with an auxiliary ADC system. Pileup refers to energy deposited in the NaI(Tl) by an additional undetected and uncorrelated particle. The integration time for the pileup ADC system was 60 ns, compared to the 200-ns integration time used with the primary ADC system. Figure 6 shows the integration periods for both ADC systems and a typical NaI(Tl) pulse. This arrangement of the integration times provided a sensitive detection of pileup pulses. Details of the algorithm that was used to reject pileup pulses are presented in Sec. V B.

#### I. Run monitor system

An independent monitoring system running on a separate computer was installed to ensure that the Crystal Box apparatus was functioning properly. This run monitor system (rms) continually scanned and recorded the readings from a set of sensors located throughout the

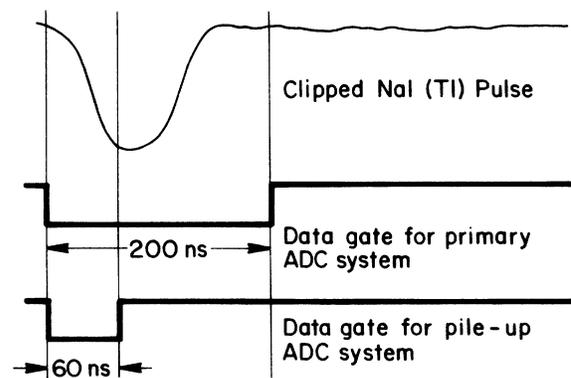


FIG. 6. Diagram of a typical NaI(Tl) signal pulse and the gating periods for the pulse-height and pileup measuring systems.

apparatus and performed an appropriate action if any sensor reading fell outside its predefined normal operating range. The rms also monitored and controlled the high voltages on the drift chamber. An autodialing modem was connected to the rms providing phone alarms.

### J. Data collection

There were three major data-taking periods, called I, II, and III. During the first, the instantaneous beam intensity was  $\sim 4.9 \times 10^6 \mu^+ \text{ s}^{-1}$  with an average duty factor of 6%. The pileup-detection system was not installed for this period. During period II, the instantaneous beam intensity was  $\sim 4.7 \times 10^6 \mu^+ \text{ s}^{-1}$  with an average duty factor of 9%. The pileup-detection system was partially installed during this period but was not used in the analysis. The pileup-detection system was fully installed and utilized during the final data-taking period, which had an average duty factor of 6.6%. Runs with different instantaneous beam intensities, varying from  $4 \times 10^6 \mu^+ \text{ s}^{-1}$  to  $9 \times 10^6 \mu^+ \text{ s}^{-1}$ , were interspersed so rate-dependent effects could be studied.

During data taking there were runs to calibrate the plastic-scintillator timing every day, and runs to calibrate the NaI(Tl) timing twice a week (see Sec. IV B 2). There were three energy-calibration periods with the drift chamber removed, the liquid-hydrogen target installed, and a  $\pi^-$  beam incident. One such period occurred during the period I, one at the end of period II, and one midway through period III. There were also occasional calibration runs with a Pu- $\alpha$ -Be source (whenever the beam went off for a long time) and special runs (such as runs triggered on a pulser and runs triggered on a single photon) at various times.

A data run took typically two hours to fill a 2400-ft, 1600-bpi (bits-per-inch) tape with  $\sim 35\,000$  events. Each data run was preceded by a short run to establish the ADC pedestals and a flasher run (see Sec. IV A 4) for NaI(Tl) channel gain corrections. A total of 1500 magnetic tapes were used in this experiment; approximately half of them contained production muon data.

## III. MONTE CARLO SIMULATION

### A. Introduction

A Monte Carlo program was written to simulate the response of the detector to many different types of events (see Table I). The Monte Carlo program followed only electrons, positrons, and photons, as these were the only particles that were detected. As described below, care was taken to ensure that this program faithfully reproduced the behavior of the detector. The program produced an output file that contained pulse heights and times for each individual detector element in the same format as for actual data; both data and Monte Carlo output files were then processed by the same programs.

### B. Event generation

#### 1. Muon decay

It was assumed that all processes initiated by muon decays originated with an unpolarized muon decaying in

TABLE I. Types of events generated with the Monte Carlo program.

Process	Trigger
$\mu^+ \rightarrow e^+ \gamma$	$e-\gamma$
$\mu^+ \rightarrow e^+ \gamma \nu \bar{\nu}$	$e-\gamma, 1-\gamma$
$\mu^+ \rightarrow e^+ \gamma \gamma$	$e-\gamma-\gamma, e-\gamma$
$\mu^+ \rightarrow e^+ e^+ e^-$	$e-e-e$
$\mu^+ \rightarrow e^+ e^+ e^- \nu \bar{\nu}$	$e-e-e$
$\mu^+ \rightarrow e^+ \nu \bar{\nu}$	$1-e$
$\mu^+ \rightarrow e^+ \gamma f$ ( $f = \text{famon}$ )	$e-\gamma$
$\pi^0 \rightarrow \gamma \gamma$	$\gamma-\gamma, 1-\gamma$
$\pi^- p \rightarrow n \gamma$	$1-\gamma$

the polystyrene target. A small residual polarization would have had a very small effect because the apparatus was symmetric about  $90^\circ$ . The muon position distribution on the target was derived from the intersection of the drift-chamber track with the target from data runs triggered on single positrons from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$ .

Single positrons from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$  were generated according to the Michel distribution,<sup>27</sup>  $F(x) = (2x - x^2)x^2$ , where  $x = 2P_e/m_\mu$ . The parameter  $x$  varies between 0 and 1. No radiative corrections were used in the generation of these positrons because the resolution of the NaI(Tl) was large enough to mask this effect and because radiated photons tended to enter the same NaI(Tl) crystals as the positron.

For  $\mu^+ \rightarrow e^+ \gamma$  the positron and photon each have energies approximately equal to one-half of the muon mass and are emitted at  $180^\circ$  with respect to each other. The energy of each particle in  $\mu^+ \rightarrow e^+ e^+ e^-$  was generated uniformly over the allowed region of phase space, consistent with conservation of energy and momentum. The final-state particle momentum distributions for  $\mu^+ \rightarrow e^+ \gamma \gamma$  were calculated from a general local interaction.<sup>28</sup> The resulting form of the distribution is

$$\frac{d^2\Gamma}{dE_1 dE_2} = K E_e E_1^2 E_2^2 (1 - \cos\theta)^2, \quad (3)$$

where  $E_e$ ,  $E_1$ , and  $E_2$  are the energies of the positron, the higher-energy photon, and the lower-energy photon, respectively,  $\theta$  is the opening angle between the photons, and  $K$  is a constant. This equation can be rewritten as

$$\frac{d^2\Gamma}{dx dy} \propto x(x-1)^2, \quad (4)$$

where  $y = (E_1 - E_2)/m_\mu$  and  $-x \leq 2y \leq x$ . A nonlocal interaction, such as a soft photon being emitted from one of the external lines from  $\mu^+ \rightarrow e^+ \gamma$ , would have a different distribution; however, the expected branching ratio would then certainly be lower than that for  $\mu^+ \rightarrow e^+ \gamma$ .

For muon inner bremsstrahlung,  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ , a coincident photon and positron were generated according to the distribution given by Fronsdaal and Uberall.<sup>29</sup> To reduce the amount of computer time needed for event generation, only positrons and photons with energies above some threshold were generated; this threshold was

typically several MeV below the hardware threshold to allow for piled-up energies and energy-resolution effects. The Monte Carlo simulation used an event generator by Sapirstein<sup>30</sup> to calculate the radiative three-electron decay  $\mu^+ \rightarrow e^+ e^+ e^- \bar{\nu} \nu$ . The distributions used to generate  $\mu^+ \rightarrow e^+ \gamma f$ , where  $f$  is a light scalar or pseudoscalar boson such as a familon, are given in Ref. 31.

## 2. Pion events

The Monte Carlo program also simulated the photons produced by stopping  $\pi^-$ 's in liquid hydrogen. The program first generated a random location in the liquid-hydrogen target for the interaction. The two processes  $\pi^- p \rightarrow \pi^0 n$  (charge exchange) and  $\pi^- p \rightarrow \gamma n$  (radiative pion capture) were generated separately. The neutral pion in the charge-exchange reaction is emitted isotropically with a unique energy (4 MeV): the photons from the  $\pi^0$  decay are then emitted back to back, isotropically in the  $\pi^0$  center of mass. The 129.4-MeV photon from radiative pion capture is emitted isotropically in the laboratory.

## C. Simulating the detector response

The heart of the Monte Carlo program is the EGS3 shower code.<sup>32</sup> Using this code, the Monte Carlo program was able to follow each particle through the detailed geometry of the detector while keeping track of additional particles produced by interactions in the detector. The program defined various regions such as the stopping target, the space between the target and the drift chamber, the drift-chamber foils and wires, the scintillator hodoscope, the NaI(Tl) crystals, and the structural parts of the detector.

It was necessary to make one improvement to the EGS3 code. When a positron or electron traversed some material in EGS3, its energy was reduced by the mean ionization energy loss. This energy loss is not correct and results in distortions when the particles traversed a thick slab of material such as the scintillator hodoscope. EGS3 was modified to calculate the charged-particle energy loss<sup>33</sup> according to a Vavilov energy-loss distribution.

The output of the EGS3 section of the Monte Carlo code was used to generate a pulse height and a time for each of the detector elements. This information was then used to determine which discriminators would have fired if the event had occurred in the real detector, and whether the event would have passed the trigger logic. The data-analysis programs were used to impose various cuts and to perform higher-level analysis on the Monte Carlo events. This step was essential to ensure that measured data and Monte Carlo-generated events were treated identically.

Several effects could not be incorporated in a general way into EGS3. Although EGS3 deals with the broadening of the energy response due to the leakage of photons out of the NaI(Tl), it cannot include the statistics of light generation and collection in the NaI(Tl) and plastic scintillator elements. Consequently, a timing resolution and an energy-dependent energy resolution were added to the signal from each element during replay. A second prob-

lem arises because there are often very-low-energy photons in an electromagnetic shower. EGS3 does not follow an individual photon with an energy below 100 keV; it deposits all of this energy in the region in which it is created. This implies that the energy lost in the inert regions of the detector was slightly overestimated; a small correction was applied during replay to remedy this effect.

The third known difference between the Monte Carlo simulation and data has to do with pileup. The amount of pileup was measured by taking special runs with a "random" trigger. The trigger was generated by a pulser, gated by the presence of the beam, and was independent of any signals in the detector. For these runs, the distribution of signals was identical to that for pileup, (i.e., due to random particles). These pileup data were used to generate the probability distribution of pileup energy for photons and for positrons as a function of location in the detector. The pileup data were then added to the Monte Carlo data. To simulate data after pileup rejection, distributions generated from the random trigger data tapes with pileup rejection were used. This latter procedure assumed that the effect of the pileup rejection scheme was independent of the presence of any "true" NaI(Tl) signal. This effect is discussed later in Sec. V B.

The energy and timing resolutions of the NaI(Tl) crystals and of the plastic scintillators had to be characterized in the Monte Carlo program. These parameters for the NaI(Tl) were determined from the data for  $\pi^0 \rightarrow \gamma \gamma$ . The parameters for the plastic scintillators were determined from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$  data. After these parameters were determined, a variety of tests were imposed to ensure that the Monte Carlo program faithfully simulated the detector. The agreement of the  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  and  $\mu^+ \rightarrow e^+ e^+ e^- \nu \bar{\nu}$  data with the Monte Carlo expectations constitute stringent tests (see Secs. V E and V G). These processes are very sharply energy dependent; any disagreement in geometry or response between the Monte Carlo simulation and data would have resulted in very different distributions. The shape of the measured positron energy spectrum from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$  agrees very well with the Monte Carlo simulation (see Sec. IV D). The Monte Carlo program accurately modeled the energy leakage as a function of positron for different crystals throughout the detector. The muon position of a particle in the NaI(Tl) was determined by an energy-weighted average of the energies deposited in the crystals (see Sec. IV C). The exponent of the energy-weighting factor was determined from the Monte Carlo program, and then compared with that found for positron data whose position was given by the drift-chamber track. The agreement of this comparison again verifies the accuracy of the Monte Carlo simulation. Finally, the Monte Carlo prediction for the energy spread into surrounding crystals for a shower agrees with observations.

## IV. DETECTOR CALIBRATION AND STABILIZATION

### A. Energy

#### 1. Hardware gains

The energy gains of all NaI(Tl) channels were equalized to within 10% by adjusting the variable-gain

amplifiers on each channel.<sup>23</sup> The 4.4-MeV  $\gamma$ -ray line from a Pu- $\alpha$ -Be source was used. The resulting gain was  $\approx 20$  mV/MeV for the CFD input. Thus the CFD threshold of 120 mV corresponded to the desired energy threshold of  $\sim 6$  MeV in the NaI(Tl) crystals. The corner crystals had two phototubes contributing to the hardware energy sums, so the individual gains of the corner crystal amplifiers were  $\approx 10$  mV/MeV. This means that each corner crystal CFD had an effective 12-MeV energy threshold.

After the hardware gains for all channels were set, one tape per quadrant of data triggered on the 4.4-MeV source photons was taken. The trigger required one and only one NaI(Tl) row with at least 2 MeV deposited. The data on each tape were preceded by a set of flasher spectra (see Sec. IV A 4). This information permitted a check of the software calibrations of the NaI(Tl) readout and established reference values for the flasher gain stabilization system.

The absolute calibration of each channel was adjusted off line in several steps. First, the shape of the nonlinear calibration function<sup>34,23</sup> was measured for each readout channel. The absolute scale of the calibration function was then set by the analysis of  $\gamma$  spectra (from  $\pi^0$  decays) collected during the runs with the hydrogen target. Finally, the energy scale was checked by the 4.4-MeV source line and the monoenergetic  $\gamma$  from  $\pi^-p \rightarrow n\gamma$ . A check of the charged-particle response was provided by the observed Michel spectrum.

## 2. Energy calibration

Once the shape and approximate scale of the calibration curves were determined, the final scale was set by referring one point on the curve to an absolute standard. The central energy of the  $\pi^0$ -decay  $\gamma$  spectrum (68.9 MeV) from  $\pi^-p \rightarrow \pi^0n$ , was chosen as the reference point because it was the calibration energy closest to the 52.8-MeV region of interest for  $\mu \rightarrow e\gamma$ .

The relevant measured quantity is the ‘‘clump’’ energy sum, and not the energy of the individual crystals.<sup>23,24</sup> A clump consists of a crystal with the largest pulse height  $> 5$  MeV and a set of neighboring crystals; the neighboring set generally consists of 24 surrounding crystals except when the high-pulse-height crystal (HPHC) is near an edge or a corner of the detector. As many as nine clumps may be defined in a given event.

The final calibration spectra were collected in a set of histograms with one histogram associated with each crystal. A crystal’s histogram was incremented when it was the HPHC of a clump. The fraction of the clump energy contained in the HPHC was required to be  $> 70\%$ , so the energy of a calibration spectrum was primarily determined by the gain of the HPHC. The appropriate adjustment of gain for each crystal was derived from the histogram associated with it.

There were 31 different symmetry regions in the detector, 25 for the face crystals and six for the corner crystals, for particles coming from the target. The Monte Carlo program was used to generate a set of reference spectra for each of these regions. For each crystal, the

measured spectrum was fit to the Monte Carlo spectrum for the appropriate symmetry group with the gain and the energy resolution as free parameters. The assumed gains for each crystal were corrected and the procedure iterated. The procedure converged within four to five iterations. Several observed and predicted calibration spectra are shown in Refs. 23 and 34.

## 3. Flasher system

The flasher system<sup>35</sup> consisted of a xenon flashtube light source and a network of plastic fiber-optic cables to transport the light to the light guide of each NaI(Tl) crystal. The electronic pulse applied to the flashtube was shaped so the light output imitated a NaI(Tl) pulse. A vacuum photodiode and a silicon photodiode monitored the flashtube light output. The vacuum photodiode also provided the start signal for the timing measurements made with the flasher system. One end of the optical fiber was located on a spherical surface such that each cable directly viewed the flashtube. The other connection was glued into a slot at the light guide of each PMT.

A direct comparison of the average energy gain of all the crystals made with the flasher system versus the average energy gain measured by positron energy spectra shows that the gains measured by the flasher changed by as much as 8% over time. A quadrant average of the energy gains determined from the positron energy spectra indicated that the flasher quadrant average gains changed by several percent with respect to each other. The individual crystal energy gains appeared to be stable within 1.5% per run.

The time offsets, as determined by the timing-peak positions, were corrected for each run using the flasher system. The quadrant average of the time peak positions without flasher corrections varied by about 500 ps over the course of a month. These time differences diminished to  $< 250$  ps using the flasher corrections.

## 4. Energy stabilization

Maintaining calibrations over several months of running for 500 channels of analog devices is a formidable task. The flasher system was to be used for the primary stabilization. Although it was able to follow relative gain shifts between crystals, an analysis of the flasher system performance showed that the monitors introduced noise when used to stabilize the overall quadrant energy gains. Thus, crystal-to-crystal gain shifts were followed with the flasher system, while gain stabilization used an analysis of the positron spectrum from the data. The stabilization was done on a quadrant-by-quadrant basis for each data run. A two-dimensional histogram of positron clump energy versus HPHC number was accumulated during the first data-analysis pass for each run. The central 25 crystals in each quadrant were summed. A fifth-order Chebyshev polynomial was fit to a ‘‘standard’’ run; subsequent runs were fit to the same polynomial with the gain shift and normalization as free parameters. This allowed a run-by-run normalization of the quadrant gains. A comparison of the stabilized gains and the gains mea-

sured in the  $1e$  data runs taken every day shows that the gains were stabilized to  $<0.25\%$ .

### B. Timing

All of the timing was done with respect to the  $I$  counter. The measured time resolution of the  $I$  counter was 350 ps (FWHM). Timing spectra were accumulated for each counter being calibrated. The spectra were then fit by the data-acquisition computer. For the on-line calibration, the fit results were analyzed to find what cable lengths were needed to bring each counter in time.

#### 1. Hardware timing

For the hardware timing, a single time-to-digital converter (TDC) was used to monitor the time of arrival at an appropriate point in the logic of signals from each member of a group of counters. The face NaI(Tl) crystals were divided into timing groups of 36 crystals each. Each crystal in a group was far enough away from other members of that group that a single shower would appear in only one crystal in the group. Timing data were taken with only one group connected to the logic at a time. In this way, the time of the signal was unambiguously assigned to one crystal for each event. The corner crystals were divided into similar groups. Each NaI(Tl) crystal was timed into the logic within  $\pm 1$  ns. The timing of the NaI(Tl) analog signals into the ADC system was checked manually to within  $\pm 5$  ns.

#### 2. Off-line timing

All PMT's were timed with respect to the  $I$  counter, with each PMT having its own TDC. For each channel, the  $t=0$  offset and gain (ns per channel) had to be found. The offset was found from each TDC spectrum in the  $(1e)\cdot I$  data coincidences between the  $I$  counter (the start signal) and the counter being calibrated (the stop signal). The gain was found by delaying the start signal by 10 ns for every other trigger, which produced a double-peaked spectrum as shown in Fig. 7. The 10-ns interval was defined by a cable. This method was used to calibrate the NaI(Tl) timing every two to three days. Between calibration runs, the flasher system was used to track drifts in the offsets. Finally, the readout of the NaI(Tl) timing system had a small crosstalk problem that was measured and removed.

The plastic-scintillator timing was calibrated daily. The guard counters were done with the same method used for the NaI(Tl). The hodoscope calibration was more complicated because there were two PMT's on each counter whose calibrations had to be coordinated. The gain of the TDC on each PMT was found as described above. The offsets were not computed from the raw data, but rather from the average (meantime) and difference of the times from the two PMT's. Only events in the downstream half of the detector were used in these spectra to minimize time-of-flight effects. The time difference data were from  $1e$  data. The time difference is a measure of the position of the hit in the scintillator. Without the  $I$  counter in the trigger, a fairly flat distribution in time

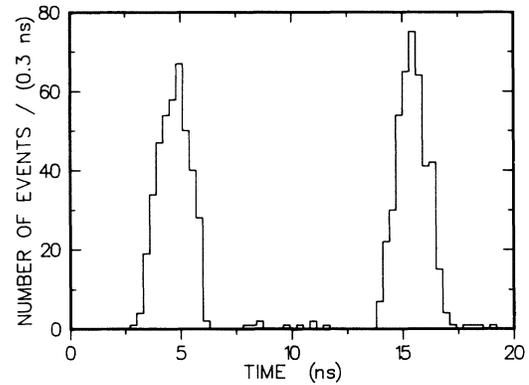


FIG. 7. A double-peaked spectrum for the relative  $I$ -counter-hodoscope-counter single phototube time from  $(1e)\cdot I$  data.

difference was obtained. Offsets were chosen that centered this distribution and the meantime peak at predetermined values.

#### 3. Time resolution: plastic scintillators

The time resolution of the hodoscope counters was derived from data with a  $(1e)\cdot I$  trigger as shown in Fig. 8. The width of the peak is the convolution of the resolutions of a hodoscope and the  $I$  counter. To separate these effects, three measurements were made. The width of the coincidence peak between the  $I$  counter and two other counters and between these two counters were measured. This last was done using cosmic rays with counters on opposite sides of the detector. Cuts and corrections were applied to minimize time-of-flight effects. Consistent results were obtained using different pairs of hodoscope counters. The results were that  $\sigma_{I\text{ counter}} = 145 \pm 26$  ps and  $\langle \sigma_{\text{hodoscope}} \rangle = 123 \pm 22$  ps at an average decay rate of  $10^5 \mu^+ \text{ s}^{-1}$ . At  $5 \times 10^5 \mu^+ \text{ s}^{-1}$ , the average resolution increased to  $\langle \sigma_{\text{hodoscope}} \rangle = 142 \pm 25$  ps.

#### 4. Final timing

The final calibration of the relative NaI(Tl)-scintillator timing was performed using coincident photons and positrons in opposite quadrants from  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ . Figure 9 shows the relative positron-photon timing for a represen-

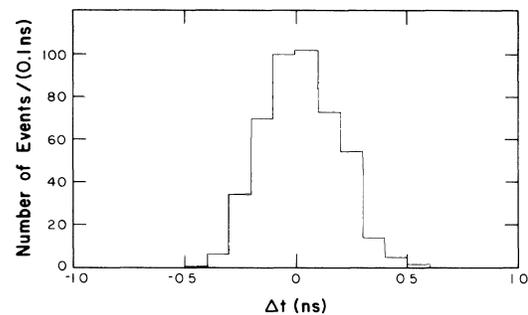


FIG. 8. A spectrum for the relative timing between the  $I$  counter and a hodoscope meantime.

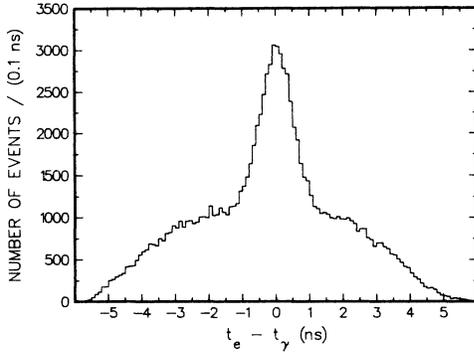


FIG. 9. The relative positron-photon timing for  $e\gamma$  triggers.

tative set of  $e\gamma$  triggers with no cuts applied. The prominent coincidence peak is due to muon inner bremsstrahlung: the broad background is due to random positron-photon triggers. The width of the peak, 1.15 ns (FWHM), is a measure of the timing resolution of the NaI(Tl) for particles with more than  $\sim 38$  MeV including all effects of time-of-flight corrections and calibration drifts.

A final correction to the positron-photon timing was applied on a quadrant-by-quadrant basis (for example, photon in the left quadrant, positron in the right quadrant) for data covering each set of timing constants. This correction was generally no larger than 200 ps.

### C. Photon conversion point

Although the trigger treated an entire NaI(Tl) quadrant as a single entity, the functional unit corresponding to a particle in the off-line analysis was a clump. It was possible for several clumps to exist within one quadrant. An estimate of the particle shower position in the NaI(Tl) calorimeter was calculated as the weighted centroid of the deposited energy distribution in a clump as described in Ref. 23.

The position resolution of the NaI(Tl) calorimeter was determined using positrons from muon decays in the stopping target. In agreement with the expectations from the Monte Carlo studies, the resolutions were  $\sigma = 1.70$  cm transverse to the beam direction, and  $\sigma = 1.85$  cm along the beam direction. The latter number is larger due to the poorer track resolution from the drift chamber in this direction.

### D. Determination of resolution functions

#### 1. NaI(Tl): energy

There were several checks used to assure that the final energy calibration was consistent over the dynamic range of interest. In all of the comparisons below, events were selected for which the HPHC of the clump was within the central  $5 \times 6$  array of one of the quadrants.

The measured detector response to the 129.4-MeV photon from  $\pi^- p \rightarrow n \gamma$  has a resolution of 7% (FWHM); the absolute peak position is correct to better than 1%. The observed spectrum for  $\pi^0$  decay events with

$\theta_{\gamma\gamma} > 175^\circ$  has a peak near 55 MeV that is quite well simulated in position and shape.<sup>23</sup> The upper edge of this peak is broadened due to the finite angular acceptance of the sample. Therefore, the peak width of 8% (FWHM) may be taken as an upper limit to the detector resolution at 55 MeV. The low-energy photon response was checked by reference to the 4.4-MeV line from the Pu- $\alpha$ -Be source. This measured peak position is correct to within 0.2 MeV.

The only well-defined spectrum available in the data to check the positron energy measurement is that from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$ . Events originating in the stopping target were selected because those muons were largely depolarized in the target material, thus minimizing any asymmetry in the angular distribution of the positron energy. The ionization energy lost in the material traversed before reaching the NaI(Tl) was added to the energy measured in the NaI(Tl) to obtain the charged-particle energy. The simulation and the data were in agreement to 1% for the positron energy calibration.

#### 2. NaI(Tl): timing

The (electron) timing resolution of individual face crystals versus the  $I$  counter was measured to be 1.10 ns (FWHM). However, the superior timing resolution of the trigger counters was always used to define the time of an electron clump. The NaI(Tl) timing information was used to define the photon clump time and to detect energy pileup (see Sec. V B). The time of a photon clump was obtained by averaging over the clump crystals that fired a CFD within 5 ns of the HPHC time. To determine the photon timing resolution, the time difference of coincident  $\gamma$ 's from  $\pi^0$  decays was measured. This distribution has a FWHM of 1.8 ns, which corresponds to 1.27 ns per photon.

#### 3. Drift chamber: vertex

The behavior of the resolution function for the vertex determination was studied with events from  $\mu^+ \rightarrow e^+ e^+ e^- \nu \bar{\nu}$ . Two different quantities were used to estimate how nearly three tracks emerged from a single point in the target. One was simply the square root of the sum of the squares of the distances between the intersection of each reconstructed trajectory with the target plane and the average of the intersection points. The second was a weighted vertex; the contribution of each track was weighted by the calculated uncertainty in the intersection point for that track. The intersection-point uncertainty was about 0.2 cm for a 40-MeV particle normal to the target plane; the uncertainty grows as the angle between the trajectory and the normal to the target increases.

Figure 10(a) shows the number of events as a function of the weighted vertex for all events with  $E + P \leq 100$  MeV and  $\Delta t_{\text{rms}} \leq 0.2$  ns (in time): Figure 10(b) shows the same distribution for  $\Delta t_{\text{rms}} \geq 1$  ns (random). The distribution for the randoms peaks at large values of the weighted vertex while the in-time events have a peak at small values of the weighted vertex as well as a contribution from random coincidences. The derived resolution

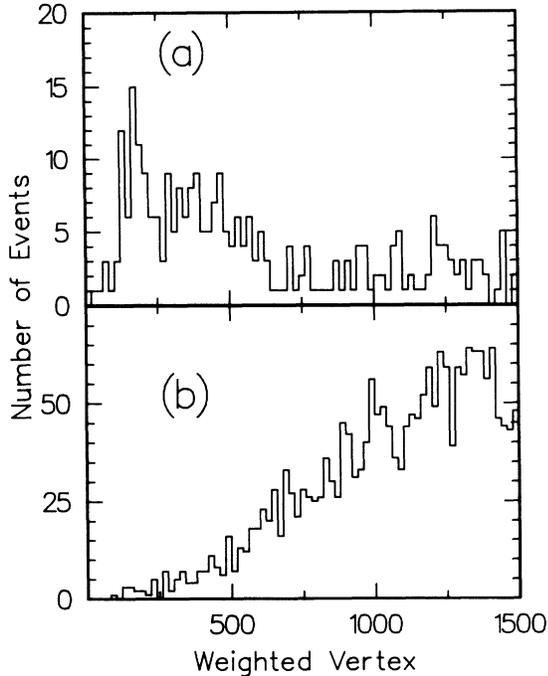


FIG. 10. The weighted vertex distribution for (a) in time and (b) random events.

function for the weighted vertex is in general agreement with, although  $\sim 15\%$  broader than, the expectation from the Monte Carlo program. The broadening is presumably due to the fact that tracks are reconstructed to a target whose position and orientation are not exactly known while the Monte Carlo program generated and reconstructed all tracks from the same assumed location of the target plane.

## V. DATA ANALYSIS

### A. Event-selection criteria

The data set for the experiment consisted of over  $2 \times 10^7$  candidate events written to magnetic tapes, corresponding to a total of  $1.4 \times 10^{12}$  muons stopped in the target. Triggers for all three muon decay modes were enabled simultaneously so the candidate events of each type were tagged and interleaved on the tapes.

The data-analysis algorithms implemented a set of criteria that were designed to remove background without eliminating any real rare muon decays. It was important to reduce the number of candidate events as quickly as possible with a minimum investment of analysis time. There were three sections of the analysis code with cuts after each section. The information available with minimum calculation was from the plastic scintillators, so cuts using this were made first. Cuts using information from the crystals were made after the clump-finding algorithms were executed. Only events that survived this stage were processed by the drift-chamber track-finding algorithms. The scintillator, crystal, and drift-chamber sections of the code were basically independent of trigger

type, but the cut algorithms depended on the mode.

The cuts using the plastic scintillators required that less than 0.25 MeV be deposited in the counters for any photon quadrant for  $\mu \rightarrow e\gamma$  and  $\mu \rightarrow e\gamma\gamma$  candidates, and that there were at least three nonadjacent hodoscope counters coincident to within 1.5 ns for  $\mu \rightarrow eee$  candidates. The NaI(Tl) information was used next to require that there were clumps that agreed with the position of the hodoscope counters that registered. The timing of the NaI(Tl) and hodoscopes were required to agree within 5 ns. The timing cut for  $\mu \rightarrow eee$  events was reimposed after a time-of-flight correction, assuming that the trajectory for each particle started at the center of the target and traveled to the clump position. Next, a good drift-chamber track was required for each electron; the trajectory had to pass within 15 cm of a clump position and had to intersect the target plane with an angle greater than  $3^\circ$ . Timing cuts were imposed again using improved time-of-flight corrections based on the vertex at the target and the trajectories from the drift chamber.

There were  $\sim 10^7$   $\mu \rightarrow e\gamma$  candidate events. Additional first-pass cuts requiring the two NaI(Tl) clump energies each be  $> 35$  MeV,  $E + P < 135$  MeV, and the opening angle between trajectories be  $> 140^\circ$  reduced the number of surviving events to  $2.6 \times 10^5$ .

The analysis for the  $\mu \rightarrow e\gamma\gamma$  mode started from a total of  $\sim 2 \times 10^6$   $e\gamma\gamma$  (three quadrant) and  $10^5$   $e\gamma$  (two quadrant) candidate events. Cuts were made in the first analysis pass requiring the NaI(Tl) clump energies to be  $> 7$  MeV and  $E + P < 135$  MeV. The number of events surviving these cuts was  $5.1 \times 10^4$ .

For the  $\mu \rightarrow eee$  mode, the hardware trigger selected a total of  $\sim 10^7$  candidates to be written on tape. The first analysis pass described above, plus cuts requiring that the energy in the NaI(Tl) was  $> 7$  MeV for each track and  $E + P < 120$  MeV, reduced the number of candidates to  $2 \times 10^4$ .

The raw data for events passing these cuts were written on tape for a second-pass analysis. The purpose of the second-pass analysis was to recalculate kinematic quantities with more accurate NaI(Tl) gains and to impose pile-up rejection, when appropriate. The second pass produced short files containing calculated quantities (such as energies, positions, and corrected times) that could be analyzed quickly.

Two additional cuts were imposed during the final analysis to eliminate photons from known background processes. The first cut, called the inner bremsstrahlung veto, discarded events in which a hodoscope counter, not struck by the positron that triggered the event, registered within 2 ns of a photon. In most of these events, the photon was produced from a  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  decay and the positron from this decay was detected in a hodoscope counter; the positron that triggered the event was in accidental coincidence with the photon. The positron from the inner bremsstrahlung decay tended to have very low energy; 60% deposited less than 2 MeV in the NaI(Tl). The drift-chamber track-reconstruction efficiency was low for these positrons due to multiple scattering. Figure 11 shows the difference in time between the extra hodoscope (for the positron) and the NaI(Tl) (for the photon)

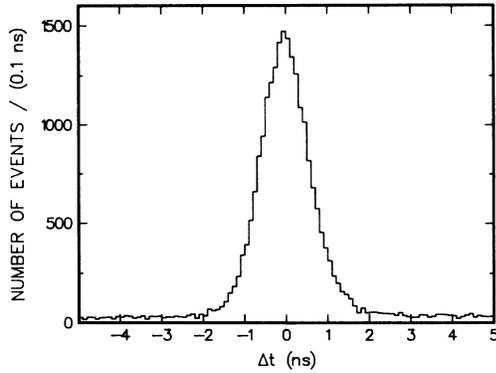


FIG. 11. Relative extra hodoscope—NaI(Tl) timing used for the inner bremsstrahlung veto.

for a sample of 177 424  $\mu^+ \rightarrow e^+ \gamma$  candidates. Requiring  $|\Delta t| < 2.0$  ns discarded 12.3% of the candidate events with a loss in detection efficiency of 0.5%.

The second cut was designed to discard photons generated by positron annihilation in flight in the drift chamber and hodoscope counters. The method was to search for partial tracks in the inner drift-chamber layers that pointed to a photon clump. A map was made of which drift-chamber wires would have been struck by a particle originating at the target for each possible HPHC. To allow for drift-chamber inefficiencies, photons were rejected if there were such hits in at least three drift-chamber layers with no more than one layer missing in the string and a hit in at least one of the two innermost layers. This cut discarded 16% (31%) of the  $\mu^+ \rightarrow e^+ \gamma$  ( $\mu^+ \rightarrow e^+ \gamma \gamma$ ) candidates and reduced the detection efficiency by less than 1%.

### B. Pileup rejection

Because pileup energy was caused by an unrelated particle, the pileup pulses occurred somewhat earlier or later than the desired pulse. The presence of undetected pileup seriously affected the ability of the detector to suppress backgrounds. For example, energy measurements were used to distinguish events due to  $\mu^+ \rightarrow e^+ \gamma$  from  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ . Inner bremsstrahlung events have positron and photon spectra that fall off very rapidly as the energy approaches 52.8 MeV. Consequently, an effect such as pileup, that added energy to an event, tended to make some fraction of the inner bremsstrahlung events indistinguishable from  $\mu^+ \rightarrow e^+ \gamma$  events.

Within a clump there were typically five or six crystals with appreciable deposited energy including any pileup energy. The probability that there was more than 1 MeV of pileup energy was typically 6% for photons and 10% for electrons. The probability was lower for photons because the hardware definition of a photon required that no charged particles passed through the plastic scintillators in the appropriate quadrant of the detector within  $\pm 10$  ns of the trigger.

A simple pileup rejector was used in the analysis of the entire data set. Any crystal within a clump had to have a crystal time within 5 ns of the time of the HPHC if it had more than 7 MeV deposited energy. If any crystal within

the clump failed this test, the entire clump was discarded because it contained pileup energy.

For the data taken in period III, an improved pileup rejection scheme was employed. This scheme was based on a second energy measuring system (Sec. II H). Pileup was detected if the ratio of the pileup energy to primary energy was outside a certain range. The 60-ns pileup gate maximized the sensitivity to signals that were not coincident with the trigger. Tests showed that low-energy (1 MeV) pileup that was out of time by more than 2 ns was detected. The pileup scheme was calibrated with 1e triggers. For each crystal, a two-dimensional histogram of the pileup energy ratio to primary energy was accumulated during the first-pass analysis. A typical histogram is shown in Fig. 12. An unpiled-up pulse had a ratio of  $\sim 270$  in the units of the figure. If a crystal was outside the normal range, it was excluded from the energy sum. In addition, if this crystal had more than 2.5 MeV deposited in it, the entire clump was discarded. This scheme was based on the typical configuration of events; because  $\sim 80\%$  of the crystals in a clump have no “real” energy in them, it was most likely for pileup energy to have occurred in a crystal that had no energy from the triggering particle.

The efficiency of the pileup rejector was determined using data triggered on a free-running pulser gated with the beam. In the absence of pileup, there would have been no energy deposited in the crystals. Consequently, these runs detected pure pileup. These runs were taken at various instantaneous beam intensities. The analysis of these data shows that the pileup rejection scheme reduced the acceptance by 5.2% for photons and 9.5% for positrons at a typical beam intensity. The probability of a photon (positron) to have more than 1.0 MeV of undetected pileup energy was 4.3% (5.7%). Figure 13 shows the pileup probability distribution for photons before and after pileup rejection. Figure 14(a) compares normalized positron energy spectra for  $\mu \rightarrow e \gamma$  candidates at instantaneous muon-stopping rates of 4.2 and 8.4 MHz; the increased pileup at the higher intensity is evident. Figure 14(b) shows the normalized spectra after pileup rejection; no appreciable differences between the spectra are apparent.

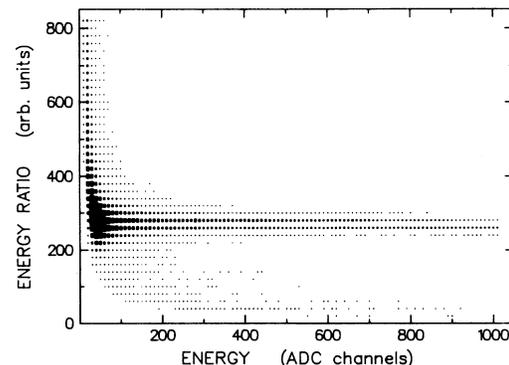


FIG. 12. The measured ratio of pileup energy to primary energy vs primary energy for a typical NaI(Tl) crystal.

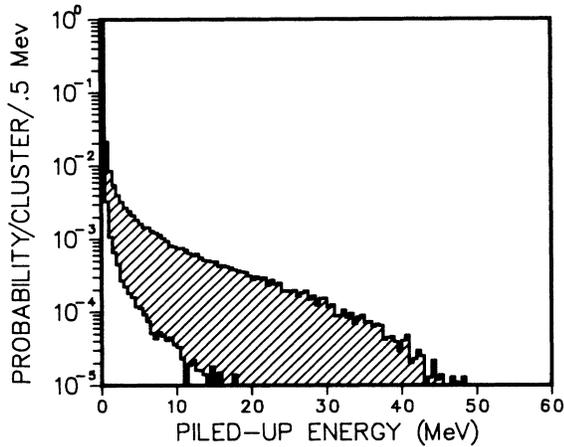


FIG. 13. The probability that a photon clump is piled up vs the amount of pileup energy before (upper curve) and after (lower curve) pileup rejection.

### C. Normalization

Measured branching ratios were calculated from the relation

$$B = \frac{N_{\text{obs}}}{N(\Omega/4\pi)\epsilon_C(\epsilon_T/b)\epsilon_D}, \quad (5)$$

where  $N_{\text{obs}}$  is the number of observed events (or the upper limit of the observed number of events),  $N$  is the total number of muons that decayed during the live time of the detector,  $\Omega/4\pi$  is the solid angle of the detector for

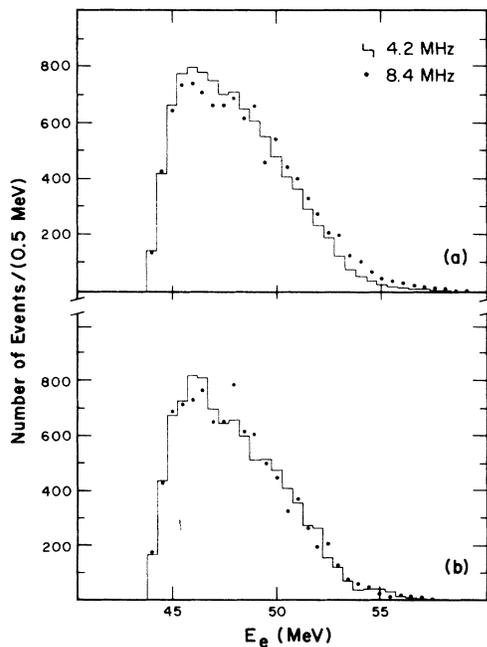


FIG. 14. Normalized spectra of the positron energy for  $\mu^+ \rightarrow e^+ \gamma$  candidates from data with instantaneous muon stopping rates of 4.2 and 8.4 MHz (a) without pileup rejection and (b) after pileup rejection.

the desired events subtended from the target,  $\epsilon_C$  is the efficiency of the data cuts,  $\epsilon_T$  is the efficiency of the trigger requirements,  $b$  is the partial branching ratio for the region of phase space over which events were generated in the Monte Carlo program, and  $\epsilon_D$  is the efficiency associated with the detector. We discuss below how each of these factors was determined for the processes of interest.

#### 1. Number of decaying muons

The total number of decaying muons for each run was determined by using the average of the number of counts in the hodoscope counter located at the center of each quadrant and then correcting for the acceptance of these counters and the live time of the experiment. The acceptance of the counters was determined with the Monte Carlo program. The program recorded the fraction of muon decays that triggered a hodoscope counter. The effects of multiple scattering, bremsstrahlung, and “back splash” from the electromagnetic shower in the NaI(Tl) were included. This method is insensitive to any residual muon polarization because of the symmetry about  $90^\circ$ . The result of this calculation is

$$N = (11.1 \pm 0.2) N_H,$$

where  $N_H$  is the sum of the number of counts in the four central hodoscope counters.

The fractional live time was defined as the fraction of the time with beam present that the detector was actually available for data collection. It was determined by the ratio of the number of counts in the four hodoscope counters gated by the experiment-ready gate and the number of counts gated only by the beam-on gate. The experiment-ready gate included the effects of dead time caused by data read out and the data-acquisition computer being busy. The fractional live time was a function of the instantaneous muon-stopping rate and was 85% averaged over the entire data-taking period. The number of muons stopped during period I was  $2.20 \times 10^{11}$ , during period II was  $1.32 \times 10^{11}$ , and during period III was  $7.97 \times 10^{11}$ , all corrected for the detector live time.

#### 2. Detector acceptance

The product,  $(\Omega/4\pi)\epsilon_C(\epsilon_T/b)$ , was found for each process from the Monte Carlo program. This product was calculated separately for the conditions of the three data-taking periods. As described previously, each type of event was generated in the Monte Carlo code and the response of the detector to each event was simulated. The events then passed through the same reconstruction programs as the data. The results of these calculations are shown in Table II.

For some processes, such as  $\mu^+ \rightarrow e^+ e^- \nu \bar{\nu}$  and  $\mu^+ \rightarrow e^+ \gamma \nu \bar{\nu}$ , events were generated over only a portion of the available phase space. Events in the excluded regions had no chance to satisfy the detector trigger requirements. In this case, the factor  $b$  had to be included to account for the fraction of phase space omitted. This factor was evaluated in the Monte Carlo code by a nu-

TABLE II. Detector acceptance calculated with Monte Carlo Program.

Process	Region of phase space generated	$b$	$\frac{\Omega}{4\pi} \epsilon_C \epsilon_T$	$\epsilon_D$
$\mu^+ \rightarrow e^+ \gamma$	All	1	0.301	0.64
$\mu^+ \rightarrow e^+ \gamma \nu \bar{\nu}$	$E_\gamma > 38$ MeV	$8.6 \times 10^{-7}$	0.0078	0.43
$\mu^+ \rightarrow e^+ \gamma \gamma$	$E_e > 38$ MeV	1	3 quadrant: 0.051	0.50
	All	1	2 quadrant: 0.013	0.62
$\mu^+ \rightarrow e^+ e^+ e^-$	All	1	0.096	0.52
$\mu^+ \rightarrow e^+ e^+ e^- \nu \bar{\nu}$	$E_e > 10$ MeV	$1.8 \times 10^{-6}$	0.0012	0.33
	$E_{e_1} + E_{e_2} + E_{e_3} > 45$ MeV			

merical integration over the allowed region of phase space.

### 3. Detection efficiency

Several detection efficiencies are included in the value of  $\epsilon_D$ . Most of these were dependent on the instantaneous muon-stopping rate and on the event type being considered. The separate factors that enter into  $\epsilon_D$  are listed in Table III. A product of the appropriate factors for each type of event was calculated for each data run. For example, for three-quadrant  $e\gamma\gamma$  triggers,  $\epsilon_D$  is given by

$$\epsilon_D^{3Q} = (\epsilon_\gamma)^2 (\epsilon_{\gamma p})^2 \epsilon_{ep} \epsilon_{DC} \epsilon_M. \quad (6)$$

The first factor  $\epsilon_\gamma$  is the efficiency of identifying a photon quadrant with no extraneous scintillator signal that would have vetoed a valid  $e\gamma\gamma$  trigger. The value of  $\epsilon_\gamma$  was obtained from the number of photon quadrants without scintillator signals when the electronics was randomly triggered during the beam gate. The square of  $\epsilon_\gamma$  appears in Eq. (6) because two different photon quadrants were required for three-quadrant  $e\gamma\gamma$  triggers. The next

two quantities  $\epsilon_{\gamma p}$  and  $\epsilon_{ep}$  are the efficiencies for discarding valid events due to the presence of piled-up photon and positron clumps, respectively. These efficiencies were different for the various data samples because of the different pileup rejection methods. The efficiencies were determined by measuring the percentage of piled-up photons or positrons as a function of the instantaneous muon-stopping rate. The quantity  $\epsilon_{DC}$  is the rate-dependent detection efficiency of the drift chamber and of the track-reconstruction algorithm. The value for  $\epsilon_{DC}$  was found from the percentage of single positron tracks reconstructed in the drift chamber for several beam rates. The final quantity  $\epsilon_M$  takes into account several miscellaneous efficiencies, such as the probability that the  $I$  counter vetoed a valid event originating from the target and that the algorithm for finding positrons annihilating in flight vetoed a good event. The efficiency  $\epsilon_M$  has a (nearly) rate-independent value of 98%. The various  $\epsilon$ 's are defined so that they are independent of each other. Thus, for example,  $\epsilon_{DC}$  in Eq. (6) is measured for electrons that are not piled up.

Even though the hardware NaI(Tl) quadrant energy

TABLE III. Detection efficiencies.

Process	$\epsilon_D$
$\mu^+ \rightarrow e^+ \gamma$	$\epsilon_\gamma \epsilon_{\gamma p} \epsilon_{ep} \epsilon_{DC} \epsilon_M$
$\mu^+ \rightarrow e^+ \gamma \nu \bar{\nu}$	$\epsilon_\gamma \epsilon_{\gamma p} \epsilon_{ep} \epsilon_{DC} \epsilon_M \epsilon_{IB}$
$\mu^+ \rightarrow e^+ \gamma \gamma$	2 quadrant: $\epsilon_\gamma (\epsilon_{\gamma p})^2 \epsilon_{ep} \epsilon_{DC} \epsilon_M$
	3 quadrant: $(\epsilon_\gamma)^2 (\epsilon_{\gamma p})^2 \epsilon_{ep} \epsilon_{DC} \epsilon_M$
$\mu^+ \rightarrow e^+ e^+ e^-$	$\epsilon_{DC}^3 (\epsilon_{ep})^3 \epsilon_M$
$\mu^+ \rightarrow e^+ e^+ e^- \nu \bar{\nu}$	$\epsilon_{DC}^3 (\epsilon_{ep})^3 \epsilon_M \epsilon_{3eN}$
Factors:	
$R$ = instantaneous $\mu^+$ rate (MHz)	
$\epsilon_\gamma$ : Photon quadrant definition = $1 - 0.0363R$	
$\epsilon_{\gamma p}$ : Probability a photon is rejected as piled-up =	$\begin{cases} 1 - 0.00381R; \text{data sets I, II} \\ 1 - 0.0108R; \text{data set III} \end{cases}$
$\epsilon_{ep}$ : Probability an electron is rejected as piled up =	$\begin{cases} 1 - 0.0106R; \text{data sets I, II} \\ 1 - 0.0197R; \text{data set III} \end{cases}$
$\epsilon_{DC}$ : Drift-chamber track-finding efficiency = $1 - 0.0076R$	
$\epsilon_{DC}^{3e}$ : Drift-chamber efficiency for $3e$ events = $1 - 0.032R$	
$\epsilon_{IB}$ : $\mu^+ \rightarrow e^+ \gamma \nu \bar{\nu}$ threshold efficiency =	$\begin{cases} 0.50; \text{data set I} \\ 0.916; \text{data sets II, III} \end{cases}$
$\epsilon_{3eN}$ : $\mu^+ \rightarrow e^+ e^+ e^- \nu \bar{\nu}$ threshold energy = 0.63	
$\epsilon_M$ : Miscellaneous efficiencies ( $\approx 0.98$ )	

threshold was set to  $\sim 37$  MeV for data set I, crystal gain drifts and crystal-to-crystal gain differences resulted in a somewhat higher effective threshold. This resulted in a low detection efficiency for  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  events. The hardware threshold was lowered by  $\sim 5$  MeV for data sets II and III, which significantly increased the inner bremsstrahlung detection efficiency. The detection efficiency, as a function of energy threshold, was determined by comparing the detected positron energy for out-of-time events, for which the positron comes from  $\mu^+ \rightarrow e^+ \nu \bar{\nu}$ , with the observed energy spectrum for 1e events with no energy threshold as discussed in Sec. V H.

#### D. Maximum-likelihood analysis

In most of the analyses described below, the maximum-likelihood method was used to estimate the number of signal events ( $n_s$ ) in the presence of a large number of background events from several sources. The likelihood function is defined to be

$$\mathcal{L}(n_s, n_{s'}) = \prod_{i=1}^N \left[ \frac{n_s}{N} P(\mathbf{x}_i) + \frac{n_{s'}}{N} Q(\mathbf{x}_i) + \frac{n_r}{N} R(\mathbf{x}_i) \right], \quad (7)$$

where  $N$  is the total number of events within the data cuts,  $n_s$  ( $n_{s'}$ ) is the estimate of the number of events from process  $s$  ( $s'$ ), and  $n_r = N - n_s - n_{s'}$  is the estimate of the number of events due to random coincidences. Each event is assumed to arise from one of these processes. The vector  $\mathbf{x}$  has components that are relative times and various kinematic quantities such as angles and particle energies.  $P$ ,  $Q$ , and  $R$  are the normalized probability distributions for events from  $s$ ,  $s'$ , and random background events, respectively. The best estimates for  $n_s$  and  $n_{s'}$  are those that maximize the likelihood for positive values of  $n_s$  and  $n_{s'}$ .

The dependences of  $P$ ,  $Q$ , and  $R$  on relative times were determined from data distributions. The dependences of  $P$  and  $Q$  on kinematic quantities were obtained from Monte Carlo programs;  $R$  was determined from out-of-time data events.

In the absence of a detected signal, the 90%-confidence-level upper limit  $n_s^{90}$  is found from<sup>36</sup>

$$\int_0^N \int_0^{n_s^{90}} \mathcal{L}(n_s, n_{s'}) dn_s dn_{s'} = 0.9 \int_0^N \int_0^N \mathcal{L}(n_s, n_{s'}) dn_s dn_{s'},$$

where, of course,  $\mathcal{L} = 0$  when  $n_s + n_{s'} > N$ .

#### E. $\mu \rightarrow e \gamma$ analysis

As described above,  $2.6 \times 10^5$  events survived the first-pass analysis. The final analysis was performed on the 17073 events that passed the second-pass pileup rejection, the inner bremsstrahlung veto, and positron annihilation algorithms and satisfied  $|\Delta t_{e\gamma}| \leq 5$  ns,  $\theta_{e\gamma} \geq 160^\circ$ ,  $E_\gamma \geq 40$  MeV, and  $E_{e\text{NaI}} > 40$  MeV, where  $E_{e\text{NaI}}$  is the energy deposited in the NaI(Tl) by the electron. The contribution from each of the three running periods appears in Table IV. Figure 9 shows  $\Delta t_{e\gamma}$ , the positron-photon relative timing, for the events from the third running period. This figure shows the broad timing distribution due to random photon-positron triggers and a coincidence peak. The width of the coincidence peak is 1.1 ns (FWHM). The majority of the events in the coincidence peak are due to  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  but any  $\mu^+ \rightarrow e^+ \gamma$  events would also be included. The random timing distribution is slightly rounded by the differing losses of efficiency for the many detector elements in the coincidence logic for large  $|\Delta t_{e\gamma}|$ .

To estimate the number of  $\mu^+ \rightarrow e^+ \gamma$  events in the complete data sample, the maximum-likelihood method was employed. The analysis was performed on the  $N = 10996$  events with  $|\Delta t_{e\gamma}| \leq 2$  ns; the remaining 6077 events were all random coincidences.  $n_s$  ( $n_{s'}$ ) is the estimate of the number of  $\mu^+ \rightarrow e^+ \gamma$  ( $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ ) events. The vector  $\mathbf{x}$  has components  $\Delta t_{e\gamma}$ ,  $\theta_{e\gamma}$ ,  $E_e$ , and  $E_\gamma$ .  $P$ ,  $Q$ , and  $R$  are the normalized probability distributions for  $\mu^+ \rightarrow e^+ \gamma$ ,  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ , and random background events, respectively. They were generated separately for each of the three running periods.

For the  $R$  and  $P$  distributions, the four coordinates  $\Delta t_{e\gamma}$ ,  $\theta_{e\gamma}$ ,  $E_e$ , and  $E_\gamma$  are statistically independent, while for the  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  distribution  $Q$ ,  $\Delta t_{e\gamma}$  is independent of the correlated coordinates  $\theta_{e\gamma}$ ,  $E_e$ , and  $E_\gamma$ . Thus, the probability distributions were factored as follows:

$$P(\mathbf{x}) = P_1(\Delta t_{e\gamma}) P_2(\theta_{e\gamma}) P_3(E_e) P_4(E_\gamma),$$

$$Q(\mathbf{x}) = Q_1(\Delta t_{e\gamma}) Q_2(\theta_{e\gamma}, E_e, E_\gamma),$$

$$R(\mathbf{x}) = R_1(\Delta t_{e\gamma}) R_2(\theta_{e\gamma}) R_3(E_e) R_4(E_\gamma).$$

The distributions  $P_1 = Q_1$  and  $R_1$  were determined by fitting the measured timing distribution (Fig. 9) to a

TABLE IV. Number of candidate events to be analyzed for  $\mu^+ \rightarrow e^+ \gamma$  for the three running periods.

Run period	NaI(Tl) quadrant trigger threshold (MeV)	Pulse-shape pileup rejection	$\mu$ stops in live time	$e\gamma$ triggers	Number of candidate events
I	$\sim 40$	No	$2.20 \times 10^{11}$	$1.24 \times 10^6$	3283
II	$\sim 35$	No	$1.32 \times 10^{11}$	$2.55 \times 10^6$	2702
III	$\sim 35$	Yes	$7.97 \times 10^{11}$	$1.03 \times 10^7$	11088
			$1.149 \times 10^{12}$	$1.40 \times 10^7$	17073

Gaussian peak and a quadratic function, respectively. The distributions  $P_2$ ,  $P_3$ ,  $P_4$ , and  $Q_2$  were obtained from the Monte Carlo simulation of the detector response to  $\mu^+ \rightarrow e^+ \gamma$  and  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ . The distributions  $R_2$ ,  $R_3$ , and  $R_4$  were determined from the 6077 data events with  $|\Delta t_{e\gamma}| > 2$  ns.

The three-dimensional distribution  $Q_2$  needed special handling because of the steep variation in the  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  matrix element. A straightforward Monte Carlo simulation of events above 38 MeV using a reasonable amount of computer time would have resulted in several bins in  $Q_2$  with at most very few events, and hence with statistical fluctuations in the probability for  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$ . This problem was solved in two stages. First, the Monte Carlo simulation was thrown over five sections of phase space, as shown in Fig. 15, to enhance the statistics in the low-probability region near  $E_e = E_\gamma = 52.8$  MeV. The five sections were combined with appropriate weights to give a distribution with better statistical significance over the entire region of interest ( $E_{e\text{NaI}} \geq 40$  MeV,  $E_\gamma \geq 40$  MeV,  $\theta_{e\gamma} \geq 160^\circ$ ). Second, the remaining fluctuations in probability were removed by smoothing the probability of a bin into the neighboring bins of this distribution. The probability-sharing weights for the central and neighboring bins were chosen to satisfy the constraint that, after three smoothing operations, the first-, second-, and third-order moments of the smoothed distribution were within 1% of those of the unsmoothed distribution. Keeping the moments constant in the smoothing prevents shifts in the spectra that would change the estimated number of inner bremsstrahlung events due to the rapid change of shapes. The smoothing weakly affected  $n_{\text{IB}}$ , the estimated number of inner bremsstrahlung events from the likelihood analysis, but did not alter  $n_{e\gamma}$ .

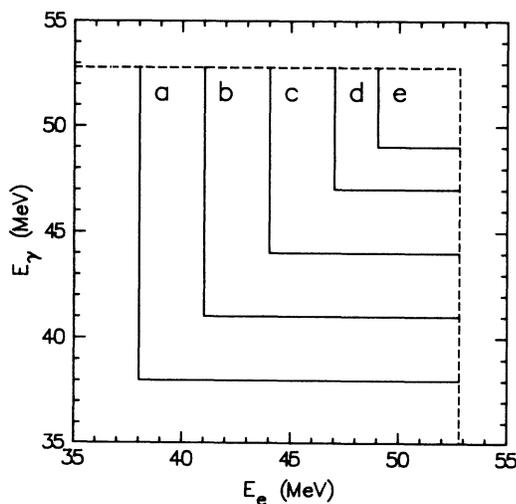


FIG. 15. The separate regions of phase space in which inner bremsstrahlung events were generated with the Monte Carlo program. Each region includes all of the area above and to the right of the boundary up to the kinematic limit. The branching ratio in each region is *a*,  $8.6 \times 10^{-7}$ , *b*,  $1.9 \times 10^{-7}$ , *c*,  $2.9 \times 10^{-8}$ , *d*,  $2.2 \times 10^{-9}$ , and *e*,  $1.65 \times 10^{-10}$ .

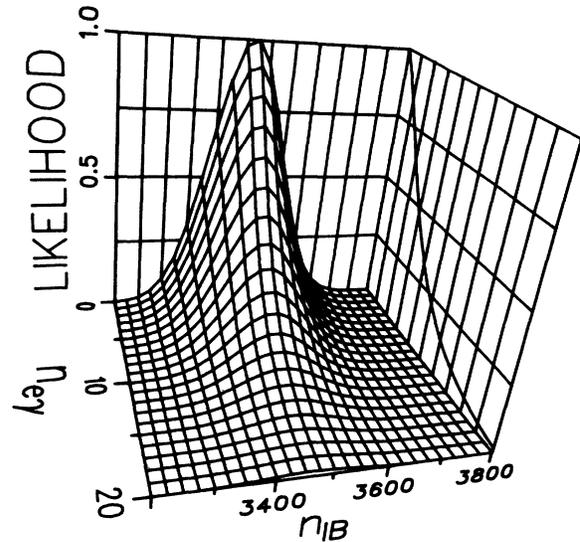


FIG. 16. The normalized likelihood function for all of the data plotted as a function of the number of inner-bremsstrahlung events and the number of  $\mu^+ \rightarrow e^+ \gamma$  events. The projected distribution on the  $n_{e\gamma}$ -likelihood plane is also shown.

For run period I, the distributions  $P_3$ ,  $P_4$ ,  $Q_2$ ,  $R_3$ , and  $R_4$  were corrected to take into account the loss in efficiency in the data due to the higher NaI(Tl) quadrant energy threshold. The correction was determined for each quadrant by comparing the  $E_e$  and  $E_\gamma$  spectra for the data from period I with those from period II.

Figure 16 shows the normalized likelihood function for all of the data. This implies  $n_{\text{IB}} = 3465 \pm 77$  and  $n_{e\gamma} = 0$ . The contributions from the three running periods are shown in Table V. The value for  $n_{\text{IB}}$  compares well with the  $3457 \pm 139$  inner bremsstrahlung events expected in the entire data set; this was derived from the total number of  $\mu$  stops in the live time of the experiment, the detection efficiency, and the fraction of phase space with  $E_{e\text{NaI}} \geq 40$  MeV,  $E_\gamma \geq 40$  MeV, and  $\theta_{e\gamma} \geq 160^\circ$ . The agreement between the observed and expected numbers of inner bremsstrahlung events is better than in Ref. 13 due to the inclusion of the residual nonlinearity of the NaI(Tl) readout system in the Monte Carlo simulation.

The trigger efficiency was much smaller for run period I (0.5) than for periods II and III (0.92) because of the higher NaI(Tl) quadrant energy threshold (Table IV). Because the hardware system detected piled-up energy in each crystal during run period III, it was possible in  $\sim 11\%$  of the cases to ignore an individual piled-up crystal in a positron or photon cluster without discarding the event. On the average, only about one-half of the energy in this crystal was due to pileup so some events near the energy thresholds were lost. The efficiency associated with this effect was about 0.92, and applied only to the data of period III.

The likelihood function distribution implies  $n_{e\gamma} \leq 11$  events (90% C.L.). Table VI shows the limits obtained separately from the three running periods, along with the derived branching-ratio limits. The final result is

TABLE V. Expected number of inner bremsstrahlung (IB) events in the data from the three running periods, and the observed number at the peak of the likelihood curve.

Run period	IB detection efficiency	IB branching ratio $\times$ acceptance	Expected number of IB	Observed number of IB
I	0.290 $\pm$ 0.031	(8.10 $\pm$ 0.16) $\times 10^{-9}$	516 $\pm$ 56	465 $\pm$ 33
II	0.596 $\pm$ 0.022	(8.10 $\pm$ 0.16) $\times 10^{-9}$	639 $\pm$ 27	705 $\pm$ 34
III	0.469 $\pm$ 0.022	(6.16 $\pm$ 0.12) $\times 10^{-9}$	2302 $\pm$ 116	2301 $\pm$ 61
All	0.449 $\pm$ 0.021	(6.69 $\pm$ 0.11) $\times 10^{-9}$	3457 $\pm$ 139	3466 $\pm$ 77

$B_{\mu e \gamma} \leq 4.9 \times 10^{-11}$  (90% C.L.). There is an additional uncertainty of 2.7% in the branching-ratio limits due to the uncertainties in determining the net  $\mu$  stops.

Many tests have been imposed to verify the accuracy of the analysis. (1) Figure 17 compares the distributions for  $E_e$ ,  $E_\gamma$ ,  $\theta_{e\gamma}$ , and  $\Delta t_{e\gamma}$  for the data and for the best maximum-likelihood fit. The agreement of the distributions show that it is a good fit. (2) The number of events from  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  agrees well with the expected number (Table V). This is an extremely sensitive test of the energy scale and thresholds as illustrated in Fig. 18. This agreement persists when the thresholds are lowered to  $E_{e\text{NaI}} \geq 38$  MeV and  $E_\gamma \geq 38$  MeV and the loss in detection efficiency at the lowest energies is included in the probability distributions (7334 $\pm$ 322 expected versus 7275 $\pm$ 104 observed from the likelihood analysis). (3) The number of inner bremsstrahlung events derived from the maximum-likelihood analysis (3466 $\pm$ 77) agrees well with the number of events obtained by fitting the peak of the timing distribution (3530 $\pm$ 100). This is sensitive to differences between  $Q(\mathbf{x})$  and the inner bremsstrahlung data. Again, the agreement is maintained at the lower-energy thresholds (7280 $\pm$ 183 in the prompt timing peak versus 7275 $\pm$ 104 from the likelihood analysis). (4) The data have been analyzed separately for the three run periods. Each period shows consistency between the expected and observed numbers of  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  events (Table V), and zero  $\mu^+ \rightarrow e^+ \gamma$  events. The consistency checks our understanding of the effects of the various hardware configurations, including differing hardware thresholds, beam intensities, and pileup rejection methods. (5) Adding 20 Monte Carlo  $\mu^+ \rightarrow e^+ \gamma$  events to the data stream causes the likelihood curve to peak at  $n_{e\gamma} = 24 \pm 8$  events, thus verifying the correctness of the likelihood analysis code. (6) For period III the data were divided and analyzed according to the four possible qua-

drant orientations. The observed numbers of  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  events agree with the expected numbers to within 5%, though the four observed numbers differ by up to 20% due to different quadrant energy thresholds. Also, each of the four likelihood curves peak near zero  $\mu^+ \rightarrow e^+ \gamma$  events with two of the distributions having their maxima at small positive values; the fluctuation about zero indicates that there is no systematic bias in the analysis.

#### F. $\mu \rightarrow e \gamma \gamma$ analysis

Candidate events for  $\mu^+ \rightarrow e^+ \gamma \gamma$  were present in two separate data streams, the three-quadrant  $e \gamma \gamma$  trigger and the opposite-quadrant  $e \gamma$  trigger. The events from the  $e \gamma$  stream had to have more energy deposited in each quadrant than the NaI(Tl) quadrant energy signal threshold due to the  $e \gamma$  trigger requirements. The  $\mu^+ \rightarrow e^+ \gamma \gamma$  candidates were extracted from the two-quadrant triggers by selecting events with at least one additional photon in either the electron or photon quadrant. There were  $2.2 \times 10^6$  candidate three-quadrant events and  $1 \times 10^5$  candidate two-quadrant events.

A total of  $2.1 \times 10^5$  events passed the first-pass cuts described above. The vast majority of the surviving events had a "split clump," an electromagnetic shower from one particle that was misidentified as being due to two particles. Crystals not included in a clump were assigned to other clump(s) even if they were contiguous with the original clump. Thus an unusually broad electromagnetic shower could be considered to be two separate clumps. One way for this to happen was for particles from the shower to emerge from the front face of the NaI(Tl) and travel to the crystals in an adjacent quadrant. Most of the remaining cuts imposed in the first data analysis pass were designed to eliminate events with split clumps

TABLE VI. The  $\mu^+ \rightarrow e^+ \gamma$  branching ratios derived from the maximum-likelihood fit to the data from the three running periods and the  $\mu$ -stop normalizations.

Run period	$\mu^+ \rightarrow e^+ \gamma$ detection efficiency	$\mu^+ \rightarrow e^+ \gamma$ acceptance	$\mu$ stop times acceptance	Observed number of $\mu^+ \rightarrow e^+ \gamma$	$\mu^+ \rightarrow e^+ \gamma$ branching ratio (90% C.L.)
I	0.663 $\pm$ 0.032	0.302 $\pm$ 0.002	(4.39 $\pm$ 0.21) $\times 10^{10}$	< 8.0	< $1.8 \times 10^{-10}$
II	0.724 $\pm$ 0.015	0.302 $\pm$ 0.002	(2.90 $\pm$ 0.06) $\times 10^{10}$	< 6.5	< $2.2 \times 10^{-10}$
III	0.632 $\pm$ 0.018	0.301 $\pm$ 0.002	(1.52 $\pm$ 0.05) $\times 10^{11}$	< 10.0	< $6.6 \times 10^{-11}$
All	0.650 $\pm$ 0.017	0.301 $\pm$ 0.002	(2.25 $\pm$ 0.06) $\times 10^{11}$	< 11.0	< $4.9 \times 10^{-11}$

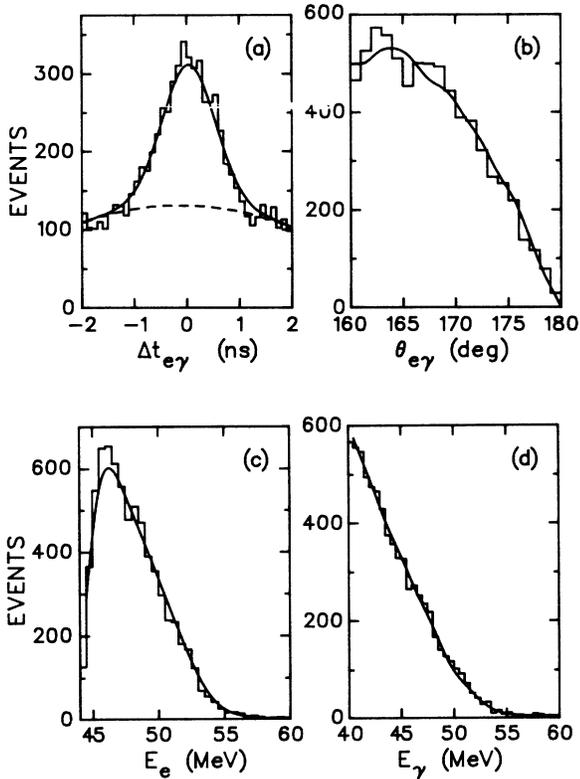


FIG. 17. Spectra from data set III for each of the quantities used in the  $\mu^+ \rightarrow e^+ \gamma$  likelihood analysis. (a) The distribution of  $\Delta t_{e\gamma}$ , the relative timing between the positron and the photon. The solid curve is the fit to these data with a Gaussian for  $\mu \rightarrow e \gamma \nu \nu$  plus a quadratic for randoms. The dashed curve is the random background in the fit. (b)–(d) The distributions of  $\theta_{e\gamma}$ ,  $E_e$ , and  $E_\gamma$ . The curves are the sum of Monte Carlo spectra for  $\mu \rightarrow e \gamma \nu \nu$  and random spectra obtained from out-of-time events with the normalizations determined from the best likelihood fit.

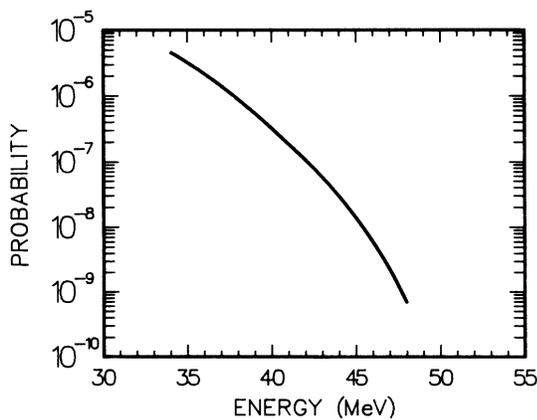


FIG. 18. Branching ratio for  $\mu^+ \rightarrow e^+ \nu \bar{\nu} \gamma$  for  $\theta_{e\gamma} > 160^\circ$  as a function of the energy threshold for  $E_e$  and  $E_\gamma$ .

without rejecting true  $\mu^+ \rightarrow e^+ \gamma \gamma$  events. Events from  $\mu^+ \rightarrow e^+ \gamma \gamma$  tend not to have soft photons but do have a large opening angle between the photons. The cuts required that the opening angle between each pair of particles was at least  $15^\circ$ , that  $E + P \leq 135$  MeV, and that the three particles were coplanar to  $\sim 25^\circ$ . A total of 50 920 events passed these cuts.

The second analysis pass corrected gain drifts and imposed the pileup rejection algorithms. In addition, two-quadrant events with the two photons in the same NaI(Tl) quadrant were eliminated. Only 1% of the  $\mu^+ \rightarrow e^+ \gamma \gamma$  events distributed according to Eq. (4) should have this topology. A total of 42 624 surviving events were written into summary files for subsequent analysis. Because these events were analyzed with the final NaI(Tl) gains, tighter energy cuts could be imposed. Accordingly, the energy in each HPHC was required to exceed 7 MeV and the energy deposited in each of the two NaI(Tl) quadrants by the three detected particles for two-quadrant events was required to exceed the measured quadrant energy thresholds. These cuts reduced the number of candidates to 7083, of which 516 were two-quadrant events.

The sum of the energies of any pair of particles must exceed one-half the muon mass for true  $\mu^+ \rightarrow e^+ \gamma \gamma$  events. Eliminating events in which a two-particle energy sum was less than 51 MeV reduced the data sample to 711 events, of which 75 were two-quadrant events. Additional cuts eliminated inner bremsstrahlung events, events with a positron annihilating in flight, and events in which either photon energy was less than 20 MeV. A total of 364 events remained, of which 33 were two-quadrant events.

Figure 19 shows the distribution of these events as a function of  $\tau = 2t_e - t_{\gamma 1} - t_{\gamma 2}$ . This timing variable is used because all known background processes are flat in  $\tau$ ; true  $\mu^+ \rightarrow e^+ \gamma \gamma$  events would peak at  $\tau = 0$  with a width of 1.5 ns (FWHM) (Refs. 14 and 35). Only those 272 events with  $|\tau| \leq 4$  ns appear in this figure. There is no evidence of a peak. An analysis of the distribution of Fig. 19 implies that there are fewer than 30 coincident events (90% C.L.) in this plot.

The final analysis utilized the kinematic properties of  $\mu^+ \rightarrow e^+ \gamma \gamma$  events as well as their timing characteristics. Two variables were defined to impose the constraints of conservation of momentum. One variable is the magnitude of the vector momentum sum within the decay plane, defined as  $p_{\parallel} = |\mathbf{p}_a + \mathbf{p}_b + \hat{\mathbf{p}}_{ab} \times (\mathbf{p}_c \times \hat{\mathbf{p}}_{ab})|$ , where  $\mathbf{p}_a$  and  $\mathbf{p}_b$  are the momenta of the two particles that are most nearly perpendicular to each other,  $\hat{\mathbf{p}}_{ab}$  is the unit vector normal to the plane defined by those two particles, and  $\mathbf{p}_c$  is the momentum of the third particle. The other variable is the cosine of the planarity angle,  $\cos \alpha = \hat{\mathbf{p}}_c \cdot \hat{\mathbf{p}}_{ab}$ , which is independent of the energy measurements. The final cuts required  $p_{\parallel} \leq 14$  MeV/c,  $|\cos \alpha| \leq 0.2$ , and  $90 \leq E_{\text{tot}} \leq 108$  MeV. These cuts would have removed 15.8% of the true  $\mu^+ \rightarrow e^+ \gamma \gamma$  events. The  $\tau$  distribution of the 20 events that survived these cuts is shown as solid bars in Fig. 19. From this distribution, there are  $2.4 \pm 3.6$ , or  $\leq 7.6$  (90% C.L.) coincident events.

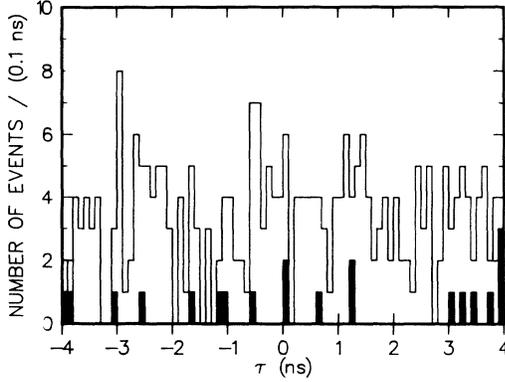


FIG. 19. Distribution of 242 events vs  $\tau = 2t_e - t_{\gamma 1} - t_{\gamma 2}$ . The solid bars are the 20 events that survived the stricter cuts.

This upper limit can be improved by examining in detail the distributions in  $p_{\parallel}$ ,  $\cos\alpha$ , and  $E_{\text{tot}}$ , as well as  $\tau$ . This was done for the nine events with  $|\tau| \leq 1.5$  ns. Figure 20 shows the distributions in these four variables for the nine data events, while Fig. 21 is for three-quadrant  $\mu^+ \rightarrow e^+ \gamma \gamma$  events as generated by the Monte Carlo program. Figure 22 shows the same distributions for background events; the  $\tau$  distribution is assumed to be flat for these events. The background distributions were obtained from the 103 events with  $2.5 \leq |\tau| \leq 5.0$  ns before the final cuts were imposed.

The final analysis for the nine surviving events used the maximum-likelihood method to estimate the number of

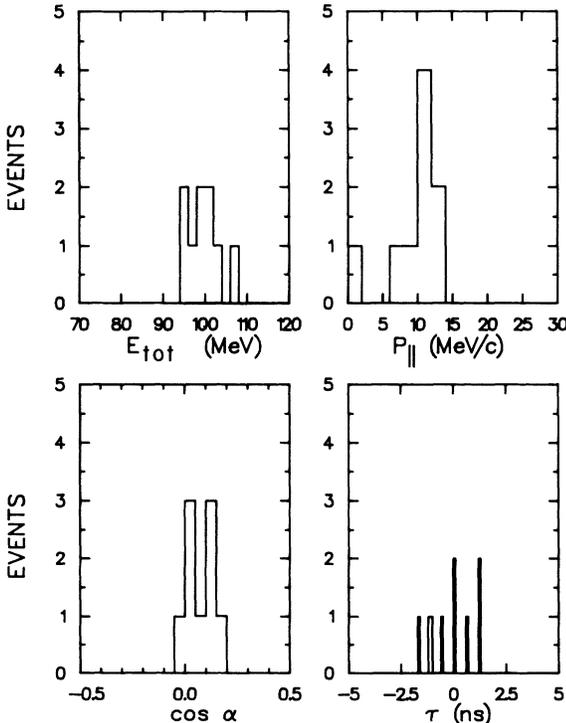


FIG. 20. Distributions of the surviving nine data events in the four analysis variables,  $E_{\text{tot}}$ ,  $p_{\parallel}$ ,  $\cos\alpha$ , and  $\tau$ .

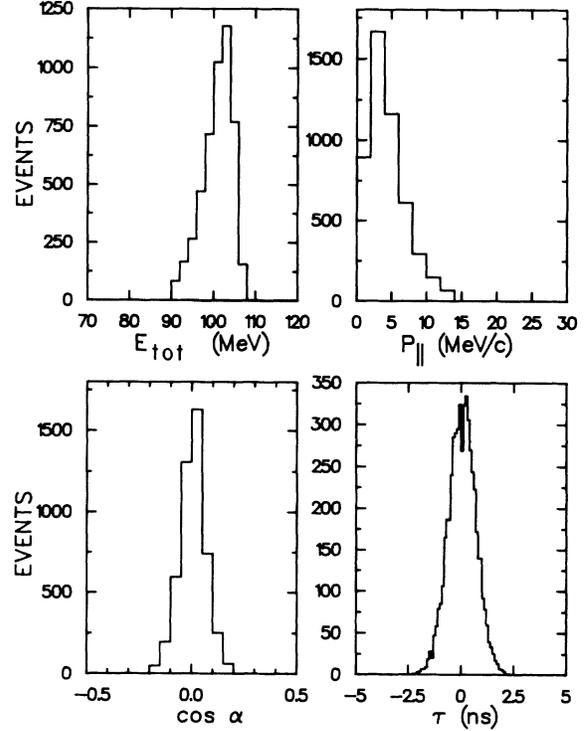


FIG. 21. Monte Carlo-generated probability distributions in the four analysis variables,  $E_{\text{tot}}$ ,  $p_{\parallel}$ ,  $\cos\alpha$ , and  $\tau$ , for the three-quadrant  $e\gamma\gamma$  events with no added piled-up energy.

$\mu^+ \rightarrow e^+ \gamma \gamma$  events. The probability densities, which were functions of the four variables described above, are shown in Figs. 20–22. The  $P(x_i)$  probability density function were determined separately for each run period to account for the differences in hardware configurations, instantaneous beam rates and pileup rejection methods; two- and three-quadrant events were also treated separately. The distributions in the variables  $p_{\parallel}$ ,  $E_{\text{tot}}$  and  $\cos\alpha$  are weakly correlated.

The resulting normalized likelihood function is shown in Fig. 23. The function peaks at  $n_{e\gamma\gamma} = 0$ . The 90%-confidence-level upper limit is  $n_{e\gamma\gamma} = 2.9$  events. This result is insensitive to moderate variations in the probability density distributions. Using the number of muons stopped on the target and the acceptance and efficiency of the apparatus, we obtain

$$B_{e\gamma\gamma} \leq 7.2 \times 10^{-11} \text{ (90\% C.L.) .}$$

The upper limit found from the analysis of the  $\tau$  distribution alone is  $B_{e\gamma\gamma} \leq 6.3 \times 10^{-10}$  (90% C.L.) before the final cuts and  $B_{e\gamma\gamma} \leq 1.9 \times 10^{-10}$  (90% C.L.) after the final cuts. Further details appear in Ref. 35.

#### G. $\mu^+ \rightarrow e^+ e^+ e^-$ analysis

For all  $\mu^+ \rightarrow e^+ e^+ e^-$  candidate events, the positron (and electron) energies were measured in the NaI(Tl), the times were measured in the scintillator hodoscope counters, and the trajectories were measured in the drift chamber. The signature for  $\mu^+ \rightarrow e^+ e^+ e^-$  is that the

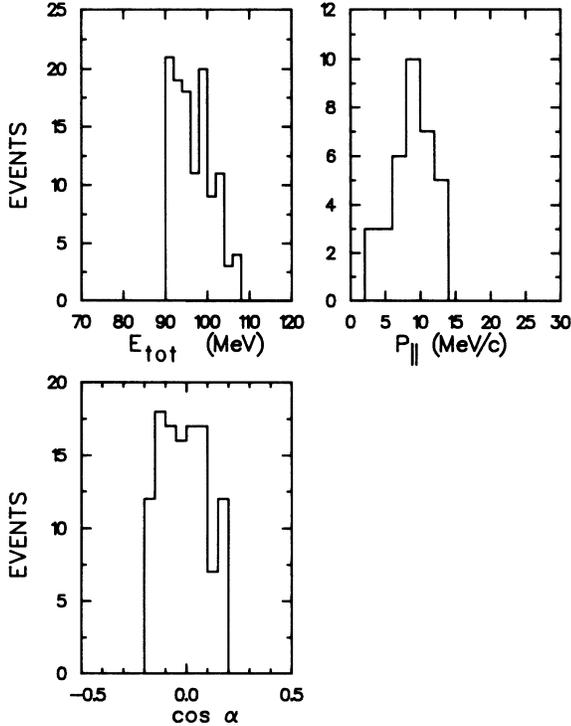


FIG. 22. Background probability distributions in three of the four analysis variables,  $E_{\text{tot}}$ ,  $p_{\parallel}$ , and  $\cos\alpha$ .

three trajectories emerge from a common vertex in the target in time coincidence, that they are coplanar, and that they obey  $E_{\text{tot}} = M_{\mu}$ , and  $|\mathbf{P}_{\text{tot}}| = 0$ .

The vast majority of  $\mu^+ \rightarrow e^+e^+e^-$  triggers were due to the detection of positrons from three uncorrelated muon decays that happened to occur at nearly the same time (randoms). Because the signs of the charges of the particles were not measured, these events had to be eliminated by requiring the above signature. Some triggers were due to  $\mu^+ \rightarrow e^+e^+e^- \nu\bar{\nu}$ . These events satisfy the vertex and timing signatures but do not satisfy the kinematic signatures. In addition, they satisfy  $E + P \leq M_{\mu}$ .

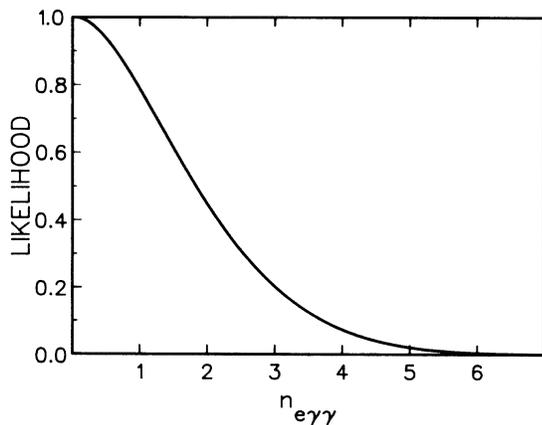


FIG. 23. The normalized likelihood function vs the number of  $\mu^+ \rightarrow e^+ \gamma \gamma$  events,  $n_{e\gamma\gamma}$ , for the nine candidate data events.

The analysis of the data taken during periods II and III is discussed here; the results of the analysis of the earlier data have already been published.<sup>37</sup> Candidate events for  $\mu^+ \rightarrow e^+e^+e^-$  were required to have  $\Delta t_{\text{rms}} \leq 0.4$  ns, an unweighted rms vertex separation  $\leq 4.2$  cm, a weighted vertex  $\leq 780$  [see Fig. 10(a)] and to be coplanar. The coplanarity cut required the absolute magnitude of the scalar triple product between the unit vectors along the three measured trajectories to be  $\leq 0.11$ .

A total of 179 events passed these cuts. Their distribution versus  $E_{\text{tot}}$  and  $|\mathbf{P}_{\text{tot}}|$  is shown in Fig. 24. Also shown is the contour that contains 95% of the  $\mu^+ \rightarrow e^+e^+e^-$  events, as calculated by the Monte Carlo program, assuming a constant matrix element. No data events fall within this contour.

The acceptance of the apparatus for  $\mu \rightarrow 3e$  within the above cuts was  $(9.6 \pm 0.2)\%$  and the detection efficiency was 0.52. Thus we obtain

$$B_{\mu 3e} \leq 4.9 \times 10^{-11} \text{ (90\% C.L.) .}$$

This result can be combined with our earlier result of Ref. 37 to yield

$$B_{\mu 3e} \leq 3.5 \times 10^{-11} \text{ (90\% C.L.) .}$$

This is the final result of this decay mode from our data. At this level, it verifies the result published from SIN-DRUM (Ref. 15) [ $B_{\mu 3e} \leq 1.0 \times 10^{-12}$  (90% C.L.)] with a different detection method.

The performance of the apparatus and the normalization were checked by studying the  $\mu^+ \rightarrow e^+e^+e^- \nu\bar{\nu}$  events. For these events, no cut on coplanarity was made because the three charged particles need not lie in a plane. Figure 25(a) shows  $\Delta t_{\text{rms}}$  for the events that passed the vertex cuts and satisfied  $E + P > 110$  MeV. These events were almost entirely randoms. Figure 25(b) shows  $\Delta t_{\text{rms}}$  for the events with  $E + P < 100$  MeV. This plot has a prominent coincidence peak superimposed on a random background. From these data, a net excess of  $400 \pm 35$  coincidence events was observed. The branching ratio for  $\mu^+ \rightarrow e^+e^+e^- \nu\bar{\nu}$  is predicted from standard electroweak theory<sup>30,38</sup> to be  $B_{\mu 3e\nu\bar{\nu}} = 3.59 \times 10^{-5}$ . This experiment detected events with electron energies greater

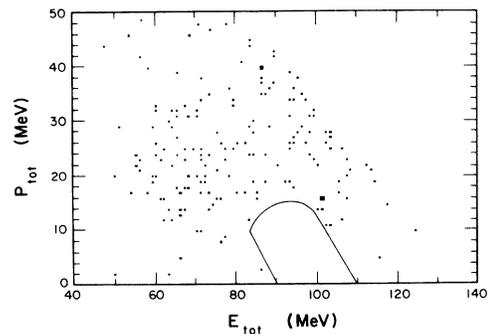


FIG. 24. Distribution of  $E_{\text{tot}}$  vs  $|\mathbf{P}_{\text{tot}}|$  for the 179  $\mu^+ \rightarrow e^+e^+e^-$  candidate events satisfying the final cuts. The contour contains 95% of the  $\mu^+ \rightarrow e^+e^+e^-$  events generated by the Monte Carlo program.

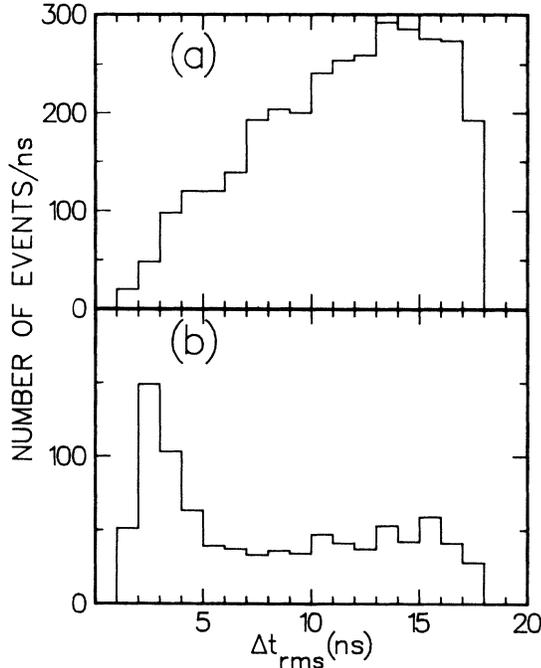


FIG. 25.  $\Delta t_{\text{rms}}$  for events with (a)  $E+P > 110$  MeV and (b)  $E+P < 100$  MeV.

than 10 MeV and  $E_{\text{tot}} > 45$  MeV: for this region of phase space, the prediction is  $B'_{\mu 3e\nu\bar{\nu}} = 1.18 \times 10^{-6}$ . The measured branching ratio is  $B'_{\mu 3e\nu\bar{\nu}} = [1.11 \pm 0.11 \text{ (stat)} \pm 0.08 \text{ (syst)}] \times 10^{-6}$ , in agreement with the theoretical prediction. The largest contribution to the systematic error is the uncertainty in the threshold of the individual NaI(Tl) discriminators: the majority of  $\mu^+ \rightarrow e^+ e^+ e^- \nu\bar{\nu}$  events had at least one particle depositing less than 10 MeV in the NaI(Tl).

#### H. $\mu^+ \rightarrow e^+ \gamma f$

The  $e^+ \gamma$  data were also used to search for light scalar or pseudoscalar bosons<sup>39–41</sup> such as familons, axions, and Majorons. Such bosons might be produced in the radiative decay  $\mu^+ \rightarrow e^+ \gamma f$ , where  $f$  is any particular such boson.

A discussion of the various Lagrangians that could generate  $\mu^+ \rightarrow e^+ \gamma f$  and of the resulting differential distributions was given in Ref. 31. Events from  $\mu^+ \rightarrow e^+ \gamma f$ , with the light boson undetected, are characterized by the fact that  $M_{\text{eff}}^2 \simeq 0$ , where  $M_{\text{eff}}^2$  is the calculated square of the effective mass of the unobserved neutral particle(s). Events from muon inner bremsstrahlung satisfy  $M_{\text{eff}}^2 \geq 0$  while random events have  $-1000 < M_{\text{eff}}^2 < 500$  MeV<sup>2</sup>. This is illustrated in Fig. 26. Strictly speaking, the limit is applicable to  $m_f < 2m_e$  since heavier bosons would be expected to rapidly decay into an  $e^+ e^-$  pair.

To maximize the sensitivity to  $\mu^+ \rightarrow e^+ \gamma f$ , the analysis was performed on data set III, which had the least undetected pileup. The final cuts were relaxed to  $E_\gamma > 38$  MeV,  $E_{e\text{NaI}} > 38$  MeV, and  $\theta_{e\gamma} > 140^\circ$ . To utilize the lower energies, the detection efficiency as a function of energy deposited in each NaI(Tl) quadrant had to be

determined and then put into the Monte Carlo program. The efficiency was determined by comparing  $E_{e\text{NaI}}$  from out-of-time events, for which the positron comes from  $\mu^+ \rightarrow e^+ \nu\bar{\nu}$ , with  $E_{e\text{NaI}}$  from 1e events with no energy threshold. Figure 27 illustrates the comparison for the top quadrant.

The number of  $\mu^+ \rightarrow e^+ \gamma f$  events in the 20015 events with  $|\Delta t_{e\gamma}| < 1.5$  ns was estimated with the maximum-likelihood method as described above. Here  $n_s$  ( $n_s$  is the estimate of the number of  $\mu^+ \rightarrow e^+ \gamma f$  ( $\mu^+ \rightarrow e^+ \nu\bar{\nu} \gamma$ ) events. The vector  $\mathbf{x}$  has components  $M_{\text{eff}}^2$  and  $\Delta t_{e\gamma}$ . Figure 26 shows the  $M_{\text{eff}}^2$  distributions for  $\mu^+ \rightarrow e^+ \gamma f$ ,  $\mu^+ \rightarrow e^+ \nu\bar{\nu} \gamma$ , randoms and for the data events, together with the distribution corresponding to the maximum-likelihood fit. The difference between the fit and data dis-

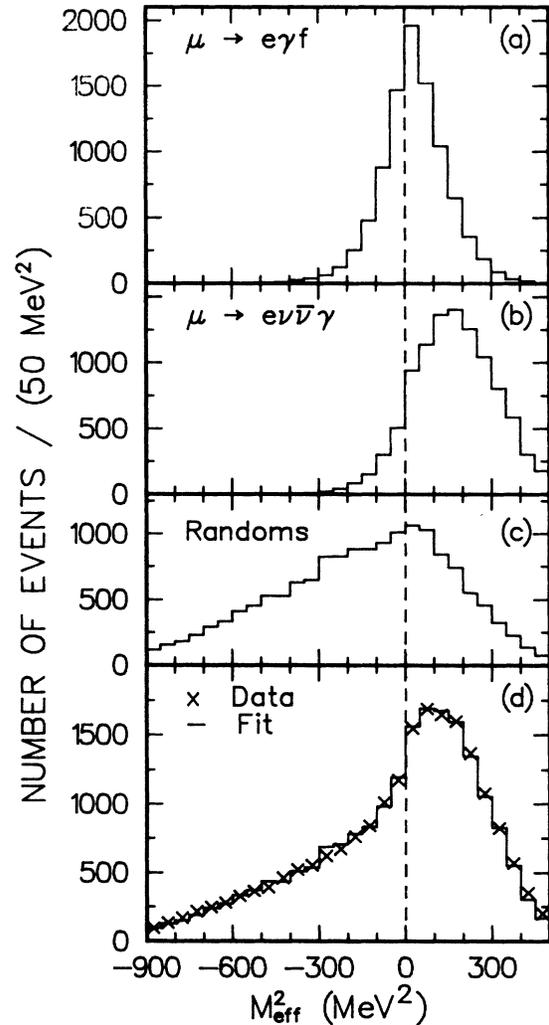


FIG. 26. The distributions for  $M_{\text{eff}}^2$ , the calculated square of the effective mass of the unobserved neutral particle(s) for events from (a)  $\mu \rightarrow e\gamma f$ ; (b)  $\mu \rightarrow e\nu\bar{\nu}\gamma$ ; (c) random events; and (d) the final data set. (a) and (b) were generated with the Monte Carlo program. Also shown in (d) is the sum of  $\mu \rightarrow e\nu\bar{\nu}\gamma$  and random events with the relative normalization determined by the best likelihood fit.

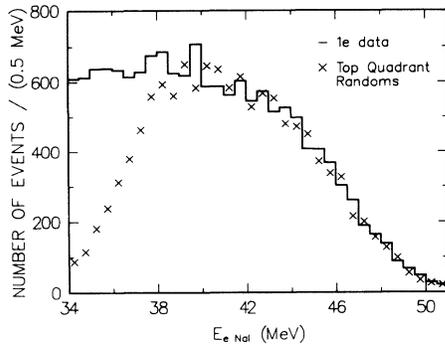


FIG. 27. A comparison of  $E_{e\text{NaI}}$ , the energy deposited in the NaI(Tl), by positrons from out-of-time events and from 1e events for the top quadrant.

tributions corresponds to a  $\chi^2$  of 30 for 27 degrees of freedom. Figure 28 shows the normalized likelihood function. The peak implies  $n_{e\gamma f}=0$ ,  $n_{\text{IB}}=7350\pm 115$ .  $n_{\text{IB}}$  here is slightly different from that quoted in Sec. V E ( $7275\pm 104$ ) due to small differences in the analyses. The  $n_{\text{IB}}$ 's agree with the  $7334\pm 118$  (stat)  $\pm 300$  (syst) inner bremsstrahlung events expected in this data sample. The systematic error is dominated by the uncertainty in the absolute experimental energy scale.

The likelihood-function distribution implies  $n_{e\gamma f} < 165$  events (90% C.L.) including contributions from systematic uncertainties. This implies  $\Gamma(\mu^+ \rightarrow e^+ \gamma f) / \Gamma(\mu^+ \rightarrow \text{all}) < 1.1 \times 10^{-9}$  (90% C.L.) over the region of phase space described above. For  $f = \text{famon}$ , this implies that  $F \geq 3.1 \times 10^9$  GeV, where  $1/F$  is the scale of the symmetry breaking. This limit is a slight improvement on the result in Ref. 31, but it is not as sensitive as the limit derived in Jodidio *et al.*,<sup>42</sup>  $F \geq 9.9 \times 10^9$  GeV. However, our result does not depend on the chirality

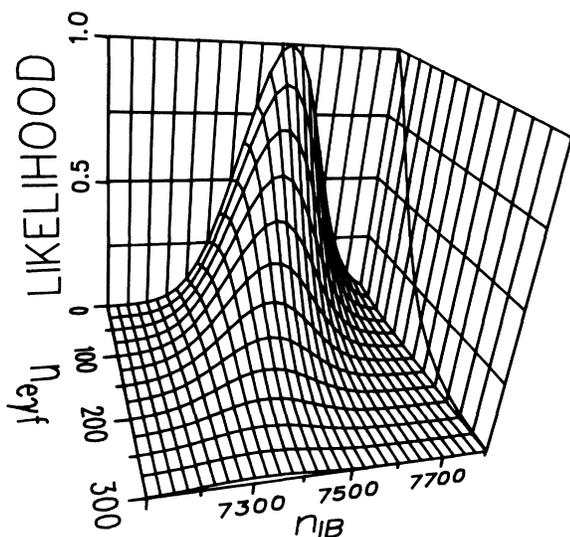


FIG. 28. The normalized likelihood function plotted vs  $n_{\text{IB}}$  and  $n_{e\gamma f}$ , the number of inner-bremsstrahlung and  $\mu \rightarrow e\gamma f$  events, respectively. The projected distribution in the  $n_{e\gamma f}$ -likelihood plane is also shown.

properties of the couplings and so extends the range of validity of conclusions based on the scale of  $F$ . Constraints for other couplings were discussed in Ref. 31; these constraints also improve by a similar small amount.

## VI. DISCUSSION

### A. Theoretical implications of the results

As indicated in the Introduction, many theoretical papers have been written on the subject of lepton-family-number-nonconserving muon decays. These usually describe detailed models with several parameters. The limits implied by our measurements on the masses of new particles in a few specific generically different models are presented below. Within the framework of these models the mass limits are well determined. In other models of the same general type, the limits are illustrative of the mass scales being probed.

Models that predict the existence of lepton-family-number-nonconserving muon decays have been developed with the following attributes: massive neutrinos,<sup>17,43</sup> left-right symmetry,<sup>44</sup> supergravity and supersymmetry,<sup>45</sup> doubly charged leptons,<sup>46</sup> extended Higgs sector,<sup>47</sup> composite particles,<sup>48</sup> horizontal gauge bosons,<sup>49</sup> and technicolor.<sup>50</sup> Some of these models are most tightly constrained by the limits on  $\mu \rightarrow 3e$  (Ref. 15), others by  $\mu^- Z \rightarrow e^- Z$  (Ref. 51), and still others by lepton-family-number-nonconserving kaon decays.

The new limit on  $\mu \rightarrow e\gamma$  provides the most stringent constraints for supersymmetric and composite models. Borzumati and Masiero<sup>52</sup> have examined  $N=1$  supergravity models. In these models, assuming the lower limit on heavy neutrinos of 20 GeV imposed by collider experiments, the new experimental limit on  $\mu \rightarrow e\gamma$  increases the lower bound of the scale of the symmetry breaking from 83 to 100 GeV. Ellis and Nanopoulos<sup>53</sup> examined the phenomenological implications of supersymmetric theories. If gravity breaks the supersymmetry,<sup>54</sup> the new limit increases the lower bound on the mass of the supersymmetric partners of the leptons from 30 to 42 GeV.

A number of authors<sup>55</sup> have studied composite models where the muon is either an excited electron or connected to the electron through dipole transitions of the constituents. The results are qualitatively similar. Using our limit on  $B_{\mu e\gamma}$  and assuming the strength of the constituent interaction has a strength of  $\alpha$ , the fine-structure constant, then the mass of the constituents in these models must be greater than  $6 \times 10^8$  GeV and the inverse size of the interaction distance is  $4.4 \times 10^6$  GeV. The uncertainty due to the coupling constant can be eliminated by combining our  $\mu \rightarrow e\gamma$  and  $\mu \rightarrow e\gamma\gamma$  results. This yields a weaker lower limit on the constituent mass of  $7 \times 10^7$  GeV.

Two papers<sup>56</sup> have studied the relationship between flavor-nonconserving processes and the electric dipole moment of the electron in some extensions to the standard model. Our limit on  $B_{\mu e\gamma}$  implies a dipole moment of  $\leq 10^{-25} e$  cm, which is an order of magnitude lower than the present direct experimental limit.<sup>57</sup>

We see no evidence for the emission of light scalar or pseudoscalar bosons in the decay  $\mu \rightarrow e\gamma f$ . The resulting limits on the couplings are discussed above and in Ref. 31. These limits are independent of the chirality properties of the coupling.

### B. Limitations of the technique

The Crystal Box experiment has taken the technique of detecting electromagnetic showers with inorganic scintillators near the practical limits for rare decay experiments. The limitations are due to singles rates, split clumps, and resolutions. Photons may be detected with lower singles rates by shielding the crystals with a magnetic field, but great cost would be required to have a large solid angle detector beyond the field. The problem of split clumps, due primarily to low-energy photons bouncing around a massive detector, limits the sensitivity of the three-body final-state experiments. The resolutions achievable with available inorganic scintillators are not adequate to suppress backgrounds to substantially lower levels.

The Crystal Box detector was, from its inception, designed to search for several rare decay modes simultaneously. This decision permitted hard-to-measure, but important, properties of the detector to be cross-checked with the different modes. For example, the measurement of the target vertex resolution in the  $3e$  mode was useful in determining the uncertainty in the assumed photon

point of origin in the modes involving photons. However, a general-purpose detector is not optimized for any one channel. Hence, somewhat more sensitive results for any particular branch could be obtained with a detector designed to measure only one mode.

### C. Conclusions

The Crystal Box experiment has been used to search for lepton-family-number-nonconserving muon decays. The branching ratios for the radiative decays  $\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu \gamma$  and  $\mu^+ \rightarrow e^+ e^+ e^- \nu_e \bar{\nu}_\mu$  have been measured and agree with theory. No evidence for the neutrinoless decays has been found. The measured upper limits are  $B_{\mu e \gamma} \leq 4.9 \times 10^{-11}$ ,  $B_{\mu e \gamma \gamma} \leq 7.2 \times 10^{-11}$ , and  $B_{\mu e e e} \leq 3.5 \times 10^{-11}$ , all at the 90% confidence level. We also find no evidence for the emission of a light scalar or pseudoscalar particle in radiative muon decays.

### ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy and the National Science Foundation. Our sincere thanks go to all who contributed their skills to this project. Many of these people have been acknowledged by name in previous publications. Rather than repeat this list, we will just add our thanks to Idan Paiss, Greg Kopp, and Pat Hansen.

- <sup>(a)</sup>Present address: Brookhaven National Laboratory, Upton, NY 11973
- <sup>(b)</sup>Present address: Department of Physics, Princeton University, Princeton, NJ 08544.
- <sup>(c)</sup>Present address: Elektrowatt Ing. AG, Zurich, Switzerland.
- <sup>(d)</sup>Present address: Lawrence Berkeley Laboratory, Berkeley, CA 94720.
- <sup>(e)</sup>Present address: Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.
- <sup>(f)</sup>Present address: Paul Scherrer Institute, CH-5234 Villigen, Switzerland.
- <sup>(g)</sup>Present address: Schlumberger Well Services, Houston, TX 77252-2175.
- <sup>(h)</sup>Present address: Lockheed Missiles and Space Company, Palo Alto, CA 94304.
- <sup>(i)</sup>Present address: Argonne National Laboratory, Argonne, IL 60439.
- <sup>(j)</sup>Present address: Department of Physics, Yale University, New Haven, CT 06520.
- <sup>1</sup>S. L. Glashow, Nucl. Phys. **22**, 579 (1961); A. Salam, in *Elementary Particle Theory: Relativistic Groups and Analyticity (Nobel Symposium No. 8)*, edited by N. Svartholm (Almqvist and Wiksell, Stockholm, 1968), p. 367; S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); **27**, 1688 (1971).
- <sup>2</sup>C. D. Anderson and S. H. Neddermeyer, Phys. Rev. **50**, 263 (1936); S. H. Neddermeyer and C. D. Anderson, *ibid.* **51**, 884 (1937); **54**, 88 (1938); Rev. Mod. Phys. **11**, 191 (1939); see also J. C. Street and E. C. Stevenson, Phys. Rev. **51**, 1005 (1937).
- <sup>3</sup>H. Yukawa, Proc. Phys. Math. Soc. Jpn. **17**, 48 (1935).

- <sup>4</sup>M. Conversi, E. Pancini, and O. Piccioni, Phys. Rev. **71**, 209 (1947).
- <sup>5</sup>E. J. Konopinski and H. M. Mahmoud, Phys. Rev. **92**, 1045 (1953).
- <sup>6</sup>J. Schwinger, Ann. Phys. (N.Y.) **2**, 407 (1957); K. Nishijima, Phys. Rev. **108**, 907 (1957); S. A. Bludman, Nuovo Cimento **9**, 433 (1958).
- <sup>7</sup>G. Feinberg and S. Weinberg, Phys. Rev. Lett. **6**, 381 (1961).
- <sup>8</sup>B. Pontecorvo, Zh. Eksp. Teor. Fiz. **37**, 1751 (1959) [Sov. Phys.-JETP **10**, 1236 (1960)]; M. Schwartz, Phys. Rev. Lett. **4**, 306 (1960).
- <sup>9</sup>G. Danby *et al.*, Phys. Rev. Lett. **9**, 36 (1962).
- <sup>10</sup>H. P. Povel *et al.*, Phys. Lett. **72B**, 183 (1977); A. van der Schaaf *et al.*, Nucl. Phys. **A340**, 249 (1980).
- <sup>11</sup>P. Depommier *et al.*, Phys. Rev. Lett. **39**, 1113 (1977); G. Azeulos *et al.*, *ibid.* **51**, 164 (1983).
- <sup>12</sup>J. D. Bowman *et al.*, Phys. Rev. Lett. **42**, 556 (1979); W. W. Kinnison *et al.*, Phys. Rev. D **25**, 2846 (1982).
- <sup>13</sup>R. D. Bolton *et al.*, Phys. Rev. Lett. **56**, 2461 (1986).
- <sup>14</sup>D. Grosnick *et al.*, Phys. Rev. Lett. **57**, 3241 (1986).
- <sup>15</sup>W. Bertl *et al.*, Nucl. Phys. **B260**, 1 (1985); U. Bellgardt *et al.*, *ibid.* **B299**, 1 (1988).
- <sup>16</sup>A. De Rújula, H. Georgi, and S. L. Glashow, Phys. Rev. D **12**, 147 (1975).
- <sup>17</sup>B. W. Lee and R. E. Shrock, Phys. Rev. D **16**, 1444 (1977).
- <sup>18</sup>H. Albrecht *et al.*, Phys. Lett. **163B**, 404 (1985).
- <sup>19</sup>A. M. Cnops *et al.*, Phys. Rev. Lett. **40**, 144 (1978).
- <sup>20</sup>A. E. Pifer, T. Bowen, and K. R. Kendall, Nucl. Instrum. Methods **135**, 39 (1976); T. Bowen, Phys. Today **38** (7), 22

- (1985); H.-W. Reist *et al.*, Nucl. Instrum. Methods **153**, 61 (1978).
- <sup>21</sup>J. H. Brewer, K. M. Crowe, F. N. Gygax, and A. Schenck, in *Muon Physics III*, edited by V. W. Hughes and C. S. Wu (Academic, New York, 1977), p. 14.
- <sup>22</sup>R. D. Bolton *et al.*, Nucl. Instrum. Methods **188**, 275 (1981); **217**, 173 (1983); **A241**, 52 (1985).
- <sup>23</sup>S. L. Wilson *et al.*, Nucl. Instrum. Methods, **A264**, 263 (1988).
- <sup>24</sup>G. H. Sanders *et al.*, Nucl. Instrum. Methods **180**, 603 (1981).
- <sup>25</sup>V. D. Sandberg *et al.*, Nucl. Instrum. Methods **A234**, 512 (1985).
- <sup>26</sup>W. Flauger, Nucl. Instrum. Methods **161**, 169 (1979).
- <sup>27</sup>C. Bouchiat and L. Michel, Phys. Rev. **106**, 170 (1957).
- <sup>28</sup>J. Dreitlein and H. Primakoff, Phys. Rev. **126**, 375 (1962); J. D. Bowman, T. P. Cheng, L.-F. Li, and H. S. Matis, Phys. Rev. Lett. **41**, 442 (1978).
- <sup>29</sup>C. Fronsdal and H. Uberall, Phys. Rev. **113**, 634 (1959).
- <sup>30</sup>J. Sapirstein (private communication). See also P. M. Fishbane and K. J. F. Gaemers, Phys. Rev. D **33**, 159 (1986).
- <sup>31</sup>T. Goldman *et al.*, Phys. Rev. D **36**, 1543 (1987).
- <sup>32</sup>R. L. Ford and W. R. Nelson, Stanford Linear Accelerator Center Report No. SLAC-PUB-210, 1978 (unpublished).
- <sup>33</sup>R. E. Mischke and R. A. Williams, Los Alamos National Laboratory Report No. LA-UR-82-1952, 1982 (unpublished).
- <sup>34</sup>S. L. Wilson, Ph.D. thesis, Stanford University, Los Alamos National Laboratory Report No. LA-10471-T, 1985.
- <sup>35</sup>D. P. Grosnick, Ph.D. thesis, University of Chicago, Los Alamos National Laboratory Report No. LA-10937-T, 1987.
- <sup>36</sup>S. L. Meyer, *Data Analysis for Scientists and Engineers* (Wiley, New York, 1975), pp. 330 and 331.
- <sup>37</sup>R. D. Bolton *et al.*, Phys. Rev. Lett. **53**, 1415 (1984).
- <sup>38</sup>D. Yu Bardin, Ts. G. Istathov, and G. V. Mitsel'makher, Yad. Fiz. **15**, 284 (1972) [Sov. J. Nucl. Phys. **15**, 161 (1972)].
- <sup>39</sup>D. Reiss, Phys. Lett. **115B**, 217 (1982); F. Wilczek, Phys. Rev. Lett. **49**, 1549 (1982).
- <sup>40</sup>R. D. Peccei and H. Quinn, Phys. Rev. Lett. **38**, 1440 (1977); S. Weinberg, *ibid.* **40**, 223 (1978); F. Wilczek, *ibid.* **40**, 220 (1978); J. E. Kim, *ibid.* **43**, 103 (1979); M. B. Wise, H. Georgi, and S. L. Glashow, *ibid.* **47**, 402 (1981); M. Dine, W. Fischler and M. Srednicki, Phys. Lett. **104B**, 199 (1981).
- <sup>41</sup>G. B. Gelmini and M. Roncadelli, Phys. Lett. **99B**, 411 (1981); Y. Chikashige, R. N. Mohapatra, and R. D. Peccei, *ibid.* **98B**, 265 (1981); Phys. Rev. Lett. **45**, 1926 (1980).
- <sup>42</sup>A. Jodidio *et al.*, Phys. Rev. D **34**, 1967 (1986).
- <sup>43</sup>T. P. Cheng and L.-F. Li, Phys. Rev. D **16**, 1425 (1977); **16**, 1565 (1977); Phys. Rev. Lett. **38**, 381 (1977); **45**, 1908 (1980); B. W. Lee, S. Pakvasa, R. E. Shrock, and H. Sugawara, *ibid.* **38**, 937 (1977); **38**, 1230(E) (1977); S. B. Treiman, F. Wilczek, and A. Zee, Phys. Rev. D **16**, 152 (1977); G. Altarelli, L. Baulieu, N. Cabibbo, L. Maiani, and R. Petronzio, Nucl. Phys. **B125**, 285 (1977); H. Fritzsch, Phys. Lett. **67B**, 451 (1977); J. D. Bjorken, K. Lane, and S. Weinberg, Phys. Rev. D **16**, 1474 (1977); W. J. Marciano and A. I. Sanda, Phys. Lett. **67B**, 303 (1977); S. M. Bilenky, S. T. Petcov, and B. Pontecorvo, *ibid.* **67B**, 309 (1977); V. Barger and D. V. Nanopoulos, Nuovo Cimento **44A**, 303 (1978).
- <sup>44</sup>J. C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974); R. N. Mohapatra and J. C. Pati, *ibid.* **11**, 566 (1975); **11**, 2558 (1975); R. N. Mohapatra and R. E. Marshak, Phys. Rev. Lett. **44**, 1316 (1980); **44**, 1644(E) (1980); R. N. Mohapatra and G. Senjanović, Phys. Rev. D **23**, 165 (1981); Riazuddin, R. E. Marshak, and R. N. Mohapatra, *ibid.* **24**, 1310 (1981); P. M. Fishbane, K. Gaemers, S. Meshkov, and R. E. Norton, *ibid.* **32**, 1186 (1985).
- <sup>45</sup>For a review of supersymmetry, see P. Fayet and S. Ferrara, Phys. Rep. **32**, 249 (1977); S. Ferrara, *ibid.* **105**, 5 (1984); P. Fayet, *ibid.* **105**, 21 (1984); C. H. Llewellyn Smith, *ibid.* **105**, 53 (1984); D. V. Nanopoulos, *ibid.* **105**, 71 (1984); A. Savoy-Navarro, *ibid.* **105**, 91 (1984); J. Ellis, *ibid.* **105**, 121 (1984).
- <sup>46</sup>F. Wilczek and A. Zee, Phys. Rev. Lett. **38**, 531 (1977).
- <sup>47</sup>J. D. Bjorken and S. Weinberg, Phys. Rev. Lett. **38**, 622 (1977); G. C. Branco, Phys. Lett. **68B**, 455 (1977); W. Konetschny and W. Kummer, *ibid.* **70B**, 433 (1977); A. Zee, *ibid.* **93B**, 389 (1980).
- <sup>48</sup>For a review of composite models, see M. Peskin, in *Proceedings of the 1981 International Symposium on Lepton and Photon Interactions at High Energies*, Bonn, edited by W. Pfeil (Bonn University Physics Institute, Bonn, Germany, 1981), p. 880; L. Lyons, Prog. Part. Nucl. Part. Phys. **10**, 227 (1983).
- <sup>49</sup>R. N. Cahn and H. Harari, Nucl. Phys. **B176**, 135 (1980); G. L. Kane and R. Thun, Phys. Lett. **94B**, 513 (1980); O. Shanker, Phys. Rev. D **23**, 1555 (1981); Nucl. Phys. **B185**, 382 (1981); D. R. T. Jones, G. L. Kane, and J. P. Leveille, *ibid.* **B198**, 45 (1982).
- <sup>50</sup>For a review of technicolor, see E. Farhi and L. Susskind, Phys. Rep. **74**, 277 (1981). See also J. Ellis, D. V. Nanopoulos, and P. Sikivie, Phys. Lett. **101B**, 387 (1981); J. Ellis, M. K. Gaillard, D. V. Nanopoulos, and P. Sikivie, Nucl. Phys. **B182**, 529 (1981); S. Dimopoulos and L. Susskind, *ibid.* **B155**, 237 (1979); S. Dimopoulos and J. Ellis, *ibid.* **B182**, 505 (1981).
- <sup>51</sup>S. Ahmad *et al.* Phys. Rev. Lett. **59**, 970 (1987).
- <sup>52</sup>F. Borzumati and A. Masiero, Phys. Rev. Lett. **57**, 961 (1986).
- <sup>53</sup>J. Ellis and D. V. Nanopoulos, Phys. Lett. **110B**, 44 (1982).
- <sup>54</sup>E. Cremmer *et al.*, Phys. Lett. **122B**, 41 (1983).
- <sup>55</sup>H. Terazawa *et al.*, Phys. Lett. **112B**, 387 (1982); Y. Tomozawa, Phys. Rev. D **25**, 1448 (1982); P. G. Esposito, Nuovo Cimento **92A**, 328 (1986).
- <sup>56</sup>S. M. Barr and A. Masiero, Phys. Rev. Lett. **58**, 187 (1987); Y. Kizukuri, Phys. Lett. **B 185**, 183 (1987).
- <sup>57</sup>H. Gould, Phys. Rev. Lett. **24**, 1091 (1970); P. G. H. Sanders and R. M. Sternheimer, Phys. Rev. A **11**, 475 (1975).