UNIVERSITY OF CALIFORNIA

Los Angeles

A Measurement of the Muonic Dalitz Decay of the Neutral Kaon

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Physics

by

Matthew Brandon Spencer

1995

The dissertation of Matthew Brandon Spencer is approved.

Robert Cousins

Thomas Muller

Jay Hauser

Simon Cherry

Katsushi Arisaka, Committee Chair

University of California, Los Angeles

1995

So long and thanks for all the fish.

Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I would like to use this opportunity to thank all of the people associated with the experiment E799 and its precursors. Time-scales are often very long in the modern era of high energy physics. The concept of analyzing data from a previous experiment and working on hardware for a newer one is not uncommon nowadays. I thank my adviser Katsushi Arisaka, for embracing this idea and supporting me through to its conclusion. This work would surely not have been possible without his support and encouragement.

I thank the spokespersons for E799, Yau Wai Wah and Taku Yamanaka who encouraged me to embark upon this analysis and continued taking a strong interest throughout the project. Special thanks to Bruce Winstein and Bob Hsiung who were enthusiastic and supportive at all times.

Bob Tschirhart was the driving force behind this work. His knowledge of all aspects of the experiment and the extraordinarily congenial working environment provided by the kaon group at Fermilab made this experience enjoyable. In this regard I would especially like to thank Hogan and Julie for sharing their office with a starving graduate student.

I would like to mention especially all of the graduate students and postdocs who were always willing to take time out to help with questions despite being under time pressures themselves. In particular Kevin McFarland and Debbie Harris whose knowledge of the analysis code and the calibration files, and the experiment in general simplified my life greatly. Eric Ramberg's considerable expertise and humor helped facilitate this work. Elliott Cheu's knowledge of all aspects of the software was invaluable.

Many people have made helpful comments on this work during analysis meetings, personal conversations, and reading drafts. I take this opportunity to thank them all individually: Katsushi Arisaka, Tony Barker, Roy Briere, Elliott Cheu, Mike Crisler, Ping Gu, Peter Haas, Debbie Harris, Bill Hogan, Bob Hsiung, John Jennings, Doug Jensen, Kevin McFarland, John Matthews, Hogan Nguyen, Tsuyoshi Nakaya, Stephen Pordes, Erik Ramberg, Ron Ray, Doug Roberts, Steve Schnetzer, Bernhard Schwingenheuer, Bill Slater, Sunil Somolwar, Gordon Thompson, Bob Tschirhart, Matt Weaver, Julie Whitmore, Bruce Winstein.

VITA

October 3, 1966	Born, Takapuna, New Zealand
1987	B.Sc., Physics Auckland University, New Zealand
1989	M.Sc., Physics Auckland University, New Zealand
1988–1989	Teaching Assistant Auckland University, New Zealand
1990–1993	Teaching Assistant University of California, Los Angeles
1993–1995	Research Assistant University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

W. H. Allen *et al.*, Survey of UHE Gamma Ray Emission in the Magellanic Clouds, Astrophysical Journal 239–248 (1993).

W. H. Allen *et al.*, Possible Observation of 100-TeV Gamma Rays from the Active Galaxy Centaurus-A, Astroparticle Physics 1, 269–276 (1993).

W. H. Allen *et al.*, Search for Point Sources of Ultra-High Energy Gamma Rays in the Southern Sky, Physical Review D48, 466–478 (1993).

I. A. Bond *et al.*, Upper Limit for Ultra High-Energy Gamma Rays from SN1987A obtained by the Cerenkov Technique at Large Zenith Angles, Astrophysical Journal L17–19 (1993).

M. B. Spencer, A Measurement of $K_L \rightarrow \mu^+\mu^-\gamma$, Presentation to the American Physical Society Division of Particles and Fields, Albuquerque, New Mexico (1994).

M. B. Spencer *et al.*, A Measurement of the Branching Ratio and Form Factor for $K_L \rightarrow \mu^+\mu^-\gamma$, Physical Review Letters 74, 17, 3323 (1995).

ABSTRACT OF THE DISSERTATION

A Measurement of the Muonic Dalitz Decay of the Neutral Kaon

by

Matthew Brandon Spencer

Doctor of Philosophy in Physics

University of California, Los Angeles, 1995

Professor Katsushi Arisaka, Chair

We report on a measurement of the decay $K_L \rightarrow \mu^+\mu^-\gamma$ from Fermilab experiment E799. We observe 207 candidate signal events with an estimated background of 10.5 ± 4.0 events and establish $B(K_L \rightarrow \mu^+\mu^-\gamma) = (3.23 \pm 0.23(stat) \pm 0.19(sys)) \times 10^{-7}$. This provides the first measurement of the $K\gamma\gamma^*$ form factor in the muonic Dalitz decay mode of the K_L.

Chapter 1

Short and Long Distance Physics

The $K_L \to \gamma\gamma$ transition has received considerable attention in the literature due to the fact that it is dominated by so-called *long distance* contributions [1, 2, 3, 4, 6, 7]. These typically involve QCD effects that are significant in many related kaon decays such as $K_L \to \mu^+\mu^-\gamma$, $K_L \to e^+e^-\gamma$, $K_L \to \mu^+\mu^-e^+e^-$, $K_L \to e^+e^-e^+e^-$, and $K_L \to \mu^+\mu^-\mu^+\mu^-$. The $\Delta I = 1/2$ rule in kaon decays is dominated by long distance processes and understanding their effects could lead to an explanation for this empirical rule. These type of QCD processes also dominate the the decay $K_L \to \mu^+\mu^-$ while the short distance contributions to $K_L \to \mu^+\mu^-$ are sensitive to Standard Model parameters including the mass of the top quark. If the long distance contributions could be calculated accurately a bound could be placed on the top quark mass. With the recent discovery of the top quark at a mass of about 180 GeV [8]

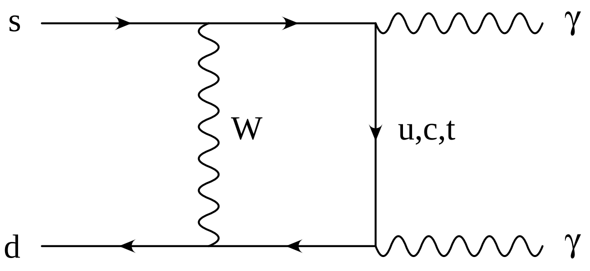

Figure 1.1: Lowest order electroweak process contributing to $K_L \to \gamma\gamma$.

this could now prove to be a good testing ground for the consistency of the standard model.

In this chapter we will discuss long distance processes and in particular the parameterization of Bergström, Massó, and Singer which is often used to estimate the $K_L \to \gamma\gamma$ transition rate when one photon is an intermediate state [1].

1.1 Physics of the $K_L \to \gamma\gamma$ Transition

The lowest order electroweak process that contributes to the $K_L \to \gamma\gamma$ rate is shown in Fig. 1.1. The amplitude for this process has been evaluated and the result is [9]

$$\begin{aligned}\langle \gamma_2(q_1,\epsilon_1)\gamma_2(q_2,\epsilon_2)|\mathcal{M}|K_L\rangle &= \frac{G_F\alpha}{\pi} i f_K \epsilon_{\mu\nu\alpha\beta}\epsilon_1^\mu \epsilon_2^\nu q_1^\alpha q_2^\beta \\ &\times \sum_{i=u,c,t} \mathrm{Re}(V_{jd}V_{js}^*)[A_j^{(i)} + A_j^{(r)}] \qquad (1.1)\end{aligned}$$

where $\langle 0|\bar{s}\gamma^{\mu}\gamma_5 d|K^0\rangle = if_K p_K^{\mu}$, ϵ_j^{ρ} is the polarization vector of the j^{th} photon, q_j^{ρ} is the 4-momentum of the j^{th} photon and V_{ij} are elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The constant f_K can be measured in the decay $K^+ \rightarrow \mu^+\nu_{\mu}$. The terms A_j are the result of integrating over the momenta in the loop of Fig. 1.1. They are functions of $x_i = m_i^2/M_W^2$ and can be found in Ref. [9]. Equation 1.1 can be written simply in the rest frame of the kaon as

$$\langle\gamma_1(q_1,\epsilon_1)\gamma_2(q_2,\epsilon_2)|\mathcal{M}|K_L\rangle = F(K_L \rightarrow \gamma\gamma)\vec{\epsilon_1} \times \vec{\epsilon_2} \cdot \hat{q}$$

where $\hat{q} \cdot \hat{q} = 1$ and [10]

$$\begin{aligned} F(K_L \rightarrow \gamma\gamma) &= \frac{M_K^2}{2}\frac{G_F\alpha}{\pi} i f_K \sum_{i=u,c,t} Re(V_{jd}V_{js}^*)[A_j^{(i)} + A_j^{(r)}] \\ &\approx 1.2 \times 10^{-7} i[A_u^{(i)} + A_u^{(r)}]. \end{aligned}$$

We have assumed here that the u-quark term dominates. The decay rate is then given by [11]

$$\begin{aligned} \text{Rate}(K_L \rightarrow \gamma\gamma) &= \frac{|F(K_L \rightarrow \gamma\gamma)|^2}{16\pi M_K}. \\ &= 8.5 \times 10^2 |A_u^{(i)} + A_u^{(r)}|^2 \ \ s^{-1} \end{aligned}$$

to be compared with the experimentally determined value of $1.1 \times 10^4\ s^{-1}$. Thus if we assume $|A_u^{(i)} + A_u^{(r)}|^2 \approx 1$ the experimentally observed rate is higher by about a factor of 10 than the contribution from the process shown in Fig. 1.1. This conclusion was originally reached by Gaillard and Lee [12] and also by Voloshin and Shabalin [13] who estimated that Eqn 1.1 can only contribute 3% to the observed rate. The integral functions in the evaluation of these processes are dominated by masses on the scale of M_W and the corresponding distance scales are $d \sim 1/M_W$. These processes are often termed *short distance* in the literature.

A possible explanation for this deficit is to extend the sum in Eqn. 1.1 over more generations $i = u, c, t, ...$ where the higher generation quarks must have higher masses. However it is shown in Ref. [9] that the largest contributions to Eqn. 1.1 are from the u-quark and decrease rapidly as the quark mass increases. Since $m_u << m_c << m_t < m_j$ it is unlikely that more generations could account for the observed rate. There are other contributions to the process as shown in Fig. 1.2: QCD radiative corrections as in the top figure (there are two related processes) and the penguin diagrams such as shown in the bottom figure. The matrix elements of these processes are unfortunately divergent and typically involve the dynamical integral [14, 15]

$$I(\mu) \quad = \quad \int \frac{d^4k}{2\pi^4} \frac{1}{k^4(k^2 - M_W^2)}$$

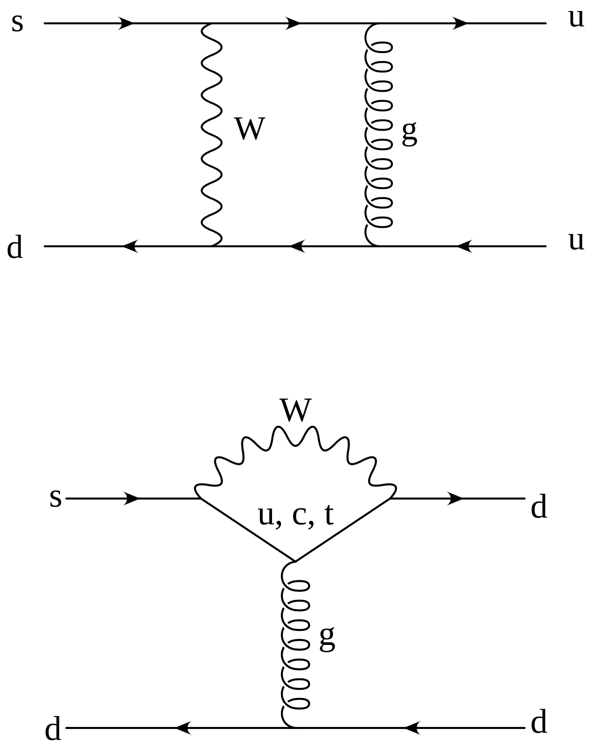

Figure 1.2: Top: QCD radiative correction to the process shown in Fig. 1.1. Bottom: strong “penguin” contribution to the $K_L \to \gamma\gamma$ rate.

$$= \frac{-i}{8\pi^2 M_W^2} \ln \left[\frac{\kappa}{(\kappa^2 + M_W^2)^{1/2}} \right]_\mu^\infty$$

where μ is a cut-off. In principle the results of the calculation must be independent of μ which is supposed to be only an artifact of the technique in calculating. As the limit on the integrals are typically assumed to be $\mu \sim m_c$ the distance scales are $d \sim 1/m_c$ and these contributions are termed *long distance.*

1.2 The $\Delta I = 1/2$ Rule

If we consider the decays $K_S \rightarrow \pi^0\pi^0$ and $K^+ \rightarrow \pi^+\pi^0$ we see quite different rates for these reactions

$$\frac{\mathrm{Rate}(K^+ \rightarrow \pi^+\pi^0)}{\mathrm{Rate}(K_S \rightarrow \pi^0\pi^0)} \approx \frac{1}{200}. \tag{1.2}$$

This is an example that is quite common in $\Delta S = 1$ transitions where processes that change the total isospin by 1/2 are always favored over transitions where the change in isospin is 3/2 [15, 16, 17]. In the above reactions the initial isospin of the kaons is always 1/2 (the K^+ and K^0 form an isospin doublet). But the final state $\pi^0\pi^0$ must be an $I = 0$ isospin state whereas the $\pi^+\pi^0$ state has isospin $I = 2$. The processes shown in Fig. 1.2 are exactly those that give rise the rates in Eqn. 1.2 and therefore a complete understanding of the $\Delta I = 1/2$ rule will also give predictions of the rates for various processes that would be dominated by the $K_L \rightarrow \gamma\gamma$ transition. Such processes are assumed to include $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow e^+e^-\gamma$.

1.3 $K_L \rightarrow l^+l^-\gamma$

The decay $K_L \rightarrow l^+l^-\gamma$ is thought to be dominated by a $K_L \rightarrow \gamma\gamma^*$ coupling where the off-shell photon subsequently decays into an l^+l^- pair. In that case an expression for the amplitude can be written without knowing the details of the process that creates the two-photon state by introducing one unknown function. This function can only depend on the invariant mass of the off-shell photon and is written $f(q^2)$, q being the 4-momentum of the off-shell photon. The amplitude for the process $K_L \rightarrow l^+l^-\gamma$ is

$$A(K_2^0 \rightarrow l^+l^-\gamma) = e\frac{f(q^2)}{q^2}\epsilon_\mu \epsilon^{\mu\nu\rho\sigma} q_{1\rho} q_{2\sigma} \bar{u}(k')\gamma_\nu v(k) \tag{1.3}$$

where k, k' are the 4-momenta of the l^+, l^-, and ϵ_μ is the polarization vector for the on-shell photon. From this one integrates over all possible final state phase space to obtain the differential decay rate as a function of q^2,

$$\frac{d\Gamma(K_2^0 \rightarrow l^+l^-\gamma)}{d(q^2/m_K^2)} \propto \frac{|f(q^2)|^2}{q^2/m_K^2}\left(1-\frac{q^2}{m_K^2}\right)^3\left(1+\frac{2m_l^2}{q^2}\right)\left(1-\frac{4m_l^2}{q^2}\right)^{1/2}. \tag{1.4}$$

The function f then describes the physics at the $K\gamma\gamma^*$ vertex. This equation was first written down by Kroll and Wada in 1955 [18]. Equation 1.4 shows immediately some of the differences between the $e^+e^-\gamma$ and $\mu^+\mu^-\gamma$ final states. A cut-off at high dilepton invariant masses, $q^2 = m_K^2$ and also a lower cut-off at $q^2 = 2m_l^2$ restricts the available phase space for muon decay mode to a region of invariant mass: $q^2 = [m_\mu^2, m_K^2]$. In the electron final state dilepton invariant masses in the range $q^2 = [m_e^2, m_K^2]$ together with a $1/q^2$ dependence gives the decay $K_L \rightarrow e^+e^-\gamma$ a much larger region of phase space to populate and hence much larger branching

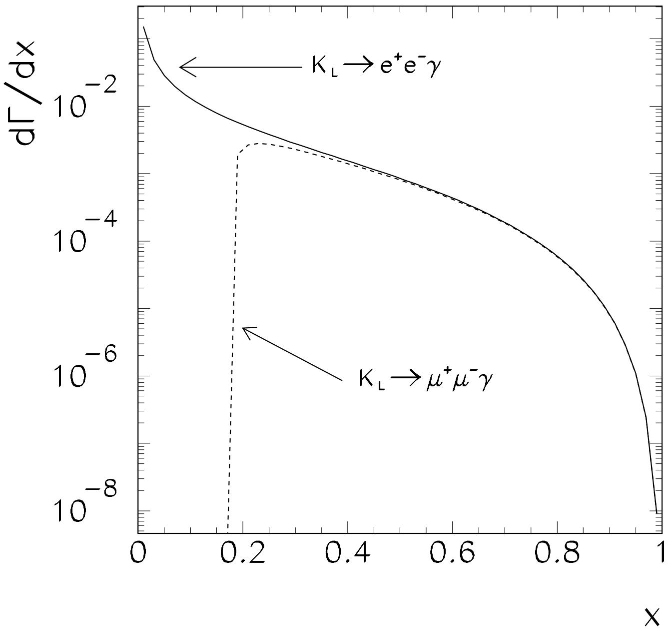

Figure 1.3: Differential decay rates for $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow e^+e^-\gamma$ ($x = (m_{l^+l^-}/m_K)^2$). $f(x) = 1$ is assumed here.

ratio than $K_L \rightarrow \mu^+\mu^-\gamma$ as shown in Fig. 1.3. The decay $K_L \rightarrow \gamma\gamma$ has the same constant of proportionality as in Eqn. 1.4 and this then cancels in the ratio of branching fractions. By assuming $|f(q^2)| = 1$ and integrating Eqn. 1.4 an estimate of the decay rate for $K_L \rightarrow l^+l^-\gamma$ can be obtained

$$\frac{\Gamma(K_L \rightarrow l^+l^-\gamma)}{\Gamma(K_L \rightarrow \gamma\gamma)} = \begin{cases} 4.09 \times 10^{-4} & l = \mu \\ 0.016 & l = e. \end{cases} \tag{1.5}$$

In the case of non-strange pseudoscalar mesons the decay $P \rightarrow l^+l^-\gamma$ contains

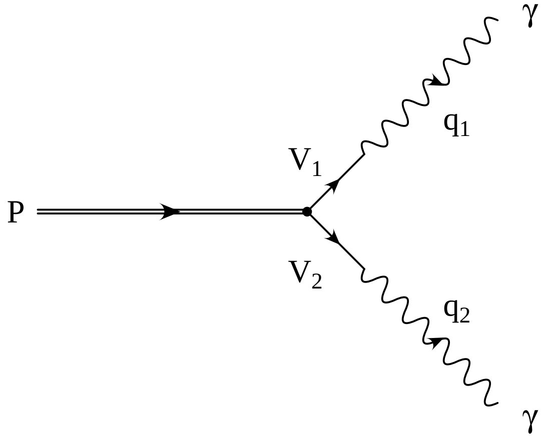

Figure 1.4: Vector Meson Dominance model for decay of non strange pseudoscalar meson, $P \to \gamma\gamma$. The meson propagators gives rise to pole terms in form factor in Eqn. 1.6.

no weak transition and a reasonably successful model exists to describe this case. The leading order contributions are electromagnetic. The two photon intermediate state is assumed to dominate, and transitions to that state occur through couplings of the pseudoscalar meson to vector mesons, V. Also $P\gamma\gamma$ and $PV\gamma$ couplings are assumed absent. With these assumptions, one has the so-called Vector Meson Dominance (VMD) model [19]. In this case the form factor for the decay of a pseudoscalar meson P, mass M, $P \to V_1V_2 \to \gamma\gamma$ is calculated from the Feynman diagram in Fig. 1.4 to be

$$f_P(q_1^2, q_2^2; M^2) \propto \frac{1}{(m_1^2 - q_1^2)(m_2^2 - q_2^2)}, \tag{1.6}$$

where m_1, m_2 are the masses of the vector mesons and q_1, q_2 are the 4-momenta of the two photons. For the analogous case in which P becomes a K meson we must take account of the decay of the strange quark and the only known mechanism for this is via the weak interaction. Bergström, Massó, and Singer have suggested that the Vector Meson Dominance model in this case must be expanded to include two contributions [1]. The first, similar to the non-strange case involves a weak transition to a intermediate meson that decays as in Fig. 1.5(a). The second contains a $KK^*\gamma$ vertex, as shown in Fig. 1.5(b). This is a new feature, as such a coupling was explicitly ignored in the analogous non-strange decay. In this model the long distance contribution to the $K_L \to \gamma\gamma^*$ transition is given by the sum of Figs. 1.5(a) and (b). The amplitude for the process described by Fig. 1.5(a) is calculated from Vector Meson Dominance to be

$$A_1 = \frac{|A_{\gamma\gamma}|}{1 - q^2/m_\rho^2}. \tag{1.7}$$

The amplitude for Fig. 1.5(b) is obtained by assuming a phenomenological Lagrangian that has accounted successfully for other kaon decay processes [1, 20, 21]:

$$\begin{aligned} A_K(q^2) &= \sqrt{2} e G_{NL} f_{K^* K \gamma} \left(\frac{m_\rho^2}{f_K f_\rho^2} \right) \\ &\times \frac{1}{1 - q^2/m_{K^*}^2} \left\{ \frac{4}{3} - \frac{1}{1 - q^2/m_\rho^2} - \frac{1}{9} \frac{1}{1 - q^2/m_\omega^2} - \frac{2}{9} \frac{1}{1 - q^2/m_\phi^2} \right\}, \end{aligned} \tag{1.8}$$

where G_{NL} is a constant that turns out to be quite close to the Fermi coupling

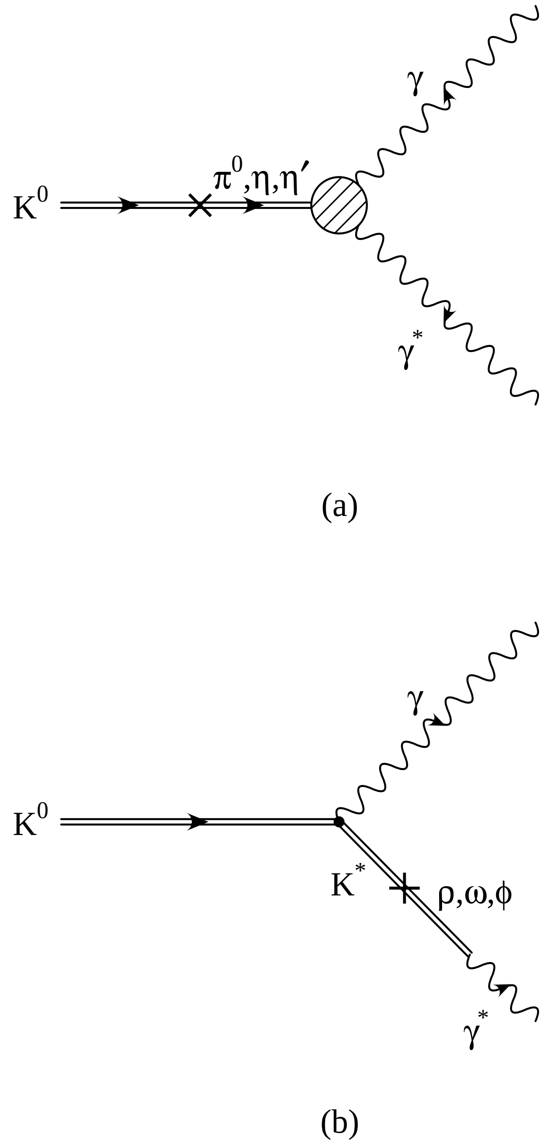

Figure 1.5: Long distance contributions to the $K_L \to \gamma\gamma^*$ transition. (a) pole term, and (b) the possible K^* contribution.

constant and f_j are known meson form factors. The total amplitude for the K_L Dalitz decay is then,

$$A_D(q^2) = A_1(q^2) + \alpha_{K^*} A_K(q^2). \tag{1.9}$$

A factor α_{K^*} is inserted here in anticipation that different models will predict different relative strengths between the diagrams in Figs. 1.5(a) and (b). We note that A_D is the form factor from Eqn. 1.3,

$$f(q^2) = \frac{A_D(q^2)}{|A_{\gamma\gamma}|} = f(q^2, 0; M_K^2), \tag{1.10}$$

and that $f(q^2)$ a limiting form of a more general form factor $f(q_1^2, q_2^2; M_K^2)$ that describes the $K_L \to \gamma^*\gamma^*$ transition where both photons are off-shell.

A strict application of the Vector Meson Dominance model for non-strange pseudoscalar mesons to the K_L Dalitz decay would ignore the K^* contribution and that would imply $\alpha_{K^*} = 0$. Sakurai developed a model using a phenomenological Hamiltonian that had pseudoscalar-pseudoscalar and vector-vector couplings with the same strength and transitions occurring with hadronic currents dominated by the lowest mass mesons [21]. In this model $|\alpha_{K^*}| = 1$. Another calculation by Bergström, Massó, and Singer attempted to use the SVZ non-leptonic weak hamiltonian which is of the form

$$H_W^{NL} = \frac{G_F}{\sqrt{2}} \sum_i c_i O_i \tag{1.11}$$

and where O_i are quark operators, and c_i are c-numbers [1, 15]. The result differs from that in Eqn. 1.8 by a factor of $1.2 \sin \theta_c$, and this corresponds to $|\alpha_{K^*}| = 0.2 - 0.3$.

A model due to Ko uses a different effective Lagrangian and the results can be summarized by the existence of additional decay channels with a set of vertices $K^0(\rho, \omega, \phi)\gamma$ which were assumed absent in VMD. In Ko's model there are no free parameters and the branching fractions for $K_L \rightarrow l^+ l^- \gamma$ are predicted unambiguously. Specifically there is a prediction that $B(K_L \rightarrow \mu^+\mu^-\gamma) \ \ /B(K_L \rightarrow \gamma\gamma) = (7.45^{+0.54}_{-0.15}) \times 10^{-4}$ [22].

Gvozdev *et al.* attempt to calculate the strength of the $K_L \rightarrow \gamma\gamma$ transition from a quark-level model and obtain a prediction for the ratio $B(K_L \rightarrow e^+e^-\gamma)/B(K_L \rightarrow \mu^+\mu^-\gamma)$ of 20 ± 0.5 [23].

1.4 Top Quark Mass and V_{td}

Figure 1.6 shows the processes that contribute to the decay $K_L \rightarrow \mu^+\mu^-$. The top figure is one of the short distance diagrams (there are two additional diagrams that involve a Z^0 boson) and the bottom figure shows the long distance contribution. The rate for the short distance processes has been calculated and the result is [4, 24, 25, 29, 30]

$$\begin{aligned} B(K_L \to \mu^+\mu^-)_{\rm SD} &= \frac{G_F^2}{2\pi^4}\frac{(1-4M_\mu^2/M_K^2)^{1/2}}{(1-m_\mu^2/M_K^2)^2}\frac{\tau(K_L)}{\tau(K^+)} \\ &\times \frac{\left(\mathrm{Re}\sum_{i=u,c,t} V_{is}V_{id}^* C_i(x_i)\right)^2}{|V_{us}|^2} B(K^+ \to \mu^+\nu_\mu) \end{aligned} \tag{1.12}$$

where $x_i = m_i^2/M_W^2$ and

$$\begin{aligned} C_t(x) &= Y_0(x) + a(\mu)Y_1(x;\mu) \\ C_c(x) &= C_{NL}(x;\mu) - B_{NL}(x;\mu) \\ C_u(x) &= 0. \end{aligned} \tag{1.13}$$

The term Y_0 is the leading order electroweak contribution,

$$Y_0(x) = \frac{x}{8}\left(\frac{4-x}{1-x} + \frac{3x}{(1-x)^2}\ln x\right),$$

and Y_1 is the QCD next-to-leading-log correction. Numerically $|Y_0| >> |aY_1|$, where $a(\mu) = \alpha_s/4\pi$; there is an explicit dependence on the renormalization scale μ. The functions C_{NL} and B_{NL} are rather complicated and can be found in Ref. [29]. Using the Wolfenstein parameterization of the CKM matrix [26] in which $V_{td} = A\lambda^3(1 -$

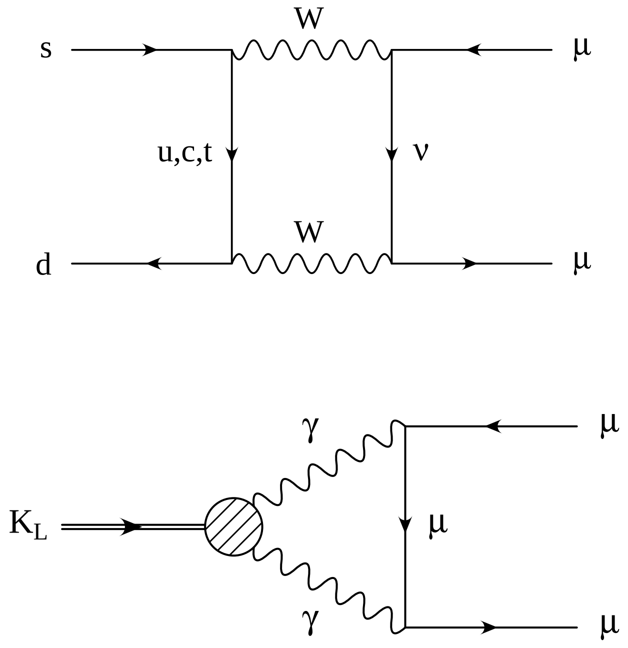

Figure 1.6: Lowest order electroweak process (top) and long distance contribution (bottom) to $K_L \to \mu^+\mu^-$.

$\rho - i\eta)$ and $V_{ts} = -A\lambda^2$ and taking $\lambda = 0.22$ we have

$$B(K_L \to \mu^+\mu^-)_{\mathrm{SD}} \;=\; 1.71 \times 10^{-9} A^4 |C_t(x_t)|^2 \left(1 - \rho + \frac{417 C_c(x_c)}{A^2 C_t(x_t)}\right)^2 . \qquad (1.14)$$

The third term in the parentheses in this equation is the result of th charm quark

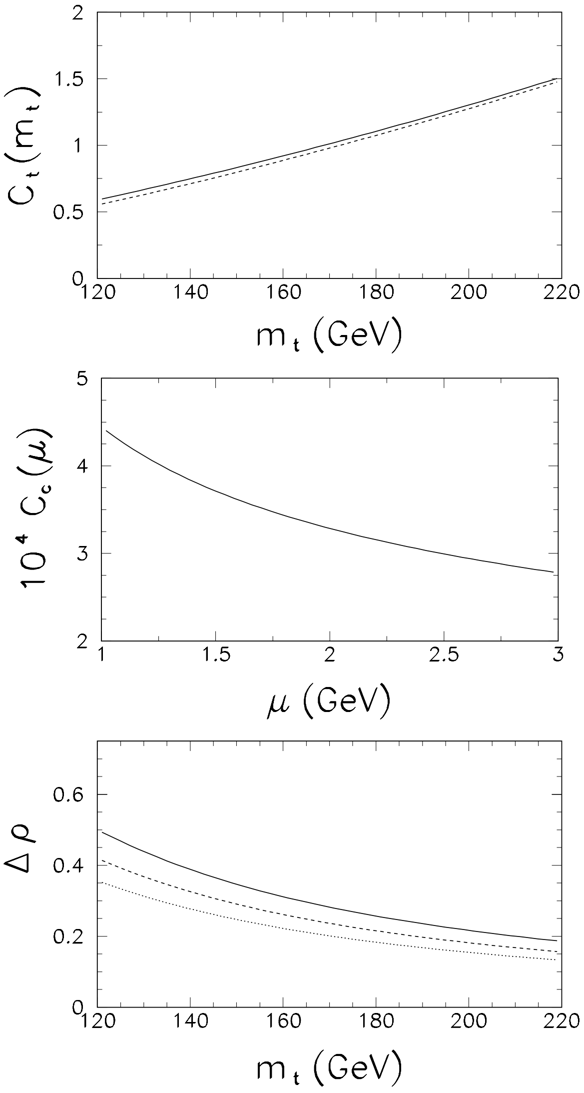

Figure 1.7: Functions used to calculate the short distance contribution to $K_L \to \mu^+\mu^-$. Top: The solid line is Y_0 and the dashed line is $Y_0 + aY_1$, showing that the QCD corrections to $K_L \to \mu^+\mu^-$ from the process involving the top quark are small. Middle: the function C_c from the charm quark contribution. Bottom: The function $\Delta\rho$, this is the charm quark contribution that modifies the expression for ρ from $K_L \to \mu^+\mu^-$. The three curves are for $A = 0.76$ (solid) 0.83 (dashed) 0.90 (dotted).

contribution. We find $C_c(x_c) = 3.83 \times 10^{-4}$ when $m_c = \mu = 1.4\ GeV$. If we define

$$\Delta\rho \equiv \frac{417 C_c(x_c)}{A^2 C_t(x_t)} = \frac{0.159}{A^2 C_t(x_t)}$$

then for $m_t = 170\ GeV$ and $A = 0.83$, $\Delta\rho = 0.23$. Thus the charm quark contribution to $K_L \rightarrow \mu^+\mu^-$ is significant. On the other hand the QCD correction to the top quark process is very small. In Fig. 1.7 we show the functions Y_0, C_t, C_c, and $\Delta\rho$.

A determination of $B(K_L \rightarrow \mu^+\mu^-)_{\mathrm{SD}}$ then constrains the CKM variable ρ in the Wolfenstein parameterization as a function of m_t

$$\rho = 1 + \Delta\rho - \sqrt{\frac{B(K_L \rightarrow \mu^+\mu^-)_{\mathrm{SD}}}{1.71 \times 10^{-9}} \frac{1}{A^2 C_t(x_t)}}. \tag{1.15}$$

To determine the contribution from the lower diagram in Fig. 1.6 a model to calculate the rate due to the two photon intermediate state is needed. The contribution comes from all photon states including off-shell photons. However a lower bound may be established by just taking the on-shell photons ($q^2 = 0$) for which the rate $K_L \rightarrow \gamma\gamma$ has been experimentally determined. This part of the two photon intermediate state is independent of the model used and is given by [11]

$$\frac{\mathrm{Rate}(K_L \rightarrow \mu^+\mu^-)_{2\gamma}}{\mathrm{Rate}(K_L \rightarrow \gamma\gamma)} = \frac{1}{2}\alpha^2 \left(\frac{m_\mu}{M_K}\right)^2 \frac{1}{\beta} \left(\ln \frac{1+\beta}{1-\beta}\right)^2, \tag{1.16}$$

where $\beta^2 = 1 - 4m_\mu^2/M_K^2$. This is called the unitarity bound and numerically it is given by

$$B(K_L \to \mu^+\mu^-)_{\mathrm{UNIT}} \equiv |\mathrm{Im}A_{2\gamma}|^2 = (6.83 \pm 0.28) \times 10^{-9},$$

where the uncertainty comes from the branching fraction for $K_L \to \gamma\gamma$ [27]. This is to be compared to the experimentally determined values

$$B(K_L \to \mu^+\mu^-)_{\mathrm{EXP}} = \begin{cases} (6.86 \pm 0.37) \times 10^{-9} & \text{BNL E791 [31]} \\ (7.9 \pm 0.7) \times 10^{-9} & \text{KEK E137 [32].} \end{cases} \qquad (1.17)$$

A general expression for the branching ratio $K_L \to \mu^+\mu^-$ can be written without assuming the two photons are real [28]

$$\frac{B(K_L \to \mu^+\mu^-)}{B(K_L \to \gamma\gamma)} = \frac{1}{2}\alpha^2 \left(\frac{m_\mu}{M_K}\right)^2 \frac{1}{\beta}|R(M_k^2)|^2 \qquad (1.18)$$

where

$$R(q^2) = \frac{i4\beta}{\pi^3 M_K^2} \int \frac{d^4k[q^2k^2 - (q \cdot k)^2] f(q^2, (q-k)^2)}{(k^2)(q-k)^2[(p-k)^2 - m_\mu^2]} \qquad (1.19)$$

and q and p are the 4-momenta of the kaon and the muon. The imaginary part of

$R(q^2)$ may be written as

$$\begin{aligned} \mathrm{Im}R(q^2) &= \mathrm{Im}R^{\mathrm{PL}}[1 - g(q^2)] \\ &= \left(-\ln\frac{1+\beta}{1-\beta}\right)[1 - g(q^2)] \end{aligned}$$

and the first term is just the model-independent (point-like) contribution from Eqn. 1.16. It is known that the remaining contribution to the imaginary part of the amplitude is small compared to this one, $|g| << 1$ (this is equivalent to ignoring contributions from other intermediate states such as $\pi\pi$ and $\pi\gamma$ [33]). The imaginary part of the amplitude is often called the *absorptive* part while the real part is referred to as *dispersive*. Bergström, Massó, and Singer have used Eqn. 1.9 to estimate ReR. They attempt to estimate the the form factor for $K_L \to \gamma^*\gamma^*$ by extrapolating the form factor for $K_L \to \gamma\gamma^*$ in two extreme cases. They determine [2]

$$-1.8 - 4.2\alpha_{K^*} < \mathrm{Re}R_{\mathrm{LD}} < -1.0 - 3.9\alpha_{K^*}$$

which corresponds to

$$(-5.0 - 11.5\alpha_{K^*}) \times 10^{-5} < \mathrm{Re}A_{\mathrm{LD}} < (-2.8 - 10.6\alpha_{K^*}) \times 10^{-5}. \qquad (1.20)$$

We may thus write

$$\begin{aligned} B(K_L \to \mu^+\mu^-) &= |\mathrm{Re}A|^2 + |\mathrm{Im}A|^2 \approx |\mathrm{Re}A|^2 + |\mathrm{Im}A_{2\gamma}|^2 \\ &= |\mathrm{Re}A_{\mathrm{LD}} + A_{\mathrm{SD}}|^2 + B_{\mathrm{UNIT}} \qquad (1.21) \end{aligned}$$

and hence

$$\Delta B \equiv B_{\mathrm{EXP}} - B_{\mathrm{UNIT}} = |\mathrm{Re}A_{\mathrm{LD}} + A_{\mathrm{SD}}|^2.$$

A measurement of $B(K_L \to \mu^+\mu^-)$ effectively constrains the contributions that can come the short distance processes, as we shall now describe. From the average value for $B(K_L \to \mu^+\mu^-)$ in Eqn. 1.17 a 90% confidence limit of $\Delta B < 6.5 \times 10^{-10}$ is obtained. In comparison, using just the Brookhaven result alone would yield $\Delta B < 5.6 \times 10^{-10}$ [31]. The lower limit of $\mathrm{Re}A_{\mathrm{LD}}$ from Eqn. 1.20 together with the maximum relative phase of the long and short distance contributions then sets the upper limit: $A_{\mathrm{SD}} < \sqrt{\Delta B} + |\mathrm{Re}A_{\mathrm{LD}}|$. As Fig. 1.8 shows, the sensitivity of A_{SD} to changes in ΔB is small in comparison to α_{K^*} which has an uncertainty of about ± 0.08. This bound on A_{SD} then implies a limit for ρ as given by Eqn. 1.15 and shown in Fig. 1.9. We note that lower values of α_{K^*} give more favorable limits on ρ. By assuming $\alpha_{K^*} = -0.25 \pm 0.07$ Ref. [31] inferred the bound $\rho > -0.6$ (see Fig. 1.10). Measurements from the B-Meson system currently constrain $|\rho| \lesssim 0.5$ [34].

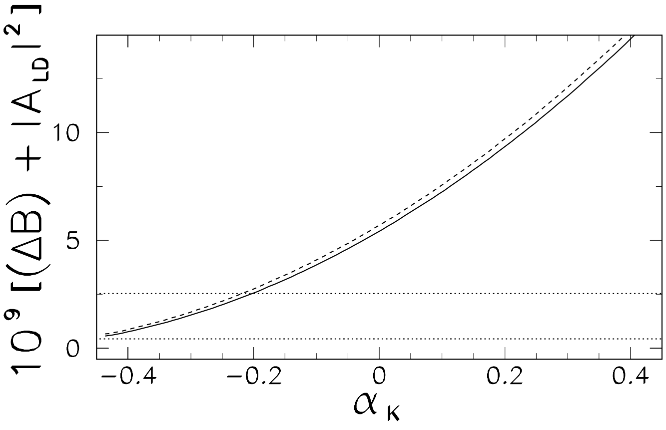

Figure 1.8: Upper limit on $\mathrm{Re}A^2_{\mathrm{SD}}$ from $K_L \to \mu^+\mu^-$. The solid line is the upper bound obtained when $|\mathrm{Re}A_{\mathrm{LD}}|^2$ is added to $\Delta B = 5.6 \times 10^{-10}$, the dashed line assumes $\Delta B = 6.5 \times 10^{-10}$. The dotted lines enclose a region that is expected from the standard model with $A = 0.83$, $m_t = 175$ GeV and $|\rho| < 0.5$

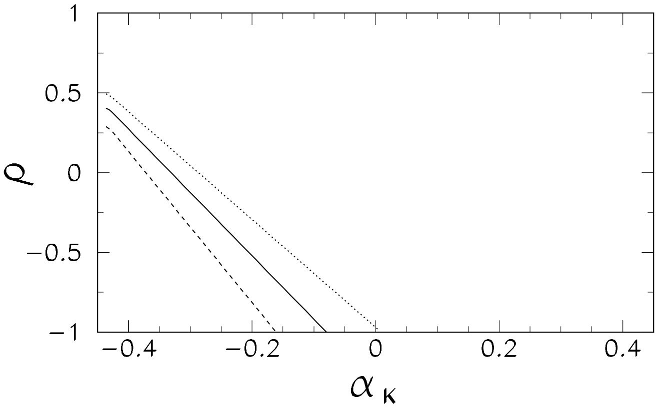

Figure 1.9: Bound on ρ from $K_L \to \mu^+\mu^-$ as a function of α_{K^*}. The dependence on α_{K^*} is due to the subtraction of the long distance contribution. The different contours correspond to $A = 0.76$ (dashed), $A = 0.83$ (solid), $A = 0.90$ (dotted). $\Delta B < 5.6 \times 10^{-10}$ and $m_t = 175$ GeV are assumed here.

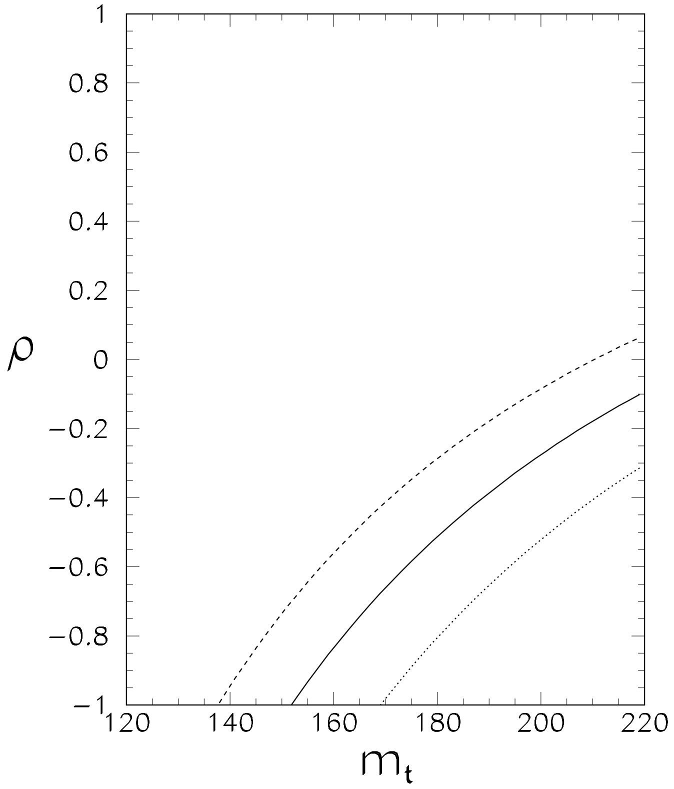

Figure 1.10: ρ versus m_t from $K_L \to \mu^+\mu^-$. The three contours correspond to $A = 0.76$ (dashed), $A = 0.83$ (solid), $A = 0.83$ (dotted). The constraint $A_{\rm SD} < 3.0 \times 10^{-9}$ leads to an allowed region above the lines. (This choice of $A_{\rm SD}$ comes from assuming $\alpha_{K^*} < -0.18$ in the model of Bergström, Massó, and Singer– see Fig. 1.8.)

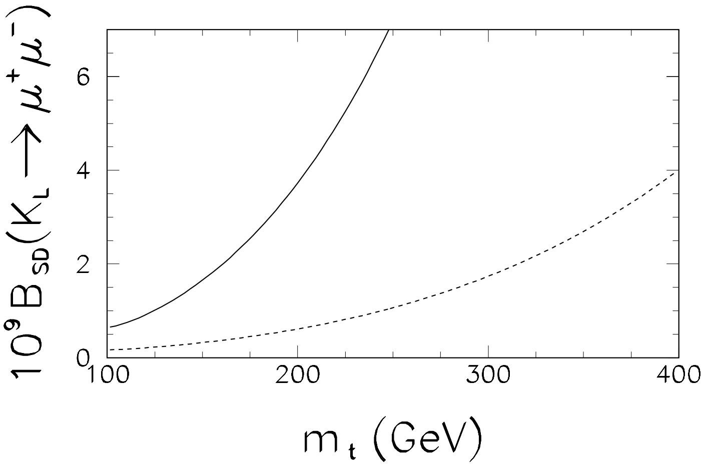

Figure 1.11: Variation of $B(K_L \rightarrow \mu^+\mu^-)$ as a function of m_t. The solid curve corresponds to $\rho = 0.5$ and the dashed curve corresponds to $\rho = -0.5$ ($A = 0.83$ is assumed here). A limit of $B_{\mathrm{SD}} < 3.0 \times 10^{-9}$ would imply $m_t < 360$ *GeV*.

Alternatively one can take the known range of $|\rho| \lesssim 0.5$ and use Eqn. 1.14 to infer a constraint on m_t [1]. Using $B_{\mathrm{SD}} < 3.0 \times 10^{-9}$ which corresponds to the 90% CL from the current world average and $\alpha_{K^*} = -0.25 \pm 0.07$ [31] the limit on the top mass is $m_t < 360$ *GeV*. (see Fig. 1.11). No lower limit on ρ can be derived as the current experimental data only imply $B_{\mathrm{SD}} > 0$. This large range of m_t can be understood because Ref. [31] quote $\mathrm{Re}(A_{\mathrm{LD}} + A_{\mathrm{SD}})^2 = (-1.0 \pm 3.7) \times 10^{-10}$, the error being dominated by the uncertainty on $B(K_L \rightarrow \mu^+\mu^-)$. We also note that as the uncertainty on ρ decreases the upper limit on m_t from $K_L \rightarrow \mu^+\mu^-$ is likely to improve.

1.5 Previous Measurements of $K_L \to \gamma\gamma^*$

In Fig. 1.12 we show the structure function $|f(x)|^2$ as a function of the invariant dilepton mass $x = q^2/m_K^2$ for various values of α_{K^*}. In the electron mode, one would typically obtain high statistics in the low-x region which would constrain a fit and then differences in α_{K^*} would be seen in the high-x region. The measurement of α_{K^*} in $K_L \to e^+e^-\gamma$ has been observed at Brookhaven [35] and CERN [36]. In Fig. 1.13 are the results from Brookhaven [35] showing a form factor consistent with a value $\alpha_{K^*} = -0.28 \pm 0.083^{+0.054}_{-0.034}$. The electron mode is complicated by radiative corrections, the decay $K_L \to e^+e^-\gamma\gamma$ where one of the γ's is very soft, being indistinguishable from the decay of interest. The Brookhaven result has a 55% correction from this effect, changing the values of α_{K^*} from -0.18 to -0.28 [35].

In Fig. 1.14 we show the complete differential decay spectrum as a function of x (Eqn. 1.4) for the $\mu^+\mu^-\gamma$ case, again for various values of α_{K^*}. Here we see the lack of data at low-x means that to a good approximation, a fit to this data is shape-insensitive, the only difference between the various values of α_{K^*} being a normalization factor. Equation 1.4 with the form factor from Eqn. 1.10 can be integrated and is a function of α_{K^*} as shown in Fig. 1.15 [1]. We note that the branching ratio for $e^+e^-\gamma$ is very insensitive to the value of α_{K^*} in contrast to the $\mu^+\mu^-\gamma$ mode.

The only previous search for the decay $K_L \to \mu^+\mu^-\gamma$ was from Brookhaven [37];

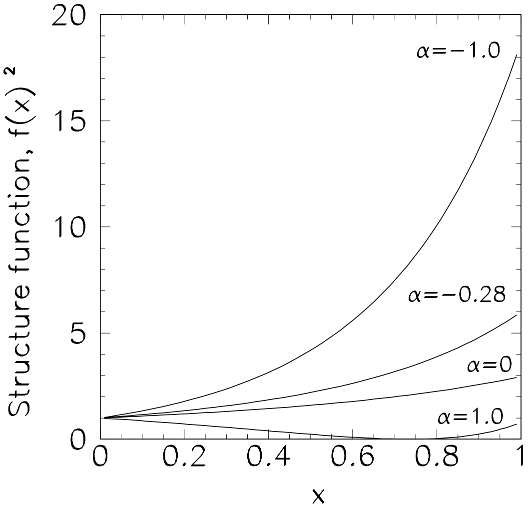

Figure 1.12: Structure function $|f(x)|^2$ for $\alpha_{K^*} = -1, -0.28, 0, +1$.

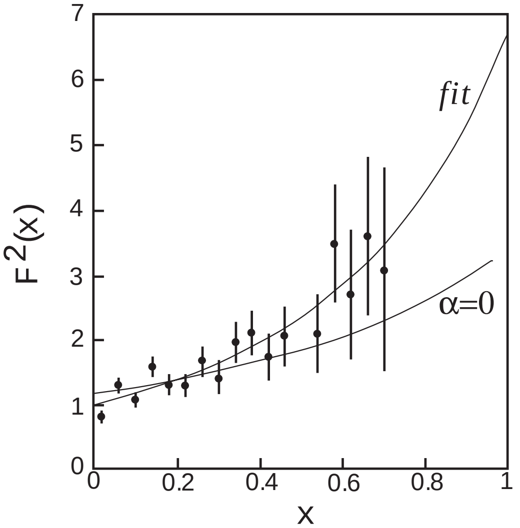

Figure 1.13: Brookhaven $K_L \to e^+e^-\gamma$ measurement of α_{K^*} The lower curve is for the case $\alpha_{K^*} = 0$ and is inconsistent with the data [35].

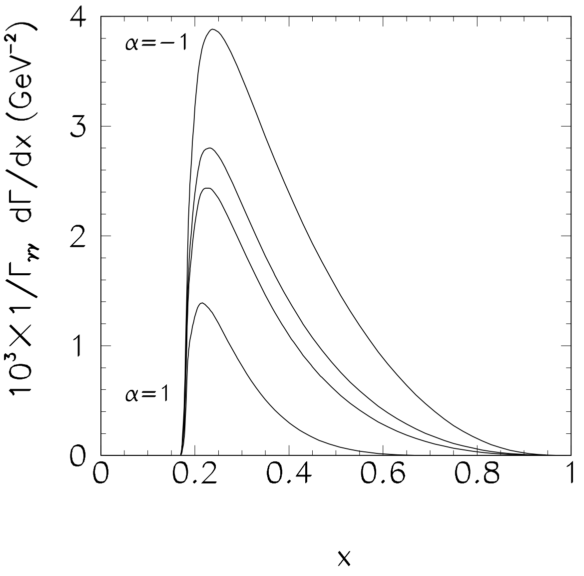

Figure 1.14: Dimuon mass distribution in the decay $K_L \rightarrow \mu^+\mu^-\gamma$ as a function of $x = m_{\mu\mu}^2/M_K^2$ for $\alpha_{K^*} = -1, -0.28, 0, +1$.

one candidate event was observed above 0.48 GeV/c^2 where 0.1 background events were expected (Fig. 1.16).

1.6 Other Physics Results from Experiment E799

Experiment E799 was spawned from a previous experiment, E731 which was designed to measure the CP violation parameter ε'. This was accomplished by simultaneously observing the decays $K_L \rightarrow \pi^0\pi^0$, $K_L \rightarrow \pi^+\pi^-$, $K_S \rightarrow \pi^0\pi^0$, and $K_S \rightarrow \pi^+\pi^-$ [38, 39]. The final result of this measurement was that $\varepsilon'/\varepsilon = (0.74 \pm 0.61) \times 10^{-3}$, consistent with zero. A similar measurement from the ex-

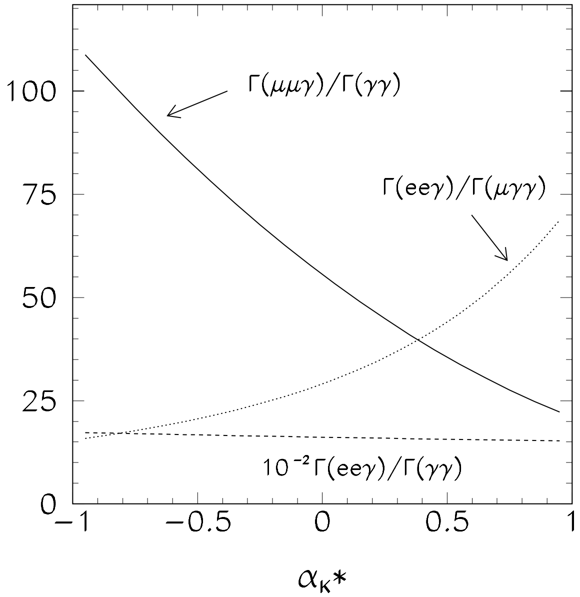

Figure 1.15: Branching ratios for $K_L \to \mu^+\mu^-\gamma$ and $K_L \to e^+e^-\gamma$ normalized to $K_L \to \gamma\gamma$ and their ratio. The parameter α_{K^*} measures the relative contribution of the K^* diagram in Fig. 1.5(b).

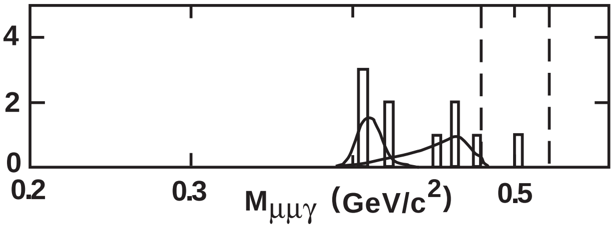

Figure 1.16: Evidence for $K_L \to \mu^+\mu^-\gamma$ from Brookhaven. One event observed where 0.1 background were expected [37].

periment NA31 at CERN obtained the result $\varepsilon'/\varepsilon = (2.3 \pm 0.65) \times 10^{-3}$ which is about 3.5 standard deviations from zero [40].

E799 was originally proposed to search for the decays $K_L \to \pi^0 \mu^+ \mu^-$ and $K_L \to \pi^0 e^+ e^-$, both of which potentially contain large CP violating amplitudes. The decay $K_L \to \pi^0 \nu\bar{\nu}$ was also a major motivation for E799. This decay mode is thought to be dominated by direct CP violating amplitudes: the electroweak penguin and the W box diagram. The Standard Model prediction for this decay mode is $0.4 - 11 \times 10^{-11}$ where the range is due to the current uncertainty in the standard model parameters η, and m_t. This is many orders of magnitude away from the best current experimental sensitivity, which comes from E799: the 90% confidence limit obtained was $B(K_L \to \pi^0 \nu\bar{\nu}) < 5.8 \times 10^{-5}$.

Table 1.1 lists the previously published results from experiment E799. The table shows that many rare decay modes of the kaon are accessible, in addition to rare pion decays and lambda hyperon physics. In addition, as many of the decays are *so* rare there also exists the possibility that before they are observed at the predicted Standard Model levels other exotic processes may begin to contribute.

Because of its genealogy the E799 detector naturally had a large acceptance for multi-body final states and it was therefore an ideal environment in which to search for many rare decays, including $K_L \to \mu^+ \mu^- \gamma$ which we will concentrate upon henceforth.

Mode	Result
$\Lambda^0 \rightarrow p^+\pi^-$	Polarization of $\Lambda, \bar{\Lambda}$ measured no energy dependence, in contrast to Σ^+, Ξ^- decays [41].
$K_L \rightarrow e^+e^-\gamma\gamma$	$B(E^*_\gamma > 5\ MeV) = (6.5 \pm 1.2\text{stat} \pm 0.6\text{sys}) \times 10^{-7}$, 58 events, 8.3 background events. This decay is a background to $K_L \rightarrow e^+e^-\gamma$ and $K_L \rightarrow \pi^0e^+e^-$. Predicted to be 5.8×10^{-7} [42].
$K_L \rightarrow \pi^0\pi^0\gamma$	90% CL = $B < 2.3 \times 10^{-4}$. Theory predicts a range of 1×10^{-8} - 7×10^{-11} [43].
$K_L \rightarrow e^+e^-e^+e^-$	$(3.96 \pm 0.78(\text{stat}) \pm 0.32(\text{syst})) \times 10^{-8}$, 27 events, 0.36 background. A form-factorless prediction via $K_L \rightarrow \gamma\gamma$ and with no radiative corrections is $(3.55 \pm 0.17) \times 10^{-8}$ [44].
$K_L \rightarrow \pi^0\mu^+\mu^-$	90% CL = $B < 5.1 \times 10^{-9}$. Standard Model predictions in range $3 - 6 \times 10^{-12}$ containing both direct and indirect CP violating amplitudes in addition to CP conserving amplitudes [45].
$K_L \rightarrow \pi^0e^+e^-$	90% CL = $B < 4.3 \times 10^{-9}$. Standard Model predictions in range $10^{-11} - 10^{-12}$ with dominant contribution being a CP violating amplitude [46].
$\pi^0 \rightarrow e^+e^-$ from $K_L \rightarrow \pi^0\pi^0\pi^0$	$B((m_{ee}/m_{\pi^0})^2 > 0.95) = (7.6^{+3.9}_{-2.8}(\text{stat}) \pm 0.5(\text{syst})) \times 10^{-8}$, signal of 9 events, 1 background event. Theory is model dependent in range $6 - 7 \times 10^{-8}$ [47].
$K_L \rightarrow \pi^0\nu\bar{\nu}$	90% CL = $B < 5.8 \times 10^{-5}$. Predicted in range $0.4 - 11 \times 10^{-11}$ which comes from uncertainty in standard model parameters η, m_t. This mode is dominated by a directly CP violating process [48].
$\pi^0 \rightarrow \mu^\pm e^\mp$	90% CL = $1/2[B(\pi^0 \rightarrow \mu^+e^-) + B(\pi^0 \rightarrow \mu^-e^+)] < 8.6 \times 10^{-9}$. Lepton flavor violation search [49].

Table 1.1: Previously published results from Experiment E799.

Chapter 2

The Detector

This chapter discusses the production, flight and decay of neutral kaons, together with the detectors used to observe the decay products.

2.1 Kaon Production

Kaons for this experiment were produced from the interactions of 800 GeV protons on a beryllium target. The protons arrived with a time structure characteristic of the Fermilab Tevatron, grouped in "buckets" of 2 ns duration, separated by 18.9 ns. These buckets last for 22 s out of each 58 s Tevatron cycle, the 22 s intervals are colloquially referred to as "spills". Each spill contained an average of 1.4×10^{12} protons. The beryllium target was a square rod with a cross-section of 3.2 $mm \times$ 3.2 mm and was one interaction length along the direction of the beam. The size of the beam-spot at the target was approximately gaussian with width 0.4 mm. A small wire chamber (SWIC) located 3 m upstream of the target was used to locate the position of the primary beam, and magnets 30 m upstream of the target were used to alter the position of the beam on the target.

The ε'/ε experiment for which the apparatus was originally designed needed two beams of neutral kaons and this requirement influenced all of the beam shaping elements downstream of the target. Figure 2.1 shows the series of beam shaping devices that followed the target area. Two considerations help to reduce spurious activity in downstream detectors. The first is to increase the neutral kaon content of the beam, and the second is to give the beams well defined edges. Charged particles are removed from the beam by "sweeping" magnets placed immediately after the target and also after other collimators in the beam line, to remove charged debris from interactions (see Fig. 2.2). Initially, the two beams were produced by a large 5.8 m thick copper collimator nine meters downstream of the target. The inner edges were defined by two horizontal "slab" collimators. The outer edges of the beams were defined by eight "jaw" collimators: 2 in the x-view and 2 in the y-view at 20 m and again at 30 m from the target. The number of photons in the beam was reduced by a lead absorber (7.6 cm, 14 radiation lengths, but only 0.44 interaction lengths) . This converted photons to electron-positron pairs that could be removed by the sweeping magnets.

The spectrum of neutral kaons emerging from protons incident on the target has been previously determined in experiment E731, the result has been parameterized by [50]

$$\frac{d^2N(K^0)}{dPd\theta} = \frac{\sin\theta}{800}\left[B_1\frac{X(1-X)^{A_1}(1+5e^{-D_1X})}{(1+P_t^2/M_1^2)^4} + B_2\frac{X(1-X)^{A_2}(1+5e^{-D_2X})}{(1+P_t^2/M_2^2)^4}\right]F(P)$$

and

$$\frac{d^2N(\bar{K^0})}{dPd\theta} = \frac{\sin\theta}{400}\left[B_2\frac{X(1-X)^{A_2}(1+5e^{-D_2X})}{(1+P_t^2/M_2^2)^4}\right]F(P)$$

with

$$A_1 = 2.924, \qquad A_2 = 6.107,$$

$$B_1 = 14.15, \qquad B_2 = 12.33,$$

$$D_1 = 19.89, \qquad D_2 = 12.78,$$

$$M_1 = 1.079 GeV/c^2, \qquad M_2 = 1.048 GeV/c^2,$$

where X is the ratio of the kaon to proton energies, P is the kaon momentum, P_t is kaon transverse momentum, θ is the production angle, and

$$F(P) = 1 + w_1P + w_2P^2 + W_3P^3 + W_4P^4,$$

$$w_1 = 6.033 \times 10^{-3}(GeV/c)^{-1}, \qquad w_2 = -4.283 \times 10^{-6}(GeV/c)^{-2},$$

$$w_3 = -1.106 \times 10^{-7} (GeV/c)^{-3}, \qquad w_4 = 1.802 \times 10^{-10} (GeV/c)^{-4}.$$

The energy spectrum of K_L's with $\theta = 4.8$ mrad is shown in Figure 2.3. The targeting angle was chosen as a compromise between maximizing the neutral kaon yield, which occurs at $\theta = 0$ and decreasing the neutron to kaon ratio of by using a larger targeting angle.

Protons incident on a target produce many different particles and neutral kaons represent only a small fraction of the incident mass-energy: Photons, neutrons, Λ^0's, Ξ^0's were present in the beam at some level. The exact content of the beams is difficult to predict and after many months of generating Monte Carlo the neutron the kaon ratio of the beams in the decay region is known to be 2:1 with about 50% error on the ratio [51]. The Λ^0's and Ξ^0's mostly decayed before reaching the effective decay volume 90 m from the target. Neutrons almost never decayed in the detector and in fact most of the K_L content also did not decay: at 50 GeV, only 5% of the K_L's decay before reaching the end of the detector.

2.2 Detector Elements

The effective decay region for the neutral kaons began at about 90 m downstream from the target. We now describe the elements of the detector starting at approximately this z-position. Figure 2.4 shows a highly compressed view of these detectors.

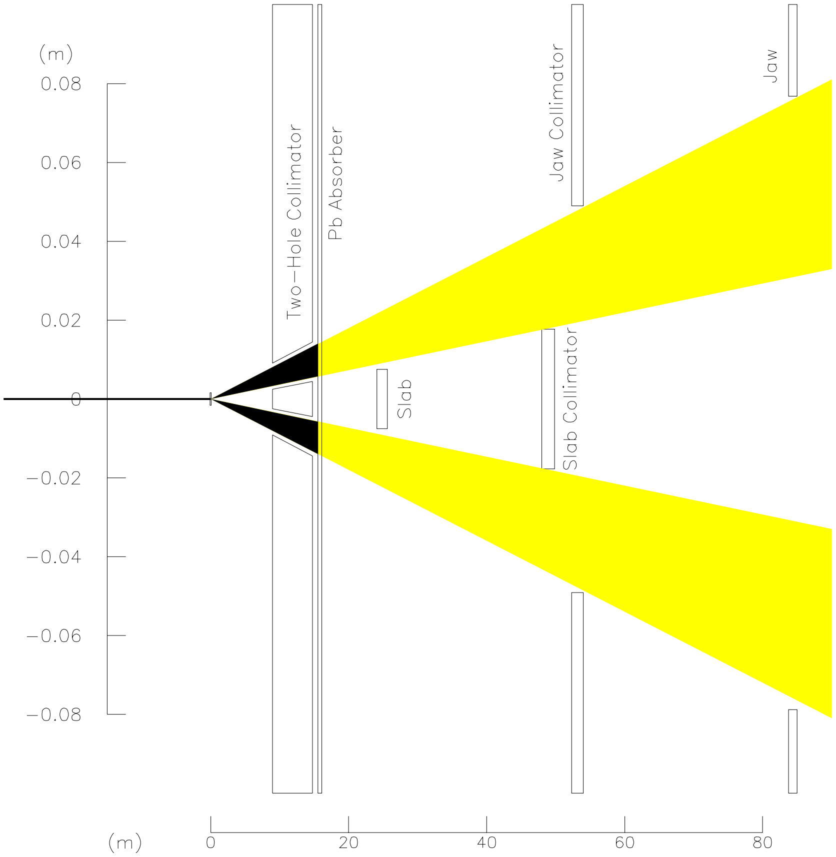

Figure 2.1: Shaping of the beams, y-view.

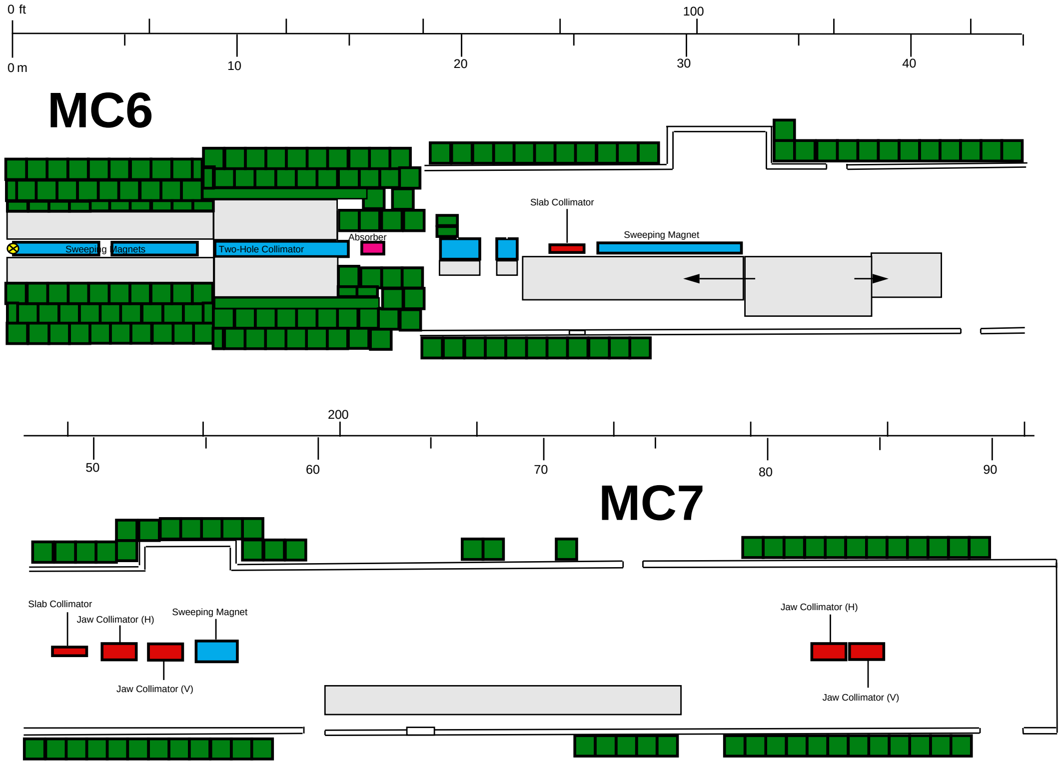

Figure 2.2: Beam line layout in region between the target $z = 0$ m) and the beginning of the effective decay region $z = 90$ m).

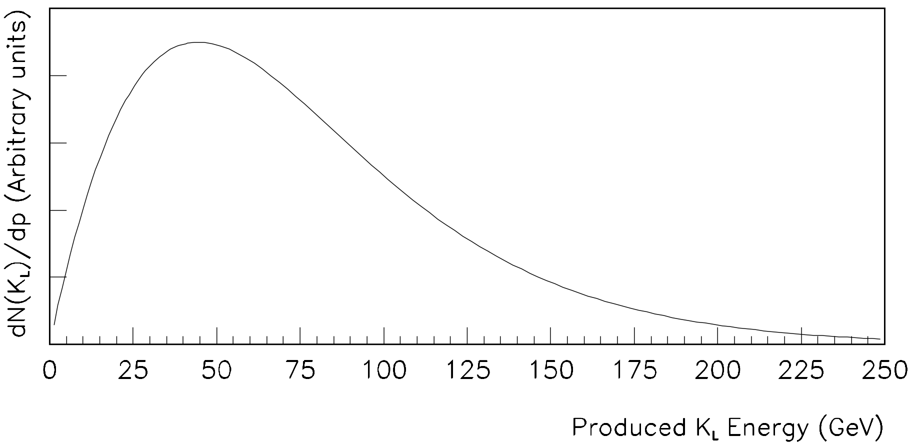

Figure 2.3: The energy spectrum of K_L for $\theta = 4.8$ mrad.

2.2.1 Magnetic Spectrometer

Measurement of the momentum and trajectories of charged particles was performed by a bending magnet and four drift chambers. Two drift chambers measured the directions of tracks upstream of the magnet and two measured the directions of tracks downstream of the magnet. The magnet was a dipole with an aperture of 2.52 m (x) × 1.46 m (y) and produced a field of 4 KGauss in the vertical direction. This imparted a transverse momentum (P_t) of about 200 MeV/c in the horizontal direction.

Figure 2.5 shows the drift chamber wire configuration. To measure x and y

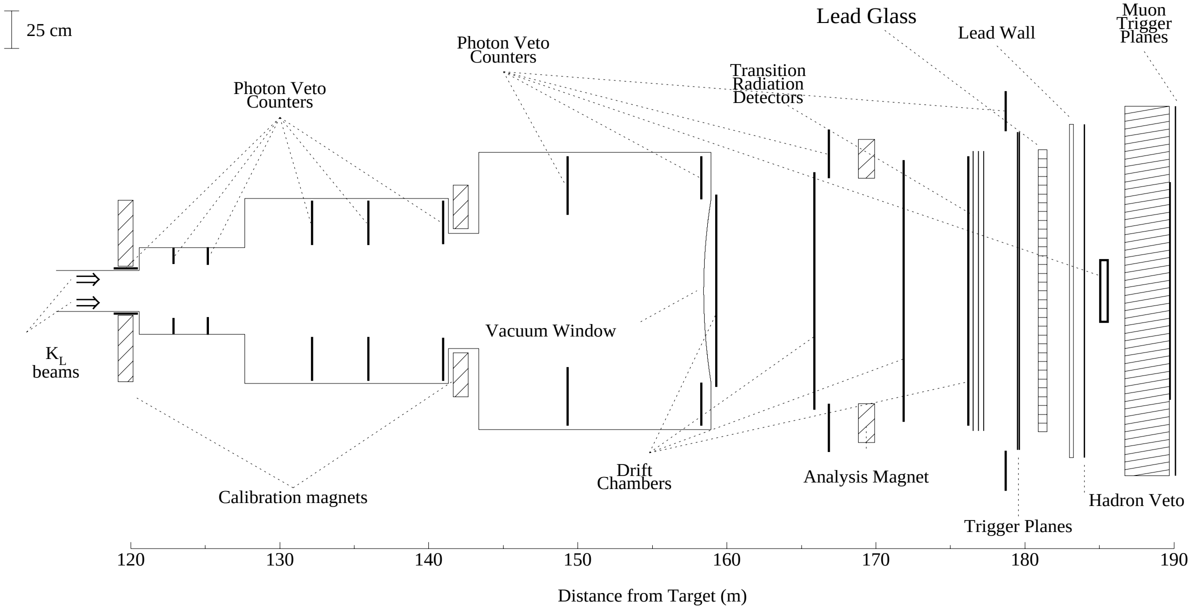

Figure 2.4: Side elevation view of selected detector elements for $z > 90$ m.

positions there were two separate planes of wires in each chamber. In each plane of wires there were two rows of sense wires surrounded by two rows of hexagonally shaped field shaping wires. These wires were contained in volume between two Aclar windows, and the entire chamber was filled with 50% Argon, 50% ethane, 1% ethanol mixture. During the spill high voltage (-2650 Volts) was applied to the field shaping wires. At this voltage, drift velocity was 50 $\mu m/ns$. The signals from the sense wires were discriminated and then split and sent to time-to-digital (TDC) converters and also to hardware track-processor modules. The TDC's were Lecroy 4291B Camac modules operating in a common stop mode, the stop being provided by the result of the level 1 trigger decision. The double pulse resolution of these TDC's was about 200 ns, meaning that for 200 ns after each pulse was received, the input to the TDC could not accept another "start". This had the advantage of excluding reflections from re-starting the TDC, but the disadvantage that dead-time was introduced.

The distribution of times recorded from a plane for all wires is illustrated in Fig. 2.6. We note the characteristic sharp edge near 240 ns resulting from the common stop mode of the TDC; values near this edge are from tracks that pass quite close to the wire. Assuming uniform illumination across cells the TDC values were converted to distances $d = x_i(t)$ for the i^{th} wire. A "sum of distances" parameter was defined from the times of two hits by by SOD $= x_i(t_1) + x_j(t_2)$. Apart from subtleties of missing hits, the requirement that the sum of the distances equal the

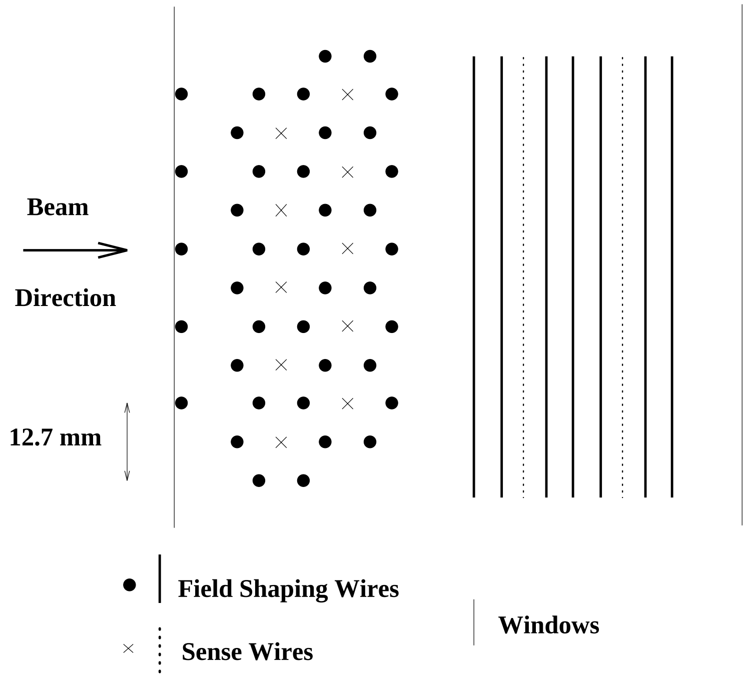

Figure 2.5: Drift chamber wires were arranged in two views as seen in the plan view. The field shaping wires were arranged in a staggered two-row hexagonal pattern, with sense wires at the centers of the hexagons. On the left is the cross section of the vertical wires which measured the x-position, and on the right are shown the horizontal wires which measured y-positions.

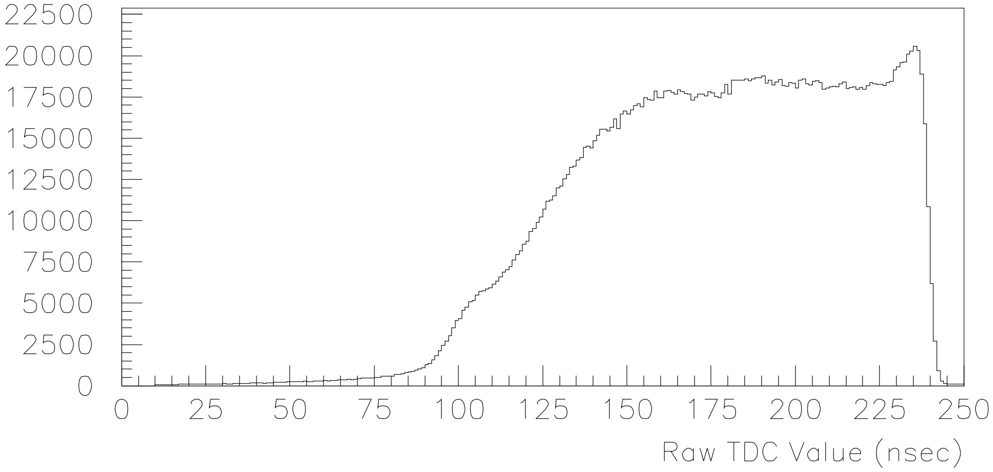

Figure 2.6: Raw distribution of TDC values for a typical plane.

cell size determined the exact location of the track. The SOD parameter is plotted for one view of a single chamber in Fig. 2.7, and indicates that the resolution for determining the position of a single track, in one view and in one chamber is typically about 100 μm.

2.2.2 Lead Glass Calorimeter

At the heart of the experiment was a lead-glass Cerenkov calorimeter. This consisted of 804 lead-glass blocks stacked in a roughly circular array, with two holes near the center through which the beams passed. Each of the calorimeter blocks was 60.2 cm long in the direction of the beam corresponding to 18.8 radiation lengths, and 5.8 cm $\times$ 5.8 cm in the dimension transverse to the beam. Individual blocks were wrapped in aluminized mylar to make them optically isolated from neighboring blocks. The

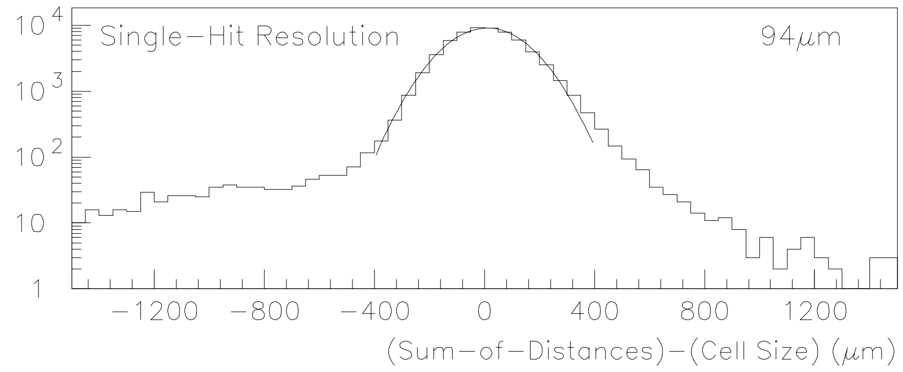

Figure 2.7: Sum of distances distribution for y view of chamber 4.

calorimeter is depicted in Fig. 2.8. The entire array was housed in a light tight, temperature controlled room. Charged particles traversing the lead-glass produced Cerenkov light that was collected by phototubes at the most downstream end of the calorimeter.

Signals were split and used for a variety of purposes. First a part of each signal was sent to a LeCroy 1885 dual-ranging analog-to-digital converter (ADC). The gate for this ADC was 100 *ns*, which was long in comparison to the time between consecutive buckets incident on the target, and therefore allowed for the possibility that energy deposits other than those associated with the in-time bucket could be included in the charge integration. For this reason, part the signals from groups of blocks were summed and sent to ADC modules that had 30 *ns* charge integration gates. This information proved to be extremely useful in the analysis for removing events that contained out-of-time energy deposits in the calorimeter. In addition

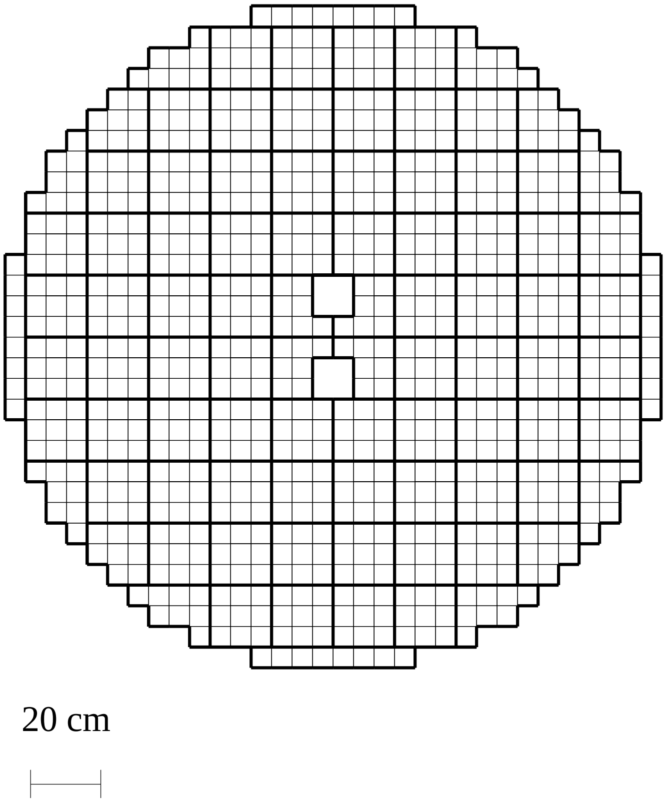

Figure 2.8: Transverse section of the lead-glass calorimeter. The bold outlines represent groups of blocks whose outputs were combined to form sums used in the trigger.

the sums from these groups of blocks were themselves added, the result being a fast calculation of the total energy deposited in the calorimeter which was used for triggering purposes.

Throughout the run the blocks, and particularly those close to the holes through which the beams passed, were subject to large doses of radiation. This deteriorated performance of the blocks due to the absorption of Cerenkov light in radiation damaged areas. Blocks near the center of the array, between the holes received as much as 23 krad integrated dose. To compensate for this, at approximately every two weeks of beam-time throughout the run the array was subjected to UV light from a 400 Watt mercury lamp for about 12 hours. Previous experiments had shown that lead-glass blocks can recover at the 90% level from this treatment.

Previous experience had shown that the phototube gains can change by up to 10% at the beginning of each spill. This "warm-up" effect of the phototubes was removed by illuminating the blocks with a low level light LED light source. Gains of the phototubes were monitored throughout the run using a Xenon flasher lamp, with a spectrum similar to that of Cerenkov radiation.

2.2.3 Photon Veto Detectors

At various z-locations there were detectors designed to veto charged and neutral decay products that were outside the acceptance region of the calorimeter. The

Detector	Aperture	Composition	z-location (m)
VA-1	Square	Scint/Pb-Lucite Sand.	122.9
VA0	Square	Scint/Pb-Lucite Sand.	125.2
VA1	Circular	Scint/Pb-Lucite Sand.	132.2
VA2	Circular	Scint/Pb-Lucite Sand.	135.9
DRA	Square*	Scint/Pb/Scint	140.9
VA4	Circular	Scint/Pb-Lucite Sand.	158.3
MA	Square	Scint/Pb-Lucite Sand.	166.8
LGA	Circular	Scint/Pb-Lucite Sand.	178.7
Iron Ring	Circular	Iron	179.3
BA	n/a	Pb/Lucite Sand.	185.0

Table 2.1: Photon Veto detector properties.(*Note: DRA had a circular outer aperture, the only veto counter of this type.)

counters were composed of various materials and came in different shapes as listed in Table 2.1

All of the veto detectors were designed to detect charged or neutral particles that would have otherwise have been lost without detection. Listed in Table 2.1 are the shapes of the inner apertures of the veto counters. The scintillator/Pb-Lucite construction of the VA detectors is illustrated in Fig. 2.9. Photon veto counters VA -1 through 4 and the Decay Region Anti (DRA) were located inside the vacuum region. The Magnet Anti (MA) covered the dead region of the magnet. The Lead Glass Anti (LGA) covered the region extending outwards from the outer edge of the calorimeter. The LGA left a gap in the acceptance between its inner edge and the outer edge of the calorimeter and a 2.9 radiation length thick iron ring was placed further upstream to cover this crack. The idea was to make photons shower so that

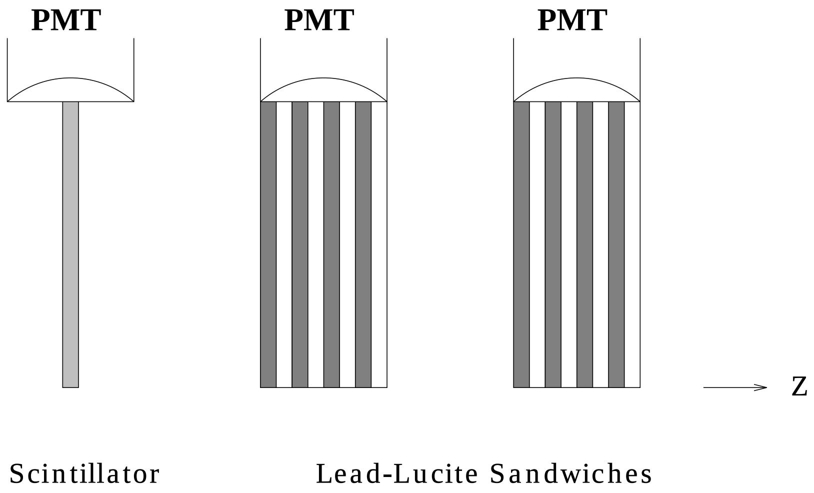

Figure 2.9: Cross section of a scintillator/Pb-Lucite sandwich veto detector.

the shower products could be detected. The most downstream veto detector was the Back Anti, positioned 5 m behind the calorimeter, its purpose was to detect particles that escaped through the beam-holes in the lead-glass. The BA was a segmented lead-lucite sandwich, 28.1 radiation lengths in depth, viewed by 72 phototubes. The segmentation meant that hadronic interactions and electromagnetic showers could be distinguished based upon the depth of the shower maximum. Figure 2.10 illustrates the coverage that the system of veto detectors provided.

2.2.4 Trigger Hodoscope Banks

For the purposes of triggering on charged tracks two planes of scintillators provided complete coverage of roughly a 2 $m \times 2$ m region directly in front of the calorimeter (see Fig. 2.4). The trigger banks, called B-bank and C-bank for historical reasons,

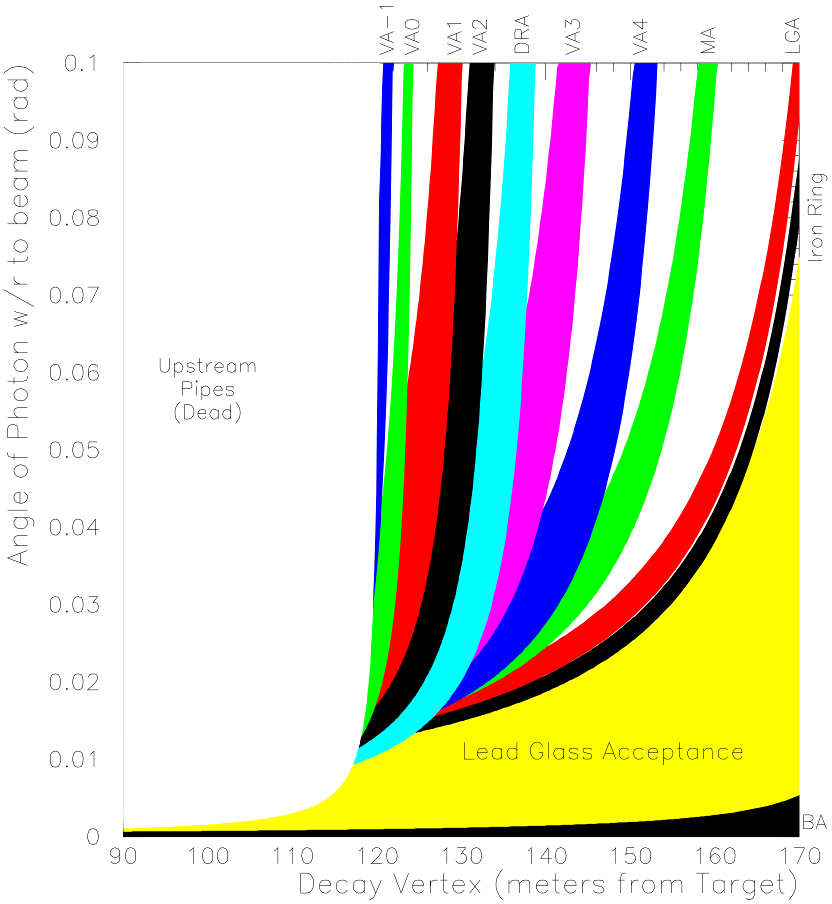

Figure 2.10: The Veto detector coverage plotted as a function distance from the target. This plot refers to photons produced in the lower beam, and traveling downwards.

are depicted in Figs. 2.11 and 2.12. This geometry allowed triggers to be formed to select different numbers of tracks. For example: at least two hits in each bank would trigger on events consistent with two charged tracks.

2.2.5 Muon Identification

There were three hodoscope banks in the detector, called MU1–3. MU2 was not used in this analysis.

MU1 was composed of 45 overlapping scintillator paddles positioned immediately behind a 10 *cm* thick lead wall and viewed by one phototube per counter (Fig. 2.13). The phototube signals were combined in an analog sum for trigger purposes. The lead wall was 0.6 interaction lengths and combined with the lead glass which was 1.7 interaction lengths the total amount of material upstream of MU1 was $2.3\lambda_I$; the same material comprised 44 radiation lengths. This geometry was optimized for previous experiments to reject hadronic shower leakage from the back of the calorimeter and at the same time retain events with electromagnetic showers. As will be seen in a later chapter, the MU1 bank could be used off-line to reject about 80% of K_L decays with at least one pion present. During E799 data taking MU1 was used in the dimuon trigger to lower the level-one rate. If the veto threshold for MU1 were set sufficiently above an average two minimum ionizing deposit, very few events with two muons would be rejected. However, the effect of putting a

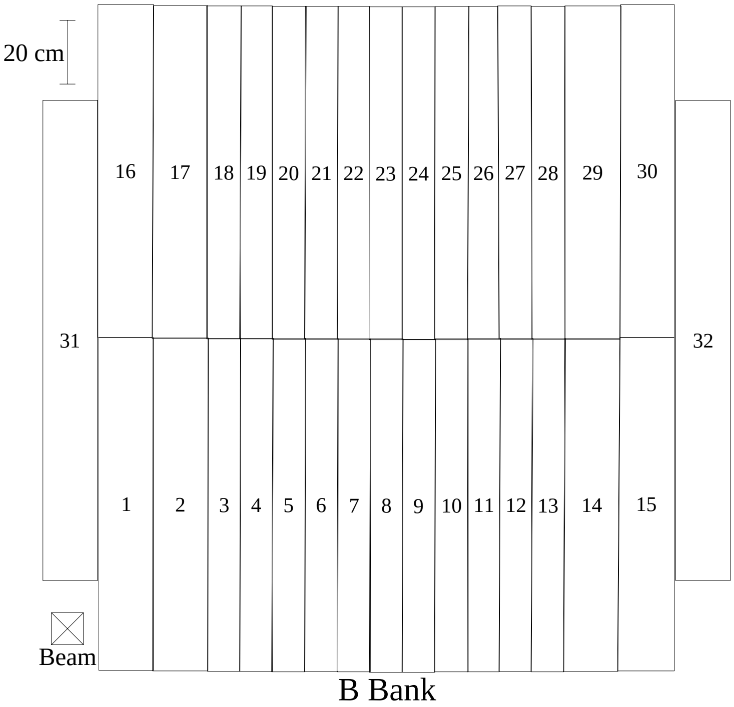

Figure 2.11: Vertical bank of trigger counters.

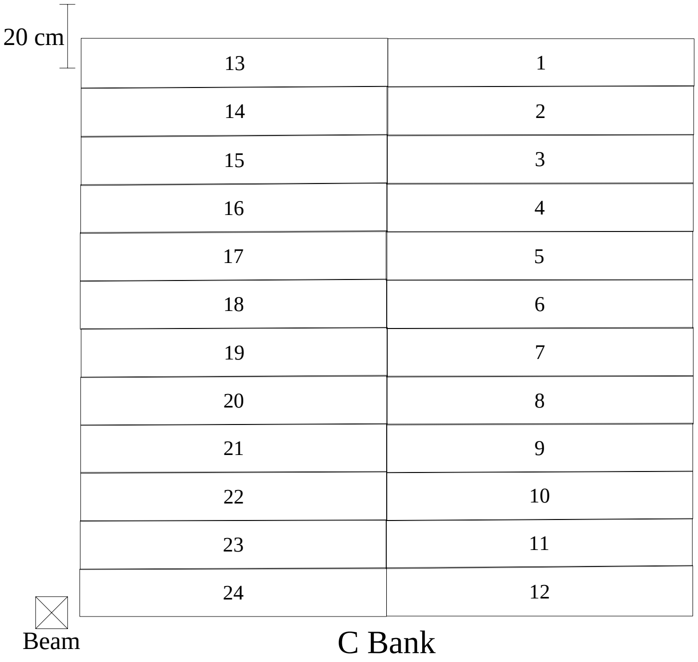

Figure 2.12: Horizontal bank of trigger counters.

counter sensitive to muons in veto are unfortunately not small and will be discussed at length in a later chapter.

MU3 consisted of 16 non-overlapping scintillator paddles viewed from the top by one phototube per counter. The bank was positioned approximately 10 cm behind 3 m (18 interaction lengths) of muon filter steel. The phototube signals were the inputs to a custom logic unit which required at least two hits, and if only two counters were above threshold they were required to be non-adjacent. The threshold for single muons to penetrate the filter steel was about 5 GeV.

2.3 Triggers

Triggering was divided into two levels, the first level was a fast decision, typically made immediately after each bucket and available after every bucket ($\sim$ 20 ns). Table 2.2 lists the detector information that was available at for first level triggering. A successful level-one trigger initiated a more complex level-two trigger of which only the following are relevant to this analysis:

- Hit Counting: A series of custom electronics boards and commercially available modules performed hit-counting on the drift chamber signals[52]. Two-track triggers required at least two hits per chamber-view.

Figure 2.13: MU1 hodoscope bank.

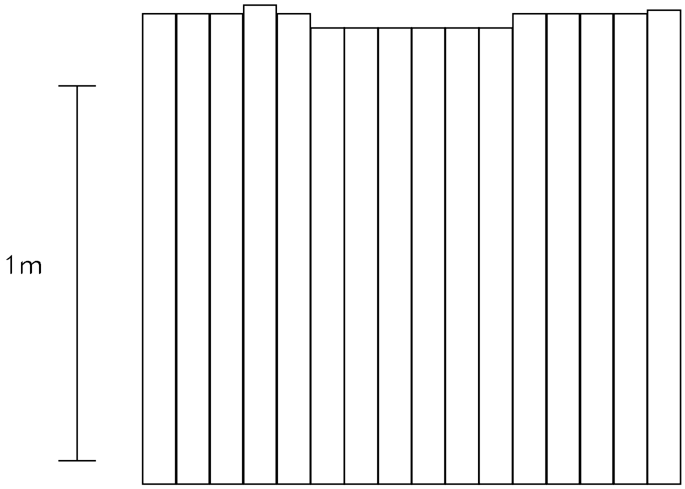

Figure 2.14: MU3 hodoscope bank.

- Hardware Cluster Finder (HCF): Another custom electronics module stored digitized hits from the lead-glass blocks that contained at least 2.7 GeV in a 30 ns charge integration gate for the data used in this analysis. Neighboring blocks were then counted as a single cluster, and the number of clusters was counted.

The result of the level-two trigger was typically available 30 μs from initiation. This time was dominated by the HCF.

The two-muon trigger was implemented to select hits in the drift chambers and trigger hodoscope banks consistent with two tracks, no photon veto counters above threshold, hits in the MU3 bank consistent with two muons, and a minimum energy deposit in the calorimeter in order to select photon candidates. Using the notation of Tab. 2.2 we write

$$2B \cdot 2C \cdot E_t^{Low} \cdot \overline{PHV} \cdot \overline{MU1} \cdot MU3 \cdot \text{HiTCnt(2track)}. \tag{2.1}$$

A minimum bias trigger was run simultaneously, and designed to collect two track events, including $K_L \to \pi^+\pi^-\pi^0$ which are used for normalization purposes. The trigger was

$$2B \cdot 2C \cdot \overline{PHV} \cdot \text{HiTCnt(2track)}. \tag{2.2}$$

A prescale of 3600 was applied to the two-track minimum bias trigger.

In addition an “accidental” trigger was constructed from two scintillator paddles forming a crude telescope directed toward the targeting region. The solid angle subtended by this telescope did not include the detector however. This trigger was designed to detect muons from the target area, and given the number of protons on target per bucket ($\sim$1000), it was assumed that this trigger was not correlated with kaon decays in the detector. All detector elements were read out for each accidental trigger. These were then overlaid upon individual Monte Carlo responses in order to account for the instantaneous ambient activity in the simulation of the detector.

Finally there were some special runs taken for calibration purposes. Notably were a series of “$B \cdot MU2$” triggers taken with a piece of copper in the beam pipes so that all particles except muons would be filtered. A similar set of $B \cdot MU2$ and $C \cdot MU2$ triggers were taken with the copper beam-stop removed. These triggers were designed to collect samples of single track muons for calibration purposes.

This does not represent an exhaustive list of triggers taken for E799, but only those relevant for this analysis. Data were collected during the fixed target run at Fermilab in 1991. In 10 weeks of data taking about 60 million two-muon triggers were written to tape from 115 runs. The total kaon flux in this time was about 30 billion for decays in the region $90 < z < 160\ m$ and for parent kaon momenta in the range $20 < |p| < 220\ GeV$.

Detector Element	Source Name	Description
B Bank	1B	one or more counters in the bank hit
	2B	two or more counters in the bank hit
	3B	three or more counters in the bank hit
	BE	one or more counters in the east half of bank hit
	BW	one or more counters in the west half of bank hit
C Bank	1C	one or more counters in the bank hit
	2C	two or more counters in the bank hit
	3C	three or more counters in the bank hit
	CU	one or more counters in the upper half of bank hit
	CD	one or more counters in the lower half of bank hit
B and C Banks	BCPH	requires a hit in a central counter of each bank
Lead-Glass	E_t Low	Rising edge above low ($\approx$ 6 GeV) threshold
	E_t High	Rising edge above high ($\approx$ 50 GeV) threshold
	Q1,Q2,Q3,Q4	one quadrant of glass above low threshold
Photon Vetoes	VA -1, VA0	a hit in the veto scintillator
	VA1-4, DRAC	a hit in the veto scintillator
	LGA	signal above threshold in the LGA lucite counters
	BA12	signal above $\approx$ 2.5 GeV in X-view of first 2/3 of BA
	BA3	signal above $\approx$ 5 GeV in X-view of last 1/3 of BA
MU1	MU1	signal above threshold in analog sum of counters
MU2	MU2	a hit in any counter
MU3	MU3	hits in two non-adjacent MU3 counters
Accidental Trigger	ACCID	a coincidence in the Accidental muon telescope
Beamhole Counters	UPBM	a hit in the counter in the upper beamhole
	DNBM	a hit in the counter in the lower beamhole

Table 2.2: Level one trigger sources.

Chapter 3

Searching for $K_L \to \mu^+\mu^-\gamma$

In this chapter we give a description of the methods used to reconstruct the track directions and the energies of kaon decay products from the digital information of the detectors. Our Monte Carlo simulation of the detector is also discussed.

3.1 Reconstructing the Event

The physics decay of interest, $K_L \to \mu^+\mu^-\gamma$, requires two tracks and an isolated deposit of energy in the calorimeter that is consistent with originating from a photon. Looking at all K_L decays with two charged tracks in the decay products, we would expect those with the largest branching fractions to be potential sources of background, namely: $K_L \to \pi^+\pi^-\pi^0$ (B=12.38%) and $K_L \to \pi^\pm\mu^\mp\nu$ (B=27.0%). For these to be mistaken as $K_L \to \mu^+\mu^-\gamma$, one or more of the decay products would have to be misidentified or lost completely in addition to having some rare and random energy deposit in the lead glass. Given that there are ~1000 protons incident on target each bucket and that on average each kaon is accompanied by about two neutrons that end up interacting in the back anti, it is not surprising that there would be such "accidental" activity present. Accidental activity is defined for us to be any energy deposits not associated with the decay of the parent kaon. Most accidental activity comes from the following sources:

- Muons from the target
- Interactions of the neutral beam particles
- Kaon decays other than the primary decay of interest.

For a $K_L \rightarrow \pi^+\pi^-\pi^0$ to fake a $K_L \rightarrow \mu^+\mu^-\gamma$ event it is necessary for one of the photons from the decay of the π^0 to pass through one of the gaps in the acceptance or for one of the photons to overlap with another cluster from a track for example. When photons are lost there is missing energy in the event and the remaining products will have an invariant mass below that of the neutral kaon. In addition we would have to misidentify both charged pions in the $K_L \rightarrow \pi^+\pi^-\pi^0$ decay as muons, which may occur either because there happens to be two hits in the muon identification hodoscope from accidental activity or because the charged pions decay into muons, or a combination of these two effects. Then one might expect that the charged tracks would not extrapolate to the hit counters in the MU3 bank. Also one may hope to reject such fake signal events by looking for pion showers in the calorimeter where muon tracks from $K_L \rightarrow \mu^+\mu^-\gamma$ events would only leave a minimum ionizing deposit. In the case of $K_L \rightarrow \pi^\pm\mu^\mp\nu$ there is already one muon present from the kaon decay so we anticipate that it will be easier for this decay to fake a signal event: if the pion is misidentified as a muon, there is a accidental deposit of energy in the calorimeter, and there is spurious hit in the MU3 bank. We label this class of events as $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$. A further kaon decay that could potentially mimic a $K_L \rightarrow \mu^+\mu^-\gamma$ decay is the decay $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{rad}$, the radiative version of $K_L \rightarrow \pi^\pm\mu^\mp\nu$.

With the expected branching ratio for $K_L \rightarrow \mu^+\mu^-\gamma$ at the level of 10^{-7} and the most probable backgrounds in the range 10^{-1}, it would not take that many misidentifications, or accidental energy deposits to completely bury a signal in background events. However, with the exception of $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{rad}$, the backgrounds can be distinguished from signal using kinematical quantities. Firstly one would expect the total momentum of the decay products transverse to the beam to be zero for real $K_L \rightarrow \mu^+\mu^-\gamma$ events. For backgrounds like $K_L \rightarrow \pi^+\pi^-\pi^0$, where a photon has to be lost in order for the event to be misidentified as signal, these events would be expected to be quite spread out in the transverse momentum parameter (when a photon overlaps with another cluster this wouldn't be the case). Also one might expect to see large energy deposits from pion showers in $K_L \rightarrow \pi^+\pi^-\pi^0$ and $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$ events which are absent in the signal mode. Finally, the invariant mass of the candidate signal events would be expected to be clustered around the kaon mass under the $K_L \rightarrow \mu^+\mu^-\gamma$ hypothesis, whereas background processes which require the loss or gain of photons would not be reconstructed at the kaon mass under the $K_L \rightarrow \mu^+\mu^-\gamma$ hypothesis.

As we do not have accurate measurements of the number of protons on target we cannot measure the number of kaons entering the detector region. However we can look for a high statistics decay mode that is similar to $K_L \rightarrow \mu^+\mu^-\gamma$ and measure the branching fraction relative to that mode. Ideally we would look for a decay mode with two muon tracks in order to reduce the systematic bias by having the analysis of the signal and normalization modes as similar as possible. However the obvious candidate, $K_L \rightarrow \mu^+\mu^-$ is another rare decay with a branching fraction $\sim 10^{-2}$ less than that expected for $K_L \rightarrow \mu^+\mu^-\gamma$. The decay $K_L \rightarrow \pi^\pm\mu^\mp\nu$ is another candidate, but the neutrino always remains undetected and establishing a sample of $K_L \rightarrow \pi^\pm\mu^\mp\nu$ events that is well measured is difficult. This leaves the most viable normalization mode as $K_L \rightarrow \pi^+\pi^-\pi^0$ which has a very small background.

With these considerations in mind, we are led to the search procedure: (i) we want to search for two well measured tracks passing through the spectrometer and to be able to measure their momenta; the resolution of the momenta can never be too good as this directly impacts the resolution of the reconstructed invariant kaon mass. (ii) We want to search the lead glass blocks for energy deposits both in order to reject pions and to identify and measure the energy of candidate photons.

3.2 Tracking

The search for tracks begins by finding all hits in the drift chambers. Given the two-layer geometry (Fig. 2.5) of the chambers these hits could be grouped accordingly:

- a pair of hits in complementary cells for which the sum of the drift distances is close to the cell size (an in-time pair)
- hit pairs for which the sum of drift distances is significantly above or below the cell size (out-of-time pairs)
- single hits adjacent to in-time pairs
- single isolated hits
- single hits between pairs

A “pair” always refers to hits in different planes. Note that groups of more than three hits are characterized by multiples pairs. In the case of ties for sum-of-distances or clusters of hits with no in-time pairs, the hits were paired off arbitrarily according to the order the raw event data was read from tapes. The method adopted to search for tracks was different for x-track segments and y-track segments.

In the x-view all possible track segments were formed between hits in chambers 1 and 2. Similarly all candidate segments were found between chambers 3 and 4. This can result in many possible tracks, and because there may be isolated hits with ambiguity in the hit position, even single hits could produce multiple track segments. To reject unphysical combinations, the upstream candidate x-track segments were required to extrapolate to within 10 cm of the beam center at the vacuum window (the beam size at that position was 9 cm) and the downstream candidate x-track segments were required to extrapolate to a position within the lead glass array. Further unphysical segments were rejected by requiring a continuous track through the spectrometer; the extrapolated x-position of downstream segments at the center-plane of the magnet had to match the extrapolated x-position of upstream segments in the same plane to within 1 cm (10σ). At this stage there may be many track candidates that share hits and a "goodness of fit" parameter is needed to choose the best combination of segments. This of course may introduce biases into the analysis that would be difficult to model correctly. However in $K_L \rightarrow \mu^+\mu^-\gamma$ we deliberately chose a normalization mode, $K_L \rightarrow \pi^+\pi^-\pi^0$, that also has two tracks. Thus in the ratio of acceptances any tracking inefficiencies will largely cancel. Then the only possible difference between the normalization mode and the signal mode will be the possibility that pions shower (an effect which is not simulated in our Monte Carlo - see below). However, there was a total of about 0.005 interaction length between the vacuum window and the lead glass, and for this analysis we may safely ignore effects this small. Therefore the exact algorithm for choosing the tracks is not as important as making it consistent between modes. The technique used was to try and quantify what a "good" track is. For example, tracks that have only in-time pairs, with no other hits are obviously the cleanest that we might hope for. Tracks with a single hit next to an in-time pair were acceptable as this may occur for example, if there was a δ-ray that produced a hit on a neighboring wire. Single isolated hits would be considered worse as cells were typically 98% efficient. Accepting a track with two such isolated hits was not considered acceptable. An out-of-time pair would be considered worse still as it could either have come from a δ-ray triggering a wire first, and the TDC dead time ensuring that the real ionization pulse went undetected, or it may come from an accidental event. There is clearly a tradeoff between track cleanliness and reconstruction efficiency. These considerations were implemented in a quantifiable way by assigning each hit or hit pair a quality number, and summing all of the quality numbers over a given track candidate (see Tab. 3.2). Then in the x-view, all of the possible track segments were ranked according to their total quality number, and only those with a quality within 2 units of the best were kept for further consideration. The track with the best offset at the magnet plane was

Classification	SOD Requirement	Quality Code
In-Time Pair	$\|\mathrm{SOD} - d_{\mathrm{cell}}\| < 1.5$ mm	0
Late Pair	$(\mathrm{SOD} - d_{\mathrm{cell}}) < -1.5$ mm	individual hits were stored with status code 1; the pair with 3
Early Pair	$(\mathrm{SOD} - d_{\mathrm{cell}}) > 1.5$ mm	3
Isolated Single Hits	n/a	2
Single Hits Next To a Pair	n/a	1
Single Hits Between Two Pairs	n/a	2

Table 3.1: Classification of chamber hits

then determined to be a good track. Any other tracks sharing hits with this track were eliminated. Remaining tracks were re-routed through this algorithm until the list of track candidates was exhausted.

The y-track finding algorithm differed because the analog of the magnet offset no longer exists when there is no bend in the track. Instead the hits from chambers 1 and 4 were used to form candidate y-segments. Then it was required that hits exist within 1.5 *cm* of the intersection of this y-segment and the hit planes in chambers 2 and 3. No y-track candidates were kept if their sum of total quality number was greater 4. Also the segment candidates were required to extrapolate to the beam in the upstream direction and the calorimeter in the downstream direction, just as for x-tracks. Tracks with single hits have an ambiguity of the hit position. For this case the χ^2 of each alternative was calculated and and the hit providing the lowest

χ^2 was used. As with the x-view the best track was chosen, this time the tracks were ranked in order of χ^2. All tracks sharing hits with this highest ranked one were rejected, and the algorithm was repeated on the remaining candidates. The tracking efficiency of events passing the trigger for finding two tracks was 80% in $K_L \rightarrow \mu^+\mu^-\gamma$ decays.

Finally, tracks in the two independent x and y planes were paired. This was done by requiring that tracks in each of the two views project to a common energy deposit in the lead glass calorimeter.

The momentum scale was found roughly by a survey of the magnetic field, which fixed the absolute momentum scale to about 0.2%. Then the kaon decay $K_L \rightarrow \pi^+\pi^-$ was used for fine-tuning as one can reconstruct the kaon mass to a very good accuracy using only the bend angle of the tracks at the plane of the magnet for this decay mode [53]. Using $\theta_{x,y} = p_{x,y}/p_z$ the relationship between track momentum and angle can be derived,

$$\left(\frac{\theta_x^u}{\sqrt{1+(\theta_x^u)^2+(\theta_y^u)^2}} - \frac{\theta_x^d}{\sqrt{1+(\theta_x^d)^2+(\theta_y^d)^2}}\right) = \frac{\Delta p_t}{p}$$

where Δp_t is a constant, and $\theta^{u,d}$ refer to the track angles upstream and downstream of the magnet. There are two major contributions to track resolution: multiple scattering, and single hit resolution. Because the momentum is proportional to the inverse of the bend angle the relative error in the momentum is equivalent to the relative error in the bend angle (θ_B) to first order: $\sigma_p/p \propto \sigma_{\theta_B}/\theta_B$. Multiple scattering has an RMS that is $\langle\theta\rangle \propto 1/p$ and therefore produces a constant term in the relative uncertainty of the reconstructed momentum. The 100 μm single hit resolution is momentum independent and the uncertainty in the reconstructed angles is a constant. Therefore single hit resolutions cause a growing uncertainty as track momentum increases. The momentum resolution for the spectrometer in E799 was given by $(\sigma_p/p)^2 = (5 \times 10^{-3})^2 + (1.4 \times 10^{-4} p/(GeV/c))^2$ [52].

3.3 Calorimeter

The lead glass calorimeter presented many problems in reconstruction. We would like to be able to search for clusters of energy in the data, and identify photons both for $K_L \rightarrow \pi^+\pi^-\pi^0$ and $K_L \rightarrow \mu^+\mu^-\gamma$. We need to be able to reconstruct the position and energy of clusters both isolated and associated with tracks.

Throughout the data taking period a series of special runs were taken in which the lead converter was removed from the upstream region of the beam line (Fig. 2.1) and also at a lower beam intensity. This enhanced the photon content of the downstream beams considerably. At a z-location of 119 m from the target a copper/titanium plate was placed in the beam (see Fig. 3.1) to convert photons [54]. Magnets (AN1, AN2) separated the resulting e^+e^- pairs and the analysis magnet directed them onto the lead glass to ensure that sufficient data was taken to calibrate each block. Off-line analysis of the tracking data was used to find events with two reconstructed tracks with at least one of them extrapolating to a cluster in the calorimeter. These electrons were then used to determine a single number for each block that maps the measured ADC value to an energy. The fundamental requirement is that on average $E/p = 1$ for electrons. Since clusters typically have leakage into neighboring blocks a fitting procedure was to used extract gains that optimized the E/p ratio. This procedure to calibrate the gains was iterated until E/p was stable.

There were many corrections to the simplified scheme described here that were required because of various non-linearities that occur at about the percent level ([52, 54]). These included

- Shower leakage: A 50 GeV electromagnetic shower could lose up to 1.5% of the shower energy from the rear of the block. In general this correction is energy dependent.
- Transverse leakage: About 2.5% of shower energy could be lost from the boundaries of a 3×3 array of blocks centered on the shower. This correction is less sensitive to shower energy.
- Light Attenuation: The Cerenkov light produced in showers is attenuated over its path length before it reaches the photocathode surface (About 1-2% per radiation length, up to 10% for blocks near the beam-holes).

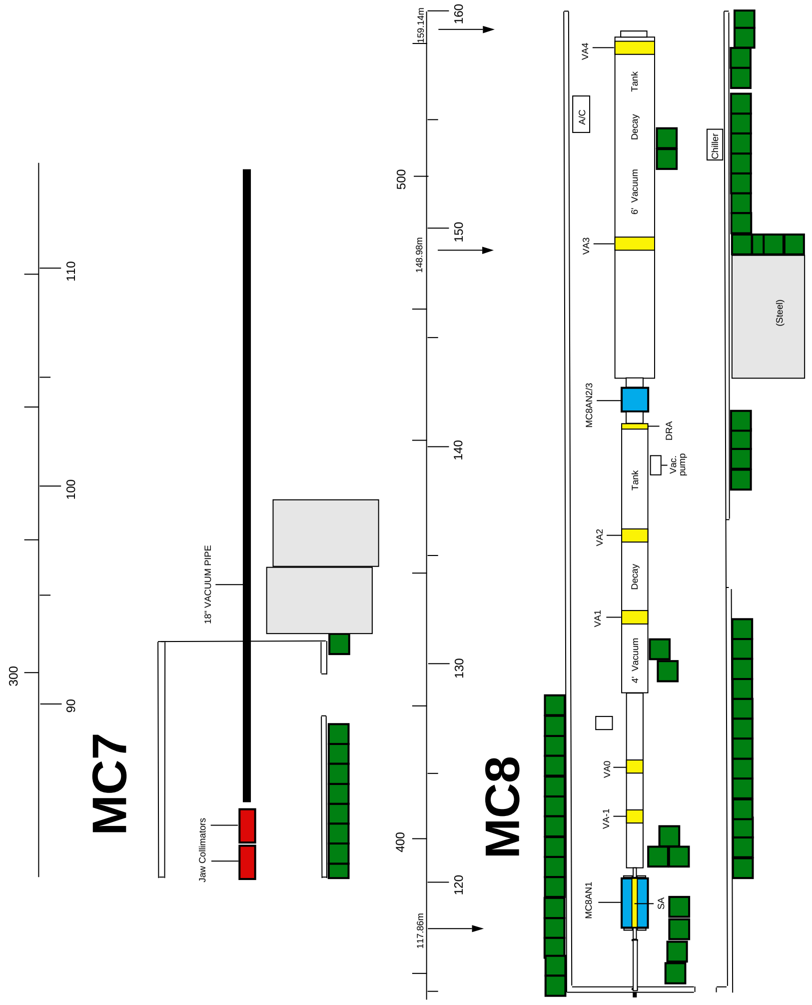

Figure 3.1: Detector Hall layout showing AN1 and AN2 magnet locations.

- Time variation: The lead glass blocks exhibited degraded performance as a function of time as a result of radiation damage. Although this was mostly reversed as a result of curing with ultra-violet light, there was still an increase in attenuation with time as the curing was incomplete.

- Shower-Type variations: Photon showers begin deeper in the lead glass than electron induced showers, and therefore the shower maximum is closer to the phototube. This increases the amount of light that a photon shower will produce at the photocathode relative to that of an electron shower. This non-linearity is compensated for somewhat because more of the shower energy is lost from the rear of the array and hence the light yield is reduced relative to an infinitely deep array. Muons on the other hand are minimum ionizing and produce light uniformly long their path. We calibrated the lead glass array using electrons (which shower with a peak at about 6-$7X_0$ out of a total of $18X_0$ of lead glass). Using this calibration the apparent muon energy deposits were larger than might have been expected by about a factor of two. This didn't affect the analysis however as we always selected events based on a comparison to one MIP deposits.

- Rate dependence: The calibration electrons were collected at a higher rate than the normal triggers. This introduces potential sources of bias between calibration and physics triggers.

The main tool used for making corrections to the initial calibration was to simulate the electron response of lead glass blocks using the EGS Monte Carlo shower generator[55]. The changes in response caused by radiation damage were taken into account by having separate calibration files for different time periods. Shower leakage corrections were made based on the Monte Carlo simulations, as was the difference in electron and photon responses. We plot the reconstructed E/p distribution for calibration electrons in Fig. 3.2. The resolution of the calorimeter is shown in Fig. 3.3 for calibration electrons. The usual form used to describe calorimeter resolutions is $\sigma_E/E = a \oplus b/\sqrt{E}$, where the energy dependent term describes fluctuations in the number of photoelectrons produced at the photocathode of the PMT and the constant term describes primarily the statistical fluctuations in shower size. The parameterization clearly fails for this calorimeter at large energies. This is attributed to a tendency for high momentum tracks to be near the center of the calorimeter where the radiation damage (and hence the absorption) was the greatest. The resolution for electrons at 10 GeV is about 4%. For photons is is slightly worse, about 5% at 10 GeV as determined from studying the photons from $K_L \rightarrow \pi^+\pi^-\pi^0$.

Clusters in the calorimeter were found by looking for local maxima of energy deposits. The algorithm searched all 3×3 arrays of calorimeter blocks with energies above 500 MeV that also contained a local maximum. The local maximum was

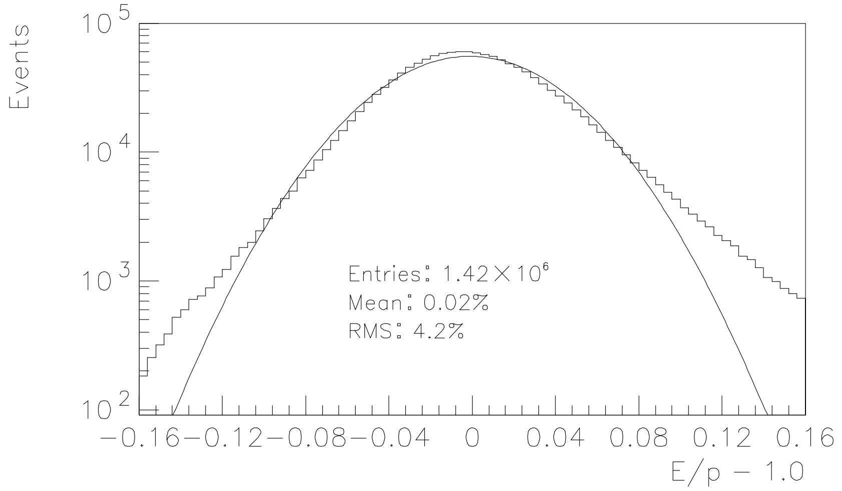

Figure 3.2: Resolution for calibration electrons [54].

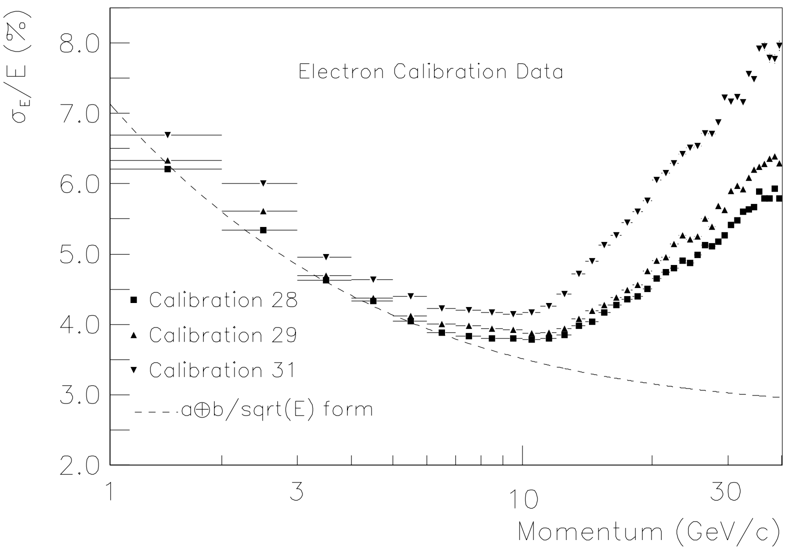

Figure 3.3: Resolution of lead glass calorimeter versus track momentum for electrons [54].

defined as a block which has a greater energy deposit than all of the surrounding eight blocks. Once clusters had been associated with tracks, the remaining clusters (which were candidate photon energy deposits in $K_L \to \mu^+\mu^-\gamma$ and $K_L \to \pi^+\pi^-\pi^0$) were required to contain a Hardware Cluster Finder seed. By imposing this requirement energy deposits that were out-of-time were suppressed as the HCF charge integration gate was only 30 ns as compared to the 100 ns gate on the lead glass ADC's. Once a cluster had been identified the position of the parent particle was estimated. For example to find the x-position of a cluster, three column energies were defined by summing down rows as shown in Fig. 3.4. A sample of electrons whose position at the calorimeter was well known from tracking was used to form a lookup table that related this column energy ratio to the distance from the center of the central block. This method gives an average position resolution of about 3 mm [54]. All of these techniques assume that energy deposits from two particles do not overlap, and extensive algorithms were developed to search for this possibility [52]. The fusion finding algorithms all searched for an anomalous pattern of energy deposition in a cluster either by summing columns and rows of lead glass responses or by looking in larger 5×5 rings. In addition there were complications when clusters were near the edges of the array which required special treatment [54].

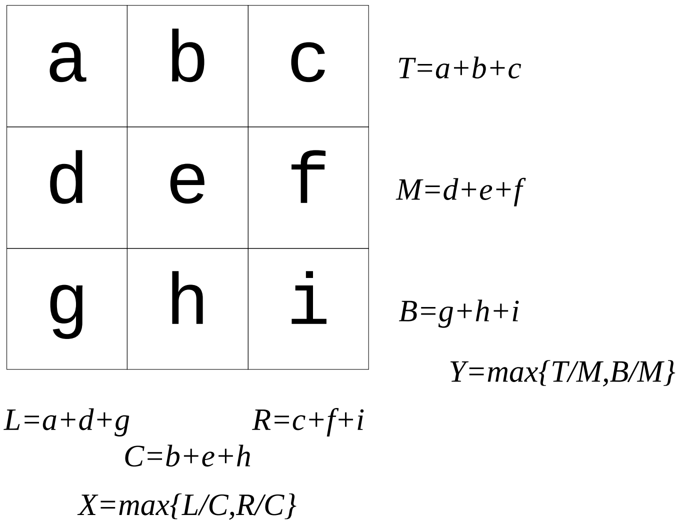

Figure 3.4: Position finding algorithm in the calorimeter [52].

3.4 Simulating the Detector

A comprehensive simulation of the detector was performed in order to understand the total acceptance [52]. The simulation evolved the $K^0/\bar{K}^0$ admixture from production at the target throughout the detector. The simulated kaons were then allowed to decay and the decay products we followed until they left the detector region, or interacted in a detector volume.

For a given sample of pure $K_L \to \pi^+\pi^-\pi^0$ decays, the detector acceptance together with a set of analysis cuts designed to isolate only clean events will reduce the size of the sample, by some factor $\mathcal{A}_{+-0}$. Similarly a sample of $K_L \to \mu^+\mu^-\gamma$ events which are passed through an analysis designed to reject background and select clean muons will be reduced by some fraction, $\mathcal{A}_{\mu\mu\gamma}$. The measured quantity BR($K_L \to \mu^+\mu^-\gamma$)/BR($K_L \to \pi^+\pi^-\pi^0$) will then be proportional to the ratio $\mathcal{A}_{+-0}/\mathcal{A}_{\mu\mu\gamma}$. Any difference between Monte Carlo generated events and data may potentially produce the wrong acceptance, especially if a similar cuts are not made in both analysis and normalization data samples. Then it becomes important for the Monte Carlo to have the correct thresholds and the correct distribution of decay products transverse to the beam direction for example. These have been studied in detail for other analyses [52, 54, 56, 57, 58]. We present several comparisons between data and Monte Carlo to illustrate the good agreement. Figure 3.5 shows the comparison between Monte Carlo and data of the z-position of the decay vertex in $K_L \to \pi^+\pi^-\pi^0$ events using the tracking information from the upstream chambers. Figure 3.6 shows the distributions of positions of the tracks when the downstream segments are extrapolated to the face of the lead glass array. The distributions of reconstructed photon positions in the calorimeter are shown in Fig. 3.7.

The proceeding chapters will examine further Monte Carlo distributions and in particular the muon system.

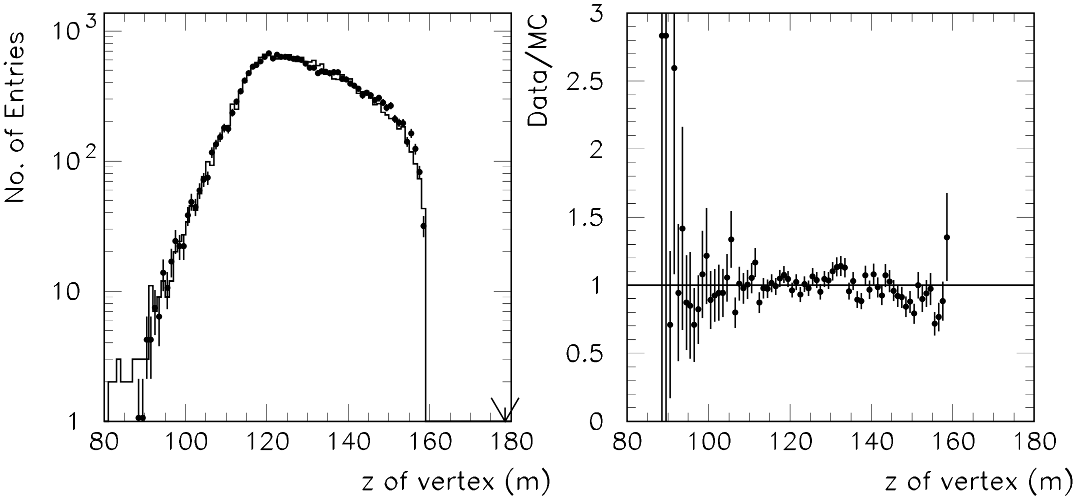

Figure 3.5: The reconstructed z-position of the vertex in $K_L \rightarrow \pi^+\pi^-\pi^0$ decays calculated from tracking information only. Points are Monte Carlo, histogram is data.

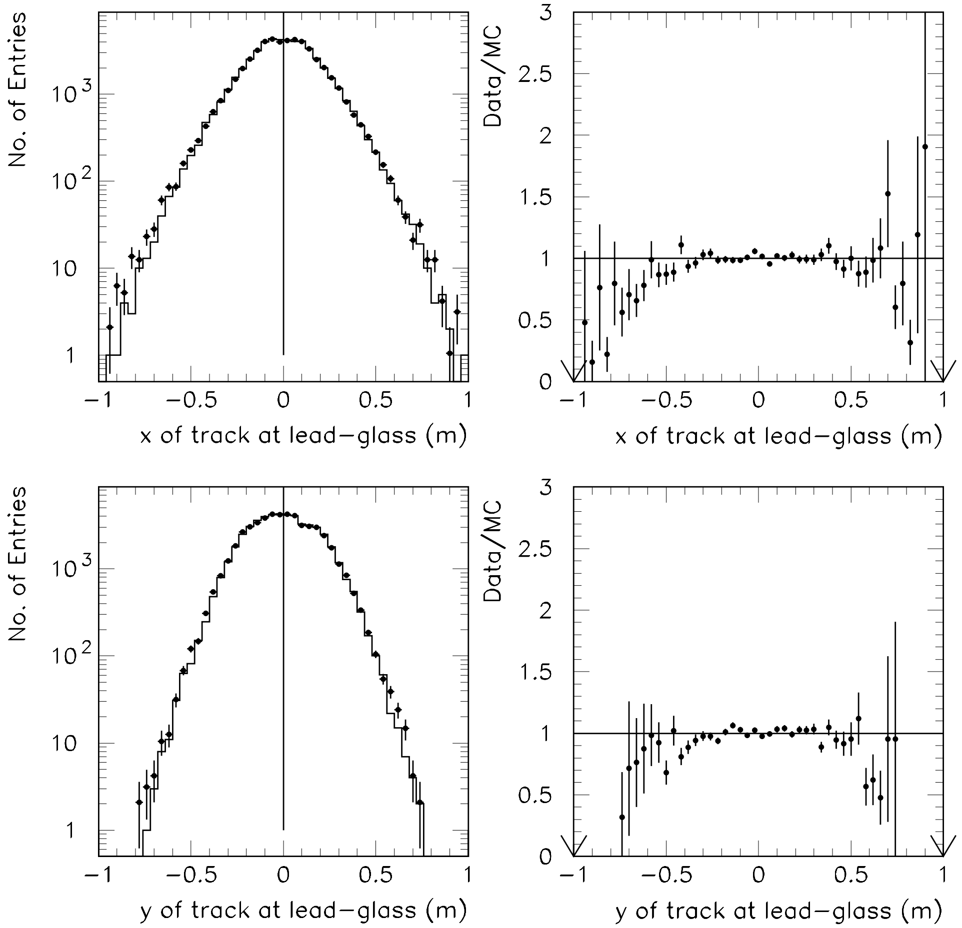

Figure 3.6: The reconstructed x- and y-positions of the tracks in $K_L \to \pi^+\pi^-\pi^0$ decays at the lead glass face. Points are Monte Carlo, histogram is data.

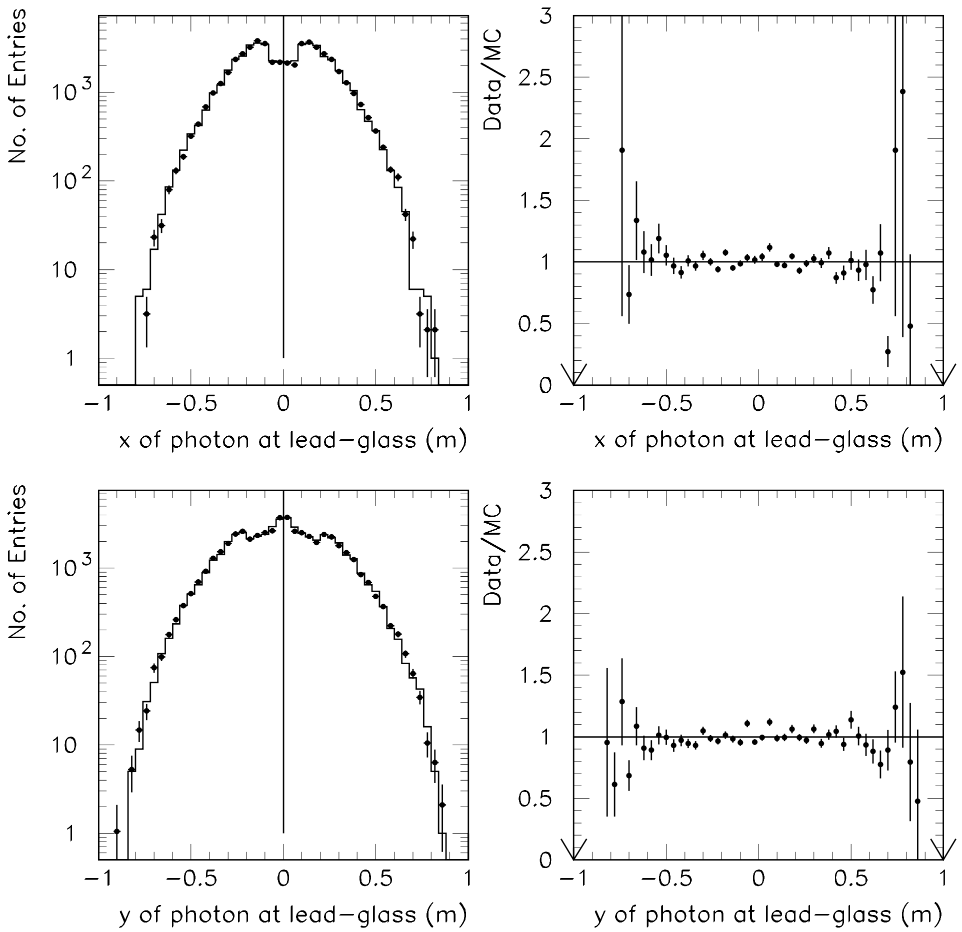

Figure 3.7: The reconstructed x- and y-positions of the photon clusters in $K_L \to \pi^+\pi^-\pi^0$ decays. Points are Monte Carlo, histogram is data.

Chapter 4

Muon Calibration

In experiments where the absolute flux of primary particles is not well known branching ratios are typically measured relative to another decay mode. A Monte Carlo simulation of the detector is used to determine the acceptance and the relative branching ratio is calculated by

$$\frac{B_{sig}}{B_{norm}} = \frac{N_{sig}}{\mathcal{A}_{sig}} \div \frac{N_{norm}}{\mathcal{A}_{norm}} \tag{4.1}$$

where N_i and $\mathcal{A}_i$ are the numbers of events observed and the acceptances respectively. The acceptances factorize into a product over the individual detectors and the result is

$$\frac{B_{sig}}{B_{norm}} \propto \frac{\mathcal{A}^1_{norm} \cdot \mathcal{A}^2_{norm} \cdot \mathcal{A}^3_{norm} \cdot \ldots \mathcal{A}^n_{norm}}{\mathcal{A}^1_{sig} \cdot \mathcal{A}^2_{sig} \cdot \mathcal{A}^3_{sig} \cdot \ldots \mathcal{A}^m_{sig}}. \tag{4.2}$$

In this equation the possibility that $n \neq m$ allows for existence of different detector elements in the analysis for signal and normalization decay modes. An important feature of this equation is that when the signal and normalization modes are similar kinematically the acceptances cancel to a large degree in the calculation, and the result becomes insensitive to the Monte Carlo simulation. This is one reason to choose a normalization mode as kinematically similar to the signal mode as possible. The converse situation is that when a certain detector is in the trigger for one of the modes but not the other the cancellation does not occur and we must understand the absolute acceptance of that detector much more accurately. This is the case for the analysis of $K_L \rightarrow \mu^+\mu^-\gamma$ where the normalization mode is $K_L \rightarrow \pi^+\pi^-\pi^0$. The normalization data was part of the minimum bias trigger and the candidate signal events were in the dimuon trigger. Because muon identification (MU1 and MU3) was included in the dimuon trigger, but not in the minimum bias trigger, we anticipate that the largest sources of uncertainty in this analysis (apart from the statistics) come from this trigger difference. In this chapter we give a detailed description of the detectors MU1 and MU3.

4.1 MU1

4.1.1 Description

The MU1 hodoscope bank (Hadron Veto) was located 2.3 meters downstream of the calorimeter behind 15 *cm* of lead (Fig. 2.4). The 45 counters were combined in a linear fan-in summing module, and a veto signal was produced if the sum exceeded a pre-determined pulse-height. The nominal threshold for the sum was such that a energy deposit equivalent to more than four minimum ionizing particles would cause an event to be rejected in the level one trigger. Thus events with hadrons showering in the calorimeter that deposit any energy in this detector were rejected and events with two muons should be accepted with some high efficiency. Unfortunately this turned out not to be the case.

4.1.2 MU1 Simulation

To calculate the dimuon acceptance of MU1 we must determine whether a given set of muons from a $K_L \rightarrow \mu^+\mu^-\gamma$ event would have set the veto bit in the trigger. As this is a statistical question, we must determine a distribution of energy deposits in the counter bank, and the probability that the trigger bit was then set. We do this by using a set of calibration muons to determine the thresholds for all of the counters individually. This data set consists of several runs in which a “beam-stop” was placed in the beam line. The beam-stop was an 18 foot long copper block of that was water cooled. This guaranteed that only muons could enter the detector region. In all we had 7 runs throughout the data taking period with a total of 2.5 million events where a single muon passed through the detector. The stability of the veto rate from MU1 is shown in Fig. 4.1. There is no evidence for shifts in the rate that are not consistent with statistical fluctuations. For each event the integrated charge of the pulse from each MU1 counter was recorded, as well as the trigger bit (i.e. whether the analog sum of the signals was above threshold).

With these beam-stop closed muons we could accurately determine the apertures of all of the detectors. We chose all the events with high momentum muons (>20 GeV) to minimize the multiple scattering in material upstream of MU1 and then looked at events in which a given counter recorded a hit (i.e. the integrated charge was above 3 ADC counts. Noise was typically only a few counts and a minimum ionizing particle deposited a most probable energy equivalent to 50 counts). Fig. 4.2 shows the efficiency for the counter as a function of distance along the counter. By differentiating this curve we can accurately define the edge of the counter. In this way the positions of all counters were measured (see Fig. 2.13).

Having found the positions of all the counters in the MU1 bank, we can now determine the thresholds by using the same set of beam-stop muons. We point the muons to a region well inside the fiducial area of each counter, and determine the

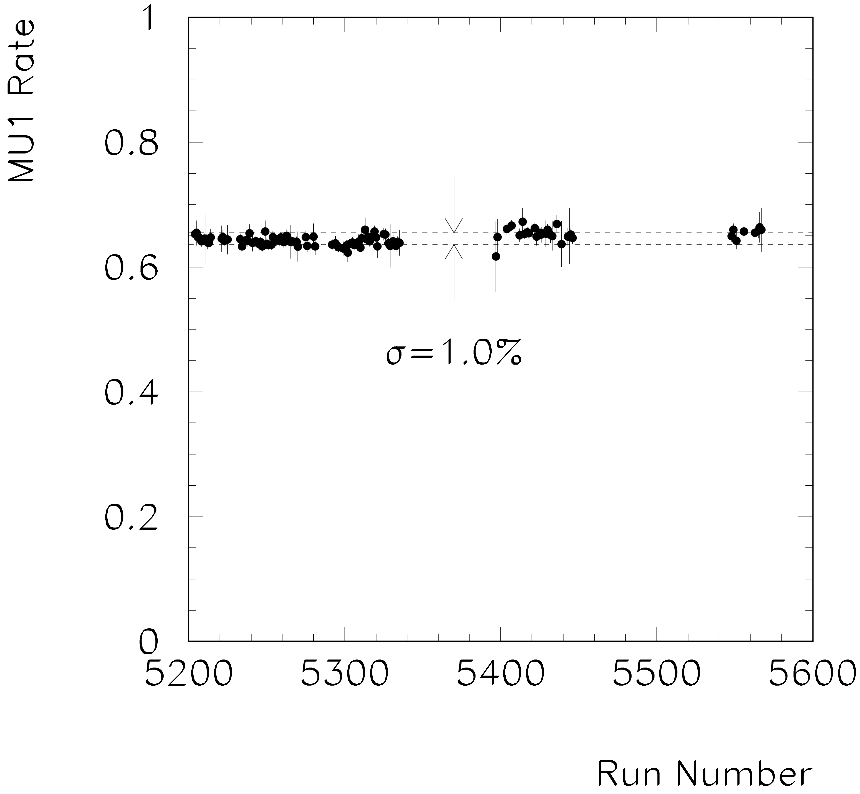

Figure 4.1: Variation of the MU1 veto rate as a function of run in the two track minimum bias data set.

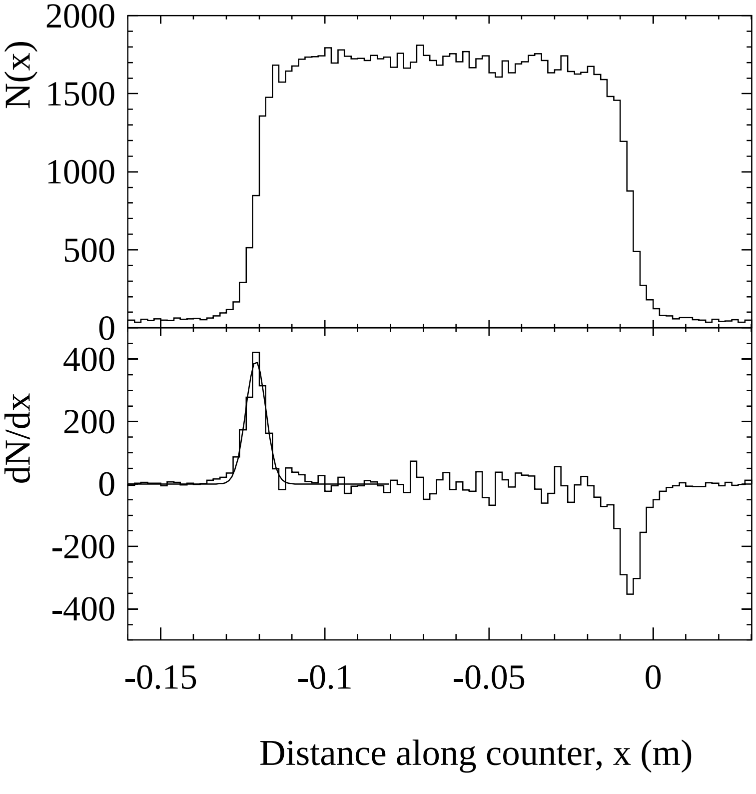

Figure 4.2: How the positions of the counter elements were determined.

distribution of pulse heights (Fig. 4.3). Also, we can determine the probability that a given pulse height will set the trigger bit. Examples of the pulse height distribution and the threshold curves for typical counters are shown in Fig. 4.4. These curves were generated in 8 bins along the length of each counter. Each counter has a different threshold due to the variation in gains of the PMT's on the counters.

In principle this is all the information necessary to simulate the MU1 hodoscope bank. We project tracks to the MU1 bank and determine which counter(s) they passed through. Then we draw randomly from the pulse height distributions shown in Fig. 4.3. If the system was linear we could divide each counter's simulated signal by its threshold, and the sum would determine the trigger bit. That is, we allow for the possibility that many counters can each produce a fraction of the threshold, but when summed they exceed it:

$$V = \sum_i \frac{(\text{pulse height})_i}{(\text{thresh})_i}. \tag{4.3}$$

If the quantity V is greater than 1 then the responses of all the counters would have summed to exceed the threshold, and the trigger veto bit would have been set.

However, the assumption of linearity is hard to justify, and probably not true. The assumption is essentially one about about pulse shape. When two muons pass through the bank at the same time the signals are added together and the resulting signal was required to be above a threshold. Also the individual signals were digitized

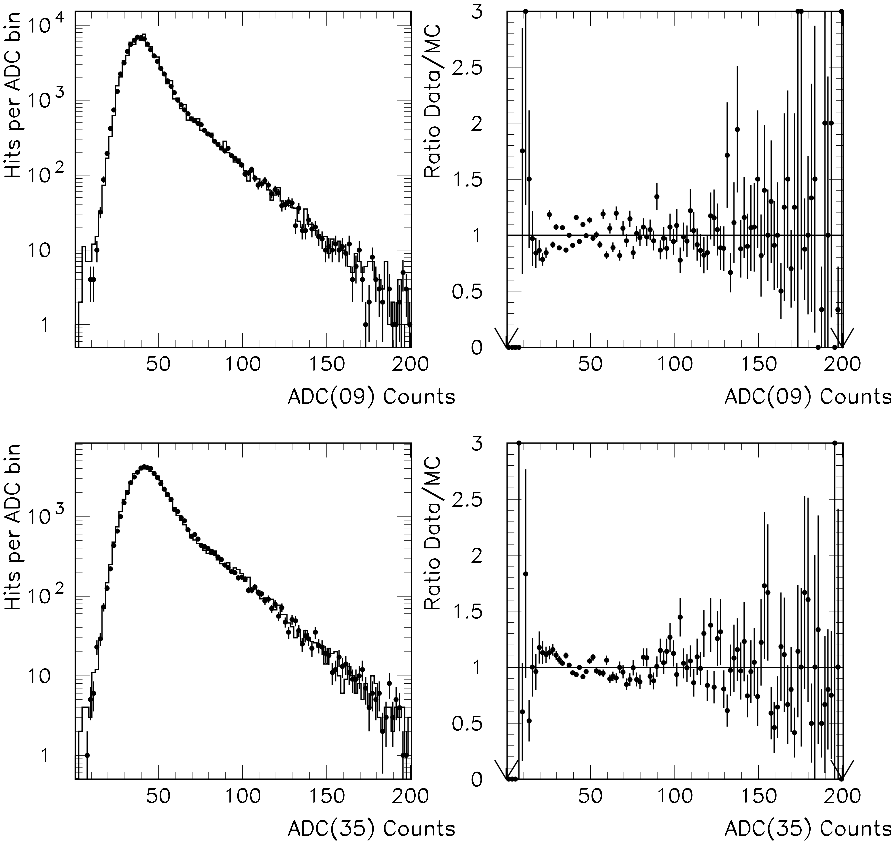

Figure 4.3: Simulated energy deposits in MU1 counter. Points are Monte Carlo histogram is data.

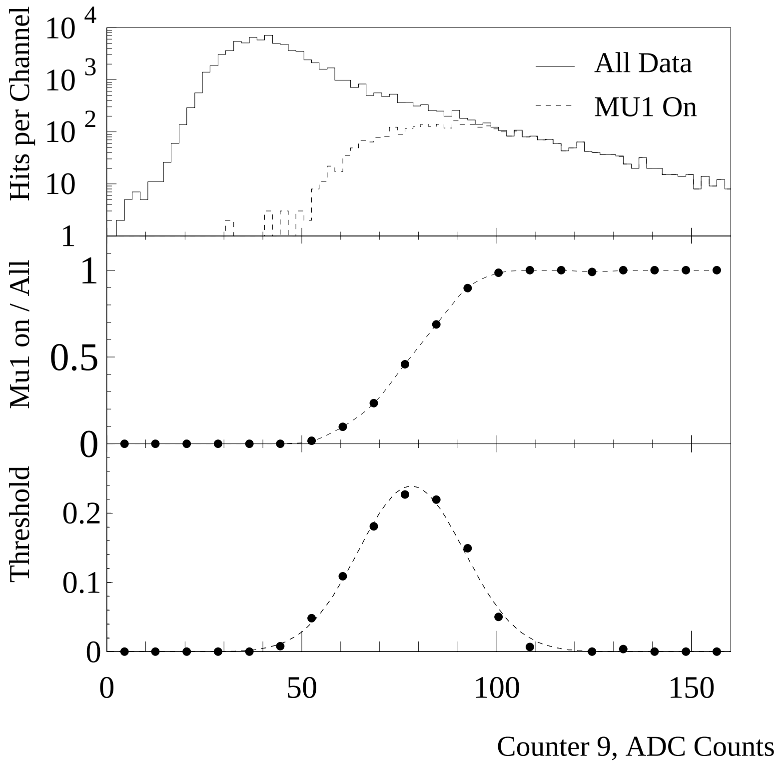

Figure 4.4: Determining the MU1 threshold for a counter. Top: Find all events where the MU1 trigger bit was set. Middle: take the ratio of the on/all to determine the efficiency as a function of ADC counts. Bottom: differentiate to find the width.

and written to tape. So the actual trigger bit was the result of a *level* threshold, but the digitized information is the charge *integral.* Two pulses that add together to exceed a level threshold may not exceed a charge integral threshold. This non-linearity between pulse height and charge integral is due to a pulse shape dependence.

The accidental environment in E799 produced many additional pulses per bucket in the MU1 hodoscope bank. The pulse height distribution of these pulses is shown in Fig.4.5. This means that to each event there is some energy added with a pulse shape much different from the minimum ionizing deposit in its time profile. This is illustrated in Fig. 4.6 where we have examined the quantity V from Eqn. 4.3 for two cases. In the first case we have taken the beam stop closed muon sample, and plotted V for all data, using the actual ADC values as well as the simulated ones. We see that there is a significant disagreement at high values of V, which we ascribe to the existence of accidental activity in the counters. To illustrate this point we note that the agreement gets much better when we require that only one of the MU1 counters be hit. Fig.4.7 shows the same plots for a sample of beam-stop closed dimuons, and the agreement is somewhat better meaning that the accidental environment affects the veto rate of MU1 less for 2-mip deposits. We note that this accidental contamination is significant even in this relatively clean environment where the beam-stop was closed, and therefore all of the activity usually associated with the beam passing through the detector region is absent. The effect can only

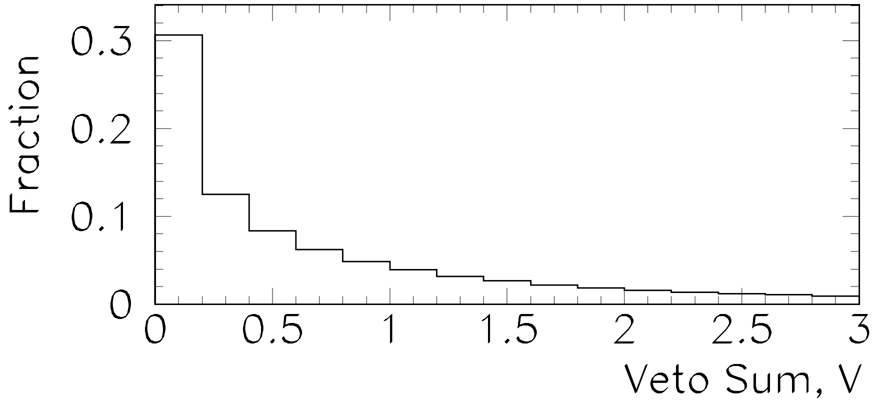

Figure 4.5: Contribution to the MU1 threshold from accidental events. The histogram is normalized to 1, so 30% of the accidental events lie in the first bin.

get larger in the beam-open data taking conditions.

Thus accidental activity plays a significant role in the rate at which MU1 rejects events. It is worth noting that in raw accidental events the veto is on in 15% of all events independently of whether a kaon decayed in the detector. Furthermore the timing of the accidental pulses relative to the leading edge of a real muon deposit is not well defined, and the pulse shape is known to be different. It then becomes impossible to overlay this accidental activity onto real dimuon events, and to predict the efficiency of this detector on a per-event basis for the Monte Carlo simulation as we have no way of re-creating the actual threshold level of the pulses. We must look for a different approach to determining the acceptance of the MU1 hodoscope

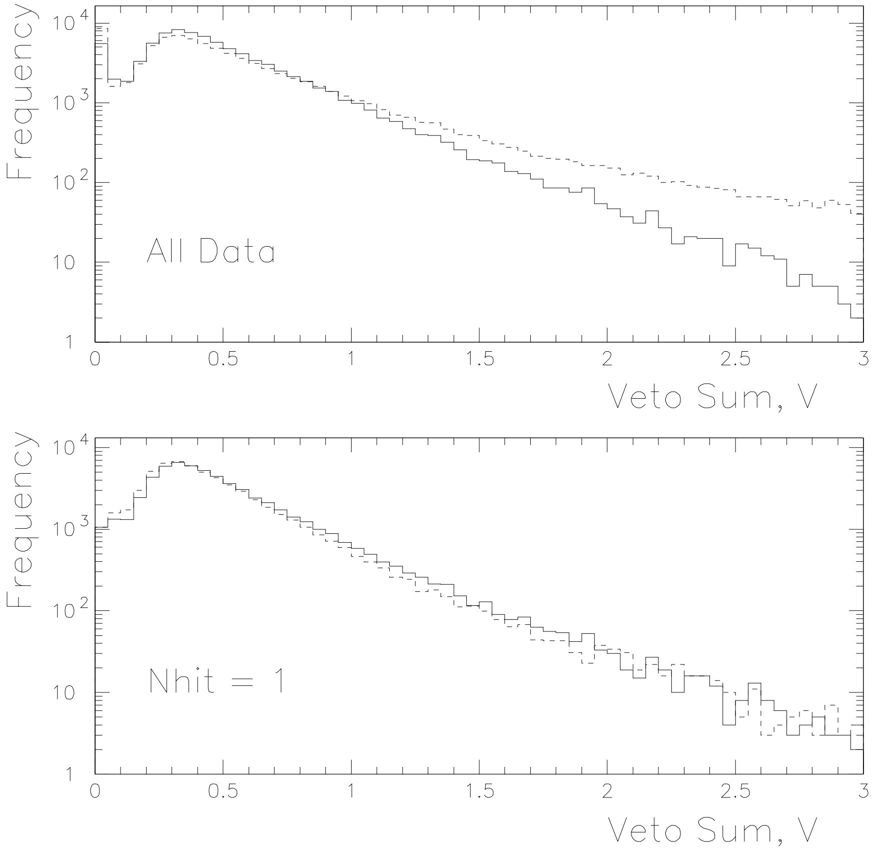

Figure 4.6: Comparison of MU1 veto rate V between data and Monte Carlo for single beam stop closed muons. Top: all data. Bottom: data with requirement that only one counter was hit. The solid line is using simulated ADC values, the dashed line uses data ADC values.

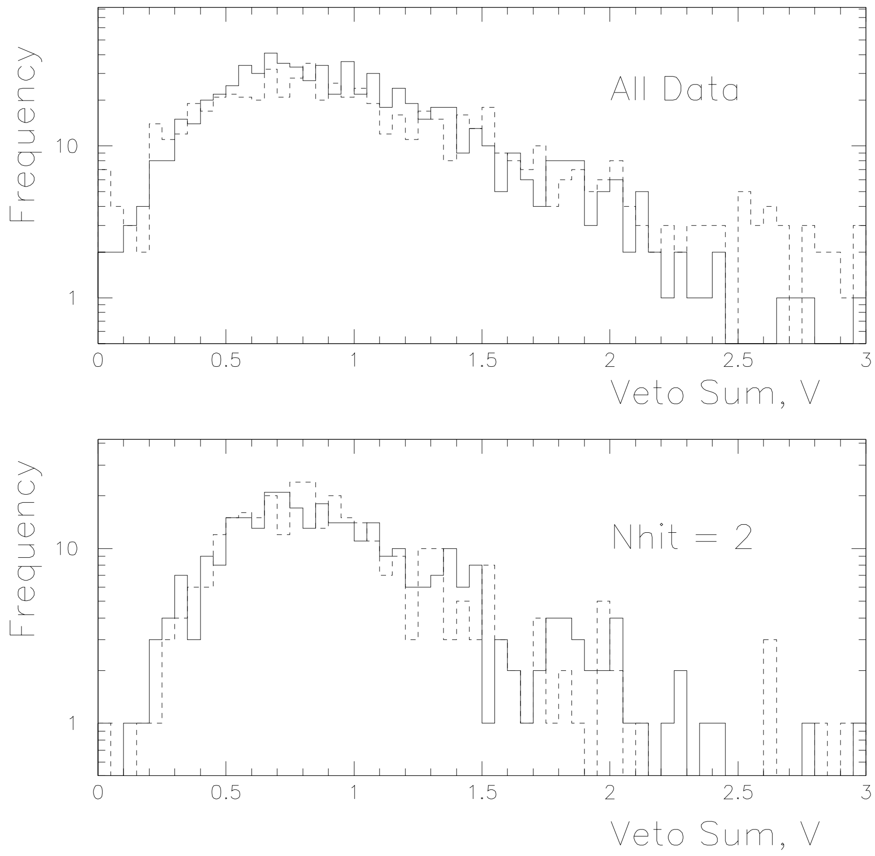

Figure 4.7: Comparison of MU1 veto rate V between data and Monte Carlo for two in-time beam stop closed muons. Top: all data. Bottom: data with requirement that only two counters were hit. The solid line is using simulated ADC values, the dashed line uses data ADC values.

bank.

4.1.3 Upper and Lower Bounds on the Efficiency of MU1

In the previous section it was shown that predicting the veto of the MU1 hodoscope bank on an event-by-event basis was not possible. We will now argue that its is possible to place upper and lower bounds on the efficiency of MU1 for rejecting events with muons in the decay products. To do this we firstly note that the data sample of $K_L \rightarrow \mu^+\mu^-\gamma$ events naturally divides into separate samples: those in which one of muons passes through the hole in the center of the counter, and those in which both tracks pass through counters. The rate at which events are rejected is much different for these two samples.

For single muons passing through a counter we showed in the previous section that if we ignore all other counters then we predict the MU1 veto rate reasonably well. The total MU1 rate will be

$$\mathrm{P(MU1)} = \mathrm{P(MIP}\ \mathit{or}\ \ \mathrm{ACC)} + \mathrm{P(MIP + ACC)} \tag{4.4}$$

P(MU1) is the probability that in any event the MU1 trigger bit is set. P(MIP *or* ACC) is the probability that the accidental event or the muon sets the bit (*i.e.* P(MIP)+P(ACC)−P(MIP)×P(ACC)). P(MIP+ACC) is the probability that the accidental energy combined with the minimum ionizing deposit set MU1. The technique we adopt is to ignore the last term in Eqn. 4.4: it is positive and so we have a lower limit on the MU1 veto rate. We find for single muons passing through the MU1 bank P(MU1)>22.1%, and for two muons P(MU1)>43.7%. (Note: this includes accidental energy deposits which cause the MU1 trigger bit to be set 15% of the time.)

To find upper bounds on the MU1 muon veto rate for the case when two muons passed through the hodoscope bank, we make use of data from the two track minimum bias trigger: $2B \cdot 2C \cdot \mathrm{P\bar{H}V} \cdot \mathrm{HiTCnt(2track)}$. The events on the minimum bias data tapes were dominated by $K_L \rightarrow \pi^{\pm}\mu^{\mp}\nu$ and $K_L \rightarrow \pi^{+}\pi^{-}\pi^{0}$. We assume a pion will deposit an average energy in the MU1 bank larger than a muon whether it showers before reaching the MU1 bank or not. The MU1 veto rate for this sample of minimum bias events will always contain at least one pion and hence we obtain an upper limit to the rate at which MU1 vetoes dimuon events. With this data we then made a series of cuts designed to enhance the fraction of events in which a pion decayed to a muon and a neutrino. The cuts also select events in which the pion did not shower in the lead glass. The cuts were: two good tracks, cluster energy below 1.5 GeV, no extra clusters in the lead glass, at least one MU2 counter hit, at least one MU3 counter hit, tracks had to project within the outer boundaries of MU1. After these cuts we then select events where there is a mismatch of the tracks at the reconstructed vertex: $\chi^2 > 10$ (Fig.4.8). With this set of cuts we determined

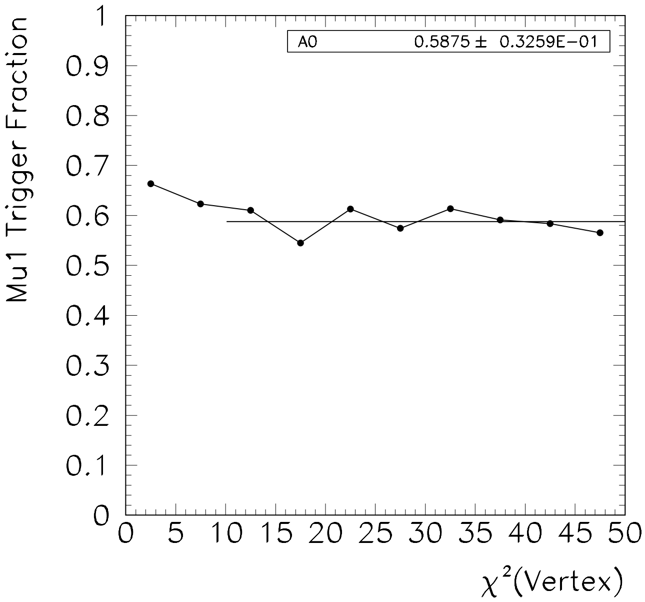

Figure 4.8: Determination of MU1 veto rate from two-track minimum bias data where both tracks extrapolate to within the area of MU1. The rate is plotted against a cut made on the vertex χ^2. The rate for two muons is determined to be below $58.8 \pm 3.3\%$.

the MU1 veto rate to be $58.8 \pm 3.3\%$ for the case when both tracks extrapolate to a region inside the MU1 bank. The uncertainty derives from the statistics of the sample after cuts. For single tracks the same quantity was $28.5 \pm 3.7\%$, however there was a better calibration data set as we now discuss.

For dimuon events in the case when one of the muons passes through the hole in the MU1 bank, we have another data set to use in calibration. This is the so-

Method	1 Mip	2 Mip
Monte Carlo	22.1%	43.7%
$1B \cdot MU2$	$25.4 \pm 1.0\%$	—
2-track Min Bias	$28.5 \pm 3.7\%$	$58.8 \pm 3.3\%$

Table 4.1: Summary of MU1 rejection rates.

called $B \cdot \mu$ triggers: $1B \cdot MU2$. With this data set we also made a series of cuts designed to reduce the pion contamination: the standard tracking cuts; no hits in the photon veto counters; less than 1.5 GeV deposited in the cluster associated with the track in the calorimeter; at least one hit in MU3; only one or two hits in MU2; track momentum greater than 7 GeV. For the events surviving these cuts the MU1 rate was $25.4 \pm 1.0\%$. Again there is contamination from pions in this sample, so this is an upper limit for the rate at which the MU1 trigger bit is set for the case when one muon passes through the MU1 bank. We summarize the MU1 rejection rates in Tab 4.1.

4.2 MU3

4.2.1 Description

MU3 consists of 16 vertical, non-overlapping scintillator slats located 30 *cm* behind the filter steel (see Fig. 2.14). Each slat, approximately 10 *cm* wide, was glued to a light guide at the top and viewed by a PMT. The signals were fed to a logic unit that applied the following condition:

$$(N_{MU3} \geq 2) \text{ and } (\text{If } N_{MU3} = 2 \text{ then } |i - j| > 1) \tag{4.5}$$

where N_{MU3} = number of counters above threshold. The “non-adjacency condition”, $|i - j| > 1$, is designed to reject single muons that emit δ-rays within ~1 *cm* of emerging from the filter steel. Ideally one would solve this problem by placing the hodoscope bank as close to the rear face of the steel as possible, but due to the irregular shape of the filter steel this was not possible.

4.2.2 MU3 Simulation

The fraction of $K_L \rightarrow \mu^+\mu^-\gamma$ events that are rejected by the MU3 hodoscope bank and would have otherwise been accepted is an important quantity for this analysis. We simulate MU3 by using the same set of beam stop muons that were used for MU1. The same method is used to map out the edges of the counters, by again pointing high momentum muons towards MU3 and asking when the counter latch was set. This time however the counters were all the same size, and were measured precisely. The wrapping tape used produced gaps between the counters that amounted to 1% of the total width. We therefore just used the high momentum muons to define the outer edges laterally, and the bottom edge of this bank. We assume that the counters are equally spaced, and place them within this region (see Fig. 2.14).

To calculate the efficiencies of the 16 counters we divided them into 5 bins along their length. Figure 4.9 shows the efficiencies of the MU3 latches for a representative sample of the 16 counters. Mostly the counters were 99% efficient, with same poorer quality ones being placed near the edges. These efficiencies were put into the Monte Carlo. The uncertainties on the counter efficiencies were typically less than 1% and due to statistics of the calibration muon sample.

4.2.3 MU3 Efficiencies

Again the stability of this bank is important as we assume the acceptance of MU3 is constant throughout the run. Figure 4.10 shows the number of MU3 latches firing throughout the run in the two-track minimum bias data. The rate is stable, and there is no evidence of any drift. We infer that the uncertainty on the average efficiency of MU3 is about 0.1%. The effect of adding MU3 to the trigger in a $K_L \rightarrow \mu^+\mu^-\gamma$ Monte Carlo sample of events is to reduce the sample after analysis cuts by 65.3% (this comes from an 82% geometrical acceptance and a 79% acceptance from the non-adjacency logic).

To demonstrate the sensitivity of the acceptance to the understanding of counter efficiencies we generated a Monte Carlo sample in which we applied random shifts in the counter efficiencies by their uncertainties on an event-by-event basis. This led to 2.5% change in the acceptance of MU3. Reducing the efficiency of

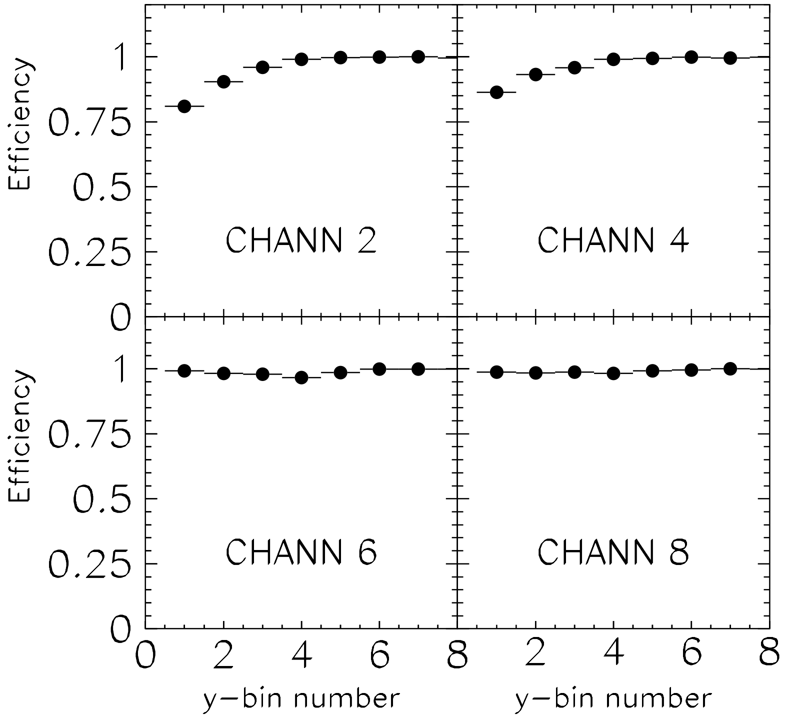

Figure 4.9: Sample of MU3 efficiencies for four counters. The PMT is at the right is these figures. Muons with high track momentum (greater than 20 GeV) were selected, and tracks were required to extrapolate within 1 cm of the counter boundaries.

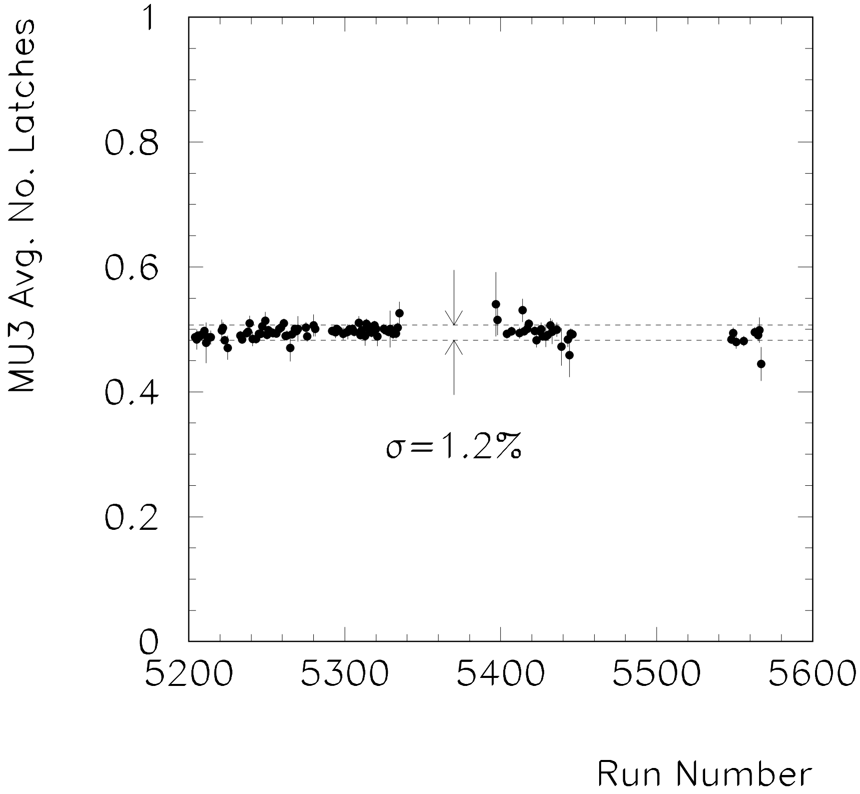

Figure 4.10: Variation of the number of MU3 latches set as a function of run in the two track minimum bias data set.

just the central counter by 10% caused the acceptance to reduce by less than 0.5%. We conclude that if for some reason the efficiencies of the MU3 counters moved within a range that is consistent with the calibration data the largest change in the acceptance (and hence the BR($K_L \rightarrow \mu^+\mu^-\gamma$)) that could be induced is 2.5%

In addition we studied the efficiency of the trigger logic to give the correct result as a function the input to the unit. For this study we used a sample of 2000 events in the beam-stop closed data where there were two tracks present. We looked for cases where the MU3 trigger logic had been satisfied in this sample. In 1% of these events the latched signals were inconsistent with with the formation of a MU3 trigger, for example there were 2 latches set directly adjacent to one another – a condition that should be impossible under the MU3 logic scheme. There were also cases where two non-adjacent latches in the MU3 bank were set – a condition that should have set the MU3 logic to true – but the MU3 logic unit was set to false. This also amount to 1% of the sample. This behavior is attributed to small timing differences at the input to the logic unit that may, for example, cause an out-of-time deposit in the hodoscope bank to be included in the logic decision. The Monte Carlo simulation of the hodoscope bank treats the signals as if they all arrived simultaneously. We conclude that events may enter the sample or be removed from the sample independently and we add the uncertainties in quadrature to give 1.7%. The MU3 acceptance is then 65.3±1.9%. We summarize the MU3 study in Tab. 4.2

Efficiency ϵ	65.3%
$\Delta\epsilon/\epsilon$ smearing	± 2.5
$\Delta\epsilon/\epsilon$ trigger logic	$\pm 1.4\%$
Total	$65.3 \pm 1.9\%$

Table 4.2: Summary of MU3 study.

Chapter 5

Calculation of the Branching Ratio

In this chapter we discuss the search for $K_L \rightarrow \mu^+\mu^-\gamma$ candidate events emphasizing the the kinematical variables used to isolate the signal. The cuts to define a well reconstructed sample of $K_L \rightarrow \pi^+\pi^-\pi^0$ normalization events are discussed.

5.1 Common Cuts

There are a large number of cuts that are common to both the $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow \pi^+\pi^-\pi^0$ data samples, these are listed in Tab. 5.1. Most of these are related to tracking as the normalization decay mode is virtually identical to the signal decay mode in this regard. The distributions of tracks for the two modes are very similar and therefore it is assumed that there are minimal biases associated with these cuts.

Removing events from runs where there was impure gas in the drift chambers was necessary to exclude data associated with one contaminated shipment of Argon-Ethane cylinders.

The tracking search has been described in a previous chapter. After two clean segments are found in the x and y views they are required to match clusters in the lead-glass. To be consistent with the decays $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow \pi^+\pi^-\pi^0$ the tracks must have opposite signs. A loose cut is made on the x-offset in the extrapolated positions of upstream and downstream track segments at the plane of the magnet (Fig. 5.1).

A source of potential disagreement is the fact that multiple scattering was not implemented in the Monte Carlo in the material between the drift chambers. This is seen in Fig. 5.2, where the lack of tails on the χ^2 distributions for track segments in the Monte Carlo evident. However as we make identical cuts in both signal and normalization samples it is more relevant to compare the same distributions between the $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow \pi^+\pi^-\pi^0$ Monte Carlo data samples. We then see that the difference is insignificant in the ratio of acceptances (Fig. 5.3).

To be consistent with having come from the decay of a single kaon, the two track candidates must form a good vertex. A loose cut was made on the χ^2 of a fit to the upstream track segments under the assumption of a common vertex (Fig. 5.4). The actual cut was momentum dependent, as multiple scattering caused the resolution to deteriorate, but it amounted to a 5σ cut at most energies. Another

No.	Cut	MC Efficiencies	
		Sig	Norm
C1	Exclude runs where chamber gas was bad or where a chamber wire not functioning	0.05	0.05
C2	Require two x and two y tracks in drift chambers	0.93	0.92
C3	Require hits in drift chambers be consistent with two tracks	0.97	0.98
C4	Require good match of x and y tracks and at least one good cluster in the calorimeter associated with a track	0.96	0.95
C5	Require tracks bend in opposite directions	0.99	0.99
C6	Drift Chamber: (i) good matching of upstream and downstream segments (1 mm), (ii) Each segment required to have good χ^2 when fit to a straight line, (iii) distance of closest approach <3 mm, (iv) vertex $\chi^2 < 10$	0.85	0.78
C7	Require all clusters be at least 0.5 blocks away from the beamholes in the lead glass	0.83	0.78
C8	Photons must not be near iron ring in front of calorimeter	0.99	0.99
C9	Tracks and photons must be away from VA0 aperture (2 cm)	0.95	0.90

Table 5.1: Cuts common to $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow \pi^+\pi^-\pi^0$ decay modes.

test of the single vertex hypothesis is the "distance of closest approach" of the upstream segments shown in Fig. 5.5.

All these tracking cuts have relative high efficiency, and combined about 80% of Monte Carlo events pass these requirements.

It was well known from previous studies that the lead glass blocks nearest the beamholes in the calorimeter suffered the worst radiation damage, and their response was very poor near the innermost regions [52, 54]. Therefore all events with energy deposits within 0.5 blocks of the holes were removed to avoid any possible bias between data and Monte Carlo. Any bias in acceptance is negligible as this is a position cut and Figs. 3.6 and 3.7 indicate the cluster position reconstruction is well understood in the Monte Carlo.

The iron ring placed upstream of the calorimeter converted some photons and the resulting e^+e^- pairs were subsequently detected. The iron ring was not part of the Monte Carlo geometry. This can be seen in Fig. 5.6 where there is a deficit of photons in $K_L \to \pi^+\pi^-\pi^0$ decays at radii greater than 0.84 m when compared to a Monte Carlo sample. To prevent any bias, events with photons that had radii greater than 0.835 m at the position of the iron ring were removed from the data set. This reduced the acceptance by 1%.

Lastly we removed all events from the sample where either a track or photon passed within 2 cm of the VA0 aperture. There was an unknown amount of material

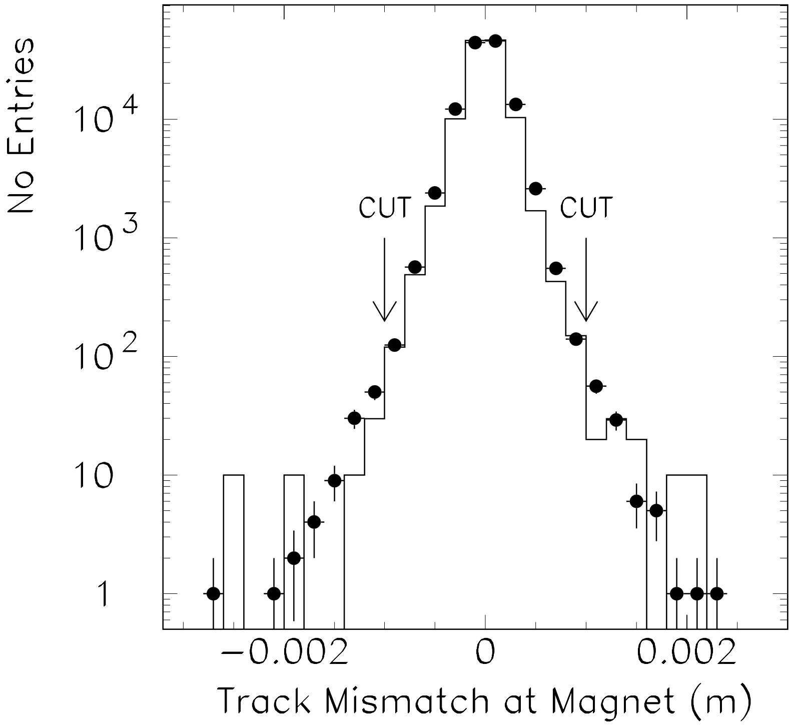

Figure 5.1: The separation between upstream and downstream track segments at the center plane of the magnet. The dots are Monte Carlo generated events and the line is data.

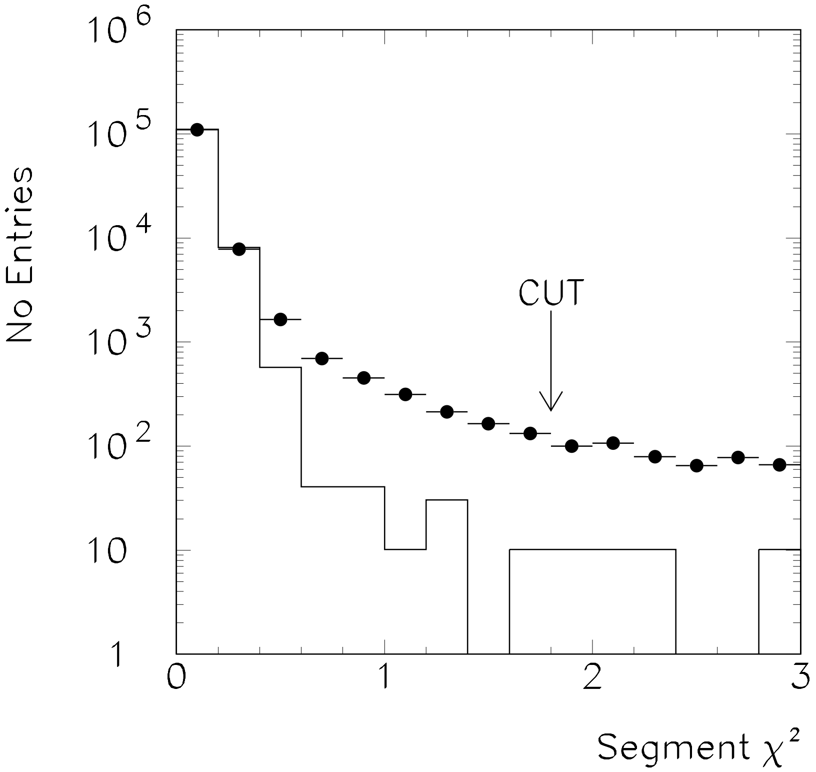

Figure 5.2: Comparison of track y-segment χ^2 between Monte Carlo (dots) and data (line). The arrow indicates the analysis cut.

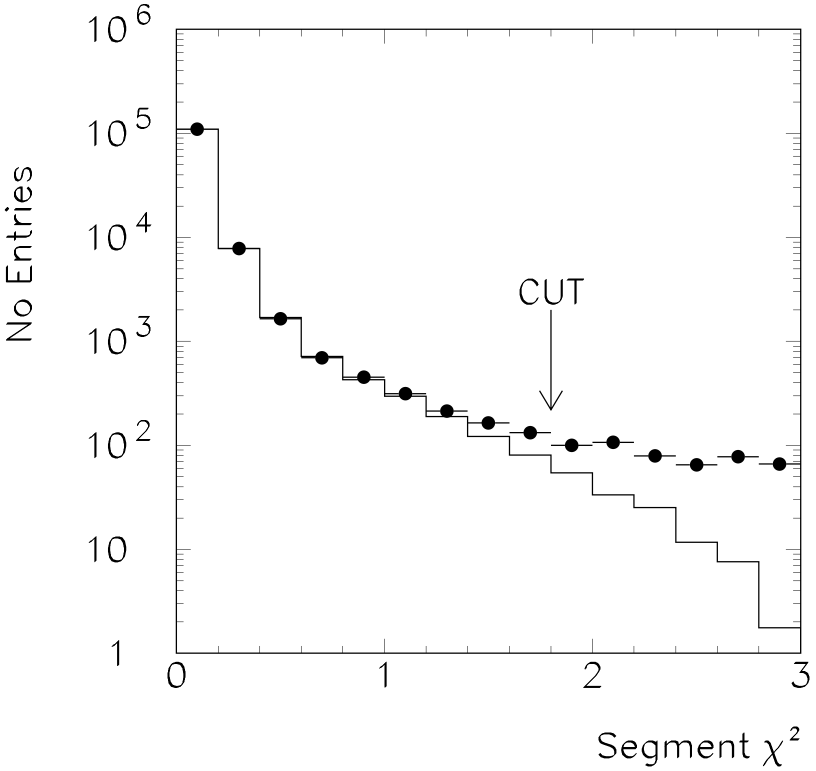

Figure 5.3: Comparison of track y-segment χ^2 between $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo (dots) $K_L \to \mu^+\mu^-\gamma$ Monte Carlo (line). The arrow indicates the analysis cut.

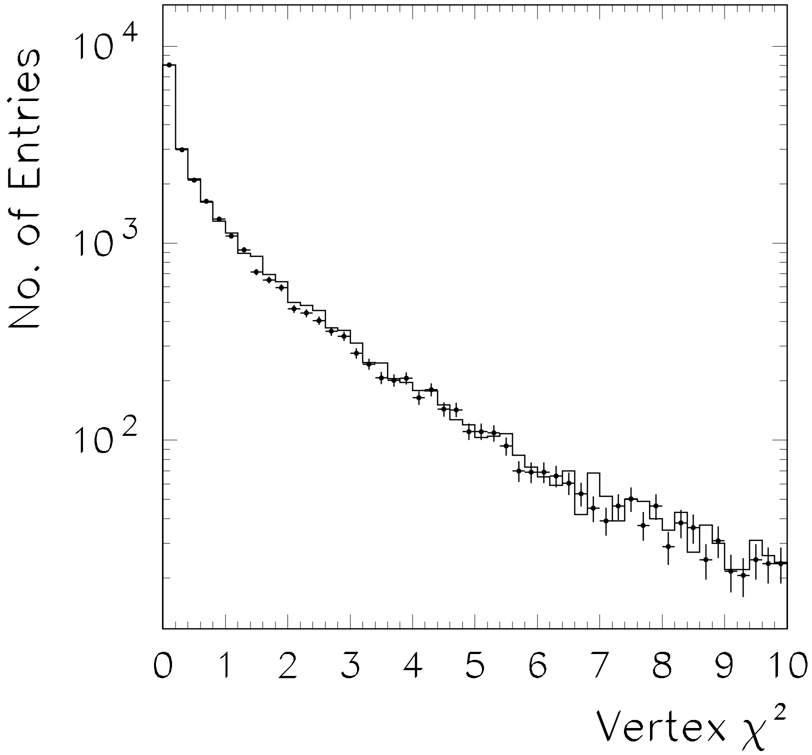

Figure 5.4: The agreement between Monte carlo (dots) and data (line) in the reconstructed vertex. In the analysis a cut at 10 was made.

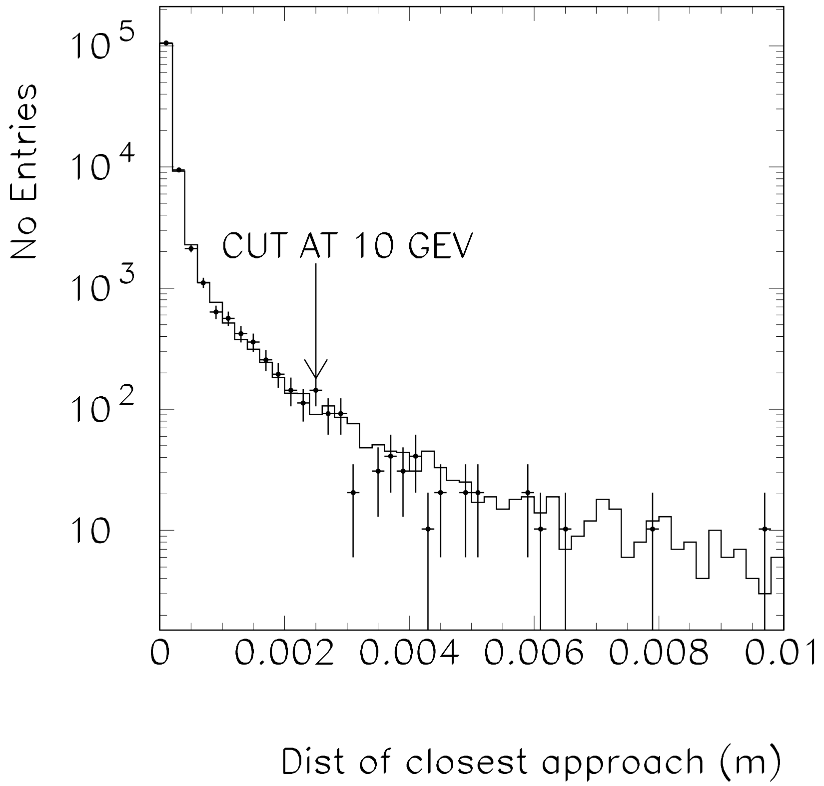

Figure 5.5: Distance of closest approach of the two upstream track segments. Dots are Monte Carlo and line is data.

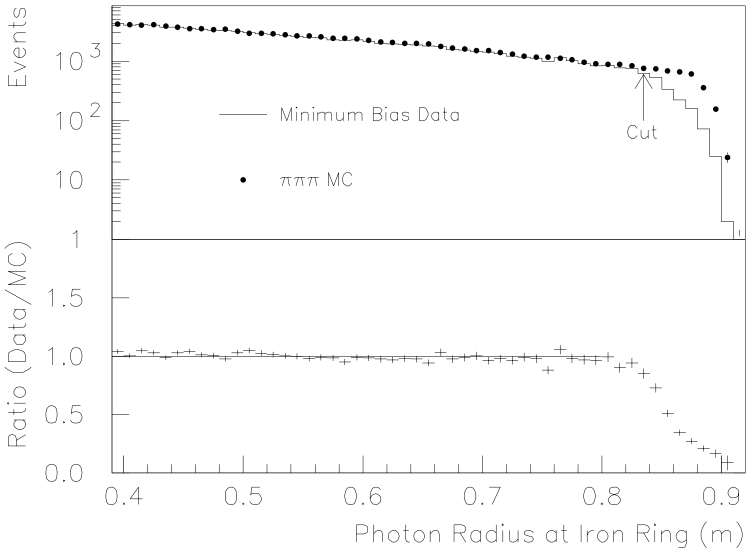

Figure 5.6: The distribution of photon radii in $K_L \rightarrow \pi^+\pi^-\pi^0$ decays for both Monte Carlo and data at the z-position of the iron ring. [54]

used in the construction of the inner edge of this detector, and this was not present in the Monte Carlo geometry. As seen in Fig. 5.7 there is a fall-off in the photon distribution in the data that is not present in the Monte Carlo, indicating the effect of the material. This cut removes possible biases due to the uncertainty in the VA0 aperture.

5.2 Backgrounds

Background identification is the next important step in trying to isolate the signal events. Before making cuts on kinematical quantities it is helpful to know the

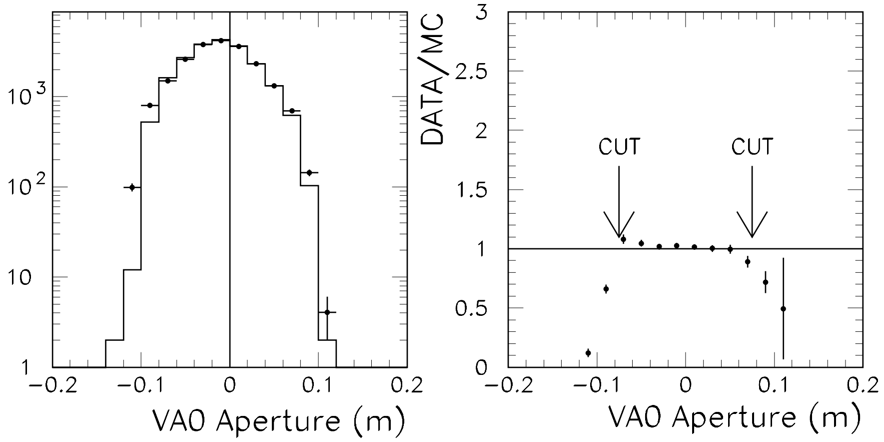

Figure 5.7: Cut on the position of the photons at the aperture of the veto counter VA0. Points are Monte Carlo and line is data.

properties of the events that are contaminating the signal region of parameter space.

A useful tool for identifying background is a "missing longitudinal momentum parameter", P_0 [59]. This is defined by assuming the event is $K_L \to \pi^+\pi^-\pi^0$ and assigning the pion mass to the two tracks. The π^0 longitudinal momentum would then be

$$P_0^2 = \frac{(M_K^2 - M_{+-}^2 - M_0^2)^2 - 4M_{+-}^2 M_0^2 - 4M_K^2 \mathcal{P}_t^2}{4(\mathcal{P}_t^2 + M_{+-}^2)} \tag{5.1}$$

where M_K is the kaon mass, M_{+-} is the invariant mass of charged tracks, M_0 is the π^0 mass and $\mathcal{P}_t$ is the transverse momentum of two tracks. An event consistent with $K_L \to \pi^+\pi^-\pi^0$ will have $P_0^2 > 0$ by definition. Only tracking information is used to obtain P_0. When this parameter is plotted versus invariant $\mu^+\mu^-\gamma$ mass for the data we see clearly three regions of population, as shown in Fig. 5.8. We developed Monte Carlo simulations of the possible two track kaon decays that could fake a $K_L \rightarrow \mu^+\mu^-\gamma$ decay: $K_L \rightarrow \pi^+\pi^-\pi^0$, $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$, $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{rad}$. Here γ_{rad} indicates a radiative photon and γ_{acc} indicates a spurious cluster in the calorimeter not associated with the kaon decay. In both $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma$ decays the pion is also misidentified as a muon, either because it doesn't shower in the calorimeter, or it decays in flight, producing the second muon, $\pi^\pm \rightarrow \mu^\mp\nu$. In the case of $K_L \rightarrow \pi^+\pi^-\pi^0$, one of the photons can overlap a track cluster or go through the beamholes in the calorimeter, and both pions must be misidentified as muons either because a pion decays, or it doesn't shower in the calorimeter and an accidental hit occurs in the MU3. Figure 5.9 shows the same plot for the candidate background Monte Carlo samples, and also for the $K_L \rightarrow \mu^+\mu^-\gamma$ Monte Carlo sample. We see that $K_L \rightarrow \pi^+\pi^-\pi^0$ reconstructs to an invariant mass far below the $\mu^+\mu^-\gamma$ mass, and is therefore not a problem. We estimate less than 0.1 of an event falls in the $K_L \rightarrow \mu^+\mu^-\gamma$ signal box region (defined below). The $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{rad}$ decay also produces very few events near the signal peak. After generating background Monte Carlo events that corresponded to the statistics of 1/2 of the run, the only potential background where any events were seen close to the $\mu^+\mu^-\gamma$ sample was $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$. Significant $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{rad}$ contributions are inconsistent with the data: by normalizing the Monte Carlo to the data sample we estimate that over 90% of the background is $K_L \to \pi^{\pm}\mu^{\mp}\nu\gamma_{acc}$. Of this background the Monte Carlo simulation showed that only 5% enters the trigger because of pion decay. The dominant contribution comes from accidental activity in the MU3 hodoscope bank that correlates with the extrapolated track of the pion.

5.3 $K_L \to \mu^+\mu^-\gamma$ Analysis

The analysis cuts for $K_L \to \mu^+\mu^-\gamma$ are listed in Tab. 5.2. Only events with exactly one lead glass cluster not associated with either track were considered. The Monte Carlo acceptance for this requirement was 95%. Since MU1 was in the trigger, and part of the estimate of the efficiency of that detector involved a simulation of the MU1 response, the tracks were required to extrapolate to within the area of MU1. Although the effect on the acceptance was minimal this cut was imposed for consistency. The three meters of steel downstream of the calorimeter (see Fig. 2.4) was designed to filter out all particles except muons. This represented 18 interaction lengths which made punch-through negligible. However there was a threshold of 5 GeV for penetration by muons. There is some straggling, and to be 95% likely that a muon will penetrate the steel, we required that all candidate muon tracks be above 7 GeV.

All potential sources of background contained charged pions, which typically

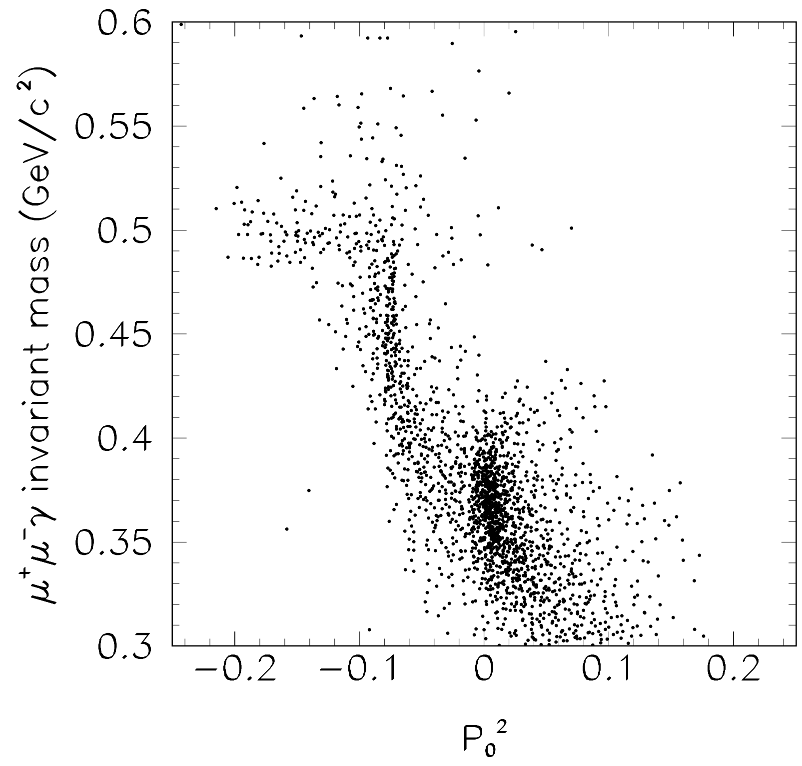

Figure 5.8: Invariant $\mu^+\mu^-\gamma$ mass versus P_0^2. Equation 5.1 defines P_0^2.

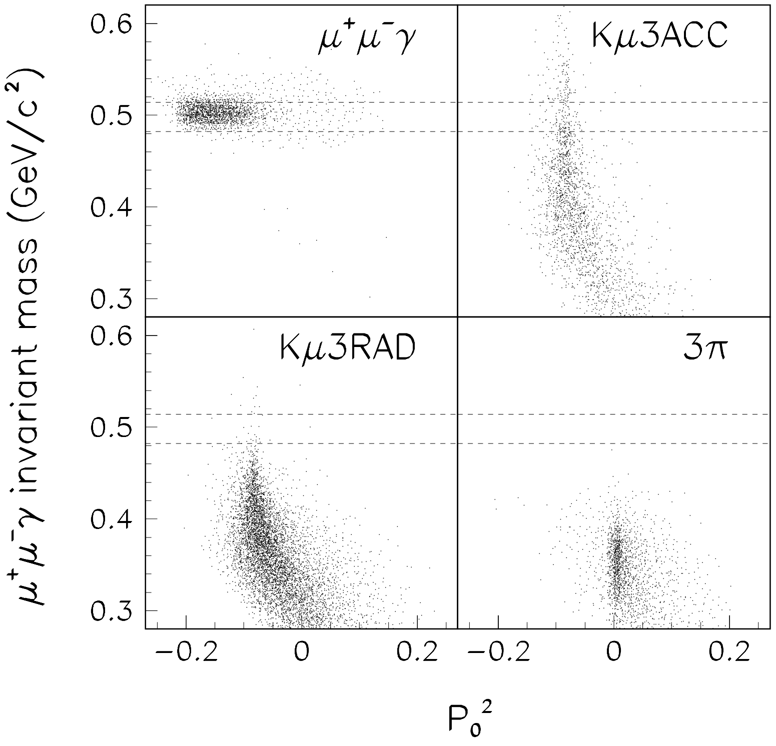

Figure 5.9: Invariant $\mu^+\mu^-\gamma$ mass versus P_0^2 for various Monte Carlo samples, as indicated. Comparing this to Fig. 5.8 clearly identifies the background as $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$. For reference the analysis cut on invariant $\mu^+\mu^-\gamma$ mass is indicated by dashed lines.

No.	Cut	Efficiencies		
		Data	MC $\mu^+\mu^-\gamma$	MC Kμ3γ_{acc}
S1	One extra in time cluster in calorimeter	0.95	0.95	0.63
S2	Tracks must extrapolate to within MU1	0.99	0.99	0.99
S3	$7 < \lvert\vec{p}_{track}\rvert < 99 \; GeV$	0.95	0.96	0.94
S4	Track Cluster Energy below 3.0 GeV	0.74	0.97	0.89
S5	Extra cluster energy $> 8 \; GeV$	0.91	0.91	0.43
S6	No. fusions of clusters using cluster fusion algorithm	0.71	0.91	0.47
S7	Track cluster - extra cluster separation $> 20 \;\; cm$	0.90	0.89	0.77
S8	Adder ratio > 0.75	0.97	0.99	0.92
S9	No. B-bank + No. C-bank latches < 12	0.94	0.99	0.91
S10	$P_T^2 < 200 \;\; MeV^2/c^2$	0.38	0.91	0.05
S11	$M_{\lvert\mu^+\mu^-\gamma} - M_K\rvert < 17 \;\; MeV$	0.01	0.86	0.11

Table 5.2: $K_L \rightarrow \mu^+\mu^-\gamma$ analysis cuts.

deposit large amounts of energy in the calorimeter when the pion showers hadronically. Figure 5.10 shows energy deposits in the calorimeter from the tracks in $K_L \rightarrow \pi^+\pi^-\pi^0$ decays are very different from minimum ionizing tracks. Events with track cluster energies above 3 GeV were removed from the sample. The number of minimum ionizing tracks rejected with this cut was 3%.

Figure 5.11 shows the distribution of photon energies for $K_L \rightarrow \mu^+\mu^-\gamma$ Monte Carlo events and also for the accidental cluster in $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$ Monte Carlo events. In the $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$ sample the isolated energy deposits are typically lower in energy than the photon energy in $K_L \rightarrow \mu^+\mu^-\gamma$ events. We removed events from the data set with extra cluster energies below 8 GeV. This removed 57% of

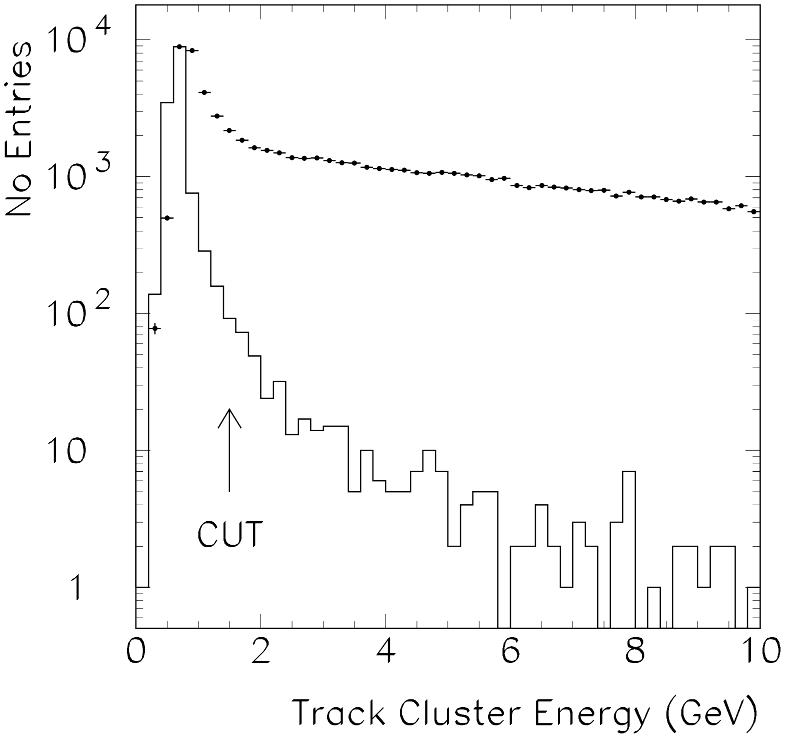

Figure 5.10: Track cluster energies. The line is the deposit of minimum ionizing particles in the lead glass. The points are from $K_L \rightarrow \pi^+\pi^-\pi^0$ tracks.

the $K_L \to \pi^{\pm}\mu^{\mp}\nu\gamma_{acc}$ Monte Carlo events, and 9% of $K_L \to \mu^+\mu^-\gamma$ events. Also, the Monte Carlo is known to do rather poorly for energy deposits below 7 GeV, and cutting away from that region removes this known bias. (This is because of residual non-linearities in the lead glass simulation; the radiation dosage of lead glass blocks near the pipe-blocks caused their attenuation of Cerenkov light to be much worse. A correction was made for this, however a small disagreement remained at low energies.)

The fusion finding algorithm identified almost half the photon clusters in the $K_L \to \pi^{\pm}\mu^{\mp}\nu\gamma_{acc}$ Monte Carlo as being inconsistent with energy deposits from single photons. This is compared to the $K_L \to \mu^+\mu^-\gamma$ Monte Carlo events for which only 9% are similarly identified. We therefore removed all events from the data sample that failed the fusion finding test.

The distance between the track positions and the extra cluster at the calorimeter was a found to be significantly different in background and signal Monte Carlo samples. Tracks tend to peak near the center of the detector, laterally, as do false candidate photon clusters. On the other hand the structure of the $K_L \to \mu^+\mu^-\gamma$ decay forces the photon to be somewhat divergent from the tracks (in the center of mass frame the photon and the sum of the momenta of the two tracks are back to back). This correlation allows us to further remove background, at the same time keeping the signal acceptance high. As shown in Fig. 5.12, if we demand that the

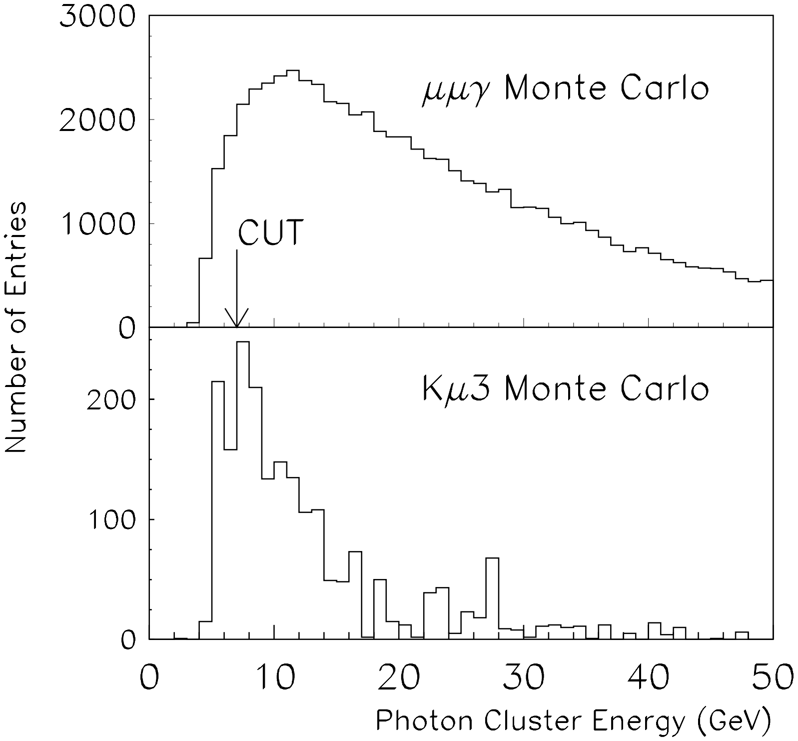

Figure 5.11: Distribution of photon cluster energies for $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow \pi^\pm\mu^\mp\nu\gamma_{acc}$ Monte Carlo events.

track-photon cluster separation be greater than 20 *cm* we can remove 33% of Monte Carlo $K_L \rightarrow \pi^{\pm}\mu^{\mp}\nu\gamma_{acc}$ events, and keep 90% of the signal events.

An attempt to remove background from out-of-time energy deposits in the calorimeter was made using the "Adder" ratios. As discussed earlier the sums of 3×3 arrays of lead glass blocks were recorded with a gate almost three times shorter than the charge integrating gates for individual blocks. The energy deposited in this smaller time interval was compared to the cluster energy derived from the longer gate. Events that recorded an excess of 25% or more energy over the adder sum in the longer gate were rejected. This was a loose cut only removing 0.1% of the $K_L \rightarrow \mu^+\mu^-\gamma$ Monte Candidates, but 8% of the $K_L \rightarrow \pi^{\pm}\mu^{\mp}\nu\gamma_{acc}$ events were removed with this cut.

In Fig. 5.13 we show the transverse momentum squared (P_t^2) versus invariant $\mu^+\mu^-\gamma$ mass for the remaining events in the sample. There is an excess of approximately 200 events near the kaon mass and extending up in P_t^2. The distribution of events predicted by the signal and normalization Monte Carlo samples are shown in Fig. 5.14. The box shown contains 77% of the sample of Monte Carlo $K_L \rightarrow \mu^+\mu^-\gamma$ events, and excludes more than 99% of the $K_L \rightarrow \pi^{\pm}\mu^{\mp}\nu\gamma_{acc}$ sample. The invariant mass distribution under the $K_L \rightarrow \mu^+\mu^-\gamma$ hypothesis for all events with $P_t^2 < 200$ MeV$^2/c^2$ is displayed in Fig. 5.15.

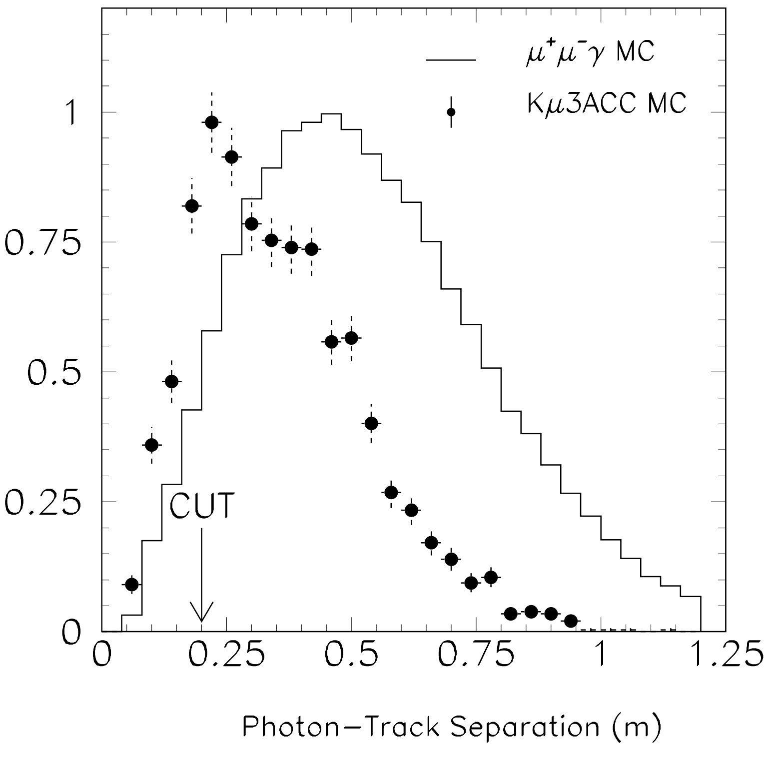

Figure 5.12: Separation of the photon cluster and the tracks at the calorimeter

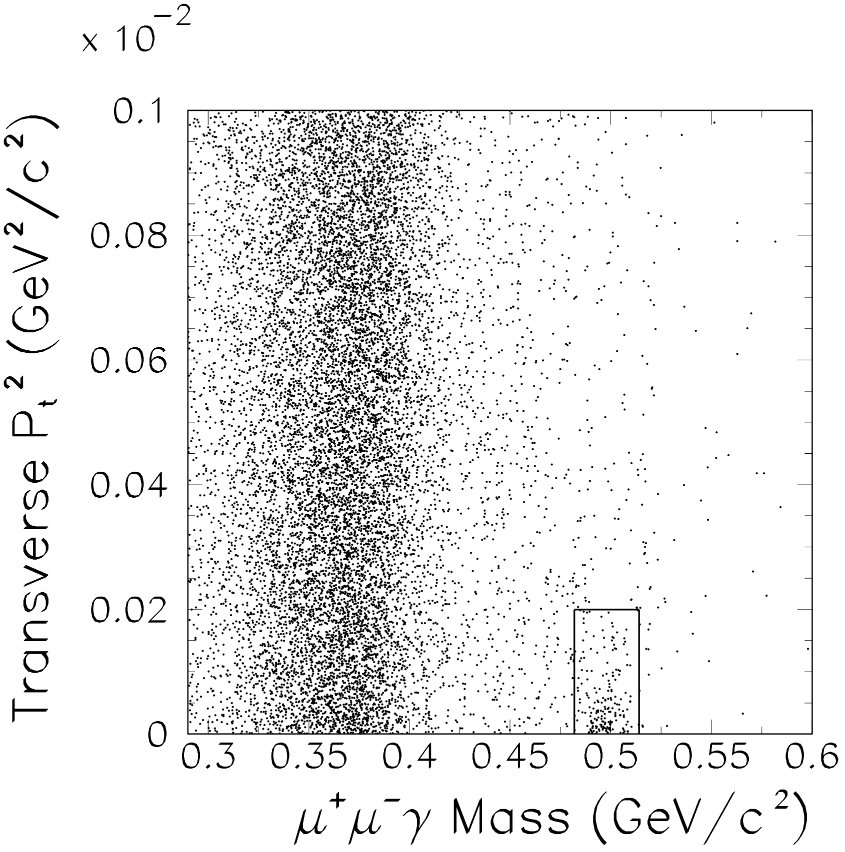

Figure 5.13: $P_t^2 - vs - \mu^+\mu^-\gamma$ invariant mass distribution for data events after cuts.

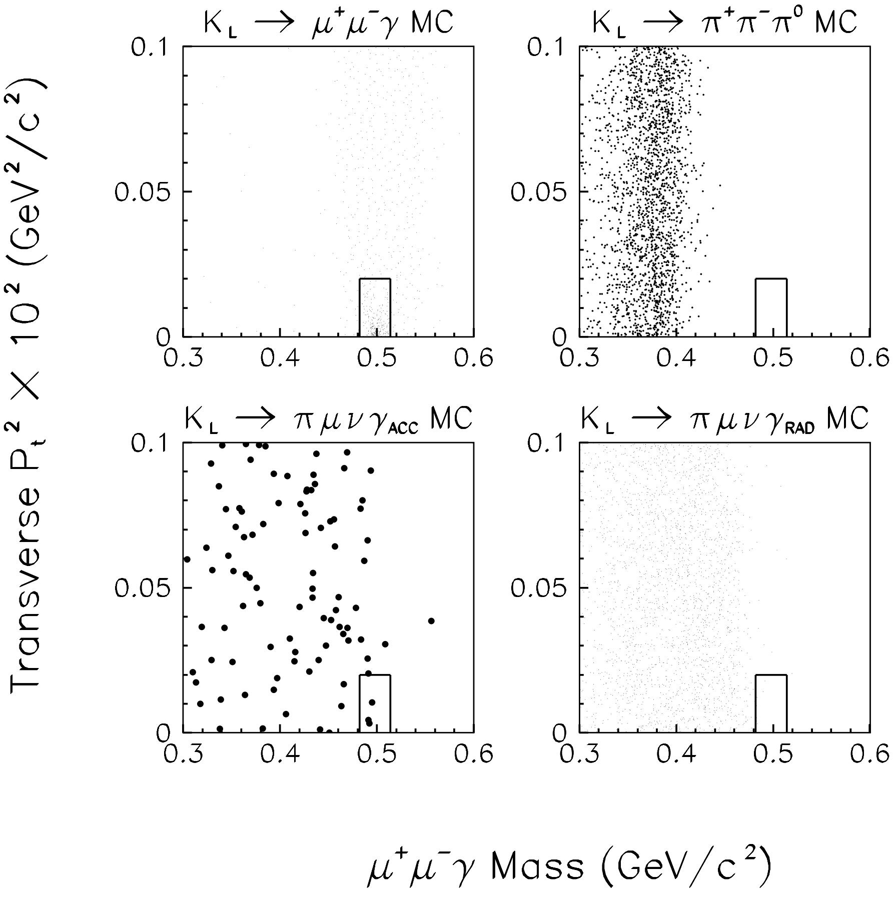

Figure 5.14: $P_t^2 - vs - \mu^+\mu^-\gamma$ invariant mass distribution for Monte Carlo events.

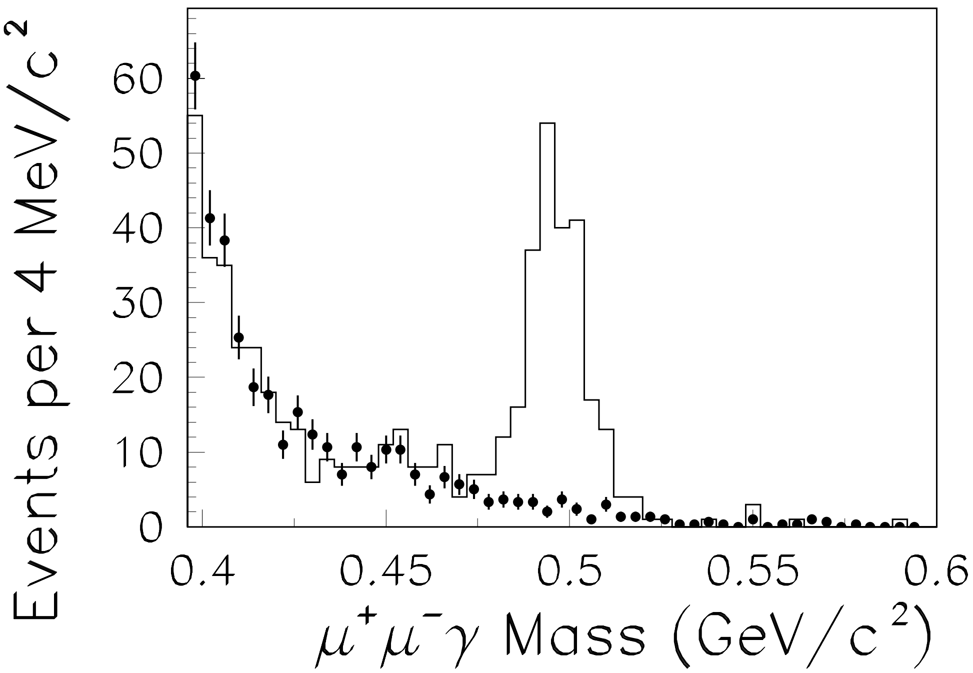

Figure 5.15: Distribution of $\mu^+\mu^-\gamma$ invariant mass after cuts. Histogram is data for events with $P_t^2 < 200$ MeV$^2/c^2$ and points are data for events with $400 < P_t^2 < 1000$ MeV$^2/c^2$, and normalized to the low P_t^2 data.

5.4 Background Subtraction

The points in Fig. 5.15 are an estimate of the background from a high P_t^2 in Fig. 5.13. We make only the assumption that the shape of the data with respect to $\mu^+\mu^-\gamma$ invariant mass is the same throughout the region $P_t^2 < 1000\ MeV^2/c^2$, and also that it is uniform with respect to P_t^2. Then we choose the region $400 < P_t^2 < 1000\ \mathrm{MeV}^2/c^2$ to obtain the distribution of background events with respect to the reconstructed $\mu^+\mu^-\gamma$ invariant mass. This enables us to obtain three times the statistical accuracy for a background shape. The region $0.425 < M_{\mu^+\mu^-\gamma} < 0.475\ GeV$ was used to normalize the shape. The effect of subtracting any residual signal events in this region is negligible. From this data we estimate there to be 10.5 background events in the signal region. Also we can extrapolate back with respect to P_t^2 and get another estimate of the background, under the assumption that it is flat with respect to P_t^2. This gives an estimate of 10.3 background events. A simple linear extrapolation of the background gives an estimate of 18 events. From these results we assign the number of background events under the mass peak to be 10.5 ± 4.0. Even with this rather conservative error on the background subtraction, the contribution to the error is much smaller than the dominant contribution to the systematic error on the branching fraction.

A Monte carlo data set was generated according to the kaon decay profile in Fig. 5.16. The number of events per run in the Monte Carlo simulation was

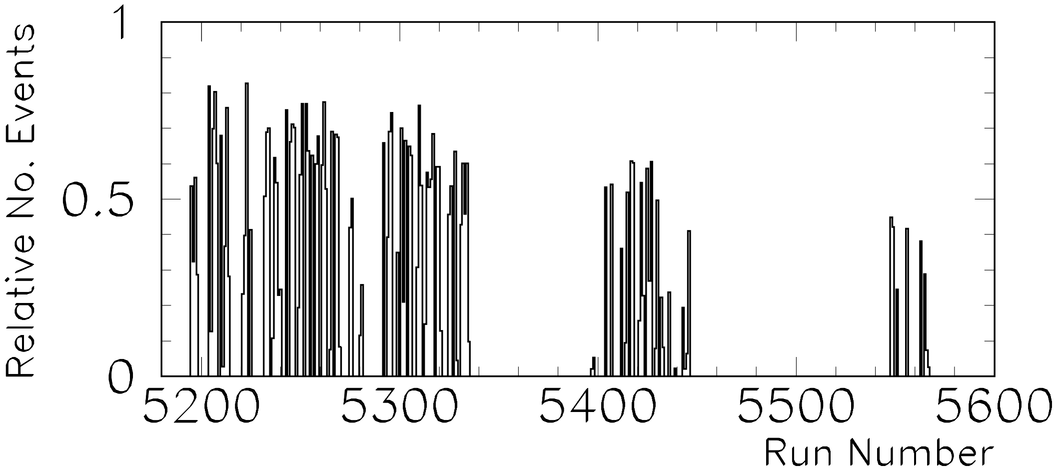

Figure 5.16: Relative number of $K_L \rightarrow \pi^+\pi^-\pi^0$ events observed per run.

determined from the $K_L \rightarrow \pi^+\pi^-\pi^0$ data sample.

In each event the results of a random accidental trigger were added to the raw information of all detectors. The accidental data were the result triggering on a random bucket. In this way we obtained a representative sample of the instantaneous ambient activity throughout the detector.

The resulting number of signal events is 190.5. In the Monte Carlo sample there were 58500 events remaining from a sample of one million generated $K_L \rightarrow \mu^+\mu^-\gamma$ decays at the final stage of the analysis implying an acceptance of 1.95%.

No.	Cut	Efficiencies	
		Data	MC
N1	Two extra in time clusters in calorimeter	0.85	0.33
N2	$5 < \|\vec{p}_{track}\| < 99\ GeV$	0.993	0.988
N3	Extra cluster energy $> 7\ GeV$	0.447	0.430
N4	Adder ratios $>\ 0.75$	0.997	0.998
N5	Track cluster - extra cluster separation $> 20\ cm$	1.00	1.00
N6	$\|M(2\gamma) - m_{\pi^0}\| < 11\ \ MeV(2\sigma)$	0.99	0.99
N7	$P_T^2 < 200\ \ MeV^2/c^2$	0.84	0.86

Table 5.3: $K_L \to \pi^+\pi^-\pi^0$ analysis cuts.

5.5 $K_L \to \pi^+\pi^-\pi^0$ Analysis

Table 5.3 lists the requirements that were imposed on the normalization data set in addition to the cuts listed in Tab. 5.1.

We required two in-time energy deposits that were isolated from clusters associated with tracks. The energy of the photon clusters was required to be above $7\ GeV$ (Fig. 5.17). This cut is slightly lower than the $K_L \to \mu^+\mu^-\gamma$ cut because the average energy of the photons is lower in this decay mode and any potential systematic biases between the signal and normalization modes is minimized in this way. This acceptance for this cut was 43%.

The adder ratio cut was identical to the $K_L \to \mu^+\mu^-\gamma$ analysis, and removed only 0.3% of the candidate normalization data sample, but 0.2% of the $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo events. To be consistent with the signal mode cuts we imposed a cut of $20\ cm$ on the distance between the track position at the face of the calorimeter and the position of the photon. The impact of this cut was negligible however.

Events with a transverse momentum squared greater than 200 MeV^2/c^2 were removed from the data set, as shown in Fig. 5.18 this quantity is simulated very well by Monte Carlo which gives us confidence that the same cut in the $K_L \rightarrow \mu^+\mu^-\gamma$ sample contains no systematic biases.

The invariant mass distribution of the two photon candidates is illustrated in Fig. 5.19. We retained only those events that were within 3σ (20 MeV) of m_{π^0}. The remaining events should have a reconstructed invariant mass close to the neutral kaon mass, and Fig. 5.20 shows that this is the case.

There were 20919 remaining events in the sample in the final stage of the analysis. In the Monte Carlo sample, 58500 events remained, from a sample of three million generated K_L decays, implying an acceptance of 1.95%

5.6 Extraction of the Branching Fraction

Because no suitable data set existed to calibrate the hodoscope bank MU1 we were forced to treat it in a different manner from the other acceptances. The method we used was to estimate upper and lower bounds for the rate at which MU1 rejected dimuon events, and to determine an average rate of rejection given those bounds.

We divided the $K_L \rightarrow \mu^+\mu^-\gamma$ events into two classes: the first class contained events where one of the tracks passed through the hole in the center of MU1 and the

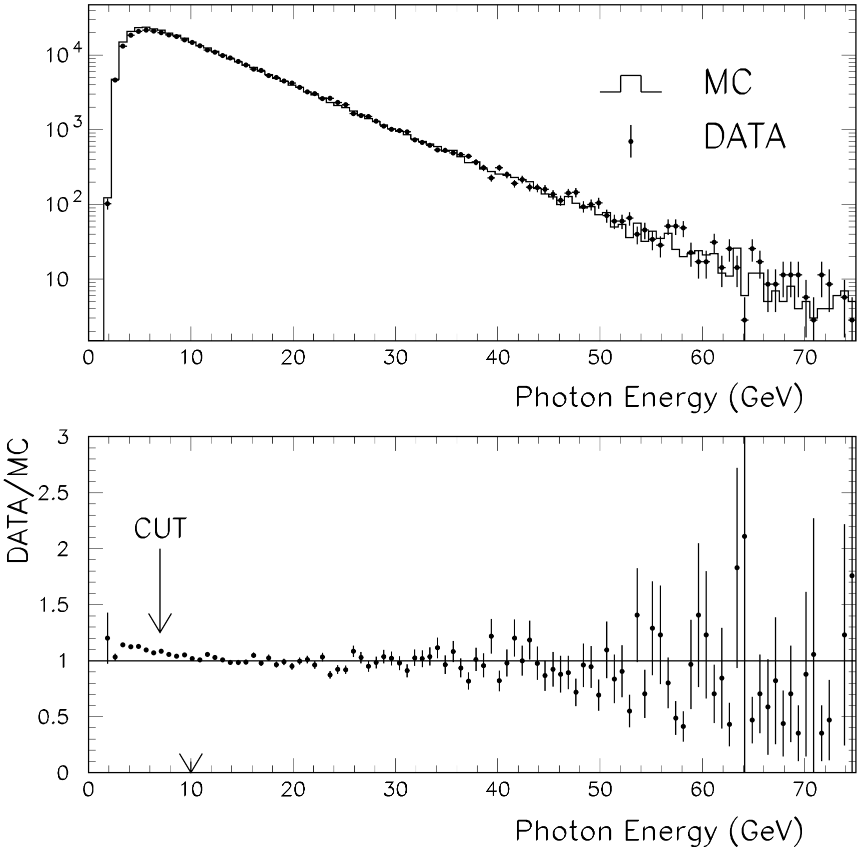

Figure 5.17: Distribution of photon energies in $K_L \rightarrow \pi^+\pi^-\pi^0$ events.

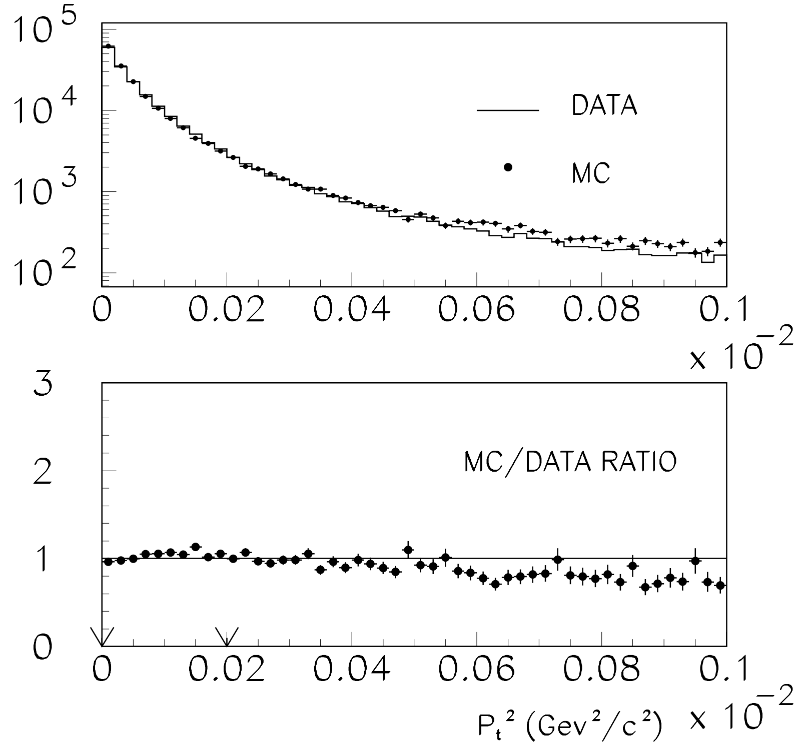

Figure 5.18: Comparison of transverse momentum squared for $K_L \rightarrow \pi^+\pi^-\pi^0$ events between Monte Carlo and data. The arrows indicate the region over which the two distributions were normalized.

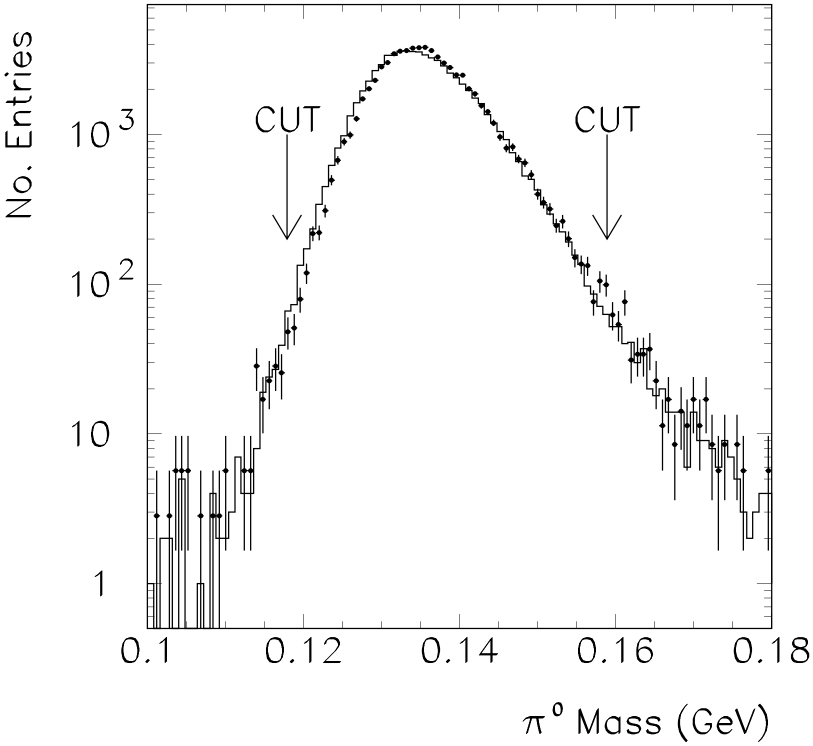

Figure 5.19: The invariant mass distribution of the the two photons in $K_L \to \pi^+\pi^-\pi^0$ candidate events (line) and the same quantity for Monte Carlo data (points).

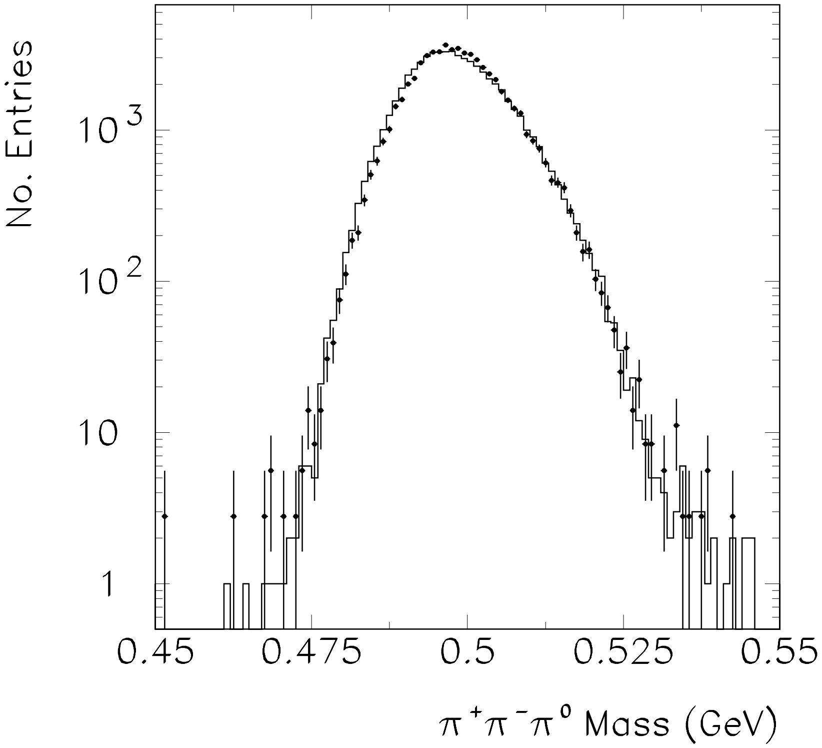

Figure 5.20: Invariant $\pi^+\pi^-\pi^0$ mass for events in the final sample of normalization data (line) compared to Monte Carlo (points).

other track passed through one of the scintillator slats; the second class consisted of events where neither track passed through the hole. These two data sets are mutually exclusive, and constitute 99% of the total data set (the other 1% of events have both tracks pass through the hole). This division was necessary because events in which two muons passed simultaneously through the hodoscope triggered the counter bank at a different rate than single muons.

Each of the two classes were then treated separately to estimate the rate at which the MU1 trigger was satisfied. If we could determine with infinite precision upper and lower bounds for these rates the answer would be obvious: we would take the average and the uncertainty on the rate would be $1/\sqrt{12}$ times their separation. However, the upper bounds were themselves uncertain, so we modeled the behavior to predict the rates and uncertainties on the rates as follows: given a lower limit for an acceptance $\mathcal{A}_L$, and an upper limit $\mathcal{A}_U \pm \sigma_U$:

- Fix the lower lower limit $\mathcal{A}_1 = \mathcal{A}_L$ and choose an upper limit $\mathcal{A}_2$ from a gaussian distribution with mean $\mathcal{A}_U$, and standard deviation σ_U.
- Draw an acceptance $\mathcal{A}^{MU1}$ randomly from a uniform distribution between $\mathcal{A}_1$ and $\mathcal{A}_2$.
- Draw a random number N from a gaussian distribution with mean N_{sig} and standard deviation $\sqrt{N_{sig}}$.

- Evaluate $\Phi_{\mu^+\mu^-\gamma} = \frac{N}{\mathcal{A}A^{MU1}}$

Here $\mathcal{A}$ is the acceptance calculated from a Monte Carlo simulation for all detectors except MU1 and we inserted it in the final step so that the quantity $\Phi_{\mu^+\mu^-\gamma}$ is the absolute flux of $K_L \to \mu^+\mu^-\gamma$ decays in the fiducial volume of our detector during the run. We repeated the above procedure many times and determined the statistical spread of the result as illustrated in Fig. 5.21.

After background subtraction the signal data set contained 103.2 events where one track extrapolated to the hole in MU1 (defined as the "IN" sample) and 93.3 events where both tracks extrapolated to the active area of MU1 (defined as the "OUT" sample).

The MU1 acceptance for the IN sample had a lower limit of 22.1% and an upper limit of $25.4 \pm 1.0\%$. We determined the $K_L \to \mu^+\mu^-\gamma$ flux from this sample to be 11010 ± 1096 which corresponds to an average MU1 acceptance of 23.7%.

The MU1 acceptance for the OUT sample had a lower limit of 43.7% and an upper limit of $58.8 \pm 3.3\%$ and we determined the $K_L \to \mu^+\mu^-\gamma$ flux from this sample to be 8790 ± 1286 and the corresponding average MU1 acceptance was 51.8%

The two values for the $K_L \to \mu^+\mu^-\gamma$ flux are statistically independent samples measuring the same quantity. The particle data group give a prescription for combining data sets $x_i \pm \sigma_i$ [27]. The weight of each measurement is $w_i = 1/\sigma_i^2$, where σ_i is the error for the i^{th} measurement and the best estimate of the quantity

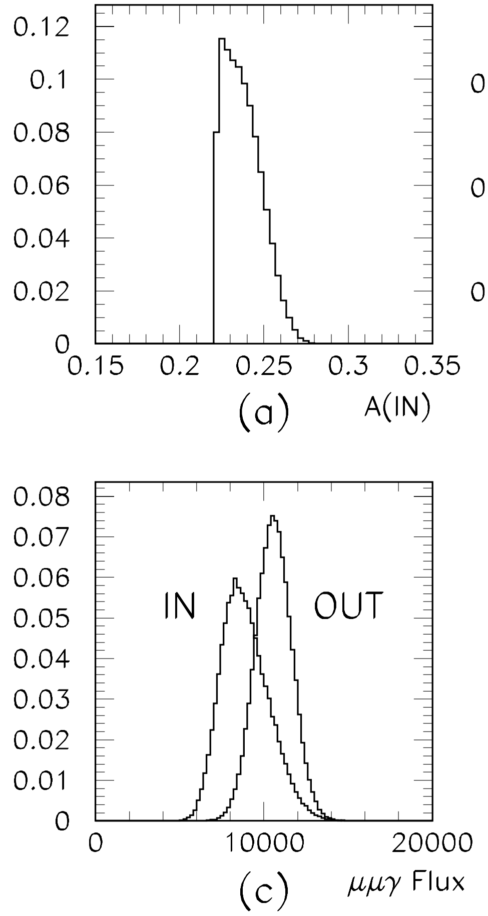

Figure 5.21: (a) The acceptance for the 'in' case, drawn randomly between the upper and lower limits for MU1 efficiency, (b) the acceptance for the 'out' case, again drawn randomly between the upper and lower limits, (c) the distributions of the $K_L \rightarrow \mu^+\mu^-\gamma$ flux for the 'in' and 'out' cases, (d) the distribution of the χ^2 statistic, with the value obtained this data.

is $\sum w_i x_i / \sum w_i \pm (\sum w_i)^{-1/2}$. The method is only statistically correct if $\chi^2/(N-1)$ is less than one, where $\chi^2 = \sum w_i(\bar{x} - x_i)^2$. For our data set, we calculate:

$$\langle \Phi_{\mu^+\mu^-\gamma} \rangle = 10075 \pm 834 \tag{5.2}$$

$$\chi^2 = 1.7 \tag{5.3}$$

However due to the non gaussian nature of the distributions we don't use tables of the χ^2 statistic, but rather use the distribution of this quantity from Fig. 5.21(d), from which we find,

$$P(\chi^2 > 1.7) = 51\% \tag{5.4}$$

and we thus conclude that the two data sets from the two classes do not differ significantly, and we combine them using this weighting method. The error of 8.28% in Eqn. 5.2 is a combination of the statistical fluctuation on 196.5 events which is 7.2% and the spread resulting from the uncertainty in the MU1 acceptance which must then be 4.2%. For the 196.5 events in the full data sample the average acceptance is thus 1.95%. In Tab. 5.4 we summarize the treatment of MU1. From the acceptances and numbers of events for the signal and normalization modes we calculate the following branching fraction

$$\frac{B(K_L \to \mu^+\mu^-\gamma)}{B(K_L \to \pi^+\pi^-\pi^0)} = \frac{196.5/0.0195}{3600 \times 20919/0.0195} = (2.61 \pm 0.19(\text{stat.})) \times 10^{-6}. \tag{5.5}$$

	IN	OUT
N($K_L \to \mu^+\mu^-\gamma$)	103.2	93.3
$\mathcal{A}$(No MU1)	1.23%	2.20%
$\mathcal{A}^{\text{MU1}}$ bounds	$[22.1\%, 25.4 \pm 1.0\%]$	$[43.7\%, 58.8 \pm 3.3\%]$
$\langle \mathcal{A}^{\text{MU1}} \rangle$	76.3%	48.2%

Table 5.4: Summary of MU1 Acceptances.

	$K_L \to \mu^+\mu^-\gamma$	$K_L \to \pi^+\pi^-\pi^0$
$\mathcal{A}$(No MU1)	3.43%	1.95%
$\mathcal{A}^{\text{MU1}}$	56.8 ± 2.4	100%
N_{EVENT}	196.5	20919
$\text{N}_{\text{MC}}/\text{N}_{\text{GEN}}$	34300/1E6	58500/3E6
Prescale	1	3600

Table 5.5: Summary of branching ratio calculation.

5.7 Systematic Uncertainties

In this section we estimate the uncertainties in the evaluation of the of the branching ratio from the analysis procedure.

Mu1 Acceptance

Due to the method described above, the error in the branching fraction due to the uncertainty in the rate at which MU1 vetoed events emerges naturally. The uncertainty in the acceptance due to the MU1 hodoscope bank was 4.5%. This dominated the non-statistical part of the uncertainty in the branching fraction.

Background Subtraction

We have already mentioned the background subtraction above. By considering three different methods for extrapolating the background event distribution under the signal peak we estimated the total number of events in the signal box to be 10.5 ± 4.0 events. The uncertainty in this number divided by the best estimate for the total number of signal events (190.5) gives the uncertainty in the branching fraction, or 2.0%.

MU3 Acceptance:

This systematic uncertainty was discussed in the previous chapter. The contributions were due to trigger logic and hodoscope inefficiencies and taken together this represents an uncertainty of 2.9%.

Photon Energy Cut:

Figure 5.22 shows how the agreement between Monte Carlo and data for the photon energies in $K_L \to \pi^+\pi^-\pi^0$. The acceptance may be wrong by up to several percent if we included all photon energies. The figure implies that acceptance for $K_L \to \pi^+\pi^-\pi^0$ can change by only about 0.3% depending on the value of this cut.

Fusion Cuts:

When all other analysis cuts were left in place and the fusion cut was imposed on a $K_L \to \mu^+\mu^-\gamma$ Monte Carlo sample, it was found that 5% of the events were consistent with having a photon cluster fused with another cluster. An estimate of the effect on the acceptance from this photon fusion requirement can be obtained from a study of $K_L \to \pi^0\pi^0\pi^0(\pi^0 \to e^+e^-\gamma)$ decays [54]. In that study 1/5th of the events rejected by the fusion cut were incorrectly identified. In this data sample five

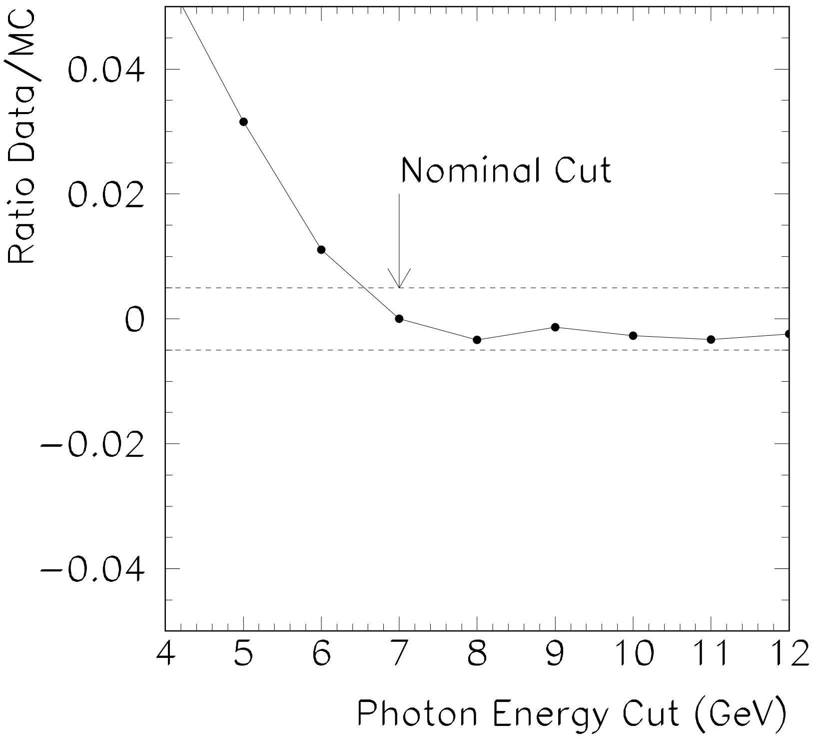

Figure 5.22: As the cut on photon energy in $K_L \rightarrow \pi^+\pi^-\pi^0$ is increased a region is entered where the ratio of data and Monte Carlo becomes stable.

photon clusters were required and there was some probability that this increased the rate of fused clusters due to their proximity. Also the $K_L \rightarrow \pi^0\pi^0\pi^0(\pi^0 \rightarrow e^+e^-\gamma)$ photons are typically closer to the center of the calorimeter where the radiation damage was worst. Both of these factors tend to make the rejection higher than $K_L \rightarrow \mu^+\mu^-\gamma$ sample which is especially clean, with two minimum ionizing deposits and a single isolated photon deposit. We conclude that the effect of photon fusion is less than 1/5th of 5%. We assign a reasonable estimate of 0.5% to this uncertainty.

The π^0-mass Cut:

We studied how varying the 3σ mass cut in the $K_L \rightarrow \pi^+\pi^-\pi^0$ analysis could affect the branching ratio, since Fig. 5.19 shows a small discrepancy between the Monte Carlo and Data in this quantity, which is consistent with a mass scale shift of about 2 MeV. By varying the cut we observed the change in the branching ratio. Figure 5.23 demonstrates that provided we keep this cut greater than 3σ we are immune to the shape disagreement, and the bias is essentially zero.

Comparison of Different Analyses:

From this analysis we determined the total kaon flux in our detector between 90 m and 180 m from the target and in a momentum range of 20 GeV to 220 GeV to be

$$\begin{aligned}
\Phi_{K_L} &= \frac{N(K_L \rightarrow \pi^+\pi^-\pi^0)}{\mathcal{A}(K_L \rightarrow \pi^+\pi^-\pi^0)B(K_L \rightarrow \pi^+\pi^-\pi^0)} \times 3600 \\
&= \frac{20919}{0.0195 \times 0.1238} \times 3600
\end{aligned}$$

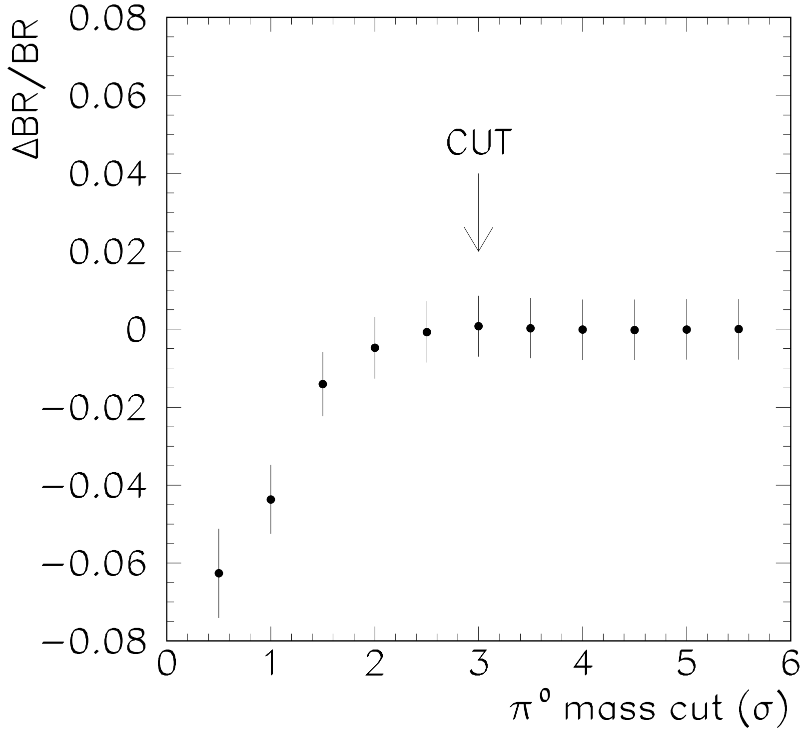

Figure 5.23: The effect on the branching ratio as the cut on π^0 mass in $K_L \rightarrow \pi^+\pi^-\pi^0$ decays is varied.

$$= \; 31.8 \times 10^9.$$

In the analysis of $K_L \rightarrow \pi^0\mu^+\mu^-$ using the same normalization data but with a different set of cuts the same quantity was determined to be 32.6×10^9 [60]. In an analysis of the decay $K_L \rightarrow \pi^+\pi^-{\pi^0}_D$ ($K_L \rightarrow \pi^+\pi^-\pi^0$ where the π^0 decays via $\pi^0 \rightarrow e^+e^-\gamma$) the flux was determined to be 31.8×10^9 [58]. We therefore have a consistent picture between this analysis and other previously published decay modes from the same experiment.

Other Analysis Cuts:

We believe that the method of taking the “double ratio” of (data)/ (Monte Carlo) for the signal and normalization modes reduces the possible biases from analysis cuts. The cases where the Monte Carlo disagrees significantly, such as in the energy reconstruction in the calorimeter have been considered separately. Cases where an analysis cut was made in the normalization mode, but not the signal mode such fusion finding cut have been considered.

The systematic uncertainties on the acceptances are summarized in Tab. 5.6. Combined in quadrature, the total systematic uncertainty is estimated to be 5.5%

MU1 Efficiency	4.2%
MU3 Efficiency	2.5%
Background Sub.	2.0%
MU3 Trigger Logic	1.4%
Fusion Cut	0.5%
Cluster Energy Cut	<0.5%
π^0-mass cut	<0.5%
Total	5.5%

Table 5.6: Summary of systematic uncertainties.

5.8 Summary

With the systematic uncertainty included the branching ratio calculated in this chapter is

$$\frac{B(K_L \to \mu^+\mu^-\gamma)}{B(K_L \to \pi^+\pi^-\pi^0)} = (2.61 \pm 0.19(\text{stat.}) \pm 0.14(\text{sys.})\ ) \times 10^{-6} \tag{5.6}$$

and using $B(K_L \to \pi^+\pi^-\pi^0) = 0.1238 \pm 0.0021$ [27] we find

$$B(K_L \to \mu^+\mu^-\gamma) = (3.23 \pm 0.23(\text{stat.}) \pm 0.19(\text{sys.})\ ) \times 10^{-7} \tag{5.7}$$

where the error on the current world average of $B(K_L \to \pi^+\pi^-\pi^0)$ has been included in the systematic error in Eqn. 5.7.

Chapter 6

$K_L \to \mu^+\mu^-\gamma$ Form Factor

This chapter discusses the form factor of the decay $K_L \to \mu^+\mu^-\gamma$. We extract measurements of this form factor from the absolute magnitude of the branching fraction, and also from a shape analysis of the invariant dimuon mass spectrum.

6.1 Introduction

The decay amplitude $K_L \to \mu^+\mu^-\gamma$ is thought to be dominated by a two-photon intermediate state with one off-shell photon that subsequently decays into a $\mu^+\mu^-$ pair, whose invariant mass is the q^2 of the photon. Then the differential decay rate with respect to the invariant mass of that photon contains the function $A_D(q^2)$. The BMS model parameterizes that function as $A_D(q^2) = A_1(q^2) + \alpha_{K^*} A_K(q^2)$, where A_1 and A_K are known functions. The distribution of events with respect to q^2 is therefore a probe of this model. But integrating the differential decay rate shows that the total rate is also a function of α_{K^*}. A check of the consistency of this model is to look at both these quantities and to measure α_{K^*} in each case.

6.2 Extracting α_{K^*} from the Branching Ratio

Using the differential decay rate given in Eqn. 1.4 and normalizing to $K_L \to \gamma\gamma$ we have

$$\Gamma_{\gamma\gamma}^{-1}\frac{d\Gamma}{dx}(K_L \to \mu^+\mu^-\gamma) = \frac{2\alpha}{3\pi}\frac{|f(x)|^2}{x}(1-x)^3\left(1+\frac{2m_\mu^2}{xm_K^2}\right)\left(1-\frac{4m_\mu^2}{xm_K^2}\right)^{1/2} \quad (6.1)$$

where $x = m_{\mu\mu}^2/m_K^2$. Equation 1.10 can be re-written to give

$$f(q^2) = \frac{1}{1-q^2/m_\rho^2} + \alpha_{K^*}\sqrt{2}eG_{NL}f_{K^*K\gamma}\left(\frac{m_\rho^2}{f_K f_\rho^2}\right)\frac{1}{A_{\gamma\gamma}} \times\frac{1}{1-q^2/m_{K^*}^2}\left\{\frac{4}{3}-\frac{1}{1-q^2/m_\rho^2}-\frac{1}{9}\frac{1}{1-q^2/m_\omega^2}-\frac{2}{9}\frac{1}{1-q^2/m_\phi^2}\right\}.$$

We use identical expressions for the meson form factor constants as in Ref. [35]

$$G_{NL} = 1.1\times 10^{-5}/m_p^2, \qquad f_{K^*\gamma}^2 = \frac{96\pi\,\Gamma(K^*\to K^0\gamma)m_{K^*}^3}{m_{K^*}^2-m_K^2} = 3.90\times 10^{-4}MeV^{-1},$$

$$f_\rho^2 = \frac{4\pi\alpha^2 m_\rho}{3,\,(\rho \to e^+e^-)} = (4.99)^2, \qquad f_{K^*} = \frac{m_{K^*}}{m_\rho} f_\rho = 5.78,$$

$$A_{\gamma\gamma}^2 = \frac{64\pi,\,(K_L \to \gamma\gamma)}{m_K^3} = (3.44 \times 10^{-12} MeV)^2, \qquad \alpha = 1/137.0,$$

we then obtain

$$\begin{aligned} f(x) = {} & \frac{1}{1-0.418x} \\ & +2.5\alpha_{K^*}\ \frac{1}{1-0.311x}\left\{\frac{4}{3} - \frac{1}{1-0.418x} - \frac{1}{9}\frac{1}{1-0.405x} - \frac{2}{9}\frac{1}{1-0.238x}\right\}. \end{aligned}$$

This expression can be inserted into Eqn. 6.1 and the only remaining free parameter is α_{K^*}. Integrating Eqn. 6.1 over the dimuon mass spectrum then gives the branching fraction for $K_L \to \mu^+\mu^-\gamma$ as a function of α_{K^*}, and the result is well approximated by a polynomial

$$\frac{B(K_L \to \mu^+\mu^-\gamma)}{B(K_L \to \gamma\gamma)} = (55.57 - 45.56\alpha_{K^*} + 11.07{\alpha_{K^*}}^2) \times 10^{-5}. \tag{6.2}$$

For reference if we assume the form factor is of the form $f(x) = 1 + \sigma x$ then the variation of the branching ratio with σ is

$$\frac{B(K_L \to \mu^+\mu^-\gamma)}{B(K_L \to \gamma\gamma)} = (40.91 + 26.88\sigma + 4.98\sigma^2) \times 10^{-5}. \tag{6.3}$$

Using , $(K_L \rightarrow \pi^+\pi^-\pi^0) = 0.1238\pm0.0021$ and , $(K_L \rightarrow \gamma\gamma) = 0.5.70\pm0.27$ our measured value of $B(K_L \rightarrow \mu^+\mu^-\gamma)$ corresponds to $B(K_L \rightarrow \mu^+\mu^-\gamma)/B(K_L \rightarrow \gamma\gamma) = (5.66 \pm 0.59) \times 10^{-4}$. Then from Eqn.6.2 we derive

$$\alpha_{K^*} = -0.018^{+0.131}_{-0.123} \qquad \text{(BR only)}. \tag{6.4}$$

6.3 Direct Measurement of the Form Factor

The form factor $f(x)$ is also directly measurable. We calculated the invariant mass of the two charged tracks in the 207 $K_L \rightarrow \mu^+\mu^-\gamma$ candidate events. Using the Monte Carlo generated data set with $|f(x)| = 1$ we were calculated the same quantity. Using the region of data $400 < P_t^2 < 1000 \quad MeV^2/c^2$ we estimated the invariant dimuon mass spectrum of the background events. This shape was normalized to 10.5 events and added to the Monte Carlo data. Figure 6.1 shows the comparison between Monte Carlo and data and also the acceptance which is a slowly varying function of the variable x. It is important to note that the effect of an uncertainty on the acceptance of the analysis cuts is uniform with respect to x and results in only an overall scaling of every bin in Fig. 6.1 by the same amount. Therefore fitting the shape is very insensitive to systematic uncertainties in the acceptance. To estimate the most likely value for α_{K^*} given this data we performed a maximum likelihood fit. The Monte Carlo distribution for a given value of α_{K^*} was normalized to 196.5

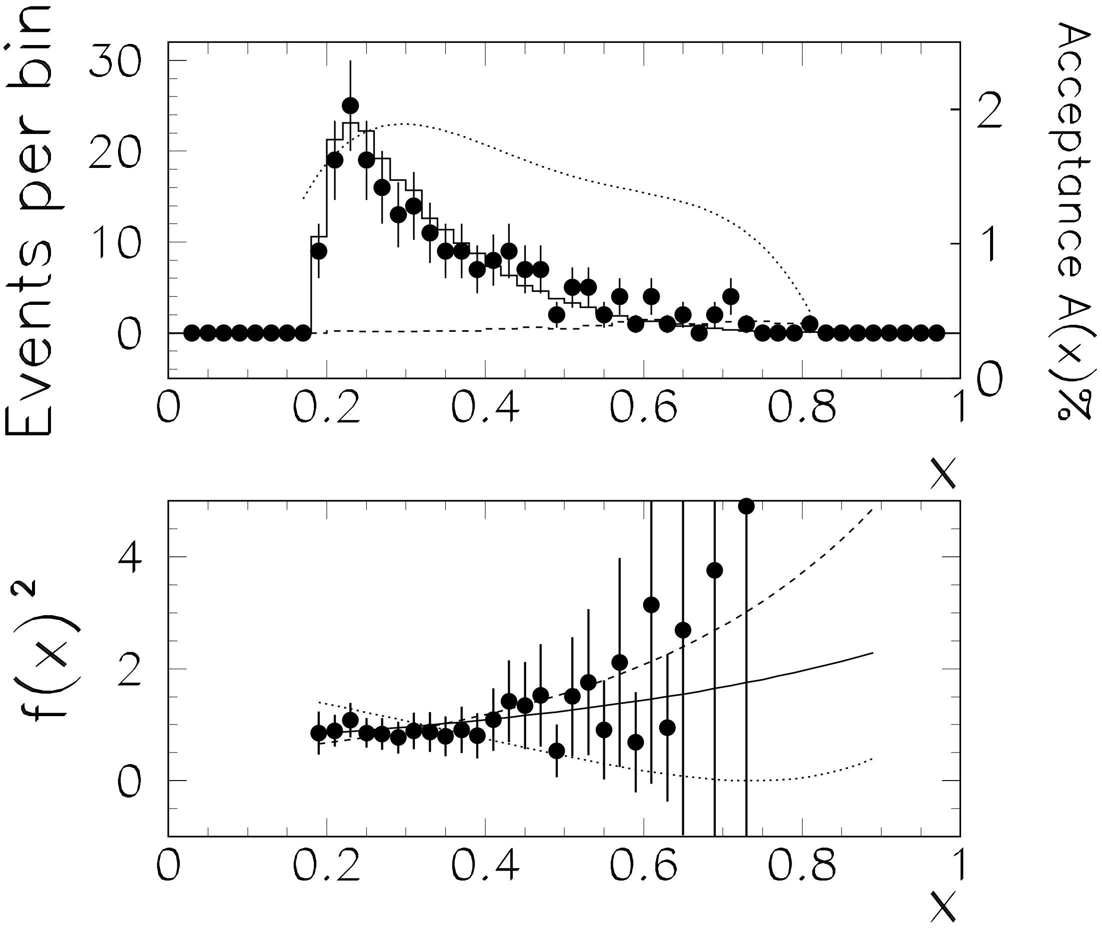

Figure 6.1: Top: Dimuon mass distribution, points are data, histogram is a Monte Carlo simulation with $f(x) = 1$, dashed line is the estimated background, dotted line is the acceptance. Bottom: the ratio of data to Monte Carlo from the top plot (*i.e.* our measured values of $f(x)^2$). The solid line is the result from a maximum likelihood fit. For illustration we also show predictions assuming $\alpha_{K^*} = -1$ (dashed) and $\alpha_{K^*} = +1$ (dotted). For these plots the normalization was fixed by the number of events.

events, and added to the background shape normalized to 10.5 events,

$$P(x;\alpha_{K^*}) \;=\; F^{MC}(x;\alpha_{K^*})\frac{196.5}{\int_0^1 F^{MC}(x)dx} + F^{BG}(x)\frac{10.5}{\int_0^1 F^{BG}(x)dx}$$

where $F^{MC}(x;\alpha_{K^*})$ is the Monte Carlo distribution of events with respect to x, $F^{BG}(x)$ is the background distribution of events and the normalization is such that $\int P(x;\alpha_{K^*})dx = 1$. We then evaluated a log-likelihood function

$$\mathcal{L}(\alpha_{K^*}) = -\log \prod_{i=1}^{207} P(x_i;\alpha_{K^*}). \qquad (6.5)$$

This function is shown in Fig. 6.2. The best estimate is $\alpha_{K^*} = -0.13$. To evaluate the error in the measurement from the likelihood fit, we assumed the correct value was -0.13 and used a simple Monte Carlo simulation to randomly draw 207 events from the distribution $P(x;-0.13)$. Then we applied the same algorithm described above to find the minimum negative log-likelihood in each of these Monte Carlo experiments. The distribution of log-likelihoods is shown in Fig. 6.3. From this we infer a 68% confidence interval is given by $-0.48 < \alpha_{K^*} < 0.14$, and hence we quote

$$\alpha_{K^*} = -0.13^{+0.27}_{-0.35} \qquad \text{(Shape Only)}. \qquad (6.6)$$

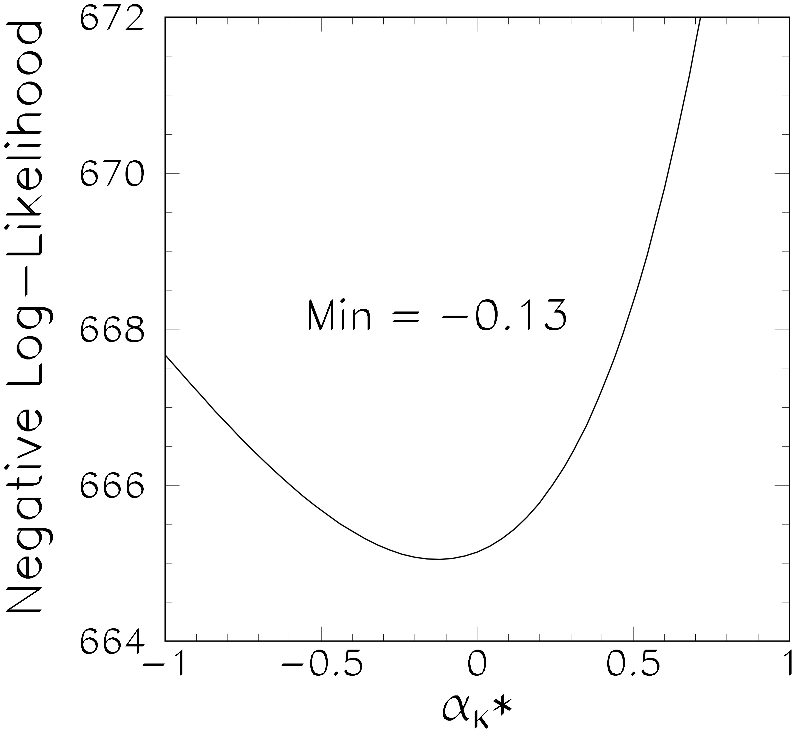

Figure 6.2: The negative log-likelihood function for the invariant dimuon mass spectrum, as a function of α_{K^*}.

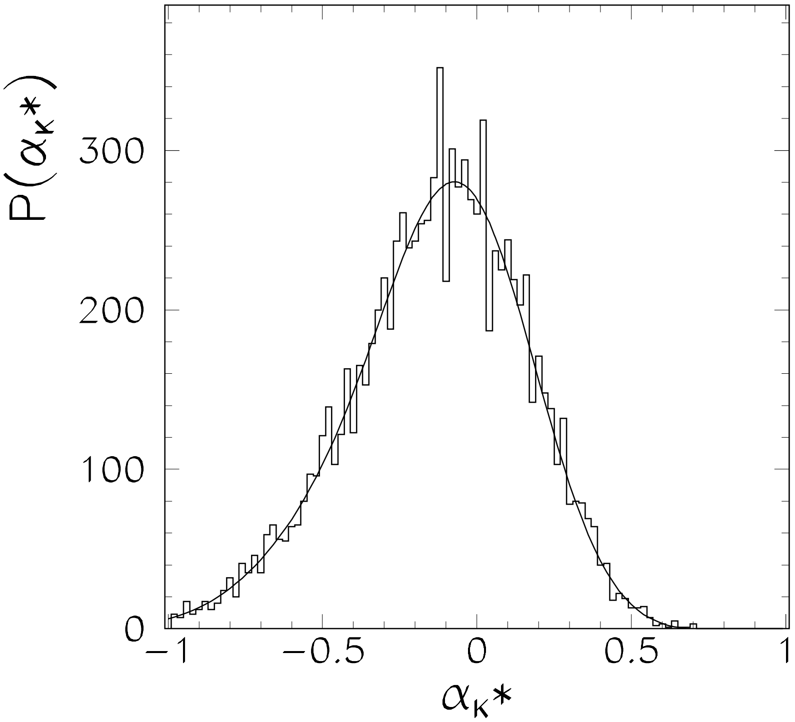

Figure 6.3: The spread in log-likelihoods from a simulation assuming α_{K^*}=-0.13.

We note that the error on this method are twice as large as those from the α_{K^*} derived from the branching ratio, emphasizing once again that most of the sensitivity to the form factor in $K_L \to \mu^+\mu^-\gamma$ is in the branching ratio rather the shape of dimuon invariant mass spectrum. The traditional method for finding confidence interval from a negative log-likelihood plot is to find the points for which $-\mathcal{L}$ rises by 1/2. From Fig. 6.2 this gives an interval of $-0.45 < \alpha_{K^*} < 0.14$ which is consistent with the interval listed above.

6.4 Combined Value for α_{K^*}

We inferred a likelihood function for α_{K^*} from the measured branching ratio only (assuming asymmetric gaussian errors for the measurement of α_{K^*} from Eqn. 6.4). We combined this with the log-likelihood function determined from the shape fit alone (Fig. 6.2). The result is shown in Fig. 6.4. Using the $\min(-\log\mathcal{L}) + 1/2$ points we determined from this curve the best estimate for α_{K^*}

$$\alpha_{K^*} = -0.028^{+0.115}_{-0.111} \qquad \text{(Combined Fit).}$$

We see that shape analysis improves the error on the high side a little, but provides little improvement on the lower side.

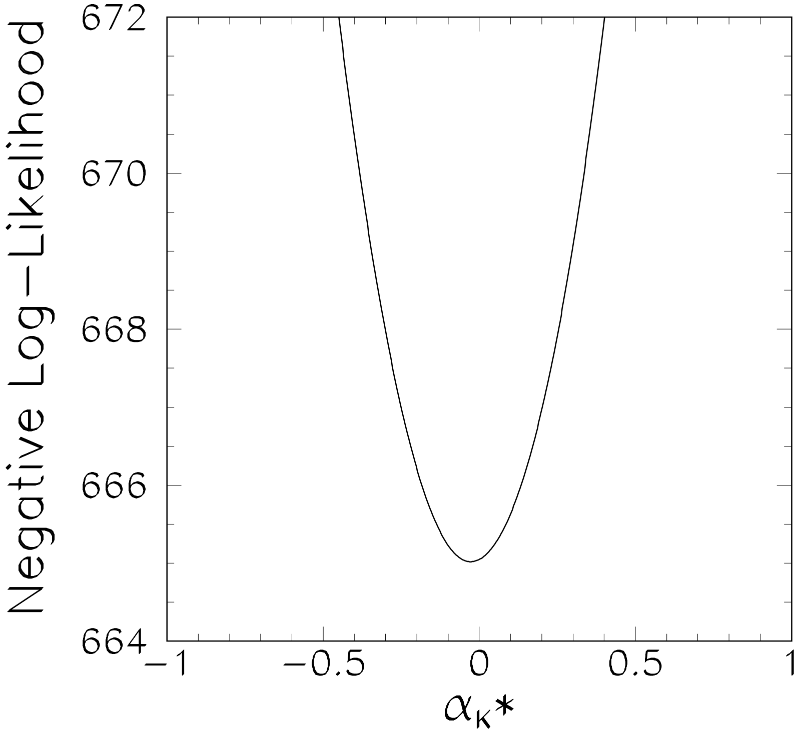

Figure 6.4: The negative log-likelihood for the combined shape fit plus branching ratio determinations of α_{K^*}, *i.e.* the likelihood from Fig. 6.2 combined with Eqn. 6.4. This likelihood curve implies central value of $\alpha_{K^*} = -0.028$. and a 68% confidence interval of $-0.139 < \alpha_{K^*} < 0.087$.

Chapter 7

Conclusion

Our measured branching fraction for $K_L \rightarrow \mu^+\mu^-\gamma$ from a sample of 207 events including 10.5 ± 4.0 background events leads to $B(K_L \rightarrow \mu^+\mu^-\gamma) = (3.23\pm0.23(stat)\pm 0.19(sys)) \times 10^{-7}$. From this we have extracted $\alpha_{K^*} = -0.028^{+0.115}_{-0.111}$.

To compare our value for the branching ratio of $K_L \rightarrow \mu^+\mu^-\gamma$ to previously published measurements of the form factor in $K_L \rightarrow e^+e^-\gamma$ we need to assume a model, and the standard in the literature has been to use the BMS model. This is illustrated in Fig. 7.1 where we display our result and also the two values for α_{K^*} extracted from $K_L \rightarrow e^+e^-\gamma$ shape analyses. We have taken all the previous results of related measurements and transformed them to an effective value of $B(K_L \rightarrow \mu^+\mu^-\gamma)/B(K_L \rightarrow \gamma\gamma)$ for comparison, as shown in Tab. 7.1.

Our result for the branching ratio is 2.7σ above the value that would be ex-

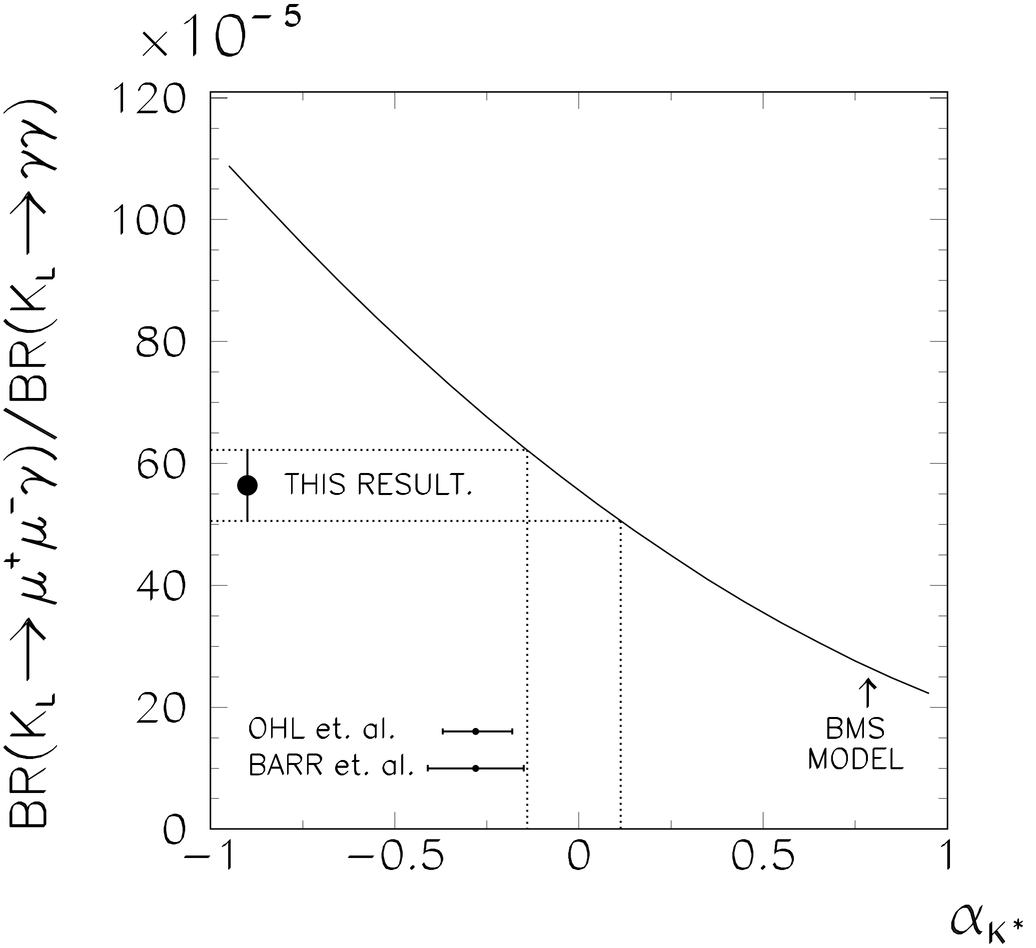

Figure 7.1: This result for the branching ratio of $K_L \rightarrow \mu^+\mu^-\gamma$, normalized to $K_L \rightarrow \gamma\gamma$ (large black dot) and the extracted range of the parameter α_{K^*} from the model of Bergström, Massó, Singer [1]. Also shown (horizontal error bars) are the extracted values of α_{K^*} from Refs. [35] and [36] for the case of $K_L \rightarrow e^+e^-\gamma$.

MODEL	$10^5 B(\mu^+\mu^-\gamma)/B(\gamma\gamma)$	Significance
• $f(x) = 1$	40.9	2.7σ
• Vector Meson Dominance [19] ($\alpha_{K^*} = 0$)	55.7	0.2σ
• BMS [1] ($\alpha_{K^*} = -0.28 \pm 0.08$)	69.2 ± 3.6	1.8σ
• Ko [22] (Eff $\mathcal{L}^{\Delta S=1}$)	$74.5^{+5.4}_{-1.5}$	2.9σ
• Gvozdev *et. al.* [23] (quark level)	44.8 ± 3.1	2.5σ

Table 7.1: Comparison of previously published works with our result of $B(K_L \to \mu^+\mu^-\gamma)/B(K_L \to \gamma\gamma) = (5.66 \pm 0.59) \times 10^{-4}$, and the significance of any differences.

pected if $f(x) = 1$ and this is the first indication of the presence of a form factor in the decay $K_L \to \mu^+\mu^-\gamma$. We see that the Vector Meson Dominance model (equivalent to ignoring the $K^0K^*\gamma$ contribution, $\alpha_{K^*} = 0$) is consistent with our result. Taking the combined result of the $K_L \to e^+e^-\gamma$ measurements as $\alpha_{K^*} = -0.28 \pm 0.08$ and transforming this to an inferred $K_L \to \mu^+\mu^-\gamma$ branching ratio we get $(69.3 \pm 3.6) \times 10^{-5}$. This is a difference of 1.8σ. The model of Ko, which includes additional $K^0(\rho\eta\phi)\gamma$ decay channels differs by 2.9σ from our result. The quark level result of Gvozdev *et al.* is $(44.8 \pm 3.1) \times 10^{-5}$ which is inconsistent with our result at the 2.5σ level. It is interesting to note that combining our result and $B(K_L \to e^+e^-\gamma) = (9.1 \pm 0.5) \times 10^{-6}$ from Ref. [27] implies

$$\frac{B(K_L \to e^+e^-\gamma)}{B(K_L \to \mu^+\mu^-\gamma)} = 28 \pm 3$$

and the BMS model with $\alpha_{K^*} = -0.28$ predicts the ratio to be 24 [3].

Assuming that the model of Bergström, Massó, and Singer is consistent we can find a world average for α_{K^*}. We find that combining our result of $\alpha_{K^*} = -0.028^{+0.115}_{-0.111}$ with the value from $K_L \to e^+e^-\gamma$ of $\alpha_{K^*} = -0.28 \pm 0.08$ [27] the average is $\alpha_{K^*} = -0.19 \pm 0.06$. This higher value for α_{K^*} leads to a slightly lower bound for ρ as seen in Fig. 1.9. For example if we take $\alpha_{K^*} = -0.13$ then the limit is $\rho > -0.80$. The bound on m_t is essentially unchanged, $m_t \lesssim 370\ GeV$ (see Fig. 1.11).

Finally in Tab 7.2 we give an indication of the improvement that may be

	Phase(I)	Phase(II)	f
Run Time	8 weeks	24 Weeks	3.00
Proton Intensity	1.2E12	5E12	4.17
Live Time	45%	70%	1.55
Beam Size	200 cm^2	290 cm^2	1.45
			~ 28

Table 7.2: Expected improvements in E799(II) relative to E799(I).

expected in the next generation of this experiment [61]. With improvements in the length of the run, the number of protons on target, the live time of the run, and the beam size, an increase by about a factor of 30 can be expected. This alone would yield $\sim$ 6000 $K_L \to \mu^+\mu^-\gamma$ events on tape. Also the acceptance is expected to be improved for the next run (for example: removing the MU1 detector from the trigger alone would improve the acceptance if $K_L \to \mu^+\mu^-\gamma$ events by a factor of almost 2). In this regime the measurement of $K_L \to \mu^+\mu^-\gamma$ would no longer be statistics limited and a detailed comparison of the models described above could be undertaken

However, within the current level of uncertainty in the experimental results the BMS model is consistent with all observed decays. Our result for $\alpha_{K^*} = -0.028^{+0.115}_{-0.111}$ is 1.8σ from the world average of $\alpha_{K^*} = -0.28 \pm 0.08$. Our result is also consistent with Vector Meson Dominance and in order to make a definitive statement about the current theoretical understanding of the $K\gamma\gamma^*$ vertex the higher statistics data samples from forthcoming experiments in both $K_L \rightarrow \mu^+\mu^-\gamma$ and $K_L \rightarrow e^+e^-\gamma$ will clearly be desirable.

Bibliography

[1] L.Bergstrom et al. Phys. Lett. 131B, 229, (1983)

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

[3] L.Bergstrom *et al.,* Phys. Lett. 249B,141 (1990)

[4] G. Belanger and C. Q. Geng. Phys. Rev. D43, 140 (1991)

[5] P. Ko. Phys. Rev. D44, 139 (1991)

[6] Paul Singer. Nuc. Phys. A527, 713c (1991)

[7] Jack L. Ritchie. Rev. Mod. Phys. 65, 1149 (1993)

[8] F. Abe *et al.*, Phys. Rev. Lett. 74, 14, 2626 (1995); S. Abachi *et al.*, Phys. Rev. Lett. 74, 14, 2632 (1995)

[9] E. Ma and A. Pramudita, Phys. Rev. D 24, 2476 (1981)

[10] We have used here: $G_F = 1.2 \times 10^{-5}\ GeV^{-2}$, $M_K = 0.497\ \ GeV$, $\alpha = 1/137$, $V_{ud} = 0.97$, $V_{us} = 0.22$, $f_K = 160\ \ MeV$. See also Ref [27] p1443

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

[12] M. K. Gaillard and B. W. Lee, Phys. Rev. D 10, 897 (1974)

[13] M. B. Voloshin and E. P. Shabalin, JETP Lett. 23, No. 2, 107 (1976)

[14] J. F. Donoghue *et. al.,* Phys. Rev. D 21, 186 (1980)

[15] M. A. Shifman et al. Nucl. Phys. B120, 316, (1977)

[16] A. Pitch, Invited Talk at ICHEP '94, Glasgow, July 1994, HEP-PH-9410208

[17] T. D. Lee and C. S. Wu, Ann. Rev. Nuc. Sci., 15, 381 (1965)

[18] N. .N. Kroll and W. Wada. Phys. Rev. 98, 1355, (1955)

[19] C. Quigg and J.D. Jackson. UCRL-Report No.18487, (1968)

[20] R. .F. Sarraga and H. J. Munczek. Phys. Rev. D4 2881 (1971)

[21] J. J. Sakurai. Phys. Rev. 156, 1508 (1968)

[22] P. Ko, Phys. Rev. D 44, 139 (1991)

[23] A. A. Gvozdev *et al.,* Phys. Lett. 292B, 176 (1992)

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

[25] T. Inami and C. S. Lim, Prog. Theor. Phys. 65, No.1, 197 (1979)

[26] L. Wolfenstein, Phys. Rev. Lett. 51, 1945 (1983)

[27] Particle Data Group, L. Montanet *et al.*, Phys. Rev. D 50, 1173 (1994)

[28] L. Ametller *et. al.,* Phys. Rev. D 30, 251 (1984)

[29] G. Buchalla *et. al.,* Nuc. Phys. B412, 106 (1994)

[30] G. Buchalla *et. al.,* Nuc. Phys. B349, 1 (1984)

[31] A. P. Heinson *et al.*, Phys. Rev. D 51, 985 (1995)

[32] T. Agaki *et al.*, Phys. Rev. Lett. 67, 19, 2618 (1991)

[33] B. R. Martin *et al.*, Phys. Rev. D 2, 179 (1970)

[34] J. Rosner, EFI 94-21 (May 1994); A. Ali and D. London, CERN-TH.708/94, HEP-PH-9409399 (1994) and referneces therein

[35] K. E. Ohl *et al.*, Phys. Rev. Lett. 65, 12, 1407 (1990)

[36] G. D. Barr *et al.*, Phys. Lett. 240B, 283 (1990)

[37] A. S. Carroll et al. Phys. Rev. Lett. 44, 8, 525 (1980)

[38] L. Gibbons *et. al.,* Phys. Rev. Lett. 70, 1203 (1993)

[39] J. R. Patterson *et. al.,* Phys. Rev. Lett. 64, 1491 (1990)

[40] G. D. Barr *et. al.,* Phys. Lett. B317, 233 (1993)

[41] E. J. Ramberg *et. al.,* Phys. Lett. B338, 31 (1993)

[42] T. Nakaya *et. al.,* Phys. Rev. Lett. 73, 2169 (1994)

[43] D. Roberts *et. al.,* Phys. Rev. D. 50, 1874 (1994)

[44] P. Gu *et. al.,* Phys. Rev. Lett. 72, 3000 (1994)

[45] D. A. Harris *et. al.,* Phys. Rev. Lett. 71, 3914 (1993)

[46] D. A. Harris *et. al.,* Phys. Rev. Lett. 71, 3918 (1993)

[47] K. S. McFarland *et. al.,* Phys. Rev. Lett. 71, 31 (1993)

[48] M. Weaver *et. al.,* Phys. Rev. Lett. 72, 3758 (1994)

[49] P. Krolak *et. al.,* Phys. Lett. B320, 407 (1994)

[50] A. J Malensek, Fermilab Report FN-341 (1981)

[51] R. Coleman, Fermi National Accelerator Lab, private communication

[52] K. S. McFarland, University of Chicago Thesis (1993)

[53] L. K. Gibbons, University of Chicago Thesis (1993)

[54] D. A. Harris, University of Chicago Thesis (1993)

[55] J. Ritchie Patterson, University of Chicago Thesis (1990)

[56] D. Roberts, University of California at Los Angeles Thesis (1994)

[57] M. Weaver, University of California at Los Angeles Thesis (1994)

[58] P. Gu, Rutgers University Thesis (1994)

[59] D. Leurs *et al.*, Phys. Rev. 133B, 1276 (1964)

[60] D. A. Harris, private communication

[61] K. Arisaka *et al.*, KTeV Design Report, Fermilab Report FN-580 (1992)