Fermilab FERMILAB-THESIS-2000-22

THE UNIVERSITY OF CHICAGO

A MEASUREMENT OF THE BRANCHING RATIO AND FORM FACTOR OF $K_L \to \mu^+\mu^-\gamma$

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
GENE BREESE QUINN

CHICAGO, ILLINOIS
JUNE 2000

Copyright © 2000 by Gene Breese Quinn
All rights reserved

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES ix

ACKNOWLEDGMENTS x

ABSTRACT xiii

1 THE PHYSICS OF $K_L \to \mu^+\mu^-\gamma$ 1
- 1.1 The Long and the Short of It 1
- 1.2 $K_L \to \mu^+\mu^-$ and ρ 3
- 1.3 $K_L \to l^+l^-\gamma$ 8
  - 1.3.1 BMS Model 9
  - 1.3.2 DIP Parameterization 11
  - 1.3.3 Other Models 12
  - 1.3.4 $K_L \to \mu^+\mu^-\gamma$ vs. $K_L \to e^+e^-\gamma$ 13
- 1.4 Previous Measurements 15
- 1.5 KTeV Measurement Technique 17

2 KTEV EXPERIMENT 18
- 2.1 KTeV Beam 18
  - 2.1.1 Primary Proton Beam 18
  - 2.1.2 Secondary K_L Beam 19
- 2.2 Vacuum Decay Region 21
- 2.3 Charged Spectrometer 23
- 2.4 Trigger Hodoscopes 25
- 2.5 Cesium Iodide Calorimeter 26
  - 2.5.1 Crystal Array 26
  - 2.5.2 Readout 27
- 2.6 Muon Identification System 28
  - 2.6.1 Pb Wall, Hadron Anti, and Muon Filter 1 28
  - 2.6.2 Muon Filter 2 and MU2 29
  - 2.6.3 Muon Filter 3 and MU3 31
- 2.7 Photon Vetoes 31
- 2.8 Transition Radiation Detectors 33
- 2.9 Accidental Counters 33

3 EVENT SELECTION . . . 34
3.1 The Trigger . . . 34
3.1.1 Level 1 . . . 34
3.1.2 Level 2 . . . 35
3.1.3 Trigger Definitions . . . 35
3.1.4 Level 3 . . . 37
3.2 Data Set . . . 38
3.2.1 The E799II Run . . . 38
3.2.2 Split . . . 38
3.2.3 Crunch . . . 39

4 EVENT RECONSTRUCTION . . . 41
4.1 Tracking . . . 41
4.1.1 Hits, Pairs, and SODs . . . 41
4.1.2 Calibrations . . . 44
4.1.3 X and Y Track Candidates . . . 49
4.2 Clustering . . . 51
4.2.1 Crystal Energies . . . 51
4.2.2 Hardware Clusters . . . 52
4.2.3 Cluster Energy Corrections . . . 52
4.3 Track Vertex Finding . . . 53

5 MONTE CARLO SIMULATION . . . 56
5.1 Event Generation . . . 56
5.1.1 Kaon Production and Decay . . . 56
5.1.2 Decay Generators . . . 57
5.2 Particle Tracing . . . 58
5.3 Detector Simulation . . . 59
5.3.1 Drift Chambers . . . 59
5.3.2 Muon System . . . 66
5.3.3 CsI Calorimeter . . . 73
5.3.4 Photon Vetoes . . . 76
5.3.5 Trigger Hodoscopes . . . 77
5.3.6 Accidental Activity . . . 77
5.4 Trigger Simulation . . . 78
5.5 Monte Carlo Samples . . . 78

6 BRANCHING RATIO ANALYSIS . . . 80
6.1 Reconstruction . . . 81
6.2 Backgrounds . . . 85
6.2.1 Global Issues . . . 85
6.2.2 $K_L \to \pi^+\pi^-\pi^0$ . . . 86
6.2.3 $K_L \to \pi^+\pi^-\gamma$ and $K_L \to \pi^+\pi^-$ . . . 86

6.2.4 $K_L \to \pi^{\pm}\mu^{\mp}\nu$ ($K_{\mu 3}$) 86
6.2.5 $K_L \to \pi^{\pm}\mu^{\mp}\nu\gamma$ ($K_{\mu 3\gamma}$) 87
6.2.6 $K_L \to \pi^{0}\pi^{\pm}\mu^{\mp}\nu$ ($K_{\mu 4}$) 87
6.3 Signal and Normalization Analysis Cuts 88
6.3.1 ID Cuts 90
6.3.2 Data Quality Cuts 96
6.3.3 Background Rejection Cuts 102
6.4 Final Samples 112
6.4.1 Normalization and Flux Measurement 112
6.4.2 Background Estimation 117
6.4.3 Signal 118
6.5 Branching Ratio Calculation 122
6.6 Systematic Errors 124
6.6.1 Muon Efficiency 125
6.6.2 Drift Chamber Effects 126
6.6.3 Photon Energy (E_γ) Uncertainty 127
6.6.4 Analysis Cuts 127
6.6.5 Other 132
6.6.6 Summary 132

7 FORM FACTOR ANALYSIS 134
7.1 Direct $f(x)$ Measurement 134
7.2 α_{K^*} Measurement 136
7.2.1 $m_{\mu\mu}$ Shape Analysis 136
7.2.2 Branching Ratio Calculation 139
7.2.3 Combined Result 143
7.3 α Measurement 143
7.3.1 $m_{\mu\mu}$ Shape Analysis 143
7.3.2 Branching Ratio Calculation 145
7.3.3 Combined Result 145

8 CONCLUSION 147

A APPENDIX 154

REFERENCES 157

LIST OF FIGURES

1.1 Diagrams for short distance contributions to $K_L \to \gamma\gamma$ 2
1.2 Diagram for long distance pole contributions to $K_L \to \gamma\gamma$ 3
1.3 Diagrams for short distance contributions to $K_L \to \mu^+\mu^-$ 4
1.4 Diagram for long distance two-photon intermediate state contribution to $K_L \to \mu^+\mu^-$ 5
1.5 Long distance vector meson pole diagrams for $K_L \to \gamma^*\gamma$ 10
1.6 Direct weak pole diagram for $K_L \to \gamma^*\gamma$ 13
1.7 $K_L \to l^+l^-\gamma$ branching ratios as a function of α_{K^*} 14

2.1 Secondary beamline components 20
2.2 Plan view of the KTeV detector 22
2.3 Drift chamber wire geometry 24
2.4 V and V$'$ trigger hodoscope banks 25
2.5 Cesium iodide calorimeter array of 3100 crystals 26
2.6 Muon identification system plan view 29
2.7 Counter maps for MU2, MU3X, and MU3Y 30
2.8 Collar Anti geometry 32

4.1 Illustration of SOD classes 42
4.2 SOD distributions for KTeV and E773 43
4.3 TDC distribution for the first X plane of DC1 45
4.4 Straight-through muon alignment of DC2 47
4.5 Diagram of the corkscrew rotation between DC1 and DC2 48
4.6 Measurement of the corkscrew rotation 48
4.7 Reconstructed target position 50
4.8 Display of CsI energy corrections in a $K_L \to \mu^+\mu^-\gamma$ event 54

5.1 e-π SOD difference 61
5.2 High-SOD and single hit inefficiency DC maps 63
5.3 $\pi^+\pi^-\pi^0$ drift chamber illumination 64
5.4 $\pi^+\pi^-\pi^0$ vertex Z distribution 65
5.5 Muon efficiencies exhibiting cracks in MU3Y 68
5.6 MU3Y local muon efficiency agreement between Monte Carlo and data 69
5.7 Muon efficiency threshold vs. track momentum 71
5.8 Muon multiple scattering vs. track momentum 72

5.9 $K_L \to \pi^+\pi^-\pi^0$ mass distributions for two punch-throughs (PP), one decay and one punch-through (PD), and two pion decays (DD) . . . 74
5.10 Evaluation of pion decay/punch-through ratio through comparisons of $K_L \to \pi^+\pi^-\pi^0$ mass distributions in data and Monte Carlo 75

6.1 Reconstructed mass in the signal and normalization 84
6.2 Energy of clusters matched to tracks in 2MULD data and $K_L \to \mu^+\mu^-\gamma$ Monte Carlo 91
6.3 Cluster energy divided by track momentum (E/p) in 2TRK data and $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo 92
6.4 Track momentum distributions for $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \mu^+\mu^-\gamma$ Monte Carlo 93
6.5 Counter separation of hits in MU3Y 95
6.6 Reconstructed kaon momentum for $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \mu^+\mu^-\gamma$ 97
6.7 Reconstructed vertex Z position for 2TRK (top) and 2MULD (bottom) data 98
6.8 Vertex X and Y positions extrapolated to the CsI from the target . 100
6.9 Separation of track position and track cluster position at the CsI . . 101
6.10 Separation of track cluster and photon cluster positions 102
6.11 Photon energy cuts for signal and normalization 104
6.12 Photon cluster shape χ^2 for signal and normalization 105
6.13 Vertex χ^2 for signal and normalization 107
6.14 X Track segment offset (OFFMAGX) for signal and normalization . 108
6.15 Ring Counter maximum energy (RCMAX) for signal and normalization 110
6.16 Collar Anti maximum energy (CAMAX) for signal and normalization 111
6.17 P_t^2 for signal and normalization 113
6.18 P_t^2 vs. $m_{\mu^+\mu^-\gamma}$ for backgrounds to $K_L \to \mu^+\mu^-\gamma$ 114
6.19 Monte Carlo mass distributions for $K_L \to \mu^+\mu^-\gamma$ and $K_L \to \pi^+\pi^-\pi^0$ 115
6.20 Reconstructed mass distribution for $K_L \to \pi^+\pi^-\pi^0$ after final cuts . 116
6.21 Components of the background to $K_L \to \mu^+\mu^-\gamma$ 119
6.22 Estimations of the background to $K_L \to \mu^+\mu^-\gamma$ 120
6.23 Reconstructed mass distribution for $K_L \to \mu^+\mu^-\gamma$ after final cuts . . 121
6.24 P_t^2 vs. $m_{\mu^+\mu^-\gamma}$ in data after final cuts 123
6.25 Muon threshold with different filter thicknesses 126
6.26 $K_L \to \pi^+\pi^-\pi^0$ photon energy spectrum 128
6.27 $K_L \to \pi^+\pi^-\pi^0$ mass distribution 129
6.28 Study of P_t^2 cut 130
6.29 Track momentum distributions for $K_L \to \pi^+\pi^-\pi^0$ 131

7.1 Dimuon mass distribution for data and Monte Carlo with no form factor 135
7.2 x distribution acceptance correction 137

7.3 Dimuon mass distributions for Monte Carlo generated with various values of α_{K^*} 138
7.4 Measurement of α_{K^*} using a negative log-likelihood calculation . . . 140
7.5 BMS model of the $K_L \to \mu^+\mu^-\gamma$ branching ratio as a function of α_{K^*} 142
7.6 Measurement of α using a negative log-likelihood calculation 144
7.7 DIP model of the $K_L \to \mu^+\mu^-\gamma$ branching ratio as a function of α . 146

8.1 Comparison of the measured $BR(K_L \to \mu^+\mu^-\gamma)$ with various model predictions 148
8.2 Comparison of the α_{K^*} measurement with previous experiments . . . 150
8.3 Probability distributions for $|\mathrm{Re}A_{exp}|^2$ and $|\mathrm{Re}A_{LD}|$ 152

LIST OF TABLES

1.1 Previous $K_L \to l^+l^-\gamma$ experimental results 16

3.1 Level 3 cuts 37
3.2 Crunch cut efficiencies 39

4.1 Drift chamber resolutions for Run 8397 45
4.2 Target and CsI survey positions 49

5.1 Selection of distributions for muon energy loss fluctuations 70
5.2 Number of generated events for all Monte Carlo samples 79

6.1 Reconstruction cut efficiencies 82
6.2 Analysis cut efficiencies 89
6.3 $K_L \to \pi^+\pi^-\pi^0$ sample sizes and calculated flux 117
6.4 $K_L \to \mu^+\mu^-\gamma$ data and Monte Carlo sample sizes 122
6.5 Statistics of the $BR(K_L \to \mu^+\mu^-\gamma)$ calculation 124
6.6 Systematic errors 133

8.1 Self-consistency check of form factor parameter measurements 149

A.1 Bin-by-bin acceptance corrected x distribution 155
A.2 Bin-by-bin form factor measurement 156

ACKNOWLEDGMENTS

I have now spent one quarter of my life working at the University of Chicago on the KTeV experiment. I can think of no other place that I would rather have been these past eight years. It is rare these days for a graduate student to have the opportunity to see an HEP experiment through from beginning to end. Above all, I thank my adviser, Yau Wah, for that privilege. Wah has given me support, guidance, instruction, friendship, and tremendous freedom to carry out my work here. It has been a pleasure to learn with him.

Being surrounded daily by exceptional physicists has enriched my graduate experience. For this I acknowledge the talents of Wah, Bruce Winstein, Ed Blucher, and the strong kaon group that they built at Chicago. Their expertise and standards have always provided an example and challenge to rigorously attack hard problems.

A large debt of gratitude goes to the outstanding group of graduate students with whom I have shared an office. Colin Bown, Eric Zimmerman, Greg Graham, Steve Bright, Val Prasad, Jim Graham, and Peter Shawhan have each made huge contributions to the KTeV experiment and to my education. Special thanks go to Colin for working with me from the first day of orientation, through first year classes, the candidacy exam, and PSD softball, up to the final preparations the day before my defense.

Nearly 100 physicists contributed to the success of KTeV. I thank all of them collectively, and extend personal thanks to a few with whom I have worked closely over the last few years. Much of this analysis was pushed forward by the work of the other core members of the Dimuon Working Group: Julie Whitmore, Amit Lath, Rick Tesarek, and Masayoshi Sadamoto. In addition, I have received great benefit working and talking with Taku Yamanaka, Bob Tschirhart, Tony Barker, Leo Bellantoni, Rick Kessler, Bob Hsiung, Elliott Cheu, Nick Solomey, and John Jennings.

KTeV could never have been realized without a wealth of tehcnical support. I thank Brian LaVoy for his ready and willing technical assistance with my hardware projects in both New Muon and KTeV Halls. I am particularly grateful for John Krider's engineering expertise. Not only were his designs and leadership crucial to the success of the TRD system, but also he taught me much of what I know about building experimental apparatus.

I am also indebted to the previous generation of graduate students from E773 and E799I who got me started in kaon study, introduced me to the endless 'features' of the KTEVANA and KTEVMC code, and left a valuable resource in the body of their work. The members of my dissertation committee, Wah, Ed, John Rosner, and David Grier, have been very helpful to me in finalizing this document and I appreciate their assistance.

I thank my family for supporting me through 24 long years of schooling and for training me in the way I should go. My deepest appreciation and thanks go to my wife, Susanne. Her patience, understanding, support, laughter, and love have seen (and sometimes pulled) me through this process which I am blessed to share with her.

Finally, I thank God the Father and Son, by and through whom was created this fascinating universe. I am grateful for every day that I am given to experience and investigate the wonders of Creation.

G. Breese Quinn
Chicago, IL
May, 2000

To Susanne

and

Ad majorem Dei gloriam

ABSTRACT

We report on the analysis of the rare decay $K_L \to \mu^+\mu^-\gamma$ using the 1997 data from the KTeV experiment at Fermilab. A total of 9327 candidate events are observed with 2.4% background. We find that $BR(K_L \to \mu^+\mu^-\gamma) = (3.70 \pm 0.04_{stat} \pm 0.07_{syst}) \times 10^{-7}$. The BMS form factor parameter α_{K^*} is measured to be $\alpha_{K^*} = -0.163^{+0.026}_{-0.027}$. In addition, we make the first experimental measurement of the parameter α from the D'Ambrosio, Isidori, and Portolés form factor, finding $\alpha = -1.55 \pm 0.09$. This measurement of α limits the CKM parameter $\rho > -0.2$.

CHAPTER 1

THE PHYSICS OF $K_L \to \mu^+\mu^-\gamma$

1.1 The Long and the Short of It

The *long distance* and *short distance* physics involved in rare kaon decays continue to be of great interest to both theorists and experimentalists [1, 2, 3, 4, 5]. Short distance refers to processes that arise from the weak interaction, including QCD corrections. Typically, they are theoretically well-understood and sensitive to standard model parameters or new physics. Long distance processes involve low-energy electromagnetic interactions leading to two-photon (intermediate) states. These interactions are poorly understood and calculations of long distance rates are difficult and highly model-dependent.[1]

Almost all rare kaon decays with final or intermediate state photons (whether real or virtual), have significant long distance contributions. Among these decay modes are $K_L \to l^+l^-$, $K_L \to l^+l^-\gamma$, $K_L \to l^+l^-l^+l^-$, and $K_L \to \gamma\gamma$. All of these modes have in common a $K_L \to \gamma^{(*)}\gamma^{(*)}$ transition, where the $(*)$ denotes the possibility that the photon is off-shell. These transitions are the source of the long distance physics in the decay modes listed, and are the key to extracting information from short distance contributions.

Perhaps the clearest example of short and long distance effects identified in a kaon decay is from $K_L \to \gamma\gamma$. The dominant short distance contribution is shown in Figure 1.1. The amplitude for this process has been calculated using a free-quark-model [6, 7, 8, 9], and is given by Ma and Pramudita [9] as

$$A(K_L \to \gamma\gamma)_{SD} = \frac{G_F\alpha}{\pi} i f_K \epsilon_{\mu\nu\alpha\beta}\epsilon_1^\mu \epsilon_2^\nu q_1^\alpha q_1^\beta \sum_{j=u,c,t} \mathrm{Re}(V_{jd}V_{js}^*)[A_j^{(i)} + A_j^{(r)}], \qquad (1.1)$$

[1] Or provider-dependent if we extend the discussion to telecommunications.

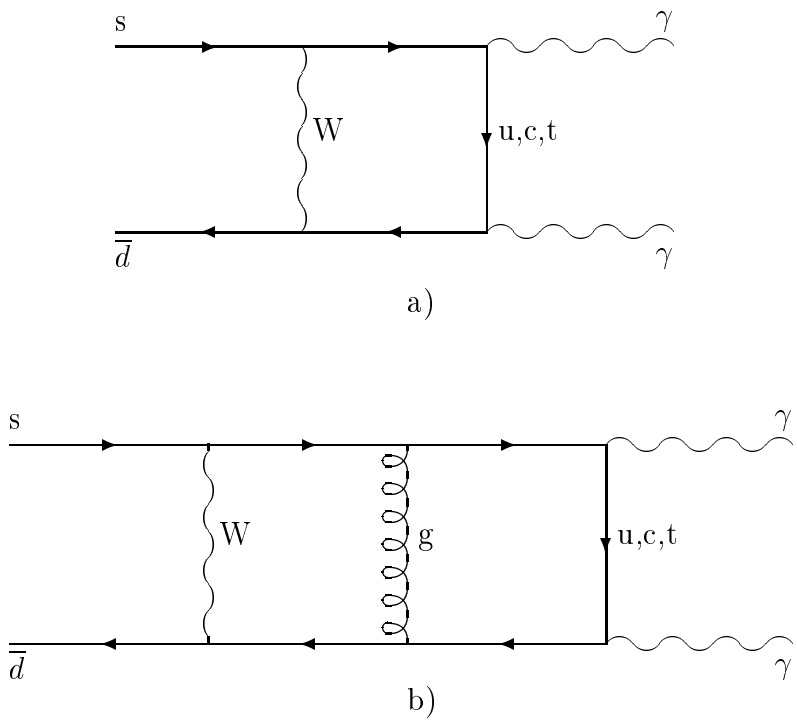

Figure 1.1: Diagrams for short distance contributions to $K_L \to \gamma\gamma$. a) the leading order electroweak box diagram, b) an example of QCD gluonic corrections (bottom) are shown.

where f_K is the kaon decay constant from $K^+ \to \pi^+\nu$, ϵ and q are the polarization and momentum vectors of the photons, V_{jk} are Cabibbo-Kobayashi-Maskawa (CKM) matrix elements [10], and A_j are loop integrals for the jth quark.[2] The result of comparing the calculated branching ratio using Equation 1.1 with the experimental value is

$$\frac{BR(K_L \to \gamma\gamma)_{SD}}{BR(K_L \to \gamma\gamma)_{EXP}} \leq 0.08.$$

Each of the free-quark-model calculations referenced found a limit at least as stringent. The natural scale of these interactions is set by the mass of the $W^\pm$ boson propagators. This leads to integration variables (and small corresponding distance scales) of order $\sim 1/m_W$, and the term 'short distance' to describe the processes.

[2]The (i) and (r) superscripts on A_j refer to one-particle irreducible diagrams such as Figure 1.1a, and one-particle reducible diagrams such as penguin diagrams that produce an effective $s\bar{d}\gamma$ vertex.

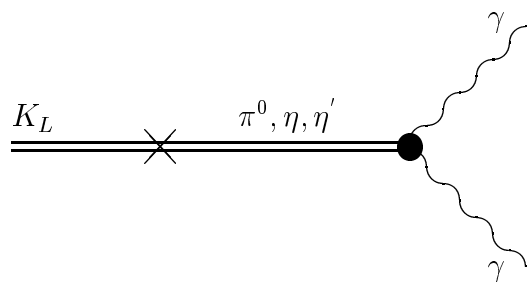

Figure 1.2: Diagram for long distance pole contributions to $K_L \to \gamma\gamma$.

Following the example of Gaillard and Lee [6], Ma and Pramudita turn to low-energy phenomenology to account for the rest of the $K_L \to \gamma\gamma$ rate. Both groups use a pseudoscalar pole model to make up the difference. The process involves pseudoscalar meson intermediate states, each contributing a pole term to the amplitude

$$A(K_L \to \gamma\gamma)_{LD} = \sum_{i=\pi^0,\eta,\eta'} \frac{A(P_i \to \gamma\gamma)\langle P_i|\mathcal{H}_W|K_L\rangle}{m_K^2 - m_i^2}, \tag{1.2}$$

where $\mathcal{H}_W$ is the non-leptonic weak interaction Hamiltonian. The pole term diagram is shown in Figure 1.2. Using known $P_i \to \gamma\gamma$ decay rates, and current algebra to derive the $K_L \to P_i$ amplitudes from $K_L \to 2\pi$, the total pole-term contribution is [9]

$$\frac{BR(K_L \to \gamma\gamma)_{LD}}{BR(K_L \to \gamma\gamma)_{EXP}} = 1.0 \pm 0.3.$$

The term long distance is applied to the low-energy pole model interactions because the meson propagators set a small mass scale on the order of 1 GeV.

1.2 $K_L \to \mu^+\mu^-$ and ρ

$K_L \to \mu^+\mu^-$ is a particularly interesting mode for studying short distance physics. The short distance contributions to this decay are generated by second-order weak interactions (Figure 1.3) [11, 12, 13]. Next-to-leading order QCD corrections have

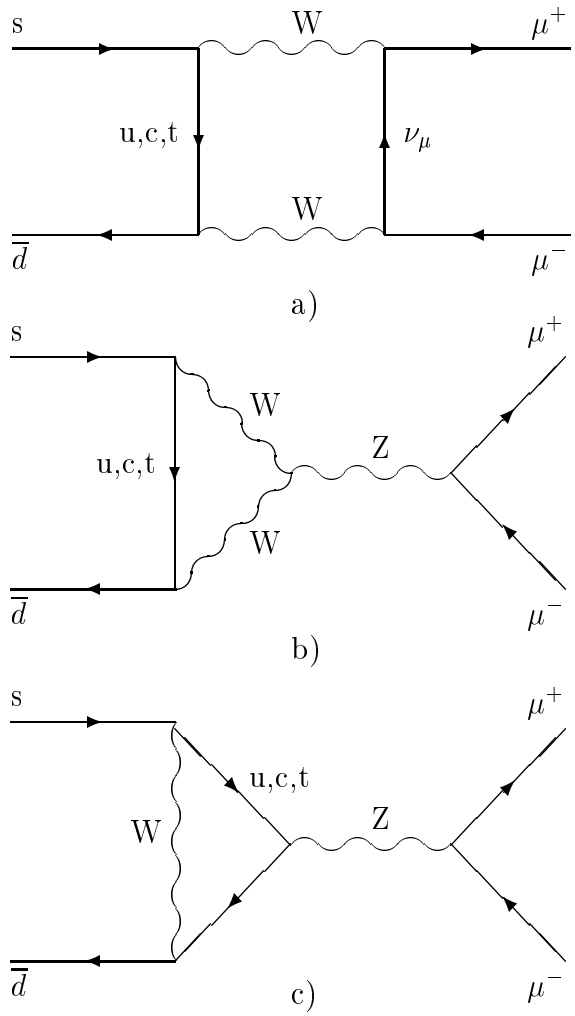

Figure 1.3: Diagrams for short distance contributions to $K_L \to \mu^+\mu^-$. The leading order electroweak box (a), and penguin (b,c) diagrams are shown without QCD corrections.

improved the theoretical estimations of these amplitudes, which are dominated by the top quark [14]. Because the top quark plays such an important role, the short distance contribution is very sensitive to the CKM matrix element V_{td}, and thus to the Wolfenstein parameter ρ [15]. In this respect, it is similar to the decay $K^+ \to \pi^\pm\nu\overline{\nu}$, which is also sensitive to V_{td}. But whereas $K^+ \to \pi^\pm\nu\overline{\nu}$ is theoretically very clean with only one observed event [16], $K_L \to \mu^+\mu^-$ has been measured to $< 3\%$ of

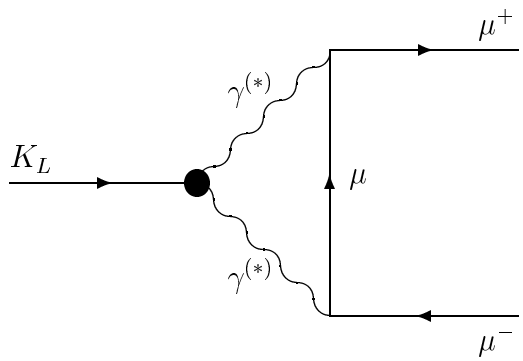

Figure 1.4: Diagram for long distance two-photon intermediate state contribution to $K_L \to \mu^+\mu^-$. $\gamma^{(*)}$ indicates that the photons can be either on or off shell.

itself but suffers from very complicated theoretical uncertainties. These uncertainties arise from the presence of long distance, two-photon intermediate states (Figure 1.4).

The branching ratio of $K_L \to \mu^+\mu^-\gamma$ can be decomposed in a simple manner according to

$$BR(K_L \to \mu^+\mu^-) \ = |\mathrm{Re}A|^2 + |\mathrm{Im}A|^2. \tag{1.3}$$

The absorptive portion of the amplitude ($|\mathrm{Im}A|^2$) represents the long distance diagram Figure 1.4 for the case where the two photons are real. In principle, other intermediate states such as $\pi\pi$ can contribute to the absorptive amplitude, but they have been shown to be negligible [17]. This component is calculated exactly [18, 19], independent of models, and is given by

$$\begin{aligned} |\mathrm{Im}A|^2 &= \frac{\alpha_{EM}^2}{2\beta_\mu}\left(\frac{m_\mu}{m_K}\right)^2\left(\ln\frac{1+\beta_\mu}{1-\beta_\mu}\right)^2 BR(K_L \to \gamma\gamma) \\ &= (7.07 \pm 0.18) \times 10^{-9} \end{aligned} \tag{1.4}$$

where $\beta_\mu = \sqrt{1 - 4m_\mu^2/m_K^2}$, and $BR(K_L \to \gamma\gamma) \ = (5.92 \pm 0.15) \times 10^{-4}$ [20]. This is called the unitarity bound and together with the measured $K_L \to \mu^+\mu^-$ branching ratio, $BR(K_L \to \mu^+\mu^-) \ = (7.18 \pm 0.17) \times 10^{-9}$ [21], limits the dispersive contribution

to

$$\begin{aligned} |\mathrm{Re}A_{exp}|^2 &= BR(K_L \to \mu^+\mu^-) \; - |\mathrm{Im}A|^2 \\ &= (1.1 \pm 1.8) \times 10^{-10} \\ &< 3.7 \times 10^{-10} \; (90\% \; \mathrm{C.L.}). \end{aligned} \tag{1.5}$$

Here the Bayesian scheme for eliminating the unphysical region has been used to set the upper limit [22]. Also, the error on the unitarity bound has been reduced by normalizing the $K_L \to \mu^+\mu^-$ rates to $K_L \to \pi^+\pi^-$ rather than to $K_L \to \gamma\gamma$ [21, 23].

Both short and long distance processes contribute to the dispersive amplitude, which can be written as

$$\mathrm{Re}A = \mathrm{Re}A_{SD} + \mathrm{Re}A_{LD}. \tag{1.6}$$

An effective Lagrangian describing the box and penguin diagrams of Figure 1.3 is used to estimate the short distance dispersive contribution [11]. Renormalization group calculations have been made accounting for QCD corrections to next-to-leading logarithm order [14]. These calculations greatly reduce theoretical uncertainties and lead to the branching ratio expression

$$|\mathrm{Re}A_{SD}|^2 = \frac{\alpha_{EM} BR(K^+ \to \pi^+\nu)}{\pi \sin^2\theta_W} \frac{\tau(K_L)}{\tau(K^+)} \frac{[\mathrm{Re}V_{cs}^* V_{cd} Y_{NL} + \mathrm{Re}V_{ts}^* V_{td} Y(x_t)]^2}{|V_{us}|^2} \tag{1.7}$$

where $x_t = m_t^2/m_W^2$ and

$$Y(x) = Y_0(x) + \frac{\alpha_s}{4\pi} Y_1(x) \tag{1.8}$$

$$Y_0(x) = \frac{x}{8}\left(\frac{4-x}{1-x} + \frac{3x}{(1-x)^2}\ln x\right). \tag{1.9}$$

$Y_0(x)$ is the pure electroweak 1-loop contribution from the top quark and $Y_1(x)$ is the 2-loop QCD correction. $Y_1(x)$ has terms similar to $Y_0(x)$, but to next-to-leading-log order. In Ref. [24], Buchalla and Buras determine that $Y(x)/Y_0(x) = 1.025$ for $m_t = 170$ GeV. Y_{NL} is the highly complicated renormalization group expression for the charm box and penguin contributions [14].

If the Wolfenstein parameterization of the CKM matrix is extended to $\mathcal{O}(\lambda^5)$

[25],

$$V = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda(1+i\eta A^2\lambda^4) & 1-\frac{\lambda^2}{2} < & A\lambda^2 \\ A\lambda^3(1-\overline{\rho}-i\overline{\eta}) & -A\lambda^2\left(1-\frac{\lambda^2}{2}(1-2\rho)+i\eta\lambda^2\right) & 1 \end{pmatrix}, \quad (1.10)$$

where

$$\overline{\rho} = \rho\left(1-\frac{\lambda^2}{2}\right) \qquad \overline{\eta} = \eta\left(1-\frac{\lambda^2}{2}\right), \quad (1.11)$$

then Equation 1.7 can be reduced to[3]

$$|\mathrm{Re}A_{SD}|^2 = 1.68\times 10^{-9} A^4 Y^2(x_t)\frac{(\overline{\rho}_0-\overline{\rho})^2}{\sigma}, \quad (1.12)$$

where

$$\overline{\rho}_0 = 1+\frac{Y_{NL}}{\lambda^4 A^2 Y(x_t)}, \qquad \sigma = \left(1-\frac{\lambda^2}{2}\right)^{-2}. \quad (1.13)$$

Finally, the approximate expression

$$|\mathrm{Re}A_{SD}|^2 = 0.9\times 10^{-9}(1.2-\overline{\rho})^2\left(\frac{\overline{m}_t(m_t)}{170\ \mathrm{GeV}}\right)^{3.1}\left(\frac{|V_{cb}|}{0.040}\right)^4 \quad (1.14)$$

is derived in Ref. [25] for $m_t = [150, 190]$ GeV. This combined with Equation 1.5 and Equation 1.6, yields a limiting expression for $\overline{\rho}$:

$$\overline{\rho} > 1.2 - \max\left[\frac{|\mathrm{Re}A_{exp}|+|\mathrm{Re}A_{LD}|}{3\times 10^{-5}}\left(\frac{\overline{m}_t(m_t)}{170\ \mathrm{GeV}}\right)^{-1.55}\left(\frac{|V_{cb}|}{0.040}\right)^{-2}\right], \quad (1.15)$$

where the relative sign between $|\mathrm{Re}A_{exp}|$ and $|\mathrm{Re}A_{LD}|$ is chosen to reflect the case of maximum destructive interference in order to set a lower bound.

It is clear from Equation 1.15 that the last remaining obstacle to extracting a limit for ρ is evaluating $|\mathrm{Re}A_{LD}|$. This requires the development of a method to handle the $K_L \to \gamma^*\gamma^*$ interaction. To begin, the two-photon intermediate state amplitude for $K_L \to \mu^+\mu^-$ can be written

[3] Using $\alpha = \frac{1}{129}, \sin^2\theta_W = 0.23, BR(K^+ \to \pi^+\nu) = 0.635$.

$$|\mathrm{Re}A_{LD}|^2 = \frac{2\alpha_{EM}\beta_\mu}{\pi^2 m_K^2} BR(K_L \to \gamma\gamma)|\mathrm{Re}R(m_K^2)|^2, \tag{1.16}$$

where

$$R(q^2) = \frac{2i}{\pi^2 q^2}\int d^4k \frac{q^2k^2 - (q\cdot k)^2}{k^2(q-k)^2[(p-k)^2 - m_\mu^2]} f(k^2, (q-k)^2), \tag{1.17}$$

and $p^2 = m_\mu^2$. The function $f(q_1^2, q_2^2)$ is the form factor describing the physics at the $K_L \to \gamma^*\gamma^*$ vertex, normalized to the $K_L \to \gamma\gamma$ amplitude [26, 27]. Equation 1.16 is model-independent, and assumes only CP conservation, and gauge and Lorentz invariance.

The challenge is to construct a suitable expression for $f(q_1^2, q_2^2)$. Various models for the form factor have been proposed, some of which will be considered in the following section. The natural mode to study the $K_L \to \gamma^*\gamma^*$ form factor is $K_L \to e^+e^-\mu^+\mu^-$, where both virtual photons convert to distinguishable lepton pairs. Unfortunately, this mode has a very small branching ratio and the largest sample consists of only 38 events [28]. The alternative is to use the high statistics modes, $K_L \to l^+l^-\gamma$. These modes probe the $K_L \to \gamma^*\gamma$ vertex, however, so any model must provide a mechanism for translating $K_L \to \gamma^*\gamma$ form factor measurements into meaningful results for the two off-shell photon case.

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

A general expression for the $K_L \to \gamma^*\gamma$ amplitude is [29, 30]

$$A(K_L \to \gamma^*(q_1, \epsilon_1)\gamma(q_2, \epsilon_2)) = iA_{\gamma\gamma} f(q_1^2)\epsilon_{\mu\nu\sigma\tau}\epsilon_1^\mu(q_1)\epsilon_2^\nu(q_2)q_1^\sigma q_2^\tau. \tag{1.18}$$

No assumptions have been made regarding the physics occurring at the vertex, other than it can only be a function of the virtual photon invariant mass, q_1^2. Therefore all of the long distance dynamics will be contained in the form factor $f(q_1^2)$. Extending the amplitude to $K_L \to l^+l^-\gamma$ and integrating over phase space in the final state gives the differential decay spectrum,

$$\frac{1}{,\,\gamma\gamma}\frac{d,\,(K_L \to l^+l^-\gamma)}{d(q^2)} = \frac{2\alpha_{EM}}{3\pi}\frac{|f(q^2)|^2}{q^2}\left(1-\frac{q^2}{m_K^2}\right)^3\left(1+\frac{2m_l^2}{q^2}\right)\left(1-\frac{4m_l^2}{q^2}\right)^{\frac{1}{2}}, \tag{1.19}$$

where the rate has been normalized to $K_L \to \gamma\gamma$. Models of the long distance physics are defined by the particular construction of the form factor, and distinctions between models will manifest as differences in the predicted decay rate. In principle, these differences could appear in both the decay spectrum relative to q_1^2, and the total integrated decay rate (i.e. branching ratio).

1.3.1 BMS Model

As was shown in Section 1.1, the $K_L \to \gamma\gamma$ rate is dominated by long distance pole contributions. Bergström, Massó, and Singer (BMS) use this as a starting point to approach the $K_L \to \gamma^*\gamma$ form factor [31].

The Vector Meson Dominance model (VMD) [18] predicts the pole term shown in Figure 1.5a. The amplitude for this process is simply

$$A_1(q^2) = \frac{|A_{\gamma\gamma}|}{1-q^2/m_\rho^2}. \tag{1.20}$$

It should be noted that this expression reduces to the $K_L \to \gamma\gamma$ amplitude as $q^2 \to 0$. This must be the case, and should be expected, since the disappearance of the vector meson pole reduces the diagram to that of Figure 1.2.

It is postulated that VMD must be extended to incorporate strangeness changing vector-vector transitions [31, 32]. This interaction produces a $K_L \to K^*\gamma$ vertex shown in Figure 1.5b. The amplitude

$$\begin{aligned} A_{K^*}(q^2) = \sqrt{8\pi\alpha_{EM}}G_{NL}f_{K^*K\gamma}\left(\frac{m_\rho^2}{f_{K^*}f_\rho^2}\right)\left(\frac{1}{1-q^2/m_{K^*}^2}\right) \\ \times\left(\frac{4}{3}-\frac{1}{1-q^2/m_\rho^2}-\frac{1}{9}\frac{1}{1-q^2/m_\omega^2}-\frac{2}{9}\frac{1}{1-q^2/m_\phi^2}\right) \end{aligned} \tag{1.21}$$

is calculated using a phenomenological octet Lagrangian from Sakurai [33].[4] The

[4] G_{NL} and the three meson form factor constants are given in Section 7.2.2.

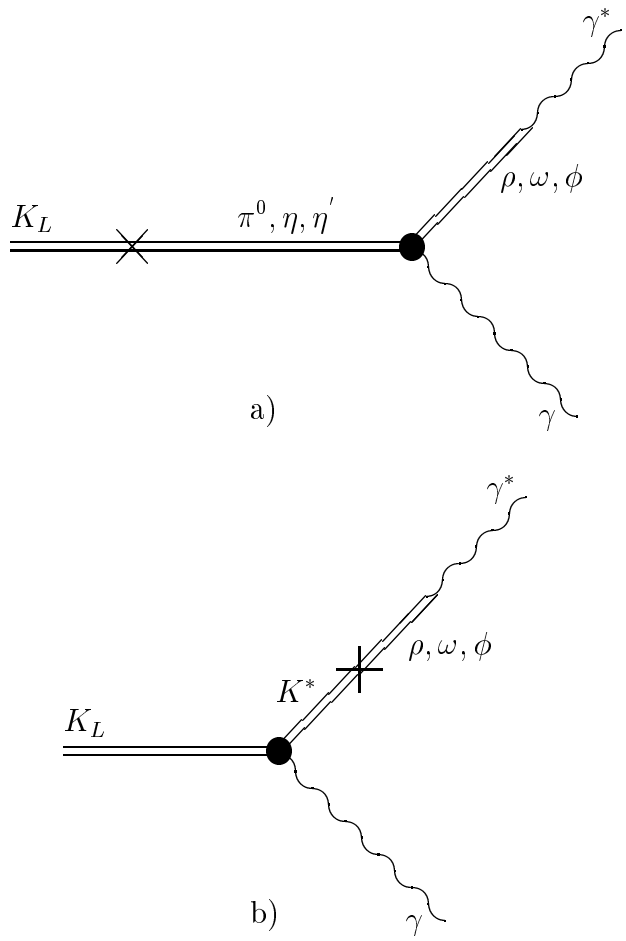

Figure 1.5: Long distance vector meson pole diagrams for $K_L \to \gamma^*\gamma$. a) standard Vector Meson Dominance contribution, b) vector-vector transition from weak interaction in strange decays.

total $K_L \to \gamma^*\gamma$ amplitude is then

$$A_{\gamma^*\gamma}(q^2) = A_1(q^2) + \alpha_{K^*} A_{K^*}(q^2), \tag{1.22}$$

where α_{K^*} is a parameter that measures the relative strength of the K^* contribution.

A strict interpretation of Sakurai's model corresponds to $|\alpha_{K^*}| = 1$. However, BMS perform a quark model calculation based on a QCD corrected Hamiltonian from Shifman, Vainshtein, and Zakharov [34]. There it is shown that penguin di-

agrams similar to Figure 1.3b,c that enhance $\Delta I = 1/2$ pseudoscalar-pseudoscalar transitions do not participate in vector-vector transitions. This lack of enhancement leads to a prediction of $|\alpha_{K^*}| = 0.2 - 0.3$.

A comparison of Equation 1.22 with Equation 1.18 shows that the BMS form factor is

$$f(q^2) = \frac{A_{\gamma^*\gamma}(q^2)}{|A_{\gamma\gamma}|}. \tag{1.23}$$

To connect this form factor to $f(q_1^2, q_2^2)$ from the $K_L \to \gamma^*\gamma^*$ vertex, BMS and Wyler evaluate $A_{\gamma^*\gamma}(q^2)$ for the case of one virtual photon originating from a vector meson ($PV\gamma$ vertex). Then they extrapolate to the case where both virtual photons are coupled to vector mesons (PVV vertex). In this way they derive bounds for the long distance dispersive amplitude [35]:[5]

$$(-5.1 - 11.0\alpha_{K^*}) \times 10^{-5} < \mathrm{Re}A_{LD} < (-2.9 - 10.2\alpha_{K^*}) \times 10^{-5}. \tag{1.24}$$

The maximum value of $|\mathrm{Re}A_{LD}|$ from either bound of this equation then can be used to set a lower limit for ρ via Equation 1.15.

1.3.2 DIP Parameterization

D'Ambrosio, Isidori, and Portolés (DIP) attempt to avoid model dependences associated with building a form factor expression. Instead of making assumptions about any particulars of the long distance processes, they parameterize the $K_L \to \gamma^*\gamma^*$ form factor in the most general way compatible with a Chiral Perturbation Theory expansion [27]:

$$f(q_1^2, q_2^2) = 1 + \alpha\left(\frac{q_1^2}{q_1^2 - m_\rho^2} + \frac{q_2^2}{q_2^2 - m_\rho^2}\right) + \beta\frac{q_1^2 q_2^2}{(q_1^2 - m_\rho^2)(q_2^2 - m_\rho^2)}. \tag{1.25}$$

The formulation does include the pole behavior characteristic of long distance physics even though no specific diagrams are assumed. By setting $q_2^2 = 0$ (i.e. making the

[5]The result is slightly different than in Ref. [35] because the input parameters have been updated as described in Section 7.2.2.

second photon real), $f(q_1^2, q_2^2)$ reduces to the $K_L \to \gamma^*\gamma$ form factor, $f(q^2, 0)$. The parameter α can be measured directly through $K_L \to l^+l^-\gamma$ and β is probed with $K_L \to e^+e^-\mu^+\mu^-$. The relation between α and the BMS form factor is found by expanding Equation 1.23 in powers of q^2/m_ρ, giving

$$\alpha = -1 + (3.1 \pm 0.5)\alpha_{K^*}. \tag{1.26}$$

Using Equation 1.25 in Equation 1.16 results in the expression

$$|\mathrm{Re}A_{LD}| = \left[\frac{2\alpha_{EM}\beta_\mu BR(K_L \to \gamma\gamma)}{\pi^2 m_K^2}\right]^{1/2} |5.25 + 3.47\alpha + 3(1 + 2\alpha + \beta)\ln\frac{\Lambda}{m_\beta}|, \tag{1.27}$$

where Λ is an ultraviolet cutoff. The last term is bounded by perturbative QCD at high q^2 to be

$$|(1 + 2\alpha + \beta)|\ln\frac{\Lambda}{m_\beta} < 0.4. \tag{1.28}$$

All that remains at this point is to obtain a measured value for α. With that in hand, bounds for $|\mathrm{Re}A_{LD}|$ and ρ can be determined.

1.3.3 Other Models

Two other models are based on pole contributions similar to those from BMS. Pure VMD ignores the $KK^*\gamma$ contribution and considers diagrams only of the type in Figure 1.5a. This is equivalent to BMS with $\alpha_{K^*} = 0$. Ko proposes the inclusion of the additional pole diagram shown in Figure 1.6 [36]. This contribution is characterized by a direct weak $KV\gamma$ vertex that is absent in BMS. Also, Ko uses a completely different chiral Lagrangian that treats the vector mesons as gauge bosons. This model has no free parameters and predicts $BR(K_L \to \mu^+\mu^-\gamma)/BR(K_L \to \gamma\gamma) = (7.45^{+0.54}_{-0.15}) \times 10^{-4}$. In contrast, both the BMS and DIP treatments predict the $K_L \to l^+l^-\gamma$ branching ratios as a function of $\alpha_{K^*}(\alpha)$. Gvozdev *et.al.* make a quark-model calculation that predicts $BR(K_L \to \mu^+\mu^-\gamma)/BR(K_L \to \gamma\gamma) = (9.9 \pm 0.4) \times 10^{-4}$ [37]. However, it is noted that varying the input parameters to the calculation can change the estimate to $\sim 8.3 \times 10^{-4}$.

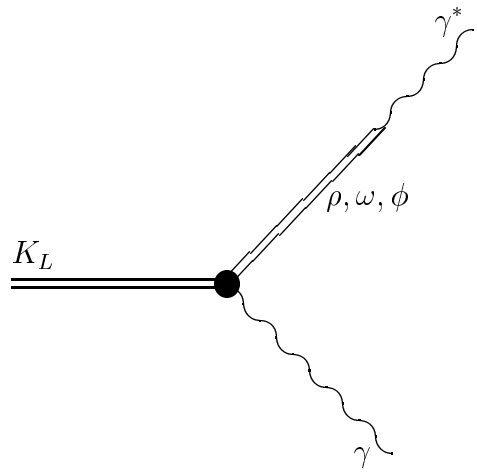

Figure 1.6: Direct weak pole diagram for $K_L \to \gamma^*\gamma$.

1.3.4 $K_L \to \mu^+\mu^-\gamma$ *vs.* $K_L \to e^+e^-\gamma$

To determine the relative merits of the $K_L \to l^+l^-\gamma$ modes for probing the $K_L \to \gamma^*\gamma$ vertex, one must consider how the form factor is measured experimentally. Equation 1.19 provides much information on this question. For the time being, ignore the form factor and consider phase space. There is a low q cutoff at twice the lepton mass. This means that for $K_L \to \mu^+\mu^-\gamma$, the invariant mass of the virtual photon is cut off much higher than for $K_L \to e^+e^-\gamma$, where the off-shell photon can extend down almost to $q^2 = 0$. Coupled with the presence of a $1/q^2$ dependence, $K_L \to e^+e^-\gamma$ will have a much higher decay rate, and the m_{ee} spectrum will be heavily weighted to low q^2.

As was mentioned earlier, one of the ways to measure the form factor is to study the distribution of $K_L \to l^+l^-\gamma$ decays with respect to q^2. In one respect, $K_L \to e^+e^-\gamma$ is hampered in this analysis by the fact that most of the data will be at low q^2. Namely, both the BMS and DIP form factors approach unity as $q^2 \to 0$. Therefore the $K_L \to e^+e^-\gamma$ data will primarily occupy the region least sensitive to the form factor parameters α_{K^*} and α.

The second way to investigate the form factor is by measuring the branching ratio, or integrated decay rate. The dearth of high q^2 $K_L \to e^+e^-\gamma$ data is especially

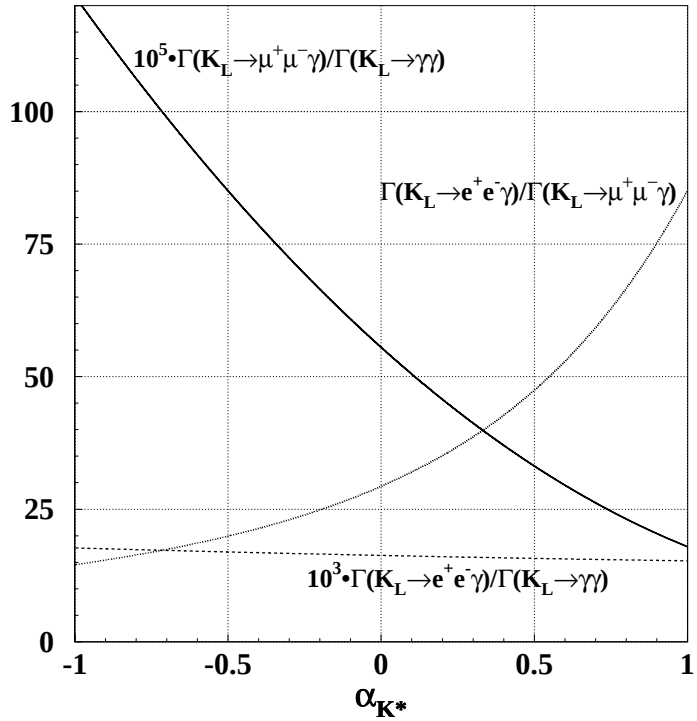

Figure 1.7: $K_L \to l^+l^-\gamma$ branching ratios as a function of α_{K^*}. Solid: $K_L \to \mu^+\mu^-\gamma$ branching ratio normalized to $K_L \to \gamma\gamma$, Dashed: $K_L \to e^+e^-\gamma$ branching ratio normalized to $K_L \to \gamma\gamma$, Dotted: ratio of $K_L \to e^+e^-\gamma$ and $K_L \to \mu^+\mu^-\gamma$ branching ratios.

acute for this measurement, because what little high mass sensitivity exists is washed out by integrating over the whole spectrum. $K_L \to \mu^+\mu^-\gamma$ does not suffer from this problem since all of the data occur in a region where the form factor terms proportional to α_{K^*}(or α) are significant. This effect can be seen in a plot of the $K_L \to l^+l^-\gamma$ branching ratios predicted by the BMS model (Figure 1.7). There it is evident that the $K_L \to \mu^+\mu^-\gamma$ branching ratio is strongly dependent upon the form factor, while the $K_L \to e^+e^-\gamma$ rate is almost completely insensitive to α_{K^*}.

An unrelated experimental difficulty compounds the challenge to measure the $K_L \to e^+e^-\gamma$ form factor. Very large radiative corrections from $K_L \to e^+e^-\gamma\gamma$ have to be made to any $K_L \to e^+e^-\gamma$ measurement [38, 39]. These corrections can be

difficult to understand and simulate.[6] This has little effect on $K_L \to \mu^+\mu^-\gamma$, where radiative corrections are minimal due to the much larger lepton masses.

In spite of the challenges, $K_L \to e^+e^-\gamma$ has historically been the primary probe into the $K_L \to \gamma^*\gamma$ vertex, based solely on statistics. However, in the current generation of experiments, many thousands of $K_L \to \mu^+\mu^-\gamma$ events are observed and statistics in the high dilepton mass region are comparable between the two modes. This coupled with form factor sensitivity in both the mass spectrum *and* branching ratio makes $K_L \to \mu^+\mu^-\gamma$ the best choice for studying long distance physics. It should be noted that the two separate form factor measurements available in $K_L \to \mu^+\mu^-\gamma$ provide an important consistency check of the different models.

1.4 Previous Measurements

Evidences of both $K_L \to l^+l^-\gamma$ modes were made by Carroll *et.al.* at Brookhaven in 1980 [40]. In that experiment, 4 $K_L \to e^+e^-\gamma$ events and 1 $K_L \to \mu^+\mu^-\gamma$ event were seen. The first form factor measurements from $K_L \to e^+e^-\gamma$ came from Brookhaven and CERN in 1990. The B845 and NA31 experiments reported nearly identical measurements of $\alpha_{K^*}$and the $K_L \to e^+e^-\gamma$ branching ratio based on samples of about 1000 events each [38, 41]. B845's result of $\alpha_{K^*} = -0.280^{+0.099}_{-0.090}$ includes a 56% radiative correction. The CERN NA48 experiment recently published a $K_L \to e^+e^-\gamma$ measurement from a sample of nearly 7000 events. Those results, including $\alpha_{K^*} = -0.36 \pm 0.06$, are consistent at the 1σ level with the previous experiments [42], and the radiative correction to α_{K^*} is 140%.

The only complete form factor analysis from $K_L \to \mu^+\mu^-\gamma$ to date was reported by the Fermilab experiment E799I in 1995 with a final sample of 197 events [43]. The measured branching ratio is $BR(K_L \to \mu^+\mu^-\gamma) = (3.23 \pm 0.30) \times 10^{-7}$. The result for α_{K^*}, combining measurements from the branching ratio and dimuon mass distribution, is $\alpha_{K^*} = -0.028^{+0.115}_{-0.111}$. The central values from the shape and branching ratio measurements are noticeably different, but well within the large errors of the

[6] A testimony to this is the fact that entire conferences are devoted to the general subject of radiative corrections (e.g. RADCOR-2000 - The 5th International Symposium on Radiative Corrections, Sep. 11-15, 2000).

Table 1.1: Previous $K_L \to l^+l^-\gamma$ experimental results.

Experiment	Year	# Events	Branching Ratio	$\alpha^{\dagger}_{K^*}$
$K_L \to e^+e^-\gamma$:				
Carroll, *et.al.*	80	4	$(17.4 \pm 8.7) \times 10^{-6}$	-
B845	90	919	$(9.1^{+0.8}_{-0.7}) \times 10^{-6}$	$-0.280^{+0.099}_{-0.090}$
NA31	90	1053	$(9.2 \pm 0.7) \times 10^{-6}$	-0.28 ± 0.13
NA48	99	6854	$(10.6 \pm 0.5) \times 10^{-6}$	-0.36 ± 0.06
$K_L \to \mu^+\mu^-\gamma$:				
Carroll, *et.al.*	80	1	$(2.8 \pm 2.8) \times 10^{-7}$	-
E799I	95	197	$(3.23 \pm 0.30) \times 10^{-7}$	$-0.028^{+0.115}_{-0.111}$ ($-0.018^{+0.131}_{-0.123}$, BR) ($-0.13^{+0.27}_{-0.35}$, Shape)
NA48	97	45	$(3.4 \pm 0.7) \times 10^{-7}$	$-0.04^{+0.24}_{-0.21}$, BR

$^{\dagger}K_L \to e^+e^-\gamma$ results for α_{K^*} are all measured from the q^2 distribution shape. $K_L \to \mu^+\mu^-\gamma$ measurements of α_{K^*} using the branching ratio or shape only are noted.

shape analysis. A $K_L \to \mu^+\mu^-\gamma$ measurement from an early 1995 NA48 run [39] uses 45 events to measure α_{K^*} from the branching ratio. That analysis finds $\alpha_{K^*} = -0.04^{+0.24}_{-0.21}$. All $K_L \to l^+l^-\gamma$ measurements are summarized in Table 1.1.

It is particularly interesting to note that the $K_L \to \mu^+\mu^-\gamma$ and the latest $K_L \to e^+e^-\gamma$ values for α_{K^*} are so different. The form factor should be identical for the two modes. It is unclear whether the difference is due to statistical fluctuation, an experimental bias between the electron and muon Dalitz modes, or a more fundamental problem.

1.5 KTeV Measurement Technique

The goal of this analysis is to make high precision measurements of the $K_L \to \mu^+\mu^-\gamma$ branching ratio, and the form factor parameters α_{K^*} and α. To do this, a high statistics sample of reconstructed $K_L \to \mu^+\mu^-\gamma$ decays is required. KTeV features a high K_L flux and very good charged track reconstruction, photon energy measurement, and muon identification. These characteristics make KTeV an ideal experiment in which to study $K_L \to \mu^+\mu^-\gamma$.

To calculate the $K_L \to \mu^+\mu^-\gamma$ branching ratio, the signal events must be normalized to another kaon decay mode with a well-known branching ratio. The final state of the normalization mode should be as similar to the signal mode as possible, in order to minimize the effects of any systematic biases in the measurement technique. Ideally, the normalization would include muons in the final state. However, there is no dimuon kaon decay with a branching ratio larger than that of $K_L \to \mu^+\mu^-\gamma$. Therefore, the normalization mode chosen is $K_L \to \pi^+\pi^-\pi^0$. This three-body decay is very similar to $K_L \to \mu^+\mu^-\gamma$, especially with regards to the charged particle kinematics. However, since there are no muons in the normalization mode, the absolute acceptance of muons in the experiment must be determined extremely well. This necessitates very good understanding and simulation of muon propagation through the detector. $BR(K_L \to \pi^+\pi^-\pi^0)$ is known to 1.6% of itself, and this sets the lower limit to the systematic error of the $BR(K_L \to \mu^+\mu^-\gamma)$ measurement.

The form factor parameters will be measured using a technique similar to that used in the E799I measurement of α_{K^*} [44]. Both α_{K^*} and α will be calculated from the branching ratio result and measured from the dimuon mass (q^2) distribution of the $K_L \to \mu^+\mu^-\gamma$ events.

CHAPTER 2
KTEV EXPERIMENT

KTeV is an experiment at Fermilab that was designed to study CP violation in the neutral kaon system. It is actually two experiments in one: E832 and E799II. The goal of E832 is to search for direct CP violation by measuring $\epsilon^{'}/\epsilon$ at the 10^{-4} level [45]. The purpose of E799II is to study CP violation and other interesting physics in rare kaon decays. This analysis is based on data from the E799II run in 1997.

Both experiments utilize two nearly parallel neutral kaon beams produced by a primary proton beam from the Fermilab Tevatron. Two beams are required for E832, which passes one of the two K_L beams through a regenerator to produce the K_S necessary for $\epsilon^{'}$ studies. For E799II, no regenerator is used, resulting in two identical K_L beams.

The KTeV detector was designed to facilitate observation and reconstruction of a variety of kaon decay final states. It includes a charged particle spectrometer with four drift chambers, a high-precision cesium iodide calorimeter, and transition radiation detectors and muon counters for particle identification.

2.1 KTeV Beam

2.1.1 Primary Proton Beam

During KTeV running, the Tevatron delivered a primary beam of 800 GeV protons. The protons are delivered in a one minute spill cycle consisting of a 40 second off-spill period in which the protons are accelerated, and 20 seconds on-spill when the 800 GeV protons are routed to the various fixed-target experiments. The beam has a 53 MHz RF structure, to which the KTeV trigger timing is synchronized. Protons

arrive every 19 ns in "buckets" which are 1-2 ns wide. KTeV typically received from 2.5×10^{12} to 5.0×10^{12} protons per spill.

The proton beam is directed onto a BeO target where it produces kaons (and many other types of particles) which are formed into a secondary beam. The target is 30 cm long, or 1.1 interaction lengths. It has a transverse cross section of 3 mm $\times$ 3 mm, and the beam spot on the target has a 250 μm RMS in both the horizontal and vertical directions. The primary protons are incident on the target at a targeting angle of 4.8 mrad below the horizontal secondary beam direction. This angle is chosen to reduce the relative neutron flux, which peaks at 0 mrad.

2.1.2 Secondary K_L Beam

Secondary neutral kaon beams are made by magnetically sweeping away charged particles downstream of the target, and collimating the remaining neutral particles into horizontal beams. The center of the target serves as the origin of the KTeV coordinate system. The positive Z axis is downstream along the secondary kaon beam direction, positive Y is the upward vertical direction, and positive X is horizontally transverse to the beam making a right-handed system (i.e., to the left facing downstream). Figure 2.1 shows the beamline configuration.

The first charged particle sweeper is a magnet just 2 m downstream of the target. This "target sweeper" primarily deflects the residual proton beam downward into a water-cooled copper beam dump below the beam. The main purpose of the remaining sweeping magnets is getting rid of muons, which would otherwise cause high-rate problems in the KTeV detector. "μ-sweep1" has a magnetic field that pushes charged particles horizontally away from the beam. Most muons, which are deflected *away* from the primary beam dump by the target sweeper, are swept to the side by μ-sweep1. Further downstream is "μ-sweep2", which also deflects particles to the sides. μ-sweep2 removes charged particles from interactions in the lead absorber or primary collimator, and residual scattered muons. The "final sweeper" magnet is located at $Z = 90$ m, just before the vacuum decay region. The job of the final sweeper is to sweep out charged particles from decays upstream in the beamline or from collimator interactions.

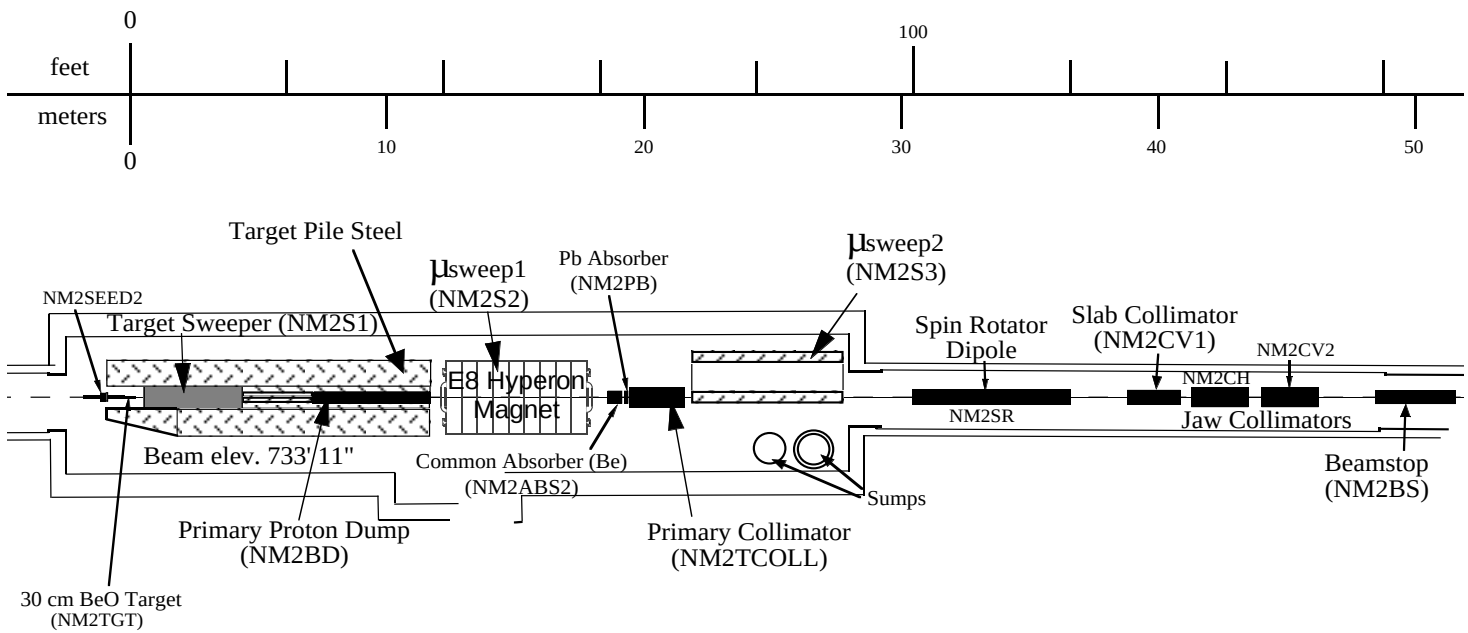

Figure 2.1: Secondary beamline components in the NM2 enclosure upstream of the KTeV detector. The common absorber is not used in E799II, the spin rotator dipole is used for hyperon physics, and the jaw collimators are for special studies.

Formation of the neutral beams begins with the lead absorber at $Z = 18.5$ m, which converts photons to e^+e^- pairs that are subsequently swept away. Beam shaping then starts with the primary collimator. The primary collimator is a 2 m long steel and brass block, with two square holes separated by 1.6 mrad relative to the target. At $Z = 28$ m (just downstream of μ-sweep2), the two nearly parallel neutral beams enter a vacuum region which extends to $Z = 159$ m, just prior to the charged spectrometer. The next beamline component is the slab collimator, which sits between the two beams. This element stops scattered beam particles from crossing between the beams. As the name implies, the defining collimator defines the kaon beam shapes. It is a tungsten block with square, tapered beam holes, the upstream edges of which define the beam cross section. For the winter E799II running, the beams exiting the defining collimator were 4.4 cm square. A larger defining collimator was used in the summer, producing beams 5.2 cm square. The slab collimator was removed for the summer E799II running when the larger beams were used.

The final beams entering the vacuum decay region typically had 25 to 50 MHz of neutral hadrons. The composition of the beams is dominated by neutrons and kaons, with neutrons predominating by 3 to 1. Hyperons make up a little less than 1% of the beams. Muon beams useful for detector calibration can be produced by placing a beamstop in the beamline to absorb all neutral and other charged particles.

2.2 Vacuum Decay Region

A 66 m long vacuum decay tank is the first part of the KTeV detector, shown in Figure 2.2. The tank is an extension of the vacuum pipe containing the K_L beams which begins at $Z = 28$ m. A vacuum decay region minimizes beam interactions with matter; with the tank pressure kept below 10^{-6} torr, the problem is virtually eliminated. The only kaon decays accepted for analysis are those which occur within the vacuum decay region extending from $Z = 93$ m to the downstream vacuum window at $Z = 159$ m. Because the K_L lifetime is so long ($\sim 5.17 \times 10^{-8}$ s [20]), only 5% of the K_L which enter the tank decay before leaving the vacuum region.

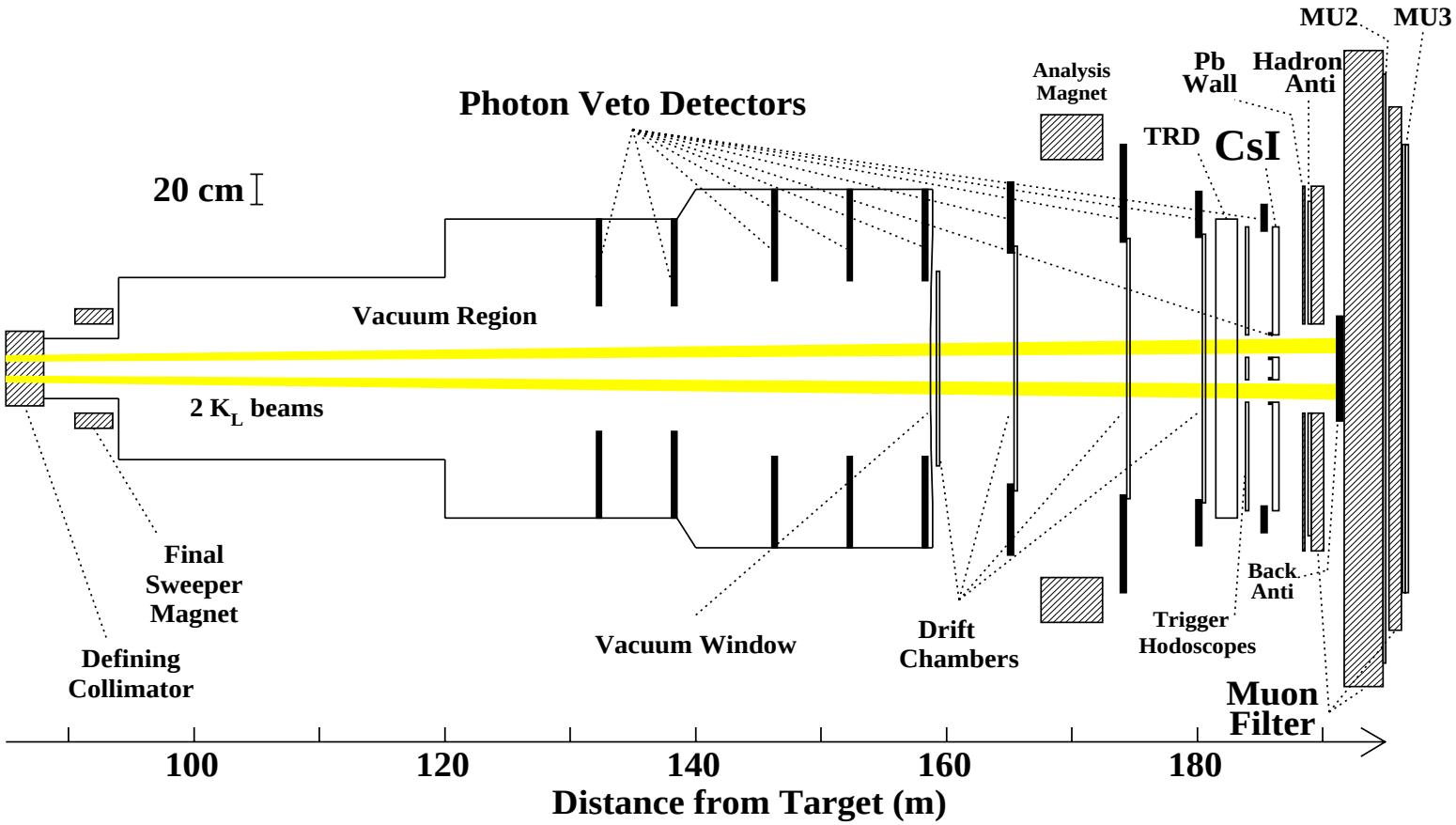

Figure 2.2: Plan view of the KTeV detector. The horizontal scale along the Z direction is compressed.

An exceedingly strong vacuum window is required to contain the sub-mtorr vacuum in such a large volume. However, the window must also be extremely thin to minimize photon conversions and bremsstrahlung. A composite material of 0.023" Kevlar laminated to 0.010" Mylar meets both requirements. The Mylar provides a tight gas seal, and the Kevlar provides the strength necessary to hold the 30 tons of differential pressure over 27.5 square feet of surface area. The total window thickness is only 0.0015 radiation lengths.

2.3 Charged Spectrometer

The KTeV charged spectrometer uses four drift chambers separated by an analysis magnet to measure charged track trajectories and momenta. Each drift chamber measures the position of charged particle tracks in the X and Y views. The analysis magnet sits between the second and third chambers (DC2 and DC3) and has a field strength of ~ 2 kilogauss oriented in the vertical direction. This field imparts a 205 MeV/c momentum kick in the horizontal direction to charged particles. Plastic bags filled with helium are placed before, between, and behind the chambers to reduce multiple scattering, photon conversions, and beam interactions.

The drift chamber frames and geometry were used in the previous experiments E731, E773, and E799I [46, 47, 48]. For KTeV, the chambers were re-strung with 0.001" diameter gold-plated tungsten sense wires, and 0.004" diameter gold-plated aluminum field wires. Aluminum was chosen for the larger field wires to reduce multiple scattering.

In Figure 2.3, the drift chamber geometry is diagrammed. Each chamber has a pair of Y view sense wire planes followed by a pair of X view sense wire planes. The sense wires are surrounded by field-shaping wires in a hexagonal geometry that provides nearly cylindrical field symmetry. The vertical X view wires shown in Figure 2.3 measure track positions along the X direction. The two planes of the pair are offset by half a wire spacing. This resolves the two-sided position ambiguity of a single hit wire.

The chambers are filled with a 50/50 mix of argon and ethane plus a small

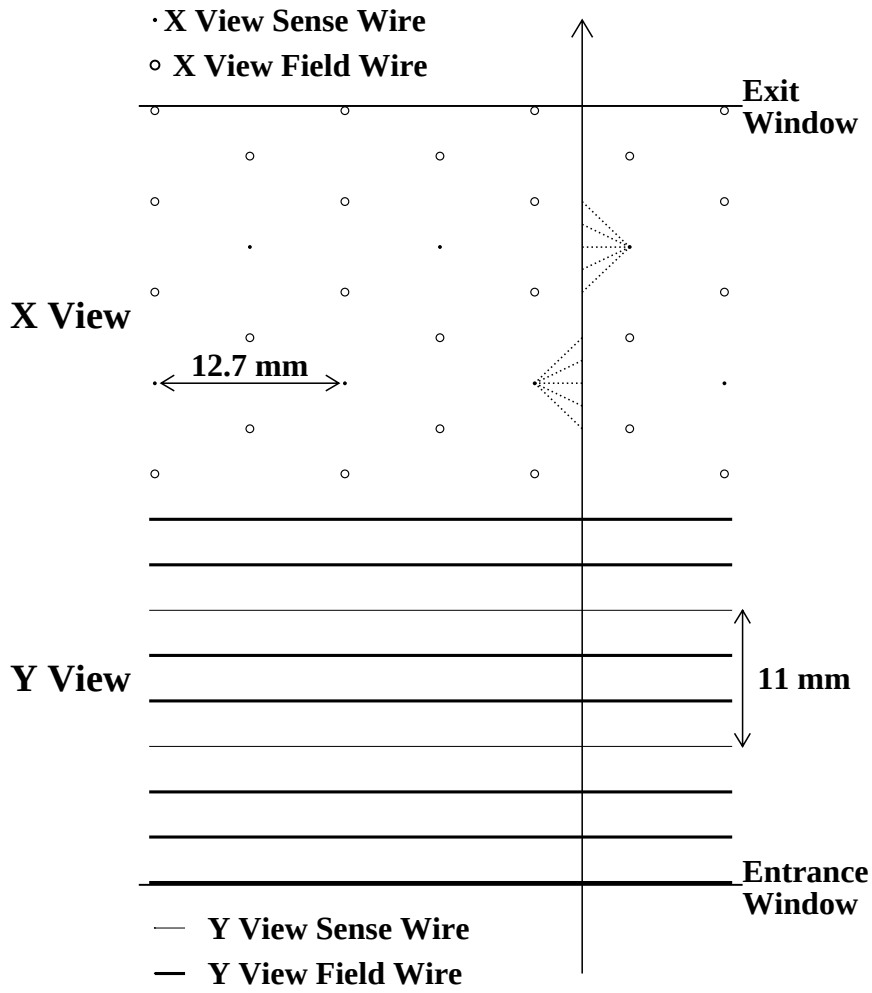

Figure 2.3: Drift chamber wire geometry. The paths of a charged particle and ionization drift electrons are represented by the arrow and dotted lines, respectively.

concentration ($\sim$ 1% by volume) of isopropyl alcohol. The alcohol slows chamber aging by absorbing damaging ultraviolet light. The drift chambers are operated at high-voltage settings of -2450 V to -2600 V on the cathode field wires. This gas mix and voltage level produces drift times of $\sim$ 50 μm/ns. The signals collected on the anode sense wires are amplified and discriminated with chamber-mounted electronics, and then digitized with LeCroy 3377 time-to-digital converters (TDCs). The TDC times have a precision of 0.5 ns. Individual hit positions are measured with a resolution of 100 μm, and the momentum resolution of the spectrometer is

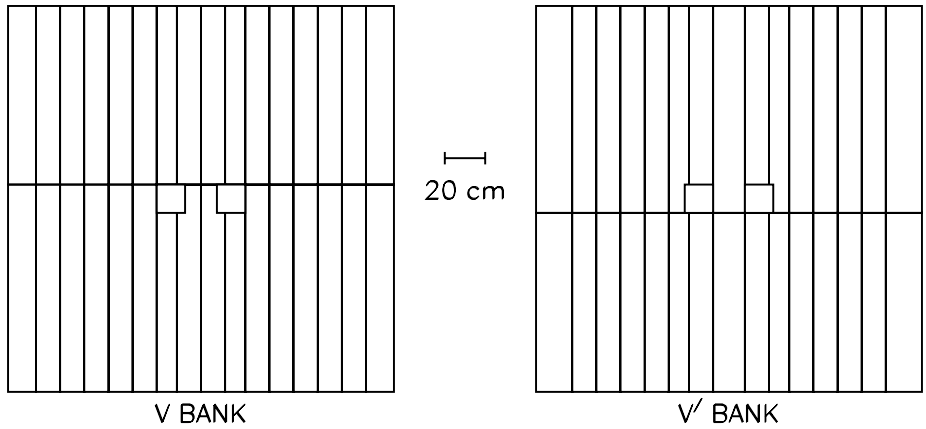

Figure 2.4: V and V$'$ trigger hodoscope banks.

$$\frac{\sigma(P)}{P} = 0.38\% \oplus 0.016\% P \tag{2.1}$$

with P in GeV/c [49].

2.4 Trigger Hodoscopes

Two planes of 1 cm thick plastic scintillator counters are used for fast triggering on charged particles. The two hodoscope banks, designated V and V', are 1.9 m square, and located 2 m in front of the cesium iodide calorimeter. The counter arrangement in Figure 2.4 shows that the banks have a split configuration, with half of the counters extending from the mid-line up, the other half from the mid-line down. Photomultiplier tubes at the top and bottom ends collect the light output from charged particle tracks. These signals are discriminated and can be used for single bucket triggering. Inefficiencies due to cracks are minimized by using different sizes and arrangements of counters in each plane, so that cracks in the two planes do not overlap. V and V' have 14 cm beam holes to reduce high hit rates and radiation damage of the calorimeter from beam interactions.

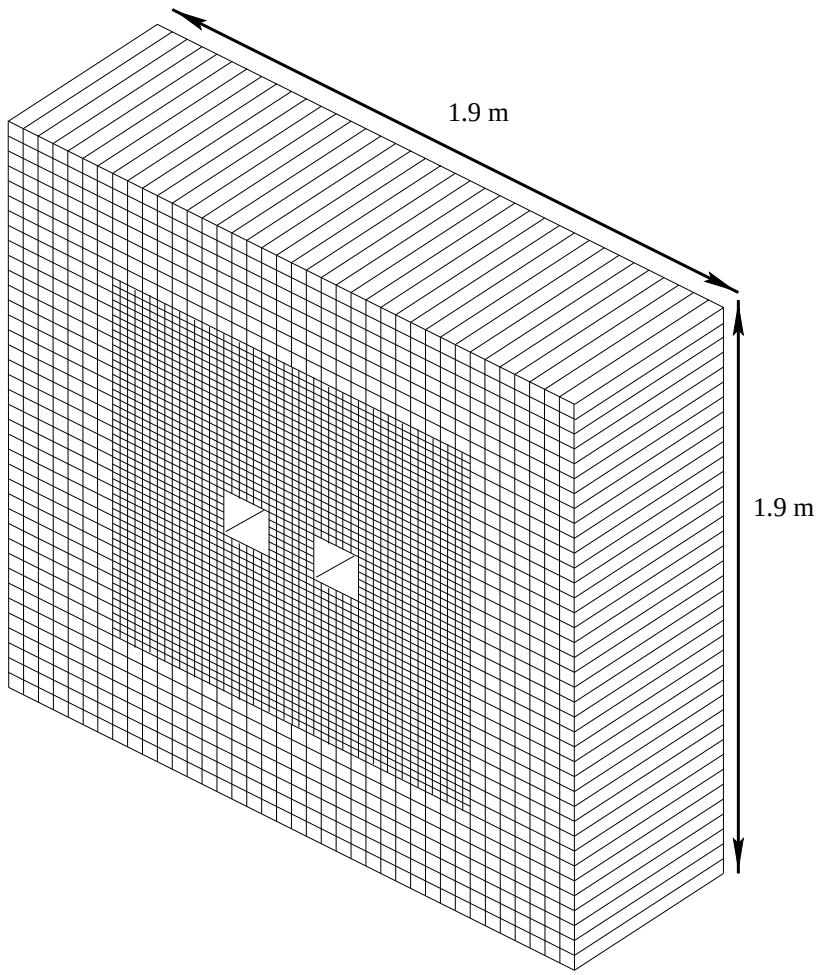

Figure 2.5: Cesium iodide calorimeter array of 3100 crystals.

2.5 Cesium Iodide Calorimeter

2.5.1 Crystal Array

Photon and electron energies are measured very precisely with the cesium iodide calorimeter (CsI). The CsI is a 1.9 m × 1.9 m array of 3100 pure cesium iodide crystals shown in Figure 2.5. The array consists of 2232 small crystals (2.5 cm square cross section) surrounded by 868 large crystals (5.0 cm square cross section). The small crystals afford better position resolution in the central region where distinguishing close energy clusters is important. There are two 15 cm square beam

holes in the CsI that allow the neutral beam to pass without incident. All of the crystals are 50 cm long, which is equivalent to 27 radiation lengths or 1.4 hadron interaction lengths. This means that electron and photon energies are completely contained in the CsI, while hadrons and minimum ionizing particles will only deposit a fraction of their energy in the calorimeter. Very effective electron and pion discrimination is made possible by this differing energy response. The CsI has photon energy resolution of

$$\frac{\sigma(E)}{E} = 0.45\% \oplus 2\%/\sqrt{E}, \tag{2.2}$$

where E is in GeV, and the position resolution is 1 mm in the small crystals and 1.8 mm in the large crystals [45].

To achieve such a good energy resolution, the crystals were processed individually to tune the light outputs for optimum performance. Each crystal was custom wrapped with black and/or aluminized 13 μm thick Mylar to make the light output uniform along the length of the crystal. The crystals are optically coupled to photomultiplier tubes (PMTs) with transparent silicone "cookies." Also included in the couplings are filters that block the slow component of light from cesium iodide.[1] Otherwise, pileup from multiple buckets would become a problem. Additionally, some crystals with very high light output have a black mask that blocks a fraction of the light, so that the output can be matched to the optimum operating range of the PMTs.

2.5.2 Readout

The PMTs are connected to special bases that digitize and buffer the PMT signals in addition to providing voltage to the tubes. This combination of a PMT with a digitizing base is called a digital photomultiplier tube (DPMT). The DPMT cards include a QIE, which is a custom charge integrating circuit. A QIE divides the PMT current from the time slice corresponding to one bucket, among eight capacitors, each receiving a different fraction of the total current (I/2, I/4, I/8,...,I/256). The

[1] Cesium iodide has a fast component of scintillation light with a 25 ns decay time and 305 nm peak wavelength, and a slow component with 1μs decay time and 480 nm peak wavelength.

capacitor voltages are then compared to reference voltages to determine which one falls in the correct digitizing range. The selected voltage is output to an 8-bit flash ADC (analog-to-digital converter) to form a digital "mantissa." A 3-bit "exponent" output from the QIE indicates the selected current range. Each QIE has four sets of identical capacitor circuits that continuously read out time slices in a round-robin fashion.

The mantissa and exponent output bits are stored in a FIFO buffer on the DPMT card. If an event passes the trigger, then the DPMT buffer is transferred to the "pipeline." The pipeline is a VME buffer and sparsification system that can be tailored to meet the various CsI output needs of each trigger.

2.6 Muon Identification System

The Muon Identification System is a catch-all title that encompasses the sequence of particle filters and scintillator counter planes downstream of the CsI. This system, shown in Figure 2.6, is designed to filter out interacting particles and identify penetrating ones. One of the filters, Muon Filter 2 (MF2), also serves as the neutral beam dump.

2.6.1 Pb Wall, Hadron Anti, and Muon Filter 1

The first three elements of the muon identification system are actually used to identify pions. The Pb Wall is a 10 cm thick wall of lead bricks that absorbs any remnants of electromagnetic showers that leak out of the CsI. It also induces hadronic showers for those hadrons (not including beam particles) that have not already showered in the CsI. Showers from the Pb Wall are detected by the Hadron Anti (HA), which is a bank of 28 non-overlapping scintillator counters. An analog sum is made from the PMT signals of the HA counters. This sum is used in many triggers to reject events with hadronic activity. Essentially, the HA provides additional $e^{\pm}$-$\pi^{\pm}$ discrimination. Consequently, it is not used in this analysis. The first Muon Filter (MF1) is a 1 m wall of steel. The purpose of MF1 is to prevent the neutral beam dump

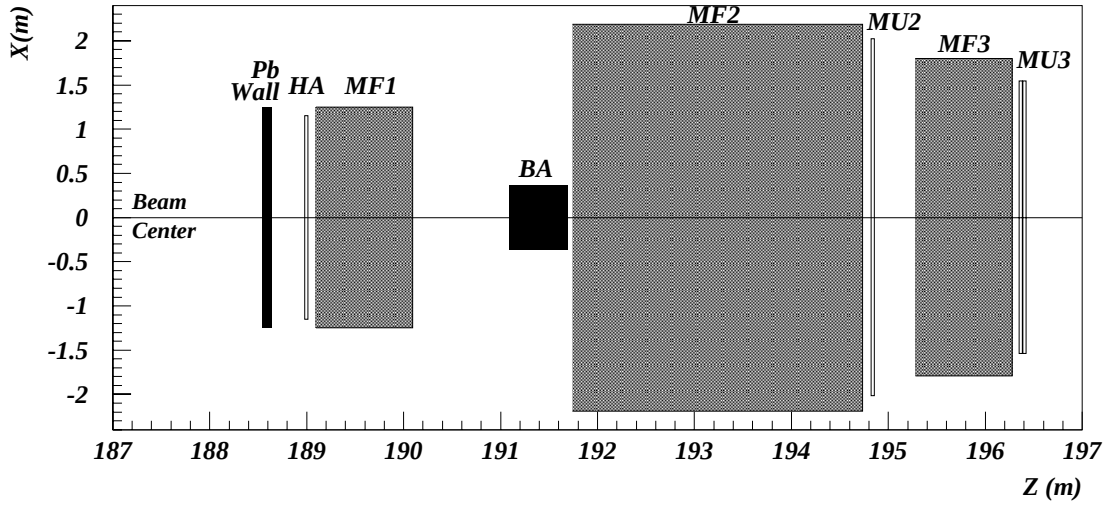

Figure 2.6: Muon identification system plan view. The three Muon Filters (MF1, MF2, MF3) are made of battleship steel.

backsplash from firing the HA. The Pb Wall, HA, and MF1 each have one hole that allows both beams to pass through.

2.6.2 Muon Filter 2 and MU2

The second Muon Filter, MF2, is 3 m deep, 3.4 m high, and 4.3 m wide - 44 cubic meters of battleship steel! MF2 is the neutral beam dump, and it stops most of the hadronic shower activity. MU2 is a 3.93 m $\times$ 2.99 m plane of 56 overlapping counters located immediately behind MF2. Figure 2.7 shows the arrangement of MU2 counters. The 15 cm wide counters are vertically oriented and overlap by 1 cm along the counter sides and along the ends near the vertical mid-line. MU2 must be hermetic because it is used as a muon veto in many triggers. The large size of the plane and the depth of MF2 were chosen to optimize muon rejection efficiency even for muons that scatter through large angles in the steel. MU2 is used as an acceptance detector for muon calibration triggers, and only in that mode is it relevant to this analysis.

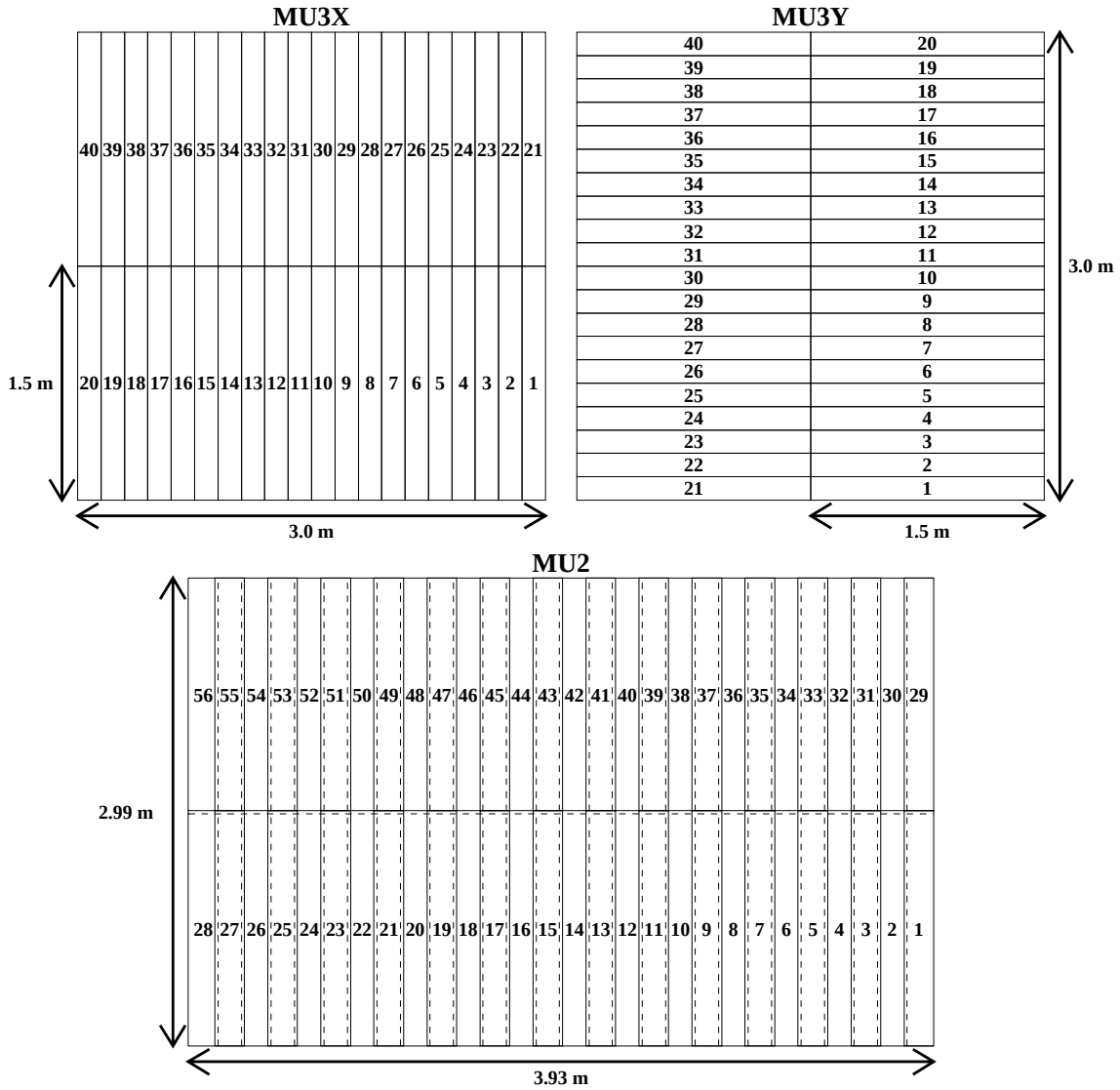

Figure 2.7: Counter maps for MU2, MU3X, and MU3Y. The dashed lines in the MU2 map indicate the counter overlap regions.

2.6.3 Muon Filter 3 and MU3

Muon Filter 3 (MF3) is an additional 1 m of steel downstream of MU2. It is designed to stop as much non-muon activity as possible that leaks through MF2, while still retaining a reasonable momentum threshold for muons to pass through. The Pb Wall and the three steel Muon Filters combine for a total of 31 interaction lengths. This amount of filtering yields a 7 GeV/c momentum threshold for muons and a 0.5% probability that shower products from a 20 GeV/c pion will leak through MF3.

MU3 consists of two planes of 40 scintillator counters each (Figure 2.7). These are acceptance detectors for triggers collecting rare decays with muons in the final state. One plane is oriented horizontally (MU3Y), and one is oriented vertically (MU3X). Used together, the two 3 m $\times$ 3 m planes give 15 cm hit resolution in the X and Y views. The counters in the MU3 banks do not overlap, so there are inefficiencies from cracks. However, requiring hits in both planes, the MU3 efficiency for muons with momentum $\geq$ 10 GeV/c is $\sim$ 99%, including effects from filtering, cracks, and the intrinsic counter and trigger inefficiencies.

2.7 Photon Vetoes

A series of detectors called photon vetoes are used to detect particles escaping from the fiducial region. These include five Ring Counters (RCs) spaced along the vacuum decay region, three Spectrometer Antis (SAs) and the Cesium Iodide Anti (CIA) preceding the final three drift chambers and the CsI respectively, and the Collar Anti (CA) shadowing the CsI beam hole edges. They are referred to as "photon vetoes" because they are designed to detect photons with energy as low as 100 MeV. All have modular designs, with each channel read out through PMTs into a LeCroy 4300 ADC. The discriminated signals are available at trigger level to veto events with particles leaving the fiducial region.

The RCs, SAs, and CIA are made of 16 radiation lengths of lead-scintillator sandwich layers. They each have square inner apertures, which define the fiducial region. The RCs have circular outer boundaries, and are located inside the vacuum decay region to prevent particles from leaving the decay volume undetected. The SAs

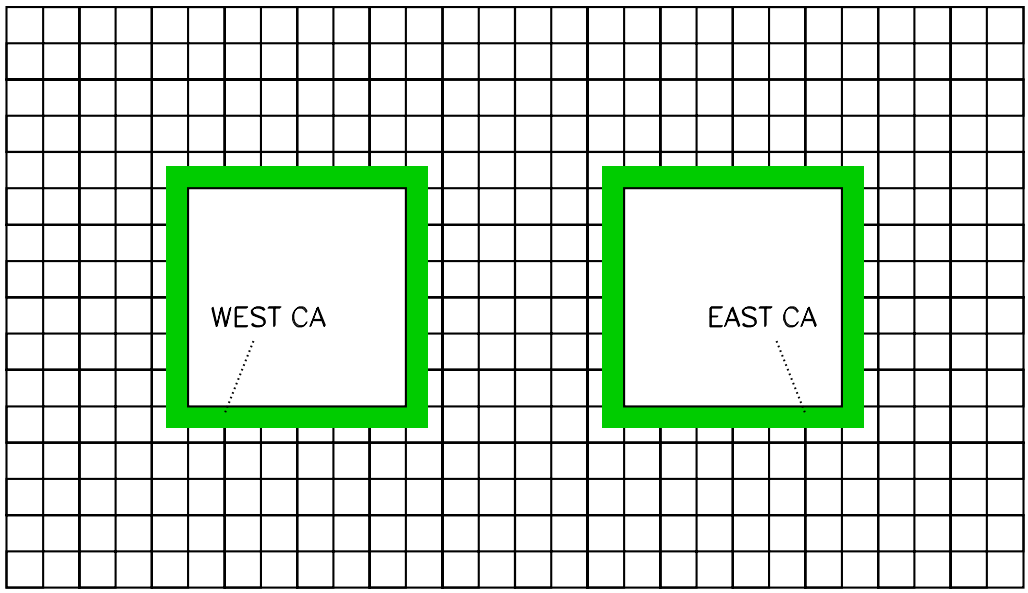

Figure 2.8: Collar Anti geometry. The position of the Collar Anti (black) relative to the beam holes (white) and inner region of the CsI (grid) is shown.

and CIA have square outer boundaries. Their inner apertures are somewhat smaller than the outer boundaries of the drift chambers and CsI that they immediately precede. SA and CIA modules are beveled to avoid crack inefficiencies. In addition to catching photons, they are particularly helpful in detecting charged particles that are kicked out of the fiducial region by the analysis magnet.

The CA defines the inner aperture of the CsI around the beam holes. It is composed of three layers of tungsten-scintillator sandwich, comprising 2.9 radiation lengths. The CA overlaps the beam hole crystals by 1.5 cm. The purpose of the CA is not so much to detect escaping particles as it is to veto events with particles that would hit the CsI too close to the edge, thereby ensuring that particle energies can be well-measured by the CsI.[2]

Another photon veto does exist: the Back Anti (BA). Located just in front of the MF2 beam dump, the BA is designed to veto events in which a photon escaped down the beam hole. It has 30 layers of lead-scintillator sandwiches divided into

[2]Another very important function of all photon vetoes dicussed thus far is to define clean apertures for the $K_L \to \pi^0\pi^0$ decay mode in the E832 $\epsilon^{'}/\epsilon$ analysis.

three sections which can distinguish between electromagnetic and hadronic (beam dump) interactions. The BA is not used in this analysis.

2.8 Transition Radiation Detectors

A system of eight transition radiation detectors (TRDs) was employed in E799II to greatly enhance $e^{\pm}$-$\pi^{\pm}$ discrimination [50]. A layer of polypropylene felt is mounted on the front of each TRD. Transition radiation X-rays are produced when highly relativistic electrons pass through the felt. Other charged particles will leave only ionization energy. The X-rays then continue on to be detected by a multiwire proportional chamber filled with a recirculating gas mixture of 80/20 Xe/CO_2. This mixture has a very short X-ray absorption length, i.e. very high X-ray detection efficiency. The different energy spectra produced by electrons and pions results in pion rejection better than 200:1 at 90% electron efficiency. However, the TRDs are not used in this analysis.

2.9 Accidental Counters

One very simple system is used to get a picture of the ambient, background detector activity that is unrelated to kaon decays. It is important to understand this activity because it can obscure kaon decay signatures in the detector. It is called "accidental" activity because it can be accidentally coincident in time with a kaon decay.

The 90° target monitor is a telescope consisting of three scintillator counters. It images the target through a hole in the target shielding aligned 90° relative to the beam direction. A coincidence of all three counters satisfies an accidental trigger, which records the detector activity occurring in that bucket. The accidental activity should be proportional to the beam intensity, but uncorrelated with actual kaon decays in the detector.

CHAPTER 3
EVENT SELECTION

$K_L \to \mu^+\mu^-\gamma$ signal events and $K_L \to \pi^+\pi^-\pi^0$ normalization events were collected and analyzed separately. There were two parts to the event selections: an online trigger and offline data reduction. The trigger consisted of two hardware stages (Level 1 and Level 2) and a software stage (Level 3). The offline data reduction included a split of the data into separate sets defined by general physics topics, and a preliminary "crunch" to remove large classes of uninteresting events.

3.1 The Trigger

The rate of kaon decays in KTeV was typically $\sim$ 1 MHz. Because the data acquisition system (DAQ) only had the capacity to process and write out events at 1 kHz, the trigger system had to be capable of reducing the event rate by a factor of 1000.[1] Additionally, trigger decisions had to be made as quickly as possible to avoid deadtime while the trigger was processing. The RF structure of the beam was 53 MHz, so that during the spill, protons buckets arrived on target every 19 ns. Ideally, most decisions would be made on the time scale of the primary proton beam structure.

3.1.1 Level 1

The first level of trigger decisions was made using all of the fastest signals available from the detector. These signals mostly came from phototubes on the photon vetoes, trigger hodoscopes, and muon counters, and total energy sum from the CsI. All of

[1] The complete trigger system is described thoroughly in Ref. [51]

these signals were processed within one bucket. A Level 1 source which took slightly longer to process was the Drift Chamber OR (DCOR). This system divided each of the planes in the first two drift chambers into 16 "paddle" regions, and applied coincidence logic to make a quick, rough count of drift chamber hits. The DCOR took 4-5 buckets to process.

All of the Level 1 sources were synchronized to the RF structure within 2 ns, and LeCroy 4508 programmable logic units performed coincidence logic on them to make trigger decisions in each bucket. This resulted in no deadtime from this stage of the trigger, which had an output rate of 100 kHz.

3.1.2 Level 2

The Level 2 hardware trigger consisted of six pattern-recognition systems of custom electronics. The two Level 2 systems used to select events for this analysis are Hit Counting and the Hardware Cluster Counter (HCC).[2] The Level 2 systems took 2.5 μs to make positive trigger decisions and another 15 μs to process accepted events. Level 2 reduced the event rate to 20 kHz; however, the Level 1 trigger was dead during Level 2 processing.

The Hit Counters were electronics that counted hits in the X and Y views of all four drift chambers. The counting logic was based on patterns of adjacent hit wires in the chambers. The HCC counted the number of clusters in the CsI based on patterns of energy deposits in 2×2 groups of crystals [52]. Triggers could make selections of 0 to 7, or more than 8 clusters based on a three-bit plus overflow HCC output. HCC decisions took longer than any other trigger processor and set the overall Level 2 decision time of 2.5 μs.

3.1.3 Trigger Definitions

There were 16 trigger definitions available for combining the Level 1 and Level 2 trigger sources to make event selections. Nine of them were for collecting the

[2]The other systems were a Y-track finder and in-time hit pair finder for the drift chambers, a TRD electron tagger, and a hyperon stiff track trigger.

various signal decay modes to be studied. The other seven were for obtaining data for calibration, systematic error studies, and some normalizations.

The hardware trigger used to collect the signal events was the Dimuon Trigger (2MULD), which was defined as

$$2MULD = GATE \bullet 2V \bullet DC12 \bullet 2MU3 \bullet PHVBAR1 \bullet 2HCY_LOOSE \bullet HCC_GE1,$$

the normalization was collected with a prescale of 500 in the Two Track Trigger (2TRK) defined as

$$2TRK = GATE \bullet 2V \bullet DC12 \bullet 2HCY_LOOSE,$$

a calibration trigger that was very important for this analysis was the Single Muon Trigger (MUON) used to collect muons during special muon only runs:

$$MUON = GATE \bullet 1V \bullet MU2,$$

and accidental activity data was collected with the ACC90 trigger

$$ACC90 = GATE \bullet ACC_90DEG,$$

where

<u>Level 1</u>

$GATE$ = on spill

$2V$ = at least 2 hits in one trigger plane and at least 1 hit in the other

$DC12$ = at least 1 DCOR hit in each view of DC1 and DC2

$2MU3$ = 2 or more hits in the X and Y views of MU3

$PHVBAR1$ = veto on all RCs except RC8, all SAs, and CIA

$1V$ = at least 1 hit in either trigger plane

$MU2$ = 1 or more hits in MU2

ACC_90DEG = three counter coincidence in the 90° target monitor.

Table 3.1: Level 3 cuts.

Description	2MULD Cut	2TRK Cut
Number of X tracks	≥ 2	≥ 2
Number of Y tracks	≥ 2	≥ 1
Number of vertex candidates	≥ 1	≥ 1
Number of track-cluster matches	2	2
Energy of track clusters	≤ 5 GeV	
Total number of clusters	≥ 3	

Level 2

$2HCY_LOOSE$ = 2 Y-view hits in each DC, but allow 1 missing hit in either DC1 or DC2

HCC_GE1 = at least 1 HCC cluster

3.1.4 Level 3

Events that were accepted by Level 2 were sent to 4 SGI Challenge machines which performed online reconstruction of events with a simplified version of the offline analysis code. In the Level 3 code, events were classified based on various reconstructed quantities, and tagged as particular decay mode candidates according to the event topologies. Level 2 output was buffered in a pipeline memory system, so that data could accumulate during the 20 seconds on-spill, and Level 3 processing could continue during the following 40 second off-spill period. Level 3 only produced deadtime if it could not finish processing the entire spill's data before the next spill began. The overall trigger deadtime averaged about 33%, and the output was written to 10 DLT tape drives at a rate of 1 kHz.

A number of basic reconstruction requirements were made at Level 3 for the 2MULD and 2TRK triggers. The Level 3 cuts for both triggers are given in Table 3.1. In the 2MULD trigger, the software searched for a reconstructed 2-track vertex with an additional cluster not associated with any track. Furthermore, the energy of clusters matched to tracks had to be less than 5 GeV, or the tracks were identified

as electrons or pions and the event was rejected. The 2TRK trigger simply looked for any reconstructed 2-track vertex.

3.2 Data Set

3.2.1 The E799II Run

The term "run" is used in many contexts for this experiment. In the most general sense, the *KTeV Run* was the period of time in which KTeV (both E832 and E799II) took hadron data between August 1996 and September 1997. The E799II data collection occurred in two segments: the *Winter Run* from January 24 through March 23, 1997, and the *Summer Run* from July 24 through September 3, 1997. The Winter and Summer Runs were preceded by Fall and Spring E832 Runs, respectively.

In the most specific case, a run refers to a period of uninterrupted data collection. In the absence of any problems, a single run lasted up to about eight hours, which is approximately the time it took for output tapes to fill, given nominal beam intensity. About half of all runs were stopped short of eight hours because of beam, detector, or data acquisition problems. The Winter Run spanned runs 8088 through 8910, and the Summer Run included runs 10463 through 10970. Prior to run 8245, there was a swapped HCC cable that corrupted some CsI data, and those runs are not included in this analysis. Many of the run numbers correspond to muon, other calibration, debugging, or aborted runs. There were 154 data runs used from the Winter Run, and 112 runs from the Summer Run.

3.2.2 Split

The 850 raw data tapes from the Winter and Summer Runs were split by trigger type and Level 3 tags into smaller sets of output tapes. There were 19 different output streams for the E799II split. About a dozen of the streams corresponded to general categories of physics interest, such as neutral events, or 4-track events. The rest of the streams produced output tape sets for calibration or accidental data [51].

All of the signal and normalization data used in this analysis was split to the

Table 3.2: Crunch cut efficiencies. Signal and normalization mode cuts are identical unless otherwise indicated.

Signal Mode Crunch Cut	$K_L \to \mu^+\mu^-\gamma$ Data	$K_L \to \mu^+\mu^-\gamma$ MC	$K_L \to \pi^+\pi^-\pi^0$ Data	$K_L \to \pi^+\pi^-\pi^0$ MC	Normalization Mode Crunch Cut
At least two X and Y tracks	0.998	1.000	0.998	0.998	At least two X and Y tracks
At least one vertex and one track-cluster match	0.985	0.991	0.864	0.976	At least one vertex and one track-cluster match
Vertex χ^2	0.903	0.993	0.966	0.973	Vertex χ^2
Track energy $\leq$ 2 GeV	0.375	0.973	0.612	0.969	Track $E/p \leq 0.95$
At least one photon cluster	0.708	0.996	0.386	0.977	At least one photon cluster
$P_t^2 \leq 250(MeV/c)^2$	0.210	0.978			

2MULD stream. In addition to accepting all 2MULD triggers, this split stream included 2TRK trigger data prescaled by an additional factor of 14. Therefore the total prescale for the 2TRK normalization sample is 7000, including the trigger prescale of 500. 2MULD was the largest of the physics split streams, consisting of 66 Winter and 64 Summer output tapes.

3.2.3 Crunch

After the split, additional or tighter cuts relative to the Level 3 selection criteria were applied to reduce the size of the data set. These cuts were made possible by the availability of improved, offline calibrations. They are detailed in Table 3.2 for the signal and normalization samples.

Some of the crunch cuts are very nearly duplicates of the ones used in Level 3. Both stages of cuts require events to have at least 2 tracks, 1 vertex candidate, track-cluster matches, and 1 cluster not matched to a track. For Level 3, only one Y-track was required for the 2TRK trigger, but 2 are necessary to survive the crunch. An additional cut in the crunch is that the fit for the vertex reconstruction

have a reasonably small χ^2. The cut on the energy of track-matched clusters is tightened from 5 GeV to 2 GeV in the signal mode data to remove $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \pi^\pm\mu^\mp\nu$ events with pion showers in the CsI. For the normalization, the ratio of track cluster energy to track momentum is required to be less than 0.95. This rejects most of the $K_L \to \pi^\pm e^\mp\nu$ events that comprise a large amount of the 2TRK data because electrons deposit all of their energy in the CsI. Requiring at least 1 photon cluster removes more $K_L \to \pi^\pm\mu^\mp\nu$ from both the 2MULD and 2TRK samples. Finally, there is a cut on the total momentum transverse to the kaon flight direction for 2MULD data. This removes a large amount of all potential backgrounds which typically have extra or missing particles involved in the reconstruction.

The crunch output consists of 6 tapes for the Winter and 5 tapes for the Summer. This is a reasonably sized data set on which to perform the full reconstruction and analysis, which are described in Chapters 4, 6, and 7.

CHAPTER 4
EVENT RECONSTRUCTION

Kaon decays are reconstructed almost entirely from drift chamber and CsI data. Charged particle trajectories and momenta are measured from tracks found in the drift chambers. CsI clusters are used for photon energy and position measurements, and for particle identification.

There are three basic stages in event reconstruction. Track candidates in the X and Y views are reconstructed from hit positions in the drift chambers. Next, CsI energy clusters are reconstructed from the signals recorded in individual crystals. Finally, information from both tracking and clustering is used to locate the decay vertex position. In all three stages, various calibrations are required to ensure accurate measurements. Upon completion of these steps, the full kinematics of the kaon decay are available for analysis.

4.1 Tracking

4.1.1 Hits, Pairs, and SODs

A hit in a drift chamber is defined as a TDC signal above threshold for a given wire. To be used for tracking, a hit must occur within an "in-time" TDC window of 235 ns. If there are multiple in-time hits, only the earliest one is used. A feature of the drift chamber geometry is that sense wire planes are arranged in complementary pairs, offset transversely by half the distance between wires in a plane. This arrangement creates cells, which are the regions between adjacent wires in complementary planes. Normally when a charged particle passes through a cell, there are TDC hits on both sense wires surrounding the cell. These hit "pairs" are the building blocks used to construct tracks.

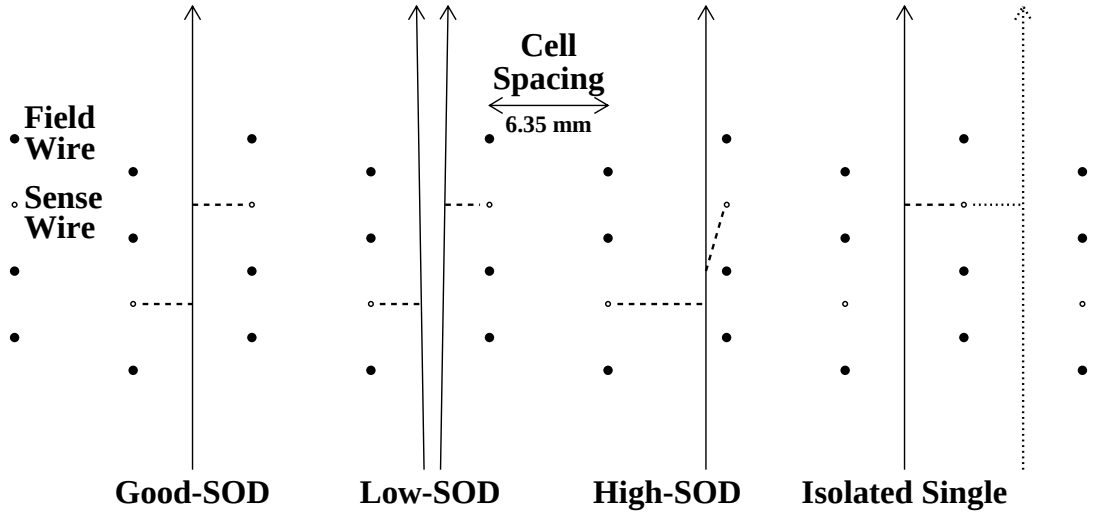

Figure 4.1: Illustration of SOD classes. Solid arrows signify charged particle trajectories; dashed lines represent measured drift distances; dotted lines indicate the ambiguity associated with isolated hits.

Each TDC signal represents the drift time for the ionization electrons to reach the wire. The drift times can be converted into drift distances using XT calibrations described in Section 4.1.2. If the drift distances for each hit in a pair are added, the sum of distances (SOD) should nominally be equal to the cell spacing of 6.35 mm. An illustration of SOD measurements is shown in Figure 4.1. A good-SOD is defined to be within 1 mm of the cell spacing to account for the plane-by-plane drift distance resolutions. However, there are several pathologies which can result in a bad SOD or an isolated single hit.

A low-SOD can occur when two tracks occupy the same cell. The close tracks could both be from products of the kaon decay. Delta rays and accidental activity can also produce low-SODs. In either case, the earliest hit on each wire comes from the track which passed closest, resulting in a SOD which is smaller than the cell size.

If the earliest drift electrons do not produce a signal large enough to pass the threshold of the chamber electronics, a high-SOD can be measured. This turned out to be a significant problem in KTeV. Insensitivities to the early electrons were

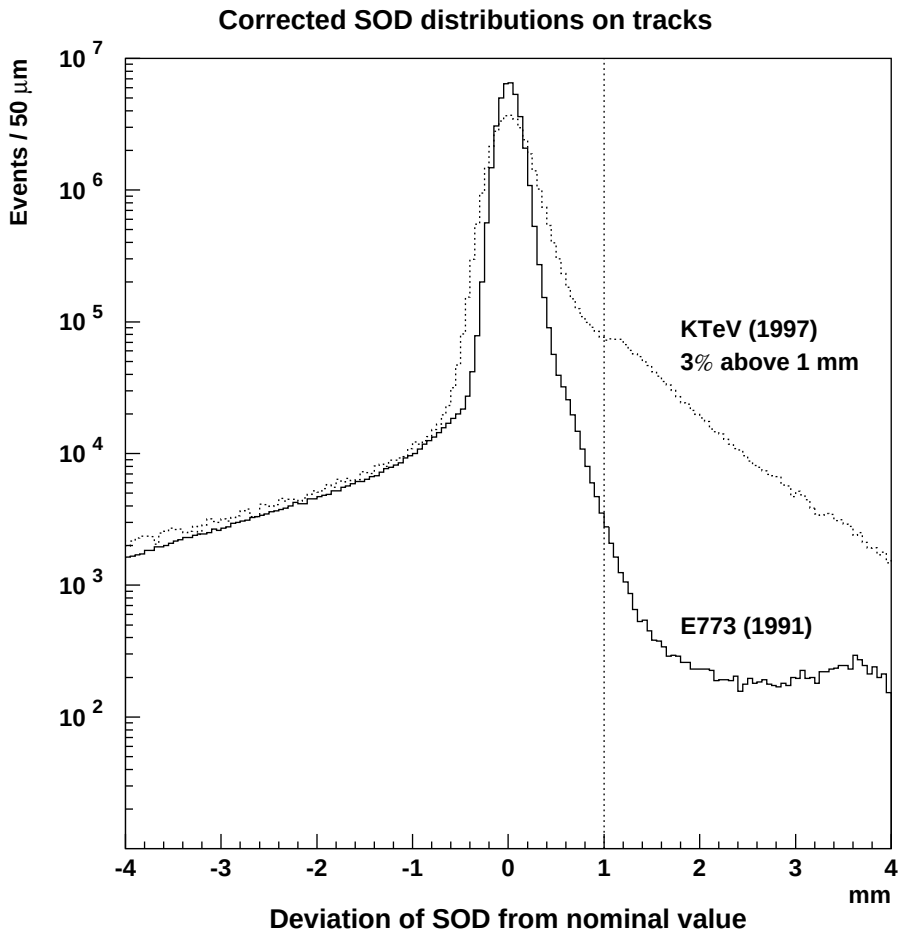

Figure 4.2: SOD distributions for KTeV and E773. The plot shows the deviation from the cell spacing of 6.35 mm.

the result of radiation damage and a combination of low gain and noise in the drift chamber electronics. Approximately 3% of all hit pairs in KTeV had high-SODs. This was a much higher level than was seen in E799I and E773, as the SOD distributions in Figure 4.2 illustrate. In addition, KTeV observed a significant component of very high-SODs (more than 2 mm larger than the cell size) which was absent in E773.

A final possibility is that a wire in only one of a pair of planes is hit. The missing hit could be the result of a defect on the wire such as contamination or kinking. For

this case of isolated single hits, there is an ambiguity as to which side of the wire the track lies.

4.1.2 Calibrations

Time-to-distance Conversions

The first of two major calibration issues related to the drift chambers is the conversion of TDC counts to drift distances. Before a time-to-distance, or $X(t)$, relation can be derived, the relative timing of individual wires must be determined. There can be differences in timing due to variations in cable lengths, in individual electronics channels, and in the global trigger signal delivered to the TDCs, for example. These differences are determined by measuring the TDC count corresponding to zero drift distance (T0) for each wire. The T0s are defined as the midpoints of the raw TDC thresholds, and are measured with $K_L \to \pi^{\pm} e^{\mp} \nu$ (K_{e3}) data. A sample TDC distribution is shown in Figure 4.3.

Once the T0 offsets are measured, maps of TDC counts to drift distances for each drift chamber plane are constructed. This XT calibration, based on K_{e3} data, assumes that the illumination of tracks is on average uniform across the cell. The earliest hits (corresponding to the largest TDC counts) are taken to occur adjacent to the wire, and the latest hits (smallest TDC counts) at the edge of the cell, 6.35 mm from the wire. The drift distance associated with the TDC count t is then calculated from the distribution of TDC hits using the expression

$$X(t) = L_{cell} \times \frac{\sum_{t_0}^{t} N(t^{'})}{\sum_{t_0}^{t_m}} N(t^{'}), \tag{4.1}$$

where L_{cell} is the cell spacing, t_0 and t_m are the earliest and latest TDC counts, and $N(t^{'})$ is the number of hits at a TDC count of $t^{'}$.

The measurement of hit position resolutions is necessary to evaluate vertex candidates. The position resolutions are measured for each complementary plane pair by performing a Gaussian fit to the central portion of the SOD distributions. A SOD is the sum of two separate drift distance measurements, so the plane pair resolution

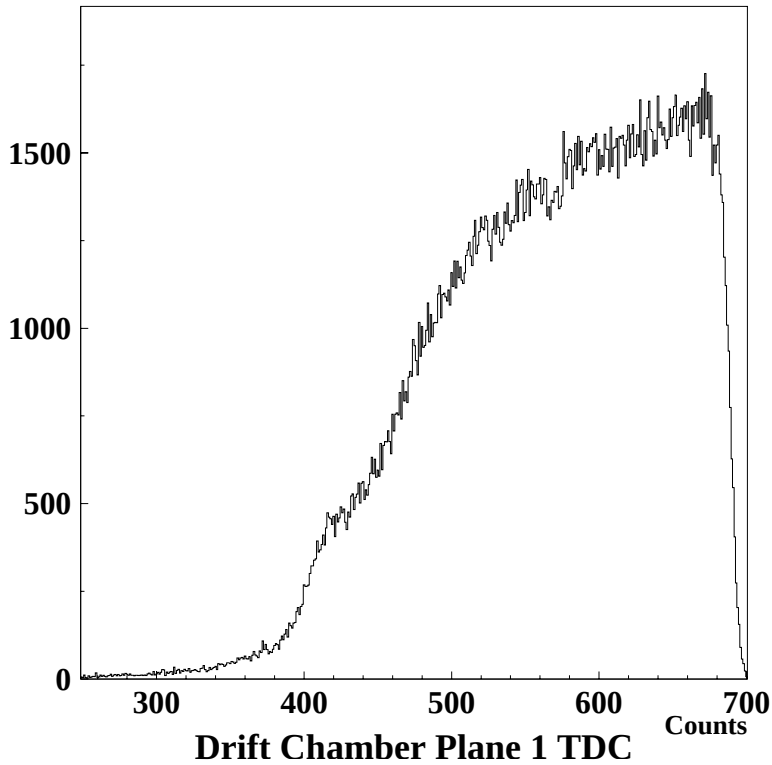

Figure 4.3: TDC distribution for the first X plane of DC1. The TDC runs in common stop mode, so the highest TDC counts correspond to the earliest hits on a wire.

Table 4.1: Drift chamber resolutions for Run 8397.

Chamber	X-view Resolution (μm)	Y-view Resolution (μm)
DC1	92.4	87.0
DC2	94.7	98.8
DC3	92.9	94.7
DC4	91.2	91.8

is related to the width of the SOD fit by

$$\sigma_{PP} = \sigma_{SOD}/\sqrt{2}. \tag{4.2}$$

The chamber resolutions for a typical run are given in Table 4.1.

Chamber Alignment

Internal alignment of the drift chambers relative to each other and external alignment of the drift chamber system with respect to the rest of the detector are crucial for good track reconstruction. The internal alignment uses muon tracks from muon runs with the analysis magnet field set to zero. These straight tracks are used to locate any two of the chambers relative to the other two which are held fixed. Then 2TRK K_{e3} data is used to measure any residual corkscrew rotation between the two chambers held fixed in the straight-through alignment. The drift chamber system is externally aligned to the survey positions of the target and CsI using 2TRK data.

Straight-through track alignment begins by choosing any two chambers to hold fixed, DC1 and DC4 for example. Muon tracks with hits in all 16 planes are selected, and tracks are constructed in each view using only the hits in the outer chambers. The location of the hits in DC2 and DC3 are compared to the hit positions expected from the tracks made using hits in DC1 and DC4. Hit position differences in one view are then plotted against the positions in the other view for each of the inner chambers. The offset and rotation of a plane pair are given by the intercept and slope of a linear fit to the data, which are shown in Figure 4.4 for DC2.

After the straight-through alignment, there can still be a residual rotation between DC1 and DC4. This rotation would manifest itself as a corkscrew rotation of the entire drift chamber system. The corkscrew rotation is measured with K_{e3} events. Charged tracks in these events are planar, since they originate from a single vertex. The vector connecting the two hits in DC1 ($\vec{\mathbf{r}}_\mathbf{1}$) is related to the analogous vector in DC2 ($\vec{\mathbf{r}}_\mathbf{2}$) through the equation

$$\vec{\mathbf{r}}_\mathbf{1} \times \vec{\mathbf{r}}_\mathbf{2} = |\vec{\mathbf{r}}_\mathbf{1}||\vec{\mathbf{r}}_\mathbf{2}| \sin\phi, \tag{4.3}$$

where ϕ is the corkscrew rotation between DC1 and DC2. The corkscrew rotation is illustrated in Figure 4.5. For small angles, the rotation can be obtained from the slope of a linear fit to a plot of $\vec{r}_1 \times \vec{r}_2$ vs. $|\vec{r}_1||\vec{r}_2|$ (Figure 4.6). This corkscrew rotation is uniform along the Z direction, thus the rotations for DC3 and DC4 are

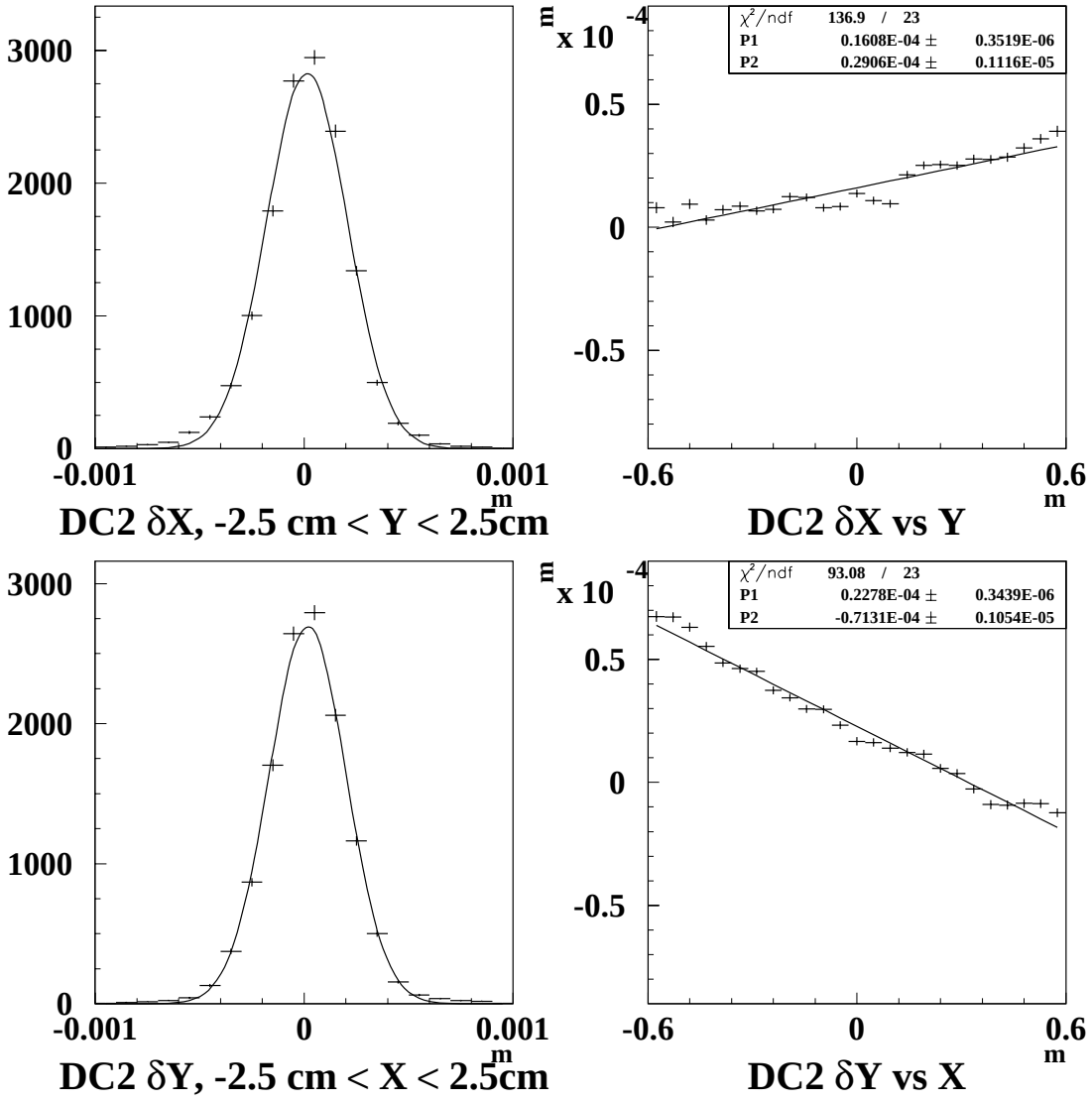

Figure 4.4: Straight-through muon alignment of DC2. In the top (bottom) graphs, the difference between the actual and expected hit position in the $X(Y)$ view are shown. The left-hand plots show the data in the central 5 cm of each view. In the right-hand graphs, the means of Gaussian fits to the data taken in 5 cm slices are plotted against the $Y(X)$ position in the chamber. The offset and rotation of each view are taken from the intercept and slope (P1 and P2, respectively) of the linear fits.

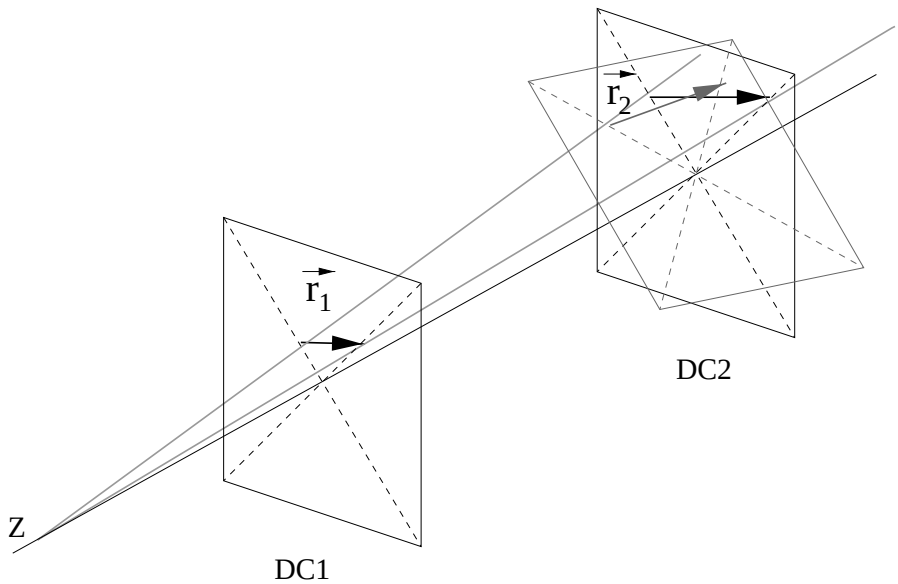

Figure 4.5: Diagram of the corkscrew rotation between DC1 and DC2. $\vec{r}_1$ and $\vec{r}_2$ are the vectors connecting the hit positions of two planar tracks originating from a common vertex in DC1 and DC2 respectively. A residual rotation will manifest as the two vectors appearing to not be parallel.

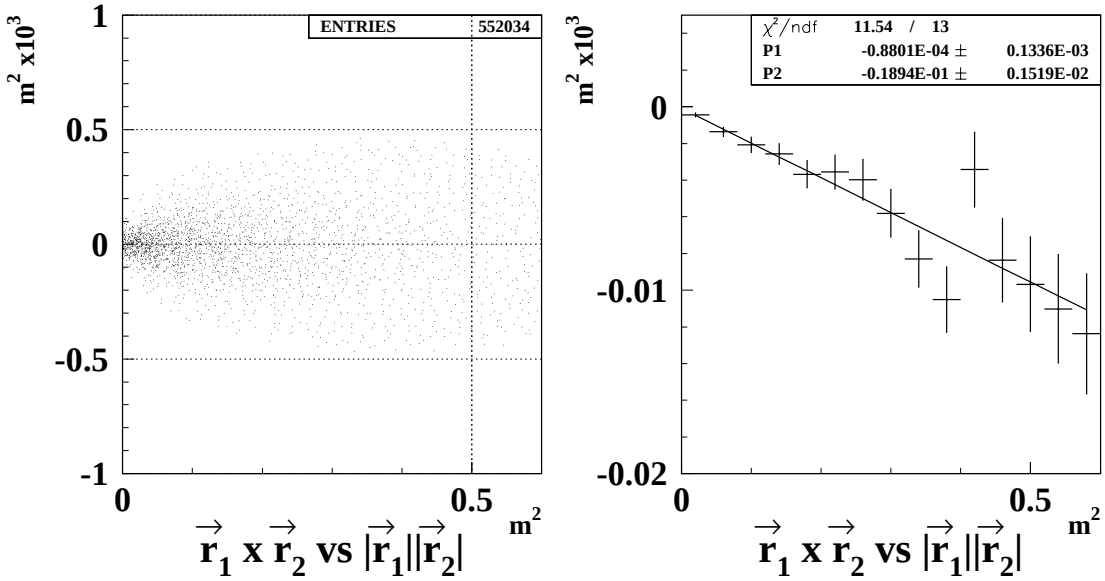

Figure 4.6: Measurement of the corkscrew rotation. The points in the right-hand graph are the means of Gaussian fits to the data taken in 0.04 m^2 slices along the abscissa. The corkscrew rotation between DC1 and DC2 is taken from the slope (P2) of the linear fit.

Table 4.2: Target and CsI survey positions.

	X (μm)	Y (μm)	Z (m)
Target	0	-305	0.000
CsI	-480	-3000	186.013

simply the rotation for DC2 scaled by the longitudinal distances of the chambers from DC1.

The final step in chamber alignment is locating the chamber system in the external reference frame defined by the target and CsI survey positions. The target location is found by reconstructing kaon decays and then projecting the total kaon momentum vector back to the target position at $Z = 0.0$ m. $K_L \rightarrow \pi^+\pi^-$ decays are used so that the kaon decay can be reconstructed completely from two tracks. The position of the kaon momentum in the $X - Y$ plane at the target Z position is shown in Figure 4.7. The means of the projections onto the X and Y axes yield the reconstructed position of the target. The position of the CsI relative to the chambers is determined in a manner similar to the straight-through muon alignment. Electron tracks from K_{e3} decays are projected to the CsI. The difference between the track position at the CsI and the center of the electron energy cluster is plotted as a function of position in the calorimeter. The intercepts and slopes of these distributions give the offsets and rotations of the chambers with respect to the CsI. The reconstructed target and CsI positions are then used to correct the positions of the chamber. This procedure is iterated until the target and CsI reconstruct to the survey positions presented in Table 4.2.

4.1.3 X and Y Track Candidates

Once all the locations of drift chamber hits have been determined from drift distances and wire positions, a list of hit "pairs" can be made. The pairs are classified according to the associated SOD, and there are limits on how many bad SOD pairs and isolated hits can be used to construct a track. Good-SOD pairs are entered into

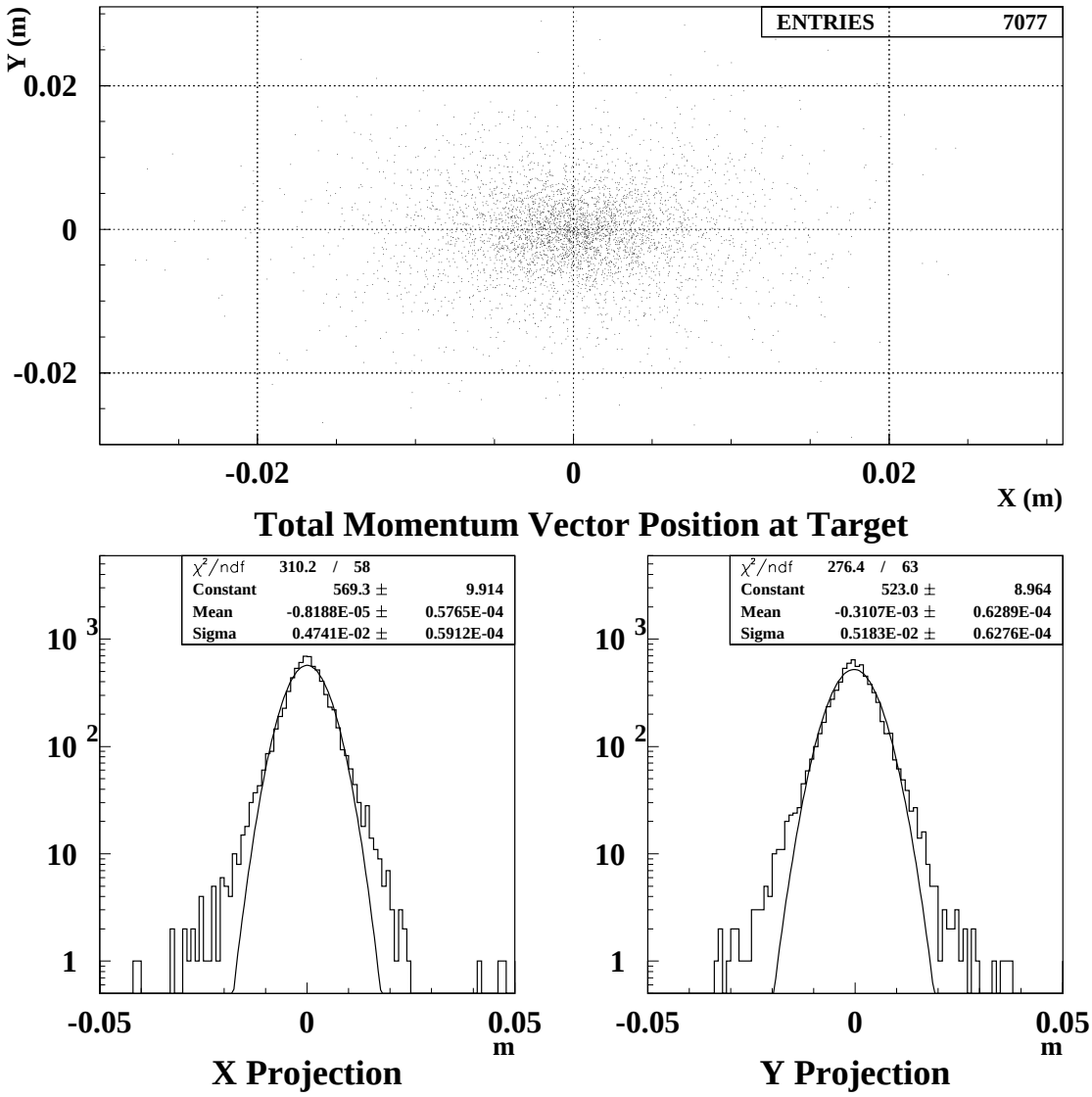

Figure 4.7: Reconstructed target position. The data represent the positions of the reconstructed kaon momenta of $K_L \to \pi^+\pi^-$ decays, projected back to the target at $Z = 0.0$ m

the list of pairs once at the average hit position. Low-SOD pairs are treated as two individual hits and one entry is made into the pair list at each hit position. Separate entries are also made for each hit in a high-SOD pair, with the restriction that only one of the hits can be used to make tracks. Because of the ambiguity associated with isolated hits, they are entered twice into the pair list, once on each side of the wire.

There is no magnet kick in the vertical direction, so Y tracks are characterized by hits that lie approximately along a straight line through all four chambers. Hits are identified that fall within a 1 cm wide "road" traversing the Y views of the chambers. A Y track is allowed to contain up to two bad SOD pairs or one bad SOD pair and one isolated hit. For all possible Y tracks, a least-square fit to a straight line is performed, and those that do not pass a maximum χ^2 cut are rejected. At least two viable Y track candidates that do not share hits must exist, or the event is rejected.

In the X view, upstream and downstream track segments are found separately. Track segments in the upstream chambers are allowed to have up to two bad SOD pairs or one isolated hit. Downstream segments may have one bad SOD pair or isolated hit. Upstream and downstream segments are matched by calculating X_{offmag}, the distance between the segments at the bend plane of the magnet. For two segments to be considered as an X track candidate, X_{offmag} must be less than 6 mm. An X track candidate is allowed to have a total of two bad SOD pairs or one bad SOD and one isolated hit. The event is rejected if there are not at least two X track candidates that do not share hits.

4.2 Clustering

4.2.1 Crystal Energies

Digitized DPMT signals are read out for the time slice containing a trigger, as well as the three subsequent time slices. To calculate the amount of energy deposited in a crystal from the DPMT signal, two calibrations are required. First, the DPMT response to varying levels of light in the crystal must be linearized by calibrating with a laser flash system. Then the energy scale relating electron energy to DPMT output

signal is measured with K_{e3} electrons. The E/p (CsI energy/track momentum) response to electrons is measured for each crystal. The gains are then calculated by finding the factor necessary to scale the measured E/p to a value of 1.0.

4.2.2 Hardware Clusters

The first step in hardware clustering is to search for local maxima in energy among all crystals which have been flagged by the HCC. The HCC flags crystals with at least 50 MeV of energy. Any such crystal that is a local maximum is considered a cluster seed block. Clusters are formed by summing the energies from a fixed number of crystals surrounding the seed block. For small crystals, a 7×7 array of crystals centered on the seed block is used. A 3×3 array is used for large crystals. If the cluster is at a boundary between large and small crystals, the large crystal array size is used. Four small crystals represent the equivalent of one large crystal in the portion of the array lying in the small crystal region.

4.2.3 Cluster Energy Corrections

There are many corrections made to the energy measurements in the CsI. Some of the corrections are made to the energies of individual crystals. The *Overlap Correction* divides the energy in a crystal between all clusters which share that crystal. The division is a function of the energy of the cluster and the position of the crystal within the cluster. The *Neighbor Correction* adjusts the energy in one cluster's crystal that is adjacent to, but not part of, another cluster. The adjustment is calculated using a transverse cluster profile obtained from GEANT studies. This correction is important for low energy clusters that neighbor high energy clusters, such as neighboring muon and photon clusters in a $K_L \to \mu^+\mu^-\gamma$ event. A *Missing Block Correction* is made to estimate the unmeasured energy for clusters near the beam holes or outer edge of the CsI. The transverse profile is used to calculate the energy in "missing blocks." Clusters near beam holes also use a *Sneaky Energy Correction*. This accounts for energy that traverses the beam hole from clusters on the other side. *Threshold Corrections* estimate the amount of energy deposited in

crystals that were below the readout threshold. Like the Overlap Correction, this one is based on the cluster energy and the position of the crystal within the cluster. Figure 4.8 shows a $K_L \to \mu^+\mu^-\gamma$ event requiring all five corrections to individual crystal energies.

Two corrections are made to the total cluster energy. The *Intra-Block Correction* deals with the non-uniform response across the face of the crystals. The cluster energies are scaled according to a map of the E/p response as a function of transverse position in each crystal. Clusters centered near crystal edges have a lower response than those near the middle of a crystal. Finally, a *Linearity Fudge Correction* is applied to correct for measured non-linear deviations of crystal response in the longitudinal direction. This correction is energy dependent since the longitudinal shower shapes vary with energy.

4.3 Track Vertex Finding

Track vertex finding begins by building a quick list of vertex candidates. The intersection point along the Z axis for each pair of X tracks and each pair of Y tracks is calculated. Uncertainties in the X and Y intersection points are allowed based on a nominal 2 mm uncertainty ascribed to each hit position along the tracks. Any instance of the ranges of an X and a Y intersection overlapping is considered a vertex candidate. The vertex Z position is the weighted average of the X and Y intersections, and the vertex $X(Y)$ position is the average position of the $X(Y)$ tracks at the vertex Z.

An attempt is then made to match the tracks from each vertex candidate to clusters in the CsI. The candidate is retained only if both tracks point to within 7 cm of a CsI cluster. For each vertex candidate both cases of pairing one of the two X tracks with one of the Y tracks are checked.

At this point, several corrections are made to the measured hit positions. The TDC times are adjusted for signal propagation along the wire. Chamber rotation corrections are made to the wire positions. Fringe fields from the magnet can extend to the drift chambers, and hits in DC2 and DC3 are corrected for this effect. Finally,

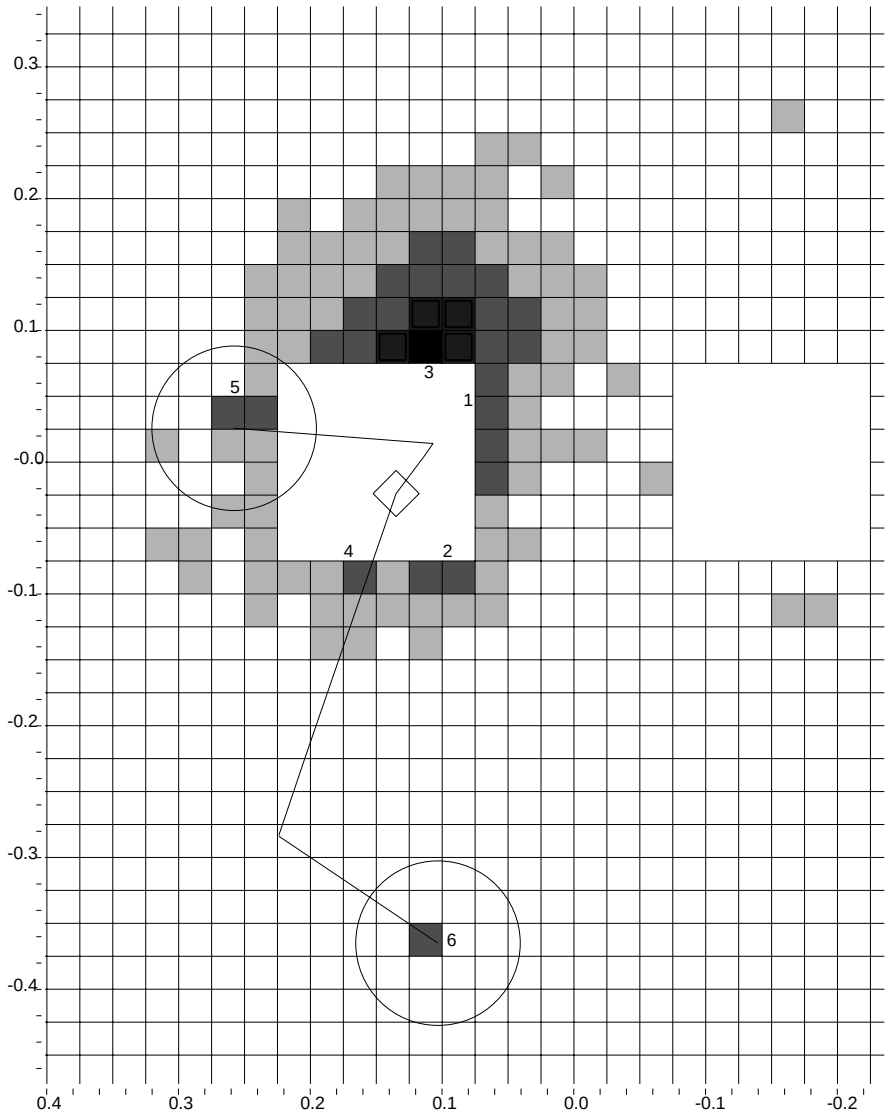

Figure 4.8: Display of CsI energy corrections in a $K_L \to \mu^+\mu^-\gamma$ event. Darker shadings represent larger crystal energies. The photon energy cluster (number 3) is centered on the black crystal, and the circled clusters (numbers 5 and 6) are the muon track clusters. Clusters 3 and 5 have some shared and some adjacent crystals that use the overlap and neighbor corrections respectively. Both clusters 3 and 5 must have corrections for missing blocks in the beam hole. A significant sneaky energy correction is made to cluster 5 due to energy from cluster 3 that crosses the beam hole.

an angular correction is applied to the measured SODs to account for the slopes of the tracks through the cells.

After all corrections have been applied, the vertex position of all remaining candidates are recalculated. A χ^2 is calculated for both the vertex position (χ^2_{vtx}) and the X and Y offsets at the magnet (χ^2_{offmag}). The best vertex candidate is selected by evaluating a combination of χ^2_{vtx}, χ^2_{offmag}, and the number of bad SOD pairs and isolated hits used for the tracks.

CHAPTER 5
MONTE CARLO SIMULATION

A Monte Carlo simulation of KTeV is used for several purposes. Primarily, it is used to calculate and study the acceptance for each decay mode. The acceptance is the reconstruction efficiency for a decay and includes effects of geometry, detector and trigger performance, and analysis cuts. The Monte Carlo is also employed to estimate the expected background in a given decay mode. Comparisons are made between distributions of various reconstructed quantities in data and Monte Carlo. These data/Monte Carlo comparisons are used to estimate systematic errors, or to make direct measurements from the distributions (such as the form factor measurements described in Chapter 7).

The four stages of the Monte Carlo are 1) generating a kaon decay, 2) tracing particles through the detector geometry, 3) simulating the detector response, and 4) simulating the trigger evaluation. The Monte Carlo makes use of an extensive database of calibration information. The calibrations are used in analysis to reconstruct decays from detector information. In the Monte Carlo they are used to simulate detector response to a generated decay. The simulated detector information is output in the same format in which the data is recorded. This enables the procedures for analyzing Monte Carlo and data events to be almost identical.

5.1 Event Generation

5.1.1 Kaon Production and Decay

A kaon generated in the Monte Carlo is characterized by its point of production, momentum, direction, and decay position. The momentum range allowed for kaons

is chosen to be 20 GeV/c to 220 GeV/c; the allowed Z range for kaon decays is 90.0 m to 160.0 m from the target. These ranges are selected because there is very little acceptance for kaons that fall outside them.

The kaon production point is selected to occur at a position within the target that is smeared to reflect the interaction depth of the target, the primary proton beam size, and the targeting angle of the beam. The initial description of the kaon is completed by selecting momentum and direction using distributions adapted from ones parameterized by Malensek [53]. The Malensek distributions are based on measurements of charged kaons produced by 450 GeV protons on a beryllium target [54]. The neutral kaon production cross sections are derived from those of charged kaons by counting valence and sea quarks involved in kaon production [47]. These momentum and angular distributions for neutral kaon production are tuned in the Monte Carlo to match the spectra measured in KTeV from $K_L \to \pi^+\pi^-$ data.

After a kaon has been generated in the target, the location of the decay can be selected. The kaon is first traced to the decay region at $Z = 90.0$ m. The event is rejected if the flight path passes through any of the collimators. Once the kaon reaches the decay region, the Z position for the decay is determined assuming a pure K_L exponential lifetime. Kaon evolution effects due to $K_L - K_S$ interference are neglected. Such interference is negligible since the K_S branching ratios to the final states of the signal, normalization, and all significant backgrounds are very small. Finally, the kaon is traced to the decay vertex position, at which point the chosen decay mode is generated.

5.1.2 Decay Generators

The $K_L \to \mu^+\mu^-\gamma$ signal mode is generated with full radiative corrections (both real and virtual) to $\mathcal{O}(\alpha_{EM}^2)$ [55, 56]. A cutoff of $m_{\gamma\gamma} > 1$ MeV/c^2 is used in evaluating the inner bremsstrahlung component of the corrections. Three different form factors are used for different portions of the analysis. The BMS form factor (Equation 1.23) is employed to generate Monte Carlo for the $K_L \to \mu^+\mu^-\gamma$ branching ratio and α_{K^*} analyses. The DIP form factor (Equation 1.25) is used to evaluate α. A form factor

of unity (flat phase space) is used in the Monte Carlo to make a direct bin-by-bin measurement of the form factor in data (see Chapter 7).

The generator for $K_L \to \pi^+\pi^-\pi^0$ calculates the decay matrix element, M, using phase space modified by the energy dependence of the K_L Dalitz plot,

$$|M|^2 \propto 1 + gu + hu^2 + jv + kv^2 \tag{5.1}$$
$$u = (s_3 - s_0)/m_{\pi^2},\ v = (s_1 - s_2)/m_{\pi^2}$$
$$s_i = (P_K - P_i)^2,\ s_0 = \tfrac{1}{3}\sum_i s_i,\ i = 1, 2, 3$$

where P_K and P_i are the kaon and pion four-vectors [20, page 449]. The π^0 is forced to decay at the kaon decay vertex. The π^0 is stopped and the daughter photons are added to the list of particles to be traced.

$K_L \to \pi^\pm\mu^\mp\nu$ events are the dominant background in this measurement. They are generated with a pure vector form factor [57]. A feature of all background mode generation is that the charged pions are always forced either to decay or to punch through the Muon Filters. These processes must occur for the background events to satisfy the trigger, and forcing them to occur in Monte Carlo speeds up event generation by a factor of 25.

5.2 Particle Tracing

Once a kaon decay has been generated, all of the daughter particles are traced through the detector geometry. Tracing particles involves propagating them along their flight directions to the Z positions of various detector elements. During each propagation step, charged pions are allowed to decay. In the case of a pion decay, tracing of the pion stops and tracing of the daughter muon begins at the decay point. If a particle travels outside the detector geometry at any point during its propagation, the particle is stopped and flagged as "lost." Monte Carlo signal events are rejected if any particles are lost, but any number of particles may be lost when generating background events.

Excluding drift chamber wire planes, there are 0.17 radiation lengths of material

from the vacuum decay region to the CsI (0.05 radiation lengths in the beam region). The material in that region includes the vacuum window, a small air gap preceding DC1, the helium bags, the drift chamber windows and wire planes, the TRDs, and the trigger hodoscopes. The interaction of particles with the matter in each object is simulated according to particle type. Photons can convert to e^+e^- pairs, in which case tracing of the photon is stopped, and tracing of the e^+e^- pair begins. Charged particles undergo multiple scattering, and electrons may also emit bremsstrahlung photons which are then traced along the electron trajectory. For each element, scattering and conversions are simulated as if the matter is located in a plane.

The analysis magnet bends charged particle trajectories horizontally. This is simulated by applying a momentum kick in the X direction at the magnet bend plane (the center of the magnet field along the Z axis). The momentum kick is approximately 205 MeV/c and is calibrated run-by-run with $K_L \to \pi^+\pi^-$ data.

Photon and electron tracing stops if those particles hit the photon vetoes or the CsI. Charged pions also stop in the CsI. Any particles other than muons that reach the back anti are stopped there. Muons are traced all the way through to the MU3 counter planes (details of muon tracing are described in Section 5.3.2).

Information about the particles or detectors can be saved at each step in tracing. The location and direction of each particle is recorded at the drift chamber positions. Particle locations, directions, and energies are saved at the TRDs and at the upstream face of the CsI. At each photon veto, trigger hodoscope, and muon plane, the energy deposits in all hit channels are determined and recorded. After tracing, all of this information is used to simulate the response of individual detectors.

5.3 Detector Simulation

5.3.1 Drift Chambers

For this analysis, it is crucial that the behavior of the drift chambers is well-modeled. Drift chamber performance affects the signal and normalization modes similarly, but is critically important in understanding background rejection. This is because the most powerful analysis cuts hinge on the quality of track and vertex reconstruction.

SODs and Drift Times

The Monte Carlo simulation of the drift chambers is simply the calculation of drift times from the hit positions stored during tracing. That calculation is based on chamber calibrations as well as our understanding of the SOD distributions.

The first step in calculating drift times is to determine the drift distances associated with each hit. For each plane, the distance between each hit and the nearest wire is found. The distance is then smeared according to a Gaussian resolution measured for SODs in that plane. At this point, corrections to the distance are made based on various SOD effects including

- discrete ionization near wires
- electron-pion ionization differences
- high-SOD and single hit inefficiencies (described in Section 4.1.1).

The corrections are detailed in the following three sections. Corrected drift distances are then converted into drift times using the XT calibrations. Finally, the wire-by-wire T0 offsets are subtracted from each drift time to arrive at the simulated drift time associated with each hit.

Discrete Ionization

The drift chamber calibrations assume a continuous average ionization density along particle paths in the chambers. However, for tracks that pass close to wires, random, discrete ionization effects become important. Discrete ionization near wires is manifested in the SOD distributions as an enhancement of the high-SOD tails. Empirical parameterizations of this effect in data were constructed and are used to correct Monte Carlo drift distance calculations. The corrections include smearing the distances toward the high side, and then scaling the distances back so that the means of the SOD distributions are centered on the cell spacing of 6.35 mm.

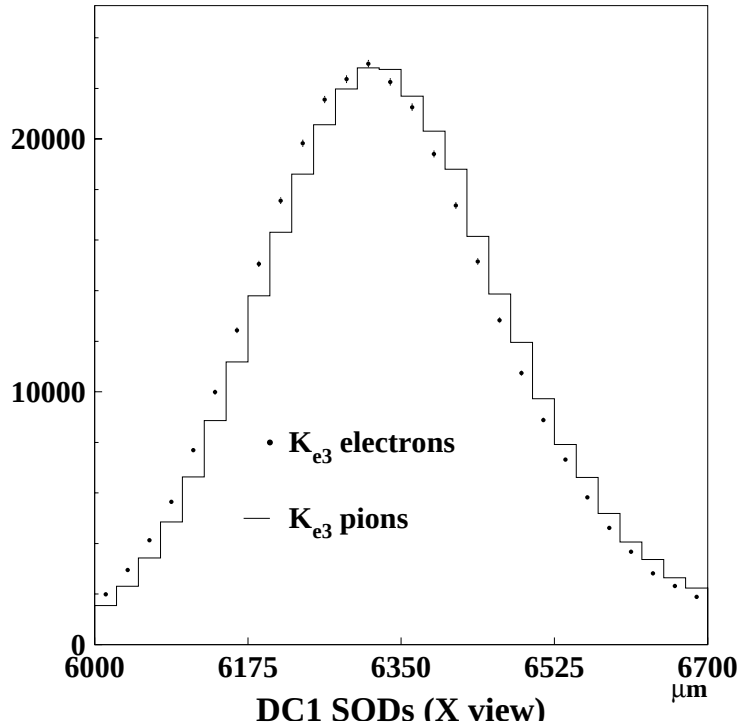

Figure 5.1: e-π SOD difference.

$e^{\pm}$-$\pi^{\pm}$ Ionization Difference

There is a slight mass dependence to ionization energy loss [20, page 144]. Any difference in ionization between particle types would result in different drift time (and SOD) distributions. This effect can be seen in Figure 5.1 as a 15 μm difference between the SOD means of electrons and pions from $K_L \to \pi^{\pm} e^{\mp} \nu$ data. Measurements of electron and pion SOD distributions were made in data and compared to the Monte Carlo distributions after discrete ionization corrections had been made. The data/Monte Carlo ratios of SOD means and widths are used as scale factors to adjust the Monte Carlo drift distances and resolutions to match those in data. Monte Carlo electrons use scale factors measured from $K_L \to \pi^{\pm} e^{\mp} \nu$ electrons. All other charged particles use constants measured from $K_L \to \pi^{\pm} e^{\mp} \nu$ pions because the small relative mass differences among those particles produce no measurable differences in ionization.

DC Maps

Even after thorough and careful drift chamber calibrations, there remained significant discrepancies between the data and Monte Carlo for drift chamber efficiencies and SOD distributions. In an effort to account for this, detailed maps of the drift chambers were generated using $K_L \to \pi^{\pm} e^{\mp} \nu$ data. Maps were made of high-SOD rates that are a result of inefficiency to early drift electrons, and of inefficiency rates corresponding to spots of wire damage. The maps are functions of position in the chamber and time in the run. Spatial binning of the maps is wire-by-wire and in 1 cm divisions along the wires. The segmentation of the map time periods averages about 6 days. Figure 5.2 shows that high-SOD inefficiencies are much higher in the beam region, and that single hit inefficiencies are localized to spots scattered throughout the chambers.

When applied, the maps resolve the data/Monte Carlo discrepancies. Prior to using the drift chamber maps, large track inefficiencies in the beam regions were not modeled by the Monte Carlo. Figure 5.3 shows how much the maps improve the simulation of track illuminations. Another quantity that suffered from poor modeling of drift chamber inefficiencies was the Z position of the reconstructed decay vertex. This was a result of the fact that the inefficiencies are much worse in the beam region. Charged particles from decays near the end of the decay volume would be much more likely to still be within the beam region at the first drift chamber. Without the map simulations, the Monte Carlo underestimated inefficiencies and thereby accepted too many events at high Z. The vast improvement in vertex Z distributions using the maps is shown in Figure 5.4.

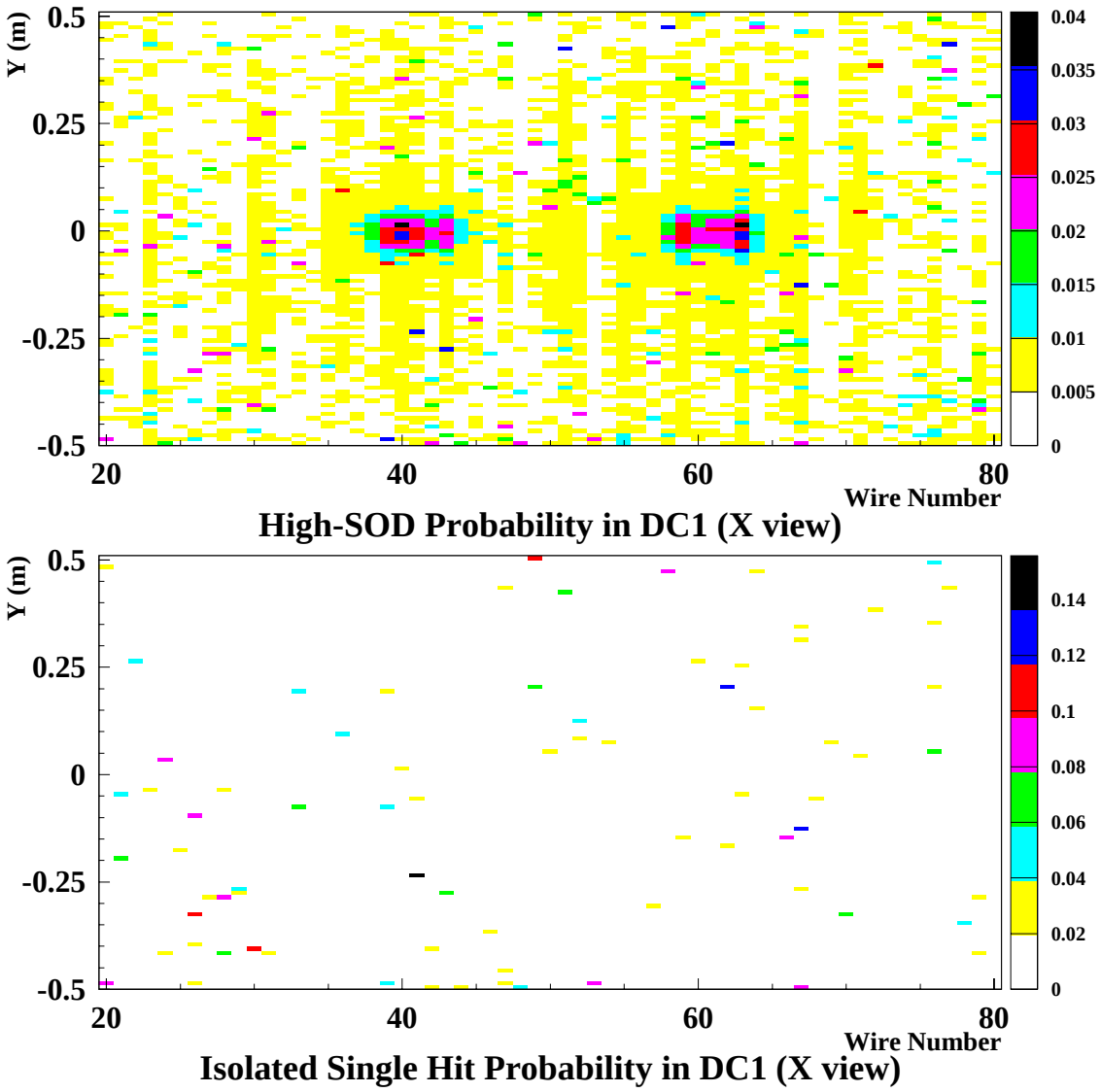

Figure 5.2: High-SOD and single hit inefficiency DC maps.

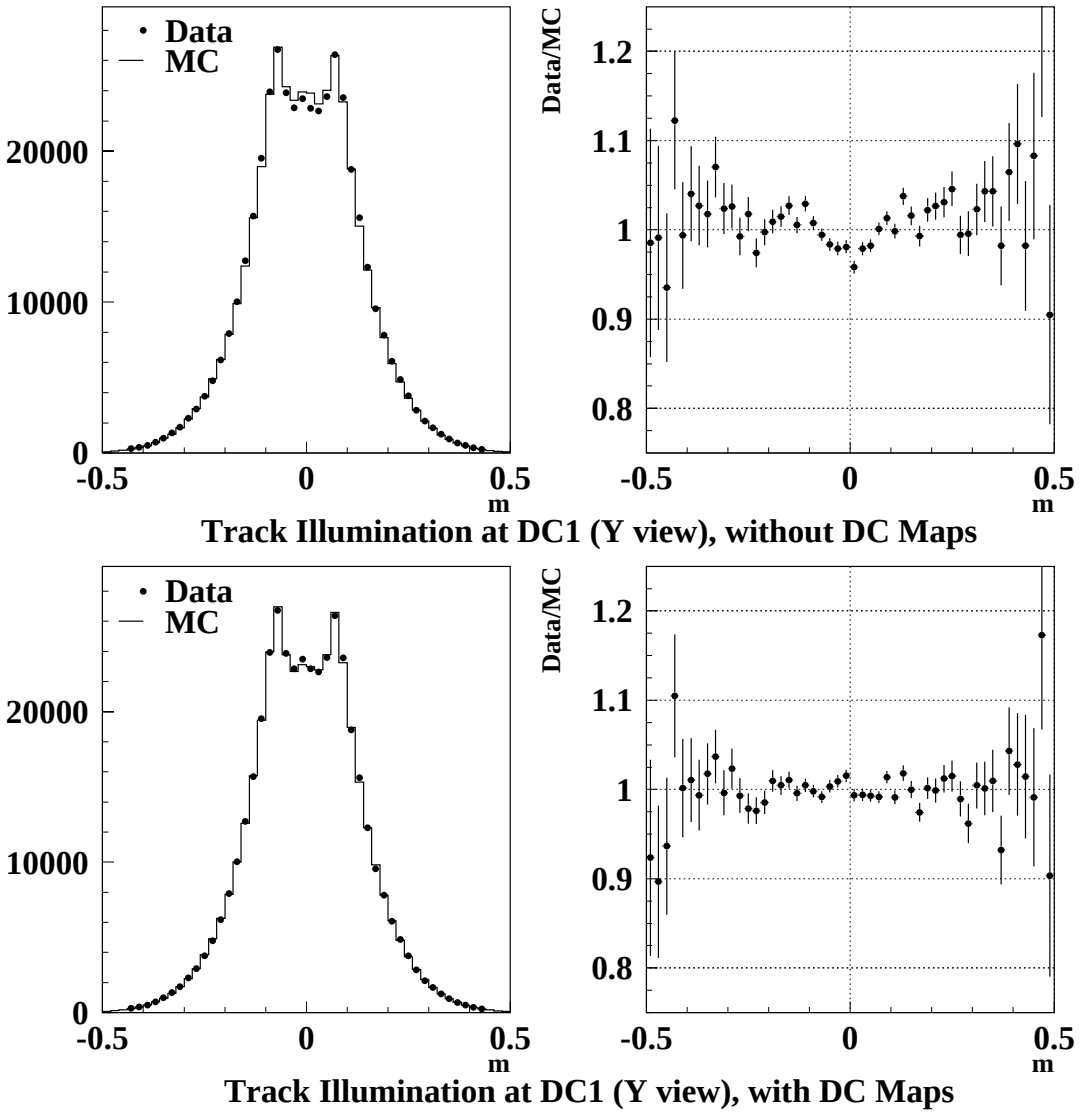

Figure 5.3: $\pi^+\pi^-\pi^0$ drift chamber illumination. The hit positions of charged pion tracks are plotted as a function of vertical position in DC1. Large track inefficiencies in the beam region (upper plots) are well-modeled by the inclusion of the DC maps (lower plots).

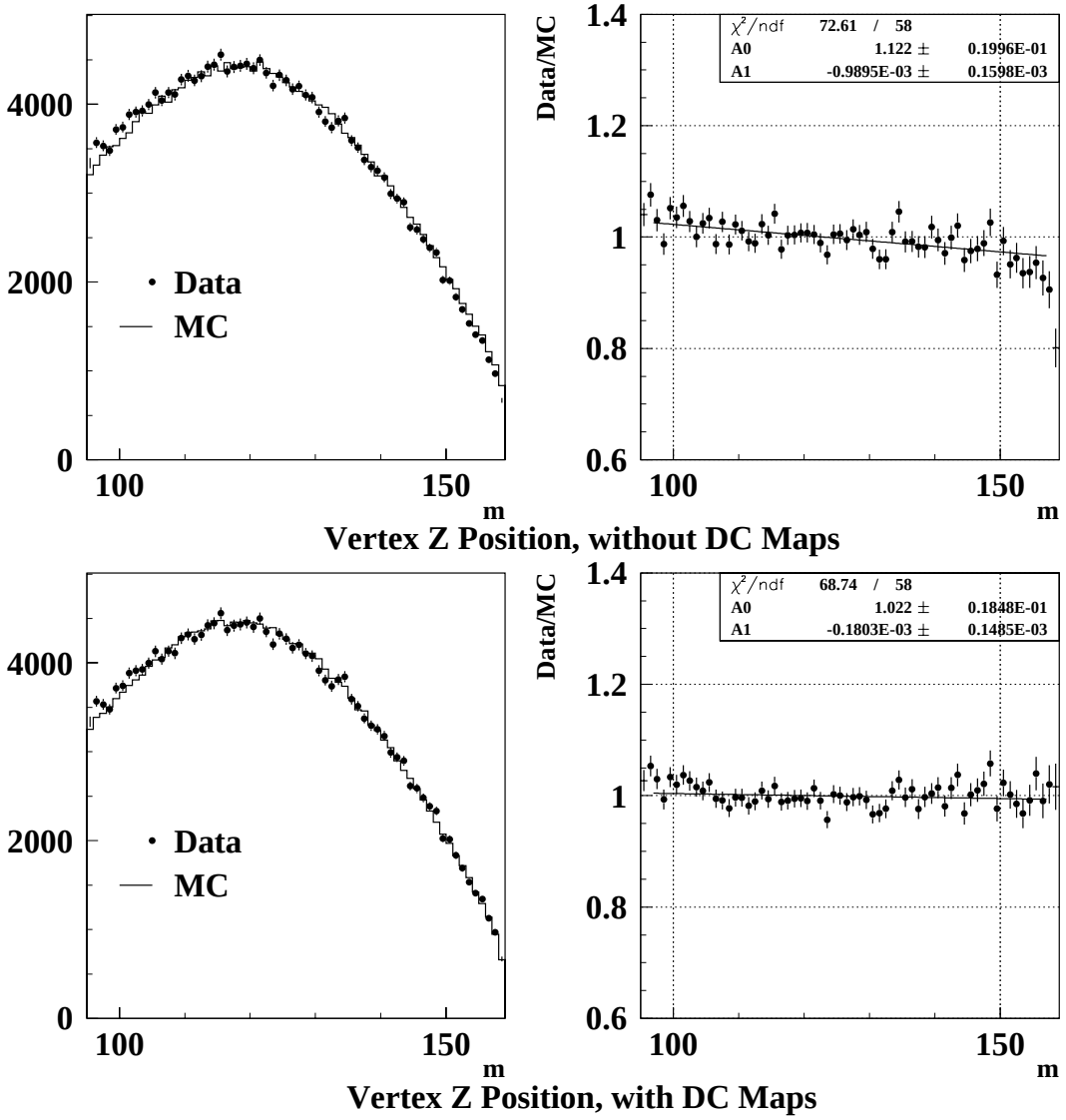

Figure 5.4: $\pi^+\pi^-\pi^0$ vertex Z distribution. Tracks from kaon decays at high Z (the downstream end of the vacuum decay region, near DC1) are more likely to be rejected due to high-SOD measurements. Before inclusion of the DC maps, this effect was not simulated and the resulting data/Monte Carlo mismatch produced a 5σ slope (upper plots). With the DC maps, this slope is almost eliminated (lower plots).

5.3.2 Muon System

The major trigger difference between the signal and normalization modes is the MU3 requirement in the Dimuon Trigger.[1] Since muon requirements are made only in the signal mode, the absolute MU3 acceptance must be determined with great precision and accuracy. This requires a very good understanding of the muon system modeling in the Monte Carlo.

Particle tracing of muons occurs in two stages. Upstream of the CsI, muons are traced in the same manner as all other (non-electron) charged particles. At the face of the CsI, muons are diverted from the standard tracing algorithm into a very detailed simulation of multiple scattering and energy loss. Muons are scattered through the CsI, Lead Wall, Back Anti, and three steel Muon Filters. Energy loss is calculated in each element as well as in the Hadron Anti, MU2, MU3X and MU3Y. Energy deposits in CsI crystals, Back Anti sections, and scintillator counters are stored for detector and trigger simulations later.

The muon multiple scattering is based on simulations using GEANT code [58], and energy loss is calculated using Bethe-Bloch energy loss with fluctuations. Muon acceptances are very sensitive to the momentum threshold, which in the Monte Carlo, is almost entirely governed by the energy loss simulation.

MU3 Cracks

Before evaluating the energy loss in detail by studying muon efficiency thresholds, inefficiencies due to cracks in MU3 must be understood. To study cracks, muon efficiencies were measured using magnet-on muon run data. The efficiency is defined as the number of single track events with a MU3 hit, divided by the total number of single track events. When plotted versus extrapolated track position at the MU3 plane, the cracks along counter sides were measured to be 0.9 mm at the PMTs and 0.3 mm at the ends, 5.6 mm at the MU3X center split, and 3.9 mm at the

[1] Other trigger differences involve CsI cluster counting and photon veto requirements. However, sensitivity to HCC and photon veto modeling is minimized since analysis cuts for both modes are placed well beyond trigger thresholds for these detectors.

MU3Y center split. Figure 5.5 shows that given these numbers, the Monte Carlo reproduces the global muon efficiency to better than 0.05%. Local efficiencies are shown by Figure 5.6 to agree better than 0.5% (1% at the center splits).

Multiple Scattering

The GEANT multiple scattering simulation consists of propagating muons through various scattering elements. Muons with energies ranging from 1 GeV to 128 GeV are scattered through CsI, lead, iron, and a composite material representing the Back Anti (consisting of sandwiched layers of scintillator, iron, and lead). Scattering element thicknesses vary from 5 cm to 400 cm. Scattering angle distributions are generated for each material at specific values of energy and thickness. Each scattering angle distribution is fit to a Gaussian, and the widths for each material are parameterized as a function of energy and thickness.

The parameterized scattering widths are used in the Monte Carlo. For a muon incident on a scattering element, the path length through the entire element is determined. Then the scattering width, σ, is calculated given the path length and the energy of the muon. Corrections are made for muons that scatter out of the element or into a beam hole. The scattering angle and lateral displacement, θ and x, are correlated, and calculated according to the Gaussian approximation of multiple scattering,

$$\theta = R_1 \sigma \tag{5.2}$$

$$x = \frac{(R_1 + R_2/\sqrt{3})\sigma t}{2.387}, \tag{5.3}$$

where R_1 and R_2 are Gaussian random numbers and t is the path length in meters [59]. The position and direction of the muon at the back of the scattering element are adjusted by θ and x, energy loss is calculated, and tracing continues to the next element.

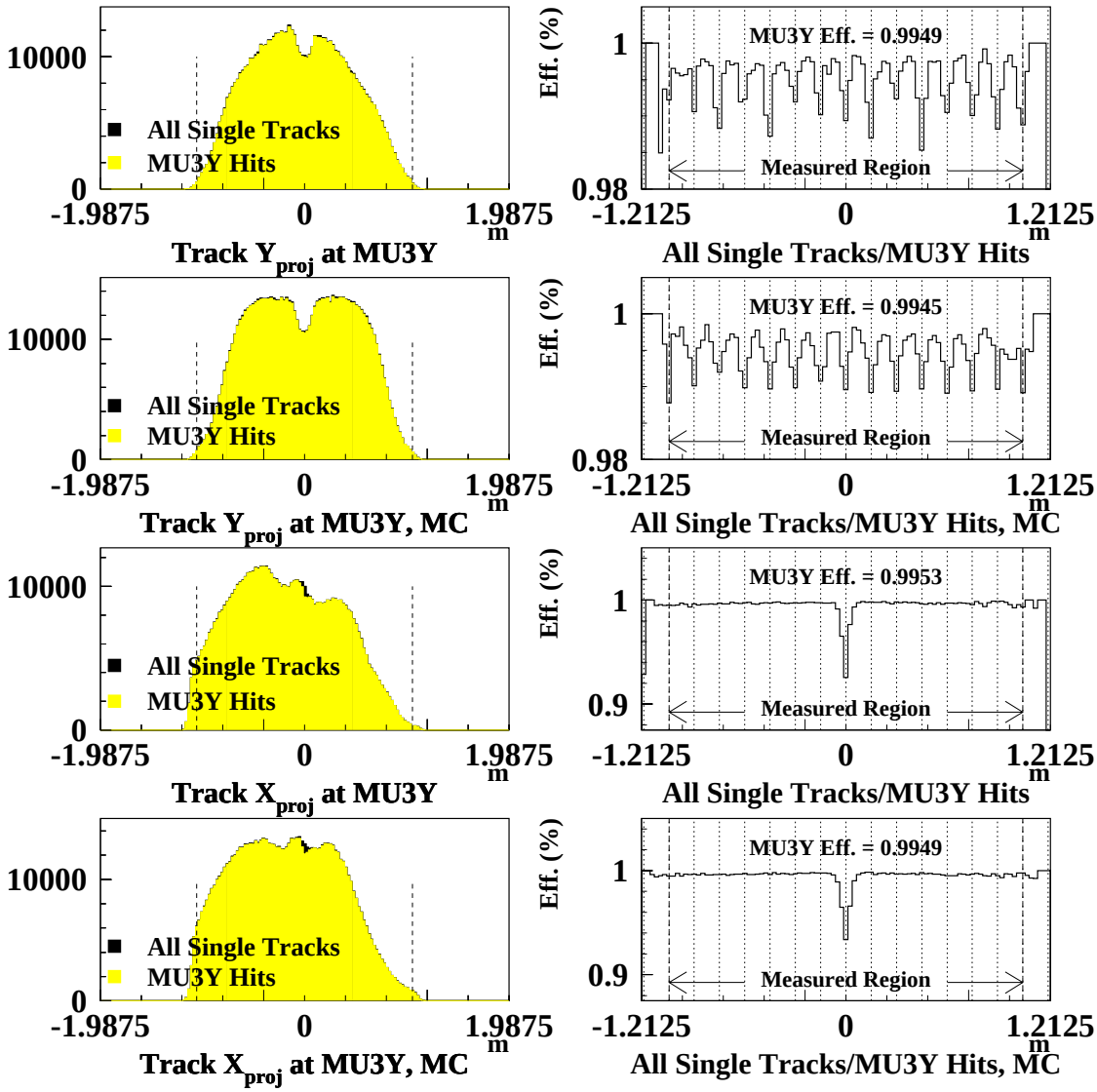

Figure 5.5: Muon efficiencies exhibiting cracks in MU3Y. Efficiencies are plotted as a function of extrapolated track position. Monte Carlo (data) is shown in plots with (without) the 'MC' designation in the title.

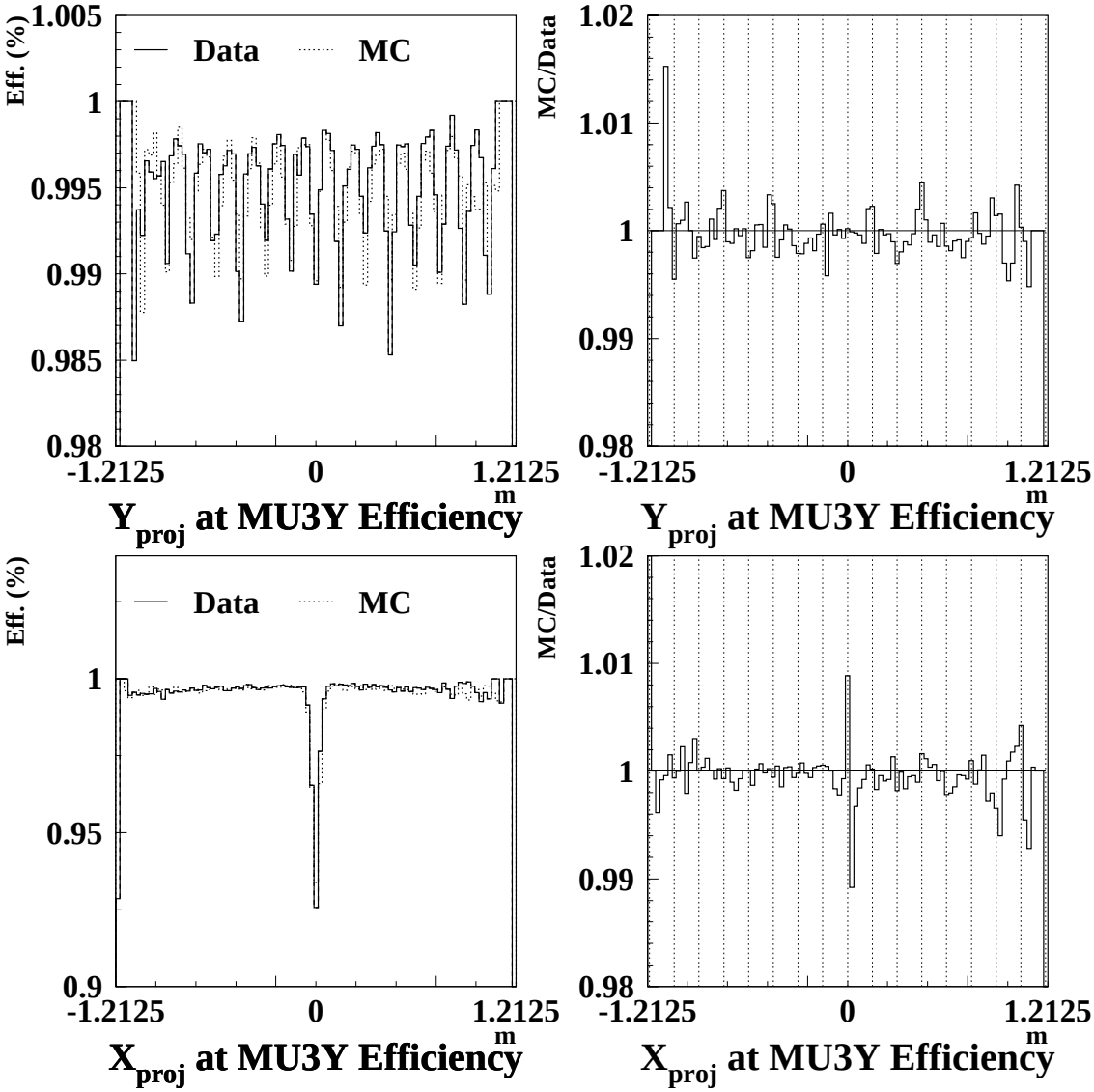

Figure 5.6: MU3Y local muon efficiency agreement between Monte Carlo and data. Efficiencies are plotted as a function of extrapolated track position. The right-hand plots are the ratios of Monte Carlo and data efficiencies.

Table 5.1: Selection of distributions for muon energy loss fluctuations.

Distribution	κ Range
Landau	$\kappa < 0.01$
Vavilov	$0.01 \leq \kappa \leq 10$
Gaussian	$10 < \kappa$

Energy Loss

With the cracks well simulated and a multiple scattering simulation in place, studies of the energy loss modeling can be made by looking at muon efficiency as a function of momentum. To determine the energy loss in a given element, Bethe-Bloch ionization energy loss is calculated [60] and random fluctuations are applied to the result.

Early Monte Carlo simulations used simple Landau fluctuations which resulted in too much energy loss for low-energy muons. The situation was improved by utilizing the full energy loss fluctuation code from GEANT. The full GEANT package still uses Landau fluctuations for the majority of the muon energy range. However, a Vavilov distribution is used for low energy fluctuations, and a Gaussian distribution for very low energies. The choice of distributions is made according to the value of the parameter

$$\kappa = \frac{153.5\rho(Z/A)t}{2m_e}(m_\mu^2 + 2m_e E + m_e^2)\frac{E^2}{p^4},^2 \tag{5.4}$$

and is summarized in Table 5.1 [59]. Figure 5.7 exhibits the excellent threshold agreement using the full GEANT package. The effectiveness of the scattering parameterization is heavily dependent on the energy loss simulation. this is because the scattering widths are functions of incident muon energy. To evaluate the multiple scattering dependence on energy loss, comparisons were made between the positions of projected muon tracks at the MU2 bank and hit counters in MU2. To avoid the

[2] ρ, Z, and A are the density in g/cm^3, atomic number and atomic weight of the scattering element, t is the path length in meters, m_e and m_μ are the electron and muon masses in GeV, and E and p are the muon energy and momentum in GeV.

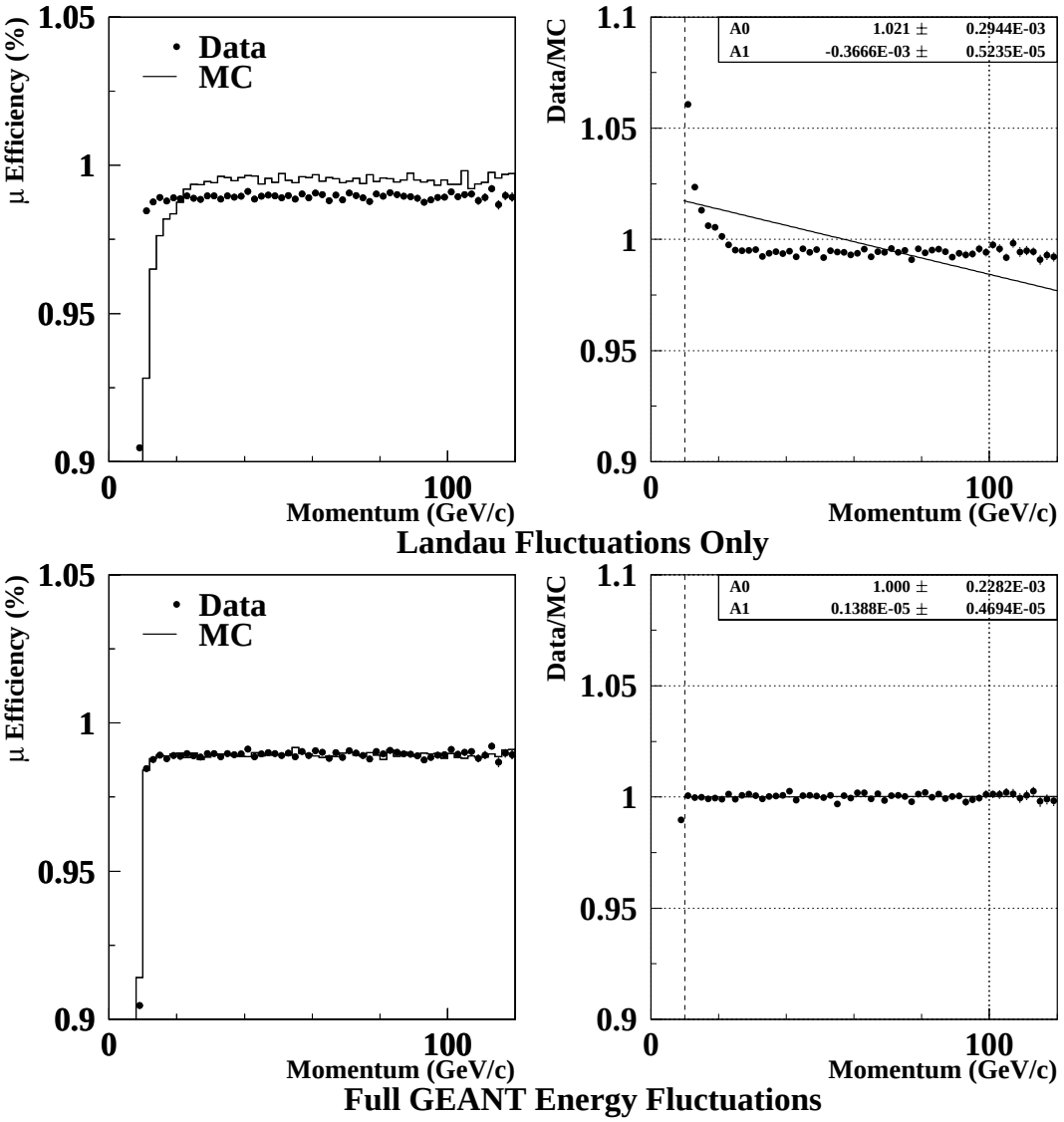

Figure 5.7: Muon efficiency threshold vs. track momentum using energy loss with simple Landau fluctuations (top) and full GEANT fluctuations (bottom).

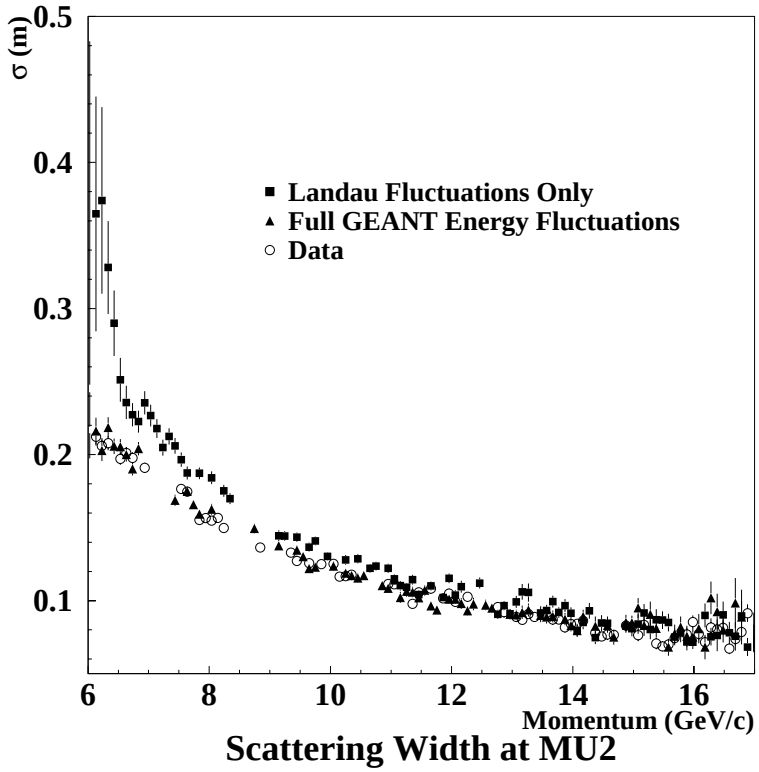

Figure 5.8: Muon multiple scattering vs. track momentum with simple Landau fluctuations and full GEANT fluctuations. σ is the width of Gaussian fits to the differences between the extrapolated muon track positions and the hit counter positions at MU2.

smearing effect due to the 15 cm counter widths, only events that fired both counters in 1 cm overlap regions of MU2 were used. The improvement to the scattering simulation using full GEANT fluctuations can be seen in Figure 5.8.

$\pi^{\pm}$ Misidentification

All of the backgrounds to $K_L \to \mu^+\mu^-\gamma$ come from kaon decay modes that contain at least one charged pion in the final state. These decays can become background when the charged pion is misidentified as a muon, either through pion decay ($\pi \to \mu\nu_\mu$) or pion punch-through.

Punch-through occurs when a pion showers in the CsI, lead wall, or steel, and MU3 is fired by shower products that leak through the filter steel. In the Monte

Carlo simulation, we only consider pion decays upstream of the calorimeter. Pion punch-through is modeled by GEANT simulations, and is defined to include true punch-throughs as well as pion decays past the upstream face of the calorimeter [61]. The probability for pion punch-through is momentum-dependent, and determined from GEANT to be

$$P = 0.00105 + 0.00017p \tag{5.5}$$

for a pion with momentum p.

Each type of pion misidentification can be studied separately in the Monte Carlo by forcing pions to either decay or punch through. The absolute acceptance of pion decays can be calculated by assigning weights to the Monte Carlo events. An event's weight is the probability for the pion to decay between the kaon decay vertex and the CsI, given the pion's momentum. Punch-throughs in the Monte Carlo are simulated with the momentum-dependence given by Equation 5.5. The relative acceptance of pion punch-throughs to decays is found by studying the kaon mass distribution of $K_L \to \pi^+\pi^-\pi^0$ events recorded in the Dimuon Trigger. Figure 5.9 shows that pion decay events are characterized by a wide mass distribution due to the smeared kinematics, while pion punch-throughs yield a much narrower peak since they have better tracking. The relative amount of decay and punch-through at Monte Carlo generation was determined by floating the ratio of decay to punch-through, until the best agreement between data and Monte Carlo was found. Using this procedure, the punch-through/decay ratio was found to be 0.18 ± 0.04 (Figure 5.10).

5.3.3 CsI Calorimeter

The two steps in the CsI calorimeter simulation are generating clusters and digitizing the signals from each crystal. Generating clusters consists of determining the amount of energy deposited by a particle, and how that energy is distributed among the crystals in the CsI. The method used to generate a cluster depends on the particle type. Digitization involves dividing the energy into readout time slices, and then converting the energy deposits into digital signals.

Photon and electron clusters in the CsI are simulated using a library of photon

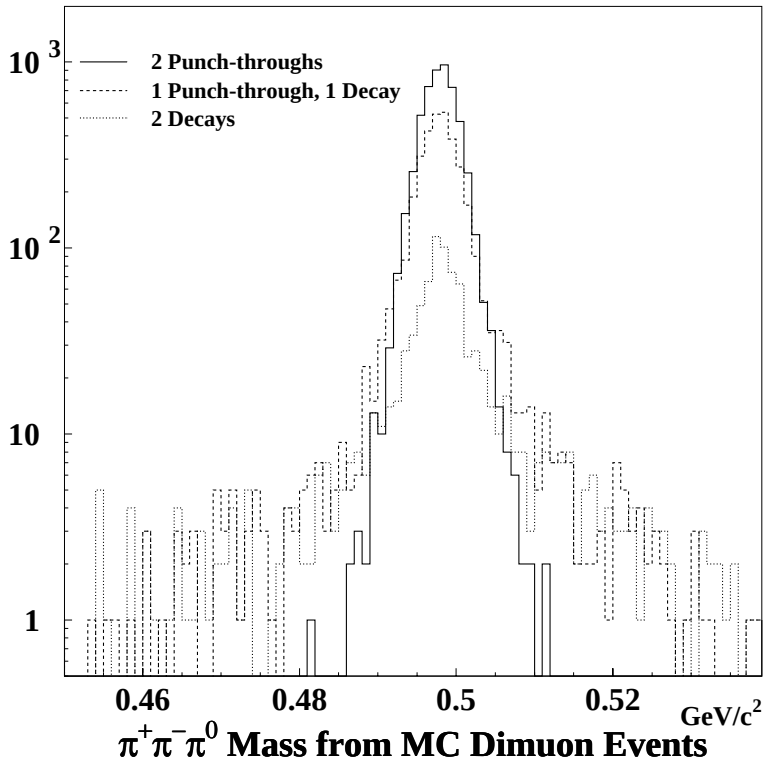

Figure 5.9: $K_L \to \pi^+\pi^-\pi^0$ mass distributions for two punch-throughs (PP), one decay and one punch-through (PD), and two pion decays (DD).

showers generated with GEANT [62]. The GEANT shower libraries were generated using a 13×13 array of small crystals, and are indexed by the incident particle's energy and transverse position within the hit crystal. Before selecting a shower from the library, the particle energy is smeared slightly to match the resolutions measured in data. Also, the transverse position is evaluated at the Z position of the shower mean within the crystal. The shower mean is a function of energy, and slightly different for photons and electrons:

$$Z_\gamma = 0.12 + 0.018 \ln E \tag{5.6}$$

$$Z_e = 0.11 + 0.018 \ln E, \tag{5.7}$$

with E in GeV and Z in meters from the upstream face of the CsI. A shower is then chosen from the library based on the adjusted energy and position. The information

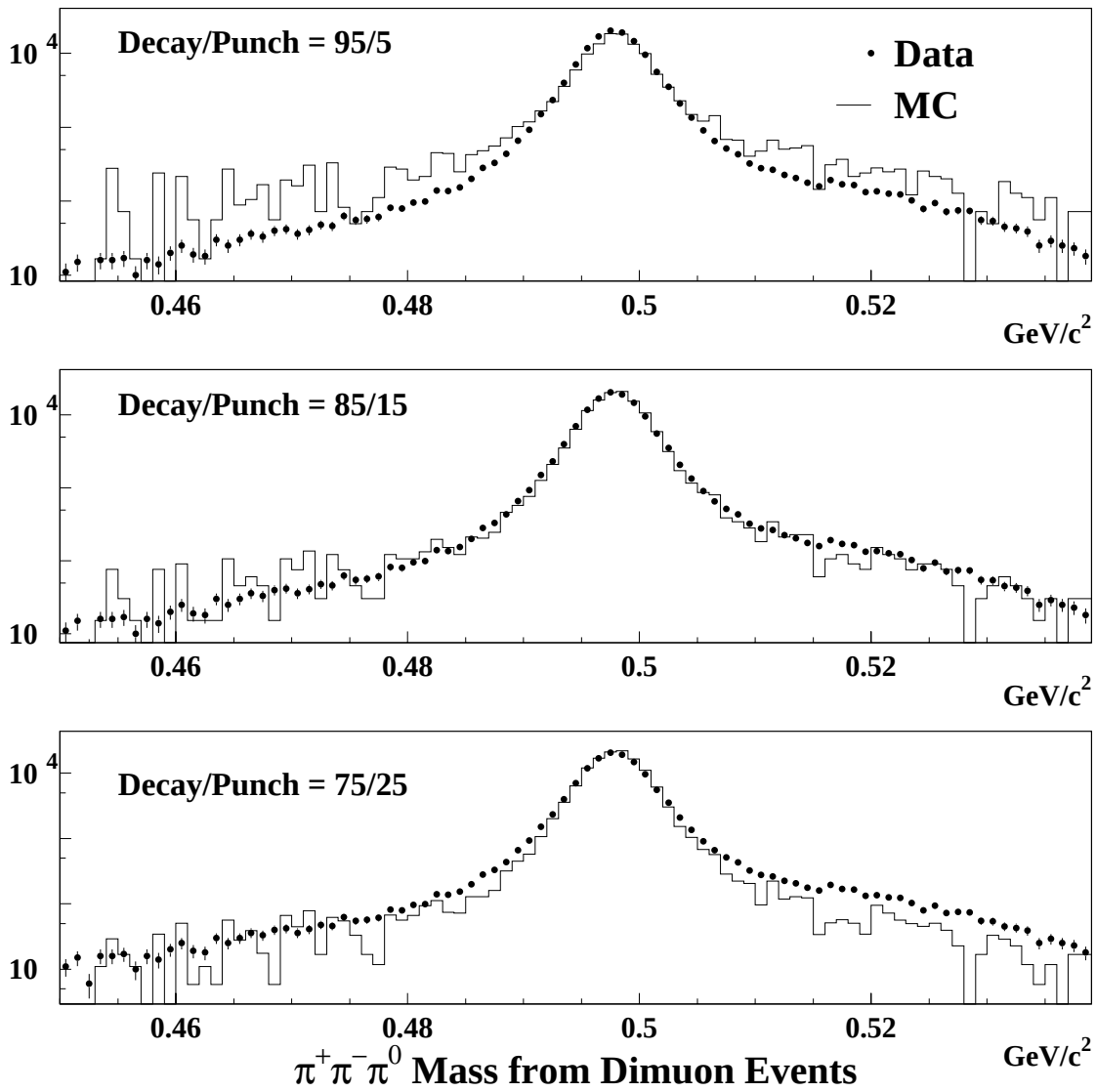

Figure 5.10: Evaluation of pion decay/punch-through ratio through comparisons of $K_L \to \pi^+\pi^-\pi^0$ mass distributions in data and Monte Carlo.

from the shower is used to divide all of the energy within the 13×13 region of small crystals centered on the hit crystal.[3] The showers are segmented in 25 bins in the Z direction for the central four blocks within a cluster, and convoluted with the longitudinal uniformity in light output measured for each crystal. Finally, cluster energy that crosses a beam hole is simulated by adding "sneaky energy" (from a separate library) to crystals on the other side of the hole.

The shower library used to generate clusters from charged pions is also built from GEANT simulations. Like its electromagnetic counterpart, this library of pion showers is indexed by energy and transverse position. However, no adjustments are made to the pion's energy or position at the face of the CsI prior to selecting a shower. In addition to distributing energy among CsI crystals, the pion showers also simulate energy deposits in the Lead Wall and Hadron Anti.

Energy deposited by a muon is confined to a single crystal, and handled by the muon simulation described in Section 5.3.2. For other minimum ionizing particles, 320 MeV is deposited in the hit crystal.

Before digitizing the signals in each crystal, the time structure of the DPMT readout must be imposed on the Monte Carlo energy deposits. The readout occurs in six time slices, and the energy in a crystal is divided among the slices according to fractions obtained from pulse shape data. The energies in each slice for each crystal are then digitized using the Q/E calibrations to calculate collected charge, and the DPMT calibrations to convert charge to digital signal.

5.3.4 Photon Vetoes

The energy deposited in photon vetoes by photons and electrons is smeared with Gaussian resolutions. The resolutions are energy-dependent and measured with photons from $K_L \to \pi^+\pi^-\pi^0$ data. In the Ring Counters, Spectrometer Antis, and Collar Anti, other charged particles deposit minimum ionizing energy with constant Gaussian resolutions measured from muon data. In the back anti, photon and electron energy is shared between the two EM sections, and hadron energy is

[3] Large crystals are treated as the sum of a 2×2 array of four small crystals.

divided among all three sections. The energy sharing fractions are obtained from $K_L \to \pi^+\pi^-\pi^0$ photons and $\Lambda \to n\pi^0$ neutrons.

5.3.5 Trigger Hodoscopes

The trigger hodoscope simulation determines which counters are hit by charged particles, accounting for the cracks between counters. To model thresholds, counter efficiencies are measured from data, and the efficiencies are implemented as probabilities for an incident particle to pass threshold.

5.3.6 Accidental Activity

There can be considerable activity in the detector that is in time, but unassociated with, a kaon decay. This "accidental" activity can stem from things such as beam interactions, muons from the beam dump, neutrons, or even a second kaon decay. Accidentals can cause events to be rejected by firing a photon veto for example. On the other hand, extra energy in the CsI could raise clusters above the HCC thresholds. This would cause events to be accepted that would otherwise be discarded. To simulate this activity, data taken from an accidental trigger is added to events generated in Monte Carlo.

The ACC90 trigger was used to collect accidental events. The data in this trigger is proportional to the beam intensity, but uncorrelated with kaon decays in the detector. These characteristics ensure that the accidental events are accurate snapshots of the underlying activity in the detector.

An accidental event is overlaid on each Monte Carlo event. Accidentals are selected only from the run in which the Monte Carlo event is generated. This is to reflect the correlation between accidental activity and beam intensity, which varies run-by-run. An overlay adds an accidental event's energies and hits for each active detector element, to the information generated for a Monte Carlo event. This occurs during individual detector simulations before thresholds and triggers are checked. After the accidental overlay, the trigger can be evaluated based on the combined detector information.

5.4 Trigger Simulation

In most cases, the L1 trigger sources are simulated by simple comparisons of energy deposits to sharp thresholds for individual detector channels. The E-total system uses more detailed efficiency and threshold curves, and the DCOR trigger sources are set using the complete list of drift chamber hits. The L1 trigger is partially evaluated immediately after particle tracing. This is to reject events that fire the photon vetoes before extensive CPU time is spent simulating the detector. After all detectors except the CsI have been simulated, all L1 sources except E-total are evaluated. If the event passes that evaluation, the CsI is simulated and the E-total L1 source is checked.

After the L1 trigger has been fully evaluated, the L2 trigger systems are simulated according to the same algorithms used in the hardware processors. The drift chamber hits are used to model the Hit-Counting system, and the HCC simulation compares cluster energies to threshold curves measured from data.

The Monte Carlo uses L1 and L2 trigger definition maps identical to those used online in hardware. In addition, the online L3 filtering software is the same code used in Monte Carlo to simulate the L3 trigger after L2 evaluation is completed. This ensures that the trigger decision process is the same for data and Monte Carlo events.

5.5 Monte Carlo Samples

Large Monte Carlo samples were generated for the normalization, signal, and background modes. The normalization $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo sample is approximately 9 times larger than the data, and the $K_L \to \mu^+\mu^-\gamma$ Monte Carlo sample is 23 times the size of the data set.

Background Monte Carlo samples are all at least twice as large as the data. However, the actual number of generated events for each background decay mode is smaller than the corresponding number of decays in the KTeV experiment. This is because the *effective* sizes of background samples are much larger due to forced pion decay or punch-through weighting. Additionally, the generation of $K_L \to$

Table 5.2: Number of generated events for all Monte Carlo samples. (D) refers to pion decay and (P) refers to pion punch-through.

Decay Mode	Branching Ratio	Generated Events Winter	Summer
Signal:			
$K_L \to \mu^+\mu^-\gamma$	$3.25 \times 10^{-7\,\dagger}$	1.37×10^6	1.15×10^6
Normalization:			
$K_L \to \pi^+\pi^-\pi^0$	$0.1256^{\,\dagger}$	2.75×10^7	2.19×10^7
Background:			
$K_L \to \pi^\pm\mu^\mp\nu$ (D)	$0.2717^{\,\dagger}$	6.06×10^7	4.57×10^7
$K_L \to \pi^\pm\mu^\mp\nu$ (P)		1.50×10^7	1.34×10^7
$K_L \to \pi^\pm\mu^\mp\nu\gamma$ (D)	$5.7 \times 10^{-4\,\ddagger}$	6.00×10^7	4.50×10^7
$K_L \to \pi^\pm\mu^\mp\nu\gamma$ (P)		1.70×10^7	1.27×10^7
$K_L \to \pi^0\pi^\pm\mu^\mp\nu$ (D)	$2 \times 10^{-5\,*}$	5.60×10^5	4.20×10^5
$K_L \to \pi^0\pi^\pm\mu^\mp\nu$ (P)		9.59×10^4	7.29×10^4
$K_L \to \pi^+\pi^-\pi^0$ (DD)	$0.1256^{\,\dagger}$	6.22×10^7	4.66×10^7
$K_L \to \pi^+\pi^-\pi^0$ (DP)		11.0×10^6	8.24×10^6
$K_L \to \pi^+\pi^-\pi^0$ (PP)		1.92×10^6	1.47×10^6
$K_L \to \pi^+\pi^-$ (DD)	$2.067 \times 10^{-3\,\dagger}$	2.52×10^6	1.87×10^6
$K_L \to \pi^+\pi^-$ (DP)		4.45×10^5	3.30×10^5
$K_L \to \pi^+\pi^-$ (PP)		7.89×10^4	5.89×10^4
$K_L \to \pi^+\pi^-\gamma$ (DD)	$4.61 \times 10^{-5\,\dagger}$	1.60×10^5	1.20×10^5
$K_L \to \pi^+\pi^-\gamma$ (DP)		2.79×10^4	2.09×10^4
$K_L \to \pi^+\pi^-\gamma$ (PP)		4.72×10^3	3.55×10^3

† [20] ‡ [5]
* estimated from $K_L \to \pi^0\pi^\pm e^\mp\nu$ and similar K^+ modes

$\pi^\pm\mu^\mp\nu$ background utilized only a small subset of accidental events that contained hardware clusters. Only 15% of accidentals contained such clusters, and this fraction is included as an additional weighting factor for $K_L \to \pi^\pm\mu^\mp\nu$ Monte Carlo.

For all modes, separate samples were generated for the winter and summer runs. A summary of all Monte Carlo samples is given in Table 5.2.

CHAPTER 6
BRANCHING RATIO ANALYSIS

Measuring the branching ratio of a given initial state decaying to a particular final state is essentially a counting experiment. In this case, our initial state is a K_L particle, and our final state is $\mu^+\mu^-\gamma$. Our task is to count how many K_L's decay in our experiment and how many times the $\mu^+\mu^-\gamma$ final state occurs as a result of a K_L decay. The branching ratio then is simply

$$BR(K_L \to \mu^+\mu^-\gamma) \ = \frac{\#K_L \to \mu^+\mu^-\gamma}{\#K_L \to anything}. \tag{6.1}$$

Of course, not all of the K_L decays that occur will be observed and recorded. The *acceptance* for a given decay mode is the fraction of those decays that are observed by the experiment and correctly identified through the analysis algorithms. The acceptance is a product of the KTeV geometry, the detector response, the trigger performance, and the efficiencies of reconstruction and analysis cuts. Acceptances are calculated using the Monte Carlo simulations described in Chapter 5.

As seen in Equation 6.1, the $K_L \to \mu^+\mu^-\gamma$ signal mode is normalized to the total number of K_L decays, also known as the K_L flux. If we rearrange that equation, it is clear that the flux can be determined by counting the number of decays to a mode with a known branching ratio. For this analysis, the normalization mode is chosen to be $K_L \to \pi^+\pi^-\pi^0$ which yields

$$BR(K_L \to \mu^+\mu^-\gamma) \ = (N_{\mu^+\mu^-\gamma}/A_{\mu^+\mu^-\gamma}) \left(\frac{N_{\pi^+\pi^-\pi^0}/A_{\pi^+\pi^-\pi^0}}{BR(K_L \to \pi^+\pi^-\pi^0)} \right)^{-1}, \tag{6.2}$$

where N is number of observed events and A is acceptance. This choice is made because $K_L \to \pi^+\pi^-\pi^0$ is copious, has a well-measured branching ratio, and has a

final state similar to the signal, especially with regard to kinematics. Because of the similarity of the two modes, many possible sources of systematic error will cancel in the ratio of acceptances.

Just because an event with an apparent $\mu^+\mu^-\gamma$ final state seems to be the result of a $K_L \to \mu^+\mu^-\gamma$ decay, "it ain't necessarily so [63]." Various analysis cuts are made to ensure that the events are the result of a kaon decay, and to reject background events from other decay modes. Finally, Monte Carlo and other techniques are used to estimate the level of background remaining in the final samples. At that point, the number of $\mu^+\mu^-\gamma$ events resulting from actual $K_L \to \mu^+\mu^-\gamma$ decays is known and the branching ratio can be calculated.

6.1 Reconstruction

Because $K_L \to \mu^+\mu^-\gamma$ and $K_L \to \pi^+\pi^-\pi^0$ have similar kinematics, event reconstruction is almost identical for the signal and normalization modes. Both involve a three-body decay, and the charged particles in each case have comparable masses. The difference in reconstructing the two modes is that the normalization has a second photon which must combine with the first photon to form a π^0 mass. Tracking, clustering, and vertexing are performed as described in Chapter 4 for both modes. A small number of cuts are applied during reconstruction to select events with a good two-track vertex and the requisite number of photon clusters (Table 6.1). Some of the reconstruction cuts are identical to those imposed during the crunch; however, there can be some additional events rejected because improved calibrations are used during the analysis.

The first cut is to remove data that was recorded during periods in which there was some problem with the detector. A 32-bit mask was used to flag a variety of detector problems on a spill-by-spill basis. In this analysis, "bad spills" were defined as those in which one of the following problems manifested:

- trigger problem (e.g. bad trigger timing)
- DPMT pedestal exponent > 0 (exponent bit is stuck, resulting in false energy readout even in pedestal events)

Table 6.1: Reconstruction cut efficiencies. Signal and normalization mode cuts are identical unless otherwise indicated.

Signal Mode Reconstruction Cut	$K_L \to \mu^+\mu^-\gamma$ Data	MC	$K_L \to \pi^+\pi^-\pi^0$ Data	MC	Normalization Mode Reconstruction Cut
Bad spills	0.969	0.974	0.969	0.982	Bad spills
At least two X and Y tracks	0.998	1.000	0.987	0.982	At least two X and Y tracks
At least one vertex and two track-cluster matches	0.995	0.999	0.962	0.992	At least one vertex and two track-cluster matches
Exactly two tracks	0.883	0.989	0.459	0.975	Exactly two tracks
Exactly one photon cluster	0.563	0.979	0.343	0.721	Exactly two photon clusters
			0.737	0.952	125 MeV/c^2 $< m_{\pi^0}$ $<$ 145 MeV/c^2

- bad DPMT capacitor (typically disables a CsI channel)
- blown QIE comparator (registers an overflow in the channel)
- dead DPMT (miscellaneous cause of death, pending autopsy)
- digital pipeline problem (corrupted CsI information)
- global CsI problem (voltage, temperature, etc.)
- ADC problem (pedestal shifts or failures)
- drift chambers (voltage trips, amplifier oscillations)
- trigger hodoscope problem (HV, timing, etc.)
- muon system problem (HV, timing, etc.)
- HCC problem (cable swapped prior to run 8245)
- Kumquat/Banana system problem (unstable readout)

- DAQ/L3 problem (corrupted data processing)
- not E799 run (some special run such as pedestal or muon)
- short run (typically less than 20 minutes ending with severe problem)
- beam problem (intensity spike)

About 3% of the data past run 8245 are removed with these cuts.[1] Less are removed from the Monte Carlo because short runs were not generated.

The reconstruction requires at least two X and Y tracks that form a viable vertex candidate. In addition, each of the two tracks forming the vertex must extrapolate to within 7 cm of separate clusters. The tracking is more restrictive here than in the crunch, in that no extra tracks are allowed: there must be exactly two reconstructed tracks. This rejects much 2TRK data consisting of accidental activity.

At this point, the reconstruction of the two modes diverges. For the signal, exactly one photon cluster is required. A photon cluster is a hardware cluster that does not have a track pointing to it. Almost half of the 2MULD data is rejected with this cut. Most of the events rejected are $K_L \to \pi^+\pi^-\pi^0$ with pion decays or punch-throughs. Exactly two photon clusters are required in the normalization. Among the data rejected by the two-photon cut are $K_L \to \pi^\pm e^\mp \nu$ and $K_L \to \pi^\pm \mu^\mp \nu$ events with an accidental cluster, and $K_L \to \pi^0\pi^0\pi^0$ with a photon conversion. The final reconstruction cut is only for the normalization mode. The invariant mass of the two photons must be within 10 MeV/c^2 of the π^0 mass.

The reconstructed kaon mass for the signal and normalization can be seen in Figure 6.1. A $K_L \to \mu^+\mu^-\gamma$ peak is already apparent at the kaon mass on top of a large background of $K_L \to \pi^+\pi^-\pi^0$ (low mass peak) and $K_L \to \pi^\pm\mu^\mp\nu$ (exponential slope under $K_L \to \mu^+\mu^-\gamma$ peak). Much of the high-side tail off of the peak in the 2TRK data is $K_L \to \pi^+\pi^-\pi^0$ events with pion decays in the spectrometer. Such decays cause mismeasurements of track momenta.

[1] All of the data before run 8245 were removed due to the swapped HCC cable.

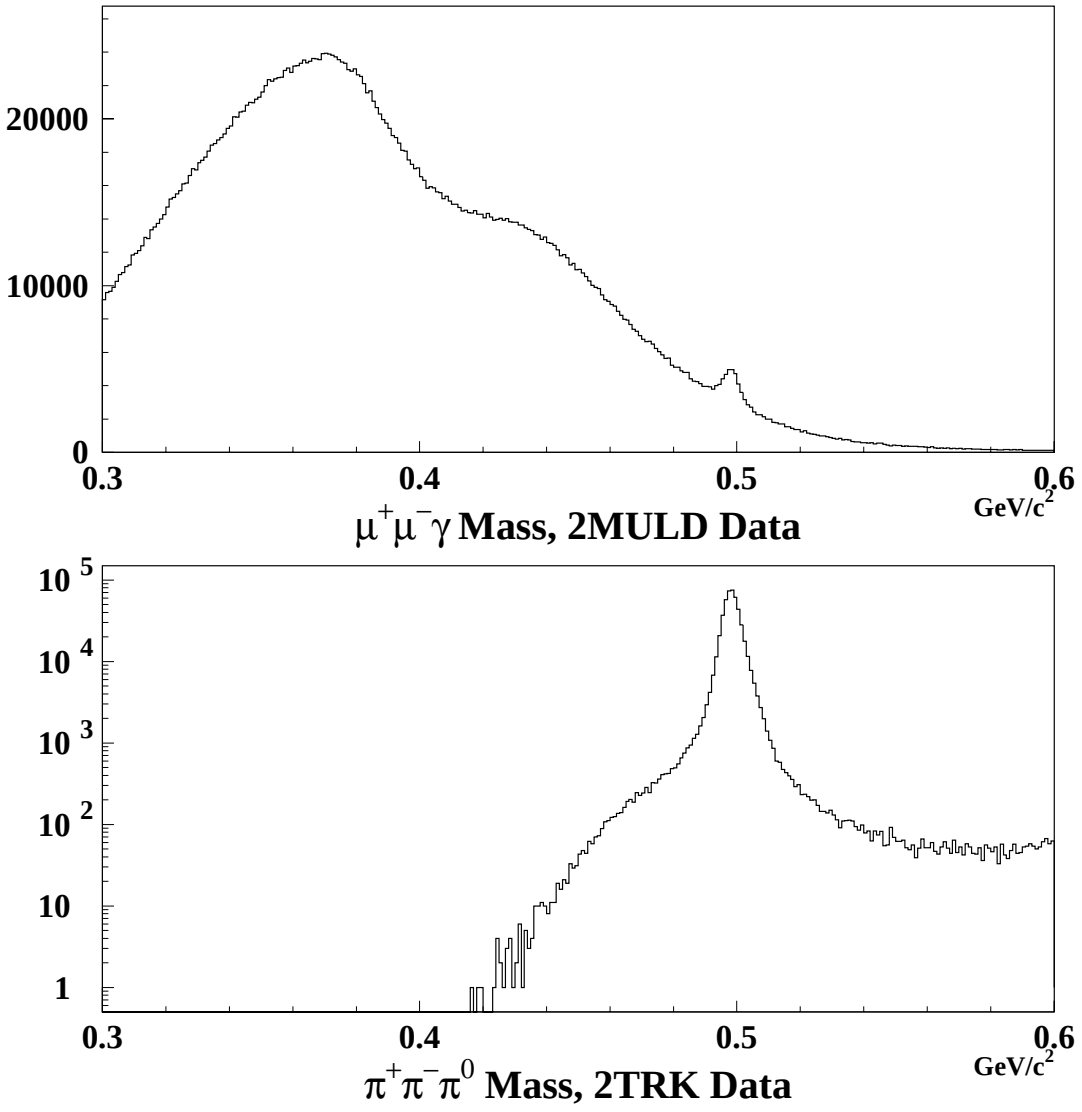

Figure 6.1: Reconstructed mass in the signal and normalization. Both plots show the data after reconstruction cuts only. In the $\mu^+\mu^-\gamma$ plot, a small $K_L \to \mu^+\mu^-\gamma$ peak is already visible at the K_L mass. The large peak at ~ 370 MeV is $K_L \to \pi^+\pi^-\pi^0$, and most of the background above 400 MeV is $K_L \to \pi^\pm\mu^\mp\nu$. Most of the high side tail in the 2TRK data is $K_L \to \pi^+\pi^-\pi^0$ events with charged pion decays. The rest of the tails on both sides comes from tracks mismeasured as a result of finite drift chamber resolution.

6.2 Backgrounds

6.2.1 Global Issues

As was mentioned in Section 5.3.2, all $K_L \to \mu^+\mu^-\gamma$ backgrounds come from kaon decays with final state charged pions which are misidentified as muons. The first of two mechanisms for this misidentification is pion decay ($\pi \to \mu\nu_\mu$). About 4% of all charged pions will decay before reaching the CsI. Pion decays can result in poorly reconstructed vertices and track segment offsets at the magnet. This will happen if the pion decay occurs upstream of or within the spectrometer. Additionally, the unobserved neutrino always carries off a portion of the original kaon energy. However, finite resolutions in the drift chambers smear the kinematics and allow some pion decays to effectively fake a muon coming from the kaon decay vertex.

Pion punch-through represents the other process for pion misidentification. A charged pion that does not decay can hadronically shower in the CsI, lead wall, or Muon Filter steel. If some of the shower products leak all the way through the filter steel and fire MU3, then the process is termed pion punch-through. For the purposes of this analysis, we define punch-through to also include the pion decays which occur past the front face of the CsI. This is to match the definitions and procedures used in the Monte Carlo simulations (see Section 5.3.2).The punch-through probability depends on pion momentum (Equation 5.5) and is about 2% for 100 GeV/c pions.

Overall, pions decay more than 5 times as often as they punch through. However, events involving pion punch-through are much more likely to reconstruct well. This is primarily due to the fact that the tracking quality tends to be much better for punch-throughs, which do not have kinks in the tracks typical of pion decays.

All backgrounds are also characterized by missing energy from the parent K_L. In each case, there are unobserved photons and/or neutrinos. Sometimes accidental activity will make up for the missing energy, but it is more likely to under-compensate or over-compensate. In the rare event that mismeasurements or accidental activity do adequately compensate for the missing energy, there will likely be a momentum imbalance in the reconstruction. For these reasons, cuts on the reconstructed

$\mu^+\mu^-\gamma$ mass and total momentum transverse to the kaon flight direction are the most powerful tools for rejecting most backgrounds.

6.2.2 $K_L \to \pi^+\pi^-\pi^0$

This copious decay accounts for the majority of the data recorded in the 2MULD trigger.[2] However, it contributes very little to the final $K_L \to \mu^+\mu^-\gamma$ background. Cutting events that have extra photon clusters means that $K_L \to \pi^+\pi^-\pi^0$ decays will only be accepted if one of the photons escapes the detector. The missing photon and 34 MeV π-μ mass difference generally result in a very low reconstructed $\mu^+\mu^-\gamma$ mass.

6.2.3 $K_L \to \pi^+\pi^-\gamma$ *and* $K_L \to \pi^+\pi^-$

$K_L \to \pi^+\pi^-\gamma$ background includes direct emission and inner bremsstrahlung components. $K_L \to \pi^+\pi^-$ can become a potential background with either external bremsstrahlung or an accidental photon. Both of these modes tend to reconstruct close to the K_L mass. For example, in the case of $K_L \to \pi^+\pi^-\gamma$ where both pions punch-through, the only offset in reconstructing the kaon mass comes from the small mass difference between pions and muons. However, the very low branching ratio of $K_L \to \pi^+\pi^-\gamma$ ensures that any tail extending up to the kaon mass would be negligible. $K_L \to \pi^+\pi^-$ with an accidental photon would be more likely to reconstruct at the kaon mass. But the small likelihood of an accidental hardware cluster (or external brem), combined with the moderately low $K_L \to \pi^+\pi^-$ branching ratio, prevents this from becoming a worrisome background.

6.2.4 $K_L \to \pi^\pm\mu^\mp\nu$ *(*$K_{\mu 3}$*)*

$K_{\mu 3}$ decays with accidental photons are by far the most serious background to $K_L \to \mu^+\mu^-\gamma$. This mode is the only potential background that requires an accidental photon in order to be accepted by the trigger. This means that the missing energy

[2]The large peak at 370 MeV/c^2 in Figure 6.1 is almost entirely $K_L \to \pi^+\pi^-\pi^0$.

carried off by the neutrinos will definitely be replaced to some extent. Hardware clusters appear in only about 2% of accidental events, and these accidental clusters tend to be low energy. But $K_L \rightarrow \pi^{\pm}\mu^{\mp}\nu$ has such a high branching fraction that a large exponential tail of this background extends well past the kaon mass (Figure 6.1). On the other hand, two missing neutrinos plus an accidental photon result in a nearly flat transverse momentum (P_t) distribution. This fact helps a great deal in rejecting $K_{\mu 3}$ background, but as will be shown later, additional cuts on the accidental photon cluster will be necessary to reduce the background to a manageable level.

6.2.5 $K_L \rightarrow \pi^{\pm}\mu^{\mp}\nu\gamma$ *(*$K_{\mu 3\gamma}$*)*

Radiative $K_{\mu 3}$'s can also be a significant background. Although they occur less frequently than their non-radiative counterparts, they do not require an accidental cluster to satisfy the trigger. These events will tend to reconstruct below the K_L mass since no accidental compensates for the missing energy carried off by the neutrino(s). However, phase space allows the decay to reconstruct very near the kaon mass if the neutrinos take away little energy. This proximity to the kaon mass can easily allow an event to slide within the signal mass window as a result of resolution smearing effects.

6.2.6 $K_L \rightarrow \pi^0\pi^{\pm}\mu^{\mp}\nu$ *(*$K_{\mu 4}$*)*

$K_L \rightarrow \pi^0\pi^{\pm}\mu^{\mp}\nu$ includes all the necessary final state particles to satisfy the 2MULD trigger, but there is no significant background from these decays. $K_{\mu 4}$ has a low branching ratio combined with no accidental compensation for significant missing energy. Therefore only a very small number of these events will reconstruct, and those that do will be characterized by a very low $\mu^+\mu^-\gamma$ mass.

6.3 Signal and Normalization Analysis Cuts

As was the case in the reconstruction, most of the cuts in the analysis are identical or very similar for the signal and normalization modes. Table 6.2 lists all of the analysis cut efficiencies in data and Monte Carlo for the signal, normalization, and $K_{\mu 3}$ background samples. The cuts are arranged into three groups. The first is a set of ID cuts used to identify certain topologies characteristic of a $\mu^+\mu^-\gamma$ (or $\pi^+\pi^-\pi^0$) final state. A series of data quality cuts are then applied in order to ensure that the physics quantities derived from detector information are accurately measured. Finally, several cuts are made for the purpose of rejecting various backgrounds.

The guiding philosophy has been to make the signal and normalization mode analyses as similar as possible. This is born from an effort to minimize systematic errors by keeping acceptance issues the same between the two modes. For example, making a tight cut on track offsets at the magnet rejects much of the background in the 2MULD data. However, this cut does nothing to improve the 2TRK sample although it throws away almost 10% of that data. Imposing this cut on the normalization sample does reduce possible systematics, though. Table 6.2 shows that this cut has almost the same efficiency in $K_L \to \mu^+\mu^-\gamma$ and $K_L \to \pi^+\pi^-\pi^0$ MC, as well as in 2TRK data. Any bias measuring track offsets in the signal mode will be mirrored in the normalization, and will cancel in the ratio of acceptances. This approach is feasible only because the normalization sample is so large that the loss of acceptance for $K_L \to \pi^+\pi^-\pi^0$ by imposing otherwise "unnecessary" cuts has no effect on the statistical error.

For the reasons given above, almost all of the analysis cuts for the signal and normalization are identical. Exceptions to this are different ID cuts on track energy, the MU3 requirement for the signal, and cutting tracks near the beam hole in the normalization. Also, the photon energy and reconstructed mass cuts are different for the two modes.

Table 6.2: Analysis cut efficiencies. Signal and normalization mode cuts are identical unless otherwise indicated.

Signal Mode Analysis Cut	$K_L \to \mu^+\mu^-\gamma$ Data	$K_L \to \mu^+\mu^-\gamma$ MC	$K_L \to \pi^\pm\mu^\mp\nu$ MC	$K_L \to \pi^+\pi^-\pi^0$ Data	$K_L \to \pi^+\pi^-\pi^0$ MC	Normalization Mode Analysis Cut
Track cluster energy < 1 GeV	0.904	0.955	0.896	0.990	0.991	Track $E/p < 0.9$
Track momentum	0.765	0.825	0.830	0.710	0.735	Track momentum
Non-adjacent MU3 hits	0.934	0.975	0.983	0.795	0.801	No tracks near beam holes
Kaon momentum	0.999	1.000	0.989	0.998	1.000	Kaon Momentum
Vertex Z position	0.897	0.931	0.914	0.865	0.933	Vertex Z Position
Vertex extrapolates within CsI beam hole	0.977	0.999	0.972	0.995	0.998	Vertex extrapolates within CsI beam hole
Track - track cluster separation at CsI	0.968	0.996	0.947	0.961	0.954	Track - track cluster separation at CsI
Track cluster - photon cluster separation	0.853	0.914	0.852	0.676	0.677	Track cluster - photon cluster separation
Photon cluster energy > 8 GeV	0.441	0.859	0.516	0.902	0.901	Photon Cluster Energy > 3 GeV
Shape χ^2	0.766	0.988	0.341	0.976	0.977	Shape χ^2
Vertex χ^2	0.732	0.964	0.623	0.957	0.954	Vertex χ^2
Track offsets at magnet	0.569	0.908	0.766	0.915	0.900	Track offsets at magnet
Ring Counter maximum energy	0.849	0.974	0.927	0.968	0.967	Ring Counter maximum energy
Collar Anti maximum energy	0.981	0.998	0.972	0.996	0.998	Collar Anti maximum energy
Transverse momentum	0.095	0.904	0.004	0.932	0.934	Transverse momentum
490 MeV/c^2 $< m_{\mu\mu\gamma} <$ 506 MeV/c^2	0.362	0.981	0.070	0.998	0.999	486 MeV/c^2 $< m_{\pi\pi\pi} <$ 510 MeV/c^2

6.3.1 ID Cuts

The first analysis cut designed to identify the $\mu^+\mu^-\gamma$ final state is a requirement that the energy of CsI clusters matched to charged tracks be less than 1 GeV. Muons typically deposit only ionization energy in the calorimeter. This produces an energy profile consisting of a MIP (minimum-ionizing particle) peak at about 400 MeV, with a Landau tail extending to higher energies. On the other hand, about 75% of charged pions produce a hadron shower in the CsI. These showers create clusters containing much more energy than a MIP would leave. The track cluster energy distributions in the signal mode data and Monte Carlo are shown in Figure 6.2. The data exhibits both the MIP and $\pi^\pm$ shower features. The crunch cut of 2 GeV is tightened to 1 GeV in the analysis thanks to the availability of improved CsI calibrations. This tightened cut removes another 10% of background, primarily that resulting from pion punch-throughs.

The normalization also employs a cut related to track cluster energy to identify $\pi^+\pi^-\pi^0$ events. A cut on track E/p is used to make sure that the charged tracks are from pions and not electrons. Pion showers are rarely contained completely within the CsI since the crystals comprise only 1.4 interaction lengths of material. As such, the energy deposited in $\pi^\pm$ clusters is usually significantly less than the track momentum measured by the spectrometer, as can be seen in Figure 6.3. Conversely, electromagnetic showers are almost always contained within the crystals which are 27 radiation lengths. Therefore, $e^\pm$ tracks typically have $E/p = 1$. In the analysis, the crunch cut of $E/p < 0.95$ is reduced to $E/p < 0.9$.

Between the front face of the CsI and the the MU3 planes sit 31 interaction lengths of CsI, lead and steel. Low energy muons (along with almost every other particle) are stopped in these filters. The momentum threshold for a muon to pass through the filters is about 7-8 GeV/c. Above 10 GeV/c, absolute muon efficiencies are greater than 98% and very well simulated (see Figure 5.7). To ensure that the charged tracks can pass the filters, and therefore be the source of hits in MU3, only tracks with at least 10 GeV/c momentum are accepted. Figure 6.4 shows that a significant amount of signal is lost due to this cut since the muon efficiency is still over 90% at 8 GeV/c. This cut is also applied to the normalization. Of course, it

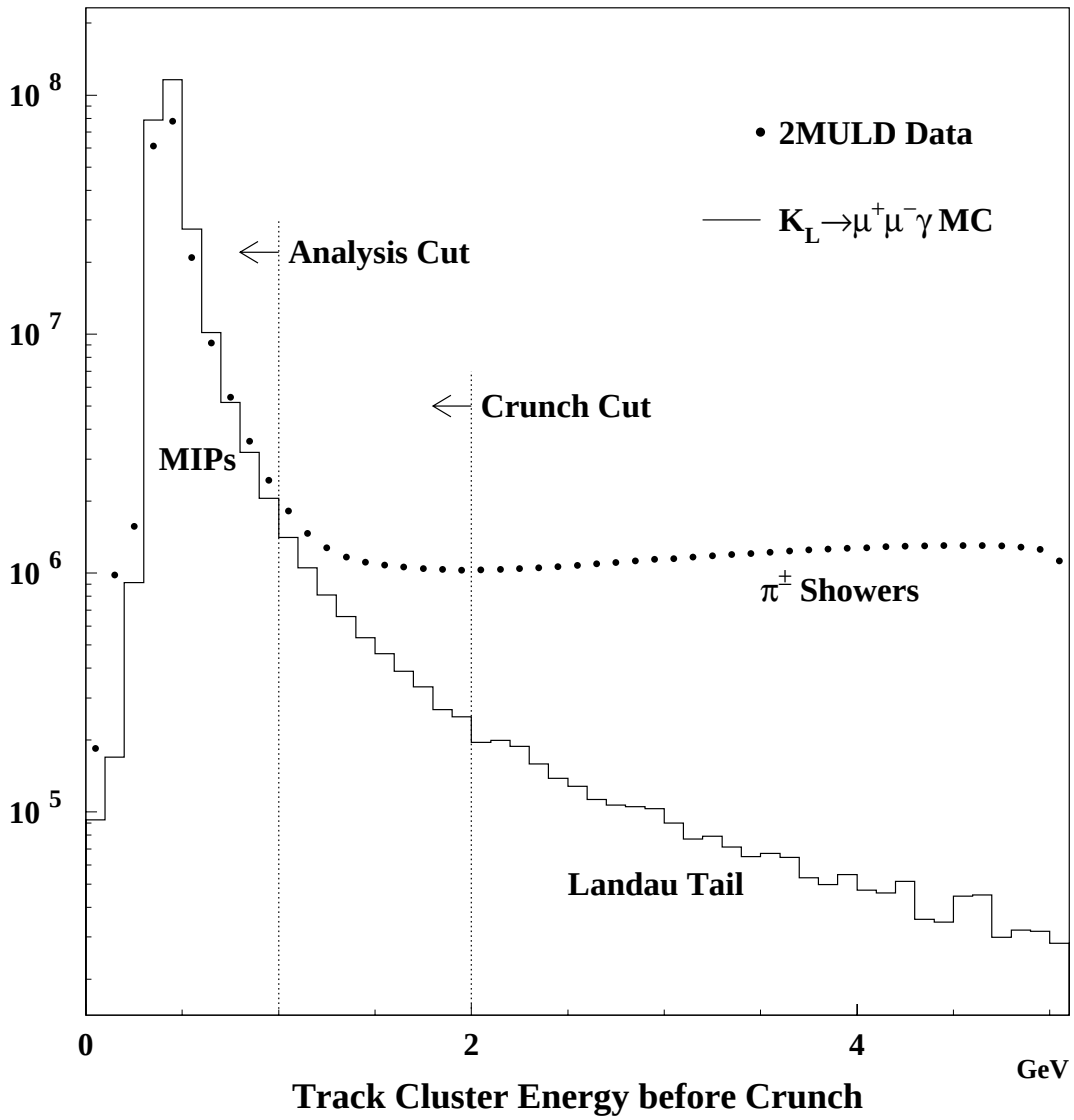

Figure 6.2: Energy of clusters matched to tracks in 2MULD data and $K_L \to \mu^+\mu^-\gamma$ Monte Carlo. The large high-energy tail in data is from pion punch-through showers in $K_L \to \pi^\pm\mu^\mp\nu$ and $K_L \to \pi^+\pi^-\pi^0$ backgrounds. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

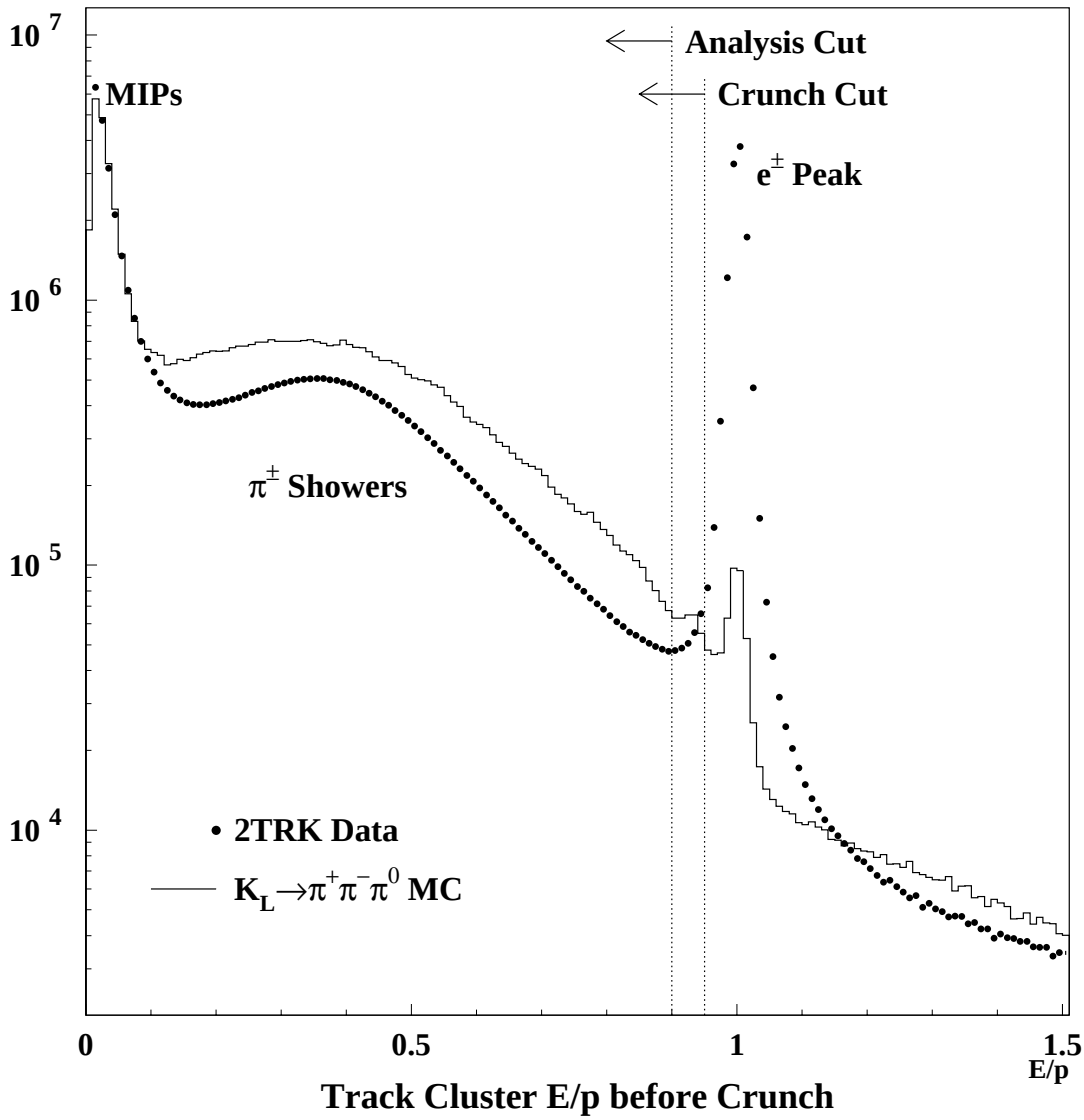

Figure 6.3: Cluster energy divided by track momentum (E/p) in 2TRK data and $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo. The large $e^\pm$ peak at E/p = 1 is from $K_L \to \pi^\pm e^\mp \nu$ electrons. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

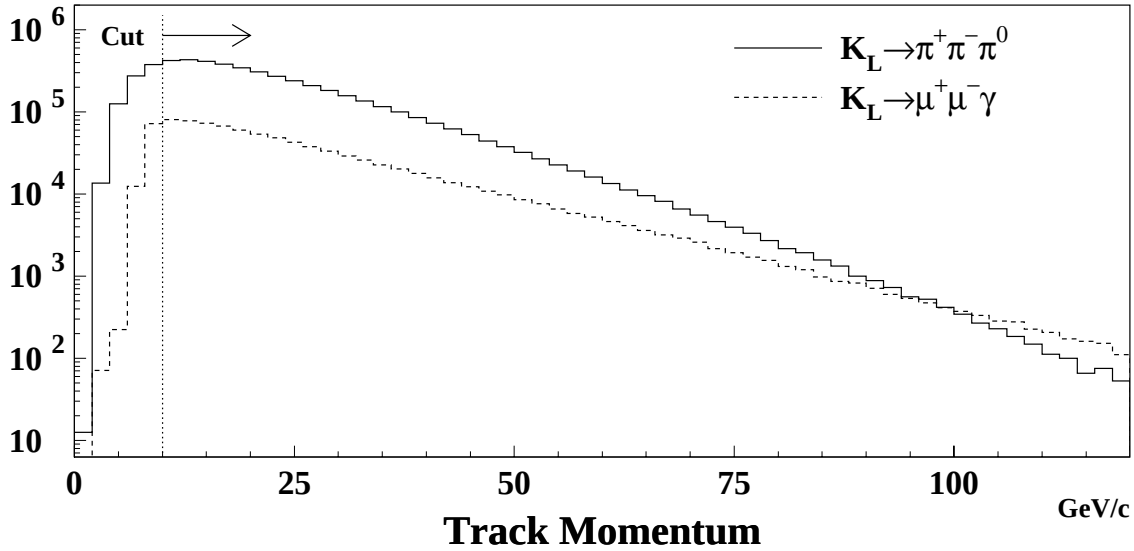

Figure 6.4: Track momentum distributions for $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \mu^+\mu^-\gamma$ Monte Carlo. The distributions shown are after reconstruction cuts have been applied. The arrow and dotted line indicate the accepted region and cut placement respectively.

has no bearing on ID or background rejection, but it serves to keep the momentum spectra for the two modes similar. This is important since the calculation of track quality quantities such as vertex χ^2 and track offsets are momentum dependent.

The 2MULD trigger requires two hits in both MU3X and MU3Y. For most events, those two hits in each plane will have been made by the two tracks already identified with greater than 10 GeV/c momentum and less than 1 GeV cluster energy. However, it is possible that one of the muons (or pions) did not reach MU3, and the second hit was caused by a δ-ray produced by the other muon. δ-rays are high-energy knock-on electrons from atomic collisions that can produce secondary ionization. If they are produced at the very back of Muon Filter 3, they can escape the steel and cause hits in MU3. These δ-rays are not simulated in the Monte Carlo, but they can be rejected from the data using a careful non-adjacent hit requirement. The presence of δ-ray events can be clearly seen in the top panel of Figure 6.5, which plots the counter separation of hits in MU3Y. In that histogram, a spike is seen at a

counter separation of 1 when there are no cuts on the MU3 hits. This is the product of δ-rays produced at the back of the filter steel that do not have space to separate very far from the parent muon before striking MU3.

Events are rejected in the analysis if the counter separation of hits is equal to 1 *or* 20 in each plane *and* there are exactly 2 hits in each plane. Separations of 20 are included because the counter number difference of counters positioned end-to-end in opposite halves of a plane is 20. Note that the event must have adjacent hits in both planes to be rejected: non-adjacent hits in just one plane are all that is needed to discount the possibility of a δ-ray event. The requirement of exactly 2 hits is made because any third hit in a plane would necessarily mean that there are at least two non-adjacent hits. The effect of this cut is presented in the center panel of Figure 6.5. There one can see the reduction of data at separations of 1 and 20.[3]

Even with the non-adjacency requirement, there is still a significant excess of events at a separation of 1 counter. This is due to the effect of extra hits in MU3. The Monte Carlo simulates fewer extra hits than are seen in data. The lack of δ-rays and the fact that the accidental events reflect somewhat less activity in MU3 than is seen in data contribute to this situation. Cutting all events with extra MU3 hits improves the data/Monte Carlo agreement at 1 counter separation. However, this is an ill-advised cut to make. Because of the disparity in extra hits between data and Monte Carlo, such a cut would bias the acceptance measurement.[4] On the other hand, it is perfectly acceptable to cut non-adjacent hit events with exactly 2 hits. In that case, extra hit events are accepted, not rejected. The two types of events that are rejected are:

- The two hits are from one muon and a δ-ray, and the other muon does not reach MU3. This case is rejected in Monte Carlo since only one MU3 hit would be recorded.

[3]The Monte Carlo is also reduced, even though δ-rays are not simulated. This is because some level of adjacency occurs simply from geometry and kinematics.

[4]Also, cutting extra hit events does nothing to address the δ-ray problem.

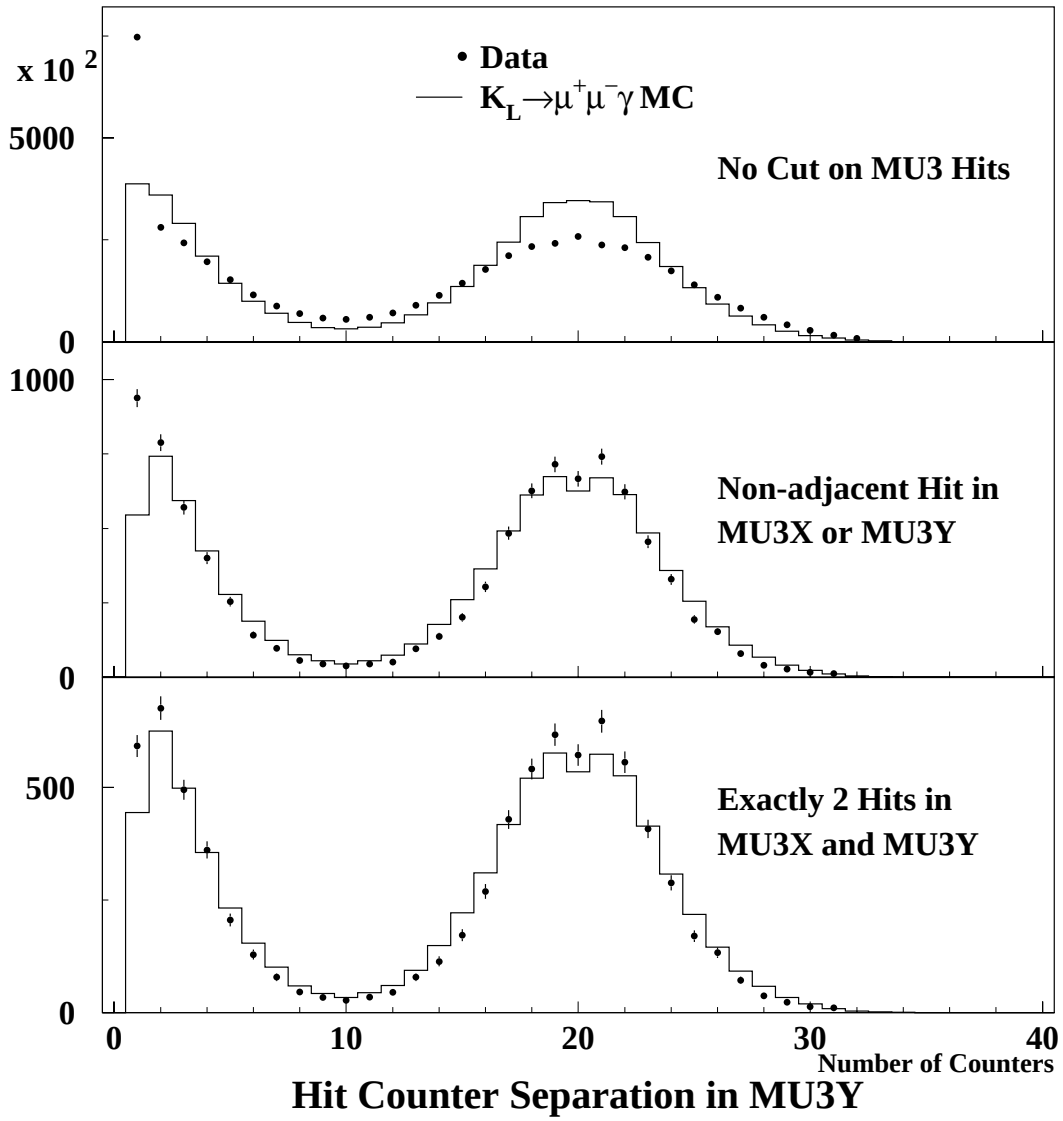

Figure 6.5: Counter separation of hits in MU3Y. The upper plot has no cut on hits in MU3. The center plot requires a non-adjacent hit in either MU3X or MU3Y. The non-adjacency requirement is also made in the lower plot, which includes only events with exactly 2 hits in MU3X and MU3Y. The second peak at 20 counters is due to the fact that the counters in the right half of MU3Y are numbered 1-20, and the left half 21-40. Counters at the same Y position in opposite halves are 20 'counters' apart, i.e. counters 1 and 21 are end-to-end.

- The two hits are from two muons that happen to be adjacent due to geometry and kinematics. This case is modeled well in the Monte Carlo.

The last ID cut is made only in the normalization mode. For $K_L \to \pi^+\pi^-\pi^0$, events are only accepted if the charged tracks extrapolate at least 2.5 cm away from the CsI beam holes. This means that tracks can not point to a CsI crystal next to a beam hole. The cut is made to ensure that $\pi^\pm$ clusters are well-measured. Strictly speaking, this is a data quality cut, but it is grouped here simply to keep the defined set of quality cuts identical for the signal and normalization.

6.3.2 Data Quality Cuts

As was described in Section 5.1.1, Monte Carlo kaons were generated with the momentum and decay vertex ranges 20 GeV/c $< P_K <$ 220 GeV/c, 90 m $< Z <$ 160 m. To make valid acceptance calculations, analysis cuts on these quantities must be at least as restrictive. The cut on the kaon momentum is the same as the Monte Carlo generation range. According to Figure 6.6, no data reconstructs with a kaon momentum below 20 GeV/c, and virtually no events have $P_K >$ 220 GeV/c after all other analysis cuts have been applied. The cut efficiencies listed in Table 6.2 verify that this cut has almost no impact on the analysis. Nor is a tighter cut warranted since the data/ Monte Carlo agreement is so good in both modes.

There is a definite need for a more restrictive cut on the vertex Z position. Two troublesome features are apparent in the Z distribution of 2TRK data shown in Figure 6.7. The most dramatic is the large spike at 159 m. This is the location of the vacuum window, and the data there represent either photon conversions or some other interaction (largely from accidentals) in the vacuum window. The second feature is the smaller spike at 94 m - the position of the final sweeper magnet. Decays occurring just downstream of the magnet are affected by the magnet's fringe field. This alteration of charged particle trajectories can cause the vertex to be reconstructed too far upstream. For these reasons, the Z cut is set to be 97 m $< Z <$ 157 m.

The Z distribution in the 2MULD data does not exhibit the same features.

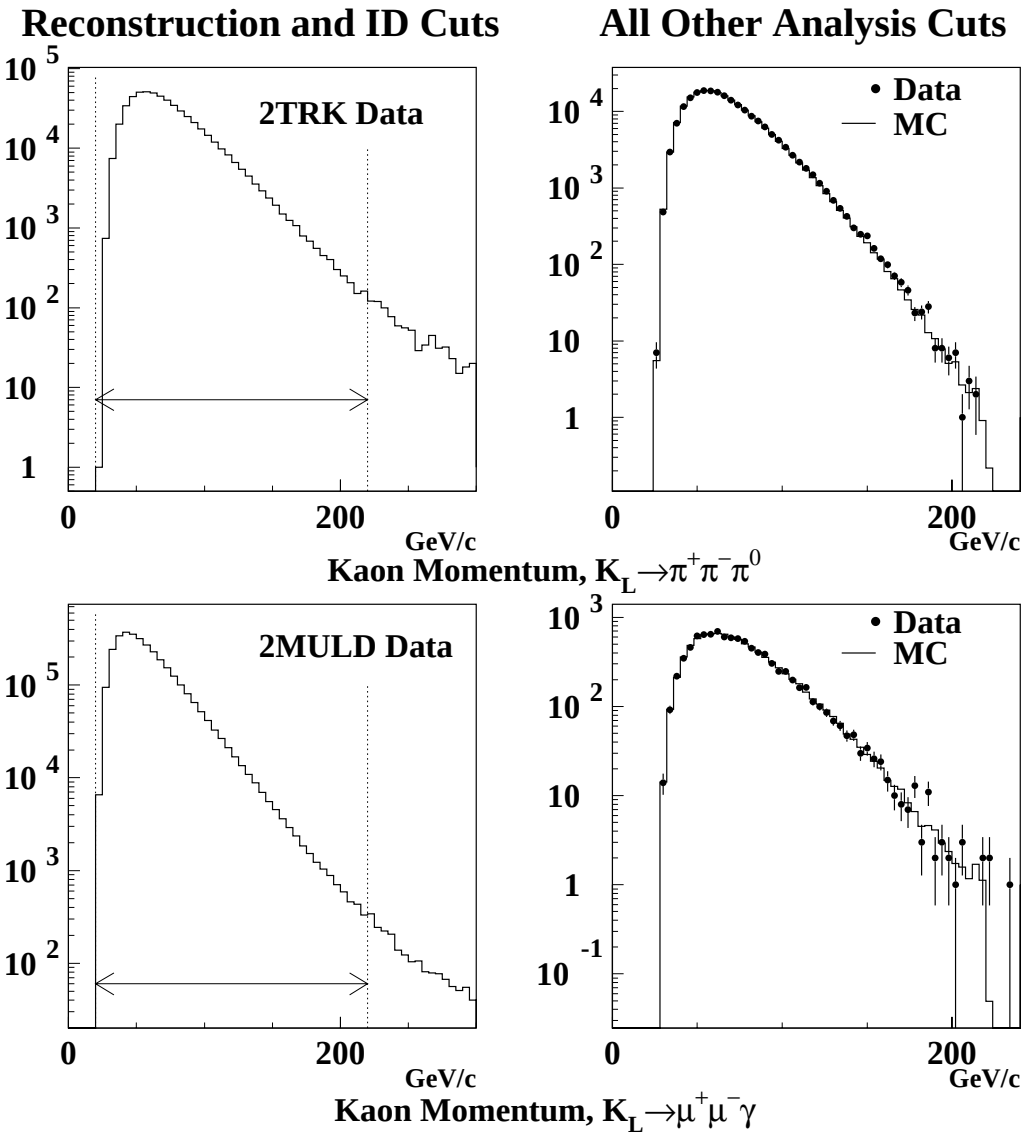

Figure 6.6: Reconstructed kaon momentum for $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \mu^+\mu^-\gamma$. The plots on the left are of the data after reconstruction and ID cuts. The right-hand plots are the data and Monte Carlo after all cuts other than the kaon momentum cut have been applied. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

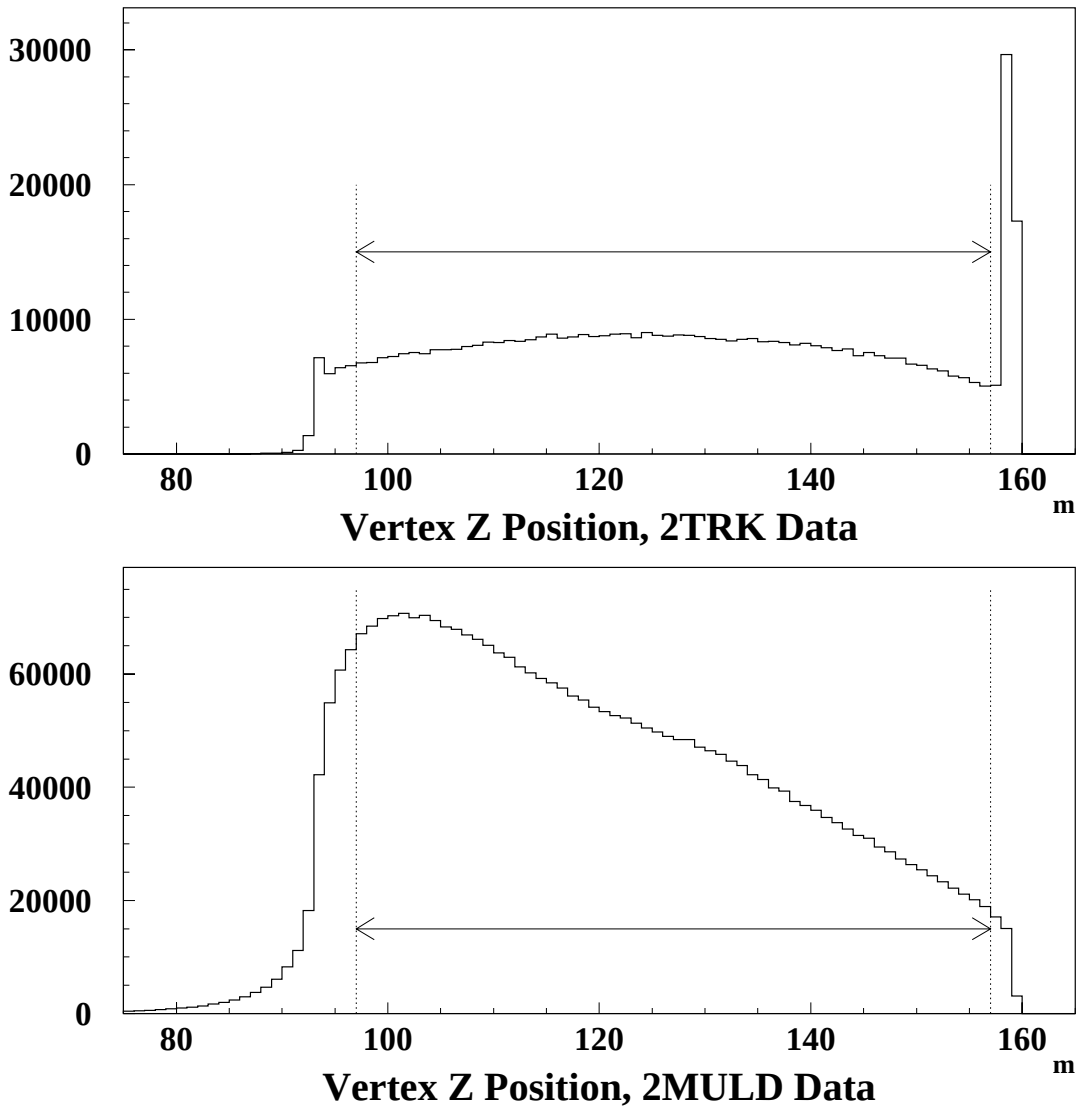

Figure 6.7: Reconstructed vertex Z position for 2TRK (top) and 2MULD (bottom) data. The plots show the data after reconstruction and ID cuts have been applied. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

The MU3 requirement alone prevents the vacuum window interactions from passing the trigger. The final sweeper magnet effect is still present, but it is masked by background events. Pion decays upstream of the analysis magnet bend plane result in measurement of a muon track momentum that is lower than the original pion momentum. This pushes the reconstructed vertex position upstream, and accounts for the Z distribution in 2MULD data being shifted to lower values.

A kaon that scatters off a collimator edge can result in a poorly reconstructed decay. In particular, the flight direction of the K_L can no longer be determined by simply tracing a line from the target to the reconstructed decay vertex. P_t calculations suffer because the transverse momentum is defined relative to that line. If a kaon passes through all of the collimators cleanly, then its flight direction is determined by the target-vertex projection, which should extrapolate to a point within the CsI beam holes. To ensure that this is the case, a cut is made on the X and Y positions of the vertex projected to the CsI ($X_{vtx}(@CsI)$, $Y_{vtx}(@CsI)$). Both positions must fall within the beam holes, i.e. 7.5 cm $< |X_{vtx}(@CsI)| <$ 22.5 cm and $|Y_{vtx}(@CsI)| <$ 7.5 cm. The distributions of the vertex projection to the CsI are displayed in Figure 6.8.

One measure of the self-consistency of the tracking and clustering is to determine how closely a track's extrapolated position at the CsI agrees with the measured position of the cluster matched to that track. Appreciable differences could indicate a poorly measured track or cluster. For $K_L \to \mu^+\mu^-\gamma$ Monte Carlo, almost all track-cluster separations are within 3.5 cm. On the other hand, the $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo exhibits a tail that extends well past this cut, and is shown in Figure 6.9. The cut efficiencies from Table 6.2 indicate that this effect accurately reflects the data.

It is also worth noting that the 3.5 cm cut efficiency for the normalization is nearly the same for $K_{\mu 3}$ MC as well as the 2MULD data, which consists predominantly of $K_{\mu 3}$ and $K_L \to \pi^+\pi^-\pi^0$ events. There are two reasons for the $4-5\%$ inefficiency. Pion decays between DC3 and the CsI will result in downstream track segments that do not precisely reflect the direction of the daughter muon that hits the CsI. Also, pion shower clusters in the CsI are very irregular in breadth and shape. The energy deposition tends to appear splattered. This irregularity makes determi-

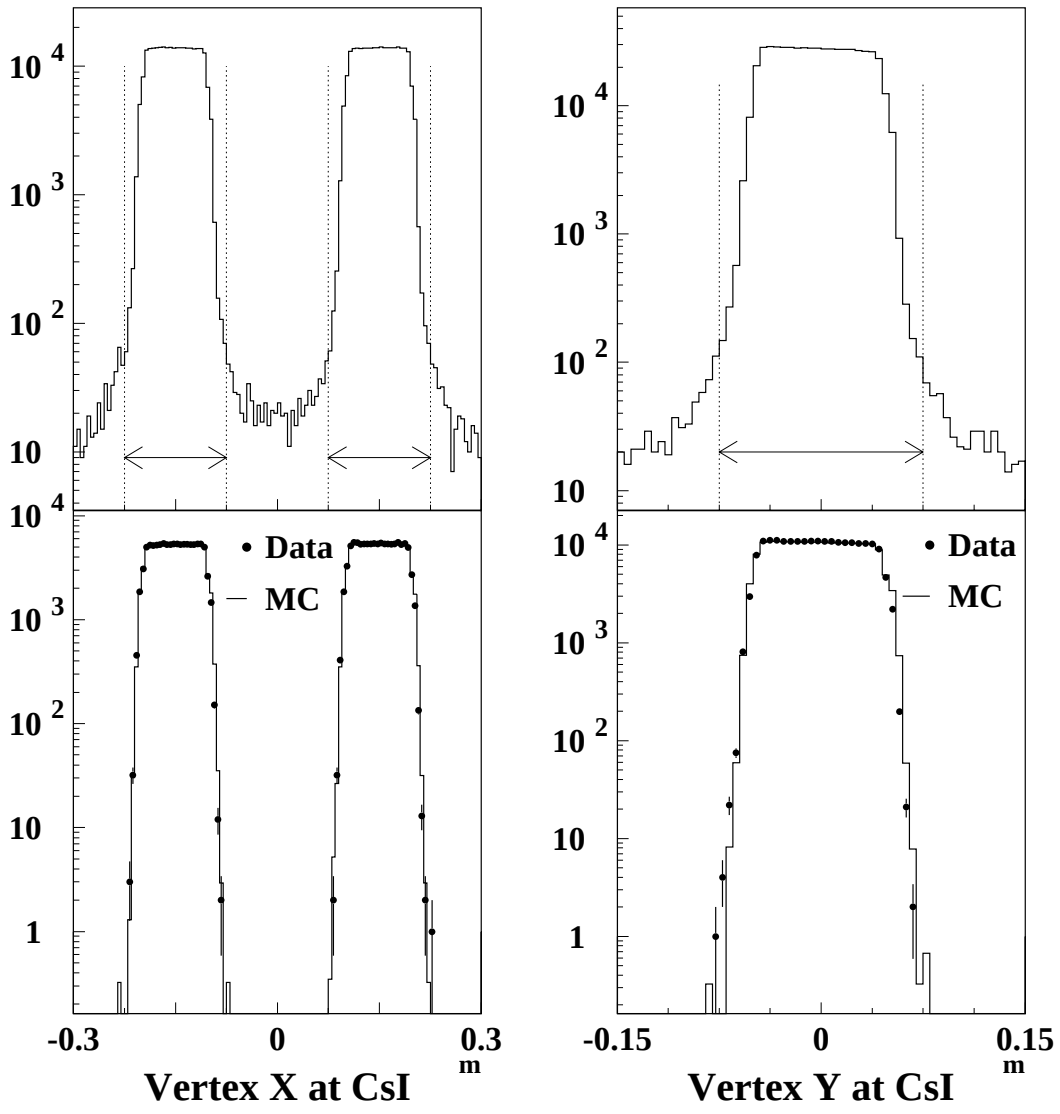

Figure 6.8: Vertex X and Y positions extrapolated to the CsI from the target. The upper plots are the 2TRK data after reconstruction and ID cuts. The lower plots are the data and $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo after analysis cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

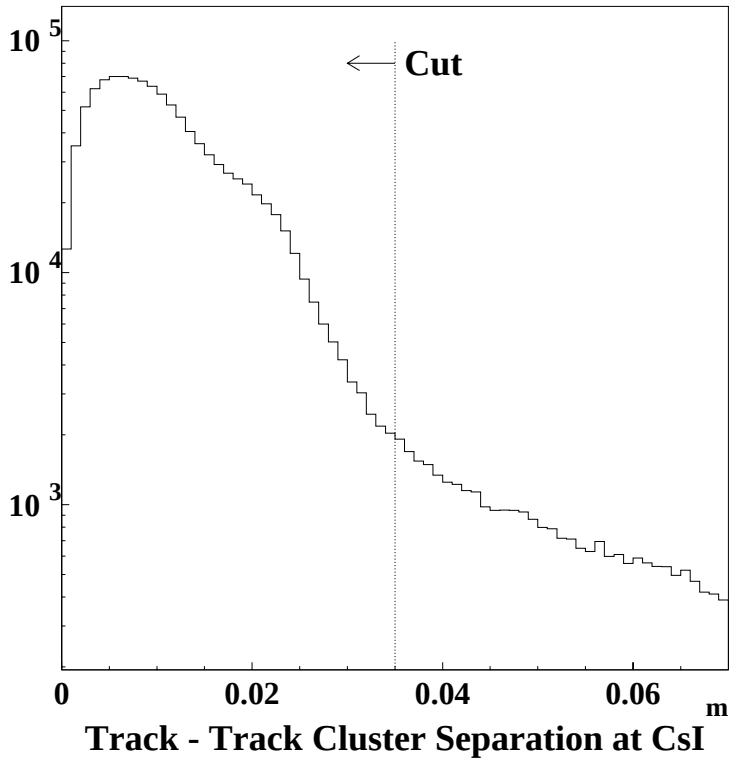

Figure 6.9: Separation of track position and track cluster position at the CsI upstream face. 2TRK data is shown after reconstruction and ID cuts. The arrow and dotted line indicate the accepted region and cut placement respectively. The bump around 2 cm is an effect of the bigger size of the large crystals.

nation of a cluster position (i.e., center) difficult, and not necessarily indicative of the incident pion's hit position.

A final data quality cut is made on the separation between track clusters and photon clusters. All photon clusters must be at least 20 cm away from any cluster matched to a track. This cut is made to guarantee that measurements of photon cluster energies are not corrupted by spurious energy associated with pion showers. This type of corruption can be a problem because the irregular energy distribution of pion showers is not simulated extremely well by either the Monte Carlo or the clustering algorithms.

Of course, this cut on track and photon cluster separations is almost exclusively relevant to the normalization mode. However, it is still applied to the signal mode

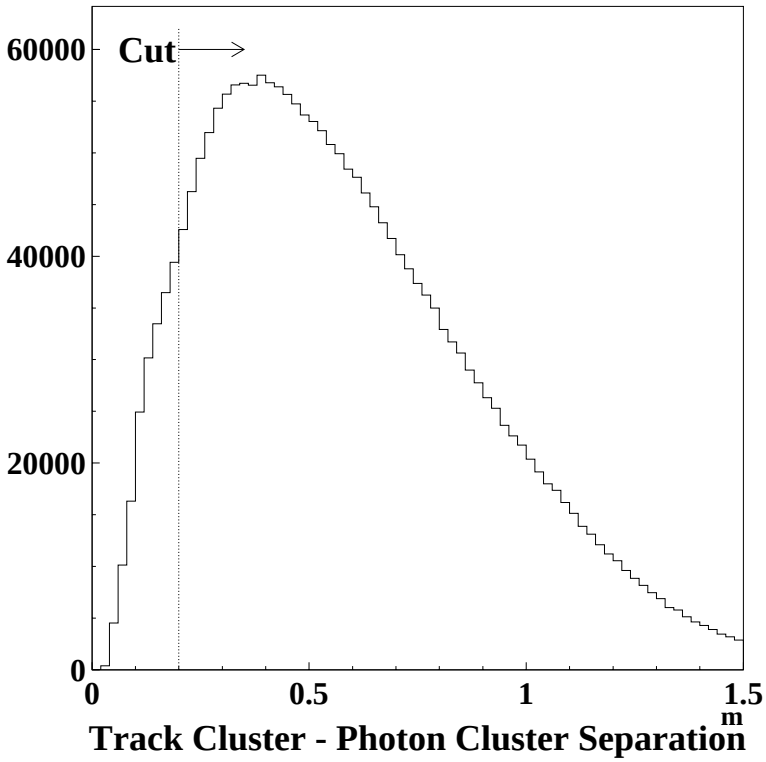

Figure 6.10: Separation of track cluster and photon cluster positions. 2TRK data is shown after reconstruction and ID cuts. The arrow and dotted line indicate the accepted region and cut placement respectively.

in spite of the loss of 9% of the data (at least it accomplishes a modest amount of background rejection there). The 2TRK data distribution of the track cluster-photon cluster separation is shown in Figure 6.10. Note that this is a four entry plot - one entry for each track-photon combination. Hence the apparent mismatch between the size of the accepted region in Figure 6.10 and the 68% cut efficiency given in Table 6.2

6.3.3 Background Rejection Cuts

Over 70% of the $K_L \to \mu^+\mu^-\gamma$ background which dominates Figure 6.1 still remains after imposing ID and data quality cuts. The prospect of rejecting that much background at first seems daunting. This is especially true for the case of $K_L \to$

$\pi^{\pm}\mu^{\mp}\nu$ with an accidental photon, where the final state observables are identical to those of the signal. Fortunately, rejection can be found almost anywhere you look.

The photon cluster is a good place to start looking for ways to eliminate background. The upper panel of Figure 6.11 shows how effectively a photon energy cut removes background from the data. The plot shows that the total background has a much softer photon energy spectrum than the signal. Photons from radiative decays are generally softer than those from direct emission. $K_L \to \pi^+\pi^-\pi^0$, $K_L \to \pi^0\pi^{\pm}\mu^{\mp}\nu$, and direct $K_L \to \pi^+\pi^-\gamma$ all have less phase space available to photons than is afforded by $K_L \to \mu^+\mu^-\gamma$. Most importantly, accidental photons (which are required for $K_{\mu 3}$ to become background) almost always have low energy. Requiring that the photon energy be greater than 8 GeV rejects almost 50% of the $K_{\mu 3}$ events in Monte Carlo.

The photon energy cut is only 3 GeV in the normalization analysis. The lower plots in Figure 6.11 make clear why this is the case. Raising the cut to 8 GeV would throw away a majority of the data.[5] The cut is not be placed any lower in order to avoid the effects of HCC threshold simulation in the Monte Carlo. This manifests as a data/Monte Carlo discrepancy at low energies.

Another powerful tool that can be used to eliminate $K_{\mu 3}$ decays is measurement of the photon cluster shower shape. Clean photon clusters generally exhibit a well-defined transverse distribution of deposited energy. The shape of a given cluster can be compared to that expected for a photon. A shape χ^2 is then calculated to measure how consistent the assumption is that the cluster was produced by a single photon. Accidental clusters often are not the result of an isolated photon. Such clusters are very likely to be produced by some other random particle not tracked by the spectrometer, or from the fusion of multiple photon clusters that were not distinguished by the clustering algorithms. The shape χ^2 distribution for these accidental clusters is relatively flat. This can be seen in the 2MULD data and background Monte Carlo distributions, which both exhibit a prominent tail extending out to very large χ^2 values (Figure 6.12). Two-thirds of the $K_{\mu 3}$ background is removed by cutting all events with photon shape $\chi^2 > 5$.

[5] This is the reason that the 8 GeV cut rejects even more 2MULD data than $K_{\mu 3}$ Monte Carlo.

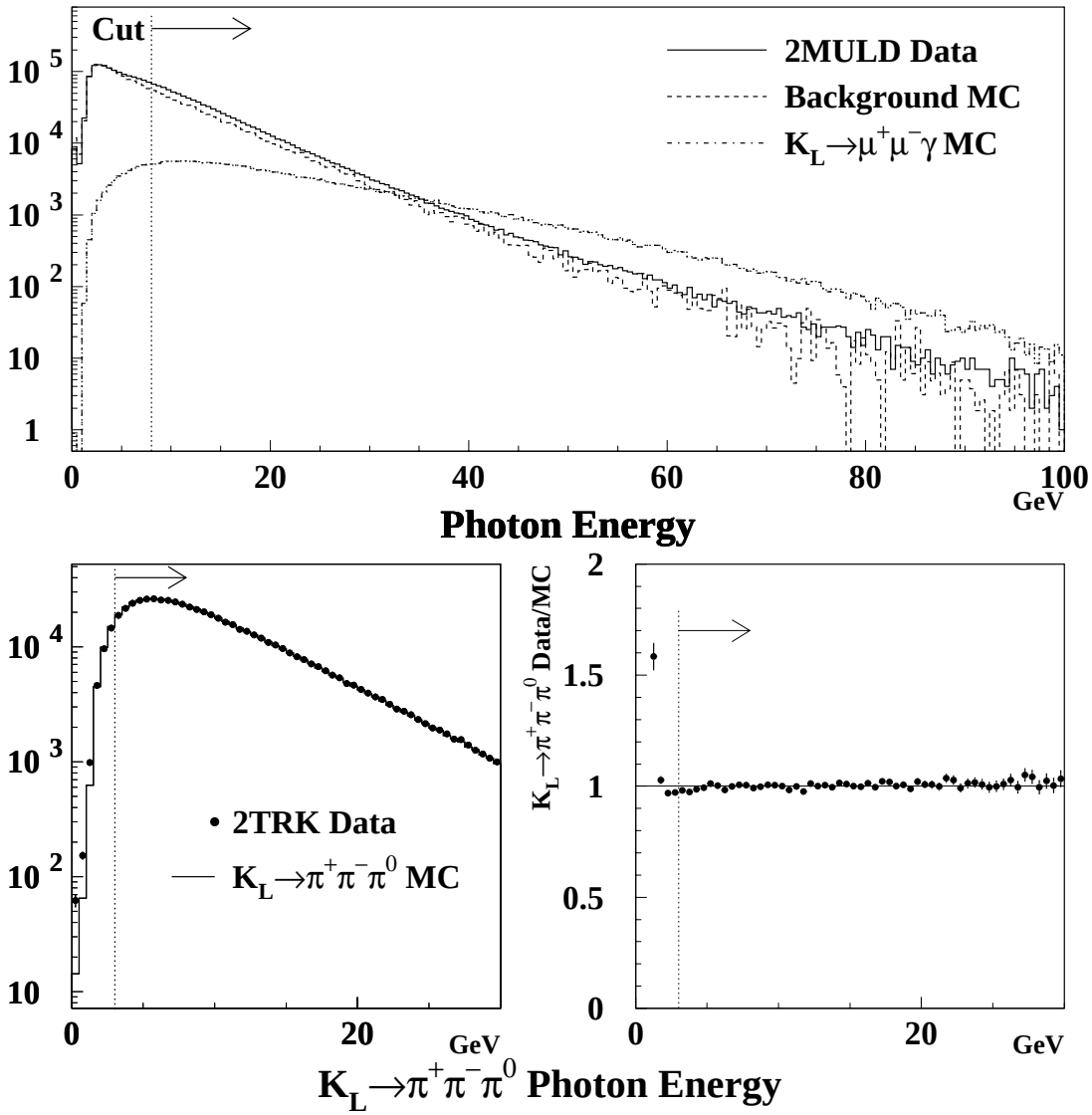

Figure 6.11: Photon energy cuts for signal and normalization. The upper plot shows the photon energy cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower plots display the cut in the normalization mode. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

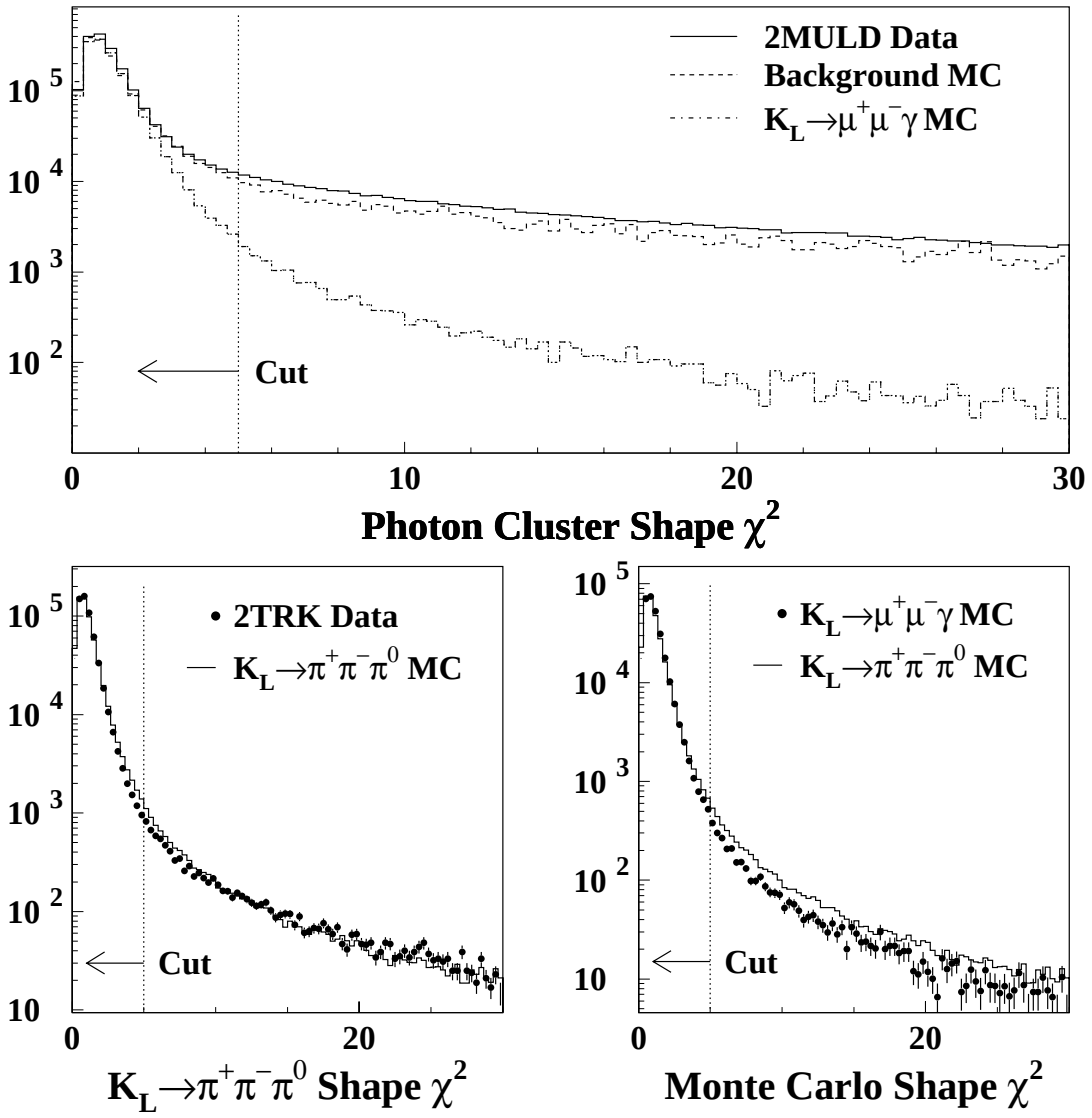

Figure 6.12: Photon cluster shape χ^2 for signal and normalization. The upper plot shows the shape χ^2 cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower left plot displays the cut in the normalization mode, and a comparison of the signal and normalization Monte Carlo is made in the lower right plot. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

Another feature of the shape χ^2 variable is that it gets worse at low photon energies. This could potentially cause a bias in the signal to normalization acceptance ratio, because the photon spectrum for $K_L \to \pi^+\pi^-\pi^0$ is softer than that for $K_L \to \mu^+\mu^-\gamma$. The lower plots in Figure 6.12 show the comparisons between $K_L \to \pi^+\pi^-\pi^0$ data and Monte Carlo, and between signal and normalization Monte Carlo. The greatest shape χ^2 difference between $K_L \to \pi^+\pi^-\pi^0$ and $K_L \to \mu^+\mu^-\gamma$ is out on the tails. This combined with the reasonably good data/Monte Carlo agreement for $K_L \to \pi^+\pi^-\pi^0$ indicates that the shape χ^2 differences should not produce a significant bias.

Track quality cuts can also be very effective in reducing background. A vertex χ^2 is formed to measure the quality of vertex fitting, and poorly reconstructed tracks can degrade the fit, leading to large values of vertex χ^2. One mechanism for forming bad tracks is using accidental chamber hits which are not associated with the actual track.[6] As one might expect, this is most likely to happen when reconstructing events that definitely contain accidental activity. This statement is validated by the fact that a third of the remaining $K_{\mu 3}$ background is rejected by requiring that the vertex $\chi^2 < 10$. A significant fraction of all other backgrounds are rejected as well, because a charged pion decay upstream of DC4 will put a kink in the charged track. The vertex χ^2 distributions in the signal and normalization modes are presented in Figure 6.13.

Cutting events with large track offsets at the magnet (offmags) also eliminates a large amount of $\pi^\pm$ decay background. Obviously, if a charged pion decays within the spectrometer, the measured upstream and downstream track segments will not be well matched at the magnet bend plane. This effect can be clearly seen in Figure 6.14, which shows large tails in X track offmags for backgrounds to $K_L \to \mu^+\mu^-\gamma$. The offmag cut is 1 mm for both X and Y tracks. This cut is relatively tight, but it eliminates nearly half of the $K_L \to \pi^+\pi^-\pi^0$ background.[7]

---

[6] Another mechanism involves the effects from high-SODs discussed in Section 4.1.1 and Section 5.3.1. This is an example of the need for good drift chamber simulation

[7] Of course, this cut is about twice as effective for rejecting two-pion background as it is for rejecting semi-leptonic modes, since in the former case there are two chances for a pion decay.

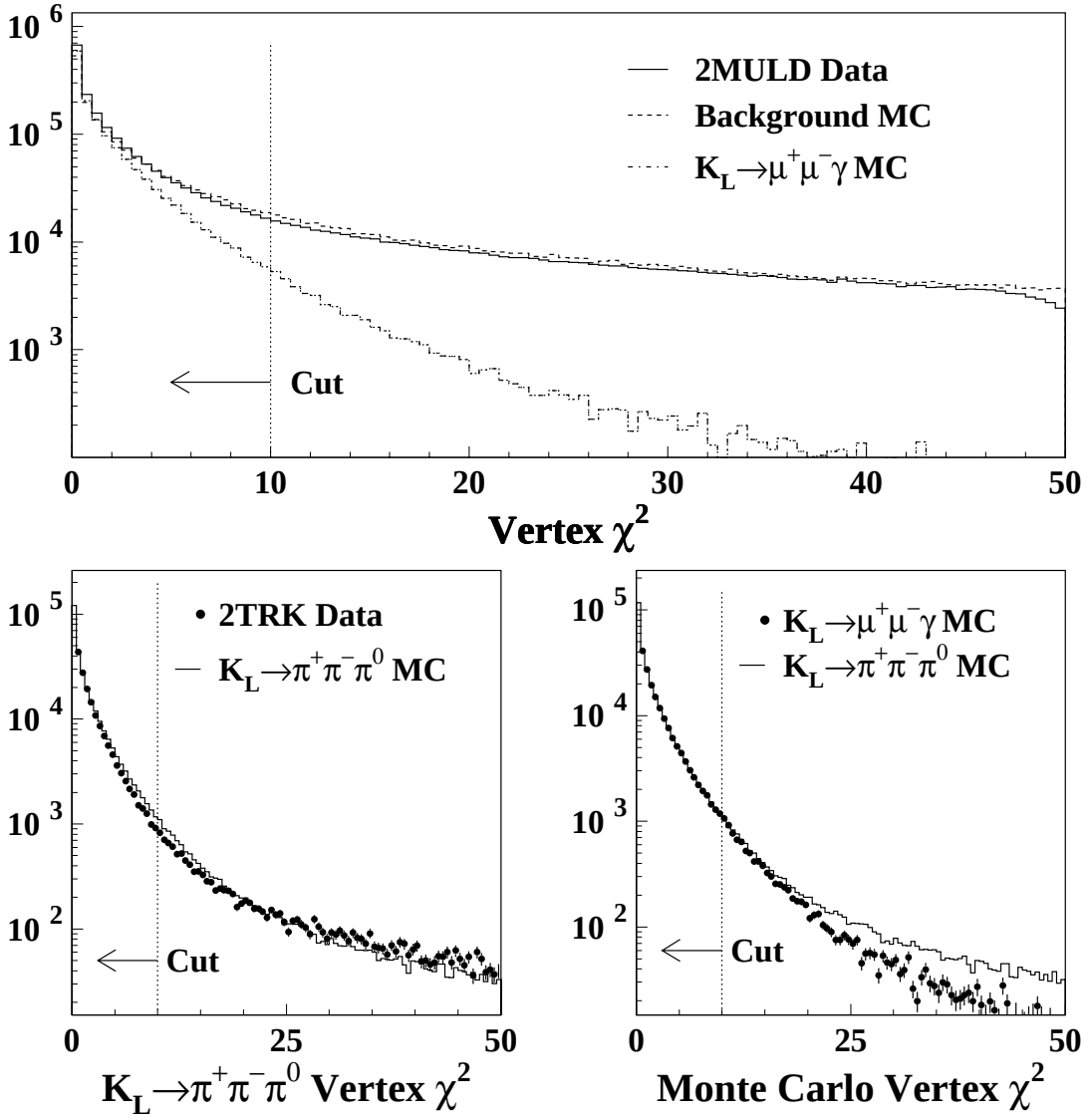

Figure 6.13: Vertex χ^2 for signal and normalization. The upper plot shows the vertex χ^2 cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower left plot displays the cut in the normalization mode, and a comparison of the signal and normalization Monte Carlo is made in the lower right plot. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

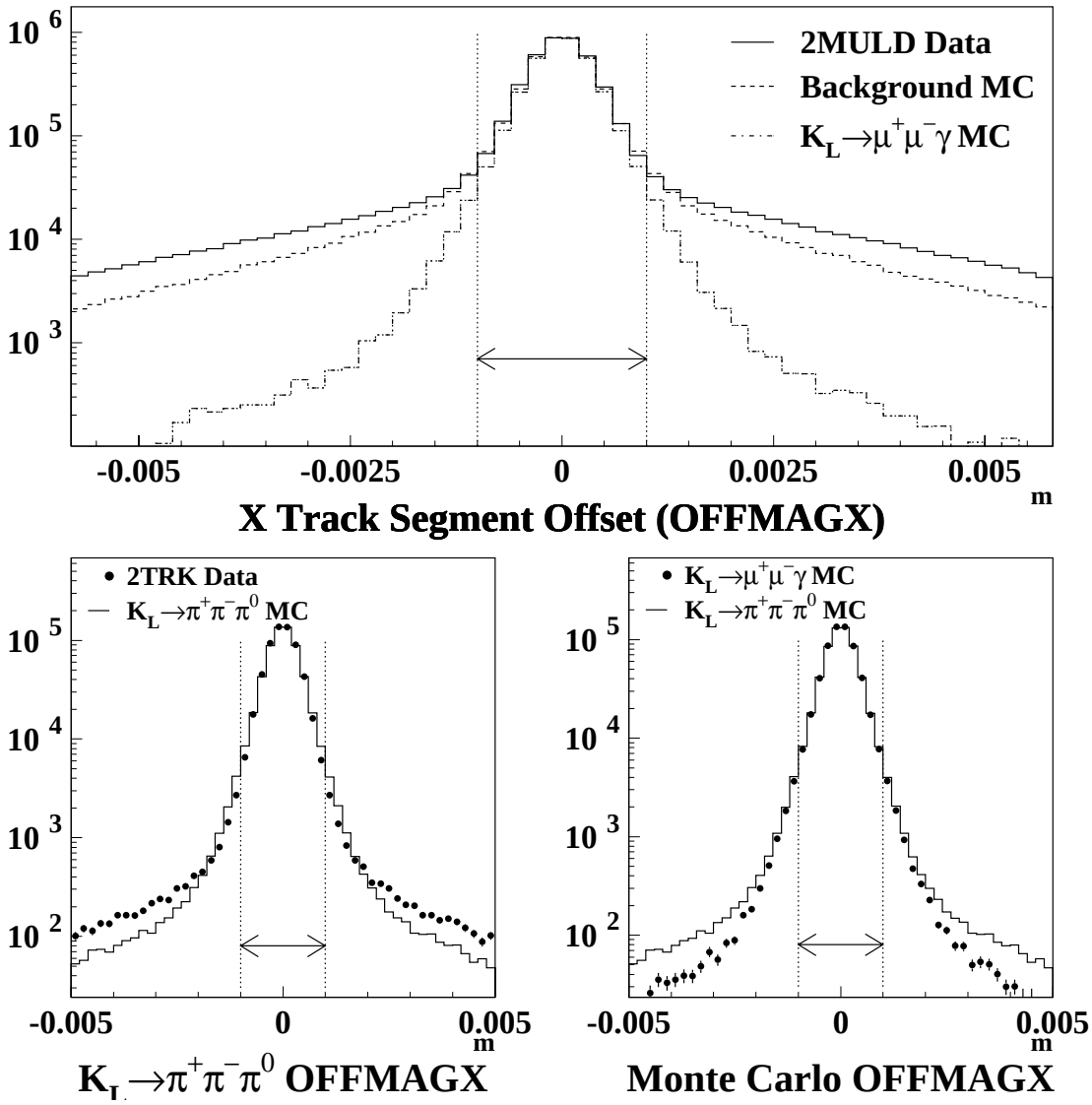

Figure 6.14: X Track segment offset (OFFMAGX) for signal and normalization. The upper plot shows the OFFMAGX cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower left plot displays the cut in the normalization mode, and a comparison of the signal and normalization Monte Carlo is made in the lower right plot. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

The $K_L \to \pi^+\pi^-\pi^0$ data/Monte Carlo agreement is not exceptionally good for either vertex χ^2 or offmag. Additionally, there are noticeable differences in the distributions of both quantities between signal and normalization Monte Carlo. These discrepancies could present a significant bias, particularly since the cuts on these quantities are relatively tight. As is the case with shape χ^2 though, the largest differences between the two modes occur on the tails. Additionally, the data/Monte Carlo disagreement seems to be similar in the signal and normalization. These two points help to allay the fear of a serious acceptance bias due to track reconstruction. A detailed examination of the systematic error associated with tracking is deferred until Section 6.6.2.

Some of the photon vetoes are useful for rejecting specific types of background events. One important characteristic of $K_L \to \pi^+\pi^-\pi^0$ background is that one of the photons from the π^0 must be lost in order for the event to pass the 1 photon cluster crunch cut. Many of these events are rejected by the photon veto requirements in the Level 1 trigger. However, at times certain photon vetoes were not in the trigger for various performance reasons. An analysis cut is made that rejects events in which the maximum energy in any Ring Counter is greater than 0.2 GeV. This cut is tighter than the 0.5 GeV trigger threshold for the Ring Counters. The tightened cut helps to eliminate a few percent of the $K_{\mu 3}$ background in which the accidental activity includes very soft photons that hit the Ring Counters. A cut on Collar Anti maximum energy was reduced from a 14 GeV trigger threshold to a 3 GeV analysis cut. The 3 GeV cut removed a relatively small number of background events with moderate Collar Anti energy deposited by either accidental activity or a π^0 photon. The distributions of maximum energy in the Ring Counters and Collar Anti are displayed in Figure 6.15 and Figure 6.16 respectively.

The final cuts on the reconstructed P_t^2 and invariant mass can be thought of as either background rejection or ID cuts. On one hand, the two cuts eliminate over 99.9% of the remaining $K_{\mu 3}$ background. On the other, this huge rejection stems from the fact that the 0.490 GeV/c^2 $< m_{\mu^+\mu^-\gamma} <$ 0.506 GeV/c^2, $P_t^2 < 0.0001$ (GeV/c)^2 signal window is perhaps the most fundamental identification for a $K_L \to \mu^+\mu^-\gamma$ event.

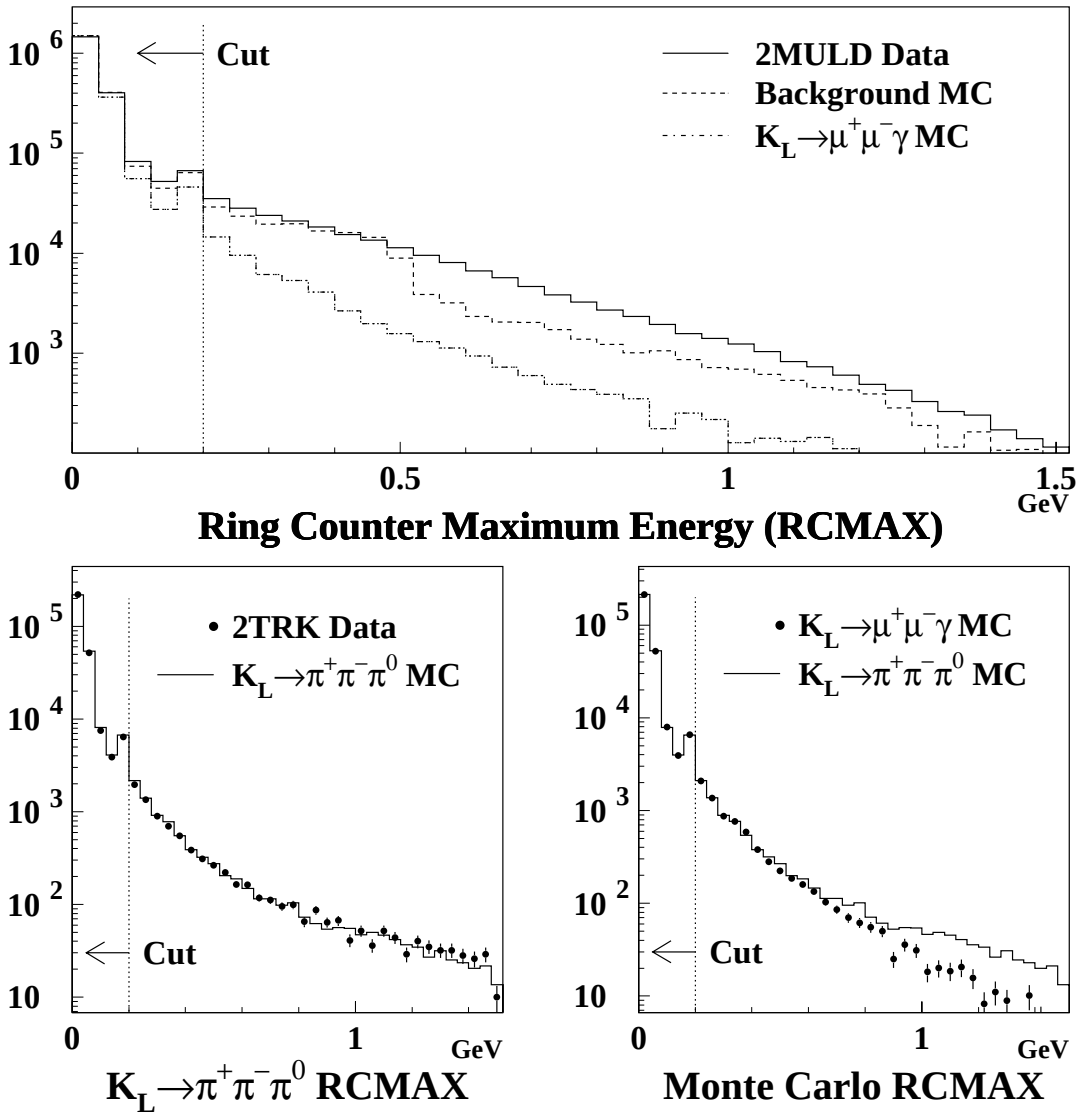

Figure 6.15: Ring Counter maximum energy (RCMAX) for signal and normalization. The upper plot shows the RCMAX cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower left plot displays the cut in the normalization mode, and a comparison of the signal and normalization Monte Carlo is made in the lower right plot. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

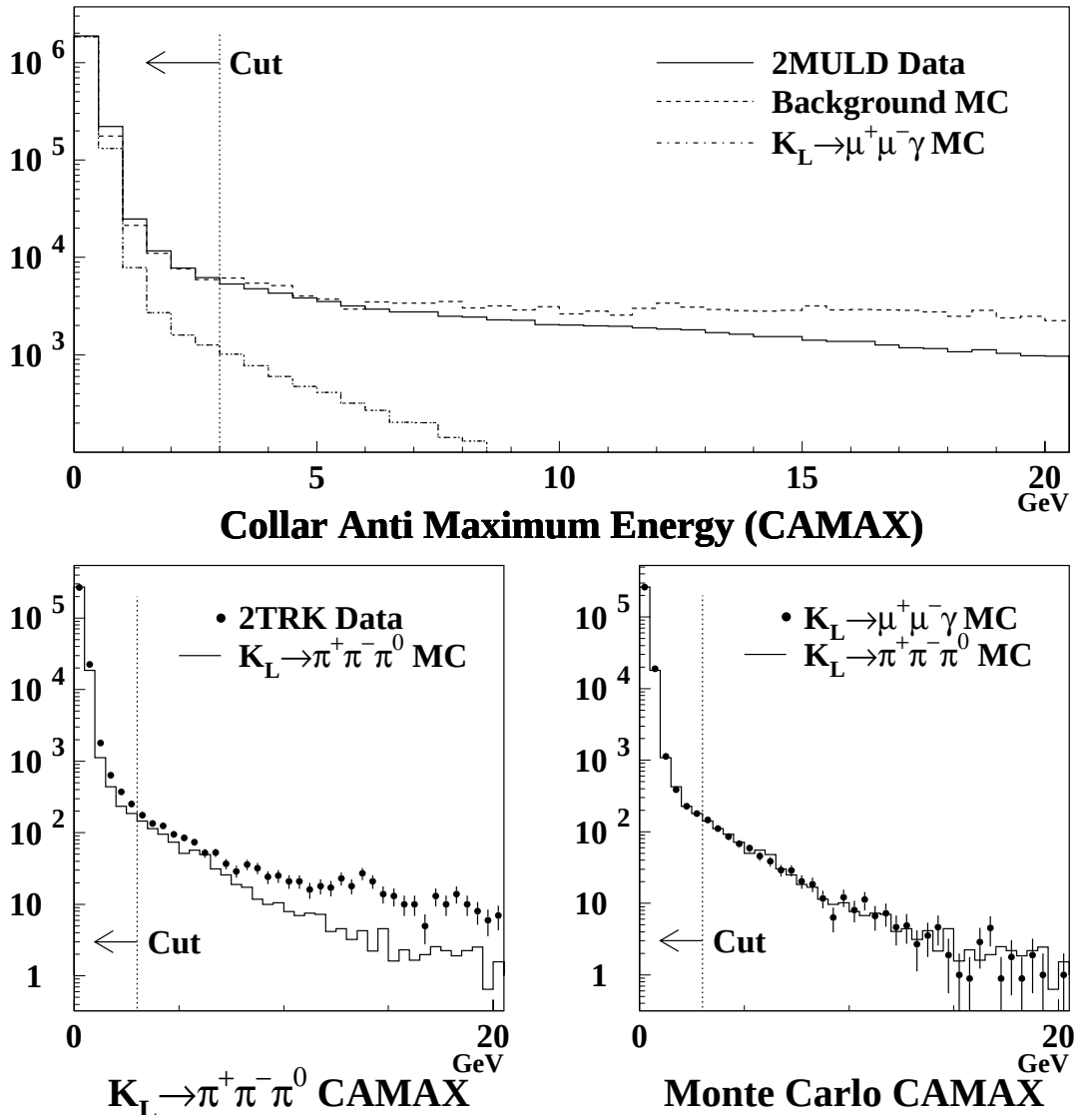

Figure 6.16: Collar Anti maximum energy (CAMAX) for signal and normalization. The upper plot shows the CAMAX cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower left plot displays the cut in the normalization mode, and a comparison of the signal and normalization Monte Carlo is made in the lower right plot. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

The most powerful of the two is the cut on P_t^2, and this is made strikingly clear in Figure 6.17. The upper panel of that figure shows that the P_t^2 distribution for background to $K_L \to \mu^+\mu^-\gamma$ is almost completely flat out to very high values. Furthermore, the small slope that does exist comes entirely from the $K_L \to \pi^+\pi^-\pi^0$ background which is not nearly as big a concern as $K_{\mu 3}$. Unfortunately, the tight cut placement lies in a region of disagreement between data and Monte Carlo, and signal and normalization P_t^2 distributions are markedly different. This is an issue that must be accounted for in systematic error studies.

A particularly illuminating view of the various background components is given in Figure 6.18. There, a plot of P_t^2 vs. $m_{\mu^+\mu^-\gamma}$ is made for each background mode. It is evident in those plots that virtually all of the background to $K_L \to \mu^+\mu^-\gamma$ comes from $K_{\mu 3}$ events.

The signal and normalization utilize different mass cuts. The mass window for $K_L \to \mu^+\mu^-\gamma$ is chosen to be $\sim 3\sigma$ (± 8 MeV) from the K_L mass. For $K_L \to \pi^+\pi^-\pi^0$, the window is extended to 5σ (± 12 MeV) because the normalization is virtually free of background. The mass resolutions in each case are taken from Gaussian fits to the Monte Carlo distributions, seen in Figure 6.19.

6.4 Final Samples

6.4.1 Normalization and Flux Measurement

The mass distributions of the 2TRK data and $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo after all other analysis cuts have been applied is shown in Figure 6.20. The data reconstructs at a slightly higher mass than the Monte Carlo (~ 0.2 MeV/c^2 higher), but otherwise the data distribution is in excellent agreement with Monte Carlo. The small shift is understood to be associated with a slight inaccuracy in the CsI calibrations, but has no significant effect on this analysis.

In principle, to calculate the flux one needs only the final number of accepted events in data and Monte Carlo, the number of generated Monte Carlo events, and the $K_L \to \pi^+\pi^-\pi^0$ branching ratio. However there is one correction that must be made to the acceptances. The $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo sample was not subjected

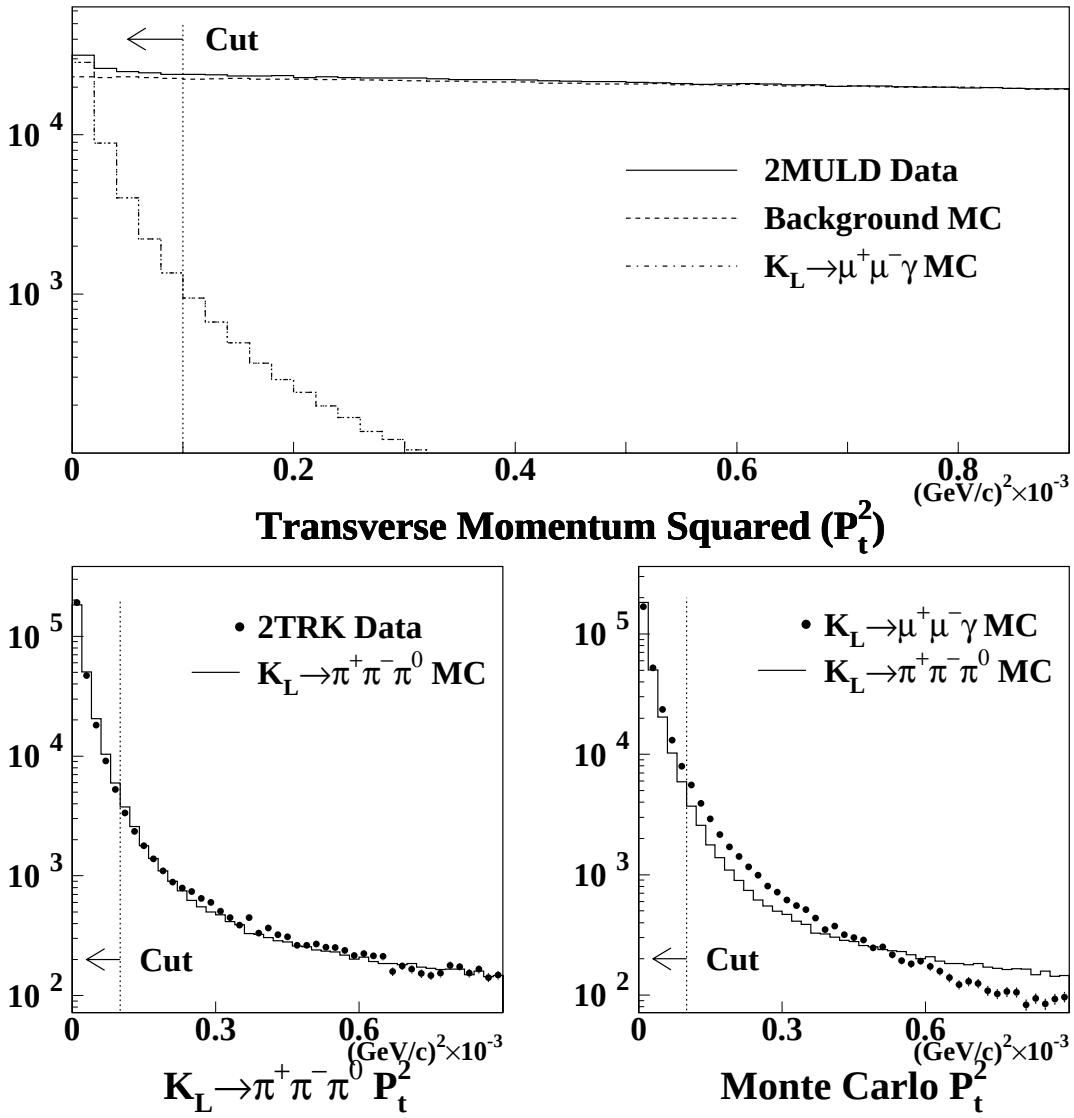

Figure 6.17: P_t^2 for signal and normalization. The upper plot shows the P_t^2 cut in the signal mode relative to the data, and background and signal Monte Carlo. Monte Carlo normalizations are arbitrary. The lower left plot displays the cut in the normalization mode, and a comparison of the signal and normalization Monte Carlo is made in the lower right plot. All plots shown are prior to any background cuts. The arrows and dotted lines indicate the accepted regions and cut placements respectively.

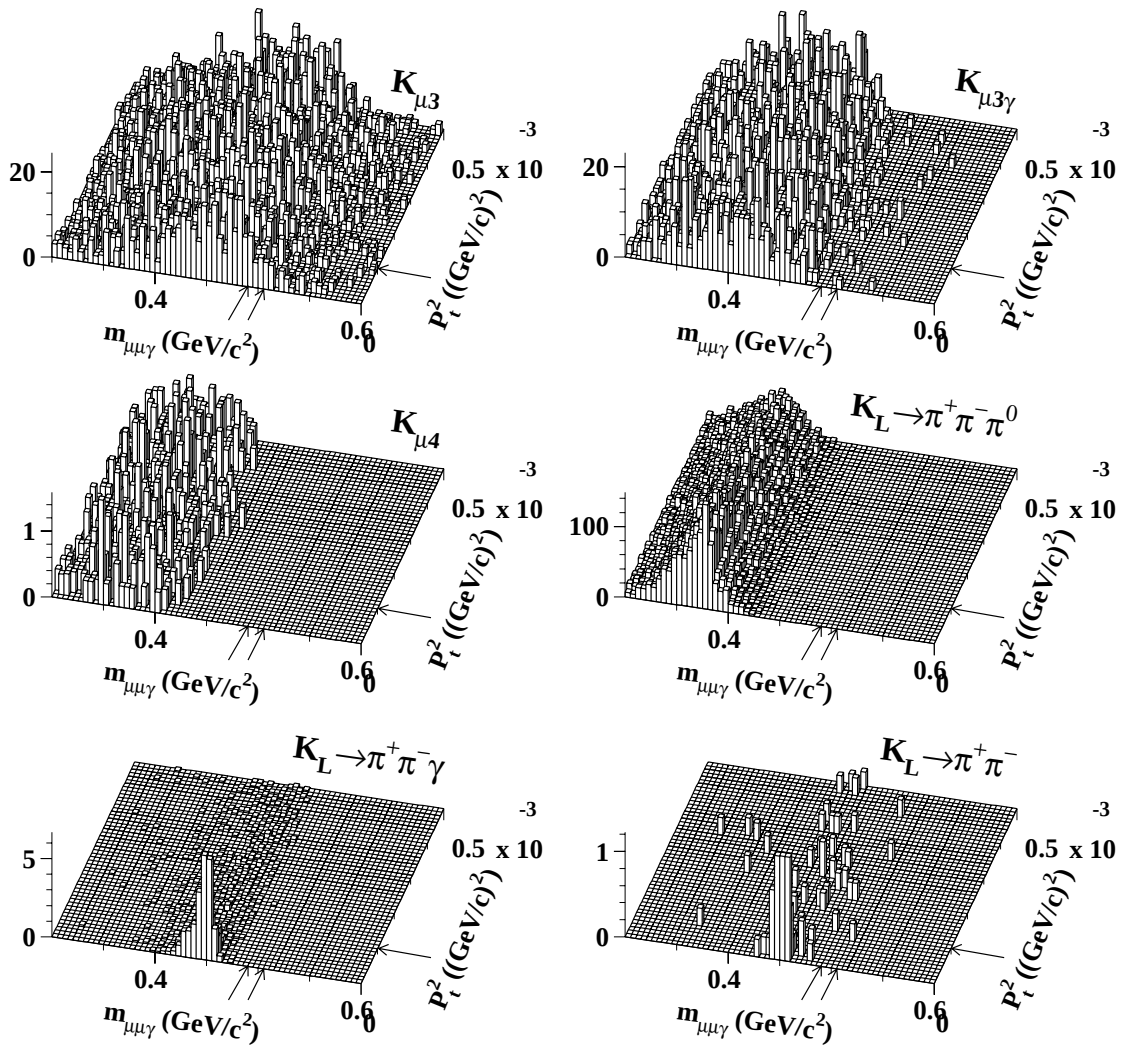

Figure 6.18: P_t^2 vs. $m_{\mu^+\mu^-\gamma}$ for backgrounds to $K_L \to \mu^+\mu^-\gamma$. The arrows indicate the signal region of $0.490\,GeV/c^2 < m_{\mu^+\mu^-\gamma} < 0.506\,GeV/c^2$, $P_t^2 < 0.0001\,(GeV/c)^2$. The plots show Monte Carlo distributions after all other cuts have been made, and each sample has been normalized to the K_L flux.

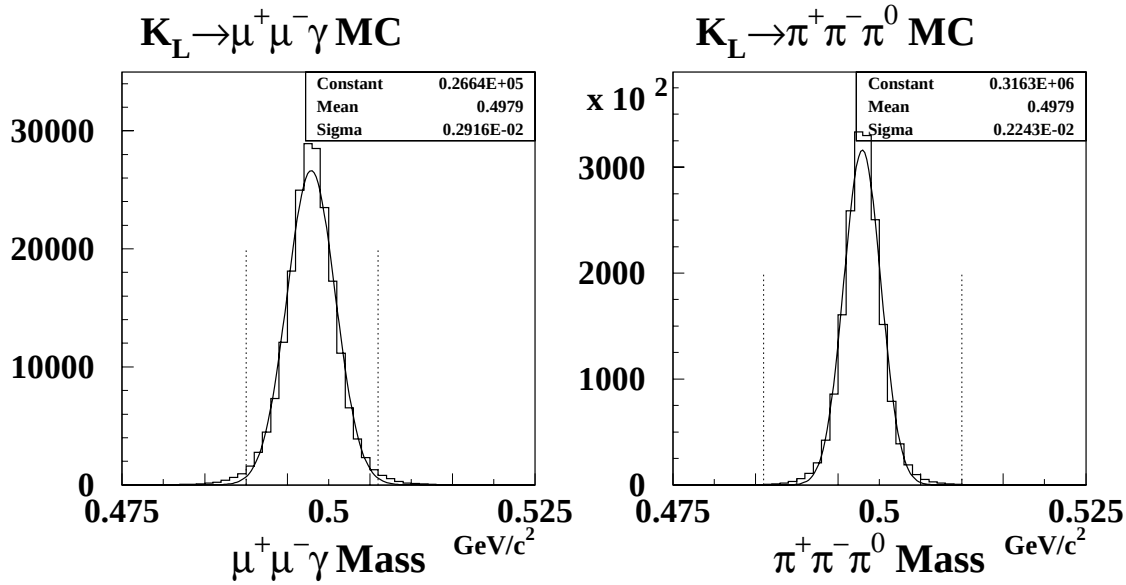

Figure 6.19: Monte Carlo mass distributions for $K_L \to \mu^+\mu^-\gamma$ and $K_L \to \pi^+\pi^-\pi^0$. The 3σ cut for $K_L \to \mu^+\mu^-\gamma$ and the 5σ cut for $K_L \to \pi^+\pi^-\pi^0$ are indicated by the dotted lines, and set by the Gaussian fit resolutions. The Monte Carlo is shown after all other analysis cuts have been applied.

to a crunch using the calibration constants available at the time of the data crunch. Because of this, rejection of events by the crunch due to early calibration inaccuracies is not simulated in the Monte Carlo. This small, but significant effect was studied by crunching and re-analyzing a subset of the $K_L \to \pi^+\pi^-\pi^0$ Monte Carlo. The ratio of the crunched to uncrunched acceptance of this subset is used as a correction factor to the overall $K_L \to \pi^+\pi^-\pi^0$ acceptance. Thus the kaon flux, Φ_{K_L}, is

$$\Phi_{K_L} = \frac{N_{\pi^+\pi^-\pi^0}}{A_{\pi^+\pi^-\pi^0}} \times \frac{PS}{BR(K_L \to \pi^+\pi^-\pi^0)}, \tag{6.3}$$

where $N_{\pi^+\pi^-\pi^0}$ is the number of $K_L \to \pi^+\pi^-\pi^0$ events in the final data sample, $A_{\pi^+\pi^-\pi^0}$ is the corrected ratio of accepted to generated Monte Carlo events, PS is the combined 2TRK trigger and split prescale of 7000, and $BR(K_L \to \pi^+\pi^-\pi^0) = 0.1256 \pm 0.0020$ [20].

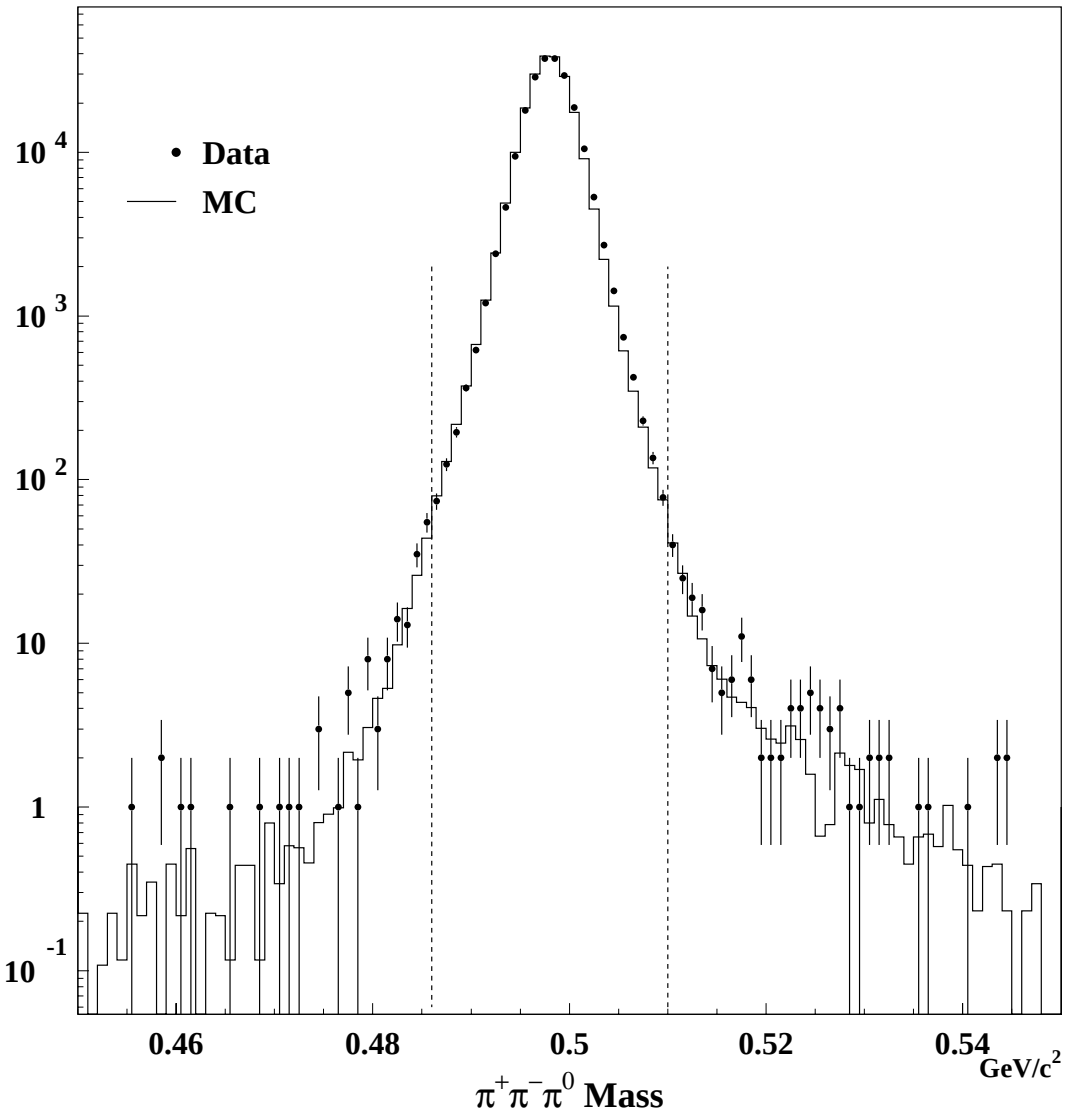

Figure 6.20: Reconstructed mass distribution for $K_L \to \pi^+\pi^-\pi^0$ after final cuts. The mass window of $0.486\,GeV/c^2 < m_{\mu^+\mu^-\gamma} < 0.510\,GeV/c^2$ is indicated by the dotted lines.

Table 6.3: $K_L \to \pi^+\pi^-\pi^0$ sample sizes and calculated flux.

	Winter	Summer
Data events passing all cuts	116327	94333
MC events passing all cuts	1002786	871256
MC events generated	27480686	21880871
Acceptance correction factor	0.9888	0.9994
Corrected acceptance	3.608%	3.980%
Flux	1.797×10^{11}	1.321×10^{11}

Separate Monte Carlo samples were generated for the Winter and Summer Runs, and the proper weighting of those samples cannot be known until the actual ratio of Winter and Summer data is determined.[8] Because it is precisely the flux measurement that determines that ratio, the K_L flux must be calculated separately for the Winter and Summer Runs. The data and Monte Carlo sample sizes are listed along with the calculated fluxes in Table 6.3. The total flux is found to be $(3.118 \pm 0.011) \times 10^{11}$, and the average acceptance for $K_L \to \pi^+\pi^-\pi^0$, weighted by the Winter and Summer fluxes, is $(3.765 \pm 0.003)\%$, where both errors are purely statistical.[9]

6.4.2 Background Estimation

To make an estimation of the final background level to $K_L \to \mu^+\mu^-\gamma$, the Monte Carlo samples of each background mode have to be normalized to the data. The most straightforward way to do this is to normalize the Monte Carlo to the absolute calculated flux. This is, in fact, what is done for all background modes except for

[8]Winter and Summer Monte Carlo samples were generated separately for a variety of reasons such as differences in detector and trigger configurations. Perhaps the most important reason was the development of architecture designed to facilitate analysis of Winter data before the Summer Run had completed.

[9]The systematic error on flux and acceptance calculations can be relatively large. This is because these calculations do not involve a ratio of acceptances, in which most systematic biases will cancel. In KTeV, flux calculations varied by as much as 20%, depending on which mode was used to determine the flux.

$K_{\mu 3}$. The $K_{\mu 3}$ case is complicated by the fact that the acceptance in that mode is completely dependent upon the presence of accidental activity. The accidental events used in the Monte Carlo underestimate the level of detector activity somewhat. To avoid sensitivity to accidental effects, the $K_{\mu 3}$ background was normalized to the sidebands around the signal mass window in the data. Final mass distributions for all background modes are shown in Figure 6.21.

Combining all normalized Monte Carlo samples, a total of 221.9 $\pm$ 14.9 background events are estimated (see top panel of Figure 6.22). Of these events, 102.7 are from the Winter Run, and 119.2 are from the Summer. The excellent agreement between data and Monte Carlo below 380 MeV/c^2 confirms that the flux calculations are accurate. In this region, the distribution is completely dominated by $K_L \to \pi^+\pi^-\pi^0$, which is normalized to the absolute flux. This data/Monte Carlo agreement shows that there is no trigger bias involved in normalizing 2MULD data to data from the 2TRK trigger, which was used to calculate the flux.

The background estimation is checked through two methods. The first method makes use of the fact that all of the data at high P_t^2 are background events. The K_L mass distribution in data from 0.0004 (GeV/c)^2 $< P_t^2 <$ 0.001 (GeV/c)^2 is plotted and scaled down by a factor of 6 to correspond to the size of the signal P_t^2 range. This method assumes that the background is flat with respect to P_t^2. This is approximately true except for the case of $K_L \to \pi^+\pi^-\pi^0$, which lies far outside the signal mass cut anyway. According to the center plot in Figure 6.22, 220 $\pm$ 6 background events fall in the mass window using this estimation.

One can also perform a simple fit to the final K_L mass distribution in data. The fit uses a Gaussian to represent the signal peak, another Gaussian for the low mass $K_{\mu 3}$, $K_{\mu 3\gamma}$, and $K_L \to \pi^+\pi^-\gamma$, and a falling exponential for the high mass $K_{\mu 3}$. A fairly good fit is achieved with this functional form, which is shown in the lower plot of Figure 6.22. The fit predicts a background of 215 $\pm$ 15 events.

6.4.3 Signal

The $K_L \to \mu^+\mu^-\gamma$ mass distribution of the 2MULD data after all other analysis cuts have been applied is shown in Figure 6.23. In the left panel of Figure 6.19 is

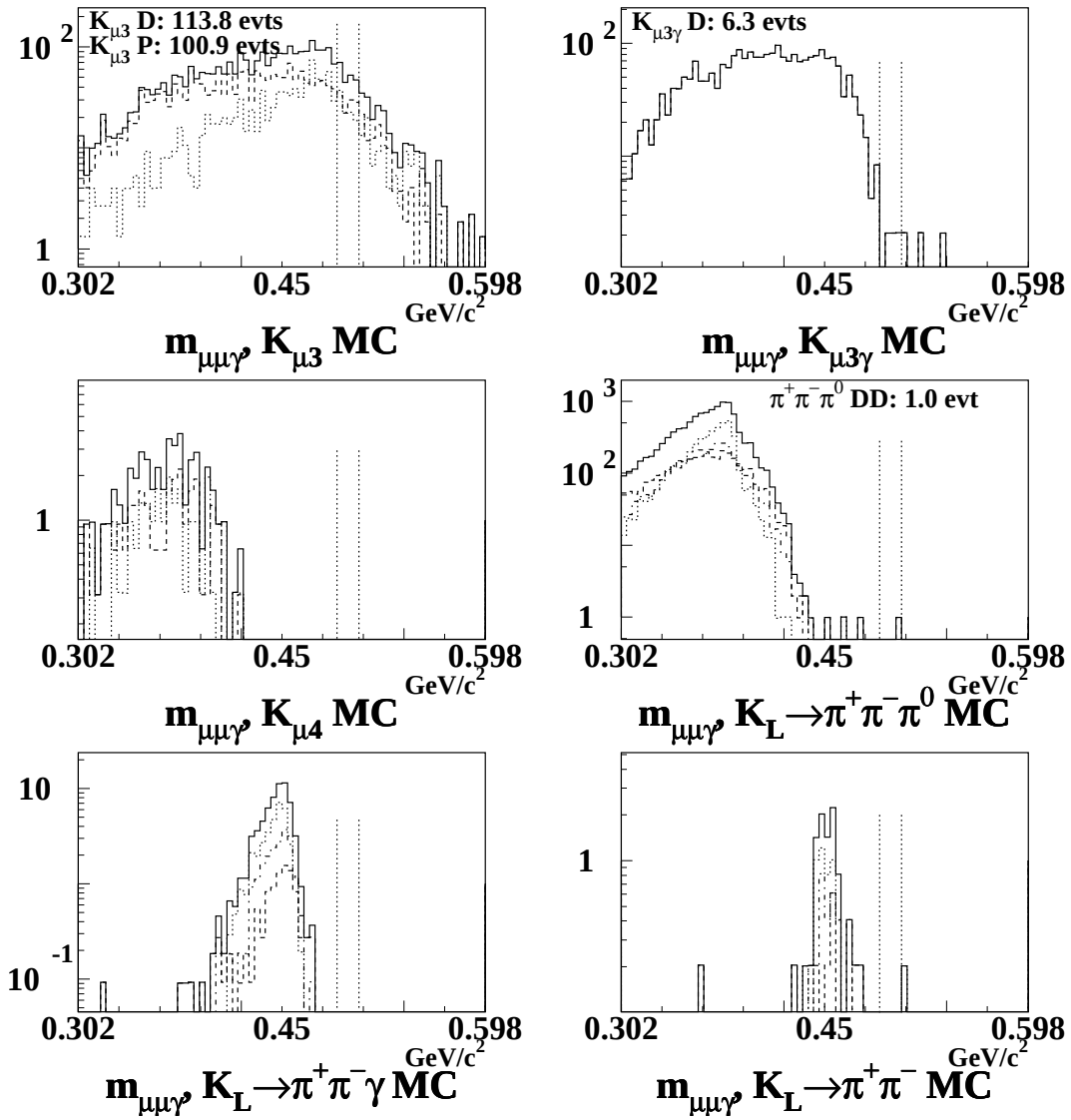

Figure 6.21: Components of the background to $K_L \to \mu^+\mu^-\gamma$. Dashed: $\pi^\pm$ decay (D); Dotted: $\pi^\pm$ punch-through (P); Dash-dotted: one $\pi^\pm$ decay, one $\pi^\pm$ punch-through; Solid: total. Vertical dotted lines indicate the mass window for $K_L \to \mu^+\mu^-\gamma$.

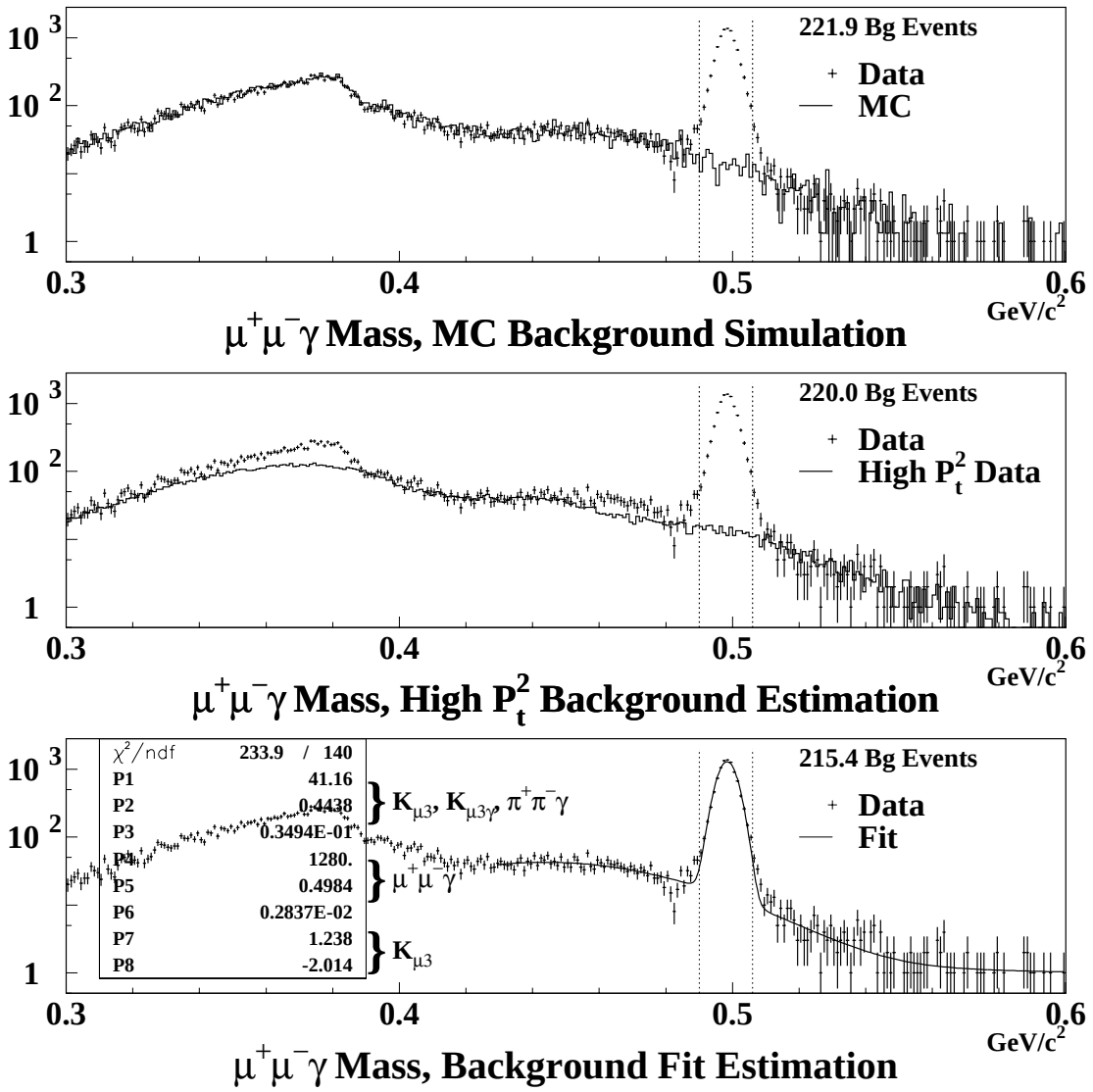

Figure 6.22: Estimations of the background to $K_L \to \mu^+\mu^-\gamma$. Top: Monte Carlo normalized to sidebands ($K_{\mu 3}$ and $K_{\mu 3\gamma}$) or measured flux (all other). Middle: data in the region $0.0004\,(GeV/c)^2 < P_t^2 < 0.001\,(GeV/c)^2$ scaled to the signal window of $P_t^2 < 0.0001\,(GeV/c)^2$. Bottom: three function fit to the data (a Gaussian for $K_L \to \mu^+\mu^-\gamma$, a Gaussian for the $K_{\mu 3}$, $K_{\mu 3\gamma}$, $K_L \to \pi^+\pi^-\gamma$ peak at 444 MeV, and an exponential tail from $K_{\mu 3}$.

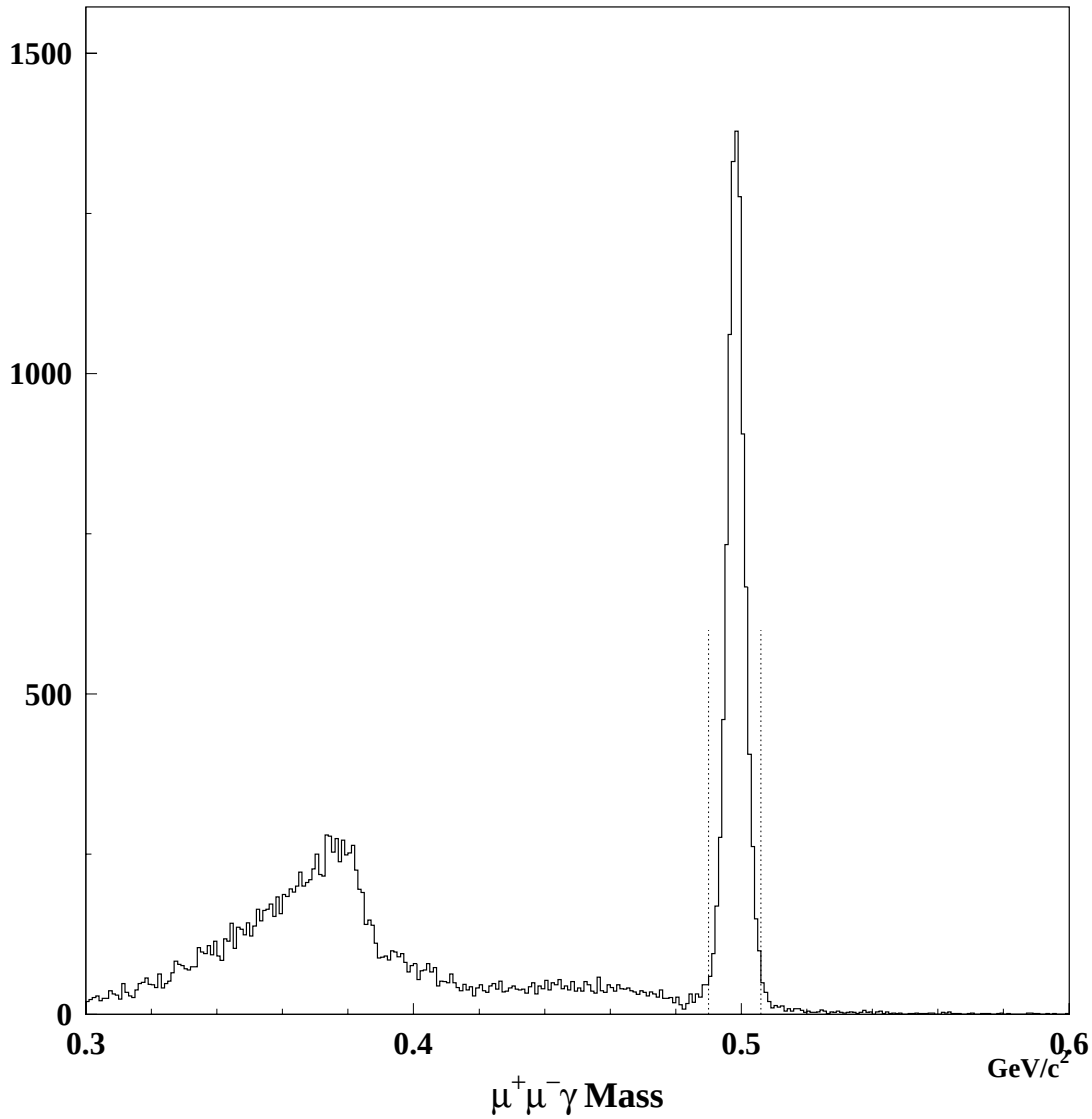

Figure 6.23: Reconstructed mass distribution for $K_L \to \mu^+\mu^-\gamma$ after final cuts. The mass window of $0.490\,GeV/c^2 < m_{\mu^+\mu^-\gamma} < 0.506\,GeV/c^2$ is indicated by the dotted lines.

Table 6.4: $K_L \to \mu^+\mu^-\gamma$ data and Monte Carlo sample sizes.

	Winter	Summer
Data events passing all cuts	5250	4077
MC events passing all cuts	108253	94814
MC events generated	1401283	1166752

found the analogous Monte Carlo distribution. The sample sizes of $K_L \to \mu^+\mu^-\gamma$ data and Monte Carlo are summarized in Table 6.4. A total of 9327 data events remain in the mass window between 490 MeV/c^2 and 506 MeV/c^2. 221.9 of these events are estimated to be background, so the final $K_L \to \mu^+\mu^-\gamma$ sample consists of 9105.1 events with a signal to background ratio of 41:1. The signal peak is especially prominent on a plot of P_t^2 vs. $m_{\mu^+\mu^-\gamma}$ (Figure 6.24), which emphasizes the paucity of background.

6.5 Branching Ratio Calculation

The $K_L \to \mu^+\mu^-\gamma$ branching ratio is calculated using the formula,

$$BR(K_L \to \mu^+\mu^-\gamma) = \frac{1}{\Phi_{K_L}} \times \frac{N_{\mu^+\mu^-\gamma} - N_{bg}}{A_{\mu^+\mu^-\gamma}}. \tag{6.4}$$

Φ_{K_L} is the flux from Equation 6.3. $N_{\mu^+\mu^-\gamma}$ is the number of $\mu^+\mu^-\gamma$ candidates observed, N_{bg} is the estimated number of background events, and $A_{\mu^+\mu^-\gamma}$ designates the signal acceptance.

Proper weighting of the Monte Carlo samples must be done to calculate the branching ratio using the entire data set. This is because the ratio of Monte Carlo events generated for the Winter and Summer Runs is not the same as the ratio of the Winter and Summer fluxes. The most straightforward way to combine the Monte Carlo is to use an average of the Winter and Summer $K_L \to \mu^+\mu^-\gamma$ acceptances, weighted by the Winter and Summer fluxes. The flux-weighted acceptance for $K_L \to \mu^+\mu^-\gamma$ is $(7.895 \pm 0.018)\%$.

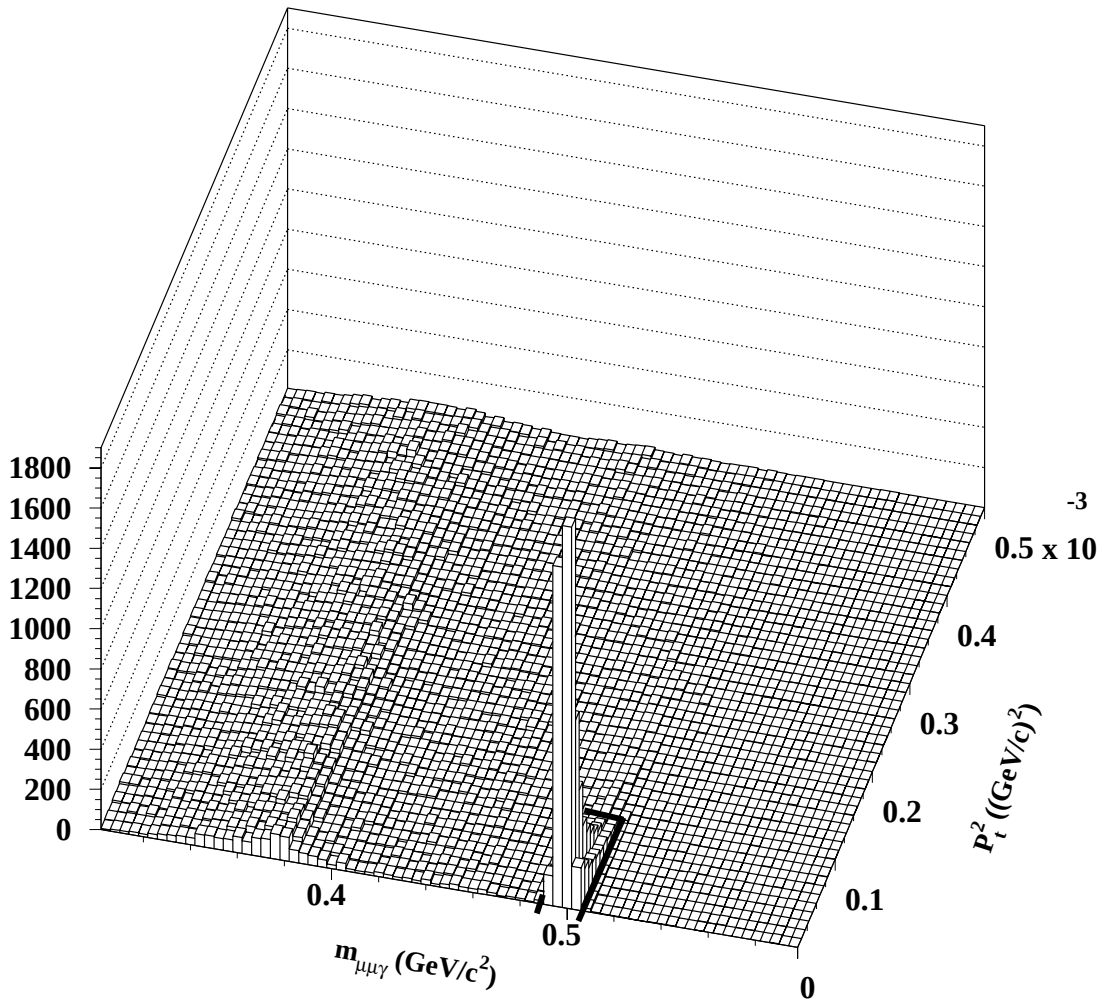

Figure 6.24: P_t^2 vs. $m_{\mu^+\mu^-\gamma}$ in data after final cuts. The signal region of $0.490\ GeV/c^2 < m_{\mu^+\mu^-\gamma} < 0.506\ GeV/c^2$, $P_t^2 < 0.0001\ (GeV/c)^2$ is indicated by the heavy black lines.

Table 6.5: Statistics of the $BR(K_L \to \mu^+\mu^-\gamma)$ calculation.

	Winter	Summer	Total
$K_L \to \mu^+\mu^-\gamma$ candidates	5250	4077	9327
Estimated background	102.7	119.2	221.9
Signal events	5147.3	3957.8	9105.1
$K_L \to \mu^+\mu^-\gamma$ acceptance	7.725%	8.126%	7.895%
K_L Flux	1.797×10^{11}	1.321×10^{11}	3.118×10^{11}
$BR(K_L \to \mu^+\mu^-\gamma)$	3.71×10^{-7}	3.69×10^{-7}	3.70×10^{-7}

After al¡l analysis cuts, there are 9327 $\mu^+\mu^-\gamma$ candidates, with an estimated background of 221.9 ± 14.9. This gives a background-subtracted signal of 9105.1 ± 96.6 events. When these numbers are inserted into Equation 6.4 with the total flux of 3.118×10^{11}, the branching ratio is calculated to be

$$BR(K_L \to \mu^+\mu^-\gamma) \ = (3.70 \pm 0.04_{stat}) \times 10^{-7}.$$

The branching ratio can also be determined separately for the Winter and Summer Runs. Those calculations yield the results

$$BR(K_L \to \mu^+\mu^-\gamma)_{Winter} = (3.71 \pm 0.05_{stat}) \times 10^{-7}$$
$$BR(K_L \to \mu^+\mu^-\gamma)_{Summer} = (3.69 \pm 0.06_{stat}) \times 10^{-7}.$$

Note that a flux-weighted average of the Winter and Summer branching ratios gives the same overall result as that obtained using the flux-weighted $K_L \to \mu^+\mu^-\gamma$ acceptance. The statistics of the branching ratio calculations are summarized in Table 6.5.

6.6 Systematic Errors

Various biases and limitations systematic to the experiment and analysis contribute to the uncertainty of the branching ratio measurement. A large part of this system-

atic error is associated with the differences between the signal and normalization modes. The most obvious of these is the fact that the muon system is critically important for $K_L \to \mu^+\mu^-\gamma$ observation, but completely irrelevant to the normalization. Therefore, any uncertainty in the efficiency for detecting muons maps directly to an error on the branching ratio. Other biases result from discrepancies between data and Monte Carlo simulation. These discrepancies become errors only to the extent that they are different in the two modes. In general, this type of systematic error is evaluated by measuring the fluctuations in the branching ratio caused by variations in analysis cuts or Monte Carlo parameters. Limitations of the measurement technique that contribute to the systematics include Monte Carlo statistics, the dependence of the $K_L \to \mu^+\mu^-\gamma$ acceptance on the form factor, and the error on $BR(K_L \to \pi^+\pi^-\pi^0)$.

6.6.1 Muon Efficiency

There are two systematic errors on the muon efficiency. The first is a systematic error assigned to the measurement of the cracks in MU3. As described in Section 5.3.2, understanding the muon system involves simulating MU3 cracks in the Monte Carlo, and then comparing the simulated muon efficiency to that found in data. The sizes of MU3 cracks were varied in the Monte Carlo to determine the range over which there is no measurable improvement in the efficiency modeling. The $K_L \to \mu^+\mu^-\gamma$ acceptance difference over this range, which is approximately 0.2 mm, is 0.5%.

The second systematic error in muon efficiency concerns the energy loss simulation. One way to gauge the effect of a misunderstanding of the energy loss is to vary the thickness of the Muon Filters in the Monte Carlo by some reasonable amount. This forces a change in the energy loss, which will alter the acceptance for muons. The amount of material in the Muon Filters was changed by 2%. This number reflects the extreme of the possible mismeasurement of the filter thickness due to gaps in the steel shielding blocks. A 0.4% change in acceptance results from the altered momentum threshold caused by such a change in the scattering material. The altered Monte Carlo threshold shown in Figure 6.25 is noticeably worse than with the nominal filter thickness. This confirms that a 2% change in thickness is a

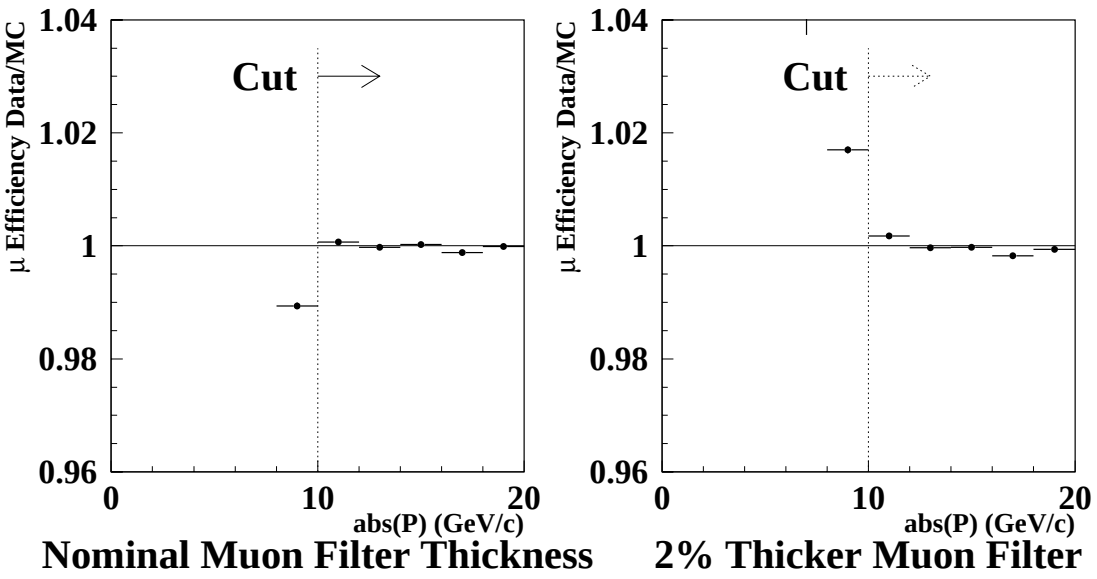

Figure 6.25: Muon threshold with different filter thicknesses. The left-hand plot shows the data/Monte Carlo ratio of muon efficiency vs. track momentum using the nominal filter thicknesses in the Monte Carlo. The ratio on the right uses Monte Carlo with the filter thickness increased by 2%.

reasonable upper limit for energy loss systematic studies.

6.6.2 Drift Chamber Effects

Drift chamber performance has a large effect on acceptance since it completely determines the ability to reconstruct tracks. Misunderstandings of efficiencies, illuminations, Z distributions, etc., can all bias the branching ratio result. Some of these effects are correlated, and moreover many will cancel in the acceptance ratio of signal to normalization. The drift chamber maps described in Section 5.3.1 address all of these issues simultaneously. A conservative error estimate for drift chamber simulation is made by varying the global weight given to the drift chamber map corrections from 1 to 0. This effectively goes from using the nominal drift chamber map simulation to not using it at all. Over that range, the largest fluctuation in the signal to normalization acceptance ratio is 0.5%.

6.6.3 Photon Energy (E_γ) Uncertainty

The photon energy spectrum is not well modeled at low energies owing to uncertainty in HCC thresholds. This is not an issue for this analysis because offline analysis cuts are made well above threshold (see Section 6.3.3). What is a concern though, is an overall slope in the ratio of $K_L \to \pi^+\pi^-\pi^0$ data and Monte Carlo photon spectra (especially since the signal and normalization have different numbers of photons). The mean photon energy in data is 0.8% higher than in Monte Carlo. Reconstructed mass peaks for all decay modes involving photons are $\sim$ 0.2-0.3 MeV/c^2 higher in data than in Monte Carlo. Mismeasurement of E_γ has long been suspected as the reason for this.

The systematic error arising from the E_γ uncertainty was estimated by shifting the photon cluster energies in Monte Carlo higher by 0.8%. This completely removes the E_γ slope shown in Figure 6.26. However, it is an overcorrection as evidenced by the $K_L \to \pi^+\pi^-\pi^0$ mass distribution seen in Figure 6.27. Even so, this shift only produces a 0.1% change in the $\pi^+\pi^-\pi^0$ acceptance and this value is assigned as a systematic error. Note that this error accounts for the correlation in uncertainties due to the E_γ distribution as well as analysis cuts on the reconstructed mass.

6.6.4 Analysis Cuts

There are two analysis cuts that produce non-negligible systematic errors. The most significant is the cut on the square of the transverse momentum, P_t^2. The 100 (MeV/c)^2 cut is fairly tight in order to remove $K_{\mu 3}$ background. Unfortunately, as was discussed in Section 6.3.3, this cut lies in the middle of significant $K_L \to \pi^+\pi^-\pi^0$ data/Monte Carlo disagreement. Compounding this problem is the fact that the signal and normalization P_t^2 distributions are markedly different. To determine the significance of this error, the P_t^2 cut was moved between 50 (MeV/c)^2 and 1000 (MeV/c)^2. The result of this study, shown in Figure 6.28, is a maximum branching ratio variation of 0.6%.

The cut on track momentum is important since it is closely tied to the muon threshold, and thus can have disparate effects on signal and normalization. The

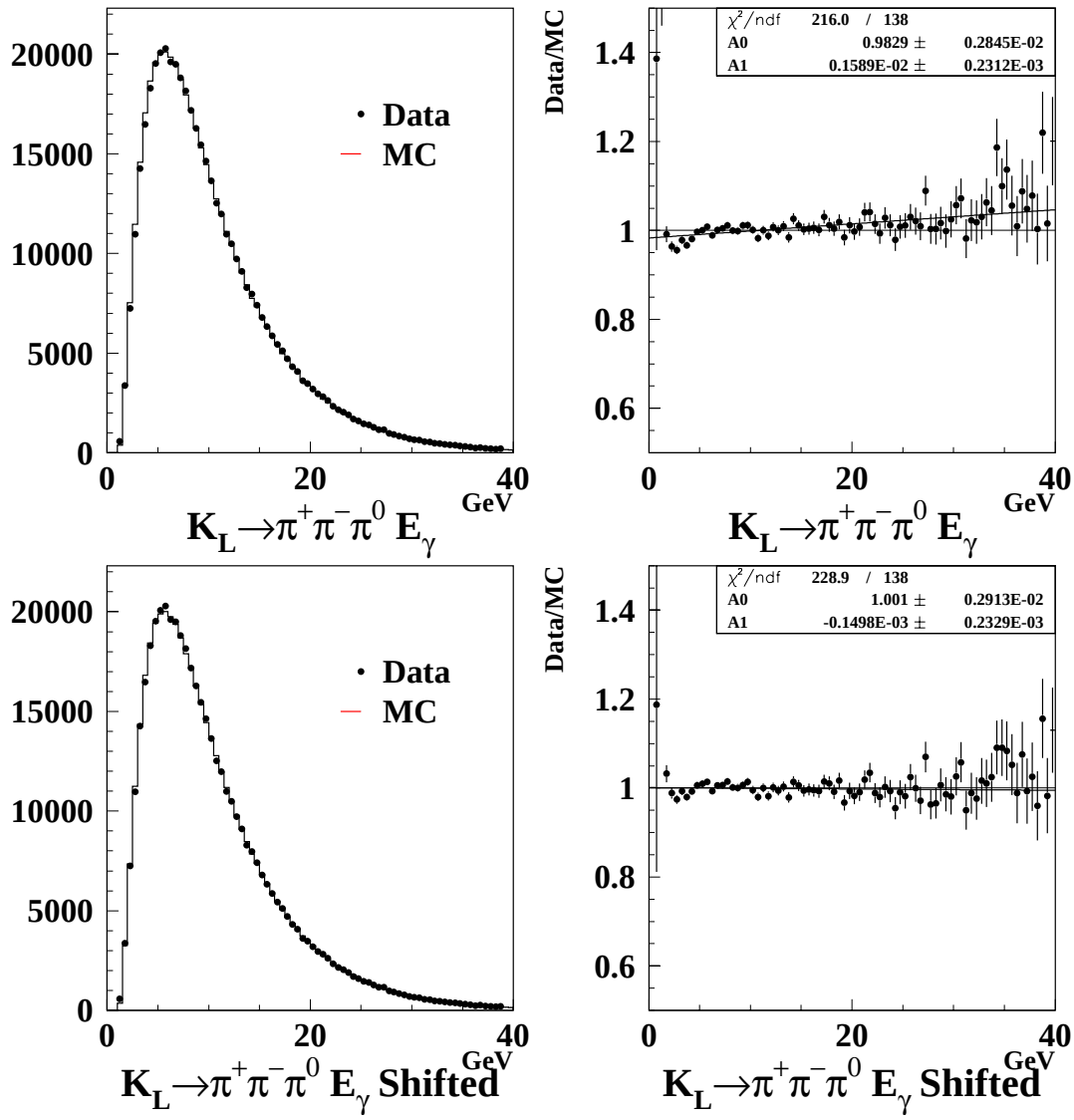

Figure 6.26: $K_L \to \pi^+\pi^-\pi^0$ photon energy spectrum. For the lower plots, Monte Carlo photon energies have been shifted 0.8% higher.

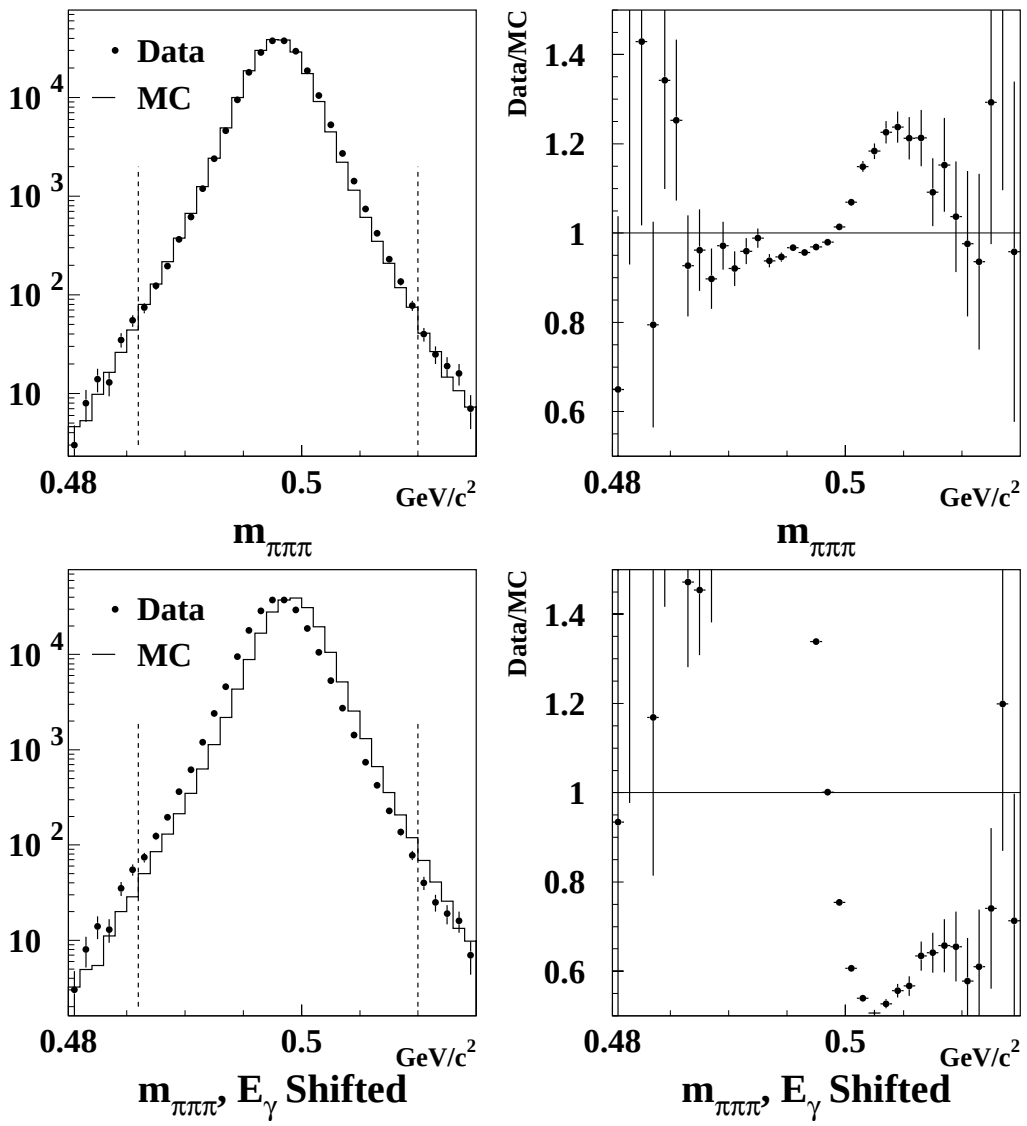

Figure 6.27: $K_L \to \pi^+\pi^-\pi^0$ mass distribution. For the lower plots, Monte Carlo photon energies have been shifted 0.8% higher. The mass window of $0.486\ \mathrm{GeV/c^2} < m_{\mu^+\mu^-\gamma} < 0.510\ \mathrm{GeV/c^2}$ is indicated by the dotted lines.

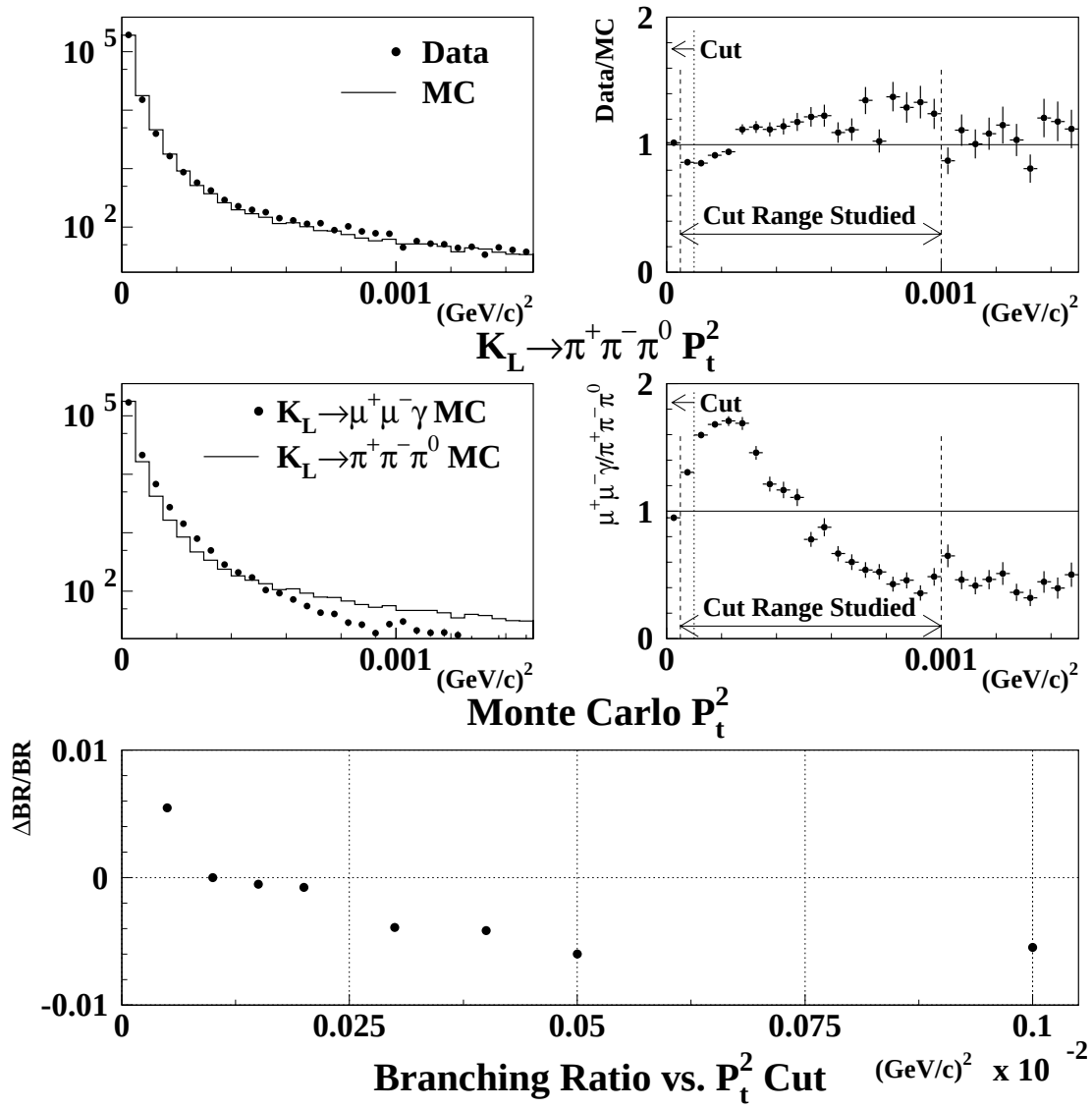

Figure 6.28: Study of P_t^2 cut. A comparison of $K_L \to \pi^+\pi^-\pi^0$ data and Monte Carlo is shown in the upper plots, and Monte Carlo for $K_L \to \mu^+\mu^-\gamma$ and $K_L \to \pi^+\pi^-\pi^0$ are compared in the center plots. The lower graph displays the fractional change in the $K_L \to \mu^+\mu^-\gamma$ branching ratio as the cut in P_t^2 is varied.

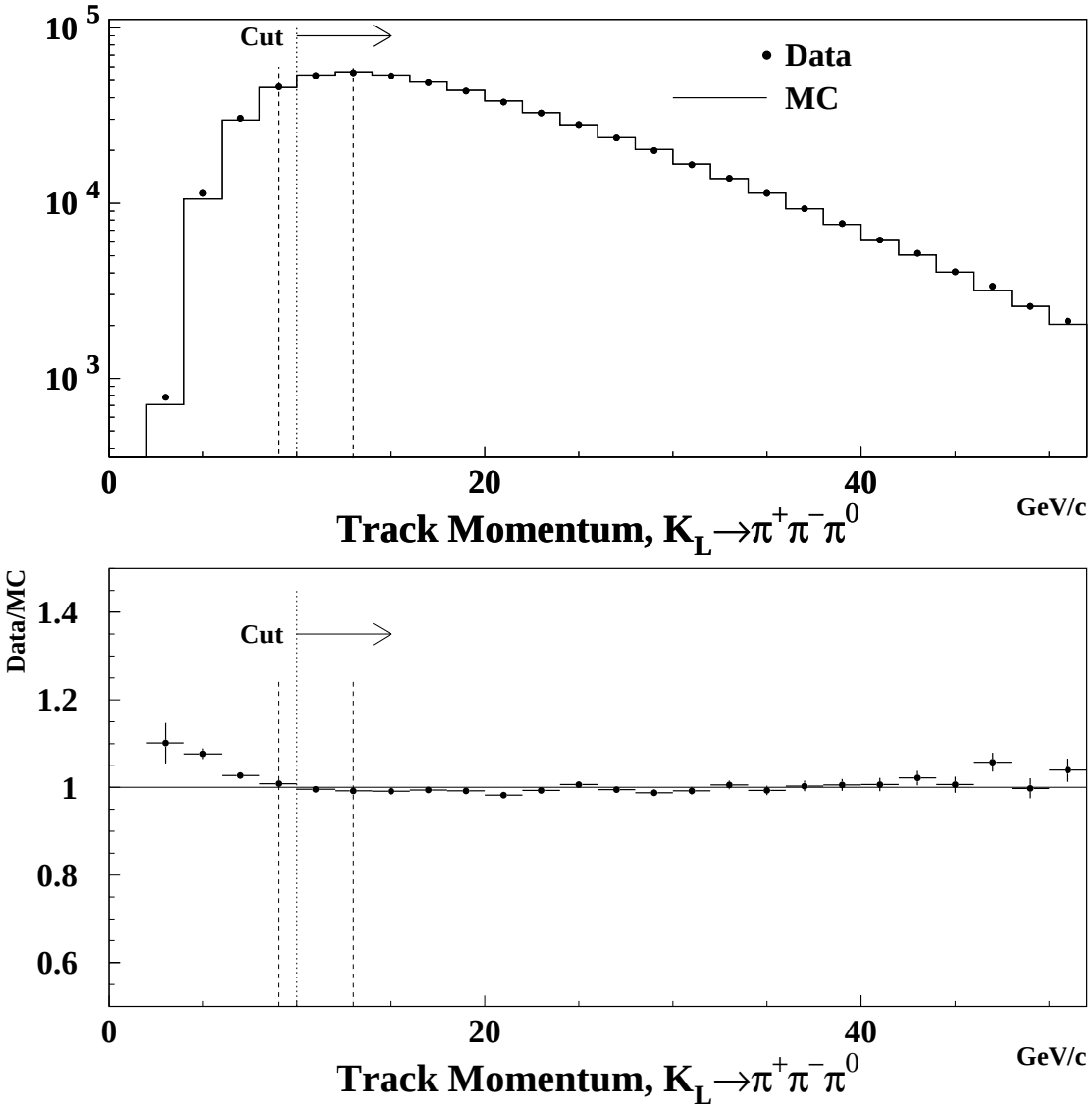

Figure 6.29: Track momentum distributions for $K_L \to \pi^+\pi^-\pi^0$. The data and Monte Carlo distributions shown are after all other analysis cuts have been applied. The arrows and dotted lines indicate the accepted regions and cut placements respectively. The range of track momentum cuts studied is indicated by the dashed lines.

analysis cut is varied from 9 GeV/c to 13 GeV/c (Figure 6.29). The cut cannot be placed any lower than 9 GeV/c, or it starts to move off of the muon threshold. Above 13 GeV/c, the statistics drop off too much to make good measurements.

Variations of the cut within this range produce 0.2% fluctuations in the branching ratio.

6.6.5 Other

The Monte Carlo background simulation is chosen as the estimate most representative of the actual background. The simulation predicts 221.9 ± 14.9 events, and the 14.9 event uncertainty translates into a 0.2% branching ratio error. $K_L \to \mu^+\mu^-\gamma$ Monte Carlo statistics contribute an additional 0.2% systematic error.

The BMS form factor was used to generate the Monte Carlo samples for acceptance studies. A choice for α_{K^*} must be made to use this form factor. The value chosen is that measured by this experiment through an analysis of the dimuon mass distribution, $\alpha_{K^*} = -0.193$ (see Section 7.2.1). The dependence of the Monte Carlo acceptance on α_{K^*} was studied by varying its value from -0.5 to $+0.1$. The 0.4% acceptance fluctuation over this range is assigned as a systematic error.

The dominant error in this measurement is the external systematic error on the $K_L \to \pi^+\pi^-\pi^0$ branching ratio. The PDG value is $BR(K_L \to \pi^+\pi^-\pi^0) = 0.1256 \pm 0.0020$. This represents a 1.6% error.

6.6.6 Summary

All systematic errors are summarized in Table 6.6, and lead to a final branching ratio result of

$$BR(K_L \to \mu^+\mu^-\gamma) = (3.70 \pm 0.04_{stat} \pm 0.07_{syst}) \times 10^{-7}.$$

Table 6.6: Systematic errors.

Source of Error	BR Error
MU3 Crack Simulation	0.5%
Muon Filter Thickness	0.4%
DC Simulation (Maps,Illum,Z)	0.5%
Photon Energy Uncertainty	0.1%
P_t^2 Cut	0.6%
Track Momentum Cut	0.2%
Background Uncertainty	0.2%
Monte Carlo Statistics	0.2%
Monte Carlo form factor	0.4%
Total from internal systematics	1.1%
$BR(K_L \to \pi^+\pi^-\pi^0)$	1.6%
Total Systematic Error	1.9%

CHAPTER 7
FORM FACTOR ANALYSIS

7.1 Direct $f(x)$ Measurement

The differential decay rate for $K_L \to \mu^+\mu^-\gamma$ is expressed in Equation 1.19 as a function only of the invariant mass of the off-shell photon, which is equivalent to the dimuon invariant mass, $m_{\mu\mu}$. By definition, the form factor $f(x)$ (where $x = (m_{\mu\mu}/m_{K_L})^2$) is also solely dependent on the dimuon mass, independent of any specific model parameterization. Therefore, the most direct measurement of the form factor possible is simply to measure the value of $f(x)$ as a function of x.

A direct $f(x)$ analysis is accomplished through a very simple study of the $K_L \to \mu^+\mu^-\gamma$ dimuon mass distribution. The ratio is taken of the x distribution from data to that from Monte Carlo generated with the form factor set to unity. This ratio is a completely model-independent measurement of $f(x)$ as a function of x. Figure 7.1 shows the x distributions after all analysis cuts for data and for Monte Carlo generated with $f(x) = 1$. Also shown is the ratio which clearly exhibits the presence of a non-trivial form factor. The bin-by-bin values of $f(x)$ are given in the Appendix.

The data x distribution has been background subtracted using high P_t^2 data. The subtraction is done by plotting x for data in the region 0.0004 (GeV/c)$^2 < P_t^2 <$ 0.001 (GeV/c)^2. This distribution is then scaled down by a factor of 6 to match the signal region size. Finally, the scaled background is subtracted from the signal distribution, which is then compared to the Monte Carlo. Two alternate methods of background subtraction were also investigated. A fit to the high P_t^2 data can be used rather than the raw data. This method is motivated by low statistics of the background distribution. The final x distribution from background Monte Carlo

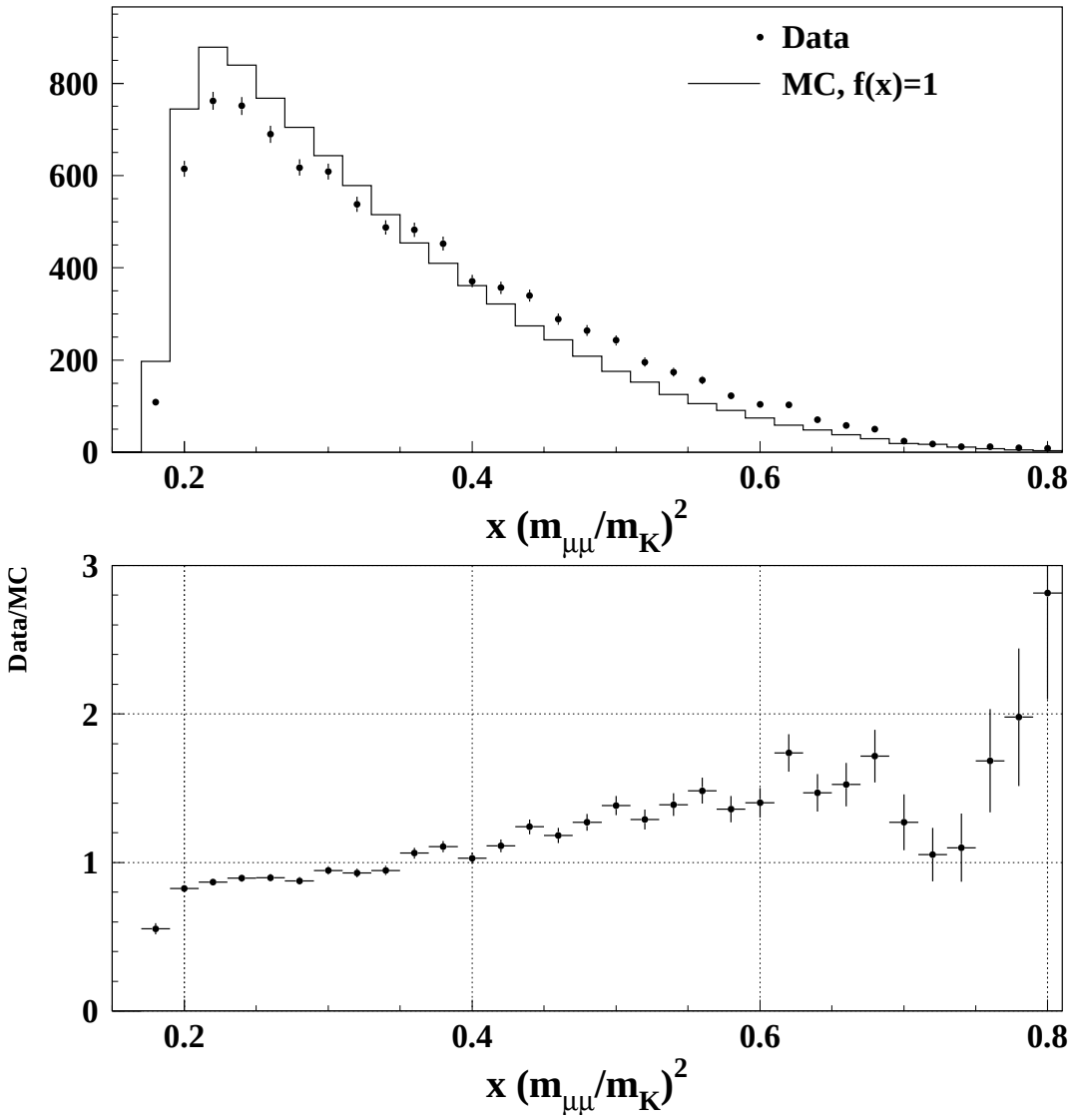

Figure 7.1: Dimuon mass distribution for data and Monte Carlo with no form factor (top). The data/Monte Carlo ratio is the measured form factor (bottom). The Monte Carlo is normalized to the total number of events in data.

can also be used for the background subtraction. Neither of these two alternate methods produced significant changes in the $f(x)$ measurement, or in any of the other measurements described later in this chapter.

An even more fundamental measurement is that of the acceptance corrected dimuon distribution. The data corrected with the bin-by-bin acceptance in x yields an x distribution that is directly proportional to the differential decay rate given by Equation 1.19. This distribution is presented in Figure 7.2 along with the $K_L \to \mu^+\mu^-\gamma$ acceptance as a function of x. The acceptance corrected data for each bin in x is listed in the Appendix.

7.2 α_{K^*} Measurement

The discussion in Section 1.3 pointed out that the BMS parameter α_{K^*} can be measured through an analysis of the dimuon mass distribution and from a calculation using the $K_L \to \mu^+\mu^-\gamma$ branching ratio. Both of these independent measurements of α_{K^*} are made, and then combined to give a final result.

7.2.1 $m_{\mu\mu}$ Shape Analysis

To measure α_{K^*} from the dimuon shape, Monte Carlo distributions of x are generated at many different values of α_{K^*}, and a negative log-likelihood (NLL) comparison of the data with each Monte Carlo distributions is then made. This method has a high discriminating power because the final data sample has large statistics, and the dimuon mass distribution in KTeV varies considerably as a function of α_{K^*} (Figure 7.3).

The first step in evaluating the likelihood of data being consistent with a given Monte Carlo set, is to construct the probability distribution,

$$P(\alpha_{K^*}; x_i) = N_{MC}(\alpha_{K^*}; x_i)\,\frac{9105.1}{\sum_{j=1}^{nbins} N_{MC}(\alpha_{K^*}; x_j)} + N_{Bg}(x_i)\,\frac{221.9}{\sum_{j=1}^{nbins} N_{Bg}(x_j)}. \tag{7.1}$$

N_{MC} is the binned x distribution from Monte Carlo generated at a particular value of

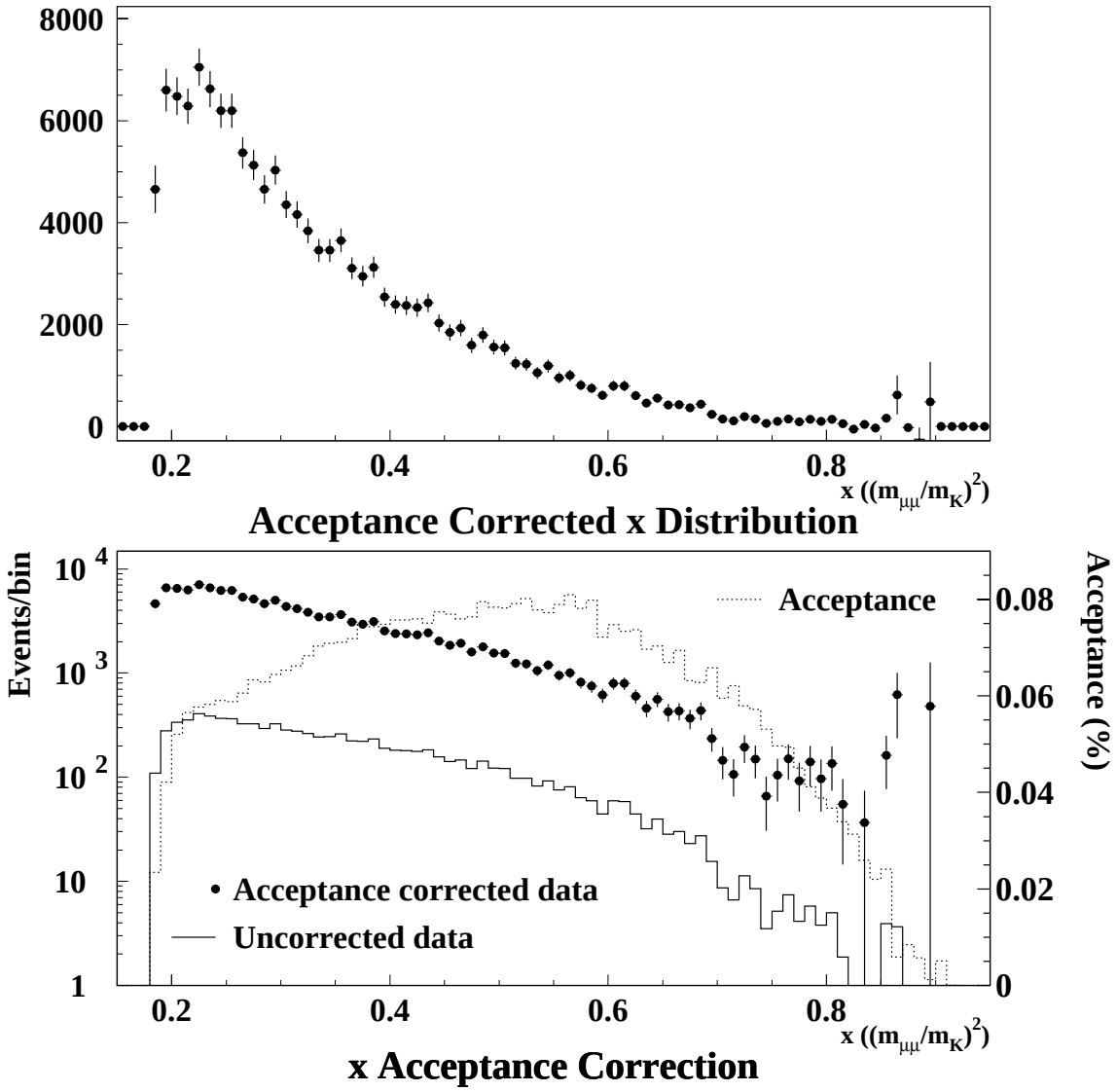

Figure 7.2: x distribution acceptance correction. The upper plot shows the x distribution from data after bin-by-bin acceptance corrections. The lower plot includes the measured x distribution before and after correction, along with the acceptance as a function of x.

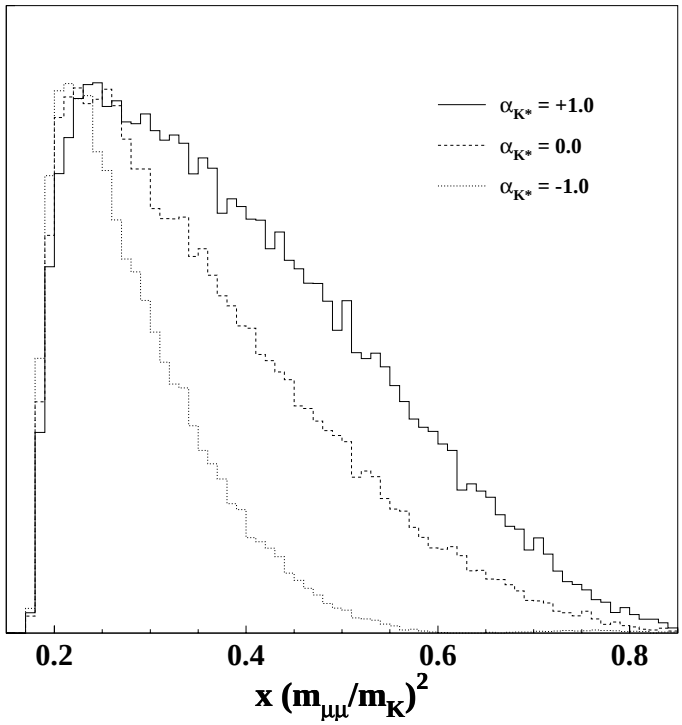

Figure 7.3: Dimuon mass distributions for Monte Carlo generated with various values of α_{K^*}.

α_{K^*} and is normalized to the 9105.1 signal events. N_{Bg} is the background distribution taken from high P_t^2 data, and is normalized to the estimated background of 221.9 events. There are 80 bins, 0.01 units wide, spanning the range $x = [0.15, 0.95]$. The overall probability is normalized such that $\sum_{i=1}^{nbins} P(\alpha_{K^*}; x_i) = 1$. The log-likelihood function then, is the log of the product of probabilities evaluated for each of the 9327 $K_L \to \mu^+\mu^-\gamma$ candidates. This is equivalent to

$$\mathcal{L}(\alpha_{K^*}) = -\sum_{i=1}^{nbins} N_{\mu\mu\gamma}(x_i) \log P(\alpha_{K^*}; x_i), \tag{7.2}$$

where $N_{\mu\mu\gamma}(x_i)$ is the number of $K_L \to \mu^+\mu^-\gamma$ candidate events in the ith bin in x.

$\mathcal{L}(\alpha_{K^*})$ was evaluated for 45 sets of Monte Carlo, generated with values of α_{K^*} ranging from -1.5 to 0.7 in steps of 0.05 units. Each Monte Carlo set consisted of

over 65000 events in the final sample. The minimum of $\mathcal{L}$ is the measured value of α_{K^*}, and the width of the function at $+1/2$ units represents the error on the measurement. Figure 7.4 shows that the result of this measurement for the entire data set is

$$\alpha_{K^*}^{Shape} = -0.193 \pm 0.035.$$

The results from the Winter and Summer data are

$$\alpha_{K^*}^{Winter} = -0.163 \pm 0.045$$
$$\alpha_{K^*}^{Summer} = -0.218^{+0.054}_{-0.057}.$$

The difference between the Winter and Summer numbers is 1.2σ from zero.

Ten thousand Monte Carlo 'Data' sets were generated with the measured value of $\alpha_{K^*} = -0.193$. For each set, 9327 events were drawn randomly from $P(-0.193; x_i)$. The NLL method is used to measure α_{K^*} for each of these sets, and the width of the α_{K^*} distribution is a check on the statistical error. A Gaussian fit to that distribution yields a result of $\alpha_{K^*} = -0.198 \pm 0.035$, the error of which is identical to that found with the NLL. One significant systematic error must be added to the statistical error on α_{K^*}: the measurement is sensitive to the cut on track momentum. Fluctuations as the cut is varied are all more negative, and the maximum shift is -0.034. Therefore, the value of α_{K^*} measured from the dimuon mass distribution is

$$\alpha_{K^*}^{Shape} = -0.193^{+0.035}_{-0.049}.$$

7.2.2 Branching Ratio Calculation

The parameter α_{K^*} is extracted from the $K_L \to \mu^+\mu^-\gamma$ branching ratio through the form factor and differential decay rate expressions. To facilitate the calculation of α_{K^*}, Equation 1.19 is written for the $\mu^+\mu^-\gamma$ case and normalized to , $(K_L \to \gamma\gamma)$. Using the form factor from Equation 1.23, and recasting the expression in terms of x gives

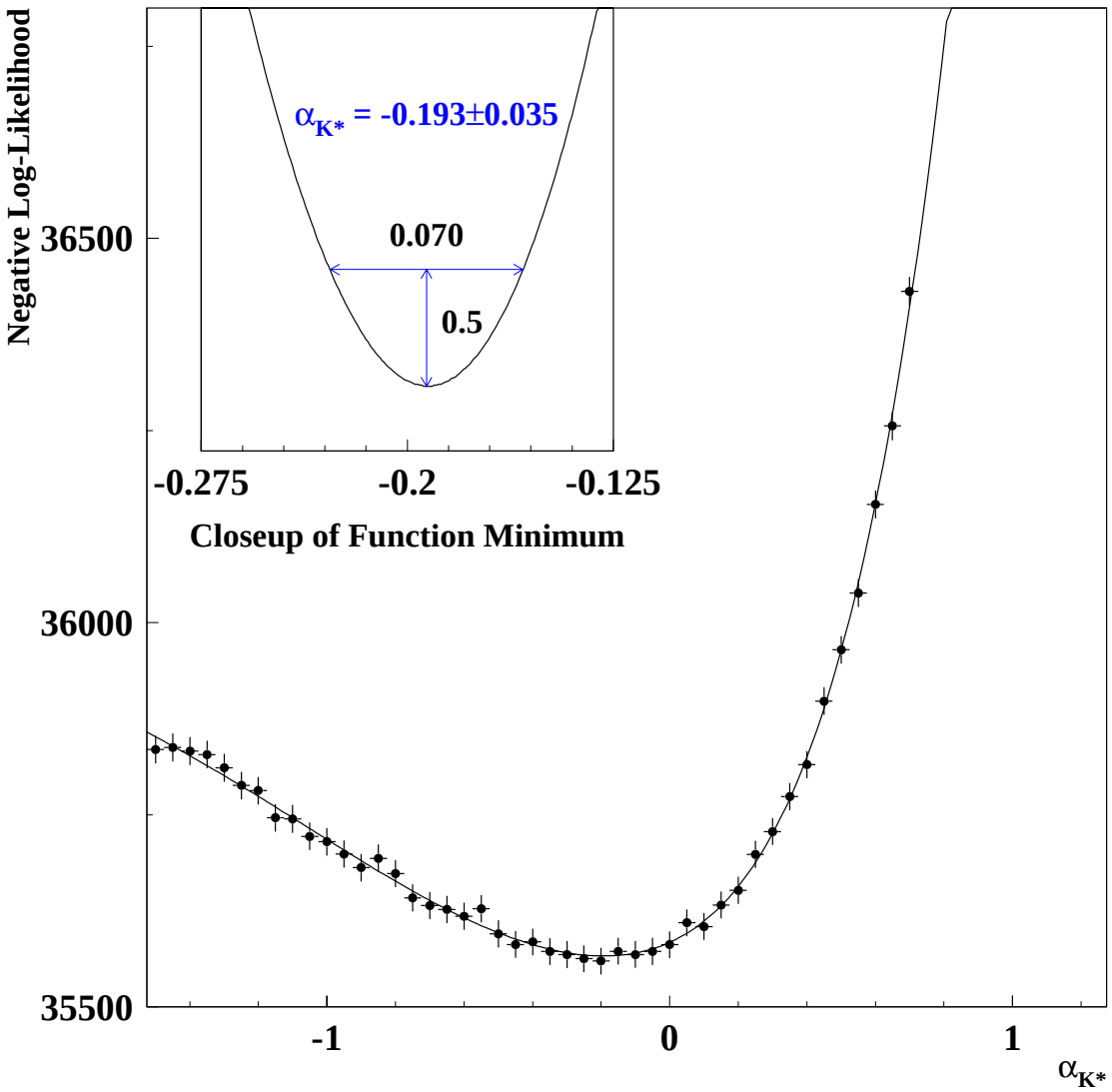

Figure 7.4: Measurement of α_{K^*} using a negative log-likelihood calculation to compare data to Monte Carlo. Monte Carlo was generated with values of α_{K^*} ranging from -1.5 to 0.7 in steps of 0.05.

$$, {}^{-1}_{\gamma\gamma}\frac{d,\,(K_L \to \mu^+\mu^-\gamma)}{dx} = \frac{2\alpha_{EM}}{3\pi}\frac{|f(x)|^2}{x}(1-x)^3\left(1+\frac{2m_\mu^2}{xm_K^2}\right)\left(1-\frac{4m_\mu^2}{xm_K^2}\right)^{\frac{1}{2}}, \quad (7.3)$$

with

$$f(x) = \frac{1}{1-x\,m_K^2/m_\rho^2} + \alpha_{K^*}\sqrt{8\pi\alpha_{EM}}G_{NL}f_{K^*K\gamma}\left(\frac{m_\rho^2}{f_{K^*}f_\rho^2 A_{\gamma\gamma}}\right)\left(\frac{1}{1-x\,m_K^2/m_{K^*}^2}\right)$$
$$\times\left(\frac{4}{3} - \frac{1}{1-x\,m_K^2/m_\rho^2} - \frac{1}{9}\frac{1}{1-x\,m_K^2/m_\omega^2} - \frac{2}{9}\frac{1}{1-x\,m_K^2/m_\phi^2}\right). \quad (7.4)$$

The expressions for the various coupling constants are

$$G_{NL} = 1.1\times 10^{-5}/m_p^2 = 1.2\times 10^{-5}\ \mathrm{GeV}^{-2}$$
$$f_{K^*K\gamma}^2 = \frac{96\pi,\,(K^*\to K^0\gamma)m_{K^*}^3}{(m_{K^*}^2 - m_K^2)^3} = (0.383\ \mathrm{GeV}^{-1})^2$$
$$f_\rho^2 = \frac{4\pi\alpha_{EM}^2 m_\rho}{3,\,(\rho\to e^+e^-)} = (5.04)^2$$
$$f_{K^*} = \frac{m_{K^*}}{m_\rho}f_\rho = 5.86$$
$$A_{\gamma\gamma}^2 = \frac{64\pi,\,(K_L\to\gamma\gamma)}{m_K^3} = (3.50\times 10^{-9}\ \mathrm{GeV}^{-1})^2$$
$$C = \sqrt{8\pi\alpha_{EM}}G_{NL}f_{K^*K\gamma}\left(\frac{m_\rho^2}{f_{K^*}f_\rho^2}A_{\gamma\gamma}\right) = 2.3.$$

It should be noted that the value calculated for C is somewhat different than the value of $C = 2.5$ used in Refs. [38, 39, 42, 43]. The evaluation of C in those references was based on information from the 1988 PDG [64]. Some experimental values have been updated since then, most notably the $K_L \to \gamma\gamma$ width. Furthermore, in each of those cases, the mass of the charged K^* was used; the neutral K^* mass is used here.

Substituting the values for the coupling constants and meson masses into Equation 7.4 yields the relation

$$f(x) = \frac{1}{1-0.418x} + \frac{2.3\alpha_{K^*}}{1-0.308x}\left(\frac{4}{3} - \frac{1}{1-0.418x} - \frac{1}{9}\frac{1}{1-0.405x} - \frac{2}{9}\frac{1}{1-0.238x}\right), \quad (7.5)$$

which is inserted into Equation 7.3. The differential decay rate can now be integrated

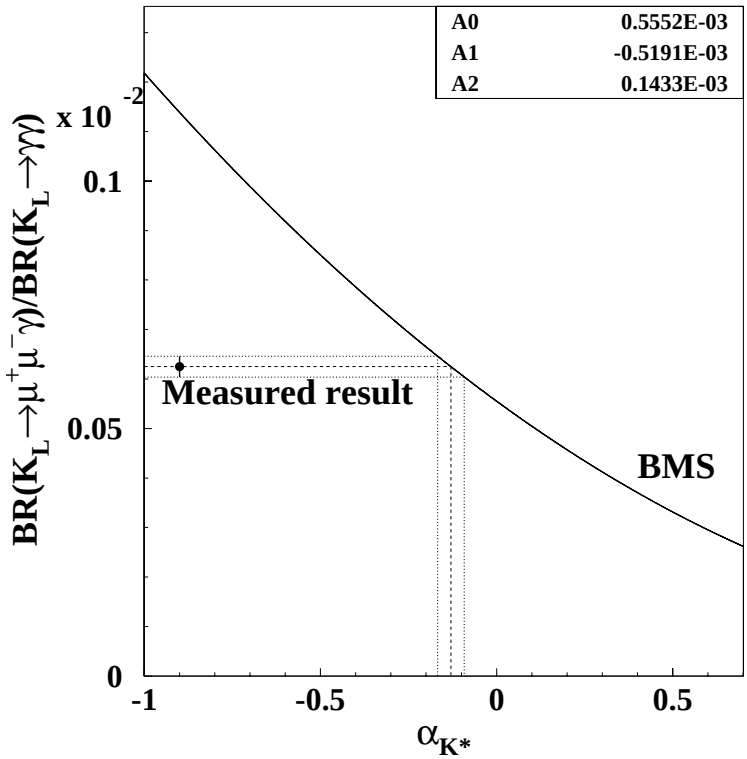

Figure 7.5: BMS model of the $K_L \to \mu^+\mu^-\gamma$ branching ratio as a function of α_{K^*}. The measured value of the $K_L \to \mu^+\mu^-\gamma$ branching ratio normalized to $K_L \to \gamma\gamma$ is shown with errors. The dashed and dotted lines extrapolate to the BMS curve to yield the measured value of α_{K^*}. The second order polynomial fit is drawn on top of the BMS curve.

over x to obtain the branching ratio relative to $K_L \to \gamma\gamma$, as a function of α_{K^*} only. The integration is performed numerically for values of α_{K^*} between -1.5 and 1.0. The result is shown in Figure 7.5, and is almost identical to the second-order polynomial

$$\frac{BR(K_L \to \mu^+\mu^-\gamma)}{BR(K_L \to \gamma\gamma)} = (14.33\alpha_{K^*}^2 - 51.91\alpha_{K^*} + 55.52) \times 10^{-5}. \tag{7.6}$$

The measured value $BR(K_L \to \mu^+\mu^-\gamma) = (3.70 \pm 0.08) \times 10^{-7}$ with $BR(K_L \to \gamma\gamma) = (5.92 \pm 0.15) \times 10^{-4}$ [20] gives the ratio

$$\frac{BR(K_L \to \mu^+\mu^-\gamma)}{BR(K_L \to \gamma\gamma)} = (62.5 \pm 2.1) \times 10^{-5}.$$

Using this result, α_{K^*} is determined both from Equation 7.6 and from the integrals directly, to be

$$\alpha_{K^*}^{BR} = -0.130^{+0.038}_{-0.037}.$$

7.2.3 Combined Result

The two independent measurements of α_{K^*} are in fairly good agreement with each other. The difference between $\alpha_{K^*}^{Shape}$ and $\alpha_{K^*}^{BR}$ is within 1.2σ of zero. A weighted average of the two values is taken using the PDG treatment for asymmetric Gaussian errors [20, page 10], leading to the final result

$$\alpha_{K^*} = -0.163^{+0.026}_{-0.027}.$$

7.3 α Measurement

The same analysis procedures used to measure α_{K^*} are followed for the measurement of the DIP form factor parameter α. The only difference is that here, the form factor in Equation 1.25 is used in place of the BMS parameterization.

7.3.1 $m_{\mu\mu}$ Shape Analysis

For this analysis, the negative log-likelihood method is used again. In this instance, $\mathcal{L}(\alpha)$ from Equation 7.2 is evaluated for 47 sets of Monte Carlo generated with the DIP form factor, with α ranging from -5.65 to 1.25 in steps of 0.15. The NLL distribution as a function of α is shown in Figure 7.6 and the minimum of that distribution is the measured value of α. The results, including the breakdown for the Winter and Summer data sets only, are

$$\alpha = -1.73 \pm 0.14.$$

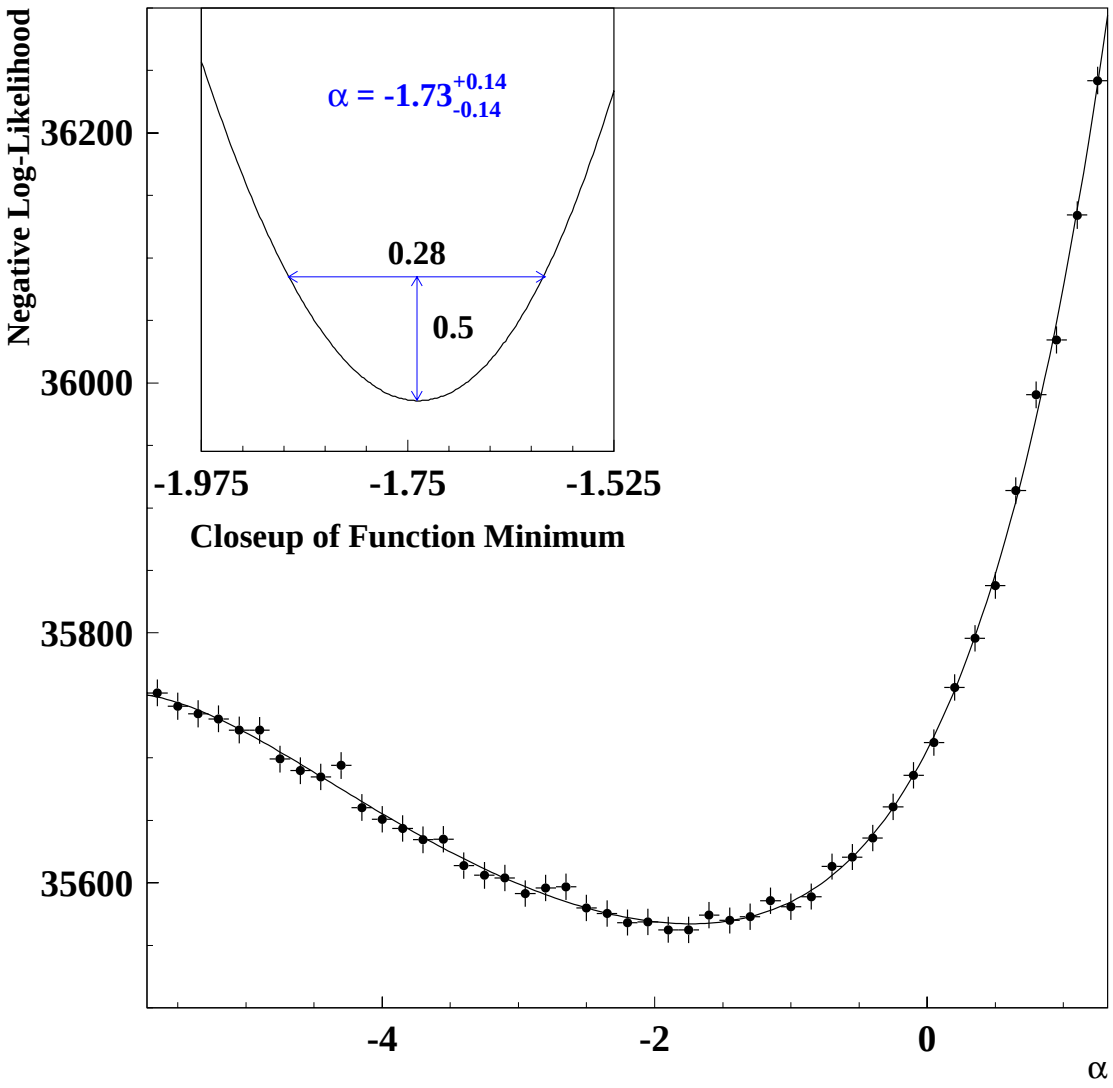

Figure 7.6: Measurement of α using a negative log-likelihood calculation to compare data to Monte Carlo. Monte Carlo was generated with values of α ranging from -5.65 to 1.25 in steps of 0.15.

$$\alpha^{Winter} = -1.66 \pm 0.18$$
$$\alpha^{Summer} = -1.99^{+0.22}_{-0.23}.$$

The difference between the Winter and Summer values is slightly greater than 1σ. There is a systematic error from the track momentum cut, similar to the one on α_{K^*}. As the cut is varied, shifts as large as -0.11 result. When this is added to the statistical error, the shape measurement of α becomes

$$\alpha^{Shape} = -1.73^{+0.14}_{-0.18}.$$

7.3.2 Branching Ratio Calculation

The expression for the differential decay used to find α_{K^*} is also used to extract α from the $K_L \to \mu^+\mu^-\gamma$ branching ratio. The form factor rewritten in terms of x is

$$f(x) = 1 + \alpha\left(\frac{x}{x - 2.40}\right), \tag{7.7}$$

and is substituted into Equation 7.3. Integrating the differential rate over x for a range of α produces the curve shown in Figure 7.7. There is no significant difference in the $K_L \to \mu^+\mu^-\gamma$ acceptance using the BMS form factor with $\alpha_{K^*} = -0.193$, or the DIP form factor with $\alpha = -1.73$. Therefore, $BR(K_L \to \mu^+\mu^-\gamma)/BR(K_L \to \gamma\gamma) = (62.5 \pm 2.1) \times 10^{-5}$ is still accurate and is evaluated against the DIP integral curve to obtain

$$\alpha^{BR} = -1.42 \pm 0.12.$$

7.3.3 Combined Result

Combining α^{Shape} and α^{BR} gives the weighted average

$$\alpha = -1.55 \pm 0.09.$$

The combined result is more than 1σ away from either of the two individual measurements, which are separated from each other by 1.6σ.

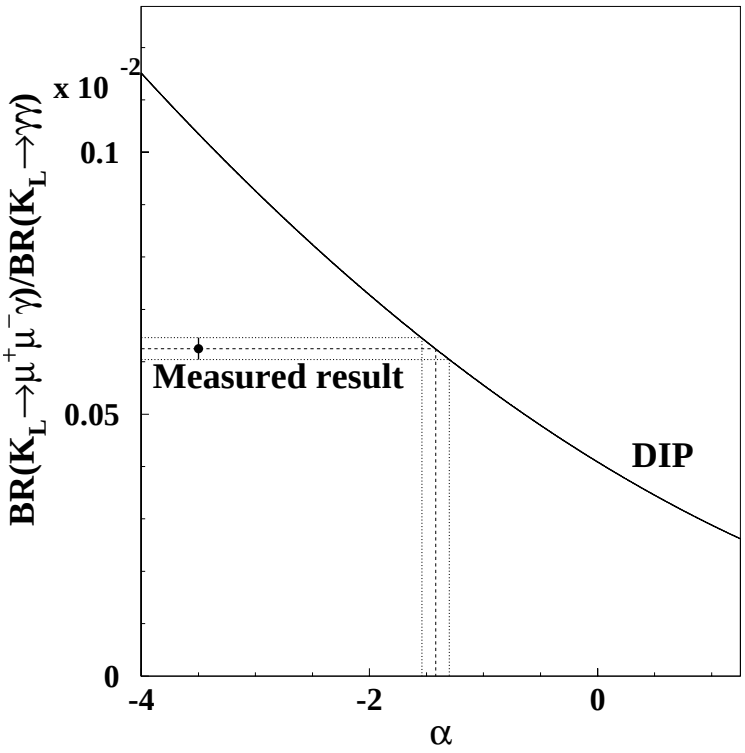

Figure 7.7: DIP model of the $K_L \to \mu^+\mu^-\gamma$ branching ratio as a function of α. The measured value of the $K_L \to \mu^+\mu^-\gamma$ branching ratio normalized to $K_L \to \gamma\gamma$ is shown with errors. The dashed and dotted lines extrapolate to the DIP curve to yield the measured value of α.

CHAPTER 8
CONCLUSION

We have measured the $K_L \to \mu^+\mu^-\gamma$ branching ratio based on a final data set of 9105 events with 2.4% estimated background. This sample is 45 times larger and has a background level two times lower than that of any previous measurement. From these events, we find the branching ratio to be

$$BR(K_L \to \mu^+\mu^-\gamma) \ = (3.70 \pm 0.04_{stat} \pm 0.07_{syst}) \times 10^{-7},$$

which is $\sim 1.5\sigma$ higher than the current world-average [20, page 461]. The total error on this result is 2.2% of the central value, making this the highest precision measurement of a rare kaon decay.[1]

The branching ratio can be compared to predictions from several models describing long distance physics. Figure 8.1 shows this measurement normalized to $K_L \to \gamma\gamma$, relative to the expectations from five different models. No model predicts the sign of α_{K^*}, and for this reason there are two entries in the plot for the Sakurai model. In principle, a prediction from the BMS model at $\alpha_{K^*} = +0.25$ should be included, but the negative value is unambiguously preferred based on its proximity to the measured value. There is no entry from the DIP parameterization because that model only puts forth a functional relation for the branching ratio relative to α; it does not make any *a priori* prediction for the value of α. The measured branching ratio is not particularly consistent with any of the models. However, the BMS model is much closer than any other, differing from the measurement by only 2.1σ. All other models shown in Figure 8.1 can be rejected based on inconsistency with the data.

[1] Only the all-pion and semi-leptonic kaon decays are known to greater precision.

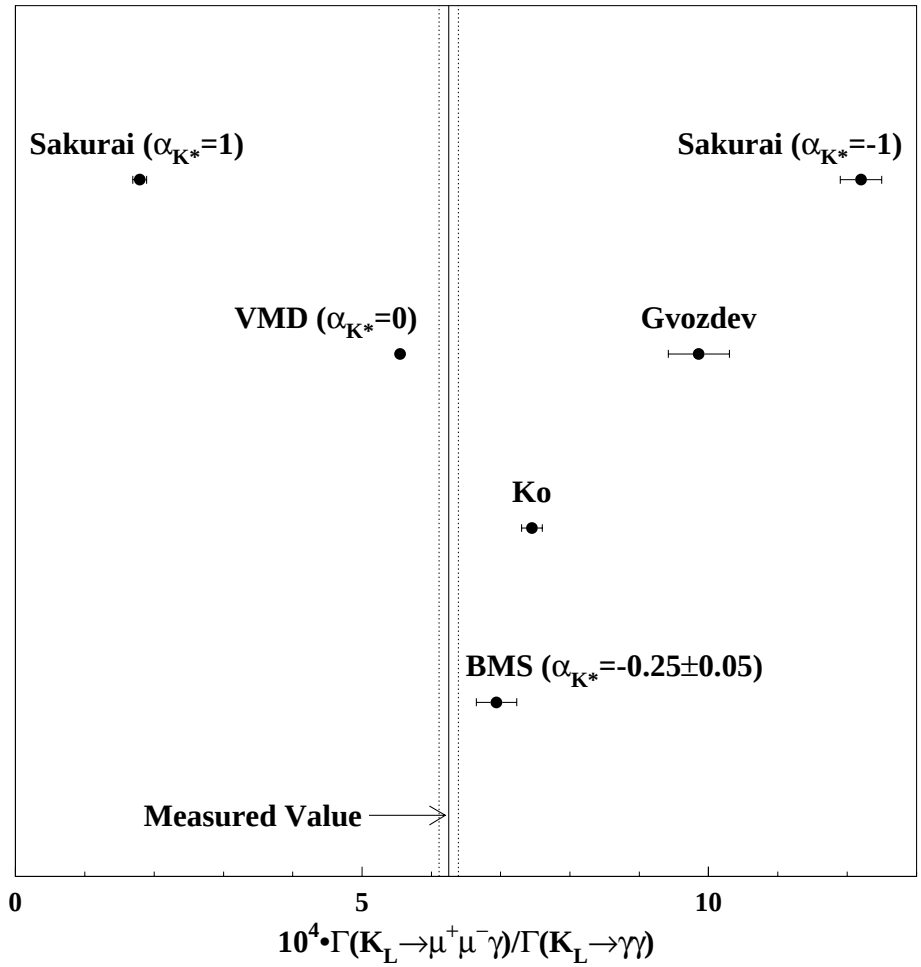

Figure 8.1: Comparison of the measured $BR(K_L \to \mu^+\mu^-\gamma)$ with various model predictions. The measured value and errors are represented by the solid and dotted lines respectively.

We have also measured parameters of the $K_L \to \mu^+\mu^-\gamma$ form factor. These include the familiar BMS α_{K^*} and the first experimental determination of α from the DIP parameterization. Each model's self-consistency can be evaluated by looking at the differences in the shape and branching ratio measurements of the form factor parameters. The two measurements for each parameter (α_{K^*} and α) are independent, and should produce the same result. The form factor measurements are summarized in Table 8.1. The BMS form factor results are slightly more consistent, although not nearly more so that a clear distinction between models can be made.

Table 8.1: Self-consistency check of form factor parameter measurements. The differences between shape and branching ratio measurements of the form factor parameters are expressed in σ.

Form Factor Model	Parameter Measurement Shape	BR	Significance of Difference
α_{K^*} (BMS)	$-0.193^{+0.035}_{-0.049}$	$-0.130^{+0.038}_{-0.037}$	1.2σ
α (DIP)	$-1.73^{+0.14}_{-0.18}$	-1.42 ± 0.12	1.6σ

Using Equation 1.26, we calculate α from α_{K^*}, and find

$$\alpha(\alpha_{K^*} = -0.163) = -1.51 \pm 0.11$$

which is quite consistent with the measured value.

KTeV is the latest of several experiments that have measured α_{K^*}. This measurement of α_{K^*} is compared to those from the previous experiments in Figure 8.2. We have measured the BMS form factor with better than twice the precision of the latest $K_L \to e^+e^-\gamma$ result. This, of course, is due to the higher proportion of $K_L \to \mu^+\mu^-\gamma$ data at high x, coupled with the ability to extract α_{K^*} from the precisely measured branching ratio. An obvious feature of Figure 8.2 is that the difference between $K_L \to e^+e^-\gamma$ and $K_L \to \mu^+\mu^-\gamma$ results hinted at by the early experiments is now established at the 3σ level. According to Ref. [30], the higher $K_L \to \mu^+\mu^-\gamma$ results for α_{K^*} are to be expected if the form factor is actually of a much simpler, linear form. This difference, together with the fact that $BR(K_L \to \mu^+\mu^-\gamma)$ is not predicted particularly well by the BMS model, begins to cast doubt on the effectiveness of the model.

We now turn to extracting short distance physics from $K_L \to \mu^+\mu^-\gamma$. Limits on the CKM parameter ρ can be determined starting with Equation 1.15. To evaluate the limit in the BMS model, a bound on $|\mathrm{Re}A_{LD}|$ is calculated by inserting $\alpha_{K^*} =$

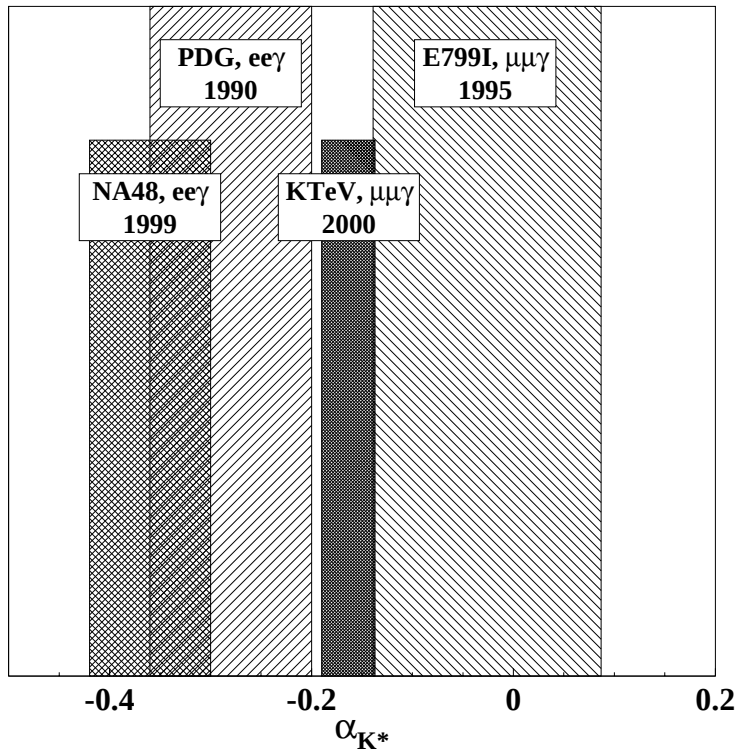

Figure 8.2: Comparison of the α_{K^*} measurement with previous experiments.

$-0.163^{+0.026}_{-0.027}$ into Equation 1.24. This produces the upper bound

$$|\mathrm{Re}A_{LD}| < 3.6 \times 10^{-5},$$

which is added algebraically to the limit on $|\mathrm{Re}A_{exp}|$ given in Equation 1.5. With this information, we evaluate the 90% confidence limits

$$\begin{aligned} \overline{\rho} &> -1.0 \\ \rho &> -1.0, \end{aligned} \tag{8.1}$$

where we have used $\overline{m}_t(m_t) = 165 \pm 5$ GeV, $|V_{cb}| = 0.0395 \pm 0.0017$, and $\lambda = 0.2195 \pm 0.0023$. This constraint is of little use since the combined limit from $|V_{ub}|$, B mixing, and the CP-violating parameter ϵ is $\rho > 0$ [65]. A limit of $\rho > -0.58$

from $K_L \to \mu^+\mu^-$ was obtained in Ref. [23] using $\alpha_{K^*} = -0.25 \pm 0.07$. Most of the difference from our result is due to that lower value of α_{K^*}.

A different process is used to obtain the ρ limit from the DIP parameterization. The perturbative QCD bound in Equation 1.28 is converted to a Gaussian distribution centered on 0, with a width equal to that of a flat distribution from -0.4 to 0.4, (i.e. $\sigma = 0.8/\sqrt{12}$) [27]. This is then added to the Gaussian distribution $\alpha = -1.55 \pm 0.09$ and substituted into Equation 1.27 to yield

$$\begin{aligned} |\mathrm{Re}A_{LD}| &= 1.61 \times 10^{-5}|0.13 \pm 0.31 \pm 0.69| \\ &= (0.21 \pm 1.22) \times 10^{-5} \\ &< 2.14 \times 10^{-5}\ (90\%\ \mathrm{C.L.}), \end{aligned} \tag{8.2}$$

where the Bayesian approach to unphysical regions has again been employed. Note that the error is now completely dominated by the theoretical uncertainty.

To derive a limit on the sum $|\mathrm{Re}A_{exp}| + |\mathrm{Re}A_{LD}|$, the measurements from Equation 1.5 and Equation 8.2 are converted into Gaussian distributions. A 2-dimensional probability distribution is then constructed by taking the product of the two Gaussians. Finally, a calculation is made to determine the contour of constant $|\mathrm{Re}A_{exp}| + |\mathrm{Re}A_{LD}|$ under which is contained 90% of the probability in the physical (positive) quadrant. The probabilty distributions are shown in Figure 8.3, and the 90% confidence limit is found to be

$$\begin{aligned} \overline{\rho} &> -0.2 \\ \rho &> -0.2. \end{aligned} \tag{8.3}$$

This limit is a large improvement over the result using BMS. It is also a significant improvement over the DIP limits reported in Refs. [21, 27], primarily because the experimental error in Equation 8.2 has been reduced by more than a factor of 2. Furthermore, this lower limit approaches that found from the combination of all other channels.

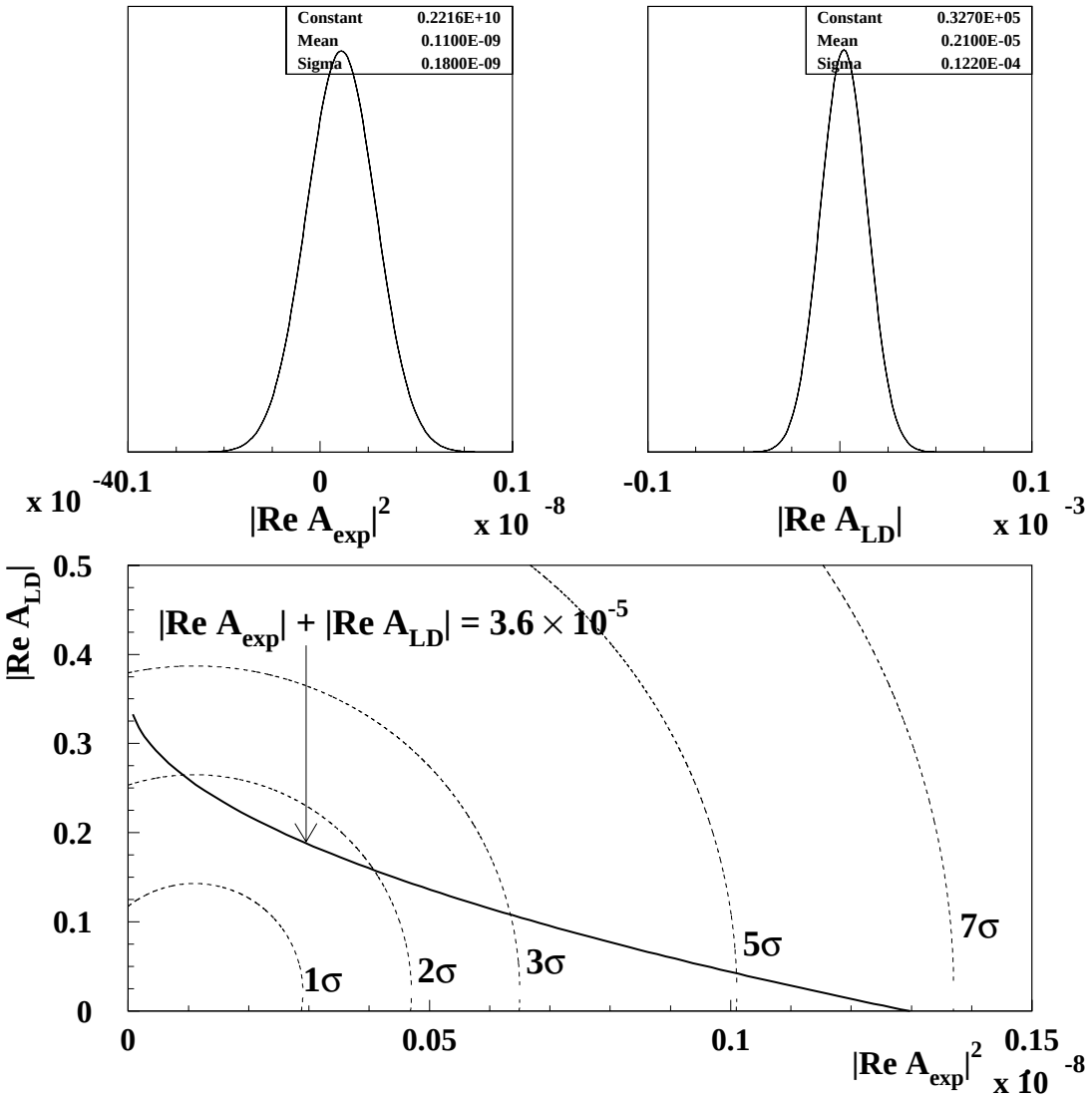

Figure 8.3: Probability distributions for $|\mathrm{Re}A_{exp}|^2$ and $|\mathrm{Re}A_{LD}|$. The upper plots show the Gaussian probability distribution for each measurement. The lower plot shows the 2-dimensional probability distribution. The dashed curves represent selected error ellipses from 1σ out to 7σ. The solid curve is the contour of constant $|\mathrm{Re}A_{exp}| + |\mathrm{Re}A_{LD}|$ under which is contained 90% of the integrated probability in the physical (positive) quadrant.

A limited amount of progress can be expected in reducing experimental errors in the above calculations. After this $K_L \to \mu^+\mu^-\gamma$ measurement and the recent E871 result for $K_L \to \mu^+\mu^-$, the most likely improvement would be a better measurement of $K_L \to \gamma\gamma$ [5]. Beyond that, any further progress would be dependent on gaining better control of theoretical uncertainties in long distance physics. The BMS model is beginning to exhibit deficiencies with regard to obtaining useful limits on ρ and consistent results between $K_L \to e^+e^-\gamma$ and $K_L \to \mu^+\mu^-\gamma$. Therefore, the DIP parameterization remains as the most attractive option available for guiding future studies in this area. The most important such study is a consistency check utilizing the $K_L \to e^+e^-\gamma$ decay mode, which is currently being conducted by KTeV.

APPENDIX A
APPENDIX

A recurring complaint among the theoretical community has been that all measurements of long distance form factors to date have been analyzed according to the BMS model only. This has made evaluation or development of other models difficult because all comparison with data had to be made through the BMS model. Here we present the (nearly) raw (almost) completely model-independent data from which the measurements in this dissertation were made.

In Table A.1, the acceptance corrected x distribution data is listed. The number of events in each 0.01 unit bin are given along with the statistical errors. This is the data represented by the upper plot in Figure 7.2. The bin-by-bin acceptance has been calculated using Monte Carlo generated with the BMS form factor. However, the Monte Carlo acceptance is largely model-insensitive especially on a bin-by-bin basis with such small binning. Therefore, this distribution accurately reflects the differential decay spectrum of $K_L \to \mu^+\mu^-\gamma$.

The bin-by-bin measurement of the form factor $f(x)$ is given with errors in Table A.2. This distribution was obtained by dividng the x distribution from data, by the x distribution from Monte Carlo generated with the form factor, $f(x) = 1$. This gives a direct measurement of the form factor as described in Section 7.1, and shown in Figure 7.1.

Table A.1: Bin-by-bin acceptance corrected x distribution.

x	N (events)	ΔN	x	N (events)	ΔN	x	N (events)	ΔN	x	N (events)	ΔN
.155	0	0	.355	3648	233	.555	956	112	.755	105	46
.165	0	0	.365	3103	213	.565	1002	113	.765	150	56
.175	0	0	.375	2948	203	.575	814	103	.775	92	46
.185	4656	465	.385	3123	210	.585	747	98	.785	141	59
.195	6597	419	.395	2539	188	.595	613	93	.795	97	51
.205	6480	370	.405	2394	181	.605	797	105	.805	136	62
.215	6284	348	.415	2374	181	.615	796	106	.815	55	41
.225	7052	366	.425	2334	179	.625	601	92	.825	−55	42
.235	6619	351	.435	2428	184	.635	458	82	.835	37	38
.245	6194	336	.445	2030	165	.645	561	91	.845	−29	37
.255	6198	337	.455	1845	158	.655	424	81	.855	163	86
.265	5369	308	.465	1931	163	.665	432	80	.865	620	384
.275	5129	294	.475	1593	147	.675	366	77	.875	−20	49
.285	4652	280	.485	1791	153	.685	437	85	.885	−253	234
.295	5031	289	.495	1556	144	.695	236	60	.895	480	792
.305	4352	265	.505	1542	143	.705	145	50	.905	0	0
.315	4160	258	.515	1239	127	.715	107	42	.915	0	0
.325	3837	243	.525	1219	125	.725	195	59	.925	0	0
.335	3454	227	.535	1054	118	.735	149	51	.935	0	0
.345	3455	226	.545	1190	126	.745	66	35	.945	0	0

Table A.2: Bin-by-bin form factor measurement.

x	$\lvert f(x)\rvert^2$	$\Delta\lvert f(x)\rvert^2$	x	$\lvert f(x)\rvert^2$	$\Delta\lvert f(x)\rvert^2$	x	$\lvert f(x)\rvert^2$	$\Delta\lvert f(x)\rvert^2$	x	$\lvert f(x)\rvert^2$	$\Delta\lvert f(x)\rvert^2$
.155	0.000	0.000	.355	1.111	0.071	.555	1.392	0.165	.755	1.304	0.589
.165	0.000	0.000	.365	1.011	0.069	.565	1.580	0.182	.765	2.113	0.811
.175	0.000	0.000	.375	1.063	0.073	.575	1.345	0.174	.775	1.492	0.756
.185	0.553	0.054	.385	1.149	0.077	.585	1.374	0.183	.785	2.581	1.131
.195	0.828	0.051	.395	1.024	0.076	.595	1.186	0.183	.795	2.079	1.118
.205	0.824	0.046	.405	1.033	0.078	.605	1.621	0.217	.805	3.838	1.860
.215	0.810	0.044	.415	1.096	0.084	.615	1.851	0.252	.815	1.821	1.386
.225	0.925	0.047	.425	1.127	0.087	.625	1.607	0.250	.825	−1.615	1.276
.235	0.899	0.047	.435	1.306	0.100	.635	1.232	0.224	.835	1.412	1.496
.245	0.890	0.047	.445	1.170	0.096	.645	1.739	0.287	.845	−1.475	1.894
.255	0.928	0.050	.455	1.135	0.098	.655	1.408	0.272	.855	12.745	8.047
.265	0.867	0.049	.465	1.231	0.104	.665	1.654	0.313	.865	87.654	98.910
.275	0.905	0.051	.475	1.155	0.107	.675	1.431	0.307	.875	−3.668	9.586
.285	0.847	0.050	.485	1.386	0.120	.685	2.062	0.412	.885	0.000	0.000
.295	0.990	0.056	.495	1.324	0.123	.695	1.629	0.428	.895	4.122	5.920
.305	0.901	0.054	.505	1.446	0.136	.705	0.911	0.316	.905	0.000	0.000
.315	0.921	0.057	.515	1.273	0.132	.715	0.706	0.278	.915	0.000	0.000
.325	0.939	0.059	.525	1.305	0.136	.725	1.479	0.455	.925	0.000	0.000
.335	0.935	0.061	.535	1.256	0.142	.735	1.530	0.541	.935	0.000	0.000
.345	0.958	0.063	.545	1.536	0.166	.745	0.651	0.353	.945	0.000	0.000

REFERENCES

[1] A. J. Buras. "CP violation and rare decays of K and B meson", Lectures given at Lake Louise Winter Institute: Electroweak Physics, Lake Louise, Alberta, Canada, hep-ph/9905437, (1999).

[2] G. Isidori. "Standard model vs new physics in rare kaon decays", Presented at 1999 Chicago Conference on Kaon Physics (K 99), Chicago, IL, hep-ph/9908399, (1999).

[3] G. Buchalla. "Theoretical update on rare K decays", Invited talk at 3rd Workshop on Physics and Detectors for DAPHNE (DAPHNE 99), Frascati, Italy, 16-19 Nov 1999, hep-ph/0002207, (2000).

[4] G. D'Ambrosio. "Theory of rare kaon decays", Invited plenary talk at the 3rd International Conference on B Physics and CP Violation (BCONF99), Taipei, Taiwan, hep-ph/0002254, (2000).

[5] S. Kettell. "Experimental results on radiative kaon decays", To be published in the proceedings of 3rd Workshop on Physics and Detectors for DAPHNE (DAPHNE 99), Frascati, Italy, hep-ex/0002009, (2000).

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

[7] D. V. Nanopoulos and G. G. Ross, *Phys. Lett.* **56B**, 279 (1975).

[8] M. B. Voloshin and E. P. Shabalin, *JETP Lett.* **23**, 107 (1976).

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

[10] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973).

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

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

[13] J. L. Ritchie and S. G. Wojcicki, *Rev. Mod. Phys.* **65**, 1149 (1993).

[14] G. Buchalla and A. J. Buras, *Nucl. Phys.* **B412**, 106 (1994).

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

[16] S. Adler et al., Preprint hep-ex/0002015, BNL E787, (2000).

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

[18] C. Quigg and J. D. Jackson, Preprint 18487, UCRL, unpublished, (1968).

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

[20] Particle Data Group, C. Caso et al., *Eur. Phys. J. C* **3**, 458 (1998).

[21] D. Ambrose et al., *Phys. Rev. Lett.* **84**, 1389 (2000).

[22] Particle Data Group, R. M. Barnett et al., *Phys. Rev. D* **54**, 164 (1996).

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

[24] G. Buchalla and A. J. Buras, *Nucl. Phys.* **B400**, 225 (1993).

[25] A. J. Buras and R. Fleischer, in *Heavy Flavours II*, edited by A. J. Buras and M. Lindner, page 65. World Scientific, (1997).

[26] L. I. Amettler, A. Bramon, and E. Massó, *Phys. Rev. D* **30**, 251 (1984).

[27] G. D'Ambrosio, G. Isidori, and J. Portolés, *Phys. Lett. B* **423**, 385 (1998).

[28] A. Lath, in *DPF 99 Proceedings of the Los Angeles Meeting*, edited by K. Arisaka and Z. Bern, pages 4–06. University of California, Los Angeles, http://www.dpf99.library.ucla.edu, (1999).

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

[30] G. D'Ambrosio and J. Portolés, *Nucl. Phys.* **B492**, 417 (1997).

[31] L. Bergström and E. Massó and P. Singer, *Phys. Lett.* **131B**, 229 (1983).

[32] R. F. Sarraga and H. J. Munczek, *Phys. Rev. D* **4**, 2884 (1971).

[33] J. J. Sakurai, *Phys. Rev.* **156**, 1508 (1967).

[34] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Nucl. Phys.* **B120**, 316 (1977).

[35] L. Bergström and E. Massó and P. Singer and D. Wyler, *Phys. Lett.* **134B**, 373 (1984).

[36] P. Ko, *Phys. Rev. D* **44**, 139 (1990).

[37] A. A. Gvozdev, N. V. Mikheev, and L. A. Vassilevskaya, *Phys. Lett. B* **292**, 176 (1992).

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

[39] V. Fanti et al., *Z. Phys.* **C76**, 653 (1997).

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

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

[42] V. Fanti et al., *Phys. Lett. B* **458**, 553 (1999).

[43] M. B. Spencer et al., *Phys. Rev. Lett.* **74**, 3323 (1995).

[44] M. B. Spencer, Ph.D. thesis (UCLA, May, 1995).

[45] P. S. Shawhan, Ph.D. thesis (The University of Chicago, December, 1999).

[46] L. K. Gibbons, Ph.D. thesis (The University of Chicago, August, 1993).

[47] R. A. Briere, Ph.D. thesis (The University of Chicago, June, 1995).

[48] K. S. McFarland, Ph.D. thesis (The University of Chicago, March, 1994).

[49] K. Hanagaki, Ph.D. thesis (Osaka University, September, 1998).

[50] G. E. Graham, Ph.D. thesis (The University of Chicago, August, 1999).

[51] E. D. Zimmerman, Ph.D. thesis (The University of Chicago, March, 1999).

[52] C. Bown et al., *Nucl. Instrum. Methods Phys. Res., Sect. A* **369**, 248 (1996).

[53] A. J. Malensek, Preprint FN-341, FN-341A (errata), Fermilab, (1981).

[54] H. W. Atherton et al., Preprint 80-07, CERN, (1980).

[55] K. O. Mikaelian and J. S. Smith, *Phys. Rev. D* **5**, 1763 (1972).

[56] K. O. Mikaelian and J. S. Smith, *Phys. Rev. D* **5**, 2891 (1972).

[57] M. V. Chizhov, *Phys. Lett. B* **381**, 359 (1996).

[58] R. Brun et al., Computer code GEANT 3.15, CERN, Geneva (1991).

[59] R. Brun et al., *GEANT User's Guide.* CERN, Geneva (1985).

[60] W. Lohmann, R. Kopp, and R. Voss, Preprint 85-03, CERN, (1985).

[61] M. Sadamoto, Ph.D. thesis (Osaka University, February, 1999).

[62] R. Brun et al., Computer code GEANT 3.21, CERN, Geneva (1994).

[63] G. Gershwin and I. Gershwin, 'Porgy and Bess' (1934).

[64] Particle Data Group, G. P. Yost et al., *Phys. Lett. B* **204**, (1988).

[65] A. Ali and D. London. "CP violation and quark mixing", To be published in the proceedings of 3rd Workshop on Physics and Detectors for DAPHNE (DAPHNE 99), Frascati, Italy, hep-ex/0002167, (2000).