A Study of the Decay $\pi^0 \to e^+e^-e^+e^-$ Using $K_L \to \pi^0\pi^0\pi^0$ Decays in Flight

by

Patrick Alan Toale

B.A., University of Colorado, 1998

M.S., University of Colorado, 2001

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Physics

2004

This thesis entitled:
A Study of the Decay $\pi^0 \to e^+e^-e^+e^-$ Using $K_L \to \pi^0\pi^0\pi^0$ Decays in Flight
written by Patrick Alan Toale
has been approved for the Department of Physics

Eric D. Zimmerman

John Cumalat

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Toale, Patrick Alan (Ph. D., Physics)

A Study of the Decay $\pi^0 \to e^+e^-e^+e^-$ Using $K_L \to \pi^0\pi^0\pi^0$ Decays in Flight

Thesis directed by Assistant Professor Eric D. Zimmerman

The rare double Dalitz decay of the neutral pion proceeds through a two-photon intermediate state, $\pi^0 \to \gamma^*\gamma^* \to e^+e^-e^+e^-$, which provides sensitivity to the physics of the $\pi^0\gamma^*\gamma^*$ coupling. The properties of this interaction are a consequence of the internal structure of mesons in general. The spatial extent of the quark-antiquark pair becomes manifest as a form factor, dependent on the momentum transfer (q^2) of the virtual photons. The double Dalitz decay is sensitive to this electromagnetic form factor in the range of q^2 from $4m_e^2$ to $M_{\pi^0}^2$. In addition, the correlation between the planes of polarization of the two photons provides a measure of the parity of the pion. This correlation is preserved in the angle between the two e^+e^- pairs. Since the pion is a pseudoscalar particle, this angular correlation can be used to limit the amount of *CP* violation at the $\pi^0\gamma^*\gamma^*$ vertex.

This thesis presents an analysis of the double Dalitz decay using $K_L \to \pi^0\pi^0\pi^0$ decays in flight. Based on a sample of 30511 of these events with less than 1 % background, the branching ratio relative to the common two-photon mode was found to be $B(4e)/B(\gamma\gamma) = (3.30 \pm 0.18) \times 10^{-5}$. These same events were used to extract the momentum dependence of the form factor, parameterized as $\alpha = 1.3 \pm 1.3$, and to set a limit on the possible CP violating component of the coupling, which in terms of a mixing angle is found to be limited to $\zeta < 1.72^\circ$ at the 90 % CL. Finally, a preliminary measurement of the branching ratio of the related radiative decay, $\pi^0 \to e^+e^-e^+e^-\gamma$, has been performed and found to be $B(4e\gamma)/B(\gamma\gamma) = (1.68 \pm 0.27) \times 10^{-6}$ based on 425 candidate events with 25 % background.

Dedication

To my wife and son. Without you I would have nothing.

Acknowledgements

I began working on KTeV in 1997 at which point the first data-taking run was well underway. Some people had already been working to bringing KTeV to life for over five years. I never met some of those people but I am grateful for their early contributions to the design and construction of a truly remarkable experiment. The work presented here really began with their foresight and ingenuity. During the seven years in which I have participated on KTeV, I have been privileged to work with an exceptional group of people. While I cannot acknowledge everyone who helped me along the way, there are a few who deserve special recognition.

I would not have survived the summer of 1997, my first contact with KTeV and Tony Barker, had it not been for Gib Lichstein who suffered under the wrath of the mallet alongside me. Gib's time on KTeV was short, but he has remained a good friend and confidant.

Peter Mikelsons and Jason LaDue are due a great deal of credit for laying the foundations of the Colorado group. Peter joined the group sometime in 1992 and was an invaluable source of information on the early phase of KTeV. I benefited directly from many of Jason's efforts; in exchange for which, he was freed from calculating radiative corrections. This thesis borrows heavily from Jason's own, a fact he kindly ignored while gratefully critiquing this work.

Jim Hirschauer gets two thanks. First, for providing both the Riv and the whiffle balls during a long, hot summer in Batavia. And second, for asking the questions that

remind me how complex the world can be.

Mike Wilking knows more about physics than any chemical engineer I have ever met. While his taste in sports is questionable, his since of truth and beauty is inspiring. Long live Poseidon.

Rune Niclasen is likely the last KTeV student at Colorado. More than anything else, I thank Rune for keeping me honest through his own stubborn questioning. One thing is obvious, Rune knows what he knows. I feel I must also apologize for leaving Rune the legacy of 10 years of KTeV at Colorado. Good luck.

The theoretical analysis presented in this thesis is due largely to the work of Hanwen Huang. I cannot thank Hanwen enough for saving me from becoming a theorist.

I must thank Eric Zimmerman for stepping up and providing focus in a time of loss and confusion. Eric has graciously adopted and cared for the remains of the KTeV group. It is funny how life works; I knew Eric briefly while he was a KTeV graduate student at the University of Chicago and now he is my advisor.

Last, but never least, is Tony Barker. Tony's greatest quality was his loyalty to his friends. Thank you Tony for having faith in me; it has enabled me to have confidence in myself.

Contents

Tables

Table

Figures

Figure

Chapter 1

Introduction

Then anyone who leaves behind him a written manual,
and likewise anyone who receives it,
in the belief that such writing will be clear and certain,
must be exceedingly simple-minded.
– Plato

The π-meson, or pion, has been an invaluable tool for unraveling the mysteries of the fundamental interactions of physics since its discovery over fifty years ago. Its existence was first postulated by Yukawa [1] in 1935 in an attempt to develop a "quanta" of the nuclear force. Shortly after Yukawa's proposal, two candidate particles were observed in cosmic rays experiments. It was not until 1947 that the identities of what are now known as the pion and the muon were finally established [2]. Three years later, the neutral pion was discovered at the Berkeley synchrotron [3]. This experiment observed pairs of photons consistent with the interpretation that they were produced by $\pi^0 \to \gamma\gamma$ decays.

The existence of the two–photon decay channel provides a means for studying the nature of the pion under the discrete symmetries of charge conjugation (C), parity (P), and time reversal (T). As pointed out by Yang [4] in 1950, the existence of the two–photon decay mode leads immediately to the conclusion that the π^0 has zero intrinsic spin. Additionally, the angle between the two polarization planes will have different correlations depending on whether the pion is a scalar (even under P) or a pseudoscalar

(odd under P) particle. Yang proposed an experiment to extract the parity of the π^0 by measuring the polarization of the two photons. His idea was expanded to include sensitivity to both C and T by Bernstein and Michel [5].

During these same years, an alternative decay mode of the neutral pion was predicted and observed. In a letter published in 1951 [6], Dalitz calculated the rate for $\pi^0 \to e^+e^-\gamma$ assuming that it proceeded through a two–photon decay in which one of the photons was virtual and converted internally into an electron-positron pair. The experimental evidence of this decay process, now known simply as a Dalitz decay, was first seen in emulsion plates exposed to the Chicago cyclotron in 1952 [7]. A number of experiments performed over the next 10 years verified Dalitz's hypothesis that the $\pi^0 \to e^+e^-\gamma$ decay resulted from internal conversion of a virtual photon [8, 9, 10].

It was soon noted that if one of the two photons could convert to an e^+e^- pair, then so could both. The resulting process, $\pi^0 \to e^+e^-e^+e^-$, is known as the double Dalitz decay. The pioneering work on this mode was published in 1955 by Kroll and Wada [11]. They worked out the differential decay rate neglecting contributions from the exchange of identical particles. They also made the first connection between the orientation of the planes formed by the two Dalitz pairs and the polarization of the intermediate photons, suggesting that the symmetries discussed previously could be extracted from the $\pi^0 \to e^+e^-e^+e^-$ final state. An experiment was soon performed by Samios at the Nevis cyclotron in 1962 [12]. Neutral pions were produced in a bubble chamber in which their decays were photographed. Some 835 thousand photos were scanned by hand to find 206 double Dalitz candidates. The branching ratio of the decay was measured to 10 % and was in complete agreement with the prediction of Kroll and Wada. The angular distribution was also studied and was found to rule out a scalar pion at $3.6\,\sigma$. A search for T-violation was inconclusive.

Today, the Samios measurements are still the only published experimental results on the double Dalitz decay of the π^0.[1] There have, however, been two important theoretical advancements made over the years. In 1973, Miyazaki and Takasugi [14] repeated the Kroll-Wada analysis, but included the contributions from the exchange diagram. More recently, the Colorado group of the KTeV experiment published a complete description of the decay process, including the effects of first-order radiative corrections [15].

The analysis presented in this thesis can be viewed as a modern rendition of the Samios experiment. The goal is to measure the branching ratio, including radiative effects, to around 1 % and to test for both CPT-invariance and CP-violation in π^0 decays. To do this, the KTeV experiment at Fermilab collected a large number of $K_L \to \pi^0\pi^0\pi^0$ events in which one of the pions decays as $\pi^0 \to e^+e^-e^+e^-$. The final data sample is roughly 30,000 events.

The theoretical background of the measurements is outlined in Chapter 2. Then a detail description of the KTeV experiment, with emphasis given to areas relevant to this analysis, is presented in Chapter 3. Next, the Monte Carlo program used to model the theory and to calculate the acceptance of the detector as well as selection cuts, is summarized in Chapter 4. Chapter 5 explains the general reconstruction techniques that turn the raw binary data into physically meaningful quantities. The actual results are then presented in the last three chapters: the $\pi^0 \to e^+e^-e^+e^-$ branching ratio analysis in Chapter 6, the CP and form factor analysis in Chapter 7, and the radiative $\pi^0 \to e^+e^-e^+e^-\gamma$ branching ratio in Chapter 8.

[1] We believe that they may be the oldest results in the Particle Data Group's review of particle properties (PDG) [13].

Chapter 2

Theory of π^0 Decays

The neutral pion interacts strongly with other hadrons, however since it is the lightest hadron it can not decay strongly. Weak decays are suppressed if not forbidden entirely, leaving just the electromagnetic channel. Since the π^0 at rest has zero angular momentum, the decay into a single photon is not allowed. However, two photons, as a system, may have zero angular momentum depending on the spin of each photon. As it turns out, the two–photon decay accounts for 98.798% of the π^0 width. All of the other observed decays can be successfully described as a two–photon decay with one or both of the photons off–shell and internally converting to an electron–positron pair. Therefore, many of the properties of the rare π^0 decays are inherited from the common two–photon decay. In this chapter, I will present a detailed analysis of the two–photon decay which will result in a general model of the $\pi^0\gamma^*\gamma^*$. I will then show how measurements of rare decay modes further constrain the model. Much of the work presented in this chapter has been previously published by Barker, *et al* [15].

2.1 The $\pi^0\gamma^*\gamma^*$ Vertex

Although the $\pi^0 \to \gamma\gamma$ decay is electromagnetic in nature, it is not a fundamental coupling in QED. The simple picture of the quark and anti-quark annihilating is complicated by the details of the hadronic bound state. Even so, there are a remarkable number of powerful constraints which, along with some mild assumptions, allow us to

express the coupling in terms of a few free parameters.

To begin with, the pion is a spin–0 particle associated with the scalar field Π and the photon is the spin–1 quanta of the electromagnetic field, represented by the field tensor $F^{\mu\nu}$. Lorentz invariance of the interaction requires the Lagrangian to be a scalar. The most general form of the interaction is therefore

$$\mathcal{L} = C_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} \Pi, \tag{2.1}$$

where $C_{\mu\nu\rho\sigma}$ is a tensor that couples the pion field to the two photon fields. There are two allowed forms of the coupling, one appropriate to a scalar pion and another for a pseudoscalar pion. This can be seen by appealing to Jackson's treatment of classical electromagnetic waves [16]. Consider the two Lorentz invariant combinations of the classical electromagnetic field, $\boldsymbol{E} \cdot \boldsymbol{B}$ and $\boldsymbol{E}^2 - \boldsymbol{B}^2$. In terms of the field tensors and the coupling, these field combinations correspond to

$$\boldsymbol{E} \cdot \boldsymbol{B} = \frac{1}{4} \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma}, \tag{2.2a}$$

$$\boldsymbol{E}^2 - \boldsymbol{B}^2 = \frac{1}{4} \left(g_{\mu\sigma} g_{\nu\rho} - g_{\mu\rho} g_{\nu\sigma} \right) F^{\mu\nu} F^{\rho\sigma}. \tag{2.2b}$$

The difference between these two combinations can be seen by considering the properties of the fields under a parity (P) transformation. The fields transform like

$$P\boldsymbol{E} = -\boldsymbol{E} \qquad\qquad P\boldsymbol{B} = +\boldsymbol{B}. \tag{2.3}$$

Therefore, the invariant combinations transform like

$$P\left(\boldsymbol{E} \cdot \boldsymbol{B}\right) = -\left(\boldsymbol{E} \cdot \boldsymbol{B}\right), \tag{2.4a}$$

$$P\left(\boldsymbol{E}^2 - \boldsymbol{B}^2\right) = +\left(\boldsymbol{E}^2 - \boldsymbol{B}^2\right). \tag{2.4b}$$

The first combination is therefore a pseudoscalar while the second is a scalar. The neutral pion is known to be a pseudoscalar particle (P–odd) and to be its own anti–particle (C–even). It should then be clear, given a pseudoscalar π^0, that the first coupling is CP–conserving while the second one is CP–violating.

Allowing for the possibility of CP–violation, the Lagrangian will be written as

$$\mathcal{L} = -\frac{\mathrm{i}}{4M}\left[\mathcal{F}_P \epsilon_{\mu\nu\rho\sigma} + \mathcal{F}_S\left(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}\right)\right] F^{\mu\nu}F^{\rho\sigma}\Pi, \tag{2.5}$$

where $\mathcal{F}_{P,S}$ are dimensionless, complex coupling coefficients that in generally depend on the momenta of the two photons. If the momentum of photon i is k_i and the pion mass is M, the quantity $x_i = k_i^2/M^2$ parameterizes how far photon i is from the mass shell. The coefficients can be decomposed into a real constant, a real function, and an imaginary phase,

$$\mathcal{F}_P = g_P f(x_1, x_2) e^{i\Psi_P}, \tag{2.6a}$$

$$\mathcal{F}_S = g_S f(x_1, x_2) e^{i\Psi_S}, \tag{2.6b}$$

where $g_{P,S}$ and $\Psi_{P,S}$ are real constants and the form factor $f(x_1, x_2)$ is real, dimensionless, and normalized such that $f(0,0) = 1$. It is useful here to introduce two parameters: a mixing angle ζ and a phase difference δ, defined by

$$g_P = \tilde{g}\cos\zeta, \tag{2.7a}$$

$$g_S = \tilde{g}\sin\zeta, \tag{2.7b}$$

$$\delta = \Psi_S - \Psi_P, \tag{2.7c}$$

where $\tilde{g}^2 = g_P^2 + g_S^2$. Ignoring an overall phase, the interaction takes its final form

$$\mathcal{L} = -\frac{\mathrm{i}}{4M}\tilde{g} f(x_1, x_2)\left[\cos\zeta\,\epsilon_{\mu\nu\rho\sigma} + \sin\zeta e^{i\delta}\left(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}\right)\right] F^{\mu\nu}F^{\rho\sigma}\Pi. \tag{2.8}$$

Before discussing the physical implications of the two couplings, I will summarize the various models typically used to describe the momentum dependence of the interaction. It should be noted that I have assumed that the momentum dependence of the two CP couplings is the same. This does not have to be the case, but the scalar term is known to be small, if not zero, and therefore whatever momentum dependence it might have is likely to be unimportant.

2.1.1 The π^0 Electromagnetic Form Factor

The electromagnetic form factor describes the departure from a point-like interaction due to the hadronic structure of the pion. The form factor must be a symmetric function of the two photon momenta and the pion momentum. It is a scalar function and takes scalar arguments. Given the three momenta available one usually uses $x_1 = k_1^2/M^2$ and $x_2 = k_2^2/M^2$. For real photons $x = 0$, while for e^+e^- pairs x has a minimum value of $x_{\min} = 4m^2/M^2$, where m is the electron mass. Additionally, the form factor is normalized by the condition $f(0,0) = 1$ and is required to vanish at some large momentum.

The π^0 form factor has been studied in a variety of theoretical frameworks, including quark triangle loops and vector meson dominance (VMD). While the calculation of quark loops was important in the context of the anomalous divergence of the axial current, the VMD models have been successful in describing the decays of both the neutral pion and kaon. These models take into account processes involving intermediate vector mesons which may decay to virtual photons. A recent VMD model, and one well suited to the double Dalitz mode, was proposed by D'Ambrosio, Isidori, and Portolés (DIP) [17],

$$f_{\mathrm{DIP}}(k_1^2, k_2^2) = 1 + \alpha\left(\frac{k_1^2}{k_1^2 - M_\rho^2} + \frac{k_2^2}{k_2^2 - M_\rho^2}\right) + \beta\frac{k_1^2 k_2^2}{(k_1^2 - M_\rho^2)(k_2^2 - M_\rho^2)}, \tag{2.9}$$

where $M_\rho \approx 770\,\mathrm{MeV/c^2}$ is the ρ meson mass. This form only includes the contribution from the ρ meson and ignores the ω and ϕ mesons. This is justified since the ρ-pole is already far removed from the physically allowed momentum range in pion decays and the other vector mesons are even more massive.

The two limiting forms offer some insight into this model. First consider the case of both photons being very nearly real, $k_i^2 \ll M_\rho^2$ for $i = 1, 2$. In that limit, the DIP form factor becomes identical to the general Taylor expansion to first order

$$f_{\mathrm{Taylor}}(k_1^2, k_2^2) = 1 - \alpha(k_1^2/M_\rho^2 + k_2^2/M_\rho^2) + \mathcal{O}(k^4). \tag{2.10}$$

Table 2.1: Recent measurements of the π^0 form factor slope parameter, $a = -\alpha\mu$.

Value	Process	Year
$+0.026 \pm 0.024 \pm 0.048$	$\pi^- p \to \pi^0 n$	1992[18]
$+0.025 \pm 0.014 \pm 0.026$	$\pi^- p \to \pi^0 n$	1992[19]
$+0.0326 \pm 0.0026 \pm 0.0026$	$e^+e^- \to e^+e^-\pi^0$	1991[20]
$-0.11 \pm 0.03 \pm 0.08$		1989[21]
$+0.032 \pm 0.004$	PDG Average	2002[13]

Clearly the parameter β is unimportant when the photons are nearly real. The other limit is where both momenta are very large. As they approach infinity, the form factor approaches the constant $1 + 2\alpha + \beta$. A reasonable model should have the form factor vanish above some very large momenta. I choose to force the form factor to vanish at infinity by fixing β such that $1+2\alpha+\beta = 0$, that is $\beta = -(1+2\alpha)$. Using this constraint the DIP form factor, in terms of x_1 and x_2, takes the form

$$f(x_1, x_2; \alpha) = \frac{1 - \mu(1+\alpha)(x_1 + x_2)}{(1 - \mu x_1)(1 - \mu x_2)}, \tag{2.11}$$

where $\mu = M^2/M_\rho^2 \approx 0.032$.

There have been several measurements of the π^0 form factor slope parameter, defined as linear coefficient in the Taylor expansion. In terms of the present notation, the slope parameter is $a = -\alpha\mu$. Table 2.1 lists the recent results taken from the PDG [13]. The indirect measurement from $e^+e^- \to e^+e^-\pi^0$ dominates the world average while the direct probes have at best uncertainties of 100%. In terms of the DIP parameter α, the world average is -1 within errors.

2.2 The $\pi^0 \to \gamma\gamma$ Decay

Consider the two–photon decay in a reference frame in which one photon travels in the $+\hat{\boldsymbol{z}}$ direction with helicity λ_1 and the other travels in the $-\hat{\boldsymbol{z}}$ direction with helicity λ_2. In that frame, the matrix element which describes the transition looks like

$$\mathcal{M}_{\lambda_1\lambda_2} = \frac{2}{M}\tilde{g} f(x_1, x_2; \alpha) \left[\cos\zeta \epsilon_{\mu\nu\rho\sigma} + \sin\zeta e^{i\delta} \left(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}\right)\right] k_1^\mu \epsilon_{\lambda_1}^{*\nu} k_2^\rho \epsilon_{\lambda_2}^{*\sigma}, \tag{2.12}$$

where k_i^μ and $\epsilon_{\lambda_i}^\mu$ are the momentum and polarization 4–vectors for photon i, $\epsilon_{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor, and $g_{\mu\nu}$ is the metric tensor. There are two allowed helicity configurations for the photons, denoted by $|{+}{+}\rangle$ and $|{-}{-}\rangle$. In the first state, both photons are right handed (right circularly polarized) and in the second they are both left handed. The calculation of the matrix element for each of the two terms is carried out in Appendix B. The results are

$$\mathcal{M}_{++} = -M\tilde{g}(\sin\zeta e^{i\delta} - \mathrm{i}\cos\zeta), \tag{2.13a}$$

$$\mathcal{M}_{--} = -M\tilde{g}(\sin\zeta e^{i\delta} + \mathrm{i}\cos\zeta). \tag{2.13b}$$

The partial width for the two allowed helicity states becomes

$$\Gamma_{++} = \frac{M\tilde{g}^2}{32\pi}\left(1 - 2\sin\zeta\cos\zeta\sin\delta\right), \tag{2.14a}$$

$$\Gamma_{--} = \frac{M\tilde{g}^2}{32\pi}\left(1 + 2\sin\zeta\cos\zeta\sin\delta\right). \tag{2.14b}$$

The decay rates to the two allowed final states must be equal if we require CPT invariance. This is because

$$C\,|{+}{+}\rangle = |{+}{+}\rangle \qquad C\,|{-}{-}\rangle = |{-}{-}\rangle \tag{2.15a}$$

$$P\,|{+}{+}\rangle = |{-}{-}\rangle \qquad P\,|{-}{-}\rangle = |{+}{+}\rangle \tag{2.15b}$$

$$T\,|{+}{+}\rangle = |{-}{-}\rangle \qquad T\,|{-}{-}\rangle = |{+}{+}\rangle \tag{2.15c}$$

Therefore, CPT will be conserved if either one of the couplings is zero or the phase difference between the two couplings is zero. Regardless of CPT conservation, the total rate of the decay is

$$\Gamma_{\gamma\gamma} = \frac{M\tilde{g}^2}{16\pi}. \tag{2.16}$$

The coupling constant $\tilde{g}$ can then be determined from the two–photon width,

$$\tilde{g} = \sqrt{16\pi\Gamma_{\gamma\gamma}/M}. \tag{2.17}$$

Table 2.2 lists the measured two–photon branching ratio and π^0 lifetime along with the derived value of the decay constant.

Table 2.2: Numerical values of the two–photon branching ratio $B_{\gamma\gamma}$ and π^0 lifetime τ from the PDG, and the π^0 decay constant $\tilde{g}$ derived from them.

Parameter	Value
$B_{\gamma\gamma}$	$(98.798 \pm 0.032)\%$
τ	$(8.4 \pm 0.6) \times 10^{-17}$ s
$\tilde{g}$	$(1.70 \pm 0.06) \times 10^{-3}$

In order to study a non–trivial form factor we must allow for one or both of the photons to be virtual and convert internally to an electron–positron pair. Such pairs are called Dalitz pairs after the first person to systematically study them, R. H. Dalitz. The single Dalitz decay allows for a measurement of the momentum dependence of the form factor when one photon is virtual, with a momentum squared in the range $4m^2$ to M^2. The next section will describe the general properties of Dalitz pairs through an analysis of the single Dalitz decay, $\pi^0 \to e^+e^-\gamma$.

2.3 The $\pi^0 \to e^+e^-\gamma$ Decay

Roughly one percent of the time one of the two photons will internally convert into an e^+e^- pair. The Feynman diagram for this process is shown in Figure 2.1. The theory of internal conversion was first studied by Dalitz [6] who predicted a tree-level branching ratio for $\pi^0 \to e^+e^-\gamma$ of 0.01185. The square of the momentum of the virtual photon is observed as the invariant mass of the Dalitz pair and takes on continuous values from $4m^2$ up to M^2. The process can therefore be used to study the functional form of the electromagnetic form factor.

The differential partial width for the single Dalitz decay has been extensively studied, most notably by Kroll and Wada [11]. In the current language it takes the form

$$\frac{1}{\Gamma_{\gamma\gamma}}\frac{d^2\Gamma_{ee\gamma}}{dxdy} = \frac{\alpha}{4\pi}f^2(x,0)\frac{(1-x)^3}{x}\sqrt{1-x_0/x}(1+y^2+x_0/x)(1+\delta(x,y)), \qquad (2.18)$$

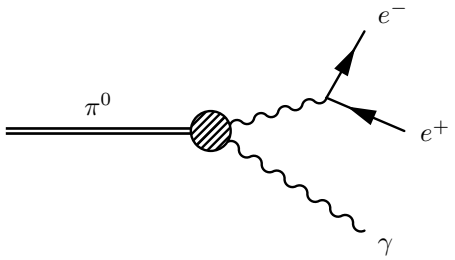

Figure 2.1: Single Dalitz Tree Level Feynman Diagram.

where the two phase space variables are defined by

$$x = \frac{(p_+ + p_-)^2}{P^2}, \tag{2.19a}$$

$$y = \frac{2\ P \cdot (p_+ - p_-)}{P^2(1-x)}, \tag{2.19b}$$

with P, p_+, and p_- representing the pion, positron, and electron momenta, respectively. The variable x_0 is the minimum value of x. The factor $\delta(x, y)$ is the leading order radiative correction, first calculated by Mikaelian and Smith [22]. The range of x is from $4m^2/M^2$ to 1, and the range of y, $-(1-x)$ to $(1-x)$. The partial width behaves as $1/x$ for small x and vanishes like $(1-x)^3$ for large values of x.

The $1/x$ dependence can be thought of as the unwillingness of the virtual photons to be too far off-shell. This is the distinguishing characteristic of Dalitz pairs. Figure 2.2 shows the distribution of e^+e^- masses predicted by the tree-level partial width, on a logarithmic scale.

While the single Dalitz process can be used to study the electromagnetic form factor when one photon is virtual, like the two–photon process, it is not sensitive to the CP nature of the pion. As discussed previously, one must measure the correlation between the planes of polarization of the two photons. The next section will show how this is accomplished in decays in which both photons are virtual.

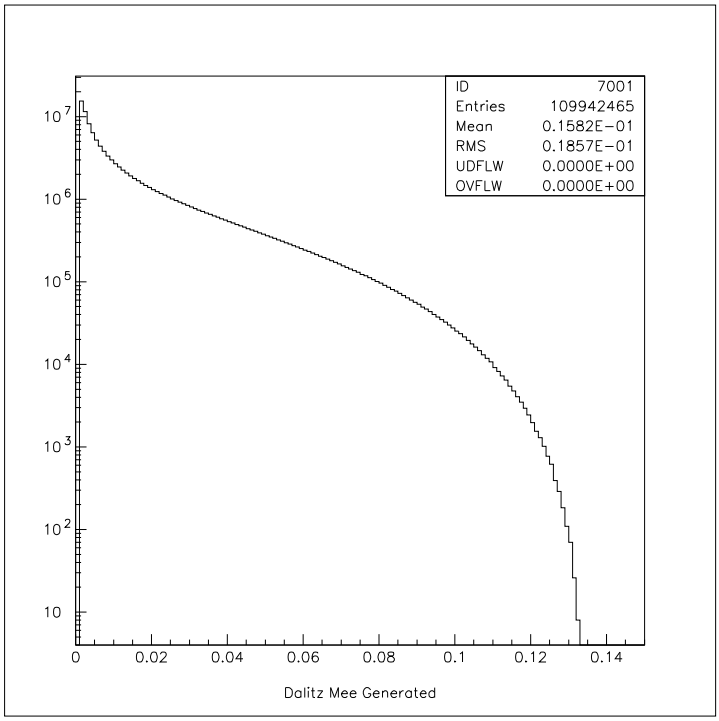

Figure 2.2: Predicted distribution of e^+e^- invariant masses in $\pi^0 \to e^+e^-\gamma$ decays shown on a logarithmic scale.

2.4 The $\pi^0 \to e^+e^-e^+e^-$ Decay

The $\pi^0 \to e^+e^-e^+e^-$ mode proceeds through the two-virtual-photon diagram shown in Figure 2.3. The final state contains two e^+e^- pairs and therefore has an exchange contribution where, for instance the two positrons are switched, as shown in Figure 2.3. As with any mode with charged particles in the final state, there are radiative graphs that must also be considered. The next section will present a tree-level analysis of the differential partial width with respect to e^+e^- masses and the angle between the two planes of the Dalitz pairs. The remaining sections will discuss the radiative decay $\pi^0 \to e^+e^-e^+e^-\gamma$ and its relationship to radiative corrections to the double Dalitz decay.

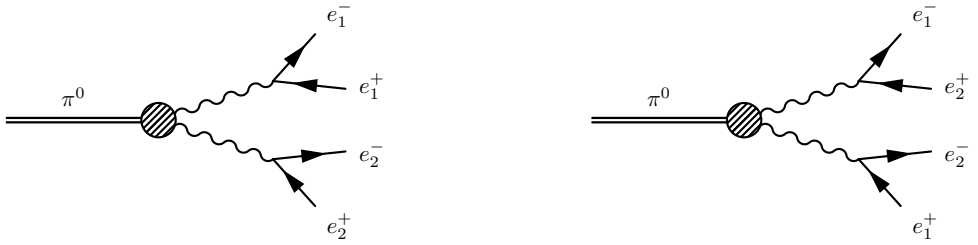

Figure 2.3: The two contributions to the double Dalitz decay at tree level.

2.4.1 Tree-Level Differential Partial Width

The differential partial width for this mode is

$$d^5\Gamma = \frac{1}{2M} \sum |\mathcal{M}_{\rm D} + \mathcal{M}_{\rm X}|^2 \, d^5\Phi_4, \tag{2.20}$$

where $\mathcal{M}_{\rm D(X)}$ is the matrix element for the direct (exchange) diagram and the sum is over the 16 helicity states. The 4–body phase space is given by

$$d^5\Phi_4 = \frac{M^4}{2^{16}\pi^6} \, \lambda \, dx_{12} \, dx_{34} \, dy_{12} \, dy_{34} \, d\phi, \tag{2.21}$$

where a factor of 1/4 has been included to account for the two pairs of identical particles in the final state. The phase space variables, along with other quantities used in this section, are defined in Appendix A. The x and y variables are equivalent to the single Dalitz phase space parameters. The ϕ variable is defined as the angle between the normals to the two planes formed by the e^+e^- pairs in the pion CM frame. The allowed physical region is defined by the following set of equations

$$x_{12} \geq 4m^2/M^2 \qquad x_{34} \geq 4m^2/M^2 \qquad \sqrt{x_{12}} + \sqrt{x_{34}} \leq 1, \tag{2.22a}$$

$$-\lambda_{12} \leq y_{12} \leq \lambda_{12} \qquad -\lambda_{34} \leq y_{34} \leq \lambda_{34}, \tag{2.22b}$$

$$0 \leq \phi < 2\pi. \tag{2.22c}$$

The squared matrix element for the double Dalitz decay, summed over final state

helicities, is

$$\sum |\mathcal{M}_D|^2 = \frac{2^8\pi^2\alpha^2\tilde{g}^2}{M^2\omega^4}\left(A\sin^2\phi + B\cos^2\phi + C\sin\phi\cos\phi + D\sin\phi + E\cos\phi + F\right), \tag{2.23}$$

where

$$\begin{aligned}
A &= \omega^2\{f_P^2\ \cos^2\zeta\ \lambda^2\left[1+(1-\lambda_{12}^2+y_{12}^2)(1-\lambda_{34}^2+y_{34}^2)\right] \\
&\quad + f_S^2\ \sin^2\zeta\ z^2\left[(1-\lambda_{12}^2+y_{12}^2)+(1-\lambda_{34}^2+y_{34}^2)\right]\}, && \text{(2.24a)}\\
B &= \omega^2\{f_S^2\ \sin^2\zeta\ z^2\left[1+(1-\lambda_{12}^2+y_{12}^2)(1-\lambda_{34}^2+y_{34}^2)\right] \\
&\quad + f_P^2\ \cos^2\zeta\ \lambda^2\left[(1-\lambda_{12}^2+y_{12}^2)+(1-\lambda_{34}^2+y_{34}^2)\right]\}, && \text{(2.24b)}\\
C &= 2\ f_P f_S\ \sin\zeta\cos\zeta\cos\delta\ \lambda\ z\ \omega^2\ (\lambda_{12}^2-y_{12}^2)(\lambda_{34}^2-y_{34}^2), && \text{(2.24c)}\\
D &= 2\ f_P f_S\ \sin\zeta\cos\zeta\cos\delta\ \lambda\ \omega^3\ y_{12}\ y_{34}\sqrt{(\lambda_{12}^2-y_{12}^2)(\lambda_{34}^2-y_{34}^2)}, && \text{(2.24d)}\\
E &= 2\ f_S^2\ \sin^2\zeta\ z\ \omega^3\ y_{12}\ y_{34}\sqrt{(\lambda_{12}^2-y_{12}^2)(\lambda_{34}^2-y_{34}^2)}, && \text{(2.24e)}\\
F &= f_S^2\ \sin^2\zeta\ \omega^4\ (1-y_{12}^2)(1-y_{34}^2). && \text{(2.24f)}
\end{aligned}$$

The integrals over the two y variables are trivial and, consequently, the distribution of the phase space in the y variables is uninteresting. The most dramatic feature of the phase space can be seen by integrating over ϕ next. The differential width with respect to the x variables looks like

$$\frac{1}{\Gamma_{\gamma\gamma}}\frac{d^2\Gamma_D}{dx_{12}dx_{34}} = \frac{\alpha^2}{18\pi^2}\frac{\lambda_{12}\lambda_{34}\lambda}{\omega^2}(3-\lambda_{12}^2)(3-\lambda_{34}^2)\left[f_P^2\cos^2\zeta\lambda^2 + f_S^2\sin^2\zeta(\lambda^2+3\omega^2/2)\right] \tag{2.25}$$

The differential decay width becomes large when both x_{12} and x_{34} are small and vanishes as either one or both becomes large.

If instead we integrate over the x variables before ϕ we arrive at the differential

Table 2.3: Results of integrating numerically over x_{12} and x_{34} while assuming $\alpha = 0$.

Integral	I_1	I_2	I_3	I_4	I_5	I_6
Value	7.2287	7.2838	14.509	15.600	15.684	0.0555

width with respect to the angle ϕ

$$\frac{1}{\Gamma_{\gamma\gamma}}\frac{d\Gamma_D}{d\phi} = \frac{\alpha^2}{12\pi^3}[I_1\ \cos^2\zeta\ \sin^2\phi + I_2\ \sin^2\zeta\ \cos^2\phi + I_3\ \sin\zeta\cos\zeta\cos\delta\ \sin\phi\cos\phi + I_4\ \cos^2\zeta + (I_5+I_6)\sin^2\zeta], \tag{2.26}$$

given in terms of integrals over x,

$$I_1 = \frac{2}{3}\iint dx_{12}\ dx_{34}\ f_P^2\ \frac{\lambda_{12}^3\lambda_{34}^3\lambda^3}{\omega^2}, \tag{2.27a}$$

$$I_2 = \frac{2}{3}\iint dx_{12}\ dx_{34}\ f_S^2\ \frac{\lambda_{12}^3\lambda_{34}^3\lambda z^2}{\omega^2}, \tag{2.27b}$$

$$I_3 = \frac{4}{3}\iint dx_{12}\ dx_{34}\ f_P f_S\ \frac{\lambda_{12}^3\lambda_{34}^3\lambda^2 z}{\omega^2}, \tag{2.27c}$$

$$I_4 = \iint dx_{12}\ dx_{34}\ f_P^2\ \frac{\lambda_{12}\lambda_{34}\lambda^3}{\omega^2}\ (3-\lambda_{12}^2-\lambda_{34}^2), \tag{2.27d}$$

$$I_5 = \iint dx_{12}\ dx_{34}\ f_S^2\ \frac{\lambda_{12}\lambda_{34}\lambda z^2}{\omega^2}\ (3-\lambda_{12}^2-\lambda_{34}^2), \tag{2.27e}$$

$$I_6 = \frac{1}{6}\iint dx_{12}\ dx_{34}\ f_S^2\ \lambda_{12}\lambda_{34}\lambda\ (3-\lambda_{12}^2)(3-\lambda_{34}^2). \tag{2.27f}$$

Once the form factor is specified the integrals can be done numerically. Table 2.3 gives the values assuming a flat form factor, $\alpha = 0$.

Combining terms with similar ϕ dependence leads to the following useful form,

$$\frac{1}{\Gamma_{\gamma\gamma}}\frac{d\Gamma_D}{d\phi} = \frac{\alpha^2}{12\pi^3}R\left(1+\kappa_1\cos 2\phi + \kappa_2\sin 2\phi\right), \tag{2.28}$$

where

$$R = (I_1/2 + I_4)\cos^2\zeta + (I_2/2 + I_5 + I_6)\sin^2\zeta, \tag{2.29a}$$

$$\kappa_1 = \frac{1}{2R}(I_2\sin^2\zeta - I_1\cos^2\zeta), \tag{2.29b}$$

$$\kappa_2 = \frac{1}{2R}I_3\sin\zeta\cos\zeta\cos\delta. \tag{2.29c}$$

The constant term R varies from 19.214 at $\zeta = 0$ to 19.381 at $\zeta = \pi/2$, a change of only 0.9 %. The coefficient of $\cos 2\phi$ changes from -0.1881 to $+0.1879$ over the same range. The size of the $\sin 2\phi$ term depends on the phase difference δ. If $\delta = 0$, the amplitude is maximum for an equal mixture of scalar-pseudoscalar couplings, at which point it is equal to $+0.1880$. In this way, again with $\delta = 0$, the magnitude of the ϕ dependence is basically a constant but the phase changes continuously as ζ is varied from 0 to $\pi/2$.

The total rate, associated with the direct contribution, is found simply by integrating over ϕ. The result is

$$\frac{\Gamma_D}{\Gamma_{\gamma\gamma}} = \frac{\alpha^2}{6\pi^2} R. \tag{2.30}$$

The rate including exchange contributions is just $\Gamma = 2\Gamma_D + \Gamma_{DX}$ since the integrated rate of the exchange graph is identical to that for the direct graph. The form of the interference term, Γ_{DX}, can be found in Reference [15]. The effect of the interference is to reduce the partial width by approximately 1 %.

2.4.2 Radiative Decays: $\pi^0 \to e^+e^-e^+e^-\gamma$

There are two channels for the radiative decay mode: the first, shown on the left of Figure 2.4, is radiation from a double Dalitz final state electron, the second, shown on the right, is a radiative single Dalitz decay with the radiated photon internally converting. There are a total of 8 diagrams that contribute, four of each of the two types.

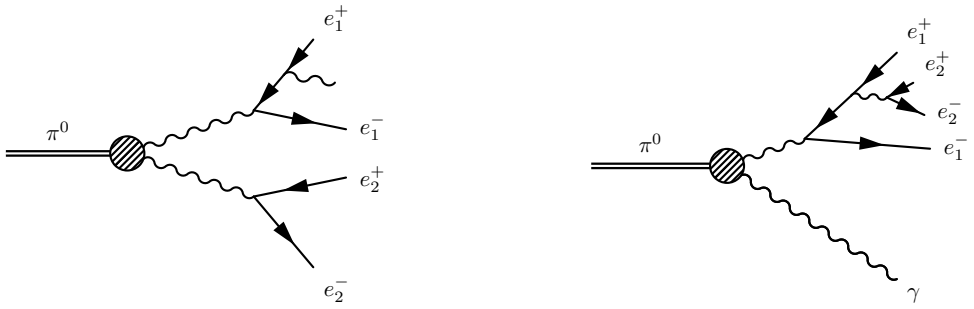

Figure 2.4: Two examples of contributions to the radiative double Dalitz decay.

On average, the Dalitz photon in the second process will be hard, with an energy near half of the pion mass in the CM frame, or 67 MeV. In contrast, the radiative photon in the first process can be arbitrarily soft. A useful quantity, closely related to the photon energy in the CM frame, is x_{4e}, where

$$x_{4e} = \frac{M_{4e}^2}{M_{4e\gamma}^2} = 1 - \frac{2E_\gamma^*}{M_{\pi^0}}. \tag{2.31}$$

The distribution of this quantity, made in a Monte Carlo simulation that is discussed in Chapter 4, is shown in Figure 2.5. The part of the distribution near $x_{4e} \sim 1$ is dominated by the soft, radiated photons from the first process, while the peak near $x_{4e} \sim 0$ is populated by the Dalitz photons of the second process. The range of x_{4e} is from $(4m)^2/M^2$ to 1.

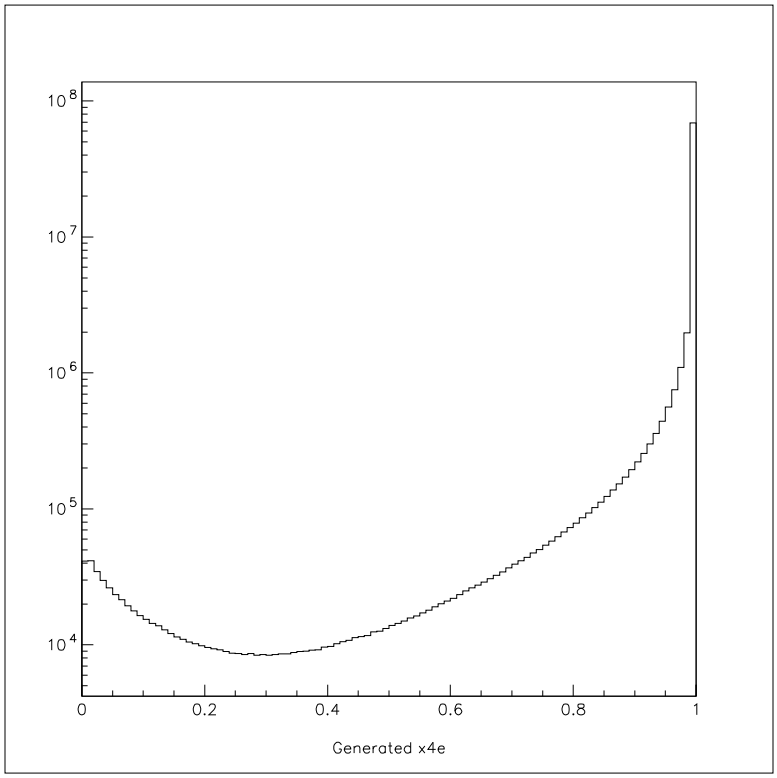

Figure 2.5: Monte Carlo distribution of x_{4e}.

The rate for the first process diverges as the photon energy approaches zero. In nature, there must be some cancellation of this divergence since the rate for any process

Table 2.4: Predicted rate for $\pi^0 \to e^+e^-e^+e^-\gamma$ process with two different infrared cutoffs imposed.

Cutoff	$\Gamma_{4e\gamma}/\Gamma_{\gamma\gamma}$
$x_{4e} < 0.9985$	6.614×10^{-6}
$x_{4e} < 0.90$	2.055×10^{-6}

must be finite. There is a cancellation which comes from other divergent, higher order, $\pi^0 \to e^+e^-e^+e^-$ graphs. These will be discussed in the next section in the context of radiative corrections. For now I will impose an infrared cutoff which can be thought of as a temporary photon mass. It will be shown in the next section that, after considering the radiative corrections, all dependence on this artifact will vanish. The Monte Carlo generator will impose a cutoff on the photon energy, in the CM frame, at 100 keV, or $x_{4e}^{\rm cut} \approx 0.9985$. The value of the cutoff was selected such as to be well below the range of experimental sensitivity, which for KTeV extends down to about 2 MeV in the pion rest frame.

The Monte Carlo predictions for the rate, using two different cutoffs, are shown in Table 2.4. The value of the observable cutoff was chosen to be to be $x_{4e} < 0.90$ based on the lower limit on the four electron invariant mass in the observed double Dalitz decays, which is roughly 128 MeV.

2.4.3 Radiative Corrections to the Differential Width

The need for radiative corrections was discussed above to deal with the apparent divergence in the radiative rate. The new processes that will be considered are the 1–loop diagrams of the double Dalitz decay. Examples of these graphs are shown in Figure 2.6.

The first diagram shows the vacuum polarization contribution, of which there are a total of four graphs. The vacuum polarization process is finite in the infrared limit, but diverges for large virtual-photon momenta. The second diagram is one of four vertex

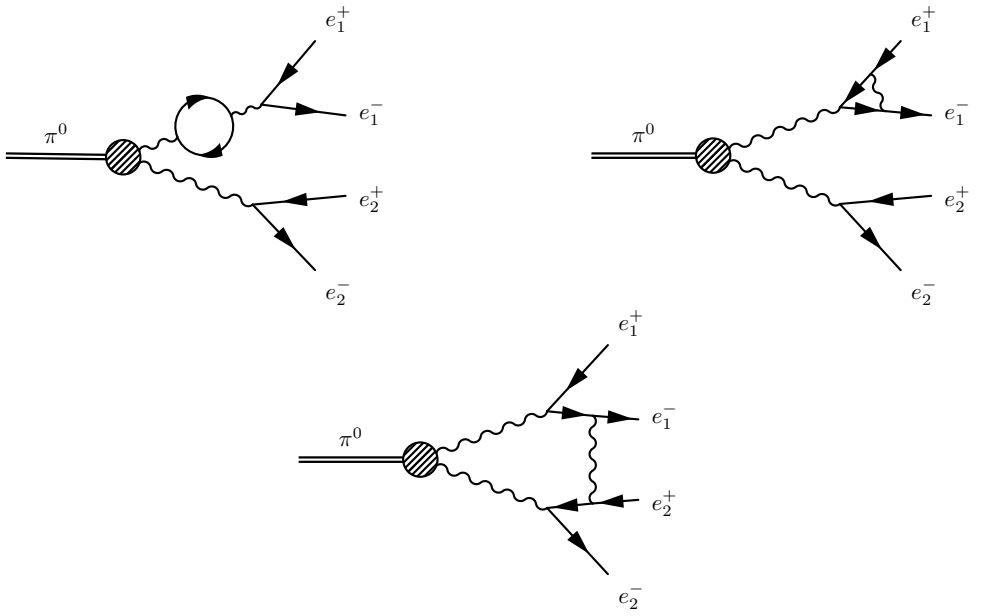

Figure 2.6: Examples of the three types of 1–loop contributions to the double Dalitz decay: vacuum polarization, vertex correction, and happy face.

correction contributions which diverge in both the infrared and the ultraviolet limits. The last process, referred to as the happy face process for many personal reasons, also diverges in both limits. There are a total of 8 happy face graphs. The combination of the two infrared divergences exactly cancel the divergence in the radiative rate, such that the total rate for the two modes, double Dalitz and radiative double Dalitz, is finite. It has been shown that this cancellation is guaranteed to occur at every order for QED (see for example Reference [23]).

The perturbation expansion of the double Dalitz matrix element, in terms of the electron charge, looks like

$$\mathcal{M} = \mathcal{M}_{\text{tree}} + \mathcal{M}_{\text{1–loop}} + \mathcal{O}(e^6). \tag{2.32}$$

The square of the matrix element is then

$$|\mathcal{M}|^2 = |\mathcal{M}_{\text{tree}}|^2 \left[1 + \frac{2\text{Re}(\mathcal{M}^*_{\text{tree}}\mathcal{M}_{\text{1–loop}})}{|\mathcal{M}_{\text{tree}}|^2}\right]. \tag{2.33}$$

The second term in braces is the first order correction to the matrix element squared.

Table 2.5: The partial width, normalized to the two–photon width.

Branching Ratio	Predicted Value
$B(4e, \text{tree})$	3.421×10^{-5}
$B(4e(\gamma), \text{all } x_{4e})$	3.536×10^{-5}
$B(4e(\gamma), x_{4e} > 0.90)$	3.331×10^{-5}
$B(4e(\gamma), x_{4e} < 0.90)$	2.055×10^{-6}

There will be three of these corrections, one for each of the three types of 1–loop graphs.

These corrections are referred to as virtual corrections, as opposed to the brem correction involving real photons. The part of the radiative rate below the infrared cutoff contains the divergent term which leads to the consideration of radiative corrections, but it also contains finite terms, some of which depend on the cutoff. The exact form of the corrections can be found in Reference [15].

The Monte Carlo prediction for the partial width at tree level and at first order are summarized in Table 2.5. The combined width is referred to as $\Gamma_{4e(\gamma)}$ to indicate that the photon may or may not be detectable.

2.5 Current Experimental Picture

Three of the results presented here have been previously measured: the double Dalitz branching ratio, the π^0 form factor, and the pseudoscalar-scalar mixing. The branching ratio and mixing have only been measured once, in the 1962 Samios experiment [12]. I will discuss that experiment in some detail, after summarizing the results on the π^0 form factor.

The slope parameter of the form factor has been measured in both the single and double Dalitz modes, with the single Dalitz results being much more precise. However, the most precise measurement to date is not from π^0 decay, but rather from π^0 production, through $e^+e^- \to \pi^0 e^+ e^-$. The q^2's accessible to this experiment were large and negative, requiring a significant extrapolation to the range accessible in π^0 decays.

The validity of this result hinges on the assumption that the physics at large, negative q^2 is related to the physics at small, positive q^2. While this may be the case, a precise measurement in the physically allowed region is desirable. Table 2.1 lists the measured values of the π^0 slope parameter.

2.5.1 The 1962 Samios Result

The only other experiment to study the double Dalitz decay of the pion was performed in 1962 at the Nevis cyclotron at Cornell University by Nick Samios and collaborators. The group recorded photographs in a bubble chamber triggered by a stopping π^-. More than 800,000 photos were taken and scanned by hand to produce a final data sample of roughly 200 $\pi^0 \to e^+e^-e^+e^-$ events. Figure 2.7 shows one such event.

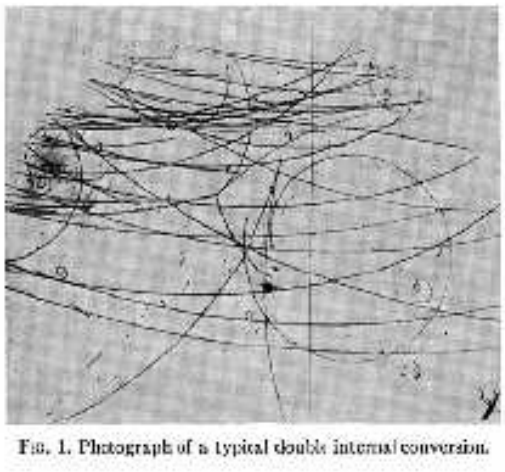

FIG. 1. Photograph of a typical double internal conversion.

Figure 2.7: Example of $\pi^- p \to \pi^0 n$ with $\pi^0 \to e^+e^-e^+e^-$ recorded in the 1962 bubble chamber experiment of Samios *et al.* The point at which the π^- stops and the two electron-positron pairs emerge is just to the upper-left of the center of the photo.

The group reported a branching ratio measurement,

$$B(\pi^0 \to e^+e^-e^+e^-)_{\text{Samios}} = (3.18 \pm 0.30) \times 10^{-5}, \tag{2.34}$$

as well as a study of the ϕ distribution. The extracted value of the $\cos 2\phi$ oscillation

was found to be $\kappa_1 = -0.12 \pm 0.15$, which is inconsistent with the hypothesis of a pure scalar pion at the $3.6\,\sigma$ level. They also attempted to measure a $\sin 2\phi$ term, with a resulting amplitude of $\kappa_2 = 0.09 \pm 0.13$.

Chapter 3

Kaons at the TeVatron

KTeV (Kaons at the TeVatron) was a fixed target experiment that collected neutral kaon decays (along with hyperon and neutral pion decays) at the Fermi National Accelerator Lab (FNAL, or more commonly Fermilab) on two separate occasions. The design goals were two–fold in order to accommodate the two experimental efforts: a thorough study of rare neutral kaon, pion, and hyperon decays, performed by E799-II, and a precise measurement of the CP violating parameter $\mathrm{Re}(\epsilon'/\epsilon)$, performed by the E832 experiment.

The two experiments had similar beam line elements and shared many of the same detector elements along with a data acquisition (DAQ) system. The data used in this analysis is from the E799-II configuration, and therefore I will focus on a detailed description of the E799-II setup.[1]

The experiment was operated in E799-II mode three times. As Figure 3.1 shows, there were two run periods in 1997 and one in 1999 and early 2000. The data collected in 1997 from the end of January to the end of March is referred to as the **Win97** dataset while data gathered between the end of July and the beginning of September is known as the **Sum97** dataset. The data collected from September 1999 through January 2000 is simply known as the **99** dataset. For reference, Table 3.1 lists the run ranges of the three periods.

[1] For a description of the E832 configuration, see for example Reference [24]

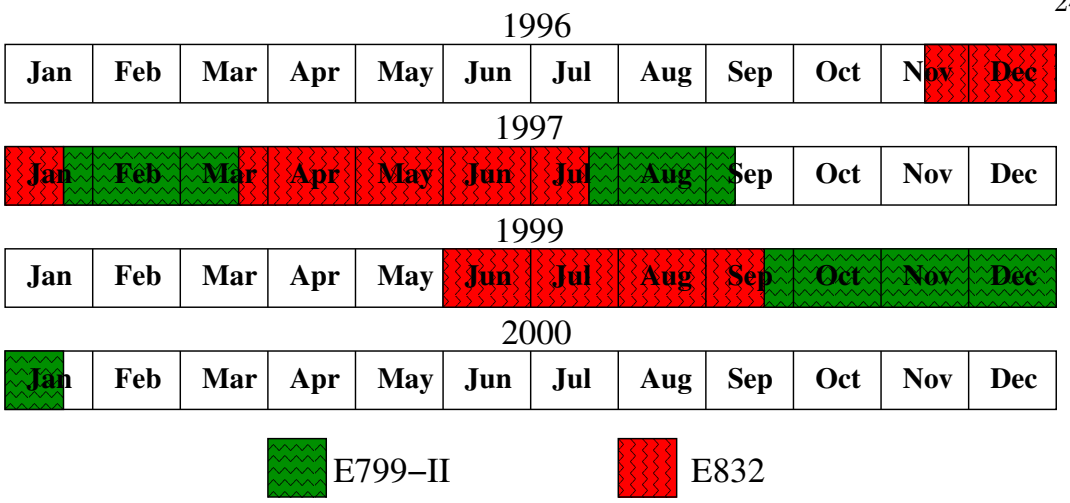

Figure 3.1: Periods of data taking.

This chapter will cover the hardware responsible for producing kaons and detecting their decay products. The first part will describe Fermilab's system of accelerators leading up to the Tevatron and the proton beam it delivers to the KTeV target. The target then converts the energy of the proton into a slew of secondary particles, including kaons. The secondary beam is filtered and shaped to form two nearly parallel beams containing mainly neutrons, kaons, and lambdas.

3.1 Fermilab

Fermilab consists of several proton/anti–proton accelerators along with both collider and fixed target experiments. Protons, taken from ionized hydrogen gas, are sent through a series of pre-accelerators including a Cockroft-Walton accelerator, a linear accelerator, and a booster. At each stage the protons gain energy and by the time they reach the booster they are energetic enough to create anti-protons on a target when in collider mode. In fixed-target mode, the protons continued on to either the Main Ring in 1997, or the newly constructed Main Injector in 1999. The protons are then accelerated to 150 GeV before being injected into the Tevatron where they will reach a final energy near 800 GeV.

Table 3.1: Run ranges for the three E799 run periods.

Dataset	First Run	Last Run
Win97	08088	08913
Sum97	10463	10970
99	14625	15548

In fixed-target mode, the protons are extracted from the Tevatron and delivered to one or more of the fixed-target beam-lines, including the Neutrino-Muon beam-line that served KTeV. The proton beam is a bunched beam, meaning that the protons are concentrated in isolated groups as they are accelerated. The bunches are between 1–2 ns long and separated by the Tevatron RF period of 19 ns. The protons are accelerated for 20 s, then kept at the 800 GeV "flat-top" energy while being slowly extracted for either 20 s (during 1997) or 40 s (in 1999), and then more protons are accelerated. During the 1997 run, the **on-spill** time was 20 s during which between 2×10^{12} and 5×10^{12} protons arrived at the KTeV target. In 1999, the on-spill time was doubled as was the number of protons per spill, making the instantaneous intensity about 1.8×10^{11} protons per second during both run periods.

3.2 The KTeV Beam-line

The KTeV beam-line begins with a **BeO target** onto which the proton beam was focused to a width of about 250 μm and an angle of 4.8 mrad. The beryllium target was rectangular with dimensions of 3 mm $\times$ 3 mm $\times$ 30 cm. This corresponded to about 1.1 interaction lengths. The KTeV coordinate system had its origin at the center of the target with the $+z$–axis pointing along the beam direction. The $+y$–axis was then defined to point vertically up leaving the $+x$–axis to point to the left as one looked downstream from the target.

Many particles were produced in the interaction of the protons with the target, almost all of which were unwanted. The charged particles, including the primary pro-

tons, could be removed with magnetic fields. Unstable neutral particles could be limited by moving the detector far enough away from the target that most would decay before reaching it. The amount of the remaining types of particles, long-lived or even stable neutrals, could be reduced through the use of absorbers that differentiate between kaons and other neutral particles, most importantly photons and neutrons. All of these techniques, along with beam collimation, were employed in the KTeV beam-line, shown in Figure 3.2.

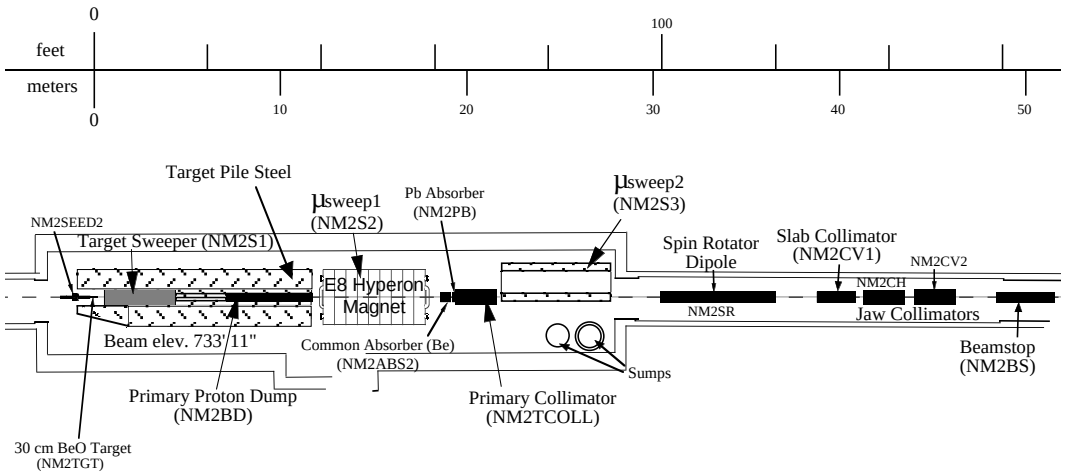

Figure 3.2: KTeV Beam-line. The NM2 enclosure houses the target, beam dump, sweeper magnets, absorbers and collimators.

A total of four sweeper magnets were employed to remove charged particles from the beam. The first three were located in the NM2 enclosure while the last was located in the NM3 enclosure just upstream of the KTeV Hall. The **target sweeper**, located between 0.4 m and 4.4 m, gave a 475 MeV/c kick in the $-y$-direction to any charged particles in the beam. The kick directed the protons into the **primary proton dump**, a 4.5 m water-cooled cooper block positioned below the beam. Starting at 12.3 m was ***μ*sweep1**, which delivered a kick of 3806 MeV/c in the $+x$-direction. Further downstream was ***μ*sweep2**, which gave a kick of either 3135 MeV/c in Win97 and 1854 MeV/c

thereafter. The final sweeper magnet was located near 90 m at the beginning of the decay volume. One additional magnet resided in the NM2 enclosure. The **spin rotator dipole**, located near 33 m, was used to alter the polarization state of neutral hyperons, like the Ξ^0 and Λ^0.

Between the two sweeper magnet was the **lead absorber** which consisted of a 3 in lead wall. The 14 radiation lengths of lead were in place to remove photons created at the target by converting them into e^+e^- pairs which could then be swept out of the beam.

The remaining elements were designed to shape the two secondary beams. The **primary collimator**, located just downstream of the absorber, consisted of a brass block with two rectangular holes. The size of the holes was 1.18×1.29 cm in Win97 and 1.62×1.73 cm for the rest of the data taking. The **slab collimator** located between the two beams to keep particles from crossing over, was only used in the Win97 period. The final collimator, located at 85 m, was called the **defining collimator** and had square holes of 4.4×4.4 cm in Win97 and 5.2×5.2 cm for Sum97 and 99.

The target was monitored by the **90° target monitor**. This detector consisted of three counters inside a small hole in the target pile perpendicular to the beam. A special trigger used the instrument to collect **accidental events** to be used by the Monte Carlo program to simulate activity uncorrelated with any kaon decays. This "noise" was composed of junk from the target, beam interactions with the detector material, and cosmic rays.

3.3 The KTeV Detector

The KTeV detector was designed to measure all of the kaon decay particles save the elusive neutrino. The relativistic boost factor for kaons with energies on the order of 100 GeV is about $\gamma \approx 200$, which sets the relative scale of the detector. The KTeV detector began 90 m downstream of the target and extended for over 100 m more. Fig-

Table 3.2: Ring counter geometry.

Counter	z Position (m)	Outer Radius (m)	Inner Aperture (m)
RC6	132.6	1.00	0.84×0.84
RC7	138.6	1.00	0.84×0.84
RC8	146.6	1.44	1.18×1.18
RC9	152.6	1.44	1.18×1.18
RC10	158.6	1.44	1.18×1.18

ure 3.3 shows a 3-D representation of the detector in E799-II mode. In this figure the scale of the z–axis is half the scale of the transverse dimensions. In E799 mode, two parallel K_L beams entered the decay region near 90 m. The decay region was evacuated and kept at a pressure of 1.0×10^{-6} Torr.

The following sections describe the individual detector elements shown in Figure 3.3 in the context of this analysis.

3.3.1 The Decay Region

The decay region consisted of an evacuated cylindrical tank starting near 90 m and ending at a Kevlar window at 158 m. The tank was kept at a pressure of 1.0×10^{-6} Torr. The diameter of the tank at its furthest upstream point was roughly 0.5 m while at the window it was 1.8 m. The vacuum window was 0.0156 % of a radiation length thick.

3.3.1.1 Ring Counters

At several locations along the length of the vacuum tank were ring shaped counters designed to detect particles leaving the decay region at high angles. These ring counters (RC) had a circular outer aperture with a square inner aperture. Each RC had 16 overlapping segments consisting of 24 lead-scintillator layers totaling 16 radiation lengths of material. Table 3.2 lists the geometry of the five RC vetoes. RC7 served as the defining aperture the veto system.

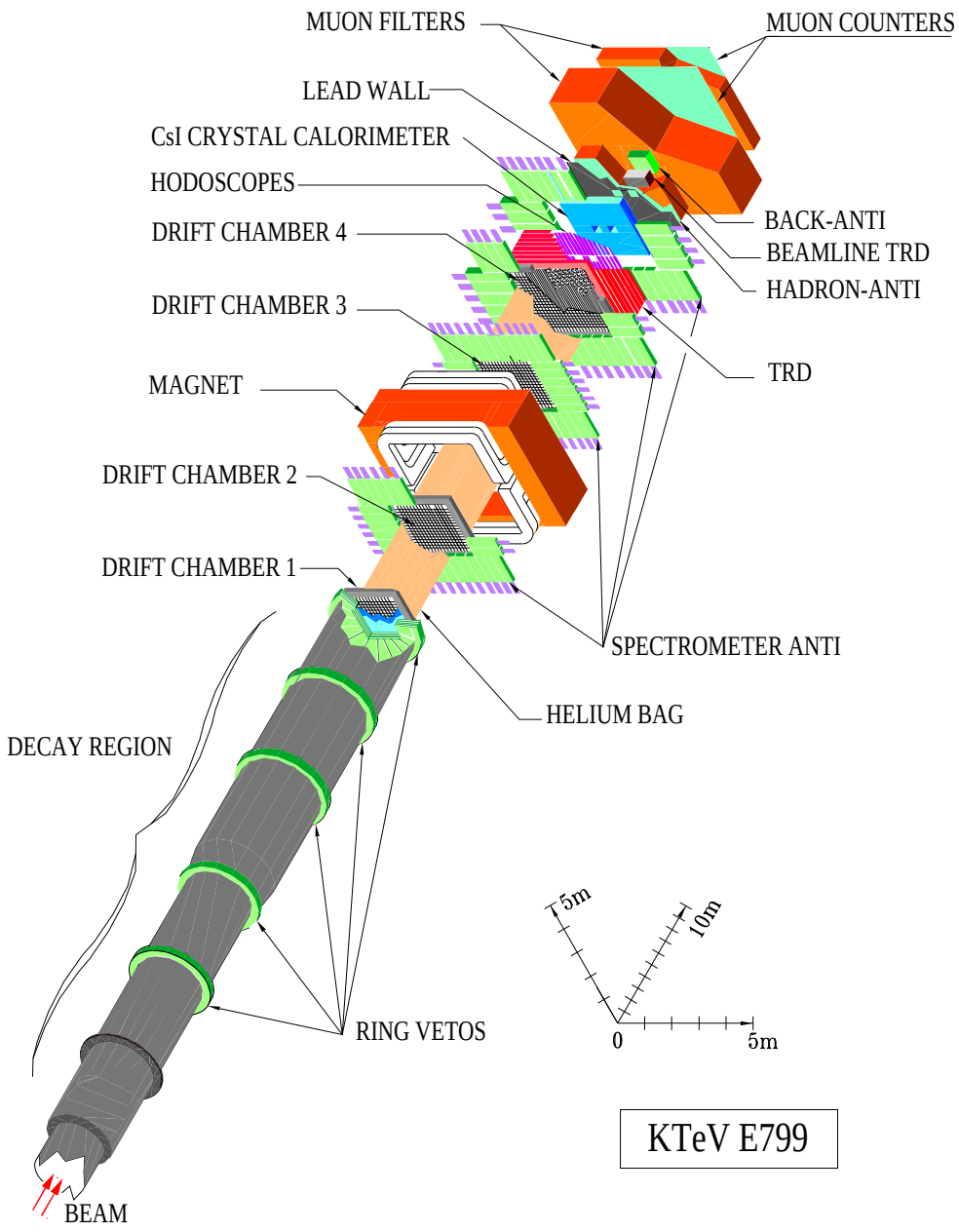

Figure 3.3: The KTeV Detector. The axis parallel to the beam defines $\hat{z}$ and is shown elongated by a factor of 2×1 relative to the transverse directions.

3.3.2 The Charged Spectrometer

The spectrometer consisted of a set of four drift chambers and an analysis magnet. The chambers provided position measurements in both x and y with a resolution

Table 3.3: Geometry of the charged spectrometer.

	z Position (m)	$x \times y$ Dimensions (m)
DC1	159.4	1.30 × 1.30
DC2	165.6	1.64 × 1.44
Magnet	170.0	2.90 × 2.00
DC3	174.6	1.74 × 1.64
DC4	180.5	1.90 × 1.90

of 100 μm. They were positioned with two upstream and two downstream of a dipole magnet which imparted a 205 MeV/c (150 MeV/c) kick in the x-direction during 1997 (1999). The magnitude of the kick was decreased in 99 in order to increase the acceptance for modes with four tracks. The polarity of the magnet was reversed periodically in order to remove any systematic bias related to detector symmetry. Table 3.3 lists the geometry of the spectrometer. A bag filled with helium was placed between adjacent chambers in order to reduce multiple scattering of tracks.

3.3.2.1 Drift Chamber

The drift chambers (DC) consisted of a set of anode "sense wires", in an argon–ethane gas, surrounded by cathode "field wires" that created large potential wells around each sense wire. Charged particles passing through the chamber ionize the gas and the potential accelerates the ionization toward the nearest sense wire where an electron avalanche occurs. By knowing the characteristic drift velocity of the ionization, a measurement of the arrival time of the ionization yields the distance at which the charged particle passed. In order to determine on which side of the wire the track passed, two planes of offset sense wires were used. The geometry of a **plane pair** is illustrated in Figure 3.4 along with a cartoon of the ionization signature of a typical charged particle.

The field wires were made of 25 μm diameter gold-plated tungsten while the sense wires were 100 μm gold-plated aluminum. The sense wires were spaced at 12.7 mm intervals, making the cell size of a plane pair 6.35 mm. The field wires were maintained

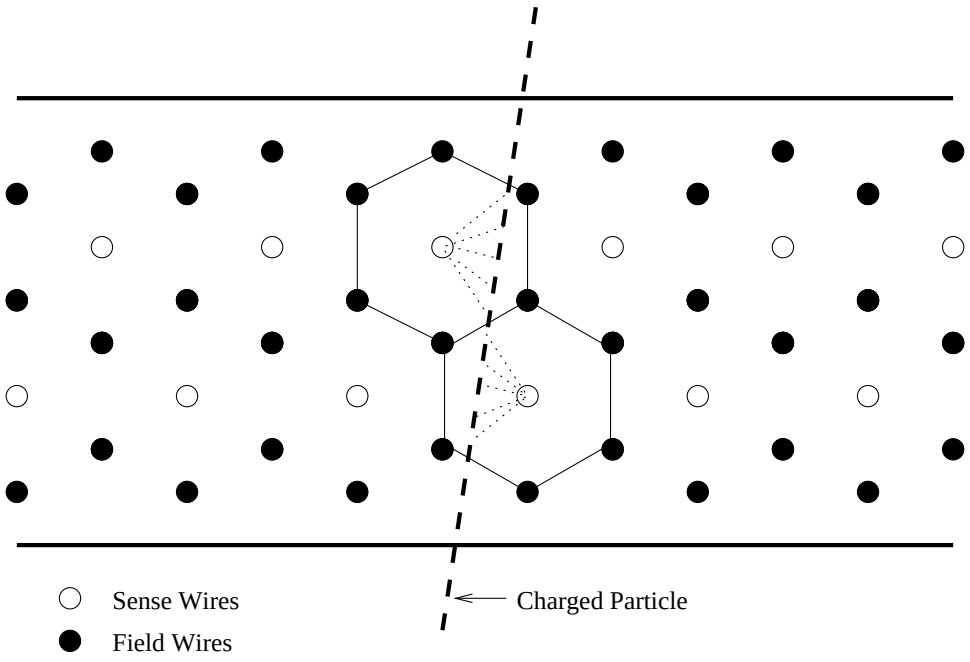

Figure 3.4: Drift chamber plane pair geometry and illustration of ionization process.

at roughly 2500 V above the sense wires, resulting in a mean drift velocity of 50 μm/ns. At this speed the maximum drift time was less than 150 ns.

Each chamber was instrumented with pre-amplifiers which relayed signals on the sense wires to a set of discriminators. The discriminated signal then started the timing window of a time-to-digital converter (TDC). The end of the window was marked by a common stop supplied by the Level 1 trigger logic described later.

3.3.2.2 Spectrometer Anti

Surrounding each of the drift chambers save DC1, were more vetoes whose purpose it was to detect particles exiting at high angles. These large detectors were designed and built in Boulder by the Colorado group. The geometry of the chambers was square and therefore so was the geometry of the spectrometer antis (SAs). Otherwise they were nearly identical to the ring counters discussed above, with 32 layers (16 radiation lengths total) of lead-scintillator. Table 3.4 lists the geometry of the three SA counters.

Table 3.4: Geometry of the spectrometer anti's.

Counter	z Position (m)	Outer Dimension (m)	Inner Aperture (m)
SA2	165.1	2.50 × 2.50	1.54 × 1.37
SA3	174.0	3.00 × 2.40	1.69 × 1.60
SA4	180.0	2.37 × 2.37	1.75 × 1.75

3.3.2.3 Spectrometer Calibration

The precision of the momentum measurement depends on two thing: how well the track positions at each chamber are known and how well the magnetic field is known. An accurate track position measurement requires both knowledge of the absolute position, in three dimensions, of all wires in all chambers as well as a map to convert drift times to distances for all wires. Both the wire positions and the conversion maps ($X(t)$) can be measured from the data. As for the magnetic field, once the track position information is calibrated, $K \to \pi^+\pi^-$ events can be used to fix the mean field strength. The fringe fields away from the bend plane were also measured in the data.

The first step to calibrating the chambers was the determination of the timing offset (T0) of every signal coming into the TDC modules. The offsets were due to signals traveling different distances, both within a chamber and also between the chambers and the TDC's. The T0 for each wire was found by studying the TDC count distribution in clean $K_L \to \pi e\nu$ ($Ke3$) events and determining the time of the early edge.

With the offsets in hand, the $X(t)$ maps could be reliably calculated. The data used for this purpose was collected during special runs called **muon runs** during which the beam stops were closed and the sweeper and analysis magnets were turned off. Assuming that these straight-through muons uniformly illuminated every cell, the time-to-distance conversion factor at time t can be found from the distribution of TDC counts before and after t.

The next step addressed the wire positions within the chamber along with the

chamber positions within the overall detector. The wire positions within the chambers were surveyed both before and after the experiment ran. The chambers were aligned with each other using data from the muon runs mentioned above. Special tracking algorithms were used to find tracks based on DC1 and DC4 only. These tracks were then interpolated to DC2 and DC3 where the difference between actual hit positions and the interpolation is partially due to misalignment of the two inner chambers with respect to the fixed outer chambers. The overall spectrometer was then aligned with respect to the target and the calorimeter using $K \to \pi^+\pi^-$ and $Ke3$ events, respectively.

After the spectrometer has been calibrated, the resolution on the track momentum is approximated by

$$\frac{\sigma(p)}{p} = 0.38\% \oplus 0.016\% \cdot p, \tag{3.1}$$

with p measured in GeV/c. The constant term is due to multiple scattering while the linear term is due to the resolution on the bend angle of the track. For a typical charged particle with a momentum of 10 GeV/c, the momentum resolution is about 0.4%.

3.3.3 The Transition Radiation Detectors

A set of eight transition radiation detectors (TRDs) was located downstream of chamber 4 in E799 mode. These detectors can be used to discriminate between charged pions and electrons since the energy of transition radiation is inversely proportional to the particle mass. At KTeV energies, electrons and pions produce x-rays in a radiator; the x-rays are detected by a multi-wire proportional chamber (MWPC). Each plane had two 15 cm square regions where the radiator was removed and the wires deadened, to reduce beam interactions. The pion rejection was roughly 300:1 with a 90% efficiency for electrons. The position of the 8 planes is listed in Table 3.5. The TRDs were not used in the trigger or off-line selection of the double Dalitz data. For a detailed description, see Reference [25].

Table 3.5: Positions of the eight transition radiation detector planes. The transverse dimensions of each plane is 2.1×2.1 m.

Plane	z Position (m)
TRD1	181.1
TRD2	181.4
TRD3	181.7
TRD4	182.0
TRD5	182.3
TRD6	182.7
TRD7	183.0
TRD8	183.3

3.3.4 Trigger Hodoscopes

The trigger used a prompt signal from a set of scintillator planes located immediately upstream of the calorimeter to identify events with charged particles that had passed through the spectrometer and would likely hit the calorimeter. These planes, known as VV', consisted of 1 cm thick paddles arranged so that paddles in different planes overlapped. There were two holes, 14 cm square, in each plane to allow the beam to pass through.

3.3.5 The Electromagnetic Calorimeter

The heart of the KTeV detector was the electromagnetic calorimeter made of 3100 individual cesium iodide (CsI) crystals. The calorimeter was designed to contain every electromagnetic shower and at the same time have sensitivity to minimum ionizing particles (MIP) like muons. This analysis relied on accurate measurements of both the energy and the position of electromagnetic clusters.

3.3.5.1 CsI Array

The calorimeter consisted of an array of 3100 pure CsI crystals. There were 2232 small crystals measuring $2.5 \times 2.5 \times 50$ cm, and 868 large crystals measuring $5 \times 5 \times 50$ cm.

The 50 cm of CsI was equivalent to 27 radiation lengths and to 1.4 nuclear interaction lengths. This means that practically all EM showers would be contained and roughly 30% of the charged pions would shower hadronically. Muons were usually minimum ionizing in the CsI and would deposit about 320 MeV of energy. The array was arranged with the small blocks in a square surrounding the two 15 cm square beam holes. The large blocks were then located around the inner square of small crystals. The whole array measured $1.9 \times 1.9 \times 0.5$ m. A sketch of the array is shown in Figure 3.5.

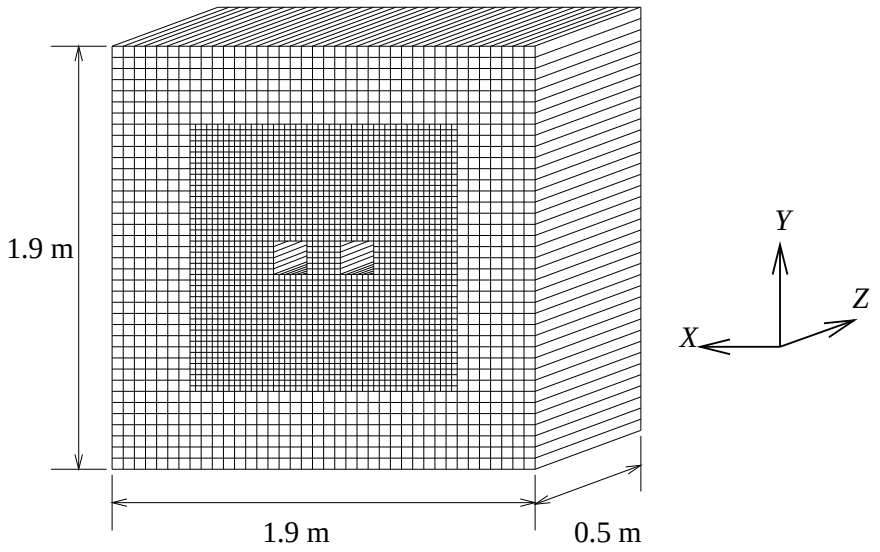

Figure 3.5: Sketch depicting the upstream face of the CsI array.

The CsI crystals produced roughly 20 photoelectrons per MeV of energy deposited [26]. The light output of the CsI crystals had two components, one fast and one slow. The fast component had a characteristic decay time of 25 ns while the decay time for the slow component was greater than 1 μs. The longitudinal response of each crystal was tuned by wrapping each block in 13 μm of mylar. Attached to each crystal was a clear rubber cookie which optically joined the back of the crystal to the front of the PMT. The cookie contained a filter to remove the slow component of the light. The output of each PMT went to a special DPMT circuit for digitization. These boards consisted of an ADC and a charge integrator and encoder (QIE). The circuit contained

a buffer that was only read out at the request of the Level 1 trigger, in which case every channel with any data was fed into a VME pipeline to await a final decision from the Level 2 trigger.

3.3.5.2 CsI Anti

At the outer edge of the calorimeter and overlapping the last row of crystals was the last photon veto, the CsI anti (CIA). The CIA had the same composition as the SAs described earlier. It was located at 185.2 m in z, with an outer and inner aperture of 2.20×2.20 m and 1.84×1.84 m, respectively.

3.3.5.3 Collar Anti

The collar anti (CA) was comprised of two square rings positioned around each of the two beam holes, overlapping the inner half of the first row of crystals. It was designed to veto events with particles hitting near the edge of the calorimeter and possibly losing a significant amount of energy down the beam hole. Each module was roughly 10 radiation lengths of tungsten-scintillator layers. Each CA had an outer dimension of 18×18 cm with an inner aperture of 15×15 cm. The nominal z position was 185.9 m.

3.3.5.4 Hadron Anti

The hadron anti (HA) was used to identify events with charged pions. It was located downstream of the calorimeter and behind a 10 cm lead wall (at $z = 189.0$ m). The wall was designed to absorb any electromagnetic showers escaping the CsI and to create hadronic showers from charged pions. These hadronic interactions were then detected by the HA, a plane of scintillator counters with holes for the beam to pass through. Behind the HA was located the first of three steel walls known collectively as the muon filters. This wall consisted of 1 m of steel and had holes for the neutral beam

to pass through. Its main purpose was to protect the HA from back-splash from the beam dump just downstream.

3.3.5.5 Back Anti

The back anti (BA) was designed to detect photons which escape by traveling through one of the beam holes in the calorimeter. It consisted of 30 lead–scintillator layers in three longitudinal modules (a total of 30 radiation lengths). This detector was not used in the trigger or off-line selection of the double Dalitz dataset.

3.3.5.6 Calorimeter Calibration

The calorimeter was calibrated in two steps. First, special runs, called **laser runs**, were performed every other day or so. A special laser supplied light to every crystal via optical fibers. The response of the DPMT to the intensity of the light was recorded. The intensity of the laser light could be adjusted to cover the full range of the electronics.

The second quantity which had to be calibrated was the conversion between the charge counted by the QIE and the energy in the crystal. This was done with the aid of electrons from $Ke3$ decays. Assuming that the spectrometer has been adequately calibrated, the charge-to-energy conversion can be tuned so that the energy in the calorimeter matches the momentum of the spectrometer, as it should for highly relativistic electrons.

The energy resolution of the calorimeter follows

$$\frac{\sigma(E)}{E} = 0.45\% \oplus \frac{2\%}{\sqrt{E}}, \tag{3.2}$$

where E is measured in GeV. For a typical cluster with 10 GeV of energy the resolution is about 0.8%. The constant factor is the result of noise, leakage, and non-uniformities while the energy dependent factor is due to the photostatistics of the scintillation light.

The position resolution is on the order of 1 mm in the small blocks and 2 mm in the large blocks.

3.3.6 The Muon System

The muon system consisted of a set of thick filters and scintillator counters. Beyond the BA was a 3 m steel block that acted as the neutral beam dump. Just downstream of that was MU2, a plane of counters similar in geometry to the VV' hodoscopes. Following MU2 was a final 1 m steel wall and two more planes of counters, known as MU3X and MU3Y. As the names imply, one was segmented in the x-direction while the other in the y-direction. The entire system accounted for nearly 30 nuclear interaction lengths, meaning that only about 0.5 % of charged pions could deposit energy in MU3.

3.4 The KTeV Trigger

The trigger had the awesome task of identifying interesting physics in a very intense environment. Approximately 1000 protons arrived simultaneously at the target every 20 ns. The resulting rate of kaon decays was about 1 MHz. The trigger was divided into three levels in order to effectively select and record the small number of interesting events while rejecting the vast amount of junk. The **level 1** (L1) trigger consisted of the fastest signals combined with simple logic and was capable of making a decision every RF bucket. Events selected by the L1 trigger were further filtered by the custom processors forming the **level 2** (L2) trigger. The slowest L2 processor took up to 2.5 μs during which the trigger was unable to consider other events. Events passing L2 were then digitized and read into buffers where they could be analyzed by the **level 3** (L3) software trigger. This program did a partial reconstruction of the event to allow for more stringent selection criteria. Events selected by L3 were written to one of ten DLT tape drives to be recorded for offline analysis.

Up to 32 different trigger configurations could be defined in the **data acquisition system** (DAQ). Sixteen of the configurations, known as **beam triggers**, were dedicated to the collection of physics modes while the other 16, known as **calibration triggers**, collected events for calibration techniques. The data used in this analysis were collected by the first trigger, called beam trigger 1 (B01) or the **2e-Nclus trigger**. The rest of this section describes in detail the elements used to define this trigger.

3.4.1 Level 1 Trigger

The level 1 trigger was synchronized with the 19 ns RF signal, meaning that only the fastest signals could be used to make a decision. These included signals from the Etot system, the VV' hodoscopes, the drift chambers, and the vetoes. The sources were synchronized with cable delays and fed into programmable logic modules which made the decision to accept or reject once every bucket. The level 1 requirement for the 2e-nclus trigger can be represented symbolically as

$$\mathrm{L1} = \mathrm{VV'} \cdot \mathrm{ETOT} \cdot \mathrm{DCOR} \cdot \mathrm{!VETO} \tag{3.3}$$

Each of these elements will be discussed in turn.

3.4.1.1 VV′

The trigger hodoscopes were designed to provide prompt evidence of at least two charged particles. Most 2–track events will hit two paddles in each view. To allow for some level of inefficiency, the trigger required that at least two paddles were hit in one view and at least one paddle in the other.

3.4.1.2 ETOT

The Etot system received signals from the PMT's on the calorimeter. The system performed an analog sum of all 3100 channels representing the total in-time energy. A

discriminator selected between 4 different thresholds. Trigger 1 used the third threshold, with a nominal value of 28 GeV. The same system was responsible for producing an HCC bit if any channel was above 1 GeV. These bits were used by the Level 2 HCC processor.

3.4.1.3 DCOR

The DCOR requirement was based on signals from the first two drift chambers. The maximum drift time of 200 ns limits the amount of information that can be gleaned from the chambers. However, when the drift time to a wire in one plane is large, the time in the complementary plane is small. The logical OR of complementary wires can then be created fast enough for L1. The wires were grouped into 16 "paddles" which were then OR'd together. Trigger 1 required at least one hit in each of the four upstream views.

3.4.1.4 !VETO

The rest of the L1 definition consisted of various vetoes. The ring counters had a nominal threshold of 500 MeV. RC8 was only used in veto for runs 8280 through 8283. The spectrometer antis were set to 400 MeV, and SA3 was only used from 8088 to 8279. The calorimeter anti and collar anti had thresholds of 400 MeV and 13 GeV, respectively. The hadron anti threshold had a nominal value of 5.6 MeV.

The muon system was used in veto for the first half of winter97, runs 8088 to 8576. Events were vetoed if there were any hits in the mu2 plane.

3.4.2 Level 2 Trigger

The Level 2 (L2) trigger received events from L1 at a rate of 100 kHz and had to reduce that by a factor of 10. The time for L2 to make a decision was about 3 μs, during which time the output of L1 was inhibited. A state machine received signals from the

various L2 processors and made decisions to either accept or reject the event. The level 2 definition of the 2e-Nclus trigger is

$$\mathrm{L2} = \mathrm{HCC} \cdot \mathrm{HCX} \cdot \mathrm{HCY} \tag{3.4}$$

The two systems responsible for forming these signals were the **hardware cluster counter** (HCC) and the **hit counting**, which was performed in both x (HCX) and y (HCY) views of the drift chambers.

3.4.2.1 Hardware Cluster Counter

The HCC element referred to the hardware cluster counting processor associated with the Etot system. The inputs to the processor were the HCC bits set at L1. Clusters were found by counting the number of right-hand turns required to trace the outline of neighboring blocks associated with the HCC bits. The algorithm considered every possible 2×2 array of crystals. The output was a count between 1 and 8 or an overflow. Trigger 1 required that the HCC find at least 4 clusters.

3.4.2.2 Hit Counting

The HCX and HCY requirements referred to the outcome of the hit counting performed by specialized modules developed and built at the University of Colorado. For a detailed description of the design and usage of the Level 2 hit counting processor, see Reference [27]. The modules came in two flavors; banana modules monitored the four upstream views while kumquat modules looked at the four downstream views. The following sections give a brief overview of each type.

Kumquats There were 35 kumquat modules which received signals from downstream views, DC3X, DC3Y, DC4X, and DC4Y. Hits that fell in a 220 ns window were accepted and counted. A mask was formed with bits set for every wire with an in-time hit. The number of tracks was counted as $N - 1$ where N is the number of continuous

bits. Isolated bits were also counted as one track. To count hits in an entire view, signals form edge wires had to be sent to multiple modules. The trigger required at least two hits in both DC3Y and DC4Y.

Bananas The banana modules had a similar purpose, but could better discriminate against out of time pairs. They did this by using TDC's to measure the correlation of times between pairs of wires. The TDC window was 464 ns long to allow for very early or late hits that may accompany an isolated single. The pair of TDC values were compared to a lookup table to determine whether the hits arrived in-time or not. The accepted in-time region could be changed by loading different tables into the modules. The shape used in E799 is shown in Figure 3.6. A mask was formed by setting bits for accepted in-time hits and the counting algorithm was identical to that used by the kumquats. The trigger required at least 1 hit in DC2X and at least 2 hits in either DC1Y or DC2Y with at least 1 hit in the other.

3.4.3 Data Acquisition

The Level 3 (L3) trigger resided in software and filtered data satisfying the trigger at levels 1 and 2. Events accepted by L2 were first digitized and then feed into buffers. The output of ADC's, along with the TDC values from the drift chambers, were read out on one of six streams. The first four streams contained all of the calorimeter data, which was sparsified, while the last two contained everything else. The buffers contained enough memory for an entire spill, about 4.6 GB in 97 and 4.9 GB in 99. The contents of the buffers were than processed by four multi-cpu Silicon Graphics Challenge computers. Each computer contained four planes which received events in parallel. The processors performed basic reconstruction to identify events with two electrons from a common vertex. The details of the algorithm are given in the Chapter 5 after general reconstruction methods are discussed. Events that satisfied the L3 trigger were sent to tape drives to be written on DLT tapes. These were high density digital tapes with a

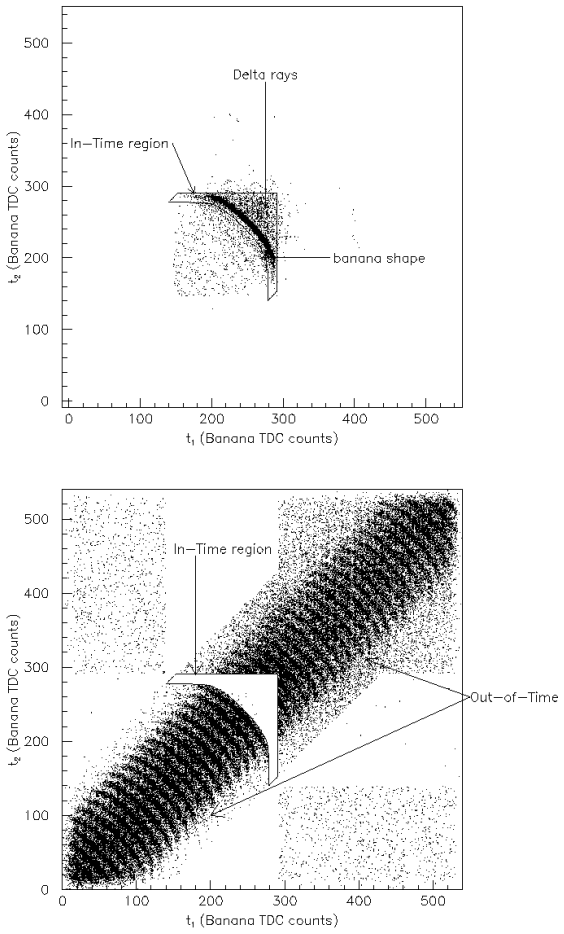

Figure 3.6: Banana TDC counts t_1 vs. t_2. The top figure contains hits used on tracks in good 2–track events. The lower figure contains all rejected hits. The in-time pair region is shown in both.

capacity between 10 and 15 GB. The L3 filter also accepted a small number of events randomly to insure a sample of unbiased data for later trigger studies.

Chapter 4

The KTeV Monte Carlo Simulation

The branching ratio analysis requires a link between the number of observed events of a given mode and the total number of such events. The ratio of these two quantities defines the average acceptance. If the acceptance were unity, all decays of a given type would be observed. In general the average acceptance is much less than unity. The three main contributors to the acceptance are detector geometry, detector response, and selection cuts made either at the trigger level or offline. To calculate the acceptance, a Monte Carlo (MC) program generates kaon decays and simulates these effects. As it turns out, for many of the $K_L \to \pi^0\pi^0\pi^0$ modes there is a 50-50 chance of detecting any given final state particle due mainly to geometry. This means that an order-of-magnitude estimate of the acceptance is simply 2^{-n}, where n is the number of final state particles.

The main parts of the KTeV simulation are kaon selection, decay generation, detector simulation, accidental activity modeling, and trigger simulation. The end result is a digitized dataset corresponding to the output of the actual L3 trigger. All of the code used to implement these effects was written in Fortran and taken together is known as **ktevmc**.

4.1 Simulating Kaons

The simulation begins with the selection of a kaon energy and direction. We do not simulate the proton or kaon beam, only individual kaon decays. The energy spectrum of the kaons produced on the KTeV target follows closely the Malensek spectrum measured from charged kaons [28]. The spectrum, in both momentum and angle, has the following parametric form,

$$\frac{d^2N}{dpd\Omega} = K\ p\ \frac{(1-x)^A(1+5\mathrm{e}^{-Dx})}{(1+p_t^2/M^2)^4}, \tag{4.1}$$

where x is the ratio of the kaon momentum to the proton momentum, and A, D, K, and M are parameters extracted from the charged kaon data. The production of neutral kaons is then assumed to be

$$\sigma(K^0) \approx \frac{1}{2}[\sigma(K^+) + \sigma(K^-)], \tag{4.2a}$$

$$\sigma(\bar{K}^0) \approx \sigma(K^-), \tag{4.2b}$$

based on quark composition. The chosen momentum is required to be between 20 and 220 GeV/c (the acceptance is practically zero outside this range). Figure 4.1 shows the distribution of generated kaon momenta.

Once a suitable energy and direction are chosen, the kaon is allowed to propagate some distance before decaying. The characteristic decay length for a 100 GeV K_L is roughly 3 km, making the distribution of decay positions nearly uniform. Only kaons decaying between $z = 90$ m and $z = 160$ m are further considered. If the kaon hits some part of the detector before decaying it may scatter or be absorbed. If it does reach the randomly chosen decay point, the $K_L \to \pi^0\pi^0\pi^0$ generator is called to produce the three pions. The π^0 is assumed to decay to two photons unless explicitly asked to decay in some other way. The next section will discuss the various methods for generating decay parameters.

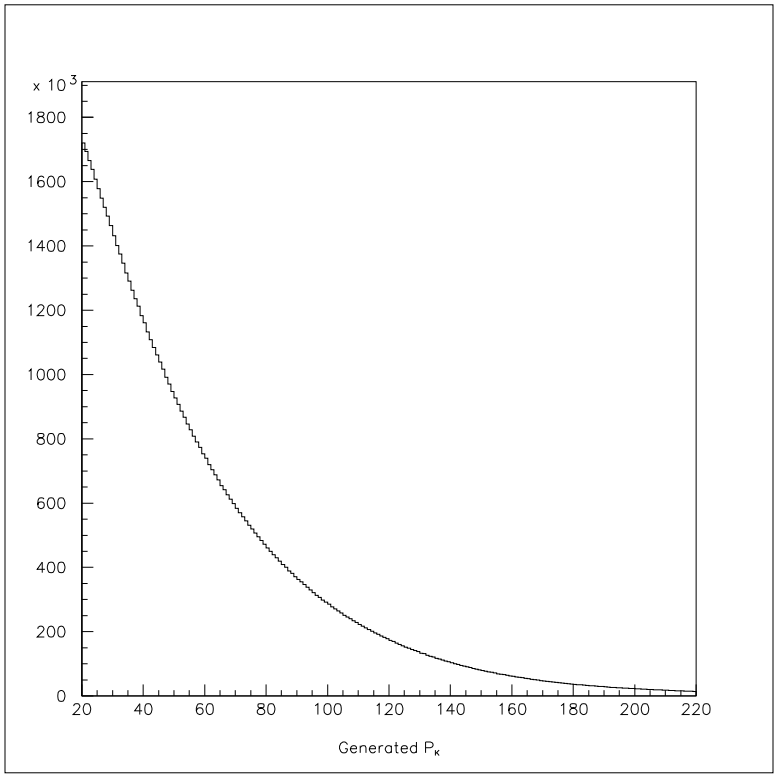

Figure 4.1: Distribution of the kaon momentum in the generated range of 20 to 220 GeV.

4.2 Decay Generators

The following sections provide some details on the individual decay generators. The generators all have the same purpose, generate a set of kinematically allowed final state 4–momenta.

4.2.1 The $K_L \to \pi^0\pi^0\pi^0$ Generator

The decay $K_L \to \pi^0\pi^0\pi^0$ is generated uniformly over the allowed phase space, in agreement with the measured form factor. The 3–body phase space is typically described by two Dalitz plot variables, such as the invariant mass of two of the three possible $2\pi^0$ states. Such a Dalitz plot is shown in Figure 4.2. The lower and upper extremes correspond to $(2M_{\pi^0})^2$ and $(M_K - M_{\pi^0})^2$, respectively. After boosting into the lab frame, a π^0 has an average energy of roughly 30 GeV. At this energy, the decay

length is less than $6\,\mu$m, which is orders of magnitude below the vertex resolution. For this reason, all π^0's are assumed to decay instantaneously.

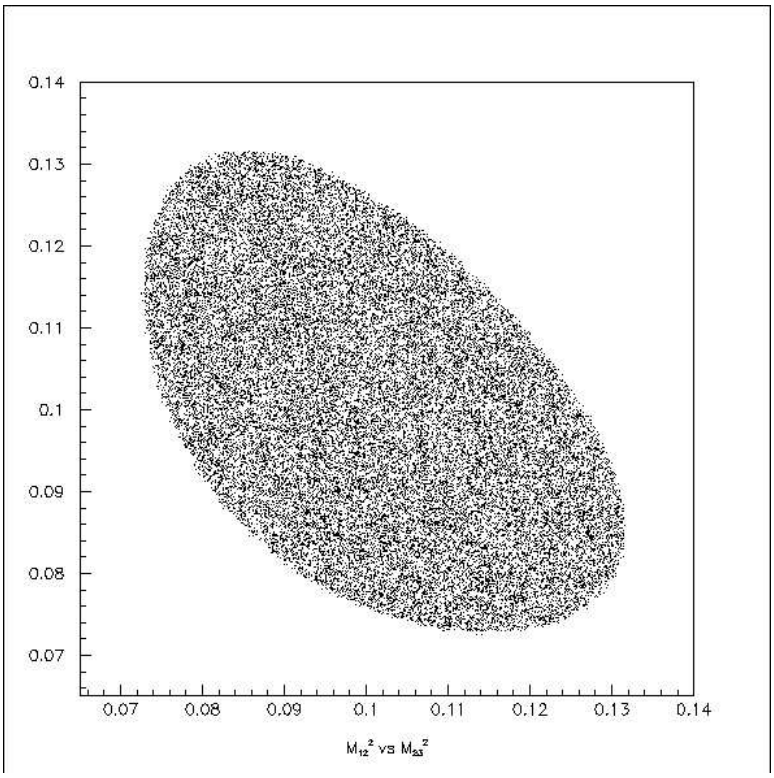

Figure 4.2: Dalitz plot for $K_L \to \pi^0\pi^0\pi^0$ decays at generation.

4.2.2 The $\pi^0 \to \gamma\gamma$ Generator

This 2–body decay is simple to generate, requiring only a direction be specified. The photons are of course back-to-back in the pion rest frame. Let one of the photons point in the direction defined by a polar angle θ and azimuthal angle ϕ. The decay is isotropic so $\cos\theta$ and ϕ should be uncorrelated and distributed uniformly over their respective ranges. The momenta of the photons are then boosted into the lab frame.

4.2.3 The $\pi^0 \to e^+e^-\gamma(\gamma)$ Generator

The single Dalitz generator produces both $\pi^0 \to e^+e^-\gamma$ decays and $\pi^0 \to e^+e^-\gamma\gamma$ decays with the predicted probability of radiation. A random number is thrown event-by-event to decide which type of decay to generate. The $\pi^0 \to e^+e^-\gamma$ generator uses the tree level matrix element with first order radiative corrections. The radiative decays are based only on the tree-level $\pi^0 \to e^+e^-\gamma\gamma$ matrix element and are generated with the invariant mass of the two-photon system above 1 MeV. The probability of radiation at this cutoff is 16.2 %.

4.2.3.1 $\pi^0 \to e^+e^-\gamma$

The differential partial width, with respect to the two–photon width, was presented in Chapter 2. The generator first randomly selects values for the two Dalitz variables, x and y. Since the distribution of the invariant mass of the e^+e^- pair is strongly peaked near threshold, the x variable is chosen according to

$$\frac{dN}{dx} = Cx^{-0.8}. \tag{4.3}$$

Then y is selected from a uniform distribution ranging from $-(1-x)$ to $(1-x)$.

The generated point is given a weight equal to its contribution to the differential partial width (normalized to the two-photon width) modified by a factor to correct for the biased selection of x values. This requires an evaluation of the tree-level matrix element, the radiative corrections, and the invariant phase space element. The radiative corrections were tabulated on a fine grid of x-y pairs. Each weight is converted to a likelihood by normalizing to the maximum weight. Another random number is picked to decide whether to keep the generated point or to discard it. If the thrown point is kept, the momenta receive a random rotation in the pion rest frame before being boosted into the lab frame. Otherwise a new point is thrown. The efficiency of the generator is defined as the number of events passing the hit-or-miss divided by the total number of

events thrown. The efficiency for the $\pi^0 \to e^+e^-\gamma$ generator is 13 %.

4.2.3.2 $\pi^0 \to e^+e^-\gamma\gamma$

The radiative single Dalitz decay is generated with the two-photon invariant mass above $1\,\mathrm{MeV/c^2}$. The 5-dimensional phase space is complicated by the soft energy spectrum of the radiated photon and the Bose symmetry of the two photons. The variables x and x_γ are thrown according to power laws, ϕ is thrown uniformly, and y and y_γ are thrown in a mix of power laws and exponentials. This variables are the same as those describing the double Dalitz final state with the e^+e^- composing one pair and the two photons the other. Again, the differential width is evaluated at the thrown point and a weight is compared to a random number in a hit-or-miss. Events that pass are rotated and boosted into the lab frame. The efficiency of the $\pi^0 \to e^+e^-\gamma\gamma$ generator is only 2.5 % due to the complicated phase space.

4.2.4 The $\pi^0 \to e^+e^-e^+e^-(\gamma)$ Generator

The double Dalitz generator is very similar to the single Dalitz generator in that both non-radiative events (with radiative corrections) and radiative events are produced in unison. Since there is at most one photon in the final state, the cutoff is expressed in terms of the photon energy in the pion CM frame. The $\pi^0 \to e^+e^-e^+e^-\gamma$ generator only throws photons with energies $E^*_\gamma > 100\,\mathrm{keV}$. The probability of radiation corresponding to this cutoff is 18.0 %. The distribution of $x_{4e} = M^2_{4e}/M^2$ at generation, including both types of events, is shown in Figure 4.3.

4.2.4.1 $\pi^0 \to e^+e^-e^+e^-$

The 5 phase space variables defined in Appendix A must be randomly chosen. The generator selects the two x values according to the same power law spectrum used by the $\pi^0 \to e^+e^-\gamma$ generator. The other variables are thrown uniformly. Both the tree-

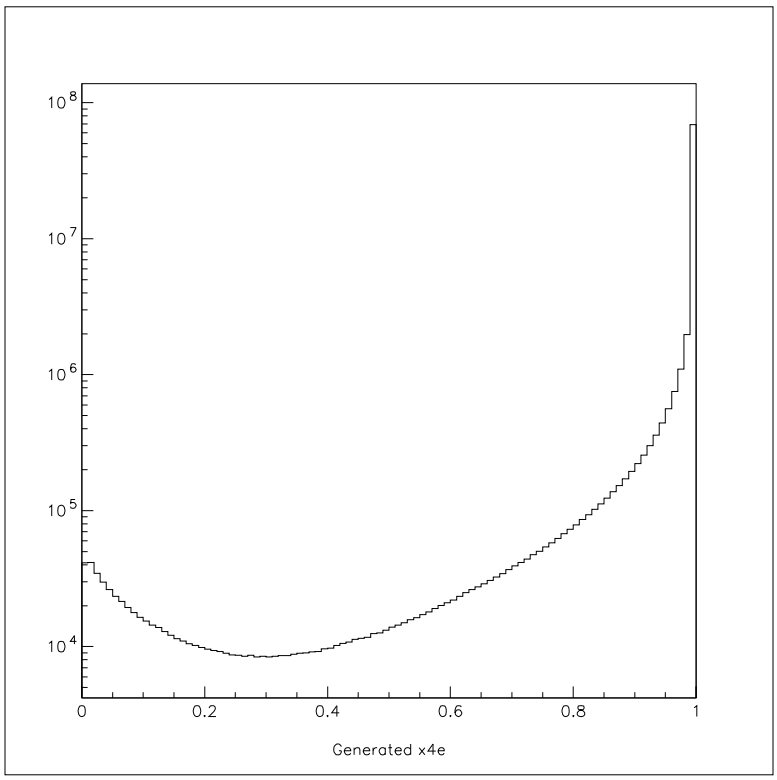

Figure 4.3: Distribution of generated x_{4e} values from both $\pi^0 \to e^+e^-e^+e^-$ and $\pi^0 \to e^+e^-e^+e^-\gamma$ events.

level amplitude and the radiative corrections are evaluated at the thrown point. The differential width is again used by a hit-or-miss selection. Events passing the selection are randomly rotated and then boosted into the lab frame.

4.2.4.2 $\pi^0 \to e^+e^-e^+e^-\gamma$

With a probability of roughly 18 %, a photon is generated with an energy greater than 100 keV. For every thrown set of phase space variables, spinors and polarization vectors are calculated and combined with Dirac matrices to form the complex amplitude of each diagram.

The distributions of the five double Dalitz phase space variables are shown in Figures 4.4 to 4.6. The plots are made from all $\pi^0 \to e^+e^-e^+e^-$ events plus $\pi^0 \to e^+e^-e^+e^-\gamma$ events with $x_{4e} > 0.9$. These distributions are based on a particular choice

of the pairing electrons with positrons. Of the two possible pairings, the one which minimizes the product of invariant masses dominates the matrix element and is the pairing used to make the plots. Furthermore, the pair within the dominant pairing with the smaller invariant mass is designated as the a pair while the other is the b pair.

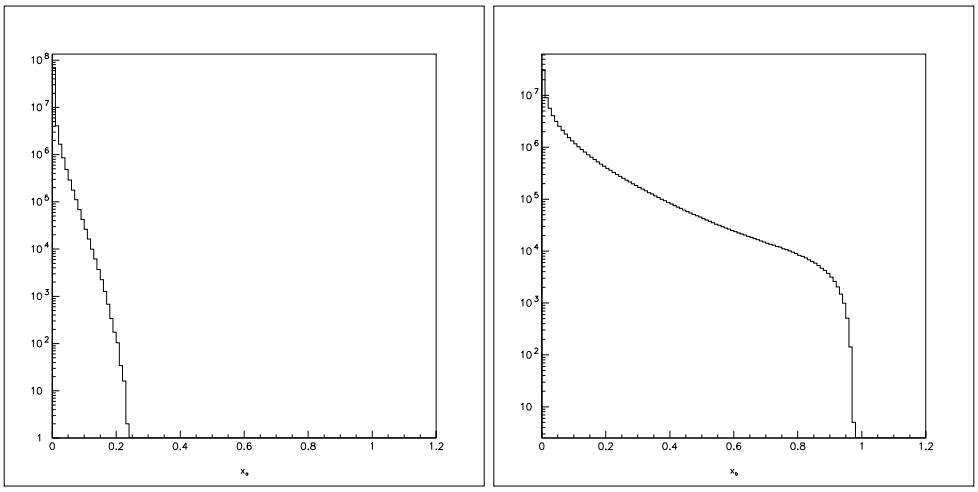

Figure 4.4: Distribution of x_a (left) and x_b (right). The plots include all $\pi^0 \to e^+e^-e^+e^-$ events plus $\pi^0 \to e^+e^-e^+e^-\gamma$ events with $x_{4e} > 0.9$.

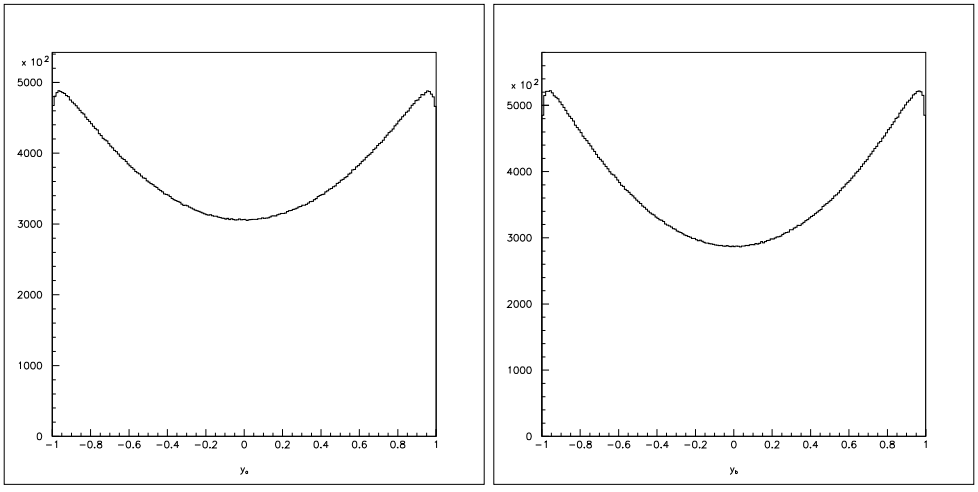

Figure 4.5: Distribution of y_a (left) and y_b (right). The plots include all $\pi^0 \to e^+e^-e^+e^-$ events plus $\pi^0 \to e^+e^-e^+e^-\gamma$ events with $x_{4e} > 0.9$.

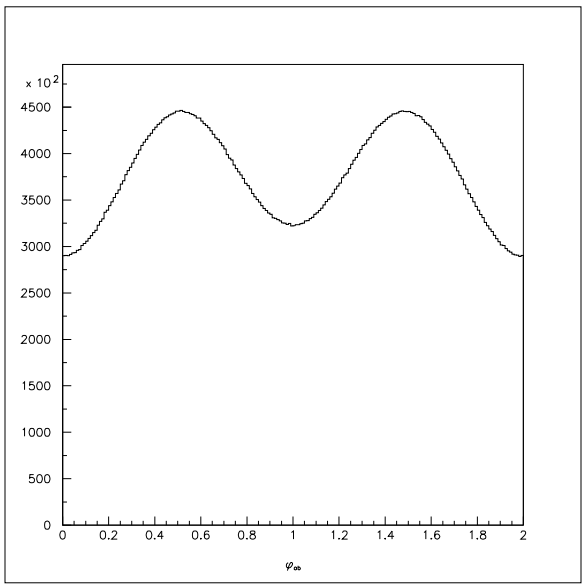

Figure 4.6: Distribution of ϕ. The plots include all $\pi^0 \to e^+e^-e^+e^-$ events plus $\pi^0 \to e^+e^-e^+e^-\gamma$ events with $x_{4e} > 0.9$.

4.3 Particle Tracing

After the generation stage there will be a list of decay particles, all located at the kaon decay point and with initial energies and directions selected by the generators. Each particle in turn is traced through regions binned in z. To begin with, a particle is propagated through the remainder of the vacuum chamber. At the z position of each RC, the transverse position of the particle is compared to the dimensions of the veto. If a particle strikes a veto, it is considered lost and all of its energy is deposited in the RC. Particles that do not hit an RC are traced to the vacuum window where they may interact. The possible interactions that may occur in material are covered below.

Beyond the vacuum window, particles are traced through each element of the spectrometer. Interactions with the chambers and helium bags are simulated. The analysis magnet is modeled as a plane containing the majority of the field with small fringe fields extending away from the plane. As charged particles travel from DC1 to DC2 and again from DC3 to DC4, their trajectories are adjusted to account for both

the Earth's field and the fringe fields. Between DC2 and DC3, charged particles are traced through a 3-D model of the field in 1 in bins in z. If a particle hits an SA veto, it deposits all of its energy and is no longer traced. Otherwise, its transverse position at each chamber plane is recorded. The response of the chambers will be discussed below in the digitization section.

The MC treats the TRD system as dead material and does not simulate the response of these detectors. Since the present analysis does not rely on the TRD's, this treatment is more than adequate. Particles are next traced to the trigger hodoscopes. Charged particles will initiate a response in any paddles which they may hit. A TDC value is calculated for every hit.

Before reaching the CsI calorimeter, each particle is traced to the CIA and the CA which are handled like the other vetoes. The tracing of electrons and photons ends at the calorimeter face. The position and energy of the particles are recorded for use by the digitization described below. The detectors downstream of the calorimeter are not simulated for the events used in this analysis.

The following sections provide details of the different ways electrons and photons can interact with material in the detector.

4.3.1 Particle Interactions

The electrons and photons in the double Dalitz final state interact electromagnetically. The electrons have a mean momentum of roughly 10 GeV/c, making multiple scattering a significant source for measurement error. Another source is energy lost to radiation, especially when it occurs upstream of the analysis magnet. External radiation is the dominant background in the radiative double Dalitz analysis. As for the photons, they can convert to electron-positron pairs, providing a background for any mode with Dalitz pairs. The amount of material in the detector is summarized in Table 4.1. For a complete discussion of the amount of material in the detector, see Reference [29].

Table 4.1: The amount of material in the various components of the KTeV detector. The amount of air in the Air Gap changed with time; 0.104 % in win97, 0.110 % in sum97, and 0.074 % in 99. The beam region of a TRD chamber is only 0.51 %.

Component	Material	X/X_0 (%)
Vacuum Window	mylar, kevlar	0.156
Air Gap	air	0.074
DC1 Upstream	mylar, wires	0.050
DC1 Downstream	mylar, wires	0.045
He Bag 2	helium	0.162
DC2 Upstream	mylar, wires	0.045
DC2 Downstream	mylar, wires	0.045
He Bag 3a	helium	0.119
He Bag 3b	helium	0.119
DC3	mylar, wires	0.090
He Bag 4	helium	0.156
DC4	mylar, wires	0.090
TRD (1 of 8)	TRD stuff	1.760
VV' (1 of 2)	scintillator	1.335

Any bremsstrahlung photons or conversion electrons that are produced due to interactions are traced through the remaining parts of the detector. Typically, interactions downstream of the magnet are less important since any particles created in an interaction will travel in the same direction, with the same energy, as the original and will be indistinguishable.

4.3.1.1 Multiple Scattering and δ–Rays

Multiple scattering refers to the complicated path a charged particle takes when passing through matter. Constant Coulombic interactions with atomic nuclei can significantly change the direction of slow moving particles. The effect has been well studied and the KTeV Monte Carlo applies the standard formula. For a given amount of material X/X_0, the distribution of probabilities for scattering by an angle θ is approximately Gaussian with a mean of zero and a width given by

$$\theta_0 = \frac{13.6 \text{ MeV}}{\beta cp} z \sqrt{X/X_0} \left[1 + 0.038 \ln \left(X/X_0\right)\right], \tag{4.4}$$

where βc, p, and z are the velocity, momentum, and charge of the particle. As charged particles exit material their direction is altered by a small angle chosen from this Molière distribution.

A δ–rays occurs when the charged particle comes close enough to an atomic electron to knock it loose. These freed electrons can occasionally travel significant distances and even produce ionization in the drift chambers. The probability of a charged particle with energy E producing a δ–ray with an energy T (greater than a cutoff T_0) is

$$P(T > T_0) = \left[15400 \frac{\text{keV cm}^3}{\text{gram}}\right] \frac{Z\rho X}{A} \frac{1}{E}, \tag{4.5}$$

where Z and A are the atomic number and weight of the material and ρ and X are the density and amount. During the DC simulation described below, δ–rays are simulated and may produce additional hits.

4.3.1.2 Bremsstrahlung

Electrons passing through matter are allowed to radiate a photon of energy $k > k_0$ with a probability given by

$$P(k > k_0) = \frac{X/X_0}{1+18\xi}\left\{2(1+12\xi)\left[\frac{k_0}{E_0} - \ln\frac{k_0}{E_0} - 1\right] + 9\xi\left[1 - \left(\frac{k_0}{E_0}\right)^2\right]\right\}, \quad (4.6)$$

where $\xi = \ln(183/Z_{\text{eff}}^{1/3})$. The probability of radiation for various parts of the KTeV detector ranges from about 1.3 % in the vacuum window to 14.8 % in a TRD chamber. If this probability is satisfied, the energy and direction of the photon are chosen from the distribution of the full cross section. Additional details can be found in Reference [27].

4.3.1.3 Photon Conversion

The probability of a photon converting as it passes through an amount of material X/X_0 is

$$P(\gamma \to e^+e^-) = 1 - \mathrm{e}^{-\frac{7}{9}X/X_0} \approx \frac{7}{9}X/X_0. \quad (4.7)$$

When a conversion does occur, the energy of the photon is divided according to the Bethe-Heitler spectrum

$$P(e_+, e_-) \propto e_+^2 + e_-^2 + \left(\frac{2}{3} - \frac{1}{9Z}\right)e_+e_-, \quad (4.8)$$

where $e_\pm = E_\pm/E_\gamma$ and $Z \sim 3.741$. The electron and positron travel in a direction equal to that of the original photon plus a small angle. The simulation of the distribution of angles is discussed elsewhere [30].

4.4 Digitization

The previous section discussed the simulation of the decay particles and matter interactions. This section will address the simulation of the detector and its response to interactions with the particles. For example, a particle is traced to a veto and deposits all of its energy. The energy is first smeared by some estimate of the resolution of the

detector in question and then converted to an ADC count for the module which was hit. A similar procedure is used to obtain the TDC count of hits in the VV' hodoscope planes. In this case, the position of the particle in a paddle relative to the photomultiplier tube (PMT) is converted to a time based on the characteristics of the scintillator. The time is then converted to a TDC value for that paddle. The simulation of the drift chambers and the calorimeter is more complex and will be described separately below. The response of the TRD chambers is not simulated at all.

4.4.1 Drift Chamber Simulation

The DC simulation takes the intercept of each charged particle at each chamber plane and creates a signal on the closest wire. This procedure starts by smearing the distance to the wire by a Gaussian resolution function. The estimate of the resolution contains two contributions: a global term based on the chamber plane, and a local term based on the region of the chamber that the hit is in. Both terms were estimated independently in multiple run periods.

Next, the high SOD tail is simulated. The SOD, or Sum-of-Distances, is defined in the Chapter 5. The effect is partly due to discrete ionization of the gas. Ideally, the first ionization electrons arriving at a sense wire were produced at the point of closest approach. However, due to the discreteness of ionization, the point of closest approach might not produce ionization. In that case, the distance measured will be bigger than the actual distance. To simulate this effect, a small distance is added to the modified hit distance. This correction is based on chamber plane, region, and run period.

The modified hit distance can now be converted to a TDC value, using the inverse of the $X(t)$ maps described in Chapter 3. The conversion formula is $t = t_0 + md$, where t_0 and m are determined from the map for the appropriate plane and run period.

The digitized hit is subjected to two additional effects: hit inefficiency and δ–rays. The probability of missing a hit is measured in the data as a function of plane, wire,

and run period. Hits are removed from the simulation based on this probability. There is also some probability of producing δ–rays which may produce an early signal in the cell in which they were produced, or create additional hits in nearby cells. If an extra hit is produced, its distance is smeared and converted to a TDC.

One additional effect is simulated after all hits in every plane have been dealt with in the above manner. Each hit produces a signal pulse with a width of about 45 ns. Starting with the earliest hit on a given wire, any other hits within this window would be obscured and are therefore deleted. This is then repeated for the next un-obscured hit on that wire and so-on.

4.4.2 Calorimeter Simulation

The CsI simulation produces clusters based on the location and energy of the particles traced to the face of the calorimeter and creates a DPMT signal for any crystal receiving energy above threshold. First the energy of the particle responsible for the cluster is smeared according to a resolution function based on seed block size, position, and incident energy.

Then the modified energy is distributed amongst nearby crystals with the aid of GEANT shower libraries. These libraries contain shower profiles, both transverse and longitudinal, binned by particle type, energy, and position within the seed block. The appropriate profiles are then rotated to create a 3-D energy distribution which is then scaled to match the energy deposited. Some corrections are applied to account for beam hole effects and transverse energy tails.

Finally, the response of the DPMT associated with each crystal above threshold is simulated. As described in Chapter 3, the DPMTs were capable of reading out every bucket, or slice. The incident energy is typically distributed over several slices. Lookup tables of actual pulse shapes are used to estimate the amount of energy in each slice.

4.5 Accidental Activity

The simulation up to this point has involved only the decay products of a single kaon decay. In reality there can be additional particles, coming either from the target or from other decays, which are uncorrelated with the kaon and its daughter particles. In order to model this noise, real data collected by the accidental trigger (described in Chapter 3) is combined with the simulated data in a process known as **accidental overlaying**. ADC values are simply added channel-by-channel. For TDC channels, accidental hits are added to the list of TDC counts and a pulse width is simulated to see if any of the preexisting hits would be obscured by the accidental hit.

The accidental overlays also provide a way to simulate time dependences on the scale of one spill. Assuming that the detector conditions are stable, spill-by-spill variations are due entirely to the beam intensity. The accidental trigger collected events at a rate roughly proportional to the beam intensity and therefore should reflect these short term variations. The following section describes how this is accomplished.

4.5.1 Simulating Time Dependence

The time dependence is divided into four scales. The largest time scale corresponds to the data taking periods; win97, sum97, and 99. Most beam line and detector geometry is stable on this scale. Next come large groups of runs or individual runs for which various calibration constants are valid. The smallest time period is the spill.

The number of $K_L \to \pi^0\pi^0\pi^0_D$ events in every run was determined and compared to the total for all runs to determine a run-by-run yield. This ratio is used by the Monte Carlo to decide how many events to generate in a given run, such that the distribution of events by run number is similar to that in the data. Events generated in a given run receive accidental overlays from that run.

The spill number of the accidental is inherited by the event. It is therefore im-

portant to have the proper spill-by-spill distribution of accidental overlays. This is accomplished in two ways: scaling the number of accidentals in each run and identifying **bad spills** in which no events will be generated. The number of accidental events in each run is scaled such that the ratio of accidental events to $K_L \to \pi^0\pi^0\pi^0_D$ events is roughly equal. Also, spills during which brief, but catastrophic, detector or trigger failures occurred are not simulated.

4.6 Trigger Simulation

The final stage of the Monte Carlo generator is the simulation of the three levels of the trigger. The hardware logic is emulated using the TDC or DPMT signals produced by various detectors. The actual online L3 software filter code is reused by the MC to simulate its efficiency. For this analysis, the requirements of the 2e-nclus trigger were imposed.

The L1 trigger uses the ADC values from the vetoes to see if any channels were above trigger threshold. TDC counts from VV' are used to check for sufficiently many in-time hits. The TDC counts from the drift chambers are combined according the DCOR logic described in Chapter 3. Lastly, the DPMT data are combined to form an ETOT sum which is compared to the threshold value.

At level 2, the HCC is simulated using inputs from the ETOT simulation. The hit counting, performed by the bananas and kumquats, is emulated using the online monitoring code.

Events that pass the L1 and L2 simulation are processed by the L3 filter code. The selection criteria are identical to what was used online and the calibration constants that are used are the same as what was used online. The final output of the Monte Carlo are the events that pass the L3 filter. The data format is identical to that of the actual data with the addition of extra bank containing information about the generated quantities.

Chapter 5

Event Reconstruction

The physical hardware responsible for producing the KTeV dataset was described in Chapter 3. This raw data, consisting mainly of ADC and TDC counts, must be converted into meaningful quantities like the momenta and positions of final state particles. This chapter will cover the software algorithms used to identify and to reconstruct events of interest. The procedures outlined below were all implemented in Fortran and bundled together, along with many other routines, in a package known simply as **ktevana**.

The main goal is to find $K_L \to \pi^0\pi^0\pi^0$ events in which one of the pions decayed as $\pi^0 \to e^+e^-e^+e^-$ and the other two decay as $\pi^0 \to \gamma\gamma$. Therefore, the selection is designed to effectively identify 8-body final states consisting of four electrons plus four photons. To do so, one must reconstruct the decay point of the kaon using the tracks left by the electrons in the drift chambers. The positions of tracks in the chambers and the energies of clusters in the calorimeter, along with the decay vertex, allow for a complete reconstruction of the final state 4-momenta. The first part of this chapter will describe the details involved in this event reconstruction.

The chapter will close with a description of the preliminary stages of software filtering, including the **L3 filter** discussed briefly in Chapter 3. The 2e-nclus trigger data was then processed by the **2e-nclus crunch** to reduce the data volume and separate events into distinct physics modes. This analysis is based on the 4–track output of the crunch, which was filtered one last time by the **4–track split** to select events with

either 8 or 9 clusters. The $\pi^0 \to e^+e^-e^+e^-$ events are from the "4e8c" sample while the $\pi^0 \to e^+e^-e^+e^-\gamma$ events are contained in the "4e9c" sample.

5.1 Reconstruction

A good $K_L \to \pi^0\pi^0\pi^0_{DD}$ event should have a well defined 4–track vertex and eight clean and symmetric clusters, four of which have tracks pointing at them from the spectrometer. The steps involved in finding the vertex and identifying the photons include constructing tracks in the spectrometer, identifying clusters in the calorimeter, and performing a global fit to connect four tracks from a single point to four clusters. Clusters without associated tracks are assumed to be due to photons.

Figure 5.1 shows the KTeV event display for a 4e8c event. The left-hand side of the plot provides information about track momenta and cluster energies. The upper right depicts the front face of the calorimeter, with cluster energies indicated by color. Below that are two views of the full KTeV detector, the first one from above and the second from the side. The spectrometer occupies the central right part of each view. The tracks are seen to bend in the x-view and not in the y-view.

This event is typical of Dalitz events. The two pairs are produced with such small opening angles that only two tracks are evident in the y-view. It is only because of the kick from the analysis magnet that the two pairs are detectable at all. The tracking code, originally optimized for $K\pi 2$ decays, had to be modified to allow multiple tracks to share the same hits in the y-view, and to allow multiple x-tracks to be paired with the same y-track.

5.1.1 Tracking

The first task is to find charged tracks in the spectrometer. The general idea is to find hits in the drift chambers that all lie on a line. Because the chambers measure x and y positions independently, we can only find track candidates in the x-z and y-z

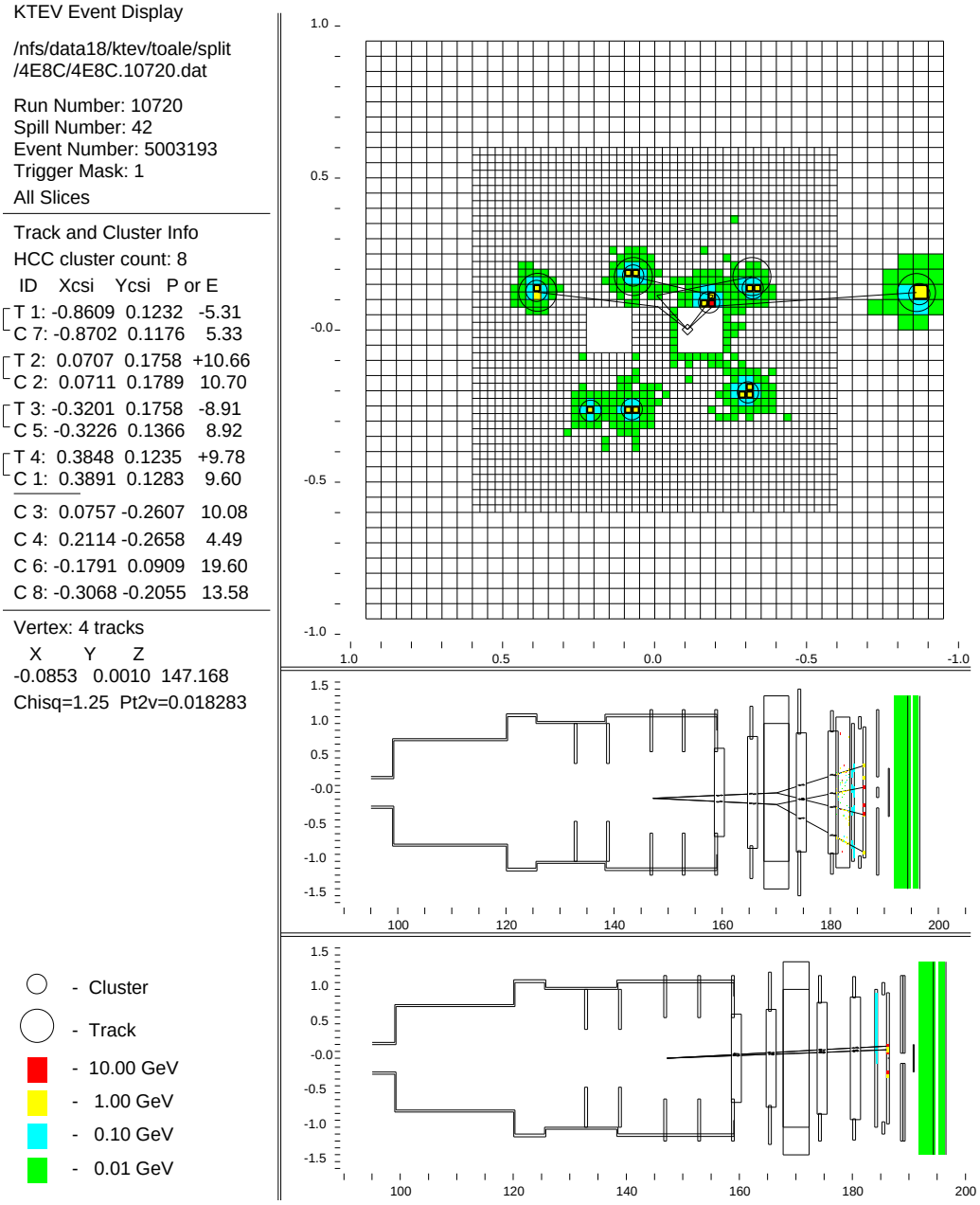

Figure 5.1: KTeV event display for a 4e8c candidate event.

planes separately. The necessary steps include finding hits in the chambers, identifying in-time pairs, and then connecting pairs to form tracks.

5.1.1.1 Hit, Pairs and Sum-of-Distances

When a charged particle passes through a drift chamber it will, under ideal conditions, produce a signal on the sense wire of any cell it passes through. This signal is referred to as a **hit**. If the corrected hit time falls within the TDC in-time window ($115 - 350\,\text{ns}$) the TDC count is converted to a drift distance using the $X(t)$ maps discussed earlier. If there are multiple in-time hits on a given wire, the earliest hit is used and all other hits are ignored. Two hits on complementary plane-pair wires form a **hit pair**. Figure 5.2 shows the distributions of number of hits per events and number of pairs per event, in both the 97 and 99 4-track data. These events have on average 6 hits per plane and 7 pairs per plane-pair.

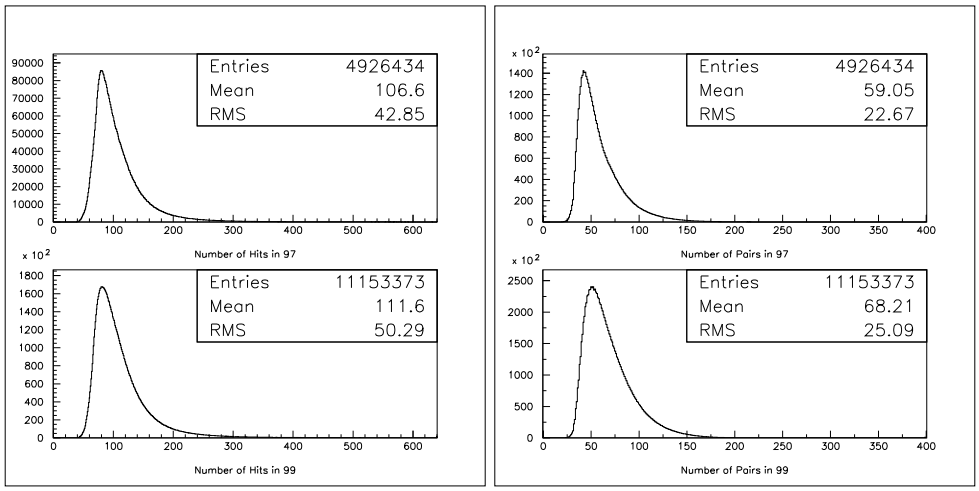

Figure 5.2: Distributions of the number of hits per event (left) and the number of pairs per event (right), in both 97 and 99 2e-nclus 4-track data.

Pairs are further classified by the **sum-of-distances** (SOD) to each of the two wires in the complementary planes. Ideally, the SOD is a constant equal to the cell spacing of 6.35 mm, however track angles and hit resolution and inefficiencies smear the SODs into a distribution that has a central Gaussian peak with a large high-side tail. Pairs with SODs within 1 mm of the nominal value are considered good pairs (a looser classification of ± 1.5 mm is used in DC3X and DC4X due to the large angles of

downstream segments). Figure 5.3 shows the distribution of SODs for all pairs along with the SODs for pairs used by tracks. The top plot is made from pairs in DC1X while the bottom one is comprised of pairs from DC3X.

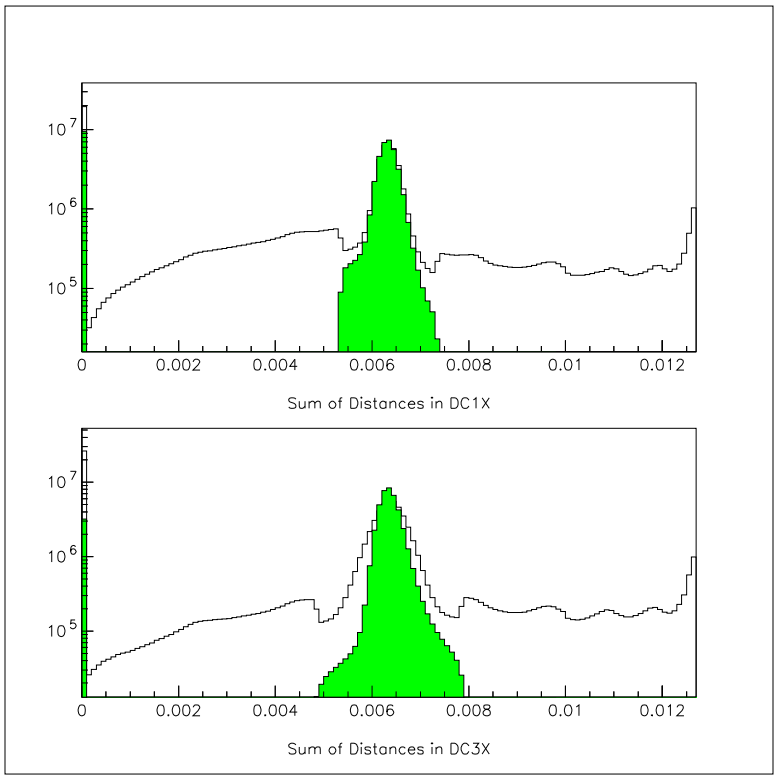

Figure 5.3: Distribution of Sum-of-Distances in the 4-track data. The top plot shows the SOD of all pairs (clear histogram) and the SOD of pairs used on tracks (green histogram) in DC1X. The bottom plot shows the same thing but in DC3X.

The low SODs are typically the result of two tracks passing through the same cell. Each wire will see two hits, one from each track. However, if there are multiple hits on a wire, only the first, or earliest, hit is used. The pair is then composed of one hit from each track and the resulting SOD is too small. High SOD pairs are expected due to discrete ionization sites, but is enhanced by inefficiencies in the detection of the first ionization. Isolated singles are due to in-time hits without a hit on complementary wires. These single hits are defined to have a SOD of zero. Singles are likely due to physical defects on the sense wires caused by radiation damage. Figure 5.4 shows a

cartoon of the different types of pairs.

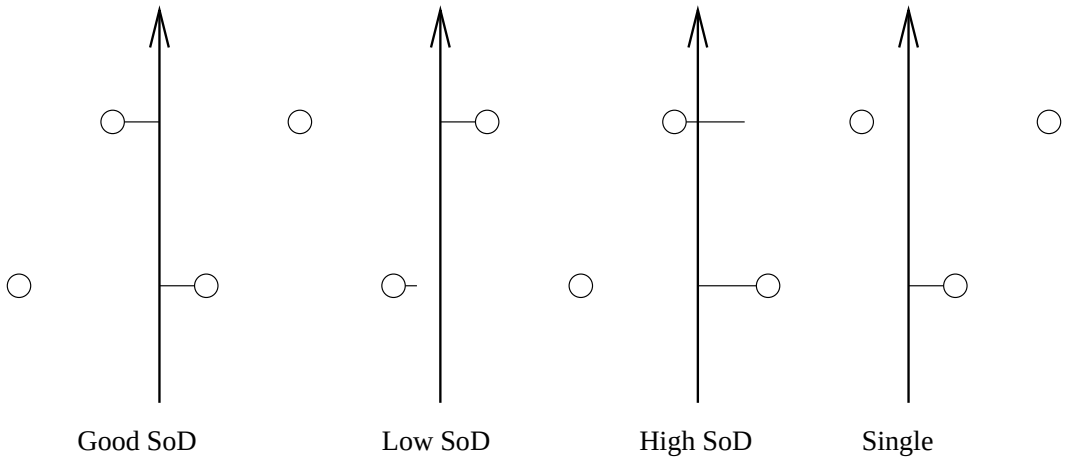

Figure 5.4: Cartoon of various pair configurations.

5.1.1.2 Tracks

Tracks are found first in the y-view using all four chambers and then in x-view as segments upstream and downstream of the analysis magnet. Since the magnet does not bend tracks in the y-view, a good track should have 4 pairs, one in each y-view, all lying on a straight line. The algorithm for finding y–tracks begins by considering all possible lines connecting each hit in DC1Y to each hit in DC4Y. For each such line, DC2Y and DC3Y are checked for hits lying close to the line. If hits are found, the sum of pair values is calculated and used to identify good tracks to save. The selection required a track to have at least two good SOD pairs and no more than one isolated single. Some amount of sharing of hits between different tracks was allowed.

The tracks in the x–view were formed in two segments, one upstream of the analysis magnet and one downstream. The upstream leg could contain as many as two bad SOD pairs or one single. The downstream leg was required to use at least one good SOD pair. An x–track was then formed by locating pairs of segments that matched to better than 6 mm at the bend plane. A full x–track could use no more than two bad

SOD's, or one bad SOD and one single. Figure 5.5 contains an expanded view of the event display focusing on the tracks in the spectrometer. Note the amount of hit-sharing in the y-view and how the two tracks in y are shared by the four tracks in the x-view.

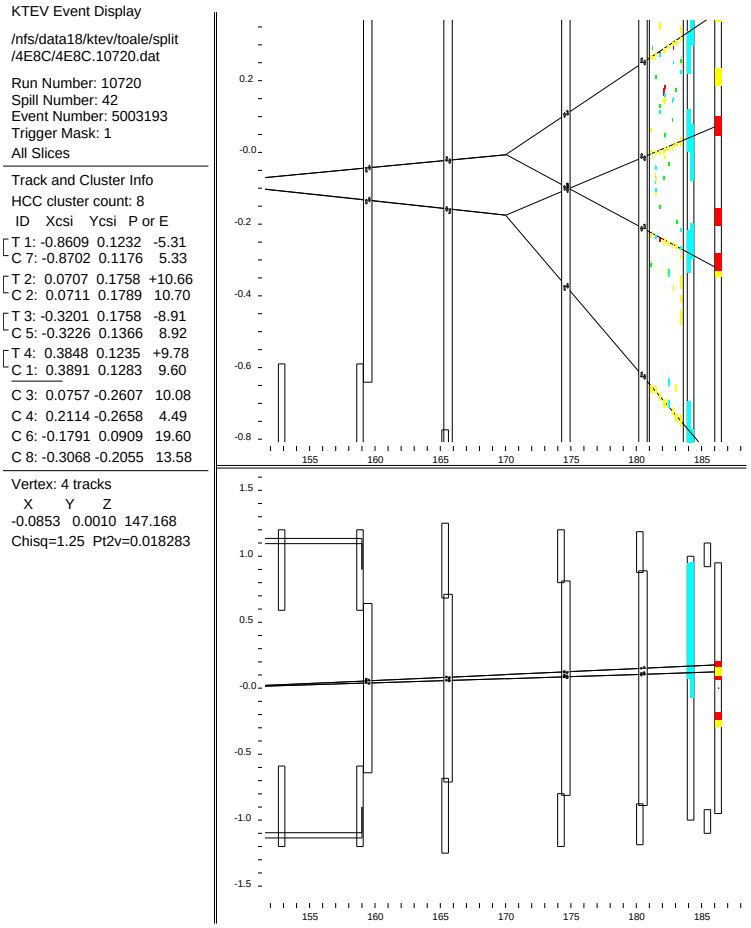

Figure 5.5: KTeV event display for a 4e8c candidate event showing the reconstructed tracks in both the x-view (top) and y-view (bottom).

5.1.2 Clustering

The clustering process reconstructs the position and the energy of both electrons and photons in the calorimeter. The cluster reconstruction algorithm can be used in two ways: **hardware clustering** based on the L2 HCC seeds, or **software clustering**

designed to find any local maximum of energy. The two methods may be used separately or in unison. For final states containing muons and charged pions, the software clustering is important for finding minimum ionizing particles and partial hadronic showers. This analysis uses only the hardware clustering method.

The hardware clustering algorithm starts with the seeds found by the L2 HCC discussed in Chapter 3. An initial energy is calculated based on the energy in the vicinity of the seed block. The energy is summed over a 7×7 grid of small crystals or a 3×3 grid of large crystals.

This energy is then used to find the transverse position of the cluster in the calorimeter. Look-up tables are used to convert between the energy distribution within the square grid of crystals to a position within the seed block. The table takes as inputs all of the ratios of the energy in the row and column containing the seed block to all other adjacent rows and columns. The output is a position with a resolution of roughly 1 mm in the small crystals and 1.8 mm in the large crystals. Figure 5.6 shows the distributions of reconstructed cluster energies in the 99 4-track data as well as the number of clusters per event in both 97 and 99.

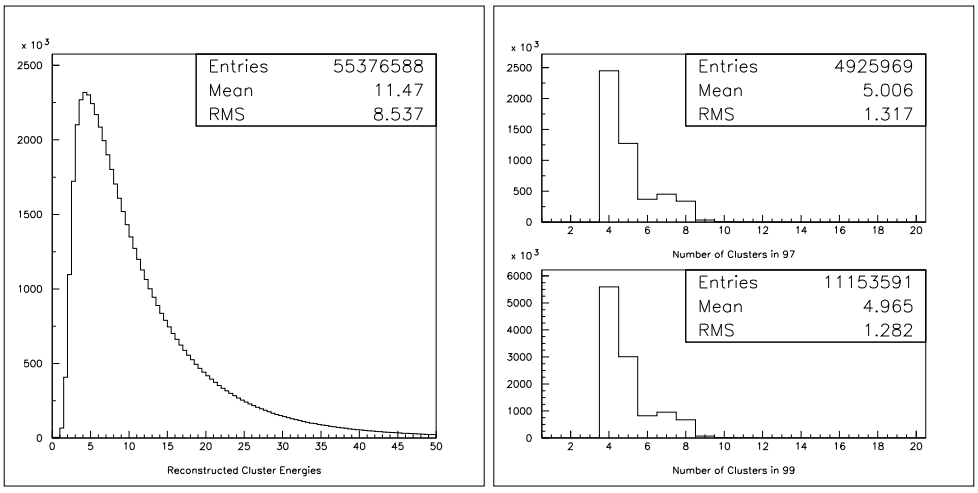

Figure 5.6: Distributions of the reconstructed cluster energies in the 99 2e-nclus 4-track data (left) and the number of clusters per event in the 97 data (top right) and in the 99 data (bottom right).

After all HCC seeds are treated in this way, a series of corrections are applied to each cluster. Some corrections pertain to the cluster as a whole while others are concerned with individual crystals within the cluster. Each correction will be discussed in detail below.

5.1.2.1 Clustering Corrections

There are five corrections applied to individual channels within clusters and three scale factors applied to cluster energies as a whole. The channel-by-channel corrections go by the names: overlap, neighbor, missing energy, sneaky energy, and threshold. The scale factors are known as: transverse, intrablock, and non-linearity. A typical cluster topology is shown in Figure 5.7. The HCC seeds are indicated by the crystals with dark borders. The two clusters shown overlap and one is adjacent to a beam hole. The corrections are intended to compensate for these common occurrences.

The overlap correction is applied when two 7×7 grids overlap. The correction simply divides the energy in the common crystals, assigning some each of the two clusters. If two clusters neighbor each other without actually overlapping, the energy in crystals along the boundary is adjusted to compensate for the presence of the other cluster. If a cluster lies near a beam hole or the outer edge of the calorimeter, a correction is applied to account for energy in the missing crystals. Next, if a cluster is close to a beam hole, a sneaky energy correction is applied to regain energy that might have passed across the beam hole into crystals on the other side. The final channel correction accounts for crystals within clusters that had energy below threshold and were therefore not read out.

The first scale factor accounts for cluster energy outside of the 7×7 array (or 3×3 in the big crystals). Next, a correction to account for the non-uniform response as a function of the intrablock position of the cluster is applied. The final correction compensates for the non-linear response of the calorimeter as a function of cluster energy.

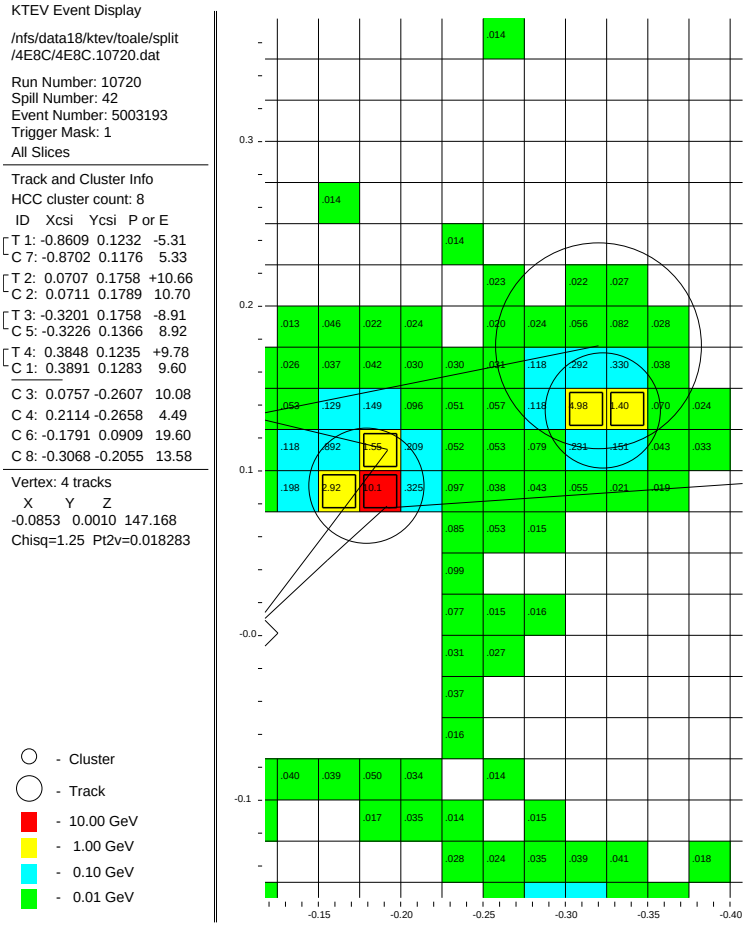

Figure 5.7: KTeV event display for a 4e8c candidate event.

5.1.3 Vertexing

The process of vertexing pulls together the tracks in both the x and y views and the clusters in the calorimeter to form a 3-D trajectory for each charged particle. The algorithm could be used to look for any number of oppositely charged pairs of particles originating at a common vertex. The data used in this analysis required the existence of a 4–track vertex.

The first step in forming a vertex is to find sets of y–tracks which intersect, within 2 mm, in the decay volume. The same is then done with the upstream x-track segments.

In order to tie the two views together, vertex candidates in each view are matched to clusters. Each track is projected to the calorimeter where its position is compared to the position of all of the clusters. If the minimum distance is less than 7 cm the track is said to match the cluster in that view. Tracks in x and in y that match to the same cluster, and project to within 1.5 cm of each other at the calorimeter, were paired together. At this point corrections are applied to the tracks to account for effects like alignment and fringe fields from the analysis magnet. For each vertex candidate, two quality parameters were calculated: the vertex χ^2 based on how well the tracks intersect at the vertex, and the offmag χ^2 based on the matching of the tracks at the magnet bend plane. The candidate with the smallest combination of these two quantities was selected as the final vertex. Figure 5.8 provides an indication of the resolution in the vertex position in z. The mean resolution in both data periods is less than 25 cm.

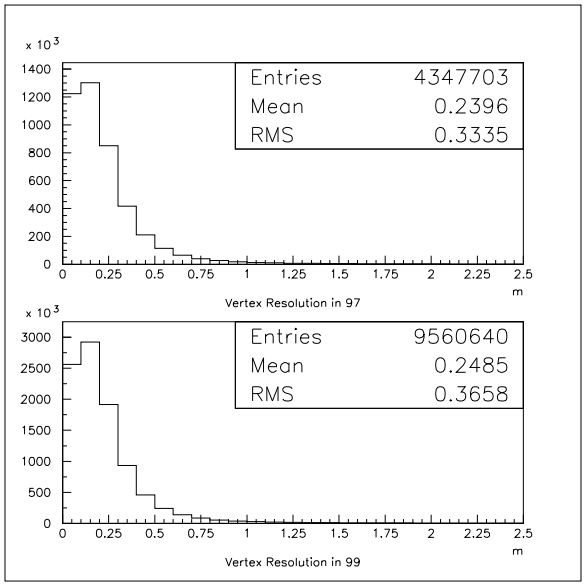

Figure 5.8: Vertex z resolution.

With the vertex position and track projections in hand, the momentum of each track can be computed. The momentum is related to the magnitude of the magnetic field, known from calibration, divided by the bend angle. For every track that matches

to a cluster we can then form the ratio of the energy of the cluster to the momentum of the track. This quantity, referred to as E/p, is a measure of the fraction of its energy a charged particle deposits in the calorimeter. It is usually near one for electrons and near zero for muons. Charged pions tend to deposit some energy, but rarely more than 80 %. Electrons can be reliably identified in this analysis by tracks with $E/p > 0.9$. The distribution of E/p for 4-track events is shown in Figure 5.9. The distribution is broader in 99 due to the degraded momentum resolution associated with the lower kick delivered by the analysis magnet.

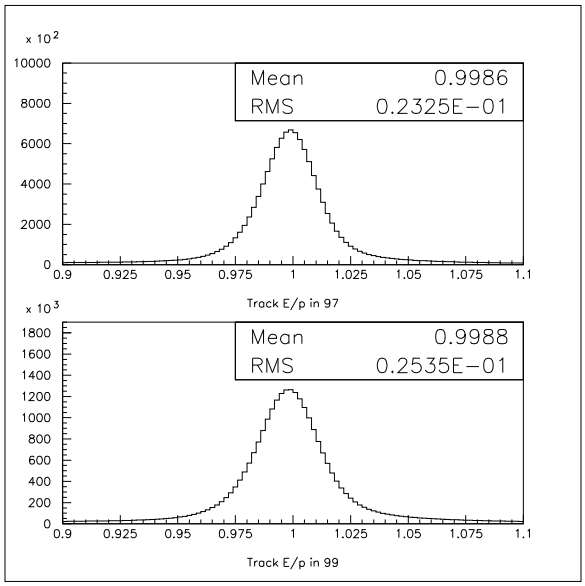

Figure 5.9: Distribution of the energy deposited in the calorimeter divided by the momentum measured in the spectrometer, for 97 (top) and 99 (bottom).

Finally, photons are identified with clusters that are not associated with any tracks. The photon momenta are calculated assuming that they originated at the charged vertex. Figure 5.10 shows a two dimensional distribution of the number of photon clusters versus the number of electron clusters. The $\pi^0 \to e^+e^-e^+e^-$ analysis requires that there are four clusters of each type.

The information available at this point is enough to select candidate events characterized by a good 4–track vertex comprised of electrons along with 4 additional photons.

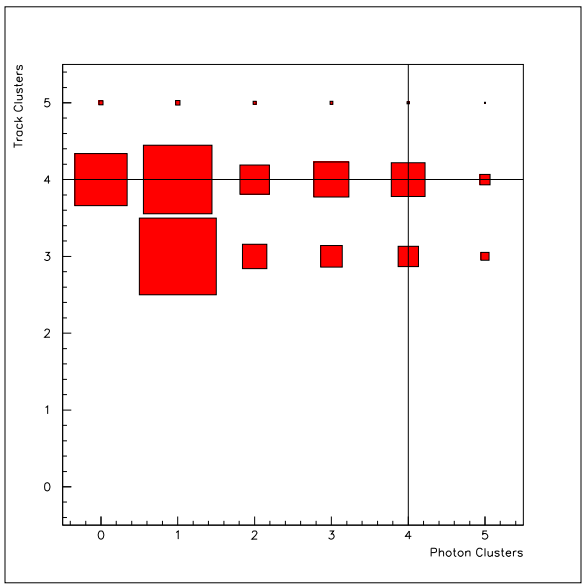

Figure 5.10: Two dimensional distribution of the number of clusters identified as photons versus the number of clusters associated with tracks.

The rest of this chapter covers the early stages of data filtering based on these concepts and algorithms.

5.2 Data Filtering

The techniques described above were first applied to the data, in a loose form, online as part of the L3 trigger. The goal was to identify the possibility that a good vertex existed and that the tracks belonged to electrons. Events tagged by 2e-nclus trigger and written to tape were eventually processed by two additional filter programs: first the 2e-nclus crunch and then the 4–track split.

5.2.1 Level 3 Filter

The L3 filter code performed a quick reconstruction, looking for evidence of two tracks with opposite charge that intersected at a common vertex and matched to two clusters in the calorimeter. A clustering algorithm similar to that used by the HCC at L2 required that there were at least 4 clusters. At this point, only a 2-track vertex was

required. The vertexing procedure will always find a 2-track vertex for events with good 4-track vertices. The fraction of $K_L \to \pi^0\pi^0\pi^0_{DD}$ Monte Carlo events that passed the L2 trigger simulation and also satisfied the L3 filter was roughly 71% in 97 and 75% in 99. Most of the lost events were cut due to insufficient matching between clusters and tracks due to tracks bent out of the detector by the analysis magnet.

The L3 code also tagged events based on a more thorough analysis. One of two of these tags were required to be set for the events used in this analysis. The **2encls** tag only required that two tracks have $E/p > 0.75$. The **k3pi0d** tag required in addition that there were at least seven clusters and that the total invariant mass was between 400 and 600 $\mathrm{MeV/c^2}$.

5.2.2 Trigger 1 Crunch

The trigger 1 crunch was a data reduction and selection process performed in Boulder, for both the 97 and 99 data, for the general use of the KTeV collaboration. The goal was to reconstruct events, apply simple selection cuts, and also to remove unnecessary components of the data stream. The output consisted of a condensed sample of high quality events, split into ten classes. The ten categories are summarized in Table 5.1.

The 4-track event selection made three requirements: a proper L3 tag must have been set, two tracks must have $E/p > 0.9$, and a 4-track vertex must be found. Figure 5.11 shows the invariant mass of the four charged particles for events satisfying these criteria. The events in blue have four tracks with $E/p > 0.9$. The peak in the blue histogram occurs at the π^0 mass. The rest of the distribution is due primarily to $K_L \to \pi^+\pi^-\pi^0$ with $\pi^0 \to e^+e^-\gamma$. The total invariant mass of the events with four electrons is shown in Figure 5.12. Here, the events in blue contain four photons while those in green have five photons. The total mass peaks at the K_L mass.

The composition of the events with four electrons and four photons can be inferred

Table 5.1: Summary of Crunch streams and the fraction of trigger 1 events saved to each in 97 and 99.

Stream	Name	Motive	97 Saved (%)	99 Saved (%)
1	L3RAND	Level 3 random accepts	0.33	0.18
2	PI0TEE	$K_L \to \pi^0\pi^0\pi^0$ with $\pi^0 \to e^+e^-$	4.20	4.42
3	3T6SC	Background studies	0.81	0.89
4	4TRACK	4–track modes ($\pi^0 \to e^+e^-e^+e^-$)	3.73	4.45
5	2T8C	$K_L \to \pi^0\pi^0\pi^0$ with $\pi^0 \to e^+e^-\gamma\gamma$	1.21	1.37
6	2PI0EE	$K_L \to \pi^0\pi^0e^+e^-$	3.01	5.92
7	3T7C	Background studies	0.64	0.64
8	3PI0D	$K_L \to \pi^0\pi^0\pi^0$ with $\pi^0 \to e^+e^-\gamma$	20.6	25.7
9	EEGGG	$K_L \to \pi^0e^+e^-\gamma$	3.19	3.70
10	EEGG	$K_L \to \pi^0e^+e^-$	1.22	3.54

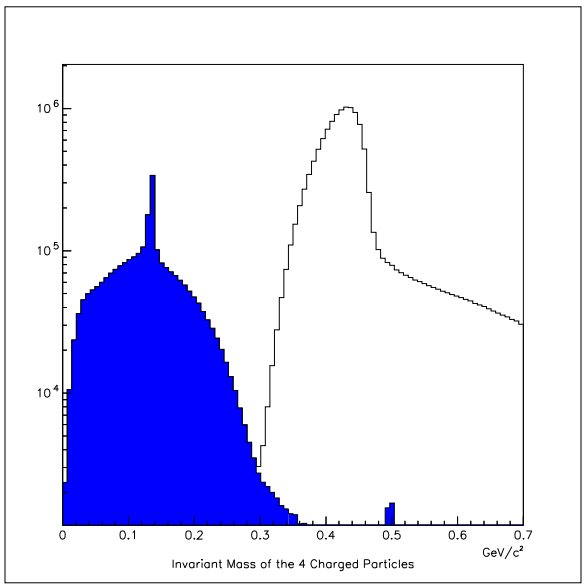

Figure 5.11: Distribution of the invariant mass of the four charged particles in events in the 4TRACK crunch stream. The events in the blue histogram have all four E/p's greater than 0.9.

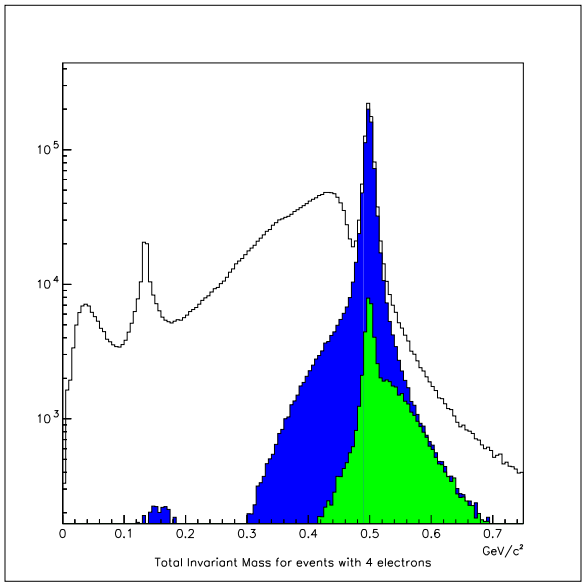

Figure 5.12: Distribution of the total invariant mass in events which have four tracks with $E/p > 0.9$. The events in the blue histogram have, in addition to the electrons, four photons (4e8c), while those in green have five photons (4e9c).

from the distribution of the invariant mass of the four electrons. Figure 5.13 shows this distribution. The continuous part of the spectrum is mainly due to $K_L \rightarrow \pi^0\pi^0_D\pi^0_D$

events, while the events in the peak at the π^0 mass are $K_L \to \pi^0\pi^0\pi^0_{DD}$ decays. The total mass versus P_T^2 for these events is shown in Figure 5.14, in which the vertical scale is logarithmic. Good $K_L \to \pi^0\pi^0\pi^0_{DD}$ events reside at the bottom of the plot with a mass near the kaon mass.

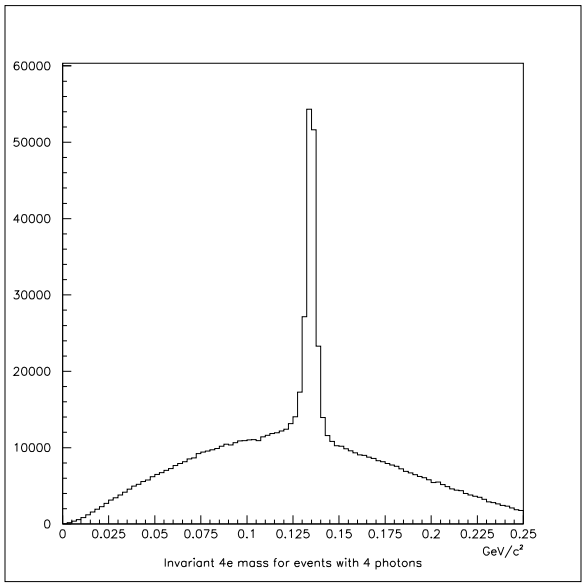

Figure 5.13: Distribution of the invariant mass of the four electrons in events with four photons.

The total number of events selected in the 4TRACK stream of the crunch is shown in Table 5.2. The total number of 4-track events identified by the crunch is just over 16 million.

5.2.3 4–Track Split

The 4–track events selected by the 2e-nclus crunch were subjected to one final stage of filtering to separate events with either 8 or 9 clusters. At this stage all four tracks of the vertex were required to match to clusters in the calorimeter. About 3.5 % of the 4-track events previously found fail to satisfy the stiffer matching requirement.

The distribution of the number of clusters in the 4-track events is shown in Figure 5.15. Roughly 6.0 % of the events contain 8 clusters while only 0.7 % have 9 clusters.

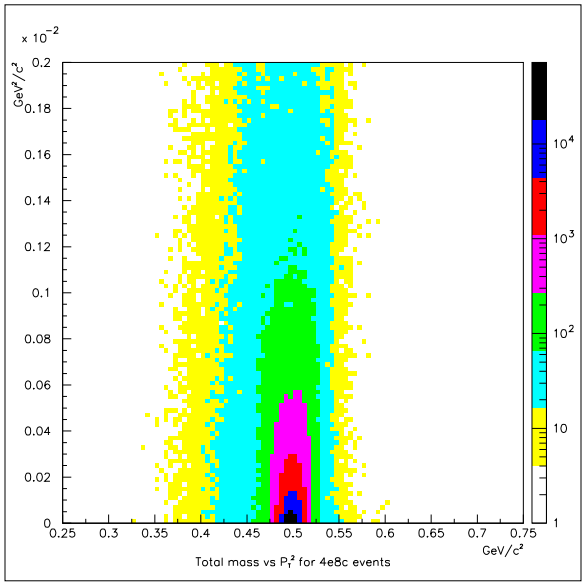

Figure 5.14: Two dimensional distribution of the total invariant mass versus the transverse momentum squared for events with four electrons and four photons.

The number of events saved with either 8 or 9 clusters is summarized in Table 5.3. The total number of 4e8c events is about 950,000, while the total number of 4e9c events is 96,000.

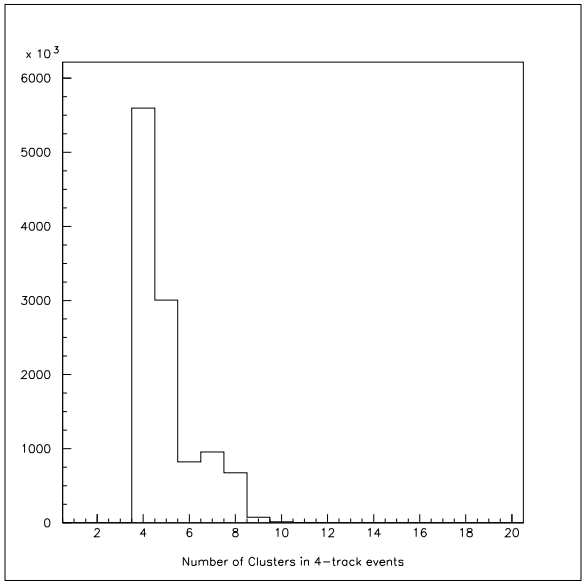

Figure 5.15: Distribution of the number of clusters in 4TRACK events.

Table 5.2: Number of events selected in the 4TRACK stream of the 2e-nclus crunch.

Dataset	N_{4TRACK}
Win97	2,830,472
Sum97	2,096,271
99	11,153,592
Total	16,080,334

Table 5.3: Number of events selected to the 4e8c and 4e9c streams by the 4-track split.

Dataset	N_{4e8c}	N_{4e9c}
Win97	183,126	16,651
Sum97	140,149	13,100
99	633,877	66,290
Total	957,152	96,041

Chapter 6

Double Dalitz Branching Ratio

This chapter discusses the branching ratio measurement for the $\pi^0 \to e^+e^-e^+e^-$ mode. The data used to extract the result are from the 4–track, 8–cluster output of the split that was described in the previous chapter. There are two modes that must be reconstructed, $K_L \to \pi^0\pi^0\pi^0_{DD}$ and $K_L \to \pi^0\pi^0_D\pi^0_D$. The first mode is referred to as the signal mode while the latter is known as the normalization mode.

The number of observed double Dalitz events is related to the branching ratio by

$$N_{DD}^{\rm obs} = 3 \cdot F_K \cdot B(K_L \to \pi^0\pi^0\pi^0) \cdot B^2(\pi^0 \to \gamma\gamma) \cdot B(\pi^0 \to e^+e^-e^+e^-) \cdot \epsilon_{DD}, \quad (6.1)$$

where F_K is the kaon flux, ϵ_{DD} is the acceptance or efficiency for reconstructing a decay, and the factor of three comes from the $3\pi^0$ combinatorics. The flux is defined as the number of kaons decaying in the fiducial region while the KTeV detector/trigger was active and ready to accept events. Since the kaons are neutral they cannot be detected directly, making it impossible to count them individually. To eliminate the dependence on the flux I can simultaneously measure two modes and calculate the ratio of events in each. The number of $K_L \to \pi^0\pi^0_D\pi^0_D$ events observed in the data is

$$N_{2D}^{\rm obs} = 3 \cdot F_K \cdot B(K_L \to \pi^0\pi^0\pi^0) \cdot B(\pi^0 \to \gamma\gamma) \cdot B^2(\pi^0 \to e^+e^-\gamma) \cdot \epsilon_{2D}, \quad (6.2)$$

where ϵ_{2D} is the acceptance for the normalization mode and the single Dalitz branching ratio is taken from the PDG. In the absence of backgrounds, the double Dalitz branching

ratio is related to the following double ratio

$$R = \frac{N_{DD}^{\rm obs}}{N_{2D}^{\rm obs}} \cdot \frac{\epsilon_{2D}}{\epsilon_{DD}} = \frac{B(\pi^0 \to e^+e^-e^+e^-) \cdot B(\pi^0 \to \gamma\gamma)}{B^2(\pi^0 \to e^+e^-\gamma)}. \tag{6.3}$$

Backgrounds will modify this form slightly.

I will first discuss the final selection criteria for both the signal and normalization modes. This will be followed by the acceptance numbers from the Monte Carlo along with specific comparisons between distributions in the data and MC. The residual background level is also predicted by the Monte Carlo.

6.1 4e8c Reconstruction

The reconstruction begins by finding tracks, clusters and a 4-track vertex, as discussed in Chapter 5. All four tracks are required to match to clusters in the calorimeter. The four clusters not associated with tracks are considered photons. Events with a 4-track vertex plus four photons are subjected to further selection criteria to identify clean samples of both the signal and normalization modes. These cuts can be separated into the following categories: data quality, trigger verification, fiducial cuts, kinematic selection, and background elimination. I will discuss each of these in turn.

6.1.1 Data Quality

The experiment was diligently monitored during data collection, however not all of the data is of the highest quality. Occasionally whole spills of data were collected under less-than-ideal conditions. Rather than try to estimate the acceptance of these **bad spills**, they are simply removed from the data and not generated in the MC. The analysis code allows a user to set a mask of bits which refer to 1 of 32 different detector problems which might have occurred during a given spill. Known problems in every run are categorized and input into the KTeV database. The masks of bad spill bits used by this analysis are summarized in Table 6.1. The distribution of bad spill bits in the 4e8c

data is shown in Figure 6.1 for each of the three datasets.

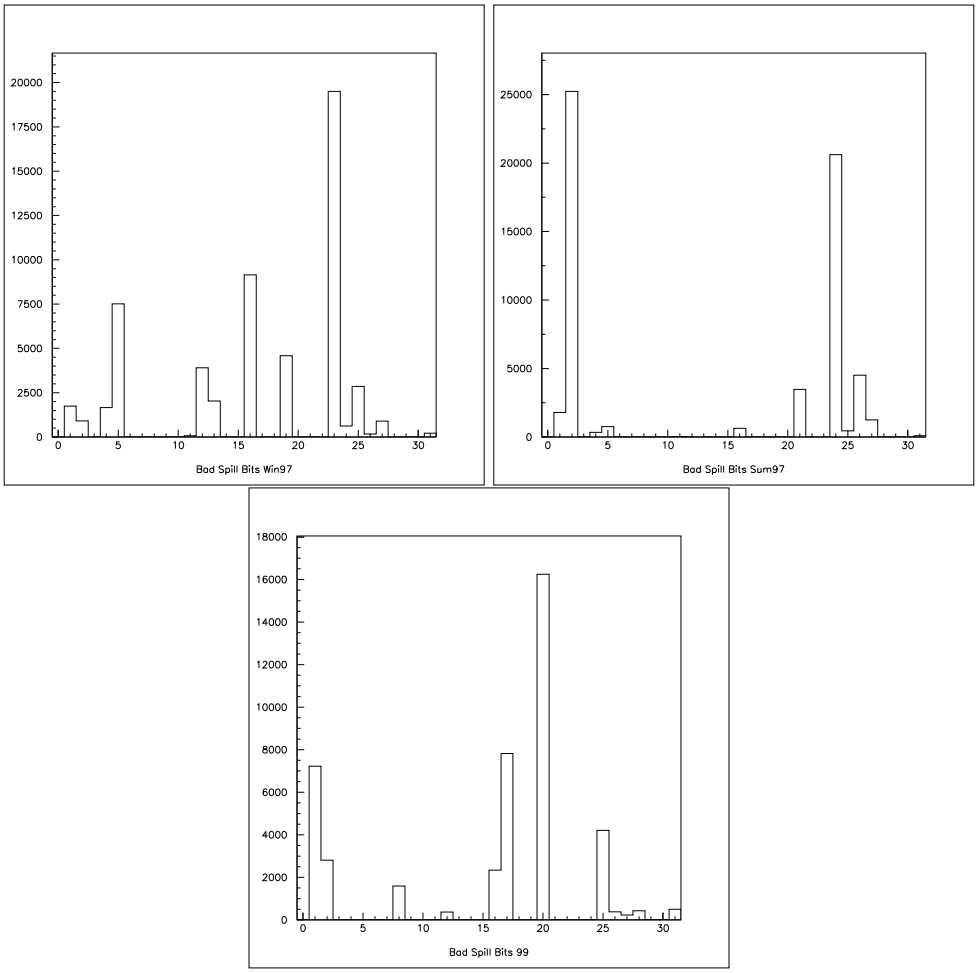

Figure 6.1: Distribution of bad spill bits in the 4e8c dataset for each of the three run periods.

6.1.2 Trigger Verification

All events are subjected to a loose trigger verification in order to remove events which require accidental coincidence to satisfy the trigger. The veto requirements are verified by cutting on the energy reconstructed in each module. The RCs, SAs, and CIA are required to have less than 300 MeV of energy in any module. The CA is required to have less than 13 GeV of energy. The muon veto was only used in a small fraction of the data and is not verified. Figure 6.2 shows the maximum energy in any RC, SA,

Table 6.1: Definition of bad spill bits and masks used to remove bad spills from the event sample. Bits 30, 31, and 32 were not used.

Bit	Name	Win97	Sum97	99
1	Trigger	1	1	1
2	DPMT Ped Exp	1	1	1
3	DPMT Cap ID	1	0	0
4	Blown QIE Comp	1	1	1
5	Dead DPMT	1	1	1
6	DPMT Ped Drift	0	0	0
7	DPMT Gain Drift	1	1	1
8	Broken Dynode	1	1	1
9	Pipe Problems	1	1	1
10	Global CsI Prob	1	1	1
11	Etot	1	1	1
12	Fera ADC	1	1	1
13	DC	1	1	1
14	Veto	1	1	1
15	V-Bank	1	1	1
16	Muon	0	0	0
17	HCC	0	1	1
18	Banana	1	1	1
19	TRD Trigger	0	0	0
20	Hyperon Trigger	0	0	0
21	DAQ	1	1	0
22	Non 799/832 Run	1	1	1
23	Short Run	1	1	1
24	Nonstd TRD HV	0	0	0
25	1 Dead TRD Plane	0	0	0
26	Few Dead Planes	0	0	0
27	TRD HV Sag	0	0	0
28	Severe TRD Prob	0	0	0
29	Beam	1	1	1
	Mask	1073EFDF	1073FFDB	1073EFDF

and CA for the 4e8c dataset.

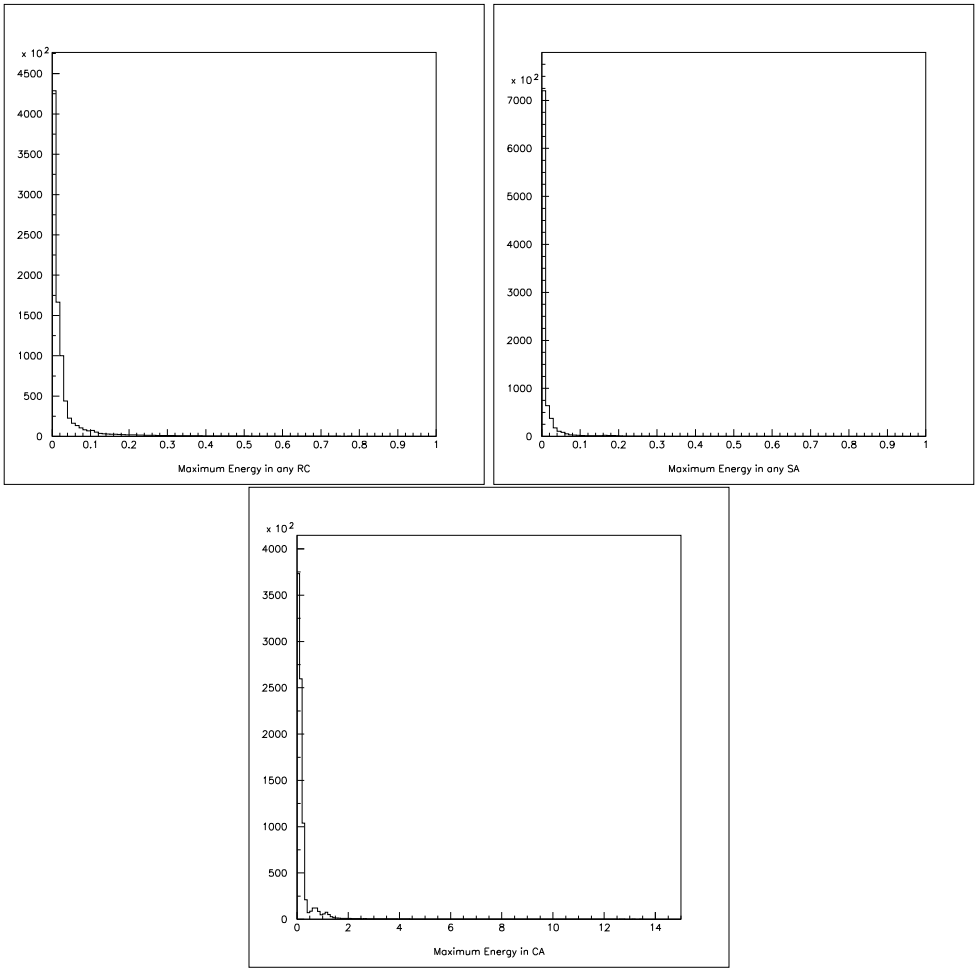

Figure 6.2: Distributions of the maximum energy deposited in, from left to right, the RC's, the SA's, and the CA's, for the entire 4e8c dataset (in GeV).

The neutral elements of the trigger are verified by requiring that the total energy of all the clusters used in the event is greater than 40 GeV. The HCC requirement is automatically satisfied since all events must have 8 hardware clusters. Figure 6.3 shows the low end of the total cluster energy. The events below the nominal trigger cut of 28 GeV contain additional energy not used in the eight hardware clusters.

The charged part of the trigger is verified by extrapolating the tracks from the vertex to the VV' hodoscopes planes and the number of paddles which should have been hit are counted. In order to account for uncertainties in the exact location of the

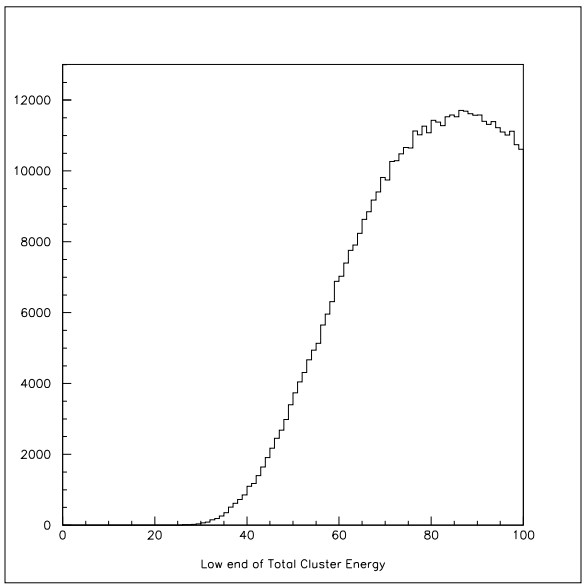

Figure 6.3: Low end of the distribution of total energy (in GeV) in the 8 hardware clusters.

edges, each paddle is expanded by 1 cm and tracks that pass through this buffer are assumed to hit both paddles. All 4e8c events pass this loose requirement. The rest of the charged trigger is assumed to be implicitly satisfied if four good tracks are found.

6.1.3 Fiducial Cuts

A series of cuts are applied to the data to ensure that the reconstruction is as robust as possible. This typically means eliminating events that may suffer from gross mis-reconstruction. These cuts include requirements on individual clusters and tracks, as well as on the vertex itself.

6.1.3.1 Cluster Requirements

There are three cuts specific to clusters in the calorimeter. Each of the eight clusters is required to have more than 2 GeV of energy. This requirement is made to remove any dependence on the HCC thresholds, nominally between $1.0 - 1.5$ GeV. The total energy in the calorimeter is required to be between 40 and 210 GeV, again to stay

away from the trigger threshold of $\sim 28\,\mathrm{GeV}$. The smallest distance between any two clusters must be more than 5 cm. Clusters with separations less than this are subject to large overlap corrections. Figures 6.4, 6.5, and 6.6 show a comparison between normalization mode data and Monte Carlo events in the three cluster variables.

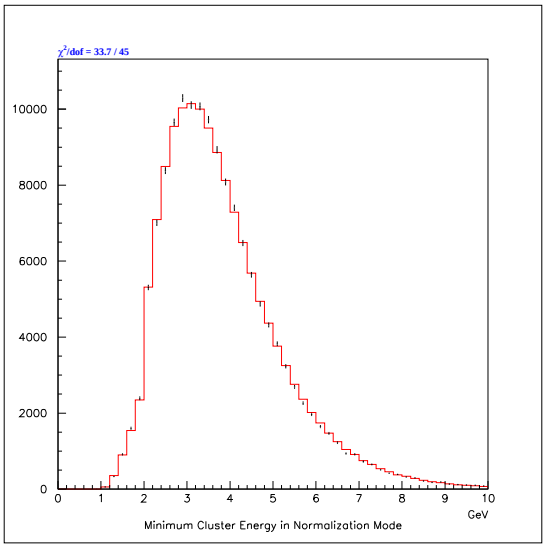

Figure 6.4: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of minimum cluster energies. All other cuts have been applied.

6.1.3.2 Track Requirements

Two cuts are applied to track parameters. The minimum track momentum must be larger than $2\,\mathrm{GeV/c}$. Below this value, the acceptance rapidly drops due to tracks bending out of the detector in the analysis magnet's field. The combination of E/p is required to reconstruct between 0.93 and 1.07. Figures 6.7, 6.8, and 6.9 show a comparison between normalization mode data and Monte Carlo events. The small discrepancy in the E/p distributions will be addressed later.

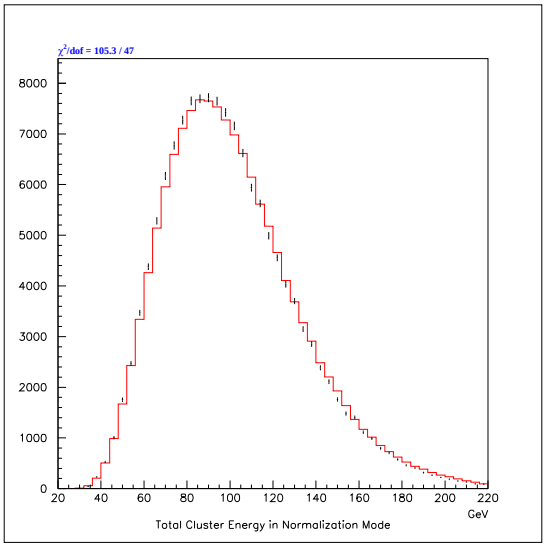

Figure 6.5: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of total cluster energy. All other cuts have been applied.

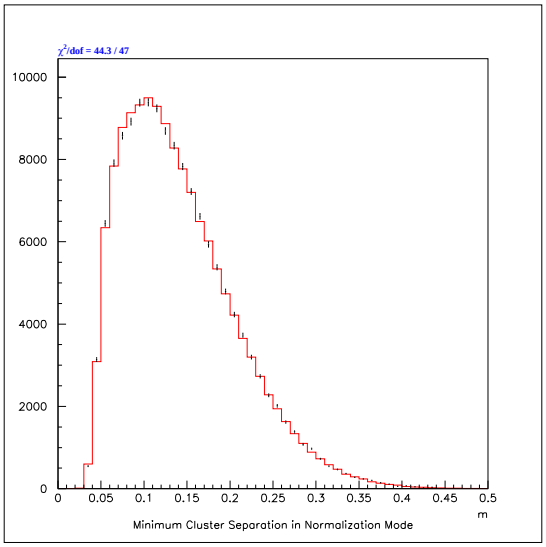

Figure 6.6: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of minimum cluster separation. All other cuts have been applied.

6.1.3.3 Vertex Requirements

Another three cuts are made to ensure the quality of the reconstructed vertex. The vertex χ^2 and magnet χ^2 are required to be less than 40 and 100, respectively. In

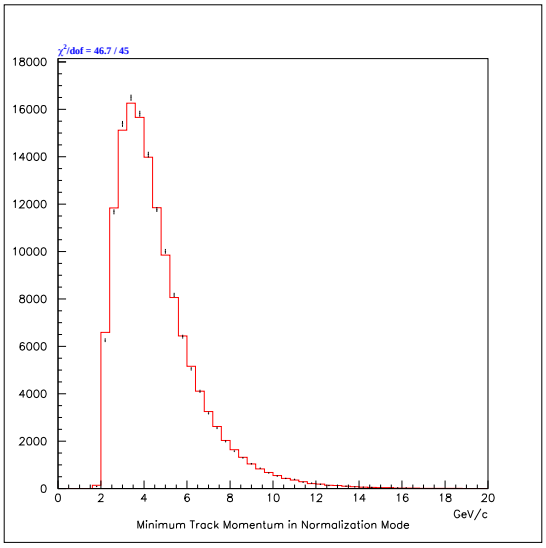

Figure 6.7: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of minimum track momentum. All other cuts have been applied.

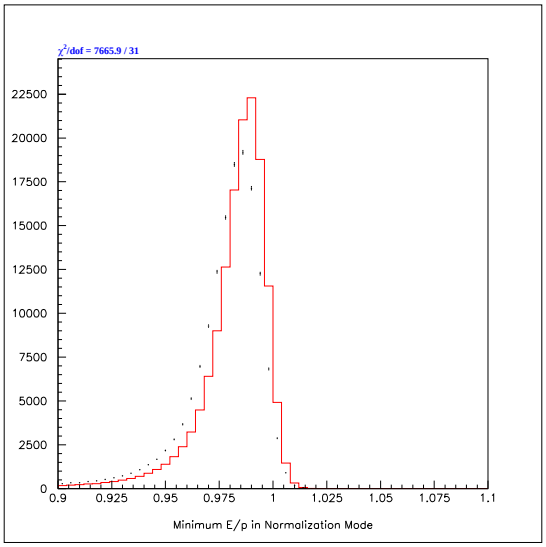

Figure 6.8: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of smallest E/p. All other cuts have been applied.

addition, the z-position of the vertex has to lie between 97 and 157 m. Figures 6.10, 6.11, and 6.12 show a comparison between normalization mode data and Monte Carlo events in the three vertex variables.

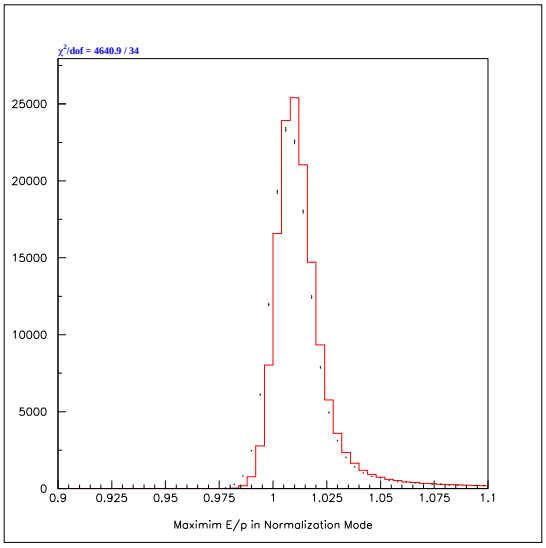

Figure 6.9: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of largest E/p. All other cuts have been applied.

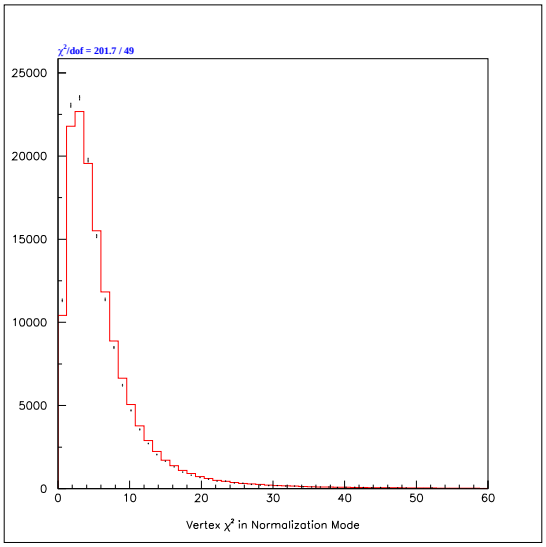

Figure 6.10: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of vertex χ^2. All other cuts have been applied.

6.1.4 Kinematic Selection

At this point each event is tested to see if it represents the complete final state of a kaon decay, and if so, if it is kinematically consistent with either a $K_L \to \pi^0\pi^0\pi^0_{DD}$ or

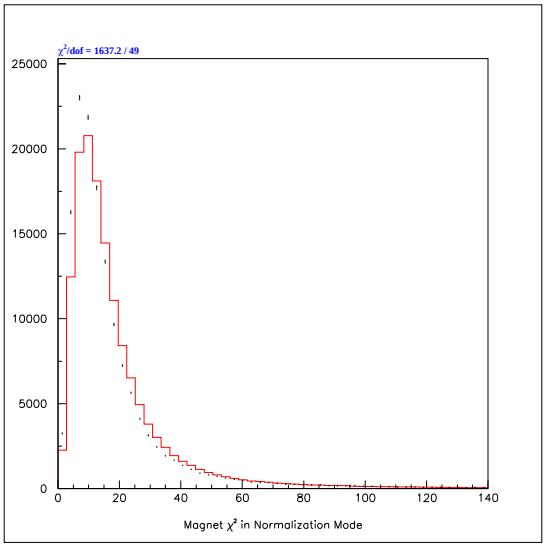

Figure 6.11: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of magnet χ^2. All other cuts have been applied.

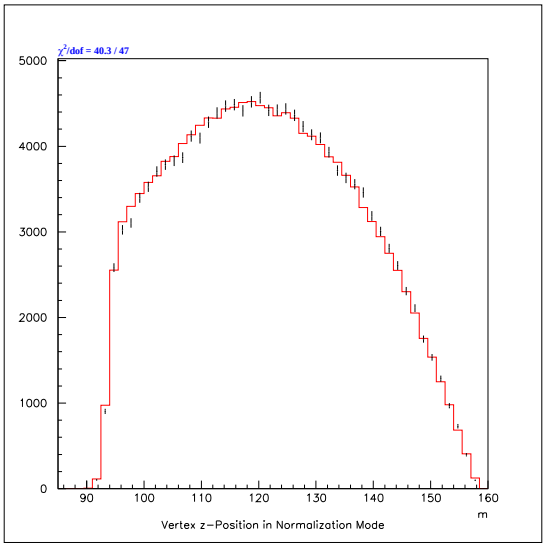

Figure 6.12: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of vertex z-position. All other cuts have been applied.

$K_L \to \pi^0\pi^0_D\pi^0_D$ decay. This is accomplished first by calculating the total invariant mass and the total 3-momentum perpendicular to the direction of the kaon. Then a pairing χ^2 is formed to distinguish between the two different $K_L \to \pi^0\pi^0\pi^0$ decays.

6.1.4.1 Kaon Cuts

The total mass must reconstruct in a window from 480 to 515 $\mathrm{MeV/c^2}$ and the square of the transverse momentum must be less than 800 $\mathrm{MeV^2}/c^4$. The transverse momentum is defined as the component of the total momentum of the reconstructed final state perpendicular to the direction of the kaon. The direction of the kaon is given by the line connecting the target to the reconstructed decay vertex. These cuts ensure that all decay products of the kaon were observed. Figures 6.13 and 6.14 show a comparison between the normalization mode data and Monte Carlo in these two quantities.

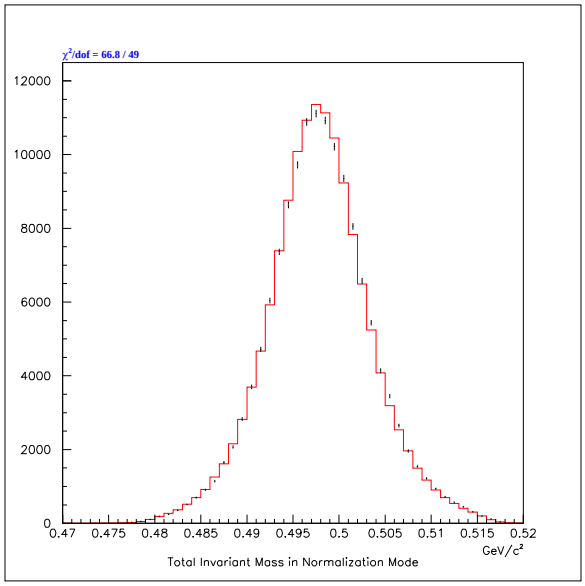

Figure 6.13: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of total invariant mass. All other cuts have been applied.

6.1.4.2 Pairing χ^2

At this point it is necessary to distinguish between the signal and the normalization modes. This is accomplished by testing two hypotheses on each event and selecting

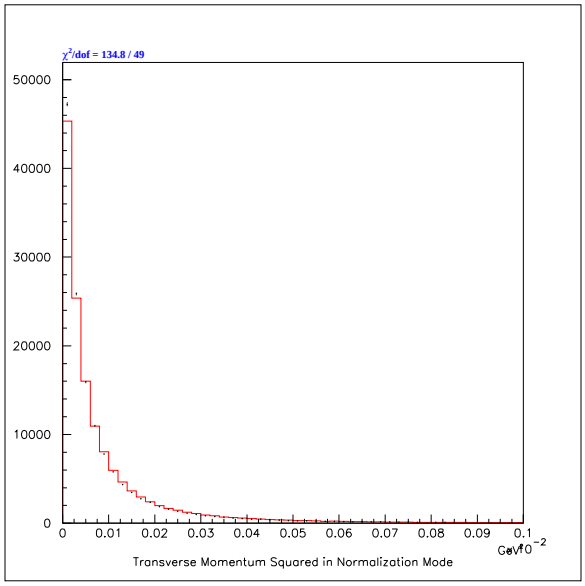

Figure 6.14: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) for the distribution of the square of the transverse momentum. All other cuts have been applied.

the more likely result. The test functions, ignoring correlations, are

$$\chi^2_{DD} = \frac{(M_{4e} - M)^2}{\sigma^2_{4e}} + \frac{(M_{\gamma\gamma 1} - M)^2}{\sigma^2_{\gamma\gamma}} + \frac{(M_{\gamma\gamma 2} - M)^2}{\sigma^2_{\gamma\gamma}}, \tag{6.4a}$$

$$\chi^2_{2D} = \frac{(M_{ee\gamma 1} - M)^2}{\sigma^2_{ee\gamma}} + \frac{(M_{ee\gamma 2} - M)^2}{\sigma^2_{ee\gamma}} + \frac{(M_{\gamma\gamma} - M)^2}{\sigma^2_{\gamma\gamma}}. \tag{6.4b}$$

The resolutions on the three masses are $\sigma_{4e} = 1.60\,\mathrm{MeV/c^2}$, $\sigma_{ee\gamma} = 1.71\,\mathrm{MeV/c^2}$, and $\sigma_{\gamma\gamma} = 1.50\,\mathrm{MeV/c^2}$. Correlations are accounted for.

For each of the final states there is an ambiguity as to the pairing of the particles into the three π^0's. For the signal mode, there are three possible combinations, all of which are tested. The most likely $K_L \to \pi^0\pi^0\pi^0_{DD}$ pairing is the one with the smallest χ^2_{DD}. The normalization mode has a 24-fold ambiguity. All 24 possibilities are considered and the one with the smallest χ^2_{2D} is selected. The best χ^2's for each hypothesis are then compared, with the smaller one determining the type of decay. The overall best χ^2 is then required to be less than 12.

Figure 6.15 shows a comparison between the normalization mode data and Monte

Carlo in the reconstruction χ^2 variable. The disagreement near the peak of the distribution is due to cluster energy mis-reconstruction at the level of 0.5 %. The effect is not a concern since the cut is placed rather far out on the tail of the distribution. This disagreement will be discussed further in the context of systematic uncertainties at the end of this chapter.

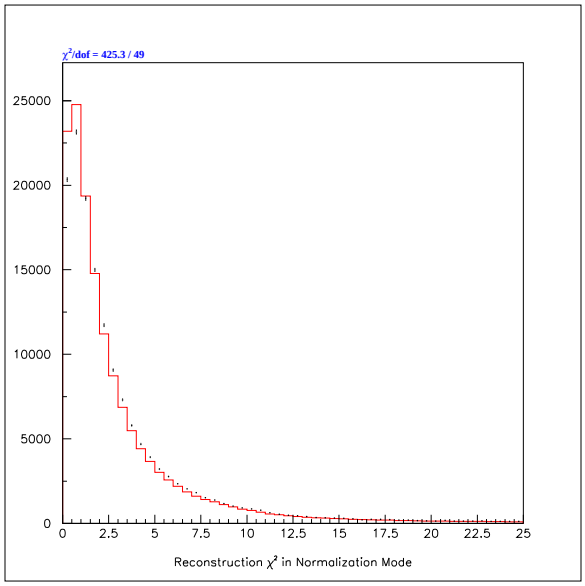

Figure 6.15: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) in the reconstruction χ^2 variable.

6.1.5 Background Elimination

The final cut that is applied to the data is required to remove the massive number of external conversion events. Photons which convert upstream of the spectrometer can mimic tracks from the vertex. The probability of a photon converting in this region is roughly 0.23 %. For a $K_L \to \pi^0\pi^0\pi^0_D$ decay the probability that one of the five photons converts is about 1.2 %, remarkably close to the single Dalitz branching ratio. The likelihood of two of the six photons in a $K_L \to \pi^0\pi^0\pi^0$ decay converting is on the order of 8×10^{-5}. Table 6.2 lists the conversion probabilities in the three run periods. The important thing to note is that the probabilities for the background processes are

Table 6.2: Conversion probabilities in the three run periods. P/γ is the probability of a photon converting upstream of the spectrometer. $P(1|5)$ and $P(2|6)$ are the probabilities of one in five, or two in six, photons converting. $P(1C)$ and $P(2C)$ are the combined likelihood of a $K_L \to \pi^0\pi^0\pi^0_D$ decay plus one conversion, or $K_L \to \pi^0\pi^0\pi^0$ decay with two conversions.

Probabilities	Win97	Sum97	99
P/γ	0.241 %	0.246 %	0.218 %
$P(1\|5)$	1.19 %	1.22 %	1.08 %
$P(2\|6)$	8.63×10^{-5}	8.99×10^{-5}	7.07×10^{-5}
$P(1C)$	8.84×10^{-5}	9.02×10^{-5}	8.00×10^{-5}
$P(2C)$	1.76×10^{-5}	1.83×10^{-5}	1.44×10^{-5}

roughly equal to those for the signal and normalization modes. The combined branching ratio for $K_L \to \pi^0\pi^0_D\pi^0_D$ decays is roughly 9.0×10^{-5}.

The distinguishing characteristic of the conversion events is the preference for very small values of the e^+e^- mass of the pair, even relative to the Dalitz pairs. This quantity is closely correlated with the track separation at the first drift chamber. The level of conversion contamination, following the cuts discussed already, can be seen in Figure 6.16. This plot shows a comparison between the normalization mode data and Monte Carlo events in the distribution of minimum track separations at DC1. The histograms have been normalized above 2 mm. The discrepancy in the first bin is due to conversion events in the data.

The striking difference in the distribution of track separations between the normalization mode Monte Carlo and single conversion Monte Carlo, passing all cuts for the normalization mode, can be seen in Figure 6.17. While the normalization mode contains a significant fraction of events with small separations, nearly the entire distribution for conversion events is below 2 mm. Only 0.26 % of single conversion events that pass all other cuts, have track separations greater than 2 mm. All remaining double conversion events are eliminated. This requirement has a significant impact on the both the signal and normalization modes as well. The efficiency of this cut for normalization

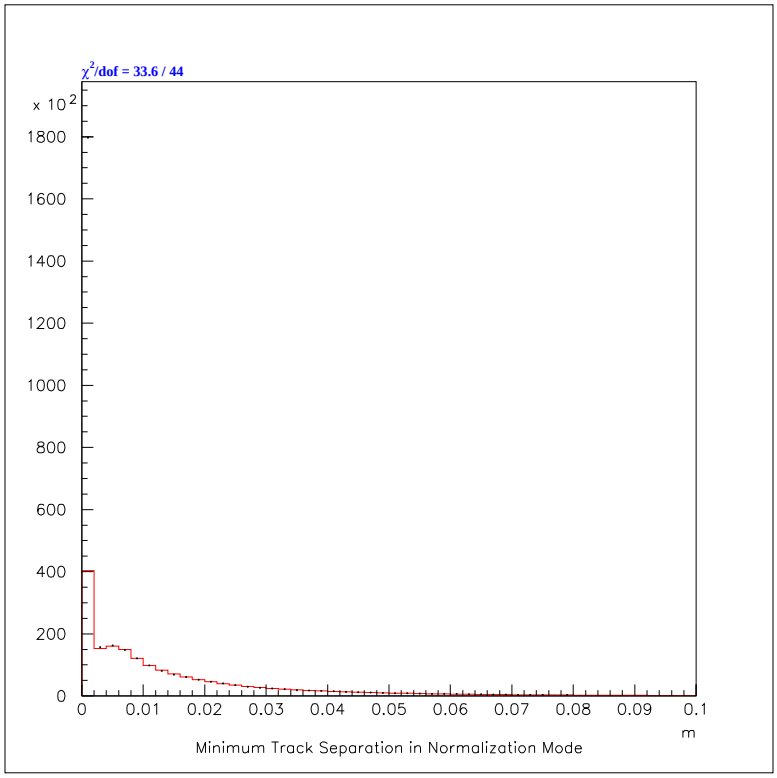

Figure 6.16: Comparison between normalization mode data (points with error bars) and normalization mode Monte Carlo (histogram) in the minimum track separation at DC1. The disagreement in the first bin is due to the presence of conversion events in the data.

mode events is just 78.4 %.

Table 6.3 lists all of the selection criteria applied to the data. The efficiency of the cuts is addressed in the next section. Systematic biases resulting from analysis cuts are discussed later in the chapter.

6.2 Acceptance Calculation

The acceptance for a given decay mode has two components, a geometrical acceptance and a selection efficiency. The geometrical acceptance is quite simple and only depends on the detector geometry and the kaon energy scale. As a rule of thumb, the geometrical acceptance is roughly 50% per particle. For an 8–body final state, a good estimate of the geometrical acceptance is $0.5^8 \sim 0.4\%$.

Table 6.3: List of selection criteria. The three sets of cuts are: fiducial, kinematic, and background.

Variable	Accepted Range
Minimum Cluster Energy	$> 2\,\mathrm{GeV}$
Total CsI Energy	$40 - 210\,\mathrm{GeV}$
Minimum Cluster Separation	$> 5\,\mathrm{cm}$
Minimum Track Momentum	$> 2\,\mathrm{GeV/c}$
E/p	$0.93 - 1.07$
χ^2_{vertex}	< 40
χ^2_{magnet}	< 100
z-Vertex Position	$94 - 157\,\mathrm{m}$
Total Invariant Mass	$480 - 515\,\mathrm{MeV/c^2}$
P_T^2	$< 800\,\mathrm{MeV^2}/c^2$
Pairing χ^2	< 12
Minimum Track Separation	$> 2\,\mathrm{mm}$

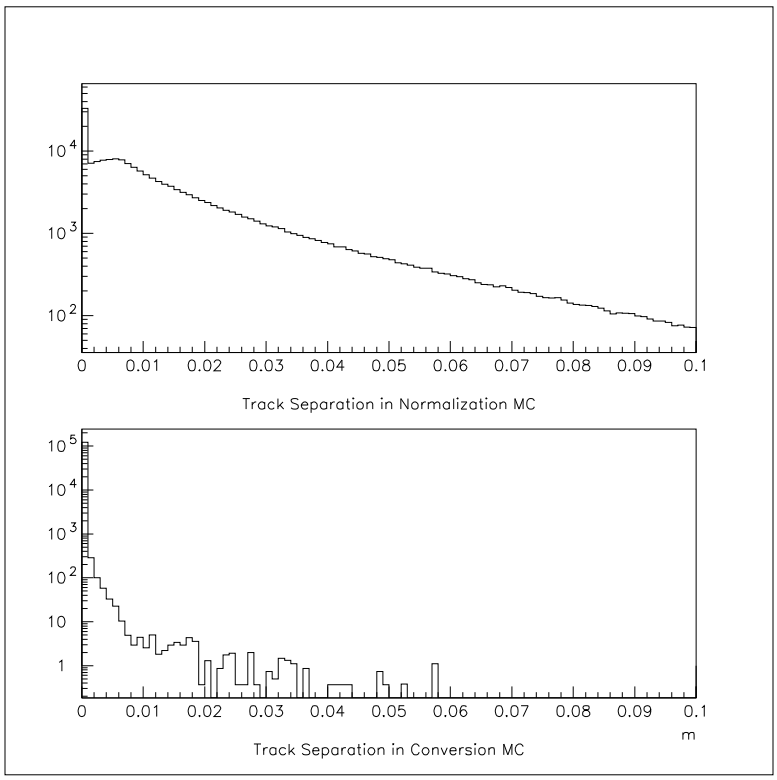

Figure 6.17: Distribution of minimum track separation at DC1 on a logarithmic scale. The top plot contains normalization mode Monte Carlo events. The bottom plot contains single conversion Monte Carlo events.

The selection efficiency is divided between the trigger and the offline analysis requirements. The Monte Carlo events must pass a trigger simulation, the 2e-nclus crunch, the 4–track split, and the selection cuts discussed in the previous section. The efficiency of these filters is the subject of the following subsections.

6.2.1 Trigger Acceptance

All three levels of the online trigger are simulated. The efficiency of each level for $K_L \to \pi^0\pi^0_D\pi^0_D$ events is shown in Table 6.4. The low acceptance at L1 is a reflection of the kaon energy distribution, much of which is below the ETOT trigger requirement. Low energy events also tend to have decay particles hit vetoes. The most restrictive element of the L1 trigger is the VV' requirement, due to low momentum tracks being

Table 6.4: Trigger efficiency for signal and normalization Monte Carlo events in the three run periods. The trigger acceptance is the product of the three efficiencies.

Dataset	Mode	L1	L2	L3	Trigger Acceptance
Win97	DD	9.3 %	92.7 %	71.1 %	6.1 %
	2D	9.4 %	93.8 %	72.3 %	6.4 %
Sum97	DD	8.7 %	94.2 %	71.2 %	5.8 %
	2D	8.9 %	94.9 %	72.8 %	6.1 %
99	DD	10.7 %	95.8 %	75.0 %	7.7 %
	2D	10.8 %	96.2 %	76.1 %	7.9 %

bent out of the detector at the analysis magnet.

The L2 and L3 efficiencies are quite high for signal events. The losses at L2 are due to either the number of hits in the y-view or the number of hardware clusters. The L3 losses are associated with the inability to find a vertex (only needs to be a 2-track vertex). Roughly 8 % of the events show no evidence of having a good vertex. The largest loss is due to inadequate matching between tracks and clusters. Close to 17 % of events with some kind of vertex fail to achieve proper matching. The overall acceptance after the trigger is slightly less than the L1 acceptance.

6.2.2 Filter Efficiencies

The two filters which were used to reduce the data volume were designed to have a high signal efficiency combined with effective background rejection. Table 6.5 summarizes the filter efficiencies and the overall acceptance after the filters, for the three run periods. The crunch code performs a full reconstruction of the event and requires a good 4-track vertex. This requirement accounts for the majority of the loss. The 4-track split only requires that the total number of clusters is equal to eight. The efficiency is a direct reflection of the distribution of number of clusters. The acceptance for events passing the split is about 0.5 % in 97 and 0.8 % in 99.

Table 6.5: Filter efficiency for signal and normalization Monte Carlo events in the three run periods. The filter acceptance is the product of the two filter efficiencies and the trigger acceptance.

Dataset	Mode	Crunch	Split	Filter Acceptance
Win97	DD	25.4 %	27.2 %	0.42 %
	2D	30.2 %	26.2 %	0.50 %
Sum97	DD	26.2 %	28.4 %	0.43 %
	2D	30.6 %	27.9 %	0.52 %
99	DD	32.5 %	27.8 %	0.69 %
	2D	36.7 %	27.3 %	0.79 %

6.2.3 Analysis Efficiency

The final selection requirements are intended to take events with the correct topology and find good candidates for the two modes of interest. The efficiency of most cuts is quite high for signal and normalization Monte Carlo events. The one exception is the track separation cut, which reduces both samples by roughly 20 %. I will define the efficiency of a cut as the number of MC events which pass all cuts divided by the number of events that pass all cuts other than the one in question. Table 6.6 summarizes the efficiencies of all analysis cuts for normalization mode MC events.

6.2.3.1 Relationship between the Acceptance and x_{4e}

In light of the radiative events with soft photons, the branching ratio is defined for a particular photon sensitivity, and therefore the acceptance must be defined accordingly. The double Dalitz branching ratio is measured for $x_{4e} > 0.9$. The acceptance for double Dalitz events is then the number of events which pass all selection requirements divided by the total number generated with $x_{4e} > 0.9$. Events that pass all cuts but were generated with $x_{4e} < 0.9$ represent a background that will be subtracted. Table 6.7 summarizes the acceptance calculation for the double Dalitz sample.

Similarly, the single Dalitz branching ratio in the quantity R must also be appropriately qualified. The results which contribute most to the world average all claim

Table 6.6: Efficiencies for each cut separately for both the signal and normalization mode Monte Carlo events.

	Efficiency (%)	
Cut	DD	2D
Minimum Cluster Energy	96.30	93.49
Total CsI Energy	99.34	99.32
Minimum Cluster Separation	96.04	95.87
Minimum Track Momentum	98.50	98.65
E/p	89.28	89.10
χ^2_{vertex}	97.29	96.90
χ^2_{magnet}	94.42	93.53
z-Vertex Position	97.70	97.55
Total Invariant Mass	84.17	87.29
P_T^2	88.84	87.28
Pairing χ^2	77.67	75.63
Minimum Track Separation	74.28	72.71
All cuts together	50.47	42.72

Table 6.7: Acceptance numbers for $K_L \to \pi^0\pi^0\pi^0_{DD}$. The fraction of events generated with $x_{4e} > 0.9$ is 96.242 % of the total. The errors on the acceptance numbers are from Monte Carlo statistics.

	Win97	Sum97	99
$N_{\rm gen},\ 4e$	18,395,530	13,039,676	30,762,515
$N_{\rm gen},\ 4e\gamma$	4,231,521	3,003,896	7,077,099
$N_{\rm gen},\ 4e + 4e\gamma$	22,627,051	16,043,572	37,839,614
$N_{\rm gen},\ x_{4e} > 0.9$	21,776,726	15,440,655	36,417,601
$N_{\rm acc}$	42,624	29,880	116,931
ϵ_{DD}	$(1.957 \pm 0.009) \times 10^{-3}$	$(1.935 \pm 0.011) \times 10^{-3}$	$(3.211 \pm 0.009) \times 10^{-3}$

Table 6.8: Acceptance numbers for $K_L \to \pi^0\pi^0_D\pi^0_D$. The errors on the acceptance numbers are from Monte Carlo statistics.

	Win97	Sum97	99
$N_{\rm gen}$	70,560,000	47,040,000	176,400,000
$N_{\rm ana}$	150,344	99,903	576,210
ϵ_{2D}	$(2.131 \pm 0.005) \times 10^{-3}$	$(2.124 \pm 0.007) \times 10^{-3}$	$(3.266 \pm 0.004) \times 10^{-3}$

to include radiative effects for all values of x_γ. That is they measure the total rate of $\pi^0 \to e^+e^-\gamma$ and $\pi^0 \to e^+e^-\gamma\gamma$. The radiative correction was discussed in Chapter 2. What this means for the acceptance is that the number of generated events includes all events of each type. Table 6.8 lists the number of events generated, number reconstructed, and the acceptance for the normalization mode.

6.3 Background Studies

There are two remaining backgrounds after all cuts have been applied. The main background is from cross-overs between signal and normalization mode events. Occasionally, a $K_L \to \pi^0\pi^0_D\pi^0_D$ event has a better χ^2 when reconstructed as $K_L \to \pi^0\pi^0\pi^0_{DD}$, or vice-versa. This effect is illustrated in Figure 6.18, which shows the distribution of the best χ^2 for both $K_L \to \pi^0\pi^0\pi^0_{DD}$ and $K_L \to \pi^0\pi^0_D\pi^0_D$ Monte Carlo events.

The predicted acceptance for crossover events was calculated from the same Monte Carlo used to get the signal and normalization acceptance quoted above. The acceptance numbers are given in Table 6.9. The acceptances are accurate to roughly 10% in 97 and 4% in 99.

Table 6.9: Monte Carlo acceptances for crossover events.

Dataset	$\epsilon_{\rm 2D}^{\rm in\ DD\ MC}$	$\epsilon_{\rm DD}^{\rm in\ 2D\ MC}$
Win97	$(2.59 \pm 0.11) \times 10^{-5}$	$(1.90 \pm 0.16) \times 10^{-6}$
Sum97	$(2.78 \pm 0.13) \times 10^{-5}$	$(2.13 \pm 0.21) \times 10^{-6}$
99	$(5.13 \pm 0.12) \times 10^{-5}$	$(3.25 \pm 0.14) \times 10^{-6}$

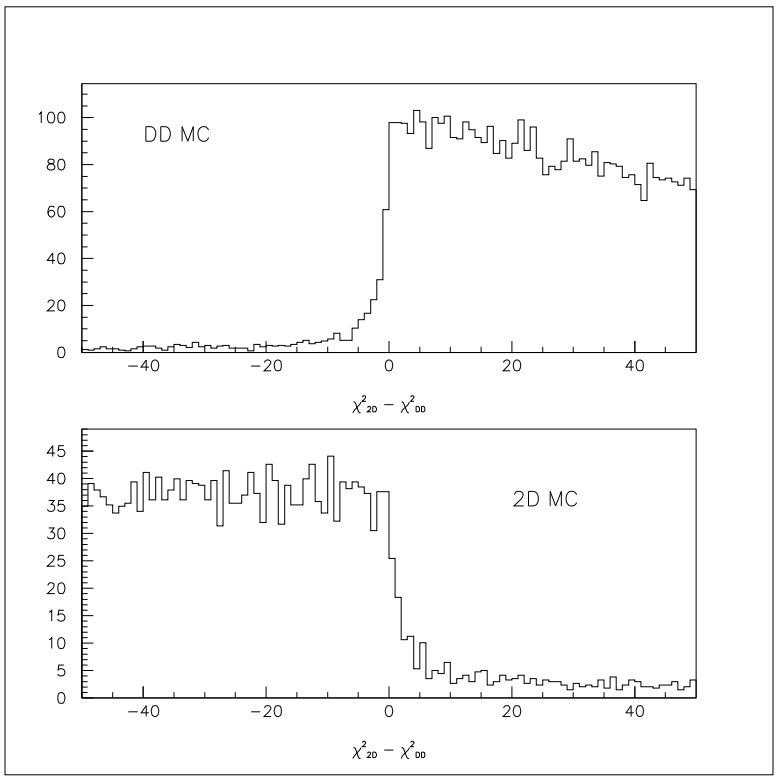

Figure 6.18: Distribution of the difference between the best pairing χ^2 assuming $K_L \to \pi^0\pi^0_D\pi^0_D$ and the best χ^2 assuming $K_L \to \pi^0\pi^0\pi^0_{DD}$, for $K_L \to \pi^0\pi^0\pi^0_{DD}$ Monte Carlo events on top, and $K_L \to \pi^0\pi^0_D\pi^0_D$ events on bottom. The difference should be positive for signal events and negative for normalization events. The crossover background is indicated by the small tails on the wrong side of zero.

The other significant background is from single Dalitz events with an external conversion upstream of the spectrometer. Even with the large reduction from the track separation, the MC predicts that this source accounts for roughly 1 % of the final event sample.

Single Dalitz backgrounds were studied by generating $K_L \to \pi^0\pi^0\pi^0_D$ events and then forcing one of the photons to convert before DC1. A randomly selected photon was required to convert in one of three places based on the number of radiation lengths: the vacuum window, the air gap, or the upstream half of DC1. The probability of this occurring naturally is given in Table 6.2. The efficiency gained by forcing conversions is the inverse of the probability. In Chapter 4, it was claimed that the quantity of air in

Table 6.10: Monte Carlo acceptance for $K_L \to \pi^0\pi^0\pi^0_D$ events reconstructed as signal and as normalization events.

Dataset	$\epsilon_{2D}^{\text{in 1D MC}}$	$\epsilon_{DD}^{\text{in 1D MC}}$
Win97	$(3.64 \pm 0.39) \times 10^{-8}$	$(8.88 \pm 1.94) \times 10^{-9}$
Sum97	$(6.41 \pm 0.64) \times 10^{-8}$	$(1.23 \pm 0.28) \times 10^{-8}$
99	$(8.38 \pm 0.36) \times 10^{-8}$	$(1.73 \pm 0.16) \times 10^{-8}$

Helium bag 1A changed with time during the 97 run. The uncertainty on the number of radiation lengths was quoted as 5%. This uncertainty feeds directly into the background prediction and will later be assigned to a systematic error in the branching ratio.

The acceptance numbers for single Dalitz events reconstructed as both signal and normalization are summarized in Table 6.10. The probability associated with forcing the conversions has been included in the acceptance factor and should not be included in the single Dalitz branching ratio in the background estimate.

Other backgrounds that were considered include $K_L \to \pi^0\pi^0\pi^0$ events with two external conversions and $\pi^0 \to e^+e^-e^+e^-\gamma$ events with $x_{4e} < 0.9$. Both of these sources have negligible contributions after all cuts.

6.4 The Double Ratio

The number of candidate events found in the data is given is Table 6.11. These numbers do include a residual background that is predicted to be less than 1 %. The modified formula for the double ratio, including background sources, is

$$R = \frac{(r\epsilon_{2D}^{2D} - \epsilon_{2D}^{DD}) + \frac{1}{\rho}(r\epsilon_{1D}^{2D} - \epsilon_{1D}^{DD})}{\epsilon_{DD}^{DD} - r\epsilon_{DD}^{2D}}, \tag{6.5}$$

where r is the ratio in observed signal to normalization events and $\rho = \Gamma_{ee\gamma}/\Gamma_{\gamma\gamma} = 0.01213$ is the PDG value of the single Dalitz branching ratio.

The calculated values of the double ratio are given in Table 6.12 along with the statistical errors and the combined result. The χ^2/dof on the average is 1.6.

Table 6.11: Number of candidate events found in the data after all cuts.

Dataset	$K_L \to \pi^0\pi^0_D\pi^0_D$	$K_L \to \pi^0\pi^0\pi^0_{DD}$
Win97	26011	5429
Sum97	19879	4152
99	95361	20930
Total	141251	30511

Table 6.12: Double ratio R and statistical error for each dataset and the combined result, which has a $\chi^2/\mathrm{dof} = 1.6$.

Dataset	R
Win97	0.2263 ± 0.0034
Sum97	0.2289 ± 0.0039
99	0.2231 ± 0.0017
Combined	0.2244 ± 0.0014

Table 6.13: Estimate of residual background levels and final event samples.

	Win97		Sum97		99	
	DD	2D	DD	2D	DD	2D
Signal	5397	25904	4123	19771	20795	94828
XO	23	71	20	59	94	332
$1C$	9	36	9	49	41	201

With the ratio R in hand, the predicted number of background events from each source can be calculated and subtracted from the number of candidate events. Table 6.13 presents the background estimates along with the number of background-subtracted signal and normalization events.

The combined number of background-subtracted events is $30,315$ for the signal and $140,503$ for the normalization. The estimated background contamination after all cuts is less than $1\,\%$ as claimed.

6.4.1 Comparisons between Data and Monte Carlo

This section presents a study of the double ratio R as a function of the different cut quantities. Each plot is the double ratio of the ratio of data to MC in the signal mode to the ratio of data to MC in the normalization mode in some variable. In all cases, the cut on the variable under study has been removed. The location of the cut is indicated.

The double ratio as a function of each of the cluster variables is shown in Figures 6.19, 6.20, and 6.21. The bin-to-bin agreement is indicated by the χ^2 of the fit to a constant. The ratio is very flat in both the minimum cluster energy and the total energy. The minimum cluster separation however shows a possible trend under $7\,\mathrm{cm}$. Disagreements of this type will be systematically treated later in this chapter.

The tracking variables are shown in Figures 6.22, 6.23, and 6.24. The agreement in the E/p distributions is remarkable considering the disagreement previously shown

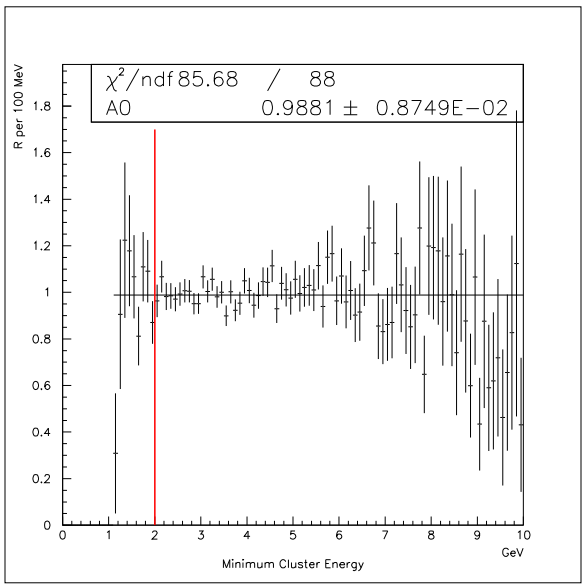

Figure 6.19: Double ratio as a function of the minimum cluster energy.

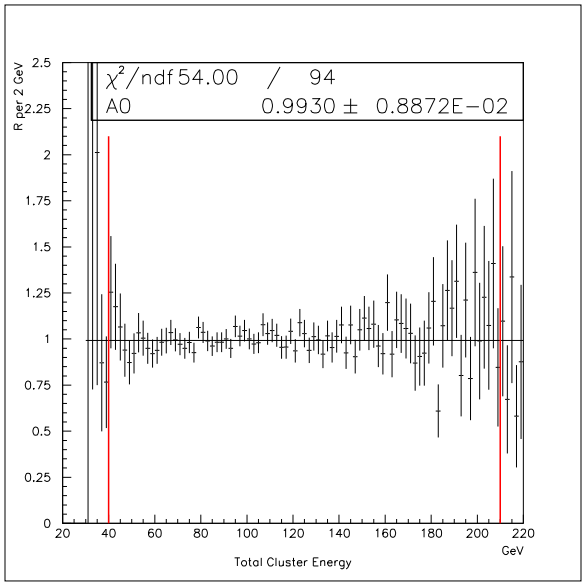

Figure 6.20: Double ratio as a function of the total energy.

for the normalization mode.

Figures 6.25, 6.26, and 6.27 display the double ratio as a function of the vertex χ^2, the magnet χ^2, and the z-vertex. The Monte Carlo apparently does the same poor job of predicting the χ^2 shapes in both modes. There is no indication of biases in the

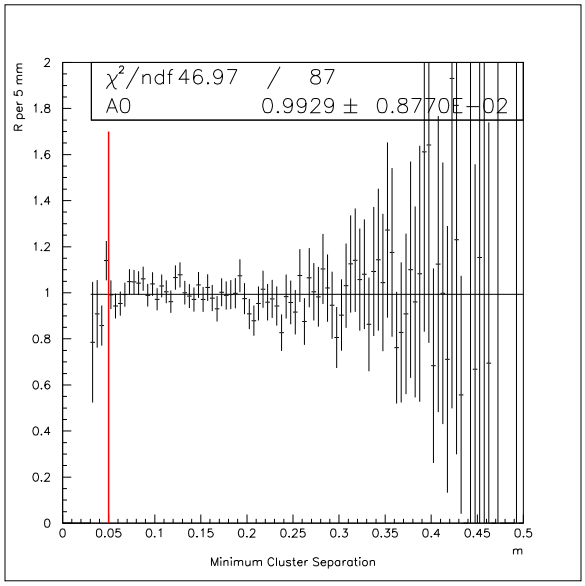

Figure 6.21: Double ratio as a function of the minimum cluster separation.

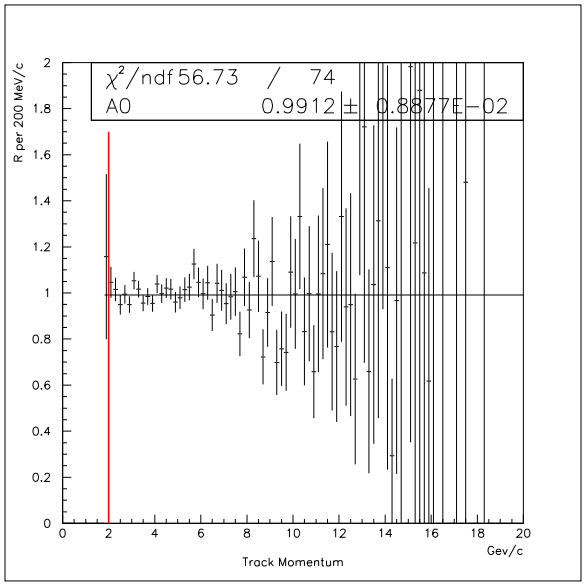

Figure 6.22: Double ratio as a function of the minimum track momentum.

position of the z-vertex.

Next, the kinematic variables are discussed. The ratios are shown in Figures 6.28, 6.29, and 6.30. While both of the kaon parameters are entirely flat, there is an indication of a downward slope in the ratio in terms of the pairing χ^2. This will be studied further

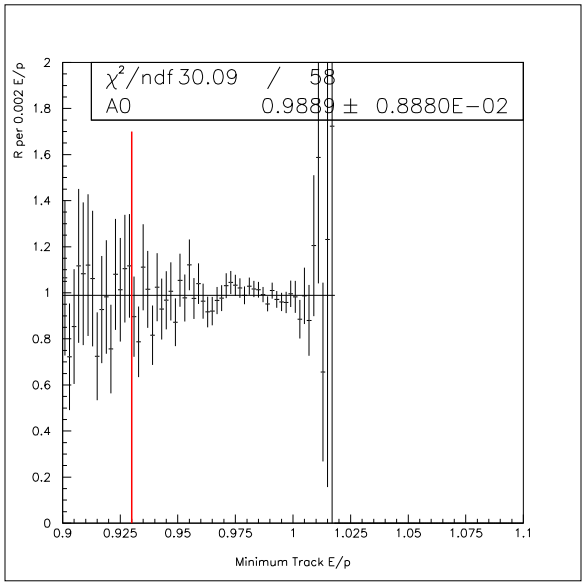

Figure 6.23: Double ratio as a function of the minimum E/p.

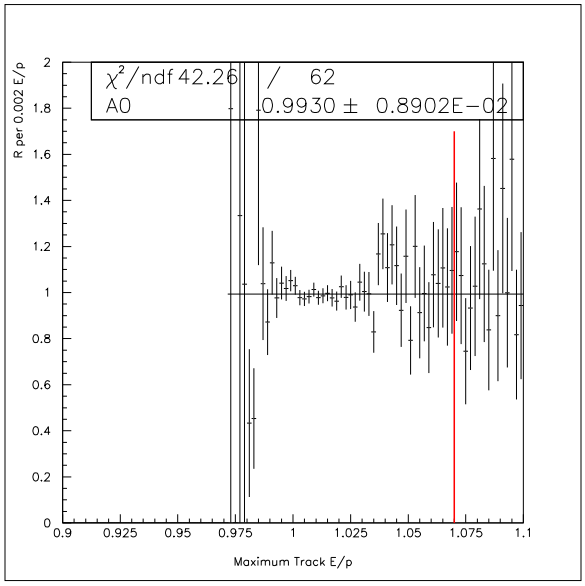

Figure 6.24: Double ratio as a function of the maximum E/p.

later in the discussion of systematic errors.

The final selection requirement is the background eliminating track separation cut. The double ratio as a function of this quantity is shown in Figure 6.31. The average is dominated by the first bin. While there is some fluctuation in the region of

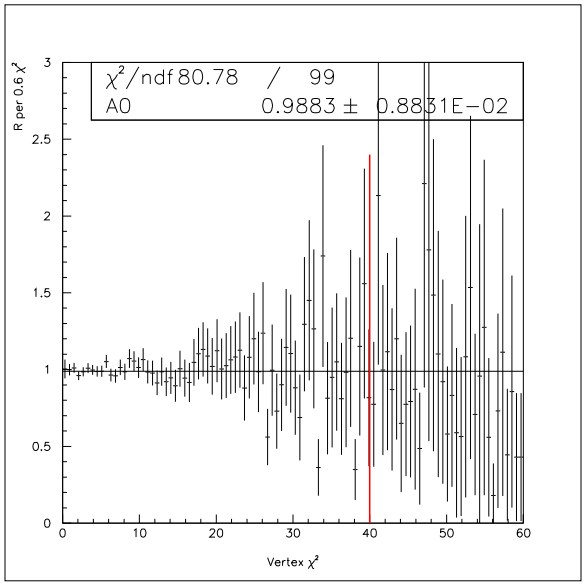

Figure 6.25: Double ratio as a function of the vertex χ^2.

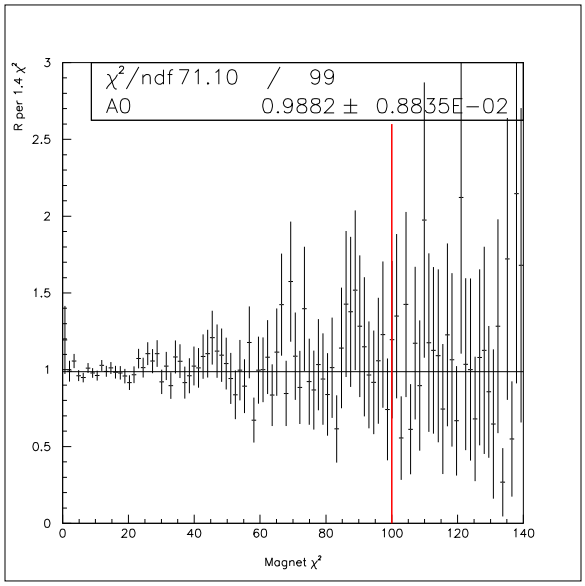

Figure 6.26: Double ratio as a function of the magnet χ^2.

the cut, there is no indication of a bias. The next section describes the many systematic studies performed, including a study of the dependence of the double ratio on the chosen values of the selection cuts.

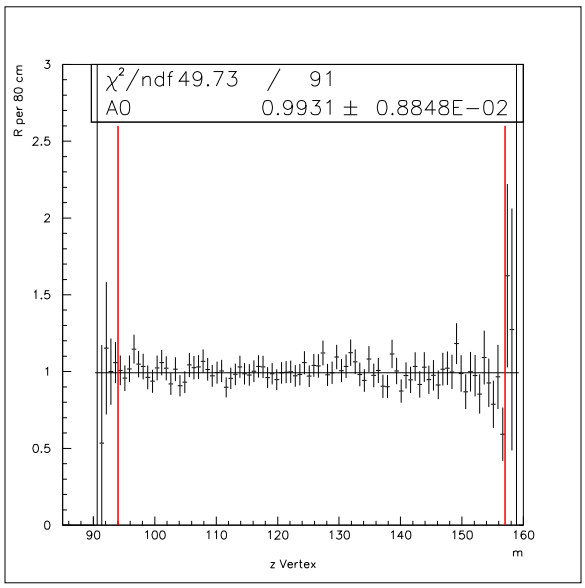

Figure 6.27: Double ratio as a function of the vertex z-position.

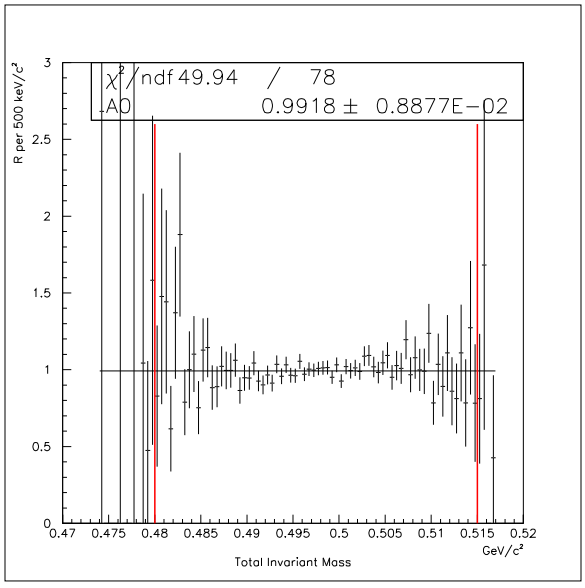

Figure 6.28: Double ratio as a function of the total mass.

6.5 Systematic Studies

The systematic studies have been divided into three categories: theory, simulation, and analysis. Sources of theoretical errors include uncertainties on the parameters of the tree-level matrix element as well as neglected higher order contributions at the

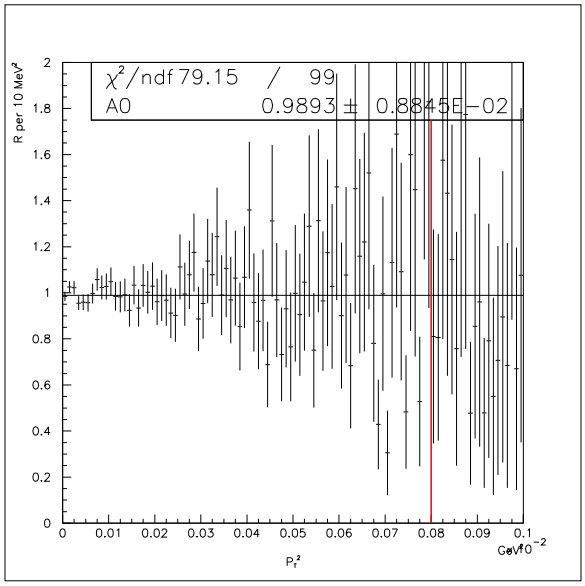

Figure 6.29: Double ratio as a function of the transverse momentum squared.

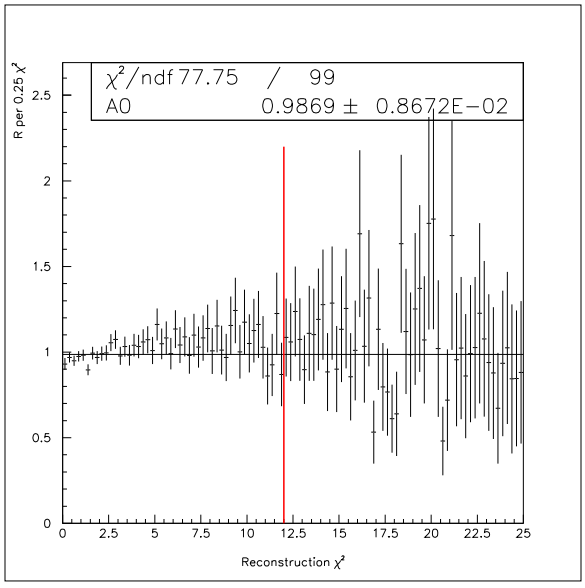

Figure 6.30: Double ratio as a function of the pairing χ^2.

2–loop level. Simulation errors are associated with details of the detector simulation such as the amount of material. Analysis uncertainties are meant to cover errors resulting from a particular choice of selection cuts. Each of these areas is explained in detail below.

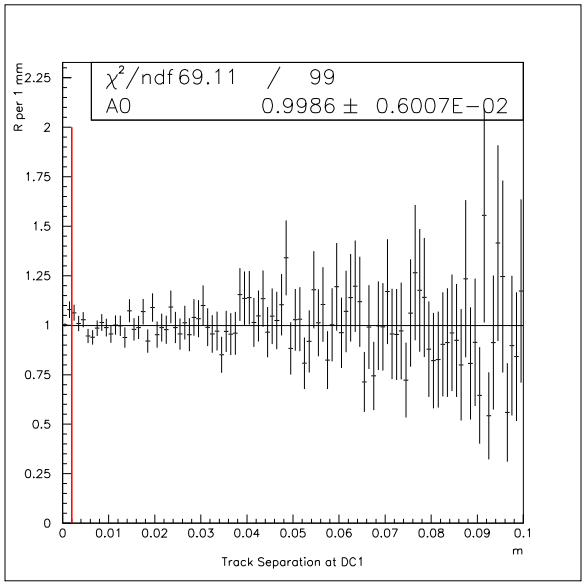

Figure 6.31: Double ratio as a function of the minimum track separation.

6.5.1 Theoretical Uncertainties

The Monte Carlo generates events based on the QED matrix elements described in Chapter 2. The signal mode requires three parameters be specified while the normalization mode depends on only one of those three. Additionally, the equations governing the processes are only known to first order. Higher order contributions, like 2–loop diagrams and radiation of multiple photons, are neglected. While the branching ratio measurement is not overly sensitive to these effects, an estimate of their possible size is desirable.

6.5.1.1 Matrix Element Parameters

The three parameters of the $\pi^0\gamma^*\gamma^*$ coupling are not precisely known. The next chapter will describe a measurement of them from the $\pi^0 \to e^+e^-e^+e^-$ data. The true branching ratio depends weakly on their exact values and the signal acceptance is modified only slightly when the parameters are varied within reason. Increasing the scalar component in the Monte Carlo events had negligible impact on the acceptance, as

did varying the form factor slope parameter within the measured errors. The systematic error is limited to less than 0.01 %.

6.5.1.2 Radiative Corrections

The impact of higher order corrections is estimated simply by taking the difference between tree-level and first-order acceptances and squaring. The rationale for this is that higher order terms are suppressed by factors of $\alpha_{\rm EM}$ and are unlikely to be significant. The change in the acceptance between lowest-order and first-order is roughly 2 %, resulting in an uncertainty due to second-order effects of approximately 0.04 %.

6.5.2 Simulation Uncertainties

The Monte Carlo simulates both the response of the decay particles to the detector material and the response of the detector to the particles. Neither of these simulations is perfect and in order to gauge the dependence of the result on the assumptions made I vary different aspects of the MC and record the change in the acceptance. For a full description of the techniques used in this section see Reference [29].

6.5.2.1 Material Estimates

The Monte Carlo acceptance depends on the amount of material simulated within the detector. As discussed previously, the number of radiation lengths for most elements is precisely known, however helium bag 1A is only known to about 5 %. The composition of the bag in the simulation was varied from all air to all helium. The ratio of acceptances changes by 3 % over the full range. The uncertainty on the ratio is then 5 % of the overall change, or 0.15 %.

6.5.2.2 Chamber Resolution

The chamber resolution enters into the reconstruction at vertexing through the χ^2 distributions. The average resolution can be inferred from the central portion of the SOD distributions. To study the effect of mis-modeling the resolution, the MC hits were smeared in the analysis in such a way as to make the SOD distributions 10 % wider. The ratio of acceptances changes by 0.84 %.

6.5.2.3 Chamber Inefficiencies

The tracking efficiency can bias any comparison of modes with different numbers of tracks or even different illuminations. The illuminations between signal and normalization are nearly identical. The inefficiencies are modeled in the Monte Carlo with a DCMAP, as discussed previously. To account for possible differences between the two modes, MC events are reweighted based on track position. The event receives a weight based on the likelihood that all four tracks would have been reconstructed given a higher DCMAP weight. Based on studies of chamber illuminations, a 20 % variation in the maps is unlikely but possible. At this level, the ratio of acceptances changes by only 0.04 %.

6.5.2.4 Calorimeter Resolution

The cluster energy resolution determines the width of the E/p and $M_{\gamma\gamma}$ distributions, among other things. To study the impact of mis-modeling this resolution in the MC, clusters in MC events are smeared in such a way as to make the E/p distribution 1 % wider. The amount of smearing was chosen by looking at the variation in the E/p width from run to run. With this amount of degradation, the ratio of acceptances changes by just 0.02 %.

6.5.3 Analysis Uncertainties

The final category of systematic effects is intended to cover biases introduced by particular choices of the selection procedure and cuts.

6.5.3.1 Variation of Cuts

Each cut was varied within reason while all other cuts were held fixed. Significant changes in the ratio R are defined as variations of more than one uncorrelated, statistical error. In the case in which the change is significant, the excess is taken as a bias. The total error is then the sum in quadrature of the individual biases. The bias due to cut variation is roughly 0.80 % in Win97, dominated by loosening the E/p cut. In Sum97, the bias is 1.49 %, due entirely to the minimum cluster separation cut. All of the biases are small in the 99 dataset and the total is just 0.19 %. The weighted average of the uncertainties in the three datasets is 0.21 %.

6.5.3.2 Monte Carlo Statistics

The calculation of the signal and normalization acceptances was based on Monte Carlo samples equal to roughly five times the data. The total number of reconstructed signal Monte Carlo events is 184865, while the number of normalization events is 826457. The combined error is then 0.26 %.

6.5.3.3 Background Estimates

The background estimates suffer from some imprecision because the background contamination is small. If the background acceptances are allowed to individually fluctuate by a factor of two, the ratio R is remains stable at the 0.15 % level.

Table 6.14: Summary of systematic uncertainties in the ratio R.

Category	Source	Error (%)
Theory	δ_{rad}	0.04
	$\mathcal{M}(\alpha, \zeta, \delta)$	0.01
Simulation	DC Res	0.84
	Material	0.15
	DC Ineff	0.04
	CsI Res	0.02
Analysis	MC Stat	0.26
	Cut Var	0.21
	Backgrounds	0.15
Total Uncertainty		0.93

6.5.4 Combined Systematic Uncertainty

A summary of systematic effects is shown in Table 6.14. The combined error is obtained by adding the individual uncertainties in quadrature.

There is one additional source of error in any measurement which relies on the measured value of the single Dalitz branching ratio, which is known to only 2.72 %. The $\pi^0 \to e^+e^-e^+e^-$ branching ratio is obtained from the double ratio by multiply by two factors of the $\pi^0 \to e^+e^-\gamma$ branching ratio, resulting in an external systematic uncertainty of 5.44 % and a combined error of 5.64 %.

6.6 Branching Ratio Result

The final result for the ratio R, with both the statistical error and an estimate of the systematic error, is

$$R = \frac{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9) \cdot B(\pi^0 \to \gamma\gamma)}{B^2(\pi^0 \to e^+e^-\gamma)}$$
$$= 0.2244 \pm 0.0014 \text{ (stat)} \pm 0.0021 \text{ (syst)} \quad (6.6)$$

The combined uncertainty on this quantity is 1.12 %. The branching ratio is quoted in two ways, using the known branching ratios of $B(\pi^0 \to \gamma\gamma) = (0.98798 \pm 0.00032)$ and

$B(\pi^0 \to e^+e^-\gamma) = (0.01198 \pm 0.00032)$. The uncertainty in the $\pi^0 \to \gamma\gamma$ branching ratio is negligible compared to other errors, especially the single Dalitz error.

The ratio of the double Dalitz rate to the single Dalitz rate squared is

$$\frac{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9)}{B^2(\pi^0 \to e^+e^-\gamma)} = 0.2271 \pm 0.0014 \text{ (stat)} \pm 0.0021 \text{ (syst)} \tag{6.7}$$

The form of the branching ratio that has been published before is relative to the two-photon decay. This ratio is found to be

$$\frac{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9)}{B(\pi^0 \to \gamma\gamma)} = (3.299 \pm 0.021 \text{ (stat)} \pm 0.030 \text{ (syst)} \pm 0.179 \text{ (ext)}) \times 10^{-5} \tag{6.8}$$

where the last error is due to the 2.72 % uncertainty in the single Dalitz branching ratio. The combined error changes from 1.12 % to 5.55 %, due to the uncontrollable external uncertainty.

The branching ratio for $x_{4e} > 0.9$ is predicted to be 94.202 % of the total rate for all values of x_{4e}. Making this theoretical extrapolation, I find

$$\frac{B(\pi^0 \to e^+e^-e^+e^-(\gamma))}{B(\pi^0 \to \gamma\gamma)} = (3.502 \pm 0.194) \times 10^{-5}. \tag{6.9}$$

6.6.1 Comparison to Previous Results

The result agrees well with the previous Samios result of $(3.18 \pm 0.30) \times 10^{-5}$. The difference between the two results is 0.28 ± 0.36. While the statistical uncertainty has improved by a factor of roughly 12, the overall error, including the single Dalitz branching ratio, is only better by a factor of 1.6.

The new analysis is also in good agreement with the preliminary KTeV results. The 97 number changed from 3.31×10^{-5} to 3.34×10^{-5}, while in 99 the result changed from 3.27×10^{-5} to 3.28×10^{-5}. The main difference between the result presented here and the preliminary KTeV result is in the definitions of the acceptances for both the signal and normalization modes. This is not a true change but rather a new definition of the observed branching ratio.

Finally, the result is in good agreement with the theoretical prediction, including radiative corrections. For $x_{4e} > 0.9$, the predicted branching ratio is 3.331×10^{-5}, while for all x_{4e} it is 3.536×10^{-5}.

Chapter 7

The $\pi^0\gamma^*\gamma^*$ Coupling

It was shown in Chapter 2 that the $\pi^0 \to e^+e^-e^+e^-$ final state can be used as a probe of the $\pi^0\gamma^*\gamma^*$ vertex. The coupling derived there has the form

$$C_{\mu\nu\rho\sigma} \propto \tilde{g} f(x_1, x_2; \alpha)[\cos\zeta \epsilon_{\mu\nu\rho\sigma} + \sin\zeta \mathrm{e}^{i\delta}(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho})], \tag{7.1}$$

where α parameterizes the momentum dependence and ζ and δ describe the relative strength and phases of the two allowed couplings. The mixing angle ζ takes values between 0 and $\pi/2$, while the phase difference δ lies between 0 and π. The point $\zeta = 0$ describes a purely CP–conserving interaction (assuming the π^0 is a pseudoscalar), while $\zeta = \pi/2$ would describe a pure CP–violating decay. In either case δ, the phase difference, is unobservable since there is no mixing. For intermediate values of ζ, the phase difference becomes a measure of CPT conservation. If CPT is conserved then the phase difference, in the absence of absorptive decay amplitudes, should be zero (or π). For π^0 decays, any absorptive process would occur at higher order in α_{EM}, therefore a large value of δ would be a clear indication of new physics.

This chapter will describe how the three parameters of the coupling may be extracted from the data. The method of likelihood maximization is well suited for multi-dimensional problems such as this one. I will discuss the general properties of likelihood functions before deriving the form specific to the $\pi^0\gamma^*\gamma^*$ coupling. This will be followed by a summary of Monte Carlo studies in which independent sets of MC

events, generated with known values of the parameters, were fit in order to identify any biases intrinsic to the method. One bias, due to the resolution on the angle between the two e^+e^- pairs, was discovered and a correction has been calculated. The rest of the chapter will cover the measurement from the KTeV dataset along with a summary of systematic studies.

7.1 Likelihood Estimation

The likelihood function is usually described as the joint p.d.f. of a given set of data [13]. In almost all cases, it is the logarithm of the likelihood function that one works with. I will therefore define the likelihood function to be

$$L(\boldsymbol{\mu}) = -\ln \prod_i^N f(\boldsymbol{x}_i; \boldsymbol{\mu}), \tag{7.2}$$

where $f(\boldsymbol{x}_i; \boldsymbol{\mu})$ is a p.d.f. describing the probability of measuring $\boldsymbol{x}_i$ given the parameters $\boldsymbol{\mu}$ and the sum runs over the observed data. I have inserted a minus sign simply because I prefer to minimize, rather than maximize, the function. The point in the parameter space at which the likelihood function takes its minimum value is then the most likely value of $\boldsymbol{\mu}$ given the observed $\boldsymbol{x}_i$. The logarithm is convenient since it turns the product into a sum,

$$L(\boldsymbol{\mu}) = -\sum_i^N \ln f(\boldsymbol{x}_i; \boldsymbol{\mu}). \tag{7.3}$$

The most likely value of $\boldsymbol{\mu}$ is calculated by finding the minimum in each parameter,

$$\frac{\partial L(\boldsymbol{\mu})}{\partial \mu_i} = 0, \tag{7.4}$$

and the uncertainties on the estimate are found by calculating the inverse of the covariance matrix, which is approximated by

$$\left(V^{-1}\right)_{ij} = -\frac{\partial^2 L}{\partial \mu_i \partial \mu_j}, \tag{7.5}$$

evaluated at the minimal point. A more general method of estimating errors is to consider changes in the likelihood function. Assuming that the true value is known,

Table 7.1: Change in likelihood function corresponding to 1, 2, and 3 σ significance for 1, 2, or 3 free parameters.

N_{par}	1	2	3
$1\,\sigma$	0.50	1.15	1.77
$2\,\sigma$	2.00	3.09	4.02
$3\,\sigma$	4.50	5.92	7.08

the difference in the likelihood function evaluated at its minimum and at the true value should be less than 1/2 approximately 68% of the time and less than 2 about 95% of the time. Table 7.1 gives the significance of changes in the likelihood function for 1, 2, or 3 free parameters in terms of standard deviations [13].

As a simple, analytic example consider a Gaussian distributed p.d.f. in one dimension,

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(x-\mu)^2/2\sigma^2}. \tag{7.6}$$

The likelihood function for a set of observed x_i is

$$\begin{aligned} L(\mu,\sigma^2) &= \sum_i^N \left[\ln\sqrt{2\pi\sigma^2} + \frac{(x_i-\mu)^2}{2\sigma^2}\right], \\ &= \frac{N}{2}\ln(2\pi) + \frac{N}{2}\ln\sigma^2 + \frac{1}{2\sigma^2}\sum_i^N (x_i-\mu)^2. \end{aligned} \tag{7.7}$$

The minimum of the likelihood in μ is given by

$$\frac{\partial L}{\partial \mu} = -\frac{1}{\sigma^2}\sum_i^N (x_i-\mu) = 0. \tag{7.8}$$

This of course is satisfied when

$$\mu = \frac{1}{N}\sum_i^N x_i. \tag{7.9}$$

The equivalent condition on σ^2 is

$$\frac{\partial L}{\partial \sigma^2} = \frac{N}{2\sigma^2} - \frac{1}{2\sigma^4}\sum_i^N (x_i-\mu)^2 = 0, \tag{7.10}$$

which is satisfied if

$$\sigma^2 = \frac{1}{N}\sum_i^N (x_i-\mu)^2. \tag{7.11}$$

Most statistics textbooks point out that the estimate for σ^2 is a biased estimator and should involve a factor of $1/(N-1)$ rather than $1/N$. Of course when N is large these two terms are equivalent. The covariance matrix for the Gaussian p.d.f. can be shown to be

$$V = -\begin{pmatrix} \sigma^2/N & 0 \\ 0 & 2\sigma^4/N \end{pmatrix}. \tag{7.12}$$

The absence of off-diagonal elements indicates that the two parameters are uncorrelated.

The simple Gaussian example has exposed three important issues. First, the value of the likelihood function at its minimum is arbitrary and depends on the size of the dataset among other things. In contrast, the minimum value of a χ^2 function tells you not only the most likely set of parameters but also how well the model actually describes the data. In any fit, it is essential to measure the goodness-of-fit since the estimate of the parameters is meaningless if the model is just wrong. In the likelihood method, it is possible to estimate the goodness-of-fit by comparing the minimum value of the likelihood function for the data to the distribution of minimum values from many, identically sized, Monte Carlo samples. The second point is that the likelihood function must be properly normalized. Leaving out the $1/\sqrt{\sigma^2}$ normalization factor in the example would have lead to a wrong conclusion. However, the factor of $\sqrt{2\pi}$ in the normalization does not affect the position of the minimum. This leads to point three: additive factors which are independent of the parameters can be ignored. With that said, we will now turn to the likelihood function appropriate for the $\pi^0\gamma^*\gamma^*$ coupling.

7.2 Likelihood Function for the $\pi^0\gamma^*\gamma^*$ Coupling

Unlike the simple example discussed above, the likelihood function for the coupling cannot be tackled analytically. Instead the function will be evaluated at select points on a 3–dimensional grid. Before writing down the likelihood function it will be convenient to make a change of variables from the bounded angles ζ and δ to the

unbounded, real variables κ and η, defined as

$$\kappa = \tan\zeta \cos\delta, \tag{7.13a}$$

$$\eta = \tan\zeta \sin\delta, \tag{7.13b}$$

The inverse transformation is

$$\zeta = \tan^{-1}\sqrt{\kappa^2 + \eta^2}, \tag{7.14a}$$

$$\delta = \tan^{-1}(\eta/\kappa). \tag{7.14b}$$

The relationship between the two sets of variables in the κ–η plane is shown in Figure 7.1.

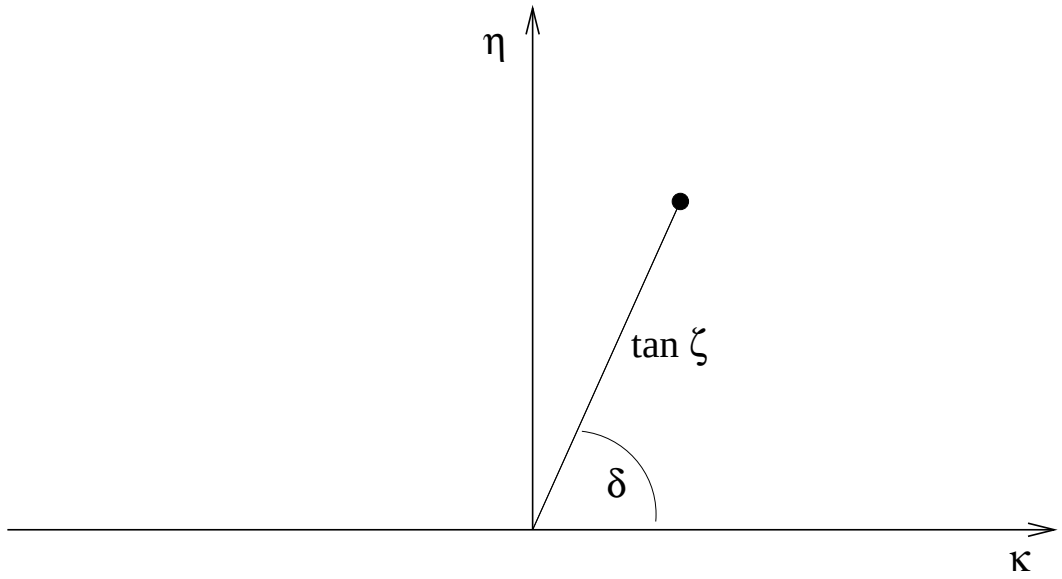

Figure 7.1: κ–η Plane.

In general, the uncertainties on the original set of parameters $\boldsymbol{\mu}$ are related to the uncertainties on the transformed set $\boldsymbol{\mu}'(\boldsymbol{\mu})$ by

$$U_{ij} = \sum_{k,l} \frac{\partial\mu'_i}{\partial\mu_k}\frac{\partial\mu'_j}{\partial\mu_l}V_{kl}, \tag{7.15}$$

where V is the covariance matrix for $\boldsymbol{\mu}'$ and U is the equivalent quantity in terms of $\boldsymbol{\mu}$.

The derivatives are to be evaluated at the minimum point. The derivatives are

$$\frac{\partial\kappa}{\partial\zeta} = \sec^2\zeta\cos\delta, \qquad \frac{\partial\eta}{\partial\zeta} = \sec^2\zeta\sin\delta, \tag{7.16a}$$

$$\frac{\partial\kappa}{\partial\delta} = -\tan\zeta\sin\delta, \qquad \frac{\partial\eta}{\partial\delta} = \tan\zeta\cos\delta. \tag{7.16b}$$

It will be shown later that the variables κ and η are uncorrelated and the following relationships can be used instead,

$$\sigma_\zeta = \frac{\sin 2\zeta}{2\zeta}\sqrt{\cos^4\delta\ \sigma_\kappa^2 + \sin^4\delta\ \sigma_\eta^2}, \tag{7.17a}$$

$$\sigma_\delta = \frac{\sin 2\delta}{2\delta}\sqrt{\sigma_\kappa^2 + \sigma_\eta^2}. \tag{7.17b}$$

In terms of α, κ, and η the coupling is

$$C_{\mu\nu\rho\sigma} = \frac{\tilde{g}f(x_1, x_2; \alpha)}{\sqrt{1+\kappa^2+\eta^2}}[\epsilon_{\mu\nu\rho\sigma} + (\kappa + i\eta)(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho})], \tag{7.18}$$

where $f(x_1, x_2; \alpha)$ is the DIP form factor of Equation (2.11). In principle, κ and η can take on any value, however the tree–level matrix element is symmetric in η (the sign of $\sin\delta$ is irrelevant) so only positive values of η will be considered.

The p.d.f. which describes the probability of measuring the data $\boldsymbol{x}_i$ given the parameters $\boldsymbol{\mu}$ is

$$f(\boldsymbol{x}_i, \boldsymbol{\mu}) = \frac{d\Gamma(\boldsymbol{x}_i, \boldsymbol{\mu})/d\boldsymbol{x}_i}{\Gamma_{\rm obs}(\boldsymbol{\mu})}, \tag{7.19}$$

where $d\Gamma(\boldsymbol{x}, \boldsymbol{\mu})/d\boldsymbol{x}$ is the differential partial width evaluated at $\boldsymbol{x}$ and $\boldsymbol{\mu}$, and $\Gamma_{\rm obs}(\boldsymbol{\mu})$ is the observed partial width, which serves as a normalization factor. The normalization term is just the integral of the numerator and basically accounts for the dependence of the acceptance on the parameters $\boldsymbol{\mu}$.

This normalization can be calculated with weighted Monte Carlo events by noting that the true decay rate is approximated by

$$\Gamma_{\rm true}(\boldsymbol{\mu}) = \frac{1}{N_{\rm gen}}\sum_i^{N_{\rm gen}} d\Gamma(\boldsymbol{x}_i, \boldsymbol{\mu})/d\boldsymbol{x}_i, \tag{7.20}$$

and that the acceptance for observing weighted events is

$$\epsilon(\boldsymbol{\mu}) = \frac{\sum_i^{N_{\rm obs}} d\Gamma(\boldsymbol{x}_i, \boldsymbol{\mu})/d\boldsymbol{x}_i}{\sum_j^{N_{\rm gen}} d\Gamma(\boldsymbol{x}_j, \boldsymbol{\mu})/d\boldsymbol{x}_j}. \tag{7.21}$$

The observed rate is then given by

$$\Gamma_{\rm obs}(\boldsymbol{\mu}) = \frac{1}{N_{\rm gen}} \sum_i^{N_{\rm obs}} d\Gamma(\boldsymbol{x}_i, \boldsymbol{\mu})/d\boldsymbol{x}_i. \tag{7.22}$$

In order to calculate the normalization factor at every trial value of $\boldsymbol{\mu}$, I can simply re–weight the MC events. To do so, I calculate the amplitude for the decay at the generated value $\boldsymbol{\mu}_{\rm gen}$ and for other values of $\boldsymbol{\mu}$. The differential rate for a given $\boldsymbol{\mu}$ can then be written as

$$d\Gamma(\boldsymbol{x}_i, \boldsymbol{\mu}) = \frac{\mathcal{A}(\boldsymbol{x}_i, \boldsymbol{\mu})}{\mathcal{A}(\boldsymbol{x}_i, \boldsymbol{\mu}_{\rm gen})} W(\boldsymbol{x}_i, \boldsymbol{\mu}_{\rm gen}), \tag{7.23}$$

where W is the event weight ($W(\boldsymbol{x}, \boldsymbol{\mu}) = d\Gamma(\boldsymbol{x}, \boldsymbol{\mu})/d\boldsymbol{x}$), and $\mathcal{A}(\boldsymbol{x}, \boldsymbol{\mu})$ is the square of the matrix element. The normalization factor now becomes

$$\Gamma_{\rm obs}(\boldsymbol{\mu}) = \frac{1}{N_{\rm gen}} \sum_i^{N_{\rm obs}} \frac{\mathcal{A}(\boldsymbol{x}_i, \boldsymbol{\mu})}{\mathcal{A}(\boldsymbol{x}_i, \boldsymbol{\mu}_{\rm gen})} W(\boldsymbol{x}_i, \boldsymbol{\mu}_{\rm gen}). \tag{7.24}$$

The likelihood function, after removing additive constant terms can be expressed as

$$L(\boldsymbol{\mu}) = -\sum_i^{N_{\rm data}} \ln \frac{\mathcal{A}(\boldsymbol{x}_i, \boldsymbol{\mu})}{\frac{1}{N_{\rm gen}} \sum_j^{N_{\rm MC}} \frac{\mathcal{A}(\boldsymbol{x}_j, \boldsymbol{\mu})}{\mathcal{A}(\boldsymbol{x}_j, \boldsymbol{\mu}_{\rm gen})} W(\boldsymbol{x}_j, \boldsymbol{\mu}_{\rm gen})}, \tag{7.25}$$

where the sum over j is over the number of observed events in the MC and the sum over i is over the number of observed events in the data sample.

In order to apply this procedure to the $\pi^0 \to e^+e^-e^+e^-(\gamma)$ data sample, I have to have a way of accounting for both non–radiative and radiative events in a consistent manner. The normalization factor must be modified to include contributions from both processes, with the acceptance appropriate for each. The combined observed rate is just

$$\begin{aligned}\Gamma_{\rm obs}(\boldsymbol{\mu}) = \frac{1}{N_{\rm gen}^{4e}} \sum_i^{N_{4e}} \frac{\mathcal{A}_{4e}(\boldsymbol{x}_i, \boldsymbol{\mu})}{\mathcal{A}_{4e}(\boldsymbol{x}_i, \boldsymbol{\mu}_{\rm gen})} W_{4e}(\boldsymbol{x}_i, \boldsymbol{\mu}_{\rm gen}) \\ + \frac{1}{N_{\rm gen}^{4e\gamma}} \sum_i^{N_{4e\gamma}} \frac{\mathcal{A}_{4e\gamma}(\boldsymbol{y}_i, \boldsymbol{\mu})}{\mathcal{A}_{4e\gamma}(\boldsymbol{y}_i, \boldsymbol{\mu}_{\rm gen})} W_{4e\gamma}(\boldsymbol{y}_i, \boldsymbol{\mu}_{\rm gen}),\end{aligned} \tag{7.26}$$

where $\boldsymbol{y}$ is the 8–dimensional phase space of the radiative decays. The event weight for $\pi^0 \to 4e$ events includes radiative corrections, which I will assume are independent of $\boldsymbol{\mu}$. In performing this calculation, I simply look to see if the MC event in question was generated as a $\pi^0 \to 4e$ or a $\pi^0 \to 4e\gamma$.

The numerator of the likelihood function is calculated using the $\pi^0 \to 4e$ amplitude, but with a large photon energy cutoff appropriate to the detector photon sensitivity. The final form of the likelihood function becomes

$$L(\boldsymbol{\mu}) = -\sum_{i}^{N_{\rm data}} \ln \frac{\mathcal{A}_{4e}^{E_{\rm cut}}(\boldsymbol{x}_i, \boldsymbol{\mu})}{\Gamma_{\rm obs}(\boldsymbol{\mu})}. \tag{7.27}$$

The value of the cutoff is chosen to correspond to a value of $x_{4e} = 0.9$, or $E_\gamma^* = 6.8\,\mathrm{MeV}$.

7.3 MC Studies

The method described above can be tested on an independent MC sample with a known value of $\boldsymbol{\mu}$. There are at least three reasons to test the fitting algorithm on Monte Carlo data. First, the MC is generated with known values of the parameters and can therefore be used to check for any biases in the method. Second, the minimum value of the likelihood function for a fixed number of events is itself a random variable and can be used as a judge of goodness-of-fit. The last reason is to study the statistical error on the extracted value of the parameters. This section will discuss all of these issues in the context of a study using five times the amount of data.

The likelihood function from one set of MC data can be sliced and plotted in either one or two of the three dimensions. Figures 7.2, 7.3, and 7.4 show the three 2–dimensional slices as contour plots. The first plot has values of α on the horizontal axis and κ on the vertical. The three contours mark the 1, 2, and $3\,\sigma$ levels appropriate for two free parameters. The crossing of the two straight lines indicates the generated point. The other two plots show the likelihood function in α–η and in κ–η, respectively.

There are two effects that can immediately be seen in the 2–D likelihood functions.

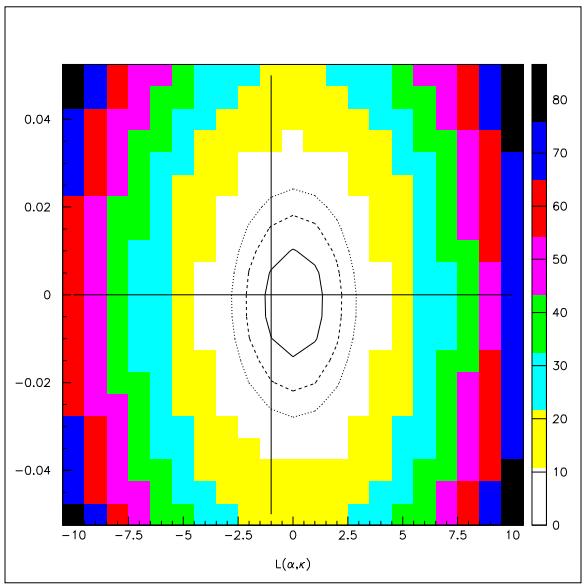

Figure 7.2: Likelihood function in two variables at the minimum of the third variable, using a Monte Carlo sample equivalent to one dataset. The two free variables are α–κ. The contours indicate the 1, 2, and 3 σ levels appropriate for two free parameters. The generated point (0,0) is given by the crossing of the two lines.

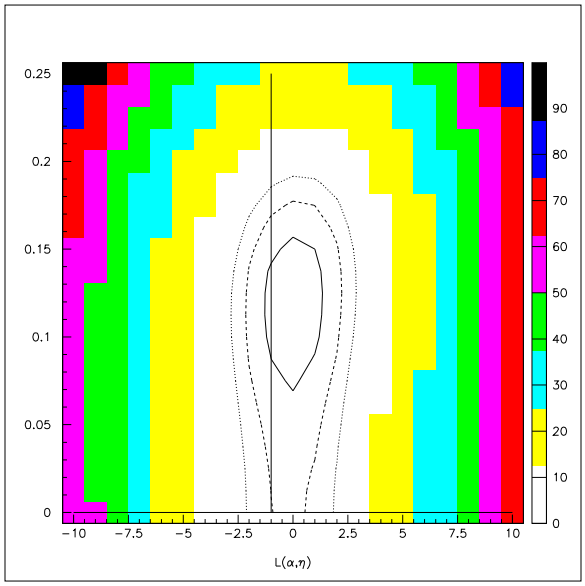

Figure 7.3: Likelihood function in two variables at the minimum of the third variable, using a Monte Carlo sample equivalent to one dataset. The two free variables are α–η. The contours indicate the 1, 2, and 3 σ levels appropriate for two free parameters. The generated point (0,0) is given by the crossing of the two lines.

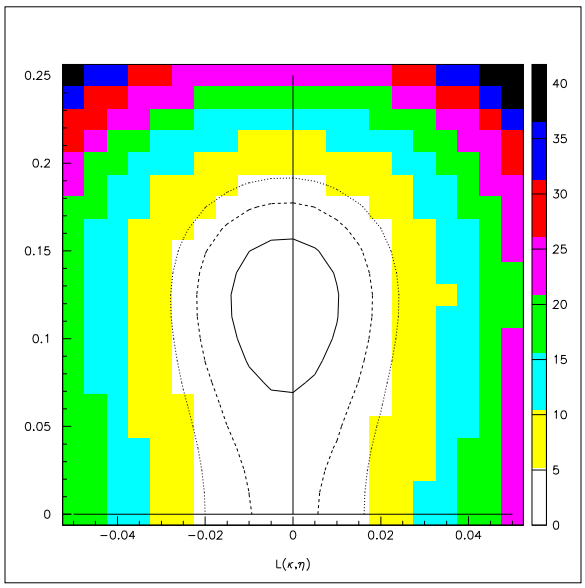

Figure 7.4: Likelihood function in two variables at the minimum of the third variable, using a Monte Carlo sample equivalent to one dataset. The two free variables are κ–η. The contours indicate the 1, 2, and 3 σ levels appropriate for two free parameters. The generated point (0,0) is given by the crossing of the two lines.

First, the parameters are appear uncorrelated. A correlation would be indicated by a tilt to the major axis of the error ellipses. The absence of such a tilt in all three slices implies that there are no correlations between the measured quantities. The second thing to notice is that the generated value of the α–κ pair falls within the 1 σ contour while the generated value of the κ–η pair does not. This bias is in the η direction and is more obvious in the 1–D slices shown in Figures 7.5, 7.6, and 7.7. These plots show slices of the likelihood function in α, κ, and η, respectively, with the other two variables fixed at their minimum values. The significance of the bias is indicated by the change in the likelihood function between its minimum and the value at the generated point. As shown in the 1-D slice, the change is approximately 2.5 which is more than 2 σ. This bias in η is attributed to the finite resolution on the angle ϕ between the two Dalitz pairs. This resolution bias is the topic of the next section.

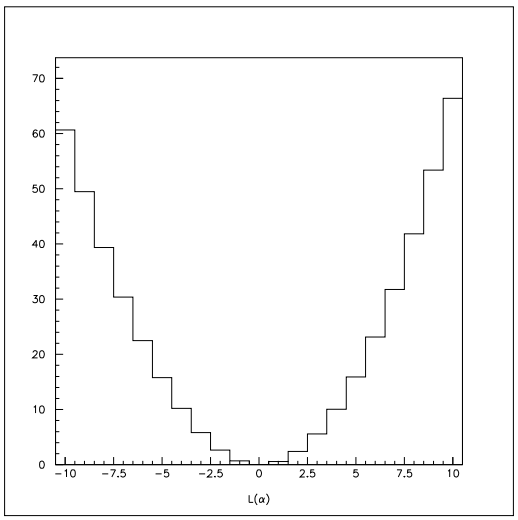

Figure 7.5: Likelihood function in one variable at the minimum of the other two variables, using a Monte Carlo sample equivalent to one dataset. The free variable is α and the generated point is $\alpha = 0$.

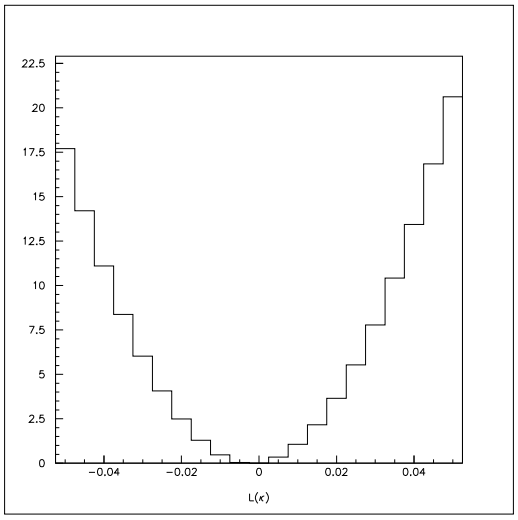

Figure 7.6: Likelihood function in one variable at the minimum of the other two variables, using a Monte Carlo sample equivalent to one dataset. The free variable is κ and the generated point is $\kappa = 0$.

7.3.1 Measurement Bias

The bias on η is created by the finite resolution on the angle ϕ. The differential width with respect to ϕ has the following form

$$\frac{d\Gamma}{d\phi} \sim 1 - A\cos 2\phi + B\sin 2\phi, \tag{7.28}$$

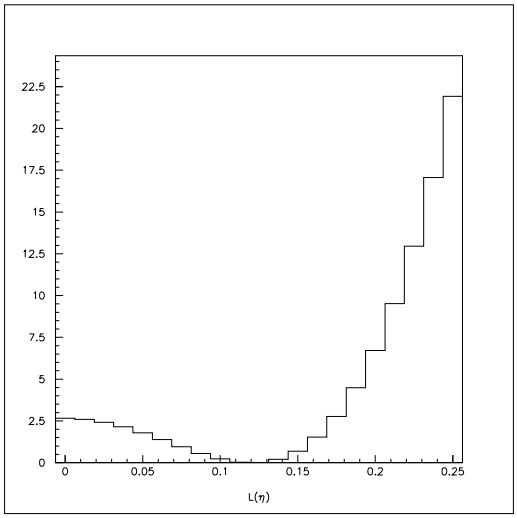

Figure 7.7: Likelihood function in one variable at the minimum of the other two variables, using a Monte Carlo sample equivalent to one dataset. The free variable is η and the generated point is $\eta = 0$.

where

$$A \sim -0.2 \cos 2\zeta, \tag{7.29a}$$

$$B \sim +0.2 \sin 2\zeta \cos \delta. \tag{7.29b}$$

The value of the κ–η pair extracted from the MC samples is consistent with $\kappa = 0$ and $\eta \neq 0$, or $\delta = \pi/2$ and $\zeta = \tan^{-1} \eta$. For small η this point is described by

$$A \sim -0.2(1 - 2\eta^2), \tag{7.30a}$$

$$B \sim 0. \tag{7.30b}$$

That is, the extracted point is consistent with a $\cos 2\phi$ term but with a reduced amplitude, which is exactly what one would expect from a random Gaussian smearing of such an oscillation. In other words, the resolution reduces the observed amplitude A and the fitter compensates by increasing η.

To remove this bias I will map out the function that relates the true value of η to the resolution–biased value extracted from the fit. In order to do this I have fit several

MC samples generated with different values of η. The extracted values, as a function of the generated values, are shown in Figure 7.8.

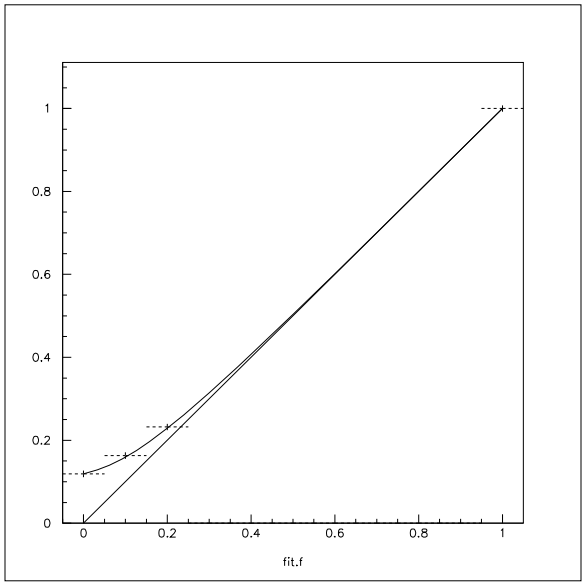

Figure 7.8: Map of extracted versus generated values of η.

The points are well described by the function

$$\eta_{\text{out}} = \eta_{\text{in}} + b\left(\frac{\mathrm{e}^{-m\eta_{\text{in}}} - \mathrm{e}^{-m}}{1 - \mathrm{e}^{-m}}\right), \tag{7.31}$$

with $b = 0.12$ and $m = 7.0$. While this function can not be inverted analytically, it is well behaved and can be inverted numerically quite easily. After the total uncertainty on the raw value of η has been computed I will transform the central value and error accordingly.

7.3.2 Goodness-of-Fit

The actual minimum value of the five likelihood values should be distributed according to a χ^2 distribution. As a test of goodness-of-fit, I can compare the minimum value of the data likelihood function to the distribution from the MC. The quantity that I use is the minimum value of the likelihood function per number of events fit. The values for the five MC samples are shown in Table 7.2.

Table 7.2: Values of minimum likelihood per number of events for the five MC samples.

MC Set	Win97	Sum97	99
0	−2.724	−2.757	−2.757
1	−2.723	−2.733	−2.708
2	−2.735	−2.732	−2.710
3	−2.729	−2.737	−2.703
4	−2.729	−2.726	−2.703

If the minimum value of the likelihood function from the data, divided by the number of data events, falls significantly outside of the range indicated in the table, it may be that the model is a poor description of the data for any choice of parameters. As will be shown in the next section, this is not the case and the model describes the data very nicely.

7.4 Extracted Values

The fit to the data was done using the same cuts described in the previous chapter for selecting $\pi^0 \to e^+e^-e^+e^-$ events. As discussed above, the value of the likelihood function at its minimum is a test of the agreement between the data and the model. The values of the minimum for the three datasets are shown in Table 7.3 and they agree quite nicely with the values from the MC, shown in Table 7.2. As an additional check on the goodness-of-fit, I can compare the distributions of the two main phase space variables between the data and the MC, shown in Figure 7.9. It is seen that the MC does a very adequate job of predicting the shape of these important quantities.

The slices of the likelihood function derived from the combined dataset are shown

Table 7.3: The number of events used in the fit along with the value of the likelihood function at its minimum divided by the number of events.

	Win97	Sum97	99
Events	5428	4152	20928
Minimum	−2.727	−2.731	−2.713

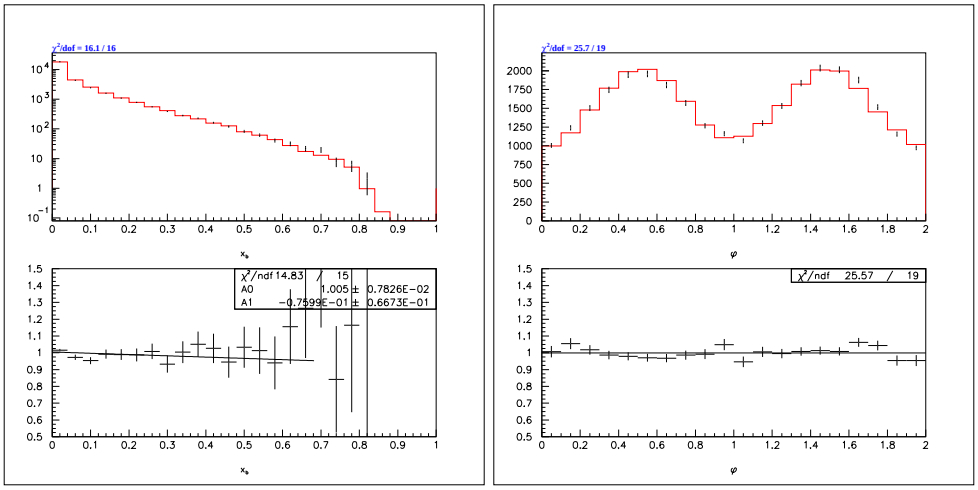

Figure 7.9: Comparison of the distributions of x_b and ϕ between the data (points) and 5 MC samples (black histogram). The ratios are data over MC.

in Figures 7.10 to 7.15. The general features are very similar to those seen in the MC distributions earlier. The apparent secondary minimum in the α–η plane has been studied and found to be an artifact of the plotting software.

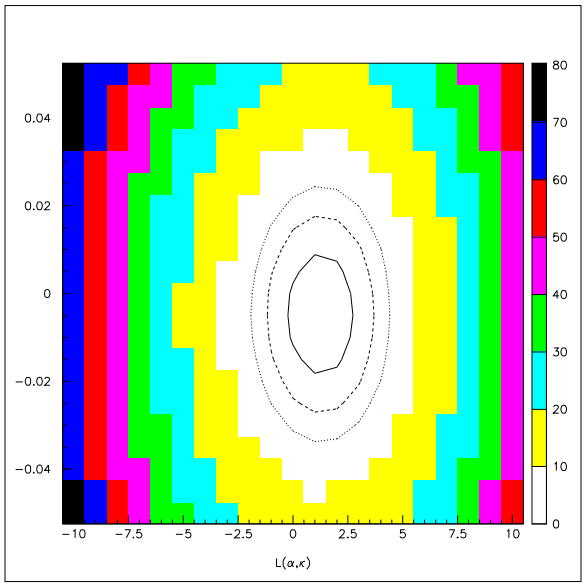

Figure 7.10: Likelihood function in two variables at the minimum of the third variable, using the complete KTeV dataset. The two free variables are α–κ. The contours indicate the 1, 2, and 3 σ levels appropriate for two free parameters.

Table 7.4 gives the raw values along with statistical errors for the three datasets

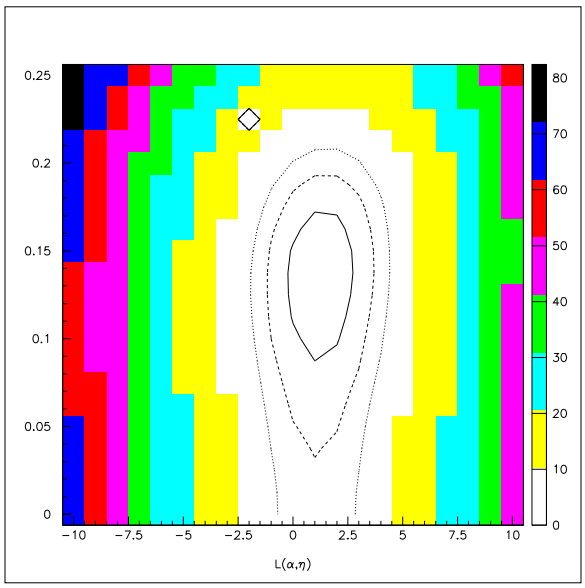

Figure 7.11: Likelihood function in two variables at the minimum of the third variable, using the complete KTeV dataset. The two free variables are α–η. The contours indicate the 1, 2, and 3 σ levels appropriate for two free parameters.

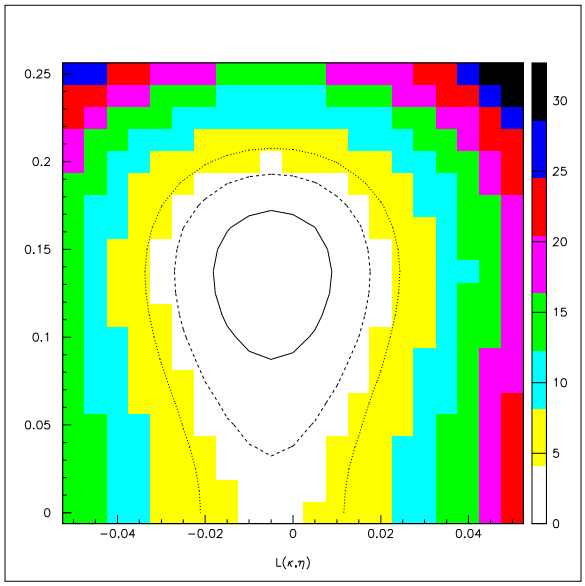

Figure 7.12: Likelihood function in two variables at the minimum of the third variable, using the complete KTeV dataset. The two free variables are κ–η. The contours indicate the 1, 2, and 3 σ levels appropriate for two free parameters.

and the weighted average. The different datasets are statistically consistent. The small χ^2 in the η average indicates that the resolution bias is a bigger effect than statistical

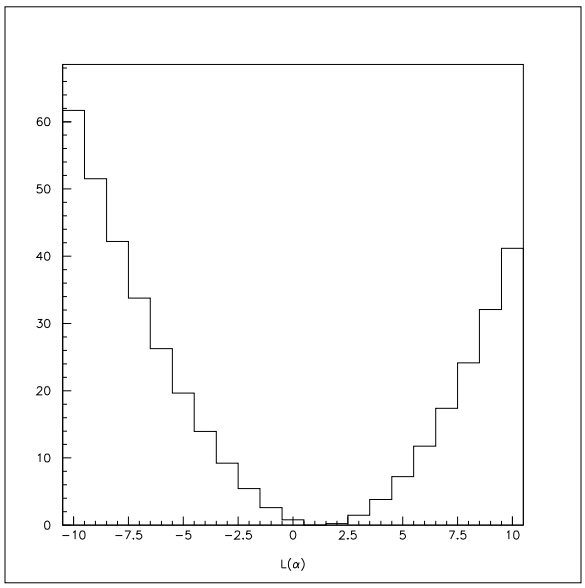

Figure 7.13: Likelihood function in one variable at the minimum of the other two variables, using the complete KTeV dataset. The free variable is α.

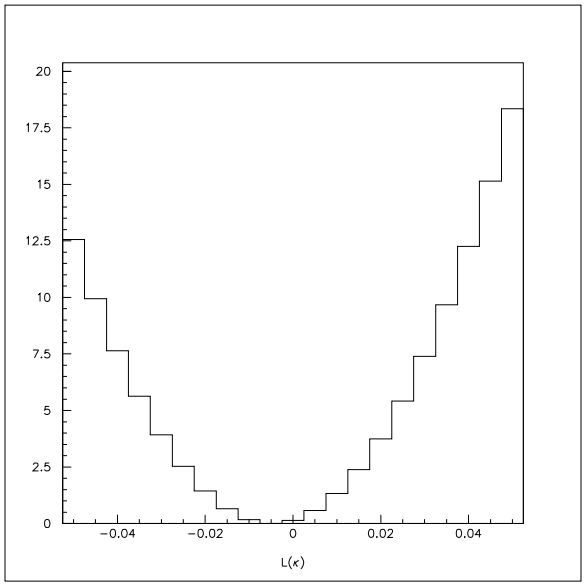

Figure 7.14: Likelihood function in one variable at the minimum of the other two variables, using the complete KTeV dataset. The free variable is κ.

fluctuations.

The next section will discuss the various systematic studies performed to access additional sources of uncertainty.

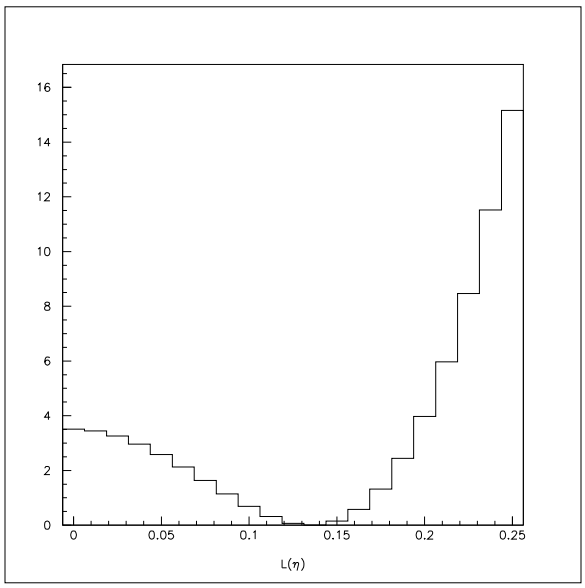

Figure 7.15: Likelihood function in one variable at the minimum of the other two variables, using the complete KTeV dataset. The free variable is η.

7.5 Systematic Studies

Many of the systematics are identical to those discussed in the previous chapter about the branching ratio. There are two unique studies discussed below, one dealing with the normalization calculation and the other dealing with the bias mentioned above.

7.5.1 Normalization

The precision of the normalization calculation increases roughly as the square root of the number of MC events used in the calculation. The main analysis uses the equivalent of five datasets to estimate the observed partial width. In order to estimate the uncertainty in this calculation, I perform the fit five times, each time using only one of the five datasets in the normalization. I then take the largest difference between the main result and any of the five fits, divided by $\sqrt{5}$, as the remaining uncertainty. Table 7.5 summarizes the error on each parameter for each dataset. The combined error on each parameter is the weighted average of the three datasets based on the number

Table 7.4: Results of the coupling parameters. The values of η are uncorrected. The errors are statistical.

Dataset	α	κ	η
Win97	-0.2 ± 2.3	-0.020 ± 0.020	0.113 ± 0.073
Sum97	-1.8 ± 2.6	-0.034 ± 0.024	0.151 ± 0.063
99	$+2.3 \pm 1.2$	-0.003 ± 0.011	0.135 ± 0.031
Combined	$+1.3 \pm 1.0$	-0.008 ± 0.009	0.135 ± 0.026
χ^2/dof	1.2	1.4	0.1

of events in the data. In each case, this MC uncertainty is less than half the statistical uncertainty.

7.5.2 Resolution Bias Correction

The bias on η due to the finite resolution on the angle ϕ is well understood. The form of the correction function given above has the correct general form but is not precisely known. To be conservative I will assign the total size of the correction as a systematic uncertainty on η. Table 7.6 lists the systematic error in each of the three run periods as well as the weighted average.

7.5.3 Cut Variation

General disagreements between the data and the Monte Carlo are assigned errors based on how the fit parameters change as the various selection cuts are varied about their nominal values. The method applied here is to change one cut to be either slightly

Table 7.5: Systematic uncertainties due to finite MC statistics in the normalization calculation.

Dataset	δ_α	δ_κ	δ_η
Win97	0.48	0.0052	0.0031
Sum97	0.55	0.0044	0.0110
99	0.22	0.0048	0.0057
Combined	0.31	0.0048	0.0060

Table 7.6: Systematic error on η due to the resolution bias.

Dataset	δ_η
Win97	0.113
Sum97	0.066
99	0.084
Combined	0.087

looser or slightly tighter than the nominal value and perform the fit. The new value is then compared to the central value using only the uncorrelated part of the errors. This is approximated by subtracting the two errors in quadrature. Any cut whose new value differs from the central value by more than one uncorrelated standard deviation is assigned a systematic uncertainty. The final error is then the sum in quadrature of all individual cut uncertainties. Table 7.7 summarizes the results of varying the 13 cuts described in the previous chapter.

This method is perhaps overly conservative since some double counting may occur when different cuts are correlated. It is also possible that some of the observed changes are actually statistical in nature and are not true systematic biases. In the end it is hoped that this method covers systematic effects such as differences in resolution and reconstruction efficiencies that are not addressed elsewhere and as such I err on the conservative side.

7.5.4 Material

There are several effects that might directly impact the shape of the different phase space distributions that feed into the fit. Among them the most important is the impact on the reconstruction of the opening angle due to multiple scattering. The uncertainty in the amount of multiple scattering is directly related to the uncertainty in the amount of material in the detector. In order to study the impact on the results of using the wrong amount of material in the Monte Carlo calculation of the normal-

Table 7.7: Summary of cut variation errors on the extracted parameters

	α			κ			η		
Cut	Win97	Sum97	99	Win97	Sum97	99	Win97	Sum97	99
1	0	0.442	0.109	0	0.0061	0.0064	0	0	0.0057
2	0	0	0	0.0001	0.0029	0	0	0	0
3	0	0.619	0	0	0	0.0026	0	0	0
4	0	0	0	0.0026	0.0083	0.0072	0	0	0
5	0	0.868	0	0	0.0056	0	0	0	0
6	0	0	0	0.0013	0	0	0	0	0
7	0	0	0.765	0	0.0128	0.0003	0	0	0
8	0	0.033	0.028	0.0015	0	0	0	0	0
9	0	0	0.026	0	0.0044	0.0014	0	0	0
10	0.748	0	0.081	0	0	0.0010	0	0	0
11	0.346	0.789	0	0	0.0025	0.0013	0	0	0
12	0	0	0.030	0.0008	0.0014	0.0019	0	0	0
13	0.084	0.231	0	0.0012	0	0	0	0	0
All	0.828	1.418	0.778	0.0036	0.0184	0.0104	0	0	0.0057
δ		0.874			0.0103			0.0057	

Table 7.8: Number of radiation lengths used to study systematic uncertainties on the parameters.

Set	Win97	Sum97	99
$-7\,\sigma$	0.00076	0.00082	0.00046
Nominal	0.00104	0.00110	0.00138
$+7\,\sigma$	0.00132	0.00138	0.00102

ization factor, I have generated two additional sets of MC events which are identical to the nominal set except for the amount of material in Helium bag 1A. The amount of air in this bag has been discussed previously in connection to the branching ratio measurement. For this study I have varied the amount of material by 7 standard deviations in both directions. Table 7.8 shows the number of radiation lengths used in the study.

The normalization was recalculated with the modified material and then used to fit the data. The impact is generally small even at $\pm 7\sigma$. To estimate a realistic uncertainty I measure the trend and then scale back to $\pm 1\sigma$. The resulting uncertainties are negligible compared to other sources of error and will not be considered further.

7.5.5 Chamber Inefficiency

The second effect that could directly impact the reconstruction is tracking inefficiencies due to high SODs and missing hits. It is known that the tracking efficiency had both a position and a time dependence due to radiation damage in the beam regions of the drift chambers. This effect is modeled in the MC by applying 2-dimensional efficiency maps measured from the data. The overall efficiency is found by comparing chamber illuminations between the data and the MC. In order to estimate the effect of this overall scale being wrong in the MC, I have generated MC events with either a much higher or a much lower scale factor. Using $K_L \to \pi^0\pi^0\pi^0_D$ events, it was found that scaling by roughly 40% in either direction created large discrepancies in the illuminations. I therefore used MC events with scale factors of 0.60, 1.00, and 1.40 to calculate

Table 7.9: Summary of Systematic Uncertainties.

Source	α	κ	η
Norm	0.310	0.0048	0.0060
ϕ Bias			0.0870
Cuts	0.874	0.0103	0.0057
Total	0.927	0.0114	0.0874

the normalization factor. Even with these unrealistic values, no significant difference in the extracted fit parameters was found.

7.5.6 Summary of Systematic Uncertainties

Other sources of systematic effects were studied and found to be insignificant. These include effects like binning and fitting of the likelihood functions and changing the physical photon energy cutoff used in the fit. Table 7.9 summarizes the important systematic uncertainties on the fit parameters.

The total systematic uncertainties are similar to the statistical errors with the exception of the correction to the η parameter.

7.6 Results

The combined results of the fit, correcting for the bias in η, are shown in Table 7.10. The first error is statistical while the second is systematic. It should be remembered that η is restricted to positive values and, since the result is consistent with zero, an upper limit is the appropriate measurement to report.

Table 7.10: Combined results of the three fit parameters including statistical and systematic errors.

Parameter	Value
α	$1.3 \pm 1.0 \pm 0.9$
κ	$-0.008 \pm 0.009 \pm 0.011$
η	$0.051 \pm 0.026 \pm 0.087$

Since η must be positive, the prescription of Feldman and Cousins [31] can be used to set an upper limit at the 90% confidence level. This results in a the following range

$$\eta < 0.20 \quad \text{at } 90\% \text{ C.L.} \tag{7.32}$$

Using the upper limit on η and the central value of $\kappa = -0.008$, the upper limit on ζ, allowing for CPT violation, is estimated to be

$$\zeta < 11.3^\circ \quad \text{allowing } CPT \text{ violation.} \tag{7.33}$$

To impose CPT conservation I use the 90 % upper limit on $|\kappa|$, calculated with the same prescription as the limit on η,

$$|\kappa| < 0.031 \quad \text{at } 90\% \text{ C.L.} \tag{7.34}$$

The corresponding limit on the mixing angle is

$$\zeta < 1.72^\circ \quad \text{requiring } CPT \text{ conservation.} \tag{7.35}$$

The possibility of large phase differences cannot be excluded. At the upper limit on η where the mixing is purely imaginary, the fraction of the CP violating to CP conserving contributions could be as large as 20 %. The same limit, but with the CPT constraint, becomes 3.1 %.

Chapter 8

Radiative Double Dalitz Branching Ratio

The radiative double Dalitz decay cannot be detected when the radiated photon energy is less than about 2 GeV in the lab frame. In the pion CM frame, a photon with an energy of 5 MeV and traveling in the same direction as the pion will be boosted to an energy of just 2 GeV in the lab frame. Photons with energies much below 5 MeV in the CM frame will never be detectable in the KTeV calorimeter. This detector threshold imposes a cutoff in the observable branching ratio for any radiative decay. The cutoff that will be used is defined in terms of $x_{4e} = M_{4e}^2/M^2$ which will be required to be less than 0.9. In terms of energy, the photon energy must be greater than 6.75 MeV in the pion CM. The equation that relates these quantities is

$$E_\gamma^* = \frac{M}{2}\left(1 - x_{4e}\right). \tag{8.1}$$

The maximum value of E_γ^* is $0.99995 \times M/2$.

The method of measuring the branching ratio is nearly identical to the method already described for the non–radiative decay. The main difference is that the normalization mode, in this case $K_L \to \pi^0\pi^0\pi^0_{DD}$, is also the largest background. The fundamental ratio in this analysis is

$$R = \frac{N_{4e\gamma}^{\rm obs}}{N_{4e}^{\rm obs}}\frac{\epsilon_{4e}}{\epsilon_{4e\gamma}} = \frac{B(\pi^0 \to e^+e^-e^+e^-\gamma, x_{4e} < 0.9)}{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9)}. \tag{8.2}$$

The $\pi^0 \to e^+e^-e^+e^-$ numbers will simply be taken from the results of Chapter 6.

The possible backgrounds to the $\pi^0 \to e^+e^-e^+e^-\gamma$ signal include radiative double-single Dalitz, single Dalitz with a conversion plus radiation, and double conversions with radiation. The largest background comes from $\pi^0 \to e^+e^-e^+e^-$ with radiation in material or with an accidental photon. It is possible to separate the internal from the external radiation because of differences in the angle at which the photon is produced at. External radiation tends to be emitted at much smaller angles, relative to the electron, than internal radiation does.

The first part of the chapter will highlight the aspects of the analysis that differ from the 4e8c reconstruction. I will then discuss the Monte Carlo predictions of the background levels and the amount of external radiation. The rest of the chapter describes the internal-external radiation separation and the branching ratio result.

8.1 Event Selection

The selection is identical to the $\pi^0 \to e^+e^-e^+e^-$ branching ratio analysis except that 9 clusters are required. The same cuts, listed in Table 6.3, are applied to the 4e9c split sample. An event χ^2 is formed to distinguish between the signal and radiative double single-Dalitz events, which has the identical final state. After all of the $\pi^0 \to e^+e^-e^+e^-$ cuts have been applied, there remains 1129 events in the combined dataset.

8.1.1 External Radiation

Radiation downstream of the magnet is not detectable since the photon and electron almost always land on top of each other in the calorimeter. The key to this analysis is therefore the ability to distinguish between internal and external radiation.

The probability of a 5 GeV electron radiating a photon with energy greater than 2 GeV is given approximately by (see Equation (4.6)), $P(E_\gamma > 2\,\mathrm{GeV}) \approx 0.84X/X_0$, where X/X_0 is the amount of material in radiation lengths. This is remarkably similar to the conversion probability for photons which goes like $(7/9)X/X_0$. The number of

radiation lengths upstream of the magnet is roughly 0.007, making the probability of radiation about 0.6 %. The probability of 1 out of 4 electrons radiating is then 2.4 %. Given the 30000 $\pi^0 \to e^+e^-e^+e^-$ events found in the 4e8c sample, one can predict that roughly 700 $\pi^0 \to e^+e^-e^+e^-$ events will be found in the 4e9c data due to hard external radiation upstream of the magnet.

One additional cut is then necessary to distinguish between the signal mode which has internal radiation and the normalization mode with external radiation. The most effective variable to perform this separation is the smallest angle between the photon and any of the electrons. The mean angle for internal radiation is 30 mrad whereas the typical angle for external radiation is only 3 mrad. Figure 8.1 shows the distribution of the logarithm of $\theta_{\min}$. The green histogram contains Monte Carlo events with internal radiation while the blue one is made from MC events with external radiation. The red histogram is the sum of the two MC samples, which may be compared to the data. While the χ^2 between the data and the sum of the two Monte Carlo samples is fine at 23 for 25 degrees of freedom, there is an obvious disagreement in the region dominated by external radiation. The interpretation of this difference in the spectrum will be discussed below in the context of backgrounds.

In order to avoid the poorly understood region of $\theta_{\min}$, a cut is made requiring that $\log\theta_{\min} > -1.2$. Figure 8.2 shows the same distribution of $\theta_{\min}$ but with the cut imposed. The number of events remaining in the data after this cut is 425. The Monte Carlo acceptance for events with $x_{4e} < 0.9$ is 5.85×10^{-4}, while the acceptance for events with $x_{4e} > 0.9$ is only 8.07×10^{-6}.

8.1.2 Backgrounds

The backgrounds that were studied included all of the same $\pi^0 \to e^+e^-e^+e^-$ backgrounds, double-single Dalitz and single and double conversions, plus $\pi^0 \to e^+e^-e^+e^-$ events with external radiation or accidental assistance. The single Dalitz and conver-

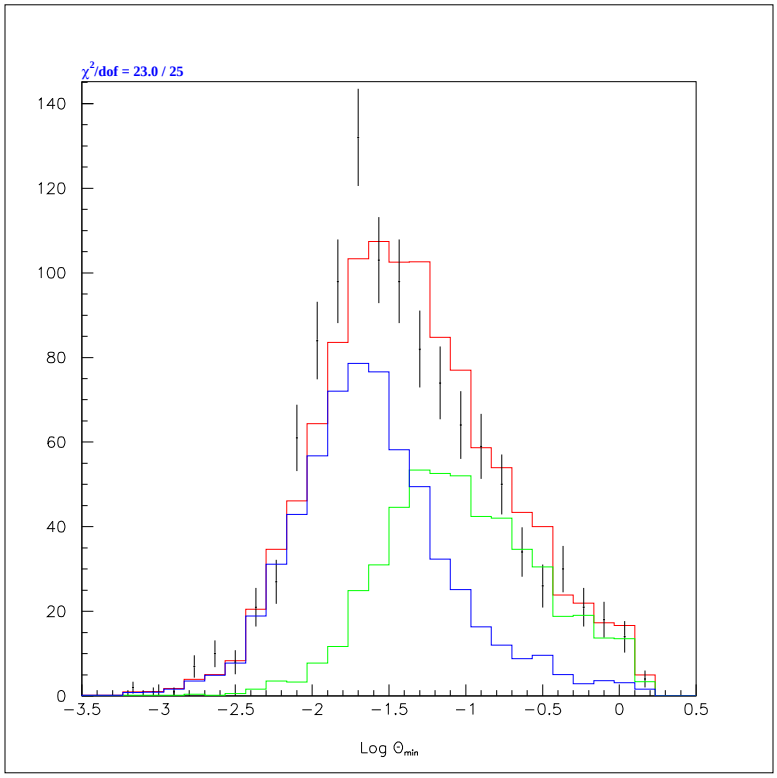

Figure 8.1: Distribution of $\log\theta_{\rm min}$. The points with error bars are the data, the green histogram is the internal radiation MC, the blue histogram is the external radiation MC, and the red histogram is the sum of the MC.

sion backgrounds are found to be negligible. The external radiation, on the other hand, dominates the sample of candidate $\pi^0 \to e^+e^-e^+e^-\gamma$ events before the cut on $\theta_{\rm min}$. Not only that, but Figure 8.1 indicates that either the external radiation is mis-modeled or there is another source of background in the data that is not simulated in the Monte Carlo.

To evaluate the impact of the discrepancy in $\theta_{\rm min}$ on the predicted background level, the amount of background is allowed to float in such a way as to minimize the differences between the data and MC. The combination that fits the data best requires that the external component of the MC be increased by 41 %. Certainly this is above and beyond any uncertainties in either the $\pi^0 \to e^+e^-e^+e^-$ rate or the probability of external radiation. It is more likely that the acceptance for external radiation is only

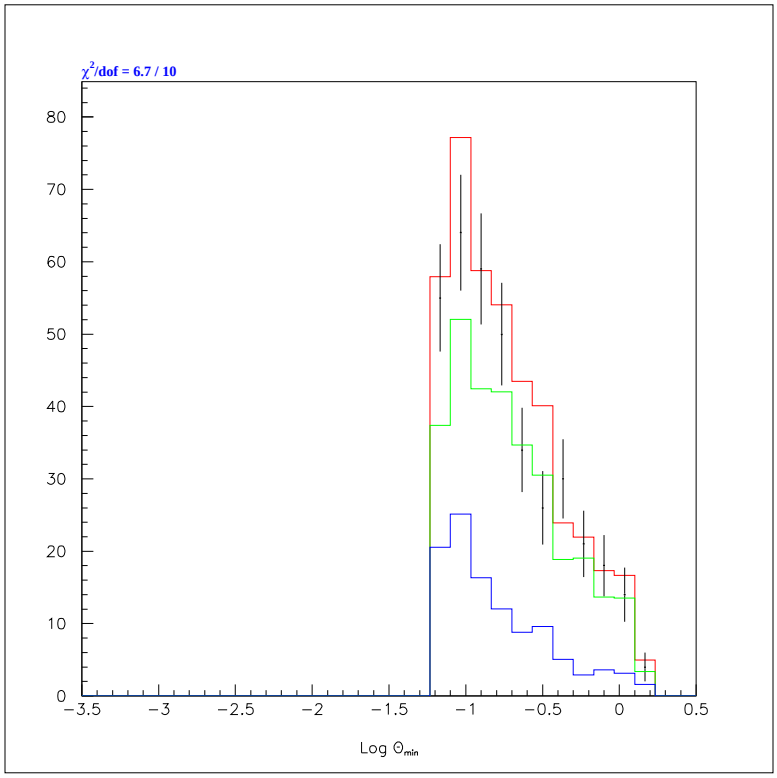

Figure 8.2: Distribution of $\log\theta_{\min}$. The points with error bars are the data, the green histogram is the internal radiation MC, the blue histogram is the external radiation MC, and the red histogram is the sum of the MC.

trustworthy at the 40 % level. This uncertainty in the background level translates into an uncertainty in the branching ratio of 15 %.

8.2 Branching Ratio

The number of observed events is assumed to depend on only two sources,

$$N_{4e\gamma} = F \cdot 3 \cdot B(K_L \to \pi^0\pi^0\pi^0) \cdot B^2(\pi^0 \to \gamma\gamma) \cdot \left[B_{4e\gamma}\epsilon_{4e\gamma} + B_{4e}\epsilon_{4e}\right], \qquad (8.3)$$

where $B_{4e\gamma}$ is defined for $x_{4e} < 0.9$ and B_{4e} is defined for $x_{4e} > 0.9$. The flux is taken from the $\pi^0 \to e^+e^-e^+e^-$ analysis, where it was found to be 5.62×10^{11}. The branching ratio is therefore given by

$$\frac{B_{4e\gamma}}{B_{4e}} = \frac{1}{\epsilon_{4e\gamma}}\left(\frac{N_{4e\gamma}}{C} - \epsilon_{4e}\right), \qquad (8.4)$$

where C is the combination of the flux and the known branching ratios, $C = 1.13 \times 10^7$.

Using the numbers quoted above, the branching ratio is found to be

$$\frac{B(\pi^0 \to e^+e^-e^+e^-\gamma, x_{4e} < 0.9)}{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9)} = 0.0504 \pm 0.0028(\text{stat}) \pm 0.0076(\text{syst}), \qquad (8.5)$$

where the systematic error is from the background estimate discussed previously. The predicted division of the data between internal and external radiation is 317 versus 108.

This mode has never been observed so there are no other measurements to compare with. The predicted branching ratio is 0.0617. This is about $1.4\,\sigma$ higher than the measured value. Given the current level of understanding this is completely acceptable.

The branching ratio with respect to the $\pi^0 \to \gamma\gamma$ decay can be found with the aid of the measured $\pi^0 \to e^+e^-e^+e^-$ branching ratio. It is found to be

$$\frac{B(\pi^0 \to e^+e^-e^+e^-\gamma, x_{4e} < 0.9)}{B(\pi^0 \to \gamma\gamma)} = (1.68 \pm 0.09(\text{stat}) \pm 0.25(\text{syst})) \times 10^{-6}. \qquad (8.6)$$

Chapter 9

Conclusions

The analysis of the combined KTeV dataset resulted in 30,511 $\pi^0 \to e^+e^-e^+e^-$ decays, nearly 150 times the 1962 experimental sample. These events were used to extract several branching ratios:

$$\frac{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9)}{B^2(\pi^0 \to e^+e^-\gamma)} = 0.2271 \pm 0.0014(\text{stat}) \pm 0.0021(\text{syst}), \tag{9.1a}$$

$$\frac{B(\pi^0 \to e^+e^-e^+e^-, x_{4e} > 0.9)}{B(\pi^0 \to \gamma\gamma)} = (3.299 \pm 0.047(\text{int}) \pm 0.179(\text{ext})) \times 10^{-5}, \tag{9.1b}$$

$$B(\pi^0 \to e^+e^-e^+e^-(\gamma)) = (3.46 \pm 0.04(\text{int}) \pm 0.19(\text{ext})) \times 10^{-5}. \tag{9.1c}$$

The last line gives the partial width of the double Dalitz decay, extrapolated to all values of x_{4e}. These results are in good agreement with the previous experimental work of Samios [12] and our own theoretical predictions [15]. While the method does introduce a large uncertainty due to the single Dalitz branching ratio, it is hoped that KTeV will eventually measure this branching ratio. While the ultimate uncertainty will be limited by systematic effects, it could likely as low as 1 %. This would improve the total error on this measurement by a factor of 3, reducing it to 1.7 %.

The same events were used to extract the momentum dependence of the π^0 form factor. The parameter α is found to be

$$\alpha = 1.3 \pm 1.3. \tag{9.2}$$

In terms of the standard slope parameter this is $a = -0.040 \pm 0.040$. This is the first

measurement of this quantity in the double Dalitz mode. While this result has the opposite sign as the PDG average, it is still within $2\,\sigma$.

Additionally, the size of any *CP*-violating contributions to the decay $\pi^0 \to \gamma^*\gamma^*$ has been limited. The most likely point in the $\kappa - \eta$ plane corresponds to $\zeta = 3.0^\circ$ and $\delta = 99^\circ$. These values imply large *CPT* and *CP* violation. However, since $1\,\sigma$ contour covers the origin in $\kappa - \eta$, the phase difference δ cannot be limited at the 68 % C.L.. If *CPT* violation is allowed the 90 % upper limit on the mixing angle is

$$\zeta < 11.3^\circ \quad \text{allowing } \mathit{CPT} \text{ violation.} \tag{9.3}$$

If however, *CPT* conservation is enforced, the limit is reduced to

$$\zeta < 1.72^\circ \quad \text{requiring } \mathit{CPT} \text{ conservation.} \tag{9.4}$$

Finally, the decay $\pi^0 \to e^+e^-e^+e^-\gamma$ has been observed and a preliminary branching ratio has been calculated. The result, based on 317 background-subtracted events, is

$$\frac{B(\pi^0 \to e^+e^-e^+e^-\gamma, x_{4e} < 0.9)}{B(\pi^0 \to \gamma\gamma)} = (1.68 \pm 0.27) \times 10^{-6}. \tag{9.5}$$

The large error is due to uncertainties in estimating the contamination from the external radiation background. At this level, the result is in good agreement with the theoretical prediction and confirms the validity of the extrapolation used above to quote the combined partial width.

Bibliography

[1] H. Yukawa. On the interaction of elementary particles. Proc. Phys. Math. Soc. Jap., 17:48–57, 1935.

[2] D. H. Perkins. Nuclear disintegration by meson capture. Nature, 159:126–127, 1947.

[3] J. Steinberger, W. K. H. Panofsky, and J. Steller. Evidence for the production of neutral mesons by photons. Phys. Rev., 78:802–805, 1950.

[4] C. N. Yang. Selection rules for the dematerialization of a particle into two photons. Phys. Rev., 77(2):242–245, 1950.

[5] J. Bernstein and L. Michel. T, P, C symmetries in the π^0 decay. Phys. Rev., 118(3):871–875, 1950.

[6] R. H. Dalitz. On an alternative decay process for the neutral π-meson. Proc. Phys. Soc. (London), A64:667–669, 1951.

[7] J. J. Lord, J. Fainberg, D. M. Haskin, and M. Schein. Narrow angle pairs of particles from nuclear interactions. Phys. Rev., 87:538–539, 1952.

[8] P. Lindenfeld, A. Sachs, and J. Steinberger. The internal pair production of γ-rays of mesonic origin; alternate modes of π^0 decay. Phys. Rev., 89:531–537, 1953.

[9] C. P. Sargent, R. Cornelius, M. Rinehart, L. M. Lederman, and K. Rogers. Diffusion cloud-chamber study of very slow mesons. Phys. Rev., 98:1349–1354, 1955.

[10] N. P. Samios. Dynamics of internally converted electron-positron pairs. Phys. Rev., 121:275–281, 1961.

[11] N. M. Kroll and W. Wada. Internal pair production associated with the emission of high-energy gamma rays. Phys. Rev., 98(5):1355–1359, 1955.

[12] N. P. Samios, R. Plano, A. Prodell, M. Schwartz, and J. Steinberger. Parity of the neutral pion and the decay $\pi^0 \to e^+e^-e^+e^-$. Phys. Rev., 126(5):1844–1849, 1962.

[13] K. Hagiwara et al. Review of particle physics. Phys. Rev., D66:010001, 2002.

[14] T. Miyazaki and E. Takasugi. Internal conversion of pseudoscalar mesons into lepton pairs. Phys. Rev., D8(7):2051–2062, 1973.

[15] A. R. Barker, H. Huang, P. A. Toale, and J. Engle. Radiative corrections to double dalitz decays: Effects on invariant mass distributions and angular correlations. Phys. Rev., D67:033008, 2003.

[16] J. D. Jackson. Classical Electrodynamics. Wiley, 2nd edition, 1999.

[17] G. D'Ambrosio, G. Isidori, and J. Portolés. Can we extract short-distance information from $B(K_L \to \mu^+\mu^-)$? Phys. Lett., B423:385–394, 1998.

[18] F. Farzanpay et al. Measurement of the slope of the π^0 electromagnetic form factor. Phys. Lett., B278:413–418, 1992.

[19] Reena Meijer Drees et al. Measurement of the π^0 electromagnetic transition form factor. Phys. Rev., D45:1439–1447, 1992.

[20] H. J. Behrend et al. A measurement of the π^0, η and η' electromagnetic formfactors. Z. Phys., C49:401–410, 1991.

[21] H. Fonvieille et al. Dalitz decay $\pi^0 \to \gamma e^+ e^-$ and the π^0 electromagnetic transistion form factor. Phys. Lett., B233:65–68, 1989.

[22] K. O. Mikaelian and J. Smith. Radiative corrections to the decay $\pi^0 \to \gamma e^+ e^-$. Phys. Rev., D5:1763–1773, 1972.

[23] Itzykson and Zuber. Quantum Field Theory. McGraw-Hill, 1980.

[24] Peter S. Shawhan. Observation of Direct CP Violation in $K_{S,L} \to \pi\pi$ Decays. PhD thesis, The University of Chicago, 1999.

[25] Eric Zimmerman. A Measurement of the Branching Ratio of $\pi^0 \to e^+e^-$ Using $K_L \to 3\pi^0$ Decays in Flight. PhD thesis, The University of Chicago, 1999.

[26] A. Roodman. The KTeV pure CsI calorimeter. In Proceedings of the VII International Conference on Calorimetry. World Scientific, 1998.

[27] Peter Mikelsons. Search for the Decay $K_L \to \pi^0 e^+ e^-$ and Study of the Decay $K_L \to e^+e^-\gamma\gamma$. PhD thesis, The University of Colorado, 1999.

[28] A. J. Malensek. Empirical formula for thick target particle production. Technical Report FN-341, Fermilab, 1981.

[29] Jason LaDue. Understanding Dalitz Decays of the K_L, in particular the decays of $K_L \to e^+e^-\gamma$ and $K_L \to e^+e^-e^+e^-$. PhD thesis, The University of Colorado, 2003.

[30] A. F. Bielajew. Improved angular sampling for pair production in the EGS4 code system. Technical Report PIRS-0287, National Research Council of Canada, 1991.

[31] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. Phys. Rev., D57:3873–3889, 1998.

Appendix A

Phase Space Kinematics

This appendix explores the phase space parameterization for 2–body, 3–body, and 4–body final states. The invariant phase space differential element for a n–body decay is defined as

$$d^D\Phi_n = \left[\prod_{i=1}^{n} \frac{1}{(2\pi)^3}\frac{d^3k_i}{2E_i}\right](2\pi)^4\delta^4(P - \sum_{i=1}^{n} p_i), \tag{A.1}$$

where k_i and E_i are the 3–momentum and energy of final state particle i and P is the parent particle 4–momentum. The dimension of the phase space, D, is just three times the number of final state particles minus four due to the momentum conserving delta function, that is $D = 3n - 4$.

A.1 2–Body Phase Space

For a 2–body decay, Equation (A.1) takes the following form,

$$d^2\Phi_2 = \frac{1}{(2\pi)^2}\frac{d^3k_1}{2E_1}\frac{d^3k_2}{2E_2}\delta^4(P - p_1 - p_2). \tag{A.2}$$

Before simplifying this expression, it will be helpful to work out some notation by considering the decay of a meson of mass M into two massive photons. To simplify matters, I will work in the CM frame and also will pick the $+\hat{\boldsymbol{z}}$–direction to coincide

with the direction of the first photon. In this frame, the three 4–momenta are

$$P^\nu = (M, 0, 0, 0), \tag{A.3a}$$

$$p_1^\nu = (E_1, 0, 0, +k), \tag{A.3b}$$

$$p_2^\nu = (E_2, 0, 0, -k). \tag{A.3c}$$

The mass of each photon can be parameterized by the dimensionless quantity $x_i = p_i^2/P^2$. It is then possible to solve for E_1, E_2, and k in terms of the quantities x_1 and x_2,

$$E_1 = \frac{1}{2}M(1 + x_1 - x_2), \tag{A.4a}$$

$$E_2 = \frac{1}{2}M(1 - x_2 + x_2), \tag{A.4b}$$

$$k = \frac{1}{2}M\sqrt{1 - 2(x_1 + x_2) + (x_1 - x_2)^2}. \tag{A.4c}$$

If both photons are real ($p_i^2 = 0$) then we recover the familiar situation $E_1 = E_2 = k = M/2$.

In order to generalize to larger final states it will be useful to define some general quantities. Let $p_{ij} = p_i + p_j$, then

$$z_{ij} = \frac{2(p_i \cdot p_j)}{p_{ij}^2} = 1 - x_i - x_j, \tag{A.5}$$

where $x_i = p_i^2/p_{ij}^2$ and $x_j = p_j^2/p_{ij}^2$ as above. I will also define

$$\delta_{ij} = x_i - x_j, \tag{A.6a}$$

$$w_{ij} = 2\sqrt{x_i x_j}, \tag{A.6b}$$

$$\lambda_{ij} = \sqrt{z_{ij}^2 - w_{ij}^2}. \tag{A.6c}$$

This last quantity can be shown to be proportional to the magnitude of the 3–momentum of either photon in the meson rest frame. So the energies and momentum of the two

photons in the CM frame can be simply expressed as

$$E_1 = \frac{1}{2} M_{12}(1 + \delta_{12}), \tag{A.7a}$$

$$E_2 = \frac{1}{2} M_{12}(1 - \delta_{12}), \tag{A.7b}$$

$$k = \frac{1}{2} M_{12}\lambda_{12}, \tag{A.7c}$$

where $M_{12}^2 = p_{12}^2$.

Returning to the phase space element, the delta function can be easily integrated over to yield

$$d^2\Phi_2 = \frac{1}{2^5\pi^2}\lambda_{12}\; d^2\Omega_1, \tag{A.8}$$

where Ω_1 is the solid angle of particle 1. Note that λ_{12} here is a constant, related to the masses of the final state photons. In the following sections one or both of the photons will appear as a particle–anti-particle pair who's combined mass is variable. Integrating over the solid angle of photon 1 yields a constant phase space contribution of

$$\Phi_2 = \mathcal{S}\frac{\lambda_{12}}{8\pi}, \tag{A.9}$$

where the symmetry factor $\mathcal{S}$ is a half if the two particles are identical and one otherwise. For two real photons $\lambda_{12} = 1$ and $\mathcal{S} = 1/2$, so $\Phi_{\gamma\gamma} = 1/(16\pi)$.

A.2 3–Body Phase Space

Now imagine that one of the massive photons converts into an electron–positron pair. The number of phase space variables here is 5, three of which will be Euler angles describing the orientation of the CM frame with respect to the lab frame. The remaining two will be chosen to describe the mass of the virtual photon and the direction of the real photon in the e^+e^- CM frame. If the positron, electron, photon, and pion momenta

are p,q,k, and P, then

$$x = \frac{(p+q)^2}{P^2}, \tag{A.10a}$$

$$y = \frac{2P\cdot(p-q)}{P^2(1-x)}. \tag{A.10b}$$

The allowed range of x is from $4m^2/M^2$ to 1 and that of y runs from $-(1-x)$ to $+(1-x)$.

The differential phase space element is

$$d^5\Phi_{ee\gamma} = \frac{1}{(2\pi)^5}\frac{d^3p}{2E_p}\frac{d^3q}{2E_q}\frac{d^3k}{2E_k}\delta^4(P-p-q-k). \tag{A.11}$$

After integrating over the uninteresting degrees of freedom, one is left with

$$d^2\Phi_{ee\gamma} = \frac{M^2}{2^8\pi^3}(1-x)\ dx\ dy. \tag{A.12}$$

A.3 4–Body Phase Space

The 4–body phase space has five dimensions (not counting the 3 Euler angles). These are again defined in terms of 2–particle subsystems. The invariant masses of the two pairs and the direction of one pair in the CM frame of the other make up the first four variables. The last one is the angle between the normals of the two planes formed by the pairs. The definitions of these quantities are

$$x_{12} = (p_1+p_2)^2/P^2, \tag{A.13a}$$

$$x_{34} = (p_3+p_4)^2/P^2, \tag{A.13b}$$

$$y_{12} = \frac{2p_{34}\cdot(p_1-p_2) - P^2\delta_{12}z}{P^2\lambda}, \tag{A.13c}$$

$$y_{34} = \frac{2p_{12}\cdot(p_3-p_4) - P^2\delta_{34}z}{P^2\lambda}, \tag{A.13d}$$

$$\phi = \arctan\frac{16\epsilon_{\mu\nu\rho\sigma}p_1^\mu p_2^\nu p_3^\rho p_4^\sigma}{P^2\lambda[P^2zy_{12}y_{34} - 2(p_1-p_2)\cdot(p_3-p_4)]}, \tag{A.13e}$$

where $\delta_{ij} = (m_i^2 - m_j^2)/(P^2x_{ij})$, $z = 1 - x_{12} - x_{34}$, $w = 2\sqrt{x_{12}x_{34}}$, and $\lambda = \sqrt{z^2-w^2}$.

The two modes discussed in this thesis with a 4–body final state are $\pi^0 \to e^+e^-\gamma\gamma$ and $\pi^0 \to e^+e^-e^+e^-$. In both cases the δ's are zero and the 4–body phase space has

the form

$$d^5\Phi_4 = \mathcal{S}\frac{M^4}{2^{14}\pi^6}\lambda\ dx_{12}\ dx_{34}\ dy_{12}\ dy_{34}\ d\phi, \tag{A.14}$$

where $\mathcal{S} = 1/2$ for $\pi^0 \to e^+e^-\gamma\gamma$ and $\mathcal{S} = 1/4$ for $\pi^0 \to e^+e^-e^+e^-$.

A.4 5–Body Phase Space

This is getting absurd. Ignoring the Euler angles, the 5–body phase space has 8 dimensions. These are simply generalized from the previous examples. There are three x variables, three y variables, and two ϕ angles. In relation to the $\pi^0 \to e^+e^-e^+e^-\gamma$ final state, they are defined as:

$$x_{12} = (p_1 + p_2)^2/P^2, \tag{A.15a}$$

$$x_{34\gamma} = (p_3 + p_4 + p_\gamma)^2/P^2, \tag{A.15b}$$

$$x_{34} = (p_3 + p_4)^2/(p_3 + p_4 + p_\gamma)^2, \tag{A.15c}$$

The phase space element is

$$d^8\Phi_{4e\gamma} = \mathcal{S}\frac{M^4}{2^{14}\pi^6}\lambda\ \lambda_{12}\ \lambda_{34}\ dx_{12}\ dx_{34\gamma}\ dx_{34}\ dy_{12}\ dy_{34\gamma}\ dy_{34}\ d\phi_{12}\ d\phi_{34}, \tag{A.16}$$

Appendix B

The Meson–$\gamma\gamma$ Couplings

In this appendix I will evaluate the two allowed couplings between a meson and two photons by explicate computation in the helicity basis. The photon momenta are defined as in Appendix A and the polarization vectors are

$$\epsilon^{+}(\pm\hat{\boldsymbol{z}}) = \frac{1}{\sqrt{2}}\,(0, 1, \pm\mathrm{i}, 0)\,, \tag{B.1a}$$

$$\epsilon^{-}(\pm\hat{\boldsymbol{z}}) = \frac{1}{\sqrt{2}}\,(0, 1, \mp\mathrm{i}, 0)\,, \tag{B.1b}$$

$$\epsilon^{0}(\pm\hat{\boldsymbol{z}}) = \frac{1}{\sqrt{k^2}}\,(k, 0, 0, \pm E)\,, \tag{B.1c}$$

where the longitudinal polarization is allowed for off–shell photons.

The two couplings of the matrix element (see Equation (2.12)) can be written as

$$\mathcal{H}_{\lambda_1\lambda_2} = H_{\mu\nu\rho\sigma} k_1^{\mu} \epsilon_{\lambda_1}^{*\nu} k_2^{\rho} \epsilon_{\lambda_2}^{*\sigma}, \tag{B.2}$$

where

$$H_{\mu\nu\rho\sigma} = \begin{cases} \frac{2}{M}\tilde{g} f(x_1, x_2) \cos\zeta \epsilon_{\mu\nu\rho\sigma} & \text{pseudoscalar coupling,} \\ \frac{2}{M}\tilde{g} f(x_1, x_2) \sin\zeta e^{i\delta} \left(g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}\right) & \text{scalar coupling} \end{cases} \tag{B.3}$$

Starting with the scalar case,

$$\mathcal{H}^{S}_{\lambda_1\lambda_2} = \frac{2}{M}\tilde{g} f(x_1, x_2) \sin\zeta e^{i\delta} \left[(k_1 \cdot k_2)(\epsilon^{*}_{\lambda_1} \cdot \epsilon^{*}_{\lambda_2}) - (k_1 \cdot \epsilon^{*}_{\lambda_2})(k_2 \cdot \epsilon^{*}_{\lambda_1})\right]. \tag{B.4}$$

Of the nine possible helicity combinations, only three are non–zero. Using the explicit

forms of the polarization vectors,

$$\mathcal{H}^S_{\lambda_1\lambda_2} = \begin{cases} -M\tilde{g}f(x_1,x_2)\sin\zeta e^{i\delta}z & \text{for } \lambda_1=\lambda_2=+, \\ -M\tilde{g}f(x_1,x_2)\sin\zeta e^{i\delta}z & \text{for } \lambda_1=\lambda_2=-, \\ +M\tilde{g}f(x_1,x_2)\sin\zeta e^{i\delta}w & \text{for } \lambda_1=\lambda_2=0, \end{cases} \tag{B.5}$$

where z and w are defined in Appendix A.

For the pseudoscalar case,

$$\mathcal{H}^P_{\lambda_1\lambda_2} = \frac{2}{M}\tilde{g}f(x_1,x_2)\cos\zeta\epsilon_{\mu\nu\rho\sigma}P^\mu\epsilon^{*\nu}_{\lambda_1}k_2^\rho\epsilon^{*\sigma}_{\lambda_2}, \tag{B.6a}$$

$$= 2\tilde{g}f(x_1,x_2)\cos\zeta k\left(\epsilon^{*1}_{\lambda_1}\epsilon^{*2}_{\lambda_2} - \epsilon^{*2}_{\lambda_1}\epsilon^{*1}_{\lambda_2}\right). \tag{B.6b}$$

Here there are only two non–zero combinations

$$\mathcal{H}^P_{\lambda_1\lambda_2} = \begin{cases} +\mathrm{i}M\tilde{g}f(x_1,x_2)\cos\zeta\lambda & \text{for } \lambda_1=\lambda_2=+, \\ -\mathrm{i}M\tilde{g}f(x_1,x_2)\cos\zeta\lambda & \text{for } \lambda_1=\lambda_2=-, \\ 0 & \text{for } \lambda_1=\lambda_2=0, \end{cases} \tag{B.7}$$

where λ is defined in Appendix A.

There are several interesting features of the two couplings. Although the transverse couplings have different kinematic dependence when the photons are off–shell, these differences are small. The significant difference in the transverse terms is the relative sign between the $|{+}{+}\rangle$ state and the $|{-}{-}\rangle$ state. This reflects the fact that the symmetric state is P–even while the antisymmetric state is P–odd. The last difference is the extra longitudinal coupling in the scalar case which only contributes when both photons are off–shell. For the decay to two real photons $\lambda = z = 1$ and $w = 0$.

Appendix C

Helicity Analysis

The matrix element for modes with Dalitz pairs has a similar structure as the matrix element of the two–photon decay. We can simply make the substitution

$$k^\mu \epsilon_\lambda^{*\nu} \rightarrow \sum_\lambda k^\mu \epsilon_\lambda^{*\nu} L_{rs\lambda}, \tag{C.1}$$

where the leptonic part,

$$L_{rs\lambda} = \frac{\mathrm{i}}{k^2} \epsilon_\lambda^\mu \Gamma_{rs,\mu}, \tag{C.2}$$

is a contraction between the polarization vector of the virtual photon and the lepton current,

$$\Gamma_{rs}^\mu = -\mathrm{i}e\bar{u}_r \gamma_{rs}^\mu v_s. \tag{C.3}$$

The leptonic part is a scalar and can therefore be evaluated in any convenient frame of reference. This appendix demonstrates such a calculation in the center of momentum frame of the Dalitz pair. First, the current, which is a vector, can be calculated in the frame in which the electron travels along the $+\hat{\boldsymbol{z}}$ direction and the positron moves in the $-\hat{\boldsymbol{z}}$ direction. The spinors in the helicity basis take a simple form in this frame

$$u_+(+\hat{\boldsymbol{z}}) = \begin{pmatrix} \Sigma \\ 0 \\ -\Delta \\ 0 \end{pmatrix}, \qquad u_-(+\hat{\boldsymbol{z}}) = \begin{pmatrix} 0 \\ -\Delta \\ 0 \\ \Sigma \end{pmatrix}, \qquad v_+(-\hat{\boldsymbol{z}}) = \begin{pmatrix} -\Delta \\ 0 \\ -\Sigma \\ 0 \end{pmatrix}, \qquad v_-(-\hat{\boldsymbol{z}}) = \begin{pmatrix} 0 \\ \Sigma \\ 0 \\ \Delta \end{pmatrix}, \tag{C.4}$$

where $\Sigma = \sqrt{E+p}$ and $\Delta = \sqrt{E-p}$. After the spinors have been contracted, the current for each of the four helicity states will be rotated into an arbitrary direction

with polar angle θ and azimuthal angle ϕ, and contracted with one of the following polarization vectors

$$\epsilon_+ = \frac{1}{\sqrt{2}}\left(0, \quad 1, \quad +\mathrm{i}, \quad 0\right), \tag{C.5a}$$

$$\epsilon_- = \frac{1}{\sqrt{2}}\left(0, \quad 1, \quad -\mathrm{i}, \quad 0\right), \tag{C.5b}$$

$$\epsilon_0 = \frac{1}{\sqrt{k^2}}\left(|\boldsymbol{k}|, \quad 0, \quad 0, \quad E\right). \tag{C.5c}$$

Therefore, there are $4 \times 3 = 12$ combinations of the helicities r, s, and λ.

The four lepton currents in the special frame are

$$\Gamma_{++} = +\mathrm{i}eM\sqrt{x_0}\left(0, \quad 0, \quad 0, \quad 1\right), \tag{C.6a}$$

$$\Gamma_{+-} = -\mathrm{i}eM\sqrt{x}\left(0, \quad 1, \quad -\mathrm{i}, \quad 0\right), \tag{C.6b}$$

$$\Gamma_{-+} = -\mathrm{i}eM\sqrt{x}\left(0, \quad 1, \quad +\mathrm{i}, \quad 0\right), \tag{C.6c}$$

$$\Gamma_{--} = -\mathrm{i}eM\sqrt{x_0}\left(0, \quad 0, \quad 0, \quad 1\right). \tag{C.6d}$$

The rotation matrix is

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ 0 & \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{C.7}$$

The currents in the rotated frame are then

$$\Gamma_{++} = +\mathrm{i}eM\sqrt{x_0}\left(0, \quad \sin\theta\cos\phi, \quad \sin\theta\sin\phi, \quad \cos\theta\right), \tag{C.8a}$$

$$\Gamma_{+-} = -\mathrm{i}eM\sqrt{x}\left(0, \quad \cos\theta\cos\phi + \mathrm{i}\sin\phi, \quad \cos\theta\sin\phi - \mathrm{i}\cos\phi, \quad -\sin\theta\right), \tag{C.8b}$$

$$\Gamma_{-+} = -\mathrm{i}eM\sqrt{x}\left(0, \quad \cos\theta\cos\phi - \mathrm{i}\sin\phi, \quad \cos\theta\sin\phi + \mathrm{i}\cos\phi, \quad -\sin\theta\right), \tag{C.8c}$$

$$\Gamma_{--} = -\mathrm{i}eM\sqrt{x_0}\left(0, \quad \sin\theta\cos\phi, \quad \sin\theta\sin\phi, \quad \cos\theta\right). \tag{C.8d}$$

The 12 lepton parts are

$$L_{++,+} = + \frac{e}{\sqrt{2}M}\frac{\eta}{\sqrt{x}}\sin\theta \mathrm{e}^{+\mathrm{i}\phi}, \tag{C.9a}$$

$$L_{++,-} = + \frac{e}{\sqrt{2}M}\frac{\eta}{\sqrt{x}}\sin\theta \mathrm{e}^{-\mathrm{i}\phi}, \tag{C.9b}$$

$$L_{++,0} = + \frac{e}{M}\frac{\eta}{\sqrt{x}}\cos\theta, \tag{C.9c}$$

$$L_{+-,+} = - \frac{e}{\sqrt{2}M}\frac{1}{\sqrt{x}}(1+\cos\theta)\mathrm{e}^{+\mathrm{i}\phi}, \tag{C.9d}$$

$$L_{+-,-} = + \frac{e}{\sqrt{2}M}\frac{1}{\sqrt{x}}(1-\cos\theta)\mathrm{e}^{-\mathrm{i}\phi}, \tag{C.9e}$$

$$L_{+-,0} = + \frac{e}{M}\frac{1}{\sqrt{x}}\sin\theta, \tag{C.9f}$$

$$L_{-+,+} = + \frac{e}{\sqrt{2}M}\frac{1}{\sqrt{x}}(1-\cos\theta)\mathrm{e}^{+\mathrm{i}\phi}, \tag{C.9g}$$

$$L_{-+,-} = - \frac{e}{\sqrt{2}M}\frac{1}{\sqrt{x}}(1+\cos\theta)\mathrm{e}^{-\mathrm{i}\phi}, \tag{C.9h}$$

$$L_{-+,0} = + \frac{e}{M}\frac{1}{\sqrt{x}}\sin\theta, \tag{C.9i}$$

$$L_{--,+} = - \frac{e}{\sqrt{2}M}\frac{\eta}{\sqrt{x}}\sin\theta \mathrm{e}^{+\mathrm{i}\phi}, \tag{C.9j}$$

$$L_{--,-} = - \frac{e}{\sqrt{2}M}\frac{\eta}{\sqrt{x}}\sin\theta \mathrm{e}^{-\mathrm{i}\phi}, \tag{C.9k}$$

$$L_{--,0} = - \frac{e}{M}\frac{\eta}{\sqrt{x}}\cos\theta, \tag{C.9l}$$

Appendix D

Likelihood Table

This appendix contains the tabulated likelihood function for the complete $\pi^0 \to e^+e^-e^+e^-$ dataset. The likelihood function was calculated on a 3-dimensional array with 21 bins in each direction. The explored range of the three parameters was

$$-10 \leq \alpha \leq +10 \tag{D.1a}$$

$$-0.05 \leq \kappa \leq +0.05 \tag{D.1b}$$

$$0.00 \leq \eta \leq 0.25 \tag{D.1c}$$

For convenience, the data are presented in a series of 2-dimensional tables, each at a fixed value of α. Within a table, each row is a bin in κ while each column is a bin in η. The translation from bin number to actual value is given by

$$\alpha(i) = -10 + (i-1) \times 1, \tag{D.2a}$$

$$\kappa(i) = -0.05 + (i-1) \times 0.005, \tag{D.2b}$$

$$\eta(i) = 0 + (i-1) \times 0.0125. \tag{D.2c}$$

The range of bins included in the tables is $5 \leq i_\alpha \leq 15$, $0 \leq i_\kappa \leq 20$, and $0 \leq i_\eta \leq 18$.

The actual number given for each bin is the change in the likelihood function between the bin in question and the bin at the minimum, scaled by 100,

$$N(\alpha, \kappa, \eta) = 100 \times [L(\alpha, \kappa, \eta) - L_{\min}]. \tag{D.3}$$

The factor of 100 is included simply for convenience. The tables are color coded to indicate points within 1,2, and $3\,\sigma$ of the minimum. With the scale factor, values less than 177 (in blue) lay within $1\,\sigma$, values less than 402 (in green) lay within $2\,\sigma$, and values less than 708 (in magenta) lay within $3\,\sigma$.

Table D.1: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 5$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	3467	3461	3446	3421	3390	3353	3315	3279	3249	3229	3224	3238	3277	3345	3447	3587	3771	4004	4288
1	3199	3194	3178	3154	3122	3086	3048	3012	2983	2963	2958	2973	3013	3081	3183	3325	3510	3743	4028
2	2962	2957	2941	2917	2886	2850	2812	2777	2747	2728	2724	2739	2779	2848	2951	3093	3278	3512	3798
3	2756	2751	2736	2712	2680	2645	2607	2572	2543	2524	2520	2536	2576	2645	2749	2891	3077	3311	3598
4	2581	2576	2561	2537	2506	2470	2433	2398	2369	2351	2347	2363	2403	2473	2577	2720	2906	3141	3428
5	2438	2433	2418	2394	2363	2327	2291	2256	2227	2209	2205	2221	2262	2332	2436	2579	2766	3001	3289
6	2326	2321	2306	2282	2251	2216	2179	2144	2116	2097	2094	2110	2151	2221	2326	2469	2656	2891	3179
7	2245	2240	2225	2201	2171	2135	2099	2064	2036	2017	2014	2031	2072	2142	2246	2390	2577	2812	3100
8	2196	2191	2176	2152	2121	2086	2050	2015	1987	1969	1965	1982	2023	2093	2197	2341	2528	2763	3051
9	2178	2173	2158	2134	2104	2068	2032	1997	1969	1951	1948	1964	2005	2075	2179	2323	2510	2745	3033
10	2192	2186	2171	2148	2117	2082	2045	2011	1982	1964	1961	1977	2018	2088	2192	2335	2522	2757	3044
11	2237	2231	2216	2193	2162	2127	2090	2056	2027	2009	2006	2022	2062	2132	2236	2378	2565	2799	3086
12	2313	2308	2293	2269	2238	2203	2167	2132	2103	2085	2081	2097	2137	2207	2310	2452	2638	2872	3159
13	2421	2416	2401	2377	2346	2311	2274	2239	2211	2192	2188	2203	2243	2312	2415	2557	2742	2975	3261
14	2560	2555	2540	2516	2486	2450	2413	2378	2349	2330	2326	2341	2380	2449	2551	2692	2877	3109	3394
15	2731	2726	2711	2687	2656	2621	2584	2548	2519	2500	2495	2509	2548	2616	2718	2858	3042	3273	3557
16	2934	2928	2913	2889	2858	2823	2785	2750	2720	2700	2695	2709	2747	2814	2915	3054	3237	3468	3751
17	3167	3162	3147	3123	3092	3056	3018	2982	2952	2932	2926	2939	2977	3043	3143	3281	3463	3693	3974
18	3432	3427	3412	3388	3356	3320	3282	3246	3215	3194	3188	3201	3237	3303	3401	3539	3719	3948	4228
19	3729	3723	3708	3684	3652	3616	3577	3541	3510	3488	3481	3493	3528	3593	3691	3827	4006	4233	4511
20	4056	4051	4036	4011	3979	3942	3904	3867	3835	3813	3805	3816	3850	3914	4010	4145	4323	4548	4825

Table D.2: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 6$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	2918	2913	2897	2871	2838	2800	2760	2721	2688	2665	2656	2667	2701	2764	2861	2996	3175	3401	3679
1	2651	2645	2629	2604	2571	2533	2493	2455	2422	2400	2391	2402	2437	2501	2598	2735	2914	3141	3420
2	2414	2409	2393	2368	2335	2297	2258	2220	2187	2165	2157	2169	2204	2268	2366	2503	2683	2910	3190
3	2209	2204	2188	2163	2130	2093	2053	2016	1984	1961	1954	1966	2001	2066	2165	2302	2482	2710	2991
4	2035	2029	2014	1989	1956	1919	1880	1842	1810	1789	1781	1793	1829	1895	1993	2131	2312	2540	2821
5	1892	1886	1871	1846	1814	1776	1737	1700	1668	1647	1640	1652	1688	1754	1853	1991	2172	2401	2682
6	1780	1775	1759	1734	1702	1665	1626	1589	1557	1536	1529	1541	1578	1643	1743	1881	2062	2291	2573
7	1699	1694	1679	1654	1622	1585	1546	1509	1477	1456	1449	1462	1498	1564	1663	1802	1983	2212	2494
8	1650	1645	1629	1605	1573	1536	1497	1460	1429	1407	1401	1413	1450	1515	1615	1753	1934	2164	2445
9	1632	1627	1612	1587	1555	1518	1479	1442	1411	1390	1383	1395	1432	1497	1597	1735	1916	2145	2427
10	1646	1641	1625	1601	1569	1532	1493	1456	1425	1403	1396	1409	1445	1510	1610	1747	1928	2157	2438
11	1691	1686	1670	1645	1613	1577	1538	1501	1469	1448	1441	1453	1489	1554	1653	1790	1971	2200	2480
12	1767	1762	1747	1722	1690	1653	1614	1577	1545	1524	1516	1528	1564	1629	1727	1864	2044	2272	2553
13	1875	1870	1854	1830	1797	1760	1722	1684	1653	1631	1623	1635	1670	1734	1832	1969	2148	2376	2655
14	2014	2009	1993	1969	1937	1899	1860	1823	1791	1769	1761	1772	1807	1871	1968	2104	2283	2509	2788
15	2185	2180	2164	2139	2107	2070	2030	1993	1960	1938	1929	1940	1975	2038	2134	2269	2447	2673	2951
16	2387	2382	2366	2341	2309	2271	2232	2194	2161	2138	2129	2139	2173	2235	2331	2465	2642	2867	3144
17	2620	2615	2599	2574	2541	2504	2464	2426	2393	2369	2360	2369	2402	2464	2559	2692	2868	3091	3367
18	2885	2879	2864	2838	2806	2768	2728	2689	2656	2631	2621	2630	2662	2723	2817	2949	3124	3346	3620
19	3181	3175	3159	3134	3101	3063	3022	2983	2949	2925	2914	2922	2953	3013	3106	3236	3410	3631	3903
20	3508	3502	3486	3461	3427	3389	3348	3309	3274	3249	3237	3244	3275	3333	3425	3554	3726	3946	4216

Table D.3: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 7$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	2464	2458	2441	2415	2380	2341	2298	2257	2221	2195	2183	2189	2219	2278	2370	2500	2673	2893	3165
1	2197	2191	2175	2148	2114	2074	2033	1992	1956	1930	1918	1925	1956	2015	2108	2239	2412	2633	2906
2	1961	1955	1939	1913	1879	1839	1798	1757	1722	1696	1685	1692	1723	1783	1876	2007	2182	2403	2677
3	1756	1751	1734	1708	1674	1635	1594	1553	1518	1493	1482	1490	1521	1581	1675	1807	1981	2204	2478
4	1582	1577	1560	1534	1501	1462	1420	1380	1346	1321	1310	1318	1350	1410	1504	1636	1811	2034	2309
5	1440	1434	1418	1392	1358	1319	1278	1239	1204	1179	1169	1177	1209	1270	1364	1496	1672	1895	2170
6	1328	1323	1306	1281	1247	1208	1167	1128	1093	1069	1058	1067	1099	1160	1254	1387	1562	1786	2061
7	1248	1242	1226	1200	1167	1128	1087	1048	1014	989	979	987	1020	1080	1175	1308	1483	1707	1982
8	1199	1193	1177	1151	1118	1079	1039	999	965	940	930	939	971	1032	1126	1259	1435	1658	1933
9	1181	1176	1159	1134	1100	1062	1021	982	948	923	913	921	953	1014	1109	1241	1417	1640	1915
10	1195	1189	1173	1147	1114	1076	1035	995	961	937	926	934	967	1027	1121	1254	1429	1652	1927
11	1240	1234	1218	1192	1159	1120	1080	1040	1006	981	971	979	1011	1071	1165	1297	1472	1694	1969
12	1316	1310	1294	1269	1235	1197	1156	1116	1082	1057	1046	1054	1086	1146	1239	1371	1545	1767	2041
13	1424	1418	1402	1376	1343	1304	1263	1223	1189	1164	1152	1160	1191	1251	1344	1475	1649	1870	2143
14	1563	1557	1541	1515	1482	1443	1402	1362	1327	1301	1290	1297	1328	1387	1479	1610	1783	2003	2276
15	1733	1727	1711	1685	1652	1613	1571	1531	1496	1470	1458	1465	1495	1554	1645	1775	1947	2167	2438
16	1934	1929	1913	1887	1853	1814	1772	1732	1696	1670	1658	1664	1693	1751	1842	1971	2142	2361	2631
17	2167	2162	2145	2119	2086	2046	2004	1964	1928	1901	1888	1894	1922	1979	2069	2197	2367	2585	2854
18	2431	2426	2409	2383	2349	2310	2268	2226	2190	2163	2149	2154	2182	2238	2327	2454	2623	2839	3107
19	2727	2721	2705	2678	2644	2604	2562	2520	2483	2455	2441	2445	2472	2527	2615	2741	2908	3123	3389
20	3053	3048	3031	3005	2970	2930	2887	2845	2807	2779	2764	2767	2793	2847	2934	3058	3224	3437	3702

Table D.4: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 8$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	2105	2099	2082	2054	2019	1977	1933	1889	1851	1821	1805	1808	1834	1888	1975	2099	2267	2481	2746
1	1839	1833	1816	1789	1753	1712	1668	1624	1586	1557	1542	1545	1571	1626	1713	1839	2007	2222	2488
2	1603	1598	1581	1553	1518	1477	1433	1390	1352	1323	1308	1312	1339	1394	1482	1608	1777	1992	2260
3	1399	1393	1376	1349	1314	1273	1230	1187	1149	1121	1106	1110	1137	1193	1281	1408	1577	1793	2061
4	1226	1220	1203	1176	1141	1100	1057	1014	977	949	934	938	966	1022	1111	1238	1407	1624	1892
5	1083	1078	1061	1034	999	958	915	873	836	808	793	798	825	882	971	1098	1268	1485	1754
6	972	966	950	923	888	847	804	762	725	697	683	688	716	772	861	989	1159	1376	1645
7	892	886	870	843	808	768	725	683	646	618	604	609	637	693	782	910	1080	1297	1566
8	843	837	821	794	759	719	676	634	597	569	556	560	588	645	734	861	1032	1249	1518
9	826	820	803	777	742	702	659	617	580	552	538	543	571	627	716	844	1014	1231	1500
10	839	833	817	790	756	715	672	630	593	566	552	556	584	640	729	856	1026	1243	1511
11	884	878	862	835	800	760	717	675	638	610	596	600	628	684	773	899	1069	1285	1553
12	960	955	938	911	877	836	793	751	714	686	671	675	703	758	847	973	1142	1358	1625
13	1068	1062	1045	1019	984	944	900	858	821	792	778	781	808	863	951	1077	1245	1461	1728
14	1207	1201	1184	1157	1123	1082	1039	996	959	930	915	918	945	999	1087	1212	1379	1594	1860
15	1377	1371	1354	1327	1292	1252	1208	1166	1128	1099	1083	1086	1112	1166	1252	1377	1543	1757	2022
16	1578	1572	1555	1528	1493	1453	1409	1366	1328	1298	1282	1284	1310	1363	1449	1572	1738	1951	2215
17	1810	1805	1788	1761	1726	1684	1641	1597	1559	1529	1512	1514	1538	1591	1675	1798	1963	2174	2437
18	2074	2068	2051	2024	1989	1947	1903	1860	1820	1790	1773	1774	1797	1849	1933	2054	2218	2428	2689
19	2369	2363	2346	2319	2283	2241	2197	2153	2113	2082	2064	2064	2087	2137	2220	2341	2503	2712	2972
20	2695	2689	2672	2644	2608	2566	2521	2477	2437	2405	2386	2386	2407	2457	2538	2657	2818	3025	3284

Table D.5: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 9$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	1844	1838	1820	1791	1754	1711	1665	1619	1577	1545	1525	1524	1545	1595	1677	1796	1958	2166	2426
1	1578	1572	1554	1526	1489	1446	1400	1355	1313	1281	1262	1261	1283	1333	1416	1536	1699	1908	2168
2	1343	1337	1320	1292	1255	1212	1166	1121	1080	1048	1030	1029	1052	1102	1185	1306	1469	1679	1940
3	1139	1133	1116	1088	1051	1009	963	918	877	846	828	828	850	901	985	1106	1270	1480	1742
4	966	960	943	915	879	836	791	746	706	674	656	656	680	731	815	937	1101	1311	1573
5	824	818	801	773	737	695	649	605	565	533	516	516	540	591	675	797	962	1173	1435
6	713	708	690	662	626	584	539	494	454	423	406	406	430	482	566	688	853	1064	1327
7	634	628	610	583	547	505	459	415	375	344	327	327	351	403	487	610	774	986	1248
8	585	579	562	534	498	456	411	367	327	296	278	279	303	355	439	561	726	937	1200
9	567	562	544	517	481	439	394	349	310	279	261	262	286	337	422	544	708	919	1182
10	581	575	558	530	494	452	407	363	323	292	275	275	299	350	434	556	720	931	1194
11	626	620	603	575	539	497	452	408	368	337	319	319	343	394	478	599	763	974	1235
12	702	696	679	651	615	573	528	484	444	412	394	394	417	468	552	673	836	1046	1307
13	809	804	786	759	723	680	635	591	550	519	501	500	523	573	656	777	940	1149	1410
14	948	942	925	897	861	819	773	729	688	656	638	637	659	709	791	911	1073	1282	1542
15	1118	1112	1094	1067	1030	988	943	897	857	825	806	804	826	875	957	1076	1237	1445	1704
16	1319	1313	1295	1268	1231	1189	1143	1097	1056	1024	1004	1003	1024	1072	1153	1271	1431	1638	1896
17	1551	1545	1527	1499	1463	1420	1374	1328	1287	1254	1234	1231	1252	1299	1379	1497	1656	1861	2118
18	1814	1808	1790	1762	1726	1683	1636	1590	1548	1515	1494	1491	1510	1557	1636	1752	1910	2115	2370
19	2108	2102	2085	2056	2019	1976	1930	1883	1841	1806	1785	1781	1800	1845	1923	2038	2195	2398	2652
20	2433	2427	2410	2381	2344	2301	2254	2207	2163	2129	2106	2102	2119	2164	2241	2354	2510	2711	2963

Table D.6: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 10$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	1681	1675	1657	1628	1589	1544	1496	1447	1403	1367	1344	1339	1357	1401	1479	1593	1749	1951	2205
1	1417	1410	1392	1363	1325	1280	1232	1184	1140	1104	1082	1077	1095	1141	1218	1333	1490	1693	1947
2	1182	1176	1158	1129	1091	1046	998	951	907	872	850	846	864	910	988	1104	1261	1465	1720
3	979	973	955	926	888	844	796	748	705	670	648	644	663	709	788	904	1062	1267	1522
4	806	800	782	753	716	671	624	577	534	499	477	474	493	539	619	735	893	1098	1354
5	665	658	641	612	574	530	483	436	393	358	337	334	353	400	479	596	755	960	1216
6	554	548	530	501	464	420	373	326	283	249	228	224	244	291	370	487	646	851	1108
7	474	468	450	422	384	341	293	247	204	170	149	146	165	212	292	409	568	773	1030
8	426	420	402	373	336	292	245	198	156	122	101	98	117	164	244	361	520	725	982
9	408	402	384	356	319	275	228	181	139	104	83	80	100	147	226	343	502	707	963
10	422	416	398	370	332	289	242	195	152	118	97	94	113	160	239	356	514	719	975
11	467	461	443	415	377	334	286	240	197	163	141	138	157	203	282	399	557	762	1017
12	543	537	519	491	453	409	362	315	272	238	217	213	232	278	356	472	630	834	1089
13	650	644	626	598	560	516	469	422	379	344	323	318	337	383	461	576	733	937	1191
14	789	782	765	736	698	655	607	560	517	482	459	455	473	518	596	710	867	1069	1323
15	958	952	934	905	868	824	776	729	685	650	627	622	639	684	761	875	1030	1232	1485
16	1159	1152	1135	1106	1068	1024	976	928	884	849	826	820	837	881	957	1070	1224	1425	1677
17	1390	1384	1366	1337	1299	1255	1207	1159	1114	1078	1055	1048	1064	1108	1183	1295	1448	1648	1898
18	1653	1647	1629	1600	1562	1517	1469	1420	1375	1339	1314	1308	1323	1365	1439	1550	1702	1901	2150
19	1947	1941	1922	1893	1855	1810	1761	1712	1667	1630	1605	1597	1611	1653	1726	1835	1986	2184	2431
20	2271	2265	2247	2218	2179	2134	2085	2035	1989	1951	1926	1917	1931	1971	2043	2151	2301	2496	2742

Table D.7: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 11$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	1620	1614	1595	1565	1525	1478	1428	1377	1330	1291	1265	1256	1269	1309	1381	1490	1641	1838	2085
1	1356	1349	1331	1301	1261	1215	1164	1114	1067	1029	1003	994	1008	1049	1122	1231	1383	1580	1828
2	1122	1116	1097	1067	1028	982	932	881	835	797	771	763	777	818	892	1002	1154	1352	1601
3	919	913	894	864	825	779	729	679	633	595	570	562	577	619	692	803	956	1154	1404
4	747	741	722	692	653	607	558	508	462	425	400	392	407	449	523	634	787	986	1236
5	606	599	581	551	512	467	417	368	322	284	260	252	267	310	384	496	649	848	1098
6	495	489	470	441	402	357	307	258	212	175	150	143	158	201	275	387	540	740	990
7	416	410	391	362	323	277	228	179	133	96	72	65	80	122	197	309	462	662	912
8	368	361	343	313	275	229	180	131	85	48	24	17	32	75	149	261	414	614	864
9	350	344	326	296	257	212	163	114	68	31	7	0	15	57	132	243	397	596	846
10	364	358	339	310	271	226	177	128	82	45	20	13	28	70	145	256	409	608	858
11	409	403	384	355	316	271	222	172	127	89	65	57	72	114	188	299	452	651	900
12	485	478	460	431	392	347	297	248	202	165	140	132	147	188	262	372	525	723	972
13	592	585	567	538	499	453	404	355	309	271	246	238	252	293	366	476	628	825	1074
14	730	724	705	676	637	591	542	492	446	408	382	374	388	428	501	610	761	958	1205
15	899	893	874	845	806	760	710	660	614	576	550	541	554	594	666	774	924	1121	1367
16	1099	1093	1075	1045	1006	960	910	860	813	774	748	738	751	790	861	969	1118	1313	1558
17	1331	1324	1306	1276	1237	1191	1141	1090	1043	1004	976	966	978	1017	1087	1194	1342	1536	1780
18	1593	1587	1568	1538	1499	1452	1402	1351	1303	1263	1236	1225	1236	1274	1343	1449	1595	1788	2031
19	1886	1880	1861	1831	1791	1745	1694	1643	1595	1554	1526	1514	1524	1561	1629	1734	1879	2070	2312
20	2210	2204	2185	2155	2115	2068	2017	1965	1916	1875	1846	1834	1843	1878	1945	2049	2193	2382	2622

Table D.8: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 12$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	1661	1655	1635	1604	1563	1515	1462	1409	1359	1317	1287	1275	1284	1319	1387	1491	1636	1827	2068
1	1398	1391	1372	1341	1300	1252	1200	1147	1097	1056	1026	1014	1023	1060	1128	1232	1378	1570	1811
2	1165	1158	1139	1108	1067	1019	967	915	866	824	795	783	793	830	898	1004	1150	1342	1585
3	962	955	936	905	865	817	766	713	664	623	594	583	593	630	699	805	952	1145	1388
4	790	784	765	734	694	646	594	542	494	453	424	413	424	461	531	637	784	977	1221
5	649	643	624	593	553	506	454	402	354	313	285	274	284	322	392	498	646	839	1083
6	539	533	514	483	443	396	344	293	244	204	176	165	176	214	283	390	538	731	976
7	460	454	435	404	364	317	266	214	166	125	97	86	97	136	205	312	460	654	898
8	412	405	386	356	316	269	218	166	118	78	50	39	50	88	158	264	412	606	850
9	395	388	369	339	299	252	201	149	101	60	32	22	33	71	140	247	394	588	832
10	408	402	383	353	313	266	214	163	115	74	46	35	46	84	153	259	407	600	844
11	453	447	428	397	357	310	259	208	159	119	90	79	90	127	197	302	449	642	886
12	529	523	504	473	433	386	335	283	235	194	165	154	164	201	270	376	522	715	958
13	636	629	611	580	540	493	442	390	341	300	271	259	269	306	374	479	625	817	1059
14	774	767	748	718	678	631	579	527	478	437	408	395	405	441	509	613	758	949	1191
15	943	936	917	887	847	799	747	695	646	604	575	562	571	607	674	777	922	1112	1352
16	1143	1136	1117	1087	1046	999	947	894	845	803	773	759	768	802	869	971	1115	1304	1543
17	1374	1367	1348	1317	1277	1229	1177	1124	1074	1031	1001	987	995	1029	1094	1196	1338	1526	1764
18	1635	1629	1610	1579	1538	1490	1438	1384	1334	1291	1260	1245	1252	1285	1350	1450	1591	1778	2015
19	1928	1922	1902	1871	1830	1782	1729	1676	1625	1581	1549	1534	1540	1572	1635	1735	1875	2060	2295
20	2252	2245	2226	2195	2153	2105	2052	1997	1946	1902	1869	1853	1858	1889	1951	2049	2188	2372	2605

Table D.9: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 13$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	1807	1800	1780	1748	1706	1656	1602	1546	1493	1448	1415	1398	1403	1434	1497	1596	1735	1920	2155
1	1544	1537	1517	1485	1443	1394	1339	1284	1232	1187	1154	1138	1143	1175	1238	1338	1478	1664	1900
2	1311	1305	1285	1253	1211	1161	1107	1052	1001	956	924	908	914	946	1010	1110	1251	1437	1673
3	1109	1103	1083	1051	1009	960	906	851	800	756	723	708	714	747	811	911	1053	1240	1477
4	938	931	911	880	838	789	735	681	630	586	554	538	545	578	642	743	885	1073	1310
5	797	791	771	739	698	649	595	541	490	446	414	399	406	439	504	605	747	935	1173
6	687	681	661	630	588	539	486	432	381	337	306	291	298	331	396	497	640	828	1066
7	608	602	582	551	510	461	407	353	302	259	227	213	220	253	318	419	562	750	988
8	560	554	534	503	462	413	360	306	255	211	180	165	172	206	271	372	514	702	940
9	543	537	517	486	445	396	343	289	238	194	163	148	155	188	253	354	497	684	922
10	557	551	531	500	458	410	357	303	252	208	176	162	168	202	266	367	509	697	934
11	602	595	576	545	503	455	401	347	296	253	221	206	212	245	310	410	552	739	976
12	678	671	652	620	579	530	477	423	372	328	296	280	287	319	383	483	625	811	1048
13	784	778	758	727	686	637	583	529	478	434	401	386	392	424	487	587	727	913	1149
14	922	916	896	865	823	774	721	666	615	570	538	522	527	559	621	720	860	1046	1281
15	1091	1084	1065	1033	992	943	889	834	782	737	704	688	693	724	786	884	1023	1208	1442
16	1290	1284	1264	1233	1191	1142	1088	1033	981	935	902	885	889	919	981	1078	1216	1400	1633
17	1521	1514	1495	1463	1421	1372	1318	1262	1210	1164	1130	1112	1116	1145	1206	1302	1439	1621	1853
18	1782	1776	1756	1724	1682	1632	1578	1522	1469	1423	1388	1370	1373	1401	1461	1556	1692	1873	2104
19	2075	2068	2048	2016	1974	1924	1869	1813	1759	1713	1677	1658	1660	1688	1746	1840	1975	2154	2384
20	2398	2391	2371	2339	2296	2246	2191	2134	2080	2033	1997	1977	1977	2004	2061	2154	2287	2465	2693

Table D.10: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 14$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	2059	2052	2032	1999	1955	1903	1847	1789	1734	1685	1649	1628	1629	1656	1713	1807	1941	2121	2349
1	1796	1789	1769	1736	1693	1641	1585	1528	1473	1425	1388	1368	1370	1397	1455	1550	1685	1865	2094
2	1564	1557	1537	1504	1461	1410	1354	1296	1242	1194	1158	1139	1140	1168	1227	1322	1458	1638	1869
3	1363	1356	1335	1303	1260	1209	1153	1096	1042	994	959	939	941	970	1029	1124	1260	1442	1673
4	1192	1185	1165	1132	1089	1038	983	926	872	825	789	770	773	801	861	957	1093	1275	1506
5	1051	1045	1024	992	949	898	843	786	732	685	650	632	634	663	723	819	956	1138	1369
6	942	935	915	883	840	789	734	677	624	577	542	523	526	555	615	711	848	1030	1262
7	863	856	836	804	761	711	656	599	545	499	464	445	448	477	537	634	770	953	1185
8	815	808	788	756	714	663	608	551	498	451	416	398	401	430	490	586	723	905	1137
9	798	791	771	739	697	646	591	535	481	434	399	381	384	413	473	569	705	888	1119
10	812	805	785	753	711	660	605	549	495	448	413	394	397	426	486	582	718	900	1131
11	857	850	830	798	755	705	650	593	539	493	457	439	441	469	529	625	761	942	1173
12	933	926	906	873	831	780	725	669	615	568	532	513	515	543	603	698	833	1014	1245
13	1039	1032	1012	980	937	887	832	775	721	674	638	618	620	648	706	801	936	1116	1346
14	1177	1170	1150	1118	1075	1024	969	912	857	810	774	754	755	782	841	934	1069	1248	1477
15	1345	1338	1318	1286	1243	1192	1137	1079	1025	977	940	920	921	947	1005	1098	1231	1410	1638
16	1544	1538	1517	1485	1442	1391	1335	1278	1223	1175	1138	1117	1117	1143	1199	1291	1424	1602	1829
17	1775	1768	1747	1715	1672	1621	1565	1507	1451	1403	1365	1344	1343	1368	1424	1515	1647	1823	2049
18	2036	2029	2008	1976	1932	1881	1825	1766	1711	1661	1623	1601	1600	1624	1679	1769	1899	2074	2299
19	2327	2320	2300	2267	2224	2172	2115	2057	2000	1950	1912	1889	1886	1910	1963	2052	2181	2355	2578
20	2650	2643	2622	2589	2545	2494	2436	2377	2320	2270	2230	2207	2203	2225	2278	2366	2493	2666	2887

Table D.11: Tabulated values of $100 \times \Delta L$ in κ versus η for $i_\alpha = 15$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	2419	2412	2391	2357	2312	2259	2200	2140	2082	2031	1990	1966	1963	1985	2038	2127	2256	2429	2652
1	2157	2150	2129	2095	2050	1997	1939	1879	1821	1770	1731	1707	1704	1727	1781	1870	1999	2174	2397
2	1925	1918	1897	1864	1819	1766	1708	1649	1591	1541	1501	1478	1475	1499	1553	1643	1773	1948	2172
3	1724	1717	1696	1663	1618	1566	1508	1448	1391	1341	1302	1279	1277	1301	1355	1446	1576	1752	1977
4	1554	1547	1526	1492	1448	1396	1338	1279	1222	1172	1133	1110	1108	1133	1188	1278	1409	1585	1811
5	1414	1407	1386	1353	1308	1256	1199	1140	1083	1033	994	972	970	995	1050	1141	1272	1448	1674
6	1304	1297	1277	1243	1199	1147	1090	1031	974	925	886	864	862	887	942	1033	1165	1341	1567
7	1226	1219	1198	1165	1121	1069	1012	953	896	847	808	786	785	809	865	956	1087	1264	1490
8	1178	1171	1151	1117	1074	1021	964	905	849	799	761	739	737	762	817	909	1040	1217	1443
9	1161	1154	1134	1101	1057	1005	947	889	832	783	744	722	721	745	800	891	1023	1199	1425
10	1175	1168	1148	1114	1071	1019	961	903	846	796	758	736	734	758	813	904	1035	1211	1437
11	1220	1213	1192	1159	1115	1063	1006	947	891	841	802	780	778	802	857	947	1078	1254	1479
12	1296	1289	1268	1235	1191	1139	1082	1023	966	916	877	854	852	876	930	1020	1150	1326	1550
13	1402	1395	1374	1341	1297	1245	1188	1129	1072	1022	982	959	957	980	1034	1123	1253	1428	1651
14	1539	1532	1512	1478	1434	1382	1325	1265	1208	1158	1118	1095	1092	1115	1168	1257	1386	1559	1782
15	1708	1700	1680	1647	1602	1550	1492	1433	1375	1325	1285	1261	1257	1279	1332	1420	1548	1721	1943
16	1906	1899	1879	1845	1801	1749	1691	1631	1573	1522	1481	1457	1453	1474	1526	1613	1740	1912	2133
17	2136	2129	2108	2075	2031	1978	1920	1859	1801	1750	1709	1683	1679	1699	1750	1836	1963	2133	2353
18	2397	2390	2369	2335	2291	2238	2179	2119	2060	2008	1966	1940	1935	1955	2005	2090	2215	2384	2603
19	2688	2681	2660	2626	2581	2528	2469	2408	2349	2296	2254	2228	2221	2240	2289	2373	2496	2665	2882
20	3010	3003	2982	2948	2903	2849	2790	2729	2669	2615	2572	2545	2537	2555	2603	2686	2808	2975	3190