Measuring the branching ratio of the rare decay

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

by

Rune Niclasen

M.S., University of Copenhagen, 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

2006

This thesis entitled:
Measuring the branching ratio of the rare decay $\pi^0 \to e^+e^-$
written by Rune Niclasen
has been approved for the Department of Physics

Eric 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.

Niclasen, Rune (Ph.D.)

Measuring the branching ratio of the rare decay $\pi^0 \to e^+e^-$

Thesis directed by Prof. Eric Zimmerman

A precise branching ratio measurement of the rare decay $\pi^0 \to e^+e^-$ has been made. The measurement was made with the rare kaon decay experiment KTeV at Fermilab where the source of π^0s was $K_L \to \pi^0\pi^0\pi^0$ decaying in flight. A total of 794 fully reconstructed $K_L \to 3\pi^0$ events consistent with two of the intermediate π^0s decaying into $\gamma\gamma$ and one into e^+e^- were collected. An estimated 53.2 ± 11.0 of these events were expected to be background. Normalizing to the π^0 Dalitz decay we found

$$\mathrm{Br}(\pi^0 \to e^+e^-, \quad \left(\frac{m_{e^+e^-}}{m_{\pi^0}}\right)^2 > 0.95) = (6.44 \pm 0.25(\mathrm{stat}) \pm 0.22(\mathrm{sys})) \times 10^{-8} \qquad (1)$$

where internal radiation, $\pi^0 \to e^+e^-(\gamma)$, was limited by the requirement $(m_{e^+e^-}/m_{\pi^0})^2 > 0.95$ which separated it from the tree level Dalitz decay, $\pi^0 \to e^+e^-\gamma$.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

Experimental measurements of neutral pseudoscalar meson decays into lepton pairs, $P \to \ell^+\ell^-$, and the comparison with theoretical predictions offer an interesting way to study long-distance dynamics in the Standard Model. The rarity of these decays, in particular, offer the possibility of probing for new physics interactions.

In the Standard Model $P \to \ell^+\ell^-$ decays have to proceed through a loop diagram since no spinless current coupling of quarks to leptons exists. One way for these decays to happen is through an electromagnetic loop suppressing the decays relative to $P \to \gamma\gamma$ by two orders of the electromagnetic coupling α. Contributions from the weak interaction are many orders of magnitude smaller for $P = \pi^0, \eta$ and can be neglected. For the $K_L \to \mu^+\mu^-$ decay the weak contribution is significant and can be calculated rather reliably. Information about the CKM matrix element $|V_{td}|$ can in principle be extracted from $\mathrm{Br}(K_L \to \mu^+\mu^-)$ if the electromagnetic contribution can be understood well enough and subtracted. Unfortunately, our theoretical understanding of meson form factors is poor, and the intriguing weak physics is hidden from us. Experimental precision measurements of $P \to \ell^+\ell^-$ decays will help guide the theoretical effort in the right direction, and this thesis provides a small part in this endeavor.

The rare decay $\pi^0 \to e^+e^-$ has been studied theoretically in some detail over the years, starting with the first prediction of the rate by Drell[1] in 1958. The decay proceeds through the $\pi^0 \to \gamma^*\gamma^*$ intermediate state and continues through a loop to form the e^+e^- final state as shown in Fig. 1.1. Helicity and two orders of α suppress the decay relative to the $\pi^0 \to \gamma\gamma$ decay,

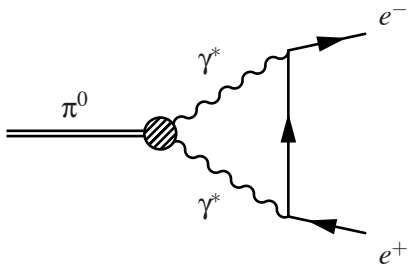

Figure 1.1: Lowest order diagram for $\pi^0 \to e^+e^-$.

leading to an expected branching ratio of about 10^{-7}.

This thesis describes how we measured the $\pi^0 \to e^+e^-$ branching ratio using the KTeV detector at Fermilab. The layout of the thesis is as follows: In the rest of this chapter we explore the $\pi^0 \to e^+e^-$ decay in more depth, looking at the current theoretical and experimental situation concerning the decay. We end the chapter with a short description of the method used to experimentally extract the $\pi^0 \to e^+e^-$ branching ratio. We move on to describe the KTeV detector and the origin of the neutral kaon beam, followed by a description of the data collection system focused on the elements and logic of the trigger system. Chapter 4 explains how the KTeV spectrometer and calorimeter were used to reconstruct events. The KTeV detector Monte Carlo is outlined in Chapter 5 along with the decay generators used for the analysis. Chapter 6 contains the main analysis work, describing the data sets used and the chain of selection cuts used to clean the sample of backgrounds. The chapter continues with a systematic error analysis and ends with the branching ratio calculation. Finally, Chapter 7 has some concluding remarks and puts the result in perspective.

1.1 $\pi^0 \to e^+e^-$ structure and theory

In order to understand better the mathematical structure of the decay we first simplify the picture by assuming a point-like interaction of the π^0 into e^+e^- (Figure 1.2). The pseudoscalar pion dictates a coupling of the form $g\ \gamma^5$ where g is some coupling strength that 'hides' the internal structure of the decay. In this approximation the decay amplitude looks like:

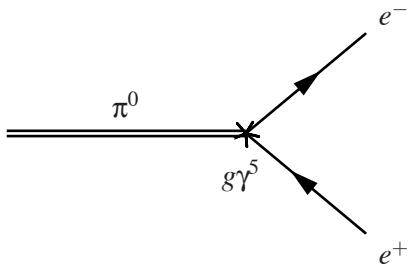

Figure 1.2: Point interaction diagram for $\pi^0 \to e^+e^-$.

$$M = g\bar{u}^s(p_-)\gamma^5 v^r(p_+) \tag{1.1}$$

where p_- and p_+ are the momenta of the outgoing electron and positron and s and r their spins. The squared amplitude becomes (suppressing momenta and spin indices in the first two lines)

$$|M|^2 = |g|^2\,[\bar{u}\gamma^5 v]\,[\bar{u}\gamma^5 v]^\dagger = |g|^2\,[\bar{u}\gamma^5 v]\,[v^\dagger(\gamma^5)^\dagger(u\gamma^0)^\dagger] \tag{1.2a}$$

$$= |g|^2\,[\bar{u}\gamma^5 v]\,[-\bar{v}\gamma^5 u] = -|g|^2\,\bar{u}\gamma^5 v\,\bar{v}\gamma^5 u \tag{1.2b}$$

$$= -|g|^2\,\bar{u}^s_\alpha\,\gamma^5_{\alpha\beta}\,v^r_\beta\,\bar{v}^r_\delta\,\gamma^5_{\delta\varepsilon}\,u^s_\varepsilon = |g|^2\,u^s_\varepsilon\,\bar{u}^s_\alpha\,\gamma^5_{\alpha\beta}\,v^r_\beta\,\bar{v}^r_\delta\,\gamma^5_{\delta\varepsilon} \tag{1.2c}$$

where in the last step we include both spin and spinor indices to be explicit. Doing the spin sums over s and r using the spinor completeness relations we get

$$|M|^2 = -|g|^2\,(\not{p}_- + m_e)_{\varepsilon\alpha}\,\gamma^5_{\alpha\beta}\,(\not{p}_+ - m_e)_{\beta\delta}\,\gamma^5_{\delta\varepsilon} \tag{1.3a}$$

$$= -|g|^2\,\mathrm{Tr}((\not{p}_- + m_e)\gamma^5(\not{p}_+ - m_e)\gamma^5). \tag{1.3b}$$

Using $\mathrm{Tr}(\gamma^5\gamma_\mu\gamma^5) = \mathrm{Tr}(\gamma_\mu) = 0$ and $\mathrm{Tr}(\gamma_\mu\gamma^5\gamma_\nu\gamma^5) = -\mathrm{Tr}(\gamma_\mu\gamma_\nu) = -4g_{\mu\nu}$ we find

$$|M|^2 = -|g|^2\,(-4g_{\mu\nu}p_-^\mu p_+^\nu - m_e^2) = |g|^2\,(4\,p_-\cdot p_+ - m_e^2). \tag{1.4}$$

If we note that

$$m_{\pi^0}^2 = (p_- + p_+)^2 = p_-^2 + p_+^2 + 2\,p_-\cdot p_+ = 2m_e^2 + 2\,p_-\cdot p_+ \tag{1.5}$$

so that $4\;p_-\cdot p_+ = 2m_{\pi^0}^2 - 4m_e^2$, then we get the final expression for the invariant amplitude squared

$$|M|^2 = |g|^2\,(2m_{\pi^0}^2 - 5m_e^2). \tag{1.6}$$

The decay rate is then simply found from

$$\Gamma(\pi^0 \to e^+e^-) \;=\; \frac{|\mathbf{p}_+|}{32\,\pi^2 m_{\pi^0}^2}\int |M|^2 \mathrm{d}\Omega \tag{1.7}$$

The contribution to the rate from the $5m_e^2$ term in M is almost 30,000 times smaller than the $2m_{\pi^0}^2$ term and can safely be neglected. Using $\mathbf{p}_+ = \sqrt{(m_{\pi^0}/2)^2 - m_e^2}$, inserting (1.6) into (1.7) and performing the trivial angular integration, we find

$$\Gamma(\pi^0 \to e^+e^-) \;=\; \frac{|g|^2}{8\pi} m_{\pi^0}\sqrt{1-4m_e^2/m_{\pi^0}^2}. \tag{1.8}$$

The contribution from the two photon intermediate state can be shown to be [2]:

$$g_{\mathrm{EM}} \;=\; -\frac{m_e}{2f_\pi}\left(\frac{\alpha}{\pi}\right)^2 R \tag{1.9}$$

where $f_\pi \approx 92$ MeV is the pion decay constant, and R, the reduced amplitude which contains the remaining dynamics, is

$$R = -\frac{2i}{\pi^2 m_{\pi^0}^2}\int d^4k \frac{q^2k^2-(q\cdot k)^2}{(k^2+i\varepsilon)((q-k)^2+i\varepsilon)((k-p)^2-m_e^2+i\varepsilon)}\, F(k^2,(q-k)^2). \tag{1.10}$$

where $p = p_+$, q is the pion four-momentum, and F is the $\pi^0\gamma^*\gamma^*$ form factor which for real photons is F(0,0)=1. Through the Adler-Bell-Jackiw anomaly, f_π is related to the $\pi^0 \to \gamma\gamma$ rate, see e.g. [3],

$$\Gamma(\pi^0 \to \gamma\gamma) \;=\; \frac{\alpha^2}{64\pi^3}\frac{m_{\pi^0}^3}{f_\pi^2}. \tag{1.11}$$

so when we take the ratio of rates

$$\frac{\Gamma(\pi^0 \to e^+e^-)}{\Gamma(\pi^0 \to \gamma\gamma)} \;=\; 2\sqrt{1-\left(\frac{2m_e}{m_{\pi^0}}\right)^2}\left(\frac{\alpha}{\pi}\frac{m_e}{m_{\pi^0}}\right)^2 |R|^2 \tag{1.12}$$

f_π cancels out. To evaluate R a model for the form factor F is needed, but we can still say something about $|R|^2$. The contribution from real photons, where $F = 1$, comes entirely from the model independent imaginary part of R,

$$\mathrm{Im}(R) \;=\; \frac{\pi}{2\beta_0}\ln\left(\frac{1-\beta_0}{1+\beta_0}\right), \quad \beta_0 = \sqrt{1-4m_e^2/m_{\pi^0}^2} \tag{1.13}$$

and since $|R|^2 \geq \mathrm{Im}(R)^2$ we get the *unitary bound* on the ratio of rates [4]:

$$\frac{\Gamma(\pi^0 \to e^+e^-)}{\Gamma(\pi^0 \to \gamma\gamma)} \geq 4.75 \times 10^{-8} \tag{1.14}$$

The constant value for F is however problematic for the real part of R which diverges logarithmically in that case. Any sensible form factor must therefore have some cutoff Λ beyond which it must fall off fast enough to cancel the divergence.

1.1.1 Helicity suppression

As already mentioned $\pi^0 \to e^+e^-$ is not only suppressed by two orders of the electromagnetic coupling but is also suppressed by the approximate helicity conservation of the interaction. The factor $(m_e/m_{\pi^0})^2$ in Equation (1.12) is less mysterious in the light of the following arguments: In the massless or high energy limit the electromagnetic interaction only couples left-handed fermions with right-handed anti-fermions or vice versa. For the back-to-back electron and positron in $\pi^0 \to e^+e^-$ this implies that their spins along the decay direction should be aligned in a $S = 1$ state. However, the neutral pion is spinless and therefore the electron and positron must have a total angular momentum of zero. Any orbital angular momentum in the final state is perpendicular to the decay direction, so the electron and positron spin projections on this axis must be zero. The two arguments conflict with each other and while angular momentum conservation is an exact symmetry helicity conservation is not for massive particles. Helicity is a good symmetry in the highly relativistic limit, but it is broken here at one of the vertices giving a suppression factor of $(1 - v/c) \cong 2m_e^2/m_{\pi^0}^2$, v being the velocity of the electron in the pion rest frame.

1.2 Form factor models and ChPT

A number of phenomenological models for the form factor have been put forward; among the most successful are the *Vector-Meson Dominance* (VMD) models. These models assume that a vector-meson intermediate state for the two photons, $\pi^0 \to VV \to \gamma\gamma$, will dominate the

decay amplitude. The form factor typically takes the form

$$F(q_1,q_2) \;=\; \frac{M_V^2}{M_V^2-q_1^2}\frac{M_V^2}{M_V^2-q_2^2} \tag{1.15}$$

where only the lightest vector mesons are relevant, $V=(\rho,\omega)$, so $M_V = 770-780$ MeV/c^2. M_V in this way provides the cutoff that make Re(R) finite. Various predictions have been made using VMD models [2][5] [6][7], all giving very consistent result for the branching ratio around $6.2-6.4\times10^{-8}$ which is about $1.3-1.4\times$ the unitary limit. Other approaches to the form factor give similar results, and as was noted by many authors the branching ratio seems to be rather insensitive to the details of the form factor model.

The phenomenological models for the form factor are all somehow ad hoc, lacking substantial theoretical support. Chiral Perturbation Theory (ChPT) provides such support. ChPT is a low energy effective theory where the pions (and eta's) are the massless Goldstone bosons resulting from the chiral symmetry breaking, and they are the fundamental degrees of freedom. An effective Lagrangian, the chiral Lagrangian, can be used to reevaluate the decay amplitude in this framework. The theory is nonrenormalizable and needs an infinite number of counterterms to balance divergences appearing when calculating loop corrections. These counterterms have to be fixed somehow to get predictions for the real part of R. At the lowest order in the chiral expansion all pseudoscalar to two leptons ($P\to\ell^+\ell^-$) branching ratios are described by the same reduced amplitude the real part being [8]:

$$\mathrm{Re}(R)(P\to\ell^+\ell^-) \;=\; \frac{1}{4\beta_0}\ln^2\left(\frac{1-\beta_0}{1+\beta_0}\right)+\frac{1}{\beta_0}\mathrm{Li}_2\left(\frac{\beta_0-1}{\beta_0+1}\right)+\frac{\pi^2}{12\beta_0}+3\ln\left(\frac{m_l}{\mu}\right)+\chi(\mu) \tag{1.16}$$

In this expression $\mathrm{Li}_2(z)=\int_z^0\frac{\ln(1-t)}{t}\mathrm{d}t$ is the di-logarithm, $\beta_0=\sqrt{1-4m_\ell^2/M_P^2}$, and $\chi(\mu)$ is the sum of the relevant counterterms at the subtraction point μ. In [9] and [8] the counterterms are fixed using measured values for the branching ratio $\eta\to\mu^+\mu^-$, and predict $\pi^0\to e^+e^-$ branching ratios of $(7\pm1)\times10^{-8}$ and $(8.3\pm0.4)\times10^{-8}$, respectively. The latter of the two used a more recent measurement for $\eta\to\mu^+\mu^-$ and should be considered the current ChPT prediction using experimental input.

The counter terms have also been estimated independent of experimental data in [10] where a lowest meson dominance approximation to large-N_c QCD was used, giving $\mathrm{Br}(\pi^0 \to e^+e^-) = (6.2 \pm 0.3) \times 10^{-8}$. This result is very similar to those from the VMD form factor models, and in this light, as pointed out in [6], the older, naive VMD models get some theoretical justification through ChPT.

1.3 Radiative corrections

Radiative corrections are important ingredients to consider in a precision measurement of $\pi^0 \to e^+e^-$. Inner bremsstrahlung gives the decay the same final state as the Dalitz decay, $\pi^0 \to e^+e^-\gamma$, and the only reason we can distinguish the two modes is because the dynamics are quite different. The Dalitz decay predominantly has a hard photon and low e^+e^- invariant mass, while the radiation from $\pi^0 \to e^+e^-$ is very soft with a high e^+e^- invariant mass. Figure 1.3 shows the distribution of the invariant e^+e^- mass for the two modes. The Dalitz mode is concentrated in the region close to $m_{ee} = 0$ and the *ee* mode near the π^0 mass. Even with this difference, about 10% of the $\pi^0 \to e^+e^-$ decays are drowned by the Dalitz decays. One can think of the two modes as two amplitudes contributing to the same final state, $e^+e^-\gamma$. Only in the regions where the interference between the two modes is small can we distinguish the two contributions.

Radiative corrections to $\pi^0 \to e^+e^-$ are treated in a paper by Bergström [11]. First order corrections are calculated assuming a model independent point-like $\pi^0\gamma^*\gamma^*$ vertex like the one introduced in Section 1.1. The virtual rate correction diagram (Figure 1.5) and bremsstrahlung corrections (Figure 1.4) are calculated as corrections to the lowest order rate, $\Gamma = \Gamma^0 + \Gamma^{\mathrm{int}} + \Gamma^{\mathrm{brem}}$. The point-like vertex assumption implicitly neglect radiation off the internal electron. Bergström argues that this is a good assumption since the loop momenta that dominate the rate are close to the pion mass leaving the electron far off-shell. Bremsstrahlung is only prominent when the particle is close to its mass shell, so this radiation contribution can safely be neglected. The interference between the Dalitz mode and the radiative $\pi^0 \to e^+e^-$ mode is discussed too

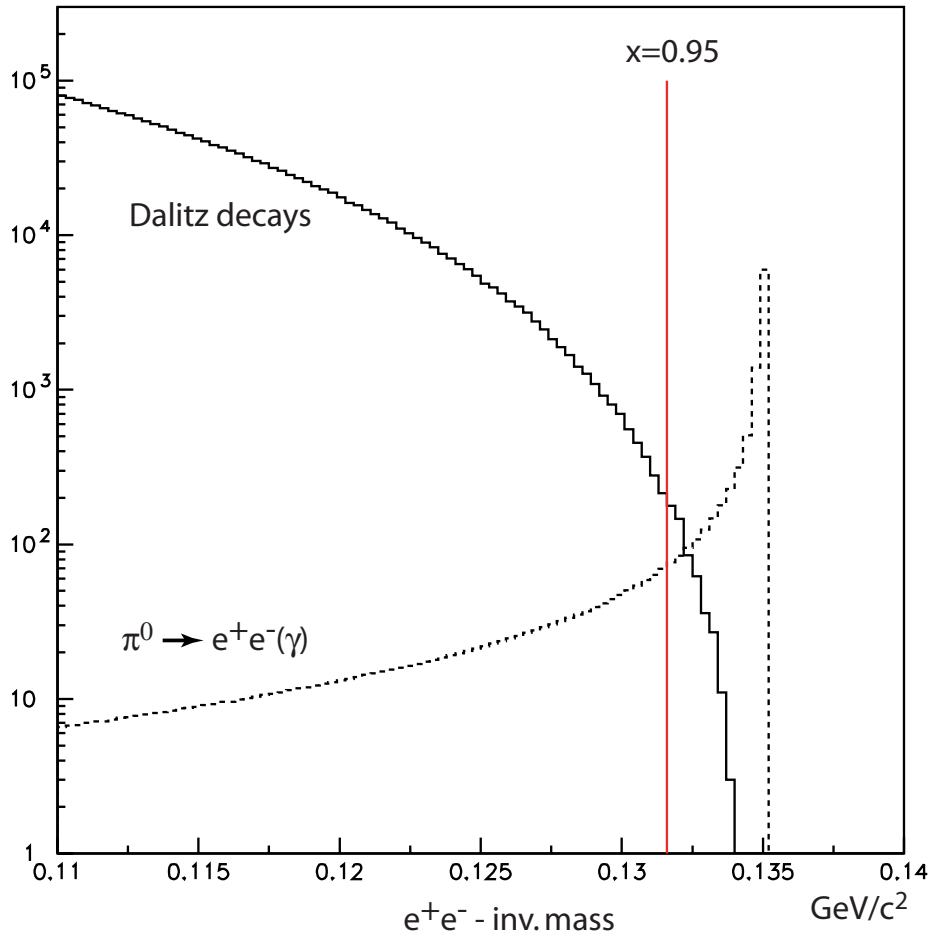

Figure 1.3: e^+e^- mass distribution for the Dalitz decay and the $\pi^0 \rightarrow e^+e^-+$ inner bremsstrahlung as the MC generator created them. The red line shows where the limit on the e^+e^--mass for signal is set.

and is shown to be small in the region of interest.

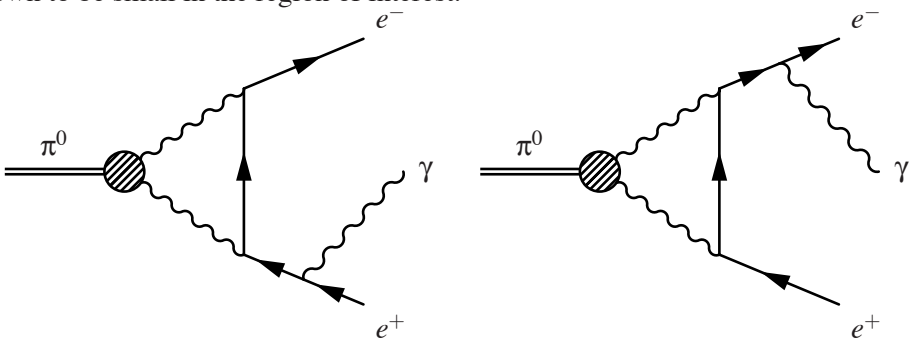

Figure 1.4: Feynman diagrams for the bremsstrahlung contribution to the radiative corrections.

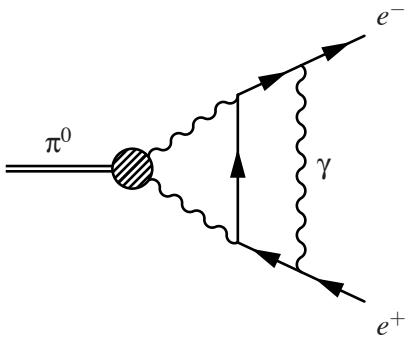

Figure 1.5: Feynman diagram for the virtual contribution to the radiative corrections.

1.3.1 Internal bremsstrahlung correction

From an experimental point of view we are most interested in the corrections from the bremsstrahlung diagrams, Figure 1.4. $\pi^0 \to e^+e^-$ events could only be accepted in a narrow region around the pion mass to avoid backgrounds, so we needed to model the rate and energy spectrum of the radiated photon in our Monte Carlo in order to accurately determine the signal acceptance. The inner bremsstrahlung correction to the differential rate provides us with this information. From Bergström's paper we find:

$$\frac{d\Gamma^{\text{brem}}}{dx} = \Gamma^0_{ee}\,\frac{\alpha}{\pi}\left\{\frac{x^2+1}{1-x}\ln\left(\frac{1+\beta}{1-\beta}\right) - \frac{2x\beta}{1-x}\right\} + O(\alpha^2) \tag{1.17}$$

where $x = \dfrac{(p_+ + p_-)^2}{m^2_{\pi^0}}$, and $\beta = \sqrt{1 - \dfrac{x_{\min}}{x}}$ with $x_{\min} = \dfrac{4m_e^2}{m^2_{\pi^0}}$. To get the rate of radiation we need to integrate Equation (1.17). This can only be done if we introduce a cutoff on x or equivalently on the photon energy, since the $1/(1-x)$ behavior makes the integral diverge logarithmically as $x \to 1$. At $x \gtrsim 0.98$ higher order radiative corrections are likely to play a role in the radiation spectrum [12], therefore a cutoff much higher than $x = 0.98$ is not very meaningful. This will be discussed further in the Monte Carlo decay generation section, Section 5.2.

1.3.2 Virtual correction and total correction

The virtual correction diagram, Figure 1.5, fixes the divergence problems from the bremsstrahlung diagrams. The same logarithmic divergence appears but with opposite sign

so as to cancel when the total correction is formed [11] :

$$\frac{\Gamma^{\text{brem}}+\Gamma^{\text{virt}}}{\Gamma^0_{ee}} = \frac{\alpha}{\pi}\left\{\frac{3}{2}\ln\left(\frac{1-\beta_0}{1+\beta_0}\right)+\frac{9}{4}\right\}+O\left(\frac{m_e^2}{m_{\pi^0}^2}\right) \tag{1.18}$$

Here $\beta_0 = \sqrt{1-x_{\text{min}}}$. Numerically the total correction to the lowest level rate is -3.4%.

1.3.3 Interference with the Dalitz decay

The quantum mechanical interference between the radiative $\pi^0 \to e^+e^-$ decay and the Dalitz decay should be included in an exact treatment of the two modes. The interference is found to be

$$\frac{1}{\Gamma^0_{e^+e^-}}\frac{d\Gamma^{\text{int}}}{dx} = \frac{-2\,\text{Re}(R)}{m_\pi^2\,|R|^2}\,\frac{(1-x)^2}{x}\ln\left(\frac{1+\beta}{1-\beta}\right) \tag{1.19}$$

which depends on the the unknown real part of R. Bergström uses this result to derive an upper limit for the interference

$$\left|\frac{1}{\Gamma^0_{e^+e^-}}\frac{d\Gamma^{\text{int}}}{dx}\right| \;\leq\; \frac{2}{\pi\ln\left(\dfrac{1+\beta_0}{1-\beta_0}\right)}\,\frac{(1-x)^2}{x}\ln\left(\frac{1+\beta}{1-\beta}\right). \tag{1.20}$$

Equality is reached when $|\text{Im}(R)| = |\text{Re}(R)|$ which as we saw in Section 1.2 is not very far from the case in most models. Integrating this upper bound on the full range of x from x_{min} to 1 gives the maximum of the absolute value of the total interference compared to the lowest order rate. This maximum is actually ~ 2.6, so the interference can be several times larger than the lowest order contribution! However most of the contribution to the integrated interference comes from low x, and in the region of interest for this measurement, $x = 0.9-1.0$, the integral is only 0.02% and therefore negligible.

1.4 Previous $\pi^0 \to e^+e^-$ measurements

The first experimental evidence for the $\pi^0 \to e^+e^-$ decay was found in 1978 by the Geneva-Saclay group [13], where about 6 events consistent with $K^+ \to \pi^+\pi^0$, $\pi^0 \to e^+e^-$ were observed. A branching ratio of $\text{Br}(\pi^0 \to e^+e^-) = (22^{+24}_{-11}) \times 10^{-8}$ was reported, a result almost 5 times higher than the unitary limit suggesting a possibly large contribution from

Re(R). A second measurement in 1983 by a LAMPF group [14] reported a branching ratio of Br($\pi^0 \to e^+e^-$)=$(17 \pm 6 \pm 3) \times 10^{-8}$ based on 59 observed events from π^0's produced through $\pi^- p \to \pi^0 n$. Another indication that the rate could be many times the unitary limit. At the time many models for the form factor had been proposed, but none of them predicted values for the branching ratio above $7 \times 10^{-8} = (1.5\times$ unitary limit), something that still holds true today. The potential for new physics contributing to Re(R) was tantalizing but no firm conclusion could be drawn due to the large experimental errors. New experiments were initiated to tackle this problem. In 1987 the OMICRON group at CERN [15] published an upper limit of Br($\pi^0 \to e^+e^-$)$<$ 53×10^{-8} which didn't change the situation, but two years later the SINDRUM collaboration [16] put a much tighter bound on the branching ratio at Br($\pi^0 \to e^+e^-$)$< 13 \times 10^{-8}$. This result agreed better with the assumption of only an electromagnetic contribution to the rate but still left a little room for speculation. In 1993 two more precise measurements were published back to back. E851 at Brookhaven [17] had seen 21 events with $\pi^0 \to e^+e^-$ decays from $K_L^+ \to \pi^+\pi^0$ resulting in Br($\pi^0 \to e^+e^-$)=$(8.0 \pm 2.6 \pm 0.6) \times 10^{-8}$, while E799-I at Fermilab [18], the predecessor of KTeV E799-II, saw 9 events and found Br($\pi^0 \to e^+e^-$)= $(8.8^{+4.5}_{-3.2} \pm 0.6) \times 10^{-8}$. Finally in 1999 KTeV E799-II [19], using the first round of data taking (the '97 dataset), published Br($\pi^0 \to e^+e^-$)=$(7.04 \pm 0.46 \pm 0.28) \times 10^{-8}$ with 275 events of which 21.4 were expected to be background. This result agreed well with most theoretical predictions and essentially shattered the remaining hope that new physics could be found in this decay.

1.5 Experimental method

A short overview of the experimental approach will be presented here - it should give the reader a sense of what's to come.

We measured the ratio of branching ratios Br($\pi^0 \to e^+e^-$)/Br($\pi^0 \to e^+e^-\gamma$) using π^0's from in flight $K_L \to 3\pi^0$ decays. The two modes will be referred to as the *signal* and *normalization* modes. The $K_L \to 3\pi^0$ decays provided a large number of π^0s owing to the large branching ratio of the $3\pi^0$ mode ($\sim 20\%$) and the fact that any of the 3 π^0's could decay to e^+e^-. The

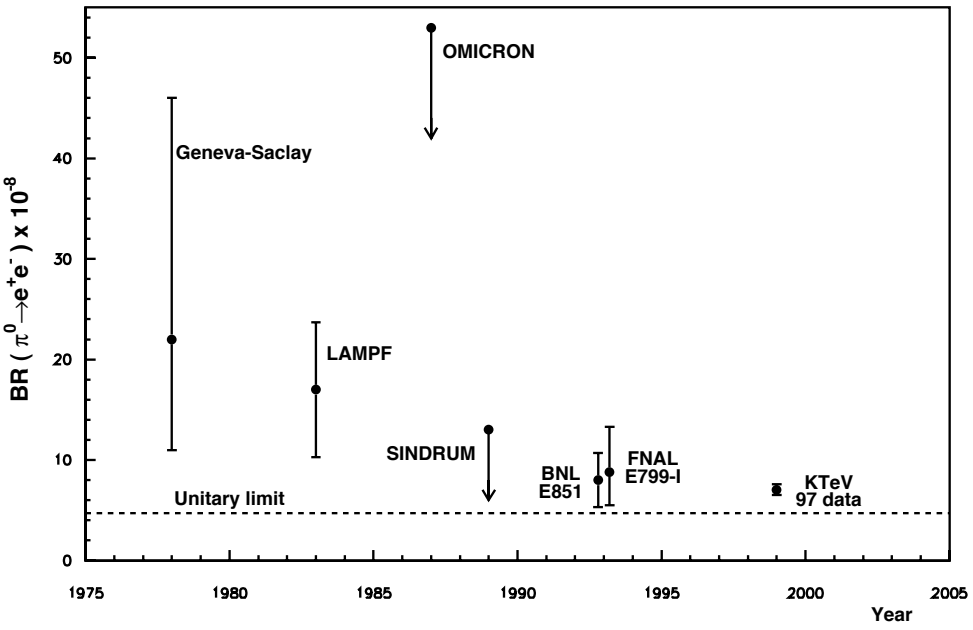

Figure 1.6: Previously measured values and upper bounds on lowest order Br($\pi^0 \to e^+e^-$). The dotted line represents the unitary limit on the branching ratio.

main advantage of the $3\pi^0$ mother decay was the large number of kinematic constraints available when fully reconstructing the $K_L \to 3\pi^0$, $\pi^0 \to e^+e^-$ decay chain. Backgrounds from other K_L decays were eliminated completely which made the analysis relatively clean. It should be noted that the continuum decay $K_L \to \pi^0\pi^0 e^+e^-$, which has the same final state as our signal, has a measured upper limit on the branching ratio of 6.6×10^{-9} [20]. In the narrow accepted region of $m_{e^+e^-}$ for our signal the background from $K_L \to \pi^0\pi^0 e^+e^-$ was negligible.

The signal mode and the normalization mode were reconstructed and analyzed in parallel from the same data sample. The similarities in modes caused the measured ratio to be insensitive to most of the detector systematics, like absolute tracking inefficiencies, absolute trigger inefficiencies and so on. The main unknown that was canceled in the ratio was the number of kaons decaying in the detector. Of course this "flux" of kaons could in principle be deduced from the analysis of any well known clean decay mode, but it turns out that the number was rather dependent on which mode was used. Absolute detector efficiencies were hard to model exactly, so a normalization mode that was topologically very similar to the signal and one that

could be reconstructed in a similar way was used.

For both modes the full K_L decay chain was reconstructed. The signal chain was

$$K_L \to \pi^0\pi^0\pi^0 \to \gamma\gamma\,\gamma\gamma\,e^+e^-$$

and the normalization chain was

$$K_L \to \pi^0\pi^0\pi^0 \to \gamma\gamma\,\gamma\gamma\,e^+e^-\gamma$$

The signature in the detector for the signal was six energy deposits in the electromagnetic calorimeter and two tracks in the charged spectrometer. For the normalization the only difference was an extra deposit of energy in the calorimeter from the extra photon.

The Dalitz normalization sample after basic event selection cuts was essentially background free and made the perfect testbed for studying the Monte Carlo simulation of the experiment. It also provided us with a tool to study possible systematic errors which was otherwise impossible with the limited statistics of the signal.

Chapter 2

The beamline and the KTeV detector setup

KTeV is the name of the detector that provided the means for the measurement described in this thesis. KTeV stands for Kaons at the Tevatron indicating that it was a Kaon experiment that ran at the Tevatron at Fermilab (Fermi National Accelerator Laboratory).

The KTeV detector had two modes of operation known by their experiment number, E799-II and E832. E832 was the main experiment designed to measure $\mathrm{Re}(\varepsilon'/\varepsilon)$. This was done by regenerating K_S's in one of two incoming beams just upstream of the decay region, resulting in a beam of K_L's and a beam of K_S's. The E799-II experiment was designed to search for the CP-violating mode $K_L \to \pi^0 e^+ e^-$, but it also had the capability to measure many other rare K_L, hyperon and pion decays. The E799-II mode was the experiment that provided data for the $\pi^0 \to e^+ e^-$ branching ratio measurement presented here. In this configuration there was no regenerator so two identical beams of K_L's were used. The only other major difference in the detector setup was the introduction of a set of Transition Radiation Detectors (TRDs) for E799-II to provide better pion-electron distinction.

Data taking for E799 was done in 3 different time periods. In 1997 there were two periods of run time known as the *winter* and *summer* runs, or collectively the *97* runs. After a number of upgrades to the detector another run was done in 1999-2000, known as the *99* run.

In this chapter the beamline and the KTeV detector components will be discussed.

2.1 The beamline

The beam of K_L's was generated by aiming the Tevatron's 800 GeV proton beam at a beryllium oxide target. The interaction in the target created a multitude of particles from which the kaons were ingeniously extracted using a series of sweeping magnets, absorbers, and collimators.

2.1.1 The proton beam

The Tevatron at Fermilab was the source of the high energy protons delivered to the Neutrino-Muon (NM) fixed target beamline in which the KTeV detector was located. The 800 GeV protons extracted from the Tevatron hit the target in *spills* of 20 s in 97 and 40 s in 99. Between spills there was an *off-spill* period of 40 s in which the Tevatron accelerated a new set of protons. During a spill the protons came in bunches 1-2 ns long with a period of 19 ns. This same frequency was used to synchronize the different detector components at KTeV.

The number of protons hitting the target in a spill was typically $2-5\times10^{12}$ during the 97 runs and was doubled (the spill was longer) for the 99 run. The instantaneous intensity remained approximately the same in all runs.

The beam was focused to a width less than $250\mu m$ as it entered the NM2 beamline enclosure and hit the 3mm$\times$3mm$\times$30cm BeO target. For the protons this presented 1.1 interaction lengths and was the production point of a wide variety of particles some of which were the desired neutral kaons.

2.1.2 The kaon beam

In order to extract the kaons from the interaction products and define the beam, NM2 was instrumented with a series of sweeper magnets, absorbers and collimators. Figure 2.1.2 shows the NM2 enclosure elements. The KTeV coordinate system also had its origin in NM2, at the center of the target. The positive z-axis extended downstream from the target towards the KTeV

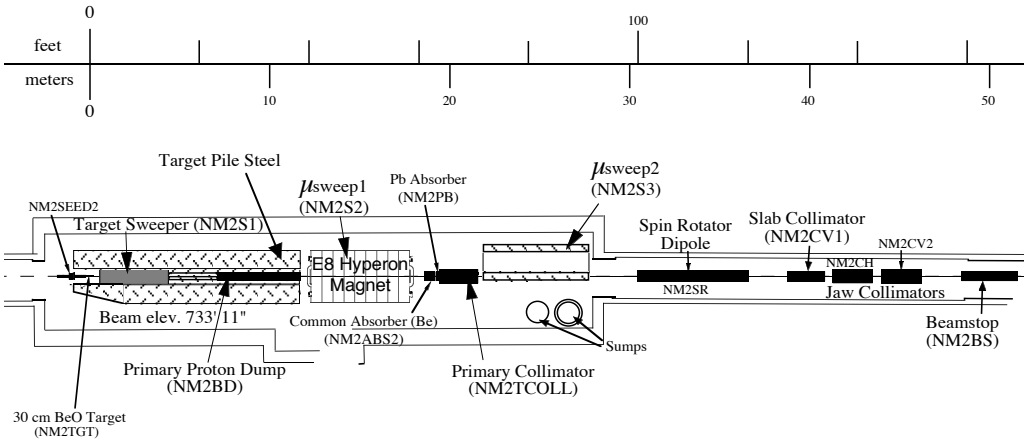

Figure 2.1: Apart from the target NM2 was instrumented with a series of sweeper magnets, absorbers and collimators.

detector, positive y was up, and positive x was to the left when looking downstream, creating a right handed coordinate system.

Right behind the target sat the first sweeper magnet, the *target sweeper*. It extended from $z = 0.6 - 4.4$ m giving charged particles a 475 MeV/c kick. The kick deflected the remaining protons from the beam down into the *primary proton dump* sitting below the beam line.

Next in the line starting at 12.3 m were two additional sweeper magnets, *μsweep1* and *μsweep2*, with a 3 inch lead wall absorber between them. The magnets delivered kicks of 3806 MeV/c and 1854 MeV/c (3135 MeV/c in Win97). The lead absorber made photons convert into e^+e^- pairs which were swept out in the second of the two magnets.

The *primary collimator*, also between μsweep1 and μsweep2, was the first beam defining component. It was a brass block with two rectangular holes in it, measuring 1.18 cm$\times$1.29 cm in the winter period and enlarged to 1.62 cm $\times$ 1.73 cm for the summer and 99 runs.

Further downstream at 33 m was the *spin rotator dipole*. Its purpose was to manipulate the polarization state of neutral hyperons. This component had no effect on the result presented here.

For the winter run a steel *slab collimator* was in place between the two beams at 38.8 m. It was meant to prevent crossovers between the two beams. It was found to have little or no

effect and was taken out for the later runs.

As the beam left the NM2 enclosure at ~ 50 m it entered a vacuum region which extended all the way through to 159 m. The vacuum was kept at $\sim 10^{-6}$ Torr.

At 85 m the *defining collimator* was used to define the edges of the two beams, giving them their final shape before they would enter the decay region. The beam holes here were 4.4 cm $\times$ 4.4 cm each for the winter runs and 5.2cm $\times$ 5.2 cm in the later runs.

Just after the defining collimator at 90m-93m a last sweeper magnet kicked out charged particles from upstream decays in the beam or from interactions in the last collimator.

At this point the beams were in the fiducial decay region where decays could be reliably reconstructed. The "purified" beam mainly consisted of two components, neutrons and K_L's, with about 3 times as many neutrons as kaons. The neutrons almost never decayed in the detector and as such did not pose a problem, they did however interact in the detector and could not be neglected. A small component of K_S's, Λ^0's and Ξ^0's were also present. The rate of K_L's at this point in the beam was about 50MHz of which close to 5% would decay in the vacuum region. The detector was sensitive to decays of kaons with energy above 20 GeV, and the distribution of decaying kaon energies can be seen in Fig. 2.2.

2.1.3 Accidental counter

Also in the beamline was the *accidental counter*. In order to simulate accidental activity in the detector a special purpose trigger was used to collect data uncorrelated with kaon decays in the detector. Three scintillator counters sat behind small holes in the shielding of the target. They were placed at a 90 degree angle and triggered on coincidences in all three detectors. A trigger collected all the activity in the detector and these "events" were used as accidental overlays in the Monte Carlo detector simulation, see also Section 5.4.3.

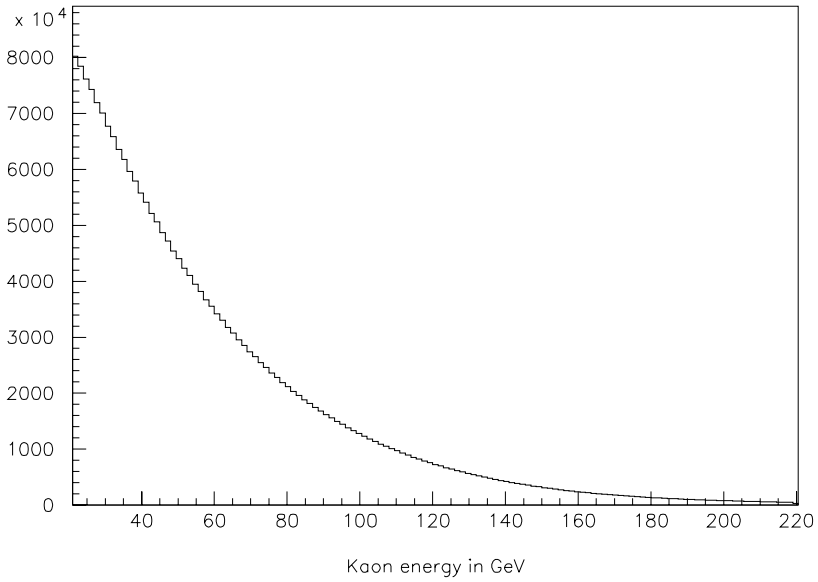

Figure 2.2: Energy distribution of decaying kaons in decay region.

2.2 The KTeV detector

Figure 2.3 and 2.4 shows the general layout of the E799 detector in two different views. In this section the important components will be described.

2.2.1 Vacuum tank and vacuum window

A 69m evacuated cylindrical tank defined the fiducial region for kaon decays. It opened up right after the defining collimator and extended downstream to the beginning of the drift chambers. On the inside it was instrumented with a set of photon vetos to reject high angle photons from decays that would otherwise escape unnoticed. These vetoes and other like ones are described in Section 2.2.6.

At the downstream end of the vacuum region was the *vacuum window*. This thin window of Kevlar laminated with mylar preserved the vacuum while decay particles could pass through the 0.16% of a radiation length of material.

Right behind the vacuum window was a volume, *helium bag 1a*, supposedly filled with

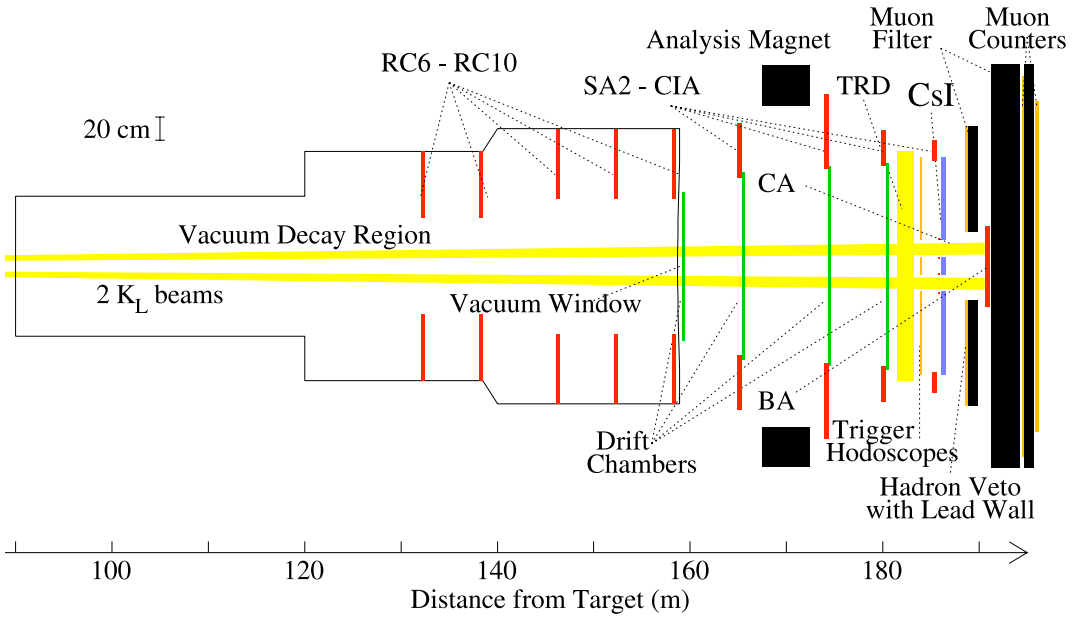

Figure 2.3: Layout of the KTeV detector in the E799-II configuration.

helium to reduce scattering and conversions. Later measurements showed that in the 97 runs it had leaked and was filled mostly with air. For the 99 run the bag was replenished with helium which stayed there for the duration of the runs. At the end of this $\sim$ 25cm volume there was a 10cm air gap separated by 1 mil of Mylar. The air gap was there to allow a massive safety shutter to be put in place in front of the vacuum window when access to the experimental hall was needed. Another sheet of 1 mil Mylar and another helium bag *helium bag 1b*, this one with actual helium in it in both 97 and 99, was placed right in front of the first drift chamber.

2.2.2 Spectrometer

The spectrometer consisted of four drift chambers (DC1-4) and a large dipole magnet imparting a kick in the x direction to charged particles. Two drift chambers were used to find tracks upstream of the magnet and the last two chambers found tracks after the magnet bent them. Helium bags were placed between the drift chambers to reduce multiple scattering and photon conversion in the spectrometer. The magnet kick was 205 MeV/c in the 97 runs and was

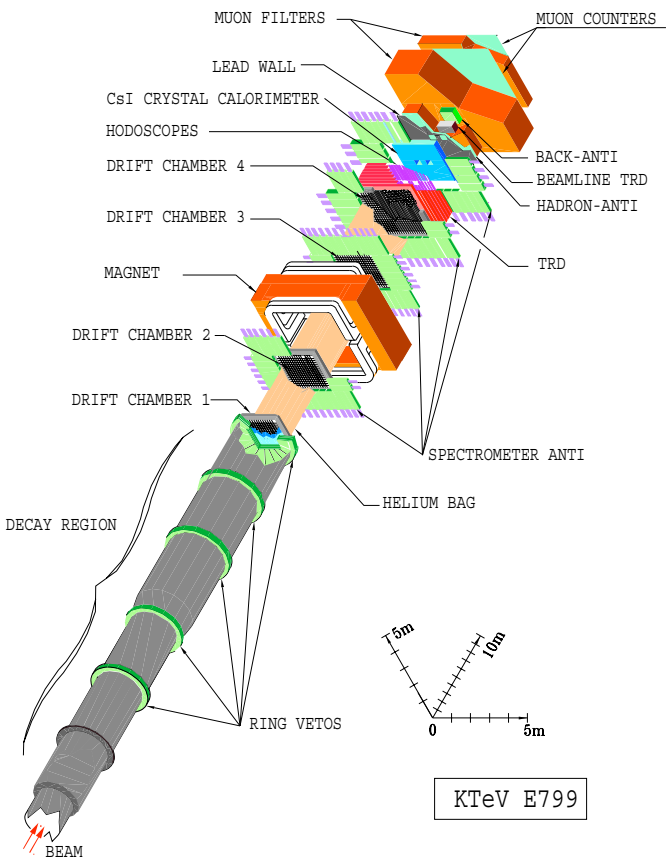

Figure 2.4: The KTeV detector in an alternative view.

lowered to 150 MeV/c for the 99 runs to increase acceptance for lower momentum tracks. The polarity of the magnet was periodically flipped to decrease possible systematic errors.

2.2.2.1 Drift chambers

The drift chambers had been used in previous experiments (E731, E773, E799-I) and apart from using new electronics they were merely restrung with new field and sense wires. Each chamber had two "views". A set of wires were strung parallel to the x axis and a set parallel to the y axis. This made the chambers capable of measuring the position of tracks in both the x view and the y view.

In each view, two layers of 100 μm gold-plated aluminum sense wires were used, all

surrounded by 25μ gold-plated tungsten field wires kept at a higher potential ($\sim$ 2500 V) to create a well for all negatively charged particles. The sense wires in the layers were staggered and the spacing between them was 12.5 mm.

When a high energy charged particle came through it would ionize the surrounding argon-ethane drift chamber gas. The electrons knocked loose saw the potential created between the field and sense wires and would drift toward the sense wires. The current pulse on the sense wire would fire the attached time-to-digital converter (TDC) registering a hit. The maximum drift times were less than 150 ns and the track position measurement had a resolution of 0.1 mm. In Fig. 2.5 an illustration of a charged particle going through one of the views in a chamber is

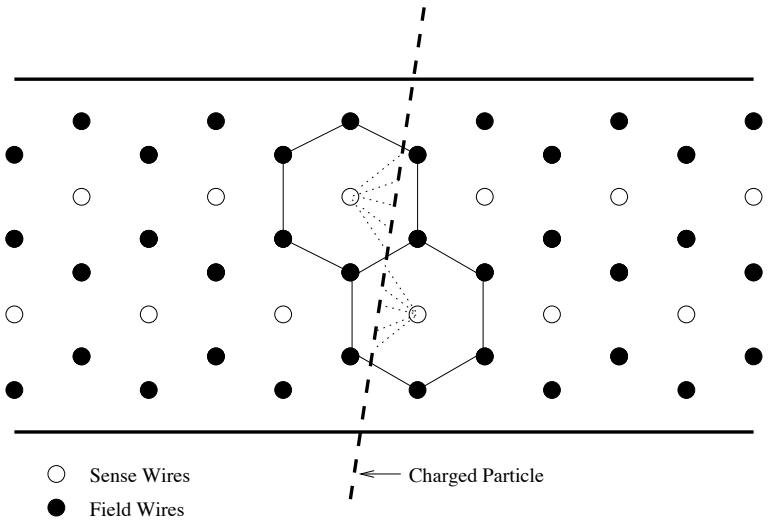

Figure 2.5: Cartoon of the sense wire/field wire configuration in a chamber view and an illustration of a typical ionization pattern.

shown. The two layers of sense wires were needed to know on which side of the wire the track went.

In its final calibrated state the spectrometer was able to measure momenta with a resolution given by:

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

with P measured in GeV/c. The constant term came mostly from scattering in the chambers and

the term linear in P was due to the smaller bend of high momentum tracks.

2.2.3 Transition radiation detector

The Transition Radiation Detector (TRD) was positioned behind DC4 at 181.1 m and was a little over 2 m long. It consisted of 8 multi wire proportional chambers (MWPCs) with a polypropylene mat in front of it. When a charged particle passed through the mat, transition radiation with energy inversely proportional to the particle mass was radiated. The energy of this x-ray radiation was measured in the MWPC effectively measuring the particle mass. This information could be used for particle ID, in particular to discriminate between electrons and pions. For a detailed description on the TRD see [21].

For the purpose of the result presented here the TRD was just a chunk of material sitting in front of the calorimeter causing tracks to scatter a bit and some photons to convert.

2.2.4 Trigger hodoscope

Between the TRD and the calorimeter, at 183.9m, two planes each made from 32 scintillating paddles were installed. This *trigger hodoscope* provided a very fast but rough indication of charged tracks in the detector, which was used in the trigger. The two planes, V and V' as they were called, were each a centimeter thick and had complete coverage of the front of the calorimeter except for two holes where the beam went through.

2.2.5 The calorimeter

The electromagnetic calorimeter was a large array of CsI crystals sitting at 186m, each crystal equipped with a photo multiplier tube (PMT) to measure the energy deposited in it from showering particles. The array consisted of 3100 crystals of two different sizes arranged as shown in Fig. 2.6. In the middle region, where most particles hit, the smallest crystals were used to increase resolution, they measured 2.5cm$\times$2.5cm$\times$50cm and were arranged to allow the two beams to pass through two 15cm$\times$15cm holes. The larger crystals surrounding them

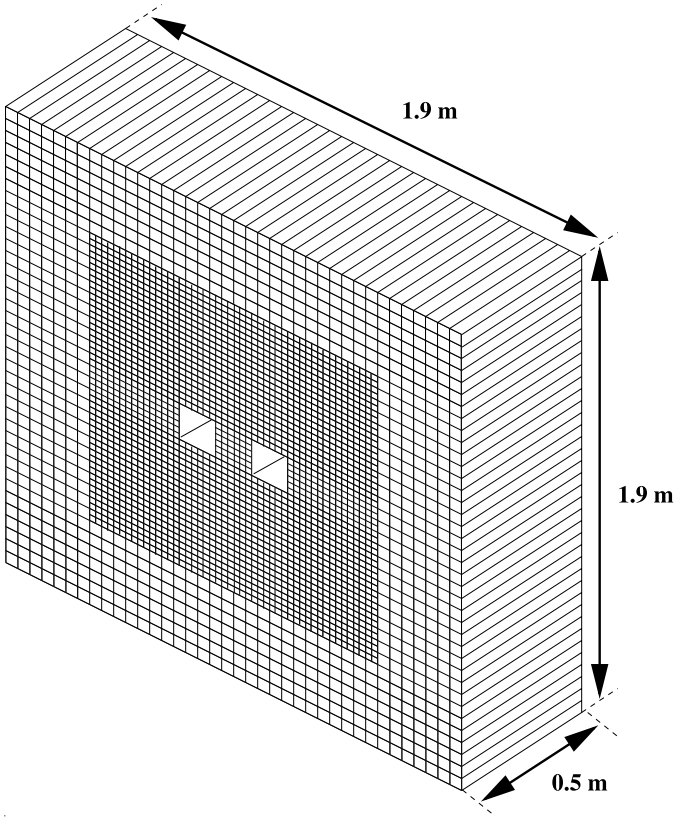

Figure 2.6: The CsI calorimeter layout.

were 5cm×5cm×50cm giving a complete array size of 1.9m×1.9m×0.5m.

The depth of the calorimeter corresponded to 27 radiation lengths causing almost every photon and electron to deposit all its energy. For other common particles like the muon and the pion the situation was different. Because of their higher masses their electromagnetic interaction through ionization was weak, leaving typically only 0.3 GeV in the calorimeter (minimum ionization). The pions also interacted strongly and with 1.4 nuclear interaction lengths of CsI the pions showered hadronically about 75% of the time but normally still only leaving a fraction of their energy.

Synchronous with the Tevatron's RF, every 19 ns, the energy deposited in each crystal would be read out and kept in a first in, first out buffer of length 32. In this way information about energy depositions in a large interval around any event was known in slices of 19 ns.

After calibration and offline corrections to cluster energies (Section 4.2.4) the calorimeter had a final energy resolution given by:

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

with energies measured in GeV. The first term came from a number of sources such as noise and non-uniformities. The photostatistics of the measured scintillation light depended on the energy deposit which gave rise to the second term.

2.2.6 Photon veto counters

The detector had a number of veto counters in place to reject events with photons escaping at high angles. Five *ring counters*, RC6-RC10, were positioned inside the vacuum decay tank. RC6 and RC7 had an inner aperture size of 0.84m×0.84m, while RC8-RC10 were 1.18m×1.18m. Each ring counter had 24 lead-scintillator layers equaling 16 radiation lengths, enough to stop most photons. Shown on the left in Fig. 2.7 is a typical RC configuration.

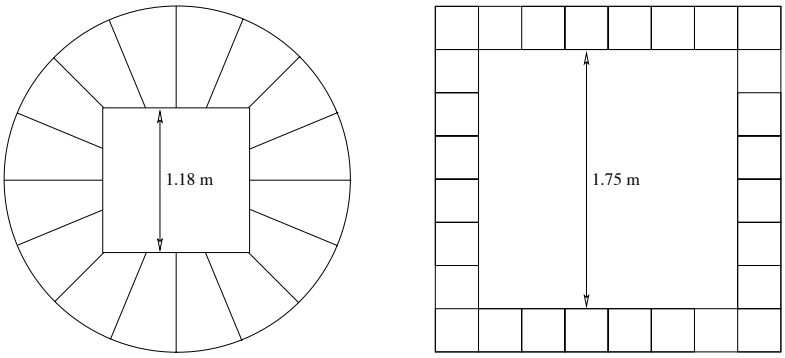

Figure 2.7: Schematic of typical Ring counters and Spectrometer antis. Shown is RC10 (left) and SA4 (right).

Further downstream around the drift chambers were the *spectrometer antis*. There was one around DC2, DC3, and DC4 named SA2, SA3, and SA4 respectively. Figure 2.7 (on the right) shows the SA4 configuration, typical for its kind. SA4 measured 1.75m×1.75m on the inside while SA2 and SA3 were somewhat smaller measuring 1.54m×1.37m and 1.69m×1.60m.

The SAs were constructed similarly to the RCs using 32 layers of lead and scintillator each 0.5 radiation lengths for a total of 16 radiation lengths.

2.2.7 Other veto counters

A number of other veto counters were used. The *collar anti* (CA) was two counters sitting right in front of the calorimeter beam holes covering parts of the inner crystals as shown in Fig. 2.8. Particles that hit close to the edge of the calorimeter could easily have energy leaking out the sides where it would be lost. Such events were undesirable and the CA's 9.7 radiation lengths of tungsten-scintillator was in place to veto such events near the beam holes. The counter was 3cm wide, the inner size being 15cm$\times$15cm.

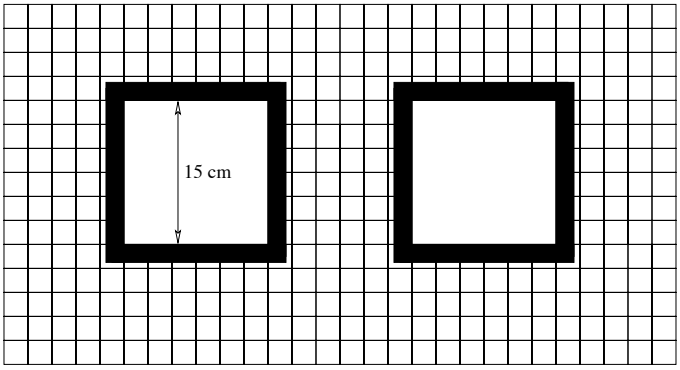

Figure 2.8: Schematic of how the Collar anti was placed around the beam holes on the calorimeter overlapping the inner CsI crystals.

Around the outside edge in front of the calorimeter was the *Cesium Iodide Anti* (CIA) which was almost identical to the SAs. It measured 1.84m$\times$1.84 on the inside and so shadowed part of the calorimeter outer edge.

Behind the calorimeter and behind a 15cm lead wall was the *hadron anti* (HA). The HA was a set of 28 scintillator paddles used to veto hadron showers. In the lead wall electromagnetic showers leaking through the back of the calorimeter were stopped, so the HA could be used to veto on showering hadrons only. The HA was covered from behind with a 1m steel wall to remove back splash from the beam dump further downstream.

A component that wasn't used in this analysis was the *back anti* (BA). This 32 layer lead-scintillator detector sat in the beam behind the 1m thick steel wall. A 60cm$\times$30cm hole in the steel allowed the beam to pass. It was there to veto particles going down the beam holes, however high rates made it almost useless for most analyses.

2.2.8 The muon system

Close to the end of the detector was a 3m steel wall used as a beam dump effectively stopping most particles except muons. Behind the wall was a plane of 56 scintillator paddles called *MU2* which was used in certain triggers to veto muons in the final state. MU2 was only used in the trigger for the first part of the winter runs for this analysis.

Behind MU2 was another lead wall, this one 1m thick, which had two sets of scintillator paddles, MU3X and MU3Y, behind it. These were used for more muon track information and were not used in this analysis.

Chapter 3

The trigger and data acquisition

When beam was being delivered to the KTeV detector the kaon decay rate was about 1 MHz in the fiducial decay region, which was too fast for the Data Acquisition (DAQ) system to record everything. The trigger system was responsible for reducing the data rate to a more manageable amount while still keeping most of the interesting kaon decays. This was done by triggering on certain coincidences of detector activity.

The selection of events was done in 3 stages, trigger level 1, 2, and 3. The first two trigger levels were done in hardware logic while the third level was done in software. The first level triggering was done continuously at 19 ns intervals with very simple logic on fast sources. If the Level 1 trigger fired, the more complicated logic of the Level 2 trigger set in using as long as 2.5μs to process, during which the Level 2 trigger did not accept Level 1 triggers. A successful Level 2 trigger would initiate the Level 3 software trigger which did simple tracking and clustering to further reduce the data rate. An event that passed all levels of triggering was written to tape for offline analysis.

Two categories of trigger definitions were used, beam triggers and calibration triggers. There were 16 different triggers in each category. The 16 beam triggers were the ones tuned to find interesting physical processes, while the calibration triggers were used to record samples of well understood decay modes which could be used to calibrate the various detector elements. The first beam trigger, *trigger 1* or the *2E-NCLUS* trigger, was used for the $\pi^0 \to e^+e^-$ measurement and it will be the focus of this chapter. A more thorough description

of the KTeV trigger system can be found in [22].

3.1 The Level 1 trigger

The Level 1 trigger (L1 trigger) was clocked with the Tevatron RF and looked for simple coincidences in the fastest detector components every 19 ns. These fast *sources* were all sent to a set of programmable logic modules through cables matched in length so that the sources were compared for the same event. This was necessary since for a particle moving at the speed of light it took several hundred nanoseconds to travel the length of the detector. For example, for a decay losing a photon which hit RC6 the rest of the decay products would hit the calorimeter, 53.6 meters further downstream, 178 ns later than the photon that hit RC6.

The Level 1 trigger accepted about a tenth of the kaon decays resulting in a Level 1 trigger rate less than 100kHz.

As an example of which trigger sources were used we look at 2E-NCLUS which was used for both signal and normalization in this analysis. The logic of the Level 1 trigger requirements for the 2E-NCLUS trigger were defined in the following way :

$$\begin{aligned} \text{TRIG1[2E-NCLUS]}_{\text{L1}} = \text{GATE} \wedge \text{2V} \wedge \text{ET\_THR3} \wedge \text{DC12} \wedge (\neg\text{MU2}) \wedge (\neg\text{HA}) \\ \wedge (\neg\text{RC\_OR}) \wedge (\neg\text{SA\_OR}) \wedge (\neg\text{CIA}) \wedge (\neg\text{CA}) \end{aligned} \tag{3.1}$$

where $\neg$ is logical 'not' and $\wedge$ is logical 'and'. The individual components are described in the following.

3.1.1 GATE

The *GATE* source was a combination of two conditions. It was set when the beam was on-spill and when beam was not sent to the adjacent beamline of the E815 experiment. The last was to avoid leakage from the E815 beamline into the KTeV detector which could cause high detector activity for short amounts of time.

3.1.2 2V

The trigger hodoscopes were used to give a prompt trigger on charged tracks in the detector. The *2V* source was set when there was evidence for two tracks, which was defined to be the case when two or more paddles in one plane were hit and one or more in the other plane.

3.1.3 ET_THR3

The Etotal system had a number of sources including *ET_THR3*. The system summed up the energies in all the 3100 CsI crystals of the calorimeter and returned four sources corresponding to four different thresholds, ET_THR1-4. In the 97 runs these thresholds were nominally held at 10, 18, 25 and 38 GeV and were changed slightly to 11, 16, 25 and 38 GeV for the 99 runs.

The Etotal system also set a bit for each calorimeter channel that had an in-time energy above a nominal 1GeV. These channels were used later in trigger level 2 as seeds for the hardware cluster count (HCC).

3.1.4 DC12

The two upstream chambers, DC1 and DC2, were instrumented with a fast triggering system called the drift chamber ORs. In each view (X and Y) the wires were grouped in groups of 16 called “paddles”. Within a paddle the logical OR of the wires were made to see if any were hit. If a wire in a paddle was hit the paddle was counted as hit. *DC12* required a hit paddle in every view of the two chambers.

The timing resolution of this source was not as good as the other Level 1 sources due to the long drift times ($\lesssim 150ns$). Since a track would normally cause two hits in a view, one with short drift time and one with a longer time, it was possible to set the DC OR sources to use a 90 ns window to trigger on.

3.1.5 Veto sources

A source from each of the photon veto counters (the RCs, the SAs and the CIA) was used in the Level 1 trigger. The *RC_OR* was set if any one of the RC counters registered an energy higher than 0.5 GeV. Similarly, the *SA_OR* was set if any of the SAs had an energy higher than 0.4 GeV. The *CIA* was set if the CIA had more than 0.4 GeV.

Another set of veto sources used the Collar Anti, the Hadron Anti, and the Muon system. The *CA* source got set if any one of the two collar antis (east or west) registered more than 14 GeV of energy, rejecting events with un-reconstructible hits near the beam holes. Events with hadrons in the final state could be vetoed using the *HA*. It required energy equivalent to 2.5 MIPs (Minimum Ionizing Particles), although for some runs in the winter the requirement was 7 MIPs. In the first part of the winter runs (8088-8576) the muon system was used in the 2E-NCLUS trigger to reject events with muons. The *MU2* source was set if one or more hits were registered in MU2. The MU3 banks were not used in trigger 1.

3.2 The Level 2 trigger

A Level 1 trigger started the Level 2 (L2) trigger processing. Programmable logic boards were employed to do a more complex hit counting in the chambers and to do cluster counting in the calorimeter. Simple track finding logic was also used, but not in the 2E-NCLUS trigger. Another factor of 10 reduction of the data volume was achieved by the L2 trigger logic. The rate of L2 triggers was about 10kHz.

The L2 requirements to the 2E-NCLUS trigger had the following logical form :

$$\text{TRIG1[2E-NCLUS]}_{\text{L2}} \;=\; \text{HCC\_GE4} \;\wedge\; \text{2HCY\_LOOSE} \;\wedge\; \text{1HC2X} \tag{3.2}$$

the 3 primitives here came out of the HCC cluster count and the chamber hit counting.

3.2.1 Hardware cluster count

Using the HCC bits set at Level 1 the cluster counter would trace out the boundary of the cluster to find the extent of each [23]; a count of separate clusters was the result. The HCC was the slowest processor at Level 2, taking upward of 2.5μs, which set the upper bound on the dead time due to processing at Level 2.

The output of the HCC was the cluster count between 1 and 7 or higher. The 2E-NCLUS requirement, *HCC_GE4*, was that four or more clusters were found.

3.2.2 Drift chamber hit counting

The Level 2 trigger used a more sophisticated way of counting hits in the drift chambers than was used at Level 1. Two types of processors were used, called the *Kumquats* and the *Bananas*. The two processor boards differed in which hits they considered as part of the event (in-time). Both boards counted the number of adjacent hits in each view in a single chamber. Single hits were counted as 1 hit and any number of adjacent hits counted as that number minus one, a pair was one count, a triplet was two, etc.

The Kumquats had the simplest logic, they simply counted hits on adjacent wires in a 205 ns window started by the L1 trigger. The Bananas were more clever in determining if hits were in-time or not. The time correlation between two hits in a pair of wires was considered and only a certain pattern was accepted. The pattern used is shown in Fig. 3.1 where pairs are plotted with in-time hits from good two track events on the left, and on the right with hits that were rejected by the bananas themselves.

The clever design of the bananas was only used in chamber 1 and 2 and only in the X view. Banana boards were installed in the Y views also, but the pattern used was square thereby not really differing from the Kumquats.

The requirement *2HCY_LOOSE* used in the 2E-NCLUS trigger needed at least two hits in each Y view except in DC1 and DC2 where one hit was allowed in one of them. *1HC2X*

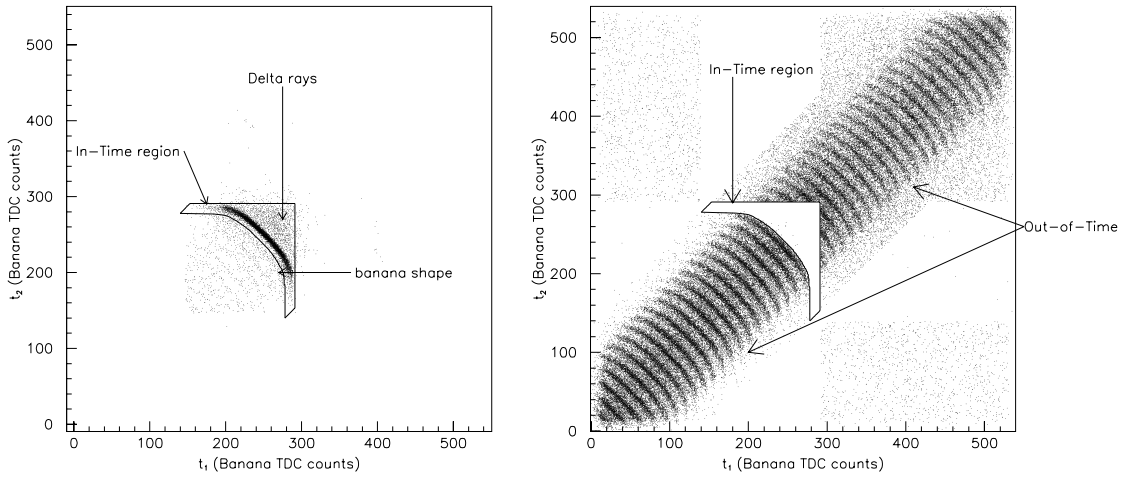

Figure 3.1: t_1 vs. t_2 plots of (on the left) hits from tracks in good two track events. The right hand plot shows how the pattern of hit times look for rejected events. The stealth bomber looking region is the in-time region used.

required one hit in the banana modules installed on the X view of chamber 3.

3.3 The Level 3 filter code

Events that passed both level 1 and 2 were digitized and run through the Level 3 filter code. Four multi-processor SGI computers (200MHz and 150MHz) did this processing which consisted of a slightly simplified version of the offline reconstruction which included tracking, clustering, and vertex finding. These general reconstruction methods are outlined in Chapter 4. Events could then be accepted based on the number of calorimeter clusters and tracks, but also on invariant masses, vertex positions and so on. A number of different *L3 tags* could be set on an event depending on which requirements it satisfied. Events that were assigned a tag by the filter code were written to tape where they were kept ready for offline analysis.

All data from a spill was buffered in 4.5 GB buffers (4.9GB for the 99 runs) before being processed. This allowed for the time between spills to be used for processing and tape writing.

The L3 trigger accepted events at a rate of 1 kHz. Over the ~ 20 sec spill this was ~ 20000 events per spill.

At this point the 2E-NCLUS trigger earned its name. It required two electron candidate tracks which formed a vertex and 4 or more clusters in the calorimeter.

Chapter 4

General event reconstruction

This chapter will describe how the raw data in terms of TDC and ADC counts were transformed into meaningful information about the particles involved in the decay.

The general reconstruction code accomplished mainly two tasks: charged track reconstruction and cluster energy and position reconstruction. The combined information was then used to find a common vertex where the particles originated.

The tracks provided information about the momenta and trajectory of charged particles while the calorimeter clusters gave information on photon energies and positions, but it also gave electron-pion distinction in the form of a cross check measurement of electron energies (E/p = 1). Typically, tracks and clusters were all that were needed to reconstruct the decay and the 4-momenta of the decay particles.

The procedures of tracking, clustering and vertexing is first reviewed. The special case of reconstructing the signal and normalization for this analysis, which required a neutral vertexing method, is discussed at the end of the chapter.

4.1 Tracking

The tracking algorithm went through a number of steps to complete. First chamber hits were paired in each view to form *hit pairs* which were the building blocks for the tracks. Tracks were then looked for in the Y view where there was no bend caused by the magnet. Then track segments from DC1 to DC2 and from DC3 to DC4 were found in the X view. Coincidences of

these segments could then be looked for to form complete tracks in X. Finally, vertex candidates were found by extrapolating back the X and Y tracks into the decay region and by requiring that they had a (loosely) consistent intersection in Z. A more in depth explanation follows.

4.1.1 Hit pairs and SODs

The information recorded in the drift chambers was hit times of ionization electrons that reached the sense wires. These times were translated into drift distances, using first a wire dependent time offset to get the drift time and then the drift velocity to give the distance. The time offset was found using a lookup table. The tracking code used only the first in-time hit on a wire, possible later hits were simply discarded.

The next step was to find pairs of hits on adjacent complementary wires. For each pair of hits their sum of drift distances could be calculated. This *sum of distances* (SOD), would for a perfectly detected track going straight through the chamber, be equal to the sense wire spacing of 6.35 mm; of course this was rarely the case for actual hits. SODs in DC1 and DC2 that were within 1 mm of the cell spacing were considered "good" SODs, in DC3 and DC4 the cut was at 1.5 mm. Out of time tracks typically had bad SODs so rejecting tracks with bad SODs were required. Even for perfectly good tracks, SODs could be low or high. Low SODs could be caused by interference with other particles passing the two wires at the same time leaving only two hits one from each track. Defects and noise in the wires could cause trouble too. Failure to detect the first ionization electrons would result in high SODs and even single hits if a hit was not detected at all. Figure 4.1 shows an illustration of the these examples. The single hits had a left-right ambiguity problem as indicated in the figure but were still used in the tracking. A plot of SOD distributions from Ke3 ($K_L \to \pi^{\pm} e^{\pm} \nu_e$) events is shown in Fig. 4.2.

4.1.2 Track candidates in the *y* view

Connecting the hit pairs was started by looking for tracks in the *y* view. The *y* view tracks had no bend from the magnet so only straight lines were looked for. For each hit pair in DC1

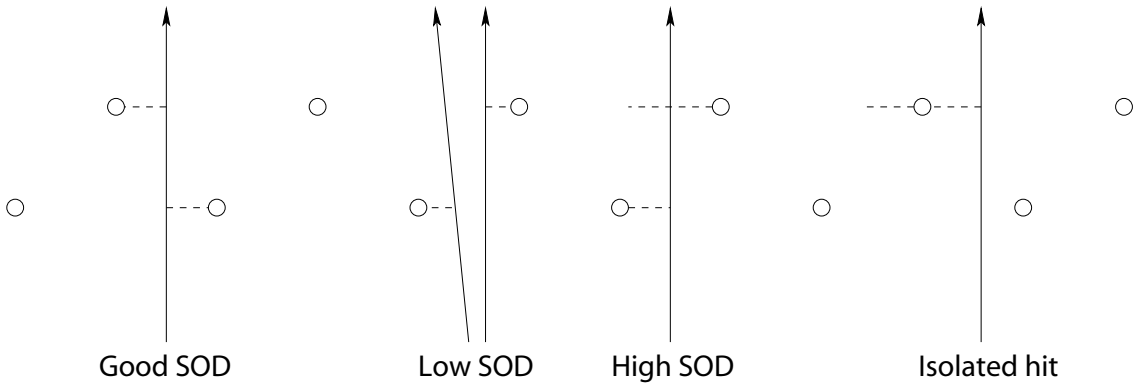

Figure 4.1: Examples of SOD types. Circles represents sense wires in a plane and the arrows tracks. Dashed lines are the inferred distances.

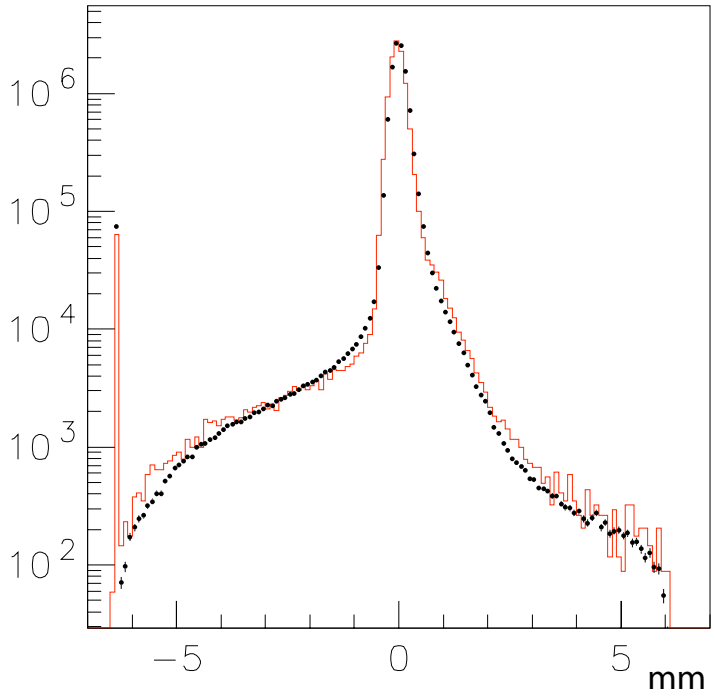

Figure 4.2: SOD distribution in Ke3 data (black) and Monte Carlo (red). The difference from the nominal wire spacing is plotted. The peak at -6.35 is from isolated hits.

and DC4 hits were looked for in DC2 and DC3 that lay close to the line joining them. Of the four hit pairs two of them were allowed to be bad SODs, or one bad SOD and a single hit. For each possible track formed from hits like this a χ^2 was calculated using the 100 μm position resolution for the hits. All track candidates with a sufficiently low χ^2 were kept and used later for matching with track candidates in the *x* view.

4.1.3 Track candidates in the *x* view

The search for track candidates in the *x* view was different. Tracks were bent at the magnet, so the procedure was to find segments in the upstream chambers and downstream chambers separately and then try to match segments up at the magnet mid plane. For the upstream segments two bad SODs were acceptable while the downstream segments were only allowed one bad SOD or one isolated hit. All combinations of an upstream and a downstream segment were considered together and if they matched to within 6 mm at the magnet mid plane and together had a maximum of two bad SODs, or one bad SOD and an isolated hit, then this *x* view track candidate was kept for later matching with the *y* view track candidates.

4.1.4 Vertex candidates

At this point one could look for vertex candidates, meaning two tracks in each view that where loosely consistent with a common intersection in Z. Every combination of four tracks like this was extrapolated back and checked for an intersection in the decay region. For the final complete tracks and vertex to be identified, energy clusters in the calorimeter needed to be found.

4.2 Cluster identification

When a particle hit the face of the calorimeter it would deposit energy in a whole cluster of CsI crystals. Figure 4.3 shows a display of a typical energy distribution in the calorimeter from a $\pi^0 \to e^+e^-$ event. The algorithm used to find these clusters and to accurately determine

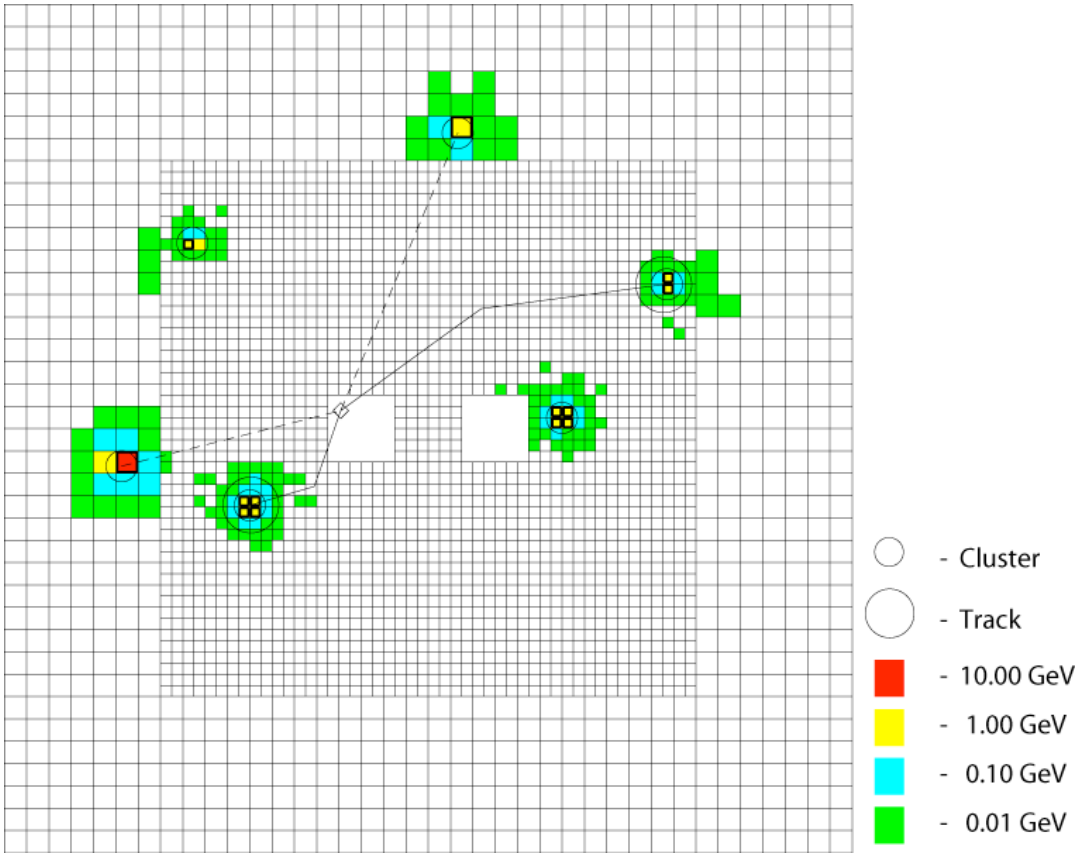

Figure 4.3: Graphical view of a $\pi^0 \rightarrow e^+e^-$ event in the calorimeter. The solid line between the two clusters indicate the projection of the tracks.

their energies is the subject of this section.

First the energy of each crystal was found by summing the PMT readouts from multiple 19 ns time slices. Starting from the in-time slice, four slices were used in the two 97 run periods while six were used in 99. The total was then converted to an energy using calibration constants measured for each crystal in special calibration runs.

4.2.1 Hardware clusters

The main cluster finding method was based on the HCC bits set in the L1 trigger. A hardware cluster *seed* was defined as a local energy maximum of the crystals that had its HCC

bit set. From each seed a cluster was formed and its raw energy calculated. The cluster was the square around the seed, 3x3 for the large blocks and 7x7 for the small blocks, and the energies in these blocks were summed to give the raw cluster energy. At the boundary between large and small blocks the large block method was used with 2x2 small blocks counting as a large block.

Hardware clusters were used for both the signal and normalization analysis.

4.2.2 Software clusters

In order to look for low energy clusters where no HCC bit was set a different method could be employed. This method was called software clustering since it didn't use the HCC information and was done exclusively in software. The algorithm searched the calorimeter for any local energy maxima excluding the ones with HCC bits set. These maxima were used as *software seeds* and clusters were formed around them similarly to the hardware clusters. A software cluster had needed a seed energy above 50 MeV and a total raw energy above 100 MeV to be kept.

4.2.3 Cluster positions

The position of each cluster could be determined quite accurately. The x and y positions were found separately by considering the ratio of energies in the seed block and the neighboring blocks in the x and y directions. A table was created from data to look up the position given these ratios and the cluster energy. The position resolution was $\sim$ 1 mm for clusters in the small blocks and $\sim$ 1.8 mm in the large blocks.

4.2.4 Cluster energy corrections

The raw summed up energy in the clusters often suffered problems from the naivety of the approach. A lot of these problems were addressed in an energy correction process. The corrections involved are summarized here.

- *Overlap* correction: Clusters that overlapped had to somehow share the energy in the overlapping blocks. This was accomplished by splitting the energy in the overlapping crystals based on the energies and positions of the two clusters in an iterative process. The cluster positions and energies were recalculated and the energy splitting was performed again until the iteration converged, or a maximum of 20 iterations had been done.

- *Neighbor* correction: Some clusters would deposit energy outside the 3x3 or 7x7 region and this energy might end up in a neighboring but not overlapping cluster. Using GEANT generated shower profiles on 13x13 arrays of small blocks the estimated amount of energy under other clusters were subtracted.

- *Missing block* correction: Clusters where the cluster region went off the edge of the calorimeter had a correction applied that added back in this missing energy. This was done using the same shower profiles that the neighbor correction used.

- *Sneaky energy* correction: If a cluster was close to the beam holes, energy could escape across the beam hole and add energy to clusters bordering the other side. The missing energy correction took care of the cluster that lost energy but the cluster across the hole had the sneaky energy correction applied subtracting off the excess energy. The sneaky energy was estimated by studying calibration Ke3 events where the electron hit close to the beam hole and the shower splashed across to the other side.

- *Threshold* correction: Energy readouts from low energy blocks were zeroed out if a given threshold energy was not met. This could lower the energies of some clusters that had low energy blocks in their region. From special runs with no block energy zeroing this effect could be estimated and energy added to the affected clusters.

- *Intra-block* correction: The CsI crystals responded differently to the same energy deposit depending on where in the crystal the particle showered. Near the edges of the

crystal the scintillation response was somewhat lower. The cluster energy was corrected with a multiplicative factor depending on where on the seed block the incoming particle struck. The effect was mapped out in 25 transverse bins by measuring E/p for calibration Ke3's in each bin for every crystal.

At a later stage in the analysis the MC was found to have a better calorimeter resolution than data. To fix this, a smearing fudge was employed in the MC to match the data resolution. The energy in each CsI crystal was smeared according to

$$E_{\text{crystal}} \rightarrow E'_{\text{crystal}} = E_{\text{crystal}}\ (1 + 0.0065 \cdot r) \tag{4.1}$$

where r is a Gaussian distributed random number centered on 0 with $\sigma = 1$. This fudge was applied just before the clusters were identified and formed. This energy smearing is discussed in more detail in Section 6.8.9.

4.3 Vertexing

Having found clusters in the calorimeter, tracks in the x and y views could now be matched without ambiguity by pointing the tracks at the clusters. From the vertex candidates found earlier each track was projected to the calorimeter face and matching with each cluster was attempted. A minimum track-cluster separation of < 7 cm was a match.

The vertex candidates where both tracks were matched to clusters went through some reprocessing where the tracks were corrected for effects like alignment of the chambers and fringe fields from the magnet. After these final corrections the best vertex candidate had to be selected. This was done by measuring how well the upstream and downstream tracks matched at the magnet bend plane, through the *offmag* χ^2, and by measuring how well the tracks matched at the vertex, through the *vertex* χ^2. These two χ^2's in combination with the number of bad SODs used in the tracks formed an overall quality variable of which the candidate with the lowest value was selected as the final vertex.

4.4 Neutral vertexing

Clusters with no matching tracks were considered photon clusters. Photon clusters were paired assuming they all came from a $K_L \to \pi^0\pi^0\pi^0$ decay chain where 2 of the 3 π^0's decayed to two photons and the last was the decay of interest. For the signal, there were 4 photon clusters that could get paired in 3 possible ways. In the normalization there was an extra photon, the Dalitz photon, and a total of 15 combinations to form the 2 π^0's with a photon left unpaired.

For each pair of photons a vertex z position could be calculated assuming they came from a π^0. For $\pi^0 \to \gamma_1\gamma_2$ with photon momenta p_1 and p_2 we have:

$$m_{\pi^0}^2 = (E_1+E_2)^2 - (\mathbf{p_1}+\mathbf{p_2})^2 \tag{4.2a}$$

$$= 2E_1E_2 - 2\mathbf{p_1}\cdot\mathbf{p_2} \tag{4.2b}$$

$$= 2E_1E_2(1-\cos\theta_{12}) \tag{4.2c}$$

The distance d between the two photon clusters and the distance to the vertex r is related to the angle θ between the photons, $r\tan\theta = d$. Since all angles are less than about 0.1 we have $r\,\theta \simeq d$ and $(1-\cos\theta) \simeq \theta^2/2$. Inserting into Equation (4.2c) we find

$$r = \frac{d\sqrt{E_1E_2}}{m_{\pi^0}} \tag{4.3}$$

For the two photon $Z_{\gamma\gamma}$ vertex we get

$$Z_{\gamma\gamma} = Z_{\mathrm{CsI}} - \frac{d\sqrt{E_1E_2}}{m_{\pi^0}} \tag{4.4}$$

The position in z of the calorimeter, Z_{CsI}, was the position of the face of the calorimeter plus the mean shower depth.

For each of the possible pairings a χ^2 was formed and minimized in order to find the best pairing. The χ^2 measured how well the two neutral vertices coincided with each other. For our purpose the e^+e^- pair should also have the same vertex z position so we form the combined *pairing* χ^2:

$$\chi^2 = \left(\frac{Z_{\gamma\gamma}-\bar{Z}}{\sigma_{\gamma\gamma}}\right)^2 + \left(\frac{Z'_{\gamma\gamma}-\bar{Z}}{\sigma'_{\gamma\gamma}}\right)^2 + \left(\frac{Z_{ee}-\bar{Z}}{\sigma_{ee}}\right)^2 \tag{4.5}$$

with a χ^2 term for each π^0. The σ's are the estimated uncertainties on the respective vertex positions. The pairing with the lowest minimized pairing χ^2 was picked and the averaged vertex z position, $\bar{Z}$, that minimizes the pairing χ^2 was used for the kinematics of the decay. The vertex x and y positions were then refitted with the constraint that the z position was fixed at $\bar{Z}$.

This procedure optimizes the momentum and vertex resolutions since it uses all the measured quantities instead of just the charged information. It has a small side effect that will be discussed in Section 6.8.1.

Chapter 5

The Monte Carlo simulation

Simulating the experiment was the single most important part of the analysis. This was where all our knowledge of the experimental setup and all the known physical interactions were input. The simulation was of a statistical nature hence the name, Monte Carlo simulation. Millions of single decays were simulated, but individually a simulated decay did not provide much useful information. The simulation only provided information on distributions of variables. This was, of course, the only thing that made sense anyway, since all the interesting physics happened through randomly distributed quantum mechanical interactions.

The Monte Carlo was used to find the detector acceptance for all decays so the total number of decays of a certain type could be inferred from how many were detected. Since the typical acceptance was around 1%, the extrapolation to the total number of decays was large and the simulation therefore had to be very accurate.

The Monte Carlo also provided useful information when trying to come up with a useful technique for reconstructing decays and eliminating backgrounds. It was the one tool to use when any issue regarding the experimental setup needed to be addressed.

For a given decay the Monte Carlo traced through a number of steps: The kaon production, the decay generation, particle tracing through the detector, accidental activity, detector response and digitization, and the trigger selection process. The output of the Monte Carlo was in the same format as the data taken from the experiment itself. Each Monte Carlo event had some extra data tagged to it though. Information only known in the simulation was included

such as the real decay vertex, and the particle momenta of each particle in the decay chain.

KTeV has a long chain of Monte Carlo/analysis versions with upgrades and bug fixes in each version. The package `KTeVANA` version 6.04 was used for all Monte Carlo generation (and general analysis).

5.1 Kaon production

The momentum and direction of initial kaons created in the target was modeled using the Malensek spectrum [24]. This spectrum, with parameters measured from charged kaon production, was adapted based on consideration of the quark composition. The spectrum was parametrized as follows,

$$\frac{\mathrm{d}^2N}{\mathrm{d}p\mathrm{d}\Omega} = Kp\,\frac{(1-x)^A(1+5e^{-Dx})}{(1+p_T^2/M^2)^4}. \tag{5.1}$$

A, D, K, and M were parameters extracted from the charged kaon data and x was the kaon momentum divided by the proton momentum. The parameters were tweaked to match the observed kaon momentum spectrum in $K_{L,S} \to \pi^+\pi^-$ decays.

Given the momentum and the direction, the kaon was traced from the target to the fiducial decay region. The process would start over if the kaon did not have a momentum in the range 20-220 GeV/c or if it interacted in the beamline and didn't make it to the decay region. The momentum range was chosen in this way since the detector acceptance was almost zero outside that range.

A kaon that made it to the decay region was 'decayed' at random on the interval z=90m-160m according to its decay probability per time. The distribution of the z-position of the decay vertex was almost uniform in this range. Only 5% of the K_L's decayed in the decay region, so the number of decays only dropped by that amount over the decay range.

5.2 Decay generators

With a kaon decaying at a given point with a given momentum the next step was to generate the decay itself. In the following the decay generators important to the analysis are described. All decays were generated in the parent rest-frame, and the decay products were then boosted into the lab frame.

5.2.1 $K_L \to 3\pi^0$

All decays used for the measurement in this thesis have the same starting point, namely the $K_L \to \pi^0\pi^0\pi^0$ decay.

This 3 body decay was generated using an almost pure phase space distribution, with a small correction from the form factor measured at E731 [25] and NA48 [26].

The three π^0s subsequently decayed into the specified final state, e.g. for the signal two $\pi^0 \to \gamma\gamma$ and the last $\pi^0 \to e^+e^-$ using the generator described below. The π^0 decays were performed instantly after the kaon decay since the pions at typical energies moved only $\sim 10\ \mu$m.

5.2.2 $\pi^0 \to e^+e^-$ generator

A custom decay generator was written for the signal mode decay. This was needed to include radiative corrections which were important for the measurement (see Section 1.3).

First the generator had to decide if a photon should be emitted in the decay, in other words should we generate $\pi^0 \to e^+e^-$ or $\pi^0 \to e^+e^-(\gamma)$? Equation (1.17) gives the differential rate for radiation relative to the lowest level rate but we needed the rate relative to the total rate including all next to leading order corrections. The total rate for radiation diverges logarithmically as $x \to 1$ and a cutoff was needed. The cutoff corresponded to a lower limit on the photon energy that was considered. The probability for radiation then depended on this cutoff. Applying the total correction (-3.4%) and a cutoff at $1-x=\delta$ we find for the radiation probability:

$$\frac{\Gamma^{\mathrm{rad}}(\delta)}{\Gamma^{\mathrm{tot}}} = \frac{1}{1-0.034}\,\frac{\alpha}{\pi}\int_{x_{min}}^{1-\delta}\left\{\frac{x^2+1}{1-x}\ln\left(\frac{1+v}{1-v}\right)-\frac{2xv}{1-x}\right\}\mathrm{d}x \tag{5.2}$$

The cutoff used was $\delta = 0.0001$ corresponding to less than a 10 keV photon, an energy which when boosted into the lab frame is not much higher than 1 MeV, significantly lower than the detector resolution. With this cutoff the probability for radiation was 40.92%. As mentioned in Chapter 1 the model for the radiative corrections used is probably only accurate up to about $x = 0.98$, suggesting a cutoff value of about $\delta = 0.01$. We picked the much lower value to have some simulation close to 1 even if it's not entirely accurate. Running the simulation with values from $\delta = 0.01 - 0.00001$ showed that the signal acceptance was not sensitive to changes in that range. Also the fraction of $\pi^0 \to e^+e^-$ events with $x < 0.95$ did not change when varying δ within that range.

If the generator chose to emit a photon, a hit or miss approach was used to pick the x-value from the differential rate:

$$\frac{d\Gamma^{\text{rad}}}{dx} = \Gamma^0_{ee} \frac{\alpha}{\pi} \left\{ \frac{x^2+1}{1-x} \ln\left(\frac{1+v}{1-v}\right) - \frac{2xv}{1-x} \right\} + O(\alpha^2) \tag{5.3}$$

For this to work a simple probability distribution that approximates the differential rate and is larger for all possible x was chosen. Simple here means one that can be generated by a direct transformation of a uniform distribution. Power law distributions have this property and x^{-1} behavior was used:

$$A(x) = \frac{1}{(\log(1-x_0) - \log(\delta))} \frac{1}{1-x} \tag{5.4}$$

The normalization was chosen so that $\int_{x_0}^{1-\delta} A(x)\mathrm{d}x = 1$. An x was then picked from the approximate distribution and based on the ratio between the approximate and the wanted distribution the event was accepted. In order to get a high acceptance rate the approximate distribution should have the same maximum as the wanted distribution. It is, however, not required to get the correct distribution. A different and an intuitive way of describing the hit or miss procedure is to pick a point in the area under the approximate distribution uniformly and keep it if it also falls in the area under the wanted distribution.

The direction of the radiated photon was not simulated but was merely aligned with the electron or the positron in the rest frame of the pion. Since the photon was never actually

reconstructed, and the electron or positron that it was aligned with would be bent in the magnet, its direction didn't matter. Another concern would be that the extra photon might veto the event in a way not simulated correct. An extra cluster or a hit in one of the photon vetoes could do this. The events that might have caused a small bias had large x since they were otherwise accepted, and for high x the photon energies were tiny, ~ 3 MeV at $x = 0.95$. Even boosted to the lab frame the energies would not exceed 1 GeV which is the threshold for a cluster seed in the calorimeter. The photon vetoes were more sensitive, the cut was at 300 MeV, but required the photon to come off at a high angle where the boost was small and not large enough to bring these soft photons above threshold.

5.2.3 $\pi^0 \to e^+e^-\gamma$

The Dalitz normalization and main background decay was simulated using the $O(\alpha_{EM}^2)$ radiative corrections calculation of Mikaelian and Smith [27] [28]. The $\pi^0 \to e^+e^-\gamma(\gamma)$ decay implementation used a cutoff on the inner bremsstrahlung in terms of a cutoff on the two photon invariant mass $m_{\gamma\gamma} > 1$ MeV/c^2. The Monte Carlo sample was generated only with $m_{e^+e^-}$ above 65 MeV/c^2. Using the Particle Data Group's average value for the π^0 slope parameter of $a = 0.032$ the $m_{e^+e^-} > 65$ MeV/c^2 tail made up 3.19% of the full decay rate. With the $m_{\gamma\gamma}$ cutoff and the $m_{e^+e^-}$ restriction the probability for inner bremsstrahlung was 28.03%.

More detail on the Dalitz generator can be found in [29].

5.3 Particle tracing

When the final decay particles and their momenta had been found they were boosted into the lab frame and traced through the detector. The tracing involved particles interacting with the material and magnetic fields in their path.

Detailed interaction with materials was started at the vacuum window and continued down to the calorimeter. In the MC this region was split into a number of volumes in z according to the material content. For each volume a separate GEANT based simulation was done,

simulating dE/dx, bremsstrahlung, multiple scattering, delta rays and photon conversions.

Charged particle interactions with the magnet field were simulated at different levels. Charged particles would receive a momentum kick as they flew from DC2 to DC3 through the magnet. The magnet field strength was regularly calibrated during the runs using $K_L \rightarrow \pi^+\pi^-$ events and the calibration values were used in the MC.

The magnet field permeated further away from the magnet and fringe field corrections were applied between the two upstream chambers and the two downstream chambers. The Earth's magnet field was also taken into account throughout the detector, in particular residual fields in the vacuum tank were measured as a function of z and included in the simulation.

Electrons and photons were traced until they either exited the fiducial region of the detector or were absorbed in a photon veto or in the calorimeter. Charged pions were traced until they escaped or hit the calorimeter and they were constantly given the chance to decay to muons (neutrinos were not traced). The weakly interacting muons were traced all the way through the muon systems at the back of the detector.

Particles created in particle interaction like δ-rays and e^+e^- pairs from conversions were also all traced, and of course tertiary decay particles were traced.

5.4 Detector response

As particles interacted with material in the active detector systems the appropriate detector simulation was run. For the photon veto systems this meant absorbing the particle hitting it, smearing the energy of the particle to reflect the resolution of the detector and then converting the energy into an ADC count. For the trigger hodoscope, a hit in a paddle was converted to a time depending on the distance to the PMT of that paddle. This time was then converted to a TDC value ready to be sent to the trigger logic.

The simulated response of the two main detector components are discussed below.

5.4.1 Drift Chamber simulation

When a charged particle was traced through a plane in a drift chamber, the nearest wire was identified and the closest distance between the particle trajectory and the wire was calculated. This distance was then smeared to match the measured hit resolution of the wire which depended on the plane and hit position within the plane. A dead-time of 45ns on the wire was simulated which would not let any other hits register for that period.

The drift distance was converted to a drift time and a TDC count using the same maps that were used to convert the other way in the analysis.

Additional tweaks were made to match data. For example, the hit efficiency was seen to be lower in data (Ke3) than in MC. Using Ke3 events, this extra inefficiency was measured for each wire and maps were created and used in the MC to match the data. More detail on these maps can be found in [30]

5.4.2 Calorimeter simulation

All photons and electrons were stopped when they hit the calorimeter and an electromagnetic shower was simulated for each. The showers were not simulated directly but were presimulated using GEANT to make up a library of shower profiles for a number of different particle energies and x-y positions in the seed crystal.

The library was used in the following manner: The incoming particle energy was first smeared to match better the detector resolution, then the energy was used to find the mean shower maximum depth in the crystal:

$$Z_{\text{shower max}}(e) = 0.11\ \text{m} + 0.18\ \text{m}\ \ln E \tag{5.5a}$$

$$Z_{\text{shower max}}(\gamma) = 0.12\ \text{m} + 0.18\ \text{m}\ \ln E \tag{5.5b}$$

with the energy measured in GeV. Photon showers started one cm further into the crystal. This was caused by the mean length the photon traveled before converting into an e^+e^- pair. Now projecting the particle to the depth of the mean shower maximum the transverse position was

found, and an appropriate shower profile could be picked (randomly in the right bin). The shower energy was then scaled to match the energy of the incoming particle.

Pions and muons were treated differently. About 30% of the pions showered hadronically and a different library of showers was created for them. The pions that did not shower (and the muons) just left a minimum ionizing energy of 320 MeV in one crystal.

The energy deposited in each crystal was converted to a charge, and a pulse shape measured from data was used to spread the energy in time so that the individual time slices read out were simulated realistically.

5.4.3 Accidental detector activity

The decay particles were not the only cause for activity in the detector although that would be ideal. All sorts of activity unrelated to the decay, such as junk from other beam interaction with the detector and cosmic rays, all acted as a constant background noise. This noise was simulated by using an overlay of accidental events measured with the trigger set off by the accidental counters (Section 2.1.3). Before the trigger simulation was done, extra hits and energy from the accidentals were added to the simulated event. This activity lead to all sorts of inefficiencies in the track finding and clustering, effects that were very important to simulate well.

5.5 Trigger simulation

The trigger definitions used in the on-line experiment were also used in the Monte Carlo, in fact the exact same files were used.

The L1 trigger was mainly simulated simply by setting bits when the various trigger thresholds were overcome and then by applying the trigger logic to the output. An exception was the Etotal source which was simulated with a realistic non-sharp threshold and inefficiencies at high energies. Similarly the HCC thresholds and inefficiencies for the individual crystals were simulated using measured values.

For the L2 trigger the exact algorithms used in the hardware were adopted and run on the MC L1 output.

The L3 software filter was again simply copied off the on-line computers and fed the output of the Monte Carlo. This also meant that the final MC output was written in the same format as the actual data.

5.6 Monte Carlo samples

A number of different Monte Carlo samples were generated for the analysis. Table 5.1 shows the samples and the number of generated events in each run period.

Sample type	Winter	Summer	99
$K_L \to \pi^0\pi^0\pi^0,\ \pi^0 \to e^+e^-$	20M	20M	20M
$K_L \to \pi^0\pi^0\pi^0_D,\ m_{e^+e^-} > 65$ MeV	350M	250M	800M
$K_L \to \pi^0\pi^0\pi^0,\ \pi^0 \to e^+e^-e^+e^-$	30M	30M	60M
$K_L \to \pi^0\pi^0_D\pi^0_D$	100M	100M	250M
$K_L \to \pi^0\pi^0\pi^0_D$ + 1 conv.	4000M	4000M	7000M
$K_L \to \pi^0\pi^0\pi^0$ + 2 conv.	1000M	1000M	2000M

Table 5.1: The samples and their sizes used in the analysis for signal, normalization, and backgrounds.

A sample of $K_L \to 3\pi^0$ events with a $\pi^0 \to e^+e^-$ was generated mainly to estimate the signal acceptance. Similarly a sample of the normalization $K_L \to 3\pi^0$'s with $\pi^0 \to e^+e^-\gamma$ where $m_{e^+e^-} > 65$ MeV was created. The normalization sample was used for many things in this analysis. As with the signal it served to find the acceptance, but it was also the main tool in identifying and estimating systematic errors from the Monte Carlo modeling. The normalization was also the main background to the signal so the same sample was used for background studies. Other backgrounds all came from $K_L \to 3\pi^0$ and included $\pi^0 \to e^+e^-e^+e^-$, $2\pi^0 \to e^+e^-\gamma$, $\pi^0 \to e^+e^-\gamma$ with a photon conversion in the material upstream of the magnet, and $K_L \to 3\pi^0$ with all $\pi^0 \to \gamma\gamma$ where two of the six photons converted. The backgrounds will be discussed

further in Section 6.5.

A couple of comments should be made for the samples with photon conversions in material. The large samples were obtained by stopping the simulation when the tracing reached the magnet if the exact number of wanted conversions had not occurred. For the double conversions this was still not enough to generate a sufficiently large sample, so a modified conversion probability was used in that case. Each interval in z from the vacuum window to the magnet had its amount of material multiplied by 10 for the purpose of finding the conversion probability through:

$$P_{\text{conv}} \;=\; 1 - e^{-\frac{7}{9}(X/X_0)} \tag{5.6}$$

X/X_0 is the fraction of a radiation length that a photon is presented with in the material. The total fraction of radiation lengths in the material before the magnet was close to 0.7% giving a probability for conversion of 0.54%. For such small values the probability is approximated well by just $P_{\text{conv}} \approx \frac{7}{9}X/X_0$, so multiplying X/X_0 by 10 was the same as multiplying the probability by 10. Since the probability was small even when enlarged it was safe to assume that two conversion was now a hundred times more likely. This factor of a hundred was used to get the normalization right for these background events.

Chapter 6

$\pi^0 \to e^+e^-$ & $\pi^0 \to e^+e^-\gamma$ reconstruction and analysis

This chapter is dedicated to describing the selection of the data sets, and how the signal and normalization candidate events were chosen from these data samples. First we discuss the initial offline data filtering called the *crunch*, then we briefly mention how data quality is checked and which requirements are put on the quality of the data for this analysis. The event reconstruction procedure and the cuts applied are then covered ending with the revelation of an unambiguous $\pi^0 \to e^+e^-$ signal. Possible systematic uncertainties in the measurement are studied before the final branching ratio is calculated at the end of the chapter.

6.1 Data sets and the 2E-NCLUS crunch

The KTeV detector was configured and operated in the E799-II configuration during three distinct time periods. The first was in the winter of 1997, the summer of 1997 was the second, and then data was taken again in the late fall and winter of 1999. The three periods are just known as the *winter*, *summer* and *99* run periods. To illustrate how the data was distributed over the run ranges, Fig. 6.1 shows the number of fully reconstructed Dalitz normalization events for each of the runs that were used in the measurement.

The data taken in each period were split to tapes depending on trigger types and level 3 tags. Trigger 1 events were then run through a data reduction procedure called the 2E-NCLUS Crunch, which did full tracking, clustering and vertexing of all events. The events were split into streams depending on the number of clusters, tracks and some kinematic constraints. The

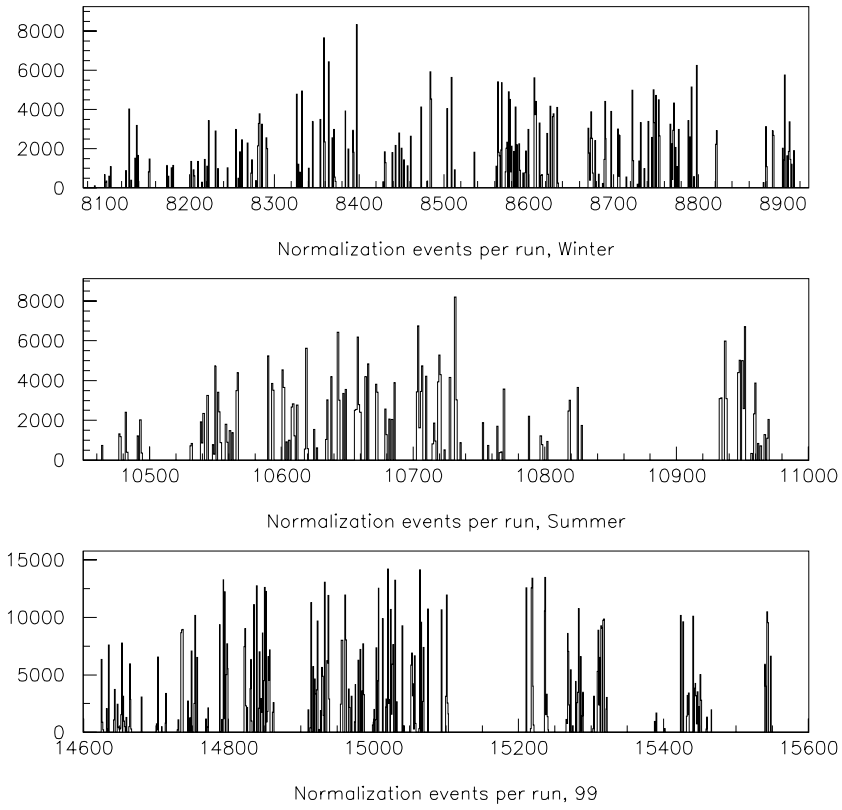

Figure 6.1: Run distribution of fully reconstructed Dalitz decays in the three run periods.

PI0TOEE stream was used for this analysis. It required six or more clusters, two tracks with $E/p > 0.9$ and an invariant mass of the e^+e^- pair greater than 70 MeV/c^2.

The total output of this stream was about 100 GB of data which constituted the basis for the analysis. The data was still split in the winter, summer and the 99 run periods and this distinction was kept all along.

6.2 Data quality selection

While the experiment was taking data, problems with the hardware would show up every now and then. If failures were found the run would be stopped and the hardware repaired or replaced, but some failures were not identified until later and data for those type of runs would

have to be selectively thrown out at a later stage.

A database was created to keep track of known problems with the hardware for each run and spill. Depending on the analysis one could then choose to use a given run/spill. E.g. for this analysis the TRDs were not used, so spills with TRD failures could be used anyway, while problems with the chambers clearly couldn't be neglected.

The database consisted of a 32 bit mask for each spill where each bit would be set if the corresponding detector component was identified as bad during the spill. Table 6.1 shows the definitions for each bit and which bits were cut on in the analysis. There are a couple of differences between the winter, summer and 99 runs as can be seen in the table.

Apart from the bit mask database, a list of spills and some entire runs were identified as bad for a variety of reasons. This list of bad runs and spills is shown in Table 6.2 and they were all cut from the analysis sample.

6.3 Reconstruction method

In this section the event reconstruction method will be discussed. The analyses of the two modes were very similar and the following applies to both modes. Any differences will be explicitly mentioned.

Events for both signal and normalization were taken from the PI0TOEE stream of the 2E-NCLUS crunch. Each event from this stream would go through a series of general analysis stages before being dubbed a signal or a normalization event candidate. First an event was checked against the bad spill mask and the set of bad runs to see if it was an event from a good quality spill (see the previous section). The next step was to run the tracking, clustering, and vertexing code in sequence as discussed in Chapter 4. The vertexing used the found tracks and clusters to look for a 2 track vertex, i.e. two oppositely charged tracks that were matched to clusters and had a common vertex. Clusters that were not matched to tracks were considered photon clusters.

With these steps done the candidate events could be selected. Six cluster events with two

Bit	Description	Win	Sum	99
1	Trigger	1	1	1
2	DPMT pedestal exponent > 0	1	1	1
3	Bad DPMT capacitor	1	0	0
4	Blown QIE comparator	1	1	1
5	Misc. dead DPMT	1	1	1
6	DPMT pedestal drift	0	0	0
7	DPMT gain drift	1	1	1
8	Broken DPMT dynode	1	1	1
9	CsI pipeline problems	1	1	1
10	Global CsI problems	1	1	1
11	E-total trigger problems	1	1	1
12	FERA ADC	1	1	1
13	Drift chambers	1	1	1
14	Photon veto	1	1	1
15	Trigger hodoscope	1	1	1
16	Muon veto/counter	1(*)	0	0
17	HCC trigger	1(**)	1	1
18	Banana trigger	1	1	1
19	TRD trigger	0	0	0
20	Hyperon trigger	0	0	0
21	DAQ/L3 trigger	1	1	1
22	non-799 run	1	1	1
23	Short run	1	1	1
24	Non- standard TRD voltage	0	0	0
25	1 dead TRD plane	0	0	0
26	> 1 dead TRD plane	0	0	0
27	TRD voltage sag	0	0	0
28	Severe TRD problems	0	0	0
29	Beam problems	1	1	1
30	unused	0	0	0
31	unused	0	0	0
32	Miscellaneous	0	0	0

Table 6.1: Description of the bad spill mask bits and how it was used. A 1 indicates that spills with this bit on was not used. 0 means that the problem was ignored. (*) The muon veto was in Trigger 1 until run 8576 after that no cut was made on muon veto problems. (**) A cable was swapped until run 8245 causing the HCC trigger to register as bad. In order to use these runs anyway the swap was simulated in the MC and no cut on that bit was made during the affected runs.

Run	Spill	Run	Spill	Run	Spill
8205	25	10742	all	10916	all
8330	55-56	10765	all	10918	all
8428	1-148	10767	28-29	10923	all
8453	120-200	10790	all	10924	all
10596	all	10909	all	10927	all
10599	all	10914	all		

Table 6.2: List of runs and spills with serious problems that were cut from the sample.

tracks pointing to two clusters were selected as signal events, and seven cluster, two track events as normalization Dalitz events.

The neutral vertexing, as described in Section 4.4, was performed, pairing up the π^0's and resulting in a somewhat modified decay vertex. The refitted vertex gave the electron and positron 4-momenta, and using the vertex position and the cluster energies the photon 4-momenta were deduced. Adding all the particle 4-momenta gave the reconstructed kaon 4-momentum, which in turn provided a reconstructed kaon mass. Another quantity found from the kaon momentum was the amount of transverse momentum unaccounted for, $|p_T|$. This was the reconstructed momentum transverse to the kaon direction, defined as the direction of the line between the target and the reconstructed decay vertex. The quantity recorded for the event was p_T^2.

6.4 Basic cuts

Before attempting to purify the two samples by cutting out backgrounds, a number of basic cuts were applied to both samples. These cuts ensured integrity of the reconstructed decay events and are described here. The cut values and efficiencies are summarized in Table 6.4 and 6.5.

- *Photon veto cuts.* The photon vetoes, the RCs, SAs, and the CIA, were in the trigger and were required to have less than 500 MeV of energy. Slightly tighter cuts were made at this stage: events with energies above 300 MeV in the RCs, the SAs or the

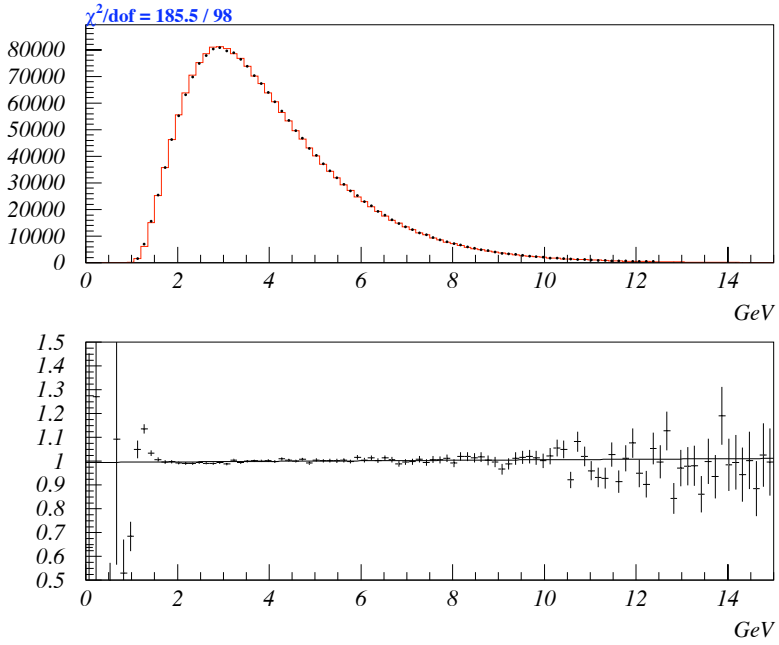

Figure 6.2: Minimum cluster energy distribution in data (dots) and MC (solid histogram) for the normalization mode.

CIA were removed.

- *Minimum cluster energy cut.* To avoid depending on the modeling of the trigger HCC block energy threshold at the nominal 1 GeV, a cut on the minimum cluster energy was made at 1.75 GeV. In Fig. 6.2 the minimum cluster energy distribution in the normalization sample is plotted for data and MC. At the low end, the MC was failing so we cut above this region.

- *Total cluster energy cut.* The Etotal trigger allowed events with more than 28 GeV of energy in the calorimeter. We cut tighter at 35 GeV to make sure to get away from any dependence on the trigger threshold. On the high side we cut events with more than 210 GeV of energy.

- *E/p cut.* For each track we required that the energy of the cluster that it pointed to had an energy within 8% of its momentum. In other words, $|E/p-1| < 0.08$ was required. The distribution of E/p in the normalization sample is plotted in Fig. 6.3.

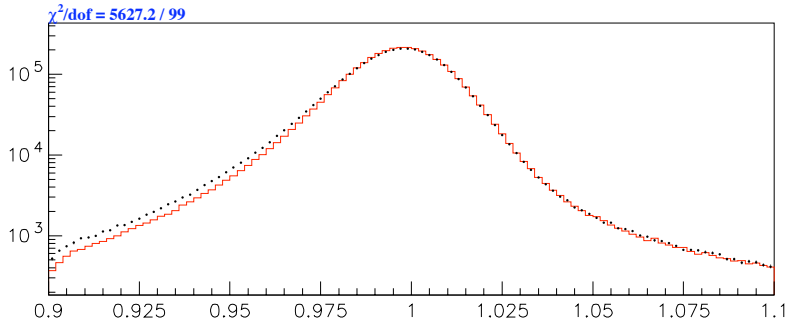

Figure 6.3: E/p distribution in data (dots) and MC (solid histogram) for the normalization mode.

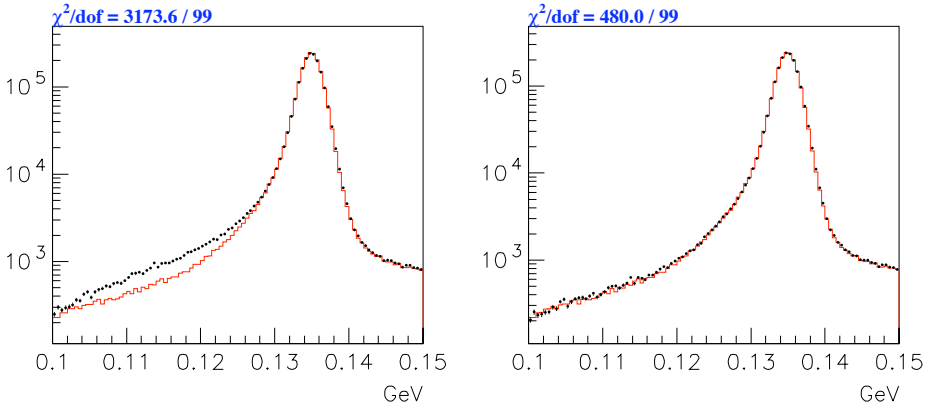

Figure 6.4: Normalization $m_{e^+e^-\gamma}$ distribution before and after the tightened CA cut. Data is the dots and MC is the solid histogram.

- *Vertex Z position cut.* Events whose reconstructed decay vertex was outside the fiducial region, be it after the vacuum window or before and in the last sweeper magnet, were cut from the sample for obvious reasons. The accepted vertex positions were between $z = 96$ m and $z = 158$ m.

- *CA energy cut.* The cut on the collar anti in the trigger at 13 GeV was tightened to 4 GeV in the analysis in order to remove a small discrepancy between data and MC seen in the $m_{e^+e^-\gamma}$ distribution in the Dalitz sample, Figure 6.4. The cut was very efficient (better than 99%) in both signal and normalization and clearly removed the excess of

data in the low side mass tail.

- *Pipe block energy cut.* Another very loose cut was used to reduce spurious events with energy lost down the beam holes. For each cluster seed in the inner ring around the beam holes, the ratio of energy in that block with its neighbor away from the hole was cut on. If the neighbor block had less than 0.05 times the energy of the seed the event was cut. The cut helped make the π^0 pairing χ^2 distribution in the normalization sample agree better on the high end tail.

6.5 Backgrounds

The reconstruction method described in the previous section allowed a number of backgrounds into the sample. The complete reconstruction of the kaon decay chain was however powerful enough to constrain the backgrounds in the candidate events to mostly come from other $K_L \to 3\pi^0$ events. The known backgrounds are described here:

High $m_{e^+e^-}$ Dalitz decays, the high $m_{e^+e^-}$ tail of the normalization mode, where the Dalitz photon was lost could reconstruct a couple of MeV high and become a serious background in the $m_{e^+e^-}$ signal region. All other backgrounds were 4 track (electron-positron) backgrounds where a positron and an electron were lost, for example by getting bent out of the fiducial region at the magnet. The remaining electron and positron could then form a $m_{e^+e^-}$ close to the π^0 mass and get into the sample. $K_L \to 3\pi^0$'s with two π^0's decaying to $e^+e^-\gamma$ or one $\pi^0 \to e^+e^-e^+e^-$ could follow this pattern. Another source of electron-positron pairs came from photon conversions in the material upstream of the chambers, in particular in the vacuum window. If any two of the photons from the $\pi^0 \to \gamma\gamma$ decays converted or if a Dalitz decay combined with a conversion, then there was a chance of the event getting into the sample. One might think that single conversions ($3\pi^0 \to 6\gamma$ with 1 conversion) could make it into the sample, but the e^+e^- mass of this type of event is highly peaked around zero and never gets as high as the pion mass. Table 6.3 lists all the identified backgrounds.

Background source decay chain		Background probability
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(\gamma\gamma)$	$(e^+e^-\gamma)$, $m_{e^+e^-} > 65$ MeV/c^2	2.22×10^{-4}
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(\gamma\gamma)$	$(e^+e^-e^+e^-)$	2.05×10^{-5}
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(e^+e^-\gamma)$	$(e^+e^-\gamma)$	8.44×10^{-5}
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(\gamma\gamma)$	$(\gamma_{\to e^+e^-}\gamma_{\to e^+e^-})$	1.91×10^{-3}
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(\gamma\gamma_{\to e^+e^-})$	$(\gamma\gamma_{\to e^+e^-})$	
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(\gamma\gamma_{\to e^+e^-})$	$(e^+e^-\gamma)$	6.96×10^{-3}
$K_L \to 3\pi^0 \to (\gamma\gamma)\,(\gamma\gamma)$	$(e^+e^-\gamma_{\to e^+e^-})$	

Table 6.3: Types of background sources. $\gamma_{\to e^+e^-}$ means a photon converting in the material upstream of the first chamber. The probabilities included the $K_L \to 3\pi^0$ branching ratio of 0.1983.

The backgrounds were simulated with the Monte Carlo. For the single Dalitz decay background the normalization sample was used, just run through the signal analysis instead. Separate samples for all the other background were generated as described in Chapter 5, Section 5.6.

To get the background normalization right, combinatorial factors and the appropriate branching ratios were used for each source. The probabilities of the individual backgrounds are shown in the second column in Table 6.3. For the background with a single conversion and a Dalitz decay only events with $m_{e^+e^-} < 65$ MeV/c^2 were considered to avoid double counting the events with $m_{e^+e^-} > 65$ MeV/c^2, which were already included in the normalization sample.

The next section describes how these backgrounds were removed from the signal sample.

6.6 Selection cuts and background

The backgrounds described in the previous section completely dominate the signal sample at this point in the analysis. The normalization mode on the other hand was relatively clean. The selection cuts described here were mainly designed to remove background from the signal sample.

Cuts were placed on most of the kinematic variables at our disposal to reduce the backgrounds, but also a special background "killer" cut was used. For each cut an efficiency for each

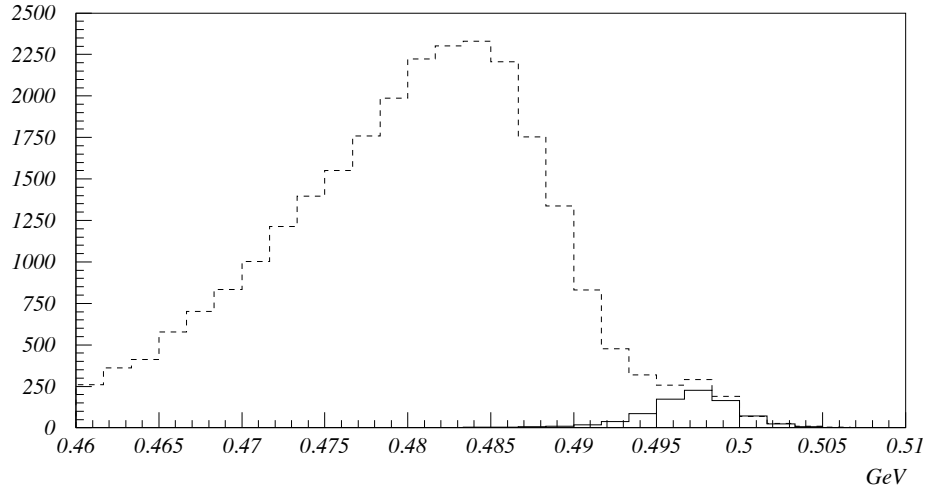

Figure 6.5: Kaon invariant mass distribution in the signal sample after all cuts except the $m_{e^+e^-}$ cut. Data is shown with dashes and signal MC as a solid histogram.

run period (win,sum,99) is quoted for both signal and normalization. The cut efficiency here is defined as the fraction of MC events that survive the cut in question after all other analysis cuts have already been made.

In Table 6.4 later in this chapter, the full set of applied cuts are summarized.

6.6.1 Kaon invariant mass cut

The reconstructed kaon mass gave us a very powerful constraint on the event candidates. Events in the signal sample were cut if the reconstructed kaon mass fell outside the interval 500 ± 10 MeV/c^2. This cut was slightly asymmetric cutting tighter on the low side of the kaon mass peak. The majority of the background had a low kaon mass, while high kaon mass backgrounds were almost nonexistent. Figure 6.5 shows this asymmetry.

In the normalization the cut was 500 ± 25 MeV/c^2. There were two reasons for the looser cut in the normalization. First the cut wasn't needed to remove backgrounds in the normalization and second the kaon mass distribution after all other cuts was somewhat broader in the normalization than in the signal. This was because in the signal a very tight cut was necessary on $m_{e^+e^-}$ which removed tails in the reconstructed kaon mass. In fact after all other

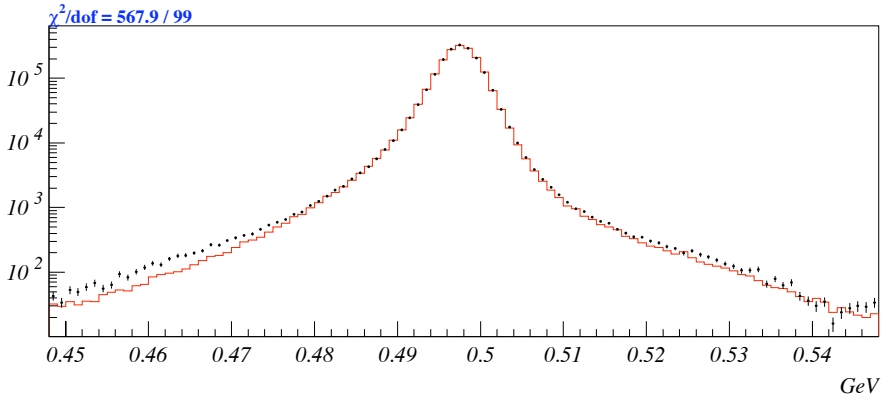

Figure 6.6: Kaon invariant mass distribution for data in dots and MC as the solid histogram, both for the normalization mode.

cuts the kaon mass cut was 99.8% efficient in both signal and normalization. A plot of the data and MC distribution in normalization after all other cuts is shown in Fig. 6.6.

6.6.2 p_T^2 cut

Another constraint from the complete reconstruction of the kaon came in the form of transverse momentum, p_T^2. The line from the target to the reconstructed vertex defined the incident kaon direction. p_T^2 measured how much reconstructed momentum was transverse to this direction. Backgrounds from events with missing particles and accidental energy could form a good kaon mass, but the transverse momentum would be off. In both signal and normalization the allowed amount was $p_T^2 < 0.001$ (or $\ln(p_T^2) < -6.9$). The natural logarithm of p_T^2 is plotted in Fig. 6.7.

6.6.3 π^0 invariant mass cut

The individual π^0 masses were also found but a cut was only put on the π^0 that decayed to either e^+e^- or $e^+e^-\gamma$. The cut was rather different in signal and normalization. Mispairings of the π^0's in the normalization caused the $m_{e^+e^-\gamma}$ distribution to have a wide tail both high and low. The low side was a bit wider due to the radiative corrections, see Figure 6.8. These events

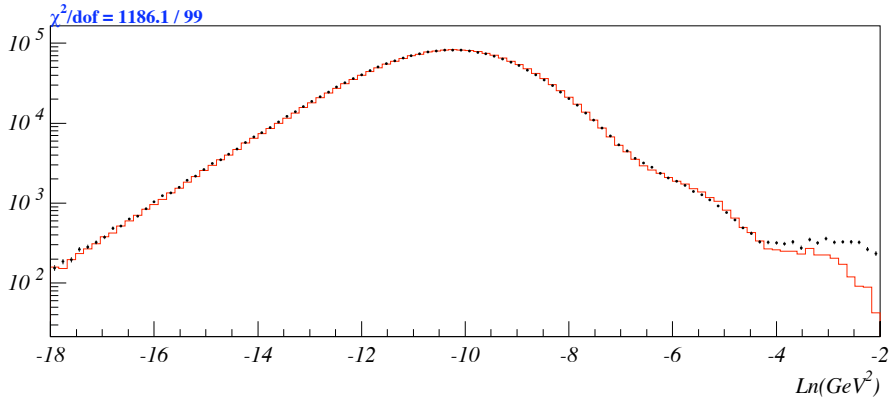

Figure 6.7: $\ln(p_T^2)$ distribution in data (dots) and MC (solid histogram) for the normalization mode.

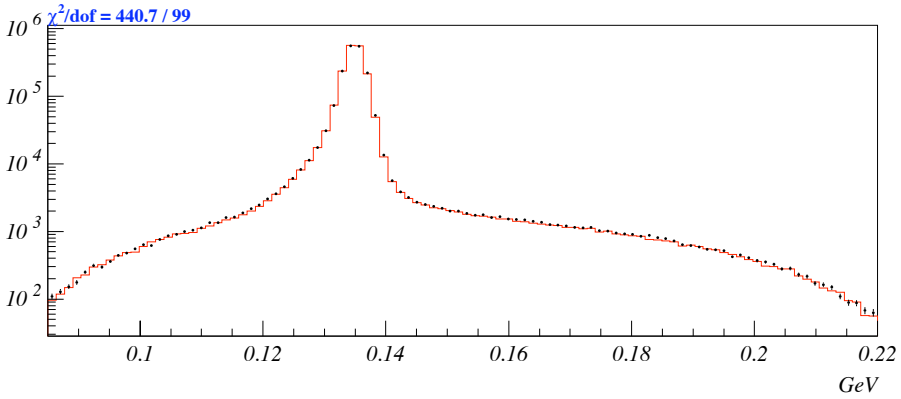

Figure 6.8: $m_{e^+e^-\gamma}$ distribution in data (dots) and MC (solid histogram).

were perfectly good Dalitz events and didn't need to be cut. The cut allowed events in the range $0.1\ \mathrm{GeV/c^2} < m_{e^+e^-\gamma} < 0.2\ \mathrm{GeV/c^2}$ and was almost 100% efficient.

For the signal the situation was quite different. Mispairings were rare so the $m_{e^+e^-}$ distribution was much narrower than the $m_{e^+e^-\gamma}$ distribution. The radiative corrections still caused a low side tail as was the case in the Dalitz decay, but here it dominated the tail. Recall that the signal was defined to be only the $\pi^0 \to e^+e^-$ decays that had $x > 0.95$. We made the cut symmetric around the mass peak ± 0.05 in x, which translates into a required mass interval: $0.1316\ \mathrm{GeV/c^2} < m_{e^+e^-} < 0.1384\ \mathrm{GeV/c^2}$. The low x events that crept into the signal region were considered background.

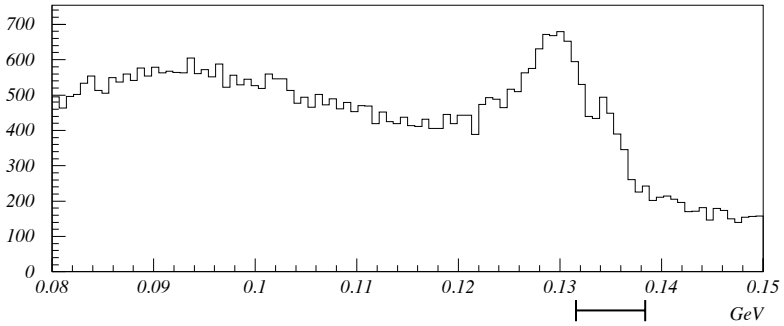

Figure 6.9: $m_{e^+e^-}$ distribution in signal data sample after basic cuts, kaon mass cut and p_T^2 cut. The signal region is indicated below the plot. The signal is obscured by background at this stage.

A plot of $m_{e^+e^-}$ in data after the cuts described is shown in Fig. 6.9. Background dominate the signal region after all the kinematic cuts. A small peak can be seen but nothing conclusive.

6.6.4 Pairing χ^2 cut

The type of background that dominate in Fig. 6.9 came from events where the electron and positron came from two different π^0's, these type of events have the π^0 pairing wrong and cutting out high pairing χ^2 events would reduce this background. Looking at Fig. 6.10 we clearly see the bulk of the events at high χ^2. The plot shows the logarithm of the χ^2 and events above $\ln(\chi^2) = \ln(20) \simeq 3$ were cut. The $m_{e^+e^-}$ distribution in data after the cut is shown in Fig. 6.11, which now only shows a peak of backgrounds on the low side of the signal region. The remaining background was mostly Dalitz decays and four track backgrounds where the two $\pi^0 \to \gamma\gamma$ had been correctly reconstructed.

In the normalization the distribution (data and MC) looked as shown in Fig. 6.12, where the logarithm of the χ^2 is plotted on both a linear and a log scale. The MC agreed well with data except on the high end tail where in the data extra events were present. The same cut was applied here as in the signal.

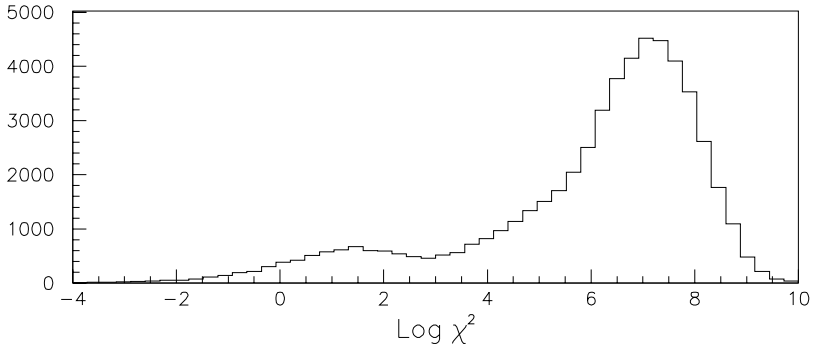

Figure 6.10: Pairing χ^2 ($\ln(\chi^2)$) distribution in signal data sample. The bulk of the distribution are mispaired backgrounds and are to the right of $\ln(\chi^2) = 3$

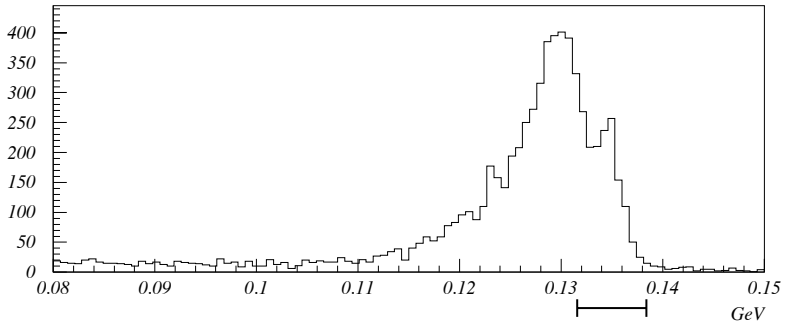

Figure 6.11: $m_{e^+e^-}$ distribution in data. The range below the plot indicates the signal region. After the χ^2 cut the signal peak was showing but still mostly covered by backgrounds.

6.6.5 Extra activity cut in the chambers

Most of the remaining four track background was eliminated using a cut on extra activity in DC2. Events with four tracks often lost one or two of the extra tracks in the magnet field, low momentum tracks would get bent out of the fiducial region and only leave hits in the first two drift chambers. The signature for these events was good SOD pairs (in-time hits), in the upstream chambers, that were not used in the final tracks.

The cut used was pretty harsh and it had to be, in order to clean up the signal region. All in-time hit pairs in one of the two views in drift chamber 2 that was not used in a track were considered. If any one of them was not too close to a track in that view the event was cut. More

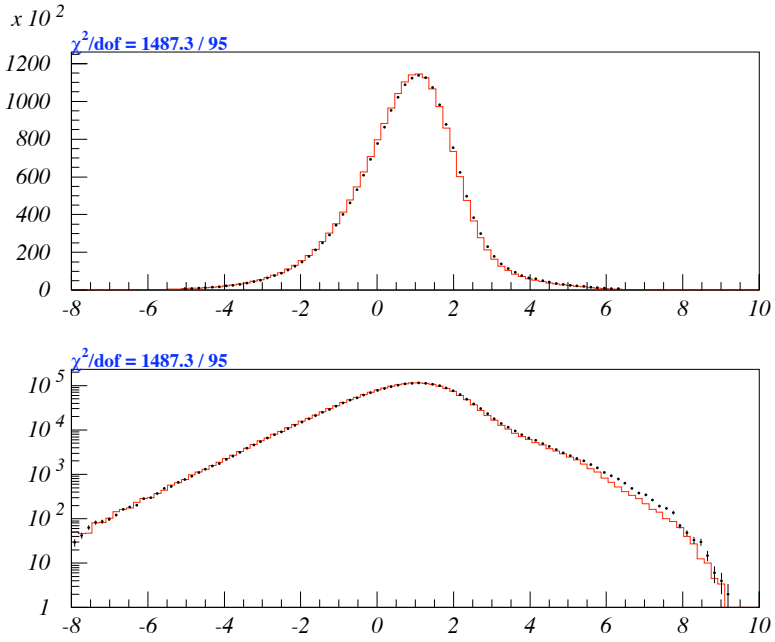

Figure 6.12: Distribution of the log of the pairing χ^2 in data (dots) and MC (solid histogram). Shown on a log scale and a linear scale.

specifically, events with an in-time hit more than 0.5 cm from any of the two tracks were cut. Figure 6.13 shows the shortest distance of in-time hits to the tracks in signal MC and background MC. We see how the signal MC was peaked at zero, which was activity created partly by the tracks themselves, while the background had very few events with activity near tracks. The same distribution in data and MC for the normalization sample is shown in Fig. 6.14. Decent agreement is seen except very close to zero where there was an excess of data.

The cut was very effective at removing the remaining four track backgrounds, as can be seen in Fig. 6.15 where the in-time activity cut has been applied.

Apart from cutting background the cut worked as an overall scaling of the signal and normalization of about 90%.

6.6.6 Signal and final background levels

A summary of the cuts applied is listed in Table 6.4. Figure 6.16 shows the $m_{e^+e^-}$ distribution of the final data sample after all cut. The background MC normalized to the measured

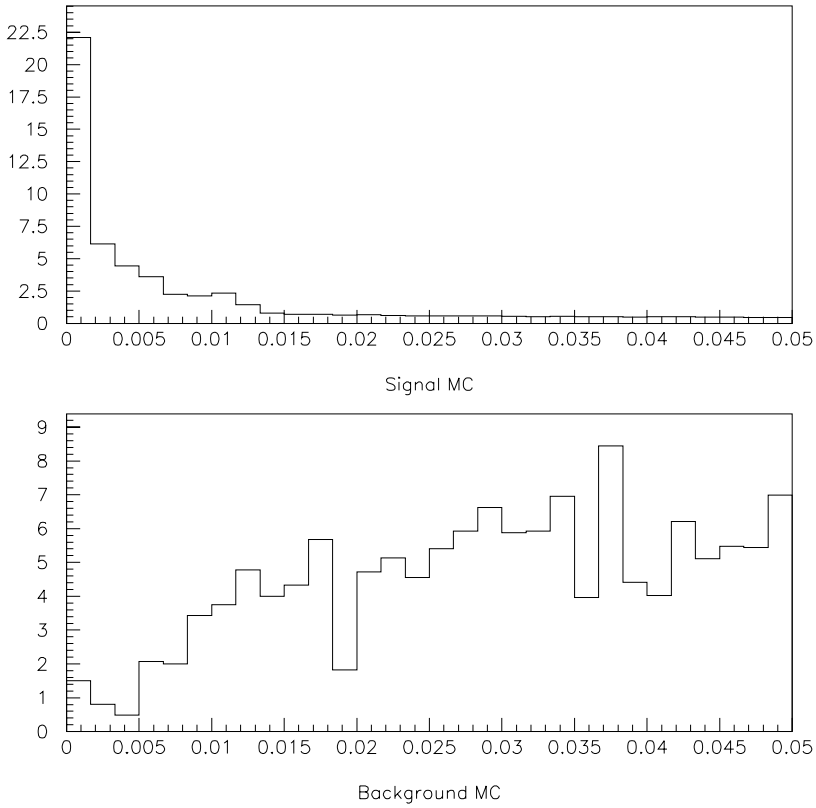

Figure 6.13: Distribution of the distance to nearest track for in-time hits in DC2. The top plot shows signal MC which peaked close to zero, and the bottom plot is the background MCs where most activity was away from the tracks. The units are meters in both plots.

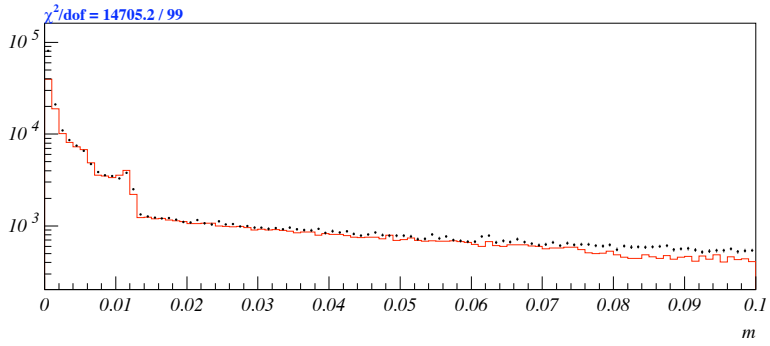

Figure 6.14: Distance to nearest track for in-time hits in DC2 for normalization MC (solid histogram) and data (dots).

Cut variable	Allowed value(s)
RC, SA, and CIA energies	< 0.300 GeV
Minimum cluster energy	> 1.75 GeV
Total cluster energy	> 35 GeV, < 210 GeV
E/p	> 0.92, < 1.08
Vertex Z	> 96 m, < 158 m
CA energy	< 4.0 GeV
Pipe block energy fraction	< 0.05
Kaon mass (signal)	> 0.490 GeV/c^2, < 0.510 GeV/c^2
Kaon mass (normalization)	> 0.475 GeV/c^2, < 0.525 GeV/c^2
p_T^2	< 0.001 GeV^2/c^2
$m_{e^+e^-}$ (signal)	> 0.1316 GeV/c^2, < 0.1384 GeV/c^2
$m_{e^+e^-}$ (crunch req.)	> 0.070 GeV/c^2
$m_{e^+e^-\gamma}$	> 0.1 GeV/c^2, < 0.2 GeV/c^2
Pairing χ^2	< 20
Extra in-time chamber act.	< 1 good SOD

Table 6.4: Summary of all offline cuts. The cuts are explained in the text.

Cut variable	Signal efficiencies	Normalization efficiencies
RC, SA, and CIA energies	(93.5%, 93.2%, 95.3%)	(93.5%, 93.2%, 95.3%)
Minimum cluster energy	(97.5%, 97.6%, 97.0%)	(96.6%, 96.6%, 95.5%)
Total cluster energy	(97.5%, 98.1%, 97.5%)	(99.2%, 99.3%, 98.9%)
E/p	(99.4%, 99.4%, 99.3%)	(99.3%, 99.3%, 99.1%)
Vertex Z	(96.9%, 96.9%, 96.8%)	(97.1%, 97.1%, 97.0%)
CA energy	(99.3%, 99.1%, 99.2%)	(99.4%, 99.2%, 99.3%)
Pipe block energy fraction	(99.8%, 99.8%, 99.9%)	(99.8%, 99.8%, 99.8%)
Kaon mass (signal)	(89.2%, 89.3%, 89.3%)	
Kaon mass (normalization)		(97.1%, 97.1%, 97.1%)
p_T^2	(98.8%, 98.2%, 98.9%)	(97.7%, 97.2%, 97.9%)
$m_{e^+e^-}$ (signal)	(92.9%, 92.8%, 91.9%)	
$m_{e^+e^-}$ (crunch req.)		
$m_{e^+e^-\gamma}$		(99.6%, 99.6%, 99.6%)
Pairing χ^2	(97.1%, 96.8%, 96.7%)	(95.7%, 95.3%, 95.2%)
Extra in-time chamber act.	(93.6%, 91.8%, 89.4%)	(93.4%, 91.8%, 89.3%)

Table 6.5: Summary of offline cut efficiencies. For each cut the efficiency is given for the three run periods (win, sum, 99).

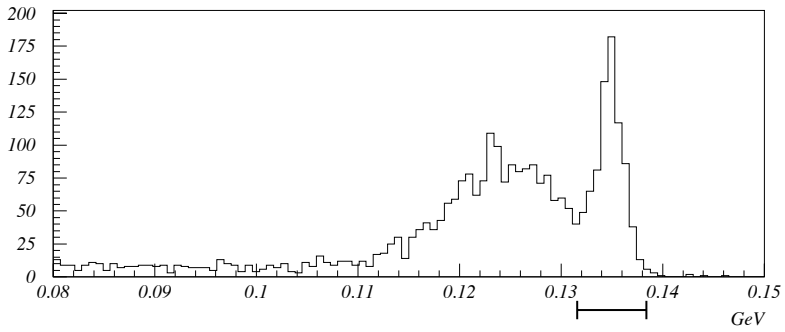

Figure 6.15: $m_{e^+e^-}$ distribution in data. The range below the plot indicate the signal region. The extra activity cut cleans the signal significantly.

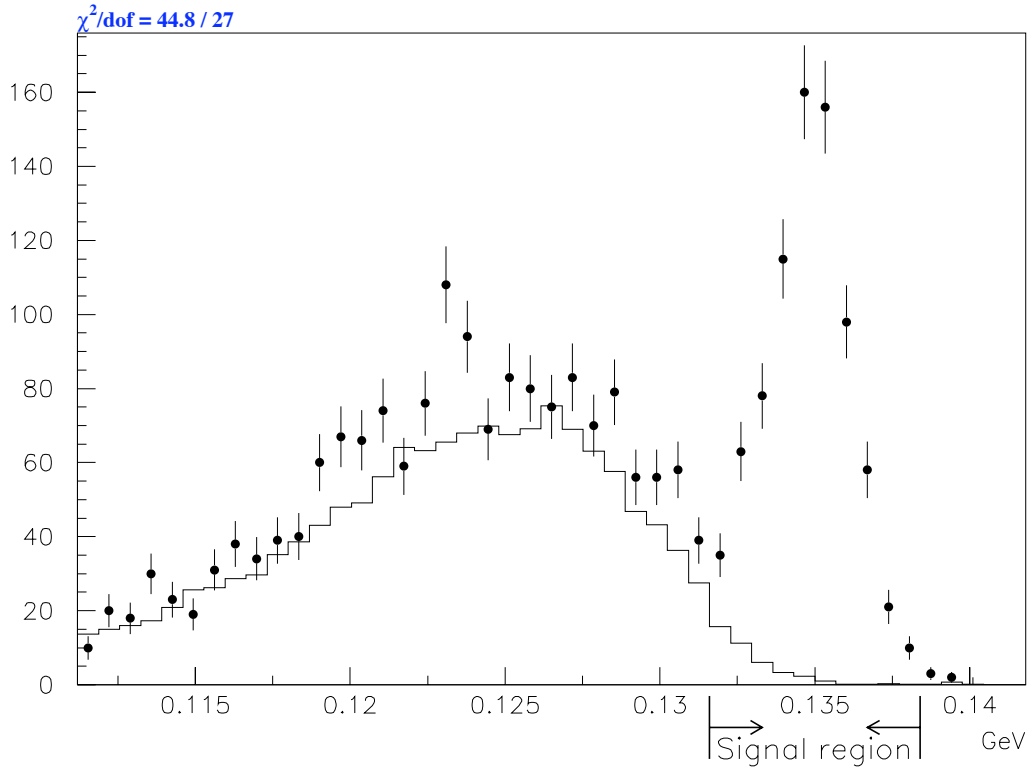

Figure 6.16: $m_{e^+e^-}$ distribution in the final data sample (Dots with errors). The background MC normalized to the measured "flux" is the overlaid histogram.

"flux" (see Section 6.9.1) is plotted on top. Table 6.6 summarizes the number of observed events after all cuts in signal and normalization, in Table 6.7 the corresponding acceptances as found by the MC are listed.

	Winter	Summer	99	Total
Signal	189	126	479	794
Background	11.4 ± 1.2	8.0 ± 0.9	24.9 ± 2.3	44.4 ± 2.7
Normalization	446137	313858	1114642	1874637

Table 6.6: Number of observed events in signal and normalization along with the initial MC estimate for the amount of background. The error on the background is from MC statistics only.

	Winter	Summer	99
Signal	$(2.941 \pm 0.004)\%$	$(2.930 \pm 0.004)\%$	$(3.139 \pm 0.004)\%$
Normalization	$(1.214 \pm 0.001)\%$	$(1.218 \pm 0.001)\%$	$(1.383 \pm 0.001)\%$

Table 6.7: Predicted overall acceptances for signal and normalization.

The remaining background, as estimated by the MC, was mostly from high $m_{e^+e^-}$ Dalitz decays, of the estimated total of 43.7 background events 34.6 of them were high e^+e^-mass Dalitz decays. Four background events came from $\pi^0 \to e^+e^-$ decays with $m_{e^+e^-}$ too low to be considered signal ($x < 0.95$). Another three background events were Dalitz decays with a conversion and the remaining fraction of the background were from the other sources.

The overall shape of the simulated background was correct, but a clear discrepancy between the data and the background MC in the sideband region, 0.11-0.13 GeV/c^2, was seen. By floating the background MC and fitting to the data on the range from 0.11 GeV/c^2 to 0.13 GeV/c^2, we found that the background level was underestimated by 20%. The $m_{e^+e^-}$ distribution is shown with the rescaled background MC overlaid in Fig. 6.17 Since we consider the 97 and 99 datasets separately when we calculated the branching ratio we show the final $m_{e^+e^-}$ distribution for the two periods in Fig. 6.18. The distributions with the background MC rescaled by 20% are plotted too. In both periods we see the same picture. Floating the background in the two periods separately indicated a 24% underestimate in 97 and 17% underestimate in 99. The difference was expected from statistics.

We found no reason that this 20% discrepancy would not be present under the signal peak,

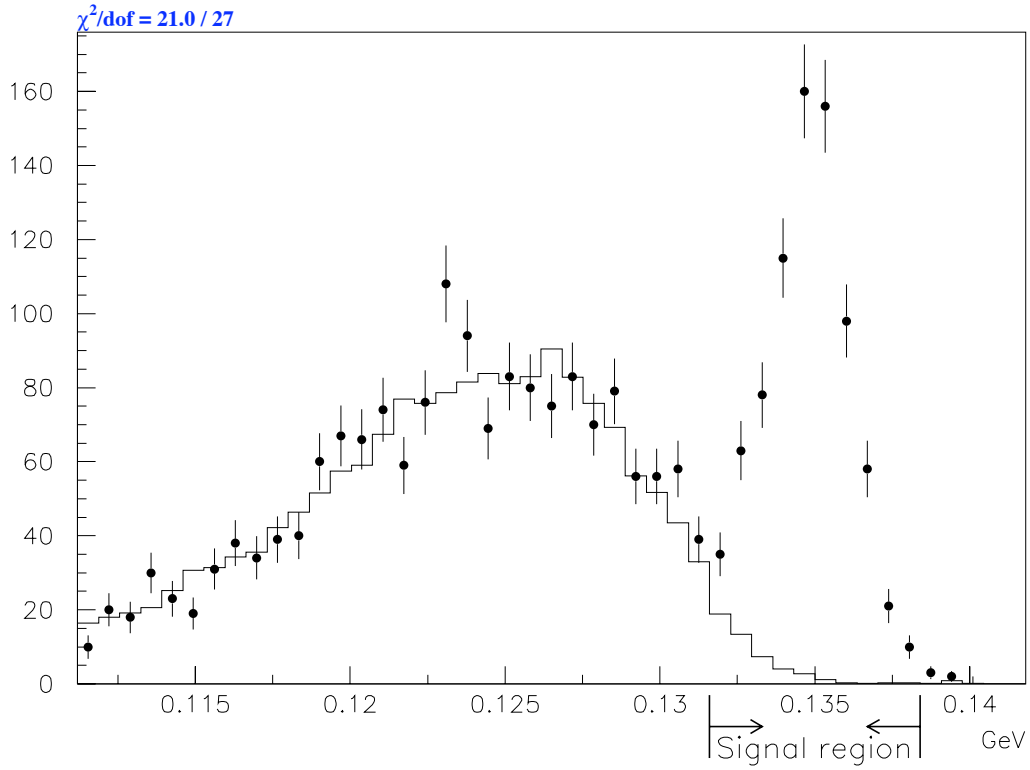

Figure 6.17: $m_{e^+e^-}$ distribution in the final data sample (dots with errors). The background MC normalized to the measured "flux", but scaled up by 20%, is the overlaid histogram.

and we therefore scaled up the background estimate by 20% to a total of 53.2 ± 3.3 events.

Due to our lack of understanding of this extra background we assigned a systematic error covering the whole background 'fix'. 8.9 events were added and we took that as the systematic uncertainty on the background. The background estimate was then 53.2 ± 9.5. The added uncertainty translated into a 1.2% uncertainty in the $\pi^0 \to e^+e^-$ branching ratio.

6.7 Background normalization

One of the biggest non-statistical uncertainties in the branching ratio measurement came from the background level not being understood properly. It would be nice to have at least a basic understanding of these extra backgrounds. Some work was done toward this end, and it's described in the following pages. It should be emphasized that the following was not part of the

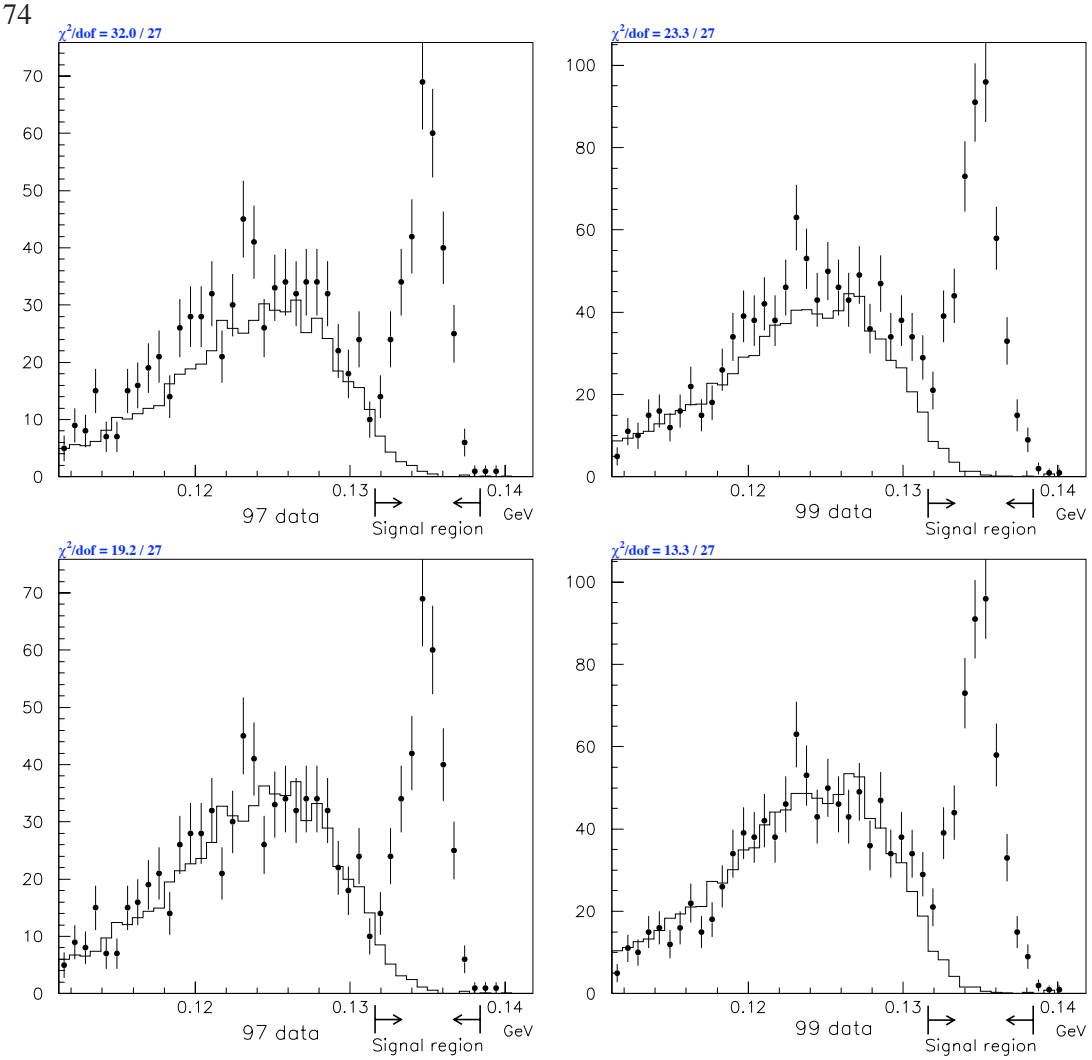

Figure 6.18: $m_{e^+e^-}$ distribution in the final data sample (dots with errors) for the 97 runs on the left and the 99 run on the right. The overlaid histograms are the background MC normalized to the measured "flux" in the top plots and scaled up by 20% in the bottom plots.

branching ratio analysis but merely a test to see if the background normalization could be better understood.

6.7.1 Extra software clusters

In an attempt to understand the discrepancy between data and Monte Carlo in the $m_{e^+e^-}$ sideband, a search for soft extra clusters were performed using *software clusters*. These were clusters that were not seeded at the trigger level but rather seeded by local energy maxima

found at the analysis stage. This allowed one to find (usually) low energy clusters that were left unnoticed by the HCC. The strategy was to reduce the background even further by cutting on these extra clusters and see if the background could be understood better. The background consisted mainly (after cuts) of high $m_{e^+e^-}$ Dalitz decays. These had a soft photon that was somehow lost, and if this happened through some effect in the calorimeter that was not simulated well, cutting events that had an extra software cluster would reduce the problem.

For each event the most energetic software cluster, if any existed, was considered. Events that had an extra in-time cluster with significant energy were identified and cut from the sample. The details of this procedure follows.

The timing χ^2 was the first thing that was looked at. It was a measure of how close the cluster was to the in-time bucket of the event, the low χ^2 clusters being in-time. Figure 6.19 shows the distribution of the natural log of the timing χ^2 of the highest energy software cluster in data and Monte Carlo. A peak in both data and Monte Carlo could be seen at low values and then there was a well simulated tail at high χ^2. Although the simulation was not very good, the peaks at low χ^2 could still be interpreted as the in-time clusters. Clusters to the right of both low peaks (at ln(3)) were considered out of time and therefore not considered as extra clusters part of the event.

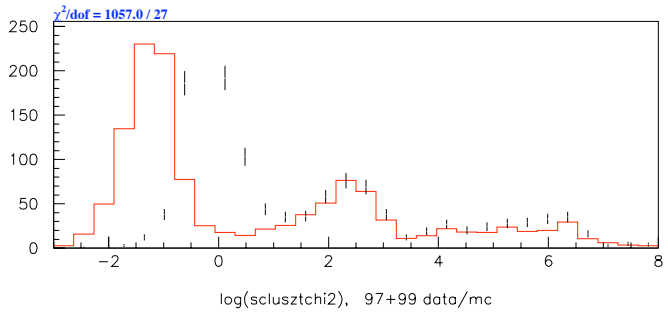

Figure 6.19: Timing χ^2 for the highest energy software cluster in data and Monte Carlo.

In Fig. 6.20 the maximum software cluster energy is plotted before and after removing

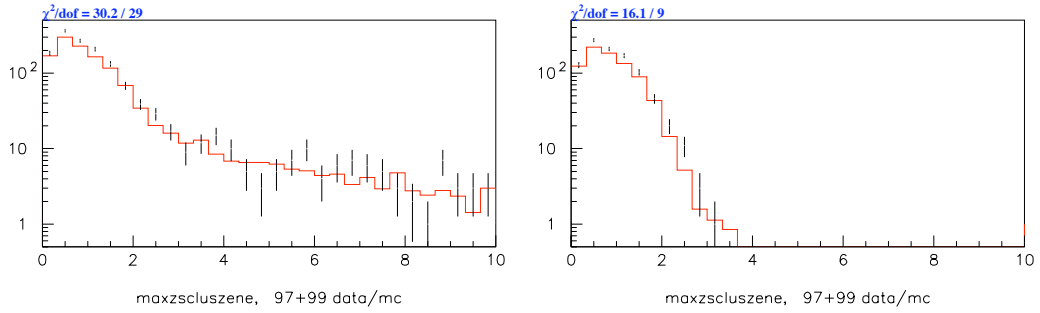

Figure 6.20: Maximum software cluster energy (in GeV) for data and MC before and after requiring low timing χ^2.

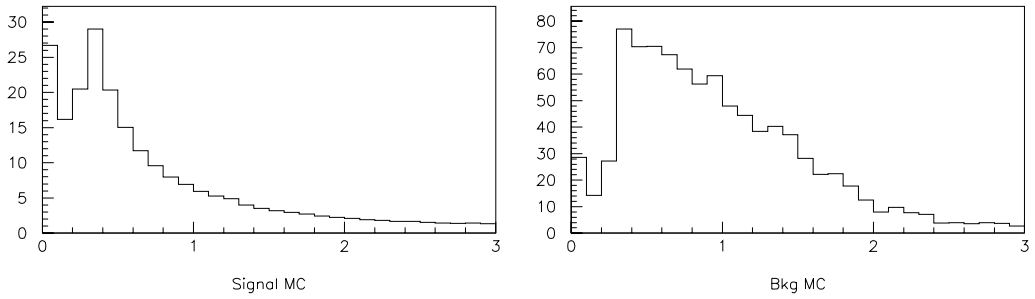

Figure 6.21: Maximum software cluster energy (in GeV) with good timing χ^2 for signal (left) and bkg (right) Monte Carlo.

the clusters with high timing χ^2. Notice how all the high energy clusters disappeared when a low timing χ^2 was required.

From looking at the maximum software cluster energy in the signal MC and the background MC, Figure 6.21, it was clear that the backgrounds had higher energy clusters and that an appropriately placed cut would remove more of the background.

Figure 6.22 shows the $m_{e^+e^-}$ distribution when events with an extra software cluster of energy > 0.5 GeV and timing $\chi^2 < 3$ were cut from the sample. The background level decreased but the $\sim 20\%$ discrepancy between data and Monte Carlo was still present, in fact it was a bit worse after the cut. Whatever was causing the discrepancy was still there and we had to look elsewhere to find the cause. The cut was considered for the analysis but it turned out that the advantage of less background was outweighed by the increased uncertainty in the overall

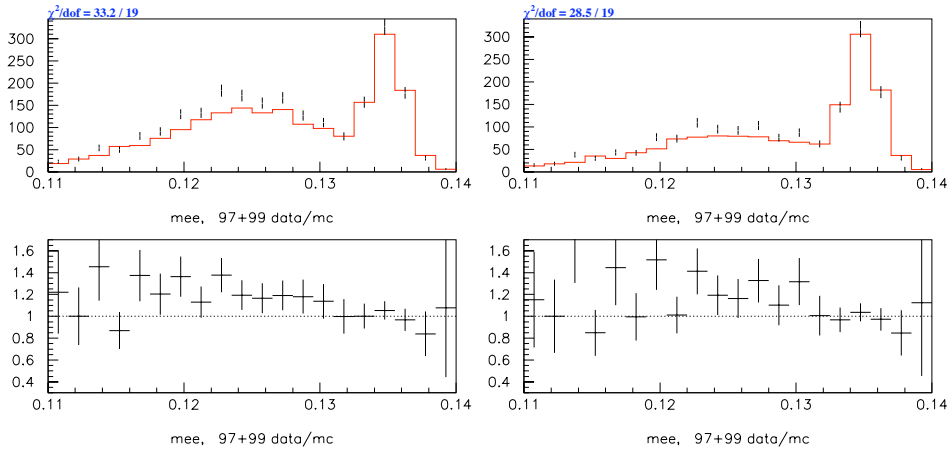

Figure 6.22: $m_{e^+e^-}$ distribution with and without the cut on in-time software clusters with significant energy.

background level. It was therefore discarded.

6.7.2 Photon veto counters

Another thing to look at that might affect the background level was the photon veto counters, the RCs, the SAs, and also the Collar Anti. If the simulation of these were not perfect, the Dalitz background, which loses a soft photon somehow, could have been affected. The nominal cut on the RCs and SAs was at 300 MeV which was somewhat tighter than the trigger requirement. Tightening the cut gradually to 30 MeV on all counters didn't have any noticeable effect on the background disagreement. Figure 6.23 shows the $m_{e^+e^-}$ distribution before and after the tightest cut. The cut just decreased the acceptance evenly in signal and background.

Cutting tighter on the CA energy turned out to be more interesting. With a cut at 150 MeV (nominal analysis cut was 4 GeV) we found what is shown in Fig. 6.24. The cut acted as an overall scale, but it also lowered the discrepancy in the sideband between data and Monte Carlo significantly. This suggested that Dalitz events where the extra photon struck close to the beam holes may not have been simulated very well.

The cut was very tight, 1/3 of the signal was lost, so it was not used in the analysis.

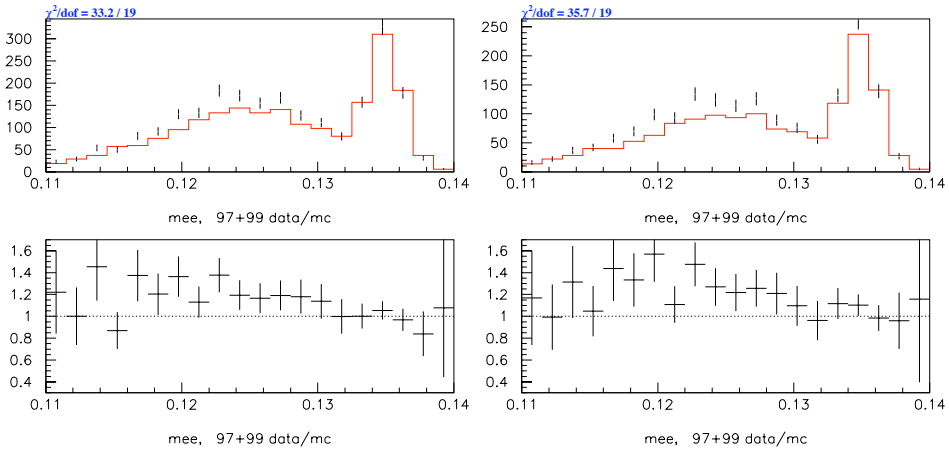

Figure 6.23: $m_{e^+e^-}$ distribution before and after tight cut on the RC and SA energies.

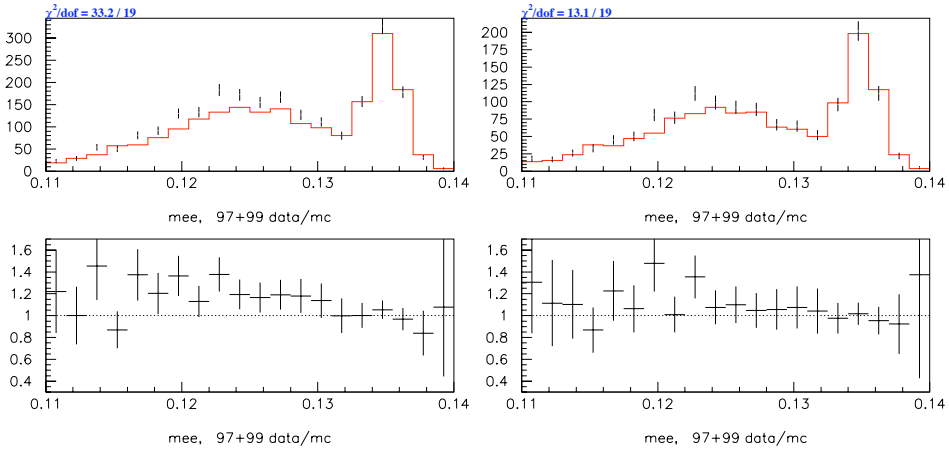

Figure 6.24: $m_{e^+e^-}$ distribution before and after a tight cut on the CA energy.

6.7.3 Varying other cuts

Varying other cuts like the kaon mass cut and p_T^2 did not change anything, only tightening the pairing χ^2 seemed to have a similar effect to the CA cut. Cutting at $\chi^2 = 3$ instead of at 20 reduced the sideband discrepancy somewhat but not quite as noticeably as the tightened CA cut, see Fig. 6.25. The plot may even be slightly deceiving since it's possible that the kaon flux got biased from the tighter χ^2 cut and raised the overall level of the MC.

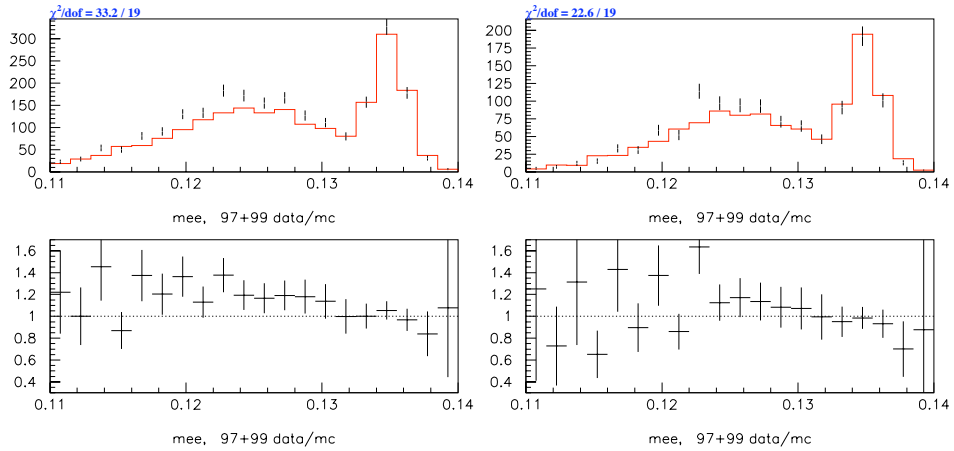

Figure 6.25: $m_{e^+e^-}$ distribution before and after tight cut on the pairing χ^2

6.7.4 Background normalization - bottom line

It was hard to argue for any conclusion on why there was this discrepancy in the $m_{e^+e^-}$ sideband. Only the tight CA cut seemed to have the effect we were looking for, suggesting that events with clusters losing energy down the beam hole (or into the CA) were not simulated well. This, of course, was not conclusive evidence and in the end we had to cover the problem with a systematic error. The study indicated no reason that the background under the signal peak was not also underestimated so the background was scaled up as quantified at the end of Section 6.6.6.

6.8 Systematic error estimates

In this section we describe potential systematic effects and the associated systematic errors on the branching ratio result. Knowing our level of ignorance about the detector setup and responses and how it affects the measurement was clearly important. Essentially all our knowledge about the detector was input in the MC, and the MC was therefore the main tool for studying systematic errors.

The aim here was to look for problems with the MC modeling which were expected to have different effects in signal and normalization. Since we have only a small sample of signal

events, we couldn't study the signal data without introducing large statistical uncertainties. We had to rely on the normalization analysis and how well the modeling was done there.

6.8.1 e^+e^- mass shift

A quick glance at the data-MC comparison of the invariant mass of the Dalitz π^0 ($m_{e^+e^-\gamma}$), left-hand plot in Fig. 6.26, shows that the mass peak in data was shifted lower than the peak in MC. Making the same plot when using only the charged vertex to reconstruct masses, no shift

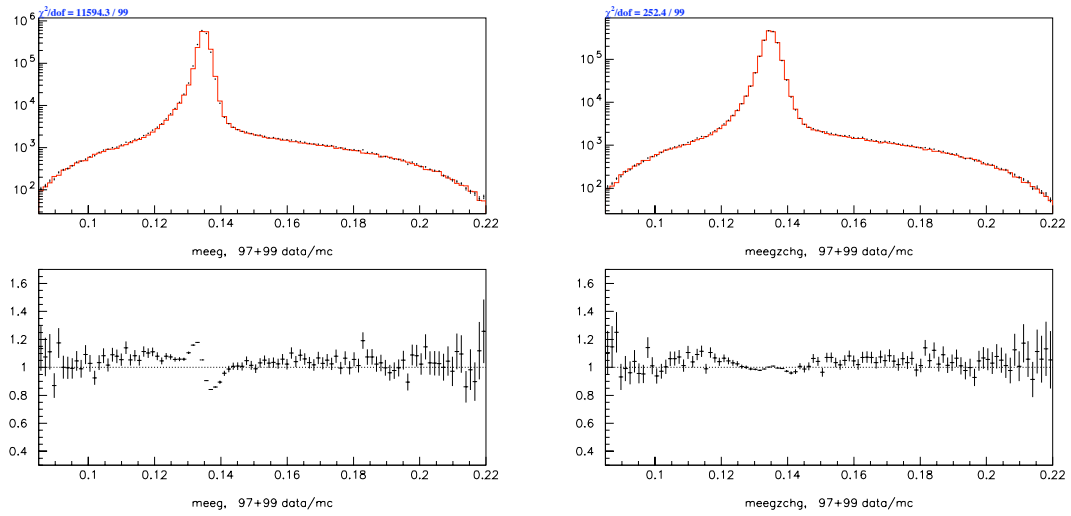

Figure 6.26: Data-MC comparison of the Dalitz π^0 invariant mass. On the left the default reconstruction is used on the right only the charged track info has been used to determine the vertex. Dots are data, the red histogram is MC. In the lower plot the data/MC ratio is shown.

was seen (right-hand plot in Fig. 6.26). This suggests that the neutral vertex information caused the mass to be shifted differently in data and MC. Gaussian fits to the peaks in data and MC showed that the data was shifted low by 0.2 MeV/c^2 away from the MC. Fits on the interval 133-137 MeV/c^2 showed the data peak to be at 134.72 MeV/c^2 and the MC peak at 134.92 MeV/c^2.

In the normalization sample this shift had no impact since the $m_{e^+e^-\gamma}$ cut was very loose, but in the signal the $m_{e^+e^-}$ cut was tight and a shift would cause trouble. Although with limited

statistics we could still make a Gaussian fit to the signal data and the signal MC to check if the shift indeed was present there too. The shift was found to be 0.15 ± 0.03, so consistent with what was observed in the normalization.

To see how this might happen we looked at the difference between the charged and the neutral vertex. This would indicate how adding in the neutral vertex information changed the vertex position and therefore $m_{e^+e^-\gamma}$. Figure 6.27 shows how the difference between the neutral

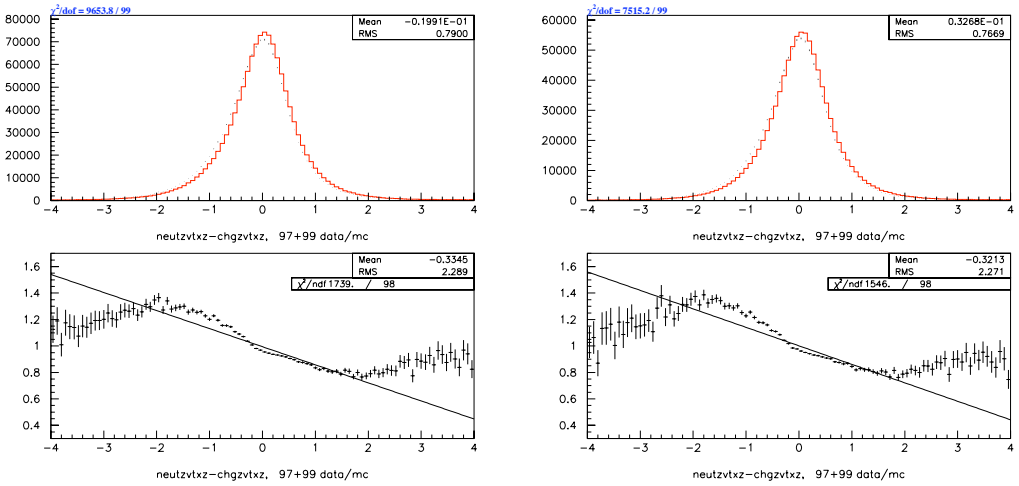

Figure 6.27: Shows the difference between the vertex z-position measured with the charged tracks and with charged and neutral tracks. The right plot has an additional cut of 75 MeV/c^2 on $m_{e^+e^-}$ reconstructed using both the charged and the neutral vertex. Black dots are data, the red histogram is MC. In the lower plot the data/MC ratio is shown.

and the charged vertex was lower in data than in MC. With an inferred vertex position further upstream the opening angle between the electron and the positron would be reconstructed smaller, thus reducing the invariant mass.

Some of the shift toward lower neutral vertex was caused by the 70 MeV/c^2 cut on the charged e^+e^--mass in the crunch. Because of the rapidly falling e^+e^--mass distribution of Dalitz decays more low neutral vertices crept into the sample than crept out at the high end. Figure 6.28 shows how the vertex difference varied with $m_{e^+e^-}$. Both data and MC had a big dip at low $m_{e^+e^-}$. The dip was well simulated but the overall shift between data and MC was

what caused the shift in $m_{e^+e^-\gamma}$. In order to remove the big dip and further study the problem we placed a cut on $m_{e^+e^-}$ at 75 MeV/c^2, reconstructed using the combined charged and neutral vertex.

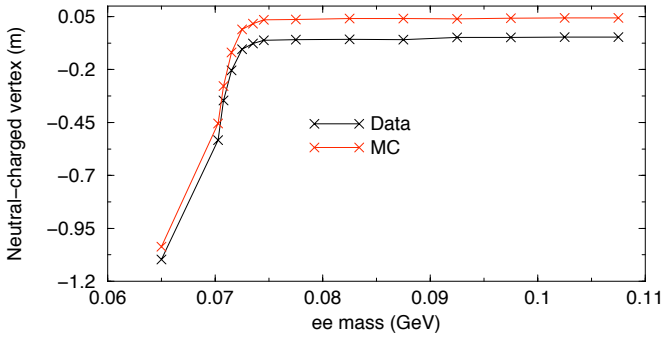

Figure 6.28: Neutral-charged vertex as a function of e^+e^--mass shown for both data and MC.

Another observation was that the vertex difference depended on both the vertex itself and the kaon momentum, see Fig. 6.29. Again we see the overall shift, but there was also a slope. Far upstream and for high kaon momenta the shift was worst, while downstream and at low kaon momenta the shift was in fact positive. An energy scale problem would have resulted in exactly such a slope. Looking at the definition for the vertex:

$$Z_{\text{vtx}} = Z_{\text{CsI}} - \frac{d\sqrt{E_1 E_2}}{m_{\pi^0}} \tag{6.1}$$

we see that the vertex is proportional to the cluster energies so an energy scale problem would have had larger effects for higher energies. The slope in the dependence on the vertex itself came from the higher geometrical acceptance for high energy events upstream.

Scaling the calorimeter block energies in data to get rid of the overall shift did not take care of the slope though. Most of the slope persisted indicating that it was not caused by a simple energy scale problem.

The (neutral-charged) vertex as a function of the z-position looks similar in MC and data except the slope was smaller in MC. If we knew the reason for the slope we might be able to figure out why it was bigger in data and from there expose the problem.

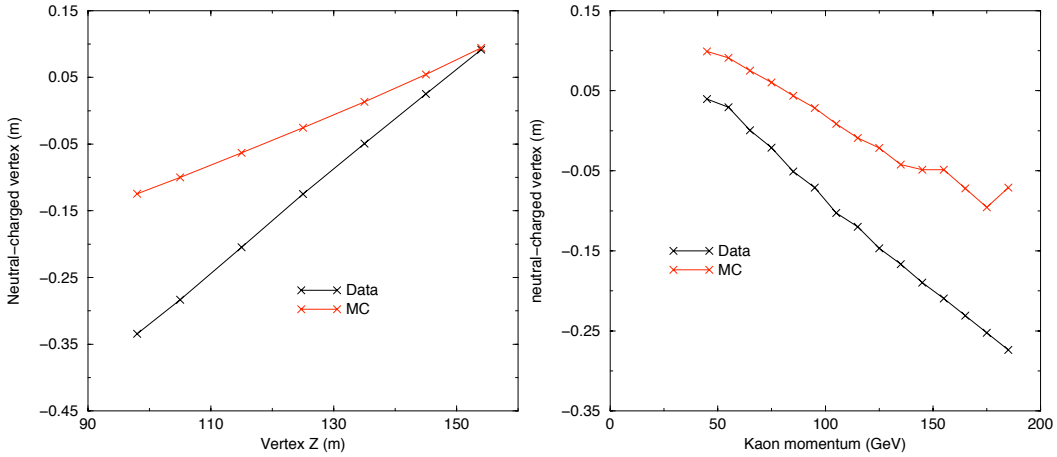

Figure 6.29: Here the neutral-charged vertex is shown as a function of the vertex Z position itself and kaon momentum.

To form the neutral vertex two measured quantities are used: the energies of the two photon clusters and the distance between the clusters. A possible explanation for the problem was that the cluster positions in the calorimeter were not calibrated well enough, unfortunately time did not permit further study of this subject and we were resigned to accept the shift.

With no solution we resorted to a fudge and assigning a systematic error to the problem and its fudged solution. The central value for the branching ratio would clearly be biased by the shift since the acceptance would be too high in the Monte Carlo, i.e. the $m_{e^+e^-}$ cut would be tighter in data than in Monte Carlo. To minimize this bias, the cut in data was lowered by 0.2 MeV/c^2 to account for the shift. The variation in acceptance when changing the cut like this in the signal Monte Carlo was 0.4%. The change was considered a systematic uncertainty to cover the fudge of shifting the cut in data.

The background was also very sensitive to the $m_{e^+e^-}$ cut so another systematic error was assigned to the uncertainty in the background estimate coming from the cut uncertainty. With the same procedure as above the background estimate increased by 5.6 event (1.9,0.9,2.8), which was considered as an additional uncertainty on the background estimate. This brought the final background estimate to 53.2 ± 11.0. The extra background uncertainty meant an un-

certainty on the branching ratio of 0.7%

Combining the two effects we notice that the two effects bias the branching ratio in opposite directions: the shift allows in more signal but also increases the background estimate. This made sure that the two effects did not add linearly but might actually cancel to some degree. Being conservative we took the maximum of the two uncertainties; a 0.7% systematic error.

6.8.2 Pairing χ^2 modeling problems

The data-MC comparison for the pairing χ^2, Figure 6.30, showed a lack of MC on the high tail and had an overall slope in the ratio. The MC includes the known four track backgrounds since these tended to have high χ^2 and contributed to the high tail in the distribution. The slope came from the resolution in the spectrometer not being modeled perfectly and was expected to be there in the signal so no bias was expected. The discrepancy in the high tail was probably some unknown background but could have been caused by mispairing effects that we didn't model well. Mispairing happened for about 4% of the normalization events, while in the signal it was rare (less than 1%). Mismodeling of the pairing would therefore not cancel in the ratio and the branching ratio measurement could potentially be biased by the full discrepancy in the high tail. As a conservative estimate of the bias introduced by cutting in this region we took the difference in the measured flux with and without the cut as a systematic error. For the winter period the difference was 0.5%, in the summer it was 0.4%, and for the 99 period 0.6%. Overall it averaged to a 0.5% systematic error.

6.8.3 Energy slope

The total cluster energy in the calorimeter (kaon energy) did not match perfectly between data and MC. A positive slope of $(0.893 \pm 0.035) \times 10^{-3}$ in the data/MC bin-by-bin ratio (Figure 6.31) indicated that the MC was slightly underestimating the kaon energy spectrum. This was a common problem in KTeV analyses, slopes in other analyses vary from $\sim +1.5 \times 10^{-3}$ to $\sim -1.5 \times 10^{-3}$. The large range from positive to negative slopes showed that it was not a

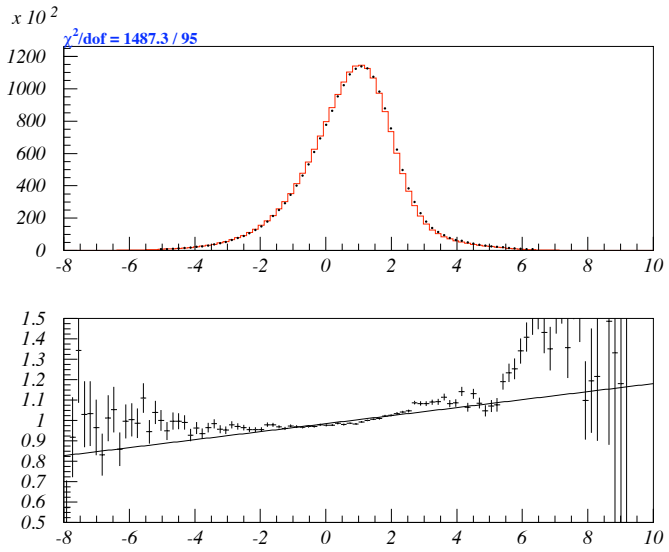

Figure 6.30: Natural logarithm of the pairing χ^2 for data (black dots) and MC (red curve). Bottom plot shows the ratio (data/MC) of the two for each bin.

problem with the kaon energy spectrum itself, since then we would have seen approximately the same slope in all decay modes. Instead the problem was probably some acceptance issue which could have been different between decay modes, in particular modes with pions in the final state seemed to have slopes very different form the pure electromagnetic final state modes. The signal and normalization of this analysis were very similar and no significant difference was expected.

As a remedy, events in the MC were re-weighted to remove the slope. The mean of the total energy distribution, $\bar{E} = 86.0$ GeV, was used as the "pivot", and the new weight of each event was

$$\text{wgt} \;=\; 1 + s\,(E_{\text{event}} - \bar{E}), \tag{6.2}$$

where s was the observed slope. For acceptance calculations the initial kaon energy distribution used in the MC was re-weighted equally in order to get the correct acceptance denominator. The denominator changed by -2.46%, making the acceptance higher in both signal and normalization. Figure 6.32 shows the effect of the re-weighting on the total cluster energy.

The re-weighting not only removed the slope in the total cluster energy distribution but

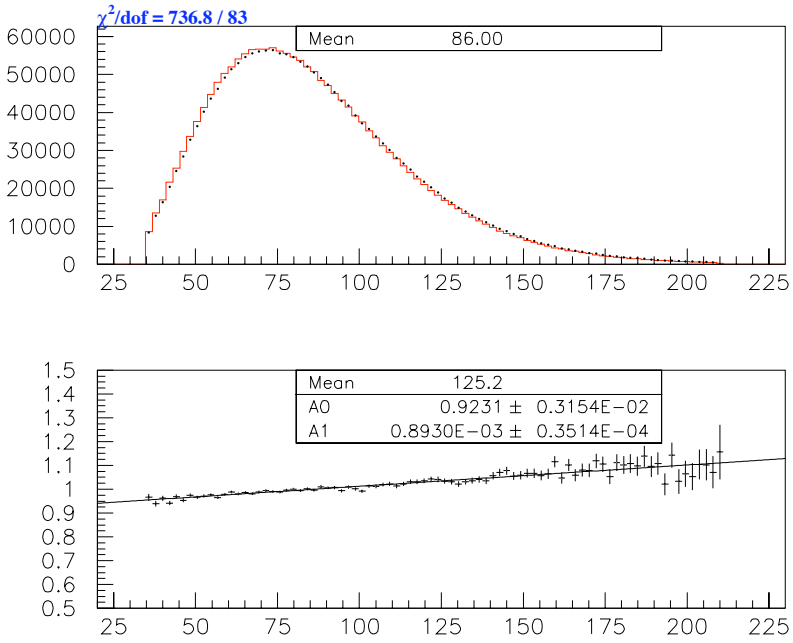

Figure 6.31: The total cluster energy had a slope. The mean of the MC kaon energy was a little too small.

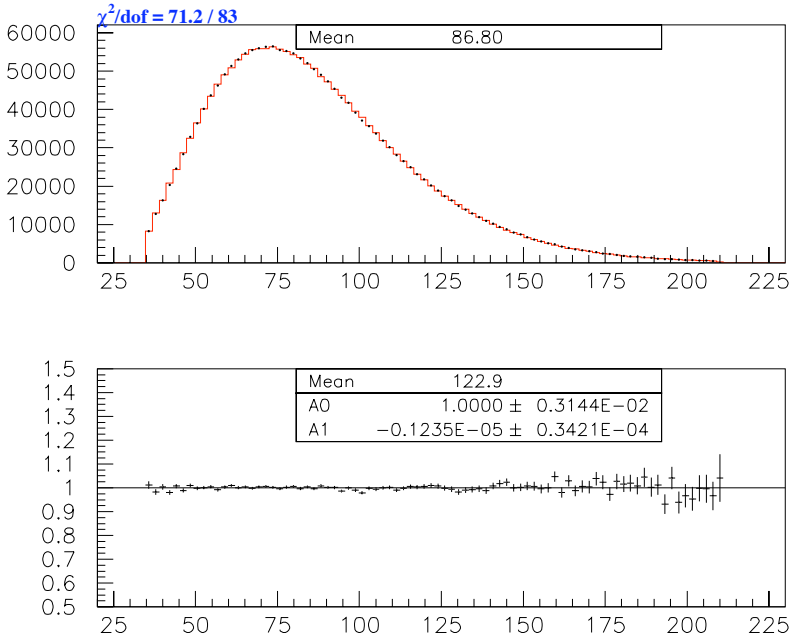

Figure 6.32: The total cluster energy after re-weighting.

in the minimum cluster energy and vertex z-position distributions as well.

The re-weighting "fudge" called for a systematic error estimate. The change in the ratio

of acceptances between signal and normalization before and after the slope correction was taken as the systematic error. It was 0.40% in 97 and 0.34% in 99, we used 0.4% in the final result.

6.8.4 e^+e^--mass edge effects in the normalization

The MC events in the Dalitz normalization sample were generated only with $m^{\rm gen}_{e^+e^-} > 65$ MeV/c^2. The PI0TOEE tag in the 2E-NCLUS crunch required $m_{e^+e^-} > 70$ MeV/c^2 which should have been tight enough to allow very little contamination into the sample from low $m_{e^+e^-}$ events that were not simulated.

In the crunch, $m_{e^+e^-}$ was reconstructed using only the charged vertex, we shall denote that by $m^{\rm chg}_{e^+e^-}$. The offline analysis used an averaged vertex (see Chapter 4 Section 4.4) to form masses, this mass we will denote $m^{\rm ave}_{e^+e^-}$.

The MC samples were not run through the same crunch as the data was, instead a crunch emulation was done where possibly the calibration was slightly different. In particular a cut on $m^{\rm chg}_{e^+e^-}$ was made at 70 MeV/c^2 which didn't necessarily result in exactly the same events being cut. This cut only had an effect in the normalization, and it was made in the region where the density of normalization events was highest. Even the slightest mismodeling of the $m^{\rm chg}_{e^+e^-}$ could have caused a bias. Figure 6.33 shows the $m_{e^+e^-}$ data-MC comparison, on the left $m^{\rm ave}_{e^+e^-}$ is plotted and on the right $m^{\rm chg}_{e^+e^-}$. There are several things to notice. The low mass tail went below 70 MeV/c^2 in $m^{\rm ave}_{e^+e^-}$, reflecting the different reconstruction method used in the final analysis (averaged vs. charged vertex). Generally the agreement was good between data and MC but in the region around 70 MeV/c^2 some disagreement was observed. This disagreement was caused by the 0.2 MeV/c^2 shift in $m^{\rm ave}_{e^+e^-}$ between data and MC. The shift was introduced when the averaged vertex was used as is discussed in Section 6.8.1. In Fig. 6.34 $m^{\rm ave}_{e^+e^-}$ is plotted with a 0.2 MeV/c^2 shift put in by hand in data. The agreement was then much better.

The distribution of $m^{\rm chg}_{e^+e^-}$ looked good but still had a small problem at the 70 MeV/c^2 cut boundary. This was the variable that was actually cut on in the data crunch and the emulated MC crunch, so the small disagreement might have caused a bias. Tightening the cut on $m^{\rm chg}_{e^+e^-}$

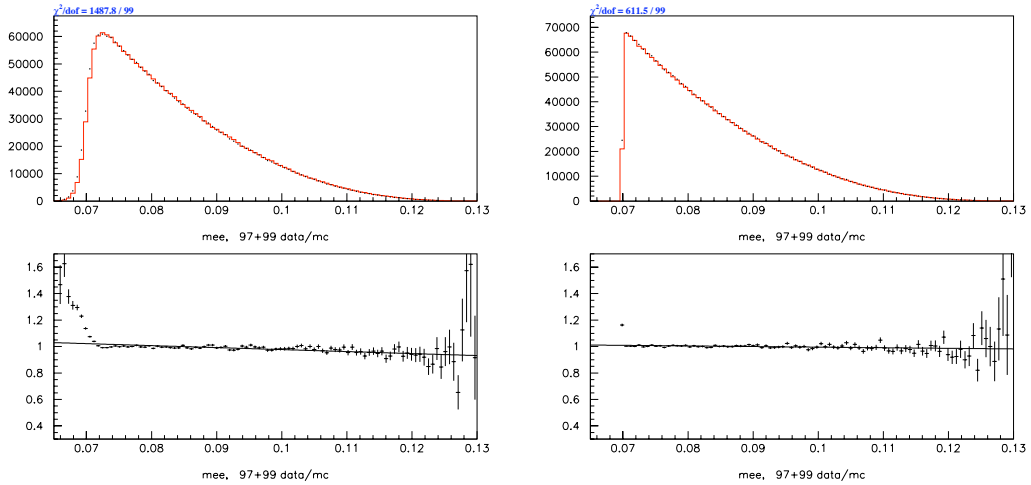

Figure 6.33: Invariant mass for the two Dalitz electrons in the normalization. On the left the mass is found using the averaged vertex, on the right only the charged vertex was used. Data is the black dots and MC is the red curve. Bottom plot shows the ratio (data/MC) of the two for each bin.

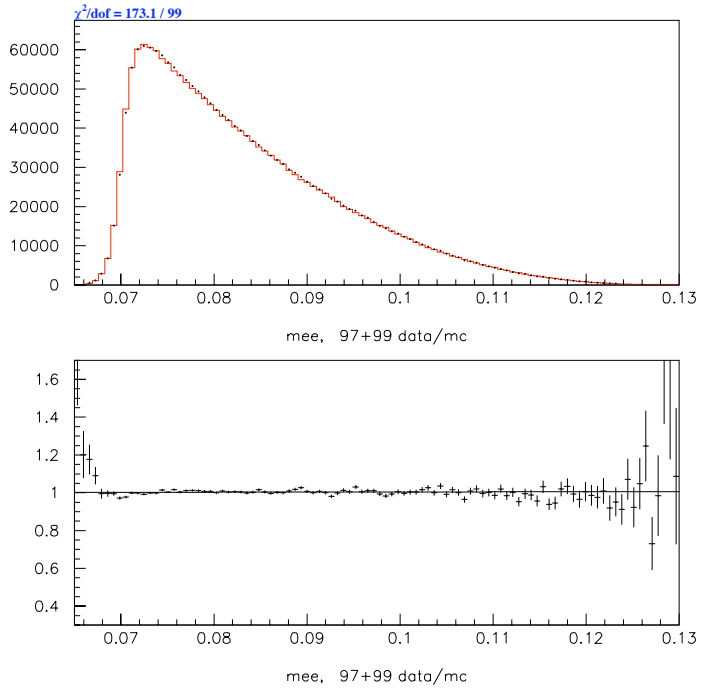

Figure 6.34: $m^{\text{ave}}_{e^+e^-}$ distribution with 0.2 MeV/c^2 mass shift fudge in data.

to 75 MeV/c^2 and observing the change in the measured flux gave an estimate of the extend of the problem. In 97 the flux changed by 0.4% and in 99 by 0.2%. An averaged 0.3% systematic uncertainty was assumed.

6.8.5 Photon cluster inefficiencies

The only topological difference between signal and normalization was the extra photon from the Dalitz decay. During the early stages of reconstruction an extra cluster from a photon made the difference between the event being a signal or a normalization event candidate, so improperly modeled photon cluster efficiencies could bias the result.

The efficiency in the calorimeter for finding clusters relied mostly on the cluster threshold. The simulation of the HCC threshold was known not to be perfect and could have caused trouble. The threshold was nominally at 1 GeV but the turn-on profile and the variations across channels were not know precisely. Cutting on the minimum cluster energy at 1.75 GeV eliminated most of the sensitivity to the simulation of the threshold. Above this cut the MC-data comparison of the minimum cluster energy was excellent, see Fig. 6.2. The measured flux did not change significantly when the cut was tightened to 2.0 GeV and all dependence on the threshold was assumed removed. No systematic error was assigned.

6.8.6 Cutting on the total invariant mass

Even though some discrepancy was seen in the data-MC comparison of the reconstructed kaon mass, Figure 6.6, no systematic error was made by cutting on it since it was very efficient in both signal and normalization. After all other cuts the efficiency for the signal mode was 99.7% in the 1997 period and 99.8% in 1999. In the normalization the efficiency was 99.8% in both periods.

6.8.7 Cutting on p_T^2

The cut on p_T^2 was very loose, for the signal it was 99.2% efficient in 1997 and 99.5% in 1999, when all other cuts were made. For normalization mode the corresponding numbers were 98.5% and 99.0%. The MC was seen to agree reasonably well with the data in this variable, Figure 6.7. The excess data in the high end tail only caused a small difference ($\sim 0.4\%$) in acceptance between normalization data and MC when the cut was made. These extra events in the data were mostly backgrounds and events where the kaon scattered in the defining collimator. In the ratio of branching ratios this would produce no significant bias.

6.8.8 Cut on extra chamber activity

The most effective of all the background cuts used was the cut on extra in-time activity in DC2. It was however a variable we knew was not very well simulated. Figure 6.35 (left) shows how a data-MC comparison of the number of extra in-time hits in DC2 looked in the normalization. On the right, using the same normalization, only events that had activity further away than 0.5 cm from the tracks are plotted, which were exactly the events that were cut from the sample. The MC sample had the known four track backgrounds included. Unsimulated

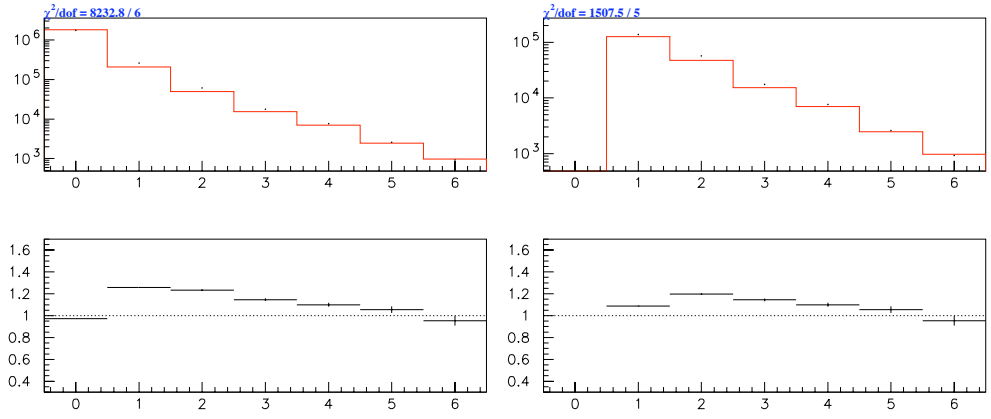

Figure 6.35: Data-MC comparisons and ratios for the distribution of extra in-time activity in DC2. On the left all events after other cuts are plotted, on the right only the events that had activity further away than 0.5 cm from the tracks are plotted.

activity was clearly present in the data, although not as pronounced in the plot of the events that were cut. Even with the clear discrepancy between data and MC we didn't expect any real bias of the branching ratio from this cut. It was based on chamber activity for which the signal and normalization were almost identical (2 tracks) and the bias would cancel in the ratio.

One thing that was different between the signal and normalization tracks was the invariant mass of the e^+e^- pair. If the cut behaved differently for very high e^+e^- masses then we might have expected a bias. We looked at the efficiency of the cut in data and MC in bins of different e^+e^- mass. Figure 6.36 shows this plot with and without backgrounds added to the MC. For

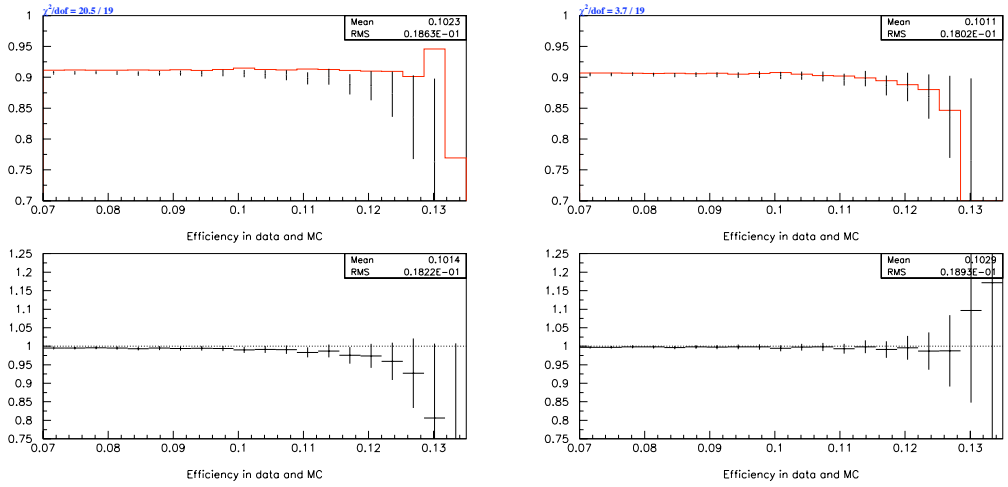

Figure 6.36: Efficiency of the extra chamber activity cut in data and MC, plotted as a function of the e^+e^- invariant mass. On the **left** no backgrounds to the normalization were included and on the **right** they were.

the most part the cut acted as just an overall scaling of about 90% in data and MC. Without the backgrounds the high mass tail did indeed look different in data and MC. Adding in the known backgrounds, though, removed any evidence of the MC failing worse in the high mass tail.

One other difference between the signal and the normalization tracks was observed. The track momenta were generally lower in the normalization, shown in Fig. 6.37 in terms of the minimum track momentum in signal and normalization. The mean momentum was about 2

GeV/c lower in the normalization. Checking for biases caused by this difference was important.

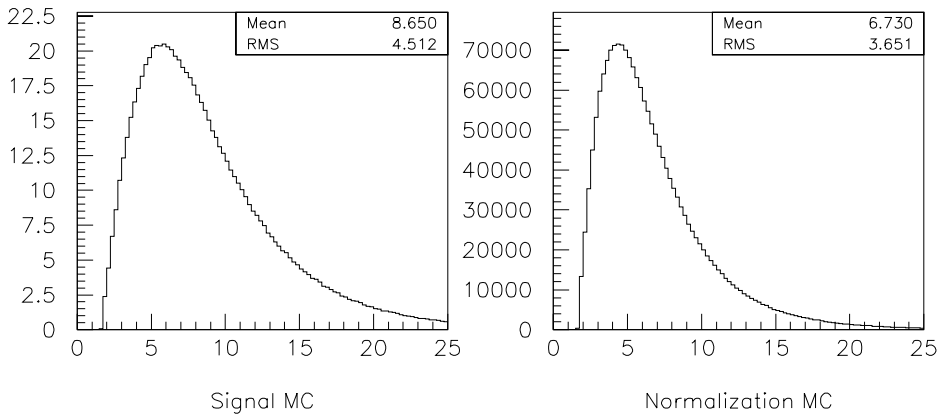

Figure 6.37: Minimum track momentum in signal and normalization MC.

We made a similar check as above in this variable, plotting the efficiency of the extra activity cut in data and MC as a function of the minimum track momentum. Figure 6.38 shows that plot with the backgrounds included. We saw no reason to be worried, the modeling was equally good across the whole range.

Overall we found no evidence that this cut would bias the result and no systematic error was assigned to this cut.

6.8.9 Energy resolution fudge

The Monte Carlo was not doing a perfect job at simulating the energy resolution in the calorimeter. This can be seen clearly in the E/p distribution of the reconstructed electrons in the normalization mode, where the width in the data was about 10% larger than the Monte Carlo predicted, see Fig. 6.39. One possible reason for this resolution problem was that the TRDs sitting right in front of the calorimeter were not simulated very well. The TRDs had a large amount of material where electrons could scatter and photons convert, and since this was not simulated too well, probably underestimating the number of interactions, it would lead to an overestimation of the cluster energy resolution.

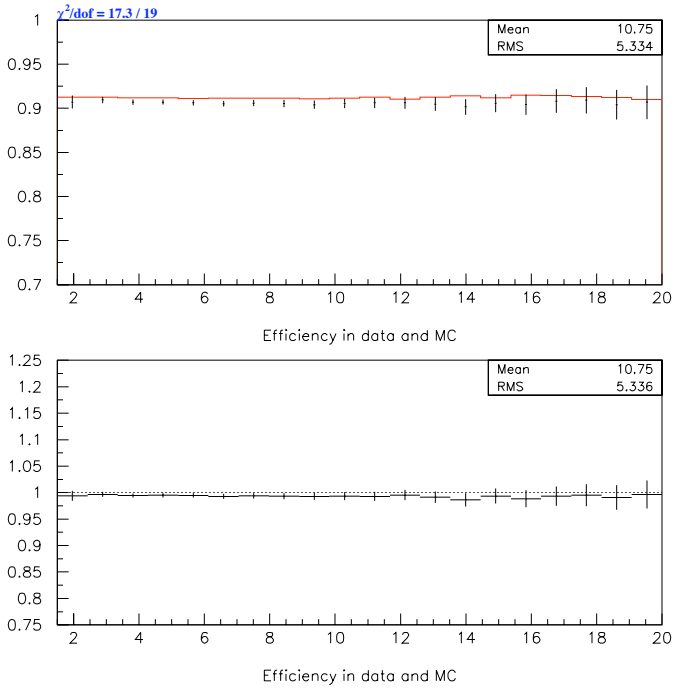

Figure 6.38: Efficiency of the extra chamber activity cut in data and MC, plotted as a function of the minimum track momentum.

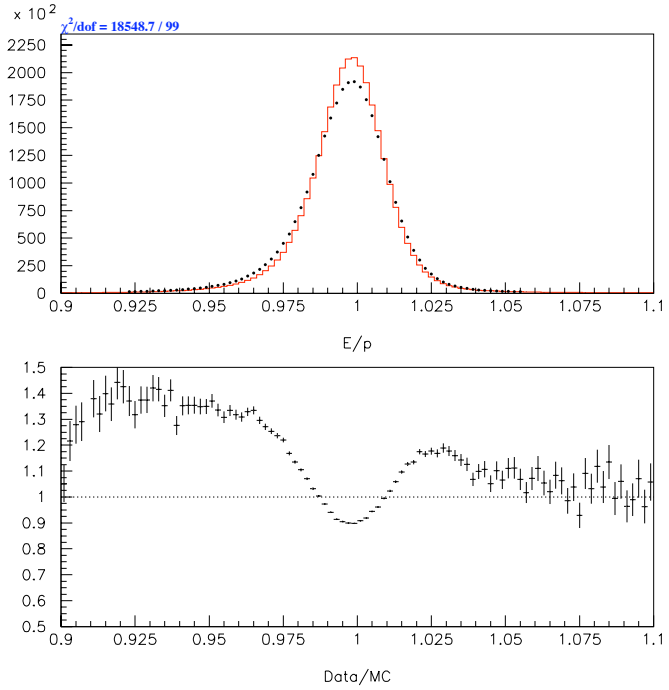

Figure 6.39: E/p distribution before applying the energy fudge in the normalization sample. The black dots are the data and the red line is Monte Carlo. Lower plot is the data/MC ratio.

In an attempt to patch the Monte Carlo an energy smearing was applied by hand. At the analysis stage each CsI crystal had its energy smeared by adding a Gaussian distributed energy to it in the following fashion:

$$E_{\text{crystal}} \rightarrow E_{\text{crystal}} \times (1 + a\, r) \tag{6.3}$$

r was a random Gaussian distributed variable with $\sigma = 1$ and a was the parameter to control the amount of smearing. In this way the resolution of E/p in the Monte Carlo could be tuned to match the data. But the resolution discrepancy showed up in other variables too. Indeed the $\gamma\gamma$-masses and the pairing χ^2 showed that the resolution was too good in the Monte Carlo, see Fig. 6.40. All of these variables got contributions from the spectrometer resolution and

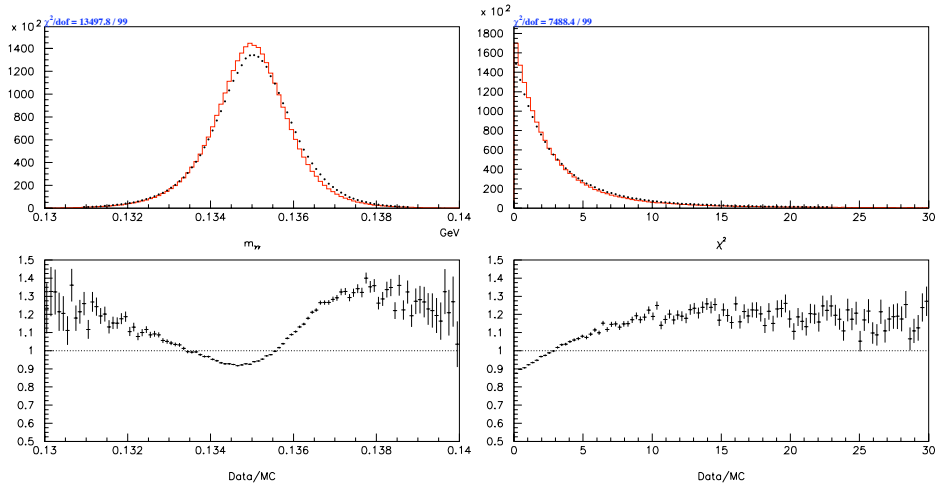

Figure 6.40: **Left**: $m_{\gamma\gamma}$ distribution before energy smearing. **Right**: Natural logarithm of the pairing χ^2 distribution before energy smearing. Both plots show normalization events in data and MC and in the lower plots the data/MC ratios in each bin.

matching their resolutions by smearing the calorimeter energies may not have been the right thing to do. A variable that only depended on calorimeter information was the neutral part of the pairing χ^2. This variable was never explicitly used in the reconstruction but was used for the purpose of finding a reasonable resolution fudge. Figure 6.41 shows the neutral pairing χ^2 for different smearing values a. There was a clear slope in the ratio between data and MC with no smearing, a clear indication that the errors used in the χ^2 were too small. As the smearing

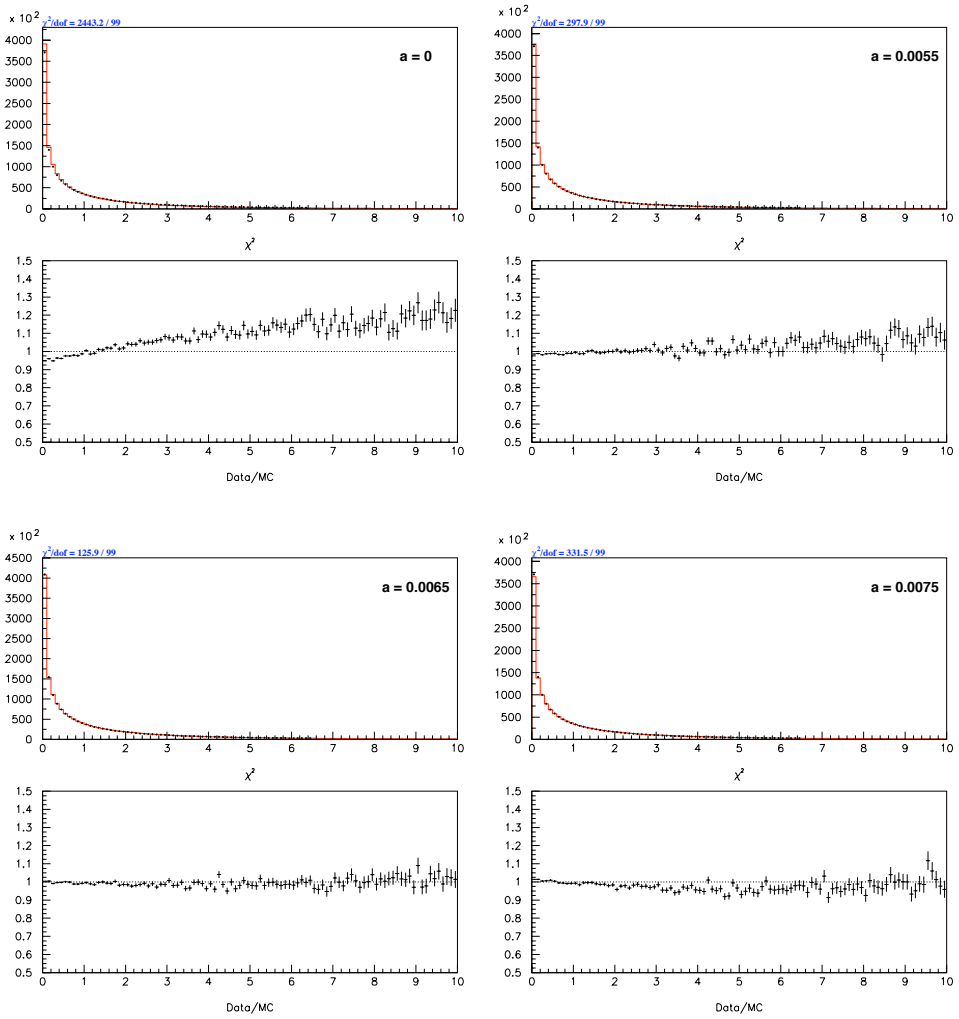

Figure 6.41: Data-Monte Carlo comparisons for the neutral χ^2 distribution with smearing parameters $a = 0$, 0.0055, 0.0065, 0.0075

was increased the χ^2 distribution got better and better. In the last plot the slope turned over indicating that the smearing was overdone. A smearing of $a = 0.0065$ was picked, based on these plots.

The other variables mentioned all suggested slightly larger smearing values, close to $a = 0.0075$. Since the smearing was put in by hand and was somewhat ambiguous, a study of the effect of varying the fudge was done. We varied the smearing from $a = 0.0055$ to $a = 0.0075$, an interval representing the uncertainty, and took the difference in the ratio of acceptances in signal and normalization as the uncertainty. The difference was $\sim 0.1\%$ and having picked any of the two other values would not have changed the branching ratio in any significant way. No systematic error was needed.

6.8.10 Monte Carlo statistics

The Monte Carlo samples only had a limited size and therefore an inherit uncertainty was present in all quantities extracted from them. This includes the signal and normalization acceptances and the background estimate.

The signal acceptances in all three run periods were extracted from a little more than $500,000$ reconstructed events giving an uncertainty of 0.14% for each acceptance. In combination (they are uncorrelated) a 0.24% uncertainty came from the signal acceptances. For the normalization acceptances the numbers were a bit higher with $\sim 600,000$ events reconstructed in the 97 periods and $\sim 1,300,000$ events in 99. The uncertainty in this case combined to 0.20%. Together the total MC statistical uncertainty in the acceptances was 0.3%.

In the background samples the number of reconstructed events were small even though most samples were many times larger than the data. The combined uncertainty on the background from MC statistics were found to be 3.2 events, which turns into a 0.4% uncertainty in the branching ratio.

6.8.11 Dalitz branching ratio and the π^0 form factor

The extrapolation from the ratio of branching ratios measured in this thesis to the final goal, the branching ratio of $\pi^0 \to e^+e^-$, used the branching ratio of the Dalitz decay and the π^0 form factor.

The Dalitz branching ratio was used to cancel the denominator in the measured ratio and the error on that measurement factors directly into the $\pi^0 \to e^+e^-$ branching ratio. We used the most current world average according to the Particle Data Group [31], $\mathrm{Br}(\pi^0 \to e^+e^-\gamma) = (1.198 \pm 0.032)\%$. The 2.7% error was then also an error on our measured $\pi^0 \to e^+e^-$ branching ratio.

The π^0 form factor was used in the Dalitz MC and played a role in finding the fraction of Dalitz events that fell in the high $m_{e^+e^-}$ tail. The tail used as the normalization. Errors from using the wrong form factor should be estimated.

The π^0 form factor is often parametrized by a simple slope away from unity

$$F(x) \;=\; 1 + ax \tag{6.4}$$

The particle data group's average for the slope $a = 0.032 \pm 0.004$ was used in the Dalitz Monte Carlo, but that number was found in a region of large space-like momentum transfer and was obtained by extrapolation assuming vector meson dominance. To account for this extrapolation we made our own estimate of the uncertainty in the form factor. It should be stressed here that we were not making a real measurement of the slope, we were just estimating what a measurement would give in order to estimate the uncertainty. A much more thorough analysis would be required to quote a real result.

A slightly different form factor $a' = a + \Delta a$ would result in a slope in the ratio of x distributions in data and MC. Assuming a and Δa were much small than 1, the slope would be $2\Delta a$. Figure 6.42 shows the $x = (m_{e^+e^-}/m_{\pi^0})^2$ distribution in data and MC. A negative slope of -0.0389 ± 0.011 was seen, suggesting $\Delta a = 0.0195 \pm 0.0055$, where the error is purely from statistics.

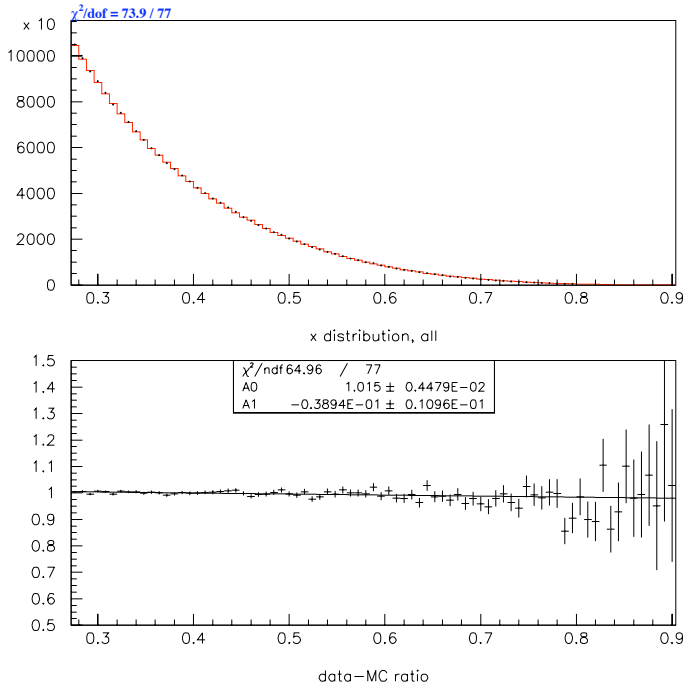

Figure 6.42: Distribution of $x = (m_{e^+e^-}/m_\pi^0)^2$ in normalization data and MC. A small slope indicates an uncertainty in the form factor.

Using a wrong form factor has two distinct effects, one is that the acceptance could be biased, the other that the extrapolation of the Dalitz x spectrum from the high tail to the full spectrum would be biased. Reweighting all events by $(1+(a+\Delta a)x)^2/(1+ax)^2$, the acceptance was found to change by about 0.1% which we could safely neglect. The same reweighting changes the fraction of events in the tail of the Dalitz distribution by 1.3% which we interpreted as the systematic error on the extrapolated result of the absolute branching ratio of $\pi^0 \to e^+e^-$.

Improved measurements of both the Dalitz branching ratio and the π^0 form factor will directly improve the systematic error on this measurement of the $\pi^0 \to e^+e^-$ branching ratio.

6.9 Calculating the branching ratio

Here we will gather the final analysis results and go through the final calculation to find the central value for the branching ratio of $\pi^0 \to e^+e^-$.

The actual measured quantity was the ratio:

$$r = \frac{\mathrm{Br}(\pi^0 \to e^+e^-,\ x > 0.95)}{\mathrm{Br}(\pi^0 \to e^+e^-\gamma,\ x > 0.2319)} = \frac{(N_{\mathrm{obs,s}} - N_{\mathrm{bkg}})/\varepsilon_{\mathrm{s}}}{N_{\mathrm{obs,n}}/\varepsilon_{\mathrm{n}}} \tag{6.5}$$

The five numbers that enter were all found in the analysis. $N_{\mathrm{obs,s}}$ was the number of observed events in the defined signal region. N_{bkg} was the expected number of background events in the signal as estimated using the Monte Carlo. ε_{s} was the signal Monte Carlo acceptance and $N_{\mathrm{obs,n}}$ and ε_{n} were the number of events and acceptance for the normalization. No external numbers enter in the calculation, since all other quantities cancel in the ratio.

For the signal acceptance there was a small twist. The decay generator generated $\pi^0 \to e^+e^-$ events with all possible values for $m_{e^+e^-}$. The signal was only the high $m_{e^+e^-}$ part of the spectrum, so the Monte Carlo acceptance was found by counting the number of events that passed all cuts and had a generated $m_{e^+e^-} > 131.56$ MeV/c^2. The number of generated events were corrected for the fraction of decays that were generated with $m_{e^+e^-} < 131.56$ MeV/c^2, that is the acceptance quoted is the acceptance for $\pi^0 \to e^+e^-$ with $x > 0.95$. The fraction above $x = 0.95$ was 89.09%, and the events below $x = 0.95$ were considered background instead of signal.

There was no estimated background for the normalization mode, since this was such a tiny fraction of the events that it could be left out entirely.

The absolute branching ratio could then be extracted from the above ratio, using the current value for the Dalitz branching ratio and correcting for the fact that only 3.19% of Dalitz events have $m_{e^+e^-} > 65$ MeV/c^2 :

$$\mathrm{Br}(\pi^0 \to e^+e^-,\ x > 0.95) = \frac{\mathrm{Br}(\pi^0 \to e^+e^-,\ x > 0.95)}{\mathrm{Br}(\pi^0 \to e^+e^-\gamma,\ x > 0.2319)} \times \mathrm{Br}(\pi^0 \to e^+e^-\gamma) \cdot 0.0319 \tag{6.6}$$

Table 6.8 summarizes the measured numbers split by the three run periods. The acceptances were almost identical in the winter and summer periods, so we combined the two periods weighted by the number of kaon decays in each. The 99 period was treated separately and the final results for the 97 runs and the 99 run was combined in the end.

	Winter	Summer	99	Total
$N_{\text{obs,s}}$	189	126	479	794
N_{bkg}	13.7 ± 1.4	9.6 ± 1.1	29.9 ± 2.7	53.2 ± 3.3
ε_{s}	$(2.955 \pm 0.004)\%$	$(2.943 \pm 0.004)\%$	$(3.153 \pm 0.004)\%$	
$N_{\text{obs,n}}$	446137	313858	1114642	1874637
ε_{n}	$(1.214 \pm 0.001)\%$	$(1.218 \pm 0.001)\%$	$(1.383 \pm 0.001)\%$	

Table 6.8: Measured values of observed event, acceptances, and backgrounds. Errors are purely from Monte Carlo statistics.

6.9.1 Measuring the kaon "flux"

Based on the Dalitz decay branching ratio and the normalization analysis, the number of kaons that decayed in the fiducial decay region can be deduced. The fiducial kaon decay region was the z-position range from 90 m to 160 m and the momentum range from 20 GeV/c to 220 GeV/c. This number of kaon decays was sometimes referred to as the "flux". The flux was found from the following expression: $F = p^{-1} \cdot N_{\text{obs,n}}/\varepsilon_{\text{n}}$, where p is the probability for $K_L \to 3\pi^0_D$ with $m_{e^+e^-} > 65$ MeV/c^2. More specifically we can write:

$$F = [\text{Br}(K_L \to 3\pi^0) \cdot 3 \cdot \text{Br}(\pi^0 \to \gamma\gamma)^2 \cdot \text{Br}(\pi^0 \to e^+e^-\gamma,\ m_{e^+e^-} > 65\text{MeV/c}^2)]^{-1} \frac{N_{\text{obs,n}}}{\varepsilon_{\text{n}}} \quad (6.7)$$

Using the numbers from Table 6.8, the recent KTeV measurement of $\text{Br}(K_L \to 3\pi^0) = 0.1945$ [32], and the fact that our MC predicts that 3.19% of Dalitz decays have $m_{e^+e^-} > 65$ MeV/c^2, we find

- $F_{\text{win}} = 1.689 \times 10^{11}$
- $F_{\text{sum}} = 1.184 \times 10^{11}$
- $F_{99} = 3.702 \times 10^{11}$

These numbers are more for reference than anything else, since they are never actually used in the analysis.

6.9.2 Crunching the numbers

Combining the winter and summer runs we found in 97 a total of 315 signal events with an expected background of 23.3 ± 1.8. The combined winter and summer acceptance was 2.950%. In the normalization sample 759,995 events were found in 97 with a combined acceptance of 1.215%. Inserting these numbers into Equation (6.5) we found

$$r_{97} \;=\; (1.581 \pm 0.096) \times 10^{-4} \tag{6.8}$$

For 99 the numbers taken directly from Table 6.8 gave

$$r_{99} \;=\; (1.768 \pm 0.086) \times 10^{-4} \tag{6.9}$$

Both numbers are quoted with only statistical errors found from,

$$(\delta r)_{\text{stat}} \;=\; r \;\frac{\sqrt{N_{\text{obs,s}}}}{(N_{\text{obs,s}} - N_{\text{bkg}})}. \tag{6.10}$$

We notice that the two results appear consistent, the two errors bars just barely touching. To be more precise we looked at the significance of the difference between the two numbers. The difference $\Delta r = (1.87 \pm 1.29) \times 10^{-5}$ was 1.4σ (statistical errors only) away from zero which did not worry us. A difference like this or bigger is expected 15% of the time.

Combining the two numbers using the statistical errors as the weights was done in the following way:

$$r_{\text{combined}} \;=\; \left(\frac{r_{97}}{(\delta r_{97})^2} + \frac{r_{99}}{(\delta r_{99})^2}\right) \Big/ \left(\frac{1}{(\delta r_{97})^2} + \frac{1}{(\delta r_{99})^2}\right) \tag{6.11}$$

and

$$\delta r_{\text{combined}} \;=\; \left(\sqrt{\frac{1}{(\delta r_{97})^2} + \frac{1}{(\delta r_{99})^2}}\right)^{-1} \tag{6.12}$$

resulting in

$$r_{\text{combined}} \;=\; (1.684 \pm 0.064) \times 10^{-4}. \tag{6.13}$$

The estimated systematic errors are summarized together with the statistical uncertainty in Table 6.9, which leads us to the final result for the measured ratio

$$\boxed{\frac{\text{Br}(\pi^0 \to e^+e^-,\;\; x > 0.95)}{\text{Br}(\pi^0 \to e^+e^-\gamma,\;\; x > 0.2319)} \;=\; (1.684 \pm 0.064 \pm 0.027) \times 10^{-4}} \tag{6.14}$$

Branching ratio uncertainties	
Statistical uncertainty	3.8%
Dalitz branching ratio	2.7%
π^0 slope parameter	1.3%
Background normalization	1.2%
$m_{e^+e^-}$ resolution	0.7%
Photon pairing χ^2 modeling	0.5%
Kaon momentum spectrum	0.4%
$m_{e^+e^-}$ cutoff in normalization	0.3%
Background MC statistics	0.4%
MC statistics for acceptance	0.3%
Internal systematic uncertainty	1.6%
External systematic uncertainty	3.0%
Total systematic uncertainty	3.4%

Table 6.9: List of uncertainties in the absolute $\pi^0 \to e^+e^-$ branching ratio.

where the first error is statistical and the second systematic. Most of the systematic error came from the uncertainty in the final background estimate of 53.2 ± 11.0.

We can now immediately translate this result into a corresponding result for the $\pi^0 \to e^+e^-$ branching ratio using Equation (6.6) :

$$\boxed{\mathrm{Br}(\pi^0 \to e^+e^-,\ x > 0.95) \ = \ (6.44 \pm 0.25 \pm 0.10(\mathrm{int}) \pm 0.19(\mathrm{ext})) \times 10^{-8}}. \qquad (6.15)$$

The first error is from signal statistics, and the other two are the systematic errors from the external sources (Dalitz parameter) and from the internal sources.

Chapter 7

Conclusions

We report a new measurement of the branching ratio of $\pi^0 \to e^+e^-$ from KTeV, a result that supersedes the old KTeV result. The branching ratio is measured to a precision of 5.1%, making it by far the most precise measurement to date. We can extract the model dependent real part of the amplitude R and compare it to the various theoretical models. Finally, we predict the branching ratios of other pseudoscalar meson decays into lepton pairs using the connection established in chiral perturbation theory.

7.1 The measured branching ratio

A total of 793 $\pi^0 \to e^+e^-$ events were observed, which included an estimated 53.2 ± 11.0 background events. Based on this sample we measured

$$\mathrm{Br}(\pi^0 \to e^+e^-, \quad \left(\frac{m_{e^+e^-}}{m_{\pi^0}}\right)^2 > 0.95) = (6.44 \pm 0.25(\mathrm{stat}) \pm 0.22(\mathrm{sys})) \times 10^{-8}. \qquad (7.1)$$

The systematic error was dominated by the uncertainty in the current best measurement of the π^0 Dalitz branching ratio.

7.2 Excess above unitary limit and Re(R)

It is common practice for theorists to predict $\mathrm{Br}(\pi^0 \to e^+e^-)$ to lowest order neglecting the radiative diagrams. The unitary limit introduced in Section 1.1 puts a model independent lower bound on this lowest order contribution to $\pi^0 \to e^+e^-$. Using the model for radiative

corrections described in Section 1.3 we can extrapolate to a result that does not discriminate radiative events. The excluded low tail in the $m_{e^+e^-}$ spectrum ($x < 0.95$) contains 10.91% of the whole spectrum, so including that and removing the -3.4% overall correction to the rate we find for the lowest order rate

$$\frac{\Gamma^0_{e^+e^-}}{\Gamma_{\text{all}}} = (7.48 \pm 0.29(\text{stat}) \pm 0.25(\text{sys})) \times 10^{-8}. \tag{7.2}$$

We can now compare with the unitary bound

$$\frac{\Gamma^{\text{unitary}}_{e^+e^-}}{\Gamma_{\text{all}}} = 4.69 \times 10^{-8}, \tag{7.3}$$

found using 0.988 for the branching ratio of $\pi^0 \to \gamma\gamma$. Expressed in terms of the unitary bound our result reads

$$\frac{\Gamma_{e^+e^-}}{\Gamma^{\text{unitary}}_{e^+e^-}} = 1.59 \pm 0.08 \tag{7.4}$$

showing that the branching ratio is more than 7 standard deviations above the unitary bound. We can also express this excess in terms of Re(R) using (1.12) and the exact model independent number Im(R)=-17.5

$$|\text{Re}(R)| = 13.5 \pm 0.9 \tag{7.5}$$

7.3 Comparison to theories

We can now compare the predictions from the models mentioned in Section 1.2 to our experimental result. In Figure 7.1 predictions from various sources are plotted together with the unitary limit and our measured result. There are essentially only two types of predictions in the plot, Vector-Meson Dominance (VMD) and Chiral Perturbation Theory (ChPT), and then a hybrid. In the VMD models the authors all agree on a branching fraction close to 6.3-6.4 $\times 10^{-8}$. Only Ametller et.al. [6] actually put an uncertainty on the prediction coming from the uncertainty of the vector-meson masses. For ChPT Savage et.al. [9] and Gomez et.al. [8] used experimental input from $\eta \to \mu^+\mu^-$ to fix counterterms in the chiral amplitude. The prediction by Gomez et.al. used essentially the same procedure as that of Savage et.al. but with a

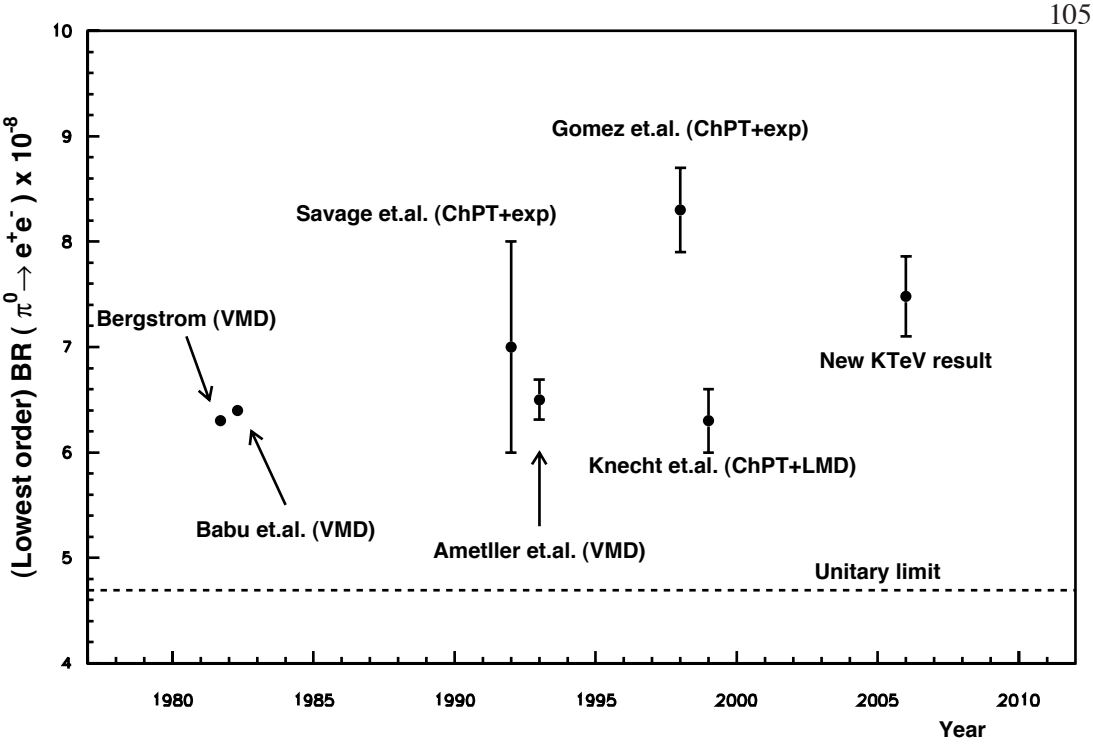

Figure 7.1: Theoretical predictions of the $\pi^0 \to e^+e^-$ branching ratio from a number of sources. Also plotted is the measured result. See the text for an explanation for the labels.

newer, more precise, measurement of the $\eta \to \mu^+\mu^-$ branching fraction. The remaining point by Knecht et.al. [10] was made using a Lowest Meson Dominance approximation, with arguments similar to the VMD form factor model, to fix the counter terms in ChPT.

7.4 Predicting other $P \to \ell^+\ell^-$ branching fractions using this measurement

As pointed out in Section 1.2, chiral perturbation theory relates all $P \to \ell^+\ell^-$ decays through a set of counterterms that are common among the decays. From our new and precise measurement of the $\pi^0 \to e^+e^-$ branching fraction we can fix these counterterms and thereby predict other $P \to \ell^+\ell^-$ branching fractions. Equation (1.12) and (1.13) holds for any $P \to \ell^+\ell^-$ decay provided we make the obvious mass substitutions:

$$\frac{\Gamma(P \to \ell^+\ell^-)}{\Gamma(P \to \gamma\gamma)} = 2\sqrt{1-\left(\frac{2m_\ell}{m_P}\right)^2}\left(\frac{\alpha}{\pi}\frac{m_\ell}{m_P}\right)^2 |R|^2 \tag{7.6}$$

and

$$\mathrm{Im}(R) \;=\; \frac{\pi}{2\beta_0}\ln\left(\frac{1-\beta_0}{1+\beta_0}\right), \qquad \beta_0 = \sqrt{1-4m_\ell^2/m_P^2}. \tag{7.7}$$

ChPT completes the picture by predicting Re(R), from Equation (1.16)

$$\mathrm{Re}(R) \;=\; \frac{1}{4\beta_0}\ln^2\left(\frac{1-\beta_0}{1+\beta_0}\right) + \frac{1}{\beta_0}\mathrm{Li}_2\left(\frac{\beta_0-1}{\beta_0+1}\right) + \frac{\pi^2}{12\beta_0} + 3\ln\left(\frac{m_\ell}{\mu}\right) + \chi(\mu). \tag{7.8}$$

Combining these and choosing a cutoff $\mu = m_\rho$ we get two solutions for the counterterms from every $P \to \ell^+\ell^-$ measurement. The authors of [8] conclude that only the overall positive solution is consistent between modes so we consider only that here.

Using the measured branching fraction for $\pi^0 \to e^+e^-$ we find $\chi(m_\rho) = 3.34^{+0.89}_{-0.96}$, and thereby derive the following branching fractions:

- $\mathrm{Br}(\eta \to \mu^+\mu^-) = 4.5^{+0.4}_{-0.2}\times 10^{-6}$ $\qquad$ $\left((5.8\pm0.8)\times10^{-6}\right)_{\mathrm{exp}}$
- $\mathrm{Br}(\eta \to e^+e^-) = 5.3^{+0.2}_{-0.2}\times 10^{-9}$ $\qquad$ $\left(<7.7\times10^{-5}\right)_{\mathrm{exp}}$
- $\mathrm{Br}(K_L \to \mu^+\mu^-) = 7.37^{+0.91}_{-0.32}\times 10^{-9}$ $\qquad$ $\left((7.27\pm0.14)\times10^{-9}\right)_{\mathrm{exp}}$

where for $K_L \to \mu^+\mu^-$ we used the prescription in [8] to include the calculable short distance contribution to the rate. All of them are in good agreement with current experimental values (values on the right), showing the power of the ChPT approach to $P \to \ell^+\ell^-$ decays.

Bibliography

[1] S.Drell. Nuovo Cim., XI:693, 1959.

[2] L.Bergström. Z. Phys., C14:129–134, 1982.

[3] Michael E. Peskin and Daniel V. Schroeder. An introduction to quantum field theory. Westview press, 1995.

[4] S.M.Berman and D.A.Geffen. Nuovo Cim., XVIII:1192, 1960.

[5] L.Bergström et al. Phys.Lett., 126B:117, 1983.

[6] Ll.Ametller, A.Bramon, and E.Massó. Phys.Rev., D48:3388, 1993.

[7] K.S.Babu and Ernest MA. Phys.Lett., 119B:449, 1982.

[8] D.Gomez Dumm and A.Pich. Phys.Rev.Lett., 80:4633, 1998.

[9] M.Savage, M.Luke, and M.B.Wise. Phys.Lett., B291:481, 1992.

[10] M.Knecht et al. Phys.Rev.Lett., 83:5230, 1999.

[11] L.Bergström. Z. Phys., C20:135–140, 1983.

[12] G.Triantaphyllou. Mod.Phys.Lett., A8:1691, 1993.

[13] J.Fischer et al. Phys.Lett., 73B:364, 1978.

[14] J.S.Frank et al. Phys.Rev., D28:423, 1983.

[15] A.G.Zephat et al. J.Phys., G13:1375, 1987.

[16] C.Niebuhr et al. Phys.Rev., D40:2796, 1989.

[17] A.Deshpande et al. Phys.Rev.Lett., 71:27, 1993.

[18] K.S.McFarland et al. Phys.Rev.Lett., 71:31, 1993.

[19] A.Alavi-Harati et al. Phys.Rev.Lett., 83:922, 1999.

[20] A.Alavi-Harati et al. Phys.Rev.Lett., 89:211801, 2002.

[21] Nick Solomey et al. Development and performance of the KTeV transition radiation detector system. Nucl. Inst. Meth., A419:637, 1998.

[22] Eric D. Zimmerman. A measurement of the branching ratio of $\pi^0 \to e^+e^-$ using $K_L \to 3\pi^0$ in flight. PhD thesis, University of Chicago, 1999.

[23] C.Bown et al. Nucl. Inst. Meth., A369:248, 1996.

[24] A.J.Malensek. Empirical formula for thick target particle production. Fermilab Report, FN341, 1981.

[25] S.Somalwar et al. Phys.Rev.Lett., 68:2580, 1992.

[26] A.Lai et al. Phys.Lett., B515:261, 2001.

[27] Mikaelian and Smith. Phys.Rev., D5:1763, 1981.

[28] Mikaelian and Smith. Phys.Rev., D5:2891, 1981.

[29] P.A.Toale. 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. PhD thesis, University of Colorado, 2004.

[30] J.R.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, University of Colorado, 2003.

[31] S.Eidelman et al. Phys.Lett., B592:1, 2004.

[32] T.Alexopoulos et al. Phys.Rev.Lett., 93:181802, 2004.