An Improved Limit on the Decay $K^+ \to \pi^+\mu^+e^-$

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Aleksey Sher

aus

Russland

Begutachtet von

Prof. Dr. Peter Truöl

Prof. Dr. Michael Zeller (Yale)

Zürich 2004

Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich auf Antrag von Prof. Dr. Peter Truöl und Prof. Dr. Ulrich Staumann als Dissertation angenommen.

Contents

Abstract

We present results of a search for the lepton family number violating decay performed by the E865 experiment at the Brookhaven National Laboratory Alternating Gradient Synchrotron (AGS). The measurement was performed during 1998 using a setup consisting of a magnetic spectrometer and particle identification detectors. No candidate events for the $K^+ \to \pi^+\mu^+e^-$ decay mode were found and an upper limit on the branching ratio at 2.2×10^{-11} was determined. Combined with the results from data collected in 1995/1996 and an earlier experiment, E777, we establish a 90% confidence upper limit on the branching ratio of 1.2×10^{-11}.

Chapter 1

Introduction

The goal of this work was a search for the lepton family number violating (LFNV) decay $K^+ \to \pi^+\mu^+e^-$. With the evidence of finite neutrino masses LFNV is no longer forbidden. Still, the Standard Model, extended to incorporate neutrino masses, predicts the $K^+ \to \pi^+\mu^+e^-$ decay at an unobservably low level. Several extensions of the standard model, however, allowed lepton family number violation at a much higher rate (as high as 10^{-10}). A Search for the $K^+ \to \pi^+\mu^+e^-$ decay is beneficial in both cases: if it is found, it signals new physics, and if not it constrains parameters of the extension models. A first search for the decay $K^+ \to \pi^+\mu^+e^-$, performed during the seventies, resulted in an upper limit of 4.8×10^{-9} [1]. After the data collected in 1995/1996 by our experiment (E865) were analyzed, and the result was combined with the results from previous searches the present upper limit B($K^+ \to \pi^+\mu^+e^-$)$< 2.8 \times 10^{-11}$ was established [2][3]. Here we discuss the analysis of the final and largest batch of the data, collected during 1998. We begin with an overview of the theoretical motivation for the $K^+ \to \pi^+\mu^+e^-$ search.

Chapter 2

Theory

Presently, the Standard Model (SM) successfully predicts experimental results at a precision level of 0.1%[4]. Although the scope of physics questions that can be addressed by the SM is far-reaching, it remains an incomplete model. One of the issues is the Gauge-Hierarchy problem which motivated the development of theoretical extensions such as Supersymmetry, Extended Technicolor, etc. These extensions often contain mechanisms for processes which violate lepton family number (LFN) conservation. However the SM either forbids the LFN violation (in the case of degenerate massless neturinos) or predicts it at an unobservably low levels (massive neutrinos). Thus an experimental observartion of a LFN violating process would be a clear indication of the physics beyond the SM.

2.1 Lepton family number

All the matter around including ourselves is constructed out of matter particles (quarks and leptons). There are six flavors (kinds) of quarks, that are arranged into three families (generations) of weak isospin doublets:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix}.$$

There are also six flavors of leptons, which are also arranged into three families of weak isospin doublets:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}.$$

In the quark sector mixing between the generations is a well established fact. However, the same mechanism did not allow the lepton family mixing, due to the assumption of a zero neutrino mass. That led to an ad-hoc rule of the lepton family number conservation. The phenomenology of this conservation law is summarized below.

There are three lepton family quantum numbers (L_e, L_μ, and L_τ) defined in the following way:

$$\begin{aligned} L_l &= +1 \text{ for } l^- \text{and } \nu_l \\ L_l &= -1 \text{ for } l^+ \text{and } \overline{\nu_l} \\ L_l &= 0 \text{ for nonleptonic particles} \end{aligned}$$

where $l = (e, \mu, \tau)$. The conservation rule of lepton family number states that the sum of each individual lepton family quantum number before and after any physical transition must be equal.

This way defined LFN conservation prohibits lepton family changing interactions such as:

$$\begin{aligned} \mu^\pm &\to e^\pm\, \gamma \\ K^+ &\to \pi^+\, \mu^+\, e^-. \end{aligned}$$

Thus the existence of LFN violating decays may uncover physical processes that lie beyond the scope of the SM.

2.2 The Gauge Hierarchy Problem and SM extensions

Along with the spin 1/2 fermions (quarks and leptons), the SM also predicts the existence of integer spin particles which are responsible for transmitting the electro-weak and strong forces. These particle fields are generators of gauge groups $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ and consists of eight gluons, the $W^\pm$ and Z, and the photon. The $W^\pm$ and Z have mass while the gluons and the photon are massless. The development of the theory which explains how mass is generated for these bosons is one of the victories in the SM. However, the mechanism which imparts masses to the bosons suggests the existence of a fundamental Higgs scalar and as we shall see below, it is the presence of this scalar particle which gives rise to the GHP.

To generate the masses for the $W^\pm$ and Z, the scalar Higgs field Φ is introduced with a corresponding Higgs potential written in terms of a mass parameter μ and quartic couplings λ as

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \frac{\lambda}{4} (\Phi^\dagger \Phi)^2. \tag{2.1}$$

The vacuum (ground) state is obtained by minimizing this potential and if the Lagrangian is re-expressed in terms of deviations from this minimum, one finds that the newly formed Lagrangian is no longer gauge invariant and contains a mass term attributed to the mass of the Higgs particle. Thus the Higgs mechanism is often called a spontaneous symmetry-breaking phenomenon.

The mass of every ordinary particle in the SM is proportional to the vacuum expectation value (VEV) of the Higgs field $\langle\Phi\rangle$. For example, quark and lepton masses are expressed as

$$m_q \sim g_Y \langle\Phi\rangle \tag{2.2}$$

while the Z and W masses are proportional to

$$g_2 \langle\Phi\rangle \tag{2.3}$$

and the Higgs mass to

$$\sqrt{\lambda}\langle\Phi\rangle. \tag{2.4}$$

Minimizing the potential above with respect to Φ, we get

$$\langle\Phi\rangle = \frac{\mu(k)}{2\sqrt{\lambda(k)}} \tag{2.5}$$

where k is the required cut-off energy for renormalizability.

One can express the renormalized mass[5] of the Higgs particle from the diagrams in Figure 2.1 as

$$\mu^2(0) = \mu^2(M) + M^2(C_1\lambda(M) + ...) \tag{2.6}$$

where M is an arbitrary chosen cut-off scale. Using Equation 2.5 and the relation[6]

$$\langle\Phi\rangle = \frac{2M_W}{g_2} \approx 246 \text{ GeV}, \tag{2.7}$$

for $\sqrt{\lambda} \sim 1$, $\mu(0)$ is on the order of 100 GeV. Choosing the Grand Unified Theory (M_{GUT}) scale $\sim 10^{15}$ as the cut-off, Eq. (2.6) can be written as

$$\mu^2(0)/M^2 = 10^{-26} = \mu^2(M)/M^2 + (C_1\lambda(M) + ...). \tag{2.8}$$

This equation says that if the ordinary particles are to be as light as they are, the parameter $\mu^2(M)/M^2$ is required to cancel the complicated series $(C_1\lambda + ...)$ up to 26 decimal places, a very unnatural condition! This point is essentially the GHP. Because the Higgs bare mass $\mu(M)$ must be tuned to provide the cancellation in order to justify the existence of a scale of $\sim 10^2$ GeV in the standard theory together with a second more fundamental scale at 10^{15} GeV, the GHP is also called the "fine-tuning" problem. Note here that GHP is linked to the introduction of an elementary scalar, namely the Higgs, because scalar mass corrections include terms that are quadratic in cut-off mass.

To deal with GHP, several extensions to the theory are formed and will be discussed in the following section.

2.2.1 Extended Technicolor

Technicolor [7, 8] (TC) is an attempt to provide electro-weak symmetry breaking via technifermions, a heavier set of fermions that carry a new quantum number called technicolor.

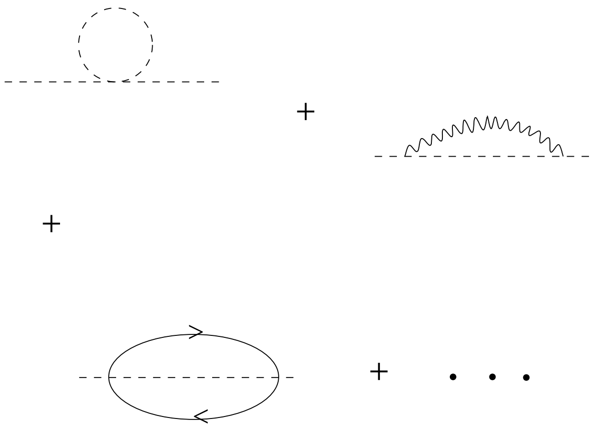


Figure 2.1: Feynman diagrams for calculating Higgs mass.

In a theory involving fermions (instead of elementary scalars such as the Higgs), the mass correction discussed previously would involve terms logarithmic in the cutoff mass and no unnatural cancellation is needed. No fine tuning is required and thus GHP is eliminated. We will now explain why technifermions and not ordinary fermions are chosen.

TC is modelled after QCD in that it is an asymptotically free gauge interaction. But now for argument's sake, let's assume the theory involves the ordinary fermions. When the interaction scale α_{TC} becomes strong, fermion condensates ($q\overline{q}$) would form and the chiral symmetry is spontaneously broken resulting in three massless pions. These condensates are the VEV and are related to the pion decay constant, f_π. Analogous to the Higgs mechanism, these pions would get "eaten up" while the $W^\pm$ and Z acquire mass and the photon remains massless. However with the value of $f_\pi = 95$ MeV, the mass of the $W^\pm$ would be

$$M_W = \frac{1}{2} g f_\pi = 30 \text{ MeV}, \tag{2.9}$$

which is three orders too small to the actual measured value of 80 GeV. In order to reconcile this mass difference, TC associates with a heavier set of fermions, the so-called technifermions. The previous symmetry breaking mechanism also applies but now the pion decay constant $f_\pi = 95$ MeV is replaced by a technipion decay constant $F_\pi = 250$ GeV. This value of F_π gives the correct masses for the $W^\pm$ and Z.

The shortcoming of TC is realized when one tries to account for the masses of quarks and leptons. Because the chiral symmetries for quarks and leptons are not explicitly broken, they remain massless. To address this problem in the framework of TC (without introducing scalars), more gauge interactions are invoked that explicitly break the chiral symmetries of quarks and leptons. This is done by combining the quarks, leptons, and technifermions into the same representation of an enlarged gauge group call extended technicolor (ETC)[9]. At high energy scales, ETC breaks the gauge symmetries and the ordinary fermions acquire mass with the presence of heavy ETC gauge bosons. Because the mass of the fermions is estimated[10] to be

$$m_{q,l} \approx \frac{g_{ETC}^2 \Lambda_{ETC}^3}{M_{ETC}^2} \approx \frac{(300\text{GeV})^3}{M_{ETC}^2}, \tag{2.10}$$

where $g_{ETC} \sim 1$ and $\Lambda_{ETC} \approx 300$ GeV, the range of M_{ETC} is about 10's of TeV.

By having all the different generations in one group, ETC also provides a mechanism for LFN changing decays. The transition between the leptons of different generations are now mediated by a horizontal ETC boson (see Figure 2.2). Because of the similarity between this decay and that of the familiar $K^+ \to \pi^0\mu^+\nu$, the decays can be compared to give an approximate mass of the ETC boson. With comparable phase space factors, the ratio of branching fractions of the two modes can be written as[11]

$$\frac{B(K^+ \to \pi^+\mu^+e^-)}{B(K^+ \to \pi^0\mu^+\nu)} = 16\frac{1}{\sin^2\theta_c}\left(\frac{g_{ETC}}{g}\right)^4\left(\frac{M_W}{M_{ETC}}\right)^4 \tag{2.11}$$

where θ_c is the Cabibbo angle, g_{ETC} and g are the ETC and weak coupling constants, and M_W is the mass of the W boson. The factor of 16 is a result of assuming that $K^+ \to \pi^+\mu^+e^-$ decay is strictly a vector (V) interaction. Re-arranging Eq. (2.11) and assuming $g_{ETC} \sim g$, one gets

$$M_{ETC} \approx 150 \text{ TeV}\left[\frac{10^{-12}}{B(K^+ \to \pi^+\mu^+e^-)}\right]^{1/4}. \tag{2.12}$$

Using the current upper limit B($K^+ \to \pi^+\mu^+e^-$) as 2.8×10^{-11}, the corresponding lower limit on the M_{ETC} is 65 TeV.

Figure 2.2: a) Feynman diagram for ETC and Horizontal Gauge mechanisms for $K^+ \to \pi^+\mu^+e^-$ decay. For ETC, replace g_H with g_{ETC}. b) Feynman diagram for $K^+ \to \pi^0\mu^+\nu$ decay.

Despite its successes, there are inconsistencies plaguing ETC. Several ETC models[10, 12, 13] have predicted rates for $K_L^0 \to \mu e$ to be above 10^{-10}, a value much higher than what is experimentally observed[14]. Normally the theory can accommodate this discrepancy by simply increasing the M_{ETC} mass scale, however, this *ad hoc* scaling is restricted by Eq. (2.10) which relates the mass of the ordinary fermions to M_{ETC}. To circumvent this problem, a scheme called "walking technicolor"[15] is invoked in which the ETC coupling constant slowly varies, allowing the ETC mass scale to be considerably increased while still generating correct masses for the ordinary fermions.

We note here that because ETC belongs to a larger group of models called Horizontal Gauge, many features from the former can be applied to the latter.

2.2.2 Other Horizontal Gauge models

The ETC or Horizontal Gauge boson which mediates the $K^+ \to \pi^+\mu^+e^-$ seen in Figure 2.2 will also couple K^0 to $\overline{K^0}$ and therefore will contribute to the $K_L^0 - K_S^0$ mass difference Δm_K. However, a discrepancy arises in that the ETC contribution is predicted to be much larger than the measured value of $\Delta m_K = 3.5\times10^{-15}$ GeV. To reconcile this disagreement, it has been suggested that there exists a symmetry which suppresses the rate of $K^0 \leftrightarrow \overline{K^0}$ transitions. This symmetry is elucidated in the following discussion.

There are Horizontal Gauge theories[16, 17, 18] which are based on a supposition of a gauge symmetry connecting different generations (families) of quarks and leptons. Specifically, the interactions between fermions of different generations are mediated by Horizontal Gauge bosons. The symmetry is illustrated by using a simple model[11] where the fermions quarks and leptons are assigned generation number G according to their family grouping. The $G = 1$ group consists of (u, d, e, ν_e) and the $G = 2$ group consists of (c,s,μ,ν_μ). The corresponding antiparticles have negative G values. Accordingly, decays are characterized by ΔG transitions depending on G values of each of the fermions in the interaction. For example the G assignments for $K^+ \to \pi^+\mu^+e^-$ fermions are

$$\begin{array}{ccccccc} K^+(u\overline{s}) & \to & \pi^+(u\overline{d}) & \mu^+ & e^- & \\ -1 & = & 0 & -2 & +1 & (\Delta G = 0) \end{array} \tag{2.13}$$

Because $G = -1$ before and after the transition, $\Delta G = 0$ and the generation number is conserved. Several other LFN violating decays which don't conserve generation number are

$$\begin{array}{cccccc} \mu^- & \to & e^+ & e^- & e^- & \\ -2 & \neq & 1 & -1 & -1 & (\Delta G = 1) \end{array} \tag{2.14}$$

and

$$\begin{array}{ccccc} \mu^- & \to & e^- & \gamma & \\ -2 & \neq & -1 & 0 & (\Delta G = 1) \end{array} \tag{2.15}$$

It is interesting to note that $K^+ \to \pi^+\mu^+e^-$ with the reverse signs also violates generation number

$$\begin{array}{cccccc} K^+(u\overline{s}) & \to & \pi^+(u\overline{d}) & \mu^- & e^+ & \\ -1 & \neq & 0 & +2 & -1 & (\Delta G = 2) \end{array} \tag{2.16}$$

and is greatly suppressed if generation mixing is small or eliminated if G is exactly conserved.

From the measured value of Δm_K one can place a constraint on the mass Horizontal Gauge boson mass (M_H) and its coupling g_H. Using an order-of-magnitude estimate[19], one gets

$$\Delta m_K \approx f_K^2 m_K \frac{g_H^2}{M_H^2} \tag{2.17}$$

where $f_K = 160$ MeV is the kaon decay constant, and m_K is the mass of K_L^0. Inserting in all values one obtains a limit

$$\frac{g_H^2}{M_H^2} \leq 2.8 \times 10^{-7}\ \text{TeV}^{-2}. \tag{2.18}$$

From this value, and eq. (2.11), replacing the subscript ETC with H (Extended TC for Horizontal Gauge), one can then estimate the upper limit on the branching fraction of $K^+ \to \pi^+\mu^+e^-$ as

$$B(K^+ \to \pi^+\mu^+e^-) \leq 2 \times 10^{-16}. \tag{2.19}$$

However $K^0 \leftrightarrow \overline{K^0}$ ($d\overline{s} \leftrightarrow \overline{d}s$) is a $\Delta G = 2$ transition and does not conserve generation number. Thus if generation mixing is not allowed *i.e.* G is exactly conserved, the Horizontal Gauge bosons do not contribute to Δm_K and the ratio in Eq. (2.18) can increase, possibly providing for a significantly larger $B(K^+ \to \pi^+\mu^+e^-)$.

2.2.3 SuperSymmetry

Like ETC, SuperSymmetry (SUSY) is introduced to deal with the subject of electro-weak symmetry breaking. However, instead of excluding fundamental scalars, the theory deals with the "unnaturalness" problem of the SM by providing cancellations for quadratic divergences. Basically, SUSY is a symmetry which relates fermions and bosons by admitting an operator Q that changes one into another

$$\begin{aligned} Q|b\rangle &= |f\rangle \\ Q|f\rangle &= |b\rangle. \end{aligned} \tag{2.20}$$

The transformation preserves all quantum numbers except spin, with a relative value of 1/2 between the two particles. This is equivalent to saying that for every ordinary particle (quarks, leptons, and gauge bosons) in the SM, there is a supersymmetric partner, generated by Eq. (2.20) that differs in spin but otherwise identical. The SUSY partners are denoted by a ˜ and are usually named by attaching *-ino* for a gauge boson, or *s-* for a fermion. Table 2.1 lists some examples of these states. One can ask why there hasn't been any observed SUSY particles in nature. This is because the symmetry is hypothesized to be broken at the scale of about 1 TeV and thus only ordinary particles have been seen.

particle	**SUSY partner**	**spin of partner**	**name**
γ	$\tilde{\gamma}$	1/2	photino
e_L	$\tilde{e_L}$	0	selectron
u_R	$\tilde{u_R}$	0	up squark
g	$\tilde{g}$	1/2	gluino
ν_μ	$\tilde{\nu_\mu}$	0	muon sneutrino
.	.	.	.
.	.	.	.
.	.	.	.

Table 2.1: Table listing some supersymmetric states.

By introducing all these new particles SUSY attempts to solve the tuning problem by providing new loops for the mass correction of the Higgs particle. As described above, every boson in a correction loop has a corresponding fermion and vice-versa, *i.e.*, for every term shown in Fig. 2.3, SUSY introduces a term of equal magnitude and opposite sign. Thus they cancel and the Higgs mass is now on the order of the square of the mass difference between SUSY fermion-boson pairs ($\approx$ 100 GeV) and the scale of symmetry-breaking for SUSY is about 1 TeV.

It is not hard to imagine that with all these additional particles, SUSY admits a mechanism for LFN changing interaction. Due to constraints set by experimental limits of decays such as $\mu \to e\gamma$, most of these mechanisms have ben ruled out. For $K^+ \to \pi^+\mu^+e^-$, the contributing Feynman box diagram is shown in Figure 2.3. Calculations[20] to probe the mass splitting between SUSY squarks or sleptons place the limit of the $K^+ \to \pi^+\mu^+e^-$ branching ratio as low as 10^{-14}.

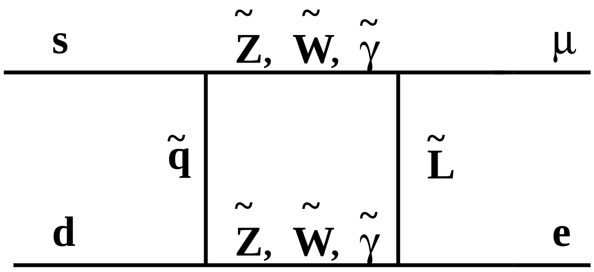


Figure 2.3: Feynman diagram for SUSY mechanism for $K^+ \to \pi^+\mu^+e^-$ decay. The supersymmetric particles are denoted by a ˜.

2.2.4 Leptoquarks

Leptoquarks are scalar or vector particles that carry both baryon number and lepton number. They were first introduced in the Pati-Salam[21] model where leptons are considered as a fourth color in an $\mathrm{SU}(4)_c$ gauge group. At a scale above the electro-weak symmetry breaking, the gauge symmetry breaks into the familiar QCD $\mathrm{SU}(3)_c$, the leptoquarks acquire mass and de-couple the leptons from the quarks. The limits on leptoquark states can be obtained directly from their production cross-sections at colliders [22] and indirectly from bounds on the leptoquark induced four-fermion interactions obtained from low energy experiments[14].

If a leptoquark couples to fermions of more than a single generation, it can induce LFN violating interactions such as $K^+ \to \pi^+\mu^+e^-$. Non-chiral leptoquarks such as the Pati-Salam leptoquarks couple simultaneously to both left- and right-handed quarks and are subjected to severe bounds from experimental constraints such as $K_L \to \mu^\pm e^\mp$. Figure 2.4 shows the mechanism for leptoquark contribution to $K^+ \to \pi^+\mu^+e^-$. Again, comparing with the $K^+ \to \pi^0\mu^+\nu$ decay, one gets

$$\frac{B(K^+ \to \pi^+\mu^+e^-)}{B(K^+ \to \pi^0\mu^+\nu)} = \frac{1}{8\sin^2\theta_c}\left(\frac{g_{LQ}}{g}\right)^4\left(\frac{M_W}{M_{LQ}}\right)^4 \tag{2.21}$$

where g_{LQ} is the leptoquark coupling constant and the M_{LQ} is the mass of the leptoquark. Re-expressing in terms of M_{LQ}, we get

$$M_{LQ} \approx 500 \text{ TeV} \left[\frac{10^{-12}}{B(K^+ \to \pi^+ \mu^+ e^-)} \right]^{1/4}. \tag{2.22}$$

If $g \sim g_{LQ}$ and using the value of $B(K^+ \to \pi^+ \mu^+ e^-)$ as 2.8×10^{-11}, M_{LQ} is approximately 220 TeV.

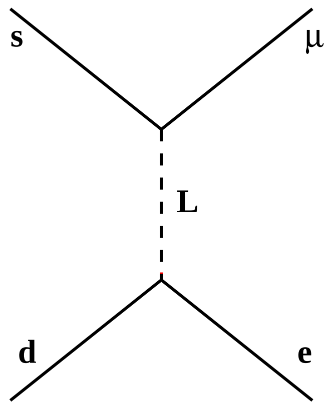


Figure 2.4: Feynman diagram for leptoquark mechanism for $K^+ \to \pi^+ \mu^+ e^-$ decay.

2.2.5 Massive neutrinos

In the quark sector of the SM, nonzero quark masses leads to quark flavor mixing in weak interactions parametrized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Similalrly with massive neutrinos, their weak eigenstates can be expressed in linear combinations of mass eigenstates allowing the different neutrino flavors to mix. This mixing would give rise to mechanisms (Figure 2.5) allowing for the decays $K_L \to \mu^\pm e^\pm$ and $K^+ \to \pi^+ \mu^+ e^-$.

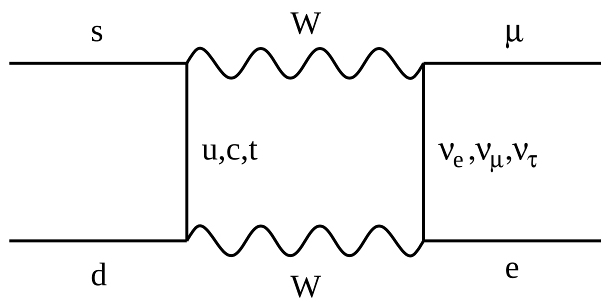


Figure 2.5: Feynman diagram allowing $K^+ \to \pi^+ \mu^+ e^-$ decay in models with massive neutrinos.

The amplitude of such process is suppressed by the low neutrino mass m_ν^2 leading to branching ratios well below 10^{-25} [23].

2.3 Other LFN violation searches

Other searches for LFN violation in kaon decays are $K_L^0 \to \mu^\pm e^\mp$[14] and $K_L^0 \to \pi^0\mu^\pm e^\mp$[25]. Both searches extend the mass reach for the same sensitivity in branching ratio as $K^+ \to \pi^+\mu^+e^-$ due to the longer K_L^0 lifetime. These decays are also allowed under the same models discussed previously (horizontal gauge symmetry, SUSY, technicolor, etc.). In addition $K_L^0 \to \pi^0\mu^\pm e^\mp$, similar to $K^+ \to \pi^+\mu^+e^-$, is sensitive to vector and scalar interactions. On the other hand, $K_L^0 \to \mu^\pm e^\mp$ probes pseudo-scalar or axial-vector couplings. Thus all three experiments complement one another. The current upper limit on the branching fraction of these null direct searches are 4.7×10^{-12} for $K_L^0 \to \mu^\pm e^\mp$ and 3.37×10^{-10} for $K_L^0 \to \pi^0\mu^\pm e^\mp$.

2.4 Outline of the experimental goal

The goal of the E865 experiment was to search for $K^+ \to \pi^+\mu^+e^-$ decay and reach single event sensitivity of 10^{-12}. To achieve it, the parent particle (K^+) had to be produced in copious amounts. The detector and the DAQ system had to handle high rates and be efficient in suppression of the possible background processes. The former and the latter requirements are often in contradiction with each other, since background would rise with intensity. In the following chapter we are giving an overview of the experimental setup, used for the $K^+ \to \pi^+\mu^+e^-$ search.

Chapter 3

The E865 Detector

The primary goal of the E865 experiment, the search for a rare decay $K^+ \to \pi^+\mu^+e^-$, heavily influenced the design of the detector. Overall detector design and particular requirements were extensively discussed in previous works [26],[27],[2]. A more detailed discussion of the apparatus and beam arrangement can be found in a recent NIM article [28]. In this thesis I will give a brief overview of the kaon beam and detector elements.

3.1 Kaon beam

Before I start describing the detector itself I should mention, that the kaon beam was created from a primary 24 GeV/c beam of protons hitting a 15cm long (transverse dimensions $5 \times 5\ mm^2$) water-cooled copper target, creating a large variety of particles, but mainly pions, protons and kaons.

The Alternating Gradient Synchrotron (AGS) at Brookhaven National Laboratory served as a source of the above mentioned proton beam, and delivered 1×10^{13} protons in a spill of 2.8s with a full cycle time of 5.1s. Before the extraction from the AGS, the proton beam was de-bunched, *i.e.* particles were smeared over the circumference of the ring to insure uniform beam intensity during the spill. After the target, the beam was momentum selected at 6 GeV/c and transported through the specially designed 28m long A2 beam line[28].

A major concern during the design of the A2 beam line was the muon halo, produced by kaon and pion decays. It was greatly suppressed by tight collimation and two foci followed by dipole magnets, which effectively swept momentum-degraded muons out of the accepted phase space. During the data taking, the beam flux in the A2 beam line was about 2×10^8 K^+, 4×10^9 π^+, and 2×10^9 protons per spill.

3.2 Detector overview

The E865 detector was located at the end of the A2 beam line that delivered the 6 GeV/c ($\pm$ 2 %) unseparated kaon beam (See Figure 3.1). The first element, placed after the beam line, was a 5m long evacuated decay region. It was estimated that roughly 5% of the kaons, produced at the target and accepted by the beam line decay in this region. Downstream of the decay region a dipole magnet (D5) separated particle trajectories by charge with negative particles going mostly to the left, and positive to the right side of the apparatus. In addition it steered the decay products from the dead region that was left in all of the detector elements for the passage of the beam.

The spectrometer consisted of four proportional wire chambers (P1-P4) and a second dipole magnet (D6).

The particle identification system consisted of two large threshold Čerenkov counters (C1 and C2). Each counter was separated into two (left and right) parts providing independent particle identification for positive and negative tracks and allowing use of different gases on the left and right sides. The left side was filled with H_2 and the right side with CH_4 (see Section 3.4.1 for details). Next element was an electromagnetic calorimeter of the Shashlik design [29]. It incorporated 600 modules (11.4cm by 11.4cm by 15 radiative length each) assembled in a 30×20 array with 18 modules in the middle, removed for beam passage. The last particle identification element was a muon system, that was located after the calorimeter and consisted of 24 planes of proportional tubes inter-spaced by iron plates.

Finally, four arrays of scintillator hodoscopes (A,B,C and D-counters) were used for the triggering purposes.

In the following sections detector elements are discussed in more detail.

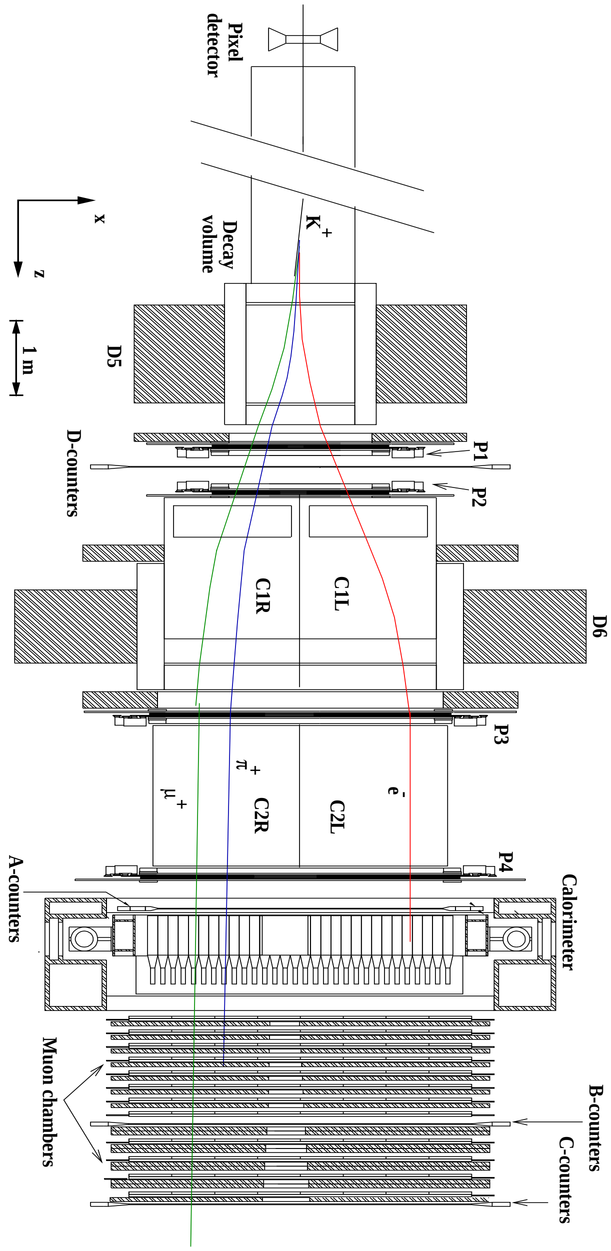


Figure 3.1: Overall view of the detector (horizontal cross section at beam height). C1/C2: gas Čerenkov counters; P1/P2/P3/P4: proportional chambers; D5/D6: dipole magnets; A/B/C/D-counters: scintillator trigger hodoscopes.

3.3 The spectrometer

3.3.1 Magnets

The first dipole magnet (D5) was used to steer charged particles away from the dead region (located in the center), which was implemented for beam passage. In addition, the first magnet (D5) separated particles by their charge (negative - on the left, positive - on the right). It had a 122 cm wide horizontal aperture and 61 cm vertical gap. Size along the beam axis was 122 cm. The field created by the magnet provided a 240 MeV/c horizontal momentum kick.

The second dipole magnet (D6) served as a spectrometer magnet. It had width of 305 cm and depth of 91.5 cm with a vertical gap of 127 cm. The D6 magnet delivered 255 MeV/c horizontal momentum kick in the direction opposite to that of D5. It bent the trajectories of the charged particles back into the acceptance region of the following detector elements (Calorimeter, and muon range system). The magnetic field components projections B_y and B_z are presented in Figure 3.2.

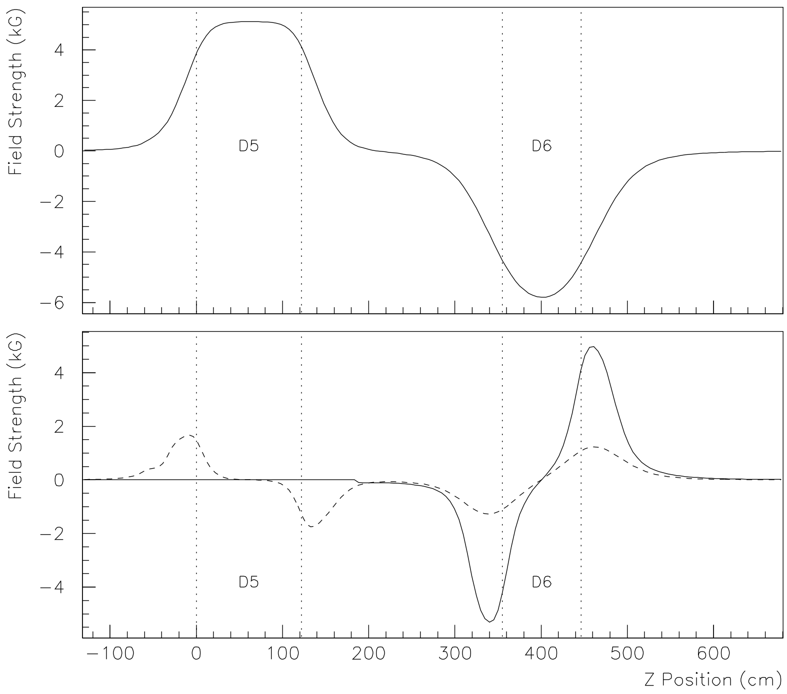


Figure 3.2: Magnetic field components B_y (top) and B_z (bottom; $y = 45$ cm: solid line; $y = 15$ cm: dashed line) along the beam direction.

3.3.2 Proportional wire chambers (PWC)

Four proportional wire chambers (PWCs) were used to determine positions of charged kaon decay products and their momentum.

The choice of proportional wire chambers was dictated by the high rates of about 50 MHz. Each chamber incorporated four sensitive planes in order to have redundancy for the unambiguous location of a hit in the high rate environment and to lessen effects of inefficiencies. A wire spacing of 2 mm was chosen as a compromise between mechanical and electrostatic limitations and requirements for spatial resolution.

The four planes had wires strung vertically, horizontally and $\pm 19.4°$ inclined with respect to vertical axis. Figure 3.3 shows the schematics of the ground plates, cathode and wire assembly.

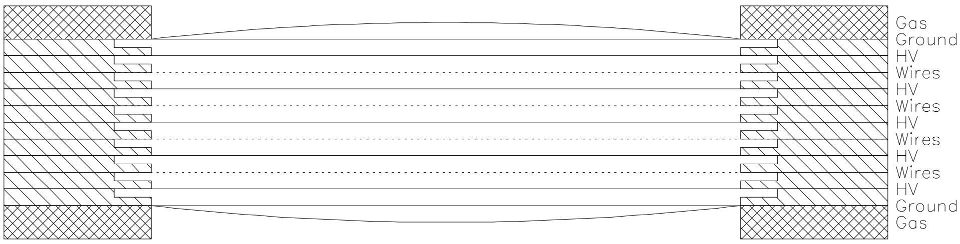


Figure 3.3: Cross-sectional view of a PWC. Dots indicate the wire planes, solid line for cathode planes, and curved lines for gas windows. The diagonal hatching represents G10 frames and the cross hatches are for support frames.

G10[1] frames supported the high voltage planes (50 μm mylar coated by 40 μm graphite) and anodes (20 μm diameter gold-plated tungsten wires). Each of the high voltage (HV) planes was divided into four areas, left, right and two small dead zones in the middle with lower high voltage, where the beam passes through at a rate of 2 GHz.

In order to equalize the electrostatic attraction between outer HV and anode planes, ground planes were placed after HV planes. Finally the outer layer was completed with gas windows (17 μm mylar and 17 μm aluminum pressed together). To make the chamber assembly air-tight, rubber O-rings were used along the G10 frames.

The gas choice was dictated by the requirement for a short drift time, asking for high electron drift velocities and high ion mobility. Aging was also a consideration and was taken into account. A 50%:50% mixture of CF_4 and isobutane was used in P1/P2;

[1] G-10-CR Natural Spauldite, Spaulding Composites Co., Tonawnada, NY 14151-0616.

and a 79.5%:20%:0.5% mixture of argon : isobutane : CF_3Br was used in larger P3/P4 chambers. The latter mixture allowed a lower operating voltage which led to reduced electrostatic forces between planes of the considerably larger chambers P3 and P4.

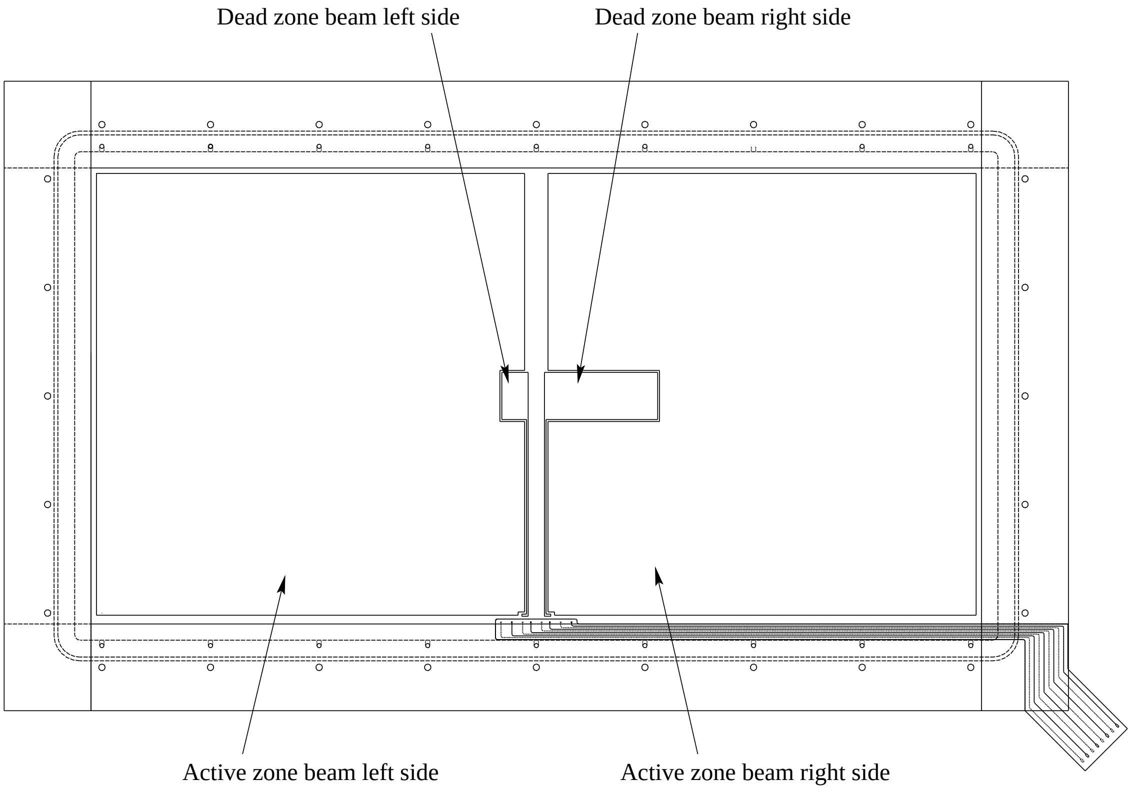


Figure 3.4: Layout of a high voltage plane (P1/P2). The HV leads connect the active regions to the HV supplies. The dimensions are given in Table 3.1.

The readout system of the PWC's (PCOS4)[28] was custom made. The readout cards, containing amplifiers, discriminators, controlled delay lines and a coincidence with an external gate were connected directly to the chambers. Each card was capable of reading out the information from 16 wires. A group of 16 cards was controlled by a motherboard (mastercard). Those motherboards were connected to the data acquisition system (DAQ) through LeCroy modules.

Table 3.1: Dimensions of active and dead regions for the PWCs.

PWC #	Active area (length × width)	Dead area (length × width)
P1	160×80 cm^2	19×9 cm^2
P2	160×80 cm^2	30×10 cm^2
P3	300×140 cm^2	55×16 cm^2
P4	320×160 cm^2	70×26 cm^2

3.4 Particle identification

3.4.1 Čerenkov counters

The threshold Čerenkov counters C1 and C2 were crucial for e^- identification and π^+, μ^+ rejection.

If a charged particle passes through a medium at a speed faster than that of light in the same medium, the particle will radiate photons. Detection of this light is the basis for allowing the Čerenkov counters to be used as particle identifiers. By taking advantage of the fact that the minimum velocity of charged particles required for producing Čerenkov light depends on the type of medium used, C1 and C2 were employed as *threshold* counters.

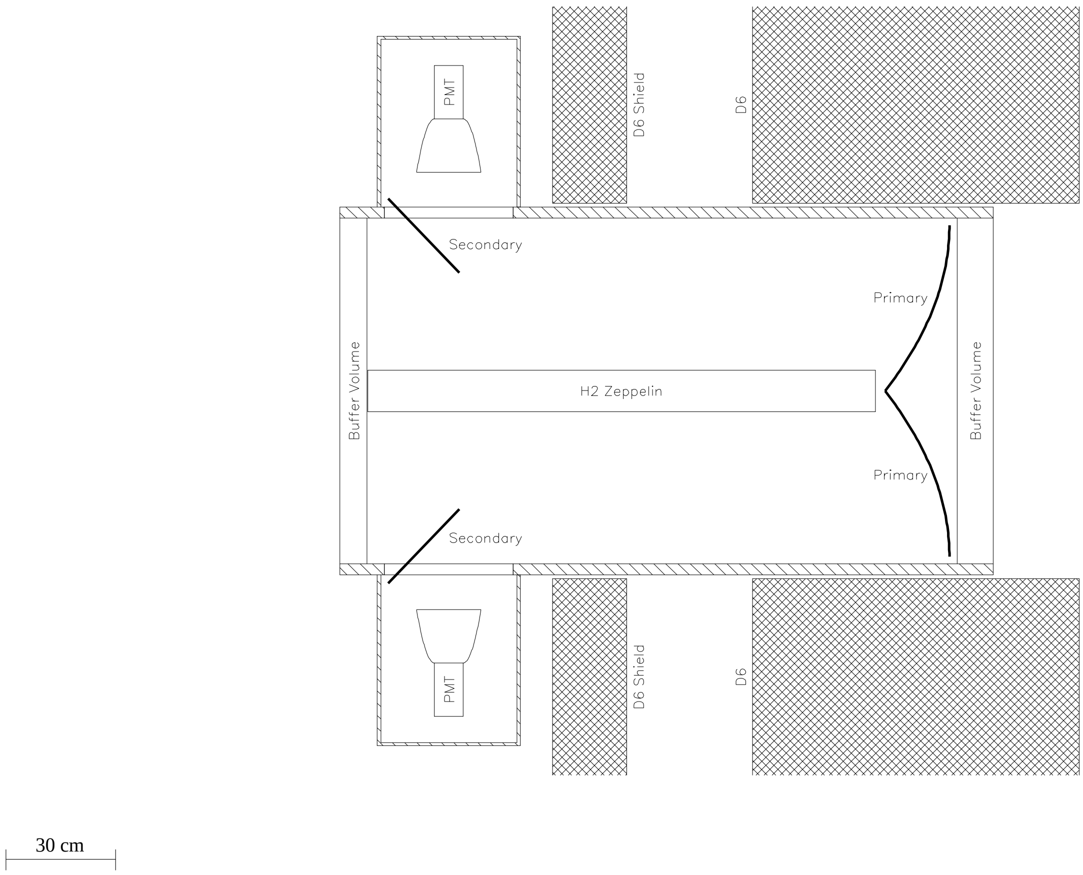


Figure 3.5: Elevation view of C1 counter on the right side. Primary and secondary mirrors directed the light towards PMTs isolated from regions of high magnetic filed. The "Zeppelin" is a H_2 filled region, where the beam passed through.

In searching for $K^+ \to \pi^+\mu^+e^-$ one must correctly identify the π^+. However, this effort is compromised by the copious misidentified e^+ tracks from $K^+ \to \pi^0\pi^+$. To minimize the misidentification of e^+ as π^+, one must detect e^+'s and veto them. To meet this demand a low threshold gas of CH_4, with Čerenkov threshold γ_T of 34.9 , is used on the right side of the counters. Any charged particle which produces Čerenkov light is identified as an e^+. Although muons with momenta greater than 3.69 GeV/c in CH_4 may produce Čerenkov light, only a small fraction of muons from $K^+ \to \pi^+\mu^+e^-$ have momenta higher than these thresholds. In either gas above, the average number of photo-electrons (p.e.'s) for positrons is around 5.8. Translating this into Poisson statistics, the average efficiency per chamber for positron detection (veto) is 99%. On the left side of the counters, where the aim is to correctly identify e^-, the main concern is π^-'s from $K^+ \to \pi^+\pi^+\pi^-$ decays. To sufficiently limit the misidentification of a π^- as an e^-, a high threshold gas medium such as H_2 at atmospheric pressure is used. With a corresponding $\gamma_T = 60$ the momentum threshold for producing Čerenkov light from muons and pions is 6.34 GeV/c and 8.37 GeV/c, respectively. With an average yield of 2.3 p.e.'s in H_2, the corresponding expected electron detection efficiency is 89%.

To separate two different gases on the left and right sides, a septum splits each counter in the middle ($x = 0$). To avoid unwanted particle production, mirror portions that overlap the beam are removed. In addition, gas balloons (zeppelins) filled with H_2 gas are inserted in the beam region on the right side (the side containing the heavier gas) of each counter to minimize beam-gas interactions. However, care is taken to prevent the zeppelins from blocking optical paths, which would cause loss of detection efficiency. An additional background source may come from delta rays produced downstream of C1 which give correlated signals in C1 and C2. To offset this effect the counters are placed inside the magnet D6 which sweeps aside these low momentum electrons.

C1 is divided into four quadrants with each quadrant again divided into two parts. A total of eight photo-multiplier tubes (PMTs) view the counter, one for each part. Mirrors are used to focus the Čerenkov photons onto PMTs to maximize the collection efficiency. Primary mirrors situated at the downstream end of the counters reflect the photons to secondary mirrors which then focus the light to PMTs that are located in regions outside of high magnetic fields. The primary mirrors are cylindrically shaped and arranged in a "venetian blind" style. To avoid degradation of the momentum resolution, one needs to minimize the amount of mass brought into the spectrometer. Thus the mirrors are made of aluminized mylar, coated with MgF to prevent oxidation and supported by a light but stiff Rohacell foam. Figure 3.5 shows an elevation view of the C1 counter.

C2 is also divided into four quadrants with each quadrant subdivided into two vertical halves (see Figure 3.6). The half closer to the $y = 0$ axis has a larger acceptance and is viewed by four PMTs while the outer half is only viewed by two PMTs. Because C2 is located outside the regions of strong magnetic field, the PMT tubes are arranged so that only primary mirrors are needed.

The initial (regular) readout electronics consisted of ADC and TDC channels[2]. However, sometimes, an unrelated preceding pulse could prevent the discriminator and subsequently the TDC to fire. It could happen if, by the time of the pulse of interest, the preceding pulse tail was still above the discriminator threshold (about 10 mV). To adress that issue, a second set of TDC channels with discriminator thresholds set above one

[2]This paragraph was added to the text after it was submitted to the University of Zurich archives

photoelectron (above 30 mV) was installed in 1997 (high threshold TDCs). And finally, in anticipation of the high intensity during the 1998 run, additional auxiliary electronics were implemented for additional degree of control. Auxiliary electronics consisted of an additional, more tightly timed ADC and TDC channel with another ADC channel gated to detect signals occuring prior to the main trigger.

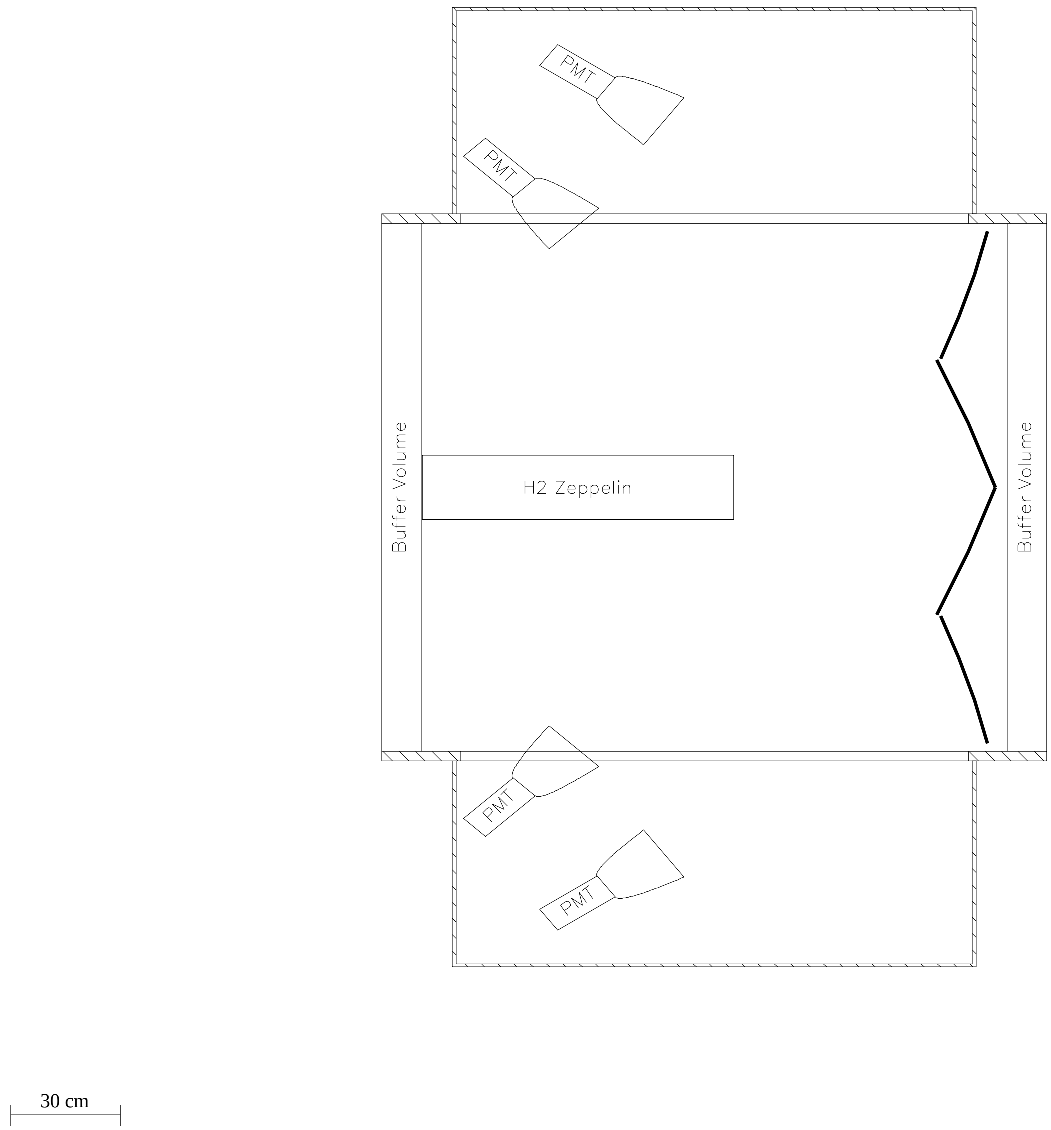


Figure 3.6: Elevation view of C2 counter on the right side. The left side is identical without the zeppelin.

3.4.2 Electromagnetic calorimeter

The electromagnetic (EM) calorimeter [29] is used to measure the energy of electrons, positrons, and photons. In addition, it is used in the first-level trigger to signal a possibility of an interesting three charged track event. This detector consists of 600 shower modules, arranged in a 30×20 rectangular array as shown in Figure 3.8. Eighteen modules in the center are removed to avoid interactions with the beam.

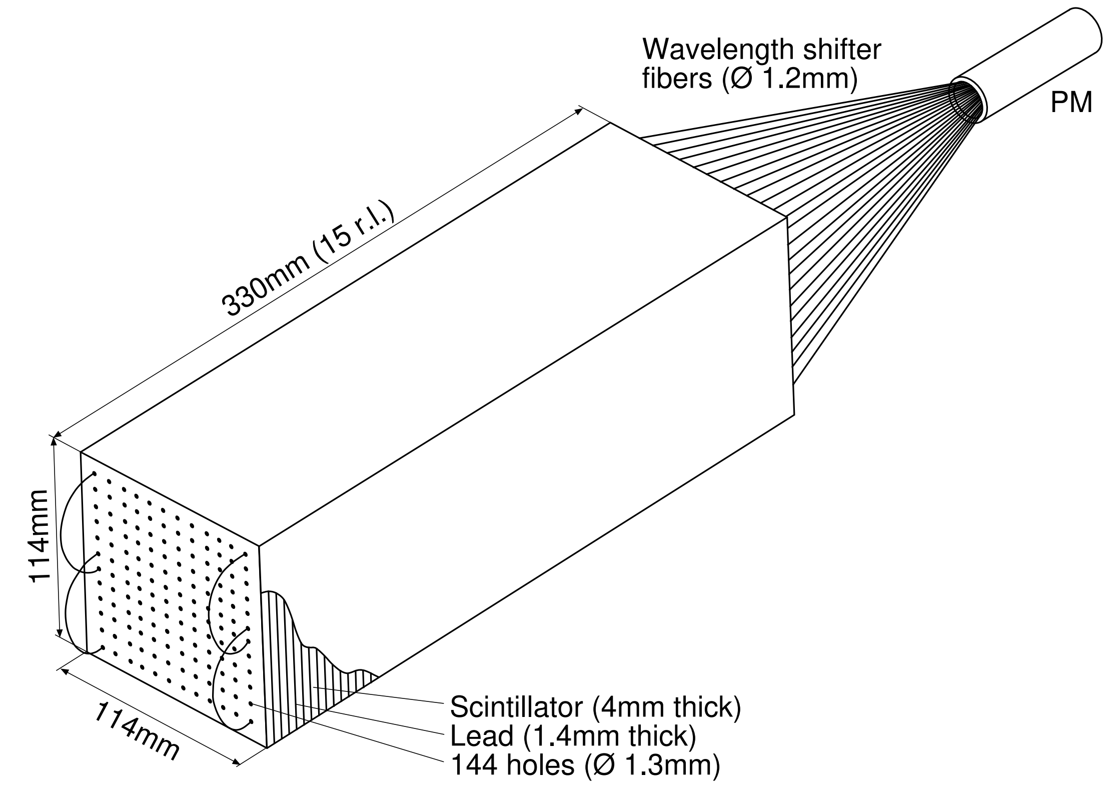


Figure 3.7: Calorimeter module.

Each module, displayed in Figure 3.7, is 33 cm long with alternating slabs of 1.44 mm thick Pb and 4 mm thick scintillator. The 60 slabs of material amount to 15 radiation lengths. In it's cross-section, each module is a square of 11.4×11.4 cm^2 corresponding to 2.5 Moliere radii. Light is collected using wavelength-shifting fibers. Starting from the back, these fibers run towards the front through drilled holes in the slabs, making a loop at the front and return to the back where both ends of fibers are coupled to a phototube. The timing resolution for shower counter signals is better than 1 ns while the energy resolution is about 6.8% $\sqrt{E(\mathrm{GeV})}$. With respect to $K^+ \to \pi^+\mu^+e^-$, the EM calorimeter's most critical role is separating electrons from pions and muons. This is accomplished by determining the ratio of the momentum of a particle to its energy (E/P). Because electrons loose all their energy in the modules, while all muons and most pions pass through, E/P values of electrons are close to unity, while those for pions and muons are lower. This feature is one of the important tools used for distinguishing these particles. However, pions can also interact strongly with the calorimeter material, creating hadronic showers yielding E/P tails which overlap the region of electron response, resulting in misidentification of pions as electrons. In such cases, one must rely on other means to separate these particles, *e.g.* using differences in their lateral shower spread (see Section 7.4.2.)

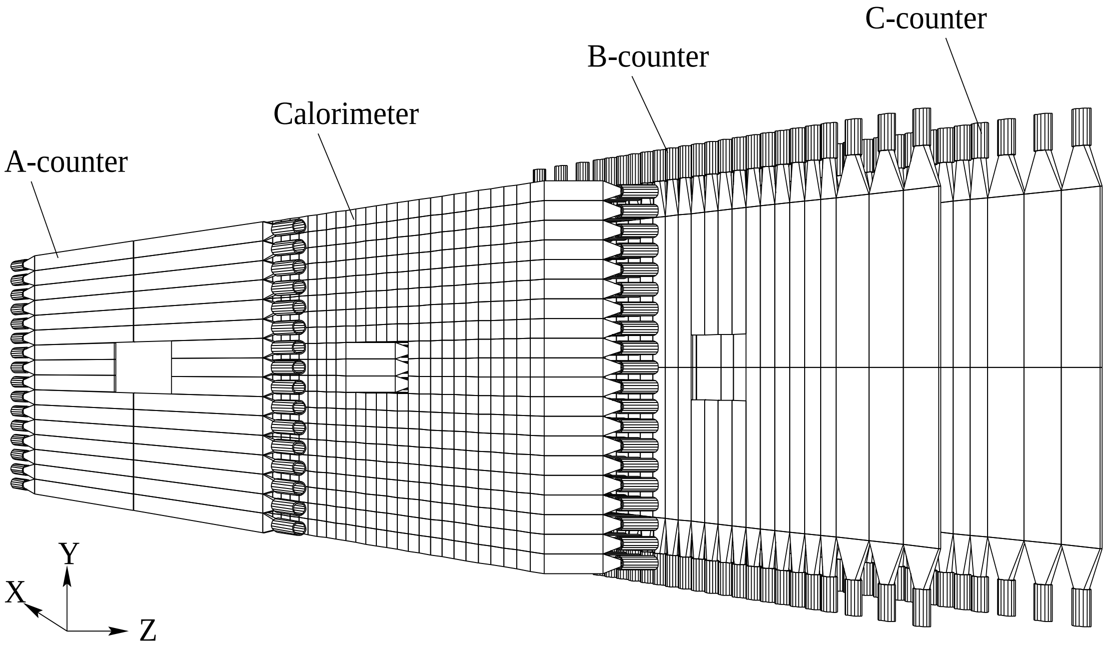


Figure 3.8: Positions of the shower counter and A,B,C - hodoscopes. For a clearer view of the shower counter, A-hodoscope is moved upstream by 1.5m in the picture.

3.4.3 Muon range system

Figure 3.9 shows the elevation view of the muon stack consisting of twelve XY chambers (M1-M12) with a steel plate behind each chamber. Each plate has a thickness of 5 cm except those behind M9 through M11 which are 10 cm thick. For muon triggering purposes the B hodoscope is situated behind M8 and the C hodoscope is placed beyond the last steel plate.

Figure 3.9: Elevated view muon stack components: chambers, steel plates, B and C hodoscopes.

Each X and Y chamber consists of 8 and 10 modules respectively where the Y modules are divided into left-right sides, five on each side. The X view originally included a ninth module in the center which coincided with the beam. However, due to the intolerable secondary particle flux from beam interactions, this module was removed. The modules contain proportional tubes with wires running vertically in modules of X chambers or horizontally for Y chamber modules. There are 32 wires (channels) per module spaced 1.3 cm apart and the tube cell containing the wires have dimensions of 1×1 cm^2 with 3 mm thick aluminum walls. The cells in each module are staggered between two planes to offset the possibility that a particle may hit the walls dividing the cells contributing to chamber inefficiency (see Figure 3.10). The gas mixture used is the same as for P1 and P2.

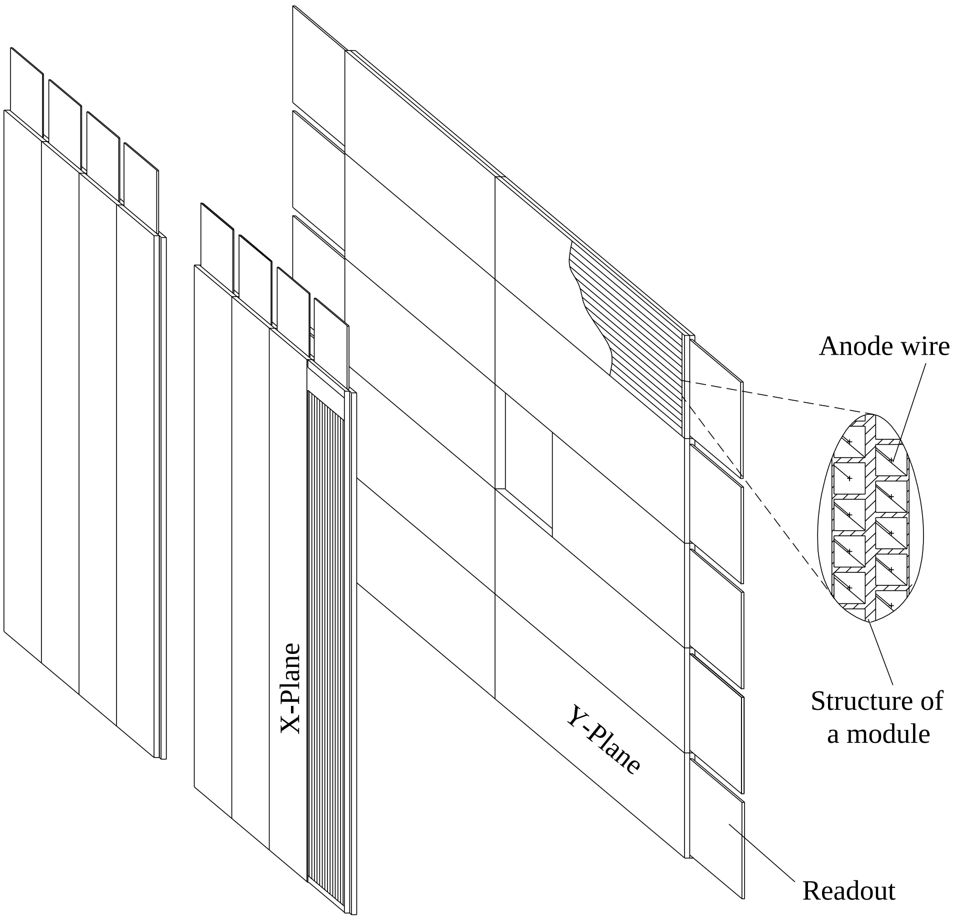


Figure 3.10: Display of X and Y muon planes. Details of the tube cells are also shown.

The main purpose of the muon stack is to separate muons and pions. This is accomplished by the fact that muons lose their energies only through ionization. With their typical momenta, they retain enough energy to range beyond the last steel plates. Pions, on the other hand, may interact strongly with the steel plates preventing them from penetrating in the range stack. By counting the number of muon chamber hits from a track and comparing this result to the number of hits expected from its momentum, one can separate muons from pions. Because an identified muon track must hit the B counter, it is required to have a minimum momentum of 800 MeV/c.

3.5 Scintillator hodoscopes

Four scintillating hodoscopes, A, B, C, and D are employed for triggering. The A hodoscope consists of 30 horizontal scintillator slats arranged in two sets, 15 each on the left and right sides (see Figure 3.11). Each slat is 1 cm thick and 11.4 cm wide to exactly overlap a row of shower counters. The 20 D slats are also arranged horizontally with 10 on each side (see Figure 3.12). The width of each slat is 8 cm. While the A counters are placed in front of the calorimeter, the D counters are situated upstream between the first two PWCs. In conjunction with the calorimeter, the A and D hodoscopes are used at the first level of the trigger hierarchy.

The arrangements of the B and C counters, each with 48 slats of scintillators, are identical. The slats are aligned vertically, with 24 on top and bottom (see Figure 3.13). Each slat is again 11.4 cm wide except for the three outermost paddles which have twice the width. Both counters are embedded in the muon stack and can be used for muon triggering.

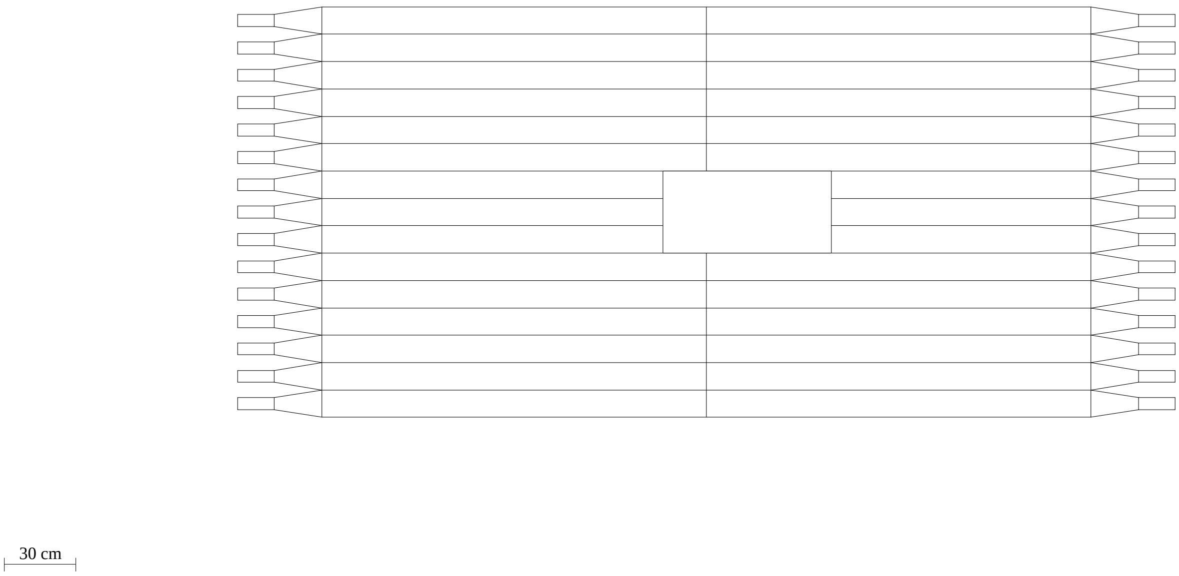


Figure 3.11: Layout of the A hodoscope looking downstream.

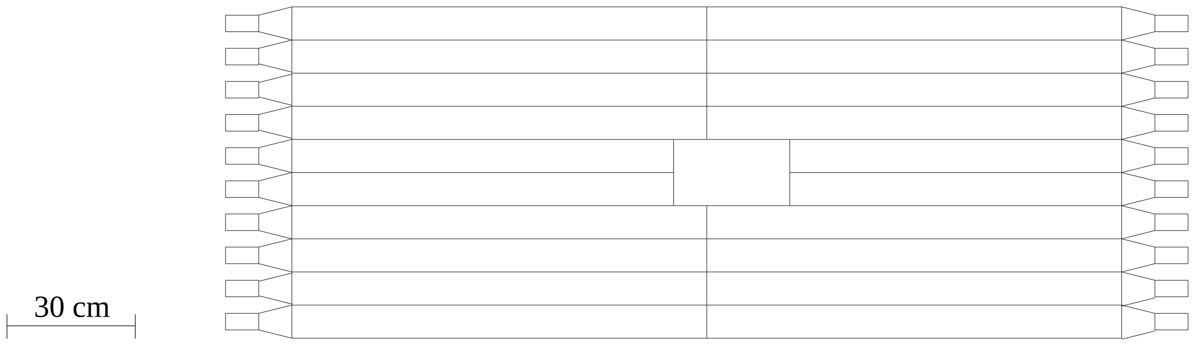


Figure 3.12: View of the D hodoscope looking downstream.

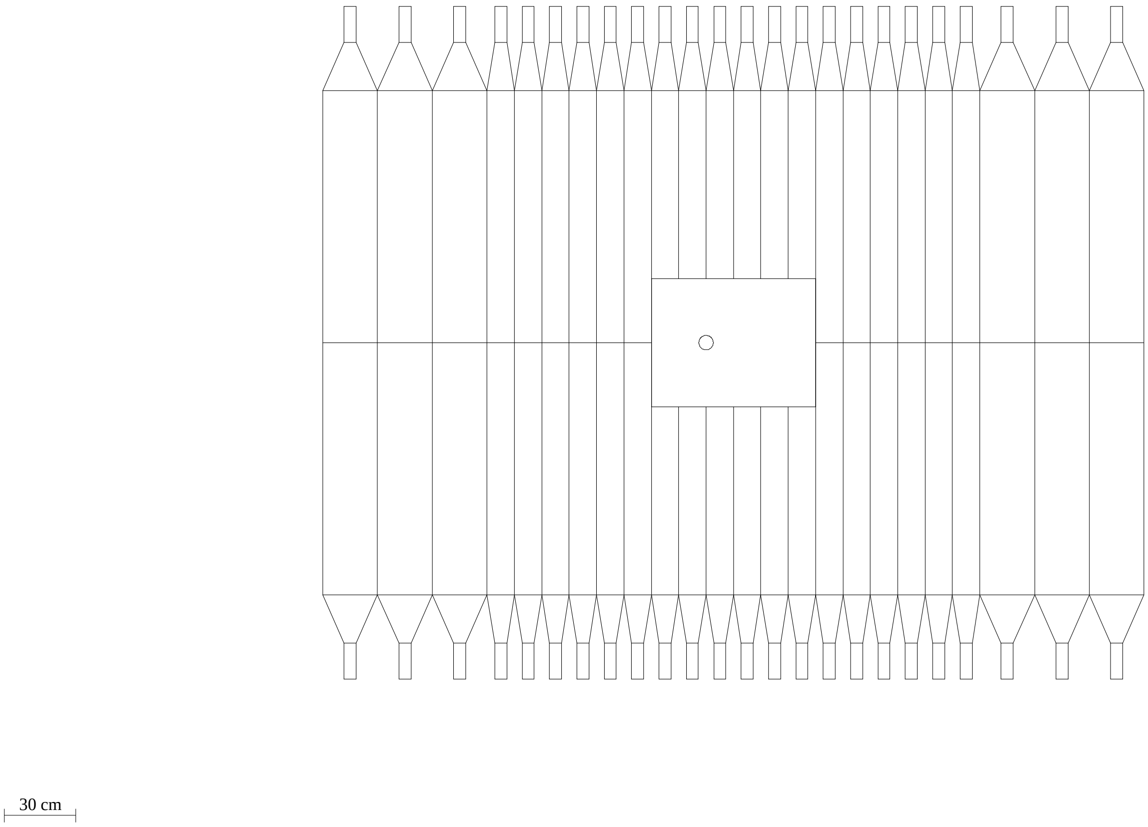


Figure 3.13: Downstream view of the B and C hodoscopes.

3.6 Trigger system

The trigger system was primarily designed to effectively filter possible $K^+ \to \pi^+\mu^+e^-$ decays, while rejecting background. During the data taking, the average rate of charged particles through the active detector parts was about 50 MHz. The data taking rate was limited by the recording speed. Data, after online triggering had to be saved. E865 used an 8mm Exabyte tape drive with a writing speed of about 400 Kb/s. With an average event size of about 2 Kb, only a few hundred events could be recorded per second.

While information was read out the experiment could not take data. Resetting PCOS4 latches or ADC/TDC modules takes up to 2μs, adding up to a total readout time of up to 15 μs. It is obvious that one would like to avoid reading out the data, when the event is not consistent with $K^+ \to \pi^+\mu^+e^-$.

That, of course, called for a fast and efficient trigger, that would arm the DAQ for readout. A multilevel trigger was constructed. The earliest and fastest trigger level would reject most of the background, while more sophisticated and slower choices were made by higher trigger levels.

3.6.1 $K^+ \to \pi^+\mu^+e^-$ trigger summary

The four trigger levels used for $K^+ \to \pi^+\mu^+e^-$ data taking had the following functions:

- **T0** : the first level trigger selected three track events by requiring at least three coincidences between the D- and A-hodoscopes and the shower counter (D*A*SH). The combination of hits in the D and A counters and the shower counter was required to be consistent with a predetermined matrix of a three track decay kinematics (one negative and two positive tracks coming from the decay volume). In order to maintain high efficiency only two hits were required in the D-counter (one on the left side, and one on the right side). The T0 trigger had a response time of 170ns after the particle would hit the A -counter, and an output rate of about 2.5kHz.

- **T1** : the second level trigger used information from the Čerenkov counters and the muon range for purpose of particle identification. In the case of $K^+ \to \pi^+\mu^+e^-$ the dedicated T1 trigger was MUE, and required light in both Čerenkov counters on the left (e^-) and a hit in the muon range (B - hodoscope) to be spatially consistent with the A*SH hit (μ^+). All T1 triggers were prescaled, except for MUE. The T1 trigger had a rate of 1.3kHz and response time of 130ns after the T0.

- **T2** : The third trigger level selected events with high invariant mass e^+e^- pairs. Information from A*SH hits, and Čerenkov counters was used to select events, consistent with low vertical separation of e^+e^- tracks (low invariant mass). Such events were pre-scaled. Events that had evidence of a e^+e^- pair , but were not satisfying the low mass criteria were passed to the next level trigger. The T2 trigger had a rate of approximately 0.6kHz and a response time of 200ns after the T1

- **T3** : the last trigger level compared the number of hits in each of the PWC plane with a predefined likelihood distribution obtained offline. This trigger rejected events that had either too few or too many PWC hits. On average, the rate of T3 was 1800 per 2.8s spill.

The trigger logic (T0...T3) was realized in hardware. In order to perform hit pattern recognition tasks custom made programmable MLU (matrix lookup units) were used. NIM units were utilized to perform more simple logical operations (AND/OR).

3.6.2 Trigger latches description

The trigger hierarchy described above allowed one to construct several trigger latches, that were useful for $K^+ \to \pi^+\mu^+e^-$ data collection and analysis.

They were constructed using the T1 trigger level, where particle identification information entered (Čerenkov counter, B - counter). At the T0, T2 and T3 levels all triggers were identical A short description of the configurations used is presented below:

1. **MUE:** CLand .and. $\overline{\text{CRor}}$ .and. B
   - trigger used for selecting $K^+ \to \pi^+\mu^+e^-$ candidate events.
2. **TAU:** (Prescaled by 10^4)
   - trigger used to select normalization sample of $K^+ \to \pi^+\pi^+\pi^-$ events.
3. **EEPS:** CLand .and. CRand (Prescaled by 20)
   - trigger used to select K^+ decays with e^+e^- pair in the final state (possible background)
4. **ČERENK:** (CLand .and. CRor) .or. (CLor .and. CRand)
   - trigger used to extract efficiency of the čerenkov counters, since it was requiring only three out of four čerenkov counters.

where:

- CLand = C1L .and. C2L
- CRand = C2R .and. C2R
- CLor = C1L .or. C2L
- CRor = C1R .or. C2R
- C1L signal above threshold (0.25 p.e.) on the left side of C1 counter
- C1R signal above threshold (0.25 p.e.) on the right side of C1 counter
- C2L signal above threshold (0.25 p.e.) on the left side of C2 counter
- C2R signal above threshold (0.25 p.e.) on the right side of C2 counter
- B - hit in the B counter, spatially consistent with a hit in SH and A counters

Chapter 4

Data Acquisition System

The design of the data acquisition system (DAQ) was based on the system used in the MEGA experiment at the Los Alamos National Laboratory [30]. It consisted of CAMAC, Fastbus, VME modules and three computers: two Digital 3000/600 Alpha Workstations and one VAXstation as shown in Figure 4.1.

CAMAC data were transferred to a 4Mb Fastbus buffer memory (LeCroy 1892) for each accepted T3 trigger. During transfer the whole system was inhibited. Time required for one event transfer averaged 2μs and depended on PWC occupancy. Fastbus ADC and TDC modules could buffer up to 7 events before the data were transferred to the buffer memory. Hence for up to six events readout of the ADC/TDC modules caused no dead time. A special counter kept track of the number of buffered events.

The data flow through the DAQ system was tailored to match the operating conditions of the AGS with the 5.1 s cycle and 2.8 s beam-on spills. Working in parallel, two fastbus smart crate controllers (FSCC) read out data into temporary memory buffers during the spill. The data transfer from memory buffers to the internal 4Mb memory of CERN host interface (CHI) was carried out when the beam was off. The CHI in turn wrote the data directly into the internal memory of two DEC Alpha workstations though a VME interface. Data transfer from CHI to workstations was executed at a maximum rate of 3.27 Mb/s [31].

After the software running in the workstations built events from the buffered data they would be transferred to a host computer (VAXstation) for taping. Data transfer rate was now limited by Ethernet connection to 450 Kb/s. With roughly 1700 events per spill corresponding to (2kb/event $\times$ 1700 events/spill) 3.4Mb/spill it was impossible to write all this data during the 5.1 s duty cycle. Additional event rate reduction was performed in workstations by means of filtering. Those on-line filters were, very much like the trigger, implemented in stages.

The filters checked buffer format and contents, performed very basic reconstruction of hits in the calorimeter and the A counter, along with reconstructing vertical coordinates of the tracks. A more detailed description of the filter code can be found in [28].

Finally a VAXstation was performing control functions and wrote events to the 8mm data tapes. A small fraction of events (about 1%) was sent to the two additional DEC stations for on-line monitoring of the detectors performance.

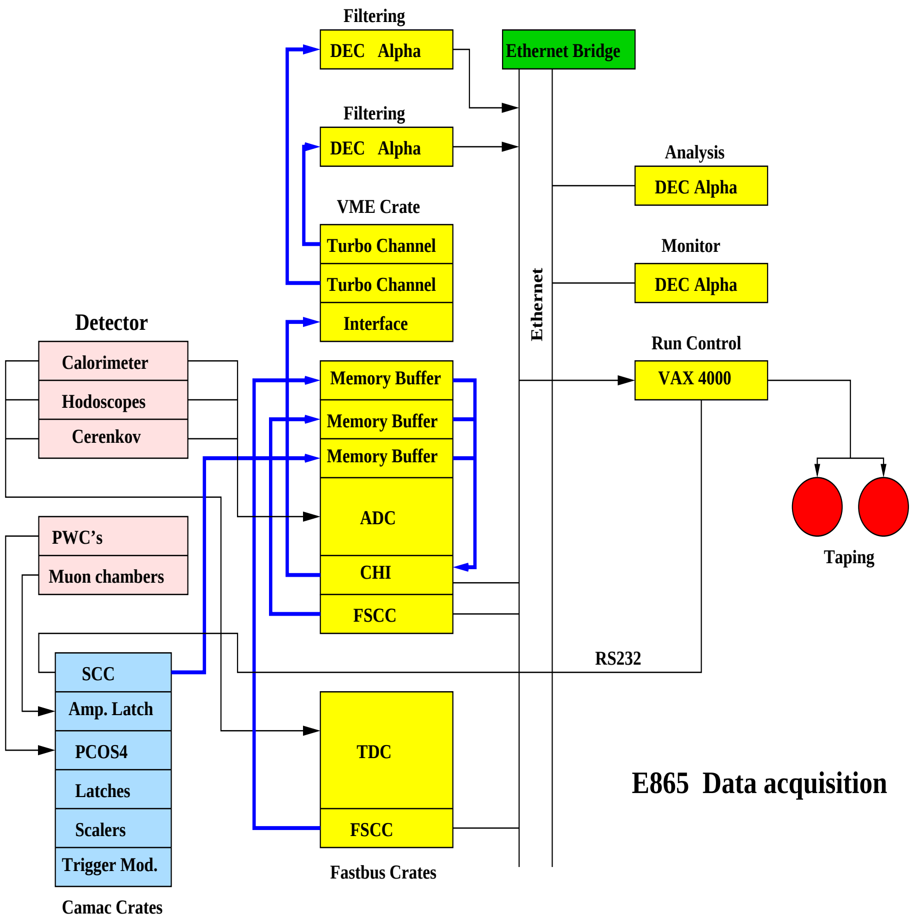


Figure 4.1: A schematic diagram of the E865 data acquisition system.The ADC modules were gated by a T1 signal, which also served as a common start for the TDC modules. Failure of the subsequent trigger levels (T2 or T3) immediately resulted in a fast clear with ADC and TDC modules ready to take data in about 1 μs.

Chapter 5

Event Reconstruction

The final $K^+ \to \pi^+\mu^+e^-$ data were collected by the E865 experiment in 1998. The running period of about five months yielded roughly twice as much data as taken in 1996. This was a consequence of an increase of the beam-on spill from 1.6s to 2.8 s, and the addition of wave length shifters on the PMTs in the čerenkov counters resulting in higher efficiency.

5.1 Overview

The signal search started with the offline analysis of the collected experimental data. Since about 500 Exabyte tapes of raw data were to be analyzed a reliable and logical procedure was needed, which proceeded in several stages.

The first pass (Pass 1) served the purpose of getting rid of the events, that could not be reconstructed. At the same time it divided the data into different streams, depending on the trigger condition to simplify the data handling.

The second pass (Pass 2) was designed to further clean up the data, using particle identification information.

Finally, after final cuts were determined and backgrounds carefully investigated and understood the data passed the third pass (Pass 3). The surviving data were analyzed using the likelihood method, which allowed to examine each event and to determine its probability for coming from a real $K^+ \to \pi^+\mu^+e^-$ decay or from a background. We describe the different steps in the following sections, beginning with the event reconstruction from the raw data.

5.2 Event building

The process of reconstructing an event followed the hardware trigger requirements and started with "clump" finding, *i.e.*, determining positions of tracks in the calorimeter based on their energy depositions. For each clump found, windows in PWC were defined and track positions were searched for in each PWC in descending order from P4 to P1. Later from these positions, tracks were formed and their momenta were determined. Using combinations of three tracks, a common vertex was reconstructed and quality of the vertex was calculated. The response of the particle identification detectors was identified for each track and kinematic variables such as invariant mass were reconstructed.

The details of these reconstruction algorithms are described in the following sections.

5.2.1 Shower counter clump finding

The search for a charged track begins with an in-time A counter slab where in-time is defined as being within 10 ns of T0. The row of shower modules directly behind each in-time A-counter slab is scanned to search for energy deposition and the module with the highest energy deposition above a minimum value of 75 MeV was considered first and labeled as a seed module. With "clump" we denote an array of three by three modules with the seed module in the center.

When a seed module was found a position dependent correction could be applied to the A counter time and the A counter and shower module coincidence could be checked again.

If this condition was satisfied, the energy deposition in the surrounding eight modules was summed to give the energy-weighted time and position of the clump.

The search for photon clumps for vetoing purposes comes later in the analysis after all tracks have been found and fitted. Aside from requiring that there be no active overlapping A counter, the algorithm for finding photon clumps was similar to that of charged-track clumps. To be of the photon type, clumps were required to have a total clump energy greater than 400 MeV and 90% of the energy in the shower must be contained in the innermost 2×2 square modules. These conditions were used to suppress backgrounds from neutrons.

5.2.2 Space points in proportional wire chambers

The next step in reconstructing tracks searched for space points determined from hits on intersecting wires in the PWCs. Due to the abundance of spurious hits, on average number being 40 per PWC, two conditions were imposed in the algorithm. The first relied on the correlation between spatial hits in the calorimeter and the PWC's. To make use of this correlation a rectangular window was projected from each clump to P4 and only wires inside this window were used to search for space points. Any space point found in P4 likewise defined a corresponding window in P3 and, similarly any space point formed in P3 defined a correlated window in P2. This process continued upstream from P2 to P1. The projection and size of the window were determined from an analysis of good $K^+ \to \pi^+\pi^+\pi^-$ events.

The second condition pertained to the number of wires to be used in the formation of space points. Although two-intersecting wires were sufficient to form a space point, the abundance of spurious hits of wires would render too many 2-wire space point combinations to be of any value and thus only space points from three or four wires combinations were kept. The algorithm started with searching for 4-wire space points and if none were found, the search continued for three wire combinations. Although a wire could be used for more than one space point, different space points could only share one common wire. For technical details of the space point reconstruction we refer to the PhD theses of Doug Bergman[27] or Stefan Pislak[26]. With a typical wire plane efficiency of 95% in P1, P2 and of 90% in P3, P4, the overall efficiencies for forming 4-wire or 3-wire space points in those chambers were 98.6% (P1,P2) and 94.8% (P3,P4).

5.2.3 Fitting tracks

The PWC space points in the correlated windows are grouped by the track finding algorithm to test for potential tracks. The tracking algorithm relied on the assumption that tracks are straight unless they are bent by magnetic fields. Thus Y-Z projection of the track is straight while the X-Z component is straight except when acted upon by the P_t-kick in the vicinity of the magnetic field. Using the middle of D6 as the median (dividing line), tracks can be divided into two parts, upstream (P1,P2 and D6) and downstream(D6,P3 and P4). Using the continuity of tracks at D6, the track quality is measured by testing the

the three upstream and downstream components of the track namely the difference in the projected X and Y positions as well as the Y slope at D6.:

The track quality was assessed using the equation

$$\chi^2 = \frac{(x^{us} - x^{ds})^2}{\sigma_x^2} + \frac{(y^{us} - y^{ds})^2}{\sigma_y^2} + \frac{(m_y^{us} - m_y^{ds})^2}{\sigma_{m_y}^2} + \frac{2(y^{us} - y^{ds})(m^{us} - m^{ds})}{\sigma_{ym_y}^2} \tag{5.1}$$

where x and y are the positions of the track at the middle of D6 from upstream(us) or downstream(ds) and m_y is the slope of the track in the y-projection. The σ's are pre-determined resolutions of the respective quantity.

The tracking algorithm begins the search from the (x, y) position of each shower clump and searches for 4 correlated space points, one in each PWC. If the χ^2 determined from Eq. (5.1) is smaller than a cutoff value, the track is kept in a track list. If no 4-PWC tracks are found, the algorithm then searches for 3-PWC tracks where the χ^2 is determined assuming that the YZ projection of the missing space point (PWC) lies on a straight line formed by the 3 present space points. Up to six tracks can be assigned to a clump since sometimes the track with a higher χ^2 may yield a better vertex fit.

Cutoff values for the track χ^2 and number of tracks to be assigned to a clump are determined by studying reconstruction efficiency of $K^+ \to \pi^+\pi^+\pi^-$ events and fixed at the values which insure optimum balance between CPU load and loss of good events.

5.2.4 Momentum fitting

Five parameters are needed to describe a track: the X and Y positions and the slopes in the XZ and YZ planes at P1, and the radius of curvature R, defined as the ratio of charge to momentum of the track. Momentum is determined by minimizing χ^2_{trk} with respect to the parameters above. The χ^2_{trk} represents the difference between the measured wire positions in the PWC's and the expected wire positions determined from the parameters of the track by propagating the track from the center of P1 using either a hard-edge approximation for the magnetic field or the measured field map. Denoting with x_i the measured wire positions, with σ_i the error in the position measurements, with α_k the five track parameters, and with $X_i(\alpha_k)$ the expected wire positions, χ^2_{trk} could be written as

$$\chi^2_{trk} = \sum_{i=1}^{16} \frac{(x_i - X_i(\alpha_k))^2}{\sigma_i^2}. \tag{5.2}$$

The sums extends over 16 wire planes and was minimized to determine the parameters α_k. From the track finding routine, a first guess of parameters, α_{k0}, is available and the χ^2_{trk} can be re-written using a first order Taylor expansion of $X_i(\alpha_k)$ about the first guess:

$$\chi^2_{trk} = \sum_{i=1}^{16} \frac{(x_i - X_i(\alpha_{k0}) - \sum_{k=1}^{5} \frac{\partial X_i(\alpha_{k0})}{\partial \alpha_k} \Delta\alpha_k)^2}{\sigma_i^2} \tag{5.3}$$

Here the $\Delta\alpha_k$ are the corrections to α_{k0} which produce the best fits for the measured track hits x_i. Minimizing and keeping only terms with first derivatives in X we get

$$\begin{aligned} \frac{\partial \chi^2_{trk}}{\partial \alpha_i} &= 2\sum_{j=1}^{16} \frac{(x_j - X_j(\alpha_{k0}))}{\sigma_j^2} \frac{\partial X_j(\alpha_{k0})}{\partial \alpha_i} - \\ & \quad 2\sum_{j=1}^{16} \frac{1}{\sigma_j^2} \frac{\partial X_j(\alpha_{k0})}{\partial \alpha_i} \sum_{k=1}^{5} \frac{\partial X_j(\alpha_{k0})}{\partial \alpha_k} \Delta\alpha_k = 0 \end{aligned} \tag{5.4}$$

In matrix form, the equation can be rewritten as

$$V_i = M_{ij}\Delta\alpha_j \Rightarrow \Delta\alpha_j = M_{ij}^{-1} V_i. \tag{5.5}$$

The most important calculation above is the initial guess, $X_j(\alpha_{k0})$. To determine this guess, tracks are propagated from P1 to P4 by either using the faster hard-edge approximation or the field map. In the latter case the step size can be matched to the desired accuracy of the tracking. To save computer time only in the last pass (pass3) the step size was reduced to 5 cm, in the first pass (pass1) 20 cm sufficed.

5.2.5 Vertex fitting

A powerful feature distinguishing real kaon decays from accidentals is the quality of the reconstructed vertex from the daughter tracks. The position of the vertex is obtained by propagating tracks, three at a time, upstream to the middle of the decay region. In propagating through D5 again either the hard-edge approximation or the field map could be used. In addition, each track was required to belong to a separate clump.

The vertex fitting algorithm then determines the vertex position as the (x, y, z) coordinate which minimizes S, the root-mean-square distance of closest approach to the vertex for the 3-track combination. The combination that produces the lowest S value was tagged. Only combinations with $S <10$ cm were kept.

The described above procedure did not take into account the errors of the reconstructed tracks' positions and as a result, the average value of S increased for the vertices found far upstream[1]. The dependence of average S on the Z position of the vertex was derived from the data (using measured $K_{\pi\pi\pi}$ events) and was used to remove such dependence (see Section 7.3.2).

In addition, in order to test the effectiveness and sensitivity of the vertex fitting algorithm, an alternative method was considered. The method of the constrained vertex fit was developed by the Pittsburgh group and utilized the track error matrix, which was obtained from the tracks' fit in the spectrometer and their propagation back to the decay region through the magnetic field of D5. Uncertainties of the field map were included in the above-mentioned error matrix. This method allowed correct treatment of correlated uncertainties of the reconstructed tracks' positions and provided the true χ^2 of the constrained vertex fit as the vertex quality indicator. It took some time to finish the implementaion of the algorithm into the offline software and by the time it was finalized, a large part of the data processing was already done using the original vertex fit. Data studies, using good $K_{\pi\pi\pi}$ events, showed that the average value of χ^2 was almost independent $(O(10^{-4}))$ of the Z position of the vertex. Due to the fact, that the use of χ^2 as a vertex qaulity indicator did provide similar rejection of accidental events as the standard S, and due to a shortage of time, required to reanalyze already processed data, the constrained fit was not used in the analysis. However, it did serve as a robustness test for the original vertex fitting algorithm. In addition, in a separate E865 1998 year data analysis, which was aimed at a precise measurement of the $K^+ \to \pi^0 e^+ \nu_e$ (K_{e3}) relative branching ratio, both vertex fitting algorithms provided same results[32].

5.2.6 Particle identification (PID)

After the kinematic reconstruction is completed, the next phase is to identify the particle type associated with each track. This was accomplished using the PID variable from the various detectors with appropriate cuts. Electrons and pions could be separated by using

[1] This and the following paragraph were added to the text after it was submitted to the University of Zurich archives

čerenkov and calorimeter PID response while pions and muons are best distinguished using the information from the muon range stack.

Each data pass required increasingly stringent conditions on the same PID variable. The general discussion presented here will be detailed below for each phase.

Čerenkov light

The number of photo-electrons in a čerenkov counter attributed to a track is determined by summing up the light in the PMTs associated with that track. If a mirror intersects the rays projected from the track, the ADCs of the PMTs matched with that mirror are included in the summation. Depending on the trajectories of the tracks, it is possible for the same PMT to contribute light to more than one track.

In addition to ADCs, a track is also characterized by čerenkov timing. After slewing correction, all timing signals with TDC values within 30 ns of T0 are accepted.

E/P

As mentioned in the track finding algorithm, each track is assigned to a shower clump in the calorimeter. Associated with the clump are the position and energy, where the latter is determined by adding the deposited energies of the 8 modules (surrounding the seed module) to that of the seed module. E/P, constructed by taking the ratio of the clump energy to the momentum of the track, is used to distinguish electrons from muons and pions (see Figure 7.1).

Muon stack

Analogous to the PWCs, the muon chambers are used to search for muon tracks in the stacks. A track in the muon stack is projected from positions in P3 and P4, providing an axis for a cone whose boundaries are limited by $\pm 3\sigma_N$ where σ_N is defined as

$$\sigma_N = \left[\left(\frac{s}{\sqrt{12}} \right)^2 + \left(\sum_{i=1}^{N} \Delta z_i \times \frac{0.014 \text{ GeV}}{P\beta} \sqrt{XR_i} (1 + \frac{\log XR_i}{9}) \right)^2 \right]^{1/2} . \qquad (5.6)$$

The first term in the equation describes the resolution of the muon chamber with a wire spacing, s, of 1.3 cm. The second term is the resolution due to multiple scattering where Δz_i is the distance between the calorimeter and the ith chamber, XR_i is the number of radiation lengths in the ith chamber, P is the momentum of the particle and β is the velocity.

It is also possible to tag muon tracks using hits in the B and C scintillation counters. From the slope of the track at P3-P4, a projected position at the B(C) hodo plane is determined and a residual (labeled b) between this position and the center of the nearest in-time B(C) slab is calculated. We define

$$\chi^2_{B(C)} = \left(\frac{b}{\sigma_{B(C)}}\right)^2 \tag{5.7}$$

where $\sigma_{B(C)}$ is the error of the muon track at the B(C) hodo plane determined from Eq. (5.6). Muon tracks are selected by placing a cut on χ^2_B (see Figure 6.1).

Chapter 6

Initial Event Selection

6.1 Pass1

Pass1 serves as the first step of the offline data analysis. The raw data, stored on 8mm Exabyte tapes were analyzed and events, that could not be reconstructed were discarded. During data collection such events were recorded because they satisfied the online trigger condition. The most important requirement was the reconstruction of a 3-track vertex. A few basic kinematic cuts were applied to clean up the sample, too. They are listed in Table 6.1.

Additionally, for future easier manageability of the data, events, that satisfied pass1 cuts, were separated into different streams, depending on their trigger condition (Refer to section 3.6.2 for Trigger latches description). The requirements along with short description of each of the pass1 streams are presented below.

- **3TRK**: contained events with 3 charged tracks with MUE trigger latch, comprising 20% of the data. This stream possibly contains events of interesting physics such as $K_{\pi\mu e}$. Later, that stream would be used for signal search.
- **TAU**: contained events with the TAU trigger latch. About 3% of surviving events were written in this stream. That stream would later be used for normalization, resolution and PID studies.
- **EEPS**: contained events with the EEPS trigger latch. About 6% of surviving events were written in this stream. That stream would be later used for PID and background studies.
- **COSMICS**: contained events with the COSMIC trigger latch and were used for calorimeter calibration using cosmic ray muons which pass from top to bottom. The events pass without any requirements and account for 1% of the data.

- **5TRK**: contained events with the five charged tracks independent of the trigger latch. That stream could contain 5 charged particles kaon decays. An example of a kaon decay contributing to this stream would be $K^+ \to \pi^+\pi^0$ followed by $\pi^0 \to e^+e^-e^+e^-$. In the search for $K^+ \to \pi^+\mu^+e^-$ decay these events were not used, thus discussion of them ends here.

Although pass1 was the simplest stage it was also a very effective one in reducing the amount of data. However, it required a lot of CPU time, namely eight months of running the raw data tapes on several Multi-Pentium CPU Linux machines. Pass1 led to a structured data set , that can be used for initial PID studies and detector performance assessment, which is required for defining PID cuts for the next stage of analysis (pass2).

Table 6.1: Summary of pass1 selection criteria.

Cut	Description of the cut
$N_{clump} \geq 3$	3 or more clumps in the calorimeter.
$N_{invtx} = 3$	3 charged tracks in vertex.
S < 10	RMS of miss distance of vertex less then 10 cm.
2 POS + 1 NEG	2 positive and 1 negative track
$P_{KAON} < 8$	Reconstructed momentum of 3 tracks is less then 8 GeV/c

6.2 Pass2

The second stage of the analysis was concentrated on the 3TRK stream from pass1, which contained possible $K_{\pi\mu e}$ candidates. During pass2 PID cuts requiring a negative track identified as an electron and two positive tracks identified as a positive pion and muon were applied.

Previous experience from the analysis of the 1996 data [2] helped and allowed to create several background oriented streams, too.

The limits for the PID cuts were chosen rather loose, in order to maintain high efficiency for possible signal events.

Main $K^+ \to \pi^+\mu^+e^-$ backgrounds

Two different types of backgrounds had to be taken into account in the $K^+ \to \pi^+\mu^+e^-$ search.

The first type of background comes from other K^+ decays, with similar kinematics, but different final set of particles. They could be tagged as $K^+ \to \pi^+\mu^+e^-$ decays, if

PID failed, and kinematic conditions were satisfied. In this category, first background process would be $K^+ \to \pi^+\pi^+\pi^-$ decay (K_{Tau}). Second process would be an array of different kaon decays with a π^0 in the final state, with latter undergoing a Dalitz decay ($\pi^0 \to e^+e^-\gamma$). We would call such kaon decays K_{Dal} decays for convenience. Both K_{Tau} and K_{Dal} can be tagged as $K^+ \to \pi^+\mu^+e^-$ decay, when different cases of misidentification occur (See Table 6.2) and kinematic requirements are satisfied. It is obvious that particle identification is crucial for suppressing such background processes.

The second type of the background is rather different, and arises from accidental combination of tracks that reconstruct into a vertex and satisfy PID requirements for $K^+ \to \pi^+\mu^+e^-$. Such events are usually characterized by poor timing and vertex quality. In addition usually a lot of *noise* (high load on detector elements, causing high number of hits in the PWC chambers, etc.) in the apparatus is accompanying such events. We would call this background - accidental events.

Table 6.2: Main $K^+ \to \pi^+\mu^+e^-$ backgrounds from other K^+ decays.

K^+ Decay Mode	Misidentification
$K^+ \to \pi^+\pi^+\pi^-$ with $\pi^+ \to \mu^+\nu$, $\pi^- \to e^-\overline{\nu}$	None
$K^+ \to \pi^+\pi^+\pi^-$ with $\pi^+ \to \mu^+\nu$	π^- as e^-
$K^+ \to \pi^+\pi^+\pi^-$ with $\pi^- \to e^-\overline{\nu}$	π^+ as μ^+
$K^+ \to \pi^+\pi^+\pi^-$	π^+ as μ^+ and π^- as e^-
$K^+ \to \pi^+\pi^0$, $\pi^0 \to e^+e^-\gamma$	e^+ as π^+ and π^+ as μ^+
$K^+ \to \pi^+\pi^0$, $\pi^0 \to e^+e^-\gamma$ with $\pi^+ \to \mu^+\nu$	e^+ as π^+
$K^+ \to \pi^0\mu^+\nu$, $\pi^0 \to e^+e^-\gamma$	e^+ as π^+

During the 1996 $K^+ \to \pi^+\mu^+e^-$ data analysis it became evident that the accidental background dominates [2]. The problem with accidental background is that it is hard to isolate and study. In order to understand accidentals better one must try to obtain the largest possible sample. In the next section we discuss how pass2 kinematic cuts were selected.

6.2.1 Pass2 kinematic selection

Taking background information into account four separate streams were formed. Besides the standard data stream (NORM) three other streams were filled, namely accidental (ACC) and Dalitz decay (DAL) events mimicking $K_{\pi\mu e}$ candidates and a stream containing $K_{\pi\mu e}$ candidates with additional cuts (PME).

Details are given in Table 6.3. PID cuts were applied in every stream. The vertex quality cut was applied only to the PME and DAL streams, because it does not reduce acceptance for these events. The cut on the kaon momentum was applied differently. For the accidental stream (ACC) a lower limit of 5 GeV/c was required, but no upper limit, since events with total momentum exceeding the beam momentum (6 $\pm$ 0.12 GeV/c) could only come from accidental track combinations. For the Dalitz decay stream (DAL) an upper limit of 7 GeV/c was applied, since the sum of the momenta is lower then the beam momentum in most cases due to unobserved decay photons.

Table 6.3: Pass2 cuts and their application for the different streams. (X - means that cut(s) was(were) applied).

Cut	NORM	ACC	DAL	PME
PID cuts ($\pi^+\mu^+e^-$)	X	X	X	X
Vertex quality cut ($S_{norm} < 5$ cm)			X	X
$P_{KAON} > 5$ GeV/c		X		X
$P_{KAON} < 7$ GeV/c			X	X

6.2.2 PID selection in pass2

The PID selection criteria for electron, positive pion and muon applied in pass2 are as follows:

Electron

To reject a π^-, but at the same time have an excellent efficiency for an e^- the following cuts were chosen.The left side of the čerenkov counters was in the trigger requirement for $K_{\pi\mu e}$ events.

- C1 p.e.$>$ 0
- C2 p.e.$>$ 0

- $|E/P - 1| < 0.5$

The first two cuts for the number of photo electrons in C1 and C2 reject a large fraction of the pions, while keeping the possible $K^+ \to \pi^+\mu^+e^-$ candidates. In order to reduce the number of misidentified pions the information from the calorimeter was used as well.

Pion

In order to effectively identify π^+ while maintaining reasonable rejection rate of e^+ the čerenkov and calorimeter information was used. Fine tuning of these cuts was however reserved to the final part of the analysis. The logic for the pass2 π^+ identification was the following:

- two out of the following three conditions:
  - C1 p.e.< 1.5
  - C2 p.e.< 1.5
  - $E/P < 0.9$
- .OR. $E/P < 0.8$

Muon

During the pass2, main goal is to have a high efficiency of a μ^+ identification, with a reasonable rejection rate for a π^+. In order to do that, information from the calorimeter and the muon counters was used. Since a muon would pass through the calorimeter, creating just a minimum amount of ionization , a cut on the energy deposited in the calorimeter was imposed In addition, the location of the B counter hit must correspond to the track, and pass an appropriate χ^2 test (See section 5.2.6 for details). Since the B counter was in the hardware trigger requirement, a cut is not deleting any events, that can be considered a signal. Finally knowing the momentum of the muon allowed to predict its range in the muon system, and compare it to the measured one (See Figure 6.1).

The muon selection cuts are presented below:

- $B_{HIT} < 25$
- Hits observed / Hits expected > 0.4
- $E_{Calorimeter} < 0.5$

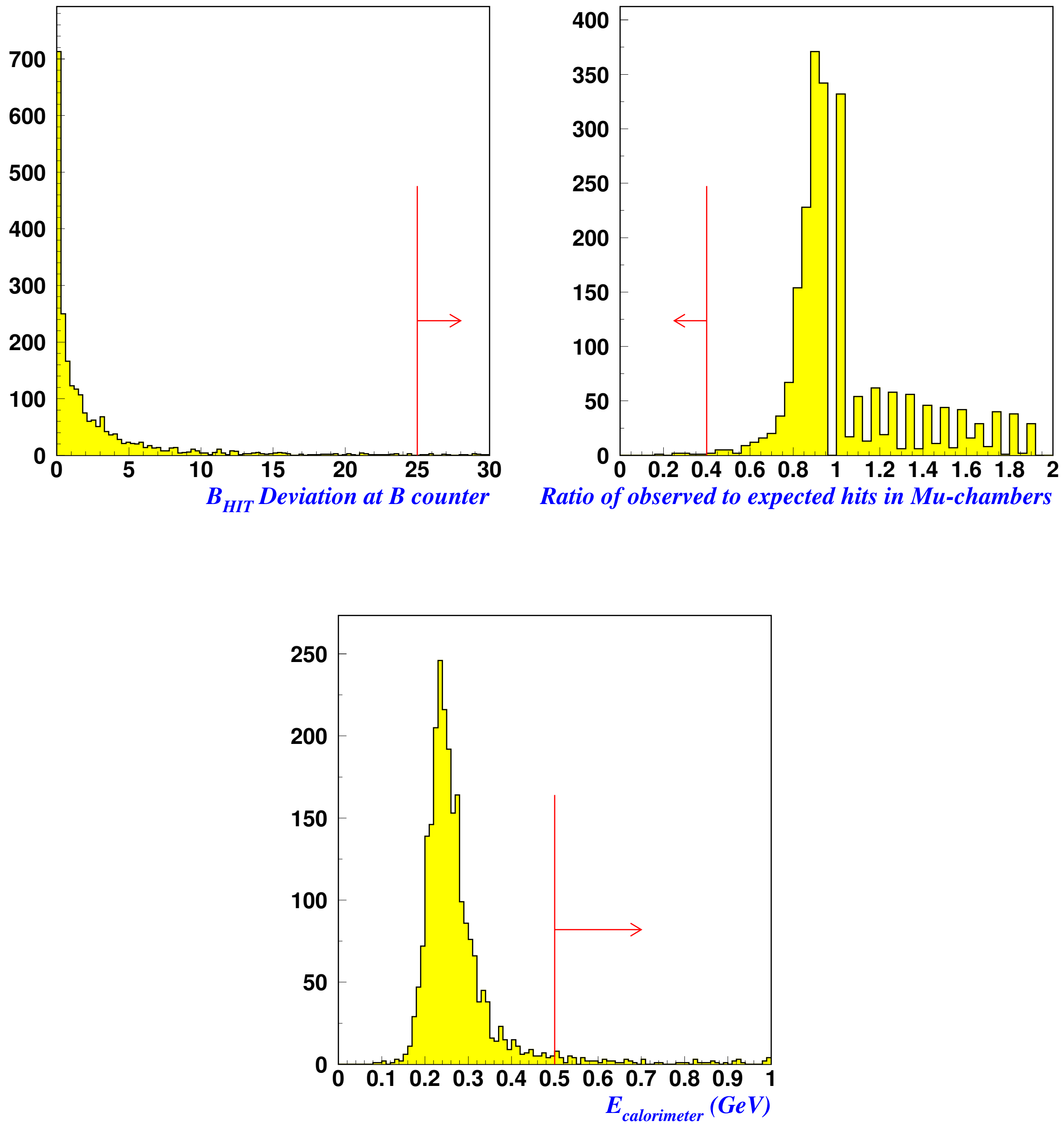


Figure 6.1: Distributions of the muon particle identification variables: A deviation of the track projection from the B counter hit (B_{HIT}); The ratio of the expected number of hits to the observed ones; The energy deposited in the calorimeter (E_{cal}).

6.3 Introduction to pass3

In pass3, all events were required to pass a final set of cuts, that will be discussed in the following chapters. These cuts were established after all possible background processes were studied, evaluated and satisfactory suppressed. In addition, the step size in the magnetic field approximation for pass3 was reduced to 5cm.

However, the final cuts applied during pass3 were not a final tool used in the search for the $K_{\pi\mu e}$ decay.

Chapter 7

Likelihood Method

7.1 Signal search strategy

In the analysis aimed at searching for a rare or forbidden decay, the method by which it is accomplished becomes important. We've chosen the so called *blind analysis* as our strategy. All backgrounds should be studied, understood and estimated. In order to check the validity of the background estimates one compares vital distributions, *i.e.* the reconstructed kaon mass $M_{\pi\mu e}$ for $K_{\pi\mu e}$ candidates and predicted backgrounds. However, one does exclude the signal region, frequently called the *box*, and looks at the distributions only outside of it. The signal region is examined at the very final stage when all backgrounds are satisfactorily understood and minimized. Such procedure allows one to perform an unbiased analysis.

For the 1996 data accidental background proved to be the strongest with a relative contribution of about 95%. The approach of simple cuts was no longer efficient in suppressing accidental events and the use of likelihood functions was suggested [2]. Following the same logic forced us to create a general likelihood function based on distributions of the relevant kinematic variables. Physical backgrounds, arising from the misidentification of particles, could be suppressed by optimizing the PID selection criteria. For that purpose likelihoods are suitable, too.

Accidental events could be characterized by poor timing and kinematic qualities, while we know that signal events would be opposite in that regard. That difference was used, when constructing likelihoods, for both signal and accidental background hypotheses. The general likelihood for signal and background (accidental events) hypotheses was used as a final selection criteria. Events, found within the signal region, were examined on an event by event basis, with the probability of being a signal or a background taken into account.

Since the likelihood method was used extensively, first in order to optimize particle identification and later to suppress the main $K_{\pi\mu e}$ background (accidental events), we continue with an overview of the method.

7.2 Overview of the likelihood method

In a cut analysis, all events that passed a defined set of cuts become indistinguishable. The likelihood method, however, allows one to extract additional information and each event can be tested against a particular hypothesis, e.g. the likeliness for a particle to be of a certain type or the decay to be of a particular mode.

For example, when simply using a set of cuts to select an electron, it is possible to misidentify a pion as an electron, if the pion satisfies all electron selection criteria. However, if one were to account for the probability of having a particular value for each PID variable, it is possible to further discriminate pions and electrons. Thus, while a pion may satisfy all PID cuts for an electron, the values for each of the PID variable will most likely fall at the edge of the various PID distributions. On the other hand, a real electron will have PID values which will more likely be in the bulk of the distributions (see Figure 7.1 as example). Taking the probabilities of an electron having a particular PID value and summing it for all PID distributions will allow one to further distinguish between a real and misidentified particle. The same method can be applied to kinematic variables and their distributions.

If a PID likelihood distribution is constructed for each particle type and a kinematic likelihood distribution is formed for each background mode, it is possible to determine the likelihood of an event to originate from a particular decay mode. Using multiple variable likelihood functions therefore allow to differentiate quantitatively between signal and background on an event-by-event basis.

7.2.1 Formulation

The conventional definition of the log likelihood function is

$$\mathcal{L}(\vec{x}) = \log P(\vec{x}) \tag{7.1}$$

where $\vec{x} = (x_1, x_2 ... x_n)$ is a vector of the measured kinematic and PID response and $P(\vec{x})$ is the n-dimensional probability density function (p.d.f.). Limited statistics from the experimental observations demands factorization of the probability:

$$P(x_1, x_2, ..., x_n) \cong P_1(x_1) \cdot P_2(x_2) \cdot ... \cdot P_n(x_n) \tag{7.2}$$

where $P_i(x_i)$ is the i_{th} probability density function (p.d.f.) constructed from the corresponding distribution and x_i is the respective measured response.

Using such an approximation one could rewrite the log likelihood in the following way:

$$\mathcal{L}(\vec{x}) = \sum_{i=1}^{n} \log P_i(x_i) \tag{7.3}$$

Generally, a p.d.f. can be represented as a continuous analytical function or a binned histogram, in which case $P(x_j)dx$ is the probability of being in the j_{th} bin with bin-width dx. In constructing a likelihood function[1], we require the probability function or histogram be normalized to unity.

While it would be ideal to use full range of a variable for the likelihood function to maximize the acceptance, it is more practical to use a finite (around 95%) range, choosing particularly the most sensitive part of the distribution. In most of the cases the range of the likelihood function would be defined by the values of the cuts selected for pass3. The number of bins should be large enough to reflect the shape of a particular distribution.

Following such an algorithm, a total likelihood function can be created for the signal hypothesis($K_{\pi\mu e}$). A possible shortcoming is that the information on the background is not yet utilized. In the next section an example of using both signal and background information together is discussed.

7.2.2 Example of e^- and π^- separation using the E/P likelihood function

Information from the shower counter can be used to separate electrons from pions. When passing through the calorimeter an electron loses its energy in the electromagnetic shower, and thus its observed energy to momentum ratio E/P is peaked around 1 within the detector resolution. The resolution depends on the electron momentum, but in the example we temporarily disregard it.

In order to select an e^- the ratio E/P is required to fall within the interval $0.65 <$ E/P < 1.5. The corresponding efficiency for selecting e^- is 98%.

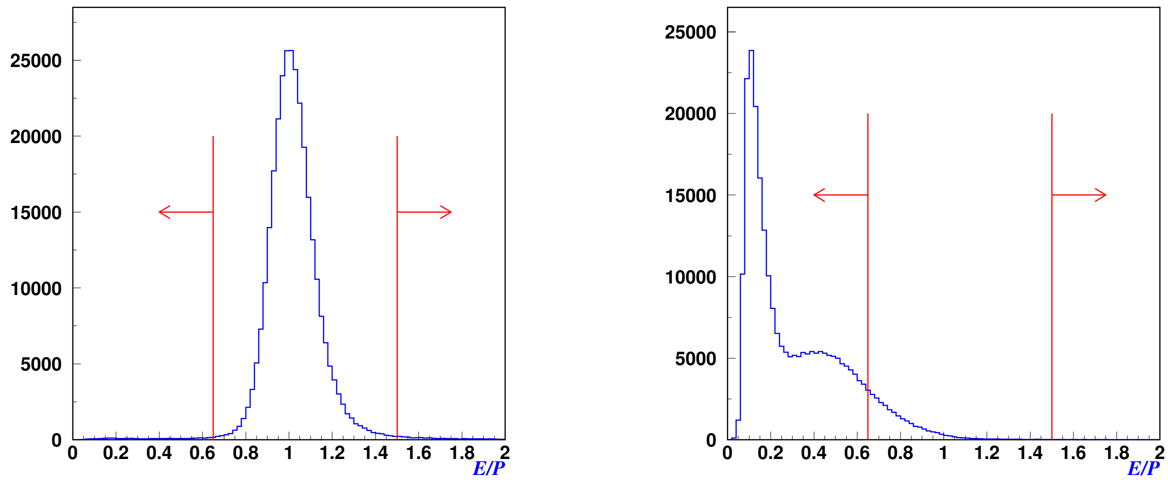


Figure 7.1: Distribution of the variable E/P for e^- (left) and π^- (right). Lines denote the particle identification cut for the e^-($0.65 <$E/P < 1.5).

[1] To avoid confusion we state here that for future references, we will use the terms likelihood functions and log-likelihood function synonymously.

A fraction of the π^- (10%) will also pass the cut. For this example we define e^- as our signal and π^- as background hypothesis and use likelihood function with the variable E/P in the range defined by the cut to separate them. The corresponding p.d.f. are shown in Figure 7.2. The likelihood that an electron candidate is really an e^- can now be calculated.

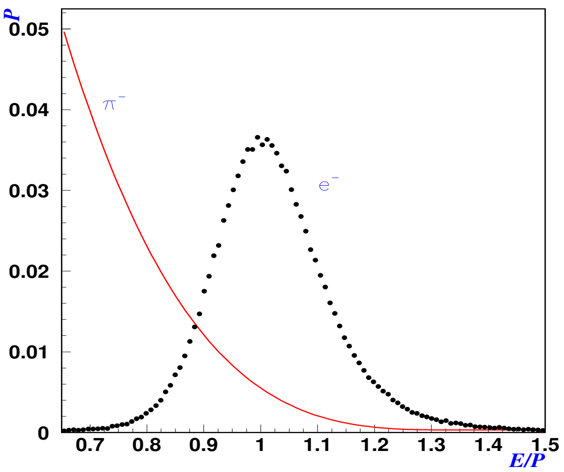


Figure 7.2: E/P probability density function (p.d.f.) for π^- (line) and e^- (points), both p.d.f.'s are normalized to 1 within the cut's range (0.65 <E/P < 1.5).

To test the efficiency of the constructed likelihood we used a sample of electrons, selected from the reconstructed $K^+ \to \pi^0\pi^+$ decays, and a sample of π^-, selected from reconstructed $K^+ \to \pi^+\pi^+\pi^-$ decays. These two samples represented signal and background in the case of e^- identification (Figure 7.1).

Plots of the likelihood distributions, the detection efficiency and misidentification probability P(π^- as e^-) as a function of the likelihood cut are presented in Figure 7.4.

While the misidentification probability (P(π^- as e^-) was reduced from the nominal value of 0.099, achieved by the cut, so far only the information about signal was used. Events with a low probability of being an e^-, namely events with E/P either lower or greater then one were rejected. But judging from the E/P distributions for e^- and π^-, events with E/P lower then one are more likely to be pions as opposed to events with E/P greater then one. It seems reasonable to compare the probability of being an e^- with the probability of being a π^- (background in that case).

Since we are working with the log-likelihoods, ratio of the probabilities can be rewritten in the following way:

$$\log \frac{\mathcal{L}_{e^-}}{\mathcal{L}_{\pi^-}} = \log \mathcal{L}_{e^-} - \log \mathcal{L}_{\pi^-} \tag{7.4}$$

The distribution of such relative likelihood (Figure 7.5) and the e^- identification efficiency versus misidentification probability curve can be seen in Figure 7.3.

Efficiencies and misidentification probabilities achieved by three different methods can be observed in Table. 7.1.

It is evident that the use of the relative likelihood ($L_e - L_\pi$ in that example) should be favored. It allows one to utilize the information both about the signal and about the background hypothesis for each event.

Table 7.1: The identification efficiency of an e^- and the misidentification probability of a π^- as an e^- from using the E/P in the calorimeter and cuts and/or likelihood method for the e^- selection.

e^- selection method	P(e^- as e^-)	P(π^- as e^-)
$0.65 <$ E/P < 1.5	0.98	0.1
$L_{e^-} >$ -6.3	0.95	0.04
$L_{e^-} - L_{\pi^-} >$ -1.5	0.95	0.03

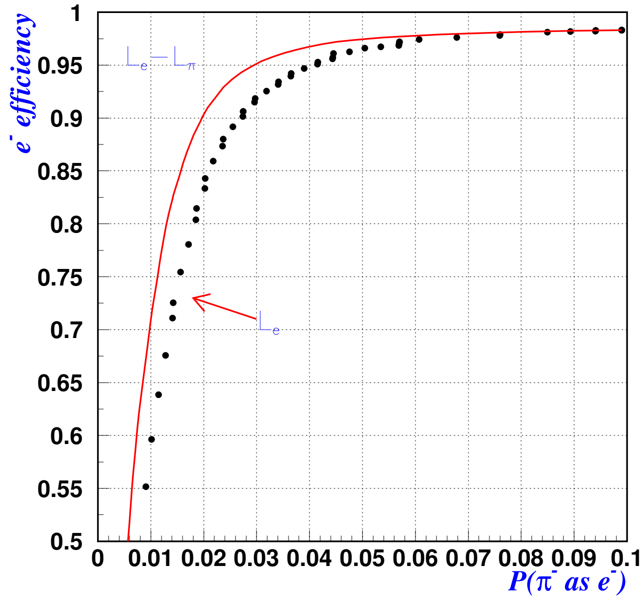


Figure 7.3: e^- identification efficiency versus misidentification probability P(π^- as e^-) using the electron E/P likelihood L_e only (dots) and the relative electron to pion likelihood $L = L_e - L_\pi$ (line).

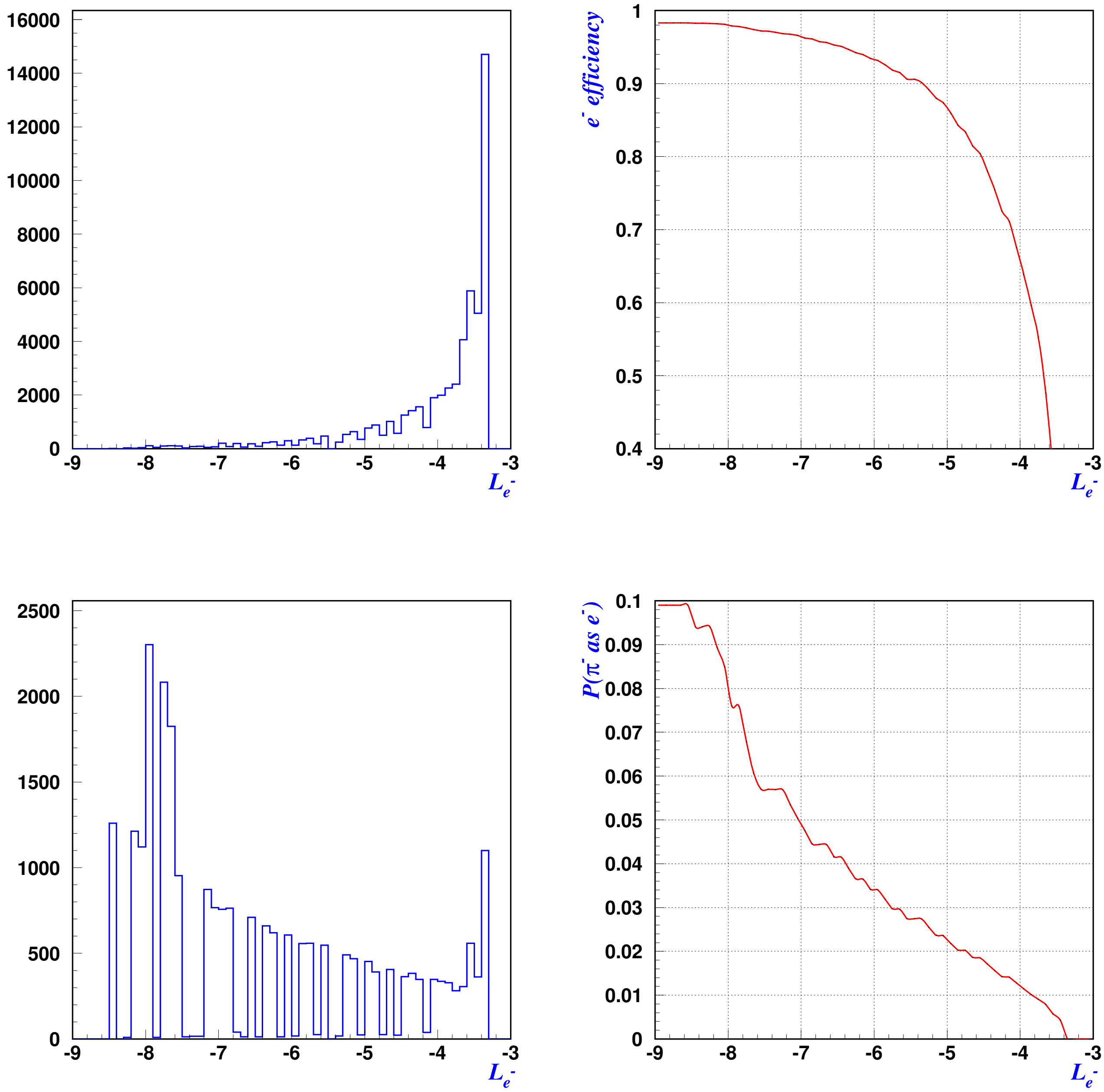


Figure 7.4: Distributions of the likelihood derived from the ratio E/P (L_{e^-}) for e^- (up left) and π^- (down left); identification efficiency (up right) and misidentification probability P(π^- as e^-) (down right) versus the likelihood.

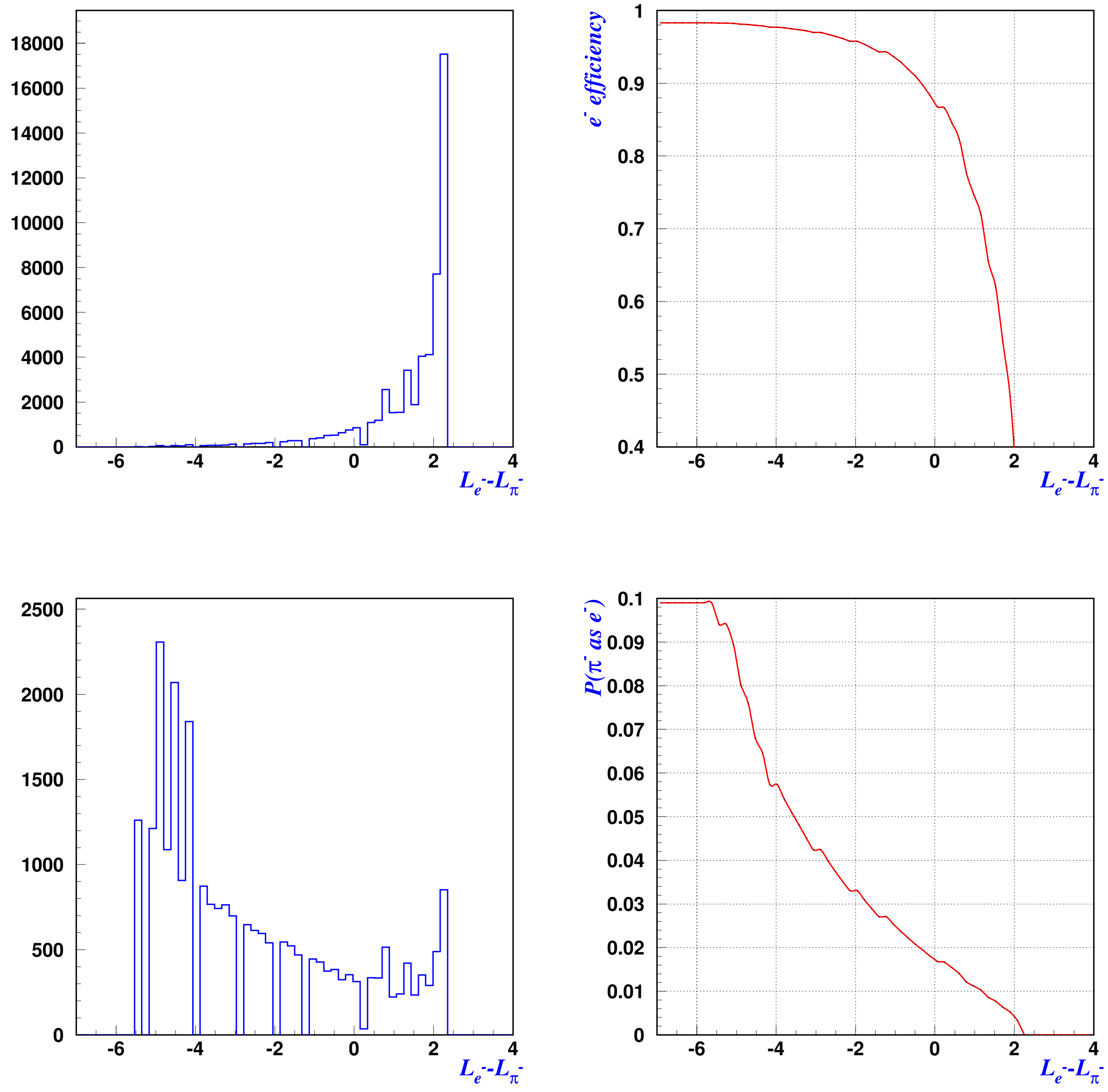


Figure 7.5: Distributions of the relative likelihood derived from the ratio E/P ($L_{e^-} - L_{\pi^-}$) for e^- (up left) and π^- (down left); identification efficiency (up right) and misidentification probability P(π^- as e^-) (down right) vs the likelihood.

7.2.3 Implementation plan

The plan to develop and test likelihood functions, which were created for particle identification optimization and suppression of the accidental background, can be summarized in the following steps:

- **Definition of the kinematic selection criteria** (Section 7.3).
- **Development of the particle identification likelihoods (L_e,L_π)**: Identification efficiencies and misidentification probabilities are determined. The likelihood functions are constructed and tested (Section 7.4).
- **Estimate of the background level caused by $\mathbf{K_{Tau}}$ and $\mathbf{K_{Dal}}$**: The set of final PID likelihood functions cuts, which would insure satisfactory suppression of such backgrounds, is determined (Sections 8.1,8.2).
- **Estimate of the accidental background level** (Section 8.4).
- **Development of a general kinematic likelihood for the $\mathbf{K_{\pi\mu e}}$ events and accidental background** (Section 9.2).
- **Test of the efficiency of the general kinematic likelihood in separating signal and accidental events** (Section 10.2).

7.3 Kinematic selection criteria

Kinematic criteria for selecting $K_{\pi\mu e}$ events were dictated by a simple task of rejecting background events, while maintaining high efficiency for signal. Preliminary data studies showed that the main $K_{\pi\mu e}$ background consisted of accidental events.

Having at least basic understanding of that background would allow one to select a reasonable set of kinematic cuts.

7.3.1 Accidental events mechanism

An accidental event by its very definition is a superposition of tracks, not all of which originated from an actual kaon decay. The simplest mechanism for accidental background event would be a combination of two tracks from a real kaon decay and one extra track from another decay, or beam background *(2+1)*. Another possibility is that all three tracks would originate from different decays/beam background *(1 + 1+ 1)*.

Accidental events are characterized by:

- The reconstructed kaon momentum is no longer constrained by the momentum of the beam kaons (6 GeV/c).

- Quality of the reconstructed vertex is poor.
- Target likelihood is poor, since the kaon momentum formed by an accidental event has a low probability to be tracked back to the target.
- Track times, determined by the time of the shower and hodoscope hits are uncorrelated with each other.
- A presence of extra clumps in the calorimeter is expected, since the accidental event is a superposition of several decays (Kaon, or beam background).

What mechanism is more frequent (*(2+1)* or *(1+1+1)*) was determined later, when all the necessary tools to accurately select accidental events were established.

7.3.2 Kinematic variables

Before revealing the set of kinematic cuts we are giving a short overview of the variables that were used.

Three track vertex quality S_{norm}

During the reconstruction of the event the quality of the fitted vertex is assessed from the r.m.s. distance of closest approach of the three trajectories to the vertex (S). From the studies of K_{Tau} decays we know the S distribution for good events. It was noticed that S depends on the z coordinate of the vertex (Z_{vertex}), *i.e.* the mean value of S increases for the decays occurring further upstream. This can be explained by the fact, that the resolution of the trajectories worsens as the track is extrapolated further upstream. In order to account for the dependence on the Z coordinate of the vertex, mean S was fitted against the Z_{vertex} with first and second order polynomials for a sample of good reconstructed K_{Tau} decays. (Figure 7.6). The second order polynomial fit yielded a better χ^2 and thus was chosen for calculating the normalized vertex quality S_{norm}

$$S_{norm} = \frac{S}{(0.3803 - 1.695 \times 10^{-3} \times Z_{vertex} + 1.518 \times 10^{-6} \times Z^2_{vertex})} \tag{7.5}$$

Two track vertex quality S_{2trk}

Since accidental events could be a result of a superposition of two tracks coming from a kaon decay and a third accidental track, it was reasonable to inspect the event two tracks vertices. We defined a two track vertex quality S_{2trk} as the shortest distance in the X-Y plane between the third track and the vertex formed by the other two tracks. S_{2trk} was defined as the maximum value for the three possible (*2+1*) combinations.

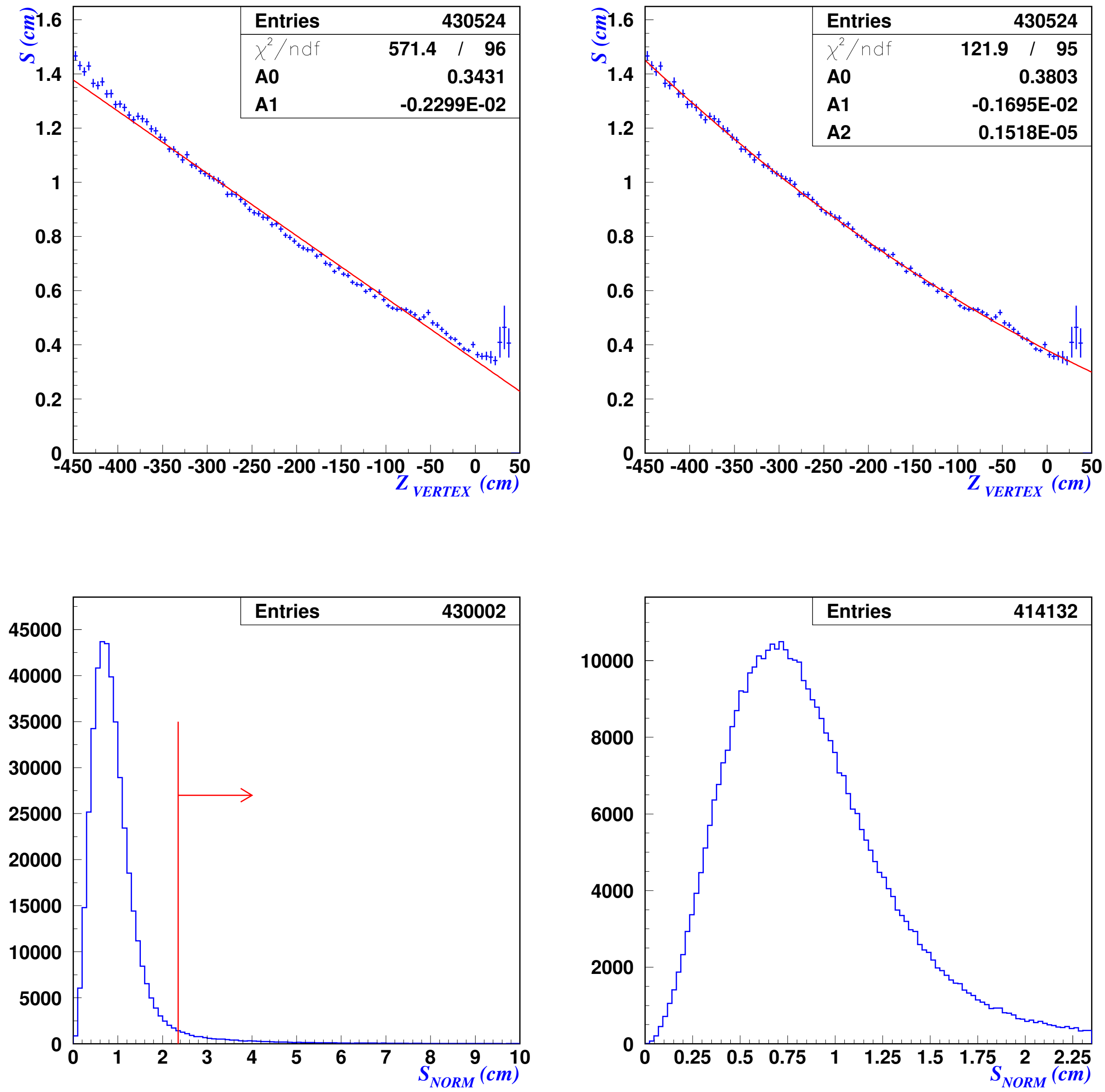


Figure 7.6: Vertex quality S versus Z coordinate of the vertex fitted with a first- (top left) and the second-order (top right) polynomial. Normalized S distribution for good K_{Tau} events (bottom left), and the same within the cut region, $S_{norm} < 2.35$ (bottom right).

Target likelihood

From the momenta of the three charged tracks one can reconstruct the kaon momentum. With this information available one can trace the kaon back to the target by using the first order beam line transfer matrix. Real kaon decays with all of the daughter particles detected are expected to successfully trace back to the target, while accidental events and K_{Dal} decays will mostly miss. The target likelihood was formed by using the five variables describing the phase space of the kaons, namely the kaon momentum P, the position x,y, and the divergence - θ_x, θ_y. It was approximated by three two dimensional distributions: x versus θ_x, y versus θ_y and P versus x (Figure 7.7):

$$\mathcal{L}_{Target} = \mathcal{L}_1(x, \theta_x)\mathcal{L}_2(y, \theta_y)\mathcal{L}_3(P, x) \tag{7.6}$$

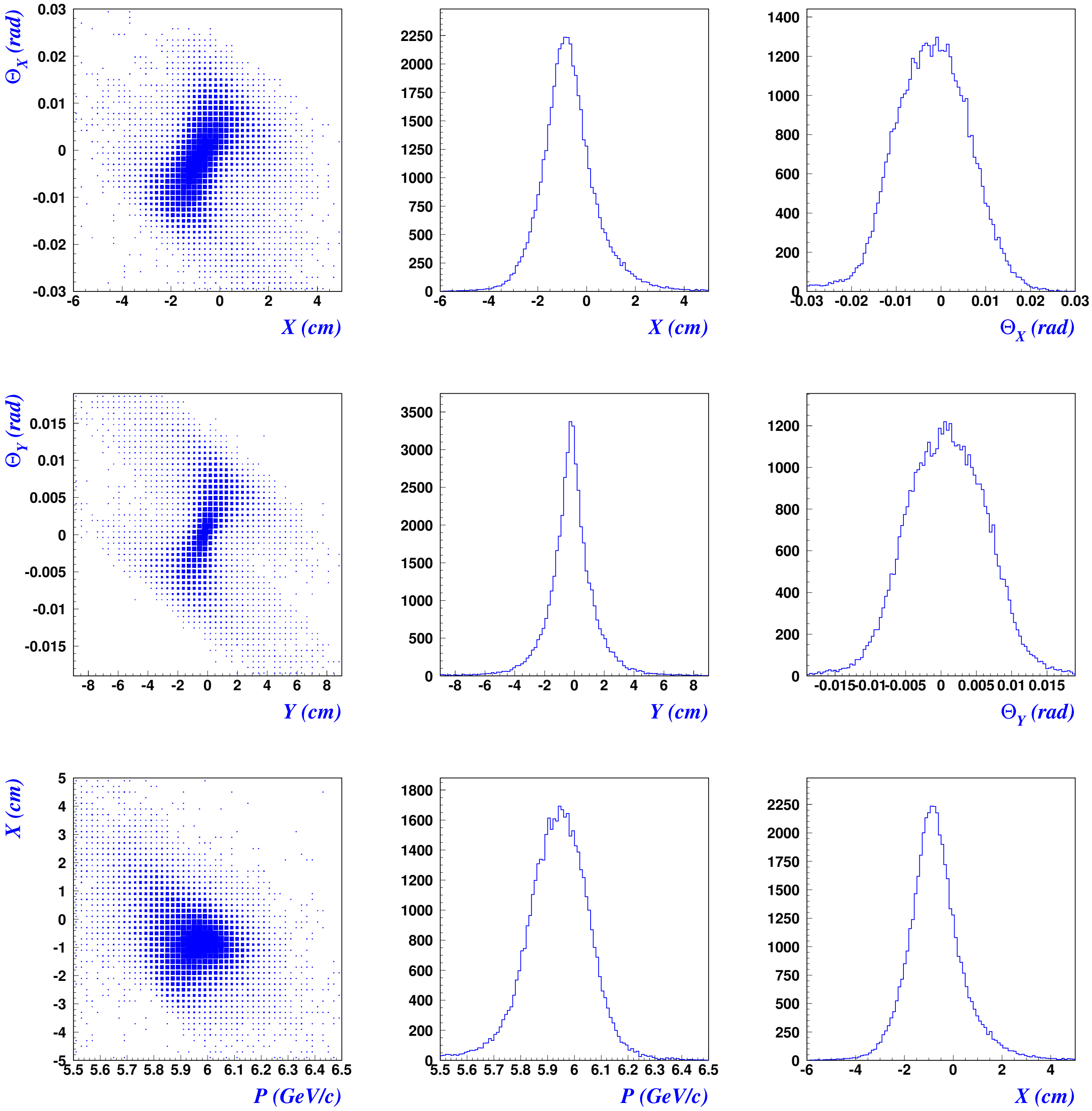


Figure 7.7: Phase space distributions of the kaons at the target. X versus θ_X; Y versus θ_Y; P versus X and corresponding projections. The distributions were obtained using good reconstructed K_{Tau} data events

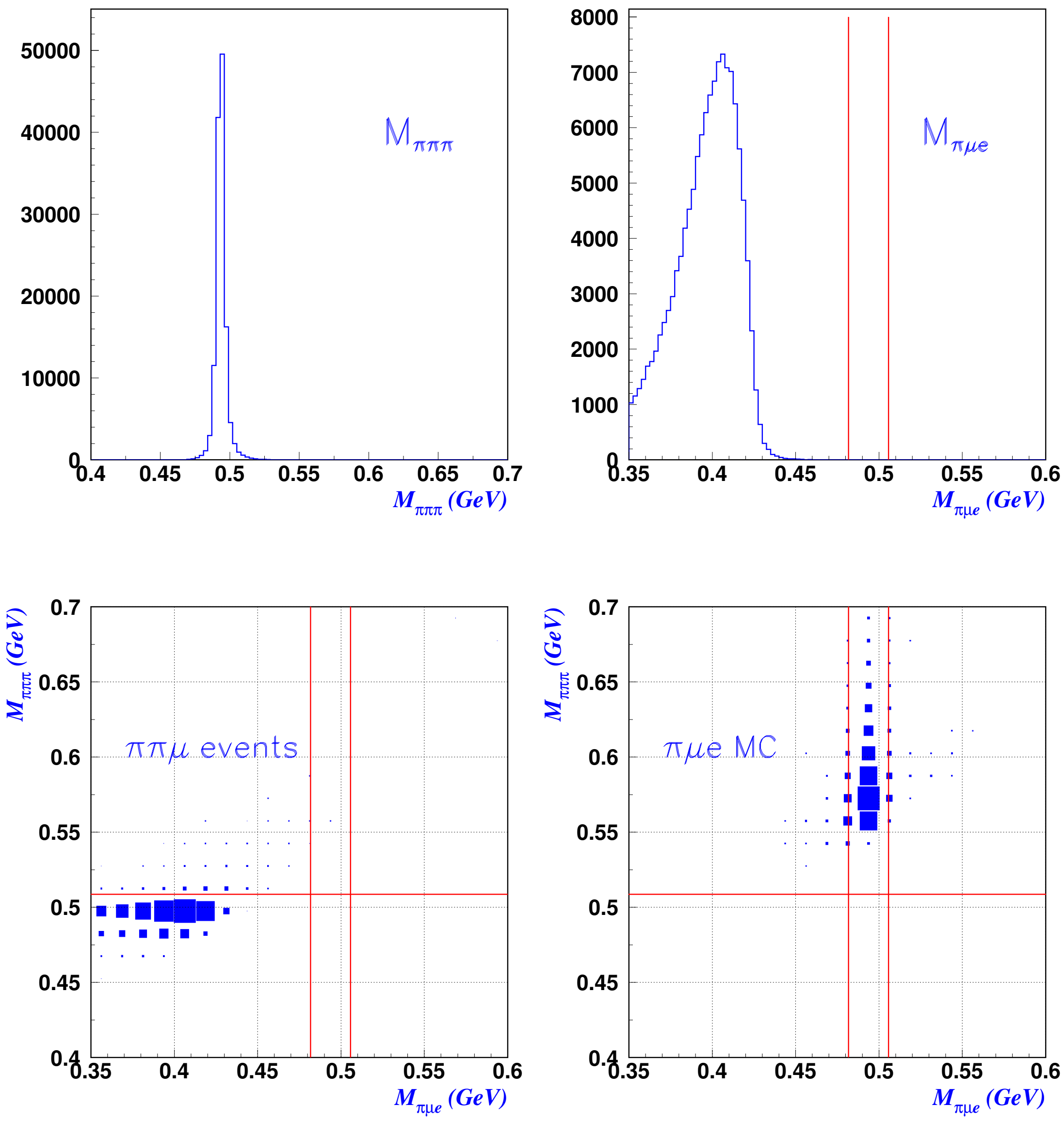


Figure 7.8: $M_{\pi\pi\pi}$ and $M_{\pi\mu e}$ mass spectrum for $\pi\pi\mu$ background events (top). Cross plot of $M_{\pi\pi\pi}$ versus $M_{\pi\mu e}$ for the same background events (bottom left) and Monte-Carlos $K_{\pi\mu e}$ events (bottom right).

Reconstructed kaon mass

From the determined momenta of the tracks, the invariant mass of the kaon was reconstructed assuming the combination of a π^+, a μ^+ and an e^- particles. For the K_{Tau} background we extracted a sample of $\pi^-\pi^+\mu^+$ events and assigned the e^- mass to a π^- track in order to calculate $M_{\pi\mu e}$ mass. For the K_{Dal} background events, a sample of $e^-e^+\mu^+$ events was extracted and the mass of the π^+ was assigned to an e^+ track.

In order to reduce the contamination from the K_{Tau} decays a cut was imposed on the $M_{\pi\pi\pi}$ reconstructed mass (See Figure 7.8 for details).

Invariant e^+e^- mass

The invariant mass of an e^+e^- pair from the K_{Dal} decays is peaked at low values and is bound by the π^0 mass of 135 MeV (Figure 7.9). For $K_{\pi\mu e}$ signal events the invariant mass of e^- and μ^+ or e^- and π^+ (in both cases both particles are assumed to have electron mass) is mostly larger then the π^0 mass. This difference is used to reject K_{Dal} background events.

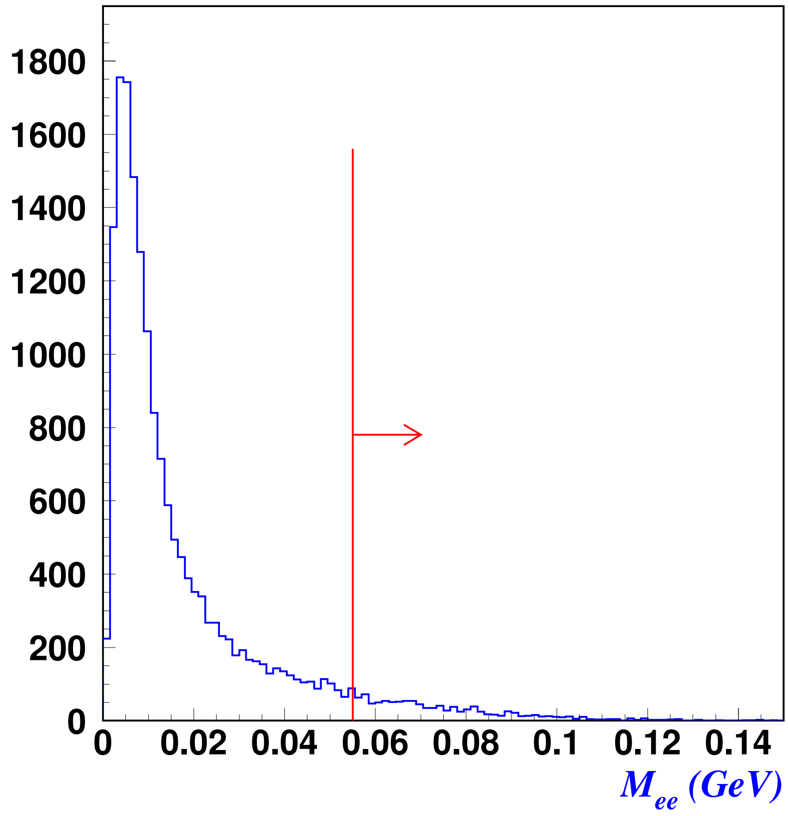


Figure 7.9: M_{ee} mass distribution for K_{Dal} background events. The cut $M_{ee} > 0.055$ GeV removes 94% of the events.

Number of photons

The number of photons, found in the calorimeter, was used to suppress K_{Dal} background, where the photon from the $\pi^0 \to e^+e^-\gamma$ decay was registered. In order to maintain high efficiency for signal events a requirement for time of the hit was also implemented (< 2ns).

r.m.s. track time

The root mean square of the track time was constructed to gauge timing quality of the event. The time of each track was calculated as the mean value of the A and shower counter times:

$$T^{i}_{\mathrm{TRACK}} = \frac{T^{\mathrm{A}}_{i} + T^{\mathrm{SHOWER}}_{i}}{2} \tag{7.7}$$

where i is changing from one to three and represents the track index. The time quality variable was defined the following way:

$$\mathrm{TrkRMS}^2 = \frac{\sum_{i=1}^{3}(T^{i}_{\mathrm{TRACK}} - T_{\mathrm{MEAN}})^2}{3} \tag{7.8}$$

where T_{MEAN} is the average track time:

$$T_{\mathrm{MEAN}} = \frac{\sum_{i=1}^{3} T_{i}^{TRACK}}{3} \tag{7.9}$$

Track χ^2 sum

For each track a χ^2, that characterized the goodness of the track fit, was formed (See section 5.2.3). In order to suppress particle decays in flight ($\pi^+ \to \mu^+\nu_\mu$) a cut on the sum of the track χ^2 was implemented.

Kinematic cuts

The list of kinematic cuts is presented below:

General kinematic cuts

- $S_{norm} < 2.35$ cm
- $S_{2trk} < 2.35$ cm
- Target likelihood > -20
- $\mid M_{\pi\mu e}$ - 0.4937 $\mid < 0.012$ GeV
- χ^2_{Track} sum < 25

Cuts designed to suppress K_{Dal} Background

- $M_{ee} > 0.055$ GeV
- N_γ=0 .or. $T_\gamma > 2$ns

Cut designed to suppress K_{Tau} background

- ($M_{\pi\pi\pi}$ - 0.4937)>0.015 GeV

Timing cuts

- TrkRMS $<$ 1 ns.
- minex gd cltm $>$ 1 ns (minimum time difference between the time of a track pair which is closest in time and the time of the extra clumps in the calorimeter)

7.4 Particle identification likelihoods

The aim of the particle identification procedure is to efficiently identify particles (e^-, π^+ and μ^+) while maintaining low misidentification probabilities (P(π^- as e^-) or P(e^+ as π^+)). Levels of acceptable misidentification probability were determined by the levels of the relative background from other kaon decays.

Below, we list the information that was available for each track from the detector elements, and was used for the PID and likelihoods construction:

- Shower counter
  - E/P (Energy to Momentum ratio)
  - shower shape (ratio of the energy deposited in the seed module to the energy sum deposited in the three nearest modules)
  - E (total energy deposited in the calorimeter)
- Čerenkov counter
  - number of p.e.'s in C1 counter
  - number of p.e.'s in C2 counter
  - time in C1 counter
  - time in C2 counter
- Muon range
  - B_{HIT} - A χ^2 related to the deviation of the B counter hit position from the track projection.
  - ratio of the number of observed and expected hits in the muon range system.

The likelihood functions for the various particular particle types were constructed from unbiased data samples of those particles. The same samples were used to determine efficiencies and misidentification probabilities.

7.4.1 E/P in the calorimeter

The energy to momentum ratio E/P served as a powerful tool to discriminate between pions/muons and electrons. It peaks around unity for electrons while for pions it peaks at the lower values for minimum ionization and has a tail extending to higher values from hadronic interactions.

For a final selection we used a set of wide cuts. Later the likelihood function provided further separation.

- π^+ : E/P < 0.9
- e^- : $0.65 <$ E/P < 1.5

Electrons

The measured E/P distribution for an electron could be successfully fitted with a mix of two Gaussian functions :

$$F_e = \mathrm{P}_1 e^{-(x-\mathrm{P}_2)^2/2\mathrm{P}_3{}^2} + \mathrm{P}_4 e^{-(x-\mathrm{P}_5)^2/2\mathrm{P}_6{}^2} \tag{7.10}$$

The mean and σ for both Gaussians depend on the particle momentum. This is a direct consequence of the momentum dependence of the resolution.

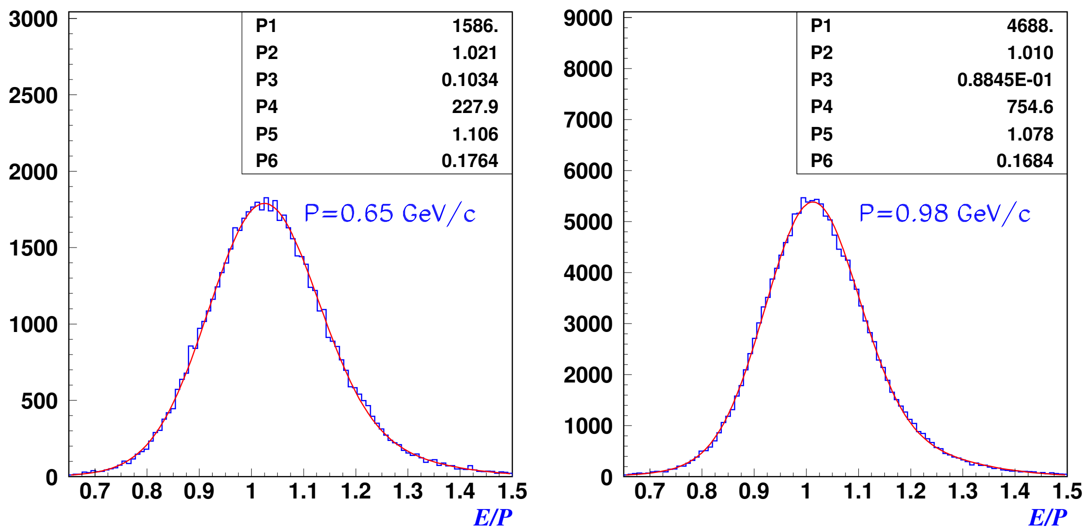


Figure 7.10: E/P distribution for electrons fitted with two Gaussians for the different momenta.

In order to extract the p.d.f. , the electron sample was divided into six momentum bins (P_e). For each momentum bin, a fit using Eq. (7.10) was performed (Figures 7.10,7.11). Then the parameters were fitted with second order polynomials of the mean momentum (Figure 7.12).

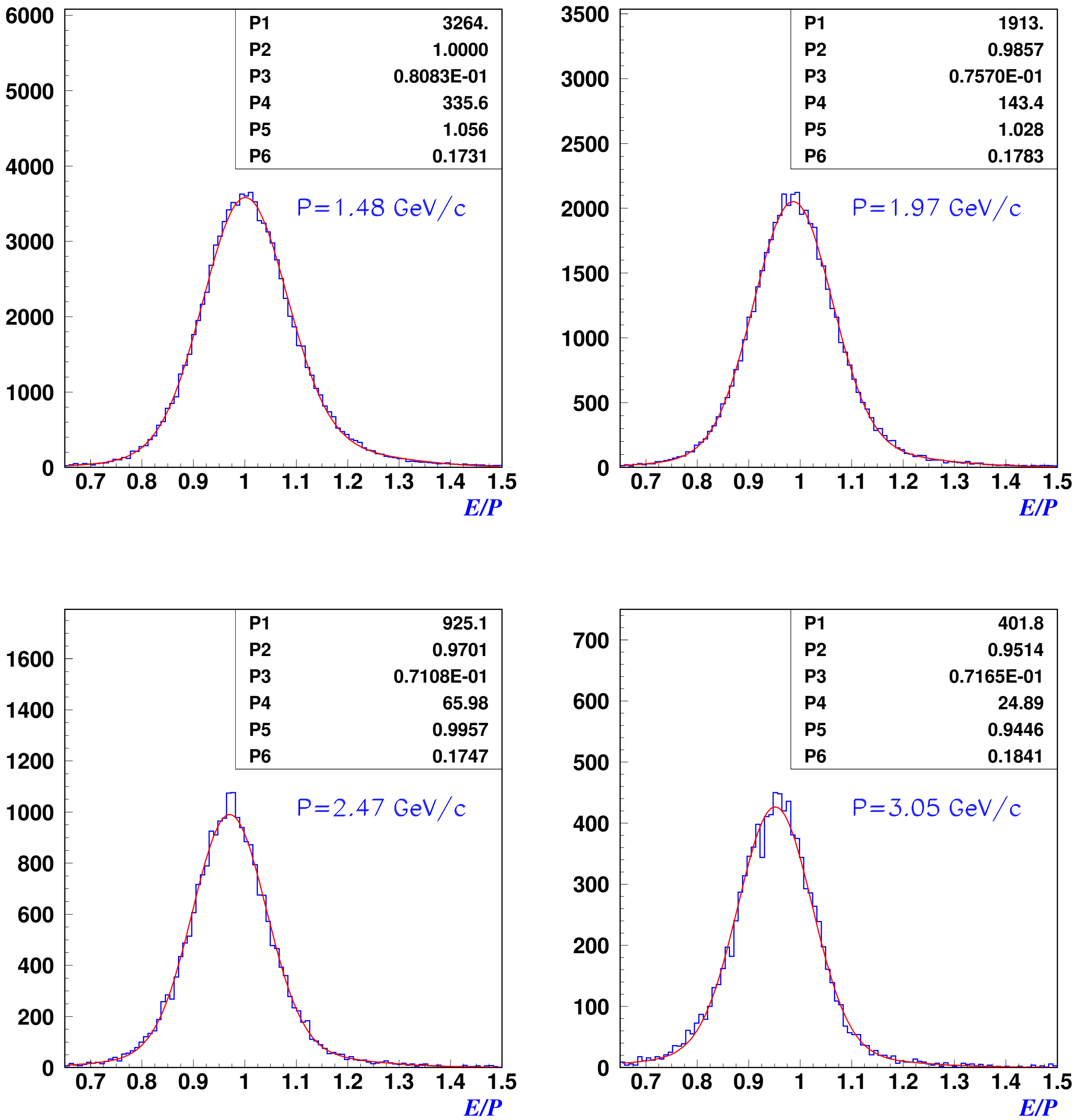


Figure 7.11: E/P distributions for electrons fitted with two Gaussians for the different momenta.

Pions

Following the same logic as with electrons, a sample of pions was first divided into five momentum bins. For each bin the E/P distribution was fitted with the sum of three Gaussians [Eq. (7.11)], (see Figures 7.13).

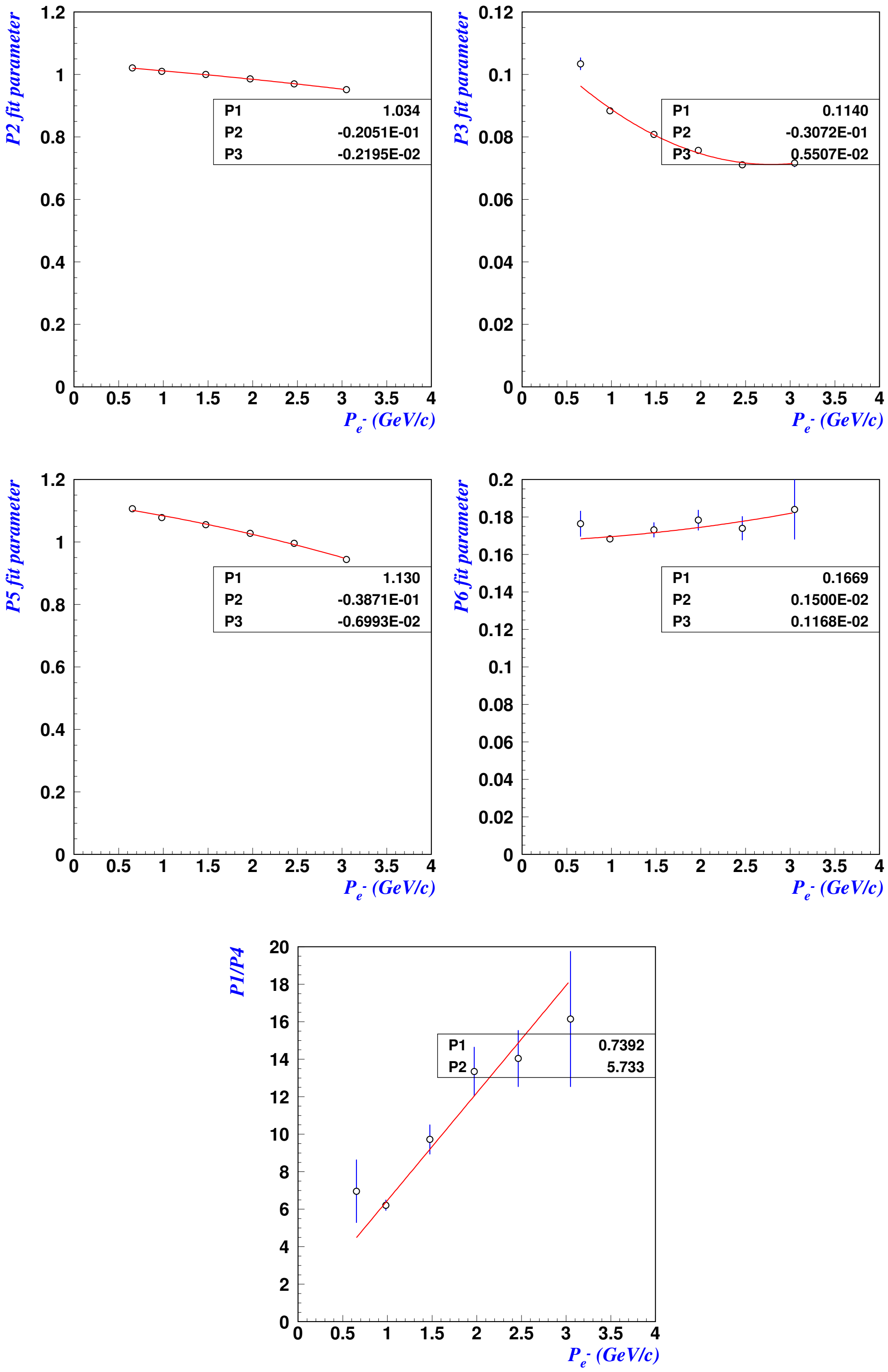


Figure 7.12: Dependence of the parameters of the fit to the electron E/P distribution on the momenta.

$$F_\pi = \mathrm{P}_1 e^{-(x-\mathrm{P}_2)^2/2{\mathrm{P}_3}^2} + \mathrm{P}_4 e^{-(x-\mathrm{P}_5)^2/2{\mathrm{P}_6}^2} + \mathrm{P}_7 e^{-(x-\mathrm{P}_8)^2/2{\mathrm{P}_9}^2} \tag{7.11}$$

The parameters of the fit were later fitted against the mean momentum of the particles P_π for each momentum bin (see Figures 7.14).

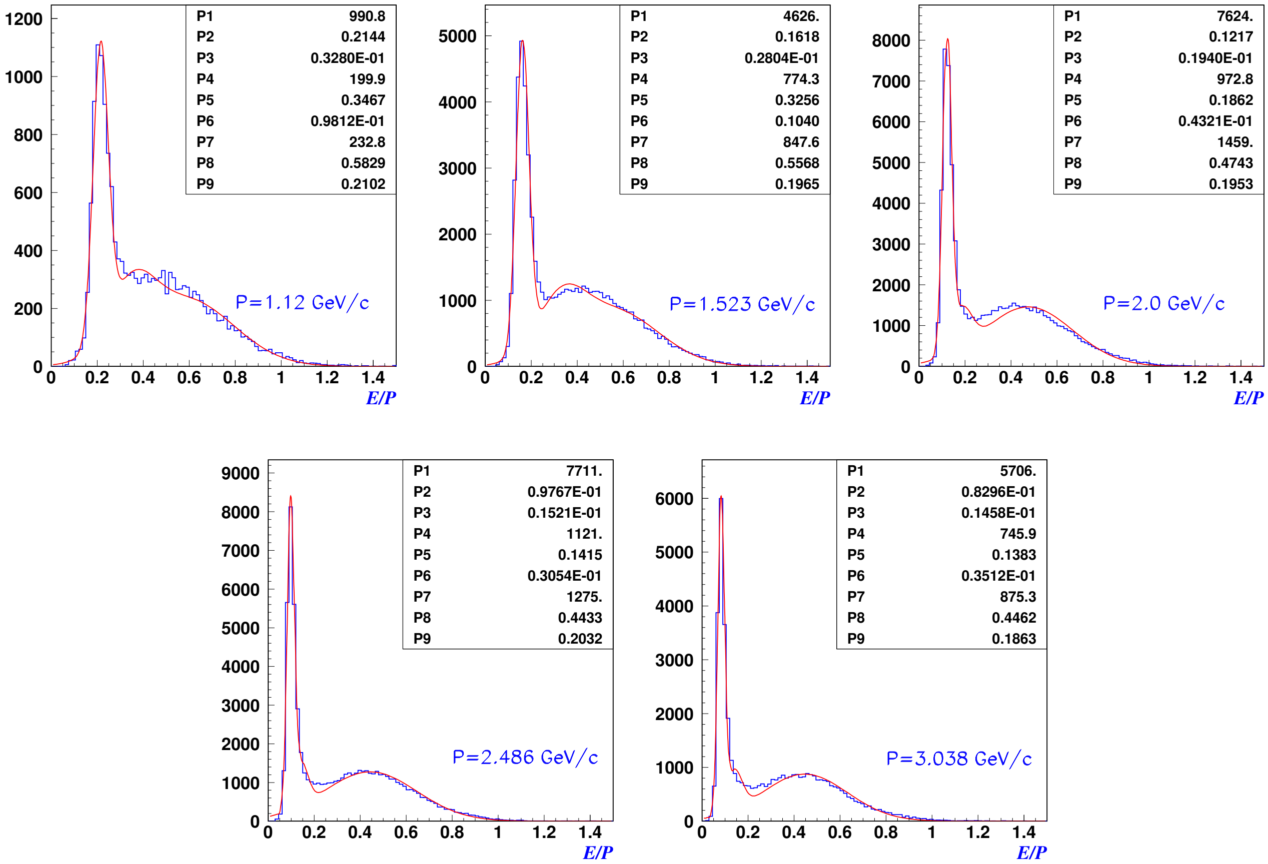


Figure 7.13: E/P distributions for pions fitted with three Gaussians for the different momenta.

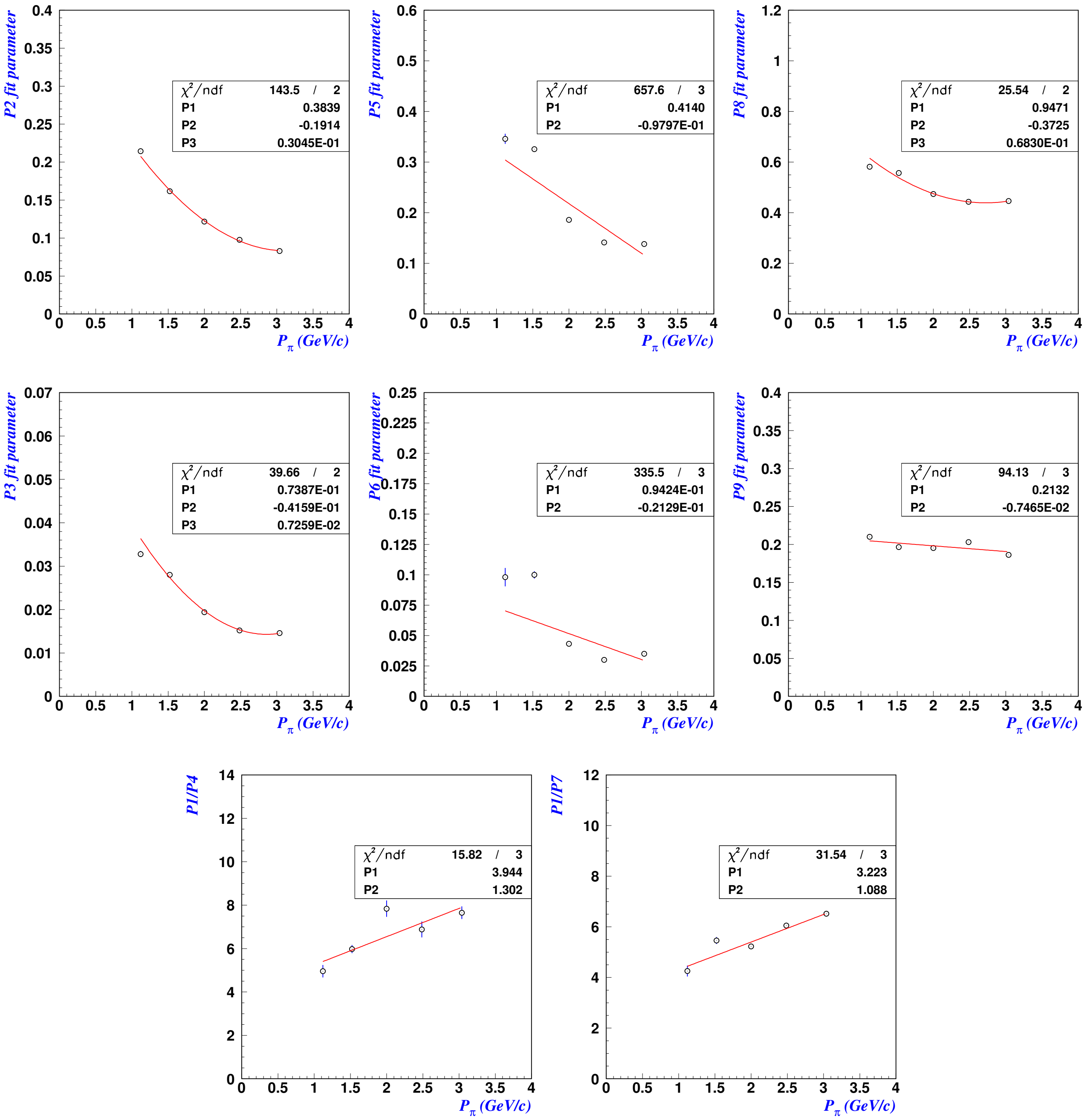


Figure 7.14: Dependence of the fit parameters of the pion E/P distribution on momentum.

7.4.2 Shower shape

Electrons produce an electromagnetic shower in the calorimeter. At the typical electron momenta in this experiment most of the energy would be contained in a single calorimeter module. Pions on the other hand can interact hadronically. This spreads the shower and the total deposited energy can be divided among neighbouring modules. This information can be useful, especially considering the fact that E/P for a pion, that interacted hadronically in the calorimeter, can reach unity and thus be similar to that one of an electron.

Obviously, the spread of energy among the modules depends on the position of the particle's entrance into the module. By using the track parameters one could determine the position of the track at the shower counter. Based on that information the hit in the module was classified into three different types. A calorimeter module was divided into 16 squares. If a particle's trajectory would pass through the four squares in the corners, a hit was classified as type three. Similarly the four squares in the center classified a hit as type one and the rest of the squares made it a type two hit (See Figure 7.15).

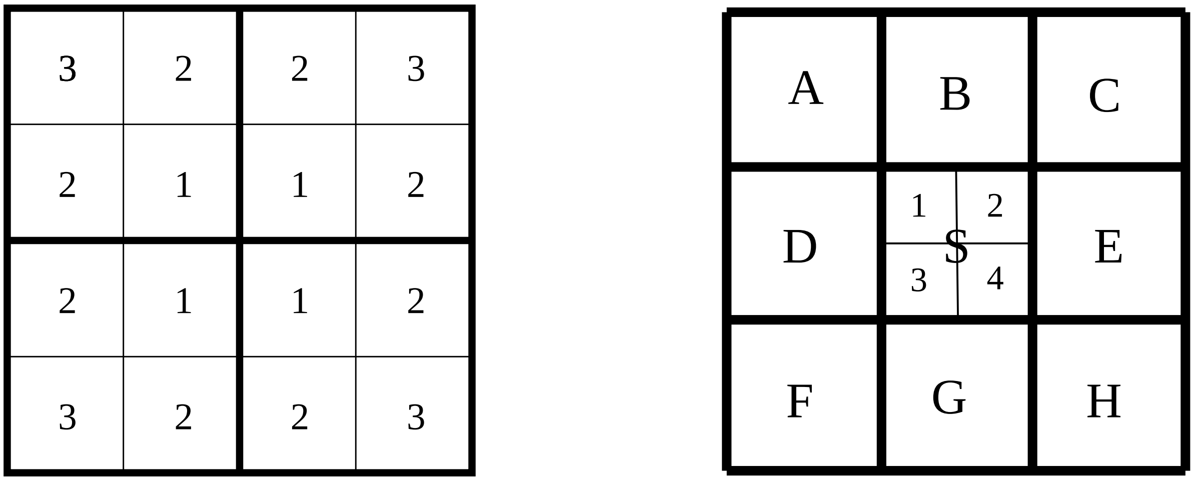


Figure 7.15: Left) A single shower module divided into 16 squares of equal area. The showers are classified by numbers 1 through 3, depending on the impact position deduced from the track. Right) An array of nine modules with the seed module S divided into four quadrants. An energy ratio is determined as the ratio of the energy of the seed module to the sum of the energies of seed and the three modules adjacent to the hit quadrant. For example, if the hit is in quadrant one, energy ratio is the ratio of the energy of the S module to the sum of the energies from S, D, A, and B modules.

The shower shape was described by a variable called *ENERGY RATIO* and defined as a ratio of the energy left in the seed module to the sum of energy left in the seed and the three most adjacent modules. Such ratio is bound by unity, which would indicate that all the energy was released in the seed module.

The newly defined variable (*ENERGY RATIO*) depends on the energy of a particle.

In order to account for that, an additional classification was implemented, based on the particle E/P ratio:

- $0 < \mathrm{E/P} < 0.45$
- $0.45 < \mathrm{E/P} < 0.9$
- $0.9 < \mathrm{E/P} < 1.5$

Separate p.d.f.'s were constructed for each E/P bin, and type of hit. For the muon p.d.f only the first E/P bin was relevant, since the majority of muons (99%) had an E/P ratio of less then 0.45.

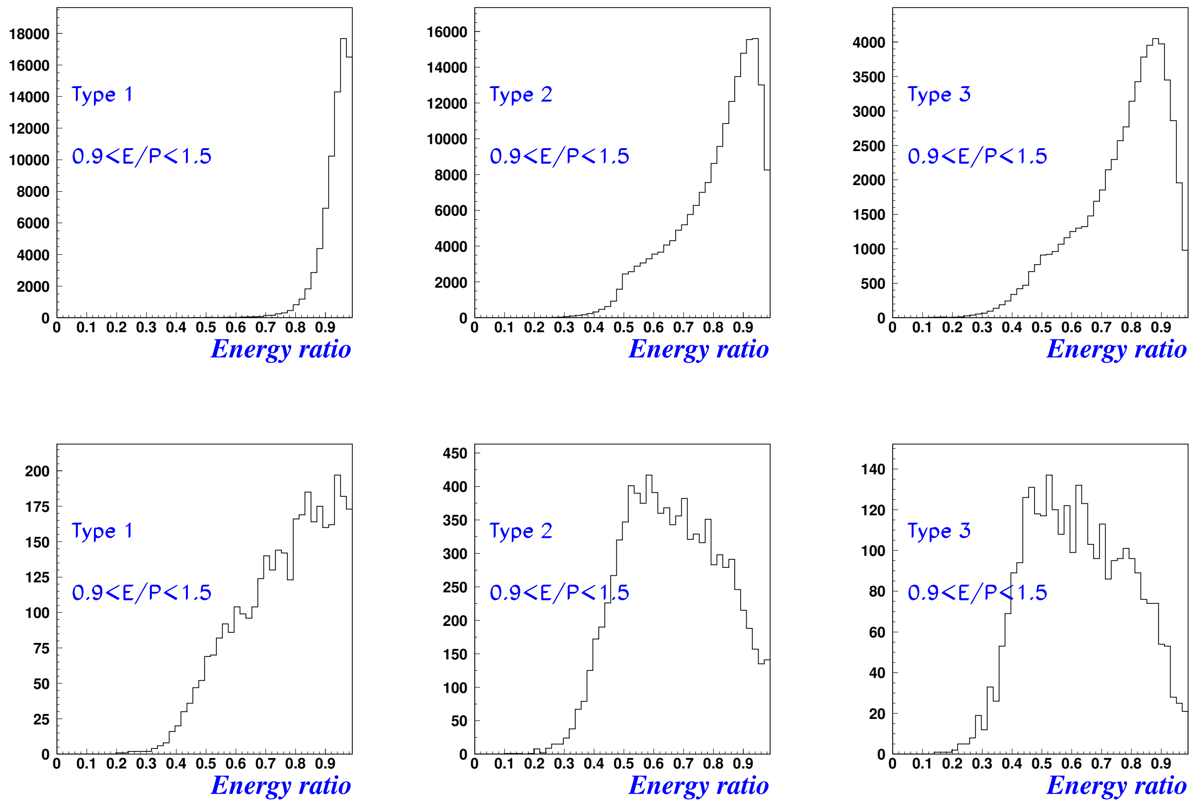


Figure 7.16: Energy Ratio distributions for E/P range of 0.9 to 1.5. Top and bottom rows represent electrons and pions respectively. Columns represent different shower hit types from 1(left column) to 3 (right column).

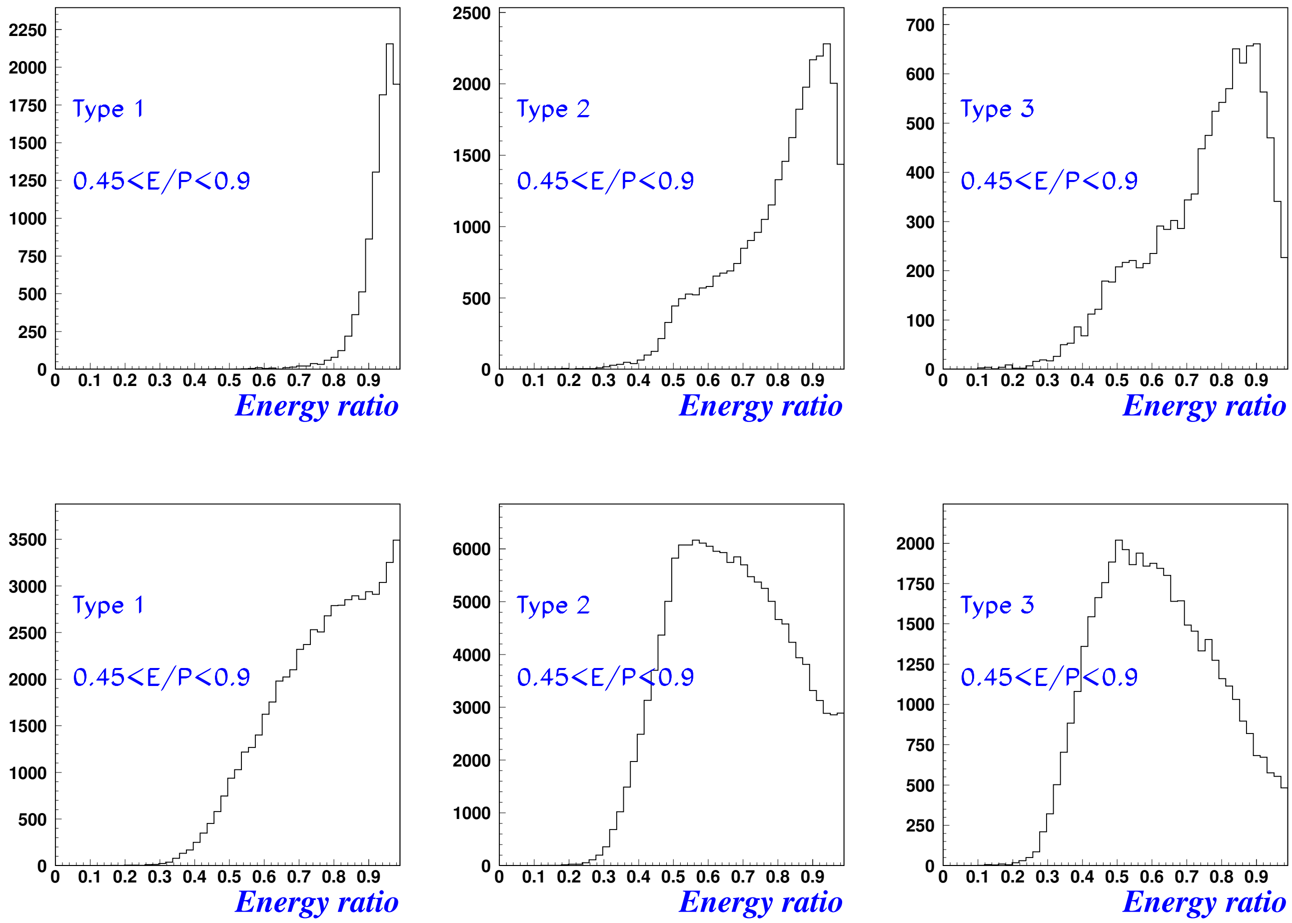


Figure 7.17: Energy Ratio distributions for E/P range of 0.45 to 0.9. Top, and bottom rows represent electrons and pions respectively. Columns represent different shower hit types from 1(left column) to 3 (right column).

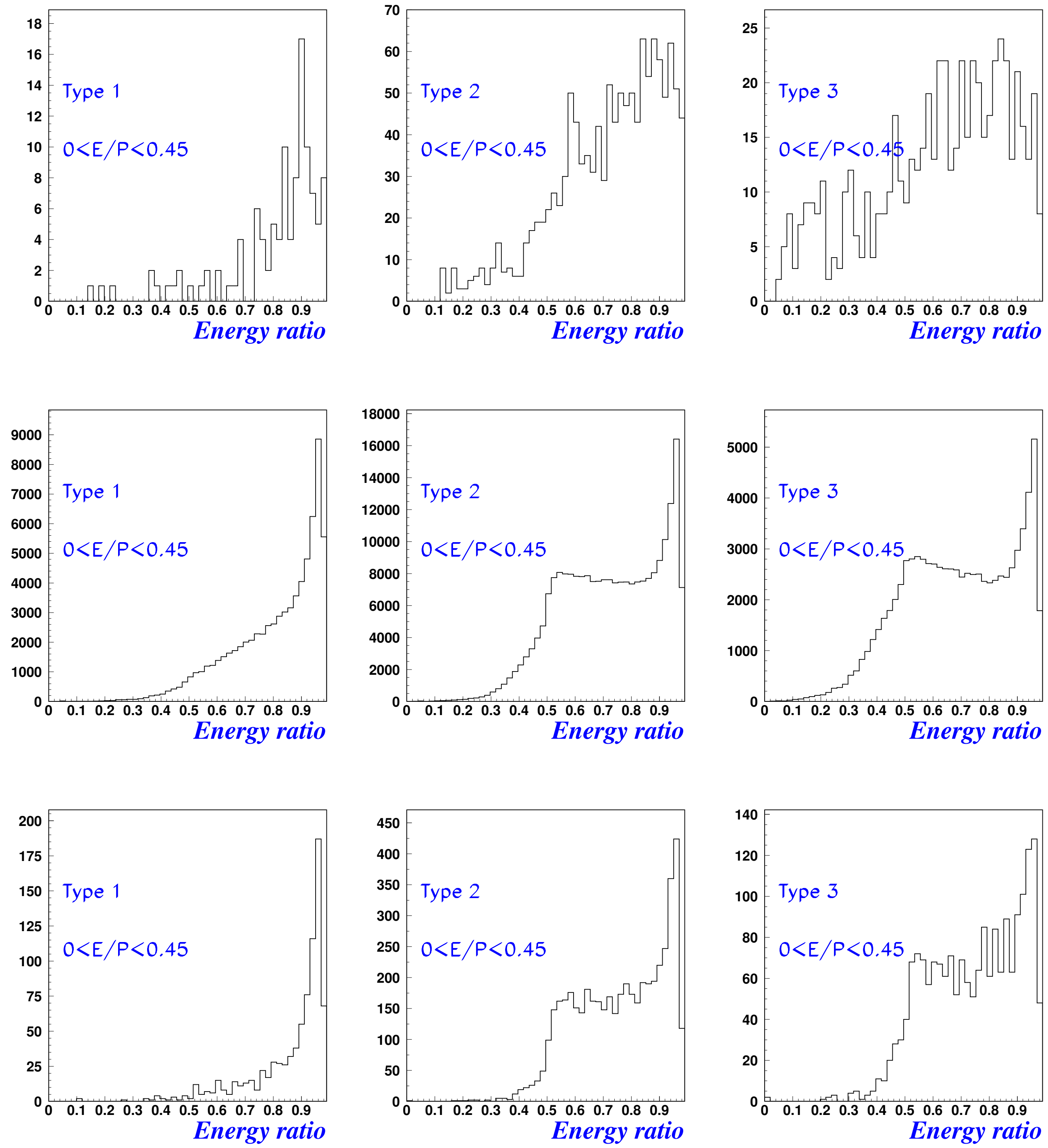


Figure 7.18: Energy Ratio distributions for E/P range of 0 to 0.45. Top, middle, and bottom rows represent electrons, pions, and muons respectively. Columns represent different shower hit types from 1(left column) to 3 (right column).

7.4.3 Čerenkov counter

The Čerenkov counter was used for identifying electrons on the left side and for rejecting positrons on the right side. A preliminary data study indicated that adequate background rejection could be achieved without invoking the auxiliary electronics (see Sect. 3.4.1)[2]. Therefore, only the "regular" ADCs and TDCs were used for particle identification. The ADC information was converted into the number of photoelectrons (p.e.) registered by the PMTs spatially associated with the tracks. The TDC information was transformed into the time of the hit in (ns). A slewing correction for the time of the hit in the Čerenkov counter was implemented using, extracted from the data, correlation between the ADC amplitude and the TDC time.

Electron and negative pion

The left side of the Čerenkov counter was used to identify electrons and reject negative pions. It was crucial to reject pions since misidentification of one as an electron would lead to a large number of K_{Tau} events being falsely identified as $K_{\pi\mu e}$. An electron was seen as a particle producing light. A pion, however, could also be identified as an electron, especially if a real electron would pass through the same region. In order to reduce such background, events with extra tracks, passing through the same region (spatially defined by the appropriate PMT coverage) as an electron candidate track, were rejected. This requirement reduced the misidentification probability P(π^- as e^-) roughly by a factor of two. CERENK triggers were used to determine the e^- identification efficiency, since they required only three out of four counters (C1L,C2L,C1R,C2R) to have a hit. To determine the misidentification probability P(π^- as e^-) a sample of π^- from $K_{\pi\pi\pi}$ decays selected from the TAU triggered data was used. We defined a set of cuts for e^- identification as follows:

- C1 p.e. > 0.3 .AND.
- C2 p.e. > 0.3 .AND.
- $|$ C1 time $| < 5$ ns .AND.
- $|$ C2 time $| < 5$ ns

Likelihood functions for e^- and π^- contained the number of photo electrons and time in both C1 and C2 counters. To create appropriate p.d.f.'s, an e^- and a π^- response in the Čerenkov counters was stored in the form of binned histograms.

Correlation between the response in C1 and C2 counters was not included in the likelihood, but we took that into account when testing the likelihood and assessing the misidentification probability.

[2]This information was added to the text after it was submitted to the University of Zurich archives

It was noticed that the time distribution of electrons was skewed to late (positive) times, even after slewing corrections were applied. We explained this effect as a consequence of coating the counter's PMTs with a wavelength shifter (WLS), which was done in order to increase efficiency. On the right side of the counter the effect was less pronounced. This can be explained by the fact that with a larger average number of photoelectrons (approximately 5.6 p.e. in C2) the reemittance delay of photons in the WLS would be less noticeable. On the left side of the counters, the average number of p.e.'s was about 2.6.

In order to check whether our assumption was true, we constructed a function, which included the decay constant of the WLS as a parameter and thus simulated a delay caused by reemittance of a photon in the WLS. The resulting function was defined as follows:

$$T_{WLS}(x) = \int_{-\infty}^{x} p_1 e^{-\frac{(y-p_2)^2}{2p_3^2}} \cdot e^{-\frac{x-y}{p_4}} \cdot dy \tag{7.12}$$

The quality of the fit to the Čerenkov time distribution was perfect (See Figure 7.19). The parameter ($P4$), which represented the decay constant of the WLS, had a fitted value of 1.5ns, in agreement with what is expected for P-Terphenyl, which was used as a WLS.

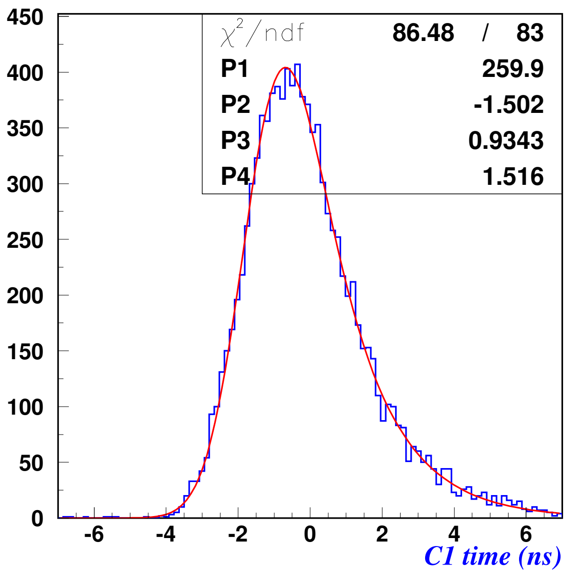


Figure 7.19: C1 time distribution of an e^- track with the number of p.e. < 1.5 fitted with the Gaussian-exponential hybrid [Eq. (7.12)].

Positive pion and positron

On the right side of the Čerenkov counter it was crucial to veto e^+, since misidentification of an e^+ as a π^+ would result in a large number of false $K_{\pi\mu e}$ events originating from the K_{Dal} decays. Any signal having associated light (> 1.2 p.e.) or time within 5ns of the trigger time in any of the Čerenkov counters on the right (C1R,C2R) was rejected.

The misidentification probability was calculated using a sample of positrons extracted from the reconstructed $K^+ \to \pi^0\pi^+$ decays, using CERENK triggers. In order to account for spatial correlations in the detector inefficiencies the counter acceptance was divided into spatial regions defined by the PMT coverage. In addition the E/P misidentification probabilities were calculated for tracks that went through the regions defined by the coverage of C2 PMTs. That allowed to calculate the overall misidentification probability for spatially correlated regions of the C1R, C2R, and the shower counters. The positive pion likelihood function was constructed using only the calorimeter information (E/P, Energy Ratio).

7.4.4 Muon system

In order to identify a muon we used the information available from the muon range and the shower counter. The main goal was to have high μ^+ identification efficiency, while keeping a reasonable rejection rate for π^+. Since a muon would act like a minimum ionizing particle in the calorimeter a cut was imposed on the energy deposited in the calorimeter. Besides, the location of the B counter hit had to be consistent with the track, and an appropriate χ^2 was formed (See section 5.2.6). Finally, knowing the muon momentum one can predict the range in the muon system. The predicted number of hits was compared to the observed number (See Figure 6.1). The set of final muon selection cuts is presented below:

- $B_{HIT} < 12$
- Hits observed / Hits expected > 0.5
- $E_{Calorimeter} < 0.45$ GeV

In order to estimate the misidentification probability P(π^+ as μ^+) we used a sample of π^+ from the reconstructed $K^+ \to \pi^0\pi^+$ decays.

7.4.5 Control samples

Table 7.2 summarizes the control samples for specific particles and the associated decay modes, along with the trigger latches.

A sample of e^- was created using the CERENK latched data by selecting reconstructed $K^+ \to \pi^0\pi^+$ decays. It allowed to extract an unbiased e^- response in the Čerenkov

Table 7.2: Trigger latches and decay modes used for clean particle samples

Particle Type	Decay mode	Trigger Latch
e^-	$K^+ \to \pi^0\pi^+$, with $\pi^0 \to e^+e^-\gamma$	CERENK
π^-	$K^+ \to \pi^+\pi^+\pi^-$	TAU
e^+	$K^+ \to \pi^0\pi^+$, with $\pi^0 \to e^+e^-\gamma$	CERENK
π^+	$K^+ \to \pi^+\pi^+\pi^-$ / $K^+ \to \pi^0\pi^+$	TAU / EEPS
μ^+	$K^+ \to \pi^0\mu^+\nu$, with $\pi^0 \to e^+e^-\gamma$	EEPS

counters. By requiring light in C1R, C2R and C2L counters the trigger condition was satisfied, thus response given in C1L was unbiased. In addition, while selecting the $K^+ \to \pi^0\pi^+$ decay, no PID cuts were applied to the negative track (e^-). That created an unbiased response in the calorimeter as well.

The e^+ sample was created also using the CERENK latched data with the selection criteria similar to the e^- sample.

Samples of π^- and π^+ were created using the TAU latched data by kinematically selecting $K_{\pi\pi\pi}$ decays. The addition of the cuts vetoing muon tracks ensured clean pion samples. A possible contamination from the K_{Dal} decays was estimated to be below 10^{-8}. In order to get the calorimeter response an additional sample of π^+ from the reconstructed $K^+ \to \pi^0\pi^+$ decays was used (EEPS trigger), too.

A sample of μ^+ was created using the EEPS trigger data by selecting reconstructed $K^+ \to \pi^0\mu^+\nu$ decays. It was impossible to completely rid the sample from the contamination of the $K^+ \to \pi^+\pi^0\pi^0$ decays and thus the efficiency of the muon cuts was determined using Monte-Carlo simulation. The efficiencies of the muon chambers, however were derived from the measured data. For a small sample of μ^+ extracted from the $K^+ \to \pi^0\mu^+\nu$ decays the efficiency of the muon cuts was found in agreement with the one determined from the Monte Carlo within 2%.

7.4.6 PID selection criteria

After the particle identification procedures are discussed, we restate the set of PID cuts designed for the pass3 phase.

PION

- C1pe $<$ 1.2 p.e.
- C2pe $<$ 1.2 p.e.
- |C1 time| $>$ 5 ns
- |C2 time| $>$ 5 ns
- E/P $<$ 0.9
- $\overline{MuonCuts}$

MUON

- C1pe $<$ 1.2 p.e.
- C2pe $<$ 1.2 p.e.
- $B_{HIT} < 12$
- Hits observed / Hits expected $>$ 0.5
- $E_{Calorimeter} < 0.45$ GeV

ELECTRON

- C1pe $>$ 0.3 p.e.
- C2pe $>$ 0.3 p.e.
- |C1 time| $<$ 5 ns
- |C2 time| $<$ 5 ns
- 0.65 $<$ E/P $<$ 1.5

The calculated efficiencies and misidentification probabilities for these cuts are given Table 7.3. The likelihood function for e^- and π^- included E/P, $ENERGY\ RATIO$, p.e. in C1 and C2, and time in C1 and C2. The likelihood function for e^+ and π^+ included just calorimeter information, namely, E/P and $ENERGY\ RATIO$.

Using appropriate particle samples (e^-, π^-, π^+, e^+) , the likelihood functions were tested and their performance can be observed in Figures 7.24,7.25. No likelihood cuts were determined since backgrounds relevant for a particular misidentification were not estimated at this point.

Table 7.3: Identification efficiencies and probability of misidentification for the PID cuts (See section 7.4.6). The symbol "$\rightarrow$" reads "identify as".

	$\rightarrow \pi^+$	$\rightarrow \mu^+$	$\rightarrow e^-$
π^+	0.804 ± 0.008	0.066 ± 0.001	-
π^-	-	-	$(8.7 \pm 2.6) \times 10^{-6}$
μ^+	-	0.79 ± 0.01	-
e^+	$(1.05 \pm 0.08) \times 10^{-5}$	$< (1.05 \pm 0.08) \times 10^{-5}$	-
e^-	-	-	0.767 ± 0.003

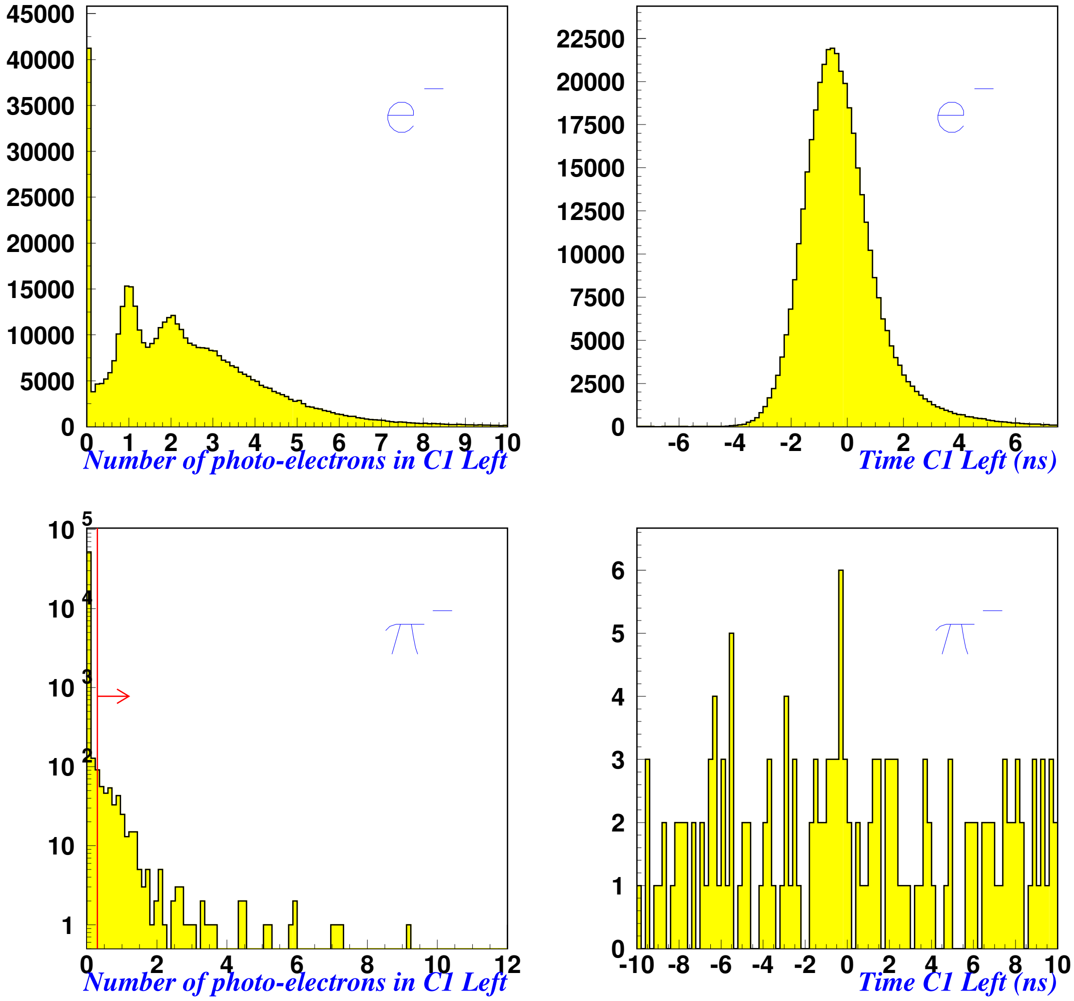


Figure 7.20: Distributions of number of photoelectrons (left) and time (right) in the C1L for e^- (up) ad π^- (bottom).

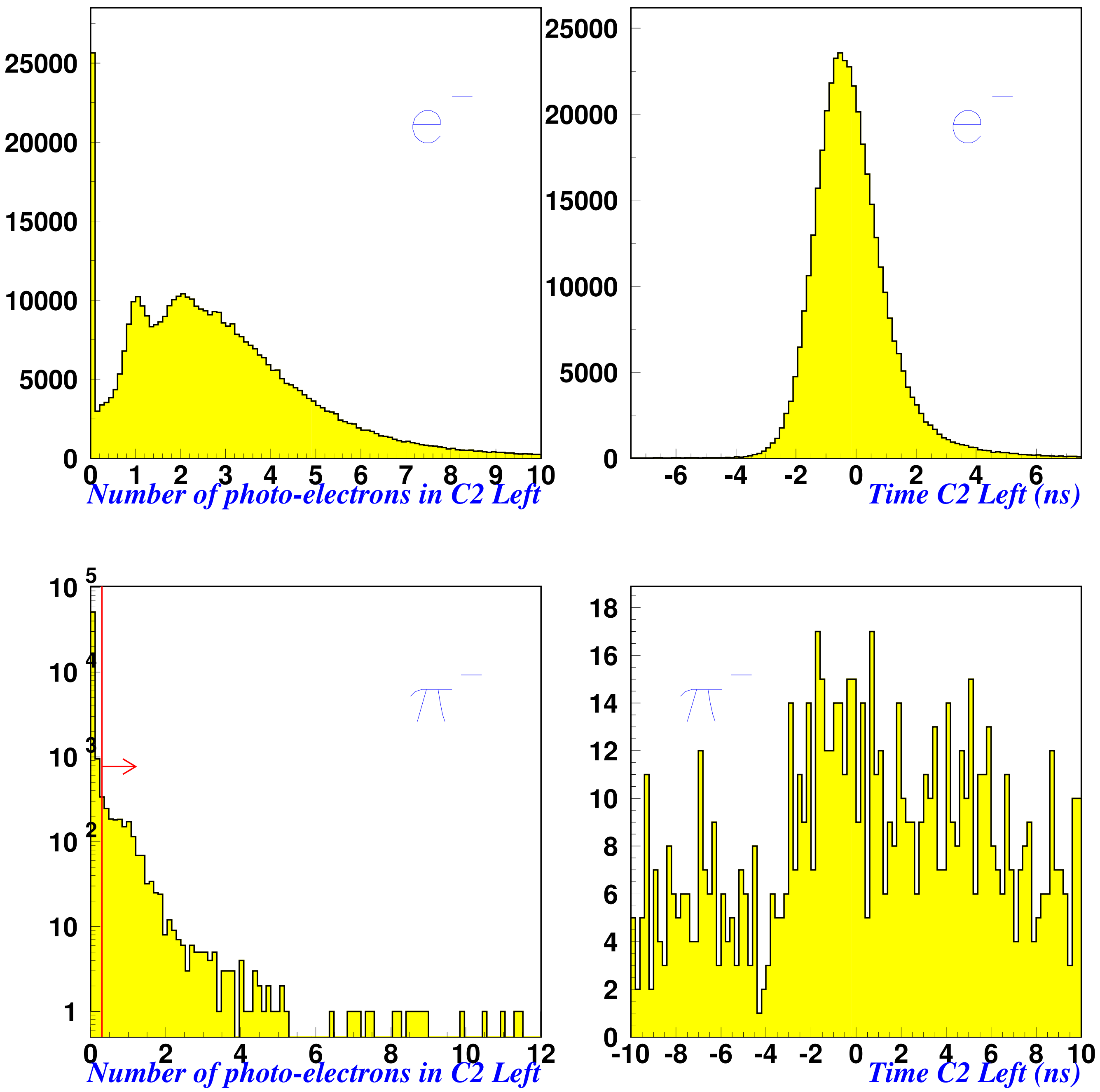


Figure 7.21: Distributions of number of photoelectrons (left) and time(right) in the C2L for e^- (up) ad π^- (bottom).

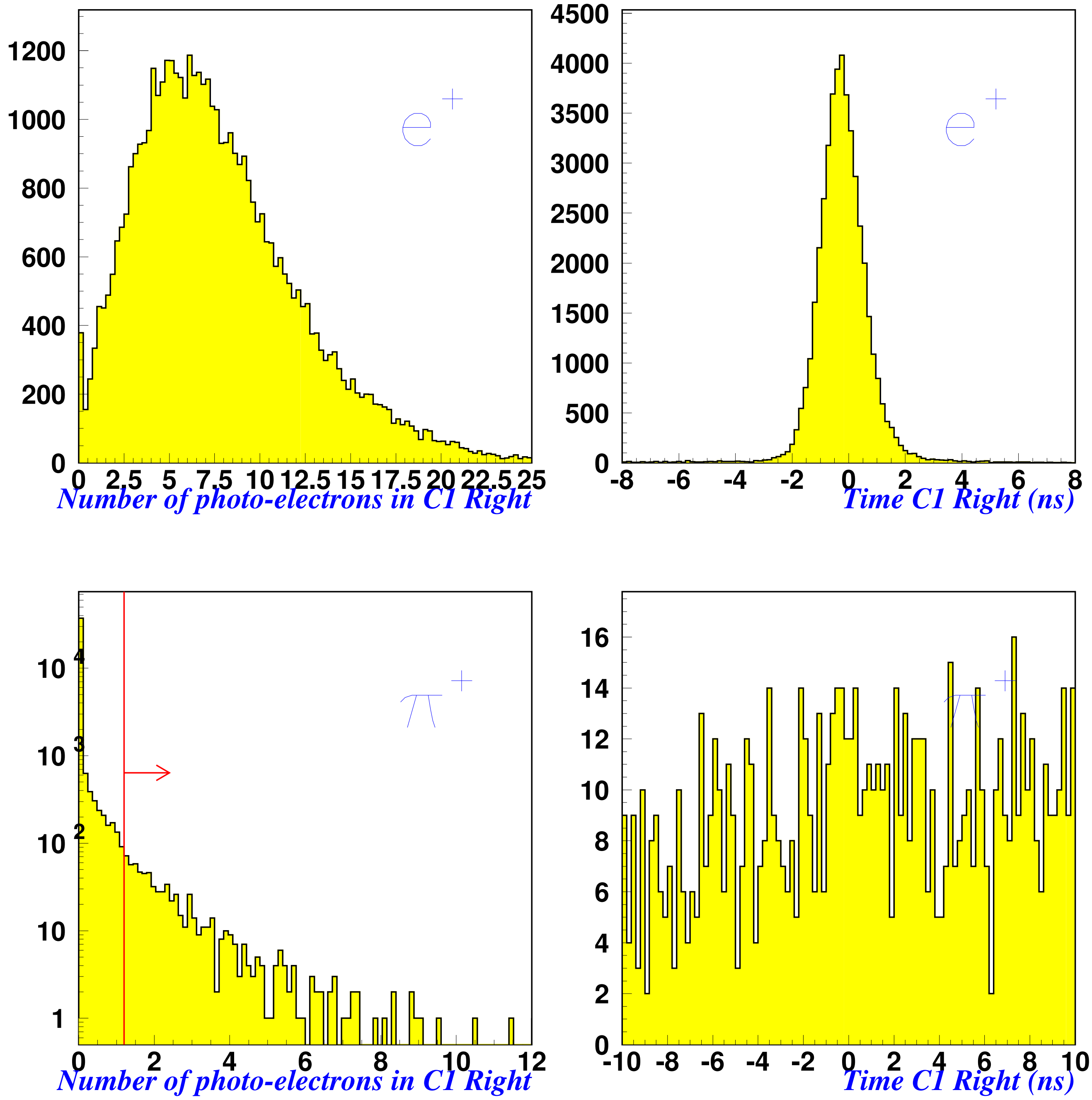


Figure 7.22: Distributions of number of photoelectrons (left) and time (right) in the C1R for e^+ (up) ad π^+ (bottom).

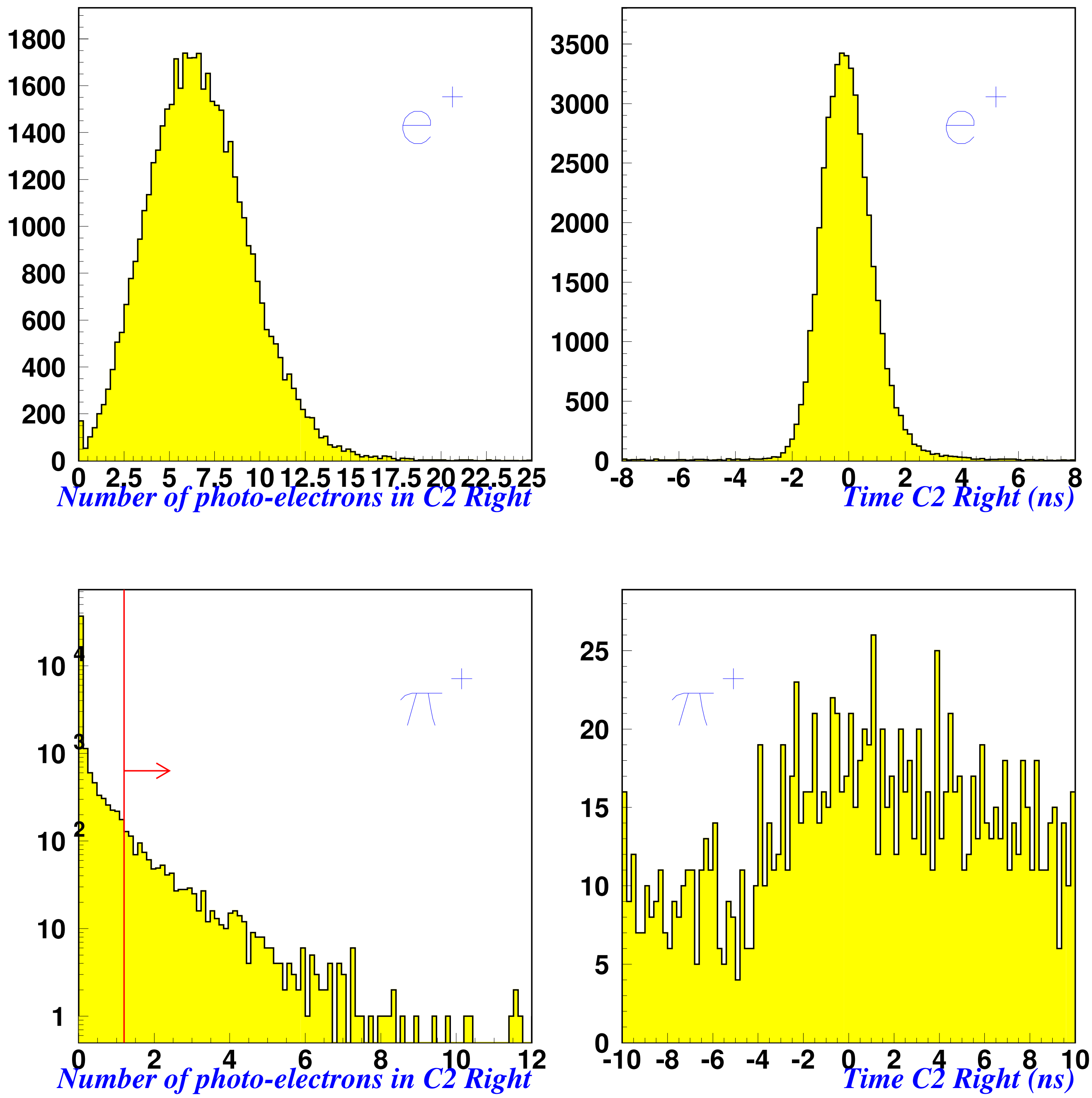


Figure 7.23: Distributions of number of photoelectrons (left) and time (right) in the C2R for e^+ (up) ad π^+ (bottom).

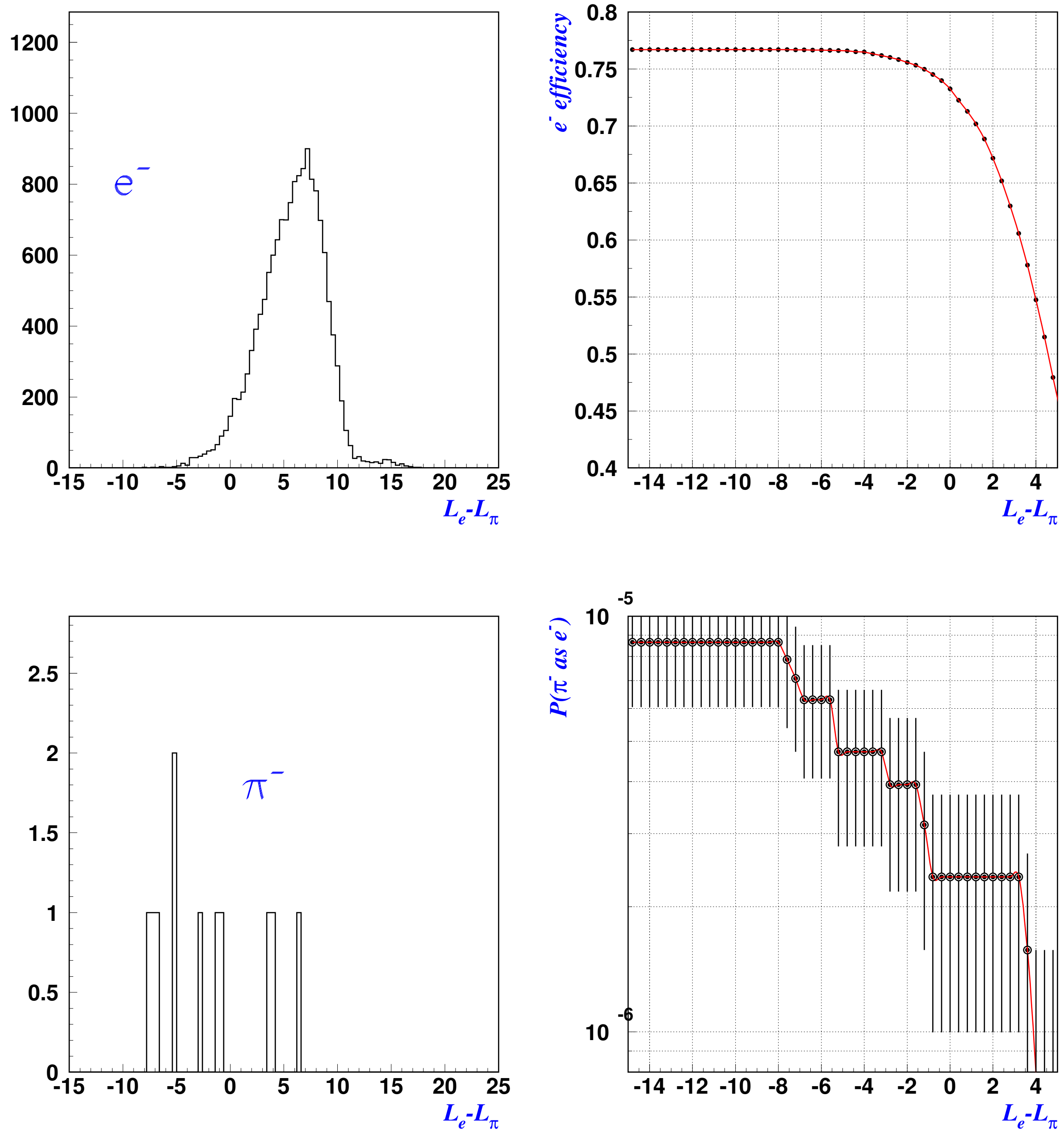


Figure 7.24: Total relative log likelihood $L_e - L_\pi$ distributions for a sample of e^-(top left) and π^- (bottom left). Efficiency of e^- identification versus log likelihood cut (top right) and misidentification probability P(π^- as e^-) versus log likelihood cut (bottom right).

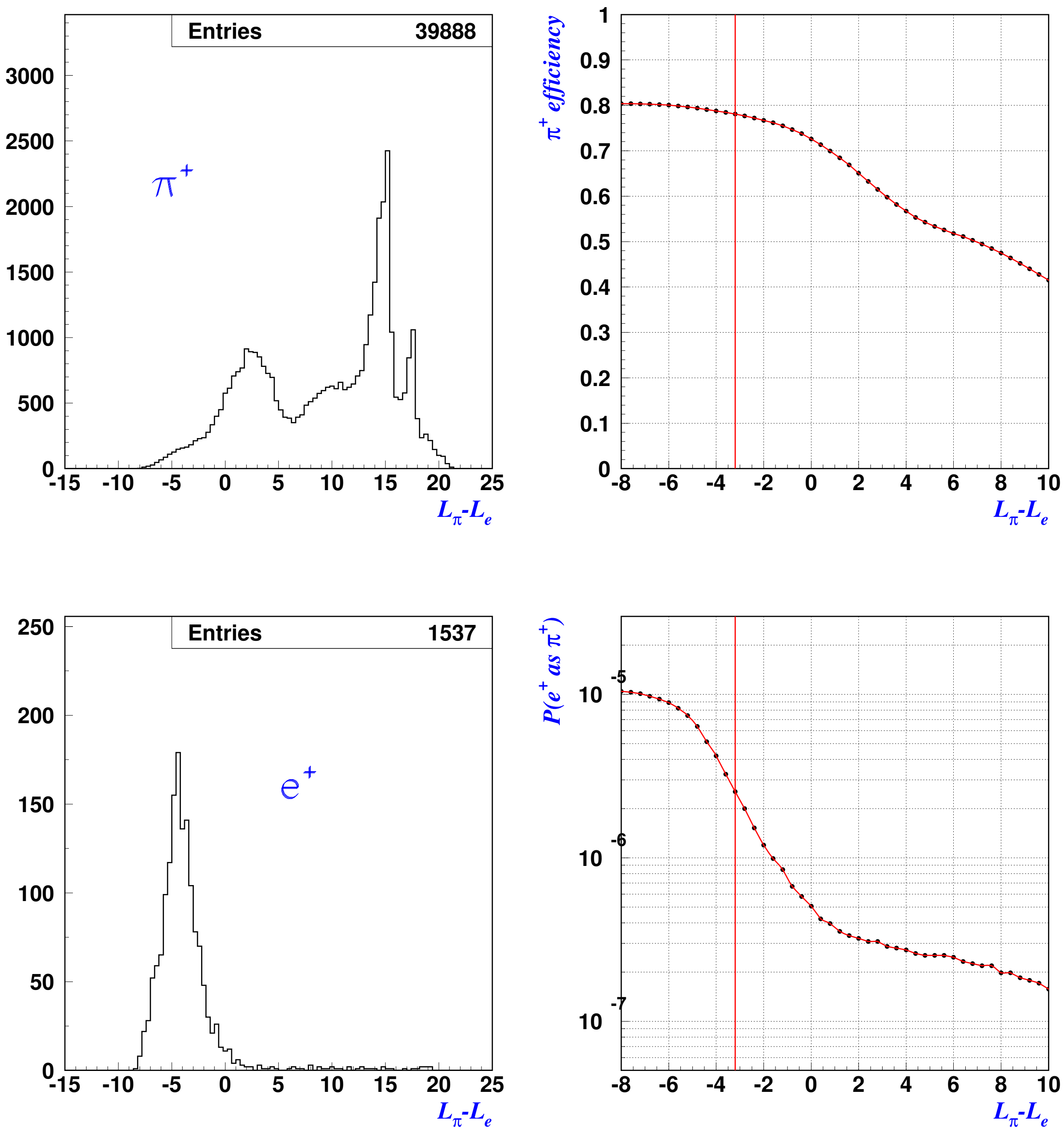


Figure 7.25: Total relative log likelihood $L_\pi - L_e$ distributions for a sample of π^+(top left) and e^+ (bottom left). Efficiency of π^+ identification versus log likelihood cut (top right) and misidentification probability P(e^+ as π^+) versus log likelihood cut (bottom right). Note: The two distinct peaks in the log likelihood distribution for π^+ (top left) are caused by the *ENERGY RATIO* likelihood, where discrimination is made between minimum ionizing pions and ones that produced hadronic showers in the calorimeter.

Chapter 8

Background Estimates

In this section we describe the estimation of the $K_{\pi\mu e}$ background, caused by other K^+ decays. The main kaon decays that would mimic $K_{\pi\mu e}$ were K_{Tau} and K_{Dal} decays (see Section 6.2).

8.1 K_{Tau} decay background

The most probable decay of K^+ with three charged particles in the final state is the $K^+ \to \pi^+\pi^+\pi^-$ decay. In order for a K_{Tau} decay to be identified as a $K_{\pi\mu e}$ it should first satisfy all kinematic requirements, where a large rejection factor comes from the reconstructed $K_{\pi\mu e}$ mass cut (Figure 7.8). In addition two particles π^+ and π^- must be misidentified, or one of the positive pions must undergo a muon decay in flight ($\pi^+ \to \mu^+\nu_\mu$), which in turn will additionally distort the kinematic quality of such event.

In order to estimate the K_{Tau} background level we analyzed the TAU triggered data and applied all final $K_{\pi\mu e}$ kinematic cuts, except for the $M_{\pi\mu e}$ reconstructed mass cut. The PID cuts were applied only to select a π^+ and a μ^+. As a result, 435 events survived that selection , with no events found in the signal region ($|M_{\pi\mu e} - 0.4937| < 0.012$ GeV). As a next step we have chosen to omit S_{norm} and S_{2trk} vertex quality cuts. Without those cuts applied, the number of surviving events increased to 1223 with three of them residing within the signal region of the reconstructed $M_{\pi\mu e}$ mass (Figure 8.1 top right). In order to check, whether efficiency of the vertex quality cuts depends on the $M_{\pi\mu e}$ reconstructed mass we determined vertex quality cuts efficiency for several $M_{\pi\mu e}$ mass bins (Figure 8.1 bottom left). The efficiency was not uniform and decreased, as $M_{\pi\mu e}$ approached the signal region. This indicated that efficiency of the vertex quality cuts is overestimated. Finally, the number of $N_{\pi\pi\mu}$ events in the signal region was extrapolated in the following way:

$$N^{bg}_{\pi\pi\mu} = N_{\pi\pi\mu} \times \epsilon_{\text{vertex}} \times PS^{\text{Hardware}} \times PS^{\text{Software}} \tag{8.1}$$

Where, $N_{\pi\pi\mu}$ is the number of K_{Tau} background events found within the signal region with no vertex quality cuts applied; ϵ_{vertex} is the efficiency of the vertex quality cuts;

PS^{Hardware} is the hardware prescale factor of the TAU triggers; PS^{Software} is the software prescale factor used in the analysis of the TAU triggers.

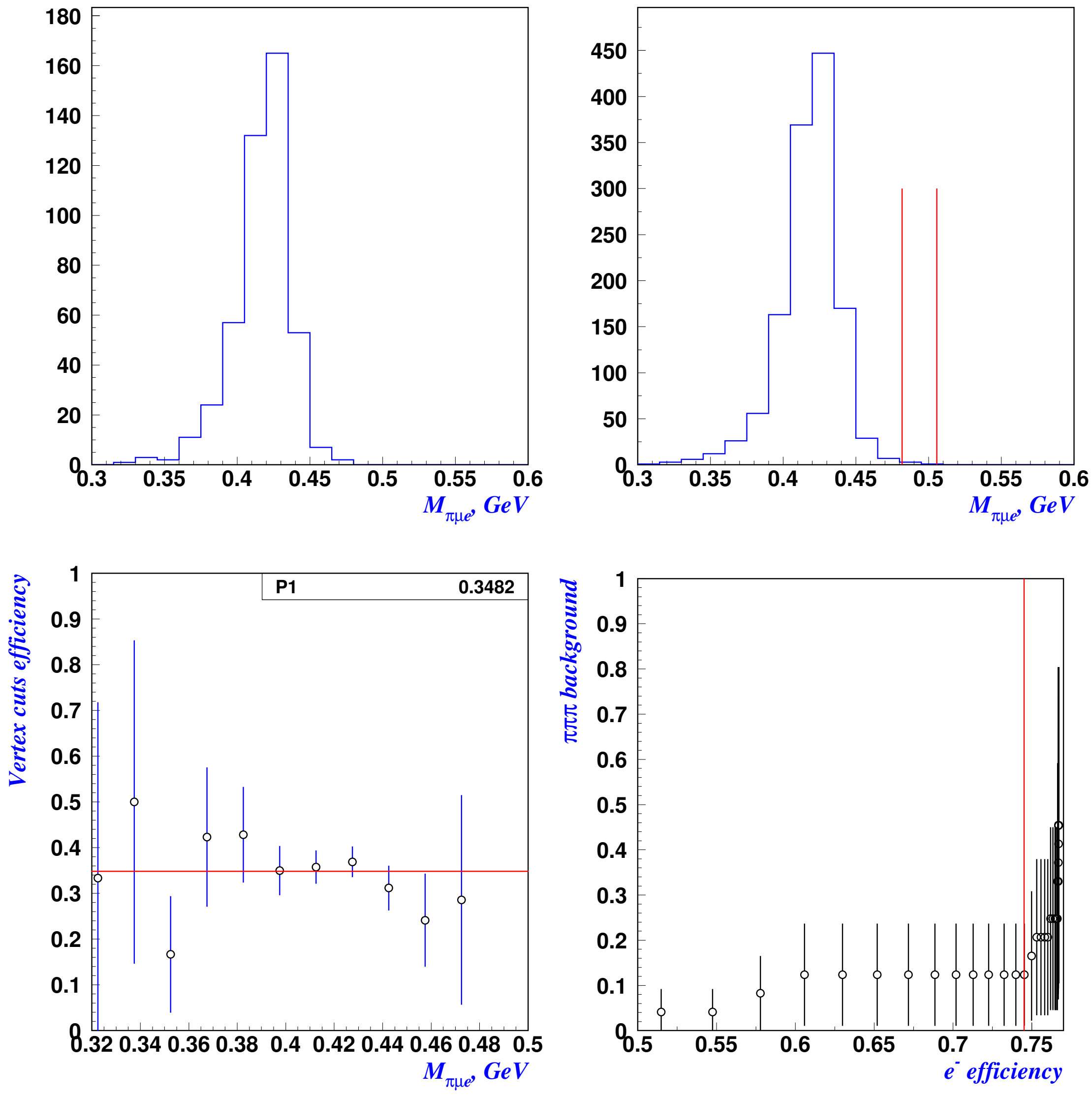


Figure 8.1: Reconstructed $M_{\pi\mu e}$ mass spectrum for K_{Tau} background with vertex quality cut applied (top left) and not applied (top right). Efficiency of the vertex quality cut versus $M_{\pi\mu e}$ reconstructed mass (bottom left). Predicted level of the K_{Tau} background versus the e^- identification efficiency (bottom right).

Putting in all the factors yielded:

$$N^{bg}_{\pi\pi\mu} = 3 \times 0.35 \times 10^4 \times 5 = 5.25 \times 10^4 \tag{8.2}$$

In order to determine the level of K_{Tau} background we multiplied the number of events ($N^{bg}_{\pi\pi\mu}$) by the corresponding misidentification probability $P(\pi^-$ as $e^-)$.

$$N^{BG}_{TAU} = N^{bg}_{\pi\pi\mu} \times P(\pi^- \text{as e}^-) \tag{8.3}$$

Since the misidentification probability depended on the PID likelihood cut, we defined such a cut by examining a cross plot of the level of K_{Tau} background versus the efficiency of e^- identification for different likelihood cuts (Figure 8.1 bottom right). The chosen likelihood cut was $L_{e^-} - L_{\pi^-} > -0.8$, and corresponded to an e^- identification efficiency of 74.5%, and a K_{Tau} background level of 0.1 ± 0.1 events (P(π^- as e^-)=$(2.4 \pm 1.4) \times 10^{-6}$).

The likelihood function proved itself useful since it effectively reduced the level of K_{Tau} background from 0.5 to 0.1 events without loosing practically any identification efficiency (2%).

8.2 K_{Dal} decay background

Another $K_{\pi\mu e}$ background is caused by $K^+ \to \pi^0\pi^+$ and $K^+ \to \pi^0\mu^+\nu$ decays, followed by a Dalitz decay of the π^0 into an e^+e^- pair and a photon. Apart from the e^+e^- invariant mass (see Section 7.3.2) the main kinematic difference between the $K_{\pi\mu e}$ and K_{Dal} decays lies in the incorrectly reconstructed direction of the kaon and its momentum, since the photon was not included in the reconstruction.

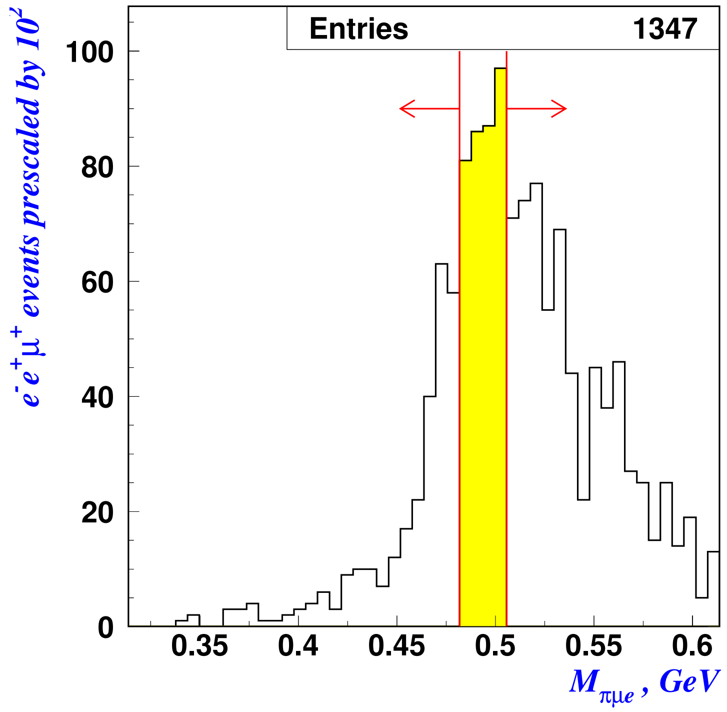


Figure 8.2: Reconstructed $M_{\pi\mu e}$ mass distribution for K_{Dal} background. Lines denote signal region.

We estimated the level of K_{Dal} background by analyzing the EEPS trigger data. All final kinematic $K_{\pi\mu e}$ cuts were applied, except for the reconstructed $M_{\pi\mu e}$ mass cut. Particle identification was required only for the μ^+ and the e^-. This effectively selected $e^-e^+\mu^+$ events. Such events would mimic $K_{\pi\mu e}$ if an e^+ was misidentified as a π^+. We again examined the $M_{\pi\mu e}$ reconstructed mass spectrum. Within the signal region we found 351 $e^- e^+ \mu^+$ events (Figure 8.2). The total number of $e^- e^+ \mu^+$ events in the signal region was extrapolated in the following way:

$$N_{ee\mu}^{bg} = N_{ee\mu} \times PS^{hardware} \times PS^{software} \tag{8.4}$$

Where, $N_{ee\mu}$ is the number of K_{Dal} background events found within the signal region; $PS^{hardware}$ is the hardware prescale factor of the EEPS trigger; $PS^{software}$ is the software prescale factor used in the analysis of the EEPS triggered data. Putting in all the right factors yielded:

$$N_{ee\mu}^{bg} = 351 \times 20 \times 5 = 3.5 \times 10^4 \tag{8.5}$$

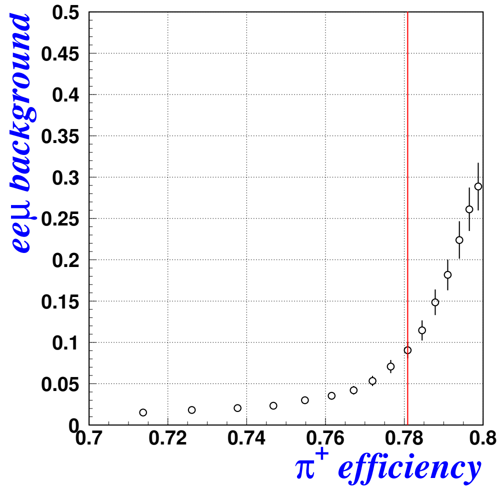


Figure 8.3: Cross plot of estimated K_{Dal} background versus π^+ identification efficiency. The line indicates selected likelihood cut $L_\pi - L_e > -3.2$, that corresponds to 78.1% π^+ identification efficiency and 0.09 K_{Dal} background level.

In order to determine the level of K_{Dal} background we multiplied the number of events ($N_{ee\mu}^{bg}$) by the corresponding misidentification probability $P(e^+ \text{ as } \pi^+)$.

$$N_{Dal}^{BG} = N_{ee\mu}^{bg} \times P(e^+ \text{as } \pi^+) \tag{8.6}$$

The likelihood function was used to reduce the misidentification probability P(e^+ as π^+) to the level ,where K_{Dal} background was at the required level of 0.1 events. The chosen likelihood cut $L_{\pi^+} - L_{e^+} > -3.2$ corresponded to the π^+ identification efficiency of 78.1 %, and the K_{Dal} background level of 0.09 ± 0.01 events ($P(e^+ \text{ as } \pi^+) = (2.56 \pm 0.5) \times 10^{-6}$).

The success of the likelihood method for PID optimization is of course mainly due to the fact, that a set of initial PID cuts was chosen to be quite loose and likelihood functions provided significant improvement by effectively tightening requirements on the variables like E/P, time in the Čerenkov counters, etc.

8.3 Accidental background

We return to the main $K_{\pi\mu e}$ background - accidental events. The preliminary stages of the analysis are now completed and accidental background can be properly studied and estimated.

Accidental backgrounds are hard to isolate and almost impossible to simulate accurately. That is why one should rely as much as possible on what little information can be deduced from the measured data. The most obvious way to select accidental $K^+ \to \pi^+\mu^+e^-$ candidates, is to look at the out-of-time events.

8.3.1 Kinematic correlations for accidental events

Since it was important for the validation of the likelihood analysis and extraction of the P.D.F's for the variables to be used, we started with an examination of the possible correlations between main kinematic variables.

Table 8.1: Conditional efficiency of kinematic cuts applied in pairs and separately for the accidental events

	none	S_{norm}	Tgtlik	S_{2trk}	N_γ	M_{ee}	$M_{\pi\mu e}$	χ^2_{trk}
S_{norm}	0.27 $\pm$ 0.01		0.31 $\pm$ 0.03	0.7 $\pm$ 0.02	0.27 $\pm$ 0.01	0.27 $\pm$ 0.01	0.27 $\pm$ 0.02	0.27 $\pm$ 0.01
Tgtlik	0.06	0.07 $\pm$ 0.01		0.07 $\pm$ 0.01	0.06	0.06	0.06	0.06
S_{2trk}	0.33 $\pm$ 0.01	0.87 $\pm$ 0.03	0.37 $\pm$ 0.03		0.33 $\pm$ 0.01	0.34 $\pm$ 0.01	0.34 $\pm$ 0.03	0.33 $\pm$ 0.01
N_γ	0.86	0.86	0.85	0.86		0.86	0.85	0.86
M_{ee}	0.79 $\pm$ 0.01	0.78 $\pm$ 0.03	0.74 $\pm$ 0.05	0.8 $\pm$ 0.02	0.79 $\pm$ 0.01		0.89 $\pm$ 0.05	0.79 $\pm$ 0.01
$M_{\pi\mu e}$	0.09	0.09 $\pm$ 0.01	0.09 $\pm$ 0.01	0.09 $\pm$ 0.01	0.09	0.10		0.09
χ^2_{trk}	0.92 $\pm$ 0.02	0.92 $\pm$ 0.03	0.93 $\pm$ 0.06	0.92 $\pm$ 0.03	0.92 90.02	0.92 $\pm$ 0.02	0.94 $\pm$ 0.05	

To trace the correlations between the kinematic variables, we selected a sample of accidental events by applying all final PID cuts and requiring $TrkRMS$ to be greater than 10ns. As a next step different kinematic cuts were applied, first individually and then in pairs (S_{norm} cut, S_{norm} cut and $Tgtlik$ cut, S_{norm} cut and S_{2trk} cut, etc.). Such a procedure allowed to check for correlations by examining conditional efficiency of the different cuts.

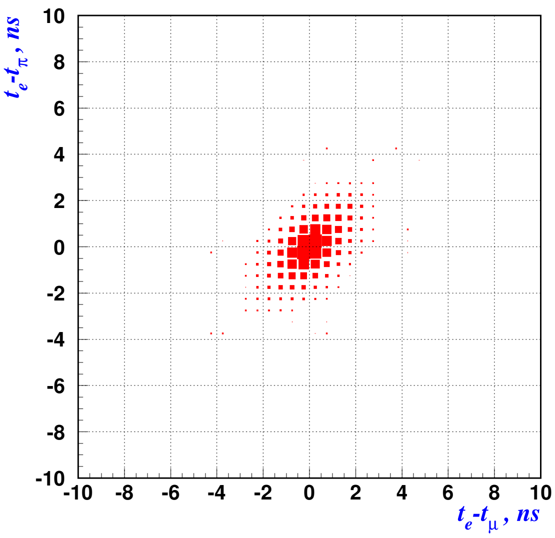
The results are presented in the Table 8.1 and show that obvious correlation was found only between S_{norm} and S_{2trk} cuts. It did not come as a surprise, since both variables are measuring quality of the vertex and should be correlated. There is also chance of correlation between the Tgtlik and the vertex quality variables S_{norm} and S_{2trk}.

It became clear that we could use kinematic variables in our general kinematic likelihood function and extract individual p.d.f's for each variable.

8.3.2 Time structure of accidental events

With a preliminary understanding of correlations between kinematic parameters of accidental events we continued our study with the examination of their time signature. The time registered by the A and Shower counters TDCs defined the time of each track. With three tracks in each event, it was sufficient to examine only two time differences to obtain full and clear picture of possible correlations.

We chose $t_e - t_\pi$ and $t_e - t_\mu$ as variables describing event's timing. The shape of the time distribution for signal events was determined from reconstructed K_{Tau} events (Figure 8.4 left).



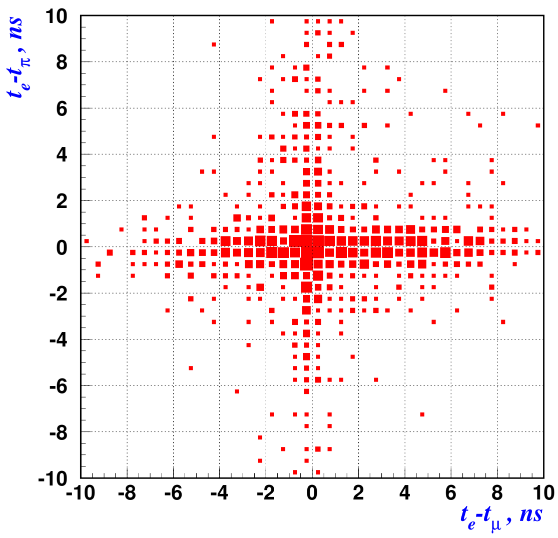


Figure 8.4: Scatter plot of differencies between track times: $t_{\pi^-(1)}$ - $t_{\pi^+(2)}$ vs $t_{\pi^-(1)}$ - $t_{\pi^+(3)}$ for the K_{Tau} events (left) and high momentum accidental events(right).

In order to select accidental events, the kaon momentum was required to be larger then 7 GeV/c and the target likelihood cut was removed along with a number of other cuts to increase the sample size.

Accidental events selected in such a manner showed a distinct *cross* in the scatter plot of our two variables (Figure 8.4 right). This indicated that accidental events had two main components. The vertical part of the cross identifies the π^+ as the additional accidental track paired with the two tracks from another kaon decay. The horizontal part identifies an accidental μ^+ track. Accidental e^- tracks were not observed, since scatter plot did not show a linear correlation between the variables. Finally, triple coincidencies (1+1+1) would appear uniformly scattered and as we expected their contribution is negligible. Using the selection cuts illustrated in Figure 8.5 we determined the ratio between the two components within the signal region approximately as two to three (40% $e\mu$+π and 60% $e\pi$+μ).

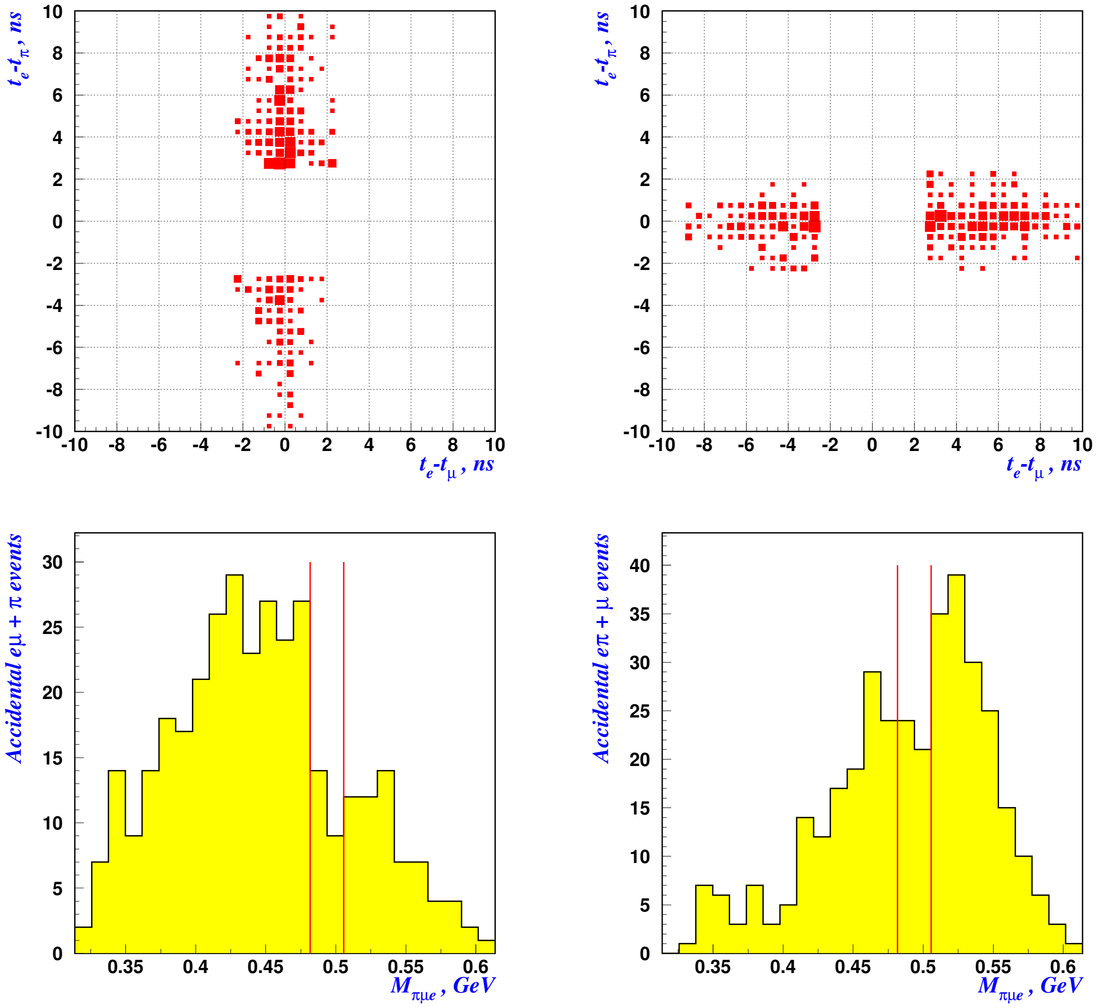


Figure 8.5: Relative time cross plots for accidental events, used to select two components: $e\mu$+π (top left) and $e\pi$+μ (top right). Reconstructed kaon mass $M_{\pi\mu e}$ for two different accidental components (bottom left - $e\mu$+π, bottom right - $e\pi$+μ). Lines on lower plots denote the signal region.

8.3.3 Kinematics of $e\pi + \mu$ accidental events.

Since the two-track part ($e\pi$) of the main accidental component ($e\pi+\mu$) could be originating from a $K^+ \to \pi^0\pi^+$ (with a $\pi^0 \to e^+e^-\gamma$) decay it was useful to understand, whether we could efficiently suppress $K^+ \to \pi^0\pi^+$ decays by using simple kinematic constraints. The $K^+ \to \pi^0\pi^+$ decay is a two body decay and that would lead to a correlation between the pion momentum, and the angle of the pion track at the vertex. The cross plot of the angle versus the momentum for a π^+ was examined for the Monte-Carlo simulated $K^+ \to \pi^0\pi^+$, correlation was indeed observed for the simulated $K^+ \to \pi^0\pi^+$ decays. (Figure 8.6). As it was expected, signal events, being a three body decay, did not show such correlation.

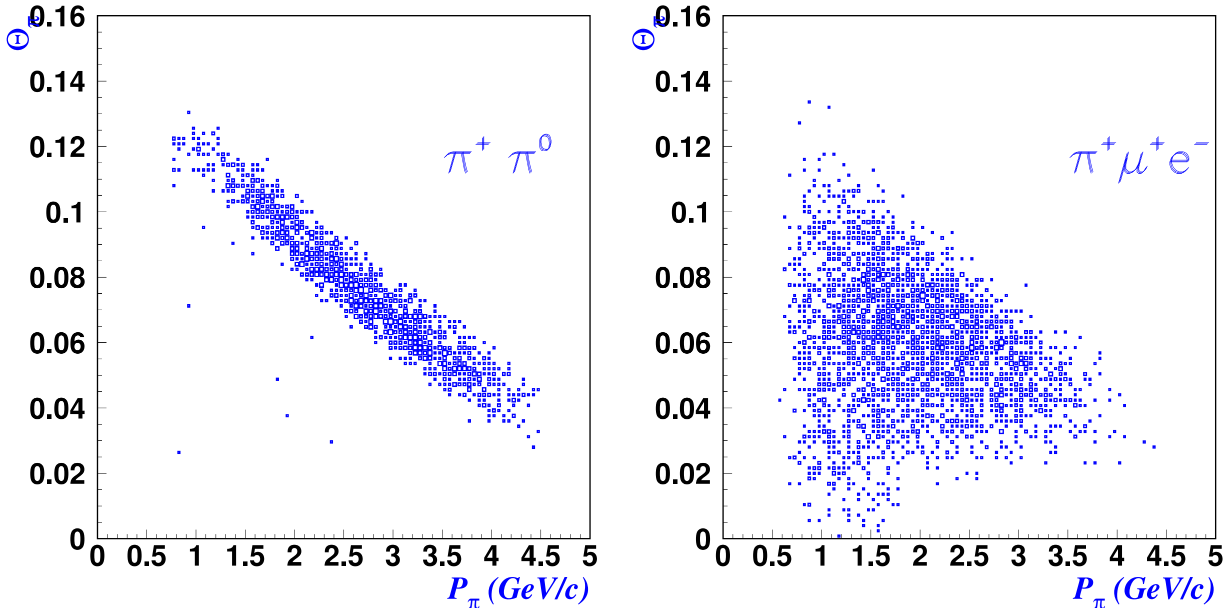


Figure 8.6: Cross plot of π^+ track angle (θ) at the vertex vs π^+ momentum (P_π) for the simulated $K^+ \to \pi^0\pi^+$ (left) and $K^+ \to \pi^+\mu^+e^-$ (right) events.

The effect was best observed when the angle-momentum dependence was projected on the axis orthogonal to the axis of almost linear correlation between the angle and momentum of a π^+ from the $K^+ \to \pi^0\pi^+$ decay (Figure 8.7 top), equivalent to using the variable $z = \theta_\pi/0.14+P_\pi/5$ (P in GeV/c and θ in rad).

With a cut at z=1 we found for out-of-time accidental events an overall 60% rejection, with 80% of the $e\pi + \mu$ and 30% of the $e\mu + \pi$ type accidentals being removed (Figure 8.7 bottom). This confirmed our hypothesis that the majority of the $e\pi + \mu$ accidentals arose from $K^+ \to \pi^0\pi^+$ decays with an unreconstructed e^+ track, paired with an accidental μ^+ track.

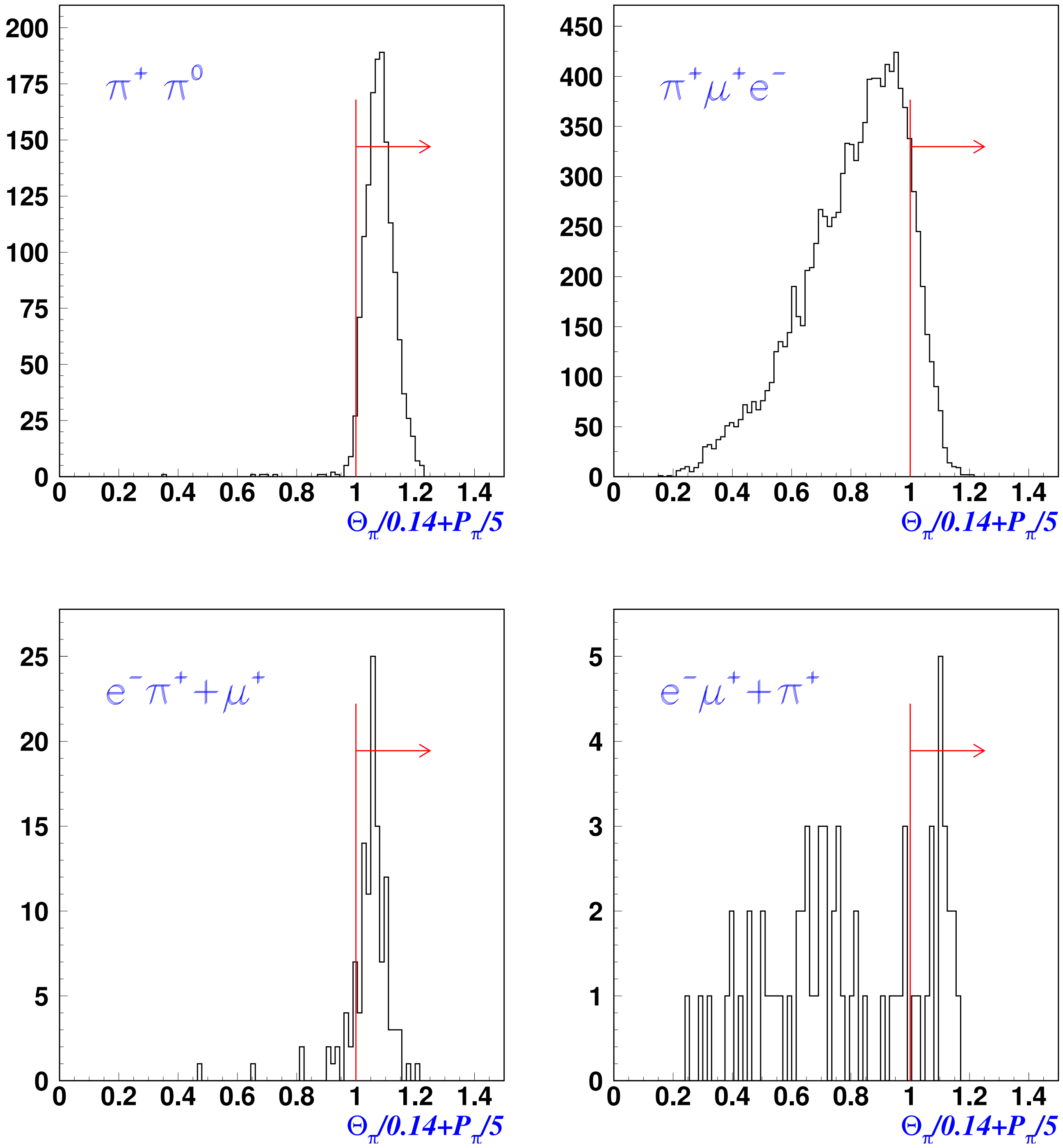


Figure 8.7: Distribution of the variable Z, defined as the sum of π^+ track angle (θ_π) at the vertex divided by 0.14 rad and the π^+ momentum (P_π) divided by 5 GeV/c, for the simulated $K^+ \to \pi^0\pi^+$ (top left), simulated $K^+ \to \pi^+\mu^+e^-$ (top right), $e\pi + \mu$ accidentals (bottom left) and $e\mu + \pi$ accidentals (bottom right).

8.3.4 Designing the new time quality estimator

With the evident time structure of accidental events (Section 8.3.2), it was logical to develop a new time-quality variable. The main consideration was to combine both accidental components ($e\pi + \mu$ and $e\mu + \pi$) into one, and to design a new time variable that would effectively reject both.

Following the logic of initial accidental time structure study we began with two variables:

- $t_a = t_e - t_\pi$
- $t_b = t_e - t_\mu$

The next step was taking the absolute value of both, since the track time difference should be symmetric.

- $t_c = |t_e - t_\pi|$
- $t_d = |t_e - t_\mu|$

And finally we wanted to combine both background components into one, which we would just call accidental background. A symmetry was found between the X and Y axis. And to use it we did the final operation with the track's time information, and produced two variables.

- $Tmin = min(|t_e - t_\pi|,|t_e - t_\mu|)$
- $Tmax = max(|t_e - t_\pi|,|t_e - t_\mu|)$

The new variable $Tmin$ reflects the time resolution, $Tmax$ indicates the maximum difference in the time of an electron track and one of the two remaining tracks. Cuts were made quite loose at the value of 3 ns for both variables (see Figures 8.8,8.9).

Without the effect of the hardware gates, one would expect $Tmax$ to be rising from zero and then being a constant. PID cuts (use of Čerenkov timing) caused $Tmax$ to gradually fall, and finally fall sharply due to the onset of the TDC gates. The variable $Tmax$ was later used in the general likelihood.

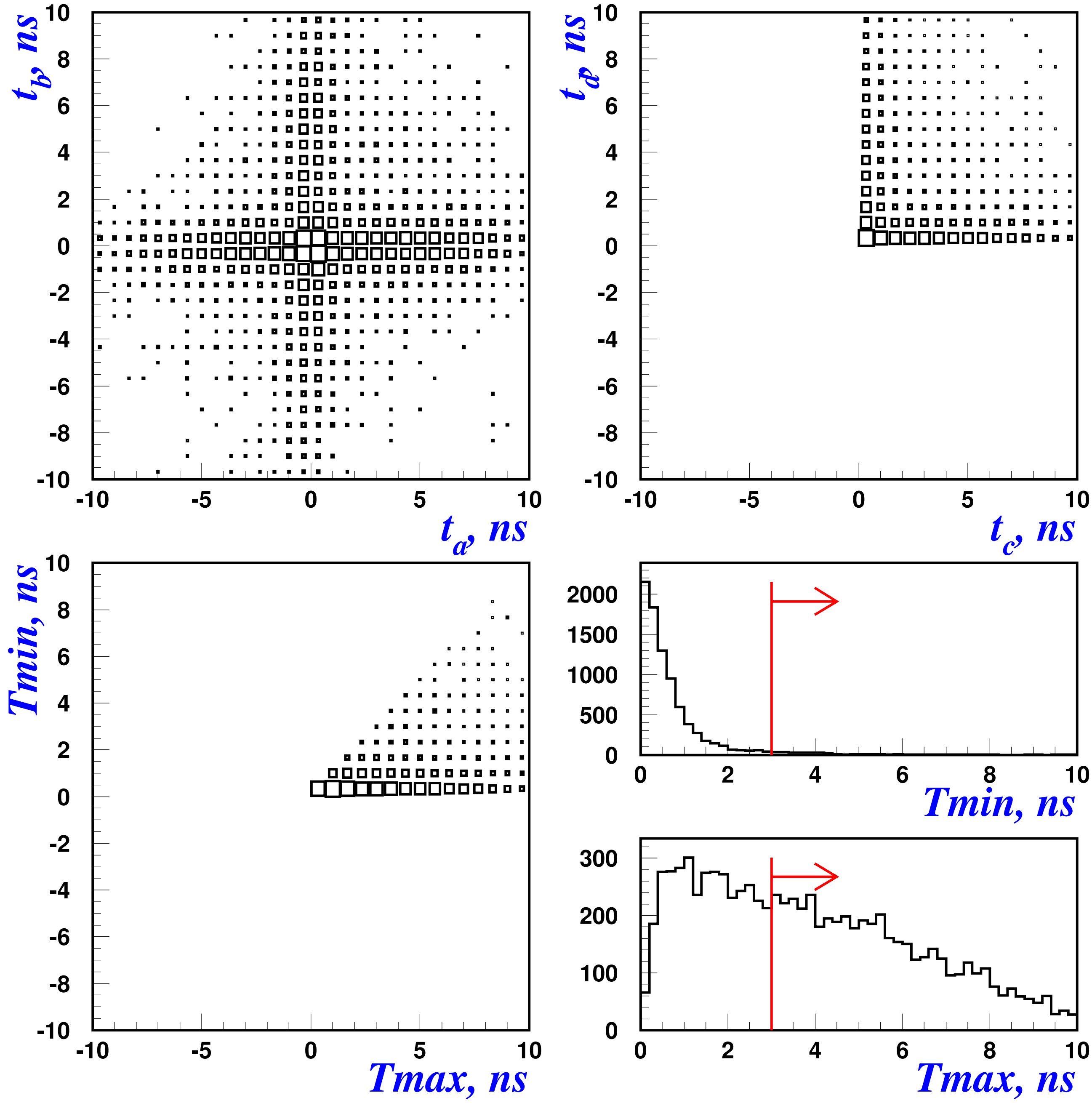


Figure 8.8: Time correlations plots for accidental events (P_{Kaon} >7 GeV/c). Lines denote final cuts. See the text for explanation of the variables.

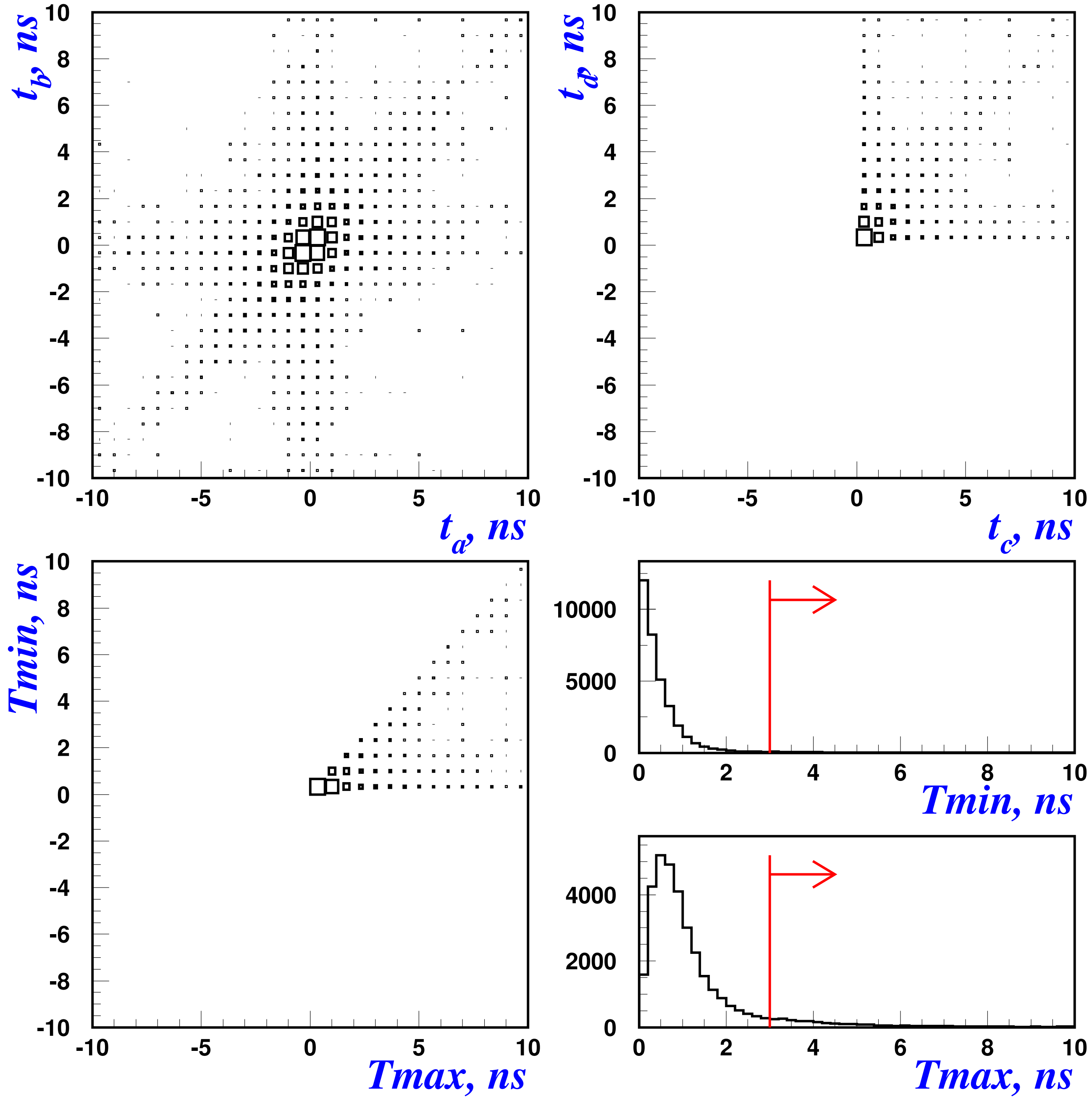


Figure 8.9: Time correlations plots for reconstructed $K^+ \to \pi^+\pi^+\pi^-$ events. Lines denote final cuts. See the text for explanation of the variables.

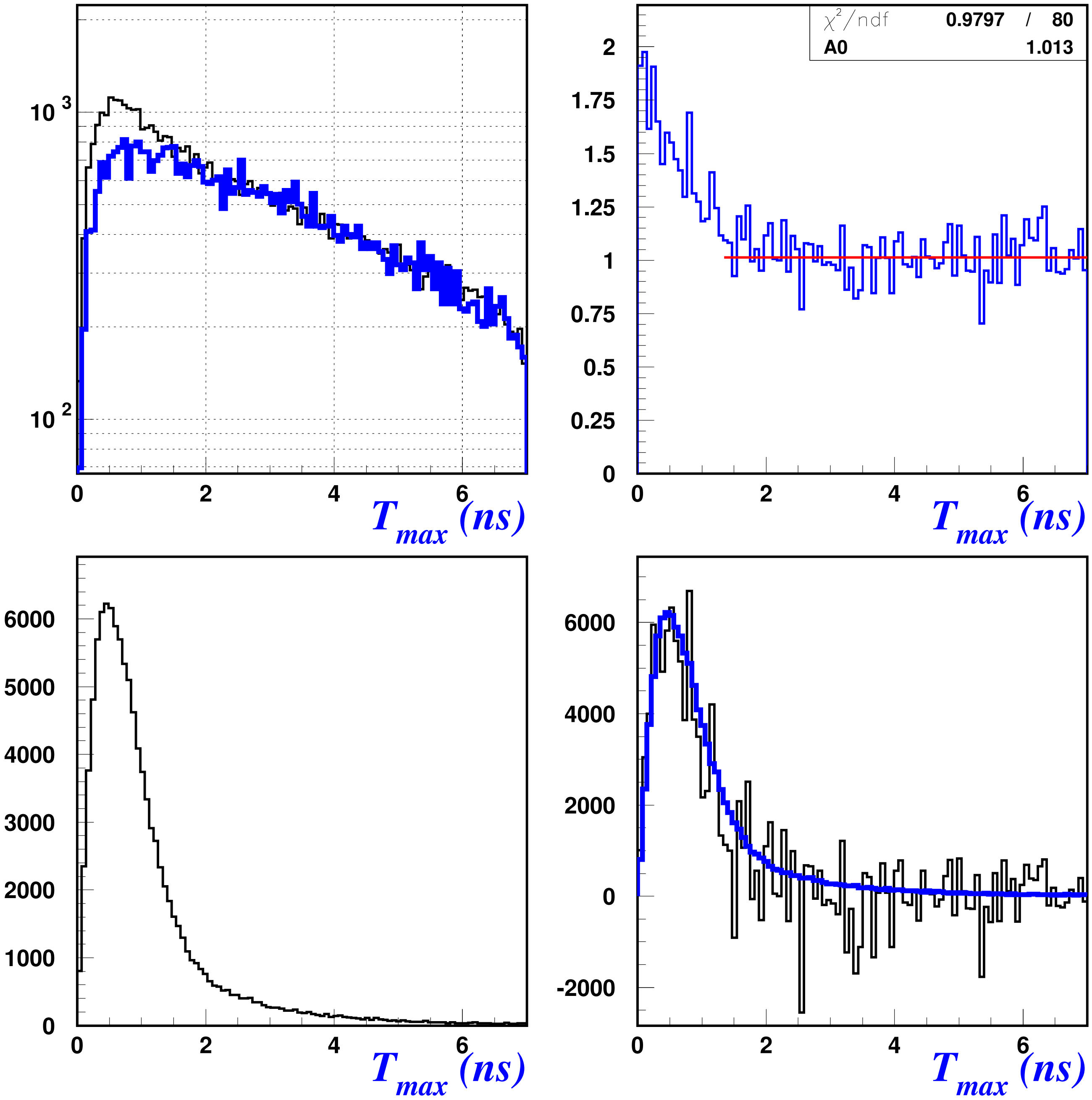


Figure 8.10: Comparison of the T_{max} distribution for high momentum accidentals and $K_{\pi\mu e}$ candidates. Top left: $K_{\pi\mu e}$ candidates(thin line) and high momentum events (bold line). Top right - ratio of the two types of events Bottom left - T_{max} distribution for good $K^+ \to \pi^+\pi^+\pi^-$ events. Bottom right - T_{max} distribution of the difference between time distribution of $K^+ \to \pi^+\mu^+e^-$ candidates and high momentum accidentals (thin line) and T_{max} distribution of good $K^+ \to \pi^+\pi^+\pi^-$ events (bold line).

8.4 Estimating accidental background

In order to estimate accidental background we used out-of-time $K_{\pi\mu e}$ candidates from the region of large $Tmax$. In order to be able to extrapolate into the signal region of in-time events ($Tmax$ <3ns) we extracted the timing distribution of accidental events by selecting high momentum $K_{\pi\mu e}$ candidates. The number of accidental background events was then approximated by:

$$N^{bg}_{ACC} = N^{out-of-time}_{acc} \times \frac{N_{acc}(Tmax < 3)}{N_{acc}(t_1 < Tmax < t_2)} \tag{8.7}$$

Where $N^{out-of-time}_{acc}$ is the number of $K_{\pi\mu e}$ candidates in the out-of-time region $((t_1 < Tmax < t_2))$, $N_{acc}(Tmax < 3)$ is the number of accidental events with $Tmax < 3$ns, and $N_{acc}(t_1 < Tmax < t_2)$ is the number of accidental events within the out-of-time region.

We will refer to the ratio between the $N_{acc}(Tmax < 3)$ and $N_{acc}(t_1 < Tmax < t_2)$ as the projection factor R:

$$R = \frac{N_{acc}(Tmax < 3)}{N_{acc}(t_1 < Tmax < t_2)} \tag{8.8}$$

However, two question arose in regard to this simple method. The first concerned the accuracy of the projection factor R, and whether the time distribution of high momentum accidentals is identical to that of those within the normal kaon momentum range. The second concerned the range of $Tmax$ to be used and how the choice would affect the background estimate. Extending the range too far would cause problems, because of lack of statistics and unclear influence of the hardware time gates. Going too low could mean that the sample would be contaminated by real kaon decays.

In order to address our first concern, we compared $Tmax$ distribution for the high momentum accidentals to that of the $K_{\pi\mu e}$ candidates outside the signal region in $M_{\pi\mu e}$ mass. The time quality cut was not applied. If high momentum events have the same time distribution as accidentals within the real kaon momentum range, then tails of the time distributions would match. If, however, the time resolution would be significantly affected by the particle momentum , a difference would be noticeable. As it can be seen on Figure 8.10 such a difference was not observed, and tails of time distribution for $K_{\pi\mu e}$ candidates and high momentum accidentals show reasonable agreement. In addition, we subtracted the time distribution of accidental events from that of the $K_{\pi\mu e}$ candidates. The observed difference was due to real kaon decays, like $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi^0\pi^+$, which, due to the particle misidentification, mimicked $K_{\pi\mu e}$ events. The observed difference satisfactorily matched the timing distribution of the reconstructed K_{Tau} events, which gave additional confidence to our choice of using high momentum accidentals to extract the timing distribution of accidental background events.

8.4.1 Kaon momentum dependence

As a next step we tested the stability of the projection factor R. We varied the P_{kaon} cut in order to check how it affects the projection factor R.

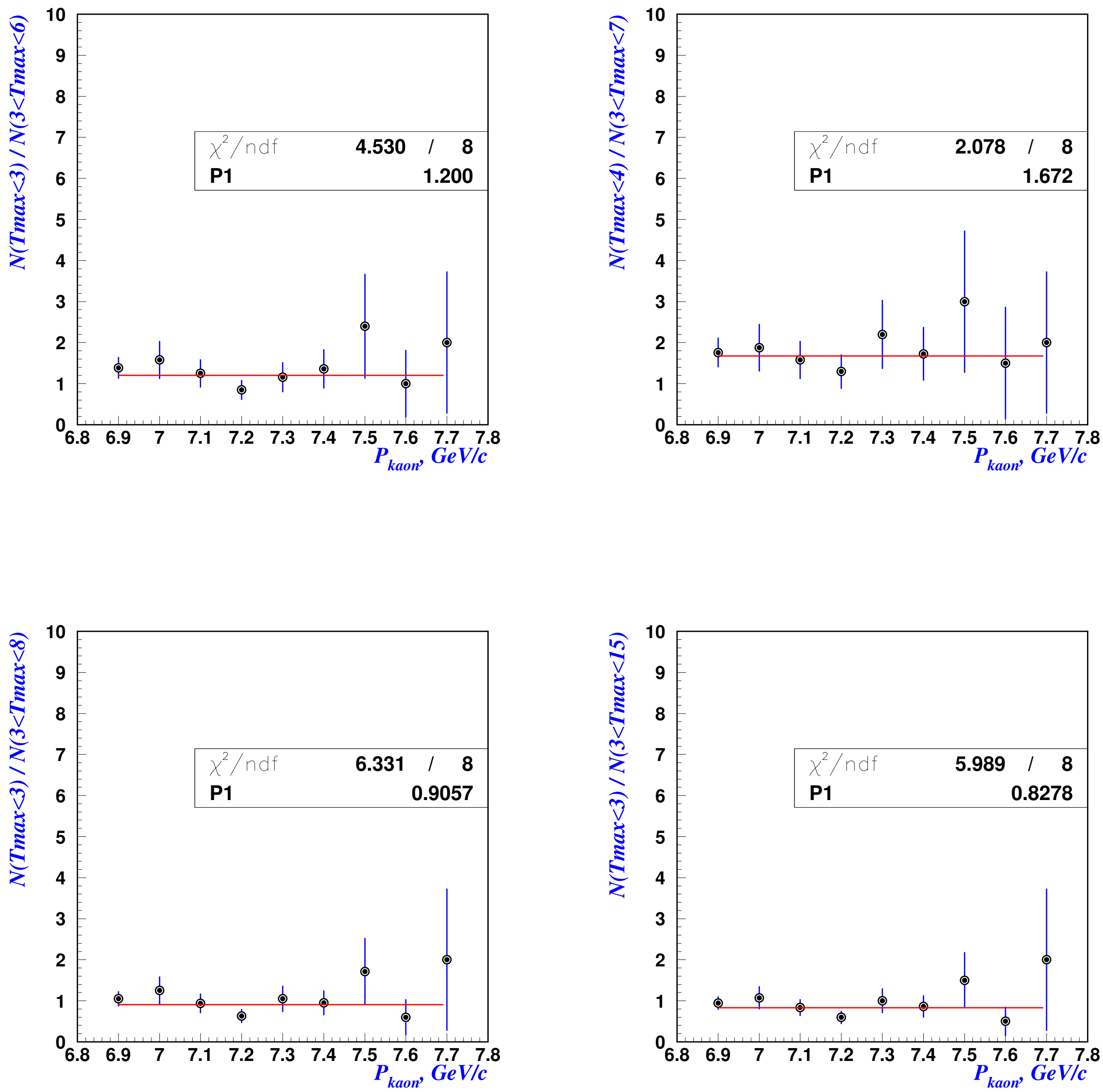


Figure 8.11: Projection factor from the out-of-time to the in-time (signal region) for the high-momentum accidental events versus the kaon momentum for different out-of-time intervals. Top left: N($T_{max} < 3$) : N($6 > T_{max} > 3$). Top right: N($T_{max} < 3$) : N($7 > T_{max} > 4$). Bottom left: N($T_{max} < 3$) : N($8 > T_{max} > 3$). Bottom right: N($T_{max} < 3$) : N($15 > T_{max} > 3$).

To select high momentum accidental events, all $K_{\pi\mu e}$ selection cuts, except Tgtlik, $M_{\pi\mu e}$ and $Tmax$ were applied. In addition P_{kaon} was required to be in one of the predefined intervals (6.9..7 GeV/c, 7...7.1 GeV/c, *etc.*). For each bin, a ratio of events with $Tmax$ <3 ns to events with $Tmax$ in the range of 3...6 ns, 4...7 ns, 3...8, and 3...15 ns was calculated and plotted versus the mean P_{kaon} value of the momentum bin. The results are shown in Figure 8.11. Within the statistical uncertainties we did not observe any dependence of the projection factor R on the momentum of the kaon (P_{kaon}).

8.4.2 Time window dependence

In a second test we varied the range of $Tmax$ and compared the following windows:

- 3 to 6 ns
- 4 to 7 ns
- 3 to 8 ns
- 3 to 15 ns

In order to reduce the statistical uncertainty. we omitted the vertex quality (S_{norm}, S_{2trk}), and reconstructed mass $M_{\pi\mu e}$ cuts.

Table 8.2 shows the number of observed accidental events for each range of the $Tmax$, as well as the number of events that in addition passed the vertex quality cuts (Scut), the reconstructed mass cut (Mcut), and both cuts simultaneously (S,M cuts). In addition, the expected number is presented, which is calculated by using an independently determined efficiency of those cuts.

Table 8.2: N - Number of $K_{\pi\mu e}$ candidates within the preselected $Tmax$ range of 3..6, 4..7, 3..8 or 3..15 ns. Scut - number of events after the vertex quality cut is applied. Mcut - number of events after the reconstructed $M_{\pi\mu e}$ mass cut is applied. S,M cuts - number of events, with the vertex and mass cuts applied. Second line in each row, labeled *expected* shows the number of expected events, which is calculated from the number of events N, by applying efficiency of the cuts.

$Tmax$	N	Scut	Mcut	S,M cuts
3..6	190±14	51±7	23±5	6± 2.4
expected		47.7±7.2.	25.8±4.8	6.5± 0.7
4..7	169±13	44±6.6	21±4.6	4±2
expected		42.4 ±6.5	23±4.3	5.8 ±0.7
3..8	267±16.3	66±8	35±5.9	8±2.8
expected		67±9.7	36.3±6.6	9.1 ±0.96
3..15	304±17.4	73±8.5	41±6.4	9±3
expected		76.3±11	41.3±7.4	10.4 ± 0.8

Table 8.3 shows the projection factor R and the estimated number of background events. As one can see, the estimate was stable for all $Tmax$ intervals.

We conclude that the expected level of accidental background was 8.2 ± 1.9 events.

Table 8.3: Table shows $Tmax$ interval in (ns) used to select accidental events, projection factor R (ratio of high momentum accidental events with $T_{max} < 3$ ns and $Tmax$ within the corresponding interval (3..6, 4..7,3..8. or 3..15)), and the number of expected background events bg.

$Tmax$	R	bg
3..6	1.2±0.26	7.8±1.9
4..7	1.67±0.36	9.6±2.4
3..8	**0.91±0.19**	**8.2 ± 1.9**
3..15	0.83±0.16	8.6 ± 1.9

Chapter 9

Signal Search

9.1 Blind analysis as a consistency check

We tested our estimates of the background processes by performing a blind analysis. The final $K_{\pi\mu e}$ selection criteria included the final kinematic (Section 7.3.2) and PID (Setions 7.4.6, 8.1, 8.2) cuts. In addition, the cut on the TrkRMS was substituited by cuts on $Tmax < 3$ and $Tmin < 3$ (Section 8.3.4). The $z < 1$ cut was added to suppress accidental background (Section 8.3.3). We excluded the signal region by reversing the cut on the reconstructed mass $M_{\pi\mu e}$:

$$|M_{\pi\mu e} - 0.4937| > 0.012 \text{ GeV} \tag{9.1}$$

The distribution for the observed $K_{\pi\mu e}$ candidates shown in Figure 9.1 exhibits a peak at lower mass which is caused by the K_{Tau} background.

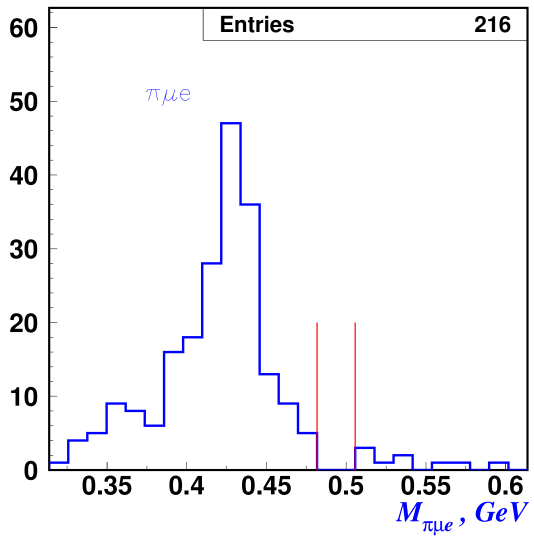


Figure 9.1: The reconstructed $M_{\pi\mu e}$ mass distribution for $K^+ \to \pi^+\mu^+e^-$ candidates, with the $M_{\pi\mu e}$ cut reversed.

Following the same algorithm as that of the background estimates we generated the $M_{\pi\mu e}$ distributions for the main background processes using measured data:

- K_{Tau} background: observed $\pi^-\pi^+\mu^+$ events scaled by the misidentification probability $P(\pi^-$ as $e^-)$ (Figure 9.3 top left);
- K_{Dal} background: observed $e^-e^+\mu^+$ events scaled by the misidentification probability $P(e^+$ as $\pi^+)$ (Figure 9.3 top right);
- Accidental background: out of time accidental events, scaled by the projection factor R (ratio of in-time to out-of-time accidental events)(Figure 9.3 bottom).

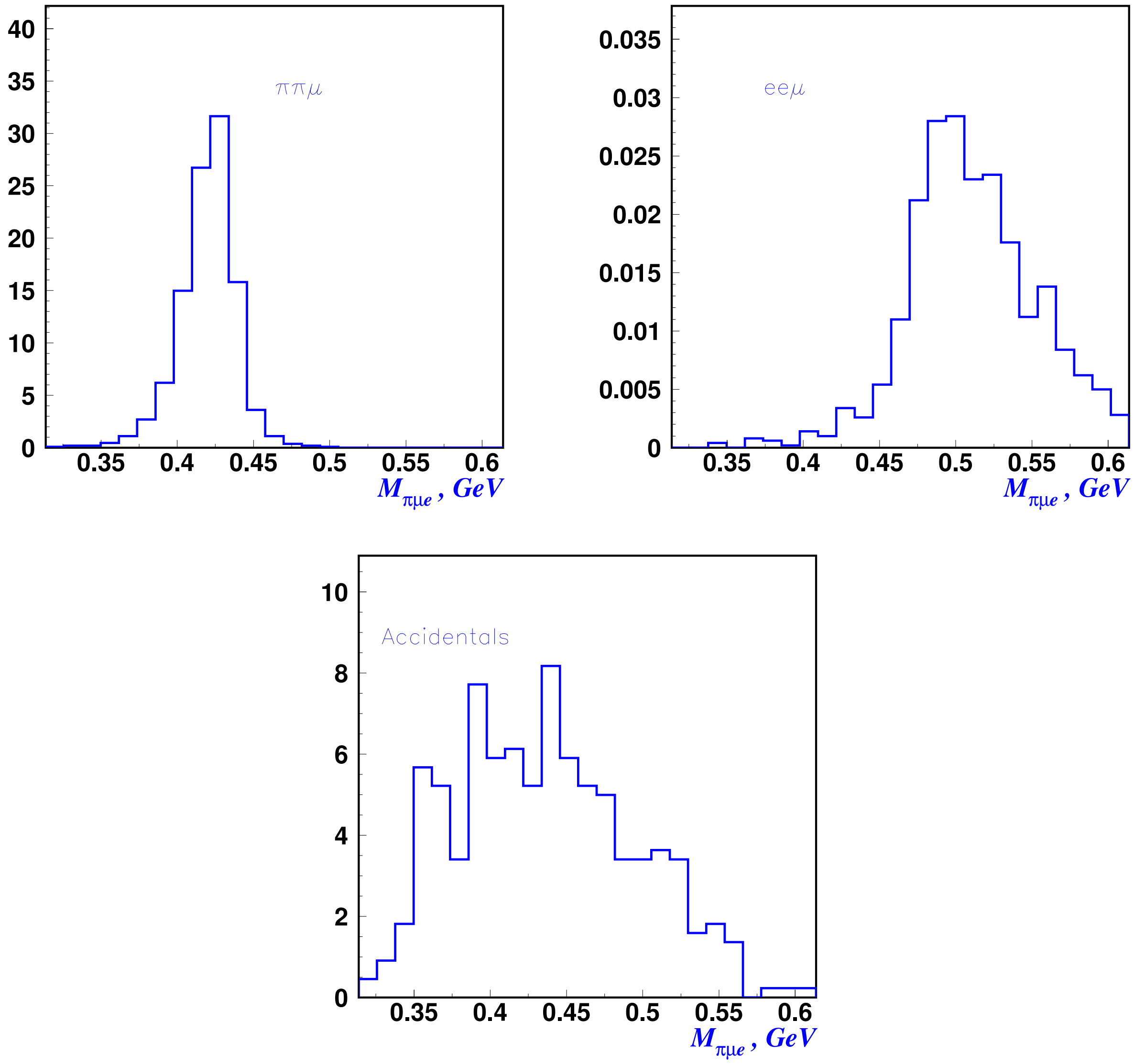


Figure 9.2: The reconstructed $M_{\pi\mu e}$ mass distribution for the estimated K_{Tau} background (top left), estimated K_{Dal} background (top right) and accidental background (bottom).

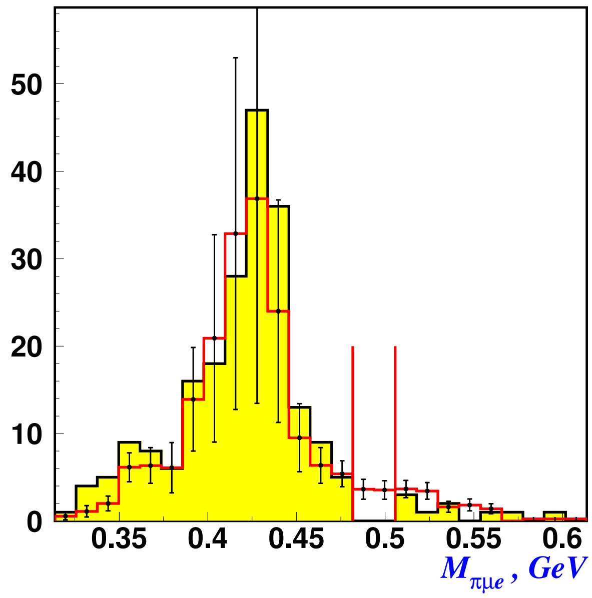


Figure 9.3: The reconstructed $M_{\pi\mu e}$ mass distribution for $K^+ \to \pi^+\mu^+e^-$ candidates (histogram) with $M_{\pi\mu e}$ cut reversed and sum of estimated backgrounds (markers with error bars). Lines denote the $M_{\pi\mu e}$ mass signal region.

Finally, the $M_{\pi\mu e}$ mass distribution of the total predicted background was compared with the distribution of the observed $K_{\pi\mu e}$ candidate events (Figure 9.3). The satisfactory agreement between the predicted and observed events served as an additional consistency check for our estimates of the background and misidentification probabilities ($P(e^+$ as $\pi^+)$ and $P(\pi^-$ as $e^-)$).

9.2 General $K^+ \to \pi^+\mu^+e^-$ likelihood

9.2.1 Overview

In order to construct the general likelihood, which was crucial for rejecting accidental background events, several kinematic and timing variables were used. If the variables are largely uncorrelated the likelihood can be factorized, which eases the creation of the probability density functions.

The following six variables, which have been defined and discussed before were included in the general likelihood:

- Target Likelihood - $Tgtlik$
- Vertex quality - S_{norm}
- Reconstructed $M_{\pi\mu e}$ mass
- sum of the track's χ^2

- $Tmax$ - maximum time difference between π^+ or μ^+ track time and time formed by the other two tracks ($e\mu$ or $e\pi$)
- $Minex\ gd\ cltm$ - minimum time difference between the time of the extra clump in the calorimeter and time of the track pair, which is closest in time

The general likelihood for $K_{\pi\mu e}$ and accidentals was then obtained as:

$$\mathcal{L}_{\pi\mu e}(x_1...x_6) = \sum_{i=1}^{6} \log P_i^{\pi\mu e}(x_i) \tag{9.2}$$

$$\mathcal{L}_{Acc}(x_1...x_6) = \sum_{i=1}^{6} \log P_i^{Acc}(x_i) \tag{9.3}$$

where $x_i = (Tgtlik,\ S_{norm},\ M_{\pi\mu e},\ \chi^2_{Track},\ Tmax,\ Minex\ gd\ cltm)$ and $P_i^{\pi\mu e}$, P_i^{Acc} are probability density functions for signal and background, respectively.

The total relative likelihood $\mathcal{L}$ for $K_{\pi\mu e}$ was defined as the difference between $\mathcal{L}_{\pi\mu e}(x_1...x_6)$ and $\mathcal{L}_{Acc}(x_1...x_6)$:

$$\mathcal{L} = \mathcal{L}_{\pi\mu e}(x_1...x_6) - \mathcal{L}_{Acc}(x_1...x_6) \tag{9.4}$$

9.2.2 Extracting the probability density functions

Since statistics was limited, we had to relax some of the cuts. Table 9.1 lists which cuts were applied (Y), reversed (U), or omitted (N) in order to extract a p.d.f. for a corresponding variable.

The distributions for each variable along with a smoothed profile, used as the p.d.f. are presented in the Figure 9.4.

Table 9.1: Application of kinematic cuts for selecting accidental samples to extract probability density functions for individual kinematic variables. Y - cut is applied. N - cut is not applied. U - cut is reversed.

variable	S_{norm}	Tgtlik	χ^2_{trk}	Timing	$M_{\pi\mu e}$	P_{kaon}
S_{norm}	Y	Y	N	U	N	N
Tgtlik	Y	Y	N	U	N	N
χ^2_{trk}	N	N	Y	U	Y	N
Timing	N	N	Y	Y	N	U
$M_{\pi\mu e}$	N	Y	N	U	Y	N

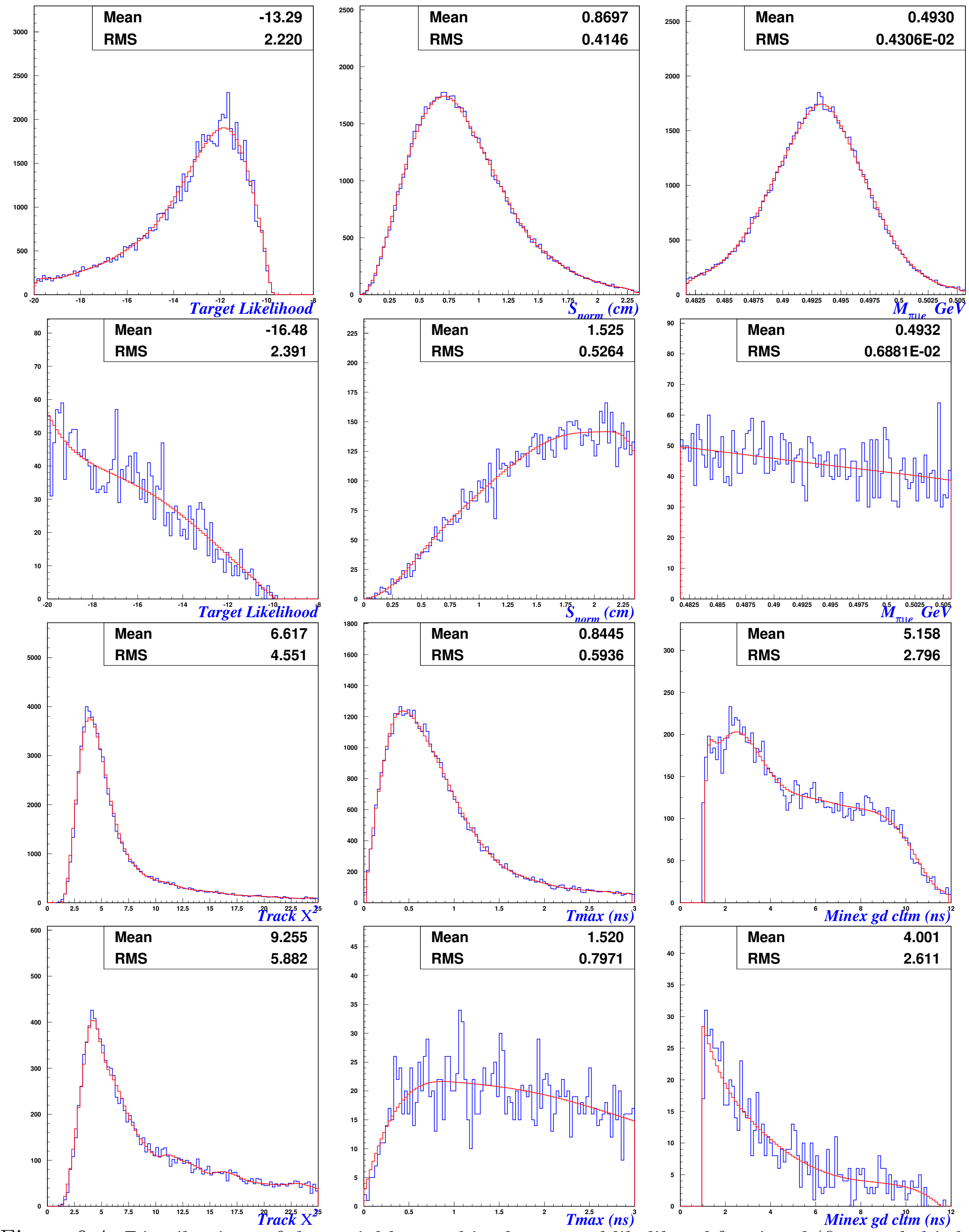


Figure 9.4: Distributions of the variables used in the general likelihood for signal (first and third row) and accidental background (second and forth row). The histogram represents measured data, and the smoothed curve was used as a p.d.f.

9.3 Likelihood templates for signal and background

In order to use the general likelihood one must obtain templates for both signal and background hypothesis.

In order to construct the signal likelihood template one could use MC simulated events, and later add timing information, derived from the data. However we decided to use the real data, namely reconstructed $K_{\pi\pi\pi}$ events. The $K_{\pi\pi\pi}$ kaon decay is very similar to the $K_{\pi\mu e}$, especially in the variables used in the likelihood function. The only difference is in the kaon reconstructed mass distribution. The mass resolution of $K_{\pi\pi\pi}$ events had to be changed by a factor of 1.725 in order to match the mass resolution of $K_{\pi\mu e}$ events, determined from the Monte-Carlo simulations.

In order to extract the $K_{\pi\mu e}$ likelihood template by using $K_{\pi\pi\pi}$ events the following steps were taken. A sample of good $K_{\pi\pi\pi}$ events was selected from the data by applying all kinematic cuts, designed to select $K_{\pi\mu e}$ events, with the exception of the $M_{\pi\mu e}$ mass cut. For $K_{\pi\pi\pi}$ events selected this way, the reconstructed $M_{\pi\pi\pi}$ mass was transformed into:

$$M_{\pi\mu e}^{NEW} = (M_{\pi\pi\pi} - 0.4937) \times 1.725 + 0.4937 (GeV) \tag{9.5}$$

The new $M_{\pi\mu e}^{NEW}$ mass was used as $M_{\pi\mu e}$ mass in the general $K_{\pi\mu e}$ likelihood. The obtained likelihood distribution can be seen in Figure 9.5.

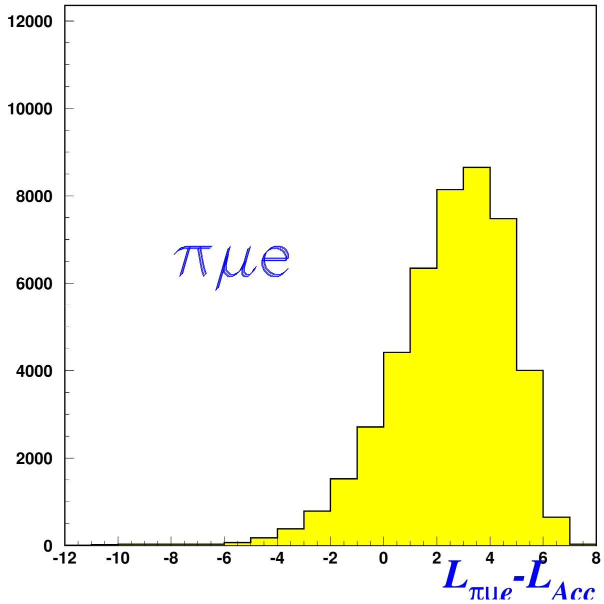


Figure 9.5: General likelihood distribution for signal hypothesis. Sample of reconstructed $K^+ \to \pi^+\pi^+\pi^-$ events with a modified kaon reconstructed mass [see Eq. (9.5)] was used.

A more challenging task was creating the likelihood template for accidental background. Our choice of selecting accidentals was quite limited. We used a sample of out-of-time (T_{max} >3 ns) $K_{\pi\mu e}$ candidate events, which passed all the other final selection cuts. Since all the six variables used in the likelihood had to be within the final cut values, the timing response was generated from a library of the measured timing response of the high momentum accidental events. The likelihood distribution, derived from the sample of out-of-time accidental events can be seen in Figure. 9.6.

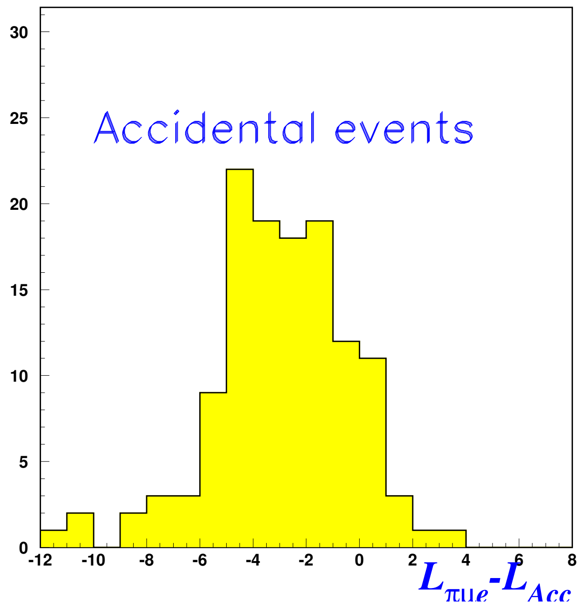


Figure 9.6: General likelihood distribution for accidental events. Sample of out-of-time accidentals was used.

When we constructed the general likelihood, using one dimensional distributions (p.d.f's) we assumed that all variables are independent. That assumption was based on our preliminary study of the accidental background (see section 8.3.1). In order to independently test that assumption we simulated response of all the six variables, used in the general likelihood, directly from their p.d.f.'s. Then we calculated the general likelihood. Simulated distributions compared favourably with the ones derived from the data (see Figure 9.7).

Finally, as a robustness test, we extracted the likelihood distribution of the likelihood events using an alternate method. Instead of using the out-of-time accidentals we decided to suppress K_{Tau} and K_{Dal} backgrounds to a negligible level and use the in-time $K_{\pi\mu e}$ candidate events to create the accidental likelihood template. Of course, the reconstructed $M_{\pi\mu e}$ mass cut was reversed, to secure the signal region. Based on the present upper limit for $K_{\pi\mu e}$ branching ratio[2] it was safe to assume that no signal events would be found outside the $M_{\pi\mu e}$ mass signal region. To suppress the K_{Tau} background we imposed the following cut: $M_{\pi\pi\pi} > 0.54$ GeV. The general likelihood was using reconstructed mass $M_{\pi\mu e}$, and thus we had to substitute $M_{\pi\mu e}$ values from a template of $M_{\pi\mu e}$ mass response of the out-of-time accidental events. Thus, our sample of accidental events was fully within the range of the final $K_{\pi\mu e}$ cuts. However, only 62 events survived, their likelihood distribution can be seen below (Figure 9.8). It also compared favourably to the one, simulated from individual PDFs.

After we extracted the templates for the signal and the background hypothesis we were ready to test the effectiveness of the likelihood. We will discuss the method we used to determine the number of signal events at a 90 % CL in the next chapter.

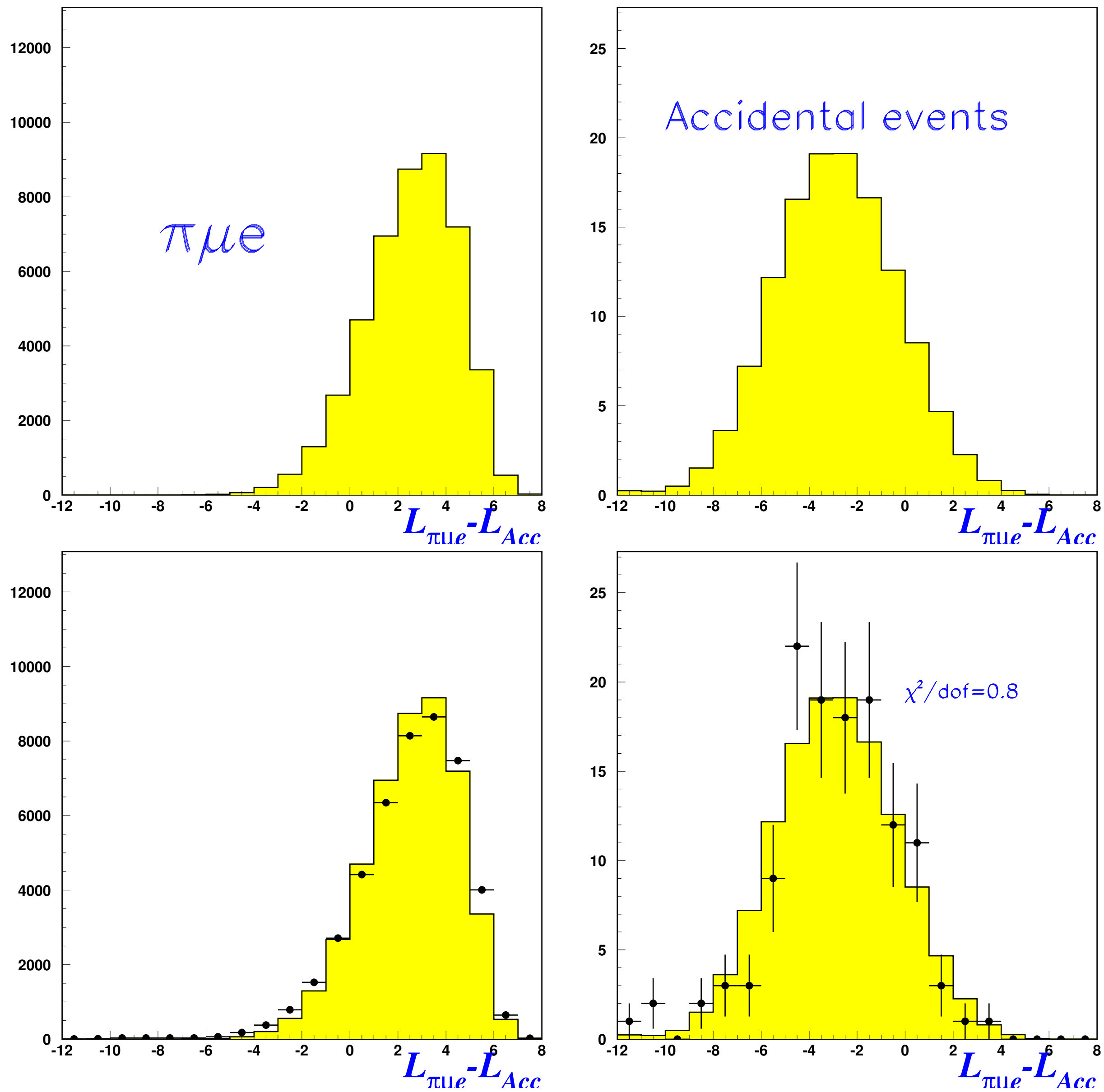


Figure 9.7: General likelihood distribution for simulated signal (top left) and accidental background (top right), and comparison with the measured data (down). Histogram represents simulated events and markers with error bars are measured data.

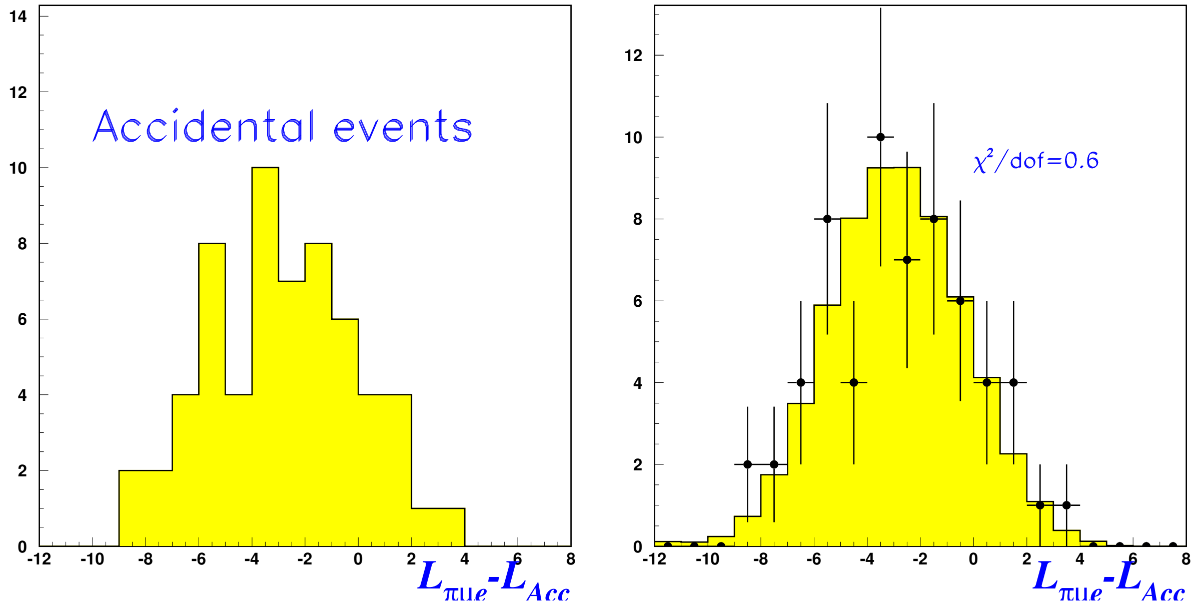


Figure 9.8: General likelihood distribution for accidental events (left) and and it's comparison with the simulation (right). Histogram represents simulated events and markers with error bars are measured data. A sample of $K_{\pi\mu e}$ candidates outside the $M_{\pi\mu e}$ mass signal region was used (see text for details).

Chapter 10

Determining the Number of $K^+ \to \pi^+\mu^+e^-$ Events

An important part of the analysis was determining the number of signal events at a given confidence level. It was especially complicated in our case where we were expecting a significant number of accidental background events (8.2 ± 1.9). The method used is discussed below, it includes those additional tests that were aimed at estimating the power of the likelihood method in suppressing the background. The systematic effects of the uncertainties in the predicted background on a final result are also discussed.

10.1 The Bayesian approach

In order to determine the number of signal events we used the Bayesian approach [24].

We defined the likelihood function for Poisson-distributed data as follows:

$$L_p(\vec{m};\vec{n}) = \prod_{i=1}^{k} \frac{m_i^{n_i} e^{-m_i}}{n_i!} \tag{10.1}$$

where

$$\begin{aligned}
k &= \text{arbitrary chosen number of bins} \\
n_i &= \text{the number of observed events in the } i_{th} \text{ bin} \\
\vec{n} &= (n_1, n_2, ..., n_k) \\
N &= \text{total number of observed events } (= \sum_{i=1}^{k} n_i) \\
m_i &= \text{the number of expected events in the } i_{th} \text{ bin} \\
\vec{m} &= (m_1, m_2, ..., m_k)
\end{aligned}$$

We parameterized the number of expected events in the i_{th} bin in the following way:

$$m_i = N_{sig}P_{\pi\mu e_i} + N_{bg}P_{acc_i} \text{ where } \sum_{i=1}^{k} P_{\pi\mu e_i} = 1 \text{ and } \sum_{i=1}^{k} P_{acc_i} = 1 \tag{10.2}$$

where N_{sig} is the number of $K_{\pi\mu e}$ events, N_{bg} is the number of background events, and $P_{\pi\mu e_i}$ and P_{acc_i} are the probabilities to be in the i_{th} bin for signal and background, respectively. Defined this way m_i depends on two parameters, N_{sig} which we want to determine and N_{bgd} which we know from our analysis. The probabilities ($P_{\pi\mu e_i}$ and P_{acc_i}) were extracted using the likelihood templates, created for the signal and background hypothesis. Under the given parameterization of m_i the initial likelihood is transformed: $L_p(\vec{m};\vec{n}) \Rightarrow L(N_{sig}, N_{bg})$.

The probability is normalized to unity:

$$P(N_{sig}, N_{bg}) = \frac{L(N_{sig}, N_{bg})}{\int_0^\infty L(N_{sig}, N_{bg})dN_{sig}} \tag{10.3}$$

To determine the upper limit on the number of signal events at the 90 % CL one starts integrating the probability $P(N_{sig}, N_{bg})$ from zero up to N_{sig}^{max}, until the 0.9 probability level is reached:

$$\frac{\int_0^{N_{sig}^{max}} L(N_{sig}, N_{bg})dN_{sig}}{\int_0^\infty L(N_{sig}, N_{bg})dN_{sig}} = 0.90 \tag{10.4}$$

By solving the Eq. (10.4) one can numerically determine the number of signal events at the 90% CL

$$N_{\pi\mu e} = N_{sig}^{max} \tag{10.5}$$

10.2 Estimating the likelihood efficiency for suppressing accidental background

To estimate our effectiveness we varied the number of accidental events according to the Poisson distribution with the mean value of 8.2. As a next step, we simulated values of the general likelihood for accidental background, by using the template that we extracted. Finally, we integrated the probability [Eq. (10.4)] until the 0.9 level was reached, and thus the limit on the number of signal events was determined. That procedure was repeated 10^4 times. All this was done assuming that there are no signal events.

After outlining the method we used to determine the number of signal events($N_{\pi\mu e}$), we note that the result would depend on the prediction of the background (N_{bg}), which was $N_{bg} = 8.2 \pm 1.9$ events. The uncertainty in the background was taken into account by adding a Gaussian function with a mean b=8.2 and standard deviation σ_b=1.9 to the probability as a factor.

$$L_p^{(1)}(\vec{m};\vec{n}) = \prod_{i=1}^{k} \frac{m_i^{n_i} e^{-m_i}}{n_i!} \times e^{-\frac{(N_{bg}-b)}{2\sigma_b^2}} \tag{10.6}$$

This method led to only a small shift (about 1%) in the mean number of signal events at a 90% CL. In order to estimate the influence of the uncertainty in the background estimate on the determined number of signal events at the 90% confidence level, we changed the simulation procedure.

We simulated the number of background events with three different mean values - 6.3, 8.2, and 10.1. Such cases corresponded to the $\pm\sigma$ and the central value of our background prediction. In the procedure to determine the number of signal events $N_{bg} = 8.2$ was still used as an estimated number of background events. The corresponding distributions (Figure 10.1) and the table that shows the effect of uncertainty in the background on the determined number of signal (Table 10.1) are presented below.

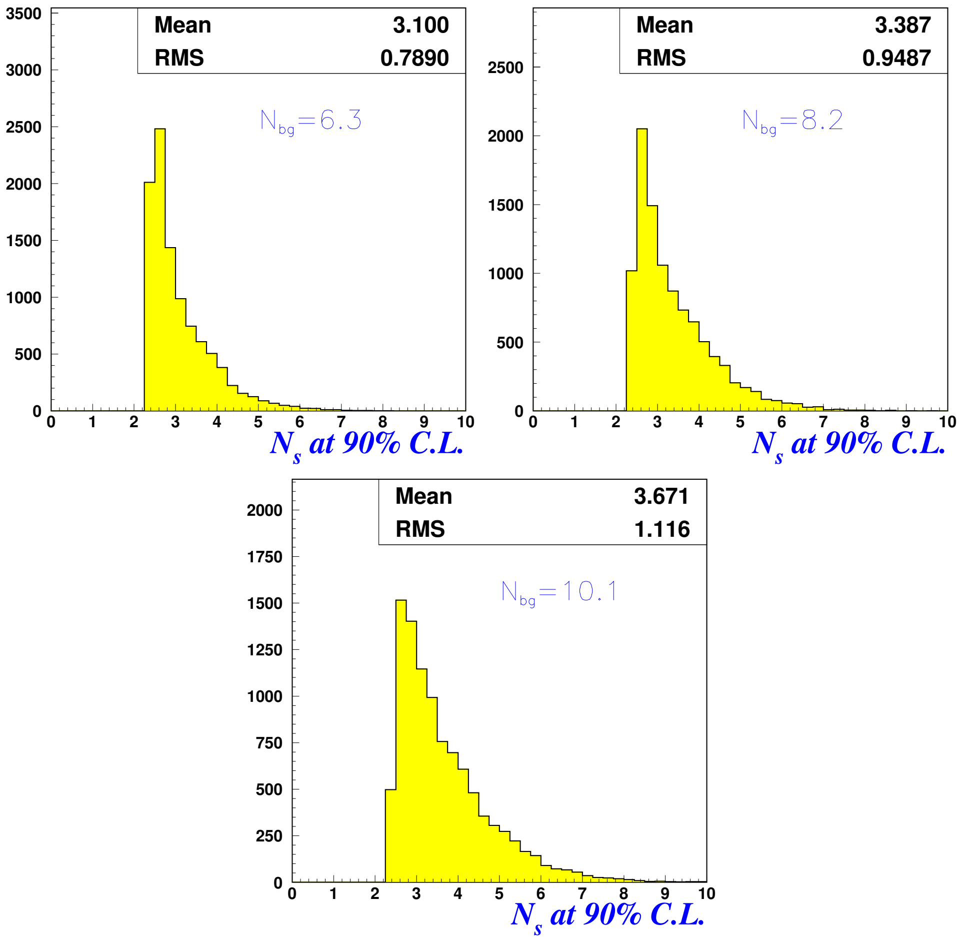


Figure 10.1: Distributions of the determined number of signal events at 90% CL for different level of accidental background put into simulation

The mean number of signal events (N_s) changed by approximately 8.5% with the change of the N_{bg} by one standard deviation. A relative statistical uncertainty in the N_s

did not change and was approximately 25% of the mean value. Such large uncertainty was caused by the fact that the number of simulated background events could fluctuate in accordance with the Poisson statistic.

Table 10.1: Systematic effect of the change in the background level on the predicted number of signal events at a 90% CL. $N_{bg}^{simulated}$ - mean value of simulated background events; $N_{bg}^{Predicted}$ - number of background events, used to determine the number of signal events: N_s - mean number of signal events; σN_s - one standard deviation of the number of signal events.

$N_{bg}^{simulated}$	$N_{bg}^{Predicted}$	N_s	σN_s
6.3	8.2	3.10 (-8.5%)	0.79
8.2	8.2	3.39 (0%)	0.949
10.1	8.2	3.67 (+8.4%)	1.12

10.3 Final likelihood analysis

After all final $K_{\pi\mu e}$ selection cuts were applied, the values of the general likelihood were calculated for all surviving events. Our background study showed that we expected 8.2 $\pm$ 1.9 accidental events. In addition, we expected that in the absence of signal events an upper limit on the number of signal events of 3.4 $\pm$ 1 (Figure 10.1) could be established. Below we present the actual result, which appeared after the examination of the signal region.

Eight events survived the final selection, consistent with our background prediction of 8.2 $\pm$ 1.9. Examination of the likelihood distribution of the surviving events revealed that they were more consistent with the background hypothesis (Figure 10.2). A simple cut on the likelihood would remove all surviving events with only a 3% loss in efficiency.

However, as it was decided beforehand, we determined the upper limit on the number of signal events at a 90% CL by using the Bayesian approach (Figure 10.3). Solving Eq.(10.4) yielded the upper limit on the number of signal events :

$N_{sig} < 2.4$ events,

which is consistent with our prediction of 3.4 $\pm$ 1.

The distributions of the variables used in the general likelihood for the observed events are presented in Figure 10.4.

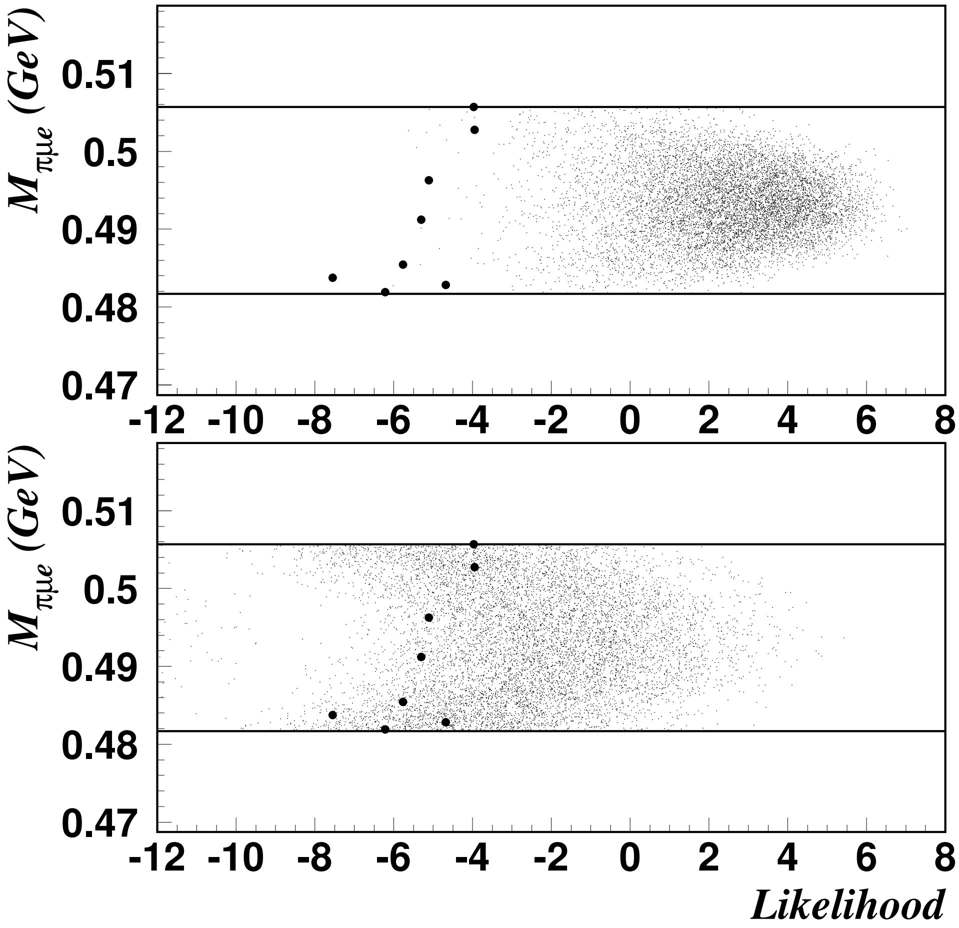


Figure 10.2: Scatter plot of E865 1998 $K_{\pi\mu e}$ data (bold dots) and simulated signal(top) and background (bottom). The abscissa is the log-likelihood of the reconstructed events, the ordinate is the invariant mass ($M_{\pi\mu e}$) of the detected particles. The horizontal lines demark the 3σ mass region

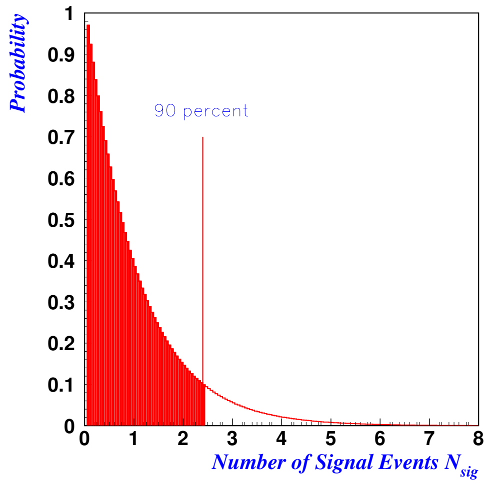


Figure 10.3: Integral of probability [Eq. (10.4)] for the number of $K^+ \to \pi^+\mu^+e^-$ events. The line denotes the 90% CL level upper limit and corresponds to 2.4 signal events.

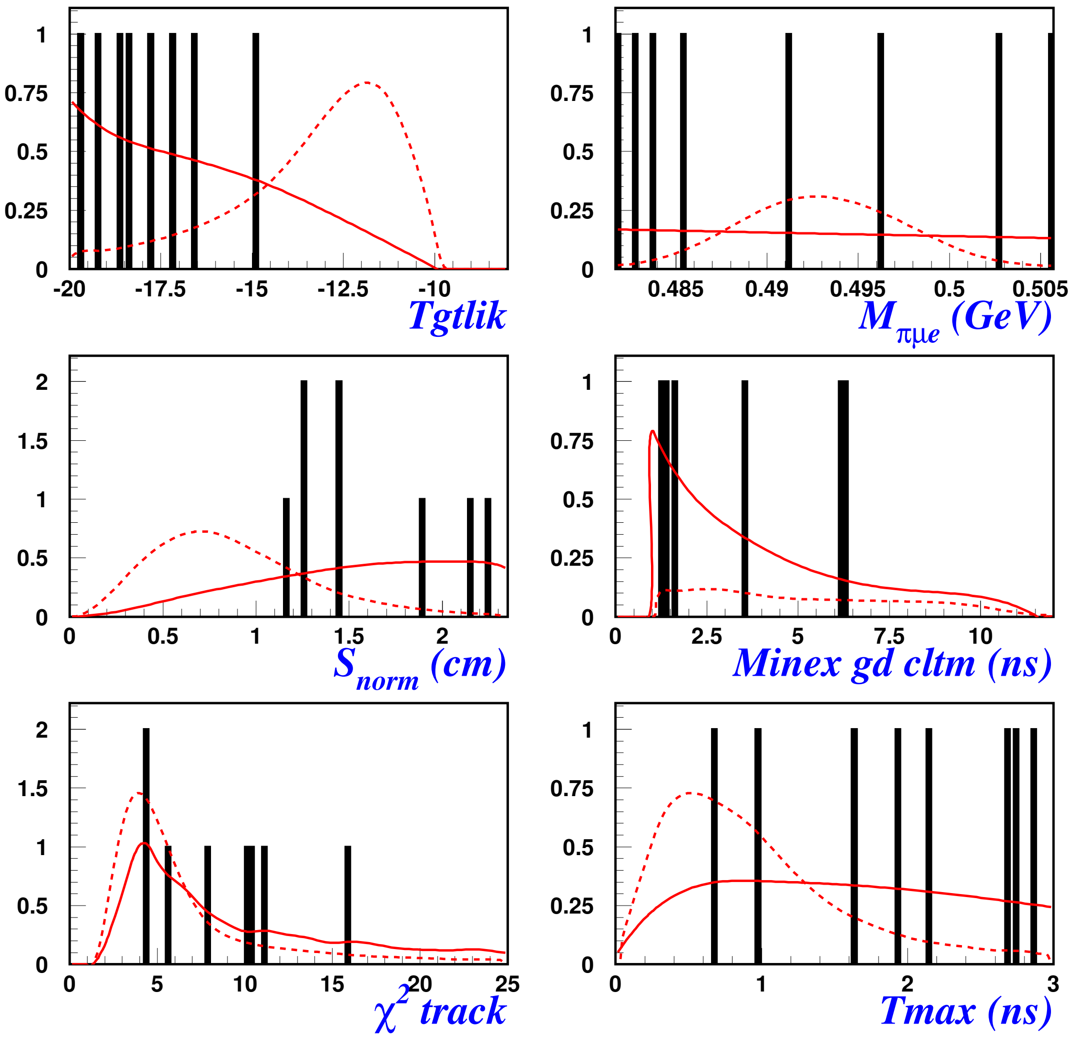


Figure 10.4: Distributions of Tgtlik, reconstructed kaon mass $M_{\pi\mu e}$, vertex quality S_{norm}, minimum time difference between extra clumps in the calorimeter and closest in time track pair (Minex gd cltm), χ^2 of a track and Tmax for eight final events. For the reference p.d.f.'s are shown for signal (dotted line) and background (solid line).

Chapter 11

Normalization and Acceptance Calculations

In order to covert the upper limit on the number of signal events into an upper limit on the $K^+ \to \pi^+\mu^+e^-$ branching ratio, the total number of K^+ decays should be determined.

To determine the total number of K^+ decays, one can try to determine the number of K^+ produced from the number of protons at the target, the beam line acceptance and decay probability in the decay region. Alternatively one can normalize to a well-understood kaon decay mode with relatively high branching ratio, which is observable in the detector. We chose the latter method, since it is clearly more accurate.

The normalization mode should be similar to the $K^+ \to \pi^+\mu^+e^-$ decay so that efficiency studies of the former could be applied. The most evident candidate was the $K^+ \to \pi^+\pi^+\pi^-$ decay which can be fully reconstructed. The disadvantage of using $K^+ \to \pi^+\pi^+\pi^-$ decays as a normalizer is the fact, that it has only π^+ in common with the $K^+ \to \pi^+\mu^+e^-$. Another candidate for the normalization mode - $K^+ \to \pi^0\pi^+$ ($\pi^0 \to e^+e^-\gamma$) decay, has two particles (π^+ and an e^-) in common with the signal. However $K^+ \to \pi^0\pi^+$ also has a photon, which is unobserved in most cases and this decay cannot be fully reconstructed. For this analysis $K^+ \to \pi^+\pi^+\pi^-$ decay was chosen to be the normalization decay mode.

With the normalization mode selected, the formula for the upper limit on the $K^+ \to \pi^+\mu^+e^-$ branching ratio at the 90 % CL is:

$$B(\pi\mu e) < B(\pi\pi\pi) \cdot \frac{N(\pi\mu e)}{N(\pi\pi\pi)} \cdot \frac{Accep(\pi\pi\pi)}{Accep(\pi\mu e)} \cdot C \tag{11.1}$$

where B denotes the branching ratio of the decay, $N(\pi\mu e)$ is the 90% CL upper limit on the number of signal events, $N(\pi\pi\pi)$ is the number of observed K_{Tau} events, adjusted for pre-scale factors, *Accep* is acceptance of the detector system, and finally C is the correction factor accounting for efficiency differences between the selection of signal ($K_{\pi\mu e}$) and normalizer ($K_{\pi\pi\pi}$) decay modes.

The acceptance was calculated using a Monte-Carlo (MC) simulation of the detector The correction factor C was determined using efficiencies determined from real data. How the different factors in Eq. (11.1) were determined is described below.

11.1 Total number of $K^+ \to \pi^+\pi^+\pi^-$ events

The number of normalization events was deduced from analyzing the pass1 TAU stream data. In order to select a clean sample of the K_{Tau} events a set of kinematic cuts was applied. Most of them were identical to the ones used in the final $K_{\pi\mu e}$ selection. This reduced the systematic uncertainties in the acceptance ratio calculation. However, some of the cuts used for $K_{\pi\mu e}$ selection were removed. The full list of cuts is shown in the Table 11.1.

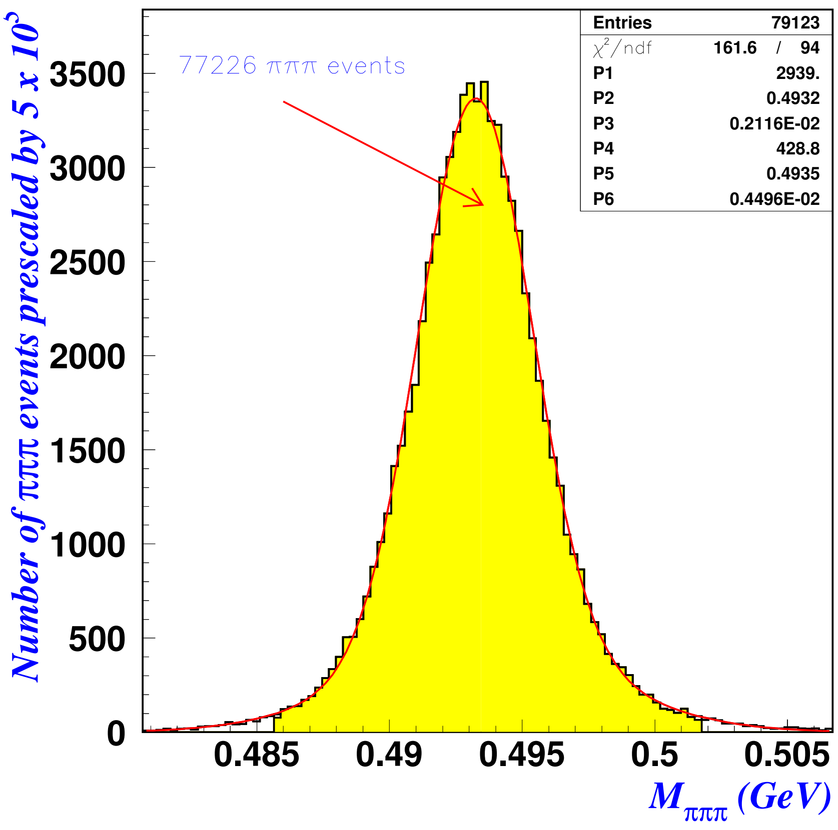


Figure 11.1: Kaon reconstructed mass distribution for normalization $K^+ \to \pi^+\pi^+\pi^-$ events sample.

The TAU trigger was prescaled by a factor of 10^4 in hardware. Due to the large statistical size and lack of disk space, during the analysis of the pass1 TAU stream, the additional software prescale factor of 50 was introduced. Applying the selection cuts to the data sample yielded 77226 surviving events (See Figure 11.1).This determined the total number of the K_{Tau} events:

$$N(\pi\pi\pi) = 77226 \times 50 \times 10^4 = 3.86 \times 10^{10} \text{ events} \tag{11.2}$$

Table 11.1: Cuts used for determining the number of $K^+ \to \pi^+\pi^+\pi^-$ events.

Variable	*Description*	*Efficiency*
S_{norm}	Normalized vertex miss distance $S < 2.35$ cm	0.927 ± 0.005
$\chi^2_{trk_i}$	$\sum_{i=1}^{3} \chi^2_{trk_i} < 25$ Quality of track reconstruction	0.967 ± 0.005
$Tgtlik$	Target likelihood > -20 and tight phase space cuts	0.923 ± 0.005
$Tmax$	Maximum time difference between π^- and π^+ track times < 3 ns	0.911 ± 0.004
$Minex\ Gd\ Cltm$	Minimum time difference between time of extra calorimeter clump and time of track pair that is closest in time. Minex Gd Cltm > 1 ns	0.956 ± 0.005
N_γ	No gamma clumps allowed within 2 ns	0.965 ± 0.005
S_{2trk}	$S_{2trk} < 2.35$ Normalized S of a two-track vertex	0.959 ± 0.005
$M_{\pi\pi\pi}$	Reconstructed invariant mass of three pions within 3σ ($\sigma = 2.5$ MeV) of $M_K = 0.4937$ GeV	0.976 ± 0.005

11.2 Acceptance definition

In this section we define the acceptance used for producing the final result. Details on the calculation and simulation will be discussed in the following sections.

The following requirements were imposed in the simulation.

- The K^+ must decay inside the decay volume and all the tracks of the daughter particles must be contained inside the pre-defined acceptance regions of the detector, given by the outer edges of detector elements.

- No decay product should pass through the dead region (beam hole) of the A counter and the calorimeter.
- All MC events were required to satisfy the hardware trigger requirement, which were simulated in the software.

11.3 Monte Carlo simulations

11.3.1 Overview

The interactions of the beam particles and the decay products with the apparatus were simulated using the program KAY [33], specifically developed for E865. It is based on the GEANT3 [34] software package from CERN. It generates the initial particle at a user defined location, and then tracks it to a random decay point. After that, decay products are tracked through the E865 apparatus, and all the information about positions, momenta, energy loss, etc. is recorded for the individual detector elements. In the next step this information is digitized. It means that from the deposited energy ADC values are calculated, positions at the PWC are converted to wire numbers, etc. Local inefficiencies were mapped and taken into account. Finally all the information was written in the standard E865 format, so that simulated events could be analyzed using the same offline procedure as measured data.

In order to save computer time, the initial simulation used an ideal efficiency map. Afterwards with the help of another program, called ZORA, the user could apply the efficiencies of individual detector parts and events would be filtered accordingly. That saved a lot of time, since the most consuming part was tracking decay products through the apparatus. Applying the efficiencies could be done much faster.

11.3.2 Defining phase space of the beam kaons

In principle the simulation could start with K^+ at the target, and then transport them along the beam line. However it will require large CPU time to track each kaon until it reaches the decay region, and in addition uncertainties in modelling the beam line would cause systematic uncertainties. This can be avoided if kaons are generated at the end of the beam line, *i.e.* the beginning of the decay region. For this one must know the phase space distribution of kaons at the beginning of the decay volume. That was achieved by selecting good K_{Tau} events from the data, and then tracking each reconstructed kaon to the beginning of the decay volume at Z = -490 cm. The phase space parameters (X,θ_X,Y,θ_Y, and P_K) were written out into a template file (RAYOUT.DAT). This method still had a disadvantage, since the phase space was smeared by the resolution of the reconstruction. During the Monte-Carlo simulation such resolution smearing would be applied again. But the effect turned out to be small and considering the fact it was a search for a rare decay, not a precision measurement of the branching ratio, we proceeded with the method described.

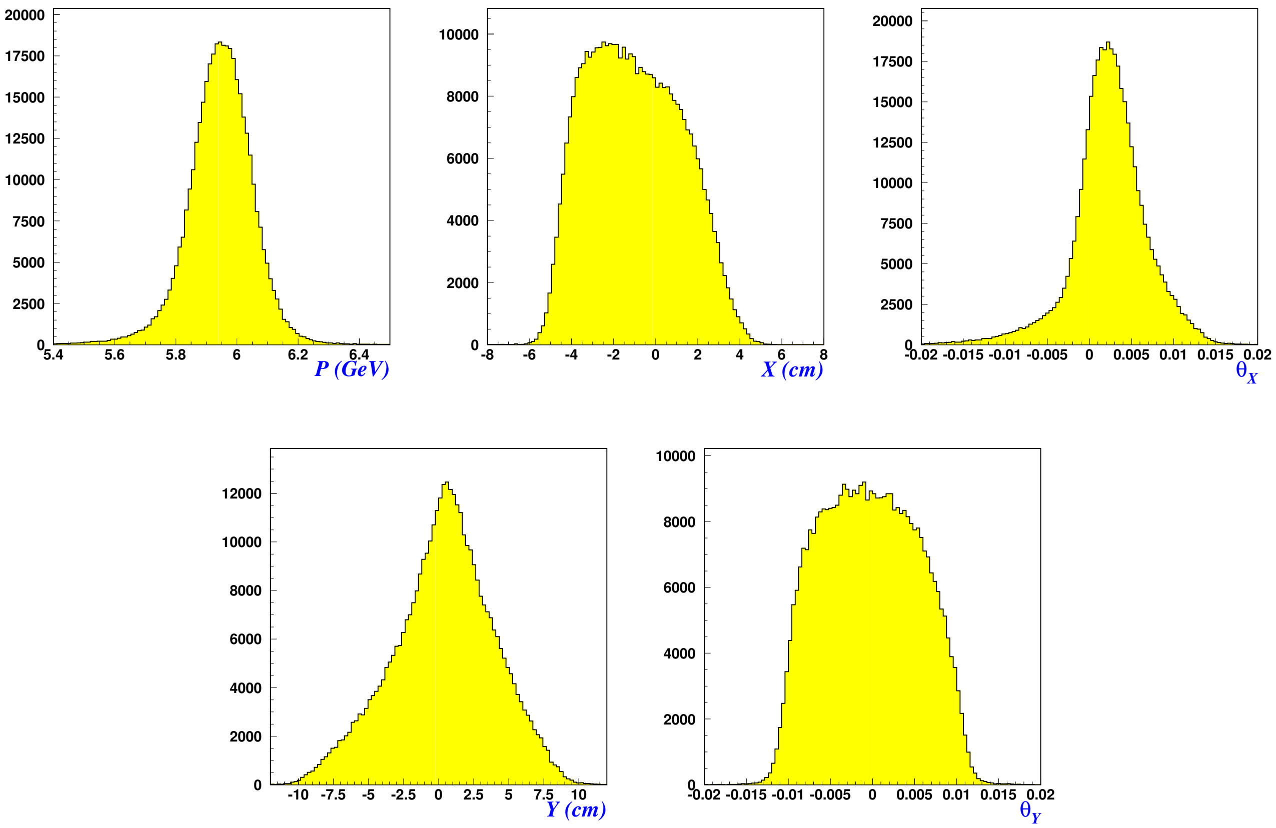


Figure 11.2: Kaon phase space distribution at Z=-490cm, derived from $K^+ \to \pi^+\pi^+\pi^-$ decays, using measured data.

11.3.3 Data and MC distributions comparison

In order to check the quality of the simulation control distributions for various kinematic variables were compared to measured data from $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi^0\pi^+$ decays. Due to the different phase space and different daughter particles, those two decays have significantly different control distributions.

The most relevant distributions like vertex quality and Z distribution, the reconstructed kaon momentum, and individual momenta of the particles, as well as some others are presented in Figures A.1-A.6. The agreement of MC and data was satisfactory and gave confidence in the acceptance ratio calculation.

For additional consistency checks we selected two groups of runs from different parts of the running period (run No.4566...4608 and No.4374...4395) and selected $K_{\pi\pi\pi}$ events from the TAU triggered data and events with a low M_{ee} pairs from the EEPS triggered data. We estimated the number of $K_{\pi\pi\pi}$ events by counting the number of low M_{ee} mass events, found after analyzing the EEPS triggered data, and applying corresponding acceptance ratio. For the calculation of the low M_{ee} mass events acceptance we simulated $K^+ \to \pi^0\pi^+$, $K^+ \to \pi^0\mu^+\nu$, $K^+ \to \pi^0 e^+\nu$, $K^+ \to \pi^+\pi^0\pi^0$ decays and weighted them by their presently known branching ratios [24]. Predictions obtained this way were in agreement with the number of observed $K_{\pi\pi\pi}$ decays within the statistical uncertainty.

Table 11.2: Number of observed and predicted $K^+ \to \pi^+\pi^+\pi^-$ events for two different run groups. Predicted number is obtained by counting low M_{ee} events from EEPS triggers and applying acceptance ratio derived from MC simulations.

Runs:	$K^+ \to \pi^+\pi^+\pi^-$ events observed	$K^+ \to \pi^+\pi^+\pi^-$ events predicted
4566...4608	$(1.85 \pm 0.01) \times 10^9$	$(1.91 \pm 0.03) \times 10^9$
4374...4395	$(1.04 \pm 0.01) \times 10^9$	$(1.02 \pm 0.01) \times 10^9$

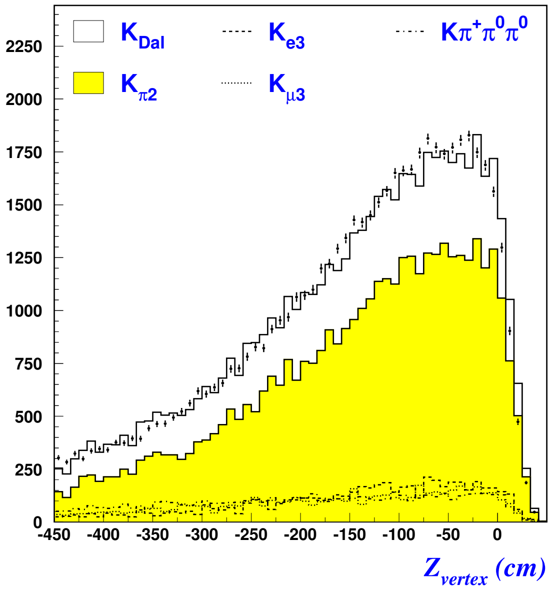


Figure 11.3: Data and Monte Carlo comparison of the distribution of the Z of the vertex for a mixture of Dalitz decays. Histograms represent MC and data is represented by markers with error bars.

Systematic uncertainty, due to some possible discrepancies in the geometrical description of the detector elements will be discussed later.

11.4 Acceptance calculation

With a reasonable agreement between Monte Carlo and Data reached we calculated acceptance for $K_{\pi\mu e}$ and $K_{\pi\pi\pi}$ decays.

For 2.5×10^5 simulated $K^+ \to \pi^+\pi^+\pi^-$ events 31600 were accepted, and that yielded:

$$\text{Accep}(\pi\pi\pi) = 0.1264 \pm 0.0008 \tag{11.3}$$

For 2.5×10^5 simulated $K^+ \to \pi^+\mu^+e^-$ events 17003 were accepted, resulting in:

$$\text{Accep}(\pi\mu\text{e}) = 0.068 \pm 0.001 \tag{11.4}$$

and

$$\frac{\text{Accep}(\pi\pi\pi)}{\text{Accep}(\pi\mu\text{e})} = 1.86 \pm 0.02. \tag{11.5}$$

To estimate the influence of alignment inaccuracies we changed the size of a dead region, which is the most sensitive parameter for both $K_{\pi\mu e}$ and K_{Tau} decays acceptance. We considered three cases, for which we defined a smaller ($\Delta X = \Delta Y$ =-0.5cm), nominal and a larger dead region ($\Delta X = \Delta Y$ =+0.5cm).

While the changes in the acceptance of the $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi^+\mu^+e^-$ decay were noticeable, the acceptance ratio showed only limited variation:

- 1.88 ± 0.02 - smaller dead region;
- 1.86 ± 0.02 - nominal dead region;
- 1.83 ± 0.02 - larger dead region.

From this study we deduce systematic uncertainty of roughly 1.5% on the acceptance ratio:

$$\frac{\text{Accep}(\pi\pi\pi)}{\text{Accep}(\pi\mu\text{e})} = 1.86 \pm 0.02(stat.) \pm 0.03(syst). \tag{11.6}$$

11.5 Correction factors

The correction factor C in Eq. (11.1) contains efficiencies to account for the difference in reconstructing $K^+ \to \pi^+\pi^+\pi^-$ and $K^+ \to \pi^+\mu^+e^-$ events. Explicitly we have

$$C = \frac{\prod_{i=1}^{n} \epsilon_i^{\pi\pi\pi}}{\prod_{j=1}^{n} \epsilon_j^{\pi\mu e}}. \tag{11.7}$$

Because most of the kinematic cuts for the $K^+ \to \pi^+\pi^+\pi^-$ sample selection were also applied to select $K^+ \to \pi^+\mu^+e^-$ events, terms in the numerator would cancel with those contained in the denominator where we assumed the efficiency values for each of these terms were the same for both samples. Eq. (11.7) reduced to

$$C = \frac{1}{\prod_{j'=1}^{n} \epsilon_{j'}^{\pi\mu e}} \tag{11.8}$$

The efficiencies ($\epsilon_{j'}^{\pi\mu e}$) listed in Table 11.3 can be combined to yield

$$C = 3.32 \pm 0.08 \tag{11.9}$$

Table 11.3: Efficiencies used to calculate correction factor.

Effic.	*Description*	*Value*
ϵ_π	π^+ detection (e^+ veto)	0.781 ± 0.004
ϵ_μ	μ^+ detection	0.79 ± 0.01
ϵ_e	e^- detection	0.745 ± 0.003
$\epsilon_{M_{ee}}$	invariant mass cut $M_{e^-\pi^+} > 55$ MeV and $M_{e^-\mu^+} > 55$ MeV	0.919 ± 0.007
ϵ_{CL}	e^- required on the left side of the Čerenkov counters	0.926 ± 0.008
ϵ_{CR}	π^+ required on the right side of the Čerenkov counters	0.891 ± 0.008
ϵ_z	reject π^+ from $K^+ \to \pi^0\pi^+$ decay	0.865 ± 0.009

Chapter 12

Conclusion

Finally we are now in the position to calculate the result for the upper limit on the branching ratio:

$$B(\pi\mu e) < B(\pi\pi\pi) \cdot \frac{N(\pi\mu e)}{N(\pi\pi\pi)} \cdot \frac{Accep(\pi\pi\pi)}{Accep(\pi\mu e)} \cdot C \tag{12.1}$$

The values of the factors in 12.1 along with their short discriptions are presented below:

$B(\pi\pi\pi)$	$K^+ \to \pi^+\pi^+\pi^-$ branching ratio	0.0558 ± 0.0003
$N(\pi\mu e)$	90% CL upper limit on the number of $K_{\pi\mu e}$ events	2.4
$N(\pi\pi\pi)$	Number of reconstructed $K^+ \to \pi^+\pi^+\pi^-$ decays	3.86×10^{10}
$\frac{Accep(\pi\pi\pi)}{Accep(\pi\mu e)}$	Acceptance ratio between $K_{\pi\pi\pi}$ and $K_{\pi\mu e}$ decays	1.86 ± 0.04
C	correction factor for efficiencies etc.	3.32 ± 0.08

Puting these values into the Eq. (12.1) yields:

$$\boxed{B(\pi\mu e) < 2.4 \times 9.0 \times 10^{-12} = 2.2 \times 10^{-11}} \tag{12.2}$$

Previous experiments and earlier E865 results showed that [2], [3]:

$$B(\pi\mu e) < 2.8 \times 10^{-11}$$

Linearly combining our result with the previous one leads to the new upper limit on $K^+ \to \pi^+\mu^+e^-$ branching ratio:

$$\underline{B(\pi\mu e) < 1.2 \times 10^{-11}}$$

Although no evidence of new physics was found, an upper limit on the $K_{\pi\mu e}$ branching ratio was improved from 2×10^{-10} to 1.2 $\times 10^{-11}$. Extension theories "phase space" was reduced.

In accordance with the previously discussed model (Section 2.2.1) the mass of the corresponding Extended Technicolor Boson is bound by the $K_{\pi\mu e}$ branching ratio to be:

$$M_{ETC} > 80 \text{ TeV} \tag{12.3}$$

Similarly in accordance with the discussed Leptoquark model (Section 2.2.4) the mass of the Leptoquark is bound to be:

$$M_{LQ} > 270 \text{ TeV} \tag{12.4}$$

Little hope is left to find Lepton Family Number Violation in the kaon decays unless it occures through a generation number conserving process. Other searches for the LFN violating processes such as $\mu^+ \to e^+\gamma$ and $\mu^- N \to e^- N$ with the projected single event sensitivity on the order of $O(10^{-14})$ and $O(10^{-16})$ respectively are planned for the future.

Appendix A

Data and MC comparison plots

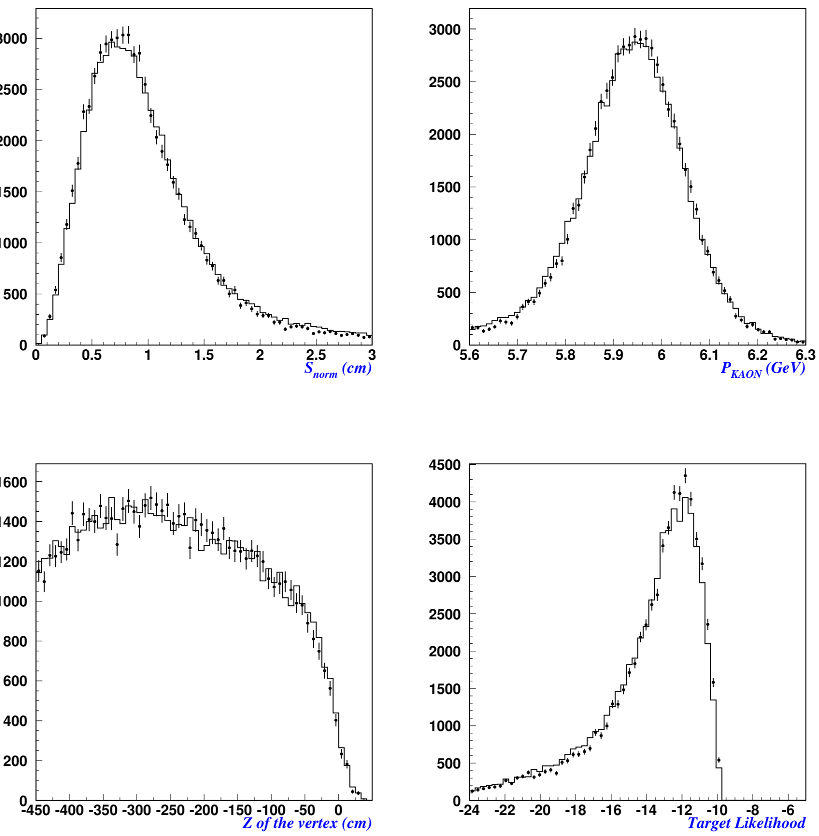


Figure A.1: Distributions of various kinematic variables for simulated (histogram) and reconstructed from the data (markers with error bars) $K^+ \to \pi^+\pi^+\pi^-$ events.

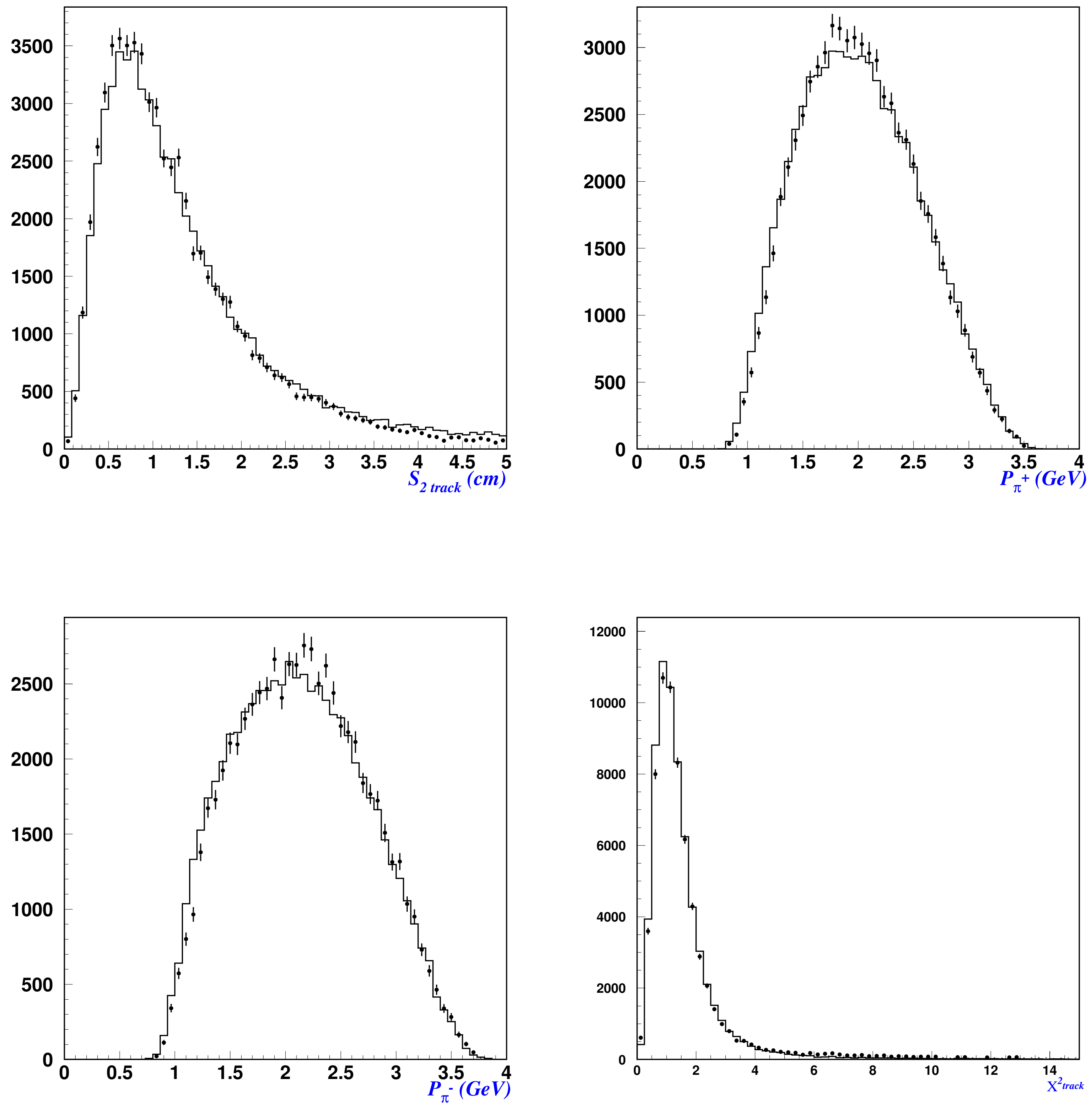


Figure A.2: Distributions of various kinematic variables for simulated (histogram) and reconstructed from the data (markers with error bars) $K^+ \to \pi^+\pi^+\pi^-$ events.

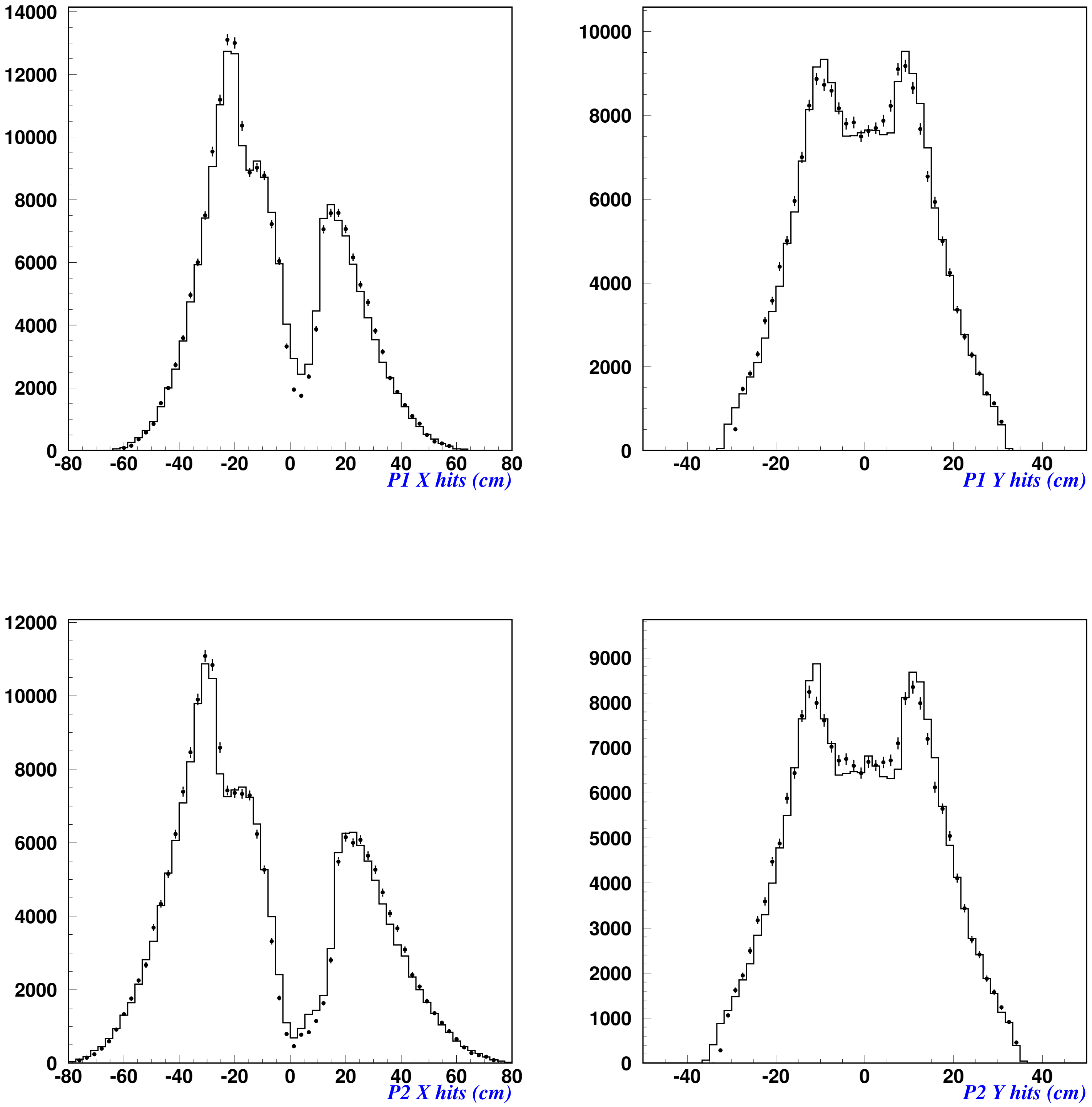


Figure A.3: Distributions of various kinematic variables for simulated (histogram) and reconstructed from the data (markers with error bars) $K^+ \to \pi^+\pi^+\pi^-$ events.

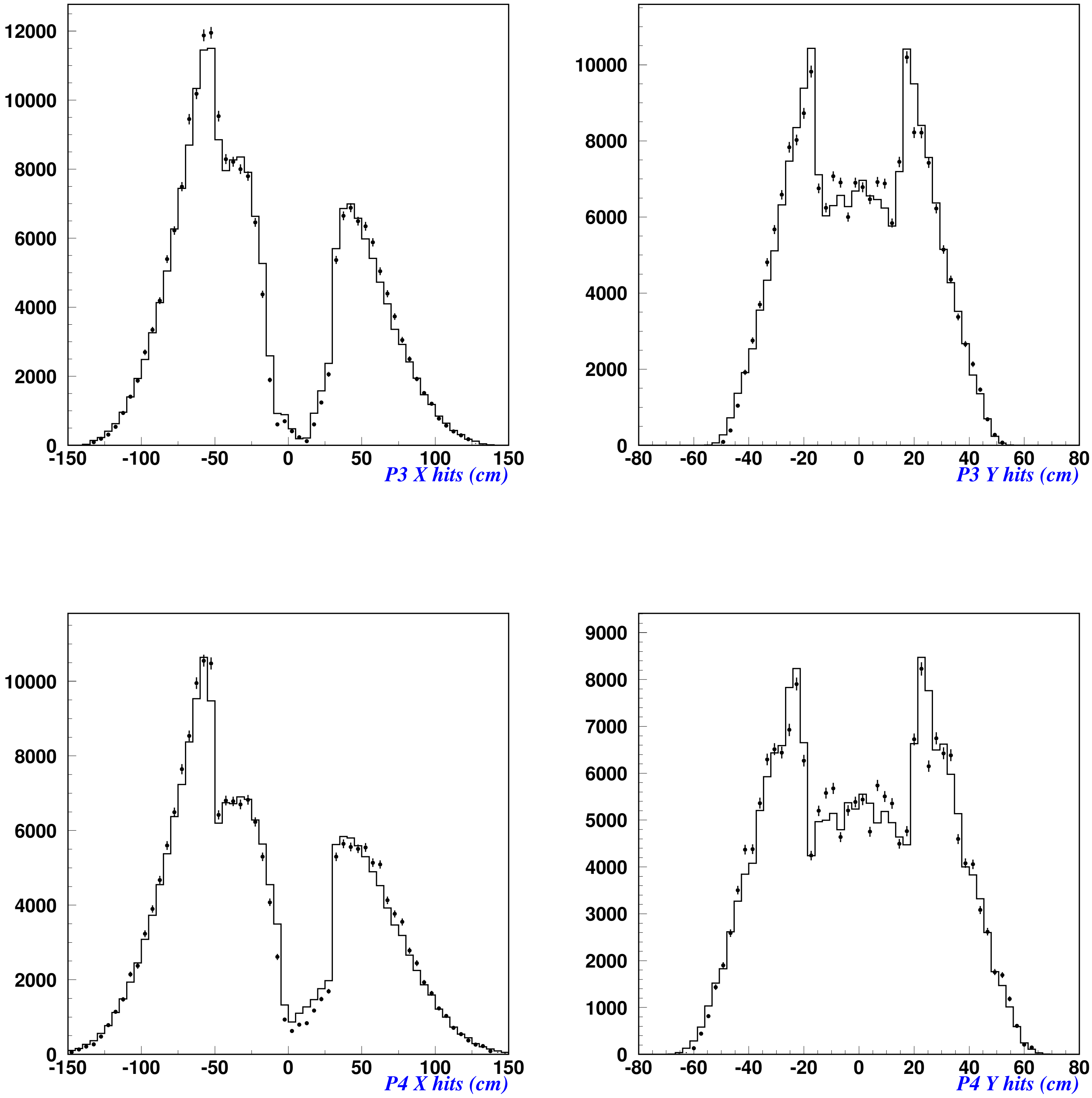


Figure A.4: Distributions of various kinematic variables for simulated (histogram) and reconstructed from the data (markers with error bars) $K^+ \to \pi^+\pi^+\pi^-$ events.

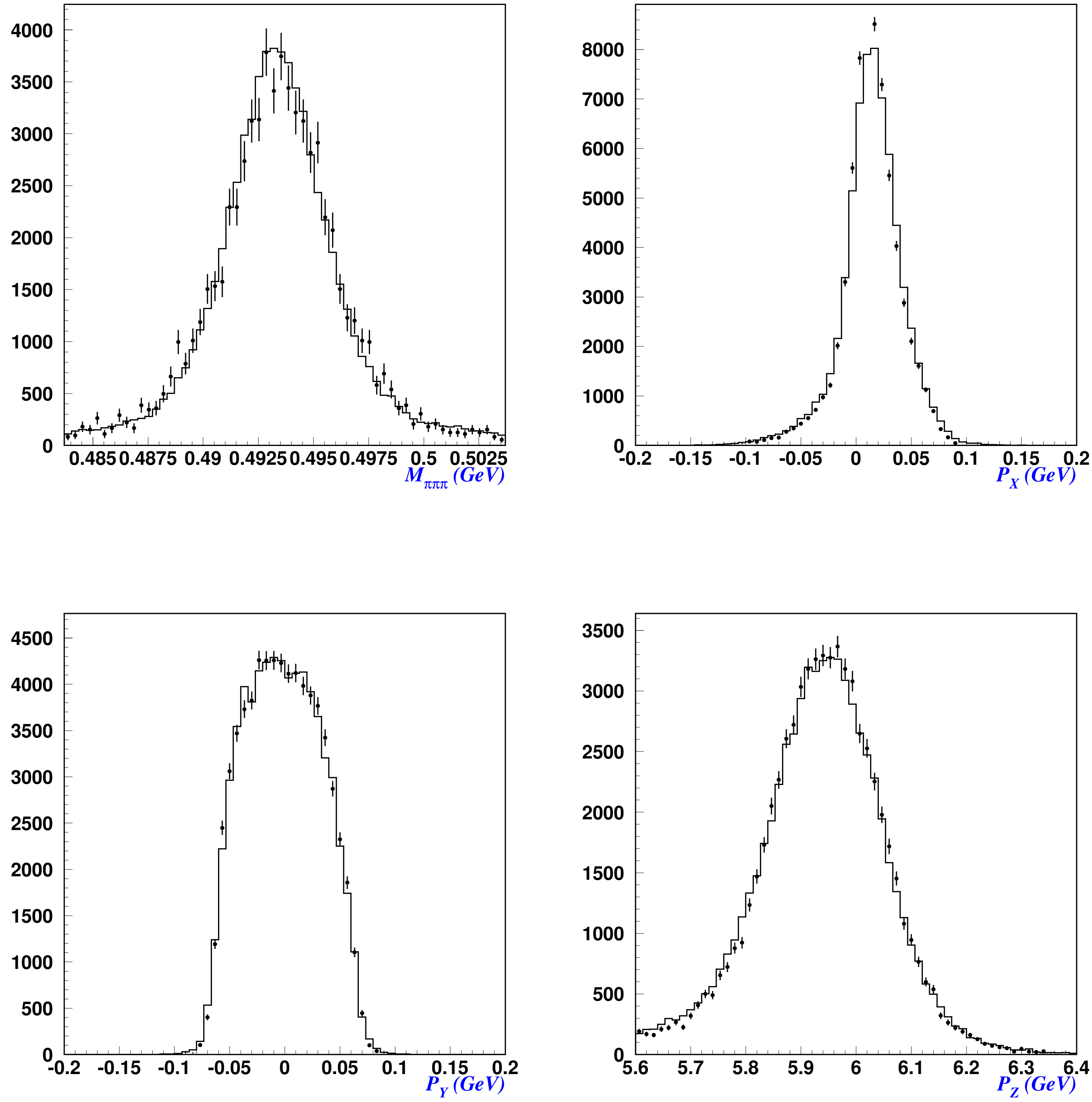


Figure A.5: Distributions of various kinematic variables for simulated (histogram) and reconstructed from the data (markers with error bars) $K^+ \to \pi^+\pi^+\pi^-$ events.

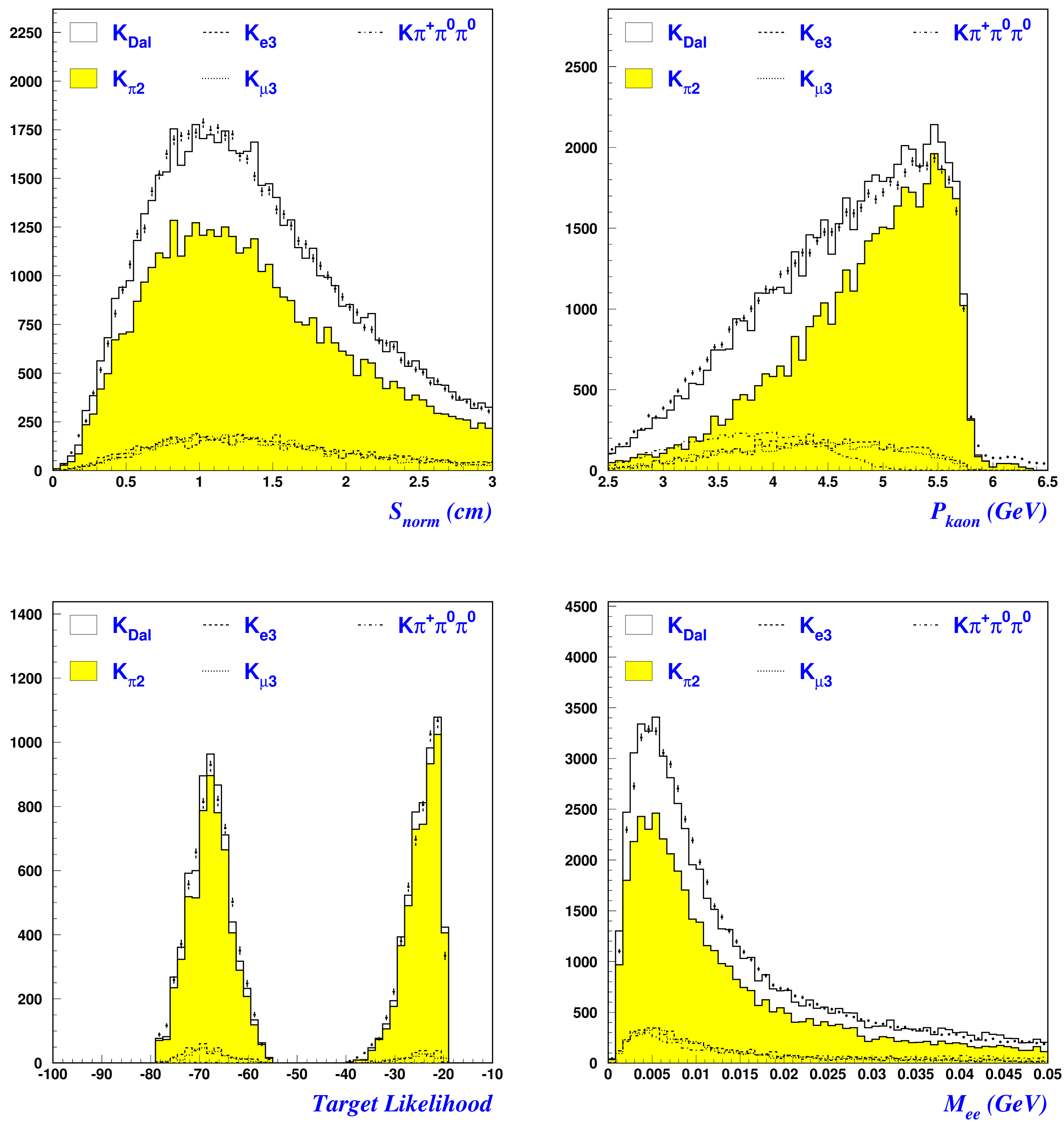


Figure A.6: Distributions of various kinematic variables for simulated (histogram) and reconstructed from the data (markers with error bars) low M_{ee} events.

Bibliography

[1] A. M. Diamant-Berger et al. Study of some rare decays of the K^+ meson. *Phys. Lett.*, 62B:485, 1976.

[2] H.Do. *An Improved Upper Limit on the Branching Ratio of the Decay $K^+ \to \pi^+\mu^+e^-$.* PhD thesis, Yale Univeristy, 2000.

[3] R. Appel, *et al.* An improved limit on the rate of decay $K^+ \to \pi^+\mu^+e^-$. *Phys. Rev. Lett.*, 85:2450, 2000.

[4] A. Pich. Precision tests of the standard model. 1997. Lectures given at the 25th Winter Meeting on Fundamental Physics, hep-ph/9711279.

[5] Leonard Susskind. The gauge hierarchy problem, technicolor, supersymmetry, and all that. *Phys. Rep.*, 104:181, 1984.

[6] G. Kane. *Modern Elementary Particle Physics.* Addison-Wesley Publishing Company, 1993. pages 105-108.

[7] Edward Farhi and Leonard Susskind. Technicolor. *Phys. Rep.*, 74:277–321, 1981.

[8] Kenneth Lane. Technicolor. 1995. hep-ph/9501249.

[9] Kenneth Lane. An introduction to technicolor. 1994. hep-ph/9401324.

[10] S. Dimopoulos and J. Ellis. Challenges for extended technicolor theories. *Nucl. Phys.*, B182:505, 1981.

[11] Robert N. Cahn and Haim Harari. Bounds on the masses of neutral generation changing gauge bosons. *Nucl. Phys.*, B176:135, 1980.

[12] Bob Holdom. Flavor changing suppression in technicolor. *Phys. Lett. B*, 143:227–232, 1984.

[13] Stephen F. King. Composite/elementary extended technicolor. *Phys. Lett. B*, 195:66–73, 1987.

[14] D. Ambrose et al. New limit on muon and electron lepton number violation from $k_L^0 \to \mu^\pm e^\mp$ decay. *Phys. Rev. Lett.*, 81:5734–5737, 1998.

[15] T. Appelquist, D. Karabali, and L. C. R. Wijewardhana. Chiral hierarchies and the flavor changing neutral current problem in technicolor. *Phys. Rev. Lett.*, 57:957, 1986.

[16] O. Shanker. Cp violation, flavor violation and fermion mass relations in some horizontal gauge theories. *Nucl. Phys.*, B185:382, 1981.

[17] O. Shanker. Flavor violation, scalar particles and leptoquarks. *Nucl. Phys.*, B206:253, 1982.

[18] Ricardo Gaitan-Lozano et al. Signals of extra gauge bosons and exotic leptons in su(6)-l x u(1)-y. *Phys. Rev.*, D51:6474–6483, 1995.

[19] Jack L. Ritchie and Stanley G. Wojcicki. Rare k decays. *Rev. Mod. Phys.*, 65:1149–1197, 1993.

[20] B. Mukhopadhyaya and A. Raychaudhuri. $K_L \to \mu e$: A possible probe of slepton mass splitting? *Phys. Rev.*, D42:3215, 1990.

[21] J. C. Pati and A. Salam. Lepton number as fourth 'color'. *Phys. Rev.*, D10:275, 1974.

[22] Johannes Blümlein. On the leptoquark interpretation of the high q^2 events at hera. 1997. hep-ph/9706362.

[23] P. Langacker, S. Uma Sankar and K. Schilcher, Large Neutrino Masses," Phys. Rev. D **38** (1988) 2841.

[24] K. Hagiwara *et al.* [Particle Data Group Collaboration], Phys. Rev. D **66** (2002) 010001.

[25] A. M. Bellavance, *Search for the lepton-flavor-number violating decay $K_L^0 \to \pi^0\mu^\pm e^\mp$ in the full E799II KTeV data set.* FERMILAB-THESIS-2002-20

[26] Stefan Pislak. *A Search for the decay $K^+ \to \pi^+\mu^+e^-$.* PhD thesis, University of Zürich, 1997.

[27] Doug Bergman. *A Search for the decay $K^+ \to \pi^+\mu^+e^-$.* PhD thesis, Yale University, 1997.

[28] R. Appel ,*et al.*, Nucl. Instr. and Meth. A **479**,349 (2002).

[29] G. S. Atoyan et al. Lead-scintillator electromagnetic calorimeter with wave length shifting fibre readout. *Nucl. Instr. and Meth.*, A320:144, 1992.

[30] D.D. Koetke, Proc. 4th Int. Conf. on the Intersections between Part. and. Nucl. Phys., Tucson (1991), W.T.H. van Oers ed., AIP Conf. Proc. **243**, 397 (1992).

[31] W. Menzel, Thesis, University of Basel (1996).

[32] A. Sher, *Positive kaon decay into neutral pion, positron, and neutrino relative branching ratio measurement in the E865 experiment at the Brookhaven National Laboratory*, PhD thesis, Univeristy of Pittsburgh, 2002.

[33] Stefan Pislak. *KAY: E865 Monte-Carlo.* Zurich, Switzerland, 1996.

[34] CERN, Geneva, Switzerland. *GEANT: Detector Description and Simulation Tool*, 1993.

Acknowledgements

Looking back at the past four years makes me feel that they were "well spent". That being said I feel the need to express my gratitude to the very people who made these years challenging, inspiring and into ones I will not forget.

I would like to thank my advisors, yes, I was lucky to have two of them. My official boss - Peter Truöl was always there to comment and check whether I was on the right track or wandering off. His ability to grasp the main point was truly amazing. He taught me a lot by providing thoughtful insight and what's also important allowed me to "made my own mistakes" when we were thinking differently. From him I learned how important it is to keep in mind the whole picture and not to focus blindly on a particular aspect. For that I'm truly greatfull.

I'm equally thankful to Mike Zeller, who was my unofficial boss and besides all the guidance and sometimes needed encouragement as well as criticism that i received from him, he was and is an example of a true physicist who besides being excited about his work tries and succeeds in understanding every aspect of it. His excitement about my work and physics was always inspiring and helping.

I can't help but express my gratitude to the E865 collaboration and the people who helped me finish this work. I want to thank Julia Thompson and David Kraus for looking into and as a consequence improving alignment settings of the E865 and thanks to Peter Lichard for creating a better magnetic field approximation. In addition I thank Julia for introducing me to the E865 and for all the helpful suggestions as well as raised questions, which I perceived as a true researcher attitude, where no stones should be left unturned.

I also would like to thank Hong Ma. During the two years spent in Brookhaven I could always count on his expertise. Our discussions were always satisfying and helpful. Same can be said about Andrej Poblaguev, a colleague who was always ready to discuss my work and provide a sound advice.

I also should thank Don Lazarus, who was always eager to help. In addition I must say that I truly enjoyed working with Robin Appel, Grigor Atoyan, Naipor Cheung, Vladimir Issakov, Heinz Kaspar, Jim Lowe and Stefan Pislak. Stefan also happened to introduce me to the software packages, used by the e865 offline, and was always thorough in answering all of my questions.

I also would like to thank people in Zurich, which were very friendly and were showing genuine interest in my work, although they were actually involved in other projects. Special thanks goes to Peter Robman and Andries van der Schaarf. I truly enjoyed working with them and learned a lot in the process. I'm also thankful to Zurich's graduate students Nicole Werner,Jurgen Krosenberg, and finally an english-speaking "intruder" from Cornell - Matt Siegler. You all made my Zurich visits more interesting and provided the sometimes needed distraction from the ntuples and .tex files.

I would like to thank my brother - Alexander, who was also responsible for my involvement with the E865. He was always there to help me with any problems I had, even beyond the scope of this analysis. He is a brother I'm very lucky to have.

I would also like to thank my first advisor Gennadij Fedotovich, with whom I made my first mistakes and made the first steps into the scientific community.

Finally, I would like to thank my parents and my grandmother whom I love very much. I surely could not have done anything without you.

Thanks also goes to my friends Kostya, Leonya, Shurick and Vasilij who despite the vast geographical separation between us remained true friends and always welcomed me.

And I would like to thank my wife Elena for putting up with me and for supporting me till the very end.