Investigating Lepton Universality
via a Measurement of the
Positronic Pion Decay Branching Ratio

Anthony Palladino, Jr.

Charlottesville, Virginia, USA

B.S., University of Connecticut, 2004

M.A., University of Virginia, 2007

A Dissertation presented to the Graduate Faculty

of the University of Virginia

in Candidacy for the Degree of

Doctor of Philosophy

Department of Physics

University of Virginia

May 2012

_____

Dinko Počanić, Dissertation Advisor

_____

Kent Paschke, Chair of Advisory Committee
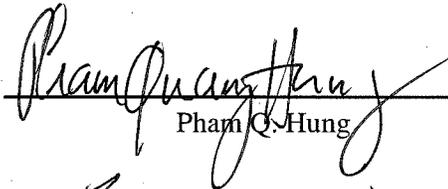
_____

Pham Q. Hung
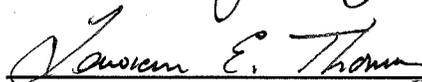
_____

Lawrence E. Thomas

Approvals

This dissertation is presented in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in the Department of Physics

by

(signature)

Anthony Palladino, Jr.

This dissertation has been read and approved by the Examining Committee:

(signature)	_(signature)_
Dinko Počanić,	Kent Paschke,
Dissertation Advisor,	Chair of Advisory Committee,
Department of Physics	Department of Physics
(signature)	_(signature)_
Phạm Q. Hưng,	Lawrence E. Thomas,
Department of Physics	Department of Mathematics
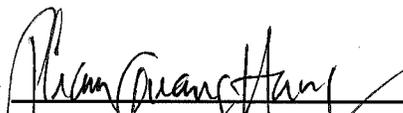
Accepted for the College of Arts and Sciences on:

26 April 2012

Date

Abstract

The PEN experiment at the Paul Scherrer Institute in Switzerland aims to measure the positronic pion decay branching ratio to an unprecedented relative precision of 0.05%. The measurement tests the existence of lepton universality and puts constraints on several theories beyond the Standard Model of particle physics. This dissertation will first describe the experimental configuration and data collection during the PEN experiment. Focus is placed on the development of data analysis tools including calibrations, event reconstruction algorithms, and a maximum likelihood analysis framework designed specifically for this experiment. The 8×10^5 $\pi^+ \to e^+ \nu_e$ events observed in 2008 were used in this study to obtain $R_{\pi_{e2}} = [1.112 \pm 0.002 \, (\text{stat.})] \times 10^{-4}$, where the central value is still intentionally blinded with an unknown multiplicative random number. Using only a small fraction of the PEN data, our statistical uncertainty in $R_{\pi_{e2}}$ is already smaller than the combined (statistical and systematic) uncertainty in the experimental world average. Including the estimated 20×10^6 additional $\pi^+ \to e^+ \nu_e$ events from 2009 and 2010 will further reduce the uncertainty on this measurement.

At times our own light goes out and is rekindled by a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us.

Albert Schweitzer

Acknowledgements

It is only through the scientific progress made by countless individuals that this work became feasible. I have made considerable effort to cite their work whenever possible. This section is reserved to acknowledge and express my sincere gratitude to those people who have been intimately involved with my work, whose contribution cannot be recognized through the citation of a published document, and without whom this work in its present form would not have been possible.

I would like to thank my thesis advisor, Dinko Počanić, for his continuous support and invaluable guidance. His knowledge of the field and dedication to his work have made a profound impact on my education. Much appreciation goes to Andries van der Schaaf for his contribution to the PEN experiment and his critical examination of my work. I would like to thank my fellow PEN Collaboration members at the University of Virginia: Emil Frlež, Maxim Bychkov, Pete Alonzi, and Martin Lehman; and my friends and colleagues: Osvaldo Gonzalez-Hernandez, Hussain Zaidi, and Luke Langsjoen; for endless discussions, several of which have influenced this work.

Many thanks go to Willi Bertl and Andries van der Schaaf for their support during the years I spent at the Paul Scherrer Institute and Universität Zürich.

Immense gratitude goes to my parents, Sara and Tony Palladino, and to my sisters, Lauren and Danielle, who have always supported my goals and provided me with the courage to pursue my dreams.

Contents

Chapter 1

Introduction

1.1 The Standard Model of Particle Physics

It has been established that the Standard Model (SM) is the authoritative description of the fundamental laws of nature. The theory is so powerful that nearly all of its predictions have been experimentally verified. It is an ever changing model that adapts to new measurements in order to adhere to the physical world.

Despite its success, the Standard Model is not complete. There exists mathematical structure in the model that is not fully understood, and there are physical phenomena which the theory cannot predict. Theorists are hard at work developing extensions Beyond the Standard Model (BSM) which attempt to explain discordant observations. This dissertation describes an experiment conducted to both test SM predictions, such as lepton universality, and narrow down the list of possible BSM theories by giving, for example, limits on the masses of hypothetical particles.

According to the current SM, all matter can be built from fundamental spin-1/2 (fermion) particles called quarks and leptons (Table 1.1(a)). The interactions between these fermions are described

Table 1.1: *Periodic Tables of Elementary Particles.* (a) *shows the fermions organized into the three generations of matter,* (b) *shows the force mediating gauge bosons, and* (c) *shows the yet to be discovered Higgs boson which gives rise to the masses of all massive particles.*

(a)

Flavor Generation

	I	II	III
Quarks	u (up)	c (charm)	t (top)
	d (down)	s (strange)	b (bottom)
Leptons	e (electron)	μ (muon)	τ (tau)
	ν_e (electron neutrino)	ν_μ (muon neutrino)	ν_τ (tau neutrino)

(b)

Force	Mediator
Strong	G (gluon)
Electromagnetic	γ (photon)
Weak	$W^\pm$, Z^0 (charged,neutral weak boson)
Gravitational	g (graviton)

(c)

Mass Generator

H (Higgs boson)

by the exchange of characteristic integer spin (boson) "force mediator" particles. There are only four known fundamental interactions that occur in nature: electromagnetic, weak nuclear, strong nuclear, and gravitational. The gauge bosons corresponding to each of these forces are listed in Table 1.1(b). The SM does not describe the gravitational interaction, which fortunately is negligible on the subatomic scale.

The SM also explains the origin of particle masses via the yet to be discovered Higgs boson. Extensive searches for the elusive Higgs particle are underway at particle physics laboratories worldwide.

1.2 The Pion

In Hideki Yukawa's 1935 paper [71] he suggested that experimentalists should search for a yet undiscovered particle that mediated the nuclear force. Using the range of nuclear forces, he predicted the mass of the new particle to be somewhere between the light electron and the heavy nucleon. It was supposed to be the force carrier, being emitted and absorbed by protons and neutrons, thereby explaining the stability of the atomic nucleus by accounting for the short-ranged attraction between its constituents. Shortly thereafter, in 1937, Carl D. Anderson and his student Seth H. Neddermeyer found a particle in cosmic radiation with the mass and disintegration properties that one would expect for the Yukawa particle [53, 54]. Their discovery was confirmed almost immediately by J.C. Street and E.C. Stevenson who gave a mass value of 130 electron masses with a 25% uncertainty [69]. Four months later, Nishina, Takeuchi, and Ichimiya published evidence of a positively charged particle with a mass from 180 to 260 electron masses [55]. The following year, Nishina et al. discovered a negatively charged particle with mass 170 ± 9 m_e and reduced the uncertainty on their estimate of the mass of the positively charged particle to 180 ± 20 m_e [56]. This intermediate-massed particle was then referred to as a "mesotron" and later shortened to meson, originating from the Greek word for intermediate, "mesos".

Figure 1.1: *A photograph of Hideki Yukawa (left) and César Lattes.*

If a meson is to explain the nuclear interaction then it should be readily absorbed by protons and nuetrons, and therefore by matter. In 1947, nearly a decade later, three Italian physicists, Marcello Conversi, Ettore Pancini and Oreste Piccioni, observed that the mesons found in cosmic radiation often passed by several hundred atomic nuclei without interacting [24]. Enrico Fermi, Edward Teller, and Victor Weisskopf, introducing the symbol μ for mesotron, explained

that the decay time of these negative particles in matter was twelve orders of magnitude longer than the time it should take the nuclear forces to capture Yukawa's particle [35]. These properties do not follow Yukawa's prediction.

Meanwhile in 1947, a group lead by Cecil F. Powell was studying tracks left by charged particles in photographic nuclear emulsion plates placed at the top of *Pic du Midi* in the French Pyrenees. Giuseppe Occhialini and César Lattes discovered that two distinct types of mesons were present in the plates. One meson would gradually slow down and stop, and at the end of its track a new meson appeared [58, 43]. Lattes travelled to the Bolivian Andes and placed several nuclear emulsion plates at the top of Mount Chacaltaya, 5500 m above sea level. Analysis of these plates revealed that one of the mesons was 30–40% heavier than the other one [44]. The primary meson (now called π meson, or pion) was heavier and disintigrated into a secondary meson (μ meson, or muon)[1]. It turns out that the muon was the known meson from the experiments of Anderson and Neddermeyer. This new, previously undiscovered, π meson, however, was shown to readily interact with nuclei and had the characteristic properties according to Yukawa's theory.

We have come a long way in our understanding of the pion since then. We now know that the pion is a meson consisting of two quarks from the first flavor generation. The pion can be charged, $\pi^{\pm}$, or neutral, π^{0}. For the remainder of this dissertation, we will restrict ourselves to the π^{+} due to experimental practicalities. Negatively charged pions are absorbed by nearby atomic nuclei with almost 100% probability due to the attractive nature of their opposite electric charges, and are therefore more difficult to deal with. On the contrary, positively charged pions are essentially shielded from nuclei by the atom's electrons, and such complications are reduced.

The π^{+}, composed of one up (u) and one anti-down $(\bar{\text{d}})$ quark, has a lifetime of 26.033 ns and

---

[1]The muon is now classified as a lepton, not a meson.

Table 1.2: *Decay modes and decay rates of the π^+ [33].*

| Decay Mode	Decay Rate
$\pi^+ \to \mu^+ \nu_\mu$	$\Gamma_1 = 0.9998770(4)$
$\pi^+ \to \mu^+ \nu_\mu \gamma$	$\Gamma_2 = 2.00(25) \times 10^{-4}$
$\pi^+ \to e^+ \nu_e$	$\Gamma_3 = 1.230(4) \times 10^{-4}$
$\pi^+ \to e^+ \nu_e \gamma$	$\Gamma_4 = 7.39(5) \times 10^{-7}$
$\pi^+ \to \pi^0 e^+ \nu_e$	$\Gamma_5 = 1.036(6) \times 10^{-8}$
$\pi^+ \to e^+ \nu_e e^+ e^-$	$\Gamma_6 = 3.2(5) \times 10^{-9}$

a mass of 139.57 MeV [33]. The pion decays when the u and $\bar{\text{d}}$ quarks annihilate, producing a

W^+ boson. This weak gauge boson then procedes to create particles along one of several allowed

modes. This process is called the weak decay mode of the pion (Table 1.2). Of particular interest

to this thesis are the 2-body decay modes, commonly denoted $\pi_{\ell 2}$ (with decay rates Γ_1 and Γ_3) in

which the π^+ produces a lepton-antilepton pair, where $\ell = e^+, \mu^+$, as shown in Figure 1.2. For the

experiment described herein, the $\pi^+ \to \pi^0 e^+ \nu_e$ decay mode is possible but very rare. Since Γ_5 is

$\sim 1/5$ of the desired uncertainty in our measurement, we'll ignore this decay mode in the initial PEN

data analysis.
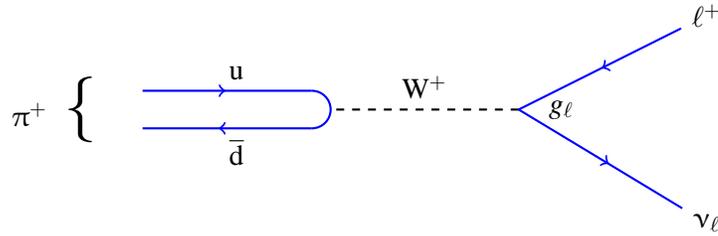
Figure 1.2: *The lowest level (tree) Feynman diagram for the $\pi^+ \to \ell^+ \nu_\ell$ decay modes.*

1.3 Review of $\pi^+ \to e^+ \nu_e$ Experiments

The $\pi^+ \to e^+ \nu_e$ decay mode was first discovered at CERN in 1958 [32, 40]. A branching ra-

tio measurement for this mode wasn't made until 1960 by Anderson et al., who measured $R_{\pi_{e2}} =$

$(1.21 \pm 0.07) \times 10^{-4}$ [7]. Four years later DiCapua et al. pushed the uncertainty down to $R_{\pi_{e2}} = (1.273 \pm 0.028) \times 10^{-4}$ [28].
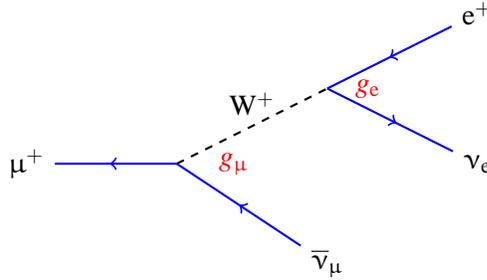
Figure 1.3: *Feynman diagram for the Michel* $\mu \to e\nu\overline{\nu}$ *decay, named after Louis Michel, who first described the decay positron's energy spectrum in detail.*

In practice, a typical measurement involves measuring the decay time and energy of the positron emitted from either the $\pi \to e$ decay or the $\pi \to \mu \to e$ sequential decay chain. The sequential Michel decay is shown in Figure 1.3, where the μ^+ emits a W^+ and is transformed into a muon antineutrino. The W^+ then creates a lepton-antilepton pair. The positron from the two-body decay, $\pi \to e\nu$, is mono-energetic at $m_\pi/2 = 69.8$ MeV. In contrast, the positron from the sequence $\pi \to \mu\nu$ followed by the 3-body decay $\mu \to e\nu\overline{\nu}$ has a continuous energy spectrum from 0 to $m_\mu/2 = 52.8$ MeV. This characteristic energy spectrum is often called the Michel spectrum, named after Louis Michel who first described and parameterized the $\mu \to e\nu\overline{\nu}$ decay in detail [50].

The first precise measurement was made in 1983 when Bryman et al. employed a NaI(Ti) crystal calorimeter sensitive to both charged particles and photons [15]. They performed two seperate analyses on the data. The first method was the so-called "2-bin" method developed by DiCapua et al. for the analysis of the 1964 experiment. The energy spectra for the positrons were collected in two identical time intervals, one starting at t_i and the other at $t_i + t_s$ after the pion stop time. Since t_s is long compared to the pion lifetime, the second interval contained essentially only positrons from

the $\pi \to \mu \to$ e chain. The branching ratio was then calculated as

$$
\begin{aligned}
R_{\pi_{e2}} &= \frac{\lambda_\mu}{\lambda_\pi - \lambda_\mu} \frac{N_{\pi \to e}\left\{1 - \exp\left[-\left(\lambda_\pi - \lambda_\mu\right)t_s\right]\right\}}{N^{(2)}_{\pi \to \mu \to e}\exp\left(\lambda_\mu t_s\right) - N^{(1)}_{\pi \to \mu \to e}} \\
&= (1.218 \pm 0.014) \times 10^{-4}
\end{aligned}
\tag{1.1}
$$

where λ_π and λ_μ are the pion and muon decay rates defined as $\lambda_\pi = 1/\tau_\pi$ and $\lambda_\mu = 1/\tau_\mu$.

Their second method was a simple fit to the decay time distribution of the positrons, obtaining the amplitudes of the $\pi \to$ e and $\pi \to \mu \to$ e components. The second method resulted in $R_{\pi_{e2}} = A_{\pi \to e}/A_{\pi \to \mu \to e} = (1.219 \pm 0.014) \times 10^{-4}$.

A more refined measurment came in 1992 by the same collaboration at TRIUMF [11]. They measured $R_{\pi_{e2}}$ by a simultaneous fit to the positron decay time spectra for events with positron energy above and below 56.4 MeV. The energies were recorded with the "TINA" NaI(Ti) 460 mm diameter $\times$ 510 mm long cylindrical crystal calorimeter. Several multiplicative systematic corrections of order $\sim$1% were applied to obtain

$$
R_{\pi_{e2}} = [1.2265 \pm 0.0034(\text{stat.}) \pm 0.0044(\text{syst.})] \times 10^{-4} \ .
\tag{1.2}
$$

The third and most recent precise measurement was performed at the Paul Scherrer Institute and published in 1993. Czapek et al. [25] considered only events with positron decay time between 7.5 ns and 200 ns after the pion stop time. They considered the total energy spectrum, i.e., the energy deposited in the target by the pion, possible muon, and the positron, plus the total energy deposited in the BGO calorimeter. The number of $\pi \to$ e events was determined using three windows in E_{Total}. The majority of $\pi \to$ e events were contained in the region $83.5 \le E_{\text{Total}} \le 101$ MeV. The amount of $\pi \to$ e events below 83.5 MeV was determined using a GEANT3 simulation. The region above
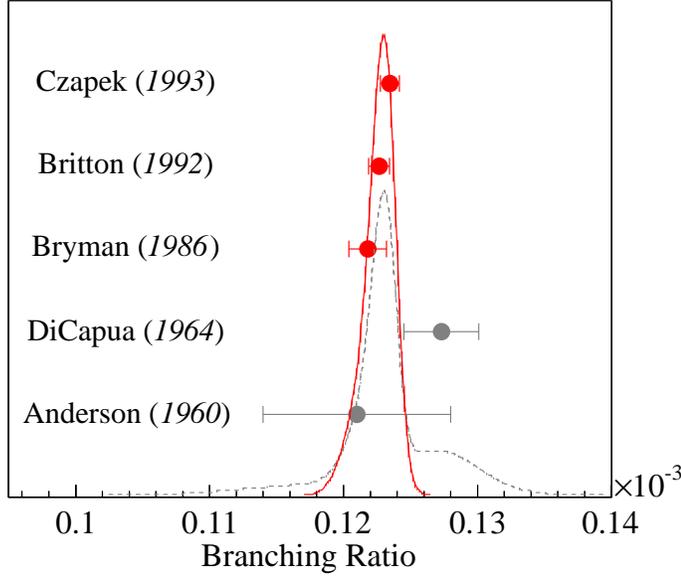
Figure 1.4: *Historical representation of $R_{\pi_{e2}}$ measurements. The solid curve is the probability distribution of the branching ratio using the three most recent measurements, as is used in the current Particle Data Group estimation of the branching ratio [33]. The dashed curve gives the probability distribution using all five measurements.*

101 MeV essentially contained only events for which the pion underwent a strong interaction in the target. The total number of pion decays was found using a normalization trigger inside the same 7.5 to 200 ns time window. They obtained the value

$$R_{\pi_{e2}} = [1.2346 \pm 0.0035(\text{stat.}) \pm 0.0036(\text{syst.})] \times 10^{-4} \ . \tag{1.3}$$

The Particle Data Group [33] combined these three most recent measurements of $R_{\pi_{e2}}$ to give the world average

$$R_{\pi_{e2}}^{\text{PDG}} = (1.230 \pm 0.004) \times 10^{-4} \ , \tag{1.4}$$

which has remained unchanged for nearly two decades.

1.4 Review of $\pi^+ \rightarrow e^+ \nu_e$ Theory

The theory behind the π_{e2} decay is described in the Standard Model with extraordinary precision. This section will examine the theoretical description in detail. More information can be found in

the report [12] and the references quoted within.

The differential decay rate for the $\pi_{\ell 2}$ decay (with $\ell = \text{e}$ or μ) can be given by,

$$d\Gamma = \frac{1}{2m_\pi} |M|^2 \frac{1}{E_\ell E_\nu} \frac{d^3 p_\ell}{(2\pi)^3} \frac{d^3 p_\nu}{(2\pi)^3} \left(2\pi^4\right) \delta^4 \left(q - p_\ell - p_\nu\right) \ , \tag{1.5}$$

where m_π is the pion mass, q, p_ℓ, and p_ν are the four-momenta of the pion, lepton, and neutrino, respectively, and M is the matrix element in the $(V - A)$ theory [49],

$$M = \frac{iG_F}{\sqrt{2}} \langle 0| \left\{V_\lambda(0) - A_\lambda(0)\right\} |\pi\rangle \bar{u}_\ell \gamma_\lambda \left(1 - \gamma_5\right) v_\nu \ , \tag{1.6}$$

where $G_F = 1.16639 \times 10^{-11} \text{ MeV}^{-2}$ is the Fermi coupling constant. Since the pion is a pseudoscalar particle and due to the lack of available axial-vector operators, we can determine on the grounds of Lorentz invariance that $\langle 0|V_\lambda(0)|\pi\rangle = 0$. Similarly, we find that $\langle 0|A_\lambda(0)|\pi\rangle = if_\pi\left(q^2\right)q_\lambda$, where $f_\pi\left(q^2\right) = f_\pi\left(-m_\pi^2\right) \equiv f_\pi = 130.7 \text{ MeV}$ is the pion decay constant.

It follows that when using the Dirac equation $u\left(\not{p} - m_\ell\right) = 0$, where $\not{p} \equiv \gamma_\lambda p_\lambda$, we obtain the matrix element

$$M = \frac{-G_F f_\pi}{\sqrt{2}} m_\ell \bar{u}_\ell \left(1 - \gamma_5\right) v_\nu \ . \tag{1.7}$$

After summing over final spin states, the differential decay rate becomes,

$$d\Gamma = \frac{G_F^2 f_\pi^2 m_\ell^2}{2m_\pi (2\pi)^2} \frac{p_\ell p_\nu}{E_\ell E_\nu} d^3 p_\ell d^3 p_\nu \delta^4 \left(q - p_\ell - p_\nu\right) \ , \tag{1.8}$$

and upon integrating over lepton momenta the total decay rate for the $\pi \rightarrow \ell \nu_\ell$ decay becomes,

$$\Gamma = \frac{G_F^2 f_\pi^2 m_\ell^2}{8\pi m_\pi^3} \left(m_\pi^2 - m_\ell^2\right)^2 \ . \tag{1.9}$$

This equation clearly demonstrates that the decay rate is proportional to the square of the lepton mass. This proportionality is the physical consequence of the $(1 - \gamma_5)$ term in (1.7), which is known as the helicity projection operator for massless leptons. This operator allows only left-handed massless particles and right-handed anti-particles. If $m_\ell = 0$, angular momentum conservation would prohibit the $\pi \to \ell \nu_\ell$ decay channels for $m_\nu = 0$. Since the positron and muon are massive, both positive and negative helicity states are mixed by an amount proportional to their mass, leading to non-zero decay rates.
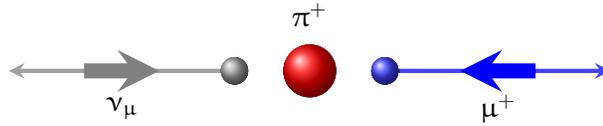
Figure 1.5: *A π^+ decaying at rest. The thin arrows represent the particle momenta, while the thick arrows represent their helicity.*

To put this situation into perspective, consider the π^+ decaying at rest as depicted in Figure 1.5. The 106.7 MeV muon wants to be right-handed; the 0.5 MeV *really* wants to be right-handed; but the $< 10^{-6}$ MeV neutrino essentially forces them to both be left-handed. Akin to polarization, we may define,

$$\begin{aligned}
\text{``Helicity Conservation''} &: \quad \frac{1}{2} + \frac{1}{2}\frac{v}{c} \\
\text{``Helicity Violation''} &: \quad \frac{1}{2} - \frac{1}{2}\frac{v}{c} \quad .
\end{aligned} \tag{1.10}$$

For $v = c$ we have zero probability of the particle "violating" helicity. Now for a given energy, the positron will have a much greater velocity than the muon due to their relative masses. Taking the ratio of the likelihood of each particle "violating" helicity we find,

$$\frac{\text{Left-handed } e^+}{\text{Left-handed } \mu^+} \approx \frac{\frac{1}{2} + \frac{1}{2}\frac{v_e}{c}}{\frac{1}{2} + \frac{1}{2}\frac{v_\mu}{c}} \approx 3.2 \times 10^{-5} \tag{1.11}$$

This severe reduction in what would otherwise be the main decay channel for the pion is known as *helicity suppression* and is of key importance to the subject of this thesis.

Considering phase-space alone, the positronic decay channel, $\pi \to e\nu_e$, would be ~ 3.3 times more likely than the muonic decay mode. Combining these two rough calculations gives an order-of-magnitude estimate of the π_{e2} branching ratio to be,

$$\frac{\Gamma(\pi \to e\nu_e)}{\Gamma(\pi \to \mu\nu_\mu)} = 3.3 \times \left(3.2 \times 10^{-5}\right) \sim 10^{-4} \ , \tag{1.12}$$

where a branching ratio is defined to be a ratio of decay rates.

Using (1.9) the branching ratio can be calculated,

$$R_{\pi_{e2}}^{\text{SM}} = \frac{\Gamma(\pi \to e\nu_e)}{\Gamma(\pi \to \mu\nu_\mu)} = \frac{f_\pi^{e2} m_e^2 \left(m_\pi^2 - m_e^2\right)^2}{f_\pi^{\mu 2} m_\mu^2 \left(m_\pi^2 - m_\mu^2\right)^2} = \frac{f_\pi^{e2}}{f_\pi^{\mu 2}} \left(1.283 \times 10^{-4}\right) \ . \tag{1.13}$$

The principle of electron-muon universality in pion decay holds under the assumptions that the basic interaction current is of the $V - A$ type if $f_\pi^e = f_\pi^\mu$.

Using the universal $V - A$ theory, Berman [8] and Kinoshita [42] showed that the branching ratio (1.13) is incomplete and requires substantial modifications due to radiative corrections. These corrections depend on m_ℓ. Radiative corrections originating from the emission of real photons, Inner Bremsstrahlung (IB), are shown in Figure 1.8. Corrections to the total $\pi \to \ell\nu_\ell$ decay rate due to virtual emission and reabsorption of photons are shown in Figure 1.7.

The direction of a charged particle changes when it scatters, and is therefore is accelerated. As a consequence it radiates. Bremsstrahlung[2] is the term used to describe radiation produced by charged particles passing through a medium. Figure 1.6 shows that the electron cannot radiate a

---

[2]Bremsstrahlung is a German word meaning braking radiation
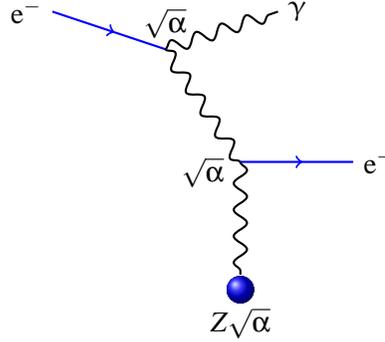
Figure 1.6: *The Feynman diagram for the Bremsstrahlung process.*

photon without exchanging a soft photon with a nearby nucleus. To be more specific, external bremsstrahlung is used to describe radiation caused by decelerations when passing through a field of atomic nuclei. Internal bremsstrahlung (IB) is used for the radiation of non-virtual quanta, i.e., photons or gluons, by particles participating in an interaction.

Including the IB, virtual corrections, and radiative corrections into (1.13), and assuming electron-muon universality, we get

$$R_{e/\mu}^{SM,RC} = \frac{\Gamma(\pi \to e\nu_e(\gamma))}{\Gamma(\pi \to \mu\nu_\mu(\gamma))} = R_{e/\mu}^{SM}(1+\delta)(1+\varepsilon) = 1.233 \times 10^{-4}\,, \qquad (1.14)$$

where the larger of the corrections is $\delta = -(3\alpha/\pi)\ln(m_\mu/m_e)$ and $\varepsilon = -0.92(\alpha/\pi)$. Another more transparent representation of the branching ratio can be given by,

$$R_{\pi_{e2}}^{SM,RC} = \left(\frac{g_e}{g_\mu}\right)^2 \left(\frac{m_e}{m_\mu}\right)^2 \frac{\left(1-m_e^2/m_\mu^2\right)^2}{\left(1-m_\mu^2/m_\pi^2\right)^2}(1+\delta R)\,, \qquad (1.15)$$

where all radiative corrections have been combined into δR, helicity-suppression is apparent in the $(m_e/m_\mu)^2$ term, and electron-muon universality would hold when $g_e/g_\mu = 1$.

Recent standard model calculations have been published by Marciano and Sirlin [48], Finke-
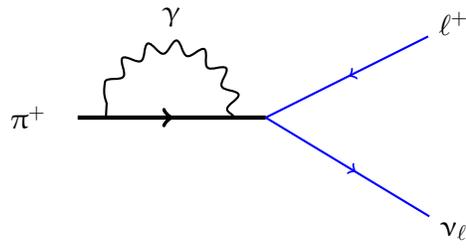
(a)
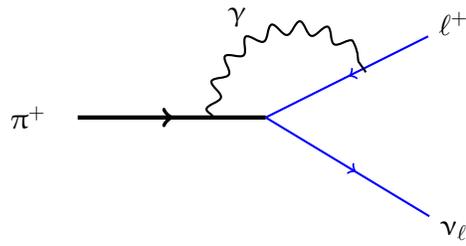
(b)
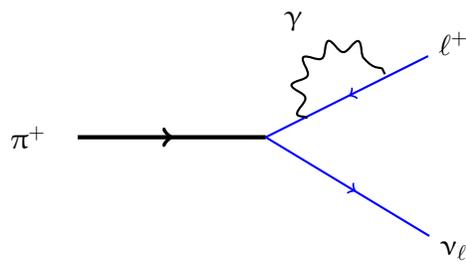
(c)

Figure 1.7: *The radiative $\pi^+ \to \ell^+ \nu_\ell$ decay processes.*
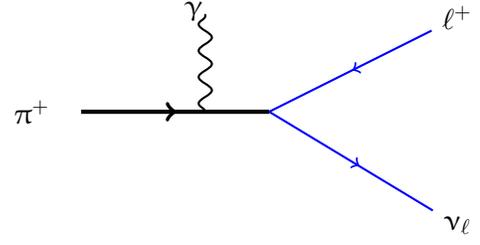
(a)
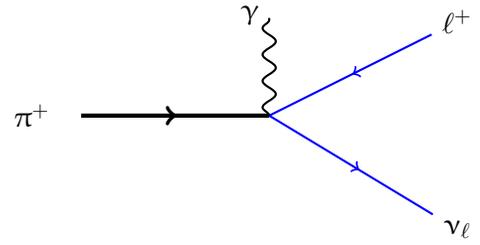
(b)
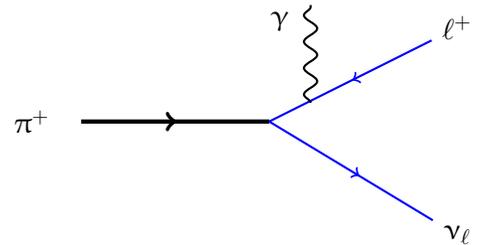
(c)

Figure 1.8: *The Inner Bremsstrahlung diagrams for the $\pi^+ \to \ell^+ \nu_\ell$ decay.*

meier [36], and Cirigliano and Rosel [22],

$$R_{\pi_{e2}}^{\text{SM,RC}} = \begin{cases} (1.2352 \pm 0.0005) \times 10^{-4} & \text{Ref. [48]} \\ (1.2354 \pm 0.0002) \times 10^{-4} & \text{Ref. [36]} \\ (1.2352 \pm 0.0001) \times 10^{-4} & \text{Ref. [22]} \end{cases} . \qquad (1.16)$$

Chapter 2

Motivation for the PEN Experiment

In comparing (1.16) and (1.4) in the previous chapter we see that the current experimental results for $R_{\pi e2}$ lag behind Standard Model calculations by an order of magnitude. This chapter will discuss several reasons motivating the PEN experiment.

2.1   Lepton Universality

The Standard Model lepton coupling constant, g_ℓ (where $\ell = $ e,μ,τ), between the weak boson and the leptons (e.g., in $\pi \rightarrow \ell \nu$ decay, Figure 1.2) is taken to be equivalent across all lepton flavors. This hypothesis, called *lepton universality* [27], is generally accepted, although there are no compelling reasons for it.

Loinaz et al. [46] have parameterized possible flavor non-universal suppressions of g_ℓ in $W\ell\nu_\ell$ coupling as,

$$ g_\ell \rightarrow g_\ell \left(1 - \frac{\varepsilon_\ell}{2} \right) \ . $$
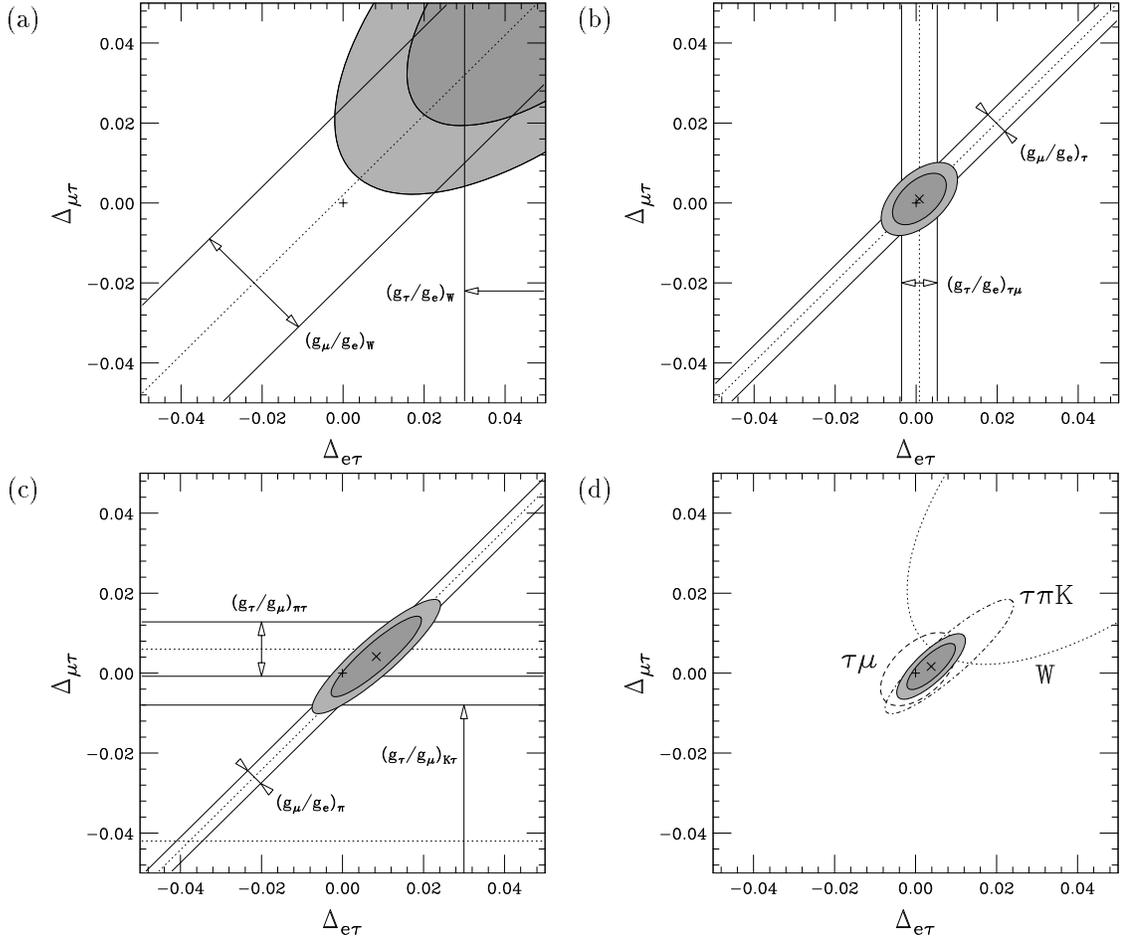
Figure 2.1: *Experimental constraints on $\Delta_{e\tau}$ and $\Delta_{\mu\tau}$ (from Loinaz et al. [46]), derived from* (a) W *decay,* (b) τ *decay,* (c) π *and* K *decay. The combination of the limits from the aforementioned decays is shown in* (d). *Improved accuracy on the limits of g_μ/g_e from π decay will reduce the allowed region to a narrower strip in* (c) *and* (d).

The linear combinations of the ε_ℓ constrained by W, τ, π, and K decays are:

$$\frac{g_\mu}{g_e} = 1 + \frac{\varepsilon_e - \varepsilon_\mu}{2} \ , \quad \frac{g_\tau}{g_\mu} = 1 + \frac{\varepsilon_\mu - \varepsilon_\tau}{2} \ , \text{ and } \ \frac{g_\tau}{g_e} = 1 + \frac{\varepsilon_e - \varepsilon_\tau}{2} \ .$$

Experimental constraints can be evaluated on $\Delta_{e\mu} \equiv \varepsilon_e - \varepsilon_\mu$, $\Delta_{\mu\tau} \equiv \varepsilon_\mu - \varepsilon_\tau$, and $\Delta_{e\tau} \equiv \varepsilon_e - \varepsilon_\tau$. Two of the three are independent, and Loinaz et al. have chosen to consider the latter two. Their plots demonstrating these experimental constraints are reproduced in Figure 2.1.

Repeating the ratio of decay rates (1.15) for convenience,

$$R^{\text{SM,RC}}_{\pi_{e2}/\pi_{\mu2}} = \left(\frac{g_e}{g_\mu}\right)^2 \left(\frac{m_e}{m_\mu}\right)^2 \frac{\left(1 - m_e^2/m_\mu^2\right)^2}{\left(1 - m_\mu^2/m_\pi^2\right)^2} \left(1 + \delta R_{\pi_{e2}/\pi_{\mu2}}\right) \, , \tag{2.1}$$

and the analogous ratio of the relevant the τ and π decay rates,

$$R^{\text{SM,RC}}_{\tau_{e2}/\pi_{\mu2}} = \left(\frac{g_\tau}{g_\mu}\right)^2 \frac{m_\tau^3}{2m_\mu^2 m_\pi} \frac{\left(1 - m_\pi^2/m_\tau^2\right)^2}{\left(1 - m_\mu^2/m_\pi^2\right)^2} \left(1 + \delta R_{\tau_{e2}/\pi_{\mu2}}\right) \, , \tag{2.2}$$

shows how the measurements of these branching ratios directly constrain the ratios of coupling constants. Using the above equations and the available experimental data one can evaluate [46],

$$\left(\frac{g_e}{g_\mu}\right)_\pi = 1.0021 \pm 0.0016 \quad \text{and} \quad \left(\frac{g_\tau}{g_\mu}\right)_{\pi\tau} = 1.0030 \pm 0.0034 \, .$$

2.2   Charged Higgs Boson

Another illustration of the reach of this precise $R_{\pi_{e2}}$ measurement is the bound we can place on the hypothetical charged physical Higgs boson. To give a value, let's consider the Lagrangian for four-fermion vector (V), axial-vector (A), scalar (S), and pseudoscalar (P), interactions that might arise from *new physics* (NP) at a scale Λ [16]:

$$\mathcal{L}_{\text{NP}} = \left[\pm \frac{\pi}{2\Lambda_V^2} \bar{u}\gamma_\mu d \pm \frac{\pi}{2\Lambda_A^2} \bar{u}\gamma_\mu\gamma_5 d\right] \bar{e}\gamma^\mu (1 - \gamma_5)\nu$$

$$+ \left[\pm \frac{\pi}{2\Lambda_S^2} \bar{u}d \pm \frac{\pi}{2\Lambda_P^2} \bar{u}\gamma_5 d\right] \bar{e}(1 - \gamma_5)\nu \, . \tag{2.3}$$

Scalar and vector interactions are already constrained to $\Lambda_S \geq 10$ TeV and $\Lambda_V \geq 20$ TeV by CKM

unitarity tests and a measurement [52] of superallowed nuclear β decay. The experiment described

in this dissertation, assuming $\Delta R_{\pi_{e2}}/R_{\pi_{e2}} \sim 0.1$ %, can probe scales of $\Lambda_P \leq 1,000$ TeV and $\Lambda_A \leq$

20 TeV. A measurement of $R_{\pi_{e2}}$ does not directly probe scalar interactions, but indirect sensitivity[1]

to $\Lambda_S \leq 60$ TeV could be obtained through pseudoscalar interactions induced by loop effects from

new scalar interactions [26, 19].

Take, for example, the charged physical Higgs boson, with couplings $\frac{g}{2\sqrt{2}}\lambda_{ud}$ to the $\bar{u}\gamma d$ pseu-

doscalar current and $\frac{g}{2\sqrt{2}}\lambda_{\ell\nu}$ to $\bar{\ell}(1-\gamma_5)\nu_\ell$, where g is the SU(2)$_L$ gauge coupling, ℓ can be e or μ,

and λ represents chirality-breaking suppression factors. The branching ratio for these new physics

scenarios can be written as

$$R_{\pi_{e2}}^{\mathrm{NP}} = R_{\pi_{e2}} \left[1 - \frac{2m_\pi^2}{m_{e(m_u+m_d)}} \frac{m_W^2}{m_{H^\pm}^2} \lambda_{ud}\lambda_{\mu\nu} \left(\frac{\lambda_{e\nu}}{\lambda_{\mu\nu}} - \frac{m_e}{m_\mu} \right) \right] \ . \tag{2.4}$$

A ± 0.1 % measurement of $R_{\pi_{e2}}$ therefore probes [16]

$$m_{H^\pm} \simeq 200 \text{ TeV} \times \sqrt{\lambda_{ud}}\sqrt{\lambda_{\mu\nu}} \left(\frac{\lambda_{e\nu}}{\lambda_{\mu\nu}} - \frac{m_e}{m_\mu} \right)^{1/2} \ . \tag{2.5}$$

If e-μ universality extends into the enlarged scalar sector, i.e., $\frac{\lambda_{e\nu}}{\lambda_{\mu\nu}} = \frac{m_e}{m_\mu}$, as in the minimal two-

Higgs doublet model, $R_{\pi_{e2}}$ will not be sensitive to $m_{H^\pm}$. In more general multi-Higgs models,

however, such a relationship is not required. In particular, $m_{H^\pm} \simeq 400$ GeV is probed for loop-

induced charged-Higgs couplings where $\lambda_{e\nu} \simeq \lambda_{\mu\nu} \simeq \lambda_{ud} \simeq \alpha/\pi$ when $\Delta R_{\pi_{e2}}/R_{\pi_{e2}}$ reaches ± 0.1 %.

---

[1]A sensitivity of $\Lambda_S \leq 60$ TeV is well beyond the capability of nuclear β decay measurements.

2.3   Massive Neutrinos

Shrock outlines a method to search for the existence of massive neutrinos using either $\pi \to e\nu$ or $\pi \to \mu\nu$ decays [67]. Britton et al.[10] state that the π_{e2} branching ratio may increase due to a relaxation of helicity suppresion. An improved measurement of the branching ratio will further confine the neutrino mixing parameter, U_{ei}. The implications of $R_{\pi_{e2}}$ on massive neutrinos is discussed further by Bryman et al.[13] and Britton et al.[9].

Also, Bryman and Numao [14] investigate limits on the existence of massive neutrinos using the plentiful $\pi \to \mu\nu$ decays.

2.4   Extra-Dimension Models

A precise measurement of $R_{\pi_{e2}}$ will constrain hypothetical extra-dimension models with strong gravity at the TeV scale [19].

2.5   Ratio of Pseudoscalar to Vector Coupling

If the $\pi \to e\nu$ decay were dominated by a pseudoscalar coupling, then the helicity suppression of the decay would vanish, and the branching ratio would be $R_{\pi_{e2}} \approx 5.5$. A difference between the Standard Model description of the decay and the best experimental results can provide an estimate of the residual pseudoscalar coupling.

Subtracting the Standard Model component from the experimental results gives bounds on C_P, the ratio of pseudoscalar to vector coupling strengths. At the 2σ level we obtain,

$$-7 \times 10^{-3} \leq \frac{C_P}{f_\pi m_e} \leq 2.5 \times 10^{-3} \, , \qquad (2.6)$$

where f_π is the pion decay constant and m_e is the electron mass. Using the model independent technique outlined by Bryman et al.[17], limits on masses of hypothetical particles for the maximal coupling can be obtained.

2.5.1 Hypothetical Leptoquark and Supersymmetric Particles

Leptoquarks are particles whose interaction vertices simultaneously involve both leptons and quarks. They appear in various extensions of the Standard Model, including technicolor models [34] and Grand Unified Theories [64, 61, 39]. In supersymmetric (SUSY) models, for each Standard Model fermion, there is a corresponding boson, and vice versa. These superpartners have the same quantum numbers as the original particles, except for spin [19].

We calculated new lower limits on the masses of pseudoscalar and vector leptoquarks and hypothetical charged Higgs particles. These calculations use the proposed uncertainty in the measurement of the $\pi^+ \to e^+ \nu_e$ branching ratio of the PEN experiment.

Neglecting radiative effects, the theoretical branching ratio according to the standard model (via the $V-A$ weak interaction) is,

$$R^0_{\pi_{e2}} = \frac{\Gamma(\pi^+ \to e^+ \nu_\mu)}{\Gamma(\pi^+ \to \mu^+ \nu_\mu)} = \left(\frac{m_e}{m_\mu}\right)^2 \left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2}\right)^2 = 1.233 \times 10^{-4} \ . \tag{2.7}$$

A measurement which disagrees with this prediction may be explained with particles contained in theories beyond the Standard Model. BSM theories with more Higgs content would contain charged Higgs bosons, there are pseudoscalar leptoquarks in theories with dynamical symmetry breaking, and Pati-Salam types of grand unified theories contain vector leptoquarks. These particles lead to processes containing pseudoscalar currents, thus creating order $1/m_{H^\pm}^2$ contributions to f^e_{PL}. These

contributions will lead to branching ratio corrections of the form [66]

$$R_{\pi_{e2}}^{\text{th}} = R_{\pi_{e2}}^{0,\text{th}} \frac{\left(1 + \frac{2m_\pi a_p}{m_e a_A} f_{\text{PL}}^{e}\right)}{\left(1 + \frac{2m_\pi a_p}{m_\mu a_A} f_{\text{PL}}^{\mu}\right)} \; . \tag{2.8}$$

But since m_e/m_μ is of order 10^{-3} and we expect f_{PL}^{μ} to be of the same order as f_{PL}^{e}, we can neglect

the second term in the denominator.

$$R_{\pi_{e2}}^{\text{th}} = R_{\pi_{e2}}^{0,\text{th}} \left(1 + \frac{2m_\pi a_p}{m_e a_A} f_{\text{PL}}^{e}\right) \; . \tag{2.9}$$

The experimental branching ratio can be written as,

$$R_{\pi_{e2}}^{\text{exp}} = R_{\pi_{e2}}^{0,\text{exp}} \pm \Delta R_{\pi_{e2}}^{0,\text{exp}} = R_{\pi_{e2}}^{0,\text{exp}} \left(1 \pm \frac{\Delta R_{\pi_{e2}}^{0,\text{exp}}}{R_{\pi_{e2}}^{0,\text{exp}}}\right) \; , \tag{2.10}$$

where $\Delta R_{\pi_{e2}}^{0,\text{exp}}$ gives the uncertainty in the measurement. The proposed relative uncertainty of the

PEN experiment is

$$\frac{\Delta R_{\pi_{e2}}^{0,\text{exp}}}{R_{\pi_{e2}}^{0,\text{exp}}} = 0.05\% \; . \tag{2.11}$$

An uncertainty of this size will further constrain f_{PL}^{e}. Assuming the BSM theories are correct,

$$R_{\pi_{e2}}^{0,\text{th}} \left(1 + \frac{2m_\pi a_p}{m_e a_A} f_{\text{PL}}^{e}\right) = R_{\pi_{e2}}^{0,\text{exp}} \left(1 + \frac{\Delta R_{\pi_{e2}}^{0,\text{exp}}}{R_{\pi_{e2}}^{0,\text{exp}}}\right) \; , \tag{2.12}$$

which means,

$$f_{\text{PL}}^{e} < \frac{m_e a_A}{2m_\pi a_p} \frac{1}{R_{\pi_{e2}}^{0,\text{th}}} \left(\sqrt{\left(R_{\pi_{e2}}^{0,\text{exp}} - R_{\pi_{e2}}^{0,\text{th}}\right)^2} + \Delta R_{\pi_{e2}}^{0,\text{exp}}\right) = 3.0 \times 10^{-7} \; . \tag{2.13}$$

Now we are able to use this constraint on f^e_{PL} to find the lower bounds on the masses of the hypothesized particles.

The charged Higgs' mass can be found from:

$$S\frac{m_t m_\tau}{m^2_{H^\pm}} \sim f^e_{PL} < 3.0 \times 10^{-7} \; , \tag{2.14}$$

where S is the mixing suppresion, taken to be 10^{-2}. The constraint is then

$$m_{H^\pm} > \sqrt{\frac{Sm_t m_\tau}{f^e_{PL}}} = 3.2 \text{ TeV} \; . \tag{2.15}$$

The contraint on pseudoscalar leptoquark mass, m_p is obtain by

$$f^e_{PL} \sim \frac{\sqrt{2}}{G_F}\left(\frac{1}{250}\right)^2 \frac{1}{2m^2_p} \; , \tag{2.16}$$

which gives,

$$m_p > \sqrt{\frac{1}{\sqrt{2}}\left(\frac{1}{250}\right)^2 \frac{1}{f^e_{PL}} \frac{1}{G_F}} = 1.8 \text{ TeV} \; . \tag{2.17}$$

The vector leptoquark's mass, m_G, is constrained by

$$f^e_{PL} \sim \frac{4m^2_W}{m^2_G} \; , \tag{2.18}$$

so we have

$$m_G > \sqrt{\frac{4m^2_W}{f^e_{PL}}} = \frac{2m_W}{\sqrt{f^e_{PL}}} = 292 \text{ TeV} \; . \tag{2.19}$$

The numerical values for the quantities used in these calculations are given in Table 2.1.

Assuming the PEN measurement gives the same central value for the π_{e2} branching ratio but

Table 2.1: Numerical values used in mass limit calculations.

Quantity	Value
m_π	139.57 MeV/c^2
m_{u+d}	10 MeV/c^2
m_e	0.511 MeV/c^2
m_t	173.8 GeV/c^2
m_W	80 GeV/c^2
$R_{0,exp}$	1.2352×10^{-4}
G_F	1.1664×10^{-5} GeV^{-2}

with a relative uncertainty of 0.05%, the lower bounds of the aforementioned hypothetical particle masses in BSM theories will be raised. In summary, the calculated mass bounds on the charged Higgs, pseudoscalar leptoquark, and vector leptoquak are: $m_{H^\pm} > 3.2$ TeV, $m_p > 1.8$ TeV, and $m_G > 292$ TeV, respectively.

2.6 Ratio of Scalar to Vector Coupling

Campbell and Maybury [20] discuss indirect constraints on the ratio of scalar to vector coupling, C_S,

$$-1.2 \times 10^{-3} \leq C_S \leq 2.7 \times 10^{-4} \ . \tag{2.20}$$

When combining experimental measurements of $R_{\pi e2}$ with limits on scalar interactions from muon capture experiments, one can provide an order of magnitude stronger limit on C_S than that which is possible from direct experimental searches [23].

2.7 Higgsino, Chargino, and Slepton Bounds and R-Parity in Supersymmetry

This section summarizes the calculations by M. Ramsey-Musolf et al. [63]. They studied the effects of supersymmetry (SUSY) on $R_{\pi_{e2}}$ in the minimally symmetric standard model (MSSM) both with and without R-parity conservation. In the R-parity conserving case, effects from SUSY loops can be of the same order of magnitude as the planned uncertainty of the PEN experiment. A deviation in $R_{\pi_{e2}}$ of

$$0.0005 \leq \left| \frac{\Delta R_{\pi_{e2}}^{\mathrm{SUSY}}}{R_{\pi_{e2}}} \right| \leq 0.001 \ , \tag{2.21}$$

due to the MSSM would imply

- a mass bound on the lightest chargino, $m_{\chi 1} \leq 250$ GeV,

- the left-handed selectron, $\tilde{e}_L$, and smuon, $\tilde{\mu}_L$, would be highly degenerate, with either $m_{\tilde{e}_L}/m_{\tilde{\mu}_L} \geq 2$ or $m_{\tilde{e}_L}/m_{\tilde{\mu}_L} \leq 1/2$,

- that at least one of $\tilde{e}_L$ or $\tilde{\mu}_L$ must be light, such that $m_{\tilde{e}_L} \leq 300$ GeV or $\tilde{\mu}_L \leq 300$ GeV, and

- the Higgsino mass parameter μ and left-handed up squark mass $\tilde{u}_L$ satisfy either $|\mu| \leq 250$ GeV or $|\mu| \geq 250$ GeV and $\tilde{u}_L \leq 200$ GeV.

2.8 Summary

In conclusion, the large difference between the precision in SM calculations and previous experimental measurements of $R_{\pi_{e2}}$ strongly motivated this new precision measurement. $R_{\pi_{e2}}$ provides the best test of electron-muon universality. The data obtained in the experiment can provide mass limits

on massive neutrinos. It can constrain extra-dimension BSM theories. Furthermore, it provides limits on hypothetical leptoquarks and supersymmetric particles.

It is also worth noting that $R_{\pi_{e2}}$ can constrain other SM extensions, such as (super)compositeness theories [19]. We will also obtain further evidence for, or against, the NuTEV anomaly [72, 73, 46]. Reaching the proposed accuracy [23] in $R_{\pi_{e2}}$ will also reduce the external systematic error in the pion beta decay branching ratio [62] to an insignificant level.

Chapter 3

The PEN Experiment

The PEN experiment, named for the process under investigation (**p**ion $\rightarrow$ **e**lectron+**n**eutrino), took place at the Paul Scherrer Institute located in Villigen, Switzerland. The goal was to measure the $\pi \rightarrow e\nu(\gamma)$ branching ratio with a relative uncertainty of 5×10^{-4} or less. This chapter describes the experimental set-up and the data collection. Initial diagnostic and development runs took place in 2007, and three data production phases occurred in the years 2008, 2009, and 2010.

3.1 Beamline

The layout of the PSI accelerator facility is shown in Figure 3.3. Protons are obtained from an ion source and DC accelerated[1] to 810 keV with the Cockcroft-Walton cascade accelerator, Figure 3.1(a). A 60 keV extraction voltage results in protons with an energy of 870 keV.

Injector-2 (Figure 3.1(b)) is a ring cyclotron with 4 sector-magnets and an extremely low injection energy of 870 keV. Its specific design provides high quality, high intensity beams of 72 MeV

---

[1]The maximum extracted energy possible from a DC high voltage accelerator based on air insulation is roughly 800 keV.

protons (37.1% of light speed). The Injector-2 cyclotron was built to replace the multi-particle,

variable energy, Injector-1 constructed by Philips, Eindhoven. Since the commissioning in 1984, it

has provided beams of 72 MeV protons to be injected into the 590 MeV ring cyclotron.
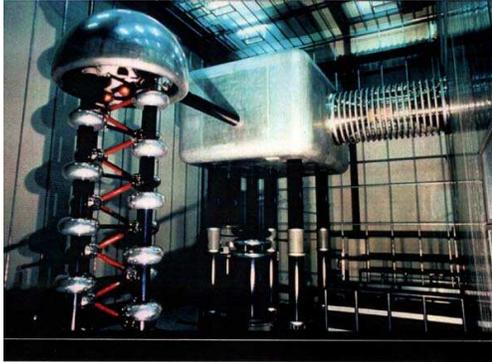
(a) *Cockcroft Walton Cascade to provide 800 keV.*

(b) *72 MeV Injector Cyclotron.*
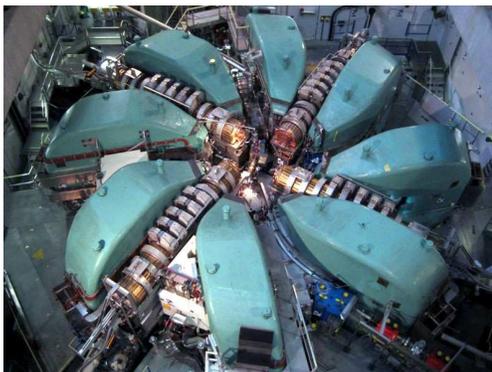
(c) *590 MeV Ring Cyclotron.*

Figure 3.1: *The three main stages of the PSI proton accelerator [2].*

The ring cyclotron is a seperated sector cyclotron with a fixed beam energy of 590 MeV (78.9% of light speed), built by PSI and commissioned in 1974. The 72 MeV beam from the Injector-2 cyclotron, enters from the back of the cyclotron and is injected into an orbit in the center of the ring. The protons are accelerated over 186 revolutions and extracted with their full energy in the foreground of Figure 3.1(c) [2]. The proton beam current is 2.2 mA DC. The accelerator frequency is 50.63 MHz, which corresponds to a time-between-pulses of 19.75 ns, with a bunch width of $\sim$0.3 ns.

Secondary beamlines at PSI provide pions and muons to experimental areas [2]. The primary beam of protons is steered towards a 4 cm long rotating cone of polycrystalline graphite[3] (Figure 3.2) where the collisions between protons and carbon nuclei occur with a center of mass energy

---

[2]The PEN experiment was situated in the πE1 area and used the πE1 beamline.

[3]Target Station E, for the French word "Epais" meaning thick.

greater than the pion mass, 140 MeV/c. Pions are created and extracted from the target in the for-

ward direction at an angle of $10°$. There are three slit systems in the beam line to control either the

beam intensity by reducing the angular acceptance of the beam, or the momentum band acceptance

and hence the momentum resolution of the transported beam.

**Figure 3.2:** *The graphite pion production target [2].*

Figure 3.3: Layout of the accelerator facility at PSI.

3.2   The PEN Detector System

The PEN detector system is basically an upgraded version of the PIBETA detector [31] that was used from 1999 to 2004 to measure the beta decay of the pion, $\pi^+ \rightarrow \pi^0 e\nu$ [62], and the radiative decay, $\pi^+ \rightarrow e^+ \nu_e \gamma$ [18]. This chapter will give a brief overview of the entire PEN detector, with emphasis given to new detector components that were added or upgraded for this experiment.
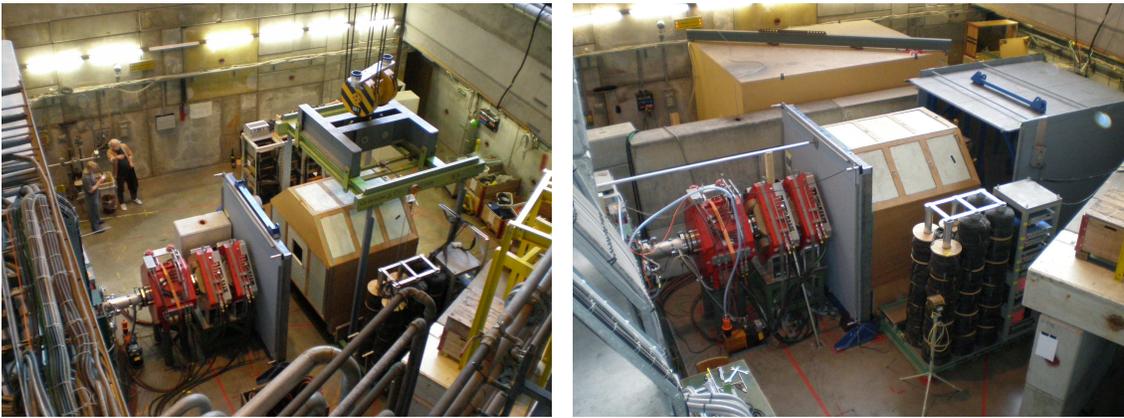
(a) *Photograph taken during the assembly of the PEN de-* (b) *PEN detector system in 2010 with the lead shielding*
tector system in 2007.   *house rolled away allowing us to view the thermal house surrounding the main detector region.*
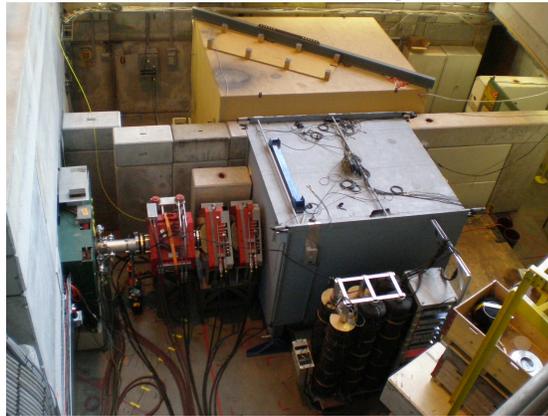
(c) *Photograph from 2007 with the lead shielding house in place.*

Figure 3.4: *Photographs of the πE1 area taken from the gallery hall (southeast) with the PEN detector system in various stages of assembly.*

3.2.1   Beam Tracking

Magnets steer the pions from the production target to the πE1 area. Bending dipole magnets and collimating slits are used to select pions with a specific momentum, typically between 70 and 85 MeV/c. The first detector the pions reach in the πE1 area is an active plastic scintillator[4] that we call the forward beam counter (BC). The thickness of the BC was increased to 3 mm from the 2 mm counter used in the PIBETA experiment[5]. It is located at the center of a beam collimator, first surrounded by a small tungsten collimator, which in turn is surrounded by a larger lead collimator. The bending dipole magnets are tuned to allow pions to pass through the 7 mm collimating hole, while positrons are stopped in the lead since they are ∼40 mm from the pions in the horizontal plane. The beam counter is wrapped in 3M VM2000 Radiant Mirror Film, a completely non-metallic multi-layer polymer film, 63.5 μm thick, to ensure all the light reaches the pair of 1-inch photomultiplier tubes attached to the top and bottom of the scintilltor. The PMTs are housed in the large lead collimator itself. Figure 3.5 shows the vacuum beam pipe with the signal and high-voltage cables for the BC PMTs entering the top and the bottom of the vacuum pipe.

After passing through the forward beam counter, the beam particles pass through a series of three focusing quadrupole magnets. The focused beam now enters the main PEN detector region where it leaves the vacuum pipe through a thin window and enters the degrader counter.

In 2008 a four-piece wedged active[6] degrader (wAD), placed 3 cm upstream of the target, was used to provide information regarding the x,y coordinates of the incoming beam particle. The idea is to have two pairs of scintillator wedges tapering from 5.0 mm to 1.5 mm with the thickness of each pair of wedges summing to a constant thickness. The ratio of scintillation light produced in each

---

[4]Bicron/Saint-Gobain: BC 408 [21]
[5]A 1 mm counter was used in the PEN development run in 2007, but resulted in an insufficient energy resolution
[6]BC 408 plastic scintillating material wrapped in 76 μm thick 3M Radiant Mirror VM 2000 reflective foil [3].

wedge of a pair is used to determine how far from the detector axis the beam particle is. A ratio of

1.0 corresponds to the detector axis, assuming the wedges of the degrader are perfectly aligned.
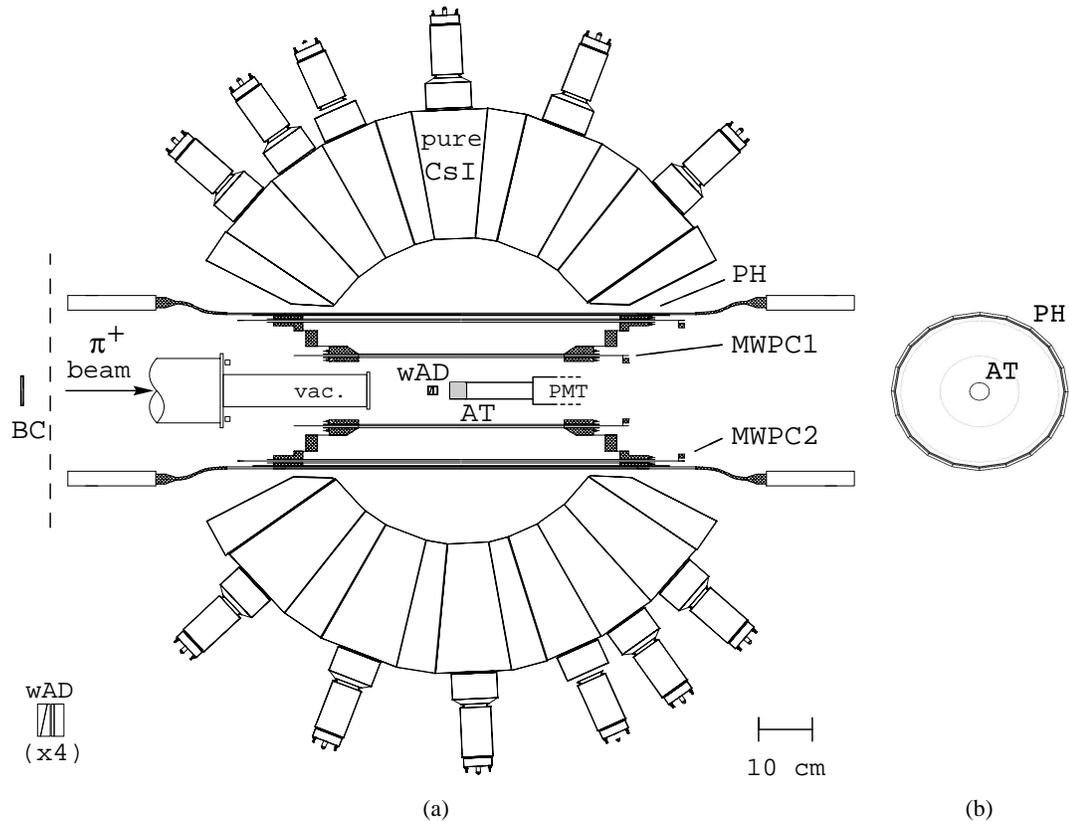
Figure 3.6: *Cross sections of 2008 PEN Detector System.* (a) *shows a slice through the y-z plane showing the wedged active degrader (wAD), active target (AT), wire chambers (MWPC1-2), plastic hodoscope (PH), and CsI calorimeter (CsI).* (b) *is a slice through the x-y plane showing the 20 hodoscope staves.*

Each degrader wedge was glued to a bent acrylic light guide and optically coupled to indi-

vidual Hamamatsu R7400U photomultiplier tube (PMT) [29]. Comparing the areas of the single-

photoelectron line with the through-going π^+ signal on a digital oscilloscope we deduced an average

light response of 160 photoelectrons/MeV for each degrader wedge. The light response of all four

wedges is digitized[7] at 2 GS/s and recorded for every event, resulting in a 0.9 ns rise time and a

2.1 ns decay time.

---

[7]Acqiris High Speed 10-bit PXI/CompactPCI Digitizer, Model DC282 [57].
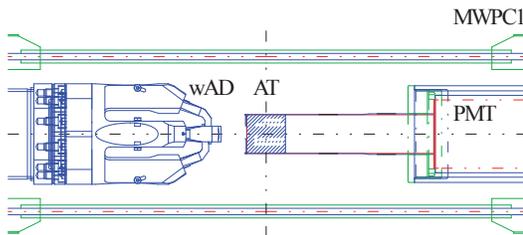
Figure 3.7: *Cross section through the central detector region indicating the locations of the inner MWPC, the wedged active degrader (wAD) and its mounting to the end of the vacuum beam pipe, as well as the air light guide between the active target (AT) and photomultiplier tube (PMT).*
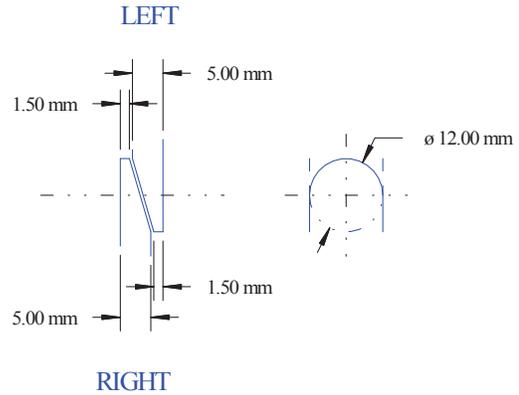

Figure 3.8: *Wedged active degrader cross section for the left and right wedges. The top and bottom wedges have the same dimensions with the exception of a 15 mm overlap diameter (as opposed to the 12 mm overlap diameter for the left and right wedges).*
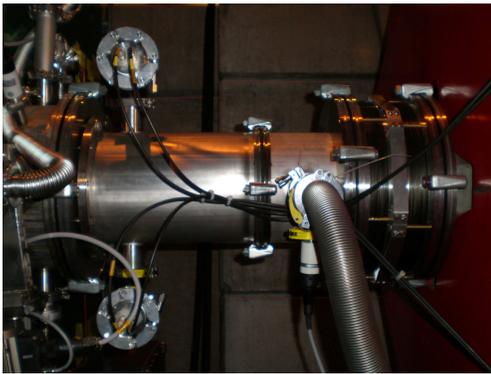
Figure 3.5: *The forward beam counter, BC, is located inside the vacuum pipe.*

This configuration requires a thicker degrader overall to ensure enough thickness at the tips of the wedges to give enough scintillation light. The drawback of this 13 mm thickness as opposed to the 5 mm single piece degrader used in the 2007 development runs is that a higher beam momentum is required to get the pion to stop in the target, resulting in more nuclear reactions in target. More material also causes an increase in multiple scattering of the pion, thus decreasing *x*,*y* position resolution.

After the beam particle passes through the degrader it enters the active target (AT). The beam momentum was chosen such that the pions will deposit enough energy in the degrader that they will come to a stop at the center of the target. The target is a cylinder made of scintillating plastic[8] with a 15 mm length and a 15 mm radius. When the pion decays at rest into a muon and a neutrino,

---

[8]Bicron/Saint-Gobain: BC 418 [21]
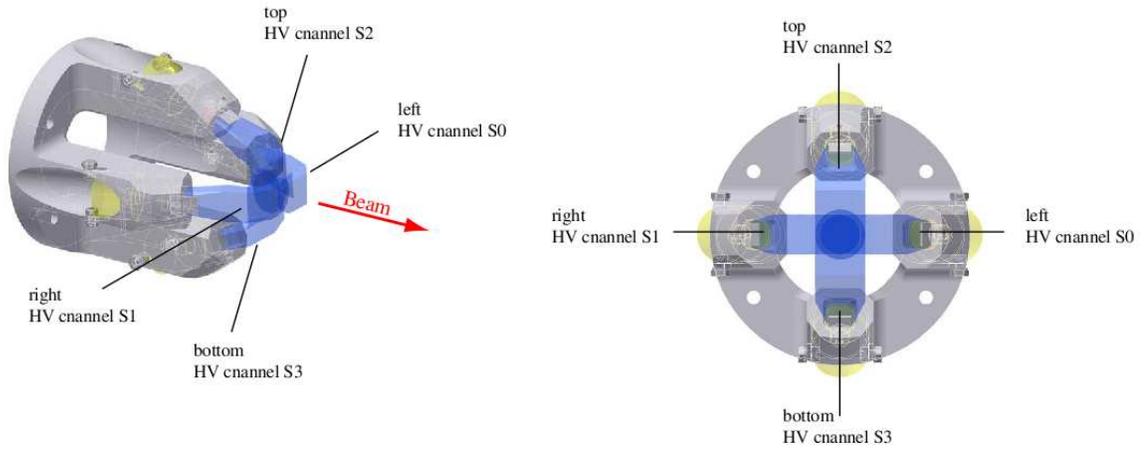
Figure 3.9: *Drawing of the wedged active degrader.*

the monoenergetic muon travels 1.38 mm, entirely inside the target. The stopped muon will subse-

quently decay into a positron that exits the target. The light collected by the PMT[9] attached to the

target via a light guide is digitized and recorded for every event.

---

[9]Hamamatsu H2431-50

(a)

(b)

(c)

(d)
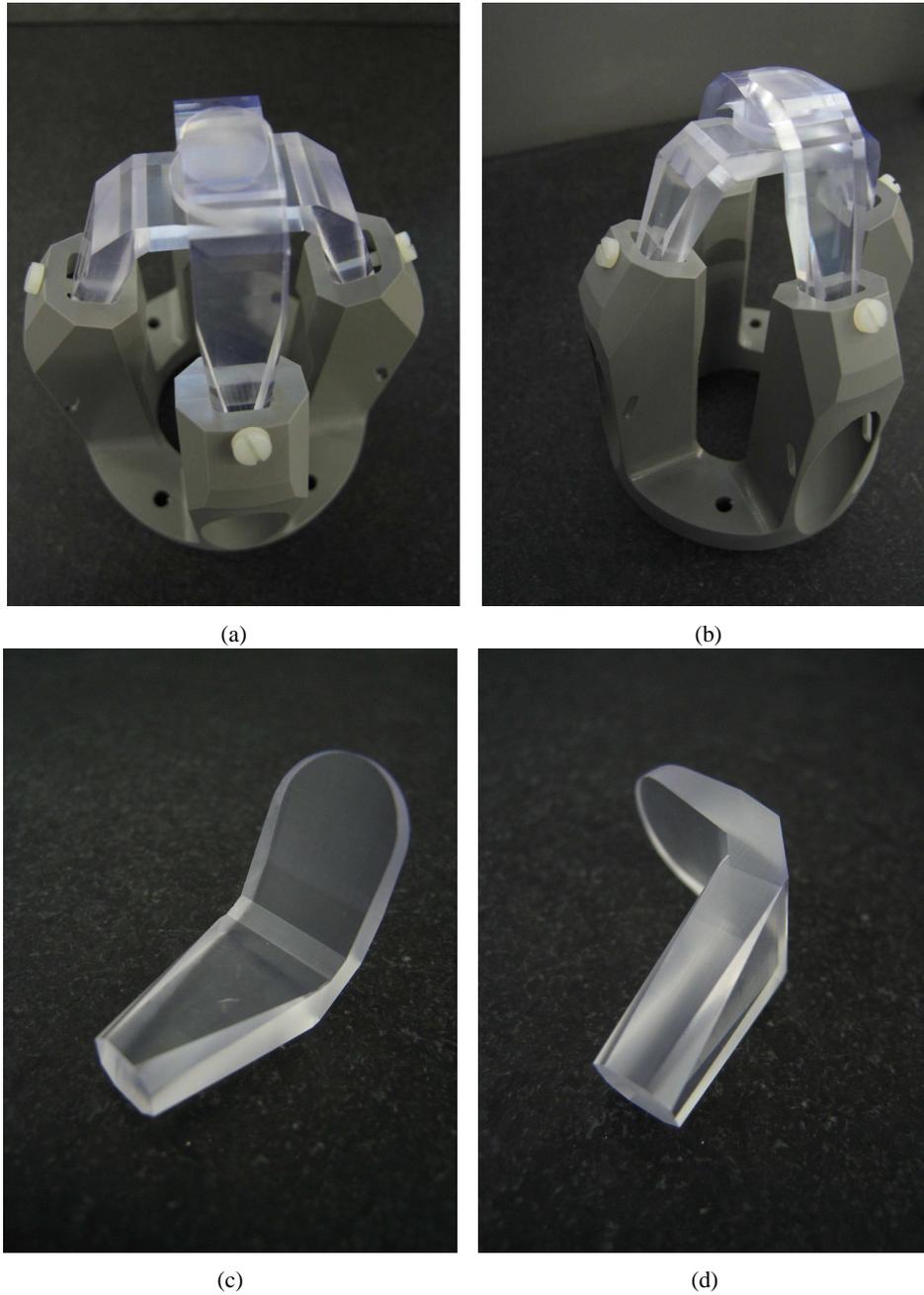
Figure 3.10: Photographs of the wedged active degrader. (a) and (b) show the four wedges together with the mounting allowing small photomultiplier tubes to attach to the ends. (c) and (d) show an individual wedge.

3.2.2 Decay Tracking

Surrounding the target are two concentric cyclindrical multi-wire proportional chambers (MWPC1-2). The inner chamber has a diameter of 12.5 cm and the outer chamber has a diameter of 25.0 cm. These wire chambers give precise tracking information for charged particles only, as neutral particles (mainly photons) are not detected.
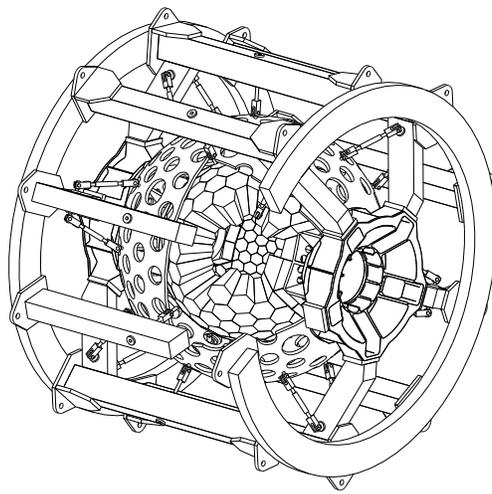

(a) *Drawing of the PEN detector system.*   (b) *Photograph of the PEN detector system.*

Figure 3.11: *Drawing (a) and photograph (b) of PEN detector system during the refurbishment in 2007. This figure shows mainly the 240 crystal CsI calorimeter, the photomultiplier tube bases, the support structure, and ducts for temperature control.*

The outer wire chamber is surrounded by twenty staves of plastic scintillator[10] which together form the plastic hodoscope (PH) system. The thickness of the staves was increased from 3.25 mm to 4.0 mm for the PEN experiment. The original staves suffered significant radiation damage and aging. Light guides and photomultiplier tubes are attached to each end of the 20 scintillating staves. As a charged particle passes through the hodoscope, time-to-digital converters (TDCs) record the time of the light pulse, and amplitude-to-digital converters (ADCs) record the energy. Since the hodoscope is cylindrical and the secondary tracks are radial, it helps to use energy loss per distance

---

[10]Bicron/Saint-Gobain: BC 408 [21]

travelled, dE_{PH}/dx_{PH}, instead of just E_{PH}. The staves are 0.4 cm thick so the distance travelled becomes $dx = 0.4\,\text{cm}/\sin\theta$, and $(dE/dx)_{PH} = E_{PH}\sin\theta/0.4\,\text{cm}$. The minimum ionizing positrons give a narrow peak, while the protons exibit quenching [45].
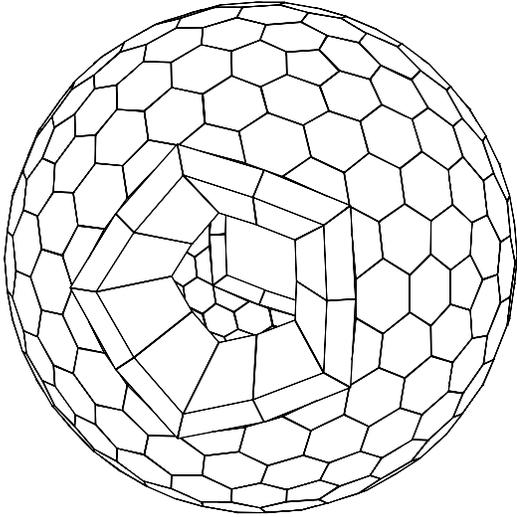
Figure 3.12: *Three dimensional view of the 240-segment pure-CsI crystal calorimeter. Ten "veto" crystals surround the upstream and downstream beam openings.*

Next in the sequence of decay tracking detectors is the 240-module pure-CsI crystal sphere shown in Figures 3.12 and 3.13. The crystal calorimeter collects energy deposited by both charged particles and photons covering a total solid angle of $0.77 \times 4\pi$ sr. The inner radius of the calorimeter is 26 cm and the axial length of the modules is 22 cm. The positron usually annihilates within about 1 radiation length, producing Bremstrahlung and a photon shower. The crystals are approximately 12 radiation lengths thick and contain most of the shower. Unfortunately, some shower photons may escape and we are unable to record the full energy.
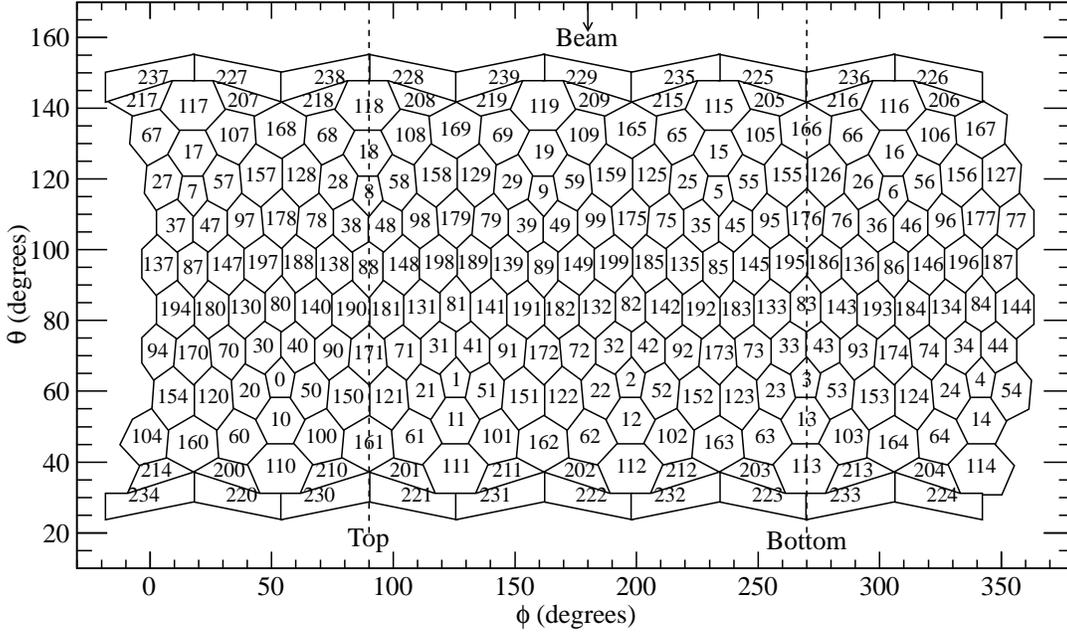
Figure 3.13: *Mercator projection of the 240 pure-CsI crystals in the calorimeter as seen from the target.*

3.2.3   Triggers

Analog PMT signals from detector elements are split into two branches, one to trigger the data

acquistion (DAQ) system, and one to digitizing electronics. This section describes the trigger logic

used to decide when to write data to disk.

A time coincidence between the forward beam counter and the two downstream degrader wedges

(right and left), adjusted for the time difference corresponding to the π^+ time-of-flight between the

two detectors, indicates that we have a pion stopping in the target. The BC, DR $\cdot$ DL, and their

coincidence signal are shown in Figure 3.14. The beam pions travel much slower than the remain-

ing beam muons and beam positrons and are easily separated by considering time-of-flight (Figure

3.15). We refer to this coincidence signal as

$$\pi\text{STOP} = \text{BC} \cdot (\text{DR} \cdot \text{DL})_{\text{High Threshold}} \; , \tag{3.1}$$

Table 3.1: The π^+ stopping rate in 2008.

πSTOP Rate (kHz)	Run Range
15	83580-84176
21	84177-84237
29	84238-84555
41	84556-84617
28	84618-84812

where "High Threshold" indicates that we required a minimum amount of energy to be deposited in the last two wedges to consider that beam particle to be a pion. Half way through the 2008 data collection[11] we required further coincidence with the target,

$$\pi\text{STOP} = \text{BC} \cdot (\text{DR} \cdot \text{DL})_{\text{High Threshold}} \cdot \text{TGT}_{\text{Low Threshold}} \cdot \qquad (3.2)$$
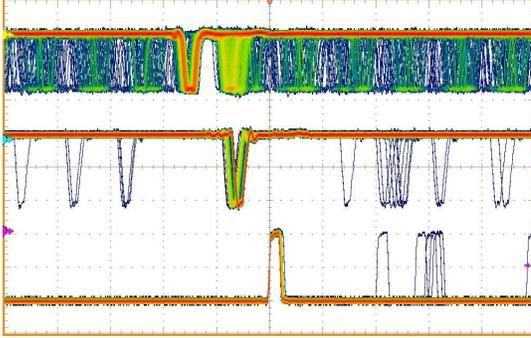
Figure 3.14: *Coincidence between forward Beam Counter and Degrader. Top trace = BC, Middle trace = (DR · DL), Bottom trace = Coincidence. Oscilloscope is triggered on the coincidence. Vertical scale: 400 mV/div. Horizontal scale = 20 ns/div.*
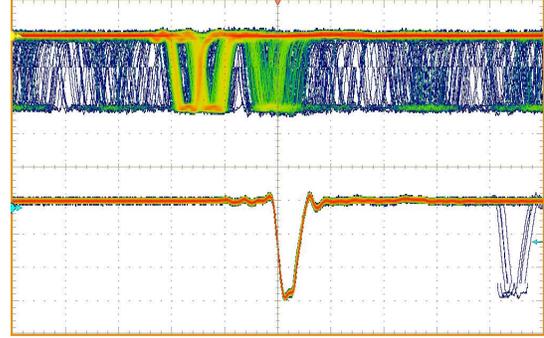
Figure 3.15: *The BC signal contains three distinct pulses. By aligning the coincidence with the degrader signal one can select the particle (pion, muon, positron – left to right). Top trace = BC, Bottom trace = (DR · DL). Oscilloscope is triggered on (DR · DL). Vertical scale: 300 mV/div. Horizontal scale = 10 ns/div.*

The πSTOP signal normally triggers the formation of a long logic signal called πGATE. The width of the gate signal determines the amount of time the electronics will record data for that event. Figure 3.16 shows the BC, (DR · DL), and πGATE signals for a single event.

---

[11]πSTOP was changed starting run 84176.

We can't just start recording the data arriving after the gate opens, otherwise we would miss the beginning of the event. Therefore the data branch is delayed with respect to the trigger branch. For the first half of 2008 data collection the data branch was delayed such that the πSTOP occured 25 ns after the gate opened. For the later half of 2008 the delay was reduced to 10 ns. The gate was open for approximately 250 ns.
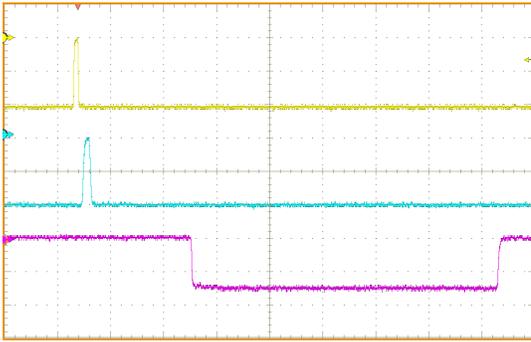
Figure 3.16: *The πGATE signal for a single event. Top trace = BC, Middle trace = (DR · DL), Bottom trace = πGATE. Vertical scale: 400 mV/div (top), 700 mV/div (middle), 500 mV/div (bottom). Horizontal scale = 50 ns/div.*
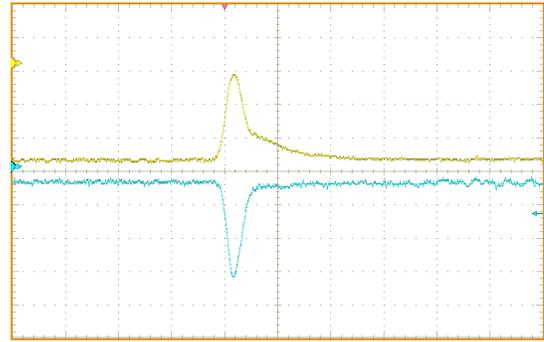
Figure 3.17: *Target pulse for a single event before and after shaping. Top trace = Original TGT signal, Bottom trace = Shaped TGT signal. The shaped TGT signal was first implemented in 2009 and is is sharper, with less tail. Vertical scale = 30 mV/div(top), 80 mV/div(bottom). Horizontal scale = 20 ns/div.*

Since the $\pi \rightarrow e$ events occur roughly a thousand times less frequently than $\pi \rightarrow \mu \rightarrow e$ events we must be smart about the events we chose to record. Every time the system accepts an event, there is an associated "dead time" while the system collects the information corresponding roughly to the 250 ns that the gate is open. If we were to record every πSTOP we see, we would collect too few $\pi \rightarrow e$ events and the experiment would take too long. The solution is to include the energy deposited on the decay side of the event into the trigger. We can then record all events with an energy above a certain threshold (predominately $\pi \rightarrow e$ events) and prescale events below the threshold (predominately $\pi \rightarrow \mu \rightarrow e$ events).

In 2008 the PEN experiment still used the field programmable gate arrays (FPGAs) as in the

PIBETA experiment and looked at the energy of the individual clusters of CsI crystals [31]. These

FPGAs took the OR of the energies in the 60 clusters of CsI crystals to determine if any of them

had energy above threshold. When that occurs in coincidence with a πGATE, a high-energy trigger

fires,

$$\mathrm{TR_{High}} = E_{\mathrm{CsI}}^{\mathrm{Above\ Threshold}} \cdot \pi\mathrm{GATE} \ . \tag{3.3}$$

For other events that don't have high enough energy deposited in the CsI calorimeter, hits in the

hodoscope are considered. The FPGA was already programmed with a 1/16 prescale factor for

the πGATE signal. Using this software to take the coincidence between a hit in the hodoscope

and prescaled πGATE gives a low-energy trigger for only 1 out of every 16 of such events. The

1/16 software prescaling wasn't enough of a data reduction so we implemented an additional 1/4

hardware prescaling to the hodoscope signal resulting in a total prescale factor of 1/64 for the low-

energy trigger,

$$\mathrm{TR_{Low}^{PS\ 1:64}} = \mathrm{PH_{Hit}^{PS\ 1:4}} \cdot \pi\mathrm{GATE^{PS\ 1:16}} \ . \tag{3.4}$$

The CsI energy threshold varied slightly throughout 2008. There are run ranges with the threshold

set to 50.9, 51.2, and 50.5 MeV. Additionally, only the 220 full-sized crystals are in the trigger logic,

so the rings of veto crystals don't contribute to the trigger.

In September of 2009 the summing of the CsI crystal energies changed from the cluster logic of

PIBETA to a new implementation involving different groups of neighbors. The 2009 high-energy

trigger also required a hodoscope coincidence. The target signal in πSTOP was also shaped to

provide a faster rise-time and to allow us to use hardware to search for a possible muon in the target,

see Appendix A.
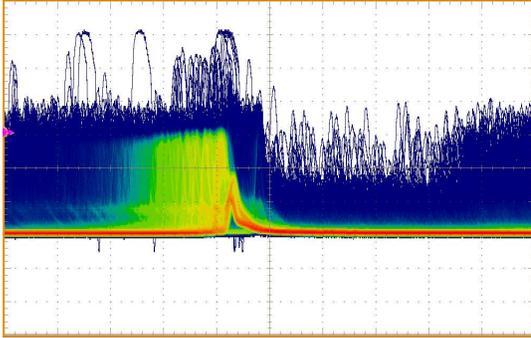
Figure 3.18: *Target signal triggered with "no muon" signal, selecting π → e events. Oscilloscope persistance for 5 seconds. The distribution is almost completely gone at 100 ns. The pronounced pulse is the positron signal and the peaks to the left of the positron are the pions. Vertical scale = 30 mV/div. Horizontal scale = 50 ns/div.*
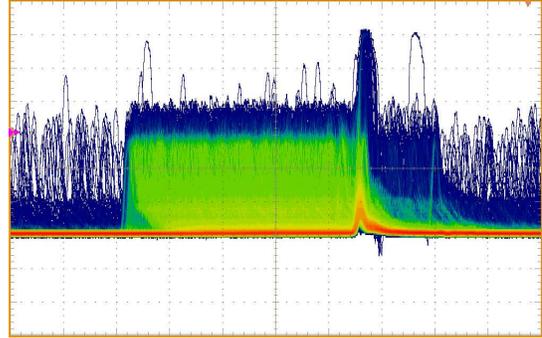
Figure 3.19: *Target signal triggered on (PH$_{Hit}$ AND πSTOP). Oscilloscope persistance for 5 seconds. The length of the distribution extends for the full πGATE width. The pronounced pulse is the positron signal and the peaks to the left of the positron are the pions. Vertical scale = 30 mV/div. Horizontal scale = 50 ns/div.*

3.2.4   Revamped Slow-Control System
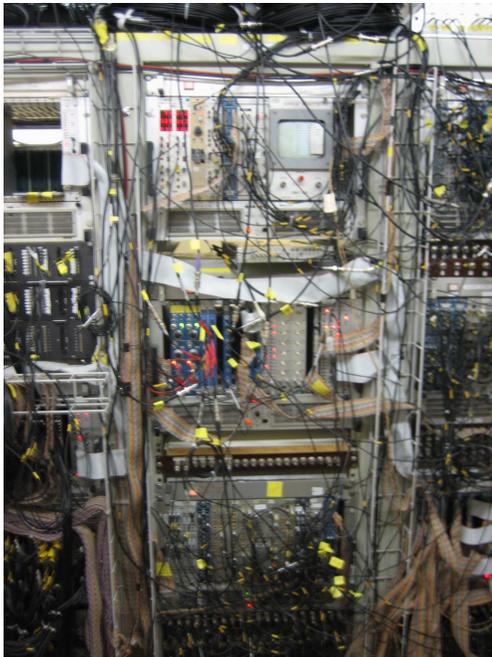
Figure 3.20: *A portion of one of the NIM electronics racks located in the climate-controlled Electronics Hut.*

The Slow-Control System was completely redesigned just prior to the 2008 data collection run. This system was responsible for controlling and monitoring high voltages, temperature, water and air flow for humidity regulation, and gas flow for wire chambers.

The heart of the slow-control system is the SCS-200, a PSI designed monitoring and control interface. It contains both voltage and current sensors as well as an Ethernet connection that allows us to change set point values remotely.

The slow-control system regulated the high voltage applied to the PMTs attached to all beam coun-
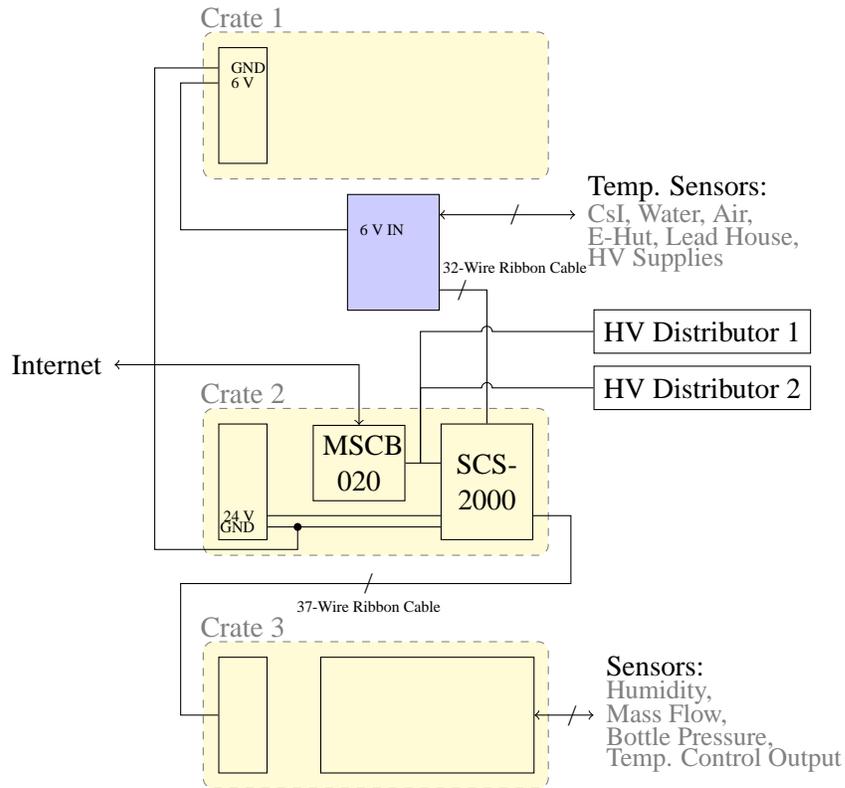
Figure 3.21: *Schematic diagram for the slow-control system.*

ters and CsI crystals. The temperature was regulated in the main detector region inside the thermal

house as well as between the thermal house and the lead shielding house. Temperature sensors were

mounted to the inner and outer surfaces of four CsI crystals, 22, 88, 118, and 174. The locations of

these crystals are shown in figure 3.13. The temperature was also regulated in the electronics hut,

and the temperature of the high voltage supplies was monitored. The main detector region (inside

the thermal house) also had humidity regulation. The SCS-200 also controlled the gas flow for the

two cylindrical MWPCs and later for the miniature time projection chamber as well.

3.3 Modifications to the Detector Subsystems in 2009-2010

I have described the experimental set up as it was in 2008. As the data collection proceded, collaborators devised new methods to make the data cleaner. In particular, the wedged active degrader was replaced with a miniature time projection chamber (mTPC), resulting in a significant increase in beam particle lateral tracking resolution. Not only was the tracking more precise, but the wAD was replaced with a much thinner single-piece degrader, reducing the effects of pion multiple scattering.

The trigger logic was completely redesigned. In September of 2009, we removed the CsI energy clustering logic as it was implemented in the days of PIBETA and replaced it with grouping CsI crystals with their neighbors. The energy threshold was then triggered on the sum of all 220 full-sized crystals directly, rather than pre-defined clusters. After reprogramming the FPGA unit (LB102) to enhance the trigger coincidences we received spurious and intermittent double pulsing of our trigger bit. We decided to replace the trigger logic with hardware rather than the software based FPGAs. Appendix A shows all of the updated trigger diagrams as a reference.

We decided to form a dedicated trigger to collect more $\pi \rightarrow$ e events in the low-energy tail region. Doing so, we implemented a "tail trigger" by accepting events with low-energy triggers and missing the 4.12 MeV muon pulse in the target waveform. We performed shaping of the target pulse in hardware (Figure A.10) which gave a faster rise time and faster fall time, allowing us to better separate a muon pulse possibly existing after the pion pulse. This logical condition was referred to as the "muon veto" in the schematic diagram shown in Figure A.11. Another improvement was the implementation of a strobe signal in order to force all triggers to have the same (Plastic Hodoscope) trigger timing, thus reducing the possibility of an associated systematic uncertainty from trigger-time walk (Figures A.7-A.9).

Chapter 4

Calibration and Temporal Stabilization

Many measured quantities in the PEN experiment change over time. The measurements are affected by external influences as well as intrinsic detector properties. We must use our knowledge of physics to stabilize the variation in these measurements, and in doing so we introduce calibration parameters. There are three types of parameters used in this analysis: Type A parameters have to be determined from a replay analysis of the data but should be constant in time (e.g., the relative rotation between the two MWPCs), Type B parameters should be constant for large sets of runs so we only store a few values and the run numbers at which they change (e.g., the target position), Type C parameters may possibly change for every run (e.g., the detector gains), so we store them in calibration offline database (.codb) files and load the values during subsequent analyses. In several cases, calibration parameters depend on the values of other parameters. Therefore, one must determine the parameters in several stages. Four calibration passes were deemed necessary to sufficiently stabilize the PEN data.

For the 2008 data set we will be concerned only with measurement data collected after the insertion of the wedged degrader, corresponding to runs 83580–84812.

4.1   Initialization

The first step in any analysis is to define the reference coordinate system: z corresponds to the beam

direction, x is beam left (left when facing downstream), and y is up. We have chosen to use the

multi-wire proportional chamber coordinates to define the origin of our system. Since we have two

such chambers, each with its own intrinsic coordinate system, it is possible for the two structures to

be physically mounted in a misaligned fashion. Therefore we have devised two parameters to check

the relative rotation, using the distance of closest approach (DCA), and longitudinal displacement

(z_0^{mirror}) of the MWPC. If the wire chambers are aligned, the distributions of each observable should

be symmetric about the origin and independent of the beam's stop position. We define,

$$\text{DCA} \equiv 2n_1 - n_2 \text{ and,} \tag{4.1}$$

$$z_0^{\text{mirror}} = \begin{cases} z_2 - 2z_1 & \text{for } z_2 > z_1 \\ 2z_1 - z_2 & \text{otherwise,} \end{cases} \tag{4.2}$$

where n_1 and n_2 are the hit wire number of the inner and outer chamber, respectively, and z_1 and z_2

are the axial position (z) determined from the inner and outer chamber, respectively. The factor of 2

arises in (4.1) because there are twice as many wires in the outer chamber as there are in the inner

chamber. The factor of 2 in (4.2) is due to the fact that the outer chamber radius is twice as large as

the inner chamber radius. If the beam stop position is not at $z_0 = 0$ then two peaks will emerge and

should be symmetric about $z_0 = 0$. The same is true with DCA for stop distributions not uniformly

distributed in azimuthal angle, ϕ. The mean values of these distributions were found to be consistent

with zero, hence the two wire chambers are not misaligned, and no corrections are necessary.

   The readout of some of the anodes on the wire chambers were not responding, and some died

during the data taking. The "dead" anodes were determined and taken into account.

The plastic hodoscope detector is used for very precise timing of the positron. Therefore it is essential to align each of the forty TDC values, from PMTs attached to the PH ends, to each other. Any mismatch in timing may be due to different cable lengths, internal signal transit times of the photo-multiplier tubes, or varied calibrations of different TDC channels in the electronics. These 40 time-offset corrections were found and range from -1.547 ns to 1.946 ns. Next we determined a coarse conversion between ADC channel and MeV for the plastic hodoscope energies. We found this value to be 0.0045 MeV/channel. The fine adjustment may vary over time depending on conditions such as applied voltage to the PMTs and is determined in the first calibration pass.

Similarly, a coarse conversion factor of 0.0381 MeV/channel was determined for the energy measurments of the CsI calorimeter ADCs. The fine, run-dependent stabilization and calibration of the CsI energies is done in its own calibration pass discussed in Section 4.3. Up to the time of writing this dissertation, it was deemed unnecessary to calibrate the CsI TDC times, as they were calibrated precisely enough during online data collection.

4.2   Calibration Pass 1

The mean range, R_π, that the pion travelled in the target is deduced using the forward-backward asymmetry in the positron energy loss in the target, and is given by the formula,

$$R_\pi = 15 \left(\frac{\overline{E}_{e^+}^{\text{up}}}{\overline{E}_{e^+}^{\text{up}} + \overline{E}_{e^+}^{\text{down}}} \right) \text{ mm}, \tag{4.3}$$

where the length of the target is 15 mm and $\overline{E}_{e^+}^{\text{up}}$ and $\overline{E}_{e^+}^{\text{down}}$, Figure 4.1, are the mean energy deposited by positrons projected onto the z axis exiting the upstream and downstream faces of the target,
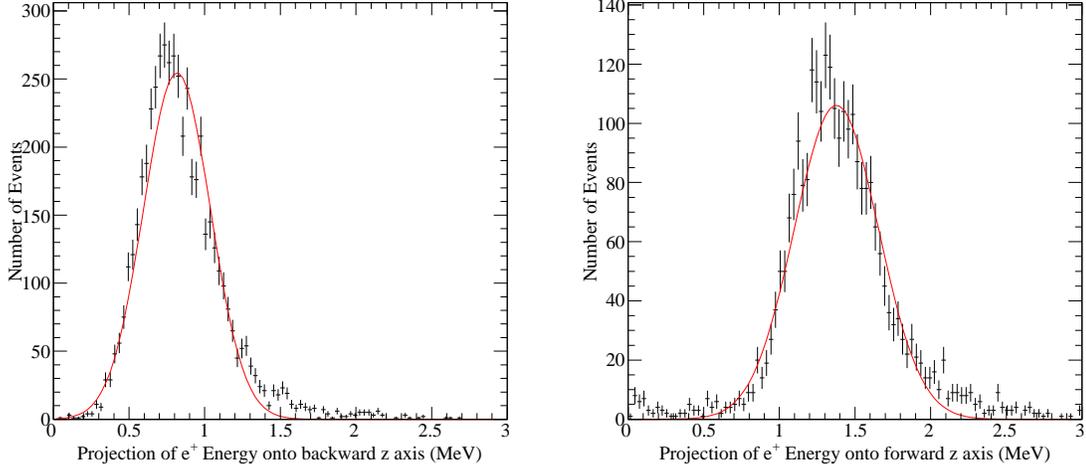
Figure 4.1: *Positron energy projected onto the z axis for trajectories going upstream (left) and downstream (right).*

respectively. The variation of R_π from run to run is shown in Figure 4.3.

The absolute stop distribution is read immediately from the $z_0^\perp$ distribution shown in Figure 4.2, where z_0 is the z coordinate of the origin of the positron track deduced from the wire chamber information. The $z_0^\perp$ distribution only contains tracks with a direction perpendicular to the beam axis, i.e., $85° < \theta < 95°$. The target center in the z direction is then $z_{\text{TGT}} = \bar{z}_0^\perp - R_\pi + L_{\text{TGT}}/2$ mm, where $L_{\text{TGT}} = 15$ mm is the length of the target along the z direction.

The center of the degrader in the z direction is also deduced using the $z_0^\perp$ distribution. The bin contents of this distribution are set to zero if the content is below a certain threshold. Then the last bin with entries before a continuous span of 20 empty bins is determined to be the downstream face of the last degrader wedge. Subtracting half the degrader thickness then gives the center of the set of degrader wedges, z_{DEG}. Figure 4.5 shows a discrete shift at run 84184.

The forty energies recorded by the ADCs on each side of the twenty plastic hodoscope staves were roughly converted to MeV before the first calibration pass. Next, any systematic effects in
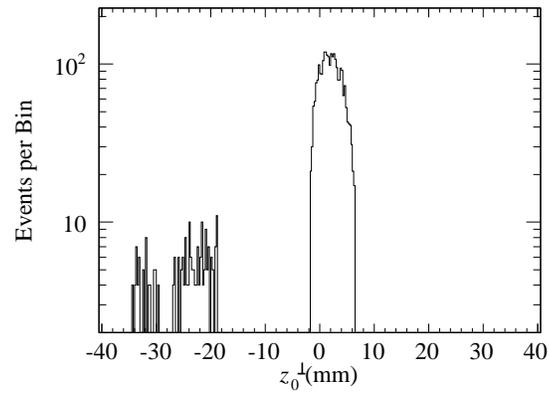
Figure 4.2: *z_0 distribution for perpendicular tracks with low entry bins set to zero.*

these gains, such as PMT high voltage drift, are corrected via an array of forty calibration parameters. As an example, Figure 4.6 shows the variations in the first two of these forty parameters.
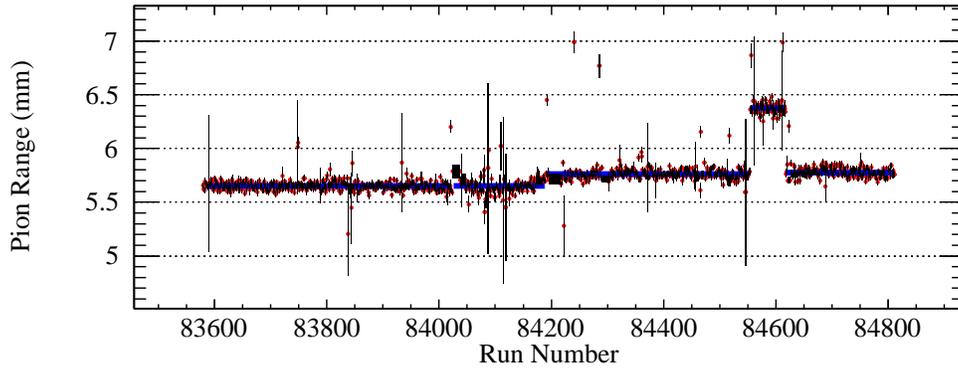
Figure 4.3: *Mean π^+ penetration depth in the target versus run number in 2008. The observed variations are due to different beam momenta.*
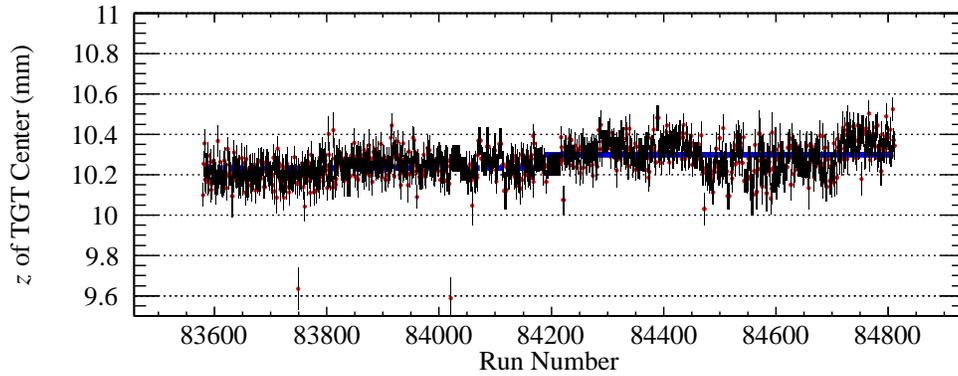
Figure 4.4: *z of the target center versus run number in 2008.*
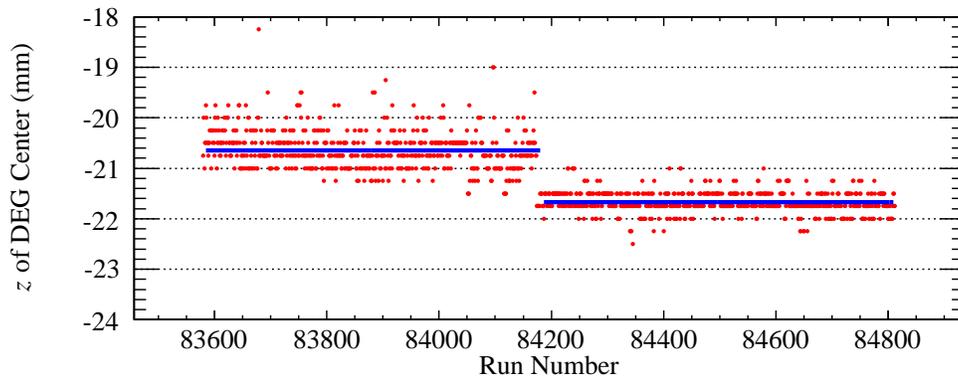
Figure 4.5: *z of the degrader center versus run number in 2008. The discrete jump occured when the degrader was moved during detector maintenance.*
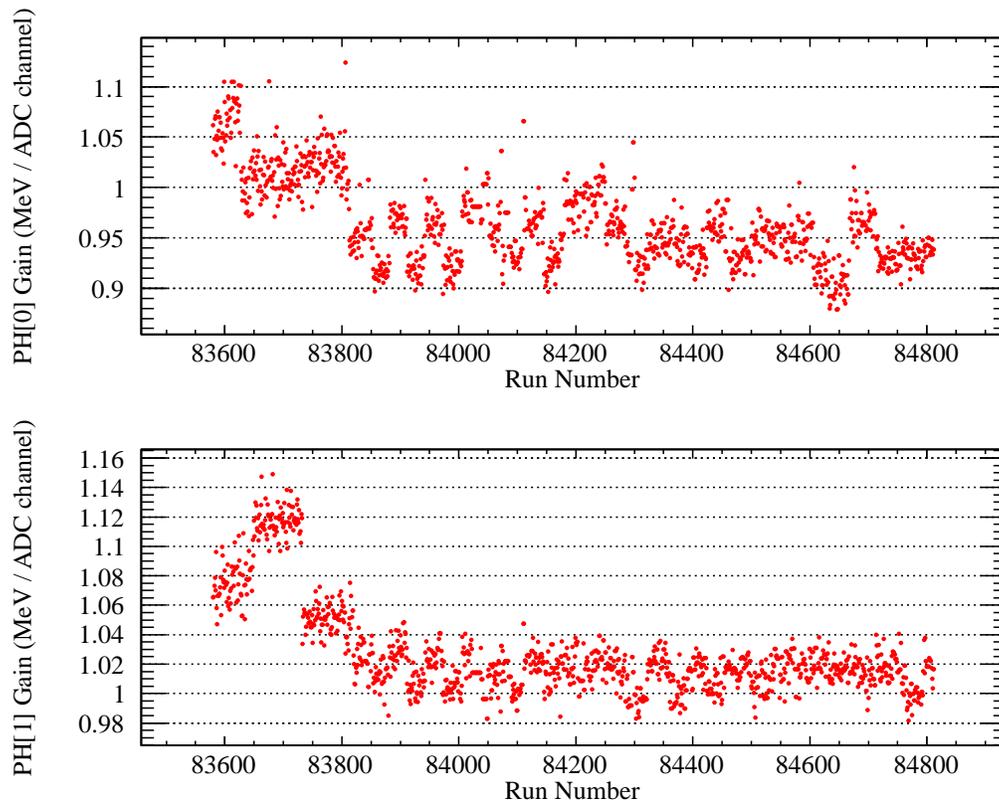
Figure 4.6: *A demonstration of the energy gain correction factors for two plastic hodoscope channels, each plotted versus run number for 2008.*

Calibration Pass 2

Once the mean penetration range of the pion in the target, R_π, has been determined for each run, we are able to calculate the momentum the pion had upon entering the target. We use the range-momentum relation empirically deduced from simulation [5, 6],

$$R_\pi = \left(\frac{p_\pi}{64.56} \right)^{3.5} \text{cm} . \tag{4.4}$$

Working backwards in steps, adding the mean thickness of each degrader wedge, we are able to find the mean energy the pion would have deposited in each wedge. Assuming the beam profile is wide enough and is roughly centered on the degrader, we can shift the mean energy in each degrader wedge, Figure 4.7, to be equal to the expected energy. Figure 4.8 shows the varation in these four gain parameters from run to run.

The next parameter is used to make sure we obtain the correct time difference between the upstream beam counter and the degrader. Since we know the distance between the beam counter and the degrader we can calculate the time it would take a photon to travel between the two counters. At our operating momentum, positrons travel very close to the speed of light. Selecting positrons in the beam and adjusting t^e_{BC} allows us to determine the time-offset correction for the upstream beam counter for each run, Figure 4.9.
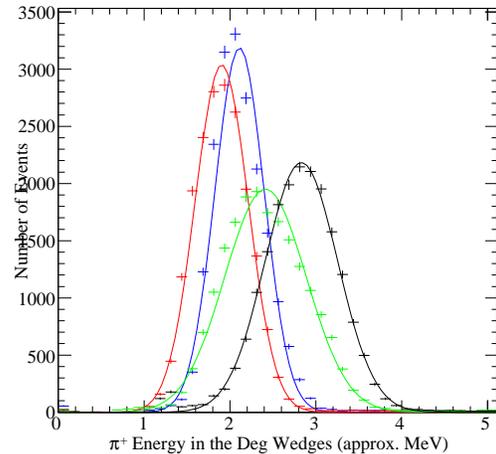
Figure 4.7: *Fits to the uncalibrated wedged degrader energies, with Bottom (red), Top (blue), Right (green), Left (black).*
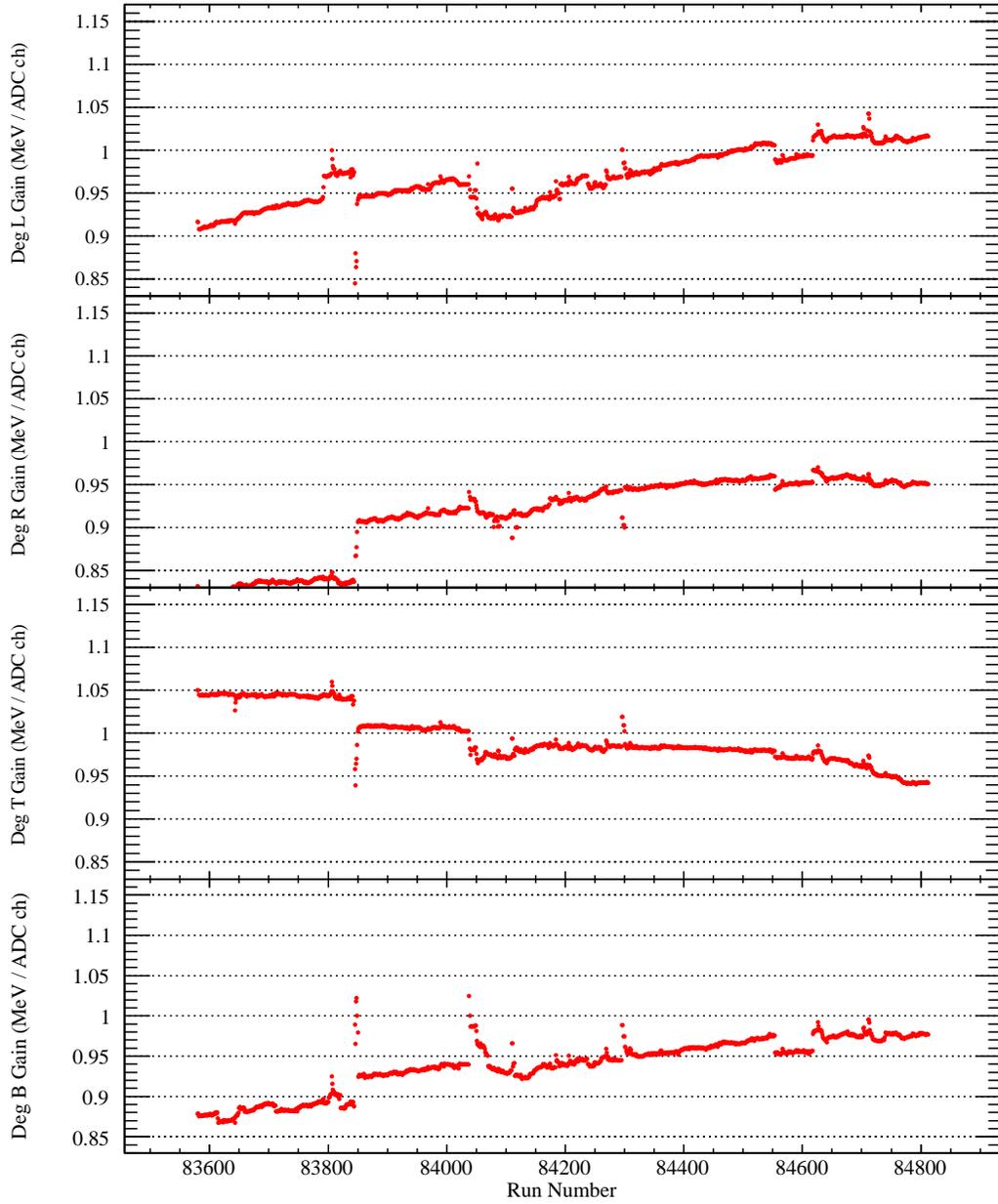
Figure 4.8: *The run dependent energy gain factors for the four degrader wedges.*
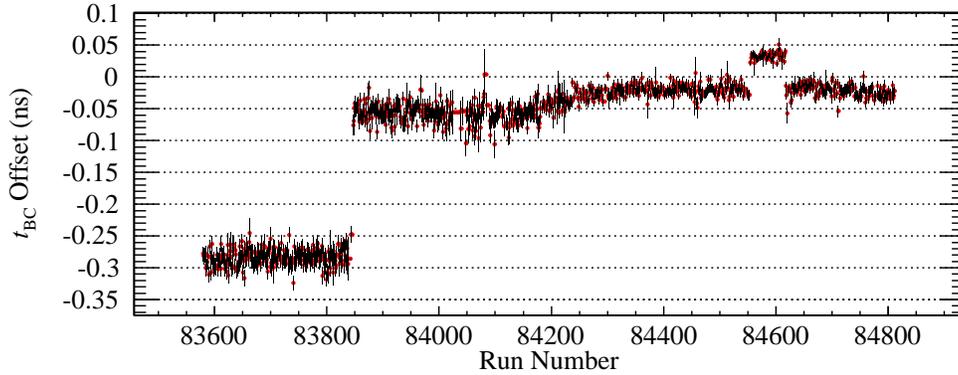
Figure 4.9: *Shift in the Beam Counter timing deduced from the Beam Counter to Degrader time-of-flight, plotted versus run number for 2008.*
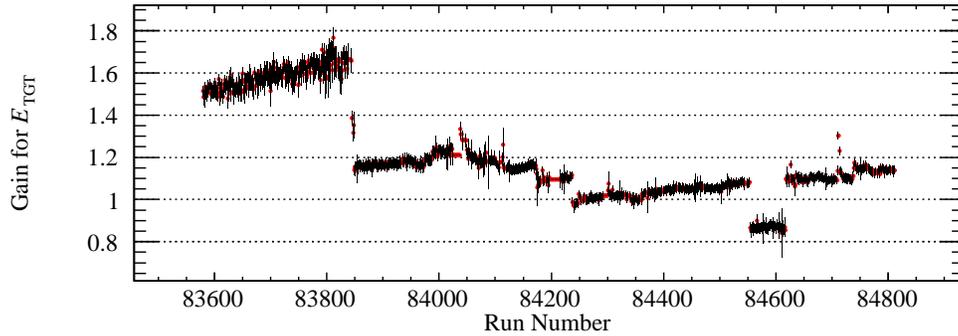

Figure 4.10: *Run-dependant target energy gain correction factor in 2008.*

The final parameter determined in pass 2 is the gain of the target energy for lower level analysis and histogram filling in the analyzer program. This target energy gain is determined such that the integral of the digitized waveform around the monoenergetic muon pulse equals 4.12 MeV. The gain variation is shown in Figure 4.10. A similar parameter is calculated again in pass 4.

After the gains of each of the degrader wedges are calibrated in pass 2, we are able to determine a Type A parameter (constant for all runs) that will correct for the light collection efficiency in the wedges. The light collected depends upon the part of the wedge through which the pion traverses and the shape of the wedge itself.

The predicted energy that the pion will deposit in the target is obtained by calculating $E_\pi^{\text{TGT}} = E_\pi - E_\pi^{\text{DEG}}$, where E_π is determined from the time-of-flight between the beam counter and the
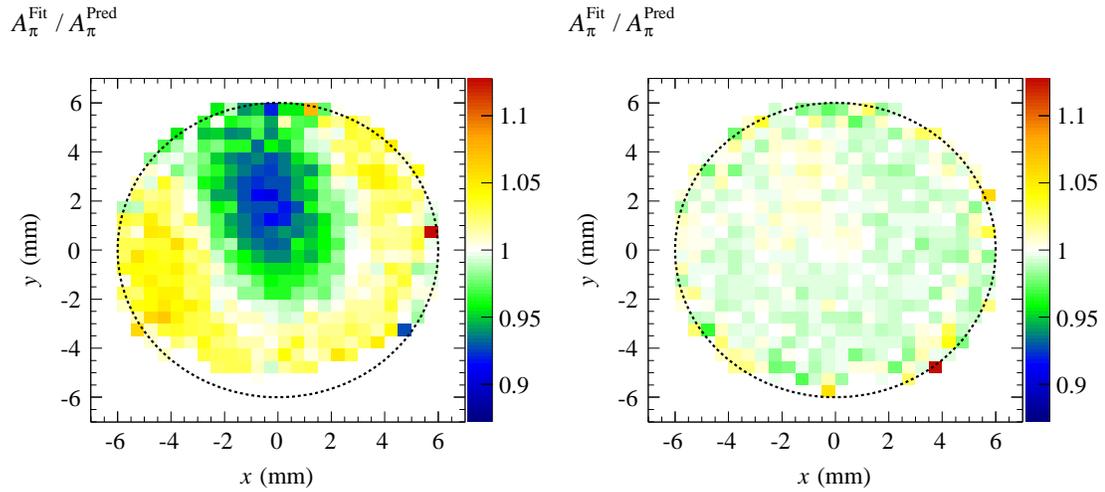
Figure 4.11: *Ratio of actual pion pulse amplitudes to predicted pion pulse amplitudes, demonstrating the extreme variation in light collection efficiency throughout the wedges of the degrader (left), and the accuracy of the predicted pulse amplitude after the correction has been applied (right). The dashed circle shows the geometrical overlap boundary among all four degrader wedges.*

degrader. This energy is then converted into the amplitude that we expect to see in the target wave-form. Taking the ratio of actual pion pulse amplitudes to predicted amplitudes for all values of x and y demonstrates the variation in light collection efficiency throughout the wedges as can be seen in Figure 4.11 (left). Including this x, y dependent correction to the degrader energy allows the pion target energy prediction to be more accurate and uniform in x, y, as shown in Figure 4.11 (right).

Calibration Pass 3

We've already determined the axial, z, coordinates of the degrader and target centers. Now we'll find the lateral offsets, x and y. To determine the center of the degrader with respect to the wire chambers, we consider only events in which the beam pion underwent a hardonic reaction in the degrader resulting in a proton track reconstructed with the MWPCs. Restricting ourselves to approximately

vertical proton tracks,

$$70 < \phi_p < 110$$

$$250 < \phi_p < 290$$

and tracks within 2 ns of the expected pion stop time, we plot the x value of the distance of closest approach to the detector axis (using wire chamber location as the reference), Figure 4.12. The mean value of this distribution gives the offset of the degrader center in the x direction for each run, Figure 4.14. The y_0 offset is found analogously from proton tracks in the horizonal directions.
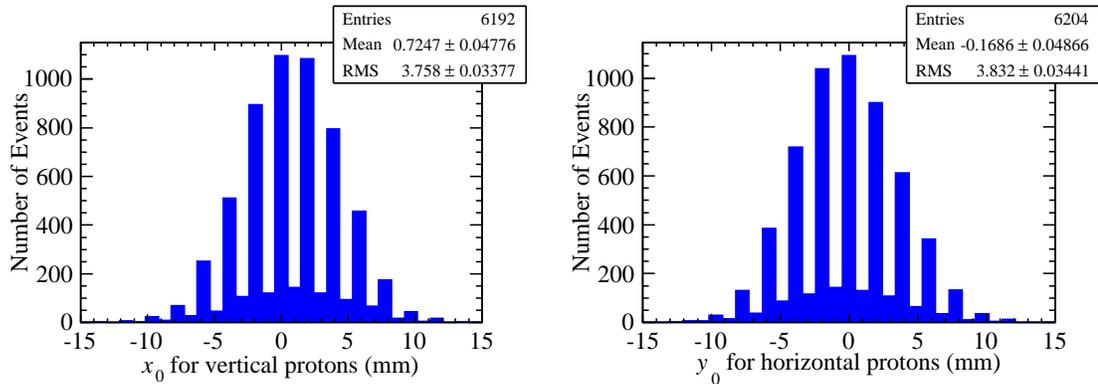
Figure 4.12: *Distributions of the x_0 of protons travelling vertically (left) and y_0 of protons travelling horizonally (right), used to determine the center of the degrader in the MWPC reference frame.*

Now that the center of the degrader is determined with respect to the MWPCs, it is trivial to find the offset of the target relative to the degrader. Again, we consider proton tracks from hadronic reactions in the degrader. This time we restrict our study to those tracks that just graze the front face of the cylindrical target, but still form tracks from two wire chamber hits. These restrictions result in the ring of events in the lateral x,y plane shown in Figure 4.13. We wrote our own circular objective function for a χ^2 minimization. For each run, we fit a circle with radius fixed at the target
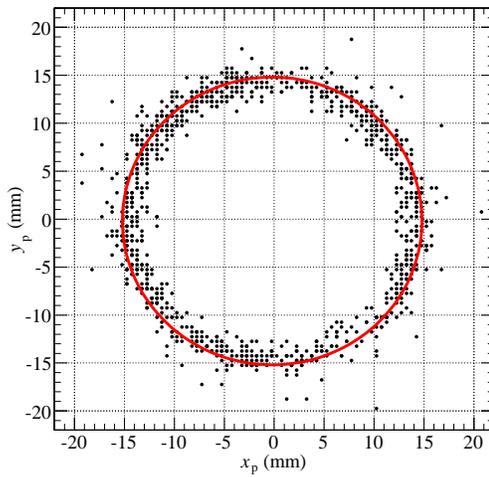
Figure 4.13: *To determine the target's position in the x,y plane, a circle with the target's radius is fit to the points where the proton tracks cross the target face.*

radius to the ring of events. The center of the circle represents the target offset with respect to the

degrader. The lateral offsets as a function of run are shown in Figure 4.15.
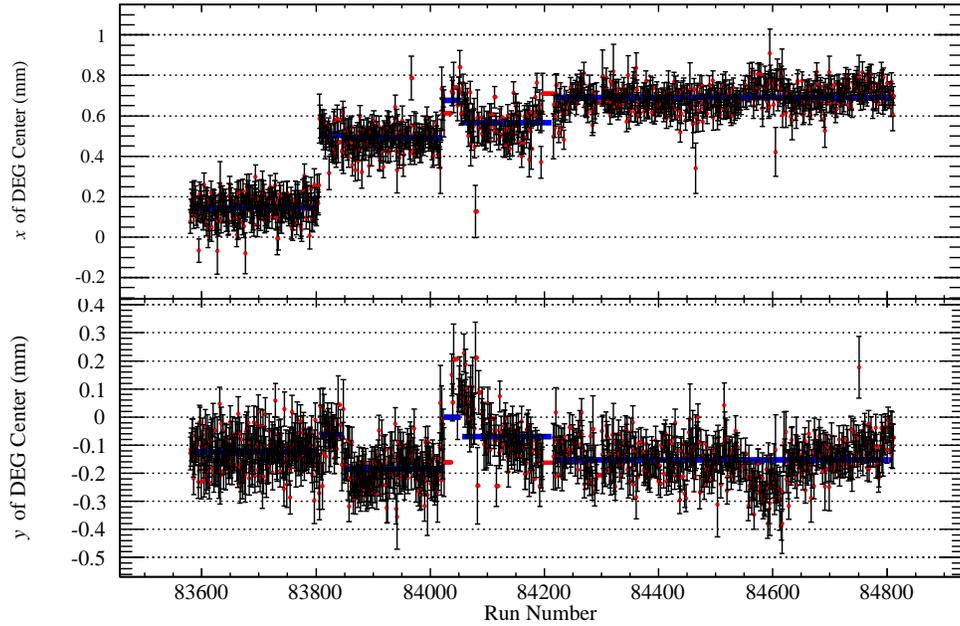
**Figure 4.14:** *Center of the degrader in the x and y directions versus run number in 2008.*
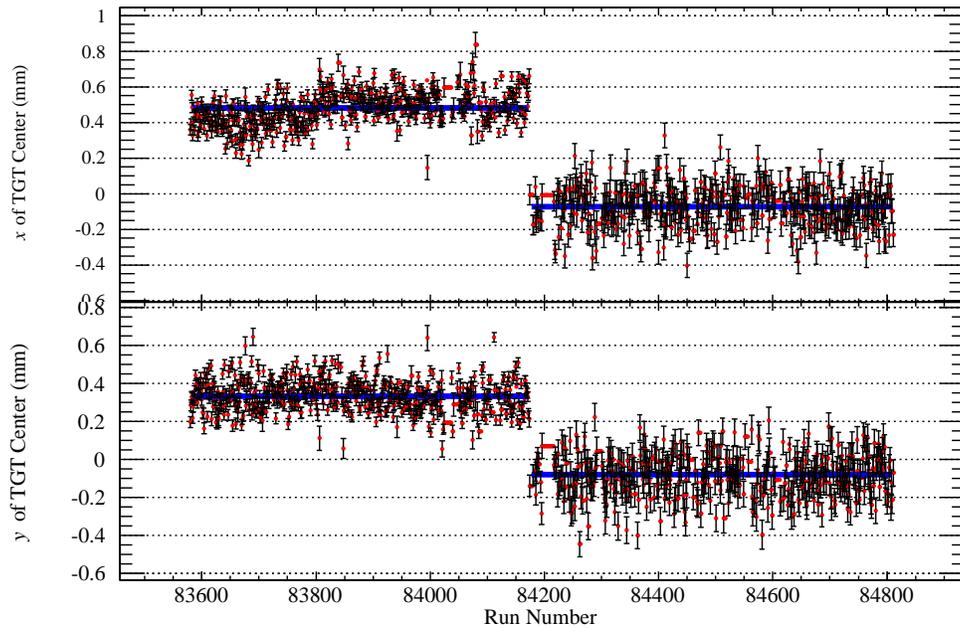
**Figure 4.15:** *Center of the target in the x and y directions versus run number in 2008.*

Calibration Pass 4

Calibration pass 4 is used to calibrate the initial values in the fit Section 5.2) to the digitized target waveform. Several fit parameters (such as t_π and t_e) are fixed to the predicted values, so having an accurate prediction is critical.

Since this pass is the first to execute fits to the waveform, we must first determine the target waveform response shape due to a charged particle. The waveform shape itself is formed mostly from collected photoelectrons and contains artifacts due to improperly terminated connections, imperfect contacts, etc., which may appear or disappear at various points during the data collection.

The task is to determine the shape of the pulse from a single event including as much of the decaying tail as possible. Pions and muons decay too quickly, and the daughter particles from the decay create peaks in the tail. Positrons trigger for the waveform recording, and occur too late in the recorded waveform to use them. We have chosen to use protons in the target, mostly created by pion-nucleon reactions in the degrader or target itself. We can cleanly select protons that occur early in the waveform, without additional pulses riding on their tails. We averaged shapes from prompt events with recorded proton target energies between 5 and 15 MeV.

To check whether the target waveform response shape changes from run to run, we used the Kolmogorov-Smirnov test. This test is a nonparametric test for the equality of continuous, one-dimensional distributions [30], thus giving us a quantitative comparison between a sample and a reference distribution. Our reference shapes were taken to be the averages of response shapes from groups of consecutive runs found to have the same response. The Kolmogorov-Smirnov probability approaches 1.0 when the distributions are the same, and 0.0 when they are very different. Using a reference waveform taken from the first group of runs in 2008, and comparing that reference waveform to every run in 2008, in Figures 4.17 and 4.18, we clearly see vast changes in the response

shape as time went on. There are seven groups of runs in 2008 with different responses as shown in Figure 4.16. The first reflection occurs after 64 ns, which suggests an imperfect impedance match at the end of a 32 ns cable; possibly a reflection back and forth to the PMT base. Another reflection is roughly at 210 ns which could originate from the positive ion signal in the PMT. Figure 4.19 shows that using the correct reference shape for each group gives a Kolmogorov-Smirnov probability of ~1.0 for the whole data set.
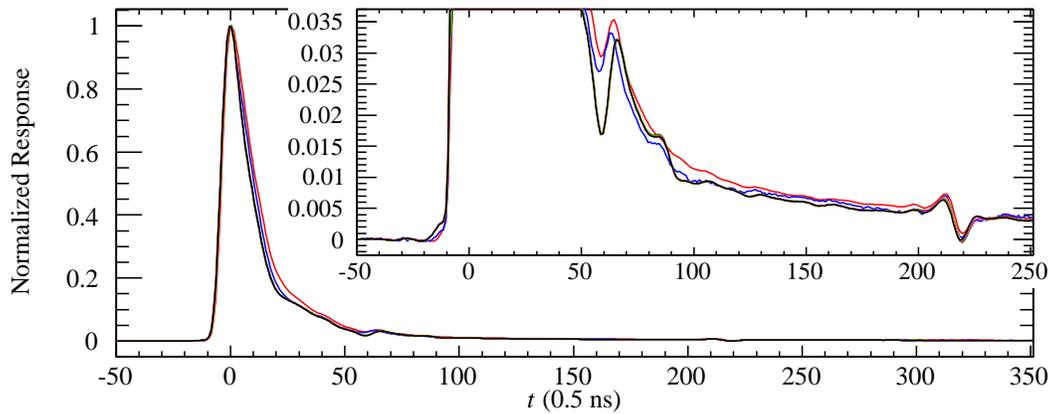
Figure 4.16: *The seven target waveform response shapes for year 2008. The inset shows the same distributions in more detail.*

Once the response shapes are known, we can design filters to transform each response into a Gaussian peak. The filtering techniques will be described in detail in Section 5.1.

In the previous calibration passes, 1 to 3, all observed quantities used to predict pulse times and amplitudes (energies) in the target waveform were calibrated. It is possible, however, to improve the predictions with an additional calibration step. This step takes into account any fluctuations of the voltage applied to the target scintillator PMT or possible modifications to the readout electronics.

We performed a three-peak fit to the filtered target waveform selecting only $\pi \rightarrow \mu \rightarrow e$ events with three well-separated pulses. This fit allows all parameters for each pulse to be free in order

to obtain the best time and amplitude of each pulse. The difference between the fit times and the predicted times gives accurate offset corrections for the pion and positron pulse times, Figures 4.20 and 4.21. The ratio of fit amplitude to predicted amplitude results in the calibration corrections to the pion and positron pulse amplitudes. In addition, we determine the scale factor by requiring the integral of the monoenergetic muon pulse to be equal to 4.12 MeV. This gain correction as a function of run is shown in Figure 4.22.

Since the filtered waveform pulses closely resemble Gaussian pulses, the option exists to fit the waveforms with analytical Gaussian pulses. As the response shapes and filters change for different run ranges, we must then determine the standard deviation, σ, of the Gaussian pulse for each particle and for each run range. The σ values are plotted versus run number in Figures 4.23–4.25.
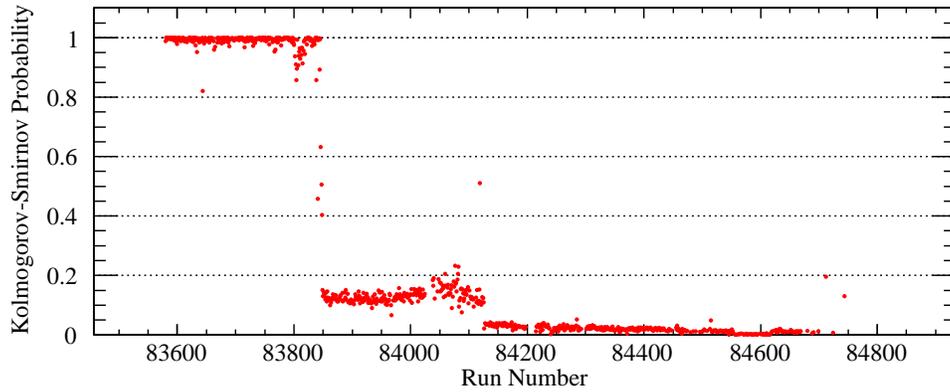
Figure 4.17: *Kolmogorov-Smirnov test for each run, using the average response shape from the first group of runs as the reference distribution.*
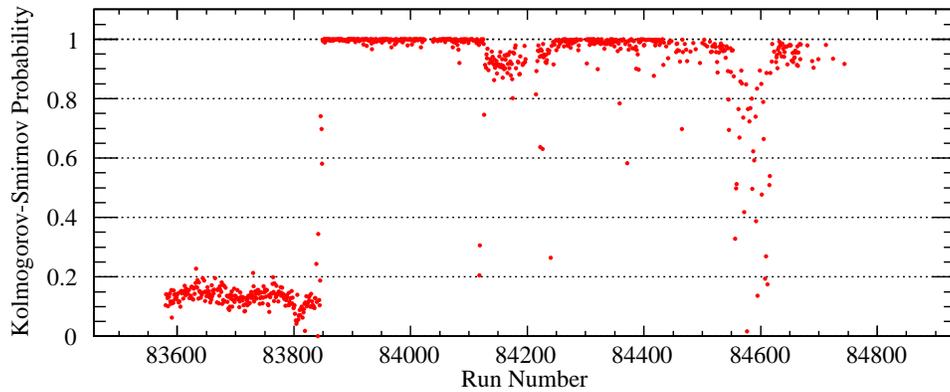
Figure 4.18: *Kolmogorov-Smirnov test for each run, using the average response shape from the third group of runs as the reference distribution. The third reference is similar to a few other later groups.*
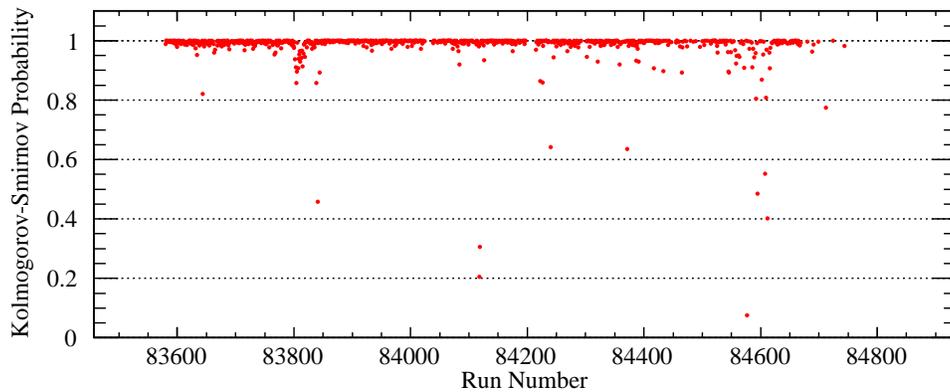
Figure 4.19: *Kolmogorov-Smirnov test for each run, using the average response shape from the appropriate group of runs as the reference distribution. Low values correspond to runs with low statistics.*
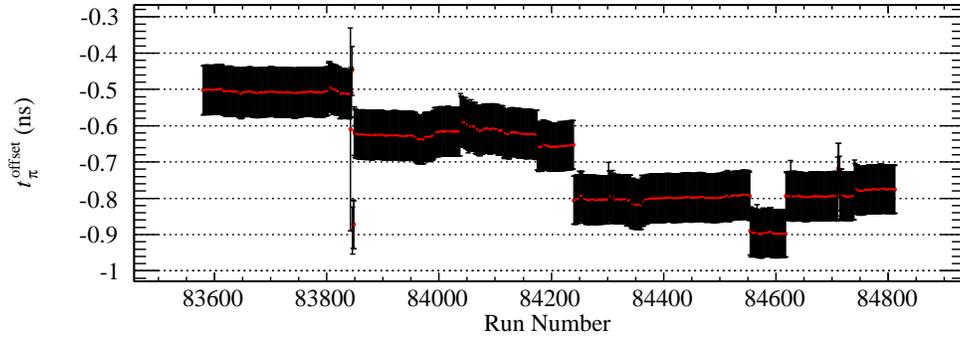
Figure 4.20: *Offset correction for t_π prediction versus run number in 2008.*
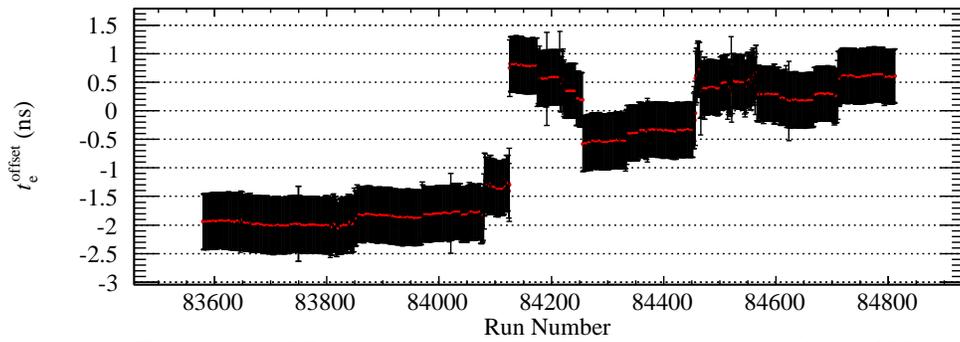
Figure 4.21: *Offset correction for t_e prediction versus run number in 2008.*
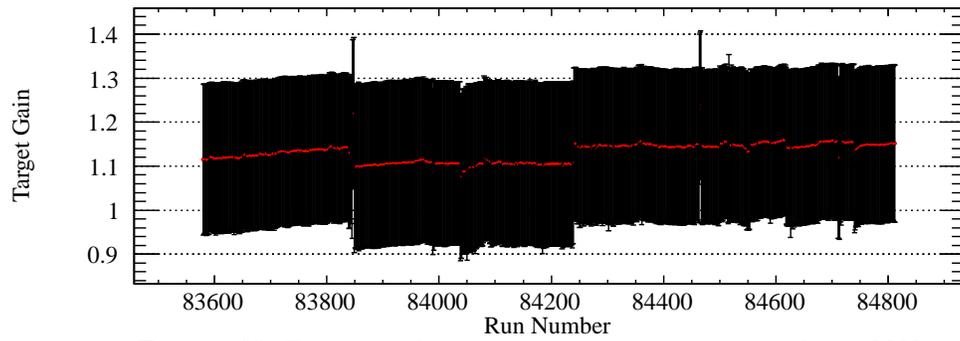
Figure 4.22: *Target waveform gain correction versus run number in 2008.*
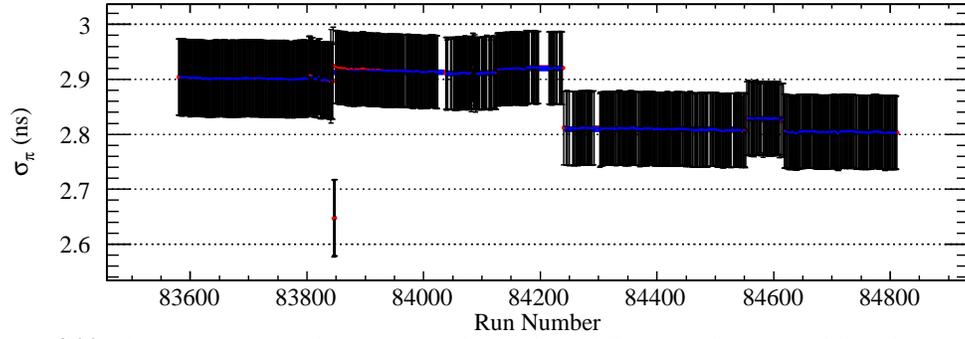
Figure 4.23: *Determination of the sigma for the analytical Gaussian best resembling the pion pulse shape.*
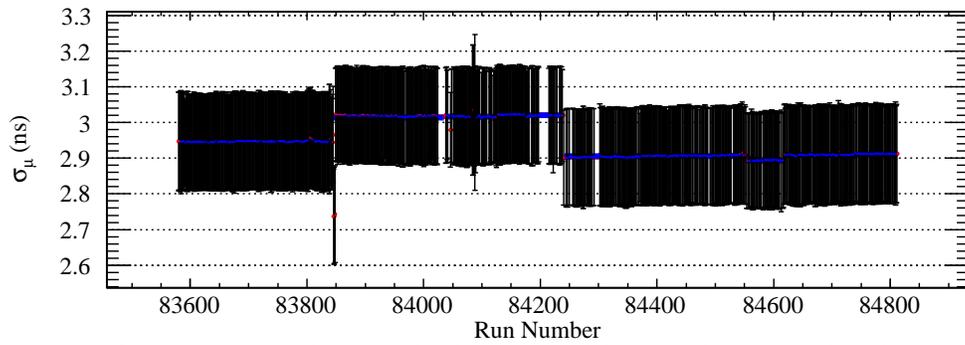
Figure 4.24: *Determination of the sigma for the analytical Gaussian best resembling the muon pulse shape.*
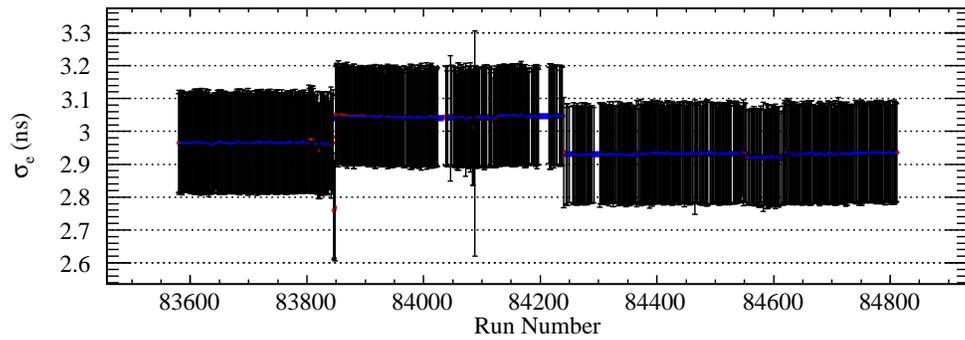
Figure 4.25: *Determination of the sigma for the analytical Gaussian best resembling the positron pulse shape.*

4.3   CsI Calorimeter Calibration

Of critical importance is the energy calibration of the CsI calorimeter. Most of the energy of the

final-state particles (usually positrons) will be deposited in the calorimter. We must therefore apply

gain correction factors to compensate for variations in high voltages applied to the photomultiplier

tube bases and variations in temperature of the crystals which affects light yield. The gain factors

translate the recorded values into physical energy deposited, in MeV.

The initial method for calibrating the CsI crystal energy gains was as follows. An uncalibrated

energy distribution for each crystal was created according to the cuts listed in Table 4.1. Two sample

distributions are shown in Figure 4.26. Since only events satisfying the high trigger logic, see (3.3),

are plotted, we have a clear indication of the neccesity of this gain calibration. The uncalibrated

energy distribution for crystal 48, Figure 4.26(a), shows the threshold around 53 MeV, while crystal

49, Figure 4.26(a), appears to have the threshold set around 50 MeV. The apparent trigger threshold

level is dependent on both the high voltage applied the the photomultiplier tubes and the light yield

of the individual crystals.
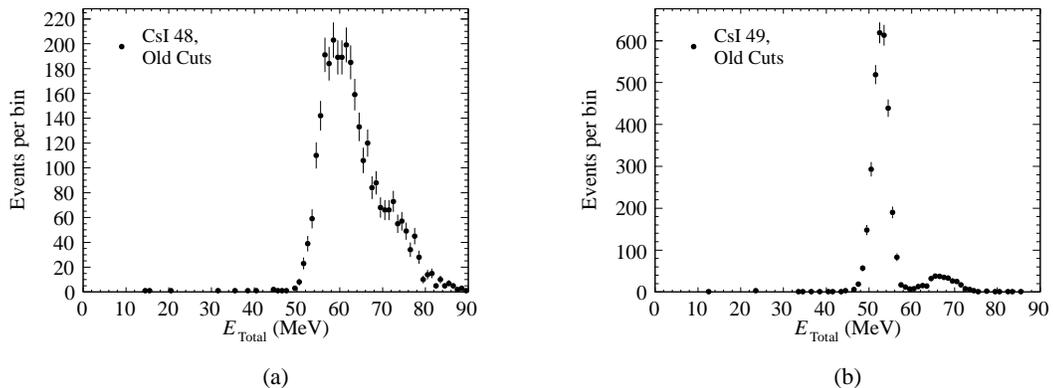
Figure 4.26: *Initial CsI energy calibration method for crystals* (a) *48, and* (b) *49, demonstrating the lack of precision in the method for crsytal 48.*

This initial method then set the bin content of all bins with bin content less than 10% of the

Table 4.1: *The old cuts for finding the upper Michel energy edge for use in calibrating the CsI crystal energies.*

Cut	Description
Crystal ID($E_e^{\text{CsI,max}}$)	Max. Energy is deposited in crystal of interest
$TR_{\text{High}} = 1$	Only high trigger events, see (3.3)
$N_{\text{DecayTrack}} = 1$	One decay track
Decay Particle ID = e$^+$	Only consider positrons
$L_e^{\text{TGT}} > 0$	Require a valid target pathlength
$E_e^{\text{PH}} > 0$	Require energy in hodoscope

maximum bin content to zero in hopes that the only remaining peak was that resulting from the upper edge of the Michel energy shape. The peak from the monoenergetic $\pi \rightarrow$ e positron was thus ignored. The resulting distribution was fit with a Gaussian and the mean value for the Gaussian fit for each crystal determined the gain correction factor to be applied such that the Michel positron upper energy endpoint was the same for all crystals.

As you can see from Figure 4.26(a), this method was not very precise. The intrinsic resolution of some crystals was just not good enough, and problems arose when the high voltage setting for a particular crystal was significantly far from an ideal setting. It also requires a stable trigger threshold.

A new method for calibrating the CsI crystal energies was devised using the target waveform analysis techniques that will be discussed in Chapter 5. Using fits to the target waveform we were able to get clean energy distributions from the monoenergetic $\pi \rightarrow$ e positron thus obtaining the best physical quantity with which to calibrate our detectors.

Now considering the same crystal 48, for the same range of runs, we imposed the $\pi \rightarrow$ e selection cuts in Table 4.2(b) to create the distribution with the sharp monoenergetic peak shown in Figure 4.27(b). Unfortunately these cuts could not completely isolate the $\pi \rightarrow$ e process. Since that was the case, the distributions were sometimes contaminated with $\pi \rightarrow \mu \rightarrow$ e events and produced the

peak of Michel events leaking into the high trigger sample. Figure 4.28(b) shows crystal 8 as an example. To overcome this drawback, we obtained a sample $\pi \to \mu \to e$ distribution using the cuts in Table 4.2(a). This $\pi \to \mu \to e$ energy distribution was then scaled down by a factor 0.003 and added to the $\pi \to e$ distribution, producing the combined distributions shown in Figures 4.27(c) and 4.28(c).

The resulting combined energy distributions were fit with a function, f, consisting of two parts: a Crystal Ball function [68, 59, 38] , f_{CB}, and an analytical approximation to the sequential $\pi \to \mu \to e$ decay energy distribution, $f_{E_{\pi \to \mu \to e}}$. The Crystal Ball function consists of a Gaussian peak with mean, $\overline{E}_{Total}$, and sigma, σ, and a power law tail. The Crystal Ball function is described in more detail in Section 6.5.1. The function describing the $\pi \to \mu \to e$ energy is,

$$f_{E_{\pi \to \mu \to e}} = N$$

$$\times \begin{cases} \exp\left(E_{Total}\right) & \text{for } E_{Total} \leq 2\,, \\[2mm] \left| (E_{Total} - p_1)^2 \left(3 - \frac{2E_{Total}}{p_2}\right) \left(1 - \mathrm{erf}\left(\frac{E_{Total} - p_3}{p_4}\right)\right)\right| & \text{for } 2 < E_{Total} \leq p_3\,, \\[2mm] 1000 \exp\left(p_3 - E_{Total}\right) & \text{for } p_3 < E_{Total}\,, \end{cases} \tag{4.5}$$

where N is the normalization, p_i are parameters governing the shape, and erf( ) is the error function. In particular, p_3 is related to the upper Michel positron energy endpoint. Typical values for the parameters are: $p_1 = 3.36$ MeV, $p_2 = 44.23$ MeV, $p_3 = 53.21$ MeV, and $p_4 = 1.37$ MeV.

The new method is much more reliable and has better fit convergence. Nevertheless, due to limited event statistics in several crystals for short run ranges between high-voltage changes we must visually inpect each fit. With this method, the mean of the Gaussian part of the Crystal Ball function, i.e., the $\pi \to e$ energy peak, is used to match the gains.

Table 4.2: *The new cuts for finding the* (a) $\pi \to \mu \to e$ *and* (b) $\pi \to e$ *energy distributions. The two distributions are then added to form a combined distribution which is fit with a Crystal Ball and Michel energy function.*

(a)

Cut	Description
Crystal ID($E_e^{\text{CsI,max}}$)	Max. Energy is deposited in crystal of interest
$\Delta\chi^2 < 0$	More like a 3-peak waveform, see (5.16)
$\chi^2_{\text{3-peak}} < 6$	Reasonable 3-peak fit
$N_{\mu\text{ candidates}} = 1$	Restrictive cut removing background
$N_{\text{DecayTrack}} = 1$	One decay track
Decay Particle ID $= e^+$	Only consider positrons
$L_e^{\text{TGT}} > 0$	Require a valid target pathlength
$E_e^{\text{PH}} > 0$	Require energy in hodoscope

(b)

Cut	Description
Crystal ID($E_e^{\text{CsI,max}}$)	Max. Energy is deposited in crystal of interest
$\Delta\chi^2 > 0$	More like a 2-peak waveform, see (5.16)
$\chi^2_{\text{2-peak}} < 6$	Reasonable 3-peak fit
$N_{\text{DecayTrack}} = 1$	One decay track
Decay Particle ID $= e^+$	Only consider positrons
$L_e^{\text{TGT}} > 0$	Require a valid target pathlength
$E_e^{\text{PH}} > 0$	Require energy in hodoscope
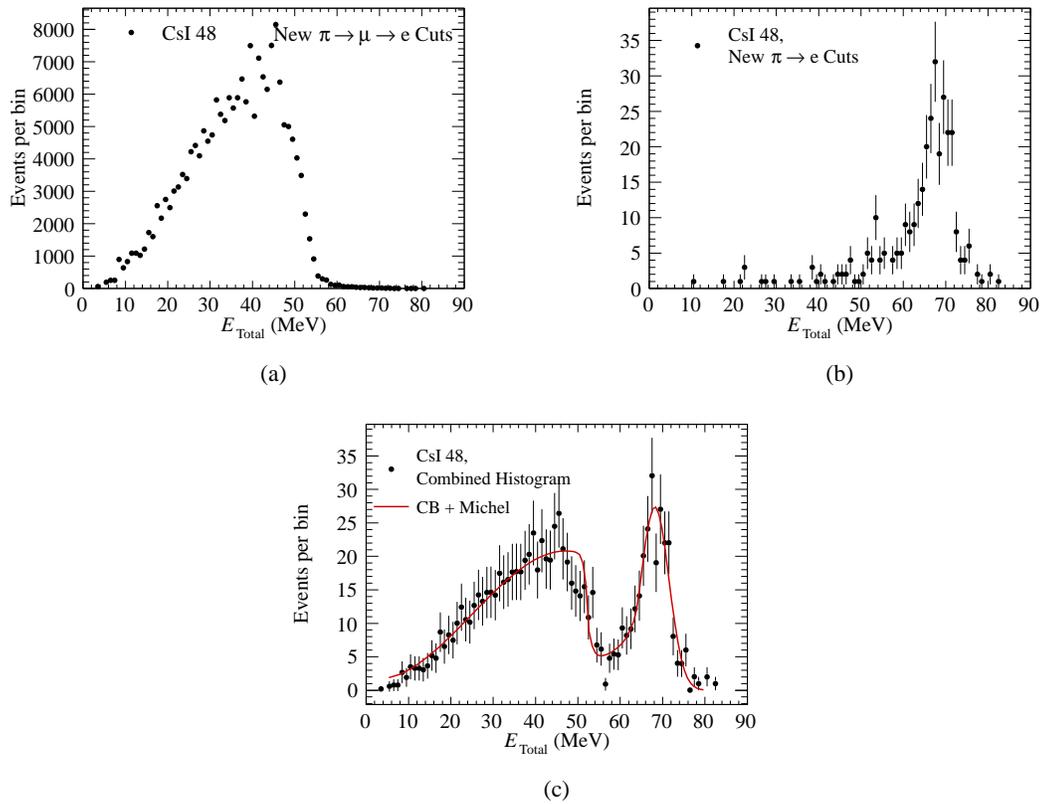
Figure 4.27: *New CsI Calibration for crystal 48.* (a) *with cuts selecting Michel events, and* (b) *with cuts selecting* $\pi \to$ e *events.* (c) *combination* (a) *and* (b) *along with a fit.*

(a)

(b)

(c)
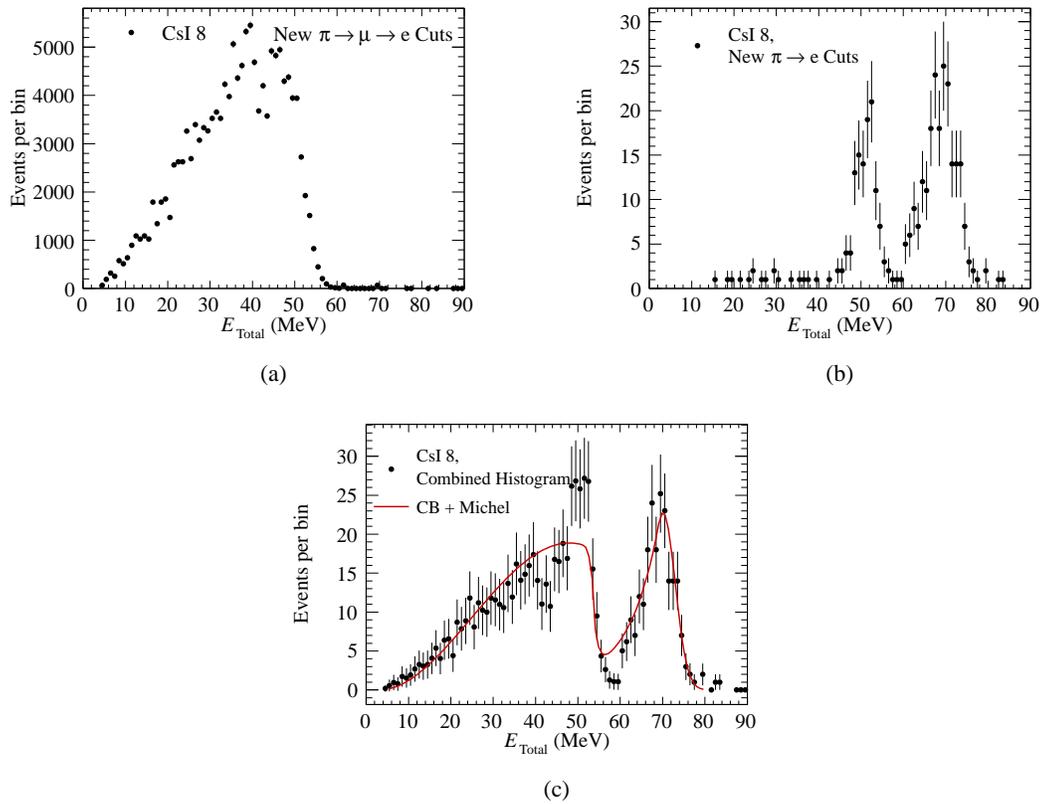
Figure 4.28: *New CsI Calibration for crystal 8.* (a) *with cuts selecting Michel events, and* (b) *with cuts selecting π → e events.* (c) *combination* (a) *and* (b) *along with a fit.*

Chapter 5

Target Waveform Analysis

Thanks to continuing advances in waveform digitizer performance and increasingly powerful data
acquistion systems it has become common practice to record complete signal traces of particle
detectors. The knowledge of the full waveform offers many advantages compared to traditional
readout schemes using charge integrating amplifiers and fixed-threshold timing discriminators cou-
pled to amplitude and time digitizers (ADCs and TDCs). Both low frequency baseline fluctuations
and high frequency noise (oscillations) can be taken into account. Particle identification by pulse
shape discrimination becomes straightforward. Deadtime-free time coincidences between different
detectors are implemented by direct binwise multiplication of waveforms. Overlapping signals re-
sulting from pulse pile up or decay sequences can be reconstructed. A large portion of this chapter's
contents is also presented in a soon to be published paper [60].

A measurement of the π_{e2} branching ratio with 0.05% relative precision requires categorical
identification of the pion decay mode. Due to the energy resolution of the PEN detector, a sig-
nificant portion of the $\pi \to e$ events overlap with the several orders of magnitude more abundant
Michel positron events. As will be discussed later (Section 6.5), we need to find precise probability

70

distribution functions for each observable in our final analysis, and for every process existing in the data. We will employ an in-depth analysis of our digitized target waveform signals.

In Section 5.1 we present the finite impulse response (FIR) filtering of waveforms. The method has been applied before to waveforms with much lower sampling frequencies, for example in the treatment of echoes in seismology [65] and acoustics [4], but is so far practically unknown in the particle physics community.

In Section 5.2 we discuss algorithms applied to the resulting filtered waveforms in order to identify overlapping signals. Fixed χ^2 fit parameters account for constraints on time and amplitude of the signals involved. Such constraints on the fit parameters may be intrinsic (as for the muon amplitude in $\pi^+ \rightarrow \mu^+ \nu_e$ decays at rest) or dependent upon additional information supplied by other detector subsystems. In Section 5.2.1 we explain the χ^2 fitting methods used in the PEN experiment analysis. We demonstrate that different classes of events can be reliably reconstructed, even when signal pulses occur simultaneously.

5.1   Waveform Filtering

Particle detectors produce signals with intrinsic shapes depending not only on the type of detector but also possibly varying with type of particle, the temperature, chemical contaminations, etc. Here we focus on fast plastic scintillators (with a main decay component of 2–3 ns). The intrinsic signal rise time is usually very short, on the time scale of the photon detector response, and can thus be ignored.

In practice, the observed analog signals may show considerable deviations from the intrinsic light response, mainly because of variations in the path length of the scintillation photons until detection with PMTs. The time it takes the scintillation light to reach the PMTs depends on the

detector geometry, the wrapping or coating, a possible wavelength shifter, and the refractive indices involved. Photons reaching the PMT photocathode are distributed statistically around the envelope of the primary detector signal and these fluctuations have to be largely integrated for optimal energy resolution.

Further distortions are caused by the response of the photocathode (transit-time spread of photomultipliers, after pulsing), amplifiers (finite bandwidth, ringing), transmission lines, or impedance mismatching. These effects show no statistical fluctuations and can thus be perfectly removed by signal filtering.

Filtering is a form of signal shaping which can be described by a convolution integral,

$$a'(t') = \int_{-\infty}^{t'} f(t'-t) \cdot a(t)\, dt \ ,$$ (5.1)

with $a(t)$ the input signal and $a'(t')$ the output signal of the filter $f(t'-t)$. Most familiar are simple passive filters based on RC networks or clip cables for which analytic expressions can be derived in many cases. Such analog filters are limited to real time applications.

For digitized signals the convolution integral turns into a sum, and (5.1) translates into a vector equation,

$$w'[n] = \sum_{m=n-M_1}^{n+M_2} f[n-m] \cdot w[m] \ ,$$ (5.2)

where vectors $w[m]$ and $w'[n]$ represent the digitized input and output waveforms, respectively. The filter array f has been truncated to M_1 forward bins and extended to include M_2 bins describing tails "running back in time" (known as acausal filtering, see below). The choice for the array boundaries M_1 and M_2 depends on the application. In case of offline software filtering the waveforms themselves are finite in most cases and care has to be taken that the recorded regions next to the signal

region of interest are sufficiently wide to accomodate the filter array.

The filter described by (5.2) is known in the literature as a finite impulse response (FIR) filter of order $M = M_1 + M_2 - 1$. FIR filters with all but two coefficients equal to zero can be used to remove the reflection caused by an incorrectly terminated transmission line. Low order electronic FIR filters are used in real time to suppress multiple reflections (ghost signals) in TV broadcasting [70]. FIR filters of order 20–40 have been realized in field-programmable gate arrays (FPGA) to remove unwanted reflections in radar signals [51].

Higher order FIR filtering is so far limited to offline data processing applications. In those cases the filter coefficients can be carefully calibrated and adjusted when conditions change.

5.1.1   Waveform Filtering in PEN Analysis

The PEN experiment data analysis requires distinguishing $\pi^+ \to e^+$ events from the $\pi^+ \to \mu^+ \to e^+$ decay chain. Digitized waveforms for a large set of pion decay events must be reliably sorted into one of the two categories in order to reveal the low-energy "tail" of the calorimeter response to the 69.3 MeV $\pi^+ \to e^+$ positrons, otherwise masked by the positrons from muon decay. In this experiment the muons live entirely inside the target. Therefore we rely heavily on digitized target waveform analysis to distinguish between the two pion decay modes.

We use data recorded by the PEN experiment to illustrate the method of filtering described in the previous section. In this experiment an ~ 85 MeV/c π^+ beam passes through a thin beam counter and a four wedge-shaped degrader counters before stopping in a plastic scintillating target. The scintillation light from each of the six detectors is recorded with waveform digitizers[1] for offline analysis. A clean measurement of the target waveform response function is obtained by selecting

---

[1]Acqiris High Speed 10-bit PXI/CompactPCI Digitizer, Model DC282. 4 Channels, each with 2 GS/s.
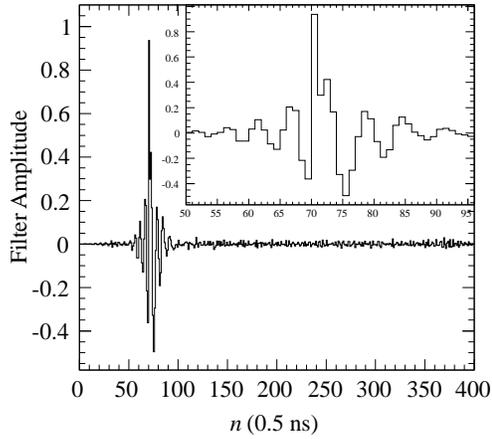
Figure 5.1: *Filter Array, f [n] used to shape target waveform.*
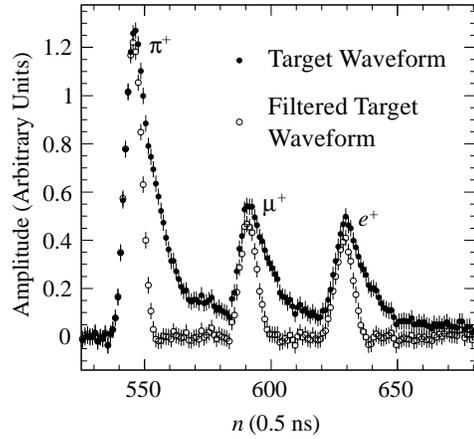
**Figure 5.2:** *Original and filtered target waveforms, showing the $\pi \to \mu \to$ e decay sequence with slightly overlapping pulses in the raw waveform, and completely isolated pulses after filtering.*

pion reactions with protons in the final state since these events lead to a single target signal without delayed decay products (see Figure 5.3). The determination of the response shape and its variation during data collection is discussed in more detail in Section 4.2.

The raw waveform w not only shows various exponential components and reflections in the trailing edge but also a small "shoulder" on the rising edge. The signal falls much slower than expected from the primary detector signal, so its shape is completely dominated by the impulse response of the system. The shoulder has been observed before in photomultiplier signals and is explained by electric cross talk in the final dynodes of the photomultiplier which reaches the anode electrode ahead of the slower main signal.

The filter coefficients were numerically optimized in an iterative procedure to produce a Gaussian output waveform w' of given sigma $\sigma_{w'}$. A Gaussian shape is advantageous for the following fitting procedures. The width of the resulting Gaussian distribution was minimized to the point where the energy resolution starts to deteriorate. As is illustrated in Figure 5.4 this situation was

reached at $\sigma_{w'} \sim 3.0$, which corresponds precisely to the σ of a Gaussian fit to the rising edge of the unfiltered pulse.
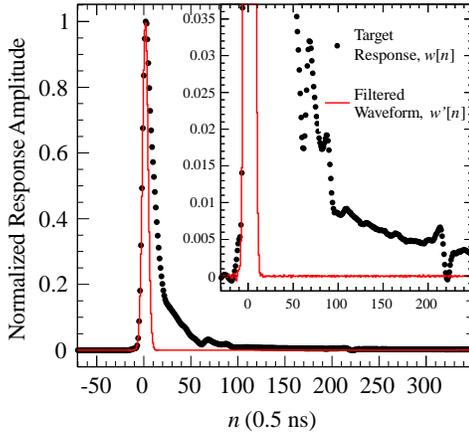
Figure 5.3: *Averaged target waveform response shape (●) and the result after filtering (—). The inset shows the reflections in the tails in more detail.*
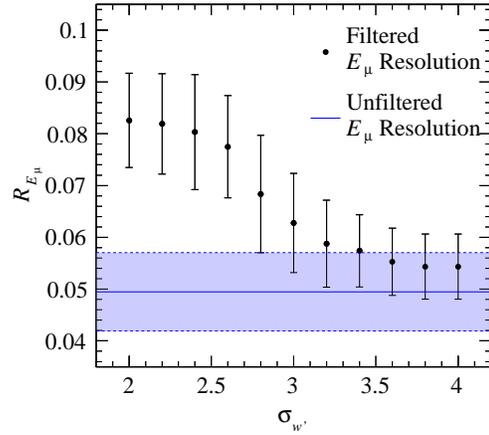
Figure 5.4: *Muon energy resolution versus sigma of the Gaussian filter output. The horizontal band illustrates the resolution ($\sigma = 4.95\%$) obtained with the raw, unfiltered waveforms.*

Figure 5.1 shows the corresponding filter array f with order 430 ($M_1 = -80$, $M_2 = 349$). Around $n = 80$ the distribution resembles the $\mathrm{sinc}(n) \equiv \frac{\sin n}{n}$ wavelet which describes the ideal low-pass filter associated with our 2 GHz sampling frequency. Whereas the periodicity of the filter coefficients is maintained over the full array, the amplitude modulation deviates strongly from the $1/n$ dependence in the sinc function, reflecting the irregularities in the trailing edge of the input waveform.

The filter, f, was developed using an iterative procedure. The initial filter was an array of zeros with the exception of one entry which was equal to 1, which when applied to the waveform did not change it. The difference, d, between the filtered response shape, w', of our system and the desired

Gaussian shape, g, was used to build the filter as in (5.3):

$$d[n] = w'[n] - g[n] \ ,$$

$$f'[n] = f[n] + (0.1)(d[n]) \ . \tag{5.3}$$

The modified filter is then applied to the response shape as in (5.2) and (5.3) are performed again. This procedure is repeated until the sum of the differences at each point in the array, $\sum_n |d[n]|$, asymptotically reaches a minimum value.
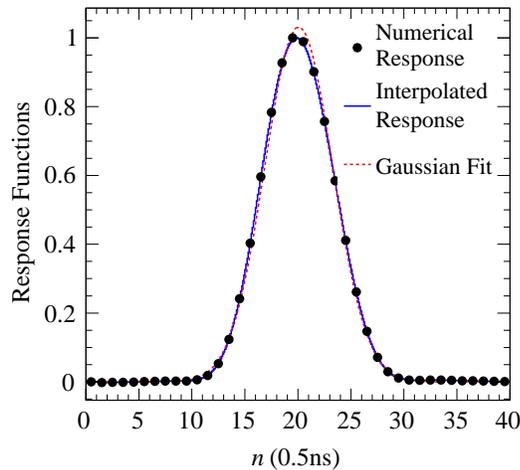
Figure 5.5: *Filtered target waveform response function and Gaussian fit, demonstrating a slight discrepancy near the peak.*

After filtering, the numerical reponse function is determined from events with well separated pulses. Initially, Many pulses are averaged; the resulting waveform is subsequently interpolated to produce a response function with a tenfold increase in the number of bins. Figure 5.5 shows the original averaged response pulse, the interpreted response pulse, and a Gaussian fit. The resulting filtered pulse shape is not exactly Gaussian, necessitating the use of a numerical response function in the minimization. The minimization process does not converge properly when the theoretical

pulse shape has discrete heights, therefore the continuous height of the pulse is calculated as

$$A(t) = A_{i=\lfloor t \rfloor} + c(t - \lfloor t \rfloor - 0.5), \tag{5.4}$$

where the coefficient

$$c = \begin{cases} \frac{A_{i=\lfloor t \rfloor + 1} - A_{i=\lfloor t \rfloor}}{2} & \text{for } t - \lfloor t \rfloor \leq 0.5, \\[2ex] \frac{A_{i=\lfloor t \rfloor} - A_{i=\lfloor t \rfloor - 1}}{2} & \text{for } t - \lfloor t \rfloor > 0.5, \end{cases}$$

and the notation $\lfloor t \rfloor$ denotes the floor function which maps the real number t to the largest integer not greater than t.

5.2 Waveform Fitting

The filtered waveform contains an unknown number of detector signals, each parametrized by their shape, time, and amplitude. Without additional constraints on amplitudes, a single pulse can always be interpreted as a pile up of an unknown number of simultaneous signals. Pile up signals may originate from a single event with more than one particle in the final state or from random coincidences (in particular with pulsed beams). For the correct interpretation additional constraints on the allowed values of the free parameters are required: the *a priori* probability distributions are no longer constant but peak at some prefered value. Signal times may be deduced from other detectors hit by the same particle. Probabilities for signal amplitudes may follow theoretical prejudice, e.g., an expected energy-loss distribution.

A further reduction of ambiguities can be achieved by considering only the signal sequences expected from a number of plausible hypotheses. Each of them would not only give the values of the associated free parameters, but also a measure of the likelihood for the observed waveform

which can be used to discriminate between these different interpretations.
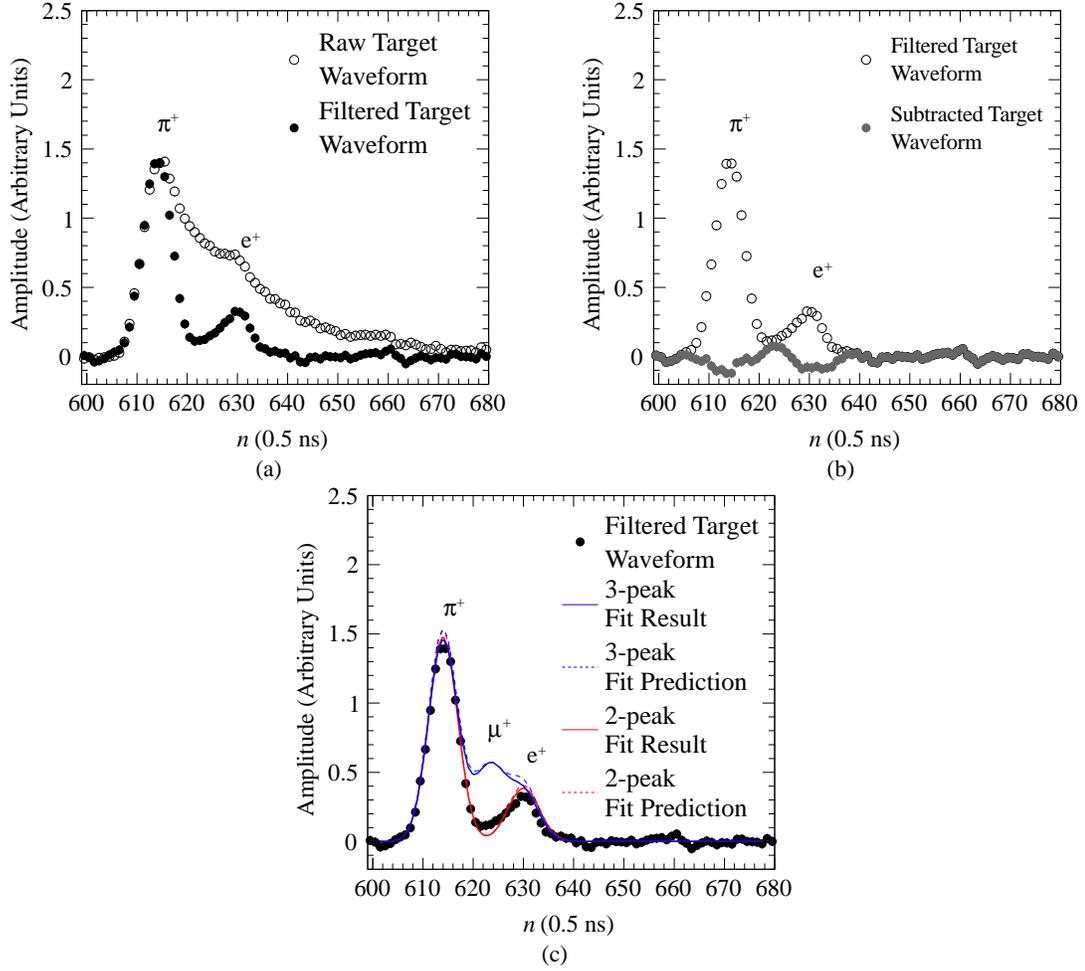
Figure 5.6: *Target waveform of a $\pi^+ \rightarrow e^+$ event.* (a) *raw and filtered waveforms,* (b) *filtered waveform before and after subtraction of predicted π and e pulses, and* (c) *fits for both the 2-peak and the 3-peak hypothesis.*

Traditionally, the waveforms are analyzed by the method of least squares which minimizes the reduced χ^2 value with respect to the values of the free parameters p_i in the model. In this case it would read:

$$\chi^2_{\text{red}} \equiv \frac{1}{N - N_f} \sum_{n=1}^{N} \left(\frac{w[n] - w^{\text{fit}}[n]}{\sigma_{w[n]}} \right)^2 \, , \qquad (5.5)$$

with N_f the number of free parameters of the hypothesis under study. Here the error $\sigma_{w[n]}$ is mostly limited by the digital noise of the digitizer with little dependence on $w[n]$. The effective number of

bits (ENOB) of present day digitizers depends strongly on the sampling rate. At frequencies above

1 GHz the ENOB drops below 8 even when the supplied number of bits can be much higher. A gen-

eralization of (5.5) can include correlations between neighboring channels. Such an approach might

give some improvement in fit performance but otherwise leads to a large increase in computing time.

5.2.1   Waveform Fitting in PEN Analysis

Most events recorded in the digitized target waveforms contain pulses which overlap. After filtering,

as in Figure 5.6(a), the two hypotheses (possible decay modes) are fit to the waveform using the

minimum χ^2 technique, as in Figures 5.6(c) and 5.7(c). The fits use information gathered in other

detectors to accurately predict and constrain the times and energies (amplitudes) of the π^+ and e^+
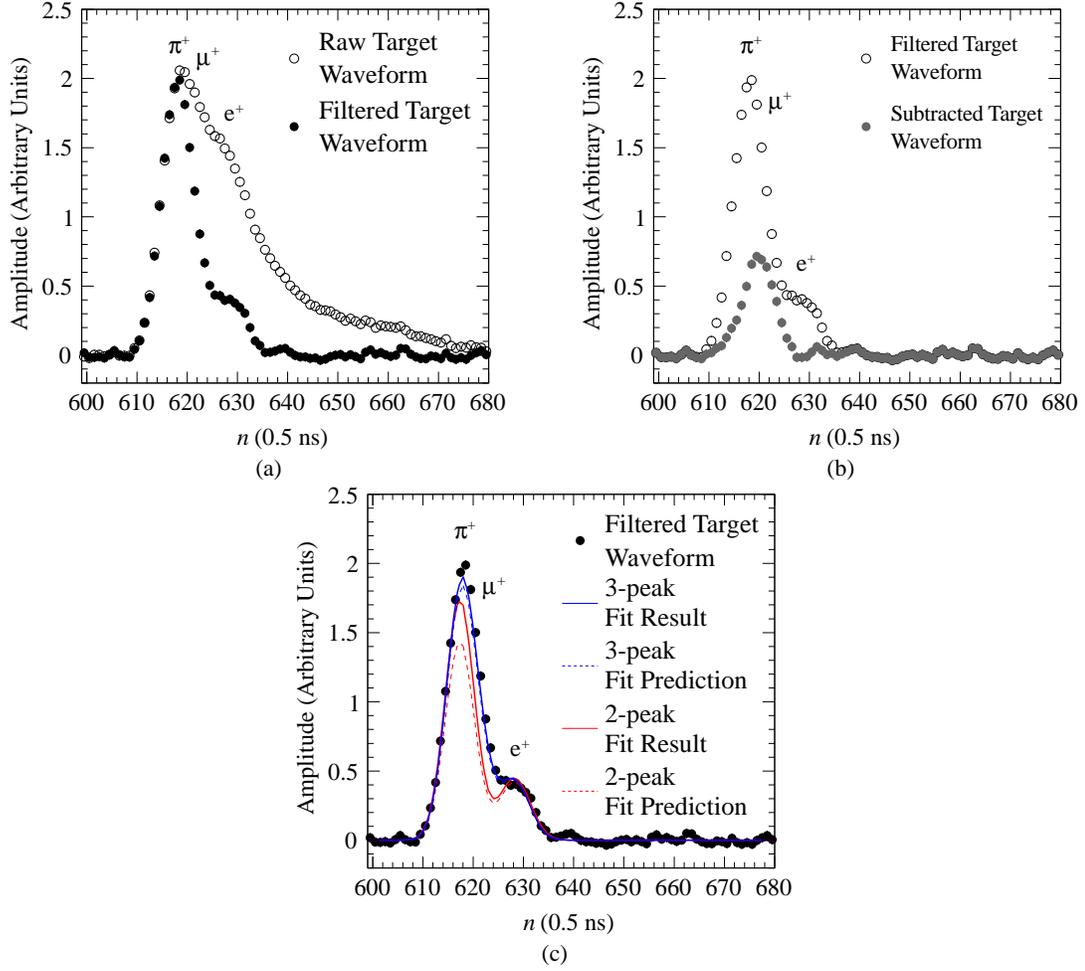
pulses.

Figure 5.7: *Target waveform of a* $\pi^+ \to \mu^+ \to e^+$ *event.* (a) *raw and filtered waveforms,* (b) *filtered waveform before and after subtraction of predicted* π *and* e *pulses and* (c) *fits for both the 2-peak and the 3-peak hypothesis.*

Predicted Parameters

For each target waveform fit (each event) there exists six initial parameters. The monoenergetic muon amplitude (4.12 MeV) is fixed to $A_{\mu^+}^{\text{Pred}}$ with $\sigma_{A_{\mu^+}} \sim 216$ keV, see Figure 5.4. This 5% measured resolution demonstrates the performance of the active target counter. Allowing the muon amplitude any freedom weakens the ability to resolve the pion and positron, especially when the pulses are overlapping. The pion and positron times, t_π^+ and t_e^+ respectively, are predicted with such high precision from external detectors ($\sigma_{t_{\pi^+}} \sim 65$ ps, $\sigma_{t_{e^+}} \sim 492$ ps) that we fix the parameters to

the predicted values.

The pion time in the waveform, t_π, is predicted as,

$$t_\pi^{\text{Pred}} = \bar{t}_\pi^{\text{wAD}} + \frac{s}{\beta_\pi c} \ ,$$ (5.6)

where s is the distance between the degrader and the pion decay vertex in the target. We measured s/c to be 0.18 ns. The most precise time of the pion in the wedged degrader is given by the pion time in each of the four wedges, weighted by the energy deposited in the wedges,

$$\bar{t}_\pi^{\text{wAD}} = \frac{1}{2} \left(\frac{E_\pi^{\text{L}} t_\pi^{\text{L}} + E_\pi^{\text{R}} t_\pi^{\text{R}}}{E_\pi^{\text{L}} + E_\pi^{\text{R}}} + \frac{E_\pi^{\text{T}} t_\pi^{\text{T}} + E_\pi^{\text{B}} t_\pi^{\text{B}}}{E_\pi^{\text{T}} + E_\pi^{\text{B}}} \right) \ ,$$ (5.7)

where L, R, T, and B, correspond to the left, right, top, and bottom wedges, respectively. The factor β_π corresponds to,

$$\beta_\pi = \frac{t_\gamma^{\text{TOF}}}{t_\pi^{\text{TOF}}} \ ,$$

where t_i^{TOF} is the time of flight of particle i between the upstream beam counter and the degrader.

The predicted positron time in the target waveform is,

$$t_{\text{e}^+}^{\text{Pred}} = t_{\text{e}^+}^{\text{PH}} - (0.6 \text{ ns}) \left(\frac{1}{\sin(\theta)} - 1 \right) \ ,$$ (5.8)

where $t_{\text{e}^+}^{\text{PH}}$ is the time the positron hits the plastic hodoscope as recorded by two TDCs[2]. The time is taken as the mean time recorded from both ends of the hit hodoscope stave. The factor of 0.6 ns is the difference in the time of flight of the positron from the target to the furthest point on the cylindrical hodoscope and from the target to the closest point on the hodoscope.

---

[2]Lecroy Model 1877 Fastbus 96 channel multihit time-to-digital converters [31]

The next two predicted quantities take into consideration the precision to which the π^+ and

e^+ amplitudes are known. The energy of the beam pion is determined from plastic scintillating

detectors in the beamline. Subtracting the energy deposited in the upstream beamline elements from

the energy determined from the pion time-of-flight between them results in an amplitude prediction

(typically ~ 13 MeV) with an uncertainty $\sigma_{A_\pi} \sim 716$ keV. The pion energy prediction is written as,

$$E_\pi^{\text{Pred}} = \gamma m_\pi - \sum_{i=\text{L,R,T,B}} E_\pi^{\text{DEG},\,i} \,, \tag{5.9}$$

where $\gamma = 1/\sqrt{1 - \beta_\pi^2}$. We found the relation between energy and amplitude for pions to be $A_\pi^{\text{Pred}} =$

$E_\pi^{\text{Pred}}/7.732$, which was constant throughout data collection and takes care of the light quenching

factor.

The amplitude of the positron pulse in the target waveform is directly related to the path length

of the positron in the cylindrical target, $L_{e^+}^{\text{TGT}}$,

$$A_e^{\text{Pred}} = k L_e^{\text{TGT}}, \tag{5.10}$$

where $k = 1.79$ MeV/cm is the coefficient including the specific energy loss, dE/dx, of the positron

and the relation between energy lost and amplitude of the pulse. Due to geometrical effects and

the nature of the algorithms used to recontruct the path length, the difference of the positron pulse

amplitude from the fit to the predicted positron pulse amplitude resembled a Vavilov distribution

smeared by photoelectron statistics. The usage of minimum χ^2 fits in data analysis assumes the

fit parameters follow normal (Gaussian) distributions. Therefore our fitting analysis will work best

when our parameters each follow a Gaussian distribution.

Taking the logarithm of the amplitude ratio makes this distribution symmetric about zero. Fur-

thermore, multiplying by the square-root of the path length helps correct for a slight path-length correlation. In total, these modifications result in symmetric predictions and allow us to calculate the uncertainty of the prediction to be $\sigma_{A_e} \sim 878$ keV, where the typical positron energy in the target is roughly 2–3 MeV. It has been found that fixing A_π and A_e to their predictions gives better results.

L_e^{TGT} is calculated from the tracking of the incoming beam pion and the outgoing decay positron. The componenets of the vertex of the pion stopped in the target, $\mathbf{v}_\pi$, are given by

$$x_\pi^{\text{wAD}} = k_{\text{LR}} \frac{E_{\text{Left}} - E_{\text{Right}}}{E_{\text{Left}} + E_{\text{Right}}} \;, \tag{5.11}$$

$$y_\pi^{\text{wAD}} = k_{\text{TB}} \frac{E_{\text{Top}} - E_{\text{Bottom}}}{E_{\text{Top}} + E_{\text{Bottom}}} \;, \text{ and} \tag{5.12}$$

$$z_\pi = \left(\frac{p_{\pi^+}}{64.56 \text{ MeV/c}} \right)^{3.5} \;, \tag{5.13}$$

with k_{LR} and k_{TB} constant, and p_{π^+} equaling the momentum of the pion as it enters the target. The positron tracks from the wire chambers give the point at which the positron exits the target, $\mathbf{v}_e$. With this information, the path length of the positron in the target becomes,

$$L_e^{\text{TGT}} = \sqrt{(\mathbf{v}_\pi - \mathbf{v}_e)^2 - D^2} \;, \tag{5.14}$$

where D is the calculated distance of closest approach between these two trajectories.

You may have noticed that this leaves only one free fit parameter[3], namely t_{μ^+}. The fit reaches convergence more often and more quickly when we are able to set the initial value of the free parameter as close as possible to the true value.

We predict the muon time in the target waveform by first subtracting the two pulses with the

---

[3]A considerable effort was put forth towards fitting with other free fit parameters. This work is described in more detail in Appendix C.
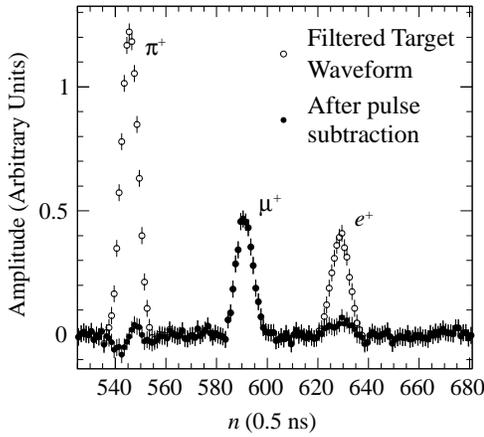
Figure 5.8: *Filtered waveform before and after subtracting predicted pion and positron pulse shapes. The subtracted waveform is scanned for a possible muon pulse to obtain a prediction for the muon time.*
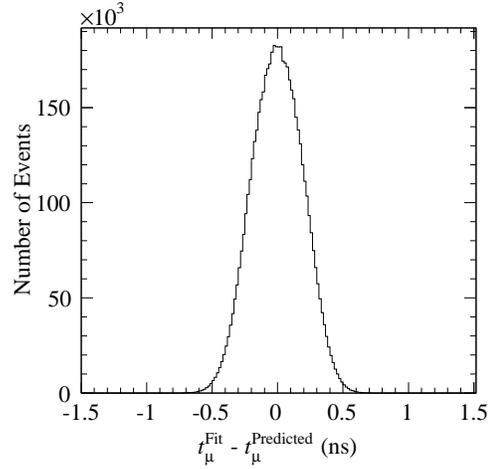
Figure 5.9: *Difference between predicted muon pulse time and the result of the fit demonstrating our predicting power.*

predicted pion and positron parameters from the filtered waveform. The resulting waveform, w'' as in Figure 5.8, is then scanned for the center bin, n_c, such that the integral of the bins surrounding n_c is closest to 4.12 MeV (the energy of the stopped muon from the $\pi \to \mu$ decay). The predicted muon time is the amplitude-weighted mean of the three bins surrounding n_c,

$$t_\mu^{\text{Pred}} = \frac{\sum_{n=n_c-1}^{n_c+1} (n)\, w''[n]}{\sum_{n=n_c-1}^{n_c+1} w''[n]} \ . \tag{5.15}$$

Minimization

The target waveform is then fit using both a two-peak ($\pi^+ \to e^+$) fit function and a three-peak ($\pi^+ \to \mu^+ \to e^+$) fit function. The minimization is executed using MINUIT2, the C++ object-oriented implementation of the ubiquitous MINUIT minimization package [41]. The "combined" algorithm is used, such that if the MIGRAD [37] method fails its first attempt, a Simplex minimization is

performed, and then MIGRAD is tried once more. Ocassionally the "combined" algorithm fails, at

which point the fit is attempted once more with the "scan" method. This method scans over the

free parameter values within the parameter limits and always returns the best parameter values. The

"scan" method is much slower, but there is no need to worry about lack of convergence. An example

showing both fits is shown in Figures 5.6(c) and 5.7(c).
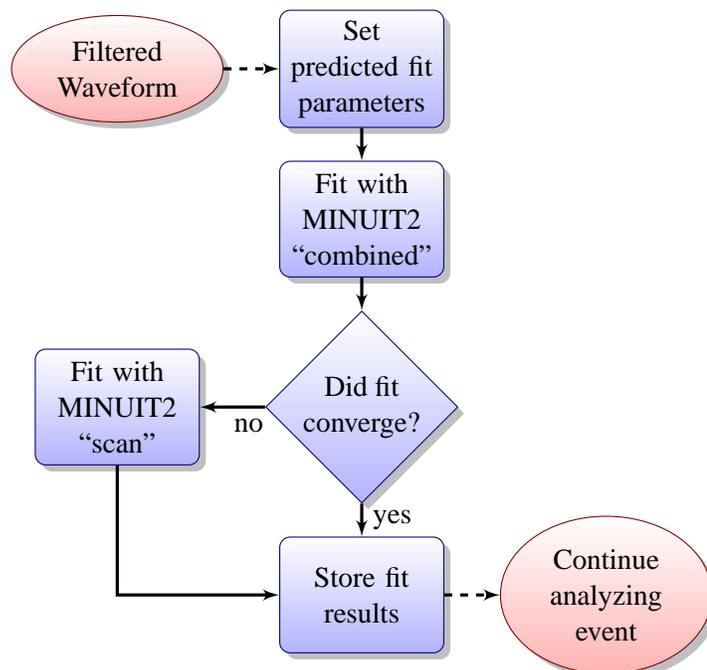
Figure 5.10: *Waveform fit flowchart demonstrating the check for convergence and the second fit attempt with the "scan" method when the "combined" method fails.*

Fitting the filtered waveforms with a series of Gaussian pulses instead of fitting the raw unfil-

tered waveforms with the known (averaged) responses has its benefits and drawbacks. Filtering,

while isolating the pulses, removes some information contained in the tails of the pulses. The filter-

ing results in only a few percent poorer temporal resolution when compared to unfiltered waveform

fits.
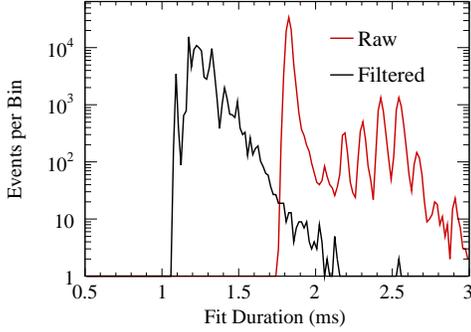
Figure 5.11: *The amount of time it takes the fits to converge for fits to raw waveforms and filtered waveforms.*
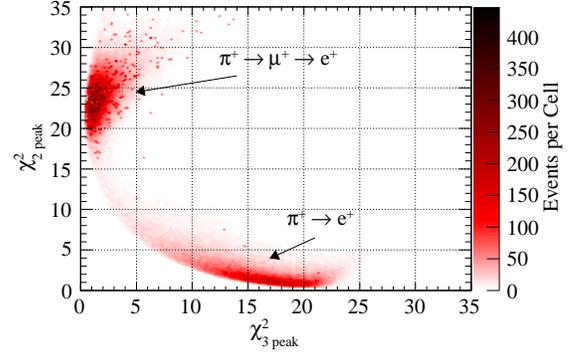
Figure 5.12: *Waveform fitting results demonstrating two distinct regions isolating the 2-peak events from the 3-peak events.*

On the other hand, there is a significant increase in the amount of time it takes each raw waveform fit to converge. For the PEN target waveforms, we see a ∼30% increase in speed when fitting filtered waveforms, Figure 5.11.

Using the Fit Results

The pion decay mode can then be determined based on the χ^2 difference between the two fit hypotheses,

$$\Delta\chi^2 = \chi^2_{3\text{-peak}} - \chi^2_{2\text{-peak}}. \tag{5.16}$$

Considering only events for which $\Delta\chi^2 > 0$ will isolate a pure sample of π_{e2} events. The sample will help determine the probability distribution function for the low-energy tail of the $\pi^+ \rightarrow e^+ \nu_e$ positron energy spectrum to be used in a maximum likelihood analysis to determine the π_{e2} decay branching ratio. Figure 5.12 clearly demonstrates the power of the χ^2 fits in isolating the decay channels. Figure 5.13 shows the characteristic Michel positron energy spectrum for negative values of $\Delta\chi^2$ and a well separated π_{e2} positron energy peak and low-energy tail.
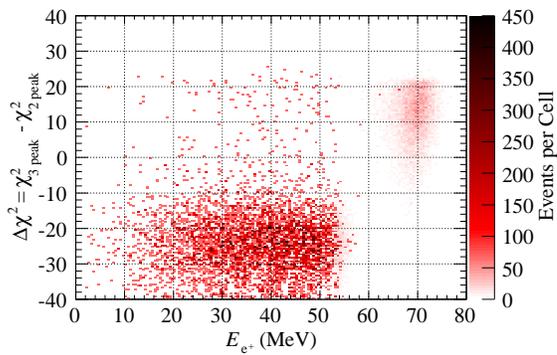
Figure 5.13: $\Delta\chi^2$ *results showing the michel positron energy spectrum for* $\Delta\chi^2 < 0$ *and the 69.3 MeV positron energy peak and low-energy tail from the* $\pi^+ \to e^+\nu_e$ *events for* $\Delta\chi^2 > 0$.
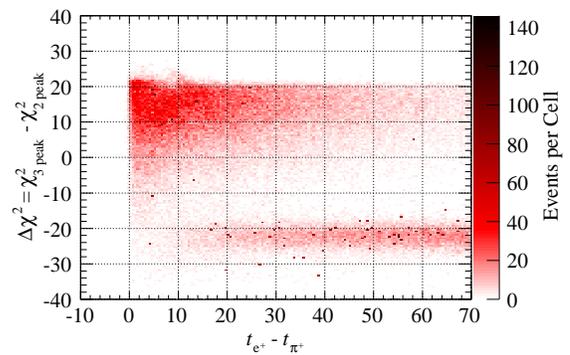
Figure 5.14: $\Delta\chi^2$ *results showing the exponential time spectrum with the pion lifetime for* $\Delta\chi^2 > 0$ *and the rising then falling spectrum characteristic of the* $\pi \to \mu \to e$ *sequence for events with* $\Delta\chi^2 < 0$.

Chapter 6

Maximum Likelihood Estimation of $R_{\pi_{e2}}$

We have chosen to use the maximum likelihood (ML) method to extract $R_{\pi_{e2}}$ from our measurement

data. In this chapter we will review the principle of maximum likelihood and develop the method in

the context of the PEN experiment.

6.1   Introduction to the Likelihood Technique

The probability of obtaining any exact value of a continuous random variable is zero since there

exists an infinite number of possible values. Therefore it is only meaningful to say the probability

that a result of an experiment lies within a certain interval. Hence we obtain a probability distribu-

tion function (PDF), $P(x)$, such that the probability to obtain a measurement between x and $x + \mathrm{d}x$

is $P(x)\,\mathrm{d}x$, while $\int P(x)\,\mathrm{d}x = 1$.

Now consider a PDF, $P(x;\theta)$, in which the shape of the distribution is not known exactly, but

can be described with the parameter θ. Both x and θ can be sets of data, $\vec{x}$, and parameters, $\vec{\theta}$,

not necessarily of the same dimension. When we make a series of N independent observations of

events, x_e, we construct the joint PDF as,

$$\mathcal{L}\left(\vec{x};\vec{\theta}\right) = \mathcal{L}\left(x_1,x_2,...,x_N;\vec{\theta}\right) = \prod_{e=1}^{N} P\left(\vec{x}_e;\vec{\theta}\right) \; , \tag{6.1}$$

which depends on both the measurements, $\vec{x}_e$, and the parameters, $\vec{\theta}$. After the measurement (or

simulation) data are collected, however, the $\vec{x}_e$ are fixed, and $\mathcal{L}$, known as the *likelihood function*,

becomes a function of $\vec{\theta}$ only. For practical purposes in analysis, it may be easier to deal with sums

than products, so we define the log-likelihood as,

$$\ell \equiv \ln \mathcal{L}\left(\vec{x};\vec{\theta}\right) = \ln \mathcal{L}\left(x_1,x_2,...,x_N;\vec{\theta}\right) = \sum_{e=1}^{N} \ln P\left(\vec{x}_e;\vec{\theta}\right) \; . \tag{6.2}$$

The *Principle of Maximum Likelihood* states that the best estimate for a parameter θ is that value

which maximizes the likelihood (or log-likelihood) function. The ML method provides a unique,

unbiased, minimum variance estimate for the parameters of interest for a large enough sample [30].

Maximum Likelihood in the PEN Experiment

A log-likelihood function for the PEN experiment analysis can be given by

$$\ell = \ln \mathcal{L} = \ln \left(\prod_{e=1}^{N} \sum_{i=1}^{M} f_i P_i(\vec{x}_e)\right) = \sum_{e=1}^{N} \ln \left(\sum_{i=1}^{M} f_i P_i(\vec{x}_e)\right) \tag{6.3}$$

where N is the total number of observed events entering the likelihood analysis, M is the number

of processes we're considering, f_i is the fraction of events in process i, $P_i(\vec{x}_e)$ is the probability to

find an event in process i given the observed values of $\vec{x}_e$ (e.g., time and energy). In this type of ML

analysis, we assume the shapes of the P_i are known exactly and replace the parameters θ_i with the

fractions of each process, f_i. Maximizing this likelihood function gives the most likely values of f_i, which we will denote $\hat{f}_i$.

The PEN measurement data contains five main processes that we will consider. The other processes, if any, occur at such a low level that they are not significant in this analysis. The two signal processes are the $\pi \to e$ and $\pi \to \mu \to e$ decays with fractions denoted as $f_{\pi_{e2}}$ and $f_{\pi_{\mu2}}$, respectively. The three background processes include: (1) accidental coincidences with fraction f_{Acc}, (2) prompt pion-nucleon reactions which result in a proton final state, with fraction f_{Had}, and (3) events in which the pion decays in flight (DIF), with the resulting muon coming to rest in the target, with fraction f_{DIF}.

About one third of the recorded events are pion-nucleon reactions in the degrader and target. In the energy window from 60 to 80 MeV we've collected more proton events than $\pi \to e$ positrons. A background of this magnitude means that we have to separate these processes better than 1:10000 (or rather understand the expected overlap at this level).

A decay-in-flight event ends up depositing a muon instead of a pion in the target at the time of the πSTOP signal. DIF kinematics are such that muons emitted forward will punch all the way through our relatively thin target, and muons emitted backward will not reach the target, or will stop in the first several μm of the target depositing a reduced amount of energy, thus making a distinct signal from a pion. Sadly, a sufficient number of muons (pions decaying sideways) will stop in the target with energy depositions closely matching those of a pion stop. These events will mimic valid 2-peak events in the target waveform, but will populate the $\pi \to e$ low-energy "tail" with bona fide Michel positrons.

The variance of $\hat{f}_i$ is given by $\mathbf{V}_{ii} = \sigma_{ii}^2$, where $\mathbf{V}$ is the covariance matrix. The elements of the covariance matrix can be calculated as the inverse of the expectation value of the second partial

derivative of the negative of the log-likelihood function,

$$\mathbf{V}_{jk}^{-1} = \mathrm{E}\left[\frac{\partial^2(-\ell)}{\partial f_j \partial f_k}\right] \quad . \tag{6.4}$$

The expectation value can be estimated in the sample mean by setting the parameters to $\hat{f}$ to give

$$\mathbf{V}_{jk}^{-1} = -\sum_{e=1}^{N} \frac{\partial}{\partial f_j} \left.\frac{P_k(\vec{x}_e)}{\sum\limits_{i=1}^{M} f_i P_i(\vec{x}_e)}\right|_{f=\hat{f}} = \sum_{e=1}^{N} \left.\frac{P_j(\vec{x}_e)P_k(\vec{x}_e)}{\left[\sum\limits_{i=1}^{M} f_i P_i(\vec{x}_e)\right]^2}\right|_{f=\hat{f}} \quad . \tag{6.5}$$

When the data have been prescaled (e.g., by a factor 64), one can say that an event had an efficiency, ε_e, of 1/64. The corresponding weight, w_e, for that event is given by $w_e = 1/\varepsilon_e = 64$. The likelihood function then takes the form

$$\ell' = \sum_{e=1}^{N} w_e \ln\left(\sum_{i=1}^{M} f_i P_i(\vec{x}_e)\right) \tag{6.6}$$

It turns out that the estimate $\hat{f}$ obtained by maximizing ℓ' is still asyptotically normally distributed about the true value [30]. Using the second derivative matrix to calculate the variance as in (6.4) is wrong, since it assumes every event has a weight of one. The correct prescription [30] is to take

$$\mathbf{V}(\hat{f}) = \mathbf{H}^{-1}\mathbf{H}'\mathbf{H}^{-1} \tag{6.7}$$

When estimated in the sample mean, the matrix elements can be calculated as

$$\mathbf{H}_{jk} = \sum_{e=1}^{N} w_e \left.\frac{P_j(\vec{x}_e)P_k(\vec{x}_e)}{\left[\sum\limits_{i=1}^{M} f_i P_i(\vec{x}_e)\right]^2}\right|_{f=\hat{f}} \tag{6.8}$$

$$\mathbf{H}'_{jk} = \sum_{e=1}^{N} w_e^2 \frac{P_j(\vec{x}_e)P_k(\vec{x}_e)}{\left[\sum\limits_{i=1}^{M} f_i P_i(\vec{x}_e)\right]^2}\Bigg|_{f=\hat{f}} \tag{6.9}$$

Now if $i = 0$ corresponds to $\pi \to \mu \to e$ and $i = 1$ corresponds to $\pi \to e$ we can propagate the

uncertainty in the ratio of branching fractions $R_{\pi_{e2}} = f_1/f_0$ as follows:

$$\left(\frac{\sigma_R}{R}\right)^2 = \left(\frac{\sigma_{f_0}}{f_0}\right)^2 + \left(\frac{\sigma_{f_1}}{f_1}\right)^2 - 2\frac{\sigma_{f_0}\sigma_{f_1}}{f_0 f_1}\rho_{f_0 f_1} \ , \tag{6.10}$$

where the correlation coefficient

$$\rho_{f_0 f_1} = \frac{\mathrm{cov}_{f_0 f_1}}{\sigma_{f_0}\sigma_{f_1}} = \frac{\mathbf{V}_{01}}{\sqrt{\mathbf{V}_{00}}\sqrt{\mathbf{V}_{11}}} \ , \tag{6.11}$$

thus obtaining the propagated relative-uncertainty squared,

$$\left(\frac{\sigma_R}{R}\right)^2 = \frac{\mathbf{V}_{00}}{f_0^2} + \frac{\mathbf{V}_{11}}{f_1^2} - 2\frac{\mathbf{V}_{01}}{f_0 f_1} \ . \tag{6.12}$$

6.2 Approaches in Likelihood Analyses

Frequentist vs. Bayesian

Experimental results can be stated using the empirical (frequentist) interpretation or the subjective

(Bayesian) interpretation.

In frequentist interpretatation, the probability is viewed as the limit of the frequency of a result

of an experiment or observation when the number of identical experiments is very large, i.e., $P(x_i) =$

$\lim\limits_{N\to\infty} (N_i/N)$. One disadvantage of this approach is that $P(x_i)$ is not only a property of the experiment,

it also depends on the "ensemble, i.e., on all N repetitions of the experiment.

Probability in the Bayesian interpretation is a subjective "degree of belief which is susceptible to modifications upon further observations. The probability is written,

$$P_{\text{posterior}}(\theta\,|x) = \frac{f(x\,|\theta)}{f_1(x)} P_{\text{prior}}(\theta)$$

A guess at the prior is subjective and therefore unscientific. Bayes Postulate says if you're completely ignorant about $P(\theta)$, then take all values of θ to be equally probable. There are several objections to this postulate, but experience has shown that $P_{\text{posterior}}$ usually converges to an identical value after several experiments, regardless of the choice of prior.

Ordinary vs. Extended

The likelihood analyses discussed so far determine the parameters describing shapes of PDFs or the relative fractions of the PDFs for processes contained in the sample. The total number of events is regarded as fixed.

Enrico Fermi decided to extended the ML method to include the total number of events as a parameter to be estimated. The extension is obtained by multiplying the ordinary likelihood function by a Poisson PDF expressing the probability of obtaining N events when the expected number is ν,

$$\mathcal{L}\left(\vec{x};\vec{\theta}\right) = \prod_{e=1}^{N} P\left(\vec{x}_e;\vec{\theta}\right) \rightarrow \mathcal{L}_{\text{E}}\left(\vec{x};\vec{\theta},\nu\right) = \frac{e^{-\nu}\nu^{N}}{N!} \prod_{e=1}^{N} P\left(\vec{x}_e;\vec{\theta}\right) \ . \tag{6.13}$$

Whether we use the ordinary maximum likelihood or extended maximum likelihood approach depends on the question we are trying to answer. If we are measuring properties of tracks from decay products of some particle and we want to determine its branching *ratio*, then the ordinary ML is prefered. If, however, we are interested in its partial decay *rates*, then the extended maximum

likelihood is required [47]. For the PEN analysis the choice is clear, we will use the ordinary likelihood.

6.3 Parameterization of the Likelihood Function

Typically the value quoted for the $\pi \to e\nu$ branching ratio is $f_{\pi_{e2}}/f_{\pi_{\mu2}}$. There is no reason why we couldn't present our result in a slightly different format. If we instead desired a result for the ratio $f_{\pi_{e2}}/(f_{\pi_{e2}} + f_{\pi_{\mu2}})$ we could parameterize the likelihood function such that we obtain that ratio directly. This parameterization would eliminate the necessity to propagate errors and would automatically include any correlations between $f_{\pi_{e2}}$ and $f_{\pi_{\mu2}}$. The parameterization of the likelihood is then given by,

$$
\begin{aligned}
\mathcal{L} = \prod_{e=1}^{N} \big[&(1-\theta_0)(1-\theta_1)(1-\theta_2)(1-\theta_3) P_{\pi_{\mu2}}(\vec{x}_e) \\
&+\theta_0(1-\theta_1)(1-\theta_2)(1-\theta_3) P_{\pi_{e2}}(\vec{x}_e) \\
&+\theta_1(1-\theta_2)(1-\theta_3) P_{\mathrm{Had}}(\vec{x}_e) \\
&+\theta_2(1-\theta_3) P_{\mathrm{DIF}}(\vec{x}_e) \\
&+\theta_3 P_{\mathrm{Acc}}(\vec{x}_e) \big] \,.
\end{aligned}
\tag{6.14}
$$

The parameters, θ_i, are defined in terms of the original fractions, f_i, as,

$$\theta_0 = \frac{f_{\pi_{e2}}}{f_{\pi_{\mu2}} + f_{\pi_{e2}}}$$

$$\theta_1 = \frac{f_{\text{Had}}}{f_{\pi_{\mu2}} + f_{\pi_{e2}} + f_{\text{Had}}}$$

$$\theta_2 = \frac{f_{\text{DIF}}}{f_{\pi_{\mu2}} + f_{\pi_{e2}} + f_{\text{Had}} + f_{\text{DIF}}}$$

$$\theta_3 = \frac{f_{\text{Acc}}}{f_{\pi_{\mu2}} + f_{\pi_{e2}} + f_{\text{Had}} + f_{\text{DIF}} + f_{\text{Acc}}} \quad . \tag{6.15}$$

Thus the best estimate of θ_0 is the goal of the PEN experiment. This parameterization forfeits the symmetry found in (6.3). The covariance matrix elements are again calculated using the expectation value of the negative of the second derivatives of the log-likelihood function, $\ell = \ln \mathcal{L}$, as in (6.4):

$$\mathbf{V}_{jk}^{-1} = \sum_{e=1}^{N} \left[\frac{1}{\mathcal{L}^2} \frac{\partial \mathcal{L}}{\partial \theta_j} \frac{\partial \mathcal{L}}{\partial \theta_k} - \frac{1}{\mathcal{L}} \frac{\partial^2 \mathcal{L}}{\partial \theta_j \partial \theta_k} \right]_{\theta = \hat{\theta}} \quad . \tag{6.16}$$

All of the required partial derivatives of $\mathcal{L}$ are given for reference in Appendix D. The corrections to the covariance matrix due to the weighted events follow as in (6.7)–(6.9) with

$$\mathbf{H}_{jk} = \sum_{e=1}^{N} \left[w_e \left(\frac{1}{\mathcal{L}^2} \frac{\partial \mathcal{L}}{\partial \theta_j} \frac{\partial \mathcal{L}}{\partial \theta_k} - \frac{1}{\mathcal{L}} \frac{\partial^2 \mathcal{L}}{\partial \theta_j \partial \theta_k} \right) \right]_{\theta = \hat{\theta}} \quad , \tag{6.17}$$

$$\mathbf{H'}_{jk} = \sum_{e=1}^{N} \left[w_e^2 \left(\frac{1}{\mathcal{L}^2} \frac{\partial \mathcal{L}}{\partial \theta_j} \frac{\partial \mathcal{L}}{\partial \theta_k} - \frac{1}{\mathcal{L}} \frac{\partial^2 \mathcal{L}}{\partial \theta_j \partial \theta_k} \right) \right]_{\theta = \hat{\theta}} \quad . \tag{6.18}$$

6.4   Including External Constraints

Certain quantities in a likelihood analysis may be determined with some precision from an external method. By external we mean using data collected possibly with other detector subsystems; the

minimum requirement being that the same data do not enter *both* the external constraint measurement and the likelihood analysis. For example, one may fit side-bands of a distribution (outside the domain of the PDF) to estimate a background. The incorporation of this external information into the PEN likelihood function takes the form,

$$\mathcal{L}_C = C(\theta_3)\,\mathcal{L} = \mathrm{e}^{-\frac{1}{2}\frac{(\theta_3 - \langle\theta_3\rangle)^2}{\sigma_{\theta_3}^2}}\,\mathcal{L} \tag{6.19}$$

Here we have measured θ_3 to be $\langle\theta_3\rangle$ with an uncertainty of σ_{θ_3}. This is the new likelihood function we wish to maximize. As mentioned earlier, numerical recipes prefer to minimize functions. Traditionally one minimizes,

$$
\begin{aligned}
-\ln(\mathcal{L}_C) \quad = \frac{1}{2}\frac{(\theta_3 - \langle\theta_3\rangle)^2}{\sigma_{\theta_3}^2} \\
- \sum_{e=1}^{N} \ln\big[(1-\theta_0)(1-\theta_1)(1-\theta_2)(1-\theta_3) \quad P_{\pi_{\mu 2}}(\vec{x}_e) \\
+\theta_0(1-\theta_1)(1-\theta_2)(1-\theta_3) \quad P_{\pi_{e2}}(\vec{x}_e) \\
+\theta_1(1-\theta_2)(1-\theta_3) \quad P_{\mathrm{Had}}(\vec{x}_e) \\
+\theta_2(1-\theta_3) \quad P_{\mathrm{DIF}}(\vec{x}_e) \\
+\theta_3 \quad P_{\mathrm{Acc}}(\vec{x}_e)\big]\,.
\end{aligned}
\tag{6.20}
$$

Following (6.4) the variance-covariance matrix becomes,

$$\mathbf{V}_{jk}^{-1} = \left[\frac{1}{C^2}\frac{\partial C}{\partial\theta_j}\frac{\partial C}{\partial\theta_k} - \frac{1}{C}\frac{\partial^2 C}{\partial\theta_j\partial\theta_k}\right]_{\theta=\hat{\theta}} + \sum_{e=1}^{N}\left[\frac{1}{\mathcal{L}^2}\frac{\partial\mathcal{L}}{\partial\theta_j}\frac{\partial\mathcal{L}}{\partial\theta_k} - \frac{1}{\mathcal{L}}\frac{\partial^2\mathcal{L}}{\partial\theta_j\partial\theta_k}\right]_{\theta=\hat{\theta}}\,. \tag{6.21}$$

Since the partial derivatives of the constraint term are,

$$
\frac{\partial \mathcal{C}}{\partial \theta_j} = \begin{cases} -\frac{\theta_3 - \langle \theta_3 \rangle}{\sigma_{\theta_3}^2} \, \mathrm{e}^{-\frac{1}{2} \frac{(\theta_3 - \langle \theta_3 \rangle)^2}{\sigma_{\theta_3}^2}} & \text{for } j = 3, \\[2ex] 0 & \text{otherwise, and} \end{cases} \tag{6.22}
$$

$$
\frac{\partial^2 \mathcal{C}}{\partial \theta_j \partial \theta_k} = \begin{cases} \left(\frac{(\theta_3 - \langle \theta_3 \rangle)^2}{\sigma_{\theta_3}^4} - \frac{1}{\sigma_{\theta_3}^2} \right) \mathrm{e}^{-\frac{1}{2} \frac{(\theta_3 - \langle \theta_3 \rangle)^2}{\sigma_{\theta_3}^2}} & \text{for } j = k = 3, \\[2ex] 0 & \text{otherwise,} \end{cases} \tag{6.23}
$$

we see that the additional term due to the constraint only modifies the $j = k = 3$ element in the variance-covariance matrix.

6.5 Probability Distribution Functions

Of critical importance in any maximum likelihood analysis are accurate probability distribution functions. The determination of the PDFs for many processes becomes tractable in that we are able to obtain one-dimensional PDFs for each observable. When the observables are independent, the total PDF for the process is just the product of all one-dimensional PDFs. We are able to make hard multi-dimensional cuts on the data to obtain a clean sample in a selected observable to accurately plot the distribution of that observable for a particular process. The preferred method of determining PDFs is to obtain them directly from measurement data. In regions where sufficient isolation of each process in measurement data is not possible, we supplement the PDF shape determination with simulated data. The following sections will discuss each observable and the corresponding PDFs in detail.

Table 6.1: *Cuts used to find* $\pi \to \mu \to$ e *energy PDF. Only the first cut is used for this analysis. The second and third cuts were found to distort the shape of the energy PDFs.*

| Cut	Description
$(t_e - t_\pi) > 4\tau_\pi$	Remove early $\pi \to$ e events
$N_{\text{Hit}}^{\text{MWPC2}} < 2$	Remove backgrounds (no longer used)
$N_{\mu \text{ candidates}} < 3$	Conservative cut removing further BG (no longer used)

6.5.1 Total Energy

The first observable we will consider is the total energy deposited in all the active detector elements by the outgoing final-state particle. In the case of pion-nucleon reactions the final-state particle is a proton, in all of the other cases we're considering it is a positron. The total energy consists of the energy deposited in the active target, the plastic hodoscope, and the CsI calorimeter,

$$E_{\text{Total}} = E_{\text{Target}} + E_{\text{PH}} + E_{\text{CsI}} \ . \tag{6.24}$$

We are not including the energy deposited as the particle crosses the MWPCs, as it is a negligible contribution.
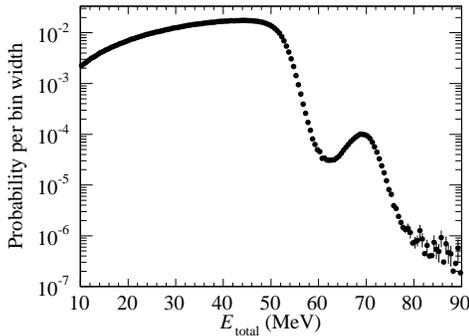
(a) E_{Total} *for all* e$^+$ *final-state processes.*
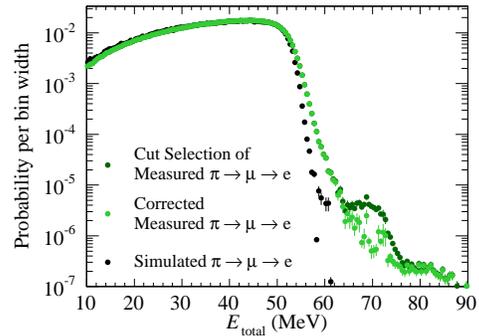
(b) *Corrected Michel PDF and simulation.*

Figure 6.1: (a) *The total energy for the four processes with a positron in the final state. Events with proton final state were removed with cuts.* (b) *PDF for Michel positron energy will be obtained first from cuts. Then the known* $\pi \to$ e *contamination in the late time window is subtracted. We see the high energy tail does not match simulation, as we cannot remove 100% of the pile-up background.*

We imposed cuts on the data to extract the shape of E_{Total} for the $\pi \to \mu \to e$ process alone, Table 6.1. The main software cut selects only events after four pion lifetimes, leaving behind many $\pi \to \mu \to e$ sequential events while removing the exponentially decaying $\pi \to e$ events. A second cut requiring at most one hit in the outer wire chamber ($N_{\text{Hit}}^{\text{MWPC2}} < 2$) would eliminate many background events, but is highly energy dependent and distorts the shape of the energy distribution. Therefore we do not use the $N_{\text{Hit}}^{\text{MWPC2}}$ cut. A third potential cut uses $N_{\mu\,\text{candidates}}$, which is a count of potential muon peak candidates obtained by scanning the target waveform for 4.12 MeV peaks. A conservative cut requiring $N_{\mu\,\text{candidates}} < 3$ would eliminate any events with very large extra pulses or more than two muon pulses, allowing us to suppress pile-up and/or accidental events.[1] After cuts, we were left with a high-energy tail consisting of events with two positrons in the same "track", pile-up events, or events with extra photons contributing to the total energy. The relative yield of pile-up events in the high-energy tail comes out correct by construction.

A significant systematic uncertainty arises due to our lack of knowledge of the shape of the E_{Total} low-energy tail for the $\pi \to e$ process. A Crystal Ball function [68, 59, 38],

$$f(E_{\text{Total}}; \alpha, n, \overline{E}_{\text{Total}}, \sigma) = N \begin{cases} \exp\left(-\frac{(E_{\text{Total}} - \overline{E}_{\text{Total}})^2}{2\sigma^2}\right) & \text{for} \frac{E_{\text{Total}} - \overline{E}_{\text{Total}}}{\sigma} > -\alpha \\ A\left(B - \frac{E_{\text{Total}} - \overline{E}_{\text{Total}}}{\sigma}\right)^{-n} & \text{for} \frac{E_{\text{Total}} - \overline{E}_{\text{Total}}}{\sigma} \leq -\alpha \end{cases} \tag{6.25}$$

was fit to the simulated E_{Total}. This function consists of a main Gaussian peak and a power-law tail to the left. N is the overall normalization, $\overline{E}_{\text{Total}}$ and σ are the mean and σ of the Gaussian peak, α

---

[1] At the time of writing this dissertation the $N_{\mu\,\text{candidates}} < 3$ cut is not imposed and the $\pi \to \mu \to e$, accidental, and pile-up processes all share the same PDF shape for E_{Total}. The differences between these PDFs will be determined in a future analysis thus removing a systematic shift in the branching ratio measurement.

(a) $\pi \to e$ *from simulation.*

(b) $\pi \to e$ *from cuts, fit, and corrected.*

(c) *After BG Subtraction.*

(d) *Uncertainty in low-energy tail.*
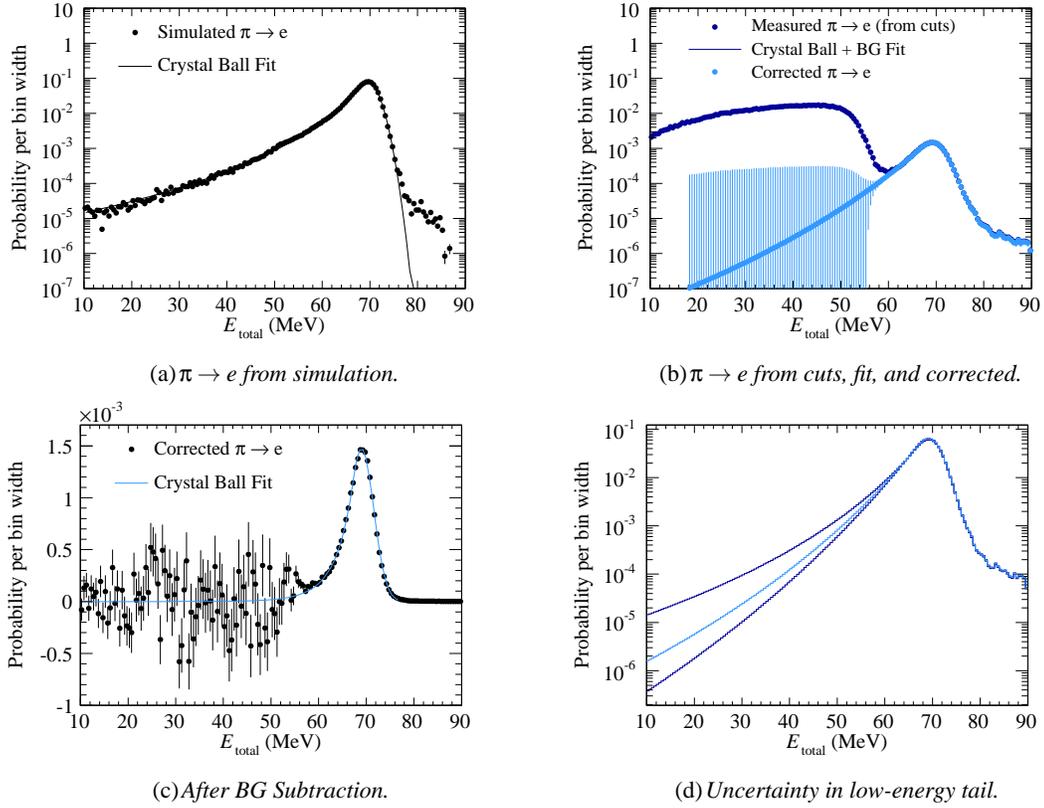
Figure 6.2: (a) *"Crystal Ball" (CB) function fitted to simulated data.* (b) *Combined function: CB + Michel shape is fit to the best "$\pi \to e$" distribution we can obtain from cuts.* (c) *Subtracting BG component and again fitting with CB function.* (d) *A demonstration of PDFs with extrema of CB fit parameter uncertainties.*

and n describe the shape of the tail and,

$$A = \left(\frac{n}{|\alpha|} \right)^{n} \exp \left(-\frac{|\alpha|^2}{2} \right) \text{ , and} \tag{6.26}$$

$$B = \frac{n}{|\alpha|} - |\alpha| \text{ .} \tag{6.27}$$

The best E_{Total} distribution for the $\pi \to e$ process we can obtain using cuts to measurement data

still contains a significant number of events due to the accidental process and $\pi \to \mu \to e$ events

as well as a high-energy tail similar to the one described above in the Michel energy distribution.

We are able to determine a PDF by fitting a function with the Crystal Ball shape combined with an

Table 6.2: *Cuts used to find $\pi \to$ e energy PDF.*

Cut	Description
$-2.0\,\text{ns} < (t_\text{e} - t_\pi) < 0.75\,\tau_\pi$	Remove late $\pi \to \mu$ events
$\left(\Delta\chi^2 > -5\right) \cup \left(N_{\mu\,\text{candidates}} < 2\right)$	Conservative cut removing further BG

analytical form of the Michel distribution[2] to the cut-based distribution as in Figure 6.2(b). We can

then obtain the energy PDF with two methods.

Using the first method, we subtract the background component as shown in Figure 6.2(c). We

then fit this subtracted distribution with the Crystal Ball function obtaining,

$$N = 0.001478 \pm 0.000005 \ ,$$

$$\overline{E}_\text{Total} = 69.023 \pm 0.014 \ ,$$

$$\sigma = 2.547 \pm 0.010 \ , \tag{6.28}$$

$$\alpha = 0.761 \pm 0.019 \ , \ \text{and}$$

$$n = 8.508 \pm 2.653 \ .$$

Note the large uncertainty on n. We set the low-energy tail of the PDF to these function values and

use the exact shape of the measurement for the peak and high-energy regions.

The second method relies heavily on simulation. The parameters governing the shape of the

---

[2]The analytical representation of the Michel energy distribution is given in (4.5) and was described in Section 4.3 where it was used to calibrate the CsI energies.

low-energy tail are fixed to the values determined in the fit to simulation, Figure 6.2(a),

$$N = 0.08332 \pm 0.00005 \ ,$$

$$\overline{E}_{\text{Total}} = 69.511 \pm 0.002 \ ,$$

$$\sigma = 1.780 \pm 0.001 \ , \tag{6.29}$$

$$\alpha = 0.670 \pm 0.001 \ , \ \text{and}$$

$$n = 5.007 \pm 0.012 \ .$$

The PDF is again taken to be the exact shape of the measurement in the peak and high-energy regions, and the PDF is set to the values of the Crystal Ball fit function in the tail region.

Using the uncertainties on each fit parameter, p, we can then set each parameter to $p \pm \sigma_p$ in the appropriate direction that either maximizes or minimizes the probability in the tail region. The results are two PDFs corresponding to $\pm 1\sigma$ fluctuations in the tail that we can use to estimate the systematic uncertainty due to our lack of knowledge about the low-energy tail shape. Figure 6.2(d) shows the PDF shapes from the first method demonstrating the huge uncertainty in the tail region, amounting to a ~ 2 % uncertainty on $R_{\pi_{e2}}$. Relying on the current set of simulated data, as in the second method, the uncertainty on the tail is drastically reduced such that our systematic relative uncertainty on $R_{\pi_{e2}}$, due to the tail, becomes ~ 0.33 %[3]. Generating a much larger set of simulated data can push this uncertainty lower. The shape of the low-energy tail from measurement is still under study, and will ultimately set the systematic uncertainty limit.

The data set that we chose to fit has prompt pion-nucleon reactions removed. The cuts used to

---

[3]Performing the minimum χ^2 fit as in Section 6.6 (Stage 1), using the "extrema" PDFs (similar to Figure 6.2(d), but from the second method), we are able to quantify our systematic uncertainty due to the low-energy tail in the E_{Total} PDF for the $\pi \to e$ process. Including this uncertainty in (6.46) we obtain $\tilde{\theta}_0^{(\chi^2)} = [1.1164 \pm 0.0025(\text{stat.}) \pm 0.0037(\text{syst.})] \times 10^{-4}$, where the systematic relative uncertainty is 0.33%.
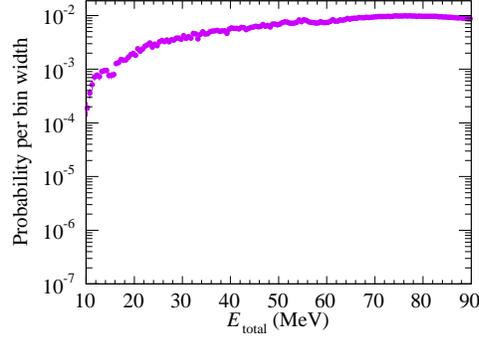
Figure 6.3: E_{Total} *for pion-nucleon reactions.*

remove these proton final-state events are given in Table 6.3, and are explained in more detail in Section 6.5.3. This cut is not 100% efficient so we created an E_{Total} PDF for hadronic reactions. The cuts used to select the proton final-state events are listed in Table 6.4, and the PDF is shown in Figure 6.3.

Minimum χ^2 Fit to Total Energy

To ensure our PDFs are reasonable, a fit to the measurement data was performed. The fit function consists of a linear combination of the PDFs for the $\pi \to \mu \to e$, $\pi \to e$, and prompt hadronic processes. For this preliminary analysis it is assumed that the $\pi \to \mu \to e$, pion decays-in-flight, and and accidental/pile-up processes have identical shapes and would be indistinguishable in a fit to E_{Total} alone. I therefore consider only their sum as a fit parameter, $N_{\pi_{\mu2}+D+A}$. The remaining parameters are the normalization of the $\pi \to e$ process, $N_{\pi_{e2}}$, and the prompt hadronic pion-nucleon reaction with the proton final state particle, N_{Had}.

Table 6.3: *Cuts used to remove events with protons in the final-state.*

Cut	Description
$\|\Delta\phi_{\text{MWPC2,PH}}\| < 1.8°$	Remove tracks through PH boundaries
$(\text{d}E/\text{d}x)_{\text{PH}} < 3$ MeV/cm	Protons deposit more energy in the PH

The fit returned $\chi^2/N_{\text{dof}} = 198.95/157 = 1.27$ with correlation matrix,

$$\rho = \begin{array}{c} \\ N_{\pi_{\mu2}+\text{D}+\text{A}} \\ N_{\pi_{e2}} \\ N_{\text{Had}} \end{array} \begin{array}{ccc} N_{\pi_{\mu2}+\text{D}+\text{A}} & N_{\pi_{e2}} & N_{\text{Had}} \\ \left(\begin{array}{ccc} 1 & 0.0119 & 7.46 \times 10^{-6} \\ 0.0119 & 1 & -6.16 \times 10^{-5} \\ 7.46 \times 10^{-6} & -6.16 \times 10^{-5} & 1 \end{array} \right) \end{array},$$

and normalizations,

$$N_{\pi_{\mu2}+\text{D}+\text{A}} = 4.93226 \times 10^8 \pm 170826.0 \ ,$$

$$N_{\pi_{e2}} = 806149.0 \pm 1754.49 \ , \text{ and} \tag{6.30}$$

$$N_{\text{Had}} = 5.61122 \times 10^{-12} \pm 48.2528 \ .$$

The fit function is shown in the top panel of Figure 6.4. Since there is no discriminatory power between the fraction of $\pi \to \mu \to$ e events, the pion decays-in-flight, and the accidental backgrounds, we are unable to quote a branching ratio from this fit alone.

Taking the difference between the measurement histogram and the fit function illustrates the regions in which our PDFs may deviate from the true shape of the distribution. There is no apparent deviation in the middle panel of Figure 6.4. The bottom panel of Figure 6.4 shows the bin-wise contribution to the χ^2, also demonstrating that a linear combination of our energy PDFs match the shape of the measured data, with the exception of a few bins in the low-statistics high-energy region.

Table 6.4: *Cuts used to select events with protons in the final state.*

Cut	Description
$-3.0 < (t_{\text{fsp}} - t_{\pi}) < 3.0$ ns	Keep only prompt events
$(\mathrm{d}E/\mathrm{d}x)_{\text{PH}} > 13.0 \exp(-0.011\, E_{\text{Total}})$	Protons deposit more energy
$\min(Q_{\text{MWPC1}}, Q_{\text{MWPC2}}) > 1000$ ADC ch.	Require a min. MWPC energy

Table 6.5: *General cuts used to suppress backgrounds in the data sample.*

Cut	Description
$-10.0 < (t_{\text{fsp}} - t_{\pi}) < 200.0$ ns	Must exist within domain of all PDFs
$10.0 < E_{\text{Total}} < 90.0$ MeV	Must exist within domain of all PDFs
$E_{\text{CsI}}^{\text{Veto}} < 2.0$ MeV	Minimal energy leakage into veto crystals
$N_{\text{beamtracks}} = 1$	Require only one beam track
Beam Particle ID = π	Require π as beam particle
$23.7 < \text{TOF}_{\text{Beam Part.}} < 24.4$ ns	BC to DEG TOF is TOF of π
$8.5 < E_{\pi}^{\text{TGT,Pred}} < 13.5$ MeV	Reasonable predicted π energy
$L^{\text{TGT}} > 0.0$ cm	Valid decay track target path length
$40° < \theta < 140°$	Decay track points to calorimeter
ID = γ for extra decay tracks	Require additional tracks to be photons
Not Cosmic Ray	Decay track isn't a cosmic ray
Not scattered beam	Beam track isn't a scattered beam particle
"P"$_{\text{pile-up}} < -1.0$	Low probability of pile-up, see (6.39)
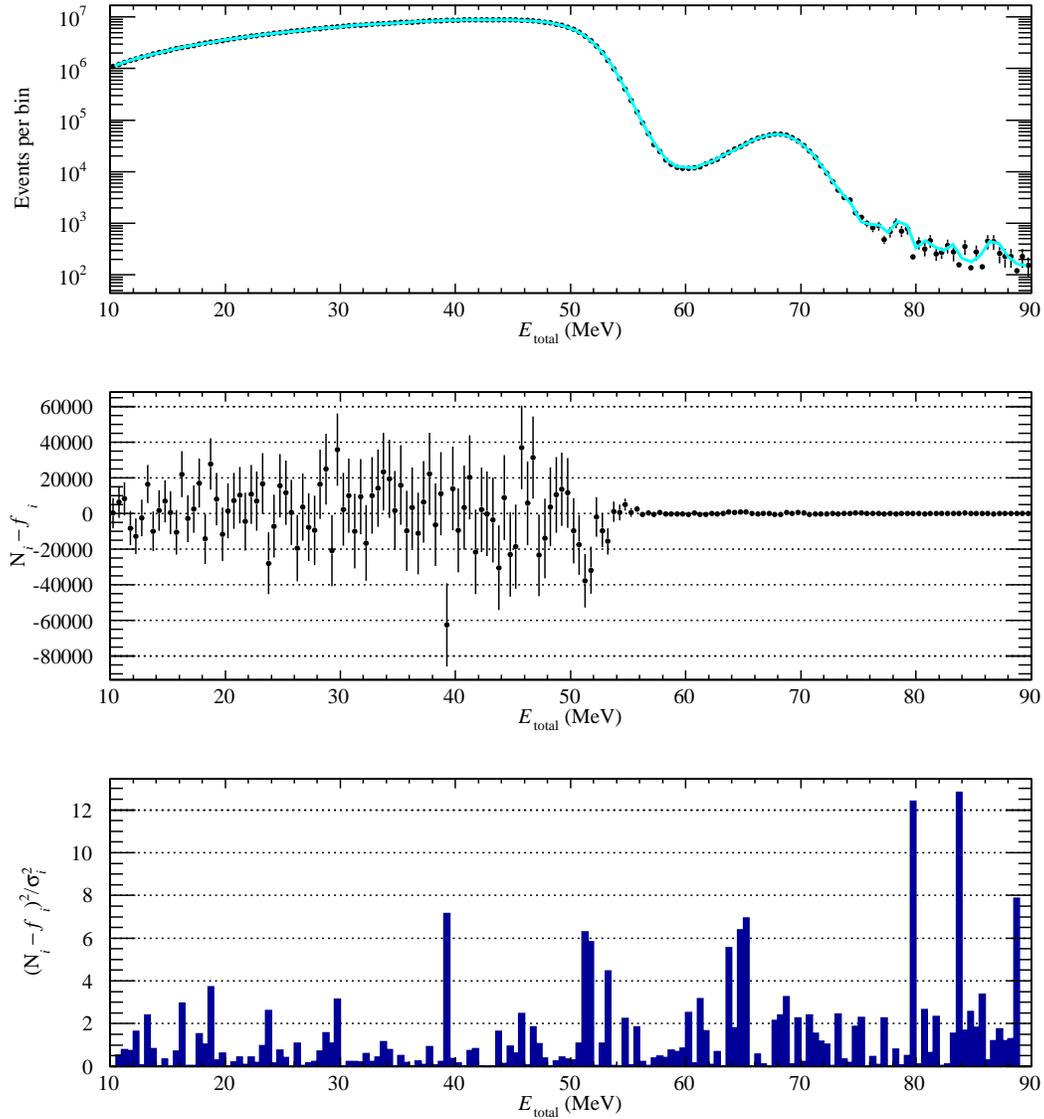
Figure 6.4: *A χ^2 fit to the E_{Total} spectrum. The fit function shown in the top panel consists of a linear combination of the PDFs for each process. The shapes of the PDFs are fixed. The relative PDF normalizations are the fit parameters. The middle panel shows the bin-wise difference between the bin content and the fit function. The bottom panel shows the bin-wise contribution to the χ^2.*

6.5.2 Decay Time

The next observable in our likelihood analysis is the time between the pion stop and the detected

outgoing final-state particle,

$$\Delta t = t_{\text{fsp}} - t_\pi \; . \tag{6.31}$$

For all five processes we will use the theoretical time distributions for the PDFs,

$$\Delta t_{\pi \to \mu \to e} = \begin{cases} 0 & \text{for} \quad \Delta t < 0 \\ \frac{1}{\tau_\mu - \tau_\pi} \left(e^{-t/\tau_\mu} - e^{-t/\tau_\pi} \right) & \text{for} \quad \Delta t \geq 0 \end{cases} ,$$

$$\Delta t_{\pi \to e} = \begin{cases} 0 & \text{for} \quad \Delta t < 0 \\ \frac{1}{\tau_\pi} e^{-t/\tau_\pi} & \text{for} \quad \Delta t \geq 0 \end{cases} ,$$

$$\Delta t_{\text{DIF}} = \begin{cases} 0 & \text{for} \quad \Delta t < 0 \\ \frac{1}{\tau_\mu} e^{-t/\tau_\mu} & \text{for} \quad \Delta t \geq 0 \end{cases} , \tag{6.32}$$

$$\Delta t_{\text{accidental}} = \frac{1}{\tau_\mu} e^{-t/\tau_\mu} \; , \text{ and}$$

$$\Delta t_{\text{hadronic}} = \delta(t - t_{\text{p}}^{\text{TOF}}) \; ,$$

where $\tau_\mu = 2197.03$ ns and $\tau_\pi = 26.033$ ns [33]. These theoretical PDFs must be smeared using a

Gaussian representing the timing resolution of our detector system. Both theoretical and smeared

functions are shown in Figures 6.6 and 6.7. Since the time signature for the prompt hadronic reaction

process is *essentially* just a delta function at the flight time of the proton after t_π we can obtain the

resolution directly. Figure 6.5(b) shows the Gaussian fit to the prompt peak resulting in an offset of

$t_{\text{p}}^{\text{TOF}} = 0.6216$ ns and a $\sigma_{\Delta t} = 0.3662$ ns. It is important to note that this value is a slight overestimate

of our timing resolution since the protons are not monoenergetic and already have a smeared time

signature even before detector resolution comes into play. Nevertheless, it is a good approximation.

(a) Δt *spectrum.*

(b) *Zoomed into the prompt region.*
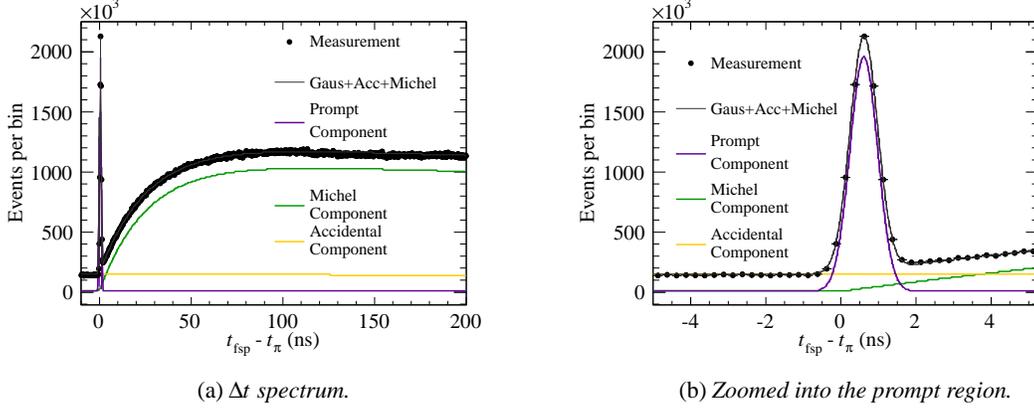
Figure 6.5: (a) *The $\Delta t = t_{\text{fsp}} - t_{\pi}$ spectrum for all five processes.* (b) *shows the prompt region in more detail. A Gaussian fit to the prompt pion-nucleon interaction peak gives an estimate of the resolution of our decay time observable.*

(a)

(b)
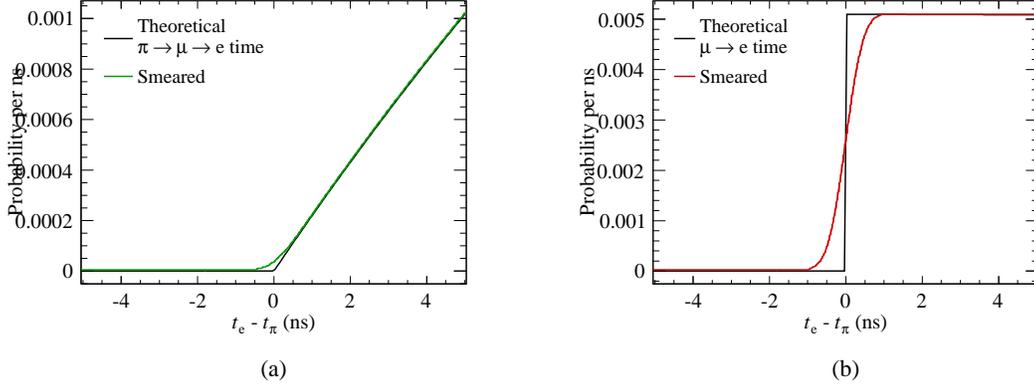
Figure 6.6: (a) $\Delta t = t_{\text{e}} - t_{\pi}$ *for the $\pi \to \mu \to$ e process,* (b) $\Delta t = t_{\text{e}} - t_{\pi}$ *for the decay-in-flight process, where* $t_{\pi} = t_{\mu} = 0$.

To verify that our PDFs properly describe the events in our data set, we performed a χ^2 fit yielding $\chi^2/N_{\text{dof}} = 1.30$. The resulting fit function, composed of a linear combination of the PDFs, is shown in Figure 6.8. The bottom panel shows the difference between the data value in each bin and the fit result, normalized to the bin contents, thus demonstrating excellent agreement throughout the entire 210 ns domain.
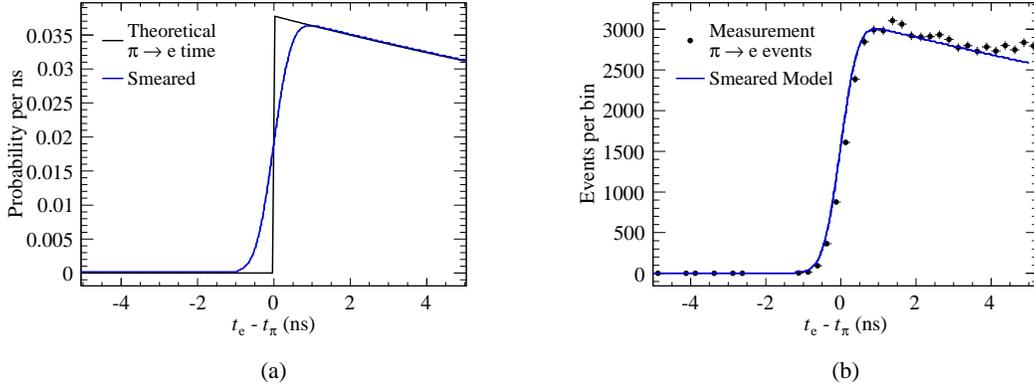
Figure 6.7: (a) $\Delta t = t_e - t_\pi$ *for the* $\pi \to e$ *process, and* (b) *shows the smeared function fit to a distribution of* $\pi \to e$ *events obtained from cuts.*

Minimum χ^2 Fit to Decay Time

To ensure a linear combination of the PDFs accurately describe the shape of the measured data we performed a minimum χ^2 fit analogous to the fit performed for E_{Total}. Using 0.25 ns bins and five degrees of freedom (the normalizations) we obtained $\chi^2/N_{\text{dof}} = 1089.59/835 = 1.30$. The corresponding correlation matrix for the fit parameters is,

$$
\rho =
\begin{array}{c}
\\
N_{\pi_{\mu 2}} \\
N_{\pi_{e2}} \\
N_{\text{Had}} \\
N_{\text{DIF}} \\
N_{\text{Acc}}
\end{array}
\begin{array}{c}
\begin{array}{ccccc}
N_{\pi_{\mu 2}} & N_{\pi_{e2}} & N_{\text{Had}} & N_{\text{DIF}} & N_{\text{Acc}}
\end{array} \\
\left(
\begin{array}{ccccc}
1 & -0.969 & 0.125 & -0.992 & 0.008 \\
-0.969 & 1 & -0.005 & 0.977 & -0.004 \\
0.125 & -0.005 & 1 & -0.109 & -0.033 \\
-0.992 & 0.977 & -0.109 & 1 & -0.078 \\
0.008 & -0.004 & -0.033 & -0.078 & 1
\end{array}
\right)
\end{array} ,
$$

and the normalizations are,

$$N_{\pi_{\mu 2}} = (4.73 \pm 0.017) \times 10^8 \ ,$$

$$N_{\pi_{e2}} = (1.30 \pm 141.08) \times 10^4 \ ,$$

$$N_{\mathrm{Had}} = (9.31 \pm 4.24) \times 10^3 \ , \tag{6.33}$$

$$N_{\mathrm{DIF}} = (3.85 \pm 19.74) \times 10^6 \ , \ \text{and}$$

$$N_{\mathrm{Acc}} = (1.71 \pm 0.015) \times 10^7 \ .$$

With a 97% correlation between $N_{\pi_{e2}}$ and $N_{\pi_{\mu 2}}$ when fitting this Δt observable alone, it doesn't make much sense to quote a value of $R_{\pi_{e2}}$ from this fit.

Figure 6.8 for Δt is analogous to Figure 6.4 for E_{Total}. We demonstrate that there is no specific region in the observable's domain that cannot be properly accounted for by a linear combination of the PDFs described in this section. It can be noted that there are systematic "wiggles" in the measured data, e.g., at $\Delta t \simeq 115$ ns and $\Delta t \simeq 155$ ns, the cause of which have not yet been determined.
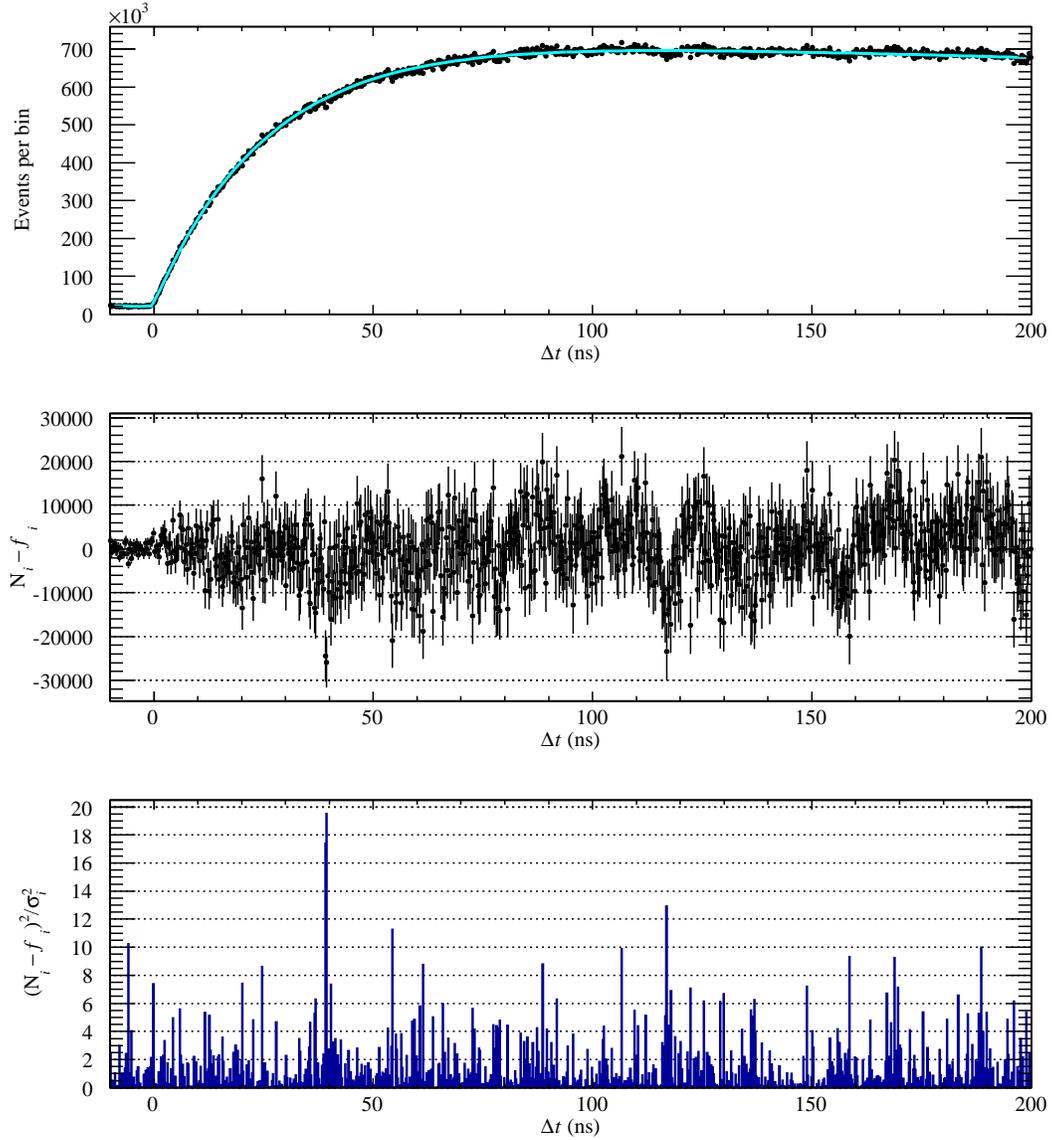
Figure 6.8: *The top panel shows a fit to the $\Delta t = t_{\rm fsp} - t_\pi$ spectrum with the PDFs for the five processes. The distribution had prompt hardonic reactions removed with cuts. The middle panel shows the bin-wise difference between the fit curve and the measurement data. The bottom panel shows the bin-wise contribution to the χ^2.*

Systematic Uncertainty in Decay Time PDFs

One major concern about the PEN experiment revolves around our knowledge of "time-zero", i.e., the pion stop time, $t = 0$ ns, and the absolute calibration of the time scale, i.e., the true value of the time-gate cutoff, $t = T$ ns. As stated in the experiment proposal, keeping the associated relative error on $R_{\pi_{e2}}$ under 2×10^{-4} imposes the requirement that the relative timing offset be known with a precision of 5 ps. Luckily, we used a single detector to calibrate both the pion and positron times, namely the active target, thus eliminating differences among trigger timing and energy threshold effects. Our decay-time observable, Δt, is formed from the prediction of the pulse positions in the target waveform and each component is calibrated to a very high precision. The uncertainties on our predicted pion and muon times are determined by allowing all of the target waveform fit parameters to vary[4] and taking the difference between the fit and predicted times. For each run we obtained a Guassian distribution centered at zero (when the predictions are calibrated properly) with some timing spread, σ, describing the random fluctuations in our measurement process. Figures 6.9 and 6.10 show $\overline{\delta t}$ and $\sigma_{\delta t}$ versus run in the later half of 2008. It should be clear that the standard deviation, σ, of the Gaussian is not indicative of our uncertainty in the value of t_π or t_e. Instead, any systematic shift of the mean away from zero would put a bias on our measurement and should be compensated for in our calibrations. The uncertainty on that mean value corresponds to our systematic uncertainty in this decay-time observable. Upon fitting a Gaussian function to $\delta t = t^{\text{Fit}} - t^{\text{Pred}}$ we find the mean of the Gaussian and the associated error in that mean,

---

[4] You may recall that the times of each pulse were fixed to the predicted values in the final analysis.

$$\overline{\delta t_\pi} = \overline{t_\pi^{\text{Fit}} - t_\pi^{\text{Pred}}} = -0.056 \pm 0.029 \text{ ps} \text{ and}$$

$$\overline{\delta t_e} = \overline{t_e^{\text{Fit}} - t_e^{\text{Pred}}} = -0.20 \pm 0.21 \text{ ps} . \tag{6.34}$$

It follows then, by summing the errors in quadrature that the systematic uncertainty in Δt (and therefore our uncertainty in $\Delta t = 0$ also known as "time-zero") is,

$$\delta(\Delta t) \leq \sqrt{\left(\overline{\delta t_\pi}\right)^2 + \left(\overline{\delta t_e}\right)^2} = 0.21 \text{ ps} , \tag{6.35}$$

which is far better than the required precision.
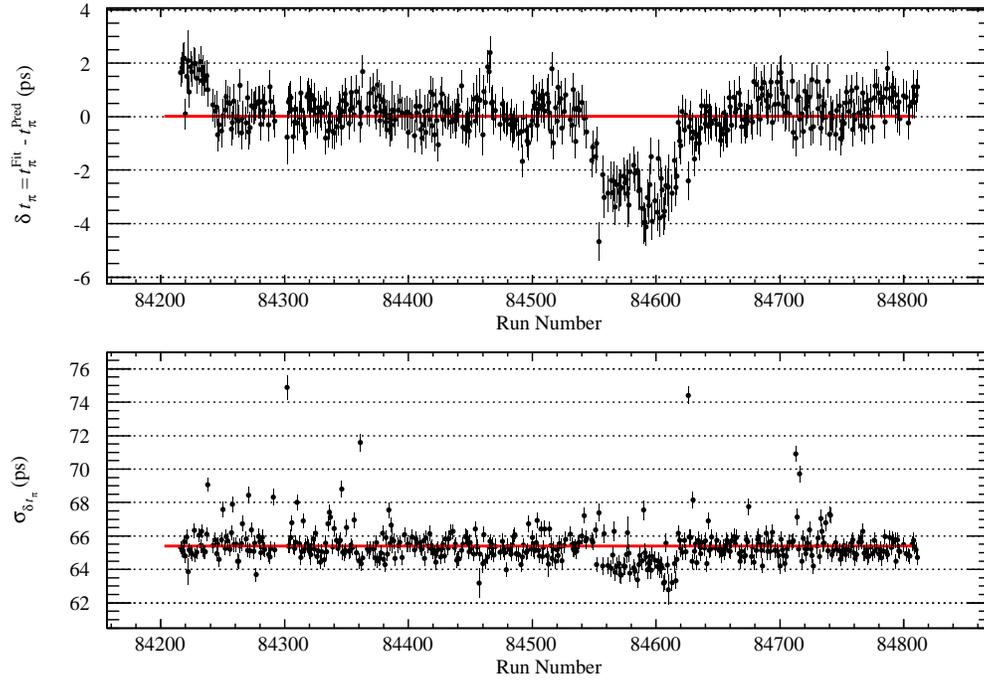
Figure 6.9: *The top panel shows the mean value of a Gaussian fit to $\delta t_\pi = t_\pi^{\mathrm{Fit}} - t_\pi^{\mathrm{Pred}}$ versus run for the later half of the 2008 dataset. The line fit is used to determine the mean and error on the mean for this set of runs. The bottom panel shows the sigma of the Gaussian fit demonstrating the random fluctuations (not bias) on this observable.*
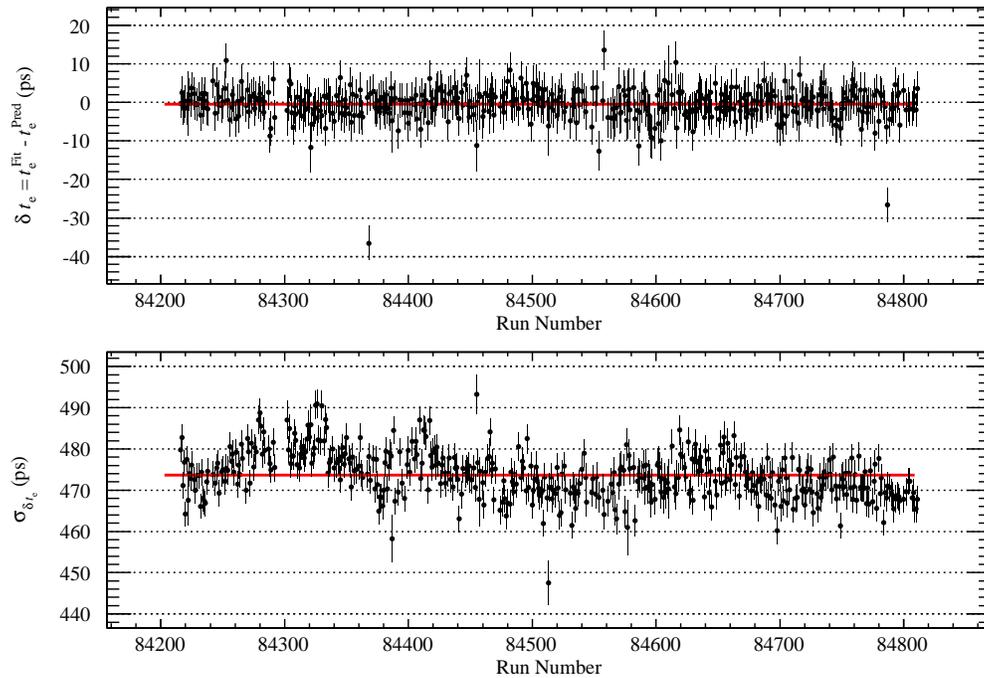
Figure 6.10: *The top panel shows the mean value of a Gaussian fit to $\delta t_e = t_e^{\mathrm{Fit}} - t_e^{\mathrm{Pred}}$ versus run for the later half of the 2008 dataset. The line fit is used to determine the mean and error on the mean for this set of runs. The bottom panel shows the sigma of the Gaussian fit demonstrating the random fluctuations (not bias) on this observable.*

6.5.3 Differential Energy Loss

The differential (or specific) energy loss is the amount of energy deposited by a particle per unit length as the particle traverses a medium. Figure 6.11 shows the correction for incident angle, θ, dependence and the clear separation of positrons with $(dE/dx)_{PH} \sim 1.82$ MeV/cm and protons with $(dE/dx)_{PH} \sim 8.2$ MeV/cm in the plastic hodoscope.

(a) Energy deposited in the plastic hodoscope.

(b) Energy deposited per unit length in the plastic hodoscope.
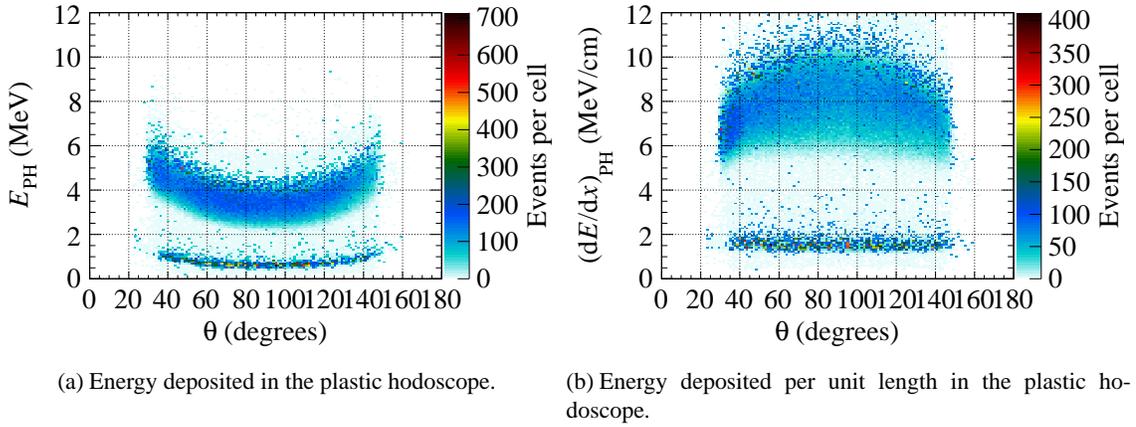
Figure 6.11: *The plastic scintillator hodoscope is useful for discriminating between positrons with dE/dx $\sim$ 1.5 MeV/cm and protons with dE/dx $\sim$ 6.5 MeV/cm.*

As this observable is identical for all positron final-state processes, its utility is limited to distinguishing between proton and positron final-state events. Since we placed cuts (Table 6.3) to remove events that underwent prompt pion-nucleon reactions, the differential energy loss observable may not be of much help. Nevertheless, some proton final-state events may leak into our data set.

The hodoscope differential energy loss is also used to select protons considering their total energy deposited. Plotting the differential energy loss versus total energy loss, Figure 6.12, we see that $(dE/dx)_{PH}$ for protons varies as a function of total energy, but is essentially constant for the minimum ionizing positrons. We are able to select protons by requiring,

$$(dE/dx)_{PH} > 13.0 \exp(-0.011\, E_{\text{Total}}) \; . \tag{6.36}$$
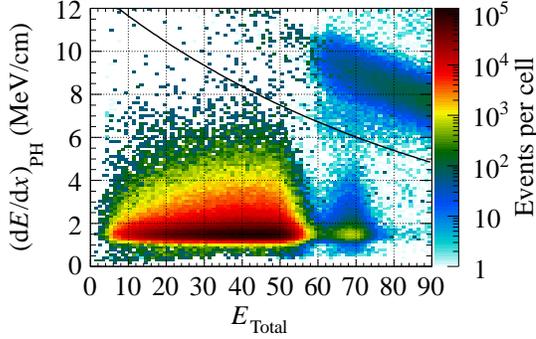
Figure 6.12: *Proton identification with energy deposition. The black curve corresponds to (6.36).*
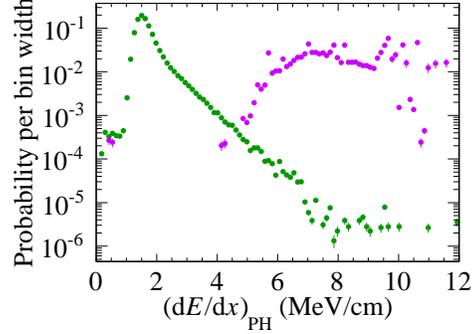
Figure 6.13: *Differential energy loss distributions from simulated positron (green) and proton (purple) final-state events.*

The observable $\Delta\phi_{\mathrm{MWPC2,PH}}$ is the difference in azimuthal angle, ϕ, of the track passing through MWPC2 and the nearest PH stave boundary. Due to the geometry of the staves at the boundary, a systematic effect occurs in which less light is produced (collected) when the particle traverses the plastic at the edge. We therefore see a dip in the differential energy loss at each stave boundary ($\Delta\phi_{\mathrm{MWPC2,PH}} \simeq 0$) as shown in Figure 6.14. Only events with $|\Delta\phi_{\mathrm{MWPC2,PH}}| > 1.8°$ were considered in the final analysis.

(a) *Measurement*

(b) *Simulated positron final-state events*

(c) *Simulated proton final-state events*
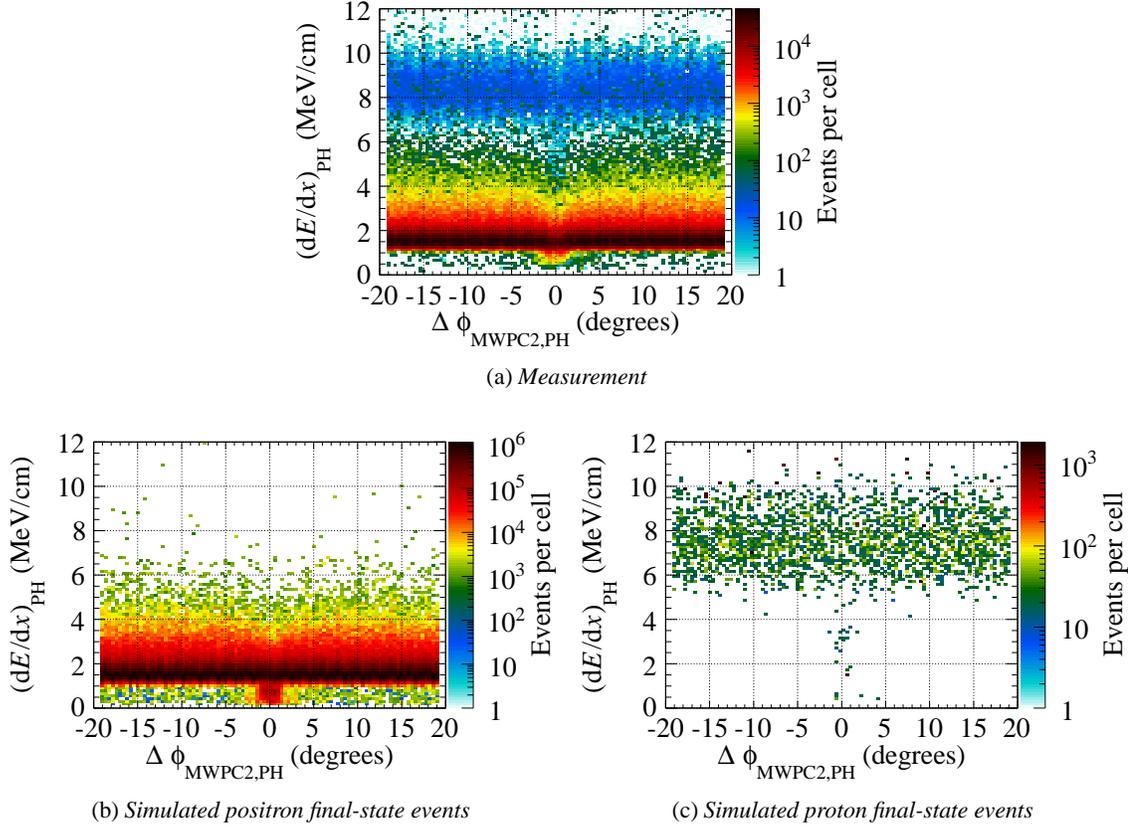
Figure 6.14: *Differential energy loss versus difference in ϕ between the track passing through MWPC2 and the nearest PH stave boundary for,* (a) *measurement data containing both positron and proton final-state events,* (b) *simulated data with positron final-state particles, and* (c) *simulated data with proton final-state particles.*

6.5.4 Target Waveform Integrals

So far we don't have an observable with a region occupied exclusively by pion decays-in-flight. A possible observable that may be used in the final analysis is the integral of the "subtracted" target waveform,

$$I_{w''} = \sum_{n=t_\pi-5\,\mathrm{ns}}^{T} w''[n] \ , \tag{6.37}$$

where the sum runs from 5 ns before the pion pulse time to the end of the recorded waveform, T, typically 180 ns after the positron pulse time ($T \simeq t_{\mathrm{e}} + 180$ ns). Figure 5.8 shows the filtered target

waveform, $w''[n]$, before and after the predicted pion and positron pulses have been subtracted.

When the pion decays in flight and the muon is emitted in the forward direction the muon is boosted forward. Sometimes the muon travels completely through the target, exiting through the downstream face. Consequently, the muon deposits more energy in the target than we expect and the predicted pusle is too small. In this case, $w''[n]$ will contain a residual pulse at t_π.

For pion decays-in-flight where the muon is emitted backwards, just the opposite happens. The muon is boosted backwards and subsequently deposits a smaller amount of energy in the target than we predict. The subtraction procedure subtracts too much and we're left with a dip in the waveform. For decays-in-flight with the muon emitted at certain angles with respect to the pion momentum the energy deposited can resemble a $\pi \to e$ decay and thus present a dangerous systematic uncertainty.

(a) *Simulation*   (b) *Measurement*
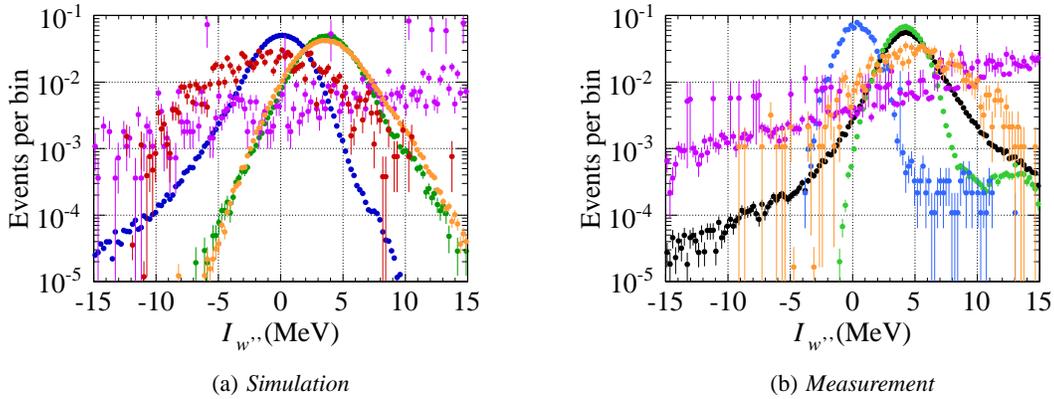
Figure 6.15: *Subtracted target waveform integral for the five main processes from* (a) *simulation and* (b) *measurement. The five processes are color coded as follows:* $\pi_{\mu2}$ = *green,* π_{e2} = *blue, hadronic = purple, decay-in-flight = red, accidental = orange, measurement (with hadronic interactions removed) = black.*

Figure 6.15 shows $I_{w''}$ for simulation and measurement. The π_{e2} process gives a peak centered at zero, since both pion and positron peaks are properly subtracted and we're left with a flat waveform. After the pion and positron pulses are subtracted from $\pi_{\mu2}$ events the muon pulse remains thus giving the peak centered at 4.12 MeV. Accidental coincidences also have a muon in the target, so this observable appears essentially the same for the $\pi_{\mu2}$ and accidental coincidence processes. As

discussed above, the distribution for decays in flight is more spread out. Once the events undergoing hadronic interactions are removed using cuts, the only events populating the left shoulder of this observable are decays in flight.

The distributions for the individual processes in Figure 6.15(b) were obtained using cuts. Both the $\pi_{\mu2}$ and π_{e2} processes exhibit a high energy shoulder resulting from extra beam contamination (beam halo), or pile-up particles. The accidental coincidence distribution was determined only using the porch $(t_e - t_\pi < 0)$ region. We don't have any way to isolate pion decays in flight, so that process is not plotted in Figure 6.15(b).

The primary goal of this section is to describe a possible observable to be included in the likelihood analysis. Considerable difficulty is involved in determining the PDFs for each process for this observable, as there is only a minimal set of cuts that won't distort the shapes. We would have to rely heavily on simulation (especially for pion decays in flight) and calculate conservative systematic error estimates.

6.5.5 Pion Decay Vertex

As discussed in the previous section, the daughter muons from pions decaying in flight are boosted either upstream or downstream. This process results in a smeared stopping distribution for the stopped muons, as opposed to the narrow stopping distribution of the stopped pions.

Figure 6.16 shows the difference between the predicted z coordinate of the pion stop position and the calculated z coordinate from the final-state particle track,

$$\Delta z = z_0^{\text{Track}} - z_0^{\text{Pred}} \ . \tag{6.38}$$

The prediction for the stopping vertex is based on the particle's momentum, obtained from time-of-flight between the forward beam counter and the degrader. The origin of the decay track is determined to be the intersection between the decay track determined from the MWPCs and the beam track determined using the wedged active degrader. We clearly see that the pion decays in flight have a much broader distribution. It is also worth noting that the distribution for events that undergo a pion-nucleon interaction is shifted since the interaction may occur before the pion comes to a complete stop. The distribution for $\pi \to \mu \to$ e events is slightly broader when compared to the distribution for $\pi \to$ e events due to the ~ 2 mm path length of the muon inside the target.
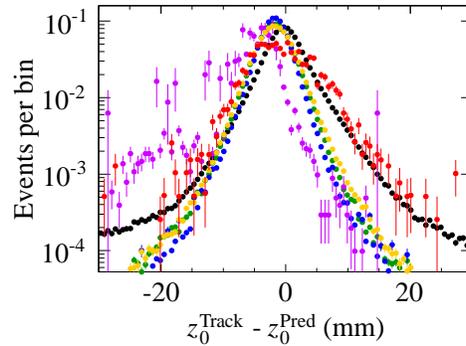
Figure 6.16: *Difference between the z coordinate of the predicted pion stop position and the calculated z from the final-state particle track. Black = measurement, and the colored data points correspond to individual simulated processes, where Green = $\pi \to \mu \to$ e process, Blue = $\pi \to$ e process, Purple = pion hadronic interactions, Orange = accidental coincidences, and Red = pion decays-in-flight.*

The stopping distribution is correlated with the energy deposited in the target. When the daughter muon is boosted forward, the recorded energy in the target will be larger and the z of the muon stop position will be larger. Figure 6.17 shows the target waveform integral versus Δz showing the correlation. Figures 6.18 and 6.19 show the distributions for each individual simulated process.

In summary, Δz might be useful to discriminate the pernicious pion decays in flight. As this observable is correlated with the target waveform integral, we should consider using two-dimensional PDFs as hown in Figures 6.18 and 6.19.

(a) *Measurement*

(b) *Measurement*
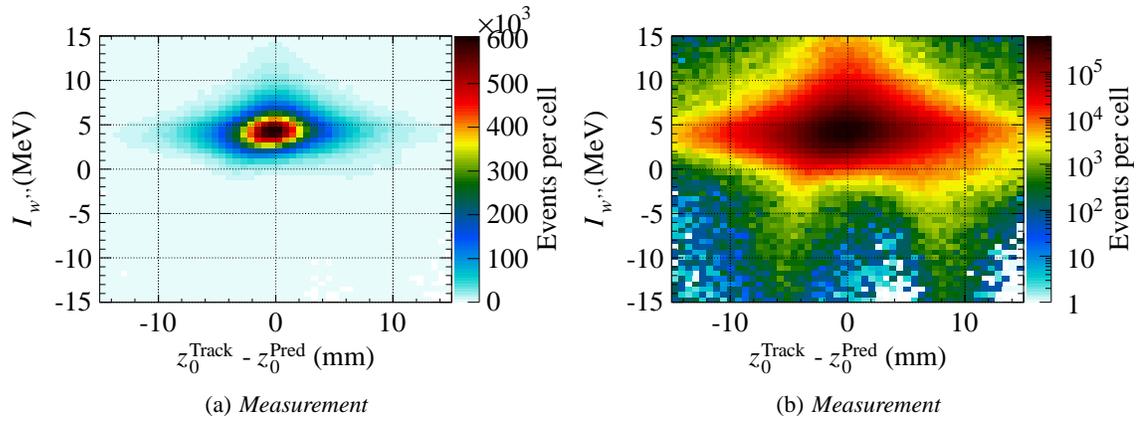
Figure 6.17: *Target waveform integral versus the difference between the z coordinate of the predicted pion stop position and the calculated z from the final-state particle track.* (a) *shows the Events per cell on a linear scale demonstrating the dominant peak due to the* $\pi \to \mu \to e$ *events.* (b) *shows the Events per cell on a logarithmic scale so we can see the structure due to the other processes.*

(a) *Simulated* $\pi \to \mu \to e$ *events*

(b) *Simulated* $\pi \to e$ *events*

(c) *Simulated pion-nucleon interaction events*

(d) *Simulated accidental coincidence events*

(e) *Simulated pion decay-in-flight events*
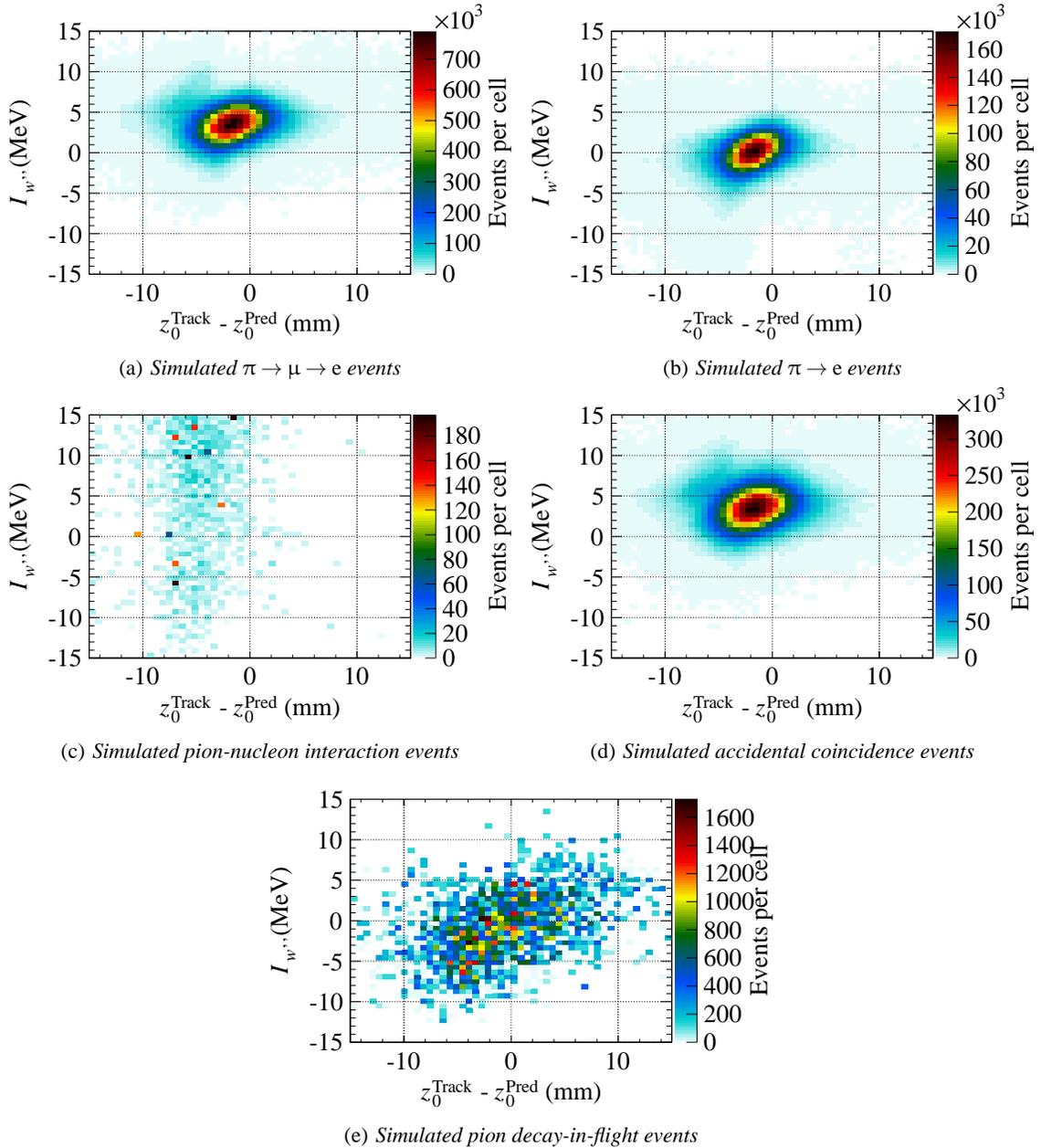
Figure 6.18: *A comparison of the five main processes using simulated data. Plotted is the target waveform integral versus the difference between the z coordinate of the predicted pion stop position and the calculated z from the final-state particle track.*

(a) *Simulated $\pi \to \mu \to e$ events*

(b) *Simulated $\pi \to e$ events*

(c) *Simulated pion-nucleon interaction events*

(d) *Simulated accidental coincidence events*

(e) *Simulated pion decay-in-flight events*
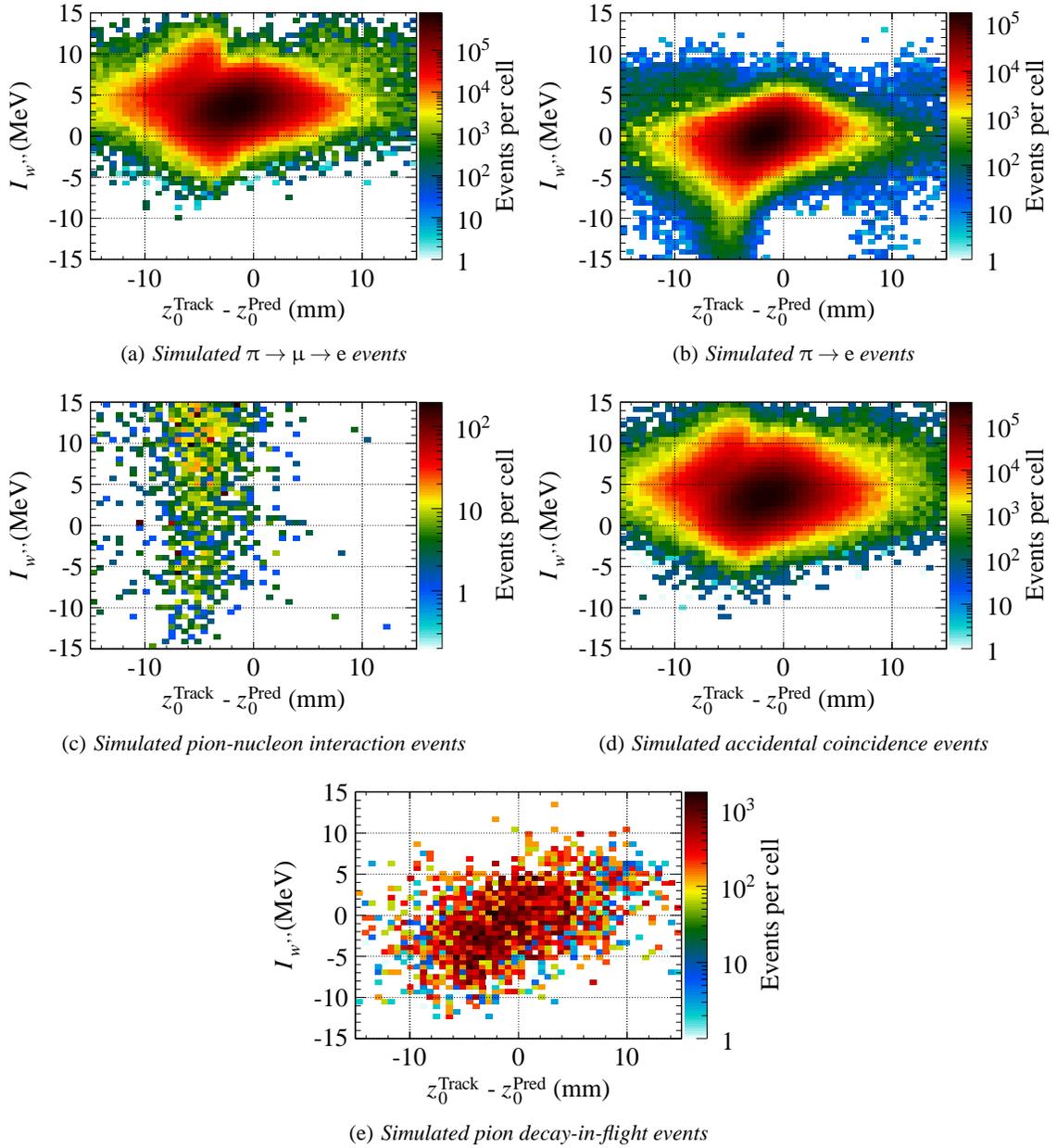
Figure 6.19: *Same as Figure 6.18 but with the Events per cell shown on a logarithmic scale.*

6.5.6 Probability of Pile-Up

Using multi-hit TDCs we were able to store many hits in the degrader from pions leading up to the actual pion hit recorded as the triggered event. The logarithm of the probability of the current event containing a muon parked in the target from a previous beam particle is given by,

$$\text{``}P\text{''}_{\text{pile-up}} = \ln\left[P_{\text{pile-up}}\right] = \ln\left[\sum_{k=1}^{\ell} e^{-|\mathrm{d}t_k|/\tau_\mu}\right]\ , \tag{6.39}$$

where ℓ is the number of TDC hits and $\mathrm{d}t_k$ is the time between the previous hit and t_π^{DEG} of the triggered event. $\text{``}P\text{''}_{\text{pile-up}}$ ranges from -13.8 for the lowest probability of having a pile-up event to ~ 0 for the highest probability of having a muon parked in the target from a previous pion. The value of -13.8 corresponds to not having any previous hits recorded in the finite range of the TDC.
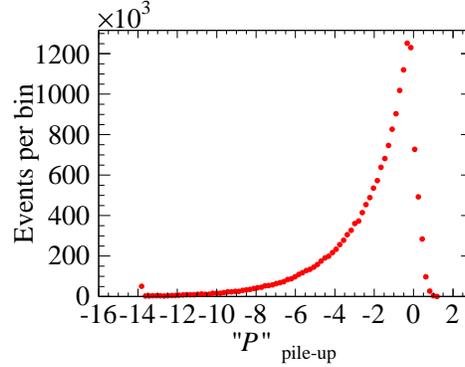
Figure 6.20: *Logarithm of the probability of a triggered event containing a muon parked in the target originating from a previous beam pion. A more negative value corresponds to a lower probability of pile-up.*

We originally considered using $\text{``}P\text{''}_{\text{pile-up}}$ as an observable in the likelihood analysis, but at this point we decided against it. For one, the PDFs would be the same for many of the processes and therefore it wouldn't add much discriminatory power to the analysis. Also, we would have to

properly simulate the previous hits in the TDC in order to maintain our philosophy that our analysis

tools do not know if the data they are analyzing is from measurement or simulation. We've decided

to just make cuts, "P"$_{\text{pile-up}} < -1.0$, to remove many pile-up events at the cost of removing some

good π_{e2} and $\pi_{\mu2}$ events.

6.6   Most Likely Estimate of $R_{\pi_{e2}}$

I've written a C++ program to determine the most likely values of the parameters θ_i. The pro-

gram uses the un-binned maximum likelihood technique and analyzes each event individually. The

program is divided into three stages.

Maximum Likelihood: Stage 1

The first stage applies all the cuts to the data that will be used in the final likelihood analysis and

fills histograms for each observable. In the example demonstrated in this dissertation I've used only

two observables, namely the decay time, Δt, and the total energy of the final-state particle, E_{Total}.

The program can be run in parallel, filling histograms for each run. Later the histograms can be

combined to create one histogram for each observable containing every event that will enter the

likelihood analysis.

The goal of the first stage is to determine the range over which we must scan each parameter

ensuring the range encompasses its most likely value. We find this range by performing a binned

minimum χ^2 fit. It is a simultaneous fit to both the decay time and total energy observables with

shared normalization parameters yielding,

$$N_{\pi_{\mu2}} = (4.7641 \pm 0.0018) \times 10^8 \text{ events,}$$

$$N_{\pi_{e2}} = (8.0591 \pm 0.0175) \times 10^5 \text{ events,}$$

$$N_{\text{Had}} = 0.00000018 \pm 48.2138 \text{ events,} \qquad (6.40)$$

$$N_{\text{DIF}} = 0.00336783 \pm 43558.1520 \text{ events, and}$$

$$N_{\text{Acc}} = (1.6803 \pm 0.0135) \times 10^7 \text{ events.}$$

This fit has a $\chi^2/N_{\text{dof}} = 1310.7/995 = 1.317$. Since we've performed a fit determining the normalizations of each process, we might as well discuss the results and present a branching ratio from this method before discussing the next stage in the likelihood analysis.

The correlation matrix from the χ^2 fit is

	$N_{\pi_{\mu2}}$	$N_{\pi_{e2}}$	N_{Had}	N_{DIF}	N_{Acc}
$N_{\pi_{\mu2}}$	1	0.0029	0.0000066	0.00060	-0.72
$N_{\pi_{e2}}$	0.0029	1	0.000031	0.000013	-0.016
$\rho = N_{\text{Had}}$	0.0000066	0.000031	1	0.0000000049	0.000011
N_{DIF}	0.00060	0.000013	0.0000000049	1	0.00016
N_{Acc}	-0.72	-0.016	0.000011	0.00016	1

and the covariance matrix is

$$\mathbf{V} = \begin{array}{c} \\ N_{\pi_{\mu2}} \\ \\ N_{\pi_{e2}} \\ \\ N_{\text{Had}} \\ \\ N_{\text{DIF}} \\ \\ N_{\text{Acc}} \end{array} \begin{array}{ccccc} N_{\pi_{\mu2}} & N_{\pi_{e2}} & N_{\text{Had}} & N_{\text{DIF}} & N_{\text{Acc}} \\ \left(\begin{array}{ccccc} 3.26 \times 10^{10} & 9.31 \times 10^5 & -7.13 \times 10^{-3} & 2.62 \times 10^3 & -1.76 \times 10^{10} \\ 9.31 \times 10^5 & 3.07 \times 10^6 & 3.24 \times 10^{-4} & 5.37 \times 10^{-1} & -3.88 \times 10^6 \\ -7.13 \times 10^{-3} & 3.24 \times 10^{-4} & 3.52 \times 10^{-5} & 7.06 \times 10^{-10} & 8.91 \times 10^{-3} \\ 2.62 \times 10^3 & 5.37 \times 10^{-1} & 7.06 \times 10^{-10} & 5.90 \times 10^2 & 5.44 \times 10^2 \\ -1.76 \times 10^{10} & -3.88 \times 10^6 & 8.91 \times 10^{-3} & 5.44 \times 10^2 & 1.81 \times 10^{10} \end{array}\right) \end{array} .$$

The branching ratio is therefore[5]

$$\boxed{R^{(\chi^2)} = \frac{N_{\pi_{e2}}}{N_{\pi_{\mu2}}} = [1.6915942 \pm 0.0037338(\text{stat.})] \times 10^{-3}}$$
$$\left(\Delta R^{(\chi^2)}/R^{(\chi^2)} = 0.0022073\right) , \tag{6.41}$$

and using the parameterization in (6.15) we have

$$\boxed{\theta_0^{(\chi^2)} = \frac{N_{\pi_{e2}}}{N_{\pi_{\mu2}} + N_{\pi_{e2}}} = [1.6915943 \pm 0.0037338(\text{stat.})] \times 10^{-3}}$$
$$\left(\Delta\theta_0^{(\chi^2)}/\theta_0^{(\chi^2)} = 0.0022036\right) , \tag{6.42}$$

where we have propagated the uncertainties using

$$\sigma_R^2 = \left|\frac{\partial R}{\partial N_{\pi_{\mu2}}}\right|^2 \mathbf{V}_{N_{\pi_{\mu2}}N_{\pi_{\mu2}}} + \left|\frac{\partial R}{\partial N_{\pi_{e2}}}\right|^2 \mathbf{V}_{N_{\pi_{e2}}N_{\pi_{e2}}} + 2\frac{\partial R}{\partial N_{\pi_{\mu2}}}\frac{\partial R}{\partial N_{\pi_{e2}}}\mathbf{V}_{N_{\pi_{\mu2}}N_{\pi_{e2}}}$$
$$= \frac{N_{\pi_{e2}}^2}{N_{\pi_{\mu2}}^4}\mathbf{V}_{N_{\pi_{\mu2}}N_{\pi_{\mu2}}} + \frac{1}{N_{\pi_{\mu2}}^2}\mathbf{V}_{N_{\pi_{e2}}N_{\pi_{e2}}} + 2\frac{-N_{\pi_{e2}}}{N_{\pi_{\mu2}}^2}\frac{1}{N_{\pi_{\mu2}}}\mathbf{V}_{N_{\pi_{\mu2}}N_{\pi_{e2}}} \tag{6.43}$$

---

[5]Normally I would truncate the least significant digits. I decided to retain them to illustrate the precision to which our various analysis methods agree.

and

$$
\begin{aligned}
\sigma_{\theta_0}^2 &= \left| \frac{\partial \theta_0}{\partial N_{\pi_{\mu2}}} \right|^2 \mathbf{V}_{N_{\pi_{\mu2}} N_{\pi_{\mu2}}} + \left| \frac{\partial \theta_0}{\partial N_{\pi_{e2}}} \right|^2 \mathbf{V}_{N_{\pi_{e2}} N_{\pi_{e2}}} + 2 \frac{\partial \theta_0}{\partial N_{\pi_{\mu2}}} \frac{\partial \theta_0}{\partial N_{\pi_{e2}}} \mathbf{V}_{N_{\pi_{\mu2}} N_{\pi_{e2}}} \\
&= \frac{N_{\pi_{e2}}^2}{\left(N_{\pi_{\mu2}} + N_{\pi_{e2}} \right)^4} \mathbf{V}_{N_{\pi_{\mu2}} N_{\pi_{\mu2}}} + \frac{N_{\pi_{\mu2}}^2}{\left(N_{\pi_{\mu2}} + N_{\pi_{e2}} \right)^4} \mathbf{V}_{N_{\pi_{e2}} N_{\pi_{e2}}} \\
&\quad + 2 \frac{-N_{\pi_{e2}}}{\left(N_{\pi_{\mu2}} + N_{\pi_{e2}} \right)^2} \frac{N_{\pi_{\mu2}}}{\left(N_{\pi_{\mu2}} + N_{\pi_{e2}} \right)^2} \mathbf{V}_{N_{\pi_{\mu2}} N_{\pi_{e2}}} \ .
\end{aligned}
\tag{6.44}
$$

Applying the acceptances[6] determined from simulation, $A_{\pi \to \mu} = 0.0271648$ and $A_{\pi \to e} = 0.411572$, with $a_i \equiv 1/A_i$, we obtain

$$
\begin{aligned}
\tilde{R}^{(\chi^2)} &= \frac{a_{\pi_{e2}} N_{\pi_{e2}}}{a_{\pi_{\mu2}} N_{\pi_{\mu2}}} \\
&= \left[1.1164924 \pm 0.0024644 (\text{stat.}) \right] \times 10^{-4} \\
&\left(\Delta R^{(\chi^2)} / R^{(\chi^2)} = 0.0022073 \right) \ .
\end{aligned}
\tag{6.45}
$$

Using the parameterization in (6.15) we have

$$
\begin{aligned}
\tilde{\theta}_0^{(\chi^2)} &= \frac{a_{\pi_{e2}} N_{\pi_{e2}}}{\left(a_{\pi_{\mu2}} N_{\pi_{\mu2}} + a_{\pi_{e2}} N_{\pi_{e2}} \right)} \\
&= \left[1.1163678 \pm 0.0024639 (\text{stat.}) \right] \times 10^{-4} \\
&\left(\Delta \theta_0^{(\chi^2)} / \theta_0^{(\chi^2)} = 0.0022071 \right) \ ,
\end{aligned}
\tag{6.46}
$$

---

[6]These acceptance values are correct to within a few percent. The acceptance values were multiplied by a random number close to 1.0, so that we blind ourselves from our actual measured value. Once the collaboration is satisfied with our analysis we will remove the multiplicative random number thus revealing our final result.

with the uncertainties propagated according to

$$
\begin{aligned}
\sigma_{\tilde{R}}^2 &= \left|\frac{\partial \tilde{R}}{\partial N_{\pi_{\mu 2}}}\right|^2 \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{\mu 2}}} + \left|\frac{\partial \tilde{R}}{\partial N_{\pi_{e2}}}\right|^2 \mathbf{V}_{N_{\pi_{e2}} N_{\pi_{e2}}} + 2 \frac{\partial \tilde{R}}{\partial N_{\pi_{\mu 2}}} \frac{\partial \tilde{R}}{\partial N_{\pi_{e2}}} \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{e2}}} \\
&= \frac{\left(a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{e2}}\right)^2}{\left(a_{\pi_{\mu 2}} N_{\pi_{\mu 2}}\right)^4} \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{\mu 2}}} + \frac{\left(a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{\mu 2}}\right)^2}{\left(a_{\pi_{\mu 2}} N_{\pi_{\mu 2}}\right)^4} \mathbf{V}_{N_{\pi_{e2}} N_{\pi_{e2}}} \\
&\quad + 2 \frac{\left(-a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{e2}}\right)\left(a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{\mu 2}}\right)}{\left(a_{\pi_{\mu 2}} N_{\pi_{\mu 2}}\right)^4} \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{e2}}}
\end{aligned}
\tag{6.47}
$$

and

$$
\begin{aligned}
\sigma_{\tilde{\theta}_0}^2 &= \left|\frac{\partial \tilde{\theta}_0}{\partial N_{\pi_{\mu 2}}}\right|^2 \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{\mu 2}}} + \left|\frac{\partial \tilde{\theta}_0}{\partial N_{\pi_{e2}}}\right|^2 \mathbf{V}_{N_{\pi_{e2}} N_{\pi_{e2}}} + 2 \frac{\partial \tilde{\theta}_0}{\partial N_{\pi_{\mu 2}}} \frac{\partial \tilde{\theta}_0}{\partial N_{\pi_{e2}}} \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{e2}}} \\
&= \frac{\left(a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{e2}}\right)^2}{\left(a_{\pi_{\mu 2}} N_{\pi_{\mu 2}} + a_{\pi_{e2}} N_{\pi_{e2}}\right)^4} \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{\mu 2}}} \\
&\quad + \frac{\left(a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{\mu 2}}\right)^2}{\left(a_{\pi_{\mu 2}} N_{\pi_{\mu 2}} + a_{\pi_{e2}} N_{\pi_{e2}}\right)^4} \mathbf{V}_{N_{\pi_{e2}} N_{\pi_{e2}}} \\
&\quad + 2 \frac{\left(-a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{e2}}\right)\left(a_{\pi_{\mu 2}} a_{\pi_{e2}} N_{\pi_{\mu 2}}\right)}{\left(a_{\pi_{\mu 2}} N_{\pi_{\mu 2}} + a_{\pi_{e2}} N_{\pi_{e2}}\right)^4} \mathbf{V}_{N_{\pi_{\mu 2}} N_{\pi_{e2}}}
\end{aligned}
\tag{6.48}
$$

Stage 1 has now given us an estimate for the normalization (6.40) of each process, $\hat{N}_i \pm \sigma_{N_i}$. We set the range for each parameter, θ_i, to $\theta_i \pm k\sigma_{\theta_i}$, where k is some constant greater than 1, chosen by hand for each range.

Maximum Likelihood: Stage 2

The second stage of the maximum likelihood analysis determines the most likely estimates of the parameters. For each event, the negative logarithm of the likelihood (6.14) is calculated for each value of the parameters, θ_i. We note that both the hadronic reaction and the decay-in-flight background fractions are essentially zero from the χ^2 fit. Therefore, to save computing time, we decided

to only have five steps each for θ_1 and θ_2, and twenty steps each for θ_0 and θ_3. That means the negative log of the likelihood is calculated $20 \times 5 \times 5 \times 20 = 1000$ times for each event. The value is added to the appropriate bin of a 4-dimensional histogram[7]. Stage 2 can also be run in parallel, producing a 4-dimensional histogram of negative log-likelihoods for each run which can be summed together.

The most likely estimates for the parameters were found to be

$$\hat{\theta}_0 = 0.00168223,$$

$$\hat{\theta}_1 = 5.06168 \times 10^{-6},$$

$$\hat{\theta}_2 = 9.32113 \times 10^{-6}, \text{ and}$$ (6.49)

$$\hat{\theta}_3 = 0.0341448,$$

and correspond to the minimum in the 4-dimensional negative log-likelihood histogram. Alternatively, we can shift the minimum to zero and exponentiate, producing the 4-dimensional likelihood distribution, in which case the best estimates correspond to the location of the maximum.

Figure 6.21 shows a three dimensional view of the projection of the likelihood onto the θ_0, θ_3 plane. The projections onto all possible two-dimensional combinations of θ_i, θ_j of the negative log-likelihood are shown in Figure 6.22, and of the likelihood are shown in Figure 6.23. The projections onto the one-dimensional θ_i are shown in Figure 6.24. The uncertainties cannot be read from these projections. Instead one must calculate the covariance matrix including the weights for each event.

---

[7]The code is written such that it is trivial to extend to an N-dimensional histogram when considering more parameters, $\theta_0...\theta_N$.
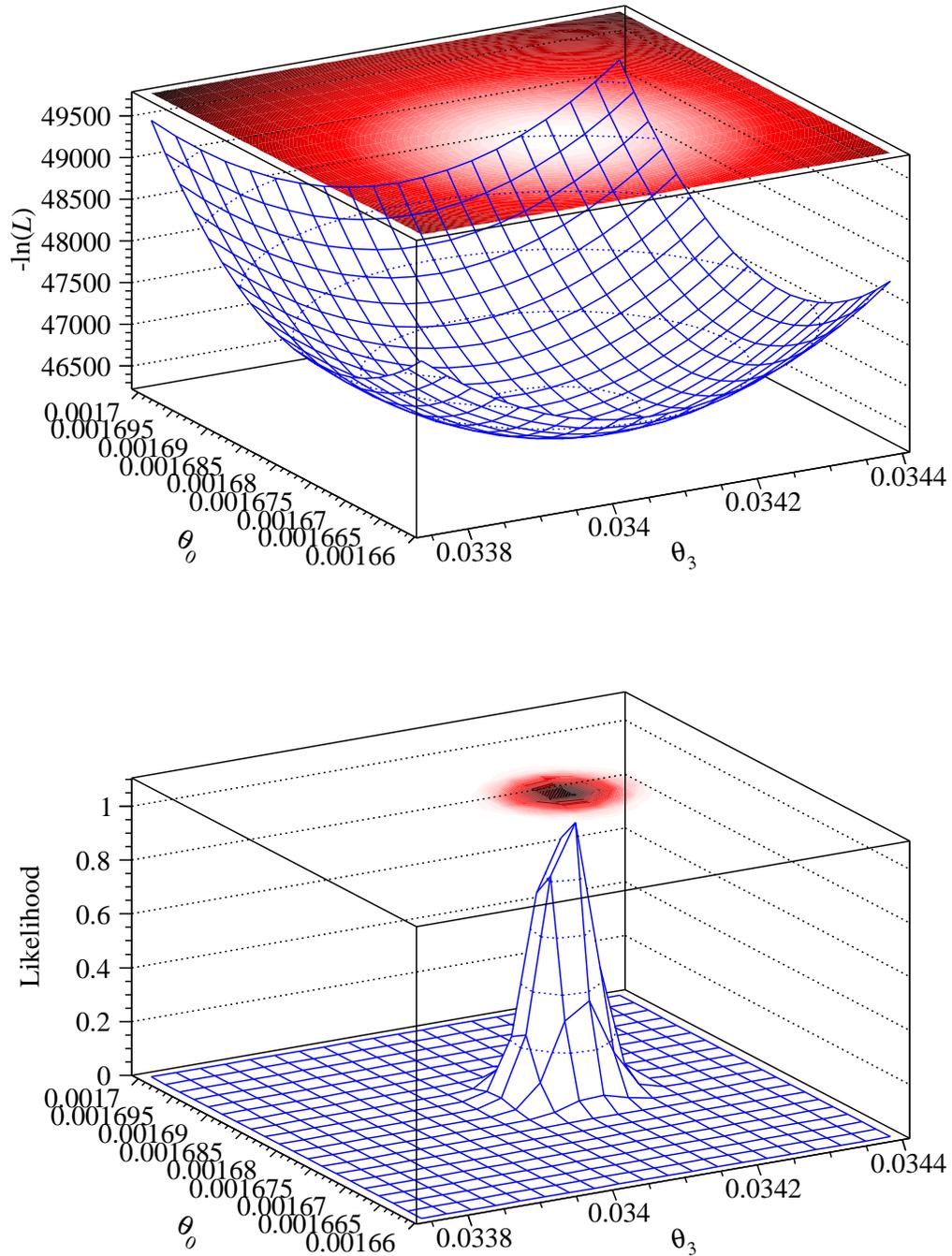
Figure 6.21: *The top panel shows the negative logarithm of the likelihood projected onto the θ_0, θ_3 plane. The bottom panel shows the likelihood itself, projected into the same plane.*
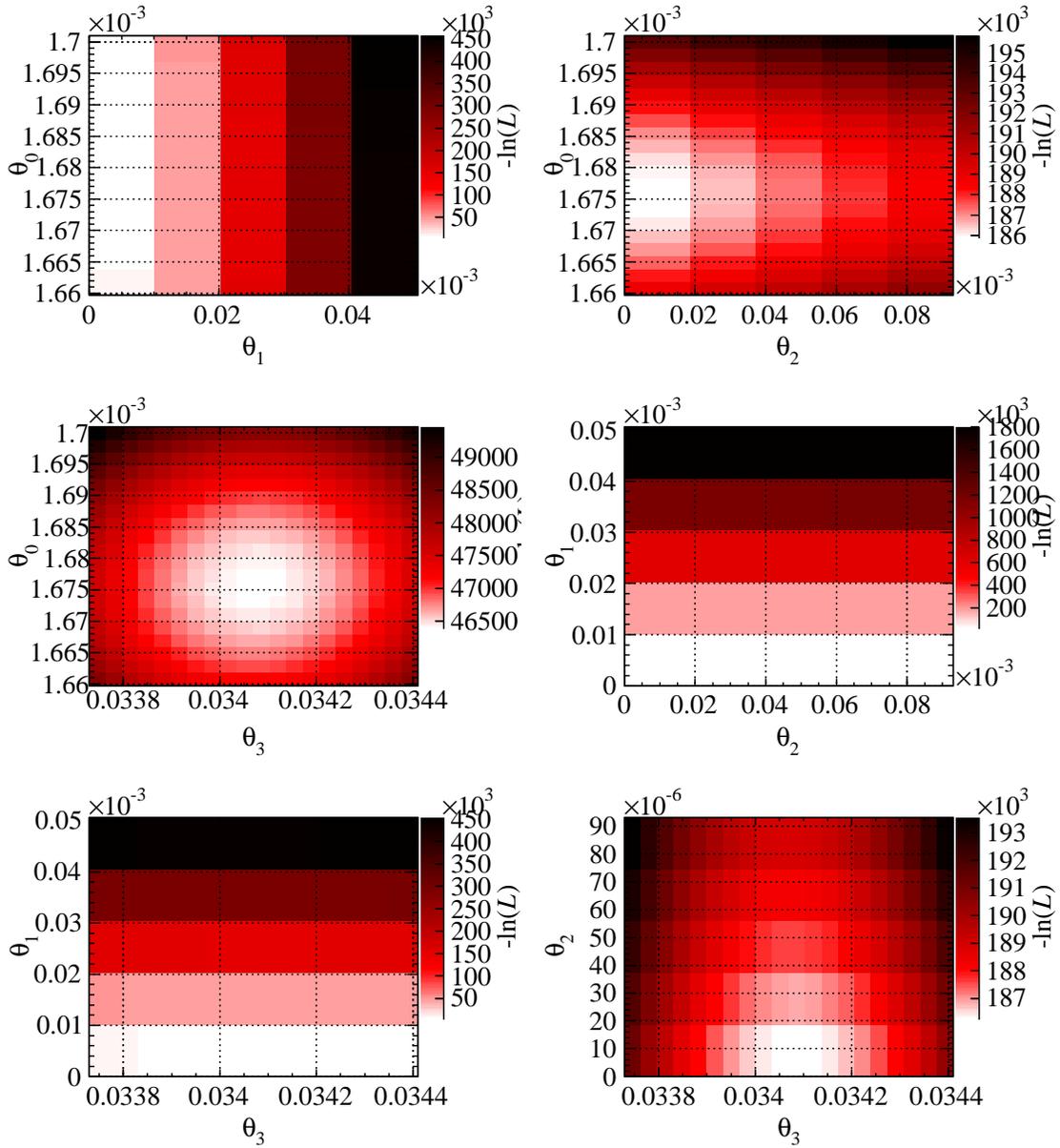
Figure 6.22: *These six panels show all possible 2-dimensional projections of the 4-dimensional negative log-likelihood result.*
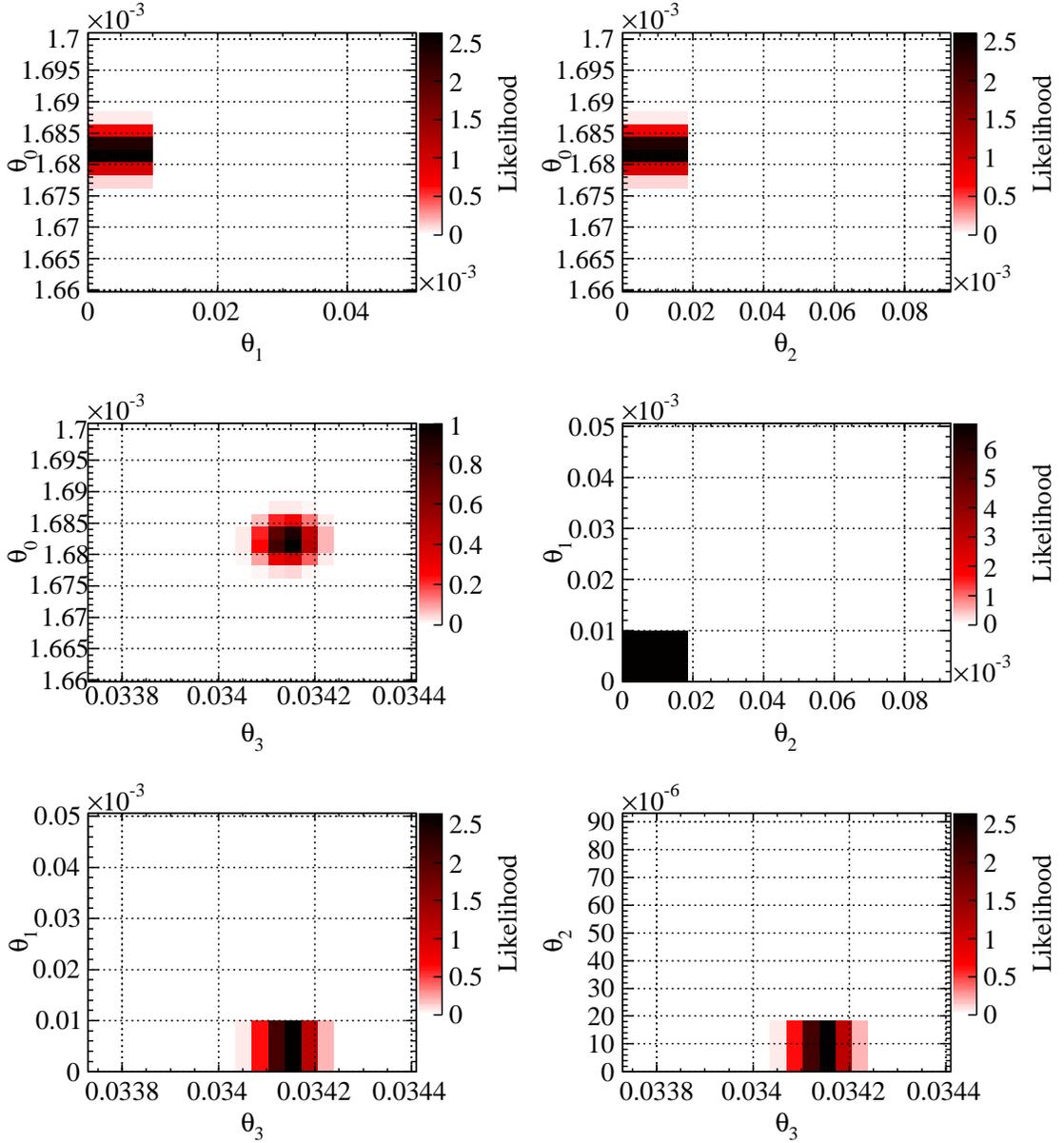
Figure 6.23: *These six panels show all possible 2-dimensional projections of the 4-dimensional likelihood result.*
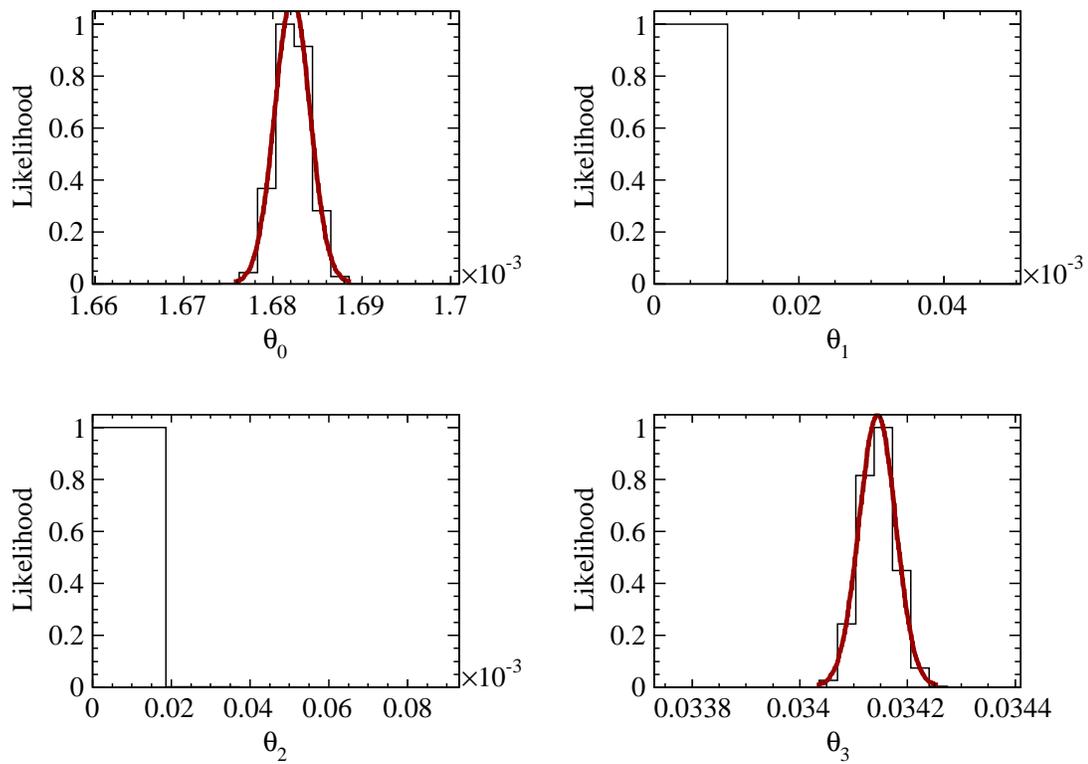
Figure 6.24: *The 4-dimensional likelihood result projected down to one dimension illustrating the most likely values of θ_0 and θ_3. The prompt hadronic reactions have been removed with cuts, so it's no surprise the entire likelihood falls in the bin containing zero for θ_1. At the moment, a PDF with clear discrimination for the decay-in-flight process has not yet been worked out, so θ_2 is ill-determined.*

Maximum Likelihood: Stage 3

The covariance matrix, (6.7), is formed from the matrices (6.16)–(6.18). The elements of these matrices are the summation of event-by-event calculations involving the best estimates, $\hat{\theta}_i$, from Stage 2. The resulting covariance matrix is

$$
\mathbf{V} = \begin{array}{c c} & \begin{array}{cccc} \theta_0 & \theta_1 & \theta_2 & \theta_3 \end{array} \\ \begin{array}{c} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{array} & \left(\begin{array}{cccc} 9.79 \times 10^{-12} & 4.46 \times 10^{-14} & -4.86 \times 10^{-11} & 9.67 \times 10^{-13} \\ 4.46 \times 10^{-14} & 6.27 \times 10^{-14} & -3.55 \times 10^{-12} & -3.40 \times 10^{-13} \\ -4.86 \times 10^{-11} & -3.55 \times 10^{-12} & 5.83 \times 10^{-7} & -9.53 \times 10^{-8} \\ 9.67 \times 10^{-13} & -3.40 \times 10^{-13} & -9.53 \times 10^{-8} & 8.83 \times 10^{-8} \end{array} \right) \end{array}.
$$

Figure 6.25 shows the decrease in $\sigma_{\theta_0} = \sqrt{\mathbf{V}_{00}}$ as the number of analyzed events increases.
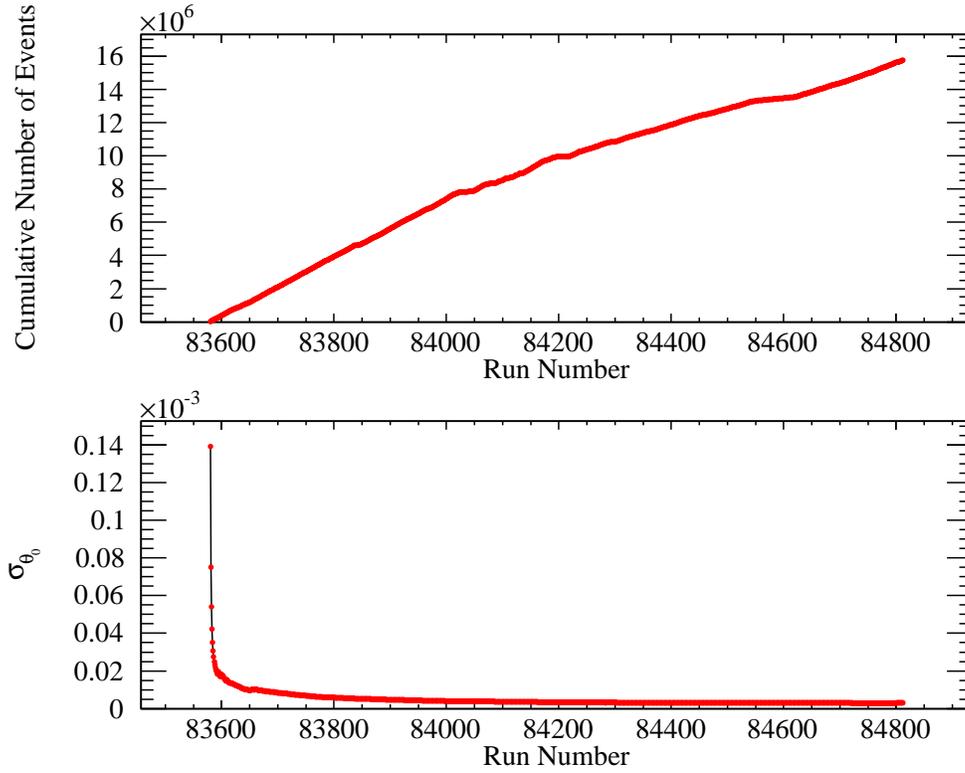
Figure 6.25: *The top panel shows the cumulative number of recorded events (unweighted) that enter the final analysis. The bottom panel shows how the statistical uncertainty in the branching ratio, σ_{θ_0}, decreases as the number of events increases.*

Maximum Likelihood Results

Once all three stages of the likelihood analysis are complete we can quote the most likely value of

the positronic pion decay branching ratio with its estimated uncertainty,

$$\theta_0^{(\text{ML})} = \hat{\theta}_0 \pm \sqrt{\mathbf{V}_{00}} = [1.6822255 \pm 0.0031295(\text{stat.})] \times 10^{-3} \tag{6.50}$$

$$\left(\Delta\theta_0^{(\text{ML})}/\theta_0^{(\text{ML})} = 0.0018603\right) .$$

Now we wish to apply the acceptances from simulation, such that

$$\tilde{\theta}_0^{(\text{ML})} = \frac{a_{\pi_{e2}}N_{\pi_{e2}}}{\left(a_{\pi_{\mu2}}N_{\pi_{\mu2}} + a_{\pi_{e2}}N_{\pi_{e2}}\right)} , \tag{6.51}$$

but we note that θ_0 implicitly contains the values of $N_{\pi_{e2}}$ and $N_{\pi_{\mu2}}$. We must re-write $\tilde{\theta}_0$ in terms of θ_0. Since

$$\frac{N_{\pi_{e2}}}{N_{\pi_{\mu2}}} = \frac{\theta_0}{1 - \theta_0} \quad \text{and} \quad \frac{a_{\pi_{e2}} N_{\pi_{e2}}}{a_{\pi_{\mu2}} N_{\pi_{\mu2}}} = \frac{\tilde{\theta}_0}{1 - \tilde{\theta}_0} \tag{6.52}$$

we can write

$$\frac{\tilde{\theta}_0}{1 - \tilde{\theta}_0} = \frac{a_{\pi_{e2}}}{a_{\pi_{\mu2}}} \frac{\theta_0}{1 - \theta_0} \tag{6.53}$$

thus obtaining

$$\tilde{\theta}_0 = \frac{a_{\pi_{e2}} \theta_0}{\left(a_{\pi_{e2}} - a_{\pi_{\mu2}} \right) \theta_0 + a_{\pi_{\mu2}}} \ . \tag{6.54}$$

To estimate the uncertainty, we use the following

$$\sigma_{\tilde{\theta}_0}^2 = \left| \frac{\partial \tilde{\theta}_0}{\partial \theta_0} \right|^2 \sigma_{\theta_0}^2 = \left| \frac{\partial \tilde{\theta}_0}{\partial \theta_0} \right|^2 \mathbf{V}_{00}$$

$$= \left| \frac{a_{\pi_{e2}} a_{\pi_{\mu2}}}{\left[\left(a_{\pi_{e2}} - a_{\pi_{\mu2}} \right) \theta_0 + a_{\pi_{\mu2}} \right]^2} \right|^2 \mathbf{V}_{00} \ . \tag{6.55}$$

The positronic pion decay branching ratio obtained with the likelihood analysis is

$$\boxed{\begin{array}{c} \tilde{\theta}_0^{(\text{ML})} = [1.1120560 \pm 0.0020650(\text{stat.})] \times 10^{-4} \\[2mm] \left(\Delta\theta_0^{(\text{ML})} / \theta_0^{(\text{ML})} = 0.0018569 \right) \ . \end{array}} \tag{6.56}$$

Chapter 7

$R_{\pi_{e2}}$ and Conclusions

This chapter summarizes the current results of an on-going effort to extract the positronic pion

decay branching ratio from the PEN experiment data. To date, only the experiment as implemented

in 2008 has been realized in a GEANT4 Monte Carlo simulation [6]. Therefore I have restricted my

analysis to the corresponding set of measurement data.

A minimum χ^2 fit resulted in a $\chi^2/N_{\mathrm{dof}} = 1.3$, illustrating good agreement between our measured

data and the probability distribution functions used to represent the processes under consideration.

The intentionally blinded branching ratio determined from the χ^2 fit is

$$\tilde{\theta}_0^{(\chi^2)} = [1.116 \pm 0.002(\mathrm{stat.})] \times 10^{-4} \tag{7.1}$$
$$\left(\Delta\theta_0^{(\chi^2)}/\theta_0^{(\chi^2)} = 0.0022 \right) .$$

An un-binned maximum likelihood analysis code was developed and properly takes into account

the weights of our events when estimating the uncertainty. The maximum likelihood estimate of the

(blinded) positronic pion decay branching ratio is

$$\tilde{\theta}_0^{(\text{ML})} = [1.112 \pm 0.002(\text{stat.})] \times 10^{-4}$$

$$\left(\Delta\theta_0^{(\text{ML})}/\theta_0^{(\text{ML})} = 0.0019 \right) \; . \qquad (7.2)$$

A summary of the results[1] from the two methods is given in Table 7.1. The un-binned maximum

likelihood technique resulted in a 0.19% relative uncertainty (statistical) which is an improvement

over the 0.22% relative uncertainty from the binned minimum χ^2 method. The statistical uncertainty

from either method is already smaller than the combined (statistical and systematic) uncertainty in

the current experimental world average [33]. The inclusion of the 2009 and 2010 data sets in the

near future will significantly reduce the statistical uncertainty of our measurement. The precision

of our measurement, therefore, already meets its design specifications.

The accuracy of $R_{\pi_{e2}}$, however, remains to be determined. Any systematic uncertainty in the

Δt observable was shown to be negligible (Section 6.5.2), but we have yet to quantify the entire

systematic uncertainty in the shape of our PDFs for the total energy. Our uncertainty in the shape of

the low-energy tail for the E_{Total} PDF for the $\pi \to e$ process currently yields a 0.33% relative uncer-

tainty on $R_{\pi_{e2}}$, which will be reduced with a larger set of simulated data and studies of the measured

distributions which are currently underway. Any systematic uncertainties due to the shapes of the

E_{Total} PDFs for the other four processes have not yet been quantified. In addition, both the χ^2 and

maximum likelihood methods currently estimate a negligible fraction of pions decaying in flight,

even though we have made no cuts to specifically remove these events. The inclusion of additional

observables, such as target waveform integrals and the pion decay vertex, should further constrain

---

[1]The central values quoted in this dissertation are still intentionally blinded (multiplied by an unknown random number) but are accurate to within a few percent of our measurement.

Table 7.1: *Summary of the positronic pion decay branching ratios from various analysis methods. Normally the least significant digits would be dropped, but they have been retained to illustrate the level at which the various methods disagree. The central values are still blinded (multiplied by an unknown factor) but are accurate to within a few percent of our measurement.*

Ratio	Value	Relative Uncertainty
$R_{\pi_{e2}}^{\mathrm{PDG}}$	$[1.230 \pm 0.004(\mathrm{comb.})] \times 10^{-4}$	0.0033
$R^{(\chi^2)}$	$[1.6915943 \pm 0.0037338(\mathrm{stat.})] \times 10^{-3}$	0.0022073
$\tilde{R}^{(\chi^2)}$	$[1.1164924 \pm 0.0024644(\mathrm{stat.})] \times 10^{-4}$	0.0022073
$\theta_0^{(\chi^2)}$	$[1.6887376 \pm 0.0037213(\mathrm{stat.})] \times 10^{-3}$	0.0022036
$\tilde{\theta}_0^{(\chi^2)}$	$[1.1163678 \pm 0.0024639(\mathrm{stat.})] \times 10^{-4}$	0.0022071
$\theta_0^{(\mathrm{ML})}$	$[1.6822255 \pm 0.0031295(\mathrm{stat.})] \times 10^{-3}$	0.0018603
$\tilde{\theta}_0^{(\mathrm{ML})}$	$[1.1120560 \pm 0.0020650(\mathrm{stat.})] \times 10^{-4}$	0.0018569

the estimation of this pernicious background process.

List of Tables

List of Figures

Appendix A

Trigger Diagrams

This appendix contains the schematic diagrams for the trigger logic as it was implemented for the PEN experiment in 2010. We implemented several upgrades in 2009 and 2010, including a shift from the software-based FPGA coincidence triggers to using logical AND and OR circuits with NIM hardware. We shaped the target pulse to produce a faster rise time and significantly reduce the tail. This hardware-based shaping of the pulse allowed us to tag muon pulses following the pion pulse in the target. The tagged muon logic signal then allowed us to implement a "muon veto" for a dedicated $\pi \rightarrow e$ low-energy tail trigger. Another improvement is the implementation of a strobe signal in order to force all triggers to have the same trigger timing, thus eliminating the possibility of an associated systematic uncertainty.
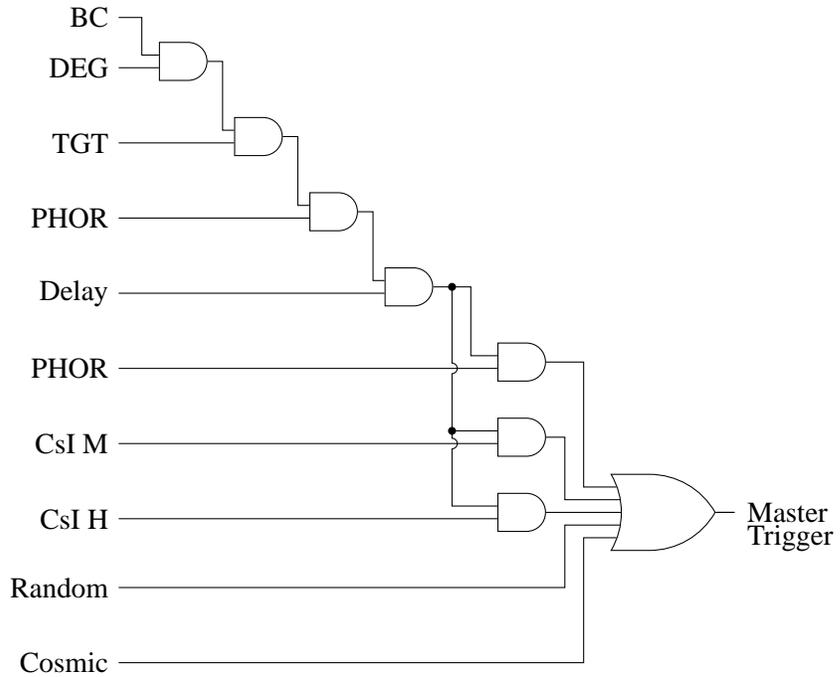
Figure A.1: *Trigger level 1.*
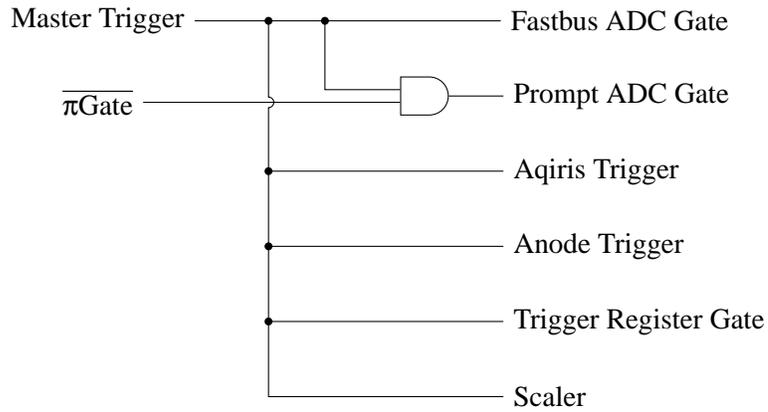
Figure A.2: *Trigger level 2.*

End of π Gate

Delayed RANDOM

Delayed COSMIC

PROMPT Gate

No Connection

D.Trigger

Discr.

TRIG RAW

CAMAC I/O

IN

Veto-IN

"Busy
Latch"

Reset   OUT

TTL-OUT

TRIG ACC

Gate for FASTBUS

Fan
Out

D.Trigger

Veto-IN

TTL-IN

CAMAC I/O

"Reset"

"LAM"

LAM

End-Mark

OUT

running
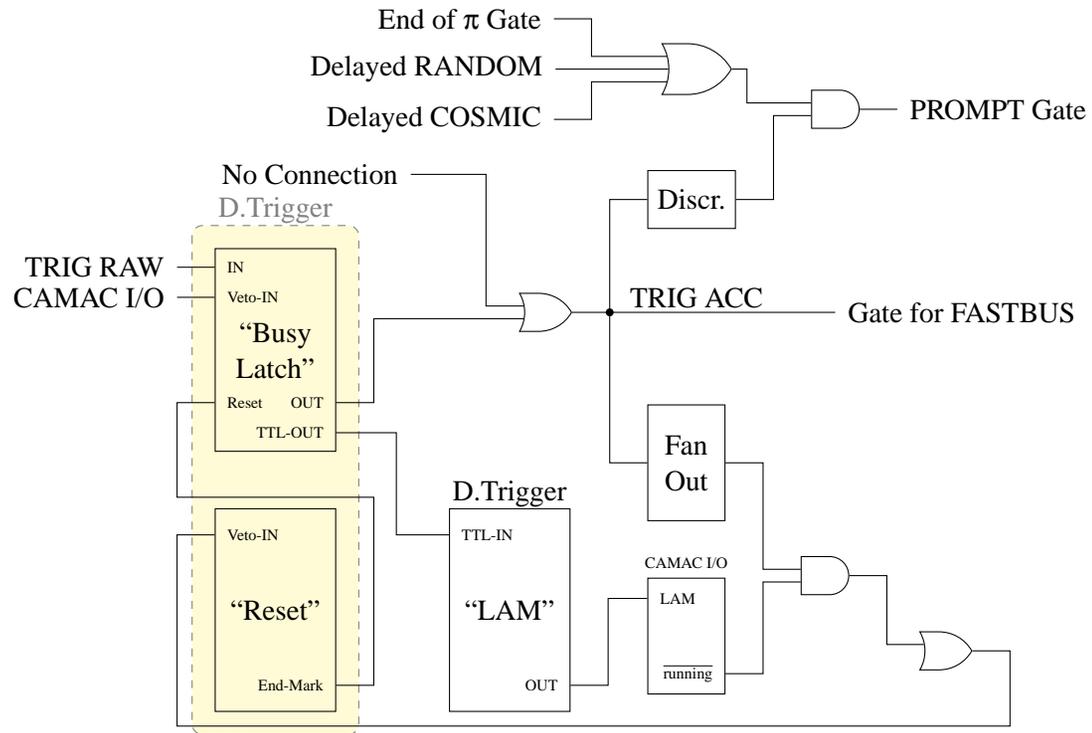
Figure A.3: *Busy-latch / Reset logic.*

Figure A.4: *PH summing.*

Many Fan In Units

CsI Signals — 240 — IN | Capacitor | Fan Out 428F | IN INV | OUT → to TDC/Scalars | OUT → to TR0/3 | OUT → to TR1 | OUT → to TR2 | OUT

Figure A.5: *CsI summing.*

Discr.

Coincidence

BC*DEG

BC — Fan Out — L | M

IN IN Overlap — BC*DEG Overlap (to TR3)

OUT

S.Shaper

DEG — Fan Out — L | M | H

πSTOP

IN IN IN OUT OUT — S.Shaper

VETO

Delay

Fan Out

TGT — Fan Out — L | M | H

πGATE*PHOR

PHOR — S.Shaper — Delay — IN IN IN

S.Shaper

PH$_{up}$ PH$_{dn}$ — Fan In — Disc. — Delay

OUT

πGate Short (to TR3)

PHOR — S.Shaper — Delay — Strobe Gate

Figure A.6: *Bottom level logic and strobe generation.*
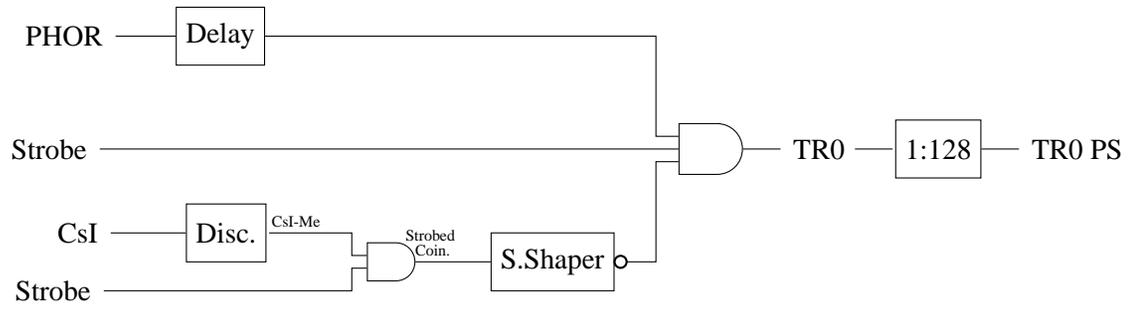
Figure A.7: *Exclusive TR0.*
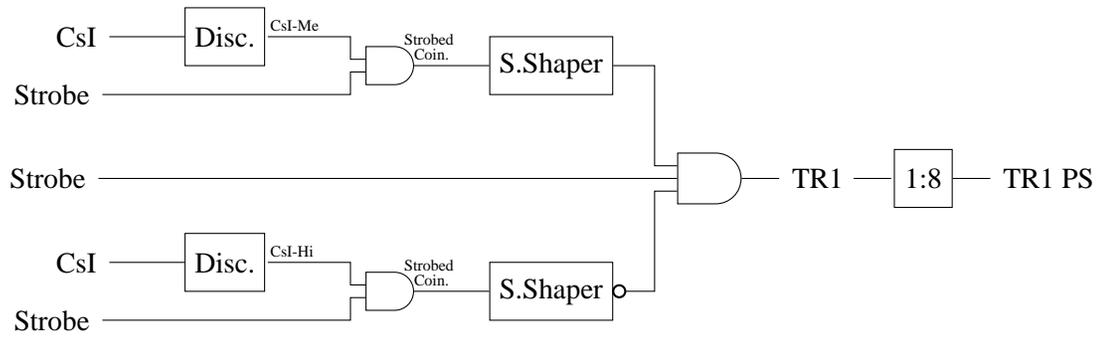
Figure A.8: *Exclusive TR1.*
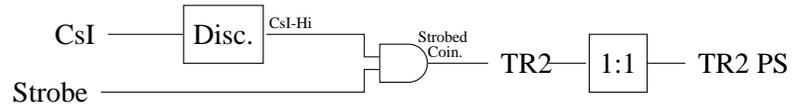
Figure A.9: *Exclusive TR2.*
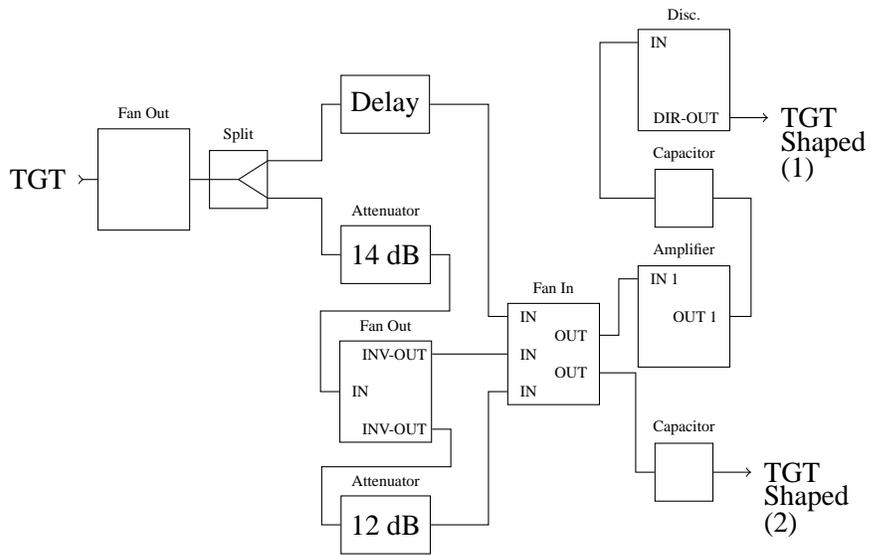
Figure A.10: *TR3 (part 1) — target clip signal.*

Figure A.11: *TR3 (part2) — Muon veto with TR3 triggers.*
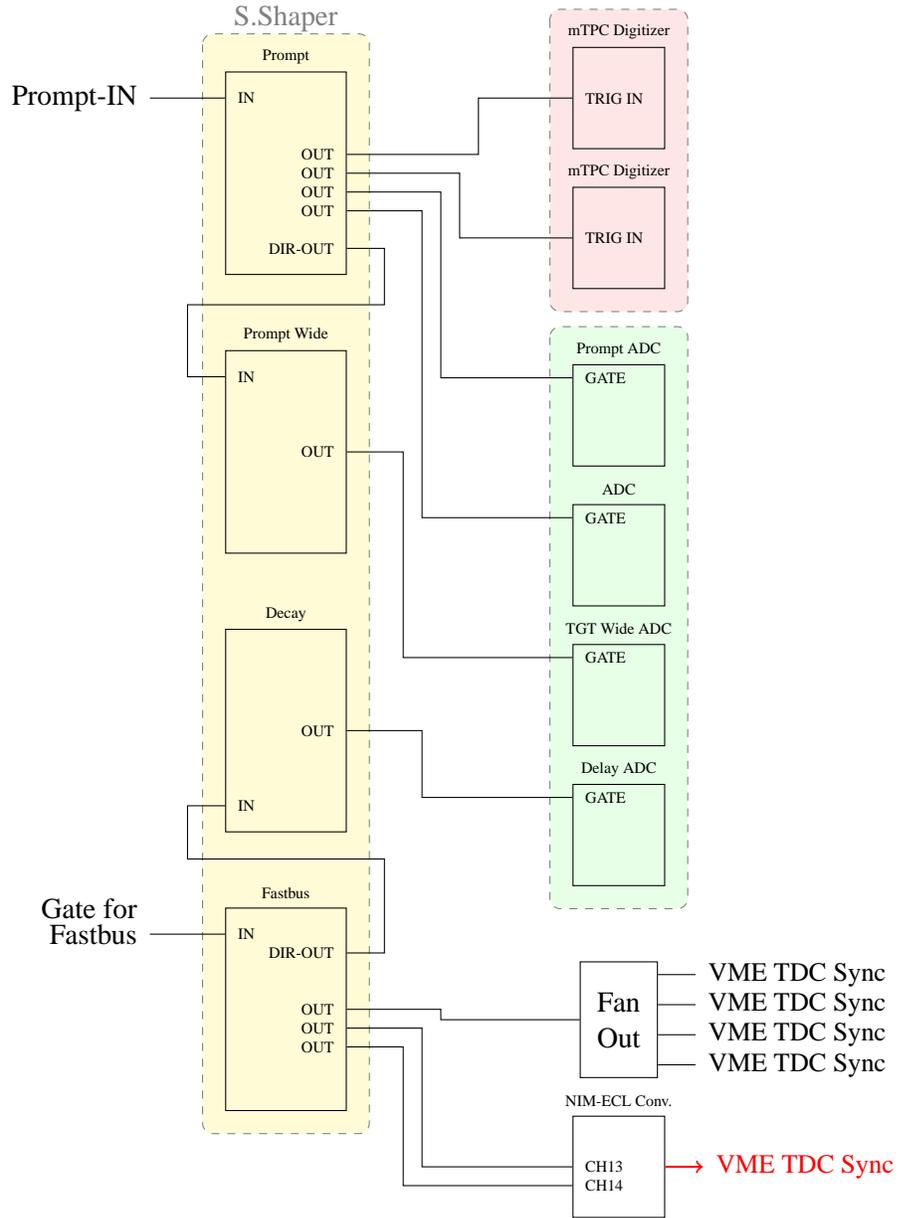
Figure A.12: *ADC gate generation.*
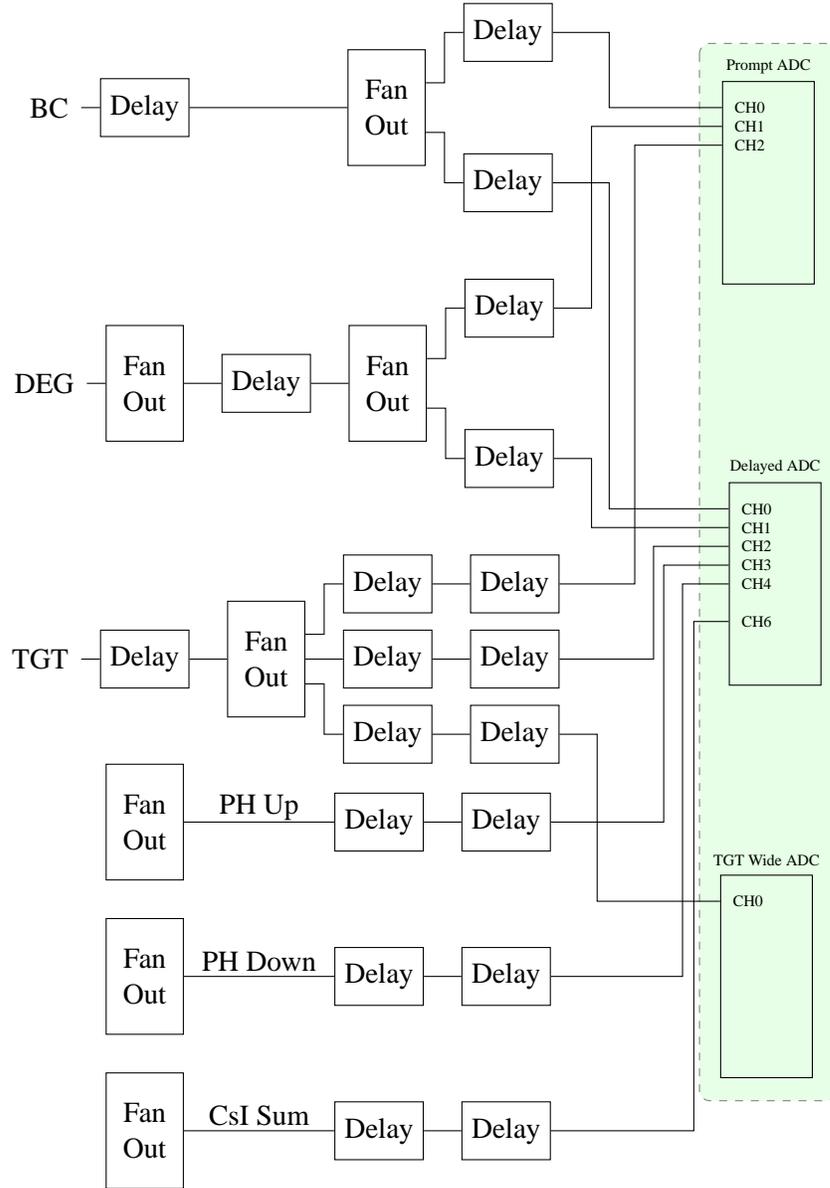
Figure A.13: *BADC logic.*

Appendix B

Data Analysis Flowchart

Data Collection: This block consists of the actual experiment as described in Chapter 3. The PEN experiment had several *frontend* computers connected to electronic data acquisition systems. Each frontend sent its recorded data for each event to a *backend* computer which in turn combined the event data and wrote data files in zipped MIDAS format (`.mid.gz`) [1]. Each data file contains 2×10^5 events (in year 2008) and corresponds to a *run* of roughly 30 minutes of data collection. In 2008 we collected approximately 2.4 TB of `.mid.gz` files in 1233 runs.

Data Generation: A GEANT4 simulation was written to produce data in exactly the same format as our measured data. This Monte Carlo simulation is described in great detail by L. P. Alonzi [6].

Calibration/Stabilization: This block is thoroughly discussed in Chapter 4.

Data Replay: Our data is "replayed" in a modular C++ program called the `analyzer`. It is in this stage where timing coincidences, energy depositions, and geometrical data points are calculated to reconstruct the beam tracks and decay tracks of the particles observed in each event. An attempt is made to identify particle type (π^+, e^+, p^+, etc.). Additionally, fits are performed to each target waveform to identify the decay mode ($\pi \to e$ or $\pi \to \mu \to e$) as discussed in Chapter 5. The

`analyzer` program can be executed in parallel on multiple processors. We typically submit an

`analyzer` job for each run. Every job produces an output file in the form of a ROOT Tree. The total

size of the .root files amounts to 200 GB for year 2008.

Data Reduction: Modern, large-scale experiments usually produce huge amounts of data of

which only a small fraction is useful for a given measurement. Another C++ program called the

`bushmaker` was written to reduce the amount of unnecessary background in our data files. The

program prunes the 200 GB of trees and creates 14 GB of bushes (still in ROOT Tree format). The

`bushmaker` is also run in parallel, with typically one job per data collection run. It is much easier,

and faster, to work with smaller files.

PDF Determination: The `findpdf` C++ program takes the bushes as input and applies cuts

to the data to isolate the various processes. The shape of the PDF for each observable, for each

process, is obtained. This program also applies theoretical descriptions to the PDFs to compensate

for low-statistics regions.

Maximum Likelihood (Stage 1): Here we fill histograms containing only the events that will

enter the likelihood maximization procedure. This calculation can be performed in parallel. Then

in the **Combine Results** stage the histograms are summed and a fit is performed to the time and

energy observables to obtain and initial guess for the range over which to scan the parameters of

interest in the maximization procedure.

Maximum Likelihood (Stage 2): The parameters of interest are scanned in the ranges obtained

in the previous stage. The most likely values of the parameters of interest are determined.

Maximum Likelihood (Stage 3): The correction to the variance-covariance matrix is applied

and a more accurate estimate of the errors is found.
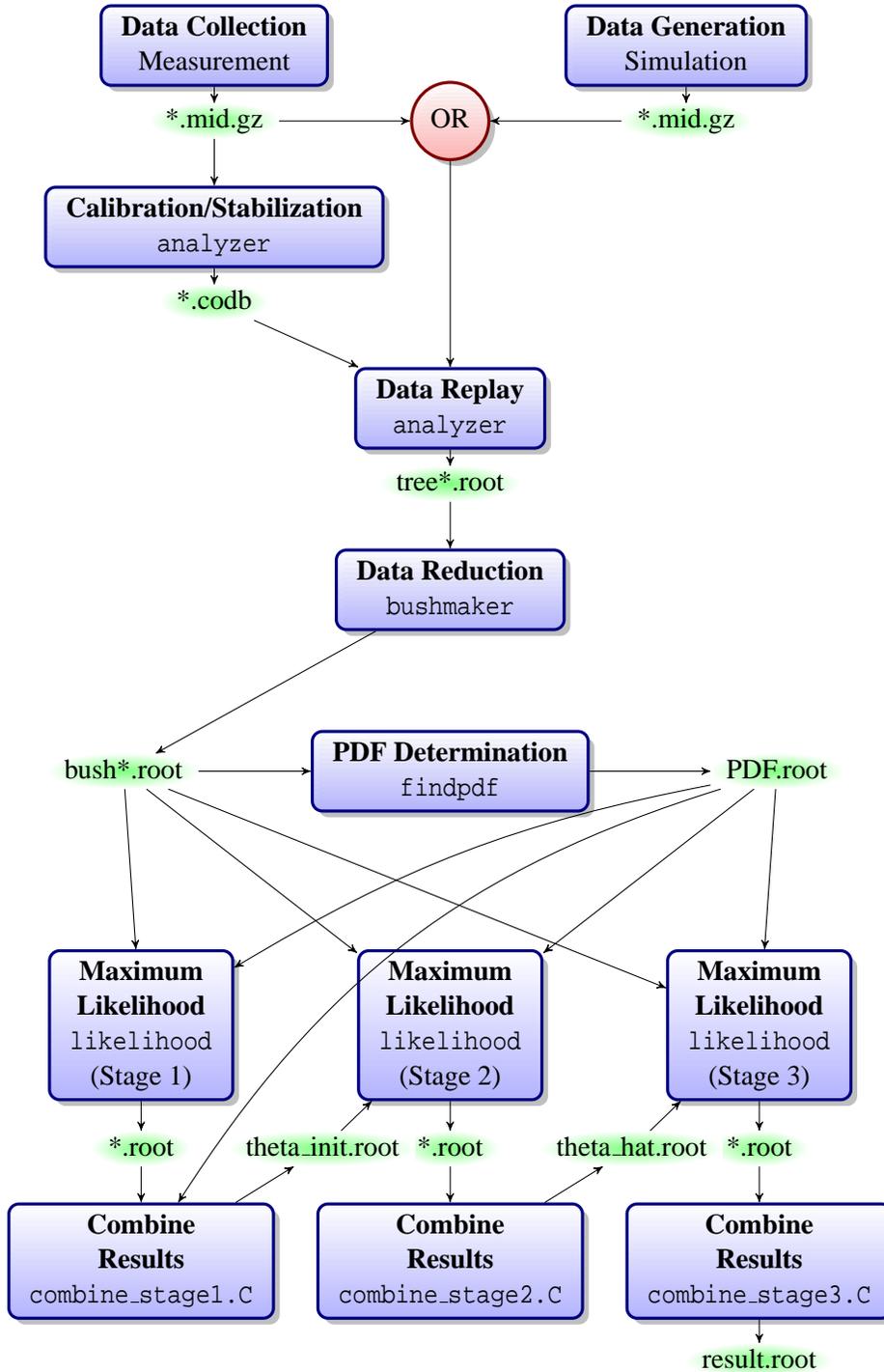
Figure B.1: *Data analysis flowchart*

Appendix C

χ^2 Minimization with Penalty Terms

When parameter constraints are available from elsewhere, such as from another detector element or even a physical constraint, we may use that information in a minimum χ^2 analysis. In our case we have several parameters in our fit that can be predicted with some certainty using multiple detectors.

We can modify the objective function, whose value gets minimized in the fit, by adding a quadratic term for each constrained parameter,

$$\chi^2 = \frac{1}{N_{\text{dof}}} \sum_{i=1}^{n} \left(\frac{\tilde{w}_i^{\text{Fit}} - \tilde{w}_i}{\sigma_{\tilde{w}}} \right)^2 + \left(\frac{E_\pi^{\text{Fit}} - E_\pi^{\text{Pred}}}{\sigma_{E_\pi}} \right)^2 + \left(\frac{E_e^{\text{Fit}} - E_e^{\text{Pred}}}{\sigma_{E_e}} \right)^2 . \tag{C.1}$$

The additional terms are known as penalty terms, since they increase the χ^2 when the fit tries to pull the parameter away from its prediction.

A modified χ^2 objective function that could be implemented in PEN target waveform analysis

is,

$$\chi^2 = \frac{1}{n - N_{\text{dof}}} \sum_{i=1}^{n} \left(\frac{A_i^{\text{Fit}} - A_i^{\text{Pred}}}{\sigma_A} \right)^2$$

$$+ \lambda_1 \left(\frac{A_\pi^{\text{Fit}} - A_\pi^{\text{Pred}}}{\sigma_{A_\pi}} \right)^2 \tag{C.2}$$

$$+ \lambda_2 \left(\frac{\sqrt{L_e^{\text{TGT}}} \log \left(\frac{A_e^{\text{Fit}}}{A_e^{\text{Pred}}} \right)}{\sigma_{A_e}} \right)^2.$$

The first term is the well known χ^2 per degree of freedom, where n is the number of analyzed waveform bins. The number of *extra* degrees of freedom, N_{dof}, is equal to 2 for a 2-peak fit, and 3 for a 3-peak fit. Only ± 20 waveform bins from each peak position are analyzed, and the number of analyzed bins is kept constant. Therefore, for events in which the pulses overlap, additional bins are included in the fit so as to fit the same number of bins for every event. σ_A gives a measure of the resolution of the digitizer. The second term takes into consideration how well the π^+ amplitude from the fit, $A_{\pi^+}^{\text{Fit}}$, agrees with our predicted π^+ amplitude, $A_{\pi^+}^{\text{Pred}}$. The uncertainty in our prediction is given by $\sigma_{A_{\pi^+}}$, and is roughly 716 keV (see Section 5.2.1). Finally, the last term in the expression increases the χ^2 as the fit positron amplitude gets further from the predicted positron amplitude. The ratio of the fit amplitude to the predicted amplitude resembles a smeared Vavilov distribution. The logarithm transforms this distribution into one symmetric about zero improving the convergence speed in the minimization process. This method is discussed in more detail in Section 5.2.1.

We included weights λ_1 and λ_2 in the additional terms. Setting the weights too small has the same effect as if the terms aren't included. If the weights are too large we are essentially fixing the parameters to the predicted values. We used a simulated waveform with a muon pulse completely overlapping with a pion pulse ($\Delta t = 0$) to determine the appropriate weight for the χ^2 contribution from the second term. We performed fits for many values of the weight, λ_1. We did a similar study

with overlapping muon and positron pulses to determine λ_2. We chose the weight values at the intersection point of the contribution from the first term and the additional terms so that we had an approximately equal contribution to the χ^2 from each term, as shown in Figure C.1 ($\lambda_1 = 0.9$, $\lambda_2 = 0.048$). There was no theoretical basis for that particular choice of weights, and we found out much later using the full Monte Carlo simulation [6] that the largest possible weights (essentially fixing the parameters) gave the best discrimination between pion decay modes.

(a)   (b)
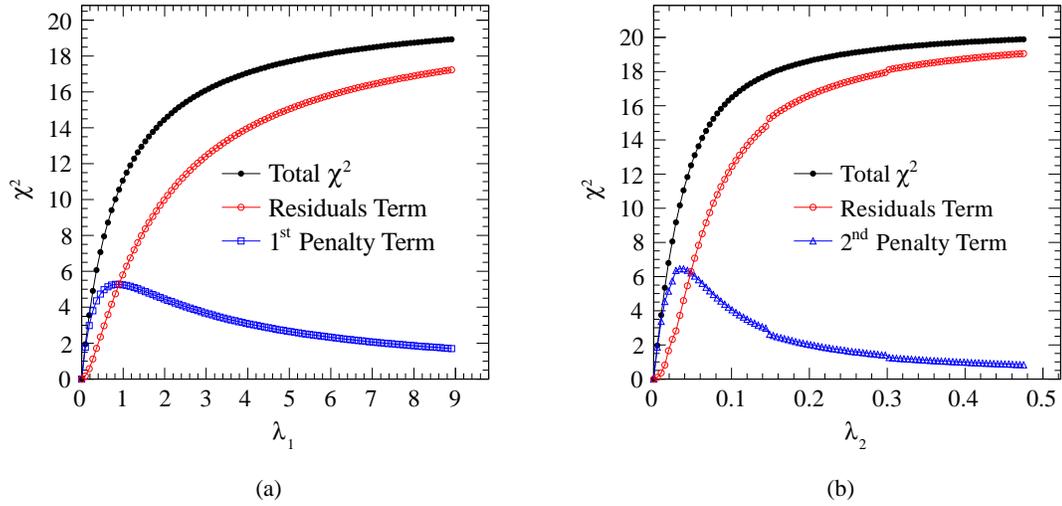
Figure C.1: *χ^2 as a function of the weights λ_1 (a) and λ_2 (b). Black is the total χ^2, red is the contribution to the χ^2 from the first term, blue is the contribution from the second term, and green shows the contribution from the third term.*

Nevertheless, to further test the fitting routine before the full Monte Carlo simulation was ready, we created a simple simulated waveform with a pion pulse at $t_\pi = 0$ ns and a positron pulse at $t_e = 25$ ns. This two-peak waveform was then fit with both the 2-peak and 3-peak hypotheses. For the 3-peak hypothesis fits, a fit was done for each predicted time of the muon pulse, t_μ^{Pred} in the range $-10 < t_\mu < 35$ ns with a 1 ns increment. These fits correspond to the bottom two panels of figure C.2.

Simulated 3-peak waveforms were created with $t_\pi = 0$ ns, $t_e = 25$ ns, and a muon pulse at 1 ns

increments between $-10 < t_\mu < 35$ ns. For these fits, the t_μ^{Pred} was set to the actual time of the pulse

in the synthetic waveform. These fits correspond to the top two panels of figure C.2.

3-Peak fit with μ^+ in WF

2-Peak fit without μ^+ in waveform
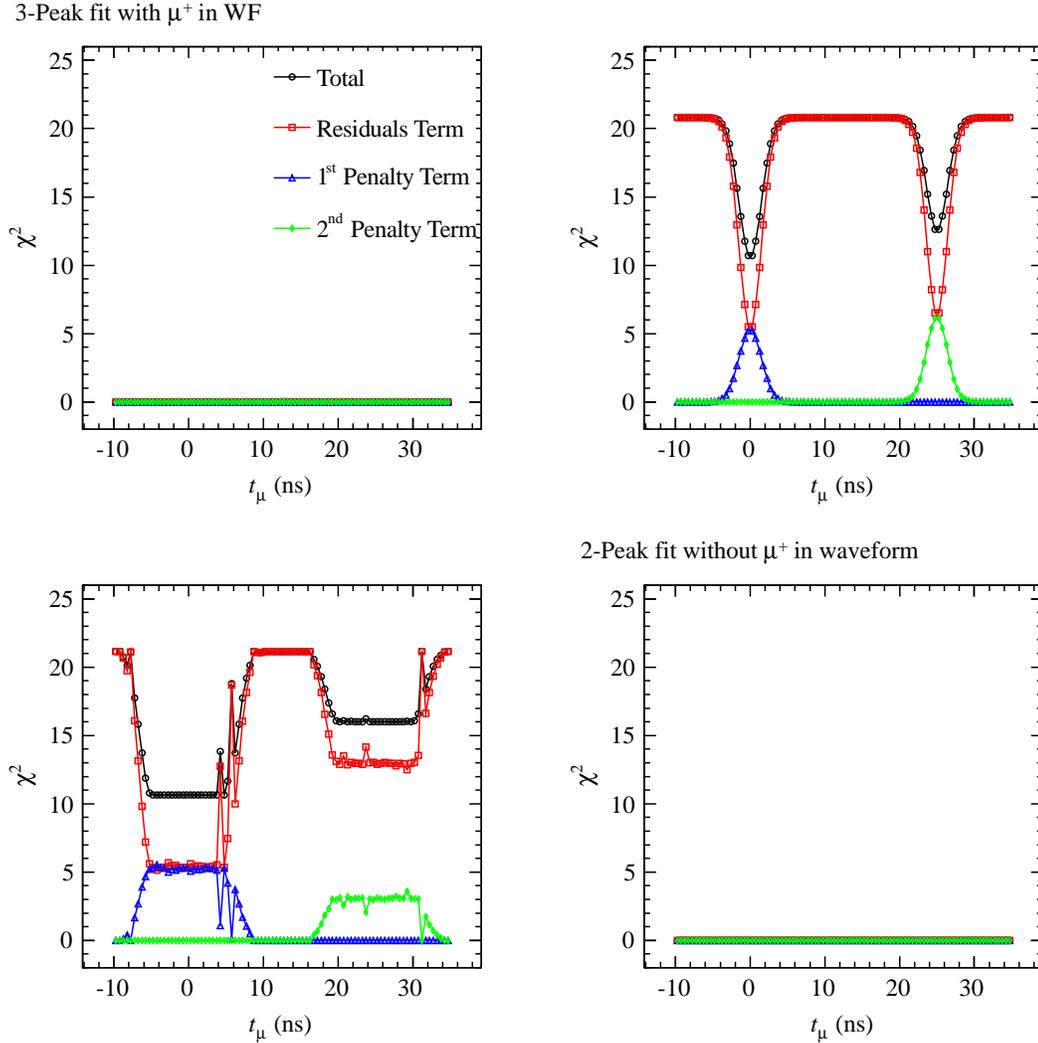
Figure C.2: χ^2 *from fits to simulated waveforms for various event types demonstrating the variation of the modified* χ^2 *as a function of pulse separation. Plotted here is not the traditional* χ^2, *but rather* $\chi^2 - 1$ *such that a good fit corresponds to zero.*

One can display the same information more compactly when using $\Delta\chi^2 = \chi^2_{\text{3-peak}} - \chi^2_{\text{2-peak}}$ as in
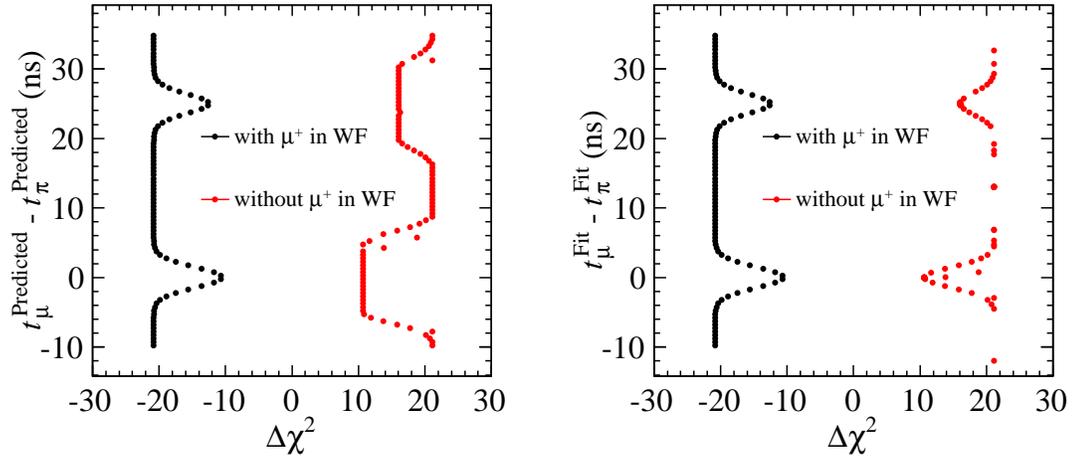
figure C.3.

Figure C.3: *Scatter plots showing the difference between the predicted (left) and fit (right) pion and muon pulse times versus $\Delta\chi^2$, demonstrating the ability of $\Delta\chi^2$ to discriminate between 2-peak ($\pi \rightarrow e$) and 3-peak ($\pi \rightarrow \mu \rightarrow e$) events even when the pulses overlap.*

Appendix D

Partial Derivatives of the Likelihood Function

$$\frac{\partial \mathcal{L}}{\partial \theta_0} = -\left(1-\theta_1\right)\left(1-\theta_2\right)\left(1-\theta_3\right)P_{\pi_{\mu 2}}\left(\vec{x}_e\right)$$

$$+\left(1-\theta_1\right)\left(1-\theta_2\right)\left(1-\theta_3\right)P_{\pi_{e 2}}\left(\vec{x}_e\right) \tag{D.1}$$

$$\frac{\partial \mathcal{L}}{\partial \theta_1} = -\left(1-\theta_0\right)\left(1-\theta_2\right)\left(1-\theta_3\right)P_{\pi_{\mu 2}}\left(\vec{x}_e\right)$$

$$-\theta_0\left(1-\theta_2\right)\left(1-\theta_3\right)P_{\pi_{e 2}}\left(\vec{x}_e\right)$$

$$+\left(1-\theta_2\right)\left(1-\theta_3\right)P_{\text{Had}}\left(\vec{x}_e\right) \tag{D.2}$$

$$\frac{\partial \mathcal{L}}{\partial \theta_2} = -\left(1-\theta_0\right)\left(1-\theta_1\right)\left(1-\theta_3\right)P_{\pi_{\mu 2}}\left(\vec{x}_e\right)$$

$$-\theta_0\left(1-\theta_1\right)\left(1-\theta_3\right)P_{\pi_{e 2}}\left(\vec{x}_e\right)$$

$$-\theta_1\left(1-\theta_3\right)P_{\mathrm{Had}}\left(\vec{x}_e\right)$$

$$+\left(1-\theta_3\right)P_{\mathrm{DIF}}\left(\vec{x}_e\right) \tag{D.3}$$

$$\frac{\partial \mathcal{L}}{\partial \theta_3} = -\left(1-\theta_0\right)\left(1-\theta_1\right)\left(1-\theta_2\right)P_{\pi_{\mu 2}}\left(\vec{x}_e\right)$$

$$-\theta_0\left(1-\theta_1\right)\left(1-\theta_2\right)P_{\pi_{e 2}}\left(\vec{x}_e\right)$$

$$-\theta_1\left(1-\theta_2\right)P_{\mathrm{Had}}\left(\vec{x}_e\right)$$

$$-\theta_2 P_{\mathrm{DIF}}\left(\vec{x}_e\right)$$

$$+P_{\mathrm{Acc}}\left(\vec{x}_e\right) \tag{D.4}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_0 \partial \theta_0} = \frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_1} = \frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_2} = \frac{\partial^2 \mathcal{L}}{\partial \theta_3 \partial \theta_3} = 0 \tag{D.5}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_0 \partial \theta_1} = \frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_0} = \left(1-\theta_2\right)\left(1-\theta_3\right)P_{\pi_{\mu 2}}\left(\vec{x}_e\right)$$

$$-\left(1-\theta_2\right)\left(1-\theta_3\right)P_{\pi_{e 2}}\left(\vec{x}_e\right) \tag{D.6}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_0 \partial \theta_2} = \frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_0} = (1 - \theta_1)(1 - \theta_3) P_{\pi_{\mu 2}}(\vec{x}_e)$$

$$- (1 - \theta_1)(1 - \theta_3) P_{\pi_{e 2}}(\vec{x}_e) \tag{D.7}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_0 \partial \theta_3} = \frac{\partial^2 \mathcal{L}}{\partial \theta_3 \partial \theta_0} = (1 - \theta_1)(1 - \theta_2) P_{\pi_{\mu 2}}(\vec{x}_e)$$

$$- (1 - \theta_1)(1 - \theta_2) P_{\pi_{e 2}}(\vec{x}_e) \tag{D.8}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_2} = \frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_1} = (1 - \theta_0)(1 - \theta_3) P_{\pi_{\mu 2}}(\vec{x}_e)$$

$$+ \theta_0 (1 - \theta_3) P_{\pi_{e 2}}(\vec{x}_e)$$

$$- (1 - \theta_3) P_{\text{Had}}(\vec{x}_e) \tag{D.9}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \theta_3} = \frac{\partial^2 \mathcal{L}}{\partial \theta_3 \partial \theta_1} = (1 - \theta_0)(1 - \theta_2) P_{\pi_{\mu 2}}(\vec{x}_e)$$

$$+ \theta_0 (1 - \theta_2) P_{\pi_{e 2}}(\vec{x}_e)$$

$$- (1 - \theta_2) P_{\text{Had}}(\vec{x}_e) \tag{D.10}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \theta_3} = \frac{\partial^2 \mathcal{L}}{\partial \theta_3 \partial \theta_2} = (1 - \theta_0)(1 - \theta_1)P_{\pi_{\mu 2}}(\vec{x}_e)$$

$$+\theta_0(1 - \theta_1)P_{\pi_{e 2}}(\vec{x}_e)$$

$$+\theta_1 P_{\text{Had}}(\vec{x}_e)$$

$$-P_{\text{DIF}}(\vec{x}_e) \tag{D.11}$$

Appendix E

Inclusion of Truncated Time PDF

It is possible to truncate the $\pi \to \mu to$e sequential-decay time PDF in order to obtain the fraction of events we would expect to see if we were to measure $t \to \infty$. That means the normalization of the $\pi \to \mu to$e sequential-decay time PDF is 1.0 when taking $t \to \infty$ but < 1.0 in our measured time window. The likelihood must be properly normalized

$$\mathcal{L} \to \mathcal{L}/\mathcal{N} \tag{E.1}$$

where,

$$\mathcal{N} = \sum_{i=1}^{M} f_i \int_{\vec{x}} P_i(\vec{x}) d\vec{x} \ . \tag{E.2}$$

Replacing $\mathcal{L}$ with $\mathcal{L}/\mathcal{N}$ in (1) will give us $\ell_{\mathcal{N}}$. Differentiating as in (6.4) yeilds

$$
\begin{aligned}
\mathbf{V}_{jk}^{-1} &= \mathrm{E}\left[\frac{\partial^2\left(-\ell_{\mathcal{N}}\right)}{\partial f_j \partial f_k}\right] \\
&= \mathrm{E}\left[-\sum_{e=1}^{N}\frac{\partial}{\partial f_j}\left(\frac{\partial}{\partial f_k}\ln\left(\frac{\mathcal{L}}{\mathcal{N}}\right)\right)\right] \\
&= \left.\sum_{e=1}^{N}\frac{P_j(\vec{x}_e)P_k(\vec{x}_e)}{\left[\sum_{i=1}^{M}f_i P_i(\vec{x}_e)\right]^2}\right|_{f=\hat{f}} - \left.\sum_{e=1}^{N}\frac{\int_{\vec{x}}P_j(\vec{x})d\vec{x}\int_{\vec{x}}P_k(\vec{x})d\vec{x}}{\left[\sum_{i=1}^{M}f_i\int_{\vec{x}}P_i(\vec{x})d\vec{x}\right]^2}\right|_{f=\hat{f}}
\end{aligned} \tag{E.3}
$$

Following the same prescription as in (6.7) we will need,

$$
\mathbf{H}_{jk} = \left.\sum_{e=1}^{N}w_e\frac{P_j(\vec{x}_e)P_k(\vec{x}_e)}{\left[\sum_{i=1}^{M}f_i P_i(\vec{x}_e)\right]^2}\right|_{f=\hat{f}} - \left.\sum_{e=1}^{N}w_e\frac{\int_{\vec{x}}P_j(\vec{x})d\vec{x}\int_{\vec{x}}P_k(\vec{x})d\vec{x}}{\left[\sum_{i=1}^{M}f_i\int_{\vec{x}}P_i(\vec{x})d\vec{x}\right]^2}\right|_{f=\hat{f}} \tag{E.4}
$$

$$
\mathbf{H'}_{jk} = \left.\sum_{e=1}^{N}w_e^2\frac{P_j(\vec{x}_e)P_k(\vec{x}_e)}{\left[\sum_{i=1}^{M}f_i P_i(\vec{x}_e)\right]^2}\right|_{f=\hat{f}} - \left.\sum_{e=1}^{N}w_e^2\frac{\int_{\vec{x}}P_j(\vec{x})d\vec{x}\int_{\vec{x}}P_k(\vec{x})d\vec{x}}{\left[\sum_{i=1}^{M}f_i\int_{\vec{x}}P_i(\vec{x})d\vec{x}\right]^2}\right|_{f=\hat{f}} \tag{E.5}
$$

again estimated in the sample mean.

Bibliography

[1] Midas data acquisition system. `http://midas.psi.ch`.

[2] Paul Scherrer Institut website. `http://www.psi.ch`.

[3] Radiant Mirror Film, VM2000. Technical report, 1990.

[4] *Handbook of Signal Processing in Acoustics(2 vol set)*. Springer, 2009.

[5] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J.J. Gmez Cadenas, I. Gonzlez, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F.W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampn, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita,

K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Oka-mura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M.G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J.P. Wellisch, T. Wenaus, D.C. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche. Geant4 — a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[6] L. P. Alonzi. *Measuring the Pion Substructure with Radiative Positronic Pion Decays*. PhD thesis, University of Virginia, 2012.

[7] H.L. Anderson, T. Fujii, and R.H. Miller. Branching Ratio of the Electronic Mode of Positive Pion Decay. *Phys.Rev.*, 119:2050, 1960.

[8] S. M. Berman. Radiative corrections to pion beta decay. *Phys. Rev. Lett.*, 1:468, 1958.

[9] D. I. Britton, S. Ahmad, D. A. Bryman, R. A. Burnham, E. T. H. Clifford, P. Kitching, Y. Kuno, J. A. Macdonald, T. Numao, A. Olin, J. M. Poutissou, and M. S. Dixit. Improved search for massive neutrinos in $\pi^+ \to e^+ \nu$ decay. *Phys. Rev. D*, 46:R885–R887, 1992.

[10] D. I. Britton, S. Ahmad, D. A. Bryman, R. A. Burnham, E. T. H. Clifford, P. Kitching, Y. Kuno, J. A. Macdonald, T. Numao, A. Olin, J-M. Poutissou, and M. S. Dixit. Measurement of the $\pi^+ \to e^+_\nu$ branching ratio. *Phys. Rev. Lett.*, 68(20):3000–3003, 1992.

[11] D.I. Britton, S. Ahmad, D.A. Bryman, R.A. Burnbam, E.T.H. Clifford, et al. Measurement of the $\pi \to$ e neutrino branching ratio. *Phys.Rev.Lett.*, 68:3000–3003, 1992.

[12] D. A. Bryman, P. Depommier, and C. Leroy. $\pi \to e\nu$, $\pi \to e\nu\gamma$ decays and related processes. *Physics Reports*, 88(3):151 – 205, 1982.

[13] D. A. Bryman, R. Dubois, T. Numao, B. Olaniyi, A. Olin, M. S. Dixit, J. M. Poutissou, and J. A. Macdonald. Search for massive neutrinos in the decay $\pi \to e\nu$. *Phys. Rev. Lett.*, 50:1546–1549, 1983.

[14] D. A. Bryman and T. Numao. Search for massive neutrinos in $\pi^+ \to \mu^+\nu$ decay. *Phys. Rev. D*, 53:558–559, 1996.

[15] D.A. Bryman, R. Dubois, T. Numao, B. Olaniyi, A. Olin, et al. New Measurement of the $\pi \to$ e neutrino Branching Ratio. *Phys.Rev.Lett.*, 50:7, 1983.

[16] Douglas Bryman, William J. Marciano, Robert Tschirhart, and Taku Yamanaka. Rare Kaon and Pion Decays: Incisive Probes for New Physics Beyond the Standard Model. *Annual Review of Nuclear and Particle Science*, 61:331–354, 2011.

[17] Douglas A. Bryman. $\pi \to e\nu$ decay: Window on the generation puzzle. *Comments Nucl. Part. Phys.*, 21:101–122, 1993.

[18] M. Bychkov, D. Počanić, B. A. VanDevender, V. A. Baranov, W. Bertl, Yu. M. Bystritsky, E. Frlež, V. A. Kalinnikov, N. V. Khomutov, A. S. Korenchenko, S. M. Korenchenko, M. Korolija, T. Kozlowski, N. P. Kravchuk, N. A. Kuchinsky, W. Li, D. Mekterović, D. Mzhavia, S. Ritt, P. Robmann, O. A. Rondon-Aramayo, A. M. Rozhdestvensky, T. Sakhelashvili, S. Scheu, U. Straumann, I. Supek, Z. Tsamalaidze, A. van der Schaaf, E. P. Velicheva, V. P.

Volnykh, Y. Wang, and H.-P. Wirtz. New precise measurement of the pion weak form factors in $\pi^+ \to e^+ \nu \gamma$ decay. *Phys. Rev. Lett.*, 103:051802, 2009.

[19] Bruce A Campbell and Ahmed Ismail. Leptonic pion decay and physics beyond the electroweak standard model. page 31, 2008.

[20] Bruce A. Campbell and David W. Maybury. Constraints on scalar couplings from $\pi^{\pm} \to \ell^{\pm} \nu_{\ell}$. *Nuclear Physics B*, 709(1-2):419 – 439, 2005.

[21] Saint-Gobain Ceramics and Plastics Inc. Scintillation Products. `http://www.detectors.saint-gobain.com/Plastic-Scintillator.aspx`.

[22] Vincenzo Cirigliano and Ignasi Rosell. Two-loop effective theory analysis of $\pi(k) \to e \overline{\nu}_e [\gamma]$ branching ratios. *Phys. Rev. Lett.*, 99:231801, 2007.

[23] PEN Collaboration. *PSI R-05-01 Experiment Proposal*. Paul Scherrer Institute, Villigen, 2006.

[24] M. Conversi, E. Pancini, and O. Piccioni. On the disintegration of negative mesons. *Phys. Rev.*, 71(3):209–210, 1947.

[25] G. Czapek, A. Federspiel, A. Fluckiger, D. Frei, B. Hahn, et al. Branching ratio for the rare pion decay into positron and neutrino. *Phys.Rev.Lett.*, 70:17–20, 1993.

[26] Sacha Davidson, David C. Bailey, and Bruce A. Campbell. Model independent constraints on leptoquarks from rare processes. *Z. Phys.*, C61:613–644, 1994.

[27] Michel Davier. Lepton universality. `http://www.slac.stanford.edu/cgi-wrap/getdoc/ssi97-002.pdf`, 1997.

[28] E. Di Capua, R. Garland, L. Pondrom, and A. Strelzoff. Study of the Decay $\pi \rightarrow e\nu$. *Phys.Rev.*, 133:B1333–B1340, 1964.

[29] Hamamatsu Electron Tube Division. H2431-50 Datasheet. `http://sales.hamamatsu.com/en/products/electron-tube-division/detectors/photomultiplier-modules/part-h2431-50.php`.

[30] W.T. Eadie. *Statistical methods in experimental physics*. North-Holland Pub. Co., 1971.

[31] E. Frlež et. al. Design, Commissioning and Performance of the PIBETA Detector at PSI. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 526(3):300–347, 2004.

[32] T. Fazzini *et al.* Electron Decay of the Pion. *Phys. Rev. Lett.*, pages 247–249, 1958.

[33] K. Nakamura *et al.* (Particle Data Group). Review of Particle Physics. *J. Phys. G*, 37:075021+, 2010.

[34] Edward Farhi and Leonard Susskind. Technicolor. *Phys. Rept.*, 74:277, 1981.

[35] E. Fermi, E. Teller, and V. Weisskopf. The decay of negative mesotrons in matter. *Phys. Rev.*, 71(5):314–315, 1947.

[36] Markus Finkemeier. Radiative corrections to pi(l2) and K(l2) decays. *Phys. Lett.*, B387:391–394, 1996.

[37] R. Fletcher. A New Approach to Variable Metric Algorithms. *Comput. J.*, 13:p. 317, 1970.

[38] J. E. Gaiser. *Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime: Appendix F*. PhD thesis, SLAC, 1982.

[39] H. Georgi and S. L. Glashow. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974.

[40] G. Impeduglia, R. Plano, A. Prodell, N. Samios, M. Schwartz, and J. Steinberger. β decay of the pion. *Phys. Rev. Lett.*, 1:249–251, 1958.

[41] F. James and M. Winkler. MINUIT. `http://www.cern.ch/minuit`, 2004. CERN.

[42] Toichiro Kinoshita. Radiative corrections to π-e decay. *Phys. Rev. Lett.*, 2:477–480, 1959.

[43] C.M.G. Lattes, H. Muirhead, G.P.S. Occhialini, and C.F. Powell. Processes involving charged mesons. *Nature*, 159:694–697, 1947.

[44] C.M.G. Lattes, G.P.S. Occhialini, and C.F. Powell. Observations on the tracks of slow mesons in photographic emulsions. *Nature*, 160:453–456, 1947.

[45] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. Springer, Berlin, 1994.

[46] Will Loinaz, Naotoshi Okamura, Saifuddin Rayyan, Tatsu Takeuchi, and L. C. R. Wijewardhana. Nutev anomaly, lepton universality, and nonuniversal neutrino-gauge couplings. *Phys. Rev. D*, 70(11):113004, 2004.

[47] Louis Lyons, Wade W.M. Allison, and Jaime Paella Comellas. Maximum likelihood or extended maximum likelihood? an example from high energy physics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 245(2-3):530 – 534, 1986.

[48] William J. Marciano and A. Sirlin. Radiative corrections to π_{l2} decays. *Phys. Rev. Lett.*, 71(22):3629–3632, 1993.

[49] R.E. Marshak and C.P. Ryan. *Theory of weak interactions in particle physics*. Interscience monographs and texts in physics and astronomy. Wiley-Interscience, 1969.

[50] L. Michel. Coupling Properties of Nucleons, Mesons, and Leptons. *Progress in Cosmic Ray Physics*, 1:127–190, 1952.

[51] Tyler J. Moeller. Fpga's for radar front-end dsp. Master's thesis, Massachusetts Institute of Technology, 1999.

[52] K Nakamura and Particle Data Group. Review of particle physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, 2010.

[53] Seth H. Neddermeyer and Carl D. Anderson. Note on the nature of cosmic-ray particles. *Phys. Rev.*, 51(10):884–886, 1937.

[54] Seth H. Neddermeyer and Carl D. Anderson. Cosmic-ray particles of intermediate mass. *Phys. Rev.*, 54(1):88–89, 1938.

[55] Y. Nishina, M. Takeuchi, and T. Ichimiya. On the nature of cosmic-ray particles. *Phys. Rev.*, 52(11):1198–1199, 1937.

[56] Y. Nishina, M. Takeuchi, and T. Ichimiya. On the mass of the mesotron. *Phys. Rev.*, 55(6):585–586, 1939.

[57] Acqiris (now Agilent Technologies). Acqiris High-Speed cPCI Digitizers. Technical report.

[58] G.P.S. Occhialini and C.F. Powell. Nuclear disintegrations produced by slow charged particles of small mass. *Nature*, 159:186–190, 1947.

[59] M. J. Oreglia. *A Study of the Reactions psi prime → gamma gamma psi: Appendix D*. PhD thesis, SLAC, 1980.

[60] A. Palladino, A. van der Schaaf, and D. Počanić. Reconstructing detector waveforms with overlapping pulses. To be published.

[61] Jogesh C. Pati and Abdus Salam. Lepton number as the fourth "color". *Phys. Rev. D*, 10:275–289, Jul 1974.

[62] D. Počanić, E. Frlež, V. A. Baranov, W. Bertl, Ch. Brönnimann, M. Bychkov, J. F. Crawford, M. Daum, N. V. Khomutov, A. S. Korenchenko, S. M. Korenchenko, T. Kozlowski, N. P. Kravchuk, N. A. Kuchinsky, W. Li, R. C. Minehart, D. Mzhavia, B. G. Ritchie, S. Ritt, A. M. Rozhdestvensky, V. V. Sidorkin, L. C. Smith, I. Supek, Z. Tsamalaidze, B. A. VanDevender, Y. Wang, H.-P. Wirtz, and K. O. H. Ziock. Precise measurement of the $\pi^+ \to \pi^0 e^+ \nu$ branching ratio. *Phys. Rev. Lett.*, 93:181803, 2004.

[63] Michael J. Ramsey-Musolf, Shufang Su, and Sean Tulin. Pion leptonic decays and supersymmetry. *Phys. Rev. D*, 76:095017, 2007.

[64] Graham G. Ross. *Grand Unified Theories*. Frontiers in Physics, 60. Benjamin/Cummings, 1984.

[65] F. Scherbaum. *Of Poles and Zeros: Fundamentals of Digital Seismology (Modern Approaches in Geophysics)*. Springer, 1996.

[66] O. Shanker. $\pi\ell 2$, $k\ell 3$ and $K^0 - \bar{k}^0$ constraints on leptoquarks. *Nuclear Physics B*, 1982.

[67] Robert E. Shrock. General theory of weak processes involving neutrinos. i. leptonic pseudoscalar-meson decays, with associated tests for, and bounds on, neutrino masses and lepton mixing. *Phys. Rev. D*, 24:1232–1274, 1981.

[68] T. Skwarnicki. *Ph.D. Thesis, DESY F31-86-02: Appendix E.* PhD thesis, DESY, 1986.

[69] J. C. Street and E. C. Stevenson. New evidence for the existence of a particle of mass intermediate between the proton and electron. *Phys. Rev.*, 52(9):1003–1004, 1937.

[70] J. H. Winters et al. Ghost cancellation of analog tv signals. *IEEE Trans. on Circuits and Systems for Video Technology*, 1(1):136, 1991.

[71] H. Yukawa. On the Interaction of Elementary Particles. I. *Proc. Physico-Math. Soc. Japan*, 17:48–57, 1935.

[72] G. P. Zeller, K. S. McFarland, T. Adams, A. Alton, S. Avvakumov, L. de Barbaro, P. de Barbaro, R. H. Bernstein, A. Bodek, T. Bolton, J. Brau, D. Buchholz, H. Budd, L. Bugel, J. Conrad, R. B. Drucker, B. T. Fleming, R. Frey, J. A. Formaggio, J. Goldman, M. Goncharov, D. A. Harris, R. A. Johnson, J. H. Kim, S. Koutsoliotas, M. J. Lamm, W. Marsh, D. Mason, J. McDonald, C. McNulty, D. Naples, P. Nienaber, A. Romosan, W. K. Sakumoto, H. Schellman, M. H. Shaevitz, P. Spentzouris, E. G. Stern, N. Suwonjandee, M. Tzanov, M. Vakili, A. Vaitaitis, U. K. Yang, J. Yu, and E. D. Zimmerman. Precise determination of electroweak parameters in neutrino-nucleon scattering. *Phys. Rev. Lett.*, 88(9):091802, 2002.

[73] G. P. Zeller, K. S. McFarland, T. Adams, A. Alton, S. Avvakumov, L. de Barbaro, P. de Barbaro, R. H. Bernstein, A. Bodek, T. Bolton, J. Brau, D. Buchholz, H. Budd, L. Bugel, J. Conrad, R. B. Drucker, B. T. Fleming, R. Frey, J. A. Formaggio, J. Goldman, M. Goncharov,

D. A. Harris, R. A. Johnson, J. H. Kim, S. Koutsoliotas, M. J. Lamm, W. Marsh, D. Mason, J. McDonald, C. McNulty, D. Naples, P. Nienaber, A. Romosan, W. K. Sakumoto, H. Schellman, M. H. Shaevitz, P. Spentzouris, E. G. Stern, N. Suwonjandee, M. Tzanov, M. Vakili, A. Vaitaitis, U. K. Yang, J. Yu, and E. D. Zimmerman. Effect of asymmetric strange seas and isospin-violating parton distribution functions on $\sin^2 \theta_W$ measured in the nutev experiment. *Phys. Rev. D*, 65(11):111103, 2002.