A Search for the Decay of Protons to $e^+\pi^0$ and $\mu^+\pi^0$

A Dissertation Presented

by

Brett Michael Viren

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Physics

State University of New York
at
Stony Brook

May 2000

State University of New York
at Stony Brook

The Graduate School

Brett Michael Viren

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of the dissertation.

Dr. Chang Kee Jung
Advisor
Professor of Physics

Dr. Luis Orozco
Professor of Physics

Dr. Robert Shrock
Professor of Physics

Dr. Maurice Goldhaber
Professor of Physics

This dissertation is accepted by the Graduate School.

Graduate School

Abstract of the Dissertation

A Search for the Decay of Protons to $e^+\pi^0$ and $\mu^+\pi^0$

by

Brett Michael Viren

Doctor of Philosophy

in

Physics

State University of New York at Stony Brook

2000

A search for proton decay to $e^+\pi^0$ and $\mu^+\pi^0$ was done with 784.9 days of data from the Super–Kamiokande detector. Background of less than 0.25 events was expected in both modes and no decay candidates were found. Partial lifetime limits of $\tau(p \to e^+\pi^0) > 2.59 \times 10^{33}$ years and $\tau(p \to \mu^+\pi^0) > 2.07 \times 10^{33}$ years are found at 90% confidence. These limits are a factor of 4.8 and 4.4 higher, respectively, than the best limits [20] set in previous experiments.

Dedicated to
my life's important women,
Pam and Yoko, thanks.

Contents

List of Figures

List of Tables

Acknowledgements

I owe a debt of gratitude to many people for helping me reach this point in my life and any I miss is only because my mind is feeble.

First, I would like thank to my advisor Chang Kee Jung. Without him, it is possible that I would not have started graduate school, certainly at least not when and where I did. I am grateful that he gave me a chance to work on a terrific experiment, hiring me almost site-unseen.

I am also happy to have Maurice Goldhaber, Luis Orozco and Robert Shrock, for all of whom I have a lot of respect. In particular, I am proud to have Maurice on this committee as he is one of the pioneers in searching for proton decay.

In this day and age essentially no experiment is done in isolation. This is especially true with large experiments like Super–Kamiokande. The experiment was ably led by Yoji Totsuka, the US group by Hank Sobel and Jim Stone and the ATMPD/HighE group by Takaaki Kajita and Ed Kearns and it has been a privilege to work with them all. All groups contributed to the experiment in one form or another, but one group deserves many accolades and that is the hosting group, ICRR from University of Tokyo.

I have also enjoyed my time in Japan working with collaborators from the other institutions. I thank my fellow students, Tomasz Barszczak for many interesting and geeky discussions, Rob Sanford, Erik Blaufuss and Matt Earl for their companionship and levity, and John Flanagan, Mark Messier and Julien Hsu for their contribution to the software used in part in this thesis. Extra thanks go to Masato Shiozawa for the enjoyable partnership when producing the first proton decay publication. Thanks to Hans-Gerd Berns for his expertise and good humor, Wojciech Gajewski and Danka Kielczewska for their patient explanations and work on the Monte Carlo I used, Bill Kropp for his able and wise leadership during construction (and the parties which followed), Mark Vagins for his unending humor and advice, Alec Habig and Kate Scholberg for their help and useful discussions and finally John Learned for showing that life is not all hard work (and for not cutting off my fingers with the coconut saw).

A special expression of gratitude belongs to the members of the Stony Brook SK/K2K group. Besides Chang Kee, I owe a large debt to the other professor of this group. Chiaki Yanagisawa helped numerous times in aspects from day to day life in Japan to understanding some point of physics.

Since professors tend to be busy, much professing fell on our three postdocs. I would like to thank Jim Hill for patiently helping me understand numerous problems as well as being a good companion. Similarly, Kai Martens and I have shared many interesting discussions. As a reader of this dissertation will notice, there are many attributions to Clark McGrew for many routines used in this analysis. While I have been challenged in my association with Clark throughout these many many years, there is no one who has so greatly and directly helped me with my work.

I was lucky (cursed?) to be the first graduate student in Stony Brook's SK/K2K group but I was even luckier to work with those students who joined after me. I would like to thank Christopher Mauger for his friendship, interesting philosophical discussions and translating all those many Japanese menus. Although in hindsight it may have been a bad idea to have two people who are so into computers in the same group, I want to thank Eric Sharkey for his clear mind, interesting geek sessions and all the help he gave me when I just couldn't come up with the best way to implement something. Matthew Malek joined our group recently and our paths seem to keep us on opposite sides of the Pacific, but the times we have overlapped he has been a source of companionship and good humor.

I have also had the pleasure to work with a steady stream of LUGs (lowly undergraduates) in our group. There have been many, and in particular I want to thank Bjorn Einarsson for a lot fun, hard work and selling me his car, Rachel Reese, Jimmy Andrello and Yuko Nakazawa for their help, and more recently Toseef Raza and Tokufumi Kato who has honored me by turning some of my analysis into a senior thesis.

I have been glad to work and play with some great people at Stony Brook. Angelika Osanna and Barry Winn started the same year as I and we are all ending the same as well. I thank them for all the help, fun and companionship and feel that I am a better person to know them. I would also like to thank Pam Burris (more on her later), Pat Peiliker, Francine Shultz, Joan Napolitano and the rest of the people who actually run the Physics department. Because of them, I knew that I was safely insulated from the twisty bureaucratic world that lies outside.

A more personal debt of gratitude goes to my parents. Pam Burris, Juan Gallardo, Mike Viren and Anne Strozier have provided me with the support and love I needed to get here. It is difficult to list all that I owe them. Besides providing half my DNA, I give thanks to Mike for helping me through college and to Anne for convincing him it was a good idea. Even though I only see them occasionally I feel like I always have a second home. It was from Mike that I learned to love playing and working with computers and his influence is a large reason I chose to study in a scientific field. I want to thank Juan for many many things. His influence is largely why I chose physics (that and because I failed most other subjects). His clear intellect and depth of knowledge is one that I seek to attain. Next, and most importantly I wish to thank Pam for giving and sustaining my life. With out her sacrifices, guidance, support and teachings I know that I would not be who and where I am now. She has taught me things which unfortunately many never learn and because of her I enjoy a deeper understanding of the people and relationships that make up my life. There is little way for me to ever repay her.

Finally, I must give my thanks to the one person who is largely responsible for the joy in my life, my wife Yoko Koshizawa Viren. When she came into my life it changed. Things that seemed crucial became trivial and I found a happiness I didn't think possible. I look forward to sharing the rest of my days with her.

Chapter 1

Introduction

This chapter starts with a basic review of concepts relating to decays in general, followed by motivation for searching for proton decay with description of various aspects of the many theories which predict it. Next, a brief explanation of why these two particular modes are studied is given. Finally, a selected history of proton decay searches is presented and the Super–Kamiokande collaboration is described.

1.1 Decaying Concepts

1.1.1 Stability

Any particle (composite or elementary) can change from one type to one or more different types unless there is something to forbid the transition, in which case the particle is considered stable. For example, a μ^- can decay into a set of three particles: a ν_μ a $\bar{\nu}_e$ and an e^-. On the other hand, an electron does not decay because electric charge must be conserved and there is no particle or collection of particles which are lighter and have an electric charge of e.

A particle can be stable depending on the local environment. For example a free neutron can decay to a proton, an e^- and an $\bar{\nu}_e$. This decay is allowed because the total energy, electric charge and all other things which must be conserved are the same before and after the decay. However, when a neutron is in a stable atomic nucleus the system as a whole will be put into a higher energy state if the neutron decays leaving a proton. The neutron could decay if it released more energy by decaying to lighter products but this avenue for decay is assumed to be disallowed for the same reasons that proton decay is assumed to be disallowed (see section 1.2.1). Searching for the decay of a proton or a neutron in a stable atomic nucleus (collectively called *nucleon decay*) are both ways to search for evidence that this assumption is wrong.

1.1.2 Definition of Decay Parameters

A particle which can decay does so after some time has passed. This time is not a single fixed value but rather varies with each decaying particle of a specific type. The value that characterizes this variance is the called the decay rate Γ and it gives the probability per unit time that a particle will decay. If there are $N(t)$ particles of a certain type at time t then the number dN which will decay in the next moment dt is,

$$dN = -N(t)\Gamma dt. \tag{1.1}$$

Solving this differential equation gives the predicted number of particles at time t as,

$$N(t) = N(0)e^{-\Gamma t}. \tag{1.2}$$

A particular type of particle then has a characteristic *lifetime* defined as $\tau = 1/\Gamma$.

Most types of particles can decay to more than one set of products. A π^0, for example, will almost always decay to two photons, but about 1.2% of the time it will decay to $e^+e^-\gamma$ and even more rarely to other products. A possible set of decay products is called decay branch or decay mode. The total decay rate is then just the sum of the individual decay rates for each mode. It is convenient to define a *partial lifetime* which is the inverse of the decay rate for the particular mode. Herein, the partial lifetime is written, for the example of a proton decaying to a e^+ and a π^0, as $\tau(p \to e^+\pi^0)$.

1.1.3 Lifetime Limits

When a type of particle decays at an observable rate then the lifetime is found by recording the time at which each particle from an observed group of particles decays. The decay rate and thus the lifetime can then be found with equation 1.2. If the decay is rare enough then it may not be measurable in this manner, in which case one must settle for finding a *limit* on the partial lifetime of the particle. This limit is a probabilistic statement about how long the lifetime must be (at some certainty) in order explain the non-observation of any decays, or an observation consistent with all events being due to background.

The measurements of a lifetime or a limit on a lifetime depends on the following quantities: the number of *candidate* events seen which are apparently due to a decay, the ability to observe the decay (the *efficiency*), the number of potentially decaying particles and the length of their observation (collectively the *exposure*) and the expected number of events not due to a decay but which can mimic the signature for the decay (the *background*). When the number of candidates is significantly above the expected background then a lifetime can be fit, if not a limit is set. This search is in the latter situation and the details of how the partial lifetime limits are set are given in chapter 5.

1.2 Motivation for Proton Decay Searches

1.2.1 The Standard Model ...

The Standard Model of particle physics postulates a set of fundamental particles, forces and rules governing their interactions. The two types of matter particles are called *leptons* and *quarks*. Table 1.1 shows their electric charges and masses. For each particle in this table there is a corresponding anti-particle which has the same mass and opposite electric charge. The masses and electric charges of the interaction carriers which mediate the coupling between leptons and quarks are summarized in Table 1.2. Leptons couple to the weak force and except for the neutrinos, the electromagnetic force as well. The quarks couple to both these forces as well as to the strong force. These three interactions are briefly described.

Leptons		
Flavor	Q	Mass
ν_e (electron neutrino)	0	< 15 eV/c^2
e (electron)	-1	0.511 MeV/c^2
ν_μ (muon neutrino)	0	< 0.19 MeV/c^2
μ (muon)	-1	105.7 MeV/c^2
ν_τ (tau neutrino)	0	< 18 MeV/c^2
τ (tau)	-1	1.78 GeV/c^2

Quarks		
Flavor	Q	Mass
u (up)	+2/3	~3 MeV/c^2
d (down)	-1/3	~6 MeV/c^2
c (charm)	+2/3	~1.3 GeV/c^2
s (strange)	-1/3	~100 MeV/c^2
t (top)	+2/3	~175 GeV/c^2
b (bottom)	-1/3	~4.3 GeV/c^2

Table 1.1: Lepton and Quark summary

The electromagnetic interaction can occur between all electrically charged particles. The force is carried by the massless and chargeless *photon* and has a coupling strength given by the fine structure constant $\alpha \approx \frac{1}{137}$. Since the photon is massless the range of the electromagnetic force is unlimited.

Force	Particle	Mass	Charge
Electro-magnetic	γ (photon)	0	0
Weak	W$^-$	80.4 GeV/c^2	-1
	W$^+$	80.4 GeV/c^2	+1
	Z^0	91.2 GeV/c^2	0
Strong	8 × g (gluon)	0	0

Table 1.2: Force carrier summary

The weak interaction can occur between all matter particles via three carriers: the electrically charged $W^{\pm}$ and the neutral Z^0. The strength of the weak force is given by $\alpha_W = \alpha / \sin^2\theta_W$ with $\sin^2\theta_W = 0.23$ for $W^{\pm}$ meditated interactions and by $\alpha_Z = \alpha/(\sin^2\theta_W \cos^2\theta_W)$ for Z^0 mediated interactions. Although the "strength" of the weak force is larger than the electromagnetic force, it appears weaker because it is effectively reduced by a factor of the square of the mass of the mediator. Since the $W^{\pm}$ have a mass of 80.4 GeV/c^2 and Z^0 has a mass of 91.2 GeV/c^2 the apparent strength of the weak force is indeed weak.

The electromagnetic and weak interaction can be combined into a single *electroweak* interaction. This was first suggested by S. Glashow and developed further by S. Weinberg and A. Salam [1, 2]. This GWS theory splits particles into left and right chiral states (corresponding to left and right handed helicity in the case of massless particles). Neutrinos are assumed massless and only left handed neutrinos are allowed. This theory has 1 gauge boson from a $U(1)_Y$ (Y is weak hypercharge) symmetry and 3 from an $SU(2)_L$ (L is left hand chirality) symmetry. These mix to give the photon, Z^0 and $W^{\pm}$. The mixing is governed by the weak (or Weinberg) mixing angle θ_W which also allows the Z and W coupling constant to be related to the electromagnetic coupling constant.

The strong interaction can occur between quarks via three strong charges, or "colors", given the names *red*, *green* and *blue*. The carriers of the strong force are called *gluons* and carry a color and an anti-color. Because the gluons have color, they can interact with themselves. There are 9 possible combination of a color and anti-color and these are arranged into a colored octet and a colorless singlet. Because of color confinement the colorless singlet should be observable. Since it has not been, the strong force can only have 8 carriers and is organized into an SU$(3)_c$ (c is color) group.

Any matter particle can couple to any other via an available force given that the coupling is not disallowed by a conservation law. For example, a coupling must conserve energy, momentum, angular momentum, charge or color. Some forces can break conservation laws which must hold for others. In weak interactions the flavor of the quarks may not be conserved, while strong and electromagnetic interactions do not change the flavor of the quark.

Most conservation laws are derived via Noether's theorem from a corresponding symmetry. Others are simply empirical. Two such empirical laws state that lepton number and baryon number is conserved in all interactions. The lepton number for leptons is +1, for anti-leptons -1, and for all other fundamental particles it is 0. Baryon number is $+\frac{1}{3}$ for quarks, $-\frac{1}{3}$ for anti-quarks and 0 for all other fundamental particles. From observations, there are no interactions seen where the lepton number or the baryon number change. However, this could be entirely accidental. The Standard Model offers no symmetry which demands this conservation. This is one of the unsatisfactory features of the standard model which is otherwise a very successful theory.

The proton is made of three quarks (two up quarks and one down) giving it a baryon number of 1. Since the proton is the lightest particle with baryon number 1 there are no other lighter particles to which the proton can decay while preserving baryon number. To find evidence of proton decay is to find evidence of baryon number non-conservation and to demand revision of the Standard Model.

1.2.2 ... and Beyond

As mentioned above, in the Standard Model the electromagnetic force and the weak force can be integrated in to a single representation, the electroweak force. The goal of Grand Unified Theories (GUTs) is to further merge these with the strong force. Most of these theories place leptons and quarks on similar footings. Because of this it is "easy" to have leptons and quarks couple which gives avenues for a proton to decay to lighter particles. Theorists are somewhat hard pressed to construct GUTs which predict baryon or lepton number violating processes which occur infrequently enough to be consistent with observation.

One of the earliest such theories places the $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ Standard Model groups into a single SU(5) group [3]. For example, this places the right handed anti neutrino ($\bar{\nu}_R$), positron (e^+_R) and the down quark (d_R, three colors) in a 5-plet and the left handed positron (e^+_L) and the up (u_L), down (d_L) and anti-up ($\bar{d}_L$) (three colors each) in a 10-plet. Within a multiplet particles can couple. This allows the proton to decay to a positron and a neutral pion as shown in Fig. 1.1. The mass scale of the force mediator X is

$M_X \sim 5 \times 10^{14}$ GeV/c^2 and the predicted partial lifetime for $p \rightarrow e^+\pi^0$ is,

$$\tau(p \rightarrow e^+\pi^0) \sim \frac{M_X^4}{\alpha_5^2 m_p^5} \sim 10^{31\pm1}\text{years}, \tag{1.3}$$

where α_5 is the strength of the unified interaction and $m_p = 938.3\text{MeV}/c^2$ is the mass of the proton. As seen in the following section, this predicted lifetime has already been shown to be incompatible with experimental observation.

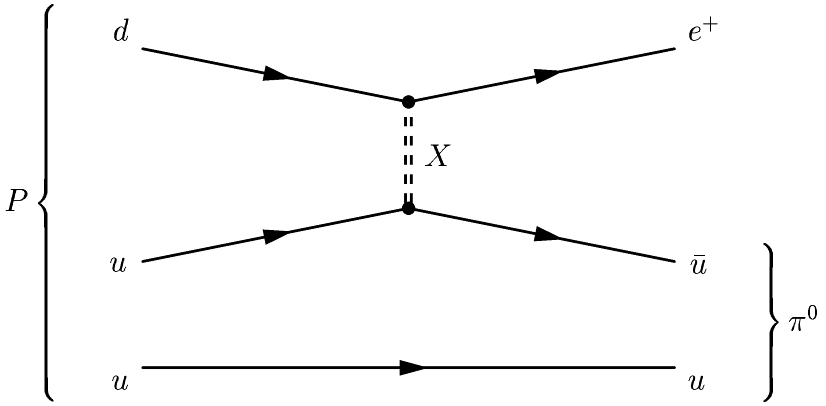


Figure 1.1: Proton Decay in SU(5).

In order to adjust for the non-observation of proton decay the minimal SU(5) model has been extended. For example, the so called *flipped* SU(5) (actually SU(5)×U(1)) rearrange which particles go in which multiplet. In doing this, the charged lepton and pion decay modes are still dominant but the predict lifetimes are now partially accessible in Super–Kamiokande. Two examples of predicted lifetimes are $\tau(p \rightarrow e^+\pi^0) \sim 10^{35\pm2}$ years [4] and $\tau(p \rightarrow \mu^+\pi^0) \sim 2 \times 10^{34\pm2}$ years [5].

Another extension employs a larger symmetry, SO(10). One such theory [6] predicts four possible partial lifetimes for proton decay via $p \rightarrow e^+\pi^0$. These lifetimes range from $10^{32.1}$ to $10^{37.7}$ years. Three classes of uncertainties on the predicted lifetimes are quoted and they are all in the exponent and all about $\pm1 \sim 2$. Again, this gives some predictions which are tantalizingly accessible by Super–Kamiokande but only partially so.

By adding additional symmetries, the non-observation of proton decay can also be accommodated. One such symmetry associates a new partner particle for each currently known particle. This so called Super Symmetry (SUSY) gives all fermions bosonic partners and *vice versa.* Except for the different spins the partners are otherwise identical. The masses of the mediators in this theory are much higher, ($M_X \sim 3 \times 10^{16}$ MeV) which causes decays like in Fig. 1.1 to occur with lifetimes around $\tau(p \rightarrow e^+\pi^0) \rightarrow 10^{38}$ years.

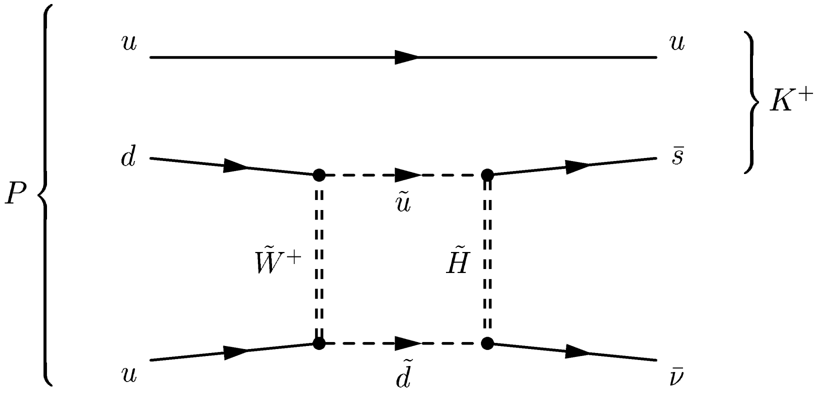


Figure 1.2: Proton Decay in Super Symmetric SU(5). The tilde ($\sim$) over the intermediate particles indicate they are a SUSY partner.

However, in this theory, the SUSY partners to the quarks, W and Higgs particle can contribute to the decay as diagrammed in Fig. 1.2. Here, the SUSY partners: sup ($\tilde{u}$), sdown ($\tilde{d}$), wino ($\tilde{W}^+$) and higgsino ($\tilde{H}$) mediate the decay. It is this mode, $p \rightarrow \bar{\nu}K^+$, which is expected to dominate the proton decay rate if SUSY is correct and lifetimes in the range of $10^{29\pm4}$ years, part of which is reachable by Super–Kamiokande, are predicted. However, as tantalizing as their predictions are, SUSY theories are not without problems. They must invoke additional symmetries, such as R-parity, $(-1)^{3(B-L)}$, to keep the proton decay rate below current observations. This low decay rate then forces a large, unnatural splitting between what should be the similar masses of the Higgs particles and their super-partners.

Recently, the exciting observation [7] of evidence for neutrino oscillation and thus neutrino mass has given theorists [8, 9] confidence in some higher order groups such as SUSY G(224) and SUSY SO(10). In them, the existence of neutrino mass is connected to the existence of baryon violating processes. It would be ironic indeed if the "accidental" discovery of neutrino mass in detectors designed for proton decay searches eventually leads to an explanation which, because of neutrino mass, demands an unstable proton, that in turn will someday be observed in the same such detectors.

1.3 The $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ decay modes and event signatures

A particular theory has a preferred mode (that which has the fastest predicted decay rate) and different theories prefer different decay modes. As discussed above, some models (*eg.* minimal SU(5) and "flipped" SU(5)) prefer $p \to e^+\pi^0$, while adding SUSY leads to preferring heavier products ($\bar{\nu}K^+$ or μ^+K^0). To be complete, an experimenter must attempt to find proton decay in all modes which are not ruled out by those conservation laws which are on a stronger footing than baryon number conservation. There are some 26 possible two body proton decay modes which can be searched for and work is in progress in Super–Kamiokande to complete these searches.

Presented here is work based on the first proton decay search [10] to be done with the Super–Kamiokande detector. Proton decay via either $p \to e^+\pi^0$ or $p \to \mu^+\pi^0$ is searched for. Besides the fact that these modes are preferred by various theories, they were chosen for study because they provide a very clear signal which is relatively easily distinguished from the atmospheric neutrino background. The distinction comes from the fact that the energy of the decay products are almost completely visible (the muon still has about 200 MeV of energy when it goes below Čerenkov threshold). Since all decay products are visible, it is easy to separate candidate events from background by demanding that the reconstructed total momentum of the products is consistent with a proton which is almost at rest (keeping in mind Fermi momentum for those protons in the oxygen nucleus) and that the invariant mass of the products is that of the proton. In comparison, most atmospheric neutrino events will have high reconstructed momentum and will have an invariant mass other than the proton mass. Further separation is done by looking for specifics to each mode, for example number and type of particles and decay electron signatures.

To describe the event signature in more detail the $p \to e^+\pi^0$ mode is focused on. Replacing the e^+ with a μ^+ makes small changes and these are pointed out. Figure 1.3(a) shows a cartoon of an idealized $p \to e^+\pi^0$ decay. In this ideal decay the positron e^+ and neutral pion π^0 exit the decay of the proton p in opposite directions. The positron initiates an electromagnetic shower and the pion decays to two photons of similar energies, each of which initiates an electromagnetic shower. The positron shower will make a single isolated Čerenkov ring and the photons from the π^0 decay will create two overlapping rings. After the π^0 decays, which is instantaneous as far as tracking in Super–Kamiokande is concerned, there can be no further decays. Any evidence of further decay electrons, for example, would invalidate an event from $p \to e^+\pi^0$

candidacy. Similarly, in the μ^+ mode exactly 1 decay electron is expected.

In Super–Kamiokande, such an ideal event might look like the event in Fig. 1.3(b). This is a Monte Carlo simulated event of a $p \to e^+\pi^0$ decay and it shows hit photomultiplier tubes (PMTs) as squares plotted as a function of $\cos\theta$ *vs.* ϕ as viewed from the fit event vertex. The size of the square symbolizes the amount of light collected in the PMT. The positron ring can be seen on the left and the two overlapping photon rings from the π^0 decay can be seen on the right. With a μ^+ as the lepton, the isolated ring would have less visible energy and sharper outer edge.

In general, real $p \to e^+\pi^0$ events will differ from this ideal picture for various reasons. For example, the pion can be absorbed or can scatter in the nucleus if the decaying proton comes from the oxygen. This will cause the two rings from the pion decay to either be missing or to not be pointed opposite of the lepton. The pion can also produce other pions or undergo charge exchange. Also, as mentioned above, the nuclear proton will in general have some momentum due to Fermi motion. This spoils the balance of the lepton and the pion momenta. In addition, the pion can decay asymmetrically so that the energy of one photon ring dominates the other and the angle between the two photon directions increases as the asymmetry increases.

Even with these complications, in comparison, other modes such as $p \to \bar{\nu}K^+$ are much more difficult to search for because some of the products, in this case the $\bar{\nu}$, will have energy that is invisible to the detector. The decay products which are visible (in this case from the decay of the K^+ into the dominant branches of $\mu^+\nu_\mu$ or $\pi^+\pi^0$) appear similar to many background events from atmospheric neutrinos since the total visible momentum is not close to zero. Searches for these modes are being conducted and results from some them are shown below in Fig 1.4 and are reported elsewhere [11].

On the surface the atmospheric neutrino induced background events to the two modes studied look similar to the expected signal. That is, they are multi ring events which produce an amount of visible energy near the proton rest mass energy. However, as mentioned above since these background events come from a neutrino scattering the products will have significant total momentum and will, in general, have invariant masses different from that of a proton. And any background containing charged pions or muons which are below Čerenkov threshold can still be tagged by looking at the number of subsequent decay electrons. As will be found in section 5.2, only one background event per mode from an effective 10 years of MC simulated atmospheric neutrinos survived the selection criteria.

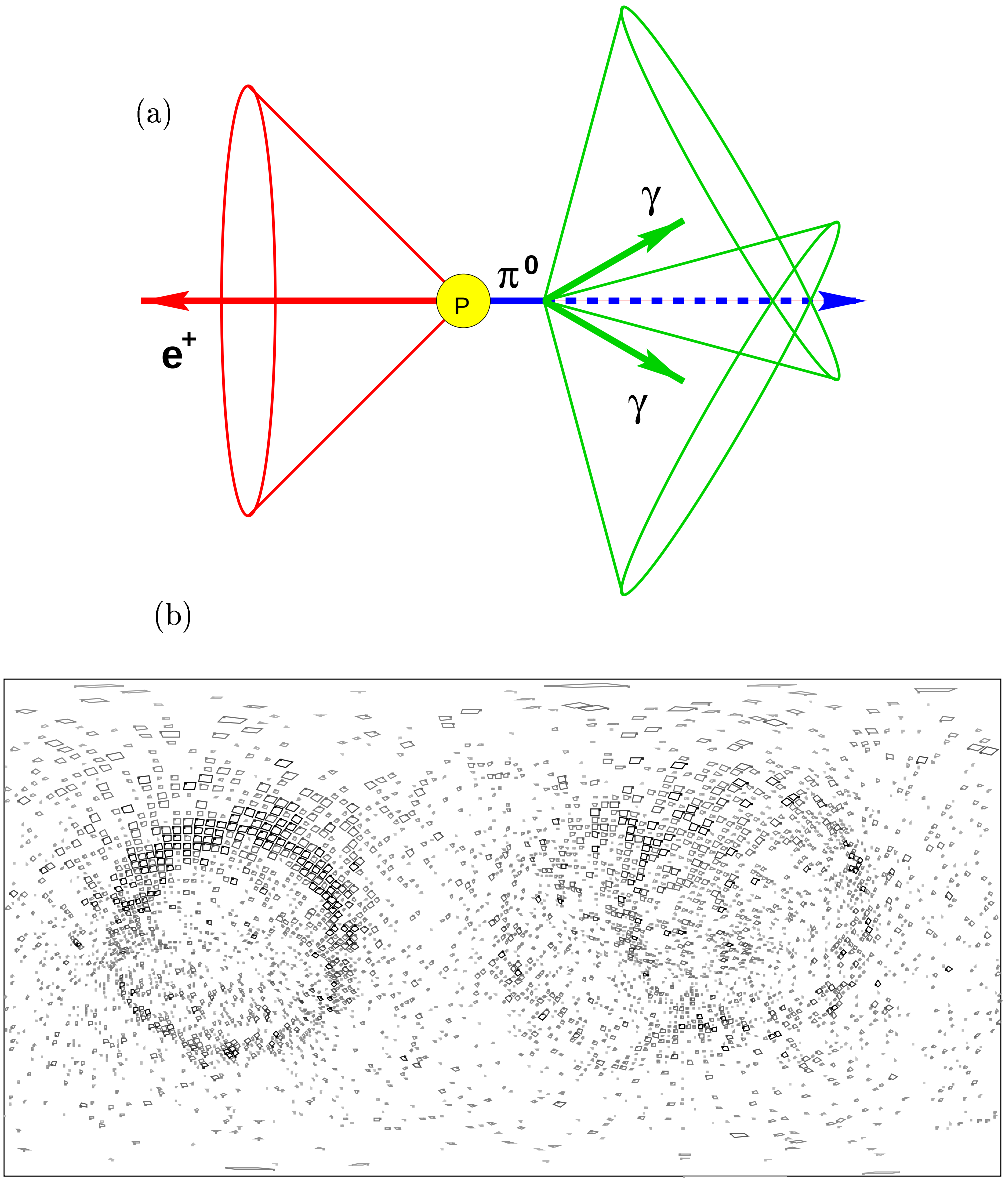


Figure 1.3: Idealized $p \rightarrow e^+\pi^0$ decay shown (a) as a cartoon and (b) as an MC simulated event.

1.4 Selective History of Proton Decay Searches

The nucleon decay searches performed before 1980 are reviewed [12] by F. Reines and J. Schultz. They break search methods down into residue-oriented and direct. Residue-oriented searches are done by looking at very old samples of ore or other material and searching for isotopes which could be produced by nucleon decay in a quantity greater than could be attributed to other processes. Direct searches strive to find nucleon decay by immediately observing the decay products. Some of these searches are mentioned briefly here.

The first example of residue-oriented searches mentioned in Ref [12] tells that in 1954 M. Goldhaber suggested that nucleon decay could leave Th^{232} in an excited and fissionable state. Comparison of the measured lifetime to that for spontaneous fission could be used to search for nucleon decay. This search was performed by G. Flerov, *et al.* [13]. He found no decays and set a limit of $\tau > 2 \times 10^{23}$ years.

Another residue-oriented suggestion, this one from S. Rosen [14] in 1975, was that nucleon decay products could accumulate in very old ores. A couple of years later, J. Evans and R. Steinberg [15] followed this suggestion and looked at Xe^{129} in tellurium ore. This isotope can be created via a chain reaction starting with either a proton decay or a baryon number violating neutron decay. Xe^{129} will collect in the ore as it does not β decay. The background for this search is Xe produced by cosmic ray interactions and this background was estimated by looking at the other isotopes of Xe. Correcting for background, they set a nucleon lifetime limit of $\tau > 1.6 \times 10^{25}$ years.

The first direct search for proton decay was made by F. Reines, C. Cowan and M. Goldhaber [16] using a 300 liter liquid scintillation detector. Analysis energy threshold was 15 MeV and the energy spectrum and intensity was consistent with the expected cosmic ray muon flux. They set a limit on the partial lifetime of $\tau > 10^{22}$ (at no particular confidence level) years for modes having high energy ionizing decay products.

The CWI (initials for Case Western Reserve University, the University of Witwatersrand and the University of California at Irvine) experiment was a later liquid scintillation detector which collected 67 ton-years of data. As reported in Ref [12], they found 6 events with a muon that stopped and decayed in the detector. They expected 0.6 interactions from cosmic ray muons and 11 atmospheric neutrino interactions and estimated that about half these interactions would be detected. They set a partial lifetime limit of $\tau > 6 \times 10^{30}$ years for modes producing a contained muon.

After 1980 four progressively larger detectors, some dedicated to nucleon decay searches, all employing direct search methods were constructed.

In 1981 the Soudan detector was constructed. This detector was an iron tracking calorimeter with just under a kton of mass located 2070 meters water equivalent underground in northern Minnesota. The Ar-CO_2 drift tubes gave the detector fine grained tracking resulting in a vertex resolution of a couple of cm. This resolution allowed photon induced showers and electron induced showers to be distinguished. In addition, proton, charged pion and muon tracks could be distinguished. Their latest results [17], for lepton plus kaon and lepton plus eta modes, found limits all around $\tau > 10^{32}$ years. Soudan 2 now finds itself in the path of the MINOS [18] neutrino beam as an auxiliary long baseline neutrino detector.

In 1982 the Irvine Michigan Brookhaven (IMB) detector [19] was built for the purpose of searching for nucleon decay. This was the first of the massive water Čerenkov detectors. The detector has gone through 3 stages, the last of which was dubbed IMB-3. It was located 1570 meters water equivalent underground in a salt mine near Cleveland, Ohio. The detector's rectangular fiducial volume contained 3.3 ktons of water which was viewed by 2048 20 cm PMTs. It detected nucleon decay candidates and their atmospheric neutrino background events by collecting Čerenkov light emitted by relativistic charged particles as they traveled through the water of the detector. More details on the use of Čerenkov light for particle detection are given in section 2.3. IMB-3 could veto cosmic ray muons, and detect electrons from muon decays. By exploiting the pattern of the Čerenkov light, it was possible to discriminate between heavy non-showering particles (muons and charged pions) and light electromagnetic showering particles (electrons and gamma rays). The final results [20] published in 1999 give limits for an extensive list of modes. In about half the modes studied by IMB there were no candidates found while the other half found a number of candidates which were all consistent with the number of expected background events. In particular, IMB-3 found no candidates for both $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ and, for these modes, they set partial lifetime limits of $\tau > 5.4 \times 10^{32}$ years and $\tau > 4.7 \times 10^{32}$ years, respectively.

Shortly after IMB, Kamiokande was built and also upgraded twice. Although containing a smaller fiducial volume holding just over a kton of water, Kamiokande had large 50 cm PMTs which achieved much better timing resolutions than those of IMB-3. The detector had just under 1000 PMTs which allowed 20% of the wall to be covered by photocathode. This made it possible to observe, in addition to atmospheric neutrino induced events and potential proton decay candidates, interactions of low energy neutrinos from the sun. With much of the same ability to search for nucleon decay as IMB-3, Kamiokande reported their search results [21] on numerous modes. Using the combined runs of the I and II detectors, they found no evidence for nucleon decay. They set partial lifetime limits of $\tau > 2.6 \times 10^{32}$ years and $\tau > 2.3 \times 10^{32}$ years for $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ modes, respectively.

In 1996 the largest water Čerenkov detector, Super–Kamiokande, was built for the purpose of searching for proton decay and studying the solar and atmospheric neutrino problems. As the name implies, this detector is very similar to a larger Kamiokande. Besides the increase in mass, it has twice the PMT coverage using improved PMTs. Details of the detector are described in chapter 2.

Proton decay limit results from Soudan, IMB, Kamiokande and some other notable experiments are summarized along with the recent results from Super–Kamiokande in Fig. 1.4[1].

1.5 The Super–Kamiokande Collaboration

The Super–Kamiokande collaboration started largely as a joining of forces of members from the Kamiokande and IMB experiments as well as a sizeable number of "third party" members. This last group includes State University of New York, Stony Brook. Super–Kamiokande currently consists of 22 institutions and about 125 physicists, mostly from Japan and the US with a few from Warsaw, Poland and Seoul, Korea. The host institution is the Institute for Cosmic Ray Research (ICRR) of Tokyo University.

Construction and maintaining the inner detector (see section 2.6) was primary responsibility of the Japanese groups while the outer detector (see section 2.7) was the responsibility of the U.S. groups. Responsibilities were shared for detector calibration, monitoring, shift taking and other duties.

1.5.1 A Tale of Two Groups

Initially, the Super–Kamiokande collaboration decided to have two separate, independent groups for the solar neutrino, atmospheric neutrino and proton decay analysis. This was done in order to provide cross checks as there was (and is) no other experiments online with the same capabilities to check our results. These two groups were dubbed *onsite* and *offsite* analysis groups and as the names imply, the division between them was largely (but not entirely) along national lines. Both groups shared a common data flow up to

---

[1] This figure was provided by C. McGrew

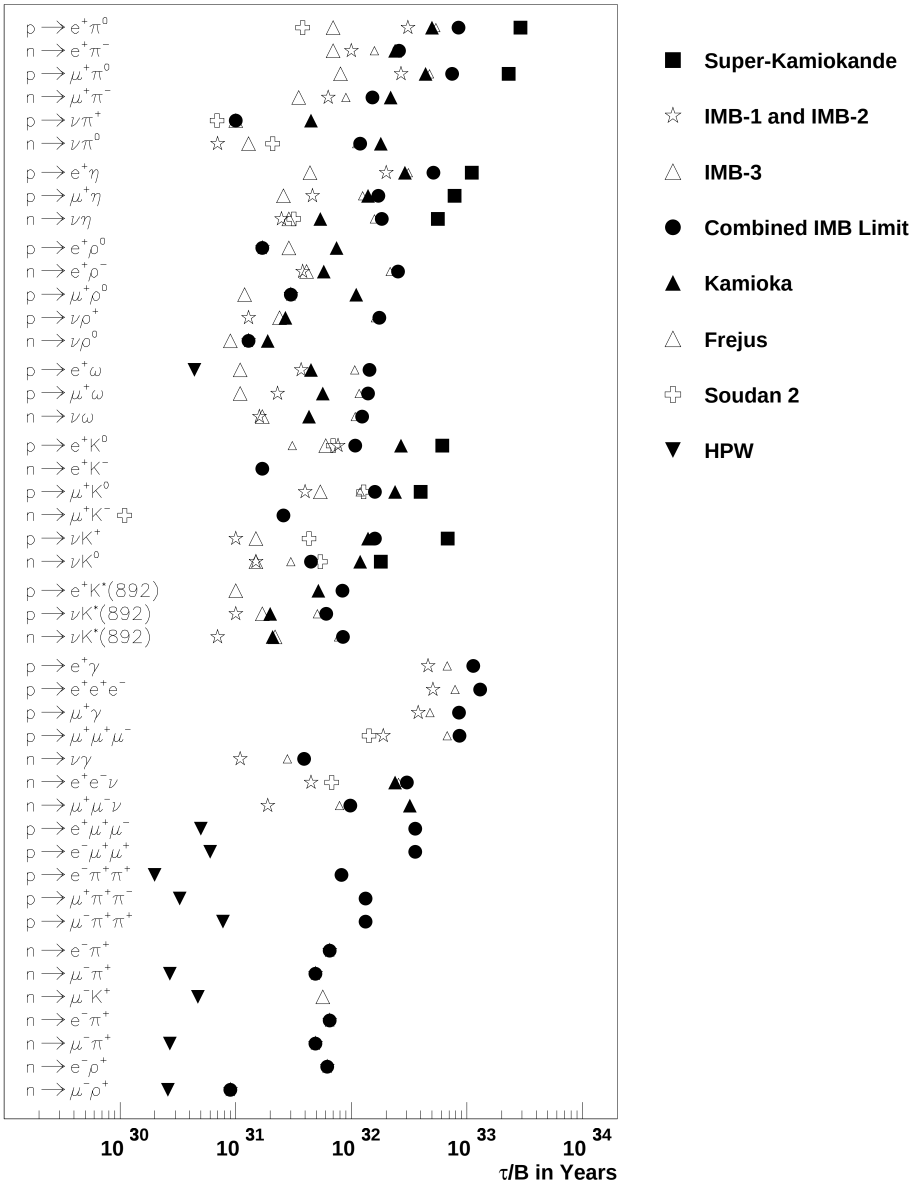


Figure 1.4: Summary of nucleon decay partial lifetime limits (τ/B) from a selection of detectors, including all modes currently measured at Super–Kamiokande.

and including initial charge and timing calibrations (see section 2.9.4). The data for the onsite group was then processed on site, while the offsite data was written to tape and shipped to Stony Brook for initial processing (see section 2.9.5) .

Both groups had large contingents from past experiments. In the case of the onsite analysis, much of the event reduction and reconstruction was adopted from the Kamiokande experiment. The offsite analysis took some of the ideas from the IMB experiment and implemented many new ones. Virtually all of the reduction and reconstruction code was written from scratch within a cohesive programming and data architecture. Because of these differences, it is a fact that the offsite analysis was not as finely tuned as much of the onsite codes, nonetheless in comparison it was simple, effective and very well understood.

Up until 400 days of data had been collected, these two independent analysis continued to show very good agreement in essentially all physics results. At 400 days, the first atmospheric neutrino results [7] were published showing this strong agreement. At that time it was decided to conserve manpower and merge the two groups. It turned out that this merger largely consisted of abandoning much of the offsite work. Since a lot of effort had already been put forth into the analysis for this thesis it was also decided to continue to use the offsite code base for this thesis. This was of mixed results. On one hand, this analysis continued to be based on well understood, flexible and proven tools. On the other, since by and large most of the offsite members moved to using onsite codes there was not much supporting manpower for this analysis.

Chapter 2

Nuts and Bolts: the Super–Kamiokande Detector

In this chapter the Super–Kamiokande detector is described. The phenomenon of Čerenkov[1] radiation and how it is used to detect particles is explained and the characteristics of different events to which Super–Kamiokande is sensitive are summarized. Finally, the flow of data starting from the PMTs and ending at the raw data sample is presented.

2.1 Location

The Super–Kamiokande detector is located at $36°25'33''$ N, $137°18'37''$ E and 371.8 m above sea level. As shown in Figure 2.1, this puts it about 250 km west of Tokyo in Gifu prefecture on Japan's Honshu island. It is near the town of Kamioka, from where the name is derived (Kamiokande is from Kamioka Nucleon Decay Experiment which, over time, has oscillated to mean Kamioka Neutrino Detector Experiment). Super–Kamiokande is in the Japanese Alps in one of the most beautiful areas of Japan. After an approximately 30 km scenic drive south from the Toyama city's bay up to Mt. Ikenoyama[2], it is situated in an active zinc mine owned and operated by the Kamioka Mining and Smelting Company. The main laboratory building and computer center, the *kenkyutou*, is in the small town of Mozumi where it sometimes seems the physicists outnumber the (very wonderful) locals. While working in the mountains, most physicists enjoy the 30 minute commute to and from the town

---

[1] This is often misspelled as both Cherenkov and Čerenkov. For this thesis, the spelling of Čerenkov is chosen.

[2] This "standard" translation is somewhat redundant as "yama" means mountain.

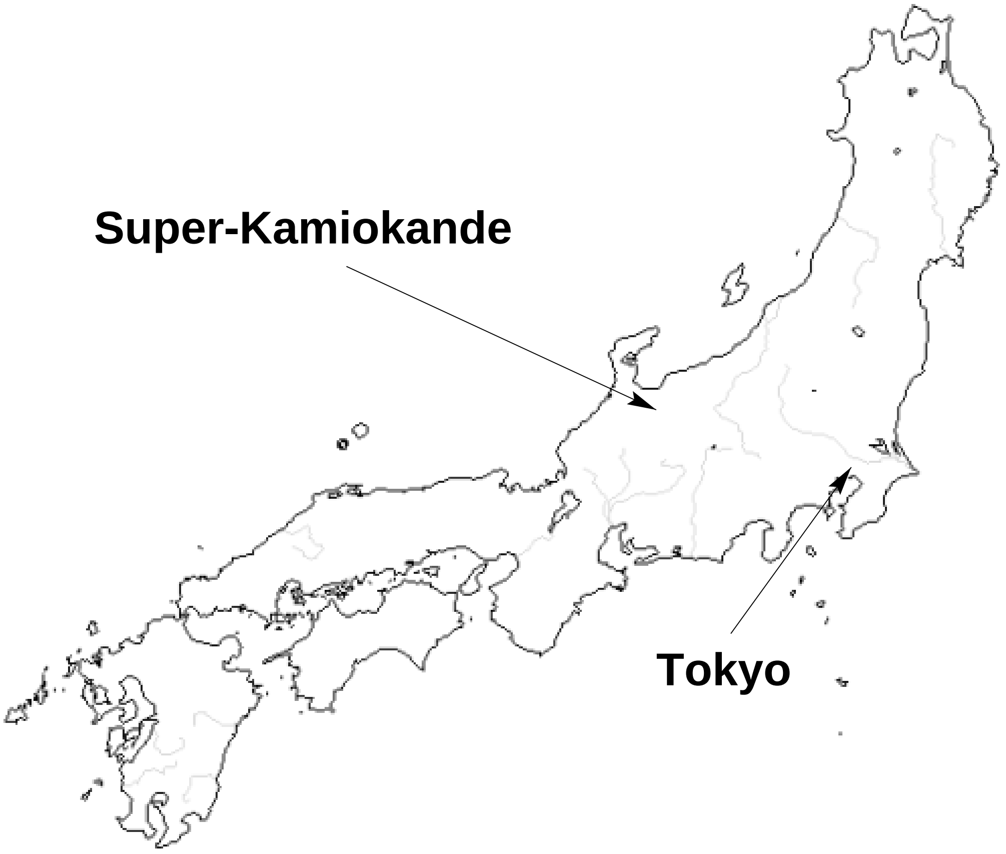


Figure 2.1: The location of Super–Kamiokande relative to Tokyo

of Osawano in Toyama prefecture. Here, a few of the US groups, including Stony Brook, maintain apartments where people who are spending a long term at the experiment call home.

2.2 Overview

Figure 2.2 shows an artists conception of the Super–Kamiokande detector and local environs. The detector is located about 2 km along a straight horizontal drift as shown in the inset slice-away. By the time the detector is reached, Mt. Ikenoyama has risen to provide an average of 1000 m (2700 m water equivalent) of rock overburden. Near the detector is a side tunnel housing

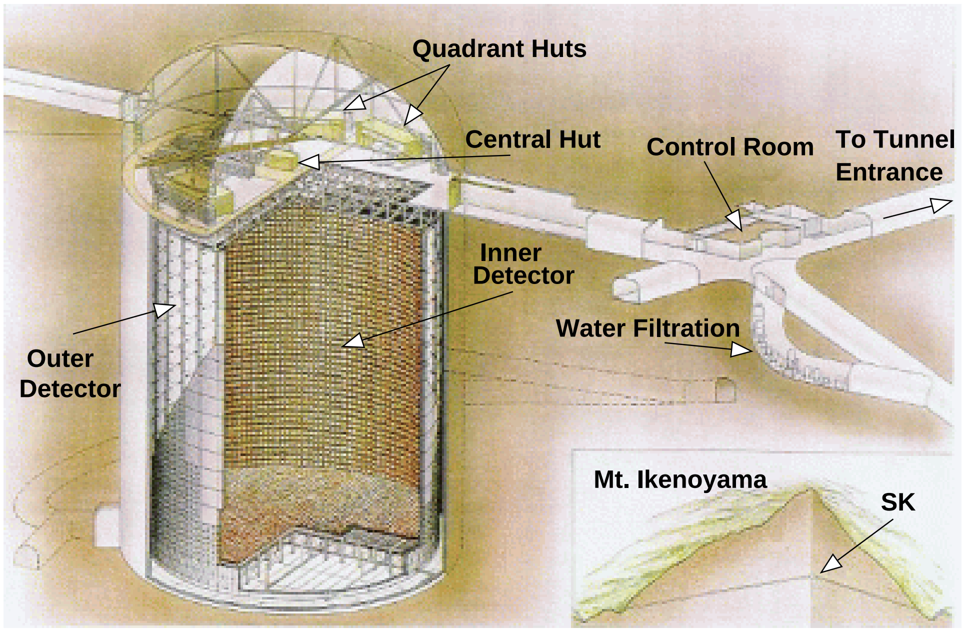


Figure 2.2: Super–Kamiokande location within Mt. Ikenoyama

the water filtration and another holding the control room where shift members monitor the detector 24 hours a day. Both the detector cavity and the control room are kept at a positive air pressure by fresh air of low radon content being piped in from outside the mine. In addition, all exposed mine rock near the detector and control room is covered in a polyurethane like material, called mine guard, to reduce the amount of emitted radon as well as dust and debris.

More detail of the actual detector and cavity can be seen in Fig. 2.3[3]. This cross section shows the cavity and dome as well as the inner detector (ID), outer detector (OD) and the dead space between the two.

2.3 Čerenkov Radiation

The primary physical phenomenon behind the Super–Kamiokande detector is called Čerenkov radiation [22, 23]. This blue light is emitted when a relativistic charged particle travels faster than the speed of light in the lo-

---

[3] Thanks to K. Martens for this figure. However, as far as we know, there isn't really a giant fish at the bottom of the detector.

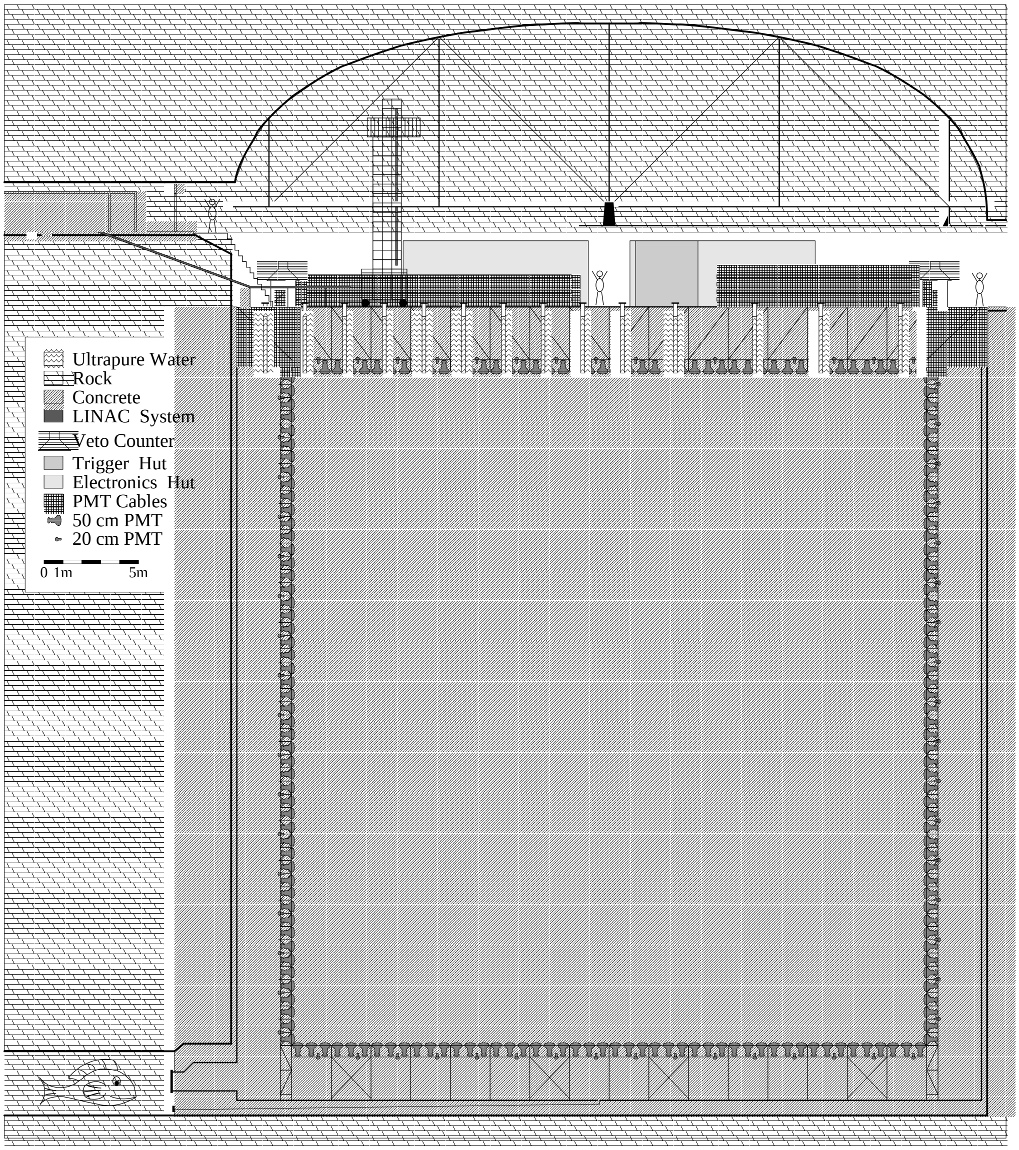


Figure 2.3: Cartoon of Super–Kamiokande detector and cavity.

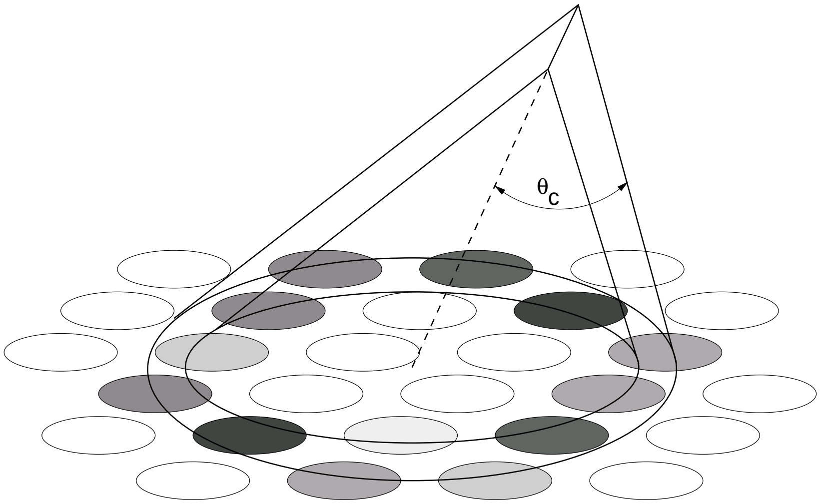


Figure 2.4: Cartoon of Čerenkov cone produced by relativistic charged particle in water and resulting ring on detector wall. The circles represent photo-detectors and their shading represents the amount of light collected in each.

cal medium. The passage of the charged particle electromagnetically disturbs the material of the medium. Upon de-excitation of the medium, radiation is emitted and constructively interferes such that a conical wavefront propagates away from the track of the charged particle as shown in Fig. 2.4. The half angle of this cone is given by,

$$\cos\theta_c = \frac{1}{\beta n(\lambda)}, \tag{2.1}$$

where β is the speed of the particle in units of c and $n(\lambda)$ is the index of refraction of the medium at a wavelength of λ. The number of photons emitted per unit length and per unit wavelength by the passage of a particle with charge $\pm e$ is,

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{\lambda}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right) = \frac{2\pi\alpha}{\lambda}\sin^2\theta_c, \tag{2.2}$$

where α is the fine structure constant.

For the water in Super–Kamiokande the index of refraction is $n \approx 1.35$, which is fairly constant over the range of wavelengths where the PMTs are sensitive. This results in a Čerenkov angle of $\theta_c = 42^\circ$ and a photon production rate of about $dN/dx \approx 575$ per cm. In simulations the photon production rates and angles are more accurately predicted by integrating over the Čerenkov spectrum as well as the photo-detector's efficiency.

By placing light sensors in the path of a Čerenkov cone and measuring the number of photons and the time at which they struck the sensors it is possible to reconstruct the position, direction and type of the original charged particle. Also, by counting the total number of photons collected it is possible to estimate the amount of energy deposited by the particle. These two issues are discussed more in Chapter 3.

2.4 Event Classes

Super–Kamiokande detects a variety of different types of events. These event classes are described below roughly in order from the lowest deposited energies of a few MeV to the highest of tens of GeV.

At the low energy end there are solar neutrino events and their primary backgrounds, radon and spalation. These events are all far below the energies of the proton decay modes studied in this thesis. In the low energy region there are also electrons from muon decays. Their energies range up to the theoretical limit of 52.8 MeV which can fluctuate to about 60 MeV due to energy resolution.

At higher energy deposition there are atmospheric neutrino induced events. Atmospheric neutrinos will primarily interact either by the charged current (CC) or neutral current (NC) weak interaction with the nuclei in the hydrogen and oxygen of the water. Interaction with the electrons of the water molecules is also possible but occurs much more rarely. CC interactions produce a single visible lepton and, if the momentum transfer is enough, one or more visible pions. NC interactions produce one or more visible pions. In both cases, the pions may be absorbed or scattered on their way out of the nucleus. Neutral pions almost always decay to two photons which initiate electromagnetic showers. If the decay is very asymmetric, one of these showers may be too dim to discern. Most of the atmospheric neutrino induced events are fully contained (FC) within the fiducial volume of the detector. If the neutrino is of high enough energy, the outgoing lepton can exit the detector. These events are called partially contained (PC) and are almost always (98%) charged current ν_μ events.

The high end of observed energies has contributions from stopping and through going muons (see below for definitions). These muons can be from either cosmic ray muons or from muon neutrinos which interact in the rock around the detector. The cosmic ray muon flux decreases to zero as the originating direction of the muon goes below the horizon. Any upward going entering muons are considered to be from neutrino interactions in the rock.

The longer through going muon tracks deposit 10 GeV of energy. More energy will be deposited if a muon undergoes bremsstrahlung.

Finally, there may be nucleon decay events which will be limited in energy to be below about 1 GeV. Figure. 2.5 shows a cartoon of these and the other higher energy event classes.

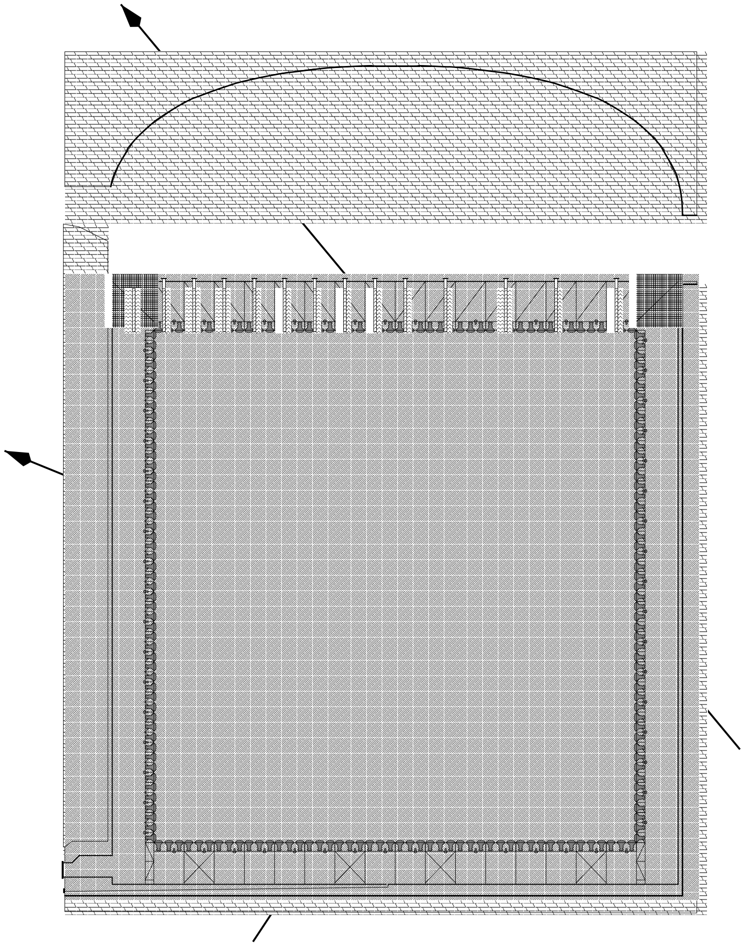


Figure 2.5: Classification of Super–Kamiokande event types: (a) contained single ring ν_{atm} induced, (b) partially contained ν_{atm} induced, (c) stopping μ, (d) through-going μ, and (e) multi ring contained (includes multi ring atm. ν induced and possible proton decay events).

2.5 Water Filtration Systems

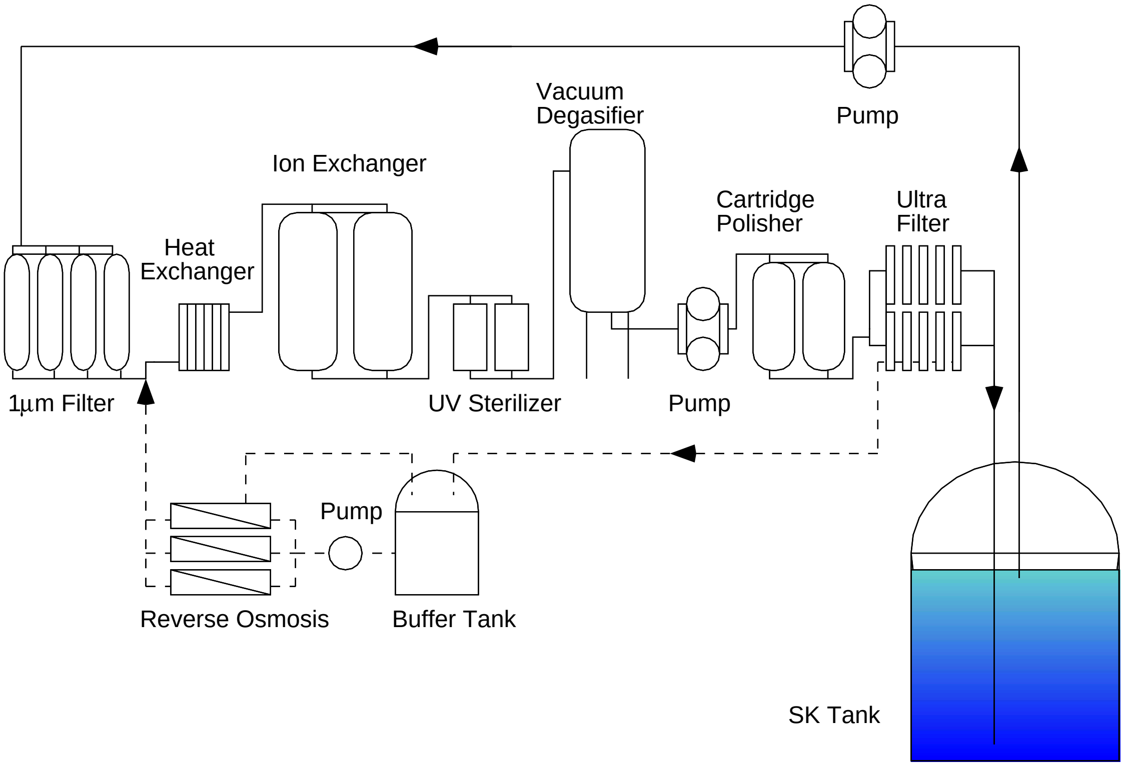


Figure 2.6: Water Filtration System

In order to detect as many of the emitted Čerenkov photons as possible, thus keeping the efficiency of the detector high, it is crucial to have extremely clear water. Super–Kamiokande arguably has some of the clearest water in the world with an attenuation length approaching 100 m for some wavelengths. In order to achieve and maintain this water purity a sophisticated water filtration system was developed.

Starting from an ample supply of water which is prefiltered as it trickles through mountain fissures, the water is sent through the system illustrated in Fig. 2.6. The various elements of the system are used to remove metal ions, dissolved oxygen and radon gases as well as more macroscopic dust particles down to sub μm sizes. Additionally, UV radiation and a temperature of 14°C kills any bacteria which may attempt to grow in the system.

Besides achieving optical clarity, there is a concerted effort to reduce the amount of radon in the detector tank. This effort has produced some of the most radon free water in the world and many theses could be devoted to just describing it. Radon contamination is a serious concern for low energy physics at Super–Kamiokande, such as solar neutrino studies. Thankfully, however, the studies presented in this thesis are affected little by the effects of radon contamination.

2.6 Inner Detector

2.6.1 Description

The inner detector (ID) is the primary volume for detecting events in Super–Kamiokande. For an event to be considered important for this thesis it must originate and stay completely within this region. The ID is a cylindrical volume 33.8 m diameter and 36.2 m high and contains 32.5 metric ktons of water. 40% of the walls are covered by photocathode from the 11146 50 cm diameter PMTs (see below) the rest by a semi-opaque black plastic used to help optically isolate the inner detector.

2.6.2 ID PMTs

A schematic of an inner detector PMT is shown in Fig. 2.7. A PMT works by turning a photon into a measurable electric pulse. It does this by first converting the photon to an electron (a so called *photo electron* or PE) through the photo-electric effect when the photon hits a deposited layer of bialkali (Sb-K-Cs) which is on the inner surface of the PMT glass. This layer is called the *photocathode.* The conversion is successful about 22% of the time (the average PMT quantum efficiency at a wavelength of $\lambda = 390$ nm). Because the photocathode is so near a large volume of water it is kept at ground. The photo electron is then accelerated to about 800 V and strikes the first dynode. Upon hitting the first dynode many electrons are liberated and accelerated through another 300 V to the second dynode. Each of these electrons continue showering as they hit 9 more subsequent dynodes each at about 100 V higher potential than the previous. After reaching the anode the shower has passed through about 2000 V and increased to about 10^7 electrons for each initial photo electron. For a single photo electron, the typical transit time from photo cathode to anode is about 100 ns with a spread of about 2.5 ns. This PMT is made by Hamamatsu, initially for Kamiokande with later design improvements [24] for Super–Kamiokande.

The PMTs are mounted to the stainless steel support structure between

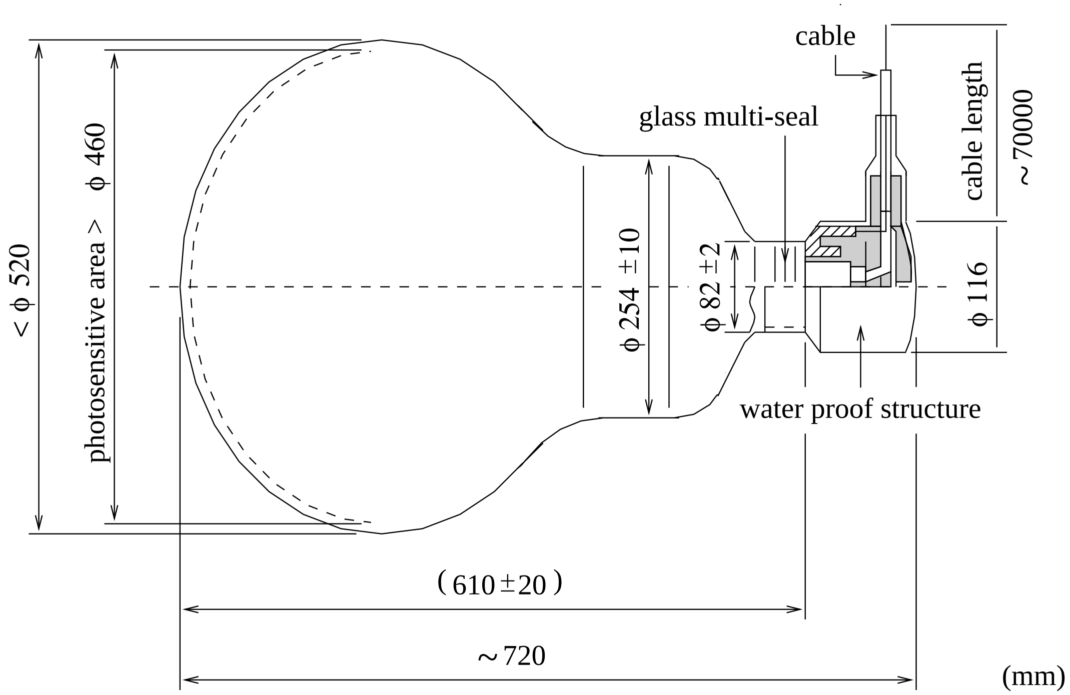


Figure 2.7: Inner Detector PMT

the inner and outer detectors by stainless steel bands. Only the photo sensitive face protrudes through the sheet of opaque black plastic and into the volume of the inner detector. The relative location of the inner PMTs is shown in Fig. 2.8. The high voltage (HV) and signal is carried through a 70 m cable between the PMT and the HV supplies and Data Acquisition (DAQ) systems.

2.6.3 Data Acquisition and Electronics

The inner detector data acquisition system is shown in Fig. 2.9. This is a simplified description of the ID DAQ, for a more full treatment see Ref. [25]. The inner detector is separated into four quadrants and data from each of these quadrants are sent to one of four quadrant huts containing digitizing electronics.

The inner detector PMT cable has an unshielded conductor for the HV, next to a shielded conductor for the signal. The signal is fed to one of 12 channels on one of the 20 Analog Timing Modules (ATM), housed in one of the 12 TRISTAN/KEK Online (TKO) electronics crates, in one of the four

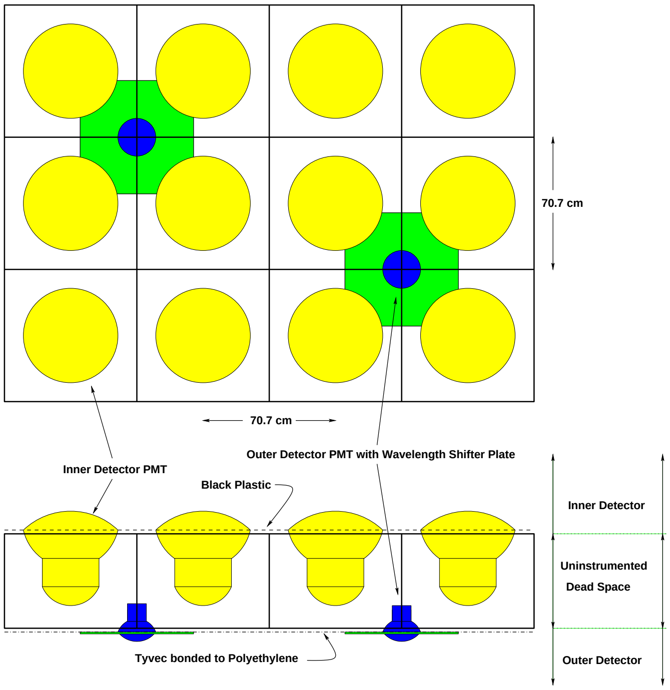


Figure 2.8: ID/OD PMT Placement.

quadrant huts. Each TKO crate also has a Super Control Header (SCH) module and a GO No Go (GONG) module. These are described below.

Each ATM channel has two sub channels, A and B, which allow recording a subsequent PMT hit, reducing dead time. In each of these sub channels are time to analog converters (TAC) and charge to analog converters (QAC). These analog values will be digitized if a global trigger is received by the GONG module. The PMT data are read out through the SCH into one of six buffers known as Super Memory Partners (SMP) which reside in one of two VME crates. The data are read out of the VME crate by a workstation, one per crate, via a VME-Sbus interface module. Each of these 8 workstations, (`sukon[1-8]`), then send their 1/8 of an event along the "online" FDDI network to a central host computer (`sukonh`) for event building.

To form the ID trigger a HITSUM signal is first made in each ATM. This signal is built by adding 200 ns pulses, one for each PMT. The HITSUM for each ATM is then added and sent to the global trigger logic (TRG VME module) which resides in the central hut. The TRG will produce a global trigger based on the inner detector if it finds 29 or more inner PMTs have fired within 200 ns. The individual triggers which cause the TRG module to

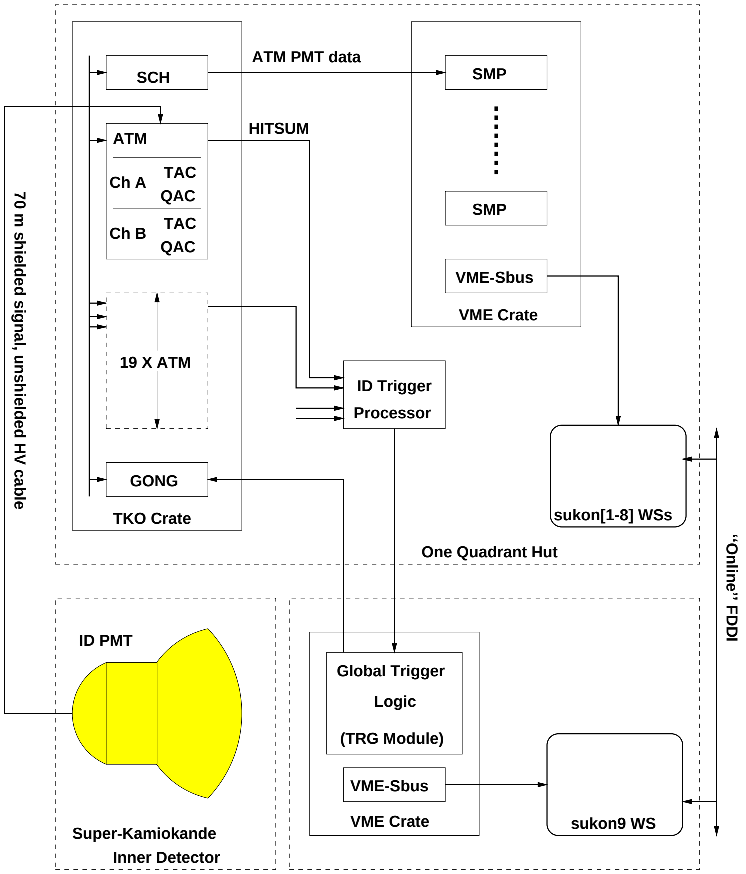


Figure 2.9: Inner Detector DAQ.

produce a global trigger is recorded and read out by a workstation (`sukon9`) and sent to `sukonh` to be included in the built event. The final event spans about 1 μs in time and varies depending on when exactly the HITSUM signal reaches the discrimination threshold.

2.7 Outer Detector

2.7.1 Description

The outer detector (OD), (or anti-detector), is primarily used as a veto against incoming cosmic ray muons. It is a cylindrical shell between 2 m thick at the wall and 2.2 m thick at the top and bottom. It completely surrounds the inner detector with 14.7 metric ktons of water. Mounted on the inner wall and facing outward are 1885 PMTs with wavelength shifter plates (WSP). The outer wall is covered with a white reflective paper like material called Tyvec. The inner wall is covered by Tyvec bonded to opaque black low density polyethylene. The Tyvec increases Čerenkov light collection and the polyethylene increases the optical separation between the OD and the rest of the detector.

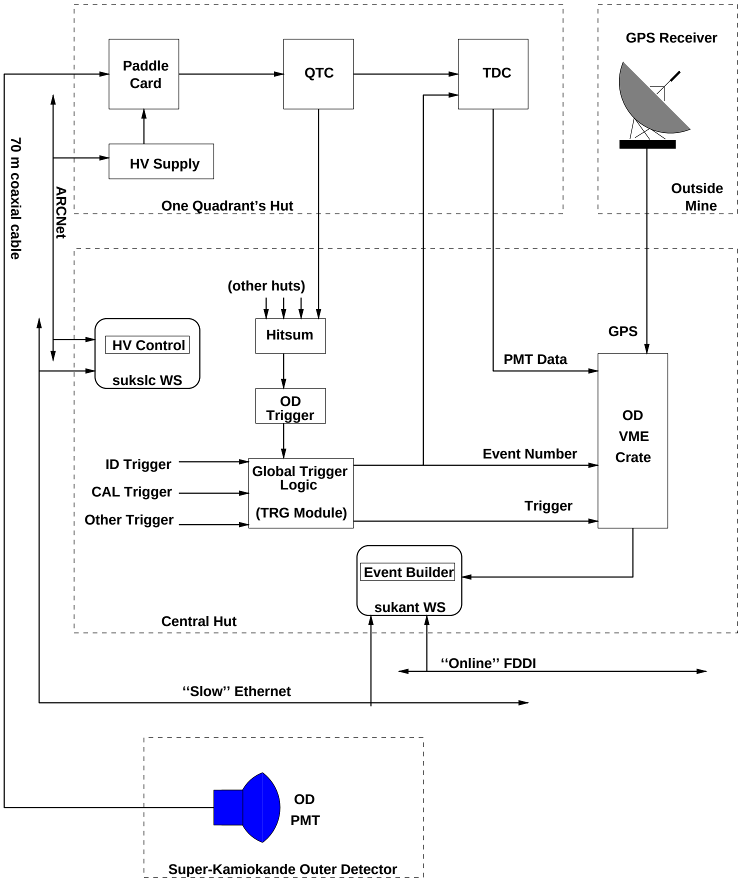


Figure 2.10: Outer Detector DAQ.

2.7.2 OD PMTs and WSPs

The photo multiplier tubes [26] and wavelength shifter plates used in Super–Kamiokande's OD are recycled from the IMB experiment [19]. These 20 cm diameter PMTs are also made by Hamamatsu however they have different characteristics than the ID PMTs. Besides being smaller, their transit time spread is 11 ns FWHM at one PE and 5 ns FWHM at 10 PE. These PMTs are also deployed more sparsely in the OD than those in the ID. For every 12 ID PMT, there are but 2 OD PMTs (see Fig. 2.8). This combined with the smaller size of each OD PMT gives only about 1% photocathode coverage. To improve the light collection, as in IMB, the PMTs have been fitted with wavelength shifting plates. These 60 cm square, 1.3 cm thick acrylic plates contain 50 mg/l flour bis-MSB and make contact with the edge of the PMTs at a close fitting hole in the center. The edges of the plates are lined with reflective aluminum coated mylar tape. When Čerenkov light hits the plate, it is absorbed and re-emitted at a wavelength more suited to detection by the PMT. These plates increase the light detection efficiency by 60% over bare PMTs, but also increase the single PE transit time spread to 15 ns FWHM.

2.7.3 OD DAQ

The outer detector data acquisition system is shown in Fig. 2.10. This is a greatly simplified description of the OD DAQ, for more details see Ref. [27]. As with the ID, the OD data are collected in chunks of quadrants. The figure shows one quadrant and the central hut OD electronics.

HV is supplied to the OD PMTs by the same 70 m coaxial cable which carries the signal. This cable goes to “paddle cards”, custom built by our LSU collaborators, where the AC signal is decoupled from the DC HV. The voltage is supplied by a LeCroy 1454 HV mainframe, one in each quadrant hut. Nominal voltage is 1800 V, however each paddle card can be supplied with voltages ranging up from 1500 V to about 2500 V. The 12 PMTs on one paddle card are selected to have matching gains for a given voltage. Each of the 4 HV mainframes are connected to the “slow control” workstation (`sukslc`) via ARCNet. This allows monitoring and setting the HV from remote locations.

After the PMT signal is stripped from the HV it goes to the charge to time converters (QTC). Both the HV and the QTCs are from our BU collaborators. The QTC will produce a pulse with a leading edge holding the PMT hit time and a width containing charge information. The QTCs also produce a HITSUM for each quadrant similarly to the ID. The HITSUM signal is the sum of a 200ns long pulse for each PMT beginning at the PMT hit time. This is joined with the other quadrant HITSUMs in the TRG module. An OD trigger occurs when the HITSUM reaches 19 PMTs hit within the 200 ns.

The QTC signal then goes to the time to digital converters (TDC), from our UW collaborators. The TDCs turn the QTC signal into digitized time and charge values. Also, the event number, sent from the global trigger electronics, is recorded in the TDCs. These data, as well as those from the other three quadrant huts, are sent to electronics in the OD VME crate in the central hut. With the help of the OD workstation (`sukant`) the 4 quadrants are assembled and sent along to "online" FDDI network to the main host workstation (`sukonh`) for merger with the ID data. The data is then transfered to the "online-offline" network as described in section 2.9.

In addition to OD PMT hits, the OD DAQ is also responsible for GPS time as well as a 50 MHz clock. This data is built into the OD header data bank.

2.7.4 Veto Hats

All of the ID and OD PMT cables must run through the top of the OD in order to exit the detector. In four places these cable bundles are about 1 m in diameter which is significantly large as to allow cosmic ray muons to enter the ID with out triggering the OD. To veto these events, scintillating paddles are placed over the exit point of these bundles. Due to their shape they have been dubbed "veto hats". Although these hats play the same role as the OD, their data are actually taken with ID electronics. These veto hats are not described in any more detail because this thesis removes the unwanted cosmic ray muons by software cuts and the veto hat data are not used.

2.8 PMT Calibration

In order to get useful physical information from the PMTs, they must be calibrated. They must all give the same gain relative to each other and digitized counts of time and charge must be converted to actual real valued numbers and the unwanted dependence of time on charge must be removed.

The relative gain of the PMTs is normalized by exposing the PMTs to a scintillator ball placed at various locations in the detector that is lit via an optical fiber by a Xe lamp. The HV supplied to individual PMTs is adjusted until all PMTs find the same relative charge (corrected by water attenuation and PMT angular acceptance, see section 3.4.3). After the adjustment is made, further exposures were done and the gains were found to have a spread of 7%.

The absolute gain is found by exposing the PMTs to low energy (~ 8 MeV) gamma rays emitted from nickel after capture of neutrons from a ^{252}Cf source. The gamma rays will produce, on average, less than one PE in a PMT. The resulting distributions of charge collected in the PMTs are used in the Monte Carlo detector simulation and gives a correspondence of about 2 pC for every PE.

It will take a different time in each PMT for a photon to produce an electric pulse. This is due to differences in the transit time between the creation of the initial PE and the development of the resulting pulse and because of differences in response of the discriminator. Transit time depends mainly on the voltage applied to the PMT. The discriminator gives pulse height dependent time differences because higher pulses tend to cross the threshold sooner. To calibrate this charge dependent timing correction, light from a fast laser pulse (~4 ns full width) at a wavelength of 384 nm is fed through an optical cable to an isotropic diffuser ball located in the tank. The time each PMT is hit, relative to the pulse is measured as a function of collected charge and a map (TQ map) is built up. This map is then applied to correct the time of each PMT hit.

2.9 Data Transfer

2.9.1 `sukonh`

As mentioned above, the inner detector data from the eight `sukon[1-8]` workstations, the outer detector data from `sukant`, and the trigger information from `sukon9` are all sent on the "online" FDDI network to `sukonh`, a Sun 4U (UltraSparc) workstation, located in the control room. With `sukonh` the event is built from the eight half quadrants of inner data, the fully built outer detector event and the trigger information. From here the data are sent along the "online-offline" FDDI network to two other workstations, `kingfish` and `sukrfm`.

2.9.2 `kingfish`

The data which is sent to `kingfish` is split off from the data which is eventually analyzed. `Kingfish`, a Sun 4M (Sparc20) workstation provided and maintained by our LSU collaborators, is used as to monitor many aspects of the data stream. It produces various histograms of such quantities as rates of the different triggers, PMT hitmaps, flasher watches (flashers are PMTs which arc and emit light, triggering the detector and are discussed more in sec. 3.2.2), calibration parameters, and others.

`Kingfish`'s other purpose is to run the online event display. This is a graphical color representation of the events. PMT time and charge information is shown as well as trigger and event header information.

2.9.3 `sukrfm`

The **`sukrfm`** Sun 4U workstation runs various programs, the most important of which is the data reformatter, from which the machine derives it's name. The reformatter converts from the packed bit level online format into one consisting of normal integers and floating point numbers. This inflated format is based on the ZEBRA [28] format. Once the data is reformatted it is sent out the mine and down the road to the kenkyutou for offline processing and storage by the **`sukop`** computer. Additionally, various monitors for offline processing, supernova watch, PMT hit quality and others run on **`sukrfm`**.

2.9.4 `sukop`

Originally, the data from the mine was received from by many Sun 4M machines named **`sukeve`** that have since been replaced with a single machine named **`sukop`**. The **`sukop`** computer is a massively parallel Sun 4U (UltraSparc Enterprise 10000) machine. It is used to apply the time and charge calibration to the PMT data, run the so called "real time"[4] data reduction processes of the onsite group and save raw and onsite-reduced data to a Sony Petasite tape library. It is also responsible for sending the data to **`sukus1`**, the US computer for copying the data for offsite analysis.

2.9.5 `sukus1`

`Sukus1`, originally a Sun 4M (Sparc20) Solaris and later an Intel x86 GNU/Linux machine was provided and is maintained by the Stony Brook group. It received the data originally via the "US-offline" FDDI network which later changed to fast Ethernet. The change from a Sun on FDDI to an Intel x86 on fast Ethernet was due to the original hardware being unreliable and expensive to maintain as well as the operating system being too limited.

The primary purpose of this machine is to run the USCOPY process which saves the data to two Digital Linear Tape (DLT) drives. Each drive holds five type IV DLTs and each DLT holds 20 GB (GB $\approx 10^9$ bytes) of compressed data. Since Super–Kamiokande data is rather incompressible, it is possible to only put about 27 GB on each tape. This, coincidently works out to almost one DLT cartridge per day of data.

---

[4] This is a misnomer as there is nothing real time about it, rather data is processed "live" from the detector instead of from some archive storage media.

Originally the USCOPY process was a rather complex set of programs which read the data from the network via sockets, shared the data via memory, and finally wrote to tape. Because it had to stay up as much as possible, a memory mapped (mmap) file was used in order to allow altering some of the programs variables while it was running. This was necessary to recover from tape hardware, network failure or other problems. After the hardware upgrade, this set of processes are now merged into a much simpler single program which doesn't need to explicitly know about the network but rather reads the data via Network File System (NFS) files.

After the data is saved to DLTs it is mailed to Stony Brook for reduction, reconstruction and further analysis, as described in chapter 3.

Chapter 3

Data Reduction and Event Reconstruction

As mentioned in section 1.5, two independent data analyses were carried out in Super–Kamiokande. This thesis is based on the offsite reduction and analysis chain. Periodically data would be shipped from the experiment and processed at Stony Brook. It first underwent data reduction to reduce the size and number of events to a manageable level. From this reduced set a contained event subsample was selected and various reconstruction algorithms were run. This chapter describes the hardware and, more importantly, the software used for the data reduction and reconstruction.

3.1 Hardware

The data reduction was performed using three DEC Alpha (21064a CPUs) workstations running OSF/1. Attached were four Digital Linear Tape (DLT) drives, one 8mm tape drive and an array of disks.

Most of the reconstruction software was developed and run at Stony Brook on various clustered GNU/Linux workstations running on Intel x86 hardware. These were chosen initially as an inexpensive alternative to X-terminals but as GNU/Linux improved, the price of PCs dropped and their performance increased we moved more and more of our work onto this platform. The move met with some resistance and difficulties, but it has come to allow us to easily, cheaply and continually upgrade our hardware. The up to date computing hardware, in turn, enables us to use ever more accurate and CPU intensive algorithms.

3.2 Data Reduction Software

The job of the data reduction processes was to remove false or uninteresting events while keeping a very high efficiency to retain classes of interesting events. Also, if there was some previously unknown class of events, the reduction should not remove them with undue prejudice. The reduction must be performed in a repeatable and reliable manner and fast enough to keep up with the input of raw data. All of the reduction (and reconstruction) software was essentially automatic and free of the systematics associated with using humans to make per-event decisions. That is, no event was rejected, saved or fit by an individual, only by an algorithm.

Most of the data reduction software was written by members of the Stony Brook group[1]. It followed four main steps, each of which was achieved by a separate program: `highe1`, `highc2`, `highc3` and `highc4`. Except for `highc3` these reduction processes were implemented as state machines. That is, during the reduction process each event exists in one particular state, it is evaluated against some conditions and then sent to another state based only on the knowledge of the current state. This allowed for flexible reduction algorithms to be developed but does lead to a somewhat complex description of the exact criteria an event must pass to be saved or rejected.

3.2.1 `highe1`

The first reduction process, `highe1`, reduced the raw data by about a factor of 30 in size and events. Besides the large reduction factor, this first level reduction attempts not to make assumptions about what a good event is, and so it will save an event if it fails to strongly categorize the event as bad.

The states in which an event would pass through are shown in Fig. 3.1. This figure does not show the exact criteria, these are shown below.

An event will be subject to the following criteria as it passed through the reduction. States are labeled as **STATE** and in each state a list of states to transition and the associated action or criteria are listed.

Begin goes to:

Save if time from preceding event is $< 100\mu s$,

Minimum Bias otherwise.

---

[1] A large percentage of this was done by C. McGrew.

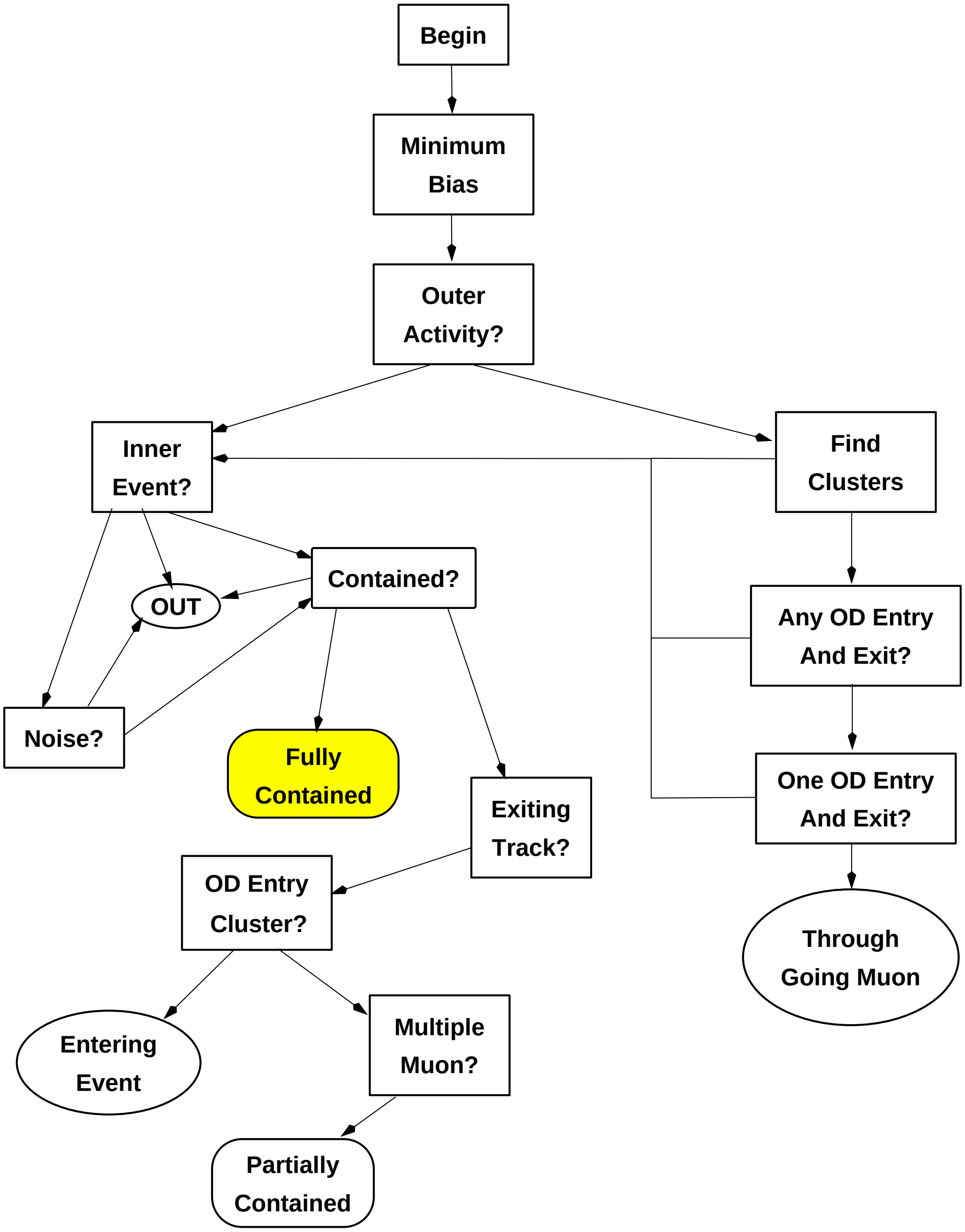


Figure 3.1: Map of `highe1` process.

Minimum Bias goes to:

Outer Activity after randomly flagging 0.1% of events.

Outer Activity goes to:

Inner Event if the # of hit OD PMTs in 10 ns window < 10,

Find Clusters otherwise.

Find Clusters goes to:

Any Entry Exit `clust()` returns > 1 OD cluster,

Inner Event otherwise.

Any Entry Exit goes to:

Inner Event if the # of hit ID PMTs < 1000,

One Entry Exit at least 2 OD clusters $>$ 10m apart are found and are within 50 ns of being consistent with a light-like separation,

Inner Event otherwise.

One Entry Exit goes to:

Thru Muon if any track joining the centroids of 2 OD clusters gives a sum of charge weighted timing residuals $<$ 5ns,

Inner Event otherwise.

Inner Event goes to:

Reject if the # of hit ID PMTs < 15 or $Q_{PMT,max}/Q_{tot} > 0.4$,

Noise if the # of hit ID PMTs < 120,

Contained otherwise.

Noise goes to:

Reject if the # of PMTs in time found by `octant()` < 15,

Contained otherwise.

Contained goes to:

Reject if the # of hit ID PMTs < 20

Reject if the charge in the 200 ns window with the most hit ID PMTs < 100 PE

Reject if the time from the previous event $< 100 \mu s$ and PMTs in-time in a 200 ns window < 20,

Reject if the charge of the PMTs in-time in a 20 ns window < 100 PE,

Save otherwise.

Exiting Track goes to:

Entry Cluster if the vertex is $<$ 6m from OD cluster,

Partially Contained after performing a `long_fit()` the goodness ¿ 50,

Entry Cluster otherwise.

Entry Cluster goes to:

Save (Entering) if the nearest OD cluster is within 6m and 50ns of entry point,

Partially Contained otherwise.

In the above, the routines `clust()`, `octant()` and `long_fit()` are mentioned. `clust()` finds clusters of PMTs temporally and spatially based on the amount of charge collected therein. It does this by treating the amount of charge in the hit PMTs similarly to elevations on a hill. It starts at the top of the hill and works down until it arrives at the bottom of the valley between hills. The algorithm starts with the PMT with the highest charge collected, called the seed, and then adds PMTs to the cluster if they pass the following criteria: (a) the PMT must be neighboring the seed, (b) the PMT must be within 50 ns of the seed and (c) the PMT must have charge less than $1 + 1.15 Q_{seed}$. Criterion (c) defines "downhill". It allows a neighboring PMT to have just slightly more charge than the seed in order to account for possible Poisson fluctuation in the number of collected PEs. If the PMT passes the criteria, it is removed from the pool of available PMTs and then has the clustering algorithm called with this new PMT in the rôle of the seed.

`octant()` attempts to find at least one point in the detector which has a significantly peaked PMT residual distribution. It does this by testing the point in the center of the detector and the 8 points in the middle of each octant around the center point. The point from these 9 points which contains the most PMTs in time is taken to be the new center and the algorithm is iterated taking the previous octant volume as the whole space. The routine terminates after a fixed number of iterations. This routine isn't a fitter, but if it finds some point in the detector from where light could have come and still be consistent with a large number of the hit PMTs' times, then it means that the event is not a noise event.

Finally, `long_fit()` attempts to find a position and direction of a Čerenkov light emitting track which is most consistent with the measured times of hit PMTs. For simplicity, it is assumed that the Čerenkov light is emitted at exactly the nominal Čerenkov angle. The fit is performed by minimizing a χ^2 built from the charge weighted difference in the time a PMT is measured to be hit and the time a PMT is expected to be hit.

3.2.2 `highc2`

After a large bulk of events which are known to be useless were pruned in `highe1` the next step in the reduction, `highc2`, is run to define a subsample of good events which are fully contained in the detector. This reduction step primarily consisted of removing flashers, entering cosmic ray muons and exiting partially contained events. Flashers are events which are caused by arcing in a PMT which is then detected by other PMTs. These events are usually very easily rejected due to their low energy, by examining either the timing of the hit PMTs or because the vertex is usually fit near the flashing PMT and thus the event was cut due to the fiducial volume cut in `highc2`. However, since flashers can occur with high frequency, many still get through these cuts. The pernicious ones, the so called "Danka" flashers,[2] were removed by examining distributions of various statistics to find clumps which could be cut away. As new types of flashers were found, different statistical criteria were needed to remove them. Although human intervention was needed to cut these events, the criteria (once chosen) were reproducible and their application was algorithmic.

The states in the `highc2` reduction process are shown in Fig. 3.2. As can be seen, this is a much more linear process. The details of this reduction are described in the same manner as the `highe1` process, above.

Begin goes to:

> **Reject** if pedestal event or is missing banks,

---

[2] I would like to thank Danka Kielczewska for providing me with the flasher list for this step, especially for the lists generated after onsite/offsite merger.

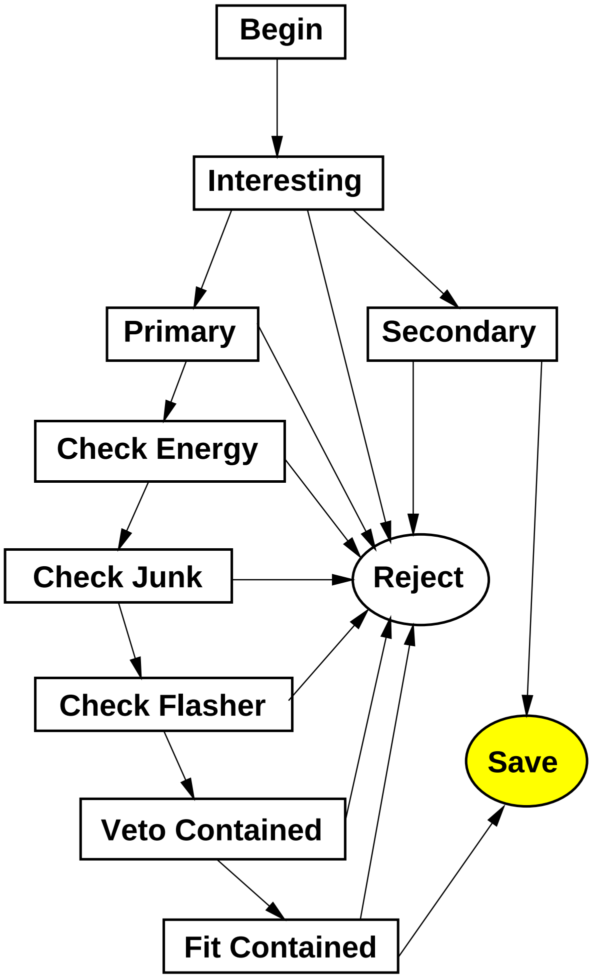


Figure 3.2: Map of `highc2` process.

Interesting otherwise.

Interesting goes to:

Secondary if time to previous event is within 100μs or `highe1` determined event to be a secondary event,

Primary if `highe1` determined event was a primary event,

Reject otherwise.

Secondary goes to:

Reject time to previous event is more than 100μs,

Save otherwise.

Primary goes to:

Reject if an OD trigger exists or if the ID or OD is dead or if pedestals are being taken in the event,

Check Energy otherwise.

Check Energy goes to:

Reject if the total positive charge $<$ 400 PE or the total negative charge $>$ 8 PE or the total negative charge $>$ 30% of the total positive charge,

Check Junk otherwise.

Check Junk goes to:

Reject if # in time ID PMT hits $<$ 65% of the # out of time ID PMT hits and # in time hits $<$ 2000,

Check Flasher goes to:

Reject if event is a "Danka" flasher,

Veto Contained otherwise.

Veto Contained goes to:

Reject if time to previous event is with in 100μs,

Fit Contained otherwise.

Fit Contained goes to:

- **Save** if # hit ID PMTs < 1000 or # ID PMTs with other neighboring PMTs which are hit close in time < 1000 of # hit ID PMTs with more than 2.5 PE < 500,
- **Reject** after performing a `long_fit()`, if more than 2 OD PMTs are found within 2m of the exit point,
- **Save** if `long_fit()` goodness > 200, the fit Z is $< +17$m or otherwise.

3.2.3 `highc3`

The third step, `highc3`, was essentially just running a high resolution vertex fitter and applying the fiducial volume cut. The vertex fitter was called Yet Another Single Track Event Fitter (`yastef`) [3]. All events which were fit further than 50 cm from the ID wall are kept.

For single ring atmospheric neutrino MC events [7], this fitter achieved a 40 cm resolution[4] and an uncorrected systematic shift in the direction of travel of +43 cm (-43 cm) for electron (muon) type events.

Using $p \to e^+\pi^0$ MC, `yastef` achieved a resolution of 25 cm with a shift of +30 cm from the true MC vertex along the fit direction. For $p \to \mu^+\pi^0$ MC, the resolution was 40 cm with about a 5 cm shift. The drastic difference in resolution is due to the fact that `yastef` would usually fit the strongest ring in the event and sometimes this is the ring due to a gamma from the π^0 decay and sometimes it is due to the muon. Thus the total resolution is a convolution of the resolution for fitting the muon in the event which has a positive offset and the resolution for fitting one of the gammas which have a negative offset.

3.2.4 `highc4`

The final step of the reduction, `highc4`, consisted of running a high precision suite of fitters based on a fitting engine called `mtfit`. This engine would estimate the amount of charge each PMT should collect given the parameters of the track or tracks assumed to be in the event. These parameters included

---

[3] Written for IMB and adapted to Super–Kamiokande by C. McGrew of Stony Brook (see Sec. 3.4.1)

[4] Since the distribution of the distances between fit vertices and MC vertices follows no simple distribution the resolution is taken to be half the shortest interval containing 68% of the events.

number and type of particles, their energy, vertex and direction. A likelihood of the given parameters would then be constructed by comparing the expected charge distribution to the measured charge distribution. A particular fitter distinguished itself by which parameters it varied as it attempts to maximize the likelihood. As before, the detailed path that an event may take is shown.

Begin goes to:

- **Process Contained** if `yastef` fit is closer than 1 m from the wall, or the total charge is less than 15000 PE or if a crude `mtfit` fit assuming the particle is a muon finds less than 50 expected hit PMTs,
- **Mtfit Muon Fine** if the `mtfit` vertex is less than 1.5 m from wall,
- **Mtfit Muon Back** otherwise.

Mtfit Muon Back goes to:

- **Mtfit Muon Fine** if after backing the vertex to the wall a fine `mtfit` fit assuming an entering muon finds less than 50 expected hit PMTs,
- **Process Contained** if this fit is worse than previous or puts vertex greater than 1.5 m from the wall,
- **Mtfit Muon Fine** otherwise.

Mtfit Muon Fine goes to:

- **Mtfit Elec Fine** if after running a fine `mtfit` fit assuming a muon event, less than 50 expected hit PMTs are found or the resulting vertex is less than 1.5 m from the wall,
- **Processes Contained** otherwise.

Mtfit Elec Fine goes to:

- **Processed Contained** if a fine `mtfit` fit assuming an electron finds less than 50 expected hit PMTs,
- **Reject** if the best `mtfit` so far finds a vertex which is within 1.5 m from the wall,
- **Process Contained** otherwise.

Process Contained goes to:

- **Reject** if the `yastef` vertex was within 1 m of the wall,

Possible Flasher after running `rispid`, the particle identification routine.

Possible Flasher goes to:

Reject if the charge weighted RMS residuals in a 20 ns window is more than 6 ns,

Save otherwise.

3.3 Sub-run Cuts and Livetime

Besides cutting individual events, data collected during the reduction was used to impose sub-run cuts. Sub-runs are a logical splittings of a run. Initially they contained about 10 minutes of runtime, and after the institution of a lowered trigger threshold (SLE) and accompanying online intelligent trigger (SKIT) this was reduced to 7 minutes. By looking at various sub-run based statistics, it was possible to remove segments of livetime which contained large amounts of obviously bad data.

Sub-runs were cut if they substantially deviated from expected runtime. Short sub-runs were removed as they were either due to aborted runs or were just the last sub-run in a run which usually contained a lot of junk events. The very long sub-runs (comprising less than 0.05% of the sub-runs) had wildly high runtimes, likely due to garbled data or other failure of the online system.

Cuts based on the ratio of the number of events per sub-run with high ID charge (> 16000 PE) and low OD hits (< 30) to high ID charge and high OD hits removes some flasher rich sub-runs as does simply cutting on the ID event rate.

From these cuts, just over a quarter million sub-runs survived from run 1681 to run 6721. This data spans from 1996/05/24 to 1998/12/30 during which the detector was alive for an integrated total of 784.9 days.

3.4 Reconstruction Software

The reconstruction software was needed to pull kinematic and geometric parameters out of just the distribution of hit PMT times and charges. It started during the reduction with the finding of the event vertex. From there, effects of light attenuation in water was removed, then a measure of the total energy and other simple parameters of the event were found. These were enough to significantly reduce the background to proton decay. From there, more detailed reconstruction was performed. This found such parameters as number of particles and their individual momenta and particle types.

3.4.1 Vertex Fitting

The `yastef` fitter starts at a rough initial vertex and direction. This rough vertex is found by calculating the local gradient in the timing residuals and jumping in the downhill direction. As the gradients decrease, so does the size of the jump. The search terminates after the jump size becomes sufficiently small or a fixed number of jumps have been taken. From this vertex the direction is estimated from the anisotropy. The anisotropy is defined as the charge weighted vector sum of the directions from the vertex to each PMT and has a magnitude which is bound by 0 and 1. It gives a rough measure of the momentum imbalance of the event, and is used later in the proton decay selection criteria.

Next, the direction which maximizes the amount of charge in a 60° opening angle cone is found, followed by movement along the track direction to bring the opening angle to the nominal Čerenkov angle and movement perpendicular to the track directions to reduce the timing residuals.

Finally, the direction is adjusted to give a ring profile (as viewed along the angle from the track direction) which is as sharp as possible, while again moving along and perpendicular to the track direction as in the previous step.

3.4.2 μ Decay Electron Counting

If an event had any muons present, either above or below Čerenkov threshold, there would be decay electrons present assuming the μ is not captured by the oxygen. When the decay electrons are energetic enough they are good tags for the presence of muons. These electrons could be found in three time regions: inside the 1 μs event window along with the light from the primary particle(s), outside the primary event window in it's own event, or in the time region just before the end of the primary event and just after the start of the next. This latter region was not well defined as different events would have different window widths. If a decay electron was in this region it tended to have a smaller efficiency for being recovered as it could have it's light split between two events.[5].

---

[5] Such events are dubbed "bye-bye events"

To find the decay electrons which created light in the same event as that from primary particles, a routine called `samedk`[6] was used. This found decay electrons by ordering the PMT times (subtracted by time of flight from the `yastef` vertex) and searching for groupings in time. The PMTs found in each group were then sent to a timing based low energy vertex fitter called `hayai`[7].

To find decay electrons which landed in their own event the routine `afterdk`[8] was used. This routine also worked by fitting a possible decay electron event with `hayai`. All PMTs which were part of a good `hayai` vertex were removed, the remainder were again fit by `hayai` and the process was repeated until no more PMTs remain.

Finally, any decay electrons which were possibly split between two events were taken care of so that they were counted only once.

3.4.3 Charge Correction

After the location of the event was found, the effects of Čerenkov light attenuation due to the water and PMT angular acceptance were removed. For this, a simple model of the water transparency was used, namely that a beam of light would be attenuated exponentially with path length and independent of wavelength and ignoring light scatter. This assumption allows one to write the expected charge for i^{th} PMT,

$$q_i = N_i a(\hat{r}_i) e^{-r_i/L}, \tag{3.1}$$

where N_i is the number of Čerenkov photons which leave the vertex and travel towards the i^{th} PMT. The function $a(\hat{r}_i)$ is the angular acceptance of the i^{th} PMT. Finally, r_i is the distance from the vertex to the i^{th} PMT and L is a characteristic *attenuation length* (more on this later).

Then the flux of Čerenkov photons can be written as,

$$\frac{dN_i}{d\Omega} = \frac{q_i e^{r_i/L}}{a(\hat{r}_i)} \frac{1}{A_\circ / r_i^2} \tag{3.2}$$

where $A_\circ$ is the cross sectional area of a PMT and $d\Omega = d(\cos\theta)d\phi$. Assuming that this flux is constant near the i^{th} PMT the total number of Čerenkov

---

[6] Written by E. Kearns and M. Messier of BU

[7] Written by J. Hsu of UCI and B. Svoboda of LSU

[8] Written by J. Flanagan of UH

photons can be estimated with,

$$N_{tot} = \int_{4\pi} \frac{dN}{d\Omega} d\Omega = \sum_i \frac{q_i e^{r_i/L}}{a(\hat{r}_i)} \frac{A_\square / r_i^2}{A_\circ / r_i^2} \cos(\theta_i), \tag{3.3}$$

where $A_\square$ is the area around the i^{th} PMT, (the 70 cm × 70 cm lattice square) and θ_i is the angle between the direction from PMT to the vertex and the direction the PMT faces.

From this, a *corrected charge* is defined,

$$Q_{cor} = \sum_{\substack{\text{intime}\\ \text{PMTs}}} \frac{q_i e^{r_i/L}}{a(\hat{r}_i)} \cos(\theta_i), \tag{3.4}$$

as being an estimator of the total number of Čerenkov photons produced in the event. The area parameters have been dropped as they just give an over all multiplicative constant. In order to remove the complication of the small amount of scattered light, the sum is restricted to those PMTs which are hit by *intime* or direct light. That is, only PMTs with times that are within a 20 ns window around the time expected for light emanating from the vertex were used.

By moving from raw charge values to vertex dependent corrected charge the effects of the water and the PMT geometry could be removed. This had the great benefit of giving a uniform detector response for various vertex locations and track directions. These two corrections are now described in more detail.

PMT angular acceptance

The function $a(\hat{r}_i)$ in eq. 3.1 gives the likelihood of a photon hitting a PMT to be detected when coming from a direction of $\hat{r}$ relative to the direction normal to the PMT. This likelihood is a mix of cross sectional shape, quantum efficiency and glass transparency. In addition, a correction at high angles of incidence was included to account for the average shadowing by neighboring PMTs. The angular acceptance correction as a function of angle between the direction normal to the PMT face and the direction from PMT to vertex is shown in Fig. 3.3. The angular acceptance was found by the calibration group and the tube shadowing was found with MC given the physical tube shape, location and quantum efficiencies.

Light Attenuation Length

The correction for light attenuation in water was the most important correction to the charge as it allowed uniform detector response independent of

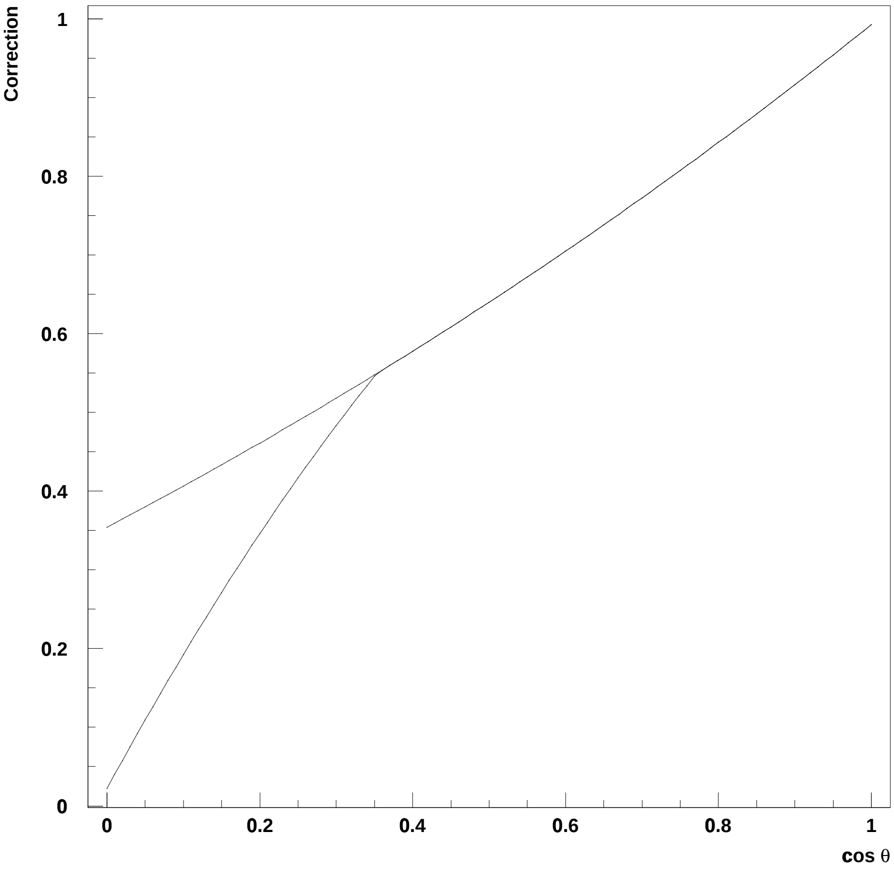


Figure 3.3: Angular acceptance correction. The lower branch is the sum of the angular acceptance and the PMT shadowing correction. The correction is a function of the angle of incidence of the photon assuming the light came for the fit event vertex.

vertex position and track direction. In general the total attenuation of a beam of light at a particular wavelength is the mix of absorption and scattering. The number of photons of a particular wavelength in a beam of light can be written as

$$N(x,\lambda) = N(0,\lambda)e^{-x/L_{tot}(\lambda)}, \tag{3.5}$$

where the total attenuation length as a function of wavelength λ is given by

$$\frac{1}{L_{tot}(\lambda)} = \frac{1}{L_{abs}(\lambda)} + \frac{1}{L_{scat}(\lambda)}, \tag{3.6}$$

and x is the distance between light source and detector.

The attenuation length measured in the Super–Kamiokande tank at various wavelengths is shown in Table 3.1. These measurements were performed by feeding a laser pulse through an optical fiber into the tank and out a diffuser ball. The ball was imaged by a CCD and monitored by a PMT. The PMT reference was necessary as the laser intensity would vary over time. The relative CCD response was measured as a function of depth of the diffuser ball. The resulting light collection, after correction for $1/r^2$ effects, was fit to an exponential function of light path distance. The inverse of the exponential fit parameter is the attenuation length.

Wavelength (nm)	Attenuation length (m)
337	67.1
400	103.1
500	34.25
580	10.32

Table 3.1: Directly Measured Attenuation Lengths.

This measurement gives an ideal attenuation length and one which is difficult and time consuming to measure. While this method gives values which are needed by the MC simulation it is not practical for tracking water quality over time as it necessitates detector dead time. Furthermore, when correcting an event, it is impossible to know the wavelength of each photon of Čerenkov light which hits each PMT, so an averaged or effective attenuation length is needed in any case.

This effective attenuation length was found by using fit[9] cosmic ray muons which enter and stop in the tank as a source of calibration light. These events

---

[9] The muons were fit using the routine `muboy` by R. Svoboda

are plentiful and are recorded continuously while the detector is on. Figure 3.4 shows an example of such an event. This display shows the unrolled detector. Each square represents a hit PMT with size and shade indicating the amount of charge. The two large circles represent the top and bottom of the ID, the rectangle represents the wall. The inset is the same for the OD. Except for an occasional bremsstrahlung these events provide a uniform line source of light which statistically covers the entire volume of the detector providing a sample of light path lengths from a few meters to just under 40 m.

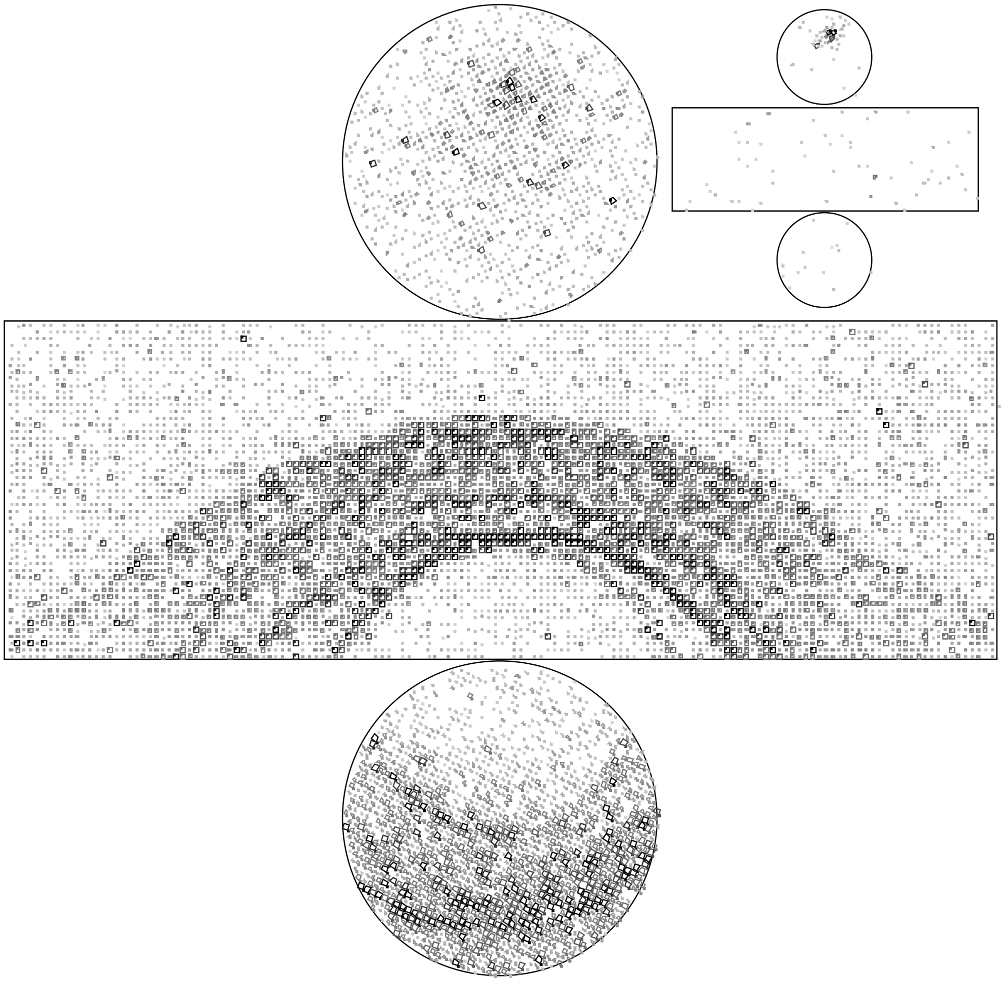


Figure 3.4: Stopping muon event display.

There are a few problems with using stopping cosmic ray muons for probing the attenuation length. They are generally more energetic than the contained events so tend to be large events, data-wise and are dropped early on in the reduction. To harvest these events, one must return to the raw data. PMT saturation can also be a problem, so PMTs which are too close to the muon track can not be used. Finally, the worse problem is that the exact attenuation length one achieves is sensitive to the exact cuts used to select a sample of muons. This last problem is turned into an advantage and it is actually exploited as shown below.

Using MC, it is possible to find the best attenuation length to use in correcting the charge by simply generating mono-energetic MC events throughout the detector and correcting their charges assuming different attenuation length. The length which causes the events to be reconstructed with uniform corrected charge is then the best for the version of the MC water parameters. This choice is shown in Fig. 3.5. With this ideal attenuation correction, the exact cuts used to find the attenuation length with cosmic ray muons are tuned to reproduce this best attenuation length correction. This method was tested with versions of the MC with different input attenuation lengths and it was found that it correctly adjusts for dirtier or cleaner water.

Attenuation length measurements for the MC and the data is shown in Fig. 3.6. Since the muon track provides a line source of light the $\frac{1}{r}$ dependence must be removed in order to see the less dramatic but more interesting exponential dependence. This is done by simply multiplying the charge collected at each distance from the fit track (labeled as DFT in the figure) by the DFT. The "slope" of this log plot gives the inverse of the attenuation length. The "Y-intercept" gives a value which is the relative gain of the detector and is used to normalize the total light output of the detector over time. The fit is done for light which travels between 10 m and 35 m. Light outside this range suffers from low statistics, PMT saturation, or problems due to PMT geometry idealizations.

To track the attenuation length over time this measurement is performed periodically on about a days worth of stopping cosmic ray muons. Since it is time consuming to make this measurement it is only done every few hundred runs which corresponds to about every two months. To find values between measurements a simple second order interpolation is used. The evolution of the attenuation length and the normalization is shown in Fig. 3.7. The upper and lower curves give the uncertainties which are dominated by systematics. As can be seen, in addition to the systematics of each measurement, the normalization has a slight systematic dependence on the attenuation length. This is due to the attenuation curves not being completely exponential. This deviation from non exponential is due to PMT saturation and slight non uniformity of the light source.

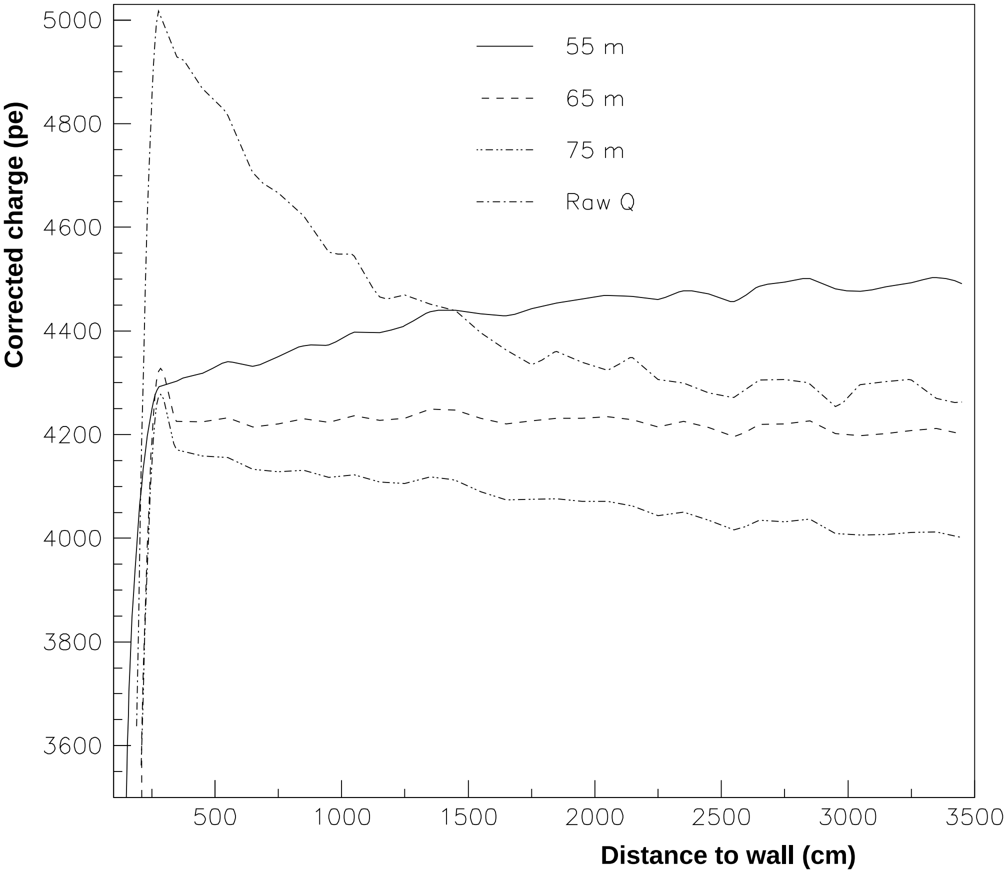


Figure 3.5: Corrected charge of 500 MeV MC electrons assuming (dash-dot) no correction, and attenuation correction with lengths of (solid) 55 m, (dashed) 65 m and (dash-dot-dot) 75 m.

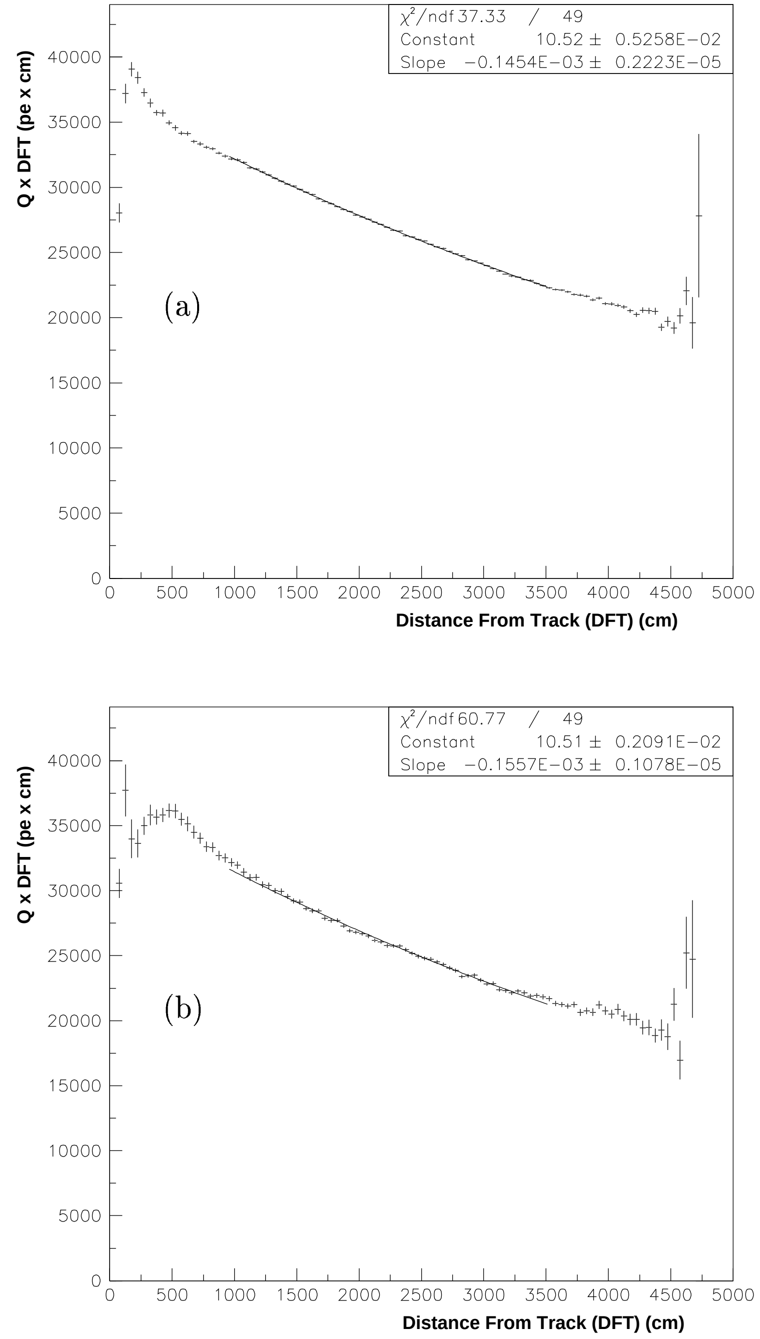


Figure 3.6: Effective attenuation length and gain normalization measurements for (a) data and (b) MC for a particular run.

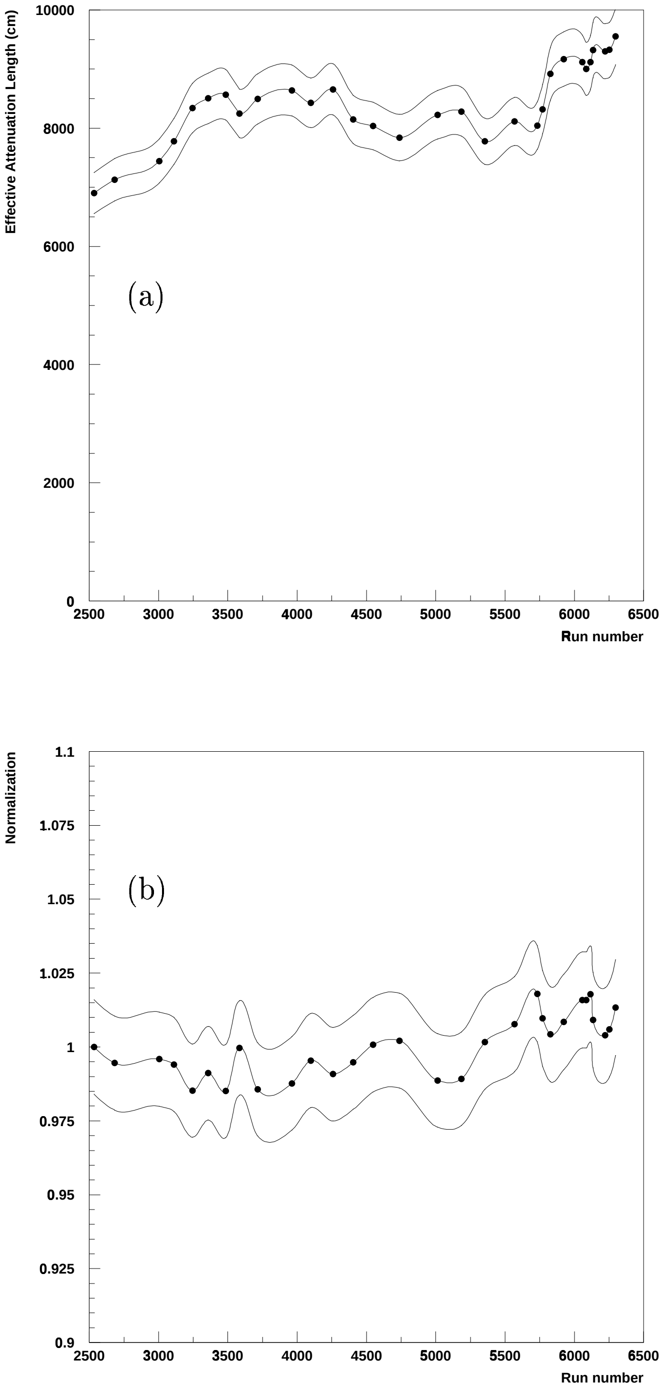


Figure 3.7: Time dependence of (a) effective attenuation length and (b) gain normalization in the data as a function of run.

3.4.4 Visible Energy (E_{vis})

From Eq. 2.2 it can be seen that, after integrating over wavelength, the number of emitted Čerenkov photons per cm of track is dependent only on the speed of the particle. Particles spend most of their time moving at a speed of $\beta \approx 1$. Muons at typically observed momenta are minimum ionizing, loosing about 1.8 MeV/cm to 2.2 MeV/cm. [29]. Electrons will cause an electromagnetic shower, but each electron and positron of the shower will also loose energy due to ionization at a roughly constant rate. Since the number of Čerenkov photons per cm and the energy loss per cm are both approximately constant, the number of Čerenkov photons is a good estimator of the energy deposited by the particle when the particle is ultra relativistic.

When the particle slows to $\beta < 1$ the number of Čerenkov photons per cm decreases, but as the particle slows, it is no longer minimum ionizing and thus looses energy more quickly. This means that the particle is either going at speed $\beta \approx 1$ or it is slower than needed to be above Čerenkov threshold and there is very little visible track while the particle is non-ultra relativistic. The total energy of the particle can then be estimated by the number of Čerenkov photons plus a relatively smaller constant energy to account for when the particle is below threshold. The former energy is called *Visible Energy* (E_{vis}).

E_{vis} is defined more precisely: The Visible Energy of an event is the energy an electron would need in order to produce the same amount of corrected charge as in the event. This definition is useful as it gives a measure of the energy which is independent of details of the event such as number of and type of particles in the event.

Without a source of mono-energetic electrons of ~GeV energies, calculation of E_{vis} with this definition must rely on MC simulations. Using MC which has had its light attenuation tuned to a particular detector time period, sets of mono energetic electrons were generated at various energies. To connect corrected charge to visible energy, the mean corrected charge of each set was plotted against the input electron energy as shown in Fig. 3.8.

The relationship is, as expected, very linear. The very small non-linearity around 300 MeV is accommodated by fitting the curve with two linear functions joined at 300 MeV. The non-linearity is due to a changing probability for a photon (coming from an electron bremsstrahlung) to pair produce. This probability increases to near one after about 100 MeV. So as more, higher energy, photons are produced there is a higher probability for all the energy of the electron to be visible instead of being lost to invisible low energy photons. The fit gives 8.33 PE per MeV below 300 MeV and 8.65 PE per MeV above. The energy resolution is $49.5\%\sqrt{E(\text{MeV})}$ with a 1.3% constant which

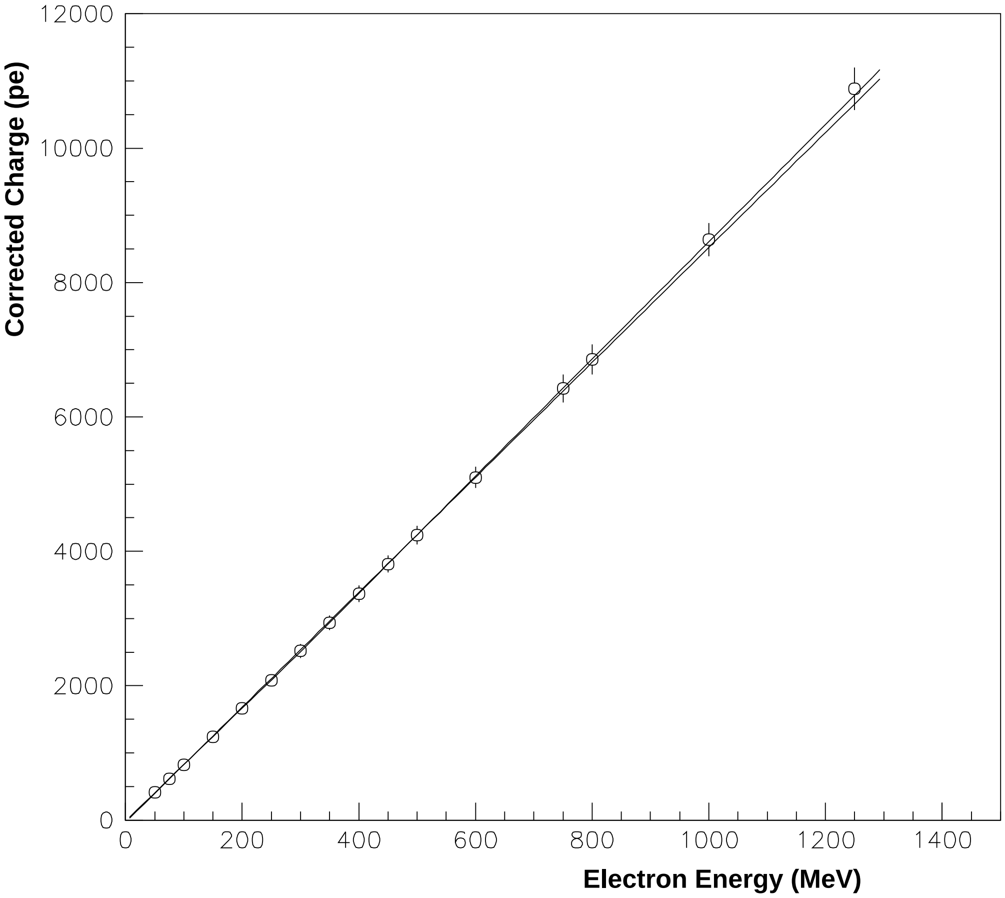


Figure 3.8: Corrected charge *vs.* total energy of MC electrons.

translates to about 3% resolution in the 0.5 - 1 GeV range.

For single electron events, the energy of the electron is precisely the visible energy. To find the energy of single muons events, the visible energy of MC muons is tabulated against their total energy and a lookup table is made. To find energies of individual particles in a multi-ring event the event undergoes *charge separation* described in Sec. 3.4.5. The per track energy (and other reconstruction parameters) is then determined as if the track was an isolated single ring event.

3.4.5 Multi-track Fitting

The bulk of atmospheric neutrino analysis makes use of single ring events as they are a very pure sample of CC neutrino interactions and allow one to infer the flavor of the interacting neutrino. In contrast, multi-ring events are expected in most proton decay searches, particularly those studied here. In order to study multi-ring events a suite of multi-track fitters was developed.

Multi-track fitting was done in several stages. Starting with a fit vertex from `yastef` (described in section 3.4.1) a set of potential track directions were found with a program called `rifle` (standing for RIng FLaging Engine[10]). Then, any potential track directions which were actually likely due to the same ring were collected with a program called simply `collect`. With this smaller set of directions the event underwent charge separation with a program called `share_flux`. This was mainly done to give a good starting point to the next step which was running a precise multi-track fitter called `mtendir` which was based on the same `mtfit` engine used in the reduction. All tracks which the `mtendir` found to be likely track directions were again submitted to `share_flux` for final charge separation. The last step was to send the surviving tracks to `mtpid` which found the particle type of the individual tracks. Each of these codes are now described in turn.

`rifle`

The basic method of rifle is to apply a Hough transform to the *charge flux* which is a value defined on a directional space, namely the directions from the fit vertex to each PMT. The charge flux is just the measured corrected charge per solid angle in each PMT and is defined as a vector pointing toward each PMT. A Hough transform in general is a mapping from a physical space to a space of parameters which model a desired shape or feature which is to

---

[10] Also because it was used to kill `MOOSE`, a previous initial track finder.

be found in real space. The mapping is the association of the value at each point in physical space with the set of points in parameter space which could contribute to the point in real space.

For `rifle`, the physical space is the set of directions from the vertex to the PMTs. The shape to look for is the ring made by the intersection of the Čerenkov cone with the unit sphere. To parameterize this ring, `rifle` makes several passes, each time assuming a different Čerenkov cone opening angle. In each pass specifying a direction is enough to give a unique ring.

The (directional) Hough space is built up from a set of discrete bins using a geodesic sphere of ∼2600 sides. The transform is done by filling each Hough direction bin with the (sometimes weighted) sum of all charge fluxes which could have arisen by a particle going in that bins direction. The resulting Hough space will then show peaks in directions which correspond to track directions in physical space.

An example of the physical charge flux values and Hough peaks from an event is shown in Fig. 3.9a and Fig. 3.9b. As can be seen, the Hough peaks are rather wide and noisy. In order to enhance the peaks and reduce noise a second Hough transform is done. The second transform will only let a flux contribute to the Hough bin which in the first transform was the highest bin to which the flux contributed. The result of such a "cleaning" transform is shown in Fig. 3.9c.

After the two Hough transforms, `rifle` will find the peaks in Hough space by clustering the bins. Clustering is done in much the same manner as `clust` described in Sec. 3.2.1. The height of each cluster gives a rough estimate of the strength of the corresponding ring.

`collect`

Usually, `rifle` will find more directions than there are particles (and will also sometime miss directions if it corresponds to a low energy particle). This is due to PMTs hit by noise or scattered light or because of "split" rings. In the case of noise and scattered light, there can be PMTs which have signals not due to light emanating from the vertex in an idealized cone. When the Hough transform is applied, these PMTs will still contribute to building the likelihood of directions, and can cause false peaks and smear true peaks in Hough space. A more arduous feature is that of "splitting" rings. Often a particle will undergo scattering or otherwise have fluctuations in the Čerenkov light output. While the overall ring is due to one particle (or shower) the ring will have substructure which will show up in the Hough space of `rifle` as two or more nearby peaks.

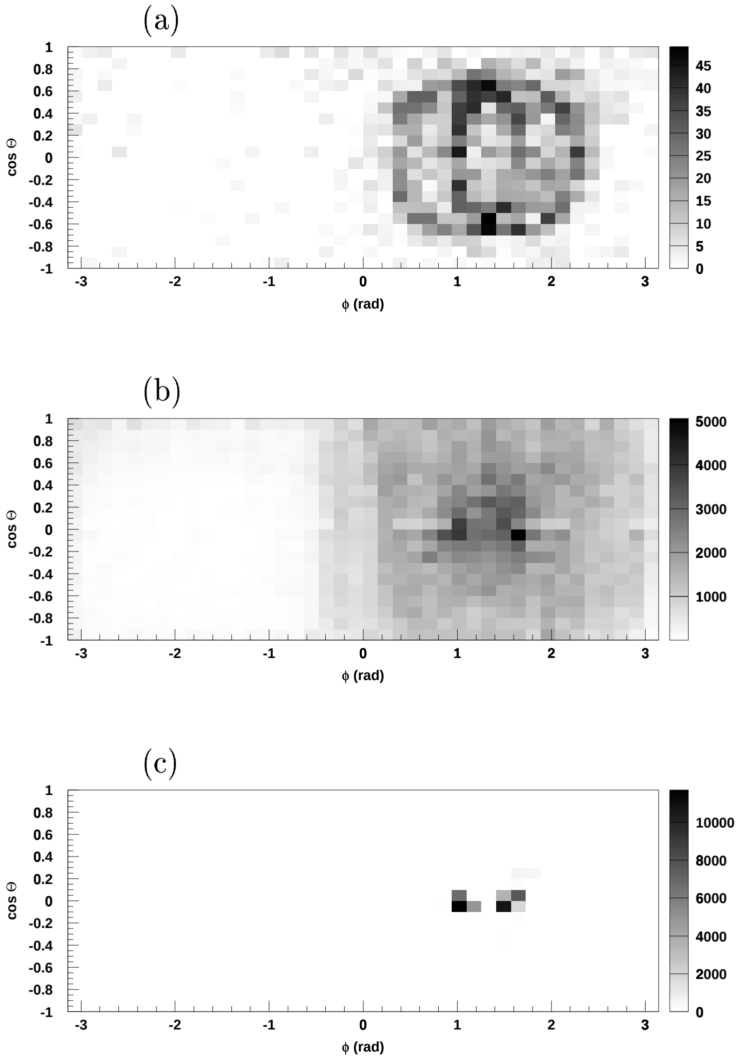


Figure 3.9: Charge of a two ring event plotted in $\cos\theta$ *vs.* ϕ in (a) real directional space, (b) Hough directional space after first transform and (c) Hough directional space after a second "cleaning" transform.

Since `rifle` makes no assumption about the physical particles in the event (other than the fixed Čerenkov cone angle in each pass) it is difficult to reduce the extra false directions without imposing assumptions on the event. This imposition is kept out of `rifle` and put into `collect`. In this program, all tracks which are within 20° are collected as one. Also, any peak which has a value in Hough space of less than 0.01 are dropped.

`share_flux`

The job of `share_flux`[11] is charge separation. That is, it finds the amount of charge in each PMT which is due to each assumed track. It does this by exploiting the fact that, ignoring fluctuations, the Čerenkov light should be azimuthally symmetric.

In `share_flux`, each given track direction has associated with it a set of *track bins*. A track bin is an annulus of solid angle centered around the track direction. A set of track bins cover the hemisphere of solid angle in the track direction and will, in general, overlap some or all of the track bins of other tracks.

The fraction of corrected charge per solid angle (the charge flux), $f_{i,\mu}$ in PMT i due to track μ is then estimated by,

$$f_{i,\mu} = \frac{\sum_{j \neq i} f_{j,\mu} q_j}{\sum_\nu \sum_{j \neq i} f_{j,\nu} q_j} \tag{3.7}$$

where q_i is the total corrected charge per solid angle in PMT i and each sum over j is limited to the PMTs contained in the track bin containing PMT i. In other words, the fraction of charge in PMT i from track μ is estimated by the ratio of the total charge from track μ (not counting the charge of PMT i) which is in the track bin containing PMT i to the sum of this same value for all tracks which have a track bin containing PMT i.

Since $f_{i,\mu}$ for a particular PMT and track is defined in terms of other $f_{i,\mu}$, Eq. 3.7 is a set of coupled equations and is solved iteratively starting out with the assumption that each PMTs charge is equally due to all tracks.

`mtendir`

The fitter `mtendir` is based on the `mtfit` engine which is described in Sec. 3.2.4. Given an event vertex and a set of possible track directions with

---

[11] The routine `share_flux` was originally written for the IMB experiment and adapted to Super-Kamiokande by C. McGrew and subsequently modified for this thesis.

initial relative track strengths from a previous run of `share_flux` (described above) `mtendir` will adjust track directions and track strengths in order to maximize the `mtfit` likelihood. Tracks which do not correspond to any ring will end up having a much reduced track strength and any of these tracks have a relative strength less than 0.1 will be cut during the subsequent and second running of `share_flux`. This removes most of the remaining false directions.

`mtpid`

Starting with the sets of $f_{i,\mu}$ found in the second running of `share_flux`, the program `mtpid` applies single ring particle identification (PID) to each track in an event. The PID was developed specifically for the offsite analysis and while the resulting distributions (see below) for the, say, sample of atmospheric neutrinos on the surface appear similar to the onsite results one should take note that the sign of the parameter is reversed.

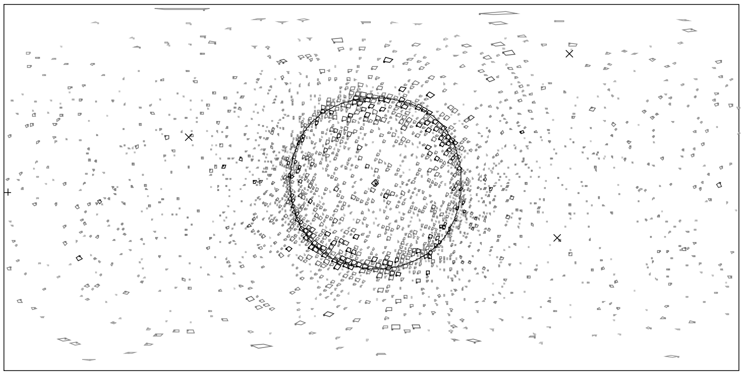
The PID is found by the routine `rispid`[12] [30] which exploits the fact that heavy non-showering particles (μ and $\pi^{\pm}$) produce rings with sharp outer edges while light showering particles ($e^{\pm}$ and γ) produce rings with a fuzzy outer edge. The differences in these two types of rings are shown in Fig. 3.10.

To quantify fuzziness, `rispid` projects the charge on to the track along the nominal Čerenkov cone. This projection is illustrated in Fig. 3.11. Next, the point which maximizes the first derivative is found and considered to be the ring edge. The ratio of the amount of projected charge in the neighborhood just outside the ring edge to the amount in the neighborhood just inside the ring edge is calculated. Since muons have very little charge outside the ring edge small ratio values are found, whereas electrons tend to have closer to equal amounts of charge just inside and just outside the ring edge and correspondingly larger ratio values are found.

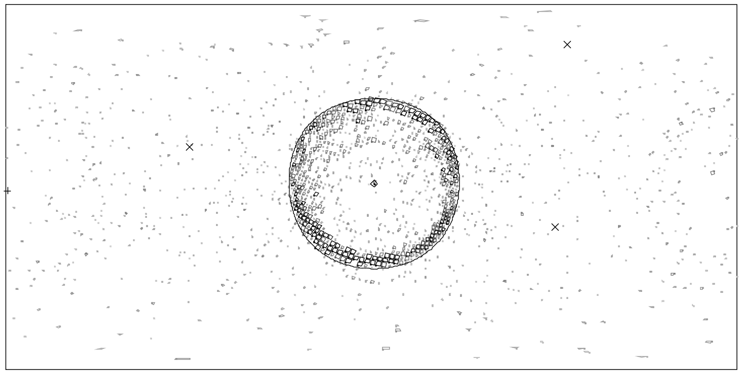
The value `rispid` finds tends to be somewhat energy dependent and awkward due to having a lower bound of zero and no upper bound. To remedy this a simple adjustment is made by a routine called `adpid` which maps `rispid` values to values between -1 and 1 and removes the energy dependence. These `adpid` values for various classes of events are shown in Fig. 3.12.

---

[12] RISing edge Particle IDentification



3.10(a)



3.10(b)

Figure 3.10: Comparison of (a) an electron like or showering ring and (b) a muon like of non-showering ring. Plots show hit PMTs distributed in $\cos\theta$ *vs.* ϕ with shade and size indicating amount of charge collected.

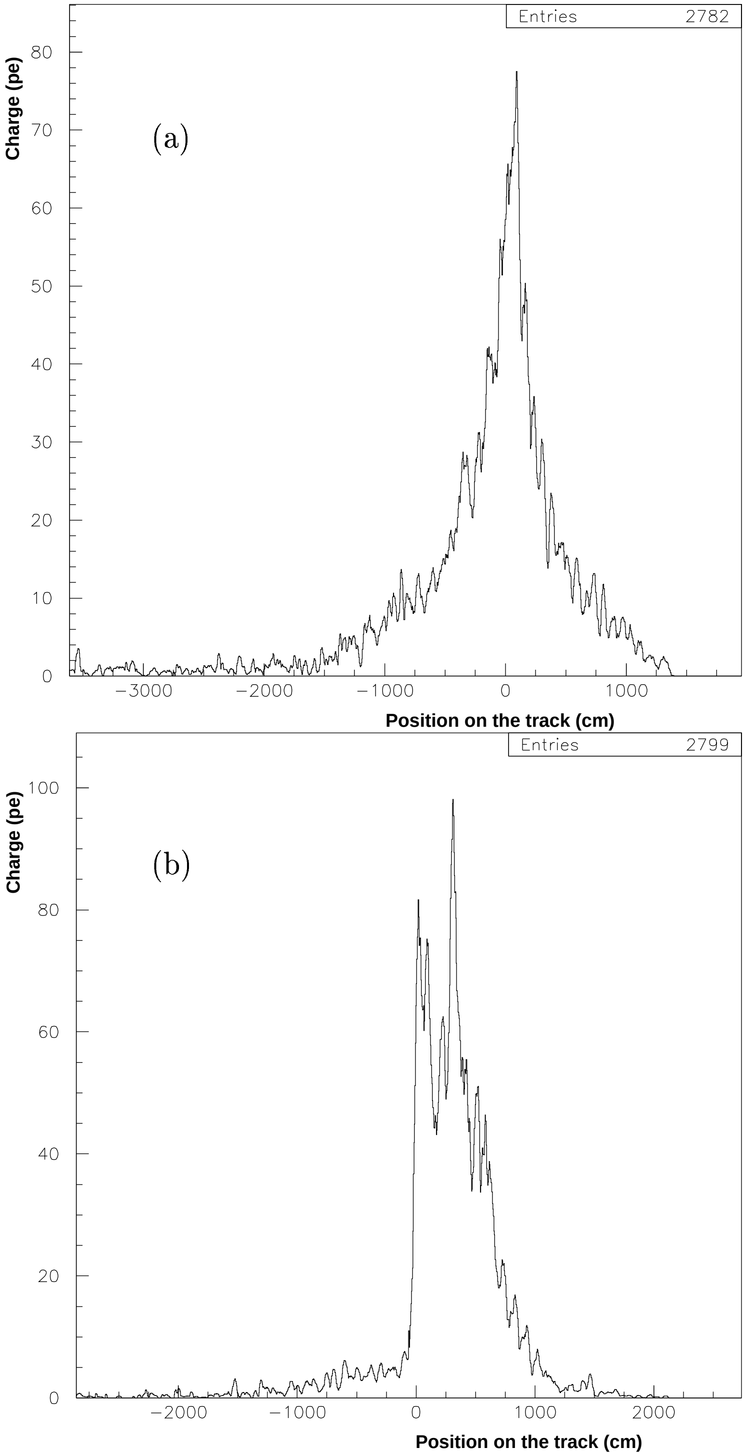


Figure 3.11: Projection of charge onto the particle track for (a) an electron like and (b) a muon like particle.

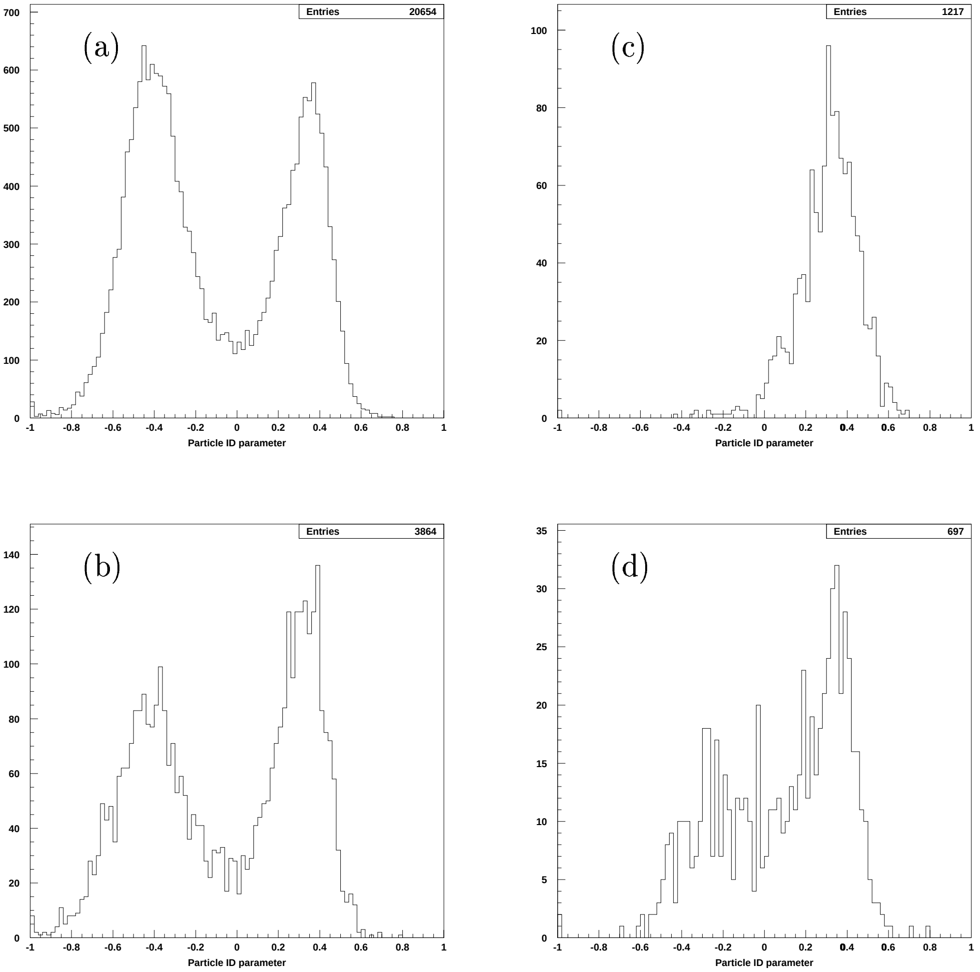


Figure 3.12: Per track adpid values for (a) single ring atmospheric neutrino MC, (b) single ring data, and (c) $p \to e^+\pi^0$ MC and (d) $p \to \mu^+\pi^0$ MC both passing visible energy and anisotropy cuts. Negative values indicate muon-like rings, positive values indicate electron-like rings.

Chapter 4

Signals, Background and Data

This chapter describes the method to separate the desired proton decay signals from the background as well as the Monte Carlo simulations of each. Cuts on the values from the reconstruction algorithms of the previous chapter are used to implement this separation.

4.1 General Method

The search for proton decay (or any other search for that matter) follows a few basic steps.

First, each decay mode must be considered for event characteristics which make it unique enough to differentiate it from any possible background.

Second, these characteristics are exploited by devising selection criteria or cuts to separate out any possible proton decay products from the background. These criteria are then optimized to give the highest efficiency while still rejecting background. This optimization can be a formal maximization of some metric based on the ratio of the number of signal events to the square root number of background events as a function of a set of parametrized cuts, or it can be done "by eye" by looking at event distributions as a function of different event parameters and choosing a simple separation criteria. In this thesis all selection criteria are determined with this latter "by eye" method.[1] Of course whichever method is taken, information contained in the actual data can never be used to determine the cuts.

---

[1] Indeed, a more formal method was tried and it was found that in the region of parameter space which optimizes the selection the metric maximized was very broad and flat and encompassed the region which would have been chosen "by eye".

The third and final step is to apply the selection criteria to an independent set of MC and data events in order to evaluate the selection criteria. This independence is of course very important. If one optimizes the selection criteria based on specific sample events it is possible to produce selection criteria which removes all the background and keeps all the signal. However it would be applicable only for the specific events used in the optimization. This independence, while adhered to in this thesis, would not be a large problem due to the rather simple cuts needed to achieve a high efficiency and low background.

For most possible nucleon decay modes accessible in Super–Kamiokande there are two main features which allow separation of this signal from the background of atmospheric neutrino induced events. These features are from the fact that the products of proton decay come from a parent which is essentially at rest and of a fixed mass, while products of atmospheric neutrino events have a substantial net momentum and a broad spectrum of invariant mass. One minor caveat which complicates this picture is that if the decaying nucleon is in the oxygen nucleus of water it is not truly at rest because of Fermi motion. The Fermi momentum is limited to below about 300 MeV/c. And in this range there is very little power in the atmospheric neutrino energy spectrum. The Fermi momentum used when simulating proton decay events is shown in Fig. 4.1. Specifics to the two modes studied in this thesis are discussed below.

In this thesis, the separation of signal and background proceeds in two stages. The first stage make use of simple and loose cuts which exploit general event parameters, while the second make more stringent cuts based on more explicit event parameters. The main reason for this two stage selection is that to make strong cuts on such event parameters as total event momentum, one needs to run CPU intensive fitters demanding long computing times (see Sec. 3.4.5).

4.1.1 Free Proton Decay

Since the pions from protons decaying in oxygen can be absorbed or scattered in the nucleus, thus creating complications in the event selection criteria, it is useful to search for decays in just the free protons of the hydrogen. While this removes 80 % of the protons from consideration, those that remain will be recovered with a much higher efficiency which allows tighter cuts to be applied and an accompanying stronger background rejection. If any candidates are found in this free proton decay search, the evidence for proton decay would be stronger than candidates found in a full search. For these reasons a sub search which considers free protons only was also performed.

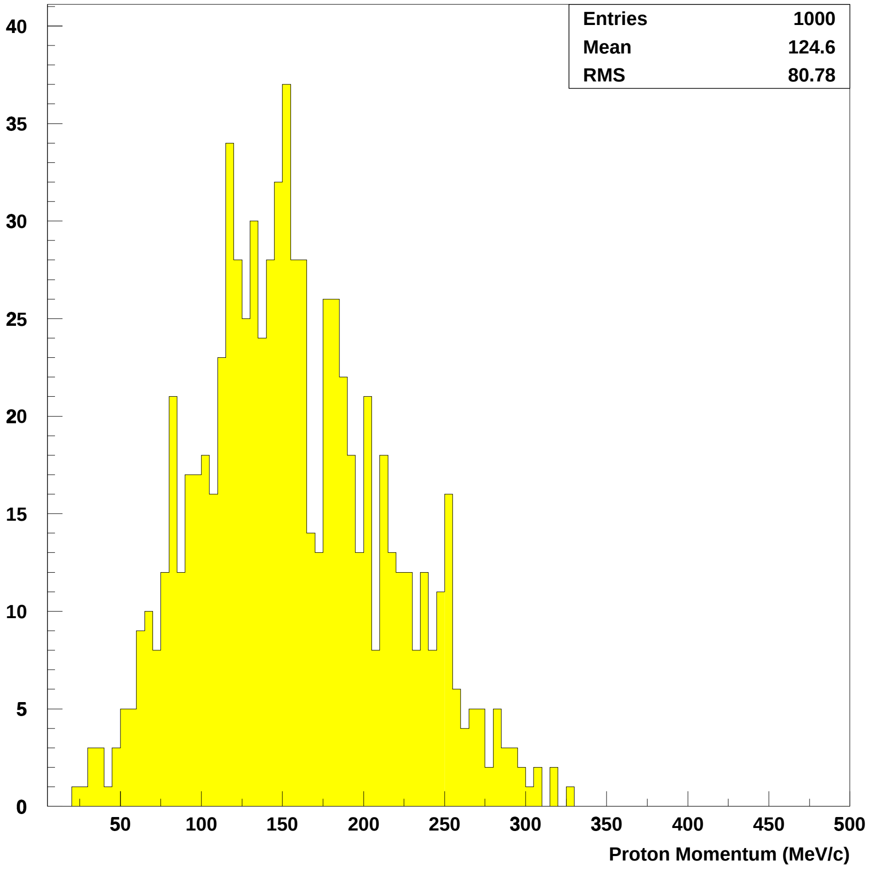


Figure 4.1: Fermi momentum used in proton decay simulations.

4.2 First Stage Event Selections

The first stage of event selection used E_{vis}, anisotropy and the number of electrons from μ decay. These three simple parameters are relatively unaffected by the systematic uncertainties which may accompany complicated and sensitive fitter based cuts. Furthermore, the cuts were chosen conservatively enough to reduce any systematics due to E_{vis} calibration or uncorrected differences in the tuning of the MC water transparency parameters. Using these parameters which do not demand intensive computing, a large fraction of background events are removed before the application of more detailed, computationally demanding fitters.

The visible energy spectrum for $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ modes as well as atmospheric neutrinos (in the range of interest) are shown in Fig. 4.2. All of the decaying proton's energy is visible for the former mode. In the latter mode, the muon goes below Čerenkov threshold with about 160 MeV total energy.

The anisotropy for the mode containing the e^+ should be very low as the visible energy of this lepton is balanced with that of the γs from the decay of the π^0. For the other mode, some of the energy of the μ^+ is not visibly manifest and is carried below Čerenkov threshold. This causes the event to be less balanced and thus with a higher anisotropy. The anisotropy for these two modes as well as atmospheric neutrinos is shown in Fig. 4.3.

Finally we can easily count the expected decay electrons for the two modes. Namely we expect none for $p \rightarrow e^+\pi^0$ and one from the μ^+ in $p \rightarrow \mu^+\pi^0$. The distributions of the number of decay electrons in these modes as well as atmospheric neutrino events are shown in Fig. 4.4. The higher deviation from the expected number of decay electrons is due to electrons from the decay of muons from the decay of charged pions which are produced via charge exchange of the π^0 in the nucleus. The lower deviation from the expected number of decay electrons is due to just the inefficiency of the algorithms which attempt to find decay electrons.

The cuts for the first stage selection for $p \rightarrow e^+\pi^0$ events are itemized below. The values in parenthesis are for the versions of the search where only free protons are considered.

1. No decay electrons. (same)
2. 740 MeV $< E_{vis} <$ 1140 MeV (860 MeV $< E_{vis} <$ 1050 MeV)
3. Anisotropy $<$ 30% (Anisotropy $<$ 16%)

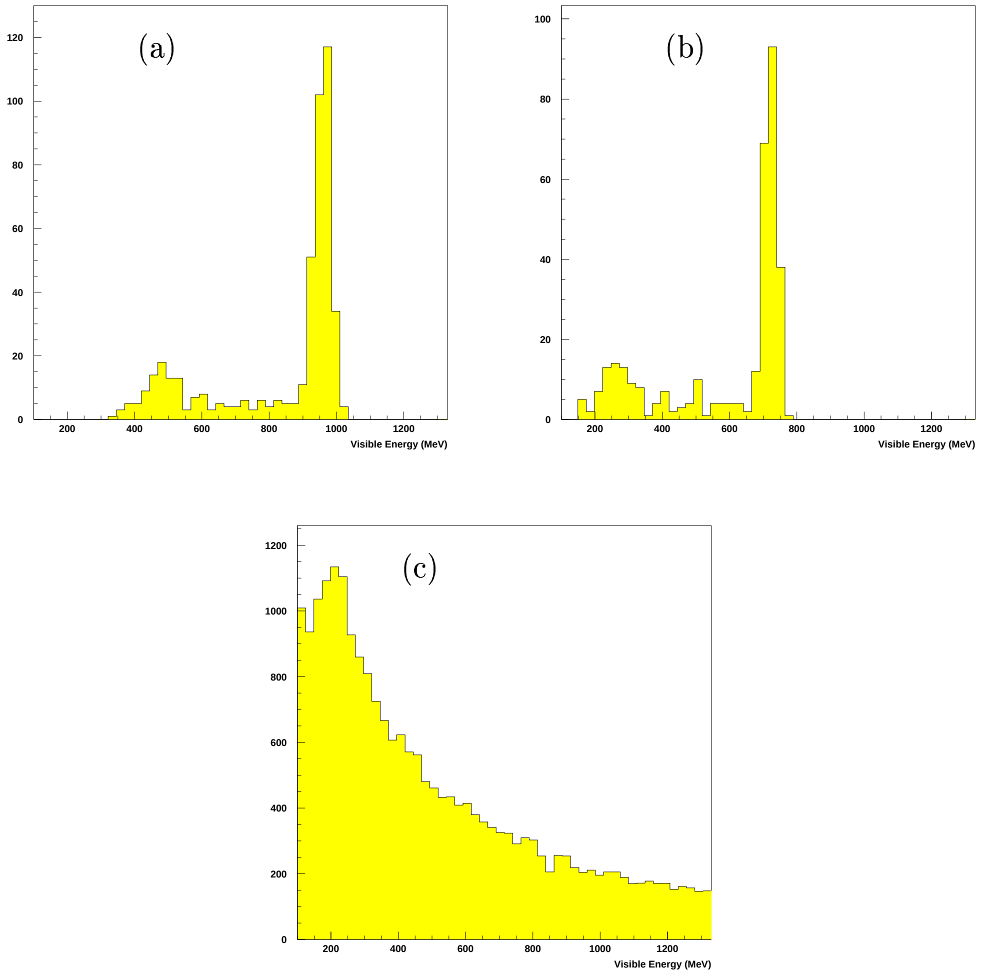


Figure 4.2: Visible Energy distributions for (a) $p \rightarrow e^+\pi^0$, (b) $p \rightarrow \mu^+\pi^0$ proton decay and (c) atmospheric neutrino Monte Carlo simulations.

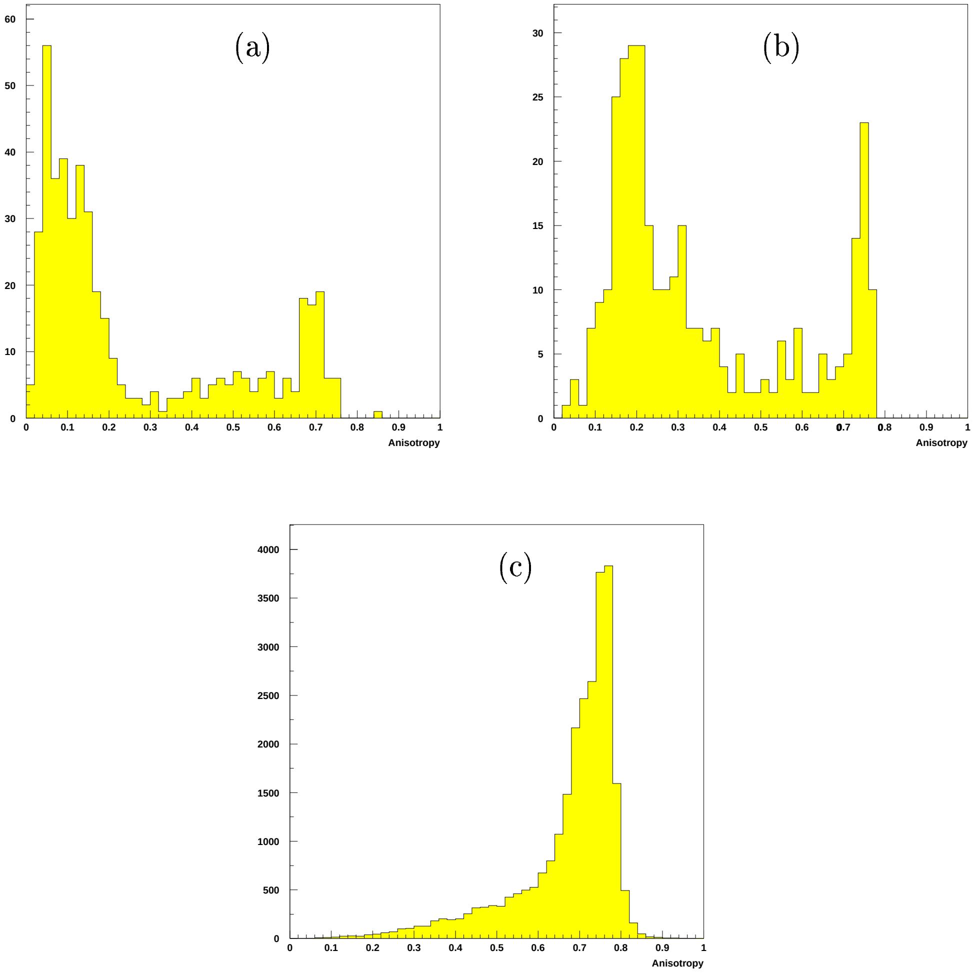


Figure 4.3: Anisotropy distributions for (a) $p \rightarrow e^+\pi^0$, (b) $p \rightarrow \mu^+\pi^0$ and (c) atmospheric neutrino Monte Carlo simulations.

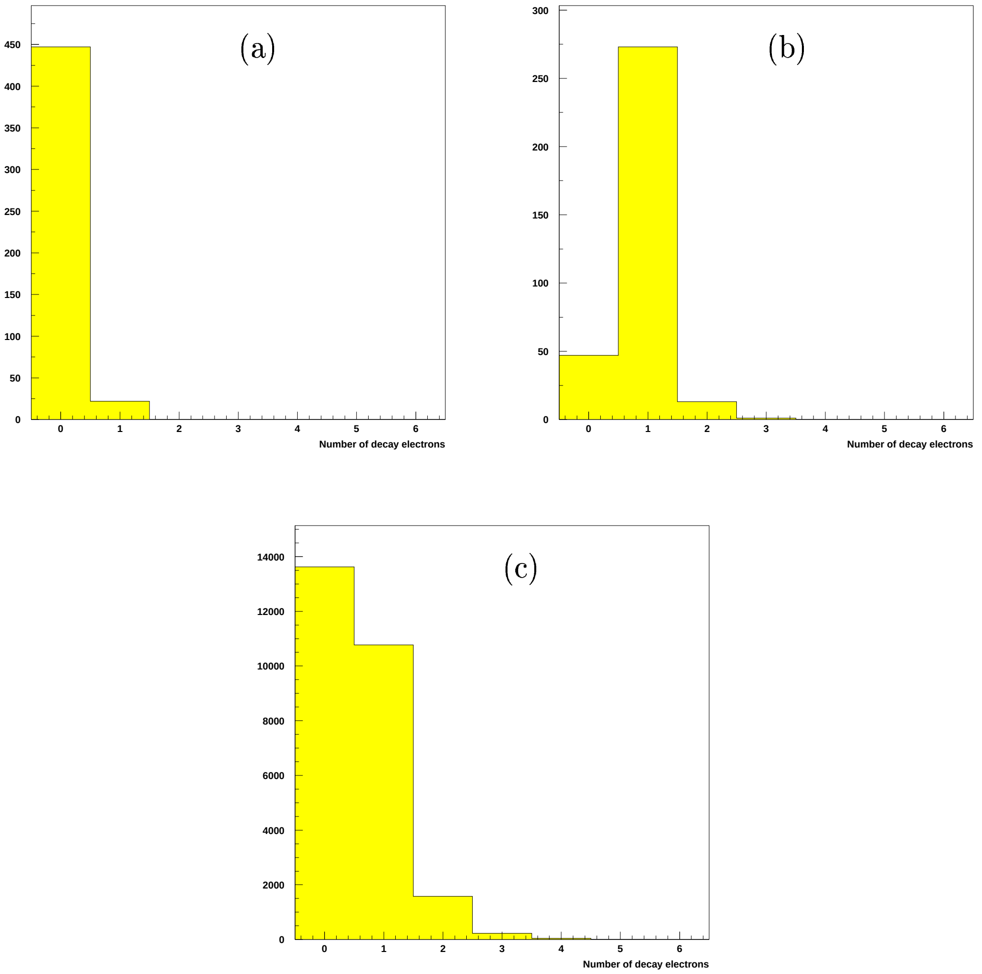


Figure 4.4: Number of decay electron distributions for (a) $p \rightarrow e^+\pi^0$ and (b) $p \rightarrow \mu^+\pi^0$ proton decay MC and (c) atmospheric neutrino MC

While those for $p \rightarrow \mu^+\pi^0$ are:

1. One decay electron (same)
2. 640 MeV $< E_{vis} <$ 840 MeV (660 MeV $< E_{vis} <$ 800 MeV)
3. Anisotropy $< 40\%$ ($8\% <$ Anisotropy $< 25\%$)

These cuts for the full search are illustrated in Fig. 4.5a and Fig. 4.6a with $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ MC events, respectively. In addition Fig. 4.5b and Fig. 4.6b show the result of these cuts on atmospheric neutrino induced MC events.

On average, the $p \rightarrow \mu^+\pi^0$ MC events have lower visible energy than the $p \rightarrow e^+\pi^0$ events due to the heavier muon. The events in both these samples show two clumps with and a connecting sparse tail. The events inside the cuts are primarily ones where the π^0 escaped the oxygen nucleus without any interaction. The broadening along the anisotropy direction is due to Fermi motion. In the $p \rightarrow e^+\pi^0$ case there is a cluster of events at anisotropy near 70%, typical for single ring electron events, while in the $p \rightarrow \mu^+\pi^0$ case the cluster is near 75%, typical for single ring muon events. These two clusters contain events in which the π^0 is absorbed in the nucleus. The connecting tail contains events in which the π^0 scattered (upper branch) or charge exchanged (lower branch).

In the case of the atmospheric neutrino MC event samples there is significant differences between the plots corresponding to each search. This is because these plots show the effect of the decay electron selection criterion. In the case of the $p \rightarrow e^+\pi^0$ selection, there are three features. The long band at an anisotropy of 70% is enriched with single ring events from CC electron neutrino interaction. The band at a visible energy of 200 MeV contains π^0 events from NC interactions. The last bin near zero energy (which is below analysis threshold) contains flasher events and low energy atmospheric neutrino induced events. After the $p \rightarrow \mu^+\pi^0$ decay electron selection criterion essentially all single ring electron events and single π^0 events are removed, leaving a single band enriched with single ring events from CC muon neutrino interactions. In both cases there is a broad and sparse field of events near the anisotropy and visible energy cuts (represented by the box). These are a mix of higher energy CC and NC atmospheric neutrino induced multi ring events with enough momentum transfered to make one or more extra pion.

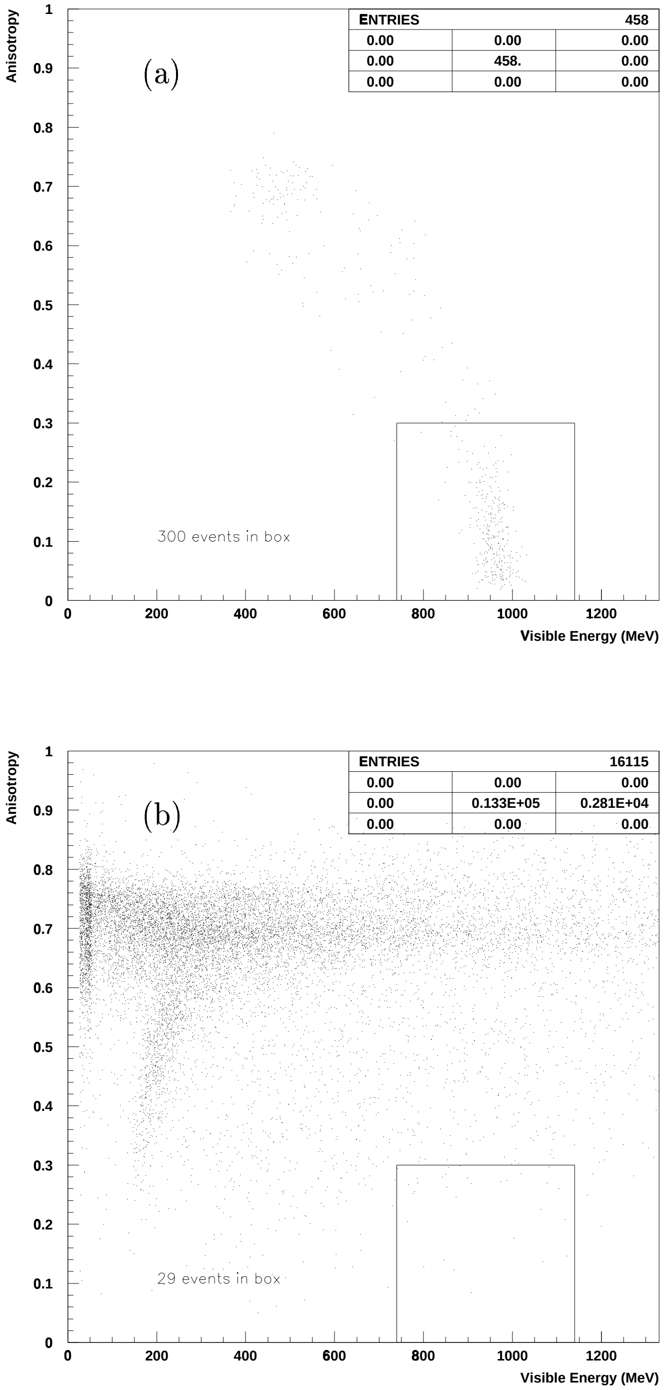


Figure 4.5: Anisotropy and Visible Energy distributions of (a) $p \rightarrow e^{+}\pi^{0}$ and (b) atmospheric neutrino MC events. In both figures events with out any decay electrons are shown. The box represents the cuts on anisotropy and visible energy

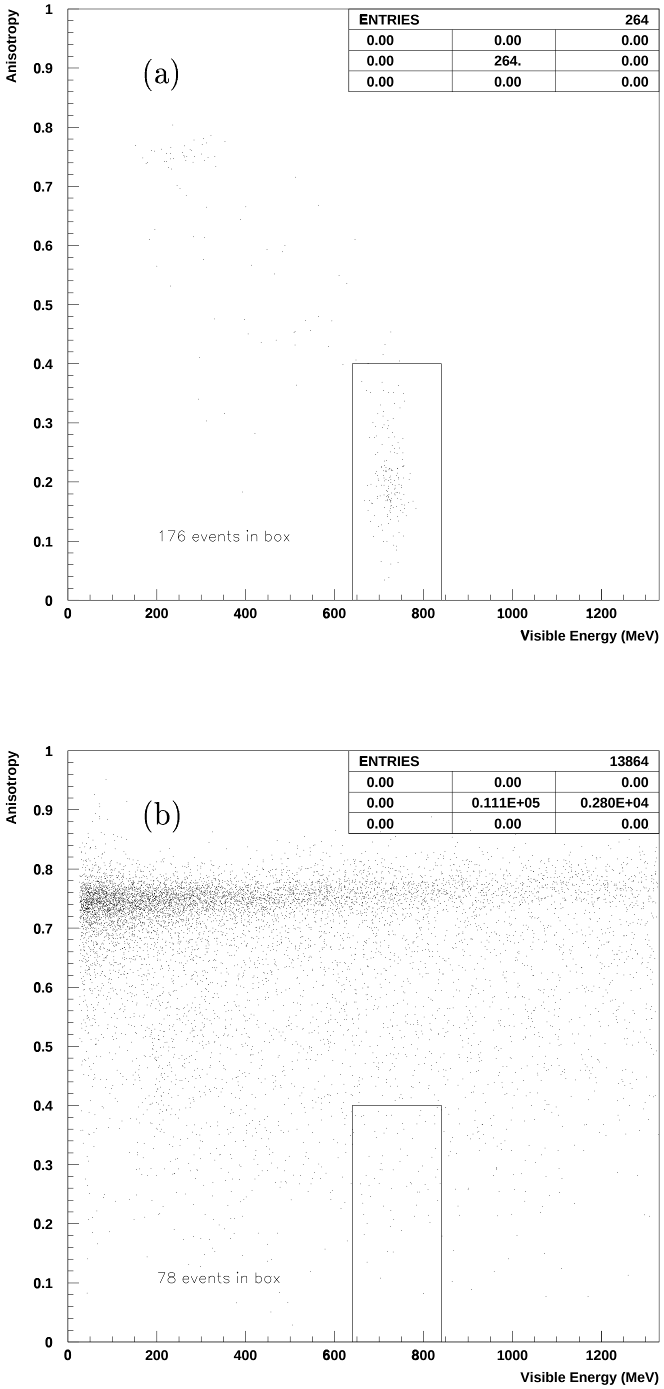


Figure 4.6: Anisotropy and Visible Energy distributions of (a) $p \rightarrow \mu^+\pi^0$ and (b) atmospheric neutrino MC events. In both figures events with exactly one decay electron are shown. The box represents the cuts on anisotropy and visible energy

4.3 Second Stage Event Selections

The second stage event selection makes use of information based on attempting to find all particle tracks in the event. Once all possible tracks are found via multi-track fitting (Sec. 3.4.5) and charge separation has been performed, it is possible to treat each track in an event as a separate single ring event. Using particle identification (PID) and track energy it is possible to find the total momentum and invariant mass of the event which provides a stronger background rejection than just anisotropy and visible energy.

The first selection criterion utilizes the number of showering (electron-like) tracks and the number of non-showering (muon-like) tracks in the event. These quantities are shown in Fig. 4.7 for events passing first stage $p \rightarrow e^+\pi^0$ cuts and in Fig. 4.8 for those passing first stage $p \rightarrow \mu^+\pi^0$ cuts.

These plots show what is expected, namely that the lepton and one or both of the photons from the π^0 decay are reconstructed in the proton decay MC events. The selection criteria at this step is to require two or three electron-like and no muon-like tracks for the $p \rightarrow e^+\pi^0$ search while one or two electron like and one muon like track is required for the $p \rightarrow \mu^+\pi^0$ search.

The final cuts in the second stage make use of reconstructed total momentum and invariant mass. If a particle is found to be showering, then it is taken to be massless. If the particle is non-showering then the muon mass (105.66 MeV/c^2) is assumed. The distributions of reconstructed total momentum and invariant mass are shown in Fig. 4.9 and Fig. 4.10 for events passing prior $p \rightarrow e^+\pi^0$ cuts and in Fig. 4.11 and Fig. 4.12 for events passing prior $p \rightarrow \mu^+\pi^0$ cuts.

The cuts for the second stage selection for $p \rightarrow e^+\pi^0$ events are itemized below. As above, the values in parenthesis are for the free searches.

1. 2 or 3 electron like rings and no muon like rings (same)
2. $P_{tot} < 300$ MeV/c ($P_{tot} < 200$ MeV/c)
3. 900 MeV/c^2 $< M_{inv} <$ 1000 MeV/c^2 (same)

and those for $p \rightarrow \mu^+\pi^0$ are:

1. 1 or 2 electron like rings and one muon like rings (same)
2. $P_{tot} < 300$ MeV/c ($P_{tot} < 200$ MeV/c)
3. 900 MeV/c^2 $< M_{inv} <$ 1000 MeV/c^2 (same)

These last two cuts are shown in Fig. 4.13 and Fig. 4.14 for the $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ searches, respectively.

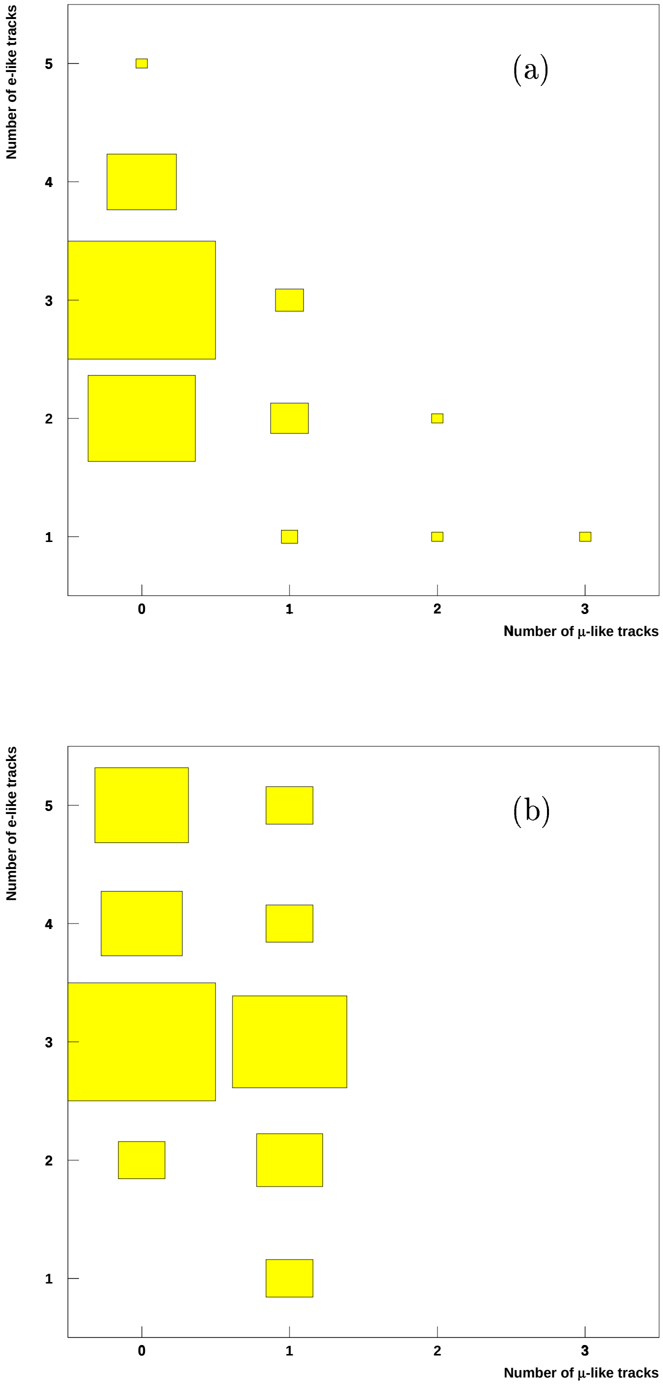


Figure 4.7: Number of showering (e-like) *vs.* non-showering (μ-like) tracks per event for (a) $p \rightarrow e^+\pi^0$ MC and (b) atmospheric neutrino MC events passing first stage $p \rightarrow e^+\pi^0$ cuts.

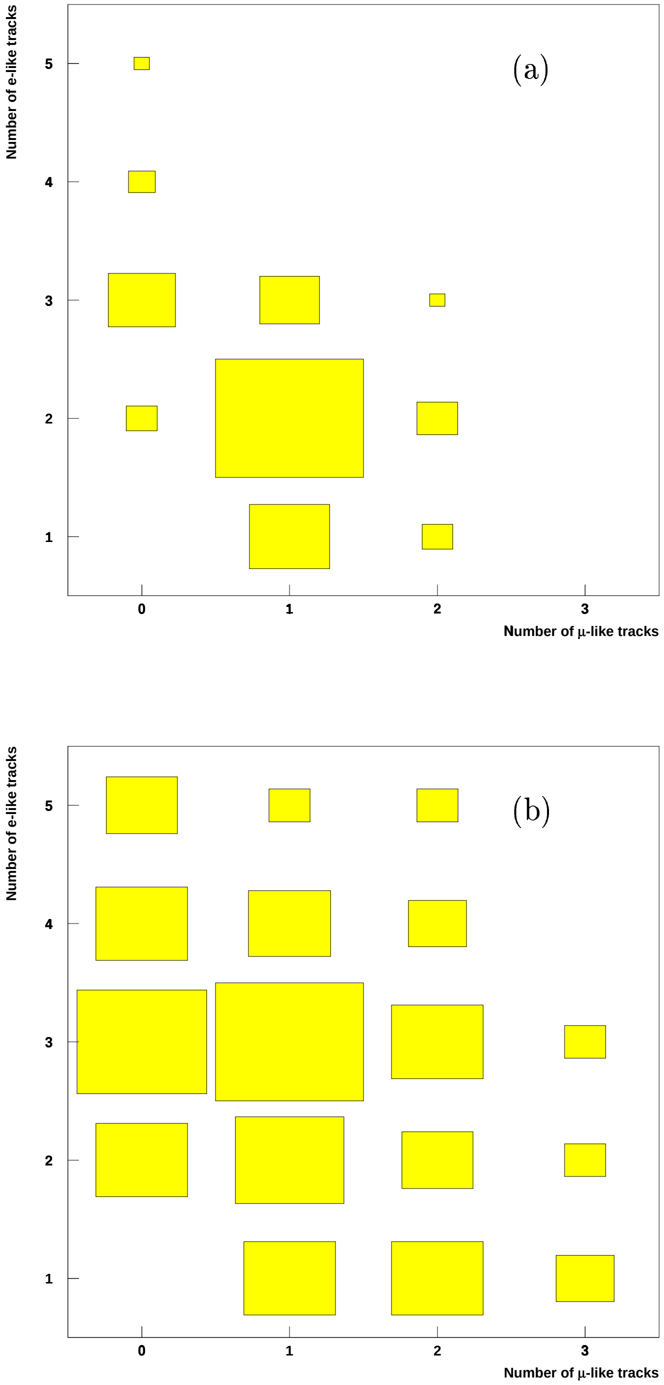


Figure 4.8: Number of showering (e-like) *vs.* non-showering (μ-like) tracks per event for (a) $p \rightarrow \mu^+\pi^0$ MC and (b) atmospheric neutrino MC events passing first stage $p \rightarrow \mu^+\pi^0$ cuts.

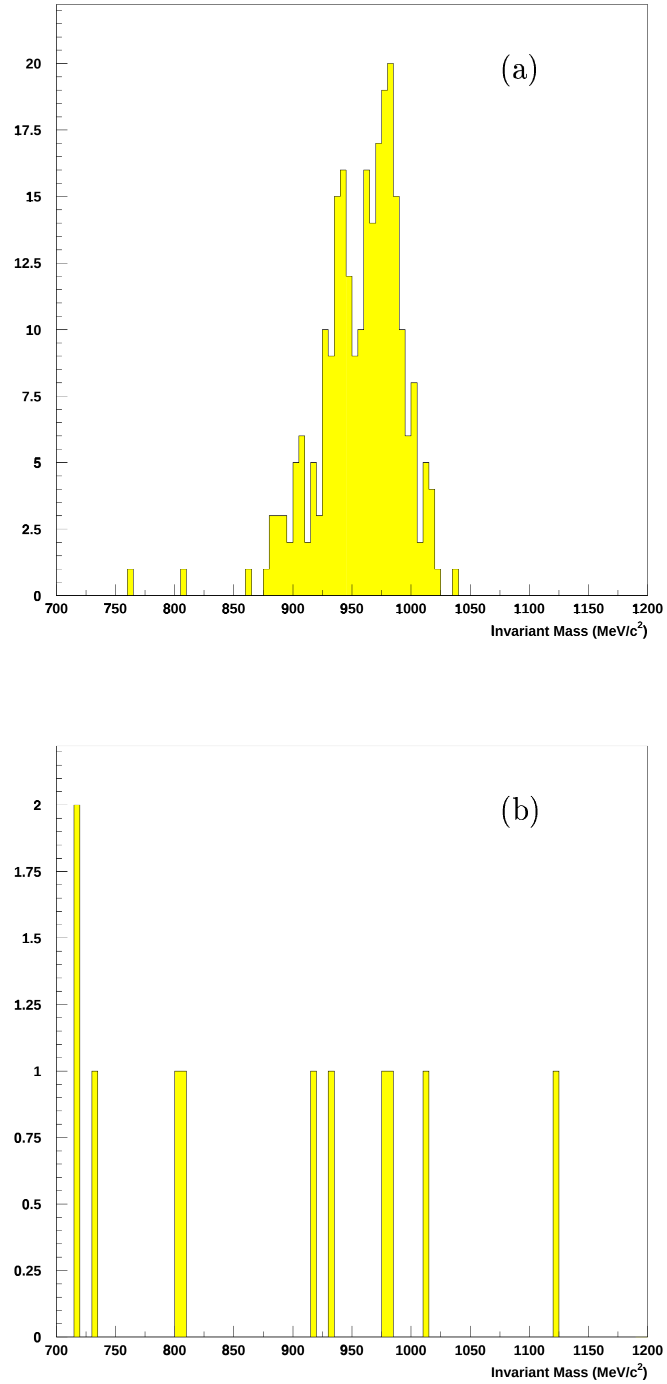


Figure 4.9: Distributions of reconstructed invariant mass for (a) $p \rightarrow e^+\pi^0$ MC events and (b) atmospheric neutrino MC events which have passed first stage cuts as well as particle identification cuts.

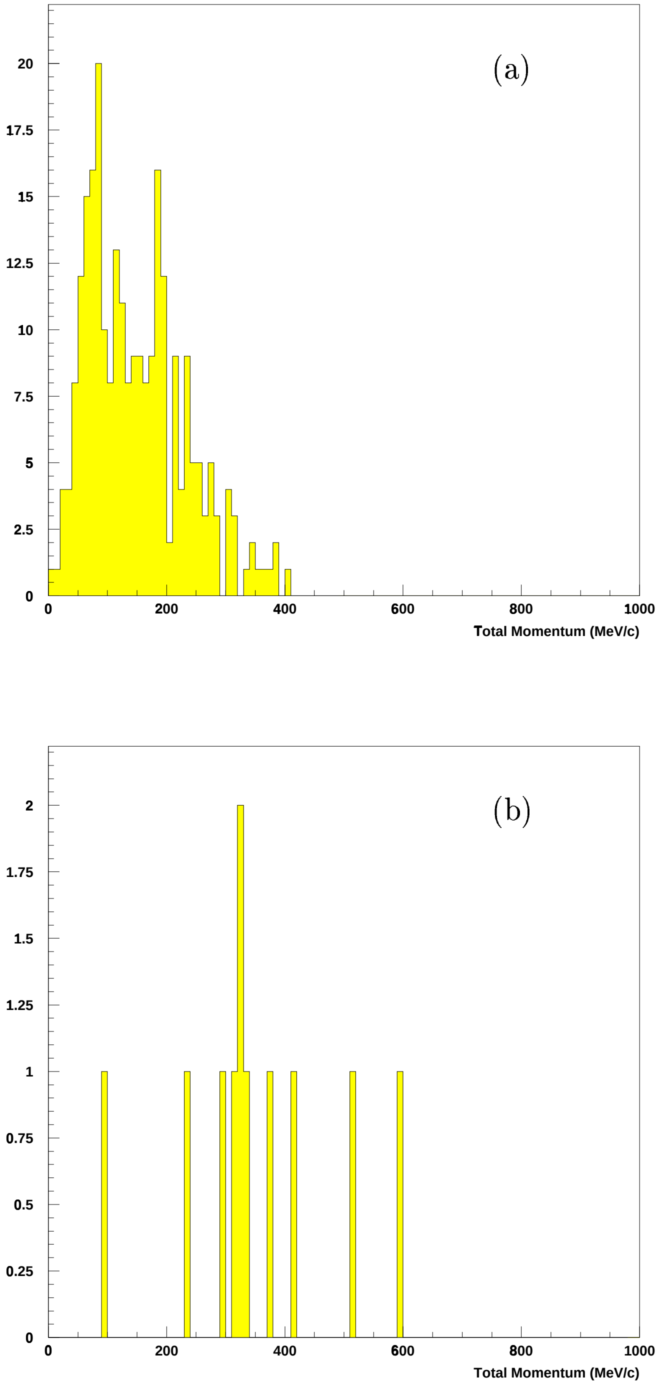


Figure 4.10: Distributions of reconstructed total momentum for (a) $p \rightarrow e^+\pi^0$ MC events and (b) atmospheric neutrino MC events which have passed first stage cuts as well as particle identification cuts.

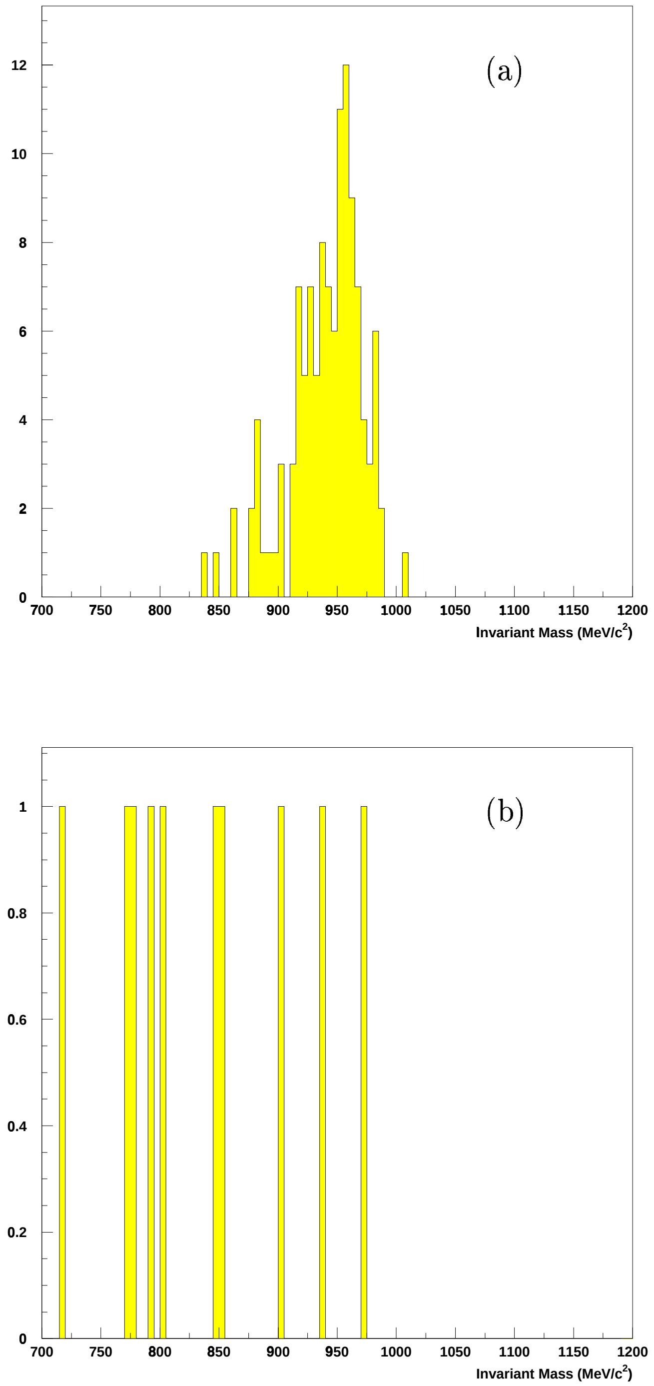


Figure 4.11: Distributions of reconstructed invariant mass for (a) $p \to \mu^+\pi^0$ MC events and (b) atmospheric neutrino MC events which have passed first stage cuts as well as particle identification cuts.

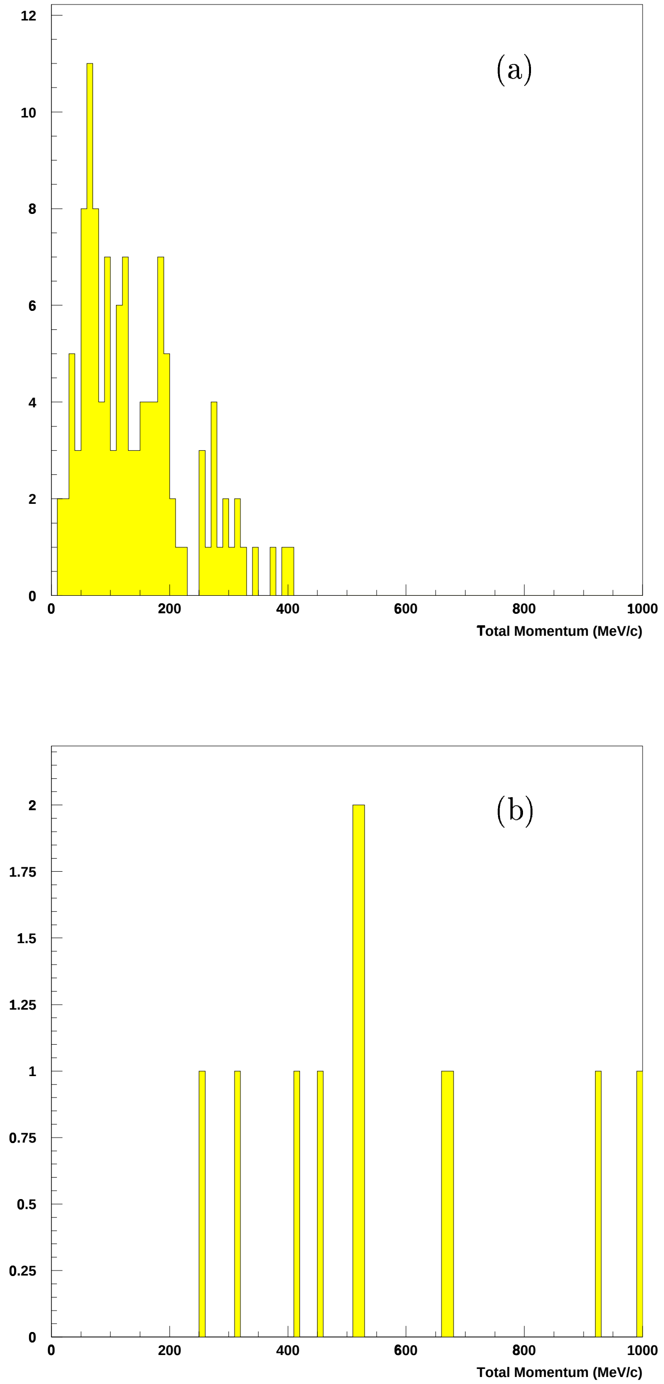


Figure 4.12: Distributions of reconstructed total momentum for (a) $p \rightarrow \mu^+\pi^0$ MC events and (b) atmospheric neutrino MC events which have passed first stage cuts as well as particle identification cuts.

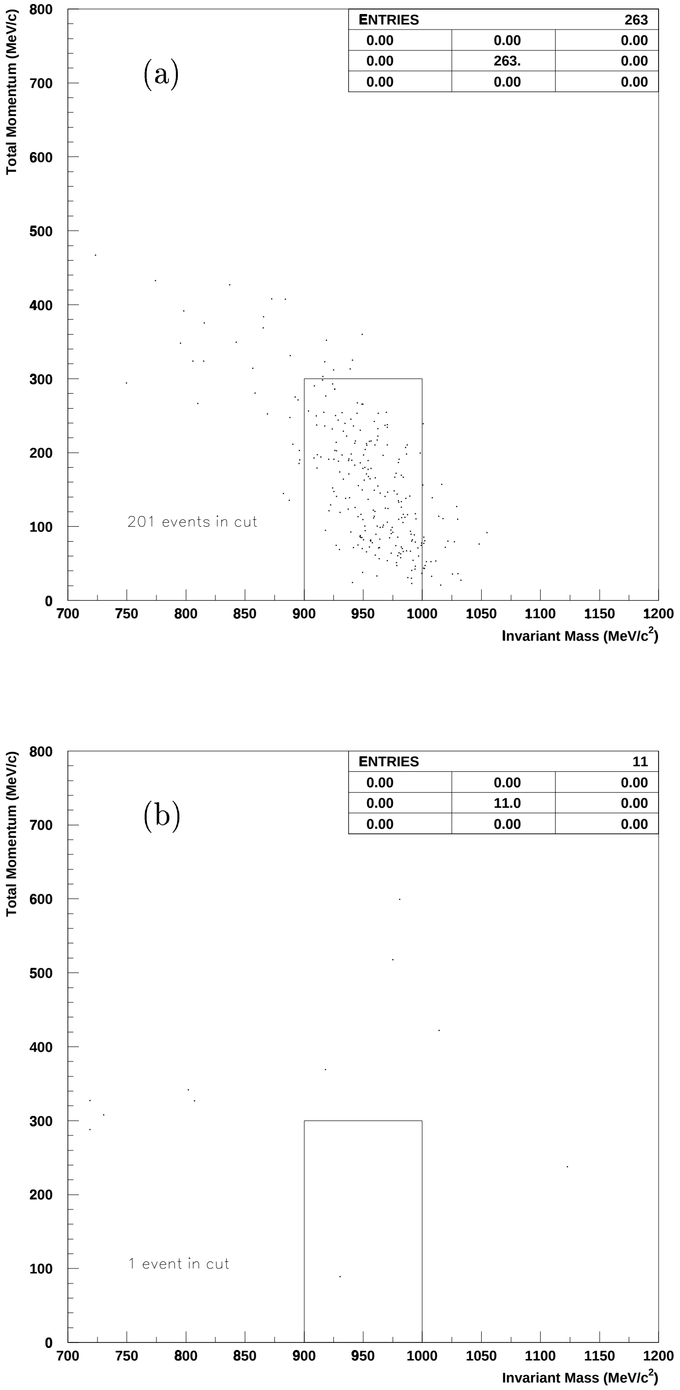


Figure 4.13: Distributions of reconstructed total momentum *vs.* invariant mass for (a) $p \rightarrow e^+\pi^0$ MC events and (b) atmospheric neutrino MC events which have passed first stage cuts as well as particle identification cuts. The box shows the cut.

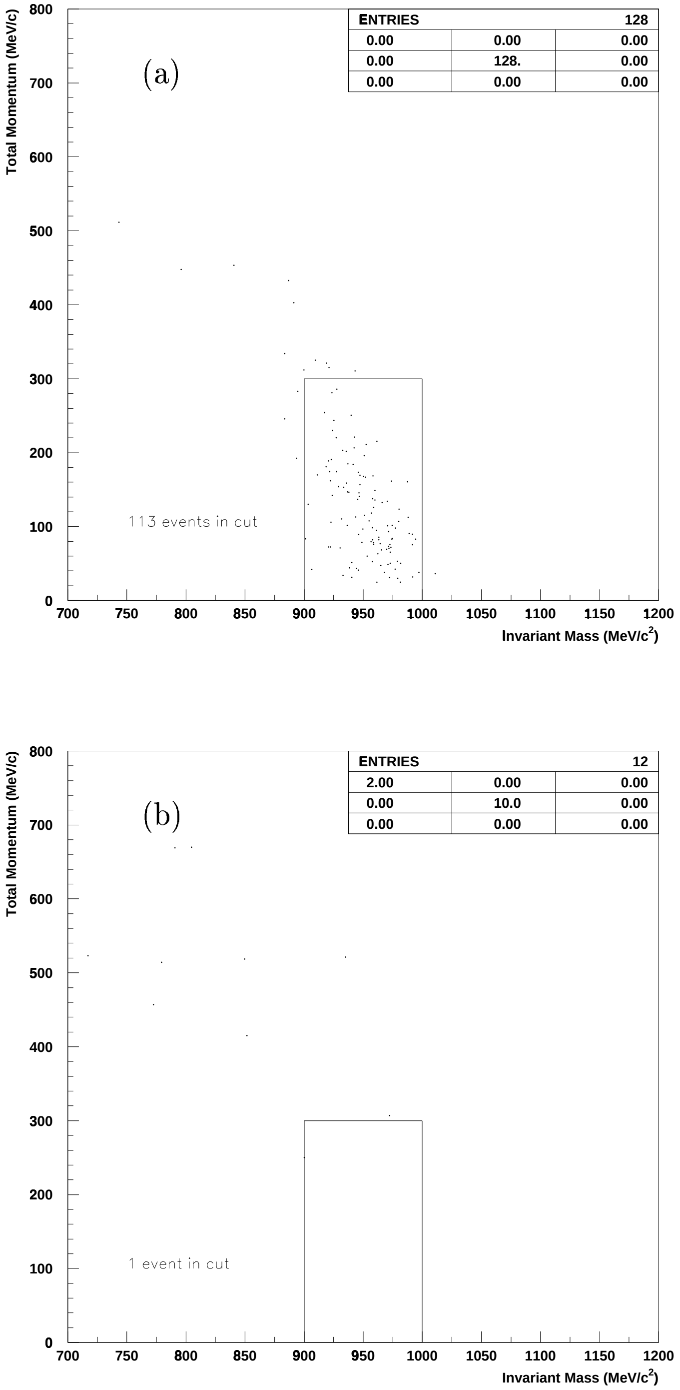


Figure 4.14: Distributions of reconstructed total momentum *vs.* invariant mass for (a) $p \rightarrow \mu^+\pi^0$ MC events and (b) atmospheric neutrino MC events which have passed first stage cuts as well as particle identification cuts. The box shows the cut.

4.4 Results

The cuts developed above using MC events samples are now applied to the data. The data is, not surprisingly, found to have very similar distributions as the atmospheric Monte Carlo. Fig. 4.15 shows the distribution of the number of decay electrons in the data. Fig 4.16 shows the distributions of data in anisotropy and visible energy space for the events passing the $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ decay electron cuts. Below each data plot the corresponding plot for the atmospheric neutrino MC sample is reproduced for comparison.

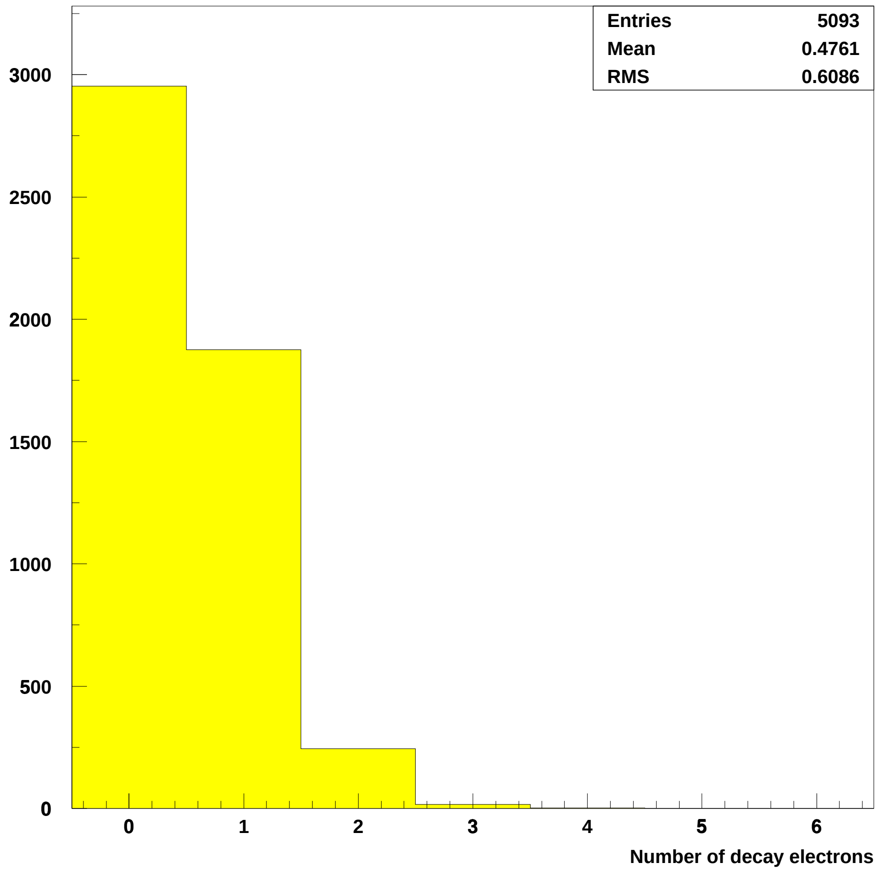


Figure 4.15: Distribution of number of decay electrons in the data.

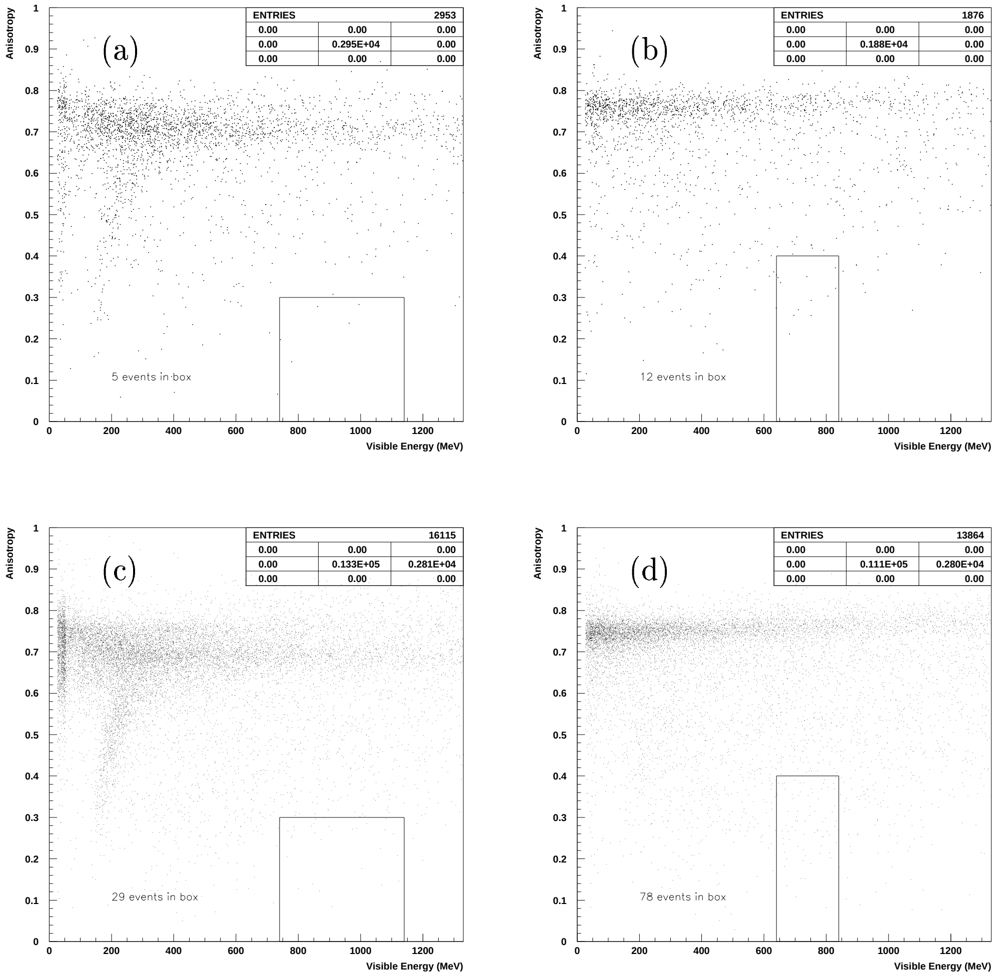


Figure 4.16: Distribution of data in anisotropy *vs.* visible energy after decay electron cuts for (a) $p \rightarrow e^+\pi^0$ and (b) $p \rightarrow \mu^+\pi^0$ selection. The box shows the cut. As a comparison these distributions for the atmospheric neutrino induced MC events passing (c) $p \rightarrow e^+\pi^0$ cuts and (d) $p \rightarrow \mu^+\pi^0$ cuts that were shown in Fig. 4.5(b) and Fig. 4.6(b) are reproduced.

For the data passing this first stage cuts, the number of showering and non-showering tracks in an event are shown in Fig. 4.17. The mass and momentum distribution of the events passing the particle identification cut are plotted in Fig. 4.18.

A summary of the number of events from the data and the various MC simulations which pass the selection criteria for the full proton decay search is shown in Table 4.1 and one for the free proton decay search is shown in Table 4.2. The row of values which is labeled with "(oscillated)" deserves some explanation.

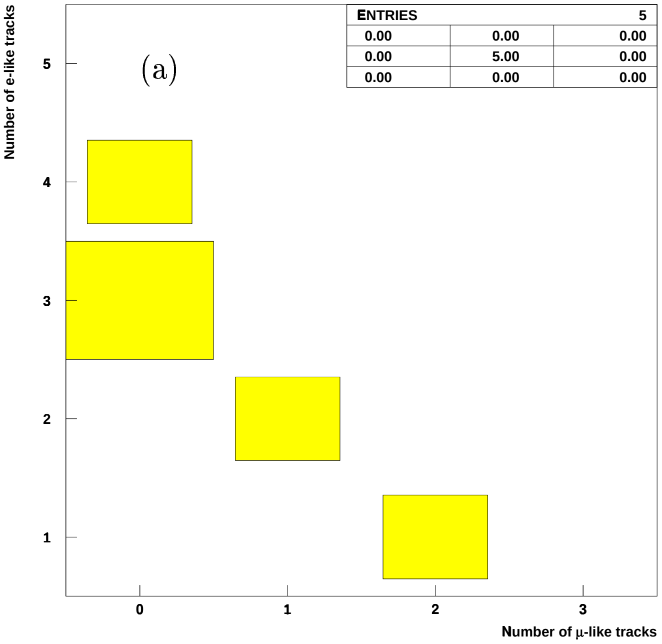
Strong evidence has been presented [31] that a ν_μ can oscillate to either a ν_τ or a sterile ν with this latter case being experimentally slightly disfavored. If a ν_μ oscillates to a ν_τ it could not create any visible events via CC interactions in Super–Kamiokande at the energy of proton decay events, but NC interactions could still be observed. If the oscillation is instead to a sterile neutrino then there will not even be a possible NC interaction to observe and the neutrino would effectively disappear from being a source of potential background. The probability that a ν_μ created in the atmosphere will be observed as a ν_μ in the detector is given by the well known equation,

$$\mathbf{P}(\nu_\mu \to \nu_\mu) = 1 - \sin^2 2\theta \sin^2\left(1.27\frac{\delta m^2 L}{E_\nu}\right), \tag{4.1}$$

where L (measured in km) is the distance between ν production and observation locations, E_ν (measured in GeV) is the neutrino energy, δm^2 (measured in eV^2) is the difference in the square of the masses of the two neutrino flavors, and θ is the mixing angle between them. The favored $\nu_\mu \to \nu_\tau$ mode is assumed and the current best fit point of $\delta m^2 = 0.0022\mathrm{eV}^2$ and $\sin^2 2\theta = 1.0$ is used.

Each atmospheric neutrino induced MC event is given a weight equal to the probability in equation 4.1. The sum of these weights is what is presented in the row marked "(oscillated)".

As can be seen, no data events survive these searches and thus no evidence for proton to decay via the $p \to e^+\pi^0$ nor $p \to \mu^+\pi^0$ modes is found. The following chapter uses these values to set a limit on the partial proton lifetime for these modes.



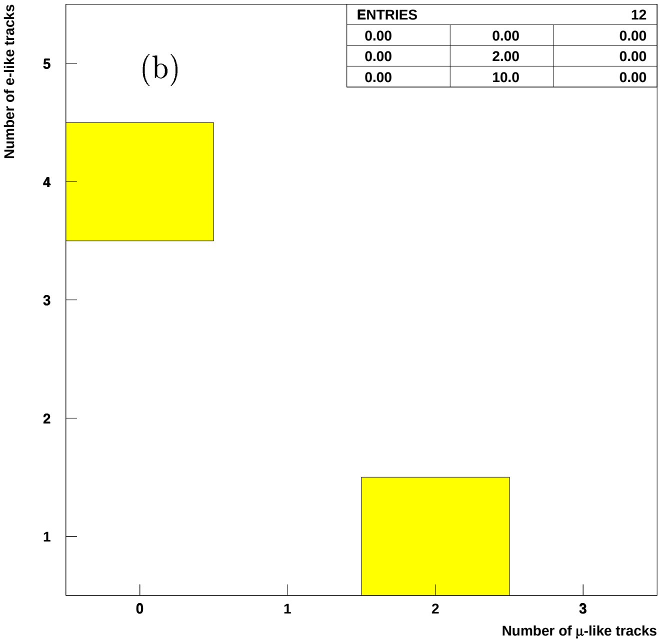


Figure 4.17: Distribution of number of electron like *vs.* number of muon like tracks in an event for data which passes (a) $p \rightarrow e^+\pi^0$ and (b) $p \rightarrow \mu^+\pi^0$ first stage cuts. Note that the low statistics make comparison to these distributions made for the MC atmospheric neutrino induced events shown in Fig. 4.7(b) and Fig. 4.8(b).

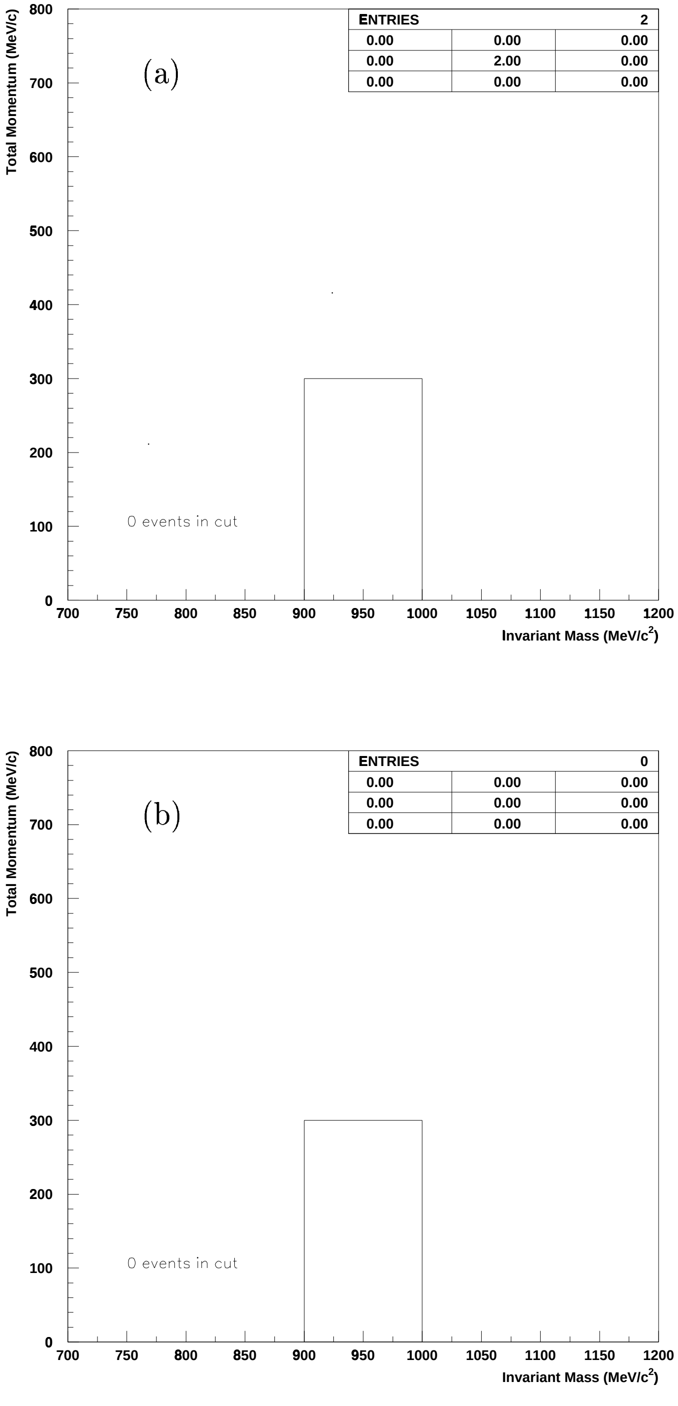


Figure 4.18: Distribution of the total momentum *vs.* invariant mass of an event for data passing (a) $p \rightarrow e^+\pi^0$ and (b) $p \rightarrow \mu^+\pi^0$ first stage and particle identification cuts. The box indicates the final cuts on P_{tot} and M_{inv}.

Sample	Tot.	Cont.	Decay $e^{\pm}$	$Anis/E_{vis}$	PID	P_{tot}/M_{inv}
Full $p \rightarrow e^{+}\pi^{0}$ cuts:						
$p \rightarrow e^{+}\pi^{0}$ MC	500	473	458	300	252	191
Atm. ν MC	62983	26263	13309	29	11	1
(oscillated)	52686	21675	12079	28	10	1
data	9100	5093	2953	5	2	0
Full $p \rightarrow \mu^{+}\pi^{0}$ cuts:						
$p \rightarrow \mu^{+}\pi^{0}$ MC	500	322	264	176	121	106
Atm. ν MC	62983	26263	11059	78	12	1
(oscillated)	52686	21675	8150	67	9	1
data	9100	5093	1876	12	0	0

Table 4.1: Summary of events passing proton decay (free and nuclear) search cuts.

Sample	Tot.	Cont.	Decay $e^{\pm}$	$Anis/E_{vis}$	PID	P_{tot}/M_{inv}
Free $p \rightarrow e^{+}\pi^{0}$ cuts:						
$p \rightarrow e^{+}\pi^{0}$ MC	200	130	130	128	73	63
Atm. ν MC	62983	26263	13309	2	1	1
(oscillated)	52686	21675	12079	2	1	1
data	9100	5093	2953	0	0	0
Free $p \rightarrow \mu^{+}\pi^{0}$ cuts:						
$p \rightarrow \mu^{+}\pi^{0}$ MC	200	136	126	120	70	66
Atm. ν MC	62983	26263	11059	15	1	0
(oscillated)	52686	21675	8150	14	1	0
data	9100	5093	1876	1	0	0

Table 4.2: Summary of events passing free proton decay search cuts.

Chapter 5

Lifetime Limits

Once the data have been reduced and the selection criteria have been chosen and applied, the actual analysis of nucleon decay searches is relatively simple. There are two possible outcomes of the selection: either a significant excess of events over background are found and evidence for proton decay is concluded or an insignificant or zero excess is found in which case a limit on the nucleon decay rate can be calculated. From the results above, this thesis is in the latter category. A somewhat general description of the method used to calculate the partial lifetime limits is shown below followed by its application to the results of the previous chapter.

5.1 Calculating the Confidence Limit

When one monitors a sample of N particles for a time period T and finds a nonzero number n of decays significantly above the expected background, the decay rate is estimated simply by,

$$\Gamma = \frac{n}{NT\epsilon}, \tag{5.1}$$

where ϵ is the detection efficiency to observe the decay products. The uncertainty of this rate can be found in the usual method of propagating uncertainty in the individual parameters.

When n is zero (or even negative) a rate can not be accurately calculated. This is because n is no longer a good estimator of the true mean. Naively one could set a decay rate of zero but this may be wrong for it may be the case that if we had observed more particles for a longer time a nonzero number of decays would have been seen. So, instead a limit on the decay rate at some confidence is set.

In simple situations, it is common to use the *Poisson Process with Background* method [32]. This simple Poisson limit is set by finding a value $\mu_{lim}(= s_{lim} + b_0)$ which is the mean of a Poisson distribution such that there would be a certain probability (usually 90%) that any value chosen from this distribution would be larger than the number of candidates n assuming an expected background b_0. This equation for s_{lim} is shown below in (5.15). To find the limit on the decay rate, s_{lim} is then substituted for n in (5.1).

However, what this method lacks is the ability to take into consideration any uncertainties in the parameters that go in to the limit, namely the efficiency ϵ, the background b and the exposure $\lambda = NT$. Because the concept of this limit is innately based on *confidence* we must somehow include the degree to which we are or are not confident in our estimated quantities.

Through the use of Bayes' theorem [33, 34] and the assignment of prior probability distributions (priors) it is possible to use the knowledge of our uncertainty in the parameters to calculate a limit which better corresponds to the confidence we quote. In this thesis the method used is directly based on Ref. [35].

5.1.1 The Method

The expected mean number of candidates μ in the data sample is modeled simply by,

$$\mu = \Gamma\lambda\epsilon + b. \tag{5.2}$$

Here, as above, Γ is the true decay rate (for a given decay mode), $\lambda = NT$ is the true exposure, ϵ is the true efficiency to detect the decay products, and b is the true mean number of background events in our sample. Because this is a counting experiment the probability to detect n events follows the Poisson distribution with Poisson parameter μ. That is, given the parameters: Γ, λ, ϵ and b, and all conditions of the experiment, I, the probability to detect n events is,

$$\mathbf{P}(n\,|\,\Gamma\lambda\epsilon b I) = \frac{e^{-(\Gamma\lambda\epsilon+b)}(\Gamma\lambda\epsilon + b)^n}{n!}. \tag{5.3}$$

Applying Bayes' theorem gives the probability for a set of parameters given a measurement

$$\mathbf{P}(\Gamma\lambda\epsilon b\,|\,nI) = A\mathbf{P}(n\,|\,\Gamma\lambda\epsilon bI)\mathbf{P}(\Gamma\lambda\epsilon b\,|\,I). \tag{5.4}$$

The constant of proportionality A will be found by insisting that $\mathbf{P}(\Gamma\lambda\epsilon b\,|\,nI)$ is normalized. Given a particular set of cuts, the decay rate, efficiency, exposure, and background are all independent of one another. Therefore, the value

$\mathbf{P}(\Gamma\lambda\epsilon b\,|\,I)$ can be separated into its constituents giving,

$$\mathbf{P}(\Gamma\epsilon\lambda b\,|\,nI) = A\mathbf{P}(n\,|\,\Gamma\epsilon\lambda bI)\mathbf{P}(\Gamma\,|\,I)\mathbf{P}(\epsilon\,|\,I)\mathbf{P}(\lambda\,|\,I)\mathbf{P}(b\,|\,I). \tag{5.5}$$

The quantities, $\mathbf{P}(\Gamma\,|\,I)$, $\mathbf{P}(\epsilon\,|\,I)$, $\mathbf{P}(\lambda\,|\,I)$ and $\mathbf{P}(b\,|\,I)$, are known as the prior probability distributions (priors) for Γ, ϵ, λ and b respectively. These distributions codify the state of knowledge of each parameter before the outcome of our experiment is known. The important subject of the choice of priors is discussed below. Given the priors, the joint probability distribution in (5.5) is marginalized (integrated) with respect to the so called nuisance variables: ϵ, λ and b. This marginalization is done because the nuisance variables are not strictly interesting. Instead, we are interested only in what the experiment can tell us about the decay rate, the distribution for which can now be written as,

$$\mathbf{P}(\Gamma\,|\,nI) = \int\!\!\int\!\!\int \mathbf{P}(\Gamma\epsilon\lambda b\,|\,nI)\, d\epsilon\, d\lambda\, db. \tag{5.6}$$

At this point the normalization constant, A, can be resolved by demanding,

$$\int_0^\infty \mathbf{P}(\Gamma\,|\,nI)\, d\Gamma = 1. \tag{5.7}$$

Once it is known how the decay rate is distributed it is theoretically trivial to calculate a limit using a definition similar to that used to find s_{lim} in the simple Poisson limit, except this time the limit on the rate is calculated directly with all uncertainties considered. Specifically the limit on the rate Γ_{lim} is simply the value for which any rate chosen from $\mathbf{P}(\Gamma\,|\,nI)$ would be smaller, at some probability equal to the confidence level (CL). This is done by solving,

$$\mathrm{CL} = \int_0^{\Gamma_{lim}} \mathbf{P}(\Gamma\,|\,nI) d\Gamma, \tag{5.8}$$

for Γ_{lim}.

5.1.2 Reduction to the Simple Poisson Limit

The method just described will reduce to the simple Poisson limit [32] when systematic uncertainties go to zero and the prior for the decay rate is taken to be uniform. To show this, $\Gamma\epsilon\lambda$ is written simply as s the mean signal expected in the data. Following the above method, the mean number of expected events is,

$$\mu = s + b, \tag{5.9}$$

and the probability to observe n events given the mean signal s and background b is,

$$\mathbf{P}(n\,|\,sbI) = \frac{e^{-(s+b)}(s+b)^n}{n!}. \tag{5.10}$$

Applying Bayes' theorem gives,

$$\mathbf{P}(sb\,|\,nI) = A\mathbf{P}(n\,|\,sbI)\mathbf{P}(s\,|\,I)\mathbf{P}(b\,|\,I). \tag{5.11}$$

In order to reduce to the simple Poisson limit, the prior for the signal $\mathbf{P}(s\,|\,I)$ is taken to be uniform, and it is assumed that the background is known with certainty, *ie.* $\mathbf{P}(b\,|\,I) = \delta(b - b_0)$. Marginalizing then gives,

$$\mathbf{P}(s\,|\,nI) = A\int_0^\infty \frac{e^{-(s+b)}(s+b)^n}{n!}\delta(b-b_0)\,db = A\frac{e^{-(s+b_0)}(s+b_0)^n}{n!}. \tag{5.12}$$

To find the limit on the signal s_{lim} at a confidence CL one must solve, after enforcing normalization and substituting into (5.8),

$$\mathrm{CL} = \frac{\int_0^{s_{lim}} e^{-(s+b_0)}(s+b_0)^n\,ds}{\int_0^\infty e^{-(s+b_0)}(s+b_0)^n\,ds}. \tag{5.13}$$

Making a change of variables of $x = s + b_0$ and using that

$$\int_{x_1}^{x_2} x^m e^{-x}\,dx = -e^{-x}\sum_{r=0}^{m}\frac{m!x^{(m-r)}}{(m-r)!}\Bigg|_{x=x_1}^{x_2}, \tag{5.14}$$

with $x_1 = b_0$ and $x_2 = s_{lim} + b_0$ for the numerator in (5.13) and $x_1 = b_0$ and $x_2 = \infty$ for the denominator, gives equation (28.40) of Ref. [32],

$$\mathrm{CL} = 1 - \frac{e^{-(s_{lim}+b_0)}\sum_{r=0}^{n}\frac{(s_{lim}+b_0)^r}{r!}}{e^{-b_0}\sum_{r=0}^{n}\frac{b_0^r}{r!}}. \tag{5.15}$$

Thus it is shown that the simple Poisson plus background limit is equivalent to a Bayesian method with a hidden implicit uniform prior for the signal and a background assumed to be perfectly known.

5.1.3 The Priors

There are two classes of priors in this method. One class contains Γ, the other the nuisance variables: ϵ, λ and b. In the former case, Γ is the parameter to be estimated with the experiment, in the latter, the parameters appear in (5.2) but are not directly interesting and are marginalized out of the final result. In both cases there is some minimum knowledge: there is some region where the parameter is physically allowed. So at a minimum, non-physical regions can be excluded by assigning them zero prior probability[1]. In addition, there is significant prior information for the nuisance parameters. In fact, explicit experiments have been performed in order to estimate their values. These estimations are imperfect due to systematic and statistical uncertainties. To handle these, the systematic uncertainty is assumed to result in a Gaussian distribution, while statistical uncertainty results in a distribution appropriate for the particular statistical process. As each prior is discussed in turn, note that the individual normalizations of the priors are ignored as they would cancel due to the application of (5.7).

Decay rate

The decay rate is the parameter which this experiment attempts to limit. Because of this, we do not want to bias the result by using a prior with anything but a minimum of information. Such a prior is dubbed a *least informative prior* (LIP). The method of uniquely and objectively determining the LIP is to maximize the *entropy* (MAXENT) of the parameter subject to any known constraints including any group symmetries appropriate to the problem [36]. In the case where it is known that a parameter must be strictly positive and nonzero, the LIP determined by MAXENT is the so called Jeffreys prior $\mathbf{P}(x \,|\, I) \propto 1/x$. It is least informative because it gives no bias to any particular scale for x. In comparison a uniform prior for x which naively seems less informative actually biases the scale towards higher values of x. Indeed, with a uniform prior $\mathbf{P}(x > x_0) > \mathbf{P}(x < x_0)$ for all x_0. However, proton decay lifetime limits have historically been calculated with the above mentioned simple Poisson limit [32], and this method, as seen in section 5.1.2, implicitly assumes a uniform prior. This same (non-LIP) uniform prior will be taken here in order that this limit would reduce to the historical one in the absence of uncertainties. Using a Jeffreys prior for the decay rate would result in a lifetime limit that is numerically larger but, of course, with a different meaning.

Explicitly then, the prior probability distribution for the decay rate is,

$$\mathbf{P}(\Gamma \,|\, I) \propto \begin{cases} 1 & \Gamma \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{5.16}$$

---

[1] In fact, I have let some of this minimum knowledge "leak" from the prior and into the limits of integration in (5.7). To be pedantic the limits should be $\{-\infty, \infty\}$ and the non-physical region of negative rate should be assigned zero probability density in the prior for the rate $\mathbf{P}(\Gamma \,|\, I)$.

In practice, the calculation is truncated at very high values of Γ in order to make the integrals more easily calculable.

Efficiency

The efficiency is estimated by generating many Monte Carlo events which simulate the decay of protons, their travel through any nuclear matter and the detector response to the Čerenkov light they emitted as they travel through the water. The number of proton decay MC events passing the cuts divided by the number generated within the fiducial volume of the detector is then the efficiency. For the search modes of this thesis, this estimate suffers uncertainty primarily due to the lack of knowledge of how the π^0 should be transported through the nucleus. This lack of knowledge has been estimated to be about 15% of the estimated efficiency. So, prior to doing the experiment, it is believed that the probability density of the true efficiency is Gaussian distributed about the estimated value. In addition, it is certain that the efficiency is above zero and is below one. This leads to a truncated Gaussian prior distribution for the efficiency,

$$\mathbf{P}(\epsilon\,|\,I) \propto \begin{cases} e^{-(\epsilon-\epsilon_0)^2/2\sigma_\epsilon^2} & 0 \leq \epsilon \leq 1 \\ 0 & \text{otherwise} \end{cases}, \tag{5.17}$$

where ϵ_0 is the estimated efficiency and σ_ϵ is the uncertainty in this efficiency.

Exposure

The exposure is easily estimated given the detector content (*ie.* water), fiducial size, and live time of our data sample. The fiducial volume and live time are known to better than 0.1%. However, there is a small systematic due to vertex fitter differences between data and MC which effectively can be attributed to an exposure uncertainty of 5%. The prior probability for the exposure is then,

$$\mathbf{P}(\lambda\,|\,I) \propto \begin{cases} e^{-(\lambda-\lambda_0)^2/2\sigma_\lambda^2} & 0 \leq \lambda < \infty \\ 0 & \text{otherwise} \end{cases}, \tag{5.18}$$

where λ_0 is the estimated exposure and σ_λ is the uncertainty in this estimate.

Background

The true mean of the real background is estimated by generating a sample of atmospheric neutrino Monte Carlo events (while considering that neutrinos oscillate), counting the number which pass the proton decay selection criteria and normalizing to the livetime of the data. Initially this procedure seems very straightforward. However, this estimation itself is a counting experiment and the number of events counted will in general be a statistical fluctuation from the true MC background mean. Furthermore, the MC only approximates reality and because of uncertainties in the atmospheric neutrino flux and neutrino–oxygen cross sections used in the MC, the true MC background mean is systematically different from the true real background mean. Given this uncertainty we can ask what the probability that the true MC background mean is some value given the true real background mean. These concepts suggest the following derivation for the background prior which is similar to calculations used in some past proton decay searches [20].

To calculate $\mathbf{P}(b \mid I)$, it is necessary to explicitly state some of the conditions represented by I. Namely the conditions that there are n_b atmospheric neutrino MC events passing the proton decay selection criteria out of C times more MC events than data events. Symbolically this can be written as, $I \equiv n_b C I'$, where I' symbolizes all conditions and information except the values of n_b and C. The method of section 5.1.1 is then followed so that the probability density function used as a background prior when analyzing the data is actually the posterior probability from the analysis of the MC "experiment".

So, by applying Bayes' theorem,

$$\mathbf{P}(b \,|\, I) = \mathbf{P}(b \,|\, n_b C I') \propto \mathbf{P}(n_b \,|\, b C I')\mathbf{P}(b \,|\, I'). \tag{5.19}$$

Here, $\mathbf{P}(b \mid I')$, is the information prior to considering the atmospheric neutrino MC. Since there is no constraint to match some historical background prior (the historical method assumes an exactly known background) the least informative Jeffreys prior is chosen for $\mathbf{P}(b \mid I')$. This will give a numerically lower lifetime limit than if a uniform prior is chosen.

There is no direct way to calculate $\mathbf{P}(n_b \mid b C I')$, as the number of atmospheric neutrino MC events passing proton decay selections do not directly relate to the true, real mean background. However, it is related to the true mean atmospheric neutrino MC background, and this in turn is related to the true mean real background. In order to include these dependencies the partition theorem [32] is used. This is briefly described. Given an event, (possible outcome of an experiment), y, from a sample space, S, and the set of continuous values, $\{x\}$, representing unique events which partition S such that the union of all the members of $\{x\}$ is S, the partition theorem says,

$$\mathbf{P}(y \,|\, I) = \int \mathbf{P}(y \,|\, x I)\mathbf{P}(x \,|\, I)\, dx. \tag{5.20}$$

This is analogous to the completeness theorem (ie. the inserting of $\mathbf{1} = \int |x\rangle\langle x|\, dx$) of Quantum Mechanics.

In this case of estimating the background, x is chosen to be the parameter b_{MC}, the true mean MC background, and equation (5.19) becomes[2],

$$\mathbf{P}(b\,|\,I) \propto \mathbf{P}(b\,|\,I') \int_0^\infty \mathbf{P}(n_b\,|\,b_{MC}bCI')\mathbf{P}(b_{MC}\,|\,bCI')\, db_{MC}. \tag{5.21}$$

The first term, $\mathbf{P}(n_b \mid b_{MC}bCI')$, contains the statistical uncertainty in the background estimation due to finite size MC sample. Note that while b appears in this term due to the algebra, the probability of getting n_b events does not depend on b. This statistical uncertainty is expressed via the Poisson distribution. The second term, $\mathbf{P}(b_{MC} \mid bCI')$, contains the systematic uncertainty in the the MC background mean. This distribution is taken as Gaussian. The full background prior is then,

$$\mathbf{P}(b\,|\,I) \propto \frac{1}{b} \int_0^\infty \frac{e^{-b_{MC}} {b_{MC}}^{n_b}}{n_b!} e^{-(b_{MC}-bC)^2/2\sigma_b^2}\, db_{MC}, \tag{5.22}$$

for $b > 0$ and zero otherwise.

5.1.4 Implementation

The problem of finding the decay rate limit then becomes the problem of solving (5.8) which for this case becomes:

$$\begin{aligned} \mathrm{CL} = \mathbf{P}(0 < \Gamma < \Gamma_{lim}\,|\,n) = & \qquad (5.23) \\ A \int_0^{\Gamma_{lim}} d\Gamma \int_0^\infty db \int_0^\infty db_{MC} \int_0^\infty d\lambda \int_0^1 d\epsilon\, I(n, \Gamma, b, b_{MC}, \lambda, \epsilon, \ldots), & \end{aligned}$$

with the integrand,

$$\begin{aligned} I(n, \Gamma, b, b_{MC}, \lambda, \epsilon, \ldots) = & \qquad (5.24) \\ \frac{b_{MC}^{n_b}(\Gamma\lambda\epsilon + b)^n}{b} \exp\left(-b_{MC} - \frac{(b_{MC} - bC)^2}{2\sigma_b^2} - \frac{(\lambda - \lambda_0)^2}{2\sigma_\lambda^2} - (\Gamma\lambda\epsilon + b)\right). & \end{aligned}$$

The factorials, $n_b!$ and $n!$ are dropped as they will cancel during normalization. The constant of normalization is found by setting $\mathrm{CL} = 1$ and $\Gamma_{lim} \to \infty$ and solving for A. It should also be noted that if the outcome of the experiment

---

[2] Here the regions of zero prior probability are explicitly stated by the limits of integration.

finds no candidate events then the limit calculation can be simplified by dropping the terms involving the background as they factor and will cancel during normalization.

Solving for an upper limit Γ_{lim} of some confidence (typically 90%), is done by minimizing, $|\mathrm{CL} - \mathbf{P}(0 < \Gamma < \Gamma_{lim}\,|\,n)|$ as a function of Γ_{lim}. This minimization is performed by the Brent [37] minimization algorithm as implemented in the GNU Scientific Library [38] (GSL).

Of course to perform the minimization a sharply peaked multi-dimensional integral must be calculated. The method found to have the most success is the application of Monte Carlo integration techniques. Specifically, using custom code called `monte` which uses the `VEGAS` [39] algorithm from the GSL. To use this method, care must be taken in some of the details. Balance between high accuracy and runtime is needed. If many more parameters were to be integrated over, this method would become less attractive.

There is currently active research in the field of statistics which uses Markov Chain Monte Carlo (MCMC) to solve similar and even more complex problems in Bayesian inference. Future work in finding limits based on the above method may make use of MCMC techniques.

5.2 Resulting Limits

From the results of the proton decay searches summarized in Tables 4.1 and 4.2 the parameters needed for the limit calculation are obtained. As can be seen, the $p \to e^+\pi^0$ mode (both the full and the free proton decay search) have a single (and the same) background MC event. This event is from a charged current electron neutrino interaction, $\nu_e + n \to e^- + p + \pi^0$ with an additional π^0 produced in the nucleus. The full $p \to \mu^+\pi^0$ search also has one background MC event. This event is a neutral current electron neutrino interaction, $\nu_e + p \to \nu_e + n + \pi^+$, again, with an additional π^0 produced in the nucleus. Because both of these background MC events are from ν_e interactions, their weight due to oscillation remains 1.0. The free $p \to \mu^+\pi^0$ search finds no atmospheric neutrino background MC events.

The efficiency for each mode is found by dividing the number of proton decay MC events which pass all cuts by the number which pass the cuts defining the contained event sample. This is 42.4% and 35.1% for the full searches and 53.1% and 49.3% for the free searches for the $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ modes, respectively.

Using the oscillated number of atmospheric neutrino MC events, 4.26 times more atmospheric neutrino MC was used than data collected. This oversampling factor is the ratio of the number of atmospheric neutrino MC events (weighted by oscillation probability) which pass the contained event cuts and the number of data events which pass the same. Assuming no oscillation, this factor is 5.16.

Finally, the exposure for the 784.9 days of running is 1.62×10^{34} proton·years for the full proton decay searches and 0.324×10^{34} proton·years for the free searches. These values are summarized in Table 5.1.

Mode	efficiency	exposure ($\times10^{34}$p·yrs)	background
$p \to e^+\pi^0$	42.4%	1.62	1
$p \to \mu^+\pi^0$	35.1%	1.62	1
$p \to e^+\pi^0$-free	53.1%	0.324	1
$p \to \mu^+\pi^0$-free	49.3%	0.324	0

Table 5.1: Summary of parameters used in the limit calculation.

Mode	Lifetime Limit (90%)	with no osc.	Poisson + Bkg meth.
$p \to e^+\pi^0$	2.59×10^{33} yrs	2.57×10^{33} yrs	2.83×10^{33} yrs
$p \to \mu^+\pi^0$	2.07×10^{33} yrs	2.07×10^{33} yrs	2.31×10^{33} yrs
$p \to e^+\pi^0$-free	0.62×10^{33} yrs	0.62×10^{33} yrs	0.68×10^{33} yrs
$p \to \mu^+\pi^0$-free	0.62×10^{33} yrs	0.63×10^{33} yrs	0.68×10^{33} yrs

Table 5.2: 90% lifetime confidence limits.

Based on no observed proton decay candidates in the data, the calculated 90% confidence limits are summarized in Table 5.2. The first column holds limits which were found by applying the method described in section 5.1.1 assuming $\nu_\mu \to \nu_\tau$ oscillation with $\delta m^2 = 0.0022$ (eV)^2 and $\sin^2 2\theta = 1.0$. The second column shows limits assuming no oscillation. The third column holds limits computed using the traditional Poisson plus background method.

Chapter 6

Conclusions and Future

This final chapter gives the conclusions from this search and outlines where nucleon decay searches are headed in the future.

6.1 Conclusions

Evidence for protons decaying via $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ was searched for in 784.9 days of data from the Super–Kamiokande detector. Less than 0.25 events due to atmospheric neutrino background were expected to have been found. No events in this data were found to be consistent with proton decay in these modes. A method based on Bayesian probability was developed to calculate a confidence limit on the partial lifetimes. This limit takes into consideration all known significant statistical and systematic uncertainties in the parameters on which the measurement depends. This calculation found partial lifetime limits of $\tau(p \rightarrow e^+\pi^0) > 2.59 \times 10^{33}$ years and $\tau(p \rightarrow \mu^+\pi^0) > 2.07 \times 10^{33}$ years at a 90% confidence assuming the background was from atmospheric neutrino induced events and that a ν_μ can oscillate to a ν_τ with the parameters $\delta m^2 = 0.0022\text{eV}^2$ and $\sin^2 2\theta = 1.0$.

The resulting $p \rightarrow e^+\pi^0$ partial proton decay lifetime limit is an increase of a factor of 4.8 and the $p \rightarrow \mu^+\pi^0$ limit is an increase of a factor of 4.4 over the previous best limits set by IMB [20]. These improved limits begin to carve out the range of predicted values of the "flipped" SU(5) and minimal non-SUSY SO(10) theories mentioned in chapter 1. In the latter case, the "central value" of Model I from Ref. [6] is even ruled out, although the distribution of this prediction is not simple and the true predicted partial lifetime may very well still be above limits set here.

More detector livetime would need to be accumulated to completely rule out this model, however the background will start to become significant which causes the limit (or the sensitivity to observe proton decay) not to increase linearly with time. For example, if the current $p \to e^+\pi^0$ search was extended for four times longer about one background event would be expected. If such a search were to still turn up no candidates the limit on the partial proton lifetime would be pushed up to $\tau_{limit} \sim 1.0 \times 10^{34}$ years, which is notably less than four times the limit found in this thesis. On the other hand, finding one candidate, which would still be consistent with expected background, would drop this limit to $\tau_{limit} \sim 7.1 \times 10^{33}$ years. This example points out that event detection and selection methods, while adequate for current background levels, do not scale well to longer detector livetimes. If the predictive theories shown above are to be ruled out or confirmed some improvements are needed. Some possibilities for such improvements are discussed in the next section.

6.2 The Future

Nucleon decay lifetime limits or, equivalently sensitivity to finding evidence of nucleon decay for modes which are background free will increase linear with exposure. One the other hand, if a mode has significant background (approaching one event or more) then the limit (or sensitivity) will only grow as the square root of the exposure. Even in the best case of background free modes, one must wait an order of magnitude longer to get an order of magnitude higher limit. With many theories allowing for decay lifetimes which are many orders of magnitude higher than current experimental limits, one must be very patient if one relies on waiting.

For the rest of us, there are several options. Improvements in event reconstruction algorithms may be able to raise efficiencies by a factor of two or three or more, particularly for modes which suffer small efficiencies from tight cuts needed to reduce background. This improved reconstruction will also tend to allow an improved background rejection which could move background limited modes out of their square root prison. If such improvements can be found then we can at least multiply a significant constant to our linear (at best) dependence on exposure.

The other, more obvious way to increase this constant is to just build a larger detector. In some ways, one could imagine building a larger version of Super–Kamiokande with an order of magnitude more mass. This idea brings up several difficulties. Such a detector would need an underground cavity of such a size that even the best rock would have trouble supporting the span. Caverns of 50 m × 50 m cross section are possible in some locations, but not much larger [40]. Another problem is the average attenuation length in

Super–Kamiokande is about 85 meters so any larger detector will have to face the difficulties of having most of the Čerenkov light being attenuated before being detected. Of course, in such an experiment, like this one, the attenuated photons can be corrected up, but Poisson statistical fluctuations will also be corrected up which will make event reconstruction more difficult. To get around this problem, PMTs could be distributed throughout the volume, with the corresponding costs which would scale by the cube instead of the square of the detector size.

Recently (23-25 September 1999) Stony Brook hosted the International Workshop on Next Generation Nucleon Decay and Neutrino Detector [41] to bring together the neutrino and nucleon decay community to discuss the problems mentioned above and others associated with building a better detector. Later, (25-26 February 2000) discussions relating specifically to nucleon decay were continued at UCI. Besides studies based on a larger water Čerenkov detector, there were suggestions for lead calorimeters and a large liquid argon detector TPC (Super Icarus) targeting kaon modes as well as a large spherical water Čerenkov detector with focusing, mirrored walls (AquaRICH).

What detector will be used for future experiments is being discussed as this dissertation is being written. So far there are many options and the relative merits of each are being explored. Whatever the decision, it is clear that there is much excitement about the prospect of detecting proton decay in my life time. I certainly hope I can learn about some future searcher's discovery of proton's decay before it is my time to do the same.

Bibliography

[1] S. Glashow, Nucl. Phys. **22**, 579 (1961)

[2] S. Weinberg S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[3] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).

[4] J. Ellis, J. S. Hagelin, S. Kelley and D. V. Nanopoulos, Nucl. Phys. **B311**, 1 (1988).

[5] I. Antoniadis, J. Ellis, J. S. Hagelin and D. V. Nanopoulos, Phys. Lett. **B231**, 65 (1989).

[6] D. Lee, R. Mohapatra, M. Parida, and M. Rani, Phys. Rev. D. **51**, 229 (1995).

[7] Y.Fukuda *et al.*, Phys. Lett. **B433**, 9 (1998).

[8] J. C. Pati, Int. J. Mod. Phys. **A14**, 2949 (1999) [hep-ph/9811442].

[9] K. S. Babu, J. C. Pati and F. Wilczek, Nucl. Phys. **B566**, 33 (2000) [hep-ph/9812538].

[10] M. Shiozawa, B. Viren *et al.* [Super–Kamiokande Collaboration], Phys. Rev. Lett. **81**, 3319 (1998) [hep-ex/9806014].

[11] Y. Hayato, M. Earl *et al.* [Super–Kamiokande Collaboration], Phys. Rev. Lett. **83**, 1529 (1999) [hep-ex/9904020].

[12] F. Reines and J. Schultz, Surveys High Energ. Phys. **1**, 89 (1980).

[13] G. Flerov *et al.*, Sov. Phys. Doklady **3**, 79 (1958).

[14] S. P. Rosen, Phys. Rev. Lett. **34**, 774 (1975).

[15] J. C. Evans, Jr. and R. I. Steinberg, Science **197**, 989 (1977).

[16] F. Reines, C. L. Cowan, Jr. and M. Goldhaber, Phys. Rev. **96**, 1157 (1954).

[17] M. C. Goodman *et al.*, *Submitted to 26th International Cosmic Ray Conference (ICRC 99), Salt Lake City, UT, 17-25 Aug 1999.*

[18] MINOS Collaboration, *The MINOS Detectors Technical Design Report*, NuMI-L-337, (1998).

[19] R. Becker-Szendy *et al.*, Nucl. Instrum. Meth. **A324** (1993) 363.

[20] C. McGrew *et al.*, Phys. Rev. **D59**, 052004 (1999).

[21] K. Hirata *et al.*, Phys. Lett. **B220**, 308 (1989).

[22] J. D. Jackson. *Classical Electrodynamics, Second Edition.* John Wiley & Sons, 1975.

[23] E. Segrè. *Nuclei and Particles, Second Edition.* Benjamin/Cummings Publishing Company, Inc., 1977.

[24] A. Suzuki *et al.*, Nucl. Inst. and Meth. **A329**, 299 (1993).

[25] Y. Koshio. *Study of Solar Neutrinos at Super–Kamiokande.* PhD thesis, University of Tokyo, 1998.

[26] R. Claus *et al.*, Nucl. Inst. and Meth. **A261**, 540 (1987).

[27] J. George. *Experimental Study of the Atmospheric ν_e/ν_μ Ratio in the Multi-GeV Energy Range.* PhD thesis, University of Washington, 1998.

[28] *The ZEBRA System*, CERN Program Library Long Writeups Q100/Q101.

[29] C. Caso *et al.*, The European Physical Journal. **C3**, 144 (1998).

[30] F. Goebel. *A Study of Particle Identification with the Super–Kamiokande Detector.* Masters thesis, State University of New York at Stony Brook, 1996.

[31] Y. Fukuda *et al.* [Super–Kamiokande Collaboration], Phys. Rev. Lett. **81**, 1562 (1998) [hep-ex/9807003].

[32] R.M. Barnett et al., Physical Review **D54**, 28.6.4 (1996).

[33] G. D'Agostini, hep-ph/9512295.

[34] T.J. Loredo, in *Maximum Entropy and Bayesian Methods*, edited by P.F. Fougère (Kluwer Academic, Dordrecht, The Netherlands, 1990). Available from:
http://d0server1.fnal.gov/users/paterno/public_html/probability/.

[35] I. Bertram, *et al.*, A Recipe for the Construction of Confidence Limits, D0 Note 2775a. Available from:
http://d0server1.fnal.gov/users/paterno/public_html/probability/.

[36] E. T. Jaynes, in *E. T. Jaynes: Papers On Probability, Statistics, and Statistical Physics* , edited by R.D. Rosenkrantz (Kluwer, Boston, 1983)

[37] R. Brent, *Algorithms for Minimization Without Derivatives*, Prentice-Hall (1973.

[38] *The GNU Scientific Library.*, Mark Galassi, talk given at Open Source / Open Science, Brookhaven National Lab., 1999. More information available from:
http://sourceware.cygnus.com/gsl/.

[39] W. Press, *et al.*. *Numerical Recipes in C*, Cambridge University Press, 1992.

[40] L. Peterson, Talk given at NNN99, SUNY at Stony Brook, 1999.

[41] For transparencies and proceedings see:
http://super.physics.sunysb.edu/NNN99/.