Chapter 1

Introduction

The standard electroweak model (SM) describes the interaction of particles (quarks and leptons) via the electromagnetic and weak forces \[\text{?} \, ?\]. This model is based on the gauge group \(SU(2)_L \times U(1)_Y\), with 4 gauge bosons (3 for \(SU(2)\) and 1 for \(U(1)\)) and 2 corresponding gauge coupling constants \(g\) and \(g'\). The input parameters of this model, required for predictive calculations and derived from measurable observables, include

1. the Cabibbo-Kobayashi-Maskawa mixing matrix. The left-handed fermion fields \(\psi_i = (\nu_i^c, l_i^c, d_i^c, u_i^c)\) of the \(i^{th}\) fermion family transform as doublets under \(SU(2)\), where \(d_i = \sum_j V_{ij} d_j\), and \(V\) is the C-K-M matrix.

2. the masses of fermions (quarks and leptons) and Higgs boson.

3. three additional parameters determining the strength of the interaction and the masses of the weak gauge bosons. A particular useful set is:

   - the fine structure constant, \(\alpha\), determined from the \(e^\pm\) anomalous mag-
netic moment, the quantum Hall effect, and other measurements.

- the Fermi coupling constant, $G_F$, determined from the muon lifetime.

- the $Z$-boson mass, $M_Z$, determined from the $Z$-lineshape scan at LEP 1.

$\alpha$, $G_F$ and $M_Z$ are the best measured quantities of electroweak physics and as such are used as inputs in all higher-order calculations. Many different experiments have been performed during the last fifty years to reduce the uncertainties.

The value of $\alpha$ is one of the most precisely determined parameters in physics, obtained from a measurement of the electron magnetic moment. The Quantum Electrodynamics (QED), the theory that describes the interaction of light and matter, provides an extremely precise prediction for the relationship between $g$ and $\alpha$, with only small, well understood corrections for short-distance physics. An improved QED calculation that includes contributions from 891 eighth-order Feynman diagrams now predicts $g$ in terms of $\alpha$ through order $(\frac{\alpha}{\pi})^4$ [8]. A new value for the electron magnetic moment

$$g/2 = 1.001 159 652 180 85 \ (76) \ [0.76 \ \text{ppt}]$$

was obtained by using a one-electron quantum cyclotron [8]. QED provides an asymptotic series relating $g$ and $\alpha$,

$$\frac{g}{2} = 1 + C_2(\frac{\alpha}{\pi}) + C_4(\frac{\alpha}{\pi})^2 + C_6(\frac{\alpha}{\pi})^3 + C_8(\frac{\alpha}{\pi})^4 + \ldots + a_{\mu\nu} + a_{\text{hadronic}} + a_{\text{weak}} \ (1.1)$$

thus the new determination of $\alpha$ is [8], [9].
\[ \alpha^{-1} = 137.035\,999\,070\,98 \pm 0.71\,\text{ppb} \]

This value is measured at momentum \( q = 0 \). In most electroweak renomalization schemes, it is convenient to define a running \( \alpha \) dependent on the energy scale of the process, with \( \alpha^{-1} \sim 137 \) appropriate at very low energy. At high energy scales (a few hundred MeV) the hadronic contributions (to vacuum polarization) become non-negligible. For example, at the energy scale of \( M_Z \), the value of \( \alpha \) is

\[ \alpha(M_Z)^{-1} = 127.918\,18 \pm 0.140\,\text{ppm}. \]

\( M_Z \), the mass of neutral \( Z \) boson, is determined from the \( Z \) lineshape scan at LEP I. It's based on the electroweak measurements performed with data taken at the \( Z \) resonance by the experiments operating at the electron-positron collider LEP. The measurements include cross-sections, forward-backward asymmetries and polarized asymmetries. The results are obtained from the average of four measurements from the ALEPH, DELPHI, L3 and OPAL experiments [?].

\[ M_Z = 91.1875\,21 \pm 0.023\,\text{GeV} \]

The Fermi coupling constant \( G_F \) is extracted from measurements of the muon lifetime, \( \tau_\mu \equiv \Gamma_\mu^{-1} \), which is a purely leptonic process and therefore very clean both experimentally and theoretically. A series of experiments, performed in the 1970s and 1980s, had produced a combined uncertainty on the muon lifetime of 18 ppm, contributing 9 ppm to the uncertainty of \( G_F \). At the time, the theoretical uncertainty was 17 ppm in which the dominant error (15 ppm) came from the second-order radiative corrections and would be completed in the following years. In 1999, the
theoretical calculation was completed, reducing the theoretical uncertainty to 0.2 ppm, making the muon lifetime the dominant uncertainty in determining $G_F$.

In the 1990s, two proposals for the next generation of experiments were submitted to Paul Scherrer Institut in Switzerland (PSI); they are MuLan and FAST experiments. The goal was to improve the precision of the muon lifetime measurement to a relative error of 1 ppm, and once again put the experimental and theoretical uncertainties at the same level.

Starting from 2003, the MuLan experiment, set up at PSI, began to work. Our MuLan experiment collected about $10^{11}$ events in 2004, and $10^{12}$ events in each of year 2006 and 2007. The result from 2004 data analysis was published in 2007. At the time, the precision on the muon lifetime, 11 ppm, reduced the relative error on $G_F$ to 5 ppm. The combined result from 2006 and 2007 data analysis will be published this year, 2010, with expected uncertainty of 1 ppm on the muon lifetime, reducing the error on $G_F$ to the sub ppm level.
Chapter 2

Theoretical Motivation

2.1 The importance of $G_F$ and $\tau_\mu$

The Fermi coupling constant, $G_F$, plays an important role in precision tests of the Standard Model of electroweak interactions. $G_F$ is one of the few quantities that is sensitive to physics at very high energy scales. $G_F$ can be related to $\alpha$, $M_Z$, $M_W$ and by inverting the relation one can predict $M_W$ through $M_Z$ which is measured much more accurately. This $M_Z - M_W$ interdependence can be then confronted with experimental value $M_W^{\text{exp}}$.

The most precise determination of $G_F$ is based on the mean lifetime of the positive muon. The importance of the muon decay is due to its pure leptonic process which is related only to the weak interaction. The lifetime of the negative muon is affected by target material due to the possibility of nuclear reaction between $\mu^-$ and nucleus.

The discovery of nuclear $\beta$ decay, characterized by a much slower decay rate than strong or weak decays, followed by the discovery of parity non-conservation in weak interaction, paved the way for the foundation of the theory of the weak interaction.
2.2 Fermi theory

The first theory of $\beta$ decay, the origin of our present theoretical ideas about all of the weak interactions, was presented by Enrico Fermi, who formulated it in close analogy to QED. In the Fermi model the weak interactions are represented by a contact interaction of four fermions, two leptons and two neutrinos, as illustrated in figure 2.1.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fermi_diagram.png}
\caption{Feynman diagrams for muon decay: the local four-fermion model as described by Fermi.}
\end{figure}
The Lagrangian, relevant for the calculation of the muon lifetime in the Fermi theory is

\[ L_F = L_{QED}^0 + L_{QCD}^0 + L_W \]

where \( L_{QED}^0 \) is the usual bare Lagrangian of QED and \( L_{QCD}^0 \) is the bare QCD Lagrangian responsible for strong interactions. The Fermi contact interaction that mediates muon decay is

\[ L_W = -2\sqrt{2}G_F \langle \overline{\Psi}_\mu \gamma_\lambda \gamma_L \Psi_\mu \rangle \langle \overline{\Psi}_e \gamma_\lambda \gamma_L \Psi_{\nu_e} \rangle \]

in which \( \Psi \) are wave functions for the leptons and their associated neutrinos, and \( \gamma_L \) denotes the Dirac left-hand projection operator.

Based on Fermi theory, the lifetime of the muon, defined as the inverse decay rate, is simply

\[ \frac{1}{\tau_\mu} \equiv \Gamma_\mu = \Gamma_\mu^0 (1 + f(\frac{m_e^2}{m_\mu^2})) \]

where

\[ \Gamma_\mu^0 = \frac{G^2 m_\mu^5}{192\pi^3} \]
\[ f(x) = -8x - 12x^2 \ln x + 8x^3 - x^4. \] (2.5)

In the above equations, the mass of the neutrinos can be neglected; the decay phase space only involves the electron and muon masses and is accounted for in \( f\left( \frac{m_e^2}{m_\mu^2} \right) \). The strength of the interaction is represented by the coupling constant \( G_\mu \). The coupling constant which describes \( \mu \) decay is very close to that describes other weak interactions such as \( \beta \) decay, and it’s believed that, within the framework of the standard model, the coupling constant which describes the strength of all weak interactions, is a universal constant, known as the Fermi coupling constant, \( G_F \). \( G_\mu \) is the best estimate of \( G_F \) because muon decay is a pure leptonic process. Comparison of the Fermi constant extracted from various measurements stringently tests the universality of the weak interactions’ strength.

### 2.3 Modern Feynman diagrams

In the modern description the weak interaction no longer takes place at a single point, but rather interacts via the propagation of a boson. In the case of positive muon decay, it takes place through the \( W^+ \) boson, as illustrated in figure 2.2. The non-zero mass of the \( W \) boson contributes to the matrix element and modifies the decay rate. But its much heavier mass than the muon mass suppresses the effect and
the correction is at the order of \( \frac{m_\mu^2}{M_W^2} \), which is about 1 ppm correction. In the modern theory the strength of the weak interaction is governed by weak coupling constant \( g_w \), which is related to \( G_F \) by

\[
\frac{G_F}{\sqrt{2}} = \frac{g_w^2}{8M_W^2}
\]

(2.6)

Figure 2.2: Feynman diagrams for muon decay: modern diagram mediated by \( W^+ \) boson.

2.4 Radiative corrections

When high precision calculation is required, the higher orders of Feynman diagram must be included. These higher order items are mostly dominated by radiative
corrections, resulted from the virtual photons. The lowest-order of the radiative correction is dominated by virtual photons between two massive particles, muon and positron, as illustrated in figure 2.3.

![Feynman diagrams for muon decay with radiative correction](image)

**Figure 2.3:** Feynman diagrams for muon decay with radiative correction.

To leading order in $G_F$ and all orders in $\alpha$ the formula for the muon lifetime, $\tau_\mu$. 
by means of $L_F$ is

$$
\frac{1}{\tau_\mu} = \frac{G_F^2 m_\mu^5}{192\pi^5} (1 + \Delta q) \quad (2.7)
$$

where $\Delta q$ encapsulates the higher order QED corrections and can be expressed as a power series expansion in the renormalized electromagnetic coupling constant $\alpha_r = e_r^2/4\pi$.

$$
\Delta q = \sum_{t=0}^{\infty} \Delta q^{(i)} \quad (2.8)
$$

The above equations are based on local four-fermion model, neglecting effects from the mediated boson. When all the corrections are included, the Fermi constant $G_F$ is related to electroweak gauge coupling $g$ by

$$
\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} (1 + \Delta r) \quad (2.9)
$$

where $\Delta r$ represents the weak-boson-mediated tree-level and radiative corrections, which have been computed to second order.

In an analogous way to $\Delta q$, $\Delta r$ can be expressed as a power series in $\alpha_r$

$$
\Delta r = \sum_{t=0}^{\infty} \Delta r^{(j)} \quad (2.10)
$$
In which

\[ \Delta r^{(0)} = \frac{3m^2}{10M^2_W} + \xi \left( \frac{m_e^2}{m_\mu^2} \right) \]  

(2.11)

is due to W propagator effects and shifts the extracted value of \( G_F \) by 0.52 ppm.

### 2.5 Errors on \( G_\mu \) is dominated by \( \tau_\mu \)

The relative error on \( G_\mu \) can be expressed as

\[ \frac{\delta G_\mu}{G_\mu} = \sqrt{\left( \frac{1}{2} \frac{\delta \tau_\mu}{\tau_\mu} \right)^2 + \left( \frac{5}{2} \frac{\delta m_\mu}{m_\mu} \right)^2 + \left( \frac{1}{2} \frac{\delta \Delta q}{\Delta q} \right)^2}, \]

(2.12)

where \( \frac{\delta \Delta q}{\Delta q} \) comes from the radiative corrections.

The error on muon mass in atomic mass units is

\[ m_\mu = 0.1134289264 \text{ (30) u [26 ppb]}, \]

and the uncertainty in the conversion factor from atomic mass units to MeV is

\[ 1 \text{ u} = 931.494043 \text{ (80) MeV [86 ppb]}, \]

so the uncertainty on the determination of \( G_\mu \) from muon mass is 0.22 ppm which is pretty small.

Before 1999, the dominant error came from QED radiative corrections to \( \Delta q \) at the level of about 15 ppm. Now the one-loop and two-loop corrections have been obtained and the relative theoretical uncertainty in the determination of \( G_F \) is onto
the order of a few tenths of a ppm and is comparable to the uncertainty from the muon mass.

The improvements on $\Delta q$ leaves the muon lifetime as the largest contributing factor to the uncertainty on $G_{\mu}$.

$$G_F = 1.166367(5) \times 10^{-5} \text{ GeV}^{-2}.$$
Chapter 3

Muon Physics

3.1 Basic Properties of the Muon

Muons are unstable elementary particles of two charge types (positive $\mu^+$ and negative $\mu^-$ having a spin of 1/2). The muon mass, about 207 $m_e$, determined most accurately by measuring the energy interval between the electronic quantum levels in the bound states. The muon lifetime, about 2.197 $\mu$s, was studied by many experiments during the past decades and current goal of the precision measurement is 1 ppm level. The muon decay mode includes normal muon decay and rare muon decays. Almost 100% of muons decay to positron (or electron) and two neutrinos. There is also rare possibility for muons to decay to positron (or electron) and $\gamma$ ($< 1 \times 10^{-8}$).

3.2 Muon production and spin polarization

Experimental studies in muon science can only be conducted when a reasonably intense muon beam of high quality is available. At very high energy (> 100 GeV) cosmic-ray muons are the only possibility, whereas at low energy, accelerator- pro-
ducing muons are almost exclusively used. Muons can only be obtained from the
decay of pions which are produced in nuclear interactions between accelerated par-
ticles and nuclear targets. The pion decay is a simple two body decay which can be
expressed as

$$\pi^+ \rightarrow l^+ + \nu_l,$$  \hspace{1cm} (3.1)

where $l = e$ or $\mu$. In these two decay modes, the decay to the much lighter positron
involves a larger phase space and therefore would appear to be favored over the decay
to a far more massive muon. But special dynamical considerations heavily suppress
the positron decay mode. The pion has spin 0, so the positron and neutrino must
emerge with opposite spins, and hence equal helicities ($h = \vec{S} \cdot \vec{p}$). The neutrino is
always left-handed, so the positron must be left-handed as well. But if positron were
massless, then (like the antineutrino) it would only exist as a right-handed particle
and the decay $\pi^+ \rightarrow e^+ + \nu_e$ would not occur at all. The positron mass is sufficiently
massless such that the decay is heavily suppressed. Thus the pion decay is the good
process for the muon production.

The calculated branching ratio is

$$\frac{\Gamma(\pi^+ \rightarrow e^+ + \nu_e)}{\Gamma(\pi^+ \rightarrow \mu^+ + \nu_\mu)} = \frac{m_e^2(m_e^2 - m_\pi^2)^2}{m_\mu^2(m_\pi^2 - m_\mu^2)^2} = 1.28 \times 10^{-4}$$  \hspace{1cm} (3.2)
In the two-body decay, the momentum of the muon is given by

\[ P_\mu = \frac{m_\pi^2 - m_\mu^2}{2m_\pi^2} c = 29.79 \text{ MeV}/c. \]  \quad (3.3)

Because the pion has spin zero the decay muons are emitted isotropically. The spin and momentum directions are always anti-parallel, and if we collect decays over only a small solid angle, we still get muons that are highly polarized.

In the MuLan experiment, we used surface muons, where the muons are emitted from the surface of production target with little energy loss. These surface muons keep the highest momentum from the decay of pion at rest and have the highest polarization compared with those pions produced in the center of the target.

### 3.3 Muonium

A high precision measurement of the muon lifetime requires free muons that are stopped completely in the target. But a positively charged muon can be bound to electrons in the target, forming a muonium atom ($\mu^+e^-$), which behaves much like a hydrogen atom. The fraction of muonium formation is highly dependent on the target material. The bound state muonium will affect the total decay rate of the free muon. Fortunately for the MuLan experiment it has been shown that the lifetime of muonium is different from that of free muon by only about 0.6 parts per billion, and would not affect the MuLan goal of 1 ppm.
3.4 Differential decay of the muon

The muon decay is a three-body decay process so that the energy of decay positrons is not uniform. Because of parity violation in the weak decay of the muon, the muon spin and the direction of the decay positrons are correlated. The differential probability for the decay positron to have a normalized energy, \( y = E/E_{max} \) \( (E_{max} = 52.8 \text{ MeV}) \), at angle \( \theta \) with respect to the spin of the muon is

\[
\frac{dP(y, \theta)}{dyd\Omega} = \frac{1}{2\pi} n^*(y)[1 - \alpha^*(y) \cos \theta], \tag{3.4}
\]

with relative event number \( n^*(y) \) and asymmetry function \( \alpha^*(y) \) given by

\[
n^*(y) = y^2(3 - 2y), \tag{3.5}
\]
\[
\alpha^*(y) = \frac{q 2y - 1}{e 3 - 2y} \tag{3.6}
\]

The above equations hold in the muon rest frame, which is the case in the MuLan experiment. \( n^* \) and \( \alpha^* \) are plotted in figure 3-1. It is clear that the probability of the muon decay to positrons with \( y > 0.5 \) is greater than that with \( y < 0.5 \) \((\frac{dP(y)}{dy} \propto n^*(y))\). The asymmetry function \( \alpha(y) \) is negative when \( y > 0.5 \), whereas it is positive when \( y < 0.5 \). So the direction of the decay positron tends to be in the same as that of the muon spin. This can be seen intuitively: at the endpoint (maximum positron energy), the two neutrinos are emitted antiparallel to the positron so that
their spins are canceled. Thus the muon spin is entirely transferred to the positron. Within the standard weak interaction theory, the positron also carries right-handed helicity. At this energy the positron velocity is so close to the velocity of light that the longitudinal polarization is always equal to the helicity.

Integrating with energy weighting, gives an average asymmetry of

$$< a > = \int_{0}^{1} n(y)a(y)dy = \frac{1}{3}$$  \hspace{1cm} (3.7)

The positive sign means that the positron is preferentially emitted with respect to the muon spin.

**Figure 3.1:** Event number and asymmetry function in the muon rest frame.
3.5 $\mu$SR effects

With an ensemble of muons stopping at the same time in the target and decaying with time, the average polarization should be considered instead of the spin of single muon. The consequence is that the $\cos \theta$ term of Eq. || is replaced by $P \cos \Theta$, where $P = \vec{P}$ is the magnitude of the average of all polarizations and $\Theta$ is the angle between the polarization and decay direction.

Uncontrolled precession of the stopped, polarized muon ensemble will cause a rotation of the angular distribution of decay positrons, and any angular variation in detector response will then modify the ideal exponential time spectrum. If uncorrected this modification would affect the muon lifetime, especially in a precision measurement. To reduce the spin-related modification, a magnetic field was imposed on the target. The purpose is to make the muons which trickle into the target during the accumulation period rotate very fast before they decay. In this way all stopped muons have spins with different directions and the net polarization will be nearly 0.

The existence of magnetic field will cause muon spin rotation and dephasing. Under an external magnetic field, particles with non-zero spin will undergo Larmor precession, with the magnitude of the spin polarization unchanged, at an angular frequency given by

$$\omega = \gamma B$$

(3.8)

where $\gamma = \frac{e \gamma}{2m}$ is the gyromagnetic ratio and $B$ is the magnitude of the magnetic
field and $g$ is approximately 2.

From this equation it's easy to express the frequency in terms of $B$, $f = \omega/2\pi \approx B \cdot 135\text{MHz/T}$. As we can see, a small magnetic field of only 3.4 mT will bring a rotation period of about 2.2 $\mu$s, equal to the muon lifetime. The Larmor precession period in the earth's magnetic field ($\approx 50 \mu$T) is about 150 $\mu$s. In the MuLan experiment, we make the muons rotate at a much faster frequency. The continuous beam ensures the ensemble of muons are accumulated in 5 $\mu$s, and in the mean time the muons rotate at different starting time, so the average of spin of all the muons that are collected is close to 0. For example, the Arnokrome-3 (AK-3) target has a very high internal magnetic field of about 0.5 T, which results in a thorough dephasing of the spin directions.

In addition to the rotation of the spin, the magnitude of the polarization also decreases with time, the so-called relaxation. There are two components describing the relaxation. The reduction of the polarization that is independent of the target material and magnetic field is described by decay time $T_1$. This longitudinal depolarization happens when the spins exchange during collisions within the medium. The other term, characterized by time $T_2$, is caused by a local inhomogeneous field inside the material. This transverse depolarization is caused by spins precessing at different rates, reducing the magnitude of the average polarization of the ensemble. There is no exact functional form for the relaxation components, so we simply replace
the constant terms with the empirical time dependence.

\[ f(t) \rightarrow f(t)[1 + aP(t) \cos < \omega > t + \phi)]. \tag{3.9} \]

where \(< \omega >\) is the mean frequency of the precesser and \(P(t)\) is the time dependent transverse spin polarization.

The initial spin polarization can also be reduced from unity when the muon comes to stop in a material. The exact mechanism of this kind of depolarization is unknown, but it is believed to be related to the multiple local field and muonium state.

3.6 rare decays

More than 99% of muons follow the normal decay to one lepton and two neutrinos. In addition there are mainly two rare decays, with one decay to positron and gamma light, and the other to 3 leptons, as showed in the following formulas:

\[ \mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu + \gamma, \quad 1.4\%, \tag{3.10} \]
\[ \mu^+ \rightarrow e^+ e^- e^+ \nu_e \bar{\nu}_\mu, \quad 0.0034\% \tag{3.11} \]

These rare decays cause “double” or tripple hits at one time but thye do not
affect our measurement because their probability is stable early to late, the main
effect is the $\chi^2$ of the overall fit.
Chapter 4

History of measurement of Muon lifetime

4.1 The basics idea of $\tau_\mu$ experiment

The importance of a precise muon lifetime measurement has encouraged several experiments with different methods during the last thirty years. The current world average is based on six experimental results. Three experiments (Balandin, Giovanetti, Bardin) that took place during the 1970's and 1980's reached an uncertainty of 20 ppm and two experiments (FAST, MuLan) in 2000's will improve it to the 1 ppm level.

All the precision measurements of muon lifetime will consider the following things:

To overcome the statistical limitation for a precision measurement of the muon lifetime, a bunch of muons were stopped in a target during a time window. The experiments listed above employed two different temporal beam structures. Giovanetti and most used a one-at-a-time method, where the average rate implied one muon entering the target every 50 $\mu$s. MuLan used a many-at-a-time method, where the continuous beam was chopped to come as a burst every some microseconds.
The one-muon-at-a-time method above allows more than one muon present at anytime, this will increase the background and reduce the sensitivity of the experiment. So we must control the beam rate at low enough to make sure the interarrival time between muons longer than the muon decay time. The pulsed beam method is a better choice because many muons can be collected in each burst, and the data rate can be much higher. But in either method there is a pile up problem related to beam rate. When more muons come, the possibility of count loss due to pile up also increases, creating a systematic error. So all experiments use a segmented detector. The more individual detectors around the target, the better.

If the stopped muons have a net spin polarization, and the spin precesses, spin decay angle correlations will create a non-perfect exponential. So some experiments have used pion beams since the pion has spin 0 and its decay is isotropic. But there is still some contamination of polarized muons in a pion beam. So all experiments applied a magnetic field to make the spin rotate fast enough to dephase the original spin orientation. In addition, the detector system is not only segmented to overcome the pile up issue, but also symmetric around the target. This way the combination of all segments will cancel the oscillation term in the decay spectrum.

4.2 History of $\tau_\mu$ experiments

Balandin used the pulsed pion beam from the Saclay Linear Accelerator. The positive pions (140 MeV/c) were stopped in a Sulfur target during 3 $\mu$s beam burst, and then
decayed to muons which decayed to positrons during the next 65 μs beam off time. The beam still contained about 5% polarized muons from pion decay in flight. The sulfur target has enough internal magnetic field to disturb the spin orientation and the symmetry of 6 detectors around the beam axis further suppresses polarization effects. The lifetime of the muon measured by Balandin is

\[ \tau_{\mu}^{+} = (2.197078 \pm 0.000073) \mu s. \]  (4.1)

The Giovanneto experiment was performed at the Tri-University Meson Facility (TRIUMF) in Vancouver, where the protons arrived at the production target in 2-5 ns bursts every 43 ns. The experiment used the \( \pi^{+}/\mu^{+}/e^{+} \) chain. The target was a water-filled container, a 48-cm long, 48-cm-diameter stainless-steel cylinder, which stopped both \( \pi^{+} \) and \( \mu^{+} \). The decay positrons were detected through their Cherenkov radiation. The time of arrival of beam particles was signaled by coincidences between two scintillation counters S1 and S2. The main part of the beam, the \( \pi^{+} \), was stopped at about 8 cm upstream from the center of the counter. The decay \( \mu^{+} \) came to rest after traveling about 0.14 cm and then decayed into \( e^{+} \). The angular distribution of these \( e^{+} \) was isotropic. Several efforts were made to reduce the systematic errors. The \( \mu^{+} \) contamination in the beam would bring spin-correlated decay positrons, which would shift with time the medium-energy region of isotropic \( e^{+} \). To reduce the spin procession, the detector was wrapped with a magnetic shield, keeping the field below 0.05 Gauss. The depolarization time of muons in water exceeds 100 μs,
so its effect can be neglected. The final result on the $\mu$ lifetime was

$$\tau = 2.19695 \pm 0.00006 \ \mu s. \quad (4.2)$$

The FAST experiment aimes to measure the muon lifetime at the 1 ppm level. A DC $\pi^+$ beam is stopped inside a plastic scintillation-fiber target, which was viewed by position-sensitive photo-multipliers (PSPM’s). The beam was operated at the rate of 1 MHz, bringing about 30 $\pi^+ - \mu^+ - e^+$ decay chains in each time slice of 30 $\mu$s.

To reduce the systematic errors from muon spin rotation, the detection of stopped $\pi^+$ was required, which forces the polarization of the muon source to be isotropic. Moreover the detector was symmetric with large solid angular acceptance. But even in the case of a properly-detected $\pi^+ - \mu^+ - e^+$ chain, the $\mu$ contamination still exists, with a corresponding $\mu SR$ effect.

The timing of the incoming beam particle was defined by a set of 3 plastic scintillation counters located upstream of the target. The target is a scintillator block, made of 48 $\times$ 32 pixels, serving as both stopping material for the pion and detector for the decay products. The first FAST result in 2007 yielded

$$\tau = 2.197083 \pm 0.000015 \ \mu s. \quad (4.3)$$

RIKEN-RAL

26
Chapter 5

Systematic Error Considerations

The basic idea of MuLan experiment is simply to count muon decays vs. time. The design of the experiment is a bit complicated, with the consideration of collecting enough muons (at least $10^{12}$ for 1 ppm measurement) in short term (no more than 2 months for the real data collection). Two main systems operate in series to perform the measurement: the muon collecting system includes kicker, target and muon corridor. the decay positron detecting system includes 340 scintillator-PMTs, WFD.

Any effect that could change the response of the systems during the measurement period would perturb the muon lifetime measurement. And most of the systematic error considerations focus on “early-to-late” effect, which can be defined as an effect that changes the magnitude from early to late. By early and late we mean the early and late in the measurement period.
5.1 Kicker Stability

During the measurement period, the kicker is turned on to bend away the continuous muon beam so that no muons can enter the Muon corridor. In reality the kicker does not shut off the beam completely so there are still some muons sneaking in the detector system (MuLan ball). These sneaky muons, low intensity but from continuous beam, will contribute to the flat distributed background in the muon lifetime spectrum. They can be characterized by the extinction factor, defined as N/B from the fit of the lifetime spectrum (where N is from the normal muon decay and B is from the flat background). The kicker operates at high voltage and any perturbations on the voltage will translate to the change in extinction, or the flat distributed background. The drift of the kicker voltage due to capacitive charging effects could result in a non-flat background. In addition, any mechanical instability (such as bending of the plates, vibrations etc.) could also lead to a change in the background during the measurement period. In one word, any change in the flat background during the beam off period has the potential to shift the measured muon lifetime.

If there were no change in the incoming muon rate when the kicker was on, the number of muons in the target would follow

\[ \frac{\Delta N}{\Delta t} = \frac{1}{\tau} N + R_{off} \]

where \( \tau \) is the lifetime and \( R_{off} \) is the muon rate during beam off.

The number of muons in the target versus time would be:
\[ N(t) = N_0 e^{-t/\tau} + R_{\text{off}} \tau (1 - e^{-t/\tau}) \]

where \( N_0 \) is the number of muons at time 0 in the measurement period.

For most systematic error consideration, the MuLan experiment assumes the simple fitting function for the lifetime spectrum, the exponent decay with flat background:

\[ N(t) = N_0 e^{-t/\tau} + B \]

and the unknown distortion of flat background has additional term of

\[ \epsilon N e^{-t/Z} \]

where \( Z \) is the time constant of the distortion and \( \epsilon \) is the magnitude relative to the muon decays.

Here there are two ingredients to study:

1. The change of the lifetime \( \Delta \lambda \) for a given magnitude \( \epsilon \) and time constant \( Z \), which can be done by the simulation.

2. The estimate of parameter \( \epsilon \) and \( Z \), which can be extracted somehow from the real data.

It was studied by Ron and Peter that, for tiny distortion (\( \epsilon \approx 1 \text{ppm} \)) the lifetime change scales linearly with \( \epsilon \).

\[ \Delta \lambda(Z) = \epsilon \cdot \frac{\Delta \lambda(0, Z)}{\epsilon_0} \]

where \( \Delta \lambda = \frac{\lambda_{\text{fit}} - \lambda}{\lambda} \) when fitting the lifetime histogram with the 3-parameter function.
It was found by the simulation that for $Z=22\mu s$ and $\epsilon=1\text{ppm}$, $\Delta\lambda = 5 \text{ ppm per } 22\mu s$.

The maximal size of $\epsilon$ can be deduced from the data. To do so, we need

1. the maximal voltage change $\Delta V$ during the measurement period, which can be deduced from our kicker trace measurements.

2. the relationship of voltage change with extinction ratio, which can be achieved from data. (extinction ratio is the background to signal ratio)

$$\epsilon = \Delta V \cdot \frac{\Delta \alpha}{\Delta V} \mid V = 25kV$$

In order to determine the maximal kicker voltage drift $\Delta V$ we measured the kicker traces with a high voltage probe both in 2006 and 2007. The kicker traces were averaged and saved with the oscilloscope.

Figure ?? shows the recorded kicker traces for 2006. The oscillation is probably from the high voltage probe parameters and is not real in terms of kicker voltage. The best fit function turns out to be an exponent plus poll (linear function).

$$g_Z(t) = A \cdot e^{-t/Z} + Bt + C.$$  

The extracted maximal voltage change for a given time constant $Z$ is

$$\Delta V(Z) = |g_Z(t \approx 5\mu s) - g_Z(t \approx 27\mu s)|,$$ where the boundaries of measurement period at $t \approx 5\mu s$ and $t \approx 27\mu s$ were adapted to each measurement.

To extract the relationship between $\frac{\Delta \alpha}{\Delta V}$ and $\Delta V$, we measured the background to signal ratio (extinction factor) at different kicker voltages. The extinction factor
can be easily obtained by fitting the lifetime spectrum at each kicker voltage setting. Figure ?? shows the kicker extinction scan for 2006 (left) and 2007 (right). They simply follow the linear relationship, especially near the operation voltage at 25 KeV. The final results for the slope of the linear fit, or \( \frac{\Delta \frac{\sigma}{N}}{\Delta V} \), are 6.3 \( \cdot 10^{-8} \) V\(^{-1}\) for 2006 and 4.2 \( \cdot 10^{-8} \) V\(^{-1}\) for 2007.  

0.1 ppm and 0.04 ppm for 2006 and 2007.

5.2 Non-Target Muon Stops

The non-target-stop muons, also called errant muons, are defined as muons that are not stopped in the target, or the center of the detector system. These errant muons can come from the muons that stop in the walls of the beam pipe upstream of the target or from the muons that are backscattered from the target. The average spin of the errant muons are not zero and will not be canceled by the symmetry of the detector system. The lifetime spectrum will be distorted by the spin precession.

We have performed two different studies on systematic error from the errant muon stops.

Before the detector (MuLan detector ball) was present, we used a four-fold coincidence between four scintillators to look at a narrow section of the beam pipe. The study estimates that there were as many as about 1 \( \times 10^{-3} \) of the total muon were stopped upstream. During the data acquisition, we also placed the parallel counters aligned with the beam pipe. These counters, being coincidenced with one tile pair of
the ball detectors, give a coarse view of the upstream muons stops in the beampipe. This study estimates that about 1% of the muons were decayed before reaching the target.

5.3 Detector Timing Stability

5.4 Detector Gain Stability

The fluctuations of energy losses by ionization of fast charged particle in a thin layer of matter was first and best described by Landau function. They give rise to a universal asymmetric probability density function characterized by a narrow peak with a long tail towards positive values.

The

5.5 Detector Gain vs. time after pulse

This study is focused on the case when there are two or more pulses in the same fill. The second pulses by themselves are harmless, having the same lifetime as the trigger pulse and therefore do not distort our measurement. However, the amplitude of the second pulse will be affected by the trigger pulse due to the response of the detector system (scintillator and PMT) and the pile up correction will be affected correspondingly.

In David’s pile up correction study, he used the pulses from different fills to build
up pile up pulses and most of these pulses are independent from each other because they are in different fills. But the real pile up is from pulses from the same fill in which the second pulse will be affected by the trigger pulse. If the amplitude of the second pulse is not stable vs. dt after trigger pulse, and we have the amplitude threshold cut, we may lose or gain pulses near the threshold area.

Figure ?? shows the amplitude of the second pulse vs. time after pulse. The black is from the trigger fills where two pulses are in the same fill, and the red is from the shadow fills where the pulses are in different fills.
Chapter 6

PSI Accelerator System

The MuLan collaboration decided to use the surface muons which are produced from the pion decay at the surface of the target. The target, made of carbon, was struck by the continuous proton beam, which are boosted in energy by the PSI acelerator.

6.1 PSI Accelerator for proton

Protons are accelerated in three steps in the PSI facility which consists of three accelerators in series. A Cockcroft-Walton accelerator, which also contains the proton source made up of hydrogen atoms, is used as the first stage. Protons are fed into Injector II of a pre-accelerator - a smaller ring cyclotron, inside which the protons reach to approximately 37% of the speed of light (equivalent to a kinetic energy of 72 MeV). The final stage the core of this facility, is the large ring cyclotron with a diameter of approximately 15 meters, in which protons are accelerated to their terminal speed of almost 80% of the speed of light (590 MeV). the PSI accelerator facility has reached a proton beam of 2200 microamperes.
Protons from accelerator hit protons (or neutrons) in a target nucleus and produce pions by the strong interaction

\[ p^+ + p^+ \rightarrow p^+ + n + \pi^+ \]  \hspace{1cm} (6.1)

The excess energy in the center-of-mass frame of the collision must exceed the rest mass of a pion, so we have a threshold energy for the incoming protons of about 290 MeV, below which no pions can be produced at all. To get a reasonable yield we need to increase the energy to 500 MeV or more. At higher energies double pion production (threshold of 600 MeV) or even higher numbers of pions are possible. In PSI, the muons originate from the decay of pions produced by the interaction of the 1.8 mA \( \sim \) 2.2 mA, 590 MeV proton beam with the nuclei of the target material.

### 6.2 Proton Target

The pion production is basically an interaction of proton beam with the individual protons or neutrons in the target, so it is independent of the atomic number \( Z \) of the target material, for a given target mass in the beam. Other unwanted effects increase with increasing atomic number or mass - neutron production by spallation or x-ray and gamma ray production by bremsstrahlung processes. Also, related to this, the beam scattering is stronger with higher atomic number nuclei. So a low atomic number is preferred.
Heat must be dissipated from the target. The nuclear collisions result in high energy particles, many of which will escape the target and end up in collimation and shielding, and energy lost by the transmitted beam is all dissipated in the target. So that suggests a material with a high melting point, and low vapor pressure as it’s likely to be in vacuum, and high thermal conductivity, and preferably low thermal expansion so the shock of turning the beam on, or even a single beam, doesn’t crack the target. So we might consider carbon (graphite) or beryllium. Lithium would melt too easily.

A higher density would actually be good to produce a compact target, and a small spot size, so we can’t use H or He gas targets which would have far too low pion yield per unit volume.

Based on the discussions above, two carbon-ring targets are chosen as muon source, as shown in figure ???. The target was assembled with metal stand and rotated with 1 Hz for cooling. The rotating wheel design is good for small proton beam spot. The carbon ring is cone like with $45^\circ$ around the rotating axis. The proton beam is parallel to the target surface and edge-on through material and muons are extracted at $10^\circ$ and $90^\circ$. 

36
6.3 Surface Muons

Pions are generated during the collisions between protons and carbon nuclei. A pion then decays to muon and neutrino after about 26 ns, by the weak interaction

\[ \pi^+ \rightarrow \mu_+ + \nu_\mu \]  

(6.2)

Surface muons originate from pions stopped close to the surface of the target. They have kinetic energies around 4.1 MeV (momentum 28 MeV/c). The pions are produced wherever the proton beam intersects the target, and start with a wide range of momenta in all directions. Some of the slower pions stop in the target, close to where they were produced. That means that if we have a small beam spot on a large target, the muon spot is still well defined. Once the stopped pions decay, the range of the 29 MeV/c muons in the target is quite short (3/4 mm) in carbon. Since the beamline can only accept a limited range of momentum, such as 10%, this means that we only collect muons from a 0.2 mm thick surface layer. Pions produced deep in a thick target will probably stop well below the surface, or if they have higher energy will escape completely, so the muon yield will go up less than linearly for thick targets. The optimum target may be a thin plate parallel to the beam, with the muons extracted normal to it.

There are two crucial characters of surface muons: Firstly, all muons generated in this manner are polarized, i.e. they all spin in the same direction. Secondly, muons
generated in this way are so slow (29.8 MeV/c) that during the experiment they come to a standstill in the sample being investigated (easily stopped in the muon target).

The pion beam is contaminated by a large number of positrons originating from the decay of $\pi^0 (\pi^0 \rightarrow 2\gamma$, and $\pi^+ \rightarrow \mu^+ \rightarrow e^+$). In order to obtain a muon beam of high purity, the positron contamination must be eliminated in the beam channel. The $e^+$ contamination can be removed by installing a DC electrostatic separator at a suitable part of the beam channel. A DC separator with vertical electric field serves not only to remove contamination positrons but also, when combined with a crossed magnetic field, to rotate the spin polarization direction of the muon beam.

Among the 6 experimental areas (piM1, piM2, piM3, piE1, piE3, and piE5) the piE3 beam line delivers pions and muons in the momentum range from 10 to 250 MeV/c. It is designed to match the optical characteristics of the low energy pion spectrometer LEPS, and it is the only beam line with a vertical bending plane, the experimental area being 6 m above the floor level. The channel views the thick target TE at an angle of 90, images more than half of the 6 cm long target, and is therefore also an excellent beam line for surface and cloud muons. Surface muons are operated in achromatic mode with momentum and intensity determined by quadruples and slits. The separator is used to distinguish muons from other unwanted background particles.
Figure 6.1: carbon target to produce muon at PSI
Figure 6.2: Beam line
Figure 6.3: Beam line 2
Chapter 7

MuLan Detector System

7.1 Basic Idea of MuLan Experiment

The basic idea of MuLan experiment is to count muon decays vs. time. The design of the experiment is a bit complicated, with the consideration of collecting enough muons (more than $10^{12}$ for 1 ppm measurement) in a short time (no more than 2 months for the data collection). The traditional DC method, detecting one muon decay at a time, would require more than one year for a 1 ppm statistical error. The MuLan method, using a time structured muon beam to measure a bunch of muon decays in each period (20 muons per 27 μs super-cycle), requires only about one month to achieve the same goal.

The design of the experiment is driven by systematic error considerations. Any effect that could change the response of the systems during the measurement period (of 22 μs cycle) may perturb the muon lifetime measurement. Most of the systematic error considerations focus on “early-to-late” effects, which can be defined as effects that change in magnitude from early to late in the measurement period. The primary
systematic concerns include:

1. muon spin precession, in view of the highly polarized muon beam,

2. multi-particle pileup,

3. muon decays outside the fiducial volume of the detector,

4. background, which is dominated by sneaky muons (unexpected muons sneaking into the detector system during the measurement period) and back-scattered muons (from the target),

5. detector gain and timing stability,

6. kicker high voltage stability.

Two main systems operate in series to perform the measurement: the muon collecting system includes the kicker, target and muon corridor; the decay-positron detection system includes 340 segmented sub-detectors. Each detector consists of scintillator, light guide, photomultiplier tubes (PMTs) and waveform digitizers (WFD). For systematic studies’ purpose, there are also various devices installed inside or around the detector system. For example, the Entrance Muon Counter (EMC wire chamber) was installed at the end of the beampipe, to monitor the beam profile after every eight hour shift. Plastic scintillators were installed near the entrance of and on the wall of the beam pipe, to monitor errant muon stops.
7.2 Detector System as a Whole

The basic setup of the whole detecting system after beam line is shown in figure 7.1. A continuous muon beam formed from pion decay is brought to our experimental area. During an accumulation period of 5 $\mu s$, a stream of approximately 20 muons is brought to rest in a thin target. The muon beam is then switched off (by the kicker), and decay positrons are recorded by a surrounding detector (the MuLan "ball") during a measuring interval lasting approximately 10 muon lifetimes (22 $\mu s$). This cycle is repeated until more than $10^{12}$ decays are recorded. The time-structured muon beam is created by means of a high-frequency, high-voltage electrostatic kicker. During the measuring interval, the Michel positrons are recorded by a highly-segmented, symmetric detector, featuring 170 independent scintillator tile pairs. Each element is read out by a photomultiplier tube (PMT), whose signal is sampled at 450 MHz by a dedicated waveform digitizer (WFD) channel. The time of arrival and energy deposited in each tile are derived from the WFD record, which was written to the tape by the data acquisition system (DAQ). The whole system is controlled by the so-called magic box, an electrical device sending time structured and ordered signals to the kicker, the WFD and the DAQ to make sure they work together consistently. Decay time histograms, constructed from coincident hits, are then fit to extract the lifetime. The entire analysis is performed blind: the analysis team is not given the frequency of the master clock until the analysis is complete. Additionally, individual analyzers report their results with a secret, personal offset. In this way, we eliminate
unconscious bias toward the world average while preparing our results.

![Diagram](image)

**Figure 7.1:** The MuLan experiment setup overview. The main system parts include: beam line, kicker, target, segmented detectors (scintillators, PMTs and WFDs), EMC and DAQ.

### 7.3 Constructing the time structured muon beam

The MuLan method requires a fast beam line kicker, which can turn the beam on and off, to impose an artificial time structure on the continuous beam. The kicker needs to run with a standard “on-off time cycle” with a fast transition time to high voltage.
The difficulty of the design of the kicker is the requirement of a voltage difference as high as 25 KV and rise and fall times of 45 ns. The high voltage is required to bend the beam away completely during the measurement period. The fast rise and fall times are designed to minimize the sneaky muons which may enter the detector system during the transition period. Sneaky muons are unwanted because they contribute to non-flat background and affect the muon lifetime measurement. Figure 7.2 shows the time structure of the kicker’s operation.

**Figure 7.2:** The MuLan kicker timeline vocabulary. The requirement of fast rise and fall time is to reduce sneaky muons during transition period. The sneaky muons during the observation period will contribute to the flat background.
The MuLan kicker consists of 2 pairs of deflector plates mechanically in series, driven by 4 MOSFET modulators operating in push-pull mode. Figure 7.3 shows a photo of the kicker set up for MuLan. The specifications for the kicker demand that the potential difference of two plates be variable up to 25 KV and the rise and fall times during transition be less than 45 ns. Each pair of plates is 75 cm long, 20 cm wide, 5 mm thick with a 2 mm radius on the edges, separated by 13 cm and housed in a vacuum beam pipe with an inside diameter of 60 cm. During the accumulation period, the kicker is uncharged, allowing muons to be delivered continuously to the experimental area. The measurement period is initiated by the charging of the kicker, which deflects muons into the downstream quadrupoles and slits. During the 2006 run period, two of the modulator stacks malfunctioned. This was overcome by short circuiting each of the two pairs of plates and driving them with only two modulators at the same potential difference of 25 kV. This increased the transition time (on and off) of the kicker pulse from 45 ns to 67 ns. With a fully operational kicker the extinction factor, $\varepsilon$, defined as the ratio of the beam rate between the accumulation and measurement periods, is measured as $\varepsilon = 890$. The $\varepsilon$ is also defined as the ratio of the number of positrons from muons stopped during the accumulation period to the pulses from the background. The higher value of $\varepsilon$ indicates a smaller background and shorter time to collect useful decay positrons. The systematic study of the kicker voltage vs. $\varepsilon$ shows that changes of $\varepsilon$ are proportional to changes of kicker voltage. The $\varepsilon$ drops by 50 when the kicker drops by 1 KV. So keeping the high value of the
voltage is crucial in the kicker setup.

Another consideration, which is more important, is about the stability of the high voltage. The kicker can't shut the beam off completely and the sneaky muons during the measurement period will contribute to the flat background. Any time varying perturbation on the kicker high voltage will result in an early to late non-flat background and the muon lifetime will be disturbed. We used a high voltage oscilloscope probe to monitor the stability of the kicker high voltage during the experiment. We also spent a few days studying the extinction factor vs. high voltage change. Results will be presented in the section on systematic errors.

7.4 AK-3 target

The primary target for the 2006 run was a thin disk of AK-3 (a ferromagnetic sheet of chromium-cobalt-iron alloy, also called Arnochrome III), which was used for data production. AK-3 was chosen in order to destroy the polarization of the muon ensemble in the target material. The AK-3 target built for MuLan is a 0.5 mm thick sheet of metal with a radius of 24.5 cm. The Arnold Engineering Company quotes the composition of AK-3 as 26 – 30% chromium, 7 – 10% cobalt, with the remainder as iron. The SRIM simulation finds a mean stopping distance of 167 μm with 23 μm straggling for the incoming muons. The AK-3 has a large internal field of about 400 mT. The muon spin precesses about the field axis with a frequency of roughly 100 MHz, which means the muon will rotate thousands of times before decaying.
Figure 7.3: The MuLan kicker photograph. There are 4 cabinets consisting MOS-FET. The 4 kicker plates are installed inside the cylinder vacuum pipe.

The AK-3 is well known from μSR studies as a polarization-destroying material. The internal magnetic field serves both to reduce the polarization at the beginning of the measurement period by dephasing, and to precess the polarization with a period much smaller than the temporal resolution of the electronics. The original polarization of the incoming muons, which is initially aligned along the beam axis, is dephased by the combination of the different muon arrival times and the fast muon spin precession. At the end of the collecting period, all we have is an unpolarized
ensemble of stopped muons. Figure 7.4 shows the dephasing property of the sulfur and AK-3 targets. The residual polarization is very small, even without dephasing. Recent μSR measurements, made by MuLan and others, found no evidence of residual polarization at time $t > 50$ ns after arrival of DC muon beam. Figure 7.5 shows the residual polarizations of muons on different target materials.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\linewidth]{figure7.4.png}
\caption{Dephasing and depolarization of muon spin by the target materials. Top: spin dephasing; bottom: spin deplorization during accumulation period.}
\end{figure}

The targets is held at the center of the detecting system. Examining the time spectrum from a detector segment on the ball, any residual spin polarization will result in a modulation of the exponential decay. The modulation in diametrically opposed detector pairs will be the opposite of each other. When the two time spectra are added together, the modulations cancel each other and the overall distortion is
Figure 7.5: The comparison of residual polarizations of 3 target materials. For the silver target we always run with a short accumulation period and the muon spin polarization is kept and asymmetry is biggest; For the sulfur and AK-3 targets we run with normal 5 $\mu$s accumulation period and due to both dephasing property and residual polarization the asymmetries are smaller. The AK-3 has little or near zero residual polarization.

minimized. If all detector segments are located symmetrically around the target, and have the same acceptance, there will be no angular bias.

Due to the importance of any residual polarization of the muon ensemble, and its possible influence on the lifetime measurement, the collaboration has performed a number of different systematic studies with different target configurations.

1. By rotating the target section of the beampipe assembly through 180° or 90°, production data were collected with AK-3 field orientations of left-to-right and
right-to-left, or top-to-bottom and bottom-to-top. The average muon spin
precession direction correspondingly changes from upwards to downward, left
and right. All results are consistent.

2. Systematic data were collected in the so-called AK-45 configuration (an ellipti-
tical AK-3 target that was oriented at a 45° angle relative to the beam axis).
The 45° pitch enabled target configurations with a longitudinal magnetiza-
tion component, i.e. a magnetization component directed either upstream or
downstream along the beam axis. Such configurations permit the study of any
effects of a longitudinal B-field component on the muon spin relaxation and
muon spin dephasing.

3. The collaboration also used the copper, silver and aluminum targets to cali-
brate the $\mu$SR amplitudes. As these metals are good conductors they should
not depolarize the muon ensemble except through dephasing. $\mu$SR was studied
with an applied B field and a short collection time to reduce dephasing. The
initial muon spin will slowly rotate in the earth’s magnetic field with a pre-
cession frequency of about 100 kHz. The large residual polarization in these
targets permits the study of the effects of the residual polarization and resulting
counting-asymmetry on the lifetime determination.

4. To study the effect of non-target stops, sets of measurements were made with
both AK-3 (polarization destroying) target and the Cu (polarization preserv-
ing) target moved upstream or downstream of the geometrical center of the MuLan ball.

7.5 Errant Muons and Muon Corridor

Any non-target muon stops may introduce systematic error into the muon lifetime measurement. These errant muons are not stopped in the center of the ball and their decay positrons will have a non-uniform distribution around the ball. If the muons are polarized, the cancellation in opposing pairs of the detectors is not perfect. Most of these errant muons stop just upstream of our target. In 2004 and 2005 configurations, muons exit the vacuum pipe through a thin mylar window, pass through the EMC wire chamber (located at the entrance of the detector ball), enter a helium-filled bag and then stike a large-diameter target. The helium bag is used to minimize the unwanted muon stops between the entrance of the ball detector and target. In 2006, a vacumm pipe was installed between the entrance and exit of the detector ball. The target was positioned in the ball center and could be rotated. The configuration required that the EMC wire chamber be placed downstream of the target (being moved to the exit of the detector ball). Muons can then only pass through to the EMC when the target is rotated out of their path. Figure 7.6 shows the beam corridor setup inside the MuLan ball detector and figure 7.7 shows two views of the vacuum pipe and rotating target.

The vacuum pipe is made of thin-walled aluminum. The wall of the pipe is lined
with the high-field AK-3 alloy. There are also two small scintillators installed near the wall of the beam pipe, to monitor the muons that are stopped in the pipe and decay positrons in the backward direction. The beam pipe can be rotated around the center of the beam axis, to study the azimuthal variation of the errant muon stops.

Figure 7.6: The setup of the detecting system with the muon corridor. The vacuum pipe was installed to reduce errant muons, the EMC was installed at the end of the pipe to monitor the beam profile and the plunging counter was installed near the entrance of the MuLan ball to study non-target stops on the pipe.
Figure 7-7: Two views of the 2006 “all-vacuum” muon corridor showing the rotatable target. The upstream portion of the beampipe is lined with AK-3 material

7.6 Sneaky Muons and EMC Wire Chamber

During the running time of the experiment, it’s important to understand the beam conditions. The ideal condition is to get a centered, narrow beam inside the detector system when the kicker is off, and to shield the whole system from the incoming muon beam when the kicker is on. If there are any sneaking muons coming in during the measurement period, it’s good to measure how many there are. So after each run shift (about 8 hours of data taking) a special run was taken without the target to stop muons.

A special detector was utilized at the end of the beam pipe to monitor the beam status. The Entrance Muon Counter (EMC), is a multi-wire proportional chamber (MWPC) recording the beam spot and the number of muons going through. The EMC is designed to measure beam position in 2 dimensions with minimal effect on the muons going through. The anode wires are 15 μm thick tungsten. The 25 μm
thick, aluminized mylar, exterior windows function both as the amplification gas container and cathode planes. A 12.5 um thick cathode plane separates the two wire planes. The wire planes are comprised of 128 wires with 2 mm pitch, read out in pairs, to give a spatial resolution of 4 mm. The intrinsic dead time per wire was measured to be 100 ns. The EMC is efficient at detecting 98% of muons and about 8% of positrons, with no ability to distinguish between the two. An image of the EMC is shown in figure 7.8 and a typical beam profile recorded by EMC is shown in figure 7.9.

A permanent magnet array, placed at the exit of the EMC, produces a central magnetic field of 11 mT. The purpose is to precess and thus dephase the polarization of muons that stop in the EMC materials, which is a relevant consideration only in 2004 and 2005. Without this magnet array, any muons stopping in the EMC would have been exposed to the earth’s magnetic field, resulting in a potential distortion of the lifetime curves. As noted earlier, in 2006 and 2007 data run, the EMC was installed downstream of the target and was used for target-open beam diagnostics once per shift.

### 7.7 Detector Ball

To detect all the positrons as efficiently as possible, we built a ball-shaped detector around the target, covering the 70% of the solid 4π angle. Pulse pileup (two pulses are too close to each other and treated as one pulse due to dead time resolution) is
Figure 7.8: A photograph of the EMC from 2006. The wire planes are housed in the octagonal structure at right.

minimized by the segmentation of the detector, which results in a relatively low peak rate per element. To cancel the effect of any residual polarization, the individual detectors are symmetrically distributed around the target. The muon detector ball is a truncated icosahedron, made up of 170 segments, with 20 hexagons and 10 pentagons (There would be 12 pentagons, but 2 pentagons are used as the entrance and exit of the beam pipe). Each detector element consists of a two-layer scintillator, inner and outer tile pair, from which coincident signals can constructed. The scintillator of each detector is triangularly shaped with the inner having a side length of 15 cm; the outer tile has sides which are 3 mm shorter. Figure 7.10 shows how individual
elements are assembled to the detector ball: from a single scintillator to a pair then into hex or pent groups and finally into a house, with support structures. Figure 7.11 is a picture of the MuLan ball in the experiment area. Overall, the MuLan ball covers roughly 75% of the 4π solid angle.

7.8 Waveform Digitizer

In 2006 and 2007 we used the waveform digitizers (WFDs) to read out PMT signals. The WFD records determine both the timing and voltage of each pulse. Having a digitized waveform for each event allows us to minimize or directly monitor a number
Figure 7.10: The MuLan ball detector elements. Individual scintillator element, light guide and PMT (top left); tile pair (bottom left); pentagon cluster of tiles (bottom middle); hex-house complete structure (top middle); MuLan ball (right).

of important systematic effects, including dead time, pile up, and gain and timing shifts. Each WFD channel consists of analog input buffers, flash analog to digital converter (FADC), FPGA, and FIFO memory. Figure 7.12 shows the WFD block diagram and a photo. The full complement of WFDs are split evenly among six VME crates, each of which is controlled and read out via a frontend control computer. WFD was also used for the HV calibration before the production run began. Figure 7.13 shows a couple of digitized WFD pulses. Normally each pulse is digitized as 24 8-bit ADC samples, which is just called a WFD block. Adjacent blocks form a WFD island. The WFD has an internal parameter to set the threshold to trigger the input pulse. After being digitized by the WFD, each pulse is represented by
a block of 8 32-bit d-words, with 6 d-words for 24 samples, 1 d-word for time and fill information and 1 zero d-word which is used in verifying data quality. All data blocks in one event (consisting of 5000 fills) are stored in the FIFO and read out by the DAQ during each event.

7.9 Data Acquisition System (DAQ)

The DAQ is based on the MIDAS system which is the standard in PSI and TRIUMF. It has a distributed, modular design with multiple frontends (FEs) and a single backend (BE) for event building and data storage. During 2006 and 2007, our DAQ efforts were focused on the followings:
1. DAQ layout: All components of DAQ were thoroughly investigated for possible bottlenecks. The main problem concerned data throughout on the BE processor, like logging data to physical storage over the network. Distributing various tasks among DAQ processors and subnetworks produced additional improvement.

2. DAQ stream compression: To store all the information for more than $10^{12}$ decay positrons, we need up to about 60 TB storage, or 120 tapes with size of 500GB. By compressing the digitized pulses on line, we can save some storage. After careful study of various compression algorithms, we decided in favor of the ZLIB library, which provides both high speed and good compression ratios.
Figure 7.13: Example pulse waveforms recorded by the WFD. The top is from the outer tile, the bottom is from the inner tile. Coincident pairs are formed from the inner and outer tiles.

3. Multi-threaded front end: Imposing the additional task of compressing the data stream on each FE processor is a significant bottleneck at very high data rates. Thus we developed a multi-threaded version of the programs which executes two tasks in parallel on dual processor machines.

4. Synchronization: The magic box, a custom electronics module which generates the fill and segment structure of the DAQ and controls the kicker, LED and laser systems, play an important role during the data acquisition. A new firmware was developed for more flexible, more universal control.

Figure 7.14 shows the schematic of the DAQ.
**Figure 7.14:** Schematic of data acquisition system (DAQ) showing backend (BE), frontends (FEs), analyzers, slow control and setup for data monitoring and control.

### 7.10 Magic Box

The magic box is crucial to control the whole detecting system. It sends a couple of signals to the kicker, the DAQ and the WFDs. The timing structures of these signals need to be organized and all devices need to cooperate to fulfill the design of the overall systems.

Figure 7.15 shows the control signals sent to devices from the magic box. The scope is triggered on the start of the data segment, which will last as long as 5000 fills. In each fill the kicker and the WFD works in phase. The start of the data segment triggers the start of the 5 μs collection period, when the kicker is off and the WFD
start gate is reset. After 5 $\mu$s the kicker is on, the system goes to 22 $\mu$s measurement period, with the WFD start gate is set. Note that the wfd starts and kicker signal are synchronised and data segments are forced to start on fill boundaries.

**Figure 7.15:** signals sent to the kicker, the WFD and the DAQ from the magic box. yellow on the top is for data segment, blue in the middle is for the kicker and pink on the bottom is the WFD start
Chapter 8

Detail of the 2006 data sets

For each of the 2006 and 2007 data run, the data acquisition system (DAQ) collected information on more than $10^{12}$ decay positrons. For each decay positron, the raw data consists of ADC samples of the analog pulse, recorded by the WFD. The raw WFD data samples then need to be fitted to get pulse time and amplitude information; here tiny pulses will be removed. The fitted pulses then need to satisfy two main requirements before forming the lifetime histograms. The pulses need to have amplitude above a threshold and they must form inner-outer coincidences to be the real signal. The pile up correction was applied during histogramming, by means of artificially built pile up pulses from adjacent fills in certain deadtime. The main lifetime histogram is fit to extract muon lifetime by a simple 3-parameter fitting function.
8.1 Muon spectrum

During the collection period, the continuous muon beam hit the target, and some decay before the end of the collection window; the number of muons in the target as a function of time is given by the solution to

\[
\frac{dN(t)}{dt} = R_{on} - \frac{1}{\tau}N(t), \tag{8.1}
\]

which gives, at the end of the collection period

\[
N = R_{on}\tau(1 - e^{-t_c/\tau}). \tag{8.2}
\]

where \(R_{on}\) is muon beam rate and \(t_c\) is the duration of the collection period.

During the 2006 run, the beam rate was about 10 MHz and \(t_c = 5\mu s\), so roughly 20 muons were in the target at the end of the collection period. The muon lifetime can be measured from the shape of the time spectrum of the decay positrons with reference to the time at the end of the kicker cycle. The time distribution of decay positrons follows an exponential law:

\[
N(t) = N_0e^{-t/\tau} + B \tag{8.3}
\]

where \(B\) is the flat background, mostly coming from sneaky muons during the mea-
suring period. This formula is good enough for most of the systematic studies. Figure 8.1 shows the time spectrum of muon collection and positron decays.

![Graph showing time spectrum of muon collection and positron decays](image)

**Figure 8.1:** Time spectrum of muon collection during accumulation period and muon decays during measurement period.

A more precise achievement of \( \tau \) should consider the pile up term, although at very small ratio but still the big effect for the 1 ppm precision measurement. This might lead to a more accurate 4-parameter fitting function:

\[
N(t) = N_0(e^{-t/\tau} + \eta e^{-2t/\tau}) + B
\]  

(8.4)

where \( \eta \) is the pile up ratio, about \( 10^{-4} \). The most precise way to get \( \tau \) is to build the pile up term artificially and subtract it from the original lifetime spectrum, what
will be left is a pure exponential term which can be simply fit by equation 8.3.

8.2 The 2006 Data Set

The data volume for the MuLan experiment depends on the precision goal of the experiment. To get a 1 ppm measurement a minimum of $10^{12}$ muon decays is required. Considering the amplitude cut below certain threshold and background, 25\% more data is in need. Each muon decay will appear as a coincidence between inner and outer detectors. For every trigger the WFD records 24 samples from the analog pulse with each sample represented by one byte. The timing information requires additional 4 bytes. The ADC samples and fill/time words are separated by 4 bytes zero word. So the total data volume for the whole run needs to be at least

$$10^{12} \times (1 + 25\%) \times 2 \times (24 + 8) = 80 \text{ terabytes}$$

This data volume was two orders of magnitude larger than the maximum storage capacity of a typical hard drive at the time. Considering the budget and efficiency, the collaboration decided to use hundreds of tapes with 400 GB size per each for the storage during the data collection when running the experiment.

The design of our data acquisition system considers two main factors: first it should follow the default data format, MIDAS (Maximum Integration Data Acquisition System), used at PSI, where we did the experiment and from which we asked for temporary space for data storage; secondly it should collect sufficient data in a reasonable time. The 2006 MuLan dataset was divided into 1.9-gigabyte compressed
data files called runs. The data was compressed to save the storage size. Each run contains about $3 \times 10^7$ muon decays, written to the disk in about 2 minutes at the typical data rate of about 30 MB/s. In all more than 30,000 runs passed the analysis cuts and were recorded to the disks and tapes in less than 2 months.

Midas is based on a modular networking capability and a central database system. The software contains a library which can be used for data transport between different computers and programs, and a set of programs for data logging and system management. Experiment configuration is stored in a fast online database and Web interface makes experiments remotely controllable. A slow control system is integrated which can be used to control high voltages and measure things like temperatures.

A typical MIDAS file consists of a number of MIDAS events which have the same structures and are stored continuously in the run file. The number of events is variable and depends on file size. Each MIDAS event contains a set of data banks and each bank corresponds to a specific device, like ADC data, scaler, and slow control data. The majority of the run file consists of decay positron data, recorded to the WFD and transferred to MIDAS; other information such as experiment control variables and slow-control data, was relatively small. In MuLan each MIDAS event contains 340 similar data banks, corresponding to 340 WFD channels. Each channel contains information about one detector tile. The data structure of the data bank for each WFD channel contains 3 parts: a header, a bunch of time-ordered WFD pulse
samples in 5000 contiguous fills and a footer, as shown in Figure 8.2. The header and footer contain checksum information about data recording of the corresponding WFD channel in the event. Typical checksum in the header includes the number of bytes requested by the DAQ and the actual number of bytes received by the DAQ. The footer tells whether the fill numbers are messed up. Each WFD pulse was sampled and stored in fixed-length data block, including: 24 8-byte ADC samples, 1 32-byte zero word and 1 32-byte fill/time word.

**Bank level checks**

![Diagram of bank level checks]

**Figure 8.2:** Type data structure of bank status which is for one channel and one event

The raw data can be analyzed both online and offline. The online analysis, reading data from the memory, is good for monitoring the experimental behavior
timely by looking at individual channels. The offline analysis, achieving data from
disk, is the final step to get the muon lifetime spectrum. To run the analyzer, we can
still follow the same MIDAS format, and analyze the data stage by stage, but this
turns out to be much bigger file size at each stage and more complicated algorithm
when considering pile up construction. The other way, which was proved to be a
better choice, is to transfer the MIDAS data format to the ROOT format and to
use the ROOT tree file. ROOT is the most popular data format system in the high
energy physics area. It has a lot of functions to organize and analyze data and is
being developed by many people in the world.

8.3 Run Selection

All channels must pass certain data validity check and if any sector fails the data
checks, the entire MIDAS event (includes 5000 fills) for that single channel will be
discarded.

The so called “golden runs”, which will be used in the final data analysis, was
selected based on two main checks: the event check and the run check.

Typical data validity check in each event based on WFD channels includes:

1. the checksum in the header: the dwords (32 bit long words) requested by the
   DAQ must equal to the dwords it receives.

2. the checksum in the tail: the fill number in the tail word must be equal to
5000, which is the maximum fill number in each event.

3. the fill number must monotonically non-decrease.

4. the time difference of adjacent WFD blocks must be equal to or bigger than 24 (the time elapsed for each WFD block), if they occur in the same fill.

5. none of the time or fill numbers maybe the hexadecimal value 0xffff, which indicates a hardware error.

Typical check in each run is listed as:

1. Quality. During the experiment, each run was marked with a quality and a description by the collaboration member(s) on shift. Such as “N” indicating a normal or good run, “C” indicating a cosmic run and “S” indicating a systematic run. A run must be marked as valid and normal before being considered as a potential golden run.

2. Firmware. We were still debugging WFD hardware issues early in the experimental run, and three different working firmwares were developed and installed to the WFD during the later time of the experiment of 2006 run. Before that, the relative time difference between inner and outer detectors for the same decay could shift by ±4 clock ticks. This problem exists in all the runs before run number 23387 and these runs will be discarded.

3. File Size. Typical file size for each run is about 1.9 GB and small file sizes
indicate a DAQ problem or man-made run stop. Golden runs must contain at least $5 \times 10^3$ muon decays or have a file size bigger than 100 MB.

4. Extinction factor. Typical extinction factor (EF, ratio of signal to the flat background) is about 890 and too small extinction factor indicates beamline or kicker problem; we required the EF be greater than 400.

5. Goodness of Fit. It happened occasionally that the DAQ recorded two copies of the data in the same run file. These runs are identified by poorer goodness of fit than normal runs and are excluded for post-production analysis. The golden runs require $\chi^2/NDF < 1.6$, while the bad runs have $\chi^2/NDF \approx 2$.

In addition to the production “golden” runs, which contain enough data for our production goals, we also performed some runs for systematic studies at different times between normal runs.

1. muon spin studies with different stopping targets. We used targets made of copper, aluminum which maintain spin polarization. We also used a tilted AK3-45 ellipse.

2. beam monitoring. Once or twice per day, after each shift, the target flap was opened and one run was taken with the EMC recording the muon beam spot.

3. long island runs. Once per day, the number of samples in WFD island was increased from 24 samples to 64 samples, and 10 runs were taken. Long island runs are useful to study the pile up and stability of the detector system.
4. asymmetries of detector acceptance. This study was performed in one day and
the procedure is simple: the whole ball detector was displaced downstream
from 0 cm to 80 cm from the original center. The target was fixed.

5. laser Study. Once or twice per day, a nitrogen laser was operated in self-trigger
mode with a repetition rate of about 37 Hz. The laser light was split and
distributed by fiber optic cables to 24 detector tiles around the ball. These
laser runs have both normal pulse and laser pulse and are included in “golden”
runs if passed the validity check. They are also used separately for timing and
gain stability of the detecting system.

8.4 Data storage

The dataset in the disks was first imported to the mass storage system (MSS) at
the National Center for Supercomputing Applications (NCSA), from which we did
the data check and data analysis work. Our data check and analysis work was based
on run by run analysis. for each run the analysis includes two parts from the raw
MIDAS data to ROOT files, where the works includes read the raw data, check the
data validity, fit the raw WFD samples to get the pulse time and height and store
them into ROOT tree files. from the ROOT tree file to ROOT histogram. where
the work includes read the fitted pulse time and amplitude for each pulse, find the
coincidence and do the time alignment for inner outer pairs, build the pile up pulses
and construct the histograms.
The analysis uses a lot of CPU times and storage space and we used the ABE clusters at NCSA which provides the parallel computing on the dataset and fast network infrastructure on the data transfer.

8.5 inner outer time alignment

To minimize the unwanted cometic pulses and beam related background like errant muons, each segment of the detector system was made up of two scintillators that are roughly normal to the direction of decay positrons. Only the coincidence pulses, which are almost from the target muon decay, will come to the final lifetime histogram. The coincidence pulses come from inner-outer tile pairs with adjacent triangle scintillators and light guides and supposed to have the same arrival time. But due to the cabling and electronics delays, signals generated by the same decay positrons do not arrive to the WFD at the same time. The tile-pair timing misalignment is on the order of a few nanoseconds. we remove these misalignments by setting one run as reference run and obtaining the mean of time difference of each tile pair as $dt$ offset. Figure ?? shows the $dt$ of one pair after offset. The mean of $dt$ is around 0 as expected. The width of $dt$ is a few clocks, there are two shoulders around the +/- 1 clock tick and this is due to the power cycle which cause the relative time between inner and outer detector to change by exactly 1 clock tick.
8.6 Amplitude Cut

To minimize the gain stability systematic error, the pulses also need pass a certain threshold. Too low amplitude pulses are the harm to the lifetime measurement because they have more noises inside. The threshold for the amplitude cut is chosen at the minimum between signal and noise region. At minimum it has flat distribution and effect of gain shift is also minimized. In 2006 data analysis we set the amplitude threshold at 60 ADC.

8.7 Artificial dead time and pile up

The fitter, which was used to fit the WFD samples to achieve time and amplitude of the pulse, has the limitation to separate two adjacent pulses. In mulan the limit is at around 3 clock ticks, which is called “artificial deadtime” or ADT. Two pulses inside the ADT will be considered as single large pulse. With the ADT cut, real pulses too close to the trigger pulses will be lost and this is called pileup.
Chapter 9

Pulse Finding and Fitting Algorithm

9.1

The Waveform digitizer (WFD) gives us access to a vast amount of information that is not available from a TDC (time to digital convertor). The WFD records the digitized pulse shape from the original analog signal. The WFD samples the input signals at constant frequency and records blocks of 8-bit samples, which we refer to as a WFD island. Typical WFD islands are shown in figure 9-1. In these cases each WFD island corresponds to one or more pulses that exceed an analog threshold. The WFD island length, 24 clock ticks in MuLan, is set by the firmware. While the time and sample values are immediately accessible, additional information like pedestal, area, rise and fall times can also be calculated.

The timing resolution on individual events is partly determined by the clock frequency at which the WFD operates. When the WFDs were introduced to MuLan experiment for the first time, back in 2005, the clock frequency was limited to 450 MHz, above which some channels did not work. Later we found that it was the
Figure 9.1: WFD islands for a pair of inner outer channels. Pulses from the inner and outer tiles will form the coincidence pulses and be taken as the real positron decay pulses.

firmware compiler issue. In 2006 the firmwares were developed 3 times during the experiment, to improve the WFD performances. The clock frequency was still set at 450 MHz to avoid bit droppings in sample data transfer.

There are many advantages of utilizing WFD. The fit of WFD pulses to predefined pulse shape helps to determine the pulses’ parameters more accurately, especially the pulse peak time, on which we achieved a resolution of 100 picoseconds. WFD data can also help to reduce systematic uncertainties. The accurate fitting of WFD pulses is good for the separation of “pileup”. Along with the laser system, the uncertainty on the timing can be obtained by studying changes in time pick of \( T \) from early to late in the measurement period. Timing shift might be caused by electronic noise or “pileup”. Pileup consists of two pulses that are close to each other (smaller than the
resolution time). We can directly monitor PMT gains throughout the run.

By themselves the WFD samples only give us coarse timing and amplitude information, achieved from the highest sample in the island, with the precision of $1/\sqrt{2}$ clock tick (about 1.6 ns) and 1 ADC count. To get much improved precision, pulse fitting is a must and constructing pulse template is the first step. The pulse template is just a predefined pulse shape, extracted from the data, that is used for the fit. Starting from the empirical observation that the PMT pulse shapes are nearly independent of pulse amplitude, we build an average pulse template for each WFD channel from a large number of good WFD islands. We can then use that template to extract individual decay event timing and pulse shape information, including pedestal, height, area, rise and fall times, etc.

To find the pulse candidate and build the average pulse shape, we followed the method used in the Brookhaven E821 g-2 experiment. The data stream is scanned and time-adjacent blocks are merged to form pulse islands. Each island is then searched for local maxima that exceed a software trigger threshold, and an initial guess is made as to the pedestal. The local maxima are labeled as pulse candidates, or prepulses, if the rise, fall, and separation from neighboring maxima are large enough. Additional heuristics are applied to handle off-scale pulses and to address islands with multiple prepulses.

We then build up the pulse templates for each channel. To do so, we run over the data, channel by channel, looking for islands with only one prepulse. For these
prepulses, we build the pseudotime histogram. The pseudotime \( \psi \) is a proxy measure of the peak position, within one clock tick, for each pulse.

\[
\psi = \frac{2}{\pi} \tan^{-1} \left( \frac{a_{\text{max}} - a_{\text{max}-1}}{a_{\text{max}} - a_{\text{max}+1}} \right),
\]

(9.1)

where \( a_{\text{max}} \), \( a_{\text{max}-1} \) and \( a_{\text{max}+1} \) are the maximum sample and the samples on its left and its right, respectively.

A typical pseudotime distribution is shown in figure 9-2. We invert this histogram to build a mapping function that takes the non-flat pseudotime into the flat sub-tick offset time: \( \tau \rightarrow t \). Using this mapping \( \tau \), we can shift each prepulse by the calculated offset so that the true peaks of all pulses line up and add the measured samples to finely binned (0.01 clock ticks or 20 ps) histograms which extend from about 20 ns before the pulse peak, to 40 ns beyond. After processing several million pulses, the true average pulse shape is obtained. If the template is insufficiently smooth, we repeat the procedure, by using the preliminary template of the previous iteration to refine our estimate of the pedestal and amplitude. Figure 9-3 shows the constructed pulse template.

Once the average pulse shape has been constructed, we can extract fit times (the pulse peak time achieved from pulse fitting) and amplitudes for all events in the data set. By comparing data to the template we can seek the time and amplitude from
Figure 9.2: Typical pseudotime distribution. Horizontal axis is pseudotime defined by \( \psi = \frac{2}{\pi} \tan^{-1}\left( \frac{t_{\text{max}} - t_{\text{max}-1}}{t_{\text{max}} - t_{\text{max}+1}} \right) \) which is approximate fine time shift in 1 clock tick coarse time resolution.

the goodness of the fit, also called \( \chi^2 \), which is computed as

\[
\chi^2 = \sum_{\text{samples}} (y_i - y_i^0)^2
\]

where \( y_i \) and \( y_i^0 \) are the ADC counts of island and template at each sample \( i \) and \( y_i^0 \) of templates are scaled by pulse height and shifted by the peak time. Within an island, the \( \chi^2 \) is good estimator of how well the pulse is fitted. The final pulse height and peak time is achieved by minimizing the \( \chi^2 \).

Different fitting methods are used depending on the waveform. Accuracy and efficiency must be balanced since more than \( 10^{12} \) pulses being processed. Brent's Parabolic Minimization is a fast routine for finding minima of functions in as few
function calls as possible. It was always attempted when one pulse is found in the island. The full Minuit fit is the most trusted fitting method but it also takes much longer time than Brent's method. So it was only used when two or more pulses are found in the island or $\chi^2$ from Brent's method is not so good. Figure 9-4 shows a template fitted pulse superimposed on the samples from a WFD island. Figure 9-5 is the fitted peak time distribution on the WFD island. The typical WFD island length is 24 clock ticks, but if there is a second pulse arrives near the island end of the first or trigger pulse, the WFD island length can be extended to 48 clock ticks. The pulse peak time has flat distribution from 8 to 12 clock ticks, reflecting the fact that we set 8 presamples before peak in the WFD firmware and the actual signal arrival time can be at any time inside 4 clock ticks which are typical WFD internal clock period. The flat top from 18 clock tick to 36 clock ticks are from events in which there is a second pulse in the same island as the trigger pulse. Figure 9-6 is the fitted amplitude distribution for all the single tiles. There is a triggering (hardware) threshold to distinguish the signals from noise (30 ADC counts in our case). In the peak range, the pulses' amplitude distribution is well described by a Landau convoluted with Gaussian distribution and there is the so-called Valley of Death (VoD) area which roughly separates true pulses from noise. There are comparatively few pulses in the valley of death, but that number rises sharply as the amplitude falls to zero. We applied a hard cutoff to minimize the effects of possible gain shifts, and this software threshold is usually chosen around the VoD, where the change of counts due to any
possible threshold shift is minimized, about 70 ADC counts in our case.

The fitter has limited ability to fit very small pulses or to separate two very close pulses, three cuts are applied in the production code. Far away from any other pulse activity, pulses with amplitudes as small as 11 ADC counts can be reconstructed with full efficiency. However near a larger trigger pulse, David’s PFA (by default) looks for no pulse with amplitude smaller than 35 ADC, nor does it keep extra pulses any closer than 3 clock ticks. Figure 9.7 shows these PFA cuts in the simulation. The main pulse has amplitude of 130 ADC counts and peak time of 9.8 clock ticks. The add-on pulses have amplitude range from 25 ADCs to 50 ADCs. The horizontal axis represents the time difference between the add-on pulse and the main pulse, the vertical axis is the probability to find 2 pulses by the PFA. There is the +/- 3 ct deadtime, inside which the add-on pulse will be discarded no matter how big it is. 35 ADC counts is the transition amplitude for the add-on pulses, below which the add-on pulses will be discarded by the fitter, and above which they are recognized by the fitter.

9.2

With the information obtained from fitted WFD pulses, systematic studies like gain shifts, pile up subtraction and timing shifts can be much improved. One concern is the pulse shape early-to-late change in the fill which could be caused by hardware problems or the effect of pileon or pile-up in our data, which is of concern because
it occurs more often at early times than at late times. Pileon refers to the direct superposition of two pulses in the dead time window of the PFA. Pile up will cause shifts of a pulse time and height due to a second pulse in the same island or in the same fill. The effect of the invisible pulses, real enough but too small in amplitude to be found by the PFA, is hard to estimate. We study the pile on effect by use of a simulation where we construct pulses from templates with different amplitudes and see how the PFA responds as we change their time separation.

The pileon includes three different effects relevant to MuLan:

1. Two pulses, both with amplitudes above the software threshold, produce a combined pulse which is also above the software threshold.

2. Two pulses, only one of which with amplitude above the software threshold, produce a combined pulse which is also above the software threshold.

3. Two pulses, neither of which with amplitude above the software threshold, produce a combined pulse which is above the software threshold.

Item 1 is straightforward - one pulse is lost and there is a timing shift on the pulse we do see. On average, that timing shift should be nearly 0. Item 2 consists of two subcases, which are shown in figure 9-8. If the pileon pulse falls in the signal region of the trigger pulse, it will contribute at some level to the amplitude of the combined pulse. However, if it falls in the pedestal region of the trigger pulse, and is not seen by the PFA it will contribute to the pedestal and therefore reduce the
computed amplitude. The contribution goes to 0 in the two transition regions, just before and just after the trigger pulse, around the time resolution of the fitter. If the errors on the samples are uniform, these two effects will exactly cancel. Item 3 is a case where we know relatively little. The PFA is very efficient for pulses with amplitude above 40 ADC counts but that efficiency drops quickly with amplitude.

We can simulate what happens when two pulses fall close together. We study the average effect of pile-on pulses on the main pulse, by moving the add-on pulses uniformly across the whole island. Consider figure 9-9, where the main pulse has an amplitude of 130 ACD counts and the pileup/pileon pulse has an amplitude of 50 ADC counts. The horizontal axis is the time difference between the add-on pulses and main pulse. The number of add-on pulses are the same in each point. This plot corresponds to item 1 mentioned above. In the signal region, the combined pulse has a pull-up amplitude and the effect of pile on is roughly symmetric around $dt=0$ (top right in figure 9-9), where the pileon and trigger pulses exactly overlap and the amplitudes simply add. The time shift of the main pulse at $dt=0$ is 0, as expected. Away from $dt=0$, the combined samples from two pulses will form a pulse shape that disagrees with the template and the shift on the fit time is manifest. The effect on the peak time is roughly anti-symmetric around $dt=0$ (top left in figure 9-9). The average timing shift in the signal region due to the add on pulse is 0.011 clock ticks. If the add-on pulse resides outside of the time resolution range (3 clock ticks in the MuLan PFA), the fitter will treat the two pulses separately, but with the same
pedestal. So the effect of an add-on pulse in this range is tiny. The average shift on
peak time in this range is as low as 0.006 clock ticks. The overall average shift on
peak time in the whole island range is only 0.009 clock ticks.

Consider figure 9-10, where the main pulse has an amplitude of 130 and the
pileup/pileon pulse has an amplitude of 10. This plot corresponds to item 2 men-
tioned above. In the signal region, the effect is roughly the same as that in case item
1. The average timing shift in the signal region is 0.005 clock ticks. In the pedestal
region, the effect of the tiny pulse is bigger than that of a small but visible pulse.
The tiny pulse has an amplitude that is below the software threshold and rejected by
the fitter. It will pull up the pedestal of the whole island. The change in amplitude
is negative and fairly constant. The transition point is around \( dt = 2.8 \) clock ticks. For
uniform errors on each sample, the effects from signal region and that from pedestal
region should exactly cancel, which is reflected in the average change in amplitude
of 0.117 ADCs that is plotted on the figure.

In practice, the cancellation is imperfect but probably adequate. The top of
figure 9-11 shows the average change in amplitude, for pileup pulses which fall over
a single island, for a variety of trigger pulse amplitudes. Even for pileon pulses as
large as \( A_p = 30 \), the net effect on a trigger pulse amplitude \( A_t = 65 \) (typical for the
Valley of Death position) is only 0.5 FADC counts, or a fractional effective change of
0.7%. With the probability of overlap (the small, unseen pulse falls inside the fitting
window) less than 1% at early times, the net fractional change in amplitude from
early to late is about $7\times10^{-5}$. As the amplitude of the pileon pulse grows, the effect on
the trigger pulse grows proportionally. The effect grows more or less linearly with
the pileon pulse height but falls as the trigger pulse grows in amplitude.

The $\delta t$ pull of pileon pulses is very small. There will be equal numbers directly
before and just following the trigger pulse and their pulls on the time of the trigger
pulse lie in opposite directions, as seen in figure 9-12. The negative and positive lobes
of the graph describe pulls to more negative and more positive time, respectively.
When the pileon and trigger pulses exactly overlap, the pull is 0. The average pull,
while normally 0.009 CT, is actually much closer to 0 CT. There is a 0.007 CT offset
in synthesizing the trigger pulse, which we checked by adding in a pileon pulse of 0
amplitude, that is, no pulse at all. So the true average pull is 0.002 CT or less.

Provided that only one pulse is detected (which is true for pileon amplitude
smaller than 35 ADC counts), the average pull remains close to 0 even as the pileon
amplitude increases. There is some scatter in the average pull but no overall trend.
With the probability of overlap (the small, unseen pulse falls inside the fitting win-
dow) at early times beling less than 1\%, the effective early to late timing shift is less
than 0.00002 CT which we can safely ignore. Only for large trigger pulses (overflow
pulses with $A>230$), is there a slight net pull at earlier times. These are shown in
bottom of figure 9-11. When the trigger pulse has amplitude of 210 ADC counts, the
average pull of pile-on pulse ($A=30$ ADCs) is -0.03 CT. Even so, the early to late
timing shift is about 0.0003 CT which is still a tiny effect.
\[0.0003c.t./9000c.t. \approx 00.3\text{ppm}\]

The situation becomes a bit murkier when the pileon pulse is right at the threshold of detectability. In this case, the asymmetry of the pileup pulses' pull is no longer perfect, as shown in figure 9.13, with a trigger pulse of amplitude 50 affected by pileup pulse of amplitude 35. There is a net pull to negative times of roughly 0.1 CT.

9.3

The pulse finding and fitting algorithm is so important to the success of the analysis that two were developed in parallel, to serve as cross checks on each other. By comparing the two methods' results on the same data set, we know much better about accuracy of our fitting method. Both PFAs use the same template fitting method and figure 9.14 shows that the pulse templates of the two fitters for the same run are identical. Figure 9.15 shows the fit time spectrum from both fitters and they also agree with each other very well.

While the two methods were based on the g-2 experiment, there are differences in some of the details. On occasion, one of the two fitters will find a pulse that the other miss it. Missing pulses will affect the lifetime measurement if they have non-muon lifetime distribution. By comparing the fitted pulses from both fitters, we can learn about how many missing pulses in either fitter and how they will affect the muon lifetime.
By comparing two fitters and searching for the missed pulses by either fitter, it’s found that most of the missing pulses are caused by a incorrect fit on overflow pulses. These overflow pulses just do not agree well with the pulse template and the fitters may fail, or produce a much lower amplitude than expected, even lower than the software threshold. Figure 9.16 and 9.17 show examples of overflow pulses that were missed due to such incorrect fits, by the two fitters.

We are interested in the percentage of these missing pulses and the effect on our lifetime measurement. Figure 9.18 and 9.19 show these extra pulses found by either fitter but missed by the other. And table ?? is the summary of these two plots.

Figure 9.16 shows the coincidence pulses that are missed by Kevin’s fitter. The two bands at amplitude around 70 ADC counts represent the software threshold regions. A tiny change on fitted amplitude would cause the pulse go below or go above the threshold. The other problem comes from those pulses well beyond the threshold cut region, usually the normal pulses in David’s fitter. Figure 9.17 shows the coincidence pulses that are missed by David’s fitter. Again, the two bands at amplitude around 70 ADC counts represent the software threshold regions. The number of missing pulses in this region is at the same level as those extra pulses in David’s fitter. (4939 vs. 5515). Most of the discarded pulses by David’s fitter are overflow pulses.

extra pulses include 3 region classified by amplitude:

1. low : either pulse of inner/outer has amplitude < 90
Table 9.1: extra pulses found by one fitter but not the other

<table>
<thead>
<tr>
<th>Fitter</th>
<th>extra(ratio)</th>
<th>low(ratio)</th>
<th>overflow(ratio)</th>
<th>normal(ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>David</td>
<td>7466(2.9E-4)</td>
<td>5515(2.2E-4)</td>
<td>624(2.5E-5)</td>
<td>1317(5.2E-5)</td>
</tr>
<tr>
<td>Kevin</td>
<td>8806(3.5E-4)</td>
<td>4939(1.9E-4)</td>
<td>3620(1.4E-4)</td>
<td>247(1E-5)</td>
</tr>
</tbody>
</table>

2. overflow: either pulse of inner/outer has amplitude > 235 (exclude low amp)

3. normal: both inner/outer pulses are in normal amplitude region (90-234)

The table above is achieved from one run with total counts of 2.54E7. The pulses missed by Kevin’s fitter is 1941 or 7.64E-5. The pulses missed by David’s fitter is 3867 or 1.52E-4. These missed pulses are only a small fraction with ratio at the level of 1E-4 and as figure 9.20 and 9.21 show, these extra pulses found by both fitters follow the muon lifetime spectrums, so they have little effect on the measured lifetime.
Figure 9.3: pulse templates, using “smeared” and “unsmeread” methods. horizontal axis is clock ticks, 24 is the typical island length; vertical axis is normalized pulse amplitude
**Figure 9-4:** example of template pulse fitting. Horizontal axis is clock ticks, 24 is the typical island length; vertical axis is WFD ADC counts

**Figure 9-5:** fittime (fitted pulse peak time position) distribution.
Figure 9.6: Amplitude distribution for all detectors before coincidence. Trigger threshold is 50-20=30 ADC counts, Valley is around 90-20=70 ADC counts (pedestal is around 20)

Figure 9.7: Probability to find pile-on pulse vs. time difference between pile-on and trigger pulse. Different color points represent different pile-on amplitudes. The pile-on pulses that are found within 3 clock ticks of the main pulse will be discarded by the fitter. Pulses that are smaller than 35 ADCs will also be discarded by the fitter.
Figure 9.8: Pileon pulses which fall in the signal and pedestal regions have very different effects on the main pulse. In the signal region it will increase the amplitude of the combined pulse and also modify the pulse shape; in the pedestal region it will increase the pedestal and therefore reduce the computed amplitude of the main pulse.
Figure 9-9: Effect of small pileon vs. time separation of pulses: The shift of time (top left), amplitude (top right) and pedestal (bottom left) of the main pulse due to the pile-on small pulses; red solid square represents single pulse found because two pulses are inside the dead time ($\sim 3$ clock ticks), and blue hole square represents the noise region when two pulses are found.
Figure 9.10: Effect of pileon vs. time separation of very small pileon pulses: the shift of time (top left), amplitude (top right) and pedestal (bottom left) of the main pulse due to the pile on tiny pulses. There is always a single pulse found no matter where the tiny pulse is placed on the island.
Figure 9.11: Average change in amplitude (upper) and pickoff time (lower) of the main pulse from pileon vs. amplitude of pile on pulse. Different colors represent different amplitudes of the main pulses.

Figure 9.12: $\delta t$ from pileon vs. time separation of pulses: small pileon pulse on big trigger pulse
Figure 9.13: $\delta t$ from pileon vs. time separation of pulses: big pileon pulse on relatively small trigger pulse
Figure 9.14: Comparison of templates from David’s and Kevin’s PFAs. For this channel, SN 1, the two templates are nearly identical.
Figure 9.15: Fittime comparison of two fitters

Figure 9.16: WFD pulse (top) that was fitted correctly by David but was fitted with too low amplitude by Kevin therefore falls below his software threshold cut. Usually this involves an overflow pulse which does not agree well with the pulse template: fit time and amplitude from David’s fitter are: 10.4 and 253.7. Fit time and amplitude from Kevin’s fitter are: 12.4 and 56.6
Figure 9.17: WFD pulse (top) that was fitted correctly by Kevin but was fitted with too low amplitude by David therefor would falls below his software thershould cut. Again this involves an overflow pulse which does not agree well with the pulse template: fit time and amplitude from David's fitter are : 9.6 and 64.8 fit time and amplitude from Kevin's fitter are : 9.8 and 456.0
Figure 9.18: amplitude of outer tiles vs. amplitude of inner tiles for those pulses found by David's fitter but missed by Kevin's fitter when cuts were placed. Beside the valley region there are some normal and overflow pulses that were fitted wrong by Kevin's fitter.
Figure 9.19: amplitude vs. amplitude of those pulses found by Kevin’s fitter but missed by David’s fitter when cuts were placed. Beside the valley region there are some overflow pulses that were fitted wrong by David’s fitter.

Figure 9.20: lifetime spectrum of extra pulses found by Kevin’s fitter. The pulses follow the muon lifetime so they have little effect on gain shift.
Figure 9.21: lifetime spectrum of extra pulses found by David's fitter. The pulses follow the muon lifetime so they have little effect on gain shift.
Chapter 10

Data Analysis

10.1 The detector timing and gain stabilities

A measurement of the muon lifetime to 1 ppm requires that over the course of the measurement period, the gain and time pickoff of our front end electronics be extremely stable. The timing and gain stabilities are thus defined as the invariance of the detector response to the timing and amplitude changes of the standard impulse during the measurement period. By standard impulse we mean those normal pulses that overcome the certain threshold and are used for productive analysis. During the long period (from run to run) there is a slow but manifest gain change, mainly due to the temperature and HV effects on the PMT. Even inside 22 $\mu$s measurement period, there are much more pulses at early time than that at late time and the recorded property of the same pulse may vary depending on when the pulse arrives, during the course from scintillator, light guide to PMT. And the real impact on muon lifetime measurement is the relative early to late time and gain shift in the 22$\mu$s measurement period instead of the long run shift.
We do not know the exact models of the timing and gain shifts during the real experiment and it’s common to assume linear changes, at least we can estimate the range of the shift.

Figure ?? shows the effect of the timing shift. The timing shift

\[ t \rightarrow t' = t(1 + \delta) \]

will cause the lifetime shift

\[ \tau \rightarrow \tau' = \tau(1 + \delta) \]

where \( \delta \) is the slope of the timing shift with time. The timing shift of 1 ppm (left plot) will cause the lifetime shift of 1 ppm (right plot).

\[ \delta t/t = 1 \text{ ppm} \]

\[ \tau = \tau(1 + 1 \text{ ppm}) \]

\[ \tau' = \tau + \delta \tau \]

**Figure 10.1:** The relationship between timing shift (left) and lifetime shift (right). It’s easy to translate the timing shift to lifetime shift, the shift on time (slope) will cause the same shift on the lifetime.

Figure ?? and ?? show the effect of the gain shift. The gain shift can be separated
to two parts: the gain change will cause counts change in each time bin (figure ??), and the counts change vs. time will cause lifetime shift. The counts change of
\[ N(t) \rightarrow N'(t) = N(t)(1 + \delta t) \]
will cause the lifetime shift
\[ \tau \rightarrow \tau' = \tau(1 + \delta \tau) \]
where \( \delta \) is the slope of the counts change with time.

The derivation of above gain shift is simple, originally we have the lifetime formula
\[ N(t) = N_0e^{-t/\tau} \]
and for each time bin the count was shifted from \( N(t) \) to \( N'(t) \), and the new formula is
\[ N'(t) = N_0e^{-t/\tau'} \]
so we have
\[ N(t)(1 + \delta t) = N_0e^{-t/\tau'} \]
or
\[ N_0e^{-t/\tau}(1 + \delta t) = N_0e^{-t/\tau'} \]
which is the same as
\[ e^{-t(1/\tau' - 1/\tau)} = 1 + \delta t \]
and finally we have
\[ \frac{\tau'}{\tau} = \delta \tau' \approx \delta \tau \]
The counts change of 1 ppm (left plot) in one lifetime will cause the lifetime shift of 1 ppm (right plot).
The gain shift can be taken as reversed threshold shift and thus will change the number of particles above certain threshold. The relationship between count change and threshold change is shown in figure 10.2 and calculated as

$$\frac{\Delta N}{N} = \frac{n \Delta h_{TH}}{N} = \frac{n}{N} \cdot h_{TH} \cdot \frac{\Delta h_{TH}}{h_{TH}}$$

where $N$ is the total counts above certain threshold $h_{TH}$ and $n$ is the count in one amplitude bin at threshold $h_{TH}$.

The relative threshold change is equal to the gain shift

$$\frac{\Delta h_{TH}}{h_{TH}} = \frac{\Delta G}{G}$$

where $G$ is the detector average amplitude which can be taken as the peak from the amplitude distribution, as shown in figure 10.2 and 10.3, either from tile positron pulses.
Figure 10.3: The relationship between threshold shift and count change. The count change is simply equal to the count at threshold times the threshold shift.

or laser pulses.

Finally, we can bring all together and calculate the gain shift as

$$\frac{\delta (\Delta N)}{\delta t} = \frac{n}{N} \cdot h_{TH} \cdot \frac{\delta (\Delta G)}{\delta t}$$

where \( \frac{n}{N} \) and \( h_{TH} \) can be achieved from amplitude spectrum of decay positrons; \( \frac{\Delta G}{\Delta t} \) can be achieved from either positron spectrum or laser pulse spectrum.

To get good precision analysis on timing and gain shift study, the utility of laser system has many advantages compared with data itself.

1. The laser source sends out single energy and laser pulses have narrower amplitude distribution and better timing resolution than the tile pulses.
Figure 10.4: The amplitude distribution for the Michael positrons from data. The horizontal axis is the ADC and the vertical axis is count. The best fit function is the Landau convoluted with Gaussian and MPV from Landau was taken as the peak value of the amplitude.

2. The laser source is isolated from the beam area so it’s not contaminated with noises that are related to beam.

3. For timing shift, we need the reference PMT as reference time 0 which is required to have better timing resolution than the ball PMT. The isolated reference PMT will be put in another no-beam-contaminated place to record laser signals.

4. The tile pulses are made up of Michael positrons of different energies and have an amplitude distribution that is characterized by landau convoluted with
**Figure 10.5:** The amplitude distribution for the reference PMT from laser pulses. The horizontal axis is the ADC and the vertical axis is count. The fit function is Gaussian and mean value from Gaussian was taken as the peak value of the amplitude.

Gaussian (see figure ??). The gain was taken as the MPV (most probable value) of the fitting of the data. The laser pulses come from laser source of narrow energy range and follow the narrower and simple Gaussian distribution and it’s much easier to get the gain from the Gaussian mean value of the fit (see figure ??).

Since 2006 the laser system was applied to the MuLan experiment to study both the timing and gain shift of the detecting system. In 2006 we took 3305 useful laser runs, in which about $5.6 \times 10^6$ laser pulses have been recorded for each of 24 laser channels.
The laser system setup is shown in figure ?? . We mounted a highly stable, nitrogen laser source outside the ball area. Laser pulse from the laser box was sent to cascaded splitter boxes on the rack of the MuLan ball, where the laser light was further distributed to 24 representative phototubes on the ball, as well as to a reference PMT and photodiode located outside the zone. During the laser run, the laser box was operated at a constant frequency of about 37 Hz, or a period of 27 ms, so roughly one laser pulse will be recorded every 1000 fills (27 ms/27 μs=1000). The arrival times of these laser pulses are randomly distributed across the whole measurement period.

10.2 Timing Stability

In the experiment we don't know about the true arrival time. So all the best we can do is to use the reference PMT as the best estimate of arrival time. Reference PMT has no beam related or electrical associated contamination and is much better on timing (δt sigma) than tile PMT. We measure the time pickoff stability over the measurement period by noting the δt between the time of a laser signal produced by a tube on the ball and that by the reference PMT. Roughly speaking, if the average δt does not change significantly over the measurement period, we can regard the time pickoff signals on the ball as stable at the same level.

To measure the gain and timing stability in the fill at the 1 ppm level, we need combine the 3000 laser runs to get the best statistics. Figure ?? shows the average δt
for SN=98 over 400 runs. The drift is roughly 0.02 clock ticks, which is only a factor 10 larger than the sorts of early-to-late time-pickoff drifts that we hope to measure and much smaller than the resolution of our individual timing measurements.

One of the practical problems one encounters in the laser analysis is the so-called start-time ambiguity. The Start signals of each WFD will not be aligned perfectly in phase, but the offset between pair of channels in each WFD will be constant (shift by 4 clock ticks). As a result, depending on exactly when Start comes with respect to each of the clocks in the two channels, there could be a difference of exactly 4 clock ticks between two \( \delta t \) measurements, even if the true \( \delta t \) were exactly the same, as seen in figure ??

As long as both the reference and tile WFDs remain powered, the internal 125 MHz clock phases will not change and the position of the peaks and \( \delta t \) mentioned above will not move. However, if one WFD is power cycled, the time 0 of that WFD can move by a multiple of the 500 MHz external clock (4 times of the internal frequency). As a result, if runs are combined across power cycles, other peaks, offset by multiples of the clock ticks, can appear in the \( \delta t \) spectra.

These multi peaks can be easily combined to one peak because they are always shifted by multiples of the clock ticks, see Fig. ??

The fit function I chose to fit the \( \delta t \) plots is the gaussian, the best method found so far. The fit was applied over a limited range (a little more than two sigmas around the combined peak, in either direction), to get a better fitting accuracy.
The times of the pulses were taken from David’s production, which used decay positron pulse shapes as a fitting template. Although there is some variation with tube, and position on the ball, the typical timing resolution is about 0.10-0.15 clock ticks (220-330 ps). Figure ?? shows the average width (sigma) of the δt for all 24 tile PMTs.

The δt plots are extremely clean. As one would expect in our very low rate experiment, where at t=0 the probability of visible pileup in 5 clock ticks window is only about 0.2%, there is no noticeable accidental background, as can be seen in figure ??.

A sample fit to the δt vs. time is shown in figure ??c. The figure of merit in the plots is the POL1 slope, which quantifies the change in δt from early to late in the fill. The numerator of the slope is the change in measurement time, in clock ticks. The denominator is also in clock ticks. The slope of $3.1 \times 10^{-8}$ implies a fractional change of 3 parts in $10^8$ over the measurement period. The error, which is roughly twice as large in this case, indicates that the fractional change over the measurement period is consistent with 0 and, at the two-sigma level, no larger than 0.12 ppm. The summary plot, figure ?? tells the same story. The mean slope is consistent with 0 and the RMS width of $6 \times 10^{-8}$ and roughly 25 degrees of freedom, indicates that the statistical error on that mean is about 5 times smaller, or $1.2 \times 10^{-8}$. The overall error, even at the two sigma level, is much less than 0.1 ppm.

In summary, the message from the timing stability studies is clear: there are no
significant time shifts over the course of the measurement period.

10.3 Gain Stability

Gain stability is measured in a similar spirit as timing stability. Extra care were
taken when combining runs. The amplitude shift in one fill is our focus point but
the amplitude is not extremely stable over many runs so we should not just simply
combine runs for amplitude vs. time study. Instead, we firstly looked for variables
that are stable over the runs.

If the fraction of the laser output to each channel isn’t stable, or there is no
correlation between the reference amplitude and those of the tubes on the ball, or
fluctuations from tile PMT photostatistics are far greater than laser output fluct-
uations, normalizing to the reference phototube amplitude on pulse by pulse basis
provides no improvement on gain stability study.

An alternative is to normalize not to the reference tube but to, say, the average
amplitude of the tube under test, taken at the end of the measurement period. As
with normalizing to the laser reference, this procedure will remove any longer-term
drift in the average amplitude (which is not of interest to us) and facilitates the
combination of runs with very different average amplitudes.

The systematic drift in laser pulse height is obvious when examined over many
runs. A plot of the ratio of the mean amplitude ratio of SN 234 with respect to the
amplitude of the reference PMT is shown in figure ???. There are big fluctuations in
the first 240 laser runs, and the ratio becomes more stable in the last 160 runs.

Pulse by pulse scatterplots of tile amplitude vs. reference PMT amplitude and tile amplitude vs. "the next pulse falling in the last histogram time bin" are shown in figure ?? and ??, respectively. Neither correlation is very impressive. Over longer time periods, say, averaging over a run, the correlation is obvious, as shown in figure ?? and ?? . The correlation for the second method is a little tighter. So in the gain stability study, we chose the amplitude ratio to that in the last time bin as our main method.

Based on the discussions above, we used two different techniques to access the gain shift - normalizing our results to the reference PMT's signal and to the same tile's signals from the last time bin of our histogram.

The typical amplitude spectrum for a single run (SN 336) is shown in figure ?? . The sigma is about 13\%, typical of the 24 tile tubes, and perfectly consistent with the roughly 80 photoelectrons we would expect to gather. If we normalize each amplitude vs. time to the amplitude in the last time bin we can eliminate some of the irrelevant variation. Figure ?? shows this ratio vs. time in one run and figure ?? shows the slope (of the ratio vs. time) vs. run, and it's manifest that the slope is very stable over 400 runs. The distribution of the average of the ratio, minus 1, is shown in figure ?? . The mean is near 0, naturally.

To see if there is any sign of an early-to-late shift in the data, we plot the means (with 1 subtracted from each) by time bin and then fit the data to a POL1. A
sample fit is shown in figure ???. The slope is small and consistent with 0. It should be remembered that the units of the slope are fractional gain per time bin, where the time bin, in this case, is about 1 muon lifetime.
Figure 10.6: amplitude spectrum for gain shift study
**Figure 10.7:** Laser system setup: laser reference PMT is outside of ball barrack, and splitters and 24 laser tile PMTs are in ball area. Figure is taken from
**Figure 10.8**: Average $\delta t$ between SN 98 and the reference PMT over the course of 400 runs

**Figure 10.9**: $\delta t$ between SN 98 and the reference PMT in one run 54330, there are two discrete groups due to WFD Start time phase shift, and they are always separated by integer number of clock ticks
**Figure 10-10:** $\delta t$ between SN 125 and the reference PMT over the course of 400 runs. The peak on the left with blue color line is a simple sum of all peaks with certain numbers of clock shifts. The gaussian fits were also applied to combined peak (blue color) and the largest peak (red color).

**Figure 10-11:** $\delta t$ sigma for all 24 tile laser channels. The width is around 0.10-0.15 clock ticks for most channels. Laser channel number 17 is worse because of the tube is not good.
Figure 10.12: $\delta t$ distribution of SN 26 over 400 runs. Although there maybe multiple bumps, there is very little background.

Figure 10.13: change in pickoff time ($\delta t$) vs. time in measurement period of SN 26 over 400 runs.
Figure 10.14: slope of $\Delta t$ vs. time for 24 tiles

Figure 10.15: Average amplitude ratio (SN 234 vs. reference PMT) vs. laser run number
**Figure 10.16:** Scatter plots of tile pulse height vs. reference PMT pulse height - for one run and selected 4 tiles.

**Figure 10.17:** Scatter plots of tile pulse height in time bin 5 vs. next tile pulse height in time bin 10 - for one run and selected 4 channel.
Figure 10.18: Scatter plots of average tile pulse height vs. average reference PMT pulse height - for 400 runs and selected 4 tiles.
Figure 10.19: Scatter plots of average tile pulse height in time bin 5 vs. average of next tile pulse height in time bin 10 - for 400 runs and selected 4 channel.

Figure 10.20: Typical amplitude spectrum for one run - from time bin 5 and SN 336
Figure 10.21: ratio vs. time in one run where ratio is the average amplitude of each time bin to that of the last time bin - SN 336

Figure 10.22: slope of ratio vs. run where slope is the POL1 fit of ratio defined in previous plot - SN 204
Figure 10.23: Typical normalized amplitude spectrum for 400 runs - from time bin 2 and SN 97
Figure 10-24: Relative gain shift vs. time (late time normalization): 400 runs, the gain shift is about $1.65 \pm 1.49 \times 10^{-4}$ per lifetime.

A histogram of all 24 slopes for the tile detectors is shown in figure ???. The mean is consistent with 0 and the RMS width of these average slopes is about $2 \times 10^{-4}$, implying an error on that mean of $4 \times 10^{-4}$ and a 95% confidence limit on the fractional gain change of $4.0 \times 10^{-5}$.

A similar set of slopes, this time using reference PMT as normalization, is presented in figure ???. The mean is once again completely consistent with 0 and the error on the average slope is nearly identical.

The fractional change in counts (dN/N) for a given change in threshold can be found via a histogram of the WFD pulse height spectrum, as can be shown in figure ???. 1 ADC change in threshold will cause about $3.0 \times 10^{-4}$ fractional change in counts.
The lifetime shift caused by the amplitude gain shift can be calculated as:

\[
\frac{\Delta N}{N} = \frac{n \cdot \Delta h_{TH}}{N} = \frac{n}{N} \cdot h_{TH} \cdot \frac{\Delta h_{TH}}{h_{TH}} = \frac{n}{N} \cdot h_{TH} \cdot \frac{\Delta G}{G} = 3.0 \times 10^{-4} \times 60 \times 4 \times 10^{-5} = 0.7
\]

which is 0.7 ppm per \( \tau \).
Figure 10.25: Relative gain shift slopes: 24 tile channels, Last time-bin self-normalization. The average gain shift is about $0.5 \pm 1.8 \times 10^{-4}$ per lifetime.
Figure 10.26: Relative gain shift slopes: 24 tile channels, Reference PMT normalization. The average gain shift is about $0.4 \pm 1.9 \times 10^{-4}$ per lifetime.
Besides the amplitude change, the pedestal shift can also be seen as changes in the effective threshold of the discriminators. When the pedestal changes, the threshold change is simply equal to the pedestal change ($dP$). So the total threshold shift by gain and pedestal is

$$dR = (dG/G) \cdot R + dP$$

![Histogram](image)

**Figure 10-27:** Pedestal vs. time for all the tile channels combined. Pedestal shift during 10 lifetime is 19.719-19.702 = 0.017 ADC.

The pedestal shift can also be taken from either the real data or the laser data. The pedestal vs. time from the real data (50% pass) is shown in figure ??_. The pedestal changes by 0.017 ADC counts across the whole fill of 10 muon lifetime, giving a 1.7e-3 change in ADC counts per muon lifetime, this would cause the fractional change in counts per lifetime of

$$3 \times 10^{-4} \times 1.7 \times 10^{-3} = 0.51ppm$$
and finally the systematic error from pedestal changes is

\[ \text{frac} \delta \tau \tau = 0.51/4 = 0.13 \text{ppm} \]
10.4 Timing Stability vs. $\Delta t$

in PFA section

10.5 Gain Stability vs. $\Delta t$

The gain stability we have talked so far is assuming there is only one pulse recorded by each detector in the measurement period. And this is true because the average number of pulse in each detector element is about $20/170 = 0.12$. But there is still the possibility that two or more pulses are recorded by single detector during the fill. And if the amplitude of shadow pulses (the pulses that happen after the trigger pulse) is affected by trigger pulse for whatever reason, then it's like a gain shift for these shadow pulses, the counts over threshold will be affected. For these shadow pulses we might lose or gain the PU reconstruction by some level.

The method for gain vs. $\delta t$ study is similar to that for gain shift study. The ratio method was used to combine all the runs and laser channels for statistics. The ratio is the mean value of the shadow laser pulses divided by the mean value of the trigger normal pulses, for each time bin, as shown in figure ??.

Figure ?? shows the gain shift of shadow pulses vs. $\Delta t$ where $\Delta t = t_{\text{shadow}} - t_{\text{trigger}}$.

From this figure we can see clearly that the shadow pulses will be affected significantly by the trigger pulses if they arrive very close to the trigger pulses.
Figure 10.28: amplitude ratio vs. $dt$ for the laser shadow pulses for SN=3 in single run 68932. \( \text{ratio} = \frac{\text{amplitude}_{\text{shadow laser}}}{\text{amplitude}_{\text{trigger normal}}} \) and $dt = t_{\text{shadow}} - t_{\text{trigger}}$.

Figure 10.29: gain shift of shadow pulses vs. $dt$ for all the laser channels. 3000 laser runs in 2006 data set. \( \text{ratio} = \frac{<\text{amplitude}_{\text{shadow laser}} >}{< \text{amplitude}_{\text{trigger normal}} >} \) and $dt = t_{\text{shadow}} - t_{\text{trigger}}$. 
10.6 Gain Stability from the Real Data

This is another approach to study the gain shift from the real data only. The method is simply to plot the MPV vs. time and get the gain change directly.

a) original MPV vs. time

Figure ?? is the original MPV vs. time plot. There are two trends in the plot: at early time (before 3500 c.t.) the target muon decays dominate and the MPV vs. time has a small slope; at late time (after 4000 c.t.) the ratio of background to target muon decays becomes bigger and bigger and the slope is much bigger too. The MPV of the two components are: MPV of pulse = 116.6 ADC and MPV of background = 117.3 ADC; they are close to each other meaning that the background may come from sneaky muons which have the similar property as the decay muons. To get the gain shift vs. time we need first subtract the background. To naive choice is the last time bin where the ratio of background is the biggest, but the last time bin has the data truncation problem because it's close to the hardware time cut. So the best choice should be a few time bin before the last one.

b) background subtracted MPV vs. time

Figure ?? shows the comparison of different backgrounds' selection. Here I did 9 loops scan to study the effect by changing the background bin from 95 to 87. Firstly it shows that the background subtracted MPV vs. time is pretty flat and we subtracted the background successfully, at least up to 8000 clock ticks.

Secondly it also shows that the tail of MPV vs. time is not flat, and very sensitive
to the background bin selection. At late time bin, where pulse/background ratio is smallest (6% at bin 95), the tail is going downward, as background bin moves earlier, the pulse/background becomes bigger (14% at bin 87), the tail is going upward. Further study shows that background subtracted MPV vs. time is not flat if the pulse bin is close, say, 10 time bins distance, to the background bin.

By fitting the background subtracted MPV vs. time by linear function and choosing fit range from 1500 clock ticks to 7800 clock ticks to skip the early oscillation and late trend, as shown in figure ??, we got a fitting result of $6.6 \times 10^{-7} \pm 2.7 \times 10^{-7}$ ADC/c.t. slope, this is about

$$6.6 \times 10^{-7} \times 1000/116 \approx 6 \times 10^{-6}$$

gainshift in one lifetime, which is smaller than that from the laser study.

**Figure 10.30**: mpv vs. time, no background subtracted. At early times the positrons dominate and slope is flat, at late times the background dominates and the slope goes up. all runs in 2006 data set.
Figure 10.31: mpv vs. time, background subtracted. The background was chosen from the last bin to the 9th to the last bin for comparison. The trend much flatter than original mpv vs. time trend.

Figure 10.32: mpv vs. time, background subtracted. 4 examples were shown here for comparison. The trend much flatter than original mpv vs. time trend.
10.7 Timing Shift from Overflow Pulses

The pulse fitter is good at normal pulses but may lose accuracy when dealing with overflow pulses which have one or more truncated ADC snaples at maximum allowed value of 255 ADCs and thus may much different from pulse template.

To study the uncertainty about the timing of overflow pulses and how it will affect the timing of the coincidence pulse, we can learn it from the time difference of the inner outer coincidence when there are overflow pulses. For each fitter, I plot the $dt = t_{overflow} - t_{normal}$ for each coincidence where the coincidence window is 6 clock ticks and the overflow pulses are defined as the ones with amplitude bigger than 230 ADC (after pedestal subtracted). Figure ?? and ?? are the $dt$ vs. time obtained from 8 runs (run 40100 to 40108), and from David’s fitter and Kevin’s fitter respectively.

The two plots give us the similar results. The $dt$ range is smaller than 0.06 clock ticks in 10 lifetimes or $\frac{0.06}{10^7} = 6$ ppm, but overflow pulses only occupy about 10% of data, so the total uncertainty is smaller than 0.6 ppm.
Figure 10.33: mpv vs. time, background bin 95 subtracted.

Figure 10.34: overflow timing shift, from David’s fitter. dt vs. time where dt=t(overflow)-t(normal).
Figure 10.35: overflow timing shift, from Kevin's fitter. \( dt \) vs. time where \( dt = t(\text{overflow}) - t(\text{normal}) \).