Towards Precision Measurement of the Neutron Lifetime using Magnetically Trapped Neutrons

A thesis presented

by

Liang Yang

to

The Department of Physics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Physics

Harvard University

Cambridge, Massachusetts

May 2006

©2006 - Liang Yang

All rights reserved.

To my parents

Thesis advisor

John M. Doyle

Author

Liang Yang

Towards Precision Measurement of the Neutron Lifetime using Magnetically Trapped Neutrons

Abstract

Preliminary results from a neutron lifetime measurement using magnetically trapped neutrons are reported. Ultracold neutrons are loaded into a 1.1 T deep Ioffe-type trap through inelastic scattering of 0.89 nm neutrons with superfluid ^{4}He. Decays of trapped neutrons are detected using the scintillation light produced by the interaction of energetic decay electrons with superfluid helium. The neutron trap lifetime after the suppression of above threshold neutrons with field ramping is measured to be 831 ± 58 s, consistent with other lifetime measurements. Studies of two leading systematic uncertainties: neutron absorption on ^{3}He and the presence of above threshold neutrons are discussed in detail. Successful development of a 3.1 T deep Ioffe trap for the next generation apparatus is also described.

Contents

Acknowledgments

First and foremost, I want to thank my advisor John Doyle for his guidance and support throughout my graduate career, for creating an exciting and enriching work environment, for sharing with me his knowledge and wisdom on physics as well as different aspects of life and for exemplifying the highest standards of intellectual pursuit.

I am grateful to Paul Huffman who introduced me to low temperature physics and taught me so many experimental techniques and life lessons. He is the first person I turn to for help with difficult technical issues, and he never failed in providing creative and practical solutions.

Thanks to my former fellow graduate students Dan McKinsey, Carlo Mattoni and Sergei Dzhosyuk. Their ingenuity and hard-work are essential to the success of the project. Their good nature and humor made even the most tedious experimental work a joyful experience. I could not have ask for better colleagues and friends. I also appreciate the help of other Doyle group members, especially Bob Michniak for spending countless hours with me in the Harvard Machine shop winding the Mark II Ioffe trap.

I appreciate the support of the NIST Neutron Interaction and Dosimetry Group and its staffs. Special thanks to Muhammad Arif for his encouragement and financial support, Alan Thompson for being our data acquisition system guru, Jeff Nico for engaging lunch conversations on possums and Chinese movies, and Pieter Mumm for careful editing of this thesis.

I would like to acknowledge the contributions of collaborators on the project. In particular, Robert Golub for his creativity and telling us everything that we did

wrong; Kevin Coakley for his guidance in the neutron trajectory simulations; and Richard Pardo and his group at ATLAS for their assistance in the helium purity measurements. I also thank Kyle Alvine, Clint Brome, James Butterworth, Laurens van Buuren, Antonio Copete, Joshua Gabrielse, Ekaterina Korobkina, Tomasz Kreft, Stephen Maxwell, Pieter Mumm, Jeremy Olson, Grant Palmquist, Chris O'Shaughnessy and Pil-Neyo Seo for their contributions to the project and fruitful discussions.

I thank my officemates at NIST, Mike Huber and Dima Pushin, whose NBA anecdotes and Russian jokes always cheered me up when the experiment did not work.

I am grateful to my parents and brother who have always believed in me. Though separated by a distance, their examples and advice have sustained me through the difficult times.

Finally, I offer my deepest gratitude to my wife, Haiyan. Your encouragement and support are the constant source of my inspiration and strength. Your love and kindness have filled my life with joy and happiness.

Chapter 1

Introduction

The properties of the neutron and its beta decay parameters have been studied by physicists for over 70 years since the discovery of the neutron by Chadwick [1]. Precision measurements of these parameters have provided us with the ability to perform very strict tests of fundamental theories.

The value of neutron lifetime τ_n in particular, is an important input into tests of the Standard Model. By combining τ_n with neutron beta decay asymmetry coefficients, one can extract the matrix element $|V_{ud}|$ in the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix; $|V_{ud}|$ is the largest term in CKM Unitarity tests [2]. In addition, the neutron lifetime plays an important role in the theory of Big Bang Nucleosynthesis (BBN). Currently, the uncertainty in the neutron lifetime measurement is the primary source of uncertainty in the theoretical estimation of ^{4}He production from the Big Bang [3].

The neutron lifetime measurements using the material storage of ultracold neutrons have so far reached an experimental precision of 10^{-3} [4, 5]. However, a 5.6σ

discrepancy between the two most precise measurements have put the quoted experimental uncertainties from storage experiments in doubt. In this thesis, a technique of measuring the neutron lifetime using magnetically trapped ultracold neutrons will be discussed. This technique, though currently has a limited precision, has the potential to overcome the limitations of the material storage measurements, and eventually reach a measurement precision of 10^{-4} [6].

1.1 Neutron Beta Decay

Neutrons and protons are the basic building blocks of nuclei. Though decay modes of the proton are predicted by Grand Unified Theories, an actual proton decay has never been observed experimentally. The current experimental limit puts the proton decay lifetime at greater than 6.7×10^{32} years [7]. Unlike the proton, a free neutron is unstable. Because the mass of the neutron ($m_n = 939.5656$MeV) is greater than the mass of the proton ($m_p = 938.2720$MeV) by 1.2946 MeV, it's energetically possible for a free neutron to decay into a proton, an electron ($m_e = 0.511$MeV), and an electron antineutrino ($m_{\bar{\nu}_e} < 1$eV) with a release of 782.6 keV in kinetic energy. Such a decay also conserves charge, baryon number and lepton number. It is indeed observed with a lifetime of approximately 15 min:

$$n \rightarrow p + e^- + \bar{\nu}_e + 782.6\,\text{keV}. \tag{1.1}$$

According to the Standard Model, the neutron decay can be modeled as one of the two down quarks in the neutron decaying into an up quark by exchanging a charged W^- boson with an electron and an electron anti-neutrino. This is shown schematically

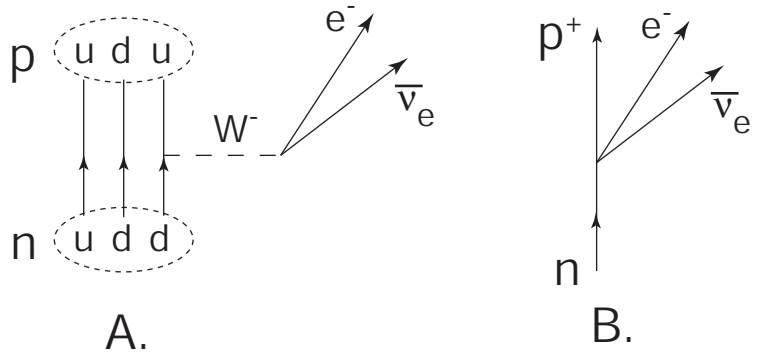

Figure 1.1: The beta decay of a neutron. A) In the Standard Model, the decay is mediated by the W^- boson. B) Due to the low energy release in neutron decay, a four fermion point interaction is sufficient to describe the process.

in Figure 1.1A. Because of the large mass of the W boson ($m_W = 80.4$GeV), its range is only on the order of 0.001 fm. Low-energy beta decay processes such as the neutron decay, can thus be adequately approximated by a point interaction of four fermions. The Feynman diagram for the point interaction decay model is shown in Figure 1.1B.

According to the weak current-current interaction theory first developed by Fermi [8] and later generalized by Feynman and Gell-Mann [9], the Hamiltonian operator of a four fermion point interaction can be written as

$$H = \frac{G_F}{\sqrt{2}} J^{c\mu} J_\mu^{c\dagger}, \tag{1.2}$$

where G_F is the Fermi coupling constant. The upper index c denotes that we only discuss charge current interactions. The weak interaction current J_μ^c can be written as the sum of hadronic and leptonic currents,

$$J_\mu^c = h_\mu^c + l_\mu^c. \tag{1.3}$$

The weak interaction currents are a combination of vector and axial-vector couplings. The leptonic current

$$l_\mu^c = \bar{\Psi}_e \gamma_\mu (1 - \gamma_5) \Psi_{\nu_e} \tag{1.4}$$

couples the leptonic doublets such as (e, ν_e), while the hadronic current

$$h_\mu^c = \bar{\Psi}_{d'} \gamma_\mu (1 - \gamma_5) \Psi_u \tag{1.5}$$

couples the quark doublets such as (d', u). Here d' is the weak eigenstate of the down quark, which is different from its mass eigenstate d. The weak and mass eigenstates are related by the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.6}$$

Because quarks are bound inside nucleons, the strong force can modify the axial coupling of the weak force, while the conservation of vector current (CVC) hypothesis postulates that the vector coupling is not affected by the strong force. Therefore, the hadronic current in the neutron decay can be written as:

$$h_\mu^c = V_{ud} \bar{\Psi}_p \gamma_\mu (1 + \lambda \gamma_5) \Psi_n = G_F^{-1} \bar{\Psi}_p \gamma_\mu (g_v + g_a \gamma_5) \Psi_n, \tag{1.7}$$

where λ is the relative axial-vector coupling constant and is related to the vector and axial-vector form factors, g_v and g_a via $\lambda = g_a/g_v$.

Applying Fermi's Golden Rule on the Hamiltonian, the neutron differential decay rate can be written as [10]:

$$\frac{d^3\Gamma}{dE_e d\Omega_e d\Omega_\nu} = \Phi(E_e) G_F^2 V_{ud}^2 (1 + 3\lambda^2) \Big(1 + a \frac{\vec{p_e} \cdot \vec{p_\nu}}{E_e E_\nu} + A \frac{\vec{\sigma} \cdot \vec{p_e}}{E_e} + B \frac{\vec{\sigma} \cdot \vec{p_\nu}}{E_\nu}\Big), \tag{1.8}$$

where $\vec{\sigma}$ is the neutron spin, E_e, E_ν, $\vec{p_e}$, and $\vec{p_\nu}$ are the energies and momenta of the electron and antineutrino respectively, and $\Phi(E_e)$ is a phase space factor. The correlation coefficients, a, A, and B in the tree-level approximation depends only on λ:

$$a = \frac{1-\lambda^2}{1+3\lambda^2} \tag{1.9}$$

$$A = -2\frac{\lambda^2+\lambda}{1+3\lambda^2} \tag{1.10}$$

$$B = 2\frac{\lambda^2-\lambda}{1+3\lambda^2}. \tag{1.11}$$

In practice, the value of λ can be obtained from any of the correlation coefficients above. However the value of B is very insensitive to λ and a is difficult to experimentally measure because the energy and momentum of the neutrino can be only measured indirectly. Measurements of A currently yield the most precise determinations of λ. The world average of λ determined from four independent measurements of A is -1.2695 ± 0.0029 [11]. The particle data group has chosen to scale the error by a factor of two to accommodate significant differences between individual measurements [11].

By averaging the decay differential cross section in Equation 1.8 over the neutron spin and integrating over the energies of the decay products, we obtain an expression for the neutron lifetime,

$$\tau_n^{-1} = \frac{m_e^5 c^4}{2\pi^3 \hbar^7} f^R G_F^2 |V_{ud}|^2 (1+3\lambda^2), \tag{1.12}$$

where f^R is a phase-space factor that includes radiative corrections and can be calculated theoretically. From Equations 1.10 and 1.12, we can see that $|V_{ud}|$ can be extracted from measurements of τ_n and A.

1.2 Motivation

1.2.1 CKM Unitarity

The CKM matrix defines the rotation of the quarks' mass eigenstates into their weak eigenstates. Proper normalization of the weak eigenstates requires the matrix to be unitary, which means $V^{\dagger}V = 1$. Non-unitarity of the CKM matrix would be a strong indication of physics beyond the Standard Model, for example a fourth generation of quarks. Experimentally, the unitarity condition of the first row, $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$, is the most sensitive test. The 2004 published Particle Data Group's (PDG) evaluations of these matrix elements are [11]:

$$|V_{ud}| = 0.9738 \pm 0.0005 \tag{1.13}$$

$$|V_{us}| = 0.2200 \pm 0.0026 \tag{1.14}$$

$$|V_{ub}| = 0.00367 \pm 0.00047 \tag{1.15}$$

By applying the unitarity condition, one obtains:

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.9967 \pm 0.0015 \tag{1.16}$$

which deviates from unitarity by 2.2 standard deviation.

Many theoretical and experimental efforts have been devoted to the study of this deviation. Recent developments in Kaon decay experiments [12] suggest a shift in $|V_{us}|$ that would close this discrepancy. Similarly, a new lifetime measurement [5] yielded a value that is significantly different from previous measurements. If this value is taken alone, it also shifts $|V_{ud}|$ sufficiently to close the discrepancy. But if both shifts are applied, there would again be a 2σ deviation from unitarity, though on the opposite

side. While the Koan result has been confirmed by two experiments [13], the neutron lifetime result hasn't been corroborated by any other experiment and this is at present not be widely accepted. These new results have generated tremendous excitement in the field in the last couple of years. Here we will briefly review the measurements that lead to determination of the first row elements in the CKM matrix.

Because of the small value of $|V_{ub}|$, both the central value and uncertainty do not significantly contribute to the unitarity test. We thus skip the discussion of the determination of $|V_{ub}|$ from B meson decay and concentrate on the other two terms in the test.

$|V_{us}|$ is determined from either Hyperon or Kaon decays. Because $|V_{us}|$ has not been the dominant uncertainty in the CKM unitarity test, the majority of experiments focused on the $|V_{ud}|$ measurement. New results on the K^+_{e3} branching ratio from Brookhaven E865 collaboration however came out in 2003 [12], yielding a value of $|V_{us}| = 0.2272 \pm 0.0030$, 2σ away from the previous world average. The PDG incorporated this result into their 2004 analysis which raised the world average value of $|V_{us}|$ to 0.2200 ± 0.0026. New measurements of K_L decay from the KTeV collaboration at Fermi lab yielded $|V_{us}| = 0.2252 \pm 0.0026$ [13]. This result has not yet been incorporated by the PDG into the world average, but will again shift $|V_{us}|$ higher. Stimulated by the experimental developments, theorists also revisited calculation of theoretical corrections. Several new radiative correction calculations now yield results differing by as much as 2σ. For example, a new analysis of hyperon decay data shifts the originally published value of $|V_{us}|$ upward to 0.2250 ± 0.0027 [14]. These results, if taken alone, could solve the CKM unitarity problem. Several experiments

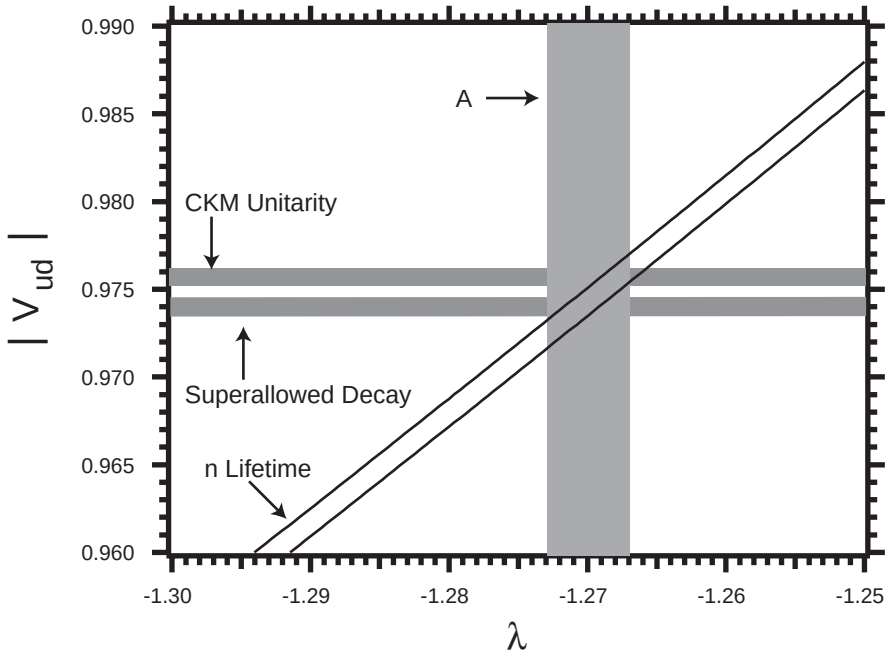

Figure 1.2: Experimental constraints on $|V_{ud}|$, including CKM unitarity requirement, measurements from superallowed beta decays, and the neutron lifetime and A (β asymmetry) measurements.

and theoretical studies are continuing to investigate this possible shift in $|V_{us}|$.

The error in the measurement of $|V_{ud}|$ provides the largest contribution to the uncertainty in the CKM unitarity test. $|V_{ud}|$ can be determined from superallowed $0^+ \rightarrow 0^+$ beta decays, neutron beta decay and rare pion decay. Figure 1.2 shows the current experimental constraints on $|V_{ud}|$ and λ.

The $|V_{ud}|$ value obtained from measurements of superallowed $0^+ \rightarrow 0^+$ beta decays currently has the smallest experimental uncertainty. In a superallowed decay between 0^+ states, the axial vector part of hadronic weak current is suppressed. Therefore, it is possible to determine $|V_{ud}|$ from measurements of the decay partial half-life t and

the statistical rate function f. Assuming conservation of vector current, one obtains,

$$V_{ud}^2 = \frac{K}{2G_F^2(1+\Delta_R^V)(1+\delta_R')(1+\delta_{NS}-\delta_C)ft}, \tag{1.17}$$

where $K = 2\pi^3\hbar^7 \ln 2/m_e^5c^4$, Δ_R^V is the nucleus-independent part of the radiative correction ($\approx 2.4\%$), δ_R' and δ_{NS} are the nucleus-dependent parts of the radiative correction ($\approx 1.4\%$), and δ_C is the isospin symmetry-breaking correction ($\approx 0.4\%$) [15]. The ft value can be measured very accurately for superallowed decays in isotopes such as ^{14}O, ^{34}Cl and ^{46}V, often better than 0.1%.

Hardy and Towner have combined all available data from 20 superallowed $0^+ \rightarrow 0^+$ decays to obtain a new evaluation of $|V_{ud}| = 0.9738 \pm 0.0004$ [16]. The contribution to the total uncertainty from the experimental uncertainty is only 0.00008. The uncertainty in $|V_{ud}|$ is dominated by theoretical corrections; Δ_R^V (0.00037) and $(\delta_{NS} - \delta_C)$ (0.00015).

Recently, a new method of measuring the Q value (decay energy) of of neclei using a Penning trap has improved the precision that one can measure f. A measurement in ^{46}V yields a new f value that is statistically inconsistent with the 7 previous Q-value measurements of ^{46}V [17]. Although the new measurement does not substantially shift the $|V_{ud}|$ value, it indicates that there could be unknown or underestimated systematic errors in at least some of the ft measurements, additional measurements in other nuclei could significantly change the present global consistency demonstrated in Hardy and Towner's evaluation. Further experimental and theoretical efforts are needed to resolve this discrepancy in Q measurements for the superallowed decays.

Free neutron decay provides an alternative system that is theoretically cleaner. The isospin symmetry-breaking correction term δ_C disappears, and the nucleus-dependent

radiative corrections $\delta_R^{'}$ and δ_{NS} are understood at the 10^{-5} level. The only large theoretical uncertainty comes from the nucleus-independent correction Δ_R^V described above. Experimentally, because neutron decay has both vector and axial vector components, one much measure both the neutron lifetime and a correlation coefficient to determine $|V_{ud}|$. The expression for $|V_{ud}|$ from neutron decay is

$$V_{ud}^2 = \frac{K/ln2}{G_F^2(1+\Delta_R^V)(1+3\lambda^2)f(1+\delta_R)\tau_n}, \tag{1.18}$$

where $\lambda = g_A^{'}/g_V^{'}$[1] and $f(1+\delta_R) = 1.71489 \pm 0.00002$ [18]. The current world average for the neutron lifetime is 888.7 ± 0.8[2]. Details of the lifetime measurements will be discussed in section 1.3.

The uncertainty in λ currently dominates the uncertainty in $|V_{ud}|$ extracted from the neutron decay. The most precise value of λ is obtained from measurements of the spin-electron asymmetry coefficient A using a highly polarized cold neutron beam. The four most precise measurements of A are inconsistent with each other. The PDG determination of the world average for A is -0.1173 ± 0.0013, with the error scaled up by a factor of 2.3 to accommodate for the inconsistency. Using this value of A value, the value of $|V_{ud}|$ obtained from neutron decay data is 0.9713 ± 0.0013 [19], in agreement with the superallowed Fermi decay result. Several experimental efforts are underway to significantly improve the measurement of A. For example, the UCNA collaboration based at Los Alamos National Laboratory hopes to improve the experimental precision of A by a factor of 3 within next few years using polarized ultracold

[1]The theory we developed in Section 1.1 does not include radiative corrections. In a more complete theory, the vector coupling constant g_V becomes $g_V^{'} = g_V(1+\Delta_R^V)^{1/2}$ and the axial coupling constant g_A becomes $g_A^{'} = g_A(1+\Delta_R^A)^{1/2}$. Since $\lambda = g_A^{'}/g_V^{'}$ is determined experimentally, the axial-vector radiative correction factor $(1+\Delta_R^A)$ is not needed for a determination of $|V_{ud}|$.

[2]This excludes a recent measurement that is 5.6σ [5] away from the average [11].

neutrons [20]. At that point, the neutron lifetime will become the limiting uncertainty in the determination of $|V_{ud}|$.

$|V_{ud}|$ can also be extracted from the rare pion beta decay, $\pi^+ \rightarrow \pi^0 e^+ \nu$. Pion decay is a simpler system theoretically than both superallowed decays and neutron decay. The $0^- \rightarrow 0^-$ pion decay is a pure vector transition, and therefore does not require nucleus-dependent corrections. But its small branching ratio ($\approx 10^{-8}$) however , makes it very difficult to measure experimentally. The value of $|V_{ud}|$ obtained from pion decay is 0.9716 ± 0.0039 [21]. Because of the large uncertainty, the pion data does not contribute to the world average for $|V_{ud}|$.

In short, to improve the unitarity test of the CKM matrix, the uncertainties in both $|V_{ud}|$ and $|V_{us}|$ need to be reduced. $|V_{us}|$ is currently limited by experimental uncertainties, while $|V_{ud}|$ is limited by the theoretical correction term Δ_R^V in superallowed Fermi decays. In fact, the same term Δ_R^V will become the theoretical limit for neutron decay and pion decay once experimental precisions in these measurements improve. Marciano and Sirlin recently proposed a method of calculating radiative corrections using Bjorken sum rules, that can reduce the theoretical uncertainty in Δ_R^V by a factor of 2 [22]. If the new result stands, the nucleus dependent corrections will become the dominant term in the extraction of $|V_{ud}|$ from superallowed decays, and measurements of neutron decay can significantly improve the accuracy of $|V_{ud}|$. An improved value of $|V_{ud}|$ obtained from neutron beta decay measurements will also serve as an important systematic check against the value obtained from the superallowed Fermi decay data.

1.2.2 Big Bang Nucleosynthesis

The neutron lifetime also plays an important role in the theory of Big Bang Nucleosynthesis (BBN). According to the big bang theory, baryons are created from the initial high-temperature and high-density quark-gluon plasma. As the universe expands and cools, light elements such as deuterium and helium (A < 12)start to form. One success of BBN theory is that it can correctly predict the abundances of light elements in the early universe. All other cosmology models give much lower predictions than the observed primordial abundance of D and ^{4}He. The neutron lifetime is an important input parameter for the BBN theory. In particular, it presently dominates the uncertainty in the theoretical calculation of ^{4}He primordial abundance [3, 23].

During the first second of the Big Bang, the weak interaction rates are higher than the universe expansion rate. Neutrons and protons are in thermal equilibrium by the reactions $n \leftrightarrow p + e^- + \bar{\nu}$, $n + \nu \leftrightarrow p + e^-$, and $n + e^+ \leftrightarrow p + \bar{\nu}$. The neutron-to-proton ratio is set by the Boltzmann factor, $n/p = e^{-Q/T}$, where $Q = m_n - m_p = 1.29$ MeV, and T is the temperature of the Universe. When the temperature drops below the "freeze-out" temperature, $T_F \sim 0.8$ MeV ($t \sim 1$ s), the rate of these weak interaction reactions become less than the expansion rate. Neutrons and protons are no longer in thermal equilibrium and the neutron-to-proton ratio is frozen out at $e^{-Q/T_F} \approx 1/6$. The exact value of T_F depends on the weak interaction matrix element $G_F^2(1 + 3g_A^2)$, which can be expressed in terms of the neutron lifetime τ_n. After the freeze-out, the neutron-to-proton ratio decreases monotonically due to neutron beta decays. After about 1 to 3 minutes ($T \sim 0.3$ to 0.1 MeV), the density of deuterium photo-disassociation photons (~ 1 MeV) drops low enough that nucleosynthesis starts

to take place. The neutron-to-proton ratio n/p at this moment is approximately 1/7. The exact ratio again depends on the neutron lifetime. During the first few minutes of nucleosynthesis, D, ^{3}He, ^{4}He and ^{7}Li are produced. Because ^{4}He is the most tightly bound nucleus among the light elements, practically all neutrons are bound into ^{4}He nuclei[3]. The mass fraction of ^{4}He, Y_p, can be easily estimated as $2(n/p)/(1+n/p) = 25\%$. More precise calculations using Monte-Carlo simulation technique yields $Y_p = (0.2491 \pm 0.0005)$. The leading uncertainty comes from the experimental uncertainty in the neutron lifetime measurement, $\delta Y_p/Y_p = 0.8\delta\tau_n/\tau_n$ [24].

The primordial ^{4}He abundance can be measured through observations of metal poor extragalactic gas clouds. The two most recent measurements report values $Y_p = (0.2452 \pm 0.0015)$ [25] and $Y_p = (0.2391 \pm 0.0020)$ [26]. The reported values are not consistent with each other. It is believed that the systematics in the experiments have been underestimated, and σ_{sys} could be as large as 0.005 [27]. This inconsistency has also led some researchers to question the accuracy of the neutron lifetime measurement. A shorter neutron lifetime such as that measured by Serebrov's group will lower the theoretical prediction of Y_p, bringing it into closer agreement with the experimental observations [28]. However, without sorting out the systematic issues in the Y_p measurements themselves, it would be premature to use it as an indication of problems in the neutron lifetime measurement.

Primordial nucleosynthesis is the most stringent test of the Big Bang theory during the first few minutes. As experimental measurements on the primordial ^{4}He abundance become more accurate, precise and accurate knowledge of the neutron lifetime will be required to improve its theoretical prediction.

[3]The ratios of D/H and ^{3}He/H are around 10^{-4}, while the ratio of ^{7}Li/H is less than 10^{-9}.

Table 1.1: A summary of measurements of the neutron lifetime used to determine the world average.

Measured lifetime (s)	Method	Reference
$886.8 \pm 1.2 \pm 3.2$	Beam, Penning trap	Dewey 03 [29]
$885.7 \pm 0.9 \pm 0.4$	UCN double bottle	Arzumonov 00 [4]
$889.2 \pm 3.0 \pm 3.8$	Beam, Penning trap	Byrne 96 [30]
882.6 ± 2.7	UCN bottle	Mampe 93 [31]
$888.4 \pm 3.1 \pm 1.1$	UCN bottle	Nesvizhevsky 92 [32]
887.6 ± 3.0	UCN bottle	Mampe 89 [33]
891 ± 9	Beam	Spivak 88 [34]

1.3 Neutron Lifetime Measurement

The current world average of the neutron lifetime is (885.7 ± 0.8) s [11]. It is determined from seven measurements using two basic techniques and these are summarized in Table 1.1. Each individual measurement agrees with the average within its respective error bar. The most recent measurement however reports a value of ($878.5 \pm 0.7_{\text{stat}} \pm 0.3_{\text{syst}}$) s, about 6.5 standard deviations away from this world average [5], (see Figure 1.3). This new value, despite being the most precise measurement to date, has not yet been included in the world average. It is deemed by the particle data group as too different from other measurements to be included; the particle data group warns that until this major disagreement is resolved, the present average for the neutron lifetime will be in doubt [11].

The neutron lifetime has been measured using three types of experimental techniques: beam, material bottle and magnetic confinement. Each has its own set of

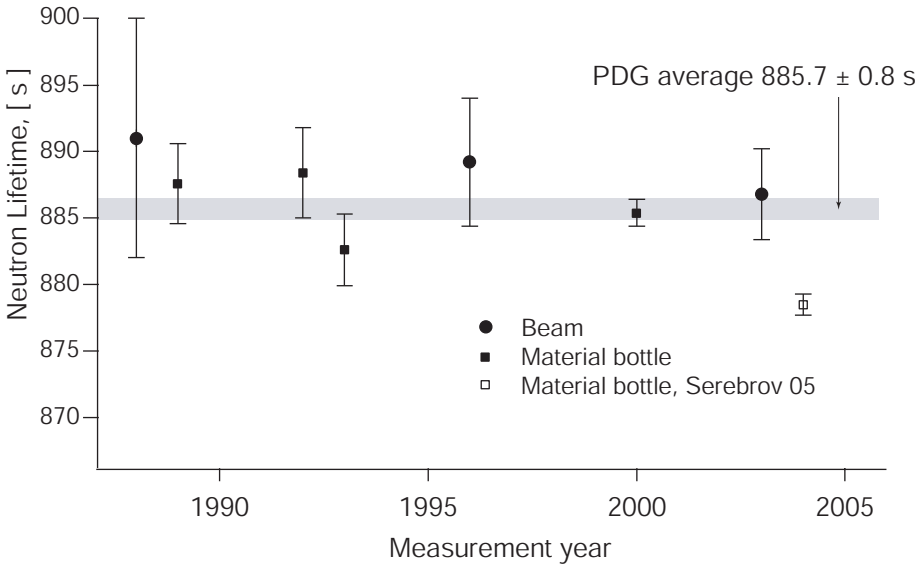

Figure 1.3: The neutron lifetime measurements with uncertainties less than 10 s. The most recent measurement by Serebrov's group is significantly different from the world average.

systematic errors. Here we will discuss each technique in detail and given specific examples of experiments utilizing these methods.

1.3.1 Beam Experiments

In beam type experiments, a cold neutron beam traverses through a well-defined decay region. The protons produced by neutrons that decay inside this region are detected[4], with rate of proton production being equal to the neutron decay rate. From the exponential decay law, $N_n = Ae{-}t/\tau_n$, it is easy to obtain

$$\frac{dN_p}{dt} = -\frac{dN_n}{dt} = \frac{N_n}{\tau_n} \tag{1.19}$$

[4]The decay electrons can also be detected, but due to large number of background electrons, most beam experiments detect protons instead of electrons.

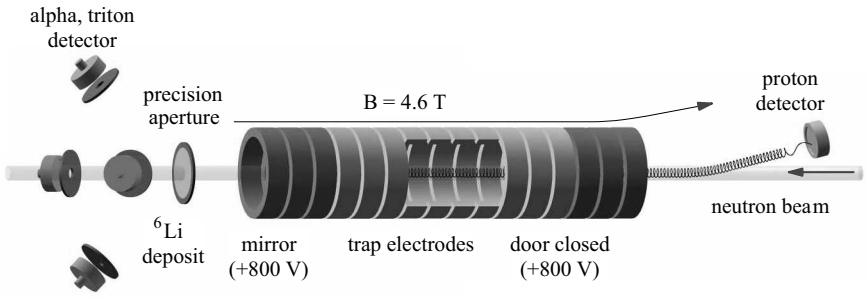

Figure 1.4: Schematic of the NIST neutron lifetime experiment [29]. Details of the experiment are discussed in the text

where N_n is the mean number of neutrons inside the decay region, N_p is the number of decay protons and τ_n is the neutron lifetime.

In the most recent beam type experiment conducted at the National Institute of Standards and Technology (NIST), a quasi-penning trap is used to trap the decay protons; a magnetic field provides radial confinement and end electrodes provide axial confinement (see Figure 1.4). The trapped protons are accelerated into the proton detector after fixed accumulation times. This trapping technique not only increases the solid angle for proton detection, but significantly suppresses background events by reducing the measurement time relative to the accumulation time. The detected decay proton rate is $\dot{N}_p = \varepsilon_p N_n/\tau_n$, where ε_p is the combined proton trapping and detection efficiency.

The average number of neutrons in the trapping volume can be expressed as

$$N_n = L \int_A da \frac{I(v)}{v}, \tag{1.20}$$

where $I(v)$ is the velocity dependent fluence rate of the neutron beam. L is the trap length, and A is the cross-sectional area of the neutron beam. In order to measure N_n, a "$1/v$" detector is used to cancel out the velocity dependent part of N_n. After

exiting the trapping region, the neutron beam goes through a thin foil deposit of ^{6}LiF. The reaction rate of ^{6}Li(n,t)^{4}He is inversely proportional to the neutron velocity v. Four silicon surface barrier detectors are used to count the decay products, alphas and tritons. Assuming the detection efficiency of the decay products for thermal neutrons ($v_0 = 2200$ m/s) is ε_0, then the detection efficiency for neutrons with velocity v will be $\varepsilon_0 v_0/v$, and the detected alpha and triton rate $\dot{N}_\alpha$ becomes

$$\dot{N}_\alpha = \varepsilon_0 v_0 \int_A da \frac{I(v)}{v}. \tag{1.21}$$

From the above equations, we obtain

$$\tau_n = \frac{\dot{N}_\alpha}{\dot{N}_p}(\frac{\varepsilon_p}{\varepsilon_0 v_0})L. \tag{1.22}$$

For a precision measurement of τ_n, each term on the righthand side of Equation 1.22 must be measured to equal precision. The trap length L is difficult to measure directly because of trap end effects. Because of this, the experiment is performed with different end-electrode distances, as the changes in L can be measured very precisely. In addition, the proton detection efficiency ε_p measurement is complicated by proton backscattering. τ_n is measured with detectors having different backscattering rates, then extrapolated to obtain the zero backscattering value. Finally, the uncertainty in thermal neutron detection efficiency ε_0 is dominated by uncertainties in the ^{6}LiF foil areal density and ^{6}Li thermal neutron cross section. The published result for the experiment is ($886.8 \pm 1.2_{\rm stat} \pm 3.2_{\rm syst}$) s, limited primarily by systematic uncertainties from ε_0 (σ_{ε_0}=2.8 s) [29]. A technique using a cryogenic radiometer to measure neutron fluence has the potential to calibrate the LiF foil absolutely to better than 0.1% [35], improving the lifetime measurement uncertainty to less than 2 s.

In beam type experiments, because both neutrons and protons need to be counted absolutely, detector efficiencies cannot be canceled and inevitably dominate the systematic uncertainty. It will be very difficult to push σ_{τ_n} below 1 s with this method.

1.3.2 Material Storage Experiments

In material storage type experiments, low energy neutrons are confined inside material bottles[5]. A consistent number of neutrons N_{ni} is loaded into the bottle. The number of remaining neutrons N_{nf} is then counted after a time t. From the exponential decay law, we obtain the neutron bottle lifetime $\tau_{bot} = t \ln(N_{ni}/N_{nf})$. As long as the neutron detector efficiency remains constant, the efficiency cancels when taking the ratio and thus does not need to be measured absolutely in contrast to the beam method. However, since the bottle lifetime τ_{bot} is now a combination of neutron decay, wall upscattering and other loss mechanisms, corrections or extrapolations need to be applied to extract the neutron lifetime from τ_{bot}.

Material storage experiments are possible due to the availability of ultracold neutron sources. Ultracold neutrons (UCN) are neutrons with such low energies (long wavelengths) that they can be totally reflected from material walls. Because the wavelength of UCN is much longer than the interatomic spacing of the material wall, their interaction with the nuclei in the wall can be approximated by an effective Fermi wall potential V,

$$V = \frac{2\pi\hbar^2}{m_n} N a, \tag{1.23}$$

where m_n is the mass of the neutron, and N and a are the nuclear number density

[5]In some cases gravity provides the vertical constraint.

and coherent neutron scattering length of the nuclei in the material [36]. A typical value of V is about 100 neV, corresponding to a UCN temperature of about 0.8 mK.

Producing and accumulating UCN is a challenging task. Current UCN experiments can only achieve a UCN density of 1 – 5 UCN/cm^3. UCN sources in development based on superthermal production promise to increase the density by 2 – 3 orders of magnitude. If realized, these sources would allow experimenters to significantly improve the statistical precision of measurements using UCN [37, 38, 39].

The neutron bottle lifetime τ_{bot} can be expressed as

$$\frac{1}{\tau_{bot}} = \frac{1}{\tau_n} + \frac{1}{\tau_{wall}} + \frac{1}{\tau_{other}}, \tag{1.24}$$

where $1/\tau_{wall}$ is the UCN loss rate due to wall collisions[6], and $1/\tau_{other}$ is the UCN loss rate due to other loss mechanisms such as UCN valve leakage, gaps in the wall coating, background gas collisions, etc. In most experiments, the loss rate $1/\tau_{other}$ can be directly studied and minimized, so the limiting systematic uncertainty typically comes from the UCN loss rate $1/\tau_{wall}$. The simple model outlined above suggests that UCN with energy less than the wall potential can be totally internally reflected, and the neutron material bottle lifetime should be very close to the neutron beta-decay lifetime. In early experiments, however, a very short bottle lifetime (around 100 – 200 s) was observed. It was soon realized that surface contamination by hydrogen is the primary contributor for the short bottle lifetime. Coating the material wall with fully-fluorinated polyester (fomblin oil) proved to be an effective way to reduce hydrogen upscattering, resulting in bottle lifetimes of 500 – 700 s [36]. Recently, a

[6]Here we assume a simple model for the wall loss. Because the wall collision rate depends on the UCN energy, and the UCN energy spectrum changes with time, the wall loss rate can be time dependent. In analysis of experiments, this effect is often corrected using Monte Carlo simulations.

new type of low temperature fomblin oil (LTF) that remains a liquid when cooled to liquid nitrogen temperature[7], was discovered. By cooling the UCN storage chamber, including the fomblin coating, storage times of over 800 s have been attained [5]. Although these bottle lifetimes are close to the neutron lifetime, some type of extrapolation to zero wall loss is still required. In most material bottle experiments, either the volume to surface ratio of the UCN bottle or the UCN energy spectrum is changed to vary the UCN wall collision rate. The neutron lifetime is obtained by extrapolating τ_{bot} to its value at a wall collision rate of zero.

Currently, the two most precise measurements of the neutron lifetime both utilize material bottle storage. In the double bottle experiment by Morozov's group [4], inelastically upscattered UCN (a significant source of trap loss) are detected by thermal neutron detectors surrounding the UCN bottles. Measurements are performed with two different bottle sizes to vary wall collision rate. They obtain a value of neutron lifetime of $\tau_n = (885.4 \pm 0.9_{\text{stat}} \pm 0.4_{\text{syst}})$ s. The leading systematic uncertainty reported arises from the calculation of the relative thermal neutron detection efficiencies between the two bottles and relies on Monte Carlo simulation.

In the gravitational trap experiment by Serebrov's group [5], UCN are confined in a material bottle with an open top. UCN with energies greater than the gravitational potential at the top of the bottle will escape from the opening. By tilting the bottle to different angles and thus lowering the height, different UCN energy spectra can be obtained. This experiment uses a low temperature fomblin oil coating and achieves storage times as high as 870 s. The published result for this experiments is $\tau_n = 878.5 \pm 0.7_{\text{stat}} \pm 0.3_{\text{syst}}$ s. The leading quoted systematic uncertainty comes from the

[7]Conventional fomblin oil will crack at -30 °C.

Monte Carlo simulation of the effective collision frequency of UCN in the bottle.

These two published results from the material bottle experiments differ from each other by 5.6σ. Although the gravitational trap result also differs from the world average by 6.5σ, there is no obvious reason to reject this result if it is not put in the context of the other measurements. Furthermore, because the UCN loss rate is much lower in the gravitational trap due to the use of low temperature fomblin coating, less extrapolation is required to obtain the lifetime value; possible systematic errors from the extrapolation should thus be smaller. Both experiments are preparing for upgrades that could reduce their uncertainties and shed light on this discrepancy in near future.

Although material storage experiments have provided the most precise measurements of the neutron lifetime to date, the large discrepancy between the two most precise measurements clearly shows that there are systematic errors that are not understood. Extrapolations and Monte Carlo simulations are the main suspects. Measurements having the same or better precisions using other methods are critical in resolving this discrepancy.

1.3.3 Magnetic Confinement Experiments

In magnetic confinement experiments, neutrons in one spin state are always pushed away from storage walls by the interaction of its magnetic moment with the magnetic field gradients. Therefore the systematics related to material wall interactions can be eliminated. Especially in light of the recent controversy in material storage experiments, a considerable research effort has been devoted to magnetic confinement

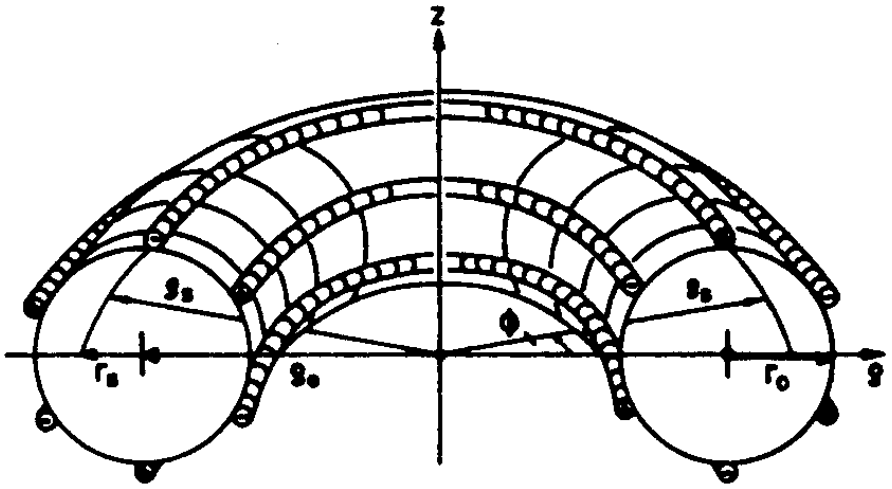

Figure 1.5: A schematic of the sextupole magnetic storage ring used in NESTOR experiment [41].

experiments. This technique was first proposed by Zeldevich in 1959[40] and has now been successfully demonstrated by several groups.

The first successful realization of this technique was known as the NESTOR experiment which used a magnetic sextupole toroidal field to confine circularly orbiting neutrons [41]. The field was created by superconducting wires arranged in the geometry as shown in Figure 1.5. The stored neutrons are not truly trapped; they have low velocities in the two transverse directions, but have a large longitudinal velocity. Although wall interactions are eliminated, betatron oscillations that couple longitudinal motion with transverse motion, can drive neutrons out of the storage ring. Neutron loss due to betatron oscillations is the leading systematic uncertainty. A statistically limited value of neutron lifetime of (877 ± 10) s was published [41].

A group led by Ezhov uses a combination of permanent magnets and gravity to

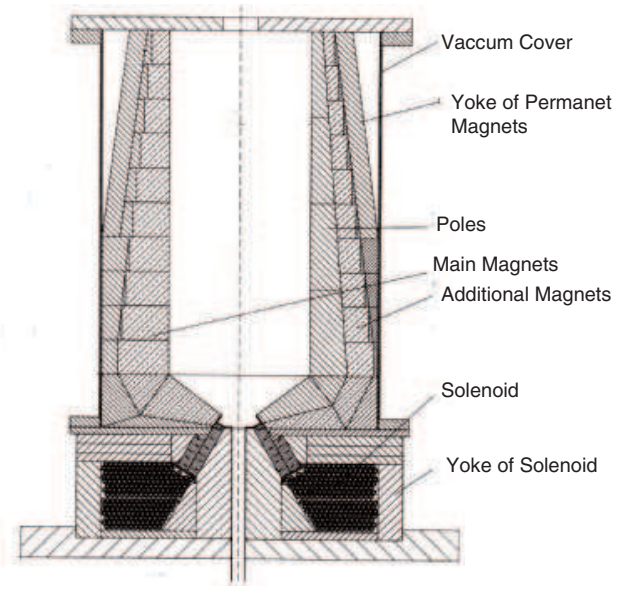

Figure 1.6: Cross-section view of the permanent magnet trap.

confine UCN [42]. A cross-sectional view of the trap is shown in Figure 1.6. Permanent magnets are arranged around the storage bottle to create a strong magnetic field gradient near the wall. UCN are confined vertically by the gravitational interaction. Neutron absorbing material is placed at the top of the trap to remove neutrons with energies higher than the trap depth. The collaboration is reporting a statistical accuracy of 4 s on their trap lifetime measurement, but the collaboration has not completed their analysis of systematic uncertainties. The main systematic uncertainty in this experiment arises from Majorana neutron spin flips which can occur in low field regions and stationary trajectories that have an energy above the trap potential.

The trapping of UCN with an Ioffe-type superconducting trap was successfully demonstrated by our group in 2000 [43]. Since then, we have improved the statistical

precision by a factor of ten, and with the next generation apparatus, a measurement of the neutron lifetime with a precision better than one second is achievable. The primary advantages of our method are the minimization of Majorana spin-flip losses, the elimination of above-threshold neutron orbits by field ramping, and the continuous observation of neutron decay events. Details of the experiment are described in section 1.4.

In addition to the experiments mentioned above, at least three other magnetic confinement experiments using various configurations of superconducting or permanent magnets are in either the proposal or development stage [44, 45, 46]. Magnetic confinement experiments currently do not reach the precision of material storage experiments, but their systematic uncertainties are believed to be easier to understand and control. It is therefore believed that this method can eventually reach a higher precision. In order to realize this potential, two challenges must be met. First, a much larger population of neutrons must be trapped in order to reduce the statistical uncertainty. And second, systematic uncertainties such as above-threshold neutron orbits need to be better understood and minimized. The UCN trapping experiment by our group is well prepared to meet these challenges as will be shown in this thesis.

1.4 Experimental Overview

The idea of using magnetically trapped UCN created inside a superfluid helium bath to measure the neutron lifetime was proposed by Doyle and Lamoreaux in 1994 [6]. An Ioffe-type magnetic trap is used to spatially confine UCN that are in the low-field-seeking state. Such a trap has no zero-field region, therefore avoiding

trap loss due to spin flips. UCN are created by the downscattering of 0.89 nm neutrons inside an isotopically pure superfluid ^{4}He bath. This so-called "superthermal process", described in Section 1.4.2, can produce a high UCN density within the trap. Neutrons travel undisturbed in the trapping region until they decay. The neutron decay events are continuously monitored by detecting the scintillation light from decay electrons passing through the helium bath. The neutron lifetime is obtained by fitting the measured rate of decay events to the exponential decay curve. This method is insensitive to the variation in the initial number of trapped neutrons from run to run, therefore eliminating another systematic uncertainty common to most other storage experiments, both material and magnetic.

1.4.1 Magnetic Trapping

The interaction between the magnetic moment $\vec{\mu}$ of a neutron and a magnetic field $\vec{B}$ is governed by the dipole interaction Hamiltonian,

$$H = -\vec{\mu} \cdot \vec{B} = \mu_n \vec{\sigma}_n \cdot \vec{B}, \tag{1.25}$$

where $\mu_n = 1.91\mu_N = 0.7$ mK/T, and $\vec{\sigma}_n$ is the Pauli spin matrix. Equation 1.25 also shows that the neutron's magnetic moment is anti-aligned with its spin. Because the neutron is a spin 1/2 particle, its spin can be in two spin states, parallel or antiparallel to the magnetic field $\vec{B}$. When the neutron's spin is parallel to the magnetic field, its potential energy in the field can be written as $H = \mu_n|\vec{B}|$. Neutrons in this spin state will seek to minimize their potential energy by moving towards low field regions. They are called the "low field seekers". Similarly, neutrons with spins antiparallel to the field are attracted to high magnetic field regions and are called the "high field

seekers".

When a neutron moves in an inhomogeneous magnetic field, its spin eigenestates change directions as the field directions change. But if the neutron moves slow enough, it can go adiabatically into the new eignestates, i.e., a neutron initially in the low field seeking spin state will stay in the low field seeking state, and its spin direction adiabatically follows the direction of the magnetic field. The probability of a spin-flip transition is proportional to $\exp(-\pi\gamma B^2/\dot{B})$, where $\gamma = 2\mu_N/\hbar$ is the gyromagnetic ratio of the neutron [47]. When the absolute value of the term in the exponent is much less than 1, the adiabatic condition will be satisfied. In the semi-classical picture, the neutron precesses around the local magnetic field at the Larmor frequency $\omega_L = \gamma B$. The adiabaticity condition is satisfied if the rate of the magnetic field direction change seen by the neutron is much less than its precession frequency

$$\frac{|d\vec{B}/dt|}{|\vec{B}|} \ll \gamma B. \tag{1.26}$$

One can reach the same adiabaticity condition in a full quantum mechanical treatment. Generally the adiabatic condition will tend to be violated if the neutron passes through zero or low-field regions.

Although Maxwell's equations do not allow the creation of a static magnetic field maximum in free space, it is possible to create a field minimum, and low field seekers can be trapped around it. In order to create a high trap depth, a large trapping volume, and to avoid zero field regions, an Ioffe type trap configuration shown conceptually in Figure 1.7 is used in the experiment. Four infinitely long wires with alternating current directions are arranged in a square configuration to generate a quadrupole field for radial confinement. Two solenoids positioned at both ends of

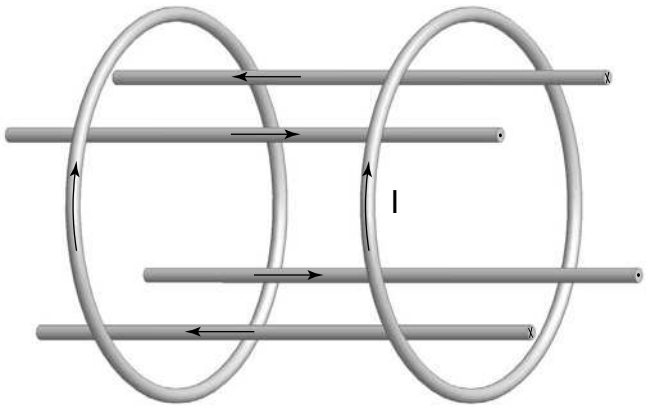

Figure 1.7: Conceptual design of an Ioffe-type trap.

trap provide axial confinement. The currents in the solenoids run in the same sense (Helmholtz configuration) to avoid creating a zero field region in the trap center; trap losses due to spin-flips is thus highly suppressed. The trap used in the recent experimental runs is 1.1 T deep with a trapping volume of 1.5 liters (see Section **??**). Approximately 4000 UCN are confined each time the trap is loaded. A next generation trap combining an accelerator type quadrupole magnet and two low current solenoids can reach a trap depth of 3.1 T with a trapping volume of 7.5 liters. We expect that the number of trapped UCN will increase by a factor of 20. The design and testing of this trap will be described in Chapter 6.

1.4.2 Superthermal Production of UCN

The neutron has a very small magnetic moment $\mu_n = 0.7$ mK/T, resulting in a very small magnetic interaction energy. Even in a 3 T deep trap, neutrons need to be cooled to below 2 mK before they can be trapped. At a typical research reactor facility,

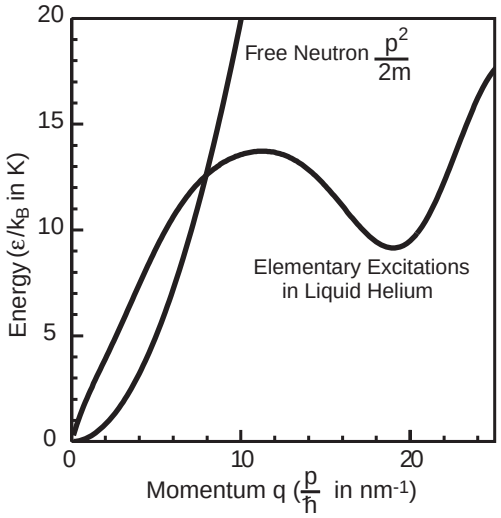

Figure 1.8: The dispersion curves of elementary excitations in liquid helium and a free neutron. The two curves intersect at two points corresponding to neutron with an energy of 0 K and 12 K.

fission neutrons are first moderated by heavy water to room temperature (T $\sim$ 300 K), then moderated to temperatures of $\sim$ 30 K using liquid hydrogen or deuterium moderators operating at $\sim$20 K. In the Leningrad ultracold neutron source, UCN were extracted from the cold moderator by the use of a vertical bent neutron guide. At the Institute of Laue-Langevin (ILL), a similar guide attached to a mechanical turbine is used to produce UCN. In these two UCN sources, neutron energies are lowered by forces derived from potentials, gravitational potential and a moving Fermi potential surface. According to Liouville's theorem, such forces cannot change the phase space density, which means the density of UCN produced by the UCN source can not exceed the UCN density that exists in the cold moderator. In practice, there is significant loss during UCN extraction and transportation, so the UCN density available in an experiment is much lower than the theoretical maxima [36].

It is possible to produce a UCN density higher than in the cold moderator using the technique of superthermal production. Because the entropy of a closed system goes down as its phase space volume decreases, a higher UCN density requires a reduction in the phase space volume and entropy of the neutron system. This can only be achieved by coupling the neutron system with another system, such as phonons in a solid or liquid, whose phase space volume and entropy increase as the UCN density increases. This allows the net entropy to increase. In addition, the thermalization time between the neutron and the other system needs to be long, so that UCN can be produced and remain out of thermal equilibrium, hence the name "superthermal".

The concept of a superthermal source based on neutron inelastic scattering in superfluid helium was first proposed by Golub and Pendlebury [48]. At low temperatures, the excitations in superfluid helium can be described by quasi-particles such as phonons and rotons [49]. Their energy-momentum dispersion curve intersects with the free neutron dispersion curve at two points (shown in Figure 1.8), corresponding to zero neutron energy and a neutron energy of 12 K. By conservation of energy and momentum, this implies that a neutron at 12 K (wavelength 0.89 nm) can down-scatter in liquid helium to near rest through the emission of a single phonon. The inverse process can be suppressed by lowering the 12 K phonon density. This density is proportional to the Maxwell-Boltzmann factor $e^{-12K/T}$, where T is the temperature of liquid helium bath. The UCN population can thus be thermally detached from the helium bath, allowing accumulation of UCN to a density as high as $P\tau$, where P is the superthermal production rate and τ is the UCN lifetime in the source [8].

[8] Solid deuterium and solid oxygen can also be used for superthermal production of UCN [37, 50]. UCN sources based on these moderators are currently under development. They aim to increase the available UCN densities to experiments by at least two to three orders of magnitude.

In our experiment, the neutron energy dissipation must occur within the trapping region due to the conservative nature of the trap. A 0.89 nm (12 K) monochromatic neutron beam passes through the liquid helium in the trapping region. UCN are produced when the 12 K neutrons downscatter to near rest via the single phonon emission process described above. UCN with energies below the trap depth and in the low-field-seeking state will become trapped. The superthermal production of UCN within the trapping region provides a mechanism for producing a high density of UCN within the trap.

We do not need to remove the liquid helium from the trapping region before measuring the neutron lifetime. the neutron beta-decay lifetime is modified at most by 10^{-5} due to the presence of the liquid [51]. Any UCN loss mechanism introduced by the presence of this helium must also be suppressed. UCN can gain energy and escape the magnetic field potential by scattering from phonons. As discussed earlier, the single phonon upscattering process is suppressed by a factor of $e^{-12K/T}$, where T is the temperature of the helium bath. Below 1 K, the two phonon upscattering process begins to dominate. Although the two-phonon upscattering rate at present has only been measured to temperatures as low as 750 mK [52], theory indicates that the rate is proportional to T^7. Extrapolation of the available data shows that at a helium bath temperature below 150 mK, the upscatter rate is less than $10^{-5}\tau_n^{-1}$. A second UCN loss mechanism is through absorption by ^{3}He atoms in the helium bath. While the neutron absorption cross section of ^{4}He is rigorously zero, ^{3}He has a very large neutron absorption cross section. The isotopic abundance of ^{3}He in natural helium is 10^{-7}. However, even such a low concentration limits the UCN lifetime in

helium to less than 1 s. The helium used in our experiment is isotopically purified by the"heat flush" technique developed by McClintock's group in England [53]. In theory, the ^{3}He isotopic abundance can be reduced to below 10^{-16} with this method, yielding a UCN absorption rate less than $10^{-5}\tau_n^{-1}$.

1.4.3 Continuous Detection of Neutron Decay Events

Neutron decay events are detected by observing the scintillation light created by the charged decay electrons as they move through the liquid helium. When an electron moves through liquid helium, it ionizes many helium atoms along its track. A large fraction of the helium ions quickly recombine into metastable He_2^* molecules that are in excited states. About half of these molecules are in the singlet state and decay within 10 ns by emitting extreme ultraviolet (EUV) photons with wavelengths ranging between 60 nm and 100 nm. About 35% of the initial electron energy goes into the production of EUV photons from these singlet decays, corresponding to approximately 22 photons per kiloelectron volt. For example, an electron with the most probable decay energy of 250 keV creates over 5000 prompt EUV photons. This strong flash of light constitutes the neutron decay signal that we detect. The other half of the He_2^* molecules are in the triplet state. They decay slowly with lifetimes ranging from microseconds up to 10 s. The photons emitted by the triple molecules therefore create a somewhat correlated weak background to the prompt signal.

Because EUV photons do not pass through normal optical windows, we frequency down-convert them to blue photons using the organic fluor tetraphenyl butadiene (TPB) coated on to a diffuse reflector (Goretex). This TPB-coated Goretex surrounds

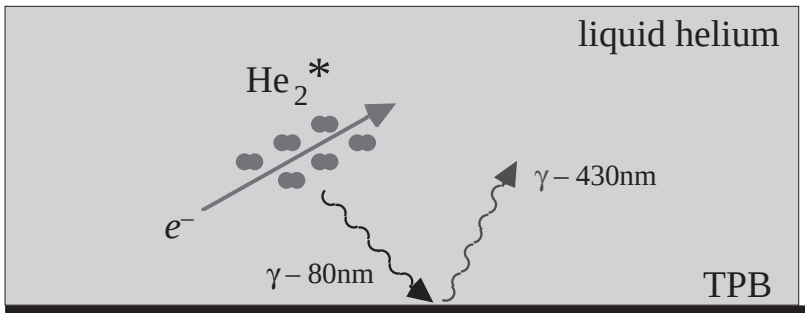

Figure 1.9: A schematic of the scintillation and energy down-conversion processes.

the trapping region. The blue photons are transported out of the apparatus through a transparent beam stop, several lightguides and optical windows, and are detected by two photomultiplier tubes (PMT) at room temperature operating in coincidence. The coincidence technique effectively suppresses background events from neutron-induced luminescence in materials surrounding the trapping region, PMT dark counts, and PMT afterpulses. After each coincidence trigger, the pulses from both PMTs are digitized and time tagged for later-analysis. The decay electron to light conversion process is shown schematically in Figure 1.9.

In traditional UCN lifetime experiments using material storage, UCN are stored for a certain time, then released and the number of remaining UCN is counted. This technique relies on the initial number of neutrons from run to run to remain constant. Our detection method allows us to observe the decay events *in situ*, and therefore measure the full decay curve. Suppose an initial number of UCN $N(0)$ is loaded in the trap, then the neutron decay count rate observed by our detector will be $N_d(t) = (N(0)\epsilon/\tau_n)e^{-t/\tau_n}$, where ϵ is the overall neutron decay detection efficiency[9].

[9]We have assumed that neutron loss mechanisms have been sufficiently suppressed such that the lifetime in the trap is dominated by the neutron lifetime.

As long as the detection efficiency remains constant during the time frame of one run, we can extract the neutron lifetime by fitting the detector count rate $N_d(t)$ to a single exponential decay curve, extracting both τ_n and $N(0)$. It is easy to see that with this technique, variations in $N(0)$ do not affect the measurement of τ_n. Furthermore, if the decay rate deviates from an exponential decay curve, it will provide an indication that additional UCN loss mechanisms or other systematic errors exist. This provideds us with a good handle on estimating systematic effects.

In our apparatus, a large number of background events are observed in addition to the neutron decay events. These events arise from ambient radiation, and neutron-induced activation and luminescence. This first of these terms gives rise to an overall constant background event rate and the neutron-induced terms are both time-dependent and correlated to the data-collection runs.

A background subtraction technique is employed to extract the neutron decay signals. To accomplish this, two kinds of runs are performed neutron trapping runs and background runs. In the trapping, or "positive" runs, the magnetic field remains on during both the trapping and observation periods. In the background, or "negative" runs, the magnetic field is on only during the observation period, so no neutrons are trapped. The difference between the data from the "positive" and the "negative" runs allows us to isolate the neutron decay data.

1.4.4 Context of this Work

The magnetic trapping of ultracold neutrons was initially demonstrated in 1999 [51] with a 1 T deep, 2 inch bore Ioffe trap (Mark I). Approximately 500 neutrons were

trapped and their decay signals were detected with 31% efficiency. The trap lifetime was measured to be 660^{+290}_{-170}.

I joined the group during the upgrade to the Mark II trap in 2000. In addition to a larger trap, a 0.89 nm neutron monochromator for background reduction and a better detection system were implemented for an upgraded apparatus. The work on the upgrades are reported in Ref. [54, 55] and also described in Chapter 4. Following the permanent installation of the monochromator and the upgrade of neutron cold source at the NIST reactor, the neutron trapping apparatus was run continuously for over a year. During that period, we discovered that the neutron trap lifetime at 300 mK (621 ± 18 s) was significantly shorter than the neutron beta-decay lifetime. We soon realized that the short trap lifetime was due to the presence of the above threshold neutrons. Using the field ramp technique to reduce the number of above threshold neutrons, we obtained a measured trap lifetime of 831 ± 58 s, which was consistent with the neutron lifetime. The analysis of these data runs can be found in Ref. [56] and in Chapter 5.

The realization that above threshold neutrons could be a significant source of systematic uncertainty prompted us to study them in detail. Both analytical and numerical methods were developed to understand the escape times of above threshold neutrons and the effectiveness of the field ramping technique. The results of these studies are described in Chapter 3.

To understand the systematic uncertainty from neutron absorption on ^{3}He, we have collaborated with Argonne National Lab for a series of accelerator mass spectroscopy (AMS) measurements of ^{3}He isotopic concentration in ultrapure helium.

Rather surprisingly, the preliminary measurements showed a much higher concentration of ^{3}He, leading to a large systematic uncertainty for the lifetime measurement. The AMS measurements are summarized in Chapter 2.

Discussions of other systematic uncertainties can be found in Section 5.6. It can be shown that all known systematic uncertainties could potentially be reduced to below 10^{-5} level.

In addition to studies of systematic uncertainties, we also initiated the development of the next generation apparatus which aims to substantially increase the statistical precision of the lifetime measurement. The design and testing of key components of the new apparatus are discussed in Chapter 6.

Finally, the road map for achieving a 10^{-4} measurement of the neutron lifetime is discussed in Chapter 7.

Chapter 2

Measurement of ^{3}He Abundance in Isotopically Pure ^{4}He

2.1 Introduction

Natural abundance helium consists predominantly of ^{4}He and a trace amount of ^{3}He. The isotopic ratio $R_{34} = {}^3\text{He}/{}^4\text{He}$ equals to 1.3 x 10^{-6} for atmospheric helium and (1 - 2) x 10^{-7} for commercial helium extracted from natural gas wells [57]. Although the presence of ^{3}He at such a low concentration does not pose any problem for the vast majority of applications, for a handful of experiments it is essential to have isotopically much purer ^{4}He. Specific examples include investigations of the supercritical velocity of negative ions at millikelvin temperature [58], the neutron electric dipole moment experiment (EDM) at Los Alamos National Lab [59] and the neutron lifetime measurement discussed here [6]. In the negative ion experiment, ^{3}He atoms tend to condense on the outside of the ions, therefore suppressing the emission

of rotons; the effect under investigation. An isotopic ratio of $R_{34} < 10^{-10}$ is necessary to prevent the condensation. In both the neutron EDM and lifetime experiments, ^{3}He atoms can absorb UCN, therefore causing unwanted neutron loss. While an isotopic ratio $R_{34} \sim 10^{-12}$ will be sufficient for the EDM experiment, the lifetime experiment puts a more stringent requirement on the isotopic ratio.

In the neutron lifetime experiment, assuming a uniform distribution of ^{3}He atoms, the UCN loss rate Γ_{abs} due to absorption by ^{3}He can be written as

$$\Gamma_{abs} = \int_v nR_{34}\sigma v\phi(v)dv, \tag{2.1}$$

where $n = 2.17 \times 10^{22}$ cm^{-3} is the number density of helium atoms, σ is the neutron absorption cross section of ^{3}He, v is the velocity of the neutron and $\phi(v)$ is the velocity distribution of the trapped neutrons. Because the neutron absorption cross section of ^{3}He is inversely proportional to the neutron velocity, σ can be expressed as $\sigma_{th}v_{th}/v$, where $\sigma_{th} = 5330$ barns is the thermal neutron capture cross section and $v_{th} = 2200$ m/s is the thermal neutron velocity. Equation 2.1 simplifies to $\Gamma_{abs} = nR_{34}\sigma_{th}v_{th}$, independent of velocity distribution. Plugging the numbers, we obtain the relative loss rate between ^{3}He absorption and beta decay ($\Gamma_\beta = 1/\tau_n$),

$$\frac{\Gamma_{abs}}{\Gamma_\beta} = 2.26 \times 10^{10} R_{34} \tag{2.2}$$

Therefore, in order to reduce the systematic correction from the ^{3}He absorption below $10^{-5}\tau_n$, an isotopic ratio $R_{34} < 5 \times 10^{-16}$ will be required for the neutron lifetime measurement.

Helium isotopic purification techniques include filtration through a superleak, differential distillation and heat flush. The superleak method filters the natural helium

through a leak so tiny that only the superfluid component of liquid helium, composed of ^{4}He, can pass through; the differential distillation methods relies on the vapor pressure difference between the two isotopes; and the heat flush method uses the mass flow of the normal component of superfluid helium to flush ^{3}He atoms away. Section 2.2 discusses these techniques in more detail. Typical isotopic ratios attainable by the superleak or distillation technique are $R_{34} < 10^{-12} - 10^{-10}$, while the heat flush technique has produced samples with indirectly measured $R_{34} < 5 \times 10^{-16}$ [53]. The ultrapure helium used in the lifetime experiment was manufactured with a continuous flow heat flush purifier developed by McClintock's group [60]. Although the apparatus is designed to produce samples with $R_{34} < 10^{-16}$, the sample isotopic ratio has never been directly measured and the best indirect limit is $R_{34} < 10^{-13}$. Therefore, a direct measurement of R_{34} is needed to confirm the purity of the samples produced by the apparatus. In addition, ultrapure helium can be easily contaminated if not handled properly, so it is necessary to check the helium isotopic ratio after the experiment to make sure no such contamination occurred. Even in the case that contamination is found, the systematic error it would introduce can be corrected to below $10^{-5}\tau_n$ level, if R_{34} can be measured with a precision better than 5×10^{-16}.

Isotopic ratio measurements with conventional mass spectroscopy have maximum sensitivities of 10^{-9}, due to interference from ions with the same charge-to-mass ratio as the less abundant isotope and are therefore unable to achieve the required precision. Accelerator mass spectroscopy (AMS) however should achieve sufficient sensitivity. Ions are accelerated to a few MeV before being detected by a particle detector. Stripping foils are often used to break up molecular ions and strip off electrons outside

atomic ions, so in the particle detector, each ion can be identified by its maximum charge and total mass, and the isotope peak under study will be clearly separated from the peaks of interfering ions. Isotopic ratio measurements with AMS can reach a maximum sensitivity of $10^{-15} - 10^{-16}$.

Since 2001, we have collaborated with Dr. Pardo's ATLAS group at Argonne National Lab to measure the isotopic purity of the ultrapure helium. The experimental facility at ATLAS is described in section 2.3, and results are presented in section 2.4. The measurements are consistent with an isotopic ratio $R_{34} \sim 10^{-12}$, higher than expected. However, we cannot yet rule out the possibility that such a high ratio is a result of natural helium contamination in the ion source; in particular large fluctuations in background ^{3}He counts are not well understood, because it is difficult to obtain the beam time needed to study the contamination and background issues carefully. Two alternative methods are proposed in section 2.5 as near term solutions for R_{34} measurements at $10^{-12} - 10^{-13}$ level.

2.2 Purification Techniques

2.2.1 Filteration through a Superleak

The filteration technique uses the vanishing viscosity of superfluid helium to separate it from impurities including ^{3}He. According to the two-fluid model of superfluidity developed by Tisza [61] and Landau [49], liquid helium below the λ point (superfluid transition temperature) can be described as containing two interpenetrating fluids, the "superfluid" which has exactly zero viscosity and the "normal fluid"

which behaves like normal viscose fluid. The fraction of superfluid increases as temperature decreases. The helium atoms in the superfluid state collectively form a single quantum state. ^{3}He atoms and other impurities cannot become part of the superfluid state, so they are carried around by the normal fluid. While the superfluid component can flow freely through extremely tiny capillaries (diameter $\sim 10^{-7}$ m) [62], the flow of the normal component is virtually stopped due to its viscosity. A superleak is made from small capillaries or porous material with pore sizes less than 1μm. It can stop the flow of normal fluid and any ^{3}He impurity, allowing only the superfluid fraction of liquid helium to go through. Therefore, filtering liquid helium below the λ point through a superleak can increase its isotopic purity significantly.

This purification technique was first demonstrated by Mezhov-Deglin in 1972. He estimated the isotopic ratio R_{34} to be less than 5×10^{-8} [63]. In 1975, Fatouros *et al.* reduced R_{34} to $(4 \pm 2) \times 10^{-10}$ by filtering superfluid helium through a Vycor superleak[1] [64]. Most recently, the neutron EDM collaboration at ILL claimed to have reduced R_{34} below 10^{-10} using a superleak made of compressed aluminum oxide powder with 20 nm grain size [65, 66].

The advantage of this technique is that it is easy to implement, and the purification can take place at the same time as the experimental cell is being filled. Its disadvantage is that the helium isotopic ratio R_{34} has a practical limit around 10^{-10}. Though theoretically it is possible to increase the isotopic purity by further reducing the superleak pore size, in practice, not only is it technically difficult to produce such a superleak, but it will also take unrealistically long time for appreciable amount of liquid helium to filter through if the pore size becomes too small. Another practical

[1] Vycor is a type of porous glass made by Corning Glass Works, Code 7930.

concern is that ^{3}He atoms can be trapped on the large effective surface area of the superleak and cause cross-contamination. In general, this technique can reduce the ^{3}He impurity level by three orders of magnitude below natural abundance, which is sufficient for most experiments. Helium with even higher isotopic purity can be obtained by the "differential distillation" or the "heat flush" technique.

2.2.2 Differential Distillation

The idea of differential distillation is to use the vapor pressure difference between ^{4}He and ^{3}He to preferentially pump away ^{3}He atoms; at low temperatures (< 2 K), the vapor pressure of ^{4}He becomes much lower than ^{3}He. In the case of dilute amounts of ^{3}He inside liquid ^{4}He, the notion of ^{3}He vapor pressure has to be modified as there is no liquid ^{3}He phase. However, it's possible to calculate the pressure of ^{3}He in the vapor phase above the liquid ^{4}He using the ^{3}He and liquid helium binding energy as well as the surface reflection probability [67]. The result shows that the helium in the vapor phase contains a much higher concentration of ^{3}He, therefore pumping away the vapor for a sufficiently long time can reduce the concentration of ^{3}He in bulk liquid ^{4}He [67].

Helium purification with differential distillation was first demonstrated by Tully in 1975 [68]. By distilling natural helium just below the lambda point, he was able to reduce the ^{3}He isotopic concentration by a factor of 500. The achievable sample purity at that temperature is limited by high ^{4}He vapor pressure as well as slow diffusion of ^{3}He atoms through the liquid ^{4}He. A recent study by the neutron EDM collaboration shows that this technique can work most efficiently between 0.3 K and

0.6 K. In this temperature regime, the phonon and roton density in superfluid ^{4}He is low enough that ^{3}He atoms can travel through liquid helium ballistically, not being limited by diffusion [69]. The temperature is also high enough that ^{3}He atoms are not trapped by the surface states of liquid ^{4}He. Such states occur at temperatures below 0.1 K. Calculation shows that at 0.5 K an isotopic purity of $10^{-11} - 10^{-12}$ can be achieved within a reasonable amount of time, though this has not been verified by experiments [67].

This purification method is also easy to implement using charcoal pumps, but the use of a ^{3}He fridge or a dilution fridge is required to reach the optimal distillation temperature. It is particularly suited for the neutron EDM experiment which requires sample purity of 10^{-12}, and whose optimal operating temperature is 0.5 K [67].

2.2.3 Heat Flush

The heat flush technique uses the mass flow of the normal fluid component of superfluid ^{4}He to flush away ^{3}He atoms. In the two-fluid model, the superfluid component cannot transport heat because it always has zero entropy, so all heat is transported by the normal fluid component. In fact, heat flow is accomplished through mass flow of the normal fluid. When a heat source is turned on inside superfluid helium, the normal fluid component flows away from the heat source to carry away entropy, while the superfluid component flows towards the heat source to conserve mass. The normal fluid flow not only drags away ^{3}He atoms close to the heat source, but also stops the back diffusion of ^{3}He. As a result, isotopically pure helium is produced around the heat source.

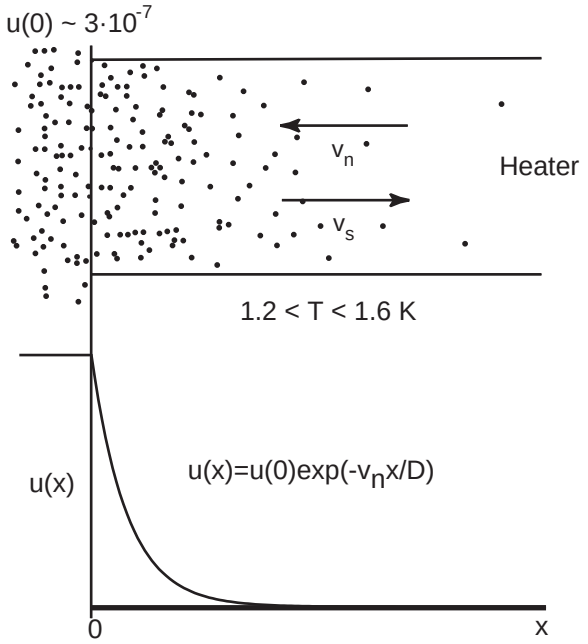

Figure 2.1: The upper graph depicts the flow of the normal and superfluid components during a heat flush of ^{3}He. The bottom graph shows the ^{3}He concentration in the flush tube under steady state condition.

The remarkable efficiency of the heat flush technique can be illustrated with a simple gas diffusion model [53]. Figure 2.1 shows an idealized schematic of the heat flush technique. A semi-infinite tube is open at one end to a liquid helium bath with a small ^{3}He concentration of u_0. Heaters establish a normal fluid flow with velocity v_n inside the tube. The flux of ^{3}He flow $\mathcal{F}$ inside the tube then can be described as gas diffusing against a moving medium (the normal fluid)

$$\mathcal{F} = -D\nabla u + uv_n, \tag{2.3}$$

where D is the diffusion coefficient of ^{3}He inside the normal fluid, and u is the number density of ^{3}He. When steady state is reached, there will be no net flow of ^{3}He atoms, so the divergence of $\mathcal{F}$ disappears ($\nabla\mathcal{F} = 0$). From the continuity of the normal fluid,

we also have $\nabla v_n = 0$. Taking the divergence on both sides of Equation 2.3, we have

$$D\nabla^2 u + v_n \nabla u = 0. \tag{2.4}$$

Provided that the tube has a uniform cross-section, and that the normal fluid flow is uniform, Equation 2.4 can be reduced to a one-dimensional form. Assuming that ^{3}He can only enter through the open end, we have boundary conditions, $u = du/dx = 0$ at $x = \inf$, and $u = u_0$ at $x = 0$. The solution of Equation 2.4 under these conditions is

$$u = u_0 e^{(-v_n x/D)}. \tag{2.5}$$

For a typical value of $D = 2 \times 10^{-3}$ cm^2/s at 1.4 K and $v_n = 1$ cm/s, we have $u(x = 1 \text{ mm})/u_0 = 3 \times 10^{-15}$. This means that the ^{3}He concentration can be reduced by 15 orders of magnitude in 1 mm distance inside the flush tube.

The heat flush technique works most efficiently for temperatures around 1.4 K [60]. As the temperature approaches the λ point (2.2 K), the normal fluid velocity v_n becomes very small for any reasonable heat inputs. Not only does this mean it takes a long time to reach the steady state, but the system will be susceptible to various instabilities such as oscillations from superfluid to normal fluid. At temperatures below 0.7 K, the density of the normal fluid drops so low that the assumption of ^{3}He atoms diffusing through a moving medium (the normal fluid) becomes inapplicable. At still lower temperatures (< 0.5 K), the mean free path of ^{3}He becomes few millimeters, comparable to the size of typical flush tubes. In this limit, the movement of ^{3}He atoms is virtually unaffected by the normal fluid, and the heat flush technique stops working completely.

The ultrapure helium production with the heat flush technique was first demonstrated by McClintock's group in 1976 [58]. The best direct measurement at that time used mass spectroscopy to place the ^{3}He isotopic ratio at $R_{34} < 4 \times 10^{-10}$. To improve this limit, the ultrapure helium is subjected to a secondary heat flush which concentrates any remaining ^{3}He into a small volume. Because ^{3}He is again not detected from helium in the concentrated sample, an indirect limit of $R_{34} < 2 \times 10^{-15}$ is inferred based on a measured concentration factor. Later, a continuous flow purifier was built by McClintock's group and an indirect limit of $R_{34} < 2 \times 10^{-13}$ was obtained [60]. The ultrapure helium used in our experiment was produced by the continuous flow apparatus during two production runs.

From the above discussion, we see that under suitable experimental conditions, it is possible to produce ultrapure helium with any desirable isotopic purity using the heat flush technique. The only drawback of the technique is that it is not easy to build a working purifier, so most groups requiring ultrapure helium simply use samples produced by McClintock's group. Furthermore, care must be taken to avoid any natural helium contamination during transportation and storage.

2.2.4 Summary

Among the three purification techniques, filtration through a superleak is the easiest to implement, but its sample purity is also the lowest. Differential distillation requires low temperatures (0.3 K to 0.6 K) to achieve high purity, so it is especially suitable for experiments that already operate at such temperatures. The heat flush technique can produce the purest samples, but the construction of the purifier is rather

complicated. Because the ^{3}He concentrations in all the ultrapure helium samples are below the sensitivity of traditional mass spectroscopy, more accurate methods of measuring trace ^{3}He concentrations, such as accelerator mass spectroscopy, need to be developed to measure the sample purity.

2.3 Accelerator Mass Spectroscopy at ATLAS

In traditional mass spectroscopy, sample atoms or molecules are first ionized inside an ion source. After a brief acceleration, the ions are sent through an analyzing magnetic field. Because an ion's deflection is proportional to its charge to momentum ratio (e/p), a position sensitive detector can be used to separate the ions based on their deflection. This is shown in Figure 2.2. There are two limitations to this technique, interference from stable isobars[2] and detector background. Firstly, because the mass of an isobar is virtually the same as the trace isotope, if both are in the same charge states, they can not be distinguished from each other by the position sensitive detector. Secondly, ions can scatter off residual gas in the chamber into wrong detector positions. This builds up into a constant background of the detector along with electronic noises. As a result of these backgrounds, the sensitivity of traditional mass spectroscopy is limited to 10^{-9}.

To overcome these two limitations, in accelerator mass spectroscopy (AMS), ions with the desired e/p ratios are accelerated to few MeV per nucleon after the low energy analyzing magnets. They then pass through a stripping foil where molecular

[2]Isobars are atoms and molecules having the atomic mass number as the trace isotope. For example, ^{14}N is a stable atomic isobar of ^{14}C, and H_3 is a molecular isobar of ^{3}He.

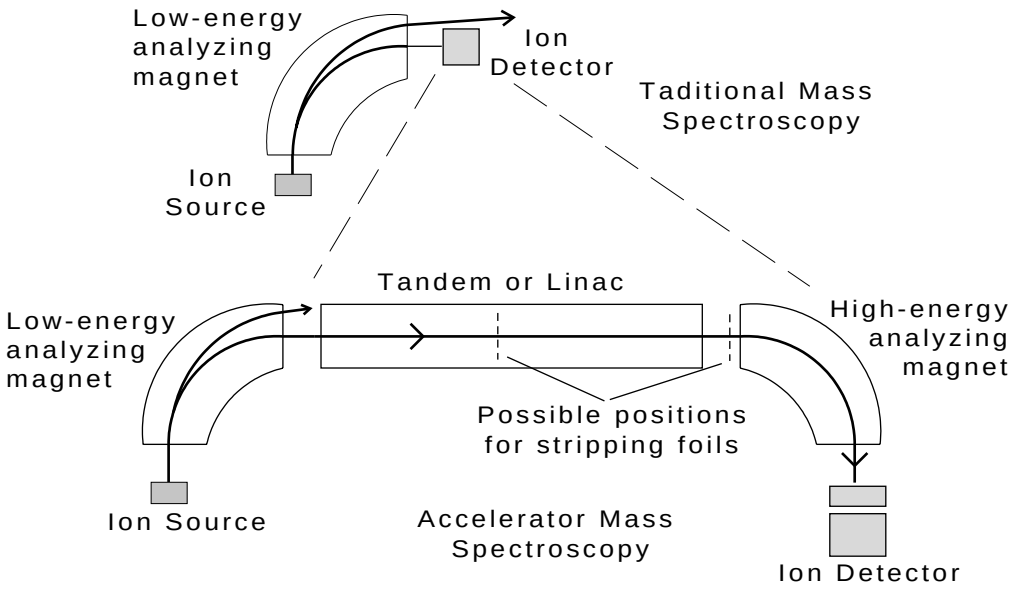

Figure 2.2: The comparison between the traditional and accelerator mass spectroscopies.

bonds are broken up and all electrons outside the nucleus are stripped away. Isobars can then be separated from the isotope of interest by their energies and stopping ranges in a particle detector. In addition, the detected ions are of such high energy that background noise from the particle detector can be totally suppressed using high thresholds. The ultimate sensitivity of AMS is usually limited by stable isobars that are not entirely separated and by total efficiency of the system which limits the observed count rate of the species of interest. However, in the case of the ^{3}He measurement, the sensitivity is mostly limited by ^{3}He contamination from natural helium in the system.

The typical AMS machine consists of three main components, ion source, accelerator and particle detector as shown in Figure 2.2. In the following sections, I will briefly discuss each component of the AMS machine at the Argonne Tandem-Linac

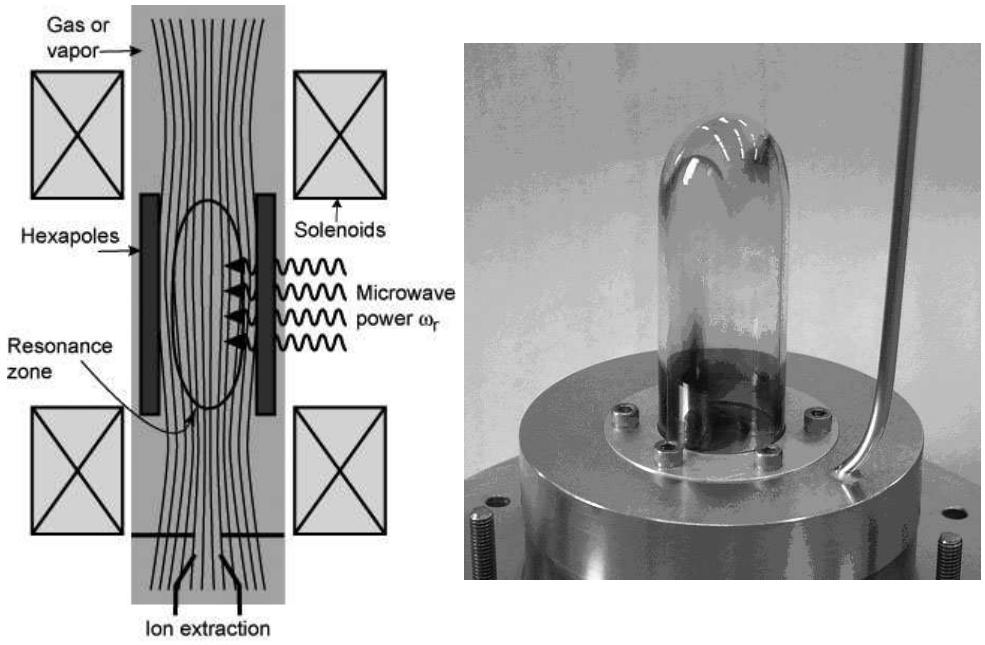

Figure 2.3: The left graph shows a schematic of an ECR source. The right graph shows the quartz tube used as the ECR source chamber during ^{3}He AMS runs.

Accelerator System (ATLAS) where the ^{3}He measurements were conducted.

2.3.1 Electron Cyclotron Resonance Source

Ion sources are categorized into negative and positive based on the charge of produced ions. In a typical negative ion source, cesium atoms are sputtered onto a sample pellet causing the release of sample atoms. The released atoms, such as ^{14}C, can pick up one electron from the cesium atoms which ionize very easily. The use of a negative ion source is preferred in some AMS measurements such as ^{12}C because isobars that cannot form stable negative ions are strongly suppressed. However, because helium atoms cannot form stable negative ions, a positive ion source based on electron cyclotron resonance (ECR) was used for the ^{3}He AMS measurement.

In an ECR source, a plasma of ions and free electrons is confined by an axial magnetic field from two solenoids and a radial magnetic field from a hexapole magnet. The strength of the field is zero at the center and increases gradually towards the walls of the source chamber. Constant field surfaces are concentric ellipsoids. Free electrons are resonantly accelerated when they cross the resonant surface where the B field is such that the electron cyclotron frequency (eB/m_e) equals the input microwave frequency ω_r. The originally neutral sample gas is ionized through collisions with energetic electrons. While freed electrons are accelerated, positive ions are extracted by extraction electrodes. The principle of ECR operation is shown schematically in Figure 2.3.

The ECR source at ATLAS injects 10 GHz microwaves, corresponding to a resonant surface of 0.36 T. Because the initial source chamber had too much natural helium contamination, a new source chamber made of cleaner material, quartz, and with a smaller volume for high pressure operation was designed and built in 2003 for the exclusive purpose of ^{3}He measurements. It indeed reduced the natural helium background from the ion source dramatically.

2.3.2 ATLAS Accelerator

At the ATLAS facility, positive ions are first accelerated by high voltage (up to 350 kV) on the ECR source mounting platform. After passing several low energy analyzing magnets, the ions go through a three-stage bunching system, which compresses and chops the DC ion current into 250 ps wide pulses. These short pulses can then be accelerated by superconducting radio frequency resonators.

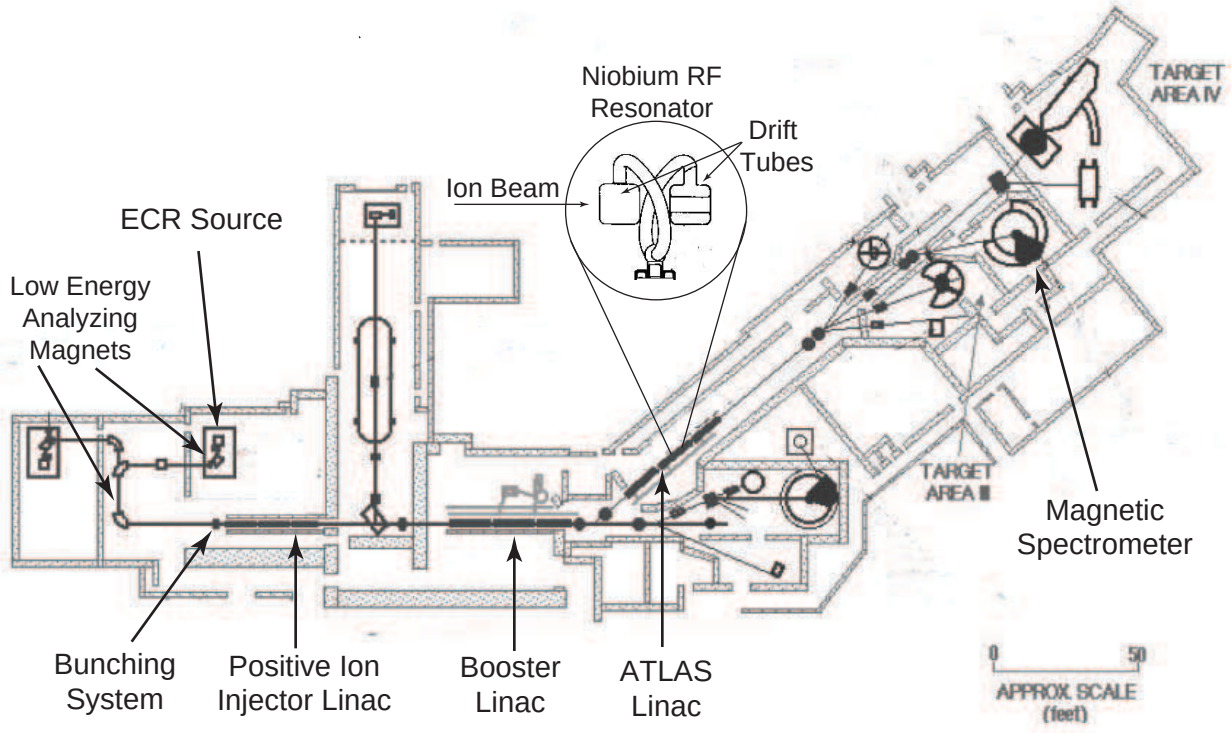

Figure 2.4: The layout of the ATLAS accelerator, and a schematic of a typical split ring superconducting resonator [70].

Compared to normal copper resonators, niobium superconducting resonators can reduce the overall power consumption at an accelerator facility by a factor of two to four and reduce the acceleration distance due to much higher field gradients. A typical split ring niobium resonator used at ATLAS is shown in Figure 2.4. A large number of electrons flow back and forth between two drift tubes through the connecting inductive ring at a frequency close to 100 MHz, creating oscillating electric fields. The phase of the resonator can be adjusted so that ion pulses always receive a boost in energy from the electric field when they pass the resonator. The resonators at ATLAS are divided into three sections (see Figure 2.4). The positive ion injector (PII) linac consists of 18

quarter wave resonators providing up to 12 MV of acceleration voltage. The booster linac and the ATLAS linac both provide up to 20 MV of acceleration voltage with 24 and 18 resonators respectively.

After acceleration, high energy ions are delivered into one of four target areas. For AMS measurements, the ions are steered into the magnetic spectrometer in target area III for further analysis.

2.3.3 Magnetic Spectrometer

The magnetic spectrometer and associated detectors at ATLAS shown in Figure 2.5 is capable of separating ions of different charge to momentum ratio (q/p), and measuring the total energy, the energy loss curve, and the time-of-flight of each ion.

In the magnetic spectrometer, high energy ions first pass through a stripping foil which breaks up molecular bonds and strips off the remaining electrons from the nuclei. The ions are then focused onto a focal plane detector based on their q/p ratios by a split-pole magnet. The split-pole design in which the poles of the magnet are divided into two sectors can provide a second-order double focusing of the ions [71], therefore increasing the resolution of the spectrometer on the focal plane.

The focal plane detector consists of two gas detectors. The first detector is filled with 5 torr of isobutane gas. Though an ion loses very little energy passing through the first chamber, the high voltage in the chamber causes avalanche ionizations to occur along its path. The avalanche electrons are quickly collected by a wire grid in the chamber. Based on the delay times between the wires, precise focal plane position and timing information of the incoming ion can be determined. The second

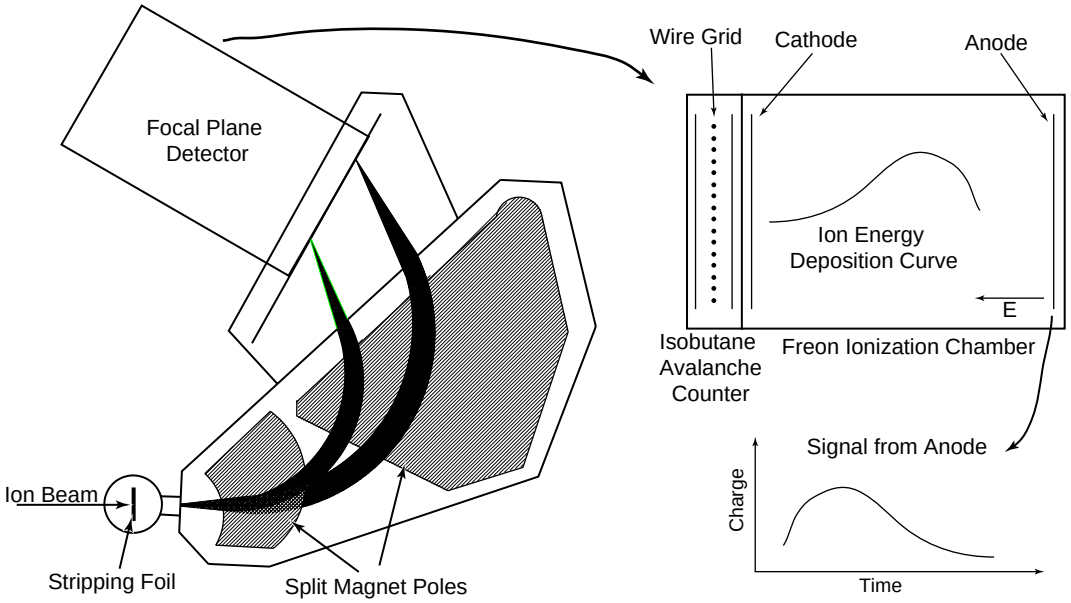

Figure 2.5: The schematic of the magnetic spectrometer at ATLAS. The left graph shows the stripping foil at the entrance and the split pole magnet which focuses the ions onto the focal plane detector. The position and timing information of an ion is detected from the isobutane avalanche counter. The energy deposition curve of the ion can be measured by the freon ionization chamber. The ionization electrons drift towards the anode. The signal on the anode is proportional to the inverse of the energy deposition curve.

chamber is filled with freon gas. It has much larger volume and higher gas pressure than the entrance chamber. An incoming ion is completely stopped in this volume. The energy of the ion is turned into electron-ion pairs along its path. An electric field in the chamber causes the ionization electrons to drift towards the anode. The time dependent signal at the anode gives information about the energy loss curve of the ion, see Figure 2.5.

The total energy, the range, and differential energy deposition of the ion are easily inferred from the curve. This information is very useful in separating the isotope of interest from interfering ions. For example, on the total energy versus focal plane

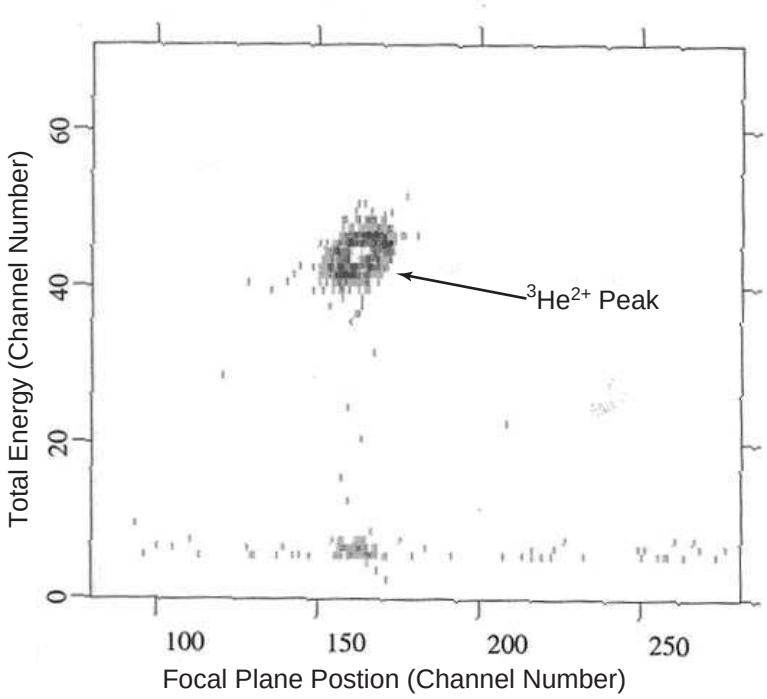

Figure 2.6: On the two dimensional plot of ion total energy versus focal plane position, the ^{3}He peak can be clearly identified, and well separated from background and other interfering isotopes.

position graph, the $^3He^{2+}$ peak is clearly identified and separated from other isotopes, see Figure 2.6.

2.4 Measurement and Analysis

In this section, the ^{3}He sample preparation and the data from four ^{3}He measurement runs at ATLAS will be discussed. We thank Dr. Pardo and his colleagues at Argonne National Lab for developing the ^{3}He AMS technique for us, and assisting us throughout the measurements. Without their help, this measurement would not have been possible.

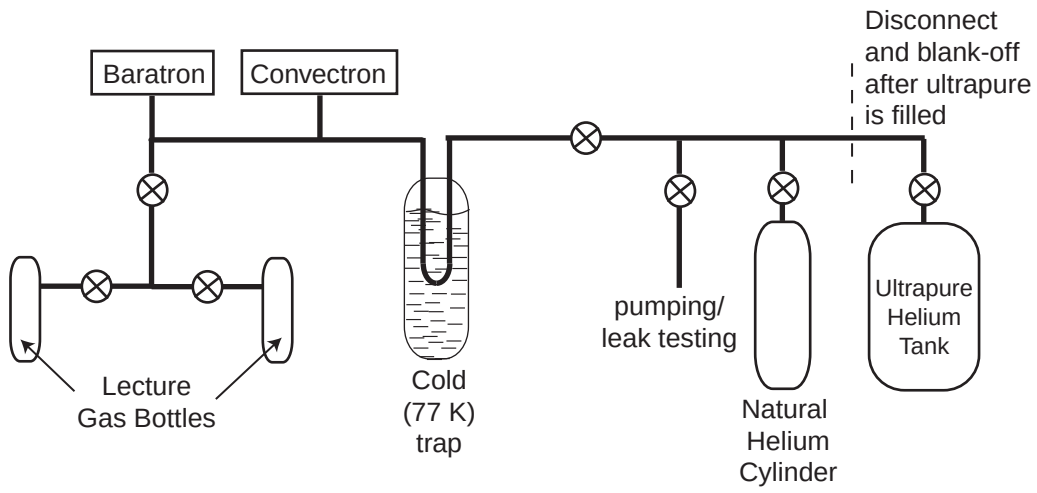

Figure 2.7: The gas handling system for preparing helium gas with various ^{3}He concentrations.

2.4.1 Sample Preparation

In addition to the ultrapure helium sample, we prepared samples with various concentrations of ^{3}He (10^{-10}, 10^{-11}, 10^{-12} and 10^{-13}) for the purpose of calibration and consistency checks.

The samples were produced by mixing a small amount of helium with high ^{3}He concentration with ultrapure helium which we assumed to have a ^{3}He concentration less than 10^{-15}. Lecture gas bottles were used to store the helium samples. Because ^{3}He concentration in air is only 6×10^{-12} (5.2 ppm of natural helium), simply pumping the pressure in the lecture bottles to below 0.1 torr[3] could reduce the ^{3}He contribution from air to less than 10^{-15} once the bottle was filled with 1 atmosphere of ultrapure helium. However, we did not know if the bottles were ever contaminated by natural helium. Therefore, each bottle was pumped down to below 10^{-4} torr, and then flushed

[3] 1 torr equals 1 mm of mercury. 1 atmosphere equals 760 torr.

with ultrapure helium one to two times. In this way the ^{3}He contamination from the bottle would be less than 10^{-15} even if we had natural helium in the bottle at the beginning, assuming ^{3}He absorption on the bottle wall is negligible.

The gas handling system for filling and mixing helium gas is shown in Figure 2.7. The baratron gauge can measure helium pressure up to 4 atmosphere, but it is not sensitive to gas pressures below 5 torr. The convectron gauge, though completely useless when the helium pressure is above 2 torr, gives accurate reading of pressure below 1 torr, down to 10^{-3} torr.

At first, four lecture bottles were filled with ultrapure helium to three atmospheres. Then, one ultrapure helium bottle was connected with a bottle filled with 1 torr of helium from a commercial helium cylinder which was assumed to have a ^{3}He concentration of 2×10^{-7}. After the gas in the two bottles completely mixed with each other, two bottles with a ^{3}He concentration of 10^{-10} are produced. Then one of the 10^{-10} sample bottles is pumped down to 200 torr, and mixed with a ultrapure helium bottle to produce the 10^{-11} sample. The 10^{-12} and 10^{-13} samples were prepared following the same procedures. It should be noted that it took about 10 minutes for the gas to completely mix, as determined by the calculated diffusion time of gas molecules. Of course, the prepared sample concentration depends on the assumptions of ^{3}He concentrations in ultrapure and natural helium. If these assumptions turn out to be incorrect, the sample concentrations would need to be modified.

2.4.2 ^{3}He AMS Measurements

Run I: (Initial test)

The initial test run I for ^{3}He AMS measurements took place in January 2001 [72]. Because using $^3He^+$ to tune the beam directly would result in contamination of the ECR ion source, a $^{12}C^{4+}$ guide beam which also has charge to mass ratio of 1/3 was used to tune the transmission of the accelerator. The tune was then scaled up by 0.5% to account for the binding energy difference between carbon and helium. The initial transmission was 8.5% inferred from the $^{12}C^{4+}$ transmission. At the end of the run, the transmission was checked directly with a $^3He^+$ beam. It had decreased to 7%. A 1 mg/cm steel absorber foil was used at the entrance of the magnetic spectrometer to stop carbon ions, but turned out to be unnecessary, because the tune shift was large enough so that the transmission of carbon ion was close to zero. The peak of fully stripped $^3He^{2+}$ ions was unambiguously identified by the focal plane detector, see Figure 2.6 for data from a sample run.

Data from this run were plagued by natural helium background. The partial pressure of helium from the ion source was measured to be 5×10^{-10} torr by a residual gas analyzer, translating to a ^{3}He partial pressure of 5×10^{-17} torr. Under normal source operation pressure (10^{-6} torr), the best limit for ^{3}He/^{4}He ratio would only be 5×10^{-11}. To reduce the background, the source was operated at a pressure as high as 10^{-1} torr. However the source operation at these high pressures was not stable. In addition, the extracted ion current was reduced by a factor of 100 at high pressures, due to the short mean free path of ions at high pressures. Figure 2.8 shows the source pressure dependence of ^{3}He/^{4}He ratio for a 10^{-14} sample. First, even at

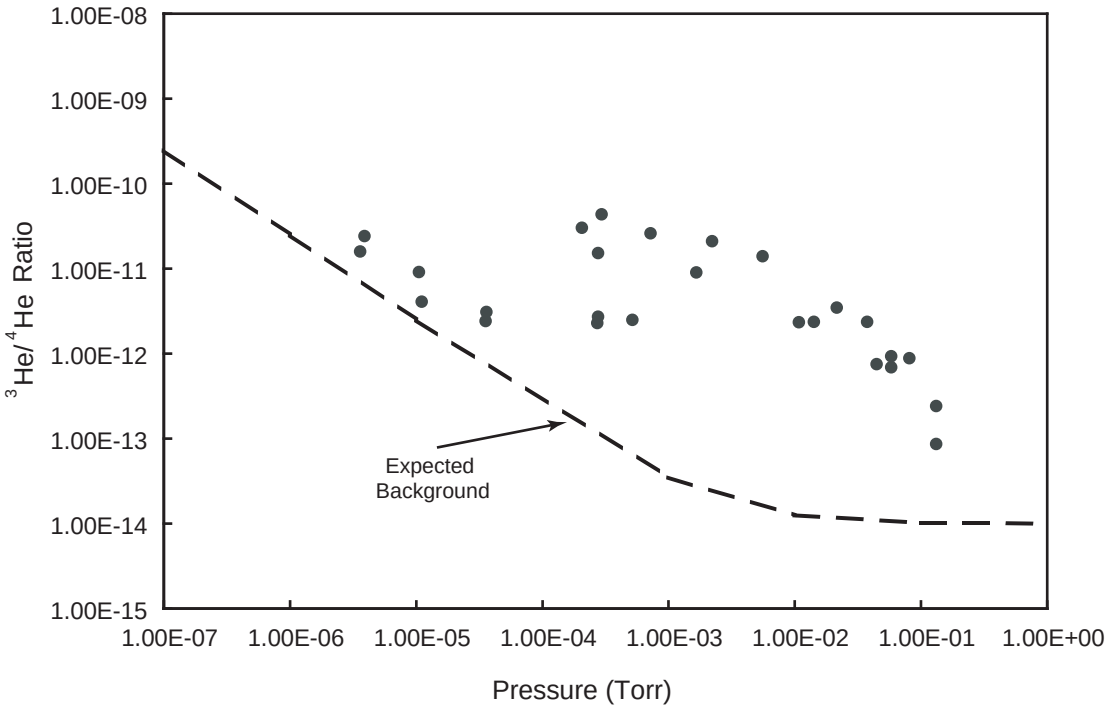

Figure 2.8: The source pressure dependence of $^3He/^4He$ ratio for AMS run in January 2000. The dashed line indicates the expected ^3He background from the ion source. The same 10^{-14} sample was used for all runs. The statistical error bar, not shown in the graph, is less than the size of the data points.

high pressures where the backgrounds were low there were significant amount of ^3He on the order of 10^{-12}. Secondly, the measured ratios of the same sample differ from each other by up to three orders of magnitude. Because the measurements were not statistically limited, the large discrepancy was attributed to systematic errors, such as large background, unstable source operation and accelerator transmission shifts.

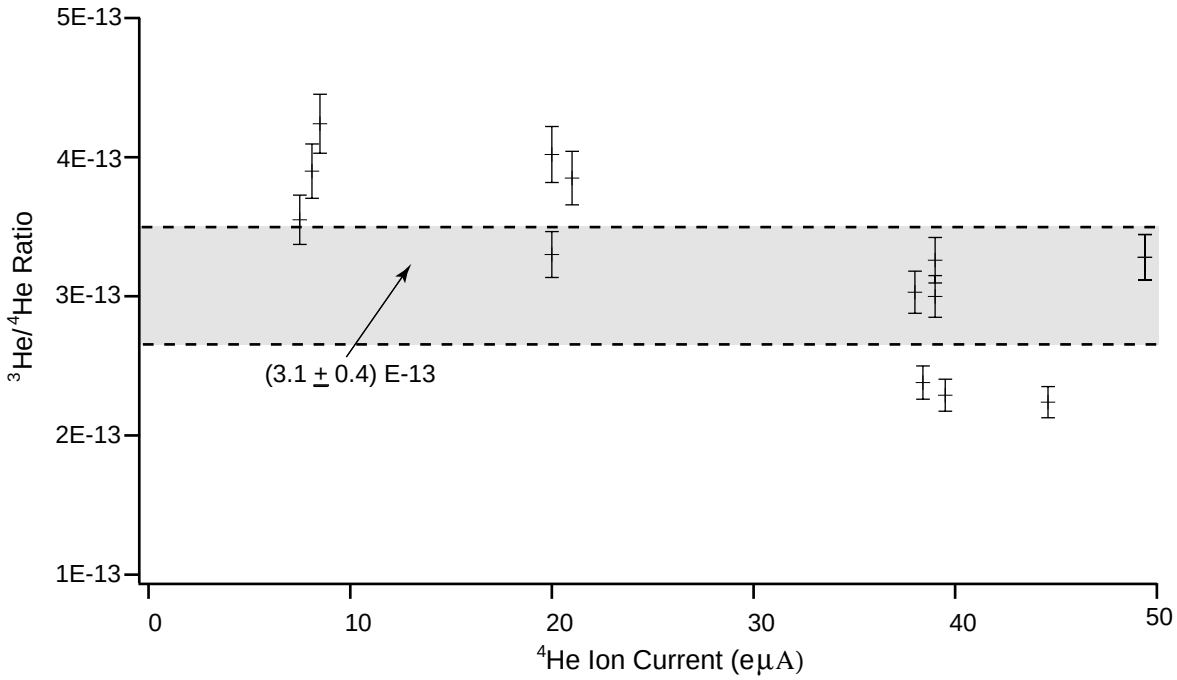

Figure 2.9: The measured $^3He/^4He$ ratio of a ultrapure helium sample at various source $^4He^+$ output current during April 2002 run. The ^3He background from the source is corrected. The shaded area is the average of all data points.

Run II: (New ECR source)

To address the background problems encountered in run I, a new ion source was developed. Because it was hypothesized that most of background came from the source chamber walls, a small-volume clean quartz tube, never exposed to natural helium, was used as the new source chamber. The hole on the extraction electrodes was reduced to 1 mm diameter to allow the source to operate at high pressures. The magnetic field configuration was also modified to move the resonance region closer to the extraction electrode, so that source could operate more reliably at high pressures [73].

The AMS run II took place in April 2002. The new ECR source was installed and

tested. The pressure inside the source could not be directly measured, but could be inferred from the extraction region pressure. The source operated stably at both low and high pressures. For example, at 4×10^{-2} torr, the source could output 100 eμA of ion current, a factor of 30 improvement over the old source. It was also discovered that H_3^+ molecular ions produced in the source could be used as a guide beam for tuning the accelerator. Its charge to mass ratio is closer to $^3He^+$ than $^{12}C^{4+}$, and it can be easily eliminated by a thin mylar foil. The initial transmission for $^3He^+$ was inferred from the H_3^{1+} beam transmission measurement to be 14%.

Again only ultrapure helium was allowed into the system at the beginning of the run to avoid ^{3}He contamination of the system. At high source $^4He^+$ output current (20 - 100 eμA), the ^{3}He count rate roughly scales with the source current as expected. But at very low source $^4He^+$ output current(< 0.2 eμA), the ^{3}He count rate stayed constant, around 5 counts/second. This could be interpreted as a constant background from the source. Correcting this constant background from the data, we obtained a set of consistent measurements. Figure 2.9 shows the $^3He/^4He$ ratio at various source $^4He^+$ output current. Averaging all data points, we get a ^{3}He concentration $(3.1 \pm 0.4) \times 10^{-13}$.

Unfortunately, a series problems occurred after the ultrapure sample measurements. First was the sparking in the magnetic spectrometer. Then a large leak from the ECR source stopped its operation. It was determined that the O-ring seal of the source chamber at the RF power input side was leaking due to too much heating. Stable source operation was never recovered in this run. Some accelerator issues also prevented a successful measurement of the transmission with $^3He^+$ at the end of run.

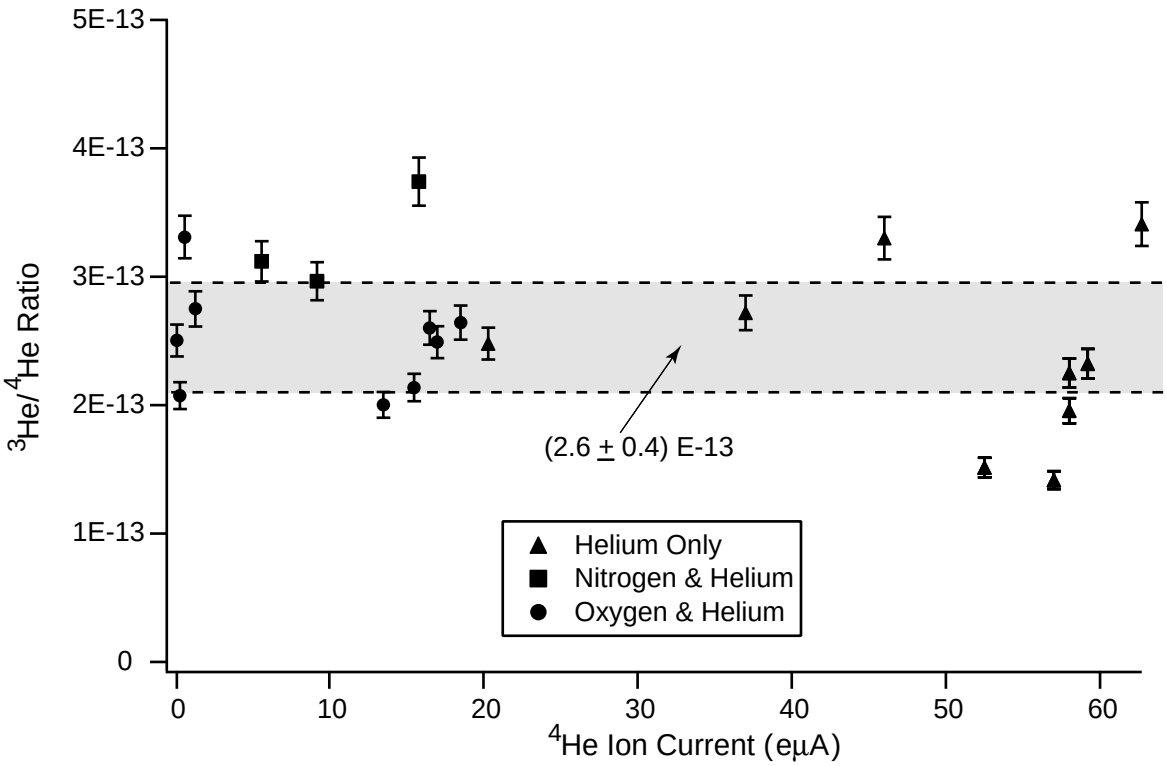

Figure 2.10: The background corrected $^3He/^4He$ ratio of the ultrapure helium sample at various $^4He^+$ current during January 2003 run. The data include helium only runs, helium plus nitrogen buffer gas runs and helium plus oxygen buffer gas runs. The shaded area is the average of all data points.

Run III: (Buffer gas operation)

Between the second and the third run, the ECR source was modified again. First the o-ring seal close to the RF power input port was eliminated to reduce the probability of an leak. Secondly, the gas handling system was rebuilt to allow the source to run with helium and a buffer gas such as nitrogen or oxygen simultaneously. Running with the buffer gas alone allowed us to study background ^3He from the source directly. A picture of the source quartz tube is shown in Figure 2.3.

The AMS run III took place in January 2003 [74]. As in run II, the accelerator

was first tuned with a H_3^+ beam. The $^3He^+$ transmission was inferred to be 24%. Initially, only ultrapure helium was introduced into the system, and the source pressure dependence of $^3He/^4He$ ratio was studied[4]. As expected, the ratio asymptotically approached a constant (3.9×10^{-13} without background subtraction). Then the source was run with pure nitrogen gas to study the background. At various nitrogen pressures, the ^3He count rate in the detector stayed constant, around 5 counts/second. An overnight ten hour run showed that the background was also constant with time. This seemed to be the same background seen in run II at very low source $^4He^+$ output current. Because here no external helium was introduced, it was a stronger evidence that we were directly measuring background ^3He coming from the source.

In the next run sequence, ultrapure helium was introduced along with the nitrogen buffer gas. The advantage of using a buffer gas was that it allowed source $^4He^+$ output current to vary while the total source gas pressure was kept more or less constant. This technique helps to reduce possible systematics caused by plasma running at different pressures. After nitrogen, oxygen was also used as a buffer gas. Initially, running oxygen as the buffer gas showed even less background than nitrogen. But at some point, the background suddenly increased dramatically and it took a long time for the system to recover.

Figure 2.10 shows the source $^4He^+$ output current dependence of the background corrected $^3He/^4He$ ratio for helium only runs, helium plus nitrogen buffer runs and helium plus oxygen buffer runs before the sudden increase in background. First, the measured ratios are independent of source $^4He^+$ output current and the type of buffer

[4]We could not measure the source pressure directly, the pressure was inferred from the pressure at the extraction region.

gas used. Secondly, the weighted average of all data ($R_{34} = (2.6 \pm 0.4) \times 10^{-13}$) is consistent with the result from run II. These are all good self-consistency checks. However, the scatter of data points is still more than their statistical uncertainties, indicating the existence of fluctuations in background or systematic uncertainties that were not accounted for.

At the end of the run, the transmission of the accelerator was again checked with a $^3He^+$ beam directly. The source was run with natural helium sample which had a known ^{3}He abundance of 3×10^{-7}. The $^3He^+$ current was reduced by a series of attenuators so that it would not overload the magnetic spectrometer. Although some accelerator problems prevented us from making a precise measurement, preliminary results showed that the transmission was about a factor of ten lower than expected. The problem with transmission measurements was confirmed in run IV.

Run IV: (Major issues discovered)

The fourth AMS run took place in August 2004 [75]. A boron nitride tube was initially used as the ECR source chamber instead of the quartz tube with the hope to reduce the ^{3}He background further[5]. However, it was quickly discovered that the background ^{3}He counts were a factor of 100 higher than before. Even after the quartz tube was reinstalled, the background ^{3}He count rate was 20 Hz, roughly four times higher than before. The background dropped to around 5 Hz after 18 - 24 hours of operation.

The ^{3}He concentrations of various samples were measured, including ultrapure

[5]Unlike quartz, cubic type boron nitride is impermeable to helium. If the background ^{3}He comes from some external source diffusing into the ECR source, the use of BN can eliminate that background.

samples, a 10^{-13} sample, a 10^{-11} sample and two ultrapure samples produced by a purifier at Los Alomas National Lab. The surprising result was that all concentrations fell in the regime of a few $\times 10^{-12}$. The measurement of the same ultrapure helium sample was higher than in run II and III by a factor of ten. Why would the measurements of the same sample be so drastically different? During one test run, the magnetic field of the ECR source was changed to move the resonance zone closer to the extraction electrodes, and an increase in the background was observed. This suggested that ^{3}He were coming off the extraction electrodes. It was also conceivable that the rate of $^3He^+$ coming off the electrodes could be proportional to the increase in source $^4He^+$ output current (i.e. plasma density). This source-current dependent background would imitate a ^{3}He impurity in the sample. Because in run IV, the electrodes could have been contaminated by ^{3}He from the boron nitride tube, while in run II and III, the electrodes were free of this contamination, it is more likely that the measurements from run II and III were closer to the true value of ^{3}He concentration in the ultrapure helium sample.

During run IV, a major problem with the accelerator transmission scaling procedure was also revealed. The transmission of the accelerator for $^3He^+$ was initially inferred from a H_3^+ beam to be 19.8%, but direct measurement with a $^3He^+$ beam later showed the transmission was only 1.3%. It was unlikely that the transmission shifted during the run by such a large factor, so this low transmission was likely the result of scaling from the H_3^+ beam to the $^3He^+$ beam. This problem most likely existed in run II and III, but was not revealed due to troubles with direct transmission measurements using the $^3He^+$ beam. If run IV was representative of the issue,

measurements from both run II and III would need to be scaled up by a factor of fifteen.

2.4.3 Summary

Since 2000, significant progress has been made in developing the technique for AMS measurement of trace ^{3}He impurities: the $^3He^{2+}$ ion peak can be clearly separated from background ions; a new ECR source can operate at high pressure with much improved stability; and a new gas handling system enables us to run buffer gas in the source to measure background directly.

Despite the progress, several problems were discovered. First, the accelerator scaling from H_3^+ beam to $^3He^+$ beam seems to be seriously flawed. The real transmission was an order of magnitude lower than expected. Secondly, we do not understand the origin and the nature of the ^{3}He background from the source. Though sometimes the background ^{3}He rate stays constant, it can at times suddenly jump up significantly. More troubling is the possibility that the background ^{3}He rate scales up with the source output current, therefore imitating the signal from ^{3}He in the sample.

Combining the results of Run II and III, we obtain a ^{3}He concentration measurement of $(4.2 \pm 1.5) \times 10^{-12}$ for the ultrapure helium sample. Here a scaling factor of 15 was included to account for the low transmission observed in run IV. The run IV data was not included here, because of possible ^{3}He contamination. However, due to problems mentioned above, this measurement is not yet conclusive, but should be viewed as a strong indication of high ^{3}He levels in the ultrapure helium.

At the 10^{-12} level, ^{3}He impurity would introduce 5% systematic correction to

the measured neutron lifetime. It will certainly need to be reduced for a precision measurement of τ_n. There is no theoretical reason to question the effectiveness of the heat flush method now. The high ^{3}He level likely comes from contamination during production, transportation or storage. It might be necessary to build a heat flush system to purify the helium again before putting it in the UCN trapping apparatus.

It is also necessary to improve the ^{3}He concentration measurement to make sure that the high level of ^{3}He seen in the sample is real. To do the AMS measurement properly, one would first tuned the accelerator with $^3He^+$ beam, then replace the whole source assembly including the source chamber and the extraction electrodes to get rid of contaminations, then proceed with careful investigations of backgrounds in various source operation conditions. Unfortunately, the whole experiment could easily take up one month of beam time at ATLAS. Given the number of experiments waiting to use the facility, it would be difficult to obtain such a long operation time for the ^{3}He AMS measurement. In the near future, two possible solutions are discussed below for measurements on the $10^{-12} - 10^{-13}$ level.

2.5 Future Prospects

2.5.1 Simplified AMS

Dr. Pardo at Argonne proposed a simplified setup for a ^{3}He AMS measurement. The central idea is that for light ions such as ^{3}He, there is not much background from isobar ions. The largest background comes from H_3 molecular ions which can be easily disassociated by a stripping foil. Therefore, it may be possible to do clear

particle identification with lower acceleration and a simpler detector. The proposed experiment is to use only the positive ion injector for acceleration up to 12 MeV, see Figure 2.4. A carbon stripping foil right after the high voltage source platform can be use to disassociate the H_3 molecules. And $^3He^+$ ions can be detected by a silicon barrier detector behind the positive ion injector linac. With this setup, a ^{3}He AMS measurement can be done at the same time as other ATLAS experiments that require negative ions. Much more beam time will be available for study of the background issues.

Initial tests of the setup in early 2005 by Dr. Pardo showed promising results [75]. The mass 3 peak was clearly seen in the detector spectrum and no higher mass was observed. To check if H_3 was in the mass 3 peak, a scatter foil could be placed in front of the silicon detector. If H_3^+ molecules existed in the beam, they would be broken up into H_1^+ and H_2^+ and show up in the mass 1 and mass 2 peaks in the spectrum. Tests showed that in order to eliminate the H_3 background, the ion beam had to be well collimated to insure its passage through the center of the stripping foil. Although this reduces the detection efficiency significantly, a measurement precision on the order of 10^{-13} could still be achieved.

2.5.2 Laser spectroscopy and atomic trapping

A laser spectroscopic method of measuring helium isotope ratio was recently developed by Dr Lu's group at Argonne National Lab [76]. The isotope line shift between ^{3}He and ^{4}He was used to distinguished the two isotopes. Because of the difficulty of obtaining a laser for the ground state to first excited state transition ($1^1S_0 \rightarrow 2^1P_1$), he-

lium atoms were first excited to the triplet metastable state (2^3S_1) in a gas discharge tube, and were then further excited via the 2^3S - 2^3P transition for resonant absorption measurement. The absorption peaks from the two helium isotopes are 65 GHz apart, much larger than the doppler width of 2GHz and the natural line width. While the resonance absorption peak of ^{4}He could be easily measured with a single laser beam, the weak absorption signal of ^{3}He was measured using frequency-modulation saturation spectroscopy [77, 78]. Lu's group has achieved a 3σ detection limit of 4×10^{-9} for the ^{3}He/^{4}He ratio. As pointed out by Lu, this limit can be improved by several orders of magnitude by installing the gas discharge cell in a Fabye-Parot cavity where the effective absorption would be increased by as much as 10^5 [79].

The helium isotope ratio can also be measured with atomic trapping techniques. Metastable helium created by RF discharge can be first laser cooled in a Zeeman slower, and then trapped in a magneto-optical trap. Careful fluorescence spectroscopy can detect a single trapped ^{3}He* atom. Lu's group used the same technique to measure the charge radius difference between ^{6}He and ^{4}He [79]. In their apparatus, the maximum ^{4}He* loading rate in the trap is 10^9/s. Assuming the same trapping efficiency for ^{3}He*, one ^{3}He* can be trapped per 1000 seconds for a 10^{-12} sample.

Both laser spectroscopy and atomic trapping methods can achieve precision of $10^{-12} - 10^{-13}$ for the ^{3}He/^{4}He ratio measurement with available technology, though it would be difficult to push it beyond 10^{-14} level, which can only be reached by AMS. However, due to current background problems and limited beam time with ^{3}He AMS measurement, a table top apparatus capable of determining the ratio at the 10^{-13} level would be very useful.

Chapter 3

Study of Above Threshold Neutrons

One of the largest systematic errors in the neutron lifetime measurement comes from the delayed escape of above threshold neutrons. During the trap loading, neutrons with energies both above and below the trap threshold are produced. The majority of the above threshold neutrons are absorbed or upscattered by the trap wall materials before data collection begins, and therefore do not affect the measurement. However, a small fraction of the above threshold neutrons in the low-field seeking state can stay in the trap for hundreds of seconds due to a combination of two mechanisms: marginal trapping and material bottling. These neutrons decaying inside the trap add a non-exponential component to the measured decay curve, therefore introducing a systematic error.

The above threshold neutrons can be ejected from the trap by lowering the trap field temporarily. In Section 3.1, a two dimensional analytical model (restricting the

neutrons to their radial motion) for studying the field ramping technique is developed. We compare two Ioffe traps, the Mark II trap which was used during the recent data collection runs (See section 4.2), and the KEK trap which will be used in the next generation apparatus (see Section 6.2). It can be shown that the KEK trap is more effective at eliminating the above threshold neutrons due to its greater trap depth. Because a three dimensional model incorporating the axial motion of the neutron cannot be solved analytically, we will only discuss the special situation when the axial motion and radial motion of the neutron can be decoupled.

The numerical simulation of neutron trajectories is complicated by the occurrence of chaotic scattering. The trajectories of above threshold neutrons with escape times longer than 5 – 10 s, in general, cannot be precisely tracked [80]. However, one can still use the simulation to study the coupling between the axial and radial motion of the neutrons. Furthermore, the simulation can be used to estimate the fraction of above threshold neutrons that escape the trap within 10 s after the trap field is ramped down. From the simulation, we can find the conditions under which total elimination of the above thresholds neutrons can be achieved. The simulation method and results are described in Section 3.2.

3.1 An Analytic Model

3.1.1 Motion of Neutrons in Ioffe Traps

An Ioffe trap consists of two solenoids for axial confinement and one quadrupole magnet for radial confinement. The trap field is not axially symmetric due to the

fringing fields from the solenoids. For the sake of simplicity, we ignore the radial field components of the solenoids in the analytic model, and assume that the field from the two solenoids does not vary in the radial direction and the quadrupole magnet does not contribute to the axial field. Then the field in the axial direction can be expressed as, $B_z = B_0 + B_1 f(z)$, where B_0 is the minimum field at the trap center, and B_1 is the radial trap depth. $f(z)$ is a function flat at the center with large field gradients at the trap ends[1]. The radial fields produced by the quadrupole magnet has a simple form $B_\rho = \sqrt{B_x^2 + B_y^2} = k\rho$. Parameters of the two Ioffe traps (the Mark II trap and the KEK trap) under study are listed in Table 3.1.

In this study, we are not concerned about neutrons in the high-field-seeking state, because they are ejected from the trap right away. We focus our attention on neutrons in the low-field-seeking state with energies higher than the trap depth, but that still stay in the trap for extended periods. The motion of such neutrons can be characterized by three frequencies, spin precession frequency ω_s, radial oscillation frequency ω_r and axial oscillation frequency ω_a. Typical parameters of a neutron with energy just above the trap threshold are shown in Table 3.1.

Firstly, because there is no zero field region in an Ioffe trap, the spin precession frequency is much higher than the other frequencies, i.e. $\omega_s \gg \omega_r, \omega_a$, everywhere in the trap. Under this adiabatic condition, the spin of the neutron always aligns with the magnetic field and the trajectory of the neutron is governed by an equation of motion,

$$m_n \frac{d\vec{v}}{dt} = -\mu_n \nabla |B|, \tag{3.1}$$

[1]An approximate function form of $f(z)$ is given in Ref [81], $f(z) = (g(z+L) + g(z-L) - 2g(L))/(g(2L) + g(0) - 2g(L))$, where $g(z) = [1 + (z/R)^2]^{-3/2}$ is the axial field of a coil.

Table 3.1: Parameters of the Mark II and the KEK Ioffe type traps and characteristic motion frequencies of a neutron with energy above the trap threshold.

Parameter	Mark II Trap	KEK Trap
Trap radius (cm)	4.3	5.5
Trap length (cm)	42	75
Trap depth (B_1) (T)	1.1	3.1
Minimum field (B_0) (T)	0.1	0.6
Radial trap gradient k (T/cm)	0.28	0.7
neutron velocity (cm/s)	~ 350	~ 600
Spin precession ω_s (MHz)	> 18	> 110
Radial oscillation ω_r (Hz)	~ 12	~ 17
Axial oscillation ω_a (Hz)	~ 2	~ 2

where m_n is the mass of the neutron and μ_n is the magnetic moment of the neutron.

Secondly, because the radial oscillation frequency is much higher than the axial oscillation frequency, i.e. $\omega_r \gg \omega_a$, an adiabatic approximation can used to describe the motion of the neutron in the center of the trap where ω_r changes slowly with time, the amplitude and frequency of the neutron radial oscillation depends only on axial coordinate z, and in the axial direction, the neutron travels through an effective potential $U(z)$ which is the sum of the trap potential and the radial kinetic energy averaged over radial oscillations [81]. In this region, the radial motion and axial motion are separated from each other and no energy mixing occurs. Near the ends of the trap, due to the large gradient of $B(z)$, the adiabatic condition that radial oscillation frequency changes slowly with time ($|\dot{\omega}_r| \ll \omega_r^2$) breaks down. In this region, though trajectories cannot be solved analytically, it can be shown numerically

that strong mixing between the axial and radial motion exists for traps where $B_1 > B_0$ [81], such as the neutron traps.

Because the exact analytical solution of the neutron motion in an Ioffe trap cannot be obtained, we will first neglect the axial motion, and study the neutron motion in a two dimensional linear trap, representing the radial trapping field from the quadrupole magnet. We then add in the axial motion, and generalize some results obtained in the radial trap. Traps with more realistic field configurations will be studied with numerical simulations.

3.1.2 Adiabatic Invariants during Field Ramp

To simplify the problem, we first ignore the motion of the neutron in the axial direction and study the radial motion of the neutron confined by a linear potential produced by the quadrupole magnet,

$$V(r) = \beta r \tag{3.2}$$

where $\beta = \mu_n k$. Since the potential is centrally symmetric, the angular momentum L is a constant of motion. The hamiltonian of the neutron in polar coordinates can be written as

$$H = \frac{{p_r}^2}{2m} + \beta r + \frac{L^2}{2mr^2}. \tag{3.3}$$

For a given total energy E, the motion of the neutron is linear at minimum angular momentum, $L = 0$, and circular at maximum angular momentum $L = m^{1/2}(2E/3)^{3/2}/\beta$. For intermediate L, the neutron trajectory is an elliptical curve that precesses around the center. Figure 3.1 shows the neutron trajectories at three different angular momenta.

When the trap field is ramped down, the neutron energy decreases and its trajectory expands. In the adiabatic approximation (field ramps slowly, i.e. $Td\beta/dt \ll \beta$, where T is the neutron oscillation period), the action variable $I = \oint p_r dr$ is an invariant [82],

$$I = 2\int_{r_{min}}^{r_{max}} \sqrt{2m(E-\beta r) - \frac{L^2}{r^2}} dr = const. \tag{3.4}$$

where, r_{min} and r_{max} are the boundary points of the radial oscillation. The total energy can be expressed as,

$$E = \beta r_{max} + \frac{L^2}{2m{r_{max}}^2}. \tag{3.5}$$

Equation 3.4 can be integrated explicitly for some neutron trajectories. For example, when $L = 0$,

$$I = 2\int_0^{r_{max}} \sqrt{2m\beta(r_{max}-r)} dr = \frac{4}{3}\sqrt{2m\beta} r_{max}^{3/2} = const., \tag{3.6}$$

we obtain that $r_{max}\beta^{1/3}$ is a constant. Using Equation 3.5, we have $E\beta^{-2/3} = const.$.

For the neutron on a circular orbit with maximum L, integral 3.4 vanishes. We notice, from the equation of motion, $mv^2/r = \beta$, which leads to $L^2 = m\beta r^3$. Again, we have $r_{max}\beta^{1/3} = const.$. Ref. [83] also discussed situations where L is close to the maximum. The radial motion of the neutron can be approximated with harmonic oscillators, and the same results are reached.

For intermediate value of L, action integral 3.4 is difficult to calculate explicitly. But we notice that if the initial trajectory is expanded to a new trajectory by a transformation $r \rightarrow (\frac{\beta_i}{\beta_f})^{1/3} r$, where β_i and β_f are the field gradients before and after the trap ramp, then the transformed orbit satisfies the action integral invariant

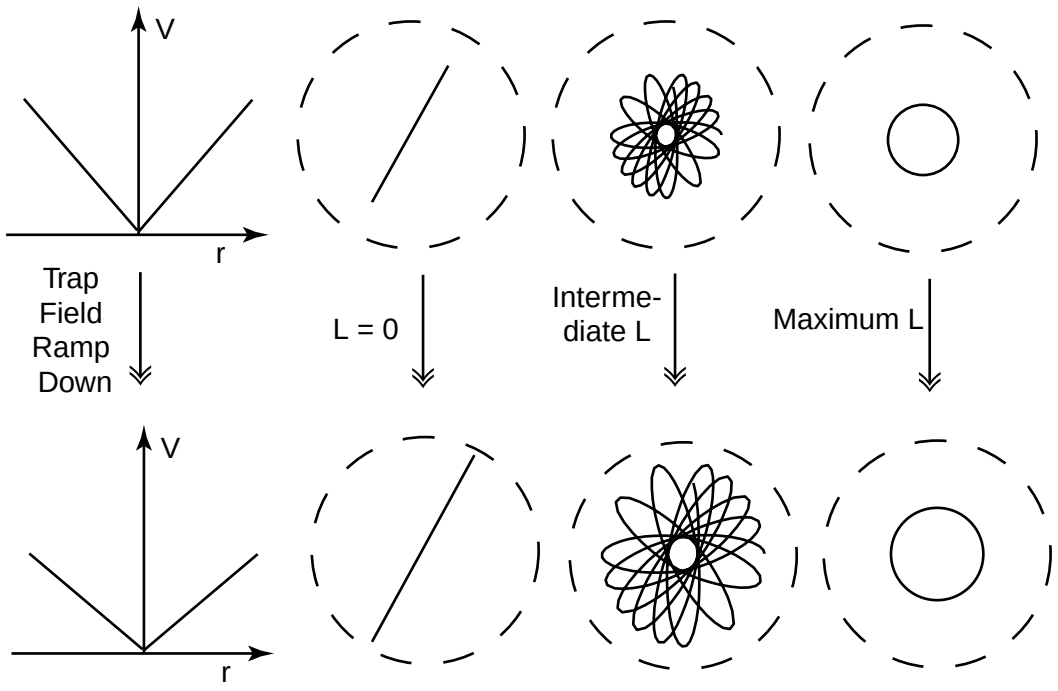

Figure 3.1: The trajectory of a neutron inside a two dimensional linear potential at various angular momenta for a given total energy. As the potential ramps down, the trajectories adiabatically expands. The new trajectory can be obtained by transformation $r \rightarrow (\beta_i/\beta_f)^{1/3} r$.

condition,

$$\begin{aligned} I_f &= 2\int_{(\frac{\beta_i}{\beta_f})^{1/3} r_{min}}^{(\frac{\beta_i}{\beta_f})^{1/3} r_{max}} \sqrt{2m(E_f - \beta_f r') - \frac{L^2}{r'^2}}\, dr' \\ &= 2\int_{r_{min}}^{r_{max}} \sqrt{2m(E_i(\frac{\beta_f}{\beta_i})^{2/3} - \beta_f(\frac{\beta_i}{\beta_f})^{1/3} r) - \frac{L^2}{r^2}(\frac{\beta_f}{\beta_i})^{2/3}}\ \ (\frac{\beta_i}{\beta_f})^{1/3} dr \\ &= I_i \qquad (r' = (\frac{\beta_i}{\beta_f})^{1/3} r) \end{aligned}$$

It can be shown that a neutron on the transformed orbit also satisfies the new Hamiltonian, therefore the transformed orbit is the solution of the neutron trajectory after adiabatic expansion, see Figure 3.1.

In summary, for a two dimensional linear trap, $r_{max}\beta^{1/3}$ and $E\beta^{-2/3}$ are two

adiabatic invariants of the system during field ramp[2].

3.1.3 Eliminating Marginally Trapped Orbits

Marginally trapped neutrons are those neutrons with energies above the trap threshold, but stay on stationary trajectories in the trap for extended periods. When the trap field is ramped down, the radii of the neutron trajectories adiabatically expand and eventually hit the trap wall, see Figure 3.1.

In the two dimensional case, the marginally trapped neutrons with the smallest orbit radii are the ones on circular orbits with energy just above the trap depth. It is easy to show that, at full field, $r = \frac{2}{3}R$, where R is the trap radius. This implies that in order to get rid of these neutrons (i.e. making their orbit radii larger than R), field must be dropped to,

$$(\frac{\beta_f}{\beta_i})^{1/3} = \frac{2}{3} \Rightarrow \frac{\beta_f}{\beta_i} = (\frac{2}{3})^3 = 0.296 \tag{3.7}$$

where β_i and β_f are the initial and final trap gradients. The final energy of the neutron after the field ramp (E_f) becomes

$$E_f = (\frac{\beta_f}{\beta_i})^{2/3} E_i = (\frac{2}{3})^2 E_i = 0.444 E_i. \tag{3.8}$$

The energy difference between this neutron and the final trap threshold is,

$$\triangle E = E_f - \beta_f R = 0.444 E_i - 0.296 \beta_i R = 0.148 \beta_i R \tag{3.9}$$

Thus, according to the analytic model developed here, when the trap field is ramped down to 29.6% of its original value, the trajectories of all neutrons with

[2]When r_{max} is greater than the trap radius, the neutron will reflect from the wall. Though the invariants no longer hold, the value of the energy E obtained from the invariant $E\beta^{-2/3}$ sets the lower bound for the neutron energy, because the neutrons that reflect from the wall experience less adiabatic cooling.

energy above the trap threshold would eventually intersect with the wall. Although the total energy of marginally trapped neutrons would decrease, their kinetic energies at the trap wall would increase to at least 14.8% of the initial trap depth. For example, in the Mark II trap, the marginally trapped neutrons would have at least 9.8 neV of kinetic energy at the trap wall after the field is ramped down to 29.6% of the maximum trap field; while in the KEK trap which has a 3.1 T trap depth, the kinetic energy will be at least 27.7 neV.

3.1.4 Eliminating Materially Bottled Neutrons

When a neutron hits the trap wall, it has a certain probability of being reflected back. The reflection probability depends on the perpendicular velocity component of the neutron and the wall potential. This material bottling effect keeps above threshold neutrons inside the trap even after they escape stationary trajectories.

The trap wall is coated with an organic fluor TPB for down-conversion of helium scintillation light. The coating can be modeled by a complex potential, $U = V - iW$ [36]. The real part of the potential V is the Fermi potential of wall material. Since the experimental cell is filled with superfluid helium, the effective potential the neutron experiences is the potential difference between TPB and liquid helium, $V = V_{TPB} - V_{He} \approx 42\text{neV} - 20\text{neV} = 22\text{neV}$. The imaginary part of the potential W describes neutron loss due to absorption and inelastic scattering. For pure TPB, W can be estimated based on its hydrogen content, $W \approx 10^{-3}\text{neV}$.

The reflection probability p_r of a neutron from the wall potential is given by [36]:

$$p_r = \frac{E_n - \sqrt{2E_n(\alpha - (V - E_n))} + \alpha}{E_n + \sqrt{2E_n(\alpha - (V - E_n))} + \alpha}, \tag{3.10}$$

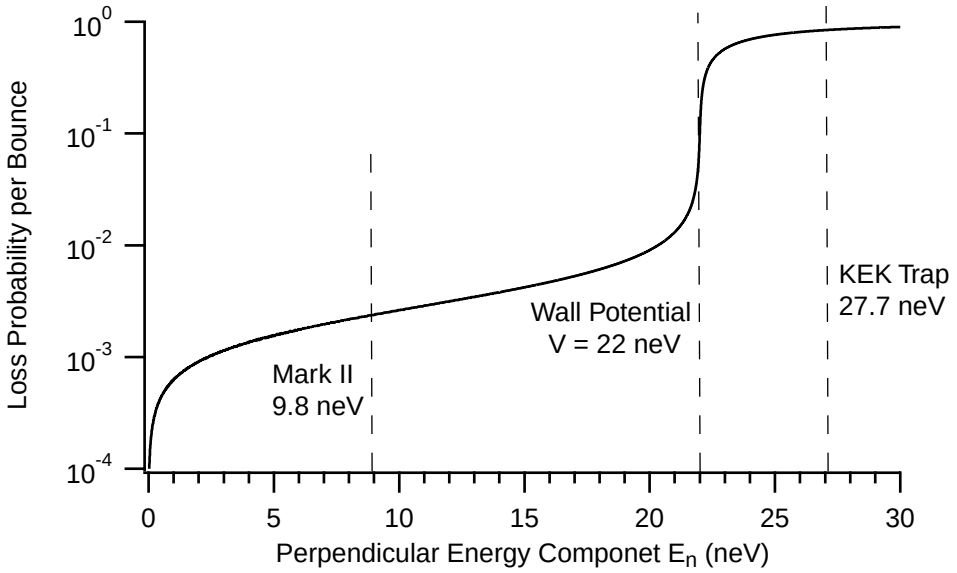

Figure 3.2: The neutron loss probability per bounce due to the trap wall potential. The horizontal axis is the perpendicular kinetic energy component of the neutron. The dotted lines denotes the real component of the wall potential V, and the minimum kinetic energy of the above threshold neutrons at the wall after the trap ramp down in both the Mark II and KEK trap.

where $\alpha = \sqrt{(V - E_n)^2 + W}$ and E_n is the perpendicular component of the neutron kinetic energy. When $E_n < V$, the loss probability per bounce of the neutron $p_l = 1 - p_r$ is on the order of $10^{-3} - 10^{-2}$, largely due to the imaginary component of the potential. When $E_n > V$, the loss probability becomes close to 100%, because the neutron energy is sufficient to penetrate into the potential barrier. Figure 3.2 shows the perpendicular energy dependence of the neutron loss probability.

As discussed in Section 3.1.3, the kinetic energy of above threshold neutrons at the trap wall increases when the trap is ramped down. The increase in kinetic energy helps the neutrons to escape the material wall potential.

In the Mark II trap, the above threshold neutrons have at least 9.8 neV kinetic

energy at the wall. When a neutron hits the wall perpendicularly, it has a loss probability of 2.4×10^{-3}. Since the cell wall surface (TPB coated on Gortex) was shown to be close to a perfect diffuse reflector from surface and light reflection measurements, we expect the incoming directions of the neutron to distribute evenly over 2π solid angle. Integrating over the solid angles for incoming neutrons, we obtain an average loss probability of 1.0×10^{-3} per bound. Given an estimated wall collision frequency of 20 Hz[3], the average lifetime of the materially bottled neutrons is approximately 50 s.

In the KEK trap, an above threshold neutron will have sufficient kinetic energy (27.7 neV) to penetrate the wall potential after the field ramp. When the incident angle of the neutron is less than $\arccos(22/27.7) = 36°$, its perpendicular energy component will be larger than the wall potential, and therefore can escape with probability close to 1. Again, integrating over the solid angle for incoming neutrons, we obtain an average loss probability of 0.071 per bounce. Given an estimated wall collision frequency of 30 Hz, we see that the average lifetime of the materially bottled neutrons is less than 0.5 s. Clearly, the deeper KEK trap is much more efficient at eliminating the materially bottled neutrons.

3.1.5 Fraction of Trapped Neutron Loss during Field Ramp

When the trap field is ramped down, a fraction of neutrons that is originally trapped will become untrapped in the shallower trap. Assume that the initial trap threshold is $E_i = \beta_i R$, and the field is ramped down from β_i to β_f. Using the adiabatic

[3]The trajectory of the above threshold neutrons intersect with the wall twice during each oscillation period, therefore average collision frequency is twice the radial oscillation frequency.

invariant $E\beta^{-2/3}$, we see that neutrons initially in the energy range $[(\beta_f/\beta_i)^{1/3}E_i, E_i]$ will now have energies in the range $[E_f, (\beta_i/\beta_f)^{1/3}E_f]$, and they can escape the trap due to collisions with the wall before the trap is ramped back up.

To estimate the fraction of trapped neutrons that become untrapped due to field ramp, we first calculate the energy distribution of trapped neutrons. Ignoring a slight asymmetry in the angular distribution of UCN production, the phase space available for UCN production with velocity v is

$$\Phi \propto v^2 \propto \sqrt{E_k}dE_k, \tag{3.11}$$

where E_k is the kinetic energy of the UCN. In the linear potential, the number of neutrons produced with energy below a threshold E_T can be expressed as,

$$N_{E_T} \propto \int_0^{E_T/\beta} rdr \int_0^{(E_0-\beta_i r)} \sqrt{E_k}dE_k \propto E_T^{3.5}. \tag{3.12}$$

If the field is ramped down to $\beta_f/\beta_i = 29.6\%$ to eliminate all stationary trajectories, only neutrons with initial energy less than $(\beta_f/\beta_i)^{1/3}E_i = 0.67\%E_i$ are still trapped[4]. Based on Equation 3.12, this constitutes only 26% of originally trapped neutrons. Certainly, the other 74% neutrons do not leave the trap right away, a fraction of them can stay in the trap for extended period due to the material bottling effect. When the trap is ramped back up, they would again become truly trapped.

Here we see a trade-off between systematic uncertainty and statistical uncertainty. As we eliminate more above threshold neutrons by ramping the trap down to lower field or keeping the trap at the low field for longer periods, we throw away a larger

[4]Because the boundary defined by the circular orbit with radius $2/3R$ which expands to R, neutrons with initial energy higher than $2/3E_i$ can collide with the wall after the field is ramped down.

fraction of originally trapped neutrons. We note that loading the trap at 29.6% of the maximum trap depth then ramping the trap field up has the same efficiency for eliminating above threshold neutrons. However, with this method, the population of truly trapped neutrons is only 16% of the neutrons that can be trapped with maximum trap field[5].

3.1.6 Addition of the Axial Motion

Because the effective potential for the axial motion is flat at the center, and rises sharply at the trap edge. The adiabatic cooling of a particle traveling in the axial direction from the field ramp is very small. To simply the discussion, we assume that the axial energy of the neutron remains the same during field ramp, which occurs in the limiting case of a square potential.

The addition of the axial motion lowers the energy associated with the radial motion. As a result, the trap field may need to be reduced further to eliminate certain stationary trajectories. If we assume that there is no mixing between the radial and axial motions, the worse-case scenario occurs when the energy distribution between the radial and axial motion is $0.83E_T$ and $0.17E_T$ respectively. In such a case, the trap field needs to be ramped down to 17% to eliminate radially circular orbits, and only 12.6% of originally trapped neutrons remain trapped at this threshold. For any other energy distributions between the radial and axial motions, a smaller field ramp amplitude is needed to eliminate all stationary trajectories. When the trap is ramped down to 17%, the minimum kinetic energy of the above threshold neutrons

[5]The UCN production rate is proportional to $E^{1.5}$.

at the trap wall would be 8.2 neV for the Mark II trap, and 24.2 neV for the KEK trap. The time it takes for the above threshold neutrons to escape the material wall potential is similar to the estimates in Section 3.1.4.

The motion of neutrons in actual Ioffe traps are more complicated. First, there is strong mixing between the axial and radial motions for deep traps ($B_1 >> B_0$). As shown in Ref. [81], a fraction of energy associated with the two types of motions would be exchanged back and forth between each other. Secondly, due to the fringing radial fields of the solenoids, the trap field is not axially symmetric, and the angular momentum L is no longer a constant of motion. Both effects actually help to eliminate stationary trajectories. As a result, the trap field will not need to be ramped down to 17% to ensure the complete elimination of marginally trapped neutrons as the simple analytical model suggests. Because of the difficulty of solving the equation of motion analytically in the full three dimensional trap, we will investigate the escape times of above threshold neutrons in actual Ioffe traps numerically.

3.1.7 Summary

We have shown with a simplified analytical model that by ramping the field down to a low threshold (17%), all theoretically possible stationary trajectories can be eliminated. The kinetic energies of the above threshold neutrons at the trap wall are increased due to the trap ramp. This increase in kinetic energies greatly help the above threshold neutrons to escape the material wall potential. A deeper trap such as the KEK trap is more suited for getting rid of the materially bottled neutrons. The drawback of the field ramp technique is that a large fraction of originally trapped

neutrons could also be thrown away during the ramp. The fraction strongly depends on the ramp amplitude.

3.2 Numerical Simulation

3.2.1 Trajectory Tracking Program

The trajectory of a neutron in the low-field-seeking state can be tracked numerically by integrating its equation of motion, a set of second order differential equation,

$$\begin{aligned} m_n \frac{d^2x}{dt^2} &= -\mu_n \frac{\partial B(x,y,z)}{\partial x} \\ m_n \frac{d^2y}{dt^2} &= -\mu_n \frac{\partial B(x,y,z)}{\partial y} \\ m_n \frac{d^2z}{dt^2} &= -\mu_n \frac{\partial B(x,y,z)}{\partial z}. \end{aligned} \tag{3.13}$$

To solve these differential equations we must be able to compute the magnetic field gradient at any point in space. For a given trap, we first compute the magnitude of the magnetic field on a three dimensional cartesian grid using the "BiotSavart" program[6]. The field magnitude at any spacial point can then be calculated from the grid points using bilinear interpolation [84], and the derivatives of the interpolation functions are taken as the field gradients [85][7]. Another approach to approximate the gradient is to fit the field on the grid to a set of special functions that satisfy the Laplace equation, $\nabla^2 B = 0$. Though the field from an earlier trap (Mark I) can be adequately approximated by a set of bessel functions [86], the fits for both the Mark

[6] A magnetic field modeling software by Ripplon Software.

[7] Our collaborator Kevin Coakley used three dimensional tensor spline functions to interpolate the field and its gradient, which ensures the continuation of interpolation functions across the grid points. Bilinear interpolation, though less accurate, is easier to implement and faster [80].

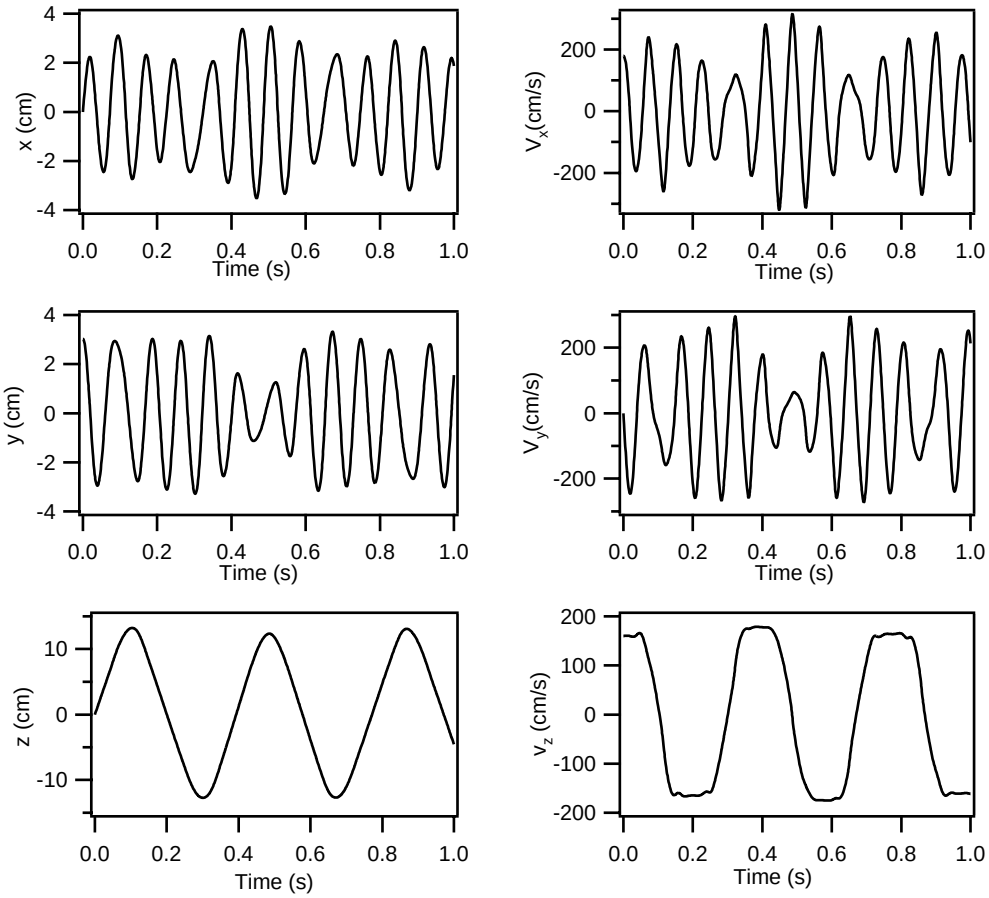

Figure 3.3: Computed trajectory for a neutron (Case A) with $E = 1.27E_T$ and escape time of 3.53 s.

II trap and the KEK trap do not converge properly, it is possible however that some other set of special functions can be found.

The initial conditions of neutrons can be either assigned manually or chosen randomly by the program. The trajectory computation then proceeds by propagating the position and velocity of the neutron through small time intervals Δt using a fourth order Runga-Kutta numerical integrator [84]. A typical trajectory in the Mark II trap is shown in Figure 3.3. The escape time of an above threshold neutron from the

trap can be determined by the time the neutron position crosses the boundary set by the trap material walls. For neutrons staying in the trap for longer than ten seconds, the escape times usually do not converge as the integrator stepsize is reduced. This numerical instability is due to the chaotic scattering behavior of the system[8]. For trajectories with converging escape times, an adaptive stepsize engine which optimizes the stepsize based the desired accuracy is implemented to reduce the computation time [84]. Results in the following sections except Section 3.2.2 are computed with the adaptive step size Rugga-Kutta integrator.

3.2.2 Chaotic Scattering

It is well known that chaotic behaviors exist in some scattering problems governed by classical hamiltonians [87]. One class of problems deals with particle escape times from the scattering region, for example a billiard ball escaping a window in the wall of a stadium shaped pool table. The escape times of marginally trapped neutrons can be dealt in the same theoretical frame.

One manifestation of chaotic scattering is its sensitivity to numerical truncation errors. The escape times for some neutron trajectories do not converge as the time step size of the numerical integrator reduces. Figure 3.4 shows the computed escape times of two sample trajectories (Case A and B) as the time step size is reduced from 10^{-3} s to 10^{-7} s. The predicted escape time converges in Case A, but not in Case B.

To study the convergence criterion for the computed escape times, one thousand neutrons within energy range $(1.05E_T - 1.60E_T)$ are randomly generated inside the

[8]A symplectic integrator which is better suited for tracking hamiltonian systems is also implemented to study the problem, but the numerical instability persisted [80].

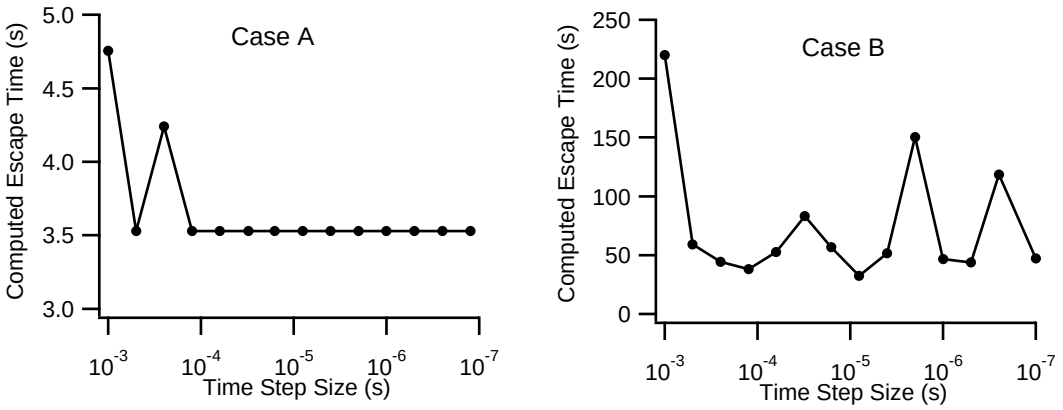

Figure 3.4: Computed escape time as a function of numerical integrator time step size for two neutron trajectories. Case A ($E = 1.27E_T$) converges to 3.53 s. Case B ($E = 1.37E_T$) does not converge, therefore exhibits chaotic behavior.

trap. The convergence of the escape time for each trajectory is determined by reducing the integration time step size from 10^{-3} s to 10^{-7} s. The results of the simulation are summarized in Table 3.2. For trajectories with escape times less than 5 seconds, the computed escape times all converge; for trajectories with escape times greater than 5 seconds, a significant fraction of computed escape times are not stable; and finally no trajectory with escape time greater than 100 seconds has a converging escape time[9]. This result agrees with Ref. [80] where it is observed that the computed escape time does not generally stabilize unless the escape time is short (less than approximately 10 s).

Among the 1000 simulated neutrons, 125 of them do not have converging computed escape times. We can bin results according to neutron energy, as shown in Figure 3.5. The number of non-converging events drops quickly with increasing energy. No such event is observed for neutron energy above $1.4E_T$, where E_T is the trap threshold.

[9]For trajectories whose escape times do not convergence, the computed escape times with the smallest time stepsize are taken as the nominal escape times for the purpose of classification.

Table 3.2: Stability of escape times for 1000 simulated neutrons in the energy range $1.05E_T - 1.60E_T$. This table lists the number of converging and non-converging trajectories in different escape time intervals.

Escape Time (s)	Converging Trajectories	Non-Converging Trajectories	Fraction of Converging Trajectories
0 - 5	812	0	100%
5 - 15	51	45	53.5%
15 - 100	12	50	19.4%
> 100	0	30	0%

The simulation results above suggest that trajectories with short escape times or high total energies do not exhibit chaotic behaviors. This agrees with an intuitive picture of chaotic scattering. Typically, chaotic behavior occurs for trajectories close to semi-stable orbits in the trap, because a slight change in initial conditions can significantly change the time a neutron spends orbiting the semi-stable orbits. Short escape times from the trap usually means that the neutron is not close to those semi-stable orbits. Chaotic behavior can also occur if the phase space available for escaping from the trap is small. In this case, when a neutron comes very close to the escape boundary, it encounters a bifurcation point, where the neutron can either exit right away or stay in the trap for extended periods before coming back to the boundary again. The existence of such bifurcation points can lead to chaotic behaviors. For neutrons with high energies, as the phase space for escaping the trap increases, the bifurcation points or semi-stable orbits described above are reduced, therefore suppressing the occurrence of chaotic scattering.

In summary, due to the occurrence of chaotic scattering, neutron trajectories cannot be reliably tracked for extended periods (> 10 s). This means that we cannot

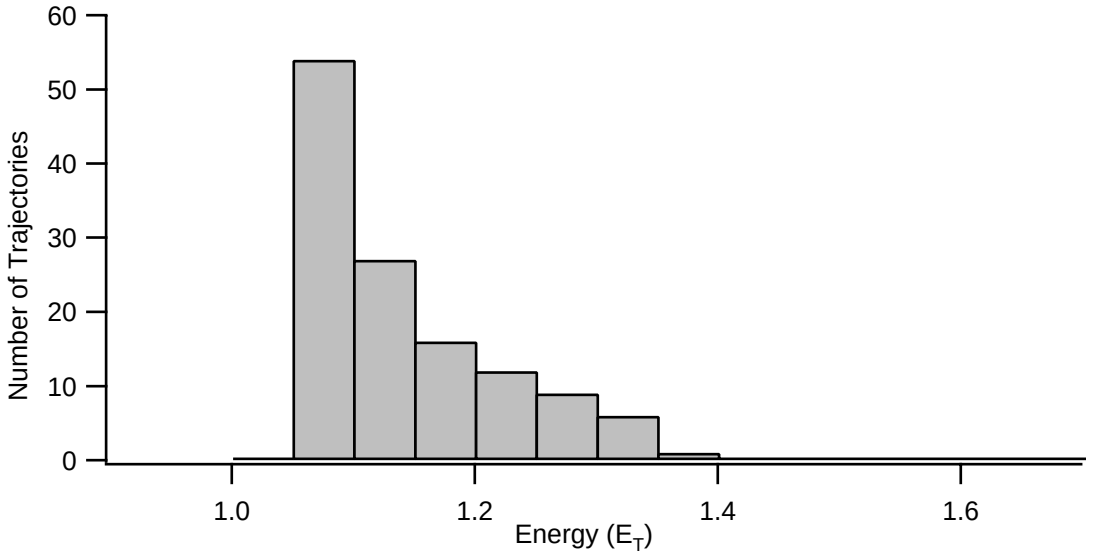

Figure 3.5: The energy histogram of non-converging trajectories from a simulation. The x axis is the neutron total energy in terms of the trap threshold E_T.

fully simulate the dynamic behaviors of above threshold neutrons, especially those with energy close to the trap threshold. However, the simulation can still be useful in several ways. First, we can use it to study the mixing between the axial and radial motions in the trap. Secondly, when the trap field is ramped down, the above threshold neutrons gain energy relative to the lowered trap depth. As the energy difference increases, all above threshold neutrons may enter the energy range where most trajectories can be reliably tracked. Numerical simulation can then be used to put a bound on the systematic uncertainty introduced by the above threshold neutrons. In the following sections, we will not calculate the convergence for each trajectory, but simply assume that trajectories with escape times less than 5 second are stable.

3.2.3 Coupling between the Axial and Radial Motion

As suggested in Section 3.1.6, the axial and radial motions in the neutron traps are strongly coupled, and this coupling allows us to more effectively eliminate marginally trapped trajectories.

We demonstrate the effect of this coupling in Figure 3.6 with two sample trajectories. Trajectory I is a circular orbit of a neutron with energy just above the trap threshold, starting out in the $z = 0$ plane. The simulation shows that angular moment is still approximately a conserved quantity. The projection of the motion in the x-y plane is close to circular. Though the axial motion was set to zero at the beginning, we still see an axial motion with a small amplitude at later times. The motion is caused by slight asymmetries in the trap field. To eliminate such an essentially two dimensional trajectory, the trap field would need to be ramped down to 29.6%, as discussed in Section 3.1.3.

Trajectory II is for a neutron with the same energy, but 17% of the total energy is assigned to the axial motion. The initial conditions of the radial motion are those of a circular orbit. If there were no coupling between the radial and axial motion, then the radial motion would stay on the circular trajectory, with an additional linear motion in axial direction. As discussed in Section 3.1.6, the trap field would need to be ramped down to 17% for the neutron on such a trajectory to escape.

However, the simulation result shows that as the trajectory propagates, its projection in the x-y plane is no longer circular. Furthermore, angular momentums is no longer a conserved quantity, and a fraction of energy is exchanged back and forth between the radial and axial motions. Both effects expand the phase space that the

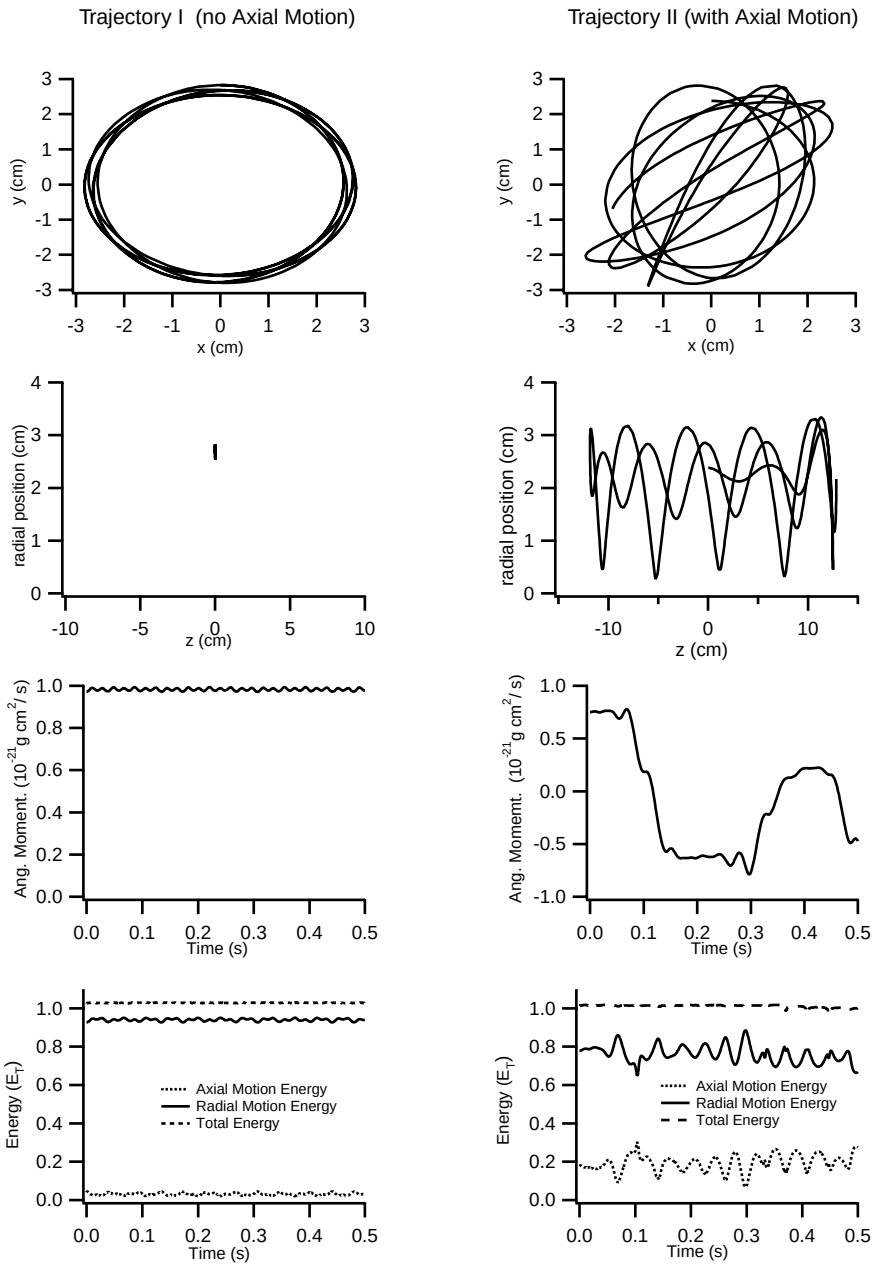

Figure 3.6: Two sample trajectories of a neutron in an Ioffe trap. Both trajectory I and II have the same total energy. Trajectory I has no axial motion, therefore angular momentum is approximately conserved. For Trajectory II, 17% of total energy is initially assigned to the axial motion. The angular momentum is no longer conserved and a fraction of energy is passed back and forth between the radial and axial motions, causing the neutron trajectory to be more ergodic.

neutron can reach, thus making it easier for the neutron to escape the trap. Indeed, although trajectory II started out on a smaller radius circle, the maximum radial position it can reach is larger than trajectory I. Ramping the field down to 29.6% therefore should be sufficient to eliminate this trajectory as well.

In summary, we have shown with an example that the addition of axial motion breaks down conserved quantities such as angular momentum, causing the trajectory to be more ergodic, hence easier to escape the trap. It is natural to hypothesize that if the trap is ramped down to 29.6% to eliminate the circular orbits with zero axial motion such as trajectory I, trajectories with non-zero axial motion such as trajectory II are also eliminated. We will check the validity of such a hypothesis using Monte-Carlo simulation.

3.2.4 Escape Times of Marginally Trapped Neutrons

In section 3.1.3, we proved that ramping the trap field down to 29.6% can eliminate all marginally trapped orbits for a two dimensional linear trap. In a three dimensional trap, because part of total energy goes into the axial motion, the proof is not applicable. However, as we have shown with an example above, the strong mixing between the axial and radial motion can help eliminate marginally trapped orbits, but would it be sufficient to eliminate all marginally trapped neutrons?

Simulations for both the Mark II and KEK trap were carried out. The trap field was set at 29.6% of its original depth. 2000 neutrons in the energy range of 20 - 48 neV for the Mark II trap and 60 - 120 neV for the KEK trap are randomly generated in the trap. Each trajectory is followed for 100 s or until the neutron exits the trap.

If a neutron does not escape after 100 s, an escape time of 1000 s is assigned.

Due to chaotic scattering, the neutron trajectory in general can not be reliably tracked beyond 5 - 10 seconds. Although the escape time of an individual neutron can not be reliably computed in general, one can reliably estimate the median escape times for marginally trapped neutrons in a particular energy interval provided that one can reliably compute escape times for at least half of the neutrons in the interval [80]. We divide the energy range into 40 equally spaced bins, and calculate the median escape time of neutrons in each bin.

The simulation results are plotted in Figure 3.7. Each dot represents a simulated neutron event, and the square symbols denote the median escape times. As discussed in Section 3.1.3, after the trap ramps down, the kinetic energies of above threshold neutrons at the trap wall are at least 14.8% of the initial trap depth. The minimum energy of the above threshold neutrons is denoted by a dotted line in the graph.

The simulation shows that the median escape time for above threshold neutrons with the minimum energy is roughly 0.2 s for the Mark II trap and 0.1 s for the KEK trap, corresponding to 1 – 2 radial oscillation periods. The majority of the above threshold neutrons ($> 99\%$) escape the trap within 1 s, corresponding to 1 – 2 axial oscillation period. Only 0.1% of above threshold neutrons escape the trap in more than 10 s.

The simulation agrees with our assumption that the coupling between the axial and radial motions helps to eliminate marginally trapped orbits, and most above threshold neutrons should exit the trap in one or two axial oscillation periods (the time period it takes for the coupling to occur). It also shows that ramping the field

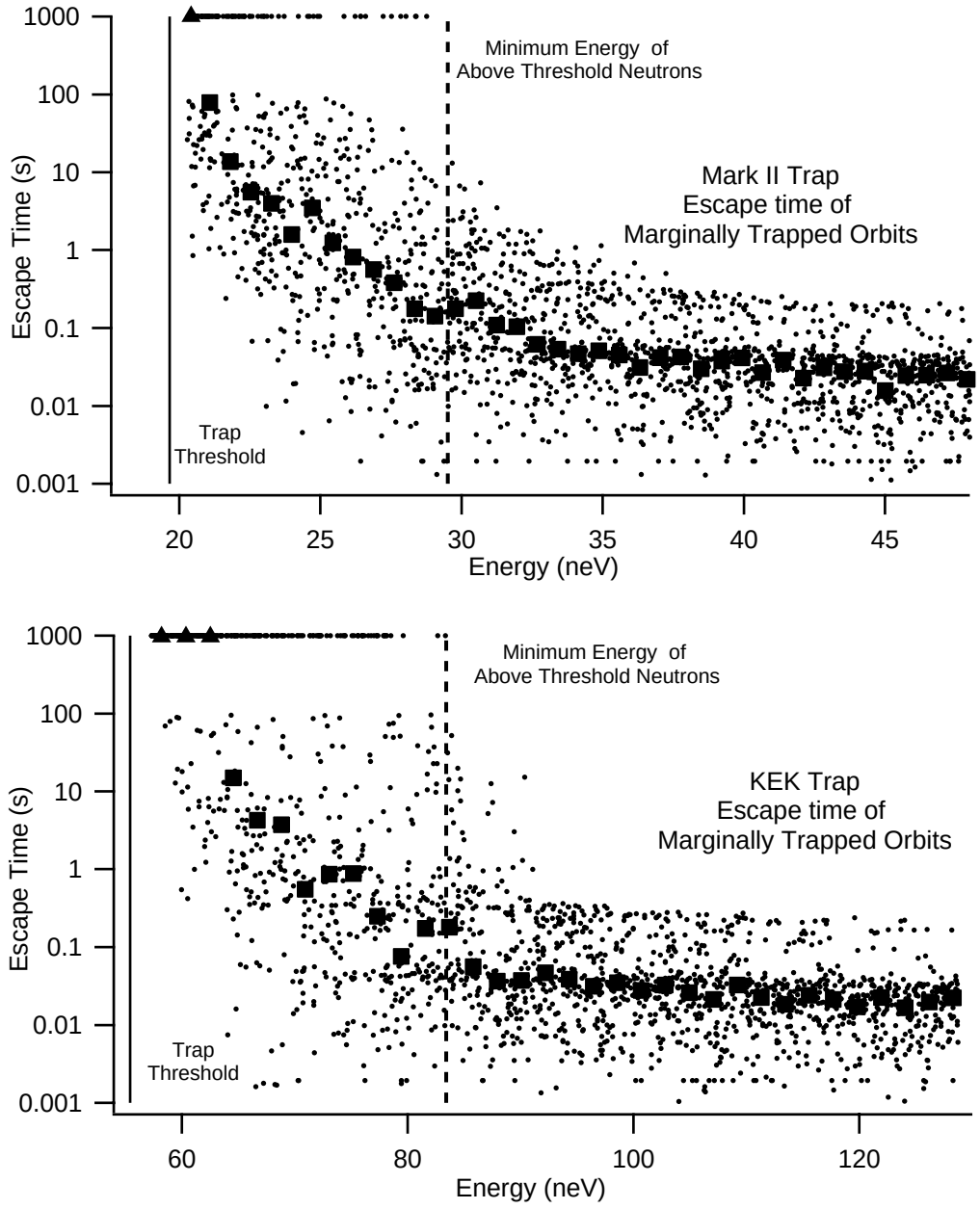

Figure 3.7: The simulated trap escape times of neutrons at 29.6% of trap depth for both the Mark II and KEK trap. The solid lines denote the trap thresholds, and the dotted lines denote the minimum energy of above threshold neutrons when the trap is lowered. The simulation halts after 100 s. Neutrons that do not escape after 100 s are assigned an escape time of 1000 s. The square symbols correspond to the median escape times less than 100 s. The triangle symbols correspond to the median escape times greater than 100 s, which can not be accurately determined.

down to 29.6% is not sufficient to eliminate all marginally trapped orbits. A tiny fraction of orbits ($\sim 0.1\%$) have escape times greater than ten seconds, for which we cannot determine the exact escape times from the simulation.

3.2.5 Escape Time from the Material Bottle

The material bottling effect of the trap can be easily added into the above simulation. When a neutron hits the cylindrical trap wall, its vertical velocity component is calculated, and Equation 3.10 is used to estimate the reflection probability of the neutron from the trap wall. We then generate a random number between 0 and 1. If the random number is less than the reflection probability, the neutron is considered to be reflected. Since the material wall is a diffuse reflector, the velocity vector of the reflected neutron is initiated with a randomly chosen direction. In the simulation, we neglect neutron reflections from two ends of the trap.

Again, simulations for both the Mark II and KEK trap are carried out. The trap field is set at 29.6% of its original depth. 2000 neutrons in the energy range of 20 - 48 neV for the Mark II trap and 60 - 120 neV for the KEK trap are randomly generated in the trap. Each trajectory is followed for 3000 s or until the neutron exits the trap[10]. If a neutron does not escape after 3000 s, an escape time of 5000 s is assigned.

The simulation results are shown in Figure 3.8. As expected, the Mark II trap, due to its small trap depth, exhibits a strong material bottling effect. For above threshold neutrons close to the minimum energy, the median of escape times is greater than

[10]Because the neutron trajectory tracking is re-initiated after each bounce, we can potentially track a neutron with many reflections reliably for a period much longer than 10 s. We thus set the program to run for a longer period.

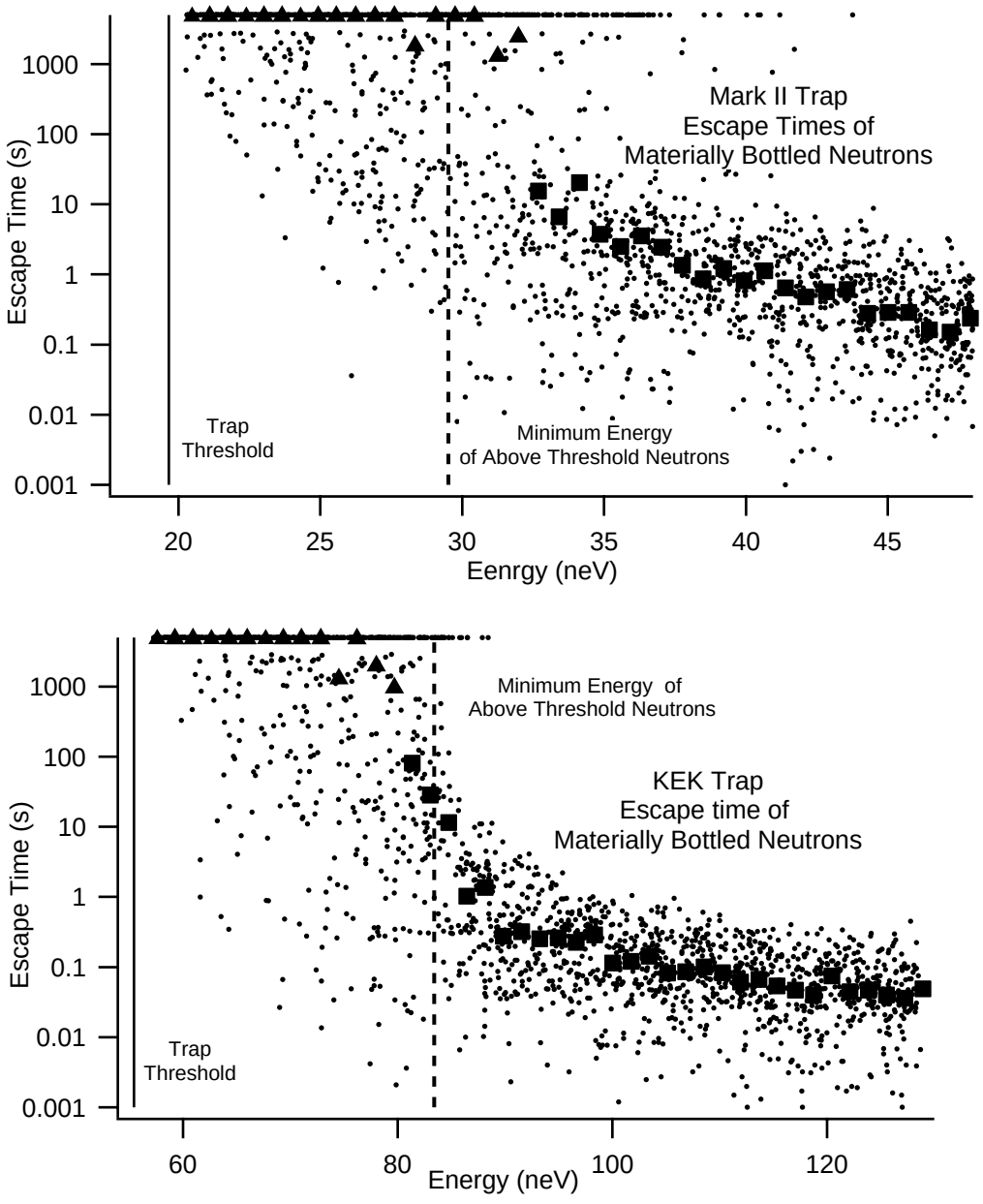

Figure 3.8: The simulated trap escape times of neutrons at 29.6% of trap depth for both the Mark II and KEK trap. The reflection from the material wall is taken into consideration. The solid lines denote the trap thresholds, and the dotted lines denote the minimum energy of above threshold neutrons when the trap is lowered. The simulation halts after 3000 s. Neutrons that do not escape after 3000 s are assigned an escape time of 5000 s. The square symbols correspond to the median escape times less than 100 s. The triangle symbols correspond to the median escape times greater than 100 s, which can not be accurately determined.

3000 s. In fact, the escape times of these neutrons likely can not be computed reliably. For the KEK trap, because the above threshold neutron will gain sufficient kinetic energy to penetrate the wall potential, the effect of material bottling is much less. The median escape time is roughly 20 s for above threshold neutrons with the minimum energy, and drops down quickly with increasing energy.

The simulation also shows that if the trap field is ramped down to 29.6% and stays at the low field for 10 s, taking into account of the material wall reflections, approximately 85% of above threshold neutrons can be eliminated for the Mark II trap, and approximately 98% of above threshold neutrons can be eliminated for the KEK trap.

3.2.6 Total Elimination of Above Threshold Neutrons

From simulations in the previous two sections, we see that a tiny fraction of above threshold neutrons still remain when the field is ramped to 29.6%. In order to completely eliminate above threshold neutrons, not only will the field need to be ramped down to lower values, but the wall potential also needs to be reduced. Figure 3.9 shows the escape times of above threshold neutrons for the KEK trap when the field is ramped down to 22% and the real part of the effective wall potential is reduced from 22 neV to 15 neV. As can be seen from the simulation plot, all above threshold neutrons in the simulation exit the trap within ten seconds when these conditions are met.

The wall material potential can be changed by doping the TPB coating with low neutron potential material such as titanium. The imaginary part of the wall potential

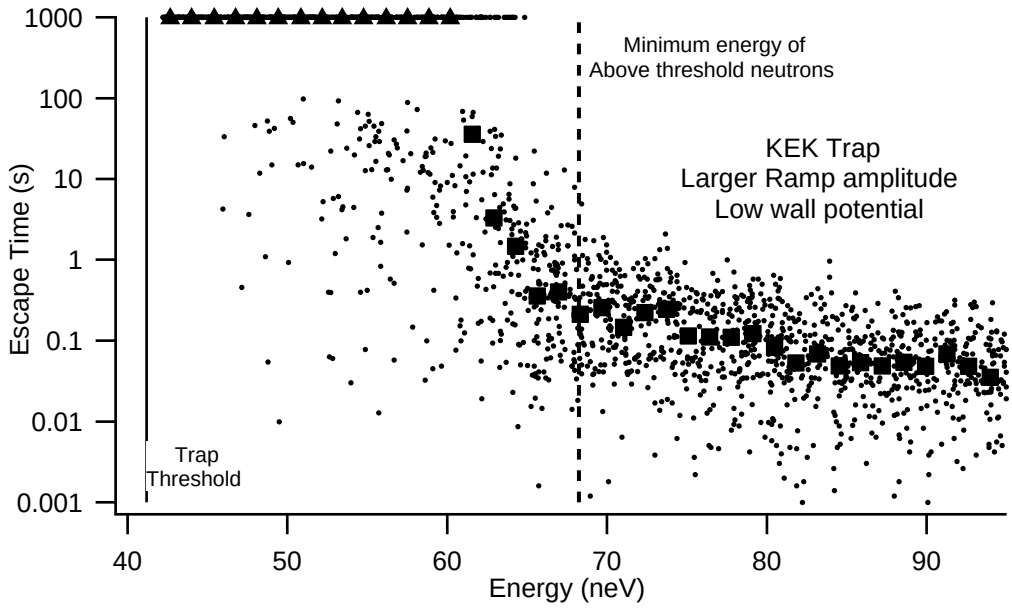

Figure 3.9: The simulated trap escape times of neutrons at 22% of trap depth for the KEK trap. The wall potential is reduced from 22 neV to 15 neV. The solid line denotes the trap threshold, and the dotted line denotes the minimum energy of above threshold neutrons when the trap is lowered. The simulation halts after 100 s. Neutrons that do not escape after 100 s are assigned an escape time of 1000 s. The square symbols correspond to the median escape time less than 100 s. The triangle symbols correspond to the median escape times greater than 100 s, which can not be accurately determined.

can be increased by doping it with hydrogenous or neutron absorbing materials, however, the imaginary component will need to be increased from 10^{-3} neV to 0.1 neV to have the same effect as reducing the real part of the potential from 22 neV to 15 neV.

Ramping the field down to 22% means that only 17% of neutrons remain truly trapped. The other 83% of originally trapped neutrons can escape the trap during the field ramp period. Assuming the field ramp takes 300 – 400 seconds, then from the simulation we estimate that up to 60% of those neutrons can escape the trap. This means that after the field is ramped back up, we would have 50% of the originally

trapped population remaining, minus neutron loss due to beta decay.

3.2.7 Summary

With the numerical simulation, we have shown that the strong mixing between the axial and radial motions of the neutron inside the Ioffe trap helps to eliminate marginally trapped orbits. In the three dimensional Ioffe trap, in order to totally eliminate the above threshold neutrons, not only will we need to ramp the field down to approximately 22%, we also need to reduce the trap wall potential to help neutrons escape the material bottling effect. We note that the results are obtained from Monte-Carlo simulations with limited number of events, and further study is needed to confirm the results.

3.3 Discussions

In this chapter, we have focused the discussion on the field ramp technique for eliminating the above threshold neutrons, because it is the easiest to implement. Two other techniques, using a neutron scraper or a RF spin flipper, can also be used to eliminate the above threshold neutrons, but both pose significant experimental challenges.

First, instead of ramping down the trap field to let the above threshold neutrons reach the trap wall, we can mechanically move a neutron absorbing surface, which we call a UCN scraper, across the trapping region to clean out the UCN energy spectrum. The advantages of this method is that the property of the neutron absorbing surface can be tailored to have very low material potential and large neutron absorption

cross-section, and not limited by the property of the organic fluor. The difficulty of the technique is the need to develop a mechanical system that can move the scraper in and out of the trapping region at below 300 mK with near zero heat generation.

Second, above threshold neutrons can be removed by sending radio frequency waves into the trapping region. When the neutron Lamor spin precession frequency is in resonance with the frequency of the RF wave, the neutron spin can be flipped, turning the neutron from the low-field-seeking state to the high-field-seeking state, and thus ejected from the trap. Calculation shows that to eliminate all the above threshold neutrons, RF power of several hundred watts would be needed. The liquid helium cell needs to be made into a high Q RF cavity to limit the heat load into the cell to less than a few miliwatts, which is not an easy task.

In summary, the field ramping technique seems to be currently the best method for eliminating the above threshold neutrons. By a combination of analytic and numerical models, we believe that we can clearly understand the systematic uncertainty related to the above threshold neutrons, and determine the conditions under which the above threshold neutrons can be totally eliminated.

Chapter 4

Experimental Apparatus

This chapter describes the neutron trapping apparatus used during the 2002 - 2003 data collection runs. Section 4.1 discusses the research reactor and the cold neutron beamlines at the National Institute of Standards and Technology (NIST). In Section 4.2, we report the design and testing of the superconducting magnetic trap. Ultracold Neutrons are produced and trapped inside a superfluid-helium-filled cell in the magnet bore. The cryogenic apparatus used to cool the helium cell to below 300 mK is described in Section 4.3. Finally, the system for detecting neutron decay events and the data acquisition system are discussed in Sections 4.4 and 4.5.

4.1 Neutron Beam

The neutron lifetime experiment is performed at the NIST Center for Neutron Research (NCNR). The NCNR operates a split-core[1] 20 MW research reactor. Neutrons

[1]In a split core reactor, fuel elements are located above and below the neutron ports. This arrangement can reduce gamma backgrounds with very little reduction in the neutron flux.

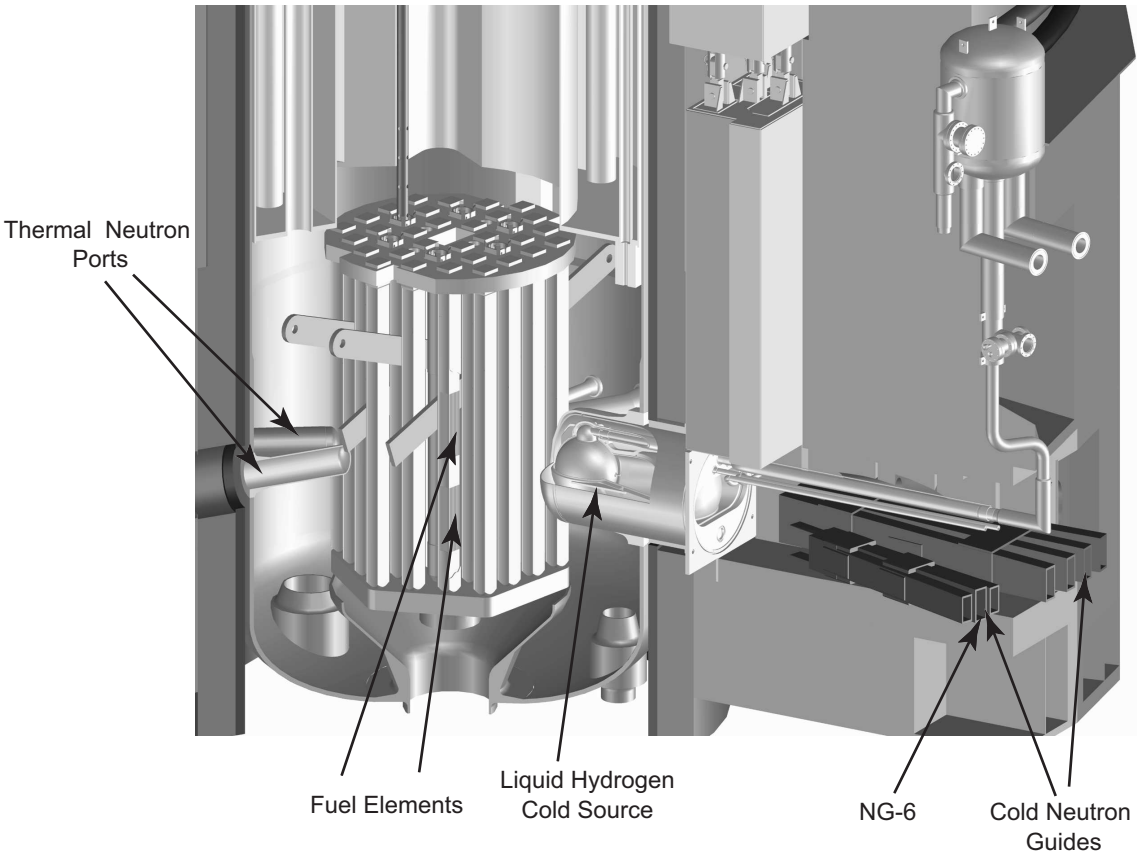

Figure 4.1: Cut-away view of the NCNR reactor core showing the thermal neutrons ports, the neutron cold source and the cold neutron guides.

are released through fission reactions of ^{235}U in the reactor core . The fission neutrons (1 MeV) are thermalized to room temperature (300 K, 26 meV) by a heavy water (D_2O) moderator surrounding the core. Nine small ports allow thermal neutrons to escape and enter a variety of experiments for material studies and neutron imaging. A 20 K liquid-hydrogen "cold source" is installed in a considerably larger tenth port, see Figure 4.1. Thermal neutrons pass through the cold source and are slowed via inelastic scattering with hydrogen atoms. The cold neutrons exit the source with a Maxwellian energy spectrum equivalent to 34 K. Eight neutron guides (NG) trans-

port the cold neutrons to experimental end-stations tens of meters away. The cold neutron guides are rectangularly shaped, 15 cm tall and 6 cm wide. The left and right surfaces of the guide are coated with ^{58}Ni, while the top and bottom surfaces are coated with m=2 supermirror. Cold neutrons with perpendicular energy less than the Fermi potential of ^{58}Ni (335 neV) will be totally reflected by the wall, and travel down the guides at the same glazing angle with minimal loss. The neutron guides are evacuated to eliminate air scattering.

Three beamlines are available for fundamental physics experiments at the exit of the NG-6 neutron guide, (see Figure 4.2). The upper 6 cm $\times$ 6 cm portion of the neutron beam exiting from the guide passes undisturbed to a 6 cm diameter hole in both a ^{6}Li neutron absorbing plastic and a tungsten collimator that defines the polychromatic (white) beam. Bismuth and beryllium blocks, cooled to 77 K, are used to reduce the direct gamma and fast neutron background emanating from the reactor core and upstream guide components. Experiments that require a high neutron flux operate on this polychromatic beamline.

The lower portion of the neutron beam from the NG-6 guide passes through three neutron monochromators before being absorbed by the neutron shielding. The first graphite monochromator (NG-6M) Bragg reflects 0.497 nm neutrons from the primary beam forming a monochromatic beam that is used for neutron calorimetry and polarization measurements [88]. The third monochromator (NG-6U) is a potassium intercalated graphite crystal. Its lattice spacing ($d = 0.874$ nm) allows the reflection of long wavelength neutrons, this monochromator is set at an angle to Bragg-reflect 0.89 nm neutrons, the optimal wavelength for superthermal production of UCN. A

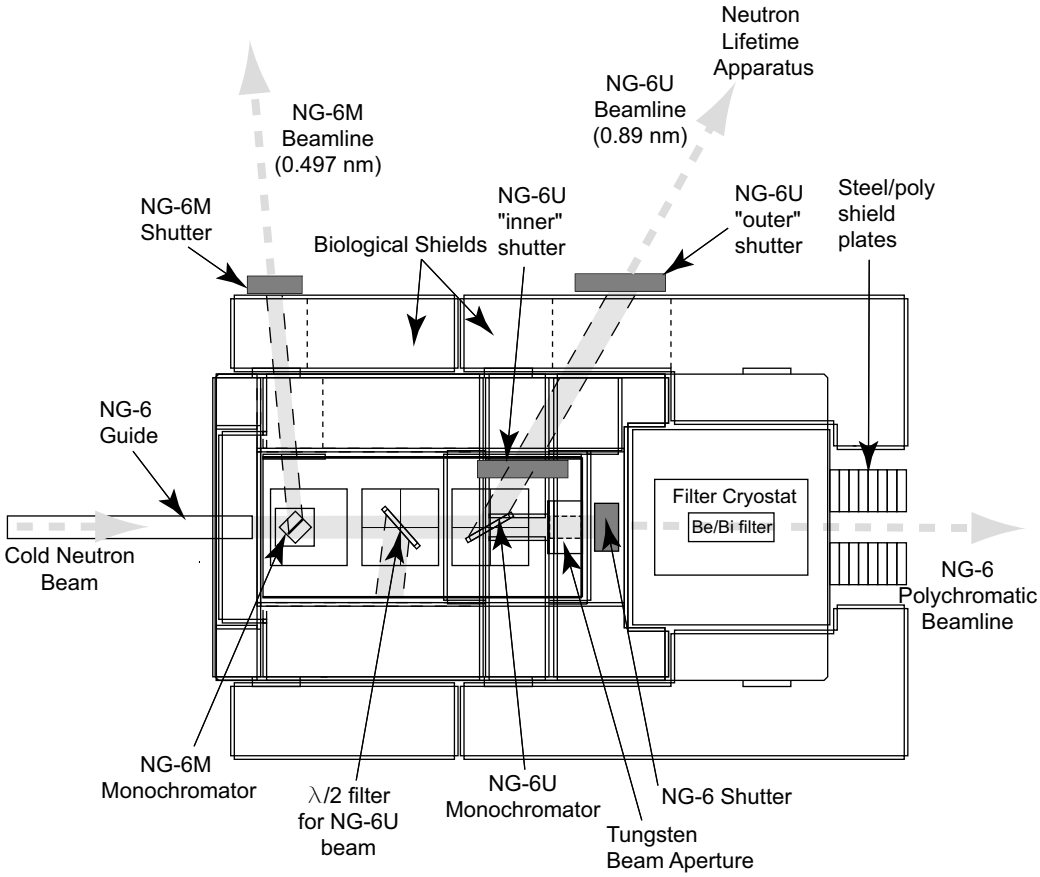

Figure 4.2: Schematic view of the cold neutron beamlines at NG-6 and the neutron shielding.

discussion of its development, installation and alignment can be found in Refs. [54] and [56]. Immediately upstream of the NG-6U monochromator is a second graphite crystal set an angle such to filter out the $\lambda/2$ component (0.45 nm) of the NG-6U beam. The neutron trapping apparatus resides on the NG-6U beamline.

The monochromators and filter cryostat are all enclosed inside biological shielding composed of steel shot and paraffin wax (see Figure 4.2. Two shutters are installed on NG-6U beamline and denoted as "inner" and "outer". The "inner" shutter is a 5 cm thick tungsten block covered with borated aluminum. The "outer" shutter consists of

a 5 cm thick steel block covered with ^{6}Li-loaded plastic. The operation of the "inner" shutter is automated by computer control. The "outer" shutter serves as a safety backup and is operated manually.

4.2 Magnetic Trap

The superconducting Ioffe-type magnetic trap is an essential component of the lifetime apparatus. Because an increase in trapping volume or trap depth translates directly to an increase in trapped neutrons, our group has been striving to build the largest and deepest trap possible with available technologies and resources for many years. This section briefly describes the Mark II trap used in the most recent experimental runs. Details of the design and quenching behavior of this trap can be found in Refs. [54] and [55]. A new trap capable of trapping twenty times more neutrons was recently tested. The design and testing of the new trap will be discussed in Section 6.2.

In the current trap, four racetrack shaped coils produce a radial confinement field. They are an experimental approximation of the four infinite current bars in the conceptual design, (see Figure 1.7). Two solenoid assemblies with identical current senses close the trap axially. Each assembly consists of a pinch coil and two bucking coils with opposite current sense. The pinch coil provides the axial confinement field. The bucking coil on the trap side increases the effective trap length, while the bucking coil on the outside decreases the fields at the turnaround regions of the racetrack coils, thus reducing the quenching probability at these high field regions. Figure 4.3 shows the configuration and current directions of the Ioffe assembly.

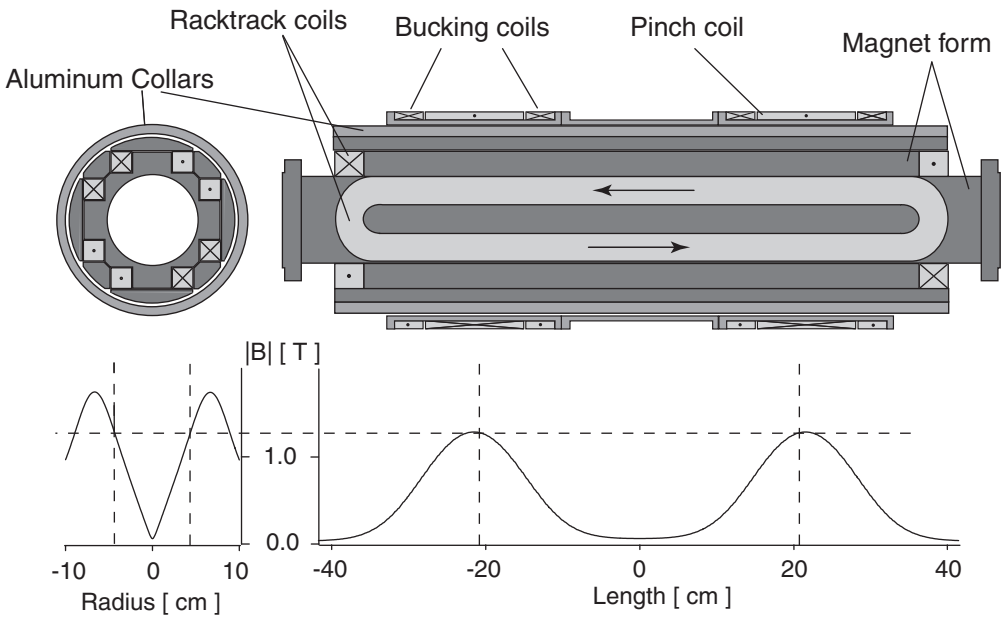

Figure 4.3: Configuration of our Ioffe type magnetic trap and its magnetic field profile.

The trap is designed to trap the maximum number of neutrons while still fitting in an existing dewar. The number of trapped neutrons scales roughly linearly with trap volume and as (trap depth)$^{3/2}$. Magnetic fields are calculated using a computer program "Biot-Savart". The optimized trapping region has a diameter of 8.4 cm and a length of 42 cm, see Figure 4.3. The designed trap depth is 2.2 T, but only 1.1 T is reached experimentally as will be discussed later.

All coils are wound with multi-filament niobium titanium (NbTi) superconducting wires[2] at Harvard University. The wires are bond in an epoxy matrix to prevent movement under Lorentz force when the magnet is energized, because even an imperceptible movement of a wire can cause sufficient frictional heating to quench the magnet. The requirements for making a good superconducting coil is to increase the

[2]Purchased from Supercon Inc. Wire diameter is 0.74 mm and the copper to superconductor ratio is 2:1.

wire packing, reduce epoxy rich regions[3] and eliminate voids in the epoxy. There are two techniques for winding epoxy bond superconducting coils, wet-winding and vacuum impregnation. During wet-winding, epoxy is applied as the coil is being wound. For vacuum impregnation, the coil is first dry wound with no epoxy, then put into a vacuum chamber. Epoxy is then introduced into the vacuum chamber and forced into the gaps between the wires. The vacuum impregnation technique is better at eliminating voids, but it involves complicated tooling and can only work using low viscosity epoxies. We chose the wet winding technique mostly for its simplicity. Epoxy[4] is applied after each layer of winding. Although the curing time of the epoxy is 24 hours, it becomes more viscous with time, so a new batch is mixed every two to three hours during winding.

Racetrack coils are more difficult to wind than solenoids. In a solenoid, wire tension during winding naturally translates into a compression force on previous layers that helps to increase wire packing and reduce epoxy rich regions. In a racetrack coil, wires in the straight sections of the coil are not under compression during winding, and have to be compressed after winding. Because epoxy applied earlier already starts to cure, it's difficult to push it out of gaps between the wires, therefore more epoxy rich regions are likely to be created in the straight sections of the racetrack coils. To prevent the crack propagations in these epoxy rich regions (a quench mechanism), a sheet of 0.5 mm thick fiberglass is placed in between each adjacent layer of the racetrack coils. Machined G-10 sheets cover the sides and faces of the racetrack coils to

[3]Because the thermal expansion coefficient of unreinforced epoxy is usually 3 - 5 times higher than the superconductor, even an epoxy rich region with a thickness of 0.5 mm can lead to cracking and magnet quench at low temperatures [89].

[4]The epoxy is Shell Resin 815C with 3140 hardener, mixed at ratio of 1:1.

protect the windings. Each current lead of the coils is made more robust by adding an additional superconducting wire to reducing the probability of a catastrophic quench occurring in the lead. In liquid helium tests, each racetrack coil reached at least 85% of the loadline (320 A) after two to six quenches [55]. Loadline is where the maximum field on the wire reaches the critical field of the superconductor.

In a quadrupole assembly, the Lorentz force tends to push the racetrack coils apart. The magnet form needs to provide sufficient radial compression to counteract the Lorentz force. This is usually done by prestressing the quadrupole assembly at room temperature with either Kevlar wrappings as in the case of our first trap [51] or stainless collars in the case of accelerator type magnet [90]. Because Kevlar creeps at low temperature [91] and compression with stainless collars requires specially designed compressor, we utilize differential thermal contraction between aluminum and titanium to achieve prestressing. The magnet form is machined out of commercially pure titanium. After the racetrack coils are put on the form, additional titanium bars are bolted onto the form to make it into a cylinder which is then machined on a lathe to have a uniform OD. Seven aluminum collars, 12 mm thick, are machined to slide around the titanium form with 25 μm clearance. When cooled down to 4 K, the aluminum collars shrink 240 μm more than the titanium form. Calculations show that the collars provide radial compression of 12.1 MPa [55], which is larger than the outward Lorentz force. The aluminum collars are designed to stay in the plastic deformation regime, so they can be thermally cycled without loss in performance.

Despite our best efforts during the design and manufacturing process, the whole magnet assembly only reached 180 A, 56% of the loadline (320 A), after repeated

quenching. The performance of the assembly is probably limited by the quadrupole assembly, because the solenoid assemblies when tested alone reached 230 A, 72% of the loadline. One possible explanation for the poor performance of both solenoid and quadrupole assemblies is that coils are not held rigidly enough. In the solenoid assemblies, the repulsive force between the pinch coils and bucking coils can stretch the aluminum form by 200 μm. Similarly, in the quadrupole assembly, the aluminum collars deform up to 200 μm under Lorentz force. A rule of thumb in the accelerator magnet community is that to produce a high performance magnet, wire movements should be limited to less than 40 μm [92]. Judging by this standard, our magnet form is not nearly strong enough. Although theoretically it's always possible to build a mechanical support that can hold the wire movements to less than 40 μm, experimental constraints such as space and available resource often makes this goal unrealistic. Under these circumstances, maybe the magnet shouldn't be expected to reach the loadline, and the magnet design should aim to be conservative, operating at 60% to 70% of the loadline. In the latest experimental run, the trap is operated at 160 A (50% of loadline), due to multiple quenches at this current.

Quench protection is necessary for large superconducting magnets. A quench occurs when a small section of the superconducting wire transits into a normal resistive state due to heating or above critical magnetic field. Ohmic heating drives the normal state wire to higher temperature which in turn leads to higher resistance and more heating. The quench also starts to propagate along the wire and across to the neighboring wires, turning more sections of the magnet normal. The entire energy stored in the magnet very quickly dissipates by ohmic heating. In a large magnet some-

times enough stored energy can be dumped in a small region of the magnet during a quench to cause structural damage or even vaporize the wires. In fact, a wire in one of our racetrack coils was vaporized during a quench and had to be replaced. After that incident, we installed a quench protection system. The basic idea is that when a quench is detected, the protection circuit quickly switches a resistor R in series with the magnet, then the energy of the magnet can be dumped into the external resistor with characteristic time $t_d = L/R$, where L is the inductance of the magnet. The method works efficiently when the resistance of the magnet does not rise much above R during the characteristic time t_d. Therefore, it generally works better with high current, lower inductance magnets. A quench is detected by observing voltage disbalance signals between matching pairs of coils. Inductive voltages cancel each other and produce zero disbalance signals, while resistive voltages from a quench do not cancel each other, and show up as a large disbalance signal. The details of the quench detection circuit and protection circuit can be found in Refs. [54] and [56].

Because some instruments in the Guide Hall such as "Spin-Echo" are very sensitive to magnetic fields, a compensating coil is implemented to reduce the fringing fields from our trap. The field from the quadrupole falls off as $1/\mathrm{R}^4$, while the field from the solenoid assemblies falls off as $1/\mathrm{R}^3$, therefore at large distance the fringing fields originate primarily from the solenoids. This far-field can be compensated to the first order by an external dipole moment with the same magnitude as the solenoidal dipole moment but pointing in the opposite direction. A dipole moment can be expressed as P = (number of current loops)×(current)×(area). The larger the area, the less amount of current or number of loops needed. The compensating coil is made from

copper bus bars, 5 cm wide and 0.6 cm thick, with maximum current capacity of 600 A. The coil is rectangular, 1.4 m wide and 3 m tall, enclosing an area 120 times larger than the solenoids. A total of six loops is needed, three on each side of the dewar supporting frame. The far-field compensating coils reduce the field at the "Spin-Echo" instrument (15 m away) from 9 mG to 0.6 mG. To ensure the proper operation of the photomultiplier tubes (PMTs)[5] used to detect scintillation light from the neutron decays, a near-field compensating coil is built to cut the field at each PMT from 26 G to below 3 G. This compensation coil is wound on a 43 cm diameter, 8.5 cm long aluminum form with 400 turns of AWG14 copper wires. The power supplies for the compensating coils are controlled by the power supply for the magnetic trap to achieve optimal cancelations at all currents. The control circuit is shown in Figure 4.4.

4.3 Cryogenic Apparatus

The cryogenic apparatus consists of a dilution refrigerator, an inverted T-shaped dewar, an ultrapure helium cell, superfluid helium link, neutron entrance windows and optical windows. The original apparatus is described in Ref. [51]. Various upgrades to the apparatus are reported in Refs. [55], [54] and [56]. Major components of the final apparatus in the latest experiment will be briefly discussed here, with an emphasis on the problems we encountered during assembly and cool-downs. Hopefully such discussion will help us identify problems to avoid in designing the next generation apparatus, see Chapter 6.

[5]PMTs do not operate in magnetic fields above 10G.

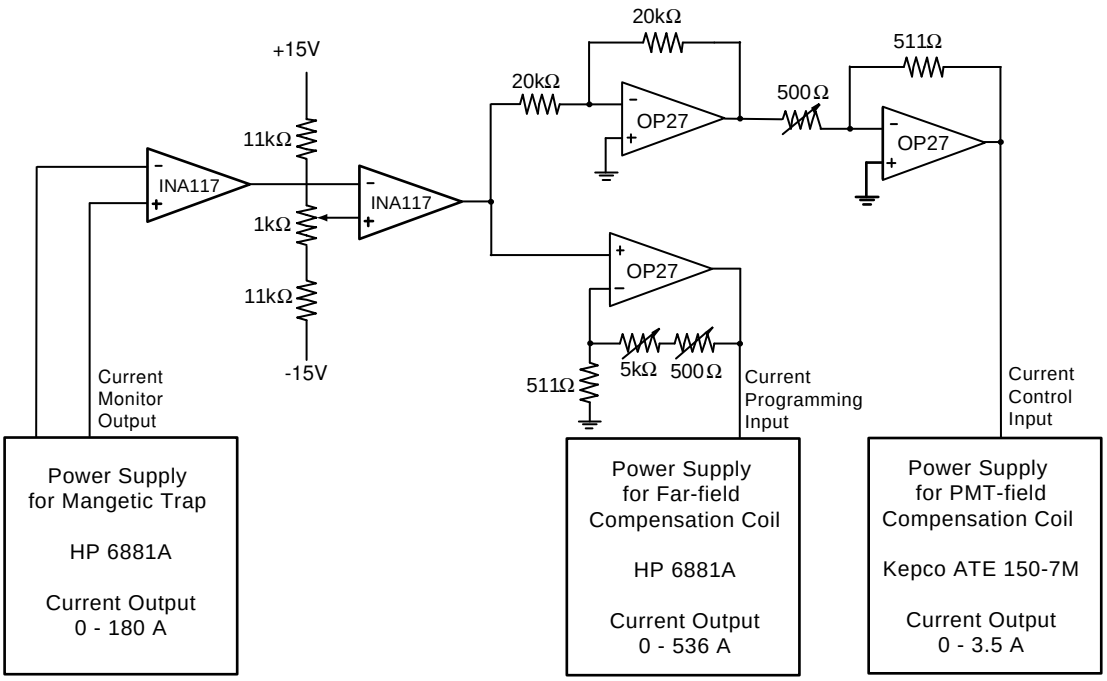

Figure 4.4: Circuit diagram for controlling both compensation coils with the magnetic trap power supply.

4.3.1 Dilution Refrigerator

To cool the ultrapure helium cell to below 300 mK, a dilution refrigerator unit is required, as it's the most commonly used technology to provide continuous cooling at temperatures below 300 mK. The idea of a dilution fridge was first proposed by H. London in 1951 [93]. As a mixture of ^{3}He and ^{4}He is cooled below a critical temperature ≈ 0.8 K, it separates into a "concentrate phase" rich in ^{3}He and a "dilute phase" rich in ^{4}He. Because the enthalpy of ^{3}He in the "concentrate phase" is less than in the "dilute phase", heat can be extracted from the system by "evaporating" ^{3}He across the phase boundary. Unlike normal ^{3}He or ^{4}He fridges where cooling power drops exponentially with temperature due to exponential drops in vapor pressures, the cooling power of a dilution fridge drops only as $1/T^2$, because the concentration

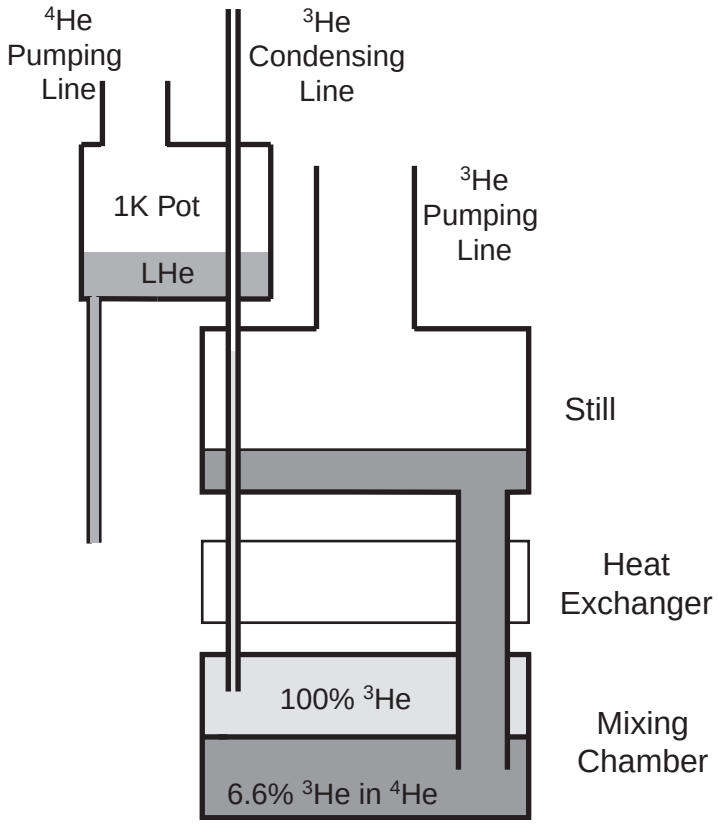

Figure 4.5: A sketch of the dilution fridge.

of ^{3}He in the dilute phase is finite even at absolute zero. A sketch of the dilution fridge unit is shown in Figure 4.5. The concentration of ^{3}He needs to be tuned so that the phase boundary occurs inside the mixing chamber. ^{3}He in the dilute phase is pumped into the still where it's further evaporated. 1 K pot is essentially a ^{4}He fridge used to precool returned ^{3}He gas. Detailed reviews on the principle and operation of a dilution refrigerator can be found in Refs. [94] and [57].

Our dilution refrigerator is specified to have a cooling power of 400 μW at 120 mK[6]. It was recently calibrated with a buffer cell attached at the bottom. The helium

[6]Kelvinox 400 manufactured by Oxford Instruments.

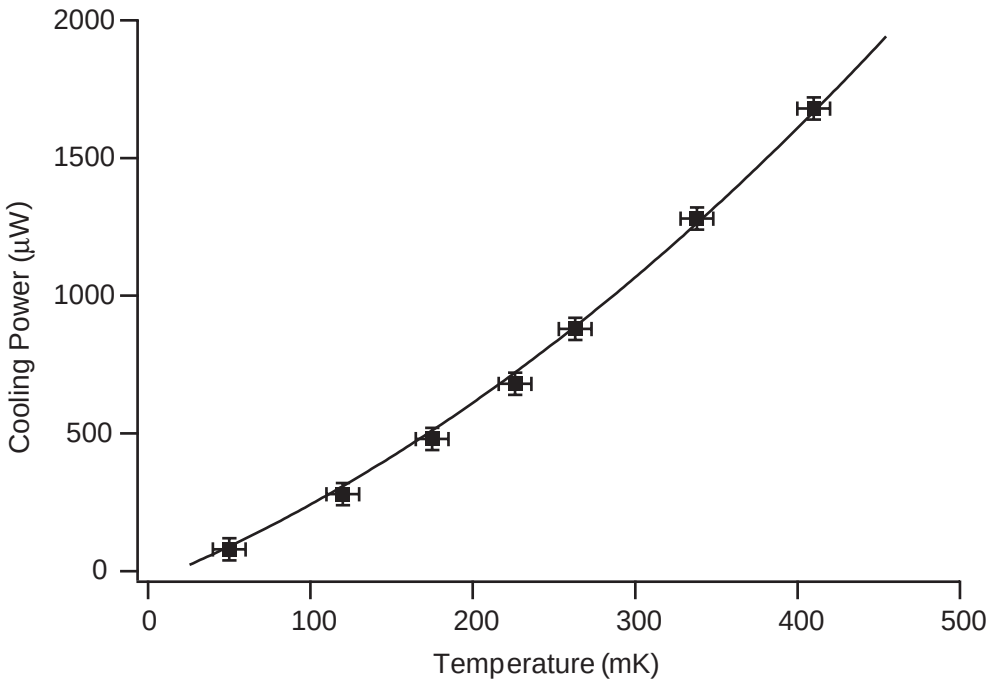

Figure 4.6: The cooling power of the dilution fridge measured in 2004.

film from the buffer cell introduces additional heat load of 56 μW [54]. The measured cooling power versus temperature curve is shown in Figure 4.6. Even after almost ten years, it still can produce 300 μW of cooling at 120 mK.

The main problem we encountered with the operation of the dilution fridge is the deterioration of its performance over 20 - 30 day periods. The pressure at the back of the helium 3 pump will slowly rise over time from 100 torr to 900 torr, at which point the mixture will flow back into the dumps and the fridge will start to warm up. Cleaning the helium cold traps or opening the second helium condensing line only helps for two to three days before the fridge warms up again. It's difficult to pinpoint a single cause. One main suspect is the residual helium exchange gas in the Inner vacuum chamber (IVC) (see Section 4.3.2). The residual gas tends to

condense on the coldest part of the fridge,i.e. the mixing chamber. After enough helium is accumulated, it can form a superfluid film that shorts the mixing chamber to warmer parts of the fridge. Evidence supporting this hypothesis is that after the fridge is warmed up to 4 K during reactor shutdowns and the IVC is pumped on for a few days, the fridge can usually come down to the base temperature again and stay there for a month before the same problem occurs again. It would be a good idea to implement more efficient pumping of IVC exchange gas in future experiments, either by lowering pumping line impedance or installing more charcoal sorbs[7].

4.3.2 Dewar

Because the dilution refrigerator has to be operated vertically, while the magnetic trap needs to be oriented along the horizontal neutron beam, a custom designed inverted T-shaped dewar is manufactured to house them both, see Figure 4.7.

The dewar has four thermal layers operating at different temperatures, an outer vacuum vessel at 300 K, liquid nitrogen shield at 77 K, liquid helium bath at 4 K and ultrapure helium cell at 300 mK. Both the 77 K and 4 K layers serve to block the blackbody radiation from 300 K. Multi-layer aluminized mylar films (super-insulation) are wrapped on the outside to reflect majority of the heat back to the warmer surfaces. The magnetic trap sits inside the helium bath. Its current leads are cooled by boil-off helium gas. Radiation baffles are used to block direct blackbody radiation into the helium bath.

The nitrogen and helium layers are mechanically supported from the top by a 25

[7]Charcoal has a very large surface area. At 4 K its helium surface sticking probability is close to 100%, so it is a very efficient helium sorption pump.

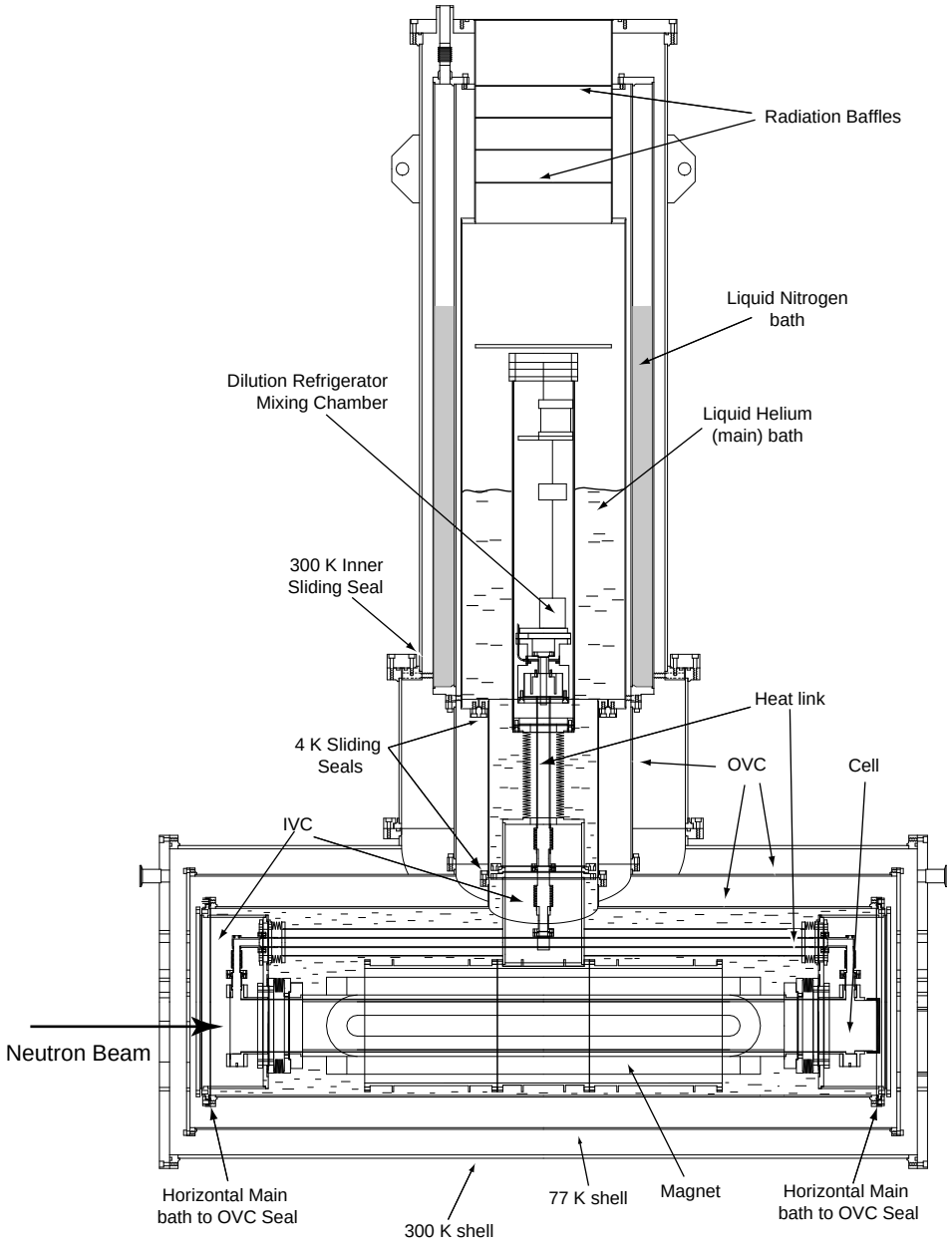

Figure 4.7: Mechanical drawing of the cryogenic apparatus. Details of the components are described in the text.

cm diameter thin-walled stainless steel tube to minimize the heat conduction from room temperature. Two vacuum spaces cut down the gas heat conduction between the layers. Between the 300 K and 4 K layer is the Outer Vacuum Chamber (OVC). Holes on the 77 K layer allow gas to flow freely across it. The space between the 4 K layer and the cell is the Inner Vacuum Chamber (IVC). It's necessary to separate the two vacuum spaces, because neon and helium exchange gases are introduced into IVC to cool the cell during the initial cool down.

Most problems associated with the dewar were leaks. The 300 K can is made of aluminum and sealed with rubber O-rings. The seals are in general very reliable. For most seals, the atmospheric pressure pushes the sealing surfaces together when the vacuum chamber is evacuated. But for the 300 K inner sliding seal (see Figure 4.7), the atmospheric pressure pushes the sealing surfaces apart with a force of 3.3 ton. Although ninety 1/4-20 screws are used to hold the two surfaces together, this seal still leaks from time to time.

The liquid nitrogen shield does not hold vacuum. It is made of aluminium for its high thermal conductivity at 77 K. Different sections of the shield are connected together by aluminum bolts to maintain the same contact pressure between connecting surfaces at low temperature.

The liquid helium layer is built with 316 stainless steel. Welded metal bellows are used to accommodate differential thermal contractions and misalignments between various parts. Because rubber loses its elasticity at low temperature, seals on the helium can are made by compressing indium wires between the sealing surfaces. Indium seals very well for small seals (< 5 inch diameter), but leaks often develop on

large seals (> 7 inch diameter). The most likely seals to leak on the helium can were two horizontal main-bath-to-OVC seals, and sliding seals at the vertical section, as indicated in Figure 4.7. It is believed that the horizontal seals are more likely to leak because of uneven cooling of the sealing flanges. During cooldown, cryogen is fed to the bottom the helium can. A large temperature gradient develops between the top and the bottom of the flanges, causing the flanges to warp and leaks to open up. On the other hand, it's believed that the sliding seals in the vertical section are more likely to leak, because of extra stresses on the seal; not only does these seals support the whole weight of the magnet, but they also experience additional torque when the horizontal seals are being tightened. In future designs, we should try to minimize such temperature gradient or external stresses on sealing surfaces.

The liquid helium main bath is first precooled using liquid nitrogen, then the liquid nitrogen is forced out using nitrogen gas, and finally the layer is cooled down by liquid helium. Most leaks develop during the liquid nitrogen cooling phase. A useful technique we developed for determining which seal leaks is to fill the main bath with liquid nitrogen, then add helium gas from the top. Because helium gas hardly diffuses through liquid nitrogen, if the leak is covered by liquid nitrogen, helium leak detector will show a sharp drop in helium leak rate. Based on the height of liquid nitrogen level in the main bath, we can determine if the leak is in the horizontal or the vertical section, and only seals in that section would need to be redone.

All leaks discussed above are leaks into the OVC. Leaks from the helium bath to IVC do not happen very often, probably because all indium seals are small. Leaks to the IVC are mostly from the cell, which will be discussed in the next section. In

general, a leak to the IVC has to be fixed because it will prevent the cell from reaching its base temperature, while a tiny leak to the OVC sometimes can be tolerated by pumping on OVC if the increase in liquid helium consumption is acceptable.

4.3.3 Experimental Cell and Heat Link

Inside the inner vacuum chamber are the experimental helium cell, the dilution refrigerator and the heatlink. The cell is made of a cupronickel tube. Two cell end caps are soft soldered to the tube. On the neutron entrance side, the cell has a 0.5 mm thick Teflon window. Although helium can diffuse through Teflon at room temperature, Teflon becomes impermeable at low temperature. Compressing the Teflon window against the end cap forms a superfluid tight seal [95]. On the other side, the cell has an optically clear acrylic window for viewing scintillation photons. The acrylic window is glued to an acrylic tube, which is glued to the outside of a thin metal fin, both with Stycast 1266 epoxy. This "snout" assembly (see Figure 4.8) is then attached to the end cap with an indium seal. The cell insert has three concentric layers. A 3.4 mm thick boron nitride layer protects the cupronickel tube from neutron activation. A 1 mm graphite layer blocks the neutron induced luminescence light from the boron nitride. The Gore-tex layer is part of the detection system. The cell insert is shown in Figure 4.11.

Several cupronickel tubes and beryllium copper bellows connect two ends of the cell with a copper buffer cell, which is attached directly under the mixing chamber, see Figure 4.8. The continuous space inside the cell, cupronickel tubes and the buffer cell can be filled with liquid helium through two 3.2 mm stainless steel fill lines. Su-

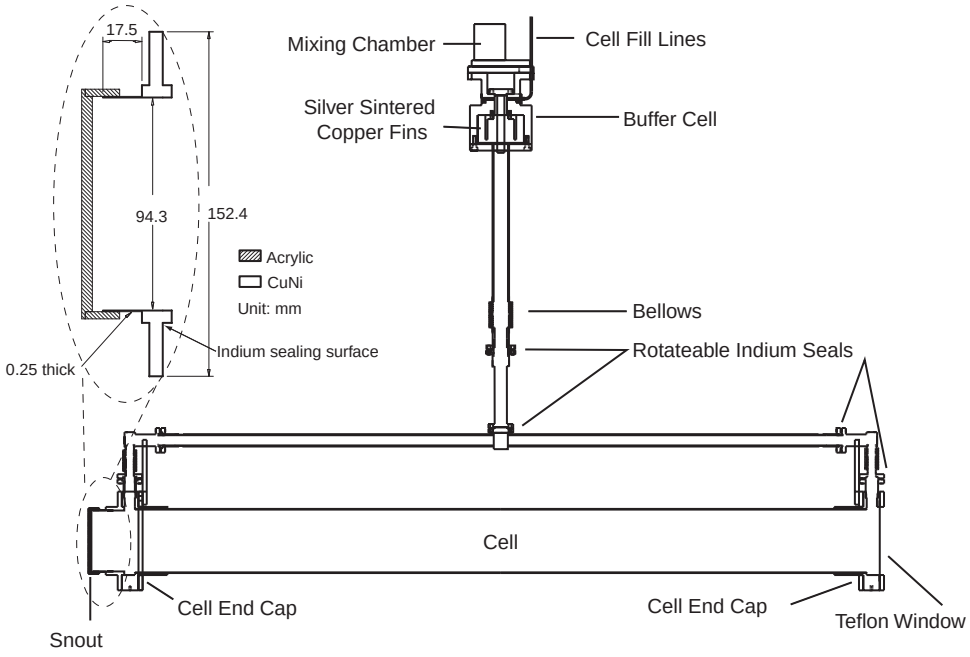

Figure 4.8: A drawing of the experimental cell, its superfluid heatlink to the mixing chamber and the buffer cell. The design of the snout is also shown in details.

perfluid helium inside the tubes is used as a heatlink between the cell and the dilution refrigerator. Not only does superfluid helium have higher thermal conductivity than any metal at temperatures below 1 K, it also does not generate eddy current heating during ramping of the magnetic field. Above 800 mK, good thermal link between the cell and the mixing chamber can be maintained by helium vapor even if the liquid helium level is below the buffer cell. Below 800 mK, the helium vapor pressure becomes too low for good thermal contact, and it's necessary to have the liquid helium level inside the buffer cell. If the level is higher than the buffer cell, liquid helium goes into the fill lines and can short the fridge. A capacitance level meter inside the buffer cell is installed to indicate the helium level. If the level meter is broken as was the case during the recent run, it was found to be better to underfill the cell than

overfill it. When overfilled, the whole cell will need to be warmed up to 2 K to pump out the extra helium. When underfilled, the cell can be filled slowly at 600 mK. The temperature of the cell ends will cool down together with the mixing chamber, when the liquid level enters the buffer cell. At temperatures below 100 mK, two copper fins with silver sinter coatings installed in the buffer cell can help overcome the Kapitza boundary resistance between liquid helium and metal [57].

Cell leaks to the IVC mostly occur on the cell soft solder joints and the acrylic to metal glue joint on the "snout". One of the solder joints has to be done with the cell inside the horizontal magnet bore. It's a difficult joint to make because melted solder tends to flow to the bottom, and can leave an uneven solder filling around the joint which opens up at low temperatures. The failure rate of this solder joint during first cooldown is roughly 30%. But once it is leak tight, it stays leak tight in subsequent thermal cycles. The glue joint on the "snout" is designed to have the acrylic tube, which contracts more than metals, shrink onto the thin metal fin whose flexibility partially compensates the differential thermal contraction. When the "snout" is made of cupronickel or stainless steel, the glue joint tends to delaminate after the first cooldown, and a leak usually opens up after two to three thermal cycles. One "snout" made of aluminum survived four thermal cycles and did not show signs of glue delamination, probably due to better matching of the thermal expansion coefficients between aluminum and acrylic. It is also important to keep the pressure inside the cell below 1.5 atmosphere during cooldowns or warmups to reduce the stress on the glue joints.

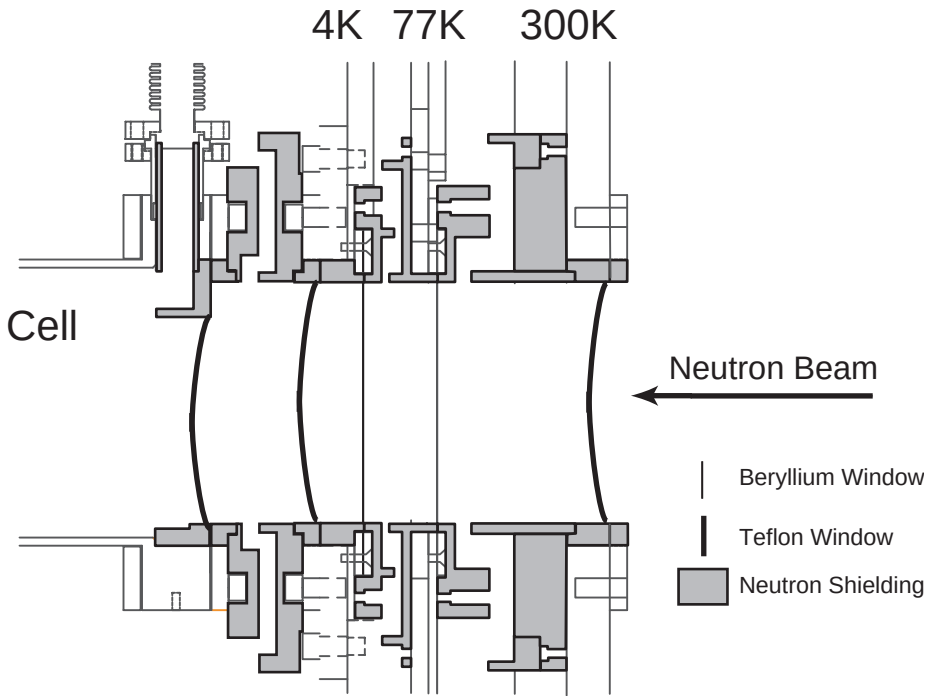

Figure 4.9: Sketch of the neutron entrance windows and the neutron absorbing BN shields used to protect flanges of the apparatus from neutron-induced activation.

4.3.4 Neutron Entrance Windows and Optical Access

Before the 0.89 nm neutrons enter the cell, they pass through a series of windows on the neutron entrance side of the dewar. Not only do the windows need to seal vacuum and block blackbody radiation, they also should have low neutron activation, low neutron induced luminescence and high neutron transmission. Teflon windows satisfy all above requirements except for blocking blackbody radiation. Separate beryllium windows, 0.05 mm thick, are installed on the 77 K and 4 K shields for that purpose. Because beryllium is extremely toxic when airborne, it is covered by a 25 μm thick teflon film on both sides. Interlocking boron nitride shielding surrounding the windows is designed to prevent neutrons from activating the cryostat.

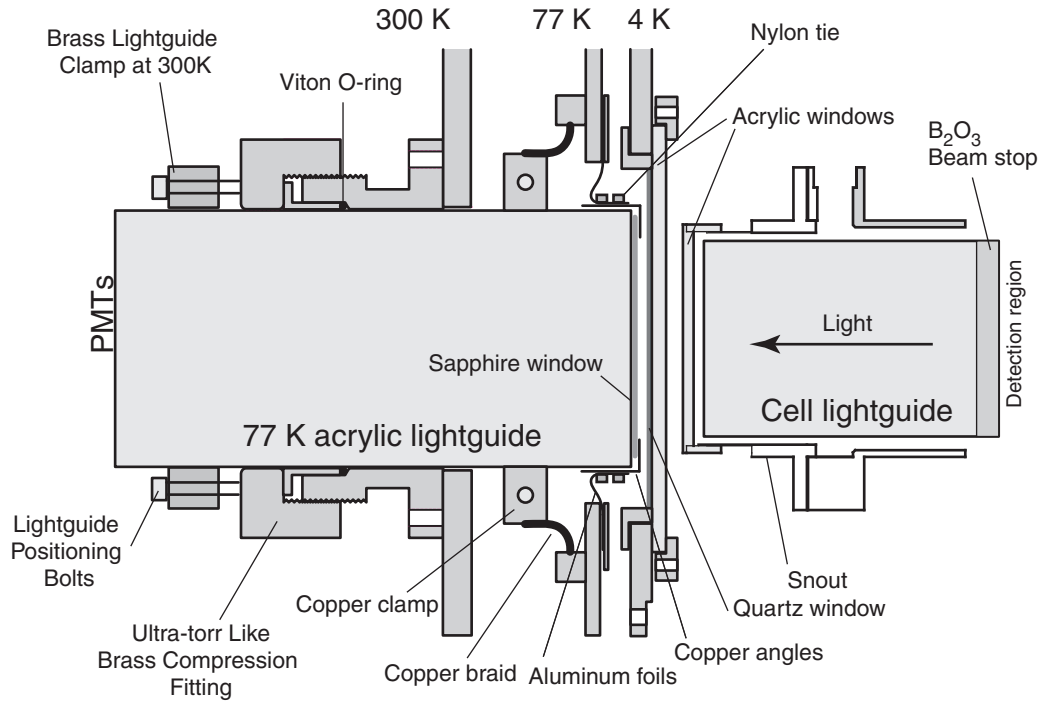

Figure 4.10: Light path from the trapping regions into the photomultipliers.

Graphite aerogel is sprayed onto the boron nitride to block the neutron induced luminescence. Because teflon windows bow inwardly under atmospheric pressure, the design of shielding pieces should take this deformation into consideration. Figure 4.9 shows the positions of the windows and the neutron shielding.

Optical access to the cell is through the other end of the dewar. Light from the cell is transported by the cell lightguide from the detection region to the snout. The light then goes through the acrylic window on the snout, quartz and acrylic windows at 4 K, a sapphire window at 77 K, and a lightguide extending from 77K to 300K, and finally splits into two photomultiplier tubes (PMTs), see Figure 4.10. The acrylic window at 4 K makes an indium seal to the 4 K stainless steel flange

separating the OVC from the IVC. It also blocks the blackbody radiation. However, acrylic is a poor thermal conductor. Without the quartz window, the center of the acrylic window heats up to 20 K, dumping large heat load onto the cell. The quartz window though has a lower cutoff wavelength for blackbody radiation, has much higher thermal conductivity. It is pressed onto the acrylic window to both block blackbody and cool the acrylic window. A single crystal quartz window would be sufficient for blocking the blackbody radiation, but it would be too fragile and too expensive for making an indium seal. The 77 K acrylic light guide is cooled by a copper clamp and aluminum foils on the outside. However due to the large diameter of the lightguide and the short distance between the 300 K and 4 K flanges, the center of the lightguide cannot be cooled adequately, and remains at above 120 K. Therefore a sapphire window that has higher thermal conductivity at 77 K is pressed onto the acrylic light guide to cool its front surface. At room temperature, a custom built ultratorr-like compression fitting seals the circumference of the lightguide with a viton O-ring, and a brass clamp and three positioning screws prevents the lightguide from moving under atmospheric pressure when the dewar is evacuated.

We had many troubles with blackbody heating when the window size was increased after the last upgrade. For a while, the blackbody radiation from the lightguide prevented us from cooling the cell below 800 mK. The sapphire and quartz windows were all added to cut down the blackbody heating. Direct cooling of the 77 K end flange also had to be implemented to help remove the conduction heating as well as blackbody heating from the lightguide [54]. It was also noticed that the heat dumped onto the 4 K windows were heating up the whole 4 K flange because of the low heat

conductivity of stainless steel. A copper plate had to be installed on top of the 4 K flange to conduct the heat to the liquid helium bath [56]. Some of these measures would not be necessary if the distances between the flanges were comparable to the lightguide diameter. For example, finite element analysis shows that the front face of the light guide can be cooled to 77 K without the sapphire window. Ideally, if detectors capable of operating at 4 K could be used to detect the neutron decay signals, we wouldn't need to worry about the blackbody radiation on the optical access side. Although blackbody heating from the neutron entrance side was not a big problem during last run. As the size of the window increases further, it may be necessary to increase the thickness of the beryllium windows to provide sufficient cooling to their centers.

4.4 Detection System

4.4.1 Overview

When a neutron decays in the trapping region, the decay electron creates a bright pulse of EUV scintillation photons. Because it is difficult to detect EUV photons at low temperature or to transmit them through windows, the helium scintillation light is converted into blue photons, then transported through a series of windows and light guides, and detected by PMTs at room temperature. Various light conversion and transportation schemes were considered and studied by our group, see Ref. [55] for a detailed account of the detector development. The detector insert used was based on an organic fluor covered diffuse reflector around the trap region and a clear

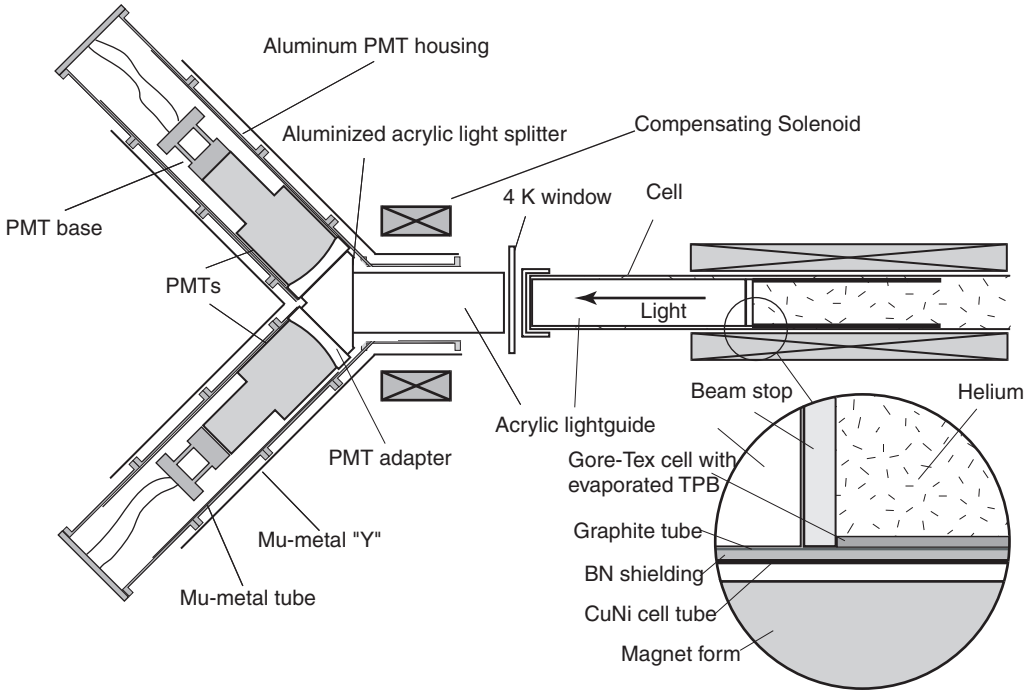

Figure 4.11: Sketch of the light detection system.

beam stop. This arrangement was found to yield the highest detection efficiency in the current trap geometry.

A sketch of the detection system is shown in Figure 4.11. Surrounding the trapping region is a 1 mm thick, 35 cm long and 8.4 cm diameter Gore-Tex tube rolled from two 17.5 cm $\times$ 28 cm sheets. The sheets are coated with 200 to 400 μg/cm^2 thick tetraphenyl butadiene (TPB) in an organic evaporator. Evaporated TPB coating exhibits a higher fluorescence efficiency[8] than the other organic fluors tested as well as sprayed TPB coating and TPB doped plastic film [96]. The measured fluorescence efficiency is (1.35 $\pm$ 0.1). Gore-Tex is chosen for its high diffuse reflection property. In fact, TPB coated Gore-Tex is characterized by NIST optical reflectivity group

[8]The ratio of emitted photons to absorbed photons.

with 430 nm light as a near ideal diffuse reflector. Because the efficiency of TPB deteriorates over time from UV exposures, the Gore-Tex insert along with the graphite and boron nitride cell inserts are assembled in dark room lit with weak incandescent light. At the end of the Gore-Tex tube sits a transparent boron oxide beam stop followed by a UVT grade acrylic cell light guide. The beam stop is hygroscopic, so its exposure to the atmosphere should be minimized. A sheet of Tyvek paper is wrapped around the cell light guide as a diffuse reflector to reflect the photons that does not undergo total internal reflection.

The EUV photons created by neutron decay electrons are frequency down converted by the TPB evaporated onto the Gore-Tex tube to blue photons peaked at 440 nm. A fraction of the blue photons reach the beam stop either directly or after several reflections off the Gore-Tex tube. After transmitting through the beam stop and the cell light guide, the blue photons pass through the 4 K windows, then are piped to room temperature by 77 K acrylic light guide. The 4 K windows and 77 K light guide are described in Section 4.3.4. At room temperature, the photons are divided by an aluminized acrylic light splitter into two Burle 8854 PMTs for detection. The PMTs are sealed inside light tight aluminum cans. The aluminum cans are overpressured with nitrogen gas to avoid the contamination of the PMTs by helium gas[9]. Because PMTs are sensitive to magnetic fields, μ-metal shielding as well as an active compensation coil are used to cancel the field from the magnetic trap. The two PMTs operate in coincidence to suppress luminescence background, which mostly consists of uncorrelated single photon events. Using a single PMT with a high

[9]Helium can diffuse through the glass window on a PMT. Once inside, helium can cause large afterpulses in the PMT.

threshold to suppress luminescence is not feasible due PMT afterpulses. Typically the probability of an afterpulse caused by a single photoelectron event is 10^{-3}. Since the luminescence rate in one PMT is 20 to 50 kHz, there will be 20 to 50 Hz large afterpulses in a single PMT, overwhelming the 2 Hz signal from the neutron decays.

4.4.2 Calibration

The detection efficiency can be calibrated by placing a ^{113}Sn source with a 364 keV beta line inside the cell. A low temperature motor system was developed to move the source across the central axis of the detection region. For most calibration runs, a single PMT was mounted at the end of the 77 K light guide. The pulse height spectrum of the PMT has a peak corresponding to light signals caused by the beta source. The position of the peak divided by the position of the single photoelectron peak gives the equivalent photoelectron numbers for 364 keV betas. Higher photoelectron number corresponds to higher detection efficiency. Figure 4.12 shows the position dependence of the photoelectron number for both PMT α and PMT β along the central axis of the detection region. Since no temperature dependence of the beta peak position is observed between 300 mK and 1.8 K, the data was taken around 1.5 K where the fridge cooling power is high enough to compensate for the heating from the motor. As shown in the data, the photoelectron number decreases as the distance between the source and the beam stop increases, because the wavelength shifted blue photons undergo, on average, more reflections off the TPB coated Gore-Tex surface. There is also a decrease in photoelectron number when the source is close to the beamstop. This is due to the decrease in the solid angle of TPB coating seen by the scintillation

EUV photons (the beam stop has no TPB coating). A simulation of the detection system using a non-imaging optics program "Guideit" produces the same position dependence for a beta source [97]. The data also show that the quantum efficiency of PMT α is about 30% higher than that of PMT β. The efficiency of the light splitter is measured to be 94% by replacing the one PMT setup with the two PMT coincidence setup. The total detection efficiency for neutron decay events requiring a certain threshold cut can be calculated by a Monte-Carlo simulation program which takes into consideration the beta spectrum of neutron decay and the measured position dependence of the detection efficiency [54, 56]. For example, the detection efficiency for a threshold of 3 photoelectrons per PMT is $48 \pm 6\%$. The largest uncertainty lies in determining the variation of detection efficiency away from the central axis, which is not directly measured.

The detection efficiency can be increased in several ways. First, it was noticed during a test run, that the boron oxide beam stop is not necessary, because the cell acrylic lightguide can act as an effective beam stop, i.e., the hydrogen in acrylic can scatter the 0.89 nm neutrons to the cell wall where they will be absorbed by the boron nitride shielding. Acrylic has both lower activation and luminescence than boron oxide. Calculations show that getting rid of the boron oxide beam stop can increase the light transmission by at least 22%; 8% from two surface reflections, and 14% from the transmission of boron oxide. Secondly, once the boron oxide is removed, the surface of the acrylic light guide can be coated with clear TPB doped polystyrene film. This will increase the detection efficiency close to the light guide significantly. Third, the light transportation efficiency to room temperature can be

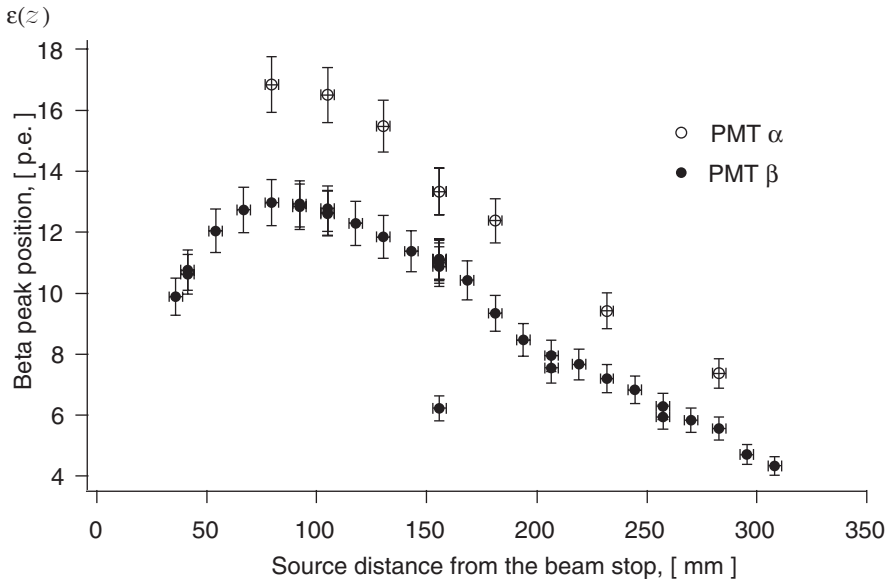

Figure 4.12: The position dependence of the detection efficiency calibrated by a 364 keV ^{113}Sn beta source moving across the central axis of the trap. The vertical axis is the equivalent photoelectron number of the peak in the pulse height spectra of a single PMT (α or β) mounted to the end of the 77 K light guide. 175 mm corresponds to the center of the trapping region.

increased. For instance, the sapphire window at 77 K can be eliminated when the geometry of the new dewar allows adequate cooling of the 77 K light guide. Polishing the lightguides and shortening distances between the lightguides and windows can also increase the detection efficiency. Fourth, PMT β can be replaced by a PMT with quantum efficiency at least as high as PMT α. Finally, if a reflective polymer film with low activation and luminescence can be placed at the end of the Gore-tex tube opposite to the cell lightguide, it can reflect the light exiting this end of Gore-tex tube back towards the cell light guide. A polymer film made by 3M is currently under our investigation.

4.4.3 Background Suppression

Since the detector is sensitive to all forms of ionization radiation, various techniques are employed to reduce backgrounds from ambient gamma rays, cosmic ray muons, and neutron induced activation and luminescence.

The ambient gamma field in the Guide Hall is the main source for the constant background in the experiment. Gamma rays can Compton scatter inside liquid helium, producing energetic electrons indistinguishable from beta decay events. Compton electrons can also cause Cherenkov radiation inside acrylic light guides. To reduce the gamma background, both the cryostat and the 300 K detection system are enclosed in a lead house with at least a four inch thick wall on each side. The neutron entrance window of the apparatus is shielded by 12 inches of lead piled on a sliding table to block direct gamma rays from the aluminum neutron monochromator holder. The sliding table can move the lead shield away during trap loading period to allow neutrons to enter the apparatus. Although the lead shielding has a large hole for the vertical section of the dewar, and many small gaps between the lead bricks, it reduces the two PMT single photoelectron coincidence rate from 280 Hz to 40 Hz. It was also observed that the coincidence rate is sensitive to scattered neutrons from nearby neutron beams. By adding sheets of boroflex and 2 inch thick polyethylene blocks on the outside of the lead house to absorb slow and fast neutrons, the coincidence rate can be further reduced by 2 - 3 Hz. To monitor changes in the ambient radiation background, a NaI detector is installed in between the two PMTs . If large changes in the ambient background were observed during an experimental run, the data of that run were be discarded.

Muons from cosmic rays can deposit their energies inside the liquid helium or the acrylic light guides to produce large flashes of light. Because typical muons have energies above 2 - 3 GeV, they will not be stopped inside the cryostat, and can be detected by plastic scintillators (muon paddles) surrounding the apparatus. Whenever the muon paddles register a scintillation event, the data acquisition system will be temporarily disabled for 280 ns to veto any scintillation events detected by the PMTs during this period. The constant background drops by 17 Hz with the muon veto added.

The time-dependent background comes from neutron beam induced activation and luminescence. As described in section 4.3, all metal parts in the neutron beam path are covered by high-purity neutron-absorbing boron nitride shielding. The luminescence from boron nitride shielding is blocked by high-purity graphite pieces. Materials inside the cell are chosen carefully after chemical analysis, neutron activation and luminescence tests [56]. Although fluorine in the teflon window and the Gore-tex can be activated, it has a very short decay lifetime of 16 s and will not interfere with the neutron lifetime measurement.

4.5 Data Acquisition System

The schematic of the data acquisition (DAQ) system is shown in Figure 4.13. The DAQ computer controls the high voltage supply for the PMTs, the inner neutron beam shutter, the slider and the inhibit line of the trigger circuit through its digital outputs. It also controls the current and voltage settings of the power supply for the magnet trap through a GPIB interface. Different experiment run schedules can be

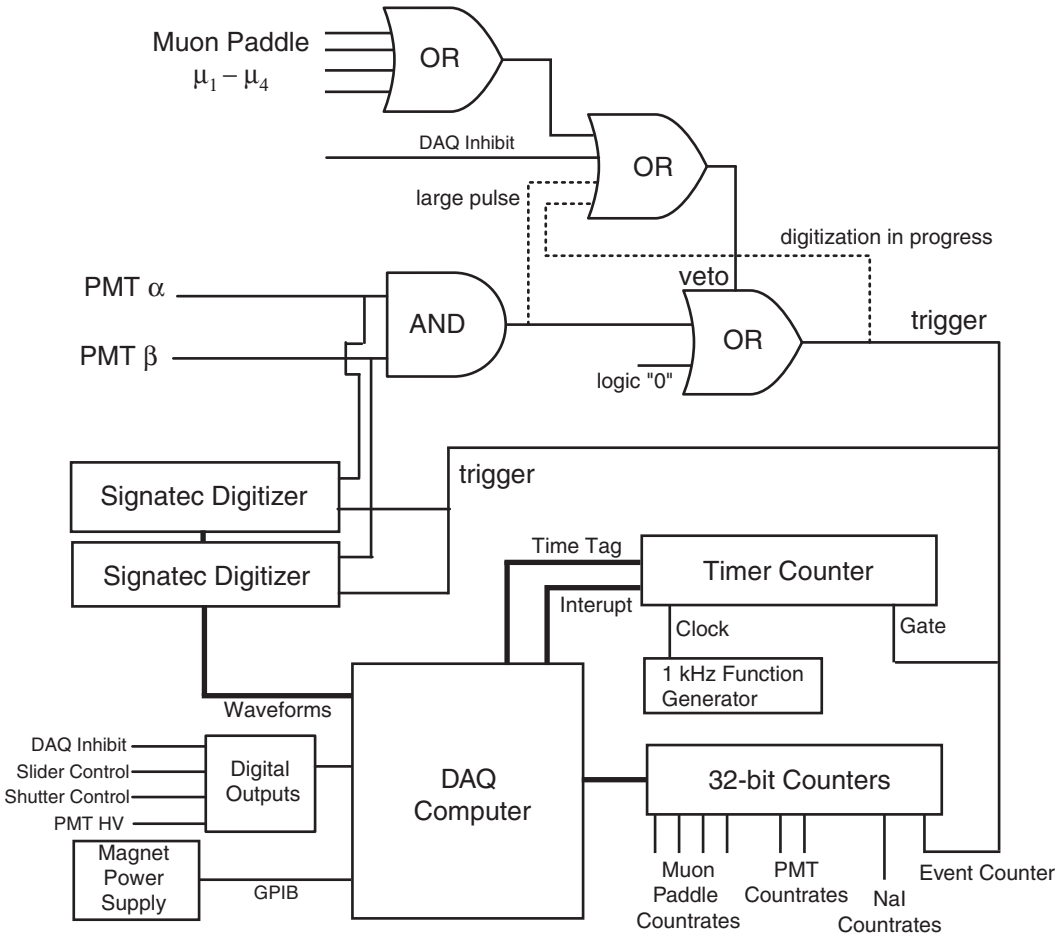

Figure 4.13: Schematics of the data acquisition system.

easily set by the DAQ software.

An event trigger is generated when both PMTs fire and the veto signal is not on. The veto is on if any of the following occur, one of the muon paddles fires, the DAQ computer sends an inhibit signal, waveform digitization is in progress, and right after large pulses are detected by the PMTs. It is observed that very large PMT pulses (amplitude greater than 1 V) are followed by many large afterpulses. These afterpulses can also set the trigger, causing the same pulse to be counted multiple

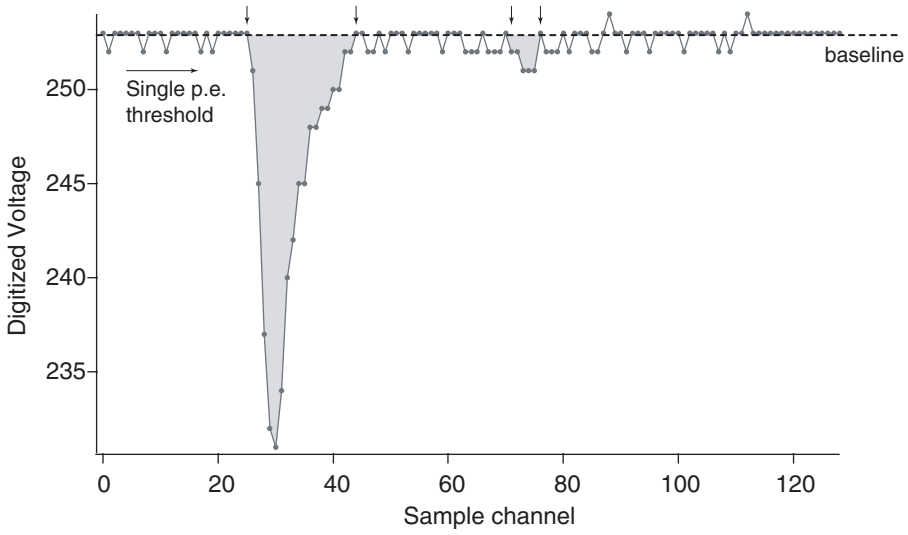

Figure 4.14: A typical digitized PMT pulse.

times. For this reason, the DAQ is disabled for 4 μs after large pulses are detected to veto these afterpulses.

When a trigger arrives, the pulses from both PMT α and β are digitized by Signatec PDA500 8-bit waveform digitizers. Each digitized waveform consists of 128 sample points, 2 ns apart. PMT pulses are usually 20 - 40 ns wide, corresponding to 10 to 20 sample points. The other sample points are used to establish the baseline. A typical digitized waveform is shown in Figure 4.14.

A 24-bit counter is used as a timer for the experiment. A 1 kHz square wave with 5 ppm frequency accuracy is sent to the clock input of the counter. The value of the counter, set to zero at the beginning of each run, represents run time in milliseconds. An event trigger sent to the gate of the counter card causes the value of the counter to be latched in an on board register and an interrupt to be sent to the DAQ computer. Upon receiving the interrupt, the DAQ computer reads out the value in the register

as the time stamp for the event. If a second trigger arrives before the first one is read out, an error flag bit is set. Matching the digitized waveforms with the time stamps is done in the analysis software.

In addition to the timer counter, eight 32-bit counters are set up to record the total event number, the count rate of each PMT, the count rate of each muon paddle and the count rate of the NaI background monitor. These counters are read out every five seconds.

A more detailed discussion of the data acquisition system can be found in Ref. [56].

Chapter 5

Results and Analysis

5.1 Overview

Data presented in this chapter were taken during six reactor cycles between 10/2002 – 8/2003. Total of 2441 data runs under various experimental conditions were recorded (see Section 5.2). Each run consists of a loading period during which the neutron beam was sent into the experimental cell, and an observation period during which the neutron beam was turned off and data acquisition system was turned on. The number of trapped neutrons builds up monotonically during the loading period, then decays exponentially during the observation period, (see Figure 5.1). Ideally, the neutron lifetime can be extracted by fitting the observed decay signals to a single exponential curve. In practice, several factors complicate the analysis, such as backgrounds and neutron loss mechanisms other than beta decay.

Backgrounds originate from ambient radiation, cosmic muons, neutron beam induced activation and luminescence. The use of PMT coincidence and muon veto

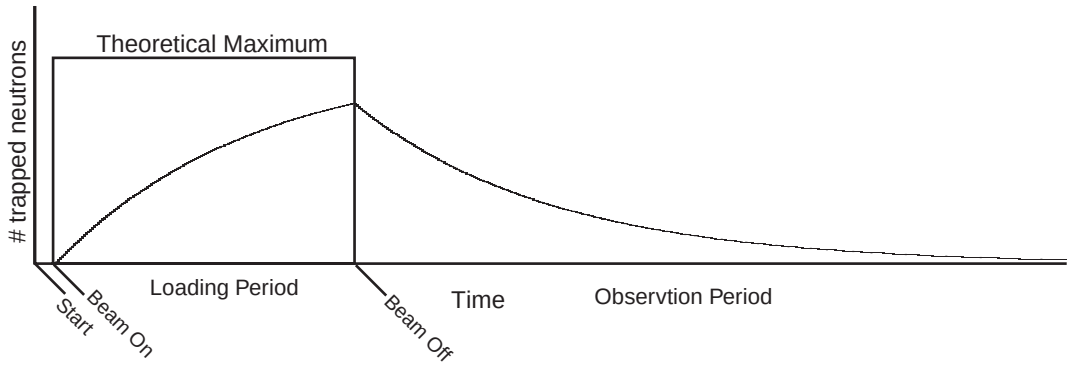

Figure 5.1: The accumulation and decay of trapped neutrons during loading and observation periods.

techniques can cut down the background rates considerably. Nonetheless, due to the limited number of trapped neutrons in each run, the signal to background ratio is still less than 1:6. In addition, the time-varying component of the background with a complicated nature exists even at high PMT pulse area cut threshold (see Section 5.4). Statistical analysis shows that when the model of the time-varying component is not well known, the trap lifetime can be better extracted by employing a background subtraction technique. We thus performed two kinds of runs, trapping (or "positive") and background (or "negative") runs. In the "positive" runs, the trap field is on during the loading period, while in the "negative" runs, the trap field is off during the loading period, (see Section 5.2 for the time sequence of each run). The count rate difference between the two types of runs can only come from decay signals of trapped neutrons or imperfect background subtraction.

The signal decay curve extracted from data runs taken at 300 mK exhibits a trap lifetime on the order of 650 s, considerably lower than the accepted value of the neutron lifetime (see Section 5.5.2). This short lifetime can be attributed to the

presence of above threshold neutrons. The data extracted from magnet ramp runs during which the field was briefly lowered to eject above threshold neutrons exhibits a trap lifetime on the order of 840 s, consistent with the neutron lifetime, though with limited statistical precision (see Section 5.5.3).

Systematic effect related to imperfect background subtraction is studied with ^{3}He runs where the ultrapure helium in the experimental cell was replaced by natural helium. The ^{3}He data we obtained is consistent with zero, which is a definitive proof that the decay signals observed from the 300 mK runs and magnet ramp runs originate not from imperfect background subtractions, but from trapped neutrons (see Section 5.5.1).

Possible neutron loss due to upscattering from superfluid helium is investigated by measuring the temperature dependence of the trap lifetime. Results are consistent with the theoretical model which predicts a T^7 dependence of the upscattering rate (see Section 5.5.4).

Finally, statistical and systematic uncertainties of the measurement are analyzed in Section 5.6.

5.2 Run schedules and data sets

In each data run, the data acquisition computer controls the operation of various experimental instruments. The control time sequences for typical data runs are shown in Figure 5.2.

During the first 100 second of each run, the neutron beam was closed and background events were recorded for diagnostic purpose. At the beginning of the trap

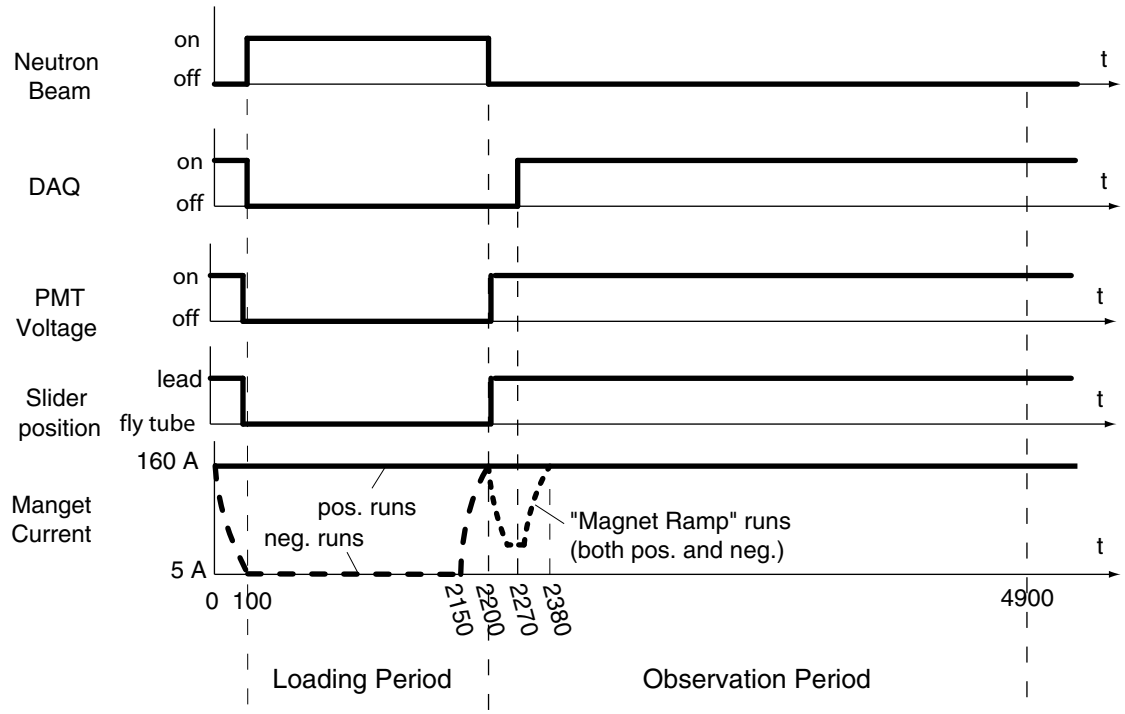

Figure 5.2: Control time sequences for typical data collection runs.

loading period which lasted 2100 s, the neutron beam shutter was opened, and the slider table (see section 4.2) moved lead shielding away to let neutrons enter the trapping region. Photomultiplier tubes (PMTs) were turned off during this period, because exposure to the intense scintillation light created by the neutron beam could potentially damage the PMTs if operated at high voltage. For positive runs, the magnet current was set at 160 A to create 1.1 T deep trap field, and the population of trapped neutron was accumulated. For negative runs, The trap field was ramped down to 5 A, and virtually no neutrons were trapped. The field was ramped back up 50 s before the end of the loading period to create the same magnetic field conditions for both the positive and negative runs during the observation period. At the

end of the the loading period, the neutron beam shutter was closed, and the PMTs were turned on. The slider table moved 12 inches of lead shielding back in front of the apparatus to block direct gamma backgrounds from the neutron monochromator. Because the data acquisition system could not handle the high initial signal rate, it was only turned on 70 seconds after the neutron beam was closed[1]. The observation period typically lasted 2700 s. In some runs, the observation period lasted 5400 s.

The run schedules for the magnet ramp runs were similar to typical runs. The only difference was that in the magnet ramp runs, the magnet field was ramped down briefly in both the positive and negative runs at the end of the loading period to eject above threshold neutrons. The whole ramp sequence took 180 s: first the magnet current was ramped down from 160 A to 48 A in 80 s, the current then remained at 48 A for 20 s, and finally the current was ramped back up to 160 A in 80 s.

The positive and negative runs were taken in an alternative manner[2]. Long term variations in backgrounds either due to external source or neutron activation of isotopes with long lifetimes could be mitigated by this procedure. Background variations was monitored with a NaI detector between the PMTs. If sudden changes in background was observed during a run, data from this run was thrown out.

Total of 2441 useful data runs were collected. They could be divided into several data sets based their experimental conditions and run schedules. A summary of the data sets is given in Table 5.1.

The "Cold" runs were all taken with the cell temperature below 300 mK. Runs

[1]The high initial rate was caused by ^{20}F activation which has a lifetime of 15.9 s. 70 s was sufficient for most of the activation signal to decay away.

[2]A typical run sequence goes like "+−+-++-" and so forth. The run sequence repeats itself with the sign reserved after each 2^n runs (n = 0, 1, 2...). The period for such a sequence is infinite.

Table 5.1: A summary of collected data runs between 10/2002 – 8/2003.

Set Name	T_{cell} (mK)	#Positive	#Negative	Comments
Cold1	< 300	358	358	
Cold2	< 300	136	136	long observation Period
Cold3	< 300	112	112	delay between runs
^{3}He Cold	< 300	109	109	natural helium in cell
^{3}He Warm	> 1000	33	33	natural helium cell
MagnetRamp	< 300	148	125	Magnet ramp schedules
500mK	500	96	93	
700mK	700	79	82	
850mK	850	161	161	

in data set "Cold2" were taken with a long observation period (5400 s) to better determine the baseline in a three parameter fit to the data. Runs in data set "Cold3" were taken with a 2700 s delay between each run. These runs were used to investigate initial discrepancy seen between the "Cold1" and "Cold2" data sets. The final results show that all three data sets are consistent with each other, and they can be combined to determined the neutron trap lifetime at 300 mK.

The "^{3}He Cold" and "^{3}He Warm" data sets were taken with natural helium inside the cell. Because ^{3}He inside natural helium could completely absorbed the trapped neutrons, these runs only measured the background. No temperature dependence of the background was observed in the data, so the two sets can also be combined during analysis.

The "MagnetRamp" runs were taken with the magnetic field lowered briefly at the end of trap loading to eject the above threshold neutrons. They are analyzed in Section 5.5.3

Data runs in data sets "500mK", "700mK" and "800mK" were taken with the experimental cell at temperatures indicated by the set name. The temperature of the cell was stabilized using resistor heaters on both cell ends controlled by a temperature feedback loop.

5.3 Data processing

As describe in Section 4.5, each data file consists of digitized PMT waveforms, time stamps, and binner data. During data analysis, the data files are first sent through a data analyzer program. The analyzer program reads in the digitized PMT waveforms and calculates the pulse areas. It then matches the time stamps to the corresponding pulse areas from both PMT α and β. The binner data which records the total number of events, the count rate of each PMT, the count rate of each muon paddle and the count rate of the NaI background monitor in five seconds time bins are also read out by the analyzer. These data are primarily used for background studies.

The calculated PMT pulse areas can be histogrammed to produce a pulse height spectrum, see Figure 5.3 for the pulse height spectrum of PMT β. The one, two and three photoelectron peaks can be clearly identified in the spectrum. However, due to the limited gain dispersion of the PMT, higher photoelectron peaks cannot be resolved.

Because the background from neutron induced luminescence can be reduced significantly by setting higher pulse areas threshold, the data from the analyzer is sent through a data cutter, where data events are selected based on required pulse area thresholds for both PMTs. A typical cutter threshold setting is "3+3+", which

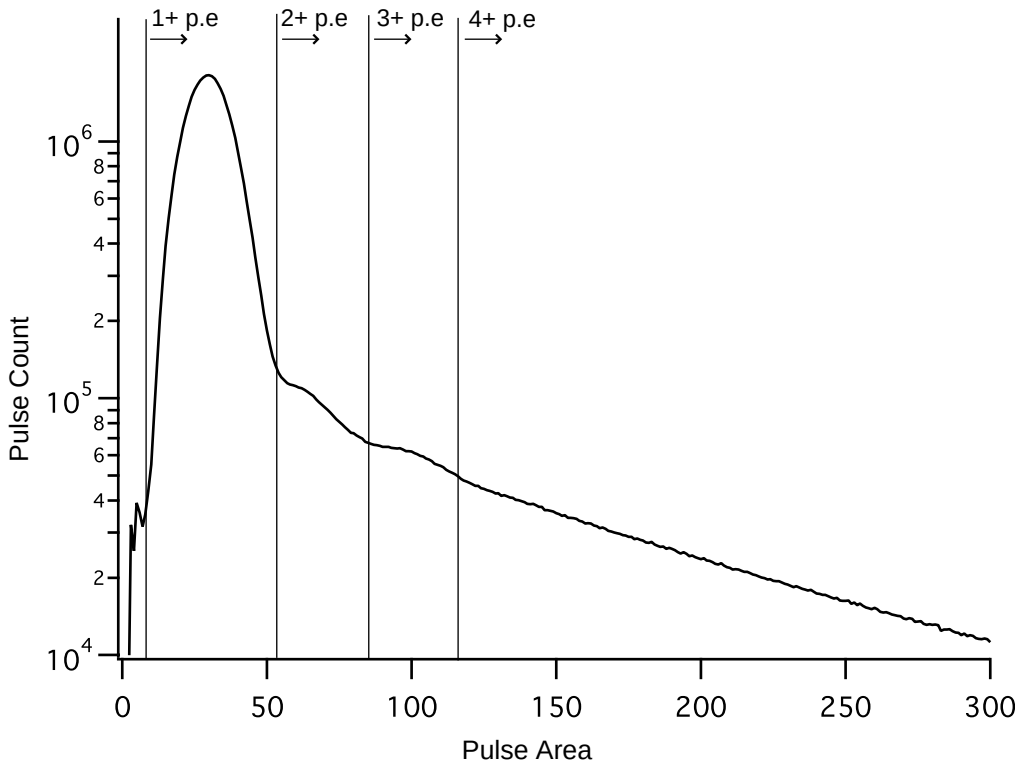

Figure 5.3: The pulse height spectrum of PMT β. The cutter thresholds for up to 4 photoelectrons are shown. The spectrum was produced by histogramming the PMT pulse areas calculated by the data analyzer from all 358 positive runs in Cold1 data set.

means at least three photoelectrons in each PMT is required for the data event. The cut thresholds up to 4 photoelectrons are indicated in Figure 5.3. The thresholds for higher photoelectron cuts such as "5+5+" are extrapolated assuming the higher photoelectron peak width is the same as the two and three photoelectron peaks.

After the data files are processed by the analyzer and the cutter, data events from the same data set can be time histogrammed together. The histograms of average positive and negative runs are subtracted to form the signal decay curve from trapped neutrons. The decay curves are then fitted to a single exponential decay model,

$y + Ae^{-t/\tau}$ using Igor Pro fitting routines to extract the neutron trap lifetime.

5.4 Backgrounds

Majority of the data recorded by the detection system are background events. Although the background subtraction technique does not require us to know the exact nature or form of the background, a clear understanding of the background can help us determine the data cut thresholds, evaluate the effectiveness of the background subtraction technique and point out ways to improve the background suppression in the future. The background can be approximately divided into a "constant" component which is caused by external radiation sources and does not show large time variation during one data run and a "time-varying" component which is induced by the neutron beam entering the cell and has a clear time dependent decay structure during one data run.

5.4.1 Constant Background

The constant background is caused by ambient ionizing radiation such as muons, gammas and neutrons interacting with liquid helium or acrylic lightguides in the apparatus. As discussed in Section 4.4.3, muon veto paddles, lead shielding and neutron absorbing materials surrounding the apparatus were used to suppressed the constant background. After the suppression, roughly 20 Hz single photoelectron coincidence rate remains[3]. At higher cut thresholds, the background rate goes down but does not go away (see Table 5.2).

[3]The single photoelectron count rate on each PMT is roughly 1 kHz, largely due to dark counts.

Table 5.2: A comparison of the constant background between reactor on and reactor off at different cut thresholds.

Cut Threshold	Reactor on (s^{-1})	Reactor off (s^{-1})
2+2+	~ 13.5	~ 3.0
3+3+	~ 10.5	~ 2.3
5+5+	~ 6.8	~ 1.6
10+10+	~ 3.3	~ 1.0

The constant background is highly correlated with the operation of the reactor. During reactor shutdowns, the constant background rate is reduced by a factor of 4.5 (see Table 5.2). This suggests that approximately 80% of background originate from ambient gamma or neutron field inside the guide hall, and it could potentially be reduced with better shielding of the apparatus. The background remaining after the reactor shutdown can be mostly attributed to naturally radioactive materials, such as Uranium, Thorium and Potassium, inside the cryostat and the lead shielding. This portion of the background could only be reduced by selecting purer material for the dewar and shielding construction.

The fraction of background events originates in the acrylic lightguides can be determined by removing the liquid helium from the cell. At low cut thresholds, more background events are seen in the lightguides due to their larger total volume and higher total mass. Because liquid helium is a much more efficient scintillator than acrylic in detecting ionization radiation, the fraction of background events inside the lightguides goes down with higher cut thresholds. At "3+3+" cut threshold, approximately 50% of background comes form the acrylic. If a low temperature detector system could be developed, acrylic lightguides could be completely eliminated and the constant background would be cut in half.

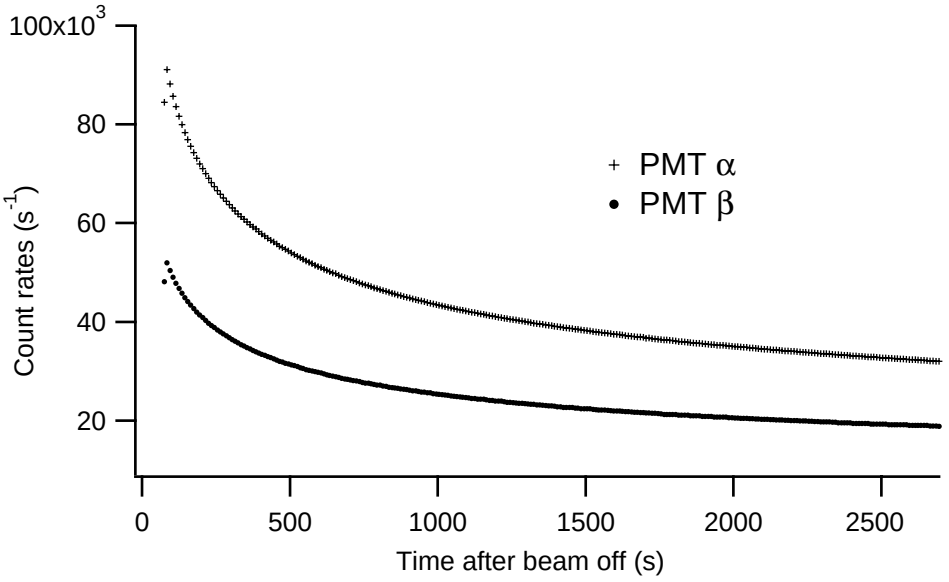

Figure 5.4: The registered single photoelectron count rates of both PMTs during the observation period. The plot is the average of 100 negative runs from "Cold2" data set.

5.4.2 Time Dependent Background

The time varying background is caused by neutrons-induced luminescence and activations. To reduce this background, materials which were exposed to the neutron beam were carefully selected to have extremely low impurities and minimal luminescence. The boron nitride neutron shielding known to emit large number of luminescence photons were all covered by thin graphite pieces to block the luminescence photons (see Section 4.4.3). Even after the implementation of these background suppression techniques, a fairly large time-varying background could be observed.

First, a large number of luminescence photons could been seen after the cell was exposed to the neutron beam. As shown in Figure 5.4, the recorded single photoelectron count rate[4] on each PMT had an initial amplitude of 60 – 100 kHz, significantly

[4]The single photoelectron count rate here means the total count rates above the 1 photoelectron

higher than the 1 – 2 kHz count rate observed before the neutron beam exposure. The singles count rate decayed slowly, and at the end of the observation period, about 20 - 40 kHz count rate still remained. The decay curves fit well to a power law decay model, in agreement with previous studies on neutron induced luminescence [56].

Because over 99.9% of neutron induced-luminescence are uncorrelated single photons, the luminescence background can be effectively suppressed by requiring a coincidence between the two PMTs. The accidental one photoelectron coincidence rate from luminescence can be estimated using the formula $R_\alpha R_\beta t_{coin}$, where R_α and R_β are the PMT single photoelectron count rate, and t_{coin} is the width of coincidence window. For a coincidence window of 35 ns, roughly a factor of 1000 reduction in luminescence background can be achieved. Figure 5.5 shows both the measured single photoelectron coincidence trigger rate ("1+1+" cut) and the calculated accidental coincidence trigger from luminescence background. At the "1+1+" cut level, approximately 50% of the background can be accounted for by accidental coincidences from the luminescence.

The luminescence background can be further reduced by setting higher cut threshold on PMT pulse areas, because the probability of the luminescence light which consists of single photons generating a large PMT pulse is fairly small. For the "2+2+" cut level (requiring more than two photoelectrons in each PMT), the accidental coincidence rate from luminescence photons can be reduced by a factor of 400. And for the "3+3+" cut level, the reduction factor is over 10^5. In practice, the luminescence background should be totally suppressed by cut thresholds higher than "3+3+".

The second source of time dependent background is neutron activation. The decay

cut threshold.

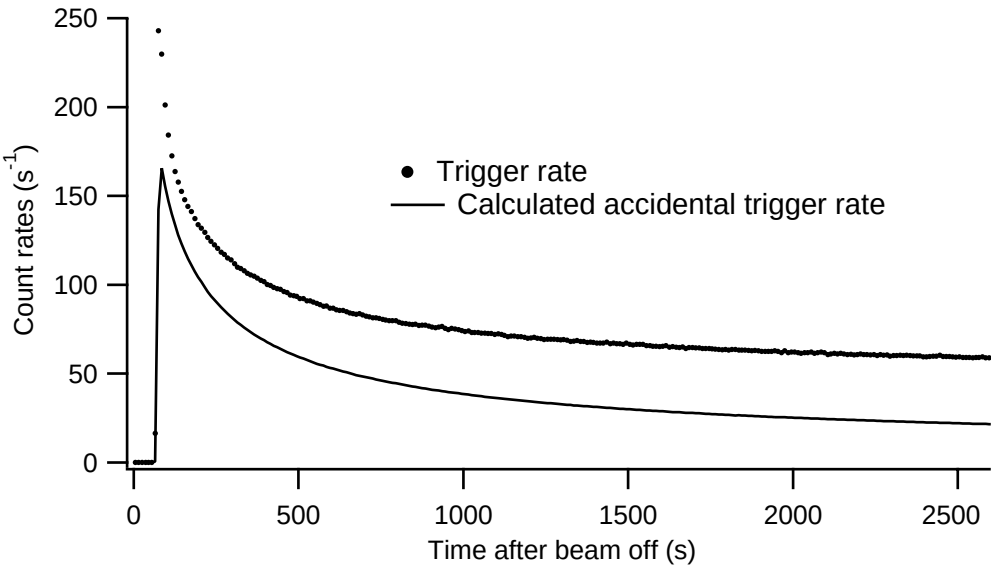

Figure 5.5: The one photoelectron coincidence trigger rate in the negative runs and the calculated accidental coincidence rate from luminescence background are shown. The plot is the average of 100 negative runs from "Cold2" data set.

gammas or charge particles from activated isotopes can produced a large number of scintillation photons inside the liquid helium or acrylic light guides. Unlike luminescence, the activation background remains at high cut thresholds. Figure 5.6 shows the PMT coincidence rate at various cut thresholds from the negative runs. A time dependent component can be clearly seen in these background data. If we assume that the time dependent signal consists entirely of neutron activations at high cut thresholds, we can fit the decay curve to a constant plus several exponential decay curves. The data do not fit well to either a single or double exponential model, but fits to a triple exponential decay model yield very good results. For example, the fit

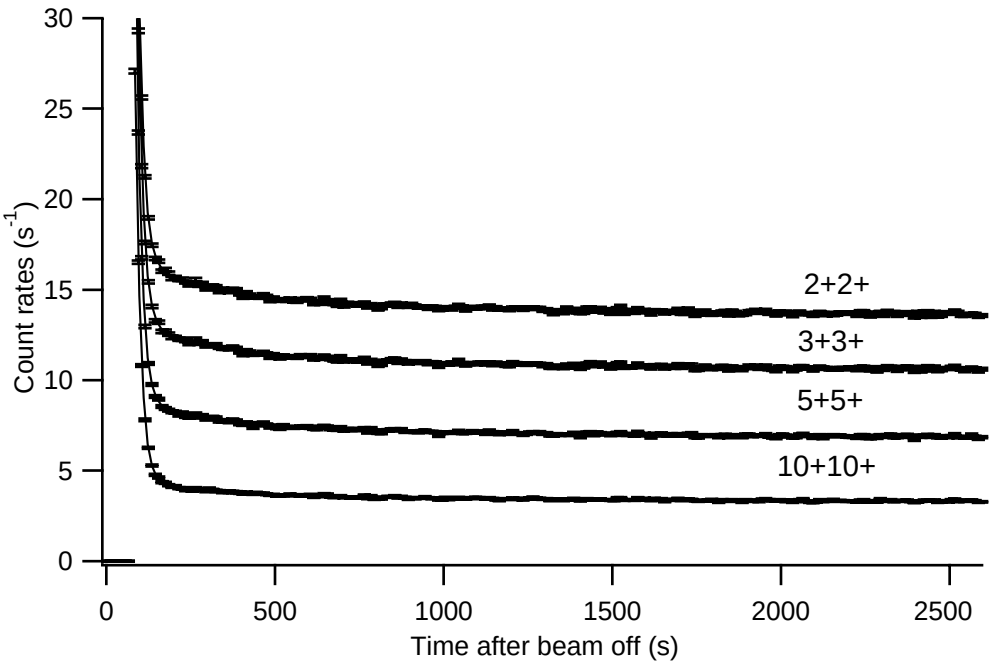

Figure 5.6: The PMT coincidence count rates from negative runs at different high cut thresholds. Both the constant background and the time dependent background from neutron activations can be seen. The plot is the average of 100 negative runs from "Cold2" data set.

to the "2+2+" data yields,

$$\begin{aligned} R_{2+2+} &= (13.5 \pm 0.1) + (6390 \pm 384)\, e^{-t/(16.0\pm0.2)} \\ &\quad +(3.2 \pm 0.2)\, e^{-t/(209+36-27)} + (1.1 \pm 0.2)\, e^{-t/(1218+845-354)} \end{aligned} \tag{5.1}$$

with $\chi^2 = 1.1$. The first exponential with a decay lifetime of 16.0 ± 0.2 s and a large amplitude is likely due to the activation of fluorine, which has a lifetime of 15.9 s. Fluorine exists in neutron entrance windows and Goretex reflector which were directly exposed to the neutron beam. The second exponential is consistent with the activation aluminum, which has a decay lifetime of 195 s. Most of the dewar and support structures is made of aluminum. Though aluminium is shielded from direct

Table 5.3: The fit results of the background count rate at different cut thresholds to the function $y + A_1 \exp(-t/\tau_F) + A_2 \exp(-t/\tau_{Al}) + A_3 \exp(-t/\tau_3)$. τ_F and τ_{Al} are the known lifetime of fluorine and aluminum. The data is the average of 100 negative runs in the "Cold2" data set and binned in 5 s bins.

Cut Thresholds	y (s^{-1})	A_1, (s^{-1})	A_2, (s^{-1})	A_3, (s^{-1})	τ_3 (s)
2+2+	13.53 ± 0.05	6609 ± 36	3.3 ± 0.2	1.20 ± 0.06	1069^{+194}_{-142}
3+3+	10.52 ± 0.04	6379 ± 30	2.7 ± 0.2	1.07 ± 0.05	1017^{+169}_{-126}
5+5+	6.80 ± 0.03	5872 ± 27	2.0 ± 0.1	0.84 ± 0.04	1043^{+184}_{-136}
10+10+	3.26 ± 0.03	4731 ± 22	1.25 ± 0.09	0.48 ± 0.02	1174^{+288}_{-193}

exposure to the neutron beam, low level of activation from scattered neutrons can be expected. The fit value for the decay lifetime of the third exponential has a very large error bar. To reduce this error bar, we can fix the first two exponentials to the known decay lifetimes of fluorine and aluminum in the fit. The fit results for the background curves at different cut thresholds are shown in Table 5.3. A likely candidate for the third exponential component is ^{13}N, which could be produced by secondary activation of ^{10}B in the beam stop[5]. ^{13}N has a lifetime of 863 s, in agreement with the fit results within 2σ. It is also likely that more than three isotopes contributed to the activation background, but a four exponential model has too many free parameters to produce a good fit.

In summary, the time dependent background from luminescence can be totally suppressed by setting high cut thresholds on PMT pulse areas. The time dependent background from activation consists of exponential decay signals of several isotopes. Our fit results identified F, Al and ^{13}N as potential sources. Activation signals from other isotopes may also present, but can not be identified from the fitting procedure.

[5]The reactions that produce ^{13}N is ^{10}B(n,α)^{7}Li, then ^{10}B(α,n)^{13}N.

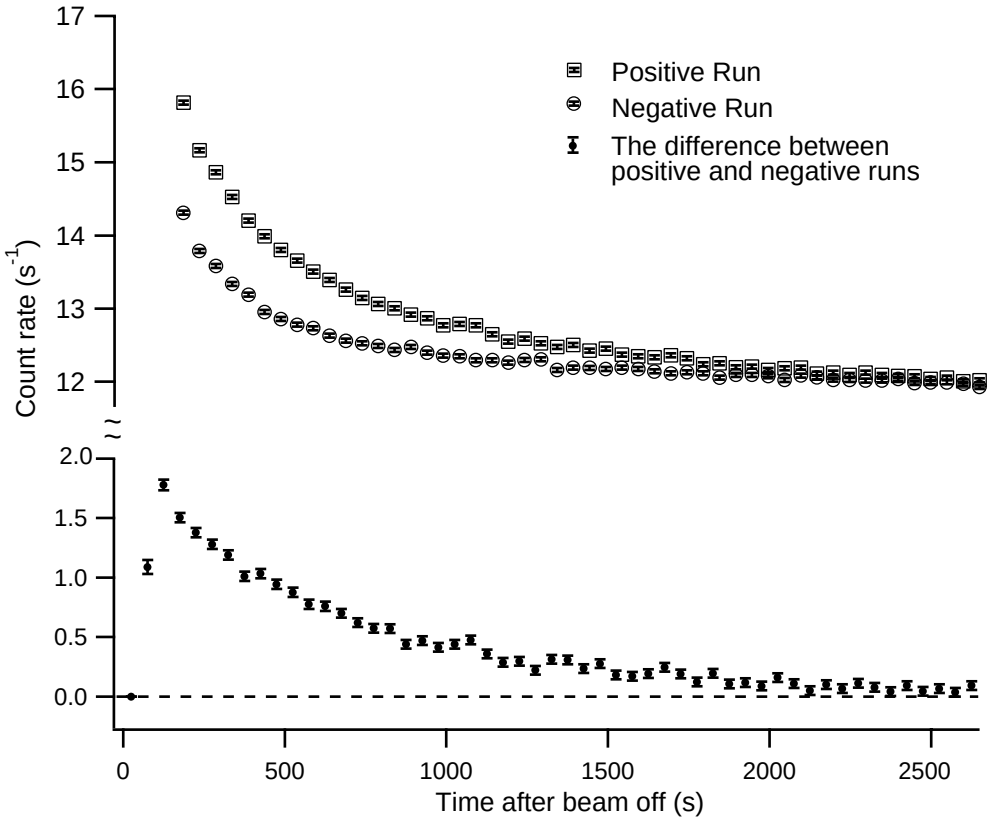

Figure 5.7: The average positive and negative runs from data set "Cold1". The data is cut at "3+3+" level, and binned in 50 s bins. Using the background subtraction technique, the decay signals from trapped neutrons can be extracted by taking the difference of the two curves.

5.4.3 Background subtraction

Due to the presence of both large constant background and a complicated time dependent background, the best strategy to extract the neutron decay signals is to measure the background and subtract it from the data. As mentioned before, two types of runs are taken, positive and negative. The difference between the two types of runs is that the magnetic field is off during the trap loading period of the negative runs. Since no neutrons are trapped, the negative runs measure the background.

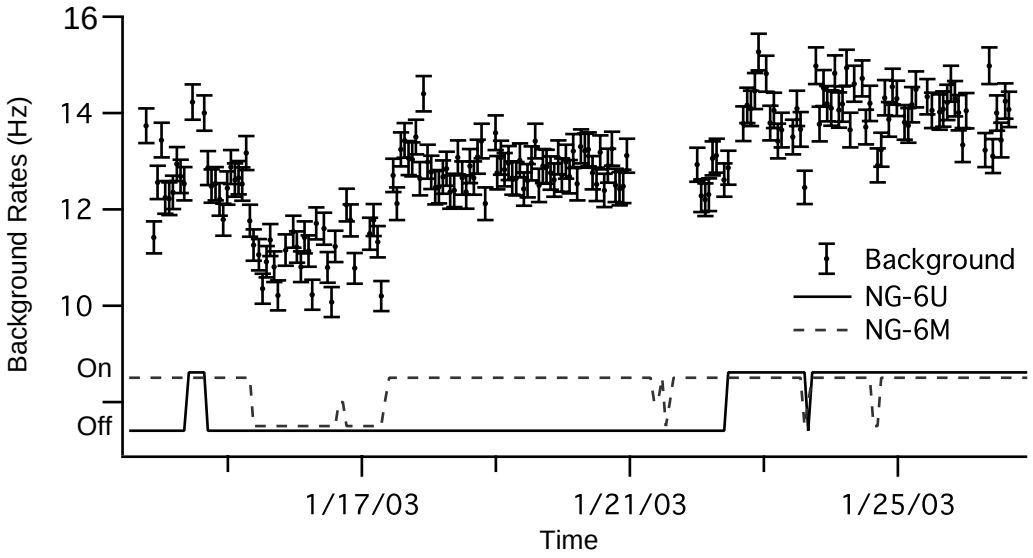

Figure 5.8: An illustration of the correlation between the constant background and the operation status of two nearby neutron beams, NG-6U and NG-6M. The background rates are estimated by the average count rate at the last 100 s of each data run at "3+3+" photoelectron cut level.

The neutron decay signals can be extracted from the difference between the positive and negative runs. Figure 5.7 shows the average positive and negative runs for the "Cold1" data set, and its difference at the "3+3+" level. The precision to which one can measure the decay lifetime of the trapped neutrons depends on how well this background subtraction works. Here we examine the subtraction for both the constant and time-varying backgrounds.

If the constant background was truly constant or varies very slowly with time, it could be canceled out with the background subtraction technique, and would not introduce any systematic uncertainty. The problem arises when there are sudden shifts in the constant background, which are not canceled by the background subtraction. In practice, sudden and large shifts of the constant background were indeed observed.

For example, the constant background could shift when the operating status of two nearby neutron beams, NG-6 and NG-6M changed. In Figure 5.8, the time variation of the constant background rate[6] is plotted together with the operating status of NG-6 and NG-6M. We can see that shifts in the background rates are clearly correlated with the operation of the two beamlines. To cancel these shifts in background, we continuously monitor the operating status of the two beamlines and match the positive runs to negative runs with the same NG-6 and NG-6M operating conditions for better background subtraction. The operating status of other instruments also has an effect on the constant background but on a much lower level, we consider them as a slow varying part of the constant background. In addition, we installed a NaI detector to monitor the constant background. If sudden background shifts were observed during a data run, this run would be thrown out.

The perfect subtraction of time dependent background is even more crucial, because residual time dependent component can cause a systematic shift in the lifetime measurement. In previous work, we have learned that luminescence centers can be partially stabilized by magnetic fields [56]. Because the magnetic field is off during the loading period for the negative runs, the luminescence signals in the positive and negative runs are significantly different. Fortunately, as discussed in Section 5.4.2, luminescence background can be suppressed by setting high cuts on PMT pulse areas. A "3+3+" cut can usually cut the luminescence background level below 10^{-3} s^{-1}. The activation signals from the positive and negative runs can also be slightly differ-

[6]The constant background rates are estimated by averaging the count rate during last 100 s of each data run at "3+3+" photoelectron cut threshold. Because majority of the time varying component of the background or trapped neutron signals have decayed away, the average rate from the last 100 s of each run is a fairly good estimate of the constant background.

ent because the magnetic focusing effect can cause the neutron beam to hit slightly different parts of the cell. The estimated focusing effect can change the distribution of the neutron beam by up to 1%. If the activated isotopes were uniformly distributed, one would expect that at least 99% of the time dependent background can be canceled out by the background subtraction. If we reduce the amplitudes of the "3+3+" fit in Table 5.3 by 100, and add them to the observed neutron decay signals (see Section 5.5), we estimate that the residual time dependent background can shift the neutron lifetime by up to 10 s. The effectiveness of the time dependent background subtraction can be measured directly by replacing the isotopically pure helium with natural helium. Such a measurement is discussed in Section 5.5.

5.5 Neutron trapping Data

The decay signals from trapped neutrons can be extracted by taking the difference between the positive and negative runs. A cut threshold requiring 3 photoelectrons in each PMT ("3+3+") can suppress luminescence background, and therefore ensure proper subtraction of the background. The difference count rates between the average positive and negative runs for 300 mK runs, magnet ramp runs and ^{3}He runs are displayed in Figure 5.9 . The fit results of the count rates to a single exponential decay model, $y + A\exp(-t/\tau)$, are summarized in Table 5.4. Implications of the data are discussed below.

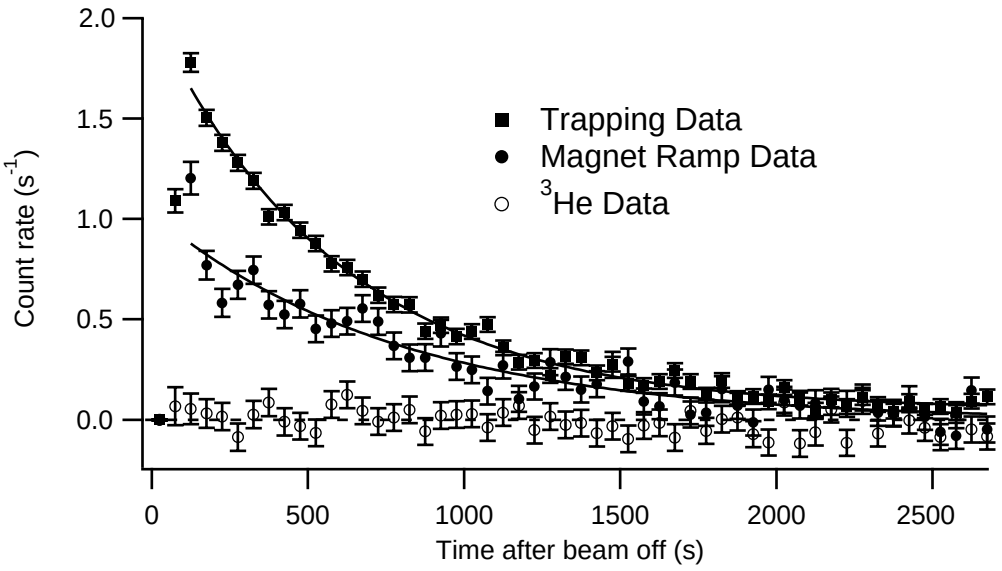

Figure 5.9: The difference count rates for the "Cold1", "MangetRamp" and "^{3}He" data sets are shown. The cut threshold is set to require at least 3 photoelectrons in each PMT. The data is binned in 50 s wide bins, and the best fit curves of the "Cold1" and "MagnetRamp" data to a single exponential model are also shown.

5.5.1 Evidence of neutron trapping

In order to demonstrate that the difference count rate between the positive and negative runs are indeed due to trapped neutrons, we can compare the 300 mK runs which had isotopically pure helium in the cell, and the ^{3}He runs which had natural helium in the cell. The trace amount ^{3}He inside natural helium has minimal effects on either luminescence or activation background, but it is sufficient to absorb all trapped neutrons within 1 s. The ^{3}He runs therefore measures any imperfect background subtraction between the positive and negative runs. The difference between the 300 mK runs and the ^{3}He runs should be entirely due to trapped neutrons.

Both the 300 mK runs and ^{3}He runs can be fitted to a single exponential decay model, see Table 5.4. While a decay signal with amplitude of 1.94 ± 0.03 s^{-1} is

Data Set	y, s^{-1}	A, s^{-1}	τ, s	χ^2
Cold1	0.04 ± 0.01	1.91 ± 0.04	$625.3^{+23.8}_{-22.1}$	1.02
Cold2	0.03 ± 0.01	1.96 ± 0.07	$586.1^{+25.8}_{-23.7}$	1.02
Cold3	0.03 ± 0.03	1.97 ± 0.06	$676.2^{+48.0}_{-42.0}$	0.98
All Cold data combined	0.04 ± 0.01	1.94 ± 0.03	$621.2^{+18.1}_{-17.1}$	0.96
MagnetRamp	-0.01 ± 0.04	1.02 ± 0.07	$817.3^{+150.4}_{-109.9}$	1.08
MagnetRamp	fixed to 0	0.95 ± 0.06	$831.1^{+58.2}_{-51.0}$	1.09
^{3}He Cold	0.1 ± 0.9	-0.1 ± 0.9	-4420^{+3910}_{-5080}	1.08
^{3}He Warm	0.1 ± 0.6	-0.1 ± 0.5	-2676^{+2860}_{-2520}	0.99
^{3}He Combined	0.04 ± 0.06	0.08 ± 0.05	1134^{+2100}_{-780}	1.02
^{3}He Combined	-0.01 ± 0.01	fixed to 0		1.02

Table 5.4: The fit results of the difference count rates for the 300 mK , MangetRamp and ^{3}He data sets to a single exponential decay model $y + A\exp(-t/\tau)$. The data is binned in 5 s bins. The range of the fit is $[150, 2680]$ seconds for the 300 mK and ^{3}He data set, and $[150, 2680]$ seconds for the magnet ramp data.

clearly seen in the 300 mK, the amplitude of the exponential fit for the ^{3}He runs $A = -0.1 \pm 0.9$ s^{-1} is consistent with zero. The difference between the two data sets is statistically significant at 20σ level. The fact that no decay signals are seen in the ^{3}He data set provides a definitive proof of trapped neutrons.

To reduce the systematic uncertainty related to imperfect background subtraction, the ^{3}He data can be directly subtracted from the trapping data. With this method, statistical precision is traded for reduced systematic uncertainty. Because the ^{3}He data at 300 mK and 1K, and from the combined data set are all consistent with zero, the fit results of the trapping data should not change after the subtraction. Furthermore, the statistical precision at the current measurement level is very limited. We

therefore chose not to perform the subtraction to extract the trap lifetime. Instead, we will estimate the possible systematic shifts from imperfect background subtraction based on the ^{3}He data, see Section 5.6.2.

5.5.2 Trap Lifetime at 300 mK

The neutron trapping data at 300 mK were taken with three slightly different time schedules. Runs from the "Cold1" data sets have the standard time schedule with 2100 s loading period and 2700 s observation period. Runs from the "Cold2" data set has a lengthened observation period (5400 s), so that the offset parameter y in the single exponential fit can be better determined. Runs from the "Cold3" data set has a delay period of 2700 s between each run, during which both the PMTs and the digitization cards were turned off. These runs were used to investigate possible systematics related to the operation of DAQ cards and the PMTs. The fit results of all three data sets shown in Table 5.4 are consistent with each other within statistical uncertainties. Thus, we can combine the 300 mK data sets together to reduce the statistical uncertainty in the fit.

The combined data sets has a fit lifetime of $621.2^{+18.1}_{-17.1}$ s, which is shorter than the accepted neutron beta decay lifetime by 15σ. This suggests that there exist neutron trap loss mechanisms other than beta decay. Furthermore, the offset parameter y fits to a positive value $y = 0.04 \pm 0.01$ s^{-1}. This suggests that the decay curve may not be a pure single exponential, but contains several exponential or non-exponential decay components.

The observed short trap lifetime and positive offset in the 300 mK data can be

explained by the presence of above threshold neutrons. As discussed in Chapter 3, a large number of above threshold neutrons can stay in the trap for hundreds of seconds due to a combination of marginal trapping and material bottling. The beta decay signals from these neutrons add a non-exponential component to the decay curve, therefore making the trap lifetime appear short.

5.5.3 Trap lifetime after field ramp

The magnet ramp runs were designed to eliminate the above threshold neutrons. As described in Chapter 4, majority of the above threshold neutrons can be ejected from the trap, if the trap field is lower briefly to 30% of its maximum value after the loading period.

The data from the magnet ramp runs, if fitted to a single exponential model, yield a much longer trap lifetime, $\tau = 817.3^{+150.4}_{-109.9}$ s. The constant offset $y = (-0.01 \pm 0.04)$ s^{-1} is consistent with zero. If the offset y is fixed to zero for the fit, the uncertainty of the lifetime can be considerably reduced, yielding $\tau = 831.1^{+58.2}_{-51.0}$ s.

The results from the magnet ramp data is consistent with the accepted value of the neutron lifetime, which confirms that above threshold neutrons were indeed the main cause of the observed short trap lifetime at 300 mK. Since most of the above threshold neutrons and a fraction of trapped neutrons were thrown away during the field ramp, the amplitude of the difference count rate in the magnet ramp runs were reduced by a factor of two comparing to the 300 mK runs. Because the figure of merit for the statistical precision of the experiment is $Signal^2/Background$, four times more magnet ramp data would be need to reach the same statistical precision

as the 300 mK data. This means that a couple of years of running time would be required for a lifetime measurement with uncertainty of ten to fifteen seconds. It was determined that the experimental resource would be better spent in the development of the next generation apparatus which can significantly increase the signal level, and statistical precision of 3 s can be achieved in one month, (see Chapter 6). The data taken for the magnet ramp runs was thus stopped when we determined that we gained sufficient understanding of the current system.

The UCN production rate can be estimated from the initial decay rate of the trapped neutrons. The magnet ramp data shows an initial decay rate of 1 s^{-1}. The actual decay rate should be higher because some trapped neutrons were thrown away during the field ramp. However, the decay rate from trapped neutrons could not be higher than the initial decay rate of 2 s^{-1} measured in the 300 mK data. Thus, we estimate the initial decay rate from trapped neutrons to be 1.5 ± 0.5 s^{-1}. Taking into account of the 48% detection efficiency at the "3+3+" cut threshold and the finite length of trap loading, we predict a UCN production rate of 3.4±1.1 UCN s^{-1}. Monte-Carlo simulation of the superthermal production of UCN predicts a UCN production rate of 6.8 ± 1.4 UCN s^{-1} [54]. The two predictions agree with each at the 2σ level.

5.5.4 Temperature dependence of the trap lifetime

The rate of neutron trap loss due to two-phonon upscattering has a theoretically predicted T^7 temperature dependence [98]. The prediction is experimentally verified for temperatures down to 750 mK [52]. To observe the effect of phonon upscattering on the neutron trap lifetime, trapping data were taken at several different helium bath

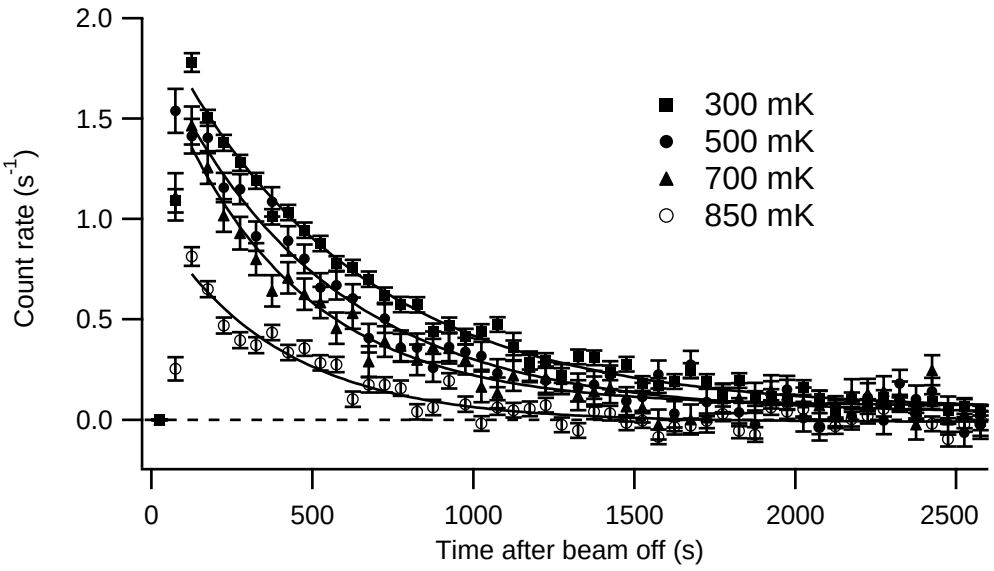

Figure 5.10: The difference countrates for the "Cold1", "500mK", "700mK" and "850mK" data sets are shown. The cut threshold is set to require at least 3 photoelectrons in each PMT. The data is binned in 50 s wide bins, and the best fit curves of the data to a single exponential model are also shown.

temperatures. Figure 5.10 shows the difference count rate from the neutron trapping data taken at 300 mK, 500 mK, 700 mK and 850 mK. The fit results of the data to a single exponential are summarized in Table 5.5.

Because the effect of phonon upscattering at 300 mK is supposed to be negligible [98], if we assume that the shortened trap lifetime at higher temperature was entirely due to phonon upscattering, the trap loss time due to phonon upscatteing τ_{upsc} can be calculated with the following formula,

$$\tau_{upsc}^{-1}(T) = \tau_T^{-1} - \tau_{0.3}^{-1}. \tag{5.2}$$

The measured upscattering loss times are shown together with previous measurements by Golub [52] and the calculated theoretical curve using Landau's Hamiltonian [98, 52] in Figure 5.11. The data agree with the theory well for temperatures below 800 mK.

Data Set	y, s^{-1}	A, s^{-1}	τ, s	χ^2
300 mK	0.04 ± 0.01	1.94 ± 0.03	$621.2^{+18.1}_{-17.1}$	0.96
500 mK	0.03 ± 0.02	1.85 ± 0.09	$518.1^{+39.6}_{-34.3}$	1.03
700 mK	0.06 ± 0.02	1.65 ± 0.11	$437.4^{+42.8}_{-35.8}$	0.98
850 mK	-0.01 ± 0.01	0.93 ± 0.09	$412.5^{+53.3}_{-42.4}$	0.96

Table 5.5: The fit results of the difference count rates at different cell temperatures. The data are are binned in 5 s bins, and fitted to a single exponential decay model $y + A\exp(-t/\tau)$. The ranges of the fits are $[150, 2680]$ seconds.

For temperatures above 800 mK, both Golub's measurements and our data show systematically higher upscattering lifetimes than the theoretical predictions. One possible explanation is that the ultrapure helium used in both measurements were not entirely void of ^{3}He. At temperatures above 800 mK, the residual ^{3}He atoms were pushed out of the cell by applied heat currents. A decrease in ^{3}He absorption rate would appear as an increase in the measured photon upscattering lifetimes.

5.6 Uncertainties

5.6.1 Statistical Uncertainty

The small number of trapped neutrons and high background rates limit the statistical precision of the experiment. The lifetime value extracted from the magnet ramp data has a statistical uncertainty of (+150/-101) s (15%). If we fix the constant offset in the fit to zero, the uncertainty can be reduced to (+58/-51) s (6%). Since the background subtraction in principle should yield zero constant offset, and we have taken measures to eliminate large fluctuations in the constant backgrounds, we believe that

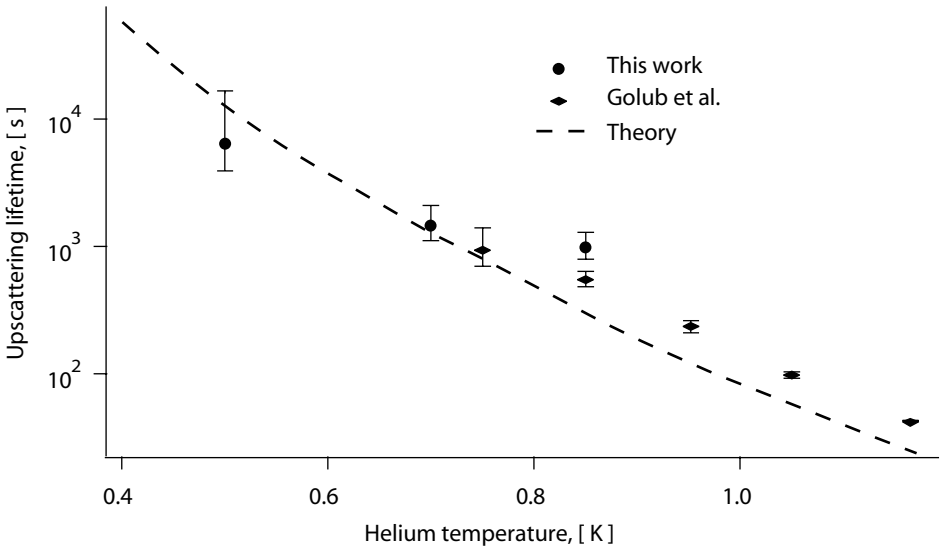

Figure 5.11: The temperature dependance of the neutron loss time due to phonon upscattering is shown. Previous measurements by Golub [52] and the theoretical prediction curve are also shown.

fixing the offset to zero is probably justified, though a more cautious approach would call for the use of the larger error bars. In the future, as the signal to background ratio increases, the difference between the two kinds of fits would disappear.

5.6.2 Systematic Uncertainties

The systematic uncertainty is currently limited by the isotopic purity of the ultra-pure helium and the presence of above threshold neutrons. In Table 5.6 , we list all known systematic effects in the current setup, and the achievable limits in the future. In the following section, we discuss each systematic effect separately.

Table 5.6: The estimated uncertainties due to systematic effects in the current measurement and the achievable limit for future measurements.

Systematic effect	Current setup (s)	Achievable limit (s)
Neutron absorption on ^{3}He	~ 70	< 0.01
Above threshold neutrons	~ 50	< 0.01
Imperfect background subtraction	~ 20	< 0.01
Detection efficiency drifts	8	< 0.01
Phonon upscattering	1	< 0.01
Neutron depolarization	< 0.1	< 0.01
Magnetic field fluctuations	< 0.1	< 0.01
Medium (helium) effects	< 0.01	< 0.01
Total uncertainty	~ 88	~ 0.01

Neutron absorption on ^{3}He

The presence of ^{3}He impurities inside the helium bath can absorb trapped neutrons, therefore shortening the neutron trap lifetime. As we have discussed in detail in Chapter 2, the large uncertainty on this systematic effect is due to the difficulty in measuring the concentration of trace amount of ^{3}He. The preliminary results we obtained from accelerator mass spectroscopy measurement gave a ^{3}He concentration of $(4.2 \pm 1.5) \times 10^{-12}$, which can result in 70 second shift in the lifetime measurement. However, as we have pointed out in Chapter 2, we can not yet rule out the possibility that this high concentration was due to contamination from background gases. With the development of better measurement techniques, we expect to significantly increase the precision of this measurement. Since there is no theoretical limit to the purity of helium that can be achieved with the heat flush method, we can always build a

purifier to further purify the helium before putting them into the experimental cell. We therefore believe that an ultimate ^{3}He concentration below 10^{-16} is attainable.

Above threshold neutrons

As discussed in Chapter 3, the presence of above threshold neutrons can make the neutron trap lifetime appear short. Majority of the above threshold neutrons can be eliminated by ramping the field down to 30%, as is confirmed by our 300 mK and the magnet ramp data. To estimate the fraction of above threshold neutrons in the current setup, we notice that the 300 mK data can be approximated by the sum of two exponentials, one with an amplitude of 0.8 s^{-1} and a lifetime of 400 s, and the other with an amplitude of 1.3 s^{-1} and a lifetime of 885 s. The component with the short lifetime can be identified with above threshold neutrons. The simulation in Section 3.2.5 shows that after the field ramp, more than 85% of the above threshold neutrons would be eliminated. Using this result, we estimate that the uncertainty introduced by the above threshold neutrons to be approximately 50 s after the field ramp. It should be noted that this is a rather crude way of estimating the uncertainty. Fortunately, in the future, with the implementation of the KEK trap, the above threshold neutrons can be completely eliminated by ramping the field down to 22% of the trap depth and reducing the effective material wall potential to 15 neV (see Section 3.2.6).

Imperfect background subtraction

The uncertainty introduced by imperfect background subtraction can be estimated from the ^{3}He data. Because any residual background after the subtraction manifests

itself in the ^{3}He data, we can fit ^{3}He data to the expected residual background model, and extract an upper limit on the background. For example, fitting the data to a single exponential yields a decay amplitude of 0.08 ± 0.05. Using this as an upper limit for an exponential decay signal, we estimate that the residual background can shift the trap lifetime by up to 20 s. In future measurements, ^{3}He data can be measured more precisely, therefore subtracted directly from the trapping data. An even better approach would be to load the magnetic trap from an external UCN source. Then the number of neutrons entering the cell is reduced by a factor of 10^7, and the time dependent background would be reduced by the same factor to a negligible level. At that point, the background subtraction technique probably would not be necessary. The decay signals can be simply fitted to a single exponential plus a constant background.

Detection efficiency drifts

The gains of photomultiplier tubes can change due to exposure to intense light. A gain shift could also occur during PMT warmup period right after it is turned on. In the current setup, the PMTs were exposed to intense light caused by the neutron beam during the trap loading period, then were turned on for the observation periods. A LED was installed to measured the gain shifts of the PMT. A 2% gain shifts was observed for the current setup, resulting in approximately 8 seconds shift in the lifetime measurement. The systematic effect of the gain shifts can be overcome by installing a *in situ* calibration system to precisely measure the gain shifts during the observation period. It should be noted that the PMT gain shifts would be highly

reduced if the trap is loaded from an external UCN source. The extremely low neutron flux would no longer produce large amount light when entering the cell. The PMTs can thus be kept on throughout the experimentation time, and should exhibit much better gain stability.

Phonon upscattering

The neutron upscattering rate from phonons has a theoretically predicted T^7 temperature dependence. Our measurements down to 500 mK, agree with the theoretical curve, though the uncertainties on the measurements are rather large. Using the measured upscattering rate at 500 mK, we expect the systematic effect due to the phonon upscattering to shift the neutron trap lifetime by roughly 1 s at the cell temperature of 300 mK. In the future if the cell temperature is lowered to 150 mK, then the systematic shift of the trap lifetime due to phonon upscattering should be below 0.01 s.

Neutron depolarization

The depolarization of trapped neutrons could occur due to Majorana spin-flip transitions, or due to spin-orbit or spin-spin interaction with charged particles in liquid helium.

The probability of Majorana spin-flip transitions is large only in the low field regions where the rate of the field change experienced by the neutron is large. In the current trap which has a minimal field of 0.1 T, a neutron on a linear trajectory going through the trap center has a spin-flip probability of (8×10^{-8}) s^{-1}, which corresponds to 10^{-4} neutron beta decay rate [51]. Since neutrons on other trajectories would have

even less spin-flip probability, we can put the uncertainty introduced by Majorana spin-flip below 0.1 s. In the KEK trap, the minimum field is 0.6 T, so the systematic shifts caused by Majorana spin-flip is well below 0.01 s.

Ionizing radiation events inside the liquid helium bath can create many electron-ion pairs, most of which quickly recombine. Those electrons or ions that escape the recombination if pass quickly in the vicinity of the trapped neutrons can cause the neutrons to depolarize due to the magnetic field produced by moving charges, or magnetic moments. The neutron and electron spin-spin interaction has the largest cross-section $\sigma_{ss} \sim 10^{-18}$ cm^2 [56]. Assuming all energy from alpha particles produced by neutron absorption on boron is deposited inside the liquid helium bath, we estimate that the highest electron density inside the bath to be $n_e \sim 10^6$ cm^{-3}. The neutron trap loss time due to charge particles therefore can be estimated as $\tau_f = 1/(n\sigma_{ss}v) \sim 10^9$ s,[7] We see that the spin depolarization due to charge particles only affect the neutron trap life time on the 10^{-6} level.

Magnetic field fluctuations

As pointed out by Ref. [47], fluctuations in the magnetic field due to either current instability or vibration of the apparatus can change the potential energy of the neutron $E_p(r,t) = E_p(r,0)(1+\epsilon(t))$. If the correlation function $K(t)$ of the fluctuation ($\overline{\epsilon(t)\epsilon(t')} = \epsilon_0^2 K(t-t')$) is not a delta function, the variation of neutron energy $\overline{(\Delta E)^2}$ will be nonzero. The variation depends strongly on the nature of the correlation function. Based on the calculated results of several functional forms of $K(t)$ in Ref. [47], we estimate that the fluctuations affect the lifetime measurement at most on the 10^{-4}

[7]The relative velocity v between the neutrons and electrons is roughly 10 m/s.

level for the current setup. This effect can be easily suppressed with better vibration isolation of the apparatus and more stable power supply.

Medium (helium) effects

The neutron decay lifetime in helium is a little different from its lifetime in vacuum. The shift in lifetime can come from either changes in the phase space available for the decay, or change in the decay matrix elements. The energies of a neutron and its decay electron are slightly higher in helium by 2×10^{-8} eV and 1.3 eV respectively, while the energy of the decay proton is slightly lower by approximately 2 eV. The overall change in the available phase space for the decay is less than 10^{-5}, therefore should not change the neutron lifetime by more than 10^{-5}. The decay matrix element can only be modified by the strong force, up to $10 - 20\%$. Since the estimated fraction of time a neutron spends close enough to the helium nuclei to be influenced by the strong force is less than 10^{-15}, possible changes in the matrix element is on the order of 10^{-16}.

Chapter 6

Development of the Next Generation Apparatus

6.1 Introduction

As shown by the latest data collection runs, the statistical precision achievable with the current apparatus is very limited. After the field ramping to eliminate the above threshold neutron, only 2000 neutrons were trapped in each run. When the background is larger than the signal, as is the case with the current setup, the statistical figure of merit is $Singal^2/Background$, so the statistical precision improves quickly with increases in the signal strength, i.e. the number of the trapped neutrons.

The number of trapped neutrons in a magnetic trap is proportional to $B_T^{3/2}V_T$, where B_T is the trap depth and V_T is the trap volume. Both signal and background increase with V_T, so the gain in the statistical figure of merit is proportional to V_T. On the other hand, an increase in the trap depth increases the number of trapped

neutrons with no increase in the background, so the gain in statistics is proportional to B_T^3. Other than waiting for better neutron sources to come online, building a deeper magnetic trap is the best way to improve the statistical precision of the experiment. A deeper trap is also more effective for eliminating the systematic uncertainty related to above threshold neutrons, see Chapter 3. Base on these considerations, the development of the next generation apparatus centers on the design and testing of a new magnetic trap with a greater trap depth.

The new trap consists of a high-current accelerator type quadrupole magnet on loan from KEK laboratory in Japan and two low current solenoids. It has a designed trap depth of 3.1 T, and is capable of trapping twenty times more neutrons. The design and testing of this KEK magnetic trap[1] is discussed in Section 6.2. Due to its high operating current, heavy weight and large volume, the KEK trap presents us with serious cryogenic challenges. First, instead of traditional vapor cooled current leads, it is necessary to use high temperature superconducting (HTS) leads capable of carrying 3400 A into 4 K to reduce the liquid helium boil-off to an acceptable level. The testing of such HTS leads is described in Section 6.3. Secondly, cryogenic posts (Section 6.4) are developed to support the over 500 kg weight of the KEK trap, while adding minimal heat load to the liquid helium bath. Finally, the KEK trap no longer fits in the existing dewar. The preliminary design of a new dewar is discussed in Section 6.5.

[1]Since the high current quardrupole magnet is an accelerator type magnet on loan from KEK lab in Japan. For lack of a better name, we have referred the new trap as the KEK magnetic trap.

6.2 The KEK Magnetic Trap

It was realized early on that an accelerator type superconducting quadrupole magnet can be combined with two solenoids to form a deep Ioffe trap. This option was not fully pursued until recently due to resource constraint. In 1999, we obtained a quadrupole magnet originally used for electron beam focusing from the KEK laboratory in Japan. The KEK magnet was tested in conjunction with a quench protection circuit in 2000. No quench was observed for current up to 2800 A. The KEK magnet was retested in 2004 with an improved quench detection circuit and quench inducing heaters. The efficiency of the protection circuit was mapped out by heater induced quenches. Meanwhile, a conservative approach was adopted for the solenoid design. The overall success probability was increased at the expense of 15% loss in signal by eliminating the bucking coils from the solenoid assemblies. Furthermore, from the study of the interaction between the quadrupole and the solenoids, it was determined that the solenoid forms needed to be specially designed to hold the solenoid windings against additional Lorentz forces caused by the field from the KEK magnet. The solenoids were wound by American Magnetics Inc. in 2005. The whole trap assembly was tested in summer 2005. It reached 90% of its design values after two quenches. This new trap has a trap depth of 3.1 T, and trapping volume of 8 l. The design issues and testing results are discussed in the following sections.

6.2.1 The KEK Quadrupole Magnet

Large superconducting dipole and quadrupole magnets used in accelerator facilities are commonly wound with large cross-section cables operating at high current.

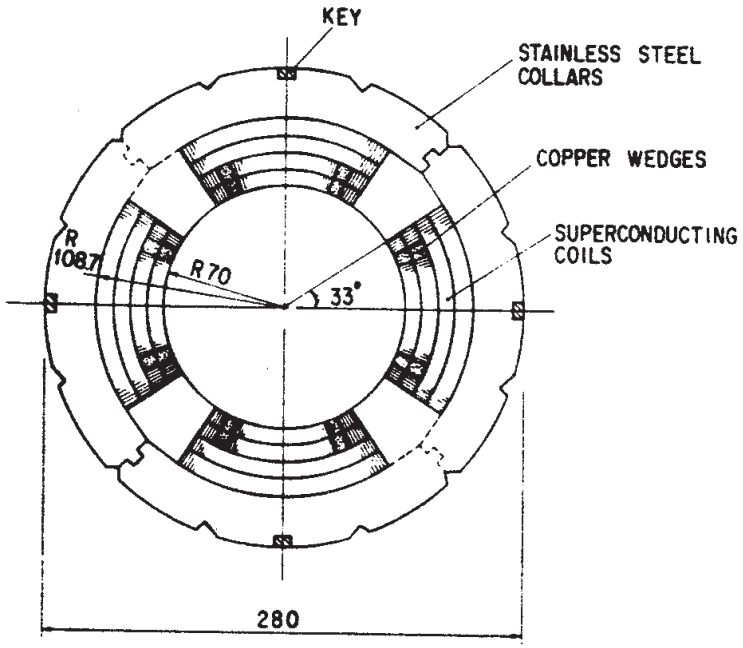

Figure 6.1: A cross-section view of the KEK quadrupole magnet.

The windings are prestressed by stainless steel collars on the outside. The main reasons for such a high current design are to limit quenching voltages to reasonable values [99] and to reduce the magnet inductance. For a low inductance magnet, the stored energy can be effectively dumped into an external resistor during a quench, therefore reducing the probability of a catastrophic quench.

The production of these accelerator type magnets is an industrial scale undertaking. Special winding machines, precision engineering, and sometimes manufacturing custom superconducting cables are necessary for the magnet fabrication. The cost is so high that it was unrealistic for our experiment to obtain a custom made quardupole magnet. Fortunately, we identified quadrupole magnets originally used

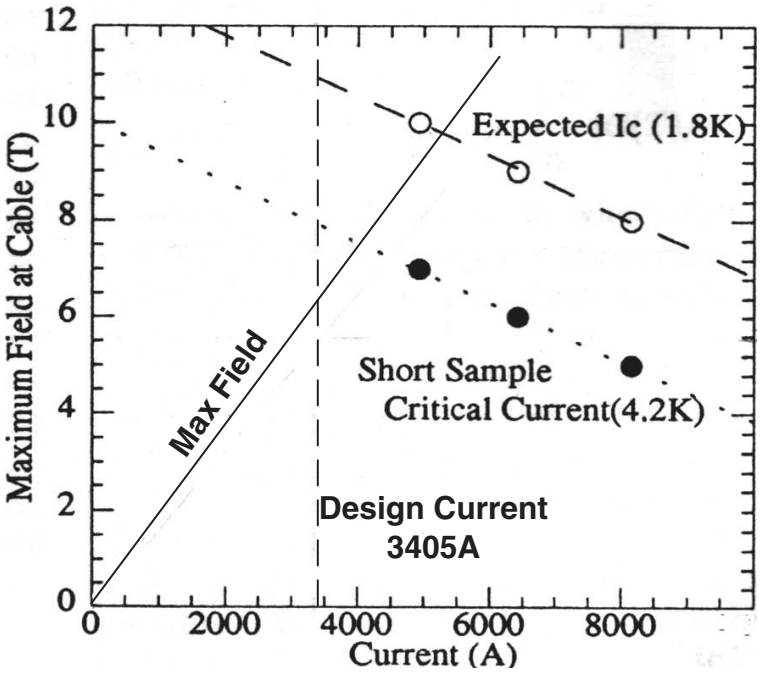

Figure 6.2: The short sample critical current of the KEK supercoducting cables at 4.2 K and 1.8 K, and the maximal field curve of the KEK magnet. The intersection of the two curves gives the critical current of the magnet. The operating current is about 85% of the critical current at 4.2 K.

for electron beam focusing at the TRISTAN accelerator at the KEK laboratory in Japan as suitable alternatives [90]. Since these magnets were not in use, the group at KEK generously loaned one magnet to us for testing.

The KEK magnet has an effective field length of 1.14 m, a bore size of 14 cm and a nominal OD of 28 cm. It consists of sixteen racetrack shaped coils, forming four concentric layers. The magnet was designed to operate at 3405 A (at 4.2 K) for a field gradient of 70 T/m with a maximum field of 4.9 T inside the magnet bore. The coils are prestressed by 30 mm thick 316LN stainless steel collars on the outside with typical radial forces of 6.9×10^5 N. A cross-section of the quardrupole magnet is shown in Figure 6.1.

The KEK magnet is wound with a keystoned cable with mean thickness of 1.27 mm and width of 9.09 mm, composed of 27 strand wires with 0.68 mm diameter. Each wire is made of approximately 2200 twisted NbTi filaments embedded in a copper matrix having a Cu:Sc ratio of 1.8. The wire strand packing factor is 0.89, which means 11% of the cable is expected to be filled by liquid helium. The direct contact of the superconductor with liquid helium provides added stability. The short sample critical current (loadline) of the cable as a function of external magnetic field at 4.2 K and at 1.8 K is shown in Figure 6.2. The maximum field in the KEK magnet is at the racetrack coil turnaround regions. The critical current of the magnet is determined by the intersection of the maximum field curve and the loadline, as shown in Figure 6.2. At 4.2 K, the critical current of the KEK magnet is about 4100 A. Notice that the critical current can be increased by roughly 30%, if the magnet operates in superfluid helium at 1.8 K.

During the initial training after its production, the KEK magnet successfully reached 4000 A after several quenches. In later test runs, the magnet was operated at 3405 A, about 85% of the loadline. No natural quenches occurred.

After obtaining the KEK magnet from Japan, we tested the magnet two times with a quench protection circuit built by our group to learn about the operation of a large magnet system as well as studying the characteristics of the protection circuit. Details of the circuits and tests will be discussed later.

6.2.2 Solenoid Design

Given the 14 cm bore size of the KEK magnet and an estimate of the space needed for implementing an experimental cell, we can expect to have an effective trapping radium of roughly 5.5 cm. At its operating current, the KEK magnet can provide 3.85 T radial trapping field. To first order, the pinch solenoids should produce an axial trapping field that matches the radial field. The parameters of the solenoids are then adjusted to maximize the total number of trapped neutrons N_n, where $N_n = k\int(B_{trap} - B)^{3/2}dV$. k is a constant related to the UCN production rate. The magnetic field B in the trap is calculated using the Biot-Savart program. B_{trap} is defined as the lowest field at which neutrons can escape the trap. This escape point is usually on the trap wall where the fringing field from a solenoid cancels the field from the quadrupole. The integral run over the volume inside the equipotential surface of B_{trap}. An additional consideration in the design is to limit the maximum field on the KEK cable to below 7 T so that the KEK magnet can operate at approximately 90% of the loadline.

In an Ioffe-type trap, bucking coils are often added to each side of the pinch solenoids to increase effective trap length and reduce peak field at the turnaround region. It was discovered that the gain from adding bucking coils to the KEK trap was modest. Figure 6.3 shows both trap designs with and without bucking coils and the expected detection efficiency[2] of the KEK trap. The addition of bucking coils allows the solenoids to be moved closer to the turnaround regions, resulting in a longer trap length and a larger trapping volume. Although the number of trapped neutrons

[2]The efficiency curve is obtained by scaling up the measured efficiency curve of the current trap by the ratio of the trap radii.

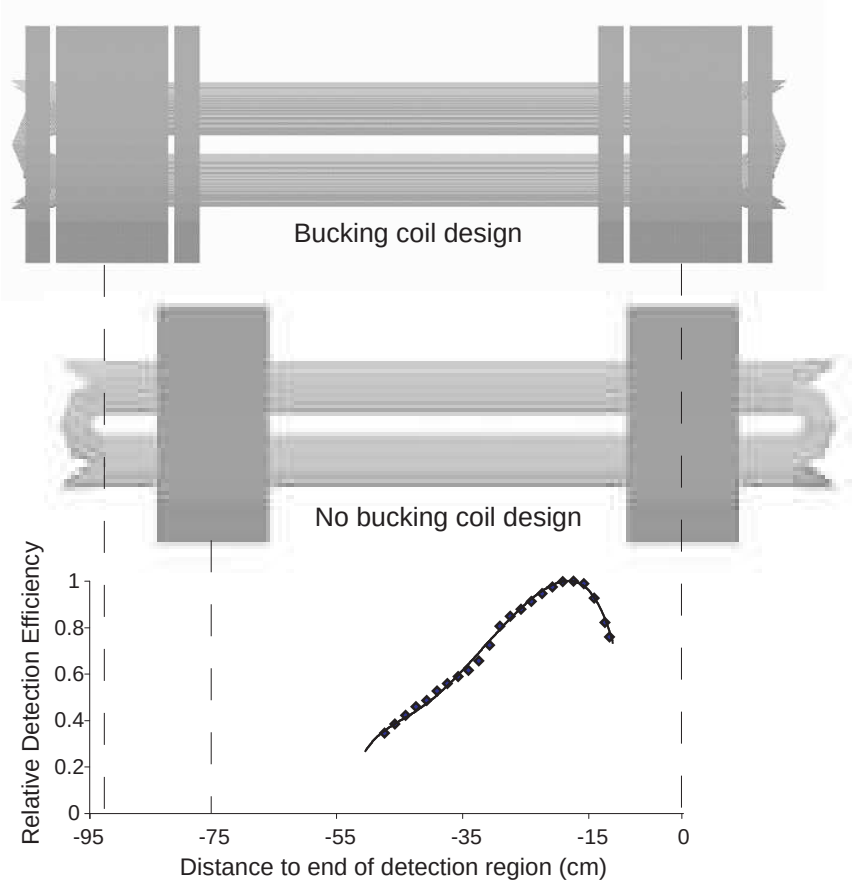

Figure 6.3: Comparison of the trap designs with and without bucking coils. The bucking coil design has a longer trap length, but the detection efficiency is very low in the added trap region.

could be increased by 50%, the increase in detectable neutron decays would be less than 15% because the detection efficiency in the increased trapping volume is close to zero. In addition, adding bucking coils would certainly increase the complexity and cost of the solenoids. There was also concern that the large repulsive lorentz force between the pinch solenoids and the bucking coils could be a cause of excessive training observed in previous traps. A very sturdy magnet form would need to be designed to counter this repulsive force. Given these considerations, we chose the

Table 6.1: Main KEK magnetic trap parameters.

Parameter	
Solenoid winding ID, cm	29.5
Solenoid winding OD, cm	36.1
Solenoid winding length, cm	18.0
Solenoid form ID, cm	28.1
Solenoid form OD, cm	39.5
Solenoid form length, cm	30.8
KEK bore size, cm	14
KEK magnet length, cm	145
Distance between solenoid centers, cm	75.0
KEK Quadrupole operating current, A	3405 (90% loadline)
KEK Quadrupole inductance, mH	58
Solenoid operating current, A	225 (75% loadline)
Solenoid inductance, H	7
Trap radius, cm	5.5
Trap depth, T	3.1

no-bucking coil design for its higher probability of success, despite the estimated 15% loss in signal.

The idea of winding the solenoids with superconducting cables and operating them at the same current as the KEK magnet was considered briefly, but it was soon realized the cost would be significantly higher than low current epoxy bond solenoids. As discussed before, low current solenoids are not well suited for quench protection. In order the reduce the probability of a catastrophic quench, the low current solenoids were designed to operate at only 75% of the loadline, although pushing the design to 90% of the loadline by shortening the solenoids could increase the trapped neutron number by an estimated 7%. The solenoids were wound by American Magnetics, Inc. on forms specially designed by our group. Table 6.1 shows the main design parameters of the KEK magnetic trap.

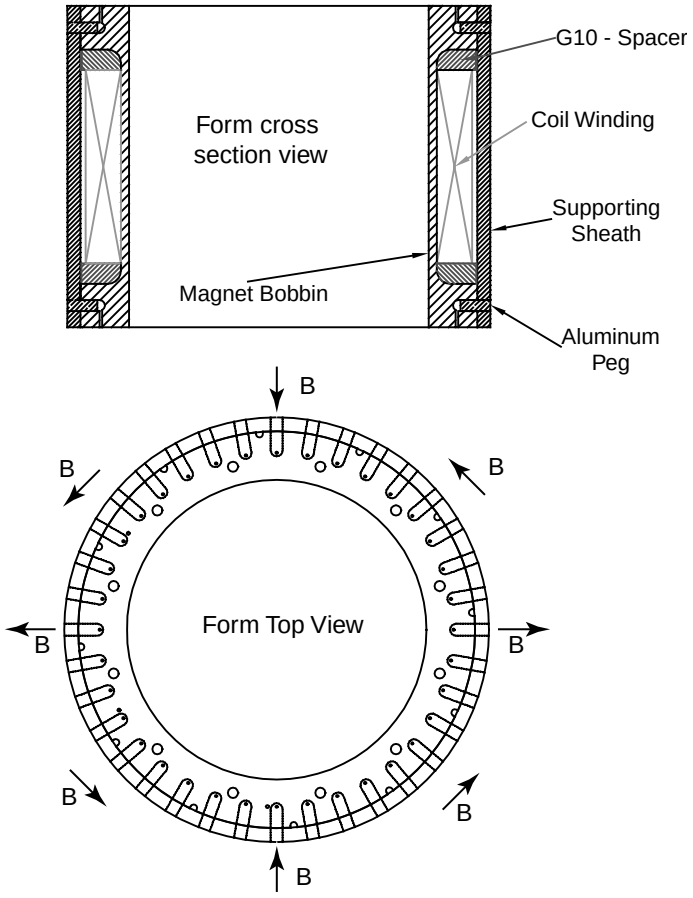

Figure 6.4: The top view and bottom view of the solenoid form design for the KEK trap. The arrow indicates the direction of the magnetic field from the KEK magnet.

6.2.3 Solenoid Form Design

When a superconducting solenoid is energized, the lorentz force pushes wires towards each other and the solenoid as a whole radially outwards. Because the solenoid form does not bear any force in stand-alone operation, it is typically made from aluminum with 1/4" wall thickness. However, in the KEK trap, the field from the quadrupole magnet yields additional lorentz forces which need to be countered by the solenoid forms.

The field from the KEK magnet at the solenoid region has a magnitude of 1 T. The direction of the field is shown in Figure 6.4. While the tangential component of the field which points in the same direction as the solenoid current exerts no lorentz force, the radial component of the field induces lorentz forces that pushes the coil towards the end flanges of the magnet form. The total force on the flanges is roughly 7.9×10^4 N. Because the field radial component changes direction for every 90 degree increment in azimuthal angle, the lorentz force also reverses its direction every 90 degrees. The net effect is to twist the coil into a saddle shape.

Several magnet form designs were considered. The form deformation and stress concentration were calculated using a finite element program ANSYS[3]. Results show that adding a supporting sheath outside the magnet form can reduce the maximal stress concentration and form displacement by a factor close to ten. The form design is shown in Figure 6.4. The sheath is connected to the form by supporting pegs. Due to the large stress concentrations in the form especially at the pegs, aluminum 2219-T87, which has a yield strength of 440 MPa[4], was chosen as the form material.

The magnet bobbins and supporting sheaths were machined at North Carolina State University. The bobbins were shipped to American Magnetics, Inc. for coil winding. After winding, the supporting sheaths were added and the support pegs were machined to fit in the side holes with 15 μm tolerance.

[3]A finite element analysis program by ANSYS, Inc..

[4]Common aluminum alloys, such as 6061, only have yield strength on the order of 66 MPa.

6.2.4 Quench Protection

At operating current, the energy stored in the KEK magnet is 336 kJ, and the energy stored in each solenoid is 180 kJ. Such large stored energies, if deposited in a small section of the magnet during a quench, can cause permanent damage to the winding, possibly even vaporizing a small chunk of wire. The magnets can be protected against these catastrophic quenches by either passive or active methods. In the active method, the stored energy is quickly dumped into an external resistor after a quench is detected. An active circuit based on this principle was built for the KEK magnet. However, this technique is only effective for low inductance magnets, and does not work for the solenoids which have inductance of 7 H each. Instead, the solenoids are protected passively by diodes across six subdivisions of the magnet. During a quench, when the voltage across a subdivision goes above a certain value (2 V), the corresponding diode goes into conduction, letting some current bypass the subdivision, therefore reducing its current density and heating. The passive technique, though cheaper to implement, has the disadvantage that it dumps all stored energy in the cryogenic environment. High pressures resulting from explosive boiling of liquid helium can cause leaks into the surrounding vacuum chamber. The active method, on the other hand, avoids dumping energy in the liquid helium bath, but is much harder to implement. Details of the quench detection and protection circuits for the KEK magnet are discussed below.

A magnet quench can be detected by observing a voltage disbalance signal between matching pairs of coils. In the KEK magnet, voltages across eight two-layer racetrack coils can be measured from voltage taps. Each voltage consists of both resistive

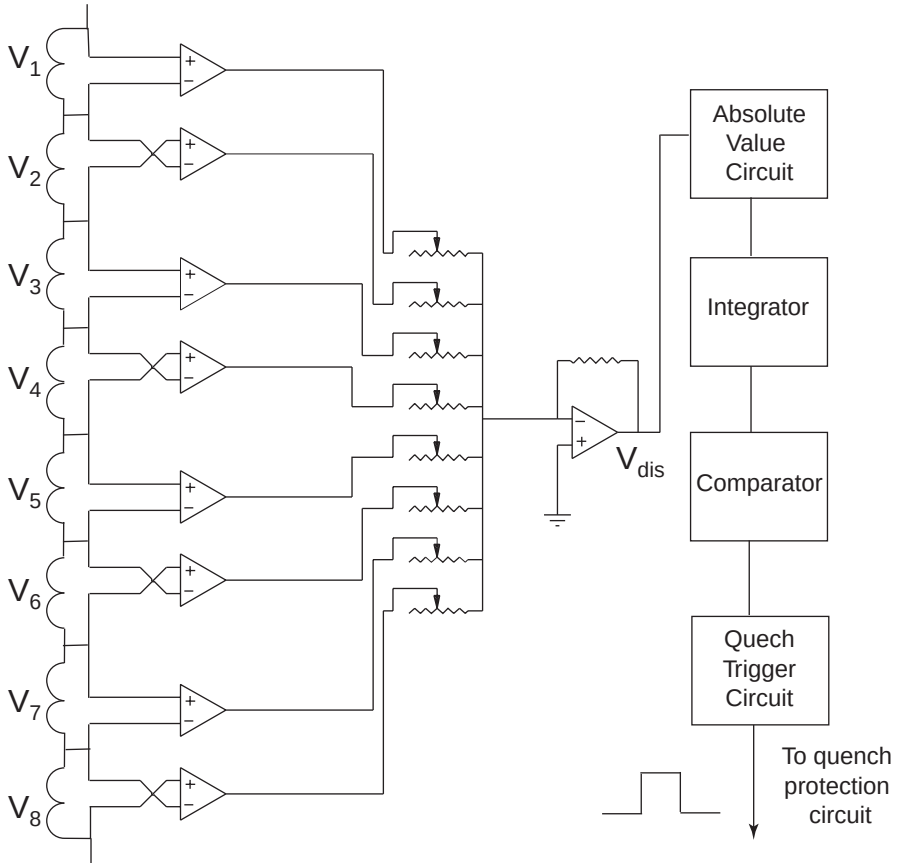

Figure 6.5: Schematic of the quench detection circuit for the KEK magnet.

and inductive components. Though the resistive component is non-zero only after a quench, the inductive component is non-zero during magnet ramps. Because the inductive voltages of two coils opposite to each other in the quadrupole (a matching pair) are almost the same due to similar amounts of magnetic flux penetration, they cancel each other in the voltage difference (with some coefficients compensating for small geometric differences). The sum of voltage differences from matching pairs forms a voltage disbalance signal, which is zero even during magnet ramps. A non-zero disbalance signal can only come from resistive voltage, is therefore a strong

indication of a quench.

The schematic of the quench detection circuit is shown in Figure 6.5. Voltages from eight two-layer racetrack coils are connected to unity gain difference amplifiers which can withstand up to 500 V between inputs and ground. Outputs of the difference amplifiers are summed up with weights set by the potentiometers to obtain a voltage disbalance signal V_{dis},

$$V_{dis} = \alpha_1 V_1 + \alpha_3 V_3 + \alpha_5 V_5 + \alpha_7 V_7 - (\alpha_2 V_2 + \alpha_4 V_4 + \alpha_6 V_6 + \alpha_8 V_8). \tag{6.1}$$

The potentiometers can be fine tuned so that the absolute value of V_{dis} is less than 10 mV during the magnet ramp. When the integrated absolute value of V_{dis} crosses a set quench threshold, a quench trigger circuit generates a 75 ms, 100 mA pulse necessary to trigger the protection circuit. The use of an integrator protects the circuit from accidental triggering by voltage spikes. Optoisolators are used in the trigger circuit to protect the detection circuit from high voltages generated during a quench.

The quench protection circuit built for the KEK magnet follows a design by Fermilab [100], in which Silicon Controlled Rectifiers (SCR)[5] are used for current switching. A schematic diagram of the circuit is shown in Figure 6.6.

The 3400 A current to the KEK magnet is supplied by four HP 6880A DC power supplies connected in autoparallel mode where the output currents from three slave supplies follow the output current of the master supply. Each power supply can output

[5]A SCR is a three terminal device (gate, anode and cathode) commonly used for fast current switching. The connection between anode and cathode normally does not conduct current, but can be switched into a conductive state by positive bias voltage (0.8 - 1.4 V)applied to the gate-cathode junction. Once in the conductive sate, SCR stays on until both gate voltage is removed and current flow stops for longer than t_q (15μs in our case).

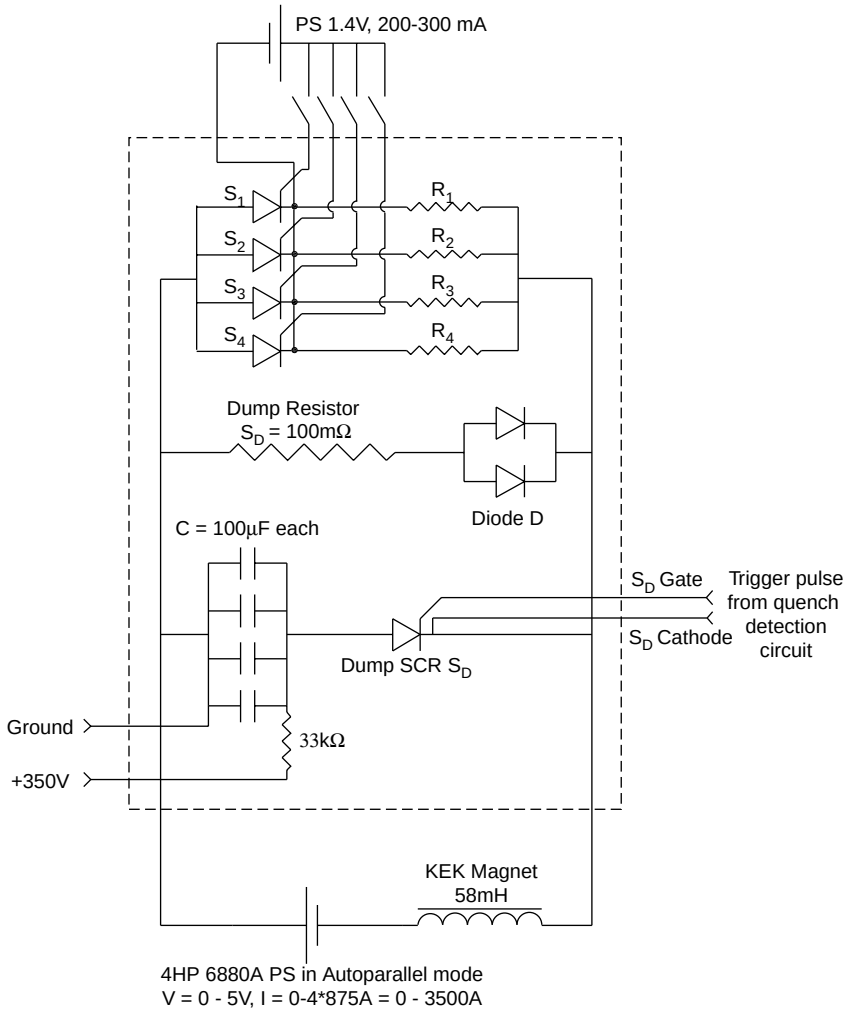

Figure 6.6: Circuit diagram of the quench protection circuit.

current up to 875 A with maximum voltage of 5 V. The output current stability can be optimized by DIP switches inside the power supplies based on inductive load of 40 mH and resistance of 100 $\mu\Omega$. The outputs of the power supplies are connected with copper (alloy 110) bus bars (1.9 cm $\times$ 10.2 cm). The same size bus bar is used to make connection with the magnet current leads. A thin layer of copper based conductive paste (Penetrox) is applied between connection surfaces to ensure good

electrical contacts.

In normal operation, as current in the magnet ramps up, four run SCRs (S_1-S_4) can be turned on manually by applying a bias voltage ($\approx$ 1.2 V) to the gate-cathode junction for 1 second. Because the voltage drop across each SCR is slightly different and SCRs have exponential I-V curves, four resistors (R_1-R_4) are put in series with the run SCRs to evenly split the current between them. Each resistor was made of a 17.8 cm long CuNi tube with a 1.3 cm ID, a 1.6 OD, and a resistance of 0.6mΩ. The current through each SCR can be measured by monitoring the voltage across its corresponding resistor. Both run SCRs and resistors have to be actively cooled, because of the large amount of energy dissipated, roughly 460 W in each resistor and 1.2 kW in each SCR at operating current. Distilled water cooled by a Neslab heat exchanger flows through the hollow centers of the resistors (CuNi tubes) and through chill blocks pressed against the anodes and cathodes of the run SCRs. Connections between chill blocks and resistors are made using flexible 1.3 cm ID tubing. Should the water flow stop, an electronic water flow sensor sends an inhibit signal that turns off the power supplies to prevent overheating the SCRs and resistors.

When a magnet quench is detected, the dump SCR (S_D) switches into a conductive state by a trigger pulse from the detection circuit to its gate. A bank of four 100 μF capacitors connected in parallel then discharges through S_D. The capacitors are initially kept at 350 V by an external high voltage power supply. The discharge current temporarily stops the current flow in the run SCRs for roughly 40 μs[6], long enough to turn all run SCRs into a non-conductive state. After the capacitors are

[6] $t = CV/I_{max} = 40$ μs. I_{max} is taken as the maximum current 3500 A, at lower operating current, the voltage across the capacitors can be reduced correspondingly.

fully discharged, current can only flow through a dump resistor, where majority of the stored energy of the magnet will be dissipated. The dump resistor is made of 1.6 cm thick 316 stainless steel plate. It has a resistance of 100 mΩ resulting in a characteristic energy dump time of $L/R = 0.6$ s. The mass of the resistor (about 30 lb) is big enough to limit its temperature rise after a quench to less then 50° C. Diodes D in series with the dump resistor is used to prevent discharge of the capacitors through the dump resistor, and a soaking reactor in series with the capacitors is used to limit $\mathrm{d}I/\mathrm{d}t$ during the initial phase of the capacitor discharge[7].

6.2.5 KEK Magnet Tests

The KEK magnet along with the quench protection circuit was tested both in 2000 and 2004. In the 2000 test, no natural quench was observed for currents up to 2800 A. In the 2004 test, heaters were installed to induce magnet quenches, and the efficiency of the protection circuit was measured. Due to the similarity in setup, the following discussion focuses on the 2004 test.

The magnet was tested in a 2.5 m tall, 40.6 cm bore vertical dewar manufactured by Precision Cryogenics. Three 1/2-13 stainless steel threaded rods support the magnet from the top flange. The current to the magnet was brought in by a pair of 3000 A vapor cooled current leads. Two ball valves were installed on the vapor leads to control the flow of helium. The magnet superconducting cable leads were soldered to NbSn extension bars attached to the the current leads with at least six

[7]A soaking reactor is basically a inductor. Here it is a made of a steel tube 12.7 cm long with 2.5 cm ID and 10.2 cm OD. A 1 mm thick slit is machined on the side of the tube to prevent remnant magnetization of steel. The limit on $\mathrm{d}I/\mathrm{d}t$ is set by the specification of S_D.

inches of overlap to reduce contact resistance. A six inch liquid helium level meter was installed at the end of the extension bar. Eight twisted pairs of thin copper wires were soldered to voltage taps of the KEK magnet, and brought out from the top through a hermetically sealed connector. A NW 80 blank-off aluminum flange normally held together by springs served as a pressure relief valve. A set of radiation shields made of 0.5 mm thick aluminum sheets was used to block blackbody radiation from the top of the assembly. Four 5Ω resistor heaters were attached to the windings of four inner coils to induce quench.

It took about 400 liters of liquid nitrogen to cool the magnet to 77 K. The magnet resistance dropped from 3.6 Ω at 300 K to 0.45 Ω at 77 K. Excess liquid nitrogen was pressurized out by helium gas through a tube extending all the way from the top flange to the bottom of the dewar. 500 liters of liquid helium was enough to cool down the magnet with some liquid accumulation in the dewar. Liquid helium boil-off was measured to be about 5 l/hr when the magnet was not energized.

During the test, no natural quench of the KEK magnet was observed with currents up to 2800 A. Quenches could be induced in the magnet cable leads by letting the liquid helium level drop below the solder joints. Quenches could also be induced in the coils by running 2 A through one of the heaters, introducing roughly 20 W. The efficiency of the quench protection circuit can be measured in three ways, by measuring the amount of boil-off liquid helium, by measuring the temperature rise of the dump resistor and by measuring the voltage decay curve across the dump resistor. All three methods were consistent with each other, however the third method had the highest precision. The voltage decay curve of the dump resistor was measured

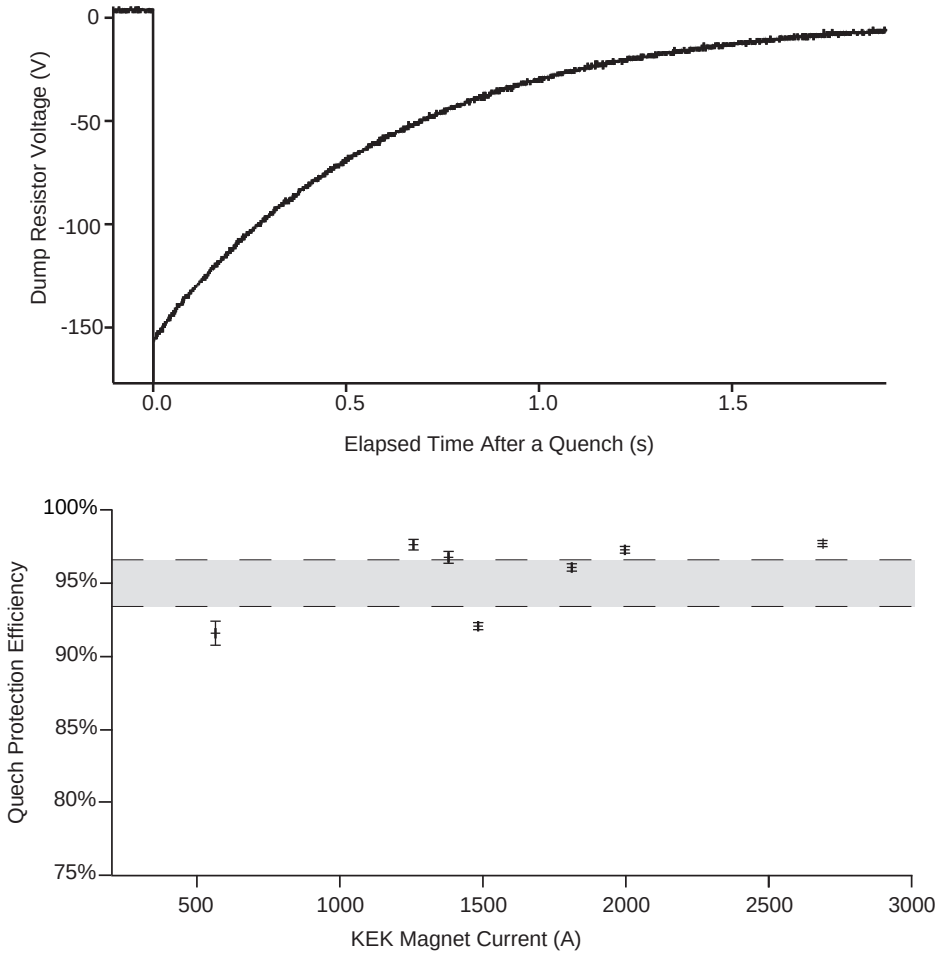

Figure 6.7: The upper graph shows the voltage decay curve across the dump resistor after an induced quench at magnet current of 1500 A. The lower graph shows the measured quench protection efficiency at various magnet currents. On average, the protection circuit has an efficiency of 95%.

by running a digital oscilloscope in single acquisition mode triggered by the quench detection circuit. Figure 6.7 shows the dump resistor voltage decay curve after an induced quench at magnet current 1500 A. The curve has a characteristic decay time of $L/R = 0.6$ s as expected. The energy dissipated in the dump resistor can be calculated as $E_{dump} = \int V^2/R dt$. The ratio of E_{dump} to total stored energy $E = 1/2LI^2$ gives the efficiency of the protection circuit at current I. As shown in Figure 6.7, the

protection circuit has an average efficiency of 95%.

6.2.6 KEK Trap Test

The KEK magnet and two solenoids were assembled and tested in summer 2005. The whole trap reached about 90% of the design current after two quenches. We were limited by vapor cooled current leads from going to higher currents in this test, but we fully expect the trap to reach its design current in future tests.

The whole trap assembly was tested in the vertical dewar described above. The test setup was also similar. In addition, a pair of 250 A vapor cooled leads was installed to bring current to the solenoids. The solenoids were connected in series with the same current sense. The solenoid leads were taped against the magnet form to reduce their movements under Lorentz force. After soldering, it was discovered that several hundred kilo-ohm resistance developed between one solenoid and its aluminum form. The most likely place for this resistance to develop was where the leads exit the magnet form. No obvious defects were found upon inspection. As a precaution, the solenoids were wrapped in kapton tape, to avoid sparking between the magnet form and the dewar, in case of a real short between solenoid windings and the form. Fortunately, this resistance became infinite upon cool down to 4 K. The resistance of the solenoid assembly itself was 316.8 Ω at room temperature, and 56.7 Ω at 77 K.

Initially, the KEK magnet and the solenoid assembly were tested individually. The KEK magnet reached 2900 A, and the solenoid assembly reached 220 A without a quench. When the KEK magnet and the solenoids were ramped up together, two quenches occurred. The first one at about 80% of the design current and the second

Table 6.2: Parameters of two quenches observed during KEK trap test.

	Quench 1	Quench 2
KEK magnet current (A)	2700	2950
Percentage of design current	79%	87%
Solenoid current (A)	190	210
Percentage of design current	84%	93%
Quench initiation place	N/A	KEK

one at about 90% of the design current. Because the KEK magnet and the solenoids were controlled by separate power supplies, their ramping rate did not exactly match each other. The exact parameters of the quenches are shown in Table 6.2. After the first quench, the voltage disbalance signal from the KEK magnet and the voltages across the solenoids were monitored by an oscilloscope. Based on the registered traces, it could be determined that the second quench originated in the KEK magnet. Though majority of the stored energy in the KEK magnet was dumped in an external dump resistor, all stored energy in the solenoids was dumped in the cryogenic environment; enough to boil off 100 - 120 liters of liquid helium. After an initial explosive boil-off, increased liquid helium boil-off was observed for ten to fifteen minutes. Because the current leads for the KEK magnet were rated for 3000 A and one lead in particular was heating up quickly at 2950 A, we did not attempt to go to higher current after the second quench. This limitation can be easily overcome in future tests with the use of higher rated current leads, either vapor cooled or the high temperature superconducting leads discussed in the next section.

It is very likely that the performance of the trap at 4.2 K will be limited by the

KEK magnet, because, in the trap design, the KEK magnet is pushed closer to its loadline (90%) than the solenoid (75%). The only data point we have so far also indicates that the quench originated in the KEK magnet. This situation makes the option of running the trap in superfluid helium very attractive. The performance of the KEK magnet is expected to increase by 30% at 1.8 K. Superfluid helium in direct contact with the superconducting cables can help stabilize the magnet against quenches. The performance of the solenoid is also expect to increase by 30% at 1.8 K. However, because the superconductors would not be in direct contact with superfluid helium, the increase in current density would make the solenoids more prone to quenches. Furthermore, without an effective protection circuit like the KEK magnet, an increase in stored energy also increases the probability of damage to the solenoids in a quench. These concerns will likely be the limiting factors of the trap performance at 1.8 K.

6.3 High Temperature Superconducting Leads

Conventional vapor cooled current leads have a thermal performance limit of 1.2 W/kA per lead [101]. To bring 3400 A current to the KEK magnet with these leads would put at least 8.2 W power into the liquid helium bath. With such a high heat load, it would be too costly to operate the trap continuously. A recently developed technology that incorporates high temperature superconductor (HTS) into current leads can reduce the heatload into 4 K by as much as a factor of ten [102], though the cost of such HTS lead is rather high[8]. We were able to borrow a pair of

[8]A pair of 3500 A HTS leads was quoted at $60,000 by HTS-110 Ltd..

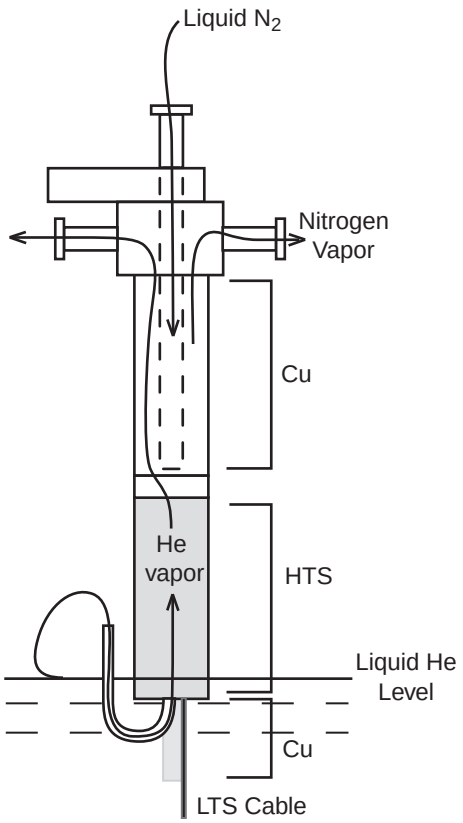

Figure 6.8: Cross section view of a Fermilab HTS lead.

5000 A HTS leads from Fermilab, which can reduce the heatload by a factor of six. The quench protection and test results of the Fermilab leads will be dicussed. We also realized that using HTS leads for the 250 A solenoids can further reduce liquid helium boil-off. The design of the low current HTS leads will also be discussed.

6.3.1 The Fermilab HTS Leads

The Fermilab HTS leads were developed to replace about 50 pairs of 5 to 6 kA current leads at Fermilab's Tevatron [103]. The pair we obtained is a prototype pair

developed by American Superconductor Corporation.

The cross-section of a Fermilab HTS lead is shown in Figure 6.8. The upper section of the lead is made of copper. Liquid nitrogen is fed directly to the copper and HTS junction through a tube at the center. This junction has to be kept below 80 K for the proper operation of the lead. Nitrogen vapor cools the copper conductor, then escapes through a vapor exhaust. The bottom section of the lead consists of parallel tapes of BSCCO-2223 (powder-in-tube) multifilamentary conductor in a silver alloy matrix [103]. Helium vapor flows through the center of the HTS section, around the copper section, before exiting a vapor exhaust. Though cooling provided by the helium vapor is not required for the lead operation, it helps to reduce the lead heatload into 4 K. A low temperature superconducting (LTS) cable is directly connected to the bottom of the lead. This junction has to be kept below 8 K for the LTS cable to stay superconducting. A six inch copper extension bar is installed next to the LTS cable. As long as the liquid helium level is above the bottom of the copper bar, sufficient cooling will be provide to the junction. At Fermilab's Tevatron, the liquid helium level is kept constant at the HTS-LTS junction.

A quench occurring in a HTS lead propagates very slowly due to its low thermal conductivity. A significant amount of heat can be dissipated in a small region before a large voltage is developed across the lead, therefore the quench detection thresholds are set very low. For the Fermilab lead, the protection criterion set by the manufacturer is that if the voltage across the HTS section goes above 1 mV or the voltage across the copper section goes above 30 mV, the current has to be stopped within ten seconds. For detecting such small voltage changes, it is particularly important

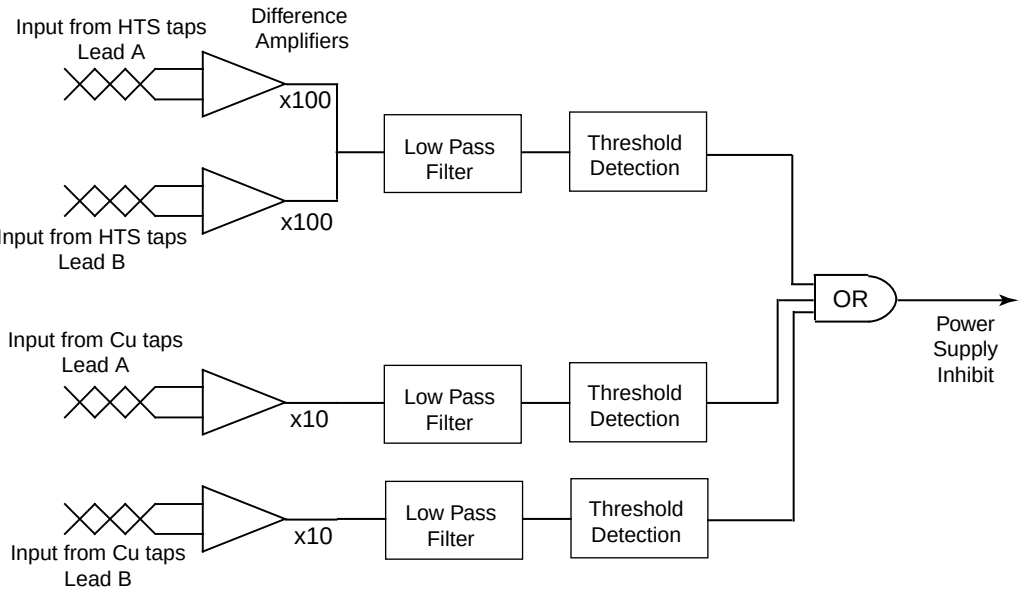

Figure 6.9: Schematic of the Fermilab HTS lead quench detection circuit.

to minimize electromagnetic interference by bringing out the voltage tap signals in twisted pairs or well shielded cables. The schematic of the HTS lead quench detection circuit is shown in Figure 6.9. The voltages across the HTS sections are first amplified by a factor of one hundred, and the voltages across the copper sections are first amplified by a factor of ten. The amplified voltages are sent through low pass filters to eliminate transient noise. If one of the voltages crosses the set threshold, the power supply inhibit signal will be triggered to ramp down the current in the magnet.

A pair of Fermilab leads were tested in a ten inch bore vertical dewar. The two leads were connected at the bottom by a LTS cable. Four HP 6880 A DC power supplies in parallel provided a maximum testing current of 3410 A. The temperatures at the Cu-HTS and HTS-LTS junctions could be monitored by resistor thermometers. Flow meters were installed at the nitrogen and helium gas exhausts to measure cryo-

Table 6.3: Operation parameters of the Fermilab HTS Leads at 3410 A.

Parameter	
Liquid helium consumption, liter/day	50
Equivalent power to 4 K, W	1.5
Liquid nitrogen consumption, liter/day	150
HTS section voltage Lead A, mV	1.8
HTS section voltage Lead B, mV	0.3
Cu section voltage A & B, mV	≈ 20

gen consumptions. We denoted the two leads as A and B. During the first cooldown, a large resistance developed at the Cu-HTS junction of Lead A. Upon warming up, we discovered that a low temperature solder joint at the junction was broken. Fortunately, Dr. Feher's group at Fermilab was able to fix the solder joint for us. The leads were successfully tested up to 3410 A during the second cooldown. Some operation parameters of the current leads at 3410 A are shown in Table 6.3. The voltage drop across the HTS section of Lead A was 1.8 mV, higher than the manufacturer's specification. The reason for this was that when Fermilab fixed the broken solder joint, they could not reattach the voltage tap to the HTS material itself, instead the tap was soldered to the copper piece right above the HTS section. The measured voltage therefore included additional voltage drop across a small section of copper. The quench detection circuit threshold was adjusted to 3 mV for Lead A. The measured liquid helium boil-off rate was 50 litres/day, in agreement with data from a Fermilab test [103]. The measured liquid nitrogen consumption was 150 liters/day, higher than expected. However, a large amount of heat loss was due to the rubber

hose connections between the liquid nitrogen dewar and the leads. If double walled nitrogen transfer tubed were used, the liquid nitrogen consumption could be reduced to below 100 liters/day.

As discussed before, the HTS-LTS junction has to be kept below 8 K. At Fermilab, this is achieved by keeping the liquid helium level constant at the junction. For the neutron lifetime apparatus, the large initial cost and additional helium consumption of an automatic filling system makes this option unrealistic. An alternative method is to use a two-stage cryocooler[9] with a cooling power of 1.5 W at 4 K to cool the junction and the LTS cable below. Furthermore, the cooling power of the first stage can be used to cool the Cu-HTS junctions of a pair of low current HTS leads for the solenoids.

6.3.2 Low current HTS leads

Using HTS instead of vapor-cooled current leads for running the 250 A solenoids can cut the lead helium boil-off from 20 l/day to 3.3 l/day. Such low current HTS leads can be implemented by connecting commercially available HTS tape leads to custom designed copper rods.

A schematic of a 250 A HTS lead is shown in Figure 6.10. A copper rod brings the current in from room temperature to the Cu-HTS junction. Then a multifilamentary HTS tape with operating current of 250 A at 64 K[10] conducts the current to 4 K. The Cu-HTS junction has to be cooled to below 64 K for the proper operation of the

[9] Janis research company sells a two-stage Gifford-McMahon type cryocooler with 1.5 W of cooling power at the second stage.

[10] The HTS tape lead was obtained from HTS-110 Ltd. at New Zealand

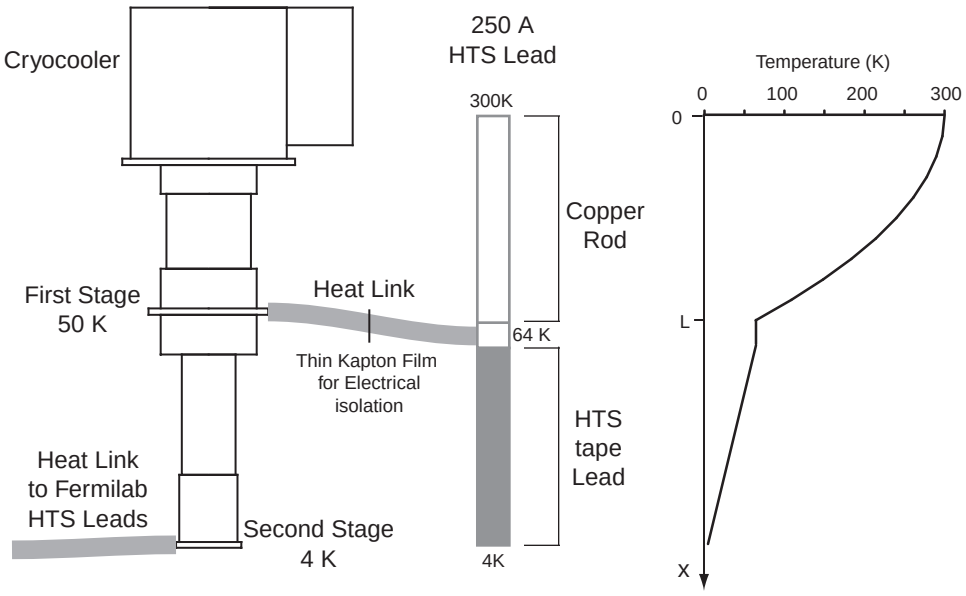

Figure 6.10: Schematic of a 250 A HTS lead. The Cu-HTS junction of the lead is cooled by the first stage of a 1.5 W cryocooler. The graph on the right shows the temperature distribution across the lead.

lead. The heat load onto the junction comes from both the ohmic and conduction heating of the copper rod. Neglecting the effect of vapor cooling, we can write the thermal balance of the copper rod as a second order differential equation,

$$d(kA(dT/dx))/dx + \rho I^2/A = 0 \tag{6.2}$$

where k is the thermal conductivity of copper, A is the cross-section of the copper rod, T is temperature, ρ is the resistivity of copper and I is current. To simplify the problem, we assume the copper rod has a uniform cross-section and k and ρ are independent of temperature. Then Equation 6.2 has a simple solution,

$$T(x) = T(0) - \frac{\rho I^2 x^2}{2kA^2} + (\frac{\rho I^2 L}{2kA^2} - \frac{T(0) - T(L)}{L})x \tag{6.3}$$

where L is the total length of the copper rod. The heat load onto the Cu-HTS junction

Q can be calculated as

$$Q = -kA\frac{dT}{dx}|_L = \frac{\rho I^2 L}{2A} + \frac{(T(0) - T(L))kA}{L}. \tag{6.4}$$

Q has a minimum value of $\sqrt{\rho k(T(0) - T(L))/2}I$ at $A/L = \sqrt{\rho/(2k(T(0) - T(L))}I$. Copper is chosen as the lead material precisely because it minimizes the product of resistivity and thermal conductivity, ρk. Using the boundary conditions $T(0) = 300$ K and $T(L) = 64$ K, the average thermal conductivity $k = 6$ W/cm K and the average resistivity $\rho = 9 \times 10^{-7}$ Ω cm, we obtain that $Q = 6.3$ W, and $A/L = 4.5 \times 10^{-3}$ cm for a 250 A lead. The results from this simplified model is very close to numerical solution of Equation 6.2 taking into full consideration of the temperature dependence of k and ρ.

The Cu-HTS junctions can be cooled by the first stage of the 1.5 W cryocooler used to cool the HTS-LTS junctions of the Fermilab leads. The first stage has 45 W of cooling power at 50 K, more than enough for the 12.6 W heat load at the 250 A lead junctions. Still thermal links between the leads and the cryocooler need to be designed carefully to avoid large temperature drops. Electrical isolation can be achieved with 5 mil thick kapton films[11] or G-10 sheets.

6.4 Cryogenic Posts

To support the KEK trap while minimizing the heatload onto the liquid helium bath, we designed and tested G-10 fiberglass based cryogenic posts. Among commonly used cryogenic materials, G-10 has the highest yield strength to thermal conductivity

[11] 3M produces a kind of thermally conductive kapton film which has twice the thermal conductivity of normal kapton.

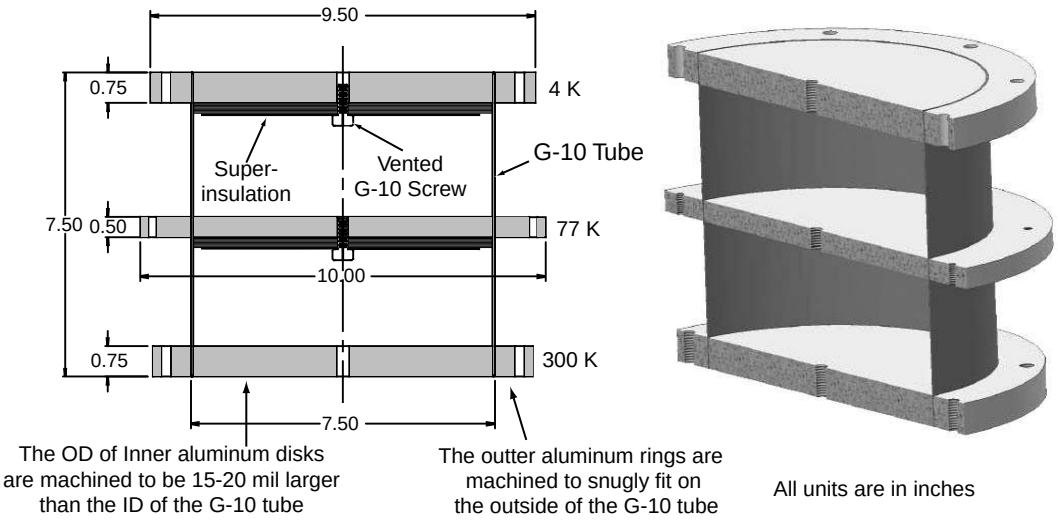

Figure 6.11: Two dimensional and three dimensional cross-section views of the G-10 based cryogenic post.

ratio, about ten times higher than stainless steel.

Our cryogenic post design closely follows the design first developed by Thomas Nicol at Fermilab. Schematic views of the post are shown in Figure 6.11. The main body of the post is a 7.5 inch long, 7.5 inch diameter G-10 tube with 1 mm wall thickness. Equally spaced aluminum flanges at 300 K, 77 K and 4 K provide lateral mechanical support for the tube and bolt holes for attachments to the dewar. All flanges were shrunk-fit onto the tube. The ID of each outer aluminum ring was machined to the exact OD of the G-10 tube, while the OD of each inner metal disk was machined to be 0.5 mm more than the tube ID. During assembly, the outer rings were fitted onto each joint first, then the inner disks were cooled down to liquid nitrogen temperature before they were fitted onto each joint. Small lips on the 300 K and 4.5 K flanges help to register them during assembly. Side surfaces of the rings and disks were roughened slightly by sand paper to increase the friction coefficient.

The 77 K and 4 K disks were covered with super-insulation to reduce blackbody radiation. The internal volume of the post was vented through holes in the metal disks and G-10 screws, because tests showed that any vent holes on the side of the G-10 tube significantly reduced the load carrying capacity of the post.

A test post was built based the above design. It was load tested up to 3000 lb at room temperature. The heatload of the post, though not directly measured, can be calculated from available G-10 thermal conductivity data. For a single post, the heatload onto the 4 K flange is 0.3 - 0.4 W.

The KEK trap can be supported from two posts, see figure 6.12. Due the thermal contraction of the 4 K can, one of the post should be allowed to move. Teflon or dry lubricant materials can be used as the post sliding surface at 300 K.

6.5 New Dewar Design

6.5.1 Two tower design

In the current inverted T-shape dewar described in Chapter 4, a ten inch bore vertical tower houses both the dilution refrigerator and the magnet current leads. Because the size of the HTS leads for the KEK trap are much larger and a 1.5 W cryocooler needs to be incorporated for cooling the leads, the same design will require a vertical tower with bore size of at least 24 inches. In an alternative two-tower design, the magnet leads and the cryocooler can be fitted into a separate 16 inch bore vertical tower, and the original vertical tower can be reused to house the dilution fridge. This design avoids large seals in the vertical section, simplifies connections in

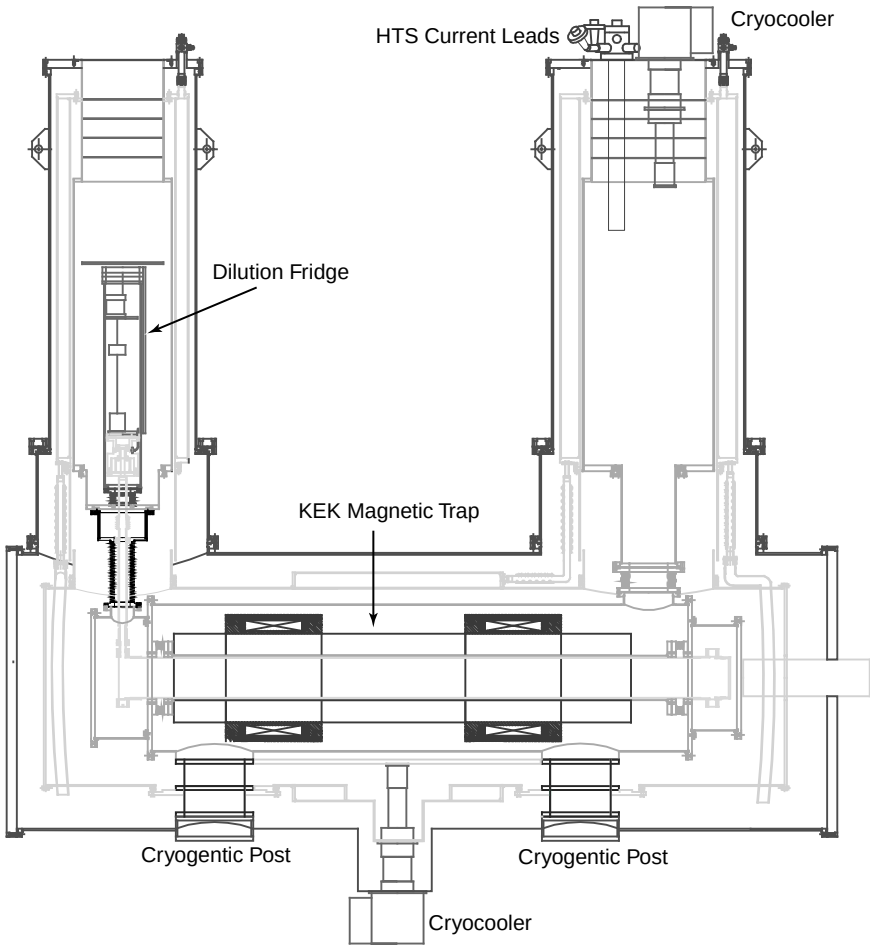

Figure 6.12: Schematic of a two tower dewar design for holding the KEK magnetic trap.

each tower and reduces the total number of indium seals at 300 mK and 4 K.

A schematic of the two tower design is shown in Figure 6.12. The 300 K cans of the vertical and horizontal sections are connected rigidly, and the whole weight of the dewar is supported from the horizontal 300 K can. The liquid nitrogen and helium bath of the vertical and horizontal sections are connected with bellows to allow for thermal contraction. The dilution fridge tower has a separate liquid helium reservoir from the magnet liquid helium bath. A magnet quench will have minimum effect on

the fridge. The horizonal 4 K can is supported by two cryogenics posts. A cryocooler can be installed to take away heatload from the posts as well as the 4 K end flanges. The cell construction, neutron entrance windows and light collection system will use the same techniques as those developed for the previous dewar, see Chapter 4.

The new design addresses several problems of the old dewar pointed out in Chapter 4. First, the horizontal section will have its own liquid nitrogen jacket which brings direct cooling to the end flanges. Secondly, the distance between 4 K and 300 K end flanges on the light collection side will be long enough to enable proper cooling of the lightguide, which can significantly reduce the blackbody radiation onto the 4 K and cell windows. Thirdly, two 1.5 W cryocoolers will be incorporated to reduce the liquid helium boil-off. Fourth, the weight of the magnet will be supported by posts from the bottom instead of by indium seals in the vertical section. Finally, the heat link between the cell and the dilution fridge is simplified. Only the neutron entrance side of the cell is connected to the mixing chamber. Such an arrangement reduces the number of indium seals at 300 mK and 4 K.

In the following section we will not go into detailed discussions of cell or indium seal designs, as they are all proven technologies. Instead we will discuss the heat load onto the cell and the 4K can, because it is essential to make sure that the cell can be cooled to below 300 mK with the simpler heat link design, and liquid helium boil-off is the major component of the experiment running cost.

6.5.2 Heat load calculations

The cell heat load comes from blackbody radiation, conduction heat from cell supports and sensor wires, neutron beam heating and eddy current heating. The estimated heat load from each source is summarize in Table 6.4.

Table 6.4: Summary of estimated cell heat loads. The eddy current calculation assumes a magnet ramp time of 200 s. Duration is the fraction of time when the heat source is on.

Heat Source	Heat input (μW)	Duration
Blackbody radiation	48	100%
cell supports and sensor wires	5	100%
neutron beam	44	40%
Eddy current in cell wall (CuNi)	1000	8%
Eddy current in buffer cell (copper)	1200	8%

From the table, we see that eddy current heating has the largest peak heat input. Fortunately, it only happens during magnet ramps which take place about 8% of the total running time. The average heat load will be 251 μW, with peak heatload of 2 mW during magnet ramps.

The cell is connected to the dilution fridge through a 1.9 cm diameter superfluid heat link. The thermal conductivity of superfluid helium is experimentally determined to be 20 d T^3 (W K^{-4} cm^{-2}) below 0.7 K (d: diameter of tube in cm) [57]. Using this formulae and the measured cooling power of the dilution fridge (see Section 4.2), we estimate that the temperature of the bulk liquid helium inside the cell to have an average temperature of 200 mK. Because of the large diameter of the cell, there will

be virtually no temperature gradient across the cell. During a magnet ramp, about 0.4 J of energy will be dumped into the cell, which will raise the temperature of the cell to about 400 mK. The cell temperature will then cool back down to 200 mK in several hundred seconds. In addition, the eddy current heating can be reduced by the use of a plastic cell or magnetic shielding at the buffer cell region. A slower magnet ramp speed will also cut down the eddy current heat, but it is not a preferred method because it also reduces the signal.

In short, the cell in the proposed design can be cooled down to below 300 mK. The only concern comes from eddy current heating. Though this can be dealt with by a slower ramping speed initially, the development of active or passive magnetic shielding near the buffer cell and a plastic experimental cell will be needed for the long term.

The heat load onto the 4 K can comes from backbody radiation, cryogenic support posts, HTS current leads, eddy current heating in the solenoid forms, dilution fridge and the two vertical towers. Table 6.5 summarizes the estimate for each heat source. The blackbody radiation is estimated assuming an emissivity of 0.1 for surfaces covered by super-insulation and 1 for the neutron and optical windows. The eddy current heating in the aluminum form is estimated assuming a magnet ramp time of 200 s. The liquid helium consumption of the dilution fridge was measured during a stand alone test.

The total heat input onto the 4 K can based on the estimates sums up to 4.7 W, or 156 l/ day. To reduce the liquid helium boil-off, two 1.5 W cryocoolers will be used in the apparatus. One connects to the HTS current leads, and the other attaches

Table 6.5: Summary of the heat load onto the 4 K can.

Heat Source	Heat input (W)	Duration	Equiv. LHe (l/day)
Blackbody Radiation	0.8	100%	27
Cryogenic Posts	0.8	100%	27
Fermilab HTS leads	1.4	100%	46
Low current HTS leads	0.1	100%	3
Eddy current in solenoids	5	8%	13
Dilution fridge	0.6	100%	20
Vertical towers	0.6	100%	20

to the cryogenic posts. If the cooling power of the cryocooler can be fully utilized, the liquid helium boil-off can be reduced to roughly 60 liter/day, a very reasonable consumption rate considering the size of the apparatus. Liquid helium transfer could then be reduced to one time a day.

6.6 Conclusion

We have so far cleared the main technical difficulties of the next generation apparatus, including a 3 T deep superconducting magnetic trap, HTS current leads for running the trap and cryogenic posts for supporting the trap. We have also developed a two tower dewar design for housing the magnet. The new design promises to be more efficient and reliable.

Chapter 7

Conclusions and Future Prospects

7.1 Conclusions

This thesis reports the progress towards a precision measurement of the neutron lifetime using magnetically trapped ultracold neutrons. The magnetic trapping apparatus was successfully operated at the 0.89 nm monochromatic neutron beamline at NIST. The trapping of ultracold neutrons is conclusively demonstrated with a statistical significance of 20σ. The observed short neutron trap lifetime 621 ± 18 s at 300 mK was found to be caused by the presence of neutrons with energies above the trap threshold. After the population of above threshold neutrons was reduced using the field ramp technique, the measured trap lifetime 831 ± 58 s is consistent with the accepted value of the neutron lifetime. Based on estimates of systematic uncertainties, a preliminary neutron lifetime measurement of $\tau_n = 831 \pm 58_{stat} \pm 88_{sys}$ can be obtained.

Works to significantly reduce both statistical and systematic uncertainties for

future measurements were also carried out. First, the two leading systematic uncertainties, neutron absorption on ^{3}He and above threshold neutrons, are studied in details. The isotopic concentration of ^{3}He inside the ultrapure helium samples was characterized using accelerator mass spectrometry at Argonne National Lab. Preliminary results indicated a ^{3}He concentration of $(4.2 \pm 1.5) \times 10^{-12}$, much higher than previously believed. Development of alternative measurement techniques and further purification of the helium with the heat flush method should help us reduce this uncertainty by several orders of magnitude. Both analytical and numerical simulations were developed to study the motion of above threshold neutrons in the trap and the effectiveness of the field ramp technique for eliminating the above threshold neutrons. We have shown that with the implementation of a deeper magnetic trap, total elimination of above threshold neutrons would become possible under certain conditions.

Second, the development of the next generation apparatus significantly improving the statistical precision is well under way. Key components of the next generation apparatus have been successfully developed and tested. A new Ioffe trap (KEK trap) consisting of an accelerator type quadrupole magnet and two low current solenoids reached 90% of its design currents. The new trap is capable of trapping twenty times more neutrons than before. High temperature superconducting current leads and cryogenic support posts necessary for the trap operation have also been successfully tested.

7.2 Future Work

The future development of the experiment can probably be divided into two phases. During phase one, the KEK trap will be operated at the NIST monochromatic beam line with the goal of performing a three second measurement. During phase two, the apparatus will be moved to new neutron facilities to further improve the statistics. A measurement with an uncertainty of 0.1 s should be achievable.

7.2.1 Phase One

Because new high flux neutron sources or high density ultracold neutron sources will not realize their full potential in the next two to three years, the NIST monochromatic beamline will remain the best available facility for the lifetime measurement for the next few years.

During this period, the development of the new apparatus incorporating the KEK trap should be completed. Assuming the same background conditions in the experimental area, a statistical precision of 3 s can be reached in 40 days of data taking. This will give us ample time to perform many systematic checks, such as measuring the neutron loss due to phonon upscattering, the elimination of above threshold neutrons and measurement of the imperfect background subtraction. The measurement of ^{3}He concentration in ultrapure helium should be continued with either the accelerator mass spectrometry or laser spectroscopy methods.

R&D work to improve the detection efficiency of the system can also be carried out. Possibilities include the use of reflective polymer to direct more scintillation light towards the PMTs, and the use of high quantum efficiency and/or low tempera-

ture detectors such as avalanche photodiodes (APD), Visible Light Photon Counters (VLPC) or Multichannel Plate based PMTs.

In phase one, we hope to obtain a statistically limited lifetime measurement with uncertainty of three seconds. Such a measurement can already help to resolve the discrepancies existing in material bottle measurements.

7.2.2 Phase Two

When new neutron sources become available and we have gained sufficient understanding of the new apparatus, we can move the KEK trap to the new facilities to further improve the statistical precision.

Two neutron sources currently under development are especially suitable for the lifetime experiment, the Spallation Neutron Source (SNS) at Oakridge, TN, and the ultracold neutron source at North Carolina State University. At the SNS, a large increase in the 0.89 nm neutrons that can be delivered into the apparatus is expected due to the use of supermirror neutron guides and the capability of bringing the apparatus right next to the exit of the neutron guide. The plan to build a separate experimental house outside of the general experiment hall for the 0.89 nm monochromatic beamline can also cut down the constant background by a factor of five. Thus, a statistical precision of 0.2 s can be achieved in 40 days of operation. The ultracold source under development at NC State University is based on the superthermal production mechanism of UCN. Because the liquid helium UCN production region would be located right next to the reactor core, resulting in a 10^3 increase in solid angle, a UCN production rate of 10^3 UCN $\mathrm{cm}^{-3}\mathrm{s}^{-1}$ would be possible. At this source, a sta-

Table 7.1: Statistical estimates for operating the KEK magnetic trap at different neutron facilities. SNS is the Spallation Neutron Source under construction in Oak Ridge, TN. NCSU is the site of the proposed dedicated UCN source.

	NIST	SNS	NCSU
Number of neutrons trapped	4×10^4	1.1×10^6	1.3×10^7
Initial neutron signal amplitude (s^{-1})	20	550	6500
Constant background, (s^{-1})	22	4	11
Time-varying background, (s^{-1})	5.5	150	0.1
Statistical error in τ in 40 days, (s)	3.0	0.2	< 0.1

tistical precision of at least 0.1 s could be achieved in 40 days of operation. Table 7.1 summarizes the expected statistical precision and backgrounds at these new facilities.

In choosing the facility for the future operation of the experiment, it is important to compare the advantages and disadvantages of the two facilities. The advantage at SNS is that very minimal modification to the apparatus is needed for its operation at the SNS beamlines. The disadvantage is that several systematic effects that have been observed in the current apparatus will be intensified. For example, the time dependent background from neutron-induced luminescence and activation will increase, the gain drifts in the PMTs would likely increase because they would be exposed to more intense light during the loading period, and the density of charge particles produced by ionizing radiation events in the cell would also increase, leading to possible depolarization of neutrons.

Operating the apparatus at the UCN source at NC state university can eliminate all the systematic effects mentioned above, because the number of neutrons entering the cell would be reduced by a factor of 10^7. However, considerable modifications to the apparatus would be needed to couple it to the UCN source. Several new

technologies will also need to be developed and incorporated into the apparatus, such as efficient cryogenic UCN transport system and windows that can seal the superfluid helium volume but let the UCN through.

In summary, the potential gain in statistical precision from the new neutron sources makes a precision measurement of the neutron lifetime at 10^{-4} level a realistic goal. Despite added complications, the lower systematic effects at the NC State UCN source makes it a more favorable venue for a precision measurement of the neutron lifetime.

Bibliography

[1] J. Chadwick. Possible existence of a neutron. *Nature* **129**, 312 (1932).

[2] K. Grotz and H. V. Klapdor. *The Weak Interaction in Nuclear, Particle and Astrophysics*. 2nd edition. Adam Hilger, Bristol, Philadelphia and New York, 1990.

[3] R E. Lopez and M S. Turner. Precision prediction for the big-bang abundance of primordial ^{4}He. *Phys. Rev. D* **59**, 103502/14 (1999).

[4] S. Arzumanov, L. Bondarenko, S. Chernyavsky, W. Drexel, A. Fomin, P. Geltenbort, V. Morozov, Yu. Panin, J. Pendlebury, and K. Schreckenbach. Neutron lifetime value measured by storing ultracold neutrons with detection of inelastically scattered neutrons. *Phys. Lett. B* **483**, 15–22 (2000).

[5] A. Serebrov et al. Measurement of the neutron lifetime using a gravitational trap and a low-temperature fomblin coating. *Phys.Lett. B* **605**, 72–78 (2005).

[6] J. M. Doyle and S. K. Lamoreaux. On measuring the neutron beta-decay lifetime using ultracold neutrons produced and stored in a superfluid-^{4}He-filled magnetic trap. *Europhys. Lett.* **26**, 253–258 (1994).

[7] Y. Hayato et al. Search for proton decay through p $\rightarrow$ antineutrino K^+ in a large water Cerenkov detector. *Phys. Rev. Lett.* **83**, 1529 (1999).

[8] E. Fermi. An attempt of a theory of beta radiation. *Z. phys.* **88**, 161 (1934).

[9] R. P. Feynman and M. Gell-Mann. Theory of the Fermi interaction. *Phys. Rev.* **109**, 193 (1958).

[10] J. D. Jackson, M. Treiman, and H. W. Wyld Jr. New limit on the D coefficient in polarized neutron decay. *Nucl. Phys.* **4**, 206 (1957).

[11] S. Eidelman et al. Review of particle physics. *Phys. Lett. B* **592** (2004).

[12] A. Sher et al. New, high statistics measurement of the $K^+ \rightarrow \pi^0 e^+ \nu$ (K^+_{e3}) branching ratio. *hep-ex/0305042* (2003).

[13] T. Alexopoulos et al. A determination of the Cabibbo-Kobayashi-Maskawa parameter $|V_{us}|$ using K_l decays. *Phys. Rev. Lett.* **93**, 181802 (2004).

[14] N. Cabibbo, E. C. Swallow, and R. Winston. Semileptonic hyperon decays. *Ann.Rev.Nucl.Part.Sci.* **53**, 39–75 (2003).

[15] I. S. Towner and J. C. Hardy. Calculated corrections to superallowed Fermi decay: New evaluation of the nuclear-structure-dependent terms. *Phys. Rev. C* **66**, 035501/13 (2002).

[16] J. C. Hardy and I. S. Towner. Superallowed $0^+ \rightarrow 0^+$ nuclear β decays: A critical survey with test of the conserved vector current hypothesis and the standard model. *Phys. Rev. C* **71**, 055501/29 (2005).

[17] G. Savard et al. $\mathcal{Q}$ value of the superallowed decay of ^{46}V and its influence on V_{ud} and the unitarity of the Cabibbo-Kobayashi-Maskawa matrix. *Phys. Rev. Lett.* **95**, 102501/4 (2005).

[18] D. Wilkinson. Fundamental interactions and nuclear structure: an impressionistic overview. *Nucl. Phys. A* **A374**, 649–664 (1982).

[19] I. S. Towner and J. C. Hardy. The evaluation of V_{ud}, experiment and theory. *J. Phys. G* **29**, 197–211 (2003).

[20] B. Tipton et al. A proposed measurement of the beta asymmetry in neutron decay with the los alamos ultra-cold neutron source. *AIP Conf. Proc.* **539** (2000).

[21] D. Počanić et al. Precise measurement of the $\pi^+ \rightarrow \pi^0 e^+ \nu$ branching ratio. *hep-ex/0312030* (2003).

[22] W. J. Marciano and A. Sirlin. Improved calculation of electroweak radiative corrections and the value of V_{ud}. *hep-ph/0510099* (2005).

[23] K. A. Olive, G. Steigman, and T. P. Walker. Primordial nucleosynthesis: theory and observations. *Phys. Rep.* **333–334**, 389–407 (2000).

[24] E. W. Kolb and M. S. Turner. *The Early Universe*. Perseus Publishing, 1990.

[25] Y. I. Izotov, F. H. Chaffee, C. B. Foltz, R. F. Green, N. G. Guseva, and T. X. Thuan. Helium abundance in the most metal-deficient blue compact galaxies: I Zw 18 and SBS 0335-052. *Astrophys. J.* **527**, 757 (1999).

[26] V. Luridiana, A. Peimbert, M. Peimbert, and M. Cervi no. The effect of collisional enhancement of Balmer lines on the determination of the primordial helium abundance. *Astrophys. J.* **592**, 846–865 (2003).

[27] K. A. Olive and E. D. Skillman. On the determination of the He abundance in extragalactic H_{II} regions. *New Astronomy* **6**, 119–150 (2001).

[28] G. J. Mathews, T. Kajino, and T. Shima. Big bang nucleosynthesis with a new neutron lifetime. *Phys. Rev. D* **71** (2005).

[29] M. S. Dewey, D. M. Gilliam, J. S. Nico, F. E. Wietfeldt, X. Fei, W. M. Snow, G. L. Greene, J. Pauwels, R. Eykens, A. Lamberty, and J. Van Gestel. Measurement of the neutron lifetime using a proton trap. *Phys. Rev. Lett.* **91**, 152302/1–4 (2003).

[30] J. Byrne, P. G. Dawber, C. G. Habeck, S. J. Smidt, J. A. Spain, and A. P. Williams. A revised value for the neutron lifetime measured using a penning trap. *Europhysics Letters* **33**, 187–192 (1996).

[31] W. Mampe, L. N. Bondarenko, V. I. Morozov, Yu. N. Panin, and A. I. Fomin. Measuring neutron lifetime by storing ultracold neutrons and detecting inelastically scattered neutrons. *JETP Letters* **57**, 82–87 (1993).

[32] V. V. Nesvizhevskiĭ, A. P. Serebrov, R. R. Tal'daev, A. G. Kharitonov, V. P. Alfimenkov, A. V. Streikov, and V. N. Shvetsov. Measurement of the neutron lifetime in a gravitational trap and analysis of experimental errors. *Soviet Physics JETP* **75**, 405–412 (1992).

[33] W. Mampe, P. Ageron, C. Bates, J. M. Pendlebury, and A. Steyerl. Neutron lifetime measured with stored ultracold neutrons. *Physical Review Letters* **63**, 593–596 (1989).

[34] P. E. Spivak. Neutron lifetime obtained from Atomic-Energy-Institute experiment. *Soviet Physics JETP* **67**, 1735–1740 (1988).

[35] Z. Chowdhuri, G. L. Hansen, V. Jane, C. D. Keith, W. M. Lozowski, W. M. Snow, M. S. Dewey, D. M. Gilliam, G. L. Greene, J. S. Nico, A. K. Thompson, and F. E. Wietfeldt. A cryogenic radiometer for absolute neutron rate measurement. *Rev. Sci. Inst.* **74**, 4280–4293 (2003).

[36] R. Golub, D. Richardson, and S. K. Lamoreaux. *Ultra-Cold Neutrons.* Adam Hilger, 1991.

[37] A. Saunders et al. Demonstration of a solid deuterium source of ultra-cold neutrons. *nucl-ex/0312021* (2003).

[38] Y. Masuda, T. Kitagaki, K. Hatanaka, M. Higuchi, S. Ishimoto, Y. Kiyanagi, K. Morimoto, S. Muto, and M. Yoshimura. Spallation ultracold-neutron production in superfluid helium. *Phys. Rev. Lett.* **89**, 284801 (2002).

[39] E. Korobkina, R. Golub, B. W. Wehring, and A. R. Young. Production of UCN by downscattering in superfluid ^{4}He. *Phys. Lett. A* **301**, 462–469 (2002).

[40] Yu. B. Zeldovich. *Sov Phys JETP* **9**, 1389 (1959).

[41] K.-J. Kügler, K. Moritz, W. Paul, and U. Trinks. Nestor: A magnetic storage ring for slow neutrons. *Nucl. Instr. Meth. A* **228**, 240–258 (1985).

[42] V. F. Ezhov, B. A. Bazarov, P. Geltenbort, N. A. Kovrizhnykh, G. B. Krygin, V. L. Ryabov, and A. P. Serebrov. Permanent-magnet trap for ultracold neutron storage. *Tech. Phys. Lett.* **27**, 64–70 (2001).

[43] P. R. Huffman, C. R. Brome, J. S. Butterworth, K. J. Coakley, M. S. Dewey, S. N. Dzhosyuk, R. Golub, G. L. Greene, K. Habicht, S. K. Lamoreaux, C. E. H. Mattoni, D. N. McKinsey, F. E. Wietfeldt, and J. M. Doyle. Magnetic trapping of neutrons. *Nature* **403**, 62–64 (2000).

[44] O. Zimmer. A method of magnetic storage of ultra-cold neutrons for a precise measurement of the neutron lifetime. *J. Phys. G.* **26**, 67–77 (2000).

[45] J. D. Bowman and S. I. Pentilla. On the measurement of the neutron lifetime using ultracold neutrons in a vacuum qaudrupole trap. *Journal of Research of the National Institute of Standards and Technology* **110**(4) (2005).

[46] K. Masuda. private communication.

[47] V. K. Ignatovich. *The physics of Ultracold Neutrons*, Volume 5 of *Oxford series on neutron scattering in condensed matter*. Oxford University Press, 1990.

[48] R. Golub and J. M. Pendlebury. Super-thermal sources of ultra-cold neutrons. *Phys. Lett.* **53A**, 133–135 (1975).

[49] L. Landau. Theory of the superfluidity of helium II. *Phys. Rev.* **60**, 356–358 (1941).

[50] C.-Y. Liu and A. R. Young. Ultra-cold neutron production in anti-ferromagnetic oxygen solid. *nucl-th/0406004* (2004).

[51] C. R. Brome. *Magnetic Trapping of Ultracold Neutrons*. PhD thesis, Harvard University, 2000.

[52] R. Golub, C. Jewell, P. Ageron, W. Mampe, B. Heckel, and I. Kilvington. Operation of a superthermal ultracold neutron source and the storage of ultracold neutrons in superfluid helium-4. *Z. Phys. B* **51**, 187–193 (1983).

[53] P. V. E. McClintock. An apparatus for preparing isotopically pure ^{4}He. *Cryogenics* **18**, 201–208 (1978).

[54] C. E. H. Mattoni. *Magnetic trapping of ultracold neutrons produced from a monochromatic cold neutron beam*. PhD thesis, Harvard University, 2002.

[55] D. N. McKinsey. *Detecting magnetically trapped neutrons: Liquid helium as a scintillator*. PhD thesis, Harvard University, 2002.

[56] S. N. Dzhosyuk. *Magnetic Trapping of Neutrons for Measurement of the Neutron Lifetime*. PhD thesis, Harvard University, 2004.

[57] F. Pobell. *Matter and Methods at Low Temperatures*. 2nd edition. Springer-Verlag, Berlin, 1996.

[58] D. R. Allum and P. V. E. McClintock. Condensation of isotopic impurities on the negative ion in He II. *J. Phys. C: Solid State Phys* **9**, L273 (1976).

[59] R. Golub and S. K. Lamoreaux. Neutron electric-dipole moment, ultracold neutrons and polarized ^{3}He. *Physics Reports* **237**, 1 (1994).

[60] P. C. Hendry and P. V. E. McClintock. Continuous flow apparatus for preparing isotopically pure ^{4}He. *Cryogenics* **27**, 131–138 (1987).

[61] L. Tisza. Transport phenomena in helium II. *Nature* **141**, 913 (1938).

[62] P. Kapitza. Viscosity of liquid helium below the λ-point. *Nature* **141**, 74 (1938).

[63] L. P. Mezhov-Deglin. An apparatus for producing pure ^{4}He. *Cryogenics* **12**, 311 (1971).

[64] P. P. Fatouros, D. O. Edwards, F. M. Gasparini, and S. Y. Shen. Isotropically pure He^4. *Cryogenics* **16**, 733–4 (1978).

[65] C. A. Baker, S. N. Balashova, J. Butterworth, P. Geltenbort, K. Greena, P. G. Harris, M. G. D. van der Grinten, P. S. Iaydjieva, S. N. Ivanova, J. M. Pendlebury, D. B. Shiers, M. A. H. Tucker, and H. Yoshiki. Experimental measurement of ultracold neutron production in superfluid ^{4}He. *Phys. Lett. A* **308**, 67–74 (2003).

[66] H. Yoshiki, H. Nakai, and E. Gutsmiedl. A new superleak to remove He^3 for UCN experiments. *Cryogenics* **45**, 399 (2005).

[67] The Neutron EDM Collaboration. A new search for the neutron electric dipole moment: Recent progress and design change. (2005).

[68] P. C. Tully. Isotropic purification of helium by differential distillation below the lambda-point. *US Bureau of Mines Report of Investigations 8054, Washington* (1975).

[69] S. K. Lamoreaux et al. Measurement of the ^{3}He mass coefficient in superfluid ^{4}He over the 0.45-0.95 K temperature range. *Europhysics Letters* **58**, 718 (2002).

[70] Web page, http://www.phy.anl.gov/atlas/index.html.

[71] J. E. Spencer and H. A. Enge. Split-pole magnetic spectrograph for precision nuclear spectroscopy. **49**, 181–193 (1967).

[72] R. Pardo. Status of the measurement of Helium-3 to Helium-4 ratios in isotopically purified helium. 2000.

[73] R. Pardo, D. P. Moehs, R. H. Scott, and R. C. Vondrasek. A modified ECR source design for ^{3}He Accelerator Mass Spectroscopy.

[74] R. Pardo. Summary of the January 2003 Helium-3 AMS run. 2003.

[75] R. Pardo. Summary of the September 2004 Helium-3 AMS run. 2005.

[76] L.-B. Wang, P. Mueller, R. Holt, Z.-T. Lu, T. P. O'Connor, Y. Sano, and N. C. Sturchio. Laser spectroscopic measurement of helium isotope ratios. *Geophysical Research Letters* **30**, 1592 (2003).

[77] G. C. Bjorklund. Frequency-modulation spectroscopy: a new method for measuring weak absorptions and dispersions. *Opt. Lett.* **5**, 15 (1980).

[78] J. L. Hall, L. Hollberg, T. Baer, and H. G. Robinson. Optical heterodyne saturation spectroscopy. *Appl. Phys. Lett.* **39**, 680 (1981).

[79] L.-B. Wang. *Determination of the Helium-6 Nuclear Charge Radius using High-resolution Laser Spectroscopy.* PhD thesis, University of Illinois at Urbana-Champaign, 2004.

[80] K. J. Coakley, J. M. Doyle, S. N. Dzhosyuk, P. R. Huffman, and L. Yang. Chaotic scattering and escape times of marginally trapped ultracold neutrons. *J. Res. Natl. Inst. Stand. Techol.* **110**, 367–376 (2005).

[81] E. L. Surkov, J. T. M. Walraven, and G. V. Shlyapnikov. Collisionless motion of neutral particles in magnetostatic traps. *Phys. Rev. A* **49**, 4478–4486 (1994).

[82] L. D. Landau and E. M. Lifshitz. *Mechanics.* 3rd edition. Butterworth-Heinemann Ltd, 1976.

[83] S. K. Lamoreaux. Note of the effect of lowering the trap potential to throw out marginally trapped neutrons. 1995.

[84] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. 2nd edition. Cambridge University Press, 1992.

[85] C. E. H. Mattoni. The precision measurement of the neutron lifetime using magnetically trapped neutrons: Marginally trapped neutrons and fluorescent time constants. Master's thesis, Harvard University, 1995.

[86] K. J. Coakley and B. Alpert. Personal communication., 1997.

[87] E. Ott and T. Tel. Chaotic scattering: An introduction. *Chaos* **3**, 4 (1993).

[88] G. L. Hansen. *A Radiometric Measurement of Neutron Flux in a Liquid ^{3}He Target*. PhD thesis, Indiana University, 2004.

[89] J. Escallier, M. Anerella, A. Ghosh R. Gupta M. Harrison A. Marone J. Muratore B. Parker W. Sampson J. Cozzolino, G. Ganetis, and P. Wanderer. Technology development for react and wind common coil magnets. *Proc. of the 2001 Particle Accelerator Conference, Chicago* , 214 – 16 (2001).

[90] K. Endo Y. Morita K. Tsuchiya, K. Egawa and N. Ohuchi. Performance of the eight superconducting quardrupole magnets for the tristan low-beta insertions. *IEEE Transactions on Magnetics* **27**(2), 1940 (1991).

[91] R. H. Ericksen. Creep of aromatic polyamide fibres. *Polymer* , 733 (1985).

[92] Sandor Feher. Private communication.

[93] H. London. In *Proceedings of the International Conference on Low-Temperature Physics*, 151. Clarendon Laboratory, 1951.

[94] O. V. Lounasmaa. *Experimental principles and methods below 1 K*. Academic Press, 1974.

[95] J. S. Butterworth, C. R.Brome, P. R. Huffman, C. E. H. Mattoni, D. N. McKinsey, and J. M. Doyle. A removable cryogenic window for transmission of light and neutrons. *Rev. Sci. Inst.* **69**, 3998 (1998).

[96] D. N. McKinsey, C. R.Brome, J. S. Butterworth, R. Golub, K. Habicht, P. R. Huffman, S. K. Lamoreaux, C. E. H. Mattoni, and J. M. Doyle. Fluorescence efficiencies of thin scintillating films in the extreme ultraviolet spectral region. *Nucl. Instr. Meth. B* **132**, 351 (1997).

[97] C. O'Shaughnessy. Guideit simulations of neutron trapping experiment.

[98] R. Golub. On the storage of neutrons in superfluid ^{4}He. *Phys. Lett.* **72A**, 387–390 (1979).

[99] M. N. Wilson. *Superconducting magnets*. Oxford University Press, 1983.

[100] A. T. Visser. A 10000 A 1000 VDC solid state dump switch. In *EPE '91. 4th European Conference on Power Electronics and Applications,* P. Ferraris, editor, 521–26, 1991.

[101] J. M. Lock. Optimization of current leads into a cryostat. *Cryogenics* **9**, 438 – 442 (1969).

[102] R. C. Niemann, Y. S. Cha, and J. R. Hull. Performance measurements of superconducting current leads with low helium boil-off rates. *IEEE Trans. Appl. Superconductivity* **3(1)**, 392 (1993).

[103] G. Citver, S. Feher, P. J. Limon, D. Orris, T. Peterson, S. Sylvester, M. A. Tartaglia, and J. C. Tompkins. HTS power leads test results. *Proc. of the 1999 Particle Accelerator Conf.* , 1420–22 (1999).