Proposal for an Experiment at the Spallation Neutron Source

Precise Measurement of the Neutron Beta Decay Parameters "*a*" and "*b*"

The *Nab* Experiment

R. Alarcon, S. Balascuta
Department of Physics, Arizona State University, Tempe, AZ 85287-1504

A. Klein, W.S. Wilburn
Los Alamos National Laboratory, Los Alamos, NM 87545

M.T. Gericke
Department of Physics, University of Manitoba, Winnipeg, Manitoba, R3T 2N2 Canada

J.R. Calarco, F.W. Hersman
Department of Physics, University of New Hampshire, Durham, NH 03824

A. Young
Department of Physics, North Carolina State University, Raleigh, NC 27695-8202

J.D. Bowman (Spokesman), T.V. Cianciolo, S.I. Penttilä, K.P. Rykaczewski, G.R. Young
Physics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

V. Gudkov
Department of Physics and Astronomy, University of South Carolina, Columbia, SC 29208

G.L. Greene, R.K. Grzywacz
Department of Physics and Astronomy, University of Tennessee, Knoxville, TN 37996

L.P. Alonzi, S. Baeßler, M.A. Bychkov, E. Frlež, A. Palladino, D. Počanić (Spokesman)
Department of Physics, University of Virginia, Charlottesville, VA 22904-4714

9 November 2007

Abstract: We propose to perform a precise measurement of a, the electron-neutrino correlation parameter, and b, the Fierz interference term in neutron beta decay, in the Fundamental Neutron Physics Beamline at the SNS, using a novel electric/magnetic field spectrometer and detector design. The experiment is aiming at the 10^{-3} accuracy level in $\Delta a/a$, and will provide an independent measurement of $\lambda = G_A/G_V$, the ratio of axial-vector to vector coupling constants of the nucleon. We will also perform the first ever measurement of b in neutron decay, which will provide an independent limit on the tensor weak coupling.

1. Scientific motivation

Neutron β decay, $n \to pe\bar{\nu}_e$, is one of the basic processes in nuclear physics. Its experimental study provides the most sensitive means to evaluate the ratio of axial-vector to vector coupling constants $\lambda = G_A/G_V$. The precise value of λ is important in many applications of the theory of weak interactions, especially in astrophysics; e.g., a star's neutrino production is proportional to λ^2. More precise measurements of neutron β-decay parameters are also important in the search for new physics. Measurement of the neutron decay rate Γ, or lifetime $\tau_n = 1/\Gamma$, allows a determination of V_{ud}, the u-d Cabibbo-Kobayashi-Maskawa (CKM) matrix element, independent of nuclear models, because Γ is proportional to $|V_{ud}|^2$, as seen in the leading order expression:

$$\Gamma = \frac{1}{\tau_n} = \frac{f^R m_e^5 c^4}{2\pi^3 \hbar^7} \left(|G_V|^2 + 3|G_A|^2\right) \propto |G_V|^2 \left(1 + 3|\lambda|^2\right) = |V_{ud}|^2 \, |g_V|^2 \, G_F^2 (1 + 3|\lambda|^2)\,, \quad (1)$$

where $f^R = 1.71482(15)$ is a phase space factor, m_e is the electron mass, $g_{V,A}$ the vector and axial-vector weak nucleon form factors at zero momentum transfer, respectively, and G_F is the fundamental Fermi weak coupling constant. While the conserved vector current (CVC) hypothesis fixes g_V at unity, two unknowns, V_{ud} and λ, remain as variables in the above expression for Γ. Hence, an independent measurement of λ is necessary in order to determine V_{ud} from the neutron lifetime. Several neutron decay parameters can be used to measure λ; they are discussed below. Precise knowledge of V_{ud} helps greatly in establishing the extent to which the three-generation CKM matrix is unitary. CKM unitarity, in turn, provides an independent crosscheck of the presence of certain processes and particles not included in the Standard Model (SM) of elementary particles and interactions, i.e., an independent constraint on new physics.

Currently, the most accurate value of the CKM matrix element V_{ud} is obtained from measurements of $0^+ \to 0^+$ nuclear β-decays, the so-called superallowed Fermi transitions [1]. However, the procedure of the extraction of V_{ud} involves calculations of radiative corrections for the Fermi transition in nuclei. Despite the fact that these calculations have been done with high precision (see [2] and references therein), it is impossible to verify the values of these nuclear corrections from independent experiments.

A problem with CKM matrix unitarity at the $2-3\sigma$ level persisted for over two decades in the first row sum; e.g., the 2002 Review of Particle Properties [3] reported values of CKM matrix elements that yield

$$\Delta \equiv 1 - |V_{ud}|^2 - |V_{us}|^2 - |V_{ub}|^2 = 0.0032 \pm 0.0014\,. \quad (2)$$

The situation changed drastically in 2003 and 2004 when a series of experiments at Brookhaven, Fermilab and CERN reported revised values of K_{l3} decay branching ratios which led to an upward adjustment, by about 2.5σ, of the CKM matrix element V_{us} [4, 5, 6]. Without getting into the details of this revolutionary development, it will suffice to note that a revised CKM unitarity check yields [7, 1]

$$\Delta = (8 \pm 15) \times 10^{-4}\,, \quad (3)$$

which indicates that, at least for the time being, the question of the CKM matrix unitarity appears to be closed. However, several questions related to V_{ud} still remain open. Firstly,

it is desirable to have an independent check of the superallowed Fermi nuclear beta decay result. Secondly, a disturbing inconsistency persists between the best results on neutron decay and those on nuclear Fermi decays, as well as within the body of the neutron decay data. Finally, by its nature, neutron decay offers redundant consistency checks whose failure can be an indication of new physics. In order to discuss the last two points we turn to the details of neutron decay dynamics.

Neglecting nucleon recoil, as well as radiative and loop corrections, the triple differential neutron decay rate is determined by the decay parameters a, b, A, B, and D, as shown:

$$\frac{dw}{dE_e d\Omega_e d\Omega_\nu} \propto p_e E_e (E_0 - E_e)^2 \times \left[1 + a\frac{\vec{p}_e \cdot \vec{p}_\nu}{E_e E_\nu} + b\frac{m_e}{E_e} + \langle \vec{\sigma}_n \rangle \cdot \left(A\frac{\vec{p}_e}{E_e} + B\frac{\vec{p}_\nu}{E_\nu} + D\frac{\vec{p}_e \times \vec{p}_\nu}{E_e E_\nu} \right) \right], \quad (4)$$

where $p_{e(\nu)}$ and $E_{e(\nu)}$ are the electron (neutrino) momenta and energies, respectively, E_0 is the electron energy spectrum endpoint, and $\vec{\sigma}_n$ is the neutron spin. The "lowercase" parameters: a, the electron–neutrino correlation parameter, and b, the Fierz interference term, are measurable in decays of unpolarized neutrons, while the "uppercase" parameters, A, B, and D, require polarized neutrons. All except b depend on the ratio $\lambda = g_A/g_V$, in the following way (given here at the tree level):

$$a = \frac{1 - |\lambda|^2}{1 + 3|\lambda|^2}, \qquad A = -2\frac{|\lambda|^2 + Re(\lambda)}{1 + 3|\lambda|^2}, \qquad B = 2\frac{|\lambda|^2 - Re(\lambda)}{1 + 3|\lambda|^2}, \qquad D = \frac{2\, Im(\lambda)}{1 + 3|\lambda|^2}. \quad (5)$$

Here we have allowed for the possibility of a complex λ, i.e., a nonzero value of D, the triple correlation coefficient, which would arise from time reversal symmetry violation in the process. Since measurements of D are consistent with zero [1], we will treat λ as real. Given that $\lambda \simeq -1.27$, parameters A and a are similarly sensitive to λ:

$$\frac{\partial a}{\partial \lambda} = \frac{-8\lambda}{(1 + 3\lambda^2)^2} \simeq 0.30, \qquad \frac{\partial A}{\partial \lambda} = 2\frac{(\lambda - 1)(3\lambda + 1)}{(1 + 3\lambda^2)^2} \simeq 0.37, \quad (6)$$

while B is relatively insensitive:

$$\frac{\partial B}{\partial \lambda} = 2\frac{(\lambda + 1)(3\lambda - 1)}{(1 + 3\lambda^2)^2} \simeq 0.076. \quad (7)$$

Experimental status of the above parameters is summarized in the Particle Data Group's review in ref. [1]. It has been true for some time that by far the best precision in extracting λ has been achieved in measurements of A, the correlation between the electron momentum and neutron spin. However, the experimental status of A and λ is far from satisfactory, as shown in Fig. 1. In both cases the error on the weighted average value was rescaled up by a factor of 2 or more because of an uncommonly bad χ^2 value and low confidence levels for the fits and extracted weighted averages. It is particularly disturbing that the most accurate measurement to date, that made by the PERKEO II collaboration [8], is in significant disagreement with the remaining world data set. This disagreement carries over

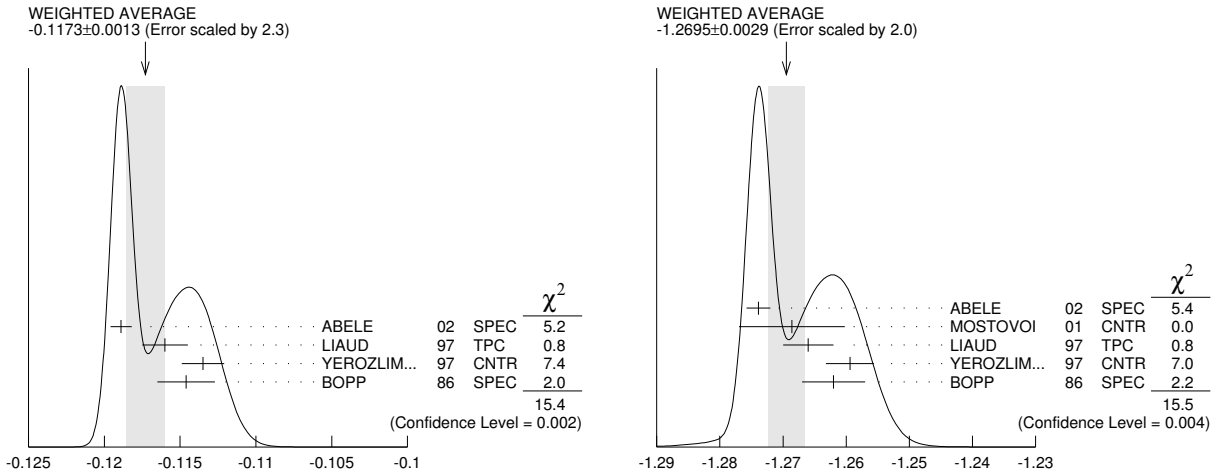

Figure 1: Particle Data Group's most recent compilation of experimental values of A (left panel) and λ (right panel), see ref. [1]

naturally to the value of V_{ud} [1]. A new measurement of A by PERKEO II confirms this finding with significantly reduced uncertainties [10].

Present inconsistencies in the value of A must, and will be, resolved by new measurements. Meanwhile, parameter a offers an independent measure of λ with comparable sensitivity and radically different systematics. The current world average value of $a = -0.103 \pm 0.004$ is dominated by two 5 % measurements [11, 12], to be compared with the 0.06 % measurement of A in PERKEO II [8].

We are proposing to make a measurement of the e–ν_e correlation parameter a with an accuracy of a few parts in 10^3 using a novel 4π field-expansion spectrometer in the FNPB line at the SNS. The spectrometer and our measurement method are discussed in detail in the next section. The experiment, which we have named 'Nab', will also produce the first measurement of b, the Fierz interference term; to date b has not been measured in neutron decay.

The Nab experiment constitutes a first phase of a program of measurements that will continue with a second-generation experiment named 'abBA', which will measure the polarized neutron decay parameters A and B in addition to a and b. Together, the two experiments form a complete program of measurements of the main neutron decay parameters in a single apparatus with shared systematics and consistency checks. The experiments are complementary: Nab is highly optimized for the measurement of a and b, while abBA focuses on A and B with a lower-precision consistency check of the a and b.

Two other experiments, currently under way, aim to improve the experimental precision of a in neutron decay, aCORN at NIST [9] and aSPECT at Munich [30]. Nab, aCORN and aSPECT all differ in important aspects of their systematics. Given the very challenging nature of the measurement of a in neutron decay, it is important to have multiple independent precise determinations of the parameter.

The scope of the Nab experiment extends well beyond resolving the remaining λ and V_{ud} inconsistencies. The relevance of precise determination of beta decay parameters, in

particular those of the neutron, to searches for signals of physics beyond the Standard Model has been recently discussed in great detail by in Refs. [13, 14]. At the proposed accuracy level, parameter a can be used to constrain certain left-right symmetric models (L-R models) as well as leptoquark extensions to the SM. The latter would also be constrained by our measurement of b which is sensitive to a tensor weak interaction that has often been linked to leptoquarks. There have also been proposals of models relying on a new anomalous chiral boson to account for a tensor interaction [15]. The sensitivities of a to, e.g., L-R model parameters $\bar{a}_{RL}$, $\bar{a}_{RR}$, etc., are competitive and complementary to those of A and B. A general connection between non-SM (e.g., scalar, tensor) $d \to ue\bar{\nu}$ interactions on the one hand, and upper limits on the neutrino mass on the other, was recently brought to light [16], providing added motivation for more precise experimental neutron decay parameters. A detailed up to date review of the supersymmetric contributions to the weak decay correlation parameters, in particular to the beta decay correlation parameters with a discussion of the theoretical implications of their precise measurement, was given by Profumo, Ramsey-Musolf and Tulin [17].

The Nab experiment brings about an interesting opportunity to perform a new independent test of the CVC hypothesis and of the absence or presence of second class currents through the measured dependence of the neutron decay parameters a (Nab) and A (abBA) on electron energy, E_e. Gardner and Zhang [18, 19] have shown that measurements of A and a at the 0.1 % accuracy along with their dependence on E_e would provide a powerful test of both the weak magnetism and induced tensor terms at an unprecedented level. Specifically, under those circumstances the error in f_2/f_1 and g_2/f_1 would be 2.5 % and 0.13λ, respectively. Here f_1, f_2 and g_2 are the vector, weak magnetism and induced tensor couplings, respectively. Presently weak magnetism and second class currents remain unresolved in nuclear beta decays. Thus, the proposed set of experiments would test with previously inaccessible precision the CVC hypothesis and presence of second class currents, as well as the very structure of the interaction terms underpinning the $V - A$ theory [18, 19]. Needless to say, this opens the way to test various models of "new" physics with increased precision.

2. Technical approach to measurement

We propose to measure the electron-neutrino correlation in neutron beta decay, a, using a novel approach. Coincidences between electrons and protons are detected in a field-expansion spectrometer. The purpose of the field expansion is to measure the magnitude of the proton momentum, p_p. The electron–neutrino correlation, a, expresses the dependence of the decay rate on the angle between the neutrino and electron,

$$\cos\theta_{e\nu} = \cos\theta_e \cos\theta_\nu + \sin\theta_e \sin\theta_\nu \cos(\phi_e - \phi_\nu)\,. \tag{8}$$

It is not necessary to measure all the above angles because $\theta_{e\nu}$ can be determined from the electron energy and the proton momentum squared. The electron energy is measured precisely in the Si detectors. The electron and neutrino momenta, p_e and p_ν, respectively, can be determined from the measured electron and proton energies. The quantity $\cos\theta_{e\nu}$ can then be determined from the proton momentum and the electron energy using

$$p_p^2 = p_e^2 + 2p_e p_\nu \cos\theta_{e\nu} + p_\nu^2\,. \tag{9}$$

The relation between proton momentum and electron energy is best illustrated in the phase space plot shown in Fig. 2. The crucial property of the plot in Fig. 2 is that phase space

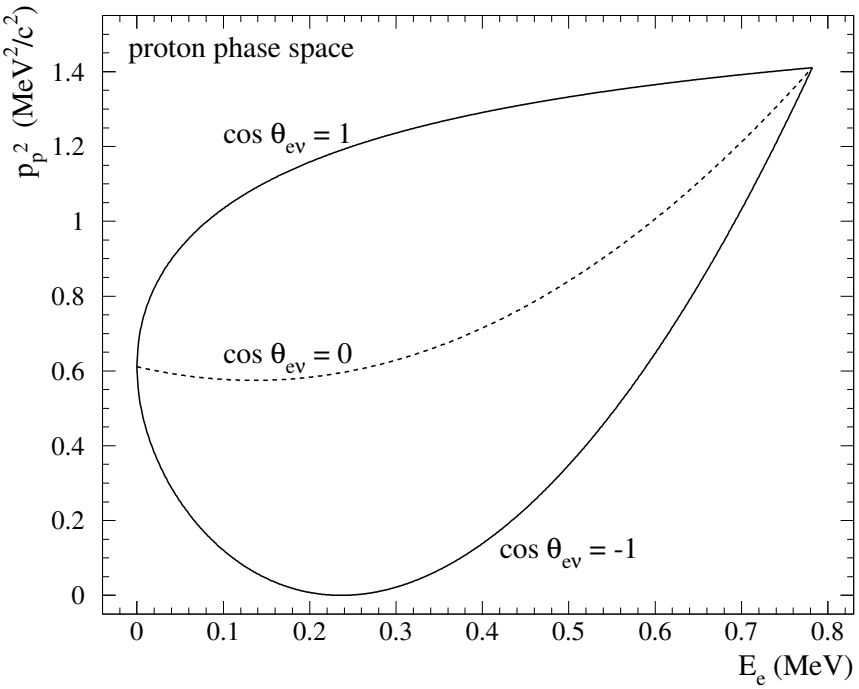

Figure 2: Available proton phase space (in terms of $p_{\rm p}^2$, proton momentum squared) in neutron beta decay as a function of $E_{\rm e}$, electron kinetic energy. The upper bound of the allowed phase space is determined by the condition that the electron and neutrino momenta are colinear, $\cos\theta_{e\nu} = 1$, while they are anticolinear, $\cos\theta_{e\nu} = -1$, at the lower bound. The central dashed parabola corresponds to the condition that e and ν momenta are orthogonal; events falling on this curve are insensitive to the correlation parameter a, while those at the upper and lower bounds exhibit maximal sensitivity to a. It is critical to note that with $a = 0$ the probability distribution of $p_{\rm p}^2$ for a given electron energy $E_{\rm e}$ would be uniform, i.e., a flat rectangular box spectrum.

distributes proton events evenly in $p_{\rm p}^2$ between the lower and upper bounds for any fixed value of E_e. Given the relationship between $p_{\rm p}^2$ and $\cos\theta_{e\nu}$, it is clear that the slope of the $p_{\rm p}^2$ distribution is determined by the correlation parameter a; in fact it is given by βa, where $\beta = v_{\rm e}/c$ (see Fig. 3. This observation leads to the main principle of measurement of a: a is determined from the slopes of the $1/t_{\rm p}^2$ distributions for different values of $E_{\rm e}$, where $t_{\rm p}$ represents the measured proton time of flight in a suitably constructed magnetic spectrometer. If $a = 0$ all distributions would have a slope of zero. Having multiple independent measurements of a for different electron energies provides a powerful check of systematics, as discussed below.

A perfect spectrometer would produce rectangular distributions of $1/t_{\rm p}^2$ with sharp edges. The precise location of these edges is determined by well-defined kinematic cutoffs that only depend on $E_{\rm e}$. However, a realistic time-of-flight spectrometer will produce imperfect measurements of the proton momenta due to the spectrometer response function, discussed

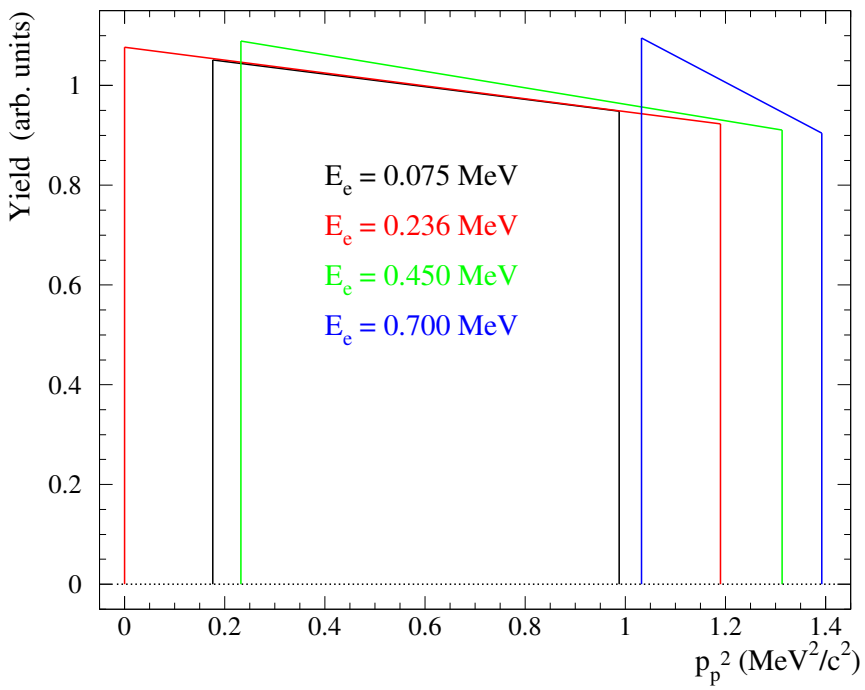

Figure 3: A plot of proton yield for four different electron kinetic energies for an ideal spectrometer. The spectrometer is assumed to have perfect time resolution. It is assumed that $t_{\rm p}/p_{\rm p} = $ const. as would be the case if the electric field were zero. The value $a = -0.105$ is assumed. If a were 0, all the distributions would have a slope of 0.

in detail in Sect. 4. The measured locations and shapes of edges in $1/t_{\rm p}^2$ distributions will allow us to examine the spectrometer response function and verify that the fields have been measured correctly.

The main requirements on a suitable magnetic spectrometer are:

1. The spectrometer and its magnetic and electrical fields must be azimuthally symmetric about the central axis, z.

2. Neutrons must decay in a region of large magnetic field. The decay protons and electrons spiral around a magnetic field line (the guiding center approximation).

3. An electric field is required to accelerate the proton from the eV-range energies to a detectable energy range prior to reaching the detector. This field will, however, impose a lower energy threshold on electron detection.

4. The momentum of the proton rapidly becomes parallel to the magnetic field as the latter expands. Between the point in z right after the field expansion and the point where the electric field begins, the proton time of flight, $t_{\rm p} \simeq lm/|\vec{p}_{\rm p}|$, and this contribution dominates the total time of flight.

For a perfect determination of the proton momentum, $p_{\rm p}$ and $\cos\theta_{{\rm e}\nu}$, the error in a

becomes

$$\sigma_a^{\min} = \sqrt{\frac{3}{N\beta_{\rm ave}^2}} = \frac{2.3}{\sqrt{N}}\,. \tag{10}$$

The reference P configuration design, described below, achieves $\sigma_a = 2.4/\sqrt{N}$, as shown in Section 4.3.

The basic concept of the spectrometer consists of a superconducting solenoid with its longitudinal axis oriented normal to the neutron beam, which passes through the center of the solenoid. The strength of the solenoidal magnetic field at the position of the neutron beam is 4 T, expanding to 0.1 T at either end of the solenoid. Inside the solenoid is a second concentric cylindrical solenoid plus cylindrical electrodes (consisting of three sections) maintaining the neutron decay region at a potential of +30 kV with respect to the ends of the solenoid where detectors are placed at ground potential.

The magnetic field strength is sufficiently high that both the electrons and protons from neutron decay are constrained to spiral along the magnetic field lines with the component of the spiral motion transverse to the field limited by cyclotron radii of the order of a few millimeters.

Thus, two segmented Si detectors, one at each end of the solenoid, view both electrons and protons in an effective 4π geometry. The time of flight between the electron and proton is accurately measured in a long, $l \sim 1.5$ meter, drift distance. The electron energy is accurately measured in the Si detectors. The proton momentum and electron energy determine the electron–neutrino opening angle. We note that by sorting the data on proton time of flight and electron energy, a can be determined with a statistical accuracy of $\sim 2.4/\sqrt{N}$, where N is the number of decays observed.

In addition to excellent statistical sensitivity, the approach has a number of advantages over previous measurements. The acceptance of the spectrometer is 4π for both particles. Thin-dead-layer segmented Si detectors as well as all other components in the apparatus, are commercially available. There are no material apertures to determine the acceptance of the apparatus. The charged particles interact only with electric and magnetic fields before striking the detectors. Coincident detection of electrons and protons reduces backgrounds, and allows the *in situ* determination of backgrounds. A time of flight spectrum is obtained for each electron energy. Different parts of the spectra have different sensitivities to a. The portions of the time-of-flight spectra that are relatively insensitive to a (cf. Fig. 2) will be used to verify the accuracy of the spectrometer response function, which is based on electric and magnetic field determinations.

Although two configurations (labeled "P" and "PZ," respectively) of the spectrometer were originally considered, only one, the former, will be used. The P configuration of the field-expansion spectrometer is designed to make the momentum of the proton inversely proportional to the proton time of flight, $|\vec{p}_{\rm p}| \propto 1/{\rm TOF}$. In the P configuration there is a small probability of order 1 % that the momentum direction will be reversed and the TOF increased.

A not-to-scale design for the P configuration of the field expansion spectrometer is shown in Fig. 4. Electrons and protons spiral around magnetic field lines and are guided to two segmented Si detectors, each having a $\sim$100 cm^2 active area. In the center of the spectrometer

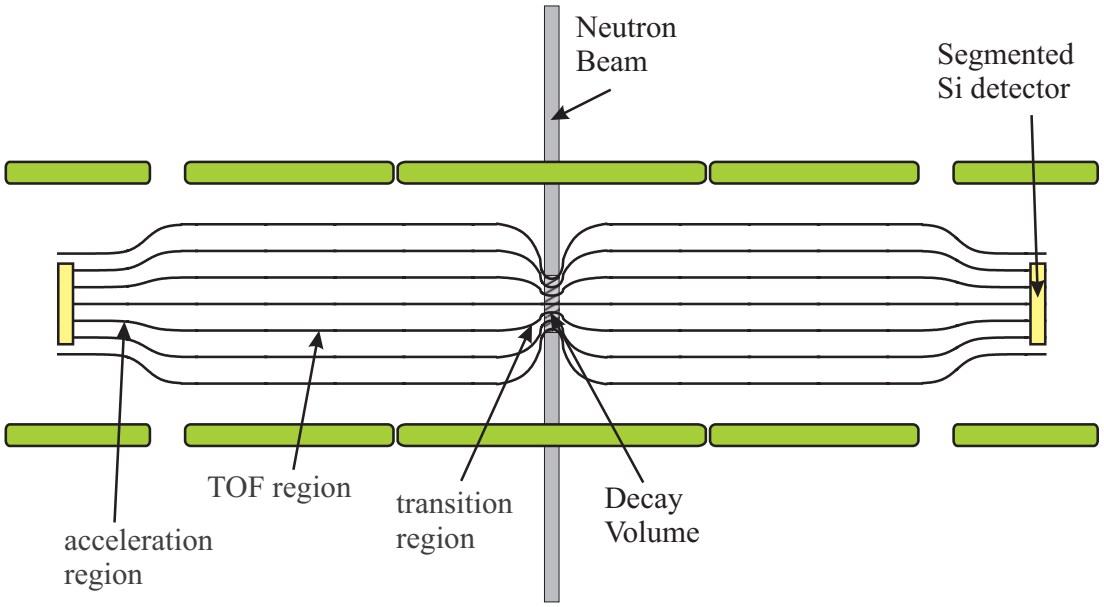

Figure 4: A schematic view of the field expansion spectrometer showing the main regions of the device: (a) neutron decay region, (b) transition region with expanding magnetic field, (c) drift (TOF) region, and (d) the acceleration region before the detector.

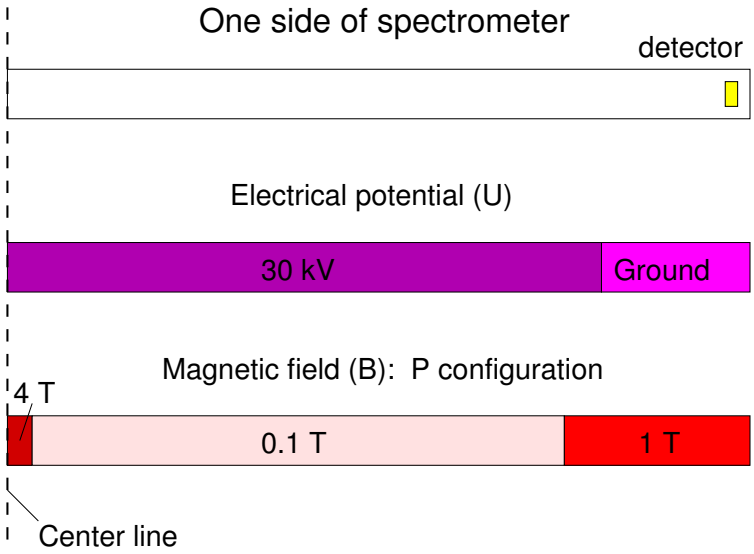

Figure 5: Scale drawing of one side of the field-expansion spectrometer for the P and PZ configurations. In the P configuration, the proton momentum is longitudinalized by the field expansion and $|\vec{p}_{\rm p}| \propto 1/{\rm TOF}$. In the PZ configuration the z component of the proton momentum is unchanged until the field expands just before the electric acceleration, and $|p_{{\rm p},z}| \propto 1/TOF$. The distance from the center of the spectrometer to either detector is 2 m.

the field strength is 4 T, in the drift region 0.1 T, and near the Si detectors 1 T (see Figs. 5 and 6).

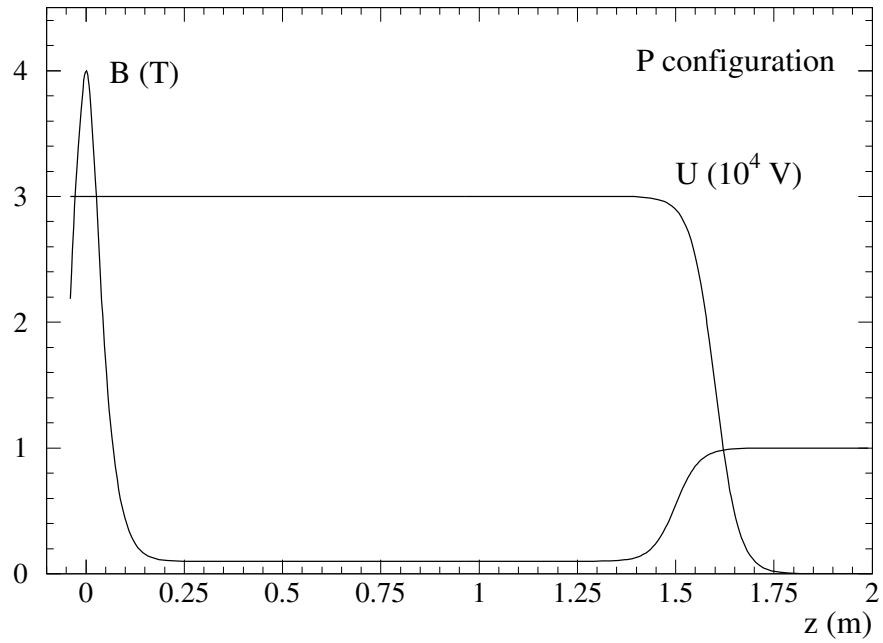

Figure 6: Electrical potential (U) and magnetic field (B) on the spectrometer axis for the P configuration.

The field expansion decreases the angle between the momentum and the magnetic field lines. The proton speed in the drift region is close to $|\vec{p}|/m_p$. The particles strike the detectors at approximately normal angles, thus reducing the probability of backscattering. An electric field is applied to the particles before they strike the Si detectors so that the protons have enough energy to be detected, while the energy of the electrons is reduced. The electric field must be applied after the magnetic field expansion so that the electrons acceptance does not depend on electron energy above a threshold. For the reference design, all electrons that have energies above 70 keV reach the detectors and deposit at least 30 keV. After the drift region the protons are electrostatically accelerated from eV–range energies to 30 keV as they cross a narrow gap in the cylindrical electrode so that the time spent between the potential change and the detector is small compared to the time spent in the drift region. Electrons may be scattered from the Si detectors, but scattered electrons are guided back to one of the detectors and eventually all of the electrons' energy is deposited in the detectors.

As the charged particle trajectories are constrained to follow the magnetic field lines, the segmented Si detectors form a projected image of the beam. The ends of the decay region are defined by the image of the beam on the detectors. The transverse migration of back scattered electrons is small because the radius of gyration is small (a few mm) and because the momentum of the electron decreases with each reflection.

These basic properties of the spectrometer greatly help in the identification of electron and proton pairs stemming from the same neutron decay. Correlated electron-proton pairs will be separated in time by several microseconds. For the favorable field configurations under study, the time of flight separation between electrons and protons exceeds about 10 μs for just several percent of the events. Furthermore, the imaging nature of the spectrometer

insures that the correlated electrons and protons impact the detector surface at the same or neighboring pixel, including, of course, the mirror pixel for events in which the two particles go in opposite directions.

Magnetic field mapping

A reliable map of the magnetic field of the spectrometer is essential for the understanding of systematic effects. While field mapping is conceptually quite simple, there are two practical issues that present a challenge for the Nab magnet: (1) choice of the field probe, and (2) accurate positioning of the probe within the magnet.

Choice of field probe

Two types of magnetic field probes are commercially available and commonly used for field mapping in high fields. These are NMR probes and select Hall effect probes. NMR probes possess extreme accuracy but wouldn't work in the high field gradients found in the Nab magnet. Hall probes are less precise and are extremely convenient to use but require auxillary calibration if significant accuracy is required. It is our intention to map the field using Hall effect probes and to carry out the necessary off-line probe calibration. This procedure has been successfully demonstrated in the aSPECT experiment [20] which employs a magnet of scale similar to that of abBA.

Accurate positioning of probe

In order to map the field the probe must be inserted into the ends of the magnet and then physically moved to different locations whose positions are accurately known. To access the entire volume of the magnet, the probe must be supported on a long member (≥ 2 m). While there is no difficultly in accurately moving one end of the support member externally to the magnet, it will be quite challenging to insure with confidence that the member is sufficiently rigid to insure that the probe's position at the other end is well known. We propose to use a different technique in which we do not depend upon the rigidity of the support. Instead we will use commercial laser ranging surveying technology to accurately measure the position of the probe itself. This will be accomplished by attaching a retro reflector to the hall probe. At each field measurement position, an accurate determination of position the probe itself, in 3 degrees of freedom, will be made. Commercial 3D laser ranging surveying instruments (known colloquially as "Total Stations") are routinely capable of reaching accuracies below 100 micron at distances of several meters. Such systems are quite costly but the SNS survey group has several and will participate in this work.

The field mapping must be done with the magnet cold. This implies that we will be required to insert a long "thimble" temporarily into the magnet from the end to allow the insertion of the probe in a room temperature environment.

We also note that it will ultimately be necessary relate the field map coordinate system to external references on the magnet to allow accurate positioning of the magnet with respect to the neutron beam. The use of the "total station" will allow this to be done with the same instrumentation as the field map.

Event and data rate

The event count rate at the SNS at 1.4 MW operation is 19.5 counts/sec/cm^3 of fiducial volume. A 20 cm^3 fiducial volume (box with $2 \times 2.5\,\text{cm}^2$ base and $h = 2$ cm) is easily attainable leading to an event rate in excess of 400 Hz. For example, the electron-neutrino

correlation can be determined with a statistical uncertainty of 0.2 % in a typical run of 7×10^5 s, or about ten days. We plan to have several such runs, thus further substantially reducing the uncertainties. The statistical uncertainty in a would be 0.0006 as compared to 0.004 in the Particle Data Listings. As discussed in the preceding section, the uncertainty in G_A/G_V in the Particle Data Listings is based on inconsistent data on A, the electron momentum–neutron spin correlation in neutron beta decay.

The hermetic nature of the electron energy measurement provides a clean and precise measurement of the electron energy spectrum, leading to an excellent determination of b, the Fierz interference term. The Fierz interference term, never before measured in neutron decay, modifies the shape of the electron spectrum. The statistical uncertainty in b is higher than that for a, because the quantity m_e/E_e is strongly correlated with the normalization of the beta spectrum for kinetic energies larger than approximately half the electron mass. The statistical uncertainty in b is given by $\Delta b_\mathrm{stat} = 10.1/\sqrt{N}$ for an electron energy threshold of 0.1 MeV. Hence, in a typical 7×10^5 s run we would expect $\Delta b_\mathrm{stat} \sim 7 \times 10^{-4}$. The $V - A$ Standard Model predicts $b = 0$. We expect to collect several samples of 10^9 events in several 6-week runs. The large event rates make it possible to study systematic uncertainties and achieve small statistical uncertainties in moderate run times.

To date, the best information on G_A/G_V has come from measurements of A, the electron–neutron spin correlation. In order to measure A it is necessary not only to determine the neutron polarization, but also which of the two detectors the electron struck first. This determination may be imperfect due to electron back-scattering. The electron–neutrino opening angle depends on the square of the proton momentum and it is therefore not necessary to determine the relative direction of the electron and proton in order to measure the electron–neutrino correlation; the TOF and electron energy are sufficient. The practical implication of combining the two directions is important. It is possible to obtain commercially segmented Si detectors with thin ion-implanted entrance windows. The large sheet resistance of the ion-implanted junction and the large rise time ($\sim$ 50 ns) make fast timing difficult. The ability to use slow Si detectors makes the experiment feasible without having to resort to new technology.

In order to optimize our design and to study the systematics in detail, we have developed a realistic Monte Carlo simulation of the spectrometer using the standard detector simulation package GEANT4 [21]. This approach allows us to test with high precision the effect of changes or uncertainties of any parameter in the apparatus, and verify the validity of our analytical calculations of the same.

While a measurement of a mainly requires the proton TOF information and uses the electron signal primarily as a time marker, the measurement of b relies entirely on a precise determination of the electron energy spectrum. In this way, the two measurements are complementary. Accurate measurements of both proton TOF and electron energy provide us with means to evaluate multiple independent cross-checks of the systematic uncertainties in both a and b.

3. The detector

The detector design is a challenging issue for any precise neutron beta decay experiment. The detector has to be able to stop and detect the full energy of 50–750 keV electrons as

well as 30 keV protons. This requires the detector thickness to be about 2 mm Si-equivalent, a very thin window technology, and very low energy threshold for detecting signals down to about 5 keV.

The very thin window/dead-layer should uniformly cover a large area of $\approx 100\,\mathrm{cm}^2$. The detector has to be segmented into about 100 elements. The segmentation has to be applied on the back side to keep the irradiated front side homogeneous. The segmentation is necessary to determine the particle position and thus identify the electron/proton trajectory. The time and spatial pattern of electron energy deposition has to be measured. The detector segmentation has to be combined with pulse processing electronics allowing for the real time signal recording with a resolution at the level of 100 ps. The low energy threshold is related to a good energy resolution, at the level of few keV for the relevant energy range of electrons and protons.

Cooled silicon detector has the optimal combination of efficiency, stability, energy resolution and timing resolution unsurpassed by other types of detector, some of which may excell in one of the above characteristics, but not in all.

The design goal, pursued in a collaboration with Micron Semiconductor Ltd., has been to build a large area segmented single wafer silicon detector, about 2 mm thick to enable stopping the electrons, and operating with a liquid nitrogen cooling at the temperature level of about 100 K. The readout will be implemented using cold-FET preamplifier and real-time digital signal processing electronics. A prototype is discussed below.

The charged-particle detectors for the Nab/abBA spectrometer will be made from 15 cm diameter, 2 mm thick silicon wafers. Charged particles will enter the detector through the junction side. Charge deposited by the particles will be collected on the ohmic side. The active area of the detector will be segmented into 127 individual elements. A sketch of the design of the segmented ohmic side of the detector is shown in Fig. 7.

A hexagonal array of detector elements is chosen for several reasons.

1. Hexagons efficiently fill the circular area of the detector,
2. they match the image of the decay volume well,
3. only three detector elements meet at a vertex, reducing the number of elements involved in a charge-sharing event, and
4. the number of adjacent elements that must be searched for the partner particle or reflected electron events is minimized.

The hexagonal detector elements in the preliminary design have sides of length $s = 5.2\,\mathrm{mm}$ and areas of $a = 0.70\,\mathrm{cm}^2$ There are several reasons for this choice. First, the maximum radius of gyration at the detector is 2.2 mm for the electrons and 2.3 mm for the protons. Therefore, the electron-proton separation on the detector can never be more than 4.5 mm. Our choice of $s = 5.2\,\mathrm{mm}$ guarantees the electron is never more than one detector element away from the proton. This means that only 14 detector elements (including conjugate elements on the opposing detector) need to be considered in constructing a coincidence event. Similarly, only 14 elements need be considered in searching for an event where an electron reflects from a detector and then stops, either in the same detector or the opposing detector. Second, the

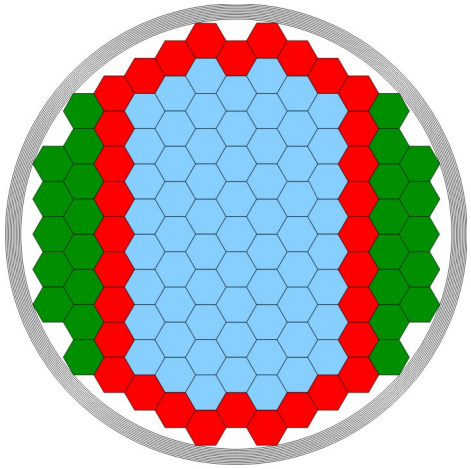

Figure 7: Design for the ohmic side of the detector. The 127 hexagons represent individual detector elements. Proton events in the interior hexagons generate a valid trigger, while the perimeter hexagons are used only for detecting electrons. The concentric circles represent the guard ring structure. Electrical contact is made to each hexagon to provide the bias voltage and collect the charge deposited by incident particles. The areas between the pixels and guard rings are electrically connected to form one additional channel.

noise gain of the preamplifier increases with detector capacitance, while the speed decreases. With our choice of $a = 0.7\,\mathrm{cm}^2$, the parallel plate capacitance of one element is approximately 6 pF. Inter-pixel capacitance and contributions from the electrical interconnects will bring the total capacitance to approximately 10 pF, which is acceptably small. Finally, the number of detector elements, 127 per detector, does not require an unacceptably large number of electronic channels.

It is important to note that, though the detector is segmented, there are no dead spaces between the detector elements. Even though there is a gap of 100 μm between the metal pads for adjacent elements, *all* charge deposited in the active volume of the detector is collected, though it may be shared among adjacent elements. This property guarantees that if a proton hits within the interior hexagons in Fig. 7, the corresponding electron must hit within the active area (interior plus perimeter hexagons) of the same (or opposing) detector. This allows the use of the proton hit as a trigger. Since the protons start with a very small energy, less than 750 eV, and are accelerated to a much higher energy, 30 keV, this trigger is much less sensitive to the kinematics of the decay than an electron trigger. This technique is only practical with large-area detectors so that there are no dead areas that can spoil the coincidence efficiency, as would exist in a tiled scheme where several smaller detectors cover the same area.

Micron Semiconductor has constructed a prototype detector that fufills all of the design criteria, with the exception of thickness. The prototype detector is 0.5 mm thick rather than the required 2 mm. The prototype is currently being tested and plans for acquiring prototypes with thicknesses of 1.0, 1.5, and 2.0 mm are in progress.

The detector will be mounted to a ceramic support, suitable for cooling to cryogenic

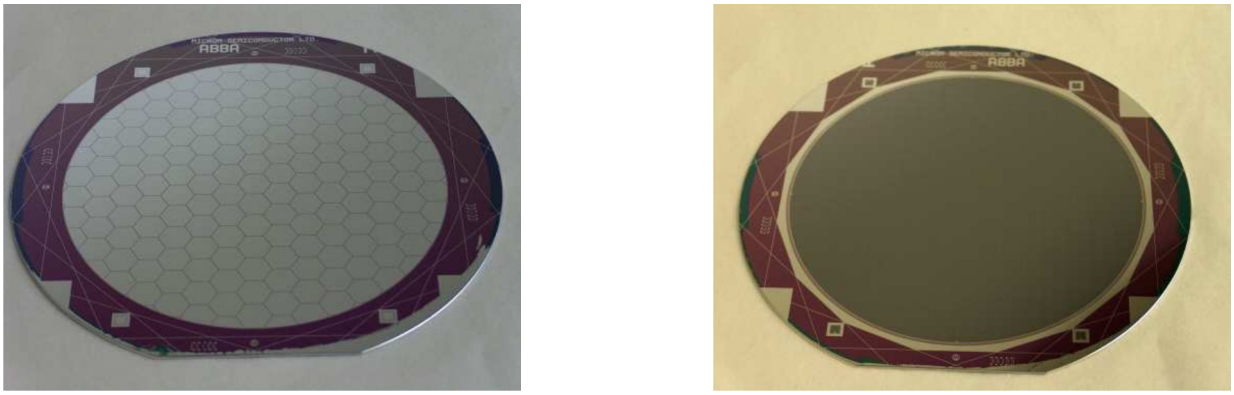

Figure 8: Photographs of the prototype abBA/Nab detector, before cutting from the 6 inch diameter silicon wafer and packaging. Charged particles enter through the junction side (left) and signals are read out from the ohmic side(right).

temperatures. Behind the ceramic support will be a circuit board with individual FETs, as well as feedback resistors and capacitors for each detector channel. Since the range of 752 keV electrons in silicon is approximately 1.7 mm, a 2 mm thickness is sufficient to stop the highest energy decay electrons.

The junction side of the detector will be formed by a thin p-implant. The total thickness of implant and metal will be equivalent to less than 100 nm of Si, resulting in < 10 keV of energy loss for 30 keV protons. The junction side will be featureless and will be held at ground potential. The ohmic side of the detector will be segmented to form the individual detector elements. The design for the segmentation consists of an array of 127 hexagonal elements, each approximately 1 cm^2 in area, as shown in Fig. 8. The active area of the detector extends to within 5 mm of the detector edge. In this boundary region there are approximately 20 guard rings that step down the applied bias voltage evenly, grading the electric field and reducing the probability of surface breakdown.

4. Dominant uncertainties in the measurement of a

In this section we will present a model which is made to study the sensitivity of our experiment to experimental imperfections. It is not refined enough to be able to describe the experimental data, but we can (and have started to) develop it further to be able to do so. The simplifications we make in the model are: (1) we neglect the time-of-flight of the electron, (2) we neglect the time the proton spends in the acceleration region, (3) we assume adiabaticity of the proton motion, (4) we assume a perfect vacuum, and (5) we assume a perfect detector apart from threshold effects. None of these approximations apply in our Monte Carlo simulations which are fully realistic apart from inclusion of backgrounds and certain details of mechanical construction of the apparatus that have yet to be specified.

The above effects become important when the model is applied to fit actual experimental data. The first assumption is based on the observation that typical values for electron and proton times of flight are $t_e = 5$ ns and $t_p = 5\,\mu$s, respectively, so that the electron time-of-flight t_e can be neglected in the first order. The second assumption is based on the fact that

the kinetic energy of the proton in the decay, expansion and time-of-flight regions is below 750 eV. In the acceleration region, this small kinetic energy is increased by the high voltage to between 30 kV and 30.75 kV. Therefore, the time the protons spend in the acceleration region is small compared to the time they need to get there. The assumptions of adiabaticity, a zero rest gas level, and the perfect detector are all good approximations. They are discussed later in this section.

We find that the analysis of our measured data depends heavily on the accuracy with which we can determine the spectrometer response function. There are two different strategies of data analysis, and we will implement both. In the first approach (Method A), we determine the shape of the spectrometer response function from theory, but leave a couple of free parameters in it which we adjust to fit the measured spectra. The second approach (Method B) relies on obtaining *a priori* as full a description of the neutron beam and electromagnetic field geometry, subsequently calculating the detection function with its uncertainties, and finally fitting the experimental data with only the physics observables as free parameters. While both methods are presented below, the first approach is also discussed in detail in Ref. [22]. We also note that Ref. [22] explicitly takes into account the time of flight of the proton in the acceleration region.

4.1. Principles of measurement and data analysis

The observables of our spectrometer are electron energy E_e and the difference of the time-of-flight of electron, t_e, and proton, t_p. The time-of-flight of the proton is given by

$$t_\mathrm{p} = \frac{f(\cos\theta_{\mathrm{p},0})}{p_\mathrm{p}}\,. \tag{11}$$

Here, $\theta_{\mathrm{p},0}$ is the initial angle of the proton relative to the magnetic field and $f(\cos\theta_{\mathrm{p},0})$ is a function given by the spectrometer which depends on the neutron beam profile and the geometry of electric and magnetic field. If we neglect the time the protons spend in the acceleration region, $f(\cos\theta_{\mathrm{p},0})$ doesn't depend on the proton momentum. If magnetic field and electric potentials are constant throughout the spectrometer, and the protons have a flight path of length l, the function f is given by

$$f(\cos\theta_{\mathrm{p},0}) = \frac{m_\mathrm{p} l}{\cos\theta_{\mathrm{p},0}}\,. \tag{12}$$

In the magnetic and electric field configuration which we have, f becomes more complicated, as θ_p, the angle between proton momentum and magnetic field depends on the position. Our electric and magnetic fields change slowly enough so that the trajectory of the proton or the electron can be calculated in the adiabatic approximation. Here, the orbital magnetic momentum is a constant of motion for the proton. From this and energy conservation we can derive that the momentum component parallel to the magnetic field at each point in the spectrometer is given by

$$p_{\mathrm{z,p}}(z) = p_\mathrm{p}\sqrt{1 - \frac{B(z)}{B_0}\sin^2\theta_{\mathrm{p},0} - \frac{e(U(z)-U_0)}{T_0}}\,. \tag{13}$$

where T_0 is the kinetic energy of the proton and B_0, and U_0 are the magnetic field and the electric potential in the decay point. The last term under the square root vanishes everywhere except in the acceleration region, so we can omit it for now. We use $\cos\theta_{\mathrm{p}}(z) = p_{\mathrm{z,p}}/p\mathrm{p}$ and arrive at

$$f(\cos\theta_{\mathrm{p},0}) = m_{\mathrm{p}} \int_{z_0}^{l} \frac{dz}{\cos\theta_{\mathrm{p}}(z)} = m_{\mathrm{p}} \int_{z_0}^{l} \frac{dz}{\sqrt{1-\frac{B(z)}{B_0}\sin^2\theta_{\mathrm{p},0}}}\,. \tag{14}$$

This function has to be modified for protons which are reflected on the magnetic field (that is the magnetic mirror effect). A proton whose initial momentum is pointing towards a magnetic field maximum will be reflected if its initial angle relative to the magnetic field $\theta_{\mathrm{p},0}$ fulfills the condition

$$\cos^2\theta_{\mathrm{p},0} < \cos^2\theta_{\mathrm{crit}} = 1 - B_0/B_{\max}\,. \tag{15}$$

$B_{\max}$ is the maximum magnetic field on the magnetic field line passing through the decay point. At the point of reflection z_{refl}, we have $\theta_{\mathrm{p}}(z_{\mathrm{refl}}) = 0$ (the square root in Eq. (13) vanishes). The Lorentz force, which was responsible for the deceleration of the proton momentum component along the z axis before the reflection will also accelerate the proton after the reflection. For reflected protons, $f(\cos\theta_{\mathrm{p},0})$ gets an extra term and we have

$$f(\cos\theta_{\mathrm{p},0}) = 2m_{\mathrm{p}} \int_{z_{\mathrm{refl}}}^{z_0} \frac{dz}{\sqrt{1-\frac{B(z)}{B_0}\sin^2\theta_{\mathrm{p},0}}} + m_{\mathrm{p}} \int_{z_0}^{l} \frac{dz}{\sqrt{1-\frac{B(z)}{B_0}\sin^2\theta_{\mathrm{p},0}}}\,.$$

We will sort the proton time-of-flights into a $1/t_{\mathrm{p}}^2$ spectrum. The observable $1/t_{\mathrm{p}}^2$ depends on p_{p}^2 through

$$p_{\mathrm{p}}^2 = f^2(\cos\theta_{\mathrm{p},0}) \cdot \frac{1}{t_{\mathrm{p}}^2}\,. \tag{16}$$

If $f(\cos\theta_{\mathrm{p},0})$ were a constant, the distribution of $1/t_{\mathrm{p}}^2$, $P_t(1/t_{\mathrm{p}}^2)$, would look like the distribution of p_{p}^2, $P_p(p_{\mathrm{p}}^2)$. Equation (10) relates p_{p}^2 and $\cos\theta_{\mathrm{e}\nu}$. Therefore, we arrive at

$$\begin{aligned} P_p(p_{\mathrm{p}}^2) = & \begin{cases} 1 + a\beta\cos\theta_{\mathrm{e}\nu} & \text{where } |\cos\theta_{\mathrm{e}\nu}| < 1 \\ 0 & \text{otherwise} \end{cases}, \\ = & \begin{cases} 1 + a\beta\frac{p_{\mathrm{p}}^2 - p_{\mathrm{e}}^2 - p_\nu^2}{2p_{\mathrm{e}}p_\nu} & \text{where } \left|\frac{p_{\mathrm{p}}^2 - p_{\mathrm{e}}^2 - p_\nu^2}{2p_{\mathrm{e}}p_\nu}\right| < 1 \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \tag{17}$$

This ideal situation would imply an infinitely wide, sudden, but adiabatic field expansion, which is a contradiction. Our field, shown in Fig. 6, is a compromise and we have to take its shape into account. We use Eqs. (14) and (16), but we still neglect the region where the proton is accelerated (the time it spends there is very small due to the acceleration). Mathematically that means, that we end the integrals at $l = 1.5\,\mathrm{m}$. Then we can treat $1/t_{\mathrm{p}}^2$ as a product of independent random variables, $1/t_{\mathrm{p}}^2 = p_{\mathrm{p}}^2 \cdot r$, $r = 1/f^2(\cos\theta_{\mathrm{p},0})$, and we can write for the $P_t(1/t_{\mathrm{p}}^2)$ distribution:

$$P_t\left(\frac{1}{t_{\mathrm{p}}^2}\right) = \int P_p\left(p_{\mathrm{p}}^2\right) \underbrace{P_r\left(\frac{1}{t_{\mathrm{p}}^2 p_{\mathrm{p}}^2}\right)\frac{1}{p_{\mathrm{p}}^2}}_{\Phi\left(\frac{1}{t_{\mathrm{p}}^2}, p_{\mathrm{p}}^2\right)} dp_{\mathrm{p}}^2\,. \tag{18}$$

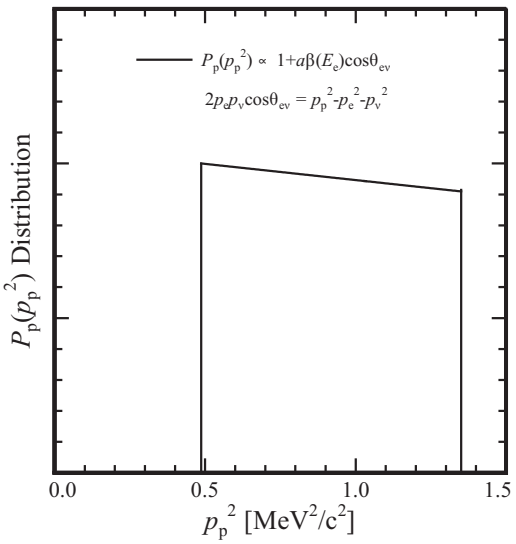

Figure 9: Distribution of $P_p(p_p^2)$ for $E_e = 550\,\text{keV}$. The slope is proportional to the neutrino–electron correlation coefficient a according to Eq. (17).

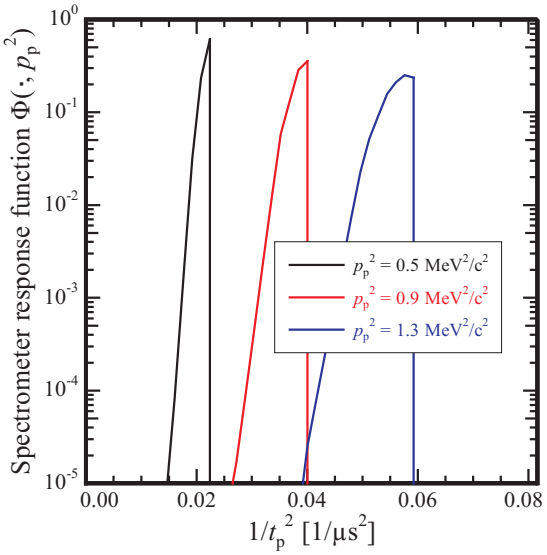

Figure 10: Nab spectrometer response function $\Phi\left(\frac{1}{t_p^2}, p_p^2\right)$, shown for different proton momenta, the magnetic field from Fig. 6 and a centered neutron beam with a width of 2 cm.

Within the framework of analysis Method B $P_r(r)$ is calculated numerically, and it is averaged over the neutron beam in the decay volume. $\Phi\left(1/t_p^2, p_p^2\right)$ is our spectrometer response function; for several given proton momenta p_p it is shown in Fig. 10.

Fig. 11 shows our expected $P_t(1/t_p^2)$ spectra. The left panel presents calculations from Eq. 18 (Method B). The right panel presents spectra generated by a full Monte Carlo simulation of the spectrometer, which doesn't rely on adiabaticity, nor neglect the electron time-of-flight. There is clear qualitative agreement.

In a more sophisticated model we could include the trigger efficiency of the detector in the spectrometer response function $\Phi\left(1/t_p^2, p_p^2\right)$. We will not do that here, instead we discuss the detector efficiency in Sect. 4.4 as a separate issue.

In the framework of the analysis Method A we will determine the spectrometer response function in a fit to the data. The high TOF and low TOF sides of the proton TOF spectra for each electron energy are primarily given by the spectrometer response. On the other hand, the slope of the central portion of the $1/t_p^2$ spectrum is determined by a, the parameter we wish to measure (see, e.g., Figs. 3, 9 and 11). Therefore we do not expect a strong correlation between the spectrometer response function and a. The relationship between t_p and pp is given by

$$t_p = \int_0^{s(l)} \frac{m\,ds}{\sqrt{p^2\left(1-(1-u^2)(B(s)/B(0))\right) + 2m_p e(U(s)-U(0))}}\,, \tag{19}$$

where s is the arclength along a magnetic field line, m_p is the proton mass, U is the electrical potential, p is the initial proton momentum, B is the magnetic field strength, and $u = \cos\theta_{pB}$ is the cosine of the angle between the proton momentum and the magnetic field line direction.

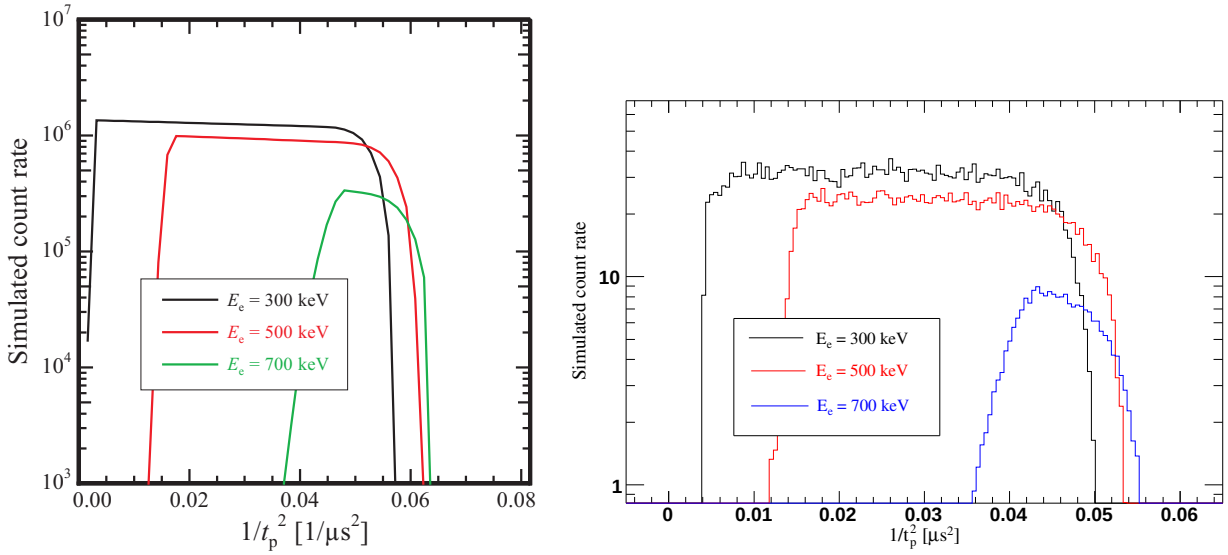

Figure 11: $P_t(1/t_p^2)$ spectra generated from one million Monte Carlo events for electron energies 300 keV (black), 500 keV (red) and 700 keV (green). In the left panel, these spectra are calculated with the theory (Method B) presented in the text. The right panel presents results of a full realistic Monte-Carlo simulation which involves minimal simplifications.

In the following discussion we use the model functions from Appendix A, which include the TOF of the proton in the acceleration region. We start with the yields at a range of electron energies like those shown in Fig. 3. We smooth these yields to calculate the yields for the model spectrometer. A p_p^2 spectrum for electron energy 0.469 MeV is shown in Figs. 21 and 22. We perform fits to these spectra using the trial function

$$Y_2(p^2) = Y\left(\alpha(p^2 - z_0)\right) + \frac{1}{2\epsilon}\left[Y((1+\epsilon)\alpha(p^2 - z_0)) - Y((1-\epsilon)\alpha(p^2 - z_0))\right] . \qquad (20)$$

This trial function includes a p^2 offset z_0, a z calibration error α, and width uncertainty ϵ. We have calculated M, the measurement matrix, at each energy. We form the yield-weighted measurement matrix by summing over energies. We then calculate the uncertainty in a by fitting all four parameters. The main result is the uncertainty in a per root event number, $\sqrt{N}\sigma_a$. Assuming perfect knowledge of the spectrometer we have $\sqrt{N}\sigma_a = 2.3$; the fitting procedure worsens this quantity to $\sqrt{N}\sigma_a = 2.6$. The error correlation matrix is

$$M = \begin{pmatrix} 1 & 0.136 & 0.247 & 0.403 \\ & 1 & 0.162 & 0.474 \\ & & 1 & 0.796 \\ & & & 1 \end{pmatrix} \begin{pmatrix} a \\ z_0 \\ \alpha \\ \epsilon \end{pmatrix} . \qquad (21)$$

The parameter most strongly correlated with a is the width of the response function.

In conclusion, for an attainable spectrometer configuration, the yield dependence on p_p^2 and E_e can be used to check the spectrometer response determined from field measurements. Carrying out these checks during the commissioning phase would enable us to validate the

measured fields. We can then use the fitting procedure to constrain the spectrometer response. It appears that the Method A fitting procedure increases the uncertainty in a only modestly. This would be a reasonable tradeoff in obtaining an independent determination of a from an *in situ* check of the measured fields.

4.2. Auxiliary asymmetry measurements

It is worth while to calculate the asymmetry of the count rates of electrons or protons in both detectors for an extended neutron beam, as we will need it for calibration purposes. The distribution of decay protons is isotropic. The proton count rate seen by each detector is given by the magnetic mirror effect. The situation is shown in Fig. 12. We define count rates in the form N_{ULU}, the count rate of protons which are produced above the magnetic field maximum (the first subscript "U"), which are emitted originally into the lower hemisphere ("L"), but which appear in the upper detector (the last "U") thanks to a magnetic mirror reflection. Then, the total count rate in the upper detector N_{U} is given by

$$\begin{aligned} N_{\mathrm{U}} = & \quad N_{\mathrm{UUU}} + N_{\mathrm{LUU}} + N_{\mathrm{ULU}} \\ = & \iiint\limits_{U+L} d^3x\rho(\vec{x}) \int\limits_0^{\pi/2} \tfrac{1}{2} N \sin\theta_{\mathrm{p},0} d\theta_{\mathrm{p},0} \\ & + \iiint\limits_{U} d^3x\rho(\vec{x}) \int\limits_{\pi/2}^{\pi-\theta_{\mathrm{crit},0}} \tfrac{1}{2} N \sin\theta_{\mathrm{p},0} d\theta_{\mathrm{p},0} \,. \end{aligned} \tag{22}$$

(23)

The count rate $N_{\mathrm{LLU}} = 0$, as in this case a magnetic mirror reflection is not possible. The critical angle $\theta_{\mathrm{crit},0}$ is the angle where the proton is reflected in the last possible moment on the magnetic field maximum. It is given by $\theta_{\mathrm{crit},0} = \arcsin\sqrt{B_0/B_{\mathrm{max}}}$, as shown in Eq. (15). The quantity $\rho(\vec{x})$ is the density profile of the neutron decays.

$$\begin{aligned} N_{\mathrm{L}} = & \quad (N_{\mathrm{LLL}} + N_{\mathrm{ULL}}) + N_{\mathrm{LUL}} + N_{\mathrm{UUL}} \\ = & \iiint\limits_{U+L} d^3x\rho(\vec{x}) \int\limits_{\pi/2}^{\pi} \tfrac{1}{2} N \sin\theta_{\mathrm{p},0} d\theta_{\mathrm{p},0} \\ & + \iiint\limits_{L} d^3x\rho(\vec{x}) \int\limits_{\theta_{\mathrm{crit},0}}^{\pi/2} \tfrac{1}{2} N \sin\theta_{\mathrm{p},0} d\theta_{\mathrm{p},0} \,. \end{aligned} \tag{24}$$

We define

$$k_\Delta = \frac{\iiint\limits_{U} \rho(\vec{x}) \cos\theta_{\mathrm{crit},0} d^3x - \iiint\limits_{D} \rho(\vec{x}) \cos\theta_{\mathrm{crit},0} d^3x}{\iiint\limits_{U+D} \rho(\vec{x}) d^3x} \,, \tag{25}$$

$$k_\Sigma = \frac{\iiint\limits_{U} \rho(\vec{x}) \cos\theta_{\mathrm{crit},0} d^3x + \iiint\limits_{D} \rho(\vec{x}) \cos\theta_{\mathrm{crit},0} d^3x}{\iiint\limits_{U+D} \rho(\vec{x}) d^3x} \,. \tag{26}$$

(27)

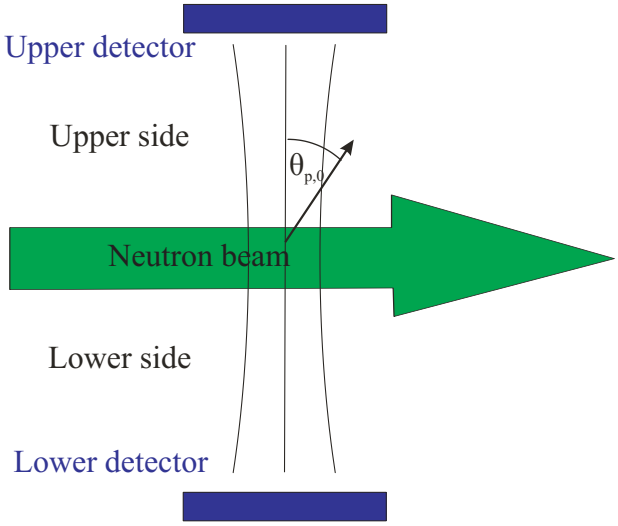

Figure 12: Sketch of the problem to calculate Up-Down-Asymmetries

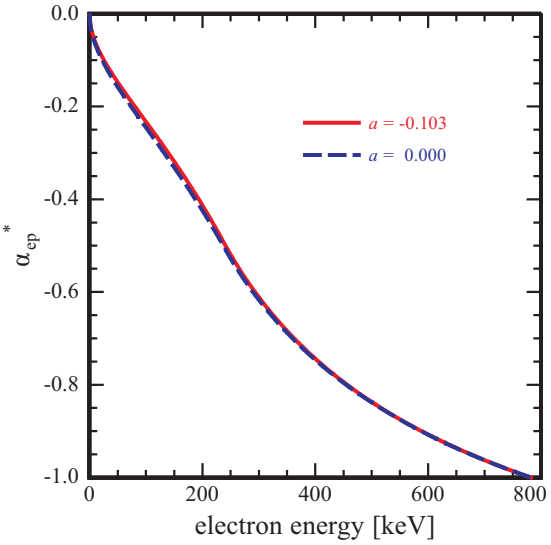

Figure 13: Asymmetry of the count rates when both electron and proton go into the same versus into the opposite detector. The asymmetry is nearly independent of the exact value of a

Then we have

$$\alpha_{\mathrm{p}}^{*} = \frac{N_{\mathrm{U}} - N_{\mathrm{D}}}{N_{\mathrm{U}} + N_{\mathrm{D}}} = \frac{k_{\Delta}}{1 + k_{\Sigma}} \,. \tag{28}$$

The magnetic field in the decay volume can be parametrized as

$$B(z) = B_0 \left(1 - (\alpha z)^2\right) \,. \tag{29}$$

In our magnetic field, $\alpha \sim 20\,\mathrm{m}^{-1}$. Then, for a uniform neutron beam with a width of $\Delta z_0 = 10\,\mathrm{mm}$ and a center of $\bar{z}_0$ with $|\bar{z}_0| \ll \Delta z_0$, we get

$$\frac{k_{\Delta}}{1 + k_{\Sigma}} = \frac{\alpha \bar{z}_0}{1 + \frac{\alpha \Delta z_0}{8}\left(1 + \left(\frac{2\bar{z}_0}{\Delta z_0}\right)^2\right)} \approx \alpha \bar{z}_0 \,. \tag{30}$$

The statistical uncertainty in the determination of α_{p}^{*} for N detected events is

$$\Delta\alpha_{\mathrm{p}}^{*} = \frac{1}{N} \,. \tag{31}$$

For the measurement of the proton asymmetry we don't identify the proton as being a second event close in detector position and time to a first event, presumably the electron. Instead, we just compare the singles event count rates with and without an electrostatic barrier of a height of $\Delta U = +1\,\mathrm{kV}$. This barrier electrode can be the outer HV electrode in Fig. 4. In this way we avoid the corrections due to uncertainties of the trigger efficiency for electrons. Due to the acceleration of the protons to 30 kV the detection efficiency of the proton detector is essentially unity. A measurement of this asymmetry allows us to center the neutron beam;

for a centered beam $\alpha_{\rm p}^*$ vanishes. In addition, by deliberately moving a diaphragm with a small slit through the beam, α can be measured with a precision of $\Delta\alpha/\alpha \sim 5 \times 10^{-3}$.

The same kind of asymmetry can be defined for electrons. The definition would then be:

$$\alpha_{\rm e}^*(E_{\rm e}) = \frac{N_{\rm U}(E_{\rm e}) - N_{\rm D}(E_{\rm e})}{N_{\rm U}(E_{\rm e}) + N_{\rm D}(E_{\rm e})} = \frac{k_\Delta}{1 + k_\Sigma}\,. \tag{32}$$

Now the count rates $N_{\rm U}(E_{\rm e})$ and $N_{\rm D}(E_{\rm e})$ are single electron count rates. Again, for our field the contribution of k_Σ can be neglected. This asymmetry serves to study the quality of the understanding of the electron trigger efficiency and background subtraction.

In addition to these single particle measurements, we can measure electron and proton and distinguish the two cases where electron and proton go into the same detector (count rate $N^{\uparrow\uparrow}(E_{\rm e})$) and into opposite detectors (count rate $N^{\downarrow\downarrow}(E_{\rm e})$). We define the asymmetry of the count rates of these two cases by

$$\alpha_{\rm ep}^*(E_{\rm e}) = \frac{N^{\uparrow\uparrow}(E_{\rm e}) - N^{\uparrow\downarrow}(E_{\rm e})}{N^{\uparrow\uparrow}(E_{\rm e}) + N^{\uparrow\downarrow}(E_{\rm e})}\,. \tag{33}$$

We can use the calculation of [30] as a starting point to compute the expected count rates to be:

$$N^{\uparrow\uparrow}(E_{\rm e}) = \tfrac{2}{4}F(E_{\rm e}) \begin{cases} f_b\left(1 - \frac{1}{2}r\right) + \frac{1}{4}a\beta\left(\frac{1}{2}r^2 - 1\right) & ;r < 1 \\ \frac{1}{2r}\left[f_b - \frac{1}{4r}a\beta\right] & ;\text{otherwise} \end{cases} \tag{34}$$

$$N^{\uparrow\downarrow}(E_{\rm e}) = f_b F(E_{\rm e}) - N^{\uparrow\uparrow}\,. \tag{35}$$

Here, $F(E_{\rm e})$ is the unpolarized electron spectrum, $f_b = (1 + bm_{\rm e}/E_{\rm e})$, and $r = p_{\rm e}/p_\nu$. The asymmetry is then written as

$$\alpha_{\rm ep}^*(E_{\rm e}) = \begin{cases} \frac{1}{2}r + \frac{a}{4f_b}\beta\left(\frac{1}{2}r^2 - 1\right) & ;r < 1 \\ \frac{1}{2r} - 1 - \frac{1}{8r^2}\frac{a}{f_b}\beta & ;\text{otherwise} \end{cases}\,. \tag{36}$$

The statistical accuracy of the measurement of the average asymmetry is $\Delta\overline{\alpha_{\rm ep}^*}/\Delta\overline{\alpha_{\rm ep}^*} = 0.68/\sqrt{N}$. This asymmetry is shown in Fig. 13. The figure also shows that $\alpha_{\rm ep}^*(E_{\rm e})$ depends only little on the value of the neutrino electron correlation coefficient a or on the Fierz interference term b. With a rough knowledge of a and b we can take the asymmetry as given and use our measurement of it as a tool to calibrate the spectrometer.

4.3. Statistical uncertainty

For the estimation of the statistical error we use the basic model (Method B) presented above. Here, the neutrino–electron correlation coefficient a is determined in a χ^2-fit to the two dimensional function (18) a normalization constant N. The fit function depends on proton time-of-flight $t_{\rm p}$ and electric energy $E_{\rm e}$. Fitting parameters are N, a, and the Fierz interference term b. Omission of b, i.e., a fit within the Standard Model, would not improve the uncertainty in a. The uncertainty in this fit is shown in the table below. We took into account several possible values for $E_{\rm e,min}$, a low energy cutoff for the electron energy due to the detection efficiency, and $t_{\rm p,max}$, a high proton time-of-flight cutoff due to accidental coincidences.

Quantity	$E_{\mathrm{e,min}} = 0$	$E_{\mathrm{e,min}} = 100\,\mathrm{keV}$	$E_{\mathrm{e,min}} = 100\,\mathrm{keV}$, $t_{\mathrm{p,max}} = 10\,\mu\mathrm{s}$	$E_{\mathrm{e,min}} = 300\,\mathrm{keV}$, $t_{\mathrm{p,max}} = 10\,\mu\mathrm{s}$
σ_a	$2.4/\sqrt{N}$	$2.5/\sqrt{N}$	$2.6/\sqrt{N}$	$3.5/\sqrt{N}$
σ_a (E_{cal}, l variable)	$2.5/\sqrt{N}$	$2.6/\sqrt{N}$		

N is the number of neutron decays. N is not restricted to the subset where electron and proton pass the respective cutoff conditions. The second line describes a fit where two parameters of the spectrometer are determined from the fit: the length of the flight path l and the slope of the energy calibration E_{cal}. The statistical uncertainty in a is affected only marginally.

4.4. Uncertainties due to the spectrometer response

Most of the time the proton needs to get from the neutron decay point to the detector it spends in the region with low magnetic field (see Fig. 6). The length of that section should be measured with an relative uncertainty of 2×10^{-5} to get a precision in a of 10^{-3}. We can take this number from a fit to the measured $P_t(1/t_{\mathrm{p}}^2)$ distribution, specially the low-t_{p} side is sensitive to this length, but we have to be careful to model the other parameters of the detector response function sufficiently well to make sure that in the fit an incorrect shape of the detector response function is not hidden by an incorrect choice of the length of the TOF region.

1. Neutron beam profile

   The position of the neutron decay gives the starting point of the proton (and electron) flight path. Neutrons which decay at the upper side of the neutron beam produce protons with shorter travel time to the upper detector and longer travel times to the lower detector. Each detector individually sees a distorted time-of-flight spectrum. A shift of the neutron beam center of $200\,\mu$m corresponds to a shift $\Delta a/a \sim 0.4\%$. This is much worse if the flight path length is fixed in the fit. It still largely cancels when a is averaged over both detectors.

   The center of the neutron beam can be precisely determined from the measurement of the asymmetry of the proton count rates α_{p}^*. The position sensitive detector allows for a possible correction due to the misalignment of the detectors. A shift of the center of the beam of $100\,\mu$m towards the upper detector would cause $\alpha_{\mathrm{p}}^* = -0.2\%$. We discussed above that we expect the accuracy of our measurement of the proton asymmetry is sufficient to extract the position of the center of the neutron beam with that accuracy.

   The width of the neutron beam in the decay volume will be about $\Delta z_0 = 20\,\mathrm{mm}$. An error of $50\,\mu$m would cause $\Delta a/a \sim 0.1\%$. But note that the measurement of the shape of the neutron beam profile is a relative measurement and can be done with much higher accuracy. Detectors with a spatial resolution of several μm were developed for ultracold neutron experiment, and can be used for cold neutrons, too [23, 24].

2. Magnetic field map

   If we use incorrect magnetic or electric field values, we will calculate the nominal TOF incorrectly, distorting the signal for a. It is very important that we know the field

expansion ratio $r_{\mathrm{B}} = B_{\mathrm{TOF}}/B_0$ well, where B_0 is the maximum of the magnetic field in the decay volume on a given magnetic field line and B_{TOF} is the magnetic field in the time-of-flight region. At the chosen magnetic field ratio $r_{\mathrm{B}} = 0.0254$, the necessary accuracy is $\Delta r_{\mathrm{B}}/r_{\mathrm{B}} = 10^{-3}$ for an uncertainty of $\Delta a/a \sim 10^{-3}$. Such an accuracy can be reached with a calibrated and temperature controlled Hall probe, as it was demonstrated in the aSPECT experiment [20].

The magnetic field in the decay volume can be parametrized as $B(z) = B_0\,(1 - (\alpha z)^2)$. Here, α has to be determined with an accuracy of 1×10^{-3}. Even if this is the accuracy of a relative measurement (neither the z offset nor the absolute magnetic field has to be known), this seems to be too demanding for a direct measurement. Fortunately the measurement of the proton asymmetry α_{p}^* while the neutron beam is shifted (a diaphragm with a slit could be used to move the beam) can be converted into a measurement of α.

In the transition region between decay volume and the drift (time-of-flight) region, a magnetic field bump (or a non-linearity of the magnetic field sensor) should not exceed a relative size of 2×10^{-3}.

3. Length of the flight path

   The effective length of the flight path ranges from the point of the decaying neutron to the onset of the electric field used for the acceleration of the proton. A shift in this path length of about $30\,\mu$m would cause the neutrino electron correlation coefficient to be off by $\Delta a/a = 0.1\%$. It is not possible to measure such an ill-defined length directly with this precision. However, the length of the flight path can be an additional fitting parameter, as discussed above. It is possible to perform a consistency check with high precision: We segment the electrodes around the decay volume and do measurements with two different flight paths, whose difference in length is precisely known. Such a technique was demonstrated recently in the NIST lifetime experiment [25].

4. Homogeneity of electric and magnetic field

   The homogeneity of the magnetic field in the time-of-flight region has been discussed before, we want to have at least 10^{-3} to give r_{B} a unique value.

   Variations of the electric potential lead to a change in the kinetic energy of the proton.

$$E_{\mathrm{p}} \;\rightarrow\; E_{\mathrm{p}} + e\Delta U \tag{37}$$

$$p_{\mathrm{p}} \;\rightarrow\; p_{\mathrm{p}} + \left(1 + \frac{e\Delta U}{2E_{\mathrm{p}}}\right) . \tag{38}$$

   If we take from the last section that the relative accuracy we need to determine the flight path length $\Delta l/l \sim 2 \times 10^{-5}$, then we need the same accuracy for the proton momentum p_{p}, which means that for an average proton energy of $E_{\mathrm{p}} \sim 400\,\mathrm{eV}$ we allow $\Delta U \sim 16\,\mathrm{meV}$. This is a condition which is much less severe than the trapping effect which is discussed next.

   If electrons or protons are produced in a minimum of the magnetic field or the electric potential, then they can be trapped. A magnetic trap confines all particles with

$|\cos\theta_{\mathrm{p}}|, |\cos\theta_{\mathrm{e}}| < \sqrt{\Delta B/B}$, where ΔB is the depth of the trap. An electrostatic trap traps protons if their energy in the longitudinal motion is smaller than the depth of the trap; the effect on the electrons is negligible due to their much higher energy. The trapping causes a bias in our sample, and it disturbs the count rate asymmetries. We can neglect the trapping if less than 10^{-4} of our events have trapped particles. This translates into the condition that $\Delta B/B < 10^{-6}$ and $\Delta U < 5\,\mu$V. The magnetic field has a strong maximum, so that there will be no minima. The condition on the homogeneity of the electric field in the decay volume is not easy to fulfill: For one, the effect of the entrance holes for the neutron beam on the electric field distribution has to be studied. We have a design in which the homogeneity is better than the specification. It is well known that the orientation-dependent work function of individual metallic grains ("patch" effect) can give rise to local electric potential variations of order several 100 mV very close to a metallic surface (In this context, "very close" means on the order of the dimensions of an individual grain). The hope is that the patch effect averages to zero when the distance to the surface becomes large, but experience from the aSPECT experiment shows that this is not necessarily the case for technical surfaces. They found which a variation of the work function which was 100 mV over a distance of 5 cm along a gold coated copper electrode. The reasons are not yet understood, impurities in or on the surface coating might be the culprit. In addition, in Ref. [26] surface charging on metallic conductors due to radiation is found and discussed. The effect can be as big as several volts, but at radiation levels which are many orders of magnitude higher than in Nab.

Our strategy is as follows: We will test our surfaces with a Kelvin probe, with which the level of local variations of surface charges and the work function can be measured with an accuracy of several meV. We can minimize the effect by considering different surface materials and treatments. We will coat the inside of the electrode, at least in the vicinity of the decay region, with evaporated gold, colloidal gold, colloidal carbon or similar material which has been shown to significantly reduce the work function inhomogeneities. Furthermore, we can test at the neutron beam if the radiation level there makes a difference. Finally, since we will not be able to measure inhomogeneities directly if their amplitude is below a meV, we will use the fact that protons which can be reflected by such small electric potentials arrive at the very end of the time of flight spectrum of the protons. Hence only the end of the time-of-flight spectrum will be distorted, and we can disregard it if necessary.

5. Rest gas

A poor spectrometer vacuum would have several consequences. Besides the technical problem of HV breakdowns we have to worry about background fluctuations and the influence of the rest gas on particle trajectories. A Monte Carlo simulation was performed with GEANT4 to determine the effect of the rest gas. The vacuum was defined as the molecular hydrogen gas at 10^{-8} torr. Fig. 14 shows the histogram of proton time-of-flight differences for one million neutron decays in the vacuum with hydrogen rest gas, and in an ideal vacuum practically devoid of rest gas ("intergalactic vacuum"). There are two effects to be noted. One, interactions with the rest gas broaden the TOF

distribution with an rms of about 3.2 ns. More importantly, the mean time of flight is increased by ∼72 ps. The effect on the electrons is less pronounced. We are proceeding with more detailed studies of the effect. While a 72 ps shift can be corrected for in the analysis, it is preferable to work on achieving a vacuum of 10^{-9} torr or better.

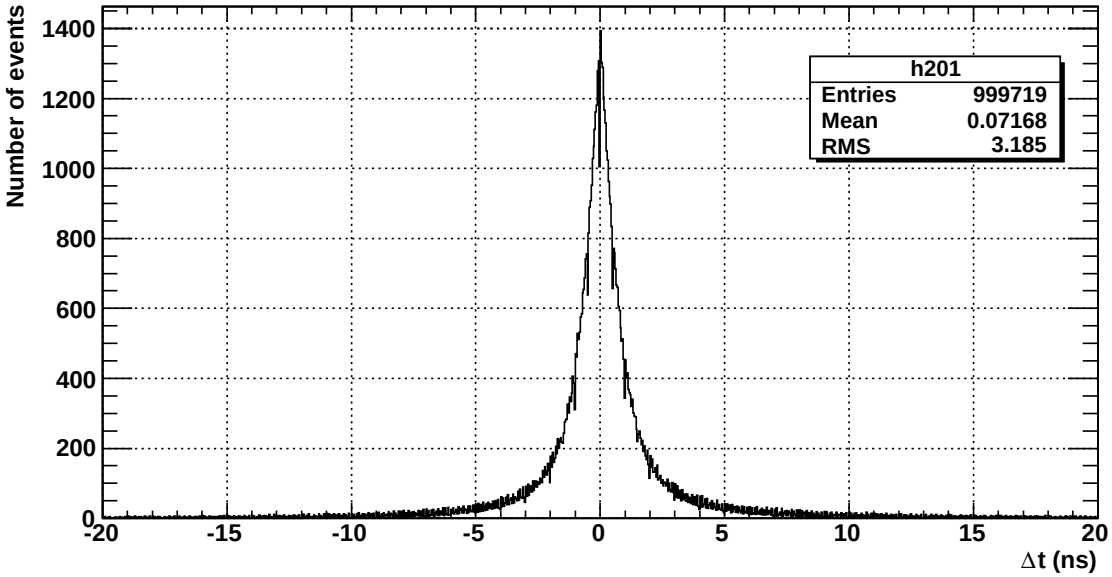

Figure 14: Effect of a rest gas at the level of 10^{-8} torr H_2 on the proton time-of-flight t_p. Plotted are the proton TOF differences, Δt, between decays in a vacuum with 10^{-8} torr of H_2, and an ideal vacuum without rest gas, for one million neutron decays in each. The mean TOF shift due to the rest gas is approximately 72 ps.

6. Doppler effect

   Our initial estimates indicate that the Doppler effect is most likely negligible since the neutron beam is transverse to the spectrometer axis. Hence, it should be possible to take it into account with sufficient precision. The Doppler effect is on the simulation agenda and will be analyzed in due course.

7. Adiabaticity

   It is not necessary that the electron and proton orbits in our spectrometer be calculable in the adiabatic approximation, but it simplifies the construction of an effective model. The analytical analysis of our uncertainties is based on the assumption of adiabaticity, and we use the Monte Carlo simulations to check that assumption. In Fig. 15 we show the distribution of the proton momentum angles with respect to the magnetic field in the drift (time-of-flight) region. There are small deviations from adiabaticity in our present beam profile, and we have to reshape the field slightly.

4.5. Uncertainties due to the detector

Another set of systematics is due to the imperfections in the proton or electron detection, as discussed below.

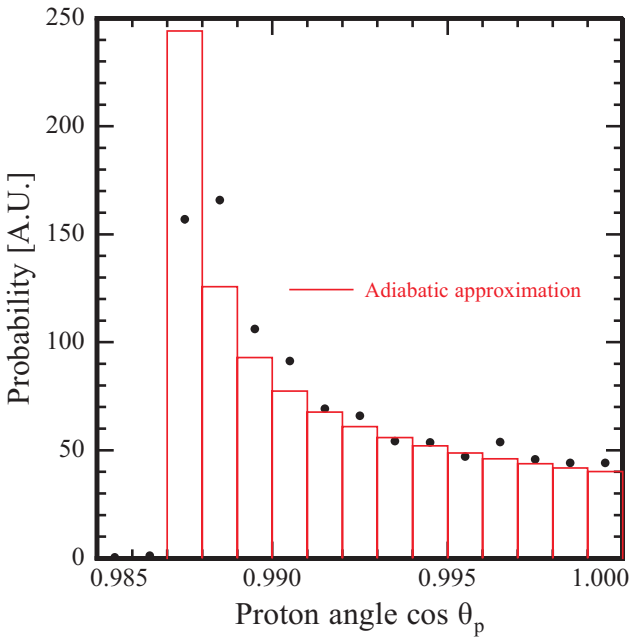

Figure 15: Comparison between the distribution of proton angles to the magnetic field in the TOF region between the Monte Carlo Simulation and the Adiabatic Model.

1. Detector alignment

   Will will align the detector with the neutron decay events. We will measure the average displacement between electron and proton if both particles go in opposite detectors, if this is non-zero, then the detectors are misaligned. We will correct the position of the detectors for this shift.

2. Electron energy calibration

   We have to understand the response of our detector to be able to extract the electron energy from the measured data. The width of the detector response function is very small in our detector ($\Delta E_e \sim 3\,\mathrm{keV}$ is our specification). The low energy tail due to backscattered electrons is usually a problem in low energy electron spectroscopy. It is strongly suppressed in our setup, as backscattered electron are either reflected back to the detector being hit first due to a reflection from the strong magnetic field in the decay volume, or they hit the second detector. The sum of the measured energy in both detectors is close to the energy of events which are not backscattered, since the energy loss in the dead layer for electrons is only a few eV. The efficiency of the detector is unity above a threshold of 70 keV (we discuss edge effects elsewhere). Therefore, we expect the uncertainty due to the shape of the detection function function to be small.

   We expect the detector to have a linear energy–channel relationship. Its slope can be determined with a set of radioactive calibration sources, which are backed by very thin ($10\,\mu\mathrm{g/cm}^2$) carbon foils and which can be moved in the decay volume in the horizontal plane to reach every point in the detector. Six possible candidates for such calibration sources are identified in [28]. A more accurate method would be to determine the slope of the energy channel relation from the measured beta spectrum. In this way, the energy calibration can be fixed with an accuracy of $0.44/\sqrt{N}$ with the number of decays N and the Fierz parameter b as additional free parameters. We have to know

the calibration to 10^{-4}.

A further check of the capability to reconstruct the electron energy correctly, and at the same time of the trigger efficiencies for electrons of both detectors, is the measurement of the directional count rate asymmetry α_e^*.

3. Trigger hermiticity

   If we lose events, the data set will be biased. We will determine a value of a from a subset of the total number of neutron beta decay events, since a does not depend uniformly on the decay phase space. We will trigger on all events in order to avoid a bias in the data sample. However, we will not be able to trigger on all electrons with energies below an energy threshold of 70 keV. The threshold poses no problem: Our fit function depends explicitely on the electron energy and the omission of low electron energy data points doesn't introduce a bias, and it doesn't hurt the statistical sensitivity substantially. There is only a small probability of backscattered electrons which produce no signal in the detector being hit first. Since this probability depends on impact angle and energy, it introduces a bias, but the size of this effect can be estimated from the measured data [27], is expected to be small.

   Protons have a kinetic energy between 30 keV and 30.8 keV when they reach the detector. The average loss in the dead layer is ~ 5 keV. The back-scattering probability for protons from Si is very small. Furthermore, backscattered protons that lose more than 800 eV in the dead layer are forced back into the same detector by the electric field. Every proton deposits ~ 25 keV in one or the other detector. We can therefore trigger on every event that deposits more than a lower threshold of, say, 10 keV, and less than an upper threshold of, say, 50 keV.

   Another problem is that in order to avoid too many accidental coincidences we have to introduce a cutoff time $t_{p,max}$, above which the identification of a proton is not reliable any more. If we assume a typical signal to background rate of 1:1, the count rate of accidental coincidences is about $40\,s^{-1}$. The signal count rate is estimated to be 200 Hz per detector, but only counts in pixels adjacent to the one hit by the electron are used for the coincidences. That means that for $t_{p,max} = 10\,\mu$s the accidental fraction is 1×10^{-4}. Large-TOF events play a limited role in the determination of a because of their relatively low statistical significance. This circumstance greatly reduces the impact of the correction due to missed large-TOF events. On the other hand, large-TOF events are valuable for checking details of the spectrometer systematics. We will therefore make the cutoff time as large as practical, and will take the accidental coincidences into account as a known correction.

4. TOF uncertainties

   We will measure the difference between the proton and electron TOF's. The accuracy of the time difference has to be $\Delta(t_p - t_e) \sim 100$ ps. It is not necessary to reach this timing accuracy for each event, but it is necessary to reach it for the average. A concern is here the walk in the detector or its trigger, that is the dependence of the trigger time on signal height and on the signal shape (if the shape is different for electron and proton). This effect can be studied off-line with calibration sources.

5. Edge effect

   The size of the decay volume in the horizontal plane is usually given by a thin diaphragm. A $10\,\mu$m aluminum foil would be sufficient to stop the protons. With our position sensitive detector we don't even need to introduce a diaphragm. We will accept only electrons which do not produce a signal in one of the outermost pixels, and will accept protons also from the outer pixels. In this way the detector acts as an effective diaphragm. The decay volume is given by the projection of the inner detector pixels along the magnetic field lines onto the neutron beam. However, electrons are not moving along the field lines, but they spiral around them, and their gyration radius $r_{\rm gyr}$ depends on the particle momentum and the magnetic field:

$$r_{\rm gyr} = \frac{p_{\rm e} \sin\theta_{\rm e}}{eB} < \frac{4\,{\rm mm}}{B\,[{\rm T}]}\,. \tag{39}$$

   The maximum value for the gyration is given here for electrons; we can only accept electrons whose associated proton will reach the detector independent of the gyration. Electrons which are produced outside the decay volume can reach the detector if they are not further outside than two gyration radii. On the other hand, some of the electrons which are produced inside the decay volume, but closer than 2 gyration radii to the edge will not be detected. The yields of these two types of events cancel exactly if the neutron beam profile were homogeneous. Our experiment will be in the region of homogeneous neutron beam intensity shortly after the neutron guide. However, the homogeneity will not be perfect. Higher electron energies tend to have larger gyration radii, and as such will take a different beam profile average. The calculation of the problem can be done analytically [34] or by Monte Carlo, in conjunction with the definition of the beam profile. Experience from other experiments shows that this systematic is important, but manageable. We have an additional advantage since our detector is pixelated and measures an image of the neutron beam *in situ*.

4.6. Backgrounds

1. Neutron beam related background

   Background events are an important consideration in the experiment. Background rates are difficult to estimate a priori as they arise from the very intense primary neutron beam, but having a low probability of producing detectable particles in the detectors. In addition, the rates depend sensitively on the design of the experiment, in particular the collimation system, shielding, and beam stop. These aspects of the experiment will have to be carefully optimized to reduce background. We can estimate the amount of background in the following way: About 3×10^{-9} of the neutrons that come out of the beam guide decay in our decay volume. The other neutrons are absorbed in the collimation system or on the beam stop by diaphragms made from ^{6}LiF. In this process, neutron and γ background is produced at the 10^{-4} level [35]. The background is further reduced by shielding ($< 1\%$), the solid angle (5×10^{-4}), and the efficiency ($< 5\%$) of our detectors, so that we expect the background rate to be smaller than the signal rate.

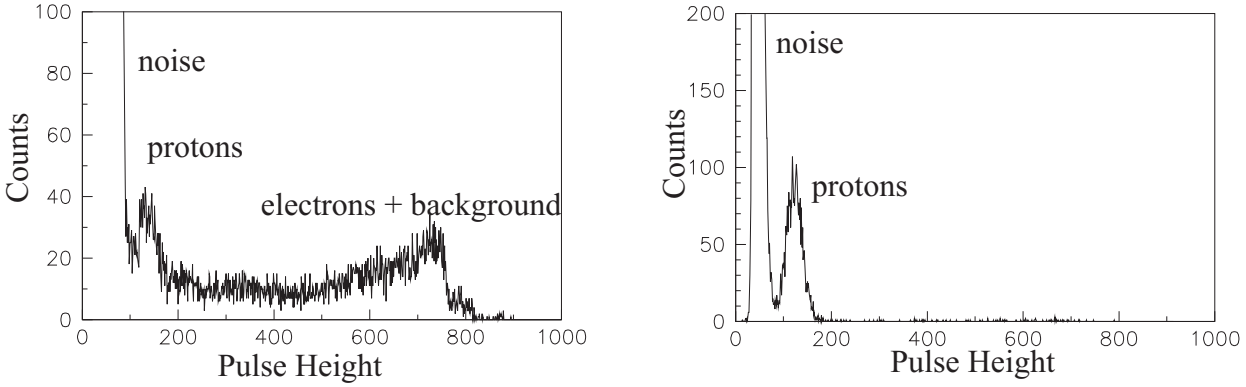

Figure 16: Plots of protons from neutron decay in singles (left) and coincidence from a test measurement at NIST. See the text for details.

An important feature of the Nab experiment is the detection of both proton and electron from the neutron decay in coincidence. We conducted a demonstration of this technique by observing electron-proton coincidences from neutron decay using a silicon surface-barrier detector coupled to the NIST lifetime apparatus. A 30 keV potential was used to accelerate protons. The detector area was approximately the size of one pixel of the Nab detector. The geometry was less ideal than the Nab configuration, as the detector was placed much closer to the neutron beam and the final collimator ($\sim$ 10 cm). Since the NIST lifetime experiment was designed to count decay protons after the neutron been was shut off, this was not a limitation for their experiment. Fig. 16 illustrates the effect of coincidence on suppressing backgrounds. The left-hand plot shows the observed singles spectrum. The 30 keV proton peak can be clearly seen, as well as electronic noise, the (distorted) electron spectrum, and background events. The electron spectrum is distorted because the detector was only 300 μm thick, too thin to stop all of the electrons. The right-hand plot shows a spectrum of proton energy in coincidence with decay electrons, greatly suppressing backgrounds. The rates of false coincidences due to these backgrounds are discussed above in the section on trigger hermiticity.

We are aware of the fact that typically there is neutron beam halo around the neutron beam which can hit surfaces in the spectrometer that are not covered with ^{6}LiF, and then cause background whose amount can exceed the direct background if the latter is shielded well. The origin of the halo is likely scattering on the edges of the diaphragms or on impurities in them. Our strategy is to cover most surfaces seen by the neutron beam with neutron absorbing material. We can do this where it matters most, i.e., in the collimation system and before the beam stop. We cannot cover the electrode system. If the halo causes too much secondary background, we can always collimate the neutron beam earlier and harder.

2. Particle trapping

Electrons can be trapped in the decay volume, expansion, and TOF regions. These regions form an electrode-less Penning trap. The Penning trap doesn't cause a problem

for electrons above our energy threshold. The longitudinalization of the electron momentum due to the magnetic field allows all of them to escape and to reach the detector. Electrons from field ionization or from rest gas intercations are a concern. Trapped particles ionize the rest gas, and the rest gas ions form a usually time-dependent background. In the worst case high voltage breakdowns and discharges can occur if the density of trapped particles becomes to high. The filling time constant due to electrons from neutron decay is long and can be neglected. Experience from other experiments (e.g, the NIST lifetime measurement [29]) shows that typically a vacuum of 10^{-8} torr has to be reached. There are strategies to remove the particles in a penning trap; wire grids around the decay volume would do it, but they cause additional systematic uncertainties. An electric field perpendicular to the magnetic field would remove particles, but would spoil the measurement of proton flight times [30]. If the filling time constant is slow enough, it can be sufficient to empty the trap from time to time, by swiping a wire through the trap, as it is planned in the neutrino mass spectrometer KATRIN [31], or by ramping down the high voltage.

Another problem are the magnetic field lines which connects two HV electrodes. The electrodes act as cathodes and on the magnetic field line in between them the electric potential has a minimum. That field configuration forms a Penning trap, but now with material endcaps, which is much more severe. If the depth of the trap is of the order of 1 kV or more it causes Penning discharges as the electrodes give rise to the well-known multiplication mechanism of a Penning discharge. The discharge starts with a positive rest gas ion which is accelerated towards the electrode. The impact produces several secondaries, which can be trapped if they collide with a rest gas atom, or due to non-adiabatic processes which prevent them from being reabsorbed at one of the cathodes. Then in turn, they ionize rest gas atoms. The remedy is a proper design of the electrodes which avoids these traps.

5. Dominant uncertainties in the measurement of b

The careful attention to understanding details of the spectrometer response to both protons and electrons discussed above also works toward reducing the systematic uncertainties in measuring b. Measuring b amounts to a precise determination of the shape of the electron spectrum and its deviation from the $b = 0$ shape predicted by the standard model. Any such deviation will be far more pronounced at low electron momenta than for high momenta. Therefore the accurate understanding of background rates (which increase with decreasing pulse size), of momentum dependence of electron detection efficiency, especially near threshold, of electron spectrum distortions due to reflections at the detector surface and elsewhere in the spectrometer, must all be controlled. The calibration methods outlined above generally bring us to the desired level. However, to date we have not produced a realistic Monte Carlo simulation of the background. Such simulations can only be taken as a guideline; eventually the backgrounds will have to be measured and dealt with *in situ* because of the complexity of the apparatus.

The statistical uncertainty in b is given in the following table, which is in agreement with the calculations of [32].

Quantity	No threshold	$E_{\mathrm{e,min}} = 100\,\mathrm{keV}$	$E_{\mathrm{e,min}} = 200\,\mathrm{keV}$	$E_{\mathrm{e,min}} = 300\,\mathrm{keV}$
σ_b	$7.5/\sqrt{N}$	$10.1/\sqrt{N}$	$15.6/\sqrt{N}$	$26.3/\sqrt{N}$
σ_b (E_{cal} variable)	$7.7/\sqrt{N}$	$10.3/\sqrt{N}$	$16.3/\sqrt{N}$	$27.7/\sqrt{N}$

Here, N is the number of neutron decays. N is not restricted to the subset where electron and proton pass the respective cutoff conditions. The statistical uncertainty of a measurement of b quickly increases if a high electron energy cutoff is chosen. The reason is that at higher electron energies a non-zero value of b looks similar to a change of the normalization N. Therefore, the fitting parameters N and b are strongly correlated.

In the second line of the table the statistical uncertainty in b is given if the energy calibration is determined from a fit to the beta spectrum. This possibility is discussed in Sect. 4. The statisctical uncertainty is not much worse than with a fixed energy calibration.

For an overall sample of 5×10^9 decays, the statistical uncertainties σ_b calculated on the basis of the above table are in the range of $\sigma_b \sim 10^{-4}$, far better than our goal of overall uncertainty of $\sim 10^{-3}$. It is clear that the main challenge in the measurement of b will be in the systematics. While we have indicated strategies for controlling the systematics, a major contribution will be the electron background, which will ultimately have to be evaluated and dealt with *in situ*.

6. Summary

We have developed a method for a simultaneous measurement of the neutron decay parameters a, the neutrino-electron correlation, and b, the Fierz interference term, with precision of a few parts in 10^3. The measurement of a would be competitive with the best measurements of the neutron decay asymmetry parameter A. Our proposed measurement of b would be the first ever in neutron decay.

We propose to make the measurements in the Fundamental Neutron Physics Beamline at the SNS. The experiment will require a dedicated 4π spectrometer with combined electric and magnetic fields, and novel Si detectors. Details of the design of the spectrometer are currently still under development, as is work on the Si detectors. Extensive analytical calculations and Monte Carlo simulations of detector response and effects on systematic uncertainties indicate that the Nab spectrometer should be fully compatible with the requirements of the other neutron decay experiments planned for the FNPB, most notably of abBA. While a broad family of electromagnetic field profiles provide satisfactory statistical sensitivities for Nab measurements, a sharply expanding field in the immediate vicinity of the decay region provides the best control of the systematics. Precise parameters of this field profile are still under study. However, the additional flexibility required to accommodate both the sharply expanding magnetic field (Nab) and the more gently expanding field (abBA), does not appear to introduce either technical roadblocks or a prohibitive increase in overall cost.

Given the planned neutron beam intensity and the spectrometer design under consideration, we propose a series of measurements after the commissioning of the abBA/Nab spectrometer. The Nab production run would take some 5000 hours and would accumulate some 5×10^9 events.

References

1. W-M Yao, et al., J. Phys. G **33**, 1 (2006).
2. J.C. Hardy, I.S. Towner, Phys. Rev. C **71**, 055501 (2005).
3. F.J. Gilman, K. Kleinknecht and B. Renk, in *"Review of Particle Physics"*, K. Hagiwara et al., Phys. Rev. D **66**, 01001-113 (2002).
4. A. Sher et al. (BNL 865 Collaboration), Phys. Rev. Lett. **91**, 261802 (2003).
5. T. Alexopoulos et al. (KTeV Collaboration), Phys. Rev. Lett. **93**, 181802 (2004).
6. A. Lai et al. (CERN NA48 Collaboration), Phys. Lett. B **602**, 41 (2004).
7. A. Czarnecki, W.J. Marciano, and A. Sirlin, Phys. Rev. D **70**, 093006 (2004).
8. H. Abele et al., Phys. Rev. Lett. **88**, 211801 (2002).
9. F.E. Wietfeldt et al., Nucl. Inst. Meth. in Phys. Res. A **545**, 181 (2005).
10. D. Mund, Ph.D. Thesis, University of Heidelberg, 2006, (http://www.physi.uni-heidelberg.de/physi/publications/doc_mund.pdf).
11. C. Stratowa et al., Phys. Rev. D **18**, 3970 (1978).
12. J. Byrne et al., J. Phys. G **28**, 1325 (2002).
13. P. Herczeg, Prog. Part. Nucl. Phys. **46**, 413 (2001).
14. N. Severijns, M. Beck, O. Naviliat-Cuncic, Rev. Mod. Phys. **78**, 991 (2006).
15. M.V. Chizhov, e-prints hep-ph/0402105, hep-ph/0411098, and references therein.
16. T.M. Ito and G. Prézaeu, Phys. Rev. Lett. **94**, 161802 (2005).
17. S. Profumo, M. J. Ramsey-Musolf and S. Tulin, Phys. Rev. D **75**, 075017 (2006).
18. S. Gardner and C. Zhang, Phys. Rev. Lett. **86**, 5666 (2001).
19. S. Gardner, in "Quark-Mixing, CKM-Unitarity," Proceedings, ed. by H. Abele and D. Mund, (Mattes Verlag, Heidelberg, 2003), arXiv:hep-ph/0312124.
20. F. Ayala Guardia, Diploma Thesis, University of Mainz, 2005.
21. http://geant4.web.cern.ch/geant4/ .
22. J.D. Bowman, technical note, http://nab.phys.virginia.edu/self_calib.pdf .
23. V.V. Nesvizhevsky et al., Nucl. Inst. Meth. A **440**, 754 (2000).
24. V.V. Nesvizhevsky et al., Eur. Phys. J. C **40**, 479 (2005).
25. M.S. Dewey et al., Phys. Rev. Lett. **91**, 152302 (2003).
26. R. Dobrozemsky, Nucl. Inst. Meth. **118**, 1 (1974).
27. M. Schumann et al., arXiv:0708.3150.
28. H. Abele, S. Baeßler et al., Phys. Lett. B **407**, 212 (1997).
29. J. Nico, private communication, (2007).
30. F. Glück et al., Eur. Phys. J. A **23**, 135 (2005)

31. J. Bonn, private communication, (2005).
32. F. Glück et al., Nucl. Phys. A **593**, 125 (1995).
33. A.P. Williams, Ph.D. thesis, University of Sussex, Brighton, 1989.
34. D. Dubbers et al., arXiv:0709.4440.
35. M.A. Lone et al., Nucl. Inst. Meth **174**, 521 (1980).

Appendix A: Proton time of flight in the acceleration region

This section addresses the question how to incorporate the acceleration region in the spectrometer response function. The presence of the electric field modifies the relationship between $t_{\rm p}$, the proton time of flight, and $p_{\rm p}$, proton momentum by introducing a dependence on both r and p^2, as seen in Fig. 17. The probability distribution function then becomes

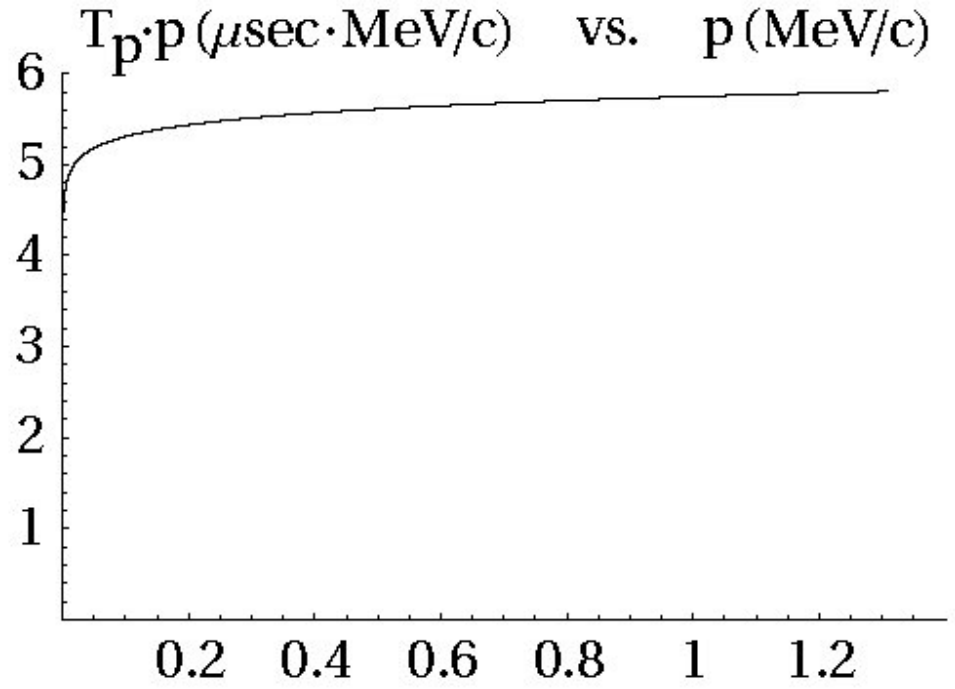

Figure 17: The product of proton time of flight and proton momentum plotted against proton momentum. If the electric field were absent, the product $t_{\rm p} \cdot p_{\rm p}$ would be constant.

$$P_t\left(\frac{1}{t^2}, p^2\right) = \int P_p(p^2)\Phi\left(\frac{1}{t^2}, p^2\right)\, dp^2\,, \tag{40}$$

where $\Phi(1/t^2, p^2)$ is the spectrometer response. The first step in dealing with this complication is to transform $t_{\rm p}$ in such a way that the resolution function is nearly independent of p^2. The strategy is to calculate the relationship of $t_{\rm p}$ on $p_{\rm p}$ for $u = \cos\theta_{pB} = 1$, i.e., for proton direction along a magnetic field line. We then solve (deconvolve) the relation between time of flight and momentum for each experimental value of p. We apply the transformation and calculate the probability density function (PDF) for a uniform distribution of u for $-1 < u < 1$, and show that to a good approximation the PDF's are independent of p. We show results for momenta that are 0.9 and 0.1 of the maximum value $p_{\rm max} = 1.188\,{\rm MeV/c}$ in Fig. 18.

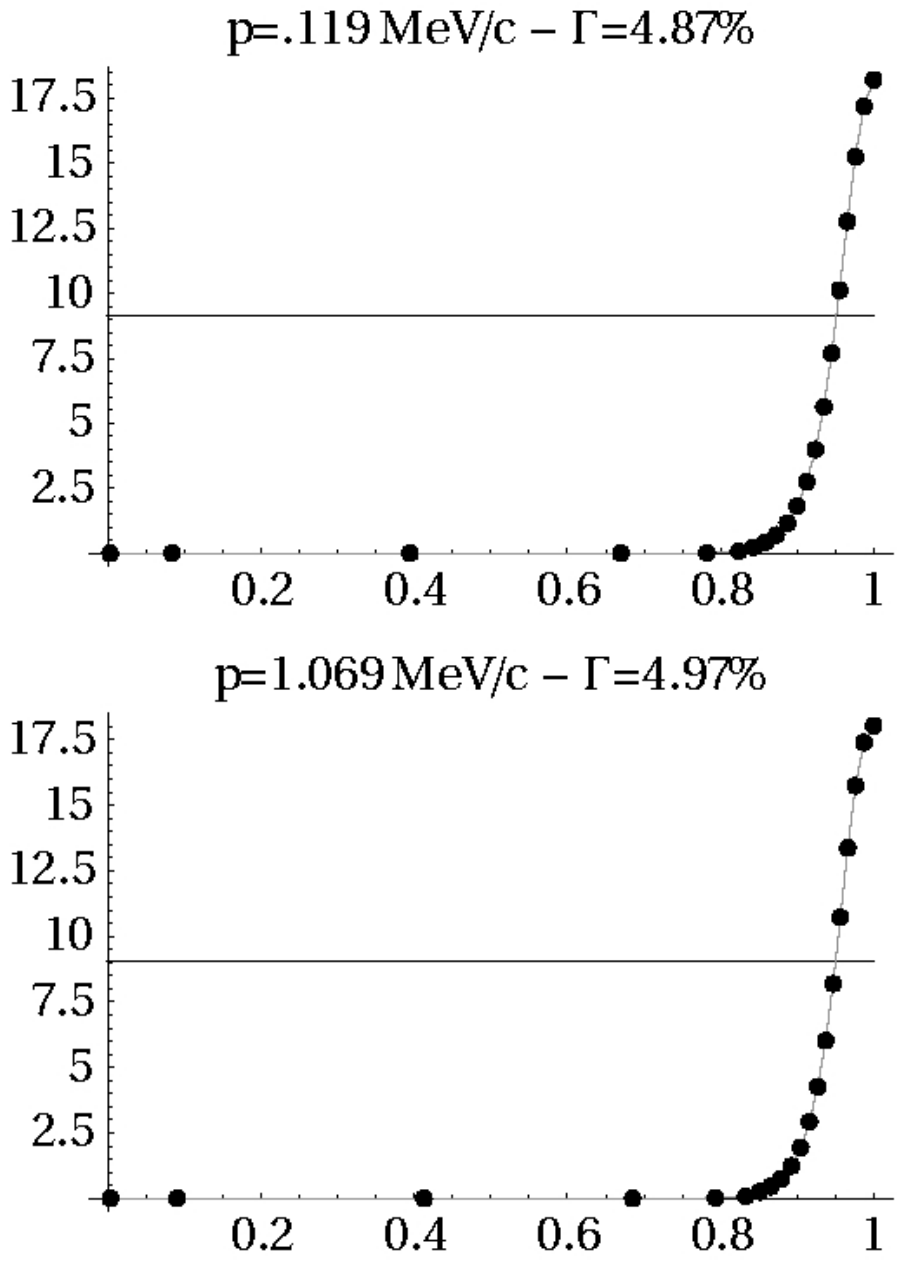

Figure 18: Spectrometer resolution function PDF of deduced proton momentum for $p = 119$ and 1,069 MeV/c. The widths differ by 2 % of their values. The variation of the product $t_{\rm p} \cdot p_{\rm p}$ before the mapping is $\sim 15\,\%$. The variation of the fractional width of $1/t_{\rm p}$ is similar. Application of the mapping reduces the width variation by a factor of 7.5 over the full range of proton momenta.

Appendix B: Optimization of the magnetic field

A narrow resolution function gives a small increase in the uncertainty in a above the uncertainty that would be obtained if the spectrometer fields were known perfectly. We investigate how the width of the resolution function depends on properties of the spectrometer such as the field expansion ratio, r_B, (the ratio of the decay field to the drift field), γ, the curvature of the decay field defined as

$$\gamma = -\frac{1}{B}\frac{d^2B}{dz^2}\,,$$

and d_z the width of the neutron beam along the spectrometer axis, z. We find that Γ, the width of the resolution, decreases as γ increases. Similarly, Γ decreases as r_B increases (see left panel of Fig. 19). For reasonable values of γ and r_B, a beam width of $d_z = 2\,\mathrm{cm}$ (which gives a count rate of ∼400 Hz) produces a small increase in the width over that from a "pencil-thin" beam and the same γ and r_B values.

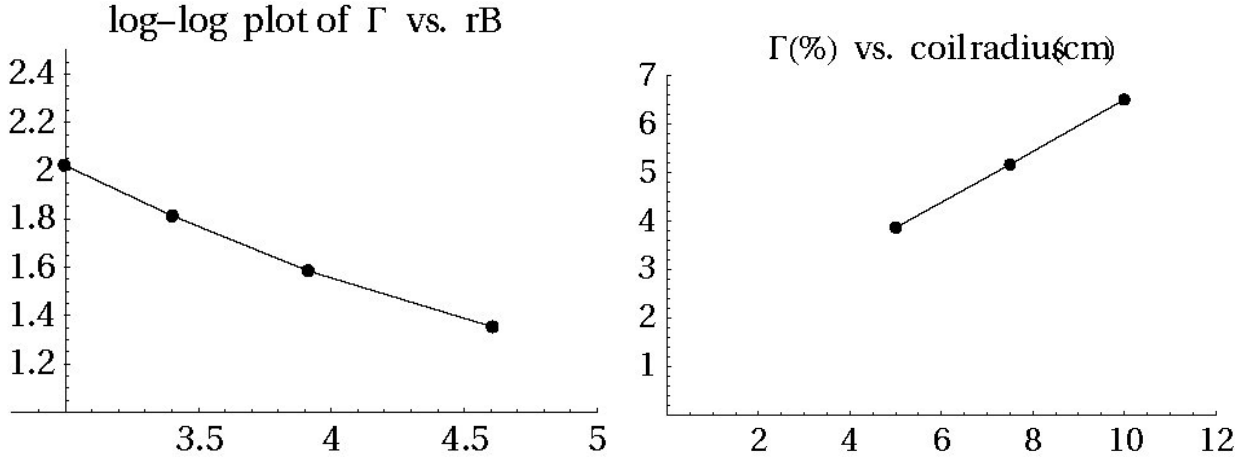

Figure 19: Left panel: Plot of the width of the response function against the expansion ratio. The width decreases as the expansion ratio increases. Right panel: Plot of spectrometer resolution against the decay region coil separation.

We have calculated the width of the resolution function for field expansion ratios r_B of 20, 30, 50, and 100, respectively. The results are shown in Fig. 19 (left). Over a range of reasonable values, the width $\Gamma \sim 1/rB^{0.4}$. This behavior can be understood from Eq. (19). The protons spend most of their time in the drift region, where $B(s)/B(0) \sim 1/r_B$. A large value of r_B decreases the influence of the $(1-u^2)$ term that causes a variation of the velocity in the drift region. If the electric field were zero, the width variation would be $\sim 1/\sqrt{rB}$.

We have studied the dependence of the width of the resolution function on the rate of decrease of the magnetic field from its maximum value at $z = 0$ to the drift region. This was done by changing the radius of the split pair of coils that create the field while keeping the ratio of the separation to the diameter constant at 0.3. Fig. 19 (right) shows a plot of the width of the resolution function vs. coil diameter for a field expansion ratio $r_B = 100$. The resolution improves as the coil diameter is decreased. This behavior can also be understood from Eq. (19). The z-component of the proton velocity is $\propto u$ before the field expansion.

For large coil radii the proton spends more time before the angle between the momentum is decreased by the field expansion. However, the coil diameter and separation cannot be decreased indefinitely because with no separation no neutrons can enter the decay region. A radius of 5 cm and a separation of 3 cm allow a beam width of $d_z = 2\,\text{cm}$ along z and a count rate of 400 Hz.

Finally, we assumed a field expansion of 100, a coil radius of 5 cm, and a separation of 3 cm, and calculated the change in the width of the spectrometer response as a function of d_z, the full neutron beam width along z, for $d_z = 0$, 1, and 2 cm. The width of the response was constant to within 5 % of its value. The PDF $Q(r)$, the factor that multiplies p2 in Eq. 19, is given in Fig. 20. The width of $Q(r)$ is 8.48 %. $Q(r)$ is the distribution of p^2, and is approximately twice as wide as the distributions of p above. Since the width of $Q(r)$ is approximately constant, we chose $d_z = 2\,\text{cm}$, the largest beam that will fit between the decay field coils, in order to maximize the count rate at ∼400 Hz.

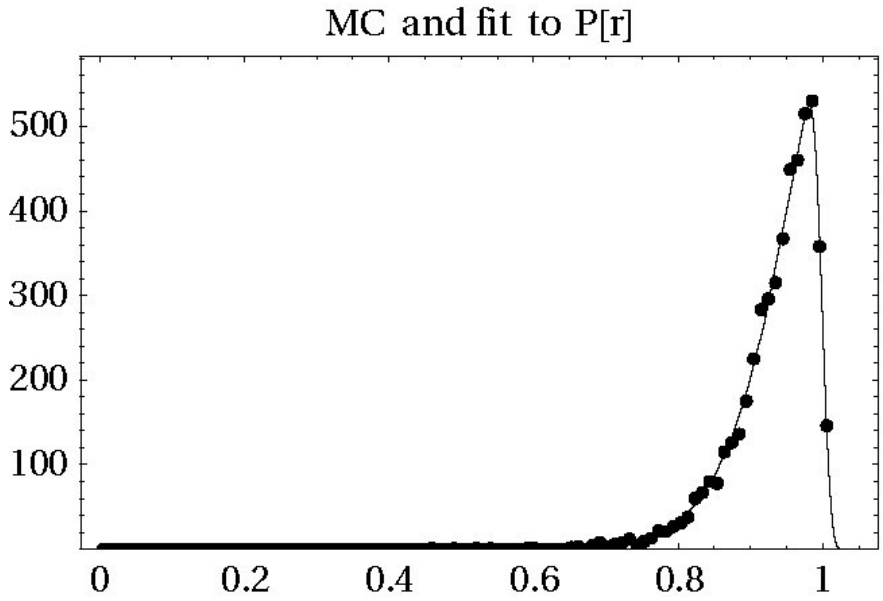

Figure 20: The probability density of $Q(r)$ in Eq. (19). The distributions of p^2 are broadened by $Q(r)$.

The best approach to minimizing the uncertainty in a is to design a spectrometer that has a narrow response function as described above. We illustrate this point by increasing the width of the response function by a factor of 3. The results are shown if Fig. 22.

In conclusion, we find that for an attainable spectrometer configuration, the yield vs. proton momentum squared and electron kinetic energy can be used to check the spectrometer response calculated determined from field measurements. If we carry out the checks in the commissioning phase we can validate the measured fields. We can then use the fitting procedure to constrain the spectrometer response. The fitting procedure increases the statistical uncertainty factor in a from 2.3 to 2.6, a small price to pay in order to obtain a robust measurement of a from an *in situ* check of the measured fields.

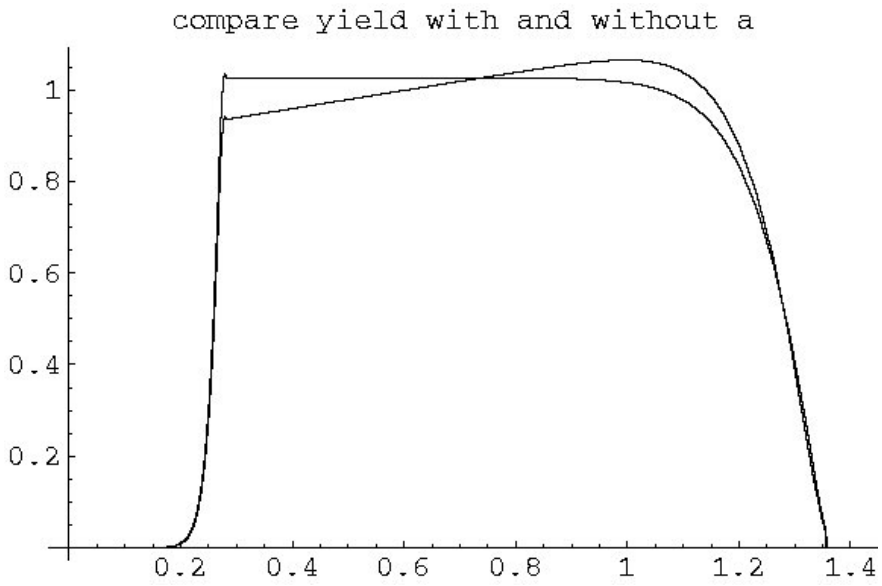

Figure 21: Calculated yield for the spectrometer response function shown in Fig. 20.

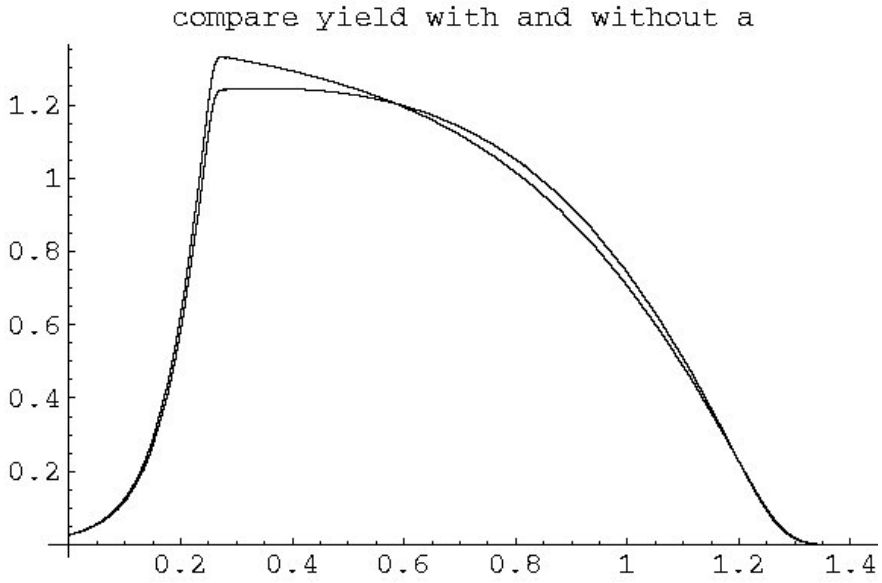

Figure 22: Calculated yield for the spectrometer response function shown in Fig. 20 broadened by a factor of 3. The features in the yield spectrum that can be used to check the spectrometer calibration, such as the breaks near $z = 0.25$ and $z = 1.3$, are washed out leading to larger uncertainty in their determination.