

Simulation of Electron Neutrino Mass Measurements by Time-of-Flight with KATRIN

Nicholas Steinbrink

Diplomarbeit

Institut für Kernphysik
Mathematisch-Naturwissenschaftliche Fakultät
Westfälische Wilhelms-Universität Münster

Mai 2012



WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

Referent: Prof. Dr. C. Weinheimer

Korreferent: Prof. Dr. J. Wessels

Heute habe ich etwas Schreckliches getan, etwas, was kein theoretischer Physiker jemals tun sollte. Ich habe etwas vorgeschlagen, was nie experimentell verifiziert werden kann.

Wolfgang Pauli

Contents

1	Introduction	7
2	The Neutrino Mass	9
2.1	Basics	9
2.1.1	The Standard Model	9
2.1.2	Standard Model Neutrinos	10
2.1.3	Neutrino oscillations	11
2.1.4	Theory of the Neutrino Mass	12
2.2	Measurement of Neutrino Masses	13
2.2.1	Motivation	13
2.2.2	Measurement Methods	15
3	The KATRIN Experiment	19
3.1	Basics of Tritium Beta Decay Experiments	19
3.1.1	The Tritium Beta Spectrum	19
3.1.2	The MAC-E-Filter	21
3.2	Design of KATRIN	24
3.2.1	Design Goals and Parameters	24
3.2.2	Main Components	25
3.3	Data Analysis of KATRIN	26
4	Simulation of a Time-Of-Flight Mode for KATRIN	29
4.1	Idea and Motivation	29
4.2	Mathematical description	30
4.2.1	TOF for a Single Electron	30
4.2.2	TOF Spectrum	32
4.3	Numerical Realization	34
4.3.1	Numerical Calculation of the TOF Spectrum	34
4.3.2	Fit Preparations	35
4.3.3	Fit routine	37
4.4	Simulation Results	38
4.4.1	Characteristics of TOF Spectra	38
4.4.2	Fit Example of Basic TOF Spectra	44
4.4.3	Measurement Time Optimization	46

5	Simulation of the Measurement Method	51
5.1	Motivation	51
5.2	Hypothetical Electron Tagging	52
5.2.1	Motivation and Possible Approaches	52
5.2.2	Mathematical Model	53
5.2.3	Simulated Effects of the Tagging Parameters	55
5.2.4	Prediction of the Electron Tagging Performance	57
5.3	Gated Filter	59
5.3.1	Principle and Mathematical Model	59
5.3.2	Simulated Effects of the Timing Parameters	60
5.3.3	Predictability of the Gated Filter Performance	62
6	Conclusion	65
6.1	Statistical Potential of a TOF Mode	65
6.2	Measurement Method - Fiction vs. Reality	67
6.2.1	General Constraints for Measurement Methods	67
6.3	Open Questions	68

1 Introduction

When the neutrino hypothesis was proposed by Pauli in 1930 it was merely an *ad hoc hypothesis* to explain the issue of missing energy in the radioactive beta decay [Pau30]. As the cumulation of such *quantitative anomalies* leads to a scientific revolution in the sense of T.S. Kuhn [Kuh62], the neutrino became later an integral part of the *standard model*, the leading paradigm in the particle physics of the last decades and one of the most successful physical theories. With the results from the Super Kamiokande Experiment in 1998, however, it became evident that neutrinos oscillate and thus have a non-zero rest mass contradictory to the standard model. Currently there are further evidences that the standard model is incomplete. These comprise for instance the existence of non-baryonic dark matter, which could be resolved in terms of cold dark matter (CDM) as lightest supersymmetric particle (LSP) or even in terms of warm dark matter (WDM) as keV sterile neutrino. Further evidences for sterile neutrinos are given by the reactor anomaly [Men11].

To find an empirically adequate successor of the standard model, a measurement of the electron neutrino mass is a vital objective. Since oscillation experiments like Super Kamiokande or SNO can only resolve the squared mass differences between the neutrino mass eigenstates, but not their absolute values, the need for a direct measurement arises. This will not only be important to obtain the mass scale of the neutrinos and thus understand their mass generation mechanism. Neutrinos played also a central role in the early universe, so knowing the neutrino masses is important to understand structure formation in cosmology.

Currently, the most promising experiment in the near future for that purpose is the KATRIN experiment built in Karlsruhe. It is designed to measure the electron neutrino mass with a sensitivity of 0.2 eV (90 % C.L.). The sensitivity of KATRIN is foremost constrained by the size of its main spectrometer and source diameter. To achieve even larger sensitivities, KATRIN or a KATRIN like experiment would have to become even larger or to be complemented with additional measurement methods.

The following thesis will investigate the scenario of using a time-of-flight (TOF) spectroscopy of the beta decay electrons in the KATRIN main spectrometer in order to determine the electron neutrino mass. A simulation has been performed to determine the statistical uncertainty on the neutrino mass when using this method. This has both been done assuming an ideal TOF measurement and for different hypothetical measurement methods, respectively.

- *Chapter 1* will reflect the basics of neutrino masses which includes the theory of

neutrino oscillation and neutrino mass generation, empirical evidences for these, the physical impact of the neutrino mass as well as an overview over the methods of measurement.

- With the KATRIN experiment, a description of a particular of such experiment is given in *Chapter 2*. It depicts the basic principles of tritium beta decay experiments, followed by the the implementation of these principles in KATRIN by its design goals and an overview of the components.
- *Chapter 3* investigates the addition of a TOF spectroscopy mode to KATRIN. It starts with the idea and principle of using TOF for neutrino mass determination, a basic mathematical model, continues with numerical methods for a determination of the statistical uncertainty and finishes with simulation results for an ideal TOF measurement.
- *Chapter 4* finally discusses two hypothetical scenarios of measuring TOF inside KATRIN and implements them with different characteristics into the simulation to evaluate their potential.
- The results are recapitulated in *Chapter 5* and a conclusion regarding the supplementation of KATRIN with TOF spectroscopy is given.

2 The Neutrino Mass

To understand the KATRIN experiment in its principle and motivation, an overview over the physics of neutrino masses has to be given. That comprises their experimental discovery, the corresponding theory and the methods how neutrino masses can be measured.

2.1 Basics

2.1.1 The Standard Model

The standard model is effectively the fundamental theory of particle physics despite the evidences for physics beyond it. It was successful in reducing the huge "particle zoo" to a finite number of elementary particles. They are results of a quantization of the underlying fields. What we call "matter" is in this model composed of spin 1/2-fermions while the interactions between them are carried out by spin 1-bosons. The fermions can be grouped into quarks (*up*- and *down*-type, respectively) and leptons (charged ones and electrically neutral *neutrinos*). For each of them there exists a corresponding anti-particle, too.¹ The number of gauge bosons is mathematically determined by the generators of the *gauge group* they correspond to. The number of fermions is not a priori limited, yet experiments suggest that for each of the fermion types stated above there are three generations [ALE89].

The interactions result from an $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge invariance of the *Lagrangian* \mathcal{L} . The $SU(3)$ group describes the strong interaction mediated by 8 different gluons. These can only couple to themselves and the quarks which have one of three different color charge states C (plus three anti-color states for the anti-quarks) each. $SU(2)$ and $SU(1)$ together form the electroweak interaction where the $SU(2)$ is represented by the weak isospin I_3 , mediated by the three W bosons W^+ , W^- and W^0 and the $U(1)$ by the weak hypercharge $Y = 2(Q - I_3)$, mediated by the B^0 boson. Due to the maximal parity violation of the weak interaction which was detected in the Wu experiment [Wu57], the $SU(2)$ bosons couple only to particles with left-handed chirality, thus the subscript L .²

¹For the neutrinos, however, it is not yet clear if neutrinos and antineutrinos are identical.

²That means that they are eigenvalues of the left-handed chirality projection operator $P_L = \frac{1}{2}(1 - \gamma_5)$ where γ_5 is one of the Dirac matrices. In the relativistic limit, the chirality coincides with the helicity which is the projection of the spin on the momentum direction.

All masses are generated by the Higgs Mechanism. This introduces a $SU(2)$ doublet field $\varphi = (\varphi^+, \varphi^0)$ that is characterized by a non-zero *vacuum expectation value* v which means effectively that the Higgs Field is excited in the vacuum. Thus, the electroweak $SU(2)_L \times U(1)_Y$ symmetry is broken to the electromagnetic $U(1)$ symmetry. In this process, the bosons and fermions get masses by coupling to the Higgs field. The fundamental B^0 and W^0 bosons superpose to the Z^0 bosons and the photon γ . After the transformation that gives the W bosons the masses and their polarization degrees of freedom, one degree of freedom of the Higgs field remains which corresponds to the Higgs particle.³

Three Generations of Matter (Fermions)				
	I	II	III	
mass→	2.4 MeV	1.27 GeV	171.2 GeV	0
charge→	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0
spin→	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
name→	u up	c charm	t top	γ photon
Quarks	4.8 MeV	104 MeV	4.2 GeV	0
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	d down	s strange	b bottom	g gluon
Leptons	<2.2 eV	<0.17 MeV	<15.5 MeV	91.2 GeV
	0	0	0	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	Z^0 weak force
	0.511 MeV	105.7 MeV	1.777 GeV	80.4 GeV
	-1	-1	-1	± 1
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
	e electron	μ muon	τ tau	W^\pm weak force
				Bosons (Forces)

Figure 2.1: Particle content of the standard model. Source: [LAG]

To summarize, all the elementary particles postulated by the Standard Model are shown in fig. 2.1.

2.1.2 Standard Model Neutrinos

It follows from the standard model that neutrinos belong to the leptons and have no electric charge, so they can interact only weakly or by gravitation. They carry a weak isospin of $I_3 = 1/2$. There is one neutrino for each fermion generation, thus there exist three neutrinos ν_e, ν_μ, ν_τ . They carry a lepton number of $L = 1$ and a lepton family number of $l_\alpha = 1$ corresponding to their generation $\alpha = e, \mu, \tau$. For anti-neutrinos, it becomes $L = -1$ and $l_\alpha = -1$, respectively. In the standard model these are strictly conserved.

³First hints for the existence of the Higgs particle have been currently found at the LHC [ATL12] [CMS12].

Historically, neutrinos were first proposed by Pauli to solve the problem of missing energy in the beta decay [Pau30]. Since in a two body decay the electron should be monoenergetic due to energy and momentum conservation, a third particle was postulated which carries away the missing energy. It could later be confirmed in the Poltergeist experiment [Rei59]. Experiments like the famous one by Goldhaber [Gol75] have furthermore shown that neutrinos have only left-handed helicity. Thus, neutrinos have to be effectively relativistic particles which means that they have a mass of almost zero. Therefore, in the standard model it is assumed that neutrinos are massless particles.

2.1.3 Neutrino oscillations

The discovery of neutrino oscillations in the past decades and the consequential violation of lepton family number conservation has shown that there is physics beyond the standard model when it comes to neutrinos. First hints for neutrino oscillations were found with the Homestake experiment which became known as the *solar neutrino problem*. By neutrino capture reactions on chlorine the neutrino flux from the sun should be measured. The measured flux was significantly lower than expected from the solar standard model [Cle98]. Neutrino oscillations became favoured with the results from the Super-Kamiokande experiment [SUP98]. Super-Kamiokande was able to detect electron and muon neutrinos by Cherenkov radiation which is caused from electrons and muons created in neutrino reactions within a large water tank. In the detection of atmospheric muon neutrinos Kamiokande showed a deficit of the expected neutrino flux from the downward hemisphere. This was interpreted as a sign of oscillations of ν_μ into ν_τ since neutrinos created on the other side of the earth travel a longer path and have thus a higher oscillation probability. Since Super-Kamiokande is not sensitive on tau neutrinos this leads to a deficit.

This behaviour can be explained by a non trivial mixing between the flavour eigenstates (ν_e, ν_μ, ν_τ) and three different mass eigenstates (ν_1, ν_2, ν_3) of the neutrinos, implying that neutrinos have masses. The mechanism which is quite similar to the quark mixing in the standard model can be described by a unitary 3×3 matrix, called the *Pontecorvo–Maki–Nakagawa–Sakata* (PMNS) matrix,

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = U_{\text{PMNS}} \cdot D_M \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \quad (2.1)$$

with

$$U_{\text{PMNS}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (2.2)$$

and

$$D_M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\lambda_2} & 0 \\ 0 & 0 & e^{i\lambda_3} \end{pmatrix}. \quad (2.3)$$

Here, s_{ij} and c_{ij} stand for $\sin\theta_{ij}$ and $\cos\theta_{ij}$ where θ_{ij} are the three mixing angles. The diagonal matrix D_M contains the Majorana phases and is deviant from unity if neutrinos are Majorana particles, i.e. identical to their own antiparticles. δ_{13} is the CP violating Dirac phase and is $\neq 0$ if neutrino oscillation violates CP conservation. When neutrinos are produced in a distinct flavour eigenstate, they are in a mixture of mass eigenstates and thus energy eigenstates. Applying the time-dependent Schrödinger equation on this state gives different frequencies in the phase evolution for each mass eigenstate contribution. It follows that neutrino flavours are oscillating with time. The oscillation probability between two flavours can be approximated by

$$P_{\nu_\alpha \rightarrow \nu_\beta} = \sum_{k,j} U_{\alpha k}^* U_{\beta k} U_{\alpha j} U_{\beta j}^* \exp\left(-i \frac{\Delta m_{kj}^2 L}{2E}\right) \quad (2.4)$$

where Δm_{kj}^2 is the squared mass difference between the mass eigenstates k and j , L the propagated distance, called *baseline* and U the mixing matrix [Giu07].

2.1.4 Theory of the Neutrino Mass

As it was shown, the theoretical description of neutrino oscillations requires that the neutrino flavours are superpositions of neutrino mass eigenstates. This implies that neutrinos are not massless. The standard model has thus to be extended by a mass-generation mechanism for the neutrinos. In the standard model Lagrangian there is no right handed chiral neutrino singlet, in contrast to the other fermions. This leads naturally to a neutrino with zero mass and only left handed helicity. In order to account for a non-vanishing neutrino mass, one could include a right handed neutrino term. Taking a simple model considering only one generation, this would result in an additional term in the Higgs lepton part of the SM Lagrangian.

$$\mathcal{L}_{\text{mass,D}} = \mathcal{L}_{\text{mass,SM}} - y_{\nu_e} (\bar{\nu}, \bar{e})_L \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}^C \nu_R + \text{H. c.} \quad (2.5)$$

This term is called *Dirac mass term*. Expanding around the Higgs vacuum expectation value v and applying a local $U(1)$ gauge transformation yields

$$\mathcal{L}_{\text{mass,D}} = \mathcal{L}_{\text{mass,SM}} - m_\nu \bar{\nu}_L \nu_R + \text{H. c.} \quad (2.6)$$

where

$$m(\nu) = \frac{y_{\nu_e} \cdot v}{\sqrt{2}}. \quad (2.7)$$

Unfortunately, since neutrino masses are very low, the Yukawa coupling constants for the neutrinos y_{ν_e} had to be several orders of magnitude smaller compared to those for the other fermions which are nearly equal.

Therefore, one can alternatively consider *Majorana mass terms*. These require, however, that neutrino and antineutrino are identical and thus violate lepton number conservation. A Lagrangian that combines Dirac and Majorana mass terms can be written in matrix form.

$$2\mathcal{L}_{\text{mass,D+M}} = \mathcal{L}_{\text{mass,SM}} - (\bar{\nu}_L, (\bar{\nu}_R)^C) \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} \begin{pmatrix} (\nu_L)^C \\ \nu_R \end{pmatrix} \quad (2.8)$$

The matrix can be diagonalized with the two eigenvalues m_1 and m_2 .⁴ For the special case of $m_L = 0$ we obtain

$$m_1 = \frac{m_D^2}{m_R} \quad m_2 = m_R. \quad (2.9)$$

This is the *seesaw mechanism type I* [Giu07]. If m_1 is low, then m_2 becomes high and vice versa. Thus, it provides a good explanation why the neutrino masses are so low. It allows us to choose m_D in the same range as all other fermions and m_R somewhere within the GUT scale which lets the *active* mass eigenvalue m_1 become low naturally. Note that m_L was set to zero since the corresponding mass term violates weak isospin and is forbidden in standard model. Choosing, however, m_L sufficiently high, one obtains the *type II seesaw*. This formalism can be extended to three active neutrino generations. The mass matrix then has the dimension $N \times N$ with $N = N_{\text{active}} + N_{\text{sterile}}$ and $N_{\text{active}} = 3$.

2.2 Measurement of Neutrino Masses

2.2.1 Motivation

As we have seen, the experimental discovery of neutrino oscillations and their quantum mechanical description force us to assume that neutrinos are massive. The measurement of neutrino oscillations allows us to determine certain oscillation parameters. Eq. (2.4) shows that by combination of different oscillation measurements at different baselines we can obtain the mixing angles and the mass differences. The results from the oscillation experiments by now are shown in table 2.4 and picture 2.2.

⁴Note that we are still concerning only one neutrino generation. The second eigenvalue means rather that the *sterile* state ν_R which does not participate in weak interaction can have an effective mass different from the *active* state ν_L . In that case, oscillation between ν_L and ν_R is possible.

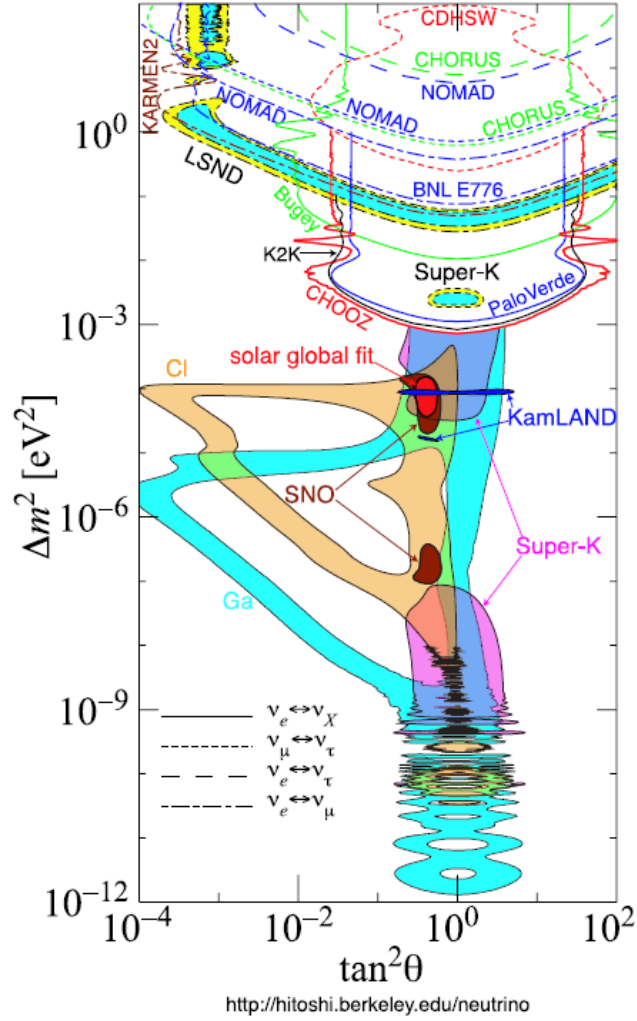


Figure 2.2: Results from the oscillation experiments for atmospheric, solar and reactor neutrinos. [Hito]

However, oscillation experiments do not provide the absolute values of the mass eigenstate and the sign of the differences, i.e. the *hierarchy*. The sign of Δm_{12} is already determined by the *MSW effect* describing neutrino oscillations in matter [Fog06]. The sign of Δm_{23} is still unknown. There are two scenarios possible for that situation which are shown in fig. 2.3. The *normal hierarchy* is described by $m_1 < m_2 < m_3$ while the *inverted hierarchy* means that $m_3 < m_1 < m_2$. However, if the masses are high enough, they are nearly equal which leads to *quasi-degenerate neutrino masses*. These are largely favoured by the type II seesaw mechanism [Ott08].

An absolute measurement of the masses is thus important in order to know which mechanism is responsible for these. A sufficiently precise knowledge of the parameters would

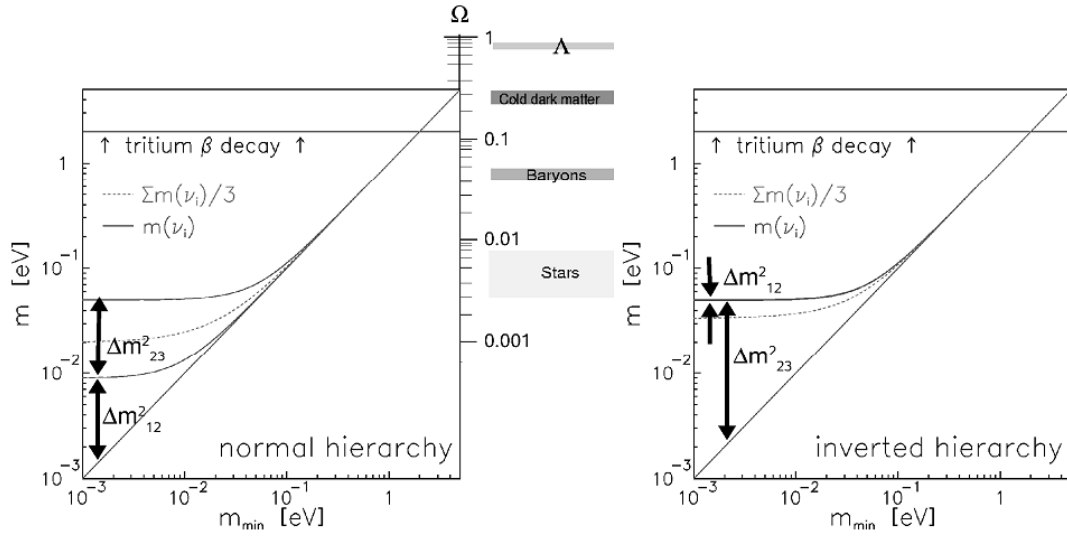


Figure 2.3: Two hierarchy scenarios for the absolute neutrinos masses which are compatible with the current mass differences obtained from oscillation experiments.

enable us to exclude a range of models. If the seesaw mechanism is valid, the masses are directly linked via m_R to the scale of new physics. The absolute value of the neutrino masses is furthermore of significance for cosmological models where neutrinos contribute to the overall energy density.

Δm_{12}^2	Δm_{23}^2	$\sin^2 2\theta_{12}$	$\sin^2 2\theta_{23}$	$\sin^2 2\theta_{13}$
$8.0^{+0.4}_{-0.3} \cdot 10^{-5} \text{ eV}^2$	$[1.9, 3.0] \cdot 10^{-3} \text{ eV}^2$	$0.86^{+0.03}_{-0.04}$	$[0.92, 1]$	0.092 ± 0.017

Figure 2.4: Current measured values for squared mass differences and mixing angles. Values from [Ott08], θ_{13} from [DAY12].

2.2.2 Measurement Methods

There are basically three methods that are able to measure the absolute masses to some extent.

It is possible to estimate an average neutrino mass based on **cosmological models**. As relativistic particles, neutrinos can escape the density fluctuations of cold dark matter and baryons. Therefore, all structures at scales lower than their free streaming length ($\approx \text{Gpc}$) are suppressed depending on the ratio between neutrino energy density Ω_ν and matter energy density Ω_m . The combination of cosmic microwave background (CMB) data and large scale structure (LSS) observations allows evaluation of the neutrino contribution to the energy density which directly depends on known neutrino density in the universe and the average neutrino mass. The most recent calculation was done in

[Ham10] with an 95 % C.L. upper limit of $\sum m_i < 1.18 \text{ eV}$.

In case neutrinos are Majorana particles, it could be possible to determine the mass in the **neutrinoless double beta decay** ($0\nu\beta\beta$). It can be figuratively viewed as two β^- decays in which the two antineutrinos annihilate each other. Besides the Majorana condition that $\nu_e = \bar{\nu}_e$, the neutrino has to be massive therefore, since due to angular momentum conservation the helicity state the neutrino was produced in has to be different from the helicity state in which it is absorbed. Hence the decay amplitude is proportional to a coherent mixture of the neutrino masses which is called *effective Majorana mass*. The Heidelberg-Moscow experiment [Kla04] reported evidence for a Majorana neutrino with effective mass within the range $\langle m \rangle = (0.1 - 0.9) \text{ eV}$ with 3σ which has been received not uncontroversially.

Beta Decay

The most important empirical method is the measurement endpoint region of the nuclear **beta decay spectrum**. The principle in itself is model-independent since it is only based on kinematic arguments. In the β^- decay

$$n \longrightarrow p + e^- + \bar{\nu}_e \quad (2.10)$$

the available decay energy is shared by the electron and the antineutrino, yielding a continuous kinetic energy spectrum of the electron. The maximal kinetic energy by the electron is equal to the decay energy in case of a massless neutrino,

$$E_{\text{max}}(m_\nu = 0) = Q_\beta = m_i - m_f - m_e, \quad (2.11)$$

where m_i and m_f are the masses of the initial and the final state, respectively. The nuclear recoil energy has been neglected. If the neutrino has a mass, the maximal kinetic energy is lowered since a neutrino needs to be produced in any case, resulting in

$$E_{\text{max}}(m_\nu > 0) = Q_\beta - m_\nu = m_i - m_f - m_e - m_\nu. \quad (2.12)$$

Precisely, one needs to insert the lowest neutrino mass state for m_ν which is $m_{\nu 1}$ for a normal hierarchy, since the decay into the neutrino flavour state ν_e can be seen as sum of the decays into the mass states. However, if the mass hierarchy is quasi-degenerate or the experimental resolution is below the mass differences, one can define an effective electron neutrino mass with

$$m_{\nu_e} = \sqrt{\sum_i U_{ei}^2 m_i^2}. \quad (2.13)$$

This may not be interpreted as flavour neutrino states having definite masses but as statistical (non-coherent) expectation value.

If the decay energy of a beta isotope is known, a precise determination of the maximal electron energy allows a measurement of the effective electron neutrino mass⁵. If the decay energy is not exactly known, the mass has to be determined along with the decay energy in an at least 2 parameter-fit of the beta spectrum.

The most important isotope in that context is tritium (^3H or T) with a low Q value that is known with 1 eV precision. The most precise upper limit by now is set by the Mainz experiment [Kra05] which measured $m_{\nu_e} < 2.3\text{eV}$ at 95 % C.L. using an electromagnetic spectrometer (*MAC-E-Filter*, see next chapter). There are also experiments planned which use cryogenic microcalorimeters and Rhenium (^{187}Re) as isotope [MAR06].

⁵which we will refer to simply as *electron neutrino mass* in the following

3 The KATRIN Experiment

We have seen that by now only the mass differences between the neutrino mass states are known but not the absolute values. These are currently only constrained by upper limits, where the most precise one comes from the Mainz experiment. The impressive results from Mainz shall in future be outreached by its successor, the *Karlsruhe Tritium Neutrino experiment* (KATRIN) [KAT04]. This chapter will start with the fundamental ideas and physics that are used in this kind of experiments. Then the KATRIN experiment is described concretely in its specifications and set-up.

3.1 Basics of Tritium Beta Decay Experiments

All tritium beta decay experiments are based on two principles.

1. The endpoint region of the beta decay spectrum of tritium is a function of the electron neutrino mass.
2. The beta spectrum can be determined by choosing threshold energies $E_{\min,i}$ for an electromagnetic filter system (MAC-E-Filter) between source and detector. Measuring the count rate for each filter setting and plotting against the threshold energy gives the integral of the beta spectrum (convoluted with the response function of the experiment).

If the response function is known, the neutrino mass can then be determined by fitting the integrated beta spectrum to the predictions, using at least $m_{\nu_e}^2$ and E_0 as fit parameters. The two main ingredients of the tritium beta experiments shall now be explained.

3.1.1 The Tritium Beta Spectrum

Tritium is an ideal candidate for a neutrino mass measurement since it has a low endpoint of $Q_T \approx 18.6 \text{ keV}$ and a short half-time of $t_{1/2} = 12.3 \text{ a}$. In the KATRIN experiment a gaseous source is used where the tritium exists in molecular form T_2 . The kinetic energy spectrum of the beta electron (*beta spectrum*) is then given as differential decay rate by

$$\frac{dR}{dE} = N \frac{G_f^2}{2\pi^3 \hbar^7 c^5} \cos^2(\theta_C) |M|^2 F(E, Z+1) \cdot p(E) \cdot (E + m_e c^2) \cdot \sum_i P_i \cdot (E_0 - V_i - E) \cdot \sqrt{(E_0 - V_i - E)^2 - m_{\nu_e}^2 c^4}. \quad (3.1)$$

The components can be itemized as follows: N is the number of mother nuclei, G_f the Fermi coupling constant, θ_C the Cabbibo angle and M the nuclear decay matrix element. $F(E, Z+1)$ denotes the Fermi function, depending on the emission energy and the charge of daughter nucleus. E_0 corresponds to the maximal kinetic energy of the electrons in case of $m_{\nu_e} = 0$ and is called simply the *beta endpoint* in the following. It deviates from the Q value of atomic tritium (2.11) by the recoil energy of the daughter molecule and the difference between molecular binding energy of the mother and daughter molecule. The best measurement based on mass differences yields $E_0 = (18574.3 \pm 1.7) \text{ eV}$ by now [Van93] [Wei93a].

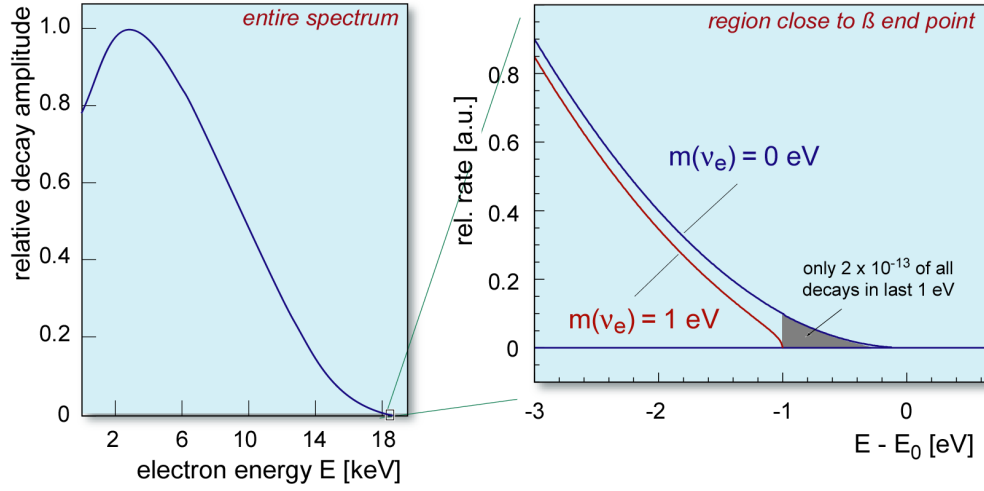


Figure 3.1: Complete differential beta spectrum of tritium and endpoint region. [KAT04]

The daughter molecules or atoms do not end up in the electronic or rotational-vibrational ground state in every case, instead, there is a distribution of different final states, which have been calculated in [Sae20]. P_i is the probability to decay into a certain final state and V_i the corresponding excitation energy. The effect of that is that the beta endpoint E_0 , which is the Q value minus the recoil energy of the daughter nucleus and corresponds to the highest possible kinetic energy of the electron if there was no neutrino mass, is shifted by V_i . The resulting spectrum is then a superposition of the spectra for each final state, weighted with the corresponding probability P_i . For the Fermi function one can use the approximation

$$F(E, Z+1) = \frac{2\pi \cdot x}{1 - e^{-2\pi \cdot x}} \cdot (1.002037 - 0.001427 \cdot v_e/c), \quad (3.2)$$

where x is a function of the electron velocity v_e ,

$$x = \frac{2\alpha_f c}{v_e} \quad (3.3)$$

and α_f the fine structure constant [Sim81]. However, near the endpoint the Fermi function is nearly constant with $F \approx 1.187$.

It can be seen (fig. 3.1) that the endpoint region which is sensitive to the neutrino mass is only a tiny spot of the whole spectrum. The last 1 eV of the spectrum produces only about $2 \cdot 10^{-13}$ of the decay events. Therefore, the need for very precise measurement technology arises.

3.1.2 The MAC-E-Filter

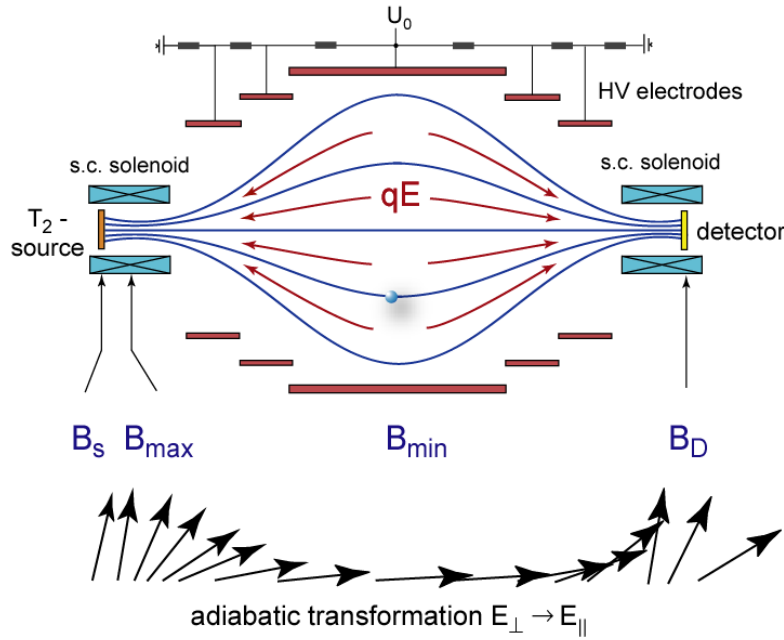


Figure 3.2: Principle of the MAC-E-Filter. The transverse momentum is transformed adiabatically into longitudinal momentum. The electron energy is then analysed by an electrostatic filter. [KAT04]

As mentioned, a so called *MAC-E-Filter*¹ is used in tritium beta decay experiments which is meant to work as high pass filter that lets only pass electrons with a kinetic energy above a threshold E_{\min} . In case of a sharp electron beam this could simply be accomplished by an electrostatic potential with a sufficient retarding potential relative to the source

$$qU_{\text{ret}} = E_{\min}, \quad (3.4)$$

where q is the electron charge. However, the decay in the source is isotropic, so an additional magnetic field is applied in order to align the electron momenta to a pre-

¹Magnetic Adiabatic Collimation with an Electrostatic Filter

ferred direction. The principle is shown in fig. 3.2. A homogeneous magnetic field is applied at the source by a superconducting coil which lets the decay electrons perform cyclotron movements around the field lines. The field strength B decreases with growing distance from the source. If the field changes are sufficiently slow the magnetic moment is conserved,

$$\mu = \frac{p_{\perp}^2}{B} = \text{const.}, \quad (3.5)$$

which leads to a transformation of the momentum component transverse to the field lines, p_{\perp} , into parallel momentum p_{\parallel} . This is called *adiabatic motion* or *guided center motion*. If another coil is applied at the detector the field strength has a minimum at the center of the experiment (referred to as *analysing plane*). If the field strength at the analysing plane B_A is sufficiently smaller than at the source B_s , the electron momenta in the analysis plane are nearly polarized with the beam axis. If now the electrostatic potential U_{ret} is applied in the analysis plane, the MAC-E filter works as a sharp high pass filter. The relative energy resolution can be derived as

$$\frac{\Delta E}{E} = \frac{B_A}{B_{\text{max}}}, \quad (3.6)$$

where B_{max} is the maximal field strength between source and analysing plane. Generally, $B_{\text{max}} \neq B_s$. Instead, a further *pinch magnet* with $B_{\text{pinch}} = B_{\text{max}} > B_s$ is placed between source and analysing plane in order to suppress electrons with large starting angles, which are likely to be scattered in the source and distort the spectrum. This set-up is shown in fig. 3.3. The maximum emission angle is then

$$\sin \theta_{\text{max}} = \sqrt{\frac{B_s}{B_{\text{max}}}}. \quad (3.7)$$

For an isotropic source, the *transmission function* which is the ratio of transmitted electrons as a function of the kinetic starting energy and the retarding potential (see fig. 3.4) can be calculated as

$$T(E, qU_{\text{ret}}) = \begin{cases} 0 & E - qU_{\text{ret}} < 0 \\ \frac{1 - \sqrt{1 - \frac{E - qU_{\text{ret}}}{E} \frac{B_s}{B_A}}}{1 - \sqrt{1 - \frac{\Delta E}{E} \frac{B_s}{B_A}}} & 0 \leq E - qU_{\text{ret}} \leq \Delta E \\ 1 & E - qU_{\text{ret}} > \Delta E \end{cases} \quad (3.8)$$

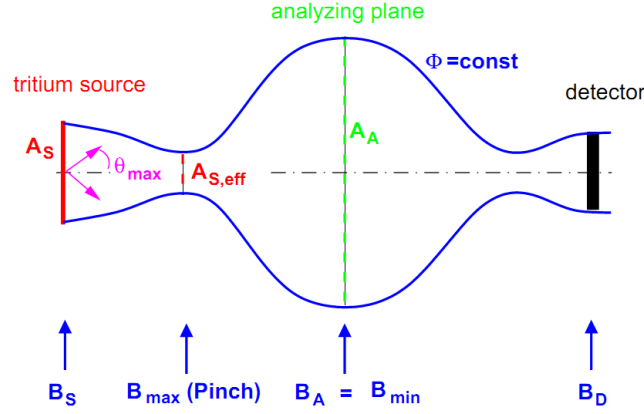


Figure 3.3: Electromagnetic set-up of the MAC-E-Filter principle used in the KATRIN experiment. The pinch magnet with B_{\max} limits the starting angle of the electrons. [KAT04]

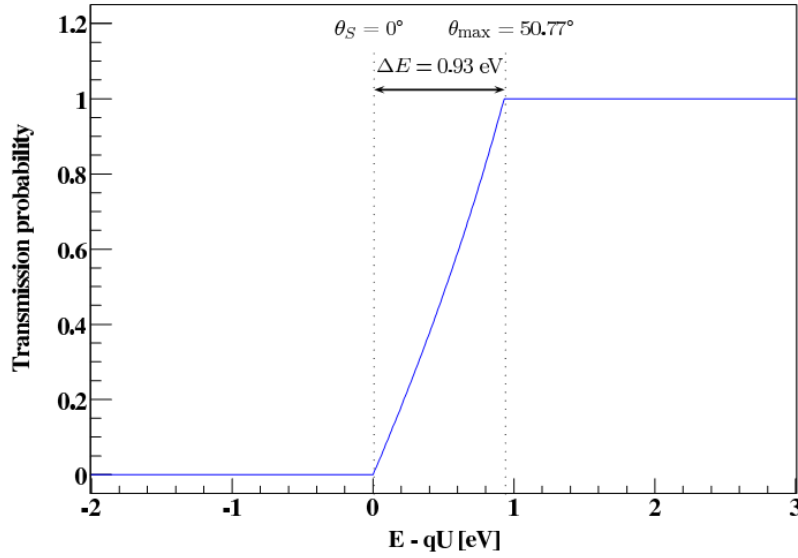


Figure 3.4: Transmission function of the KATRIN main MAC-E-Filter. The transmission probability is 0 below the potential barrier and reaches 1 at $E - qU_{\text{ret}} = \Delta E$. [Wol08]

3.2 Design of KATRIN

3.2.1 Design Goals and Parameters

The goal of KATRIN is to reach a sensitivity on m_{ν_e} of 0.2 eV at 90 % C.L. for three years net measurement time. This is about one order of magnitude more precise than the Mainz experiment. It touches already the regime where hierarchical and degenerate mass scenarios can be distinguished. In order to reach such precision, a high technological level is necessary.

The MAC-E-Filter is ought to work with a maximal magnetic field of $B_{\max} = 6$ T which is roughly the limit of technical managability. The source field was chosen to $B_S = 3.6$ T, resulting in a maximum starting angle of $\theta_{\max} = 50.77^\circ$. In order to reach an optimal resolution (3.6), the magnetic field in the analysis plane has to be set as low as possible. However, from (3.5) follows that the lower the magnetic field is, the greater the flux tube radius of the electrons will become. Therefore, the field is bounded below by the radius of the main spectrometer in the analysis plane. In KATRIN, a main spectrometer with a diameter of $D_A = 9$ m is used, allowing a minimal field of $B_A = 3$ G. According to (3.6) this corresponds to a filter resolution of $\Delta E < 1$ eV near the beta endpoint.

To account for sufficient statistics, a high tritium decay rate of $1.8 \cdot 10^{11}$ Bq is required. Since the event rate very close to the endpoint is low anyway (~ 0.2 Hz within the last 10 eV of the spectrum), KATRIN needs to be optimized for an ultra low background of < 10 mHz. Therefore, amongst others, an ultra high vacuum (UHV) of $< 10^{-11}$ mbar inside the components is necessary.

The set-up of KATRIN is shown in fig. 3.5 . In the following, the main components of KATRIN shall be explained.

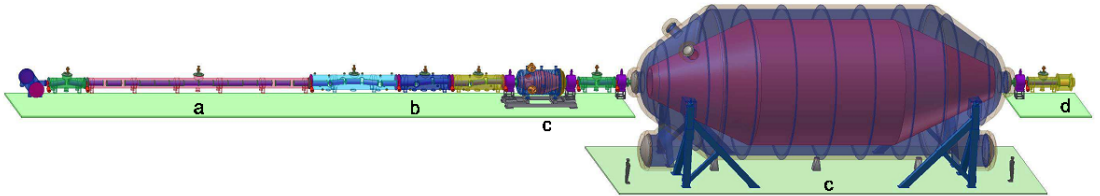


Figure 3.5: Set-up of KATRIN with main components. a) WGTS, b) transport and pumping section c) pre-spectrometer and main spectrometer, d) detector. [KAT04]

3.2.2 Main Components

WGTS and Transport Section

The beta decays take place in the *Windowsless Gaseous Tritium Source*. The desired tritium activity of $A_T = 1.8 \cdot 10^{11}$ Bq (both hemispheres) is provided by a tritium column density of $\rho d = 5 \cdot 10^{17}$ molecules/cm² within a diameter of $D_S = 90$ mm and with a gas purity of $\epsilon_A = 0.95$ at a temperature of 27 K. The tritium is injected in the middle of a 10 m long tube and exhausted at the ends with turbomolecular pumps (DPS1-R and DPS2-F) where it is repurified and returned to the cycle.

Since the knowledge of the column density is one of the main contributions to the systematic uncertainty of KATRIN, it is important to hold the pressure and other parameters in the source constant and monitor them precisely.

To prevent residual tritium molecules from getting into the spectrometers and causing background, the remaining tritium has to be removed. The transport section consists of two units, namely a differential pumping section (DPS2-F) and a cryogenic pumping section (CPS). The first is performed by a further turbomolecular pump while the latter uses cryogenic trapping on the surface. In both sections, the electron flux is guided with bending magnets on a slightly curved line while the neutral tritium molecules keep largely a constant moving direction.

Spectrometers

Before the electrons enter the main spectrometer they have to pass through a pre-spectrometer which makes use of the MAC-E-Filter principle. The pre-spectrometer operates with a retarding potential in the range of -18.3 keV and with an energy resolution of about 75 eV. This procedure blocks all electrons except those within the last ~ 300 eV of the spectrum. Since the electrons below that threshold carry no substantial neutrino mass information, this is an effective means of reducing the electron flux, thereby minimizing background from inelastic scattering on gas molecules.

The main spectrometer thereupon is the central component which carries out the physically significant analysis of the electron energies by the MAC-E-Filter principle. It is a tank with a length of 23.3 m and a diameter of about 10 m in which the electron energies are analyzed. Two superconducting coils at the entrance and the exit with 4.5 T each and an air coil around the tank for fine-tuning provide an inhomogeneous magnetic field. It is minimal in the center of the analysis plane, with a strength of 3 G, thereby transforming the transverse momentum of the electrons into longitudinal momentum. The vessel hull is held on the negative HV retarding potential U_{ret} which blocks all electrons that do not have the sufficient kinetic energy in the analysis plane.

Since the retarding potentials form the abscissa of the data points in the measured integrated beta spectrum, it is extremely important that the HV is held constant and permanently monitored. This will be on one hand done by direct measurements using

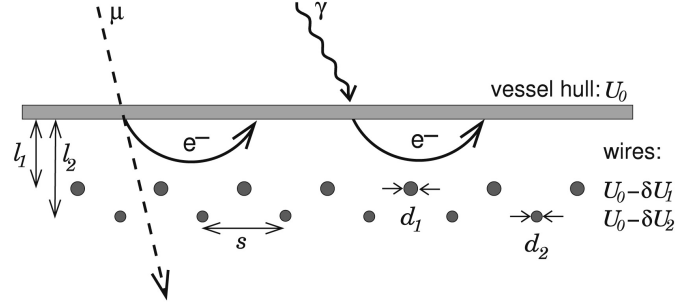


Figure 3.6: Principle of the main spectrometer wire electrode. Background electrons from the vessel hull are reflected by a slightly more negative potential. [Val10]

a high precision voltage divider and on the other hand by an additional monitor spectrometer. This one calibrates the HV by measuring a monoenergetic electron source like ^{83}Kr using the same voltage as the main spectrometer on the vessel hull while varying the source potential. The existing main spectrometer from the Mainz experiment is used for that means.

A complex wire electrode system inside the main spectrometer is used for reduction of background caused by charged particles from collisions of cosmic radiation with the vessel hull or intrinsic radioactivity. These particles are reflected by the wires which lie closely to the vessel hull and have a slightly more negative potential than the vessel hull. The principle is shown in 3.6. To reduce secondary background from the wire electrode itself it has shown to be most efficient if two wire layers are used, an inner one with thick wires and an outer one with thin wires and a more negative potential.

Detector

At the end of the beam line, a silicium detector is used to measure the number of transmitted electrons. It consists of 148 pixels in a 'dart board' like geometry (fig. 3.7). Since the electric potential and the magnetic field are not perfectly constant in the analysis plane it is important to know which way an electron has taken through the main spectrometer. Effectively, this is accomplished with such a position-sensitive detector where each pixel corresponds to a different transmission function (3.8). For further background reduction, the detector will have an energy resolution of < 1 keV, a post acceleration of 30 kV and a muon veto. An efficiency of $> 90\%$ is proposed as well as a time resolution of < 50 ns for TOF measurements.

3.3 Data Analysis of KATRIN

As already mentioned in the beginning, the bottom line of the data measured by KATRIN is the countrate $R_T(U_i)$ for each applied retarding potential U_i . This is the

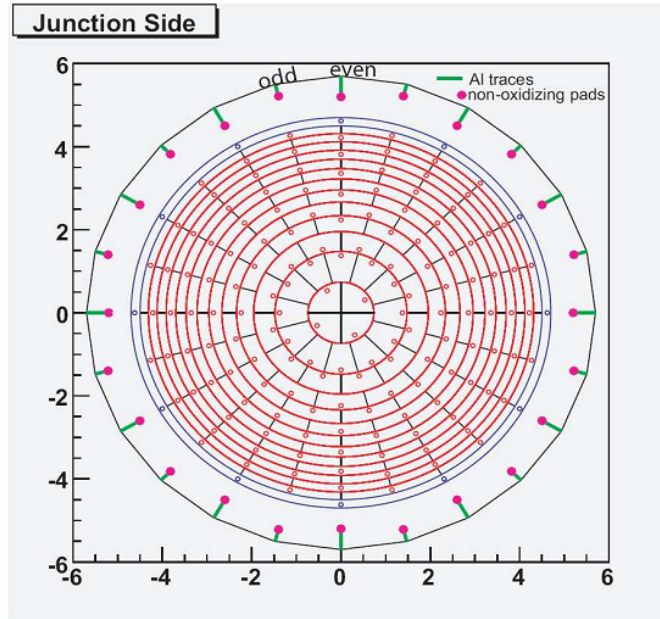


Figure 3.7: Layout of the KATRIN detector. [Ste07]

resulting fit function from which the neutrino mass is extracted by fitting the predicted expectation values depending on m_{ν_e} to the measured data. The function is a convolution of the differential beta spectrum with the response function of the experiment plus a constant background b ,

$$R_T(U) = \left(\int_U^{E_0 - m_{\nu_e} c^2} \frac{dR}{dE}(E) T'(E, U) dE \right) + b, \quad (3.9)$$

where the response function is given by

$$T'(E, U) := T(E, U) \otimes f_{loss}(\Delta E) \quad (3.10)$$

as the convolution of the transmission function (3.8) with the energy loss function (fig. 3.8). The latter represents the probability that an electron loses energy between ΔE and $\Delta E + \delta E$ in inelastic collisions within the source.² By a χ^2 minimization on (3.9), using as parameters

- the squared electron neutrino mass $m_{\nu_e}^2$,
- the beta endpoint E_0 ,
- the signal amplitude S and

²This function is discussed in detail in the next chapter as it contributes to the TOF spectrum.

- the background rate b ,

the neutrino mass can be determined if it is above the 90 % C.L. sensitivity lower limit of 0.2 eV.

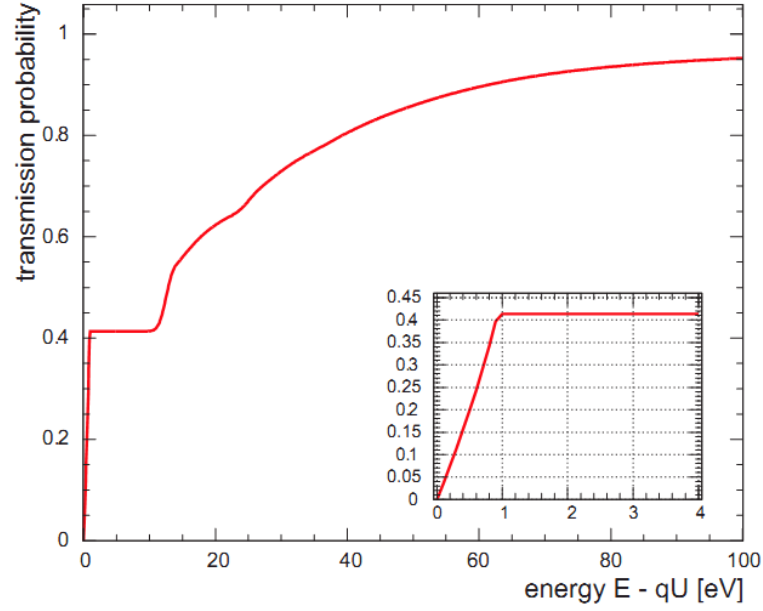


Figure 3.8: Response function of KATRIN as convolution of the transmission function and the energy loss function. [KAT04]

4 Simulation of a Time-Of-Flight Mode for KATRIN

4.1 Idea and Motivation

It was shown that in the KATRIN experiment the electron neutrino mass is to be determined by measuring the integrated beta spectrum. The count rate is scanned as a function of the retarding energy below which all electrons are blocked by the MAC-E-Filter. That function depends explicitly on $m_{\nu_e}^2$ and can finally be fitted.

An alternative idea is to use *time of flight (TOF) spectroscopy* to measure the neutrino mass. The TOF of a beta decay electron through the main spectrometer of KATRIN is a function of the kinetic energy and the emission angle. Since the MAC-E-filter slows down the electrons near the retarding energy, the TOF measurement will be very sensitive on subtle energy differences near that point. By measuring the TOF distribution of the electrons which we call the *TOF spectrum* one can reconstruct the parameters determining the beta spectrum, including $m_{\nu_e}^2$.

While the standard method measures just a number for each retarding energy (the count rate), the TOF spectroscopy mode obtains thus a two-dimensional dataset. This gain of information allows useful optimizations. In principle it would be sufficient to measure only at a single retarding energy near the endpoint, though a mixture of retarding energies might be more sensitive. Since the systematic uncertainty on $m_{\nu_e}^2$ grows as the retarding energy is decreased, the goal is to minimize the amount of measured data contributing from energies far from the beta endpoint. The TOF method could in theory provide this and concentrate the measurement on a few retarding energies near the endpoint, each of them delivering a full TOF spectrum from which $m_{\nu_e}^2$ can be disentangled.

Motivated by that, simulations were performed in order to determine whether this procedure would reduce the statistical uncertainty on $m_{\nu_e}^2$, particularly when measuring only near the endpoint to minimize the systematic uncertainty. This was achieved by the following steps.

1. Simulating a model of the TOF spectrum of the KATRIN events for certain parameters.
2. Simulating virtual datasets of TOF spectra as they would be measured in a real experiment assuming a certain neutrino mass and beta endpoint.

3. Fitting the neutrino mass square $m_{\nu_e}^2$ that was put into (2) by using the models from (1).
4. Determining the statistical uncertainty on $m_{\nu_e}^2$ that would result by this method, particularly for different combination of retarding energies and measurement time distributions inside them.

4.2 Mathematical description

To state the TOF spectrum as a function of certain parameters we need to develop a mathematical model of it within the context of the KATRIN main spectrometer geometry. The important parameters determining the TOF spectrum for a fit are the neutrino mass square $m_{\nu_e}^2$, the beta endpoint E_0 and the retarding potential U_{ret} of the MAC-E-Filter. A suitable model should be an explicit function of these. The first step towards a mathematical expression of the TOF spectrum is to derive the TOF for a single event.

4.2.1 TOF for a Single Electron

Within the main spectrometer the principle of adiabatic motion where the magnetic moment is held constant (3.5) is valid in good approximation. Using a simplified geometry, we take only the field on the z axis into account and neglect drifts. The B field is then only a function of the z coordinate. Then, the transverse momentum of an electron can be derived as a function of its z position.

$$p_{\perp}^2(z) = p(0)^2 \cdot \sin^2 \theta(0) \cdot \frac{B(z)}{B(0)}, \quad (4.1)$$

where $p(0)$, $\theta(0)$ and $B(0)$ are the total momentum, emission angle and total magnetic field at the source $z = 0$. The fraction reflects the role of the adiabatic magnetic field geometry of the MAC-E-filter: As the field $B(z)$ grows, the transverse momentum is converted continuously into longitudinal momentum. The relativistic energy of the electron is given by the energy-momentum-relation

$$E_{\text{rel}}^2(z) = p_{\parallel}^2(z)c^2 + p_{\perp}^2(z)c^2 + m_e^2c^4. \quad (4.2)$$

Since the total energy $E_{\text{tot}} = E_{\text{rel}} + E_{\text{pot}} = E_{\text{kin}} + m_e c^2 + E_{\text{pot}}$ is always conserved, we can state the relativistic energy as a function of z .

$$E_{\text{rel}}(z) = E_{\text{rel}}(0) - E_{\text{pot}}(z) + E_{\text{pot}}(0) = E_{\text{kin}}(0) + m_e c^2 - q \cdot \Delta U(z), \quad (4.3)$$

where $\Delta U(z)$ is the difference of the retarding voltage at source and at z and q the electron charge.¹ Note that this equation is only valid for those z positions that can be

¹By convention, $q \cdot \Delta U$ is positive.

reached by the electron. If there is a $\Delta U(z_{\max})$ with $q\Delta U(z_{\max}) = E_{\text{kin}}(0)$, then the electron cannot reach all positions $> z_{\max}$ and thus, (4.3) is not defined at those positions. Combining eqs. (4.1), (4.2) and (4.3), we derive an expression for the longitudinal momentum as a function of z only in terms of the fields B and ΔU :

$$\begin{aligned} p_{\parallel}^2(z)c^2 &= (E_{\text{kin}}^2(0) + 2E_{\text{kin}}(0) m_e c^2) \left(1 - \sin^2 \theta(0) \cdot \frac{B(z)}{B(0)}\right) \\ &+ q^2 \Delta U^2 - 2q\Delta U (E_{\text{kin}}(0) + m_e c^2). \end{aligned} \quad (4.4)$$

The TOF is now determined by integrating the reciprocal parallel velocity $1/v_{\parallel} = \gamma m/p_{\parallel} = E_{\text{rel}}/p_{\parallel} c^2$ over the measurement path.

$$\tau(E_{\text{kin}}(0), \theta(0)) = \int dz \frac{1}{v_{\parallel}} = \int_{z_{\text{start}}}^{z_{\text{detector}}} dz \frac{E_{\text{kin}}(0) + m_e c^2 - q\Delta U}{\sqrt{p_{\parallel}^2(z)c^2} \cdot c}. \quad (4.5)$$

The lower bound z_{start} of the integration interval depends on where the start signal time is measured. For any measurement method (see next chapter) it is recommended to perform the measurement in the entrance of the main spectrometer. Significant TOF differences between electrons with different energies occur due to the retardation in the main spectrometer, as stated in the last chapter. Thus, the path of the electrons through the main spectrometer is the important one in terms of neutrino mass measurement.

The integral (4.5) is somewhat simplified since it does not integrate over the full path of the center motion of the electron (which is congruent with the field lines) but only along the z axis. It is important to understand what this simplification implies. In the real KATRIN experiment there is a difference of ≈ 1.1 eV between the potential at the center and at the border of the analysing plane [Zac09]. Thus, each radial path through the main spectrometer has its own transmission function. The KATRIN detector is a multi-pixel detector to determine which radial path the electron has taken and thus to apply the correct transmission function for the analysis. To integrate only on the z axis means that we pretend all detector events to occur in a central pixel. That would be incorrect for a TOF fit of real data, where we would need to simulate the TOF spectrum for each detector pixel. But it is sufficient for a principle study if the statistics do not differ between central and border detector pixels.

We end up with an equation stating the TOF for a single electron as a function of its kinetic starting energy $E_{\text{kin}}(0)$, its starting angle $\theta(0)$, given a certain electric potential $\Delta U(z)$ and magnetic field $B(z)$ on the z axis. As the adiabatic approximation (3.5) is valid through the whole transport section, the coordinate of the tagger is arbitrary. The emission angle $\theta(0)$ is automatically transformed to its correct value at the tagger position $\theta_{z=z_{\text{tagger}}}$ by (4.4) because only the ratio of local and source magnetic field $B(z)/B(0)$ matters but not the field evolution between $B(0)$ and $B(z_{\text{tagger}})$.

4.2.2 TOF Spectrum

Equation (4.5) presumes a fixed kinetic starting energy $\epsilon := E_{\text{kin}}(0)$ and starting angle $\theta := \theta(0)$ as arguments. However, these parameters are statistically distributed within KATRIN. Moreover, the distribution of the kinetic starting energies corresponds in first order to the beta spectrum which contains the information about m_{ν_e} . As mentioned in the beginning, the distribution of flight times is therefore also a function of m_{ν_e} . It will be the key quantity which is measured in a TOF spectroscopy set-up. In order to be capable numerically, the TOF spectrum will be determined as a *binned* spectrum where the differential TOF spectrum $\frac{dR}{dt}$ is discretized into bins of constant length Δt and integrated over each bin j .

$$F(t_j) := \int_{t_j}^{t_j + \Delta t} dt \frac{dR}{dt}(t) \quad (4.6)$$

According to (4.5), the TOF is fully determined by the kinetic starting energy and starting angle. The number of counts in each bin is therefore equivalent to the number of events with energies ϵ and angles θ which lead to a TOF within that bin:

$$\begin{aligned} F(t_j) &= \iint_{\{(\epsilon, \theta) | t_j \leq \tau(\epsilon, \theta) \leq t_j + \Delta t\}} d\theta d\epsilon \frac{d^2 R}{d\theta d\epsilon} \\ &= \int_0^{\theta_{\max}} d\theta \int_{\epsilon(t_j, \theta)}^{\epsilon(t_j + \Delta t, \theta)} d\epsilon g(\theta) \left(\frac{dR}{d\epsilon} \right)_\theta \\ &:= \int_0^{\theta_{\max}} d\theta g(\theta) \int_{\epsilon_j(\theta)}^{\epsilon_{j+1}(\theta)} d\epsilon \left(\frac{dR}{d\epsilon} \right)_\theta \end{aligned} \quad (4.7)$$

[Cow98]. Here, $g(\theta)$ refers to the probability distribution of the starting angles and $\left(\frac{dR}{d\epsilon} \right)_\theta$ to the differential decay rate as a function of the kinetic energy ϵ , given a certain θ . By the separation in the last step, the rate in each bin becomes an integral over a specific interval of the kinetic energy spectrum, weighted with the angle distribution. The correct intervals are given by the relation between energy, angle and TOF (4.5). As the emission in the source is isotropic, a sine law applies for the angle distribution:

$$g(\theta) = c_\theta \cdot \sin \theta. \quad (4.8)$$

The normalization constant c_θ is given by the condition that the integral over the distribution for all allowed angles must be 1. Since (3.7) restricts the angles due to the field design to $\approx 50.77^\circ$, it gives

$$c_\theta^{-1} = \int_0^{\theta_{\max}} d\theta \, g(\theta) = 1 - \cos \theta_{\max} . \quad (4.9)$$

The spectrum of kinetic energies $(\frac{dR}{d\epsilon})_\theta$ is at first order given by the beta decay spectrum $\frac{dR}{dE}$ in (3.1) which is angle-independent. However, some electrons lose energy in inelastic scattering processes with the daughter molecules. Furthermore, the raw decay rate $\frac{dR}{dE}$ decreases due to several losses inside the experiment. The latter one can be represented by a further normalization constant c_ϵ while the first one is taken into account by convolving the beta spectrum with the total energy loss function, which gives

$$\begin{aligned} \left(\frac{dR}{d\epsilon}\right)_\theta &= c_\epsilon \cdot \left(\frac{dR}{dE} \otimes f_{\text{loss},\theta}\right) \\ &= c_\epsilon \cdot \left(p_0(\theta) \cdot \frac{dR}{dE} \Big|_{E=\epsilon} + \sum_{n=1}^{\infty} p_n(\theta) \cdot \frac{dR}{dE} \otimes f_n\right) \end{aligned} \quad (4.10)$$

where the f_n is the energy loss function of scattering order n which is defined recursively through the single scattering energy loss function f_1 as

$$f_n = f_{n-1} \otimes f_1 \quad \text{for } n > 1 . \quad (4.11)$$

The function $f_1(\Delta E)$ is the probability density of losing an energy ΔE at a singular scattering. The functions of f_n can then correspondingly be interpreted as the same for n -fold scattering. In this equation, all changes of the angle of the electron during scattering are neglected. p_n is the probability that an electron is scattered n times. If we neglect again changes of the angle, it is a function of the emission angle θ and given by a Poisson law

$$p_n = \frac{\lambda^n(\theta)}{n!} e^{-\lambda(\theta)}, \quad (4.12)$$

where expectation value λ is given in terms of the column density ρd and the mean free column density ρd_{free} by

$$\lambda^n(\theta) = \frac{\rho d}{\rho d_{\text{free}} \cdot \cos \theta} = \frac{\rho d \cdot \sigma_{\text{scat}}}{\cos \theta}. \quad (4.13)$$

The normalization constant in (4.10) is given as a product of several loss factors if $\frac{dR}{dE}$ in (3.1) is normalized correctly to the average number of tritium molecules in the source.

$$c_\epsilon = \frac{\Delta\omega}{2\pi} \cdot \frac{1}{2} \cdot \epsilon_{\text{pipe}} \cdot \epsilon_{\text{det}}. \quad (4.14)$$

Here, $\Delta\omega/2\pi = 1 - \cos\theta$ is the accepted forward hemisphere solid angle. A reduction factor of 2 applies since $\frac{dR}{dE}$ refers to the decay rate into both hemispheres while in KATRIN only the forward hemisphere contributes. ϵ_{pipe} is the reduction of the signal rate due to inelastic scattering with the flux tube and ϵ_{det} the detector efficiency.

The normalization constant can also be expressed in terms of the signal rate S which is given by the tritium activity A_T as

$$S = c_\epsilon \cdot A_T = c_\epsilon \cdot \int_0^{E_0 - m_{\nu_e} c^2} dE \frac{dR}{dE}. \quad (4.15)$$

With the design goal of $A_{T2} \approx 1.8 \cdot 10^{11}$ Bq, $\epsilon_{\text{det}} \approx 0.9$ and assuming $\epsilon_{\text{pipe}} \approx 0.8$, this gives $S \approx 2.38 \cdot 10^{10}$ Hz. The relation above is helpful if one does not use the normalized beta spectrum (4.10) but a proportional expression $\frac{dR^*}{dE}$. As one can easily derive, the normalization constant c_ϵ^* to get the correct kinetic energy spectrum is then

$$c_\epsilon^* = S \cdot \left(\int_0^{E_0 - m_{\nu_e} c^2} dE \frac{dR^*}{dE} \right)^{-1}. \quad (4.16)$$

4.3 Numerical Realization

4.3.1 Numerical Calculation of the TOF Spectrum

As mentioned, a first goal towards the aim of determining the statistical uncertainty on $m_{\nu_e}^2$ with the TOF method, is to be able to calculate the TOF spectrum for each possible and reasonable set of $m_{\nu_e}^2$, E_0 and U_{ret} . The integral over the angle distribution and kinetic energy spectrum in (4.18) has to be solved, somehow. Since $\tau(E, \theta)$ is not invertible analytically, the integral must be solved numerically.

To do so, the integral over the angle distribution can be turned into a sum. Therefore, equidistant points θ_i with $\theta_{i+1} = \theta_i + \Delta\theta$ are chosen. The differential weights $g(\theta)$ then have to be replaced by integral weights A_i , given by

$$A_i = \int_{\theta_i}^{\theta_{i+1}} d\theta g(\theta). \quad (4.17)$$

The binned TOF spectrum (4.7) becomes then

$$F(t_j) \approx \sum_i A_i \int_{\epsilon_j(\theta_i)}^{\epsilon_{j+1}(\theta_i)} d\epsilon \left(\frac{dR}{d\epsilon} \right)_{\theta_i}. \quad (4.18)$$

For $\Delta\theta \rightarrow 0$, the expression is exact. Even for larger discretizations, the result is quite accurate since, unlike in a Riemann sum, the weights $A(\theta_i)$ are exact integrals instead of squares. The remaining inaccuracy comes mainly from the energy integral bounds, given as a function of θ_i , but it shows that due to the smooth nature of the TOF spectra, a superposition of a few single-angle TOF spectra is sufficient. In practice, values of $\Delta\theta \sim 1^\circ$ showed accurate results.

The fact that TOF spectrum (4.18) is binned is quite useful numerically. If we have only a limited number of TOF bins, then we can evaluate the integral bounds $\epsilon_j(\theta_i)$ numerically within reasonable performance. Since the angles are discretized now, we are only left with a finite number of tuples (t_j, θ_i) for which the limits $\epsilon_j(\theta_i)$, given according to (4.5) by $t_j = \tau(\epsilon_j, \theta)$, have to be determined.

The evaluation of the bounds has been performed with the Van Wijngaarden-Dekker-Brent method and the integral in (4.5) was solved numerically by Romberg's method [Pre02]. The one-dimensional field maps $\Delta U(z)$ and $B(z)$ in (4.5) have been obtained by the KATRIN simulation tools *magfield* [Glu06] and *elcd3_2* [Glu04] using a modestly simplified geometry that contains the most important coils and electrodes in the main spectrometer. For additional performance, the integral in (4.18) has been further separated into the energy loss orders (4.10). The integrals for all orders $n > 0$ have been calculated by taking the Riemann sum over a pre-calculated table which contains the results of the convolutions in (4.10).

If the spectrum is sufficiently smooth and the bin sizes are small, one can reconstruct the differential TOF spectrum (4.7) from the binned one (4.18) by the approximation

$$\left. \frac{dR}{dt} \right|_{t_j} \approx F(t_j) \cdot \Delta t \quad (4.19)$$

for the sampling points t_j and by interpolation for the other ones. However, for neutrino mass fits one should use (4.18).

4.3.2 Fit Preparations

As mentioned above, the task is to use the modelled TOF spectrum, obtained by (4.18), to perform a fit on some simulated test data to determine the statistical uncertainty $\sigma(m_{\nu_e}^2)$. Several steps need to be performed in preparation for this purpose.

Pre-computation of models and interpolation

Though the routine to solve the integrals and calculate the TOF spectrum was optimized in the course of the development it is anyhow too slow to allow to be calculated in real time during a fit. A solution is to pre-compute a matrix of models with different parameters $((m_{\nu_e}^2)_k, (E_0)_l)$ for each retarding energy $(U_{\text{ret}})_m$ that shall be taken into account. By means of a 2-dimensional interpolation routine we can then obtain the spectrum continuously for each possible combination $(m_{\nu_e}^2, E_0)$. To increase accuracy, it makes sense to choose the sampling points near the region of interest. The interpolation was performed in this simulation by a bi-cubic natural spline [Pre02]. Even the possibility of fitting without interpolation was implemented but since it takes too long it was mainly used to check if the results from the interpolation are correct.

Creation of data

In order to perform a fit, somehow data have to be simulated by a Monte Carlo method, based on the core parameters $(m_{\nu_e}^2, E_0, U_{\text{ret}})$ as well as additional parameters like a constant background. That implies that for each fit a certain neutrino mass and beta endpoint are arbitrarily assumed which are later reproduced by the fit. The datasets should be distributed realistically as they would be measured in an experiment and converge against certain expectation values when averaging a high number of them. A simple approach of dataset creation is to roll the number of counts in each bin as a Poisson distributed random number, where the expectation values μ_j for each bin are taken from the theoretical binned count rate (4.18) as

$$\mu_j = \left(\frac{S}{S_0} \cdot F(t_j) + b \cdot \frac{\Delta t}{t_{\text{max}}} \right) \cdot t_{\text{meas}}, \quad (4.20)$$

where t_{meas} is the supposed measurement time and b a constant background rate. It is assumed that the total background rate b_{raw} stretches equally over the complete spectrum with a maximal TOF of t_{max} . S/S_0 is the relative signal amplitude, expressed in terms of the signal rate (4.15). Thereby, S_0 is the estimated signal rate, as defined in (4.15), while S is an arbitrary signal rate. The reason for this rescaling factor is that S_0 is not known exactly in the experiment, so that S needs to be added as free fit parameter.

Eq. (4.20) is also the fit function. This method has the disadvantage that it is self-consistent, i.e. if the model is calculated wrong then the data are wrong as well. Therefore, an additional Monte Carlo method where the TOF is rolled for each single event based on the underlying distributions in (4.5) has been implemented. That method is too slow to allow a large number of fits but is useful to check consistency.

4.3.3 Fit routine

As soon as the models can be accessed and reasonable datasets are there, the fits can be performed. As in the standard mode of operation, there are again four fit parameters:

1. the neutrino mass square $m_{\nu_e}^2$,
2. the beta endpoint E_0 ,
3. the signal rate S and
4. the background rate b .

As mentioned, all these parameters are 'known' in that way that arbitrarily certain values for the creation of the Monte Carlo data are set. The aim is, thereby, to know how accurate a fit reproduces these values, measured by the statistical parameter uncertainty. The signal rate as relative amplitude and the background have to be fitted, too, since they are not sharply defined.

The retarding potential U_{ret} is not a fit parameter since it can be controlled very precisely in the experiment. Anyhow, it is an important quantity that influences the shape of the spectrum and thus the sensitivity. It is expected that spectra at high retarding potentials near the endpoint are very sensitive to the neutrino mass but hard to be decorrelated from the endpoint. Since spectra at lower retarding potentials are still sensitive to the endpoint, it is likely that measurements at different retarding energies have to be combined. To account for this, the possibility of fitting datasets from different retarding potentials simultaneously has been included.

Two minimization methods for the fit have been implemented. The first one is a modified chi square, given by

$$\chi^2 = \sum_j \frac{(\mu_j - n_j)^2}{n_j}. \quad (4.21)$$

Actually, in some bins within the TOF spectrum the prediction of the model is $\mu_j = 0$, especially for bins which fall below the minimal TOF. To account for this, the entries and predictions of all bins with $n_j < 5$ are respectively summed up in a 'residual bin'. The total number of D.O.F. is thus variable.

Alternatively a maximum likelihood method has been implemented based on Poisson statistics in which the negative log-likelihood function

$$-\log L = -\log \prod_i \frac{\mu_i^{n_i} \cdot e^{-\mu_i}}{n_i!} = -\sum_i n_i \cdot \log \mu_i - \mu_i - \log(n_i!) \quad (4.22)$$

is minimized. For unknown reasons, the log-likelihood minimization showed a better performance in practice.

To combine measurements from different retarding potentials, the corresponding single-potential chi square or log-likelihood functions simply have to be summed up. The fits are performed with the MINUIT package [MIN]. The parameters are fitted by the MIGRAD routine and parameter error analysis performed by the MINOS tool. The errorbars correspond to $\chi^2 \pm 1$ and $\log L \pm 0.5$.

To obtain a symmetric χ^2 or $\log L$ parabola for neutrino masses near zero, there must also be an expression for a negative $m_{\nu_e}^2$ that fades continuously into the physical spectrum at $m_{\nu_e}^2 > 0$. To accomplish this, to each term in the sum of the beta spectrum (3.1) a factor

$$f_i = \left(1 + \frac{m_{\text{eff}}}{\epsilon_i} e^{-(1+\epsilon_i/m_{\text{eff}})} \right) \quad (4.23)$$

is applied in case of $m_{\nu_e}^2 < 0$ and $\epsilon_i + m_{\text{eff}} > 0$ and with $\epsilon_i = E_0 - V_i - E$ and $m_{\text{eff}} = \sqrt{-m_{\nu_e}^2}$ [Wei93b].

4.4 Simulation Results

4.4.1 Characteristics of TOF Spectra

A typical TOF spectrum, as it is obtained by (4.19) and normalized according to (4.16) is depicted in Fig. 4.1. The obtained spectrum is characterized by the following details:

- There exists a minimal TOF t_{min} . This corresponds to the maximum kinetic emission energy that an electron can have, given by $E_{\text{max}} = E_0 - m_{\nu_e}$.
- From t_{min} on, a steep slope begins, leading soon to a maximum somewhat above t_{min} , followed again by a long, slow fall. The maximum can be explained by the fact that the higher the energy becomes, the lower the number of electrons is, due to the shape of the end of the beta spectrum, whereas the 'TOF energy density', i.e. the interval size of the energy that corresponds to a certain TOF bin, increases. These effects balance each other, leading to a maximum somewhere in the middle.
- There is no maximal TOF. The closer the energy of an electron is to the retarding potential, the slower it will be. That means that electrons with an energy infinitesimally above the retarding energy will have an infinite TOF.

As it can be shown, changing the core parameters will have the following effects:

- In the last part of the beta spectrum, a variation of the endpoint E_0 will result approximately in a shift of the beta spectrum on the TOF axis (Fig. 4.2). That means for the TOF spectrum that t_{min} will be shifted to the left in case of an increase of E_0 and to the right in case of a decrease of E_0 since the maximum kinetic emission energy changes. The rest of the spectrum will be stretched or squeezed with respect to that. This implies also a slight shift of the maximum

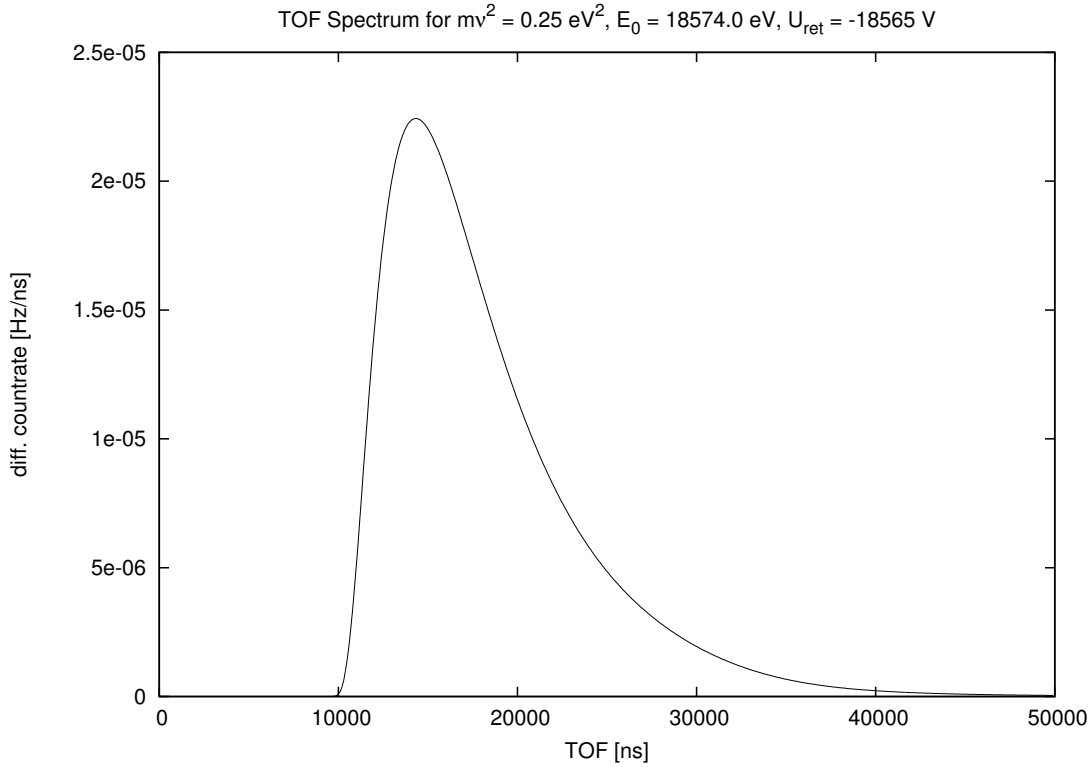


Figure 4.1: *Single TOF spectrum*

which is evident due to the fact that the distribution of electrons for lower energies will be higher but the 'TOF energy density', that only depends on U_{ret} , stays the same. Furthermore, when increasing the endpoint the total count rate, the integral over the whole TOF spectrum, increases as well. This is also evident since if the endpoint is increased, more events with emission energies between qU_{ret} and E_0 are produced.

- If the neutrino mass square $m_{\nu_e}^2$ is changed, the main signature is as well a shift of the minimal TOF (Fig. 4.3). Because of that, it is not easy to decorrelate the neutrino mass from the endpoint. However, while a change of the endpoint has effects on the whole spectrum, a change of the neutrino mass affects mainly the last few eV of the beta spectrum (see Fig. 4.5). That means that for the TOF spectrum the beginning of the spectrum is more affected. As it can be seen, the effect concentrates actually on the very first few bins. A shift of t_{min} due to a certain increase of the neutrino mass seems to be less than the shift cause by the corresponding decrease in endpoint. Of course, however, t_{min} is identical in both cases but the beginning of the curve is more flattened when increasing the neutrino mass. Due to the reciprocal behaviour of the TOF with respect to the energy, the quite sharp cut-off in the beta spectrum from the neutrino mass (fig. 4.5) leads

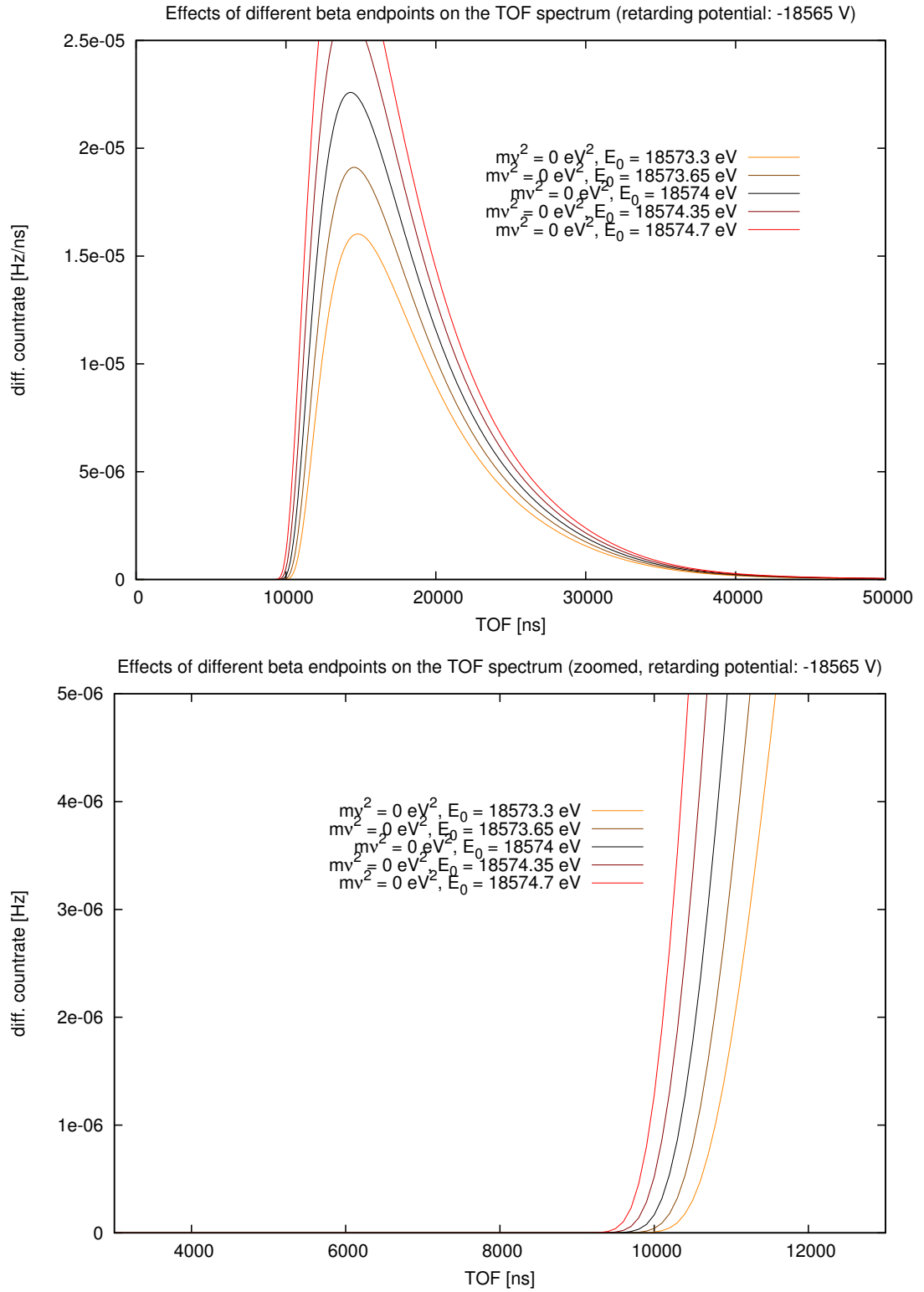


Figure 4.2: Effects on the TOF spectrum for different endpoints.

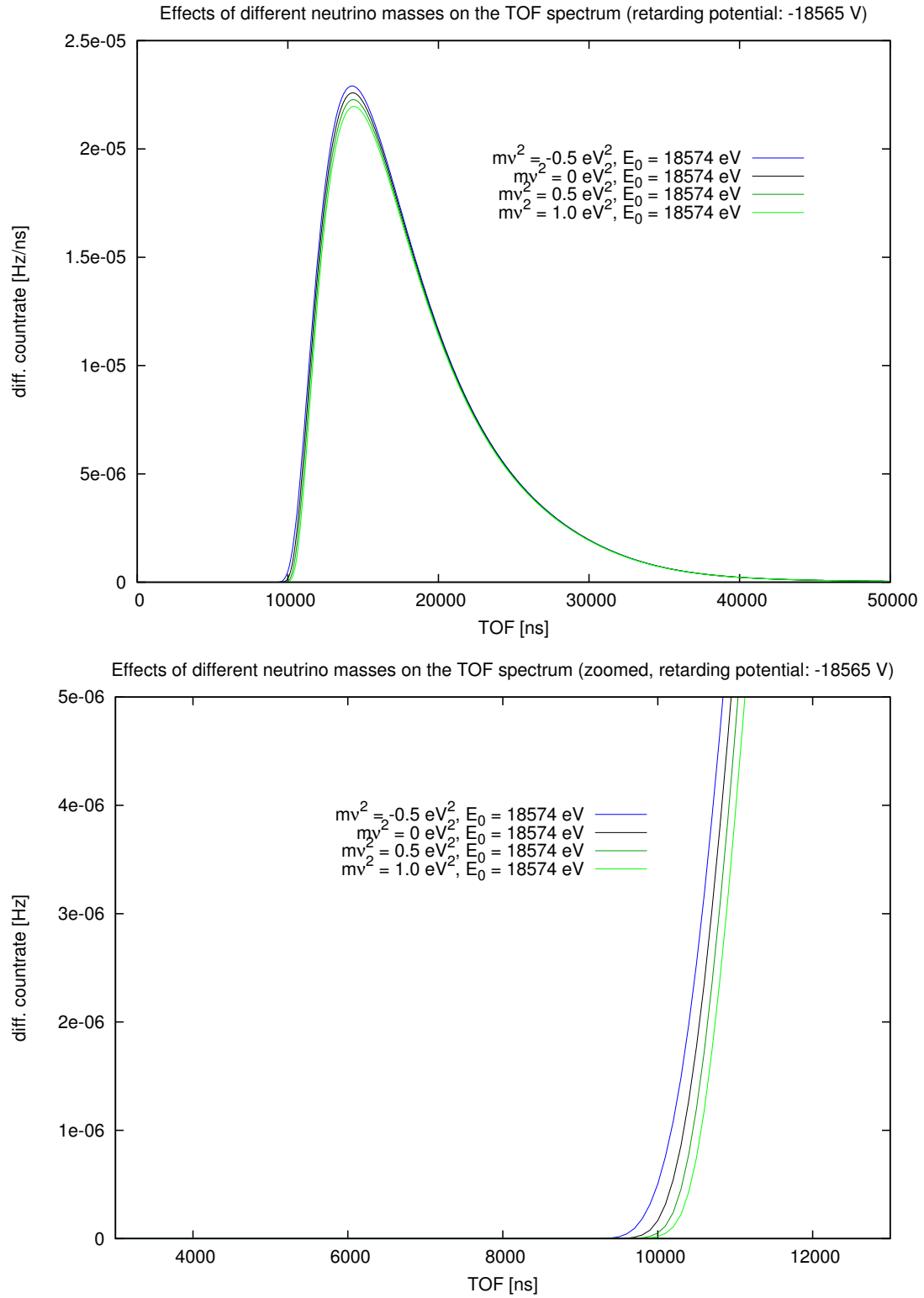


Figure 4.3: Effects on the TOF spectrum for different neutrino masses.

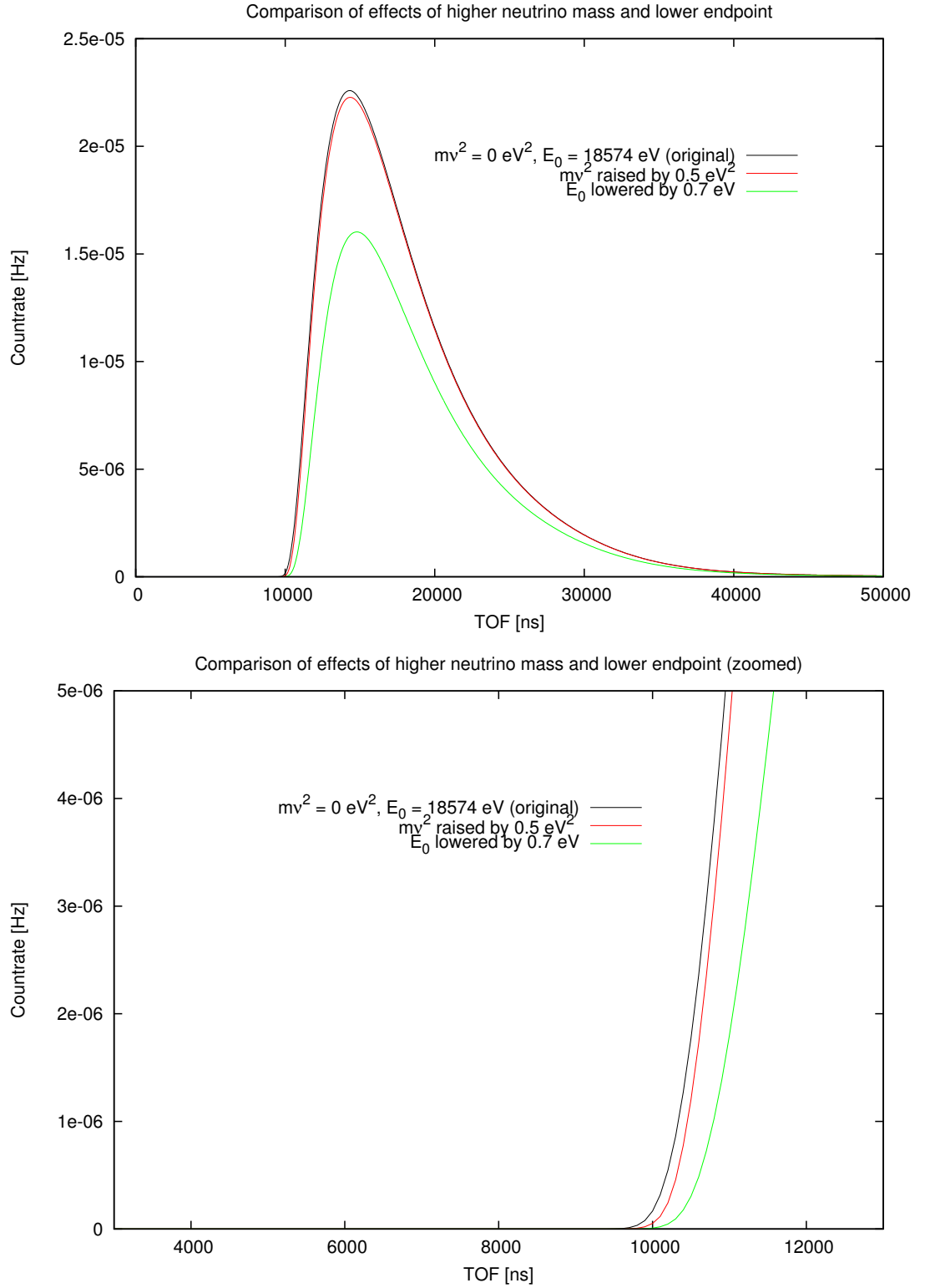


Figure 4.4: Comparison of the effects of higher neutrino mass and lower endpoint on the TOF spectrum.

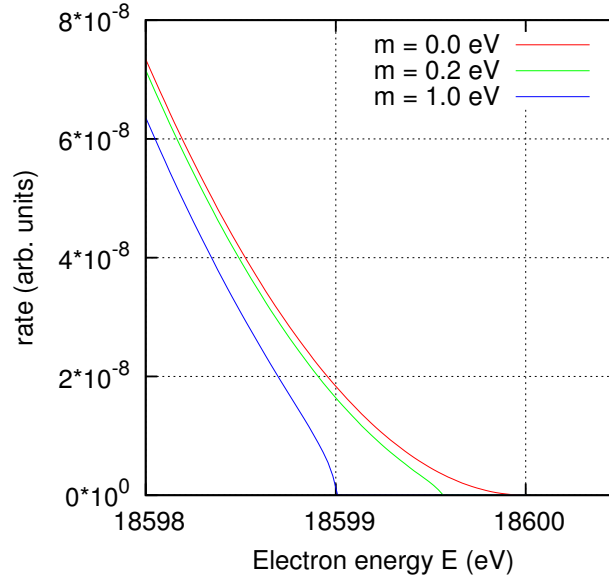


Figure 4.5: *Effects of the neutrino mass on the last part of the beta spectrum.* [Zac09]

to a slighter beginning of the TOF spectrum. A higher retarding potential leads to a clearer distinction (fig. 4.6) since the effects of the neutrino mass are mainly visible in the last few eV of the beta spectrum. Therefore, it seems to optimal to measure at retarding potentials near the endpoint. However, the count rate is rather low there and it is difficult to decorrelate the neutrino mass square from the endpoint, so measurements from lower retarding energies should presumably be added to the data.

Changing the retarding energy produces quite a similar effect as changing the endpoint. Here, the beta spectrum does not change but instead the 'TOF energy density' since the de-accelerating potential changes. The retarded velocity that the electrons have in the analysis plane of the main spectrometer is the main contribution to the TOF. The retarded velocity depends mainly on the surplus energy, i.e. the energy difference between potential and emission energy. The change of the endpoint leads approximately to a shift to the left or right, the change of the retarding potential however to a shift of the 'point of reference'. Thus, for the distribution of surplus energies which mainly governs the TOF spectrum, it basically does not matter if the retarding potential or the endpoint is changed.

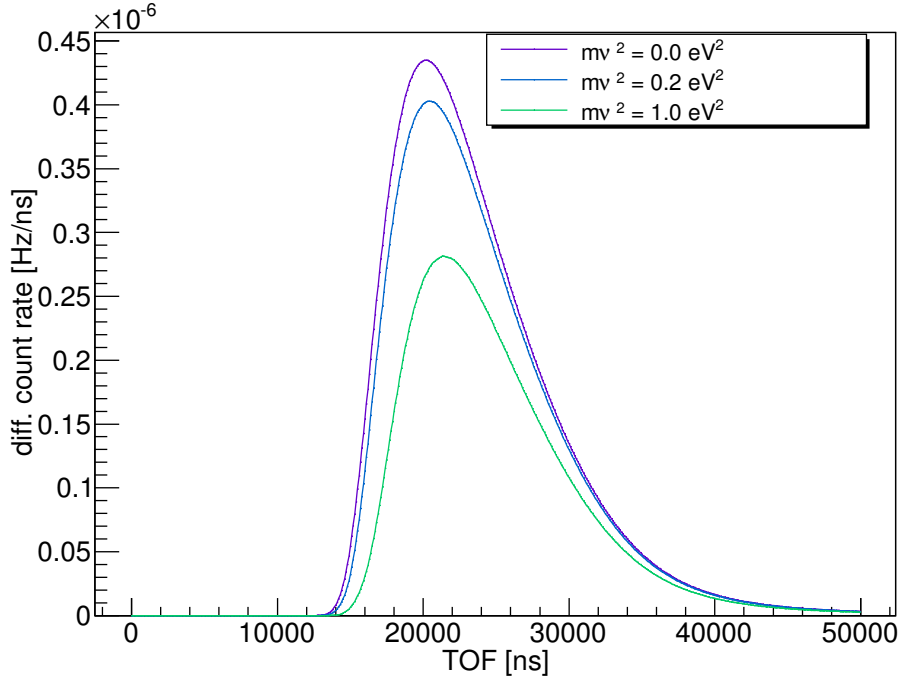


Figure 4.6: Effects on the TOF spectrum for different neutrino masses at a high retarding potential (18570 eV). The endpoint is $E_0 = 18574.0$ eV.

4.4.2 Fit Example of Basic TOF Spectra

It was shown that there are clear signatures for the neutrino mass inside the TOF spectrum. The obvious question is thus how significant they are. As mentioned in the beginning, the way to determine the sensitivity on $m_{\nu_e}^2$ is to simulate realistic TOF data by Monte Carlo techniques and to fit these with the models they were obtained.

Fig. 4.7 shows an example for such a fit. The data were generated by the Poisson law using (4.20). The random numbers are obtained by ROOTs *TRandom3* generator [ROOT]. For the creation of the data, it was assumed arbitrarily that $m_{\nu_e}^2 = 0.5$ eV² and $E_0 = 18574.0$ eV. The background was set to $b = 0$ and the signal rate to the estimate from (4.15). To show the principle of fitting several datasets simultaneously, three different retarding potentials have been taken for the data creation and the fit. These are -18565 V, -18560 V and -18555 V. Furthermore, it was assumed that each dataset is measured approximately one year, yielding $3.2 \cdot 10^7$ s. The maximum likelihood method (4.22) was used to perform the fit. For the interpolation the sampling points were spaced exponentially where the distance to the center grows by a factor of ~ 1.6 at each step, starting with 0.001 eV for the beta endpoint and 0.001 eV² for the neutrino mass squared. As it can be seen, the assumed neutrino mass squared of $m_{\nu_e}^2 = 0.5$ eV² is met by the fits estimated quite accurately by $(m_{\nu_e}^2)_{\text{est}} = 0.5001^{+0.0057}_{-0.0062}$ eV². The uncertainty is slightly unsymmetrical but to an acceptable extent.

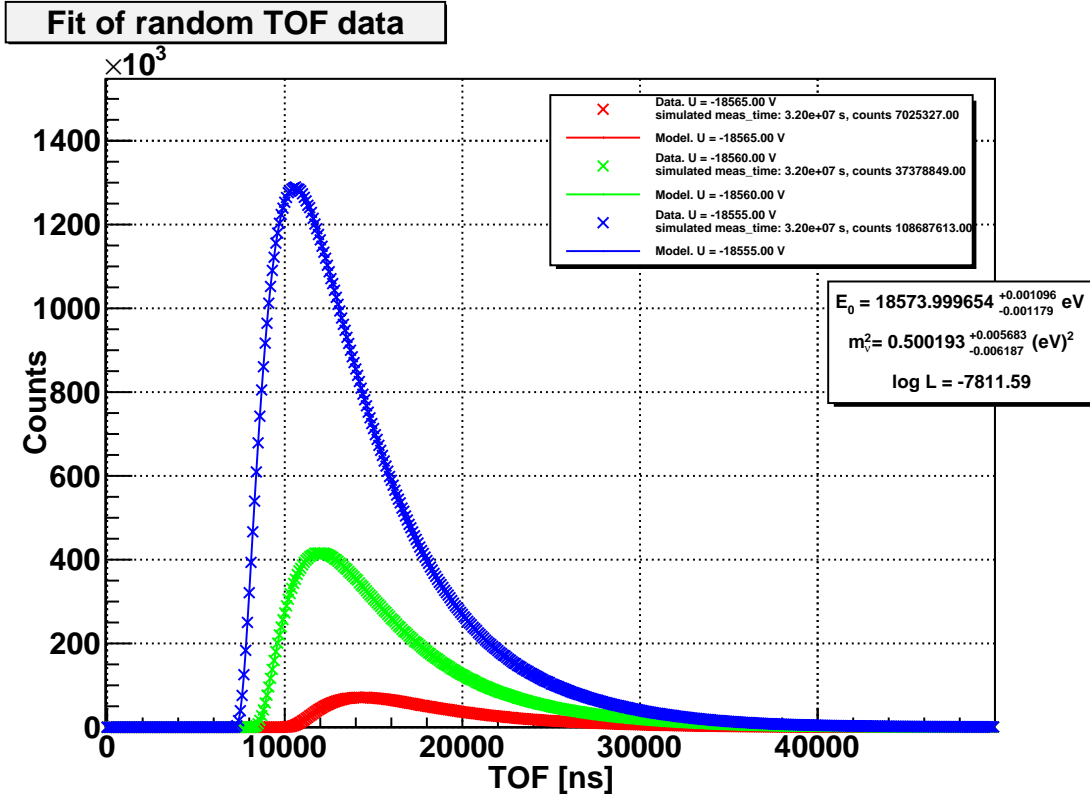


Figure 4.7: Example of a fit of MC created TOF data for $m_{\nu_e}^2 = 0.5 \text{ eV}^2$ and $E_0 = 18574.0 \text{ eV}$, using three datasets with retarding potentials -18565 V, -18560 V and -18555 V and assuming 1 y measurement time for each dataset.

Fig. 4.8 shows a fit with the same parameters but with $m_{\nu_e}^2 = 0 \text{ eV}^2$. Here, the fitted estimate is $(m_{\nu_e}^2)_{\text{est}} = 0.0019^{+0.0043}_{-0.0042} \text{ eV}^2$. It is obvious that the statistical uncertainty is lower than in the case of $m_{\nu_e}^2 = 0.5 \text{ eV}^2$. This can be explained by the fact that the shift of the beta endpoint is governed not by $m_{\nu_e}^2$ but by m_{ν_e} . This shift might be the main contribution to the signature in the TOF spectrum, even though the change of the phase space in the beta spectrum by a non-vanishing m_{ν_e} has an influence, too. As one can obtain from the zoomed area of fig. 4.3, the difference between two spectra decreases with growing $m_{\nu_e}^2$.

The result for $m_{\nu_e}^2 = 0$ reflects the *statistical sensitivity on the squared neutrino mass* for the TOF mode. That is the minimal $m_{\nu_e}^2$ that can be detected within 1σ , assuming that there is no additional systematic uncertainty. This quantity can be compared with the corresponding one for the KATRIN standard mode of operation (fig. 4.9). Yet, the improvement factor depends on the point of comparison, as the performance of the standard mode depends on a certain choice of parameters as well. There is no a priori preferred point of comparison. Instead, what needs to be compared is relative to the formulation of a question. The value achieved by the TOF example is about 6 times

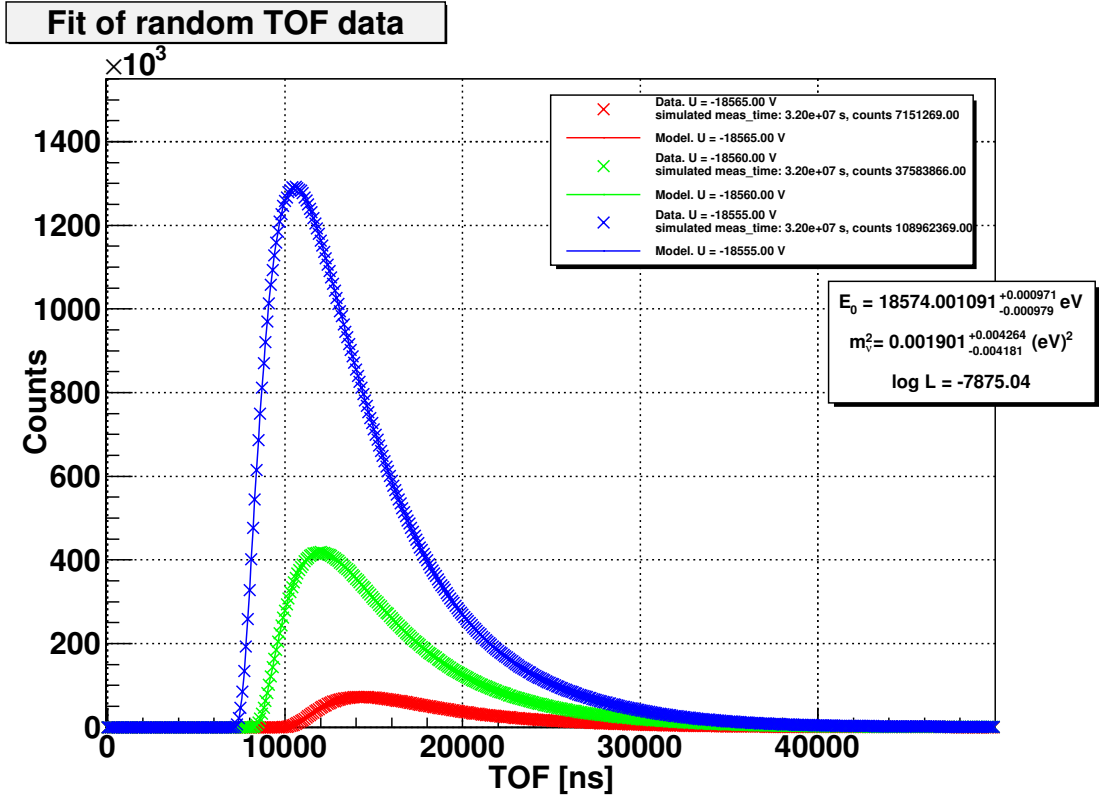


Figure 4.8: Example of a fit of MC created TOF data for $m_{\nu_e}^2 = 0 \text{ eV}^2$ and $E_0 = 18574.0 \text{ eV}$, using three datasets with retarding potentials -18565 V , -18560 V and -18555 V and assuming 1 y measurement time for each dataset.

smaller than the standard mode with no measurement time optimization (red line) and 10 mHz background. However, by a measurement time optimization, the result will improve for the standard mode (blue line). That might as well apply for the TOF mode but not necessarily by the same factor. Furthermore, the results for the standard mode improve with shrinking background (purple curve). Hence the additional question arises how good a TOF measurement scores in comparison to a sole background reduction, especially if a TOF measurement method intrinsically reduces background (see next chapter).

4.4.3 Measurement Time Optimization

The results above correspond to a uniform distribution of measurement time among the retarding potentials. For a further comparison with the standard mode for optimized measurement time distributions (fig. 4.9, blue and purple line), a measurement time optimization for the TOF mode itself has to be performed as well. Moreover, such an optimization probes which combination of retarding potentials are particularly sensitive

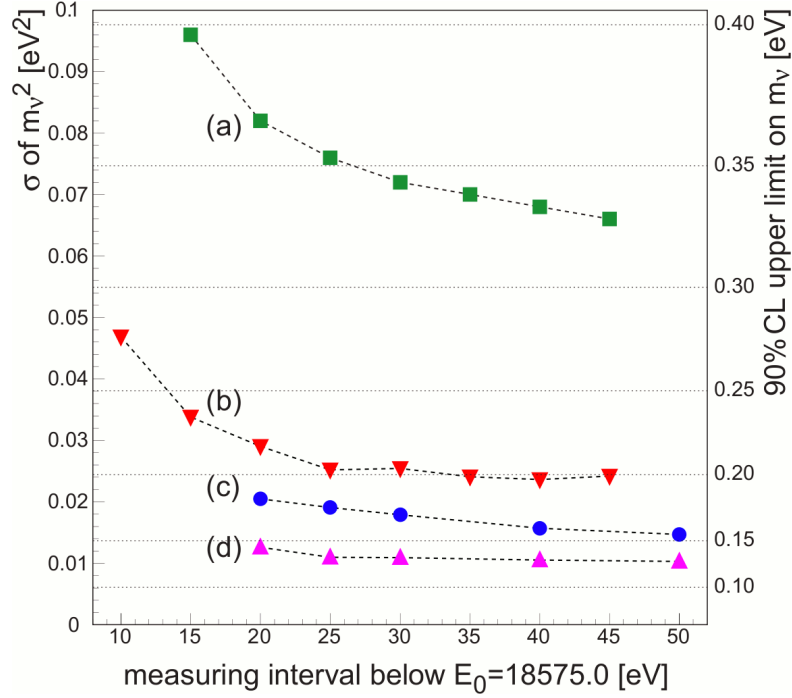


Figure 4.9: Statistical uncertainty of the KATRIN standard mode as a function of the measurement interval, i.e. the lowest retarding potential with respect to the beta endpoint. The left side of the y axis reflects the statistical uncertainty on $m_{\nu_e}^2$ and the right side the 90 % C.L. upper limit of m_{ν_e} . The green curve is based on the values for a main spectrometer with only 7 m diameter, the red curve on the reference values in the KATRIN design report, the blue curve on the latter plus measurement time optimization and the purple curve on the latter with a further reduction of the background by a factor 10.

on the neutrino mass. Thus, it can be checked if a minimization of the evaluation interval, i.e. the lowest retarding potential with respect to the beta endpoint, is possible. That would not only reduce the statistical but also the systematic uncertainty.

To achieve a certain degree of independence from random numbers and an behaviour of the statistical uncertainty as a function of the measurement time distribution that is as monotonous as possible, the following method was used. Let us assume a certain total amount of available measurement time t_{meas} and split it into n_{units} units of a constant measurement time Δt_{meas} :

$$t_{\text{meas}} = n_{\text{units}} \cdot \Delta t_{\text{meas}} \quad (4.24)$$

For each retarding potential n_{units} different Poisson datasets were generated. The total number of units n_{units} can now be distributed among the retarding energies, represented by a distribution vector \vec{n} . For example, $\vec{n} = (10, 0, 0, 2)$ means 10 units on the first

potential and 2 on the last. A certain number of units n_i for a certain potential means then that the first n_i generated datasets for that potential are summed up by summing the entries in each bin of the spectrum. That provides more stability than scanning the parameter space with unique datasets for each call. Otherwise, for instance, if some measurement time parameter is varied and the other left constant, there would be in each case a completely different spectrum for the constant one. The summed datasets are finally used for a fit. This is done for all possible permutations of \vec{n} that grant the unitary condition of $\sum_i n_i = n_{\text{units}}$. The statistical uncertainty is then obtained as a function of a discretized measurement time distribution.

Such a measurement time optimization was performed again for the parameters $m_{\nu_e}^2 = 0 \text{ eV}^2$ and $E_0 = 18574.0 \text{ eV}$ using four retarding potentials -18565 V, -18560 V, -18555 V and -18550 V. The background was again neglected and set to zero. A total measurement time t_{meas} of 3 years was assumed here, divided into n_{units} 12 units, representing 3 months each. That yields a total number of 454 possible permutations. The fit has been done with the modified chi squared minimization (4.21). The most important results are shown in table 4.10.

lowest potential [V]	distribution type	distribution vector \vec{n}	mean stat. error [eV ²]
-18550	uniform	(3, 3, 3, 3)	0.0040
-18550	optimized	(8, 3, 0, 1)	0.0038
-18555	uniform	(4, 4, 4, 0)	0.0041
-18555	optimized	(9, 1, 2, 0)	0.0040
-18560	uniform	(6, 6, 0, 0)	0.0043
-18560	optimized	(8, 4, 0, 0)	0.0043
-18550	single	(0, 0, 0, 12)	0.0044
-18555	single	(0, 0, 12, 0)	0.0044
-18560	single	(0, 12, 0, 0)	0.0045
-18565	single	(12, 0, 0, 0)	0.0050

Figure 4.10: Statistical uncertainty on $m_{\nu_e}^2$ (arithmetic mean of upper and lower bound) of uniform and optimized distributions depending on the lowest retarding potential for 3 years simulated measurement time, $m_{\nu_e}^2 = 0 \text{ eV}^2$, $E_0 = 18574.0 \text{ eV}$ and at most four retarding potentials -18565 V, -18560 V, -18555 V and -18550 V as well as for single retarding potentials.

In principle, it seems to be optimal to maximize the measurement time only at one retarding potential. However, as expected and confirmed, it is hard to decorrelate the neutrino mass from the endpoint in that case. Therefore, it is optimal to enrich the measurements with further data from different retarding potentials. It seems that the uncertainty is slightly smaller the further away from the endpoint additional potentials are added. This effect, anyhow, is small. Is it sufficient to add only one additional retarding potential rather close to the existing one, as shown by the distribution (8, 4, 0, 0) in

table 4.10. That might have beneficial consequences. Compared to the standard mode of operation where measurements at lower retarding potentials are necessary (cf. fig. 4.9), for a TOF mode, a lower bound of 15 eV below the endpoint is sufficient. Since the systematic uncertainty of KATRIN increases significantly at lower retarding potentials, it may be possible to improve even the systematics with a TOF mode.

Anyhow, the results found in table 4.10 suggest a possible improvement of the statistical uncertainty on $m_{\nu_e}^2$ of up to a factor 5, if compared to the measurement time optimized standard mode (fig. 4.9, blue curve). Compared to the scenario of a standard mode with only 1 mHz background (purple curve), the improvement is still up to a factor 3. The comparison is affected by the fact that there only exist only reference values for an optimized standard mode at ≥ 20 eV below the endpoint. Altogether, the benefit by measurement time optimization is not as big as in case of the standard mode, where it is about 40 - 50 %. There may be two explanations. On one hand, the optimization algorithm used in this simulation was more simple due to stability and performance. On the other hand, it is plausible that in case of an integrated beta spectrum where the measured function is itself composed only by the retarding potentials and the corresponding count rates the choice of the right measurement points is more critical than in case of the TOF spectroscopy where for each potential a full TOF spectrum is obtained.

Note that the improvement potential by the TOF as stated above is purely hypothetical and a best-case scenario. Each TOF measurement will have technical limitations, depending on the method. There will especially be a limited time resolution and an influence on the background level. The background can no longer be neglected as done by now. The next chapter is therefore dedicated to a consideration of these effects.

5 Simulation of the Measurement Method

5.1 Motivation

In the last chapter the examinations were of purely theoretical nature. No specific measurement method was assumed, instead, the only premises were the Tritium beta decay spectrum, the basic characteristics of the KATRIN experiment and an a priori assumption that for all electrons the 'start' and the 'stop' times are known with infinite precision. Obviously, such a condition cannot be fulfilled in reality. Yet, there is so far no direct way to measure the TOF of an electron within this context. While the measurement of a 'stop' signal requires a detector with a very good time resolution - a non-trivial but feasible task - the measurement of the 'start' signal is complicated. A direct measurement would require to detect incoming single electrons with minimal interference so that the already subtle neutrino mass signature is not erased from the TOF spectrum.

Two measurement alternatives are discussed in this chapter. The first one is the hypothetical direct on-the-fly detection of single electrons which will be called *electron tagging* in the following. The method is not applicable with the current technical means. However, some possible research lines will be brought up. Foremost yet, a simulation has been performed in which different characteristics of an electron tagger are modelled and integrated into the previous simulations from the last chapter. This gives a picture of the requirements that such a method will have to fulfil in order to outrange the standard mode of operation.

The second alternative has in contrast already been discussed in the past since it is in general compatible with the MAC-E-Filter set-up and not hard to be implemented. The idea is to periodically chop the electron beam so that it is pulsed. The method will be called *gated filter* in the following, motivated by the gate from signal processing. Especially the KATRIN design is well suitable for it since the pre spectrometer retarding voltage could be used as gate voltage. The method does not provide a direct measurement of each electrons starting time like in the tagging method but imposes a statistical distribution of starting times. It will be shown that it is mathematically comparable to tagging with a broad time uncertainty. For the pulsed filter a simulation has been performed, too, which should find out if a sufficiently sensitive TOF neutrino mass measurement can be accomplished with it.

5.2 Hypothetical Electron Tagging

5.2.1 Motivation and Possible Approaches

It is obvious that a direct measurement of the start time of each electron is the most effective way of measuring flight times. As it was already assumed following (4.5), an ideal position for a such a 'tagger' would be in the entrance of the main spectrometer. The reason is that the physically interesting retardation takes place in the main spectrometer and furthermore that by the pre spectrometer the signal rate is reduced to $\sim 10^3$ Hz before that. All electrons that pass this position would induce a 'start' signal in the tagger and would then open a short time window preferably in the order of some μ s. Within this time frame, a signal in the detector is expected. If the electron can not pass the potential barrier no event is taken up and nothing happens. If, however, the electron passes and hits the detector the time of the 'stop' signal is measured and the TOF stored which is simply given by the difference.

That set-up might furthermore serve as an intrinsic background veto. Events in the detector are only accepted in case of a prior signal in the tagger. This means that all background particles can only contribute within the short time frame opened by the tagger signal. If the signal rate is low enough, which could possibly be maintained by the pre-spectrometer, the detector will only be alert in a small fraction of the total measurement time. The background will then be reduced by that amount. Such a set-up would even benefit in case of driving KATRIN in the standard mode of operation.

Due to the requirement of detecting electrons with minimal interference the producibility of such a tagger is highly questionable. Two lines of research could be imagined which may contribute to that task. The first one would use layers of ultra-cold Rdyberg atoms. Such an attempt has been presented in [Jer10] and could theoretically be utilized for TOF spectroscopy. The idea is based on the principle that the states in highly excited atoms lie so close to each other that a minimal disturbance is sufficient to change their state. This might for instance be the 53s state of rubidium which could be excited to 53p by a passing electron. The 53s states can be de-excited radiately by stimulated Raman adiabatic passage (STIRAP) while the 53p states can be ionized by an electric field ("field ionization") and counted. That would allow to determine if an electron has passed.

A different approach could utilize electrical pickups similar to as used in Schottky mass spectroscopy (SMS) [Fran95]. The passing of a charged particle induces an image charge in such a detector which can be amplified and measured. In beams with multiple particles the *Schottky signal* of the amplified detector signal corresponds to statistical fluctuations of the spatial particle distribution. In a storage ring where the beam passes the pickup periodically the frequency components of the signal can be determined by Fourier transformation and thus the orbital period. The challenge for the linear beam of KATRIN would be to recognize the signals just for a single passage of an electron.

5.2.2 Mathematical Model

Let us assume that there exists a method of detecting electrons without changing their momentum. In that case there are some universal characteristics that could describe the effect of such a method on the measured TOF spectrum.

Time resolution σ_t

We assume a Gaussian error distribution with standard deviation $\sigma_{t,\text{tag}}$ for the starting time measurement. Furthermore, a gaussian time uncertainty $\sigma_{t,\text{det}}$ of the detector adds which we neglected by now for simplicity reasons. Since the convolution of two Gauss distributions is again a Gauss distribution we can express both error distributions by a single one with $\sigma_t = \sigma_{t,\text{tag}} + \sigma_{t,\text{det}}$. The new TOF spectrum (4.7) is then given by the convolution with the error distributions,

$$\frac{dR}{dt}_{\sigma_t} = \frac{dR}{dt} \otimes N(0, \sigma_t) \quad (5.1)$$

where $N(0, \sigma_t)$ is the symmetric normal distribution with a standard deviation of σ_t . According to the KATRIN design goals, $\sigma_{t,\text{det}} < 50 \mu\text{s}$ while the resolution of the tagger is unknown. However, it should be in the same order for a sufficiently precise measurement.

Tagging rate r_{tag}

Only a small fraction of start signals registered by the tagger is followed by a subsequent stop signal in the detector, producing a TOF event. It makes not much sense to speak of 'false' tagger signals since most of them are caused by real electrons that do not fulfil the transmission condition of the main spectrometer later. However, other sources producing additional tagger signals are supposable, too. These could comprise intrinsic effects, penning traps between pre and main spectrometer or background electrons created in the main spectrometer.

A high rate of tagging signals is itself not fatal because background is actually induced in the TOF spectrum only if a tagger signal is followed by a false event in the detector. As mentioned in the beginning, this is a fundamental advantage of a 'tagging-based' TOF method. However, if the tagging rate is high, the detector will be in the alert, background-sensitive state more often. The background in (4.20), which was unconsidered by now, can now be specified as function of the tagging rate,

$$b_{\text{tag}} = \eta \cdot b_{\text{raw}}. \quad (5.2)$$

As it was mentioned in the beginning of the chapter, most of the gross background rate b_{raw} (which is equal to the KATRIN design goal of 0.01 Hz if no additional background

sources are considered) is blocked by the intrinsic veto, being suppressed to a fraction η . If we assume that after each tagging signal the detector will be active for a time t_{\max} (which sets an upper limit to the measurable TOFs) then η is the probability that a tagging signal occurs within a time interval of length t_{\max} . It can be derived by a Poisson law where the expected number of tagger signals per time interval is given by $t_{\max} \cdot r_{\text{tag}}$.

$$\eta = p_{\text{tag}}([0, t_{\max}]) = 1 - \exp(-t_{\max} \cdot r_{\text{tag}}) \stackrel{t_{\max} r_{\text{tag}} \ll 1}{\approx} t_{\max} \cdot r_{\text{tag}} \quad (5.3)$$

A second effect of a high tagging rate is an increased pile-up probability. That is the case when a tagger signal is followed by an additional tagger signal within the same time frame, so the starting time corresponding to an event at the detector can not be distinguished. All pile-up events have to be rejected by the DAQ since the starting time can not be determined. It causes a loss of counts, which can be described by a reduced signal rate (4.15). The pile-up probability is equal to η , following from its definition in (5.3). That gives

$$S_{\text{eff}} = S \cdot (1 - \eta) \quad (5.4)$$

with S as old signal rate and S_{eff} as new effective signal rate. In the reference design of KATRIN the electron flux will be reduced by the pre-spectrometer to $\mathcal{O}(10^3 - 10^4 \text{ Hz})$. Assuming $t_{\max} = 50 \mu\text{s}$, that would result in $\eta \approx 0.04 - 0.40$, entering regions in which the tagging rate might become critical. Therefore, for tagging-based TOF spectroscopy the pre-spectrometer retarding potential should be increased which would reduce the tagging rate. The minimal electron flux which could be achieved by that without distortion of the measured region of the spectrum depends on the measurement interval and the transmission properties of the pre-spectrometer which were investigated in [Frae10]. The tagging rate might also become critical in case of a penning trap between pre- and main spectrometer [Hil11]. A solution of that problem is vital not only for the standard mode but also for the TOF mode.

Tagging efficiency ϵ_{tag}

Not all electrons are registered by the tagger. The intrinsic tagging efficiency ϵ_{tag} is the fraction of electrons that are detected by it. Besides the count-rate losses due to pile-up discussed above, the signal rate S is reduced by ϵ_{tag} ,

$$S_{\text{eff},\epsilon} = S \cdot (1 - \eta) \cdot \epsilon_{\text{tag}}. \quad (5.5)$$

Additionally, all electrons that pass the potential barrier but have not been unregistered by the tagger can not be distinguished from background any longer. If such a missing tag

is accompanied by another tag, a false starting time is recorded. That can be described as effective background source b_ϵ proportional to the count rate which adds to the normal KATRIN background. The gross background rate in (5.2) is then given by

$$b_{\text{raw}} = b_0 + b_\epsilon = b_0 + (1 - \epsilon_{\text{tag}}) \cdot R_T(qU_{\text{ret}}) \quad (5.6)$$

as a function the tagging efficiency and the integral count rate R_T from (3.9). While b_0 is < 10 mHz, b_ϵ is likely a moderate fraction of the count rate and will be the main contribution to the background in the tagging-based TOF spectroscopy mode. From (5.2) and (5.6) one can derive the condition that the tagging method has less background than the standard mode as

$$\eta \cdot \left(1 + \frac{R_T}{b_0} (1 - \epsilon_{\text{tag}}) \right) \leq 1. \quad (5.7)$$

While by now in (4.20) just the net additive background was treated as a fit parameter, this role is now adopted by b_0 which is in contrast to b_ϵ constant and not dependent on the count rate.

5.2.3 Simulated Effects of the Tagging Parameters

The effects of the quantities introduced in the last section on the statistical uncertainty of $m_{\nu_e}^2$ are shown in fig. 5.1. The results can be interpreted as follows.

The timing seems to be uncritical even for resolutions that are big compared to the one of the detector. For resolutions in the range up to $1 \mu\text{s}$ the error seems to increase only by about 10 %. Only for larger values beginning at about $1.5 \mu\text{s}$ the uncertainty increases significantly. The dependence is quadratic in good approximation. That behaviour might correspond well to the scale of the TOF spectrum where the neutrino mass sensitive part stretches from the beginning of the curve to the maximum over some μs .

The effects of the tagging rate are moderate for scales on which the pile-up probability is sufficiently low. For $r_t < 1000 \text{ Hz}$ the uncertainty increases quite steeply due to the sudden appearance of some background. This behaviour scales with a square root, enlarging $\sigma(m_{\nu_e}^2)$ by about 15 % at 1000 Hz and 30 % at 5000 Hz. For larger scales however the uncertainty increases faster again due to the increasing pile-up probability which saturates at 1 for high tagging rates. Therefore, an asymptotic exponential growth adds to the square root dependence.

The maximal TOF t_{max} was chosen to be $50 \mu\text{s}$ in these simulations. The part of the TOF spectrum which is cut off is therewith $< 0.5\%$. A lower t_{max} is supposable, as even the introduction of a minimal TOF t_{min} . These would in first order stretch the function shown in the center of fig. 5.1 on the TOF axis, allowing larger tagging rates. However, in case of a lower t_{max} the shifted cut-off would again have effects on the statistical

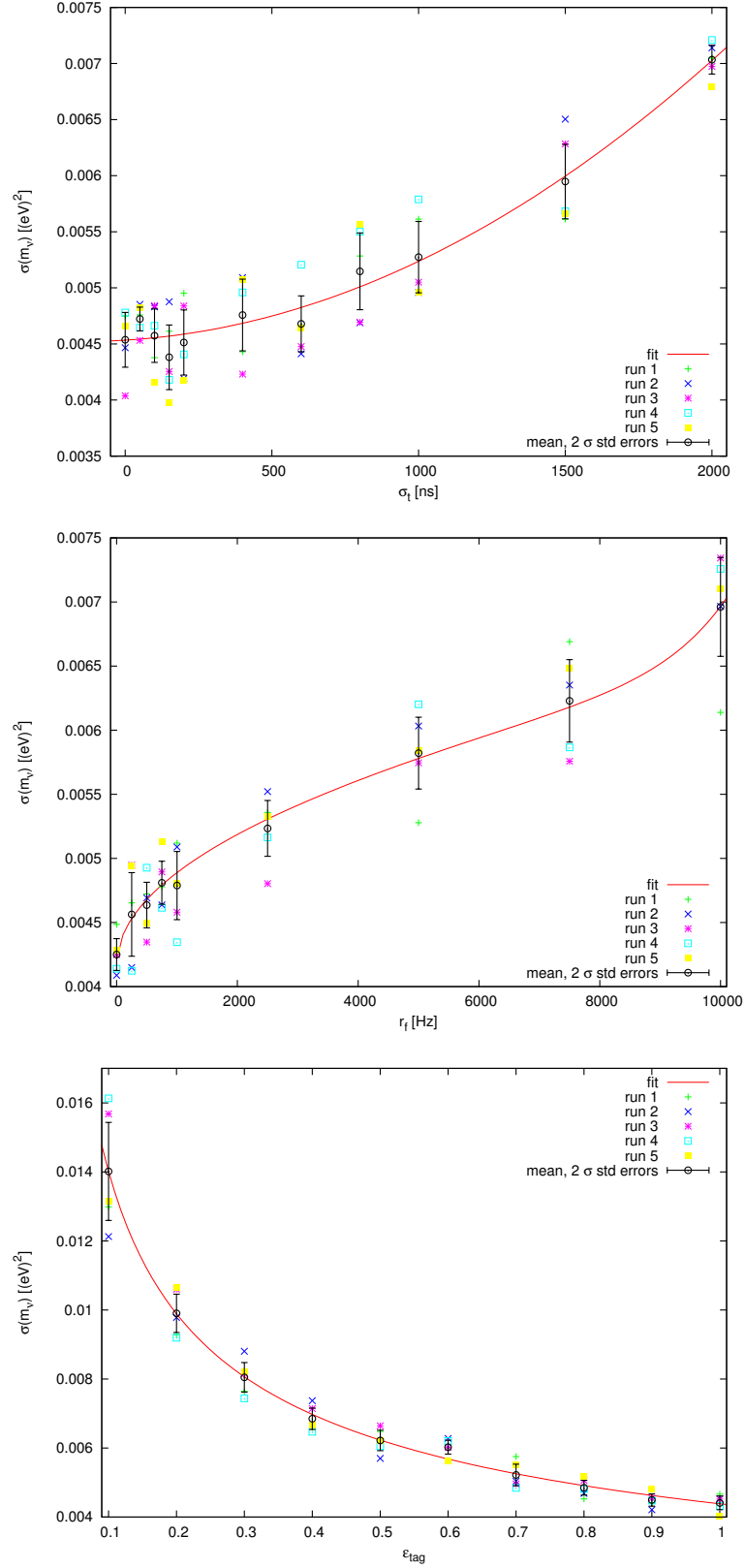


Figure 5.1: Effects of different time resolutions (top), tagging rates (center) and tagging efficiencies (bottom) on the statistical uncertainty of $m_{\nu_e}^2$. All results are based on 3 years simulated measurement time with a distribution of $2/3$ and $1/3$ respectively on 18565 V and 18560 V. b_0 was assumed to be 10 mHz, $E_0 = 18574.0$ eV and $m_{\nu_e}^2 = 0$. In every plot the results from 5 different simulation runs have been averaged. The other two parameters are set to their ideal value, each.

uncertainty and induce further background by the same mechanism as with a reduced tagging efficiency which needed to be simulated, too.

Concerning the tagging efficiency the simulation shows a clear dependence scaling with $1/\sqrt{\epsilon_{\text{tag}}}$. Here, the effect is purely based on the reduction of the effective signal rate as the tagging rate is set to zero whereby no background can arise. A slightly different behaviour might appear in the more physical case of a non-vanishing tagging rate.

5.2.4 Prediction of the Electron Tagging Performance

To summarize, the effects of the timing resolution is less critical if it is sufficiently low. Only on higher scales the relative growth becomes large. The tagging rate and the tagging efficiency should be optimized carefully. Both are in first order proportional to the effective background in the TOF spectrum and especially the tagging efficiency limits the event rate.

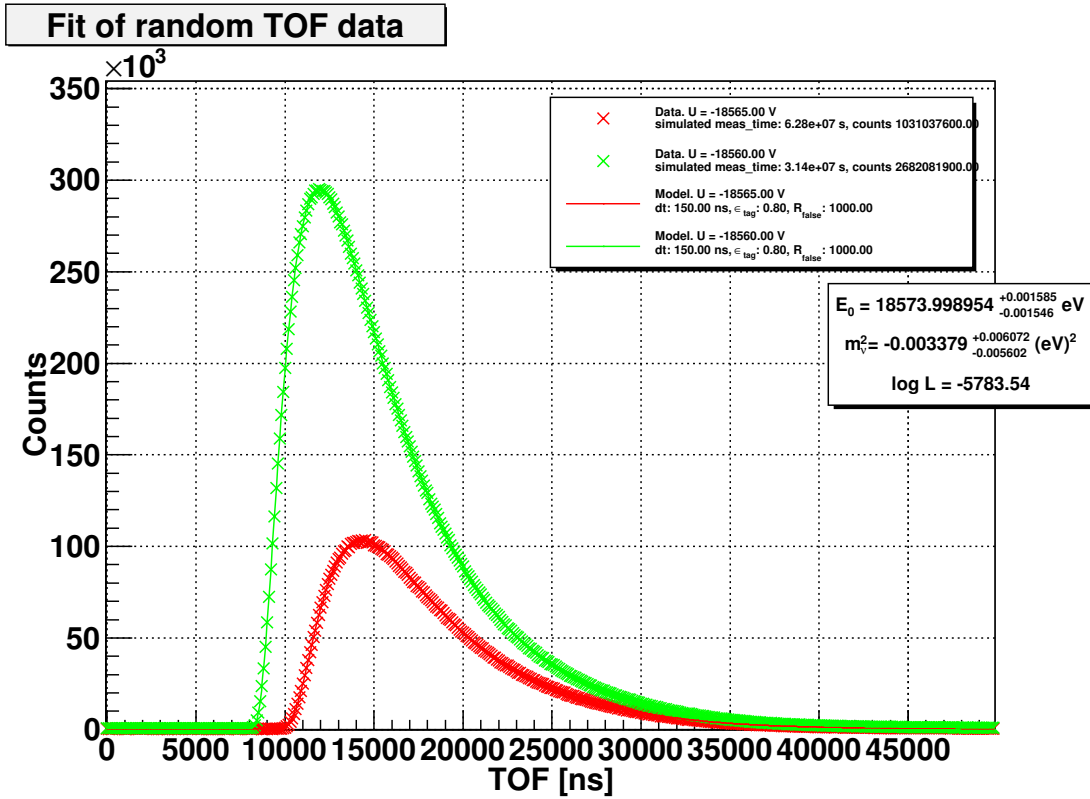


Figure 5.2: Example of a fit of data using an electron tagger. Tagging parameters are $\sigma_t = 150$ ns, $r_{\text{tag}} = 1000$ Hz and $\epsilon_{\text{tag}} = 0.8$. The result is based on 3 years simulated measurement time with a distribution of 2/3 and 1/3 respectively on 18565 V and 18560 V. b_0 was assumed to be 10 mHz, $E_0 = 18574.0$ eV and $m_{\nu_e}^2 = 0$.

An example of a fit based on values for σ_t , r_{tag} and ϵ_{tag} that might be 'realistic' is shown in fig. 5.2. The reference value of $\sigma_{\text{stat}}(m_{\nu_e}^2) \approx 0.0043 (\text{eV})^2$ for that measurement time distribution without measurement simulation declines by $\sim 40\%$. Compared to the standard mode with measurement time optimization, the improvement factor would in this example decrease to 3 - 4 depending on the point of comparison (see 4.9, blue line). A measurement optimization using the same tagging parameters confirmed the behaviour found in the measurement time analysis of the last chapter even for the tagger case.

If parameters as in the example above are imaginable then it is not hard to figure out that the background suppression condition (5.7) might easily be violated. One should therefore take in mind the following conclusions.

- It reveals that the intrinsic background suppression mechanism which the tagger is supposed to entail is already easily disturbed by the fact that a limited tagging efficiency creates further background. The closer investigation shows that this benefit should not be taken for granted.
- For $\epsilon_{\text{tag}} = 1$, eq. (5.7) is a priori fulfilled. While the reduction of η might be limited by the pre spectrometer, the best way to avoid background is thus to optimize the tagging efficiency. Whenever research in this area gets serious, this should be the primary goal.
- However, the matter is not that crucial insofar as the effective background induced by $\epsilon_{\text{tag}} < 1$ is proportional to the count rate. Measurements close to the endpoint which are usually very background-critical due to their low count rate are not as much afflicted as measurements further away.

One might consider a comparison to the standard mode with a reduced background, too, as the 'electron tagger' could also be driven in the standard mode whilst keeping its background-reducing capabilities. However, there exist no reference values, since the background reduction by the tagger depends on the count-rate and thus on the retarding potential. For the parameters in fig. 5.1, using eqs. (5.3) and (5.6), the background is reduced by ~ 5 at 18565 eV and ~ 1 at 18560 eV. The simulated result is anyhow double as good as in the case of a standard mode with 1/10 background (4.9, purple line).¹ That allows a central conclusion regarding TOF spectroscopy. It seems that the benefit of a TOF mode comes essentially from the gain of information by measuring a differential spectrum and only at second order from a reduced background.

¹ $\sigma_{\text{stat}}(m_{\nu_e}^2)$ is assumed to scale with the background rate as $\sim b^{1/5}$ in the standard mode [KAT04].

5.3 Gated Filter

5.3.1 Principle and Mathematical Model

Since the development of a tagger is unrealistic for the near future, the question arises if an already existing measurement method could provide TOF measurements. A method that has been discussed in the past is to create a pulsed electron beam instead of a continuous one. [Bon99] However, a somewhat different approach is presented in the following which varies in the following points.

- The MAC-E-TOF mode in the past was in the first instance intended to be used for a neutrino mass measurement but for auxiliary measurements, as for instance spectral anomalies or studies of the energy loss.
- The past approaches proposed to combine the TOF measurement with the MAC-E-Filter to a *band-pass filter*. The arrival times in the detector are recorded and all too fast and too slow electrons are rejected. As in the standard mode of operation, a single count rate is obtained for each retarding potential which however does not correspond to the whole integral from the retarding energy to the endpoint but only to a narrow energy interval. However, for a neutrino mass measurement, especially the fast electrons are important and should not be rejected. Instead, as stated in the last chapter, the aim is to obtain a full TOF spectrum for each retarding potential, i.e. to have a *TOF spectroscopy* mode. The arrival times should not just be used for filtering but provide themselves the TOF spectrum.

To obtain a pulsed beam, the continuous flux of beta decay electron has to be chopped or *gated* by some means. This could be achieved by a HF modulation of the source or the pre-spectrometer potential. While pulsing the source potential is the 'classical' approach, in a two-spectrometer set-up like KATRIN it is more recommended to pulse the retarding potential of the pre-spectrometer. In that case the TOF spectrum is not distorted by the transport section and the retardation of the pre-spectrometer. Furthermore, the HV pulse levels do not need to be exact. It is only necessary to have one setting for full transmission and one for zero transmission.

A pulsed flux can in the simplest case be described by two timing parameters. The first one is the HF period t_r with which the flux is gated. After each period the detector clock is reset. The second one is the time t_s in which the gate is open in each period. The ratio of t_s and t_r gives the *duty cycle*

$$\Delta = \frac{t_s}{t_r}. \quad (5.8)$$

In contrast to an electron tagger, the flight times can not be measured for each single electron. However, this is not categorically necessary since we are not interested in single TOFs but in the TOF spectrum. A pulsed beam allows no direct measurement of the starting times but restricts them to certain intervals of length t_s . As one can imagine,

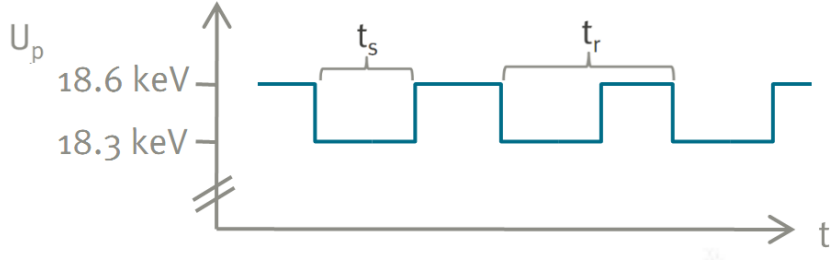


Figure 5.3: Timing parameters of the gated filter. *X axis: time. Y axis: pre-spectrometer retarding voltage. At 18.3 keV all electrons of the interesting endpoint region are transmitted while at 18.6 keV all electrons are blocked. The levels are examples, other values could be used as well.*

for $t_s \rightarrow 0$ and sufficient period lengths t_r one gets infinitesimally sharp starting times with infinitesimally low luminosity. If t_s is extended, the luminosity increases and the starting times are afflicted with an uncertainty. The uncertainty is given by a uniform probability distribution in the interval $[0; t_s]$. As in the case of the tagger the measured TOF spectrum is then given by the convolution with the detection time and the starting time distribution,

$$\left(\frac{dR}{dt} \right)_{t_s} = \frac{dR}{dt} \otimes N(0, \sigma_d) \otimes U(0, t_s). \quad (5.9)$$

If the detector clock is periodically reset with times t_r then some electrons with flight times $> t_r$ might hit the detector not in this but in the next period. Therefore, the final measured spectrum is a superposition of all contributing time distributions (5.9), shifted by multiples of t_r and finally cut off at 0 and t_r :

$$\left(\frac{dR}{dt} \right)_{t_s, t_r}(t) = \begin{cases} 0 & t < 0 \\ \sum_{n=0}^{\infty} \frac{dR}{dt}_{t_s}(t + n \cdot t_r) & 0 \leq t \leq t_r \\ 0 & t > t_r \end{cases} \quad (5.10)$$

As $> 99.5\%$ of the flight times lie within $\sim 50 \mu s$, all contributions with $n \cdot t_r \gtrsim t_s + 50 \mu s$ can be neglected. The resulting measured time spectrum (5.10) should not be confused with the physical TOF spectrum. Precisely spoken, it is rather a detection time spectrum of the detector clock which is reset periodically.²

5.3.2 Simulated Effects of the Timing Parameters

An exemplary measured TOF spectrum according to (5.10) is shown in fig. 5.4. The two characteristics mentioned that describe a simple periodic gate show clear signatures

²This holds true for an electron tagger as well. The resets correspond then to the tagger signals.

in the curve. The uniform error distribution imposed on the spectrum by the uniform uncertainty (5.9) leads to a clear broadening of the shape. Since the time precision is lower than in the tagger case the broadening is stronger visible. In addition, the smearing with the step function leads to steeper edges than the Gaussian one. A clear sign of the detector resets are the residuals from former gate periods at the beginning of the spectrum. The effects of the timing parameters on the spectral shape allow some preliminary predictions on the effects of the performance:

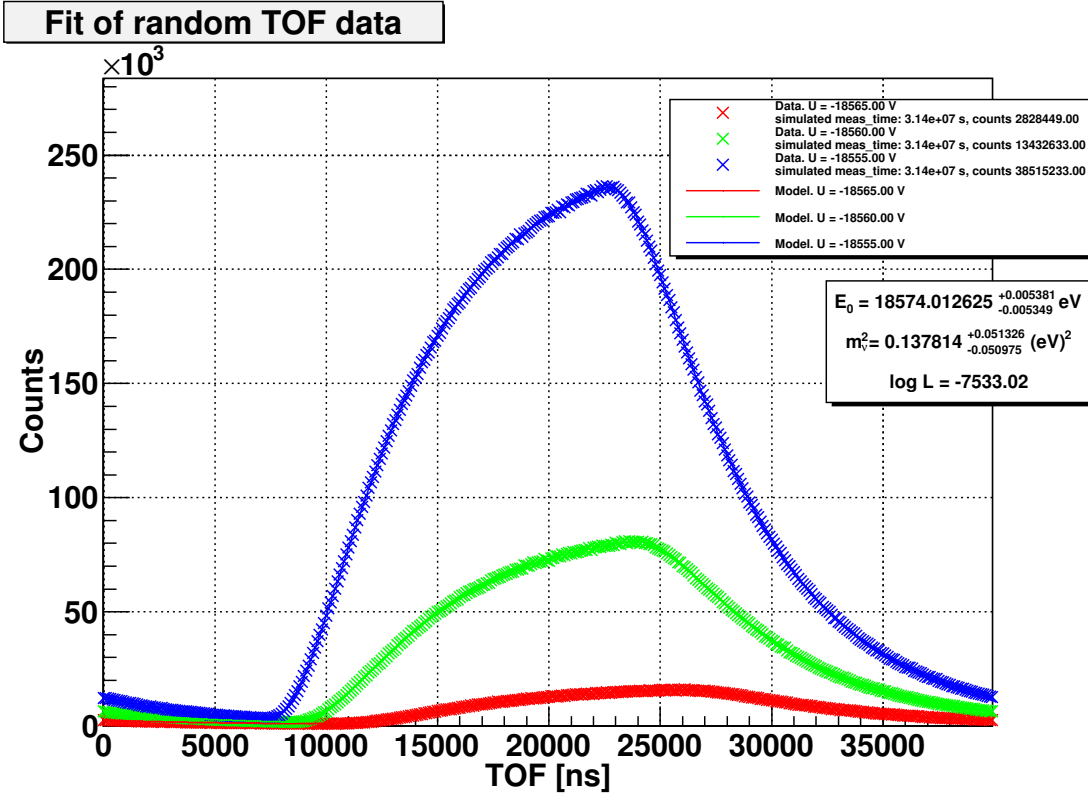


Figure 5.4: Example of a fit of a TOF spectrum with a gated filter. Timing parameters are $t_r = 40 \mu\text{s}$ and $t_s = 15 \mu\text{s}$. On the left side of the spectrum the residuals from earlier cycles can be seen which emerge continuously from the end of the spectrum. The peaks exhibit effects of the convolution with a step function. The results are based on 3 years simulated measurement time with a 1 year each at -18565 V, -18560 V and -18555 V. b_0 was assumed to be 10 mHz, $E_0 = 18574.0 \text{ eV}$ and $m_{\nu_e}^2 = 0$.

- For constant t_s , reducing t_r will increase the duty cycle. However, more residuals from former periods contaminate the spectrum. That acts like a non-uniform background. Duty cycle and residual contributions need to be balanced.
- For constant t_r , reducing t_s will reduce the time uncertainty. In contrast, the duty cycle will be reduced, resulting in a lower count-rate. Here, the timing and the duty cycle need to be balanced.

An optimization of t_s and t_r on basis of the measurement time distribution found to be optimal in the last chapter has been done. The effects on $\sigma_{\text{stat}}(m_{\nu_e}^2)$ are plotted in fig. 5.5. The significance seems to be moderate. Only extreme values of t_s with respect to t_r produce clearly a higher uncertainty. It seems that lower values of t_r are slightly preferred. That might be due to better combinations of duty cycle and timing allowed by that. However, also an accumulation of convergence problems occurred in that case as larger parts of the spectra get superposed. Since the improvement at lower t_r is small it seems senseful to leave t_r at $\sim 40 \mu\text{s}$ where the residuals that reach the beginning of the spectrum at τ_{min} are negligible.

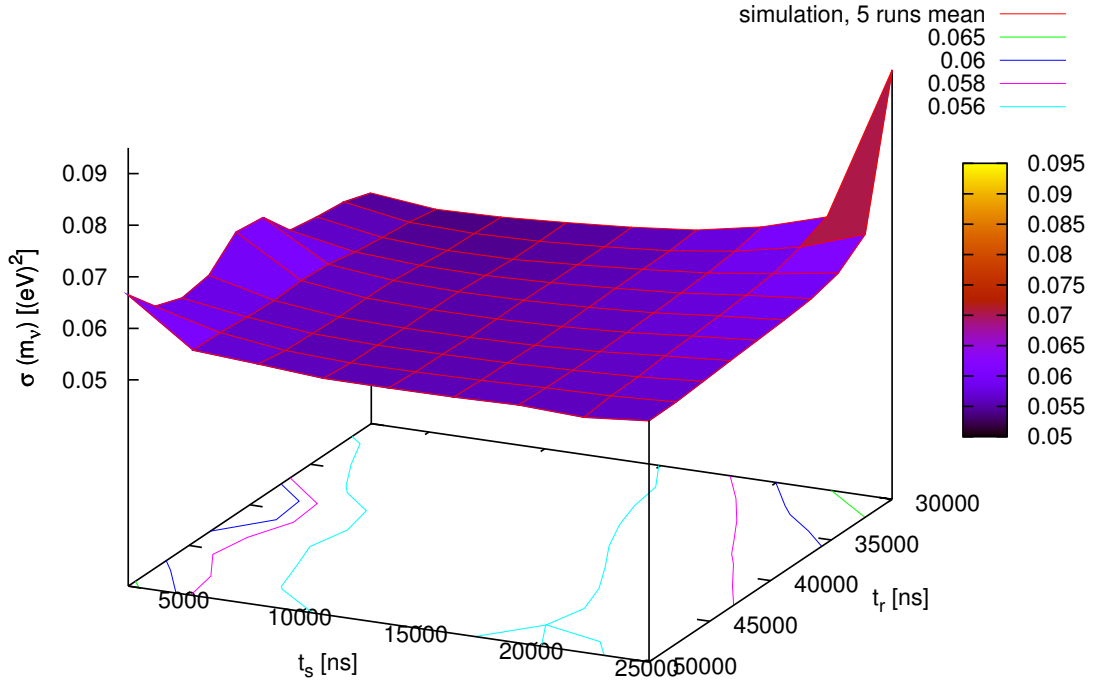


Figure 5.5: Effects of different period times t_r and gate opening times t_s on the statistical sensitivity of $m_{\nu_e}^2$. The results are based on 3 years simulated measurement time with a distribution of 2/3 and 1/3 respectively at -18565 V and -18560 V. b_0 was assumed to be 10 mHz, $E_0 = 18574.0 \text{ eV}$ and $m_{\nu_e}^2 = 0$. For each grid point the results from 5 different simulation runs have been averaged.

5.3.3 Predictability of the Gated Filter Performance

The minimal values of $\sigma_{\text{stat}}(m_{\nu_e}^2)$ in fig. 5.5 lie in the range of $\sim 0.05 (\text{eV})^2$, about 2 - 2.5 times worse than the statistical sensitivity of the standard mode. That would disqualify the gated filter. However, there are more parameters which can be optimized

than a global choice of t_r and t_s . As the gated filter is most cases less time-sensitive than a tagger would be, it makes sense to retreat from the paradigm of using as few and as high retarding potentials as possible. It makes more sense to establish a KATRIN-like measurement time distribution with more retarding potentials, most densely spaced around the endpoint.

As an example the following retarding potentials have been chosen: -18555, -18560, -18565, -18567.5, -18570 and -18575 V. The latter one is above the endpoint that was assumed, thus being sensitive to the background level. It is supposed, that for higher retarding potentials a higher duty cycle makes sense. The filter settings near the endpoint consist in any case of only low energetic electrons, so an additional separation makes less sense than for lower potentials. Rather, the increase of duty cycles brings additional count-rate which is crucial near the endpoint. Therefore, a second measurement time optimization was performed where the duty cycle t_s/t_r is as well a free parameter for each retarding potential. That gives 12 free parameters. An optimization via the method applied in the last chapter is not possible since on the one hand the spectra can not be simply added if the duty cycle is modified. On the other hand the evaluation of all possible permutations makes no sense for 12 parameters. Therefore, a brute force approach was chosen. Each parameter is scanned successively and set to its local optimum by fitting the average uncertainty to a step function and determining the minimum. This is repeated until the improvements per iteration are sufficiently small.

(duty cycle, measurement time fraction) at						$\sigma_{\text{stat}}(m_{\nu_e}^2)$ [eV ²]
-18555 V	-18560 V	-18565 V	-18567.5 V	-18570 V	-18575 V	
(0.5, 1/6)	(0.5, 1/6)	(0.5, 1/6)	(0.5, 1/6)	(0.5, 1/6)	(0.5, 1/6)	0.033
(0.4, 1/13)	(0.6, 1/13)	(0.6, 2/13)	(0.4, 1/13)	(1.0, 4/13)	(1.0, 4/13)	0.025

Figure 5.6: Statistical uncertainty of $m_{\nu_e}^2$ (arithmetic mean of upper and lower bound) for uniform and optimized distributions of a pulsed filter setup for 3 years simulated measurement time with $m_{\nu_e}^2 = 0 \text{ eV}^2$, $E_0 = 18574.0 \text{ eV}$ and -18555 V as lowest retarding potential. t_r was held constant at $40 \mu\text{s}$.

The results for a uniform measurement time distribution with duty cycles set to 0.5 each and for the optimum after 5 iterations are shown in table 5.6. The optimum is plotted in 5.7. It can be seen that already by the choice of six retarding potentials the results from fig. 5.5 improve by roughly a factor 2. The optimization of duty cycles and measurement times provides a further improvement of $\sim 30 \%$. By that the order of magnitude of the standard mode is reached but not exceeded.

The impact of these results should not be overrated since the retarding potentials that have been chosen were arbitrary, the minimization method was primitive and t_r , which can also be set separately for the retarding potentials, has been held constant. However, it seems not very likely the gated filter will reach a result that outreaches the standard mode significantly. Minimal improvements would not justify the technical expenses that would be spent for the implementation of the gated filter and could anyhow be accomplished by prolonging the running period of KATRIN. Hence, the simulation suggest that

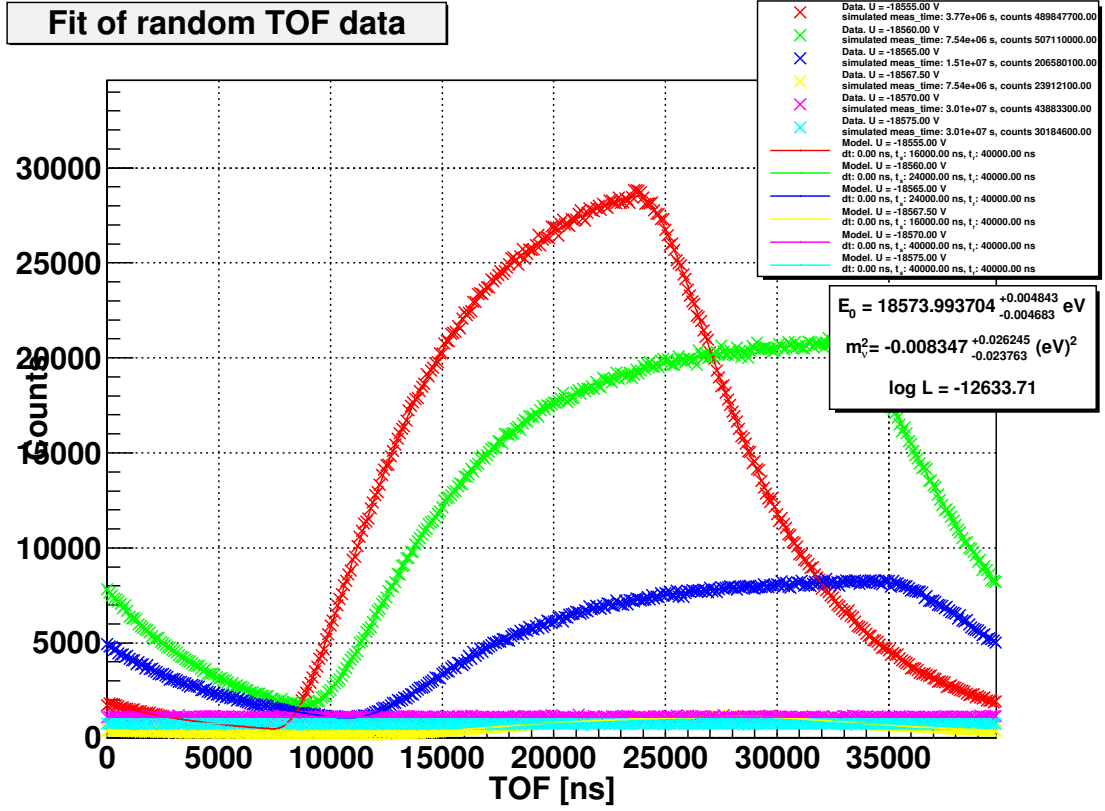


Figure 5.7: Fit of TOF spectra for the measurement time optimum with the gated filter from table 5.6.

for an appropriate TOF spectroscopy method an electron tagging method or comparable would be necessary.

6 Conclusion

The investigations described in the last two chapters throw interesting light on TOF spectroscopy as means of measuring the electron neutrino mass. The results of that study shall be discussed for the TOF method in general and for the investigated measurement scenarios.

6.1 Statistical Potential of a TOF Mode

A TOF spectroscopy mode could in principle provide significant improvements in the statistical neutrino mass sensitivity over KATRIN's standard mode of operation. The values from table 4.10 are best case results since they have been obtained with zero background and infinite time resolution. Moreover, they are based on optimized measurement time distributions. Hence, they can be interpreted as reflecting the *maximal statistical capability* of a TOF measurement. Thereby, the study revealed the following aspects.

- The results are already good at high retarding potentials while KATRIN in contrast needs to measure at least down to 20 eV below the endpoint. Thereby, it is completely sufficient to consider only two retarding potentials, for instance -18565 V and -18560 V. Even a singular potential delivers a full TOF spectrum and thus produces acceptable results that may still outrange the standard mode.
- This suggests that even the systematic uncertainty can be reduced with a TOF mode since it is especially critical at lower retarding potentials.
- The benefits of measurement time optimization, though existent, seem not as strong for the TOF mode as for the standard mode. The main reason may possibly coincide with the first point named. Since already single retarding potentials provide a full TOF spectrum in the TOF mode, the optimization of the choice of retarding potentials and the corresponding measurement times may be not as critical as for the standard mode.

A quantitative analysis of the improvement potential of the TOF mode relative to the standard mode depends on which points are compared on fig. 6.1. The choice should be guided by the study interest the comparison refers to. If the goal is to find out how much a TOF mode could maximally profit over the current design, it may be most plausible to compare the 'optimal' point of KATRIN with the 'optimal' one of the TOF mode. For KATRIN a good compromise between systematics and statistics

may be a lower bound of 30 eV below the endpoint. For this point the reference value of $\sigma_{\text{stat}}(m_{\nu_e}^2)_{\text{SM}} = 0.018 \text{ eV}^2$ is valid. For the TOF mode the best compromise may be the case of two retarding potentials -18565 V and -18560 V with 2/3 and 1/3 of the measurement time, respectively. This will in the optimal case of negligible background and apparative distortions obtain a statistical uncertainty of $\sigma_{\text{stat}}(m_{\nu_e}^2)_{\text{TOF}} \approx 0.0044 \text{ eV}^2$ for three years measurement time. It corresponds to a maximal improvement of the statistical sensitivity on $m_{\nu_e}^2$ of a factor 4 and on the absolute value m_{ν_e} of a factor 2.

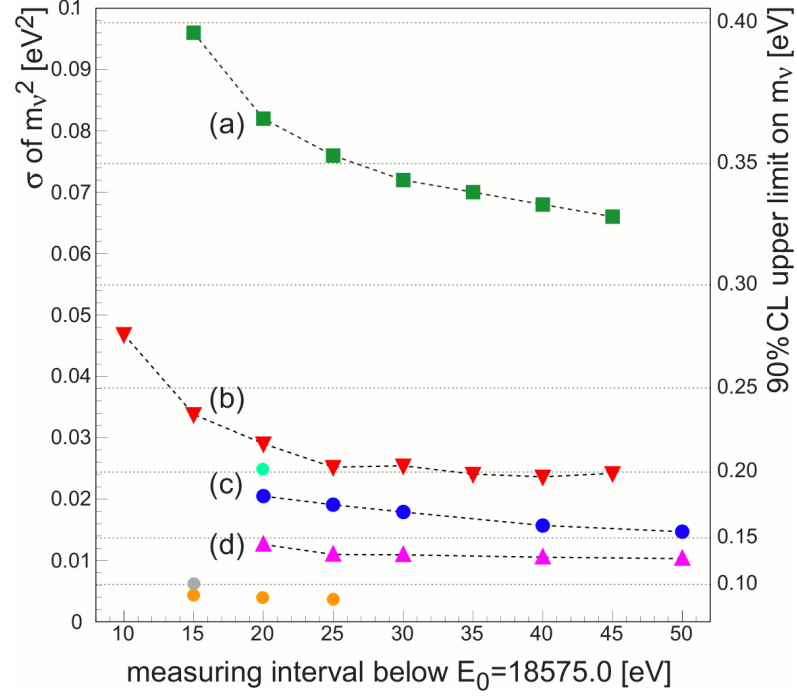


Figure 6.1: Statistical uncertainty of the standard mode as a function of the measurement interval from fig. 4.9 with simulation values for the TOF mode added. *Orange:* TOF with no background and infinite time resolution. *Turquoise:* TOF with gated filter. *Gray:* TOF with tagger using $\sigma_t=150 \text{ ns}$, $r_{\text{tag}}=1 \text{ kHz}$, $\epsilon_{\text{tag}}=0.8$. All TOF results measurement time optimized.

That comparison however does not take into account the systematic uncertainty, which will likely improve by the high retarding potential choice as well. A raw comparison of the statistical uncertainties without any consideration about the systematics should refer to identical measurement intervals in fig. 6.1. The value is then dependent on the interval. Some hypothetical TOF measurements methods like 'electron tagging' furthermore might also reduce the background level in the standard mode of operation. For any given parameterization of such a case the arising alternative of using the standard mode with a limited background¹ should as well be compared to the full TOF spectroscopy.

¹such an example with a background reduction of 10 corresponds to the purple line in fig. 6.1

6.2 Measurement Method - Fiction vs. Reality

The discussion of measurement method embraced two scenarios of which the first is rather a hypothetical thought experiment, though suggesting a potentially interesting field of research, and the second a more realistic concept which could already be implemented by the current technical design of KATRIN. The idea of the electron tagger provides a good performance in the simulations as long as the technical parameters are held moderate. Imaginable orders of magnitude for these values provide a result which is roughly up to 50 % worse than without measurement simulation but still outranges the standard method of operation. Within that context it could furthermore be shown that a TOF spectroscopy mode with good time resolution is in itself superior to a sole background reduction of the standard mode. Anyhow, the electron tagger might in principle block some of the background that KATRIN has to deal with. However, a limited tagging efficiency can induce additional background by untagged electrons. Therefore the tagging rate and especially the tagging inefficiency of any 'electron tagger' needs to be minimized.

In contrast, a gated filter set-up produces a result roughly half as good as the standard mode as long as the TOF spectroscopic paradigm of measuring at only few and high retarding potentials is applied. The incorporation of further retarding potentials along with an optimization of the duty cycles separately for each retarding potential lifts the result at least in the same order of magnitude of the standard mode. It can not be excluded that a more dedicated parameter optimization could make the statistical sensitivity slightly exceed the $0.018 (\text{eV})^2$ of the standard mode, although there are no hints towards that direction. A reduction of the systematics seems possible, too. It will however be smaller, if compared to an electron tagger, due to the decreased spectral sensitivity at the higher retarding potentials. On the other hand, a gated filter set-up does not alter the energy of the electrons.² That might be the case with a tagger and add to the systematics there.

To summarize - while the simulations suggest significant improvements by TOF spectroscopic neutrino mass measurements in principle, it is not clear if an appropriate measurement method can be utilized to manifest that improvement.

6.2.1 General Constraints for Measurement Methods

The field of TOF spectroscopy is certainly also open for other types of measurements. The question can be stated in how far the above findings for the hypothetical tagging technique and the gated filter also allow conclusions on the capability of any other method. In general, all methods are in some way characterized by a timing performance, a detection efficiency and an influence on the background level. The tagger performance depended mostly on the efficiency and on the level of background suppression. In contrast, the gated filter suffers from an intrinsic anticorrelation of efficiency

²Losses by synchrotron radiation are negligible due to the adiabaticity of the MAC-E-Filter.

and timing: increasing the duty cycle reduces the duration of the off-periods and allows more electrons to pass but smears the start time even more.

Fig. 6.2 reflects the approximate requirements that any TOF measurement method has to fulfil in order to compete with the standard mode of operation. The statistical sensitivity is plotted as a function of a gaussian time uncertainty σ_t , which was already discussed for the hypothetical electron tagger and a simple overall efficiency ϵ acting just as a factor for the signal rate (4.15), without second-order effects as in the tagger discussion. The background has been modelled as simple additive constant without any further assumptions. The scenario of additional background by the TOF measurement (as possible with the tagger in case of bad characteristic) was neglected. Instead, the plot has been performed for the case of no background and for the case of the standard mode background $b_0 = 10$ mHz. As measurement time distribution the obtained optimum for 18565 V as lowest retarding potential (table 4.10) has been taken.

As far as a TOF measurement method can approximately be described by these general parameters, the plot gives an estimate of its statistical power. This is valid for instance for any tagger as well as the gated filter in case of constant duty cycles, but not for the gated filter with optimized duty cycles. It may be noted that the performance can be improved by a second measurement time optimization based on a model of the measurement method. The experience from the simulations suggests that the benefit of such a measurement time optimization is the greater the less sensitive the method is. However, such an optimization can not change the order of magnitude unless further parameters which extend the validity context of fig. 6.2 are taken into account.

6.3 Open Questions

Several points have still been left open. The study incorporated only the statistical uncertainty but not the systematic. It was mentioned that a reduction of the measurement time at low retarding potentials will likely reduce the systematics, though, it should be verified quantitatively. That might be accomplished in case of the usual KATRIN systematics. If the TOF measurement however entails its own yet unknown systematics, for instance a small but not negligible influence on the electron energies, the predictability will be limited.

In case a TOF measurement might get serious, a more accurate study with less approximations has to be performed. That would comprise especially the calculation of flight times for outer detector pixels. The beta spectrum and the field geometries should be taken into account with more details. Besides, deviations from the adiabaticity should be considered which would require full particle tracking. Such a precise calculation would not only be useful for TOF spectroscopy but also for integrity measurements by TOF with an electron gun [KAT12]

The results for the gated filter can not be taken as the last word since the parameter optimization has been somewhat exemplary by now. A more sophisticated optimization

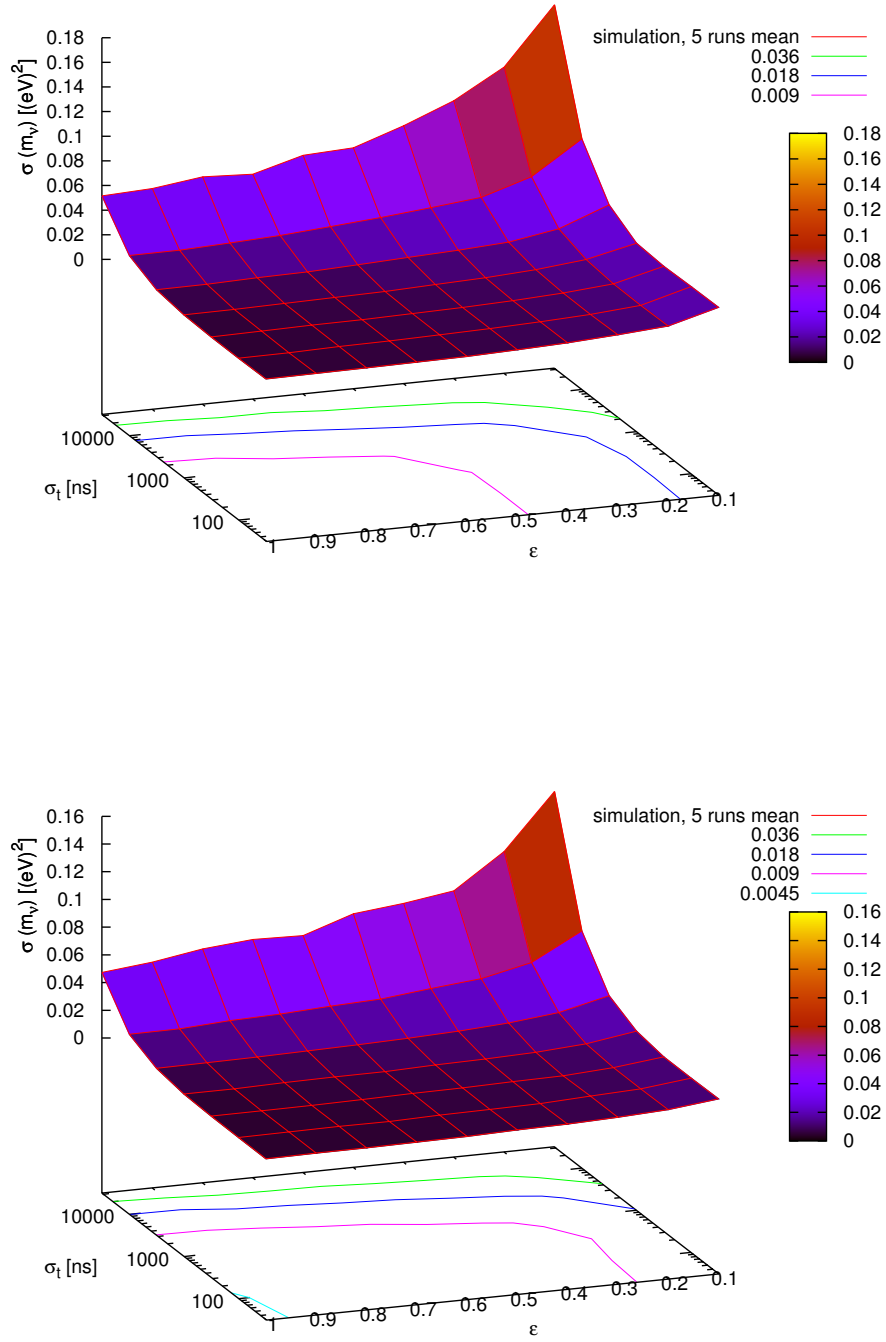


Figure 6.2: Statistical sensitivity on $m_{\nu_e}^2$ as a function of a gaussian time resolution σ_t and overall efficiency ϵ for 10 mHz constant additive background (above) and no background (below) for 3 years measurement time and a $(2/3, 1/3)$ measurement time distribution with retarding potentials -18565 V and -18560 V. The dark blue line reflects the reference value for the standard mode of KATRIN of $\sigma(m_{\nu_e})_{\text{stat}} = 0.018 (\text{eV})^2$.

routine would enhance the accuracy and could take into account more parameters like a higher number of retarding potentials, the period length t_r separately for each of these or more complex gate functions. However, the optimization of $\gtrsim 20$ parameters is by no means trivial. In first instance it is a matter of performance. A dedicated method for multi parameter minimizations like Markov Chain Monte Carlo (MCMC) [Sej11] should be utilized. In addition, the statistical spread of $\sigma_{\text{stat}}(m_{\nu_e}^2)$ that complicates the procedure needs to be reconsidered in the method.

While the gated filter method is probably too insensitive for neutrino mass measurements it might get interesting when it comes to new physics. A hot topic currently discussed is the possible existence of keV sterile neutrinos which are serious dark matter candidates [Veg11]. These might be visible by a kink in the spectrum some keV below the endpoint. The gated filter might resolve the low and medium energy range of the beta spectrum more precisely than an integrating spectrometer. In this region the count-rate is no issue. Yet, the integration would be disturbing since in every case the high energetic spectral parts contribute to the measured count-rate. Moreover, the DAQ would be unable to cope with the high event rate. So a reduction by a gate with a sharp timing allows a clear distinction between high and low energetic electrons and might enhance the sensitivity. That shows that spectral anomalies which are hidden in the integrating spectrometry might potentially get visible with TOF spectroscopy.

Bibliography

- [ALE89] ALEPH Collaboration, D. DeCamp et al., Phys. Lett. B **231**-4, 519 (1989). *Determination of the number of light neutrino species.* doi:10.1016/0370-2693(89)90704-1
- [ATL12] ATLAS Collaboration, Phys. Lett. B **710**, 49 (2012). *Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7\text{ TeV}$ with the ATLAS detector at the LHC.* doi:10.1016/j.physletb.2012.02.044
- [Bon99] J. Bonn, L. Bornschein, B. Degen, E.W. Otten, Ch. Weinheimer, Nucl. Instr. Meth. A **421**, 256 (1999). *A high resolution electrostatic time-of-flight spectrometer with adiabatic magnetic collimation.* doi:10.1016/S0168-9002(98)01263-7
- [Cle98] B. T. Cleveland et al., Astrophys. J. **496**, 505 (1998). *Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector.* doi:10.1086/305343
- [CMS12] CMS Collaboration, Phys. Lett. B **710**, 26 (2012). *Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7\text{ TeV}$.* doi:10.1016/j.physletb.2012.02.064
- [Cow98] Glen Cowan, *Statistical Data Analysis*. Oxford University Press (1998). Ch. 1, 6, 7.
- [DAY12] DAYA BAY Collaboration, F.P. An et al., Preprint. *Observation of electron-antineutrino disappearance at Daya Bay.* arXiv:1203.1669v2 [hep-ex]
- [Fog06] G.L. Fogli, E. Lisi, A. Marrone, A. Palazzo, Prog. Part. Nucl. Phys. **57**, 742 (2006). *Global analysis of three-flavor neutrino masses and mixings.* doi:10.1016/j.ppnp.2005.08.002
- [Fran95] B. Franzke et al., Phys. Scr. **1995**, 176 (1995). *Schottky mass spectrometry at the experimental storage ring ESR.* doi:10.1088/0031-8949/1995/T59/021
- [Frae10] F. Fränkle, Ph.D. Thesis, Karlsruher Institut für Technologie (KIT) (2010).
- [Giu07] C. Guinti, C.W. Kim, Fundamentals of Neutrino Physics. Oxford Unity Press (2007). Ch. 3, 6, 7.
- [Glu04] F. Glück, KATRIN internal document (2004). BSCW: *EMD - Programs - electric fields*
- [Glu06] F. Glück, KATRIN internal document (2006). BSCW: *EMD - Programs - magnetic fields*

- [Gol75] M. Goldhaber, L. Grodzins, A.W. Sunyar, Phys. Rev. **109**-3, 1015 (1975). *Helicity of Neutrinos*. doi:10.1103/PhysRev.109.1015
- [Ham10] J. Hamann, S. Hannestad, J. Lesgourgues, C. Rampf, Y.Y.Y. Wong, JCAP **1007**, 22 (2010). *Cosmological parameters from large scale structure - geometric versus shape information*. arXiv:1003.3999v2 [astro-ph.CO]
- [Hito] <http://hitoshi.berkeley.edu/neutrino>. As cited in [KAT04].
- [Hil11] B. Hillen, Ph.D. Thesis, Westfälische Wilhelms-Universität Münster (2011).
- [Jer10] M. Jerkins, J.R. Klein, J.H. Majors, F. Robicheaux, M.G. Raizen, New J. Phys. **12**, 43022 (2010). *Using Cold Atoms to Measure Neutrino Mass*. doi:10.1088/1367-2630/12/4/043022
- [KAT04] KATRIN Collaboration, FZKA Scientific Report 7090 (2004). *KATRIN Design Report*.
- [KAT12] KATRIN Collaboration, Internal Report (2012). *SDS Commissioning Measurements*. Revision 12332, ch. 7.2.
- [Kla04] H.V. Klapdor-Kleingrothaus et al., INFN Annual Report 2003 (2004). *Search For Neutrinoless Double Beta Decay With Enriched ^{76}Ge 1990-2003 – HEIDELBERG-MOSCOW-Experiment* arXiv:hep-ph/0404062v1
- [Kra05] Ch. Kraus et al., Eur. Phys. J. C **40**, 447 (2005). *Final Results from phase II of the Mainz Neutrino Mass Search in Tritium β Decay*. doi:10.1140/epjc/s2005-02139-7
- [Kuh62] T.S. Kuhn, *The Structure of Scientific Revolutions*. University of Chicago Press (1962), 2nd ed. 1970
- [LAG] LAGUNA homepage. http://www.laguna-science.eu/images/stories/standard_model_of_elementary2.gif
- [MAR06] MARE - Microcalorimeter Arrays for a Rhenium Experiment, proposal (2006)
- [Men11] G. Mention et al., Phys. Rev. D **83**, 73006 (2011). *The Reactor Antineutrino Anomaly*. arXiv:1101.2755 [hep-ex]
- [MIN] MINUIT 2 Homepage. <http://seal.web.cern.ch/seal/MathLibs/Minuit2/html/>
- [Ott08] E.W. Otten, C. Weinheimer, Rep. Prog. Phys **71**, 86201 (2008) *Reports on progress in physics: Neutrino mass limit from tritium β decay*. doi:10.1088/0034-4885/71/8/086201
- [Pau30] W. Pauli, *Collected Scientific Papers*. Eds. Kronig & Weisskopf, Vol. 2, 1316 - 1317.
- [Pre02] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C++*. Cambridge University Press (2002), 2nd ed. 2003. Ch. 3, 4, 5.
- [Rei59] R. Reines and C.L. Cowan, Phys. Rev. **113**, 273 (1959). *Detection of the Free Neutrino: a Confirmation*. doi:10.1126/science.124.3212.103

- [ROOT] ROOT Homepage. <http://root.cern.ch>
- [Sae20] A. Saenz, S. Jonsell, P. Froelich, Phys. Rev. Lett. **84**, 242 (2000). *Improved Molecular Final-State Distribution of HeT^+ for the β -Decay Process of T_2* . doi:10.1103/PhysRevLett.84.242
- [Sej11] A. Sejersén Riis, S. Hannestad, C. Weinheimer, e-Print (2011). *Analysis of KATRIN data using Bayesian inference*. arXiv:1105.6005v1 [nucl-ex]
- [Sim81] J.J. Simpson, Phys. Rev. D **23**, 649 (1981). *Measurement of the β -energy spectrum of ^3H to determine the antineutrino mass*. doi:10.1103/PhysRevD.23.649
- [Ste07] M. Steidel, KATRIN internal document, 2006
BSCW: Review of Detector Design SUMMER 07
- [SUP98] Super-Kamiokande Collaboration, Y. Fukuda et al., Phys. Rev. Lett. **81**, 1562 (1998). arXiv:hep-ex/9807003
- [Van93] R.S. VanDyck Jr., D.L. Farnham, P.B. Schwinberg, Phys. Rev. Lett. **70** (1993) 2888. *Tritium-helium-3 mass difference using the Penning trap mass spectroscopy*. doi:10.1103/PhysRevLett.70.2888
- [Val10] K. Valerius, Prog. Part. Nucl. Phys. **64**-2, 291 (2010). *The wire electrode system for the KATRIN main spectrometer*. doi:10.1016/j.ppnp.2009.12.032
- [Veg11] H.J. de Vega, O. Moreno, E. Moya de Guerra, M. Ramon Medrano, N. Sanchez, e-Print (2011). *Search of keV Sterile Neutrino Warm Dark Matter in the Rhenium and Tritium beta decays*. arXiv:1109.3452v2 [hep-ph]
- [Wei93a] C. Weinheimer, Ph.D. Thesis, Johannes Gutenberg-Universität Mainz (1993).
- [Wei93b] C. Weinheimer et al., Phys. Lett. B **300**, 210 (1993). *Improved limit on the electron-antineutrino rest mass from tritium β -decay*. doi:10.1016/0370-2693(93)90355-L
- [Wol08] I. Wolff, Diploma Thesis, Westfälische Wilhelms-Universität Münster (2008).
- [Wu57] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, R.P. Hudson, Phys. Rev. **105**, 1413 (1957). *Experimental Test of Parity Conservation in Beta Decay* doi:10.1103/PhysRev.105.1413
- [Zac09] M. Zacher, Diploma Thesis, Westfälische Wilhelms-Universität Münster (2009).

Eigenständigkeitserklärung

Ich versichere, dass ich diese Arbeit selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und Zitate kenntlich gemacht habe.

Acknowledgement

I'd like to thank everyone who supported me during, before, at or despite my work. This is in particular:

Hanna Sandberg
Johanna Steinbrink
Axel Steinbrink
Ursula Steinbrink
Thea Herweg
Fritz Herweg
Christa Müller

All current and former members
of the AG Weinheimer and associated.

Especially:
Jan D. Behrens
Michael Zacher
Daniel Winzen
Christopher Kranz
Volker Hannen
Matthias Prall
Anna Sejersen Riis
Björn Hillen
Miroslav Zboril
The Kaff0r Troops

For the second review:
Johannes Wessels

And last but not least:
Christian Weinheimer

I beg for pardon for everyone which I have forgot. Mail bombs may be sent to the Institut für Kernphysik, Room 213.