Fakultät für Physik

First Sub-Electronvolt Direct Neutrino Mass Measurement with the KATRIN Experiment

Christian Robert Karl

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:
apl. Prof. Dr. Norbert Kaiser

Prüfer der Dissertation:
Prof. Dr. Susanne Mertens
Hon.-Prof. Dr. Allen C. Caldwell

Die Dissertation wurde am 07.06.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 05.07.2022 angenommen.

Abstract

The mass of the neutrino is an important parameter in astroparticle physics and cosmology. While neutrino oscillation experiments have proven that neutrinos have mass, they are not sensitive to its absolute scale. The **Ka**rlsruhe **Tri**tium **N**eutrino (KATRIN) experiment is designed to measure the effective electron anti-neutrino mass m_ν with a sensitivity of $200\,\mathrm{meV}$ at $90\,\%$ confidence level (CL) using the kinematics of tritium β-decay. First data used for neutrino mass analysis was taken in spring 2019, with a second measurement phase following in autumn of the same year.

This thesis explains a novel approach to infer the neutrino mass confidence interval from the data based upon full Monte Carlo propagation of uncertainty. We show that our approach gives consistent results with the well-known nuisance parameter method in all cases currently used by the KATRIN collaboration.

Applying our Monte Carlo propagation approach to the data of the first two measurement campaigns, we find a combined best fit value of $m_\nu^2 = 0.12^{+0.32}_{-0.33}\,\mathrm{eV}^2$ including all relevant statistical and systematic uncertainties. This leads to the first sub-electronvolt laboratory measurement of the neutrino mass, limiting it to $m_\nu < 0.8\,\mathrm{eV}$ ($90\,\%$ CL).

In addition, we describe a future proof analysis method by approximating the KATRIN physics model with a neural network (NN). Our NN model shows no significant bias at the statistical sensitivity expected for KATRIN and is able to reproduce the results of analysing the first two measurement phases while reducing the computation time by several orders of magnitude. This makes it a promising approach to analyse future KATRIN neutrino mass data.

Kurzfassung

Die Masse des Neutrinos ist ein wichtiger Parameter in der Astroteilchenphysik und Kosmologie. Neutrinooszillationen zeigen zwar, dass Neutrinos nicht masselos sind, können aber die absolute Masse nicht messen. Das **Ka**rlsruhe **Tri**tium **N**eutrino (KATRIN) Experiment wurde konstruiert, um die effektive Elektron-Antineutrino-Masse m_ν mit einer Sensitivität von $200\,\mathrm{meV}$ bei einem Konfidenzlevel (KL) von $90\,\%$ zu messen. Dafür werden die kinematischen Eigenschaften des β-Zerfalls von Tritium genutzt. Erste Daten für die Neutrinomassen-Analyse wurden im Frühjahr 2019 erfasst. Im Herbst desselben Jahres fand eine zweite Messphase statt.

Diese Arbeit erklärt einen neuen Ansatz, um das Neutrinomassen-Konfidenzintervall aus den Daten zu bestimmen, welcher auf der vollen Monte Carlo Fortpflanzung von Unsicherheiten basiert. Wir zeigen, dass unser Ansatz konsistente Ergebnisse mit der bekannten *nuisance parameter* Methode liefert. Dies gilt für alle Fälle, die aktuell von der KATRIN Kollaboration in der Analyse verwendet werden.

Wenn wir unseren Monte Carlo basierten Analyse-Ansatz auf die Daten der ersten beiden Messkampagnen anwenden, ergibt sich einen Bestwert von $m_\nu^2 = 0.12^{+0.32}_{-0.33}\,\mathrm{eV}^2$ unter Berücksichtigung aller relevanter statistischer und systematischer Fehler. Daraus resultiert die erste Labor-basierte Messung der Neutrinomasse unter einem Elektronvolt, mit dem entsprechenden Limit von $m_\nu < 0.8\,\mathrm{eV}$ ($90\,\%$ KL).

Zusätzlich erläutern wir einen zukunftssicheren Analyseansatz, bei dem das KATRIN Physikmodell mit einem neuronalen Netz (NN) approximiert wird. Unser NN Modell zeigt keinen signifikanten Bias bei der erwarteten statistischen Sensitivität von KATRIN und kann die Ergebnisse der Analyse der ersten beiden Messkampagnen reproduzieren. Dabei wird die Rechenzeit um mehrere Größenordnungen reduziert. Dies macht unseren Ansatz vielversprechend für die Analyse zukünftiger KATRIN Neutrinomassen-Daten.

Contents

Chapter 1

Introduction

The neutrino is the second most abundant elementary particle in our universe after the photon. As such, it plays an important role in cosmology, especially in the field of structure formation in the universe. Its peculiar elementary properties also make it an interesting research topic in particle physics. Thus, the field of neutrino physics is a highly active one in astroparticle physics, bridging the smallest and the largest scales known.

Wolfgang Pauli first postulated the existence of the neutrino in 1930 to explain the continuous electron energy spectrum observed in β-decay, in contrast to the sharp peaks seen in α- and γ-decays. As the postulated particle does not interact electromagnetically or with the strong nuclear force, the task of discovering experimentally was extremely challenging and could only be fulfilled more than twenty years later by Cowan and Reines in 1956.

Over the past 60 years, our understanding of the neutrino has improved substantially, and it is now well embedded in the standard model (SM) of particle physics. We know that it comes in three flavours, each as a partner to the one of the three charged leptons (electron, muon, tau). In addition, neutrino oscillations have shown that it is possible for neutrinos to change their flavour with a probability that depends on their energy and distance travelled. This oscillation is only possible if each flavour eigenstate is a quantum-mechanical superposition of three independent mass eigenstates, with a different mass value for each mass eigenstate. Therefore, we can conclude that at least two neutrino mass eigenstates must be non-zero and thus that neutrinos are massive particles.

However, while neutrino oscillations are sensitive to the mass difference of the mass eigenstates and can thus provide a lower bound on neutrino mass, they are not sensitive to the absolute scale and cannot provide an upper bound. Measuring this absolute mass scale has been a major effort of the past decades, with three independent approaches leading the way: Cosmological measurements are sensitive to the sum of the three neutrino masses, for example via their impact on structure formation in the universe. Neutrinoless double beta decay ($0\nu\beta\beta$) experiments aiming to prove that the neutrino is its own anti-particle, i.e. a Majorana

particle, can relate the half-life of 0νββ-decay to the neutrino mass value. Finally, it is possible to infer the neutrino mass directly from the kinematics of β-decay via its shape impact on the electron energy spectrum at highest energies. Combining the results of the three approaches, we can safely conclude that the mass scale of the neutrino is at or below the electronvolt scale, more than five orders of magnitude lighter than the second lightest particle, the electron.

The currently leading experiment to measure m_ν using the kinematics of β-decay is the **Ka**rlsruhe **Tri**tium **N**eutrino (KATRIN) experiment built at the Karlsruhe Institute of Technology (KIT). It is designed to measure the effective electron anti-neutrino mass m_ν with a sensitivity of $200\,\mathrm{meV}$ at $90\,\%$ CL. To achieve its goal, it combines a highly luminous windowless gaseous tritium source with the same measurement principle of its predecessors in Mainz and Troitsk: magnetic adiabatic collimation with an electrostatic filter (MAC-E filter). This setup allows energy resolutions on the electronvolt-scale when measuring an integral β-spectrum.

Combining the electronvolt-scale resolution with the high statistical power of the KATRIN experiment, we can see that it requires a very precise model description and data analysis framework to infer m_ν from the measurement. This includes the description of the β-decay, the experimental response as well as any statistical and systematic uncertainties. Implementing such an analysis framework, applying it to the first two measurement campaigns to infer the confidence interval of m_ν, and making it future proof for upcoming neutrino mass campaigns is the task of this thesis.

This thesis is structured in the following manner: Chapter 2 gives a general introduction to neutrino physics with a focus on the task of measuring the absolute neutrino mass scale. Next, chapter 3 explains the KATRIN experiment in detail, derives a model of the expected count rate and describes the sources of systematic uncertainty. Chapter 4 then focuses on the analysis strategies applied to infer the neutrino mass from the data. As such, it describes the general KATRIN likelihood function, methods for data combination, parameter inference, and uncertainty treatment. In chapter 5 we focus on a novel analysis method developed over the course of this thesis based upon Monte Carlo propagation of uncertainty, benchmark it against the nuisance parameter approach, and describe its advantages and disadvantages compared to other approaches pursued by the KATRIN collaboration. We then make use of this analysis method in chapter 6 to analyse the first two measurement campaigns, leading to the first sub-electronvolt direct neutrino mass measurement. Chapter 7 then explains our novel neural network based analysis approach developed within the frame of this thesis, its accuracy and computational improvements, and proving that it provides valid approach for future neutrino mass analysis. Finally, we recapitulate our results in chapter 8.

Chapter 2

Neutrino physics

Over the past century, understanding the properties of the neutrino has been a major effort in particle physics. This chapter gives an overview of the history, current standing and remaining unknowns of neutrino physics.

2.1 Discovery of the neutrino

The history of the neutrino begins with its postulation by Wolfgang Pauli in 1930 [1]. To explain the continuous energy spectrum of the electron emitted in β-decay, see fig. 2.1, Pauli proposed an additional electrically neutral, spin-$\frac{1}{2}$ particle. This particle is emitted together with the electron and the decay energy is shared between the two as well as the much heavier daughter nucleus. This energy sharing in a three body decay solves the mystery of the continuous electron energy spectrum.

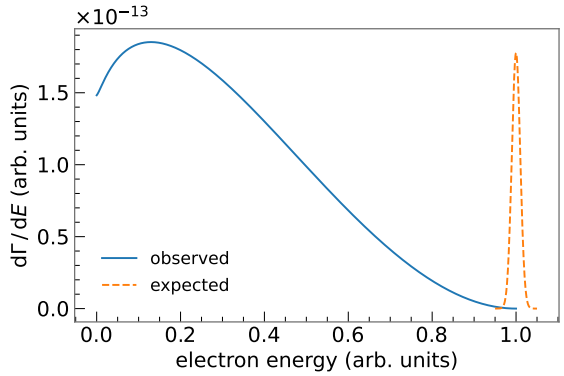


Figure 2.1: Electron energy spectrum of β-decay. The observed continuous spectrum (blue solid line) is in contradiction to the narrow peak expected from a two body decay (orange dashed line).

However, already Pauli described that it is challenging, perhaps impossible, to detect such a particle experimentally. Thus, it should take more than twenty years until the existence of the neutrino is confirmed by Cowan and Reines in 1956 [2]. In their experiment, they dissolved approximately $40\,\mathrm{kg}$ of cadmium chloride in two large water tanks with a total volume of roughly $200\,\mathrm{l}$ to detect the incoming (electron) anti-neutrino flux from the nearby Savannah river plant. These incoming anti-neutrinos can react with protons in the water via inverse β-decay,

$$\bar{\nu}_e + p^+ \rightarrow n + e^+, \tag{2.1}$$

creating a neutron n and a free positron e^+. The positron instantly annihilates with an electron creating two coincident γ-rays. These γ-rays are absorbed by a liquid scintillator whose emitted light is subsequently detected by photo multiplier tubes (PMTs). The neutron is captured by the dissolved cadmium several microseconds later, giving off a third detectable γ. This unique coincidence of the two γ-rays from e^+e^- annihilation, followed by a single γ from the neutron capture is the clear signature of incoming $\bar{\nu}_e$ particles.

Only six years later in 1962 a neutrino flavour corresponding to the muon was discovered by Lederman, Schwartz, and Steinberger at the Brookhaven National Laboratory [3] proving the doublet structure of leptons: each charged lepton (e, μ) has a corresponding uncharged neutrino (ν_e, ν_μ). With the discovery of the τ lepton in 1975 [4], it was clear that there should also be a third neutrino flavour. This third neutrino was experimentally detected in 2000 by the DONUT collaboration [5].

This discovery completes the leptonic part of the current standard model (SM) of particle physics: three charged leptons (e, μ and τ) as well as the corresponding neutrinos (ν_e, ν_μ and ν_τ). The charged leptons participate in the electromagnetic and weak interaction, while the neutrinos only interact weakly. This separates the leptons from the other fermionic spin-$\frac{1}{2}$ particles in the SM, the quarks, which also interact via the strong force. The SM is completed by the force carrier bosons: the gluon (strong interaction), photon (electromagnetism), Z- and W-boson (weak interaction) as well as the Higgs-boson (Higgs field). A summary of all standard model particles is shown in fig. 2.2. For a full theoretical description of the SM we refer to [6].

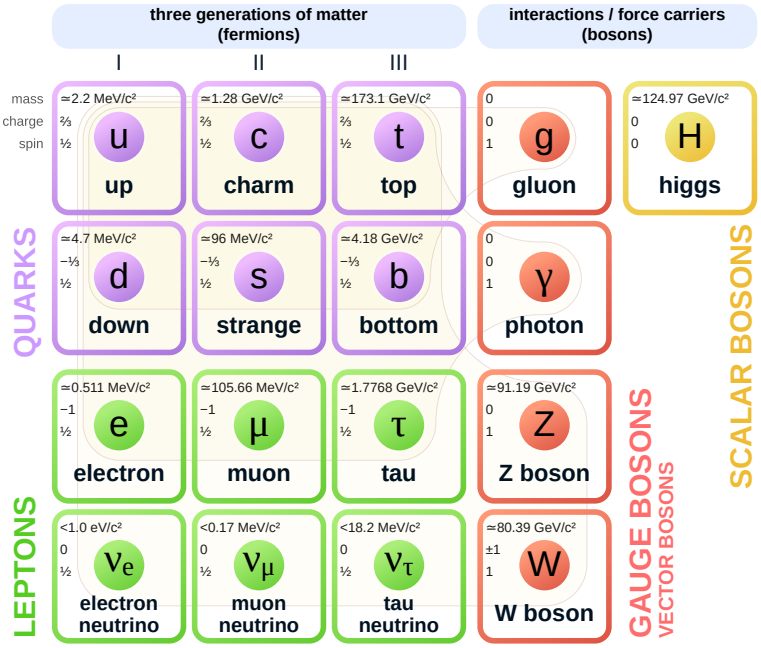


Figure 2.2: Standard model of particle physics [7]. Particles are split into matter fermions (left) and force carrier bosons (right). The fermions are further split into quarks (top) and leptons (bottom). This thesis focuses on the uncharged fermions, the neutrinos, namely the investigation of their absolute mass scale.

2.2 Neutrino oscillations

While the discovery of the electron neutrino solved the puzzle of the continuous energy spectrum of the electrons emitted in β-decay, another question soon arose in the neutrino sector. The Homestake experiment, designed in the late 1960s to measure the flux of neutrinos coming from nuclear fusion in the sun, detected only one third of the expected flux [8, 9]. This discrepancy of measurement compared to the theoretical predictions of the standard solar model (SSM) became known as the *solar neutrino problem*.

While a common opinion was that the SSM predictions were wrong, another solution involves the phenomenon of *neutrino oscillations* first predicted by Pontecorvo in the late 1950s [10]. This phenomenon arises from the mixing of neutrino flavour and neutrino mass eigenstates, i.e. a neutrino of a specific flavour α (e, μ, τ) is a quantum-mechanical superposition of mass eigenstates i (1, 2, 3). The mixing is parametrized by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix $U_{\alpha i}$:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}. \tag{2.2}$$

While neutrinos interact weakly in their flavour eigenstate ν_α, they propagate space as plane waves in their mass eigenstates ν_i. Thus, we can write their time propagation as

$$|\nu_i(t)\rangle = e^{-i(E_i \cdot t - \vec{p}_i \cdot \vec{x})} \cdot |\nu_i(0)\rangle \tag{2.3}$$

with their energy E_i, the evolved time t, their momentum $\vec{p}_i$ and the position change $\vec{x}$. Using $|\vec{p}_i| \gg m_i$ and $t \approx L$, we can rewrite

$$E_i \cdot t \approx \left(E + \frac{m_i^2}{2E} \right) \cdot L \tag{2.4}$$

as well as

$$\vec{p}_i \cdot \vec{x} \approx E \cdot L \tag{2.5}$$

and thus

$$|\nu_i(t)\rangle = e^{-i \frac{m_i^2}{2} \cdot \frac{L}{E}}. \tag{2.6}$$

The probability of a neutrino created in state α with energy E to be detected in the state β after travelling the distance L is then given by

$$P_{\alpha \to \beta} = |\langle \nu_\beta(L) | \nu_\alpha \rangle|^2 . \tag{2.7}$$

Let us now discuss the reduced case of only two neutrino states for simplicity. In this case, we can write the mixing matrix U as

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{2.8}$$

with a single mixing angle θ. Therefore, we have

$$|\nu_\alpha\rangle = \cos\theta\,|\nu_1\rangle + \sin\theta\,|\nu_2\rangle \tag{2.9}$$

$$|\nu_\beta(L)\rangle = -\sin\theta\,|\nu_1(L)\rangle + \cos\theta\,|\nu_2(L)\rangle\,, \tag{2.10}$$

and we can calculate the oscillation probability using eq. (2.7) to be

$$P_{\alpha\to\beta} = 2\sin^2\theta\cdot\sin^2\left(\frac{\Delta m^2_{21}L}{4E}\right) \tag{2.11}$$

with the squared mass difference $\Delta m^2_{21} = m^2_2 - m^2_1$.

This simplified two neutrino approach is often a valid approximation, and already shows us three important features of the oscillation:

1. The probability is only non-zero if the masses m_1 and m_2 are different. Thus, at least one of the two must have non-zero mass in case an oscillation is observed.

2. A non-zero mixing angle θ is needed in addition to the mass difference.

3. The oscillation frequency depends on $\frac{L}{E}$, i.e. the distance travelled and the neutrino energy. This property can be used when designing experiments to detect neutrino oscillations.

When investigating the three neutrino case, more parameters than a single mixing angle θ and mass splitting Δm^2_{21} are required:

- there are three mixing angles θ_{12}, θ_{23}, and θ_{13},
- two mass splittings Δm^2_{21}, and Δm^2_{31} as well as
- a CP-violating phase δ_{CP}.

For a full derivation of the three neutrino case we refer to [11]. However, the important fact that the oscillation probability is only non-zero if $\Delta m^2_{ij} = m^2_i - m^2_j$ is non-zero still holds.

In 1998 the Super-Kamiokande (Super-K) collaboration was the first to report evidence for the oscillation of so-called *atmospheric neutrinos* [12]. These neutrinos are produced in hadronic showers initiated by the collision of cosmic rays with particles in the atmosphere. Super-K observed a discrepancy of the incoming flux of muon neutrinos between upcoming ν_μ which had travelled through the earth and ν_μ coming directly from the atmosphere. The measured flux difference is consistent with two-flavour $\nu_\mu \to \nu_\tau$ oscillation, and therefore the first evidence for non-zero neutrino mass.

Trying to solve the solar neutrino problem, the Sudbury Neutrino Observatory (SNO) experiment [13] took a closer look at the solar neutrino flux using a $1000\,\text{t}$ heavy-water (D_2O)

Cherenkov [14] detector. The experiment was designed to be able to discriminate the incoming electron neutrino flux as well as the total neutrino flux. This was done using the following interactions

$$\text{charged current (CC):}\ \nu_e + \mathrm{D} \rightarrow p + p + e^- \tag{2.12}$$

$$\text{neutral current (NC):}\ \nu_\alpha + \mathrm{D} \rightarrow p + n + \nu_\alpha \tag{2.13}$$

$$\text{electron elastic scattering (ES):}\ \nu_\alpha + e^- \rightarrow \nu_\alpha + e^- \tag{2.14}$$

where the CC interaction is only sensitive to the ν_e flux, while the NC and ES interactions are sensitive to the total neutrino flux. The CC interaction is only sensitive to electron neutrinos, as the charged lepton is produced in the process and solar neutrinos have energies below the mass of the muon and the tau. In their data they observed a deficit in the electron neutrino flux just like the Homestake experiment, but no deficit in the total neutrino flux. Thus, the SSM predictions are correct, however the electron neutrinos change flavour on their way to earth. For completeness, we would like to mention that the flavour change of solar neutrinos observed by SNO is dominated by the Mikheyev–Smirnov–Wolfenstein (MSW) effect [15, 16] which modifies the oscillation probability of neutrinos travelling through matter of varying density.

Over the past years the three mixing angles as well as the absolute value of the mass splittings have been measured precisely by various experiments. What remains to be determined in the oscillation space is the sign of Δm^2_{31} which defines the mass hierarchy ($m_3 > m_1$: normal ordering, $m_3 < m_1$: inverted ordering) as well as whether the CP-violating phase δ_{CP} is non-zero.

2.3 Measurement of the absolute neutrino mass scale

While neutrino oscillations have proven that at least two neutrino mass eigenstates have non-zero rest mass, they cannot asses the absolute mass scale, but only the squared differences of the mass eigenstates Δm^2_{ij}. This chapter briefly describes the various approaches currently taken to answer this question and discusses their advantages and disadvantages.

2.3.1 Cosmology

In cosmology the evolution of the universe since short after the Big Bang can be modeled by the Lambda cold dark matter (ΛCDM) model. In ΛCDM, our universe consists of three major components: dark energy described by the cosmological constant Λ, cold dark matter (CDM), which is heavily searched for experimentally, and regular matter.

As neutrinos are the second most abundant particles in the universe, they play an important role on cosmological scales. Cosmological measurements are sensitive to the energy density of neutrinos Ω_ν as well as the number of active flavours. The energy density of neutrinos is then proportional to the sum of neutrino masses

$$m_\Sigma = \sum_i m_i. \tag{2.15}$$

Accurate measurements of the total neutrino mass involve their imprint on the cosmic microwave background (CMB) as well as on structure formation in the early universe [17]. Due to their relatively low mass, neutrinos behave as hot dark matter (HDM) and dilute structures below their free streaming length. If all dark matter was HDM (m_Σ is large), one therefore expects a top-down generation of the universe: galaxy clusters form before galaxies, in contrast to observation where galaxies are seen to be older than clusters.

The most precise estimations come from combining multiple cosmological measurements and assuming base ΛCDM, limiting $m_\Sigma < 0.12\,\mathrm{eV}$ ($95\,\%$ CL) [18]. When extending the ΛCDM model by five additional parameters, thus significantly increasing the freedom of the cosmological model, the upper limit loosens to $m_\Sigma < 0.515\,\mathrm{eV}$ ($95\,\%$ CL) [19].

2.3.2 Neutrinoless double beta-decay

Another measurement sensitive to the absolute neutrino mass is neutrinoless double β- decay (0νββ). In regular two neutrino double β^--decay (2νββ), a nucleus X decays into a daughter nucleus Y^{2+}, two electrons and two electron anti-neutrinos:

$$X \to Y^{2+} + 2e^- + 2\bar{\nu}_e. \tag{2.16}$$

In case the neutrino is its own anti-particle, a so-called Majorana particle, it is possible for the two neutrinos to cancel one another within the decay process and they do not show up in the final state:

$$X \to Y^{2+} + 2e^-. \tag{2.17}$$

In this case, the two electrons share almost the complete decay energy, with the much heavier daughter nucleus only receiving a small recoil. This energy peak can be searched for experimentally as signature of 0νββ. In case observed, the decay rate (inverse of the lifetime) $\Gamma_{0\nu\beta\beta}$ can be related to the effective Majorana neutrino mass $m_{\beta\beta}$ via

$$\Gamma_{0\nu\beta\beta} = \frac{1}{T_{0\nu\beta\beta}} = G_{0\nu} \cdot |M_{0\nu}|^2 \cdot |m_{\beta\beta}|^2 \tag{2.18}$$

with the phase space factor $G_{0\nu}$ and the nuclear matrix element $M_{0\nu}$. The effective Majorana neutrino mass is defined as the coherent sum of the neutrino mass eigenstates i over the PMNS

matrix elements U_{ei}:

$$|m_{\beta\beta}| = \left| \sum_{i=1}^{3} U_{ei}^2 \cdot m_i \right| . \tag{2.19}$$

This simple relation between $m_{\beta\beta}$ and $T_{0\nu\beta\beta}$ only holds in the *standard scenario* in which the decay is mediated by a light Majorana neutrino.

Therefore, if 0νββ is observed:

- the neutrino is a Majorana particle, and
- the absolute neutrino mass scale can be inferred from the measured lifetime. However, the exact relation depends on the mediator model.

This unique combination of information that can be derived from observing 0νββ makes it a heavily pursued effort in neutrino physics. Up to date the process has not been measured and therefore only lower limits on the lifetime (upper limits on $m_{\beta\beta}$) have been set. The leading limits of the corresponding experiments and their used isotopes are:

- $m_{\beta\beta} < 0.079\,\mathrm{eV}$ to $0.180\,\mathrm{eV}$ ($90\,\%$ CL) from the GERDA collaboration [20] using ^{76}Ge. LEGEND [21] will continue this search, building on the expertise of its predecessors GERDA as well as the MAJORANA demonstrator [22].
- $m_{\beta\beta} < 0.075\,\mathrm{eV}$ to $0.350\,\mathrm{eV}$ ($90\,\%$ credible interval (CI)) from the Cuore collaboration [23] using ^{130}Te.
- $m_{\beta\beta} < 0.036\,\mathrm{eV}$ to $0.156\,\mathrm{eV}$ ($90\,\%$ CL) from the KamLAND-Zen collaboration [24] using ^{136}Xe. This effort will be continued in KamLAND2-Zen [25].

The large interval of each $m_{\beta\beta}$ limit comes from uncertainties on the calculation of the nuclear matrix element and will be discussed further in section 2.3.4.

2.3.3 Electron energy spectrum of beta-decay

The final and most direct option to measure the absolute mass scale of the neutrino we would like to discuss investigates the electron energy spectrum of β-decay. We already saw in fig. 2.1 that this energy spectrum led to the postulation of the neutrino in the first place. We will now take a closer look at this process and how to measure the neutrino mass using it.

In a β^--decay a nucleus X decays into a daughter nucleus Y, an electron e^- and an electron anti-neutrino $\bar{\nu}_e$:

$$X \to Y^+ + e^- + \bar{\nu}_e. \tag{2.20}$$

The amount of energy released in the decay Q is shared between the decay products. A low recoil energy E_{rec} is passed to the much heavier daughter nucleus. In the region of electrons with highest energy this recoil energy can be considered constant, and the remaining so-called endpoint energy

$$E_0 = Q - E_{\text{rec}} = E + E_\nu \tag{2.21}$$

is split between the electron (E) and the neutrino (E_ν). This energy also corresponds to the maximal energy the electron can receive if the neutrino is massless.

The differential decay rate depending on the electron energy can be written as

$$\begin{aligned}\frac{\mathrm{d}\Gamma}{\mathrm{d}E} =& \frac{G_{\text{F}}^2 \cos\theta_{\text{C}}^2}{2\pi^2} \cdot |M_{\text{nuc}}|^2 \cdot F(Z', E) \cdot p \cdot (E + m_e) \cdot (E_0 - E) \cdot \\ & \cdot \sum_i |U_{ei}|^2 \cdot \sqrt{(E_0 - E)^2 - m_i^2} \cdot \Theta(E_0 - E - m_i)\end{aligned} \tag{2.22}$$

with the Fermi constant G_{F}, the Cabibbo angle θ_{C}, the nuclear matrix element M_{nuc}, the Fermi function $F(Z', E)$ which accounts for the electromagnetic interaction of the outgoing electron with the daughter nucleus, and the Heaviside function Θ which ensures energy conservation [26]. The Fermi function in its classical form is given by

$$F(Z', E) = \frac{2\pi\eta}{1 - \exp(-2\pi\eta)} \tag{2.23}$$

with the atomic charge of the daughter nucleus Z' and the Sommerfeldt parameter

$$\eta = \frac{\alpha Z'}{\beta} \tag{2.24}$$

defined by Z', the finestructure constant α and the relativistic β factor. A commonly used approximation of its fully relativistic form is

$$F(Z', E) \cdot \left(1.002037 - 0.001427 \frac{p \cdot m_e}{E + m_e}\right) \tag{2.25}$$

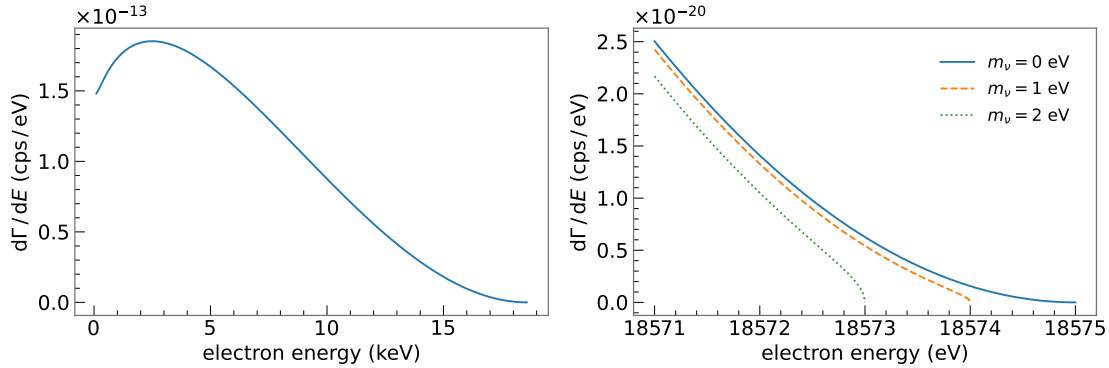
using the classical $F(Z', E)$ [27].

In principle this decay rate is a superposition of three functions: one for each neutrino mass eigenstate with mass m_i. However, in practice experiments are not sensitive to the mass splitting, and only measure the effective electron (anti-) neutrino mass,

$$m_\nu = \left(\sum_{i=1}^{3} |U_{ei}|^2 \cdot m_i^2\right)^{\frac{1}{2}}, \tag{2.26}$$

which we denote with m_ν in this thesis instead of the also often used m_β. This simplifies eq. (2.22) to

$$\begin{aligned}\frac{\mathrm{d}\Gamma}{\mathrm{d}E} =& \frac{G_{\text{F}}^2 \cos\theta_{\text{C}}^2}{2\pi^2} \cdot |M_{\text{nuc}}|^2 \cdot F(Z', E) \cdot p \cdot (E + m_e) \cdot (E_0 - E) \cdot \\ & \cdot \sqrt{(E_0 - E)^2 - m_\nu^2} \cdot \Theta(E_0 - E - m_\nu).\end{aligned} \tag{2.27}$$



(a) full energy range

(b) endpoint region

Figure 2.3: Electron energy spectrum of tritium β-decay over the full energy range (a) and zoomed into the endpoint region (b). The neutrino mass manifests itself as a spectral distortion near the endpoint.

This energy dependence of the differential decay rate is shown over the full energy range in fig. 2.3a and for different values of m_ν in the endpoint region in fig. 2.3b. We can identify a clear shape effect of the neutrino mass in the changing curvature of the spectrum near the endpoint. This is the signal experiments measuring m_ν are looking for.

The currently leading experiment in this field is the **Ka**rlsruhe **Tri**tium **N**eutrino (KATRIN) experiment which is designed to measure m_ν with a sensitivity of $200\,\mathrm{meV}$ at $90\,\%$ CL. We will discuss KATRIN in detail in chapter 3. An upcoming experiment using the β-spectrum of atomic tritium is Project 8 [28]. They make use of cyclotron radiation emission spectroscopy (CRES) to measure the electron energy spectrum and target a sensitivity of $40\,\mathrm{meV}$ ($90\,\%$ CL). Another related kinematic approach to infer m_ν makes use of electron capture instead of β-decay. In this field, the electron capture in holmium (ECHo) experiment using ^{163}Ho and a calorimetric energy measurement is leading, targeting a sub-electronvolt sensitivity [29].

2.3.4 Comparison and synergy of different measurements

First and foremost we would like to mention that the three different approaches to assess the absolute neutrino mass scale measure different parameters (m_Σ, $m_{\beta\beta}$, m_ν) and rely on different models and should therefore be seen as complementary. A summary of the parameter space for these different neutrino mass observables, along with current and possible constraints from the corresponding experiments, is shown in fig. 2.4. Here, the blue and orange regions are the allowed regions from the best-fit values of a global fit to the oscillation data [30, 31].

Starting with cosmology, we can see that the current measurements are extremely sensitive, especially when assuming the base ΛCDM model. Here, the result is already close to excluding

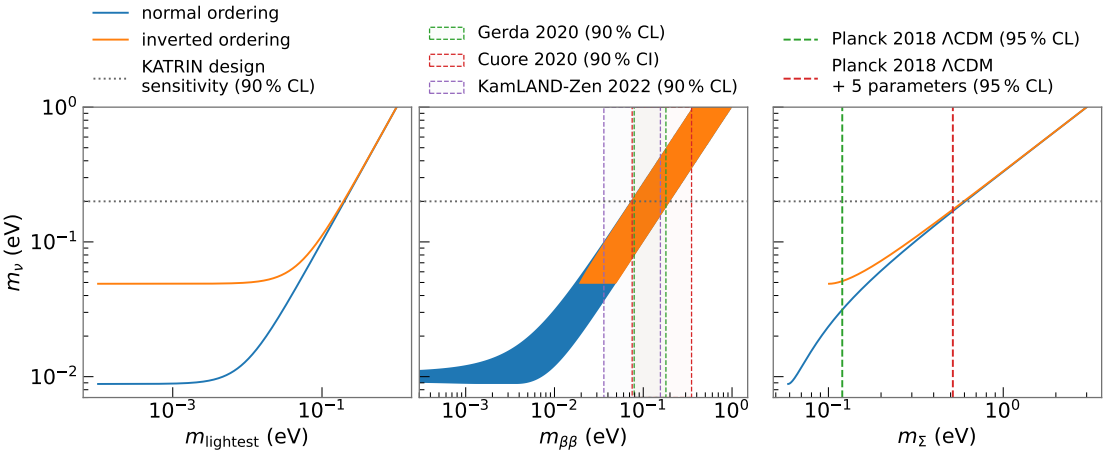


Figure 2.4: Summary of neutrino mass values from the different measurement principles. The orange and blue regions come from the best-fit values of a global fit to the oscillation data [30, 31]. The gray-dotted horizontal line is the design sensitivity of the KATRIN experiment which would exclude the region above. The vertical dashed lines are current upper limits from double beta decay [20, 23, 24] and cosmology [18, 19], each excluding the region to the right of the line.

the inverted ordering, where $m_\Sigma \gtrsim 0.1\,\mathrm{eV}$ is implied by the oscillation data. In addition, cosmological measurements do not rely on the Majorana or Dirac nature of the neutrino. However, especially at highest sensitivities, the results depend on the underlying cosmological model (ΛCDM) and parametrization. This reliance on a cosmological model and the interplay of scales of completely different order of magnitude arguably makes cosmology the most indirect approach to asses the neutrino mass scale.

Next, we observe that 0νββ is also quite sensitive to the mass scale, with the most stringent upper limits approaching the exclusion of inverted mass ordering. However, we also see that the allowed parameter space from oscillation is quite large and it is even possible for $m_{\beta\beta}$ to be zero in the normal ordering due to cancellations by the complex CP- and Majorana-phases. The measured upper limits in $m_{\beta\beta}$ are given as intervals, due to the described uncertainties on the nuclear matrix element. In addition, by definition, 0νββ relies on the Majorana nature of the neutrino. Therefore, neutrinoless double β-decay can be seen primarily as a method to infer the neutrino nature, with $m_{\beta\beta}$ as secondary parameter.

Finally, we see that inferring m_ν from regular β-decay is less sensitive than the other approaches, with even the design sensitivity of KATRIN not approaching the possibility to discriminate between the mass ordering. However, it can be seen as the most direct method as it does not require a cosmological model and does not rely on the neutrino nature.

Chapter 3

The KATRIN experiment

The KATRIN experiment is the currently leading direct neutrino mass measurement using the kinematics of β-decay. It is designed to measure the effective electron anti-neutrino mass

$$m_\nu = \left(\sum_i |U_{ei}|^2 \cdot m_i^2 \right)^{\frac{1}{2}} \tag{3.1}$$

with a sensitivity of $200\,\mathrm{meV}$ at $90\,\%$ CL using the same measurement principle as its predecessors in Mainz [32] and Troitsk [33].

This chapter explains this measurement principle [26] and the experimental setup of KATRIN [34, 35]. Afterwards, a model for the expected integral β-spectrum is derived [27] and the various sources of systematic uncertainty are explained.

3.1 Measurement principle: MAC-E filter spectroscopy

The electron energy spectrum depicted in fig. 2.3 shows two of the main properties an experiment measuring the neutrino mass from β-decay must fulfil. Due to the extremely low count rate in the endpoint region where the effect of nonzero neutrino mass is most visible, a highly luminous source is required. The emitted high-energy electrons must then be measured with an energy resolution on the electronvolt-scale to resolve the small neutrino mass imprint.

To this end, the electrons are guided from the source to an electrostatic filter with magnetic adiabatic collimation (MAC-E filter) as shown in fig. 3.1. In the filter, the electrons run against an applied retarding voltage U with the corresponding retarding energy qU where $q = -1\,e$ denotes the electric charge of the electron. Only electrons with a kinetic energy parallel to the electric field lines $E_\parallel$ larger than qU pass this filter. These electrons are subsequently

guided to the detector and counted. Varying the applied retarding energy results in an integral measurement of the β-spectrum.

As the retarding energy only filters out electrons depending on $E_{\parallel}$, it must be guaranteed that the total kinetic energy of the electron $E_{\text{kin}} = E_{\parallel} + E_{\perp} \approx E_{\parallel}$. The residual energy perpendicular to the retarding energy $E_{\perp}$ defines the energy resolution of the MAC-E filter.

To optimize the energy resolution, a specific magnetic field configuration is required. The electrons emitted in the source are guided adiabatically, which including relativistic effects means under conservation of

$$\frac{p_{\perp}^2}{B} = \frac{E_{\perp}(\gamma+1)\cdot m_e}{B} = \text{const} \tag{3.2}$$

along the beamline with the relativistic gamma factor $\gamma = \frac{E+m_e}{m_e}$, the varying magnetic field B and the momentum (energy) perpendicular to the electrostatic potential $p_{\perp}$ ($E_{\perp}$). Defining the pitch angle θ, we can write $E_{\perp} = E\sin^2\theta$ and correspondingly $E_{\parallel} = E\cos^2\theta$.

One can thus decrease $E_{\perp}$ by decreasing the magnetic field strength B. The optimal configuration is to have the lowest magnetic field exactly in the plane with maximum retarding energy qU where the electrons are filtered. This plane is called the analyzing plane with the corresponding magnetic field B_{ana}. After passing the analyzing plane, electrons are focused onto the detector with an increased B_{det}. Using the conservation condition defined in eq. (3.2), we can calculate the energy resolution ΔE for a specific electron energy E as

$$\Delta E = E\cdot\frac{B_{\text{ana}}}{B_{\text{src}}}\frac{\gamma+1}{\gamma_{\text{ana}}+1}\sin^2\theta_{\text{max}} \tag{3.3}$$

with the magnetic field in the source where the electrons are emitted B_{src} and the maximum acceptance angle θ_{max}.

In principle it is possible to set the maximum magnetic field B_{max} in the source where the electrons are emitted. However, this leads to the acceptance of electrons with all angles, especially those with very large angles and thus long travelled distances in the source. This is unfavourable due to increased scattering with the tritium gas in the source. Therefore, the maximum magnetic field is set after the source and electrons with large angles are reflected by the magnetic mirror effect. The maximum acceptance angle

$$\theta_{\text{max}} = \arcsin\sqrt{\frac{B_{\text{src}}}{B_{\text{max}}}} \tag{3.4}$$

is determined by these magnetic fields. Therefore, the energy resolution of the MAC-E filter is given by

$$\Delta E = E\cdot\frac{B_{\text{ana}}}{B_{\text{max}}}\frac{\gamma+1}{\gamma_{\text{ana}}+1}. \tag{3.5}$$

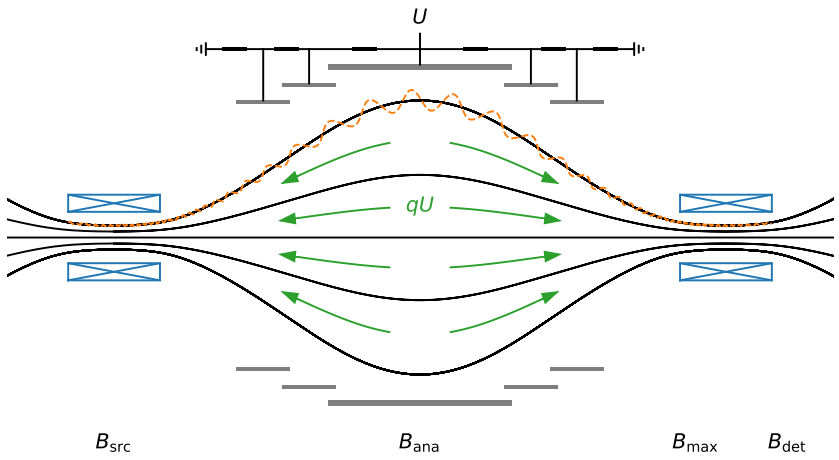


Figure 3.1: Working principle of a MAC-E filter. Electrons (orange dashed) are guided along the magnetic field lines and run against the retarding energy qU (green arrows).

KATRIN was designed to achieve an energy resolution of $0.93\,\mathrm{eV}$ for electron energies near the endpoint of tritium $E \approx 18.6\,\mathrm{keV}$ [26] with the magnetic field settings $B_{\mathrm{max}} = 6\,\mathrm{T}$ and $B_{\mathrm{ana}} = 3 \times 10^{-4}\,\mathrm{T}$. Electrons up to an angle of $\theta_{\mathrm{max}} \approx 51°$ are transmitted with $B_{\mathrm{src}} = 3.6\,\mathrm{T}$.

An important fact to keep in mind is that the magnetic flux $\phi = B \cdot A$ is constant. Therefore, the spectrometer including the analyzing plane must be much larger than the source. We will see later, that this fact also lead to different magnetic field settings in the performed KATRIN measurement campaigns for background mitigation purposes reducing the energy resolution to about $2.8\,\mathrm{eV}$.

3.2 Experimental setup

As described in the previous section, KATRIN uses the MAC-E filter principle to measure an integrated tritium spectrum. An overview of the experimental setup is displayed in fig. 3.2. We will now go over each part of the beamline, starting from the source and working our way towards the detector system, concluding with various monitoring devices.

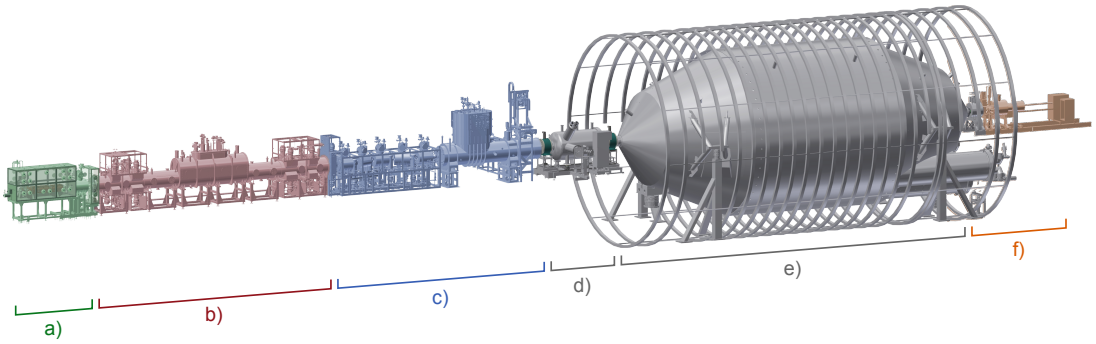


Figure 3.2: Beamline of the KATRIN experiment, figure provided by Leonard Köllenberger.
a) rear section
b) windowless gaseous tritium source
c) transport section
d) pre-spectrometer
e) main spectrometer
f) detector system

Windowless gaseous tritium source

Following the principle of the Mainz and Troitsk experiments, KATRIN uses molecular tritium as β-emitter for its advantageous properties:

- Its endpoint energy of approximately $18.6\,\mathrm{eV}$ is the second lowest of all isotopes which undergo β-decay. This is technically advantageous as it allows operating the MAC-E filter at lower retarding voltages which are easier to control and stabilize. In addition, a lower endpoint leads to relatively more counts in the endpoint region in which the shape distortion induced by the neutrino mass is maximal.

- Molecular tritium T_2 can be used in its gaseous state at low temperatures. The gaseous state avoids complicated solid state effects and low temperatures reduce broadening via the Doppler effect, leading to an overall reduction of systematic uncertainties.

- The β-decay of tritium is super-allowed, leading to a relatively low half-life of $\mathrm{T}_{1/2} = 12.6$ years. This allows high rates at low source densities. On top of this, the nuclear matrix element is energy independent and simple to calculate.

β-decay of molecular tritium is described by

$$\mathrm{T}_2 \rightarrow (\mathrm{HeT})^+ + e^- + \bar{\nu}_e. \tag{3.6}$$

In the KATRIN experiment, molecular tritium gas with a high purity of $\epsilon_\mathrm{T} > 95\,\%$ is kept in a closed loop system [36] and continuously injected into the windowless gaseous tritium

source (WGTS) [37] via 250 holes at its center. The WGTS itself is a $l_{\text{src}} = 10\,\text{m}$ long tube with a diameter of $d_{\text{src}} = 90\,\text{mm}$ where the injected gas can stream freely to both ends. It is then pumped away by multiple turbo molecular pumps (TMPs) and re-injected into the loop system.

The number of gas molecules N_{mol} in the source is described by the column density ρd which is the gas density integrated over the source length. Using the definition of the column density, the total number of gas molecules is given by

$$N_{\text{mol}} = A_{\text{src}} \cdot \rho d \tag{3.7}$$

with the source area

$$A_{\text{src}} = \left(\frac{d_{\text{src}}}{2}\right)^2 \pi. \tag{3.8}$$

As the source is almost fully composed of hydrogen isotopologues with two atoms, the number of atoms $N_{\text{atom}} = 2 \cdot N_{\text{mol}}$.

Electrons emitted in the source are guided along the beamline using magnetic field lines with the source magnetic field set to B_{src}. As described in section 3.1, only electrons with starting angles smaller than θ_{max} reach the detector to avoid long paths in the source and thus large scattering effects. In addition, the magnetic fields are configured to only map the inner part of the source tube to the detector and avoid the outermost region to exclude electrons that scatter on the beam tube. This reduces the source area to A_{eff} and thus the total number of atoms within the volume visible to the detector to

$$N_{\text{atom,eff}} = 2 \cdot A_{\text{eff}} \cdot \rho d. \tag{3.9}$$

When combining this with the solid angle of the acceptance, we can define an efficiency describing how many of the tritium decays are actually mapped towards the detector:

$$\frac{A_{\text{eff}}}{A_{\text{src}}} \cdot \frac{1 - \cos\theta_{\text{max}}}{2} \tag{3.10}$$

which amounts to $15.3\,\%$ assuming the design magnetic flux of $191 \times 10^{-4}\,\text{T}\,\text{m}^{-2}$ and the design magnetic field $B_{\text{src}} = 3.6\,\text{T}$. Note that this differs slightly from the configuration used in any neutrino mass campaign.

The number of tritium atoms is then given by multiplying the total number of atoms with the *tritium purity*

$$\epsilon_{\text{T}} = c_{\text{T}_2} + \frac{1}{2} c_{\text{HT}} + \frac{1}{2} c_{\text{DT}} \tag{3.11}$$

given by the relative concentrations $c_i \in [0, 1]$ of the tritium isotopologues T_2, HT and DT and thus reducing the total number of tritium atoms to

$$N_{\text{T}} = 2 \cdot A_{\text{src}} \cdot \rho d \cdot \epsilon_{\text{T}} \tag{3.12}$$

and the effective number mapped to the detector

$$N_{\text{eff}} = 2 \cdot A_{\text{eff}} \cdot \rho d \cdot \epsilon_{\text{T}}. \tag{3.13}$$

Using the half life of tritium and the number of tritium atoms, we can estimate the total activity in the source

$$\Gamma_{\text{src}} = N_{\text{T}} \cdot \frac{\ln 2}{\text{T}_{1/2}}. \tag{3.14}$$

Inserting the maximal column density of $5 \times 10^{21}\ \mathrm{m}^{-2}$ and a tritium purity of $95\ \%$, the maximal activity is approximately 10^{11} decays per second.

Transport section

While electrons should be guided towards the spectrometer section, tritium molecules and ions must be rejected. The transport section fulfils these two tasks with a combination of methods, always adiabatically guiding the electrons along the magnetic field lines of superconducting magnets.

First component of the transport section is the differential pumping section (DPS) [38]. The beam tube is formed in a chicane shape to avoid a direct line of sight from the source to the spectrometer. Tritium gas is then removed by a set of TMPs reducing the overall flow by at least seven orders of magnitude. As charged ions follow the magnetic field lines just like the electrons, another method to reject these is applied: a dedicated set of ring and dipole electrodes with a positive voltage ranging from $5\ \mathrm{V}$ to $200\ \mathrm{V}$ to block the positively charged ions.

Following the DPS is the cryogenic pumping section (CPS) [39]. Similar to the DPS, it is arranged in a chicane shape. Instead of using TMPs, it makes use of cryo-sorption of tritium on a condensed argon frost layer on the gold-plated beam tube wall. This method reduces the tritium flow by another seven orders of magnitude, leading to an overall rejection factor of more than 10^{14}.

Spectrometer section

After passing the transport section, the electrons reach the spectrometer section consisting of the pre- and the main spectrometer. Both follow the MAC-E filter principle described in section 3.1, but fulfil slightly different tasks.

The smaller pre-spectrometer is typically operated at a retarding voltage of about $300\ \mathrm{V}$ less than the main spectrometer. It thus acts as a pre-filter, rejecting the majority of low energy

electrons not interesting for the energy measurement. This was originally designed to reduce the overall background rate [34]. However, the field configuration of two MAC-E filters operated in sequence leads to a Penning trap between the two spectrometers. This Penning trap builds up an increased amount of stored electrons over time, and can only be cleaned with a set of so-called *Penning wipers* installed specifically for this purpose [40]. In practice, this leads to an additional background component which depends on the time between the operation of the Penning wipers. For this reason, the pre-spectrometer will no longer be operated in upcoming neutrino mass measurement campaigns and was turned off starting from the end of the fourth measurement phase. This background will be discussed in more detail in section 3.4.

Following the pre-spectrometer is the $23\,\mathrm{m}$ long main spectrometer with a diameter of $10\,\mathrm{m}$. This large size is required to achieve an energy resolution on the electronvolt scale with a low magnetic field in the analyzing plane, while fully mapping the relatively large source area. The final energy selection of the electrons is performed here, and all electrons with sufficient energy to pass the electrostatic filter are then re-accelerated and re-focused towards the end of the main spectrometer. To this end, an ultra precise high voltage divider with ppm accuracy and excellent reproducibility on the $10\,\mathrm{meV}$ scale [41–43] ensures precise energy filtering as well as easy data combination. The main spectrometer air coil system [44, 45], consisting of 14 low-field correction system (LFCS) and 16 earth magnetic-field correction system (EMCS) coils, together with the beam line magnets defines the MAC-E filter setup. The flexibility of individually setting each air coil allows for complex magnetic field setups, such as the shifted analyzing plane discussed in section 3.5.2.

Detector system

All electrons leaving the main spectrometer undergo an additional post-acceleration of roughly $10\,\mathrm{kV}$ before hitting the focal plane detector (FPD) [46]. The FPD is a 148 pixel silicon p-i-n detector with an energy resolution of approximately $1.4\,\mathrm{keV}$ full-width half-maximum (FWHM) per pixel. This energy resolution is sufficient to perform a region of interest cut, rejecting a majority of natural backgrounds at lower energies. The pixels are arranged in a dart-board shape shown in fig. 3.3, each with an area of $44\,\mathrm{mm}^2$. This pixel arrangement allows taking radial and azimuthal effects into account in the analysis.

Rear section

The rear section serves two main purposes: house multiple calibration devices and provide a defined electric potential to the beginning of the tritium source [47].

Enclosing the source at the front of the rear section is the rear wall. Its main purpose is to provide a well-defined surface towards the source plasma which determines the starting

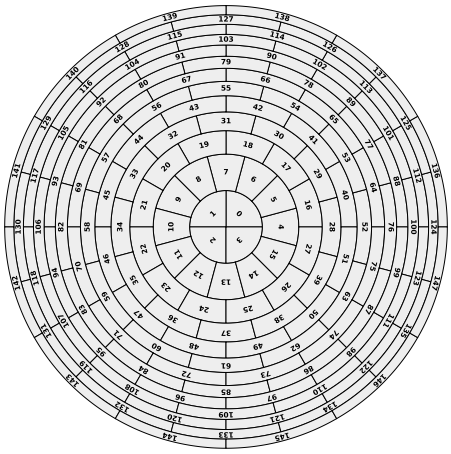


Figure 3.3: Pixel arrangement of the focal plane detector.

potential of the β-electrons and will be covered in more detail in section 3.4. The rear wall is a $145\,\mathrm{mm}$ wide gold-coated stainless steel disk. To manipulate the surface potential and thus the source plasma, a bias voltage of up to $\pm 500\,\mathrm{V}$ can be applied, although in practice the voltage rarely exceeds $\mathcal{O}(1\,\mathrm{V})$.

To measure the gas density in the source and energy loss effects of electrons via scattering, a monoenergetic electron gun (egun) is housed in the rear section [48, 49]. Monoenergetic electrons are produced via the inverse photoelectric-effect and accelerated to energies slightly above those of the β-electrons. These electrons then travel through the full beamline of the experiment before they reach the detector, allowing a precise measurement of response and scattering effects described in section 3.3.

Another device located in the rear section is the β-induced X-ray spectrometry (BIXS) activity monitor [50]. Electrons that hit the rear wall and are absorbed in the gold layer emit Bremsstrahlung. The emitted X-ray photons are then detected using two off-axis silicon drift detectors. As the number of X-ray photons is proportional to the number of emitted electrons, which in turn is proportional to the source activity, it can be monitored at the $0.1\,\%$ level.

Forward beam monitor

An additional activity monitoring device is the forward beam monitor (FBM) located between the transport and the spectrometer section [51]. At this point, it measures the incoming flux of all β-electrons. The FBM is designed to achieve a statistical uncertainty on the β-rate of $<0.1\,\%$ after $60\,\mathrm{s}$ of measurement time with an absolute rate stability of $0.1\,\%$ on the $24\,\mathrm{h}$ scale.

For the first neutrino mass measurement, this task was fulfilled by a 2-pixel silicon p-i-n

detector with an energy resolution of approximately $2\,\mathrm{keV}$ FWHM. After the first neutrino mass measurement, a new 7-pixel TRISTAN [52] prototype silicon drift detector (SDD) was installed as FBM improving the energy resolution to $<350\,\mathrm{eV}$ FWHM [53]. While fulfilling the design requirements, both devices observe an internal long-term drift leading to a drift of the measured rate on the percent-level per month.

3.3 Modelling the expected count rate

Next, we derive a model describing the expected electron count rate at the detector as a function of the retarding potential [27].

3.3.1 Molecular tritium β-spectrum

We start with the electon energy spectrum for an allowed β-decay derived in section 2.3.3 given by eq. (2.27) and shown in fig. 2.3. This formula could be applied directly to the β-decay of atomic tritium. However, since KATRIN uses tritium in its molecular form, the daughter molecule $(\mathrm{HeT})^+$ can be in a rotational, vibrational or electronic excited state f. This requires additional energy which is no longer shared by the emitted electron and neutrino. Introducing the energy of a particular state V_f, the corresponding probability P_f and $\epsilon_f = E_0 - E - V_f$, our differential β-spectrum now reads as

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E} = \frac{G_\mathrm{F}^2 \cos\theta_\mathrm{C}^2}{2\pi^2} \cdot |M_\mathrm{nuc}|^2 \cdot F(Z', E) \cdot p \cdot (E + m_e) \cdot \sum_f P_f \cdot \epsilon_f \sqrt{\epsilon_f^2 - m_\nu^2} \cdot \Theta(\epsilon_f - m_\nu). \quad (3.15)$$

This final state distribution (FSD) can be calculated theoretically [54] and is shown in fig. 3.4. We can see a rather narrow distribution of states centered at $1.75\,\mathrm{eV}$ with a width of $0.44\,\mathrm{eV}$, the so-called electronic ground state, followed by electronically excited states with energies larger than $19\,\mathrm{eV}$. The total probability to decay into a ground state is $P_\mathrm{gs} = 57.4\,\%$. Figure 3.5 shows the effect of the FSD on the β-spectrum in the endpoint region:

- The spectrum is shifted to the left due to the positive mean energy of the electronic ground state.
- A lower overall rate is observed due to the $57.4\,\%$ probability to decay into a ground state, excited states are only visible much deeper in the spectrum.

The FSD depends on the exact molecular shape, and thus differs for the tritium isotopologues T_2, HT and DT present in the KATRIN source. This can be described by interpreting P_f as the product of the theory calculation given a single isotopologue and the fraction of the source

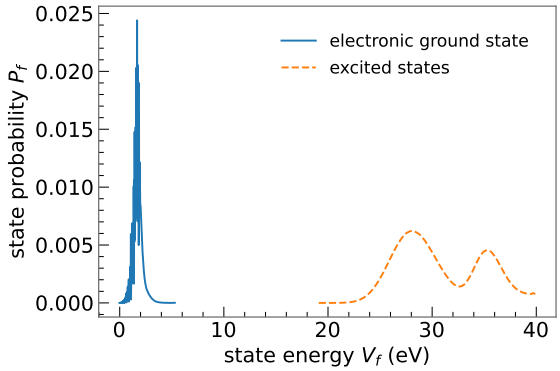


Figure 3.4: Final state distribution of $(HeT)^+$ with the electronic ground state (blue solid line) and electronically excited states (orange dashed line).

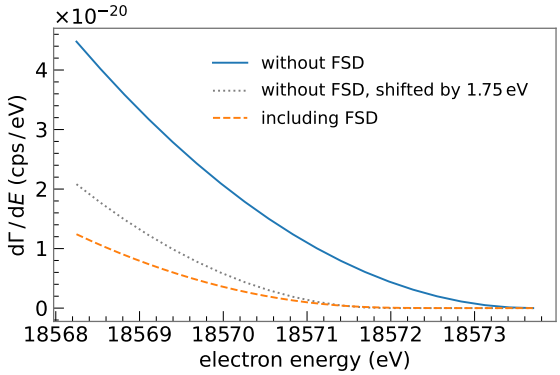


Figure 3.5: Tritium β-spectrum without final states (blue solid line) and including the final state distribution of $(HeT)^+$ shown in fig. 3.4 (orange dashed line). An endpoint value of $E_0 = 18\,573.7\,\mathrm{eV}$ and a neutrino mass of $m_\nu = 0\,\mathrm{eV}$ is assumed.

molecules made up by the corresponding isotopologue, essentially tripling the total number of final states in the sum.

In addition to the final state effect, the thermal motion of tritium molecules in the source leads to a Doppler broadening of the β-spectrum. In the non-relativistic approximation, this can be described by convolving eq. (3.15) with a normal distribution

$$g(\Delta E) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{D}}} \cdot \exp\left(-\frac{\Delta E^2}{2\sigma_{\mathrm{D}}^2}\right) \tag{3.16}$$

with standard deviation corresponding to the Doppler broadening width

$$\sigma_{\mathrm{D}} = \sqrt{2E\, k_{\mathrm{B}} T \frac{m_e}{m_{\mathrm{T}_2}}} \tag{3.17}$$

which is defined by the energy of the electron E, the Boltzmann constant k_{B}, the source temperature T and the mass ratio of electron and the tritium molecule $\frac{m_e}{m_{\mathrm{T}_2}}$. Inserting typical values of $E = E_0 \approx 18.6\,\mathrm{keV}$ and $T = 30\,\mathrm{K}$ $(90\,\mathrm{K})$ leads to a broadening of $\sigma_{\mathrm{D}} = 93.5\,\mathrm{meV}$ $(161.9\,\mathrm{meV})$. This changes the differential spectrum to

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E) \to \int_{-\infty}^{+\infty} g(E-\epsilon)\frac{\mathrm{d}\Gamma}{\mathrm{d}E}(\epsilon)\,\mathrm{d}\epsilon. \tag{3.18}$$

The impact of an (exaggerated) broadening on the β-spectrum in the endpoint region is shown in fig. 3.6 and compared to the effect of inserting unphysical negative neutrino mass squared values into the formula describing the differential spectrum. The broadening shifts the spectrum to higher energies, and this shape-effect can be described very well by negative m_ν^2 values. This observation can be explained by the relation $\Delta m_\nu^2 = -2\sigma^2$ derived for example in [55] which relates an unnacounted broadening variance σ^2 to a negative shift of the neutrino mass squared value.

In practice, it is often favourable to not describe an additional broadening by a convolution of the full β-spectrum with the Gaussian kernel, but instead emulate it in the final state distribution. To this end, every discrete final state with probability P_f and energy V_f is replaced by a Gaussian of the form

$$P_f \cdot \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{broad}}} \cdot \exp\left(-\frac{(E-V_f)^2}{2\sigma_{\mathrm{broad}}^2}\right). \tag{3.19}$$

For practical reasons, the continuous sum of all Gaussian distributions is then re-binned with a bin width of $0.1\,\mathrm{eV}$ or finer. The effect this has on the FSD is shown in fig. 3.7. Note that this additional broadening width σ_{broad} must not be limited to the Doppler broadening σ_{D}, but can include other effects that effectively lead to a broadening in the energy scale such as time-wise changes.

On top of the effects described, there are various theoretical corrections to the β-spectrum going beyond the basic Fermi theory of β-decay. The only contribution significant and understood well enough to be included in the current KATRIN analysis are radiative corrections

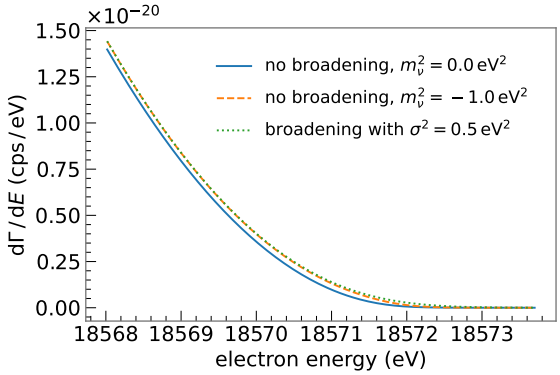


Figure 3.6: Impact of a non-zero broadening on the tritium spectrum compared to negative neutrino mass squared values. The broadenened spectrum (green dotted line) is shifted to the right compared to the unbroadened one (blue solid line) when both assume $m_\nu^2 = 0.0\,\mathrm{eV}^2$. Inserting the unphysical value of $m_\nu^2 = -1.0\,\mathrm{eV}^2$ into the differential spectrum (orange dashed line) leads to a very similar shape as the broadening.

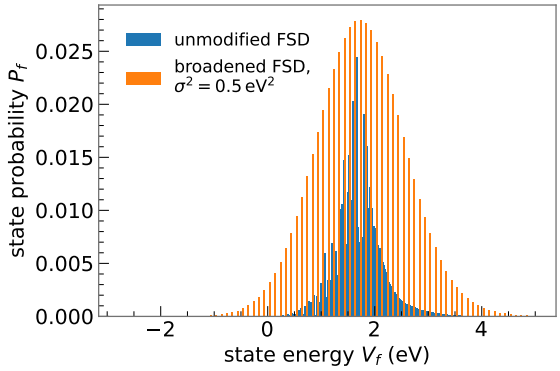


Figure 3.7: Effect of emulating a broadening in the FSD on the electronic ground state. One can see a widening of the distribution, as well as a slightly reduced number of states due to the binning of $0.1\,\mathrm{eV}$.

related to electromagnetic effects involving contributions from virtual and real photons. The multiplicative correction factor G is shown in fig. 3.8.

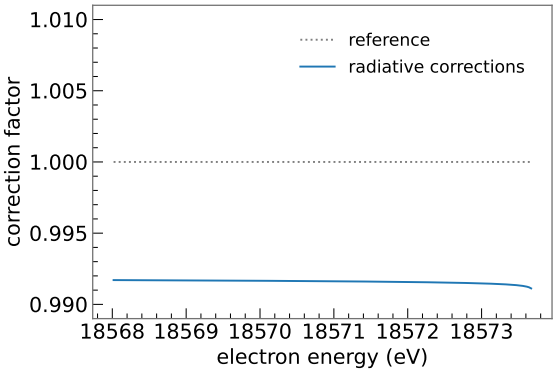


Figure 3.8: Impact of radiative corrections on the differential β-spectrum in the endpoint region.

3.3.2 Experimental response function

In addition to the β-spectrum, we need to take into account the experimental response function R which describes the probability of an electron with energy E to pass the electrostatic filter set at the retarding energy qU. In an ideal apparatus, this would be a simple step function

$$R_{\text{ideal}}(E; qU) = \begin{cases} 0 & E < qU \\ 1 & E \geq qU. \end{cases} \tag{3.20}$$

However, in a realistic setup, there are various effects that modify this transmission probability.

Transmission properties of the MAC-E filter

As described in section 3.1, the retarding voltage set in the analyzing plane is only sensitive to the kinetic energy of the electron parallel to the electric field lines. Thus, any remaining transversal energy may lead to the rejection of electrons with a total kinetic energy greater than qU. We now extend our calculations to derive the transmission condition depending on the starting angle of the electron, and from this the transmission function of the MAC-E filter.

At any given point z, an electron can only continue moving against the electrostatic potential if

$$E_{\parallel}(z) > 0. \tag{3.21}$$

To retrieve the transmission condition at any point in z, we consequently need an expression for $E_{\parallel}(z)$. We define the starting potential of the electron as qU_{src}, its initial kinetic energy E with corresponding gamma factor γ and pitch angle θ in the given magnetic field B_{src}. Using the adiabatic invariant defined in eq. (3.2), we can write

$$E_{\parallel}(z) = E - E_{\perp}(z) + qU_{\text{src}} - qU(z) \tag{3.22}$$

$$= E - E\sin^2\theta \frac{B(z)}{B_{\text{src}}} \frac{\gamma+1}{\gamma(z)+1} + qU_{\text{start}} - qU(z). \tag{3.23}$$

Inserting this into eq. (3.21), we retrieve

$$E - E\sin^2\theta \frac{B(z)}{B_{\text{src}}} \frac{\gamma+1}{\gamma(z)+1} > qU(z) - qU_{\text{src}}. \tag{3.24}$$

In the MAC-E filter, the minimal magnetic field B_{ana} and maximum retarding energy qU coincide in the analyzing plane. Any electron passing this plane therefore also passes the electrostatic filter. We thus have to evaluate eq. (3.24) for the values at the analyzing plane to calculate if an electron passes the filter. Without loss of generality, we set $qU_{\text{src}} = 0$. This leads us to the transmission function depending on electron energy and starting angle:

$$T(E,\theta;qU) = \begin{cases} 1 \text{ if } E - E\sin^2\theta \frac{B_{\text{ana}}}{B_{\text{src}}} \frac{\gamma+1}{\gamma_{\text{ana}}+1} > qU \\ 0 \text{ else} \end{cases} \tag{3.25}$$

with the gamma factor in the analyzing plane $\gamma_{\text{ana}} = \frac{E-qU+m_e}{m_e} \approx 1$.

To calculate the probability of any electron with energy E to overcome qU, we have to integrate eq. (3.25) over θ given an angular distribution of the electrons emitted in the source. Assuming an isotropic distribution, our integral reads

$$T(E;qU) = \int_{\theta=0}^{\theta=\theta_{\text{max}}} T(E,\theta;qU)\sin\theta\,\mathrm{d}\theta. \tag{3.26}$$

$$= \int_{\cos\theta_{\text{max}}}^{1} T(E,\cos\theta;qU)\,\mathrm{d}\cos\theta \tag{3.27}$$

As one always integrates over a constant value (zero or one depending if electrons with this $\cos\theta$ pass the filter or not), this essentially simplifies to finding the maximum starting angle theta that allows passing the filter for a given E:

$$E\left(1 - (1-\cos^2\theta)\frac{B_{\text{ana}}}{B_{\text{src}}}\frac{\gamma+1}{\gamma_{\text{ana}}+1}\right) > qU \tag{3.28}$$

$$\Leftrightarrow \cos\theta > \sqrt{1 - \frac{E-qU}{E}\frac{B_{\text{src}}}{B_{\text{ana}}}\frac{\gamma_{\text{ana}}+1}{\gamma+1}}. \tag{3.29}$$

Note that for $E < qU$ this condition can never be fulfilled and it is fulfilled for all accepted angles if $E > qU + \Delta E$ with the energy resolution ΔE from eq. (3.5). This leads to the

transmission function for isotropic electrons being

$$T(E;qU) = \begin{cases} 0 \text{ if } E < qU \\ 1 - \sqrt{1 - \frac{E-qU}{E}\frac{B_{\text{src}}}{B_{\text{ana}}}\frac{\gamma_{\text{ana}}+1}{\gamma+1}} \text{ if } qU < E < qU + \Delta E \\ 1 - \sqrt{1 - \frac{B_{\text{src}}}{B_{\text{max}}}} \text{ if } E > qU + \Delta E. \end{cases} \tag{3.30}$$

This representation of the transmission function still includes the effect that not all electrons can make it passed the filter due to the maximum acceptance angle. Usually, the solid angle is treated independent of the transmission function, giving us

$$T(E;qU) = \begin{cases} 0 \text{ if } E < qU \\ \frac{1-\sqrt{1-\frac{E-qU}{E}\frac{B_{\text{src}}}{B_{\text{ana}}}\frac{\gamma_{\text{ana}}+1}{\gamma+1}}}{1-\sqrt{1-\frac{B_{\text{src}}}{B_{\text{max}}}}} \text{ if } qU < E < \Delta E \\ 1 \text{ if } E > qU + \Delta E. \end{cases} \tag{3.31}$$

The resulting transmission function of the MAC-E filter for isotropic electrons is compared to the ideal step function in fig. 3.9a assuming $B_{\text{max}} = 4.24\,\text{T}$, $B_{\text{src}} = 2.52\,\text{T}$, and $B_{\text{ana}} = 0.63\,\text{mT}$ at a retarding energy of $qU = 18\,545\,\text{eV}$.

At this point it is worth mentioning that in a real apparatus the field settings in the analyzing plane qU and B_{ana} depend on the radius r and to a lesser extend on the angle ϕ, leading to a pixel dependence of the transmission function. This is depicted in fig. 3.9b assuming the field settings for the second neutrino mass campaign. When pixels are not treated independently in the analysis, these individual functions are averaged leading to an additional broadening of the effective transmission function.

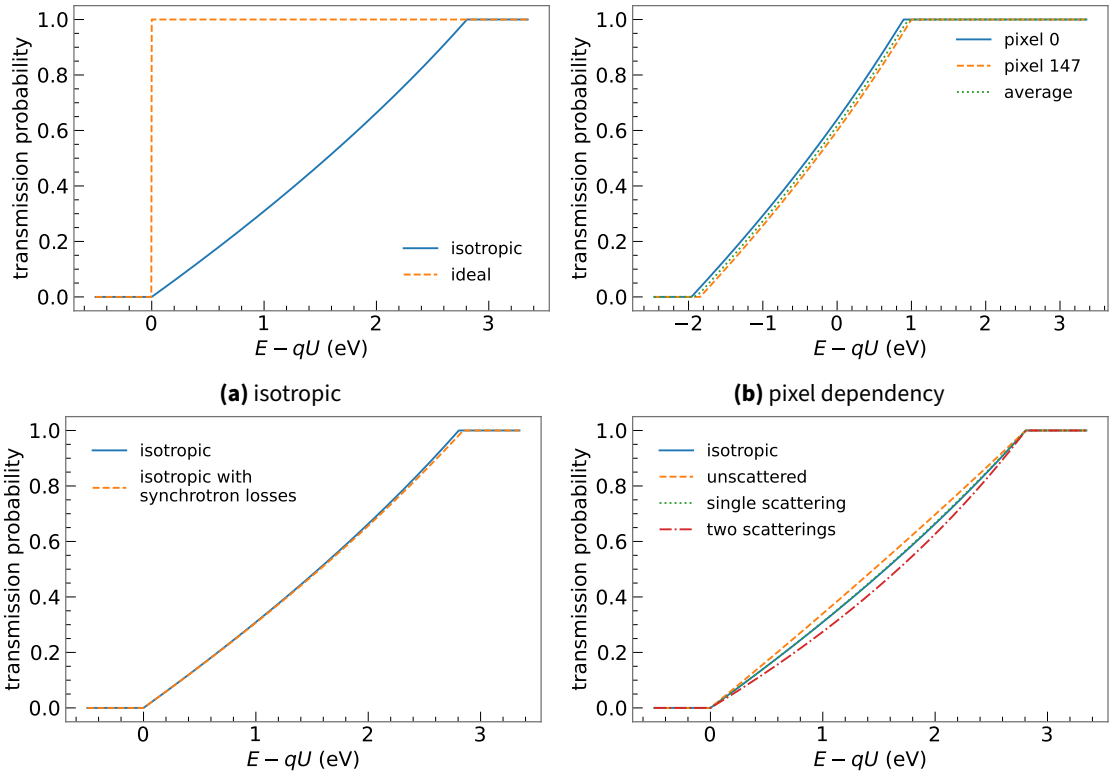
Synchrotron radiation losses

Another effect we must take into account are energy losses of the electron due to synchrotron radiation during its path along the beamline. The synchrotron loss of an electron travelling a distance l at magnetic field B with pitch angle θ is given by

$$\Delta E_{\text{sync}} = -\frac{\mu_0}{3\pi c}\frac{e^4}{m_e^3} \cdot B^2 \cdot E \cdot \sin^2\theta \cdot \gamma \frac{l}{v\cos\theta} \tag{3.32}$$

with the permeability of free space μ_0 and the speed of light c.

In the KATRIN experiment, these losses mainly take place in the transport section with a minor contribution from the source. To good approximation, we can assume an effective model for the average synchrotron loss. To this end, we need the average distance an electron travels in the source which is $\frac{l_{\text{src}}}{2} \approx 5\,\text{m}$, the length of the transport section $l_{\text{tr}} \approx 14\,\text{m}$ and the



(a) isotropic

(b) pixel dependency

(c) effect of synchrotron losses

(d) effect of angular changes due to scattering

Figure 3.9: Various transmission functions describing the probability of an electron with kinetic energy E to pass the retarding energy qU of the electrostatic filter. A clear broadening of the function compared to an ideal filter can be seen due to the MAC-E filter transmission (a). In a real apparatus, the retarding energy qU and magnetic field B_{ana} depend on the radius and thus on the pixel number. This is shown for the settings applied in the second measurement campaign of KATRIN along with the average transmission over all pixels (b). Synchrotron radiation losses lead to a slight degradation of the energy resolution (c) while the non-isotropic angular distribution due to scattering changes the curvature (d).

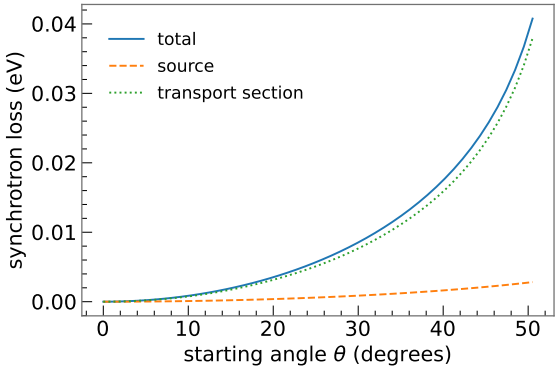


Figure 3.10: Synchrotron energy loss of an electron with different starting angles in the source.

corresponding magnetic fields B_{src} and B_{tr}. The synchrotron loss for electrons with different starting angles in both sections is shown in fig. 3.10. Note that the pitch angle in the transport section is different from the starting angle due to the varying magnetic field.

To include these effects in our transmission function, we must modify eq. (3.25) to include the energy loss:

$$T(E,\theta;qU) = \begin{cases} 1 \text{ if } E + \Delta E_{\text{sync}}(E,\theta) - E\sin^2\theta \frac{B_{\text{ana}}}{B_{\text{src}}} \frac{\gamma+1}{\gamma_{\text{ana}}+1} > qU \\ 0 \text{ else.} \end{cases} \tag{3.33}$$

Finding the maximum starting angle that allows passing qU for a given E is now no longer possible analytically and must be performed with a numerical root search. The effect of the synchrotron losses on the transmission function is shown in fig. 3.9c: a slight degradation of the energy resolution.

Scattering effects

In addition to the transmission effects, an electron can lose energy by inelastic scattering on tritium molecules in the source. To accurately describe this effect, we need two components: the probability of an electron to scatter s-times P_s and a probability density function that describes the amount of energy lost in this process, the so-called energy loss function $f_s(\epsilon)$.

The probability of an electron to scatter s times depends on the cross section σ_{inel} and the effective amount of gas it passes through. This amount of gas is described by the column density ρd, the starting position z and also on the starting angle of the electron as electrons with larger angles travel longer distances in the source due to their synchrotron motion. We summarize this in the quantity

$$\lambda(z,\theta) = \frac{z \cdot \rho d}{\cos\theta} \tag{3.34}$$

where z will range from 0 to 1 and the electron can thus see from no gas at all ($z = 0$ if it was emitted at the very front of the source) up to all of it ($z = 1$ in case it was emitted at the very end of the source). As scattering events are independent and we can assume to excellent approximation that the angle of the electron is unaffected by the scattering [56], the scattering probability can be described by a Poissonian:

$$P_s(z,\theta) = \frac{(\lambda(z,\theta)\cdot\sigma_{\text{inel}})^s}{s!}\exp\left(-\lambda(z,\theta)\cdot\sigma_{\text{inel}}\right). \tag{3.35}$$

To calculate the average scattering probability of an electron, we integrate eq. (3.35) over z and θ, once again assuming an isotropic angular distribution:

$$P_s = \frac{1}{1-\cos\theta_{\text{max}}}\int_0^{\theta_{\text{max}}}\sin\theta\int_0^1 P_s(z,\theta)\,\mathrm{d}z\mathrm{d}\theta. \tag{3.36}$$

The energy-dependent inelastic scattering cross section is given by

$$\sigma_{\text{inel}}(T) = \frac{4\pi a_0^2}{T/R}\cdot\left[M_{\text{tot}}^2\cdot\ln\left(4c_{\text{tot}}\cdot\frac{T}{R}\right)+\delta_{\text{E}}\right] \tag{3.37}$$

with the Bohr radius a_0, the non-relativistic kinetic energy of the electron $T = \frac{1}{2}m_e\beta^2$ given the relativistic beta factor $\beta = 1-\frac{m_e^2}{(m_e+E)^2}$, the Rydberg energy R, $M_{\text{tot}}^2 = 1.5356$ for tritium [57, 58], and $c_{\text{tot}} = 1.18$ [59]. δ_{E} describes the relativistic and $1/E^2$ correction and can be calculated as follows:

$$\delta_{\text{E}} = \frac{\gamma_{\text{tot}}^2}{T/R}+\delta_{\text{rc}}, \tag{3.38}$$

$$\delta_{\text{rc}} = -M_{\text{tot}}^2\left[\ln\left(1-\beta^2\right)+\beta^2\right] \qquad \text{(relativistic correction)}, \tag{3.39}$$

$$\gamma_{\text{tot}}^2 = 2\left[-\frac{7}{4}+\ln\left(\frac{E_i}{T}\right)\right] \qquad (1/E^2 \text{ parameter}). \tag{3.40}$$

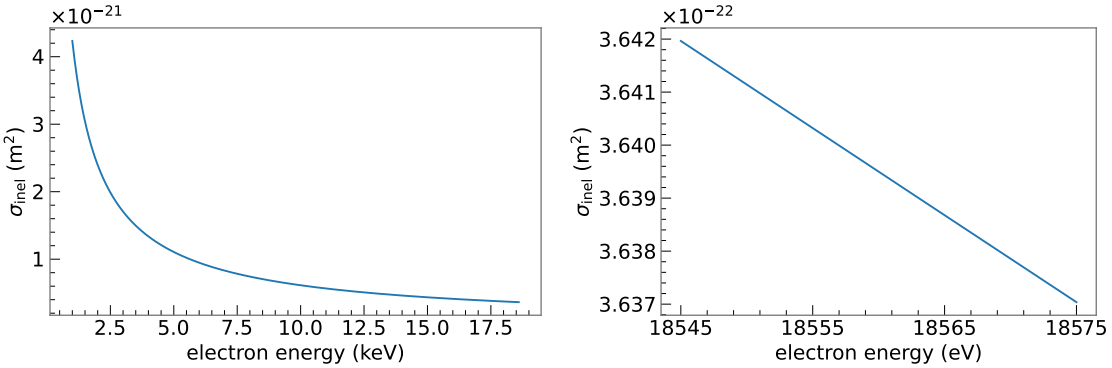
Figure 3.11 shows this cross section over a large energy range (a) as well as in the energy interval interesting for KATRIN (b).

The scattering probabilites for electrons with kinetic energy $E = 18\,575\,\text{eV}$ are shown in fig. 3.12. Clearly, scattering effects increase with increasing gas density. However, unscattered electrons are still the most abundant for all column densities possible at KATRIN.

Having derived the scattering probabilities, the next step is to define the energy loss function. Electrons that do not scatter, lose no energy. Their energy loss function is thus a simple Dirac delta function

$$f_0(\epsilon) = \delta(\epsilon). \tag{3.41}$$

Higher order scatterings $s > 1$ can be described by convolving the energy loss function for one scattering $s-1$-times with itself. To retrieve $f_1(\epsilon)$ dedicated measurements with the electron



(a) large energy range

(b) KATRIN energy interval

Figure 3.11: Inelastic scattering cross section depending on the electron energy. While it shows a clear increase to lower energies (a), there is only a small linear change in the energy interval interesting for KATRIN (b).

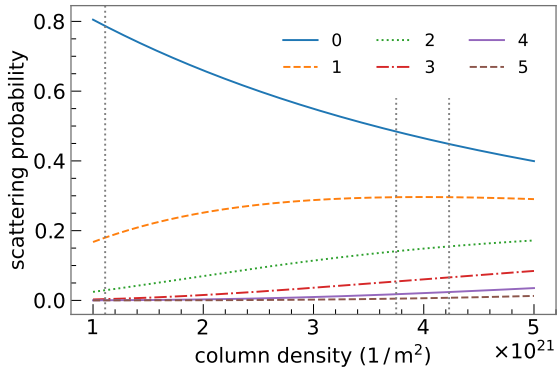


Figure 3.12: Probability of an electron to scatter s times in the source depending on the column density. As expected, scattering effects increase with increasing number of gas molecules. The gray vertical dotted lines indicate the column density of the 1st campaign ($1.11 \times 10^{21}\,\mathrm{m}^{-2}$), the 2nd campaign ($4.23 \times 10^{21}\,\mathrm{m}^{-2}$) and the reference for future campaigns ($3.75 \times 10^{21}\,\mathrm{m}^{-2}$).

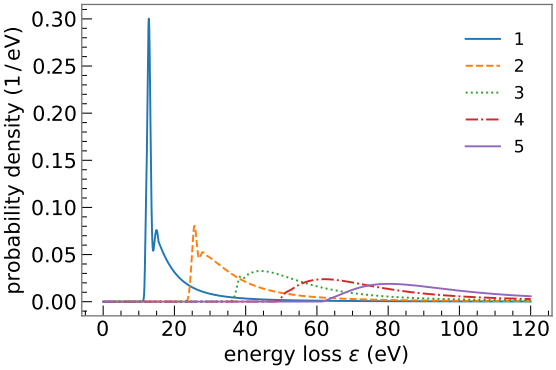


Figure 3.13: Energy loss function for s scatterings describing the probability of an electron losing the energy ϵ in the s-fold scattering process [49]. Higher orders are retrieved by convolving the function for a single scattering $s-1$-times with itself.

gun were performed fitting the semi-empirical function

$$f_1(\epsilon) = \begin{cases} \sum_{j=1}^{3} a_j \exp\left(-\frac{(\epsilon-\mu_j)^2}{2\sigma_j^2}\right) & \text{for } \epsilon \leq E_i \\ \frac{f_1(E_i)}{f_{\mathrm{BED}}(E_i)} \cdot f_{\mathrm{BED}}(\epsilon) & \text{for } \epsilon > E_i, \end{cases} \tag{3.42}$$

which is a sum of three Gaussians with amplitude a_j, mean μ_j, and standard deviation σ_j for energies below the ionization energy $E_i = 15.486\,\mathrm{eV}$ [60] of T_2 and the binary-encounter-dipole (BED) function [61] above, to this measurement data [49]. The resulting energy loss function for electrons scattering on T_2 is shown in fig. 3.13.

Combined response function

To calculate the response function $R(E;qU)$ of the experiment, the transmission function $T(E;qU)$ is convolved with the energy loss function $f_s(\epsilon)$ for each scattering, weighted by the scattering probability P_s:

$$R(E;qU) = \int_0^{E-qU} T(E-\epsilon;qU) \cdot \sum_{s=0}^{\infty} P_s \cdot f_s(\epsilon)\,\mathrm{d}\epsilon. \tag{3.43}$$

In practice, between five and ten scatterings are considered, depending on the analysis window. Scattering more often is highly unlikely, less than 5×10^{-6} at maximal column density, and electrons lose too much energy to pass the electrostatic filter for neutrino mass measurement intervals near the endpoint.

While we mentioned that the angular change of an electron due to scattering is negligible, there is another related effect that should be taken into account. Electrons with large angles are

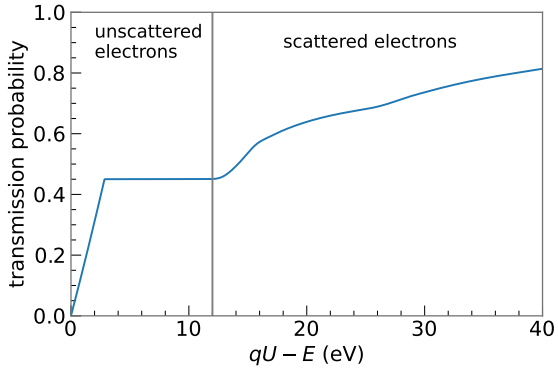


Figure 3.14: Combined response function describing the probability of an electron with energy E to pass the retarding energy qU. The gray line splits the unscattered electrons (left), where only the transmission effects of the MAC-E filter are relevant, from the ones that have scattered at least once (right) and thus also the energy loss function plays a role.

likely to scatter more often. Therefore, on average, these electrons need a larger surplus energy to pass the electrostatic filter. To include this into the calculation, the scattering probability

$$P_s(\theta) = \int_0^1 P_s(z, \theta)\, \mathrm{d}z \tag{3.44}$$

is included when integrating the transmission condition from eq. (3.33) over the pitch angle

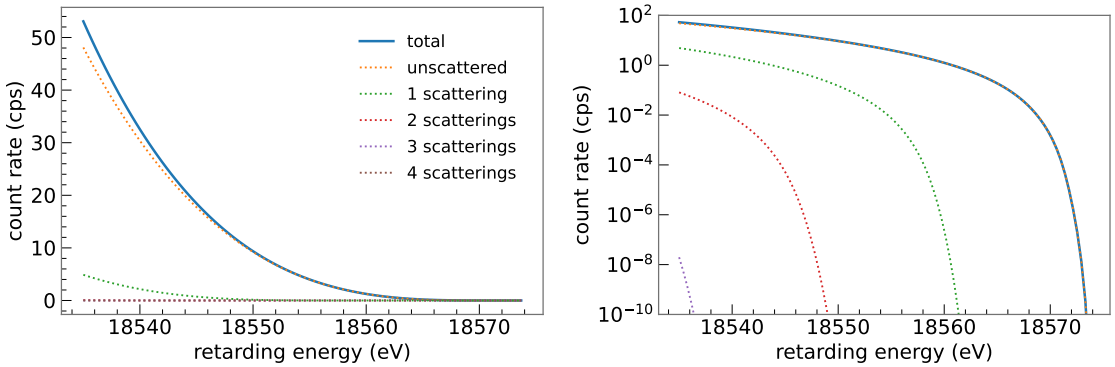
$$T_s(E; qU) = \int_{\theta=0}^{\theta_{\max}} T(E, \theta; qU) \sin\theta P_s(\theta)\, \mathrm{d}\theta, \tag{3.45}$$

leading to a dependence of the transmission function on the number of scatterings. The impact this has on the transmission function for different scatterings is displayed in fig. 3.9d. Note that now both finding the maximum starting angle that allows passing qU for a given E and the actual integral evaluation are non-trivial and need to resort to numerical evaluations.

Taking this into account, the response function is slightly modified to

$$R(E; qU) = \int_0^{E-qU} \sum_{s=0}^{\infty} P_s \cdot T_s(E - \epsilon; qU) \cdot f_s(\epsilon)\, \mathrm{d}\epsilon. \tag{3.46}$$

It is most important to include this effect for unscattered electrons, since these are by far the most abundant as will be shown in the next section. The combined response function can be seen in fig. 3.14.



(a) regular scale

(b) logarithmic scale

Figure 3.15: Integrated tritium spectrum according to eq. (3.47) in regular (a) and logarithmic (b) scale assuming $\rho d = 4.2 \times 10^{21}\,\mathrm{m}^{-2}$. The blue solid line is the total integrated spectrum, while the other colored dotted lines indicate the contributions of the individual number of scatterings. The sum is clearly dominated by unscattered electrons ($> 99\,\%$ 20 eV below E_0; $\approx 90\,\%$ 40 eV below E_0), followed by those which scattered once.

3.3.3 Integrated spectrum

To complete our signal model $I(qU)$, the differential tritium β-spectrum is integrated over the experimental response function:

$$I(qU) = C \cdot \int_{-\infty}^{+\infty} \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E) \cdot R(E; qU)\,\mathrm{d}E \tag{3.47}$$

with the constant prefactor

$$C = N_{\mathrm{eff}} \frac{1 - \cos\theta_{\mathrm{max}}}{2} \cdot \epsilon_{\mathrm{detector}} \tag{3.48}$$

including the effective number of tritium atoms, c.f. eq. (3.13), the solid angle and a constant detector efficiency $\epsilon_{\mathrm{detector}}$.

This integrated spectrum is shown in fig. 3.15 including the contributions of the individual number of scatterings.

3.3.4 Total model of the expected counts

In addition to the signal model, we must add a constant background rate r_{bg} to the integrated spectrum to calculate the expected count rate

$$r(qU) = I(qU) + r_{\mathrm{bg}}. \tag{3.49}$$

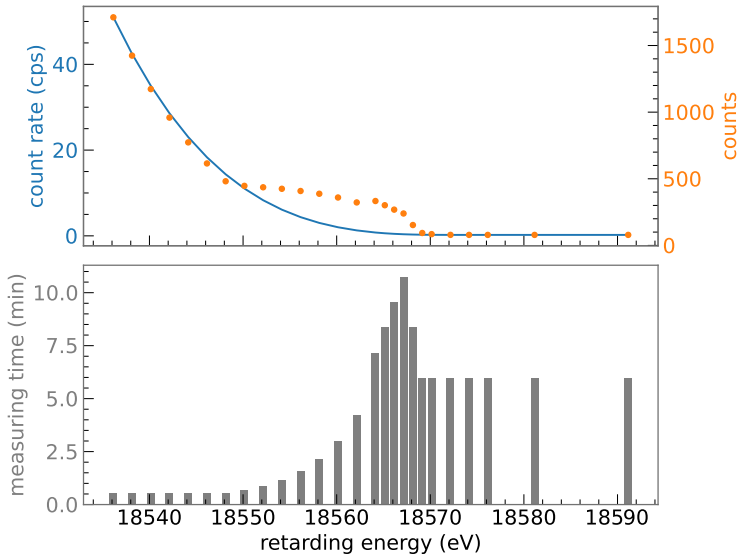


Figure 3.16: Distribution of counts for a single spectral scan (orange dots) given our model rate calculation (blue solid line) and the measurement time distribution (gray bars).

Any corrections to this flat background model are considered systematics and will be discussed in more detail in section 3.4.

During a neutrino mass campaign, the same set of retarding energies $\{qU_i\}$ is scanned repeatedly. In each scan, the time t_i is spent at the corresponding qU_i. Typically between 30 to 40 retarding energies are scanned with a total scan time of two to four hours. How much time is spent at each qU_i is defined by the measuring time distribution (MTD) which is optimized for maximal neutrino mass sensitivity [62]. Accounting for this, our expected number of counts μ_i can be defined as

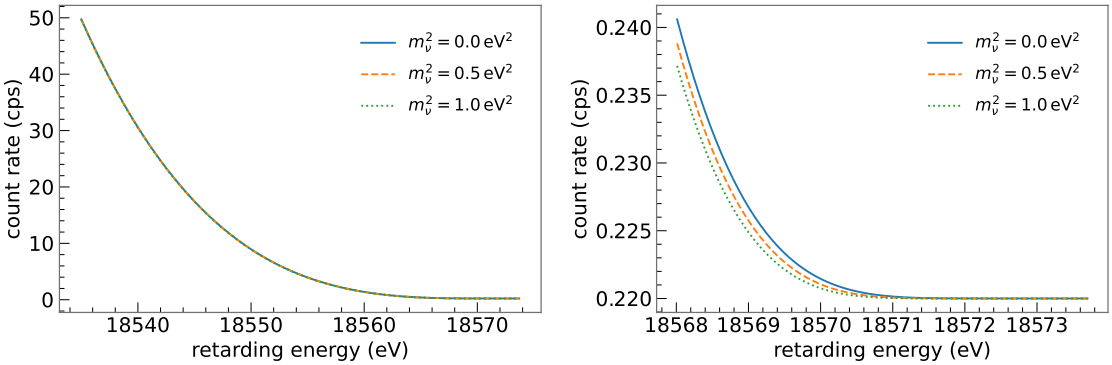
$$\mu_i = \mu(qU_i, t_i) = r(qU_i) \cdot t_i = r_i \cdot t_i. \tag{3.50}$$

Figure 3.16 shows an example for the distribution of counts given our model rate and the measuring time distribution.

3.3.5 Neutrino mass imprint

To conclude this section, we revisit the imprint the neutrino mass leaves and take a closer look how the neutrino mass signal compares to the endpoint and is affected by the background rate.

Figure 3.17 displays the total model rate for different values of m_ν^2, similar to fig. 2.3b which displays the impact on the plain differential β-spectrum. Over the full measurement interval



(a) full measurement interval

(b) endpoint region

Figure 3.17: Model count rate for different neutrino mass squared values given a background rate of $r_{\mathrm{bg}} = 220\,\mathrm{mcps}$. In the full measurement interval (a) the spectra can hardly be distinguished requiring a zoom into the endpoint region (b).

a neutrino mass squared of $1.0\,\mathrm{eV}^2$ is nearly indistinguishable from the reference without neutrino mass. Zooming into the endpoint region, the difference becomes visible, however it is not clear how much this is different from a plain shift of the energy scale or endpoint value.

To investigate this, fig. 3.18 shows the relative model rate

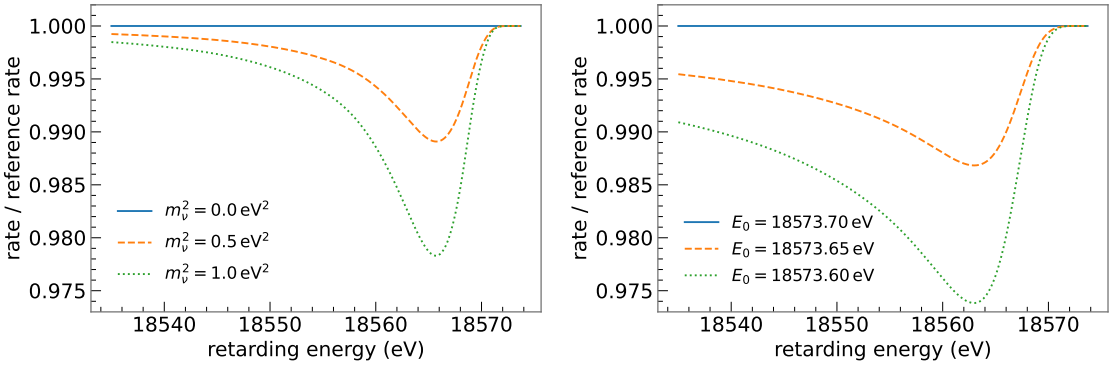
$$\frac{r(m_\nu^2, E_0)}{r(m_\nu^2 = 0.0\,\mathrm{eV}^2, E_0 = 18\,573.7\,\mathrm{eV})} \tag{3.51}$$

for different m_ν^2 and E_0 values. Both a positive m_ν^2 and a lower E_0 value lead to relatively less counts in the signal region. This shows as a dip structure in the relative rate where the dip related to a positive m_ν^2 value is sharper than the one of a reduced E_0. This allows distinguishing these two parameters in an analysis, but the clear similarity of the two effects explains the well known large correlation of m_ν^2 and E_0 for neutrino mass measurements from β-decay [26].

The position of the largest impact, i.e. the minimum in figs. 3.18a and 3.18b, as well as its scale depends on the background value: In the integrated spectrum, the effect is largest close to the endpoint, leading to a reduction of the change towards lower retarding energies. However, the non-zero background rate means that the signal to background ratio worsens towards higher retarding energies. This worsening of the signal to background ratio leads to a decrease of the m_ν^2 impact which at some point overpowers the increase of the effect in the β-spectrum. It can be shown that the point of maximum impact is approximately where the signal to background ratio equals one. To display this, fig. 3.19 shows the relative rate

$$\frac{r(m_\nu^2 = 1.0\,\mathrm{eV}^2, r_{\mathrm{bg}})}{r(m_\nu^2 = 0.0\,\mathrm{eV}^2, r_{\mathrm{bg}})} \tag{3.52}$$

for different background rates r_{bg}. From this figure it is clear that lowering the background rate is of high priority for the KATRIN experiment.



(a) neutrino mass squared

(b) endpoint

Figure 3.18: Shape effect of a non-zero neutrino mass squared on the model rate (a) compared to a lower endpoint value (b).

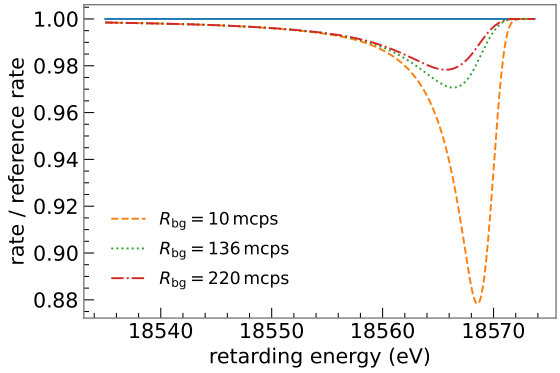


Figure 3.19: Shape effect of $m_\nu^2 = 1.0\,\mathrm{eV}^2$ on the model rate for different background rates. The orange dashed line corresponds to the design background, the red dashed line to the background for the second measurement campaign and the green dotted one to the background rate achieved with an optimized field setting described in section 3.5.2.

3.4 Sources of systematic uncertainty

In this section the various sources of systematic uncertainty present in the KATRIN experiment are described.

Residual tritium on the rear wall

Over time, tritium can accumulate on the rear wall by T_2 hitting the rear wall and forming various tritiated molecules. The tritium in these molecules can later decay, leading to a superposition of the signal β-spectrum coming from the decay of T_2 with the β-spectrum of tritium on the rear wall [63].

The activity depends on the integrated flux of circulated tritium gas and therefore increases over time if no measures are taken. Since the total flux was rather low, the residual tritium was not an issue for the first two neutrino mass campaigns, but must be taken into account for future analyses. This can be done by measuring the residual β-spectrum with an empty source and including it in the neutrino mass analysis by adding it to our model rate expectation defined in eq. (3.49). Typically, the rear wall spectrum resembles the β-spectrum of T_2 to good approximation with a slightly higher endpoint $E_{0,\mathrm{RW}}$. Thus, the systematic uncertainties can arise from uncertainties on the rear wall activity, the measured $E_{0,\mathrm{RW}}$ and the accuracy of the β-spectrum shape itself.

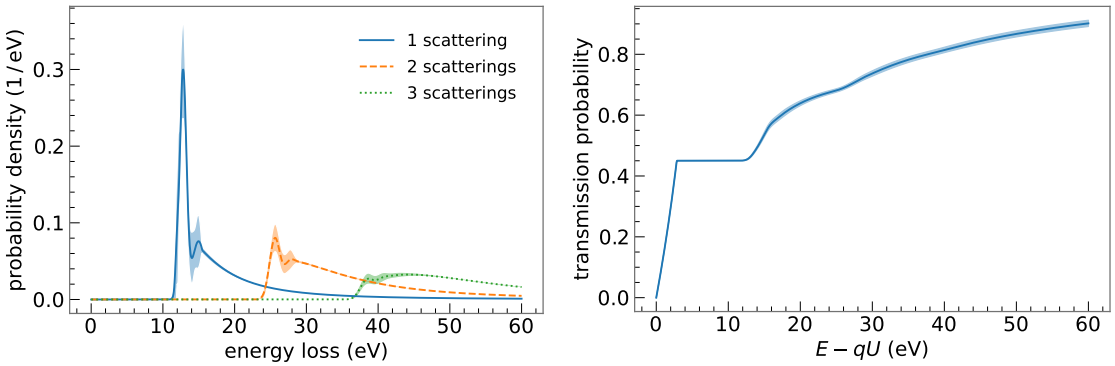
To lessen the accumulation effect, the rear wall is regularly cleaned using ozone reducing the activity by $>95\,\%$.

Gas density in the source

The overall gas density in the source, described by the column density ρd, affects the signal model in two ways:

- The overall number of tritium molecules is not known exactly. This only changes the overall normalization of the model rate, and does not impact the shape. Therefore, it has no impact on the neutrino mass measurement.

- The scattering probabilities described by eq. (3.35) have an uncertainty. This directly impacts the response function shape and must therefore be taken into account as uncertainty in any m_ν analysis.



(a) energy loss **(b)** response function

Figure 3.20: Uncertainty band of the energy loss function for one to three scatterings (a) and its propagation into the response (b). The uncertainty is increased by a factor of 100 for better visibility.

Due to its impact on the scattering probabilities, a precise estimate of the absolute value of $\rho d \sigma_{\text{inel}}$, the column density times the inelastic scattering cross section, is required. For a single campaign, the average ρd value is determined by combining several measurements with the electron gun with a continous monitoring device. Each egun measurement effectively scans the shape of the response function, see fig. 3.14, by changing the retarding voltage of the filter while keeping a fixed egun electron energy E. Fitting this response function gives an accurate estimate of $\rho d \sigma_{\text{inel}}$ which is used to anchor the value of any continuous montioring device. The monitoring device is then used to calculate ρd values for every (sub-) scan during a measurement campaign and therefore defines the time evolution of ρd from which the average column density is estimated [64].

Energy loss function

All parameters describing the energy loss function, c.f. eq. (3.42), come with a fit uncertainty described by their covariance K_{eloss}. This directly propagates into an uncertainty on the response function shape. This impact of K_{eloss} on the energy loss function as well as the response function is shown in fig. 3.20. The energy loss function as well as its uncertainty was measured in situ with a combination of differential and integral egun measurements [49].

Molecular final states

Describing the uncertainty of the theoretical calculation of the molecular final state distribution is non-trivial as the theory prediction does not come with uncertainties. For the first two neutrino mass campaigns, a conservative estimate was made.

To this end, the newer theoretical calculations of *Saenz et al.* [54] were compared to earlier ones performed by *Fackler et al.* [65]. While the ground state variance matches to $\mathcal{O}(1\,\%)$, the excited states show larger discrepancies. This difference is taken as estimate for the FSD uncertainty [66]. The corresponding procedure should:

- vary the ground state variance by $1\,\%$,
- vary the excited states variance by $4\,\%$, and
- impact the total ground state probability P_{gs} by $1\,\%$.

This is achieved by fluctuating the probability of each state P_f uncorrelated with the neighbour state with defined standard deviation where the standard deviation depends on whether the f belongs to the ground state or is an excited state. The defined variance uncertainties are achieved by fluctuating each ground state probability with an uncertainty of $4\,\%$ and each excited state probability with $18\,\%$ uncertainty. After the bin-to-bin fluctuation, the ground and excited states are re-normalized to their initial probabilities P_{gs} and P_{es}.

The new ground state probability P'_{gs} is then drawn randomly according to its uncertainty of $0.01\,P_{\text{gs}}$ and all state probabilities are modified accordingly:

$$P_f \to P'_f = \begin{cases} P_f \cdot \frac{P'_{\text{gs}}}{P_{\text{gs}}} & \text{if } f \text{ is a ground state} \\ P_f \cdot \frac{P_{\text{es}}+P_{\text{gs}}-P'_{\text{gs}}}{P_{\text{es}}} & \text{if } f \text{ is an excited state.} \end{cases} \tag{3.53}$$

It can be convenient to define a quantity F, the *FSD onset* [67], such that:

$$\begin{aligned} F := +1 \quad &\Leftrightarrow P'_{\text{gs}} = 1 \quad \wedge \quad P'_{\text{es}} = 0 \\ F := 0 \quad &\Leftrightarrow P'_{\text{gs}} = P_{\text{gs}} \quad \wedge \quad P'_{\text{es}} = P_{\text{es}} \\ F := -1 \quad &\Leftrightarrow P'_{\text{gs}} = 0 \quad \wedge \quad P'_{\text{es}} = 1. \end{aligned} \tag{3.54}$$

Using the onset, the probabilities are modified as follows:

$$P'_{\text{gs}}(F) = \frac{(1+F)P_{\text{gs}}(P_{\text{gs}}+P_{\text{es}})}{(1+F)P_{\text{gs}}+(1-F)P_{\text{es}}}, \tag{3.55}$$

$$P'_{\text{es}}(F) = \frac{(1-F)P_{\text{es}}(P_{\text{gs}}+P_{\text{es}})}{(1+F)P_{\text{gs}}+(1-F)P_{\text{es}}}. \tag{3.56}$$

This relationship is shown in fig. 3.21. An uncertainty on P_{gs} of $1\,\%$ corresponds to an uncertainty on F of $\sigma_F = 0.012$ using Gaussian error propagation.

Figure 3.22 shows the impact of the uncertainty on the final states distribution. Both the ground state and the excited states uncertainty is dominated by the bin-to-bin fluctuation which is much more clearly visible in the excited states, where the overall uncertainty is larger

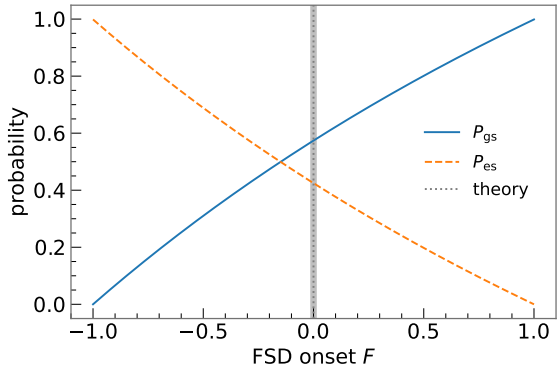
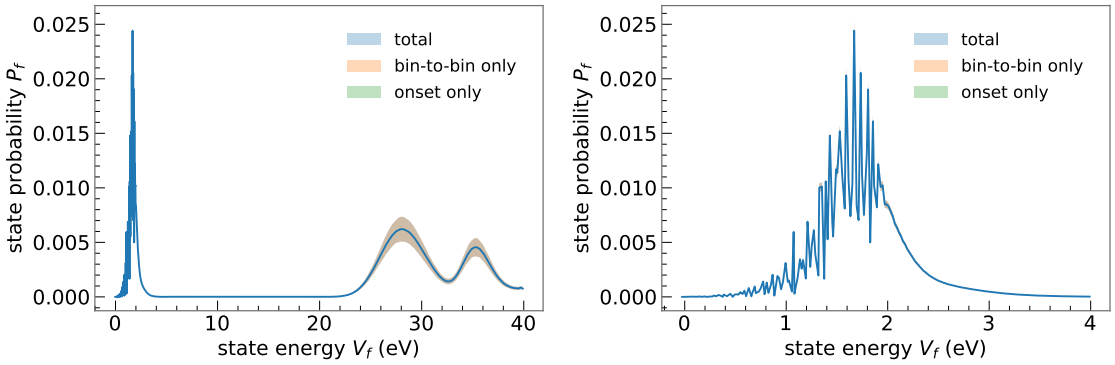


Figure 3.21: Relationship between the FSD onset F and the ground (blue solid line) and excited states (orange dashed line) probability P_{gs} and P_{es} respectively. The gray dotted line indicates the theory prediction, and the gray band the uncertainty corresponding to a $1\,\%$ uncertainty on P_{gs}.



(a) KATRIN energy interval

(b) ground state

Figure 3.22: Uncertainty band of the final states distribution for the full KATRIN energy interval (a) and a zoom-in to the ground state (b). The total uncertainty is fully dominated by the bin-to-bin fluctuations designed to impact the ground state and excited states variance of the FSD.

by design. As the excited states only contribute a minor part to the signal in the KATRIN energy interval used for the first two measurement campaigns, this is not an issue for the m_ν analysis. However, this procedure is known to overestimate the uncertainty of the FSD [68] and it will therefore be revised for future analyses with higher sensitivity to m_ν.

Activity fluctuations

Any instability in the gas density, the composition of the tritium isotopologues, the magnetic fields impacting the acceptance angle or the detector efficiency directly impact the constant prefactor C of our signal model in eq. (3.47). As both the magnetic fields and the detector are very stable over time, this effect is dominated by fluctuations of the source activity.

Any change of the source activity between the measurement points i leads to an overdispersion of the measured counts N_i compared to plain statistical fluctuations. Ideally, one would measure the relative activity at every point A_i and correct the model per-point accordingly $C \to A_i \cdot C = C_i$. The uncertainty on the individual A_i can then be included as additional uncertainty on the measured counts.

There are two methods to calculate A_i:

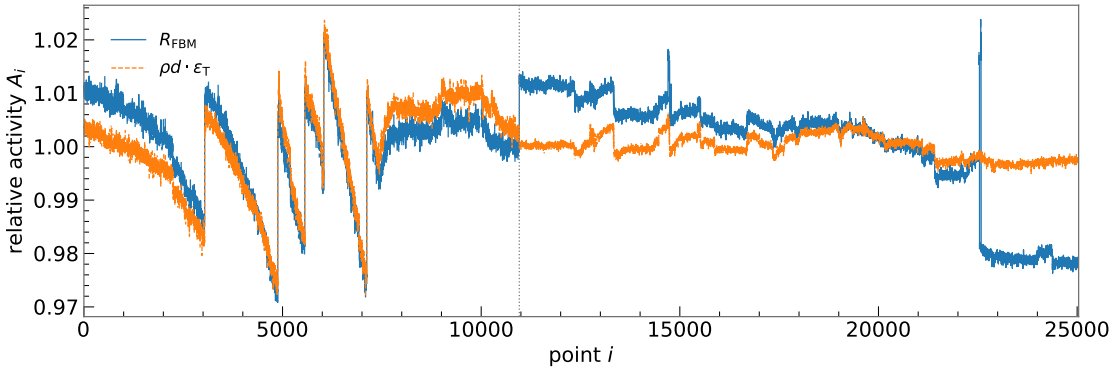
- Use the product of the activity parameters measured directly: $A_i \propto \rho d_i \cdot \epsilon_{\mathrm{T},i}$, or
- make use of an activity monitoring device such as the rate of the FBM: $A_i \propto R_{\mathrm{FBM},i}$.

Both methods are shown for the first two neutrino mass measurement phases in fig. 3.23a.

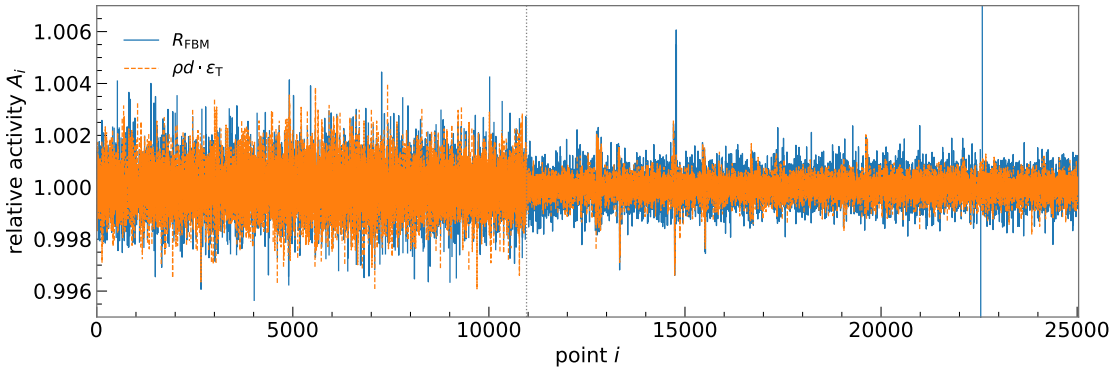
However, in practice, the accuracy of the measured A_i is not clear due to the drift of the FBM and the slow control measurements missing for some points i with short measurement time. Therefore, to avoid any artificial bias, the model is not corrected and only the uncertainty on A_i is taken into account.

During a measurement campaign, typically all points at the same set voltage can be combined (see section 4.2.2 for more details). Therefore, any activity change between scans averages out exactly during combination and is not of interest for uncertainty propagation. Instead, only the changes within a single scan are relevant. Figure 3.23b shows the relative activity, normalizing the mean activity of each scan to 1. Intuitively, the uncertainty of each point i can be increased by the standard deviation σ of the detrended activity changes as shown in fig. 3.24. When combining n points, the effect is averaged out, therefore the added uncertainty on the combined point is the error of the mean

$$\frac{\sigma}{\sqrt{n}}. \tag{3.57}$$



(a) overall changes



(b) scan wise changes removed

Figure 3.23: Relative activity during the first (left of the gray dotted line) and second measurement campaign using the product of the activity parameters (orange dashed line) and the FBM rate (blue solid line). For both campaigns a downwards drift of the FBM rate compared to the slow control measurement is visible. (a) displays the overall changes including long-term effects while (b) removes any changes from scan to scan and only displays the variation within each scan, therefore showing relative changes smaller by almost one order of magnitude. The jump in the FBM relative activity around point $22\,500$ is related to large noise after a PC reset [69].

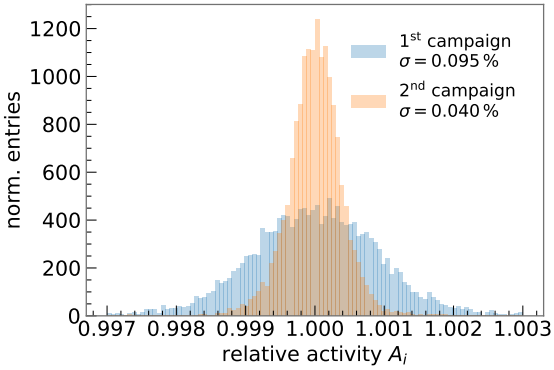


Figure 3.24: Projection of the relative activity without scan wise changes shown in fig. 3.23b using the slow control values as estimate for the first (blue) and second (orange) campaign.

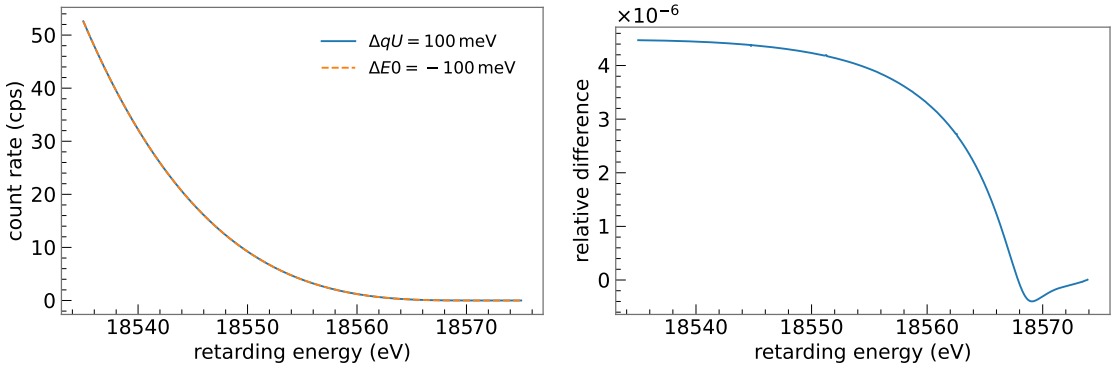
As the activity can be related to a linear drift, such as over multiple periods during the first measurement campaign, c.f. fig. 3.23a, correlations between neighboring points can occur. Therefore, instead of using the plain error of the mean when combining n points, one has to calculate the covariance over n scans for all points $\frac{K_{\text{activity}}}{n}$. It should be noted that this correlated uncertainty is avoided when including the effect of an activity drift in the model. The optimal treatment for upcoming measurement phases is currently under discussion.

Source electric potential

The electric starting potential of the emitted β-electrons is defined by the cold magnetized source plasma [35]. This plasma is created by the large number of free charges. As calculated using eq. (3.14), there are up to 10^{11} decays in the source emitting this amount of electrons and the corresponding $(\text{HeT})^+$ ions. In addition to these primary particles, each electron emits approximately 50 secondary electrons by inelastic scattering on gas molecules in the source [70]. This plasma leads to multiple effects that need to be taken into account.

The first effect to consider is any spatial inhomogeneity of the potential. Radial, and in principle also azimuthal, differences can be included into the model due to the pixel segmentation of the detector. A difference in the starting potential effectively leads to a different retarding potential per pixel qU_{px}. However, as shown in fig. 3.25, a shift in the retarding potential by ΔE is well approximated by a shift of the endpoint with opposite sign $-\Delta E$. Therefore, radial and azimuthal inhomogeneities of the starting potential can be absorbed by a pixel dependent endpoint $E_{0,\text{px}}$.

This is not possible though for any longitudinal inhomogeneities [70]. In principle, a distribution of the starting potential depending on the longitudinal starting position $\Delta qU(z)$ could be taken into account by including it in the model and averaging over all z via integration. How-



(a) absolute **(b)** relative

Figure 3.25: Comparison of the integrated β-spectrum once shifted in the retarding energy by $100\,\mathrm{meV}$ and once in the endpoint by $-100\,\mathrm{meV}$. In the absolute overlay (a) no difference is visible. The relative differences are on the 10^{-6} level (b).

ever, this is neither convenient from a modelling point of view, nor is it currently possible to predict or measure $\Delta qU(z)$ accurately. Instead, the distribution is modelled via two effective parameters: a potential broadening with variance σ_z^2 and a shift of the energy loss function ϵ_z. While any distribution different from a Dirac delta function intuitively introduces a broadening, the ϵ_z parameter is less obvious. It comes from the fact that electrons starting at the back of the source have a higher probability to scatter as they travel longer distances through the gas. This is shown by the distribution of electrons over z for different scatterings in fig. 3.26.

Therefore, if $\Delta qU(z)$ is not symmetric but has a gradient over z, electrons which scatter see a different average starting potential than those that do not. This can be modelled by shifting the energy loss function $f_1(\epsilon) \to f_1(\epsilon - \epsilon_z)$ and the effect of this *eloss shift* parameter on the response function is displayed in fig. 3.27.

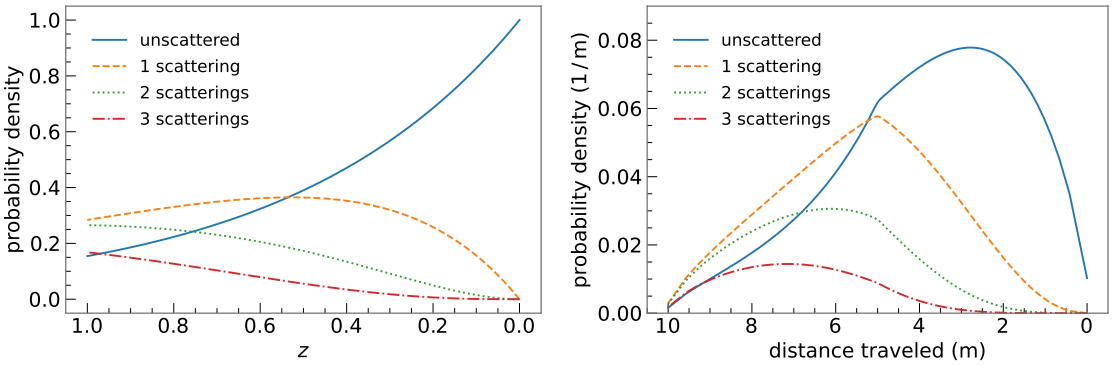
The longitudinal broadening σ_z^2 can be measured using $^{83\mathrm{m}}$Kr and related to an upper limit on ϵ_z via

$$|\epsilon_z| < \frac{\sigma_z}{k} \tag{3.58}$$

where k is an empirical parameter depending on the column density and the measurement interval as both affect the electron distributions shown in fig. 3.26 [70]. The resulting two-dimensional probability distribution is shown in fig. 3.28.

In addition to the spatial inhomogeneities, there can be short-term fluctuations of the overall plasma potential. These can be modelled by including their variance $\sigma^2_{\text{short-term}}$ as an additional broadening in the model. In a measurement, this is indistinguishable from σ_z^2 and one would simply measure the sum of both broadening variances.

Finally, the source electric potential can drift over longer periods of time. While this is not an issue when only analyzing single scans, it must be taken into account when combining



(a) fraction of gas seen

(b) real distance traveled

Figure 3.26: Longitudinal distribution of electrons for different scatterings assuming a column density of $4.2 \times 10^{21}\,\mathrm{m}^{-2}$. As expected, electrons that travel a longer distance in the source, i.e. those from the back of the source, scatter more often.

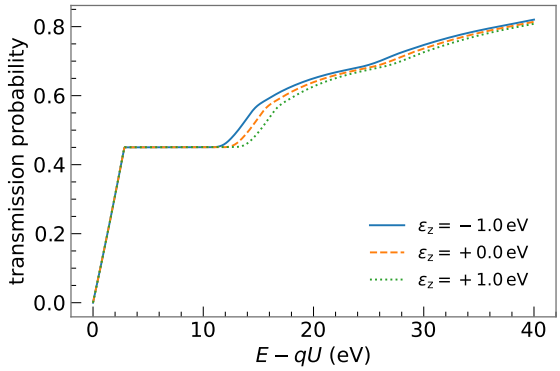


Figure 3.27: Response function for different energy loss shifts. By definition, the transmission part is not affected, and only the scattering part is shifted by ϵ_z

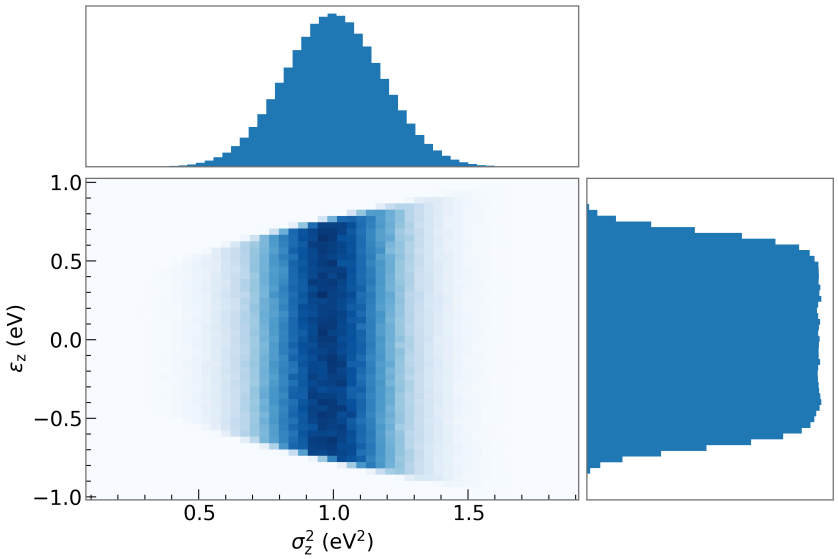


Figure 3.28: Probability density function of the longitudinal plasma parameters. The broadening variance σ_z^2 is sampled from a normal distribution with mean of $1\,\mathrm{eV}^2$ and standard deviation $0.2\,\mathrm{eV}^2$. For each sampled $\sigma_{z,i}^2$, the corresponding $\epsilon_{z,i}$ is sampled from a uniform distribution with range $\left[-\frac{\sigma_{z,i}}{k}, +\frac{\sigma_{z,i}}{k}\right]$ and $k = 1.3$.

data from longer time periods such as a full measurement campaign. To asses this effect, the energy scale of the experiment must be monitored over the campaign. One can then compare the variance of the measured distribution $\sigma^2_{\text{measured}}$ with the one expected from statistical fluctuations σ^2_{stat} to retrieve the overdispersion related to the long-term drift

$$\sigma^2_{\text{long-term}} = \sigma^2_{\text{measured}} - \sigma^2_{\text{stat}}. \tag{3.59}$$

One straightforward approach to asses the time evolution of the energy scale is to evaluate the endpoint value for each individual scan within the measurement campaign.

To summarize, the source electric potential can be described by two parameters:

- a total broadening with variance $\sigma^2_{\text{src}} = \sigma^2_{\text{z}} + \sigma^2_{\text{short-term}} + \sigma^2_{\text{long-term}}$ which can for example be emulated in the final state distribution,
- and the energy loss shift ϵ_{z} describing the longitudinal inhomogeneity modifying the response function.

Maximum and source magnetic field

The maximum and source magnetic field affects our model calculation in multiple ways. Firstly, the solid angle and thus the overall signal amplitude is different. While this normalization effect does not impact the neutrino mass analysis, the scattering probabilities also differ due to the angular integration boundary changing. This difference in scattering probabilities changes the shape of the response function and must be taken into account when propagating uncertainties.

In addition, both magnetic fields also impact the transmission function already in the isotropic case as shown in eq. (3.31) leading to another shape effect to consider when propagating uncertainties. The impact of B_{max} and B_{src} on the transmission function is shown in fig. 3.29. While B_{max} directly defines the energy resolution according to eq. (3.5), B_{src} only has a minor impact on the transmission shape.

Analyzing plane fields

Next we consider the analyzing plane fields, namely the retarding potential qU and the magnetic field B_{ana}. Any overall shift of qU can be absorbed by the endpoint E_0 as discussed for the source electric potential and shown in fig. 3.25 and is thus not relevant for uncertainty propagation. However, there could also be a variation of the retarding potential within a single pixel with variance σ^2_{qU}. While this effect is negligible for the field setting used for the first two measurement campaigns, it needs to be considered for sub-sequent measurement campaigns in the *shifted analyzing plane* setting discussed in detail in section 3.5.2. In case the potential broadening needs to be included, it can be treated equivalently to the source potential broadenings and the variances are simply added.

Like B_{max}, the magnetic field in the analzing plane defines the energy resolution of the MAC-E filter, c.f. eq. (3.5), and therefore directly impacts the shape of the transmission function as shown in fig. 3.29.

Background overdispersion

To understand any possible background overdispersion, we have to first understand the background processes at KATRIN [71].

The dominant background component scales with the volume of the flux-tube in the spectrometer. Its characteristics are in good agreement with low energy electrons generated within the main spectrometer volume. In case these are created within the flux tube after the analyzing

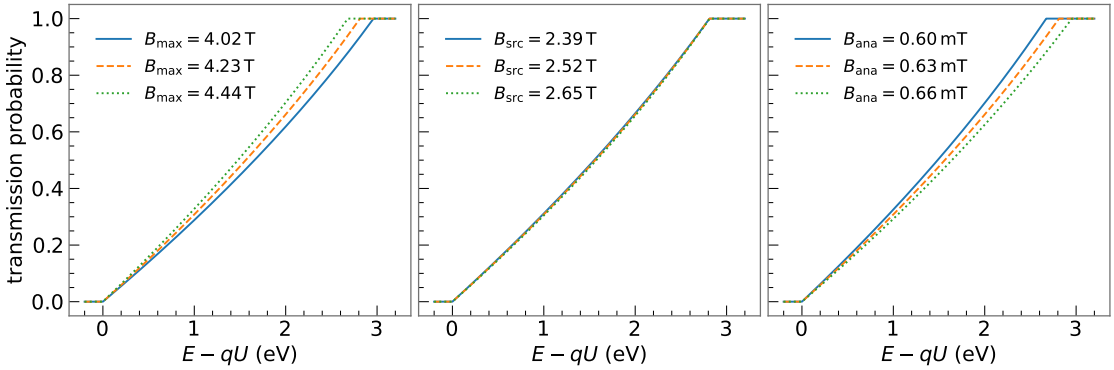


Figure 3.29: Impact of maximum, source and analyzing plane magnetic field on the transmission function. If not stated otherwise, the values assumed are $B_{\text{max}} = 4.23\,\text{T}$, $B_{\text{src}} = 2.52\,\text{T}$, and $B_{\text{ana}} = 0.63\,\text{mT}$.

plane, they are accelerated towards the detector and counted, indistinguishable from the signal electrons. A possible candidate to generate these electrons are highly excited atoms, called *Rydberg atoms* [72], generated by sputtering in the main spectrometer surface induced by the decay-chain of ^{210}Pb in the spectrometer surface. These Rydberg atoms can then release their electrons with energies typically $< 100\,\text{meV}$ via ionization possibly induced by the black body radiation of the main spectrometer surface. This process is Poisson distributed as the sputtering events are independent, the underlying rate is constant over time and each sputtering event creates at most a single electron that reaches the detector.

Another background comes from the intrinsic background of the FPD, dominated by environmental radiation and thus simply increasing the Poissonian background rate.

The final major background component is related to stored high-energy electrons in the main spectrometer [73]. Electrons with sufficient transversal energy created within the main spectrometer are trapped due to the magnetic bottle created by the MAC-E filter field configuration. Storage conditions can mainly be broken via multiple scatterings of the electron with the residual gas in the spectrometer. This is highly unlikely due to the ultra high vacuum conditions, leading to storage times up to several hours. During the storage, the high energy electron creates secondary electrons by ionization of the residuals gas with each scattering process. These low-energy electrons are then eventually guided to the detector, analogous to the Rydberg electrons, if they were created within the flux-tube past the analyzing plane. As a single primary electron creates multiple secondaries during its storage, this leads to an elevation of the background during this time, and a time-correlation of this background rate. Therefore, this process is not Poissonian and leads to an overdispersion compared to the expected Poisson variance.

To extract the overdispersion, one can measure the background distribution and model it with

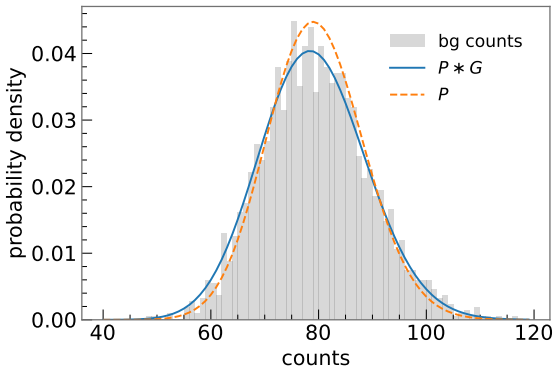


Figure 3.30: Overdispersion of the measured background counts modelled by the convolution of a Poisson distribution with a Gaussian $P * G$ compared to a plain Poissonian P.

a convolution of a Poisson distribution P with a Gaussian distribution G:

$$(P * G)(N, t; \lambda, \sigma_{\text{np}}) = c \int_0^\infty \frac{\mu^N \exp(-\mu)}{N!} \cdot \frac{1}{\sqrt{2\pi\sigma_{\text{np}}^2}} \cdot \exp\left(-\frac{(\lambda \cdot t - \mu)^2}{2\sigma_{\text{np}}^2}\right) \mathrm{d}\mu \quad (3.60)$$

with the Poisson rate λ, the Gaussian standard deviation σ_{np}, the measured counts N with corresponding measurement time t and the factor c ensuring normalization. The parameters λ and σ_{np} are then extracted from the background data using a distribution fit as visualized in fig. 3.30. This Gaussian widening must then be taken into account when analyzing neutrino mass data, effectively leading to an increased statistical uncertainty.

Background retarding energy dependence

While the background rate is assumed to be constant over the retarding energy for the small measurement interval in neutrino mass campaigns, measurements over an interval spanning several keV show a clear dependence of $r_{\text{bg}}(qU)$ [71]. Therefore, it is difficult to exclude any dependence also in the small retarding energy interval. To approximate any shape effect in this small qU interval, one can extend r_{bg} from a constant to a linear function

$$r_{\text{bg}}(qU) = r_{\text{bg}} + b_{qU} \cdot (qU - 18575) \quad (3.61)$$

with the background retarding energy slope parameter b_{qU}.

This slope parameter can be measured with an empty source [74, 75] and then included in the neutrino mass analysis.

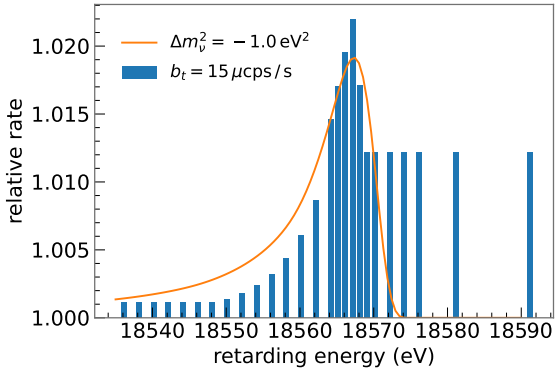


Figure 3.31: Effect of the Penning background compared to a shift of m_ν^2 to more negative values using the measurement time distribution and model parameters similar to the second measurement campaign.

Penning induced background

Another background systematic is related to the Penning trap between pre- and main spectrometer described in section 3.2. Discharges of this Penning trap create positive ions which are accelerated into the main spectrometer. These ions then create background electrons via ionization of residual gas or by creating Rydgerg atoms. The number of background electrons is proportional to the number of electrons stored in the trap. Electrons in the trap accumulate approximately linearly over time for short time intervals [40], leading to a linear increase of the background rate:

$$r_{\mathrm{bg}}(t) = r_{\mathrm{bg}} + b_t \cdot t \tag{3.62}$$

with the Penning slope parameter b_t. As the trap is emptied when the voltage set point is changed, the average background rate within a measurement point of duration T is

$$r_{\mathrm{bg}}(T) = \frac{1}{T}\int_0^T r_{\mathrm{bg}} + b_t \cdot t \,\mathrm{d}t = r_{\mathrm{bg}} + \frac{1}{2} b_t \cdot T. \tag{3.63}$$

Therefore, the Penning induced background rate directly follows the measuring time distribution. Since the MTD is optimized to follow the neutrino mass signal, this background component can emulate a shift of m_ν^2 to more negative values as shown in fig. 3.31.

Detector systematics

Final component to discuss are systematics related to the focal plane detector [76, 77]. All electrons creating an energy deposition that falls within a predefined region of interest (ROI), $14\,\mathrm{keV}$ to $32\,\mathrm{keV}$ for the first two measurement phases [43, 66], are counted. Therefore, any

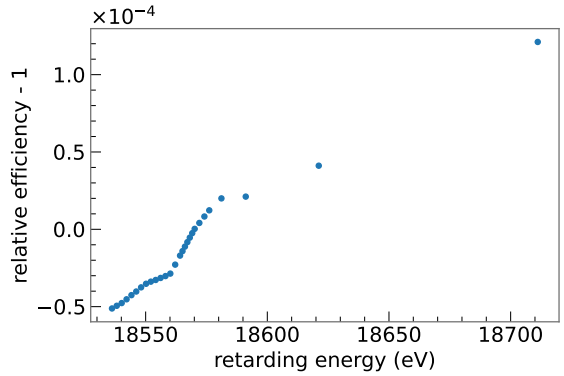


Figure 3.32: Relative detector efficiency depending on the retarding energy set at the main spectrometer for pixel zero of the second neutrino mass campaign. The effect itself is on the 10^{-4} scale with uncertainties at least another order of magnitude smaller.

effect that changes the fraction of electrons within the ROI depending on the voltage set point directly impacts the spectral shape. There are four known effects of this kind:

- The ROI is not shifted according to the retarding energy. Therefore, the comparably small change of electron energies with different retarding potential leads to a qU dependent ROI coverage. This leads to a change of the detector efficiency $\epsilon_{\text{detector}}(qU)$ smaller than $1\,\%$ per keV which is modelled by a linear slope.

- Pile-up occurs when multiple electrons hit the detector at the same time and are not resolved as individual events. Instead of measuring two events with energies E_1 and E_2, a single event with energy $E_1 + E_2$ is detected. If $E_1 + E_2$ exceeds the ROI, neither of the two electrons is counted. The probability for pile-up to occur depends on the total rate, and therefore also on the retarding potential.

- In case of back-scattering, an electron that hits the detector is reflected instead of being absorbed and, if not reflected back by the post acceleration or spectrometer fields, is subsequently lost to the source. The probability for this to occur depends on the retarding energy as electrons have a higher probability to escape to the source for lower qU values.

- Finally, the FPD resolution and gain can drift over time, effectively also leading to a dependency of the detection efficiency on the voltage set point.

The region of interest coverage as well as pile-up is corrected for using a qU dependent detector efficiency correction shown in fig. 3.32 while the effect of back-scattering and resolution drifts is not included in the model. The uncertainties and biases related to these effects are too small to be of any importance for the first two neutrino mass campaigns and will not be discussed further over the course of this thesis.

3.5 Neutrino mass measurement phases

We now briefly describe the KATRIN neutrino mass measurement phases, starting with the first two campaigns followed by any subsequent ones in an updated field setting.

3.5.1 First two campaigns

The first two KATRIN neutrino mass (KNM) measurements took place in spring and autumn 2019 respectively. This section gives a brief overview of the experimental settings and of the statistics accumulated with these two campaigns. A detailed analysis of the data will take place in chapter 6. Table 3.1 gives a brief summary of the quantities discussed in the following paragraphs.

The first measurement campaign also was the first time the KATRIN system was exposed to large amounts of tritium. This caused radioactive reactions of tritium with the so far unexposed inner metal surface of the injection capillary [78], leading to a downward drift of the gas density in the source and therefore also the total activity. This effect was too large to handle at nominal gas density, but could be counteracted by repeatedly increasing the injection pressure at a reduced gas density of about $22.2\,\%$, keeping the overall stability within $\pm 2\,\%$. Both the downwards trend and the upward jumps after an increase of the injection pressure are visible in fig. 3.23a.

Over a time period of roughly one month, a total of more than 300 neutrino mass scans were taken out of which 274 so-called *golden* scans were selected for the final analysis. The time spent in the analysis interval of the neutrino mass, amounts to $521.7\,\mathrm{h}$ and a total of $2.0\,\mathrm{million}$ electrons were collected in this time with an average background rate of $292\,\mathrm{mcps}$.

Before the second measurement campaign, two major efforts were taken to improve the signal to background ratio. Firstly, the main spectrometer was baked at high temperatures to reduce the Rydberg background [71, 79]. In addition, the system was flushed continuously with large amounts of tritium, exposing all components to the radioactive gas, and thus stabilizing the system [80, 81]. Both measures were successful and the gas density could be increased up to $84.5\,\%$ of the maximum with a decreased average background of $220\,\mathrm{mcps}$. With these parameters, a total of 361 golden scans were taken over a period of one and a half months. This works out to $694.3\,\mathrm{h}$ spent in the analysis interval with a total of $4.3\,\mathrm{million}$ electrons collected, more than double the number of the first campaign [43].

Table 3.1: Overview of the first two neutrino mass measurement campaigns.

	1st campaign	2nd campaign
date	2019-04-10 to 2019-05-13	2019-09-27 to 2019-11-14
number of *golden* scans	274	361
time spent in analysis interval	521.7 h	694.3 h
gas density	22.2 %	84.5 %
background rate	292 mcps	220 mcps
total electrons in analysis interval	2.0 million	4.3 million

3.5.2 Sub-sequent measurement phases using the shifted analyzing plane

After the first two measurement campaigns, a short campaign was run to test a new field setting meant to reduce the overall background rate, the so-called shifted analyzing plane (SAP) [82]. As the dominant background component scales with the volume in the fluxtube after the analyzing plane, an effective mitigation strategy is to reduce this volume. In the SAP setting, this is done by shifting the analyzing plane from the center of the main spectrometer downstream towards the detector as shown in fig. 3.33. Using this technique, the background rate was decreased by a factor of two. However, the special field setting leads to a larger inhomogeneity of B_{ana} and qU over the detector as shown in fig. 3.34 compared to the *symmetric* setting used for KNM1 and KNM2.

This larger inhomogeneity makes it absolutely necessary to split the model into multiple parts, one for each *patch* with similar field values, leading to a more complex analysis. In addition, the strong potential gradient also introduces the need of an additional broadening σ^2_{qU}.

The field setting, as well as its implications on the data analysis, were first tested during KNM3a and could be compared directly to the subsequent measurement phase, KNM3b, in the symmetric setting. As no major issues were discovered, all further measurement campaigns were taken in the SAP setting. An overview of the settings of all measurement campaigns completed before 2022 is summarized in table 3.2 and the cumulative number of electrons is shown in fig. 3.35.

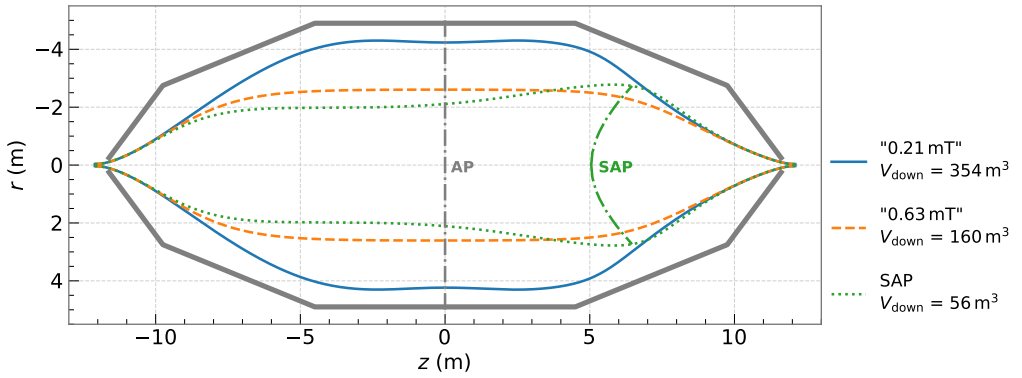
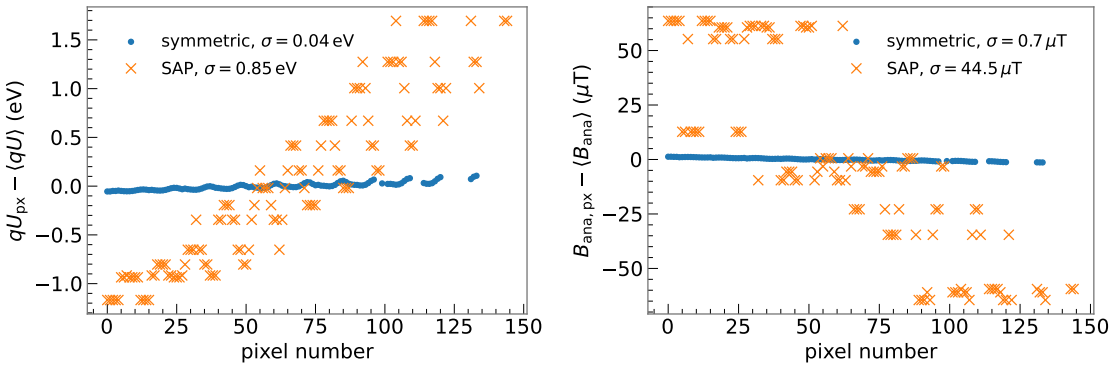


Figure 3.33: Fluxtube and analyzing plane for different field settings. Both the $0.21\ \mathrm{mT}$ and the $0.63\ \mathrm{mT}$ are in the symmetric setting where the analyzing plane is in the center of the main spectrometer (gray dashdotted line). The $0.21\ \mathrm{mT}$ one is the design setting, however the $0.63\ \mathrm{mT}$ setting was used for the first two neutrino mass campaigns to already reduce the fluxtube volume for background mitigation purposes. In the SAP setting, the analyzing plane is shifted towards the detector and therefore further reduces the volume of the fluxtube after the analyzing plane (green dashdotted line). Figure adapted from [82].



(a) retarding potential **(b)** magnetic field

Figure 3.34: Retarding potential (a) and magnetic field (b) variation in the analyzing plane for the symmetric and shifted analyzing plane setting. For both fields the inhomogeneity in the SAP setting is at least one order of magnitude larger.

Table 3.2: Overview of field setting and gas density for neutrino mass measurement campaigns completed before 2022.

campaign	field setting	gas density
KNM1	symmetric	$22.2\,\%$
KNM2	symmetric	$84.5\,\%$
KNM3a	shifted analyzing plane	$41.3\,\%$
KNM3b	symmetric	$75.1\,\%$
KNM4	shifted analyzing plane	$75.4\,\%$
KNM5	shifted analyzing plane	$75.5\,\%$
KNM6	shifted analyzing plane	$74.1\,\%$

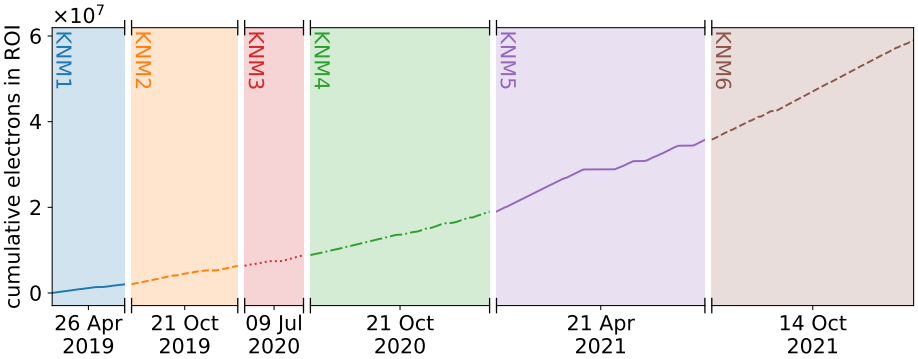


Figure 3.35: Cumulative electrons detected in the analysis interval of the neutrino mass measurement phases completed before 2022. Campaigns after KNM2 are currently being analyzed (January 2022) and the analysis interval as well as the scan selection may still change. Figure adapted from Alessandro Schwemmer.

Chapter 4

Analysis strategies

Next, we discuss the general analysis strategies used to infer the neutrino mass value and uncertainty from the KATRIN neutrino mass scans.

4.1 General KATRIN likelihood

Widely used methods to infer parameter information from measurement data using a model μ depending on parameters $\boldsymbol{\theta}$ make use of the likelihood function $\mathcal{L}(\mu(\boldsymbol{\theta}); \text{data})$ which describes how well this set of parameters $\boldsymbol{\theta}$ describes the measured data. This section derives the likelihood function for KATRIN which will subsequently be used to infer information on m_ν^2 from the data.

As described in section 3.3.4, KATRIN measures the number of counts N_i at a given retarding energy qU_i with measurement time t_i. These counts can be predicted by the model $\mu_i = \mu(qU_i)$ as defined in eq. (3.50). Excluding systematic effects and activity drifts, this count expectation is constant within each measurement point. In addition, the individual decays are statistically independent. Therefore, the probability mass function (PMF) describing how probable it is to measure N_i counts given the expectation μ_i is the Poisson distribution

$$P(N_i; \mu_i) = \frac{\mu_i^{N_i}}{N!} \cdot \exp\left(-\mu_i\right). \tag{4.1}$$

As the individual measurements at different qU_i are independent of one another, the joint PMF of all points $\boldsymbol{i}$ with voltage set points $\boldsymbol{qU}$, times $\boldsymbol{t}$, counts $\boldsymbol{N}$ and corresponding model predictions $\boldsymbol{\mu}$ is the product of the individual Poissonians

$$P_{\text{joint}}(\boldsymbol{N}; \boldsymbol{\mu}) = \prod_i P\left(N_i; \mu_i\right). \tag{4.2}$$

This gives us a function that describes the probability to measure $\boldsymbol{N}$ given our model prediction $\boldsymbol{\mu}$. In practice, the measured counts $\boldsymbol{N}$ are known and are used to infer information on the model rate $\mu(\boldsymbol{\theta}; qU, t)$ which depends on the parameter values to be inferred. Therefore, the relationship in eq. (4.2) is flipped to arrive at the likelihood

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{N}, \boldsymbol{qU}, \boldsymbol{t}) = \prod_i P(N_i; \mu(\boldsymbol{\theta}; qU_i, t_i)). \tag{4.3}$$

As the likelihood is in general not normalized over $\boldsymbol{\theta}$, it cannot be taken as a probability density function (PDF) as-is, but as a measure of *how likely* a specific set of parameters $\boldsymbol{\theta}$ is compared to another one.

For a large number of counts, the Poisson distribution can be approximated with a Gaussian

$$G(N_i; \mu_i) = \frac{1}{\sqrt{2\pi N_i}} \exp\left(-\frac{(N_i - \mu_i)^2}{2N_i}\right) \tag{4.4}$$

with mean μ_i and standard deviation $\sqrt{\mu_i} \approx \sqrt{N_i}$.

In KATRIN, the parameter of interest to infer from the data is the neutrino mass squared m_ν^2. In addition, the endpoint value, also absorbing any potential offsets, E_0, the overall signal normalization A_S and the constant background rate r_bg must be inferred directly from the data. This gives us the basic model used for a likelihood based analysis

$$\mu(\boldsymbol{\theta}; qU, t) = A_\mathrm{S} \cdot C \cdot \int_{-\infty}^{+\infty} \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E; m_\nu^2, E_0) \cdot R(E; qU)\, \mathrm{d}E + r_\mathrm{bg} \tag{4.5}$$

using the model derived in section 3.3.4 and the parameters $\boldsymbol{\theta} = \{m_\nu^2, E_0, A_\mathrm{S}, r_\mathrm{bg}\}$. At this point we note that it is convenient to fit m_ν^2 instead of m_ν as the likelihood is roughly Gaussian in m_ν^2 which allows using certain simplifications as will become clear in section 4.3.

4.2 Data combination

The likelihood derived in eq. (4.3) assumes the measurement of a single spectrum with i points. In practice, the KATRIN data is segmented both spatially due to the 148 pixels of the FPD and in time as multiple scans are taken. This section discusses how the data can be combined into a single likelihood, while ensuring that the model prediction is still sufficiently accurate.

4.2.1 Detector pixels

First, we discuss combining the spectra from the 148 pixels of the FPD. In principle, each measured spectrum is independent of the other spectra, allowing us to simply extend the

likelihood product by another dimension

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{N}, \boldsymbol{qU}, \boldsymbol{t}) = \prod_{\text{px}} \prod_{i} P(N_{\text{px},i}; \mu(\boldsymbol{\theta}_{\text{px}}; qU_{\text{px},i}, t_i)) \tag{4.6}$$

where the outer product loops over all pixels and the inner one over all voltage set points. In general, some parameters, such as the magnetic field in the analyzing plane B_{ana} and the retarding potential qU depend on the pixel number, and others, like the neutrino mass squared, are equal for pixels. Therefore, the overall parameters $\boldsymbol{\theta}_{\text{px}}$ differ for the pixels, and also the model prediction is calculated for each pixel individually. Consequently, this approach is referred to as the *multi-pixel* combination.

While the multi-pixel combination is the most general, it is also the most difficult and computationally expensive as it has the most data points and model evaluations possible. It is always favourable to reduce the dimensionality as much as possible, while not significantly harming the model accuracy. To do so, the data spectra of multiple pixels within a *patch* are combined into a single spectrum by adding the counts and averaging the times. The model prediction of the patch is then given by sum of the individual model predictions

$$\mu_{\text{patch}} = \sum_{\text{px}} \mu_{\text{px}}. \tag{4.7}$$

which can also be written as

$$\sum_{\text{px}} \mu_{\text{px}} = n_{\text{px}} \cdot \langle \mu_{\text{px}} \rangle \tag{4.8}$$

using the average over all pixel models $\langle \mu_{\text{px}} \rangle$. An exact calculation of this average is to average the individual functions. However, if the differences of the individual functions do not deviate significantly from a linear function, this can be approximated by averaging the parameter values and only evaluating a single model:

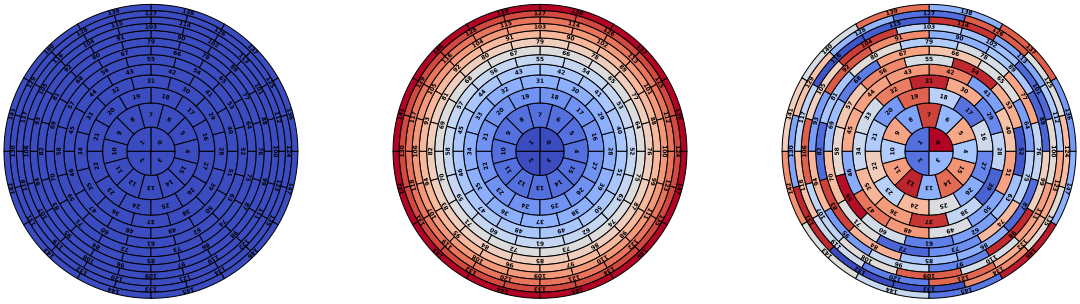
$$\langle \mu_{\text{px}} \rangle = \langle \mu(\boldsymbol{\theta}_{\text{px}}) \rangle \approx \mu(\langle \boldsymbol{\theta}_{\text{px}} \rangle). \tag{4.9}$$

A hybrid approach where the parameters are averaged for parts of the model, such as the differential β-spectrum, and the functions for other parts, such as the transmission function, is often a good compromise.

Using our patch-wise model prediction, our likelihood product is now

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{N}, \boldsymbol{qU}, \boldsymbol{t}) = \prod_{\text{patch}} \prod_{i} P(N_{\text{patch},i}; \mu_{\text{patch}}). \tag{4.10}$$

In the extreme case where each patch consists of a single pixel, we are back at eq. (4.6). The other extreme is to group all pixels into a single *uniform* detector patch. As radial inhomogeneities are typically one order of magnitude larger than azimuthal ones, grouping pixels into rings or ring-like patches is a valid middle ground. Different patch segmentations used in KATRIN data analysis are visualized in fig. 4.1.



(a) uniform **(b)** ring **(c)** pixel

Figure 4.1: Visualization of various detector patch segmentations.

4.2.2 Scans during a single measurement campaign

During a measurement campaign, the β-spectrum is repeatedly measured by scanning the same set of voltage points qU_i defined by the measuring time distribution as shown in fig. 3.16. Each scan typically takes between two and four hours, and there are several hundreds of scans during each campaign as summarized in table 3.1.

In full analogy to the pixel segmentation, we can write the complete likelihood product by adding another dimension

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{N}, \boldsymbol{qU}, \boldsymbol{t}) = \prod_s \prod_i P(N_{s,i}; \mu(\boldsymbol{\theta}_s; qU_{s,i}, t_{s,i})) \tag{4.11}$$

where the outer product now loops over the individual scans s and the inner one over the individual voltage set points i.

During a single measurement campaign, all slow control parameters such as the gas density or the magnetic field values, are typically kept constant with excellent stability at or below the percent level. In addition, the reproducibility of the high voltage set points is excellent at the parts-per-million level [41, 42, 83]. Therefore, combining the data by summing the counts and measurement time while averaging the model parameters is an excellent approximation. This approach to combine all scans into a single spectrum is referred to as *stacking* in KATRIN.

In case the MTD is changed during a measurement campaign, but all other slow control parameters remain stable, the individual points $qU_{s,i}$ can no longer be grouped simply by their index i, but a more involved clustering algorithm is required. While in the regular stacking case, when the MTD consists of n_{MTD} points, the stacked spectrum also always consists of n_{MTD} points, the clustering of m MTDs, each with $n_{\text{MTD},m}$ points leads to a combined spectrum with

$$\max_m(n_{\text{MTD},m}) \leq n_{\text{clustered}} \leq \sum_m n_{\text{MTD},m} \tag{4.12}$$

points. Note that in this *clustering* approach still a single averaged model is used, which is only evaluated for more retarding potentials. Clustering spectra that all follow the same MTD leads to the same result as the naive stacking.

4.2.3 Measurement campaigns

While the slow control parameters are kept stable during a single measurement campaign, this must not be true for different campaigns. This can for example be seen in the drastically different gas density between the 1st and 2nd measurement phase, c.f. table 3.1. Therefore, clustering or stacking spectra from different periods is not possible in general.

Instead, the individual periods can be clustered (stacked), but the overall likelihood must be expanded by the *period* dimension

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{N}, \boldsymbol{qU}, \boldsymbol{t}) = \prod_{\text{period}} \prod_{i} P(N_{\text{period},i}; \mu(\boldsymbol{\theta}_{\text{period}}; qU_{\text{period},i}, t_{\text{period},i})) \tag{4.13}$$

where the outer product now loops over all measurement campaigns and the inner one over the clustered (stacked) points within this campaign. This combination approach is significantly less harmful than treating each scan independently as indicated in eq. (4.11) since there are only a few periods, but several hundreds of scans in each period.

4.2.4 Combined likelihood function

Using the methods described in the previous sub-sections, we can now write out a generic likelihood function to combine KATRIN data:

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{N}, \boldsymbol{qU}, \boldsymbol{t}) = \prod_{\text{period}} \prod_{\text{patch}} \prod_{i} P(N_{\text{period,patch},i}; \mu(\boldsymbol{\theta}_{\text{period, patch}}; qU_{\text{period,patch},i}, t_{\text{period},i})). \tag{4.14}$$

This allows for model parameters to differ between each measurement campaign as well as each detector patch. Counts are summed over patches and over all scans within a period for each point i. The measurement time is equal for all detector patches, but summed over the individual scans within a period. Finally, the retarding energies are averaged within a patch and clustered within a campaign, where the corresponding set points that fall into the same cluster are also averaged.

Any extreme case can be seen as a limit of this likelihood function. For example, analyzing a single scan with a uniform detector simply means that the two outer products only have one factor. On the other extreme, one could associate each individual scan with its own period and

each pixel with its own patch. In practice, one will typically group all measurement campaigns into their own period, and either use a uniform detector or a radial segmentation for the patches.

4.2.5 Effects of model averaging

To conclude this section, we briefly discuss the effects of averaging on the model and why it is in general not possible to average over all periods and detector patches.

When averaging the input parameter values, one effectively neglects a broadening effect. If this broadening is in the retarding energy dimension, it directly translates to a negative m_ν^2 bias. In any case, the model is not a direct representation of the data, which leads to worse $-\log\mathcal{L}$ values. However, when neither the m_ν^2 nor the $-\log\mathcal{L}$ value is significantly biased, averaging parameter values significantly speeds up the calculation time as the number of points in the likelihood and the number of model evaluations is reduced.

Averaging the corresponding functions instead of the parameter values includes the broadening effect into the model. This means in general m_ν^2 and $-\log\mathcal{L}$ remain unbiased. But, the broadening leads to a worse effective (energy) resolution thus slightly worsening the overall sensitivity. The computational benefit is less compared to directly averaging the parameter values as the number of model evaluations is not reduced, only the number of points in the likelihood is.

When constructing the combined likelihood, the general approach is to average as much as possible without biasing m_ν^2 and $-\log\mathcal{L}$.

4.3 Parameter inference and interval estimation

Having defined the likelihood function, we can now make use of it to estimate which parameters best describe our data (parameter inference) as well as which set of parameters is allowed given the measurement (interval estimation).

4.3.1 Method of maximum likelihood

We start with probably the most widely applied method based on the likelihood function: the method of maximum likelihood. As the likelihood $\mathcal{L}(\boldsymbol{\theta})$ is a measure how probable a given set of parameter values $\boldsymbol{\theta}$ is given the measured data, we can find the most likely values

by maximizing $\mathcal{L}(\boldsymbol{\theta})$. This is mathematically equivalent to minimizing $-\log\mathcal{L}(\boldsymbol{\theta})$ which is numerically favourable for two reasons: First, the product of many numbers less than one quickly leads to numerical zero, whereas the sum of multiple numbers of $\mathcal{O}(1)$ is numerically stable. Secondly, without loss of generality, most algorithms are designed for minimization, not maximization.

At this point it is worth pointing out that minimizing $-\log\mathcal{L}$ is equivalent to the well known χ^2 minimization in case the PDF of each data point is a Gaussian:

$$-\log\mathcal{L} = -\log\prod_i \frac{1}{\sqrt{2\pi N_i}}\exp\left(-\frac{(N_i-\mu_i)^2}{2N_i}\right) \tag{4.15}$$

$$= -\sum_i \log\frac{1}{\sqrt{2\pi N_i}}\exp\left(-\frac{(N_i-\mu_i)^2}{2N_i}\right) \tag{4.16}$$

$$\propto -\sum_i \log\exp\left(-\frac{(N_i-\mu_i)^2}{2N_i}\right) \tag{4.17}$$

$$= \frac{1}{2}\sum_i \frac{(N_i-\mu_i)^2}{N_i} \tag{4.18}$$

$$= \frac{1}{2}\chi^2. \tag{4.19}$$

Knowing the most likely values of $\boldsymbol{\theta}$ by minimizing $-\log\mathcal{L}(\boldsymbol{\theta})$, we can now ask which other values are generally compatible with the data. In Frequentist probability, this question is answered by finding the confidence interval of a given parameter θ at a specific confidence level (CL) α. The confidence interval is designed such that α of the intervals cover the true value of θ would the measurement be repeated infinitely often. Note that it by design cannot make any statement on the outcome of a specific measurement, but is only valid in the context of an ensemble with infinite samples.

With this definition of the confidence interval, we can derive a Monte Carlo method to approximate the estimation of it. Given our best-estimate of the parameter values and the PDF of each data point in the spectrum, we can generate N random statistically fluctuated data sets. We then fit each of these samples and store the best fit values $\boldsymbol{\theta}_{\text{sample}}$. The distribution of each of the parameter values describes the statistical uncertainty and we can cut $1-\frac{\alpha}{2}$ samples from the left and right to retrieve the central confidence interval.

While this Monte Carlo method is general and always applicable, it is also computationally involved as it requires at least several thousands of minimizations. Therefore, one often makes use of Wilks' theorem [84] which states that the distribution of the test statistic

$$-2\log\left(\frac{\mathcal{L}(H_0)}{\mathcal{L}(H_1)}\right) = 2\cdot(\log\mathcal{L}(H_1)-\log\mathcal{L}(H_0)) =: 2\Delta\log\mathcal{L} \geq 0 \tag{4.20}$$

asymptotically approaches the χ^2_n distribution with the null and alternative hypothesis H_0 and H_1 respectively. The number of degrees of freedom n is equal to the difference in dimensionality of the two hypotheses, i.e. the difference in number of free parameters.

One notable example is the $\chi^2_1(x)$ distribution where the cumulative distribution function (CDF) includes $1\,\sigma$ of the values for $x \leq 1$, $2\,\sigma$ for $x \leq 4$ and so forth. In general, to retrieve the value ΔD that corresponds to a confidence level α, one applies the inverse CDF of the χ^2_n distribution for the corresponding degrees of freedom n to α.

In our case of interval estimation, we identify the null hypothesis with the parameter values that best describe the data and the alternative hypothesis with another set of values which we test to see if they fall into our confidence interval. For example, say we have the best fit values $\boldsymbol{\theta}_{\text{best}}$ and are interested in the $1\,\sigma$ confidence interval of a specific parameter θ. Applying Wilks' theorem, this means that we search for values of θ such that $2\Delta \log \mathcal{L} = 1$ Note that the likelihood must be maximized with respect to the remaining parameters $\boldsymbol{\theta}_{\text{other}}$ for each θ to account for possible correlations. This means we have transformed the full Monte Carlo method into a root searching problem of the form

$$D(\theta) - \Delta D = 2 \cdot (\log \mathcal{L}(\theta, \boldsymbol{\theta}_{\text{other,best}}) - \log \mathcal{L}(\boldsymbol{\theta}_{\text{best}})) - \Delta D \stackrel{!}{=} 0 \tag{4.21}$$

where $\Delta D = 1$ for the $1\,\sigma$ interval. This method of maximizing the likelihood with respect to the remaining parameters while scanning along the parameter of interest is referred to as the *profile likelihood*.

To apply the described methods to KATRIN data, we need to define likelihood, model and the inferred parameters:

- Our likelihood is defined by eq. (4.14) using the Poisson (if applicable in its Gaussian approximation) distribution for the measured counts at given retarding energies and measuring times.
- The model is the integrated β-spectrum combined with a background rate.
- In the simplest case, the only parameters with respect to which we need to minimize our likelihood, are those which always need to be inferred from the data, i.e. $\theta = \{m^2_\nu, E_0, A_\text{S}, r_\text{bg}\}$. Typically, m^2_ν is the only parameters shared over period and patch, and E_0, A_S and r_bg are independent for each period and patch, resulting in $1 + 3 \cdot n_\text{patch} \cdot n_\text{period}$ free parameters.

As a first step, we fit our data to retrieve the most likely values of m^2_ν, E_0, A_S and r_bg and the corresponding model prediction as shown in fig. 4.2 together with the normalized residuals [1].

[1] The normalized residuals are defined as the data value minus the model prediction divided by the error of the data point. For Gaussian-distributed data this is simply $\frac{y_\text{data} - y_\text{model}}{\sigma_\text{data}}$. In case of Poisson-distributed data, it can be defined as $2 \cdot \left[N_\text{data} \log\left(\frac{N_\text{data}}{N_\text{model}}\right) + N_\text{model} - N_\text{data} \right]$.

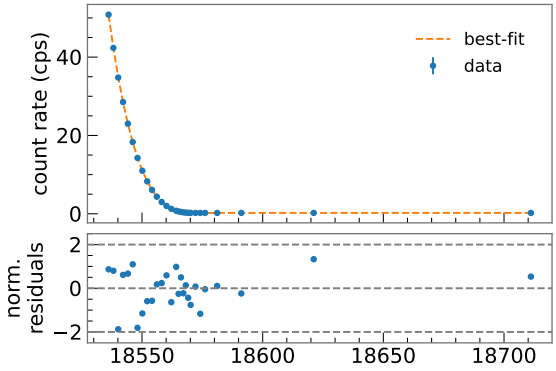


Figure 4.2: Best fit using the maximum likelihood method with normalized residuals on a dummy data set.

Next, we obtain the confidence interval for our parameter(s) of interest, in the KATRIN case m_ν^2, either via the Monte Carlo method or using the profile likelihood as shown in fig. 4.3. In this case, the $1\,\sigma$ interval is identical, showing good applicatility of Wilk's theorem. This can be underpinned by performing a coverage test as shown in fig. 4.4. Here, statistically randomized data is once fitted with a free m_ν^2 value, and once with m_ν^2 fixed to the MC truth of $0\,\mathrm{eV}^2$. When Wilks' theorem holds, as is the case in this study, the distribution of the $\Delta 2 \log \mathcal{L}$ values follows the χ_1^2 distribution.

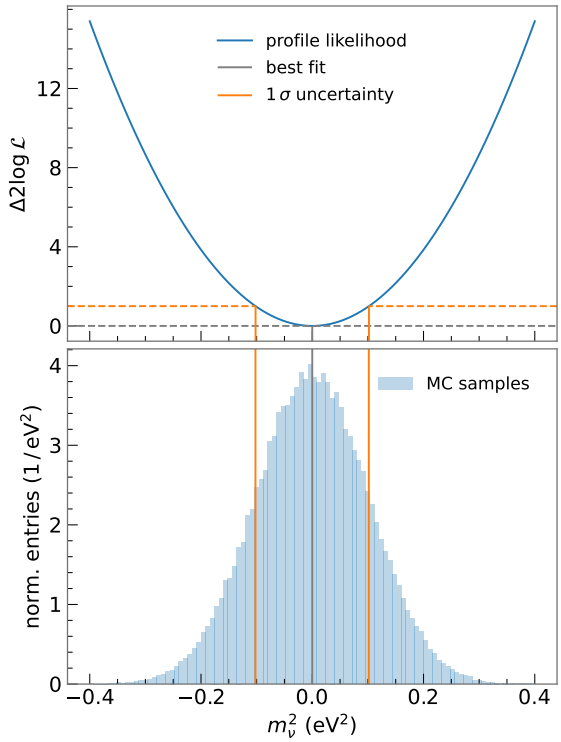


Figure 4.3: $1\,\sigma$ confidence interval of the neutrino mass squared for a dummy dataset using the Monte Carlo method (bottom) and the profile likelihood (top). In this case, the $1\,\sigma$ interval is identical, showing good applicability of Wilks' theorem.

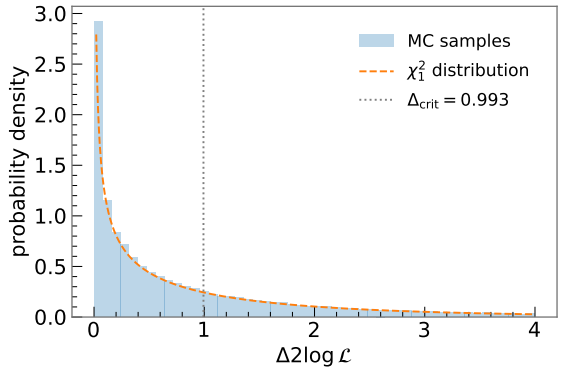


Figure 4.4: Coverage test comparing the $\Delta 2 \log \mathcal{L}$ of Monte Carlo samples to the χ^2_1 distribution. As the two match, Wilks' theorem is fulfilled in this simple case study. The gray dotted line indicates the critical $\Delta 2 \log \mathcal{L}$ value corresponding to the $1\,\sigma$ interval which is compatible with 1 as expected.

4.3.2 Systematic uncertainties

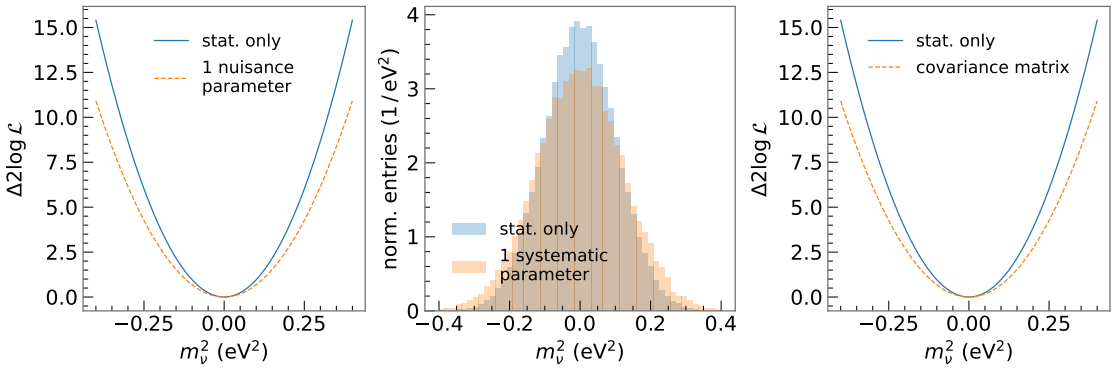
So far we have only considered the statistical error of the individual data points in our interval estimation. In addition to statistical uncertainties on the data, there can be systematic uncertainties on the model estimation. These can often be described by additional parameters $\boldsymbol{\theta}_{\text{syst}}$ where their best-knowledge value $\boldsymbol{\mu}_{\text{syst}}$ comes with a non-negligible uncertainty. We can see that many of the systematic uncertainties in the KATRIN experiment described in section 3.4 are of this form. Typical examples are the column density, the magnetic fields or the source potential parameters, but also background parameters such as the retarding energy and the Penning slope. In the most simple and common case, each parameter uncertainty is just a Gaussian standard deviation $\boldsymbol{\sigma}_{\text{syst}}$, but in principle any PDF can describe each parameter. We now explain three different approaches to include this type of systematic uncertainty in a Frequentist analysis frame.

Nuisance parameters

The most common and classical approach is the so-called *nuisance parameter* or *pull term* method. Here, in addition to the free parameters that must be estimated from the data, the systematic parameters $\boldsymbol{\theta}_{\text{syst}}$ are included in the likelihood maximization. To consider the knowledge of their PDF, the likelihood is multiplied with each PDF. We have therefore introduced n_{syst} new parameters in our optimization, but also n_{syst} points in our likelihood, one for each PDF. Therefore, the number of degrees of freedom has not changed, but the additional freedom to modify $\boldsymbol{\theta}_{\text{syst}}$ can increase the width of our interval estimation for the parameter(s) of interest.

In addition to increasing the width of the estimated confidence interval, as $\boldsymbol{\theta}_{\text{syst}}$ are included in the likelihood it is possible to learn from the data and improve the estimate on $\boldsymbol{\theta}_{\text{syst}}$. This can lead to a change of the best estimate $\boldsymbol{\mu}_{\text{syst}}$ as well as a narrowing of the width of their PDF. In case $\boldsymbol{\mu}_{\text{syst}}$ changes, the best fit value of our parameter(s) of interest will also change accordingly.

In principle, the nuisance parameter method can be used both with the Monte Carlo approach for interval estimation as well as the profile likelihood. However, in practice, the profile likelihood is more common. Figure 4.5a compares the profile likelihood m_ν^2 for the same dummy dataset as in fig. 4.3 once with statistical uncertainties only, and once including an additional nuisance parameter. One can nicely see the expected widening from including an additional systematic uncertainty.



(a) nuisance parameter **(b)** Monte Carlo method **(c)** covariance matrix

Figure 4.5: Effect of including an additional model uncertainty in the form of one parameter with a Gaussian uncertainty. In this simple case the nuisance parameter (a), Monte Carlo (b), and covariance matrix (c) method give identical results of $0.000 \pm 0.121\,\mathrm{eV}^2$ compared to the result with statistical uncertainties only of $0.000 \pm 0.102\,\mathrm{eV}^2$.

Monte Carlo method

Next, we will discuss how to include systematic uncertainties by extending the Monte Carlo method. As optimizing and utilizing this approach for KATRIN was a major effort over the course of this thesis, all details are given in chapter 5 and this paragraph only serves as a basic overview.

In addition to sampling the data points randomly from their PDF, one can also sample the model parameters $\boldsymbol{\theta}_{\text{syst}}$ according to their corresponding PDFs that describe their systematic uncertainty. This additional variation will lead to an overall widening of the distribution of our fit parameters. However, as all fits are performed on Monte Carlo data, it is not possible to learn from the data. Thus, the best fit as well as the knowledge on $\boldsymbol{\theta}_{\text{syst}}$ cannot change.

Therefore, we must modify the method slightly. When updating our model in each iteration, in addition to fitting the random Monte Carlo spectrum, we also fit the data with our updated model. This data fit gives us a likelihood value which describes the compatibility of our updated model with the data. To include this information, we weight the distribution of our fit parameters by this likelihood value. We can also weight our sampling distribution of each $\boldsymbol{\theta}_{\text{syst}}$ with this likelihood to gain information on the parameters with systematic uncertainties from the data. In section 5.2 we will show that this weighting leads to identical results as the nuisance parameter method.

The effect of introducing an additional systematic uncertainty in the form of a single parameter on our fit parameter distribution using our dummy data set is shown in fig. 4.5b. As with the nuisance parameter approach, the overall distribution is widened. The updated best fit value can be calculated using the weighted median of our m_ν^2 samples and is also identical to the

one of the nuisance parameter method.

Covariance matrix

The final method to include systematic uncertainties we want to discuss is the *covariance matrix* approach [85]. To be able to apply this method, our data points must follow a Gaussian PDF thus allowing us to use the χ^2 minimization. Defining the residual vector $\boldsymbol{r} = \boldsymbol{N} - \boldsymbol{\mu}$ and the diagonal variance matrix $\boldsymbol{V}_{\text{stat}} = \text{diag}\boldsymbol{N}$, we can re-write our χ^2:

$$\chi^2 = \sum_i \frac{(N_i - \mu_i)^2}{N_i} = \boldsymbol{r}^\intercal \boldsymbol{V}_{\text{stat}}^{-1} \boldsymbol{r}. \tag{4.22}$$

Next, we calculate the covariance matrix that describes our systematic uncertainties. For this, similar to the full Monte Carlo approach described in the previous paragraph, we sample random values of $\boldsymbol{\theta}_{\text{syst}}$ according to their PDF. However, instead of performing a full fit with the updated model, we simply evaluate the prediction for each data point given this model $\boldsymbol{\mu}_{\text{sample}}$. After calculating $n_{\text{samples}} \gtrsim 1000$ of these n_{points} model predictions, we end up with a $n_{\text{samples}} \times n_{\text{points}}$ matrix of values. We can now calculate the $n_{\text{points}} \times n_{\text{points}}$ covariance matrix $\boldsymbol{V}_{\text{syst}}$ along the n_{samples} dimension which describes the distribution of the model predictions introduced by the variation of $\boldsymbol{\theta}_{\text{syst}}$ in their multi-normal approximation.

This covariance matrix can now be included in our representation of the χ^2 minimization shown in eq. (4.22):

$$\chi^2 = \boldsymbol{r}^\intercal \left(\boldsymbol{V}_{\text{stat}} + \boldsymbol{V}_{\text{syst}}\right)^{-1} \boldsymbol{r}. \tag{4.23}$$

For this approach to be valid:

- Our data points must follow a Gaussian distribution to good approximation, otherwise the log-likelihood cannot be described by a generalized χ^2.
- The calculated model predictions from the variation of $\boldsymbol{\theta}_{\text{syst}}$ must be well described by the multi-normal distribution so the calculated covariance matrix accurately describes the uncertainty. This is typically the case when the sampling distributions are (multi-) normal and the introduced change of the model prediction is approximately linear at each point. For relatively small systematic uncertainties this is the case as the model variation can be described by the first two terms of its Taylor expansion.

Applied to our dummy dataset, the covariance matrix method gives the exact same results as the nuisance parameter approach as shown in fig. 4.5c.

4.3.3 Frequentist belt construction

So far, all estimations we have made were done on the fit parameter values $\boldsymbol{\theta}_{\text{fit}}$. This means, all confidence intervals we have constructed were constructed under the assumption that our best fit values $\boldsymbol{\theta}_{\text{fit}}$ are also the true values $\boldsymbol{\theta}_{\text{true}}$. However, the question we are usually interested in is *which true parameter values could have lead to our fit result?* not *which other parameter values are compatible with our measurement if our fit result is the truth?*

This difference becomes especially clear in case there are physical boundaries on some parameters. For KATRIN, the parameter of interest m_ν^2 can by definition only be positive. In our Frequentist fit however, we allow for negative m_ν^2 values. Estimating a confidence interval for the true value of m_ν^2 can clearly not be done by assuming a negative fit result to be the truth. Instead, we need to modify our confidence interval construction using the *Frequentist belt construction*.

To construct the belt, we follow this procedure:

1. Assume a true value of the parameter of interest, say $m^2_{\nu,\text{true}} = 0.0\,\text{eV}^2$.

2. Using this value for the parameter of interest, the PDF of all data points, and the best knowledge of all other parameters generate an Asimov Monte Carlo spectrum representing the data.

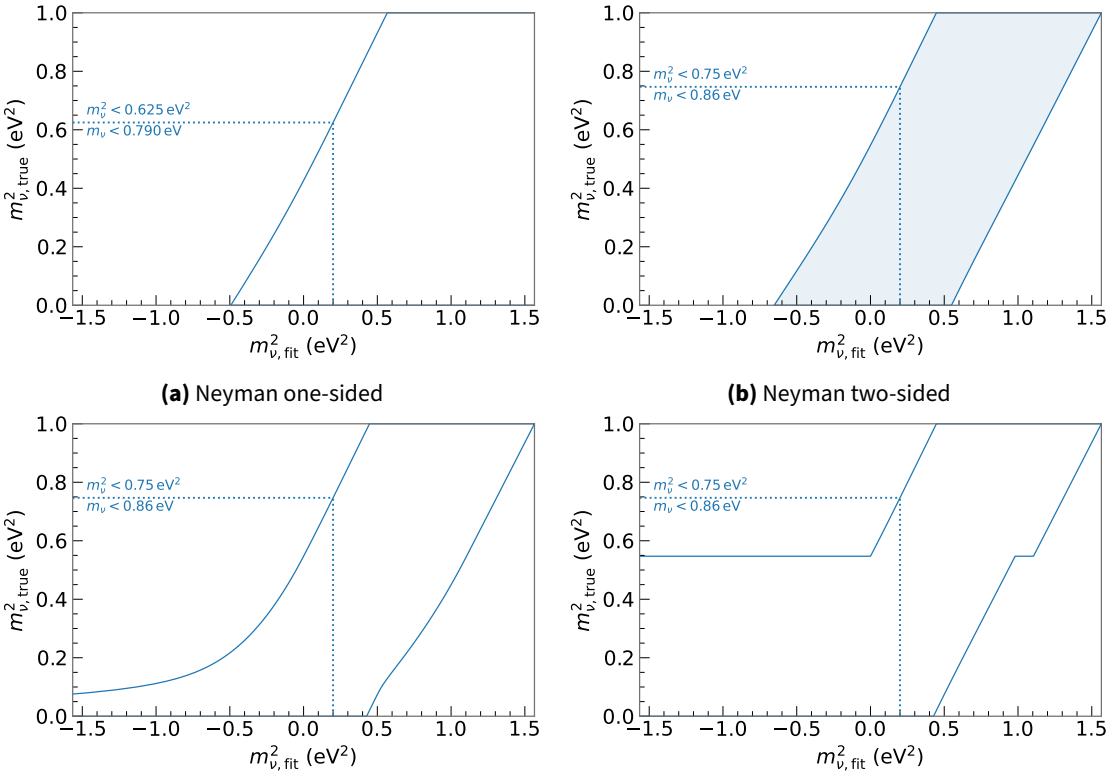
3. Using this underlying spectrum, calculate the profile likelihood or Monte Carlo samples of $m^2_{\nu,\text{fit}}$ using one of the methods in section 4.3.2.

4. Accept a percentage of the $m^2_{\nu,\text{fit}}$ according to your confidence level α and an ordering principle that defines how to sort the $m^2_{\nu,\text{fit}}$ values. This gives a lowest and highest value for $m^2_{\nu,\text{fit}}$ given this $m^2_{\nu,\text{true}}$. The exact approach differs for each of the belt construction methods described in the following paragraphs.

5. Repeat steps 1-4 for different values of $m^2_{\nu,\text{true}}$.

6. Find which values of $m^2_{\nu,\text{true}}$ are compatible with your best fit of the data. This is done by checking for which $m^2_{\nu,\text{true}}$ our given $m^2_{\nu,\text{fit}}$ lies within the accepted interval.

A graphical representation of various belt constructions is shown in fig. 4.6. By construction, all belts have correct *coverage*, meaning the derived interval in $m^2_{\nu,\text{true}}$ from the specific $m^2_{\nu,\text{fit}}$ will contain the true value in α cases for any $m^2_{\nu,\text{true}}$.



(a) Neyman one-sided

(b) Neyman two-sided

(c) Feldman Cousins

(d) Lokhov Tkachov

Figure 4.6: Comparison of various belt construction methods on a dummy dataset assuming $m^2_{\nu,\mathrm{fit}} = 0.2\,\mathrm{eV}^2$ which leads to an upper limit in all cases. The one-sided Nemyan approach leads to empty intervals for $m^2_{\nu,\mathrm{fit}} \lesssim -0.49\,\mathrm{eV}^2$ and by definition cannot lead to a closed interval. The two-sided Neyman method switches to a closed interval for $m^2_{\nu,\mathrm{fit}} \gtrsim 0.55\,\mathrm{eV}^2$ but can also give an empty interval for $m^2_{\nu,\mathrm{fit}} \lesssim -0.65\,\mathrm{eV}^2$. Both the method of Feldman Cousins and Lokhov and Tkachov switch between limits and closed intervals seamlessly and fully avoid empty intervals.

Neyman

In the classical Neyman construction [86] the $m^2_{\nu,\text{fit}}$ values are simply ordered by their value. One can then pick any fraction α of these values, but the only two options generally used are the one-sided Neyman construction where the lower (upper) α are accepted and the two-sided Neyman construction where the central α are accepted. For a typical $90\,\%$ CL, this means we reject the upper (lower) $10\,\%$ for the one-sided approach, and the lower and upper $5\,\%$ for the two-sided approach. Note that these values can once again be calculated using the Monte Carlo method from individual samples or using the cumulative of the χ^2_1 distribution and Wilks' theorem to calculate them from the profile likelihood.

The one-sided and the two-sided Neyman construction are shown in fig. 4.6a and fig. 4.6b respectively. We can see disadvantages for the two approaches when trying to measure the neutrino mass squared. Both approaches can lead to empty confidence intervals for sufficiently negative $m^2_{\nu,\text{fit}}$ values. This happens in $1-\alpha$ of the cases for the one-sided and $\frac{1-\alpha}{2}$ of the cases in the two-sided construction by definition. Quoting an empty interval however is unsatisfactory, especially after performing a multi-year physics experiment. The one-sided Neyman method also has the disadvantage of never leading to a closed interval in $m^2_{\nu,\text{true}}$, no matter how large $m^2_{\nu,\text{fit}}$.

Feldman-Cousins

To solve these issues, Feldman and Cousins introduced their unified approach which seamlessly swaps between limits and closed intervals and avoids empty intervals [87]. Instead of ordering by the plain value of $m^2_{\nu,\text{fit}}$, a new test statistic

$$\Lambda(m^2_{\nu,\text{fit}}) = \frac{\mathcal{L}(m^2_{\nu,\text{true}})}{\mathcal{L}(m^2_{\nu,\text{fc}})} \tag{4.24}$$

is introduced where $m^2_{\nu,\text{fc}} = \max(m^2_{\nu,\text{fit}}, 0)$ is the best-fit value bound to the physical region. In practice, as $\log \mathcal{L}$ is minimized, it is often favourable to use

$$\log \Lambda(m^2_{\nu,\text{fit}}) = \log \mathcal{L}(m^2_{\nu,\text{true}}) - \log \mathcal{L}(m^2_{\nu,\text{fc}}) \tag{4.25}$$

instead.

The confidence interval is now chosen such that

1. α of the $m^2_{\nu,\text{fit}}$ are contained, and
2. $\Lambda(m^2_{\nu,\text{lower}}) = \Lambda(m^2_{\nu,\text{upper}})$.

As this is significantly more involved than the initial Neyman method, let us walk through the process for both the Monte Carlo and the profile likelihood method to calculate these intervals starting with the Monte Carlo approach.

For a given $m^2_{\nu,\text{true}}$, many random Monte Carlo spectra are generated. For each spectrum, we calculate $\log \Lambda$ according to eq. (4.25) by once fitting with $m^2_\nu = m^2_{\nu,\text{true}}$ and once with a free neutrino mass squared to obtain $m^2_{\nu,\text{fit}}$. In case the best-fit is negative, we fix m^2_ν to zero and run a third optimization. These two or three minimizations allow us to calculate the test statistic for the given sample $\log \Lambda_\text{sample}$. We now order our samples by $\log \Lambda_\text{sample}$ and accept the fraction α with the lowest $\log \Lambda$. This is shown assuming a unit Gaussian in fig. 4.7a. The effect of the introduced test statistic is clearly visible: For positive values of $m^2_{\nu,\text{fit}}$, the usual parabola form of $\log \mathcal{L}$ is obtained. This flattens to a linear function for negative $m^2_{\nu,\text{fit}}$ and goes towards flat zero for $m^2_{\nu,\text{true}} \to 0$. This is explained by the fact that $\log \mathcal{L}(m^2_{\nu,\text{true}}) =: \log \mathcal{L}(0)$ as well as $\log \mathcal{L}(m^2_{\nu,\text{fc}}) =: \log \mathcal{L}(0)$ for negative best fit values.

Next, let us investigate how to construct the Feldman Cousins interval from the profile likelihood [88]. For each $m^2_{\nu,\text{true}}$ we first calculate the profile likelihood on the Asimov spectrum as usual. In addition, we always need the profile assuming $m^2_{\nu,\text{true}} = 0$ which we subtract from the current likelihood profile for negative $m^2_{\nu,\text{fit}}$. The resulting $\log \Lambda$ profiles are shown in fig. 4.7b once again assuming a unit Gaussian. The lower and upper bounds in $m^2_{\nu,\text{fit}}$ are now calculated by performing a root search which ensures $\log \Lambda_\text{lower} = \log \Lambda_\text{upper}$ and finds the approriate $m^2_{\nu,\text{lower}}, m^2_{\nu,\text{upper}}$ such that

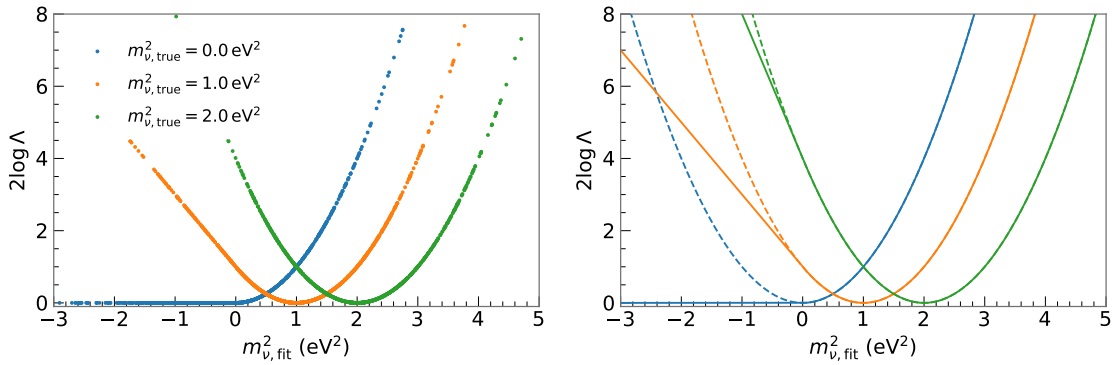
$$0.5 \cdot \left(\text{CDF}_{\chi^2_1}(2 \log \mathcal{L}(m^2_{\nu,\text{lower}})) + \text{CDF}_{\chi^2_1}(2 \log \mathcal{L}(m^2_{\nu,\text{upper}})) \right) = \alpha \tag{4.26}$$

once again making use of Wilks' theorem for coverage. Note that in addition to Wilks' theorem, we also assume that the likelihood value for $m^2_{\nu,1}$ given Asimov data generated with $m^2_{\nu,2}$ is the same as the likelihood value for $m^2_{\nu,2}$ given Asimov data generated with $m^2_{\nu,1}$ by making use of the profile likelihood for a single truth, instead of generating multiple samples and inserting the truth.

The resulting belt is shown in fig. 4.6c where we can see that empty intervals are avoided and upper limits naturally turn into closed intervals for large $m^2_{\nu,\text{fit}}$.

Lokhov-Tkachov

While the method of Feldman and Cousins solves all needs of a physicist from a statistical point of view, the feature of improved upper limits for more negative $m^2_{\nu,\text{fit}}$ values can be seen critically for direct neutrino mass measurements where neglected systematics often lead to more negative $m^2_{\nu,\text{fit}}$. Therefore, this is avoided in the method of Lokhov and Tkachov specifically designed for this type of experiment [89].



(a) Monte Carlo method

(b) Profile Likelihood

Figure 4.7: Test statistic using the Method of Feldman and Cousins as defined in eq. (4.25) for the unit Gaussian once using Monte Carlo samples (a) and once from the profile likelihood (b). Both cases show the test statistic for three different true neutrino mass squared values. For the profile likelihood, the dashed line indicates the unmodified profile likelihood and the solid line the test statistic retrieved by subtracting the profile likelihood for $m^2_{\nu,\text{true}} = 0$.

For all $m^2_{\nu,\text{true}}$ below the *sensitivity* of the experiment, a one-sided Neyman construction cutting $1 - \alpha$ from the right is performed. Above the sensitivity, a regular two-sided Neyman construction is applied instead. The sensitivity is the $m^2_{\nu,\text{true}}$ value for which the lower bound of the two-sided Neyman construction coincides with zero.

In fig. 4.6d we can see the effects of this construction. Negative $m^2_{\nu,\text{fit}}$ do not improve the upper limit compared to the sensitivity. The hard switch to a two sided construction above the sensitivity leads to a kink in the belt at this value. Like the method of Feldman and Cousins, empty intervals are avoided and closed intervals come naturally for large $m^2_{\nu,\text{fit}}$.

4.3.4 Bayesian inference

In contrast to our Frequentist methods described in the previous sections, Bayesian inference assigns direct probability values to the parameters measured from the specific dataset. To do so, it makes use of Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{4.27}$$

to rewrite our likelihood

$$\mathcal{L}(\boldsymbol{\theta}) = P(\text{data}|\boldsymbol{\theta}) = \frac{P(\boldsymbol{\theta}|\text{data})P(\text{data})}{P(\boldsymbol{\theta})} \tag{4.28}$$

$$\Rightarrow P(\boldsymbol{\theta}|\text{data}) = \frac{P(\text{data}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\text{data})} \tag{4.29}$$

$$= \frac{\mathcal{L}(\boldsymbol{\theta})P(\boldsymbol{\theta})}{\int \mathcal{L}(\boldsymbol{\theta})P(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}} \tag{4.30}$$

into an expression that defines the probability of our parameter values given the measured data. As we have discussed the likelihood in detail in the previous parts, we only briefly discuss the other components of eq. (4.30) now [90].

- $P(\boldsymbol{\theta})$ is the prior distribution that allows incorporating any existing knowledge on $\boldsymbol{\theta}$. This could for example be any PDF describing systematic uncertainties, but can also be left *uninformative*, most commonly by using a constant distribution, the so-called *flat prior*.

- $P(\boldsymbol{\theta}|\text{data})$ is the posterior distribution, retrieved by updating the existing prior knowledge with the information contained in the likelihood, typically via a Markov Chain Monte Carlo algorithm. This distribution is the complete result containing all information of the Bayesian analysis as it is the direct joint PDF of our measured parameters $\boldsymbol{\theta}$. To, for example, retrieve the result for a single parameter, one would marginalize the posterior distribution by integrating it over all other parameters. This one dimensional distribution is then used to construct any *credible intervals*, some common ones being the central or one-sided credible interval at a given credibility level (such as $90\,\%$ or $1\,\sigma$).

- Finally, the factor $\int \mathcal{L}(\boldsymbol{\theta})P(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}$ ensures normalization of our posterior probability distribution.

In a KATRIN neutrino mass analysis, one would typically use flat priors in E_0, A_S and r_bg and include any systematic uncertainties on other parameters with the corresponding prior. The prior to use for the neutrino mass is often a topic for discussion, but the most common options are flat in m_ν^2 to directly compare to a Frequentist analysis, or flat positive in m_ν^2 to ensure only values in the physical region.

For completeness, we note that Bayesian and Frequentist intervals contain different information and cannot be compared directly in general, although they may match for simple cases. This thesis focuses on a Frequentist analysis, only making brief use of Bayesian inference for a proof-of-concept in section 7.5.1.

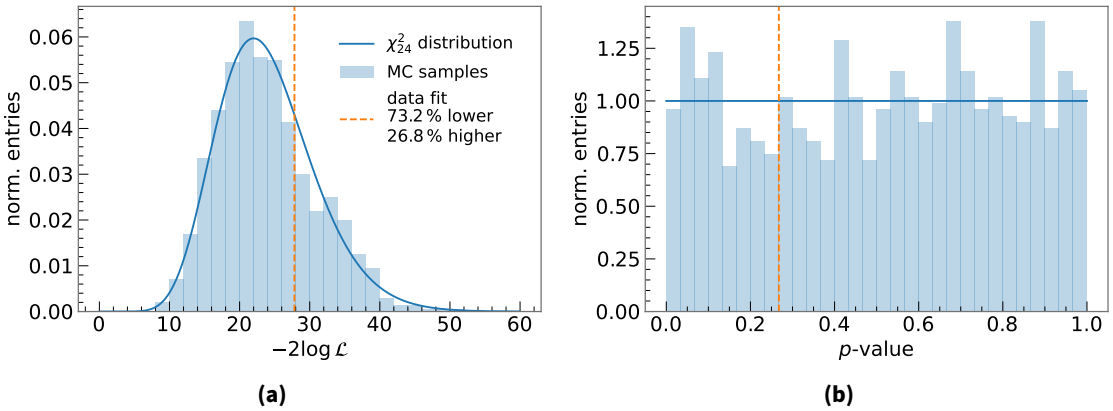


Figure 4.8: Distribution of $-2\log\mathcal{L}$ (a) and p-values (b) assuming an underlying χ^2_{24} distribution and a random best fit value (orange dashed line). The fraction of $-2\log\mathcal{L}$ values above the data fit value define the p-value.

4.4 Goodness of fit

Another question typically asked is *how well does the model describe the data*? This can be answered by a *goodness of fit* test statistic. So far, we have only used the likelihood function to find the set of parameters that best describe the data, assuming an underlying model. However, this best estimate may still not describe the data at all if our model assumption is wrong.

To get an estimate if this is the case, we once again make use of a Monte Carlo method. Assuming our model, we generate statistically randomized spectra and fit each of them. For each sample, we store $-2\log\mathcal{L}_{\text{sample}}$. This gives us the expected distribution of $-2\log\mathcal{L}_{\text{sample}}$ if our model describes the data, where higher values describe the data worse by definition. Comparing the value we get from our real data fit with the expected distribution, we can calculate how probable it is to get a $-2\log\mathcal{L}$ value as large as the one of our data fit if our model would describe the data. This quantity is referred to as the p-value of the fit and it follows a uniform distribution between zero and one. The absolute $-2\log\mathcal{L}$ values and their relation to the p-value is visualized in fig. 4.8.

For a χ^2 minimization, $-2\log\mathcal{L} =: \chi^2$ follows a χ^2_{ndof} distribution with ndof degrees of freedom defined as the number of data points minus the number of free parameters. Therefore, instead of estimating our $-2\log\mathcal{L}$ distribution, we can directly make use of χ^2_{ndof} and calculate the p-value from the cumulative. This calculation can often be used as approximation, even when the χ^2 minimization is not performed directly.

4.5 Treatment of negative neutrino mass squared values

Any negative value of m_ν^2 is unphysical. However, in a Frequentist analysis, one typically allows for negative values of $m^2_{\nu,\text{fit}}$ and only avoids the unphysical region when constructing the confidence belt by assuming $m^2_{\nu,\text{true}} \geq 0\,\text{eV}^2$. Negative $m^2_{\nu,\text{fit}}$ then lead to upper limits on $m^2_{\nu,\text{true}}$ or empty intervals, depending on the specifics of each method.

In any case, to allow $m^2_{\nu,\text{fit}} < 0\,\text{eV}^2$ in the fit, the part of the phase space factor of eq. (3.15) including m_ν

$$\sqrt{\epsilon^2 - m_\nu^2} \cdot \Theta(\epsilon - m_\nu) \tag{4.31}$$

must be adapted. The simplest adaptation is to only set $m_\nu = \max(0, m_\nu)$ in the Heaviside function and otherwise keep the formula as-is. One can observe that this leads to a slightly asymmetric likelihood, with $\sigma_\text{lower} > \sigma_\text{upper}$ for $m_\nu^2 \approx 0\,\text{eV}^2$. This means one will always have to at least calculate the profile likelihood to retrieve the asymmetric errors, and possibly perform a coverage study to ensure applicability of Wilks' theorem, but does not introduce any fundamental statistical issues. This is especially true, as a Frequentist belt construction is performed in the last step to convert $m^2_{\nu,\text{fit}}$ to an interval in $m^2_{\nu,\text{true}}$.

However, to simplify the analysis, one can opt to modify the phase space for negative m_ν^2 such that the likelihood is symmetric around zero by construction. One such extension has been used to analyse the Mainz experiment [32]. It uses the phase space

$$\left(1 + \frac{\mu}{\epsilon} \exp\left(-1 - \frac{\epsilon}{\mu}\right)\right) \sqrt{\epsilon^2 - m_\nu^2} \cdot \Theta(\epsilon + \mu) \tag{4.32}$$

with $\mu = -k m_\nu^2$ where the factor k depends on the specifics such as the statistical sensitivity of the experiment and must be derived from a Monte Carlo simulation [62].

As the negative m_ν^2 regime is unphysical in any case, the choice is up to the individual analysis how to extend the phase space into it, as long as it is use consistently and correct error estimation and coverage is ensured. All results in this thesis use the nearly unmodified phase space with asymmetric errors in m_ν^2.

In a Bayesian analysis, it is straightforward to exclude the unphysical regime by setting a prior on m_ν^2 which is zero for $m_\nu^2 < 0\,\text{eV}^2$.

Chapter 5

Monte Carlo propagation of uncertainty

The Monte Carlo method to propagate uncertainty briefly described in section 4.3 was developed in collaboration with Martin Slezák and optimized for the usage in KATRIN in the course of this thesis. This chapter will give a detailed description of the methods, a validation by means of comparing to the nuisance parameter approach and a description of how to practically apply the technique in KATRIN [91, 92].

5.1 Method description

Basic idea of the Monte Carlo propagation of uncertainty is to repeat the full fit procedure, with different parts randomized in each step. In this sense, it is based on the Frequentist concept of probability, as it creates thousands of randomized experiments that never took place to derive confidence intervals for the parameters of interest.

5.1.1 Statistical uncertainty

Propagating statistical uncertainty using Monte Carlo samples is a well known ensemble method. The general procedure is to generate an Asimov spectrum which represents the data to our best knowledge. Therefore, it must include the exact values for the retarding energies and measuring time, but also the best estimate for the model parameters such as the magnetic fields, the gas density in the source, and the neutrino mass squared.

While the measuring time distribution and the *slow control* parameters are known or measured using external calibrations, the fit parameters $\{m_\nu^2, E_0, A_\mathrm{S}, r_\mathrm{bg}\}$ can only be estimated from

the data. Therefore, the first step is to fit the data with these parameters free (and whatever segmentation over time and detector is used) and get a best estimate of their central values. With all model parameters estimated, the Asimov spectrum can be calculated.

This Asimov spectrum is now fluctuated statistically according to the PDF of the data points. Each randomized spectrum is then fit with the same fit parameters, storing the values of each sample $\boldsymbol{\theta}_{\text{sample}}$. The resulting distribution of each parameter is a direct result of the statistical fluctuation of the data points and thus represents the statistical uncertainty of each parameter. Once can then derive quantities such as the central $1\,\sigma$ interval and the corresponding σ_{lower} / σ_{upper} uncertainties from these distributions.

The best fit does not change from the randomization as no additional information is drawn from the data compared to the initial fit. The quantity of the resulting distributions that corresponds to the best fit is the median, which is not affected from statistical fluctuations. This corresponds to the intuition of half the fluctuations moving the fit parameter upwards and half of them downwards. Using the mean is not a good estimate, as the lower and upper distribution width does not necessarily coincide, i.e. $\sigma_{\text{lower}} \neq \sigma_{\text{upper}}$ in general. In fact, the KATRIN likelihood using the differential spectrum as-is for negative m_ν^2 values shows $\sigma_{\text{lower}} > \sigma_{\text{upper}}$ for $m^2_{\nu,\text{best}} = 0\,\text{eV}^2$.

Let us summarize the steps again:

1. Fit the data to get an initial estimate on the free parameters.
2. Generate an Asimov spectrum using the best knowledge of the data and the best fit of the free parameters from 1.
3. Statistically randomize the Asimov spectrum according to the PDF of the data points.
4. Fit the randomized spectrum to get the fit parameter values $\boldsymbol{\theta}_{\text{sample}}$.
5. Repeat steps 2 to 4 to get the distribution of $\boldsymbol{\theta}_{\text{sample}}$.
6. Derive confidence intervals from the distribution, the best fit is unaffacted and described by the median of the distribution.

5.1.2 Systematic uncertainty

Next step is to discuss the Monte Carlo propagation of systematic uncertainties. In this case, we have an uncertainty on the model parameters themselves, which can be described by some PDF. Our goal is to propagate the uncertainty of the model parameters into an uncertainty on our fit parameters via propagation of distributions.

To this end, we first need an initial estimate of our fit parameters and a corresponding Asimov spectrum as in the previous case with statistical uncertainty. Now, instead of randomizing the data points in the Asimov spectrum, we vary the model according to the systematic uncertainty and fit the unvaried Asimov spectrum. The resulting distribution in our free parameters from fitting with different models comes from the initial systematic uncertainty on the model itself. At this point it makes sense to mention that this is equivalent to generating a new Asimov spectrum with varied model parameters, and fitting it with the inital unvaried model, as long as the sampling PDF is symmetric.

With this method, the initial systematic uncertainty is propagated to our fit parameters. However, the data itself also has information on the parameters with systematic uncertainty. To include this, we fit every model with a given set of values for the systematic parameters to the data and store the likelihood value $\mathcal{L}_{\text{sample}}$. This likelihood is a measure how well this specific model describes the data. Therefore, we will weight our fit parameter distribution with the corresponding likelihood value to include information from the data. In case the calibration is much more precise than any information in the data, the likelihood values will be nearly identical and have no impact on the resulting distribution. This is very often the case for KATRIN.

Putting this together gives us the following prescription to propagate systematic uncertainties:

1. Fit the data to get an initial estimate on the free parameters.

2. Generate an Asimov spectrum using the best knowledge of the data and the best fit of the free parameters from 1.

3. Randomize the model according to the systematic uncertainties.

4. Fit the Asimov spectrum to get the fit parameter values $\boldsymbol{\theta}_{\text{sample}}$.

5. Fit the data spectrum to get the likelihood value $\mathcal{L}_{\text{sample}}$.

6. Repeat steps 2 to 5 to get the distribution of $\boldsymbol{\theta}_{\text{sample}}$ and corresponding distribution of $\mathcal{L}_{\text{sample}}$.

7. Derive confidence intervals from the parameter distribution weighted by the likelihood value, the best fit is the weighted median of the distribution.

5.1.3 Total uncertainty

To get the total uncertainty which includes statistical and systematic uncertainties, we only have to combine the two approaches. In each iteration we both statistically randomize the Asimov spectrum and change the model according to the systematic uncertainties. This gives us the total recipe to run the Monte Carlo propagation of uncertainty:

1. Fit the data to get an initial estimate on the free parameters.
2. Generate an Asimov spectrum using the best knowledge of the data and the best fit of the free parameters from 1.
3. Statistically randomize the Asimov spectrum according to the PDF of the data points.
4. Randomize the model according to the systematic uncertainties.
5. Fit the randomized spectrum to get the fit parameter values $\boldsymbol{\theta}_{\text{sample}}$.
6. Fit the data spectrum to get the likelihood value $\mathcal{L}_{\text{sample}}$.
7. Repeat steps 2 to 6 to get the distribution of $\boldsymbol{\theta}_{\text{sample}}$ and corresponding distribution of $\mathcal{L}_{\text{sample}}$.
8. Derive confidence intervals from the parameter distribution weighted by the likelihood value, the best fit is the weighted median of the distribution.

An example of the three steps applied is shown in fig. 5.1. In this case, the total uncertainty is clearly dominated by the statistical uncertainty with only a small contribution of systematic uncertainties.

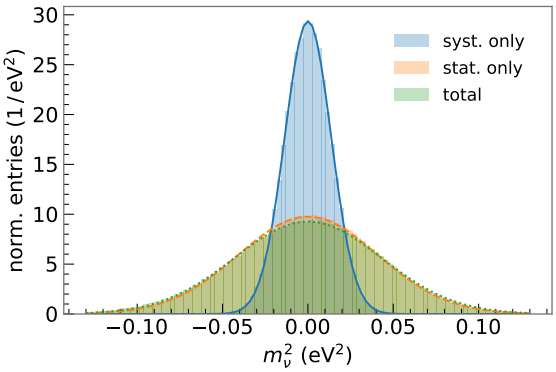


Figure 5.1: Example of Monte Carlo propagation of uncertainty on a dummy dataset including only systematic uncertainties (blue), only statistical uncertainties (orange) and both (green). As expected, the distribution with the total uncertainty is wider than the two individual ones. In this case, the total is dominated by statistical uncertainty.

5.2 Comparison with nuisance parameter method

In this section, the Monte Carlo propagation of uncertainty is validated by comparing various cases to the nuisance parameter method.

First, the case of statistical uncertainty only is investigated. In this case, the Monte Carlo method is by definition equivalent to the confidence intervals retrieved via profile likelihood as long as Wilks' theorem holds or the appropriate $\Delta 2 \log \mathcal{L}$ is derived from a coverage study. In fig. 5.2 the statistical uncertainty using both methods is compared for different measuring times. As expected, the two match, and Wilks' theorem is valid for this simplified KATRIN case.

Next, we include a single systematic uncertainty in the form of one parameter with a Gaussian uncertainty which is more stringent than the data. This is shown in fig. 5.3a for different uncertainties on the background retarding energy slope b_{qU}. In fig. 5.3b this is extended by adding an additional uncertainty on the gas density in the source ρd. For all cases, the total uncertainty of the nuisance parameter method as well as the Monte Carlo propagation match.

To check the effect of weighting with the likelihood value, we now increase the Gaussian uncertainty on b_{qU} until it becomes larger than the uncertainty when estimating the background energy slope from the data. In this case, we expect the uncertainty on m_ν^2 to converge towards the uncertainty as if b_{qU} was included as free parameter in the fit. This is exactly what is shown in fig. 5.4a: both the nuisance parameter as well as the weighted MC propagation learn from the data and the uncertainty converges. However, if we neglect the weighting in the MC propagation, the uncertainty on m_ν^2 diverges with the included systematic uncertainty.

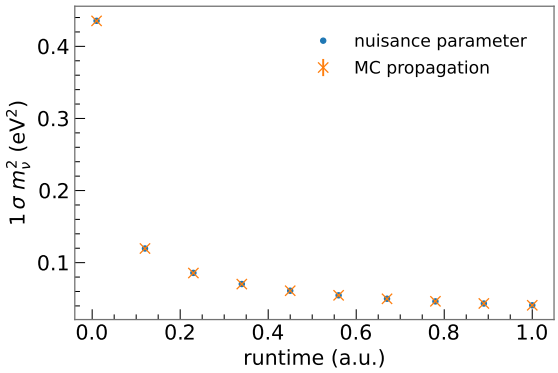
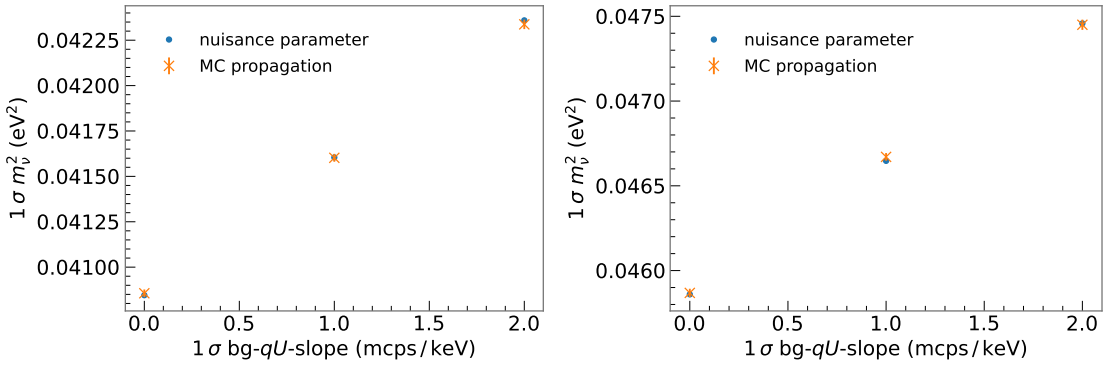


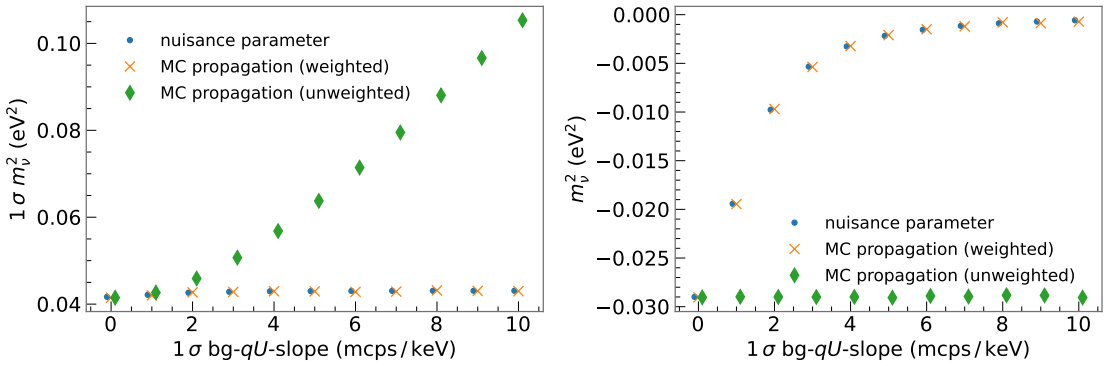
Figure 5.2: Statistical uncertainty from the profile likelihood assuming Wilks' theorem and the Monte Carlo propagation method for different measuring times. The two match perfectly for this simplified KATRIN dummy dataset.



(a) single systematic uncertainty

(b) two systematic uncertainties

Figure 5.3: Comparison of the total uncertainty on m_ν^2 using the nuisance parameter and the MC propagation approach. In both cases, the systematic uncertainty on the background energy slope is increased from $0.0\,\mathrm{mcps/keV}$ to $2.0\,\mathrm{mcps/keV}$. While this is the only systemtatic parameter with uncertainty in (a), (b) also includes a $0.5\,\%$ uncertainty on the gas density in the source. In all cases, the nuisance parameter and MC propagation give consistent results.



(a) uncertainty **(b)** best-fit

Figure 5.4: Impact of weighting the MC propagation samples with the likelihood value on the uncertainty (b) and uncertainty (a). The model value of the qU-slope is centered at $3.0\,\mathrm{mcps/keV}$ with the MC truth at $0.0\,\mathrm{mcps/keV}$ resulting in a biased m_ν^2 in a fit with statistical uncertainty. The corresponding points are slightly shifted with respect to one-another for better readability. Without weighting (green), the best-fit remains biased even if the uncertainty covers the true value. The uncertainty also keeps increasing, as the samples do not "learn from the data". With weighting (orange) both the best-fit and the uncertainty are in good agreement with the nuisance parameter method (blue).

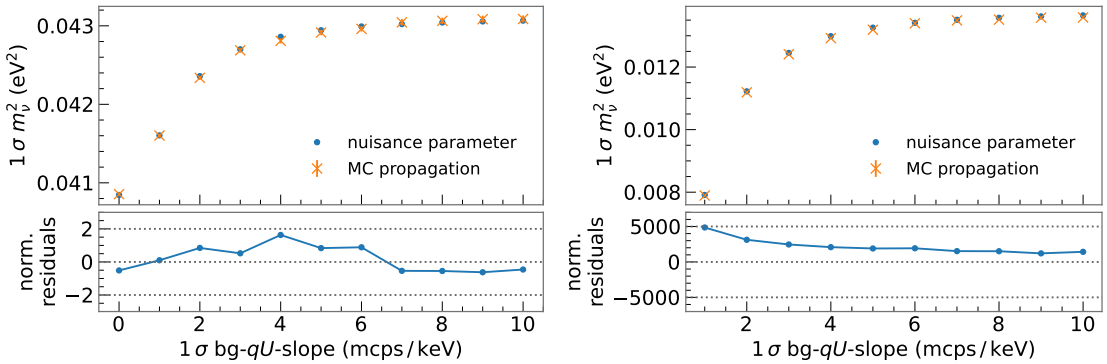
In fig. 5.4b we do something similar, but observe the impact on the best fit value of m_ν^2. Here, we bias the central value of the systematic parameter in our model, in this case again the background energy slope b_{qU}, to $3.0\,\mathrm{mcps/keV}$ while the Monte Carlo truth is at $0.0\,\mathrm{mcps/keV}$. This biases the central value of m_ν^2 for constraints on b_{qU} as stringent as the bias. However, when the constraint become more loose, we expect to be able to recover by learning from the data and fitting both $b_{qU} \approx 0.0\,\mathrm{mcps/keV}$ and thus $m_\nu^2 \approx 0.0\,\mathrm{eV}^2$. This is exactly what we observe for the nuisance parameter method and the weighted MC propagation. Once again, neglecting the weighting and thus the possibility to learn from the data leads to a wrong result, here a continous bias on m_ν^2.

In our final step, we come back to the uncertainty on m_ν^2 displayed in fig. 5.4a, and take a closer look by displaying the residual

$$\frac{\sigma_{\mathrm{mc}} - \sigma_{\mathrm{nuisance}}}{\Delta_{\sigma_{\mathrm{mc}}}} \tag{5.1}$$

where the uncertainty on the estimate from the Monte Carlo propagation $\Delta_{\sigma_{\mathrm{mc}}}$ comes from the finite sample size and is calculated using a bootstrapping method [93]. The total uncertainty, meaning the intrinsic statistical uncertainty of the data set as well as the included systematic uncertainty from b_{qU}, is shown in fig. 5.5a. The two approaches agree well within the uncertainty of the MC method. However, when comparing only the systematic uncertainty in fig. 5.5b we can see a small, but statistically very significant, deviation. As the statistical and the total uncertainty agrees, we blame this on inaccuracies when estimating the systematic uncertainty of the nuisance parameter approach using

$$\sigma_{\mathrm{syst}} = \sqrt{\sigma_{\mathrm{total}}^2 - \sigma_{\mathrm{stat}}^2} \tag{5.2}$$



(a) total uncertainty

(b) systematic uncertainty

Figure 5.5: $1\,\sigma$ uncertainty on m_ν^2 when propagating an uncertainty on the background energy slope using the nuisance parameter (blue) and Monte Carlo propagation (orange) approach.

which only holds for perfectly Gaussian uncertainties. This is not given in the KATRIN case, as the lower and upper uncertainty on m_ν^2 differ in general.

To conclude, we have shown that the Monte Carlo propagation of uncertainty leads to consistent results with the nuisance parameter approach both for statistical uncertainty only and when including additional systematic uncertainties. This holds for various cases, and weighting the MC samples with the corresponding likelihood also allows learning from the data exactly as the nuisanance parameter method does.

5.3 Advantages and downsides compared to other uncertainty treatments

To conclude this chapter on the Monte Carlo propagation of uncertainty, we want to discuss its advantages and disadvantages compared to the other methods to include systematic uncertainties described in section 4.3.2, namely the nuisance parameter and covariance matrix approach.

An advantage of the Monte Carlo method is that it makes very little assumptions and has very little requirements on the underlying probability distributions. It does not assume Wilks' theorem as all confidence intervals are constructed directly from the parameter distributions, not using $\Delta 2 \log \mathcal{L} = 1$. In addition, it does not require the PDF of the data points to follow a Gaussian, which is required for the covariance matrix. Furthermore, the full PDF of the systematic uncertainties is propagated and not reduced to its covariance on the data point. Like for the covariance matrix, any model variation can be propagated without the need to parametrize it in form of a fit parameter. One example where this is needed is the bin-to-bin

variation of the molecular final state distribution described in section 3.4.

In the MC propagation approach it is easy to infer the impact of individual systematics, as it is possible to perform the propagation of systematic uncertainties without also propagating the statistical uncertatinty. Thus, one does not have to rely on approximations like $\sigma^2_{\text{syst}} = \sigma^2_{\text{total}} - \sigma^2_{\text{stat}}$ to calculate a breakdown of individual systematic uncertainties.

Another KATRIN specific advantage is that the expensive response function, see section 3.3.2, can be pre-calculated for every fit. This is possible as any response changing systematic parameter is only varied once before the fit, but then kept fixed at this specific value during optimization. This goes hand in hand with the fact that the number of free parameters is kept minimal, only those that must be estimated from the data are included in the optimization, not any of the constrained nuisance parameters. These advantages are shared with the covariance matrix, where the corresponding covariance matrix is calculated once beforehand, and the systematic parameters are then no longer varied within the final fit.

However, the whole process is still numerically expensive as several thousands, or even hundreds of thousands of fits are performed to achieve good accuracy on the estimated intervals. This difference is especially large when comparing with the covariance matrix method, as here only several thousands of spectra are calculated, and then a single fit can be performed. The nuisance parameter approach also only requires a single fit, however the response function must be re-calculated during the fit for changing systematic parameters.

In addition to the remaining numerical costs, it is difficult to get a final likelihood value from the MC propagation. While this is typically not an issue from a goodness of fit point of view, as the KATRIN analysis is typically dominated by statistical uncertainty, this makes it difficult to apply the method of Feldman and Cousins without approximation.

These possible downsides were not an issue for the first two measurement campaigns and we applied the Monte Carlo propagation of uncertainty to analyse these data sets [43, 66, 78].

Chapter 6

First sub-electronvolt direct neutrino mass measurement

A major milestone of this thesis was the analysis of the first two KATRIN measurement campaigns, leading to the first direct neutrino mass measurement limiting m_ν to less than one electronvolt ($90\,\%$ CL) [43]. This chapter describes the datasets, our analysis method to infer m_ν^2 and a corresponding confidence interval in m_ν, and concludes with a critical discussion of our results.

6.1 Data from the first two measurement campaigns

As described in section 3.5.1, the first two measurement campaigns took place in spring and autumn 2019 respectively. From an analysis point of view, the main difference between the two campaigns is the increased gas density (from $22.2\,\%$ to $84.5\,\%$) and thus signal rate as well as the reduced background (from $292\,\mathrm{mcps}$ to $220\,\mathrm{mcps}$) of the 2$^{\mathrm{nd}}$ measurement phase. These two improvements are nicely visible in the measured spectra shown in fig. 6.1.

Figure 6.2 shows the measuring time distribution for both campaigns. The shift to the right of the peak, which is always located in the region most sensitive to the neutrino mass, also reflects the higher signal to background ratio. Another change from KNM1 to KNM2 is the inclusion of a background point above $18\,700\,\mathrm{eV}$. This single point can constrain any background retarding energy slope well, but poses a risk as it is the only point to do so. Therefore, any fluctuation of this point will strongly influence the value of the background retarding energy slope. For each campaign, an interval of roughly $40\,\mathrm{eV}$ below the endpoint is included in the analysis, resulting in 27 and 28 considered points respectively.

For both campaigns a slightly different set of 117 pixels was selected out of the total of 148 possible focal plane detector pixels. The corresponding pixel selection is shown in fig. 6.3.

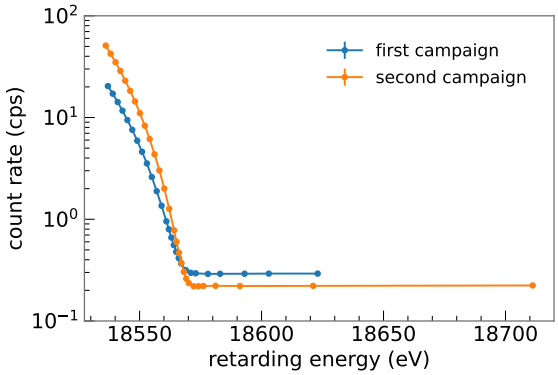


Figure 6.1: Data spectrum of the first (blue) and second (orange) measurement campaign. In both cases, all considered scans and pixels are summed up. The error bars are too small to be visible at this scale. One can clearly see the improved data quality in the second campaign with a higher signal rate as well as a lower background.

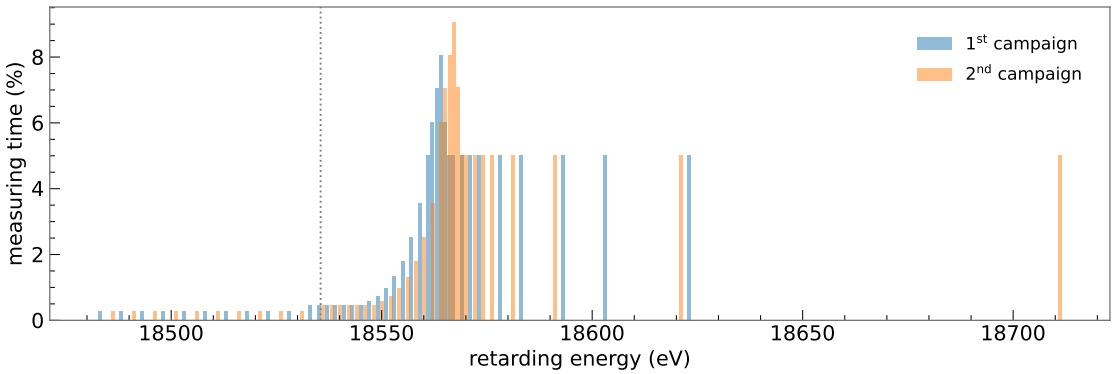
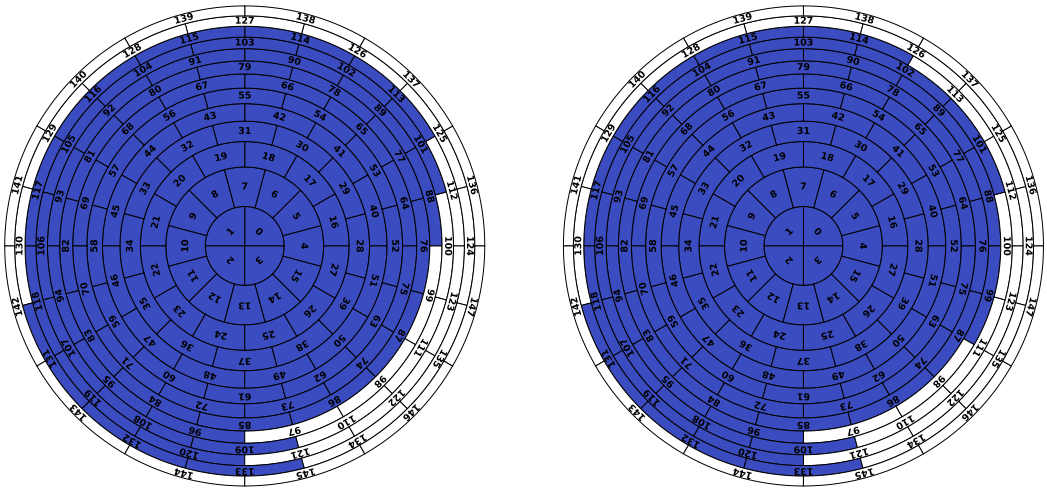


Figure 6.2: Measuring time distribution for the first two neutrino mass campaigns. The gray dotted line indicates the nominal analysis interval where all points to the right are included. Any monitoring points more than $100\,\mathrm{eV}$ below the endpoint are excluded for better visualization.



(a) first campaign **(b)** second campaign

Figure 6.3: Pixel selection for the first (a) and second (b) measurement campaign.

Reasons to reject a pixel include it having a deteriorated energy resolution or being noisy, being shadowed by the forward beam monitor, and not seeing the full source due to misalignment [94, 95].

During the second measurement campaign, the effects of the source electric potential are non-negligible for the first time. This means, the rear wall bias voltage plays an important role as described in section 3.2. Over the course of this measurement phase, there were hints for a radial inhomogeneity of the potential, seen in a slight radial dependency of the endpoint. This was counteracted twice by setting the rear wall voltage to a different value as shown in fig. 6.4 to compensate for the work function difference of the rear wall and beam tube surfaces. With both surfaces at the same potential, the electric potential in the source is the most homogeneous radially and longitudinally. For the first campaign, the bias voltage was constant over most of the time, with only a short period in the beginning set to a slightly lower value. The effect of having three distinct rear wall voltages in KNM2 on m_ν^2 will be discussed in section 6.7.3.

To conclude this section, we take a look at the stability of the source slow control parameters: namely the column density and the tritium purity shown in fig. 6.5. The gas density in the first campaign clearly shows the downward drift caused by radiochemical reacions of tritium with the injection capillary. The spikes caused by counteracting the drift via increasing the injection pressure keep the overall spread manageable with a standard deviation of $0.80\,\%$. In the second campaign, the gas density was much more stable with a standard deviation of $0.13\,\%$, despite being almost four times higher, showing that the continous flushing was successful. During KNM1 there is a slight upwards trend of the overall tritium purity with

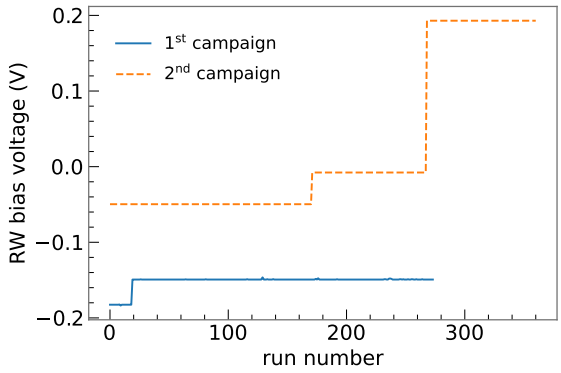
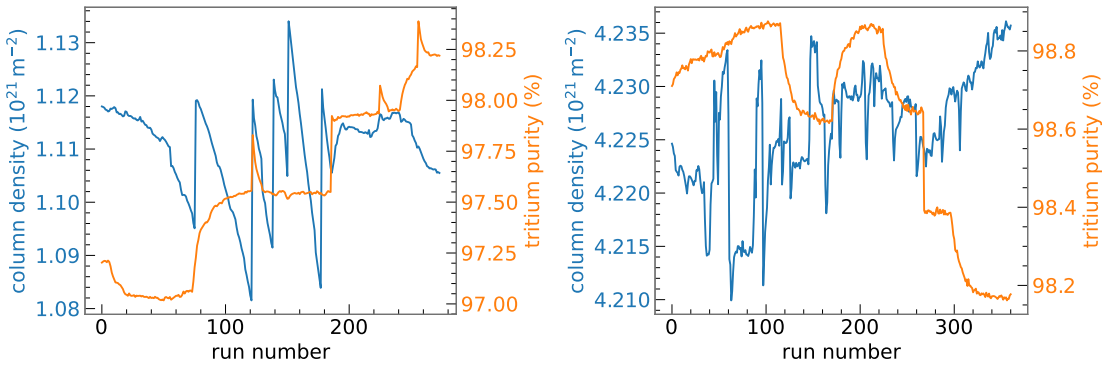


Figure 6.4: Rear wall bias voltage for the first two measurement campaigns.



(a) first campaign

(b) second campaign

Figure 6.5: Evolution of the column density and tritium purity over the first (a) and second (b) neutrino mass campaign. The downward drift of the gas density counteracted by frequent increases of the injection pressure during KNM1 is cleary visible, leading to an overall larger spread.

an overall standard deviation of $0.39\,\%$. In KNM2, the tritium purity slowly decreases over the course of the campaign, but is also more stable with a standard deviation of $0.24\,\%$. We will see in the uncertainty breakdown that the activity changes are not of concern for either campaign.

6.2 Fit parameter stability

Next, we discuss the stability of the fit parameters other than m_ν^2, namely the endpoint, the signal normalization and the background rate. The stability of the endpoint is crucial for the neutrino mass analysis, as any variation of the energy scale leads to a broadening which translates to a more negative m_ν^2 value if unaccounted. Observing the stability of A_S is also useful, as any change hints to incomplete knowledge of the activity parameters such as the

gas density and the tritium purity. Keeping a constant background rate is not critical for the m_ν^2 fit as it completely averages out as a linear parameter, but will be discussed here for completeness.

First step is to check the time evolution of these parameters. This is done by fitting each scan independently with statistical uncertainty only and the neutrino mass fixed to zero in the regular analysis interval. The resulting parameter values over scan are shown in fig. 6.6.

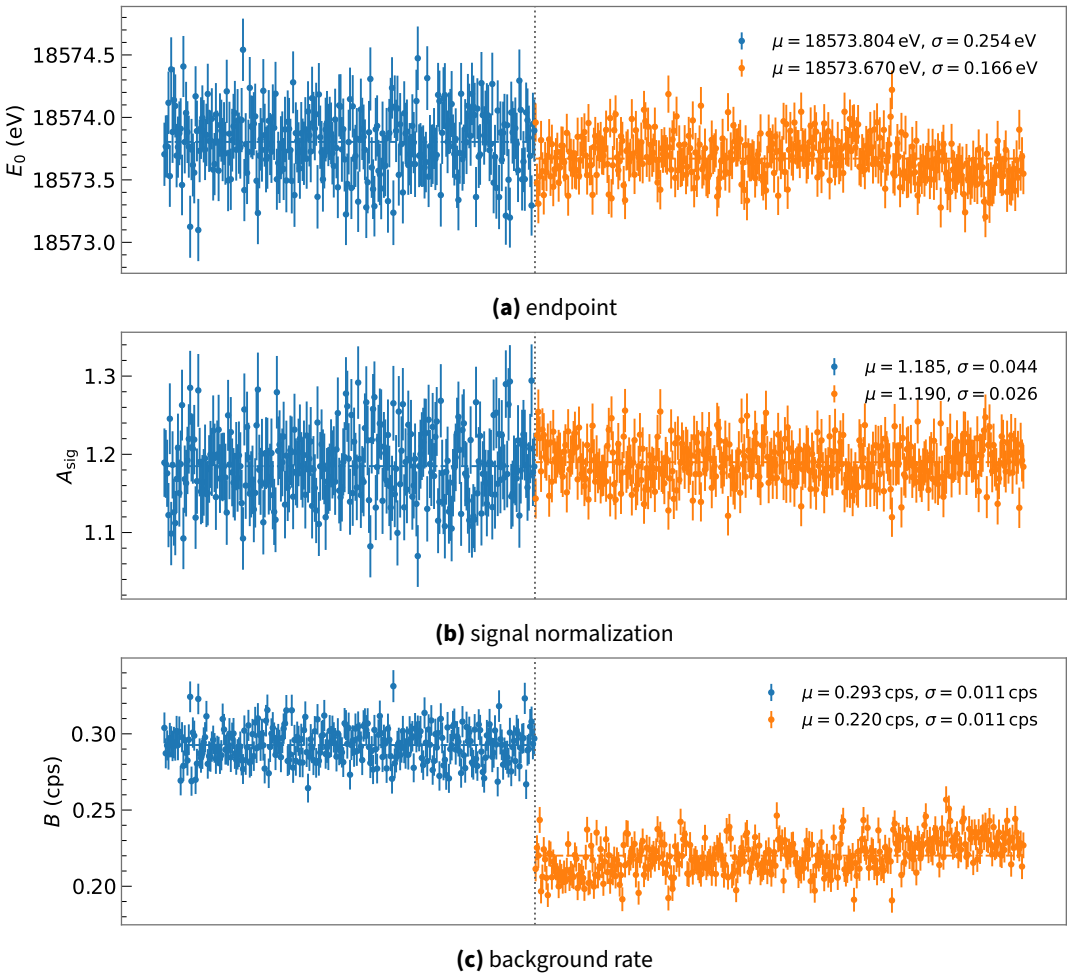
The endpoint (fig. 6.6a) shows no obvious pattern during the first campaign. In the second one, there are various indications for a drift, especially during the later part of the campaign. In both measurement phases the endpoint evolution indicates a slight overdispersion of $50\,\mathrm{meV}$ and $58\,\mathrm{meV}$ respectively. This is small compared to the expected sensitivity and was neglected for KNM1, while it was explicitly taken into account in the analysis of KNM2 as will be discussed in section 6.4.

During both campaigns the signal normalization is stable over time (fig. 6.6b) and shows no significant overdispersion compared to the statistical uncertainty. Note that we find a value of approximately 1.2 which is significantly different from the expectation of 1.0. Main reason for this is an underestimation of the fluxtube mapped to the detector [96], which only affects the normalization and does not lead to a shape effect or neutrino mass bias. The background rate (fig. 6.6c) overfluctuates in both campaigns and shows a clear increase over the course of the second campaign. One important factor for this increase is that the main spectrometer was baked out directly before KNM2. The surface conditions on the spectrometer then change over time, leading to an increased efficiency for creating Rydberg atoms and thus also an increase of the Rydberg background component.

In addition to the time evolution, we also investigate the radial dependency of these three fit parameters. This is done by fitting all scans stacked, but each ring-wise detector patch individually. The resulting parameter values over ring are visualized in fig. 6.7.

While the endpoint over ring (fig. 6.7a) shows no clear pattern in KNM1, it tends to lower values for the outer rings in KNM2. This is enforced by the fact that the p-value of $58.2\,\%$ for a constant E_0 is reasonable for the first campaign, but it is on the low end with $3.9\,\%$ for the second one. However, in either case the overall E_0 variation is small with a standard deviation of $58\,\mathrm{meV}$ and $43\,\mathrm{meV}$ respectively.

The signal normalization is stable over ring for each campaign (fig. 6.7b), underpinned by a p-value of a constant of $89.8\,\%$ or $49.2\,\%$. The background rate clearly increases to the outer part of the detector (fig. 6.7c).

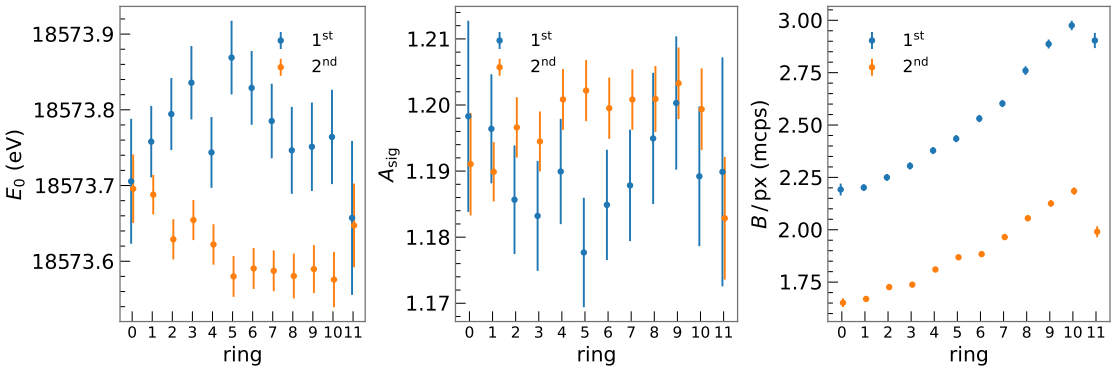


(a) endpoint

(b) signal normalization

(c) background rate

Figure 6.6: Fit parameters over scan for the first (blue) and second (orange) measurement campaign. The dashed lines indicate the corresponding mean value. Both the mean μ and the standard deviation σ are given in the legend.



(a) endpoint **(b)** signal normalization **(c)** background rate per pixel

Figure 6.7: Fit parameters over detector ring for the first (blue) and second (orange) measurement phase. The background rate was normalized with the number of pixels in each ring.

6.3 Data combination and likelihood

For the first measurement phase, we can combine all detector pixels as well as all individual scans into a single spectrum due to the low statistical power, as well as the low source activity and thus small effect of the source potential distribution from the plasma. Therefore, we end up with a single likelihood function with the regular four free parameters $\{m_\nu^2, E_0, A_S, r_{bg}\}$ for the final neutrino mass fit.

For the second campaign it was decided to include the properties of the source electric potential. This was based on the hypothesis that the higher source activity also increases the plasma effects on the source potential distribution as discussed previously. To account for any radial potential distribution, the detector pixels were grouped into rings with individual endpoints instead of treating them in a uniform manner. Any change of the potential in time was included as an additional broadening with an uncertainty. Thus, the second campaign is segmented into twelve parts, namely the twelve detector rings. This leads us to a likelihood function consisting of twelve patch likelihoods as described in section 4.2.1. The neutrino mass squared as parameter of interest is shared over all rings, but endpoint, signal normalization, and background rate are not, giving a total of $1+3\cdot 12 = 37$ free parameters in the likelihood.

Therefore, we have two individual likelihood functions, one with four free parameters for the 1st campaign and one with 37 for the 2nd campaign (not accounting for possible additional nuisance parameters describing systematic uncertainties).

As both campaigns run at significantly different settings, especially the gas density, the spectra cannot be stacked or clustered. A full analysis with a combined likelihood and shared neutrino mass as described in section 4.2.3 is the most general approach. However, as both measurement phases are strongly dominated by their statistical uncertainty, a simplified approach

where both are analysed individually and the m_ν^2 results are combined at a later stage is also viable as we will show later.

6.4 Treatment of systematic uncertainties

We now discuss how systematic uncertainties described in section 3.4 are included for the neutrino mass analysis of the first two campaigns. We will stick to the same order which roughly follows the beamline of the KATRIN experiment.

As the experiment had not been exposed to large amounts of tritium, yet, an additional beta spectrum from implanted tritium on the rear wall does not play an important role in these measurement phases. It was therefore not included in the neutrino mass analysis.

In both measurement phases, the absolute column density was calibrated using several egun measurements. For the first campaign, continous monitoring was performed using the throughput sensor, while the second one made use of a krypton capillary sensor [64, 97]. For the 1$^{\text{st}}$ measurement phase, the relative uncertainty amounted to $0.85\,\%$ leading to $\rho d = 1.11 \pm 0.01 \times 10^{21}\,\mathrm{m}^{-2}$. In the 2$^{\text{nd}}$ campaign, the higher overall gas density allowed a more precise estimation, with a relative uncertainty of $0.25\,\%$ or $\rho d = 4.23 \pm 0.01 \times 10^{21}\,\mathrm{m}^{-2}$. As response parameter affecting the scattering probabilites, the uncertainty on the column density is included using the MC propagation method.

The energy loss function was measured in situ for each campaign. The covariance K_{eloss} matrix describing the uncertainty of the nine eloss parameters was used as sampling distribution for the Monte Carlo propagation of uncertainty. For each campaign, a slightly different measurement of the parameters and thus a slightly different uncertainty matrix was used. Both pre-date the version published in [49], but are not significantly different at the given sensitivity on m_ν of about $1\,\mathrm{eV}$ ($90\,\%$ CL).

The uncertainty on the molecular final states was incorporated by fluctuating them bin-to-bin and sampling the FSD onset parameter as described in section 3.4 using the Monte Carlo method.

For both measurement phases, the activity measurements are taking from the product of the activity parameters ρd and ϵ_{T} as the FBM showed a significant drift over time. To calculate the activity for subruns without corresponding slow control measurement, it was linearly interpolated between the previous and next measurement point. From these activity measurements, the overall activity covariance matrix $\frac{K_{\text{activity}}}{n}$ can be calculated. This was included directly into the fit using the covariance matrix approach for the second campaign. In the analysis of the first campaign, the activity was sampled directly from the underlying data points using the Monte Carlo propagation method.

As the gas density of KNM1 was relatively low, any plasma inhomogeneities on the source electric potential only have a minor effect and were not included in the analysis. Any possible biases due to neglecting this effect will be discussed in section 6.7.1. With the increased gas density of the second campaign, the plasma effects become relevant. The longitudinal broadening together with the broadening induced by short-term fluctuations is determined from multiple Krypton calibration campaigns and then scaled to the KNM2 gas density resulting in $\sigma_z^2 + \sigma^2_{\text{short-term}} = 0.0124 \pm 0.0161\,\text{eV}^2$. This value also serves as input for the energy loss shift parameter ϵ_z via the inequality in eq. (3.58) assuming $k = 1.3$. The resulting sampling distribution shown in fig. 3.28 is used as input for the Monte Carlo propagation. In addition, the long-term broadening was inferred from the time evolution of the endpoint in a fit range that excludes the neutrino mass measurement interval [98]. This adds another broadening input of $\sigma^2_{\text{long-term}} = 0.0058 \pm 0.0014\,\text{eV}^2$ to be sampled independently of the other source potential parameters.

The uncertainties on the magnetic fields are estimated by comparing simulation to measurements [99–101]. For the 1$^{\text{st}}$ campaign, the following values (uncertainties) are used:
$B_{\text{src}} = 2.52\,\text{T}\,(2.5\,\%)$, $B_{\text{max}} = 4.23\,\text{T}\,(0.2\,\%)$, $B_{\text{ana}} = 6.31 \times 10^{-4}\,\text{T}\,(1.0\,\%)$,
while for KNM2 new measurements could improve the estimates to
$B_{\text{src}} = 2.52\,\text{T}\,(1.7\,\%)$, $B_{\text{max}} = 4.24\,\text{T}\,(0.1\,\%)$, $B_{\text{ana}} = 6.31 \times 10^{-4}\,\text{T}\,(1.0\,\%)$.
All magnetic field parameter uncertainties are incorporated with the usual MC propagation method.

In both campaigns, the background overdispersion was measured directly from the background data of the neutrino mass scans. This results in a Gaussian standard deviation of $\sigma_{\text{np}} = 10\,\text{mcps}$, or an overdispersion of $6\,\%$, for KNM1 and a Gaussian standard deviation of $\sigma_{\text{np}} = 12\,\text{mcps}$, corresponding an overdispersion of $11\,\%$, for KNM2. For the first campaign this uncertainty was only included in the randomization of the spectra during the Monte Carlo propagation. For each point, the mean background rate defining the Poissonian PDF is randomized according to the overdispersion, before performing the Poissonian randomization of the counts. In the second campaign, a Gaussian likelihood was used allowing to directly increase the statistical uncertainty of each point in the likelihood as defined by σ_{np}.

For these two measurement phases, the background energy dependence was constrained by a background slope measurement from the first tritium campaign [71]. The central value was neglected, but the uncertainty used as $1\,\sigma$ uncertainty on b_{qU} giving an input of $b_{qU} = 0.00 \pm 4.74\,\text{mcps}\,\text{keV}^{-1}$ over the full detector. This uncertainty was divided into the corresponding uncertainty of each detector ring for the 2$^{\text{nd}}$ campaign. In any case, the background energy slope parameter is included in the model and the uncertainty propagated using the MC method.

The Penning induced background was discovered just before the unblinding of KNM2. Therefore, it was not known and not included during the analysis of the first measurement phase. However, it was retrospectively measured to be $b_t = -2.2 \pm 4.3\,\mu\text{cps}\,\text{s}^{-1}$, showing no significant value for this campaign, possibly due to the lower gas density. The effect of including

Table 6.1: Summary of systematic uncertainties and their treatment.

systematic	input 1st	treatment 1st	input 2nd	treatment 2nd
rear wall	none	neglected	none	neglected
ρd ($10^{21}\,\mathrm{m}^{-2}$)	1.11 ± 0.01	MC prop	4.23 ± 0.01	MC prop
energy loss	9 params	MC prop	9 params	MC prop
molecular states	custom	MC prop	custom	MC prop
activity fluct.	data driven	MC sampling	data driven	cov mat
$\sigma_z^2 + \sigma^2_{\text{short-term}}$ (eV^2)	none	neglected	0.0124 ± 0.0161	MC prop
ϵ_z (eV)	none	neglected	0.000 ± 0.061	MC prop
$\sigma^2_{\text{long-term}}$ (eV^2)	none	neglected	0.0058 ± 0.0014	MC prop
B_{src} (T)	2.520 ± 0.063	MC prop	2.520 ± 0.043	MC prop
B_{max} (T)	4.230 ± 0.008	MC prop	4.239 ± 0.004	MC prop
B_{ana} (10^{-4} T)	6.311 ± 0.063	MC prop	6.308 ± 0.063	MC prop
σ_{np} (mcps)	10	MC prop	12	likelihood
b_{qU} ($\mathrm{mcps\,keV}^{-1}$)	0.00 ± 4.74	MC prop	0.00 ± 4.74	MC prop
b_t ($\mu\mathrm{cps\,s}^{-1}$)	-2.2 ± 4.3	neglected	3 ± 3	MC prop
detector	none	neglected	none	neglected

this value into the analysis will be discussed in section 6.7.1. For the 2nd campaign, the corresponding slope was measured to be $b_t = 3 \pm 3\,\mu\mathrm{cps\,s}^{-1}$. Like the background energy slope, the Penning slope parameter is included in the model and propagated using Monte Carlo propagation.

Detector related systematics play only a very minor role at the electronvolt scale sensitivity, and were thus neglected for the analysis of the first two campaigns.

All systematic uncertainties and their treatment are summarized in table 6.1.

6.5 Blinding strategy

Before coming to the neutrino mass analysis, we briefly discuss the blinding strategy pursued by the KATRIN collaboration, meant to avoid any self-bias by the analysis teams.

In a first step, unfluctuated Monte Carlo copies, so called *MC twins*, of each neutrino mass scan are created. The twin contains all information of the real scan, the same slow control values, retarding energies and measuring time. The endpoint is fixed to a representative value, usually $18\,573.7\,\mathrm{eV}$, and A_S and r_{bg} are set to represent the true signal and background rates. For each scan, the neutrino mass is set to $m_\nu^2 = 0.0\,\mathrm{eV}^2$.

The complete analysis chain must then be finalized on the Monte Carlo twins. This allows for two types of sanity check:

1. The analysis must recover the true value of $0.0\,\mathrm{eV}^2$, showing that it is unbiased.

2. The overall m_ν^2 uncertainty, as well as a breakdown showing the impact of the individual components, is estimated by multiple independent analysis teams on their MC copies. To be allowed to continue in the unblinding process, each team must implement all relevant effects and come to consistent results.

After completing the analysis on the Monte Carlo copies, the next step is to analyse the real data with a blinded model. The model is blinded by changing the width of the ground state of the molecular final state distribution. As this directly impacts the overall broadening of the model, it shifts the neutrino mass by $\approx -2(\sigma^2_{\text{true}} - \sigma^2_{\text{blind}})$ with the true ground state variance σ^2_{true} and the unknown variance of the blinded FSD σ^2_{blind}. The value of σ^2_{blind} is picked randomly and not shared with the analysis teams. This allows analysing the true data independently and comparing the best fit as well as uncertainty on m_ν^2, without revealing the actual fit result when using the correct model.

If m_ν^2 best fit and uncertainty of the different teams match on data with blinded model, the analysis is performed with the true model, revealing the final results.

6.6 Neutrino mass analysis

In this section we describe our approach to infer the neutrino mass from the data of the first two measurement campaigns using mainly the Monte Carlo propagation of uncertainty to include systematic effects. The results have been published in [43, 66, 78].

As a first step, we fit the data including only the uncertainties from Poisson statistics and the background overdispersion as well as a uniform detector segmentation for comparability. Figure 6.8 shows the data spectrum, best fit, residuals and measuring time distribution of this analysis. The higher signal rate and lower background rate of the second campaign is again clearly visible. The normalized residuals show no clear structure in the signal region in either case, with a slight hint for a positive background slope in the second campaign that will be discussed in section 6.7.4. For each campaign, the goodness of fit is excellent with a χ^2 (p-value) of 21.7 ($54.1\,\%$) and 28.4 ($24.4\,\%$) respectively.

Next, we segment the data of the second campaign into twelve detector rings as planned for the final neutrino mass analysis, and perform a fit with uncertainties from Poisson statistics and the background overdispersion. The data spectrum, model and normalized residuals of this ringwise fit is displayed in fig. 6.9. Not all rings have the same count rate, as the number of

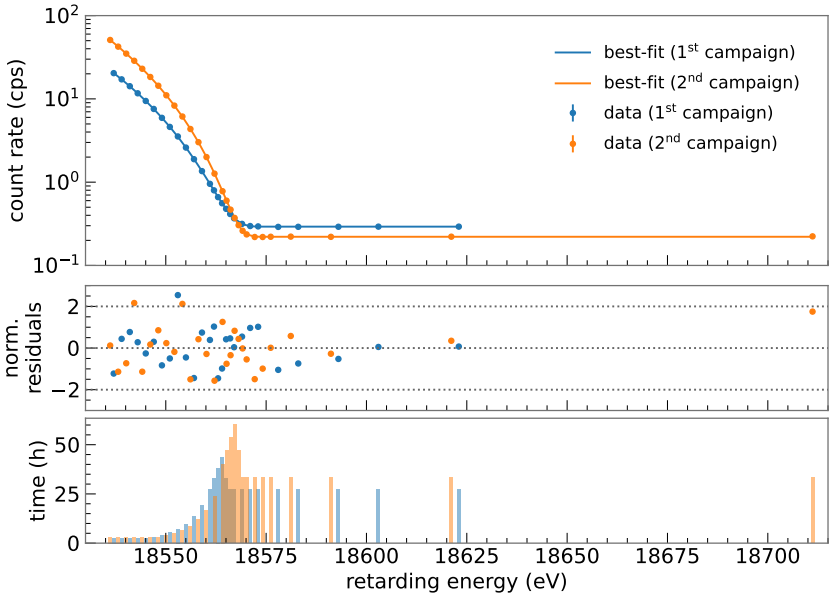


Figure 6.8: Data spectrum and best fit (top), normalized residuals (center) and measuring time distribution (bottom) for the first (blue) and second (orange) measurement campaign. Both analyses include the uncertainty from Poisson statistics and the background overdispersion and make use of a uniform detector segmentation for visualization and comparability.

pixels in the active pixel selection is not the same for each ring. There is no clear dependency of the normalized residuals over qU or ring, indicating that our model describes the data. This is underpinned by the χ^2 value of 291.4 at 299 degrees of freedom, resulting in a p-value of $61.3\,\%$.

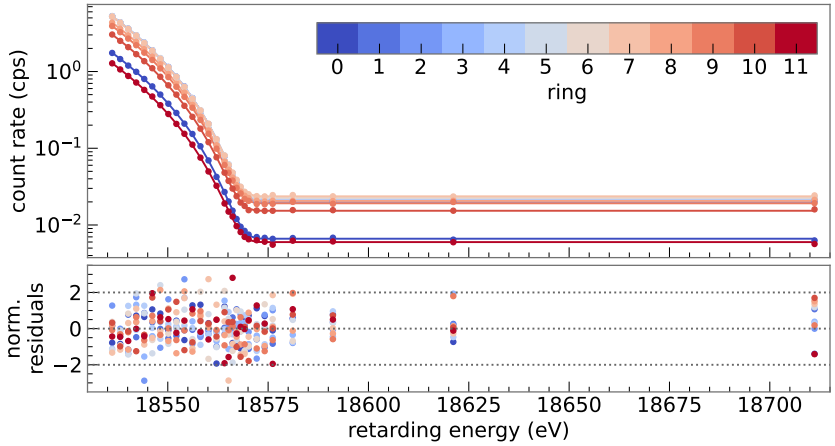
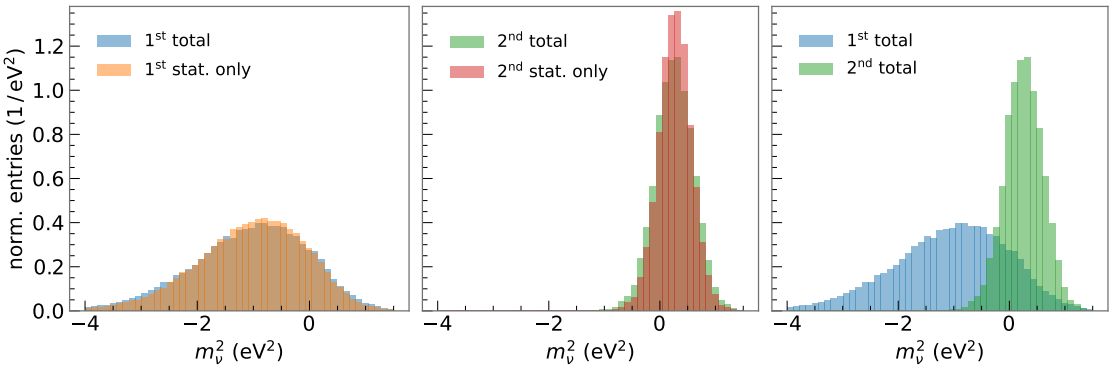


Figure 6.9: Data spectrum and best fit (top) as well as normalized residuals (bottom) for the second measurement campaign in the used ringwise segmentation. In the normalized residuals, no structure is visible over retarding energy or ring, indicating that our model describes the data well. Like fig. 6.8, this fit includes the uncertainty from Poisson statistics as well as the background overdispersion.

Having shown that our model describes the data, we come to the actual neutrino mass analysis, which we perform using Monte Carlo propagation of uncertainty. Here, we start with an analysis using the uncertainty from Poisson statistics only. Therefore, we first fit each campaign to get the best fit estimate and the corresponding model. Afterwards, we propagate the statistical uncertainty by fitting randomized spectra according to the Poisson uncertainty and our best model estimate. The resulting distribution in m_ν^2 is shown in fig. 6.10a and fig. 6.10b as the *stat. only* case. For KNM1, we retrieve $m_\nu^2 = -0.96^{+0.89}_{-1.05}\,\mathrm{eV}^2$, a roughly $1\,\sigma$ fluctuation to the negative assuming $m_\nu^2 = 0\,\mathrm{eV}^2$. The second campaign yields $m_\nu^2 = 0.27^{+0.29}_{-0.29}\,\mathrm{eV}^2$, now a $1\,\sigma$ fluctuation to the positive. We can see that the statistical uncertainty in m_ν^2 improved by more than a factor of three going from KNM1 to KNM2.

Next, we also include all systematic uncertainties in the analysis as described in section 6.4. This gives the *total* m_ν^2 distribution in figs. 6.10a and 6.10b. In both cases, the distribution only widens slightly and the best fit value remains almost unchanged, resulting in $m_\nu^2 = -0.98^{+0.95}_{-1.12}\,\mathrm{eV}^2$ and $m_\nu^2 = 0.26^{+0.34}_{-0.34}\,\mathrm{eV}^2$. Therefore, the 1$^{\mathrm{st}}$ (2$^{\mathrm{nd}}$) campaign stays with an approximately $1\,\sigma$ negative (positive) m_ν^2 best fit, showing no significant deviation from $0\,\mathrm{eV}^2$. The m_ν^2 distribution from each measurement phase is compared directly in fig. 6.10c. This clearly shows the about three times lower overall uncertainty of KNM2 with a much narrower distribution. All m_ν^2 fit results are summarized in table 6.2.

After discussing the one dimensional m_ν^2 distribution, we take a look at the two dimensional m_ν^2 - E_0 distribution of each campaign in fig. 6.11. An obvious feature is the large correlation between m_ν^2 and E_0 of 0.97 and 0.93 for the 1$^{\mathrm{st}}$ and 2$^{\mathrm{nd}}$ measurement phase respectively. It is interesting to observe that the one dimensional projections of each parameter show significant overlap of the distribution from each campaign, but the two dimensional distribution is



(a) first campaign **(b)** second campaign **(c)** comparison

Figure 6.10: Neutrino mass squared distribution from Monte Carlo propagation of uncertainty for the first two measurement campaigns. Both campaigns are dominated by statistical uncertainty as seen in (a) and (b). As expected, the distribution for the second measurement phase is much narrower (c).

Table 6.2: Best fit m_ν^2 with $1\,\sigma$ uncertainty in eV^2 for the first two measurement campaigns, once including only the uncertainty from Poisson statistics, and once with all considered uncertainties. For a full summary of all fit parameters we refer to appendix A (tables A.1 to A.4).

campaign	Poisson statistics only	total uncertainty
KNM1	$-0.96^{+0.89}_{-1.05}$	$-0.98^{+0.95}_{-1.12}$
KNM2	$+0.27^{+0.29}_{-0.29}$	$+0.26^{+0.34}_{-0.34}$

clearly split. However, this is not unexpected as the absolute endpoint value depends on the energy scale of KATRIN, which includes any work function and source potential effects. These have an overall uncertainty of approximately $500\,\mathrm{meV}$ [78], and may well be different after a spectrometer bake out, source warm up, and running at a different gas density. In addition, KNM2 was analyzed with a new calibration of the high voltage [102] which alone accounts for a difference of about $140\,\mathrm{meV}$ in E_0.

To understand how the overall uncertainty is composed, especially which systematic uncertainties play the most important role, we now calculate an uncertainty breakdown. This is done by only activating one specific uncertainty in the Monte Carlo propagation at a time, and calculating the average $1\,\sigma$ uncertainty in the resulting in m_ν^2 distribution. In case the uncertainty is not included via the Monte Carlo method, we fall back to $\sigma^2_{\mathrm{syst}} = \sigma^2_{\mathrm{tot}} - \sigma^2_{\mathrm{stat}}$.

The resulting uncertainty breakdown is displayed in fig. 6.12. All numbers are also summarized in appendix A (tables A.5 and A.6). We can see that both campaigns are clearly dominated by the uncertainty from Poisson statistics, followed by the effect of the background overdispersion. For KNM1, the first non-statistical effect is the background retarding energy dependence with an impact of $0.066\,\mathrm{eV}^2$, followed by the gas density ($0.052\,\mathrm{eV}^2$), the magnetic fields ($0.048\,\mathrm{eV}^2$)

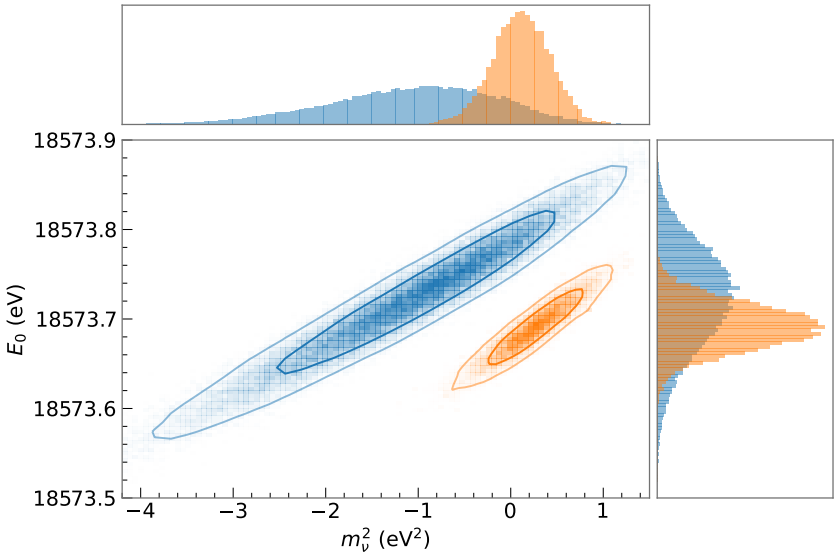
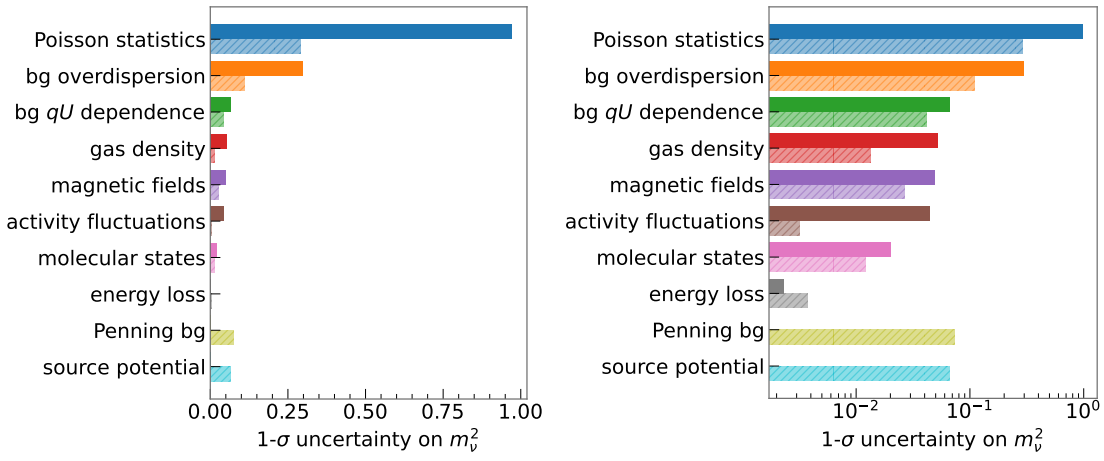


Figure 6.11: 2-dimensional neutrino mass squared - endpoint distribution from Monte Carlo propagation of uncertainty for the first (blue) and (second) measurement phase as well as their 1-dimensional projections.



(a) linear scale **(b)** logarithmic scale

Figure 6.12: Uncertainty breakdown for the first (regular fill) and second (hatched) measurement campaign. Both are dominated by statistical uncertainty, followed by background effects. Other systematic uncertainties only have a minor impact on the total error budget.

and the activity fluctuations ($0.044\,\mathrm{eV}^2$). The other uncertainties only play a minor role with values $< 0.03\,\mathrm{eV}^2$. Adding all non-statistical uncertainties (everything except for Poisson statistics and background overdispersion) in squares gives $0.11\,\mathrm{eV}^2$. This only increases the $0.97\,\mathrm{eV}$ Poisson uncertainty by less than $1\,\%$ when adding the two in squares, clearly showing how statistics dominated the first campaign is.

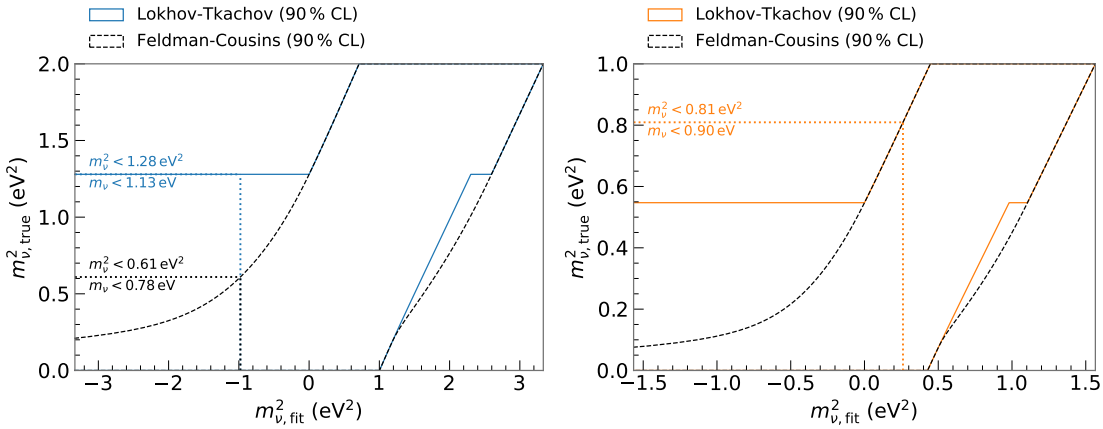
In KNM2, the dominant non-statistical effects are the two new ones: the Penning background ($0.074\,\mathrm{eV}^2$) and the source electric potential ($0.066\,\mathrm{eV}^2$). These are then followed by the same effects as in the first campaign, namely the background qU dependence ($0.041\,\mathrm{eV}^2$) and the magnetic fields ($0.026\,\mathrm{eV}^2$). The other effects all contribute less than $0.02\,\mathrm{eV}^2$ to the final budget. Here, adding the non-statistical uncertainties in squares results in a total effect of $0.11\,\mathrm{eV}^2$. This increases the Poisson uncertainty by $7.2\,\%$. Thus, the second measurement phase is still clearly dominated by statistical uncertainty, but not as much as KNM1 due to the improved statistics and new systematic effects considered.

To complete the individual analyses, we construct the confidence belt to retrieve the $90\,\%$ CL confidence interval in m_ν^2 using the method of Lokhov and Tkachov (LT) [89] as baseline and the one of Feldman and Cousins (FC) [87] for reference. The resulting belts are shown in fig. 6.13.

We see that for KNM1 the LT method yields the sensitivity limit of $m_\nu^2 < 1.28\,\mathrm{eV}^2$, as $m^2_{\nu,\mathrm{fit}}$ is negative. This results in an upper limit of $m_\nu < 1.13\,\mathrm{eV}$ ($90\,\%$ CL). Using FC, the upper limit is much stricter due to the negative $m^2_{\nu,\mathrm{fit}}$, yielding $m_\nu^2 < 0.61\,\mathrm{eV}^2$ and thus $m_\nu < 0.78\,\mathrm{eV}$ ($90\,\%$ CL).

For the second measurement campaign, we lie in the regime where LT and FC give the same upper limit of $m_\nu^2 < 0.81\,\mathrm{eV}^2$ and therefore $m_\nu < 0.90\,\mathrm{eV}$ ($90\,\%$ CL). This is therefore the first sub-electronvolt limit of a direct neutrino mass measurement. Due to the positive best fit value, the sensitivity limit of KNM2 is slightly more strict than the result derived from the data ($m_\nu^2 < 0.55\,\mathrm{eV}^2, m_\nu < 0.74\,\mathrm{eV}$).

Finally, we complete the neutrino mass analysis of KNM1 and KNM2 by combining the results from the two campaigns. As we have shown that both are clearly dominated by (uncorrelated) uncertainties of statistical nature, we perform a simple combination instead of running a combined fit at this stage. This is done by multiplying the m_ν^2 distributions resulting from the Monte Carlo propagation as shown in fig. 6.14. We then derive the combined result from the product distribution as usual: the best fit as the median, and the central $1\,\sigma$ interval by integrating from both sides. This results in a combined fit result of $m_\nu^2 = 0.12^{+0.32}_{-0.33}\,\mathrm{eV}^2$. We translate this into an upper limit of $m_\nu < 0.8\,\mathrm{eV}$ ($90\,\%$ CL) using the method of LT and assuming a Gaussian distribution of $m^2_{\nu,\mathrm{fit}} = m^2_{\nu,\mathrm{true}} \pm 0.32\,\mathrm{eV}^2$ This is only a slight improvement compared to the result of using only KNM2 where the limit was $m_\nu < 0.9\,\mathrm{eV}$ ($90\,\%$ CL).



(a) first campaign **(b)** second campaign

Figure 6.13: Frequentist confidence belt using the methods of Lokhov-Tkachov [89] and Feldman-Cousins [87] for the first (a) and second (b) measurement campaign. For both KNM1 and KNM2 an upper limit is derived, corresponding to $m_\nu < 1.13\,\mathrm{eV}$ (90 % CL) and $m_\nu < 0.90\,\mathrm{eV}$ (90 % CL) respectively using the Lokhov-Tkachov construction. As $m^2_{\nu,\mathrm{fit}}$ is roughly $1\,\sigma$ negative for the 1st campaign, and $1\,\sigma$ positive for the 2nd one, the limit by Feldman-Cousins for KNM1 is actually more strict than the one for KNM2, $m_\nu < 0.78\,\mathrm{eV}$ (90 % CL) vs. $m_\nu < 0.90\,\mathrm{eV}$ (90 % CL), despite having a significantly worse sensitivity.

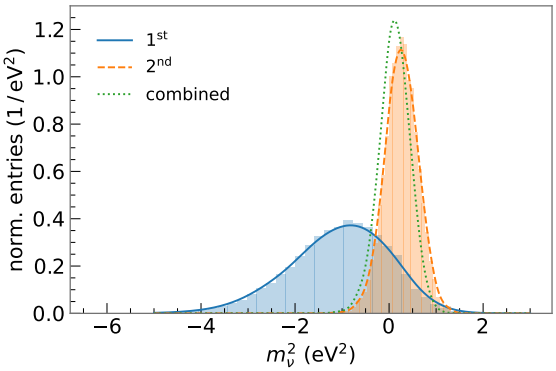


Figure 6.14: Combination of the MC propagation results of the first and second campaign. The combined result (green dotted line) is retrieved by multiplying and normalizing the individual two.

6.7 Further discussion

We will now discuss the results further, checking if any effect significantly changes our result for the neutrino mass.

6.7.1 Neglected effects during the first measurement campaign

First, we investigate the impact of including the effects that were neglected during the analysis of the first measurement campaign as published [66, 78] on the best fit value of m_ν^2. The change of the sensitivity by including additional systematic uncertainties is negligible at the statistics of KNM1. Our reference best fit value used for comparison is from the analysis including only the uncertainty of Poisson statistics yielding $m_\nu^2 = -0.96\,\mathrm{eV}^2$.

Background overdispersion in likelihood

Here we compare our reference fit to one which includes the background overdispersion in the likelihood, instead of only using it for sampling later in the MC propagation. We can see that this updates the best fit to $-0.975\,\mathrm{eV}^2$ compared to the initial $-0.96\,\mathrm{eV}^2$ when using only Poisson statistics in the likelihood.

All following values include the background overdispersion in the likelihood.

Penning background

When including the Penning background as a nuisance parameter with $b_t = -2.2 \pm 4.3\,\mu\mathrm{cps}\,\mathrm{s}^{-1}$, the best fit shifts to a slightly more negative value of $m_\nu^2 = -1.074\,\mathrm{eV}^2$. This is expected, as neglecting a positive b_t value would shift m_ν^2 to a more negative value, so neglecting an underfluctuation does the opposite.

Scattering transmission

During the initial analysis of the 1$^{\mathrm{st}}$ campaign, an isotropic transmission function was used without accounting for the effect that scattered electrons on average have larger starting angles, see section 3.3.2. Including this effect, the best fit also shifts slightly to the negative: $m_\nu^2 = -1.047\,\mathrm{eV}^2$.

Table 6.3: Impact of the neglected effects on the best fit of the neutrino mass squared for KNM1. As the source potential effects are upper bounds [103], they lead to an interval in m_ν^2 instead of a single updated value.

effect	details	m_ν^2 (eV^2)
background overdispersion	included in likelihood	-0.975
Penning background	$b_t = -2.2 \pm 4.3\,\mu\mathrm{cps/s}$	-1.074
scattering transmission	included in model	-1.047
eloss shift	$\lvert\epsilon_z\rvert < 66\,\mathrm{meV}$	$[-1.021, -0.937]$
source potential broadening	$\sigma_z^2 + \sigma^2_{\text{short-term}} < 0.0055\,\mathrm{eV}^2$	$[-0.975, -0.963]$
all effects combined		$[-1.20, -1.01]$

Source electric potential

Any effects of the source electric potential were not included, as the calibration measurements at the time were inconclusive and the effect was assumed to be small at the given sensitivity and low gas density. However, upper limits on the order of magnitude of the plasma effects have later been derived [103].

These upper bounds can be propagated into an allowed interval of m_ν^2. Combining both the effect of the energy loss shift and the potential broadening, the allowed m_ν^2 interval is $-1.02\,\mathrm{eV}^2$ to $-0.92\,\mathrm{eV}^2$, a $\mathcal{O}(5\,\%)$ effect compared to the statistical uncertainty. The interval width is largely dominated by the effect of the energy loss shift, which alone accounts for $-1.02\,\mathrm{eV}^2$ to $-0.94\,\mathrm{eV}^2$.

All effects combined

Including all neglected effects except for the source electric potential, we get a best fit value of $m_\nu^2 = -1.15\,\mathrm{eV}^2$. When including the bounds of the plasma effect, the allowed interval in m_ν^2 is $-1.20\,\mathrm{eV}^2$ to $-1.01\,\mathrm{eV}^2$. Therefore, the best fit shifts at most by $-0.24\,\mathrm{eV}^2$ compared to our reference of $-0.96\,\mathrm{eV}^2$, roughly $20\,\%$ of the statistical uncertainty. All effects and their impact are summarized in table 6.3.

This interval of best fit values is inserted into our Frequentist confidence belts ($90\,\%$ CL) in fig. 6.15. The derived upper limit in m_ν^2 is not affected when using the method of Lokhov and Tkachov [89], the main KATRIN result, as the sensitivity is quoted for negative $m^2_{\nu,\text{fit}}$. In case of Feldman-Cousins [87], the upper limit becomes more stringent from $0.52\,\mathrm{eV}^2$ to $0.56\,\mathrm{eV}^2$ compared to the value used for publication of $0.61\,\mathrm{eV}^2$. This would translate to $m_\nu < 0.72\,\mathrm{eV}$ to $0.75\,\mathrm{eV}$ instead of the published $m_\nu < 0.8\,\mathrm{eV}$.

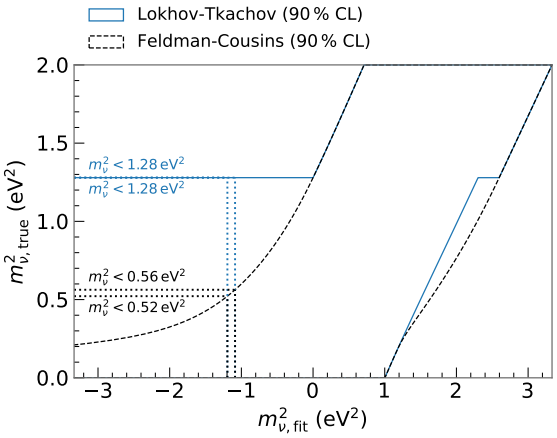


Figure 6.15: Impact of the changed best fit value from neglected effects during the first measurement campaign on the derived upper limit.

6.7.2 Effect of detector segmentation

For the first campaign, a uniform detector segmentation was used for the analysis, while the second one made use of ring-wise patches. In this part we discuss how the best fit would change in case the other option (ring-wise KNM1, uniform KNM2) would have been picked. For the following discussion we include all neglected effects in the analysis of the first measurement phase. All results are summarized in table 6.4.

We can see that the uncertainties are (almost) unaffected by the segmentation. The best fit shifts to more positive values by approximately $10\,\%$ of the uncertainty when switching from uniform to ring-wise in KNM1. In contrast, it is moved to a value lower by roughly $10\,\%$ of the uncertainty in KNM2. The overall small changes, as well as opposite signs, speak for the fact that the radial segmentation has no significant impact on the fit in any case. This is expected from a potential point of view, as the endpoint shown in fig. 6.7a also has no strong radial dependency in either campaign.

Table 6.4: Neutrino mass squared best fit and uncertainty in eV^2 using updated inputs for KNM1 depending on the measurement campaign and detector segmentation.

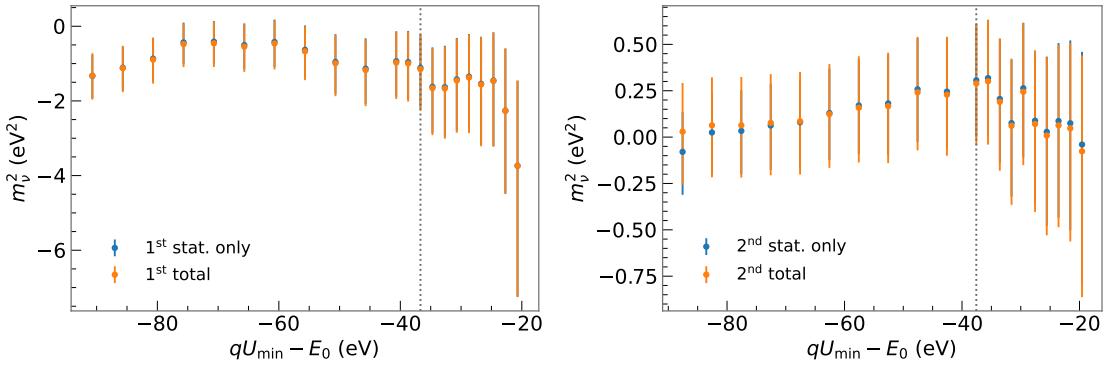
campaign	uniform	ring-wise
KNM1	$-1.14^{+0.90}_{-1.08}$	$-1.02^{+0.89}_{-1.06}$
KNM2	$+0.29^{+0.33}_{-0.33}$	$+0.26^{+0.34}_{-0.34}$

6.7.3 Neutrino mass dependencies

This part discusses the dependency of the fitted m_ν^2 on various segmentations of the data, looking for any anomalies.

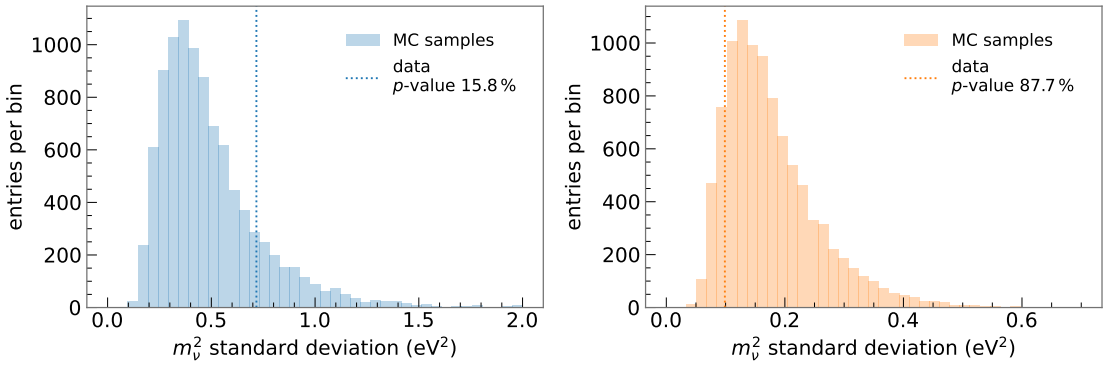
First, we take a look at m_ν^2 depending on the fit interval in qU by performing a qU *scan*. The fit interval is varied by moving the gray dotted line in fig. 6.2 to the left (right) to include more (less) points. The resulting dependency of m_ν^2 is shown in fig. 6.16. Model inaccuracies can lead to a significant variation of m_ν^2 depending on the fit range, such as a wrong modellation of the energy loss or the final states distribution. However, judging if our result is significantly different from a constant is difficult by eye as the points are highly correlated. This comes from the fact that each point to the left of another has exactly the same points in the m_ν^2 fit, except for one additional one.

Therefore, we performed a Monte Carlo simulation with $m_\nu^2 = 0.0\,\mathrm{eV}^2$ and statistically randomized spectra with the statistics of the corresponding neutrino mass campaign. For each randomized spectrum, we repeated the qU scan. As a proxy for the deviation from a constant, we calculate the standard deviation of the m_ν^2 values. We can then compare the data result to the ensemble values as shown in fig. 6.17 and calculate the probability for finding a standard deviation as large as the one in the true data. This leads to a p-value of $15.8\,\%$ for KNM1 and $87.7\,\%$ for KNM2. None of them stand out as particularly low, indicating that the qU scan results are not alarming concerning an invalid model.



(a) first campaign **(b)** second campaign

Figure 6.16: Dependency of the neutrino mass squared on the analysis interval. Shown is the best fit m_ν^2 depending on the lowest retarding energy in the analysis interval qU_{min} for the first (a) and second (b) campaign. For better visibility of the size of the analysis interval, an endpoint value of $E_0 = 18\,573.7\,\text{eV}$ is subtracted from qU_{min}. The gray dotted line indicates the analysis interval used in our main analysis.



(a) first campaign **(b)** second campaign

Figure 6.17: Ensemble test checking the overdispersion of the qU scan on the first (a) and second (b) campaign using the total uncertainty. For each statistically randomized spectrum, a qU scan is performed and the standard deviation of the neutrino mass squared values is calculated. The resulting distribution of standard deviations is then compared to the one from the qU scan on the data (dotted line).

Next, we split our data into four parts: KNM1 and KNM2a-c where KNM2a-c corresponds to the three different rear wall voltage set points during the second measurement campaign. The m_ν^2 evolution is displayed in fig. 6.18, showing no significant deviation from a constant. Calculating the χ^2 value of the weighted mean, using the average of the lower and upper error for each point, gives $\chi^2 = 2.94$ at three degrees of freedom, corresponding to a p-value of $40\,\%$. Assuming the null hypothesis of zero neutrino mass, we calculate $\chi^2 = 3.34$ at four degrees of freedom, or a p-value of $50\,\%$.

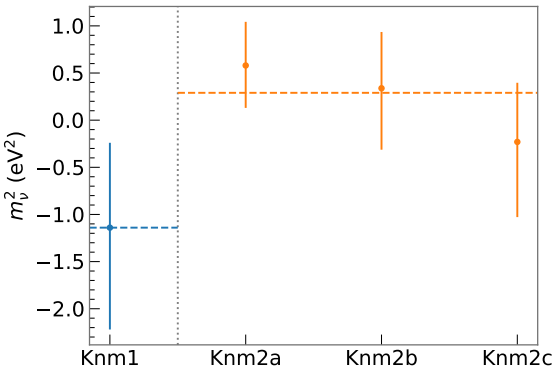


Figure 6.18: Best fit neutrino mass squared for different periods using a uniform detector segmentation and updated KNM1 inputs. The second campaign is split into three parts, corresponding to the three different rear wall voltage set points. The dashed lines indicate the best fit over the whole campaign.

Finally, we also split our data over the detector: once radially into four *pseudo rings* and once azimuthally into four slices. The segmentation as well as the resulting m_ν^2 dependency is shown in fig. 6.19 and fig. 6.20 respectively.

Neither campaign shows a clear radial dependency of m_ν^2, with a slight downward trend for KNM1 and a slight upward trend for KNM2. Fitting a constant to the m_ν^2 values gives a χ^2 (p-value) of 3.26 ($35.4\,\%$) and 3.89 ($27.4\,\%$).

When segmenting into azimuthal slices, there is no significant deviation from a constant for the second campaign with $\chi^2 = 0.88$ and the corresponding p-value of $83.1\,\%$. For the first campaign, the zeroth slice (top right) shows a larger m_ν^2 value than the three other slices. This is reflected in a worse χ^2 for the fit with a constant of 10.68 yielding a p-value of $1.4\,\%$. While this is not yet critical, it is on the lower end of acceptable values. Calculating the difference in m_ν^2 between slice zero and the other three gives $\Delta m_\nu^2 = 8.0 \pm 2.5\,\mathrm{eV}$, an effect slightly larger than $3\,\sigma$. However, there is no physical explanation for such an effect, and it has not been observed in any of the following neutrino mass campaigns with higher sensitivity.

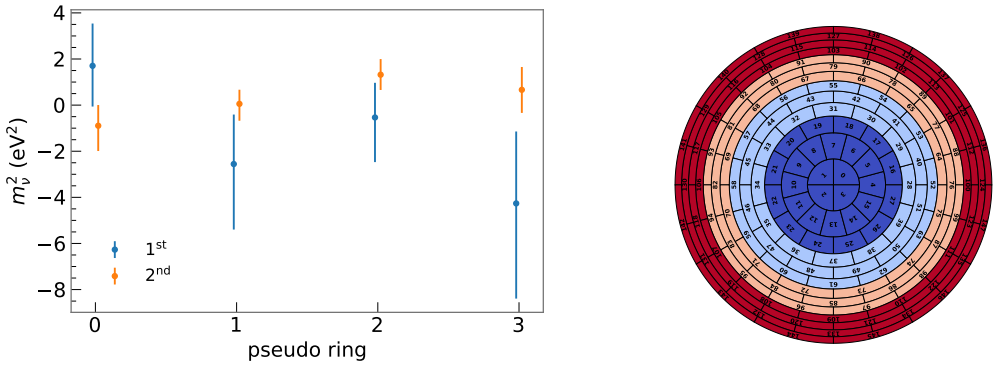


Figure 6.19: Best fit neutrino mass squared over the four pseudo rings displayed on the right for the first two measurement campaigns. Numbering starts in the inner part of the detector. The fit includes uncertainties from Poisson statistics and the background overdispersion.

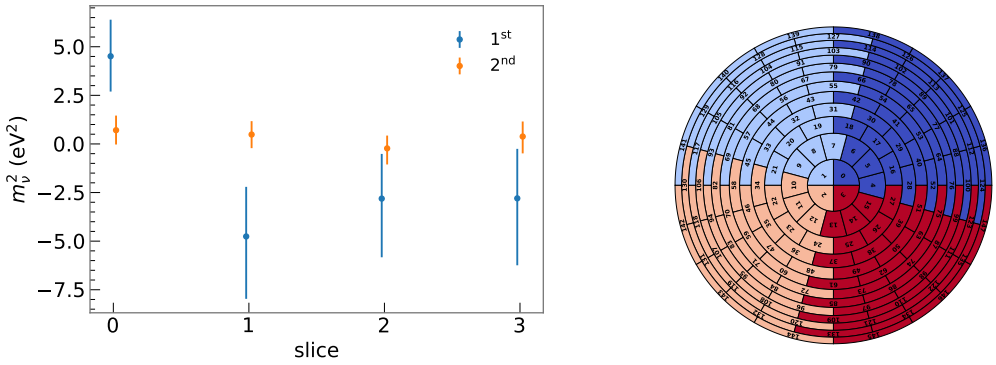


Figure 6.20: Best fit neutrino mass squared over the four azimuthal slices displayed on the right for the first two measurement campaigns. Numbering starts in the top right and continues in mathematically positive direction. The fit includes uncertainties from Poisson statistics and the background overdispersion.

6.7.4 Impact of the background systematic parameters

Finally, we want to discuss the impact of the background systematic parameters, namely the background retarding energy slope b_{qU} and the Penning slope b_t on the fit result. To this end, we discuss three cases each:

1. The parameter value is fixed to zero.

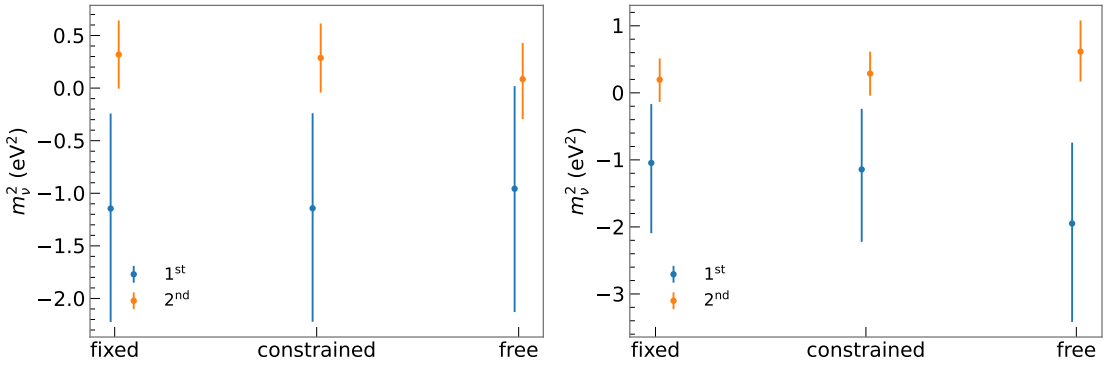
2. The parameter is included as nuisance parameter in the fit and constrained by a pull-term.

3. The parameter is completely free in the fit.

For each case, we include all other sytematic uncertainties, and keep a uniform detector segmentation. The results are visualized in fig. 6.21. Logically, the uncertainty in m_ν^2 increases with the increased freedom in the fit in each step. Therefore, we will focus on the development of the central value in this discussion.

Freeing the background retarding energy slope in KNM1 leads to $b_{qU} = -16 \pm 40\,\mathrm{mcps\,keV^{-1}}$, an insignificant negative value. This shifts the fitted m_ν^2 to slightly more positive values. In KNM2 we fit $b_{qU} = 25 \pm 12\,\mathrm{mcps\,keV^{-1}}$, a roughly $2\,\sigma$ positive value that reduces the fitted m_ν^2 by about $0.2\,\mathrm{eV^2}$. This positive slope is mainly driven by the point above $18\,700\,\mathrm{eV}$ with large lever-arm and a rate over-fluctuation (see fig. 6.8). At $2\,\sigma$, this is not yet critical for the m_ν^2 result, but this potential issue of depending on a single point for b_{qU} will be avoided by not having a single point with large lever-arm in future measuring time distributions.

Keeping the Penning slope completely free in the fit is tricky as it is highly correlated with the neutrino mass squared, and thus blows up the corresponding uncertainty. Nevertheless, it is interesting to observe how extreme this effect is. For the first campaign, a best fit of $b_t = -29 \pm 19\,\mathrm{\mu cps\,s^{-1}}$ drives the fitted m_ν^2 down to $m_\nu^2 = -1.95^{+1.21}_{-1.47}\,\mathrm{eV^2}$. The opposite sign can be observed in the second measurement phase, where $b_t = 17 \pm 13\,\mathrm{\mu cps\,s^{-1}}$ moves the fitted m_ν^2 up to $m_\nu^2 = 0.62^{+0.46}_{-0.45}\,\mathrm{eV^2}$. For both campaigns, the m_ν^2 value is shifted by roughly $1\,\sigma$ of the initial uncertainty with constrained b_t. This makes it even more clear that the Penning slope is critical for any future neutrino mass analysis. Therefore, this effect is avoided in measurement campaigns after KNM4 completely by turning off the pre-spectrometer.



(a) background qU slope

(b) Penning slope

Figure 6.21: Neutrino mass squared fit result depending on the treatment of the background qU slope (a) and the Penning slope (b) parameter.

6.8 Conclusion

We conclude this chapter by summarizing the main results of the analysis of the first two measurement campaigns, and highlighting the efforts made over the course of this thesis.

To analyse the first neutrino mass campaign, we introduced a novel approach to include systematic effects, the full Monte Carlo propagation of uncertainty described in chapter 5, to the KATRIN collaboration. Along with this, a new code base for data handling, modelling and analysis building upon the work performed in [104] was developed. This set of tools was successfully applied to KNM1, leading to a limit of $m_\nu < 1.1\,\mathrm{eV}$ (90 % CL) using the method of Lokhov and Tkachov [89] and including all relevant systematic uncertainties. Using the MC propagation method, we performed an uncertainty breakdown to show that this campaign is fully dominated by statistical uncertainties. Further tests have shown that the neutrino mass result is not changed significantly by effects identified after publishing our initial results [43, 66, 78].

For the second measurement campaign, running at higher gas density, we split our model into twelve detector rings, therefore significantly increasing the strain on our analysis tools. In addition, some new systematic effects were identified or became relevant at the increased gas density, mainly the source electric potential and the Penning background. With these new and all other relevant systematic uncertainties, we limit the neutrino mass to less than $0.9\,\mathrm{eV}$ (90 % CL) using only the data of KNM2. When including the first campaign, this slightly improves to $0.8\,\mathrm{eV}$ (90 % CL). This corresponds to the first direct neutrino mass measurement below $1\,\mathrm{eV}$.

Chapter 7

Future proof KATRIN analysis using a neural network

After analyzing the first two measurement campaigns using the Monte Carlo propagation method, the second major effort of this thesis was to develop a future proof KATRIN analysis using a neural network based approach. The results were published in EPJ-C [105].

This chapter discusses the need for a new analysis method, explains our neural network based approach, validates its performance and applies it to various realistic KATRIN scenarios.

7.1 Computational limitations of current analysis methods

First, we discuss why all of the current analysis methods require excessive computational power and would thus benefit from a fast model calculation.

Likelihood evaluation

The most important point is that the evaluation of the KATRIN likelihood described in section 4.1 in itself is computationally expensive. There are two main reasons for this.

First of all, the model evaluation alone is expensive due to the amount of nested numerical estimations such as integration and root search. The integrated spectrum defined in eq. (3.47) numerically integrates the differential β-spectrum over the response function. In the β-spectrum, we sum over hundreds of final states. However, the really expensive part of the integrand is the response function. In itself, it already consists of the convolution of the transmission function with the energy loss function for s scatterings, summing over up

to 10 of these convolutions as shown in eq. (3.46). To calculate the energy loss function for s scatterings, we need to convolve the initial energy loss function for a single scattering $s-1$ times with itself. The transmission function includes an additional integration to include the effect of scattering on the angular distribution (eq. (3.45)), and a root search to include the energy loss of electrons via synchrotron radiation (eq. (3.33)). We can easily see how putting all this together leads to a very expensive model evaluation. While parts can be pre-calculated and interpolated depending on the analysis, we will see in the following paragraphs that this is not sufficient without more advanced methods.

In addition to the model being expensive in itself, the likelihood can be split into various parts depending on the detector and time segmentation of the data analysis as shown in eq. (4.14). This can lead to several hundred, or even thousands of model evaluations for a single likelihood call.

This is especially valid for upcoming data analyses. There are already four measurement phases taken in the shifted analyzing plane setting (KNM3a, KNM4, KNM5 and KNM6) and three in the symmetric field setting (KNM1, KNM2, KNM3b) as shown in table 3.2. The much higher statistics will no longer allow the simple combination approach pursued for the first two campaigns as time correlated systematics become more relevant with the increased statistics. In addition, the SAP setting does not allow for a uniform detector, but requires the segmenting into 14 patches. This gives us $14 \cdot 4 + 1 \cdot 3 = 59$ segments for the data already taken, with multiple SAP campaigns still planned for the future. We will show in the following parts that this already at the edge of feasibility using the regular numerical model.

Nuisance parameter method

In the nuisance parameter method, all parameters with an uncertainty are included in the minimization, and constrained by their corresponding pull term. This has two important impacts on the computational requirements. Firstly, the dimensionality of the optimization problem has increased which leads to more likelihood calls needed in general. On top of this, no part of the model which depends on the free parameters can be pre-calculated fully. This is especially true for the expensive response function as it depends for example on the magnetic fields, gas density and parameters of the energy loss function.

Analyzing a single measurement campaign, the detector segmented into 12 rings or 14 patches, already requires on the order of 10^{10} evaluations of the integral spectrum when including the required estimation of derivatives. Using the regular numerical model, this can take up to one CPU year with drastic parallelization being difficult due to the serial nature of minimization.

The computational requirements scale at least quadratically with additional measurement campaigns when performing a combined fit. This is the case as both the number of data points

and the number of free parameters increase linearly. Therefore, this method is unfeasible for final KATRIN analysis using the unmodified numerical model.

Monte Carlo propagation method

In the Monte Carlo propagation approach, all parameters with systematic uncertainty are sampled and kept fixed during the fit. This allows pre-calculating the expensive response function for each sample, but requires performing thousands of fits. Thus, the overall process also becomes computationally expensive when each individual fit already requires a lot of computing time.

We already noticed this being an issue during the analysis of the second measurement campaign. Our final result with all uncertainties included, see section 6.6, included approximately $100\,000$ samples which took a total of about 50 CPU years. In contrast to the nuisance parameter approach, the MC method is embarrassingly parallel. However, at some point it makes no sense to try to solve the problem via brute-force.

This can be seen by estimating the required computation time for a final KATRIN analysis with MC propagation. The scaling of computing time remains quadratic in the number of campaigns, as our free fit parameters and number of data points still increase linearly. Using 15 campaigns, and the time for KNM2 as reference for a single campaign, we estimate $\approx 15^2 \cdot 50 = 11\,250$ CPU years. Even using a computing cluster with a very generous 1000 nodes and 32 cores each, this still amounts to a real time of

$$\frac{11\,250\,\text{CPU years}}{1000 \cdot 32\,\text{CPUs}} \approx 128\,\text{d} \tag{7.1}$$

for a single analysis.

Full Bayesian sampling

In a full Bayesian sampling approach, all parameters with an uncertainty are varied in every step, much like in the nuisance parameter method. Therefore, the response function can also not be pre-calculated. In general, sampling is more expensive than minimizing, shown by about 10^{11} model evaluations for sampling a single measurement campaign segmented into 12 patches, which makes the full Bayesian method infeasible already for the analysis of KNM1 and 2.

Instead, a model variation technique similar to the MC propagation was pursued [43]. Once again, the parameters with systematic uncertainties are randomized before now running a full Markov Chain Monte Carlo chain. Finally, the samples from all chains are combined to

incorporate the effect of systematic uncertainty. This approximated hybrid method requires at least the computing time of the MC propagation, and can therefore also not be scaled as is.

Covariance matrix approach

The covariance matrix approach is in a different situation as the approaches described above. While it is not limited due to computational reasons, it requires the approximation of a normal likelihood function instead of a Poissonian likelihood to be valid. This is no longer the case when the data is strongly segmented, as is the case for 14 detector patches and short measurement campaigns. Therefore it is not clear whether it is possible to pursue this approach to analyse the campaigns in the shifted analyzing plane configuration.

7.2 Approximating the KATRIN model with a neural network

As we have seen, all current analysis methods would benefit greatly from a fast calculation of the full integrated tritium spectrum $I(\boldsymbol{\theta}_{\text{spec}}; qU)$ as described in eq. (3.47). Here $\boldsymbol{\theta}_{\text{spec}}$ denotes any parameters that impact the integrated spectrum, such as m_ν^2 and E_0, but also the response parameters like the magnetic fields and the gas density.

One possible solution to this problem is to pre-calculate I for multiple samples of $\boldsymbol{\theta}_{\text{spec}}$ and to use multi-dimensional interpolation to then retrieve I for arbitrary values of $\boldsymbol{\theta}_{\text{spec}}$. Unfortunately, the high dimensionality in the spectral parameters of $\mathcal{O}(10)$, combined with the stringent accuracy requirements of KATRIN, makes traditional interpolation algorithms such as cubic splines or (k-) nearest-neighbor unfeasible.

Our solution to this interpolation problem is to make use of a NN. We now describe the architecture, sample generation and training of the NN before we proceed with analyzing its behaviour.

Network structure

We would like our network to predict the integral spectrum within a given range of parameter values $\boldsymbol{\theta}_{\text{spec}}$. For this, we must pass $\boldsymbol{\theta}_{\text{spec}}$ as input to the neural network. To allow the NN to learn correlations and spectral shape effects, we use the full spectrum $r(\boldsymbol{\theta}_{\text{spec}}; qU) = I(\boldsymbol{\theta}_{\text{spec}}; qU) + r_{\text{bg}}$. This leaves us with one input node for each parameter in $\boldsymbol{\theta}_{\text{spec}}$ and one output node for each qU-point in our spectrum.

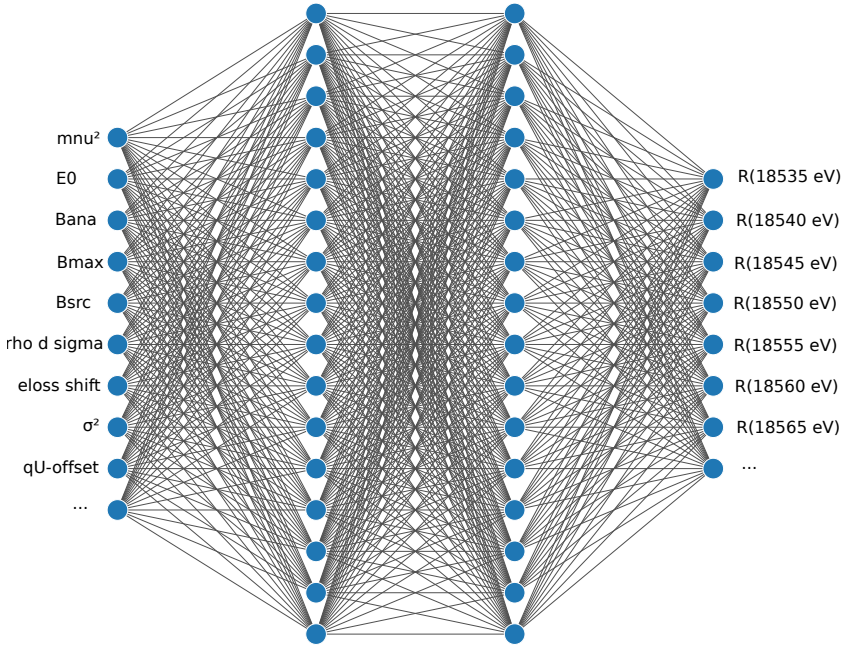


Figure 7.1: Structure of our neural network. We use a total of four layers: an input layer with one node for each parameter, two fully-connected hidden layers with 128 nodes (amount reduced for visualization) and an output layer with one node for each point in the spectrum.

To connect the input and output layer, we insert two fully-connected hidden layers, each with 128 nodes. For these hidden layers, we use the *mish* activation function,

$$\mathrm{mish}(x) = x \cdot \tanh\left(\mathrm{softplus}(x)\right) \tag{7.2}$$

$$\mathrm{softplus}(x) = \ln\left(1 + e^{x}\right) \tag{7.3}$$

while we use *softplus* as activation for our output layer. Both activation functions are differentiable, allowing us to use the network model with a gradient-based minimization algorithm. The *softplus* function ensures the output of our net is always positive.

The structure of our NN is summarized in fig. 7.1 and the activation functions are displayed in fig. 7.2.

Sample generation

Before we can train our network, we first need a large set of training data. To this end, we sample our input parameters $\boldsymbol{\theta}_{\text{spec}}$ using the N-dimensional R_2-method [106] as it provides excellent coverage of the parameter space. For each sample, we first retrieve N pseudo-random numbers distributed uniformly between zero and one. These samples and their projection are visualized in fig. 7.3a and fig. 7.3b respectively. We then transform these uniformly distributed numbers

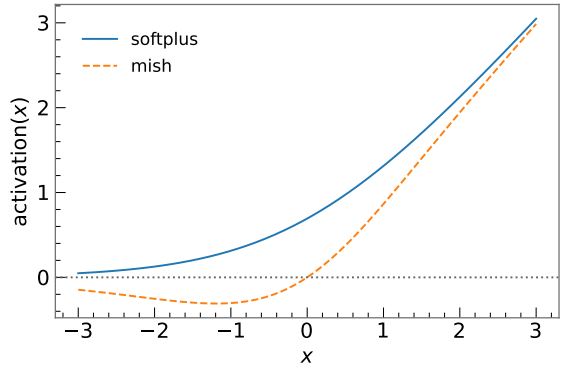
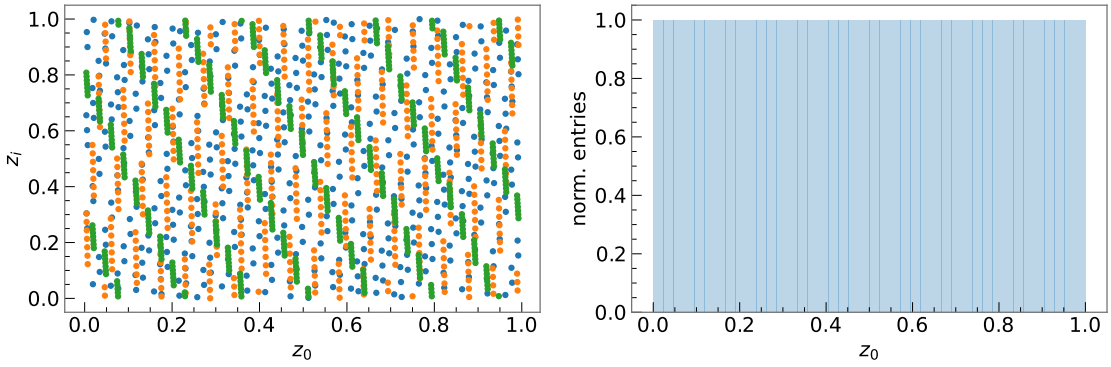


Figure 7.2: Activation functions of our neural network. Both *mish* and *softplus* are continuos and differentiable. The *softplus* activation ensures positive output values.



(a) first 500 samples in the z_0-z_i-plane for $i = \{1, 2, 3\}$

(b) projection of z_0 for $500\,000$ samples

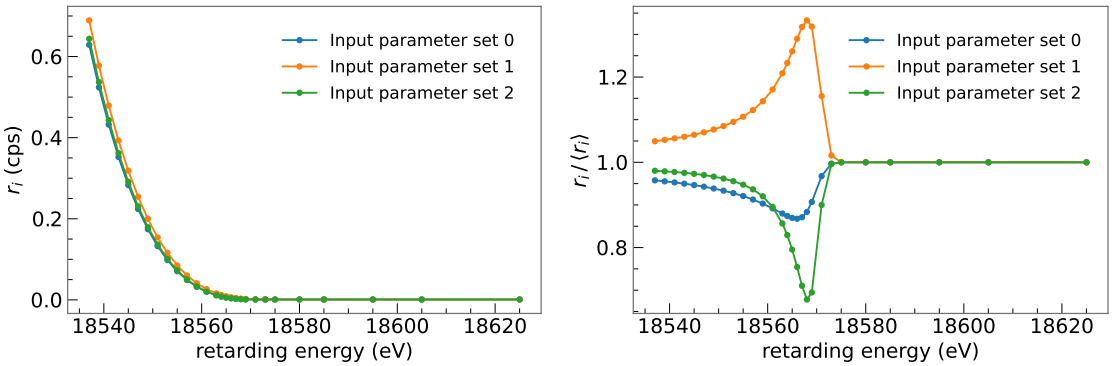
Figure 7.3: Uniformly distributed samples using the 9-dimensional R_2-method [106].

to the required parameter distribution. Usually, samples are generated covering the expected 1, 3 and $5\,\sigma$ range of each parameter using both a normal and a uniform distribution over this range. The $1\,\sigma$ interval is determined either by the systematic uncertainty for constrained parameters, or by the statistical sensitivity for the unconstrained ones.

After converting the samples into our parameter space, we can calculate the integrated spectrum I for the given parameters $\boldsymbol{\theta}_{\text{spec}}$. Repeating this process up to a few million times for each range gives the full dataset we can use for training. Typically, generating the several million samples required to train our NN takes about one day on a regular computing cluster.

Training

For training, we perform two additional transformations to our sample data. There is no need for decorrelation as the R_2-method provides uncorrelated samples for each dimension, but we



(a) raw spectra

(b) spectra divided by sample mean

Figure 7.4: Sample spectra used to train the neural network. The individual spectra differ in the input parameter values used. While all follow the same shape (a), the shape effect is clearly visible after dividing by the sample mean (b).

normalize all input parameters to have a mean of zero and a standard deviation of one. As all tritium spectra follow roughly the same shape, we divide the output rates by the sample mean at each point in qU. This allows the network to learn the more prominent changes, instead of having to learn small changes on top of an underlying tritium spectrum as shown in fig. 7.4. The division by the sample mean is also the reason for adding a constant background rate r_{bg} to the training samples, as we thus avoid dividing by zero above the endpoint. In addition, the shape effects resemble the data more closely, where the background rate is also non-zero.

During training, we optimize the weights of the neural network to minimize our loss function

$$\text{loss}(\text{weights}) = \left\langle \left(C_{\text{i}} - C_{\text{pred,i}}(\text{weights})\right)^2 \right\rangle \tag{7.4}$$

with the true rate change of each sample $C_{\text{i}} = \frac{r_i}{\langle r_i \rangle}$ and the corresponding prediction of the neural net $C_{\text{pred,i}}$. For this optimization problem, a custom interface of scipy's [107] L-BFGS [108, 109] minimizer to keras was implemented. As it is not feasible to process the full training dataset at once, we split it into batches smaller than $1\,000\,000$ spectra. We then train the network in an iterative approach:

1. Start with a minimizer tolerance of 10^{-3}.
2. Select a random batch of training samples.
3. Perform the weight optimization and store the weights.
4. Decrease the minimizer tolerance and repeat steps 2 and 3 using the stored weights as starting values.
5. Repeat the process until the loss stabilizes, this typically takes about 40 iterations with a

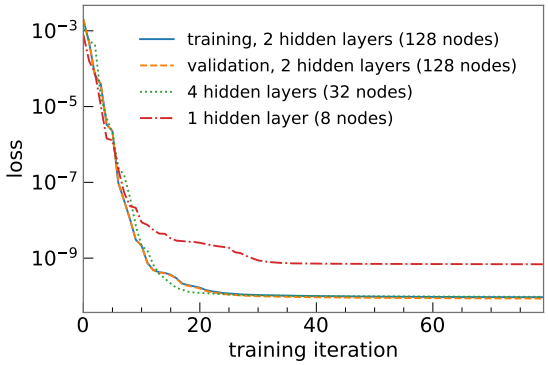


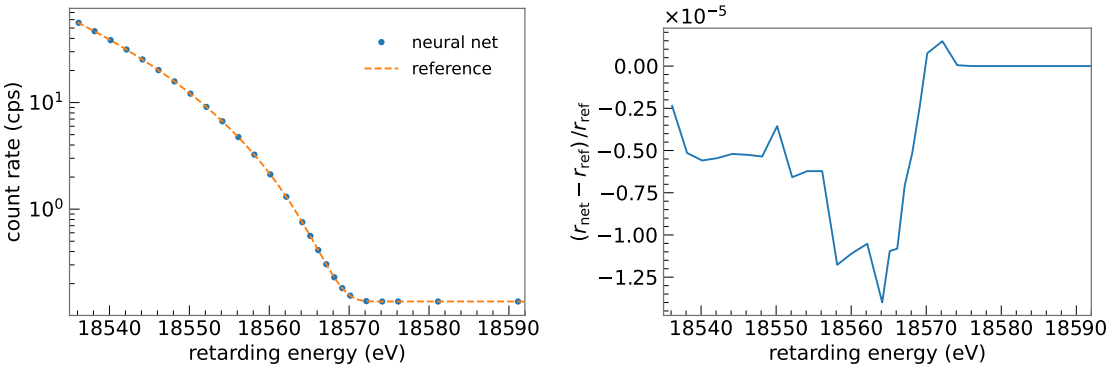
Figure 7.5: Loss over training iteration for different network architectures. The loss on the training (blue solid) and validation (orange dashed) dataset of our reference network follow one-another perfectly, indicating that there is no overfitting. Another network with a similar number of free weights (green dotted) converges to the same final loss as our reference. However, oversimplifying the NN (red dash-dotted) leads to an increase in the achievable loss and thus to a worse performance.

final minimizer tolerance around 10^{-20}.

To check for so-called *overfitting*, we split our samples into two parts: the actual training dataset which is used for the optimization procedure ($90\,\%$ of the samples) and a validation dataset ($10\,\%$) the net never sees during training. In each iteration, we evaluate our loss function both on the training and on the validation dataset. If the loss improves on the training dataset, but stays stagnant or even increases on the validation dataset, one would speak of overfitting. The evolution of our loss function for different training iterations is shown in fig. 7.5. First, we can see that both the training and the validation loss of our reference implementation with two hidden layers, each with 128 nodes, follow one-another perfectly, so there is no indication of overfitting. Both converge to a loss of approximately 10^{-10} after 40 iterations. A network with a slightly different architecture, but a similar number of free weights, converges to the same loss. However, if the underlying net is too simple, we see a significant increase in the achievable loss as shown by the NN using one hidden layer with eight nodes. This shows that the specific architecture we pick does not impact the final result, as long as the freedom in the NN parametrization is sufficiently large to pick up all features in our training data.

The complete training process is completed within a few hours on a single GPU. An important point to make is that the sample generation and training only have to be performed once, and the resulting trained network can then be used for multiple different analyses, such as a complete uncertainty breakdown, fits with different data segmentation and so forth.

Our next step is to compare the behaviour of the neural net based model with the reference, ensuring the NN learns the spectral features correctly.



(a) spectral overlay

(b) relative difference

Figure 7.6: Comparison of the output count rate of the neural net r_{net} and the reference model r_{ref}. The simple overlay (a) shows a perfect match. Displaying the relative difference (b) indicates a 10^{-5} effect. Note that the spectra are taken as-is using the same input parameters, no fit is performed.

7.3 General behaviour of our neural network based model

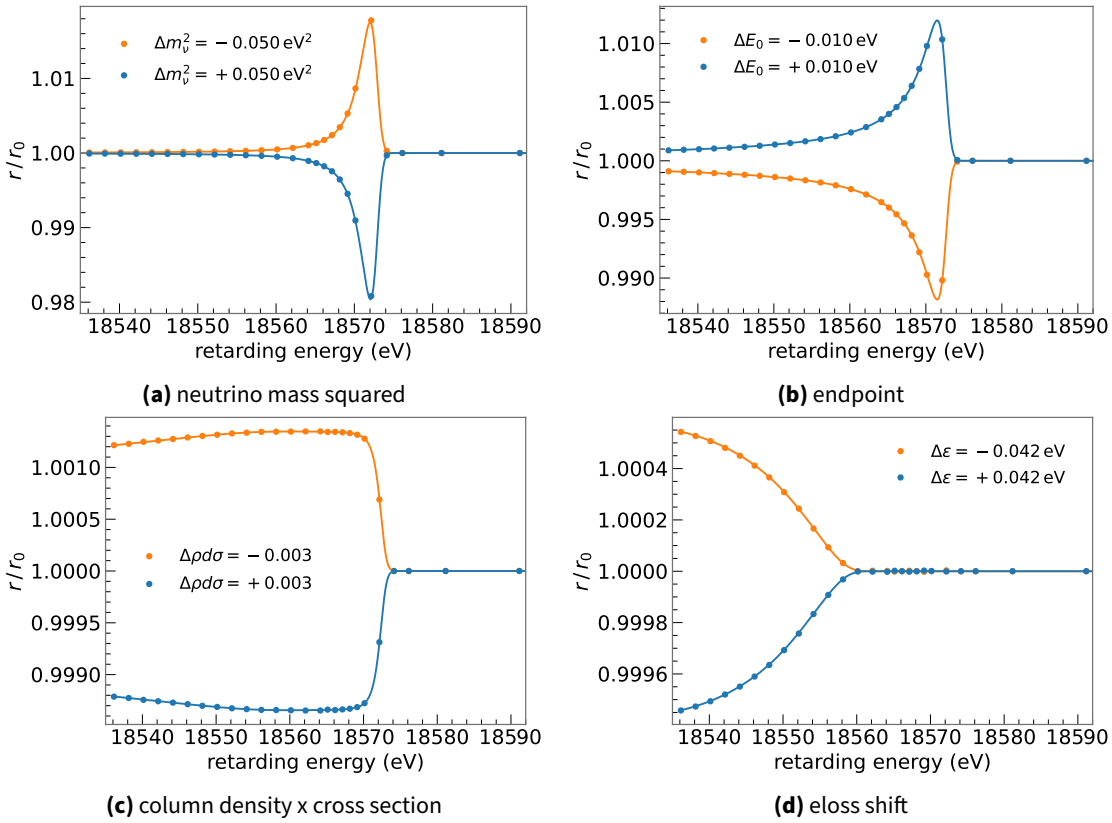
To analyze the general behaviour of our neural network based model, we initially compare the count rate of the neural net r_{net} with the output of the reference model r_{ref} in fig. 7.6 assuming $\boldsymbol{\theta}_{\text{spec}}$ similar to those in the 2$^{\text{nd}}$ measurement phase. A direct overlay of the two spectra (fig. 7.6a) shows no visible difference. Therefore, we take a closer look in fig. 7.6b by displaying the relative difference. We can see that the maximum deviation is on the order of 10^{-5}, a small effect compared to relative statistical error

$$\frac{\sigma_i}{r_i} > 10^{-3} \,\forall\, i \tag{7.5}$$

of the KNM2 data (stacked, uniform detector).

For the direct comparison, we assumed the best estimate of our parameter values, i.e. the values at which the sampling distributions of the neural net are centered. We denote this expected rate with r_0 where fig. 7.6b has shown $r_{0,\text{net}} \approx r_{0,\text{ref}}$. Now, we check the accuracy of our neural network when varying a parameter value to calculate the rate r away from r_0. To do so, we display the relative change $\frac{r}{r_0}$ in fig. 7.7 for both our NN and the reference model.

The relative change is displayed with a varied neutrino mass squared (fig. 7.7a), endpoint (fig. 7.7b), column density (fig. 7.7c), and eloss shift (fig. 7.7d). In all cases, the points stemming from our NN calculation lie on the line which denotes the estimate using our reference model. We can thus conclude that our neural net learns the shape effect of each parameter, and is able to accurately calculate the model rate in different parts of the sample space.



(a) neutrino mass squared

(b) endpoint

(c) column density x cross section

(d) eloss shift

Figure 7.7: Impact of changing various parameter values of the integral spectrum. We display the rate with an updated parameter value divided by the initial rate for the NN (points) compared to the reference model (solid line). All points lay on the line, showing that the net learns the dependence of the spectral rate on each parameter.

7.4 Comparison with existing model on a 1000-day Asimov spectrum

As our NN is able to learn the model rate and the effect of individual parameters accurately, the next step is to perform a direct comparison with the existing model on a 1000-day Asimov spectrum.

To this end, we generate a Monte Carlo dataset with the reference model. For this dataset, we assume the measurement time distribution and slow control parameter values of KNM2 as well as the fit parameter values summarized in table 7.1. First step of our comparison is to fit this dataset with our neural network, assuming a uniform detector segmentation and using only the regular four fit parameters $\{m_\nu^2, E_0, A_\mathrm{S}, r_\mathrm{bg}\}$, thus neglecting any systematic uncertainties. This simplification allows us to compare to the reference model in detail, as the analysis is computationally feasible in both cases.

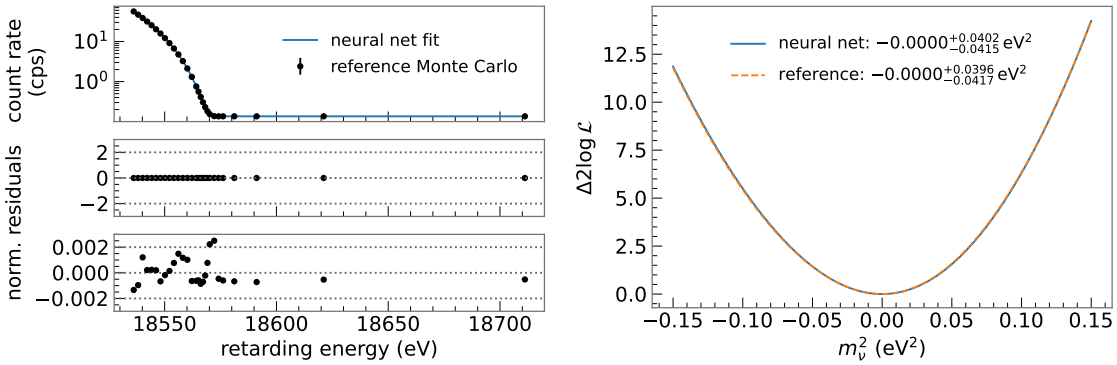
Figure 7.8a shows the neural net fit as well as the resulting normalized residuals. There is no structure in the residuals when scaled to the usual $1\,\sigma$ level. When zooming in, we see residuals on the $0.002\,\sigma$ level, about three orders of magnitude smaller than the statistical uncertainty, showing our neural network can describe the data perfectly. This perfect description is underpinned by the m_ν^2 bias of less than $10^{-5}\,\mathrm{eV}^2$, and exact recovery of all other fit parameters, as shown in table 7.1.

In addition to the recovery of all central values, we check if the neural network is also able to estimate the uncertainty correctly. Therefore, we perform a profile likelihood in m_ν^2 on the Asimov dataset with the NN and the regular analysis framework. The two are overlaid in fig. 7.8b. The profile likelihood of the neural net gives $m_\nu^2 = 0.0000^{+0.0402}_{-0.0415}\,\mathrm{eV}^2$, while the reference is $m_\nu^2 = 0.0000^{+0.0396}_{-0.0417}\,\mathrm{eV}^2$. Thus, the difference in the lower (upper) error is $2\times10^{-4}\,\mathrm{eV}^2$ ($-6\times10^{-4}\,\mathrm{eV}^2$), less than $2\,\%$ of the statistical uncertainty. We therefore conclude that the neural network biases neither the central value, nor the $1\,\sigma$ interval significantly at final KATRIN statistics on an Asimov dataset.

To ensure this statement also holds on statistically randomized data, we perform an ensemble

Table 7.1: Parameter recovery of the Asimov cross-fit. Our neural network recovers all true parameters with the bias in m_ν^2 being less than $1 \times 10^{-5}\,\mathrm{eV}^2$.

fit parameter	true value	recovered value
$m_\nu^2\,(\mathrm{eV}^2)$	0	-9.8×10^{-6}
$E_0\,(\mathrm{eV})$	18 573.700	18 573.700
A_S	1.18	1.18
$r_\mathrm{bg}\,(\mathrm{mcps})$	136	136



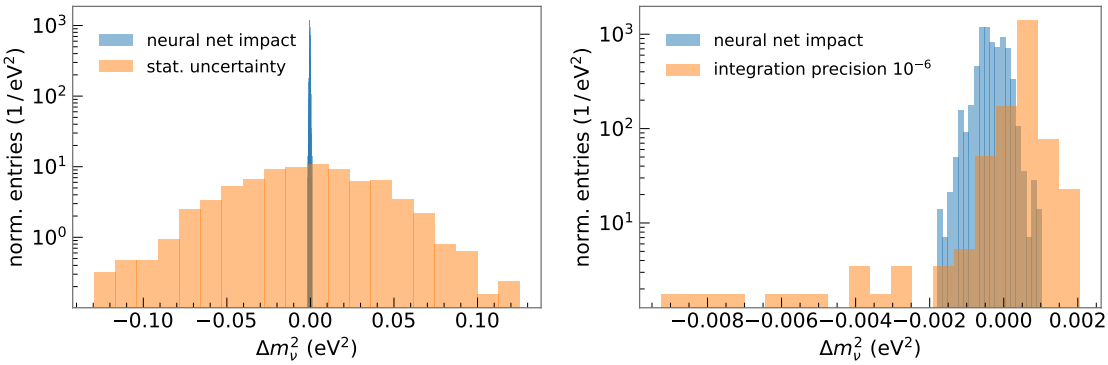
(a) spectral fit with residuals

(b) m_ν^2 profile likelihood

Figure 7.8: Fit of 1000 day Asimov spectrum generated with the reference code using both the reference and the neural network.
(a) displays the neural net fit (blue solid line) and the Monte Carlo data (black points) on the top. The two lower plots show the normalized residuals, once in a regular scale where no structure is visible (middle) and once with a zoom by a factor of 1000.
(b) compares the m_ν^2 profile likelihood of the neural net (blue solid) to the profile likelihood using the regular analysis framework (orange dashed). The overlay shows an excellent match, with the 1-σ uncertainty differing on the $10^{-4}\,\mathrm{eV}^2$ level.

test. Accordingly, we randomize the 1000 day Asimov spectrum with its Poisson statistics to generate 1000 randomized spectra. We then fit each of these spectra with the neural net and with the reference analysis, storing the best fit value of the neutrino mass squared $m^2_{\nu,\text{net}}$ and $m^2_{\nu,\text{ref}}$ respectively. The resulting distribution of the difference $\Delta m^2_\nu = m^2_{\nu,\text{ref}} - m^2_{\nu,\text{net}}$ is shown in fig. 7.9 as the blue histogram. We use this distribution to infer the bias (median) of $-3.2 \times 10^{-4}\,\mathrm{eV}^2$ and width (central $1\,\sigma$ interval) of $3.8 \times 10^{-4}\,\mathrm{eV}^2$. Both quantities show that the impact of using our NN is negligible compared to the statistical uncertainty of $\approx 0.04\,\mathrm{eV}^2$, as displayed in fig. 7.9a.

In addition to comparing to the statistical uncertainty, we compare to the intrinsic numerical noise of the regular analysis framework. Therefore, we also fit the 1000 randomized spectra with a reduced integration precision of 10^{-6} compared to the reference of 10^{-8}. This *reduced* integration precision is actually the value used so far, also for the analysis of the first two neutrino mass campaigns. The resulting distribution of the difference is compared to the impact of the NN based analysis in fig. 7.9b. We can see that the biases ($4.2 \times 10^{-4}\,\mathrm{eV}^2$ and $-3.2 \times 10^{-4}\,\mathrm{eV}^2$) are comparably large, but the analysis with reduced integration precision actually leads to more outliers and thus an increased width compared to the NN ($7.6 \times 10^{-4}\,\mathrm{eV}^2$ instead of $3.8 \times 10^{-4}\,\mathrm{eV}^2$). Our neural net based analysis is thus numerically at least as accurate as our regular model at current precision.



(a) compared to statistical fluctuations

(b) compared to noise of 10^{-6} integration precision

Figure 7.9: Ensemble test on 1000 statistically randomized Monte Carlo spectra. Each spectrum is fit with the regular analysis framework (10^{-8} integration precision) and the neural network. The difference of the fitted m_ν^2 values is then compared with the distribution of m_ν^2 from the statistical fluctuation (a) and to the difference of the reference analysis when using 10^{-6} integration precision (b), as used for previous analyses.

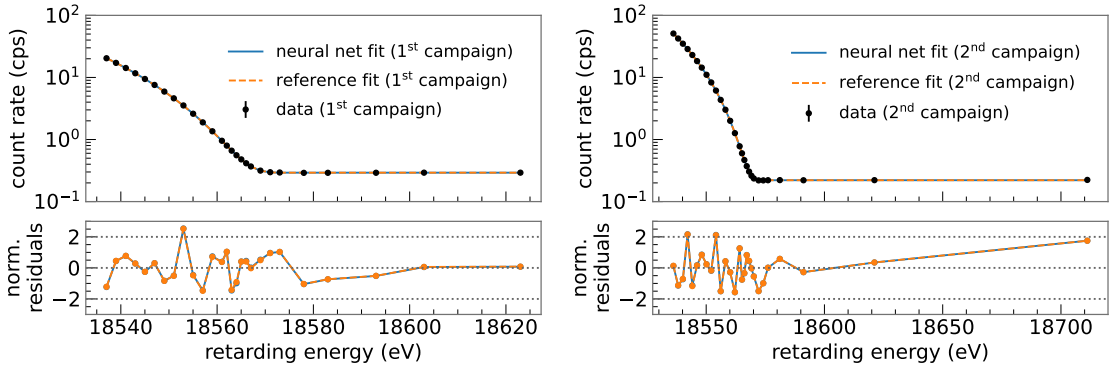
7.5 Application to realistic KATRIN scenarios

Having shown that the neural network based analysis is sufficiently accurate for final KATRIN, we next apply it to realistic KATRIN scenarios. Here, we start with a reanalysis of the first two measurement campaigns reproducing the results in [43, 66, 78]. Next, we go beyond what was published by further segmenting the data in a combined fit as well as performing a proof of concept for a full Bayesian analysis. Finally, we prove the feasibility of our approach on a Monte Carlo dataset representing final KATRIN.

7.5.1 Analysis of the first two measurement campaigns

As a first step of our NN based analysis of the first two measurement campaigns, we reproduce the published results discussed in detail in chapter 6. For the analysis, we make use of the nuisance parameter method to include parametrized systematic uncertainties. The background overdispersion is included in the likelihood, and uncertainties on molecular final states and activity fluctuations are included with the covariance matrix method.

Initially, we fit both campaigns with a uniform detector segmentation including uncertainties from Poisson statistics and the background overdispersion. Note that this analysis includes all effects neglected during the first campaign as discussed in section 6.7.1. The resulting best fit model as well as the normalized residuals are displayed in fig. 7.10 for both the neural net and the reference. We can see that the two fits show the exact same residual structure for each of the two campaigns, indicating that our NN behaves exactly like the reference model.



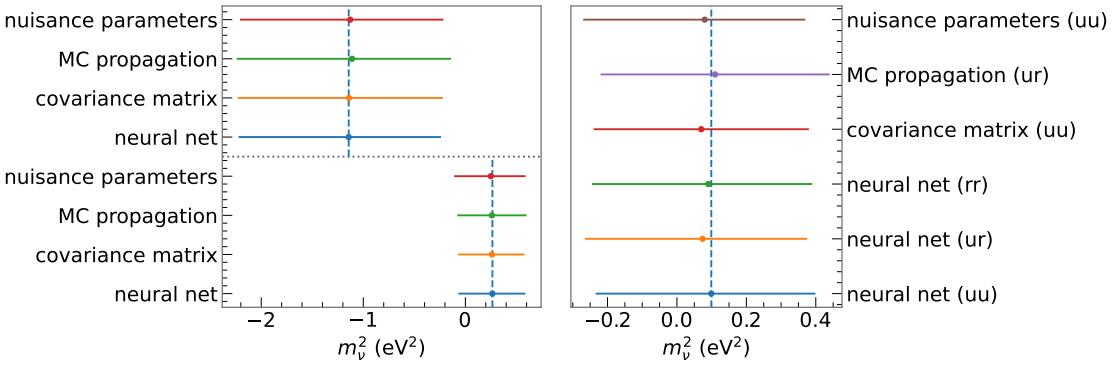
(a) first campaign **(b)** second campaign

Figure 7.10: Comparison of spectral fit for the first (a) and second (b) measurement phase using the reference framework (orange dashed) and the neural network (blue solid). The fits include statistical uncertainty from Poisson statistics and the background overdispersion. The normalized residuals shown at the bottom of each plot match perfectly for the two analysis methods, indicating that the neural network behaves exactly like the regular analysis framework.

We then proceed to fit each campaign individually, including all systematic uncertainties, KNM1 using a uniform detector and KNM2 in the radial segmentation, as performed in our reference analysis. For the first campaign, we retrieve $m_\nu^2 = -1.14^{+0.90}_{-1.08}\,\mathrm{eV}^2$ and for the second $m_\nu^2 = 0.26^{+0.32}_{-0.33}\,\mathrm{eV}^2$, in excellent agreement with the results discussed in sections 6.6 and 6.7.1. A comparison with the results published in [43] is displayed in fig. 7.11a. We can see that the NN results are fully compatible with those published, fluctuating on the same percent level as the other results with respect to one another.

In addition to the analysis with all uncertainties, we also perform an uncertainty breakdown of the second campaign with the neural network and compare it to the one presented in section 6.6. The result is shown in fig. 7.12. We can see that the overall distribution of uncertainties is well reproduced by the neural network. Small differences are expected due to the difference in treatment (e.g. the nuisance parameter approach requires subtracting in squares) and ensemble uncertainty on the results from MC propagation. One bar where the difference in treatment is especially visible is the background qU dependence: Here the best fit value changes to a value closer to zero when running the nuisance parameter approach. This leads to a larger σ_{lower} and thus a larger average uncertainty. We do not observe this in the MC propagation with systematic uncertainties only.

To go beyond our results in chapter 6, we now perform a combined fit of the first two campaigns using the neural network. As proof of concept, we use three different detector segmentations: both uniform (uu), KNM1 uniform and KNM2 ring-wise as in the individual analyses (ur) and both ring-wise (rr). Fitting either of the two campaigns ring-wise in a combined analysis had not been done before due to the computational requirements, and as it was not necessarily required. The results are displayed in fig. 7.11b, where we also compare to the published results. Here, MC propagation (ur) is our result where the m_ν^2 distributions are combined after



(a) first (top) and second (bottom) campaign

(b) combined analysis of the first two campaigns

Figure 7.11: Comparison of the fit results when analysing the first and second neutrino mass campaign individually (a) and combined (b). The results using the neural net (blue) are consistent with the ones retrieved using the various published analysis methods [43, 66, 78] on the percent level.

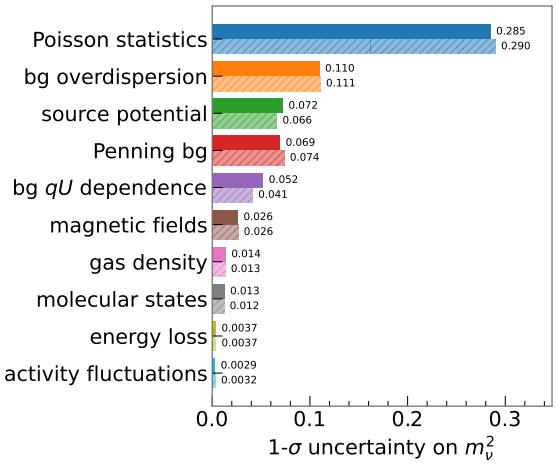


Figure 7.12: Comparison of uncertainty breakdown for the second measurement campaign using the reference framework with Monte Carlo propagation of uncertainty (hatched fill) and the neural network with nuisance parameters (regular fill).

fitting individually, and nuisance parameters (uu) as well as covariance matrix (uu) are the results published in [43]. Once again, the results using the neural network are within the usual spread of different analysis methods on the percent level.

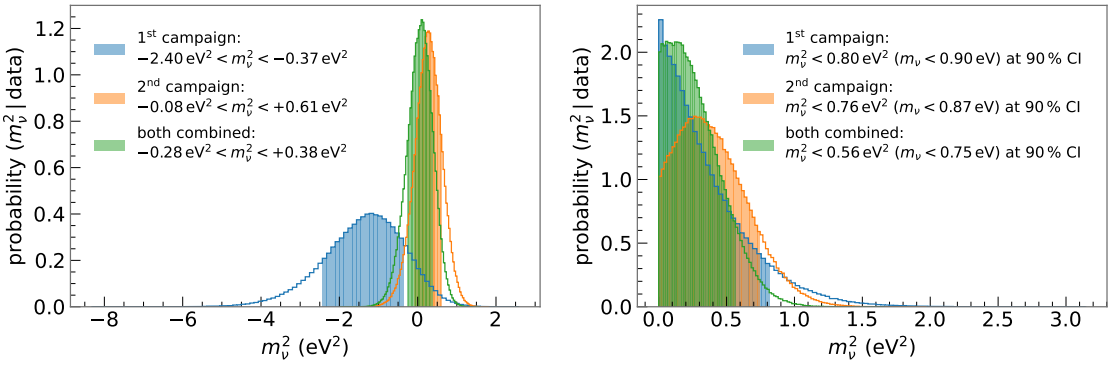
At this point we want to discuss the computing time of the different analyses performed with the neural network. All times are given as CPU time on a single core and should only serve as rough reference. Calculating the best fit and uncertainty of a single campaign using the nuisance parameter method and profile likelihood takes less than one minute (about four minutes) with a uniform (ring-wise) detector segmentation. For the combined fit, we estimate two minutes (uu), eight minutes (ur) and 17 minutes (rr) for the various segmentations. This should be compared with the 50 CPU years of our Monte Carlo propagation based analysis of KNM2. An improvement over several orders of magnitude for the final fit. The computing times are summarized in table 7.2. We can also see the re-usability of the neural network start to come to play. For each of the analyses we used the same two trained networks, one for the first and one for the second campaign.

To conclude our analysis of the first two campaigns, we provide a proof of concept for a full Bayesian analysis using the NN. To this end, we include all parameters that were included via nuisance parameters in the Frequentist analysis, as sampling parameters with the corresponding priors. For the sampling, we make use of the *emcee* [110] python package which uses a Markov Chain Monte Carlo (MCMC) ensemble sampler under the hood.

We then sample the likelihood of the first and second campaign individually, as well as the combined likelihood of the two. In each case, we use a uniform detector segmentation for simplicity. The resulting posterior using a fully flat prior in m_ν^2 is shown in fig. 7.13a, while fig. 7.13b shows the posterior using a flat positive prior on the neutrino mass squared. From the posterior with m_ν^2 constrained to the physical region, we derive an upper limit on the neutrino mass by integrating the posterior to $90\,\%$ yielding

- $m_\nu^2 < 0.80\,\mathrm{eV}^2$ ($m_\nu < 0.90\,\mathrm{eV}$) for the first campaign,
- $m_\nu^2 < 0.76\,\mathrm{eV}^2$ ($m_\nu < 0.87\,\mathrm{eV}$) for the second campaign,
- $m_\nu^2 < 0.56\,\mathrm{eV}^2$ ($m_\nu < 0.75\,\mathrm{eV}$) for both combined.

These results are once again consistent with the Bayesian analysis using the model variation technique published in [43, 66]. All our results completed on a single core within a day, and consisted of $2\,\mathrm{million}$, $2.4\,\mathrm{million}$ and $4.2\,\mathrm{million}$ samples in the posterior for the first, second and combined analysis.



(a) flat prior in m_ν^2

(b) flat positive prior in m_ν^2

Figure 7.13: Posterior in m_ν^2 using a Bayesian MCMC for the first campaign (blue), second campaign (orange) and both combined (green) once with a flat prior in m_ν^2 (a) as well as limiting the parameter space to positive m_ν^2 values (b).

7.5.2 Feasibility study for final KATRIN dataset

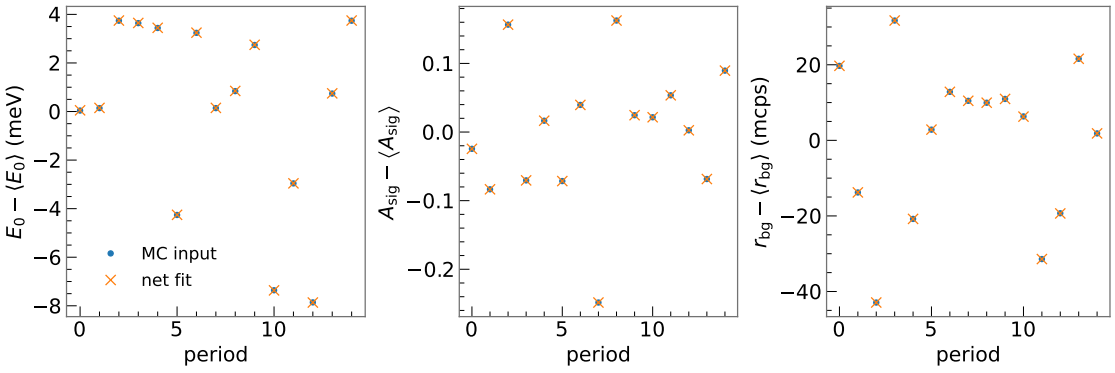
Finally, we prove the feasibility of the analysis of final KATRIN results using our neural network. To do so, we generate an Asimov Monte Carlo dataset segmented similarly to what is expected from the data taking in the coming years: 15 periods, each in the shifted analyzing plane setting requiring 14 detector patches.

For each of the 15 periods, we assume a measuring time of $60\,\mathrm{d}$, a true neutrino mass of zero and a column density of $3.75\,\mathrm{m}^{-2}$. To have some model variation over the periods, we slightly vary the endpoint, normalization and background as displayed in fig. 7.14 with mean values of $18\,573.7\,\mathrm{eV}$, 1.17 and $136\,\mathrm{mcps}$ respectively.

We then fit these 15 periods in a combined fit including all systematic parameters using the nuisance parameter method. In the fit we then have:

- 1 neutrino mass squared shared over all patches and periods,
- the Penning background, source magnetic field, column density, eloss shift and overall plasma broadening shared over all patches, but distinct for each period,
- and an endpoint, signal normalization, background rate, background retarding energy slope, maximum magnetic field and magnetic field in the analyzing plane individual for each patch and period.

This then amounts to a total of $1 + 15 \cdot 5 + 15 \cdot 14 \cdot 6 = 1336$ parameters in the minimization, and should easily be sufficient for a final KATRIN analysis.



(a) endpoint **(b)** normalization **(c)** background rate

Figure 7.14: Input parameter values (blue dot) and neural network fit result (orange cross) for each of the 15 periods in our Asimov Monte Carlo dataset emulating the data segmentation of final KATRIN. All parameters are recovered perfectly by our neural network.

Table 7.2: Computing times for different analyses using the neural network. This includes the best fit as well as an asymmetric error scan for the $1\,\sigma$ errors in m_ν^2 including systematic uncertainties with the nuisance parameter approach. All computations are run on a single core and should only serve as rough estimates to prove the feasibility of our approach.

analysis	computing time (CPU)
first campaign (uniform)	$< 1\,\mathrm{min}$
second campaign (12-ring)	$4\,\mathrm{min}$
both combined (uniform - uniform)	$2\,\mathrm{min}$
both combined (uniform - 12-ring)	$8\,\mathrm{min}$
both combined (12-ring - 12-ring)	$17\,\mathrm{min}$
KATRIN final MC (15 x 14-patches)	$1\,\mathrm{d}$

The combined fit completes in roughly one day including the interval estimation of the neutrino mass squared which gives $m_\nu^2 = 0.000^{+0.047}_{-0.049}\,\mathrm{eV}^2$, perfectly recovering the input value of zero and showing reasonable uncertainties. This perfect parameter recovery is also the case for all other parameters, as shown in fig. 7.14 for E_0, A_S and r_bg, where we average over the patches for E_0 and A_S and sum over patches for r_bg.

For this type of analysis, the preparation of all used neural networks, most likely one per period, would take several weeks at most. This can already be done for each campaign individually, when it is first analyzed, and then re-used for all subsequent combined analyses.

Together with the excellent parameter recovery and duration of less than one day on a single CPU, we conclude that a final combined fit with several hundreds of free parameters is both computationally and numerically feasible with our approach. Thus, this neural network based method fulfills all requirements needed to analyze the final KATRIN dataset in a simultaneous

fit.

Chapter 8

Conclusion

In this thesis we presented a novel uncertainty treatment, based upon full Monte Carlo propagation of uncertainty, for KATRIN analyses. We have shown that it gives consistent results with the well-established nuisance parameter method in all cases relevant for KATRIN. Additionally, it allows pre-calculating the expensive experimental response function, and is embarrassingly parallel. This makes it a powerful method, especially when combined with a computing cluster. One unique advantage of the Monte Carlo propagation method is that it allows calculating the impact of systematic uncertainties without including the statistical uncertainty, thus removing the need to subtract the total and the statistical uncertainty in squares. It also allows propagating unparameterized uncertainties and does not rely on a Gaussian likelihood function.

We next apply our developed analysis method and framework to the neutrino mass data of the first two campaigns, KATRIN neutrino mass (KNM) 1 and KNM2. Including all relevant systematic uncertainties, we fit $m_\nu^2 = -0.98^{+0.95}_{-1.12}\,\mathrm{eV}^2$ (KNM1) and $m_\nu^2 = 0.26^{+0.34}_{-0.34}\,\mathrm{eV}^2$ (KNM2). The first result is a roughly $1\,\sigma$ under-fluctuation, the second a $1\,\sigma$ over-fluctuation, when assuming a neutrino mass of zero. Therefore, we do not observe any evidence for a non-zero neutrino mass, and derive an upper limit using the method of Lokhov and Tkachov. This yields $m_\nu < 1.1\,\mathrm{eV}$ ($90\,\%$ CL) for the first campaign, and $m_\nu < 0.9\,\mathrm{eV}$ ($90\,\%$ CL) for the second one. Combining the two results by multiplying the distributions retrieved via Monte Carlo propagation further improves the limit slightly to $m_\nu < 0.8\,\mathrm{eV}$ ($90\,\%$ CL). Thus, we achieve the first sub-electronvolt direct neutrino mass measurement using data of the KATRIN experiment. The results of this thesis were published in PRL [78], PRD [66] and Nature Physics [43] representing one of the three official KATRIN analyses.

Investigating the uncertainty breakdown of each campaign shows that both are strongly dominated by statistical uncertainties, mainly the intrinsic Poisson statistics, followed by the background over-dispersion effect. In addition, we show that the fitted m_ν^2 value does not depend on the analysis interval or detector segmentation of the final fit for either campaign. Effects initially neglected in the analysis of KNM1 only have a minor impact on the best fit, and

do not change the limit derived using the method of Lokhov and Tkachov.

While a powerful tool for analysing individual campaigns, we show that our current approach using Monte Carlo propagation of uncertainty and the regular numerical model quickly runs into computational limits when performing a combined fit of multiple campaigns. To mitigate this issue, we developed a novel model approximation based upon a neural network (NN). This NN is trained to learn the spectral shape in dependence of our model input parameters such as the neutrino mass squared, the endpoint, or the magnetic fields. We prove that the NN model is able to learn this dependency and does not show any indication of over-fitting. When applying the neural network based analysis to a 1000-day Monte Carlo dataset, we observe a bias smaller than $1 \times 10^{-4}\,\mathrm{eV}^2$ in an Asimov cross-fit and of $-3.2 \times 10^{-4}\,\mathrm{eV}^2$ in an ensemble test. Both are significantly smaller than the expected statistical sensitivity of KATRIN on the order of $0.02\,\mathrm{eV}^2$ to $0.04\,\mathrm{eV}^2$, proving that the NN is sufficiently precise to be applied for KATRIN data analysis.

Next, we used our NN to the re-analyse the data of the first two neutrino mass campaigns and find results consistent with the conventional analysis frameworks. As proof of concept, we showed a full Bayesian analysis as well as a combined fit of both periods, each segmented into twelve detector rings. Finally, we performed a feasibility study on a dataset representing the data segmentation expected for final KATRIN analysis: 15 independent campaigns, each split into 14 detector patches due to the shifted analyzing plane configuration. Our NN was able to fit this dataset in one large combined fit involving 1336 parameters in the minimization. The best-fit together with an asymmetric error scan converges within a single day on a single core. Combined with the preparation of the neural networks, taking at most several weeks, and performable during the analysis of each campaign beforehand, we conclude that this approach is viable to analyse the final KATRIN dataset, and recommend further usage for the upcoming neutrino mass analyses. The results of this work were published in EPJ-C [105].

Bibliography

[1] Wolfgang Pauli. “Pauli letter collection: letter to Lise Meitner”. Typed copy. URL: https://cds.cern.ch/record/83282.

[2] C. L. Cowan et al. “Detection of the Free Neutrino: a Confirmation”. In: *Science* 124.3212 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103. eprint: https://www.science.org/doi/pdf/10.1126/science.124.3212.103. URL: https://www.science.org/doi/abs/10.1126/science.124.3212.103.

[3] G. Danby et al. “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos”. In: *Phys. Rev. Lett.* 9 (1 July 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[4] M. L. Perl et al. “Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation”. In: *Phys. Rev. Lett.* 35 (22 Dec. 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[5] K. Kodama et al. “Observation of tau neutrino interactions”. In: *Physics Letters B* 504.3 (2001), pp. 218–224. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(01)00307-0. URL: https://www.sciencedirect.com/science/article/pii/S0370269301003070.

[6] Matthew Robinson. *Symmetry and the Standard Model*. Springer, New York, NY, 1993. ISBN: 978-1-4419-8267-4. DOI: https://doi.org/10.1007/978-1-4419-8267-4.

[7] MissMJ and Cush. *Standard Model of Elementary Particles*. Licensed under the Creative Commons Attribution 3.0 Unported license. Title removed as information is conveyed in caption. 2019. URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 04/28/2022).

[8] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. “Search for Neutrinos from the Sun”. In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[9] Bruce T. Cleveland et al. “Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector”. In: *The Astrophysical Journal* 496.1 (Mar. 1998), pp. 505–526. DOI: 10.1086/305343. URL: https://doi.org/10.1086/305343.

[10] B. Pontecorvo. “Mesonium and Antimesonium”. In: *Soviet Journal of Experimental and Theoretical Physics* 6 (Jan. 1958), p. 429.

[11] J W F Valle. “Neutrino physics overview”. In: *Journal of Physics: Conference Series* 53 (Nov. 2006), pp. 473–505. DOI: 10.1088/1742-6596/53/1/031. URL: https://doi.org/10.1088/1742-6596/53/1/031.

[12] Y. Fukuda et al. “Evidence for Oscillation of Atmospheric Neutrinos”. In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[13] Q. R. Ahmad et al. “Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory”. In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[14] P. A. Cherenkov. “Visible emission of clean liquids by action of γ radiation”. In: *Doklady Akademii Nauk SSSR* (1934).

[15] L. Wolfenstein. “Neutrino oscillations in matter”. In: *Phys. Rev. D* 17 (9 May 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.2369.

[16] S. P. Mikheyev and A. Yu. Smirnov. “Resonance enhancement of oscillations in matter and solar neutrino spectroscopy”. In: *Yadernaya Fizika* 42 (Jan. 1985), pp. 1441–1448.

[17] Julien Lesgourgues and Sergio Pastor. “Neutrino Mass from Cosmology”. In: *Advances in High Energy Physics* 2012 (2012), pp. 1–34. DOI: 10.1155/2012/608515. URL: https://doi.org/10.1155%2F2012%2F608515.

[18] Planck Collaboration et al. *Planck 2018 results. VI. Cosmological parameters*. 2021. arXiv: 1807.06209 [astro-ph.CO].

[19] Eleonora Di Valentino, Alessandro Melchiorri, and Joseph Silk. “Cosmological constraints in extended parameter space from the Planck 2018 Legacy release”. In: *Journal of Cosmology and Astroparticle Physics* 2020.01 (Jan. 2020), pp. 013–013. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/01/013. URL: http://dx.doi.org/10.1088/1475-7516/2020/01/013.

[20] M. Agostini et al. “Final Results of GERDA on the Search for Neutrinoless Double-β Decay”. In: *Phys. Rev. Lett.* 125.25 (2020), p. 252502. DOI: 10.1103/PhysRevLett.125.252502. arXiv: 2009.06079 [nucl-ex].

[21] N. Abgrall et al. “The Large Enriched Germanium Experiment for Neutrinoless Double Beta Decay (LEGEND)”. In: *AIP Conf. Proc.* 1894.1 (2017). Ed. by Osvaldo Civitarese, Ivan Stekl, and Jouni Suhonen, p. 020027. DOI: 10.1063/1.5007652. arXiv: 1709.01980 [physics.ins-det].

[22] S. I. Alvis et al. “A Search for Neutrinoless Double-Beta Decay in ^{76}Ge with 26 kg-yr of Exposure from the MAJORANA DEMONSTRATOR”. In: *Phys. Rev. C* 100.2 (2019), p. 025501. DOI: 10.1103/PhysRevC.100.025501. arXiv: 1902.02299 [nucl-ex].

[23] D. Q. Adams et al. “Improved Limit on Neutrinoless Double-Beta Decay in ^{130}Te with CUORE”. In: *Phys. Rev. Lett.* 124.12 (2020), p. 122501. DOI: 10.1103/PhysRevLett.124.122501. arXiv: 1912.10966 [nucl-ex].

[24] S. Abe et al. “First Search for the Majorana Nature of Neutrinos in the Inverted Mass Ordering Region with KamLAND-Zen”. In: (Mar. 2022). arXiv: 2203.02139 [hep-ex].

[25] H. Ozaki and A. Takeuchi. “Upgrade of the KamLAND-Zen mini-balloon and future prospects”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 958 (2020). Proceedings of the Vienna Conference on Instrumentation 2019, p. 162353. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2019.162353. URL: https://www.sciencedirect.com/science/article/pii/S016890021930943X.

[26] E W Otten and C Weinheimer. “Neutrino mass limit from tritium β decay”. In: *Reports on Progress in Physics* 71.8 (2008), p. 086201. URL: http://stacks.iop.org/0034-4885/71/i=8/a=086201.

[27] M. Kleesiek et al. *β-Decay Spectrum, Response Function and Statistical Model for Neutrino Mass Measurements with the KATRIN Experiment*. 2018. eprint: arXiv:1806.00369.

[28] Project 8 Collaboration et al. *The Project 8 Neutrino Mass Experiment*. 2022. DOI: 10.48550/ARXIV.2203.07349. URL: https://arxiv.org/abs/2203.07349.

[29] K. Blaum et al. “The electron capture 163Ho experiment ECHo”. In: *Journal of Low Temperature Physics* 176 (May 2014), pp. 876–884. DOI: 10.1007/s10909-014-1187-4.

[30] *NuFit 5.1 (2021)*. URL: http://www.nu-fit.org/.

[31] Ivan Esteban et al. “The fate of hints: updated global analysis of three-flavor neutrino oscillations”. In: *Journal of High Energy Physics* 2020.9 (Sept. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep09(2020)178. URL: http://dx.doi.org/10.1007/JHEP09(2020)178.

[32] Ch Kraus et al. “Final results from phase II of the Mainz neutrino mass search in tritium β decay”. In: *The European Physical Journal C - Particles and Fields* 40.4 (2005), pp. 447–468. ISSN: 1434-6052. DOI: 10.1140/epjc/s2005-02139-7. URL: https://doi.org/10.1140/epjc/s2005-02139-7.

[33] V. N. Aseev et al. “Upper limit on the electron antineutrino mass from the Troitsk experiment”. In: *Phys. Rev. D* 84 (11 2011), p. 112003. DOI: 10.1103/PhysRevD.84.112003. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.112003.

[34] KATRIN Collaboration. *KATRIN design report 2004*. Tech. rep. 51.54.01; LK 01; Auch: NPI ASCR Rez EXP-01/2005; MS-KP-0501. Forschungszentrum, Karlsruhe, 2005. 245 pp.

[35] M. Aker et al. “The design, construction, and commissioning of the KATRIN experiment”. In: *Journal of Instrumentation* 16.08 (Aug. 2021), T08015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/08/t08015. URL: http://dx.doi.org/10.1088/1748-0221/16/08/T08015.

[36] Michael Sturm et al. “Kilogram scale throughput performance of the KATRIN tritium handling system”. In: *Fusion Engineering and Design* 170 (2021), p. 112507. ISSN: 0920-3796. DOI: https://doi.org/10.1016/j.fusengdes.2021.112507. URL: https://www.sciencedirect.com/science/article/pii/S0920379621002830.

[37] Florian Heizmann, Hendrik Seitz-Moskaliuk, and KATRIN collaboration. “The Windowless Gaseous Tritium Source (WGTS) of the KATRIN experiment”. In: *Journal of Physics: Conference Series* 888.1 (2017), p. 012071. URL: http://stacks.iop.org/1742-6596/888/i=1/a=012071.

[38] S. Lukić et al. “Measurement of the gas-flow reduction factor of the KATRIN DPS2-F differential pumping section”. In: *Vacuum* 86.8 (2012), pp. 1126–1133. ISSN: 0042-207X. DOI: https://doi.org/10.1016/j.vacuum.2011.10.017. URL: http://www.sciencedirect.com/science/article/pii/S0042207X11003800.

[39] W. Gil et al. “The Cryogenic Pumping Section of the KATRIN Experiment”. In: *IEEE Transactions on Applied Superconductivity* 20.3 (June 2010), pp. 316–319. ISSN: 1051-8223. DOI: 10.1109/TASC.2009.2038581.

[40] F Fränkle et al. “Penning discharge in the KATRIN pre-spectrometer”. In: *Journal of Instrumentation* 9 (July 2014), P07028–P07028. DOI: 10.1088/1748-0221/9/07/P07028.

[41] M. Arenz et al. “Calibration of high voltages at the ppm level by the difference of $^{83\mathrm{m}}$Kr conversion electron lines at the KATRIN experiment”. In: *Eur. Phys. J. C* 78.5 (2018), p. 368. DOI: https://doi.org/10.1140/epjc/s10052-018-5832-y. arXiv: 1802.05227 [physics.ins-det].

[42] Th. Thümmler, R. Marx, and Ch. Weinheimer. “Precision high voltage divider for the KATRIN experiment”. In: *New J. Phys.* 11 (2009), p. 103007. DOI: https://doi.org/10.1088/1367-2630/11/10/103007. arXiv: 0908.1523 [physics.ins-det].

[43] M. Aker et al. “Direct neutrino-mass measurement with sub-electronvolt sensitivity”. In: *Nature Physics* (Feb. 2022). DOI: 10.1038/s41567-021-01463-1. URL: https://www.nature.com/articles/s41567-021-01463-1.

[44] Ferenc Glück et al. “Electromagnetic design of the large-volume air coil system of the KATRIN experiment”. In: *New Journal of Physics* 15.8 (Aug. 2013), p. 083025. DOI: 10.1088/1367-2630/15/8/083025. URL: https://doi.org/10.1088/1367-2630/15/8/083025.

[45] M. Erhard et al. “Technical design and commissioning of the KATRIN large-volume air coil system”. In: *Journal of Instrumentation* 13 (Dec. 2017). DOI: 10.1088/1748-0221/13/02/P02003.

[46] J.F. Amsbaugh et al. “Focal-plane detector system for the KATRIN experiment”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 778 (2015), pp. 40–60. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2014.12.116. URL: http://www.sciencedirect.com/science/article/pii/S0168900215000236.

[47] Martin Babutzka. “Design and development for the Rearsection of the KATRIN experiment”. PhD thesis. 2014.

[48] Behrens, J. et al. “A pulsed, mono-energetic and angular-selective UV photo-electron source for the commissioning of the KATRIN experiment”. In: *Eur. Phys. J. C* 77.6 (2017), p. 410. DOI: 10.1140/epjc/s10052-017-4972-9. URL: https://doi.org/10.1140/epjc/s10052-017-4972-9.

[49] M. Aker et al. “Precision measurement of the electron energy-loss function in tritium and deuterium gas for the KATRIN experiment”. In: *Eur. Phys. J. C* 81.7 (July 2021). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09325-z. URL: http://dx.doi.org/10.1140/epjc/s10052-021-09325-z.

[50] Marco Röllig. “Tritium analytics by beta induced X-ray spectrometry”. 51.03.01; LK 01. PhD thesis. 2015. DOI: 10.5445/IR/1000054050.

[51] A. Beglarian et al. *Forward Beam Monitor for the KATRIN experiment*. 2021. arXiv: 2101.11495 [physics.ins-det].

[52] Susanne Mertens et al. “A novel detector system for KATRIN to search for keV-scale sterile neutrinos”. In: *Journal of Physics G: Nuclear and Particle Physics* 46.6 (May 2019), p. 065203. DOI: 10.1088/1361-6471/ab12fe. URL: https://doi.org/10.1088/1361-6471/ab12fe.

[53] Korbinian Urban. *Application of a TRISTAN Silicon Drift Detector as Forward Beam Monitor in KATRIN*. 2019.

[54] Alejandro Saenz, Svante Jonsell, and Piotr Froelich. “Improved Molecular Final-State Distribution of HeT^+ for the β-Decay Process of T_2”. In: *Phys. Rev. Lett.* 84 (2 2000), pp. 242–245. DOI: 10.1103/PhysRevLett.84.242. URL: https://link.aps.org/doi/10.1103/PhysRevLett.84.242.

[55] Martin Slezák. “Monitoring of the energy scale in the KATRIN neutrino experiment”. PhD thesis. Charles University in Prague, 2015. URL: http://www.katrin.kit.edu/publikationen/phd-Martin_Slezak.pdf.

[56] Stefan Groh. “Modeling of the response function and measurement of transmission properties of the KATRIN experiment”. PhD thesis. 2015.

[57] Ferenc Glück. *Theoretical total inelastic e-H2 cross section*. KATRIN internal. June 11, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/a/ab/SigmaineleH2Theor.pdf (visited on 01/21/2022).

[58] Ferenc Glück. *Theoretical total inelastic cross section: isotopic dependence and accuracy*. KATRIN internal. July 2, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/3/36/SigmainelTheorH2D2T2.pdf (visited on 01/21/2022).

[59] J. W. Liu. “Total cross sections for high-energy electron scattering by H_2 ($^1\Sigma_g^+$), N_2 ($^1\Sigma_g^+$), and O_2 ($^3\Sigma_g^-$)”. In: *Phys. Rev. A* 35 (2 Jan. 1987), pp. 591–597. DOI: 10.1103/PhysRevA.35.591. URL: https://link.aps.org/doi/10.1103/PhysRevA.35.591.

[60] P. Weck, B. Joulakian, and P. A. Hervieux. “Fivefold differential cross section of fast $(e, 2e)$ ionization of H_2, D_2, and T_2 by a Franck-Condon approach”. In: *Phys. Rev. A* 60 (4 Oct. 1999), pp. 3013–3019. DOI: 10.1103/PhysRevA.60.3013. URL: https://link.aps.org/doi/10.1103/PhysRevA.60.3013.

[61] Yong-Ki Kim, José Paulo Santos, and Fernando Parente. “Extension of the binary-encounter-dipole model to relativistic incident electrons”. In: *Phys. Rev. A* 62 (5 Oct. 2000), p. 052710. DOI: 10.1103/PhysRevA.62.052710. URL: https://link.aps.org/doi/10.1103/PhysRevA.62.052710.

[62] Marco Kleesiek. “A Data-Analysis and Sensitivity-Optimization Framework for the KATRIN Experiment”. PhD thesis. 2014.

[63] Matthias Weidenthaler. *Systematic Uncertainties of the KATRIN Neutrino Mass Measurement Associated with Beta Decays on the Rear Wall of the Experiment*. 2022.

[64] Christoph Köhler. *Determination of the Column Density in the KATRIN Beamline with Electrons from the Photo-Electric Source*. 2019.

[65] O Fackler et al. “Accurate theoretical beta-decay energy spectrum of the tritium molecule and its neutrino mass dependence”. In: *Physical review letters* 55.13 (Sept. 1985), pp. 1388–1391. ISSN: 0031-9007. DOI: 10.1103/physrevlett.55.1388. URL: https://doi.org/10.1103/PhysRevLett.55.1388.

[66] M. Aker et al. “Analysis methods for the first KATRIN neutrino-mass measurement”. In: *Phys. Rev. D* 104 (1 July 2021), p. 012005. DOI: 10.1103/PhysRevD.104.012005. URL: https://link.aps.org/doi/10.1103/PhysRevD.104.012005.

[67] Martin Slezák. *Final state distribution (FSD) onset*. 2018.

[68] Alejandro Saenz, Christian Weinheimer, and Alexey Lokhov. *Status of the molecular FSD and its uncertainties*. KATRIN internal. Oct. 7, 2021. URL: https://indico.scc.kit.edu/event/2476/contributions/9115/attachments/4492/6763/katrin_cm41_p.pdf (visited on 01/25/2022).

[69] Martin Descher, Enrico Ellinger, and Karol Debowski. *ELOG - FBM Troubleshooting*. KATRIN internal elog. Nov. 7, 2019. URL: https://ikp-neu-katrin.ikp.kit.edu/elog/KNM2/103 (visited on 01/25/2022).

[70] Moritz Machatschek. “A Phenomenological Theory of KATRIN Source Potential Systematics and its Application in Krypton-83m Calibration Measurements”. PhD thesis. 2020.

[71] Anna Schaller. “Characterization and mitigation of the background in KATRIN”. PhD thesis. 2020.

[72] Indrek Martinson and L.J. Curtis. *Janne Rydberg - his life and work*. eng. 2005. DOI: 10.1016/j.nimb.2005.03.137. URL: https://lup.lub.lu.se/search/files/4641948/3409138.pdf.

[73] S. Mertens et al. “Background due to stored electrons following nuclear decays in the KATRIN spectrometers and its impact on the neutrino mass sensitivity”. In: *Astroparticle Physics* 41 (2013), pp. 52–62. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2012.10.005. URL: https://www.sciencedirect.com/science/article/pii/S0927650512001892.

[74] Fabian Block. *Characterization of the Background in the KATRIN Experiment*. 2018.

[75] Florian Fränkle. *Background retarding potential dependence in SAP configuration*. KATRIN internal. Apr. 15, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/9/9a/SAPqUSlopeAnalysisReport.pdf (visited on 01/26/2022).

[76] Sanshiro Enomoto. *Detector Count Corrections for First Tritium*. KATRIN internal. Feb. 3, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/d/d2/FirstTritiumDetectorSystematics_20190203.pdf (visited on 01/27/2022).

[77] Christoph Wiesinger. *Detector systematics in neutrino mass analysis*. KATRIN internal. July 13, 2021. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/6/6e/20210713_DetectorSystematics_AnalysisCall.pdf (visited on 01/27/2022).

[78] M. Aker et al. “Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN”. In: *Phys. Rev. Lett.* 123 (22 Nov. 2019), p. 221802. DOI: 10.1103/PhysRevLett.123.221802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.221802.

[79] Dominik Hinz. *Background measurements after bake-out and baffle cooling*. KATRIN internal. Sept. 10, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/3/30/BackgroundMeasurements20190910.pdf (visited on 01/28/2022).

[80] Michael Sturm. *Tritium Loops*. KATRIN internal. Nov. 4, 2019. URL: https://fuzzy.fzk.de/bscw/bscw.cgi/d1305735/95-TRP-3702-P1-MSturm.pdf (visited on 01/28/2022).

[81] Loops Operator. *Return of tritiated gas after end of tritium exposure of Loop system*. KATRIN internal. Aug. 26, 2019. URL: https://ikp-neu-katrin.ikp.kit.edu/elog/Loops-Main/161 (visited on 01/28/2022).

[82] Alexey Lokhov et al. “Background reduction at the KATRIN experiment by the shifted analyzing plane configuration”. In: *preparation* (2022).

[83] M. Erhard et al. “High-voltage monitoring with a solenoid retarding spectrometer at the KATRIN experiment”. In: *JINST* 2014.9 (2014), P06022. DOI: https://doi.org/10.1088/1748-0221/9/06/P06022.

[84] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360. URL: https://doi.org/10.1214/aoms/1177732360.

[85] Lisa Schlüter. *Development of New Methods to Include Systematic Effects in the First Tritium Data Analysis and Sensitivity Studies of the KATRIN Experiment*. 2019.

[86] Jerzy Neyman. “Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability”. In: *Philosophical Transactions of the Royal Society A* 236 (1937), pp. 333–380.

[87] Gary J. Feldman and Robert D. Cousins. “A Unified approach to the classical statistical analysis of small signals”. In: *Phys. Rev. D* 57 (1998), pp. 3873–3889. DOI: https://doi.org/10.1103/PhysRevD.57.3873. arXiv: physics/9711021 [physics.data-an].

[88] Till Moritz Karbach. *Feldman-Cousins Confidence Levels - Toy MC Method*. 2011. arXiv: 1109.0714 [physics.data-an].

[89] A. V. Lokhov and F. V. Tkachov. “Confidence intervals with a priori parameter bounds”. In: *Phys. Part. Nucl.* 46.3 (2015), pp. 347–365. DOI: https://doi.org/10.1134/S1063779615030089. arXiv: 1403.5429 [physics.data-an].

[90] Martin Ha Minh. *Bayesian Analysis of the First Data of the KATRIN Experiment*. 2018.

[91] Christian Karl, Susanne Mertens, and Martin Slezák. “Analysis of First KATRIN Neutrino-Mass Data”. In: *8th International Pontecorvo Neutrino Physics School*. JINR Publishing Department, 2020. ISBN: 978-5-9530-0538. URL: http://theor.jinr.ru/~neutrino19/files/neutrino2019_proceedings.pdf.

[92] Christian Karl, Susanne Mertens, and Martin Slezák. “Treatment of systematic effects in the KATRIN experiment by Monte Carlo propagation of uncertainty”. In: *progress* (tbd).

[93] Bradley Efron and R.J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 1993.

[94] Florian Fränkle and Anna Schaller. *Pixel Selection for KNM1*. KATRIN internal. May 28, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/e/eb/AnaCallPixelCut190528.pdf (visited on 02/27/2022).

[95] Florian Fränkle. *Pixel Selection for KNM2*. KATRIN internal. Jan. 10, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/3/3b/AnaCallPixelCut190924.pdf (visited on 02/27/2022).

[96] Jan Behrens. *Flux tube calculation in SSC*. KATRIN internal. Feb. 15, 2022. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/6/6d/AnaCall_20220214_FluxCalculationSSC.pdf (visited on 04/11/2022).

[97] Fabian Block. “Determination of electromagnetic fields and tritium column density for neutrino mass analysis with KATRIN”. PhD thesis. 2021.

[98] Vikas Gupta et al. *Global broadening determination with E_0-fit method in KNM2*. KATRIN internal. May 28, 2020. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/9/93/E0_broadening_reference.pdf (visited on 02/22/2022).

[99] Jan Behrens and Fabian Block. *KATRIN First Tritium Reference Report Systematic Uncertainty of Magnetic Field in the Analyzing Plane*. KATRIN internal. Apr. 3, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/9/92/FT_syst_uncertainty_Bana_030419.pdf (visited on 02/22/2022).

[100] Ferenc Glück. *Uncertainty of the WGTS and PCH magnetic fields*. KATRIN internal. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/e/e4/WGTSPCHReport.pdf (visited on 02/22/2022).

[101] Jan Behrens et al. *KNM2 Systematic Uncertainties Uncertainties of magnetic field in analyzing plane, pinch magnet and WGTS*. KATRIN internal. May 25, 2021. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/3/3d/KNM2_magFields_sysUncFinal.pdf (visited on 02/22/2022).

[102] O Rest et al. "A novel ppm-precise absolute calibration method for precision high-voltage dividers". In: *Metrologia* 56.4 (July 2019), p. 045007. DOI: 10.1088/1681-7575/ab2997. URL: https://doi.org/10.1088/1681-7575/ab2997.

[103] Moritz Machatschek et al. *Plasma investigations during KNM1*. KATRIN internal. Sept. 10, 2019. URL: https://ikp-katrin-wiki.ikp.kit.edu/katrin/images/3/39/GKrS_KNM1_report_v1.0.pdf (visited on 02/24/2022).

[104] Christian Karl. *Analysis of First Tritium Data of the KATRIN Experiment*. 2018.

[105] Christian Karl, Philipp Eller, and Susanne Mertens. "Fast and precise model calculation for KATRIN using a neural network". In: *Eur. Phys. J. C* 82.5 (2022), p. 439. DOI: 10.1140/epjc/s10052-022-10384-z. arXiv: 2201.04523 [hep-ex].

[106] Martin Roberts. *The Unreasonable Effectiveness of Quasirandom Sequences*. 2018. URL: http://extremelearning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/ (visited on 11/05/2021).

[107] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[108] Jorge Nocedal. "Updating Quasi-Newton Matrices with Limited Storage". In: *Mathematics of Computation* 35.151 (1980), pp. 773–782. ISSN: 00255718, 10886842. URL: http://www.jstor.org/stable/2006193.

[109] Dong C. Liu and Jorge Nocedal. "On the limited memory BFGS method for large scale optimization". English (US). In: *Mathematical Programming* 45.3 (Dec. 1989), pp. 503–528. ISSN: 0025-5610.

[110] Daniel Foreman-Mackey et al. "emcee: The MCMC Hammer". In: *Publications of the Astronomical Society of the Pacific* 125.925 (Mar. 2013), pp. 306–312. ISSN: 1538-3873. DOI: 10.1086/670067. URL: http://dx.doi.org/10.1086/670067.

List of figures

List of tables

List of acronyms

BED binary-encounter-dipole . . . 33
BIXS β-induced X-ray spectrometry . . . 21
CDF cumulative distribution function . . . 65
CI credible interval . . . 10
CL confidence level . . . ii
CMB cosmic microwave background . . . 9
CPS cryogenic pumping section . . . 19
CRES cyclotron radiation emission spectroscopy . . . 12
DPS differential pumping section . . . 19
ECHo electron capture in holmium . . . 12
egun electron gun . . . 21
EMCS earth magnetic-field correction system . . . 20
FBM forward beam monitor . . . 21
FPD focal plane detector . . . 20
FSD final state distribution . . . 22
FWHM full-width half-maximum . . . 20
KATRIN **Ka**rlsruhe **Tri**tium **N**eutrino . . . ii
KNM KATRIN neutrino mass . . . 54
LFCS low-field correction system . . . 20
MCMC Markov Chain Monte Carlo . . . 130
MSW Mikheyev–Smirnov–Wolfenstein . . . 8
MTD measuring time distribution . . . 36
NN neural network . . . ii
PDF probability density function . . . 59
PMF probability mass function . . . 58
PMNS Pontecorvo-Maki-Nakagawa-Sakata . . . 6
PMT photo multiplier tube . . . 4
ROI region of interest . . . 52
SAP shifted analyzing plane . . . 55

Appendices

Appendix A

Summary tables

This appendix chapter contains various summary tables of results which are too long to be included in the main text.

Table A.1: Best fit for all fit parameters of the first measurement campaign with uncertainty from Poisson statistics only. The uncertainty has been symmetrized for simplicity.

parameter	fit result
r_{bg}	$0.2923 \pm 0.0007\,\mathrm{cps}$
E_0	$18\,573.728 \pm 0.058\,\mathrm{eV}$
m_ν^2	$-0.963 \pm 0.971\,\mathrm{eV}^2$
A_{S}	1.191 ± 0.006

Table A.2: Best fit for all fit parameters of the first measurement campaign with total uncertainty. The uncertainty has been symmetrized for simplicity.

parameter	fit result
r_{bg}	$0.2923 \pm 0.0007\,\mathrm{cps}$
E_0	$18\,573.727 \pm 0.061\,\mathrm{eV}$
m_ν^2	$-0.978 \pm 1.038\,\mathrm{eV}^2$
A_{S}	1.185 ± 0.038

Table A.3: Best fit for all fit parameters of the second measurement campaign with uncertainty from Poisson statistics only. The uncertainty has been symmetrized for simplicity. The number within the round bracket indicates the ring number for ring-wise parameters.

parameter	fit result
m_ν^2	$0.274 \pm 0.290\,\mathrm{eV}^2$
$r_{\mathsf{bg}}\,(0)$	$0.0066 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(1)$	$0.0201 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(2)$	$0.0208 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(3)$	$0.0209 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(4)$	$0.0218 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(5)$	$0.0225 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(6)$	$0.0227 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(7)$	$0.0237 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(8)$	$0.0206 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(9)$	$0.0192 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(10)$	$0.0153 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(11)$	$0.0060 \pm 0.0001\,\mathrm{cps}$
$E_0\,(0)$	$18\,573.766 \pm 0.049\,\mathrm{eV}$
$E_0\,(1)$	$18\,573.769 \pm 0.033\,\mathrm{eV}$
$E_0\,(2)$	$18\,573.708 \pm 0.034\,\mathrm{eV}$
$E_0\,(3)$	$18\,573.730 \pm 0.034\,\mathrm{eV}$
$E_0\,(4)$	$18\,573.699 \pm 0.033\,\mathrm{eV}$
$E_0\,(5)$	$18\,573.653 \pm 0.034\,\mathrm{eV}$
$E_0\,(6)$	$18\,573.667 \pm 0.034\,\mathrm{eV}$
$E_0\,(7)$	$18\,573.656 \pm 0.035\,\mathrm{eV}$
$E_0\,(8)$	$18\,573.655 \pm 0.036\,\mathrm{eV}$
$E_0\,(9)$	$18\,573.665 \pm 0.038\,\mathrm{eV}$
$E_0\,(10)$	$18\,573.647 \pm 0.041\,\mathrm{eV}$
$E_0\,(11)$	$18\,573.722 \pm 0.057\,\mathrm{eV}$
$A_{\mathsf{S}}\,(0)$	1.122 ± 0.007
$A_{\mathsf{S}}\,(1)$	1.120 ± 0.005
$A_{\mathsf{S}}\,(2)$	1.126 ± 0.005
$A_{\mathsf{S}}\,(3)$	1.125 ± 0.005
$A_{\mathsf{S}}\,(4)$	1.131 ± 0.005
$A_{\mathsf{S}}\,(5)$	1.132 ± 0.005
$A_{\mathsf{S}}\,(6)$	1.129 ± 0.005
$A_{\mathsf{S}}\,(7)$	1.132 ± 0.005
$A_{\mathsf{S}}\,(8)$	1.131 ± 0.005
$A_{\mathsf{S}}\,(9)$	1.133 ± 0.005
$A_{\mathsf{S}}\,(10)$	1.130 ± 0.006
$A_{\mathsf{S}}\,(11)$	1.114 ± 0.009

Table A.4: Best fit for all fit parameters of the second measurement campaign with total uncertainty. The uncertainty has been symmetrized for simplicity. The number within the round bracket indicates the ring number for ring-wise parameters.

parameter	fit result
m_ν^2	$0.261 \pm 0.340\,\mathrm{eV}^2$
$r_{\mathsf{bg}}\,(0)$	$0.0066 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(1)$	$0.0201 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(2)$	$0.0208 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(3)$	$0.0209 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(4)$	$0.0218 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(5)$	$0.0225 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(6)$	$0.0227 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(7)$	$0.0236 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(8)$	$0.0206 \pm 0.0002\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(9)$	$0.0192 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(10)$	$0.0153 \pm 0.0001\,\mathrm{cps}$
$r_{\mathsf{bg}}\,(11)$	$0.0060 \pm 0.0001\,\mathrm{cps}$
$E_0\,(0)$	$18\,573.764 \pm 0.055\,\mathrm{eV}$
$E_0\,(1)$	$18\,573.765 \pm 0.039\,\mathrm{eV}$
$E_0\,(2)$	$18\,573.702 \pm 0.040\,\mathrm{eV}$
$E_0\,(3)$	$18\,573.730 \pm 0.040\,\mathrm{eV}$
$E_0\,(4)$	$18\,573.696 \pm 0.040\,\mathrm{eV}$
$E_0\,(5)$	$18\,573.654 \pm 0.040\,\mathrm{eV}$
$E_0\,(6)$	$18\,573.666 \pm 0.040\,\mathrm{eV}$
$E_0\,(7)$	$18\,573.661 \pm 0.040\,\mathrm{eV}$
$E_0\,(8)$	$18\,573.656 \pm 0.042\,\mathrm{eV}$
$E_0\,(9)$	$18\,573.664 \pm 0.043\,\mathrm{eV}$
$E_0\,(10)$	$18\,573.652 \pm 0.047\,\mathrm{eV}$
$E_0\,(11)$	$18\,573.715 \pm 0.064\,\mathrm{eV}$
$A_{\mathsf{S}}\,(0)$	1.123 ± 0.025
$A_{\mathsf{S}}\,(1)$	1.121 ± 0.024
$A_{\mathsf{S}}\,(2)$	1.128 ± 0.025
$A_{\mathsf{S}}\,(3)$	1.126 ± 0.024
$A_{\mathsf{S}}\,(4)$	1.132 ± 0.025
$A_{\mathsf{S}}\,(5)$	1.133 ± 0.025
$A_{\mathsf{S}}\,(6)$	1.130 ± 0.025
$A_{\mathsf{S}}\,(7)$	1.132 ± 0.025
$A_{\mathsf{S}}\,(8)$	1.132 ± 0.025
$A_{\mathsf{S}}\,(9)$	1.134 ± 0.025
$A_{\mathsf{S}}\,(10)$	1.131 ± 0.025
$A_{\mathsf{S}}\,(11)$	1.116 ± 0.026

Table A.5: Uncertainty breakdown for the first neutrino mass campaign.

effect	$1\,\sigma$ uncertainty on m_ν^2 (eV^2)
Poisson statistics	0.97
bg overdispersion	0.30
bg qU dependence	0.066
gas density	0.052
magnetic fields	0.049
activity fluctuations	0.044
molecular states	0.020
energy loss	0.0023

Table A.6: Uncertainty breakdown for the second neutrino mass campaign.

effect	$1\,\sigma$ uncertainty on m_ν^2 (eV^2)
Poisson statistics	0.29
bg overdispersion	0.11
Penning bg	0.074
source potential	0.066
bg qU dependence	0.041
magnetic fields	0.026
gas density	0.013
molecular states	0.012
energy loss	0.0037
activity fluctuations	0.0032

Table A.7: Uncertainty breakdown for the second neutrino mass campaign using our neural network.

effect	$1\,\sigma$ uncertainty on m_ν^2 (eV^2)
Poisson statistics	0.28
bg overdispersion	0.11
source potential	0.072
Penning bg	0.069
bg qU dependence	0.052
magnetic fields	0.026
gas density	0.014
molecular states	0.013
energy loss	0.0037
activity fluctuations	0.0029

Appendix B

Code repositories

In this appendix we summarize the most important code repositories used and developed over the course of this thesis in a bullet-point fashion.

Numerical Calculus

Main author: Martin Slezák
Gitlab link: https://nuserv.uni-muenster.de:8443/slezak/NumericalCalculus
Purpose: A collection of algorithms to perform differential and integral calculus as well as further numerical calculations.
Main language: C++

Fitter

Main author: Martin Slezák
Gitlab link: https://nuserv.uni-muenster.de:8443/slezak/Fitter
Purpose: Interface for fitting models to data by maximum likelihood estimation including uncertainty estimation. Main language: C++
Note: Builds upon *Numerical Calculus*.

Fitrium

Main author: Christian Karl
Gitlab link: https://nuserv.uni-muenster.de:8443/karlch/Fitrium
Purpose: Modelling and fitting tools for the KATRIN experiment. Provides the full KATRIN

model, an interface for data reading and analysis as well as various Monte Carlo tools. Main language: C++
Note: Builds upon *Numerical Calculus* and *Fitter*.

Netrium

Main author: Christian Karl
Gitlab link: https://nuserv.uni-muenster.de:8443/karlch/Netrium
Purpose: Toolkit to analyse KATRIN data using a neural network. Provides the chain to train a neural network and use it for KATRIN data analysis.
Main language: python

NetriumGensamples

Main author: Christian Karl
Gitlab link: https://nuserv.uni-muenster.de:8443/karlch/netriumgensamples
Purpose: Toolkit to generate samples for *Netrium* using *Fitrium*.
Main language: C++

Kasper

Main author: KATRIN Collaboration
Gitlab link: https://nuserv.uni-muenster.de:8443/katrin-git/kasper
Purpose: General KATRIN analysis and simulation package.
Main language: C++
Note: Only used during this thesis to retrieve and read the KATRIN data contained in *run summary* files.

RS2HDF5

Main author: Christian Karl
Gitlab link: https://nuserv.uni-muenster.de:8443/karlch/RS2HDF5
Purpose: Converter for KATRIN *run summary* files in the KATRIN-specific *ktf* format into the HDF5 format.
Main language: C++

Webtrium

Main author: Christian Karl
Gitlab link: https://nuserv.uni-muenster.de:8443/karlch/Webtrium
Purpose: Online analysis tools using *Fitrium*. Provides near-time fit results during a measurement campaign accessible via webtrium.mpp.mpg.de.
Main language: python

Acknowledgements

I would like to thank everyone who helped me during my doctoral studies, in particular:

- Prof. Dr. Susanne Mertens for giving me the chance to work on this thesis and her incredible supervision with lots of fruitful discussions and permanent support,
- Dr. Martin Slezák for his constant support concerning coding, physics and more as well as for developing the *Fitness Studio*,
- Dr. Christoph Wiesinger and Alessandro Schwemmer from our analysis team for proofreading my thesis and continuing our Monte Carlo propagation and neural network based methods,
- Dr. Philipp Eller for his help in developing the neural network based analysis,
- Dr. Frank Edzards for the great atmosphere in our office and valuable support on any upcoming questions,
- Prof. Debarghya Ghoshdastidar from the TUM IT department and his group as well as Ulf Mertens, Sebastian Neubauer, and Felix Wick from Blue Yonder for the fruitful discussions on neural networks,
- all fellow PhD, Master's and other students in the TRISTAN group at MPP for a great time and support on various topics,
- the KATRIN collaboration for the great working environment and the opportunity to work on KATRIN data,
- the free software community for helping me improve my coding skills and providing an innumerable amount of great tools,

and of course my family, girlfriend and friends for their permanent encouragement making this thesis possible at all.