Search for Lorentz Violation at the KATRIN Experiment

Master thesis at the faculty of physics
of the
Ludwig-Maximilians-Universität München

submitted by
Johannes Gerald Wickles
born in Lichtenfels

Munich, April 15, 2021

Supervisor (LMU): Prof. Dr. Thomas Kuhr
Supervisor (MPP): Prof. Dr. Susanne Mertens

Abstract

Lorentz invariance is a central part of modern physics. However, some theories that try to combine quantum field theory and general relativity, argue that Lorentz invariance might be violated and is not an exact symmetry. To test Lorentz violation (LV), a Standard Model extension was developed, which includes LV operators. Many of those are already constrained by experiments. However, for the a^{μ}-type Lorentz violation there are no current limits in the neutrino sector. The Karlsruhe Tritium Neutrino (KATRIN) experiment, which aims to measure the effective mass of the electron anti-neutrino, can be used to investigate effects caused by this Lorentz violation mode. In this thesis the sensitivity of the KATRIN experiment to this type of Lorentz invariance violation is studied. In particular a possible oscillation and shift of the endpoint of the Tritium spectrum is investigated. Using a Bayesian as well as a frequentist analysis it is shown that the KATRIN experiment has the sensitivity to set the first limit on this LV parameter.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

MAX-PLANCK-INSTITUT FÜR PHYSIK

Suche nach Lorentzverletzung am KATRIN Experiment

Masterarbeit an der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Johannes Gerald Wickles
geboren in Lichtenfels

München, den 15. April 2021

Betreuer (LMU): Prof. Dr. Thomas Kuhr
Betreuerin (MPP): Prof. Dr. Susanne Mertens

Zusammenfassung

Lorentzinvarianz ist ein zentraler Bestandteil der modernen Physik. Manche Theorien, die Quantenfeldtheorie und Relativitätstheorie vereinen möchten, suggerieren jedoch, dass Lorentzinvarianz keine exakte Symmtrie ist und gebrochen sein könnte. Um eine mögliche Lorentzverletzung (LV) zu testen, wurde eine Erweiterung des Standardmodells eingeführt, die LV-Operatoren enthält. Die Größe vieler dieser Operatoren wurde bereits durch Experimente eingeschränkt. Für die a^{μ}-Typ Lorentzinvarianz Verletzung (LV) im Neutrino-Sektor gibt es jedoch bis dato keine experimentelle obere Grenze. Das Karlsruhe Tritium Neutrino (KATRIN) Experiment, welches das Ziel hat, die Masse des Elektron-Neutrinos zu bestimmen, kann genutzt werden, um den Wert dieser Art von LV-Operator einzugrenzen. In dieser Arbeit wird die Sensitivität von KATRIN auf diesen Operator untersucht. Insbesondere wird dabei eine mögliche globale Verschiebung und Oszillation des Endpunkts des Tritium-Spektrums erforscht. Mittels einer frequentistischen und einer Bayesschen Analyse zeigt diese Arbeit, dass das KATRIN Experiment die Sensitivität besitzt, eine erste obere Grenze für diese Art der a^{μ}-Typ LV zu setzen.

Contents

Chapter 1

Introduction

In the year 1930, the physics community was puzzled by the spectra of the β-decays of ^{6}Li nuclei which were observed to be continuous. However, at this time only the daughter nucleus and the electron were observed as decay products. Thus energy-momentum conservation dictated a discrete spectrum. In a letter to his colleagues Wolfgang Pauli proposed the idea of introducing a neutral particle as a further decay product that has a large penetration length and is, therefore, difficult to measure. This particle, which originally was called "neutron", is now known as "neutrino" ν and explains the continuity of the β-spectra [1].
In modern physics, symmetry and conservation laws play a central role. A theoretical model obeys a specific symmetry if it does not change under a specific transformation. According to Noether's theorem, continuous symmetries imply conserved quantities [2]. Therefore theories can be constructed in a way that they agree with observed conservation laws - like the introduction of a neutrino was a result of applying the energy-momentum conservation in the β-decay [3].
A symmetry that is part of the Standard Model (SM), as well as general relativity, is called Lorentz symmetry, which describes the invariance under linear transformations of space-time. Originally this kind of transformation was studied in the context of electrodynamics as symmetry of the Maxwell-equations [4]. Later Einstein could derive this symmetry based on the foundations of his special relativity. In particular he assumed that the speed of light is constant in every reference frame. This basic assumption paved the way for Lorentz invariance as a central part of modern physics.
Even though first tests of this symmetry dating back to Michelson and Morley [5] or Kennedy and Thorndike [6] could not find a deviation from Lorentz invariance, there has been an increased interest in probing this symmetry in recent years. Inspired by theoretical suggestions that Lorentz invariance might be an effective symmetry and violated at high energies, the interest in experiments testing the prediction of Lorentz symmetry has risen. In particular, approaches to combine quantum field theory and general relativity have resulted in new theoretical frameworks like string theory and loop quantum gravity. Some of those theories argue that Lorentz invariance might be violated and not an exact symmetry [7]. Lorentz invariance is related to CPT-symmetry due to the CPT-theorem which states that the combination of charge (C), parity (P) and time (T) conjugation is preserved. Therefore Lorentz violating (LV) operators can also be CPT-violating.
Since those violations are expected for the high energy regime, which cannot be reached by current particle accelerators, an effective field theory including LV was introduced - the so-called "Standard-Model Extension" (SME) [8]. This tool is used to probe LV as modifications in the energy regime of the Standard Model.
A candidate for a messenger of physics beyond the Standard Model is the neutrino. Historically, investigating the neutrino led to an improved understanding of physics, e.g. in the

context of the beta-decay [1]. In particular the examination of neutrino oscillation brought evidence of physics beyond the Standard Model [9]. Therefore it might be of interest to use neutrinos to study deviations from exact Lorentz symmetry. In the SME all LV-operators for neutrino propagation are specified [10]. Most of them are studied by oscillation experiments. However, the so-called oscillation-free modes cannot be studied by those experiments and are usually investigated by time-of-flight experiments. Nevertheless, there are four LV operators that cannot be analyzed by that kind of approaches, but e.g. in decay-experiments [9]. In particular one can focus on anisotropic effects of LV in experiments by changing their direction to investigate the anisotropy. In the Karlsruhe Tritium neutrino experiment (KATRIN) this change is introduced by the rotation of the earth. Therefore it is of large interest to examine whether the β-decay shows variations with sidereal time.
In this thesis the effects of LV on the endpoint of the electron spectrum of Tritium decay are analyzed. Especially, the sensitivity of the KATRIN experiment for such an analysis will be studied. In chapter 3 the theoretical foundations of LV will be examined. Additionally the model for anisotropic effects on the endpoint will be discussed. In chapter 4 and chapter 5 the development of an analysis strategy to determine the influence of LV effects on the endpoint is described. Finally, this framework is also applied to the data of the first neutrino mass measurement of KATRIN and a prognosis for future results is given (see chapter 6).

Chapter 2

The neutrino and the KATRIN experiment

2.1 Discovery of the neutrino

As discussed previously Wolfgang Pauli suggested the existence of the "neutrino" to solve the puzzle about the form of the β-decay spectrum [1]. Building on this idea Enrico Fermi developed a quantum theory of the β-decay that includes the neutrino. He suggested that the neutron n decays in a three-body-decay into an electron e^-, a proton p^+ and an (anti)neutrino $\bar{\nu}$ (see eq. (2.1)). This theory explains the continuous spectrum of the β-decay via energy-momentum conservation because the individual energies of the decay products are not fixed [11].

$$n \rightarrow p^+ + e^- + \bar{\nu}. \tag{2.1}$$

It took more than 20 years until the existence of the neutrino was confirmed by Clyde L. Cowan, Frederick Raines, et al.. Their idea was to use the (potential) neutrino flux of a nuclear power plant as a neutrino source and to detect the neutrinos by their reaction with protons in the target tank. This reaction creates neutrons and positrons e^+ (see eq. (2.2)). The decay products were detected using a pair of γ-pulses. The first photon pulse is due to the deceleration and the annihilation of the positron. A delayed pulse is recognized as the moderated neutron is captured by Cadmium Cd (see eq. (2.3)) implanted in the tank. Cowan and Raines observed a signal which was dependent of the power of the nuclear power plant and showed good agreement with the predicted flux by eq. (2.1) and therefore provided the first confirmation of the existence of the neutrino [12].

$$\nu + p^+ \rightarrow e^+ + n \tag{2.2}$$

$$^{133}\mathrm{Cd} + n \rightarrow {}^{134}\mathrm{Cd} + \gamma \tag{2.3}$$

Throughout time the neutrino became a fundamental part of the model that describes the elementary particles - the Standard Model (SM). In the SM leptons are grouped into families of different flavors, namely electron, muon and tauon flavors (e, μ, τ). SM neutrinos are predicted to be massless particles [13]. However, in the next section it will be shown that neutrino oscillation experiments suggest that neutrinos do have a mass.

2.2 Neutrino oscillation and mass

Neutrinos are e.g. produced in the fusion processes of the sun. The standard solar model predicts a neutrino flux due to the decay of ^{8}B. This prediction inspired Raymond Davis Jr. to detect the solar-neutrinos in the famous "Homestake experiment". He used 390 000 l of liquid tetrachloroethylene in an underground laboratory in the Homestake mine. The location below the ground was used to reduce background from ^{37}Ag produced by cosmic rays. The idea of the experiment is that the neutrinos produce argon (see eq. (2.4)), which is constantly removed from the tank and therefore provides information about the neutrino flux. As a result of the experiment Davis Jr. concluded that the observed neutrino flux is about a factor seven smaller than the one expected by the solar model [14].

$$^{37}\mathrm{Cl} + \nu_e \rightarrow {}^{37}\mathrm{Ag} + e^- \tag{2.4}$$

One of the explanations for this effect is the so-called neutrino oscillation, meaning that neutrinos change their flavor while traveling over distances. Subsequently, many researchers observed neutrino oscillation using experiments in which neutrinos travel over (large) distances. Some experiments e.g. looked at atmospheric neutrinos from hadronic showers of cosmic rays. The ratio of the electron and muon neutrino flux was calculated by theory and compared with the measurements. The probability that a neutrino ν_a oscillates to ν_b after traveling a distance L depends on its energy E_ν, the so-called "mixing-angle" θ and the mass difference Δm^2 of the neutrino flavors [15].

$$P_{a\rightarrow b} = \sin^2 2\theta \sin^2\left(\frac{1.27\Delta m^2\,(\mathrm{eV}^2)\,L(\mathrm{km})}{E_\nu(\mathrm{GeV})}\right) \tag{2.5}$$

The explanation for the neutrino oscillation arises from the fact that for neutrinos the flavor eigenstates are different from the mass eigenstates. Each flavor is a composite of different mass eigenstates described via the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [16, 17]

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}. \tag{2.6}$$

The Kamiokande experiment showed that neutrino oscillations exist. According to eq. (2.5) this implies a non-zero mass difference and thus neutrino mass [15]. However, oscillation experiments are only sensible to Δm^2. Therefore a different kind of experiment is needed to determine the absolute neutrino mass. Such an experiment is described in the following section.

2.3 The KATRIN experiment

The KArlsruhe TRItium Neutrino (KATRIN) experiment is designed to measure the effective electron (anti-)neutrino mass [18]:

$$m\,(\nu_e) = \sqrt{\sum_{i=1}^{3} |U_{ei}|^2\, m_i^2}. \tag{2.7}$$

After about three years of data taking the KATRIN collaboration aims to achieve a sensitivity of [19]

$$m\,(\nu_e) < 0.2\,\mathrm{eV} \quad (90\%\,\mathrm{C.L.}). \tag{2.8}$$

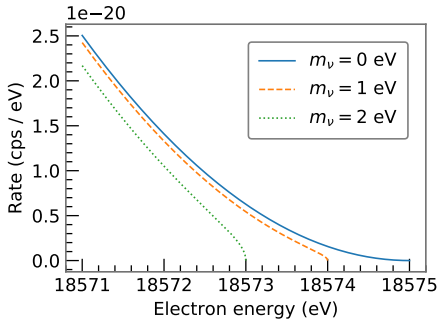


Figure 2.1: This figure shows differential β-spectra for different values of m_ν^2. The shape of the spectrum close to the endpoint is distorted due the influence of the neutrino mass. Therefore KATRIN aims to measure this distortion. Figure from [18].

In its approach to limit the neutrino mass by a direct kinetic method, the KATRIN experiment is similar to its predecessors in Troitsk [20] and Mainz [21].
Like in the previous experiments, the β-decay of Tritium is used to measure the neutrino mass as described in section 2.3.1. Section 2.3.2 gives an overview of the experimental setup of the KATRIN experiment. Its model and first results are presented afterward.

2.3.1 Tritium β-decay

Aiming to measure the neutrino mass via a direct kinetic method the KATRIN experiment utilizes the electrons emitted by Tritium β-decays:

$$^3\mathrm{H} \rightarrow {}^3\mathrm{He}^+ + e^- + \bar{\nu}_e \tag{2.9}$$

The signature of a non-zero neutrino mass would manifest itself in a distortion of the energy spectrum of the emitted electrons. In fig. 2.1 this spectral distortion is illustrated for different neutrino masses. The effect is strongest close to the endpoint E_0 of the spectrum as the count rate increases rapidly below the endpoint with $\frac{dN}{dE} \propto (E - E_0)^2$. Therefore mostly this narrow region is analyzed [19].
Using Fermi's rule one can obtain the differential spectrum of the Tritium β-decay as in eq. (2.10). $F(Z, E)$ is the Fermi function which describes the final electromagnetic interaction of electron and daughter nucleus with charge Z. E, p and m_e correspond to the energy, momentum and mass of the electron. The Heaviside function Θ guarantees that energy is conserved while the constant C includes the nuclear matrix element M_{nuc}, the Fermi constant G_{F} and Cabibbo angle θ_{C} [18, 22].

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E} = C \cdot F(Z,E) \cdot p \cdot (E + m_e) \cdot (E_0 - E) \cdot \sqrt{(E_0 - E)^2 - m_\nu^2} \cdot \Theta(E_0 - E - m_\nu) \tag{2.10}$$

$$C = \frac{G_{\mathrm{F}}^2 \cos^2 \theta_{\mathrm{C}}}{2\pi^3} \cdot |M_{\mathrm{nuc}}|^2 \tag{2.11}$$

To precisely model the spectrum, some changes to the differential spectrum become necessary. One has to include e.g. Doppler broadening and the distribution of the molecular final states, as discussed in [18, 23].

2.3.2 Experimental overview

The KATRIN experiment, which was built to measure the electron spectrum of the Tritium decay, has a total length of about 70 m. The beamline (see fig. 2.2) can be divided into several

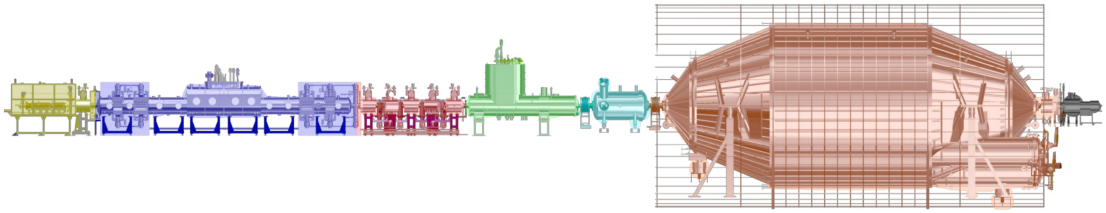


Figure 2.2: This figure shows the beamline of the KATRIN experiment and its sections. Figure from [18].
a) Rear Section
b) Windowless Gaseous Tritium Source
c) Differential pumping section
d) Cryogenic pumping section
e) Prespectrometer
f) Main Spectrometer
g) Focal plane detector

parts, which fulfill their specific function. In the following the different components will be described.

Rear Section

The rear section has mainly calibration and monitoring purposes. It provides a reference potential for the plasma in the source via a conducting surface, namely the rear wall. Furthermore it is equipped with monitoring and calibration tools. For example it monitors the activity of the source via beta-induced X-ray spectroscopy and includes an electron source for calibration [24].

Windowless Gaseous Tritium Source

A Windowless Gaseous Tritium Source (WGTS) is the standard source for the electrons in the long-time measurements of KATRIN. This section provides ultracold ($T = 27\,\mathrm{K}$) molecular Tritium with high purity ($> 95\,\%$) in a 10 m long stainless steel tube. The gas is injected through capillaries and by controlling the pressure, the column density ρd of the source can be adjusted relative to its nominal value of $5 \times 10^{17}\,\mathrm{cm}^{-2}$. The electrons, which are produced in the β-decays, are adiabatically guided to the end of the tube via the source magnetic field B_S [19].

Transport Section

The transport section has two main purposes. Firstly, it adiabatically guides the electrons from the WGTS to the spectrometer. Secondly, it reduces the Tritium gas flow by about twelve orders of magnitude. This is necessary to prevent residual Tritium gas in the spectrometer, which would increase the background. The transport section consists of a Differential Pumping Section (DPS) and a Cryogenic Pumping Section (CPS). The DPS uses multiple turbomolecular pumps to reduce the gas flow by five orders of magnitude. The CPS acts as an approximately 3 m long cold trap. The electrons are guided through this trap by magnets

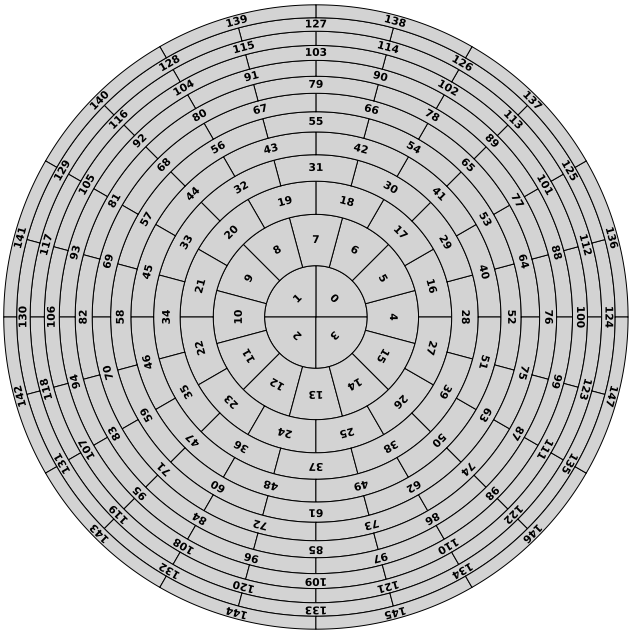


Figure 2.3: This figure shows the Focal plane detector (FPD), which consists of 148 pixels. Their arrangement is illustrated in this plot.

while the neutral gas is trapped. After passing the transport section the electrons enter the spectrometer while the flow of residual Tritium gas is reduced by a factor of 10^{12} [25].

Pre- and Main Spectrometer

The energy of the electrons from the source is analyzed with MAC-E filters (see section 2.3.3) in the pre-spectrometer (PS) and the main spectrometer (MS). The PS is operated at a fixed retarding voltage to reduce the background by filtering out electrons deep in the spectrum as well as ions. The MS can change its retarding voltages to measure an integrated spectrum of the electrons close to the endpoint with a resolution of about 1 eV. The spectrometers are operated at low pressures smaller than 1×10^{-11} mbar to minimize background from ionization of residual molecules. Additionally an inner wire electrode, that has a slightly more negative potential than the vessel, is attached in the MS. Its purpose is to suppress the number of electrons emanating from the walls that enter the inner part of the spectrometer [19].

Focal plane detector

The focal plane detector (FPD) rounds out the beamline. Behind the superconducting pinch magnet, which completes the MAC-E filter, the electrons are further accelerated by about 10 keV using a post-acceleration-electrode to reduce the background from β- and γ- radiation close to the detector. Using another superconducting magnet the electrons are guided to the FPD which is a multipixel silicon detector. It consists of 148 pixels with an equally large surface as shown in fig. 2.3. Each pixel counts the number of electrons, which allows to include radial and azimuthal variations in the experimental response function in the analysis [19, 25].

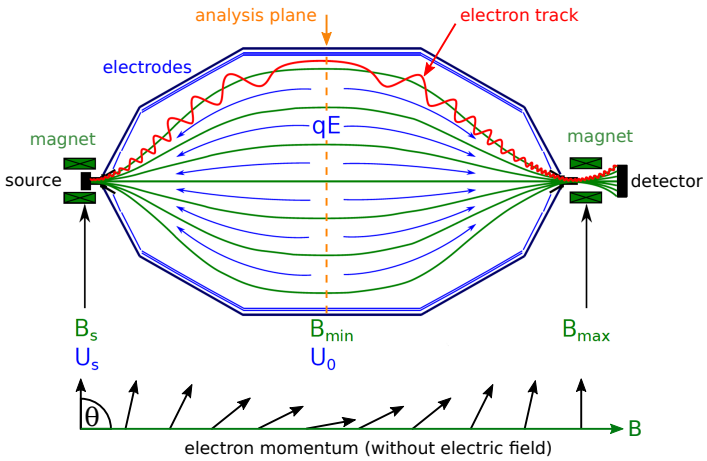


Figure 2.4: This figure illustrates the MAC-E filter principle. The transverse motion of the electrons is transformed into logitudinal motion. The electric potential at the analysis plane filters the electrons which energy is larger than the retarding potential qU. Figure from [26].

2.3.3 MAC-E filter

Similar to its predecessors in Mainz and Troitsk the KATRIN experiment uses a Magnetic Adiabatic Collimation in combination with an Electrostatic Filter (MAC-E filter) as it combines high resolution with high luminosity. This is important in order to measure the endpoint region of the Tritium spectrum [19].
The electrons transported from the source are guided in the spectrometer by the magnetic field and perform cyclotron motion around the field lines (see fig. 2.4). In the spectrometer the magnetic field drops slowly by many orders of magnitude. The resulting magnetic gradient transforms the transverse cyclotron motion into longitudinal motion. As the magnetic field varies slowly the transfer from transverse energy $E_\perp$ to longitudinal motion is adiabatic and the magnetic moment is conserved [19]:

$$\mu = \frac{E_\perp}{B} = \text{ const.} \tag{2.12}$$

This means that the electron motion is transformed until the electrons move almost perfectly aligned to the magnetic field lines. This beam of parallel electrons hits an electric potential such that only electrons that have enough energy can pass this barrier and are re-accelerated towards the detector. In this way the spectrometer acts as a high-pass filter. Varying the potential of the electrode is used to measure the spectrum. From eq. (2.12) follows that the energy resolution is defined by the minimal magnetic field B_{min} and the maximal one B_{max} [19]

$$\frac{\Delta E}{E} = \frac{B_{\text{min}}}{B_{\text{max}}}. \tag{2.13}$$

2.3.4 Model of the integral spectrum

The model of the Tritium beta decay integral spectrum consists of several parts. The differential form of the decay according to Fermi is shown in section 2.3.1. This spectrum is combined with an experimental response function. Additionally other modifications (e.g. for background, Doppler broadening) are incorporated into the model of the integral spectrum.

Response function

The MAC-E filter functions as a high-pass filter, meaning only electrons with sufficiently large longitudinal energy reach the detector. This property is described by the so-called transmission function. In the case of an ideal high-pass filter, the response is a step-function [18]

$$R_{\text{ideal}}\left(qU, E\right) = \begin{cases} 0, & \text{if } E < qU \\ 1, & \text{if } E \geq qU \end{cases} . \tag{2.14}$$

However, the MAC-E filter does not perfectly convert the perpendicular momentum into a longitudinal one, meaning that electrons that have a total energy larger than the retarding potential but start at larger angles, may not reach the detector. Using eq. (2.13) a relativistic transmission function can be derived as [18]

$$T(qU, E) = \begin{cases} 0, & \text{if } E < qU \\ \frac{1-\sqrt{1-f\cdot\frac{B_{\text{S}}}{B_{\text{min}}}\cdot\frac{E-qU}{E}}}{1-\sqrt{1-\frac{B_S}{B_{\text{max}}}}}, & \text{if } qU \leq E \leq qU\frac{f\cdot B_{\text{max}}}{f\cdot B_{\text{max}}-B_{\text{min}}} \\ 1, & \text{if } E > qU\frac{f\cdot B_{\text{max}}}{f\cdot B_{\text{max}}-B_{\text{min}}} \end{cases} \tag{2.15}$$

with the relativistic factor

$$f = \frac{\frac{E-qU}{m_e}+2}{\frac{E}{m_e}+2}. \tag{2.16}$$

This transmission function is depicted in fig. 2.5a.
However, there is another process that changes the response function. Inelastic scattering of electrons with Tritium molecules inside the source results in a major energy loss. Thus the response function is modified by combining the transmission function with the scattering probability and the corresponding energy loss function [18]. This leads to the response function which is illustrated in fig. 2.5b.
Additionally, in order to reduce the number of electrons that have a long path-length within the Tritium source and are likely to scatter often, the magnetic mirror effect is used to create an acceptance cone. This allows only electrons within a specific angle θ_0 to arrive at the spectrometer. The acceptance angle is defined by the magnetic field of the source B_S and the maximal magnetic field in the spectrometer B_{max} via [19]

$$\sin\theta_{\text{max}} = \sqrt{\frac{B_{\text{S}}}{B_{\text{max}}}}. \tag{2.17}$$

Integral spectrum

The KATRIN experiment measures an integrated spectrum $I(qU)$. In order to model the expected rate, the differential spectrum $\frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E)$ is integrated over the response function $R(qU, E)$ [18]:

$$I(qU) = C \cdot \int_{qU}^{E_0} \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E) \cdot R(qU, E)\mathrm{d}E \tag{2.18}$$

with

$$C = N_{\text{eff}} \cdot \frac{1-\cos\theta_{\text{max}}}{2} \cdot \epsilon_{\text{detector}} \tag{2.19}$$

where N_{eff} is the effective number of Tritium atoms in the source, θ_{max} is the solid angle and $\epsilon_{\text{detector}}$ describes the detector efficiency.
A background rate B, which is constant in the retarding potentials, is added such that the model reads

$$\Gamma(qU) = I(qU) + B = C \cdot \int_{qU}^{E_0} \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E) \cdot R(qU, E)\mathrm{d}E + B. \tag{2.20}$$

Model parameters

This model is used to fit the data. The fit has four free parameters, the normalization of the spectrum, the square of the neutrino mass m_ν^2, the background B and the endpoint E_0 [18]. At this point it is worth mentioning that KATRIN's parameter of interest is the neutrino mass squared which distorts the shape of the spectrum. However, for LV analysis in this thesis the important parameter is the endpoint. Additionally one should notice that it is an experiment-dependent quantity as it includes e.g. off-set potentials.

2.3.5 FITRIUM software

For modeling and fitting the data of the KATRIN experiment the software "FITRIUM" is used. It can be applied to fit the data to different kinds of spectra, like the integral spectrum (see section 2.3.4). FITRIUM is developed and maintained by Christian Karl. Besides fitting the spectrum it is also capable of doing Monte Carlo (MC) simulations and MC propagation to include systematic uncertainties [18].
To receive the best fit the likelihood function $\mathscr{L}$ of a model $\mu(\vec{\theta})$ is maximized with respect to the parameters $\vec{\theta}$ [18].

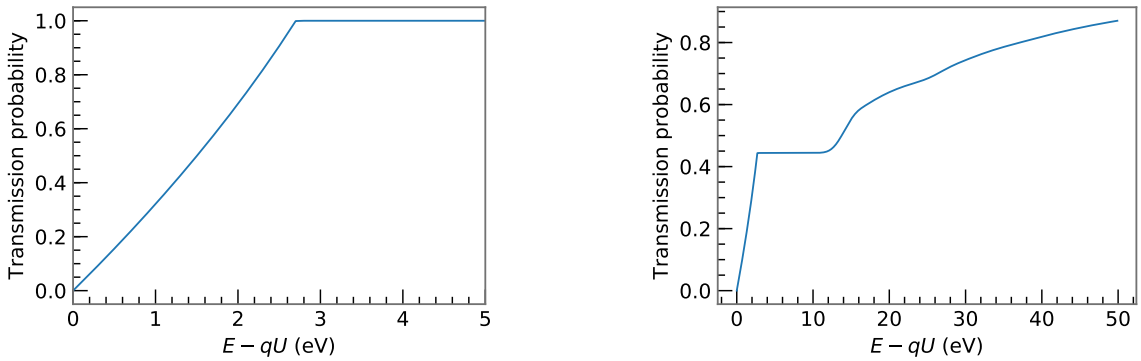
$$\mathscr{L} = \mathscr{L}(\mu(\vec{\theta}); x) = \mathscr{L}(\vec{\theta}; x) \tag{2.21}$$

In the KATRIN experiment the number of counts N follows a Poisson distribution. However, for large numbers of counts the law of large numbers predicts that the likelihood is asymptotically described by a normal distribution with the model predictions μ

$$P(N \text{ counts}, \mu) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{(N-\mu)^2}{2N}}. \tag{2.22}$$

Instead of maximizing the likelihood it is common to minimize $-\ln \mathscr{L}(\vec{\mu}; \vec{N})$, which is equivalent of minimizing the χ^2 for a Gaussian likelihood

$$-\ln \mathscr{L}(\vec{\mu}; \vec{N}) \propto \sum_i \frac{(N_i - \mu_i)^2}{2N_i} =: \frac{1}{2}\chi^2(\vec{\mu}; \vec{N}) \tag{2.23}$$



(a) Transmission function

(b) Response function

Figure 2.5: These plots show the transmission function on the left and the experimental response function on the right. The transmission function takes into account the fact, that not the full transversal energy of the electrons is converted into longitudinal motion. The response function also includes effects like scattering in the source and energy losses. Figures from [18].

where the index i indicates the i^{th} data point.
One approach to include systematic uncertainties is the "Nuisance Parameter Method" which is also called "pull-term method". In this technique, the systematic parameters are added as further fit parameters. However, to incorporate some constraints the likelihood function is modified by a pull-term, which only depends on the systematic parameters. Usually this is done by including a Gaussian to the likelihood which constrains the systematic parameters in their expectation values and uncertainties. However, this is computationally expensive as the systematic parameters are included as additional fit parameters which need to be minimized [18].
An alternative method is using covariance matrices $V_{\text{tot}} = V_{\text{stat}} + V_{\text{sys}}$ where the χ^2 gets modified like

$$\chi^2(\vec{\mu}; \vec{N}) = (\vec{\mu} - \vec{N})^T V_{\text{tot}}^{-1} (\vec{\mu} - \vec{N}). \tag{2.24}$$

Besides statistical uncertainty the covariance matrix includes also systematic uncertainties. These are included by calculating several spectra sampling the systematic parameters from their distribution and calculating the covariance matrix of those spectra [18].
A third method uses MC propagation of uncertainties. The values of the systematic parameters are randomly drawn from their distribution. Then the data is fitted using these fixed values. This process is repeated several times to retrieve the distribution of the fit parameters. More details about the different treatments of the systematic uncertainties can be found in [18].

2.3.6 First neutrino mass measurement campaign

In spring 2019 the first measurement campaign, KATRIN neutrino mass 1 (KNM-1), took place. The integral spectrum was scanned multiple times over a period of one month. Each scan had a duration of about 2 h consisting of various subscans at different retarding potentials qU. Within a subscan the retarding potential is fixed and only the electrons with energy large enough to overcome this barrier are counted by the detector. The measurement time was distributed in a way that more time was spent at higher retarding potentials. This region is most sensitive to the distortion of the spectrum which is caused by the neutrino mass (see fig. 2.6). Furthermore the scan-direction was alternated between up-scans (increasing the retarding potentials during a run) and down-scans (decreasing the retarding potentials during a run) to compensate for drifts of the system [27].
Based on data-quality 274 runs ("golden-runs") were selected. The KATRIN collaboration also defined a list of 117 out of 148 pixels which are included in the analysis. The excluded pixels were shadowed or noisy. Figure 2.6 shows the spectrum and best fit of KNM-1. The error bars are enlarged by a factor of 50 for illustration purposes. The residuals show no obvious pattern which supports the statement that this measurement phase was dominated by statistical uncertainty. The fit parameters of the model are the ones described in section 2.3.4. The covariance method, as well as MC propagation, were used to include the systematic uncertainties of the experiment. The final fit results showed a squared neutrino mass of [27]

$$m_\nu^2 = \left(-1.0^{+0.9}_{-1.1}\right) \text{eV}^2. \tag{2.25}$$

Using the methods of A. V. Lokhov and F. V. Tkachov (see [28]) an upper limit for the neutrino mass was derived [27]:

$$m_\nu \leq 1.1\text{eV (90\% C.L.)} \tag{2.26}$$

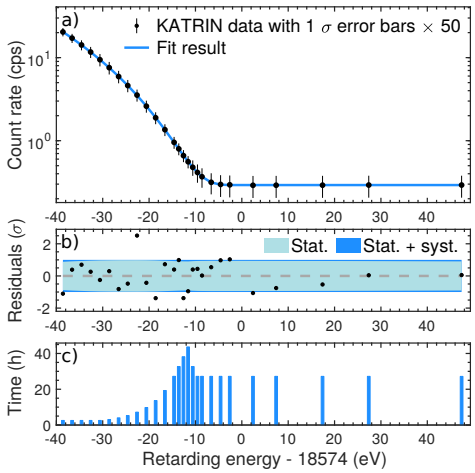


Figure 2.6: This figure shows the fit result of the first measurement campaign of KATRIN (KNM-1). a) shows the integral spectrum of the electrons. The the error-bars are increased by a factor of 50 to make them visible. b) shows the normalized residuals and the 1σ-band. In c) the measurement time distribution is illustrated. Figure from [27].

2.3.7 Probes for physics beyond the Standard Model

Even though the main purpose of KATRIN is to measure the neutrino mass, the experiment also offers insights into different fields of physics, especially non Standard Model effects. The KATRIN experiment used e.g. the data of KNM-1 to limit the parameter space of a hypothetical sterile neutrino [29]. The sterile neutrino is an extension of the Standard Model by a neutrino that does not take part in the electro-weak interaction. Inter alia sterile neutrinos are motivated by the so-called "reactor anomaly" [30], as they could explain the fact that short-baseline experiments (smaller 100 m) using nuclear reactors observed less electron neutrinos than expected [31].
Another modification to the SM, which is of interest in Grand Unified Theories, is the violation of Lorentz invariance. Assuming that this symmetry is no exact symmetry but a low-energy approximation, Lorentz violation (LV) can change the spectrum of the β-decay [9, 19]. The effects of LV at the KATRIN experiment and the sensitivity of KATRIN to limit LV is the topic of this thesis.

Chapter 3

Lorentz violation theory and the KATRIN experiment

Lorentz symmetry is a central principle of relativity and all accepted fundamental descriptions of physics. The examination of the propagation of light played a crucial role in establishing rotation and boost symmetry of Lorentz invariance. A famous experiment was performed by Michelson and Morley, which searched for anisotropic effects in the velocity of light. Other experiments, like the Kennedy-Thorndike-experiment, investigated the boost invariance. Both experiments could not find a deviation of Lorentz symmetry [5, 6, 32].
However, it is believed that the Standard Model is only an effective theory and the underlying fundamental theory is suppressed by the Planck mass m_P in the order of $m_W/m_P \approx 10^{-17}$. To find those effects experiments would need a high sensitivity. Therefore one attempts to search for deviations of the symmetries in the SM, like the violation of Lorentz invariance. In order to describe this theoretically, one can extend the SM minimally by including Lorentz invariance violation terms in the Lagrangian [8].
Many quantum field operators that violate Lorentz invariance have been studied by interferometrical methods using the effects of neutrino oscillation (see e.g. [33, 34]). However, these experiments cannot test the so-called oscillation-free modes. Operators of that kind have been studied by examining the neutrino propagation, e.g. in time-of-flight experiments. Nevertheless four Lorentz violation operators are not measurable by this method, as they are "countershaded" [35]. Their size can be determined by an analysis of their effect on the spectrum of the β-decay [9].
The following chapter will briefly discuss the theoretical foundation of LV in the neutrino sector and describe the effects of LV in the KATRIN experiment.

3.1 Theoretical development of Lorentz violation operators

To understand the origin and effects of LV operators, the following section will focus on a derivation in an educational case.

3.1.1 Fermions

As the focus of this work is on neutrinos one can have a look at non interacting fermions for simplicity. In this case all Lorentz-violating terms in the action are fermion bilinears.

Typically, one can define a $2N$- dimensional spinor-multiplet as

$$\Psi_A = \begin{pmatrix} \psi_a \\ \psi_a^C \end{pmatrix} \tag{3.1}$$

with $A \in [1, 2N]$. In general one can write LV and CPT violation as a modification of the Lagrangian density with terms that are not invariant under such transformations [10]:

$$\mathcal{L} = \frac{1}{2}\bar{\Psi}_A \left(\gamma^\mu i\partial_\mu \delta_{AB} - M_{AB} + \widehat{\mathcal{Q}}_{AB}\right) \Psi_B + \text{h.c.} \tag{3.2}$$

In this equation the first term is the standard kinetic term and M_{AB} is a general mass matrix. The third term is the one that includes the LV. $\widehat{\mathcal{Q}}_{AB}$ is a matrix with 2×2 dimensions in spinor and $2N \times 2N$ dimensions in flavor space and includes also derivatives like $i\partial_\mu$ [10].
At this point it is obvious that the LV term modifies the Dirac equation like,

$$\left(p \cdot \gamma \delta_{AB} - M_{AB} + \widehat{\mathcal{Q}}_{AB}\right) \Psi_B = 0. \quad . \tag{3.3}$$

Multiplying with γ_0 from left and using the energy eigenstate equation one finds a definition for the Hamiltonian:

$$\left(E\delta_{AB} - H_{AB}\right) \Psi_B = \gamma_0 \left(p \cdot \gamma \delta_{AB} - M_{AB} + \widehat{\mathcal{Q}}_{AB}\right) \Psi_B = 0 \tag{3.4}$$

where $E = p_0$. Therefore one can identify the total effective Hamiltonian

$$H_{AB} = \gamma_0 \left(\vec{p} \cdot \vec{\gamma} \delta_{AB} + M_{AB} - \widehat{\mathcal{Q}}_{AB}\right) = (H_0)_{AB} + \delta H_{AB} \tag{3.5}$$

consisting of a conventional part $(H_0)_{AB}$ and a Lorentz violating perturbation δH_{AB}. Thus, up to leading-order the effective Hamiltonian is [10]

$$H_{AB} = (H_0)_{AB} - \gamma_0 \left(\widehat{\mathcal{S}}_{AB} + i\widehat{\mathcal{P}}_{AB}\gamma_5 + \widehat{\mathcal{V}}^\mu_{AB}\gamma_\mu + \widehat{\mathcal{A}}^\mu_{AB}\gamma_5\gamma_\mu + \frac{1}{2}\widehat{\mathcal{T}}^{\mu\nu}_{AB}\sigma_{\mu\nu}\right) |_{E\to E_0}\,. \tag{3.6}$$

The LV terms are defined as [10]

$$\begin{array}{lll} \widehat{\mathcal{S}}_{AB} = \widehat{e}_{AB} - \widehat{m}_{AB}, & \widehat{\mathcal{P}}_{AB} = \widehat{f}_{AB} - \widehat{m}_{5\text{AB}}, & \widehat{\mathcal{V}}^\mu_{AB} = \widehat{c}^\mu_{AB} - \widehat{a}^\mu_{AB} \\ \widehat{\mathcal{A}}^\mu_{AB} = \widehat{d}^\mu_{AB} - \widehat{b}^\mu_{AB}, & \widehat{\mathcal{T}}^{\mu\nu}_{AB} = \widehat{g}^{\mu\nu}_{AB} - \widehat{H}^{\mu\nu}_{AB}. & \end{array} \tag{3.7}$$

For a more detailed derivation and definition of those terms, refer to appendix A and [10].

3.1.2 Neutrinos

(Anti-)Neutrinos, which are a fermionic decay product of the Tritium decay, are chiral fermions. Therefore the formalism will be projected on left-handed fields. It is convenient to split the mass matrix into left or right handed parts $M = m_L P_L + m_R P_R$. The Majorana and Dirac parts can be isolated as

$$m_R \mathcal{C} = \begin{pmatrix} L & D \\ D^T & R \end{pmatrix}. \tag{3.8}$$

L and R are left- and right-handed Majorana-type mass matrices, respectively, while D is Dirac-type [10].
For projected theories it is common to uses the so-called "Weyl-basis" including Weyl-spinors $\Phi_\text{W} = \begin{pmatrix} \phi \\ \phi^C \end{pmatrix}$ with $\phi^C = i\sigma^2\phi^*$ [10]. As previously discussed, one can describe small LV theory as a perturbation and the effective net Hamiltonian is

$$h_\text{eff} = (h_\text{eff})_0 + \delta h \tag{3.9}$$

with $\delta h = \frac{1}{|\vec{p}|}\begin{pmatrix} \widehat{a}_{\text{eff}} - \widehat{c}_{\text{eff}} & -\widehat{g}_{\text{eff}} + \widehat{H}_{\text{eff}} \\ -\hat{g}^{\dagger}_{\text{eff}} + \hat{H}^{\dagger}_{\text{eff}} & -\widehat{a}^{T}_{\text{eff}} - \widehat{c}^{T}_{\text{eff}} \end{pmatrix}$.

Here

$$\begin{aligned} \widehat{a}_{\text{eff}} &= p_\mu \widehat{a}^{\mu}_{L} - \widehat{e}_l + 2i\epsilon_\mu \epsilon^{*}_{\nu} \widehat{g}^{\mu\nu}_{l} \\ \widehat{g}_{\text{eff}} &= i\sqrt{2} p_\mu \epsilon_\nu \widehat{g}^{\mu\nu}_{M+} + \sqrt{2}\epsilon_\mu \widehat{a}^{\mu}_{l} \end{aligned} \tag{3.10}$$

are CPT-odd while

$$\begin{aligned} \widehat{c}_{\text{eff}} &= p_\mu \widehat{c}^{\mu}_{L} - \widehat{m}_l + 2i\epsilon_\mu \epsilon^{*}_{\nu} \widehat{H}^{\mu\nu}_{l} \\ \widehat{H}_{\text{eff}} &= i\sqrt{2} p_\mu \epsilon_\nu \widehat{H}^{\mu\nu}_{M+} + \sqrt{2}\epsilon_\mu \widehat{c}^{\mu}_{l} \end{aligned} \tag{3.11}$$

are CPT-even.

3.1.3 The flavor-blind oscillation-free mode

In the case of a flavor-blind model one has three copies of eq. (3.9) which can be diagonalized for this case to

$$h^{\text{fb}}_{\text{eff}} = \begin{pmatrix} C & S \\ -S^{*} & C \end{pmatrix} \begin{pmatrix} E^{\text{fb}}_{+} & 0 \\ 0 & E^{\text{fb}}_{-} \end{pmatrix} \begin{pmatrix} C & -S \\ S^{*} & C \end{pmatrix} \tag{3.12}$$

with $C = \sqrt{\frac{\lambda + \widehat{a}_{\text{eff}}}{2\lambda}}$, $\quad S = \frac{\widehat{g}_{\text{eff}}}{\sqrt{2\lambda(\lambda + \widehat{a}_{\text{eff}})}}$ and eigenvalues

$$E^{\text{fb}}_{\pm} = |\vec{p}| + \frac{|m_l|^2}{2|\vec{p}|} - \frac{\widehat{c}_{\text{eff}}}{|\vec{p}|} \pm \frac{\lambda}{|\vec{p}|}. \tag{3.13}$$

The expanded form of the effective coefficients are [10]:

$$\begin{aligned} \widehat{a}_{\text{eff}} &= \textstyle\sum_{djm} |\vec{p}|^{d-2} Y_{jm}(\hat{p}) \left(a^{(d)}_{\text{fb}}\right)_{jm} \\ \widehat{c}_{\text{eff}} &= \textstyle\sum_{djm} |\vec{p}|^{d-2} Y_{jm}(\hat{p}) \left(c^{(d)}_{\text{fb}}\right)_{jm} \\ \hat{g}_{\text{eff}} &= \textstyle\sum_{djm} |\vec{p}|^{d-2}{}_{+1} Y_{jm}(\hat{p}) \left(g^{(d)}_{\text{fb}}\right)_{jm} \end{aligned} \quad , \tag{3.14}$$

where

$$\begin{aligned} \left(a^{(d)}_{\text{fb}}\right)^{*}_{jm} &= (-1)^m \left(a^{(d)}_{\text{fb}}\right)_{j(-m)} \\ \left(c^{(d)}_{\text{fb}}\right)^{*}_{jm} &= (-1)^m \left(c^{(d)}_{\text{fb}}\right)_{j(-m)}. \end{aligned} \tag{3.15}$$

In the framework of a flavor-blind and oscillation-free model there is no mixing between the generations which implies $\left(g^{(d)}_{\text{eff}}\right) = 0$ and results in the dispersion relation [10]:

$$E^{\text{of}}_{\nu} = |\vec{p}| + \frac{|m_l|^2}{2|\vec{p}|} + \sum_{djm} |\vec{p}|^{d-3} Y_{jm}(\hat{p}) \left[\left(a^{(d)}_{\text{of}}\right)_{jm} - \left(c^{(d)}_{\text{of}}\right)_{jm}\right] \quad . \tag{3.16}$$

It is worth mentioning that for $d = 3$, the operator $\left(c^{(d)}_{\text{of}}\right)_{jm}$ does not exist and this equation modifies to:

$$E^{\text{of}}_{\nu} = |\vec{p}| + \frac{|m_l|^2}{2|\vec{p}|} + \sum_{djm} Y_{jm}(\hat{p}) \left(a^{(3)}_{\text{of}}\right)_{jm} \quad . \tag{3.17}$$

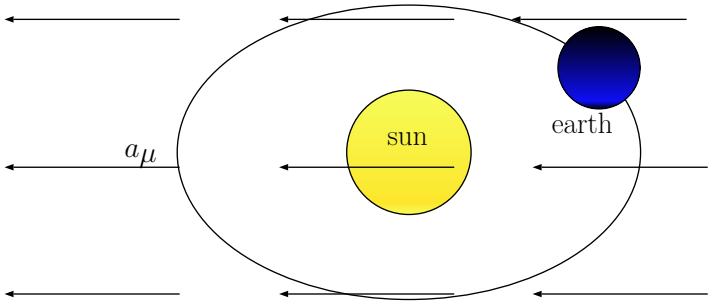


Figure 3.1: This figure illustrates the a^μ-type LV theory. It could be understood as an background field, in which the earth is moving and rotating.

3.2 Measuring neutrino operators using electrons

The previous section gave an insight into how the introduction of LV can modify the wave-function of the particles. Of course, LV does not only play a role in non-interacting fermions but influences also the interaction processes. This section will give a brief explanation of how this fact can be used to limit neutrinos' LV parameters using the spectrum of the electrons of a β-decay.

3.2.1 General idea

The KATRIN experiment is a neutrino experiment. However, it measures the spectrum of the emitted electrons in the beta-decay of Tritium. Furthermore it is known that the angular correlation of electrons and neutrinos in the neutron decay is weak [36]. Therefore a natural question is, why KATRIN can be sensitive to Lorentz violation in the neutrino sector. In the following, a general explanation based on a private discussion with Ralf Lehnert is given to make the search plausible before the next section will deal with the theoretical framework in more depth.
As many LV operators of the minimal SM extension have been limited to be small [37], this thesis focuses on an effective theory in which all LV operators other than the neutrino's a^μ are treated as zero. As it is visible from eq. (3.2), this LV can be treated as an a^μ-type which is a modification of the Lagrangian for spin-$\frac{1}{2}$ fermions

$$\delta L^a_{\mathrm{SME}} = -\bar{\psi} a^\mu \gamma_\mu \psi. \tag{3.18}$$

This is equivalent to a term arising from a background vector field a^μ (see fig. 3.1) that changes the dispersion relation but also the wave function of the neutrinos [38]. As the LV is quadratic in the fields the effects can be seen as affecting the free propagation, which changes the external legs in the calculation of the Feynman diagram. If one, therefore, wants to calculate the decay rate of the Tritium-decay one observes a modified matrix element δM due to the LV. Thus the decay rate is

$$d\Gamma \propto [M_0 + \delta M]^2 \, dp^3 dk^3 dq^3 \tag{3.19}$$

where p, k, q are the momenta of the electron, neutrino, and nucleus respectively.
If one does the calculation in general one can receive terms like

$$\Gamma \propto \int \langle T|H|e\nu\mathrm{He}|T|H|e\nu\mathrm{He}\rangle \, d\Pi \propto \int [\alpha a^\mu p_\mu + \beta a^\mu k_\mu + \gamma a^\mu q_\mu] \, d\Pi, \tag{3.20}$$

where α, β, γ are factors that have to be determined in a detailed calculation and Π is the phase space. Note that due to the scalar transformation of the decay rate $a^\mu p_\mu$, $a^\mu k_\mu$, $a^\mu q_\mu$

are the only choice for relevant terms including LV to first order.
Therefore isotropic nucleus and neutrino momentum distributions make the second and third term vanish. Due to the finite acceptance angle of the electrons ($\theta_0 = 51° < 180°$) there is still a term which is proportional to the LV parameter and therefore enables the KATRIN experiment to measure the effects of the neutrino's LV parameter.
In general the effect of LV consists of an isotropic part a^0 and another (partly) anisotropic part $\vec{a}$. In laboratories where the Tritium is approximately at rest, the four-momenta of Tritium and the electron are largely isotropic. Thus the effect of the anisotropic part is suppressed compared to the isotropic part. The exact value of the suppression is currently recalculated by Ralf Lehnert and myself. The values of the effect, which were e.g. stated in [9] made, inter alia, use of assumptions on the correlation between the momenta of the electrons and neutrinos. Those are not applicable for a Tritium decay like in KATRIN, which will be explained in the next section in more detail.

3.3 LV model for the KATRIN experiment

In the previous section a brief overview about LV was presented as well as an explanation why an experiment like KATRIN, that measures electrons from a β-decay, can be sensitive to LV in the neutrino sector. The following section will focus on the appearance of a possible signal and the model which will be used in the context of examining effects of LV at KATRIN.

3.3.1 Crosscheck of current theory

However, before going into the illustration of a possible signal, one has to develop the theoretical prediction, which describes the effects of LV on the Tritium spectrum. Therefore this section focuses on currently published models, their assumptions and discusses the applicability for the KATRIN experiment.
In 2013 Díaz, Kostelecký and Lehnert published a prediction how a^μ-type LV would influence the endpoint of the Tritium spectrum [9]. To determine a limit on the "countershaded" anisotropic oscillation-free mode they suggest making use of the acceptance angle in Tritium-decay experiments which use a MAC-E filter. Within their paper the argue that LV modifies the neutrino four-momentum $q^\mu = (\omega, \vec{q})$ and therefore the neutrino's phase space $d^3q = f(\omega)d\omega d\Omega$, where the neutrino function is [9]

$$f(\omega) \approx \omega^2 - \frac{1}{2}m_\nu^2 - 2\omega\delta\omega. \tag{3.21}$$

The term $\delta\omega$ contains LV operators, which can be described by [9]

$$\delta\omega = -\sum_{jm} e^{im\omega_\oplus T_\oplus}\mathcal{N}_{jm}\left(a_{\text{of}}^{(3)}\right)_{jm} \tag{3.22}$$

where $\omega_\oplus$ and $T_\oplus$ is the sidereal frequency and time. The components of the three-dimensional oscillation-free mode in spherical decomposition are noted as $\left(a_{\text{of}}^{(3)}\right)_{jm}$. The factor $\mathcal{N}_{jm}$ contains the direction of the neutrinos and the location of the experiment relative to the sun-centered celestial equatorial frame [9].
In general, the differential Tritium-spectrum close to the endpoint can be described by [39]

$$\frac{d\Gamma}{dT} = C\left[(\Delta E)^2 - \frac{1}{2}m_\nu^2\right] \tag{3.23}$$

where C is approximately constant, m_ν is the neutrino mass and $\Delta E = E_0 - E$ describes the difference of the energy of the electron to the endpoint of the spectrum. Díaz argues further that the endpoint gets modified due to LV $E_0 \to E_0 + \delta E$ with [39]

$$\delta E = \frac{1}{\Delta\Omega}\int_{\Delta\Omega} d\Omega_{\bar{\nu}} \left(a_{\text{of}}^{(3)}\right)^{\text{lab}}. \tag{3.24}$$

Starting from this the paper uses the acceptance angle of the experiment to calculate δE to be [9]

$$\begin{aligned}\delta E =& \frac{1}{\sqrt{4\pi}}\left(a_{\text{of}}^{(3)}\right)_{00} - \sqrt{\frac{3}{4\pi}}\cos^2\frac{1}{2}\theta_0 \sin\chi\cos\xi\left(a_{\text{of}}^{(3)}\right)_{10} \\ & - \sqrt{\frac{3}{2\pi}}\cos^2\frac{1}{2}\theta_0\left[\sin\xi\,\text{Im}\left(\left(a_{\text{of}}^{(3)}\right)_{11} e^{i\omega_\oplus T_\oplus}\right)\right. \\ & \left. + \cos\chi\cos\xi\,\text{Re}\left(\left(a_{\text{of}}^{(3)}\right)^*_{11} e^{-i\omega_\oplus T_\oplus}\right)\right]\end{aligned} \tag{3.25}$$

where θ_0 is the aperture of the acceptance cone, χ the colatitude of the experiment and ξ the angle of the magnetic field at the source measured east of local north. This formula shows a global shift due to $\left(a_{\text{of}}^{(3)}\right)_{00}$ and $\left(a_{\text{of}}^{(3)}\right)_{10}$ as well as an oscillation caused by $\left(a_{\text{of}}^{(3)}\right)_{11}$. They therefore suggest to search for an oscillation of the endpoint to limit the anisotropic $\left(a_{\text{of}}^{(3)}\right)_{11}$. By doing so, however, they assume that the acceptance angle which maps a part of the electrons of the decay to the detector is the same as the one for the neutrinos. Thus they imply a directional correlation between electron and neutrino. Yet, it is known that the angular correlation of electrons and (anti-)neutrinos in the Tritium decay is similar to the one of neutron decay, and therefore only very weakly correlated [36, 40]. A more appropriate way would be to argue that the neutrinos are distributed isotropically. This, however implies $\theta_0 = \pi$ which vanishes all terms but $\left(a_{\text{of}}^{(3)}\right)_{00}$ in eq. (3.25).

3.3.2 New model for LV at KATRIN

Thus, such an approach is not applicable for the KATRIN experiment. To solve this issue, we collaborated with Ralf Lehnert to create a model, that is suitable for an experiment like KATRIN. In this context the beta-spectrum including a^μ-type LV was derived [41].
For doing so one assumes that the Lagrangian gets modified by the term

$$\delta L^a_{\text{SME}} = -\bar{\psi} a^\mu \gamma_\mu \psi \tag{3.26}$$

which is quadratic in the spinors and therefore modifies the free propagation [41].
In general the a^μ-type LV can be applied for all fermions in the Tritium decay process. Thus, one expects a_T^μ, a_H^μ, a_e^μ and a_n^μ for the Tritium, the Helium, the electron and the neutrino respectively. Via redefinitions of the spinors one can "move" the a^μ-type LV from one species to the other. Therefore one expects the final spectrum not to be dependent on a_T^μ, a_H^μ, a_e^μ and a_n^μ individually, but rather on a sum of those [41].
To calculate the spectrum one can follow the well-known procedure of Feynman rules with a slight modification of the external legs due to LV

$$\begin{aligned} 0 &= \left[\left(k'^\mu - a_T^\mu\right)\gamma_\mu - M_T\right)\right] u_T'^{(s)}\left(k'\right), \\ 0 &= \left[\left(l'^\mu - a_H^\mu\right)\gamma_\mu - M_H\right)\right] u_H'^{(s)}\left(l'\right) \\ 0 &= \left[\left(p'^\mu - a_e^\mu\right)\gamma_\mu - m_e\right)\right] u_e'^{(s)}\left(p'\right) \\ 0 &= \left[\left(q'^\mu + a_n^\mu\right)\gamma_\mu + m_n\right)\right] v_n'^{(s)}\left(q'\right) \end{aligned} \tag{3.27}$$

where k'^{μ}, l'^{μ}, k'^{μ}, p'^{μ} and q'^{μ} are the physical momenta of the Tritium, Helium, electron and neutrino [41].
By carefully applying the modified Feynman rules the decay rate can be calculated to be [41]

$$\frac{d\Gamma}{dE_e} = \frac{\left(G_F V_{ud} g_V\right)^2}{(2\pi)^5 E_T(k)} |\vec{p}| \int_{\Delta\Omega} d\Omega_e \underbrace{\int \frac{d^3q}{E_n(q)} \frac{d^3l}{E_H(l)} \delta^{(4)}(k+a-l-p-q) \frac{|\bar{M}(k,l,p,q)|^2}{\left(4G_F V_{ud} g_V\right)^2}}_{\equiv R} . \quad (3.28)$$

The decay rate depends on the momenta of the particles and the sum of the a^{μ}, namely $a^{\mu} \equiv a_T^{\mu} - a_H^{\mu} - a_e^{\mu} + a_n^{\mu}$. Using the fact, that R is a scalar, it becomes clear, that it can only contain terms like $k \cdot p, k \cdot a,$ and $p \cdot a$ [41].
This enables the calculation of R and therefore the decay rate. As an expansion of the spectrum close to the endpoint is difficult since the spectrum is not defined beyond the endpoint, it makes sense to expand the endpoint in terms of LV a^{μ}. This leads to [41]

$$E_m^a(\hat{p}) \simeq E_m + \left(1 - \frac{E_m}{M_T}\right) a^0 + \frac{\sqrt{E_m^2 - m_e^2}}{M_T} \hat{p} \cdot \vec{a} \quad (3.29)$$

where E_m is the conventional endpoint without LV. This equation makes it obvious that the endpoint has an isotropic as well as an anisotropic part. The latter changes under rotation of the reference frame. To have a final formula for the KATRIN experiment, eq. (3.29) has to be "averaged" over the acceptance cone, as KATRIN measures all electrons with a direction within this region. However, since the acceptance angle is smaller than 180° the averaged effect stays larger than zero.

$$\begin{aligned} \int_{\Delta\Omega} d\Omega_e F(\vec{p} \cdot \vec{a}) &\simeq \Omega_e F(0) + F'(0) \int_{\Delta\Omega} d\Omega_e (\vec{p} \cdot \vec{a}) \\ &= \Omega_e F(0) + F'(0) \pi (\hat{z} \cdot \vec{a}) |\vec{p}| \sin^2 \kappa \\ &\simeq \Omega_e F\left(\Omega_e^{-1} \pi |\vec{p}| \hat{z} \cdot \vec{a} \sin^2 \kappa\right) . \end{aligned} \quad (3.30)$$

In this equation $\Omega_e \equiv 2\pi(1 - \cos\kappa)$ with κ the angle that defines the cone in $\hat{z}$-direction. Therefore eq. (3.28) gets a global factor $\Delta\Omega$ and the electron momentum $\vec{p}$ gets replaced by an "average" $\bar{p}\hat{z} \equiv \Omega_e^{-1} \pi |\vec{p}| \hat{z} \sin^2 \kappa$. Thus the change of the endpoint becomes [41]

$$\bar{E}_m^a(\hat{z}) \equiv E_m + a^0 + M_T^{-1} \Omega_e^{-1} \pi \sqrt{E_m^2 - m_e^2} \hat{z} \cdot \vec{a} \sin^2 \kappa. \quad (3.31)$$

This formula is in the laboratory frame. However, coordinate-system dependent quantities like vectors should be expressed in a reference frame. For LV one usually transforms the equations into the sun-centered celestial equatorial frame [9]. The endpoint difference consists of two reference frame dependent parts

$$\Delta E_m = a'^0 + M_T^{-1} \Omega_e^{-1} \pi \sqrt{E_m^2 - m_e^2} \hat{z} \cdot \vec{a'} \sin^2 \kappa \quad (3.32)$$

where the prime marks the laboratory frame. In general a'^0 is boosted by the rotation velocity with which the KATRIN experiment is rotated by the earth.

$$a'^0 = \gamma a^0 - \gamma \vec{\beta}_{rot} \cdot \vec{a} \quad (3.33)$$

The rotation velocity in the sun-centered frame changes its direction with the rotation of earth. Applying the rotation matrix on $\vec{\beta}_{rot}$ one receives

$$a'^0 = \gamma a^0 - \gamma \left|\vec{\beta}_{rot}\right| \left[\cos(\omega_{\oplus} T_{\oplus}) a^2 - \sin(\omega_{\oplus} T_{\oplus}) a^1\right] \quad (3.34)$$

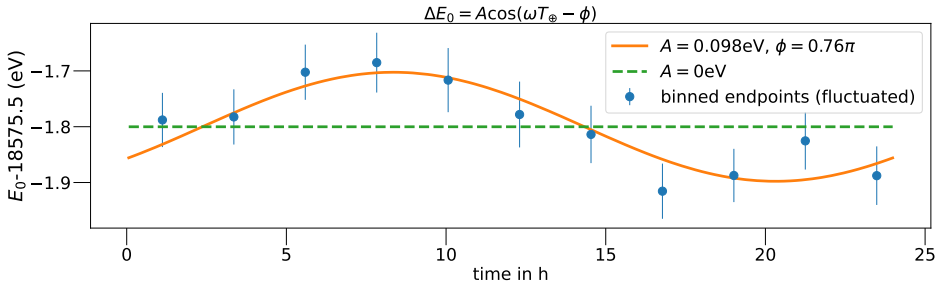


Figure 3.2: This figure shows fluctuated endpoints fitted from MC data using an amplitude $A = 0.098\,\text{eV}$ and phase $\phi = 0.76\pi$ in blue. The endpoints are binned using a binning of 2 h for illustration. The orange line shows the predicted endpoints for the same LV.

where $\omega_\oplus$ is the sidereal frequency and $T_\oplus$ sidereal time.
For the term $\hat{p}\vec{a}'$, the vector $\vec{a}$ is transformed (from the sun-centered frame) into a laboratory frame with the x-axis pointing south, the y-axis pointing east, and the z-axis pointing vertically upwards. This is done with the transformation matrix given in [42]:

$$\vec{a} \to \begin{pmatrix} \cos\chi\cos\omega_\oplus T_\oplus & \cos\chi\sin\omega_\oplus T_\oplus & -\sin\chi \\ -\sin\omega_\oplus T_\oplus & \cos\omega_\oplus T_\oplus & 0 \\ \sin\chi\cos\omega_\oplus T_\oplus & \sin\chi\sin\omega_\oplus T_\oplus & \cos\chi \end{pmatrix} \vec{a} \tag{3.35}$$

The vector $\vec{z}$ in the laboratory frame can be transformed into this coordinate system via

$$\hat{z} \to \begin{pmatrix} -\cos(\xi) \\ \sin(\xi) \\ 0 \end{pmatrix} \tag{3.36}$$

where ξ is the angle of the experiment east of the local north. Therefore the change of the endpoint becomes

$$\begin{aligned} \Delta E_0 =& \gamma a^0 + B\sin\chi\cos\xi a^3 + \cos(\omega_\oplus T_\oplus)\left[(-\beta_{rot} + B\sin\xi)a^2 - B\cos\xi\cos\chi a^1\right] \\ &+ \sin(\omega_\oplus T_\oplus)\left[(\beta_{rot} - B\sin\xi)a^1 - B\cos\xi\cos\chi a^2\right] \end{aligned} \tag{3.37}$$

with $\beta_{rot} \approx \gamma\left|\vec{\beta}_{rot}\right|$ and $B = M_T^{-1}(2\pi(1-\cos\kappa))^{-1}\pi\sqrt{E_m^2 - m_e^2}\sin^2\kappa$. However, this calculation is work-in-progress while writing this thesis and the correct prefactors of of a^μ have to be taken with caution. Therefore the following section will discuss a more generic model.

3.4 Possible signal at the KATRIN experiment

As shown before, the effects of LV on the KATRIN experiment can be classified into effects on the endpoint of the β-spectrum caused by the isotropic or the anisotropic part.

3.4.1 Isotropic effects

The isotropic part of a^μ consists of a^0 but also of a component of $\vec{a}$, namely [9]

$$a_{\text{isotropic}} \propto \cos\xi\sin\chi\, a^3 \tag{3.38}$$

where ξ is the angle relative to east of local north and χ the co-latitude of the experiment [41]. The values for KATRIN can be found in table B.1. It is important to remember that

$a_{\text{isotropic}}$ is suppressed relative to a^0.
The two parameters are independent of the rotation of the earth. A measurement of the endpoint using all runs combined corresponds to the average over time. Therefore the oscillation part vanishes and the time-averaged endpoint change becomes

$$\langle \Delta E_0 \rangle_t = k \times a^0 + l \times \cos\xi \sin\chi a^3 \tag{3.39}$$

where k and l are the prefactors according to the LV theory.
Therefore the parameters a^0 and a^3 introduce a global shift of the endpoint. In order to estimate those parameters the endpoint relative to its theoretical value will be analyzed.

Calculation of an experimental independent endpoint

The value of the fitted endpoint depends on the experiment, e.g. on the electric potential present in the gaseous tritium source. A way to find a value that is comparable with the theory is the Q-value, the energy released in the Tritium-beta decay. It is e.g. dependent on the mass difference between ^{3}H and ^{3}He, the recoil energy, the work function $\Delta\Phi$ of the spectrometer and the source potential. The Q-value of the KATRIN experiment can be related to the endpoint by [43]

$$Q = E_0 + 1.720\,\text{eV}(+\Delta\Phi) \tag{3.40}$$

This formula makes clear that a global shift of the endpoint does also shift the Q-value, which is independent of the experiment. Thus the idea of this analysis is to compare the measured Q-value of KATRIN with its theoretical value and interpret possible shifts as a result of the isotropic effects of LV.

Solution for degeneracy

Yet, the KNM-1 data provides only one (time-averaged) Q-value that can be compared to the theoretical one. But LV suggests two parameters that can introduce a shift. To solve this degeneracy one needs a second experimental Q-value or endpoint. Furthermore it is important to mention that one cannot just split the data-set of KATRIN to have multiple averaged Q-values. This has two main reasons. Firstly, the assumption of a time-averaged Q-value becomes worse the fewer runs at different times are performed. Secondly, the degeneracy can only be solved if the prefactors in eq. (3.39) differ. Therefore one should use the Q-value of different experiments that are located at different coordinates. In section 6.2 the results of KATRIN's predecessor in Mainz will be used together with the results of KNM-1 to limit the isotropic parts of a^μ.

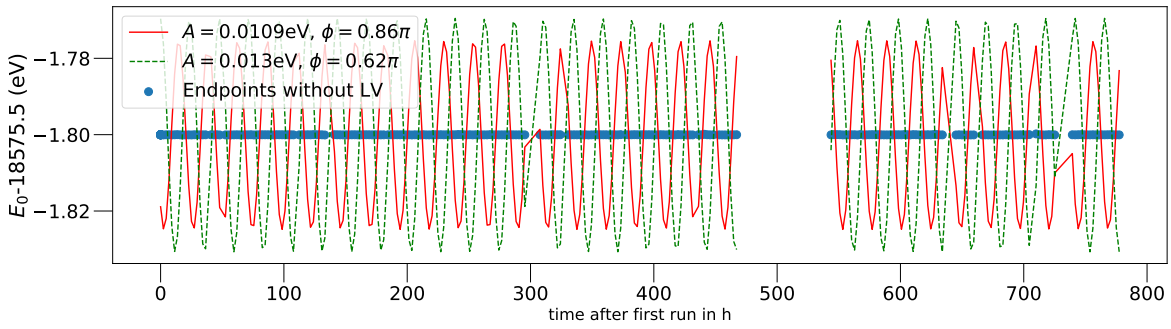
3.4.2 Non-isotropic effects

As the KATRIN experiment performs runs with a duration of about 2 h, it has also a temporal resolution of the endpoint and one can also examine time-dependent changes due to LV. There are 3 major effects of the a^μ-theory that are time-dependent.
Firstly, the earth is moving around the sun once a year and therefore performs a rotation within the field a^μ in this time. Therefore one expects to see an oscillation of the endpoint with a period of one year due to the anisotropic $\vec{a}$. However, the KNM-1 runs were performed within a few weeks. Therefore this single measurement phase is not sensitive to such effects [41].
Secondly, the earth is rotating around its own axis once a day (23 h 56 min). Thus one expects to see an oscillation due to the anisotropic part of a^μ with a period of one day. As KNM-1

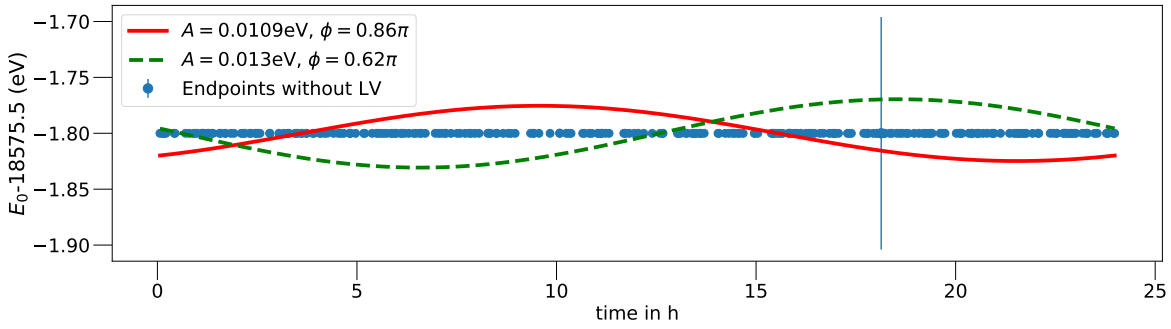
was performed with 2 h runs, KATRIN is sensitive to such effects.
Thirdly, the experiment is also moving with a finite velocity. Therefore also boost-effects of a^μ in the laboratory frame are expected. The main boost is due to the rotation velocity of the earth. KATRIN is not located at the poles and therefore experiences a velocity which is of the order $1 \times 10^{-6}c$. As the earth is rotating and therefore the direction of the velocity, that causes the boost, changes, one expects also an oscillatory effect proportional to the isotropic part a^0. This effect is suppressed by approximately 1×10^{-6} and therefore often neglected. However, depending on the relative suppression between the isotropic and the anisotropic part in the KATRIN experiment, the oscillation effects due to the boost can be relevant. Nevertheless, both cause oscillations with a fixed frequency. Therefore one expects an oscillation of the endpoint of the Tritium spectrum with a period of 23 h 56 min. Since the exact value of the suppression of the anisotropic part is not finally calculated at the time of writing, this thesis will analyze a general oscillation, which is described by

$$\Delta E_0 = A \cos(\omega T_\oplus - \phi) \tag{3.41}$$

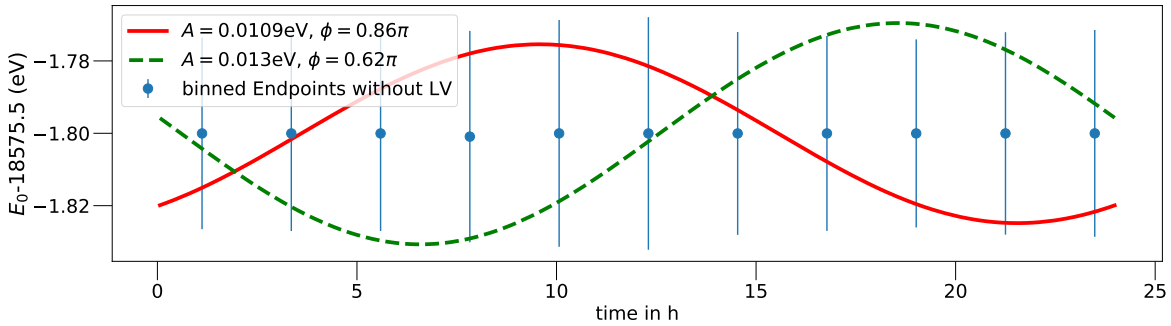
where A describes the amplitude of the oscillation and ϕ its phase. ω is the sidereal frequency with the corresponding sidereal time $T_\oplus$ (c.f. table B.1). A possible signal is illustrated in fig. 3.2.
In the analysis, we have a look at the (Monte Carlo) data of KNM-1 with the corresponding time of the runs (c.f. fig. 3.3a). For the actual analysis, we use a data set where the endpoints of the runs are plotted corresponding to their sidereal starting time (modulo one day) as in fig. 3.3b. The model eq. (3.41) is used to fit the amplitude and phase of the oscillation and therefore determine or limit the size of A and ϕ. The next section will discuss the connection of A and ϕ to LV operators.



(a)



(b)



(c)

Figure 3.3: **(a)** This figure shows the endpoints of the "golden-runs" of KNM-1 (Monte Carlo) in blue drawn at their starting time relative to the first run of this run list. The red straight line indicates the theoretical predicted endpoint for A =0.0109 eV and $\phi = 0.86\pi$ at each time while the dashed green line shows the prediction for A =0.013 eV and $\phi = 0.62\pi$.
(b) This plot illustrates that different real and imaginary parts of the operator correspond to a different phase and amplitude. The endpoints of the "golden-runs" of KNM-1 (Monte Carlo) are drawn in blue at their sidereal starting time modulo one day. For one data point in blue the uncertainty is indicated by the blue line. The analysis was done in that way - using sidereal time.
(c) This figure shows the binned endpoints of the "golden-runs" of KNM-1 (Monte Carlo) in blue drawn at their sidereal starting time modulo one day. The binning corresponds to the average duration of a run. The combined uncertainty is shown as the blue lines. The binning is only used for illustration purposes.

3.5 Model parameters

In general the effects of LV, that cause an oscillation can be described by

$$\Delta E_0 = a\sin(\omega T_\oplus) + b\cos(\omega T_\oplus) \tag{3.42}$$

where the amplitudes a, b depend on the LV model. However, the sum of a sine with a cosine can be rewritten as:

$$a\sin(\omega T_\oplus) + b\cos(\omega T_\oplus) = A\cos(\omega T_\oplus - \phi) \tag{3.43}$$

where A=$\sqrt{a^2+b^2}$ and $\phi = \arctan(a/b)$. Thus by searching for an oscillation with frequency ω one can limit the LV parameters.
Hereby, one has two natural sets of fit parameters, that one can choose. Firstly, one can use the prefactors of the sine and cosine a and b (in the following called "two-amplitude" description). However the signal one would see in the KATRIN experiment is an oscillation with an amplitude A and phase ϕ. Thus the approach using this description is more illustrative and therefore used in the following.
Applying the preliminary calculation of eq. (3.37) we see the following relations between a, b and a^1, a^2:

$$\begin{aligned} a =&(\beta_{rot} - B\sin\xi)a^1 - B\cos\xi\cos\chi a^2 \\ b =&(-\beta_{rot} + B\sin\xi)a^2 - B\cos\xi\cos\chi a^1 \end{aligned} \tag{3.44}$$

Therefore the amplitude of the oscillation is given by

$$A = \sqrt{a_1^2 + a_2^2}\sqrt{B^2\cos^2\chi\cos^2\xi + (\beta_{rot} - B\sin\xi)^2} \tag{3.45}$$

where for the sake of notation the Lorentz indices are written as subscripts. Sometimes the LV-operators are expressed and cited in spherical decomposition [10, 37]. For better comparison with other LV operators, the transformation is performed with the preliminary prefactors derived in the section above. In spherical decomposition the components can be written as [10]

$$\begin{aligned} a^0 =&\frac{1}{\sqrt{4\pi}}a_{00}^{(3)} \\ a^1 =&\sqrt{\frac{3}{2\pi}}\,\mathrm{Re}\left(a_{11}^{(3)}\right) \\ a^2 =&-\sqrt{\frac{3}{2\pi}}\,\mathrm{Im}\left(a_{11}^{(3)}\right) \\ a^3 =&-\sqrt{\frac{3}{4\pi}}a_{10}^{(3)}. \end{aligned} \tag{3.46}$$

One can trivially see that the amplitude is proportional to the modulus of $a_{11}^{(3)}$:

$$A = \sqrt{\frac{3}{2\pi}}\left|a_{11}^{(3)}\right|\sqrt{B^2\cos^2\chi\cos^2\xi + (\beta_{rot} - B\sin\xi)^2} \tag{3.47}$$

Therefore the limit of the amplitude A can be used to derive a first limit on $a_{11}^{(3)}$. In spherical decomposition the general form of the modified endpoint due to LV is

$$\begin{aligned} \Delta E_0 =&\gamma\frac{1}{\sqrt{4\pi}}a_{00}^{(3)} - \sqrt{\frac{3}{4\pi}}B\sin\chi\cos\xi a_{10}^{(3)} \\ &+ \frac{3}{\sqrt{2\pi}}\cos(\omega_\oplus T_\oplus)\left[(\beta_{rot} - B\sin\xi)\,\mathrm{Im}\left(a_{11}^{(3)}\right) - B\cos\xi\cos\chi\,\mathrm{Re}\left(a_{11}^{(3)}\right)\right] \\ &+ \frac{3}{\sqrt{2\pi}}\sin(\omega_\oplus T_\oplus)\left[(\beta_{rot} - B\sin\xi)\,\mathrm{Re}\left(a_{11}^{(3)}\right) + B\cos\xi\cos\chi\,\mathrm{Im}\left(a_{11}^{(3)}\right)\right]. \end{aligned} \tag{3.48}$$

Chapter 4

Analysis strategy

To avoid biases the KATRIN collaboration relies on a blinding strategy. Besides other measures, Monte-Carlo data is used to establish the analysis procedure and to determine the sensitivity before using real data. In this procedure, it is very important to look at systematic uncertainties and check for statistical problems in advance [27]. Following this spirit, this chapter focuses on the analysis flow used to find limits on A and ϕ (c.f. fig. 4.1). The general idea is to fit an oscillation to the KNM-1 (Monte Carlo) data and to determine the value of the LV parameter A and ϕ using eq. (3.41). Therefore the data is fitted with FITRIUM to receive a set of endpoints at different times. Afterward, this set is compared to the theoretical model eq. (3.41). The methods used are a frequentist χ^2-grid-search and a Bayesian MCMC, which will be discussed below.

4.1 Model and effective time

Before coming to the analysis methods it is important to define the model which should be fitted to the data. As discussed in section 3.4 the expected model is an oscillation of the endpoint that is generally described by

$$E_0 = C + A\cos(\omega T_\oplus - \phi). \tag{4.1}$$

where C is the averaged endpoint and therefore the baseline of the oscillation.
The goal of the analysis is to limit the parameters A and ϕ. The averaged endpoint is hereby fitted as a nuisance parameter. Furthermore it is crucial to know the time $T_\oplus$ of the fitted endpoint.
In the previous section, the endpoints were plotted at the starting time of the corresponding run. But the endpoint, that is fitted, is the endpoint of a whole run with a finite duration of approximately two hours. The naive approach to use a non-weighted average over the duration of a run as interpretation of the fitted endpoint was discarded since one expects different points in the spectrum to have a different impact on the fitted endpoint.
The most accurate way of taking this effect into account would be to include the Lorentz violation at the level of the spectral fit. However, this is rather time-consuming and would not lead to a significant improvement in the sensitivity. Instead, we choose an approach based on an "effective time" for each endpoint:
To have data to compare with, runs for different A and ϕ were simulated with an endpoint according to the Lorentz violating theory (eq. (3.41)) for each subrun. Furthermore, a broadening due to the finite duration of the subruns was included[1]. Afterward these simulated runs

[1]The broadening is $\frac{1}{t_2-t_1}\int_{t_1}^{t_2} dt(\Delta E_0)^2$, where t_1, t_2 are the start and the end of the subrun, while ΔE_0 is the difference of the endpoint due to LV according to eq. (3.41).

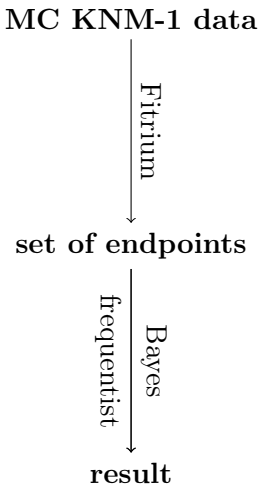


Figure 4.1: In this illustration, the analysis flow is presented. The MC data of KNM-1 is fitted using the FITRIUM software which generates a set of endpoints with corresponding times. This set is then used to limit A and ϕ of the oscillation. To do so a Bayesian approach as well as a frequentist one is used.

were fitted with FITRIUM to get a set of endpoints like one would find it in the KATRIN data assuming this given LV operator components.
Looking at this data it seems to be a good approximation to describe the fitted endpoint by the theoretical endpoint at an "effective time", which is shifted from the start of the run (c.f. fig. 4.2). This effective time is then fitted using the generated data[2].
Thus the model, that will be compared with the fitted endpoints in order to determine the size of A and ϕ, is the theoretical equation eq. (3.41) at this effective time. In this context, the uncertainty of the effective time is considered as the width of a Gaussian pull-term. Its size will be discussed in section 5.1.2.

4.2 FITRIUM analysis

The FITRIUM software is used to fit the endpoints of the KMN-1 Monte Carlo data (without LV). The statistics of one single run are low. We, therefore, fixed the neutrino mass to zero and performed a uniform fit. Hereby the counts of all pixels are added and fitted with an averaged model [18]. Especially for the MC data with zero neutrino mass, this is a good approximation but also reasonable for the real data. Additionally the correlations between the normalization and the fitted endpoint were marginalized. A study showed that the discrepancy between the endpoints using a pull-term and using a fixed neutrino mass is small. Therefore we stick to the analysis using a Tritium-decay model with zero neutrino mass.

4.3 Frequentist analysis

After creating a set of endpoints with corresponding times a χ^2-grid-search with 80×80 grid points is performed to limit A and ϕ. For each grid point (A, ϕ) the χ^2 is minimized with the nuisance parameters average endpoint and effective time. The latter is included with a

[2] As the measurement time distribution (MTD) of up and down scans differ, they are treated differently. This means there is an effective time for up scans and an effective time for down scans.

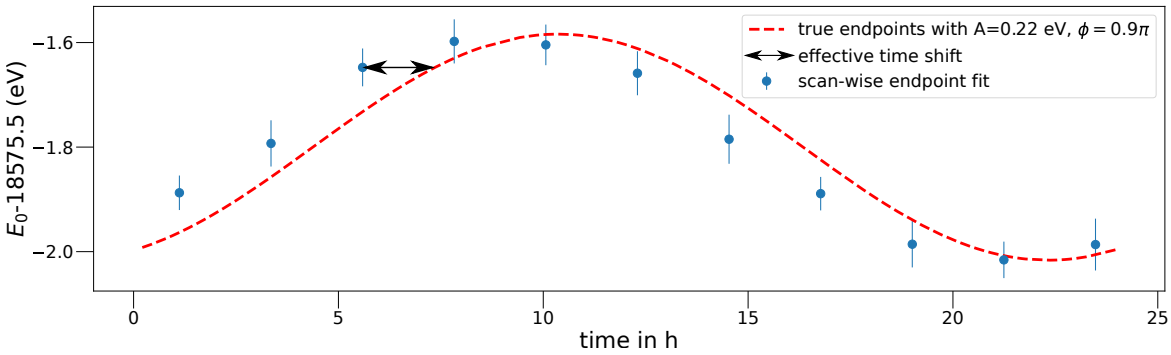


Figure 4.2: This figure shows the endpoints of up-scans of the "golden-runs" of KNM-1 (Asimov) with A =0.22 eV and $\phi = 0.9\pi$ in blue. For better visualization they are binned using a binning of 2 h and depicted at their sidereal starting time modulo one day. The red dashed line is the theoretical prediction for the endpoint using the starting time. By comparing the blue endpoints with the red dashed line one can already guess that they would fit well if they were not evaluated at the starting time but at an effective time. The black arrow indicates the effective time, which is shifted by 87 min from the starting time of the runs. The value of 87 min is the result of a fit using the blue endpoints and a free effective time in the theoretical model with A =0.22 eV and $\phi = 0.9\pi$. Due to their different measurement time distribution (MTD) up- and down-scans are treated separately.

pull-term with its uncertainty which is described in section 5.1.2.
Afterward, the confidence limits are found by using a likelihood-ratio as test statistic,

$$\Lambda\left(A,\phi\right) = \frac{\mathcal{L}\left(A,\phi\right)}{\mathcal{L}_{\text{best}}}. \tag{4.2}$$

Applying Wilks' theorem

$$\Delta\chi^2 \equiv -2\ln\Lambda\left(A,\phi\right) = \chi^2\left(A,\phi\right) - \chi^2(\text{best fit}) \tag{4.3}$$

follows the χ^2-distribution for two degrees of freedom [44]. Thus grid points that have a $\Delta\chi^2 > \chi^2_c$ are excluded. The critical $\chi^2_c = 4.61$ for 90% C.L. for two degrees of freedom [45].

4.4 Coverage study

Wilks' theorem is only applicable under specific conditions and it is known to become invalid for values close to boundaries. Thus it is important to check if one can use Wilks' theorem to set an exclusion curve. Therefore you have to look at the seeming boundaries. It might be suspected that a fit that leads to a negative A is not physically meaningful and one could try to impose a boundary at zero amplitude. However, if you have a look at the description of the LV theory you can easily notice that a fit result with a negative amplitude is the same as the one with the positive amplitude and a phase, that is shifted by π due to the periodicity[3]. This means that there is no physically motivated boundary present. Thus as long as the distribution of the best fits follows the χ^2-distribution as suggested by Wilks' theorem, the critical $\chi^2_c = 4.61$ can be used to set an exclusion curve.
To test if Wilks' theorem is applicable for this analysis, we varied the endpoints within their

[3]Consider the example $(A,\phi) = (-1,0)$. The LV theory says this is $-\cos(\omega T_\oplus) = \cos(\omega T_\oplus + \pi)$. Thus $(A,\phi) = (-1,0) \cong (1,\pi)$. The negative amplitude is not nonphysical but another description.

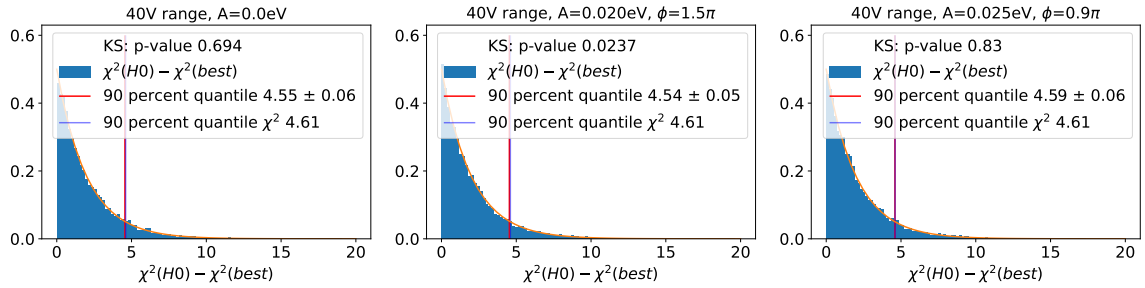


Figure 4.3: In this figure you can see the histogram of $\chi^2(A, \phi) - \chi^2$(best fit) in blue bars. The orange line shows the theoretical χ^2 distribution for two degrees of freedom. Its 90 % quantile is indicated by the purple line at approximately 4.61. The 90 % quantile of the histogram data is shown by the red straight line. The indicated uncertainties correspond to bootstrap errors. For all of these exemplary values of LV Wilks' theorem holds. This can also be seen by the Kolmogorov–Smirnov test (KS) which suggests a good agreement between the χ^2 distribution and the distribution of $\chi^2(A, \phi) - \chi^2$(best fit). Therefore we conclude that usage of Wilks' theorem is acceptable.

uncertainty to create a set of fluctuated data ($\sim$ 5000 samples). In the next step we plotted the histogram of the difference

$$\chi^2(A, \phi) - \chi^2(\text{best fit}) \tag{4.4}$$

and compared it with the χ^2 distribution for two degrees of freedom (c.f. fig. 4.3). This procedure was done for several values of LV. All of them showed a good agreement between the distributions and quantiles which indicates that Wilks' theorem is applicable [44].
Note: The applicability of Wilks' theorem is known to be limited in the case of a sterile neutrino search in short-baseline oscillation data. In this case the sought-after oscillatory signature of a sterile neutrino is always preferred in the fit because it can describe part of the statistical fluctuations. However, in the case of the LV search the frequency of the oscillation is fixed, and therefore this issue is absent in the presented analysis.

4.5 Bayesian analysis

For the Bayesian analysis a MCMC with flat priors from 0 to ∞ for A and from 0 to 2π for ϕ are used. We applied the implementation of a two-stages Delayed Rejection Adaptive Metropolis (DRAM) of "pymcmcstat" [46]. This method combines two powerful modifications of the well-known Metropolis-Hastings algorithm (Delayed Rejection (DR) and Adaptive Metropolis sampling (AM)). The basic idea of this adaptions is stated in appendix C and in more detail in [47].
As already mentioned in the section above negative values for the amplitude can be rejected as they correspond to points "mirrored" due to the periodicity of the representation. This means all regions, except the one that is restricted to positive values for A and to the range from 0 to 2π for ϕ, are just describing the same physical parameter. Therefore we restrict the analysis to positive values for the amplitude in order to make use of the advantages of the Bayesian framework and increase the number of posterior samples in the region of interest. The exclusion plot is found using the joint posteriors.

Chapter 5

KATRIN Sensitivity to LV

In this chapter the sensitivity of the KATRIN experiment to LV is studied. In particular, the analysis strategy that was presented in the previous chapter is evaluated and sensitivity limits are calculated.

5.1 Uncertainties

A crucial part of the analysis is investigating and estimating uncertainties. There are statistical uncertainties due to the finite number of counts at each retarding potential, and systematic uncertainties introduced by slow-control parameters, like magnetic fields. Additionally, there are also uncertainties on the model, like the effective time.
In the following the uncertainties of the endpoint fit and the uncertainties concerning the effective time model will be discussed.

5.1.1 Uncertainties of the endpoint fit

The general fitting of the LV model involves the idea that each fitted endpoint has an uncertainty which is included in the final fit of the LV induced amplitude and phase (either using the Frequentist or the Bayesian framework).

Treatment of uncertainties

A major difference compared to other KATRIN analyses is that not the absolute values of the endpoints but the changes from run to run are of interest. Therefore we consider only the time-stability of the parameters.
The uncertainty of the endpoint fit consists of a statistical part and a systematic part, which considers the stability of the slow-control parameter. If one would include uncertainties that are stable over time, the introduced shifts between the different runs were correlated. But as we are only interested in the change of the endpoints between the runs we exclude such time-independent uncertainties.
The influence of the stability is calculated using the MC propagation method of FITRIUM as described in section 2.3.5. The values that are propagated are listed in table 5.1.

Table 5.1: This table shows the stability of the different uncertainties that were considered in the course of this analysis. If they were not included in the MC propagation it is also stated.

systematic uncertainty	value	comment
Column density ρd	0.8 %	
B_{max} stability	0.0038 %	
B_{source} stability	0.1 %	
B_{ana} stability	0.3 %	
Non-Poissonian background	0.0116 cps	added to statistical uncertainty in squares
Plasma e-loss shift	66 meV	
Plasma spectral broadening	80 meV	
Subrun activity fluctuation	dependent on fit range	included via covariance-matrix
Final state distribution	independent of time	not included
Isotopologue concentration	low impact	not included

Activity fluctuation

One uncertainty in the endpoint fit that is not considered using the MC propagation method, the fluctuation of the activity. This systematic uncertainty concerns changes of the activity of the source on the subrun-level and can therefore not be treated as run-to-run fluctuation. In order to estimate the effect of subrun-activity-fluctuations, we used a data-driven covariance matrix approach. Using the subrun-activity of the different runs of KNM-1 a covariance matrix was calculated and used to fit the endpoints. When one compares these fits to the ones without the covariance matrix, one can observe an increase in the uncertainty of the endpoint when including activity fluctuations. Table 5.2 shows the averaged influence of the activity fluctuation on the uncertainty of the endpoints σ_{fluc} and compares it to the statistical uncertainty σ_{stat}. This influence is not neglectable for larger fit ranges, which include retarding potentials further away from the endpoint region. For the 40V-range[1] the influence of this uncertainty is in the order of the other systematic uncertainties (see fig. 5.1). Therefore we argue that even this data-driven approach, which might be not totally perfect[2] is sufficient, as the analysis is statistically dominated, which can be seen in fig. 5.1 and fig. 5.7. Due to this fact and because the 40V-range is also used for the neutrino mass analysis of KNM-1 as systematic effects are well under control in this region of the spectrum, the complete analysis will be performed using this fit range.

Table 5.2: This table shows the influence of the intensity drift on the uncertainty of the endpoint fit for the three different fit ranges. The 40V-range includes retarding potentials larger than 18 535.5 eV. The 99V-range includes potentials larger than 18 476.5 eV while the full range includes all potentials measured in KNM-1

	$\sigma_{\text{fluc}} = \sqrt{\left\langle \sigma_{\text{E}_0}^{(\text{fluc})} \middle\vert \sigma_{\text{E}_0}^{(\text{fluc})} \right\rangle^2 - \langle \sigma_{\text{E}_0} \vert \sigma_{\text{E}_0} \rangle^2}$	$\frac{\sigma_{\text{fluc}}}{\sigma_{\text{stat}}}$
40V range	0.0057 eV	2.3 %
99V range	0.0078 eV	5.5 %
full range	0.012 eV	12.3 %

[1] The so-called "40V-range" includes all retarding potentials higher than 18 535.5 eV.

[2] The covariance matrix is calculated using a combination of all runs, even though we use it for individual runs (not stacked). Thus this description is an approximation. It assumes the same drift uncertainty for all fitted runs.

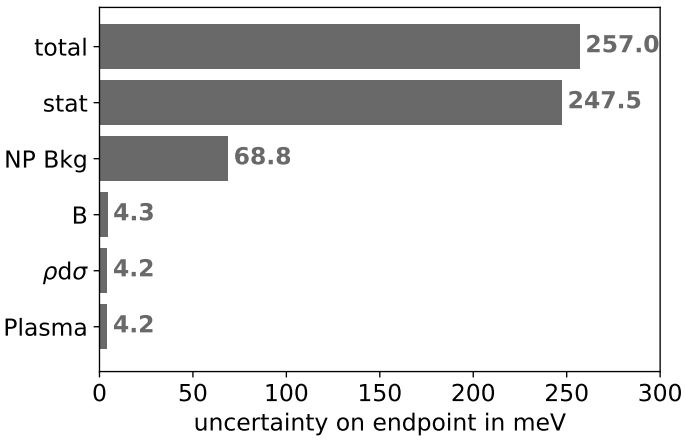


Figure 5.1: This plot shows the systematic breakdown for the 40V fit ranges. The uncertainties used are the stability of the uncertainty (c.f. table 5.1). Activity fluctuations are included via the covariance matrix. One can clearly see that the statistical uncertainty is dominating.

5.1.2 Uncertainties of the effective time

As explained in chapter 4 the analysis procedure is twofold. After assessing the uncertainties of the endpoint fit, this section will focus on the uncertainty related to the model of the effective time (see section 4.1). Hereby the goodness of the model is examined and afterward the uncertainty of the effective time and its influence on the final analysis regarding the LV parameter is evaluated.

Model of effective time in dependence on LV

In order to estimate how good the effective time model performs, we calculated the χ^2 for different values of LV (see fig. 5.2). It is important to note that we used (Asimov) MC data and therefore expect $\chi^2 = 0$. The analysis shows that the effective time model gets worse with larger values of LV. Nevertheless for small LV in the sensitivity region of KATRIN the model shows reasonably good values of χ^2.
Furthermore table 5.3 shows, that the model with an effective time outperforms the model using the starting time of the runs. We therefore conclude that the "effective time"-model can be used as a good approximation.

Table 5.3: This table shows the χ^2 values for the fit using the model with effective time and the model without effective time for $A \approx 0.02\,\mathrm{eV}$. Also the effective times for up and down scans are noted including their uncertainty according to the fit.

	40 V
χ^2_{up} (effective time)	5.8×10^{-5}
χ^2_{up} (no effective time)	0.05
χ^2_{down} (effective time)	3.3×10^{-5}
χ^2_{down} (no effective time)	0.02
effective time (up-scan)	$(87.0 \pm 0.2)\,\mathrm{min}$
effective time (down-scan)	$(51.9 \pm 0.2)\,\mathrm{min}$

Another point that is important for the performance of the effective time model is the fact, that one assumes that the effective time is independent of the LV parameter. To check this we

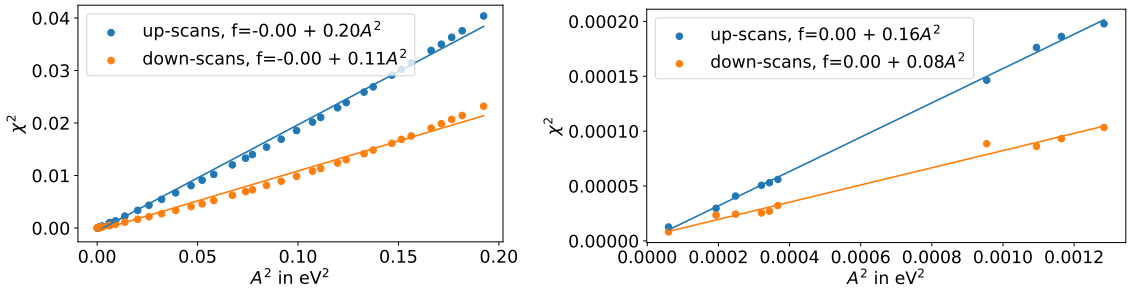


Figure 5.2: These figures show the behavior of χ^2 of the fit for different LV parameters. The right plot is zoomed into the region of smaller LV. For convenience, a linear function is fitted to illustrate the behavior.

created many data sets with different values of the LV parameter and fitted the effective time as described above. Figure 5.3 shows the behavior of the effective time for different values of LV, which indicates, that it is stable and therefore can be used as a constant. At this point, it is important to note that the phase of the oscillation is not defined for no LV. Therefore the values close to zero were neglected.

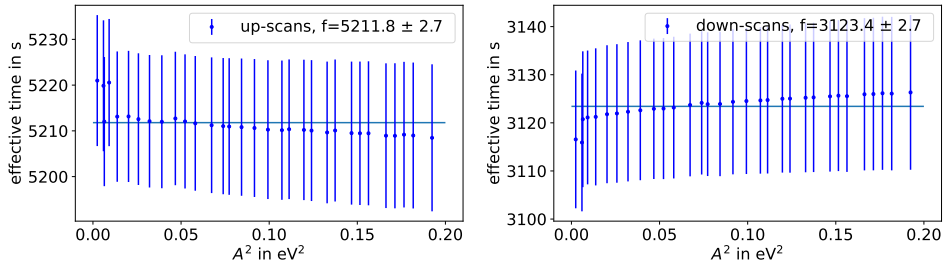


Figure 5.3: This figure shows the fitted effective times for different values of LV. The left plot shows the effective time for up scans, the right one for down scans. In both of the plots the effective time is stable to a reasonable amount.

Uncertainty of effective time

After the performance of the model was examined the uncertainty of the effective time will be studied. The effective time is found by fitting the oscillation for a given LV parameter to the simulated MC data. This fit has an uncertainty that can be obtained from the covariance matrix of the fitter. However, it remains unclear if this uncertainty is valid over the full oscillation. Thus, to have a different approach to estimate the uncertainty, the residuals of the effective time were calculated in fig. 5.4.
The residuals are larger at the extrema of the oscillation as the oscillation becomes flat in this region and small differences in the endpoint value correspond to larger residuals in the effective time. The width of the distribution of the residuals is used to calculate the uncertainty of the effective time. It exceeds the one of the fit (c.f. fig. 5.5). To be conservative this uncertainty will be used and a later analysis shows that this does not harm our sensitivity.
In the fit of the LV parameter we use the effective time as a nuisance parameter with a Gaussian pull-term. The effective time corresponds to a shift in time and therefore a shift in the phase of the oscillation. This means that its uncertainty is correlated to an uncertainty of the phase. However, the phase of the oscillation ranges from 0 to 2π in approximately 24 h,

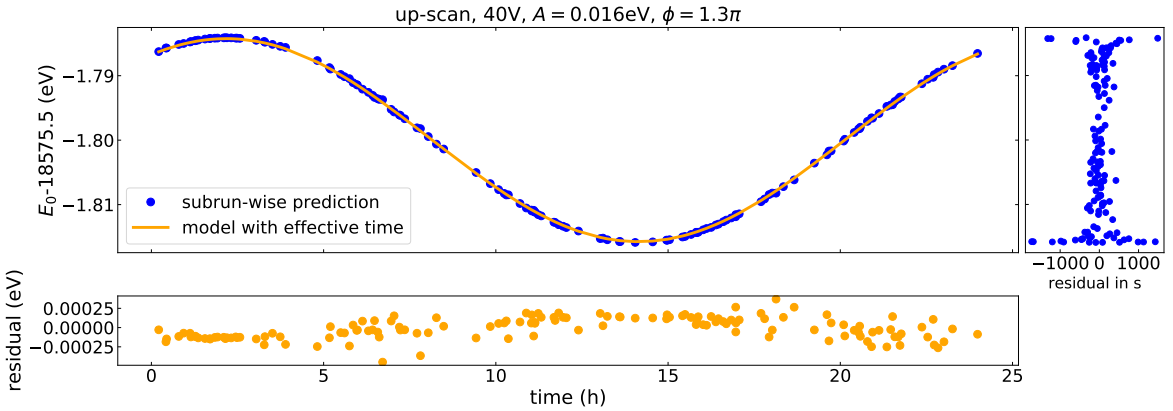


Figure 5.4: This figure shows the generated endpoints for $A \approx 0.016\,\text{eV}$ and $\phi \approx 1.3\pi$ in blue. The model using the effective time estimates the endpoints as shown in orange. The residuals of the endpoints are small as one can see below the figure. On the right-hand side, the residuals in time are shown. As expected the residuals increase close to the extrema because at these points even small discrepancies lead to larger discrepancies in the time residuals while not harming the estimated size of the endpoints much.

while the uncertainty of the effective time is a few minutes, which avoids the problem of this degeneracy. Figure 5.6 shows that the effect of the pull-term is tiny and has, therefore, no negative influence on the analysis.
To summarize, this section has shown that the effective time model is a good approximation to the data. The used effective time is stable for different values of LV and can be included in the model with an uncertainty that has only a minor effect on the final result.

5.1.3 Breakdown of systematic effects on the exclusion curve

In the previous section the uncertainties of the endpoint fits and the effective time were shown. For the final analysis of the LV parameter these uncertainties are considered together. As the influence of the uncertainty of the effective time is negligible, fig. 5.7 shows only the breakdown for the systematics of the endpoint fit.
In the systematic breakdown of the endpoint fit uncertainty (fig. 5.1) one can see that the statistical uncertainty is by far the largest contribution to the total uncertainty budget. There-

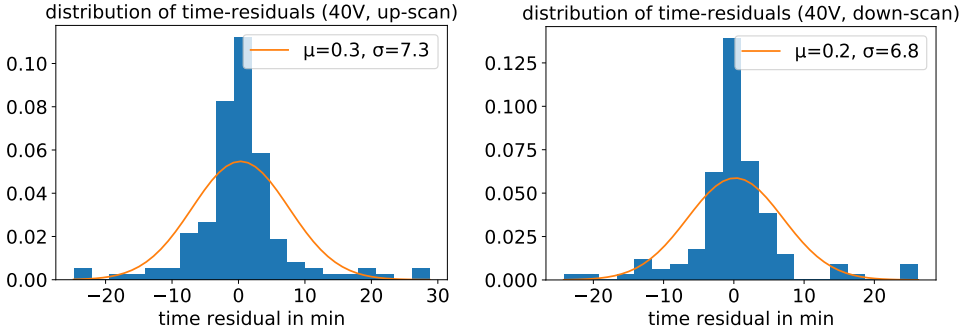


Figure 5.5: This figure shows the distribution of the residuals of the effective time as shown in fig. 5.4. A Gaussian was fitted to have an estimate of the uncertainty which is around 7 min, which is larger than the uncertainty received by the fit.

fore one could suspect that also the sensitivity on A and ϕ is mostly statistically dominated. In order to support this statement in fig. 5.7 the sensitivity for using only the statistical uncertainty and for using the total uncertainty are drawn next to each other and one can only spot a minor difference. Therefore we can conclude that the analysis is mostly statistically dominated.

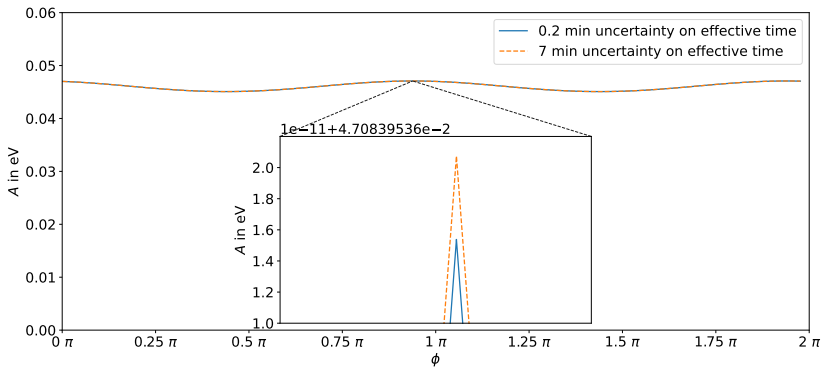


Figure 5.6: This figure shows two different exclusion curves for no LV. The blue one uses the 0.2 min uncertainty on the effective time received by the fit. The orange one uses 7 min from the residuals of fig. 5.5. In the plot there is no difference visible. To show that there is a tiny difference part of the plot was zoomed-in. This result shows that the difference in the sensitivity between using the larger or smaller uncertainty on the effective time is negligible.

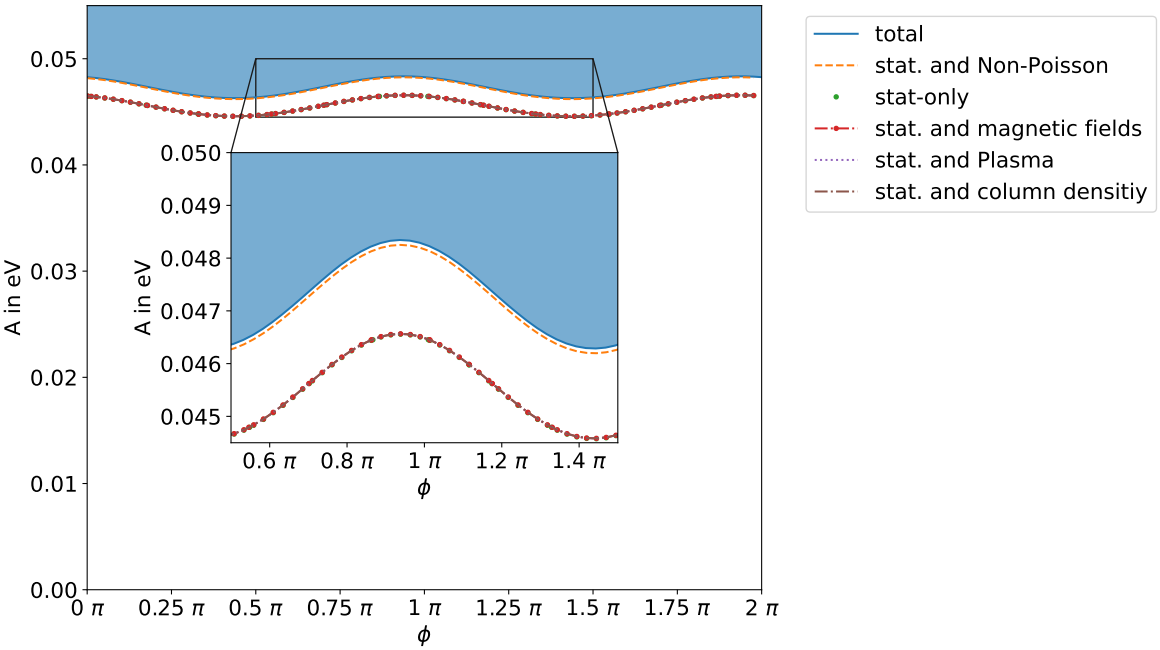


Figure 5.7: This figure shows the influence of the different uncertainties on the 90% C.L. exclusion curve for the case of no LV in the frequentist approach. The shaded area is excluded. In the zoomed-in plot it becomes clear that if one considers the non-Poissonian uncertainty additionally to the statistical uncertainty, the difference to the total uncertainty becomes small. As the non-Poissonian uncertainty scales with the statistical uncertainty, this reinforces that the analysis is statistically dominated.

5.2 Exclusion

Using the complete set of uncertainties that were discussed in the previous section one can perform the exclusion for the LV induced oscillation. Figure 5.8 shows the sensitivity curve, which illustrates the parts of the parameter space of A and ϕ that can be excluded. The sensitivity curves are plotted using the Frequentist as well as the Bayesian analysis. It is clearly visible that in the case of no LV the sensitivity of KATRIN is limited to exclude A but hardly the phase ϕ, as it is not defined for the case of no LV. Furthermore, a clear difference between the Bayesian credibility and the frequentist confidence exclusion is visible. This will be discussed in the following.

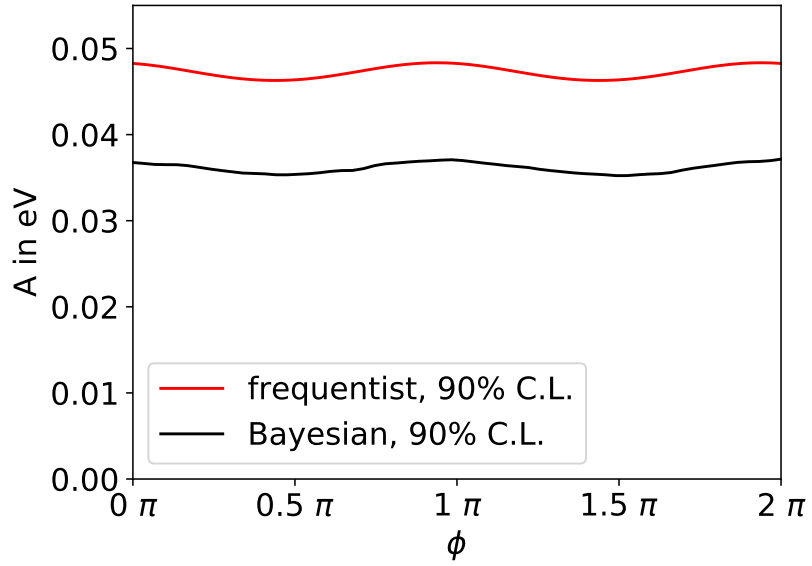


Figure 5.8: In this plot, the 90% exclusion curve for the Bayesian MCMC (black) and frequentist χ^2 (red) are shown for the case of no LV. The used fit range includes potentials higher than 18 535.5 eV. The exclusions are created using the joint posterior or a grid-search. The visible difference is explained in section 5.2.1.

5.2.1 Difference between Bayesian and frequentist exclusion

In general the frequentist and the Bayesian approaches rely on different interpretations of probability and it is therefore known and obvious that credibility regions[3] which are derived in the Bayesian framework and confidence levels of the frequentist methods do not have to agree in general (see e.g. [48, 49]).
The frequentist interpretation of a confidence interval is the following: If an ensemble of several similar experiments, that measure the same parameter a, calculate a confidence interval at 90% C.L. it is understood that 90% of the calculated intervals contain the true value of a [49]. On the other hand the Bayesian framework assigns a probability distribution to the true parameter a. A Bayesian credibility interval of 90% C.L. states therefore that the true value of is a within this interval with a probability of 90% [48].
In particular, the Bayesian framework is dependent on the used priors, which describe the a priori degree-of-belief that the parameter has a specific value. Generally priors are not

[3] In particle physics the term confidence level is also used for Bayesian exclusion. However, to avoid confusion, this thesis will stick to credibility regions.

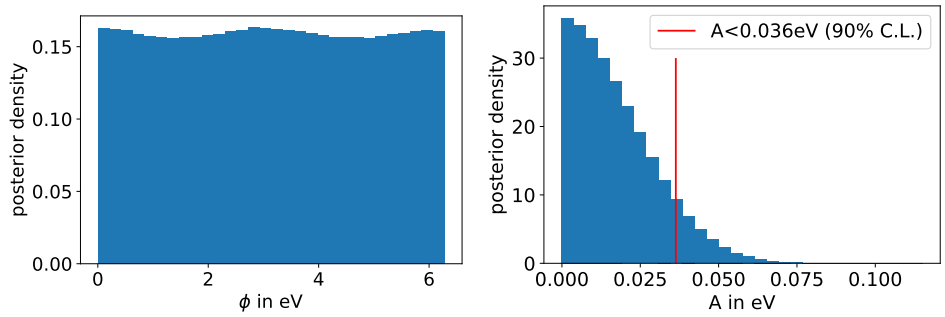


Figure 5.9: This figure shows the distribution of the marginalized posteriors. In the left plot it is visible that for no LV the posteriors of the phase are almost flat. Therefore no limit or exclusion is determined. For the amplitude in the right plot the limit can be found to be $A \leq 0.036\,\text{eV}$ at 90% credibility level.

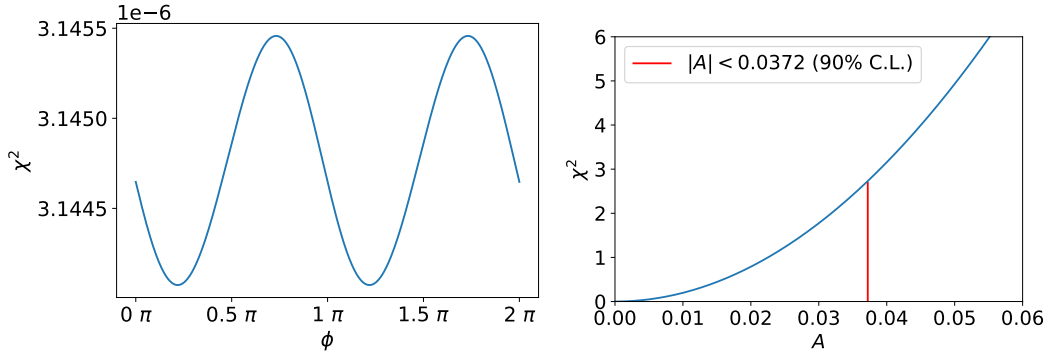


Figure 5.10: This figure shows the χ^2 of the profiled parameter. In the left plot it is visible that for no LV the posteriors of the phase are almost flat and close to $\chi^2 = 0$. Therefore no limit or exclusion is determined. For the amplitude in the right plot the limit can be found to be $A \leq 0.037\,\text{eV}$ at 90% confidence level.

invariant under reparameterization. In section 5.5.4 this obstacle will be discussed. As a detailed discussion on the controversy between supporters of the Bayesian interpretation of probability and the frequentist one is beyond the scope of this thesis, all central conclusions are performed using the latter as it is more common in particle physics.

5.2.2 Sensitivity limits on A and ϕ

Besides a joint exclusion of A and ϕ one can also marginalize them. In the Bayesian approach the posteriors of the corresponding parameter are used for the derivation. In the frequentist approach the χ^2 is minimized with respect to the other parameters and Wilks' theorem is used for setting the limit. This means, for limiting A, ϕ is treated as a nuisance parameter. As the phase is not defined in the case of no LV, it is senseless to calculate a limit for it. Nevertheless, the profile likelihood and the posterior distribution are plotted in fig. 5.9 and fig. 5.10.
For the Bayesian MCMC we can conclude that $A \leq 0.036\,\text{eV}$ at 90% credibility level. The limit of the frequentist method is slightly larger and quotes $A \leq 0.037\,\text{eV}$ at 90% confidence level.

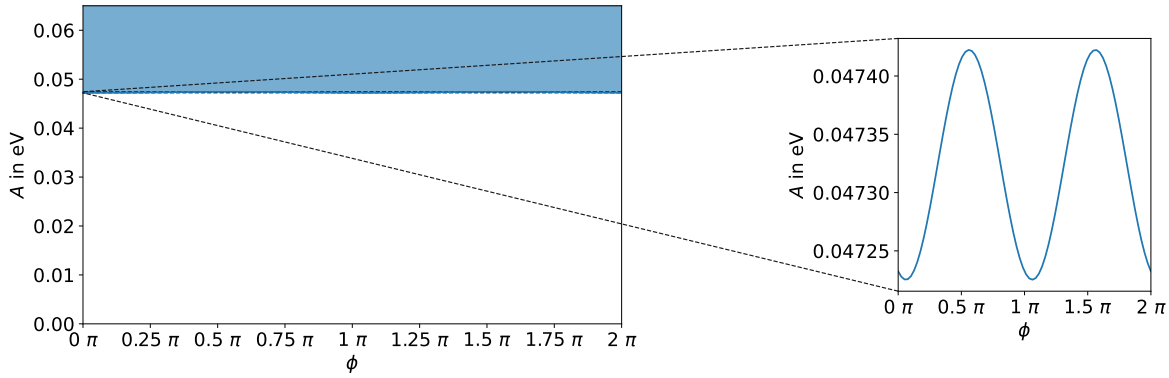


Figure 5.11: This figure shows the exclusion curve for no LV if one distributes the endpoints equidistantly in time. The strength of the "wiggling" structure in the phase is reduced compared to fig. 5.7.

5.2.3 Explanation of the shape

In e.g. figs. 5.8 and 5.9 an oscillatory structure in ϕ is visible. This effect can be explained by the following argumentation.
From fig. 3.3b it is visible that the endpoints are not distributed equidistantly in time. Therefore values of the phase are preferred that describes the region of a "cluster" of endpoints better. In order to demonstrate this effect, we redistributed the MC endpoints equidistantly in time. The result shows a much smaller amplitude of the "wiggling" structure in the exclusion plot (see fig. 5.11).
However, a small structure is still visible which can be explained by the model of LV $\Delta E = A\cos(\omega t - \phi)$. If one assumes that all N endpoints in the period T have the same value and uncertainty σ the χ^2 becomes

$$\chi^2 = \frac{A^2}{\sigma^2}\left[\cos^2(0-\phi) + \cos^2\left(\frac{2\pi}{N} - \phi\right) + \cos^2\left(2\frac{2\pi}{N} - \phi\right) + \dots\right] \tag{5.1}$$

$$= \frac{A^2}{\sigma^2}\sum_{k=0}^{k=N}\cos^2\left(\frac{2\pi}{N}k - \phi\right) \tag{5.2}$$

$$= \frac{A^2}{2\sigma^2}(N + \cos(2\phi) + 1). \tag{5.3}$$

This explains why an oscillatory behavior in the χ^2 distribution for different phases with a period of π is predicted. Therefore one expects the behavior observed in fig. 5.11.

5.2.4 Parameter Recovery

After the values and shapes of the sensitivity exclusion curves are explained, it is important to keep in mind, that these use Monte Carlo data with no LV as input. However, it is of large interest to test whether the used analysis framework is capable to recover values of LV parameters, that are non-zero and therefore an oscillation with non-zero amplitude. We therefore created data sets of endpoints for different values of LV as described in section 5.1.2. This data was afterward used to fit the LV parameters, which were compared to the ones used for generating the data. Table 5.4 shows that the analysis has the power to recover the parameters.
Furthermore the coverage study of the previous section was not only performed for the case of no LV but repeated for different values of LV and it showed that Wilks' theorem and therefore

the exclusion method is applicable (c.f. fig. 4.3).
These results show that the KATRIN experiment can search for LV induced oscillations using the methods explained above.

Table 5.4: This table shows some exemplary best fits for different LV. Those were received using the frequentist method. For the case of no LV the phase is expected to be equally good (almost) everywhere.

original A	original ϕ	best fit A	best fit ϕ
0 eV	not defined	7.44×10^{-7} eV	1
0.020 eV	3.98	0.019 eV	3.98
0.0255 eV	2.47	0.0256 eV	2.47

5.3 Projection for final KATRIN

As KNM-1 was the first campaign of the KATRIN experiment the statistics are limited. Thus it is interesting to investigate the sensitivity of all (planned) KATRIN measurements to the anisotropic part of a^μ.
While KNM-1 was performed with a column density of $1.10 \times 10^{21}\,\mathrm{m}^{-2}$ the total KATRIN is expected to run at $5 \times 10^{21}\,\mathrm{m}^{-2}$ and a measurement time that corresponds approximately to 750 d [19]. As also different settings and run-lengths are planned, a correct estimate of the sensitivity of KATRIN to LV is tedious and hardly possible. We therefore decided to do a rough estimate for which we used the sensitivity of 750 d of KMN-1 like runs. Assuming a duration of approximately 2 h this corresponds to about 9000 runs. We therefore created a set of 9000 endpoints, which we afterward fitted using the same frameworks as described before. Therefore the expected improvement is $\sqrt{\frac{9000}{274}} \approx 5.7$. This can also be seen in fig. 5.12, which shows that $A \lesssim 0.008$ eV for final KATRIN.

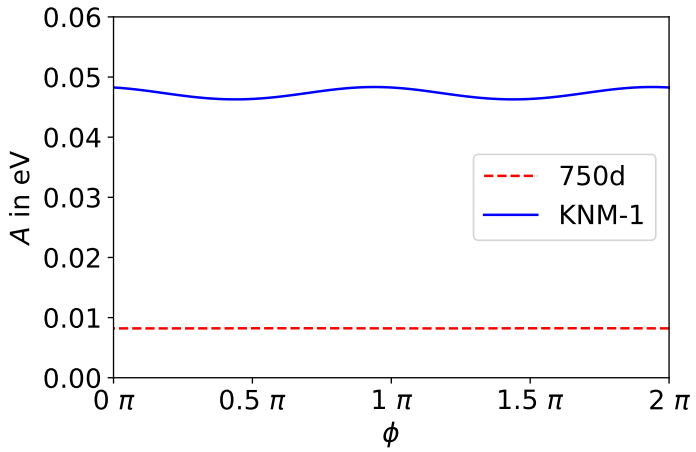


Figure 5.12: This figure shows the difference in the stat-only sensitivity of KNM-1 (blue) and full KATRIN (red) to LV. For the full KATRIN estimation, 750 d KNM-1 like data was used.

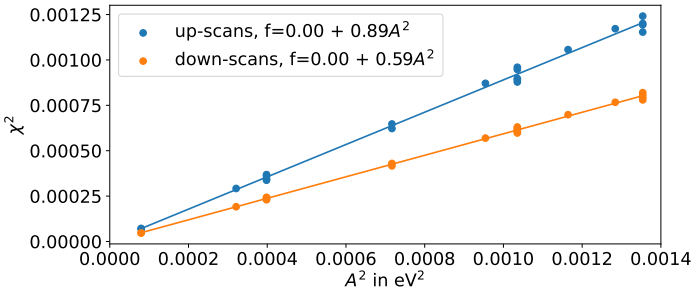


Figure 5.13: This figure shows the χ^2 of 274 3 h runs (split in up and down runs) in dependence of A^2. To illustrate the dependence a linear fit is performed and is indicated by the orange and blue lines. For this plot MC data was used.

5.4 Influence of run length

As mentioned in the previous section the KATRIN collaboration plans to change the length of the runs. This has also an impact on future LV analysis. An increased run length improves the statistics of a single run. However, the number of runs that can be performed per day is reduced. Furthermore larger runs might worsen the temporal resolution and the model of effective time. This chapter focuses on an exemplary increase of the run duration by a factor of 1.5. In particular, it will concentrate on the impacts on the LV analysis.

5.4.1 Effective time

The effective time is trivially influenced by a longer run-time which also means extended subrun durations. To estimate the effective time for the enlarged runs, we use the same method as in section 4.1 and calculate MC data for different values of LV and fit them with FITRIUM. Afterward, these endpoints are compared to the theoretical model with a free effective time. The residuals in the effective time can be used to estimate the uncertainty of the effective time. Table 5.5 shows the differences between the KMN-1 like runs (2h runs) and the ones which run-times are increased by a factor of 1.5 (3h runs). It is visible that not only the effective times become larger but also the difference between the effective time of the up- and the down-scans increases. Furthermore one can notice that the uncertainty of the effective time rises.

Table 5.5: This table shows the effective time for KNM-1 like runs (2h runs) and runs that were increased by a factor of 1.5 (3h runs). Also the differences in the uncertainty of the effective time is displayed.

	2 h	3 h
effective time (up)	87 min	130 min
effective time (down)	52 min	79 min
uncertainty	6 min to 7 min	12 min to 13 min

5.4.2 Effective time model

A rise in the uncertainty could indicate a worse performance of the effective time model. To validate the quality of this model, the χ^2 was calculated for different values of LV. Table 5.6

shows the goodness of the model. In order to compare the cases with different statistics, it is useful to introduce the reduced $\chi^2_R = \frac{\chi^2}{\#d.o.f.}$ which normalizes the χ^2 by the number of degrees of freedom. For illustration, fig. 5.13 shows the χ^2 of the model.
This comparison clearly shows that the performance is better for the shorter run-duration. Yet, this model is still outperforming a model using the start times by orders of magnitudes. Additionally fig. 5.14 shows that the effective time is still a constant for different values of LV. Nevertheless one could consider implementing the LV model into the maximum-likelihood fit of the KATRIN data.

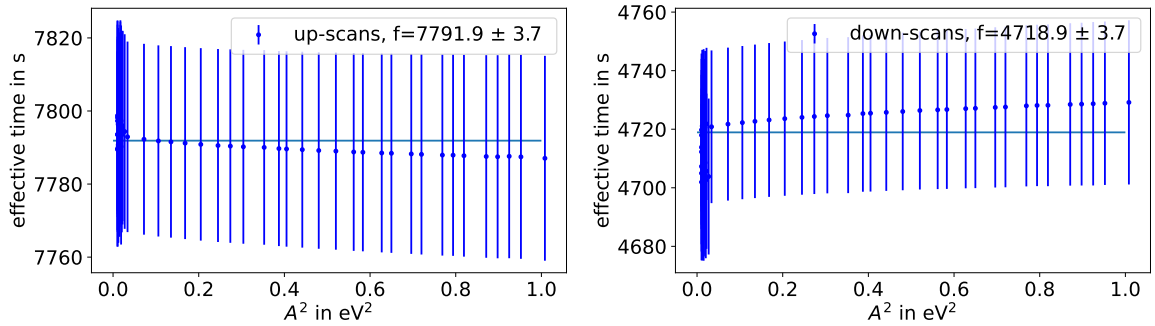


Figure 5.14: This figure shows fitted effective times for different values of A^2 for up and down scans. One should notice that for $A^2 = 0$ no phase is defined and therefore no effective time can be fitted. Furthermore no clear dependence is visible.

Table 5.6: This table shows the reduced χ^2_R of the effective time model for KNM-1 like runs (2h runs) and runs that were increased by a factor of 1.5 (3h runs) for $A = 0.023\,\mathrm{eV}$.

	2 h	3 h
χ^2_R effective time (up)	3×10^{-6}	1×10^{-5}
χ^2_R effective time (down)	1×10^{-6}	1×10^{-5}

5.5 Analysis in two-amplitude description

As explained in section 3.5 there are other choices to describe the oscillation. In particular one can use the "two-amplitude" description. Even though the description using phase and amplitude seems more intuitive as it directly reflects the effect one can see, it is interesting to investigate the differences that might occur using the "two-amplitude" model. As the procedure is in most parts identical to the one shown before, this section will focus on the slight differences.

5.5.1 Model and effective time

The measured and fitted endpoints are not influenced by the different parameterizations. Therefore the effective time stays unchanged but the parameterization of the fitted model becomes

$$\Delta E_0 = a\cos(\omega T_\oplus) + b\sin(\omega T_\oplus). \tag{5.4}$$

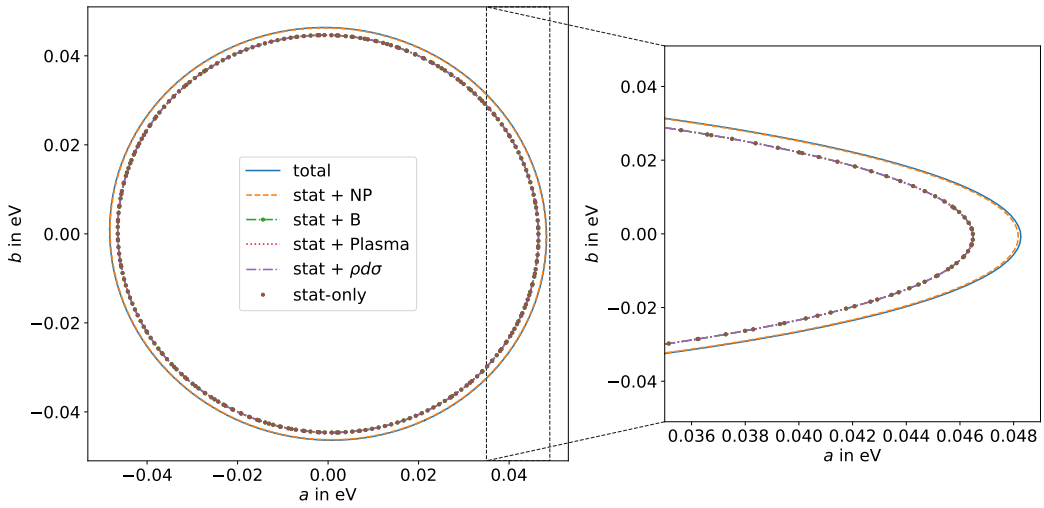


Figure 5.15: This figure shows the influence of the different uncertainties on the 90% C.L. exclusion curve for the case of no LV in the "two-amplitude" description. The area outside of the ellipses is excluded. In the zoomed-in plot on the right it becomes clear that if one considers the non-Poissonian uncertainty additionally to the statistical uncertainty, the difference to exclusion for the total uncertainty budget becomes small.

This leads to the same fitting procedure with the difference that a and b are the fitted parameters. In the Bayesian approach, a further change is necessary. While it made sense to restrict the prior of the amplitude to be positive in the previous description, the prefactors a, b part can be negative and positive. Therefore flat priors from $-\infty$ to ∞ are used for them.

5.5.2 Breakdown of systematic effects

While the systematic uncertainties of the endpoint fits and the effective time are not affected by a new parameterization, the breakdown and exclusion curves trivially are. Therefore fig. 5.15 shows the exclusion-curves for the different uncertainties. This figure shows that the analysis is statistically dominated as the exclusion-curve for non-Poisson and statistical uncertainty is almost identical with the one of the full uncertainty budget. A further difference to the description with A and ϕ is immediately visible. Due to the fact, that a and b are both well defined for no LV ($a = b = 0$), the contours are closed.

5.5.3 Coverage study and parameter recovery

The applicability of Wilks' theorem was performed analogously to the description using A and ϕ by calculating

$$\chi^2\,(a, b) - \chi^2(\text{best fit}) \tag{5.5}$$

of fluctuated data and comparing the distribution to the χ^2 distribution for two degrees of freedom. Figure 5.16 shows that Wilks' theorem is applicable for this analysis. It was also tested for different values of LV and suggests that the limits of a and b can be found by using the critical χ^2_c according to Wilks' theorem.
Additionally it was examined whether the framework is able to recover non-zero LV parameters. Table 5.7 suggests a very good agreement between the recovered best fits a and b and the true data. Therefore also the analysis using the "two-amplitude" description can be performed to limit the LV which induces the endpoint oscillation.

Table 5.7: This table shows some exemplary best fits for different LV. Those were received using the frequentist method in the "two-amplitude" description.

original a	original b	best fit a	best fit b
$0\,\mathrm{eV}$	$0\,\mathrm{eV}$	$6.25 \times 10^{-7}\,\mathrm{eV}$	$5.09 \times 10^{-7}\,\mathrm{eV}$
$1.32 \times 10^{-2}\,\mathrm{eV}$	$1.44 \times 10^{-2}\,\mathrm{eV}$	$1.27 \times 10^{-2}\,\mathrm{eV}$	$1.40 \times 10^{-2}\,\mathrm{eV}$
$2.02 \times 10^{-2}\,\mathrm{eV}$	$4.51 \times 10^{-2}\,\mathrm{eV}$	$2.04 \times 10^{-2}\,\mathrm{eV}$	$4.55 \times 10^{-2}\,\mathrm{eV}$

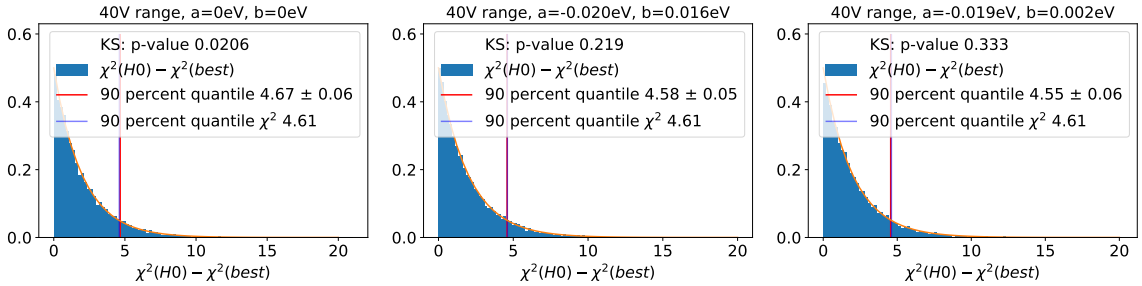


Figure 5.16: These plots show the distribution of $\chi^2\,(a, b) - \chi^2(\text{best fit})$ in blue for different values of LV. The orange curve is the theoretical χ^2 distribution for two degrees of freedom. The vertical lines indicate the 90% quantiles. The uncertainty corresponds to the bootstrap-error. The distributions and their quantiles agree well with the prediction of Wilks' theorem.

5.5.4 Exclusion and comparison of the descriptions

At this point the question arises whether the two descriptions lead to the same interpretation. The exclusion curve of the "two-amplitude" description can be transformed into an exclusion in A and ϕ by using the relation

$$A = \sqrt{a^2 + b^2} \tag{5.6}$$

$$\phi = \arctan(a/b) \tag{5.7}$$

However, before doing the transformation one should notice that the Bayesian and frequentist exclusions agree numerically quite well in "two-amplitude" description. Figure 5.17 shows the 90% credibility and confidence exclusions. Thus one expects them to have a similar exclusion in A and ϕ.
Figure 5.18 shows the exclusion of the different descriptions in A and ϕ. It is visible that

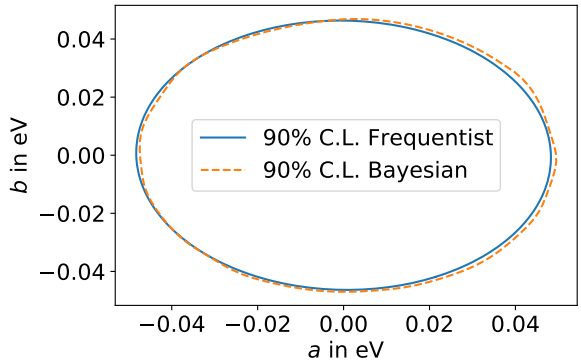


Figure 5.17: This figure shows the 90% C.L. exclusion curves for MC-data with no LV in "two-amplitude" description. It is visible that the two frameworks agree within numerical precision.

no difference is observable between the two descriptions for the frequentist method. The difference in the Bayesian exclusions is due to the different priors that were used in the two approaches. A flat prior in a and b does not translate into a flat prior in A and ϕ, which is illustrated in fig. 5.19.
The fact that priors are generally not independent of the parameterization is known. In the last century a discussion between the subjective interpretation of Bayesian inference and the usage of so-called "non-informative" priors took place. In this context, inter alia, Harold Jeffreys tried to develop a concept to select priors according to a convention. He is in particular known for a method that suggests priors which are invariant under different parameterizations.
A general transformation of a prior p_α from parameters α to α' can be written as [50, 51]:

$$p_{\alpha'} = p_\alpha \left\| \frac{\partial \alpha}{\partial \alpha'} \right\| \tag{5.8}$$

As Fisher's information $I(\alpha)$ transforms like [52]

$$I\left(\alpha'\right) = \frac{d\alpha}{d\alpha'} I(\alpha) \frac{d\alpha^{\mathrm{T}}}{d\alpha'}, \tag{5.9}$$

a parameterization-independent choice of the prior π is given by the so-called Jeffreys' prior [52]

$$\pi(\alpha) \propto |I(\alpha)|^{1/2} \tag{5.10}$$

where $|I(\alpha)|$ is the determinant of the Fisher information matrix in the parameterization α.
In the following a test will be performed to check if the flat prior in "two-amplitude" description is a Jeffereys' prior. Therefore we use the χ^2 of the LV model as basis. We furthermore assume for the calculation N equidistant data points in the period of 2π.
The Fisher matrix can be calculated as [53]

$$F_{\alpha\beta} = \sum_l \frac{1}{\mathrm{Var}[\Delta \hat{E}_0(t)]} \frac{\partial \Delta E_0^{\mathrm{theory}}\left(t, \{\bar{\lambda}_\gamma\}\right)}{\partial \bar{\lambda}_\alpha} \frac{\partial \Delta E_0^{\mathrm{theory}}\left(t, \{\bar{\lambda}_\gamma\}\right)}{\partial \bar{\lambda}_\beta} \tag{5.11}$$

with $\{\bar{\lambda}_\gamma\} = \{a, b\}$ and $\{\bar{\lambda}_\gamma\} = \{A, \phi\}$ respectively.
Assuming a constant uncertainty of the endpoints $\mathrm{Var}[\Delta \hat{E}_0(t)]$ the determinants can be cal-

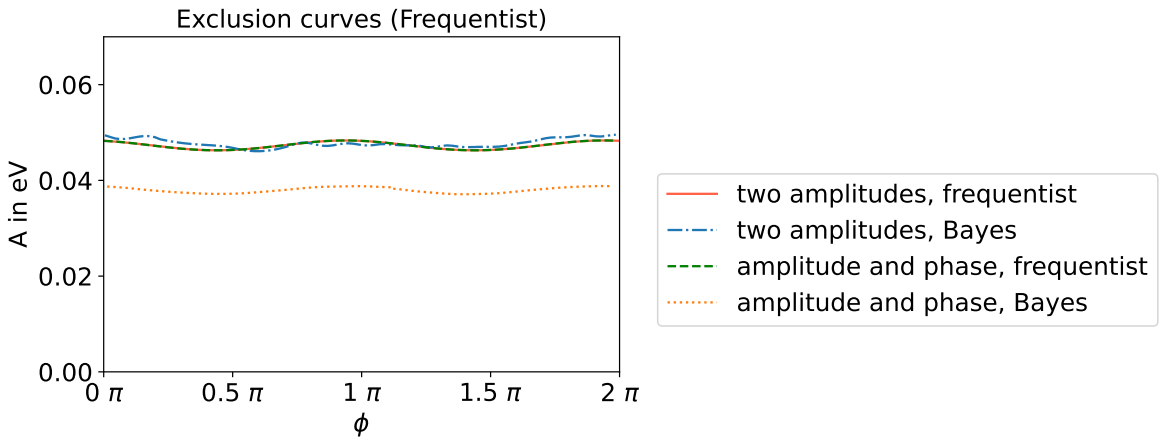


Figure 5.18: This figure shows the exclusion curves in A and ϕ for the descriptions using amplitude and phase as well as the "two-amplitude" description. The frequentist exclusions agree very well in both descriptions. The Bayesian analysis shows a difference. Namely the limit derived with a description in A and ϕ is lower than the limit of "two-amplitude" description. However, the difference in the limits is only a few 10 meV.

culated to be:

$$|I(a,b)| \propto C \propto 1 \tag{5.12}$$

$$|I(A,\phi)| \propto C \times |A|^2 \propto A^2 \tag{5.13}$$

It can be easily checked that this is consistent with the Jacobian determinant of the transformation.
Thus the flat priors in A and ϕ are indeed Jeffreys' priors while the flat priors in the "two-amplitude" description are not, which explains the difference between them. A usage of Jeffreys' prior is shown in appendix D. However, the choice of Jeffreys' priors is not always used. Even Jeffreys himself defined cases in which he wished to deviate from this "general rule" [54]. Additionally, the frequentist framework is more common in particle physics. Therefore the final analysis of the data will be performed using the frequentist framework.

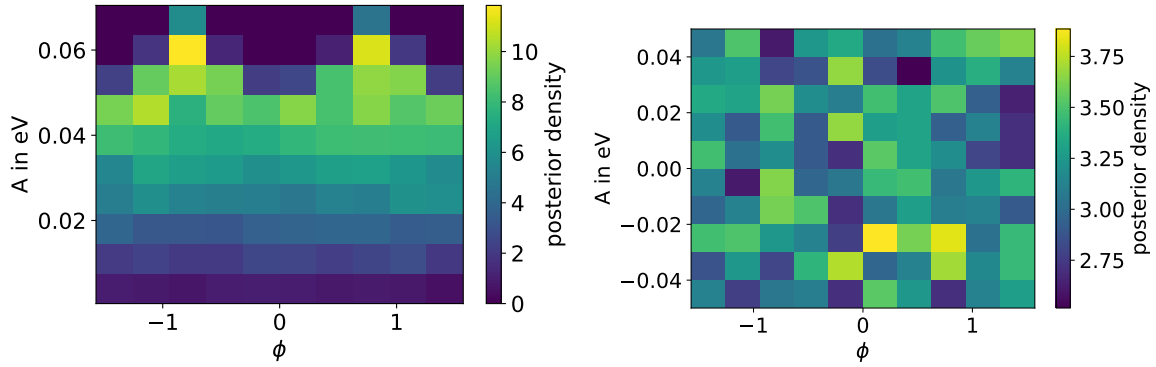


Figure 5.19: The left plot shows flat priors in a and b that are transformed into amplitude and phase. On the right flat priors in A and ϕ are shown for comparison. This illustration makes clear, that the different descriptions lead to different priors and therefore different exclusion curves. Furthermore the prior distribution that is flat in a and b prefers larger values of A which explains the larger exclusion curve.

Chapter 6

LV search based on KNM-1 data

In the previous sections the sensitivity of the KATRIN experiment on effects to the anisotropic neutrino LV parameters was studied. In this section the procedures of chapter 4 and chapter 5 will be applied to KNM-1 data and a first limit will be concluded.

6.1 Limit on non-isotropic LV

6.1.1 Data

The analysis framework of chapter 5 was studied with Monte Carlo data. After those tests had been carried out the same framework is applied to the data of the first neutrino mass campaign KNM-1.
The 274 "golden-runs" are fitted uniformly and analyzed with the model of effective time in the frequentist framework to calculate the limit on A and ϕ. Hereby the following procedure was performed:

1. The endpoint of each KNM-1 run is fitted including the covariance matrix to account for activity-fluctuations on the subrun-level. This creates a set of endpoints that is associated with the time of the corresponding run.
2. The uncertainties of the slow-control parameters are propagated to assign an uncertainty to each fitted endpoint.
3. The model using the effective times and its uncertainties are fitted to the data-set of endpoints, times and endpoint uncertainties. The effective time and its uncertainty are determined by an MC-study as described in chapter 5.
4. A χ^2-grid scan is used to determine the exclusion curve for A and ϕ according to Wilks' theorem.

6.1.2 Uncertainty breakdown

The uncertainties on the endpoints were calculated in the same way as in section 5.1.1. Figure 6.1 shows the breakdown of an exemplary run. Similar to the MC study, it is visible that the endpoint fit is dominated by statistical uncertainty and non-Poisson background. Furthermore one can see a good agreement with the breakdown of the MC data. The propagation of all uncertainties is computationally expensive and the sensitivity study shows that the difference of the exclusion curves for total uncertainty and statistical with NP uncertainty

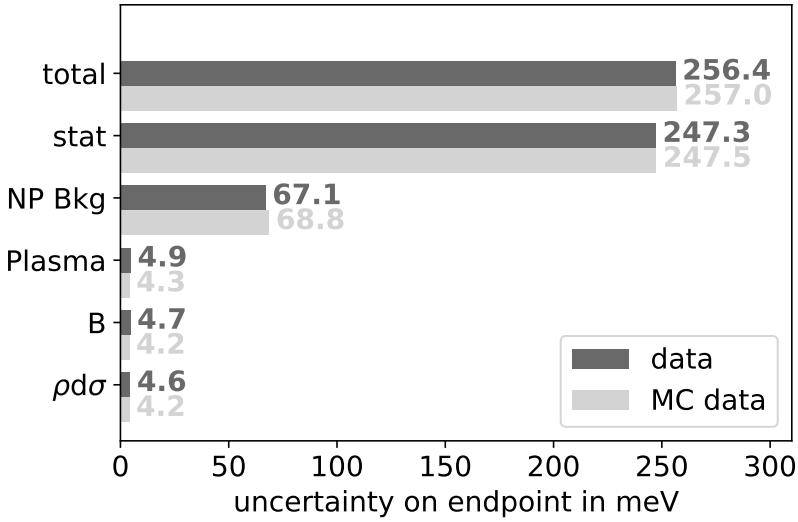


Figure 6.1: This figure shows the systematic breakdown of one exemplary run (51443) for data and MC. It is visible that the uncertainty budget is dominated by the statistical and NP uncertainty.

are tiny. Therefore, in the course of this thesis the systematic uncertainties are added to the statistical and NP ones in squares.[1]
If this is applied to the analysis, this leads to the exclusion-curves which are depicted in fig. 6.2. They suggest that the analysis is indeed statistically dominated.

Exclusion on non-isotropic LV

The exclusion limit is obtained by a χ^2-grid search. The uncertainties of the endpoints include both statistical and systematic contribution (see fig. 6.2). It is visible that the best fit is not equal zero (see table 6.1). However, one can calculate the difference of the best fit to the Null-hypothesis and $\Delta\chi^2 = \chi^2(A = 0) - \chi^2(\text{best fit}) = 1.86$, which corresponds to a 0.9σ-fluctuation and is therefore not significant at 90% C.L.. This can also be seen in fig. 6.2 as the contour is not closed at this confidence level. Figure 6.3 illustrates the fit result.

Table 6.1: This table summarizes the best fit results of A and ϕ with data of KNM-1.

	A	ϕ
stat-only	$3.12 \times 10^{-2}\,\text{eV}$	0.79π
stat-only + NP	$3.05 \times 10^{-2}\,\text{eV}$	0.78π
total	$3.05 \times 10^{-2}\,\text{eV}$	0.78π

Marginalized exclusion limits

Similar to section 5.2.2 it is possible to calculate a "marginalized" limit on A. Therefore we use a profile likelihood, whereby for different values of A, χ^2 is minimized with respect to the nuisance parameters and ϕ. Using this method leads to a more stringent limit than in the

[1] Hereby the systematic uncertainties of multiple runs were tested and showed good agreement. Due to the fact that the analysis is statistically dominated, those values were used for the analysis with all runs.

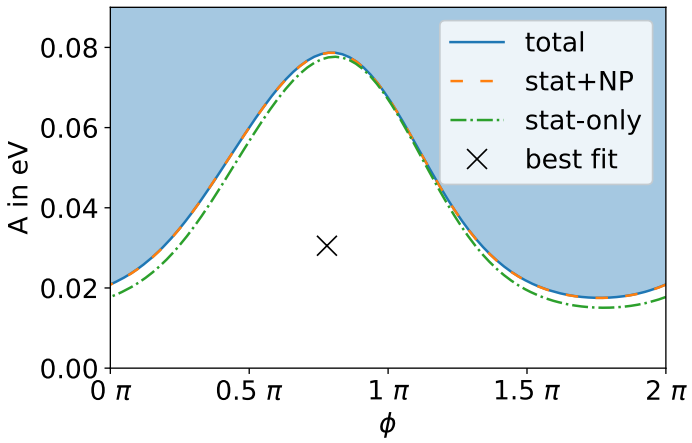


Figure 6.2: This figure shows the exclusion curves using only statistical uncertainties (green), statistical and NP uncertainty (orange) and all uncertainties (blue). The colored area is excluded by the analysis. A fluctuation, that is not significant at 90% C.L., is visible.

maximal value of A in the two-dimensional grid-scan:

$$\boxed{|A| \leq 0.068\,\text{eV}}\ (90\%\ \text{C.L.}). \tag{6.1}$$

On the other hand, the maximal value of A in the grid-scan with the total uncertainty budget is found to be $|A| < 0.079\,\text{eV}$ at 90% C.L.. Applying eq. (3.47) one can transform the former limit into spherical decomposition.
Using the values of table B.1, one receives $B = M_T^{-1}(2\pi(1-\cos\kappa))^{-1}\pi\sqrt{E_m^2 - m_e^2}\sin^2\kappa \approx 4.7 \times 10^{-5}$ and $\kappa = \theta_0/2$. Furthermore the rotation-velocity is $\beta_{rot} \approx r_\oplus \omega_\oplus \sin\chi \approx 1 \times 10^{-6}$ (see table B.1). This leads to a preliminary limit of

$$\boxed{\left|a_{11}^{(3)}\right| \lesssim 2.7 \times 10^{-6}\,\text{GeV}}\ (90\%\ \text{C.L.}). \tag{6.2}$$

The prefactors B and β_{rot} are not yet confirmed by theorists and might change. Therefore this limit in $a_{11}^{(3)}$ needs to be taken as a preliminary and illustrative example of how one can use the limit on A to restrict the size of the actual LV operators.

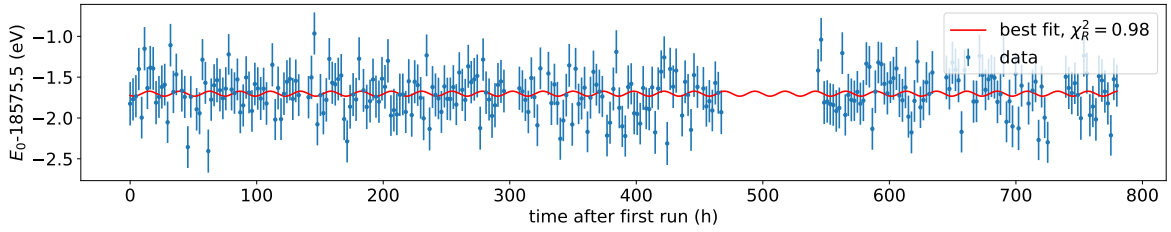


Figure 6.3: This figure shows the endpoints of the KNM-1 runs and the fitted oscillation. The best fit is not significant at 90% C.L..

6.2 Limit on isotropic LV

Besides the anisotropic part, there are also isotropic effects on the Tritium spectrum due to LV. These can be studied by examining the time-averaged endpoint of the Tritium-spectrum, meaning a combination of all 274 KNM-1 runs in a combined fit.

6.2.1 Values for the KATRIN and the Mainz experiment

As explained in section 3.4.1 at least two different experiments are needed to constrain the isotropic parts of LV. For this thesis the values of KATRIN and its predecessor in Mainz will be used, as both experiments make or made use of a MAC-E filter and the decay of Tritium. The coordinates of the experiments, as well as their acceptance angle, are listed in table B.1. Hereby one can notice that the co-latitude of Mainz and Karlsruhe is similar. However, the orientation to the local north and the acceptance angle of both experiments differ, which allows for this analysis.
The Mainz experiment quotes $E_{0,\text{theo}} = (18\,574.3 \pm 1.7)\,\text{eV}$ for the theoretical and $E_{0,\text{exp}} = (18\,576.6 \pm 0.2)\,\text{eV}$ for the experimentally determined endpoint [21].
For the KATRIN experiment the Q-values were provided and therefore the difference of the endpoints corresponds to the difference of the Q-values. In the publication of the results of KNM-1 the KATRIN collaboration refers to a theoretical value of $Q_{\text{theo}} = (18\,575.72 \pm 0.07)\,\text{eV}$ while the experimental result reads $Q_{\text{exp}} = (18\,575.2 \pm 0.5)\,\text{eV}$ [27].

6.2.2 Calculated limits

We make use of the different acceptance angles of the two experiments, solve the system of equations and perform a Gaussian error propagation, which leads to

$$k \times a^0 = (4.4 \pm 3.0)\,\text{eV} \tag{6.3}$$
$$l \times a_3 = (-7.7 \pm 4.8)\,\text{eV} \tag{6.4}$$

where k and l are KATRIN's prefactors due to LV. To check if the best fit is comparable with the Null-hypothesis of no LV, one can calculate the χ^2 of the best fit and use the $\Delta\chi^2 = \chi^2(0,0) - \chi^2(\text{best fit}) = 2.9$ which corresponds to a 1.2σ fluctuation. Thus this best fit is not significant at the 90 % C.L..
If one does the calculation for the 90% confidence level assuming a Gaussian uncertainty the limits are

$$\begin{aligned} -0.5\,\text{eV} < k \times a^0 < 9.4\,\text{eV} \\ -15.7\,\text{eV} < l \times a_3 < 0.3\,\text{eV}. \end{aligned} \tag{6.5}$$

Applying eq. (3.48), we can figure out that $k = \gamma \approx 1$ while $l \approx B \approx 4.7 \times 10^{-5}$. For better comparison a^0 and a^3 will be transformed into spherical decomposition via

$$\begin{aligned} a^0 &= \sqrt{\frac{1}{4\pi}}\, a_{00}^{(3)} \\ a^3 &= -\sqrt{\frac{3}{4\pi}}\, a_{10}^{(3)}. \end{aligned} \tag{6.6}$$

The found limits at 90% C.L. are therefore

$$\boxed{-1.9 \times 10^{-9}\,\text{GeV} < a_{00}^{(3)} < 3.3 \times 10^{-8}\,\text{GeV}}$$
$$\boxed{-1.3 \times 10^{-5}\,\text{GeV} < a_{10}^{(3)} < 6.8 \times 10^{-4}\,\text{GeV}} \tag{6.7}$$

In contrast to the limits in [9] these are derived using the updated formula from section 3.3.

6.2.3 Limitations and improvements of this analysis

The limits presented above depend on the suppression factors, the values of which are still to be calculated and published by the theorists (at the time of writing this thesis).
The presented result relies only on two different experiments. The quality of the exclusion could be improved by further experiments. Additionally future measurement campaigns of the KATRIN experiment would provide better Q-values and therefore increase the sensitivity. Another optimization would include a reevaluation of the Mainz data using new results for the theoretical prediction. However, this is beyond the scope of this thesis.

6.3 Influence on neutrino mass search

As LV is an effect that is not included in the standard analysis of the KATRIN neutrino mass fit, this negligence could bias the neutrino mass result. In this section we will estimate the shift on the m_ν^2 due to the anisotropic effects of LV. This is done using the formula [22]

$$\left\langle \delta m^2 \right\rangle = -2 \left\langle (\delta E_0)^2 \right\rangle \tag{6.8}$$

where δE is the broadening of the spectrum caused by the LV-induced endpoint oscillation. For the purpose of calculating this, one has to find the variance of the oscillation caused by the LV. This variance is proportional to the amplitude of the LV A via

$$\boxed{\left\langle \delta m^2 \right\rangle = -A^2}. \tag{6.9}$$

Therefore it is obvious that the neutrino mass bias is proportional to the fitted A. We will use the maximal value of the 90% C.L. of our derived limit to estimate the influence on the neutrino mass search.

$$\boxed{\left\langle \delta m^2 \right\rangle \approx -6 \times 10^{-3}\,\mathrm{eV}^2} \tag{6.10}$$

This result is approximately 25% of the statistical uncertainty aimed in the TDR [19] and 0.5 % of the uncertainty of KNM-1 result. Thus LV seems not to harm the search for the neutrino mass of KNM-1. This can be also estimated using the sensitivity of KNM-1 and of the expected final KATRIN experiment (c.f. table 6.2).

Table 6.2: This table shows the influence of LV on the neutrino mass. It therefore assumes a LV in the order of 90% C.L. of the sensitivity using a data set without LV and compares it to the uncertainty of KNM-1 and the statistical uncertainty of the TDR.

data	$\Delta \mathrm{m}_\nu^2$	$\frac{\Delta \mathrm{m}_\nu^2}{\Delta \mathrm{m}_{\nu,\mathrm{KMN}-1}^2}$	$\frac{\Delta \mathrm{m}_\nu^2}{\Delta \mathrm{m}_{\nu,\mathrm{TDR}}^2}$
KNM-1 (40 V)	$-2 \times 10^{-3}\,\mathrm{eV}^2$	0.2 %	11 %
750 d KNM-1 (40 V)	$-6 \times 10^{-5}\,\mathrm{eV}^2$		0.3 %

Chapter 7

Summary and Conclusion

In this thesis the capability of the KATRIN experiment to determine a^μ-type LV is investigated. The anisotropic part of the a^μ-LV theory can manifest itself in an oscillatory behavior of the endpoint of the Tritium spectrum with sidereal frequency. The KATRIN experiment is able to fit the endpoint of the spectrum for each run, which has a duration of two hours. Therefore it provides the possibility to analyze its change over the time of the first measurement campaign and can be used to examine a^μ-type LV.
The endpoint, that is fitted for a single run, is an average over the duration of the run. However, a non-weighted average is not a sufficient approximation. In the course of this work a description of an "effective time", to which each endpoint corresponds to, was developed using MC simulations. The effects of different values of LV on the effective time were evaluated and it has become evident that this description is applicable as the size of the effective time is constant for different LV-values. Furthermore it was shown that the performance of this model does not harm the analysis for reasonable values of LV.
In the search for LV, the set of endpoints at the effective time was used to fit a sidereal oscillation with free amplitude A and phase ϕ. In this fit each endpoint has an uncertainty which consists of a statistical and a systematic part. Since the analysis focuses on the change of the endpoint from one run to another, the stability of the uncertainties was analyzed. As a result this thesis shows that the uncertainty budget of the endpoint fits is mainly statistically dominated for the first scientific run of KATRIN.
To investigate the size of a possible oscillation of the endpoints, a frequentist method, namely a χ^2-grid-scan in A and ϕ, was used. In order to constrain the amplitude A Wilks' theorem was used. As the limited applicability of the theorem is known, a detailed coverage study was performed to confirm its suitability for this analysis. Besides that, a Bayesian analysis using an MCMC was performed. In particular the usage of different priors was discussed and the results of the Bayesian and frequentist analysis showed reasonable agreement.
The investigation revealed that the the first measurement campaign of the KATRIN experiment has the sensitivity to limit A

$$|A| \leq 0.04\,\text{eV}\ (90\%\ \text{C.L.}). \tag{7.1}$$

As an additional part of the sensitivity studies the analysis framework was also applied to cases with non-zero LV. It was capable to recover the tested LV values.
After those MC studies, the framework for determining anisotropic effects of LV was applied to the data observed in the first neutrino mass measurement of KATRIN (KNM-1). The best fit, that was found in the analysis, is not significant at 1σ confidence level. Using the presented analysis one can limit the amplitude of an oscillation of the endpoint to be

$$\boxed{|A| \leq 0.068\,\text{eV}}\ 90\%\ \text{C.L.}. \tag{7.2}$$

This can be transformed into a preliminary first limit in terms of the LV parameter $a_{11}^{(3)}$ [37]:

$$\boxed{\left|a_{11}^{(3)}\right| \lesssim 2.7 \times 10^{-6}\,\mathrm{GeV}} \text{ (90\% C.L.).} \tag{7.3}$$

Previously published formulas for the transformation into LV parameters are not applicable for the KATRIN experiment. Therefore a new version was used. Even though I helped with cross-checking the results, they still need to be published and verified by the time of writing this thesis.

In a further investigation the effects of LV on the neutrino mass search at KATRIN were studied and it was shown that the introduced bias is in the sub-percent level compared to the uncertainty of KNM-1. Assuming an increased statistics of a factor of 5.7 and a stable endpoint over the entire run time of KATRIN, the sensitivity to LV would improve to be $A \lesssim 0.008\,\mathrm{eV}$.

Besides anisotropic effects a^{μ}-type LV can also cause isotropic effects. Those can be studied by integrating out the temporal oscillation. This is done using the averaged endpoint of all KNM-1 runs. In this case, LV is expected to cause a global shift of the endpoint. In order to evaluate the value of those isotropic LV parameters, one can use the difference between the experimentally determined Q-value and its theoretical prediction. Together with the results from KATRIN's predecessor in Mainz, we were able to provide an limit for the isotropic parts.

Summa summarum, this thesis presented an analysis framework for the search for a^{μ}-type LV at the KATRIN experiment. This framework was thoroughly tested and used to show the capability of KATRIN to set a first limit for those LV parameters. Using the data set of the first neutrino mass measurement campaign a preliminary limit was found. Once the exact prefactors of the LV operators have been cross-checked by theory, the analysis presented here can lead to the first results constraining the size of those LV parameters. Furthermore, an increase of runs would allow KATRIN to be more sensitive to LV effects on the sidereal basis and search for yearly effects. Possibly this could give new insights into Lorentz invariance violation in the context of physics beyond the Standard Model.

Bibliography

[1] W. Pauli, "Pauli letter collection: letter to Lise Meitner." Typed copy.

[2] E. Noether, "Invariant Variation Problems," *Gott. Nachr.* **1918** (1918) 235–257, arXiv:physics/0503066.

[3] A. Zee, *Quantum Field Theory in a Nutshell; 1st ed.* Nutshell handbook. Princeton Univ. Press, Princeton, NJ, 2003. https://cds.cern.ch/record/706825.

[4] H. A. Lorentz, *La Théorie Électromagnétique de Maxwell et Son Application Aux Corps Mouvants*, pp. 164–343. Springer Netherlands, Dordrecht, 1936. https://doi.org/10.1007/978-94-015-3447-5_4.

[5] A. A. Michelson and E. W. Morley, "On the relative motion of the Earth and the luminiferous ether," *American Journal of Science* **s3-34** no. 203, (1887) 333–345. https://www.ajsonline.org/content/s3-34/203/333.

[6] R. J. Kennedy and E. M. Thorndike, "Experimental establishment of the relativity of time," *Phys. Rev.* **42** (Nov, 1932) 400–418. https://link.aps.org/doi/10.1103/PhysRev.42.400.

[7] D. Mattingly, "Modern Tests of Lorentz Invariance," *Living Reviews in Relativity* **8** no. 1, (Sep, 2005) . http://dx.doi.org/10.12942/lrr-2005-5.

[8] D. Colladay and V. A. Kostelecký, "Lorentz-violating extension of the standard model," *Physical Review D* **58** no. 11, (Oct, 1998) . http://dx.doi.org/10.1103/PhysRevD.58.116002.

[9] J. S. Díaz, V. A. Kostelecký, and R. Lehnert, "Relativity violations and beta decay," *Phys. Rev. D* **88** (Oct, 2013) 071902. https://link.aps.org/doi/10.1103/PhysRevD.88.071902.

[10] V. A. Kostelecký and M. Mewes, "Neutrinos with Lorentz-violating operators of arbitrary dimension," *Phys. Rev. D* **85** (May, 2012) 096005. https://link.aps.org/doi/10.1103/PhysRevD.85.096005.

[11] E. Fermi, "Versuch einer Theorie der β-Strahlen. I," *Zeitschrift für Physik* **88** no. 3, (Mar, 1934) 161–177. https://doi.org/10.1007/BF01351864.

[12] C. L. Cowan, F. Reines, F. B. Harrison, *et al.*, "Detection of the Free Neutrino: a Confirmation," *Science* **124** no. 3212, (1956) 103–104. https://science.sciencemag.org/content/124/3212/103.

[13] B. Povh, K. Rith, C. Scholz, *et al.*, *Teilchen und Kerne*. Springer Berlin Heidelberg, 2014. https://doi.org/10.1007/978-3-642-37822-5.

[14] R. Davis, D. S. Harmer, and K. C. Hoffman, "Search for Neutrinos from the Sun," *Phys. Rev. Lett.* **20** (May, 1968) 1205–1209. https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[15] **Super-Kamiokande** Collaboration, Y. Fukuda, T. Hayakawa, E. Ichihara, *et al.*, "Evidence for Oscillation of Atmospheric Neutrinos," *Phys. Rev. Lett.* **81** (Aug, 1998) 1562–1567. https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[16] B. Pontecorvo, "Mesonium and anti-mesonium," *Sov. Phys. JETP* **6** (1957) 429.

[17] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles," *Progress of Theoretical Physics* **28** no. 5, (11, 1962) 870–880, https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf. https://doi.org/10.1143/PTP.28.870.

[18] C. Karl, "Analysis of First Tritium Data of the KATRIN Experiment," Master's thesis, Technical University Munich, https://www.katrin.kit.edu/publikationen/mth-karl.pdf, 2018.

[19] KATRIN Collaboration and KATRIN Collaboration, "KATRIN design report 2004," tech. rep., Forschungszentrum, Karlsruhe, 2005. 51.54.01; LK 01.

[20] V. N. Aseev, A. I. Belesev, A. I. Berlev, *et al.*, "Upper limit on the electron antineutrino mass from the Troitsk experiment," *Physical Review D* **84** no. 11, (Dec, 2011) . http://dx.doi.org/10.1103/PhysRevD.84.112003.

[21] C. Kraus, B. Bornschein, L. Bornschein, *et al.*, "Final results from phase II of the Mainz neutrino mass searchin tritium β decay," *The European Physical Journal C* **40** no. 4, (Apr, 2005) 447–468. http://dx.doi.org/10.1140/epjc/s2005-02139-7.

[22] E. W. Otten and C. Weinheimer, "Neutrino mass limit from tritium β decay," *Reports on Progress in Physics* **71** no. 8, (Jul, 2008) 086201. https://doi.org/10.1088/0034-4885/71/8/086201.

[23] M. Kleesiek, J. Behrens, G. Drexlin, *et al.*, "β -Decay spectrum, response function and statistical model for neutrino mass measurements with the KATRIN experiment," *The European Physical Journal C* **79** no. 3, (Mar, 2019) 204. https://doi.org/10.1140/epjc/s10052-019-6686-7.

[24] M. Babutzka, *Design and development for the Rearsection of the KATRIN experiment.* PhD thesis, 2014.

[25] M. Arenz, W.-J. Baek, M. Beck, *et al.*, "First transmission of electrons and ions through the KATRIN beamline," *Journal of Instrumentation* **13** no. 04, (Apr, 2018) P04020–P04020. https://doi.org/10.1088/1748-0221/13/04/p04020.

[26] K. M. Altenmüller, *Search for sterile neutrinos in beta decays.* Dissertation, Technische Universität München, München, 2019.

[27] M. Aker, K. Altenmüller, M. Arenz, *et al.*, "Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN," *Physical Review Letters* **123** no. 22, (Nov, 2019) . http://dx.doi.org/10.1103/PhysRevLett.123.221802.

[28] A. V. Lokhov and F. V. Tkachov, "Confidence intervals with a priori parameter bounds," *Physics of Particles and Nuclei* **46** no. 3, (May, 2015) 347–365. https://doi.org/10.1134/S1063779615030089.

[29] M. Aker, K. Altenmüller, A. Beglarian, *et al.*, "Bound on 3+1 Active-Sterile Neutrino Mixing from the First Four-Week Science Run of KATRIN," *Physical Review Letters* **126** no. 9, (Mar, 2021) . http://dx.doi.org/10.1103/PhysRevLett.126.091803.

[30] G. Mention, M. Fechner, T. Lasserre, *et al.*, "Reactor antineutrino anomaly," *Phys. Rev. D* **83** (Apr, 2011) 073006. https://link.aps.org/doi/10.1103/PhysRevD.83.073006.

[31] C. Giunti and T. Lasserre, "eV-Scale Sterile Neutrinos," *Annual Review of Nuclear and Particle Science* **69** no. 1, (2019) 163–190. https://doi.org/10.1146/annurev-nucl-101918-023755.

[32] V. A. Kostelecký and M. Mewes, "Signals for Lorentz violation in electrodynamics," *Phys. Rev. D* **66** (Sep, 2002) 056005. https://link.aps.org/doi/10.1103/PhysRevD.66.056005.

[33] **LSND** Collaboration, L. B. Auerbach, R. L. Burman, D. O. Caldwell, *et al.*, "Tests of Lorentz violation in $\overline{\nu}_\mu \to \overline{\nu}_e$ oscillations," *Phys. Rev. D* **72** (Oct, 2005) 076004. https://link.aps.org/doi/10.1103/PhysRevD.72.076004.

[34] J. Díaz, T. Katori, J. Spitz, and J. Conrad, "Search for neutrino–antineutrino oscillations with a reactor experiment," *Physics Letters B* **727** no. 4-5, (Dec, 2013) 412–416. http://dx.doi.org/10.1016/j.physletb.2013.10.058.

[35] V. A. Kostelecký and J. D. Tasson, "Prospects for Large Relativity Violations in Matter-Gravity Couplings," *Phys. Rev. Lett.* **102** (Jan, 2009) 010402. https://link.aps.org/doi/10.1103/PhysRevLett.102.010402.

[36] J. Byrne, P. G. Dawber, M. G. D. van der Grinten, *et al.*, "Determination of the electron anti-neutrino angular correlation coefficient a_0 and the parameter $|\lambda| = |G(A)/G(V)|$ in free neutron beta decay from measurements of the integrated energy spectrum of recoil protons stored in an ion trap," *J. Phys. G* **28** (2002) 1325–1349.

[37] V. A. Kostelecký and N. Russell, "Data tables for Lorentz and CPT violation," *Reviews of Modern Physics* **83** no. 1, (Mar, 2011) 11–31. http://dx.doi.org/10.1103/RevModPhys.83.11.

[38] D. Colladay and V. A. Kostelecký, "CPT violation and the standard model," *Physical Review D* **55** no. 11, (Jun, 1997) 6760–6774. http://dx.doi.org/10.1103/PhysRevD.55.6760.

[39] J. S. Díaz, "Tests of Lorentz Symmetry in Single Beta Decay," *Advances in High Energy Physics* **2014** (2014) 1–7. http://dx.doi.org/10.1155/2014/305298.

[40] Y. Akulov and B. Mamyrin, "Half-life and fT1/2 value for the bare triton," *Physics Letters B* **610** no. 1, (2005) 45–49. https://www.sciencedirect.com/science/article/pii/S0370269305001784.

[41] R. Lehnert, "Beta Spectrum with Lorentz Violation." Private communication, 2021.

[42] V. A. Kostelecký and M. Mewes, "Signals for Lorentz violation in electrodynamics," *Physical Review D* **66** no. 5, (Sep, 2002) . http://dx.doi.org/10.1103/PhysRevD.66.056005.

[43] R. Sack, *Measurement of the energy loss of 18.6 keV electrons on deuterium gas and determination of the tritium Q-value at the KATRIN experiment.* [electronic ed.] Auflage, 2020. http://nbn-resolving.de/urn:nbn:de:hbz:6-59069498754.

[44] S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Annals Math. Statist.* **9** no. 1, (1938) 60–62.

[45] **Particle Data Group** Collaboration, K. Olive *et al.*, "Review of Particle Physics," *Chin. Phys. C* **38** (2014) 090001.

[46] P. R. Miles, "pymcmcstat: A Python Package for Bayesian Inference Using Delayed Rejection Adaptive Metropolis," *Journal of Open Source Software* **4** no. 38, (2019) 1417. https://doi.org/10.21105/joss.01417.

[47] H. Haario, M. Laine, A. Mira, and E. Saksman, "DRAM: Efficient adaptive MCMC," *Statistics and Computing* **16** no. 4, (Dec, 2006) 339–354. https://doi.org/10.1007/s11222-006-9438-0.

[48] J. VanderPlas, "Frequentism and Bayesianism: A Python-driven Primer," 2014.

[49] R. Cousins, "Why isn't every physicist a Bayesian?," *American Journal of Physics* **63** (1995) 398–410.

[50] R. Kass and L. Wasserman, "The Selection of Prior Distributions by Formal Rules," *Journal of the American Statistical Association* **91** (1996) 1343–1370.

[51] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **186** no. 1007, (Sept., 1946) 453–461. https://doi.org/10.1098/rspa.1946.0056.

[52] C. P. Robert, N. Chopin, and J. Rousseau, "Harold Jeffreys's Theory of Probability Revisited," *Statist. Sci.* **24** no. 2, (05, 2009) 141–172. https://doi.org/10.1214/09-STS284.

[53] S. Dodelson and F. Schmidt, "14 - Analysis and inference," in *Modern Cosmology (Second Edition)*, S. Dodelson and F. Schmidt, eds., pp. 401–431. Academic Press, second edition Auflage, 2021. https://www.sciencedirect.com/science/article/pii/B9780128159484000206.

[54] H. Jeffreys, *Theory of Probability*. Clarendon Press, 2 Auflage, 1948.

Appendix A

Derivation of the fermion's LV operators

As the focus of this work is on neutrinos one can have a look at non interacting fermions for simplicity. In this case all Lorentz violating terms in the action are fermion bilinears. Generally one can define a $2N$- dimensional spinor-multiplet as

$$\Psi_A = \begin{pmatrix} \psi_a \\ \psi_a^C \end{pmatrix} \tag{A.1}$$

with $A \in [1, 2N]$. In this case the general Lorentz and CPT violation form of the Lagrange density is given by [10]:

$$\mathcal{L} = \frac{1}{2}\bar{\Psi}_A \left(\gamma^\mu i\partial_\mu \delta_{AB} - M_{AB} + \widehat{\mathcal{Q}}_{AB}\right)\Psi_B + \text{h.c.} \tag{A.2}$$

The first term describes the kinetic part and M_{AB} is a general mass matrix. The third term consists mainly of the Lorentz violating matrix $\widehat{\mathcal{Q}}_{AB}$, which is 2×2 in spinor and $2N \times 2N$ in flavor space and includes derivatives like $i\partial_\mu$ [10].
For further convenience one can use the hermiticity of the Lagrange density and rewrite the mass matrix as $M_{AB} = m_{AB} + im_{5AB}\gamma_5$, where m and m_5 are $2N \times 2N$ matrices. Because of the redundancy in the definition of Ψ[1] one gets the following relations [10]:

$$m = \mathcal{C}m^T\mathcal{C}, \quad m_5 = \mathcal{C}m_5^T\mathcal{C} \tag{A.3}$$

Even though a Lorentz violating operator $\widehat{\mathcal{Q}}_{AB}$ can, in general, be dependent on the spacetime position, the following section focuses on operators that conserve momentum and energy[2] to avoid further complexity beyond Lorentz violation. This assumption might neglect e.g. soliton solutions. However, one can think of this model as an approximation for dominant forces [10]. A possible decomposition for $\widehat{\mathcal{Q}}_{AB}$ is an expansion in the basis of the Dirac matrices γ_I [10],

$$\begin{aligned} \widehat{\mathcal{Q}}_{AB} =& \sum_I \widehat{\mathcal{Q}}^I_{AB}\gamma_I \\ =& \widehat{\mathcal{S}}_{AB} + i\widehat{\mathcal{P}}_{AB}\gamma_5 + \widehat{\mathcal{V}}^\mu_{AB}\gamma_\mu + \widehat{\mathcal{A}}^\mu_{AB}\gamma_5\gamma_\mu + \frac{1}{2}\widehat{T}^{\mu\nu}_{AB}\sigma_{\mu\nu}. \end{aligned} \tag{A.4}$$

[1] This implies $\Psi^C = \mathcal{C}\Psi, \quad \mathcal{C} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

[2] Momentum and energy conservation implies that the action is invariant under spacetime translations and therefore $\widehat{\mathcal{Q}}_{AB}$ is spacetime independent.

In this equation $\widehat{\mathcal{Q}}^I_{AB}$ is a derivative dependent $2N \times 2N$ matrix. Its derivative dependence can be expressed via an expansion in operators of mass dimension d:

$$\widehat{\mathcal{Q}}^I_{AB} = \sum_{d=3}^{\infty} \mathcal{Q}^{(d)I\alpha_1\alpha_2\ldots\alpha_{d-3}}_{AB} p_{\alpha_1} p_{\alpha_2} \ldots p_{\alpha_{d-3}} \tag{A.5}$$

where $p_\mu = i\partial_\mu$ and $\mathcal{Q}^{(d)I\alpha_1\alpha_2\ldots\alpha_{d-3}}_{AB}$ has mass dimension $d-4$ [10].
We can split $\widehat{\mathcal{Q}}_{AB}$ into operators of even and odd mass dimension in the following way:

$$\gamma^\nu p_\nu \delta_{AB} - M_{AB} + \widehat{\mathcal{Q}}_{AB} = \widehat{\Gamma}^\nu_{AB} p_\nu - \widehat{M}_{AB} \tag{A.6}$$

The odd part $\widehat{M}_{AB}$ and the even part $\widehat{\Gamma}^\nu_{AB} p_\nu$ can also be expressed in the basis of Dirac matrices [10]:

$$\begin{aligned}
\widehat{\Gamma}^\nu_{AB} =& \gamma^\nu \delta_{AB} + \widehat{c}^{\mu\nu}_{AB}\gamma_\mu + \widehat{d}^{\mu\nu}_{AB}\gamma_5\gamma_\mu + \widehat{e}^\nu_{AB} + i\widehat{f}^\nu_{AB}\gamma_5 + \frac{1}{2}\widehat{g}^{\kappa\lambda\nu}_{AB}\sigma_{\kappa\lambda} \\
\widehat{M}_{AB} =& m_{AB} + im_{5AB}\gamma_5 + \widehat{m}_{AB} + i\widehat{m}_{5AB}\gamma_5 + \widehat{a}^\mu_{AB}\gamma_\mu + \widehat{b}^\mu_{AB}\gamma_5\gamma_\mu + \frac{1}{2}\widehat{H}^{\mu\nu}_{AB}\sigma_{\mu\nu}
\end{aligned} \tag{A.7}$$

As $\widehat{\Gamma}^\nu_{AB} p_\nu$ is contracted with p_μ in equation A.6 it makes sense to introduce the following contracted operators:

$$\widehat{c}^\mu_{AB} = \widehat{c}^{\mu\nu}_{AB} p_\nu, \quad \widehat{d}^\mu_{AB} = \widehat{d}^{\mu\nu}_{AB} p_\nu \quad \widehat{e}_{AB} = \widehat{e}^\nu_{AB} p_\nu, \quad \widehat{f}_{AB} = \widehat{f}^\nu_{AB} p_\nu, \quad \widehat{g}^{\kappa\lambda}_{AB} = \widehat{g}^{\kappa\lambda\nu}_{AB} p_\nu \tag{A.8}$$

Using eq. (A.4), this can be simplified as [10]

$$\begin{aligned}
&\widehat{\mathcal{S}}_{AB} = \widehat{e}_{AB} - \widehat{m}_{AB}, \quad \widehat{\mathcal{P}}_{AB} = \widehat{f}_{AB} - \widehat{m}_{5\mathrm{AB}}, \quad \widehat{\mathcal{V}}^\mu_{AB} = \widehat{c}^\mu_{AB} - \widehat{a}^\mu_{AB} \\
&\widehat{\mathcal{A}}^\mu_{AB} = \widehat{d}^\mu_{AB} - \widehat{b}^\mu_{AB}, \quad \widehat{\mathcal{T}}^{\mu\nu}_{AB} = \widehat{g}^{\mu\nu}_{AB} - \widehat{H}^{\mu\nu}_{AB}.
\end{aligned} \tag{A.9}$$

Appendix B

Location values of the KATRIN and Mainz experiments

Table B.1: This table shows the used values of the KATRIN and Mainz experiment according to [9].

Quantity	KATRIN value	Mainz value
ω	$2\pi/23\,\mathrm{h}56\,\mathrm{min}$	$2\pi/23\,\mathrm{h}56\,\mathrm{min}$
θ_0	51°	62°
χ	41°	40°
ξ	16°	-65°
$\beta_{rot} = r_\oplus \sin\chi$	1×10^{-6}	1×10^{-6}
p_e	$139\,\mathrm{keV}$	$139\,\mathrm{keV}$
m_e	$511\,\mathrm{keV}$	$511\,\mathrm{keV}$

Appendix C

DRAM

DRAM combines two powerful modifications of the well-known Metropolis-Hastings algorithm (Delayed Rejection (DR) and Adaptive Metropolis sampling (AM)). An n-stages DR improves the sampling to start faster by allowing to try n-times to move away from the current position in the Markov-Chain. Assuming that q_i is the proposal at the i^{th} stage this means that the acceptance probability at this stage is [47]

$$\alpha_i\left(x, y_1, \ldots, y_i\right) = 1 \wedge \left\{ \frac{\pi(y_i) q_1(y_i, y_{i-1}) q_2(y_i, y_{i-1}, y_{i-2}) \ldots q_i(y_i, y_{i-1}, \ldots, x)}{\pi(x) q_1(x, y_1) q_2(x, y_1, y_2) \ldots q_i(x, y_1, \ldots, y_i)} \frac{[1-\alpha_1(y_i, y_{i-1})][1-\alpha_2(y_i, y_{i-1}, y_{i-2})] \cdots [1-\alpha_{i-1}(y_i, \ldots, y_1)]}{[1-\alpha_1(x, y_1)][1-\alpha_2(x, y_1, y_2)] \cdots [1-\alpha_{i-1}(x, y_1, \ldots, y_{i-1})]} \right\}. \tag{C.1}$$

The idea of AM is to create a Gaussian proposal function with a covariance matrix that can be calculated by a recursion relation without expensive computational power.
DRAM combines the two methods and was shown to be ergodic and to protect against over and under calibrated proposals [47].

Appendix D

Jeffreys' prior in amplitude and phase

In a description using amplitude A and phase ϕ the Fisher information can be calculated to be

$$|I(A,\phi)| \propto C \times A^2 \propto A^2. \tag{D.1}$$

This implies that the Jeffrey's prior is

$$\pi \propto A. \tag{D.2}$$

For the sake of simple implementation, this can be transformed to a flat prior in A^2 as $\int dAA \propto A^2$.
Performing this analysis leads to the posterior distribution displayed in fig. D.1. The upper limit in this case is a similar exclusion curve like in the case of the "two-amplitude" description (see fig. D.2). It is worth mentioning that the amplitude of zero corresponds to a single point in the "two-amplitude" description and therefore has zero measure. This is analogously translated in the fact that the Jeffreys' prior is proportional to A which is zero for no LV.

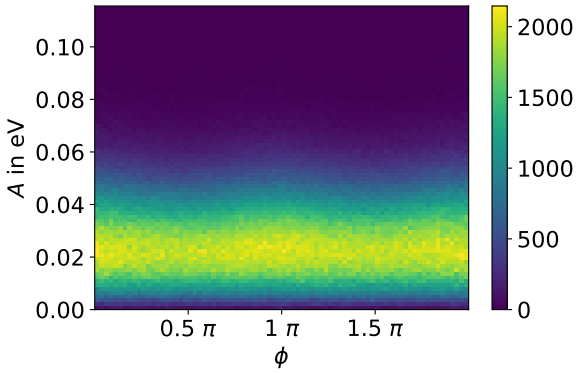


Figure D.1: This figure shows the posterior distribution in amplitude and phase starting with a Jeffrey's prior for MC data with no LV.

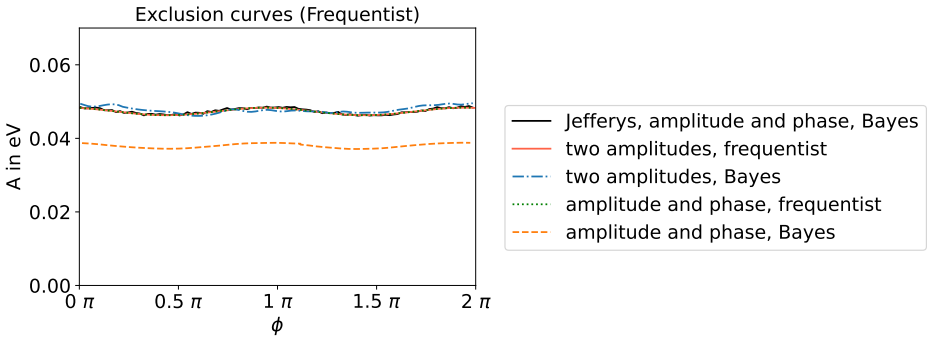


Figure D.2: This figure shows the exclusion curves in amplitude and phase for the Bayesian and frequentist methods. In the case of the Bayesian analysis in A and ϕ the exclusion curves were calculated for flat and Jeffreys' priors. It is visible that the frequentist methods agree as well as the Bayesian one with Jeffreys' prior does. This behavior is expected by construction.

Appendix E

Abbreviations and acronyms

CPS Cryogenic Pumping Section

DPS Differential Pumping Section

FPD Focal Plane Detector

KNM-1 KATRIN neutrino mass 1

KATRIN Karlsruhe Tritium Neutrino experiment

LV Lorentz invariance violation

MAC-E Magnetic Adiabatic Collimation in combination with an Electrostatic Filter

MC Monte Carlo

MCMC Markow Chain Monte Carlo

MS Main Spectrometer

NP Non-Poissonian

PS Pre-spectrometer

SM Standard Model

SME Standard Model extension

WGTS Windowless Gaseous Tritium Source

List of Figures

List of Tables

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit mit dem Titel

Search for Lorentz Violation at the KATRIN Experiment

selbständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

München, 15. April 2021

Johannes Gerald Wickles