Technische Universität München
Fakultät für Physik

Bachelor's Thesis in Physics

Using KATRIN Data to Search for Light Sterile Neutrinos with the Neutrino Mass as a Free Parameter

Suche nach leichten Sterilen Neutrinos mit KATRIN Daten und der Aktiven Neutrinomasse als freien Parameter

Sebastian Münkel[1]

17th July 2023

[1] s.muenkel@tum.de, Fakultät für Physik, TU München, James-Franck-Str. 1, 85748 Garching b. München

Primary reviewer: Prof. Dr. Susanne Mertens
Secondary reviewer:
Supervisors: Xaver Stribl, Christoph Köhler, Dr. Thierry Lasserre

Contents

Introduction

The existence of sterile neutrinos would be a minimal extension to the standard model of particle physics (SM) that has been suggested as a solution to some observed anomalies in neutrino physics. The **KA**rlsruhe **TRI**tium **N**eutrino (KATRIN) experiment is capable of providing strong constraints on the parameter-space of eV-scale sterile neutrinos. As KATRIN nears its final measurement campaign, the collected dataset has grown to the point where a final analysis of the full dataset with conventional modeling of the spectrum is not feasible anymore. As such, Neural Networks (NNs) as a fast and efficient tool to calculate the integral spectrum have been employed to future-proof the analysis of both active and sterile neutrino parameters.

The sterile neutrino manifests itself as a kink-like signature in the β-decay electron energy spectrum, at energies below the endpoint equal to the mass of the sterile eigenstate. Without considering light sterile neutrinos, the tritium beta decay spectrum is modeled with four free parameters: the total normalization, the endpoint energy, the background, and the effective (active) neutrino mass. To account for light sterile neutrinos, an additional fourth neutrino branch is added to the model, described by a new mass state and mixing angle. In the vanilla analysis, the mass of the active neutrino is set to zero. In more comprehensive studies, the mass of the active neutrino is allowed to vary freely and correlates with the mass of the fourth neutrino (mainly sterile). This particular case was not yet implemented in the Netrium neutral network software and is the main topic of this bachelor thesis.

Natural units ($\hbar = c = 1$) are used throughout this work.

Chapter 1

Neutrino Physics

1.1 Neutrinos in the Standard Model

Neutrinos are very light, electrically neutral fermions. They were proposed by Wolfgang Pauli in 1930 in response to unexpected measurements of the β decay electron energy spectrum [1]. At the time, the theory of β decay included only two decay products: the daughter nucleus and a single electron. Therefore, conservation of momentum and energy points to a mono-energetic energy spectrum for each decay product. However, the observed energies of β-decay electrons were in a continuous spectrum which led to two main options to resolve this discrepancy.

The first solution is to disregard energy conservation or charge conservation — conservation laws which had both held for decades prior [2, 3]. The second solution is to introduce a third particle which would take some of the energy and momenta of the decay. A particle that matches the observations would have to be neutral, in order to preserve charge conservation, and massless or very light. The particle that fit these criteria was the newly postulated neutrino.[1]

Neutrinos in the SM are assumed to be massless, charge-less and to come in three flavors [4]. Since it is electrically neutral and only couples via the weak interaction, it is very hard to detect directly.

It is found in various decay and scattering processes, usually accompanied by a lepton. Each flavor corresponds to the lepton-family it couples to. Since there are three charged leptons (electron e^-, muon μ^- and tau lepton τ^-), the neutrino flavors are associated accordingly (electron-neutrino ν_e, muon-neutrino ν_μ, and tau-neutrino ν_τ), along with the corresponding antileptons.

---

[1] The neutrino was originally called "neutron" as it's postulation (1930, Wolfgang Pauli) predates the discovery of the neutral baryon, the neutron (1932, James Chadwick). The name "neutrino" was later popularized by Enrico Fermi.

In this SM view of the neutrino only left-handed neutrinos exist and parity is maximally violated, as shown in experiments by Goldhaber and Wu [5, 6].

1.2 Neutrino Oscillations

While the SM assumes that the neutrino is massless, experimental results from appearance and disappearance studies have shown that at most one neutrino mass-eigenstate can be massless [7–11]. More notably, they also imply that the mass-eigenstates of the neutrino do not match the flavor-eigenstates.

This ability of the neutrino to change flavor as it travels is called *neutrino oscillations* and was first proposed by Bruno Pontecorvo in 1958 [12]. Neutrinos are produced in weak interactions as pure flavor-eigenstates but propagate as their mass eigenstates. If the mass-eigenstates were the same as the flavor-eigenstates, no flavor-appearance or disappearance would occur. Each neutrino would effectively travel as a pure flavor-eigenstate, unaffected by the distance traveled. Furthermore, oscillations also imply that the three mass eigenstates have three distinct masses. The difference in masses is the only way that a coherent superposition of mass eigenstates (corresponding to a specific flavor eigenstate) could change into a different superposition of mass eigenstates, or even become fully decoherent as they travel. This is because each mass eigenstate travels at a speed determined by it's energy and individual mass.

The theory behind these observations is given by the Pontecorvo Maki Nakagawa Sakata (PNMS)-matrix [13].

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \cdot \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.1}$$

This mixing matrix is assumed to be unitary and provides an expression of every flavor-eigenstate $|\nu_\alpha\rangle$ as a superposition of mass-eigenstates $|\nu_i\rangle$

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i} \, |\nu_i\rangle \tag{1.2}$$

and vice versa

$$|\nu_i\rangle = \sum_{\alpha \in \{e,\mu,\tau\}} U^*_{\alpha i} \, |\nu_\alpha\rangle \, . \tag{1.3}$$

Furthermore, a propagating mass-eigenstate is given as a function of its mass, energy and the distance travelled

$$|\nu_j(L)\rangle = \exp\left(-i \cdot (E_j t - \vec{p}_j \cdot \vec{x})\right) |\nu_j(L=0)\rangle \,. \tag{1.4}$$

It follows then, that the probability to observe the flavor transition ($\alpha \rightarrow \beta$) at a given energy and distance is given by

$$\begin{aligned} P_{\alpha\rightarrow\beta} &= |\langle \nu_\beta | \nu_\alpha(L)\rangle|^2 \\ &= \left| \sum_j U^*_{\alpha j} U_{\beta j} \exp\left(-i \cdot \frac{m_j^2 L}{2E}\right) \right|^2 \end{aligned} \tag{1.5}$$

which can be approximated by a two-flavor scenario if only two generations participate significantly

$$U_{\text{two-flavor}} = \begin{pmatrix} U_{\alpha 1} & U_{\alpha 2} \\ U_{\beta 1} & U_{\beta 2} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{1.6}$$

$$P_{\alpha\rightarrow\beta,\alpha\neq\beta} = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \tag{1.7}$$

where θ is the mixing angle between the two flavors, and $\Delta m^2 = m_2^2 - m_1^2$ the mass difference between the two mass-eigenstates.

This is the case if one of the mass differences Δm is significantly larger than the others for distances L and energies E such that $\Delta m^2 L/E = \mathcal{O}(1)$ [14]. Here it also becomes clear that neutrino oscillations imply mass differences between the eigenstates, thereby indicating that at most one massless eigenstate can exist.

1.3 Sterile Neutrinos

Section 1.1 mentions that according to our understanding of the weak interaction, it only couples to the left-handed components of particles and parity is violated maximally. As a result, all neutrinos that have been detected so far have been left-handed neutrinos or right-handed antineutrinos. A minimal extension to the SM would be to introduce right-handed neutrinos. These would not couple via the weak interaction but could interact via their mass or via flavor oscillations and are therefore referred to as *sterile neutrinos.*

Since there is no strong evidence for their existence yet, the mixing angle between the active and sterile branch of neutrino flavors is assumed to be very small. Furthermore, the so-called seesaw mechanism provides motivation to assume that the sterile neutrino is significantly heavier than the active neutrino. The seesaw mechanism would not only provide a reason for the mass of these sterile eigenstates to be a lot heavier than the mass of the active branches, it could also explain the unusually small mass of the active neutrino in relation to the other massive fundamental particles [15]. With these assumptions, the conditions to use the two-flavor approximation seen in equation 1.7 are met, and the active and sterile neutrino can be summarized into one branch each. In principle multiple sterile neutrino flavors could exist, but for this work we will summarize the interaction between the known flavors and a potentially sterile flavor with a two flavor model.

1.4 Experimental Limits

1.4.1 Neutrino Mass

There are multiple ways to measure the active neutrino mass. The three main ones are cosmology, the search for neutrinoless double beta decay ($0\nu\beta\beta$), and the kinematics of β-decay. Each of these measure slightly differently weighted sums of the neutrino mass eigenvalues m_i.

Cosmological probes place limits on the total mass by looking at structure formation in the early universe [16]. A heavier neutrino washes out structures more as it travels, leading to a measurable impact on the structure formation. Current limits published by the Planck mission [17] place the total sum of mass eigenstates at

$$m_\nu = \sum_i m_i < 0.54\,\mathrm{eV}. \tag{1.8}$$

$0\nu\beta\beta$ searches place lower limits on the half-life of $0\nu\beta\beta$, which in turn depends on the effective Majorana mass of the neutrino [18]

$$\frac{1}{T_{0\nu\beta\beta}} \propto m_{\beta\beta}^2. \tag{1.9}$$

The GERmanium Detector Array (GERDA) experiment [19] places limits for the effective Majorana mass $m_{\beta\beta}$ at

$$m_{\beta\beta} = \sum_i U_{ei}^2 m_i < 79 - 180\,\mathrm{meV}. \tag{1.10}$$

Regular beta decay experiments have the benefit of being more model independent than the aforementioned methods. Current results from the KATRIN collaboration place limits on the effective neutrino mass at $m_\beta < 0.8\,\mathrm{eV}$ [20] with

$$m_\beta = \sqrt{\sum_{i=1}^{3} |U_{ei}|^2 m_i^2}. \tag{1.11}$$

The measurement principle behind this measurement is discussed in detail in Sec. 2.1.

1.4.2 Neutrino Oscillation Parameters

The parameters that govern neutrino oscillations – the mixing angles $\sin^2\theta_{ij}$ and the mass differences $\Delta m^2_{ij} = m^2_i - m^2_j$ – are determined through appearance and disappearance experiments. The flavor fractions in Fig. 1.1 are calculated using Eq. 1.5, based on current experimental limits of the oscillation parameters.

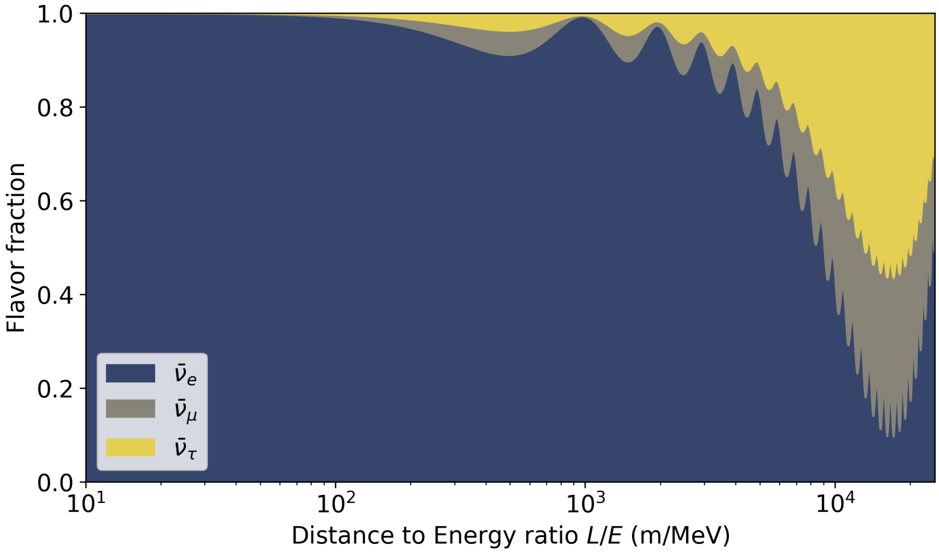


Figure 1.1: Probability to detect each active neutrino flavor as a function of the distance to energy ratio L/E, assuming an electron antineutrino is generated at distance zero and propagates through a vacuum. Based on a figure from [21].

Fig. 1.1 also shows the validity of the two-flavor approximation. For distance to energy ratios $L/E \sim \mathrm{km\,MeV^{-1}}$ the mixing is well approximated by summarizing the muon antineutrino $\bar{\nu}_\mu$ and the tau antineutrino $\bar{\nu}_\tau$ as one combined flavor. Reactor neutrinos tend to have energies in the MeV range [22], so for reactor experiments this figure shows the approximate disappearance ratios of electron antineutrinos $\bar{\nu}_e$ as function of distance L in meters.

1.4.3 Sterile Parameters

Many of the experiments that primarily focus on the neutrino's active branch parameters can also provide constraints on the mass of the sterile mass eigenstate m_4 and the mixing $\sin^2(2\theta_{e4}) = |U_{e4}|^2$ between the electron flavor-eigenstate and the sterile mass-eigenstate.

Table 1.1: Neutrino flavor oscillation parameters, taken from the PDG [23]

Parameter	Value	Comment
$\sin^2\theta_{12}$	0.307(13)	KamLAND+global solar; 3ν
$\sin^2\theta_{23}$	0.547(21)	Assuming inverted ordering
$\sin^2\theta_{23}$	0.545(21)	Assuming normal ordering
$\sin^2\theta_{13}$ (10^{-2} units)	2.18(7)	
Δm^2_{12} (10^{-5} eV^2)	7.53(18)	KamLAND + global solar + SBL + accelerator: 3ν
Δm^2_{23} (10^{-3} eV^2)	−2.546(400)	Assuming inverted ordering
Δm^2_{23} (10^{-3} eV^2)	2.453(340)	Assuming normal ordering
δ (π rad)	1.36(17)	CP violating phase

Fig. 1.2 shows the impact of different combinations of sterile parameters on the transition probability as a function of distance to energy ratio L/E. For an eV-scale sterile neutrino the typical oscillation wavelength is much shorter than that of any transition between the active flavors. As such, anomalies in the measured flux of reactor electron antineutrinos $\bar{\nu}_e$ at short distances can be explained by a sterile neutrino signal [25].

The Neutrino-4 experiment claims the existence of an $m_4 = 7.3\,\text{eV}$ sterile neutrino with a mixing angle θ_{14} of approximately 20 degrees [24]. The black topmost line in Fig. 1.2 shows an approximation of the sterile to active neutrino oscillation as predicted by this result.

This result is in direct opposition to exclusion limits published by other collaborations [26]. Other claims of a sterile neutrino based on a Reactor Antineutrino Anomaly (RAA) lie partially in the area that the KATRIN experiment is sensitive in.

Figs. 1.3 and 1.4 show some current results as well as future sensitivities for light sterile neutrino searches. Overall the parameter-space for a sterile neutrino that matches the observed sterile signals of the Baksan Experiment on Sterile Transitions (BEST) and Neutrino-4 experiments [24, 25] is coming into tension with numerous other results that do not find a sterile signal. Future results from KATRIN will provide results that are independent of the reactor based experiments, and will thus contribute to resolving this tension.

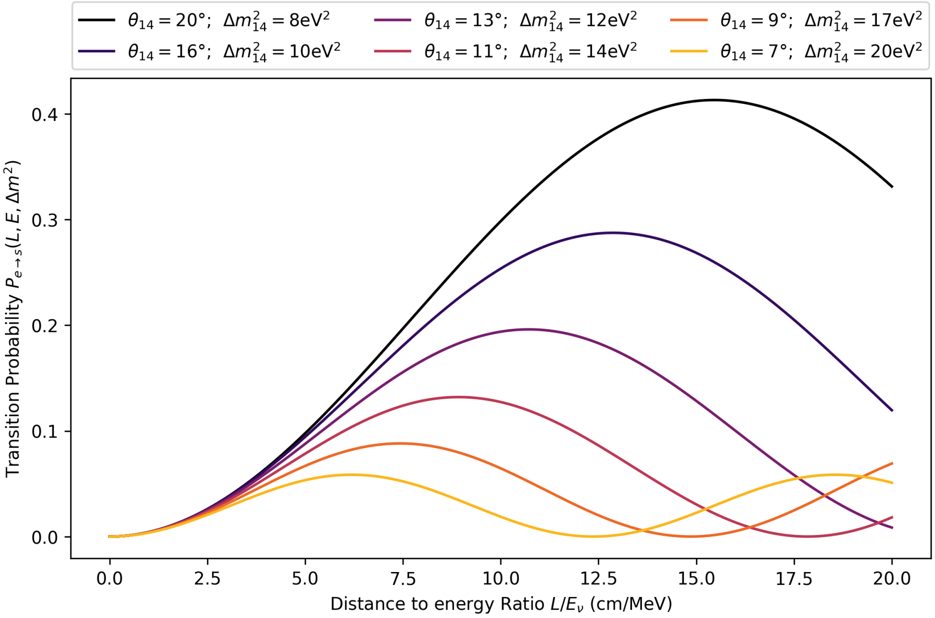


Figure 1.2: Transmission probabilities for an electron (anti-)neutrino to turn into a sterile (anti-)neutrino for various sterile oscillation parameters. The topmost line approximates the best fit results of the Neutrino-4 experiment [24]. Note how the wavelength of the oscillations for reactor neutrinos would be on the sub-meter scale. With many neutrino detector designs requiring very large detector volumes, this makes direct detection of active to sterile neutrino oscillations difficult.

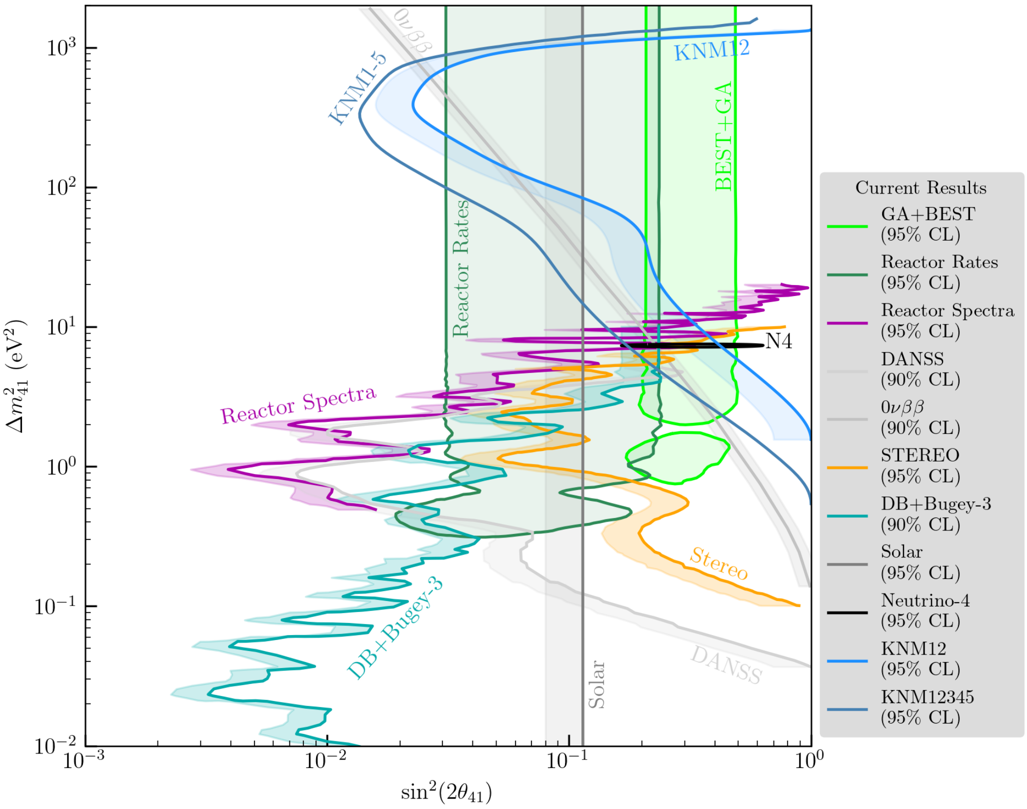


Figure 1.3: Comparisons of various sterile neutrino results. Fully shaded regions denote claims that a sterile neutrino with the respective mass m_4 and mixing $\sin^2(2\theta_{41})$ exist. Lines with a shaded fringe are exclusion contours, with the shaded area pointing towards the area that is excluded. The blue line labeled KNM1-5 is a predicted sensitivity of the combined first five measurement campaigns of the KATRIN experiment, courtesy of Xaver Stribl. Other contours taken from [24–27] using WebPlotDigitizer [28].

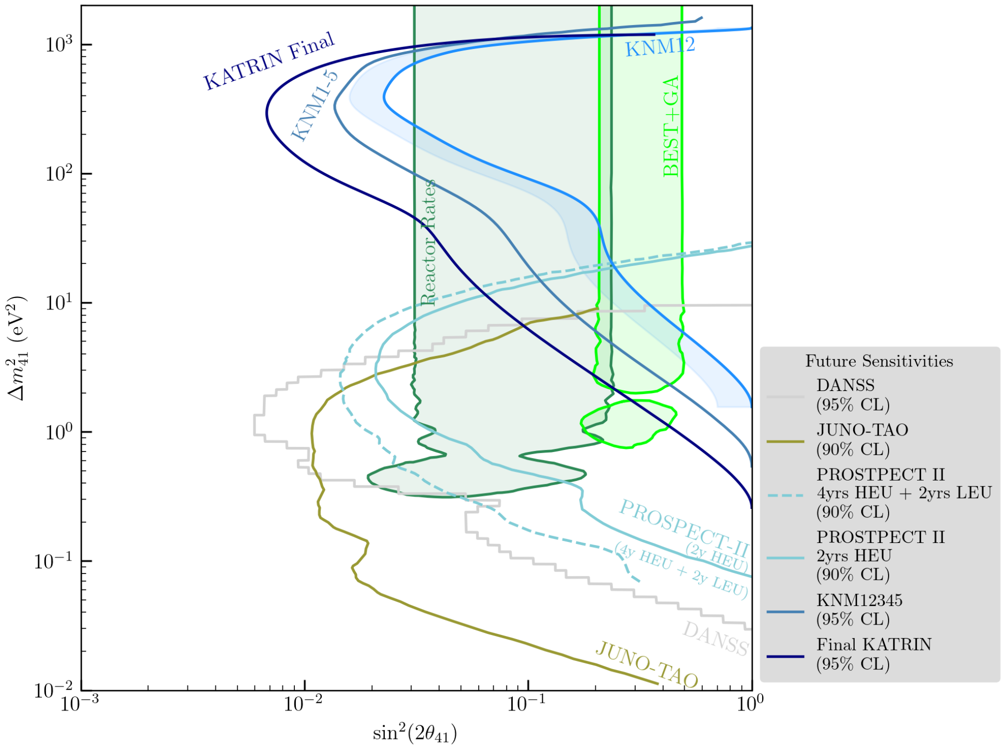


Figure 1.4: Comparisons of various predicted sterile neutrino sensitivities, alongside the RAA and KATRIN results for reference. Fully shaded regions denote claims that a sterile neutrino with the respective mass m_4 and mixing $\sin^2(2\theta_{41})$ exist. Lines with a shaded fringe are exclusion contours, with the shaded area pointing towards the area that is excluded. Lines without any shading are predicted sensitivities. The blue line labeled KNM1-5 and KATRIN Final are predicted sensitivities of the combined first five measurement campaigns of the KATRIN experiment, and of the final sensitivity, courtesy of Xaver Stribl. Other contours taken from [10, 25–27, 29] using WebPlotDigitizer [28].

Chapter 2

The KATRIN experiment

This chapter introduces the KATRIN experiment which forms the basis of this analysis. After discussing the underlying physics and the measurement principle, the experimental setup is described briefly.

2.1 Measurement principle

The KATRIN experiment aims to determine the properties of the electron antineutrino by performing high precision measurements of the β decay spectrum of tritium

$$\mathrm{T}_2 \longrightarrow (^3\mathrm{HeT})^+ + \mathrm{e}^- + \bar{\nu}_e + \mathrm{Q}. \tag{2.1}$$

Q is the energy released via the decay, which then gets divided up to the three decay products. For convenience, we introduce the *endpoint energy* E_0 available to the neutrino and electron

$$E_0 = Q - E_{\mathrm{rec}} = E + E_\nu \tag{2.2}$$

as the difference between the total decay energy Q and the energy of the recoiling $(^3\mathrm{HeT})^+$ nucleus E_{rec}. It is equal to the sum of the energy E_ν of the electron antineutrino $\bar{\nu}_e$ and the energy E of the electron e^-.

Equation 2.2 shows that the electron energy is maximal when the neutrino is emitted at rest. In this case $E_\nu = m_\nu = E_0 - E_{\mathrm{max}}$. This shows that the endpoint of the electron energy spectrum is sensitive to the mass of the neutrino. Specifically, we can calculate the differential rate of the β decay $d\Gamma/dE$ as a function of electron energy

$$\begin{aligned}\frac{d\Gamma}{dE} = \sum_i |U_{ei}|^2 \ \frac{G_F^2 \cos^2\theta_C}{2\pi^3} &\cdot |M_{\text{nuc}}|^2 \cdot F(Z', E) \cdot p \cdot (E + m_e) \\ &\cdot (E_0 - E) \cdot \sqrt{(E_0 - E)^2 - m_i^2} \cdot \Theta(E_0 - E - m_i) \end{aligned} \tag{2.3}$$

with

G_F	Fermi constant,	θ_C	Cabbibo angle,
M_{nuc}	nuclear matrix element,	$\lvert U_{ei}\rvert$	PNMS element,
F	relativistic Fermi function,	Z'	atomic charge of daughter nucleus,
E	kinetic energy of the e$^-$,	p	electron momentum,
m_e	electron mass,	E_0	endpoint energy,
Θ	Heaviside step function,	m_i	neutrino mass-eigenvalue [13, 30].

Using the definition for the effective active neutrino mass m_β from Eq. 1.11 we can approximate Eq. 2.3 in the $3 + 1\nu$ framework as

$$\begin{aligned}\frac{d\Gamma}{dE} \approx \ &\frac{G_F^2 \cos^2\theta_C}{2\pi^3} \cdot |M_{\text{nuc}}|^2 \cdot F(Z', E) \cdot p \cdot (E + m_e) \cdot (E_0 - E) \cdot \\ &\left[\left(1 - |U_{e4}|^2\right) \sqrt{(E_0 - E)^2 - m_\beta^2} \cdot \Theta(\dots) + |U_{e4}|^2 \cdot \sqrt{(E_0 - E)^2 - m_4^2} \cdot \Theta(\dots) \right].\end{aligned} \tag{2.4}$$

The expected spectrum for a sterile neutrino is therefore simply the superposition of two spectra for active neutrinos of different masses. The sterile mass m_4 manifests itself as a kink in the spectrum, exactly m_4 (eV) below the endpoint E_0. Fig. 2.1 shows how different combinations of active and sterile neutrino parameters impact the spectrum at the kink and at the endpoint. It should be noted that the actual measurement is complicated by the fact that there is a background rate of comparable size to the signal at the endpoint of the spectrum.

The next section will discuss how the KATRIN experiment measures this spectrum at its endpoint.

2.2 Experimental Setup

Fig. 2.2 shows an overview of the experimental setup of the KATRIN experiment.

2.2.1 Rear System

The rear wall (RW) is the most important component of the rear system, as it separates the rear system from the WGTS. It is there to properly close

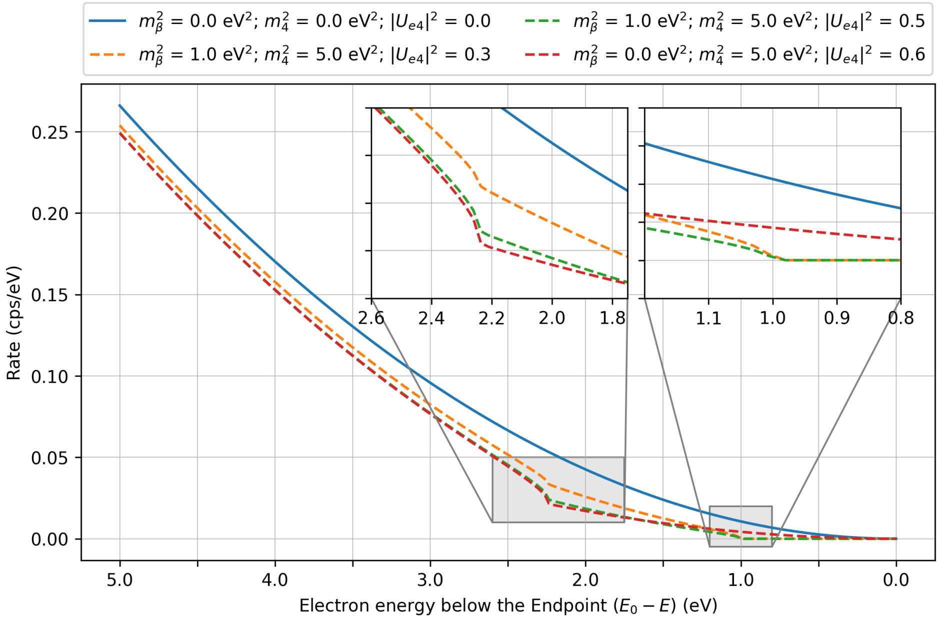


Figure 2.1: Impact of different combinations of active mass m_β, sterile mass m_s and sterile mixing $|U_{e4}|^2$ on the kink-like signature deeper in the differential energy spectrum as well as the endpoint.

off the ultra high vacuum (UHV) system of the windowless gaseous tritium source (WGTS), while also controlling the plasma potential of the tritium gas. Controlling the plasma potential is of critical importance, since any β electrons emitted will also experience any potential present here, and thus either receive or lose energy accordingly. Any inhomogeneity in the surface potential of the RW or the beam tube could therefore lead to spectral distortions in the final measurement. In order to stabilize the potential of the T_2 plasma, the RW is kept at a constant voltage and illuminated with UV radiation. This causes additional photoelectrons to be released from the RW itself, which enhances the conductivity of the T_2 and maintains its quasi-neutrality [32].

Due to this influence of the potentials in the plasma, both the RW and the plasma are considered sources of uncertainty in the final measurement.

The other main feature of the rear system is that it contains the calibration and monitoring section (CMS). The CMS is located behind the RW and contains multiple electron guns and electron detectors. These are used to monitor the flux of β decay electrons and various calibration sources. It has access to the main WGTS via a hole in the RW which allows the electron guns to be fired through.

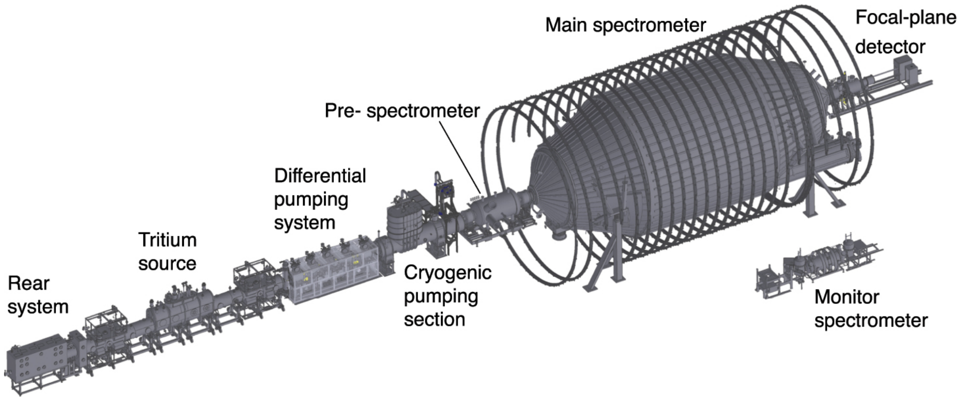


Figure 2.2: Engineering drawing of the KATRIN experimental setup with the main sections labelled. Electrons are released in β decays in the windowless gasseous tritium source (WGTS). These electrons are guided towards the focal plane detector (FPD). Along the way, the electrons are separated from the T_2 molecules in the two pumping sections. The pre-spectrometer is in essence a coarse high pass filter to reduce the rate of background events. The 10m wide main spectrometer is crucial in determining the energy of signal electrons. Electrons are counted in the FPD. Taken from [31].

2.2.2 Tritium Source

The WGTS houses a large amount of tritium on the gram scale. Due to its low half-life of 12 years [33], T_2 plasma it is an exceptionally active source of β-decay electrons. The tritium gas is pumped in a closed loop. It enters the windowless source tube through a series of small holes at the center, and is pumped out at both ends [34].

This leads to an inhomogeneous density profile that peaks in the middle. Understanding this column density ρ_D is important because β-decay electrons will scatter on the T_2 molecules and lose some energy which can distort the spectrum. This energy loss is measured and modeled, such that it can be correctly accounted for in the final spectrum.

In a similar vein, the temperature and density need to be kept stable in order to maintain a predictable impact on the spectrum.

2.2.3 Differential and Cryogenic Pumping Systems

The differential and cryogenic pumping systems – also together known as the transport section – reduce the T_2 flow rate by more than 12 orders of magnitude [32] while transporting the electrons adiabatically towards the spectrometer.

The differential pumping section uses a combination of turbo-molecular pumps to reduce T_2 flow by 5 orders of magnitude. The differential pumping section is unable to reduce the flow by 12 magnitudes by itself. Likewise, the cryogenic pumping section is unable to handle the initial T_2 pressure. As such the two complement each other, and the cryogenic pumping section reduces the flow by another 7 orders of magnitude by using a cold argon frost system which is held at 3 K to 4 K [32].

2.2.4 Pre- and Main Spectrometer

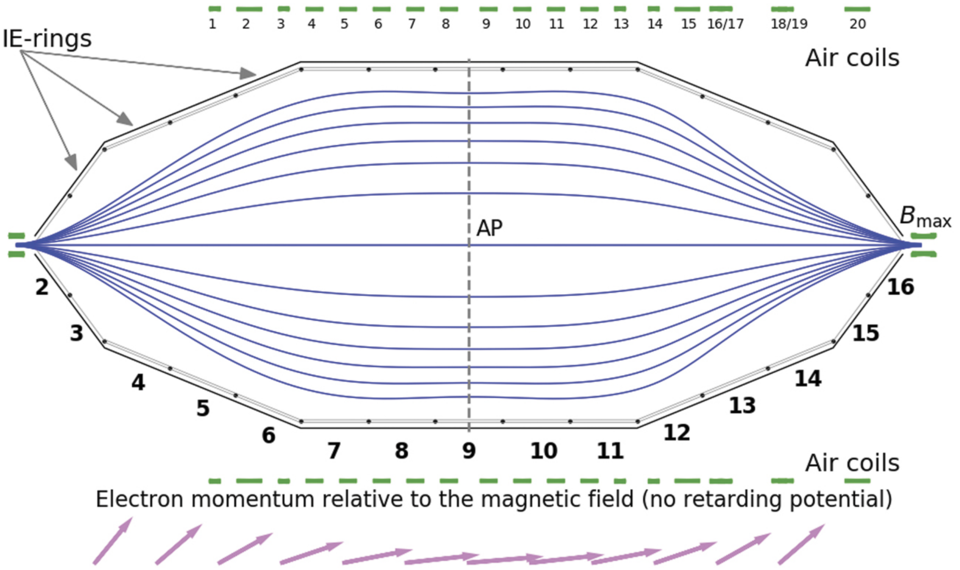


Figure 2.3: The magnetic adiabatic collimation with electrostatic filtering (MAC-E) filter acts as a high-pass energy filter. Electrons move adiabatically along the magnetic field lines drawn in blue. The magnetic fields are very strong at either end and very weak in the middle. This collimates the electron momenta to bring them in parallel to the electric retarding potential at the middle of the MAC-E filter, the analyzing plane (AP). Only electrons with energies great enough to pass the retarding potential reach the detector, thus allowing the measurement of the integral spectrum by measuring the rate of electrons at various retarding potentials. Figure taken from [32].

The working principle of the two large-volume MAC-E filters is depicted in Fig. 2.3. The pre-spectrometer is used to reduce the background rate, by providing a coarse first high pass filter, thus reducing the rate arriving at the main spectrometer significantly. However, this also leads to the formation of a Penning trap in between the two spectrometers. Penning discharges are a significant source of background events. Furthermore, electrons from the spectrometer surface can be induced by ambient radiation sources.

The main spectrometer filters the collimated electrons by having a precisely set retarding potential. How well the incoming electrons are collimated directly depends on the ratio between the magnetic fields at the entrance and at the analyzing plane, such that

$$\frac{\Delta E}{E} = \frac{B_{\mathrm{min}}}{B_{\mathrm{max}}}. \tag{2.5}$$

2.2.5 Focal-Plane Detector

The FPD consists of an array of 148 pixels. Since the main spectrometer provides a high pass filter, the FPD merely needs to count the incoming electrons in order to be able to measure an integral spectrum.

Further details on all the components can be found in the design report [32].

Chapter 3

Analysis

This chapter discusses the impact of adding the active neutrino mass m_β as a free parameter to the model of the integral spectrum. Adding another parameter will decrease the sensitivity of the model in certain parts of the parameter-space, however fixing it to zero can also bias the result of the sterile analysis. The following sections will introduce the analysis method, explain the procedure and the different test cases that were investigated, and compare some ways to find a compromise between accuracy in the model and a loss of sensitivity. Finally the impact of the active neutrino mass on the other fit parameters and the systematics will be discussed. The basis of this analysis is the second KATRIN measurement campaign.

3.1 Netrium

Netrium is an analysis toolkit developed by Christian Karl in order to speed up the analysis of the final KATRIN dataset by about three orders of magnitude compared to conventional methods. The KATRIN physics model is approximated using a NN which is trained on one to ten million sample β-decay electron energy spectra. These spectra are generated using the analytical model described in Sec. 2.1.

3.1.1 Neural Networks

NNs are a class of computational models which are inspired by the architecture of our brains. The theoretical groundwork for the models we use today was developed in the late 19th century [35, 36], with first implementations being developed in the mid 20th century [37] and deployed in the 90s [38]. Using NN for calculations and data classifications in physics has been done right from the start [39], but has seen a significant increase in usage in recent years [40, 41] as a result of advancements in computer hardware [42].

In basic terms a NN is a way to generate a fast and precise interpolation function, even for high-dimensional input or output. Extrapolation is also possible, but often inaccurate and thus not reliable [42]. However since we can generate sample energy spectra at any given point in the parameter-space, extrapolation should never be necessary, so long as the correct set of training data is chosen.

3.1.2 Implementation for KATRIN

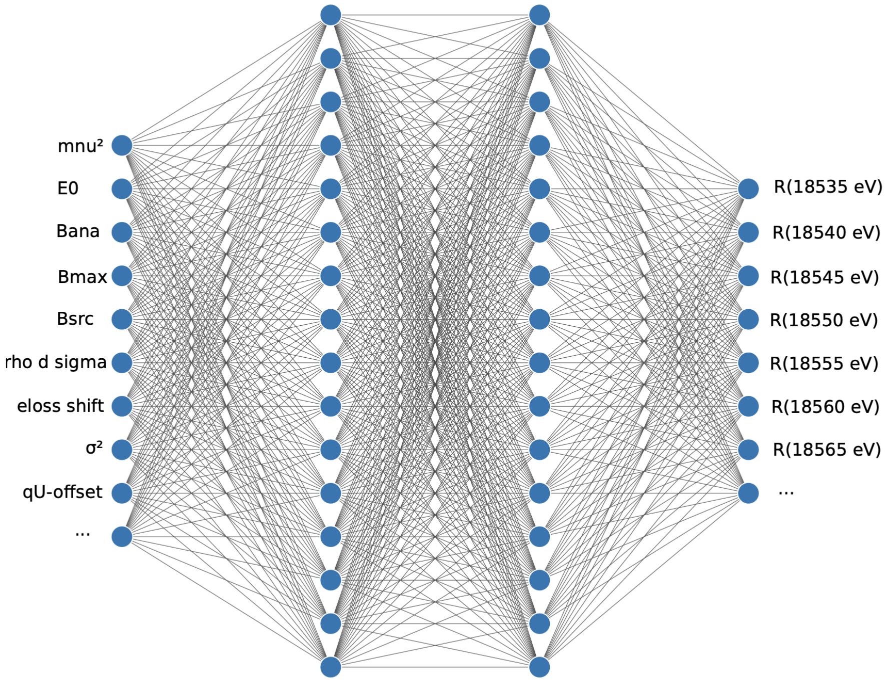


Figure 3.1: Structure of the NN used in the Netrium analysis toolkit. A total of four layers are used: one input layer with one node per parameter, two fully connected hidden layers with 128 nodes each, and one output layer with one node for each point in the spectrum. Figure taken from [43].

Fig. 3.1 shows the general architecture of the NN used in Netrium. This NN is used to precisely approximate the β-decay integral count rate

$$R(qU) = \int_{qU}^{E_0} \frac{d\Gamma}{dE} \cdot f(E; qU) dE \tag{3.1}$$

where qU is the retarding energy and $f(E; qU)$ is the response function as defined in [44]. The response function gives the probability that an electron with energy E is going to pass the retarding energy qU of the MAC-E-filter. This is then combined with a normalization factor A_{sig} and a constant background B to form the expected model rate

$$\mu(\theta; qU) = A_{\text{sig}} \cdot R(\theta; qU) + B \tag{3.2}$$

where θ denotes the parameters that can impact the spectrum. The distributions for these parameters are shown in Tab. 3.1. In the following sections the only parameter distribution that will be changed is the active neutrino mass squared m_β^2.

Table 3.1: Distributions for each of the input parameters that are used to generate the samples. In order to be able to isolate the impact of different choices for the active neutrino mass squared, all other parameter distributions were kept the same.

Parameter	Distribution	μ	σ
ActiveNeutrinoMassSquared	*varies*	*varies*	*varies*
SterileNeutrinoMassSquared	Log	0.010 000 00	9000.000 000 00
SterileNeutrinoMixingAngle	Log	0.000 100 00	0.600 000 00
Endpoint	Uniform	18 573.400 000 00	18 574.000 000 00
MagneticFieldSource	Normal	2.506 500 00	0.006 000 00
RhodSigma	Normal	1.538 519 64	0.005 951 59
ElossShift	Normal	0.000 000 00	0.022 950 00
KNM1Amp1	Normal	0.032 800 00	0.001 209 00
KNM1Pos1	Normal	11.918 900 00	0.008 331 00
KNM1Sig1	Normal	0.183 600 00	0.006 970 00
KNM1Amp2	Normal	0.295 700 00	0.000 676 81
KNM1Pos2	Normal	12.804 600 00	0.002 121 90
KNM1Sig2	Normal	0.467 700 00	0.002 178 60
KNM1Amp3	Normal	0.075 750 00	0.000 368 08
KNM1Pos3	Normal	14.967 700 00	0.004 051 30
KNM1Sig3	Normal	0.907 000 00	0.012 777 00
MagneticFieldAna	Normal	0.000 630 75	0.000 006 31
MagneticFieldMax	Normal	4.239 126 57	0.004 239 13
FSDBroadeningVariance	Normal	0.005 968 42	0.001 161 45
RetardingEnergyOffset	Normal	−1.906 834 19	0.001 000 00

3.2 Choice of Training Data

The Netrium analysis toolkit needs a large number of sample KATRIN β-decay spectra (sample-sets) in order to train NNs that we can use in our analysis. These sample spectra will have to vary in their parameter inputs over ranges that match the expected values to be used as model inputs in our fit. For example, if the true best fit value for m_β^2 at a given point is $10\,\text{eV}^2$ but the NN has only been trained on values that range from $-5\,\text{eV}^2$ to $5\,\text{eV}^2$, inputing values close to $10\,\text{eV}^2$ will not produce accurate spectra as the model will be

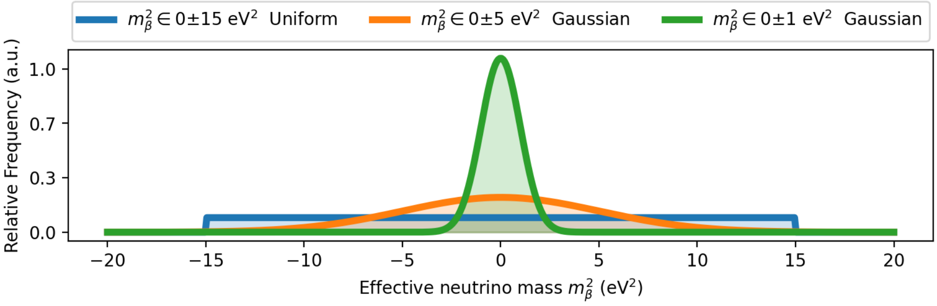


Figure 3.2: Distribution functions for the different sets of sample spectra used in this work.

unable to extrapolate that far. In this case it would not be possible to tell if the best fit is at $10\,\mathrm{eV}^2$ because the fit at that value is inaccurate. Therefore the performance of the NN and the best fit parameters that are determined by the NN are highly dependent on the sample-sets used.

There is a tradeoff between breadth and precision here. A NN that has been trained on a very wide range of input values will be able to account for a wider range of scenarios. However if we know that the most reasonable value for a fit parameter across the entire parameter space is in a specific region we can gain precision by restricting the sample-set on a narrow distribution around this specific region. Furthermore, avoiding unnecessary or unreasonable parameter inputs for the training-set saves on computational resources, as the generation of these sample-sets is very computationally expensive.

Fig. 3.2 shows the distributions of the active mass m_β^2 in the different sample-sets used in this section. Previous publications [27] show that the preferred value for m_β^2 when freely fitted can range from $-10\,\mathrm{eV}^2$ to $10\,\mathrm{eV}^2$. This is why in initial testing, two broad distributions were chosen to evaluate the impact of uniform vs Gaussian distributions. A table showing more granular values for parameters of these sample-sets can be found in the appendix, in Tab. A.1.

3.3 Sensitivity Curves

Sensitivity curves are calculated by determining the agreement of the model spectra with the null hypothesis spectrum for different combinations of input parameters.

More specifically, the null hypothesis is given by an Asimov spectrum. For our fit parameters, expected values are well known for most of the nuisance parameters. Nuisance parameters in this case are all of the parameters that impact the spectrum, such as endpoint energy E_0, background, normalization, column density ρ_0 etc. The Asimov spectrum is an unfluctuated electron β-decay energy spectrum with all of the input parameters set to their expected

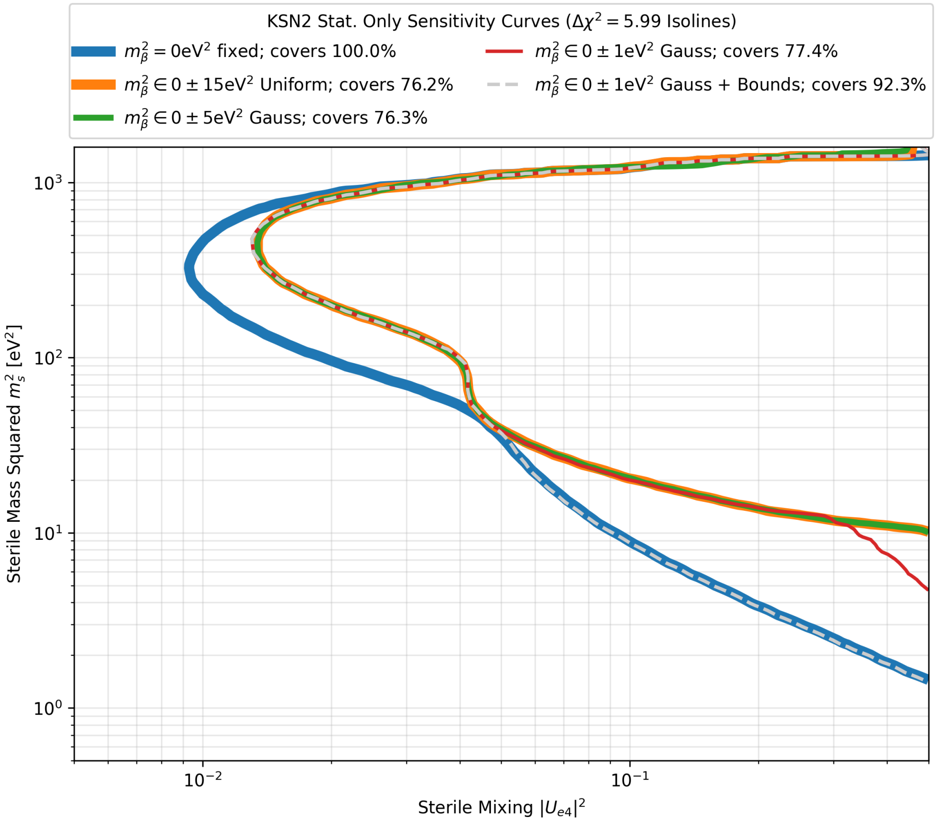


Figure 3.3: Sensitivity curves for all of the different active neutrino mass sampling distributions. The range indicated for the neutrino mass m_β^2 refers to the distribution of values that were used to train the net. In order to quantify the losses and gains in sensitivity, the coverage of each model's sensitivity curve is given. Coverage in this context refers to the amount of log-log area covered for $|U_{e4}|^2 \leq 0.5$ in comparison to the presently used method of keeping the active mass m_β^2 fixed to $0\,\mathrm{eV}^2$. The NNs with free neutrino mass agree very well for high sterile masses, although they lose some sensitivity compared to the model with a fixed neutrino mass. For low masses, the NN that is trained on a 0 ± 1 Gaussian distribution and then bound to positive active neutrino masses in the χ^2 minimization coincides with the model that has the active neutrino mass fixed to zero.

value [45]. For this thesis, the active mass m_β^2 is also a nuisance parameter, with an expected value of $0\,\mathrm{eV}^2$. The sterile parameters m_4^2 and $|U_{e4}|^2$ are also set to zero.

Results from Wilks [46] and Wald [47] show that the deviation from the null hypothesis $\Delta\chi^2 = \chi^2 - \chi^2_{\mathrm{NH}}$ approximates the log-likelihood ratio $-2\log\Lambda$ asymptotically. The most relevant takeaway from this is that the parameter-space with $\Delta\chi^2 \geq 5.99$ corresponds to an exclusion with 95 % confidence level (C.L.).

As such, the sensitivity curves are $\Delta\chi^2\left(m_4^2, |U_{e4}|^2\right)$ isolines, under the assumption of Wilks' theorem. Since we are looking at an Asimov set, $\chi^2_{\mathrm{NH}} = 0$ by construction. The other summand $\chi^2\left(m_4^{2\prime}, |U_{e4}|^{2\prime}\right)$ is then the deviation between the Asimov spectrum and the best fit of that data when fixing the sterile parameters to $\left(m_4^2, |U_{e4}|^2\right) = \left(m_4^{2\prime}, |U_{e4}|^{2\prime}\right)$

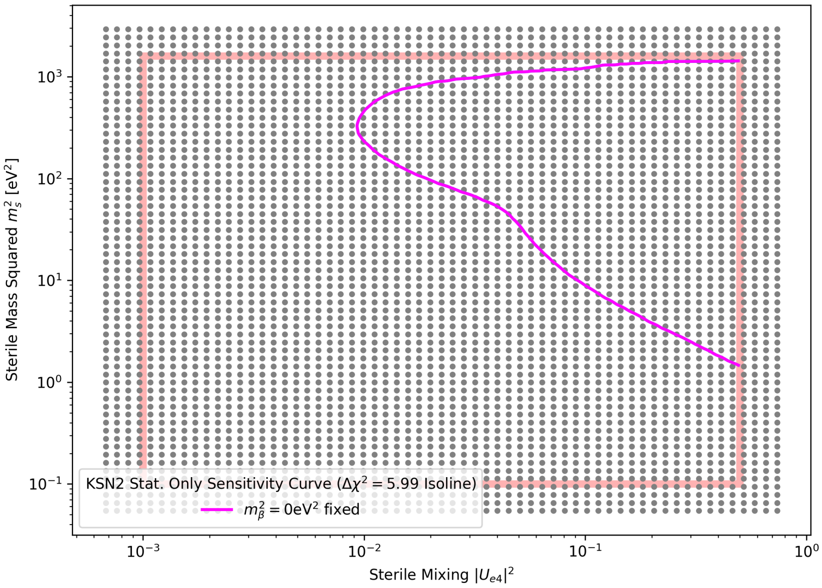


Figure 3.4: Grid scan evaluation points used throughout this work. The region of interest is marked with the red box. Overall this is a 61x61 evenly spaced grid in log-log space. 53x53 of the points lie within the region of interest.

Each NN gets used to probe an evenly spaced grid scan in the log-log region of interest in the sterile parameter-space. Fig. 3.4 shows the grid points used in these grid scans, alongside the actual region of interest as a red box, and the sensitivity curve for the $m_\beta^2 = 0\,\mathrm{eV}^2$ fixed case for reference. The grid intentionally goes slightly beyond the actual region of interest, since being able to look slightly beyond when investigation patterns and behaviors in the parameters can be quite helpful. Furthermore, it improves interpolation of values along the edges.

At each point in this grid the sterile parameters are fixed accordingly. Next a χ^2 minimization is done to find the other parameters that produce a model spectrum that best fits the Asimov spectrum. In this χ^2 minimization the model as described in Sec. 3.1.2 is used. After doing this for all of the networks the $\Delta\chi^2$ maps – and thus the sensitivity curves – can be compared.

There is very good agreement between the sensitivity curves produced by the different NNs for most of the parameter space, as seen in Fig. 3.3. As expected, there is some overall loss in sensitivity, as the free neutrino mass can impact the sterile signal. This loss in sensitivity is quantified as a percentage of the log-log parameter-space covered relative to the model which keeps the active mass fixed to zero.

The deviation of the sensitivity curve for the NN trained on active masses $m_\beta^2 \in 0 \pm 1\text{eV}^2$ (Gaussian) at sterile masses $m_4^2 \sim 10\,\text{eV}^2$ and sterile mixing $|U_{e4}|^2 \geq 0.3$ corresponds to the region in which the best fit of the active neutrino mass values strays from the minimum of the sample-set.

In order to better illustrate this behavior, the best fit values for the active mass m_β^2 for values between $-10\,\text{eV}^2$ to $10\,\text{eV}^2$ are shown as a contour map in Fig. 3.5. The dashed lines represent the sensitivity curves of the model in question, as well as the fixed case for reference. The minimum of the sample-set used was at $-5.3\,\text{eV}^2$, the maximum at $5.89\,\text{eV}^2$.

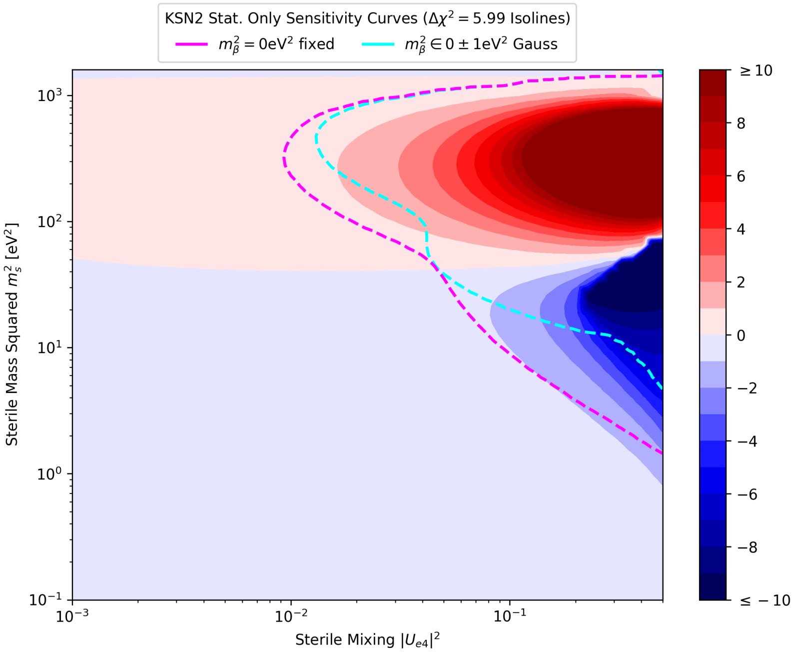


Figure 3.5: Contour map showing the best fit values for the active mass m_β^2 across the entire grid scan for the 0 ± 1 eV^2 (Gauss) net. The kink-shaped deviation of the sensitivity curve for the 0 ± 1 eV^2 (Gauss) case at $m_4 \sim 10\,\text{eV}^2$ coincides with the points where the best fit tends towards values that exceed the sampling range of this NN's training data. This is because the NN becomes inaccurate at points where a parameters best fit result lies outside of its training range, thus increasing the $\Delta\chi^2$ more than in the other NNs. An analogous deviation for positive active masses $m_\beta^2 \geq\sim 6\,\text{eV}^2$ is hidden because here the sterile signal at high masses m_4^2 causes an increase in $\Delta\chi^2$ for all of the NNs.

This difference is because the sample generation samples randomly within this distribution, so the actual minima and maxima are slightly arbitrary. Since it is only negative mass values that can dampen the sterile neutrino signal, this deviation in the χ^2 map from this NN to the others does not manifest in the positive active mass region.

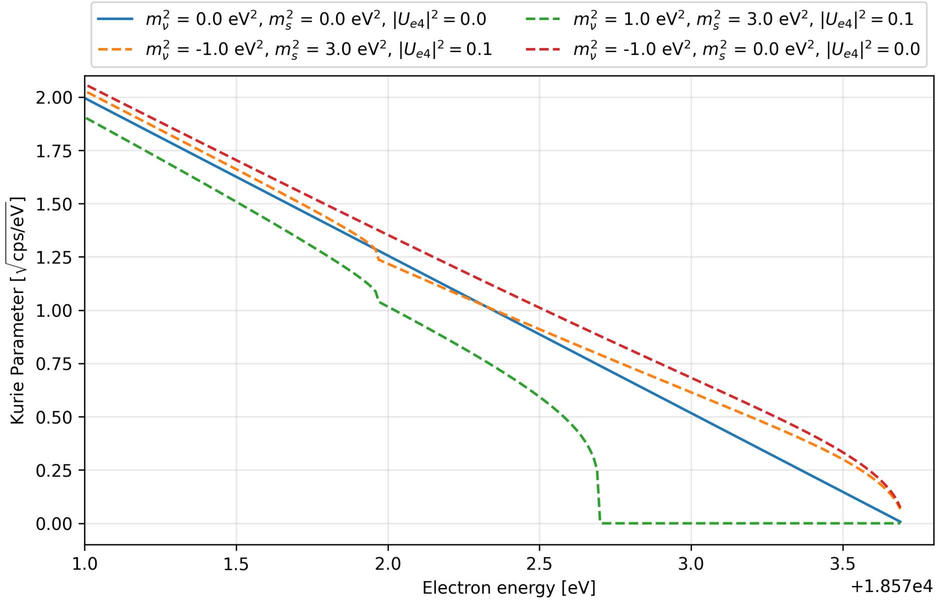


Figure 3.6: Kurie Plot [48] to show the impact of various combinations of the sterile mass m_s, the active mass m_β and the mixing ratio $|U_{e4}|^2$. Note how the inclusion of a negative active mass compensates the impact of a sterile signal and makes it difficult to distinguish from the null hypothesis, especially if the endpoint is varied as well.

Apart from this deviation at active neutrino masses that exceed the sample-set, the m_β^2-free NNs agree on the distribution of the best fit values.

3.4 Bounds

Due to the degeneracy of the active and sterile neutrino branch it could be useful to restrict the active neutrino mass values to regain some sensitivity. This degeneracy is shown in Fig. 3.6. Note how the null hypothesis with no neutrino at all in blue looks very similar to the orange dotted signal with a negative active mass m_β^2 and positive sterile mass m_s^2. Sensitivity studies that compare the deviation from the model χ^2 between a given combination of sterile parameters and the null hypothesis will therefore unreasonably conclude that there is less sensitivity in regions where the negative mass can compensate in this manner. Therefore, disallowing negative values in the fit is a reasonable approach to regain some sensitivity.

Setting bounds for the active mass m_β^2 to be greater or equal to zero results in considerable gains in lost sensitivity. As shown in Fig. 3.3 the sensitive area of the parameter-space for a free neutrino mass relative to the case where m_β^2 is fixed is at roughly 76 %. By setting a lower bound of $m_\beta^2 \geq 0\,\mathrm{eV}^2$ we retain 92 % of the sensitivity relative to the fixed case. Note that these comparisons are arbitrary, but they quantify the fact that restricting the active neutrino mass squared m_β^2 breaks some of the degeneracies between the active and the sterile neutrino masses. This gain is especially large in the lower sterile mass region $m_4^2 < 40\,\mathrm{eV}^2$.

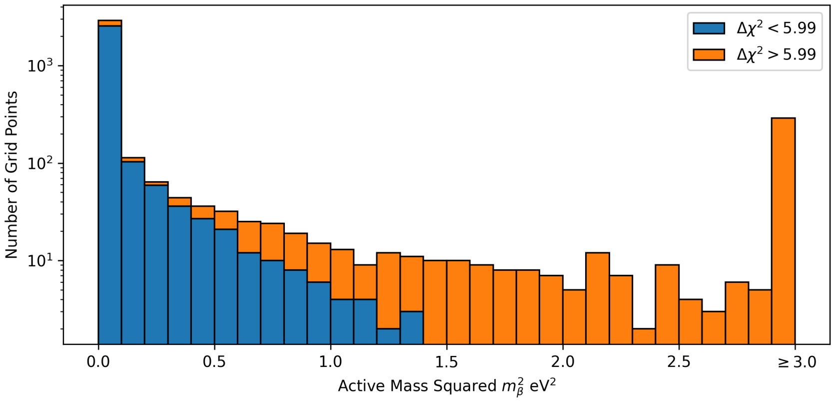


Figure 3.7: Stacked histogram showing the distribution of m_β^2 best fits for the NN trained on samples over $m_\beta^2 \in 0 \pm 1$ eV^2 Gauss. This model was evaluated using bounds such that $m_\beta^2 \geq 0$ eV^2.

Fig. 3.7 shows the distribution of the fitted values for m_β^2 for the NN trained on the narrow Gaussian set of active masses, using the lower bound of $m_\beta^2 \geq 0\,\mathrm{eV}^2$. The best fit values for m_β^2 outside of the sensitivity curve never exceed $1.5\,\mathrm{eV}^2$. Even in the region within the sensitivity curve, the majority of values never exceed $2\,\mathrm{eV}^2$. As the model tries to compensate for a significant sterile signal, high active neutrino masses are fitted. When fitting to real data, the active mass could be restricted within a reasonable range given by other experiments.

However, this choice of restriction will always be arbitrary which makes settling on a specific restriction difficult.

For the purposes of the rest of this analysis, the different scenarios are used to highlight different aspects. The fixed case will be referenced throughout as a reference to compare against. The model trained on the narrow Gaussian distribution without bounds is useful because this comparison of sensitivities has shown that a narrow distribution is enough to reproduce the sensitivity curves produced by a much wider distribution, while being sufficient to represent the expected values from a real fit. The addition of bounds seems like an easy way to compromise between a high sensitivity while avoiding any potential inaccuracies in the sterile analysis due to underestimating the impact of the active neutrino mass. The model trained on the broad uniform sample-set will be used to compare what a truly freely fitted neutrino mass model will do in comparison to the fixed case.

3.5 Impact of Individual Fit Parameters

Adding the active mass as a new fit parameter also impacts the behavior of the other fit parameters. Here the four parameters that get fitted by the NNs as part of a statistics only grid scan get analyzed individually. If there is any difference in the correlation between the sterile parameters and the fitted input parameters, then an attempt at an explanation is given.

3.5.1 Mass

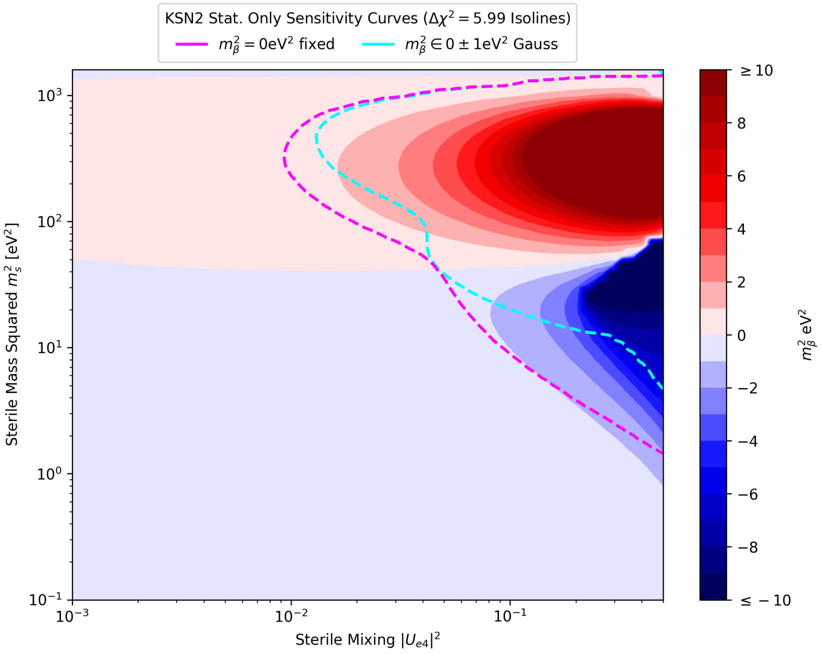
In the low mass region, the fit compensates for the sterile distortion of the signal by adding a negative active mass. This phenomenon is outlined in Fig. 3.6, and is especially visible in Fig. 3.9. This works up to sterile masses of approximately $50\,\mathrm{eV}^2$. For sterile masses above this point, the overall signal becomes so dampened that the best way to bring it in agreement with the null hypothesis is to increase the endpoint energy E_0, and eventually the normalization as well.

3.5.2 Endpoint

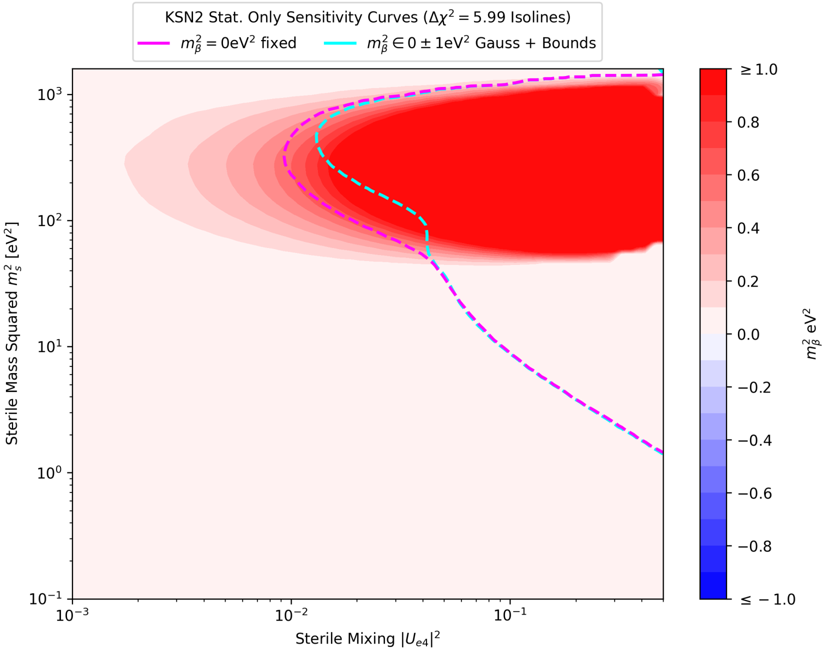
As indicated by the example spectra in Fig. 2.1, and by the Kurie plot in Fig. 3.6, the addition of a neutrino to the spectrum causes a dampening of the spectrum. Since the fit tries to match the Asimov spectrum without any neutrinos, it has to compensate for this dampening. An increase in the endpoint energy E_0 corresponds to an increase in the overall energy available to the electron. Thus an increase corresponds to a mitigation of this aforementioned dampening, which improves the fit.

The addition of a sterile neutrino with masses m_4 of up to $\sim 30\,\mathrm{eV}$ means a much more significant dampening of the spectrum needs to be compensated for in order to fit the Asimov spectrum. Fig. 3.10 shows the contour map of the fitted endpoint energies E_0. The correlation between the endpoint E_0 and the active neutrino mass m_β has already been demonstrated in previous publications [20] and is clearly present here too. Note how the endpoint energies E_0 reach a higher maximum in the grid scan with the free neutrino mass model. This is because adding a positive neutrino mass m_β^2 can lead to a better fit in the overall shape of the spectrum, at the cost of more dampening. This additional dampening is compensated with additional endpoint energy E_0.

Fitting higher and higher endpoint energies E_0 for increasing sterile masses squared m_4^2 only works up to a certain point. As the sterile mass m_4 approaches the limit of the measurement window – which is $40\,\mathrm{eV}$ below the predicted



(a)



(b)

Figure 3.8: Contour map of the fitted background count rate for the model in which the active neutrino mass is fixed to zero (a), and for the model in which the active neutrino mass is free, which is trained on a wide uniform distribution $m_\beta^2 \in 0 \pm 5$ eV^2 (b). Note the fact that the scales on the masses are different, in order to highlight the sub-eV structure at positive mass ranges.

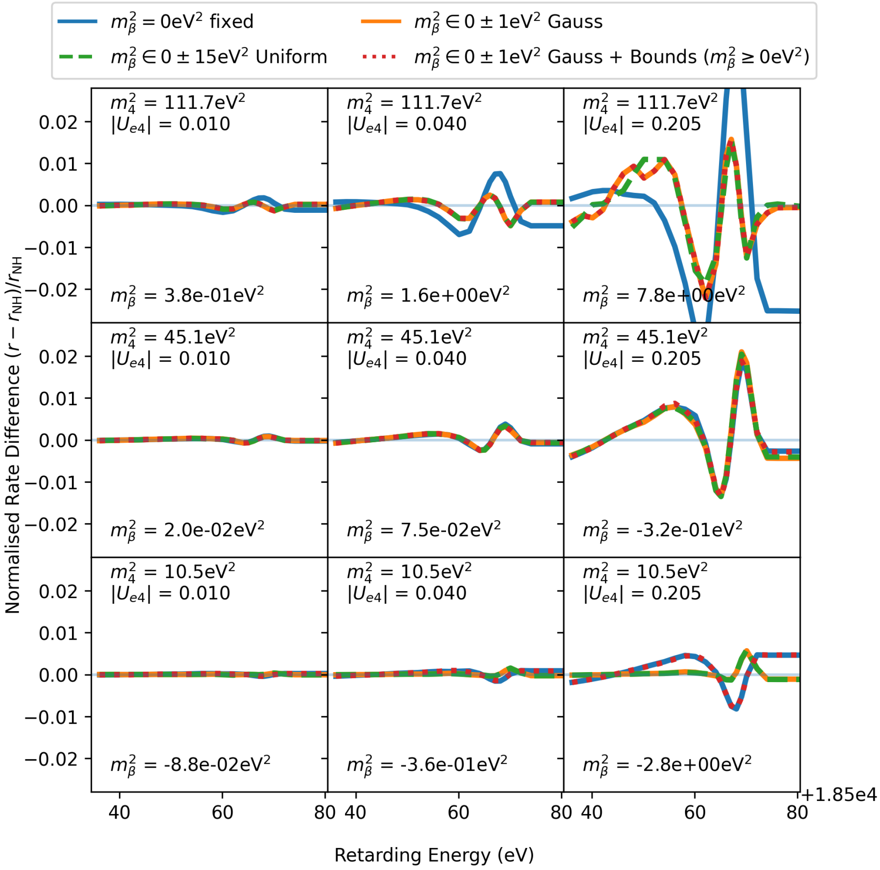
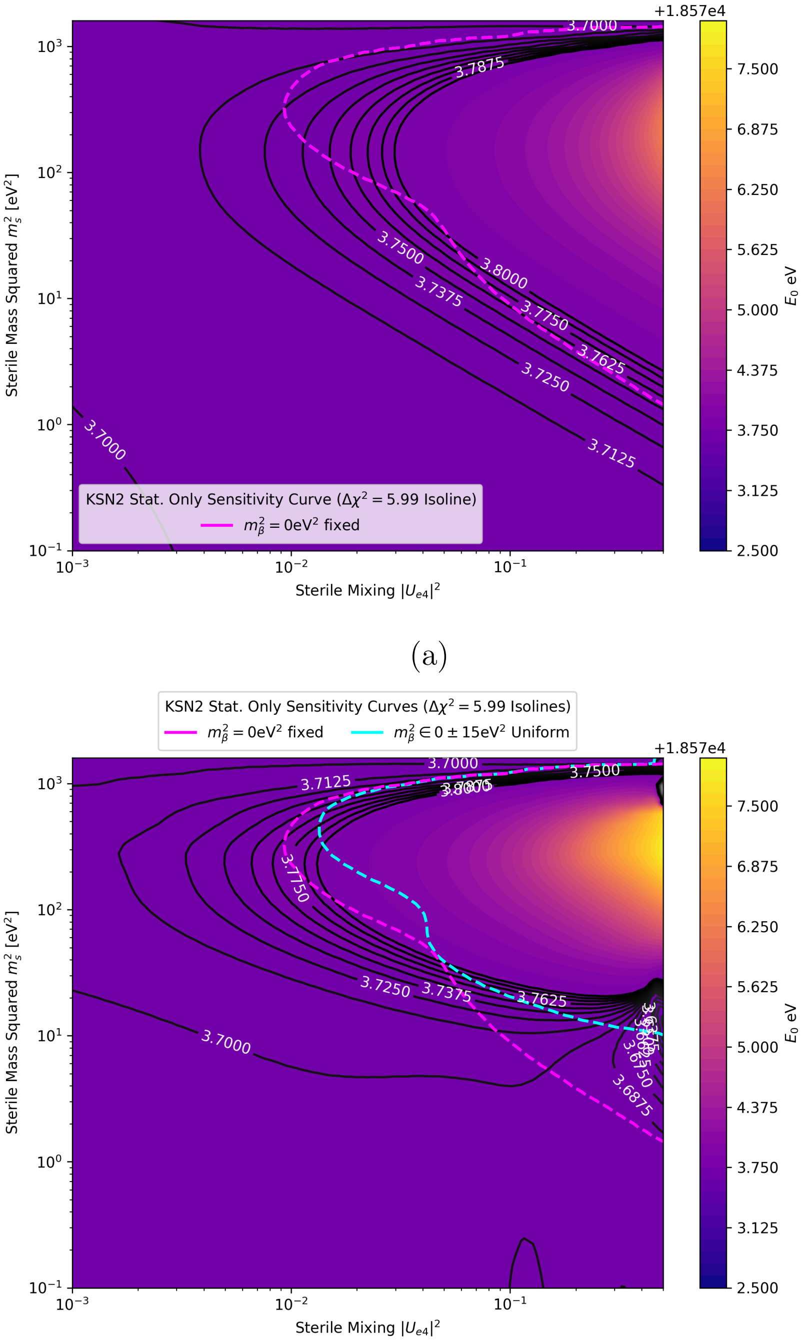


Figure 3.9: Individual best fit spectra relative to the null hypothesis, at 9 selected points in the grid scan. Since these are relative values, a perfect match to the null hypothesis would be a perfectly flat line. The listed values for the active mass m_β^2 refer to the best fit value as provided by the $m_\beta^2 \in 0 \pm 15\text{eV}^2$ (uniform) net. Note how the degeneracy between the active and sterile neutrino signals leads to an overall dampened signal – especially at high sterile mass m_4^2 and/or large mixing $|U_{e4}|^2$. This corresponds to a decreased $\Delta\chi^2$ and thus a reduced sensitivity at these points. Also note the difference in the spectra produced by the narrow and wide distributions of active neutrino masses m_β^2 as a result of the NN being forced to extrapolate for high sterile masses and mixings, and thus losing some accuracy.

endpoint energy E_0 – most of the signal gets dampened very significantly, to the point where increasing the normalization by a large amount yields better fits than changing the neutrino parameters.

3.5.3 Normalization

The fitted normalization has little variation throughout most of the parameter space. However at very high sterile masses $m_4^2 \sim 1000\,\text{eV}^2$ with high sterile



(a)

(b)

Figure 3.10: Contour map of the fitted endpoint energies E_0 for the model in which the active neutrino mass is fixed to zero (a), and for the model in which the active neutrino mass is free, which is trained on a wide uniform distribution $m_\beta^2 \in 0 \pm 5$ eV^2 (b). The fitted endpoint energies reach a higher maximum in the case of a free neutrino mass, and they become significantly modified in comparison to the fixed mass case. Especially the distortion at sterile mass ranges of $\sim 20\,\text{eV}^2$ seems interesting and could warrant further investigation.

mixings $|U_{e4}| \sim 0.5$, it becomes the most impactful fitting parameter, as the spectrum is dampened over most of the measurement range. Since this dampening for points between the endpoint energy E_0 and m_4 electronvolt below the endpoint $E_0 - m_4$ is homogenous, the multiplicative normalization can restore a spectrum that matches the null hypothesis.

The addition of the active neutrino mass as a free parameter does slightly modify the normalization isolines, but the overall profile remains the same.

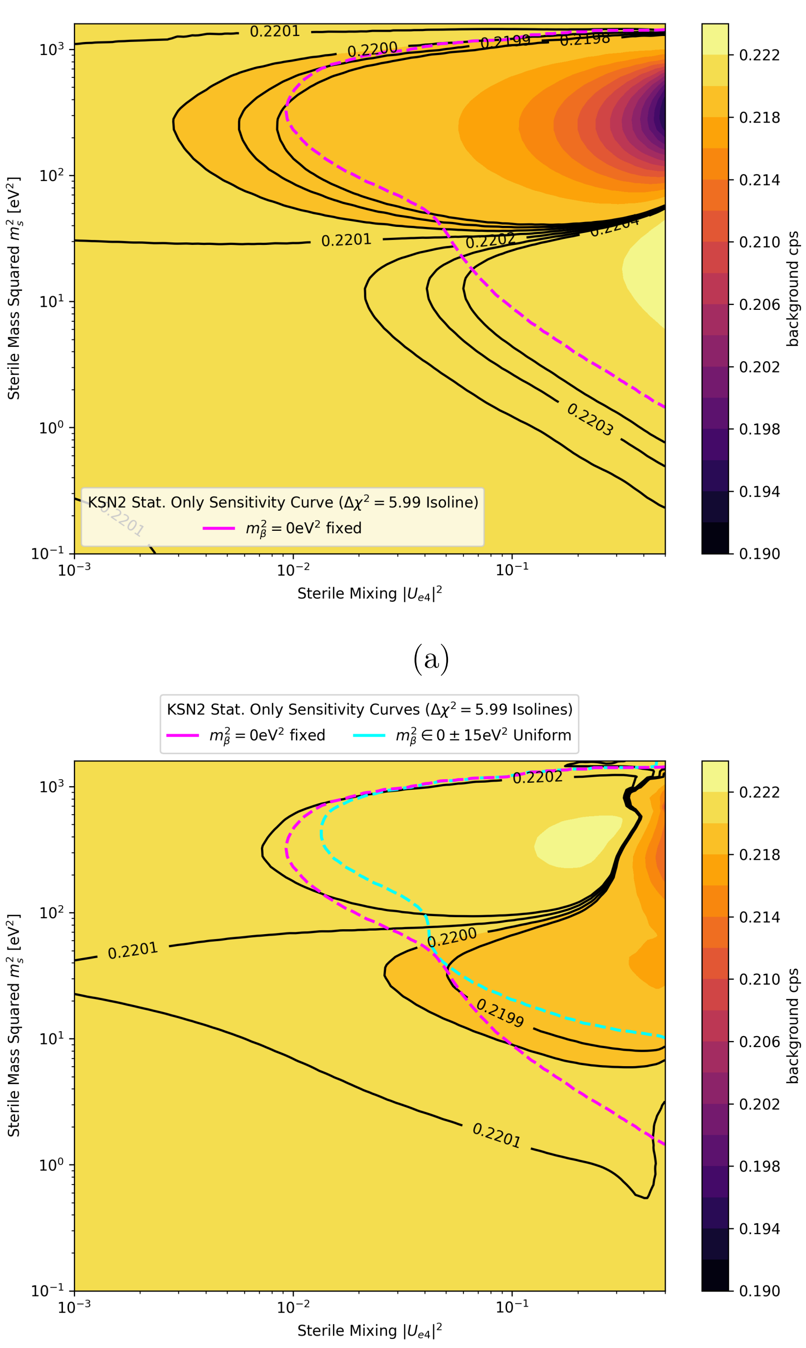
3.5.4 Background

The background is correlated to the endpoint in these fits, as the measurement includes points beyond the simulated endpoint energy E_0. In the cases where increasing the endpoint energy leads to better agreement at points slightly deeper in the spectrum, this would lead to an overshoot beyond the true endpoint. At points beyond the endpoint of the spectrum, only the background would contribute to the rate. Therefore, increasing the fitted endpoint energy E_0 corresponds to decreasing the background rate to compensate for the measured values above the true endpoint.

Interestingly, the addition of the active neutrino mass as a free parameter completely changes the shape of the contour. It appears that the sterile mass m_4 and the background are strongly correlated in the fixed active neutrino mass model, while the opposite is true for the free active neutrino mass models: the two parameters seem weakly anti-correlated.

3.6 Systematics

The inclusion of systematics has a measurable impact on the sensitivity of the model, as seen in Fig. 3.12. The impact is well behaved and relatively small. Further investigations could be done to see how the individual systematics have changed across the parameter-space due to the addition of the active neutrino mass as a free parameter.



(a)

(b)

Figure 3.11: Contour map of the fitted background count rate for the model in which the active neutrino mass is fixed to zero (a), and for the model in which the active neutrino mass is free, which is trained on a wide uniform distribution $m_\beta^2 \in 0 \pm 15$ eV^2 (b).

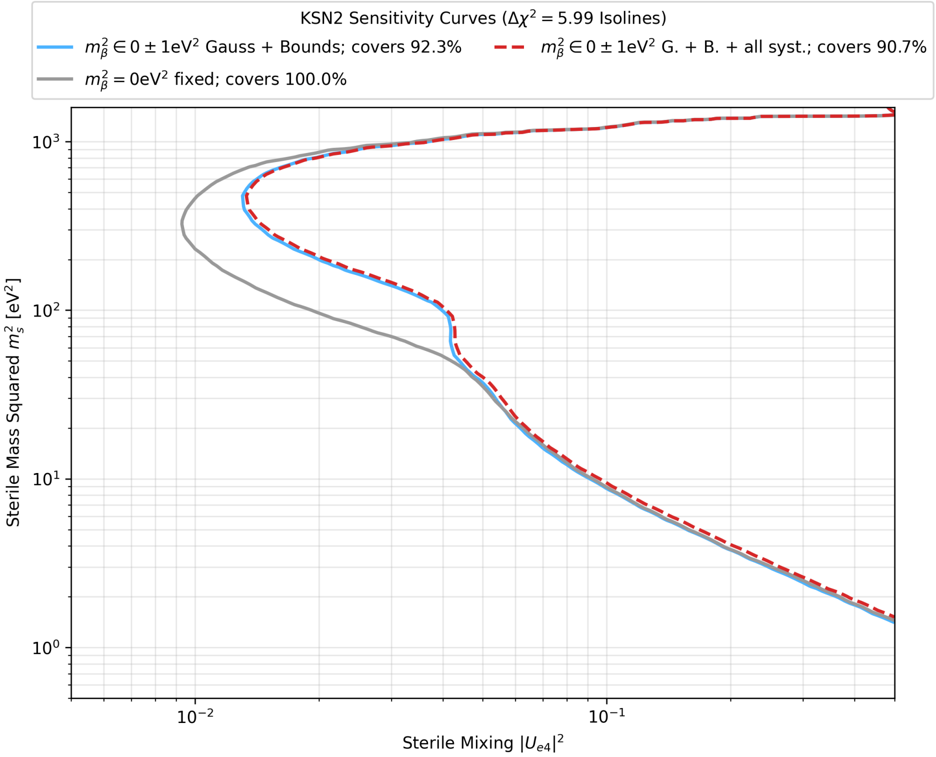


Figure 3.12: Sensitivity curves for the model with a fixed active neutrino mass $m_\beta^2 = 0\,\mathrm{eV}^2$ shown in gray, and with a free active neutrino mass where the model has been trained on a narrow Gaussian distribution $m_\beta^2 \in 0 \pm 1\ \mathrm{eV}^2$ shown in blue. In the case of the free active neutrino mass, the χ^2 minimization has been restricted to positive values $m_\beta^2 \geq 0\,\mathrm{eV}^2$. The two solid lines refer to sensitivity curves where only statistical uncertainties are considered. The dashed red line shows the sensitivity curve when all systematics are considered as well. The coverage percentage refers to the amount of log-log area covered in the parameter-space, relative to the model based on the fixed active neutrino mass. As can be seen, the systematics have a measurable but small impact on the sensitivity of the model with a free active neutrino mass.

Chapter 4

Conclusion

In addition to providing sub-eV limits on the effective neutrino mass m_β, the KATRIN experiment is capable of providing competitive constraints in the search for eV-scale sterile neutrinos. In the analysis of the sterile parameter-space, the effective neutrino mass m_β needs to be considered, as it can lead to degeneracies due to dampening and mimicking of a sterile mass m_4^2 signal.

The analysis of β-decay spectra measured at the KATRIN experiment has recently begun to use neural networks (NNs) as a method of fast and precise spectrum simulation. The aim of this work is to implement methods for freely fitting the active neutrino mass within the Netrium analysis toolkit. The dataset used in this thesis is based on the second run of KATRIN measurements in 2019 (referred to as KSN2).

These NNs require large amounts of sample β-electron energy spectra as an upfront computational expense in order to benefit from their efficiency in subsequent analyses. The precision of the NN is improved by having a dense sampling around the best fit value. This requires more samples which are computationally expensive. Furthermore a broader distribution allows the same NN to cover a broader area of parameter-space accurately. In order to determine the most effective way to sample the active neutrino mass to achieve a good trade-off in precision, cost efficiency and flexibility, multiple NNs were trained with different sample distributions. The four sample distributions that were used to train the NNs differed in the distribution of the active neutrino mass squared m_β^2. The distributions used were: $m_\beta^2 = 0$ eV^2 fixed, $m_\beta^2 \in 0 \pm 15$ eV^2 uniform, $m_\beta^2 \in 0 \pm 5$ eV^2 Gaussian, and $m_\beta^2 \in 0 \pm 1$ eV^2 Gaussian. The NNs were then analyzed regarding their sensitivity and their impact on the correlations between the other fitting parameters.

Leaving the active neutrino mass free confirms that the exclusion contours are mainly affected in the sterile mass m_4^2 region, below $40\,\mathrm{eV}^2$, resulting in a significant loss of sensitivity. This loss in sensitivity below sterile masses squared $40\,\mathrm{eV}^2$ can be reduced by introducing restrictions to the active neutrino mass. As an example, this work examined the impact of restricting the active neutrino mass squared to positive values $m_\beta^2 \geq 0\,\mathrm{eV}^2$. Implementing this restriction mitigates most of the sensitivity loss, however it is an arbitrary choice. Furthermore allowing negative active neutrino masses squared accounts for statistical underfluctuations of the spectrum.

When choosing the distribution to train the net with, a narrow distribution necessitates restricting the χ^2 minimization from above with some external limit, as the NNs become inaccurate for extrapolations beyond the parameter-space of their sample-spectra.

Based on first testing, the addition of the active neutrino mass as a free parameter has no adverse impacts on the systematic uncertainties. The changes to the systematic parameters that impact the spectrum were studied in more detail. Just as with the neutrino mass analysis, there is a significant correlation between fitted mass of the sterile neutrino and the endpoint energy E_0. This correlation becomes stronger when adding the active neutrino mass as a free parameter. In the case of keeping the active mass fixed, the sterile mass and the background count rate are anti-correlated, especially for high sterile mixings $|U_{e4}|^2 > 0.05$. The addition of the active neutrino mass as a free parameter seems to change that correlation to the point where the two parameters are correlated rather than anti-correlated for high sterile mass-regions. The correlation between the sterile parameters and the signal normalization is also slightly modified by the addition of the active neutrino mass, but remains fundamentally the same.

This work has outlined how to implement a free active neutrino mass in Netrium for the analysis of light sterile neutrinos in KATRIN. The best sample distribution appears to be a superposition of two distributions centered around zero. One being broad uniform distribution to avoid extrapolation at sterile masses m_4^2 below $10\,\mathrm{eV}^2$ and over $100\,\mathrm{eV}^2$, and the other being a narrow Gaussian in order to retain high precision around the expected true value of m_β^2. The Netrium analysis toolkit is now ready to be used for further complementary analyses with a free neutrino mass.

Appendix A

Additional Figures and Tables

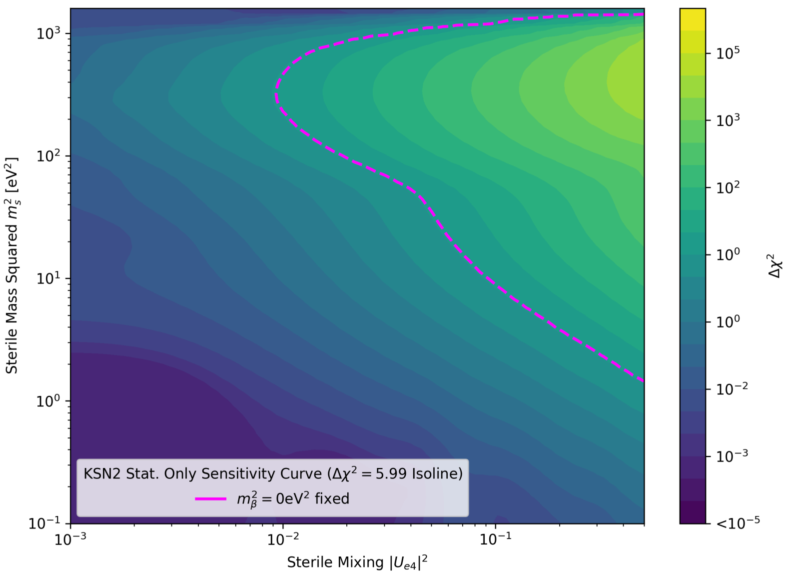


Figure A.2: $\Delta\chi^2$ map for the case with the active neutrino mass fixed to zero. Colorbar values were chosen to maintain consistency between plots.

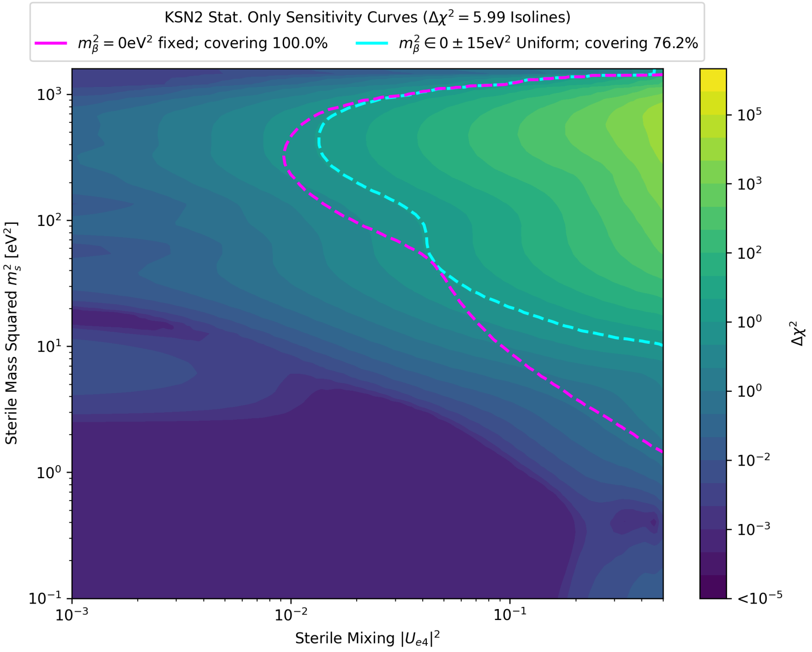


Figure A.3: $\Delta\chi^2$ map for the case with the active neutrino mass free, using a neural net trained on a broad uniform sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

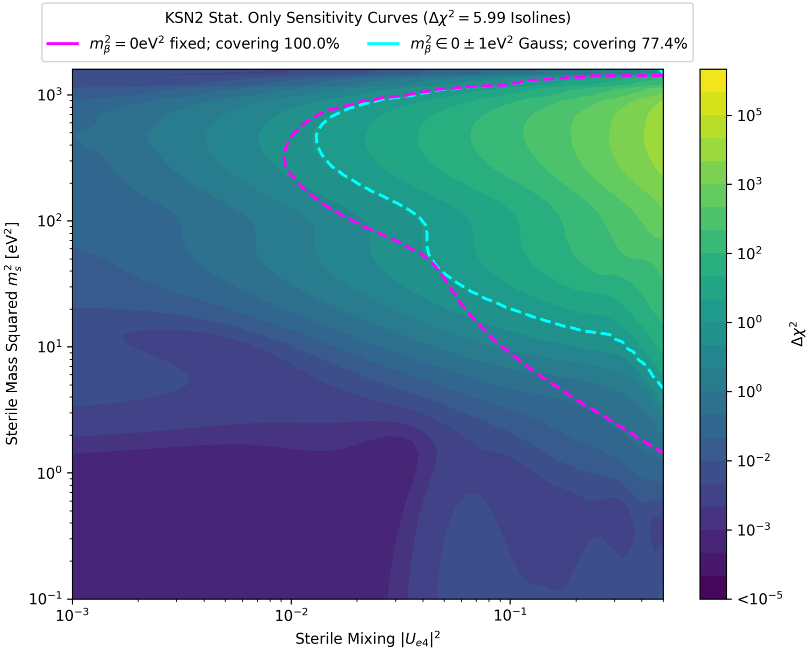


Figure A.4: $\Delta\chi^2$ map for the case with the active neutrino mass free, using a neural net trained on a narrow Gaussian sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

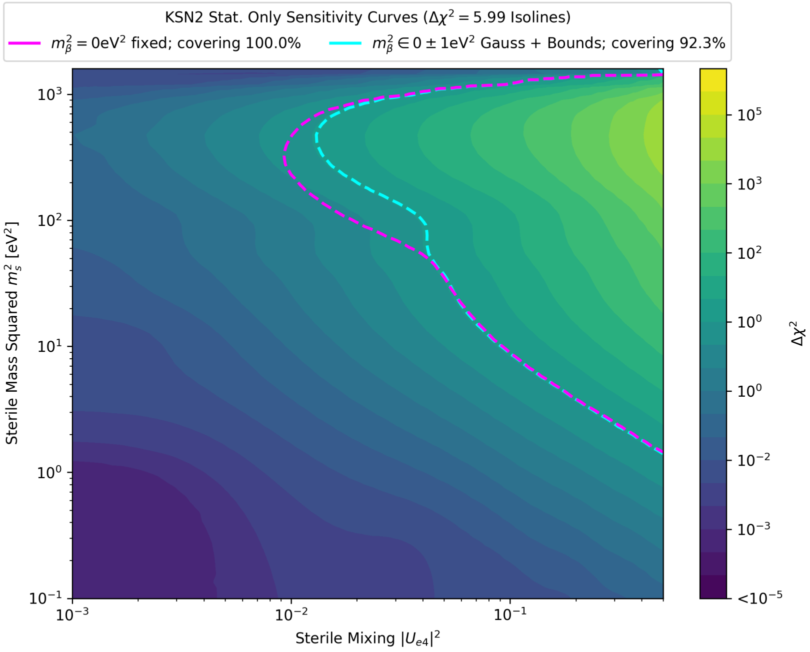


Figure A.5: $\Delta\chi^2$ map for the case with the active neutrino mass free, using a neural net trained on a narrow Gaussian sample-range of active mass squared values, bounded to positive active mass squared values. Colorbar values were chosen to maintain consistency between plots.

Table A.1: Active mass squared m_β^2 distributions for the sample sets used in training the nets

Name	Samples	Distribution	min (eV^2)	max (eV^2)	μ (eV^2)	σ (eV^2)
KSN2	2 400 000	Linear Uniform	-3.50×10^{-4}	3.50×10^{-4}	-2.27×10^{-9}	2.02×10^{-4}
KSN2 Gauss 0±5	4 600 000	Linear Gaussian	-2.79×10^{1}	2.56×10^{1}	-4.02×10^{-5}	5.00
KSN2 Uniform 0±15	4 800 000	Linear Uniform	-1.50×10^{1}	1.50×10^{1}	9.28×10^{-5}	8.66
KSN2 Gauss 0±1	4 800 000	Linear Gaussian	-5.30	5.89	3.66×10^{-5}	1.00

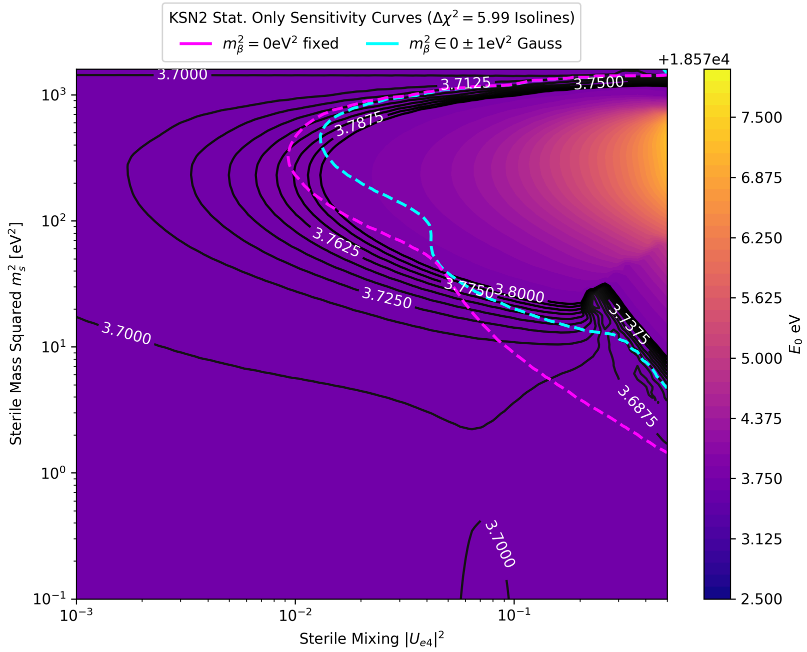


Figure A.6: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a narrow Gaussian sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

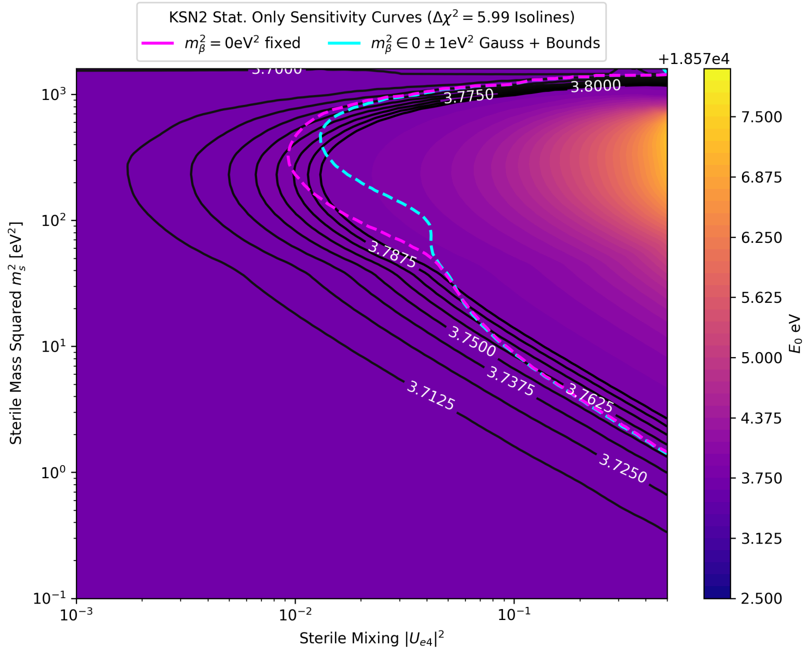


Figure A.7: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a narrow Gaussian sample-range of active mass squared values, bounded to positive active mass squared values. Colorbar values were chosen to maintain consistency between plots.

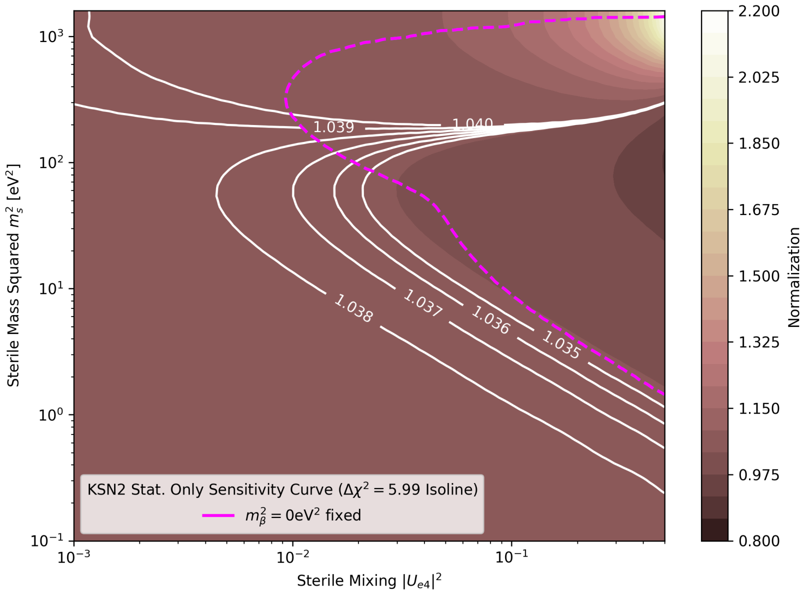


Figure A.8: Contour plot for a fit parameter, with a selection of isolines included, using the active mass fixed case. Colorbar values were chosen to maintain consistency between plots.

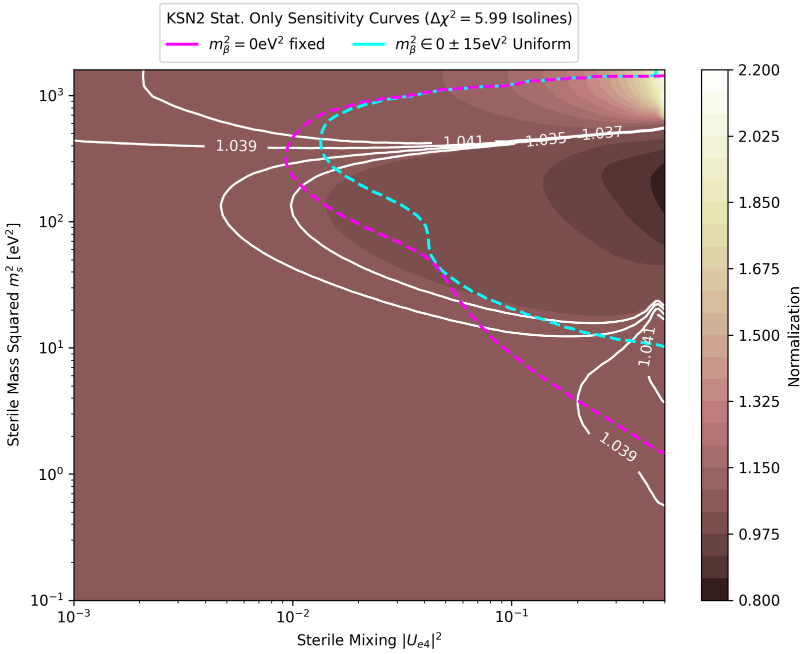


Figure A.9: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a broad uniform sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

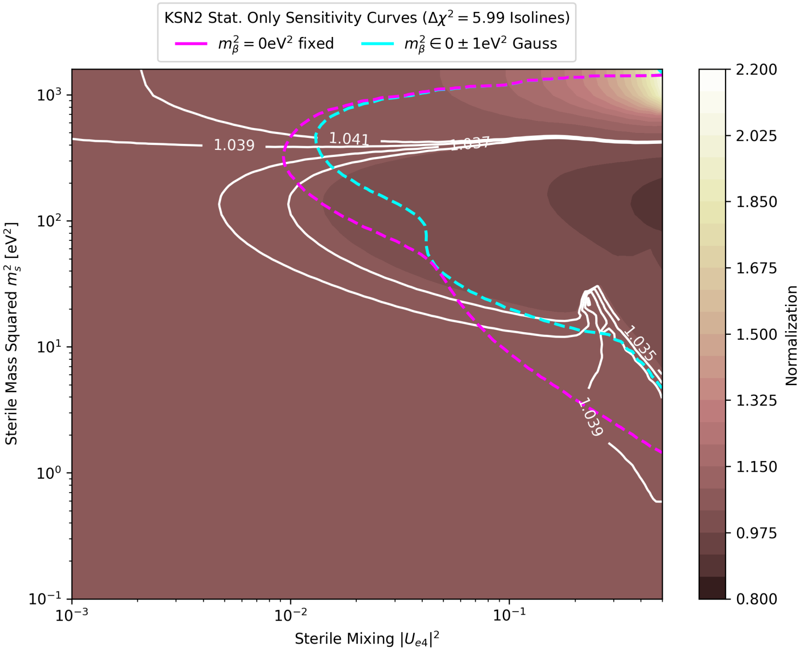


Figure A.10: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a narrow Gaussian sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

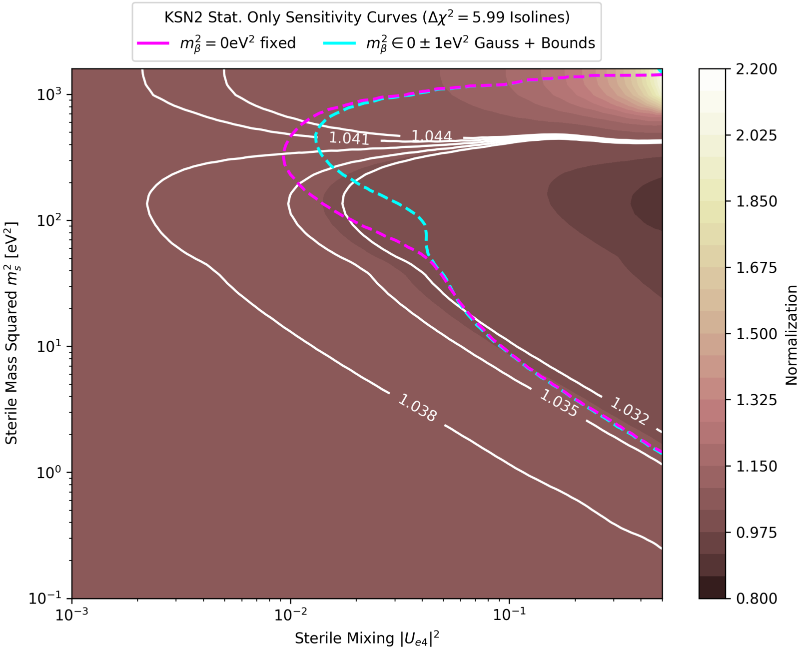


Figure A.11: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a narrow Gaussian sample-range of active mass squared values, bounded to positive active mass squared values. Colorbar values were chosen to maintain consistency between plots.

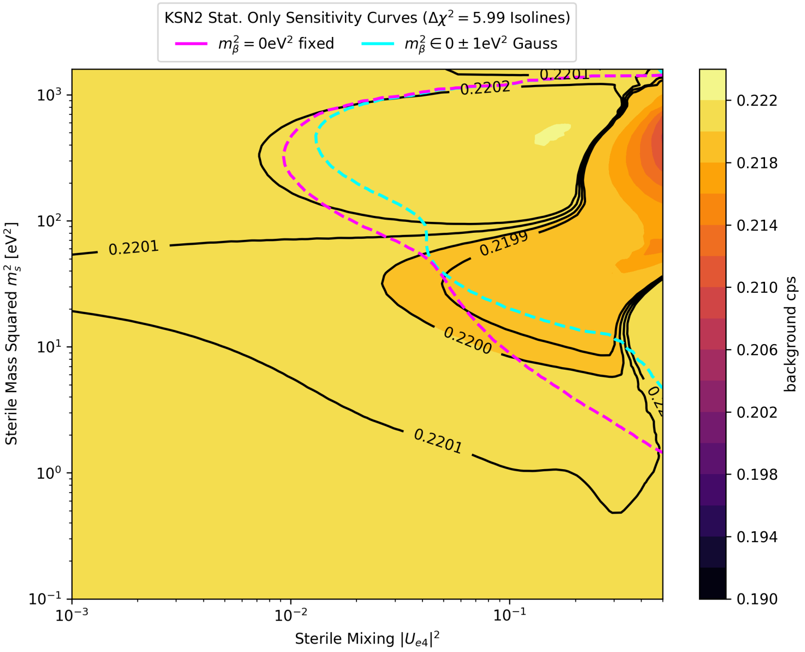


Figure A.12: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a narrow Gaussian sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

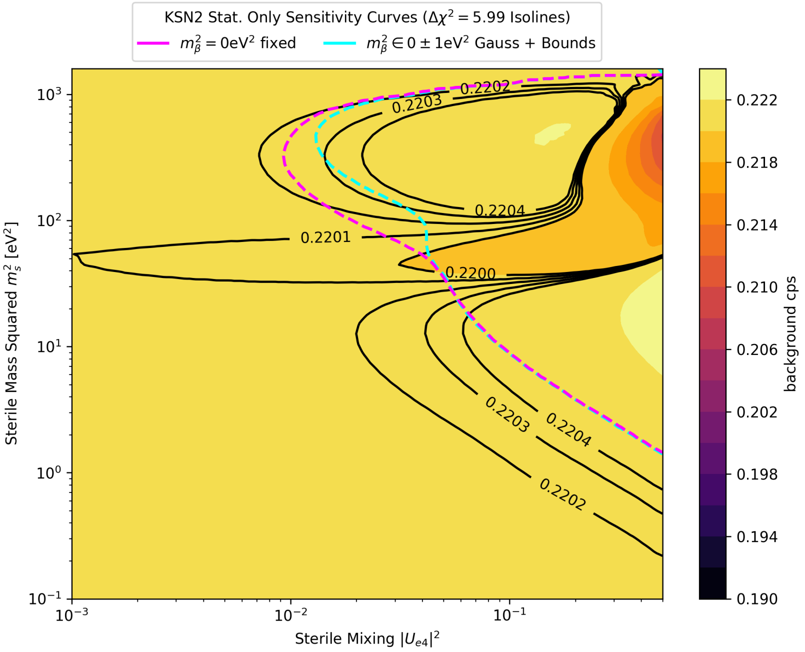


Figure A.13: Contour plot for a fit parameter, with a selection of isolines included, using a neural net trained on a narrow Gaussian sample-range of active mass squared values. Colorbar values were chosen to maintain consistency between plots.

Bibliography

1. Pauli, W. *Pauli letter collection: letter to Lise Meitner* `https://cds.cern.ch/record/83282`.
2. Franklin, B. *Letter to Cadwallader Colden on Public Affairs, Military Desertion, and Electrical Research* June 1747. `https://franklinpapers.org/`.
3. Mayer, J. R. XLVIII. Remarks on the forces of inorganic nature. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **24,** 371–377. eprint: `https://doi.org/10.1080/14786446208643372`. `https://doi.org/10.1080/14786446208643372` (1862).
4. Workman, R. L. & Others. Review of Particle Physics. *PTEP* **2022,** 083C01 (2022).
5. Goldhaber, M., Grodzins, L. & Sunyar, A. W. Helicity of Neutrinos. *Phys. Rev.* **109,** 1015–1017. `https://link.aps.org/doi/10.1103/PhysRev.109.1015` (3 Feb. 1958).
6. Wu, C. S., Ambler, E., Hayward, R. W., Hoppes, D. D. & Hudson, R. P. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.* **105,** 1413–1415. `https://link.aps.org/doi/10.1103/PhysRev.105.1413` (4 Feb. 1957).
7. Davis, R., Harmer, D. S. & Hoffman, K. C. Search for Neutrinos from the Sun. *Phys. Rev. Lett.* **20,** 1205–1209. `https://link.aps.org/doi/10.1103/PhysRevLett.20.1205` (21 May 1968).
8. Ahmad, Q. R. *et al.* Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.* **87,** 071301. `https://link.aps.org/doi/10.1103/PhysRevLett.87.071301` (7 July 2001).
9. Fukuda, Y. *et al.* Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.* **81,** 1562–1567. `https://link.aps.org/doi/10.1103/PhysRevLett.81.1562` (8 Aug. 1998).
10. Abe, Y. *et al.* Indication of Reactor $\overline{\nu}_e$ Disappearance in the Double Chooz Experiment. *Phys. Rev. Lett.* **108,** 131801. `https://link.aps.org/doi/10.1103/PhysRevLett.108.131801` (13 Mar. 2012).
11. Ahn, J. K. *et al.* Observation of Reactor Electron Antineutrinos Disappearance in the RENO Experiment. *Phys. Rev. Lett.* **108,** 191802. `https://link.aps.org/doi/10.1103/PhysRevLett.108.191802` (19 2012).
12. Pontecorvo, B. INVERSE *beta* PROCESSES AND NONCONSERVATION OF LEPTON CHARGE. *Zhur. Eksptl'. i Teoret. Fiz.* **34** (1958).

13. Maki, Z., Nakagawa, M. & Sakata, S. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics* **28,** 870–880. ISSN: 0033-068X. eprint: `https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf`. `https://doi.org/10.1143/PTP.28.870` (Nov. 1962).
14. Kayser, B. *NEUTRINO PHYSICS AS EXPLORED BY FLAVOR CHANGE* May 2002. `https://pdg.lbl.gov/2002/neutrino_mixing_s805.pdf`.
15. Mohapatra, R. N. & Goran, S. Neutrino Mass and Spontaneous Parity Nonconservation. *Phys. Rev. Lett.* **44,** 912–915. `https://link.aps.org/doi/10.1103/PhysRevLett.44.912` (14 1980).
16. Dvorkin, C. *et al. Neutrino Mass from Cosmology: Probing Physics Beyond the Standard Model* 2019. arXiv: `1903.03689 [astro-ph.CO]`.
17. Planck Collaboration. Planck results. *Astronomy and Astrophysics* **641,** A6. `https://doi.org/10.1051%2F0004-6361%2F201833910` (2020).
18. Patrignani, C. Review of Particle Physics. *Chinese Physics C* **40,** 100001. `https://dx.doi.org/10.1088/1674-1137/40/10/100001` (2016).
19. Agostini, M. *et al.* Final Results of GERDA on the Search for Neutrinoless Double-β Decay. *Phys. Rev. Lett.* **125,** 252502. `https://link.aps.org/doi/10.1103/PhysRevLett.125.252502` (25 2020).
20. Aker, M. *et al.* Direct neutrino-mass measurement with sub-electronvolt sensitivity. *Nature Physics* **18,** 160–166 (Feb. 2022).
21. Vogel, P., Wen, L. & Zhang, C. Neutrino oscillation studies with reactors. *Nature Communications* **6,** 6935 (2015).
22. Hayes, A. C. & Vogel, P. Reactor Neutrino Spectra. *Annual Review of Nuclear and Particle Science* **66,** 219–244. eprint: `https://doi.org/10.1146/annurev-nucl-102115-044826`. `https://doi.org/10.1146/annurev-nucl-102115-044826` (2016).
23. Zyla, P. *et al.* Review of Particle Physics. *Prog. Theor. Exp. Phys.* **2020.** Particle Data Group, 083C01 (2020).
24. Serebrov, A., Ivochkin, V., Samoilov, R. *et al.* First Observation of the Oscillation Effect in the Neutrino-4 Experiment on the Search for the Sterile Neutrino. *Jetp Letters* **109,** 213–221 (2019).
25. Barinov, V. V. *et al.* Search for electron-neutrino transitions to sterile states in the BEST experiment. *Physical Review C* **105.** `https://doi.org/10.1103%2Fphysrevc.105.065502` (2022).
26. Akindele, O. A. *et al. High Energy Physics Opportunities Using Reactor Antineutrinos* 2022. arXiv: `2203.07214 [hep-ex]`.
27. Aker, M. *et al.* Improved eV-scale sterile-neutrino constraints from the second KATRIN measurement campaign. *Phys. Rev. D* **105,** 072004. `https://link.aps.org/doi/10.1103/PhysRevD.105.072004` (7 Apr. 2022).
28. Rohatgi, A. *Webplotdigitizer: Version 4.6* 2022. `https://automeris.io/WebPlotDigitizer`.

29. Danilov, M. *New results from the DANSS experiment* 2022. arXiv: `2211.01208 [hep-ex]`.
30. Fermi, E. Versuch einer Theorie der β-Strahlen. I. *Zeitschrift für Physik* **88,** 161–177. ISSN: 0044-3328. `https://doi.org/10.1007/BF01351864` (Mar. 1934).
31. Aker, M *et al.* KATRIN: status and prospects for the neutrino mass and beyond. *Journal of Physics G: Nuclear and Particle Physics* **49,** 100501. `https://dx.doi.org/10.1088/1361-6471/ac834e` (Sept. 2022).
32. KATRIN collaboration, T. *et al.* The design, construction, and commissioning of the KATRIN experiment. *Journal of Instrumentation* **16,** T08015. `https://dx.doi.org/10.1088/1748-0221/16/08/T08015` (2021).
33. Lucas, L. L. & Unterweger, M. P. Comprehensive Review and Critical Evaluation of the Half-Life of Tritium. *Journal of Research of the National Institute of Standards and Technology* **105,** 541–549 (2000).
34. Heizmann, F., Seitz-Moskaliuk, H. & the KATRIN collaboration. *The Windowless Gaseous Tritium Source (WGTS) of the KATRIN experiment* in *XXVII International Conference on Neutrino Physics and Astrophysics (Neutrino2016)* **888** (IOP Publishing, 2017), 012071.
35. Bain, A. *Mind and Body: The Theories of Their Relation* (D. Appleton and Company, New York, 1873).
36. James, W. *The Principles of Psychology* (H. Holt and Company, New York, 1890).
37. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* **65,** 386–408 (1958).
38. Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86,** 2278–2324 (1998).
39. Denby, B. Neural networks and cellular automata in experimental high energy physics. *Computer Physics Communications* **49,** 429–448. ISSN: 0010-4655. `https://www.sciencedirect.com/science/article/pii/0010465588900045` (1988).
40. Radovic, A., Williams, M., Rousseau, D. *et al.* Machine learning at the energy and intensity frontiers of particle physics. *Nature* **560,** 41–48 (2018).
41. Duarte, J. & Vlimant, J.-R. in *Artificial Intelligence for High Energy Physics* 387–436 (WORLD SCIENTIFIC, Feb. 2022). `https://doi.org/10.1142%2F9789811234033_0012`.
42. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
43. Karl, C. *First Sub-Electronvolt Direct Neutrino Mass Measurement with the KATRIN Experiment* PhD thesis (2022).
44. Karl, C., Eller, P. & Mertens, S. Fast and precise model calculation for KATRIN using a neural network. *The European Physical Journal C* **82.**

https://doi.org/10.1140%2Fepjc%2Fs10052-022-10384-z (May 2022).

45. Cowan, G., Cranmer, K., Gross, E. & Vitells, O. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C* **71.** https://doi.org/10.1140%2Fepjc%2Fs10052-011-1554-0 (Feb. 2011).
46. Wilks, S. S. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics* **9,** 60–62. ISSN: 00034851. http://www.jstor.org/stable/2957648 (2023) (1938).
47. Wald, A. Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. *Transactions of the American Mathematical Society* **54,** 426–482 (1943).
48. Kurie, F. N. D., Richardson, J. R. & Paxton, H. C. The Radiations Emitted from Artificially Produced Radioactive Substances. I. The Upper Limits and Shapes of the β-Ray Spectra from Several Elements. *Phys. Rev.* **49,** 368–381. https://link.aps.org/doi/10.1103/PhysRev.49.368 (5 1936).

Acronyms

$0\nu\beta\beta$ neutrinoless double beta decay. 6

BEST Baksan Experiment on Sterile Transitions. 8

C.L. confidence level. 23
CMS calibration and monitoring section. 15

FPD focal plane detector. 16, 18

GERDA GERmanium Detector Array. 6

KATRIN **KA**rlsruhe **TRI**tium **N**eutrino. 1, 6, 8, 10, 11, 13, 14, 16, 19, 21, 35, 36

MAC-E magnetic adiabatic collimation with electrostatic filtering. 17, 20

NN Neural Network. 1, 19–28, 30, 35, 36, 50

PNMS Pontecorvo Maki Nakagawa Sakata. 4, 14

RAA Reactor Antineutrino Anomaly. 8, 11
RW rear wall. 14, 15

SM standard model of particle physics. 1, 3–5

UHV ultra high vacuum. 15

WGTS windowless gasseous tritium source. 14–16

List of Figures

List of Tables