Technical University of Munich
School of Natural Sciences

Master's Thesis in Nuclear, Particle and Astrophysics

Sterile Neutrino Search with Neural Networks at the KATRIN Experiment

Suche nach Sterilen Neutrinos mit Neuronalen Netzen am KATRIN Experiment

Luca Fallböhmer

10 January 2025

Primary reviewer: Prof. Dr. S. Mertens
Secondary reviewer: Prof. Dr. L. Heinrich
Supervisor: Dr. A. Onillon

Abstract

Right-handed neutrinos, also known as sterile neutrinos, with masses in the range of a few keV could pose as a suitable dark matter candidate. One way to experimentally access keV-scale sterile neutrinos is via a laboratory measurement of the β-decay spectrum. An admixture of a heavy neutrino mass eigenstate to the electron neutrino, which is created in β-decay, would result in a kink-like spectral distortion of the β-decay spectrum. With the TRISTAN detector upgrade, the KArlsruhe TRItium Neutrino (KATRIN) experiment plans to extend its measurement range to probe the entire tritium β-decay spectrum for such a sterile neutrino signature down to the parts per million level. The main goal of this thesis is to show that Neural Networks (NNs) are a powerful method to detect this signature. In this thesis, multiple NNs were trained to do Binary Classification (BC) on tritium β-spectra, including statistical fluctuations and optionally experimental effects or other perturbations. NNs are demonstrated to achieve a statistical sensitivity to the sterile parameters that is comparable to that of more conventional χ^2-based methods, corresponding to an active-to-sterile neutrino mixing angle down to $\sin^2(\theta) = 2 \times 10^{-7}$ for one year of measurement at 95 % confidence level (CL). It is demonstrated that the NN approach is largely insensitive to experimental effects and their uncertainties, as well as modelling inaccuracies that result in smooth spectral distortions.

Contents

Chapter 1

Introduction and Objectives

When Fritz Zwicky observed the velocity dispersion of galaxies in the Coma cluster back in 1933, he noticed that the motions he observed could only be explained by a large amount of mass that could not be seen. He termed the invisible matter "Dunkle Materie", which translates to Dark Matter (DM) [1]. Over the years, observables on many scales such as galaxy rotation curves, gravitational lensing effects and anisotropies in the Cosmic Microwave Background (CMB) provided more evidence for the existence and abundance of DM. Current results estimate that around 84 % of all matter in the universe is DM [2]. While there are hints for the particle nature of DM, e.g. from observations of the Bullet cluster [3], its true nature is still largely unclear. Promising particle candidates include Weakly Interacting Massive Particles (WIMPs), axions and sterile neutrinos [4]. These particles, however, are not included in the Standard Model of Particle Physics (SM) and are a sign of physics beyond the SM.

Additionally, the observations of neutrino oscillations indicate a non-vanishing neutrino mass, opposite to the predictions of the SM. By extending the SM with additional right-handed neutrinos, also known as sterile neutrinos, both the non-vanishing neutrino mass and the true nature of DM can be explained. Within the Neutrino Minimal Standard Model (νMSM), two sterile neutrinos with masses exceeding 1 GeV are proposed to account for the small masses of the active neutrinos [5]. The mass of the third sterile neutrino, however, needs to lie in the keV-range or slightly above to account for the observed dark matter density in the universe [6]. A sterile neutrino does not participate in the weak, strong, and electromagnetic interaction. However, their signature can be detected in a laboratory-based experiment due to their mixing with the three (active) left-handed neutrino flavors.

With the TRISTAN detector upgrade, the KATRIN experiment [7] will provide an opportunity to search for keV-scale sterile neutrinos. KATRIN is designed and currently operates to measure the neutrino mass through beta spectroscopy of the gaseous tritium decay. The latest result yielded the world leading upper limit on the neutrino mass of $m_{\bar{\nu}_e} < 0.45\,\mathrm{eV}$ (90 % CL)[1]. After the neutrino mass measurement campaign is completed at the end of 2025, the KATRIN beamline will be

[1] Here, and in the rest of this thesis, natural units ($\hbar = c = e = 1$) are used.

modified and upgraded with the TRISTAN detector system. While the kinematic endpoint of the tritium β-electron energy spectrum was investigated with an integral measurement for the neutrino mass campaigns, the search for keV sterile neutrino will be conducted through a differential measurement covering nearly the entire tritium spectrum. The TRISTAN detector and new data aquisition (DAQ) system are equipped to handle the very high rates that result from the increased energy range and provide their own energy resolution. In comparison to the neutrino mass measurement, which is limited by statistical uncertainties due to the low count rates at the endpoint, the keV sterile neutrino search is expected to be dominated by systematic effects and their uncertainties.

The sterile neutrino signal is a distinct step-like distortion in the β-spectrum. KATRIN has the statistical power to probe such distortions at the parts-per-million level. Such signals can be overshadowed by the systematic effects and the uncertainties associated to them, which act on the spectrum at the percent level and therefore require very accurate modelling and robust uncertainty estimation. The current modelling approach is constrained by excessive computational requirements and inherent inaccuracies, which makes the sterile neutrino search on measured data extremely challenging. Thus, the purpose of this thesis was to develop and investigate a "direct" search for the spectral distortion introduced by a sterile neutrino using NN, which are known for their ability to discriminate complex patterns in data [8, 9].

Chapter 2 of this work will give an overview on neutrino physics and sterile neutrino dark matter. Then, chapter 3 introduces the KATRIN experiment and TRISTAN detector upgrade, as well as the relevant systematic effects and their modelling. Chapter 4 provides an overview on NNs and the analysis methods used in this work. Chapter 5 then conducts sensitivity studies to assess how suitable NNs are for a direct search of the sterile neutrino signature.

Chapter 2

Neutrino Physics

The following chapter provides a short overview of the discovery of the neutrino through electron β-decay (section 2.1). It then continues with the theory of neutrino mixing and oscillations (section 2.2), which show that the neutrino is not a massless particle. How the neutrino masses could be generated and measured is discussed next (section 2.3). A sterile neutrino is a possible explanation for the neutrino mass and can also serve as a dark matter candidate, which is further explored in the final part of this chapter (section 2.4).

2.1 Historical Overview

The first observation of the energy spectrum of electrons emitted in radioactive β-decay by J. Chadwick in 1914 [10] marks the dawn of neutrino physics. Back then, β-decay was thought to be a two-body decay where the decaying nucleus only emits an electron, which can then only receive a fixed decay energy due to energy-momentum conservation. Thus, the energy spectrum was expected to be monochromatic. Chadwick, however, measured a continuous spectrum, which appeared to be violating energy-momentum conservation, as seen in figure 2.1. The possibility of explaining the observed spectrum by considering secondary effects that broaden a monoenergetic line was ruled out several years later [11, 12]. W. Pauli postulated the solution to this conundrum in 1930 [13]: By introducing a third electrically neutral particle that is created along with the electron, β-decay turns into a three body process. The energy of the decay is distributed among the decay products, explaining the continuous energy spectrum of the electron. The third particle is the neutrino, and is very evasive due to its lack of charge and would likely not manifest in calorimetric measurements, while carrying away the missing decay energy. It should also have spin-1/2 to conserve angular momentum. The discovery of further lepton generations as well as the concept of lepton number conservation led to today's name for this particle, electron antineutrino ($\bar{\nu}_e$).

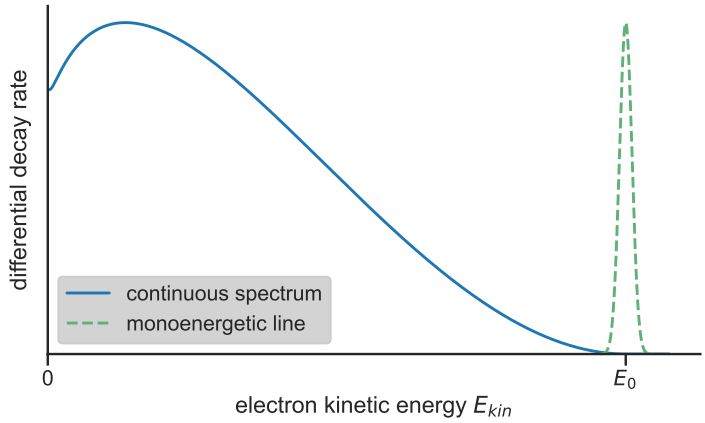


Figure 2.1: The continuous spectrum of the electron kinetic energy in β^--decay. In contrast to that, the monoenergetic line spectrum at the kinematic endpoint energy E_0 is shown in green, which would be expected if β-decay were a two body process.

Theory of β-decay

Following Pauli's suggestion, the first theoretical description of β-decay was given by E. Fermi in 1934 [14]. Here, a neutron is converted into a proton or vice versa, and electrons and neutrinos (or their antiparticles) are created in a charge and lepton number conserving manner. In practice, a neutron-rich nucleus N with atomic number Z decays to a more stable state by turning a neutron into a proton via β^--decay. During the decay process, an electron e^- and electron antineutrino $\bar{\nu}_e$ are created:

$$ {}^{A}_{Z}\mathrm{N} \longrightarrow {}^{A}_{Z+1}\mathrm{N}'^{+} + \mathrm{e}^- + \bar{\nu}_e + Q. \tag{2.1} $$

The decay energy Q is the surplus energy after subtracting the masses of the electron and its antineutrino from the difference in binding energy of the mother and daughter nucleus and the mass difference of neutron and proton. This energy is distributed randomly to the daughter particles and manifests as their kinetic energy. If, by chance, the neutrino is created almost at rest, the electron obtains the highest possible kinetic energy denoted by $E_0 = Q - E_{\mathrm{rec}}$, where E_{rec} is the recoil energy the daughter nucleus receives. The kinematic endpoint E_0 is shifted towards lower energies in the case of a massive electron antineutrino, as Q will be further reduced by the neutrino mass, thus lowering the highest possible kinetic energy for the electron, as seen in figure 2.2. The reverse of this process also exists and is called β^+-decay. Here, a proton in a nucleus turns into a neutron whilst emitting a positron and an electron neutrino.

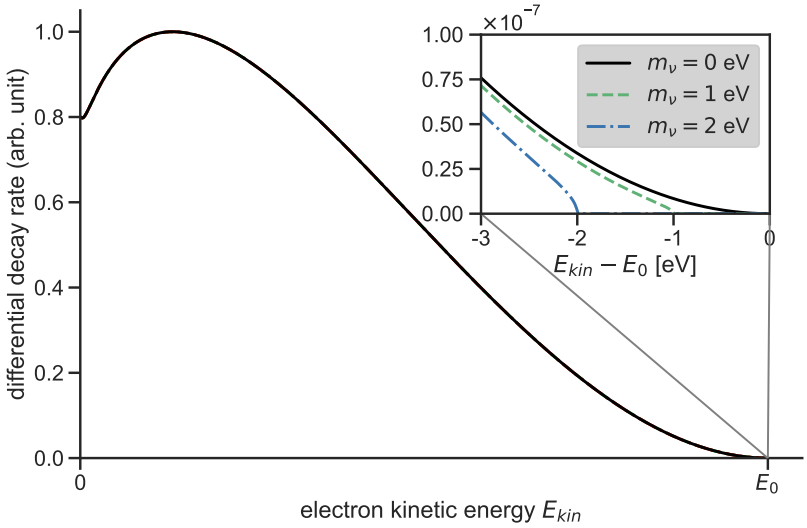
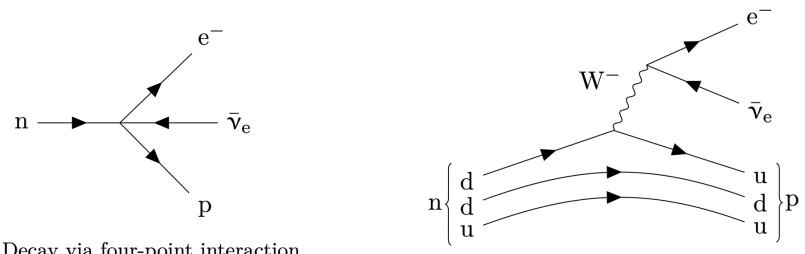


Figure 2.2: To account for an effective neutrino mass m_ν, the kinematic endpoint of the electron β-decay spectrum is shifted towards lower energies, and the spectral shape close to the respective endpoint is distorted.

In detail, Fermi describes β^--decay as a four-point interaction, meaning the spinor fields that represent the incoming and outgoing particles directly couple with each other without a mediating virtual boson. The modern understanding of this decay is different, although Fermi's four-point interaction can be seen as the low energy limit of the Glashow-Weinberg-Salam (GWS) model, when the energies are not high enough to resolve the propagator of the $W^\pm$-Bosons. The GWS model is part of the SM of particle physics and describes the electroweak sector. It shows that β^--decay can be seen as the transition of a down-quark in a neutron to an up-quark via an interaction with a W^--Boson, converting the neutron into a proton. The virtual W^--Boson then creates an electron and an electron antineutrino, as can be seen in figure 2.3.

The Weak Interaction

The weak interaction is mediated by the massive $W^\pm$- and Z-Bosons, and thus β-decay is also a consequence of the weak force, as well as every other neutrino production mechanism. In 1957, the Experiment by C. S. Wu et al. discovered that the weak force is not invariant under parity transformation of the spatial coordinates [16]. In simpler words, the physics of a system are not the same when mirrored. The weak force is the only fundamental interaction with this property. An isolated system's



(a) Decay via four-point interaction

(b) Decay mediated by W-boson

Figure 2.3: The difference between Fermi's four point interaction theory (a) and its modern interpretation (b) for the first order Feynman diagrams of neutron β-decay. In the latter, the decay happens through the transition of a down-quark into an up-quark via emission of a virtual W^- boson, which then decays into a lepton pair. The left picture can be seen as an approximation of the right for energies much smaller than the W-boson mass. Picture taken from [15].

overall parity is the product of the parities of its subsystems, and was thought to be conserved when the system interacts. But, according to experimental results and the theoretical considerations from Lee and Yang [17], the weak force only allows particles with left-handed chirality or antiparticles with right-handed chirality to couple via the W-Boson. This built-in dependence on right- or left-handedness is not indifferent to mirror reflections, and therefore maximally violates parity. This discovery also had a big influence on the understanding of the neutrino in the SM. In 1958 M. H. Goldhaber et al. proved the helicity of the electron neutrino is negative [18], meaning only left-handed neutrinos are produced in the SM. Massless particles, such as the neutrino in the SM, move with the speed of light, and thus their chirality and helicity are identical. The reason being that it is not possible for an observer to boost to a reference frame moving faster than the spinning massless particle, which would make the particle appear to move backwards and thus reverse its helicity. As neutrinos are exclusively produced in the weak interaction, they are always left-handed, and consequently anti-neutrinos are always right-handed. In the SM, a massless right-handed neutrino would not be able to interact with other SM particles and would thus be called *sterile*.

In Fermi's theory, the weak interaction is described only using a pure vector current coupling to fermions, which preserves parity. Thus, to describe the interaction, the "V-A" theory was introduced [19, 20]. It incorporates vector and axial vector currents which have opposing parity and thus maximally violate parity.

Discovery of the Neutrino

The electron antineutrino was finally experimentally discovered more than twenty years after its first conception, by the groups of C. L. Cowan Jr. and F. Reines [21–23]. They carried out their measurements at the Savannah river nuclear reactor, which provides a large flux of electron antineutrinos to make up for how little they interact. The electron antineutrinos $\bar{\nu}_e$ from the fission products interact with the water target with dissolved cadmium chloride ($CdCl_2$) and produce a positron e^+ and a neutron n via inverse β-decay:

$$\bar{\nu}_e + p \longrightarrow e^+ + n. \tag{2.2}$$

The positron decelerates and annihilates with an electron almost immediately, creating two characteristic annihilation gamma rays, which are detected by the two scintillation detectors that the target is sandwiched in between. The water moderates the neutron until it is captured by the Cd, releasing delayed coincident de-exitation radiation. In combination, these two events made it possible to discriminate strongly against background and allowed to detect the electron antineutrino for the first time with significant confidence.

Soon after Pauli's hypothesis of the neutrino, the muon, which carries the second lepton flavor, was discovered in 1936. Back then, the question arose if there exists one universal neutrino associated to all lepton flavors or if there are three different neutrino flavors corresponding to the three different leptons, as neutrinos are also produced in other weak processes like pion decay:

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu. \tag{2.3}$$

This question remained unanswered until 1962, where M. Schwarz, L.M. Lederman and J. Steinberger discovered the muon neutrino while studying cosmic radiation [24]. They found that the neutrinos in the decay of charged pions, which were created as a consequence of shooting a 15 GeV proton beam at a beryllium target, produced muons via inverse beta decay,

$$\bar{\nu}_\mu{}^+ \, p \longrightarrow n + \mu^+ \tag{2.4}$$

but not electrons. The next advance came in 1975, where the tau lepton was discovered [25], which naturally led to the postulation of a tau neutrino. The tau neutrino would remain undiscovered until the turn of the millennium, when in 2000 the DONUT experiment at Fermilab finally confirmed the existence of the last missing lepton in the standard model [26]. The reason this was known being the 1989 ALEPH experiment at CERN, which showed that due to the width of the Z-Resonance, there have to be three active light neutrino flavors that couple to the Z-Boson [27].

Solar Neutrino Problem

The weak interaction is termed "weak" because its coupling constant is orders of magnitudes lower compared to that of the electromagnetic force, which itself is overshadowed by the strong force. Consequentially, only about 1 neutrino in every 10^{12} undergoes an interaction with the nucleons in fluids used to detect neutrinos. To circumvent this limitation, enormous fluxes of neutrinos and large detection chambers are needed for direct neutrino observation. A perfect candidate for a large neutrino flux is the sun. These solar neutrinos are created from fusion processes in the sun's core, mainly via the pp-chain [28]. It consists of multiple subprocesses, which can be summed up in the following reaction:

$$4\,\mathrm{p} + 2\,\mathrm{e}^- \longrightarrow {}^4\mathrm{He} + 2\,\nu_e + 26.73\,\mathrm{MeV}. \tag{2.5}$$

Thus, only electron neutrinos are created and propagate from the centre of the sun to the earth. The solar neutrino flux was measured in 1968 by R. Davis Jr. using inverse beta decay reactions with ^{37}Cl and counting the number of ^{37}Ar atoms [29] that are created:

$$\nu_e + {}^{37}\mathrm{Cl} \longrightarrow {}^{37}\mathrm{Ar} + \mathrm{e}^-. \tag{2.6}$$

Only approximately one-third of the expected 1.7 interactions per day were observed [30]. This large deficit of electron neutrinos was coined the "solar neutrino problem" after cross-validation with other experiments [31–33]. The pivot in this problem was that Davis' experiment was only sensitive to electron neutrinos and not other neutrino flavors.

2.2 Neutrino Mixing and Oscillations

The Sudbury Neutrino Observatory experimentally solved the solar neutrino problem in 2001 by performing a measurement that is sensitive to not only the electron neutrino but also the two other neutrino flavors [34]. They discovered that the large fraction of solar electron neutrinos that was missing in the previous experiments was made up by the other neutrino flavors, which means that some electron neutrinos ν_e changed into muon neutrinos ν_μ and tau neutrinos ν_τ on their way towards the earth. With their measurement, they confirmed the hypothesis that neutrinos can oscillate between flavors during their propagation over long distances. These oscillations were later also observed on different length-scales by several other experiments using neutrinos from nuclear reactors, cosmic ray interactions in the atmosphere and particle accelerators [35–37]. The phenomenon of neutrino oscillations and consequentially neutrino mixing is essential to understand why neutrinos have mass and how a possible sterile neutrino signature looks like in a β-decay spectrum.

2.2.1 Theory of Neutrino Oscillations

Neutrinos oscillate between flavors because their flavor eigenstates $|\nu_l\rangle$ (l = e, μ, τ) are not equal to their mass eigenstates $|\nu_i\rangle$ (i = 1, 2, 3), which are the eigenstates neutrinos propagate through space in. The mass eigenstates are a quantum mechanical superposition of the flavor eigenstates and vice-versa. They can each be described using the corresponding mixing amplitudes U_{li}:

$$|\nu_i\rangle = \sum_l U_{li} |\nu_l\rangle \qquad \text{and} \qquad |\nu_l\rangle = \sum_l U_{li}^* |\nu_i\rangle . \tag{2.7}$$

From this description, it is easy to see that the mass eigenstates are simply rotated with respect to the flavor eigenstates. This rotation is described by the complex 3×3 Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, with the amplitudes U_{li} as entries [38]:

$$\begin{pmatrix} \nu_\mathrm{e} \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \underbrace{\begin{pmatrix} U_\mathrm{e1} & U_\mathrm{e2} & U_\mathrm{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}}_{\text{PMNS - matrix}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} . \tag{2.8}$$

Assuming this matrix to be unitary and that only three neutrino mass eigenstates exist, the PMNS matrix can be parametrized using three different mixing angles (θ_{12}, θ_{13}, θ_{23}) and a complex phase δ, called the Dirac phase. As neutrinos only partake in the weak interaction, the creation and detection of a neutrino corresponds to a measurement of the flavor state, leading to a collapse of the superimposed wave function into one of the flavor eigenstates. A neutrino is created in one of its flavor eigenstates $|\nu_l\rangle$. It then propagates through time (and space). Assuming a plane wave ansatz for the Schrödinger equation adds complex phases to the mixing amplitudes of the mass eigenstates $|\nu_l\rangle$, which is made up of [38]:

$$|\nu_l(t)\rangle = \sum_i U_{li} e^{-iE_i t} |\nu_i\rangle . \tag{2.9}$$

Thus, each neutrino mass eigenstate picks up a species-dependent phase dictated by its dispersion relation $E_i = \sqrt{p_i^2 + m_i^2} \approx p + \frac{m_i}{2E}$. Assuming the neutrino to be relativistic, it is proportional to their mass eigenvalue m_i^2. This is what ultimately enables the flavor oscillations. As relativistic neutrinos travel close to the speed of light c, we can set time $t \simeq L$, with L being the distance the neutrino travelled (and $c = 1$). Pulling all this together, the transition probability between two flavor states l and k after a distance L becomes [38]:

$$\begin{aligned} P_{\nu_l \to \nu_k}(L, E) &= |\langle \nu_k(L) | \nu_l(0) \rangle|^2 \\ &= \sum_{ij} e^{-i(m_i^2 - m_j^2)L/(2E)} U_{ki} U_{kj}^* U_{li} U_{lj}^* . \end{aligned} \tag{2.10}$$

This probability only vanishes if the masses m_i and m_j are either zero or equal. It is an experimental fact that neutrino oscillations occur between all three flavors [39–42]. As a consequence, at least two of the three neutrino flavors have a different and non-zero mass. The amplitude of the transition is dictated by the PMNS matrix elements, which are themselves described by the mixing angles. When simplifying the treatment to two neutrino flavor eigenstates connected via a single mixing angle θ and with mass difference $\Delta m^2 = m_j^2 - m_i^2$, equation 2.10 simplifies to [43]:

$$P_{\nu_l \to \nu_k}(L, E) \approx \sin^2(2\theta) \sin^2\left(1.27 \cdot \Delta \mathrm{m}^2[\mathrm{eV}^2] \cdot \frac{L[\mathrm{km}]}{E[\mathrm{GeV}]}\right). \tag{2.11}$$

Here, it becomes apparent the transition probability is sinusoidal as a function of traveled distance L and directly connected to the size of the mass splittings. Unfortunately, as visible from equations 2.10 and 2.11, neutrino oscillation experiments are only sensitive to the mass differences and not the absolute neutrino masses. In conclusion, neutrino oscillation experiments provide information on the mixing angles, Dirac phase and mass squared differences, which are evidence of a non-zero neutrino mass. The question now being where the neutrino masses come from, and how they can be experimentally probed, which will be covered in the next section.

2.3 Massive Neutrinos

To formulate possible explanations for the neutrino masses it is important to first understand where other particles get their masses from and see why the neutrino is taken to be massless in the SM.

2.3.1 The Standard Model

The SM includes two types of particles: The matter particles (fermions) and the force carriers that mediate the interaction between the matter particles (bosons), as can be seen in figure 2.4. The former all have spin 1/2 and are represented by spinor fields, while the latter have to be subdivided into the vector bosons and a scalar boson. The vector bosons have spin 1 and are described by vector fields, and the scalar boson, more precisely called Higgs boson, has spin 0 and is described by a scalar field. It is the Higgs bosons which gives mass to all elementary particles except the photon, gluon and neutrinos, by a mechanism explained in the next section. A particle in the SM can be seen as a quantized excitation of their respective field. Therefore, the SM is fundamentally a quantum field theory. It is constructed on symmetries, which lead to conserved quantities via E. Noether's theorems [44]. Mathematically, a symmetry is an operation that when performed on the Lagrangian of a system leaves it invariant. Of particular interest are local gauge symmetries, which dictate

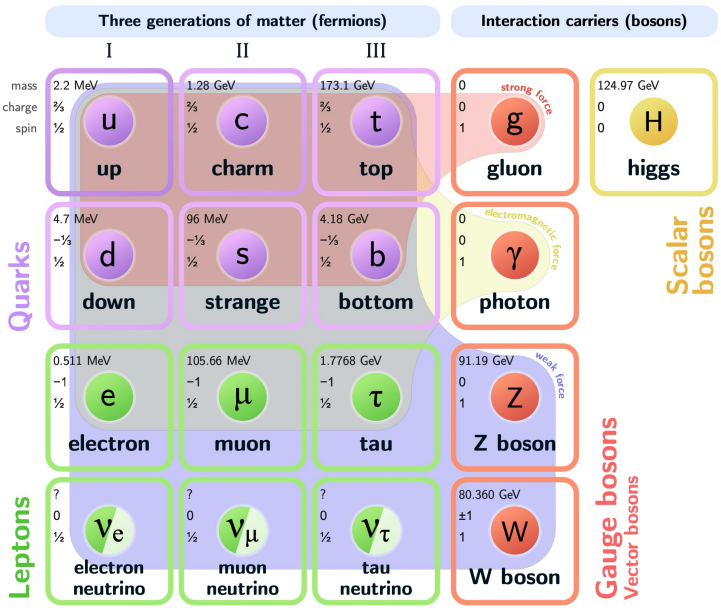


Figure 2.4: The different particles included in the Standard Model of particle physics. The Interactions and their mediating bosons are indicated by the shaded regions. Because the weak interaction does not conserve parity, only left-handed neutrinos exist, which is indicated by the half circles. Figure adapted from [15].

the form of the interaction between matter and force carriers. The only interaction that breaks the parity symmetry is the weak force, which will be further explained in the following section.

2.3.2 Electroweak Force and Masses in the SM

The weak force is the only force that does not conserve parity and only interacts with left-handed particles, as discussed in 2.1.

Chiral Structure of the Weak Force

To incorporate this dependence on the chirality of the particles, the SM is constructed as a chiral theory, where the (spinor fields of the) left-handed fermions are grouped in left-handed doublets and right-handed singlets. For the leptons, this results in the

following structure:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}_L \qquad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}_L \qquad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}_L \qquad (2.12)$$

$$e_R \qquad \mu_R \qquad \tau_R$$

The doublets are susceptible to $SU(2)$ transformations and therefore the weak interaction and are thus defined to be left-handed. Because the charged leptons also partake in electromagnetic interactions, which conserve parity and consequently are not chiral, they also have singlet representations which are invariant under $SU(2)$ transformations and therefore defined as right-handed. The right-handed singlets do not partake in weak interactions.

The symmetries of the SM can also be broken, and there are different ways to break a symmetry. An important example would be the inclusion of a mass term for a lepton l to the SM Lagrangian, which would have the general shape

$$\mathcal{L}^l_{\text{mass}} = -m_l(\bar{l}_L l_R + \bar{l}_R l_L). \qquad (2.13)$$

This term would not be invariant under a $\mathrm{SU}(2)_{I_3} \times \mathrm{U}(1)_{Y_W}$ gauge symmetry, which, if conserved, gives rise to the electroweak gauge vector fields (i.e. vector bosons) and interactions with the weak isospin I_3 and weak hypercharge Y_W as conserved Noether charges. This is because the fields l_L and l_R belong to different $SU(2)_L$ representations and have different $U(1)_Y$ hypercharges. Thus, is there any other way to introduce mass terms to the SM Lagrangian without breaking this symmetry? The key to this problem is so-called Spontaneous Symmetry Breaking (SSB). SSB happens when only the vacuum configuration, meaning the field configuration with the minimal energy (also called ground state), is not invariant under a certain symmetry.

The Higgs Mechanism

In 1964, Higgs, Englert and Kibble introduced the Higgs mechanism, which provides a method of giving gauge fields a mass while maintaining gauge invariance, all made possible by SSB [45–47]. Later, Glashow, Weinberg and Salam applied the Higgs mechanism to the aforementioned $\mathrm{SU}(2)_{I_3} \times \mathrm{U}(1)_{Y_W}$ gauge symmetry, marking the birth of the electroweak theory [48–51]. As the name suggests, this theory unifies the electromagnetic and weak interactions above a certain energy threshold, characterized by the transition temperature $T_c \approx 160\,\mathrm{GeV}$. Below this threshold, SSB takes place, meaning the $\mathrm{SU}(2)_{I_3} \times \mathrm{U}(1)_{Y_W}$ gauge symmetry is broken until only the $U(1)_{\text{em}}$ symmetry of quantum electrodynamics (QED) remains. This process is also called electroweak symmetry breaking. Pivotal for the Higgs mechanism is the

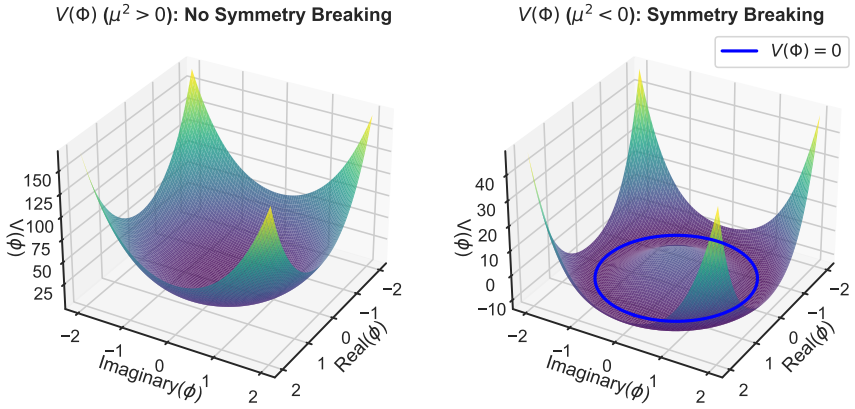


Figure 2.5: The shape of the Higgs potential $V(\Phi)$ for different signs of μ^2. For negative μ^2, a degenerate family of minima is created, indicated by the blue circle. SSB takes place when a particular point on the blue circle is chosen as a minimum.

shape of the Higgs potential, which depends on the complex scalar Higgs doublet Φ:

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \qquad \text{with} \qquad \Phi = \begin{pmatrix} \phi_0 \\ \phi_1 \end{pmatrix} \tag{2.14}$$

with $\mu^2, \lambda > 0$, and $\phi_{0,1}$ denoting complex scalar Higgs fields. The potential is as shown in figure 2.5, where the potentials minima, corresponding to the vacuum state, are at $\Phi = 0$ and a degenerate family of minima lie on a circle on the complex plane of ϕ with modulus $|\langle 0| \phi |0\rangle| = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}$. The SSB takes place when a particular direction in the internal $SU(2)$ doublet space of the scalar Higgs field ϕ is chosen for the minimum of the Higgs potential. The Higgs doublet Φ can then be expanded around its new vacuum expectation value $v \approx 246\text{GeV}$ and written as:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \rho(x) \end{pmatrix}. \tag{2.15}$$

Lepton Mass Generation

The Higgs field Φ after electroweak symmetry breaking can now be used to write gauge invariant couplings linking l_L and l_R. These new additional terms in the

Lagrangian are called Yukawa interactions. As an example, considering electrons ($l = \mathrm{e}$), the Yukawa interaction term becomes:

$$\begin{aligned}\mathcal{L}^{\mathrm{e}}_{\mathrm{Yukawa}} &= -\frac{\lambda_e}{\sqrt{2}}\begin{pmatrix}\bar{\nu}_L & \bar{e}_L\end{pmatrix}\begin{pmatrix}0\\ v+\rho\end{pmatrix} e_R - \frac{\lambda_e}{\sqrt{2}}\,\bar{e}_R\begin{pmatrix}0 & v+\rho\end{pmatrix}\begin{pmatrix}\nu_L\\ e_L\end{pmatrix} \\ &= \underbrace{-\frac{\lambda_e}{\sqrt{2}} v\left(\bar{\mathrm{e}}_L \mathrm{e}_R + \bar{\mathrm{e}}_R \mathrm{e}_L\right)}_{\mathcal{L}^{\mathrm{e}}_{\mathrm{mass}}} \underbrace{-\frac{\lambda_e}{\sqrt{2}}\rho\left(\bar{\mathrm{e}}_L \mathrm{e}_R + \bar{\mathrm{e}}_R \mathrm{e}_L\right)}_{\mathcal{L}^{\mathrm{e}}_{\mathrm{int}}},\end{aligned} \tag{2.16}$$

with λ_{e} being the Yukawa coupling for electrons, resulting in an electron mass of $m_{\mathrm{e}} = \frac{\lambda_{\mathrm{e}} v}{\sqrt{2}}$. The product of the $SU(2)$ left-handed fermion doublet with the Higgs doublet results in a $SU(2)$ singlet, making the individual terms invariant under $SU(2)_{I_3}$ transformations. By defining the weak hypercharge of Φ to be $Y_W = 1/2$, the hypercharges of the individual terms all sum to 0, making them also invariant under $U(1)_{Y_W}$ transformations. This mechanism is rigorously tested at the LHC, and the experimental data agrees very well with this prediction [52]. Noticeably, there are *no mass terms for the neutrino*, leading to it being treated as a *massless* particle in the SM.

2.3.3 Mechanisms for Neutrino Mass Generation

A neutrino mass would introduce beyond the SM (BSM) terms into the Lagrangian. There are different approaches to adding such a mass term.

Dirac mass term

The straightforward method of generating neutrino masses is to allow them to couple to the Higgs field by introducing three right-handed $SU(2)$ neutrino singlets:

$$\left(\nu_{\mathrm{e}}\right)_R \qquad \left(\nu_\mu\right)_R \qquad \left(\nu_\tau\right)_R \ . \tag{2.17}$$

If the neutrino fields ν_l are then treated as a spin-1/2 Dirac field (like e.g. the electron), then a mass term may be generated identically to the other leptons in the SM by a Yukawa term and the Higgs mechanism:

$$\mathcal{L}^{\nu}_{\mathrm{Dirac}} = -\frac{\lambda_\nu}{\sqrt{2}} v\left(\bar{\nu}_L \nu_R + \bar{\nu}_R \nu_L\right). \tag{2.18}$$

The resulting term is called Dirac term, with the resulting mass being called Dirac mass $m_D = \lambda_\nu v/\sqrt{2}$. Because these right-handed neutrinos do not couple electromagnetically or to neutral or charged weak currents (being singlets under $SU(2)$) nor do they carry color charge, they are referred to as *sterile neutrinos*. However, they do couple to the Higgs field, which provides them a mass and consequently

allows them to interact gravitationally. The only way for these sterile neutrinos to interact with other SM particles would be through mixing with the active neutrino flavors, by expanding the PMNS matrix, which will be further explained in section 2.4. The main problems with this approach being that the right-handed neutrino has not been observed in nature so far, and that e.g. direct neutrino mass measurements like the KATRIN experiment constrain the effective antineutrino mass to be below 0.45 eV (90 % CL) [53], resulting in a Yukawa coupling constant many orders of magnitude lower than that of other fermions like the electron or top quark $\lambda_{\nu_l} \sim \lambda_\mathrm{e} \cdot 10^{-6} \sim \lambda_t \cdot 10^{-12}$.

Majorana Mass Term

Alternatively, a neutrino mass term could be introduced without relying on a right-handed neutrino. The prerequisite here being that neutrinos are their own antiparticles and can thus be described via a spin-1/2 Majorana field. Neutrinos can be their own antiparticles because they do not carry electric charge. This allows the construction of mass terms that else would violate lepton number conservation:

$$\mathcal{L}^{\nu_L}_\mathrm{Majorana} = -\frac{m_L}{2}\left(\bar{\nu}^C_L \nu_L + \bar{\nu}_L \nu^C_L\right). \tag{2.19}$$

The superscript C denotes charge conjugation, which is the operation that turns particles into antiparticles and vice-versa.

The neutrino being a Majorana particle does not exclude the existence of right-handed neutrino singlets, thus also the following Majorana mass term is possible:

$$\mathcal{L}^{\nu_R}_\mathrm{Majorana} = -\frac{m_R}{2}\left(\bar{\nu}^C_R \nu_R + \bar{\nu}_R \nu^C_R\right). \tag{2.20}$$

This term in particular does not require any interactions with the Higgs field, allowing the Majorana mass m_R to be arbitrarily large.

Again, the problem with this approach is that there is still no experimental evidence that the neutrino is actually a Majorana particle. A possible way to test this is by observing the so-called neutrinoless double-β decay, which will be explained briefly in section 2.3.4.

Seesaw Mechanism

The Dirac and Majorana mass terms above do not exclude each other and can be combined and written compactly [38]:

$$\mathcal{L}^{\nu}_\mathrm{mass} = \mathcal{L}^{\nu}_\mathrm{Dirac} + \mathcal{L}^{\nu_L}_\mathrm{Majorana} + \mathcal{L}^{\nu_R}_\mathrm{Majorana} = -\frac{1}{2}\begin{pmatrix}\bar{\nu}_L & \bar{\nu}^C_L\end{pmatrix}\begin{pmatrix}m_L & m_D \\ m_D & m_R\end{pmatrix}\begin{pmatrix}\nu^C_R \\ \nu_R\end{pmatrix} + h.c. \tag{2.21}$$

Using this mass term, the neutrinos Yukawa coupling can be made more compatible with those of other particles while keeping the neutrino mass small. In order to achieve this, the simplifying assumption that $m_L = 0$ and $m_D \ll m_R$ is made. In this scenario in particular, by setting $m_L = 0$ it is assumed that the active neutrino is not a Majorana particle. The eigenvalues of the resulting mass matrix then become

$$m_1 = \frac{m_D^2}{m_R} \qquad \text{and} \qquad m_2 = m_R \left(1 + \frac{m_D^2}{m_R^2}\right) \approx m_R \ , \tag{2.22}$$

leading to a small effective mass of the active neutrino m_1 suppressed by the large Majorana mass associated with the sterile neutrino state m_R. This means the bigger m_R, the smaller m_1 gets, hence the name *seesaw mechanism*. To achieve active neutrino masses on the sub-eV scale, the sterile Majorana mass needs to be $\mathcal{O}(100\,\mathrm{MeV})$ or well above, depending on the details of the theory [54, 55]. This is only one of several seesaw mechanisms, more detailed information on the other types can be found in [56].

2.3.4 Neutrino Mass Measurements

As discussed in section 2.2.1, neutrino oscillation experiments are only sensitive to the squared mass differences of the single flavors, and not the absolute mass scale. There are mainly three ways to determine the absolute neutrino mass scale experimentally and many experiments, several of which will be briefly mentioned in the following.

Cosmology

Cosmological structure formation in the early universe is heavily influenced by the neutrino, due to their excess after the Big Bang. By acting as Hot Dark Matter (HDM), the relic neutrinos washed out structures on scales smaller than their free streaming length, which depends on their mass [57]. The CMB reflects the anisotropy of the matter distribution, and the respective scale at which these anisotropies arise depends on the neutrino mass. Thus, by measuring the multipole spectrum of the CMB a limit can be put on the sum of the neutrino masses m_i:

$$m_{\mathrm{tot}} = \sum_i m_i < 0.39\,\mathrm{eV}\ (95\,\%\ \mathrm{CL}), \tag{2.23}$$

additionally including observations from lensing and baryonic acoustic oscillations further reduces the limit to $m_{\mathrm{tot}} < 0.11\,\mathrm{eV}$ (95 % CL) [2]. However, these limits largely depend on dataset selection, cosmological model and analysis, making direct laboratory-based measurements essential.

Neutrinoless Double β-Decay

A neutrinoless double β-decay would be a direct indicator of the neutrino being a Majorana fermion, with the half-life of the decay relating to the absolute neutrino mass scale. Double β-decay happens in a select number of atoms in which single β-decay is forbidden due to energy conservation, but the β-decay of two nucleons (here e.g. neutrons) at once is allowed:

$$(\mathrm{Z}, \mathrm{A}) \longrightarrow (\mathrm{Z}+2, \mathrm{A}) + 2\,\mathrm{e}^- + 2\,\bar{\nu}_e. \tag{2.24}$$

This decay of two neutrons into protons raises the proton number Z by two and creates two electrons and electron antineutrinos. If, however, the neutrino is a Majorana particle and thus its own antiparticle, the two neutrinos are virtually exchanged in the decay. As a result, no neutrino would be emitted, leading to the name of neutrinoless double-β-decay:

$$(\mathrm{Z}, \mathrm{A}) \longrightarrow (\mathrm{Z}+2, \mathrm{A}) + 2\,\mathrm{e}^-. \tag{2.25}$$

To detect this decay, the continuous double-β-decay spectrum is measured, searching for a peak at the corresponding Q-value. So far, no such signal has been observed, with the current leading half-life limit for the decay being set to $T_{1/2} > 1.8 \cdot 10^{26}\,\mathrm{s}$ (90 % CL) by the GERDA experiment [58] using the double-β decay of ^{76}Ge. The half-life then relates to the effective Majorana mass of the neutrino $m_{\beta\beta}$, which is a coherent sum of the neutrino mass eigenstates. Here, the complex Dirac phase and the two additionally introduced Majorana phases may lead to interference effects. The above-mentioned half-life constrains the effective Majorana mass [58]:

$$m_{\beta\beta} < 0.18\,\mathrm{eV}\ (90\,\%\ \mathrm{CL}). \tag{2.26}$$

Direct Kinematic Measurement

A popular method for a direct kinematic measurement of the neutrino mass is beta spectroscopy. Beta spectroscopy provides a model-independent and laboratory-based way of measuring the neutrino mass. As previously outlined in section 2.1, a non-zero neutrino mass would leave a distortion at the kinematic endpoint of the β-decay spectrum, as seen in figure 2.2. The observable in this case being the effective electron neutrino mass

$$m_{\nu_\mathrm{e}} = \sqrt{\sum_i |U_{\mathrm{e}i}^2| m_i^2}. \tag{2.27}$$

This is because current experiments only posses an energy resolution of $\mathcal{O}(\mathrm{eV})$, which is too coarse to resolve the small mass splittings. The KATRIN experiment currently

sets the lowest limit on the effective electron antineutrino mass using molecular Tritium as β-emitter [53]:

$$m_{\bar{\nu}_e} < 0.45\,\text{eV}\ (90\,\%\ \text{CL}). \tag{2.28}$$

The experiment will be explained in more detail in section 3.

2.4 Sterile Neutrinos and Dark Matter

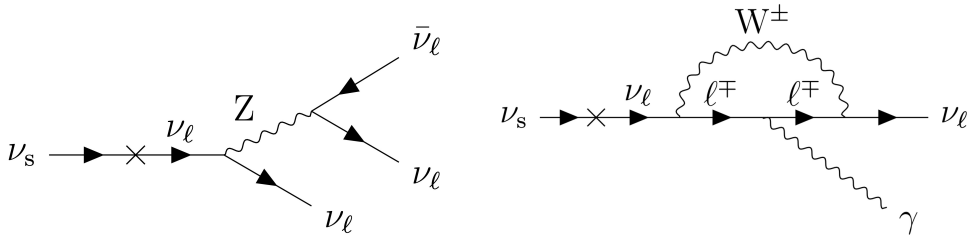
Motivated by the missing neutrino masses in the SM, so-called sterile neutrinos can be introduced. These are beyond the SM particles, which do not interact directly via any SM force. Right-handed neutrinos would be an example of such sterile neutrinos. Since all other SM particles exist with both right- and left-handed chirality, they seem to be a natural and minimal extension to the SM. Nonetheless, they are flexible enough to explain numerous open questions. Very heavy sterile neutrinos of up to $\mathcal{O}(10^{15}\,\text{GeV})$ could explain the small masses of the active neutrino flavors using the seesaw mechanism outlined in 2.3.3, as well as give insight on the matter-antimatter asymmetry in our universe [59]. Very light sterile neutrinos ($\mathcal{O}(\text{eV})$), on the other hand, provide a solution to anomalies found in neutrino oscillation experiments conducted over short distances, like the reactor anomaly [60]. However, the focus of this thesis lies on sterile neutrinos with masses on a keV-scale, as they are viable candidates for both Cold and Warm Dark Matter (WDM), depending on their production mechanism in the early universe [61, 62].

2.4.1 Dark Matter

Around 84 % of matter in our universe is DM [2]. As the prefix "dark" implies, this type of matter does not interact with photons. Its presence, however, can be clearly detected due to its gravitational interactions on many scales. DM explains the rotation curves of spiral galaxies, the velocity dispersion in galaxy clusters, gravitational lensing effects, the structure formation in the early universe and resulting temperature anisotropies of the CMB [4]. The true nature of dark matter is still up to debate. The observation of the bullet cluster effect however strongly points towards the particle nature of DM [63]. There are many possible candidates, like supersymmetric particles, extra-dimensional particles or axions [4], but the focus of this thesis will be on keV-scale sterile neutrinos.

2.4.2 KeV-scale Sterile Neutrinos

As mentioned above, keV-scale sterile neutrinos are sufficiently heavy and do not interact directly with SM particles, making them excellent candidates for WDM. WDM would alleviate shortcomings of most Cold Dark Matter (CDM) models on galactic



(a) Primary decay channel **(b)** Secondary decay channel

Figure 2.6: Leading Feynman diagrams of sterile neutrino decay channels. The primary decay channel of a sterile neutrino ν_s into three (active) neutrinos (a). Secondary decay channel of a sterile neutrino ν_s to a lighter (active) neutrino flavor under the emission of an X-ray. Figures taken from [6, 65]

sub-Mpc scale by explaining the smaller number of observed satellite dwarf galaxies and the density profiles of low mass galaxies (cusp-core problem) [64]. The sterile neutrino is made tangible by introducing additional mass eigenstates that mix with the active neutrino flavors. Like in section 2.2.1, the mixing amplitude can be characterized by a mixing angle or equivalently its squared sine, $\sin^2(\theta)$. This mixing allows the heavy steriles to decay into active neutrinos. The main decay channel at energies below twice the electron's mass is $\nu_s \longrightarrow \bar{\nu}\,\nu\nu$, dominating the sterile neutrino's lifetime. The Feynman diagram of the decay is shown in figure 2.6. To be DM, sterile neutrinos have to be stable on a timescale larger than that of the universe, limiting the frequency of this decay channel, and thus also the size of the mixing amplitude of active-to-sterile mixing. An alternative, less dominant decay channel would be the radiative decay $\nu_s \longrightarrow \nu\gamma$ as illustrated in figure 2.6, leading to a monoenergetic X-ray line at half the original sterile neutrinos mass. The non-observation of these lines also constrain the mixing amplitudes. Some X-ray telescopes have detected an otherwise unidentified X-ray line at 3.5 keV, which would correspond to a sterile with $m_4 \approx 7\,\text{keV}$ [66]. However, more recent studies exclude the DM origin of this line [67]. Again, the constraints on the sterile parameter space from cosmology are often model dependent, making laboratory searches strongly necessary. The main focus of this thesis is on such a laboratory-based search for keV-scale sterile neutrinos using the KATRIN experiment, which will be thoroughly explained in the following chapter.

Chapter 3

The KATRIN Experiment and TRISTAN Project

The **KA**rlsruhe **TRI**tium **N**eutrino (KATRIN) experiment was designed to measure the kinematic endpoint of the Tritium β-decay spectrum to search for the effective electron antineutrino mass $m_{\bar{\nu}_e}$. The measurement is still ongoing, with recent results setting the current world-leading upper limit $m_{\bar{\nu}_e} < 0.45\,\text{eV}$ (90 % CL) and a possible final sensitivity of $m_{\bar{\nu}_e} < 0.3\,\text{eV}$ [53] after 1000 days of data taking.

The effect of a $m_{\bar{\nu}_e} > 0$ leads to a shape distortion at the kinematic endpoint energy E_0 of the spectrum. In this region only a small percentage of decays happen. Therefore, an extremely luminous gaseous Tritium source achieving an activity of up to $10^{11}\,\text{Bq}$ [7] is needed. With such a strong source, still only ($\mathcal{O}$(1 events/sec)) arrive at the detector. To be sensitive to an $m_{\bar{\nu}_e}$ in the eV mass range, a Magnetic Adiabatic Collimation with Electrostatic filtering (MAC-E) spectrometer with an energy resolution of $\mathcal{O}$(1eV) is being used [7]. To achieve the sensitivity target, the KATRIN neutrino mass experiment plans to continue data taking until 2026 for a total amount of 1000 days of β-spectrum measurements. After the neutrino mass measurement is completed the KATRIN collaboration plans to upgrade the KATRIN detector and DAQ section to perform a keV-scale sterile neutrino search.

For measuring deeper into the Tritium β-decay spectrum, the Focal Plane Detector (FPD) currently in use can not handle the higher count rates. Additionally, the energy resolution of the FPD is not good enough to be used in a differential measurement. This is why once the neutrino mass measurements are concluded, the FPD will be exchanged with the TRISTAN detector along with other changes to the KATRIN beamline. This is called the TRISTAN upgrade. The TRISTAN detector is a Silicon Drift Detector (SDD), which provides a better energy resolution and can handle higher rates, enabling the search for sterile neutrinos with $m_4 \leq 18\,\text{keV}$ at a potential sensitivity of $|U_{es}|^2 = \sin^2(\theta) \sim 10^{-6}$ for the mixing angle [68].

The following chapter starts with the introduction of the measurement principles for a neutrino mass and sterile neutrino measurement (section 3.1). Next, it outlines the KATRIN experiment (section 3.2). Afterwards, the TRISTAN detector upgrade and beamline changes for a sterile neutrino measurement will be detailed in

section 3.3. Finally, the most important systematic effects and how they are modelled for TRISTAN using the package TRModel will be explained in section 3.4.

3.1 Tritium β-Decay

As outlined in section 2.3.4, the KATRIN experiment utilizes the decay of gaseous molecular Tritium (T_2) into Helium-3 (^{3}He) as a source of β-decay electrons:

$$T_2 \longrightarrow T^3He^+ + e^- + \bar{\nu}_e + Q(T), \tag{3.1}$$

with Q(T) being the surplus energy stemming from the mass difference between T and He^+. This energy is then distributed among the three decay products, accounting for the neutrino's and electron's kinetic energy and mass, and the recoil energy of $^3He^+$. The recoil energy is relatively small because $^3He^+$ is heavy compared to the other decay products, peaking at 1.72 eV at the kinematic endpoint $E_0 = 18.6\,\text{keV}$ [69]. Using T_2 has several advantages [70]. The decay is super-allowed, meaning it has a low half life of 12.33 years and thus a comparatively large activity, allowing to gather more statistics faster. Additionally, the kinematic endpoint energy for T_2 is low enough to emphasize the neutrino mass signature and allow for spectrometry using the MAC-E filter. Using Fermi's Golden Rule, the shape of the nuclear β-electron energy spectrum can be calculated [70], from which emerges that

$$\left(\frac{\mathrm{d}\,\Gamma}{\mathrm{d}E}\right)_\beta \propto \sum_i |U_{ei}|^2 \sqrt{E_\nu^2 - m_i^2} \qquad \text{with} \qquad E_\nu = E_0 - E. \tag{3.2}$$

This means that the measured β-decay spectrum is a superposition of the spectra of the individual neutrino mass eigenstates m_i, with the squared PMNS-Matrix elements U_{ei}, or equivalently mixing-angles, as respective weights. This is illustrated in figure 3.1.

As the neutrino mass differences Δm^2 are much smaller than the experimental resolution of $\mathcal{O}(1\,\text{eV})$, the individual spectra can not be resolved [71]. Instead, the spectral shape is simplified to an effective weighted neutrino mass:

$$m^2_{\bar{\nu}_e} = \sum_i |U^2_{ei}| m_i^2 \,. \tag{3.3}$$

The KATRIN experiment has managed to set an upper limit of $m_{\bar{\nu}_e} < 0.45\,\text{eV}$ (90 % CL) [53].

Searching for a sterile neutrino signature

If there were now an additional sterile neutrino eigenstate m_4, with a mass in the keV range, that mixes with the three masses of the active neutrinos under an angle

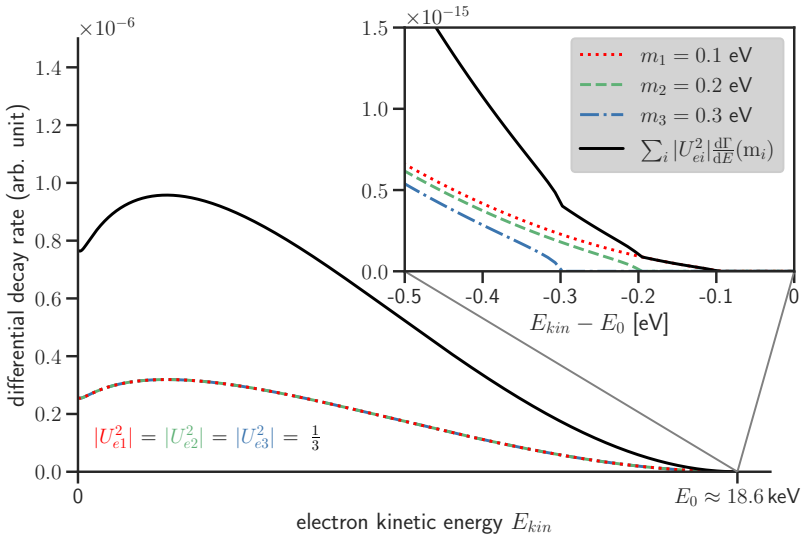


Figure 3.1: The total Tritium β-electron energy spectrum (black) as sum of the individual differential decay spectra of the neutrino mass eigenstates $\frac{\mathrm{d}\Gamma}{\mathrm{d}E}(\mathrm{m}_i)$ scaled by their mixing amplitudes $|U_{ei}|^2$ (red, green and blue). For KATRIN, the individual neutrino masses m_i are too small to be resolved experimentally. This simplifies the observed spectral shape to an effective weighted neutrino mass, differing from the one shown above. The neutrino masses and mixing amplitudes are chosen arbitrarily for illustrative purposes.

of $\sin^2(\theta)$, an additional term would appear in the differential decay rate. Like in figure 3.1, this results in a spectrum scaled by the mixing amplitude $|U_{\mathrm{e}4}^2|$ and with a kinematic endpoint at $E_4 = 18.6\,\mathrm{keV} - m_4$ being added to the β-decay spectrum. This is shown in figure 3.2. Thus, a keV sterile neutrino leaves a kink-like signature on the tritium β-decay spectrum, with the location depending on its mass m_4 and the size of the effect depending on the sterile-to-active mixing amplitude $\sin^2(\theta)$. An experiment using tritium β-decay is therefore sensitive to sterile neutrino masses $m_4 < 18.6\,\mathrm{keV}$. With the TRISTAN detector upgrade for the KATRIN experiment the collaboration explores the possibility of reaching a sensitivity to mixing angles down to $\mathcal{O}(10^{-6})$, which surpasses current laboratory limits by more than two orders of magnitude.

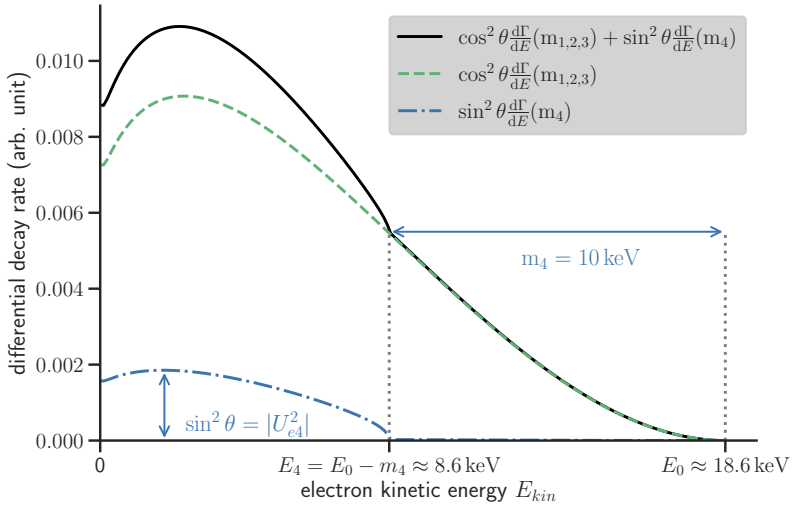


Figure 3.2: The differential Tritium β-electron energy spectrum when adding a sterile neutrino mass eigenstate m_4. The additional eigenstate mixes with the active neutrinos' mass eigenstates $m_{1,2,3}$ under an amplitude $|U_{e4}|^2 = \sin^2(\theta)$. This adds a decay branch (blue) with a kinematic endpoint at $E_4 = E_0 - m_4$. The KATRIN experiment is possibly sensitive to $m_4 \in (0, 18.6)\,\mathrm{keV}$. The "kink"-like shape distortion the sterile neutrino eigenstate introduces to the spectrum changes depending on the mass and mixing amplitude of the sterile neutrino mass eigenstate. The magnitude of $\sin^2(\theta)$ is exaggerated for illustrative purposes and m_4 is set to an exemplary value of 10 keV.

3.2 KATRIN Experiment

The goal of the KATRIN experiment before the TRISTAN detector upgrade is to search for an effective electron antineutrino mass by measuring the kinematic endpoint of the Tritium β-spectrum. Parts of the experimental setup discussed in the following will change for the TRISTAN detector upgrade. These changes are highlighted in section 3.3. The KATRIN experiment is composed of five main parts. Electrons are produced in the Windowless Gaseous Tritium Source (WGTS) via Tritium β-decay, detailed in section 3.2.1. The electron emission angle is isotropic, thus half of the β-electrons travel downstream and half upstream, guided by magnetic fields. The upstream electrons meet the Rear Wall. (RW), where the beamline terminates, which is outlined in section 3.2.2. The downstream electrons pass through the transport and pumping section explained in section 3.2.3, where the remainder of the gaseous tritium is pumped away. They then enter the Main Spectrometer. (MS),

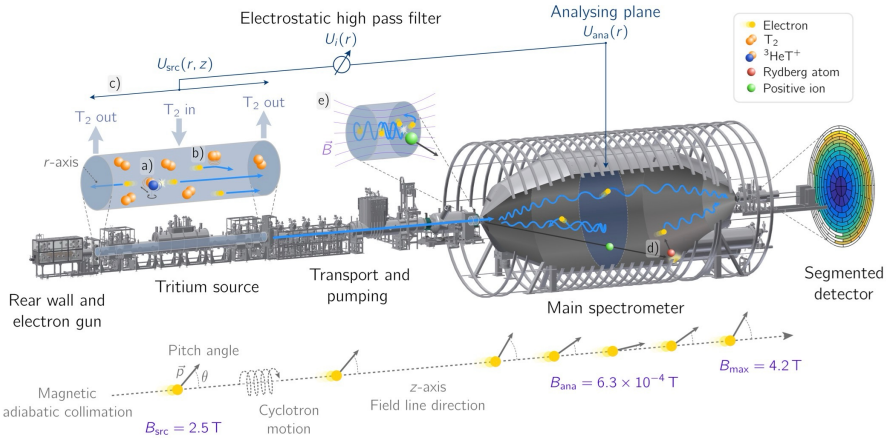


Figure 3.3: Illustration of the KATRIN beamline and its working principle. In the upper part the different components of the experiment are shown. In the lower part the electron momentum with respect to magnetic field line is given. The blue lines illustrate possible paths of the electrons inside the experiment. Figure taken from [72].

where all electrons below a certain threshold are filtered, covered in section 3.2.4. Finally, the remaining electrons arrive at the detector, which will be detailed in section 3.2.5. The entire beamline and working principle is illustrated in figure 3.3.

3.2.1 Source Section

The main component of the source section is the WGTS. The WGTS is a 10 m long tube with a diameter of 9.0 cm filled with high-purity molecular tritium gas. The reference column density is 5.0×10^{17} molecules/cm^2, resulting in an activity of 1.7×10^{11} Bq. The tritium decays and β-electrons are adiabatically guided towards the RW section and the transport section using a homogeneous magnetic field of up to 3.6 T, generated by superconducting solenoids. The tritium can not be contained in the source using a solid window since the interaction of the electrons with the material would lead to an energy distortion larger than the imprint of the neutrino mass. As a consequence, the tritium molecules can freely diffuse from the middle of the tube, where they are injected. They travel towards both ends, where they are pumped out. This ensures an ultra-high vacuum at the spectrometer. The resulting density profile can be seen in figure 3.4. The constant gas injection keeps the tritium purity stable at $> 95\,\%$. The WGTS is enclosed by a cryostat, cooling the tube and gas down to a constant 30 K. This reduces Doppler broadening through the thermal motion

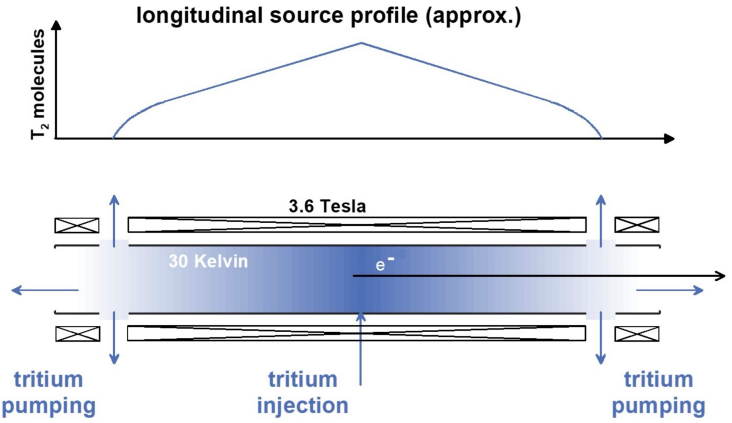


Figure 3.4: The setup of the WGTS, with an approximate longitudinal tritium density distribution (top). By injecting low temperature tritium at the center and pumping out the tritium that drifted towards the ends, the density profile is kept constant. Figure taken from [7].

of the gas molecules and therefore the tritium throughput necessary for maintaining the right tritium density at relatively low pressure [73]. Both the temperature and the column density are kept stable with fluctuations of less than 0.1 %/h [74].

3.2.2 Rear Wall Section

The creation of charged particles in β-decay and subsequent processes like scattering in the source leads to the development of a cold and strongly magnetized plasma with an electric plasma potential at the meV scale. This starting potential needs to be homogeneous such that it does not interfere with the electron energy measurement. The RW can be used to determine and manipulate the starting potential. The magnetic field lines from the upstream end of the WGTS are mapped onto the RW, leading to charged particles being transported to the RW. By suitably modifying the RW potential, taking the WGTS and RW work functions into consideration, homogeneous plasma conditions can be established. For the neutrino mass measurements, a gold-plated steel disk with a diameter of 14.5 cm held at a small bias voltage is used as a RW. Because the electrons in tritium β-decay are created with isotropically distributed momenta, only half of the electrons in the WGTS have momenta pointed towards the detector, with the other half being pointed towards the RW. This leads to many electrons reaching the RW and backscattering at the RW disc. These electrons lose energy in the backscattering process. They are subsequently filtered out by the MS when set to a high potential. Behind the RW disc, there is also the Calibration and Monitoring System (CMS), including an electron gun as

well as a mount for radioactive calibration sources. These can be used to measure the transmission function and perform other calibration measurements.

3.2.3 Transport Section

The transport section fulfills two purposes. On the one hand, it adiabatically transports the β-electrons from the WGTS downstream to the MS using magnetic fields of up to 4 T. On the other hand, it stops tritium gas from entering the spectrometer section by reducing the tritium flux to below $10^{-14}\,\mathrm{mbar}\cdot 1/\mathrm{s}$. This prevents tritium molecules from decaying in the spectrometer section and thus adding additional background.

Two pumping sections achieve such a low tritium flux: The Differential Pumping Section (DPS) and the Cryogenic Pumping Section (CPS). After leaving the WGTS, electrons and tritium molecules first travel through the DPS. It consists of five beam tubes and superconducting magnets that are tilted by 20° with respect to each other. This makes tritium molecules hit the walls more often, increasing the likelihood for them to be pumped away by the four turbomolecular pumps, whilst the light electrons are guided through the corners by the magnetic field. The DPS reduces the tritium flux by at least seven orders of magnitude.

The CPS lies further downstream, right after the DPS and removes the remaining tritium flux. The CPS follows the same chicane pattern as the DPS with six turns. As the turbomolecular pumps cannot further reduce the tritium flux, the CPS is cooled to 3 K with liquid helium to create a cold trap. The efficiency of this tritium trap is increased further by coating the inside with an argon frost layer to which the gas molecules adhere upon impact. After passing through the DPS and CPS, the β-electrons move on to the spectrometer section.

3.2.4 Spectrometer Section

The spectrometer section is divided into the Pre-Spectrometer (PS) and the MS, where the latter is used for β-scanning. The Main-Spectrometer is 23.38 m long and has a diameter of 9.8 m at its widest point. It is operated at an ultra-high vacuum of $10^{-11}\,\mathrm{mbar}$. Both spectrometers utilize the principle of MAC-E, which is illustrated in figure 3.5 [75, 76]. In a nutshell, the MAC-E filter only allows β-electrons above a certain kinetic energy set by the maximum retarding potential to pass, with the rest being reflected back. The electrons kinetic energy manifests in their momentum, which is distributed isotropically upon their creation in the WGTS, meaning only a fraction of its total momentum is parallel to the magnetic field lines. When employing a retarding potential, only the projection of the electron momentum that is parallel to the electric retardation field is affected. Thus, using only a retarding potential to block electrons below a certain energy would result in

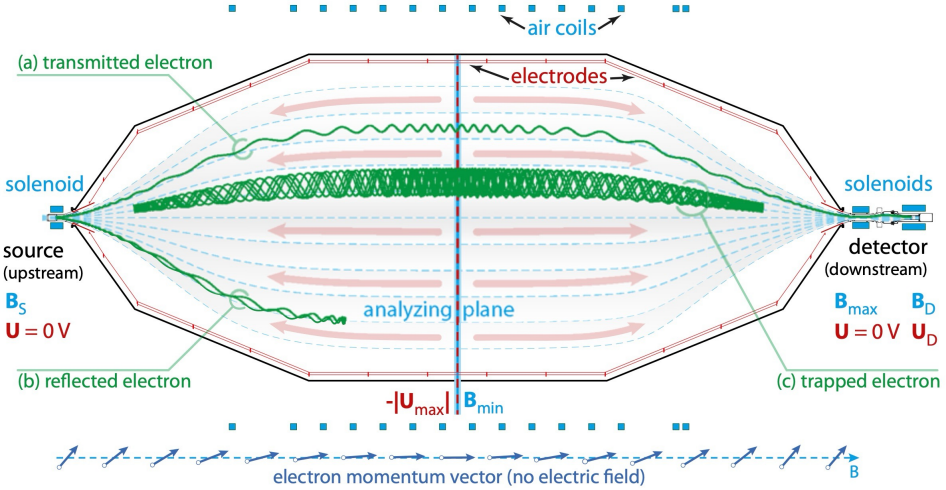


Figure 3.5: Illustration of the MAC-E filter principle utilized in the MS of KATRIN. The β-electrons are guided adiabatically along the magnetic field lines, performing cyclotron motion. The magnetic field is drastically reduced towards the analyzing plane. This leads to the alignment of the electron momentum with the magnetic field and thus converts most of their transversal momentum into parallel momentum, as visualized by the blue arrows. Here, the electric retarding potential is maximal, allowing only electrons with sufficient parallel momentum to pass (track a). Electrons with insufficient parallel momentum are reflected (track b), while electrons that are created inside the spectrometer volume may remain trapped by the magnetic mirror effects at both ends (track c). Figure taken from [7].

many more electrons being blocked than intended. This is because only the ones with high enough kinetic energy and momentum that is close to parallel to the field lines by chance will be able to pass. Here, the "MAC" part of the MAC-E filter comes into play. Using magnetic adiabatic collimation, the electron momenta can be aligned with the magnetic field lines without affecting the total energy of the electrons.

Electrons travelling along a magnetic field perform cyclotron motion, where the component of its momentum that is perpendicular ($p_\perp$) dictates the cyclotron radius. In an adiabatic motion in a magnetic field, the magnetic moment μ of the electron's cyclotron motion is conserved:

$$\mu = \frac{e|\vec{l}\,|}{2m_e} = \frac{p_\perp^2}{2m_e B} = \frac{E_\perp}{B} = \text{const.} \tag{3.4}$$

where $E_\perp$ is the transversal kinetic energy of the electron. It can be written in terms of the pinch angle $\theta = \angle(\vec{p}, \vec{B})$ between the electron momentum $\vec{p}$ and the magnetic

field $\vec{B}$ as $E_\perp = E \cdot \sin^2(\theta)$. If the magnetic field strength B is reduced, $p_\perp$ also has to decrease in order for μ to be conserved. Consequentially, the pinch angle is decreased and which leads to an increase in $E_\parallel = E \cdot \cos^2(\theta)$. Thus, to achieve that $E_\parallel \approx E$, the field strength at the entrance of the spectrometer should be as large as possible, while the field strength in the analyzing plane, where the electric retarding potential is maximal, should be as small as possible. It would make sense to use the smallest possible magnetic fields to maximize this effect. There is a caveat, however, because the magnetic flux

$$\Phi = \int \vec{B} \cdot \mathrm{d}\vec{A} \tag{3.5}$$

needs to be conserved. Consequently, in parts of the beamline where the magnetic field decreases, the area of the cross-section of the flux, also called flux tube, needs to increase. Therefore, the lower the minimal magnetic field strength is in the analyzing plane, the larger the diameter of the spectrometer has to be.

The energy resolution of the MAC-E filter is therefore determined by the remaining perpendicular component of the momentum $p_\perp$, or equivalently $E_\perp$, which depends on the minimum and maximum magnetic field strengths that are applied:

$$\Delta E = E \cdot \frac{B_{\mathrm{min}}}{B_{\mathrm{max}}}. \tag{3.6}$$

Using the KATRIN design field values of the MS, $B_{\mathrm{max}} = 6\,\mathrm{T}$ at the pinch magnet located directly at the entrance to the MS, and $B_{\mathrm{min}} = B_{\mathrm{ana}} = 3 \times 10^{-4}\,\mathrm{T}$ located at the middle of the spectrometer in the analyzing plane, results in an energy resolution of 2.7 eV at the tritium kinematic endpoint [7]. Magnetic flux conservation requires that the MS has a radius of $\sim 5\,\mathrm{m}$. The complete magnetic and electric field setup along the beamline can be seen in figure 3.6.

The magnetic adiabatic collimation effect is also relevant when considering the propagation of an electron from the source to the pinch magnet. As the source magnetic field strength B_{src} is typically smaller than the pinch magnetic field B_{max}, the electron momentum perpendicular to the magnetic field lines $p_\perp$ will increase (and analogously $p_\parallel$ will decrease) along its way to the pinch magnet. If now the electrons initial pinch angle at the source is large enough, it is possible for the pinch angle to grow beyond 90° during its propagation to the pinch magnet and the electron to be reflected. This is called the magnetic mirror effect. Thus, there is a maximum initial angle θ_{max} that an electron can have and still be transmitted towards the detector:

$$\theta_{\mathrm{max}} = \arcsin\left(\sqrt{\frac{B_{\mathrm{src}}}{B_{\mathrm{max}}}}\right). \tag{3.7}$$

In the current KATRIN field setting, $\theta_{\mathrm{max}} = 50.4°$ [73]. The PS was used to lower the flux of electrons arriving at the MS by pre-filtering at a retarding potential

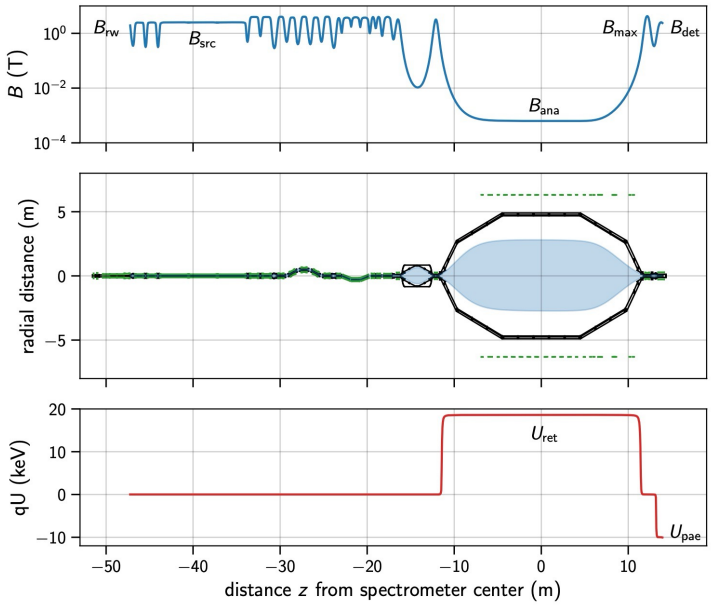


Figure 3.6: The electromagnetic field configuration of KATRIN. The upper plot shows the magnetic field strength, the bottom plot shows the electric potential. The middle plot shows a cross-section of the beamline geometry, viewed from the top. Here, beam tube elements and electrodes are shown in black, magnets in green and the mapping of the flux tube to the detector in blue. Figure taken from [15].

of 10 kV, but its operation has led to unwanted magnetic trapping effects between the spectrometers [7]. Electrons caught in these traps need to scatter and therefore loose energy to escape. Electrons that have an initial pinch angle lower than $\theta_{\max}$ and a kinetic energy above the electrostatic retarding potential are transmitted to the detector, which is detailed in the next section.

3.2.5 Detector Section

After passing through the analyzing plane, the β-electrons are guided towards the detector section and get an energy boost of 10 keV by a Post-acceleration Electrode (PAE), before they hit the FPD. The PAE reduces detector systematics like backscattering and shifts the spectrum to an energy region with lower background. The FPD consists of 148 silicon pin diodes, which are arranged in 12 rings counting 12 pixels each, plus an additional 4 pixels in the center, as visible in figure 3.3. The total rate limit for the FPD is approximately 10^6 cps, and the energy resolution is approximately 1.5 keV at full-width half maximum (FWHM) [7]. In the effective elec-

tron antineutrino neutrino mass measurement with KATRIN the FPD only counts electrons and is therefore sufficient.

3.2.6 Latest KATRIN Results

In the latest KATRIN neutrino mass publication, the results of the first two physics runs are included. Here, the spectral shape of the integral spectrum was obtained by measuring the count rate for 39 different MS retarding potential qU_{ret} settings in the interval $[E_0 - 300\,\mathrm{eV}, E_0 + 135\,\mathrm{eV}]$ around the tritium endpoint energy E_0. In the final 40 eV of the integral tritium spectrum, a total of 3.7×10^6 β-electrons were detected.

To obtain the squared effective antineutrino mass $m^2_{\bar{\nu}_\mathrm{e}}$ and kinematic endpoint energy E_0, a spectral fit is performed on the measured data with a model prediction. Only 28 of the original 39 points in the interval $[E_0 - 40\,\mathrm{eV}, E_0 + 135\,\mathrm{eV}]$ were used for the spectral fit, with the remaining points at lower energies serving as an indicator for the activity stability. The integral spectrum model is calculated by convolving the theoretical differential decay spectrum $\frac{\mathrm{d}\,\Gamma}{\mathrm{d}E}$ with the instrumental response function $R(E - qU_{\mathrm{ret}})$:

$$\dot{N}(qU_{\mathrm{ret}}; E_0, m^2_{\bar{\nu}_\mathrm{e}}) = A_\mathrm{s} \cdot N_{\mathrm{T,eff}} \int_{qU}^{E_0} \frac{\mathrm{d}\,\Gamma}{\mathrm{d}E} \cdot R(E - qU_{\mathrm{ret}})\mathrm{d}E + R_{\mathrm{bg}}, \tag{3.8}$$

with the signal amplitude A_s and background rate R_{bg} treated as nuisance parameters. The normalization factor $N_{\mathrm{T,eff}}$ describes the effective number of tritium atoms in the source. The response function represents the transmission probability of an electron from the source to the detector given a surplus energy $E - qU_{\mathrm{ret}}$, and depends on various systematic parameters. More detailed information on the modelling of the response function, can be found in [69].

The resulting best fit value or the squared effective antineutrino mass $m^2_{\bar{\nu}_\mathrm{e}} = -0.14^{+0.13}_{-0.15}\,\mathrm{eV}^2$ results in an upper limit for the mass of $m_{\bar{\nu}_\mathrm{e}} < 0.45\,\mathrm{eV}$ at 90 % CL [53]. This is the currently world leading limit on the effective neutrino mass, and can be seen in comparison to other experiments' results in figure 3.7. With a final dataset including 1000 days of live measurements, the KATRIN experiment can reach a sensitivity of 0.3 eV at 90 % CL [53]. After these measurements are concluded, the TRISTAN detector upgrade will be installed for the keV-scale sterile neutrino search.

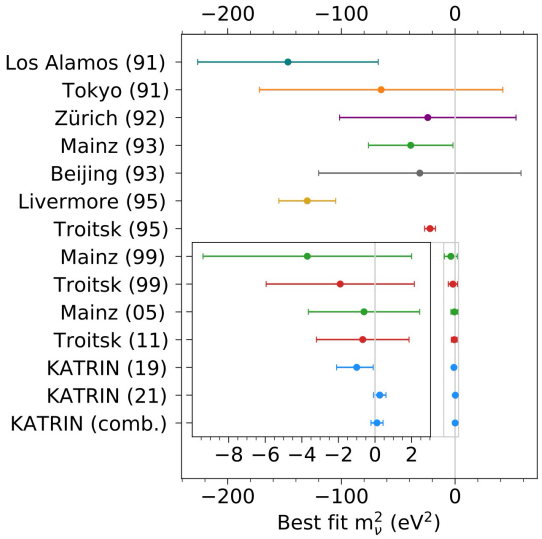


Figure 3.7: Comparison of the best-fit values from the first two KATRIN neutrino mass campaigns with previous neutrino mass experiments. Figure taken from [73].

3.3 keV-Scale Sterile Neutrino Search with the TRISTAN Detector Upgrade

The goal of the TRISTAN detector upgrade of the KATRIN collaboration is to search for the signature of a keV-scale sterile neutrino signature in the β-electron energy spectrum of tritium (see 3.1).

The KATRIN experiment is a good fit for the keV-sterile neutrino search, as proven by a commissioning run performed form 2018. Here, the tritium β-electron energy spectrum was measured from the kinematic endpoint down towards $E_0 - 1.6\,\mathrm{keV} = 17\,\mathrm{keV}$, thus being sensitive to sterile neutrinos with a mass of up to 1.6 keV. No sterile neutrino signal was found and an exclusion limit on the sterile-to-active mixing amplitude of $\sin^2(\theta) < 5 \cdot 10^{-4}$ (95 % CL) was determined [77]. This run was performed in integral measurement mode, with the MAC-E filter turned on and providing the energy resolution by lowering the retarding potential iteratively. The main problem when measuring deeper into the spectrum are the increasing count rates. Thus, for the FPD to be able to handle rates still relatively close to E_0, the column density had to be reduced accordingly.

To be sensitive to larger sterile neutrino masses, the measurement range has to be increased further into the tritium β-spectrum. Therefore, in the TRISTAN setting, the MAC-E filter is set to a much lower retarding potential. As a consequence, the rate of electrons arriving at the detector will be many orders of magnitude higher than

the FPD can handle, facilitating the need for an alternative detector. Additionally, because the MAC-E filter is not varied in a differential measurement, the such a detector has to provide an energy resolution itself, as will be further detailed in the following section.

The TRISTAN SDD combines both of these aspects, and was developed by the KATRIN collaboration for the keV-scale sterile neutrino search. Accompanying it are other changes to the KATRIN beamline and operation, which are explained in the following section.

The TRISTAN Configuration

The elementary change to the KATRIN setup is the addition of a new detector and DAQ system, which will mainly measure the differential and not the integral spectrum. That means the energy resolution of the β-electron energy spectrum is provided by the detector itself and not by varying the retarding potential of the MS and measuring the respective count rates. It was demonstrated that with an energy resolution of 300 eV at a β-electron energy of 20 keV [68] no sensitivity loss will be observed, in contrast to the FPD's aforementioned energy resolution of ≈ 1.5 keV at an electron energy of 18.6 keV and no significant improvement towards higher energies [78].

Additionally, the MS will operate on a very low retarding potential in the range of approximately 500 eV, with the exact value still up to debate, and thus allow measuring much deeper into the tritium β-decay spectrum, as illustrated in figure 3.2. This is needed to detect larger sterile neutrino masses m_4, because the sterile neutrino signature shifts further away form the kinematic endpoint for increasing m_4, as visible in figure 3.2. Looking at the spectrum, the differential decay rate increases by several orders of magnitude moving from higher to lower β-electron energies. The TRISTAN detector system is designed to handle count rates of up to $\mathcal{O}(10^8$ CPS). The detector is segmented into $\mathcal{O}(1000)$ pixels, which lowers the electron count rate each pixel needs to handle to $\mathcal{O}(10^5$cps) [68]. To meet these requirements, the TRISTAN detector is designed using the SDD technology.

Silicon Drift Detectors

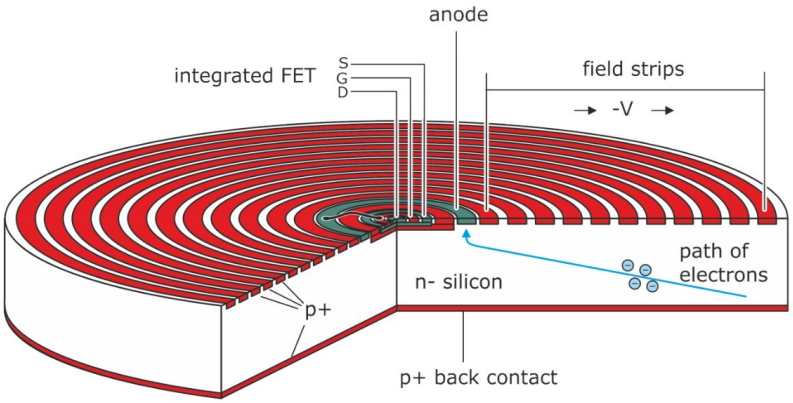
SDDs consist of the semiconductor material silicon, doped with electron donor atoms ("n-type"). An SDD is based on the principle of sideward depletion. It uses a slightely n-doped silicon wafer as a base material. To create the depletion zone the entrance window side is heavily p-doped. On the opposite side, where the electronic readout lies, a small heavily n-doped anode is placed in each pixel. A negative bias voltage is applied to create a large depletion zone, which serves as

the active detector volume. Energy depositions of incident ionizing particles, such as a β-electron, create electron-hole pairs in the depletion zone. The TRISTAN detector has the goal to measure β-electrons, which will therefore be focus in the following. The incident β-electrons enter from the solid p-doped back-contact and create electron-hole pairs in the depletion zone. The number of electron-hole pairs depends on the energy the incident electron deposits in the depletion zone and the material average energy needed to create a single electron-hole pair. These electrons have to be guided to the anode to create a measurable signal. This is achieved by p-doped junctions on the opposite side (the detector readout side) that are segmented into concentric drift rings with a small n-doped anode in the middle, which is connected to the read-out electronics. The drift rings are biased such that they generate an electric field with a strong component parallel to the surface, causing the electrons to drift towards the anode, whereas the holes are collected at the back contact or the drift rings. The working principle of an SDD is shown in figure 3.8.

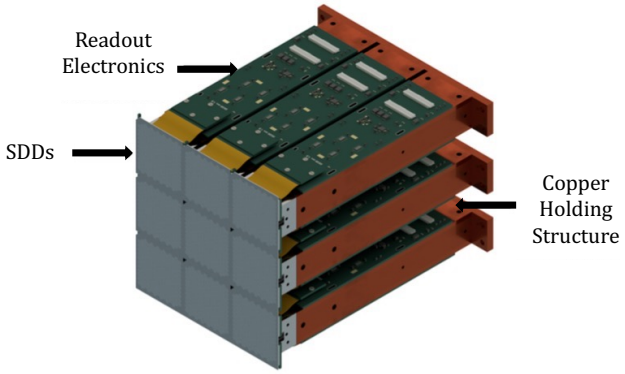
The small physical dimensions of the anode allow the detector to only have a small capacitance almost independent of the detector area, resulting in a better energy resolution and faster rise times compared to a conventional PIN-diode. This enables the construction of large pixels while maintaining an optimal energy resolution close to the Fano limit, while the fast rise times allow better event separation and thus enables the detector to handle higher rates. In addition, the electronic noise can be reduced by cooling down the detector. As shown in [68], an electronic noise of $9.9\,e_{\mathrm{rms}}$ can be achieved with temperatures as high as $-35\,^{\circ}\mathrm{C}$. This is sufficient for the keV-scale sterile neutrino search [79].

Multiple SDD pixels are grouped onto a detector chip, where the limiting factors are complexity and cost [80]. A single TRISTAN detector chip, also called detector module, consists of 166 pixels. The pixels have a hexagonal shape and cover a $4\,\mathrm{cm} \times 4\,\mathrm{cm}$ area. Central to the design of the detector modules is the non-planar geometry with a front-facing SDD chip, allowing for the side-by-side arrangement of modules. The detector for the first measurement phase will consist of 9 detector modules, as visible in figure 3.8. This offers a total of 1264 obstruction free pixels. From these, 936 are inner pixels that should receive an even illumination with electrons. To handle the high count rates the new detector is capable of, a new data acquisition (DAQ) system is currently being developed.

In addition to the new detector system and the MS operation at low retarding potentials, some other changes to the beamline will be employed.



(a) Schematic drawing of a cylindrical silicon drift detector with an exemplary electron drift path towards the anode shown in blue. The p-doped back contact and the drift rings that create the electric field are shown in red. The collecting anode and the integrated amplifying transistor in the center of the device are shown in green. Figure from [80].



(b) CAD drawing of the planned arrangement of nine TRISTAN detector modules, containing 166 pixels each. Figure from [81].

Figure 3.8: Schematic drawing and illustration of the actual implementation of an SDD in TRISTAN.

Further Changes

To lower the impact of systematic uncertainties on the TRISTAN sensitivity, the following additional changes are made to the KATRIN beamline [15, 82]:

- the magnetic field settings are being optimized to reduce systematics like detector and rear wall backscattering,
- the column density will be reduced to around 1 % of the nominal value to keep an input rate per pixel of 10^5 CPS,
- a new post acceleration electrode allowing acceleration voltages of up to 20 keV is being developed to reduce detector systematics like backscattering, pileup and dead layer effects,
- and finally, other materials and structures are being considered for the rear wall to reduce the contribution of the spectrum stemming from the RW to the normal tritium spectrum. Because the MS will be operated at low retarding potentials, many of the electrons backscattering at the rear wall will be able to reach the detector even after loosing some energy, and the backscattering probability at the current gold-plated rear wall is very high. Consequentially, the normal tritium spectrum and the spectrum emitting from the RW superimpose, reducing the sensitivity to the sterile neutrino signature.

An overview of the different systematic effects will be given in the next section.

3.4 Systematic Effects and their Modelling

When measuring the tritium spectrum using the TRISTAN setup, significant deviations from the raw theoretical predictions will be observed. This is due to experimental effects that distort the decay spectrum. To describe these effects, new parameters are introduced, which usually carry systematic uncertainties. Thus, on top of the statistical uncertainties each measured data point is subject to these systematic uncertainties. Unlike statistical uncertainties, which decrease with more data, systematic uncertainties do not necessarily decrease with longer measurement times.

By conducting a measurement deep into the tritium spectrum, even with the activity of the source being reduced by a factor of approximately 200, many more events will be detected than in the endpoint measurement, pushing the statistical sensitivity limit down very quickly [83]. The systematic uncertainties however are not mediated directly by larger rates. The latest investigation shows that the systematic uncertainties reduce the sensitivity of TRISTAN by at least one order of magnitude

if the same beamline settings as for the neutrino mass measurement are used [84]. A prediction of the statistical sensitivity, compared to the previous KATRIN result and other limits is shown in figure 3.9. Therefore, it is fundamental that the systematic

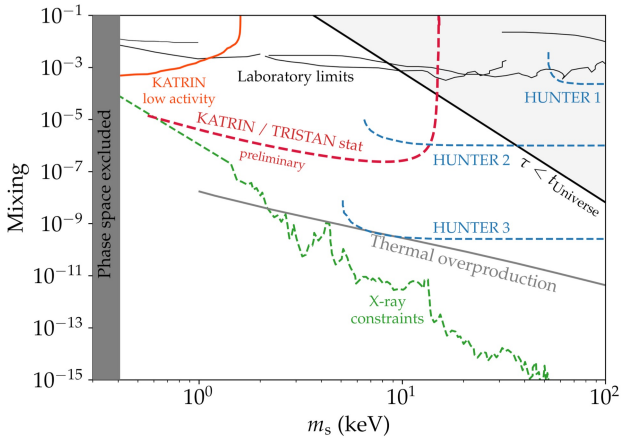


Figure 3.9: Statistical sensitivity prediction for the TRISTAN detector upgrade, compared to previous laboratory and cosmological limits and the predicted sensitivity of the HUNTER project [85]. Plot adapted from [84].

effects and their uncertainties themselves, as well as the effect they have on the differential decay spectrum and TRISTANs final sensitivity limit are understood.

In the following, several of the systematics will be detailed:

Tritium Decay

The decay energy of β-electrons follows the predictions of Fermi's theory of weak decay, with a few corrections.

- Atomic corrections: charge shielding effects and nuclear recoil, both subject to theoretical uncertainties, must be considered in the calculation of the differential energy spectrum.
- Final state distribution: the tritium decay daughter molecules may be excited to rovibrational final states with a certain probability, which is energy dependent.

Rear Wall

Due to the isotropic distribution of the β-electron momenta, half of the electrons created in the source head towards the rear wall instead of the detector. Electrons

coming back from the rear wall are also more likely to scatter on their way to the detector.

- Backscattering: the electrons hitting the rear wall can scatter back while loosing a part of their energy and changing their pitch angle. In the neutrino mass measurement where the MS is operated at a high retarding potential qU_{ret}, electrons with an initial energy close to 18.6 keV are typically reflected. For the keV-scale sterile neutrino search KATRIN is operated at much lower qU_{ret}. Thus, depending on the electrons pinch angle and remaining energy after backscattering, it can still overcome the MS and reach the detector and create a misleading signal falsifying a possible sterile signature. The backscattering probability is given by the backscattering coefficient. A gold-plated rear wall has a backscattering coefficient of $\approx 50\,\%$, leading to a very sizeable imprint on the spectrum and this being one of the dominant sytematics [15]. To reduce this systematic, the possibility of a Beryllium RW (backscattering coefficient of 3.4 %) and modifications on the magnetic configuration are currently being investigated [15].
- tritium decay on RW: as not every single tritium molecule can be pumped away, there is a non-negligible accumulation of tritium on the rear wall. These molecules are probable to be bound in hydrocarbons [86], which affects their final state distribution. Cleaning the rear wall will mitigate this effect, but can not be performed in short time intervals.

Source

The main systematic effects in WGTS stem from β-electrons scattering on other particles or molecules.

- T_2 scattering: the β-electrons starting location in the WGTS mainly follows the tritium gas density distribution (see figure 3.4). Thus, they still have to travel through a certain portion of tritium gas before most of it is pumped away, with multiple scattering processes possibly taking place. The electrons can partake in elastic scattering, molecular excitation and electron impact ionization, all making them loose energy and altering their pitch angle and thus falsifying the resulting spectrum.
- Magnetic traps: dips in the magnetic field can trap electrons with sufficiently large pitch angles, leading to an increased number of scatterings and thus distorting the energy spectrum [15, 82].

Transport and Spectrometer

During their propagation through the beamline, electrons move through varying magnetic field strengths, undergoing cyclotron motion as described in section 3.2.4.

- Magnetic Mirrors: the magnetic mirror effect can reflect or trap electrons depending on their pitch angle, leading to a higher scattering probability for the trapped electron [15].
- Synchrotron energy loss: electrons moving in magnetic fields emit energy in form of synchrotron radiation. For electrons only travelling once through the beamline, this energy loss is negligible compared to the energy resolution of the detector [82]. However, many electrons undergo scatterings and reflections, making them pass multiple times through the beamline.
- Transmission function: the transmission function describes the probability of an electron with certain energy and pitch angle being transmitted through the MS, given a certain retarding potential. In the differential mode, this can be very well approximated by a step function with a cut at the retarding potential energy, as the 2.7 eV width of the transmission edge [7] is negligible compared to the energy resolution of the detector of 300 eV [80]. Thus, small instabilities in the retarding potential are negligible for a differential measurement.
- Non-adiabaticity: the high kinetic energies of some electrons allows them to move through the MS without exactly following the magnetic field lines, leading to a possible unwanted loss of transmission [87].

Detector and Readout

Further distortions to the spectrum are introduced by the detector and signal readout systematics, which need to be described in detail for a differential measurement. For a more detailed account of these systematics, the following works should be considered [81, 88–90].

- Backscattering and back-reflection: electrons can backscatter at the detector after only depositing a fraction of their total energy. This effect is illustrated in figure 3.10. Depending on their pitch angle, the backscattered electrons can be reflected back onto the same or even a different pixel by the magnetic mirror effect, possibly being counted as a separate event.
- Dead Layer: electrons can deposit part of their energy in a non-sensitive volume at the entrance window of the detector, called dead layer. The effect of a dead layer is illustrated in figures 3.10 and 5.6.

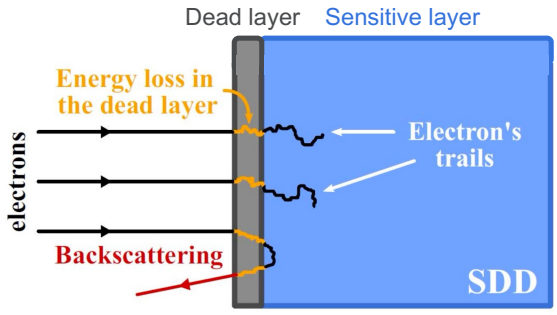


Figure 3.10: Illustration of the interaction of β-electrons with the detector volume. The fully sensitive detector volume is shown in blue, whereas the completely insensitive volume called dead layer is shown in grey. The energy loss in the dead layer region is illustrated by the yellow part of the electron trail. Figure adapted from [91].

- Charge sharing: an incident electron can deposit its energy in multiple adjacent pixels, possibly being counted as separate events. The impact of charge sharing on the spectrum is shown in figure 5.10.
- Fano noise: the creation of electron-hole pairs in the depletion zone of the SDD is a statistical process and thus subject to random fluctuations between events. This is called Fano noise [92], and leads to an intrinsic limit to the energy resolution of SDDs.
- Pileup: Two or more consecutive incident electrons on a short timescale are registered as one event with the sum of their respective energies being assigned to it. This is the main reason the count rate per pixel is limited to 10^5 CPS.

The full detector response for a monoenergetic electron beam is illustrated in figure 3.11.

Modelling with TRModel

The TRModel was developed to make a prediction of the expected measured spectrum by holistically combining all relevant systematic effects. Its primary goal is to provide a sensitivity estimation and to perform hardware optimization tests for the TRISTAN project [15, 93]. It models most of the significant experimental effects mentioned above. This is done by convolving the differential tritium β-decay spectrum with experimental response functions, each modelling an experimental effect. The response function gives the probability of measuring a β-electron energy $\hat{E}$, when the actual energy was E (when integrating over the different electron pitch

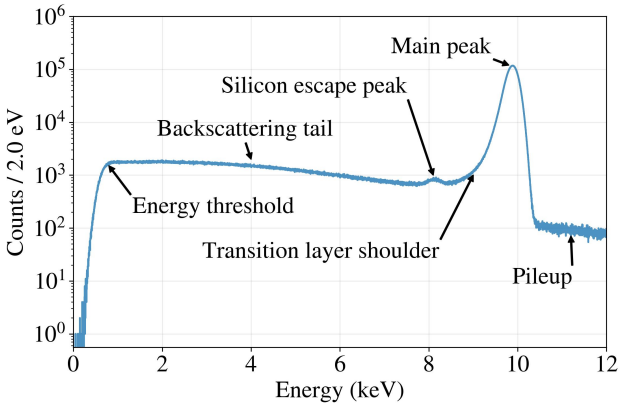


Figure 3.11: Response of the TRISTAN Detector to a monoenergetic electron beam. The primary peak of the 10 keV electron beam is slightly shifted to lower energies due to incomplete charge collection near the entrance window, which also causes a low-energy transition layer shoulder. Photon emission following electron ionization can result in a silicon escape peak, while backscattered electrons exiting the detector produce a backscattering tail. The detection threshold marks the point where noise and physical events become indistinguishable, and pileup effects distort the spectrum, contributing to counts above the actual electron energy. Figure adapted from [89].

angles). In an ideal experiment, the response function would be a delta function with the measured energy equaling the actual energy.

In the actual experiment, the tritium β-decay spectrum is not a continuous distribution, but a histogram. Therefore, the discretized version of the response function, a response matrix, has to be used. The energy histogram measured by the experiment $\vec{S}_{\text{exp}}$ can then be calculated using matrix multiplication of the response matrix R with the binned original differential decay spectrum $\vec{S}_{\text{theo}}$ [15]:

$$\vec{S}_{\text{exp}}(\hat{E}) = R(\hat{E}, E; \vec{p}) \times \vec{S}_{\text{theo}}(E; m_4, \sin^2(\theta)). \tag{3.9}$$

The response matrix depends on a set of nuisance parameters $\vec{p}$ and is an $m \times n$ matrix. The binned theoretical spectrum $\vec{S}_{\text{theo}}$ is an n-dimensional vector containing the number of events in each bin, and also depends on the sterile signature, which is dictated by the sterile neutrino mass m_4 and active-to-sterile mixing amplitude $\sin^2(\theta)$. The binned measured spectrum is an m-dimensional vector corresponding to the measured histogram, where the bin width and number of bins m do not have to match those of the binned theoretical spectrum.

The complete response matrix describing the entire experiment can be factorized

into response matrices describing the N individual experimental effects [15]:

$$R(\hat{E}, E; \vec{p}) = R_N(\hat{E}_N, \hat{E}_{N-1}; \vec{p}_N) \times R_{N-1}(\hat{E}_{N-1}, \hat{E}_{N-2}; \vec{p}_{N-1}) \times \cdots \times R_1(\hat{E}_1, E; \vec{p}_1). \tag{3.10}$$

The individual response matrices R_i may depend on their own set of nuisance parameters $\vec{p_i}$ which are all concatenated into $\vec{p}$. They are numerically calculated using analytical descriptions or Monte Carlo (MC) simulations, depending on the experimental effect they model. Magnetic adiabatic collimation and the magnetic mirror responses for example can be calculated analytically, but increasingly more complex effects like detector backscattering have to rely on simulations.

Challenges

With the TRISTAN detector upgrade, signal amplitude down to the parts per million scale ($\sin^2(\theta) \sim 10^{-6}$) can be probed. Nonetheless, systematic effects act on the spectrum at the percent level, making the impact of the systematic effects significantly larger than the sterile neutrino signature. The sterile neutrino signature is a very characteristic step-like distortion, but it is a shape-only effect, requiring very accurate modelling and a robust uncertainty and correlation estimation for the considerably larger systematic effects.

Many response matrices are created with MC simulations, which are a type of simulation that uses random sampling to numerically approximate the prediction of an observable [94]. Consequently, they are prone to statistical fluctuations and thus contain statistical uncertainty. To alleviate these, numerous random samples have to be used, a rule of thumb being to have at least 10 to 100 times the actual number of samples. This becomes a problem when using the model for a deep spectrum measurement like in TRISTAN, with an expected statistic of $\mathcal{O}(10^{15})$ β-electrons [83]. Here, the computing times for the simulations become infeasible, while the accuracy of the simulation is hard to evaluate and corrections, like from calibration measurements, are difficult to apply [15]. A rough time estimate from [15] states that it would take over 300 years to complete the simulation of the MC dataset with equal statistics as a one-year measurement.

Another challenge is posed by the bin widths of the response matrices. More coarse binning leads to less accurate responses, while finer binning increases the memory consumption of the matrix, as well as the computation time required for the convolution. This considerably limits how small the bin widths can be.

In the light of these limitations, performing a fit of the sterile neutrino parameters on measured data using the TRModel is currently not feasible. Thus, other analysis

channels have to be pursued, that do not directly depend on a parametric fit with TRModel. One Idea is to use the ability of NN to discriminate complex patterns in data to detect the sterile neutrino signature. In this work, the search for the sterile neutrino signature using NN will be outlined, and the first sensitivity studies will be discussed.

Chapter 4

Neural Networks

The challenges of using TRModel outlined in section (3.4) make testing other analysis channels worthwhile. Neural Networks (NNs), a type of Machine Learning (ML) model, are a candidate for a direct search of the spectral distortion a sterile neutrino would add to the kinetic energy spectrum of tritium β-electrons in TRISTAN data. They have been used widely in particle physics due to their ability to discriminate complex patterns in data [8, 9]. In this thesis, multiple NN are trained on simulated data modified with different systematic effects. Their goal is to classify the samples into those that contain a sterile neutrino signature and those that do not. This is called Binary Classification (BC) and is different from directly inferring the sterile neutrino mass and active-to-sterile mixing angle of the sterile neutrino signature that a sample contains. The NNs are therefore agnostic on the specific values of these sterile parameters. However, their sensitivity to a sterile neutrino signature as a whole can be judged using the method outlined in section 4.3. Thus, training with BC as a target serves as a good first step when assessing the usefulness of NN for detecting a sterile neutrino signature in TRISTAN. In addition, the robustness of the NNs when dealing with systematic uncertainties and model inaccuracies is tested in chapter 5.

This chapter begins by introducing the fundamental principles behind doing ML with NNs and seeks to demystify the topic by providing a clear mathematical foundation rooted in statistical principles (section 4.1). It then moves on to explain how BC works and what the learning algorithm of the NN looks like (section 4.2). Next, it introduces the different NN architectures that are compared against each other in this thesis (section 4.2.1). Afterwards different techniques to optimize the training of the NNs are discussed (section 4.2.2). The chapter then continues with the data the models are trained on, as well as the relevant preprocessing steps (section 4.2.3). Finally, the analysis strategy is outlined, where the traditional approach is contrasted with the NN approach used in this thesis (4.3).

4.1 Fundamental Principles

To understand why NN work and how they could be a strong candidate for detecting a sterile neutrino signature in TRISTAN data, it is necessary to introduce the mathematical backbone of NN. In the usual statistical setting, a closed-form mathematical model is used to describe the data. Here, real-world data x is interpreted as a realization of an underlying concept z. An example would be the concept of a non-zero neutrino mass, which is manifested in the measured integral tritium β-decay spectrum with a shifted kinematic endpoint. The connection between the concept and its realization is captured in the conditional probability distribution $p(x|z)$.

In this framework, the observed data for a specific latent concept (x, z) emerges from both the distribution of concept values, $p(z)$, and their specific realizations, $p\,(x|z)$:

$$(x, z) \sim p\,(x, z) = p\,(x|z)\,p(z).$$

Bayes' theorem then provides a foundation for inferring the latent concept z given observed data x, allowing us to compute the target distribution $p(z|x)$ (also called *posterior*):

$$p\,(z|x) = \frac{p\,(x|z) \cdot p(z)}{p(x)}. \tag{4.1}$$

Various statistical techniques, such as Markov Chain Monte Carlo or parametric fitting, can aid in this inference. However, in many cases, a robust mathematical model is either incomplete or unavailable, leaving only observed data samples x, presumed to originate from the joint distribution $x \sim p(x, z)$, accessible for analysis. An example of this would be the search for a sterile neutrino signature in TRISTAN data, where the modelling of $p(\text{data}\,|\,m_4, \sin^2(\theta))$ is difficult or computationally not feasible. The approach now uses a setup that requires only sample inputs. This is made possible by framing the statistical inference task, which is to approximate $p(z|x)$, as a learning problem. Learning here is defined as a search through a predefined space of different hypotheses, which could contain a good approximation of the solution. The specific form of this hypothesis space is determined by the NN architecture. The search is guided by a performance function that depends on the data.

4.2 Binary Classification with Neural Networks

To produce a useful NN, a model architecture has to be chosen, as well as a performance objective that leads to the desired outcome and a learning algorithm that does the footwork of iteratively updating the network. In the context of TRISTAN, the goal is to approximate the posterior distribution $p(m_\mathrm{s}, \sin^2(\theta)\,|\,\text{binned β-spectrum})$ by a NN (denoted by $q_{\vec{\theta}}(z|x)$, the notation will be explained shortly). In this thesis, a simpler first step is discussed. Here, the goal is only to discuss the general

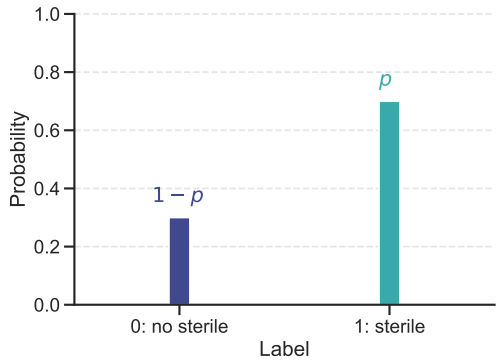


Figure 4.1: The Bernoulli Distribution $\mathcal{B}(z|p)$ used for BC. A NN learns to assign a probability p for a given data sample to have the label 1. In the case of this thesis, the label 0 belongs to spectra that do not contain a sterile neutrino signature, while the label 1 belongs to spectra that do contain one.

sensitivity of NN to a sterile neutrino signature. This is achieved using Binary Classification (BC). BC means that β-decay spectra (samples x) are given a label z, where

$$z = \begin{cases} 1, & \text{if the sample } x \text{ contains a sterile neutrino signature} \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

The goal of the NN is to assign the correct label to its input.

Search Space

The posterior distribution $p(z|x)$ is approximated by a parametrized variational family of densities. In simpler words, the approximation is chosen from a set of probability distributions (the variational family) defined by adjustable parameters (the NN). As mentioned above the task at hand is BC, because the NN is trained to distinguish two classes of input data. Thus, for BC purposes the Bernoulli Distribution $\mathcal{B}(z|p)$ is used. Here, z denotes the class label, where $z \in \{0, 1\}$, and p the probability of obtaining the outcome with label 1. It is depicted in figure 4.1

The probability p is then taken to be a function of the data x depending on learnable parameters θ, resulting in $p = p_\theta(x)$. This is the NN. It maps the N dimensional input space $\mathbb{R}^N$ to $[0, 1] \in \mathbb{R}$. In general, it can be seen as a composition of different functions, how many and which kind of functions is dictated by the model architecture.

Performance Objective

The performance objective is to approximate the posterior by the density parametrized using the NN. This is equivalent to minimizing the distance between the posterior and the density parametrized by the NN. A distance measure commonly used in ML is the forward Kullback-Leibler Divergence (KL Divergence) [95]:

$$D_{\mathrm{KL}}(p||q) = \int \mathrm{d}x\, p(x) \log \frac{p(x)}{q(x)}. \tag{4.3}$$

It measures the discrepancy of the two distributions p and q when looking through the lens of one of them. In practice, defining $q_\theta(z|x) \coloneqq \mathcal{B}(z|p_\theta(x))$, the performance functional, also called Loss functional $L(\theta)$, is given by

$$\begin{aligned} L(\theta) &= D_{\mathrm{KL}}(p(z|x)||q_\theta(z|x)) \\ &= -\int \mathrm{d}x\, p(z|x) \log q_\theta(z|x). \end{aligned} \tag{4.4}$$

In the second line of equation 4.4, it was used that the $p(z|x) \log p(z|x)$ term of the KL Divergence does not matter for the minimization as it is simply a constant. Unfortunately, the integral over D_{KL} is not computable. To solve this, one can try to approximate it empirically by sampling $z \sim p(z|x)$ and summing over the resulting terms. However, $p(z|x)$ is precisely what we try to approximate and do not know, so sampling from this distribution is not possible. To fix this, instead of finding a good approximation of the posterior $q_\theta(z|x)$ for a single datum x, the problem is solved for all data at once by calculating the expected value of $L(\theta)$ with respect to x (i.e. taking the integral over $p(x)$ with respect to x). Doing this avoids having to sample $z \sim p(z|x)$ to approximate D_{KL}, and instead sample from $(x, z) \sim p(x, z)$. This is called amortized variational inference and results in the loss function

$$\begin{aligned} \mathbb{E}_x L_\theta(x, z) &= -\int p(x, z) \log q_\theta(z|x) \\ &\approx -\frac{1}{N} \sum_{(x,z)\sim p(x,z)} \log q_\theta(z|x) = \hat{L}. \end{aligned} \tag{4.5}$$

Here, the integral over $p(x, z)$ is approximated by the sum over labeled samples $(x, z) \sim p(x, z)$. The resulting $\hat{L}$ is called (empirical) Cross Entropy Loss and provides a simple performance measure for training inference tasks on labeled data (x, z).

Learning Algorithm

To learn the optimal parameters $\vec{\theta}$ that minimize $\hat{L}$, Stochastic Gradient Descent (SGD) is used. The gradient of $\hat{L}$ at the current position with respect to the learnable

parameters $\vec{\theta}$ is calculated and used in an update step to iteratively approach a minimum of $\hat{L}$. The update step is given by

$$\vec{\theta}_{n+1} = \vec{\theta}_n - \lambda \cdot \nabla_\theta \hat{L} \tag{4.6}$$

In this work, a high-performance variation of SGD called Adam is used [96]. As apparent from equation 4.6, to compute an update step the gradient of the scalar loss function $\hat{L}$ with respect to the networks parameters $\vec{\theta}$ is needed. To compute this efficiently, a technique called backpropagation is used. For an overview on the history of backpropagation and a visual explanation of how it works, refer to [97, 98].

4.2.1 Model Architectures

The model architecture gives the hypothesis space its form. It defines the complexity and number of parameters of the model. All the following architectures will be compared against each other in different scenarios regarding their sensitivity to a sterile neutrino signature in chapter 5. The goal is to infer which architecture is the most sensitive to a sterile neutrino and how complexity affects performance. The architectures and their explanations are loosely based on [99].

Multilayer Perceptron

This architecture serves as a benchmark as it is the most straightforward. Nonetheless, the Multilayer Perceptron (MLP) already proves to be very powerful, as can be seen in chapter 5.

The MLP is composed of L layers: $f = f^{(L)} \circ \cdots \circ f^{(1)}$. The layers with $l < L$ are called *hidden layers*. The l-th layer defines a function that maps a d_{l-1}-dimensional input vector called *features* to a d_l-dimensional output $f^{(l)} : \mathbb{R}^{d_{l-1}} \to \mathbb{R}^{d_l}$. A layer takes on the form

$$\vec{f}^{(l)}(\vec{u}) = a^{(l)}\left(W^{(l)}\vec{u} + \vec{b}^{(l)}\right) \tag{4.7}$$

where the a denotes a non-linear function, also called *activation*, which is applied element-wise. The matrix $W^{(l)} \in \mathbb{R}^{d_l \times d_{l-1}}$ is called *weight matrix* and $\vec{b}^{(l)} \in \mathbb{R}^l$ is called *bias vector*. The parameters of the network are then compromised of the full collection of weight matrices and biases:

$$\vec{\theta} = \left(W^{(1)}, \ldots, W^{(L)}, \vec{b}^{(1)}, \ldots, \vec{b}^{(L)}\right). \tag{4.8}$$

Even with a single hidden layer ($L = 2$), an MLP is a universal function approximator [100].

The activation functions are the secret to the expressiveness of NNs. Simply stacking multiple linear layers without activations would effectively result in just

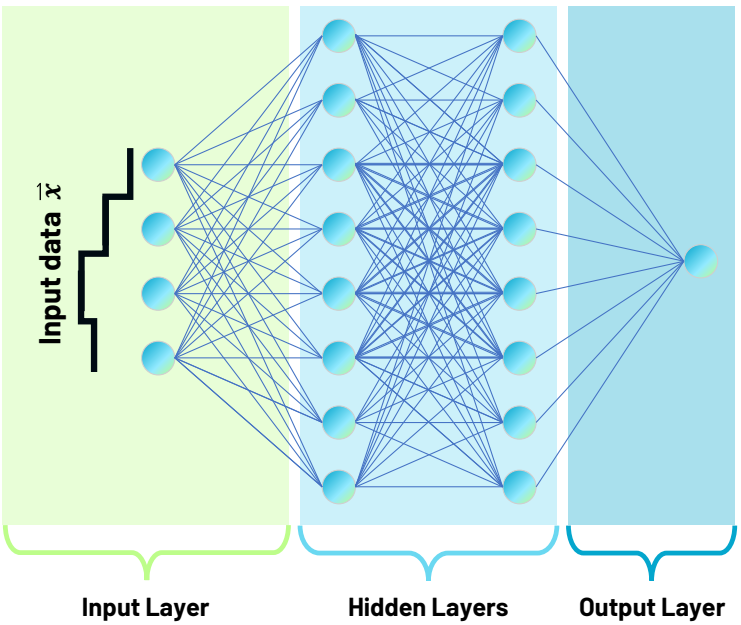


Figure 4.2: Illustration of a Multilayer Perceptron (MLP). The circles denote the individual neurons, with the blue lines denoting the weight that is connecting every pair. An array x, for example a binned β-spectrum, is fed to the input layer. The dimensionality of the input layer, here meaning the number of neurons it contains, has to match the dimension of the input array x. This input is then propagated through the hidden layers of the network to the output layer. In the case of binary classification the output is only one-dimensional, as illustrated with the single neuron above.

another linear layer, as linear combinations of linear combinations are still just linear. The popular choice of activation function for a deep NN is the Rectified Linear Unit (ReLU):

$$a_{\text{ReLU}}(u) = \begin{cases} x, & \text{if } u > 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.9}$$

In the case of binary classification, the final activation $a^{(L)}$ should be a sigmoid. This is to map the output to be in the interval $[0, 1]$, the standard case for BC. The full MLP model architecture is schematically shown in figure 4.2.

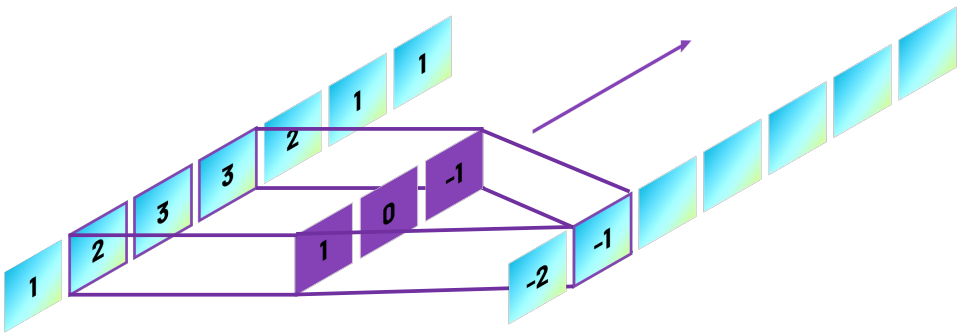


Figure 4.3: Illustration of a one dimensional convolution used in a CNN. The purple boxes depict a kernel of size three with the respective learned filter weights sliding over the input.

Convolutional Neural Network (CNN)

The MLP in general does not assume any spatial relationships between two entries in the input array. Spatial here refers to the space of kinetic energies of the β-electrons. To change this, a subclass of MLPs, the Convolutional Neural Networks (CNNs) is introduced [99]. They are most commonly used for image and time series classification [101]. A CNN consists of multiple filters (also called kernels) that are convolved with the input signal. The parameters of these kernels are learned by the network. The convolution involves striding over the input array and calculating the inner product at each step, as can be seen in figure 4.3. In this sense, a CNN can be understood as a fully connected MLP with shared weights. The convolution process is characterized by the *kernel size*, which dictates how many weights the NN can learn in each kernel, as well as the *stride*, and *padding*, which control at which positions in the array the inner product is calculated and whether to expand the input array by a fixed amount of entries when the kernel would extend beyond the last entry. The one dimensional convolution process is visualized in figure 4.3.

In practice, a CNN includes multiple filters, resulting in a multivariate representation of the signal with dimension equal to the number of filters used. The goal is to learn multiple discriminative features useful for the classification task. These discriminative features are learned by following up the convolutional layers with a classifier (e.g. an MLP), with a *global pooling layer* in between the two. The pooling layer aggregates the entire array of the preceding layer in each dimension into a single real value by either choosing the maximum or calculating the average value of the entries in the array. The entire architecture can be seen in figure 4.4. To summarize, the convolutional layers serve to construct a more descriptive representation of the input which can subsequently be classified as usual by an MLP.

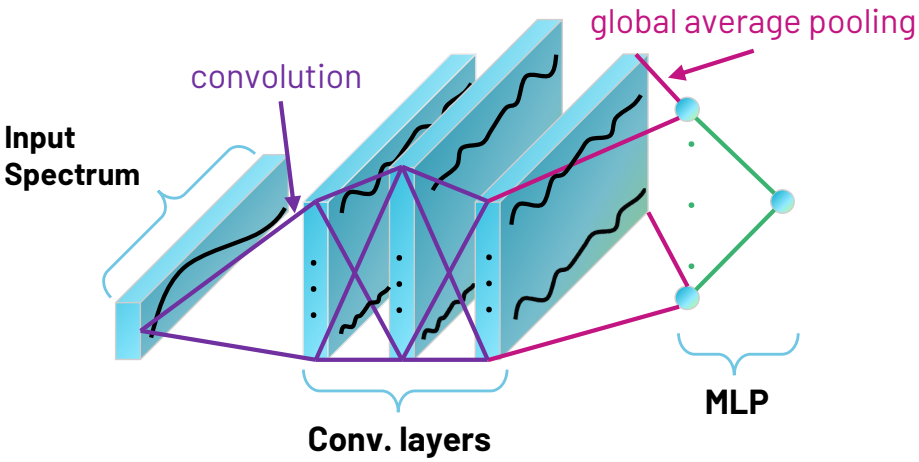


Figure 4.4: Illustration of a convolutional neural network (CNN) with three convolutional layers. After the final convolutional layer the signal is pooled in each channel dimension. For a classification task, the global pooling layer is followed by a fully connected MLP.

Residual Neural Networks

Finally, to see if deeper architectures (i.e. architectures that stack more layers on top of each other) are more performant, Residual Neural Networks (ResNets) are introduced as an architecture.

Deeper NN architectures can handle more complex tasks but suffer from vanishing gradients, as the gradient propagation becomes more and more complicated with increasing depth. To control the gradient flow through the network on an architectural level, residual or "skip" connections are utilized. The output of each block is added element-wise to the original signal, leading the single blocks to effectively learn an "update" on the input signal. These residual connections allow the gradient to flow directly to each block, simplifying the updates to their weights. Residual connections are most prominently applied to deep CNNs, but can also be used for other deep models like MLPs, and are realized as shown in figure 4.5.

4.2.2 Supplementary Training Techniques

The following techniques are employed to ensure a smooth training process and improve the NNs sensitivity to a sterile neutrino signature, as well as their generalization capabilities.

As mentioned above, a common problem deep NNs face is that of exploding and vanishing gradients. To prevent or alleviate this the input data x should be normal-

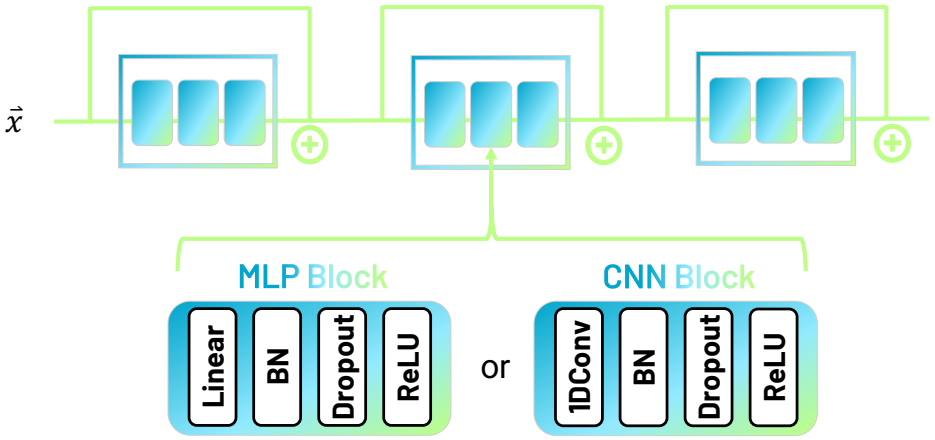


Figure 4.5: Sketch of a residual neural network ResNet with nine total layers. After each three-layer block the update computed by the block is added to the original input, utilizing a so-called "skip" or "residual" connection. Depicted here are two examples of blocks: a CNN and an MLP Block. The BN and Dropout boxes denote additional layers that improve the generalization capabilities of the NN as outlined in section 4.2.2.

ized to have a mean of zero and unit variance

$$x \to \frac{x - \mathbb{E}_x[x]}{\mathrm{Var}(x)}, \tag{4.10}$$

where $\mathbb{E}_x[x]$ is the mean input datum. This does not affect the data distribution but rather corresponds to a choice of units. Additionally, the weights w (and biases) of the network have to be initialized in a certain manner to control the scale of the gradients. In this thesis, this is achieved using Xavier normal initialization, in which the variance of the gradients is kept in check for both a forward and a backward pass through the NN by sampling w from a normal distribution with zero mean and standard deviation based on the number of input and output features ($N_{\text{in}}, N_{\text{out}}$):

$$w_{\text{NN}} \sim \mathcal{N}(\mu = 0, \sigma^2 = \frac{2}{N_{\text{in}} + N_{\text{out}}}). \tag{4.11}$$

Another recurring problem for neural networks is that they overfit the data they are trained on. This means that the model simply memorizes the input data instead of the latent concept. To diagnose this issue, the dataset is split into a training and a validation part. The validation data is not used to update the models weights during training, but to evaluate its performance on previously unseen data. If the models' performance differs largely between the two datasets, it is a strong indicator that

it is overfitting. To combat this, regularization techniques that force the model to learn the general latent concept instead of the data itself can be used. The following techniques are leveraged in this thesis:

- **Batch Normalization Layers** normalize the input of each layer by scaling and shifting it based on the mean and variance of a mini-batch of the training data. This prevents the mean and covariance of the data representations in hidden layers to evolve during training which could pose challenges for the downstream layers. Thus, adding batch normalization layers helps accelerate and regularize training. For more detailed information refer to [102].
- **Dropout Layers** randomly zero some elements of the input with a certain probability. They drop neurons only during training of the model and alleviate overfitting by preventing complex co-adaptations of neurons, forcing each neuron to learn to detect a generally helpful feature [103].
- **Weight Decay** adds the L2 norm (meaning the square root of the squared sum) of all the weights to the loss function, which will make each iteration of the NN try to minimize the model weights in addition to the loss. Doing this will keep the weights as small as possible preventing overfitting and exploding gradients.

Parametrizing the Model

It is possible to add additional inputs to the NN to give the model more context. An encoding of the values of the parameters of interest, the sterile neutrino mass m_4 and active-to-sterile mixing angle $\sin^2(\theta)$, can be appended to the binned differential β-spectrum and passed to the model. The values correspond to the actual sterile neutrino signature in the spectrum. The same is also done, however, to spectra that do not contain a sterile signature. This technique is called "parametrizing" the model:

$$\text{NN}(\text{binned spectrum}) \rightarrow \text{NN}_p(\text{binned spectrum}, m_4, \sin^2(\theta)) \tag{4.12}$$

Parametrizing a NN allows it to adapt to specific locations in the sterile parameter space, effectively training a distinct model for each location. This leads to an increase in the NNs sensitivity to the sterile neutrino signature, as illustrated in figure 5.1.

Hyperparameter Optimization

The hyperparameters of a ML model are configuration settings that are set before the training itself and are not learned directly from the data. They define the architecture and training dynamics of a model, influencing its performance and ability

to generalize. The architectural hyperparameters include for example the number of hidden layers, the number of neurons for each hidden layer, and the use of batch norm and dropout layers. Hyperparameters that affect the training dynamics of the model include the learning rate, which controls the step size of weight updates, the batch size, which defines the number of training samples processed before the model updates its weights, the training epochs, which specify how many times the entire dataset is passed through the model, and the weight decay parameter, which controls the weight of the L2 norm in the loss function.

A large part of the work for this thesis was invested into hyperparameter optimization. Initially, a grid search through the space of hyperparameters was used. The hyperparameters were split into architectural parameters and training dynamic parameters. First an initial guess for the architectural parameters was used to scan through the grid of the training dynamic parameters, in the order: learning rate $\rightarrow$ batch size $\rightarrow$ weight decay parameter. Then the training dynamic parameters were fixed, and the grid scan was done for the architectural parameters in the order: number of hidden layers $\rightarrow$ number of neurons $\rightarrow$ dropout probability. This cycle was repeated iteratively until the sensitivity to the sterile neutrino signature using the NN did not improve any further.

However, through further research and discussion, this method turned out to be very ineffective, with a grid scan proven to perform worse than even a random search [104]. Thus, Bayesian hyperparameter optimization with a hyperband strategy was chosen due to its robustness and efficiency compared to other hyperparameter optimization strategies [105]. Bayesian optimization uses probabilistic models, such as a Tree-structured Parzen Estimator, to intelligently search the hyperparameter space by balancing exploration and exploitation. Hyperband complements this by allocating computational resources dynamically, evaluating configurations with increasing fidelity (here training iterations) and discarding underperforming candidates early. In conclusion, the Bayesian optimization proposes promising configurations (based on a few initial random configurations), while hyperband efficiently manages resource allocation [105, 106].

4.2.3 Training and Evaluation Data

The backbone of each ML model is the data used to train it. To ensure optimal sensitivity of the NN to the sterile signature, the size and contents of the dataset were carefully chosen. The entire training dataset is split into multiple parts to cross-check the NN performance and generalization capabilities. Thus, the *training dataset* refers to the dataset used for training and the dataset used for validation of the NN combined. As an abbreviation, the *evaluation dataset* refers to the dataset

that is used for drawing the sensitivity contour as explained in section 4.3.1, and should not be confused with the validation dataset.

For the training process, an 80/20 split of training data and validation data was chosen. The full training dataset used in this thesis contains 1.2×10^6 samples x, each with a label z. The samples x consist of differential tritium β-decay spectra, divided into energy bins with a bin width of 100 eV and stored in arrays with N_i entries each. The binned β-spectra are generated from the theoretical prediction discussed in section 3.1. To save computational resources and time, the probability density function of the β-electrons from tritium β-decay is evaluated at the center of the energy bin instead of integrating over the whole bin width. However, bin-integration will be added in the future. This does not impact the predictive power of the NN. Half of the spectra in each dataset contain no sterile signature (i.e. $z = 0 \leftrightarrow m_\mathrm{s} = 0, \sin^2(\theta) = 0$), these are called *reference spectra*. The other half contains sterile neutrino signatures that vary from spectrum to spectrum, these spectra are called *sterile spectra*. The sterile parameters for the individual sterile spectra in the training dataset are drawn as follows:

$$\begin{aligned} m_\mathrm{s} &\sim \mathrm{Uniform}(0\,\mathrm{keV}, 18.6\,\mathrm{keV}) \\ \sin^2(\theta) &\sim \mathrm{Log}(10^{-1}, 10^{-3}). \end{aligned} \tag{4.13}$$

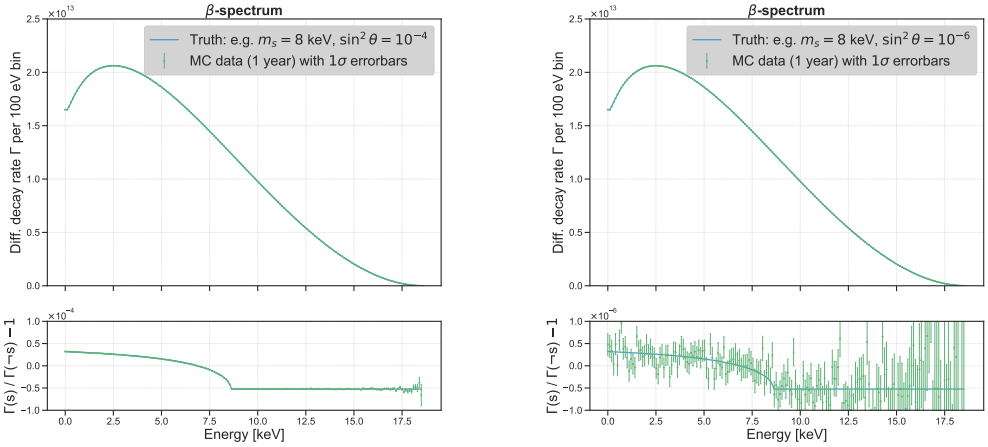
Multiple testings have shown that it is sufficient and even boosts performance to only introduce samples with active-to-sterile mixing angles down to $\sin^2(\theta) = 10^{-3}$. Each spectrum should emulate a histogram for a TRISTAN measurement period of one year, amounting to the collection of 2×10^{15} electrons. Therefore, bin-wise Poisson noise corresponding to an overall normalization of 2×10^{15} counts is added to each spectrum, as shown in figure 4.6.

The dataset used for training is processed multiple times by the NN during the training process. An *epoch* is one complete pass through the entire dataset, while a *batch* is a subset of data processed in a single iteration. The networks weights are updated after each batch. Batches enable efficient training by updating model weights multiple times per epoch, balancing computational cost and convergence stability.

Preprocessing

After the generation process, each sample is passed through a preprocessing chain, where different modifications can be applied, depending on which study the model is used for:

- The samples can optionally be convolved with (different) **response matrices**



(a) $\sin^2(\theta) = 10^{-4}$ **(b)** $\sin^2(\theta) = 10^{-6}$

Figure 4.6: Examples of a MC β-spectrum (shown in green) and the underlying theoretical truth (shown in blue) for different active-to-sterile mixing angles. Below, the relative difference between the truth with- and without a sterile neutrino signature is shown, in addition to the MC realization. The MC contains a total of 2×10^{15} electrons, corresponding to one year of data taking.

to incorporate systematic effects and study the effect of systematic uncertainties.

- A **shape factor** can be added to the samples to mimic a general shape distortion.
- Additionally, more general **perturbations** that mimic modelling inaccuracies can also be added to the samples.

A combination of the above is also possible. Finally, the array entries in each sample are shifted to have a mean of zero and a variance of one. This final preprocessing step is the same for every sample regardless of its former modifications. An example of the preprocessing chain can be seen in figure 4.7.

4.3 Analysis Strategy for Sensitivity Studies

The sterile neutrino sensitivity refers to the experiment's capability to exclude the existence of a sterile neutrino signature of a given magnitude in the measured spectrum. The goal of this thesis is to infer if a NN-based ML approach is sensitive to

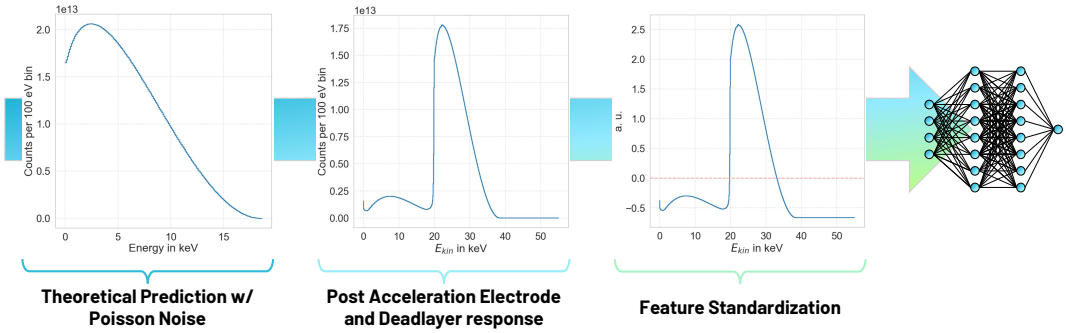


Figure 4.7: Exemplary preprocessing chain. The MC data sample (stemming from the theoretical prediction with bin-wise Poisson noise) is convolved with the PAE response and the detector dead layer response. Before passing the resulting sample to the NN, the mean of the data sample is shifted to zero and its variance is scaled to one. This step is often called feature standardization.

the sterile neutrino signature and robust under systematic uncertainties and unexpected distortions. In the future, sterile parameter regression using this approach is planned, but for now only binary classification is used. The latter is equal to a hypothesis test which is agnostic to the actual values of the sterile mixing angle and mass. In this regard the NN-based approach presented in this thesis is different from a parametric fit, which would also try to infer the parameters governing the sterile neutrino signature. The following two sections will explain how sensitivity contours with respect to the sterile parameters are drawn in the case of a parametric fit and the NN- and BC-based approach, as well as a quick explanation of the underlying package developed for this thesis.

Maximum Likelihood Analysis

The method of maximum likelihood is a technique that is used to estimate parameter values from a given set of n measurements of a random variable x, which is assumed to follow a known probability density function (p.d.f.) $f(x, \vec{\theta})$. The p.d.f. $f(x, \vec{\theta})$ can depend on multiple parameters $\vec{\theta} = (\theta_1, \ldots, \theta_m)$ of which at least one is unknown. The likelihood function

$$\mathcal{L}(\vec{\theta}) = \prod_{i=1}^{n} f(x_i, \vec{\theta}) \tag{4.14}$$

is maximized by the parameters $\vec{\theta}_{\text{fit}}$ that best describe the measured data [94]. Because working with a sum is more numerically stable than working with a product, and modern optimization algorithms are designed for minimization, the negative logarithm of $L(\vec{\theta})$ is used as a minimization target [107]. This scenario is commonly

used and referred to as "neg-log-likelihood minimization". If $f(x, \vec{\theta})$ is a normal distribution, the minimization of $-\log(\mathcal{L})$ is equivalent to a χ^2 minimization [94], because here

$$-2 \cdot \log(\mathcal{L}) = \chi^2 \tag{4.15}$$

For the KATRIN experiment after the TRISTAN detector upgrade, the random variables x_i are the number of counts N_i measured within a given energy range of the differential tritium β-electron energy spectrum $\frac{\mathrm{d}\,\Gamma}{\mathrm{d}E}$, called an energy bin i. The model prediction μ_i for an energy bin i is then given by the integral of $\frac{\mathrm{d}\,\Gamma}{\mathrm{d}E}$ over the corresponding energy range i, and depends on the sterile neutrino parameters of interest (m_4, $\sin^2(\theta)$):

$$\mu_i = A \cdot \int_i \frac{\mathrm{d}\,\Gamma}{\mathrm{d}E}(m_4, \sin^2(\theta))\,\mathrm{d}E \tag{4.16}$$

Here, A denotes an overall normalization factor which is handled as a nuisance parameter in the fit. The differential decay spectrum in theory also depends on the neutrino mass m_ν. The inference of m_ν is not the goal of the TRISTAN detector upgrade, and thus m_ν is set to zero for all the following analyses.

The probability to measure N_i counts for an energy bin i given a model prediction μ_i is then given by the Poisson distribution:

$$f(N_i, \mu_i) = \frac{\mu_i^{N_i} \exp(-\mu_i)}{N_i!} \tag{4.17}$$

As N is very large in each energy bin for a measurement time of one year with the TRISTAN detector upgrade ($N_i \geq \mathcal{O}(10^8)\,\forall i$), the Poisson distribution approaches a normal distribution in each energy bin [94]. This allows the use of χ^2 minimization as described in equation 4.15.

Sensitivity Gridscan

To evaluate the sensitivity of the KATRIN experiment with the TRISTAN detector upgrade to keV-scale sterile neutrinos, a grid scan is performed over the two-dimensional parameter space spanned by the active-to-sterile mixing amplitude $\sin^2(\theta) = |U_{\mathrm{e}4}|^2$ and the sterile neutrino mass m_4 of size $k \times l$. The index k corresponds to the logarithmically spaced values for $\sin^2(\theta)$ and the index l for the linearly space values for m_4. The goal is to use the $\Delta\chi^2$ test statistic to draw a sensitivity contour at a certain CL. The test statistic measures the compatibility between data and a model prediction, which depends on a hypothesis. An example for a hypothesis is the existence of a sterile neutrino with a specific mass m_4 and active-to-sterile mixing angle $\sin^2(\theta)$. The "no-sterile" hypothesis with $\sin^2(\theta) = 0$ is used as a null hypothesis.

At each grid point a χ^2 minimization is performed. Here, the sterile neutrino mass m_4 and mixing angle $\sin^2(\theta)$ are kept fixed at a value corresponding to the grid point, resulting in a minimal χ^2 value χ^2_{kl}. Additionally, at each grid point the χ^2-value for the null hypothesis $\chi^2_{kl}(\sin^2(\theta) = 0)$ is calculated. The best fit value out of all the χ^2_{kl} is then determined by $\chi^2_{\mathrm{BF}} = \min_{kl} \chi^2_{kl}$. The difference $\Delta\chi^2_{kl} = \chi^2_{kl} - \chi^2_{\mathrm{BF}}$ is used to draw the sensitivity contour when $\Delta\chi^2_{kl} = \Delta\chi^2_{\mathrm{crit}} = 5.99$ at 95 % CL for two degrees of freedom [94]. The validity of this approach relies on whether Wilks' theorem holds across the entire sterile parameter space, which is assumed to be valid for the purposes of this thesis. For the TRISTAN detector upgrade, the expected statistical error on the counts in each energy bin of the spectrum is small. Consequentially, the distribution of counts in each energy bin can be assumed to be normally distributed. In that case, the $\Delta\chi^2$ test is equal to the likelihood-ratio test, which is the most powerful hypothesis test according to the Neyman-Pearson Lemma [108, 109].

The χ^2_{kl} can be modified to include not only the statistical uncertainties of the bin contents but also systematic uncertainties and correlations between the energy bins. This is done using covariance matrices V [15]:

$$\chi^2_{kl}(\vec{p}) = \left(\vec{\Gamma}_{kl}(\vec{p}) - \vec{\Gamma}_{\mathrm{ref}}\right)^T V^{-1} \left(\vec{\Gamma}_{kl}(\vec{p}) - \vec{\Gamma}_{\mathrm{ref}}\right). \tag{4.18}$$

Here, the vector $\vec{\Gamma}_{kl}$ denotes the binned differential tritium spectrum with the detected rate in each energy bin for a sterile neutrino signature corresponding to the grid point kl, depending on a set of systematic parameters $\vec{p}$. The model evaluated under the null hypothesis ($\sin^2(\theta) = 0$) and a fixed set of systematic parameters is denoted by Γ_{ref}. The covariance matrix can be sub-divided into statistical and systematic contributions:

$$V = V_{\mathrm{stat}} + V_{\mathrm{syst}}. \tag{4.19}$$

The statistical uncertainties in each bin scale with the measurement time. The sensitivity contours constructed from a grid scan that result for different measurement times when only taking the statistical uncertainties for each bin into account are shown in figure 4.8.

4.3.1 Neural Network Method

The NN method also utilizes a grid scan. At each grid point kl a fixed number of spectra $N_{\mathrm{GP}} = 10^5$ that contain a sterile neutrino signature with a sterile neutrino mass m_4 and mixing angle $\sin^2(\theta)$ corresponding to the grid point are generated. They are each computed from the model prediction, with Poisson noise added to each bin. Additionally, each spectrum can be further modified with an individual

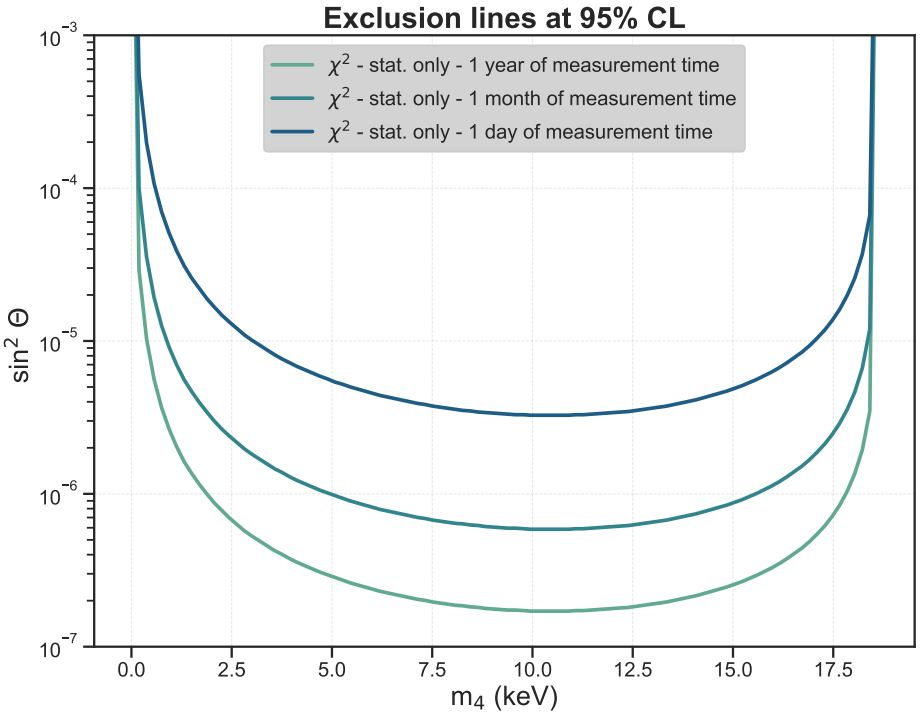


Figure 4.8: Statistical exclusion contours for different measurement times with the TRISTAN detector upgrade. The contours were constructed using a grid scan over the sterile parameter space. At each grid point, a $\Delta\chi^2$ test is performed. At the grid points where the value crosses the threshold value of $\Delta\chi^2_{\text{crit}} = 5.99$ for two degrees of freedom at 95 % CL, the contour is drawn. The shorter the measurement time, the fewer electrons are collected. Consequentially, the statistical sensitivity decreases.

response matrix to account for systematic uncertainties. For example, each of the spectra can contain an individual response for different deadlayer widths, mimicking the uncertainty in the actual value of the deadlayer. This is used to demonstrate the NNs generalization capabilities beyond systematic uncertainties. These spectra are passed through the model. The outputs of the model will be distributed because no two spectra are alike due to the inherent Poisson noise in each energy bin of each spectrum. The output is used to fill a histogram. The resulting histogram is called "sterile histogram" $\mathcal{H}_{s,kl}$.

Next, N_{GP} additional spectra corresponding to the null hypothesis, i.e. containing

no sterile neutrino signature ($\sin^2(\theta) = 0$), are generated and again passed through the model. The resulting outputs are used to create the "null histogram" $\mathcal{H}_{0,kl}$. The expectation value of $\mathcal{H}_{s,kl}$ is used as test statistic s for each grid point. Based on observations, the model outputs for each grid point kl are assumed to be normally distributed. Therefore, the test statistic s is equal to the mean of $\mathcal{H}_{s,kl}$. As threshold t, the 95th percentile of $\mathcal{H}_{0,kl}$ is used. The sensitivity contour is drawn when $s = t$. This corresponds to the case where the model is not able to distinguish between the expected spectrum containing a sterile neutrino signature and the reference spectra with a 95 % CL. The whole process is illustrated in figure 4.9.

4.3.2 Code Implementation

As ML framework, the PyTorch library with GPU accelerated training via the Apple MPS or Nvidia CUDA backend was used [110–112]. The package facilitating the data generation, model training and evaluation was developed specifically for this thesis and can be found here [113]. For Hyperparameter Optimization, the package SMAC3 was used [106]. It implements Bayesian Hyperparameter Optimization with a Hyperband Strategy.

To ensure reproducibility, `torch.manual_seed`, `torch.mps.manual_seed` and `torch.cuda.manual_seed` were fixed throughout the whole training and evaluation process. For the MPS backend however, this was insufficient with some fluctuations in training outcomes occurring. This needs to be investigated further in the future.

The data generation and preprocessing processes are parallelized, with optional just-in-time (jit) compilation of the relevant theory functions using numba. However, the overhead from the compilation itself proves to be too large to amortize with the speed-up it provides. Generating 10^6 statistically fluctuated samples takes 36 s[1]. The NN training averages at 1.85 epochs per second for a batch size of 25000. Currently, the main bottlenecks of the code are the hyperparameter optimization process and the generation of the sensitivity contours using the NN method. Drawing a sensitivity contour using 1000 samples to generate the test statistic and threshold values at each grid point, for a grid with 10^4 total grid points, takes around 10 minutes. Adding modifications to the dataset increases the computation time by up to a factor of 1.5.

[1] Calculations performed on an Apple M3 Pro MacBook Pro with a 12-core CPU and 18-core GPU.

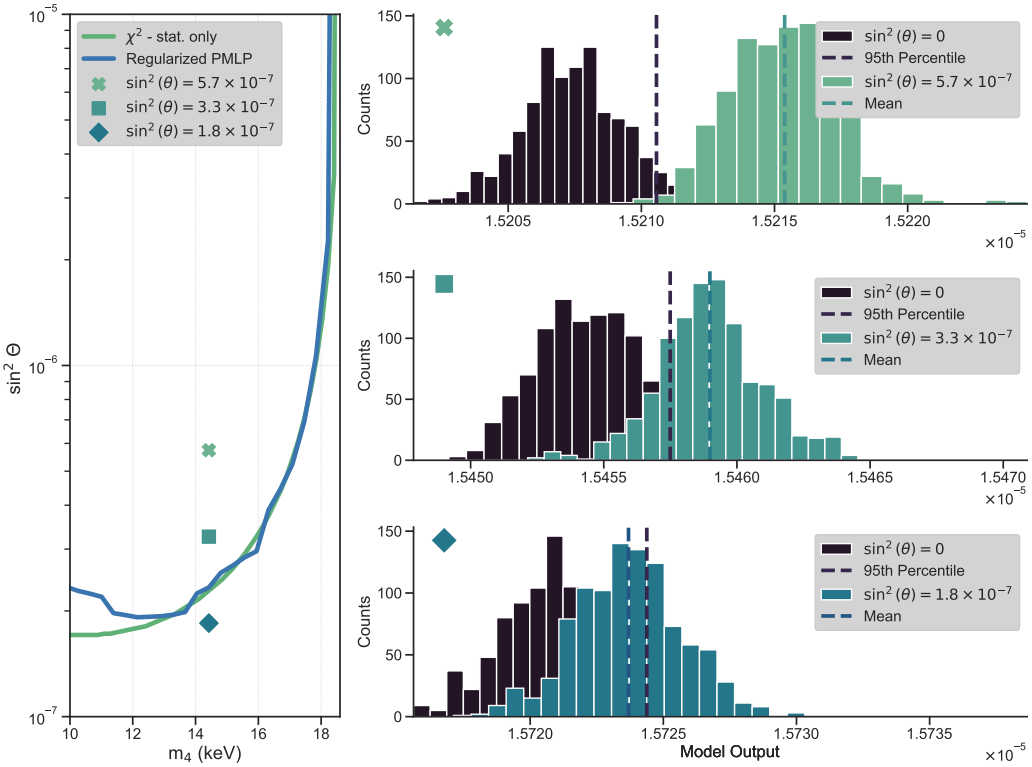


Figure 4.9: Illustration of the grid scan process to construct sensitivity contours for NNs. The left-hand plot shows a part of the sterile parameter space with two exclusion lines. The green line corresponds to the χ^2 exclusion line at 95 % CL for spectra only containing statistical uncertainty. The blue exclusion line is constructed using the algorithm outlined in section 4.3.1. On the three plots on the right, the exclusion line construction algorithm for the three highlighted grid points is shown explicitly. At each of these points 10^5 statistically fluctuated spectra containing the null hypothesis ($\sin^2(\theta) = 0$) are passed through the model and the outputs are used to fill a histogram (shown in black). Identically, 10^5 spectra containing a sterile neutrino signature equivalent to the grid point are used to create the colored histograms. As a test statistic, the mean of the model outputs for the sterile spectra is used. The 95th percentile of the null hypothesis histograms is used as a threshold for the test statistic. Test statistic and threshold get closer for smaller $\sin^2(\theta)$, until they eventually cross. The blue exclusion line is drawn at the point of crossing. This process is repeated for the entire sterile mass range, and with more grid points on the mixing angle axis than shown in this plot.

Chapter 5

Sensitivity Studies

This chapter investigates the sensitivity of the NNs of this work to sterile neutrino signatures that are accessible to the tritium spectrum measurement with the KATRIN experiment when equipped with the TRISTAN detector. Furthermore, it evaluates the robustness of the NNs under systematic uncertainties and its generalization capabilities for model inaccuracies.

The sensitivity of the NN quantifies its ability to exclude sterile neutrinos with specific masses m_4 and active-to-sterile mixing angles $\sin^2(\theta)$ based on the measured energy spectrum. The primary goal of this thesis is to illustrate the sensitivity of different NN architectures to the sterile signature in spectra containing statistical and systematic uncertainties. This is achieved using Binary Classification (BC), which means that the parameters of interest (m_4 and $\sin^2(\theta)$) are not inferred directly, in stark contrast to, for example, a parametric fit. First, the chapter covers the sensitivity to the sterile neutrino signature when only accounting for statistical uncertainty (section 5.1). The most performant NN is then used for further studies regarding the impact of detector related effects and the uncertainties involved in the parameters associated to their modelling. To model generic uncertainties, different shape factors with known parametrization are applied to the spectra and the robustness of the NNs even under nonphysical distortions is tested. Finally, the robustness of the NN under arbitrary unknown perturbations and unknown shape factors is investigated.

5.1 Model Performance with Statistical Uncertainties

As a baseline for performance evaluation, all the NNs were trained to classify statistically fluctuated spectra. The performance of a NN here is given by how close the sensitivity contour is to the statistical limit for each sterile neutrino mass m_4. The goal of this section is to demonstrate that NNs are very close to being as sensitive to the sterile neutrino signature as a $\Delta\chi^2$ test. Their sensitivity is evaluated using a grid scan over the sterile parameter space. The results of this study are depicted in figure 5.1, and the hyperparameters used for their training can be found in the appendix A.

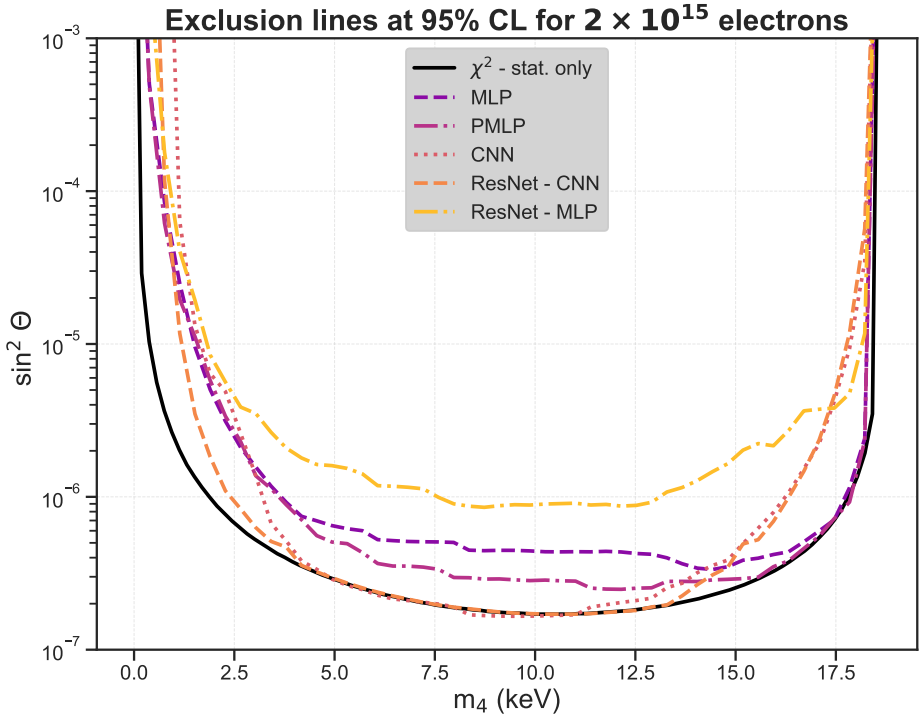


Figure 5.1: Statistical sensitivity contours for different NN architectures across the sterile parameter space at 95 % CL. The contours have been constructed using a grid scan and subsequent interpolation between grid points. They are each prone to statistical fluctuations. The CNN-based architectures reach the best sensitivity for sterile masses $m_4 \in [1, 13]$, while the MLP-based architectures dominate for large sterile masses.

The performance agrees well with the statistical limit set by the $\Delta\chi^2$ test, but none of the NN achieve the same level across the entire accessible sterile neutrino mass range. The closer the sensitivity contour when evaluated with a NN is to the statistical limit, the more sensitive the method is to the sterile neutrino signature. The most performant models for sterile masses in the range of $m_4 \in [2, 14]\,\mathrm{keV}$ are the shallow CNN and the deep CNN using residual connections (ResNet- CNN). For larger sterile masses $m_4 > 14\,\mathrm{keV}$, the MLP and Parametrized Multilayer Perceptron (PMLP) exhibit higher sensitivity. The deep MLP employing residual connections (ResNet- MLP) performs approximately half an order of magnitude worse than its shallower counterparts, potentially due to suboptimal hyperparame-

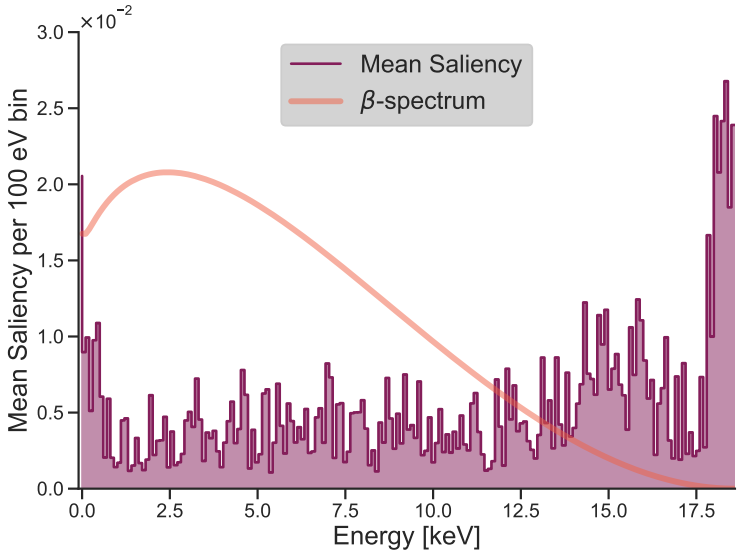


Figure 5.2: The saliency map averaged over all input spectra (purple), superimposed with the mean input spectrum (orange). The saliency map highlights the most important regions in an input that contribute to the NNs prediction. The higher the saliency value for an energy bin, the larger the influence of that area on the decision.

ter optimization, which has to be further investigated. Furthermore, all the models performances lack in the lower sterile range when compared to the statistical exclusion line.

A possible explanation for this phenomenon can be found by looking at the regions which influence the NNs output the most. The input regions most influential in the network's decision-making process are highlighted in a saliency map. The map is created by calculating a backward propagation pass for an input spectrum (i.e. calculating the gradient of the network with respect to the input) and saving the gradient for each energy bin. The average saliency map, created by averaging the saliency map of each evaluation spectrum, can be seen in figure 5.2. For illustrative purposes, it is superimposed on the average evaluation input spectrum. Based on the saliency map, it is apparent that the network pays close attention to the final few energy bins of the spectrum. Furthermore, the NN also gives the regions around $E_{\text{kin}} \approx 15\,\text{keV}$ and $E_{\text{kin}} \approx 0\,\text{keV}$ a higher importance weight, with the rest of the spectrum being less important for the decision. This illustrates that the sterile neutrino signature for smaller sterile masses is harder for the network to distinguish. Additionally, the NN seems to mainly base its output on relative rate differences

between the beginning and end point of the spectrum, with intermediate energy bins contributing only marginally.

The statistical fluctuations in each energy bin of the training and evaluation data propagate to the model output. As the sensitivity contours depend on the model output, they are also prone to statistical fluctuations. To investigate the impact of the statistical fluctuations in the evaluation data on the sensitivity contours, the evaluation dataset was bootstrapped. Bootstrapping involves resampling the evaluation dataset. A grid scan is then performed for each resampled dataset with the pre-trained model. This is repeated $N_B = 100$ times. The resulting contours are averaged and the standard deviation for the critical $\sin^2(\theta)$ value for each m_4 is calculated to draw a 1σ error band. The results for the PMLP can be seen in figure 5.3.

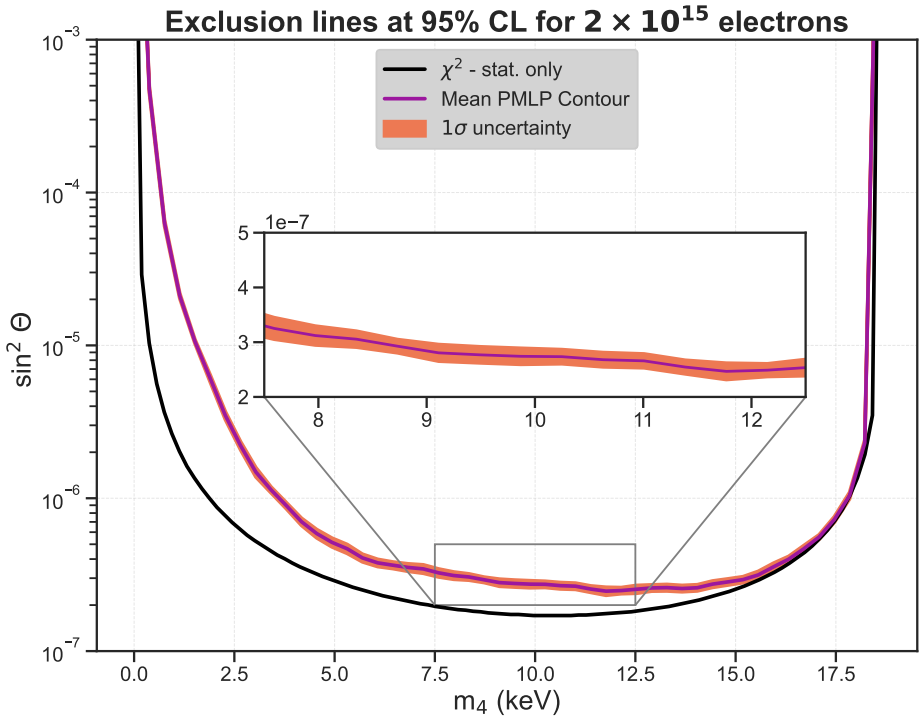


Figure 5.3: The mean contour and 1σ uncertainty band for the PMLP. The contour was calculated by bootstrapping the evaluation dataset. In detail, this means the spectra used to evaluate each grid point were resampled $N_\mathrm{B} = 100$ times and passed to the pre-trained model. The 100 resulting contours are averaged, and the standard deviation is used as an uncertainty.

In conclusion, the NN approach is extremely sensitive to a sterile neutrino signature in statistically fluctuated data, down to active-to-sterile mixing angles of $\sin^2(\theta) = 2\times10^{-2}$. The models only show a sensitivity loss compared to the statistical exclusion limit when it comes to smaller sterile masses.

The following section will show that this sensitivity loss can be alleviated by introducing a PAE response to the datasets, which in particular also boosts the PMLPs performance to be comparable to the CNN, as shown in figure 5.5. Thus, for the following studies, the PMLP is chosen as representative model and only the results for the PMLP will be shown. An additional reason to choose the PMLP as the representative model for this proof-of-concept study is its shorter training time and simpler architecture.

5.2 Detector related Effects

The spectrum measured by the TRISTAN detector will not simply be a statistically fluctuated β-electron energy spectrum like in figure 4.6. There are numerous experimental effects that modify the resulting spectrum. The most dominant effects at the time of writing this thesis are expected to be stemming from the RW, and detector related effects. This section demonstrates that the NN approach is largely insensitive to systematic uncertainties connected to detector related effects. In order to achieve this, the detector related effects will be introduced iteratively to all datasets in this section. In each iteration, the sensitivity using a PMLP is investigated for different realizations of the parameter carrying the systematic uncertainty. This is done to check whether the sensitivities evaluated with NNs are diminished by the addition of systematic effects. An overview of the detector related effects discussed in this thesis and their uncertainties is in table 5.1.

Parameter	Symbol	reference value	Uncertainty
Post acceleration energy	E_{pae}	10 keV	-
Dead layer parameter	λ	58 nm	2 nm
Charge cloud width	w_{cc}	20 µm	1 µm

Table 5.1: Parameter values and their systematic uncertainties involved in the modelling of detector related effects. The values were taken from [15].

Response Matrices

To add the experimental effect to the data, the same response matrix approach used in the TRModel, outlined in section 3.4, was chosen. Thus, detector related effects are

modelled in form of response matrices that can be convolved with a binned spectrum. The response matrices are taken from the catalog developed for the TRModel [93]. In this thesis, electron pitch angle related effects are ignored. All electrons are assumed to arrive perpendicular to the detector, corresponding to a pitch angle of 0°. This reduces the response matrices down to two dimensions: The input energy E_{in} in the first and the output energy E_{out} in the second dimension.

Systematic Uncertainties

The description of experimental effects often depends on parameters that carry a systematic uncertainty. The focus of this thesis also lies on the effect of these systematic uncertainties on the NN performance. The systematic uncertainty of a parameter p describing an experimental effect is modelled by drawing the value of the parameter from a normal distribution centered at the reference value v, with the standard deviation given by its uncertainty u:

$$p \sim \mathcal{N}(\mu = v,\, \sigma = u). \tag{5.1}$$

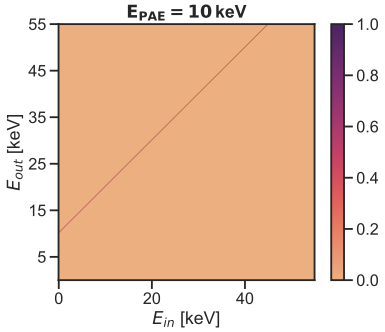
Two approaches can be used to introduce systematic uncertainties of parameters describing experimental effects to the analysis chain when using NNs.

- Drawing the value of p for each spectrum in the training data individually from $\mathcal{N}$ and convolving the spectrum with response matrix corresponding to the resulting value.
- Using only the reference value v of the parameter p for the entire training data set.

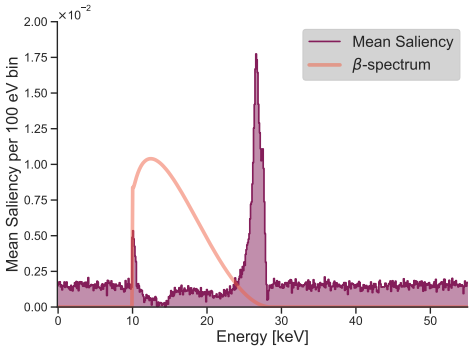
The evaluation is then performed and compared for different values of the parameter p, starting with the reference value, and n standard deviations away from the reference value $\mu \pm n \times \sigma$.

Post Acceleration Electrode Response

The first experimental effect added to the datasets is the PAE response. The PAE voltage is set to 10 kV, and no systematic uncertainty is considered. The reason for it being discussed here nonetheless is that the sensitivity loss due to systematic effects and uncertainties in the parameters describing detector related effects is mitigated by using a PAE [15]. Therefore, the PAE response is added to the data before the other response matrices. The PAE response simply shifts the input energies by $E_{\text{PAE}} = 10\,\text{keV}$ and can be seen in figure 5.4. The size of this response matrix is chosen such that the output data is compatible with the other responses.



(a) The 2D response matrix for a PAE set to 10 kV. An input energy of $E_{\text{in}} = i\,\text{keV}$ is shifted to an output energy of $E_{\text{out}} = (i+10)\,\text{keV}$.



(b) The mean saliency map for a model trained on statistically fluctuated data with a PAE response of 10 kV.

Figure 5.4: Effects of adding a PAE response to the data.

By adding this response an interesting effect was discovered: extending the range of interest beyond 18.6 keV boosts the model performance significantly in the lower sterile mass region, as demonstrated in figure 5.5. A possible explanation can be found again by looking at the mean saliency map. The region that influences the models decision the most is close to the kinematic endpoint of the spectrum (shifted by $E_{\text{PAE}} = 10\,\text{keV}$) as shown in figure 5.4. Extending the range of interest beyond the shifted endpoint and to the left of the beginning of the spectrum, where the differential decay rate is strictly zero, still influences the decision of the NN. Picking up on the explanation from section 5.1, these points then serve as additional points the NN can base a relative difference to the endpoint on. However, these points do not fluctuate (as they are strictly zero) and thus serve as a good baseline for the model. This makes it easier to detect the sterile signatures for smaller sterile masses and consequentially boosts the model performance significantly in the lower sterile mass region and slightly in the medium mass region. The magnitude of E_{PAE}, which dictates how far the spectrum is shifted towards higher energies, does not influence the model performance, further solidifying this explanation (see appendix B). Furthermore, the amount of zero entries itself (dictated by the size of the PAE response matrix) does influence the model performance slightly (see appendix B). This theory has to be studied further to reach a distinctive conclusion.

For the following studies, the PMLP is chosen as representative model and only the results for the PMLP will be shown. This is because after extending the range-of-interest beyond the kinematic endpoint, like with a the PAE response, the PMLP

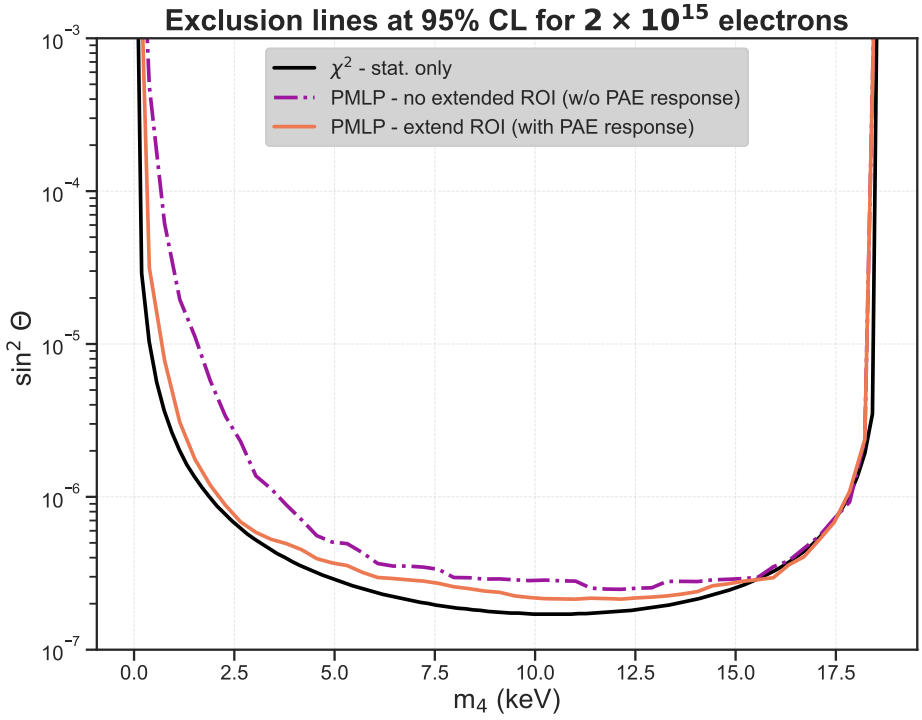


Figure 5.5: In this figure, the PMLP sensitivity to the sterile neutrino signature when training on statistically fluctuated data before and after adding a PAE response are compared. The PAE response increases the range-of-interest (ROI). This improves the NNs sensitivity to smaller sterile neutrino masses significantly.

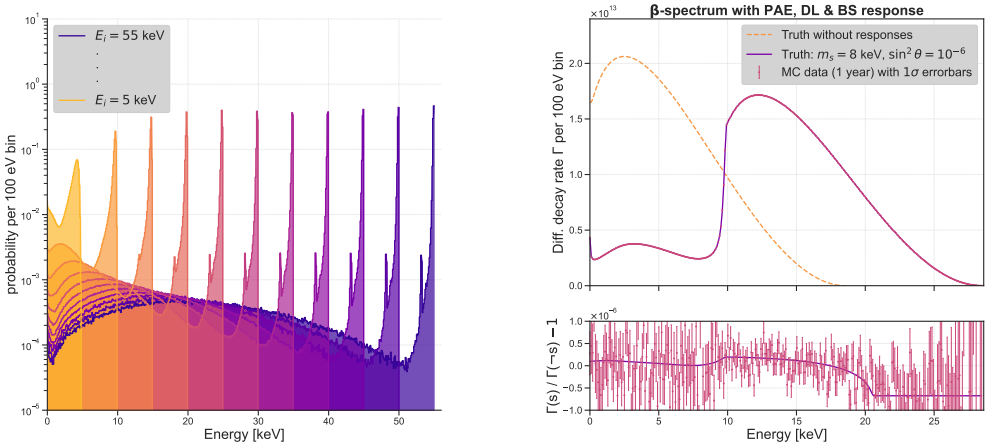
performance is comparable if not better than that of the CNN based architectures. Furthermore, the PMLP is faster to train and utilizes a simpler architecture.

Detector Dead Layer

The detector dead layer width and electron backscattering are the first two detector effects that are investigated. An explanation of these two effects can be found in section 3.4. Both are responsible for energy deposition losses. The uncertainty associated with the probability of the electron backscattering at the detector is neglected in this work. Please refer to [15] for more information on the construction of the response matrices. The detector dead layer is described by the parameter λ, which indicates the width of the dead layer and carries a systematic uncertainty (see table 5.1). The electron backscattering and detector dead layer response is shown in

figure 5.6.

A PMLP is now trained on data containing the response with the dead layer param-



(a) Detector response (DL + BS) for different electron input energies.

(b) Spectrum after adding the DL + BS response for $\lambda = 58\,\text{nm}$.

Figure 5.6: Illustration of the detector dead layer and electron backscattering and back-reflection response. Shown is the spectrum after applying the dead layer response for different input energies (a), and one exemplary spectrum after applying the response (b). The detector response for different input energies exhibits a peak at the input energy, which is then broadened to the left from dead layer energy loss. It also contains a silicon escape peak 1.74 keV below the main peak and a backscattering tail extending to lower energies. The backscattering tail is clearly noticeable in the spectrum itself for energies smaller than 10 keV.

eter λ at its reference value of 58 nm. It is then evaluated on a dataset containing responses with $\lambda_{\pm n\sigma}$ it has not seen during training. The PMLPs performance is shown in figure 5.7. The sensitivity contour is robust under changes in the dead layer parameter of up to 2σ, corresponding to a dead layer parameter of $\lambda_{+2\sigma} = 62\,\text{nm}$. Here, the performance decreases for small sterile-to-active mixing angles $\sin^2(\theta)$. A possible explanation being that in this case the variance of the model output gets larger for small $\sin^2(\theta)$. This makes increases the range of fluctuations in the mean. Only with larger deviations from the reference value it starts to lose its predictive power. The effect that varying λ has on the data is illustrated in figure C.1.

When drawing $\lambda \sim \mathcal{N}(\mu = 58\,\text{nm}, \sigma = 2\,\text{nm})$ for each spectrum in the training dataset individually and then using a fixed value $\lambda_{\pm n\sigma}$, the contour for $\lambda_{\pm 2\sigma}$ gets

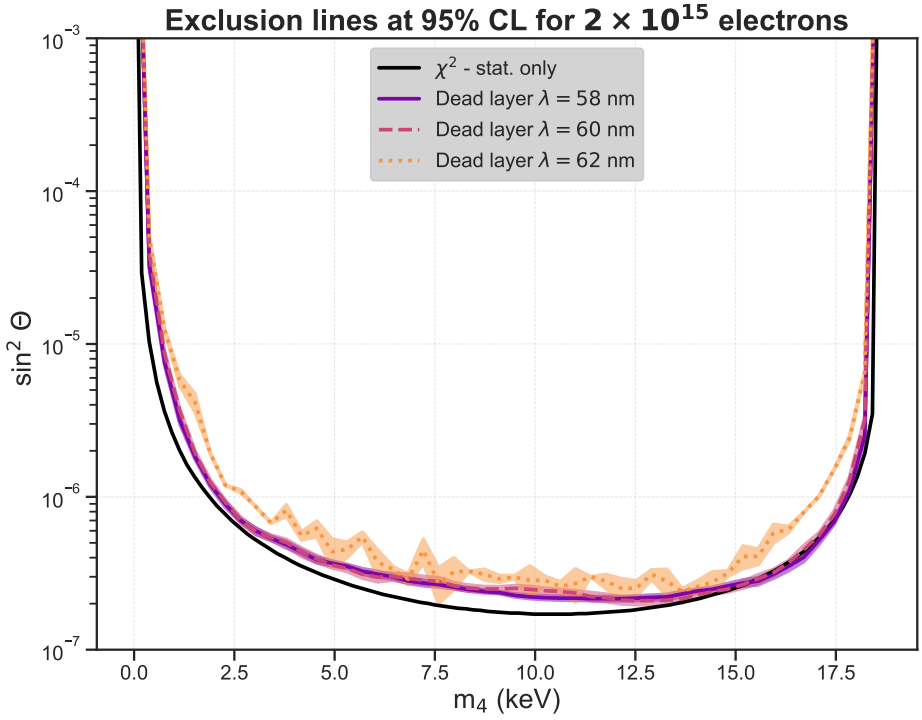


Figure 5.7: Sensitivity contours for various dead layer parameters (λ). The model is trained using data corresponding to the reference value $\lambda = 58\,\text{nm}$ but evaluated for other λ values to account for systematic uncertainty. For λ values offset by twice the uncertainty ($\sigma_\lambda = 2\,\text{nm}$), model performance breaks down completely for $\lambda_{\pm 3\sigma} = 64\,\text{nm}$. This behavior is symmetric for $\lambda < 58\,\text{nm}$. Uncertainty bands are derived by resampling the evaluation dataset $N_B = 10$ times.

smoother. This is illustrated in figure 5.8. In that case, however, the training budget had to be doubled to 40 epochs to achieve a comparable performance, whereas a larger training budget in the former case does not lead to a better performance. The higher variance resulting from varying λ-values in the training dataset also leads to a slight performance decrease for smaller sterile-to-active mixing angles $\sin^2(\theta)$.

In conclusion, the model is robust even when evaluating it on data containing previously unseen dead layer responses with λ values of up to 2σ away from the reference value. Incorporating the systematic uncertainty by varying λ in the training dataset makes model performance stay consistent even further away from the reference value,

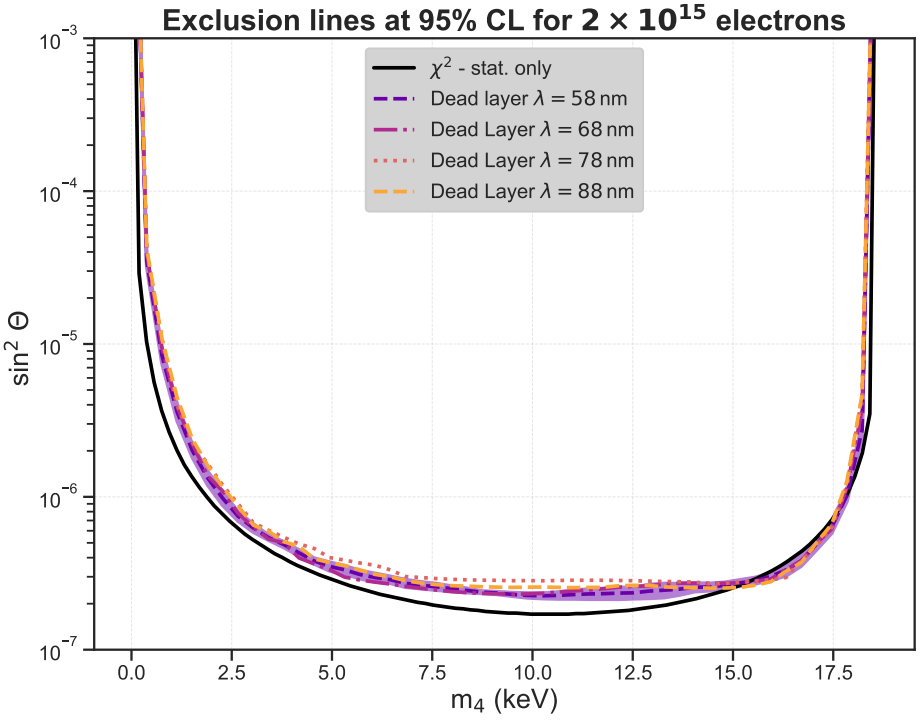


Figure 5.8: The effect of introducing the systematic uncertainty to the training dataset. The model is trained using data with dead layer parameters $\lambda \sim \mathcal{N}(\mu = 58\,\text{nm}, \sigma = 2\,\text{nm})$, and evaluated on datasets with fixed $\lambda_{\pm n\sigma}$. A small sensitivity loss is observed when the dead layer width is significantly decreased $\lambda > 68\,\text{nm}$. Nonetheless, the significance of this loss is difficult to quantify precisely due to the statistical uncertainty of the evaluation.

at the cost of a larger training budget and a slight performance decrease. Thus, the latter strategy should be used for parameters with large systematic uncertainties, to guarantee a more predictable model performance. A comparison with the χ^2-based approach from equation 4.18 is shown in figure 5.9.

Charge Sharing

The TRISTAN SDD consists of 166 hexagonal pixels per module. When incident electrons strike near a pixel boundary, the generated charge is shared between adjacent pixels. The proportion of charge shared with neighboring pixels is determined by the proximity of the hit position to the pixel boundary and the lateral spread

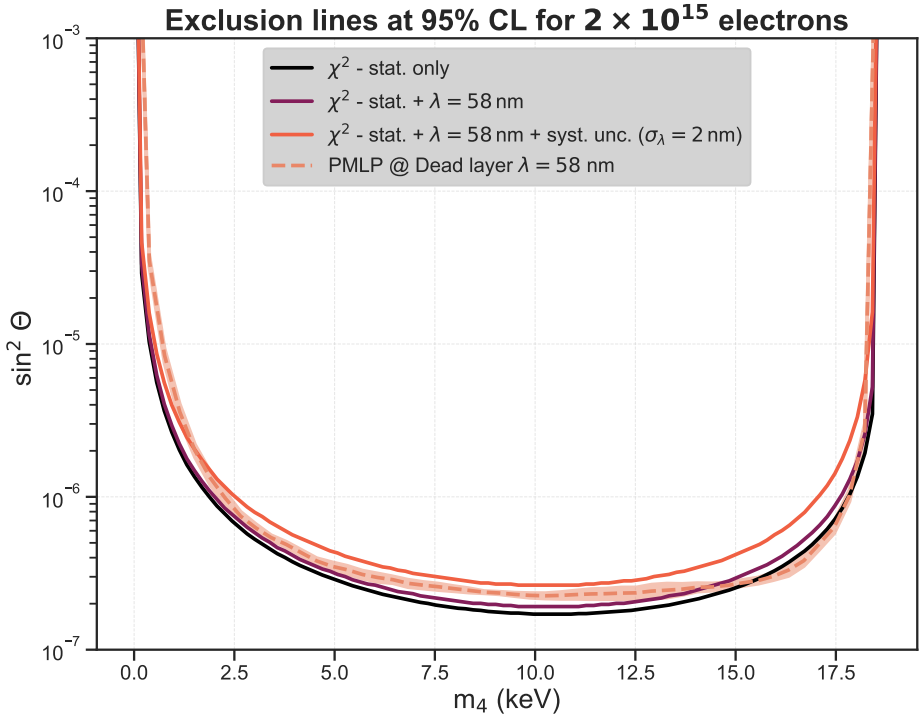
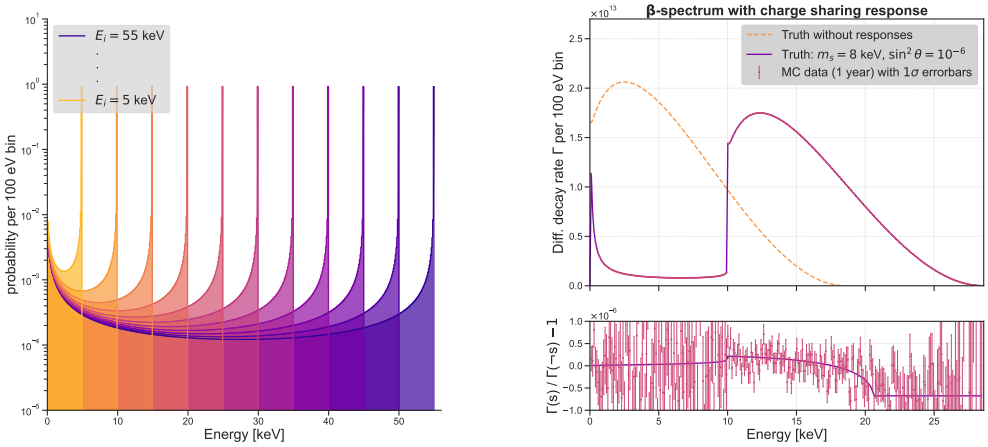


Figure 5.9: Comparison between the sensitivity contours of the NN-based and the χ^2-based approaches. Shown is the contour of the PMLP trained on data with dead layer parameters $\lambda \sim \mathcal{N}(\mu = 58\,\text{nm}, \sigma = 2\,\text{nm})$ and evaluated on a dataset with a fixed $\lambda = 58\,\text{nm}$ (dashed line). Adding the dead layer response for $\lambda = 58\,\text{nm}$ to the dataset results in the purple contour. Incorporating a systematic uncertainty of $\sigma = 2\,\text{nm}$ via a covariance matrix results in the orange contour.

of the charge cloud w_{cc}. The latest measurements report a charge cloud width of $w_{\text{cc}} = (16.3 \pm 0.2)\,\mu\text{m}$ [90]. Nonetheless, to ensure consistency with previous studies, the value currently used in the TRModel and reported in table 5.1 is applied in this study. The hexagonal shape of the pixel is neglected during modelling as the pixel radius $r_{\text{px}} = 1.5\,\text{mm}$ is much larger than w_{cc} [15]. The charge sharing response is shown in figure 5.10.

The NN is trained on data with $w_{\text{cc}} = 20\,\text{nm}$, and evaluated on data with w_{cc} shifted by up to three times the uncertainty $\sigma_w = 1\,\text{nm}$. Figure 5.11 demonstrates that the NN is very robust under unexpected deviations from the reference value



(a) Detector response (CS) for different electron input energies.

(b) Spectrum after adding the charge sharing response for $w_{\mathrm{cc}} = 20\,\mathrm{nm}$.

Figure 5.10: Demonstration of the detector charge sharing response. Shown is the resulting spectrum after the response for different input energies (a), and one exemplary spectrum after applying the response (b). For these results, a charge cloud width of $w_{\mathrm{cc}} = 20\,\mathrm{nm}$ and detector pixel radius of $r_{px} = 1.5\,\mathrm{mm}$ were used.

for w_{cc}. When introducing the systematic uncertainty to the training dataset, i.e. sampling $w_{cc} \sim \mathcal{N}(\mu = 20\,\mathrm{nm}, \sigma = 1\,\mathrm{nm})$ for each sample in the training dataset, the performance of the NN stays robust for charge cloud width values of up to $w_{cc} = 40\,\mathrm{nm}$, as shown in figure 5.12. The effect that varying w_{cc} has on the data is illustrated in figure C.1. From this it is apparent that the impact of varying the charge sharing width w_{cc} on the sterile neutrino signature is very small. A comparison with the χ^2-based approach from equation 4.18 is shown in figure 5.17.

In conclusion, the NN performance stays consistent for effects that do not strongly affect the sterile neutrino signature, especially if they are encoded in the training dataset.

Energy Resolution

To model a general energy resolution of the detector, each spectrum in the dataset is subject to Gaussian smearing with a standard deviation of σ_{ER}. The model is trained on β-spectra with a bin width of 100 eV and no Gaussian smearing. It is evaluated on spectra containing $\sigma_{\mathrm{ER}} = [100, 200, 300]\,\mathrm{eV}$. The results are shown in

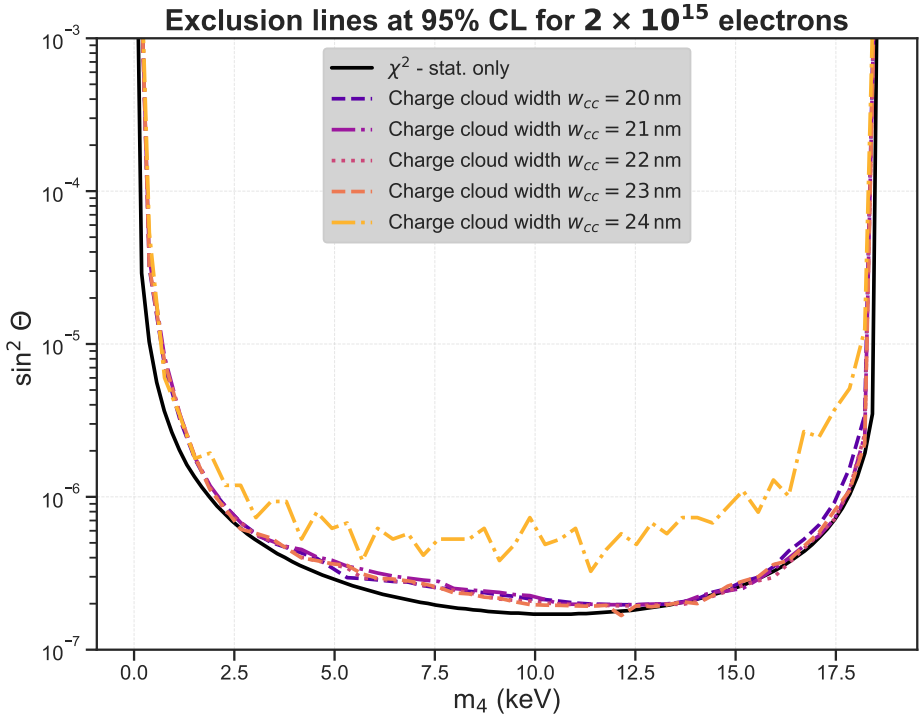


Figure 5.11: In this figure, the sensitivity contours for various charge sharing widths (w_{cc}) are shown. The model is trained using data corresponding to the reference value $w_{\mathrm{cc}} = 20\,\mathrm{nm}$ but evaluated for other w_{cc} values to account for systematic uncertainty. The model performance completely breaks down for shifts of 4 nm. This behavior is identical for decreasing w_{cc}.

figure 5.14.

The NN performance is only slightly affected by differing energy resolutions. This illustrates that the model can optionally be trained without having to take the total energy resolution of the detector into account, as long as it does not deviate from the bin width by more than 200 eV. To ensure optimal performance, however, the detector's energy resolution should be fixed to the actual value for the training dataset.

In conclusion the NN method is not strongly affected by the systematic uncertainties discussed above. Similarly, as shown in Figure **??**, χ^2-based methods are also only slightly impacted by these systematic effects and the corresponding systematic

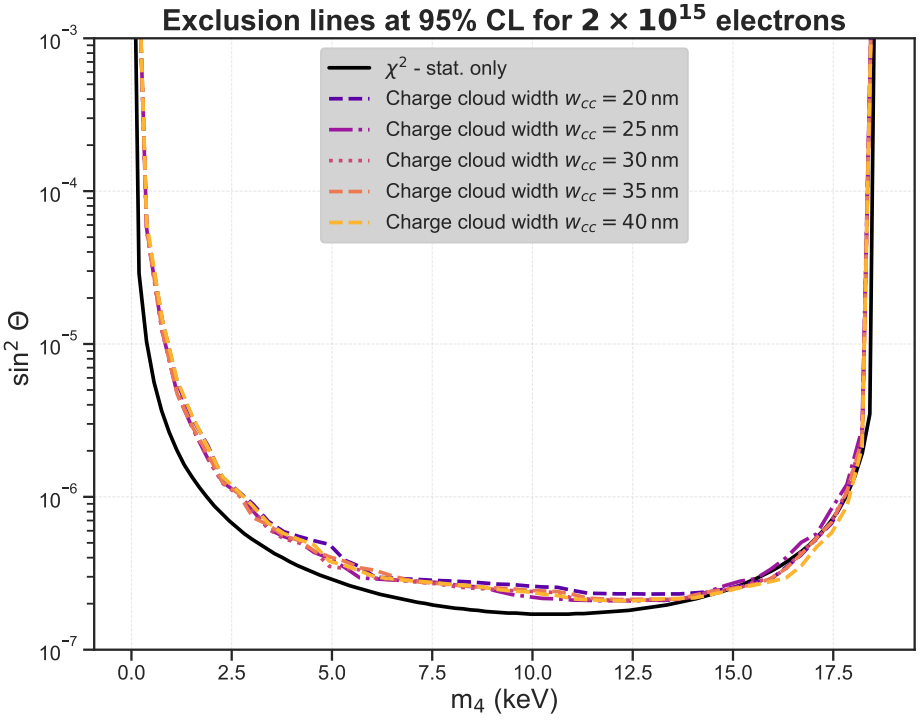


Figure 5.12: In this figure, the sensitivity contours for various charge sharing widths (w_{cc}) are shown. The model is trained using data with charge sharing parameters $w_{cc} \sim \mathcal{N}(\mu = 20\,\mathrm{nm},\, \sigma = 1\,\mathrm{nm})$, and evaluated on datasets with in fixed $w_{cc,\pm nu}$. The model performance stays robust even for shifts of up to 20 nm, at the overall cost of some sensitivity.

uncertainties. If larger deviations from the reference value are expected, it helps to introduce the systematic uncertainty in the training dataset, at the cost of a larger training budget and a slight performance decrease. Here, the network's performance surpasses the χ^2 method in the case of the detector dead layer systematic.

5.3 Effect of Modelling Inaccuracies

The goal of this section is to conduct a qualitative analysis of the impact of more general shape uncertainties and biased training data on the sensitivity to the sterile neutrino signature with NNs.

The general shape uncertainty serves as a proxy for potential systematic uncer-

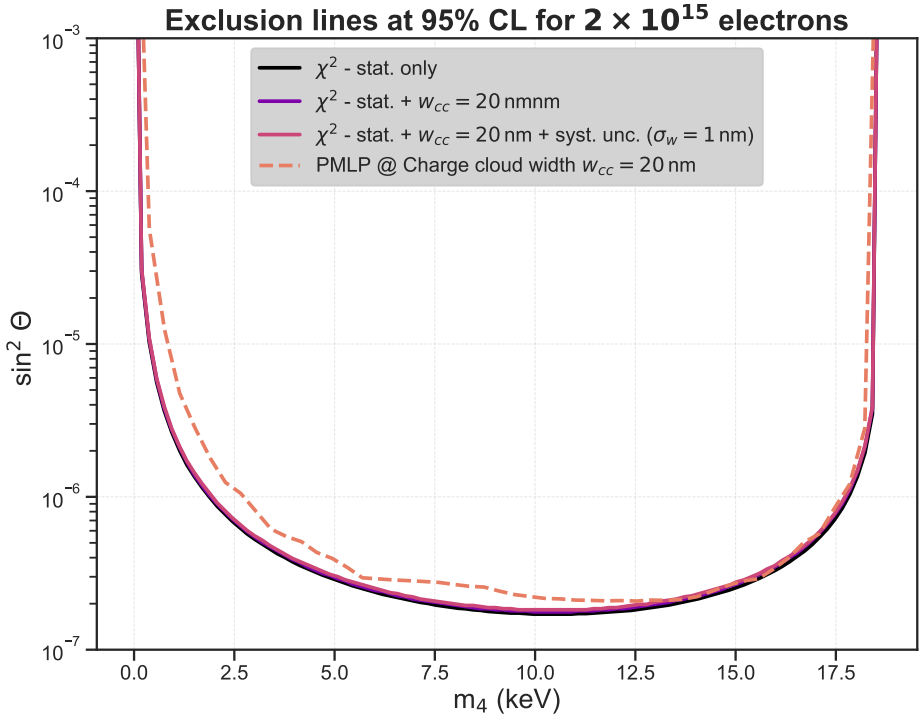


Figure 5.13: Comparison between the sensitivity contours of the NN-based and the χ^2-based approaches. Shown is the contour of the PMLP trained on data with charge cloud width parameters $w_{\mathrm{cc}} \sim \mathcal{N}(\mu = 20\,\mathrm{nm}, \sigma = 1\,\mathrm{nm})$ and evaluated on a dataset with a fixed $w_{\mathrm{cc}} = 20\,\mathrm{nm}$ (dashed line). Adding the charge sharing response for $w_{\mathrm{cc}} = 20\,\mathrm{nm}$ to the dataset results in the purple contour. Incorporating a systematic uncertainty of $\sigma = 1\,\mathrm{nm}$ via a covariance matrix results in the red contour.

tainties that could impact the shape of the tritium β-decay spectrum, particularly in the form of a smooth energy-dependent function.

The NN-based approach is not strongly affected by systematic uncertainties, if the parameters that encode a systematic effect are varied in the training dataset. However, small modifications of the shape that are previously unknown or that stem from modelling inaccuracies can not be treated this way. Therefore, the case of using biased training data is investigated. "Biased" training data means that the training data and evaluation data are fundamentally different. This is achieved by applying a small and smooth random perturbation to only the evaluation dataset.

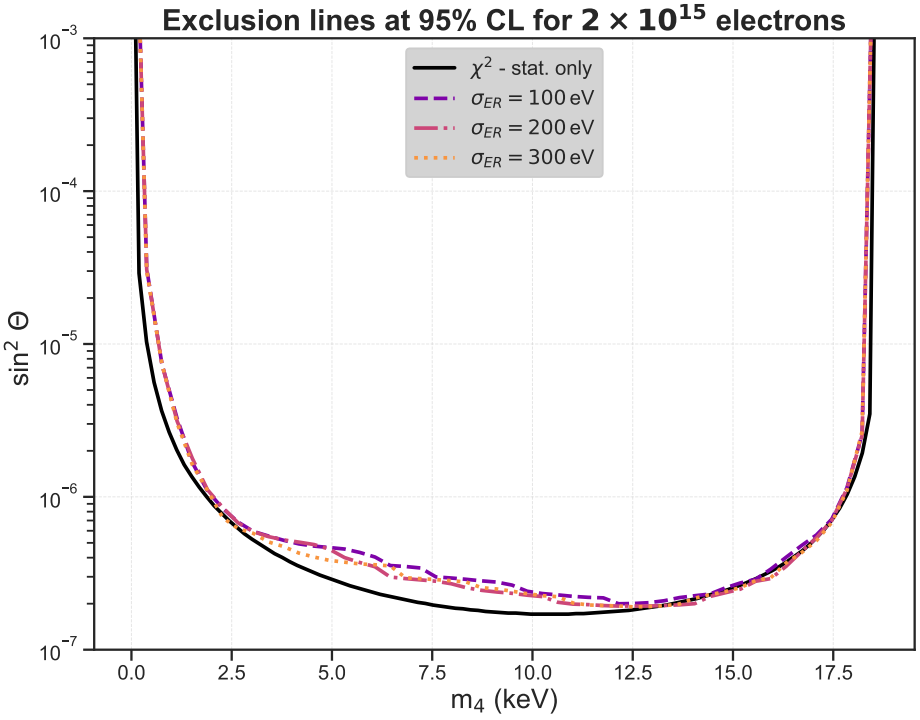


Figure 5.14: In this figure, the sensitivity contours for three different energy resolutions are shown. The model is trained on β-spectra with a bin width of 100 eV, but without any gaussian smearing. Each β-spectrum in the evaluation dataset is convolved with a gaussian filter with width σ_{ER}.

5.3.1 Shape Uncertainty

First, the impact of a systematic uncertainty affecting the shape of the tritium β-decay spectrum is studied in a more generic way. For this purpose, a shape factor polynomial is used to model a systematic effect, with its coefficients being drawn from a normal distribution centered at zero to model a systematic uncertainty. It is demonstrated that the sensitivity evaluated with NNs is not strongly dependent on the precise knowledge of the spectral shape and therefore the sensitivity does not degrade when evaluated under smaller systematic uncertainties that affect the shape of the spectrum, provided that the changes to the tritium β-decay spectrum are smooth functions of energy.

Each spectrum in a dataset is multiplied with a polynomial function of the electron

kinetic energy E_e

$$p(x) = 1 + \sum_{i=1}^{n} c_i \cdot x^i, \qquad x = \frac{E_e - E_0}{E_0}, \tag{5.2}$$

where E_0 is the kinematic end point of the spectrum, n the degree of the polynomial and $c_i \in \mathbb{R}$ the coefficients. Because of the currently ongoing KATRIN neutrino mass measurement, previous efforts to understand the spectral shape have primarily focused on the region near the endpoint. Therefore, uncertainties on the spectral shape are assumed to increase at lower β-electron energies. To simulate the impact of a spectral shape uncertainty on the sensitivity, the spectral shape is allowed to vary within a band of width σ, as illustrated in figure 5.15.

Optionally, for each sample in the training data, the coefficients c_i can be drawn from a normal distribution with a mean of zero and a standard deviation of $\sigma = 0.01$. With the current hyperparameters and training budgets it was not possible to achieve a robust result for larger σ. For each sample in the evaluation data, the c_i are set to a fixed value. In both cases, after multiplication with a shape factor polynomial each spectrum is rescaled to be in agreement with the total measurement statistics of 2×10^{15} β-electrons for one year of measurement time. This affects the shape of the uncertainty band shown in figure 5.15.

The results of the PMLP can be seen in figure 5.16. The model performance stays robust for $c_i < 0.05$. Larger values start to have an effect on the NNs sensitivity to the sterile neutrino signature. A comparison with the χ^2-based approach from equation 4.18 is shown in figure 5.17.

5.3.2 Biased Model

The goal of this study is to conduct a qualitative analysis of the impact of using biased training data on the NNs performance. To conduct this study, the NN is trained on statistically fluctuated data containing a PAE response. A perturbation is applied to the evaluation data (bias), and the NNs sensitivity contour is drawn.

Generating perturbations

Because the concept of a "biased model" is broad, the focus is set on a specific subset of random perturbations. By adopting a simple yet generic shape distortion function, the relevance of the analysis is preserved while ensuring it remains simple and interpretable. In this study, the random perturbations $b(x)$ are represented as a

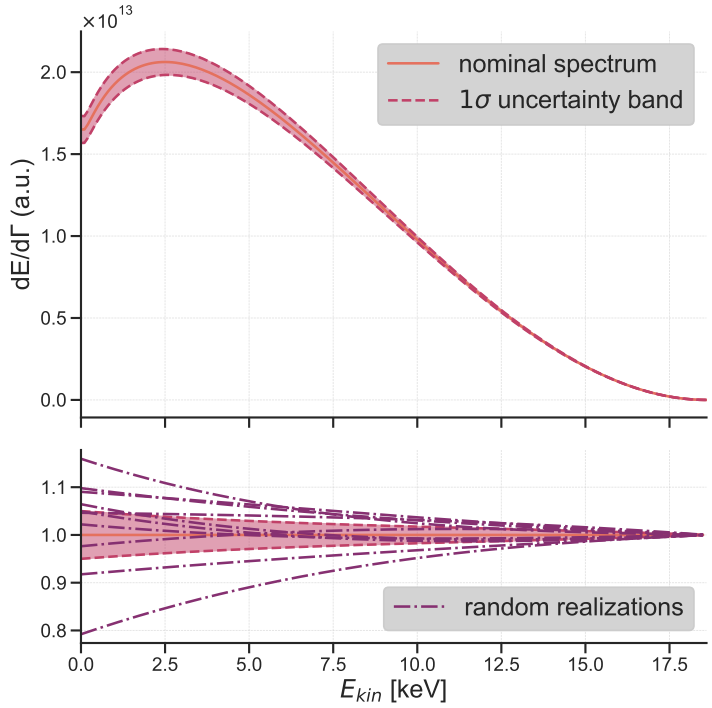


Figure 5.15: This figure illustrates an assumed 1σ shape uncertainty of the tritium β-decay spectrum. The dashed lines represent the spectrum after multiplication with a shape factor polynomial of degree three. For the linear, quadratic and cubic coefficients an uncertainty of $\sigma = \pm 5\,\%$ is used. This means that coefficients used for scaling the training data are sampled from a normal distribution with a width of 1σ. Shown below are random realizations of the shape factor polynomial applied to a constant function.

sum of normal distributions

$$b(x) = N \cdot \sum_{i=1}^{n} A_i \cdot \exp \frac{-(x-\mu_i)^2}{2\sigma_i^2}, \qquad x = E_e, \tag{5.3}$$

with the number of normal distributions n, the amplitude A_i, mean μ_i and standard deviation σ_i of the individual distributions, and the overall normalization of the perturbation N as tunable parameters. Additionally, a constraint on the smoothness and monotonicity can be applied. The arbitrary shape function samples with and without a smoothness and monotonicity constraints chosen for this study can be seen in figure 5.18. Figure 5.19 shows the impact of differing the perturbation amplitude parameter N for this specific shape function on a sterile neutrino signature.

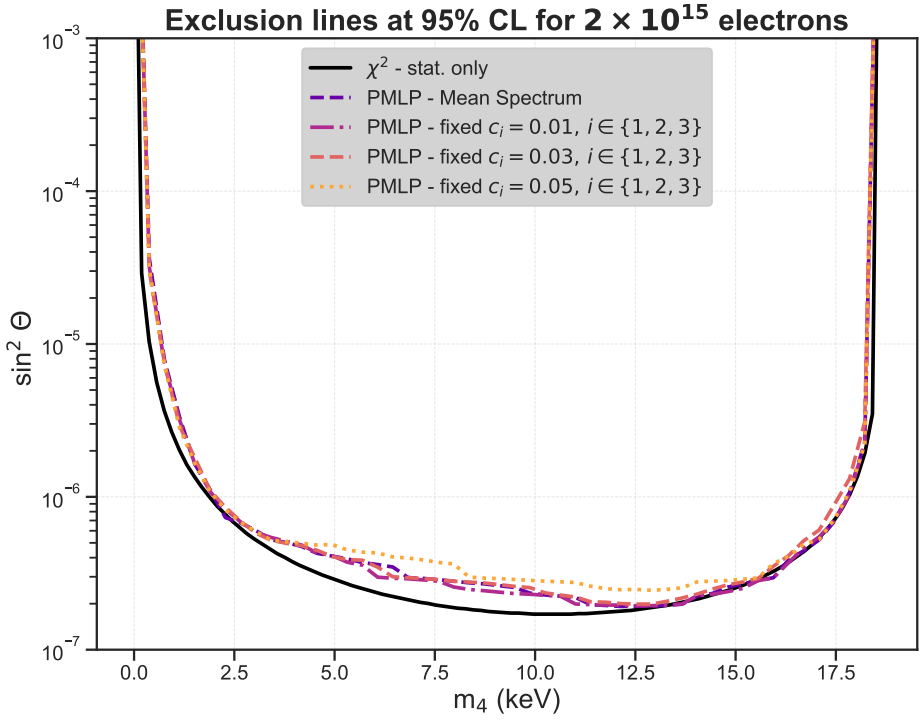


Figure 5.16: Effects of adding a third degree shape factor polynomial to the evaluation dataset. The contour is drawn for the shape factor coefficients c_i at different fixed values. The performance on the mean spectrum is shown as a reference. The performance starts to deteriorate for shape factors with $c_i > 0.05$.

Neural network performance

The NN is trained only on data with statistical uncertainty and a PAE response, not containing the shape functions. The particular shape functions shown in figure 5.18 are used for the evaluation of the NN. They are applied to each spectrum in the evaluation dataset. This is repeated for three different total amplitudes $N = \{10^{-7}, 10^{-6}, 10^{-5}\}$. The impact of the amplitude parameter N on the sensitivity to the sterile neutrino signature using the PMLP is shown in figure 5.20. The PMLP is largely unaffected by perturbations with an amplitude of up to 10^{-5}, even though they significantly change the sterile neutrino signature. Furthermore, the sensitivity is not affected when considering a shape function without constraints, as also illustrated in figure 5.20.

Repeating this for different realizations of the random shape functions produces

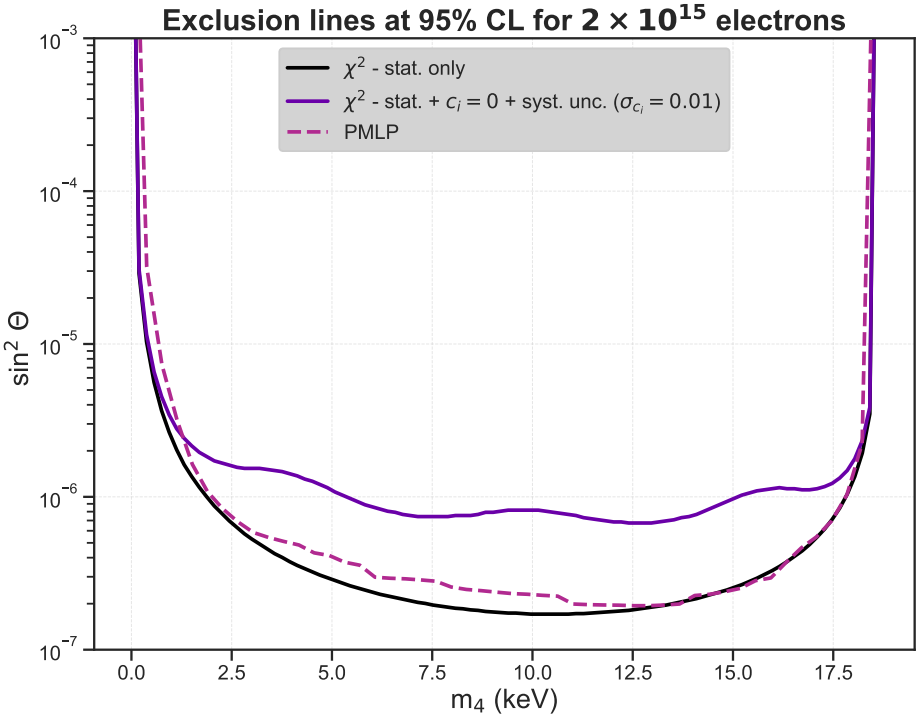
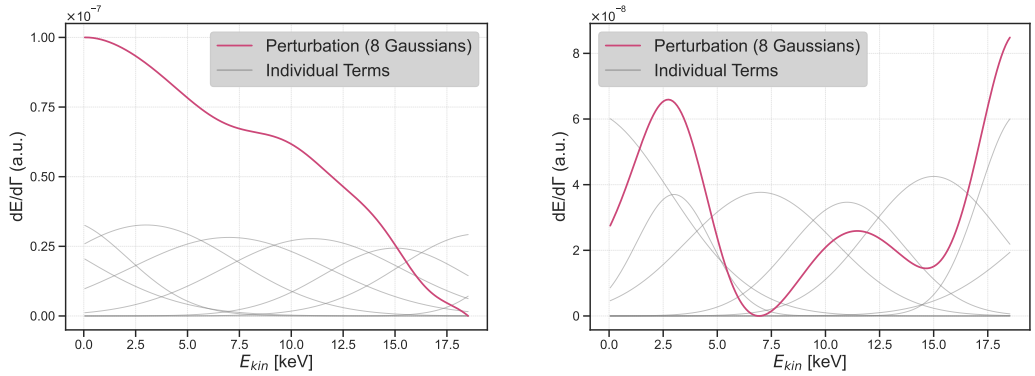


Figure 5.17: Comparison between the sensitivity contours of the NN-based and the χ^2-based approaches. Shown is the contour of the PMLP trained on data with shape factor coefficients $c_i \sim \mathcal{N}(\mu = 0, \sigma = 0.01)$ and evaluated on a dataset with a fixed $c_i = 0$ (dashed line). Incorporating a systematic uncertainty of $\sigma = 0.01$ via a covariance matrix results in the purple contour.

an identical result. A comparison with a χ^2-based approach was not feasible due to fit-convergence issues.



(a) With smoothness and monotonicity constraint.

(b) Without smoothness and monotonicity constraint.

Figure 5.18: Example of random perturbations constructed from 8 different gaussian distributions shown in gray. The amplitudes of the gaussians are not to scale.

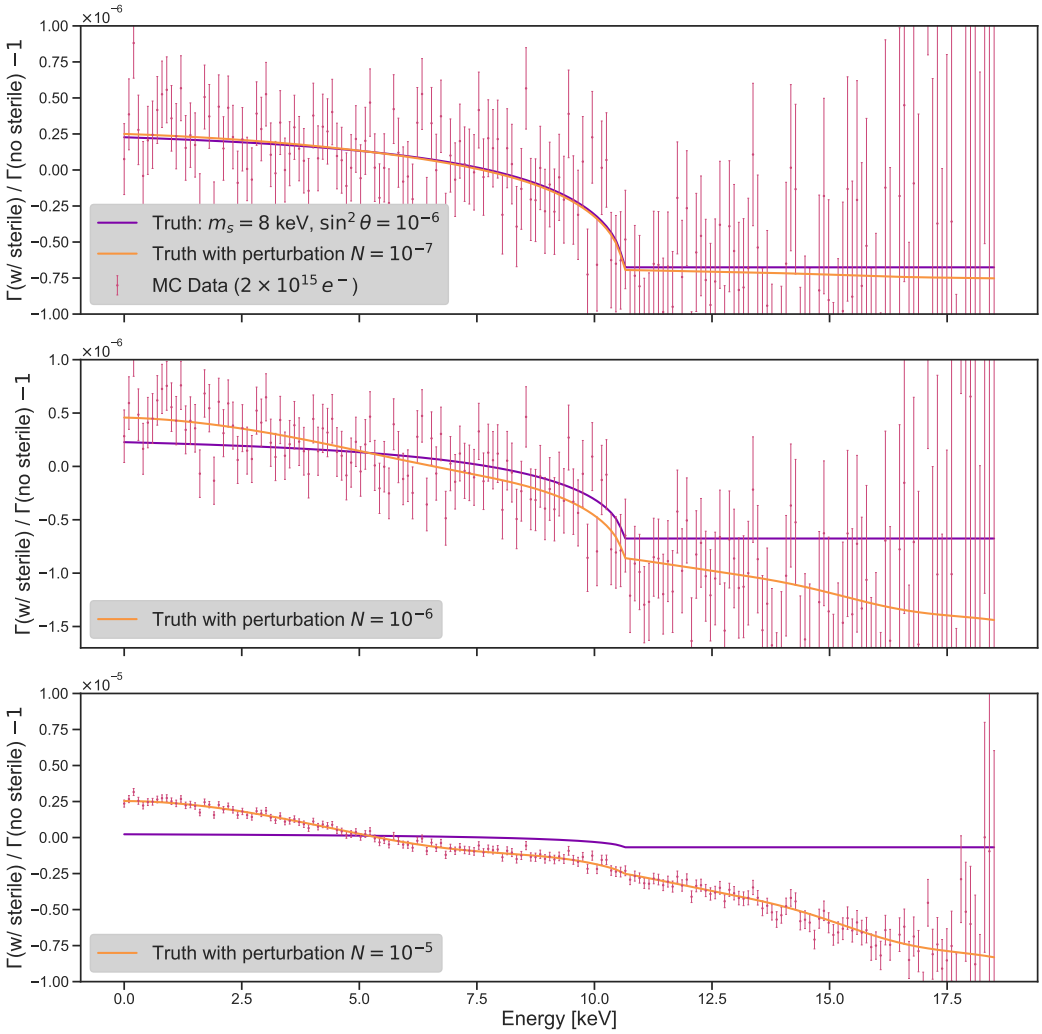


Figure 5.19: This figure illustrates the impact of the bias function amplitude N on the sterile neutrino signature. The theoretical sterile signature (violet) is modified by perturbation amplitudes varying between the three plots (orange). The perturbed theoretical spectrum is used to generate MC data (red). A NN trained on unperturbed spectra is then evaluated on spectra with specific perturbation amplitudes.

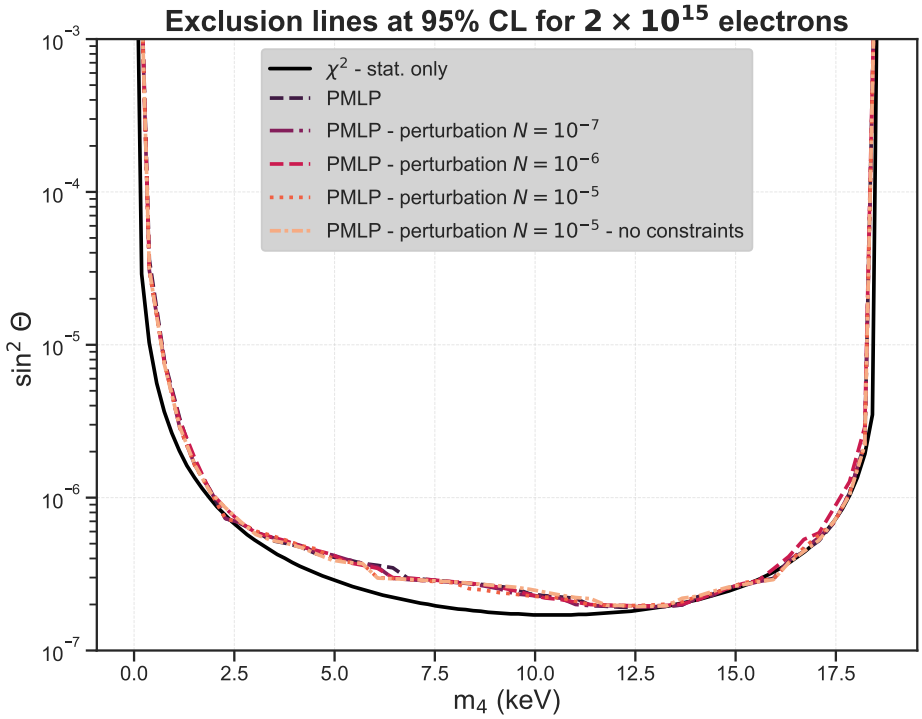


Figure 5.20: The sensitivity contours using a PMLP for different normalizations N of the two bias functions shown in figure 5.18. The PMLP is not impacted by the bias function and all three normalization values considered in this analysis. The effect of the bias function normalization on the sterile neutrino signature is illustrated in figure 5.19. The constraints on the perturbation do impact the network's performance, as illustrated by the exclusion line for a perturbation without constraints.

Chapter 6

Conclusion and Outlook

This thesis has demonstrated that various neural networks (NNs) are a powerful and effective tool for detecting the signature of a keV-scale sterile neutrino in the tritium β-electron energy spectrum. To that end, a library was developed to handles the training and evaluation of different NN architectures.

The NN method achieves a statistical sensitivity to the sterile neutrino signature comparable to the χ^2-method for statistically fluctuated data, which is close to an active-to-sterile mixing angle of 2×10^{-7} for a total measurement time of one year. The statistical limit is not strongly affected when experimental effects are included in the data.

The results demonstrate that the sensitivity to a sterile signature using the NN approach is not strongly affected by systematic uncertainties of detector related effects. This was achieved by evaluating the NN performance on mock data for which the parameters describing detector systematics were perturbed according to their nominal values by multiples of their systematic uncertainties.

Furthermore, it is shown that the exact shape of the β-spectrum does not need to be known to a precision level close to the size of active-to-sterile neutrino mixing for the NN approach to be sensitive to a sterile neutrino signal. This was accomplished by introducing a shape uncertainty to the β-spectrum, by which the NN was not strongly affected.

Finally, it is established that the NN approach is not strongly affected by unexpected smooth perturbations with amplitudes of up to two orders of magnitude above the active-to-sterile neutrino mixing. This was demonstrated by adding a shape function consisting of a sum of gaussians with a certain overall amplitude to only the evaluation dataset, for which the NN performance stayed consistent.

However, the work in this thesis is just the first step to using a NN-based approach for the analysis of measurement data from the TRISTAN detector upgrade of KATRIN. Further development is necessary in three main areas.

First, the experimental effects still missing must be incorporated. So far, only general shape distortions and detector related effects have been taken into account. Significant experimental effects inflicted with large systematic uncertainties like the electron backscattering at the rear wall of KATRIN have yet to be addressed. The NN

then needs to be trained with all experimental effects enabled and evaluated with all systematic uncertainties taken into account, which could affect the performance heavily and require larger datasets and training budgets. After this step, the NN could be used as a flag for false positives / false negatives accompanying the traditional analysis.

Second, sterile neutrino parameter regression must be implemented. Currently, the NNs perform only BC, categorizing input spectra as either containing or not containing a sterile neutrino signature. To achieve a fair comparison with traditional parametric fitting methods, the NNs must be extended to infer the sterile neutrino mass (m_4) and the sterile-to-active mixing angle ($\sin^2(\theta)$) for a given input.

Third, the open questions from this work need to be addressed. For example, the positive impact of including the post-acceleration electrode (PAE) response on NN performance requires further investigation. Preliminary insights into this effect, derived using saliency maps of the NNs, have been developed in this thesis, but further studies are required to confirm those findings.

Chapter 7

List of Acronyms

DM Dark Matter

CMB Cosmic Microwave Background

WIMP Weakly Interacting Massive Particle

SM Standard Model of Particle Physics

KATRIN KArlsruhe TRItium Neutrino

GWS Glashow-Weinberg-Salam

PMNS Pontecorvo-Maki-Nakagawa-Sakata

SSB Spontaneous Symmetry Breaking

CMB Cosmic Microwave Background

HDM Hot Dark Matter

WDM Warm Dark Matter

CDM Cold Dark Matter

FPD Focal Plane Detector

SDD Silicon Drift Detector

WGTS Windowless Gaseous Tritium Source

RW Rear Wall.

MS Main Spectrometer.

PS Pre-Spectrometer

CMS Calibration and Monitoring System

DPS Differential Pumping Section

CPS Cryogenic Pumping Section

MAC-E Magnetic Adiabatic Collimation with Electrostatic filtering

PAE Post-acceleration Electrode

DAQ data aquisition

MC Monte Carlo

NN Neural Network

ML Machine Learning

BC Binary Classification

MLP Multilayer Perceptron

PMLP Parametrized Multilayer Perceptron

CNN Convolutional Neural Network

ResNet Residual Neural Network

p.d.f. probability density function

CL confidence level

Appendix A

Neural Network Hyperparameters

The values of the hyperparameters were derived based on the results from the Bayesian Hyperparameter Optimization tool SMAC3 [106]. The F1-Score on the validation dataset was used as the optimization target. 20 training epochs were used as the maximum training budget. The duration of one hyperparameter optimization process was set to 16 hours, regardless of the model size. This could have impacted the results for especially the deeper architectures, with longer runs being planned for the future. For studies in sections 5.2, 5.3.1 and 5.3.2, the maximum training budget was adjusted, but the other hyperparameters were kept the same.

Model Architecture	Training Epochs	Batch Size	Initial Learning Rate	Weight Decay
(P)MLP	20	25000	10^{-3}	10^{-6}
CNN	20	1000	5×10^{-5}	10^{-6}
ResNet - MLP	20	128	3×10^{-3}	8×10^{-7}
ResNet - CNN	20	6000	2×10^{-3}	-

Table A.1: Training-related hyperparameters for all models shown in figure 5.1.

Model Type	CNN	(P)MLP
Hidden Layers	3	2
Ch. / Lin. Dim.	[128, 256, 128]	[256, 256]
Kernel Sizes	[16, 16, 16]	-
Batch Norm Layers	[✓, ✓, ✓]	[✓, ✓,]
Dropout Prob.	-	–

Table A.2: Architecture-related hyperparameters for the MLP and CNN shown in figure 5.1.

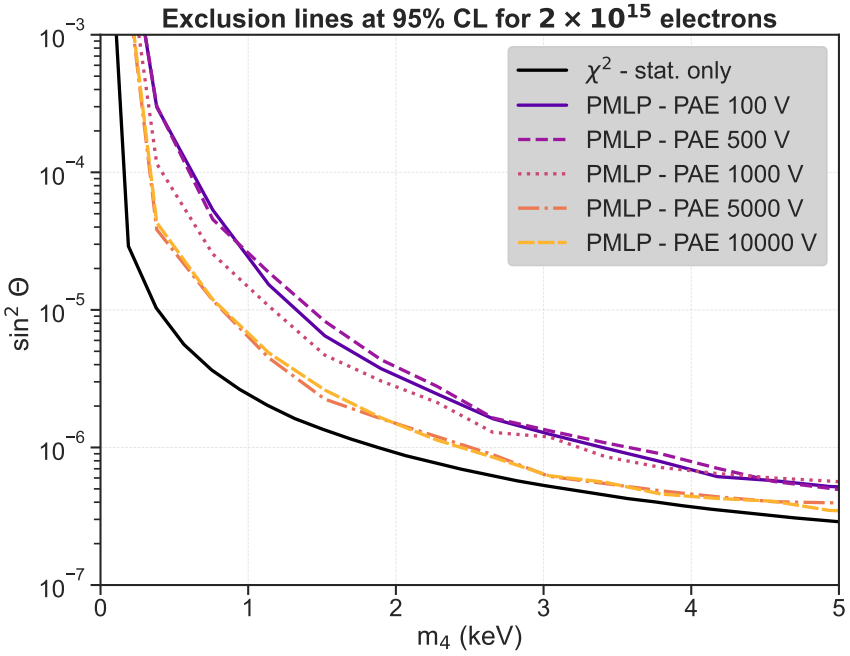
ResNet Block Type	CNN	MLP
Block Length	3	3
Ch. / Lin. Dim.	$[137, 61, 96]$	$[n_{\text{bins}}, 128, 128]$
Kernel Sizes	$[1, 5, 7]$	-
Batch Norm Layers	[✓, ✓, ✓]	[✓, ✓, ✓]
Dropout Prob.	-	$[0.5, 0.5, 0.5]$
Number of Blocks	3	3

Table A.3: Architecture-related hyperparameters for the ResNet MLP and ResNet CNN shown in figure 5.1.

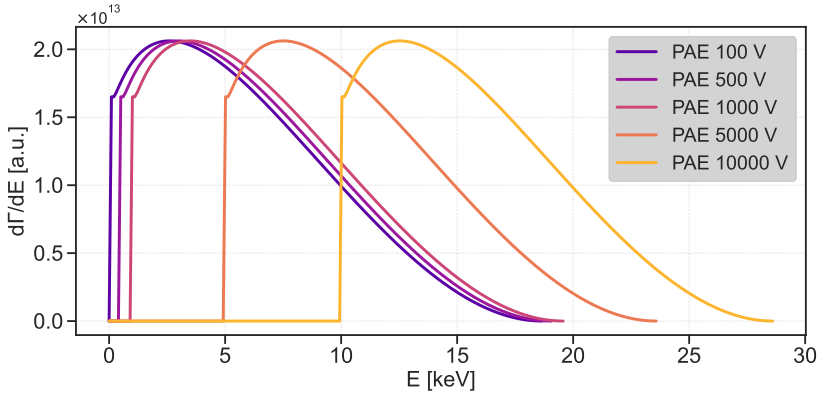
Appendix B

Post Acceleration Electrode Response Investigations

This chapter discusses the investigation regarding the effect of the PAE presented in section 5.2 in more detail. In particular, the effect of adding a PAE response without also expanding the energy range further than the shifted endpoint energy (figure B.1), as well as adding no PAE response and just extending the energy range beyond the endpoint (figure B.2) is investigated. This results in appending entries with a value of zero either to the beginning or the end of the statistically fluctuated spectrum. The model performance gradually improves for both cases, but saturates after a certain value. This effect has to be investigated further to reach a conclusive explanation.

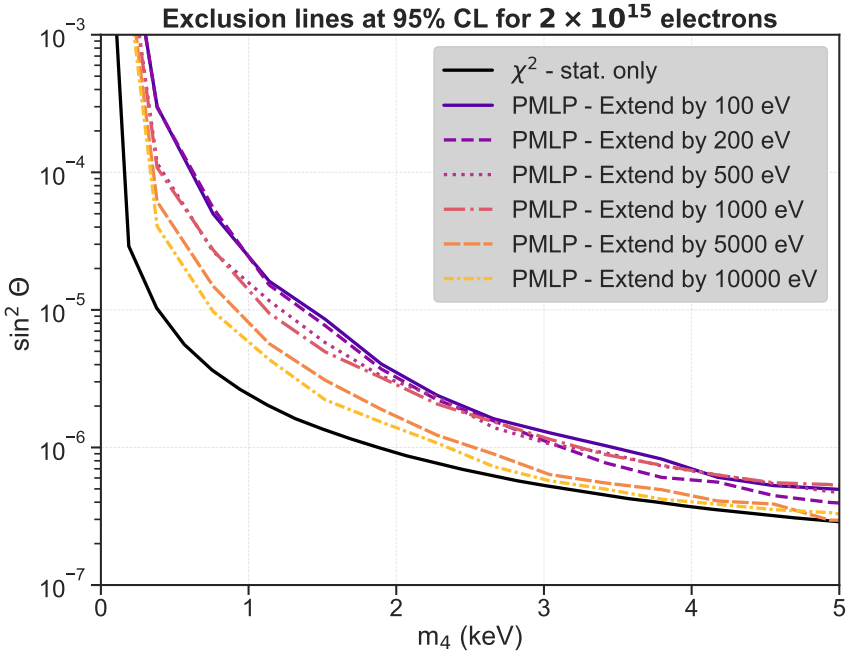


(a) Sensitivity contours for different PAE responses without increasing the energy range beyond the shifted kinematic endpoint of the tritium β-spectrum.

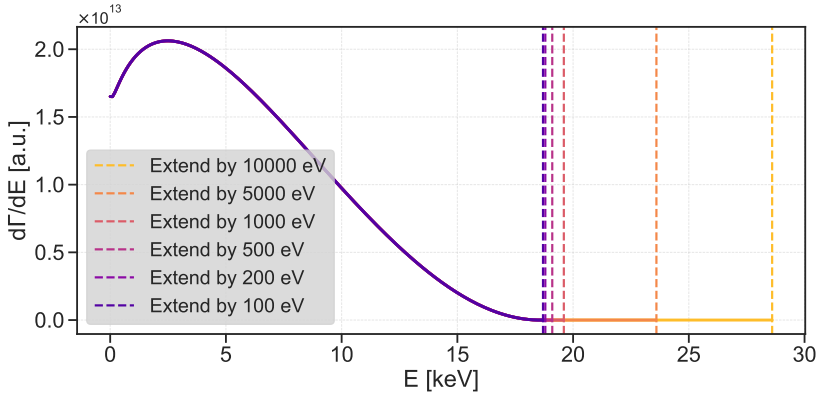


(b) Illustration of the effect of a PAE response without an additional energy range increase beyond the shifted kinematic endpoint on the data.

Figure B.1: Increasing the PAE to larger values increases the PMLPs sensitivity to smaller sterile masses significantly. This effect saturates at a PAE voltage of 5 kV.



(a) Impact on the sensitivity of extending the range of interest beyond the kinematic endpoint of the tritium β-spectrum.

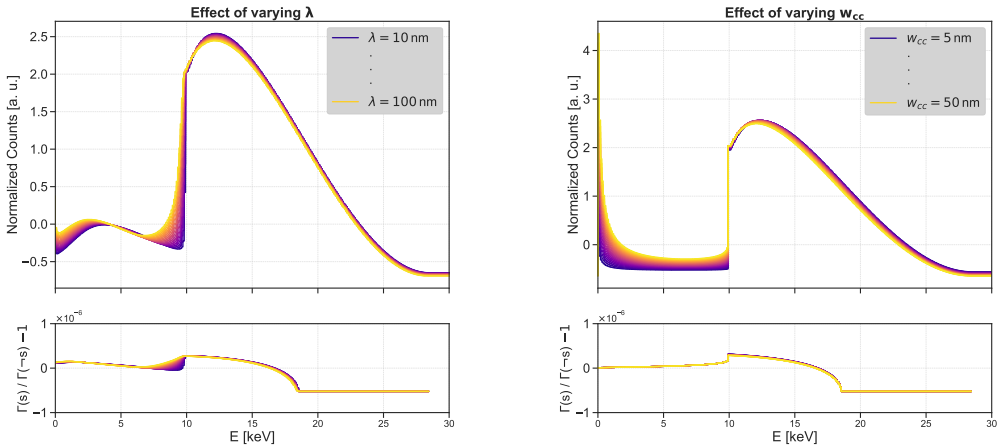


(b) Illustration of the effect of extending the range of interest beyond the kinematic endpoint on the data.

Figure B.2: Increasing the energy range increases the PMLPs sensitivity to smaller sterile masses significantly. This effect saturates after an increase of 10 keV.

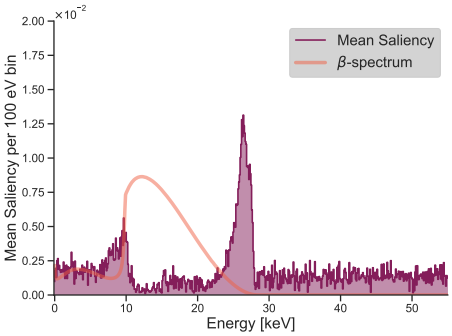
Appendix C

Supplementary Plots

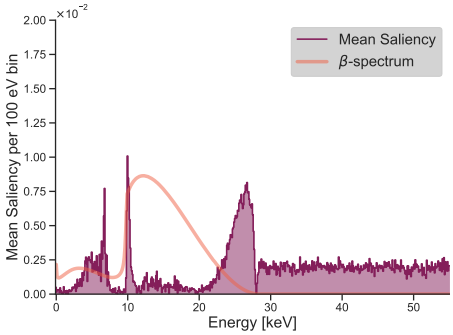


(a) Varying the dead layer parameter λ. **(b)** Varying the charge cloud width w_{cc}

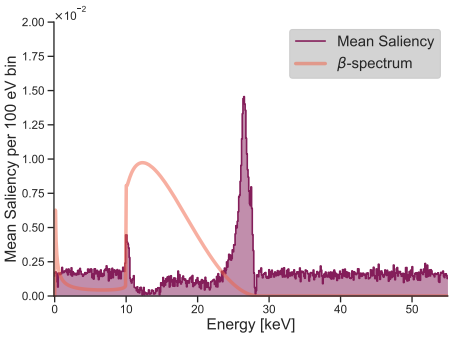
Figure C.1: Illustration of the impact of the dead layer parameter λ and the charge cloud width w_{cc} on the β-decay spectrum after feature standardization (upper part), as well as the effect on the sterile neutrino signature (lower part).



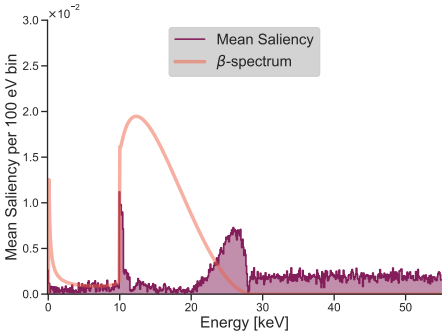
(a) Training and evaluation on data with fixed $\lambda = 58\,\text{nm}$



(b) Training on data with $\lambda \sim \mathcal{N}(\mu = 58\,\text{nm}, \sigma = 2\,\text{nm})$, evaluation on fixed $\lambda = 58\,\text{nm}$.



(c) Training and evaluation on data with fixed $w_{cc} = 20\,\text{nm}$



(d) Training on data with $w_{cc} \sim \mathcal{N}(\mu = 20\,\text{nm}, \sigma = 1\,\text{nm})$, evaluation on fixed $w_{cc} = 20\,\text{nm}$.

Figure C.2: Mean saliency maps for various scenarios.

Bibliography

[1] F. Zwicky. »Die Rotverschiebung von extragalaktischen Nebeln.« In: *Helv. Phys. Acta* 6 (1933), pp. 110–127. DOI: 10.1007/s10714-008-0707-4.

[2] M. Tristram et al. »Cosmological parameters derived from the final Planck data release (PR4).« In: *Astronomy; Astrophysics* 682 (Jan. 2024), A37. ISSN: 1432-0746. DOI: 10.1051/0004-6361/202348015. URL: http://dx.doi.org/10.1051/0004-6361/202348015.

[3] Scott W. Randall et al. »Constraints on the Self Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657 56.« In: *The Astrophysical Journal* 679.2 (June 2008), pp. 1173–1180. ISSN: 1538-4357. DOI: 10.1086/587859. URL: http://dx.doi.org/10.1086/587859.

[4] Gianfranco Bertone, Dan Hooper, and Joseph Silk. »Particle dark matter: evidence, candidates and constraints.« In: *Physics Reports* 405.5–6 (Jan. 2005), pp. 279–390. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2004.08.031. URL: http://dx.doi.org/10.1016/j.physrep.2004.08.031.

[5] A. Boyarsky et al. »Sterile neutrino Dark Matter.« In: *Progress in Particle and Nuclear Physics* 104 (Jan. 2019), pp. 1–45. ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2018.07.004. URL: http://dx.doi.org/10.1016/j.ppnp.2018.07.004.

[6] Basudeb Dasgupta and Joachim Kopp. »Sterile neutrinos.« In: *Physics Reports* 928 (Sept. 2021), pp. 1–63. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2021.06.002. URL: http://dx.doi.org/10.1016/j.physrep.2021.06.002.

[7] M. Aker et al. »The design, construction, and commissioning of the KATRIN experiment.« In: *Journal of Instrumentation* 16.08 (Aug. 2021), T08015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/08/t08015. URL: http://dx.doi.org/10.1088/1748-0221/16/08/T08015.

[8] Dan Guest, Kyle Cranmer, and Daniel Whiteson. »Deep Learning and Its Application to LHC Physics.« In: *Annual Review of Nuclear and Particle Science* 68.1 (Oct. 2018), pp. 161–181. ISSN: 1545-4134. DOI: 10.1146/annurev-nucl-101917-021019. URL: http://dx.doi.org/10.1146/annurev-nucl-101917-021019.

[9] Giuseppe Carleo et al. »Machine learning and the physical sciences.« In: *Reviews of Modern Physics* 91.4 (Dec. 2019). ISSN: 1539-0756. DOI: 10.1103/revmodphys.91.045002. URL: http://dx.doi.org/10.1103/RevModPhys.91.045002.

[10] J. Chadwick. »Intensitätsverteilung im magnetischen Spectrum der β-Strahlen von radium B + C.« In: *Verhandl. Dtsc. Phys. Ges.* 16 (1914), p. 383. URL: https://cds.cern.ch/record/262756.

[11] Charles Drummond Ellis and W. A Wooster. »The average energy of disintegration of radium E.« In: *Proc. Roy. Soc. Lond. A* 117.776 (1927), pp. 109–123. DOI: 10.1098/rspa.1927.0168.

[12] Lise Meitner and Wilhelm Orthmann. »Über eine absolute Bestimmung der Energie der primären β-Strahlen von Radium E.« In: *Zeitschrift für Physik* 60 (1930), pp. 143–155. URL: https://api.semanticscholar.org/CorpusID:121406618.

[13] Wolfgang Pauli. »Pauli letter collection: letter to Lise Meitner.« Typed copy. 1930. URL: https://cds.cern.ch/record/83282.

[14] E. Fermi. »An attempt of a theory of beta radiation. 1.« In: *Z. Phys.* 88 (1934), pp. 161–177. DOI: 10.1007/BF01351864.

[15] M. Descher. »Differential spectrum modeling and sensitivity for keV sterile neutrino search at KATRIN.« PhD thesis. KIT Faculty of physics Karlsruhe, 2023.

[16] C. S. Wu et al. »Experimental Test of Parity Conservation in Beta Decay.« In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[17] D. Griffiths. *Introduction to Elementary Particles*. 2nd ed. Wiley-VCH, 2008.

[18] M. Goldhaber, L. Grodzins, and A. W. Sunyar. »Helicity of Neutrinos.« In: *Phys. Rev.* 109 (3 Feb. 1958), pp. 1015–1017. DOI: 10.1103/PhysRev.109.1015. URL: https://link.aps.org/doi/10.1103/PhysRev.109.1015.

[19] E. C. G. Sudarshan and R. E. Marshak. »Chirality Invariance and the Universal Fermi Interaction.« In: *Phys. Rev.* 109 (5 Mar. 1958), pp. 1860–1862. DOI: 10.1103/PhysRev.109.1860.2. URL: https://link.aps.org/doi/10.1103/PhysRev.109.1860.2.

[20] R. P. Feynman and M. Gell-Mann. »Theory of the Fermi Interaction.« In: *Phys. Rev.* 109 (1 Jan. 1958), pp. 193–198. DOI: 10.1103/PhysRev.109.193. URL: https://link.aps.org/doi/10.1103/PhysRev.109.193.

[21] F. Reines and C. L. Cowan. »Detection of the Free Neutrino.« In: *Phys. Rev.* 92 (3 Nov. 1953), pp. 830–831. DOI: 10.1103/PhysRev.92.830. URL: https://link.aps.org/doi/10.1103/PhysRev.92.830.

[22] C. L. Cowan et al. »Detection of the Free Neutrino: a Confirmation.« In: *Science* 124.3212 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103. eprint: https://www.science.org/doi/pdf/10.1126/science.124.3212.103. URL: https://www.science.org/doi/abs/10.1126/science.124.3212.103.

[23] F. Reines et al. »Detection of the Free Antineutrino.« In: *Phys. Rev.* 117 (1 Jan. 1960), pp. 159–173. DOI: 10.1103/PhysRev.117.159. URL: https://link.aps.org/doi/10.1103/PhysRev.117.159.

[24] G. Danby et al. »Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos.« In: *Phys. Rev. Lett.* 9 (1 July 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[25] M. L. Perl et al. »Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation.« In: *Phys. Rev. Lett.* 35 (22 Dec. 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[26] K. Kodama et al. »Observation of tau neutrino interactions.« In: *Phys. Lett. B* 504 (2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0. arXiv: hep-ex/0012035.

[27] J. A. Bagger (editor). »Precision electroweak measurements on the Z resonance.« In: *Physics Reports* 427.5 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006.

[28] H. A. Bethe and C. L. Critchfield. »The Formation of Deuterons by Proton Combination.« In: *Phys. Rev.* 54 (4 Aug. 1938), pp. 248–254. DOI: 10.1103/PhysRev.54.248. URL: https://link.aps.org/doi/10.1103/PhysRev.54.248.

[29] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. »Search for Neutrinos from the Sun.« In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[30] John N. Bahcall, Neta A. Bahcall, and Giora Shaviv. »Present Status of the Theoretical Predictions for the ^{37}Cl Solar-Neutrino Experiment.« In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1209–1212. DOI: 10.1103/PhysRevLett.20.1209. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1209.

[31] M. Cribier. »Results of the whole GALLEX experiment.« In: *Nucl. Phys. B Proc. Suppl.* 70 (1999). Ed. by A. Bottino, Alessandra Di Credico, and Piero Monacelli, pp. 284–291. DOI: 10.1016/S0920-5632(98)00438-1.

[32] J. N. Abdurashitov et al. »Measurement of the solar neutrino capture rate in SAGE.« In: *Nucl. Phys. B Proc. Suppl.* 118 (2003). Ed. by F. von Feilitzsch and N. Schmitz, pp. 39–46. DOI: 10.1016/S0920-5632(03)01302-1.

[33] C. Arpesella et al. »Direct Measurement of the ^{7}Be Solar Neutrino Flux with 192 Days of Borexino Data.« In: *Phys. Rev. Lett.* 101 (9 Aug. 2008), p. 091302. DOI: 10.1103/PhysRevLett.101.091302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.101.091302.

[34] Q. R. Ahmad et al. »Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory.« In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.011301.

[35] Y. Fukuda et al. »Evidence for Oscillation of Atmospheric Neutrinos.« In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[36] F. P. An et al. »Observation of Electron-Antineutrino Disappearance at Daya Bay.« In: *Phys. Rev. Lett.* 108 (17 Apr. 2012), p. 171803. DOI: 10.1103/PhysRevLett.108.171803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.171803.

[37] K. Abe et al. »Observation of Electron Neutrino Appearance in a Muon Neutrino Beam.« In: *Phys. Rev. Lett.* 112 (6 Feb. 2014), p. 061802. DOI: 10.1103/PhysRevLett.112.061802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.061802.

[38] K. Zuber. *Neutrino Physics*. 3rd. ed. CRC Press, 2020.

[39] K. Abe et al. »Solar neutrino results in Super-Kamiokande-III.« In: *Phys. Rev. D* 83 (5 Mar. 2011), p. 052010. DOI: 10.1103/PhysRevD.83.052010. URL: https://link.aps.org/doi/10.1103/PhysRevD.83.052010.

[40] B. Aharmim et al. »Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory.« In: *Phys. Rev. C* 88 (2 Aug. 2013), p. 025501. DOI: 10.1103/PhysRevC.88.025501. URL: https://link.aps.org/doi/10.1103/PhysRevC.88.025501.

[41] M. G. Aartsen et al. »Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data.« In: *Phys. Rev. D* 91 (7 Apr. 2015), p. 072004. DOI: 10.1103/PhysRevD.91.072004. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.072004.

[42] K. Abe et al. »Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV.« In: *Phys. Rev. D* 97 (7 Apr. 2018), p. 072001. DOI: 10.1103/PhysRevD.97.072001. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.072001.

[43] D. H. Perkins. *Particle Astrophysics*. 2nd ed. Oxford University Press, 2008.

[44] E. Noether. »Invariante Variationsprobleme.« ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: http://eudml.org/doc/59024.

[45] Peter W. Higgs. »Broken Symmetries and the Masses of Gauge Bosons.« In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[46] F. Englert and R. Brout. »Broken Symmetry and the Mass of Gauge Vector Mesons.« In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[47] T. W. B. Kibble. »Symmetry Breaking in Non-Abelian Gauge Theories.« In: *Phys. Rev.* 155 (5 Mar. 1967), pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554. URL: https://link.aps.org/doi/10.1103/PhysRev.155.1554.

[48] S. L. Glashow. »The renormalizability of vector meson interactions.« In: *Nuclear Physics* 10 (1959), pp. 107–117. DOI: 10.1016/0029-5582(59)90196-8.

[49] A. Salam and J. C. Ward. »Weak and electromagnetic interactions.« In: *Nuovo Cim* 11.4 (1959), pp. 568–577. DOI: 10.1007/BF02726525.

[50] Steven Weinberg. »A Model of Leptons.« In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[51] Abdus Salam. »Gauge unification of fundamental forces.« In: *Rev. Mod. Phys.* 52 (3 July 1980), pp. 525–538. DOI: 10.1103/RevModPhys.52.525. URL: https://link.aps.org/doi/10.1103/RevModPhys.52.525.

[52] Serhat Ordek, the ATLAS, and On behalf of CMS Collaborations. »Measurements of Higgs couplings to fermions and bosons at the LHC.« In: *PoS* PANIC2021 (2022), p. 418. DOI: 10.22323/1.380.0418.

[53] M. Aker et al. *Direct neutrino-mass measurement based on 259 days of KATRIN data*. 2024. arXiv: 2406.13516 [nucl-ex]. URL: https://arxiv.org/abs/2406.13516.

[54] Valerie Domcke et al. »MeV-scale seesaw and leptogenesis.« In: *Journal of High Energy Physics* 2021.1 (Jan. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep01(2021)200. URL: http://dx.doi.org/10.1007/JHEP01(2021)200.

[55] Tsutomu Yanagida. »Horizontal Symmetry and Masses of Neutrinos.« In: *Prog. Theor. Phys.* 64 (1980), p. 1103. DOI: 10.1143/PTP.64.1103.

[56] R N Mohapatra et al. »Theory of neutrinos: a white paper.« In: *Reports on Progress in Physics* 70.11 (2007), pp. 1757–1867. ISSN: 1361-6633. DOI: 10.1088/0034-4885/70/11/r02. URL: http://dx.doi.org/10.1088/0034-4885/70/11/R02.

[57] G Drexlin et al. »Current Direct Neutrino Mass Experiments.« In: *Advances in High Energy Physics* (2013), pp. 1–39. DOI: 10.1155/2013/293986.

[58] M. Agostini et al. »Final Results of GERDA on the Search for Neutrinoless Double-β Decay.« In: *Phys. Rev. Lett.* 125 (25 Dec. 2020), p. 252502. DOI: 10.1103/PhysRevLett.125.252502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.252502.

[59] Laurent Canetti and Mikhail Shaposhnikov. »Baryon asymmetry of the Universe in the νMSM.« In: *Journal of Cosmology and Astroparticle Physics* 2010.09 (Sept. 2010), pp. 001–001. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2010/09/001. URL: http://dx.doi.org/10.1088/1475-7516/2010/09/001.

[60] K. N. Abazajian et al. *Light Sterile Neutrinos: A White Paper*. 2012. arXiv: 1204.5379 [hep-ph]. URL: https://arxiv.org/abs/1204.5379.

[61] Qirong Zhu et al. »Baryonic impact on the dark matter distribution in Milky Way-sized galaxies and their satellites.« In: *Monthly Notices of the Royal Astronomical Society* 458.2 (Feb. 2016), pp. 1559–1580. ISSN: 1365-2966. DOI: 10.1093/mnras/stw374. URL: http://dx.doi.org/10.1093/mnras/stw374.

[62] R. Murgia et al. »"Non-cold" dark matter at small scales: a general approach.« In: *Journal of Cosmology and Astroparticle Physics* 2017.11 (Nov. 2017), pp. 046–046. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/11/046. URL: http://dx.doi.org/10.1088/1475-7516/2017/11/046.

[63] Douglas Clowe et al. »A Direct Empirical Proof of the Existence of Dark Matter.« In: *The Astrophysical Journal* 648.2 (Aug. 2006), pp. L109–L113. ISSN: 1538-4357. DOI: 10.1086/508162. URL: http://dx.doi.org/10.1086/508162.

[64] Kevork Abazajian, George M. Fuller, and Mitesh Patel. »Sterile neutrino hot, warm, and cold dark matter.« In: *Phys. Rev. D* 64 (2 May 2001), p. 023501. DOI: 10.1103/PhysRevD.64.023501. URL: https://link.aps.org/doi/10.1103/PhysRevD.64.023501.

[65] R. Adhikari et al. »A White Paper on keV sterile neutrino Dark Matter.« In: *Journal of Cosmology and Astroparticle Physics* 2017.01 (Jan. 2017), pp. 025–025. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2017/01/025. URL: http://dx.doi.org/10.1088/1475-7516/2017/01/025.

[66] Kevork N. Abazajian. »Sterile neutrinos in cosmology.« In: *Physics Reports* 711–712 (Nov. 2017), pp. 1–28. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2017.10.003. URL: http://dx.doi.org/10.1016/j.physrep.2017.10.003.

[67] S Bhargava et al. »The XMM Cluster Survey: new evidence for the 3.5-keV feature in clusters is inconsistent with a dark matter origin.« In: *Monthly Notices of the Royal Astronomical Society* 497.1 (June 2020), pp. 656–671. ISSN: 1365-2966. DOI: 10.1093/mnras/staa1829. URL: http://dx.doi.org/10.1093/mnras/staa1829.

[68] Susanne Mertens et al. »A novel detector system for KATRIN to search for keV-scale sterile neutrinos.« In: *Journal of Physics G: Nuclear and Particle Physics* 46.6 (May 2019), p. 065203. ISSN: 1361-6471. DOI: 10.1088/1361-6471/ab12fe. URL: http://dx.doi.org/10.1088/1361-6471/ab12fe.

[69] M. Kleesiek et al. »β-Decay spectrum, response function and statistical model for neutrino mass measurements with the KATRIN experiment.« In: *The European Physical Journal C* 79.3 (Mar. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6686-7. URL: http://dx.doi.org/10.1140/epjc/s10052-019-6686-7.

[70] E W Otten and C Weinheimer. »Neutrino mass limit from tritium β decay.« In: *Reports on Progress in Physics* 71.8 (July 2008), p. 086201. ISSN: 1361-6633. DOI: 10.1088/0034-4885/71/8/086201. URL: http://dx.doi.org/10.1088/0034-4885/71/8/086201.

[71] S. Navas et al. »Review of Particle Physics.« In: *Phys. Rev. D* 110 (3 Aug. 2024), p. 030001. DOI: 10.1103/PhysRevD.110.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.110.030001.

[72] M. Aker et al. *First direct neutrino-mass measurement with sub-eV sensitivity.* 2021. arXiv: 2105.08533 [hep-ex]. URL: https://arxiv.org/abs/2105.08533.

[73] The KATRIN collaboration. »Direct neutrino-mass measurement with sub-electronvolt sensitivity.« In: *Nat. Phys* 18.2 (Feb. 2022), pp. 160–166.

[74] Florian Priester, Michael Sturm, and Beate Bornschein. »Commissioning and detailed results of KATRIN inner loop tritium processing system at Tritium Laboratory Karlsruhe.« In: *Vacuum* 116 (2015), pp. 42–47. DOI: 10.1016/j.vacuum.2015.02.030.

[75] A. Picard et al. »A solenoid retarding spectrometer with high resolution and transmission for keV electrons.« In: *Nucl. Instrum. Meth. B* 63.3 (1992), pp. 345–358. DOI: 10.1016/0168-583X(92)95119-C.

[76] V. M. Lobashev and P. E. Spivak. »A METHOD FOR MEASURING THE ANTI-ELECTRON-NEUTRINO REST MASS.« In: *Nucl. Instrum. Meth. A* 240 (1985), pp. 305–310. DOI: 10.1016/0168-9002(85)90640-0.

[77] M. Aker et al. *Search for keV-scale Sterile Neutrinos with first KATRIN Data.* 2022. arXiv: 2207.06337 [nucl-ex]. URL: https://arxiv.org/abs/2207.06337.

[78] J.F. Amsbaugh et al. »Focal-plane detector system for the KATRIN experiment.« In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 778 (Apr. 2015), pp. 40–60. ISSN: 0168-9002. DOI: 10.1016/j.nima.2014.12.116. URL: http://dx.doi.org/10.1016/j.nima.2014.12.116.

[79] Daniel Siegmann et al. »Development of a silicon drift detector array to search for keV-scale sterile neutrinos with the KATRIN experiment.« In: *J. Phys. G* 51.8 (2024), p. 085202. DOI: 10.1088/1361-6471/ad4bf8. arXiv: 2401.14114 [physics.ins-det].

[80] TRISTAN Group. »Conceptual Design Report: KATRIN with TRISTAN modules.« In: 2022. URL: https://www.katrin.kit.edu/downloads/TRISTAN__Technical_Design_Report%20(10).pdf.

[81] C. Bruch and S. Mertens. »Characterisation of a 166-Pixel TRISTAN Detector Module in the KATRIN Monitor Spectrometer.« Master's Thesis. Technical University of Munich, 2023.

[82] A. Huber. »Analysis of first KATRIN data and searches for keV-scale sterile neutrinos.« PhD thesis. KIT Faculty of physics Karlsruhe, 2021.

[83] S. Mertens et al. »Sensitivity of next-generation tritium beta-decay experiments for keV-scale sterile neutrinos.« In: *Journal of Cosmology and Astroparticle Physics* 2015.02 (Feb. 2015), pp. 020–020. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/02/020. URL: http://dx.doi.org/10.1088/1475-7516/2015/02/020.

[84] Anthony Onillon. »Neutrino mass measurement and sterile neutrinos search with the KATRIN experiment.« In: *EPJ Web of Conferences* 282 (2023). DOI: 10.1051/epjconf/202328201011. URL: https://doi.org/10.1051/epjconf/202328201011.

[85] C. J. Martoff et al. »HUNTER: precision massive-neutrino search based on a laser cooled atomic source.« In: *Quantum Sci. Technol.* 6.2 (2021), p. 024008. DOI: 10.1088/2058-9565/abdb9b.

[86] Rear Wall Working Group. *Rear wall residual tritium spectrum wiki page*. 2024. URL: https://iap-katrin-wiki.iap.kit.edu/katrin/index.php/RW_residual_tritium_spectrum.

[87] P. vogt and S. Mertens. »Transmission properties of the KATRIN main spectrometer (when operated) for a keV-scale sterile neutrino search.« Master's Thesis. Technical University of Munich, 2023.

[88] S Mertens et al. »Characterization of silicon drift detectors with electrons for the TRISTAN project.« In: *Journal of Physics G: Nuclear and Particle Physics* 48.1 (Dec. 2020), p. 015008. ISSN: 1361-6471. DOI: 10.1088/1361-6471/abc2dc. URL: http://dx.doi.org/10.1088/1361-6471/abc2dc.

[89] Daniela Spreng et al. *Investigation of Electron Backscattering on Silicon Drift Detectors for the Sterile Neutrino Search with TRISTAN*. 2024. arXiv: 2405.12776 [physics.ins-det]. URL: https://arxiv.org/abs/2405.12776.

[90] C. Forstner and S. Mertens. »Characterization of a TRISTAN Sil- icon Drift Detector Array with a Laser System.« Master's Thesis. Technical University of Munich, 2023.

[91] Matteo Gugiatti. »Development of a large-Matrix SDD-based radiation detector for beta-decay spectroscopy in neutrino physics.« PhD thesis. Politecnico di Milano, 2022.

[92] U. Fano. »Ionization Yield of Radiations. II. The Fluctuations of the Number of Ions.« In: *Phys. Rev.* 72 (1 July 1947), pp. 26–29. DOI: 10.1103/PhysRev.72.26. URL: https://link.aps.org/doi/10.1103/PhysRev.72.26.

[93] Martin Descher et al. *TRModel*. Version 1.0.0. 2023. URL: https://nuserv.uni-muenster.de:8443/katrin-git/tristan/trmodel.

[94] G. Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.

[95] S. Kullback and R. A. Leibler. »On Information and Sufficiency.« In: *The Annals of Mathematical Statistics* 22.1 (1951), pp. 79–86. ISSN: 00034851. URL: http://www.jstor.org/stable/2236703 (visited on 10/28/2024).

[96] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization.* 2017. arXiv: 1412.6980 [cs.LG]. URL: https://arxiv.org/abs/1412.6980.

[97] Jürgen Schmidhuber. »Deep learning in neural networks: An overview.« In: *Neural Networks* 61 (Jan. 2015), pp. 85–117. ISSN: 0893-6080. DOI: 10.1016/j.neunet.2014.09.003. URL: http://dx.doi.org/10.1016/j.neunet.2014.09.003.

[98] Grant Sanderson. *What is backpropagation really doing?* 2017. URL: https://www.youtube.com/watch?v=Ilg3gGewQ5U.

[99] R.L. Workman et al. *Review of Particle Physics: 2022.* Vol. 2022. Notes. Oxford: Oxford University Press, 2022. DOI: 10.1093/ptep/ptac097. URL: https://cds.cern.ch/record/2836514.

[100] Halbert White Kurt Hornik Maxwell Stinchcombe. »Multilayer feedforward networks are universal approximators.« In: *Neural Networks* 2.5 (1989), pp. 359–366. ISSN: 0893-6080. DOI: 10.1016/0893-6080(89)90020-8. URL: https://doi.org/10.1016/0893-6080(89)90020-8.

[101] Hassan Ismail Fawaz et al. »Deep learning for time series classification: a review.« In: *Data Mining and Knowledge Discovery* 33.4 (Mar. 2019), pp. 917–963. ISSN: 1573-756X. DOI: 10.1007/s10618-019-00619-1. URL: http://dx.doi.org/10.1007/s10618-019-00619-1.

[102] Sergey Ioffe and Christian Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* 2015. arXiv: 1502.03167 [cs.LG]. URL: https://arxiv.org/abs/1502.03167.

[103] Geoffrey E. Hinton et al. *Improving neural networks by preventing co-adaptation of feature detectors.* 2012. arXiv: 1207.0580 [cs.NE]. URL: https://arxiv.org/abs/1207.0580.

[104] J. Bergstra and Y. Bengio. »Random Search for Hyperparameter Optimization.« In: *Journal of Machine Learning Research* 13 (2012). URL: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf.

[105] Stefan Falkner, Aaron Klein, and Frank Hutter. *BOHB: Robust and Efficient Hyperparameter Optimization at Scale.* 2018. arXiv: 1807.01774 [cs.LG]. URL: https://arxiv.org/abs/1807.01774.

[106] Marius Lindauer et al. »SMAC3: A Versatile Bayesian Optimization Package for Hyperparameter Optimization.« In: *Journal of Machine Learning Research* 23.54 (2022), pp. 1–9. URL: http://jmlr.org/papers/v23/21-0888.html.

[107] Christian Robert Karl. »First Sub-Electronvolt Direct Neutrino Mass Measurement with the KATRIN Experiment.« PhD thesis. Munich, Tech. U., 2022.

[108] Volker Blobel and Erich Lohrmann. *Statistische und numerische Methoden der Datenanalyse*. Creative Commons, 2012. URL: https://www.desy.de/~sschmitt/blobel/ebuch.html.

[109] Jerzy Neyman and Egon Sharpe Pearson. »Ix. on the problem of the most efficient tests of statistical hypotheses.« In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231 (694-706 1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009.

[110] Adam Paszke et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. 2019. arXiv: 1912.01703 [cs.LG]. URL: https://arxiv.org/abs/1912.01703.

[111] Apple. *Apple Metal Performance Shaders for PyTorch*. 2024. URL: https://developer.apple.com/metal/pytorch/.

[112] Nvidia. *Nvidia Compute Unified Device Architecture (CUDA) for PyTorch*. 2024. URL: https://developer.apple.com/metal/pytorch/.

[113] Luca Fallböhmer. 2024. URL: https://github.com/lucafally/masterthesis.git.