ETP-KA/2019-06

Karlsruhe Institute of Technology

High rate systematic effects for keV sterile neutrino searches at KATRIN

Master Thesis

Martin Descher

Department of Physics
Institute of Experimental Particle Physics

Reviewer:	Prof. Dr. Guido Drexlin
Second reviewer:	Prof. Dr. Marc Weber
Advisor:	Marc Korzeczek

Karlsruhe, 14. June 2019

Erklärung der selbständigen Anfertigung der Masterarbeit

Hiermit versichere ich wahrheitsgemäß, die Masterarbeit mit dem Titel

»*High rate systematic effects for keV sterile neutrino searches at KATRIN*«

selbstständig verfasst, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde sowie die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Martin Descher
Karlsruhe, den 14. Juni 2019

Als Prüfungsexemplar genehmigt von

Prof. Dr. G. Drexlin
Karlsruhe, den 14. Juni 2019

Abstract

The TRISTAN project of the KATRIN experiment aims to search for the signature of keV sterile neutrinos in the β-decay spectrum of tritium. For a three year measurement of the differential β-electron energy spectrum at an event rate of 10^8 counts per second (cps) a sensitivity for sterile mixing angles $\sin^2\theta < 10^{-6}$, within the accessible sterile neutrino mass range, is projected.

In order to achieve this goal, an upgrade of the KATRIN detector and data acquisition system is necessary, as it currently cannot provide the required energy resolution and high rate capabilities for a sufficiently precise measurement of the differential tritium spectrum. A silicon drift detector pixel array with ~ 3500 pixels is envisioned for accomplishing this task. Furthermore, a data acquisition system (DAQ) capable of handling up to 10^5 cps per channel is essential.

At these high event rates, the rate dependent systematic effects arising from data acquisition reach a magnitude where they require thorough modeling. In this work, the systematic effects of dead time and signal pileup are investigated. This includes a detailed study of their origin, the development and verification of analytical and simulation-based modeling approaches, as well as the discussion of correction and mitigation methods. For modeling the shape of the pileup spectrum a convolution approach is proposed and verified with simulated data. Based on this model the impact of pileup on the final sensitivity for keV sterile neutrinos is estimated. The study yields a sensitivity reduction of a factor ~ 0.5 in the most sensitive sterile neutrino mass range. Moreover, it is demonstrated that almost all sensitivity can be recovered through post acceleration of the β-electrons before they hit the detector.

Additionally, the DAQ system used for data readout with the current 7-pixel TRISTAN prototype detector is characterized with regard to its high rate capabilities. A detector emulator is employed for studying pileup and the associated minimum resolution time for resolving successive events. The minimum resolution time is found to be sufficiently low to ensure a pileup event fraction below 1 % for the expected rate per pixel. Dead time and its connection to input rate estimation are investigated as well. It is established that the default input rate estimation significantly underestimates the rate, which can be explained and corrected through consideration of event losses due to preamplifier resets. The proposed correction method provides a more reliable input rate estimation up to 10^5 cps.

Contents

CHAPTER 1

Introduction

Neutrinos are a well-established sector of the standard model of particle physics, however many questions regarding neutrinos still remain unanswered. This includes the determination of their absolute mass. Before the discovery that neutrinos change their flavor state through neutrino oscillations, they were thought to be massless. The neutrino oscillation mechanism requires mass differences between the three known neutrinos and thus, at least two non-zero masses are required. Nevertheless, neutrino masses are extremely small. So small, that no experiment has succeeded in determining the absolute mass yet.

The KATRIN experiment (**Ka**rlsruhe **Tri**tium **N**eutrino experiment) is a high precision β-decay experiment which aims to measure the effective electron neutrino mass with a sensitivity of 0.2 eV(90 % confidence level). This would improve upon previous experiments by one order of magnitude. The experiment consists of a highly active, gaseous tritium source and a MAC-E type spectrometer to select β-decay electrons with kinetic energies above a set minimum energy. The remaining electrons are focused on a silicon drift detector to count them and an integral spectrum is acquired by setting the spectrometer to different minimum energies. In this spectrum, a non-zero neutrino mass would lower the maximum kinetic energy available to the decay electrons, and furthermore lead to a distinct distortion of the spectrum near the endpoint. Thus, only the endpoint region is of interest for a neutrino mass measurement.

Using the KATRIN setup it is also possible to search for the existence of sterile neutrinos with masses in the keV range. These non-interacting neutrinos are candidates for warm dark matter and could play a role in explaining the smallness of the neutrino masses. Through neutrino mixing, they would contribute to the β-decay and leave a kink signature as well as a spectral distortion in the differential tritium spectrum. Within the context of warm dark matter models, the existence of such sterile neutrinos is already experimentally constrained to the ppm-level. Attempting to detect the spectral contribution in a β-decay pushes the experimenters understanding of model and experimental parameters to their limit.

For measuring the differential tritium spectrum with KATRIN, the spectrometer can be set to a low filter energy. This allows most of the decay electrons to reach the detector, increasing the event rate by a factor 10^8. Without the spectrometer the detector itself is

used to determine the electron energies, which requires a good intrinsic energy resolution. Leveraging the high luminosity of KATRIN's source enables acquisition of sufficient statistics to reach ppm-sensitivity within three years data taking. However, KATRIN's current detector and readout system is not suited to this task, since it is optimized for operation in the endpoint region where the event rate is low. Due to the detector technology it cannot provide the necessary precision in energy determination.

For this reason, the TRISTAN project (**TR**itium **I**nvestigation on **ST**erile (**A**) **N**eutrinos) was founded. It is centered around the construction and characterization of a new detector system that will be able to handle rates up to 10^9 counts per second (cps) while maintaining an excellent energy resolution.

For detector and front-end electronics the optimization of energy resolution is paramount. The biggest challenge for the back-end data acquisition system (DAQ system) is the rate requirement. When confronting DAQ systems with high rates, two systematic effects are of particular importance: dead time, which results in an overall loss of recorded events, and signal pileup, referring to the effect of multiple events being registered as one, leading to a distortion of the measured spectrum. Both effects scale with the event rate, making them evermore important for a ppm-level sterile neutrino search with TRISTAN. The magnitude of dead time and pileup can be reduced, to a certain extent, by using suitable hardware. Understanding and modeling these effects is crucial for TRISTAN and this work is an effort to advance in this direction.

Chapters 2 to 4 serve as an introduction to neutrino physics, the KATRIN experiment, and the TRISTAN project. In the subsequent chapters, the focus is laid on data acquisition. First, some basic principles of data acquisition for silicon drift detectors are explained in chapter 5. Building on this, dead time and pileup are investigated in chapter 6 which includes a detailed study of their origin, the development and verification of modeling approaches, as well as the discussion of correction and mitigation methods. Furthermore, an analytic model for pileup is put into the context of a sterile neutrino search and its sensitivity impact is quantified.

On the experimental side, the TRISTAN prototype DAQ system is tested with a detector emulator with regard to its capabilities for pileup rejection and rate detection. These measurements are also used to review the previously developed models and the results are presented in chapter 7.

CHAPTER 2

Neutrino physics

Neutrinos are uncharged, very light elementary particles that rarely interact with other matter, nevertheless they play an important role in nuclear, high-energy, and astrophysics. They were abundantly produced during the Big Bang resulting in a cosmic background of low energetic relic neutrinos carrying information about the earliest moments of our universe.

A large neutrino flux is also constantly emanating from the sun as a product of nuclear fusion within its core. On Earth roughly $6.5 \cdot 10^{10}$ neutrinos of solar origin pass through each square centimeter per second [Bel04], traveling unnoticed and nearly unhindered through us as well as the solid ground below our feet. On rare occasions they interact via the weak nuclear force, but even then they exhibit behavior unlike any other elementary particle of the standard model, as they seem to be missing a right-handed counterpart due to the combination of near masslessness and lack of electrical or color charge.

Since their postulation and discovery in the early 20th century the field of neutrino physics has developed vastly. Three neutrino flavors were discovered, completing the structure of three generations within the standard model of particle physics. Contrary to prior belief, the observation of neutrino flavor oscillations confirmed that neutrinos carry mass. However, due to its smallness the absolute mass scale has yet to be determined. Other questions about neutrinos remain as well. For example, whether they can act as their own antiparticle and the question of the existence of their non-interacting right-handed partners, the sterile neutrinos.

This chapter provides an introduction to aspects of neutrino physics with a focus on neutrino masses and sterile neutrinos. As a basis, their discovery, incorporation within the standard model, and the topic of neutrino oscillations are summarized in sections 2.1 to 2.3. Building on this, methods for probing the absolute mass scale and possible extensions of the standard model with sterile neutrinos are discussed in sections 2.4 and 2.5.

2.1 Postulation and discovery

On the 4th of December 1930, the existence of neutrinos was postulated by Wolfgang Pauli in a letter to a society of nuclear physicists in Tübingen [Pau30]. What he first named "neutron" later became known as the neutrino when Enrico Fermi shaped this idea into a theory about the particle's interactions [Fer34].

The postulation of the neutrino was Pauli's attempt to explain the shape of energy spectra from radioactive beta decay. Today we know that in such a decay, or more specifically a β^--decay, a neutron converts to a proton, an electron, an electron antineutrino and a fixed amount of energy distributed to all these daughter particles. This process is typically represented in the reaction scheme

$$\mathrm{n} \rightarrow \mathrm{p} + \mathrm{e}^- + \overline{\nu}_\mathrm{e} \ . \tag{2.1}$$

In 1930 it was known that an electron is emitted in this process, but neutrinos, being electrically neutral and their other interactions being weak as well, could not be detected. If only two particles were produced in this process, namely the daughter nucleus and the electron, both would receive the same amount of energy due to the conservation of momentum. Then the electron energy spectrum would be a monoenergetic line. This, however, is in contradiction to experiments that show a continuous distribution of energies between zero and a maximum energy E_0, as shown in figure 2.1. Some physicists were ready to abandon the law of conservation of energy to explain this conundrum. Pauli, on the other hand, suggested that an additional, electrically neutral particle is emitted which remained undetected by the experimenters.

Due to the neutrinos low interaction cross section they were often referred to as "ghost" particles and it took until 1956 for the first direct observation of the neutrino by the Cowan-Reines neutrino experiment [RC53]. Quite fittingly, this experiment was often referred to as project poltergeist and it looked for the signature of inverse beta decays induced by electron antineutrinos. A large neutrino flux was needed to compensate for the low interaction cross section and it was provided by a nuclear reactor. This idea proved to be excellent and reactor neutrinos now play a significant part in modern neutrino oscillation experiments.

Over the remaining half of the 19th century two additional neutrino flavors were discovered corresponding to the second and third generation of particles in the standard model of particle physics. The muon neutrino ν_μ was discovered in 1962 at the AGS accelerator in Brookhaven [Dan62], and the tau neutrino ν_τ was observed in 2000 by the DONUT experiment [Kod01]. All three neutrinos were appropriately labeled according to the particles they couple to in charged current weak interactions, namely the electron, muon and tau lepton.

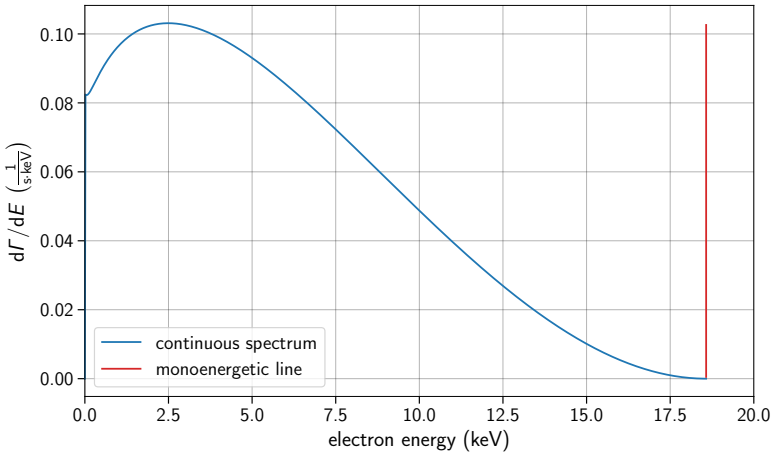


Figure 2.1: Normalized energy spectrum of beta decay electrons as described by Fermi (blue curve). Here shown in case of the tritium decay with an endpoint energy of $E_0 = 18.6\,\mathrm{keV}$. The red line exemplifies the expected monoenergetic line if no neutrinos were involved.

2.2 Neutrinos in the standard model

The standard model [Tan18] is the most successful theory of particle physics to date and has yet to be replaced by a more complete theory. It is a quantum field theory that describes the properties of the 17 known fundamental particles and their interactions through three of the four fundamental forces. Gravity has not yet been successfully integrated into a coherent theory of particle physics and on the scale of elementary particles, it is very weak compared with the other forces.

The particles of the standard model can be grouped into categories (see figure 2.2): the matter particles quarks and leptons, the force carrier gauge bosons, and the Higgs boson. The quarks and leptons are also labeled as fermions since they each have a spin of $1/2$ in contrast to the bosons that have integer spins. Additionally, each fermion has a corresponding antiparticle which is its charge conjugated partner.

Quarks and leptons are subdivided into three generations or flavors of increasing mass scales. Most known matter of our universe is made of particles from the lightest generation, namely the up quark, down quark and the electron. Up and down quarks form protons and neutrons within atomic nuclei, and electrons constitute their atomic shells. As stable particles photons and neutrinos are also very abundant in our universe. In contrast, the other matter particles of higher generations are unstable and decay into particles of the first generation.

Forces are mediated by gauge bosons that stem from invariances of the standard model Lagrangian density toward local gauge transformations. The photon transmits the electromagnetic force between electrically charged particles, gluons transfer the strong force between color-charged particles confining quarks and gluons into bound states, and Z and W bosons mediate the weak force which is responsible for flavor changing phenomena such as the beta decay and quark mixing. Furthermore, the electromagnetic

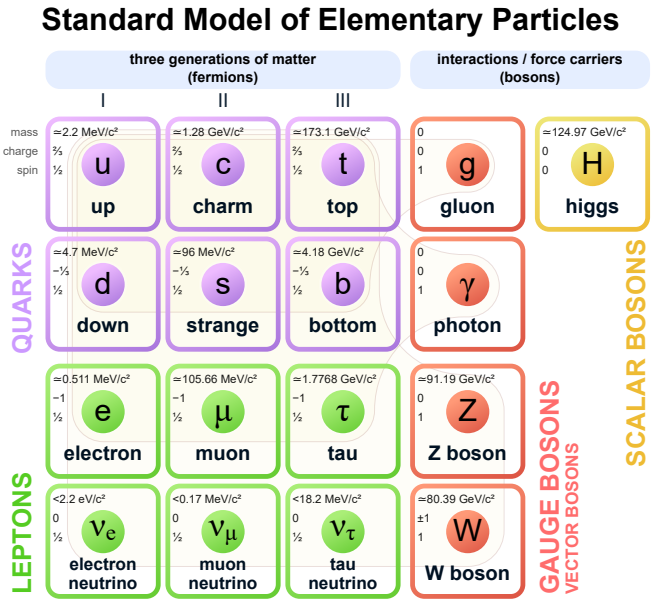


Figure 2.2: Particles of the standard model [Mis06].

and the weak force are unified into the electroweak force, connecting the massless photon and the massive Z and W bosons through one combined gauge symmetry. This electroweak symmetry is spontaneously broken at low energies through the Higgs mechanism which explains the non-zero masses of the Z and W. More specifically this is achieved by adding a term containing a scalar field, the "Higgs field", to the standard model Lagrangian. The particle corresponding to this field is the Higgs particle and as an important missing piece of the standard model it was discovered in 2012 at the Large Hadron Collider by the CMS and ATLAS experiments [CMS12] [ATL12].

There are three neutrinos in the standard model. They carry neither electrical nor color charge and only interact via the weak force which makes them notoriously hard to detect. They were assumed to be massless until the discovery of neutrino oscillations by the Super-Kamiokande experiment in 1998 [Fuk98]. These flavor changing oscillations imply a mass splitting between the neutrino states, but give no information on the absolute mass scale. This absolute mass scale is very low in comparison to the masses of other elementary particles. So low that even to this date no experiment has succeeded in performing a direct measurement of neutrino masses. Instead, these experiments provide upper limits on the neutrino mass scale.

Neutrinos are closely related to the phenomenon of parity violation as demonstrated by the Wu experiment [Wu57] and the Goldhaber experiment [GGS58] in 1957 and 1958 respectively. These experiments proved that neutrinos are left-handed (spin opposed to the direction of travel) and antineutrinos right-handed (spin aligned in direction of travel). This implies that charged current weak interactions, in which beta decay neutrinos are

produced, do not act in the same way on parity-transformed systems. On the contrary, they maximally violate this symmetry, unlike the electromagnetic and strong forces which conserve parity. This led to the conclusion that W bosons only couple to left-handed particles and right-handed antiparticles.

Thus, the theory of the electroweak interaction is constructed as a chiral theory, splitting particle fields into left-handed and right-handed components. The matter particles of the standard model are then represented as left-handed doublets and right-handed singlets of the weak isospin - with one exception. As there is no evidence for the existence of right-handed neutrinos and left-handed antineutrinos they are not included in the standard model. Furthermore, this leads to neutrinos being massless in this context. For massive particles, it is possible to detect particles in the parity-flipped state compared to the initial state of charged current production. Left-handed and right-handed states are allowed to mix if there is a frame of reference where the helicity is flipped. This introduces a term containing the flipped state to the particles state with a coefficient $\propto \frac{m}{E}$ where m is the mass and E the energy of said particle [Tan18]. For massless particles, this term is equal to zero and they are not allowed to change handedness through propagation. Contrary to this, for the neutrino this term must vanish if the right-handed state does not exist, meaning they have no mass.

If right-handed neutrinos existed, they would not couple to standard model particles and would only manifest themselves through neutrino oscillations and gravity. Hence they are often referred to as sterile neutrinos [Vol02]. They pose a compelling extension to the standard model and could help explain neutrino masses and their large disparity compared to the other fermion masses. It is very natural to assume the existence of right-handed neutrinos when all other particles have a right-handed counterpart. The inclusion of neutrino masses in an extended standard model also raises questions on the Dirac or Majorana nature of neutrinos. The masses of all charged fermions are represented by Dirac mass terms. Since neutrinos only carry the weak isospin it is possible to add another type of mass term, the Majorana mass term, to the standard model Lagrangian without violating charge conservation [Moh07].

2.3 Neutrino oscillations and mixing

Neutrino oscillations are transitions of neutrino flavors from one state to another through propagation over distances. A neutrino produced as an electron neutrino in a weak process, in a reactor for instance, may change its flavor to a muon neutrino upon detection a couple of kilometers away from the reactor. This phenomenon is only possible for massive neutrinos. The idea of massive neutrinos and their oscillations was proposed by Bruno Pontecorvo in 1968 [Pon68], inspired by the deficit in the neutrino flux originating from the sun as measured by the Homestake Experiment in the late 1960s [DHH68].

The observation of neutrino oscillations is one of the most important physics discoveries in recent decades and there is evidence from many different experiments studying solar, atmospheric, accelerator, and reactor neutrinos [Tan18]. The first definite observation of neutrino oscillations was achieved by the Super-Kamiokande Collaboration in 1998

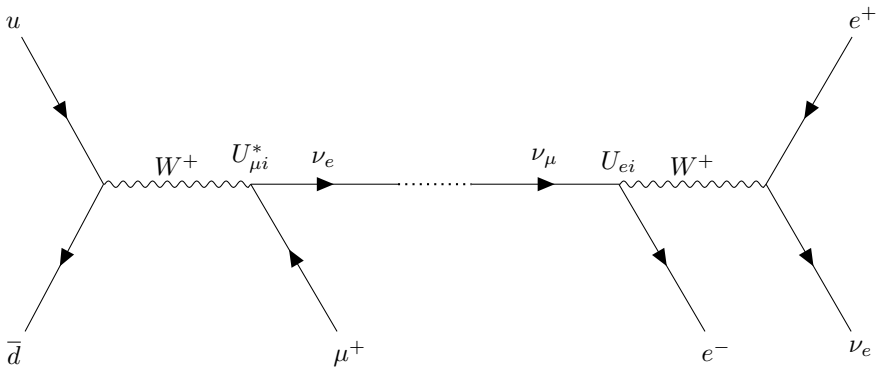


Figure 2.3: Representative Feynman graph for a neutrino oscillation. A muon neutrino is produced in a pion decay and propagates as a superposition of mass eigenstates. The propagated state is projected onto the electron neutrino eigenstate when it is detected in an electron scattering process.

[Fuk98] and confirmed by the SNO Collaboration in 2001 [PS01], which led to the award of the 2015 Nobel Prize for Physics to Takaaki Kajita [Kaj15] and Arthur McDonald [McD15].

The mathematical description of neutrino oscillations is given in terms of transition and survival probabilities. For example, this can refer to the probability for survival of accelerator muon neutrinos $P(\nu_\mu \to \nu_\mu)$ when they are detected through charged current muon production, or to the disappearance probability $P(\nu_e \to \nu_{\mu,\tau})$ of reactor electron neutrinos when they transition to the other flavors and are not registered by an inverse beta decay detector. These transitions are possible for massive neutrinos where the eigenstates of propagation (mass eigenstates) are not the same as the eigenstates of weak interaction (flavor eigenstates). This is similar to quark mixing where the strong eigenstates differ from the weak eigenstates resulting in mixing between quark generations. The two sets of neutrino eigenstates are connected by the Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS matrix) $U_{\ell i}$. Similarly to the Cabibbo–Kobayashi–Maskawa matrix (CKM matrix) from quark mixing, this is a unitary matrix describing a rotation from one set of eigenstates to another.

With this matrix the flavor eigenstates $|\nu_\ell\rangle$ can be expressed as a linear combination of the mass eigenstates $|\nu_i\rangle$ or vice versa with the conjugate transposed matrix.

$$|\nu_\ell\rangle = \sum_{i=1}^{3} U_{\ell i} \, |\nu_i\rangle \qquad i = 1, 2, 3 \;\text{ and }\; \ell = e, \mu, \tau \tag{2.2}$$

In case of three neutrino flavors the PMNS matrix has 3×3 complex elements and can be parametrized by three Euler angles $\theta_{12}, \theta_{23}, \theta_{13}$ and up to three CP violating phases, depending on the Dirac or Majorana nature of neutrinos [Tan18].

The mass splitting and the PMNS matrix elements are probed with solar, atmospheric, accelerator and reactor neutrinos by a multitude of experiments that measure transition

and survival rates. In two flavor neutrino mixing the probability for a transition from flavor state α to flavor state β would be

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta | \nu_\alpha(t) \rangle|^2 = \sin^2(2\theta) \cdot \sin^2\left(\frac{1.27 \cdot \Delta m^2_{ij}[\mathrm{eV}^2] \cdot L[\mathrm{km}]}{4E[\mathrm{GeV}]}\right) , \qquad (2.3)$$

where θ is the mixing angle, L is the distance between creation and detection, E is the energy of the neutrino and $\Delta m_{ij} = m_i^2 - m_j^2$ is the mass splitting. A consequence of this equation is that neutrino oscillations are only possible if the masses of neutrinos differ from each other ($\Delta m_{ij} \neq 0$) and if the generations mix ($\theta \neq 0$). Two flavor mixing is a good approximation for $\nu_\mu \to \nu_\tau$ and $\nu_e \to \nu_{\mu,\tau}$ transitions, but generally it will be a superposition of multiple oscillation terms in case of three or more flavors [Tan18].

Results from many experiments are frequently combined by the NuFit-group [Est19; Est18] which currently yields

$$\Delta m^2_{12} \approx 7.4 \cdot 10^{-5}\,\mathrm{eV}^2 \quad \text{and} \quad |\Delta m^2_{23}| \approx 2.5 \cdot 10^{-3}\,\mathrm{eV}^2 \qquad (2.4)$$

for the two mass splittings. The sign of Δm^2_{23} is still unknown. Furthermore, oscillation experiments are only sensitive to mass splittings and not the absolute mass scale. This suggests three possibilities for the hierarchy of the mass eigenstates. The normal hierarchy where $m_1 \ll m_2 < m_3$, the inverted hierarchy $m_3 \ll m_1 < m_2$ and the quasi-degenerate mass spectrum $m_1 \approx m_2 \approx m_3$, where the mass scale is on the order of 0.1 eV and much larger than the splittings.

2.4 Neutrino mass measurements

The absolute neutrino mass scale is so low that no experiment has been able to measure it yet. Instead, the experiments deliver upper limits in the eV range. There are currently three approaches to measure and constrain the neutrino mass scale [Dre13].

- **Cosmological constraints**
  Neutrinos and their masses are connected to structure formation in the early universe. They they were plentifully produced during decoupling of the electroweak force after the Big Bang and due to their low mass neutrinos have a large free streaming length, acting as hot dark matter. Therefore, they provide gravitational wells for matter clustering only on large scales while decreasing the formation of galaxies. By measuring the structure of the early universe, which is imprinted in the cosmic microwave background, and the structure of today's universe, determined by galaxy surveys, the sum of the neutrino masses can be tightly constrained. The most recent result by the Planck collaboration yields [Agh18]

  $$\sum_j m_j < 0.12 - 0.54\,\mathrm{eV}, \quad 95\,\%\ \mathrm{CL}\ . \qquad (2.5)$$

  This limit is highly model dependent and varies with the choice of included datasets, where the tightest bound is achieved via inclusion of baryon acoustic oscillation measurements from galaxy redshift surveys.

- **Neutrinoless double beta decay** ($0\nu\beta\beta$)
  The neutrinoless double beta decay is a hypothetical decay of a nucleus emitting two beta electrons at the same time without emitting neutrinos. This process is possible if neutrinos are Majorana fermions, meaning they can act as their own antiparticles. The two neutrinos involved in the neutrinoless double beta decay are only virtual particles and cancel out as shown in figure 2.4. The corresponding decay rate is proportional to the absolute square of the effective electron neutrino Majorana mass m_{ee}, which is the coherent sum of all neutrino masses weighted with PMNS matrix elements:

  $$\Gamma \propto m_{ee}^2 = \left| \sum_{i=1}^{3} U_{ei}^2 m_i \right|^2 . \tag{2.6}$$

  Here, in contrast to the pure Dirac neutrino picture, the PMNS matrix contains two additional Majorana phases for a total of three CP violating phases. The decay rate is also proportional to the nuclear matrix element which introduces a model dependent theoretical uncertainty. A combination of experimental results from 2015 sets the limit $m_{\beta\beta} < 0.130 - 0.310\,\mathrm{eV}$ depending on the choice of calculation for the nuclear matrix element [Guz15].

- **Direct neutrino mass measurement**
  The most reliable method of determining the neutrino mass, in terms of model independence, is a direct measurement based solely on kinematics. This technique relies on the fundamental relationship $E^2 = p^2 + m^2$. By determining the energy and momentum of a neutrino or the kinetic energy of electrons in a β-decay, the neutrino mass scale can be measured or constrained.

  One possibility is a time-of-flight measurement of neutrinos from supernovae. The momentum is connected to the time delay between light and neutrino arrival and the energy can be obtained from the amount of signal in the detector. Data from SN1987A collected by Kamiokande-II, IMP and Bakasan yield the mass limit $m_\nu^2 < 5.8\,\mathrm{eV}$ [PRV10], which is only slightly dependent on the supernova model.

  Another more sensitive method is precision beta spectroscopy. Here, the neutrino mass is determined by precisely measuring the energy spectrum of β-electrons in the endpoint region, where the electron receives the full decay energy and the neutrino has momentum equal or close to zero. This method is sensitive to the effective electron antineutrino mass

  $$m_{\nu_e}^2 = \sum_{i=1}^{N_\nu} |U_{ei}^2| m_i^2 , \tag{2.7}$$

  which is an incoherent weighted sum over all squared neutrino masses and, opposed to m_{ee} from $0\nu\beta\beta$-experiments, does not contain the CP violating phases. The lowest limit from this method is set by the Troitsk experiment [Ase11]

  $$m_{\overline{\nu}_e} < 2.05\,\mathrm{eV} , \quad 95\,\%\ \mathrm{CL} . \tag{2.8}$$

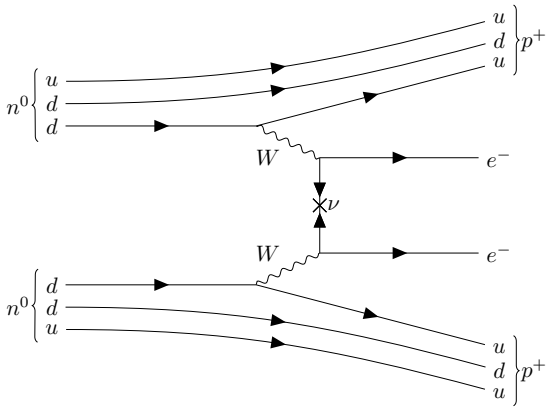


Figure 2.4: Feynman diagram of the neutrinoless double beta decay. This process is only possible if neutrinos are Majorana fermions and act as their own antiparticle.

With $m_{\overline{\nu}_e} < 2.3\,\text{eV}$ at 95 % CL the Mainz experiment achieved a similar result [Kra05] and the next generation of laboratory experiments is working on improving these limits. This includes the tritium beta decay experiments KATRIN and Project 8 which will measure the electron antineutrino mass, as well as complementary experiments ECHo, HOLMES and NuMECS determining the electron neutrino mass from the ^{136}Ho electron capture spectrum [Mer16]. The KATRIN experiment, which uses the MAC-E filter technique, is the direct successor of the Troitsk and Mainz experiments and aims to reach a sensitivity of $\sim 0.2\,\text{eV}$ at a 90 % CL [KAT05].

2.5 Sterile neutrinos

Sterile neutrinos represent a natural and well-motivated extension of the standard model. Currently, the standard model includes the three active neutrinos ν_e, ν_μ, and ν_τ, which have all been directly observed through their interactions and this exact number of light, active neutrinos is supported by measurements of the invisible Z boson decay width [ALE06]. As mentioned in section 2.2, neutrinos are represented as part of the left-handed electroweak lepton doublets. However, unlike other particles their counterpart, the right-handed electroweak singlet, is missing. Adding right-handed neutrinos can contribute to explaining a variety of open problems. It would fill the gap in the standard model fermion field structure, open a path to introduce neutrino masses, and in some theories explain the large mass discrepancy between active neutrinos and the other fermions through the seesaw mechanism.

These right-handed neutrinos are often referred to as “sterile”, since they would not couple to the other particles through any of the standard model interactions. This makes them natural candidates for dark matter depending on their mass and if there is a viable mechanism of production in the early universe. Although a direct interaction with standard model particles is only possible via gravity, sterile neutrinos would still

mix with the active flavors and possibly leave signatures in oscillation and spectroscopy experiments.

The methods of theoretically constructing the enhanced neutrino sector are plentiful and leave many choices such as the number of sterile neutrinos, their mass and the inclusion of Majorana mass terms [Vol02]. The following list highlights three cases as given in [Adh17], which are distinguished by the sterile neutrinos mass m_s:

- $m_s \gg$ TeV
  A typical method for explaining the small scale of the neutrino masses is the minimal type I seesaw mechanism [Boy19]. This scenario introduces both Dirac and Majorana neutrino mass terms where the Dirac mass m_D is on the scale of the other fermion masses and the Majorana mass M is much larger. Diagonalizing the corresponding mass matrix naturally leads to three small mass states and one or multiple very heavy sterile neutrinos with masses above the TeV-scale.

  Not only would this answer the question of the large mass discrepancy between neutrinos and the other fermions, but it could also provide an answer to the generation of baryon asymmetry in the early universe through its implications on leptogenesis [Tan18].

- $m_s = \mathcal{O}$(eV)
  The oscillation between active and sterile neutrinos with additional mass states in the eV-range is a proposed solution to the measured excess of electron neutrino appearances at LSND and MiniBooNE [Agu18]. Other measurements in favor of this scenario are the electron antineutrino reactor anomaly, which constitutes a significant lack of antineutrinos detected in short baseline reactor experiments [Men11], and the gallium anomaly where a deficit of neutrino flux was detected in radioactive source calibrations of the GALLEX and SAGE experiments [Tan18]. There are however tensions with a multitude of other appearance and disappearance experiments, which significantly constrain the parameter space of active to sterile oscillation parameters in case of a model with three active neutrinos and one sterile neutrino [Tan18; An14; Alm18].

  Additional mass states in the eV scale would also leave signatures in the form of kinks in beta decay spectra and can be probed with the KATRIN experiment.

- $m_s = \mathcal{O}$(keV)
  As first proposed by Scott Dodelson and Lawrence M. Widrow [DW94], sterile neutrinos in the keV range are viable candidates for dark matter. They carry no standard model charge and can act as either warm or cold dark matter depending on the specifics of the production mechanism. They could, for instance, be produced via oscillations at high temperatures in the early universe, but unlike WIMPs, they would not enter thermal equilibrium due to their low interaction strength [BRS09]. Due to their non-zero coupling to active flavors, keV sterile neutrinos are technically unstable and decay to a photon and an active neutrino. With sufficiently low mixing

angles, however, their lifetime is on the order of the age of the universe. Cosmic X-rays and other cosmological observables strongly limit the allowed parameter space for sterile neutrinos in the high mass region, but do not exclude the low keV region as shown in figure 2.5 [Boy19].

A popular extension of the standard model which includes keV sterile neutrinos is the neutrino minimal standard model (νMSM)[ABS05]. It introduces no new physics other than three sterile neutrinos. One at keV mass, which represents the dark matter candidate, and two at roughly the GeV-scale which ensure small active neutrino masses through the seesaw mechanism. Furthermore, these heavy states can generate the observed baryon asymmetry of the universe through oscillation-induced leptogenesis [AS05].

A direct search for keV sterile neutrinos with laboratory experiments is challenging but possible. They would leave signatures in beta decay and electron capture spectra [Adh17], however due to their extremely low mixing this effect would be very small. The TRISTAN project, a part of the KATRIN experiment, aims to probe the sterile parameter space up to masses of $\sim 18.6\,\mathrm{keV}$ and down to mixing angles $\sin^2\theta < 10^{-6}$ [Mer18]. This region is largely excluded by existing X-ray constraints, but direct and model independent searches with laboratory conditions are worthwhile nonetheless.

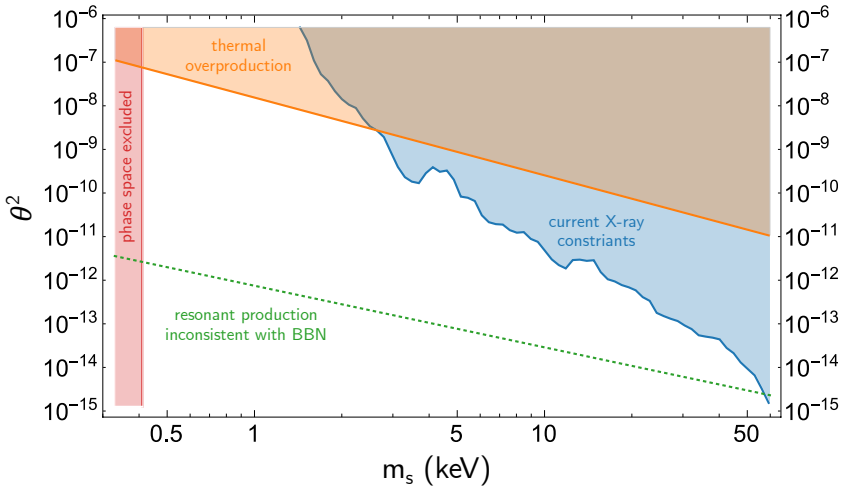


Figure 2.5: Constraints on the keV sterile neutrino parameter space. The figure is based on [Boy19] and shows which combinations of sterile masses m_s and mixing angles θ are excluded. Only the most stringent limits on keV sterile neutrinos are shown. They are comprised of phase space considerations due to Pauli's exclusion principle (red), limits from non-observation of X-rays from sterile neutrino decays (blue), and the overproduction of sterile neutrinos at large mixing angles compared to the observed dark matter density (orange). The orange line matches the parameter values where the correct amount of dark matter is produced in thermal equilibrium. Below this line, sterile neutrinos can be produced in the appropriate amount through resonantly enhanced thermal production (see [Boy19]). This mechanism breaks down below the dotted green line, because it would be inconsistent with Big Bang nucleosynthesis.

CHAPTER 3

The KATRIN experiment

Almost a century after Pauli's discussion of beta decay and the postulation of the neutrino, beta decay is still relevant in modern physics. The energy spectrum of beta decay electrons allows direct determination of the neutrino mass based solely on kinematics.

Over the last decades, the experimental techniques have improved drastically and the beta spectrum can be measured with unprecedented precision. Building on the experience of former direct neutrino mass measurements Troitsk and Mainz, the KATRIN Experiment (**Ka**rlsruhe **Tri**tium **N**eutrino Experiment) [KAT05] aims to improve upon the current neutrino mass limit by an order of magnitude to achieve a sensitivity of $m_\nu \sim 0.2\,\mathrm{eV}$ at a 90 % confidence level. It does so by measuring the integral beta spectrum of tritium with the largest MAC-E filter ever built. The experiment uses molecular tritium gas as a radioactive source, because of its high activity and low endpoint energy of 18.6 keV. Handling gaseous tritium requires secure infrastructure and radiation expertise. KATRIN is therefore located at the Tritium Laboratory Karlsruhe (TLK) at the Karlsruhe Institute of Technology (KIT) to leverage existing infrastructure. The KATRIN collaboration finished the commissioning of the full apparatus in 2018 and is now taking data.

This chapter discusses the signature of non-zero neutrino masses in tritium β-decay and provides an overview of the experimental setup. The information presented in this chapter is mainly based on [KAT05; Are16; Kle18; Ams15].

3.1 Neutrino mass signature in tritium β-decay

Beta decay is a phenomenon of the weak nuclear force. In β^--decay a nucleus emits an electron and an electron antineutrino through the W-boson as shown in figure 3.1. In this process a down quark transitions to an up quark changing the neutron to a proton. Furthermore, a fixed amount of energy is released which is distributed to all daughter particles. The KATRIN experiment uses molecular tritium as a source of decay electrons.

$$\mathrm{T}_2 \rightarrow \mathrm{T}\,{}^3\mathrm{He}^+ + e^- + \overline{\nu}_e \ . \tag{3.1}$$

The energy spectrum of the decay electrons is adequately described by Fermi's effective theory of the beta decay [Fer34; Mer15a]. In the case of only one neutrino mass eigenstate m_ν, it yields the differential decay rate

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E} = C \cdot F(E,Z) \cdot p \cdot (E + m_e) \cdot (E_0 - E)\sqrt{(E_0 - E)^2 - m_\nu^2} \ , \tag{3.2}$$

where E denotes the electrons kinetic energy, p its momentum and m_e its mass. The endpoint energy, i.e. the maximum energy the electron can receive as it would be for $m_\nu = 0$, is $E_0 \approx 18.6\,\mathrm{keV}$ in tritium decay. The Fermi function $F(E,Z)$ is a theoretical correction is which takes the Coulomb interaction between electron and daughter nucleus into account. The constant

$$C = \frac{G_F^2}{2\pi^3}\cos^2\theta_C |M|^2 \tag{3.3}$$

contains the Fermi constant G_F, which represents the effective weak coupling, the Cabibbo angle θ_C and the nuclear transition matrix element M. The latter is energy independent since tritium decay is superallowed. This means T and He^+ have a common ground state and additional shape factors can therefore be neglected.

For multiple neutrino mass eigenstates m_i the spectrum will be a superposition of spectra weighted by the corresponding mixing amplitude $|U_{ei}|^2$ to the electron neutrino state

$$\frac{d\Gamma}{dE} = C \cdot F(E,Z) \cdot p \cdot (E + m_e) \cdot (E_0 - E) \sum_{i=1}^{3} |U_{ei}|^2 \sqrt{(E_0 - E)^2 - m_i^2} \ . \tag{3.4}$$

Although the KATRIN main spectrometer has an excellent energy resolution below 1 eV, it is not enough to resolve the much smaller mass splittings (see equation 2.4). Instead the experiment is sensitive to an effective mass m_{eff} because

$$\sum_{i=1}^{3} |U_{ei}|^2 \sqrt{(E_0 - E)^2 - m_i^2} \approx \sqrt{(E_0 - E)^2 - m_{\mathrm{eff}}^2} \quad \text{for} \quad \frac{m_i^2}{(E_0 - E)^2} \ll 1 \tag{3.5}$$

$$\text{where} \quad m_{\mathrm{eff}}^2 = \sum_{i=1}^{3} |U_{ei}|^2 m_i^2 \ , \tag{3.6}$$

which can be shown through Taylor expansion [Kra05]. The effect of a non-zero effective mass on the differential spectrum is depicted in figure 3.2.

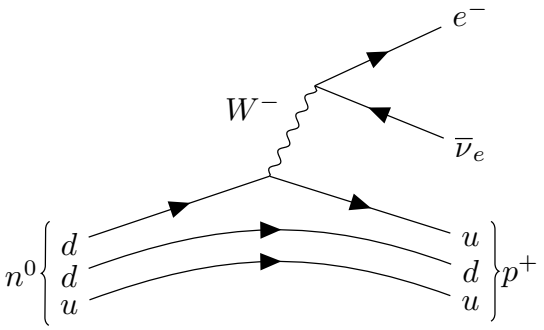


Figure 3.1: Feynman graph of a β^--decay. A down quark within the neutron transitions to an up quark under emission of an electron and an electron antineutrino through the exchange of a virtual W boson. With the two spectator quarks a proton remains.

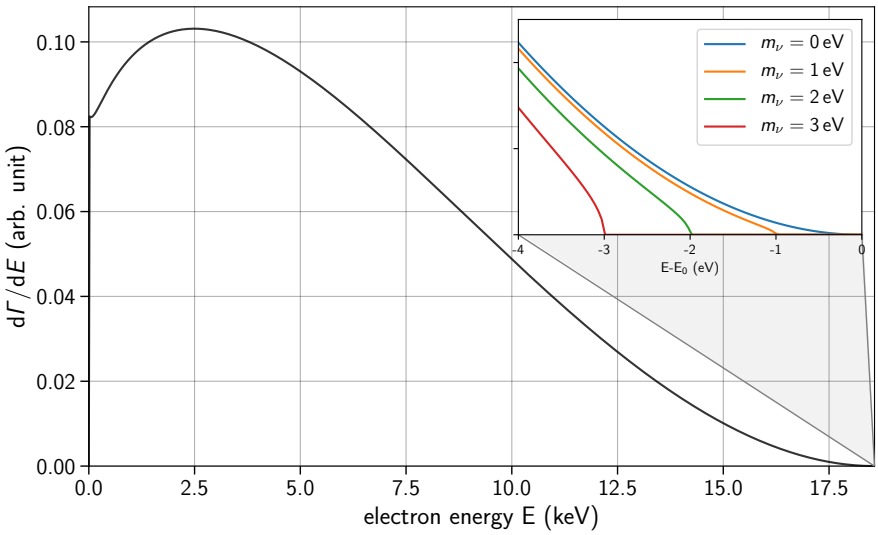


Figure 3.2: Differential energy spectrum of β-electrons from tritium decay. The neutrino mass signature is visible in the endpoint region and demonstrated for several mass values in the $0 - 3\,\mathrm{eV}$ range. The effective endpoint is shifted below E_0, which is the endpoint for $m_\nu = 0$, and the spectral shape near the endpoint is altered.

3.2 Experimental setup

The KATRIN experiment measures the integral tritium beta spectrum by fulfilling two general tasks. The first is to provide a very active and highly stable radioactive source of β-electrons, for which KATRIN uses molecular tritium. The other is to precisely measure the energy of these electrons which is realized with an electrostatic retardation spectrometer. These tasks are reflected in the experimental setup by separation into the source and transport section (STS) and the spectrometer and detector section (SDS). Together they form the roughly 70 m long KATRIN beamline as shown in figure 3.3. In the following, an overview of the beamline is given and sections 3.2.1 to 3.2.4 provide more detailed descriptions for the source, the transport section, the spectrometer, and the detector. Some additional focus is put on the detector section in order to provide a comparison to the TRISTAN detector upgrade discussed in chapter 4.

Tritium is injected in the windowless gaseous source (WGTS) where it is held at a constant level of activity and circulated in a closed loop. The source is not separated from the rest of the beamline by any windows in order to avoid β-electron energy loss from passing through a solid. The decay electrons are magnetically guided along the beamline in helical cyclotron motion.

When they exit the source they enter into the transport section, comprised of a differential pumping section (DPS) and a cryogenic pumping section (CPS), where gas and ions are removed and the pressure lowered to $< 10^{-10}$ mbar. This extremely high vacuum (XHV) is required to reduce scattering of β-electrons off molecules which would result in energy loss. Furthermore, tritium is not allowed in the spectrometer and detector section for background and radiation safety reasons.

Upstream of the WGTS is the rear section (RS). It is used for monitoring the WGTS and for defining a reference potential for all electrodes. Furthermore, it is equipped with an electron gun which is used for characterization of the whole beamline. Rear section, source, retention components, and monitoring systems form the STS.

Downstream of the transport section, the spectrometer and detector section begins. The SDS is comprised of a pre-spectrometer, the main spectrometer, the monitor spectrometer and the focal plane detector (FPD). All three spectrometers are electrostatic filters of the MAC-E type, which stands for magnetic adiabatic collimation and electrostatic retardation. The working principle is further discussed in 3.2.3, but essentially they act as high-pass filters letting only electrons of some minimum energy pass.

The pre-spectrometer is mainly used as a pre-filter to reduce the flux of β-electrons entering the main spectrometer, but among other applications, it is also used during commissioning and maintenance measurements. The heart of KATRIN is the main spectrometer with its 23 m long and 10 m high vacuum vessel and it is used for the precise determination of β-electron energies. Due to its size, the main spectrometer achieves a much better energy resolution than pre-spectrometer and the spectrometers of predecessor experiments. For an accurate measurement of the retarding potential the former spectrometer of the Mainz experiment, now referred to as monitor spectrometer, is repurposed as a monitoring device.

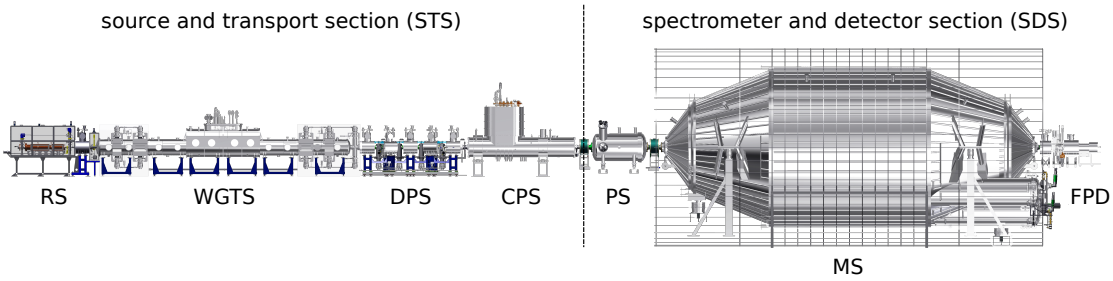


Figure 3.3: The KATRIN beamline can be subdivided into eight sections. From left to right these are the rear section (RS), the windowless gaseous tritium source (WGTS), the differential pumping section (DPS), the cryogenic pumping section (CPS), the pre-spectrometer (PS), the main spectrometer (MS) and the focal plane detector (FPD). Along the beamline β-electrons are guided with magnetic fields generated by superconducting solenoids and air coil systems. In parallel to the beamline the monitor spectrometer, which is not shown here, is used for an accurate determination the retarding potential of the main spectrometer.

Electrons that pass the main spectrometer are magnetically focused onto the focal plane detector where they are counted. By setting the main spectrometer to different retarding potentials and measuring the events, the integral β-spectrum is scanned.

3.2.1 Windowless gaseous tritium source

The windowless gaseous tritium source (WGTS) is a 10 m long steel tube with a diameter of 90 mm. Molecular tritium gas is injected through capillaries in the middle and pumped off with a total of eight turbomolecular pumps at the ends which results in an approximately triangular density profile (see figure 3.4). The gas is circulated in the inner loop system and its molecular content is continually monitored with laser Raman spectroscopy. By removing decay products and adding fresh tritium the purity is kept above 95 %. The injection rate controls the number of tritium molecules in the source, which is usually expressed as a column density ρd in units of molecules per cm^2 of the source cross section area. The column density is an important parameter in the theoretical model of the KATRIN apparatus as it defines the rate of β-electrons and the amount of scattering in the source. The nominal value for ρd is $5 \cdot 10^{17}\,\mathrm{cm}^{-2}$ which corresponds to an activity of roughly 10^{11} Bq. A source stability of less than 0.1 % is needed in order to keep the systematic contribution low enough to reach design sensitivity. To keep gas pressure and flow rate low, and for reducing uncertainties in the energy of β-electrons due to Doppler broadening, the WGTS is kept at 30 K during operation. The whole source is encased by superconducting solenoids creating a 3.6 T magnetic field directed along the beam tube. This field guides all electrons that are emitted in the forward direction toward the spectrometer section. Electrons that are emitted in the backward direction scatter off the rear wall and lose too much energy to pass the main spectrometer if reflected back toward the SDS. The magnetic field strength is also an important parameter for modeling, because its ratio to the field strength in front of the main spectrometer defines the acceptance angle (see section 3.2.3).

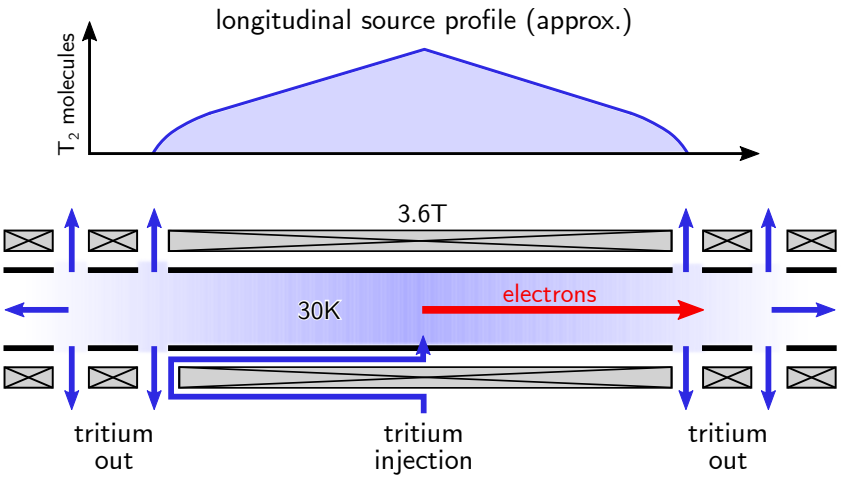


Figure 3.4: Schematic drawing of the WGTS and the longitudinal tritium density profile. The graphic is adapted from [KAT05].

3.2.2 Transport section

In addition to the pumps at the ends of the WGTS, two additional pumping stages are used to retain gas within the source and transport section and prevent it from reaching the spectrometers. The combined retention system is designed to keep the tritium flow into the pre-spectrometer lower than 10^{-14} mbar l/s. The first stage is a differential pumping section (DPS) which relies on turbomolecular pumps with pumping speeds > 2000 l/s for H_2. The second stage is a cryogenic pumping section where gas molecules are adsorbed on a surface cooled to 4.5 K with liquid helium. To improve the trapping probability the surface is covered with a thin layer of argon frost.

Due to the low pressure, gas moves through the transport section with molecular flow. More molecules can therefore be retained by not allowing a direct line of sight to the spectrometer section. Instead, the beamline is bent at eight occasions so that molecules have a higher chance of rebounding off the wall toward an adjacent pump. Meanwhile, β-electrons are guided around the bends by magnetic dipole fields generated by superconducting solenoids.

As well as producing electrons, the source also produces charged ions through tritium decay and scattering processes. Similar to electrons, ions are magnetically guided along the beamline, but have to be retained in the STS. This is achieved with a combination of ring electrodes, to create potential barriers that block positive ions, and dipole electrodes, for drifting ions to the walls of the beam tube.

3.2.3 Spectrometer section and response function

The KATRIN tritium source is an isotropic source and as such β-electrons are emitted in all directions equally. To measure the kinetic energy of these electrons with an electrostatic retardation field, it is necessary to align the electrons momenta with the electric

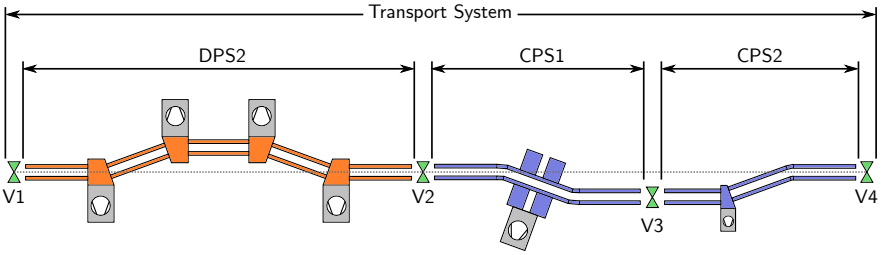


Figure 3.5: Schematic view of the transport system adapted from [KAT05].

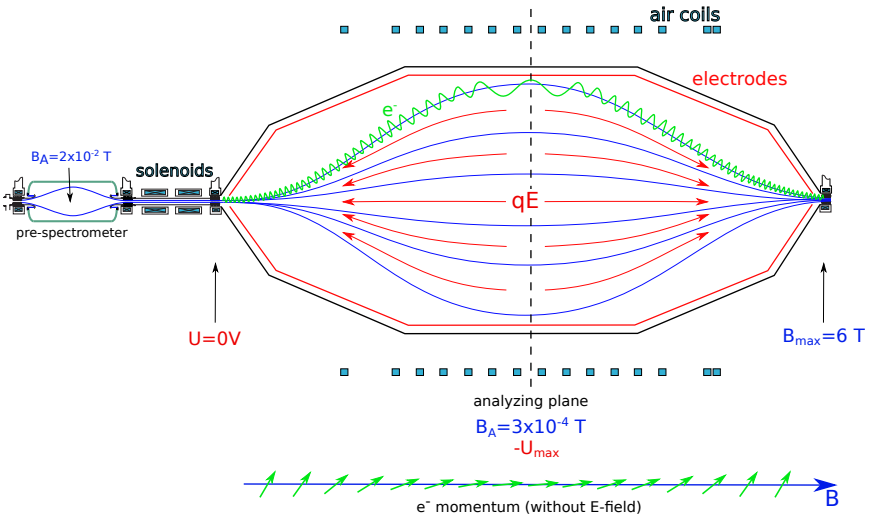
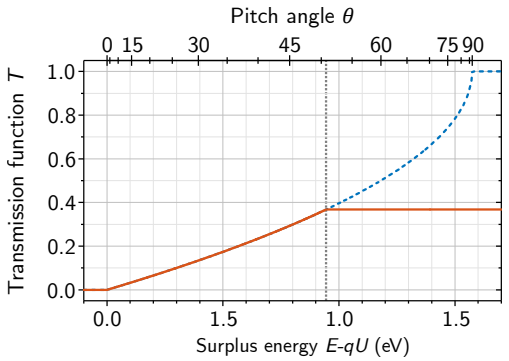
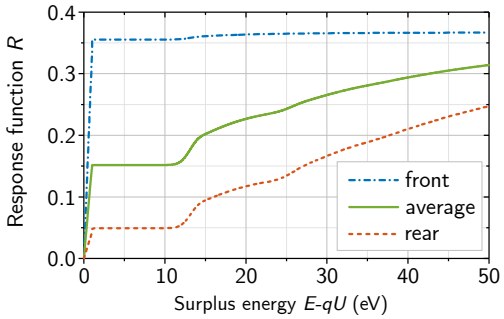


Figure 3.6: Working principle of the MAC-E filter. The graphic is adapted from [KAT05; Are16]. Electrons from the source (green) are guided along the magnetic filed lines (blue) and travel against the electric field (red) generated by electrodes inside the spectrometer.



(a) Transmission probability for β-electrons. With (red line) and without (blue dotted line) magnetic reflection for pitch angles larger than $\theta_{\max} = 50.8°$.



(b) Response function for β-electrons depending on their starting point in the WGTS. The height of the plateau corresponds to the probability for no scattering.

Figure 3.7: Transmission function of the main spectrometer and response function of the full apparatus, taken from [Kle18]. They specify the fraction of transmitted electrons as a function of the surplus energy over the retarding potential qU. The response function is a convolution of the transmission function with an energy loss function weighted with scattering probabilities.

field lines. This is achieved with the MAC-E filter technique which stands for magnetic adiabatic collimation and electrostatic retardation. This principle is illustrated in figure 3.6. Electrodes along the walls inside the vessel create an electric retarding potential which has its maximum in the so-called analysis plane in the center of the spectrometer. Two superconducting solenoids at the ends of the spectrometer tank generate an inhomogeneous magnetic guiding field that is strong at the ends and several orders of magnitude weaker in the analysis plane. The magnetic field inside the pinch magnet, located at the end of the main spectrometer, is referred to as $B_{\max}$ and the analysis plane magnetic field as B_{A}. A system of air coils is used to precisely shape the magnetic flux tube inside the main spectrometer and cancel the Earth's magnetic field as well as remnant magnetic fields from surrounding ferrous metals.

Electrons from the source travel on helix trajectories along the magnetic field lines. It is useful to split their kinetic energy E into transverse and parallel components $E_\perp = E \sin^2\theta$ and $E_\parallel = E\cos^2\theta$ where $\theta \sphericalangle(\vec{p}, \vec{B})$ is called the pitch angle. Their transverse energy is then transformed into momentum parallel to the magnetic field lines due to the magnetic field gradient. This can be motivated using the adiabatic invariant

$$(\gamma + 1)\ \frac{E_\perp}{B} = \text{const.}\ , \tag{3.7}$$

which remains constant for small changes in magnetic field strength. Here, $\gamma = \frac{E}{m_{\mathrm{e}}} + 1$ is the relativistic Lorentz factor which is approximately equal to one since the kinetic energy is much smaller than the electron mass m_{e}. As B gets smaller toward the analysis plane, $E_\perp$ has to decrease accordingly and gets transferred to parallel momentum due to energy conservation $E = E_\parallel + E_\perp$ in adiabatic transport. In the analysis plane,

where almost all the electron's momentum is directed parallel to the magnetic field, only electrons with $E_{\parallel} > qU$ (q: electron charge, U: retarding voltage) can pass the electric potential barrier. As a result, the spectrometer acts as an integrating high-pass filter with a step-like transmission function as shown in figure 3.7a. After crossing the analysis plane the electrons get reaccelerated and their pitch angle reverts to its original state. The spectrometers energy resolution depends only on the ratio of magnetic fields in the analysis plane and the pinch magnet.

$$\frac{\Delta E}{E} = \frac{B_{\mathrm{A}}}{B_{\mathrm{max}}} \ . \tag{3.8}$$

The limiting factor for MAC-E filter resolution is the necessary size to achieve large differences in magnetic field strength. The pre-spectrometer has an energy resolution of about 300 eV at E_0, which is sufficient for limiting the β-electron flux into the main spectrometer. The much larger main spectrometer has an unprecedented sub-eV resolution of $\Delta E = 0.93\,\mathrm{eV}$ at E_0.

In order to observe a neutrino mass signal in the measured integral β-spectrum, a theoretical model of the full apparatus, which includes all relevant systematics, is necessary. This model can be expressed as a convolution of the differential β-spectrum $\frac{d\Gamma}{dE}$ with a response function $R(E, qU)$. Neglecting some details, the number of events recorded at a given retarding potential qU is

$$N = N_{\mathrm{T}} t_U \int_{qU}^{\infty} \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E, E_0, m_\nu) \cdot R(E, qU)\,\mathrm{d}E \ + \Gamma_{\mathrm{b}} t_U, \tag{3.9}$$

where N_{T} is the number of tritium atoms in the source, t_U the time spent at one specific retarding potential, and Γ_b the background rate. The response function itself is again a convolution of the transmission function with an energy loss function which describes the energy that is lost from β-electron scattering off gas molecules in the source. The resulting function is depicted in 3.7b. For a more precise description including corrections to the β-spectrum, source volume and detector segmentation, and radial inhomogeneity of the retarding potential see [Kle18]. To reduce effects from electron gas scattering, electrons with high pitch angles are suppressed because they travel much longer distances within the source and are therefore more likely to collide with gas molecules. This suppression is achieved by reflection through the magnetic mirror effect at the strong field of the pinch magnet. β-electrons emitted at angles less than the acceptance angle θ_{max} are not reflected. It is given by

$$\sin\theta_{\mathrm{max}} = \sqrt{\frac{B_{\mathrm{S}}}{B_{\mathrm{max}}}}, \tag{3.10}$$

where B_S denotes the magnetic field in the WGTS. With standard operation parameters $B_S = 3.6\,\mathrm{T}$ and $B_{\mathrm{max}} = 6.0\,\mathrm{T}$ equation 3.10 yields $\theta_{\mathrm{max}} = 50.8°$.

3.2.4 Focal-plane detector

Electrons passing the potential barrier of the main spectrometer are guided towards the focal-plane detector [Ams15]. A sketch of the system surrounding the detector is shown in figure 3.8. The detector itself is a monolithic 148-pixel silicon p-i-n diode wafer. It is located within a magnet that adjusts the magnetic flux tube to fit the 90 mm diameter of the active detection area. Furthermore, it is encased by a lead and copper radiation shielding and a cosmic muon veto system. The front face of the detector is operated in UHV conditions because it is directly connected to the main spectrometer vacuum with a gate valve. A post-acceleration electrode directly in front of the wafer can be used to increase the energy of the electrons by up to 30 keV in order to shift the measured spectrum into a region with less background.

As depicted in figure 3.9, the 148 equal-area pixels are arranged in twelve concentric circles around four bull's eye pixels, each ring consisting of twelve pixels. Arranging the pixels in this matter allows for corrections of radial electric and magnetic inhomogeneity in the analyzing plane. Each pixel maps onto a corresponding section of the WGTS and the pixels can slightly vary in performance. Thus, the pixels are regarded as independent and each records its own specific tritium spectrum during scans. These scans are later combined in the analysis.

When charged particles hit the substrate of a pixel they scatter and generate a charge cloud that is electrically drifted towards the contacts. These are connected to wires with pogo pins and the wires are fed through a flange to a separate high vacuum chamber where front-end electronics modules are located. Here, the detector signal is collected using charge-sensitive preamplifiers. The modules are separated from the UHV of the detector because they are not sufficiently vacuum safe due to their outgassing. The signal is then passed to the ambient-air electronics outside the vacuum chamber and shielding, where it is further amplified and sent to the data acquisition system (DAQ) via plastic-optical-fiber links.

The DAQ hardware is custom made and modeled after existing hardware used for the Pierre Auger Observatory. The amplified detector signal is fully digitized with analog to digital converters (ADC) and processed with digital filtering algorithms running on FPGAs (field-programmable gate arrays). Two stages of trapezoidal filtering are employed to shape the event pulses for noise reduction, triggering, and energy readout from the pulse height. The DAQ software ORCA (Object-oriented Real-time Control and Acquisition) offers three modes of operation that differ in the amount of recorded information and thus data size. A "trace"-mode allows recording of full signal waveforms, an "energy"-mode provides only event energies and timestamps, and a "histogram"-mode groups events into energy bins.

An example of a spectrum is depicted in figure 3.10a. In a scan of the tritium β-spectrum, the events for each retarding potential setting are counted within a certain region of interest of the detected energy spectrum. Since the energy measurement is performed by the main spectrometer, the energy resolution of the detector itself is not as important. For standard KATRIN operation a resolution of $1.5 - 2.5$ keV is sufficient.

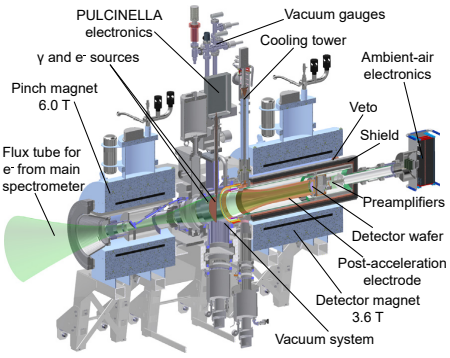


Figure 3.8: Components of the focal-plane detector system. Taken from [Ams15].

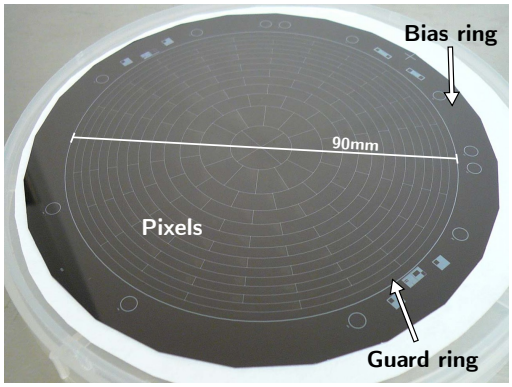
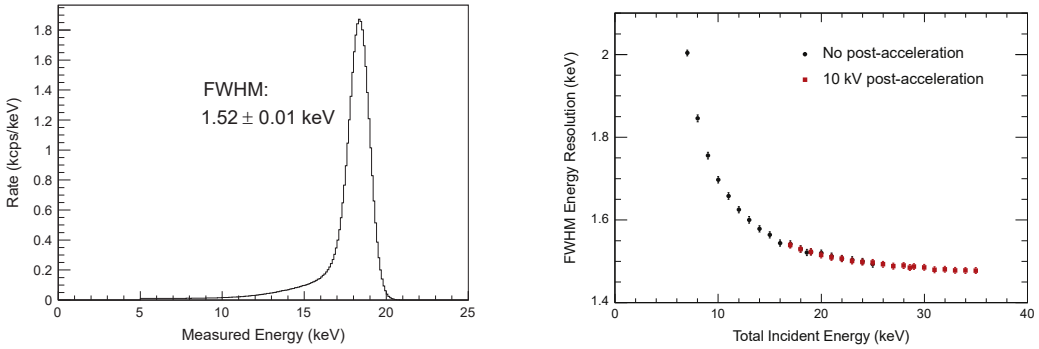


Figure 3.9: Backside of the detector wafer with visible segmentation into 148 pixels [KAT17].

Another key characteristic of a detector and readout system is its timing resolution as it plays an important role in time-of-flight (ToF) measurements [Ste13]. In this type of measurement, the energy of an electron is determined by the time it takes to travel a given distance along the KATRIN beamline. Electrons that barely pass the potential barrier of the main spectrometer are significantly slower than higher energetic ones and their difference in time-of-flight can be resolved. One application is the measurement of the energy loss function with electrons from an electron gun (eGun) located in the rear section. Here, the pulsed eGun provides a start signal and the time difference to their arrival at the FPD is measured.

The timing resolution of the FPD is on the order of 100 to 400 ns, depending on the electron energy and the shaping parameters of the DAQ. This is shown in figure 3.11a, where a pulsed photon source is used for measuring the width of the distribution of time differences between electron emission and detection for several shaping length settings. The timing uncertainties stemming from the detector itself, such as charge collection in the p-i-n diode and FET response time, are much lower than the 50 ns ADC sample spacing and therefore not as significant [Ams15].

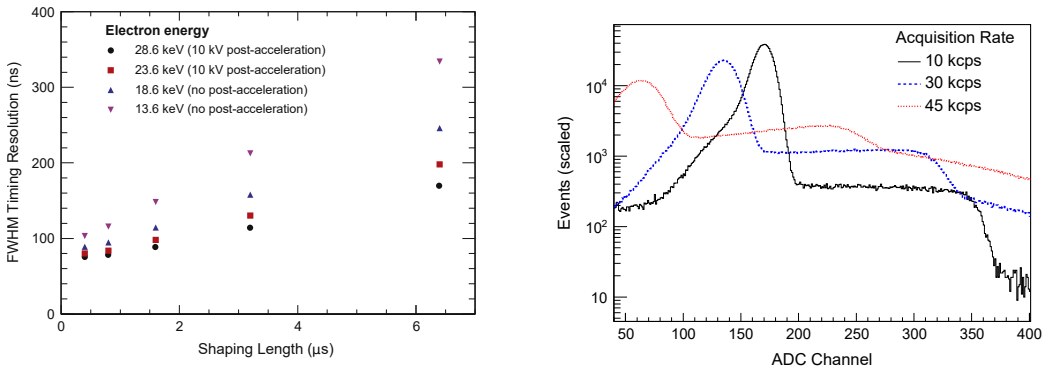
The high rate behavior of the FPD is demonstrated in figure 3.11b. Pileup effects significantly deteriorate the measured energy spectrum for rates in the kcps-range. A broad spectral component next to the main peak is developed due to peak pileup and the spectrum is shifted to lower energies due to tail pileup. Both effects are discussed in detail in chapter 6. For neutrino mass runs the expected rate is much lower and high rate effects do not play a major role. However, for extensive sterile neutrino searches the detector needs to be able to handle rates of $\sim 10^8$ cps. Hence, the FPD system is unsuitable for this task and needs to be upgraded. This upgrade is presented in the following chapter.



(a) Global energy spectrum for 18.6 keV electrons.

(b) Mean FWHM resolution of the FPD from electrons at two post-acceleration settings.

Figure 3.10: Energy measurement performance of the FPD as determined by [Ams15]. The measurements were performed over 146 of 148 channels since two channels were defective at the time.



(a) Timing resolution of the FPD system depending on the pulse shaping length of the DAQ.

(b) High rate behavior of the FPD for 18.6 keV photoelectrons at three different rates. The spectra are normalized to equal area for easier comparison.

Figure 3.11: Timing and high rate performance of the FPD as determined by [Ams15].

CHAPTER 4

The TRISTAN project

The TRISTAN project (**TR**itium **I**nvestigation on **ST**erile (**A**) **N**eutrinos) is an extension of the KATRIN physics program scheduled to take place after the KATRIN neutrino mass runs are completed. The aim is to search for the signature of sterile neutrinos with masses in the keV range by measuring the tritium β-decay spectrum over a large energy range. For this purpose, the energy threshold set by the main spectrometer is chosen much lower than for neutrino mass measurements, where only the endpoint region is scanned. Leveraging the high intensity of the KATRIN tritium source enables the accumulation of high statistics within three years of data taking.

Although it is to some degree possible to search for sterile neutrinos with the nominal KATRIN operation, the full potential of the tritium source cannot be realized with the current detector system. Using the high source activity of $\approx 10^{11}$ Bq, within three years sufficient data can be taken to probe mixing angles down to $\sin^2\theta < 10^{-6}$ [Mer18]. However, this requires a detector that is able to measure rates of $10^8 - 10^9$ counts per second (cps) without significant degradation of the acquired spectrum. Furthermore, measuring the differential spectrum, instead of the integral one, yields much higher sensitivity. The primary option to achieve this is to utilize the detector itself for the energy measurement which requires a much better resolution than the KATRIN FPD offers. The TRISTAN group is working on an upgrade of the detector and readout system in order to achieve these goals.

This chapter provides an introduction to the TRISTAN project by describing the measurement principle and statistical sensitivity. Afterward, the detector and DAQ requirements, and the preliminary detector design are presented.

4.1 Sterile neutrino search in tritium β-decay

As discussed in section 3.1, the theoretical tritium β-decay spectrum is a superposition including all neutrino mass eigenstates m_i, weighted by mixing amplitudes $|U_{ei}|^2$ (see equation 3.4). If sterile neutrinos exist, mix with the active neutrinos via oscillation, and are lighter than the endpoint energy $E_0 = 18.6\,\mathrm{keV}$, they contribute to this superposition and leave a signature in the spectrum. In the case of one sterile neutrino in the keV mass range, the resulting spectrum can be written as

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E} = \cos^2\theta \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(m_{\mathrm{eff}}) + \sin^2\theta \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(m_{\mathrm{s}}) \ , \tag{4.1}$$

where m_{eff} is the effective electron antineutrino mass as in section 3.1. In this context it is reasonable to set $m_{\mathrm{eff}} = 0$, because the endpoint cannot be resolved with enough precision in a deep spectrum scan. The mass of the sterile neutrino is denoted as m_{s} and the corresponding mixing amplitude to the electron neutrino state $|U_{\mathrm{es}}|^2$ is expressed as an Euler angle θ. An exaggerated example of this superposition is shown in figure 4.1. The sterile neutrino produces a contribution on top of the original spectrum in the kinematically allowed region. Next to the overall change in shape, the endpoint of the sterile contribution produces a characteristic kink at $E_{\mathrm{kink}} = E_0 - m_{\mathrm{s}}$.

The experimental setup of KATRIN offers three modes of operation to search for keV sterile neutrinos [Mer15a]:

- **Integral:** Similarly to nominal neutrino mass runs, a measurement of the integral beta spectrum can be performed by scanning with the main spectrometer. The key difference is the scan depth. Nominal neutrino mass scans range down to 200 eV below E_0. During deeper scans the rate increases dramatically, resulting in a loss of detection efficiency and an increase of systematic uncertainties of the fit model. Furthermore, the kink signature and the spectral shape modification due to sterile neutrinos are less pronounced than in the differential spectrum (see figure 4.2). Nevertheless, during the first tritium runs several deep scans to 1.6 keV below the endpoint were performed which will already yield first results and help with the development of analysis strategies.

- **Differential:** By setting the spectrometer to a fixed, low retarding potential and measuring the electron energy with the detector itself, a differential tritium spectrum can be attained. This mode is in principle much more sensitive to sterile neutrinos, but it requires a significant improvement of the detector's energy resolution. The goal is a resolution of $\approx 300\,\mathrm{eV}$ FWHM at 20 keV opposed to the $\sim 2.5\,\mathrm{keV}$ FPD resolution, which would obscure any sign of a kink. From a modeling perspective, this mode is a big challenge since previously unaccounted systematics, particularly detector effects that change the shape of the measured spectrum, need to be understood in detail [Mer18].

- **Time-of-flight:** A hybrid time-of-flight (ToF) mode based on the gated-filter method, as described in [Ste13], is considered as well. According to this method,

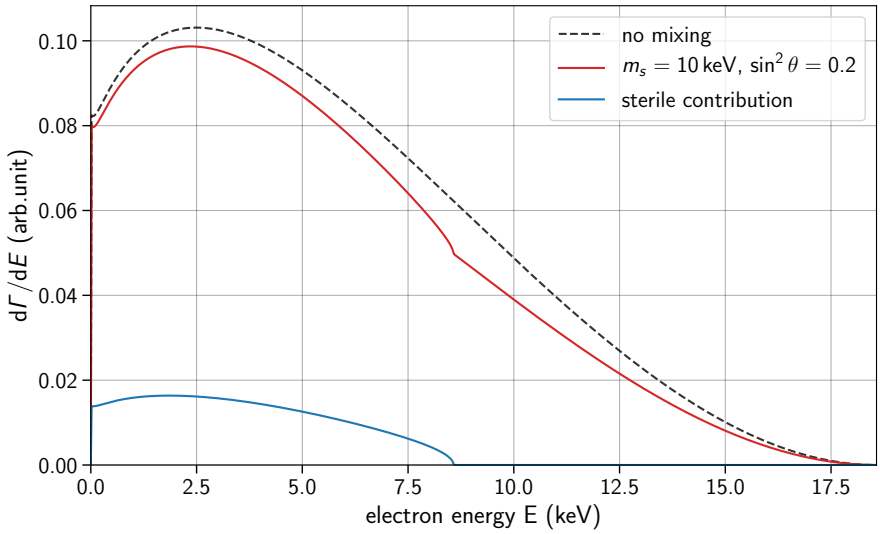


Figure 4.1: Signature of a keV sterile neutrino in the differential tritium β-spectrum for an exaggerated mixing angle of $\sin^2\theta = 0.2$. For low energies the sterile neutrino causes a spectral contribution which ends at the energy $E_{\mathrm{kink}} = E_0 - m_{\mathrm{s}}$, where a characteristic kink emerges.

a gating signal is applied to the pre-spectrometer in order to define sharp time intervals where electrons can pass. Such an interval, if chosen comparatively small, can serve as a start time. The arrival of corresponding electrons at the detector mark stop times for their flight duration. "Slow" electrons with energies slightly above the main spectrometer retarding potential exhibit measurable differences in flight duration. The flight duration of higher energetic electrons cannot be resolved, causing a cutoff for low flight times. The accessible energy region therefore only extends into a small region above the retarding potential, but the whole spectrum can be measured by scanning in the same way as in an integral measurement.

A combination of these modes can be employed for cross-checking because each is prone to different kinds of systematic effects. Regardless of the specific method, a measurement of the full spectrum, or at least as deep below the endpoint as possible, is needed to cover a large sterile mass range. Lowering the main spectrometer retarding voltage in order to let less energetic electrons pass, vastly increases the rate at the detector. The full rate at nominal column density is $\approx 1.5 \cdot 10^{10}$ cps with consideration of the acceptance angle. Compared to the integral rate at 100 eV below the endpoint (several kcps), the increase is larger than a factor 10^6. Although the amount of tritium in the source can be decreased to lower the rate, a sufficiently high rate is necessary to meet the design goal of probing mixing amplitudes $\sin^2\theta < 10^{-6}$ [Mer18]. This can be achieved with a rate of $\sim 10^8$ cps at 100-fold reduced column density and three years of effective data taking. As an added benefit, scattering in the tritium source is also reduced by lowering the column density. Overall, the rate requirements pose a major challenge to the detector and data acquisition system. The current detector system is not capable of handling these high

rates, as it suffers from severe pileup for rates above several kcps (see section 3.2.4 and [Ams15]).

To summarize, the main requirements for a detector system suitable for sterile searches are:

- Capability of handling high rates without major deterioration of the measured energy spectrum. The goal is good performance up to 10^8 cps.
- Excellent energy resolution with $\Delta E < 300\,\text{eV}$ FWHM at 30 keV to reduce kink smearing.

For the analysis of the data, two complementary approaches have been proposed: A spectral fit method [Mer15a], similar to the KATRIN neutrino mass analysis, and a wavelet decomposition [Mer15b]. The spectral fit requires an accurate model of the entire KATRIN apparatus, which can be expressed as a convolution of the theoretical β-spectrum with all relevant processes that change the detected spectrum. Known systematic uncertainties can be included as nuisance parameters. This model can then be evaluated for different mixing hypotheses and compared to the data, for example by calculating the χ^2 value

$$\chi^2 = \sum_{ij} (N^i_{\text{obs}} - N^i_{\text{exp}}) V^{-1}_{ij} (N^j_{\text{obs}} - N^j_{\text{exp}}) \ . \tag{4.2}$$

Here, N^i_{obs} is the observed number of events in energy bin i, N^i_{exp} the expected number from the model and V^{-1}_{ij} the inverse covariance matrix containing statistical errors and bin-to-bin correlations. An exclusion contour, similar to figure 2.5, can be obtained by scanning the (m_s, θ)-parameter space and requiring $\chi^2 > 5.99$ for an exclusion with 95 % confidence. When done correctly, this method is very sensitive because it takes full advantage of the spectral distortion below the kink. A study by [Mer18] yields a statistical sensitivity of $\sin^2\theta < 10^{-6}$ for a dataset of 10^{16} β-electrons, which corresponds to three years of data taking at 100-fold reduced column density. The corresponding exclusion contour is depicted in figure 4.3.

The wavelet decomposition, as described in [Mer15b], is a complementary approach designed to specifically search for a kink signature. Similar to a Fourier transform, the measured spectrum is expressed as a superposition of wave packets. The frequency power spectrum then reveals the location and intensity of potential kinks. Judging by statistical sensitivity alone, the wavelet approach is not as powerful as the spectral fit, because the spectral distortion below the kink does not contribute. The sensitivity strongly depends on the energy resolution which smears the kink, but it is robust towards systematics that can be expressed as smooth modifications of the tritium spectrum. The major advantage of the wavelet approach is its model-independence, since no precise knowledge of the spectrum is needed.

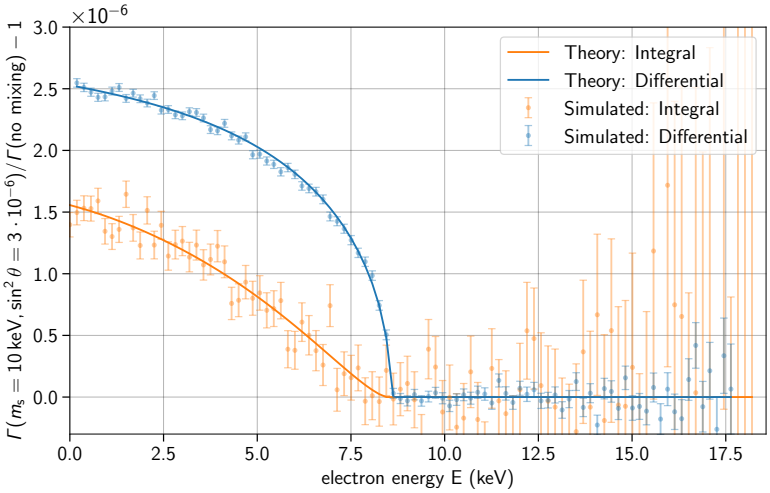


Figure 4.2: Comparison of the keV sterile neutrino signature for a differential and an integral measurement. A β-electron energy spectrum containing a sterile neutrino with $m_s = 10\,\mathrm{keV}$ and $\sin^2\theta = 3 \cdot 10^{-6}$ is shown. For increased visibility of the kink and spectral distortion, it is normalized to the expected spectrum in the "no-mixing" case. The simulated data points are drawn from a Poisson distribution around the expectation and correspond to an estimated dataset of 10^{16} β-electrons for three years effective data taking at $\sim 10^8$ cps. The statistical errors are larger for the integral spectrum, because the measurement time is equally distributed over all data points.

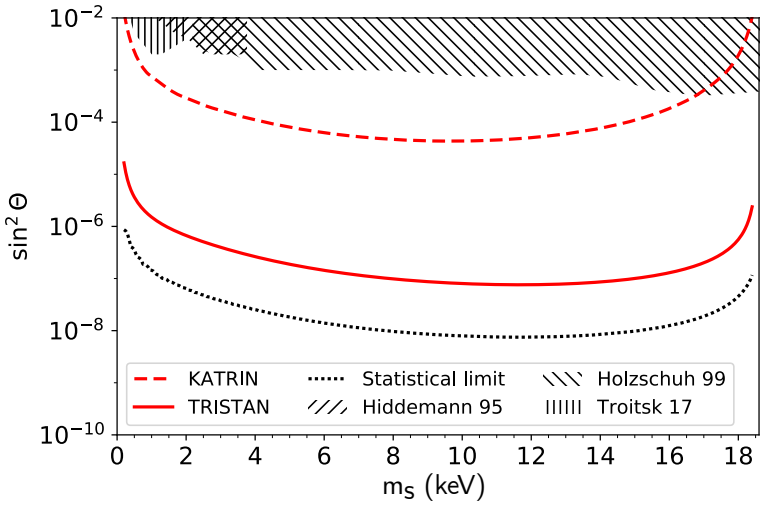


Figure 4.3: Statistical exclusion limits (95 % C.L.) and sensitivities taken from [Mer18]. The dashed line shows the sensitivity of KATRIN with the current detector system. The solid red line represents the sensitivity of TRISTAN with an estimated dataset of 10^{16} β-electrons for three years effective data taking at 100-fold reduced column density. The dotted line shows the statistical limit when the full column density is used. Systematic effects are not included. The shaded area shows existing exclusion limits by other laboratory experiments.

4.2 Detector design

The TRISTAN project uses Silicon Drift Detector (SDD) technology, because of its excellent energy resolution, high rate capabilities, scalability, and suitability for use in strong magnetic fields. The preliminary detector design is shown in figure 4.4. The detector is 20 cm in diameter and composed of 21 rectangular modules with 166 SDD pixels each. For maximum area coverage, the 3 mm pixels are hexagonal and arranged in a honeycomb pattern. The modular design allows for easier scaling during R&D and for swapping defective parts without replacing the full detector. Each detector wafer is mounted on a copper block holding structure with two front-end electronics boards on the sides which carry charge sensitive preamplification ASICs. For reduction of electronic noise, the copper blocks are cooled to $-30°\mathrm{C}$.

In the following, some defining characteristics of the detector system are discussed.

Number of pixels
A simple and effective method for detecting particles at high rates is distributing the total flux over a large area and many separate channels. With a total expected rate of 10^8 cps at least 1000 pixels are required to reach a manageable 10^5 cps per channel. The number of pixels is ultimately limited by cost and technical feasibility. The TRISTAN detector will therefore have about 3500 pixels, an increase of factor ~ 23 compared to the 148 pixels of the KATRIN FPD.

Pixel and detector size
Pixel and detector size are optimized in regards to charge-sharing between neighboring pixels, energy resolution, and backscattering and backreflection of electrons. Minimizing charge-sharing requires the pixels to be as large as possible without compromising the energy resolution. The probability for electrons backscattering off the detector, leaving only a fraction of their energy, is governed by their pitch angle (see section 3.2.3). This, in turn, depends on the magnetic field B_{det} enveloping the detector and the post acceleration voltage. Backscattered electrons can be recovered by back reflection at the high magnetic field of the pinch magnet. Preferably these back reflected electrons should hit the same pixel they were scattered at, or at least neighboring pixels. This requires a sufficiently large magnetic field in order to get cyclotron radii close to the pixel diameter. Another constraint is that the overall detector size needs to fit the area A of the magnetic flux tube guiding the β-electrons. The magnetic flux $\Phi = B \cdot A$ is a conserved quantity and defined by the source magnetic field and diameter. To adjust the flux tube to the detector size, B_{det} has to be adjusted accordingly. A simulation-based optimization of these intertwined requirements yields 3 mm pixel size, a total detector area $A = \pi \cdot (10\,\mathrm{cm})^2 \approx 314\,\mathrm{cm}^2$, and $B_{\mathrm{det}} = 0.7\,\mathrm{T}$ [Mer18; Kor16].

Energy resolution
As mentioned in the previous section, an energy resolution $\Delta E < 300\,\mathrm{eV}$ at 30 keV is necessary to resolve the sterile kink signature. With resolutions close to the Fano-limit for silicon, SDDs are capable of fulfilling this requirement even for high count rates. SDDs

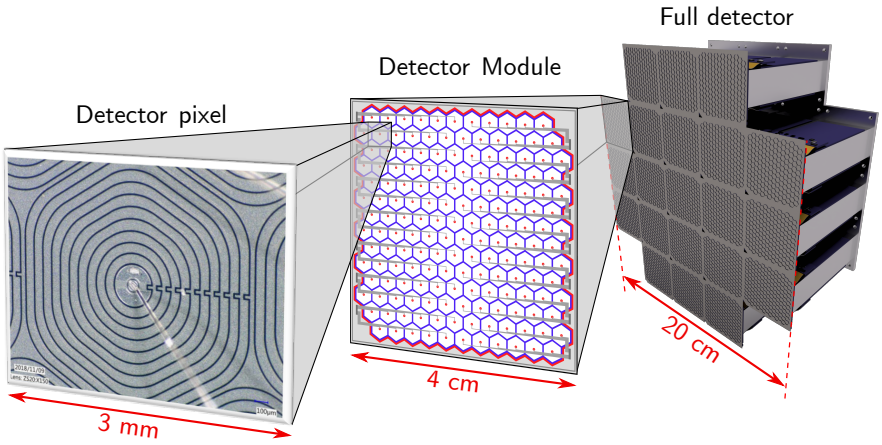


Figure 4.4: Preliminary TRISTAN detector design. The detector is subdivided into 21 modules with 166 SDD pixels each. The detector wafers are mounted on cooling blocks made of copper, which also carry front-end electronics boards on two sides.

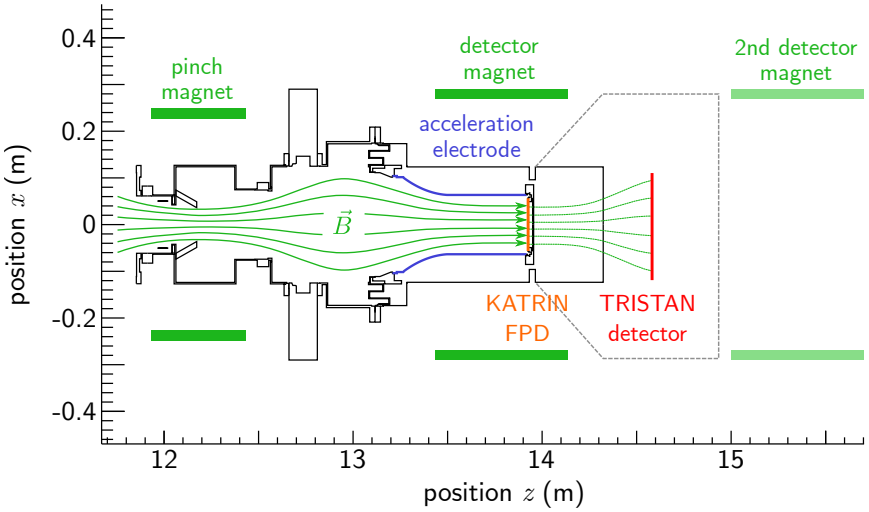


Figure 4.5: Placement of the TRISTAN detector in comparison to the KATRIN FPD. Adapted from [Mer18]. Unlike in the KATRIN case, the TRISTAN detector and front-end electronics are located within the same UHV to avoid feedthroughs between pixels and preamplifaction stage. A second detector magnet can be installed behind the detector in order to get field lines perpendicular to the detector surface.

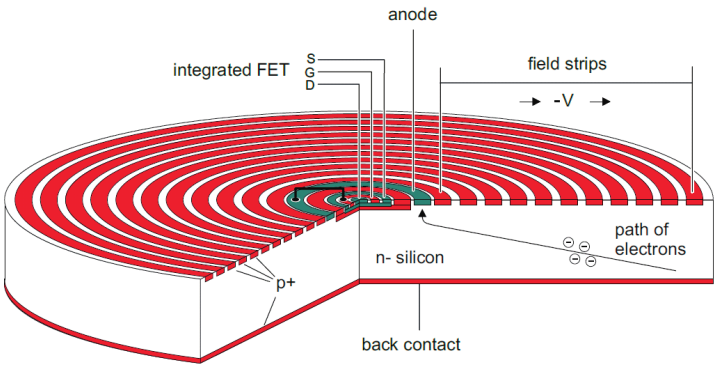


Figure 4.6: Structure of a silicon drift detector [Tho15].

are typically used in X-ray spectroscopy. Applying them to keV electrons introduces the systematic effect of energy loss in the thin layer of non-active material, primarily oxidized silicon, on the detector surface. This region is usually referred to as a dead layer or entrance window. It should be noted that the dead layer cannot be regarded as a homogeneous bulk of non-active material as the name suggests, but rather a transition region with gradually increasing activeness. Photons deposit their energy very close to their point of interaction. Hence, they are likely to either be detected at full energy if they interact past the dead layer, or fall below the detection threshold because most energy is lost in the dead layer. Electrons, on the other hand, are charged particles and will always lose some of their energy while traversing the entrance window. This shifts the measured energy slightly downwards and degrades the energy resolution since the amount of lost energy is a statistical quantity. For an optimal energy resolution, a dead layer $< 100\,\mathrm{nm}$ is necessary [Mer18].

Electronic readout

Each detector pixel is a small silicon drift "chamber" as originally proposed by Emilio Gatti and Pavel Rehak [GR84]. Such a pixel consists of a bulk of n-doped silicon with a p-doped back contact on the entrance window side and an anode on the other side (see figure 4.6). The anode is surrounded by p-doped field strips or drift rings. A high negative voltage is applied between back contact and anode which depletes the p-n junction. When ionizing particles hit the detector, electron-hole pairs are formed. By applying successively larger negative voltages from the inner to outer drift rings, an electric field is created that guides the electrons toward the anode. By drifting the electrons, the active detector area is much larger than the almost point-like anode, which results in a low capacitance and good energy resolution. In order to reduce electronic noise, an n-type junction gate field-effect transistor (nJFET) is integrated directly into the pixel with its gate connected to the anode. It acts as the first amplification stage allowing for a couple of centimeters distance between the pixel and the front-end preamplification ASIC. This alleviates technical challenges for the design of a large pixel array.

Vacuum safety
Since the detector is operated in ultrahigh vacuum conditions, a low outgassing rate of all components within the chamber is necessary. The total outgassing rate ultimately limits the pressure that can be reached. Furthermore, outgassed molecules such as water and carbohydrates can condense on the cooled detector surface and form an additional dead layer. Electronic parts are especially prone to large outgassing and materials must be selected accordingly. For this reason, the KATRIN FPD front-end electronics are located in a vacuum separated from the detector and connected via electrical feedthroughs. For TRISTAN this approach would prove difficult due to the large number of channels. The currently favored option is reducing the outgassing through material selection and verification with dedicated outgassing measurements.

Data acquisition
The amplified detector signals are fed through to the ambient air electronics and passed to a data acquisition system (DAQ). For the reduction of nonlinearities in the energy measurement [Dol17], the signal waveform is fully digitized with an analog-to-digital converter (ADC) and processed by a digital pulse processor (DPP). The latter employs digital filters to shape the signal pulses for a more precise energy readout. A trigger logic for on-line rejection of pileup and charge-sharing events can be implemented, which reduces the amount of output data and increases the throughput. An ADC sampling rate of $\sim 100\,\mathrm{MHz}$ enables the DAQ to exploit the full potential of the SDD in terms of timing resolution and pileup rejection. The DAQ system that will be used for TRISTAN is not yet finalized. During the ongoing prototype stage with 7-pixel detectors, commercial systems are being investigated and an in-house solution is developed in parallel.

4.3 Prototype detector setup

While the development of a full 166-pixel module is ongoing, the current stage of the TRISTAN R&D revolves around 7-pixel detector prototypes produced by the Semiconductor Laboratory of the Max-Planck-Society (HLL) in Munich [HLL]. For this detector, special production steps are taken in order to minimize the entrance window thickness. The electronics and the data acquisition system used for operating the detector are provided by the company XGLab from Milano, Italy [XGLa]. A photo of the prototype setup is shown in figure 4.7.

The prototype system's main purpose is the characterization of detector properties such as dead layer and energy resolution in order to examine the aptitude of the SDD pixel array design for keV sterile neutrino search. Meanwhile, it is also utilized for studying other systematics such as charge-sharing, backscattering, and pileup. These studies lay the groundwork for constructing a detailed description of the detector and DAQ response function which represents the final part of the model for the experiment. Such a detector response function can be convolved with the spectrum of incident electrons for calculating the expected output spectrum. With a complete model of the experimental apparatus, the observed data can be compared to the model expectation and checked for accordance with various sterile mixing parameters (see section 4.1). In chapter 6 several modeling approaches for the rate dependent systematic effects dead time and pileup are presented.

Overall, the prototype achieves very good benchmarks with an energy resolution of 139 eV FWHM at 5.9 keV for photons from an ^{55}Fe source [Mer18]. The optimal energy resolution is realized with cooling to -30° C and a comparably short peaking time of $\approx 1\,\mu$s for the trapezoidal pulse shaper. A short peaking time is crucial for TRISTAN because it is related to the dead time window after an event trigger. At high rates per pixel of $\sim 10^5$ cps, even a short peaking time of 1 µs already leads to $\sim 18\,\%$ dead time, which represents a significant loss of statistics. For a detailed discussion of dead time see chapter 6.

Apart from the characterization with ^{55}Fe photons, the prototype system is currently being tested with other particle sources. These include electron guns, an electron microscope [Sie19], radioactive sources such as ^{241}Am and $^{83\mathrm{m}}$Kr, and a compact electrostatic proton accelerator. Furthermore, the system was successfully utilized for pilot sterile neutrino searches at KATRIN's predecessor experiment in Troitsk.

The 7-pixel prototype used for this thesis differs slightly from the final pixel design as it has a 2 mm pixel diameter, opposed to the 3 mm design diameter, and it has no integrated nJFET. Instead, the charge sensitive preamplifier ASIC (CUBE by XGLab) containing the FET is located on the detector board and is bonded directly to it (see figure 4.8). A bias board is used for operating preamplifier and detector. It supplies the detector with the voltages for the back contact and back frame as well as the inner and outer drift rings. The preamplifier is an active reset type and needs external circuitry for resetting its voltage if the dynamic range is exhausted. Reset threshold voltage, offset voltage, reset inhibit duration, and the detector voltage supply can be adjusted with

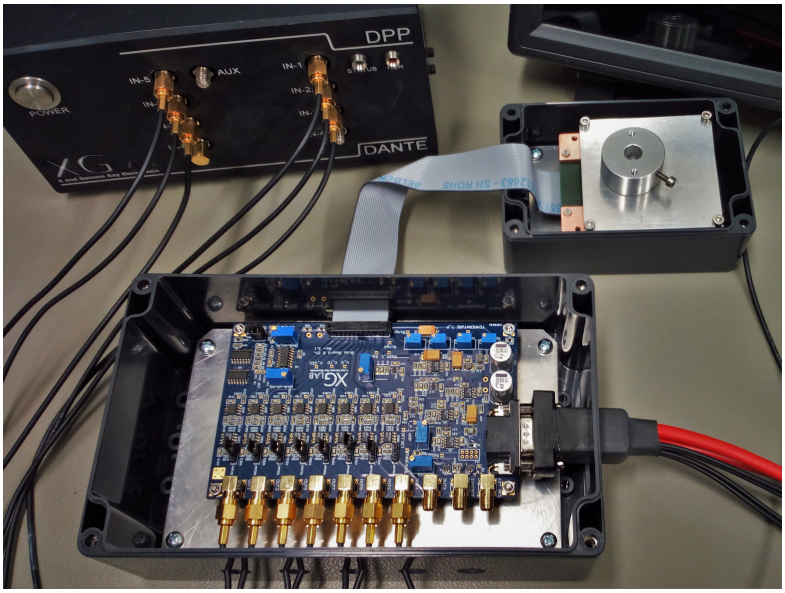


Figure 4.7: Seven pixel prototype detector setup at ambient air conditions for radioactive source measurements. The detector board (top right) is shielded with an aluminium aperture and irradiated with ^{55}Fe photons. A ribbon cable connects the detector to the bias board, which is supplied with voltage through a D-sub connector from the right. The channels are linked to the DANTE DAQ system (top left) with SMA connectors.

potentiometers on the board. It also includes further amplification and buffering stages before passing the signal waveforms on to the DAQ system.

The DAQ system is an 8-channel digital pulse processor (DPP) called DANTE. For each channel a 16-bit ADC samples the waveform at 125 MHz for an 8 ns sample spacing. The sampled waveform is analyzed with an FPGA that employs trapezoidal filters for pulse shaping and triggering. The pulse processor is connected to a PC via USB or Ethernet and the DAQ firmware can be configured either with a GUI or an API. The measured event energies and timestamps, as well as other information such as live time and reset count, are transmitted to the PC. Since it is a major topic of this thesis, a more detailed description and characterization of the DANTE DPP follows in chapter 7.

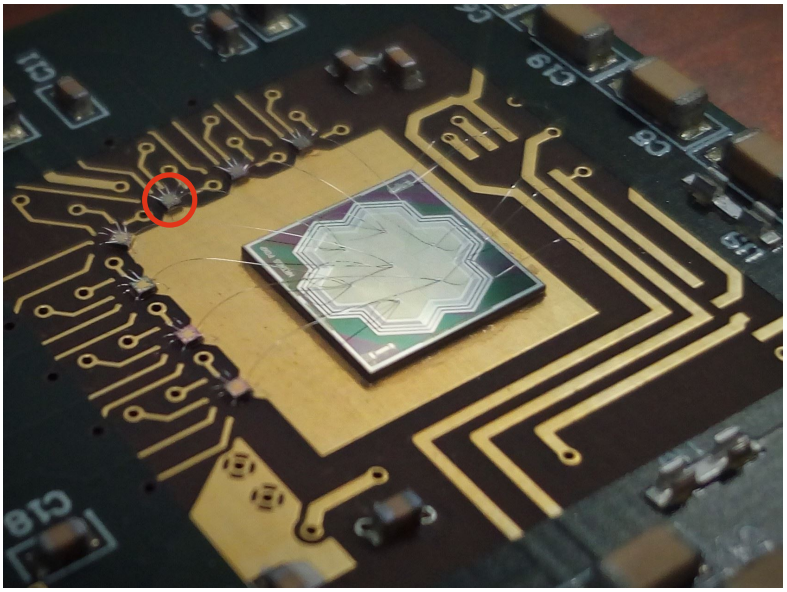


Figure 4.8: Backside of the detector board without its protective cover. The pixel anodes are bonded directly to the adjacent CUBE preamplifier ASICs (one marked in red).

CHAPTER 5

Data acquisition

At a basic level, data acquisition is a method of extracting information about features of a sensor signal that may already be amplified and shaped by the preceding circuitry. In modern data acquisition systems, it is common to fully digitize the signal waveform and process it with digital filters for determining features such as the pulse height and timing information. In physics, the pulse height is often related to an individual particle's energy, whereas timing information allows separating different particle signals from each other. While it is possible to record the waveform as it is and perform a detailed off-line analysis for extracting these features, the amount of data would quickly exceed the transmission, computing, and storage capacities for detectors with a large number of channels such as the TRISTAN SDD array. Realistically, on-line processing is required to reduce the amount of data to a small number of interesting features like energy and timestamps of events. Field programmable gate arrays (FPGAs) are a natural choice for this task because they provide high bandwidth and throughput while remaining configurable unlike an application-specific integrated circuit (ASIC). A device that combines the analog-to-digital conversion (ADC) with the FPGA-based waveform analysis is referred to as a digital pulse processor (DPP) or digitizer.

This chapter discusses some key concepts of pulse height analysis through digital pulse processing. This is the foundation for a discussion of rate dependent systematics in chapter 6 and the characterization of the DANTE DPP in chapter 7.

5.1 Signal waveform

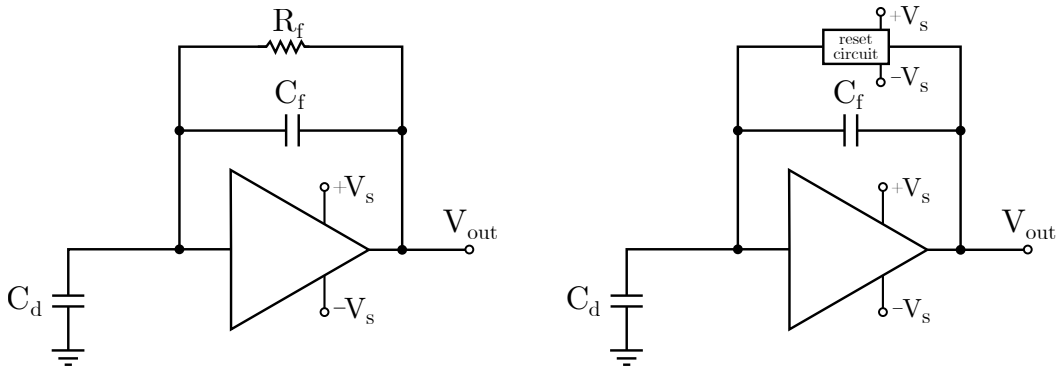
The waveform created from the amplified detector signal is highly dependent on the type of preamplifier. For particle detectors, these are commonly charge-sensitive types that collect the charge created in the detector on a capacitance in the feedback loop of an amplifier (see figure 5.1). The output signal is then the integral of the detector current. Events in the detector cause steps in the output signal which can be further amplified and then processed by the DAQ. As the event energy is proportional to the charge created in the detector, and therefore the capacitor voltage, the energy is encoded in the step height. It should be noted that these steps have a non-zero rise time which is determined by the detectors charge collection time and intrinsic capacitance. This signal rise time is usually defined as the time it takes for the signal to increase from 10 % to 90 % of the step amplitude and it is on the order of several 10 ns for SDDs.

For sustainable operation, the amplification circuit requires some mechanism for resetting the feedback capacitor voltage. In RC-preamplifiers the capacitor is continuously discharged via a parallel resistor in the feedback loop as depicted in figure 5.1a. Hence, they are often referred to as *continuous reset preamplifiers*. After the step caused by an event, there is an exponential decrease of the pulse with a characteristic time constant $\tau = R_\mathrm{f} C_\mathrm{f}$, where R_f and C_f are the feedback resistance and capacitance. On the one hand, this constant needs to be much larger than the rise time of the step in order to acquire a correct energy reading. On the other hand, for high rates, successive pulses will overlap and increase the mean voltage to an equilibrium level. Therefore, τ must be chosen low enough to prevent saturation of the signal. Typical time constants are in the 10 µs-range. A simulated example of a continuous reset waveform is shown in figure 5.2a.

Another possibility for resetting the preamplifier voltage is a quick active discharge controlled by an additional circuit (see figure 5.1b). Such a reset signal can, for instance, be provided in the form of a periodic pulse, or when some voltage threshold is crossed. Accordingly, these preamplifiers are called *active or pulsed reset preamplifier*. Since there is no exponential tail as for a continuous reset preamplifier, the pulses are just steps in the voltage, but a tail can still be introduced by an AC-coupling in the subsequent signal chain. Apart from the event steps and the negative reset edge, the example waveform depicted in figure 5.2b exhibits a linear increase caused by non-zero leakage current through the detector.

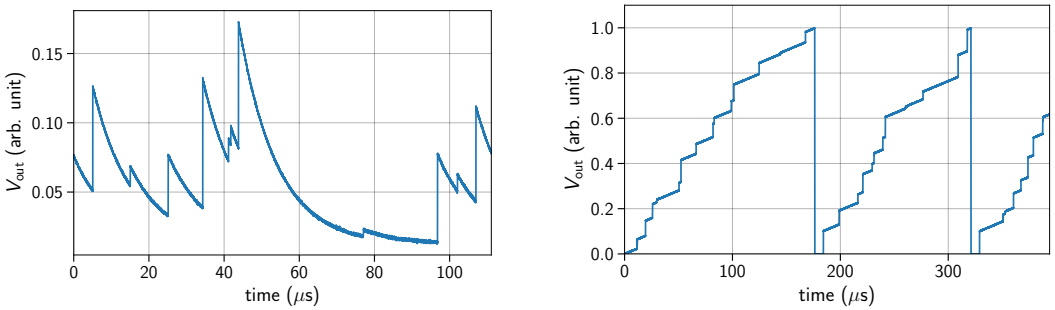
While the preamplifier used for the KATRIN FPD is of the continuous reset type, the CUBE ASIC used for the 7-pixel TRISTAN prototype setup is an active reset preamplifier. It resets the voltage when any of the seven channels detects the crossing of a reset threshold which can be adjusted via the bias board.

Regardless of the type of preamplifier, in digital pulse processing the signal waveform is digitized by an ADC with a sufficiently high resolution and sampling frequency for the specific application. The pulse heights are extracted by shaping the pulses and applying a triggering and readout algorithm. The following section introduces the most common pulse shaping filter employed in digital pulse height analysis, namely the trapezoidal filter.



(a) Continuous reset. **(b)** Active reset.

Figure 5.1: Schematic circuit diagrams for charge-sensitive preamplifiers.



(a) Continuous reset with $\tau = 10\,\mu s$. **(b)** Active reset when V_{out} reaches 1.

Figure 5.2: Simulated waveforms of charge-sensitive preamplifiers for a rate of 10^5 cps and energies drawn from a flat distribution.

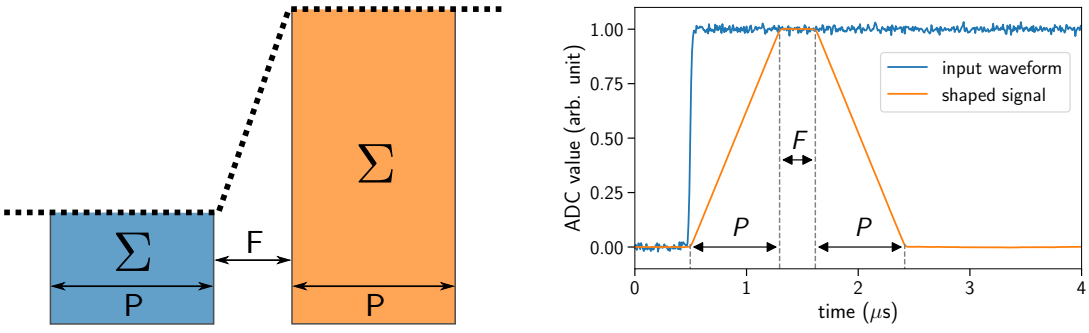
5.2 Trapezoidal filter

For a reliable determination of the event energies from the waveform, the pulses need to be shaped. Digital pulse processors usually rely on the trapezoidal filter. This filter is designed to turn a step signal into an isosceles trapezoid with a defined width and the same height as the step. Simultaneously, it smoothens noise which improves the energy resolution. It does so by averaging the sampled signal over a given range within the *peaking time* P. It then calculates the difference of two such regions which are separated by a gap of F samples, which is called the *flattop time*. This is repeated for every sample point. If the signal is expressed by a vector of samples $x[i]$, where i is the sample number, the shaped signal $y[i]$ can be calculated according to

$$y[i] = \frac{1}{P}\left(\sum_{j=0}^{P} x[j] - \sum_{j=P+F}^{2P+F} x[j]\right) \quad . \tag{5.1}$$

Here, the two sums represent the averaging regions. Normalizing with $\frac{1}{P}$ corresponds to a gain equal to one which ensures the same height of input signal and shaped pulse. A visual representation of this equation is given in figure 5.3a, and figure 5.3b shows a simulation of a pulse, shaped from a step, as it would occur for an active reset preamplifier. The latter makes it evident where peaking and flattop time get their names because they represent the time it takes for the pulse to reach full height and the length of the trapezoid's upper side.

Both parameters require optimization relating to the specific application. Since noise is averaged during the peaking time, its length has an impact on the energy resolution which improves for longer peaking times. However, a short shaped pulse is desirable for minimizing pileup and dead time. While the flattop time can be chosen much smaller than the peaking time, some gap between both averaging regions is necessary to accommodate for non-zero signal rise times. As discussed in chapter 6, the flattop can have implications on the pileup spectrum's shape as well.



(a) Schematic drawing of the trapezoidal filter's working principle.

(b) Simulated trapezoidal filter response to a step waveform.

Figure 5.3: Trapezoidal filter working principle and response.

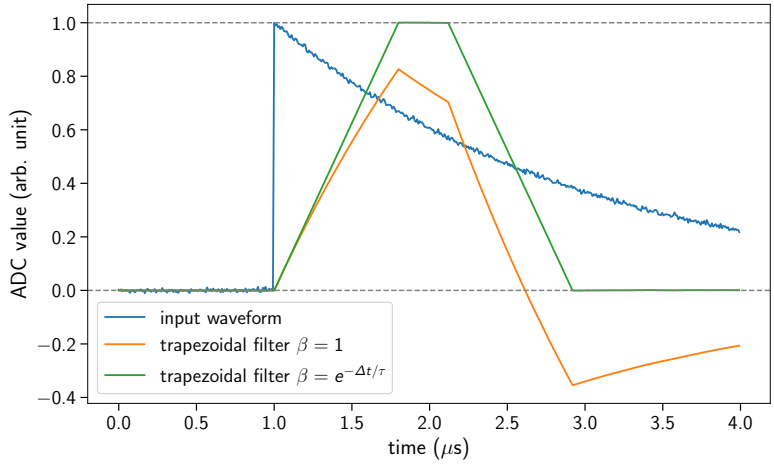


Figure 5.4: Impact of a pole-zero correction on the shaped pulse. For demonstration purposes the time constant of the pulse tail is chosen very small ($\tau = 2\,\mu s$). If no pole-zero correction is applied ($\beta = 1$) the shaped pulse is distorted, doesn't reach full height and undershoots.

Another representation of the trapezoidal filter is by its transfer function in the Z-transform domain as derived by Zbigniew Guzik and Tomasz Krakowski [GK13]:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{P}(1 - z^{-P}) \cdot (1 - z^{-(P+F)}) \cdot \frac{1 - \beta z^{-1}}{1 - z^{-1}} \cdot \frac{z^{-1}}{1 - z^{-1}} \ . \tag{5.2}$$

This form is useful for implementation in simulation code and is used for all waveform simulations in this work. When the transfer function is expanded and ordered by powers of z^{-n}

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_m z^{-n}}{a_0 + a_1 z^{-1} + \cdots + a_n z^{-n}} \ , \tag{5.3}$$

where the coefficients a_j in the denominator and b_j in the numerator can be used for a very efficient recursive calculation of $y[i]$ through

$$y[i] = \frac{1}{a_0} \left(\sum_{j=0}^{m} b_j x[i-j] - \sum_{k=1}^{n} a_k y[i-k] \right) \ . \tag{5.4}$$

Albeit very similar to equation 5.1, there are two key differences. One is the aforementioned recursiveness and the other is the factor $\frac{1-\beta z^{-1}}{1-z^{-1}}$ in equation 5.2, which corresponds to a pole-zero cancellation. For pulses with an exponential tail, the trapezoidal filter in the form of equation 5.1 produces a distorted pulse and slightly underestimates the energy, because the slope right after the pulse is steeper than before the pulse. A demonstration of this is shown in figure 5.4. If β is set equal to $e^{-\Delta t/\tau}$, where τ is the time constant of the pulse tail and Δt the sample spacing, this effect is corrected. For $\beta = 1$, which corresponds to an infinite time constant, equation 5.4 is equivalent to 5.1.

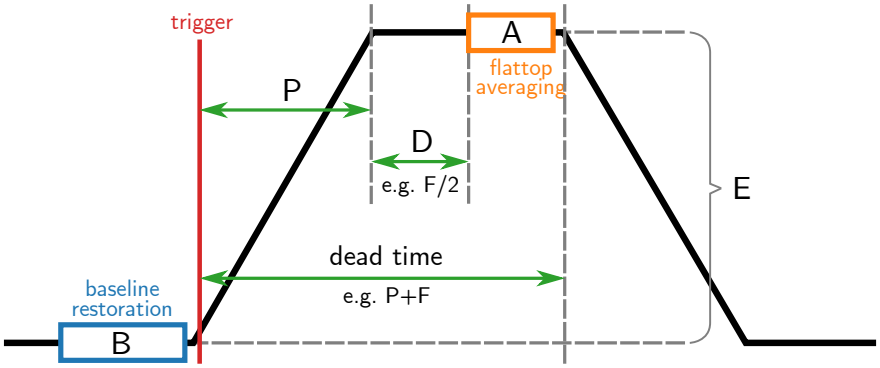


Figure 5.5: Schematic of the energy readout from the trapezoidal filter pulse. If the trigger is much faster than the energy filter, a pileup rejection window or artificial dead time can be implemented to discard events where pulses overlap in the energy filter (see chapter 6).

5.3 Trigger and Energy readout

The event energy can be attained through measuring the height of the shaped pulse by reading the ADC value during the flattop. For increased energy resolution multiple samples of the flattop can be averaged, for example by a moving average filter. This removes high frequency noise which can remain after the initial averaging by the trapezoidal filter [GK13]. The trapezoid height must be measured in reference to the baseline which is not usually at zero. Leakage current through the detector, as well as noise pickup from external sources and amplification circuitry, cause the baseline to shift and it does not necessarily stay at a constant level during operation. The detector can heat up, for instance, which increases leakage current. A dynamic method of *baseline restoration* is required. One such method described in [GK13] is to determine the baseline with a moving average during times where the channel is below a certain threshold and therefore considered not busy. The energy is then the difference of flattop height and baseline in front of the trapezoid.

All of these calculations require a reference time that specifies a precise point during the shaped pulse. This is provided by a trigger signal which simultaneously serves as an event timestamp. While trigger implementations can vary, several characteristics are desirable:

- a low detection threshold for measuring less energetic events,
- resistance to noise triggers,
- high accuracy of the trigger position in relation to the shaped pulse (including independence of the event energy) which translates to timing resolution, and
- low dead time or short busy state which determines the minimum resolution time for detecting consecutive pulses.

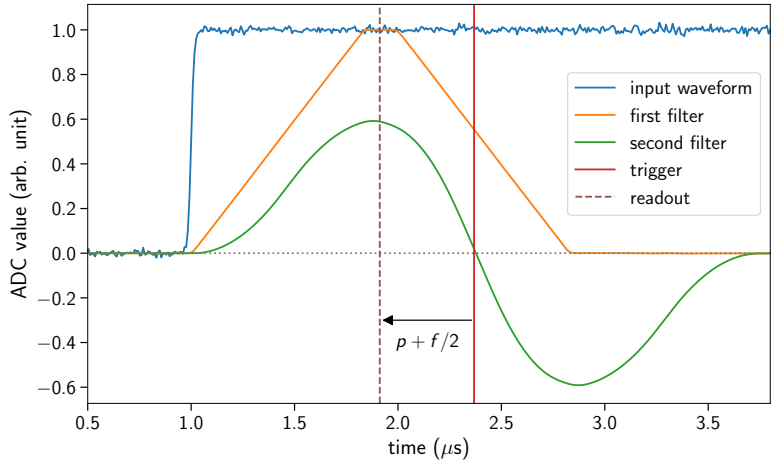


Figure 5.6: Demonstration of the bipolar trigger. If a zero-crossing is detected in the second stage, and the first filter is above a threshold, the energy is read out from the first stage at $p + f/2$ before the trigger. Here, p and f refer to the settings of the second stage trapezoidal filter.

In the following, two trigger algorithms are outlined. Both involve a second trapezoidal filter stage. In order to create a distinction to the parameters of the energy filter, which are represented by the uppercase letters P and F, the lowercase letters p and f are chosen for the trigger filter.

Bipolar trigger

The KATRIN DAQ system uses a combination of zero-crossing and threshold triggering with two successive trapezoidal filter stages [Ams15]. Figure 5.6 shows a simulation illustrating this trigger. The first filter is used for the energy readout as usual. A second filter with half the peaking time of the first stage is applied to the shaped signal. Similar to a derivative, this results in a bipolar pulse with a zero crossing which is positioned at $p + f/2$ after the middle of the first trapezoid, where p and f refer to peaking and flattop time of the second stage.

If a zero crossing is detected in the second stage, the value in the first stage is compared to a threshold. If the threshold is crossed, an energy readout is triggered. By relying on the already shaped signal for triggering and additionally requiring a zero crossing, this method is presumably resistant to noise triggers. A downside is the length of time the second filter stays in a busy state as it takes longer than the first filter stage for returning to the baseline. As the peaking times have to be sufficiently large for a precise energy readout, this results in a high minimum resolution time, rendering the system prone to signal pileup from consecutive events. Since the rate is low for nominal KATRIN operation and only the number of events in a certain region of interest below the tritium endpoint is relevant, reducing signal pileup is not as important. For TRISTAN, however, this trigger would be unsuitable due to the much higher rate.

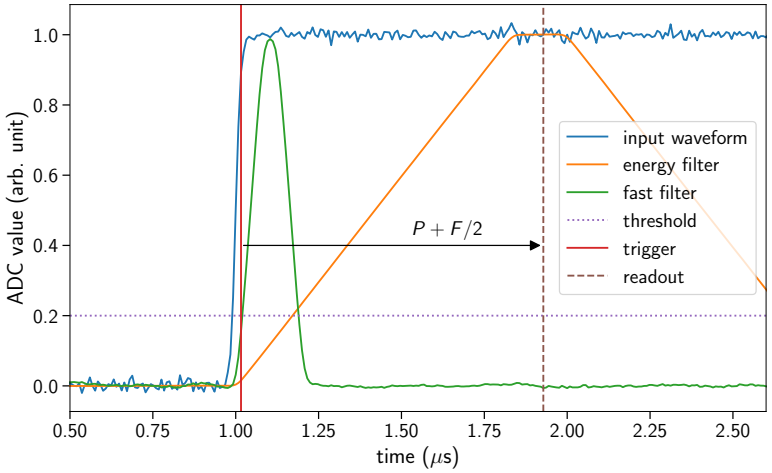


Figure 5.7: Demonstration of the fast filter trigger. If the fast filter crosses a given threshold, the energy is read out from the energy filter at $P + F/2$ after the trigger time, where P and F refer to the energy filter settings.

Fast filter trigger

Trigger speed and minimum resolution time can be improved significantly by using a separate and much faster filter in parallel to the energy readout filter. This filter is not used for an energy readout so its pulse shape can be much shorter. If two successive events are close enough to each other that their signals cause an overlap in the energy filter, they can still be detected as separate events in the fast filter. This way, pileup can be rejected very effectively.

The DANTE DAQ system, discussed in chapter 7, relies on a fast trapezoidal filter for triggering. Its peaking time and flattop time can be chosen much shorter than the parallel energy filter, resulting in a sharp pulse. The simulation shown in figure 5.7 depicts the general principle of the trigger. If the fast filter crosses a threshold, an energy readout is triggered. With this filter setup, the trigger position is slightly dependent on the event energy because trapezoids from higher energetic events will reach the threshold more quickly. For sufficiently low fast filter settings, this effect is negligible. However, if higher peaking times are required for the fast filter, the flattop time of the energy filter should be increased accordingly to ensure that the energy is still read out at the maximum of the shaped pulse.

Other systems, such as those described in [Tin08; GK13], rely on an RC-CR^2 filter for triggering. This produces a bipolar pulse shape similar to the second filter stage of the KATRIN DAQ. However, this pulse can be chosen to be much shorter than the energy filter pulse, allowing for pileup rejection. Its zero crossing is independent of the event energy and therefore provides a reliable reference time. When compared to the fast trapezoidal filter trigger, the drawbacks of this filter may include less resistance toward noise triggers and a wider pulse shape for similar averaging length.

5.4 Trace level simulation

For characterization and research of DAQ systematics, a trace level simulation is of major importance. Here, the term *trace* refers to a time series of digital voltage values in ADC units. In this context, the waveform denotes the trace produced by the ADC in the first DAQ stage. By applying digital filters to the waveform, traces of subsequent stages can be calculated. A first version of such a simulation was developed in the scope of this work since existing frameworks did not include the specific triggers discussed in the previous section.

The simulation framework is structured as a collection of separate methods for generating traces, applying digital filters to traces, triggering, and applying pileup rejection logic. These building blocks can be chained together to construct a signal and data acquisition chain as required. This open structure allows for direct access to all intermediate stages which is useful for building a detailed understanding of DAQ systems and DAQ related systematics.

Some of these simulation methods are inspired by the *Pulse Train Simulation* by David C Radford and the detector electronics simulation tool *DRIPS* by Sanshiro Enomoto. For convenience the code is implemented in Python [Pyt] and sped up by relying on the packages NumPy [Num] and SciPy.signal [Sci] where possible.

Several examples of simulation results are shown in the previous sections and the following chapters. In particular, in chapter 6 the simulation is used for verifying dead time and pileup models. As a reference, the simulation stages and algorithms are briefly discussed in this section.

Waveform generation
Waveforms can be initialized by filling a blank trace with values corresponding to event energies at given timestamps. Integrating this trace with a digital CR-filter results in a series of tail pulses with a time constant τ according to the CR-Filter settings. Setting $\tau = \infty$ yields flat steps as is the case for active reset preamplifiers. A constant charge deposition from leakage current can be added prior to the integration. For flat steps, this results in a characteristic upward slope as discussed in section 5.1. Alternatively, the waveform can be generated as a sum of sigmoid step functions for simulating non-zero signal rise times.

Noise
Two options allow for the simulation of electronic noise. Current noise can be emulated by adding fluctuations to the constant charge deposition due to leakage current. This results in variations to the slope after integration by the digital CR-filter. Voltage noise can be superimposed onto the integrated waveform by adding Gaussian random values.

Energy filter
In order to calculate the trace used for the extraction of pulse heights, the waveform is processed by a trapezoidal filter of peaking time P and flattop time F. It is implemented according to the transfer function from equation 5.2. Pole-zero compensation can be activated by specifying the parameter β.

Baseline restoration
A simple baseline restoration algorithm is implemented with a moving average filter which is applied to the energy filter trace. If an event causes the energy filter trace to jump above a given freeze threshold, compared to the baseline, the baseline is frozen until the energy filter trace returns to the baseline.

Trigger
Event timestamps and energy readout positions are determined by a trigger. The simulation includes the two trigger implementations mentioned in the previous section.

- Bipolar trigger (see figure 5.6):
  A timing filter trace is generated by applying a second stage trapezoidal filter to the energy filter trace. Peaking time p and flattop time f are set to half of the values of the energy filter settings P and F. A downward zero-crossing at t_{zero} and an energy value above a given threshold at $t_{\text{zero}} - p - f$ are required for generating a trigger timestamp. For each trigger timestamp the energy value is stored.

- Fast filter trigger (see figure 5.7):
  A fast trapezoidal filter with peaking time $p < P$ and flattop $f < F$ is calculated from the waveform in parallel to the energy filter. An upward threshold crossing causes a trigger at t_{trigger} and the energy filter is read out at $t_{\text{trigger}} + P + F/2$.

Pileup rejection
If activated, basic pileup rejection based on trigger dead time is performed by calculating time differences between trigger timestamps and comparing these to a set trigger dead time τ_{dead}. If such a time difference falls below τ_{dead}, both events are rejected.

CHAPTER 6

Rate dependent systematics

Most parts of the TRISTAN setup can be modeled as independent in relation to the rate of β-electrons reaching the detector. For example, although the column density is proportional to the activity and therefore the rate, for modeling systematic effects such as source scattering only the particle densities are parameters of interest. The actual rate of electrons exiting the source can be ignored. But for the detector and DAQ system, the incident rate is crucial since high rates can cause wrong energy readings and event loss. Simply put, if the rate of incident particles is high, a significant fraction of events will be very close in time and cannot be resolved. They can be lost in dead time after a previous event trigger, or be so close together that they cause pileup events with overlapping signal pulses, leading to events with increased energy readout. Since TRISTAN operates at rates of up to 10^5 cps per channel, rate dependent effects reach levels that can not be neglected and must be considered as a contribution to the systematics.

This chapter covers three rate dependent systematics arising from data acquisition: event losses due to dead time, tail pileup, and peak pileup. The discussion includes the magnitude of the effects, how they may be corrected or mitigated, and the investigation of modeling approaches. In the last section, these models are utilized to estimate the impact of peak pileup on TRISTAN's sensitivity for keV sterile neutrinos in the context of a spectral fit approach.

6.1 Overview

A fundamental assumption for the discussion of rate dependent effects is the randomness of radioactive decays. Independent events which are distributed randomly in time are called Poissonian, since the probability for a number n of events within a fixed time interval Δt is described by the Poisson distribution

$$P_{\text{Poiss}}(n; r, \Delta t) = \frac{(r \cdot \Delta t)^n}{n!} \cdot e^{-r \cdot \Delta t} \ . \tag{6.1}$$

Here, r is a constant rate and $r\Delta t$ corresponds to the expectation value as well as the variance of the distribution. For TRISTAN this assumption is presumably valid because the source activity is kept stable and adiabatic transport and acceleration of β-electrons in the magnetic and electric fields of transport and spectrometer section have no influence on time differences between events. The Poisson distribution is closely related to the distribution of waiting times between events, which is given by

$$f(\Delta t) = r \cdot e^{-r \cdot \Delta t} \ . \tag{6.2}$$

When $f(\Delta t)$ is integrated over a time interval $[\tau, \infty)$, it yields the probability for two events being further apart than τ

$$P(\Delta t > \tau) = \int_{\tau}^{\infty} r \cdot e^{-r \cdot \Delta t} \mathrm{d}\Delta t = e^{-r \cdot \tau} \ . \tag{6.3}$$

Since $P_{\text{Poiss}}(n = 0) = e^{-r \cdot \Delta t}$, this result can be interpreted as the probability to observe no event within the time frame between the events. This result is relevant, because it translates to the magnitude of peak pileup events and event losses from dead time as shown in the following.

Assuming a system with a timing trigger, that operates faster than the pulse shaper for energy readout, the distribution of time differences can be split into three domains:

- $\Delta t < \tau_{\text{min}}$:
  If events are so close in time that the timing filter cannot distinguish them and only triggers once, they will cause an unresolved peak pileup event. In the following, this dead time of the timing filter is referred to as minimum resolution time and labeled as τ_{min}. Since the energy filter pulses are much wider than τ_{min} they will overlap and lead to an increased energy readout. However, the energy of the pileup event is not necessarily the sum of both energies, which is demonstrated in section 6.4.2. The probability for two events to form one pileup event can be calculated by

  $$P_{\text{pileup}} = \int_{0}^{\tau_{\text{min}}} f(\Delta t) d\Delta t = 1 - e^{-r \cdot \tau_{\text{min}}} \ . \tag{6.4}$$

  Since pileup events themselves can fall into dead time of yet another event, this probability has to be multiplied by the factor from equation 6.5 in order to determine the number of recorded pileup events. This is shown in figure 6.2 for several τ_{min}. For triple, quadruple or higher orders of pileup, P_{pileup} can simply be raised to the power $n - 1$, where n is the number of contributing events.

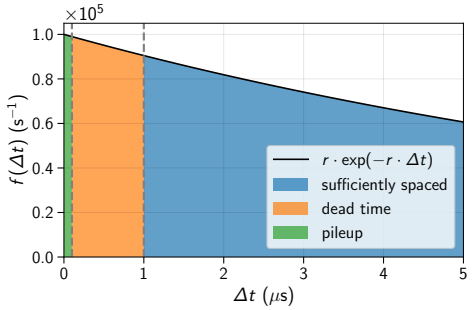


Figure 6.1: Distribution of times between events for a rate $r = 10^5$ cps. They are split in domains by minimum resolution time τ_{min}, and trigger dead time τ_{dead}. $\tau_{\text{min}} = 100\,\text{ns}$, and $\tau_{\text{dead}} = 1\,\mu\text{s}$ are realistic choices based on the performance of the DANTE DPP.

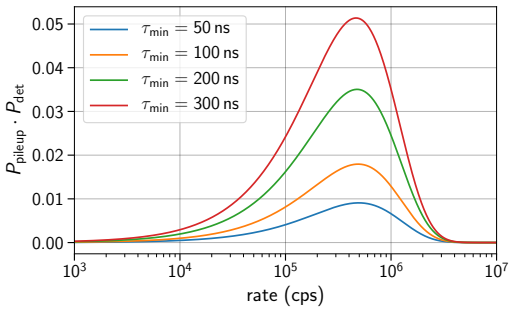


Figure 6.2: Pileup probability multiplied by detection probability for $\tau_{\text{dead}} = 1\,\mu\text{s}$ and different minimum resolution times. This probability determines the number of pileup events in the pileup spectrum. For an achievable $\tau_{\text{min}} \approx 100\,\text{ns}$, pileup is on the scale of $\approx 1\,\%$ for 10^5 cps which is the expected rate per pixel for TRISTAN.

- $\tau_{\text{min}} < \Delta t < \tau_{\text{dead}}$:
  In this domain, events are sufficiently spaced for separate triggers by the timing filter, but their pulses in the energy filter can still overlap. If read out, this causes wrong event energies. So instead the pulse processor can discard the events, which represents a form of pileup rejection. In the following, these events are referred to as resolved peak pileup events and the length of the rejection window as trigger dead time τ_{dead}. The latter must be at least as long as the peaking time for correct pulse heights, but can also be longer depending on the specific readout implementation. In the DANTE system it is set to $\tau_{\text{dead}} = P + F$, where P and F are the peaking and the flattop time of the energy filter.

- $\tau_{\text{dead}} < \Delta t$:
  For time differences larger than the trigger dead time the event pulses do not overlap, or at least not significantly, and the energies are acquired correctly. Since each event can either be too close to the preceding or the following event, the detection probability P_{det} for an event to not fall into the trigger dead time is the square of the result from equation 6.3

$$P_{\text{det}} = (e^{-r\cdot\tau_{\text{dead}}})^2 = e^{-r\cdot 2\tau_{\text{dead}}} \ . \tag{6.5}$$

  This equation suggests that a system, where both events are rejected if they are too close, behaves like a system with an effective dead time $\tau_{\text{eff}} = 2\tau_{\text{dead}}$ that extends before and after an event.

6.2 Dead time

Dead time usually refers to the time period after a detection system registers an event, where the system regenerates and is incapable of registering a following event. In general, the dead time period can be influenced by all parts of the electronic readout and data acquisition chain. For example, the detector itself is limited by the charge collection time, the ADC gives a lower bound on the time resolution through its sampling frequency, and pulse shaping requires the signals to be at least a peaking time apart for proper event detection. The latter is the dominant source of dead time because the peaking time is naturally much larger than ADC sample spacing and charge collection time.

On another note, the detector and ADC dead times are true hardware limitations, while the pulse processing dead time can be regarded as artificial. It depends on the implemented algorithm and is chosen by the programmer in order to avoid pileup events.

Paralyzability
Logically, even events that occur during the trigger dead time of a preceding event need to cause a dead time window of their own, since otherwise, the energy filter pulse can overlap with yet another event. Through this extension of dead time, the system can be completely paralyzed for extreme rates. Accordingly, this behavior is labeled *paralyzable* or *extendable* dead time. Non-paralyzable systems or mixtures of both types do exist, but they are irrelevant in the scope of this work. For further discussion see [UP18].

Recovery
For pulsed reset preamplifiers, an additional source of dead time has to be considered. Although very short, each reset takes a certain amount of time. For the prototype setup described in section 4.3 the reset causes a sharp downward step followed by a small period where the voltage remains at ground level. The length of this inhibit pulse can be tuned via the bias board. Afterward, the voltage begins to rise again and after some stabilization, the system is ready for measuring events. In order to exclude the reset region the pulse processor triggers a dead time called *recovery period*, here labeled τ_{rec}, upon detecting a negative edge.

Live time correction
While the energy of events falling into dead time cannot be recovered, an estimation of the true input count rate (ICR) is still possible by keeping track of the total dead time t_{dead}, or inversely the live time t_{live}, by counting the number of samples where the channel is in a dead state. The ICR can then be calculated from the output count rate (OCR), which is the number of accepted events divided by the total data taking time t_{total}.

$$\mathrm{ICR} = \mathrm{OCR} \cdot \frac{t_{\mathrm{live}}}{t_{\mathrm{total}}} \tag{6.6}$$

This correction method is crucial in an integral measurement mode, because the rate and therefore the total dead time increases for lower retarding potentials.

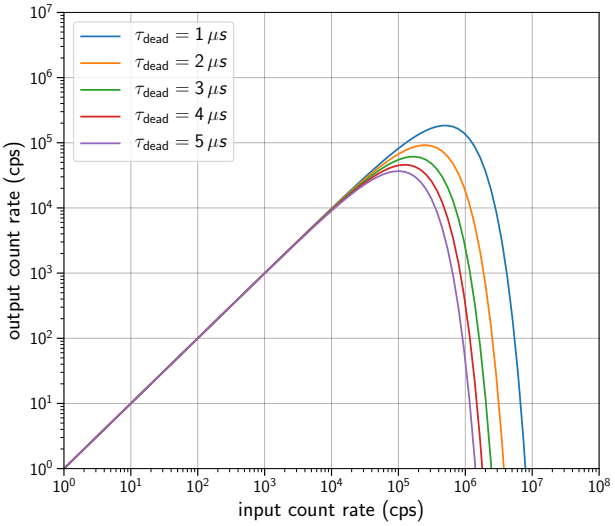


Figure 6.3: Detected count rate (OCR) over true count rate (ICR) for a paralyzable dead time model with various trigger dead times. Reset recovery periods are neglected. At roughly 10^5 cps the detected rate starts to drop significantly and drops to zero for extreme rates where the system is completely paralyzed. The peak of the curve signifies the rate setting for optimized statistics.

Dead time model

Since each event is equally likely to occur within dead time, the impact on a differential spectrum measurement is simply an energy independent efficiency factor. This factor can be measured by determining the live time as described above. For cross checks it is useful to have an analytic model. In the case of an active reset preamplifier, there are two independent contributions, namely the trigger dead time τ_{dead} and the recovery period τ_{rec}. In the paralyzable case the total live time can be calculated by

$$t_{\text{live}} = (t_{\text{total}} - n_{\text{resets}} \cdot \tau_{\text{rec}}) \cdot e^{-r \cdot 2\tau_{\text{dead}}} \ , \tag{6.7}$$

where n_{resets} is the number of resets as counted by the DPP. The first factor is simply the time where there is no reset period and the second factor is the fraction of sufficiently spaced events (see equation 6.5). Both contributions are multiplied since they are independent. The dead time fraction is then

$$\epsilon_{\text{DT}} = \frac{t_{\text{total}} - t_{\text{live}}}{t_{\text{total}}} = 1 - (1 - \frac{n_{\text{resets}} \cdot \tau_{\text{rec}}}{t_{\text{total}}}) \cdot e^{-r \cdot 2\tau_{\text{dead}}} \ . \tag{6.8}$$

In section 7.3.2 this model is compared to the dead time retrieved from the DANTE DPP. It should be noted that noise triggers cause dead time as well, which is not included in this model. Noise reduction plays a crucial role in decreasing the overall dead time.

6.3 Tail pileup

Tail pileup is the effect of energy underestimation for pulses with an exponential tail when pole-zero corrections are neglected. As already demonstrated in figure 5.4 pulses can be severely distorted if the pulse tail time constant τ is on the scale of the peaking time P. The main effect is a pulse undershoot which causes a mean downward baseline shift and therefore a smaller energy readout if not corrected. For sufficiently large τ, much larger than P, this effect is not as severe for single pulses. However, for high rates the pulses add up and the baseline shift accumulates. This can lead to a significant underestimation of event energies if the mean time difference between events $\overline{\Delta t} = 1/r$ is smaller than the pulse tail time constant ($\overline{\Delta t} < \tau$). The acquired spectrum is then shifted to lower energies, where the magnitude depends on the peaking time P, the flattop time F, the time constant τ, the mean event energy $\langle E \rangle$, and the rate r. An analytical model from [Eno14] yields the mean shift

$$\overline{\Delta E} = - \langle E \rangle \cdot \frac{(P+F)/\tau}{\exp(\frac{1}{r\tau}) - 1} \ . \tag{6.9}$$

Figures 6.4 and 6.5 show the effect of tail pileup on the energy filter trace and spectrum acquisition. For this demonstration, a trace level simulation based on the methods described in section 5.4 was performed. The pulses are generated with a time constant $\tau = 100\,\mu\text{s}$ at a rate $r = 2 \cdot 10^5\,\text{cps}$ with heights drawn from a Gaussian distribution ($\mu = 1.0$ and $\sigma = 0.05$). The trapezoidal energy filter is set to $P = 1.12\,\mu\text{s}$ and $F = 0.32\,\mu\text{s}$ and a bipolar trigger, as shown in figure 5.6, is used for acquisition. The blue spectrum in figure 6.5 shows how tail pileup shifts the spectrum to lower energies and it is in good agreement with the prediction from equation 6.9. Furthermore, due to baseline fluctuations the peak width slightly increases to a fit width of $\sigma = 0.68$.

Possible corrections for tail pileup include pole-zero correction and baseline restoration which are shown as well. All three acquisitions are performed on the same simulated waveform. For the baseline restoration method, a small shift remains, which is due to the non-zero freeze threshold that is needed for the algorithm described in section 5.4.

The original FPD DAQ system lacks such a correction (see figure 3.11b), due to technical reasons and because it is not necessary for the low rates during neutrino mass runs. However, the capability for high rate measurements can be useful for calibration and commissioning measurements. A current DAQ upgrade includes a baseline independent bipolar energy readout from the second filter stage as a correction method.

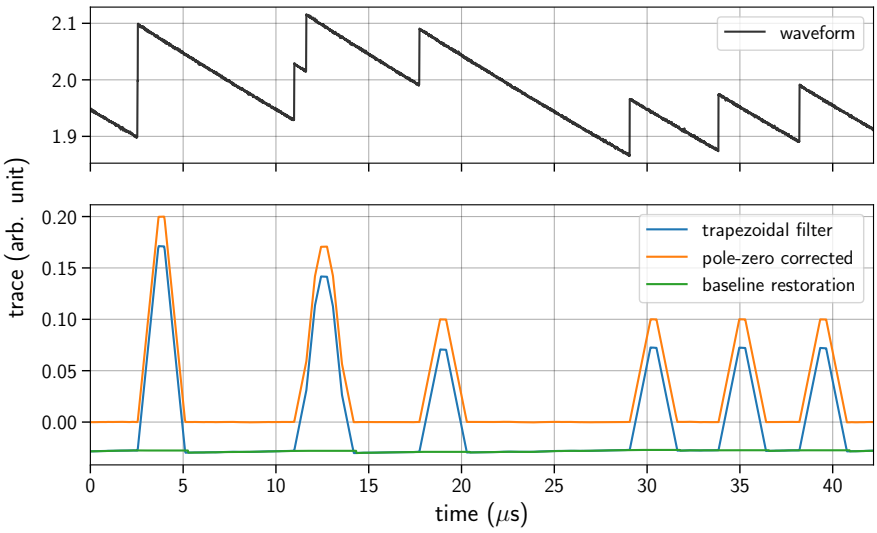


Figure 6.4: Excerpt of a trace level simulation demonstrating the baseline shift due to tail pileup (blue line). The pole-zero corrected trace (orange line) and the restored baseline (green line) are illustrated as well.

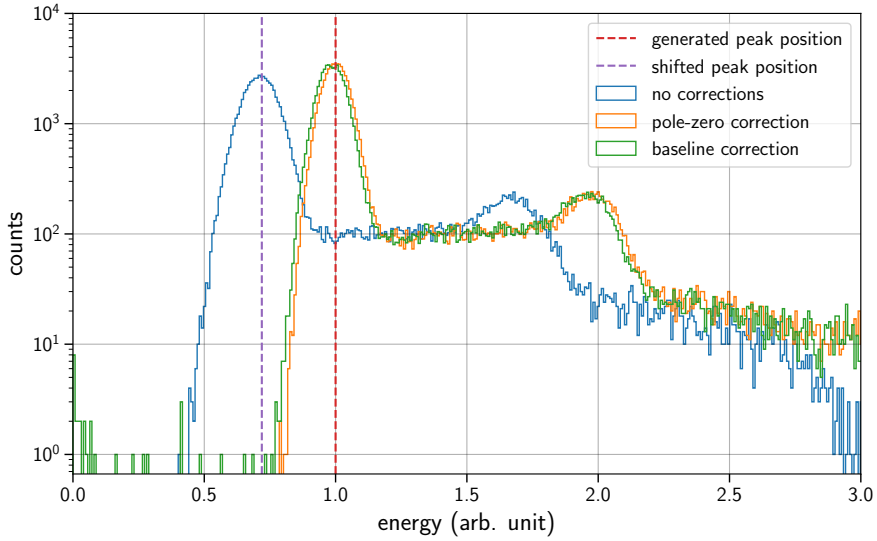


Figure 6.5: Spectra from a trace level simulation with exponential pulse tails. The shifted peak position (purple dashed line) indicates the prediction from equation 6.9. Applying pole-zero cancellation or baseline restoration corrects for these effects.

6.4 Peak pileup

In general, peak pileup occurs when the time between two events is smaller than the peaking time of the energy filter ($\Delta t < P$). The shaped pulses will then overlap to the extent that they produce a single pulse. The height of this combined pulse is dependent on the time difference and lies between the sum of both energies and the single energies. In a spectrum measurement, peak pileup leads to a spectral component reaching up to twice the maximum energy of the input energy spectrum. Depending on pileup rejection and minimum energy of the input spectrum it may overlap with the main spectrum. Its overall height depends on rate, trigger dead time, and the minimum resolution time. In the following, this spectral component is referred to as pileup spectrum.

6.4.1 Minimum resolution time

For a trapezoidal threshold trigger, the minimum resolution time can be derived from a geometric construction as shown in figure 6.6. Two successive pulses can only be detected as separate if the trace between the pulses reaches the trigger threshold H for at least one sample. The lowest intermediate trace point depends on the time difference Δt between events and rises if Δt decreases. Hence, the criterion for $\Delta t = \tau_{\min}$ is that this point touches the threshold. The time period $\delta t = P \cdot \left(1 - \frac{H}{E_2}\right)$ (for $E_2 < E_1$) is obtained from simple geometry and introduces a threshold and energy dependence. For $E_2 > E_1$ the picture is flipped and δt depends on E_1 instead. In total, the derivation yields

$$\tau_{\min} = F + P + \delta t = F + P \cdot \left(2 - \frac{H}{\min(E_1, E_2)}\right) \ . \tag{6.10}$$

By taking the minimum of E_1 and E_2 both cases are considered. Since this argumentation only holds true for well-shaped pulses, it's validity is limited. Large signal rise times and noise fluctuations deform the trapezoid and for small settings of P and F the pulse may not even reach full height. Nonetheless, it is useful as an approximation and can be compared to simulations and measurements.

A simple method for measuring the minimum resolution time is by determining the fraction of unresolved pileup events when the true rate is known and the pileup spectrum can be clearly separated from the main spectrum. The number of events in the pileup spectrum compared to the total number of input events N_{evt} depends on the live time fraction and the pileup probability from equation 6.4

$$\frac{N_{\text{pileup}}}{N_{\text{evt}}} = \frac{t_{\text{live}}}{t_{\text{total}}} \cdot \left(1 - e^{-r \cdot \tau_{\min}}\right) \ . \tag{6.11}$$

The live time fraction can be determined in-situ from the spectrum by replacing it with the fraction of events in the main spectrum, since $\frac{N_{\text{main}}}{N_{\text{evt}}} = \frac{t_{\text{live}}}{t_{\text{total}}}$, which leads to

$$\tau_{\min} = -\frac{1}{r} \cdot \ln\left(1 - \frac{N_{\text{pileup}}}{N_{\text{main}}}\right) \ . \tag{6.12}$$

Inversely, knowledge of $\tau_{\min}$ enables the estimation of the height of the pileup spectrum for any given input rate which can, for instance, be used as a constraint in a spectral fit.

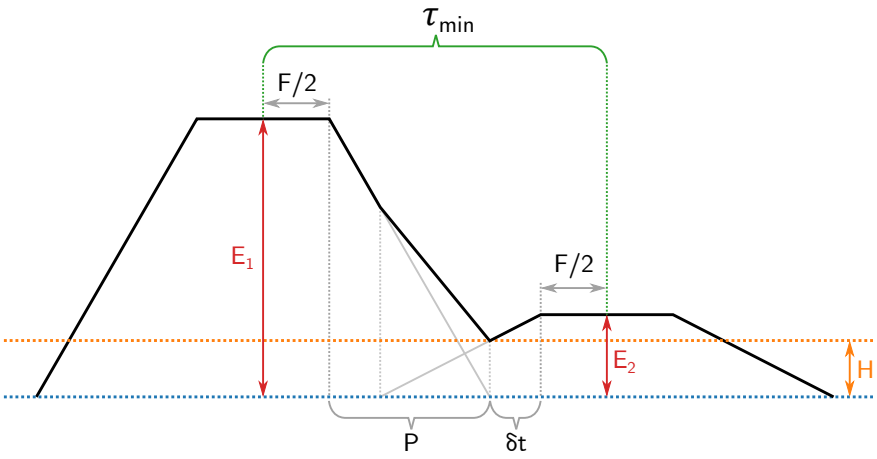


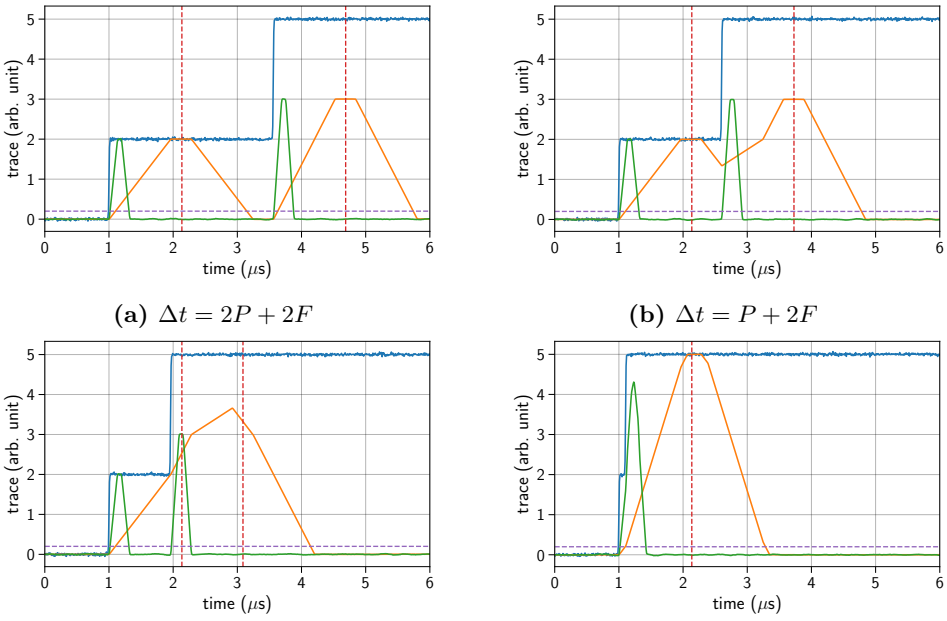
Figure 6.6: Geometric derivation of the minimum resolution time for a trapezoidal threshold trigger. The drawing catches the exact time difference between events where the trace reaches down to the trigger threshold H in between pulses.

6.4.2 Pileup energy

A natural assumption for the energy of a pileup event is that it is simply the sum of both energies. While this is true in some cases, it is not in general. In figure 6.7 this is demonstrated using trace simulations. Depending on the time Δt between the events, the pulses may only partially overlap and lead to a combined pulse of intermediate height. Furthermore, the readout position likely does not coincide with the highest point of the pileup pulse except for very small time differences. This example also shows how a fast filter with slimmer pulses helps to detect such events for rejection. However, as soon as these overlap as well, there will be pileup if no further pileup rejection mechanisms are employed.

In figure 6.8 the same simulation is expanded to a full range of event time differences and the event energy is acquired according to the fast filter trigger described in section 5.4. For Δt less than half the energy filter flattop F the pulse height is the sum of both event energies. For larger Δt the acquired energy gets lower and linearly transitions to the energy of the separate event as soon as the second pulse no longer overlaps with the readout position $P+F/2$. Furthermore, the simulation reveals the point where the system is able to separate both events allowing for pileup rejection. The minimum resolution time coincides well with the expectation from equation 6.10, however, the agreement worsens slightly for larger signal rise times as shown in figure 6.8b. For correct energy readout of both events, the trigger dead time τ_{dead} must be chosen to be larger than at least $P + F/2$. In the DANTE DPP, it is set to $\tau_{\text{dead}} = P + F$ which allows for some leniency for late triggers and larger rise times.

Ideally, a pileup event readout should always yield the sum of both energies, which allows for much easier modeling of the pileup spectrum. This can be achieved in two ways. One is to increase the flattop of the energy filter in order to increase the flat section for



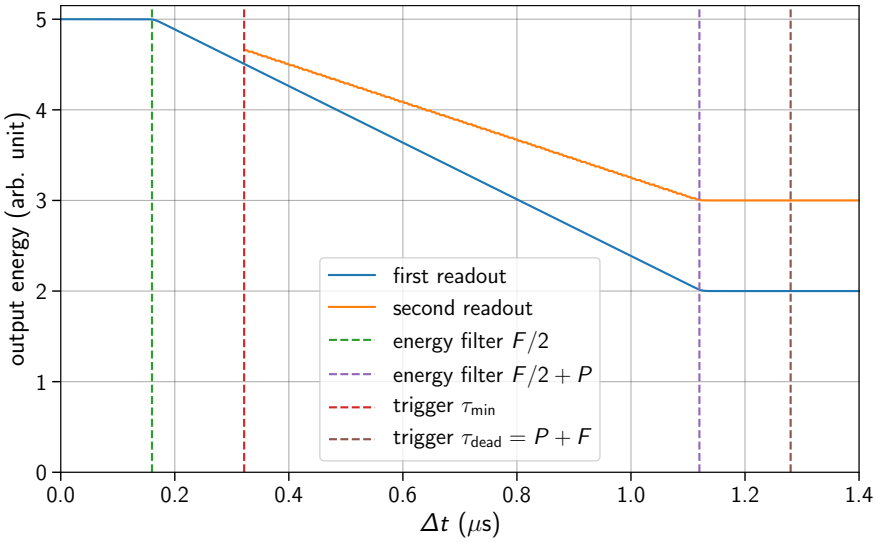
(a) $\Delta t = 2P + 2F$ **(b)** $\Delta t = P + 2F$

(c) $\Delta t = P$ **(d)** $\Delta t = F/3 < \tau_{\min}$

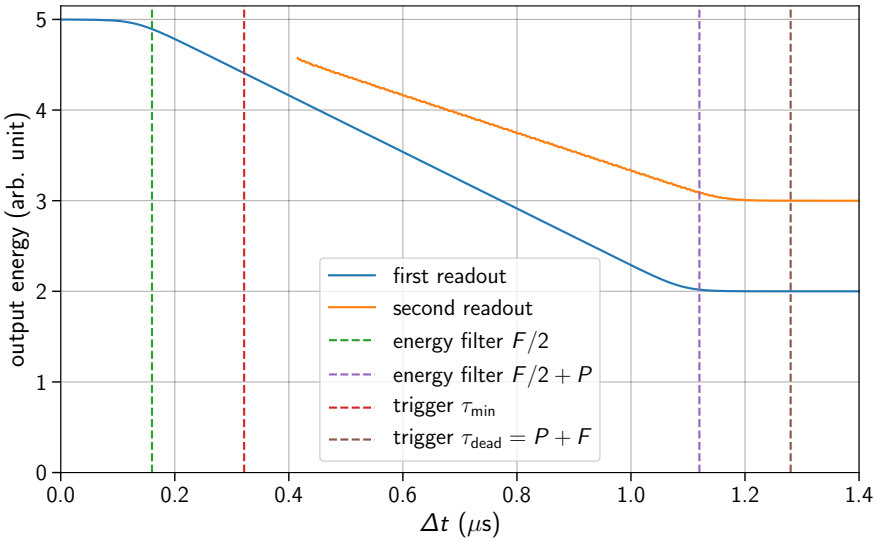
Figure 6.7: Trace level simulation for demonstrating pileup. A step waveform for two events is generated for several time differences Δt and shaped by a fast filter for triggering and a slower filter for energy readout.

low Δt. On the downside, this will also increase the dead time. Another possibility is to decrease the minimum resolution time, which has the added benefit of lowering the number of pileup events overall. Naturally, it can be lowered by decreasing peaking and flattop time of the fast filter. However, the minimal peaking time is limited by the noise level, since voltage fluctuations can then more easily cause a trigger. A certain flattop is necessary in order to account for the signal rise time, because for very low flattops the fast filter pulse is less likely to reach the trigger threshold. It should also be noted that the pileup energy can already start to drop below $\Delta = F/2$ for larger signal rise times as shown in figure 6.8b, so it is not always sufficient to lower $\tau_{\min}$ below this threshold.

A more sophisticated method of decreasing the resolution time is pulse shape analysis of the fast filter pulse. One metric is, for example, the area under the fast filter pulse which can be much larger for pileup events compared to regular events in the case of a partial pulse overlap. By estimating the area for a regular pulse, pileup pulses can be rejected for values that are too high. The area is correlated to the effective rise time of the combined pulse and this method is therefore labeled *maximum rise time rejection*.



(a) Low signal rise time $t_{\mathrm{rise}} = 16\,\mathrm{ns}$.



(b) High signal rise time $t_{\mathrm{rise}} = 100\,\mathrm{ns}$.

Figure 6.8: Energy readout for pileup events from a trace level simulation. Two pulses with heights $E_1 = 2$ and $E_2 = 3$ are generated for each time difference Δt and a fast filter threshold trigger and readout algorithm is applied. The red line represents the geometric minimum resolution time as given by equation 6.10.

6.4.3 Modeling the pileup spectrum

For building an understanding of pileup on the FPGA level, the trace simulation is a very useful tool. However, an analytical description of the pileup spectrum is desirable as well, because it can easily be used for fitting data with significantly less computational effort. This analytic model is comparatively simple if $\tau_{\min}$ is sufficiently low for guaranteeing that the pileup energy of two events is given by

$$E_{\text{pileup}} = E_1 + E_2 \ . \tag{6.13}$$

Since every energy of the input spectrum $\frac{\mathrm{d}\Gamma}{\mathrm{d}E}$ can pile up with any other energy, E_{pileup} is generally a sum of two random numbers with $\frac{\mathrm{d}\Gamma}{\mathrm{d}E}$ being their probability density function. The pileup spectrum can therefore be expressed as a convolution of the spectrum with itself:

$$\left(\frac{\mathrm{d}\Gamma}{\mathrm{d}E}\right)_{\text{pileup}} = \int \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E-\varepsilon) \cdot \frac{\mathrm{d}\Gamma}{\mathrm{d}E}(E)\ \mathrm{d}\varepsilon \ . \tag{6.14}$$

This is demonstrated in figure 6.9 for the simulation of a flat input spectrum reaching from $E_{\min} = 0.5$ to $E_{\max} = 1.0$. As long as equation 6.13 holds true, the pileup spectrum corresponding to double pileup reaches from $2E_{\min}$ to $2E_{\max}$. Therefore, in this example the pileup spectrum and the main spectrum are well separated. As input 10^6 events were simulated and read out with the fast filter trigger described in section 5.4, and the detected events are binned in 200 energy intervals. For the expectation the flat input spectrum is convoluted with itself for the pileup spectrum, and the pileup spectrum itself is convoluted with the input once again for the triple pileup spectrum. These functions are then scaled to the expected number of events resulting from pileup probability (equation 6.4) and detection probability (equation 6.5). Integration over an energy interval leads to the expected number of events in the respective bin i

$$N_{i,\text{exp}} = N_{i,\text{main}} + N_{i,\text{pileup}} + N_{i,\text{triple}} \ , \tag{6.15}$$

where

$$N_{i,\text{main}} = N_{\text{evt}}\ e^{-2r\tau_{\text{dead}}} \int_{\text{bin } i} \mathrm{d}E \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}E}\right)_{\text{input}} \ , \tag{6.16}$$

$$N_{i,\text{pileup}} = N_{\text{evt}}\ e^{-2r\tau_{\text{dead}}}\ (1-e^{-r\tau_{\min}}) \int_{\text{bin } i} \mathrm{d}E \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}E}\right)_{\text{pileup}} \ , \tag{6.17}$$

$$N_{i,\text{triple}} = N_{\text{evt}}\ e^{-2r\tau_{\text{dead}}}\ (1-e^{-r\tau_{\min}})^2 \int_{\text{bin } i} \mathrm{d}E \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}E}\right)_{\text{triple}} \ . \tag{6.18}$$

Here, τ_{dead} is the sum of energy filter peaking and flattop time. $\tau_{\min}$ is calculated according to the geometric formula 6.10 for the mean input energy $\langle E \rangle = 0.75$. Taking the mean is not necessarily correct because each pileup energy can be composed of different sets of input energies which would reflect in the energy dependence of $\tau_{\min}$. Still, for the purpose of this demonstration this approximation works well and results in good agreement between convolution and simulation even without fitting any parameters. A larger dataset may however reveal more tensions and the influence of quadruple pileup.

Table 6.1: Simulation and convolution parameters for figure 6.9.

General parameters		Filter settings	
Events	10^6	Energy filter P	0.96 µs
Rate	10^5 cps	Energy filter F	0.96 µs
Sample spacing	8 ns	Fast filter p	0.20 µs
Threshold	0.1	Fast filter f	0.08 µs

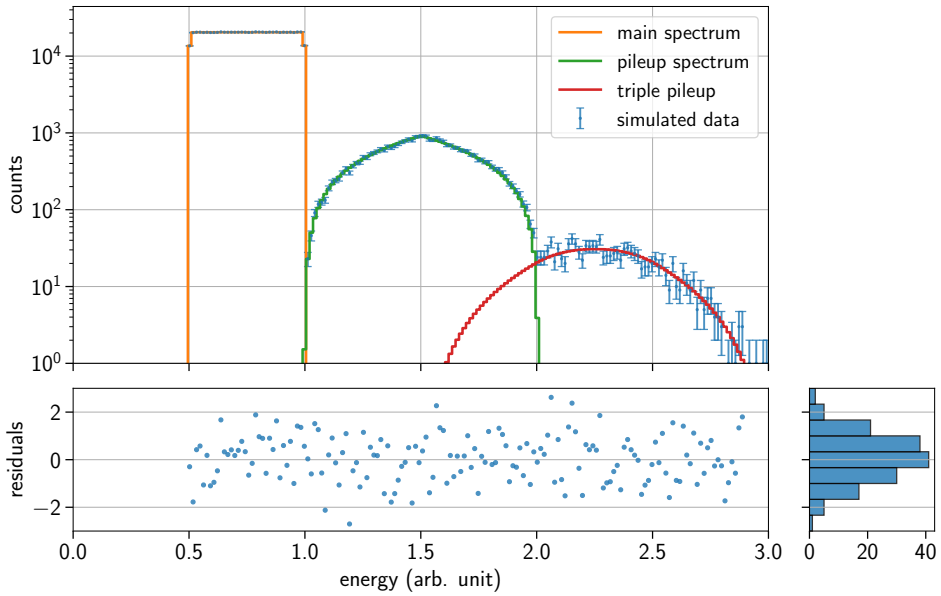


Figure 6.9: Comparison of trace level simulation and convolution approach with residuals normalized to the standard deviation from the square root of expected events. The analytical expectation is not fit. Nonetheless, it is in good agreement with the simulation at $\chi^2/\text{ndf} = 152.6/160 = 0.954$.

6.5 Impact on keV sterile neutrino sensitivity

Using the models from this chapter it is possible to estimate the impact of pileup on TRISTAN's sensitivity for keV sterile neutrinos. For this study, tail pileup is neglected because it can be corrected as demonstrated in section 6.3. Furthermore, it is not a concern for the prototype detector setup which relies on an active reset preamplifier with DC-coupling and is therefore tail pileup free. However, dead time and particularly peak pileup need consideration.

Dead time reduces the detected rate resulting in an overall loss of statistics which translates to a reduction of sensitivity. Due to its rate dependence, it can provide an argument for rate optimization as shown in figure 6.3.

Peak pileup poses a challenge since it causes an additional spectral component which can overlap with the tritium spectrum. The pileup spectrum can hide a potential kink caused by a sterile neutrino in the low statistics region near the endpoint. In addition, the overlap causes a kink of its own which may interfere with the kink search of the model independent wavelet analysis approach (see section 4.1). With an expected magnitude of roughly 1 %, as shown in figure 6.2, pileup cannot be neglected in a spectral fit. In order to achieve a good χ^2-value, pileup needs to be considered in the model, for example with the convolution approach or a Monte Carlo simulation.

For the purpose of this study, pileup is described with the method from the previous section and put into the context of the spectral fit approach mentioned in section 4.1. The procedure is based on the investigations from [Mer15a; Dol16] and relies on determining a χ^2-value for testing sterile mixing hypotheses. A dataset without a sterile neutrino is generated from the theoretical, differential tritium spectrum and compared to a null hypothesis which is the expectation for a given sterile mass m_s and mixing amplitude $\sin^2\theta$.

The data is binned in 0.2 keV energy intervals ranging from 0 keV to 60 keV. A large upper bound is required to account for pileup and to allow for simulating the influence of the post acceleration potential which shifts the spectrum to higher energies. The χ^2-value is calculated according to

$$\chi^2 = \sum_{ij} (N^i_\mathrm{obs} - N^i_\mathrm{exp}(m_\mathrm{s}, \theta)) V^{-1}_{ij} (N^j_\mathrm{obs} - N^j_\mathrm{exp}(m_\mathrm{s}, \theta)) \ , \tag{6.19}$$

where N^i_obs is the "observed" number of events in bin i without sterile neutrino and $N^i_\mathrm{exp}(m_\mathrm{s}, \theta)$ is the expectation for a given mixing scenario. V^{-1}_{ij} represents the inverse covariance matrix containing statistical errors and bin-to-bin correlations from systematic effects.

Following the procedure described in section 6.4.3 N^i_obs and N^i_exp are are taken as a sum of the main spectrum and the pileup spectrum

$$N^i = N^i_\mathrm{main} + N^i_\mathrm{pileup} \ . \tag{6.20}$$

Triple pileup is neglected, since its contribution is roughly two orders of magnitude lower than the contribution from double pileup. N^i_{main} and N^i_{main} are again given by

$$N^i_{\text{main}} = N_{\text{evt}}\, e^{-2\tau_{\text{dead}} r/n_{\text{pix}}} \int_{\text{bin } i} \mathrm{d}E \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}E}\right)_{\text{input}} , \tag{6.21}$$

$$N^i_{\text{pileup}} = N_{\text{evt}}\, e^{-2\tau_{\text{dead}} r/n_{\text{pix}}}\, (1 - e^{-\tau_{\text{min}} r/n_{\text{pix}}}) \int_{\text{bin } i} \mathrm{d}E \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}E}\right)_{\text{pileup}} . \tag{6.22}$$

Several parameters are chosen and their values are listed in table 6.2. N_{evt} denotes the total number of events given by total rate $r = 10^8\,\text{cps}$ multiplied with a measurement time of three years. Since the β-electron flux is distributed over all detector pixels, the rate entering detection and pileup probabilities is divided by the number of pixels n_{pix}.

For N^i_{exp} the input spectrum is taken as a sum of active and sterile contributions and normalized to unity after addition. This ensures that the total number of events for mixing and no mixing are equal. In practice, the normalization would be fit alongside a multitude of other parameters such as the endpoint, background rate, and nuisance parameters. However, since this approach is simplified the fit can be avoided with proper normalization. In order to reduce computation time, the pileup spectrum is only calculated once for the case of no mixing, because variations resulting from the sterile contribution are negligible for small $\sin^2\theta$.

With the covariance matrix approach statistical and systematical uncertainties, such as pileup in this study, can be calculated as separate contributions and combined via addition

$$V = V_{\text{stat}} + V_{\text{pileup}} . \tag{6.23}$$

The statistical error is taken as the square root of observed events in bin i leading to a diagonal matrix

$$V^{ii}_{\text{stat}} = N^i_{\text{obs}} . \tag{6.24}$$

V_{pileup}, on the other hand, introduces off-diagonal elements as well. These describe bin-to-bin correlations stemming from uncertainties in the description of the pileup spectrum. For expressing these uncertainties, several parameters determining position and height of the pileup spectrum are varied around their mean value for 1000 toy pileup spectra. This includes variations of τ_{min} and r, to account for calibration uncertainty and limited rate stability. Both have an influence on the height of the pileup spectrum. Two additional parameters are introduced in order to emulate uncertainties in the determination of pileup energies. An energy offset ΔE slightly shifts, and an energy scaling factor S_E stretches the pileup spectrum in E-direction. An error on the pileup energy measurement is mainly motivated by the underestimation of pileup energy as discussed in section 6.4.2. For instance, underestimation can occur if the minimum resolution time is too high or if the signal rise time is very large (see fig. 6.8b). A conceivable energy dependence of τ_{min} is, however, neglected. The magnitudes of parameter variation are noted in table 6.2 as well. To ensure convergence of the covariance matrix a sufficient number of toy datasets is necessary. $n = 1000$ is chosen because it is computationally manageable and a comparison with matrices from $n = 100$ toy datasets shows no noticeable differences.

Table 6.2: Parameters used for sensitivity calculation and their variations for estimating the pileup covariance matrix.

Parameter	Variable	Value	Variation
Rate	r	10^8 cps	1 %
Measurement time	t_{meas}	3 y	-
Number of events	N_{evt}	$\approx 9.5 \cdot 10^{15}$	-
Number of pixels	n_{pix}	3486	-
Trigger dead time	τ_{dead}	1 µs	-
Minimum resolution time	τ_{min}	200 ns	5 %
Post acceleration	E_{PAE}	$0 - 20$ keV	-
Pileup energy shift	ΔE	0 keV	0.2 keV
Pileup energy scale	S_E	1	1 %

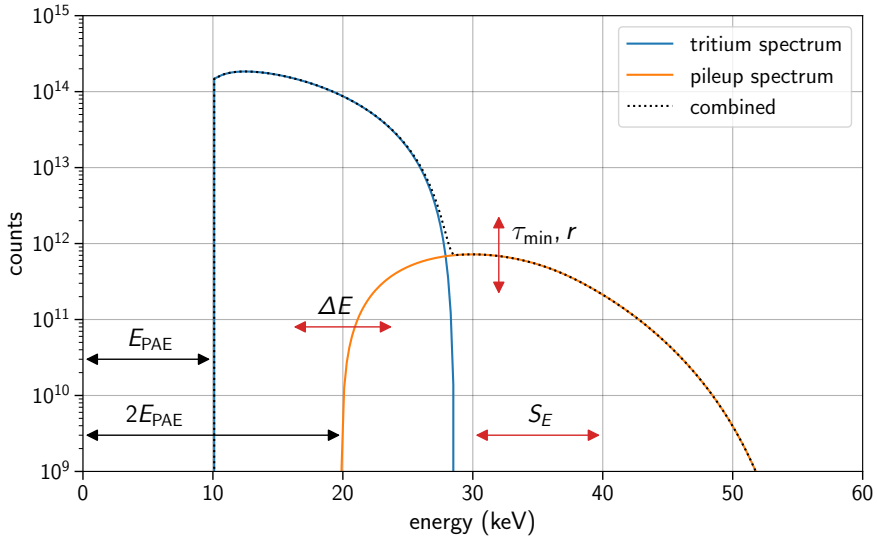


Figure 6.10: The tritium spectrum and its pileup spectrum from convolution. The spectrum is shifted in E-direction with a post acceleration setting $E_{\text{PAE}} = 10$ keV, as a consequence the pileup spectrum is shifted by double this amount. Red arrows illustrate the variations applied to the pileup spectrum for estimating the covariance matrix.

Following the procedure described in [Sta], the covariance matrix can be estimated from the $n = 1000$ toy datasets in accordance with the covariance estimator

$$V_{ij} = \frac{1}{n} \sum_{k=1}^{n} (N_k^i - \overline{N^i}) \cdot (N_k^j - \overline{N^j}) \ . \tag{6.25}$$

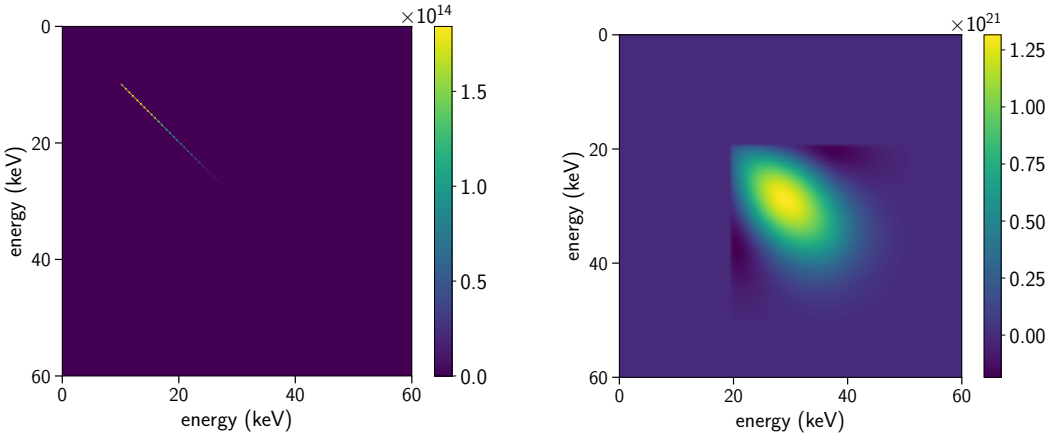
Here N_k^i refers to the number of events in bin i for dataset k and $\overline{N^i}$ is the mean number of events in bin i averaged over all datasets.

With all parts of the χ^2 from equation 6.19 incorporated, a scan of a 50×50 grid of the $(m_\mathrm{s}, \sin^2\theta)$-parameter space is performed. By requiring $\chi^2 > 4.61$, a 90 % confidence level exclusion contour is obtained.

Figure 6.12 shows the resulting sensitivity based on the parameter choice given in table 6.2. For comparison, the sensitivity is also computed for the case of only statistical errors contributing to the covariance. It yields that by including pileup, the sensitivity is slightly reduced over the full range of sterile masses when no post acceleration is applied. It may be noted that the parameters τ_min, ΔE, S_E, and their variations are chosen rather conservatively. Furthermore, the introduction of the pileup energy shift and the scaling factor is arbitrary and might not represent the real impact of faulty pileup energy estimation. In the future, a more detailed and complete pileup model, which depends solely on physical parameters, could be used and the parameter values and their variations can be constrained with calibration measurements.

The calculation is repeated for two post acceleration settings $E_\mathrm{PAE} = 10\,\mathrm{keV}$ and $E_\mathrm{PAE} = 20\,\mathrm{keV}$. The corresponding curves are also shown in figure 6.12. By activating post acceleration the tritium spectrum is shifted to higher energies and since the pileup energy is modeled as the sum from two separate events, the pileup spectrum is shifted by twice the amount. Effectively, the pileup spectrum is shifted out of the tritium spectrum and sensitivity is recovered. For low post acceleration settings, the overlap of tritium and pileup spectrum is limited to energies near the endpoint which corresponds to sensitivity loss for low sterile masses while sensitivity for high masses is unaffected. If the post acceleration energy is larger than the endpoint energy $E_0 = 18.6\,\mathrm{keV}$, pileup spectrum and tritium spectrum are fully separated.

As a conclusion, post acceleration can be a powerful tool for mitigating pileup. However, this method is based on the correct estimation of pileup energies which requires a very low minimum resolution time achieved through pileup rejection mechanisms. Furthermore, the input spectrum used in this study is idealized. The effectiveness of shifting pileup out of the tritium spectrum is compromised by the spectral shape when other systematics are considered. Energy loss in the detector entrance window and backscattering cause a low energetic tail which also contributes to pileup events. This tail reaches down to the detection threshold independent of the post acceleration setting, so some level of pileup overlap is unavoidable. In addition to pileup, the post acceleration influences other systematics as well. This includes the energy resolution, which worsens for higher energies, and ADC-nonlinearities [Dol17], which favor high post acceleration settings. So the final value of the post acceleration needs to be optimized with all relevant systematics in mind.



(a) V_{stat} **(b)** V_{pileup}

Figure 6.11: Representation of the covariance matrices. The covariance describing statistics is diagonal whereas the pileup covariance matrix has both diagonal and off-diagonal elements. For this calculation the post acceleration is set to $E_{\mathrm{PAE}} = 10\,\mathrm{keV}$ which results in a shift of the pileup spectrum and corresponding matrix elements to above $2E_{\mathrm{PAE}} = 20\,\mathrm{keV}$.

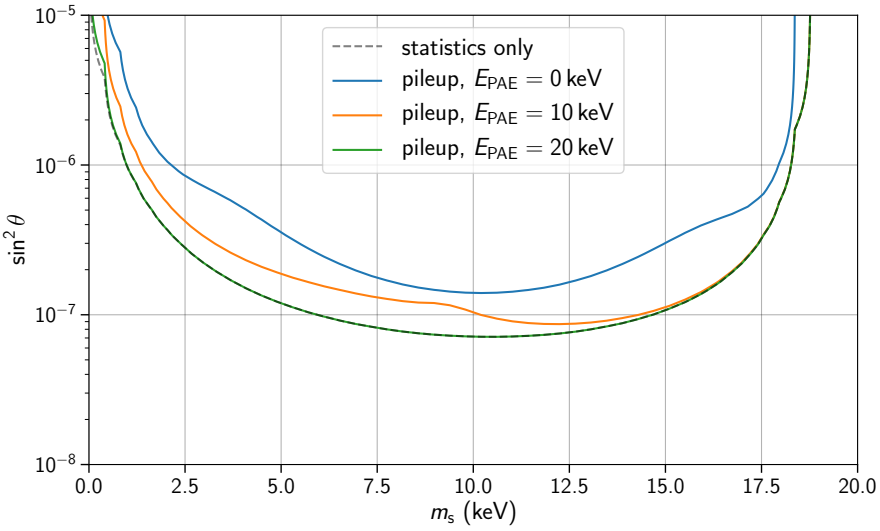


Figure 6.12: Estimated 90 % confidence level contours demonstrating the effect of peak pileup on TRISTAN's sensitivity for keV sterile neutrinos. When post acceleration is activated, the pileup spectrum is shifted out of the tritium spectrum and sensitivity is recovered.

CHAPTER 7

Characterization of the DANTE digital pulse processor

The 7-pixel prototype TRISTAN detector is mainly used in conjunction with the 8-channel DANTE Box by XGLab [XGLb]. It is constructed as a daisy chain of eight DPP_4552 boards which are designed for use with pulsed reset type preamplifiers such as the CUBE ASIC [XGLc]. The information presented in this chapter is based on the hardware manual revision 2.2, detector emulator characterization measurements and all measurement results are in reference to firmware version 321.276.3633.

The goal of this chapter is the characterization of the DANTE DPP with regard to high rate behavior. The tests concern the determination of minimum resolution time for pileup detection, rate measurement accuracy, maximum rise time rejection efficiency, as well as a cross check of the dead time implementation for various parameter settings.

For these measurements a digital detector emulator, the DT5810B by CAEN [CAE], is used. It can generate waveforms similar to those from the amplified SDD signal with either fixed pulse heights or variable pulse heights drawn from spectra. The rate can be set to constant or random event intervals in a large range of rates up to 30 Mcps, which makes it well suited for this study. Unfortunately, in pulsed reset mode it is not able to generate a slope for emulating leakage current and variations to the signal rise time are not possible either. Therefore, the prototype setup waveform cannot be replicated exactly. In order to keep this chapter concise, a detailed description of the emulator is omitted. For more information the manual is available at [CAE].

7.1 General description

Each pulse processor board is equipped with a 16-bit ADC running at 125 MS/s for digitization of the DC-coupled preamplifier waveform at an 8 ns sample spacing. An FPGA on the board performs digital filtering for pulse shaping and employs a trigger and readout algorithm for acquiring event energies and timestamps. The general principles of these methods are discussed in chapter 5. While this chapter provides a deeper insight into the logic applied in the DANTE FPGA, figure 7.1 shows a simplified schematic of the filter and readout structure.

Connection to a PC is established either via USB or Ethernet. It is used for data transfer and firmware updates, as well as the configuration of the digital filters and other parameters. The graphical user interface for operating the DANTE Box offers four main modes of data acquisition:

- **List mode**: Acquires event energies and timestamps.
- **Spectrum mode**: Measured event energies are binned in a histogram and event timestamps are not recorded resulting in a significant reduction of output data size.
- **Waveform**: The unfiltered, digitized input waveform is acquired for a set length of time.
- **Listwave mode**: Event energies and timestamps are acquired and for each event, a snippet of the waveform is recorded.

All modes also provide additional logging output. Among many variables, this includes filter settings, live time and dead time, output count rate, estimated input count rate and detected number of preamplifier resets.

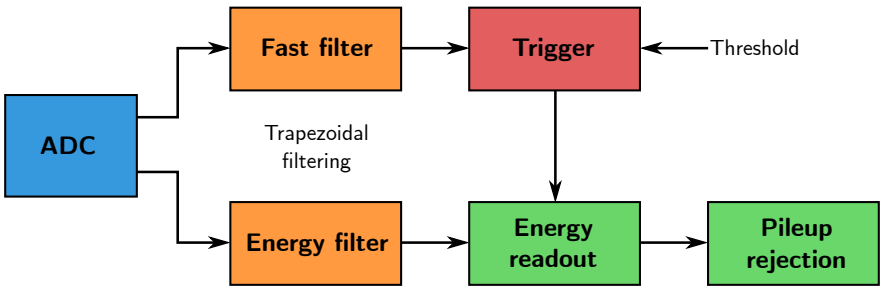


Figure 7.1: Schematic of the DANTE filter and readout structure.

In the current firmware ten parameters are required for configuring the DPP boards:

- trigger threshold,
- fast filter peaking and flattop time,
- energy filter peaking and flattop time,
- recovery period length after reset detection,
- baseline samples for baseline restoration,
- maximum rise time for improved pileup rejection,
- zero peak rate for acquisition of baseline points as a calibration reference, and
- digital gain for adjusting bin edges in spectrum mode.

Except for digital gain and zero peak rate, all parameters require optimization. The trigger threshold determines the detection threshold, but a sufficiently high threshold is necessary in order to avoid noise triggers which cause dead time. Increasing the fast filter peaking time also helps with noise trigger reduction and its flattop can account for the signal rise time. As discussed in section 6.4.1, all three of these parameters are connected to the minimum resolution time which favors lower fast filter settings.

The energy filter peaking time can be optimized in regard to the overall energy resolution. A flattop is again required to account for signal rise time and a low flattop is preferable for the reduction of trigger dead time. However, as demonstrated in section 6.4.2 it can be increased to ensure a better estimation of pileup energy.

The recovery needs to be long enough to avoid faulty event triggers due to features in the signal waveform during resets and can easily be chosen by inspecting the waveform.

In regards to baseline restoration, the manual recommends choosing a number of baseline samples close to the energy filter peaking time. Specifics about the baseline restoration algorithm are unfortunately undisclosed by the manufacturer.

For improved pileup rejection, the DANTE DPP offers a maximum rise time setting. Pileup events exhibit a longer effective rise time, so by rejecting those events above the maximum rise time threshold, a significant amount of pileup can be avoided in addition to the rejection through trigger dead time. For an explanation of the latter see chapter 6. The maximum rise time setting effectively reduces the minimum resolution time, but if chosen too low it can reject regular events as well. Correct and false positive rejections are studied in section 7.3.4 of this chapter.

7.2 Reset event loss

As discussed chapter 6, relying on active reset preamplifiers introduces another source of dead time, next to trigger dead time, for covering the recovery period after a reset. This alone is not enough to account for event loss caused by resets, because it does not incorporate events that trigger a reset. This effect is schematically demonstrated in figure 7.2. The events that push the voltage over the reset threshold happen before the recovery dead time is initiated so they are not corrected for by live time correction. Unfortunately, they cannot be registered as regular events either, since the reset happens virtually instantaneously after the threshold is crossed, leaving no room for proper event detection. Furthermore, it has to be considered that resets can be initiated due to the ramp slope as well. With the assumption of a constant leakage current the number of resets N_r can be estimated by dividing the total accumulated voltage from leakage and events by the ramp height U_ramp:

$$N_\mathrm{r} = \frac{t}{U_\mathrm{ramp}} \left(\left(\frac{\mathrm{d}U}{\mathrm{d}t} \right)_\mathrm{leak} + \langle U_\mathrm{evt} \rangle \cdot r \right) . \tag{7.1}$$

Here, $\langle U_\mathrm{evt} \rangle$ is the mean voltage generated per event, $\left(\frac{\mathrm{d}U}{\mathrm{d}t}\right)_\mathrm{leak}$ the slope due to leakage current, r the true rate and t the measurement time. According to this model the fraction of resets involving an event loss is rate dependent, which is shown in figure 7.3. However, when normalized to the total number of events $N_\mathrm{evt} = rt$, it reveals that the relative event loss only depends on the fraction of event step height and ramp amplitude:

$$\frac{N_\mathrm{r,evt}}{N_\mathrm{r}} = \frac{\langle U_\mathrm{evt} \rangle}{U_\mathrm{ramp}} . \tag{7.2}$$

This equation can be applied for correction of the estimated input count rate, which could be useful for integral spectrum measurements.

In the case of detector emulator measurements, reset event losses are much more pronounced. Since the emulator is incapable of generating the leakage slope, every reset causes an event loss. The estimated rate can then simply be corrected by adding the number of resets to the event count, which is demonstrated in section 7.3.1.

For differential spectrum acquisitions it has to be considered that higher energetic events are more likely to cause a reset resulting in an energy dependent efficiency. This is shown in figure 7.4 with the emulation of a flat spectrum. Generalizing equation 7.2 leads to the efficiency

$$\varepsilon_\mathrm{reset} = 1 - \frac{U_\mathrm{evt}(E)}{U_\mathrm{ramp}} , \tag{7.3}$$

which can be used for correction.

These correction methods apply to single channel measurements. In the case of 7-channel measurements with the full prototype setup, it has to be considered that all channels are reset simultaneously as soon as one of them reaches the threshold. Therefore, in multichannel operation reset event loss can be expected to be less severe.

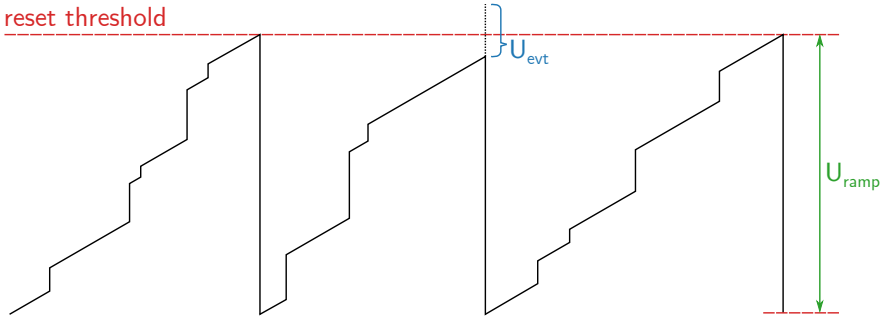


Figure 7.2: Example for the loss of an event that triggers a preamplifier reset. If a reset is caused by the leakage current slope instead, no event is lost.

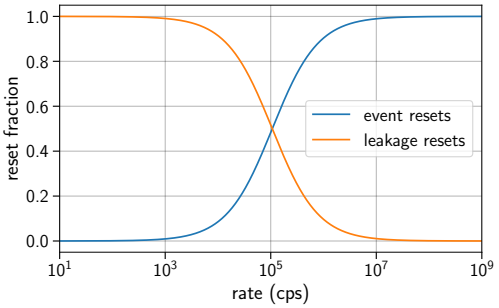


Figure 7.3: Fraction of resets caused by events or leakage current as implied by the model from equation 7.1. The fractions are dependent on rate and leakage slope. Here the value $\left(\frac{\mathrm{d}U}{\mathrm{d}t}\right)_{\mathrm{leak}} = 70\,\mathrm{ADU/\mu s}$ is used, which is determined from regular prototype setup operation.

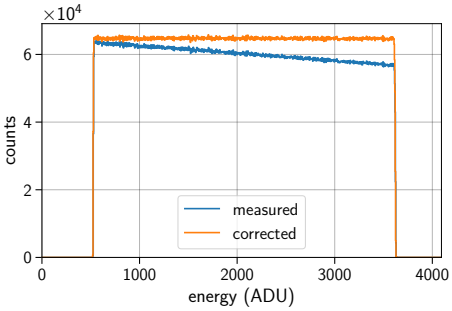


Figure 7.4: Emulation of a flat spectrum demonstrating the energy dependence of reset event losses. In order to avoid pileup, constant event time intervals are used. Higher event energies are more likely to cause a reset, so the measured spectrum exhibits a slope. Dividing by the efficiency given in equation 7.3 corrects for this.

7.3 Emulator measurements

This section summarizes several emulator measurements that were conducted in order to evaluate the high rate behavior of the DANTE DPP.

7.3.1 Rate estimation

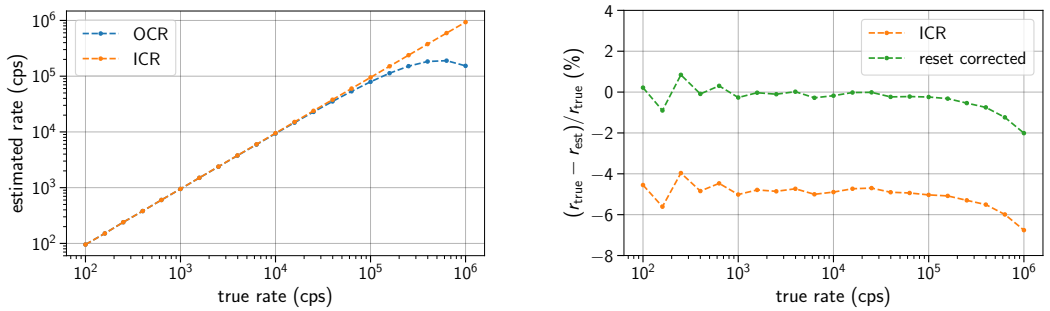
For each measurement, the DANTE DPP provides the output count rate (OCR) and an estimate of the input count rate (ICR). The OCR is calculated from the number of accepted events divided by measurement time and the ICR is obtained according to the live time correction explained in section 6.2. The detector emulator is well suited for checking the performance of these estimations, because it generates signals at stable and precise rates. For this measurement, a monoenergetic line spectrum is emulated with rates in the range of $10^2 - 10^6$ cps. The resulting OCRs and ICRs are shown in figure 7.5a. Upon closer inspection, the ICR slightly underestimates the true rate by a constant, relative shift as demonstrated in figure 7.5b. Applying the reset event correction from section 7.2 corrects for this shift. For extreme rates above 10^5 cps the rate drops even further, however, this is not explained by reset event loss.

7.3.2 Dead time cross check

Several emulation measurements are conducted in order to cross check the analytic dead time model from section 6.2.

First, a list mode measurement of a monoenergetic spectrum with 10^5 cps is performed for verifying that the minimal time difference between events is never smaller than the expected trigger dead time $\tau_{\text{dead}} = P + F$, where P is the peaking time and F the flattop of the energy filter. For this measurement $P = 0.8\,\mu\text{s}$ and $F = 0.128\,\mu\text{s}$ are chosen. The histogram of time differences in figure 7.6 confirms that no accepted event time difference is shorter than τ_{dead}.

In a second step, the dead time output of the DANTE DPP is compared to the expectation from equation 6.8. Again, a line spectrum is emulated and multiple spectra are acquired for different settings of the parameters of interest that are present in the expectation formula. Namely, these are the energy filter peaking time and flattop time, recovery length and rate. Figure 7.7 shows the results and demonstrates that measurement and expectation match well.



(a) OCR and ICR over true rate.

(b) ICR underestimation and correction by considering reset event loss.

Figure 7.5: Rate detection performance of the DANTE DPP as measured from emulated line spectra at different rates.

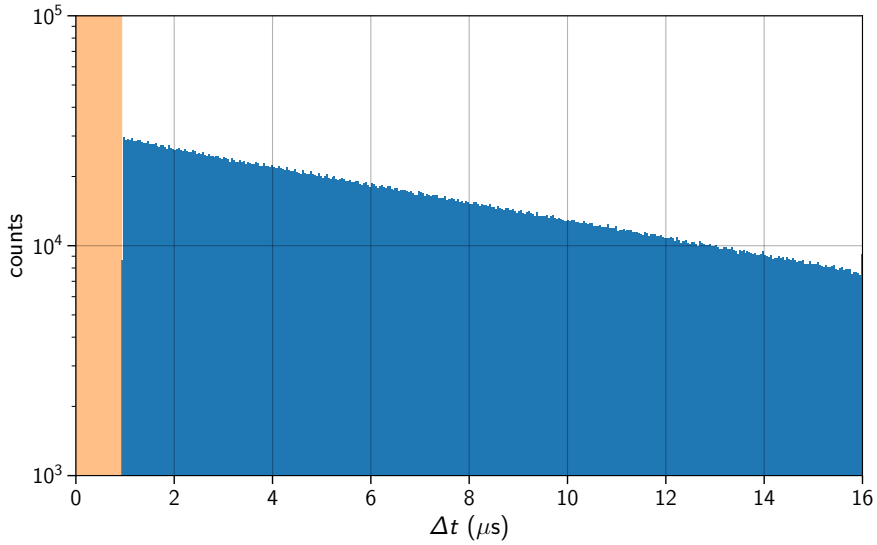


Figure 7.6: Distribution of time differences from a list mode measurement. A monoenergetic line spectrum is emulated with 10^5 cps and as expected no output events are closer in time than the trigger dead time marked in orange.

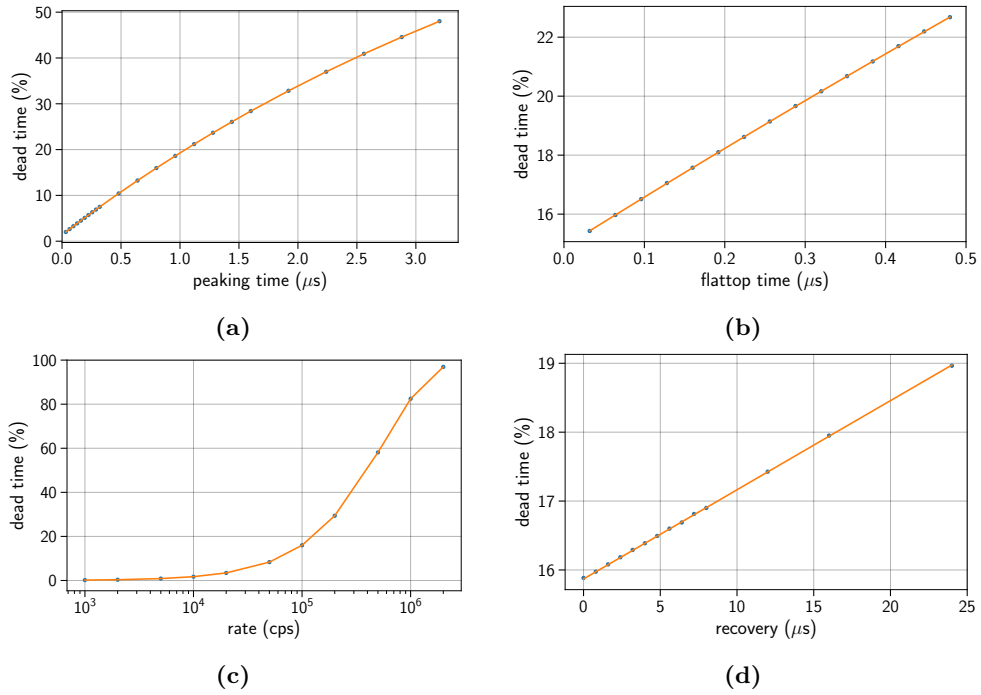


Figure 7.7: Cross check of the dead time model from equation 6.8. The measured dead times (blue dots) match the expectation (orange lines) very well.

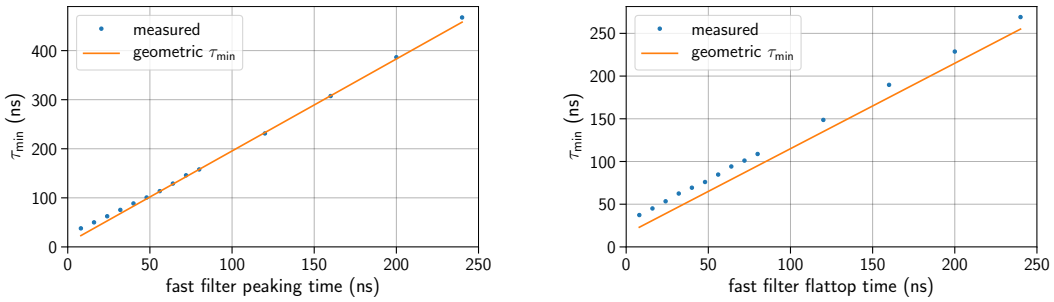
7.3.3 Minimum resolution time

The minimum resolution time τ_{min} of the DANTE trigger can be inferred from the fraction of unresolved pileup events following the procedure outlined in section 6.4.1. A measurement of this property is useful for predicting the amount of pileup at any given rate. For this purpose, a line spectrum is emulated at sufficiently high energy for proper separation of the main peak and the pileup spectrum, even if the pileup energy can be significantly underestimated for large τ_{min}. Since the rate is precisely known from the emulation setting $r = 10^5$ cps, the minimum resolution time can be calculated by

$$\tau_{\text{min}} = -\frac{1}{r} \cdot \ln\left(1 - \frac{N_{\text{pileup}}}{N_{\text{main}}}\right) \quad . \tag{7.4}$$

This is done for a range of fast filter peaking time and flattop settings. In figure 7.8 the results are compared to the expectation from the geometric formula from equation 6.10. Measurement and expectation are in agreement, however there is some tension for low peaking times. This is clearly apparent in figure 7.8b, where the peaking time was fixed at one 8 ns sample.

Overall, with fast filter settings less than ≈ 100 ns, which is well within the realm of realistic operation parameters for room temperature conditions, resolution times below 200 ns are achievable. With good noise reduction $\tau_{\text{min}} < 100$ ns seems realistic even without further pileup rejection mechanisms.



(a) Peaking time scan with fixed flattop at $F = 8\,\text{ns}$

(b) Flattop scan with fixed peaking time at $P = 8\,\text{ns}$

Figure 7.8: Measurement of the minimum resolution time for several fast filter settings (blue dots) with a comparison to the expectation from equation 6.10 (orange line).

7.3.4 Maximum rise time rejection efficiency

Pileup rejection can be significantly improved by enabling the maximum rise time setting. If enabled the DPP analyzes the shape of the fast filter pulse and determines an estimated rise time. If this value is larger than a specified setting, the event is rejected. Since pileup events exhibit an effectively larger rise time than regular events, they are primarily discarded. However, occasionally regular events can falsely be tagged as pileup, especially for low rejection settings.

In this measurement, both correct and incorrect rejection efficiencies are determined in the simple case of a monoenergetic spectrum. The spectrum is emulated at a rate of $10^5\,\text{cps}$ for different maximum rise time settings and once with it disabled. Correct and incorrect rejection efficiencies are calculated from the number of events in pileup and main peak compared to the case of no rejection

$$\varepsilon_{\text{true}} = 1 - \frac{N_{\text{pileup}}}{N_{\text{pileup,off}}} \quad \text{and} \quad \varepsilon_{\text{false}} = 1 - \frac{N_{\text{main}}}{N_{\text{main,off}}} \,. \tag{7.5}$$

The results are shown in figure 7.9 which demonstrates that up to $\approx 90\,\%$ of pileup events can be rejected without removing events from the main spectrum. But it should be noted that this measurement is performed under idealized conditions and for a real measurement the performance may differ significantly. Firstly, the rise time of the emulated pulses is presumably very uniform. Rise times from a real detector signal may fluctuate more, due to different drift paths in the SDD, which would lead to more false-positive rejections. Secondly, if maximum rise time rejection is used it should be investigated whether it introduces other systematic effects such as an energy dependence. Lastly, the estimated rise time value is not an output variable due to throughput limitations. And since the specific implementation is unknown it is unclear how well the estimate represents the true rise time.

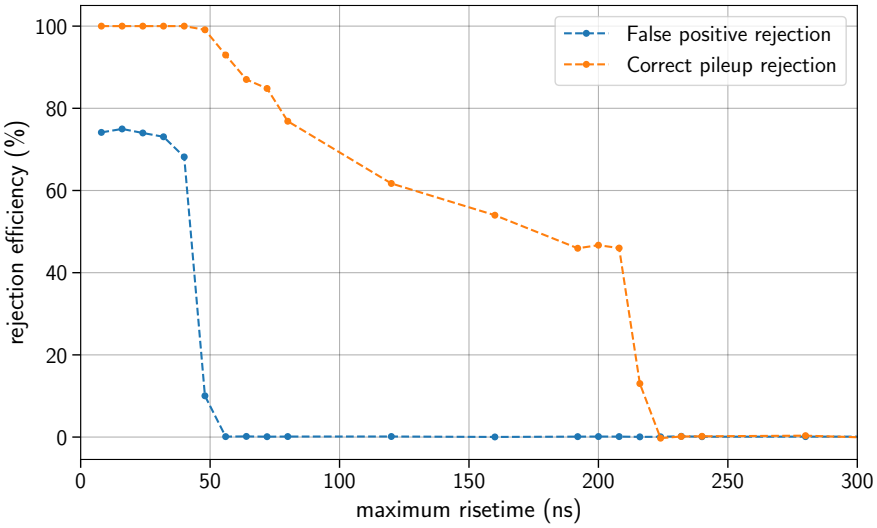


Figure 7.9: Measurement of the maximum rise time pileup rejection efficiency as well as the fraction of false positive rejection of regular events. Calculated from emulated the line spectra.

CHAPTER 8

Conclusion

The KATRIN apparatus, and in particular its high luminosity β-electron source, offer a unique opportunity to go beyond the neutrino mass measurement and expand the physics program to search for signatures of keV sterile neutrinos. With a precise measurement of the differential tritium spectrum, the TRISTAN project aims to reach a sensitivity to sterile mixing amplitudes of $\sin^2\theta < 10^{-6}$. In order to achieve this goal, an excellent intrinsic energy resolution and capability of handling count rates on the order of 10^8 cps are fundamental requirements for the detector system. This necessitates an upgrade of the detector system, where the high rate requirement mainly concerns the number of detector pixels and the data acquisition system.

While the DAQ for TRISTAN is not yet finalized, the current prototype detector system is utilized to investigate rate dependent systematics as a pilot study. For this purpose, the DAQ system DANTE, the main system used for acquisitions with the prototype detector, is characterized with regard to its high rate capabilities. A detector emulator is employed for studying peak pileup and the associated minimum resolution time, as well as dead time and its connection to input rate estimation.

The DANTE system provides two-fold pileup rejection through the combination of a fast trigger and maximum rise time rejection. The measurements show that sufficiently low resolution times can be achieved to limit the fraction of pileup events to below 1 % at TRISTAN's design rate of 10^5 cps per pixel. In terms of rate determination accuracy, the default input count rate estimation from live time correction is slightly off, as it does not account for preamplifier event reset losses. A correction method is proposed which enables a more reliable rate acquisition up to 10^5 cps.

In addition to these measurement results, this work focuses on exploring the details of rate dependent systematics on a more general level, allowing for application to other systems. Modeling and correction approaches for dead time, tail pileup, and peak pileup are discussed and checked with a trace level simulation software that was developed in the course of this investigation. Furthermore, the dead time and minimum resolution time models are successfully tested with the DANTE system and the detector emulator.

A key result in regard to pileup is the determination of pileup event energy and how faulty pileup energies can be avoided. For modeling of the pileup spectrum, a convolution approach is proposed, which is successfully verified with a trace level simulation to a

high degree of accuracy. This convolution model is then utilized to estimate the impact of pileup on the final sensitivity of the TRISTAN project. The study yields a sensitivity reduction over the entire probed sterile neutrino mass range and a reduction factor of ~ 0.5 in the central mass region.

Furthermore, this investigation demonstrates that the impact of pileup can be mitigated by using post acceleration to separate the pileup spectrum from the tritium spectrum. This recovers almost all sensitivity if the post acceleration energy is higher than the tritium endpoint energy. The effectiveness of this method is compromised by effects that extend the pileup spectrum to lower energies. These include energy loss in the detector entrance window, backscattering, and the underestimation of pileup energies when no, or insufficient, pileup rejection mechanisms are employed.

CHAPTER 9

Bibliography

[Bel04] A. Bellerive. „Review of Solar Neutrino Experiments“. *International Journal of Modern Physics A* 19.08 (Mar. 30, 2004), pp. 1167–1179.
DOI: 10.1142/S0217751X04019093.

[Pau30] W. Pauli. *Pauli letter collection: letter to Lise Meitner*. Dec. 4, 1930.
URL: http://cds.cern.ch/record/83282.

[Fer34] E. Fermi. „Versuch einer Theorie der β-Strahlen. I“. *Zeitschrift für Physik* 88.3 (Mar. 1, 1934), pp. 161–177.
DOI: 10.1007/BF01351864.

[RC53] F. Reines and C. L. Cowan. „Detection of the Free Neutrino“. *Physical Review* 92.3 (Nov. 1, 1953), pp. 830–831.
DOI: 10.1103/PhysRev.92.830.

[Dan62] G. Danby et al. „Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos“. *Physical Review Letters* 9.1 (July 1, 1962), pp. 36–44.
DOI: 10.1103/PhysRevLett.9.36.

[Kod01] K. Kodama et al. „Observation of Tau Neutrino Interactions“. *Physics Letters B* 504.3 (Apr. 12, 2001), pp. 218–224.
DOI: 10.1016/S0370-2693(01)00307-0.

[Tan18] Particle Data Group Collaboration. „Review of Particle Physics“. *Physical Review D* 98.3 (Aug. 17, 2018).
DOI: 10.1103/PhysRevD.98.030001.

[Mis06] MissMJ. „Standard Model of Elementary Particles“. June 27, 2006.
URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 04/01/2019).

[CMS12] CMS Collaboration. „Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC“. *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61.
DOI: 10.1016/j.physletb.2012.08.021.

[ATL12] ATLAS Collaboration. „Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC“. *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29.
DOI: 10.1016/j.physletb.2012.08.020.

[Fuk98] Super-Kamiokande Collaboration. „Evidence for Oscillation of Atmospheric Neutrinos“. *Physical Review Letters* 81.8 (Aug. 24, 1998), pp. 1562–1567.
DOI: 10.1103/PhysRevLett.81.1562.

[Wu57] C. S. Wu et al. „Experimental Test of Parity Conservation in Beta Decay“. *Physical Review* 105.4 (Feb. 15, 1957), pp. 1413–1415.
DOI: 10.1103/PhysRev.105.1413.

[GGS58] M. Goldhaber, L. Grodzins, and A. W. Sunyar. „Helicity of Neutrinos“. *Physical Review* 109.3 (Feb. 1, 1958), pp. 1015–1017.
DOI: 10.1103/PhysRev.109.1015.

[Vol02] R. R. Volkas. „Introduction to Sterile Neutrinos“. *Progress in Particle and Nuclear Physics* 48.1 (2002), pp. 161–174.
DOI: 10.1016/S0146-6410(02)00122-9.

[Moh07] R. N. Mohapatra et al. „Theory of Neutrinos: A White Paper“. *Reports on Progress in Physics* 70.11 (Oct. 11, 2007), pp. 1757–1867.
DOI: 10.1088/0034-4885/70/11/R02.

[Pon68] B. Pontecorvo. „Neutrino Experiments and the Problem of Conservation of Leptonic Charge“. *Journal of Experimental and Theoretical Physics* 53.5 (May 1968).
URL: http://www.jetp.ac.ru/cgi-bin/e/index/r/53/5/p1717?a=list (visited on 04/03/2019).

[DHH68] R. Davis, D.S. Harmer, and K.C. Hoffman. „Search for Neutrinos from the Sun“. *Physical Review Letters* 20.21 (May 20, 1968), pp. 1205–1209.
DOI: 10.1103/PhysRevLett.20.1205.

[PS01] A.W.P. Poon and SNO Collaboration. „Neutrino Observations from the Sudbury Neutrino Observatory“. *AIP Conference Proceedings* 610 (Oct. 7, 2001), pp. 218–230.
DOI: 10.1063/1.1469931.

[Kaj15] T. Kajita. „The Nobel Prize in Physics 2015“. 2015.
URL: https://www.nobelprize.org/prizes/physics/2015/kajita/diploma/ (visited on 04/14/2019).

[McD15] A.B. McDonald. „The Nobel Prize in Physics 2015“. 2015.
URL: https://www.nobelprize.org/prizes/physics/2015/mcdonald/diploma/ (visited on 04/14/2019).

[Est19] I. Esteban et al. „Global Analysis of Three-Flavour Neutrino Oscillations: Synergies and Tensions in the Determination of $\Theta 23$, δCP, and the Mass Ordering“. *Journal of High Energy Physics* 1/2019.1 (Jan. 11, 2019), p. 106.
DOI: 10.1007/JHEP01(2019)106.

[Est18] I. Esteban et al. „NuFIT 4.0“. 2018.
URL: `http://www.nu-fit.org/` (visited on 04/14/2019).

[Dre13] G. Drexlin et al. „Current Direct Neutrino Mass Experiments“. *Advances in High Energy Physics* 2013 (2013), pp. 1–39.
DOI: `10.1155/2013/293986`.

[Agh18] Planck Collaboration. „Planck 2018 Results. VI. Cosmological Parameters“ (July 17, 2018). arXiv: `1807.06209`.

[Guz15] P. Guzowski et al. „Combined Limit on the Neutrino Mass from Neutrinoless Double-β Decay and Constraints on Sterile Majorana Neutrinos“. *Physical Review D* 92.1 (July 1, 2015), p. 012002.
DOI: `10.1103/PhysRevD.92.012002`.

[PRV10] G. Pagliaroli, Fernando Rossi-Torres, and Francesco Vissani. „Neutrino Mass Bound in the Standard Scenario for Supernova Electronic Antineutrino Emission“. *Astroparticle Physics* 33.5-6 (June 1, 2010), pp. 287–291.
DOI: `10.1016/j.astropartphys.2010.02.007`.

[Ase11] V. N. Aseev et al. „Upper Limit on the Electron Antineutrino Mass from the Troitsk Experiment“. *Physical Review D* 84.11 (Dec. 8, 2011), p. 112003.
DOI: `10.1103/PhysRevD.84.112003`.

[Kra05] Ch Kraus et al. „Final Results from Phase II of the Mainz Neutrino Mass Searchin Tritium β-Decay“. *The European Physical Journal C - Particles and Fields* 40.4 (Apr. 1, 2005), pp. 447–468.
DOI: `10.1140/epjc/s2005-02139-7`.

[Mer16] S. Mertens. „Direct Neutrino Mass Experiments“. *Journal of Physics: Conference Series* 718 (May 2016), p. 022013.
DOI: `10.1088/1742-6596/718/2/022013`.

[KAT05] KATRIN Collaboration. „KATRIN Design Report 2004“. Forschungszentrum, Karlsruhe, Feb. 2005.
DOI: `10.5445/IR/270060419`.

[ALE06] ALEPH Collaboration et al. „Precision Electroweak Measurements on the Z Resonance“. *Physics Reports* 427.5-6 (May 2006), pp. 257–454.
DOI: `10.1016/j.physrep.2005.12.006`.

[Adh17] R. Adhikari et al. „A White Paper on keV Sterile Neutrino Dark Matter“. *Journal of Cosmology and Astroparticle Physics* 2017.01 (Jan. 13, 2017), pp. 025–025.
DOI: `10.1088/1475-7516/2017/01/025`.

[Boy19] A. Boyarsky et al. „Sterile Neutrino Dark Matter“. *Progress in Particle and Nuclear Physics* 104 (2019), pp. 1–45.
DOI: `10.1016/j.ppnp.2018.07.004`.

[Agu18] MiniBooNE Collaboration. „Significant Excess of ElectronLike Events in the MiniBooNE Short-Baseline Neutrino Experiment“. *Physical Review Letters* 121.22 (Nov. 26, 2018), p. 221801.
DOI: 10.1103/PhysRevLett.121.221801.

[Men11] G. Mention et al. „Reactor Antineutrino Anomaly“. *Physical Review D* 83.7 (Apr. 29, 2011), p. 073006.
DOI: 10.1103/PhysRevD.83.073006.

[An14] Daya Bay Collaboration. „Search for a Light Sterile Neutrino at Daya Bay“. *Physical Review Letters* 113.14 (Oct. 1, 2014), p. 141802.
DOI: 10.1103/PhysRevLett.113.141802.

[Alm18] STEREO Collaboration. „Sterile Neutrino Constraints from the STEREO Experiment with 66 Days of Reactor-On Data“. *Physical Review Letters* 121.16 (Oct. 17, 2018), p. 161801.
DOI: 10.1103/PhysRevLett.121.161801.

[DW94] Scott Dodelson and L. M. Widrow. „Sterile Neutrinos as Dark Matter“. *Physical Review Letters* 72.1 (Jan. 3, 1994), pp. 17–20.
DOI: 10.1103/PhysRevLett.72.17.

[BRS09] A. Boyarsky, O. Ruchayskiy, and M. Shaposhnikov. „The Role of Sterile Neutrinos in Cosmology and Astrophysics“. *Annual Review of Nuclear and Particle Science* 59.1 (Oct. 21, 2009), pp. 191–214.
DOI: 10.1146/annurev.nucl.010909.083654.

[ABS05] T. Asaka, S. Blanchet, and M. Shaposhnikov. „The νMSM, Dark Matter and Neutrino Masses“. *Physics Letters B* 631.4 (Dec. 29, 2005), pp. 151–156.
DOI: 10.1016/j.physletb.2005.09.070.

[AS05] T. Asaka and M. Shaposhnikov. „The νMSM, Dark Matter and Baryon Asymmetry of the Universe“. *Physics Letters B* 620.1 (July 28, 2005), pp. 17–26.
DOI: 10.1016/j.physletb.2005.06.020.

[Mer18] S. Mertens et al. „A Novel Detector System for KATRIN to Search for keV-Scale Sterile Neutrinos“ (Oct. 15, 2018). arXiv: 1810.06711.

[Are16] KATRIN Collaboration. „Commissioning of the Vacuum System of the KATRIN Main Spectrometer“. *Journal of Instrumentation* 11.04 (Apr. 7, 2016), P04011–P04011.
DOI: 10.1088/1748-0221/11/04/P04011.

[Kle18] M. Kleesiek et al. „β-Decay Spectrum, Response Function and Statistical Model for Neutrino Mass Measurements with the KATRIN Experiment“ (June 1, 2018).
DOI: 10.1140/epjc/s10052-019-6686-7.

[Ams15] J. F. Amsbaugh et al. „Focal-Plane Detector System for the KATRIN Experiment“. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 778 (Apr. 1, 2015), pp. 40–60.
DOI: 10.1016/j.nima.2014.12.116.

[Mer15a] S. Mertens et al. „Sensitivity of Next-Generation Tritium Beta-Decay Experiments for keV-Scale Sterile Neutrinos“. *Journal of Cosmology and Astroparticle Physics* 2015.02 (Feb. 17, 2015), pp. 020–020.
DOI: 10.1088/1475-7516/2015/02/020.

[Ste13] N. Steinbrink et al. „Neutrino Mass Sensitivity by MAC-E-Filter Based Time-of-Flight Spectroscopy with the Example of KATRIN“. *New Journal of Physics* 15.11 (Nov. 11, 2013), p. 113020.
DOI: 10.1088/1367-2630/15/11/113020.

[KAT17] KATRIN Collaboration. „KIT - KATRIN -Detector“. May 15, 2017.
URL: https://www.katrin.kit.edu/84.php (visited on 04/24/2019).

[Mer15b] S. Mertens et al. „Wavelet Approach to Search for Sterile Neutrinos in Tritium β-Decay Spectra“. *Physical Review D* 91.4 (Feb. 24, 2015), p. 042005.
DOI: 10.1103/PhysRevD.91.042005.

[Kor16] M. Korzeczek. „eV- & keV-Sterile Neutrino Studies with KATRIN“. May 2016.
URL: https://www.katrin.kit.edu/publikationen/MKorz-MA-eV-keV%20sterile%20Neutrinos.pdf.

[GR84] E. Gatti and P. Rehak. „Semiconductor Drift Chamber — An Application of a Novel Charge Transport Scheme“. *Nuclear Instruments and Methods in Physics Research* 225.3 (Sept. 1, 1984), pp. 608–614.
DOI: 10.1016/0167-5087(84)90113-3.

[Tho15] K. Thompson. „Technology Focus: Silicon Drift Detectors“. Jan. 15, 2015.
URL: https://www.thermofisher.com/blog/mining/technology-focus-silicon-drift-detectors/ (visited on 05/07/2019).

[Dol17] K. Dolde et al. „Impact of ADC Non-Linearities on the Sensitivity to Sterile keV Neutrinos with a KATRIN-like Experiment“. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 848 (Mar. 11, 2017), pp. 127–136.
DOI: 10.1016/j.nima.2016.12.015.

[HLL] HLL MPG. „Halbleiterlabor of the Max-Planck-Society“.
URL: http://www.hll.mpg.de/ (visited on 04/29/2019).

[XGLa] XGLab. „XGLab - X and Gamma Ray Electronics“.
URL: https://www.xglab.it/ (visited on 04/29/2019).

[Sie19] D. B. Siegmann. „Investigation of the Detector Response to Electrons of the TRISTAN Prototype Detectors“. Apr. 15, 2019.
URL: https://www.katrin.kit.edu/publikationen/2019_Master_Thesis_Siegmann_TRISTAN_Detector_Response_Final.pdf.

[GK13] Z. Guzik and T. Krakowski. „Algorithms for Digital γ-Ray Spectroscopy“. *Nukleonika* 58 (Jan. 1, 2013), pp. 333–338.
URL: https://www.researchgate.net/publication/285919105_Algorithms_for_digital_g-ray_spectroscopy.

[Tin08] C. Tintori. „Digital Pulse Processing in Nuclear Physics“ (Aug. 8, 2008), p. 21.
URL: https://www.caen.it/documents/News/20/WP2081_DigitalPulseProcessing_Rev_2.1.pdf.

[Pyt] Python Software Foundation. *Python Language Reference*. Version 3.6.8.
URL: https://www.python.org/ (visited on 05/22/2019).

[Num] NumPy Developers. *NumPy Package Reference*. Version 1.15.4.
URL: https://www.numpy.org/ (visited on 05/22/2019).

[Sci] SciPy community. *Scipy.Signal Package Reference*. Version 1.2.1.
URL: https://docs.scipy.org/doc/scipy/reference/signal.html (visited on 05/22/2019).

[UP18] S. Usman and A. Patil. „Radiation Detector Deadtime and Pile up: A Review of the Status of Science“. *Nuclear Engineering and Technology* 50.7 (Oct. 1, 2018), pp. 1006–1016.
DOI: 10.1016/j.net.2018.06.014.

[Eno14] S. Enomoto. *Simple Analytical Models of Signal Pile-Up*. Apr. 24, 2014.

[Dol16] K. Dolde. „Detector and Read-out Development to Search for Sterile Neutrinos with KATRIN“. Master Thesis. KIT, Karlsruhe, Mar. 31, 2016.
URL: https://www.katrin.kit.edu/publikationen/mth_dolde.compressed.pdf.

[Sta] Stat Trek. „Covariance Matrix“.
URL: https://stattrek.com/matrix-algebra/covariance-matrix.aspx (visited on 05/27/2019).

[XGLb] XGLab. „DANTE Digital Pulse Processor“.
URL: https://www.xglab.it/dpp-digital-pulse-processor.shtml (visited on 05/14/2019).

[XGLc] XGLab. „CUBE Preamplifier“.
URL: https://www.xglab.it/low-noise-preamplifier.shtml (visited on 05/28/2019).

[CAE] CAEN. „Digital Detector Emulator DT5810“.
URL: https://www.caen.it/products/dt5810/ (visited on 05/29/2019).

Danksagung

Über das letzte Jahr habe ich viel gelernt. Ich konnte einiges an Erfahrung über die Arbeit im wissenschaftlichen Betrieb sammeln, vielerlei praktische Fähigkeiten erlernen oder verbessern, und durfte einfach eine wunderbare Zeit in einer Arbeitsgruppe mit außerordentlich herzlicher und angenehmer Atmosphäre genießen. Deswegen möchte ich Prof. Guido Drexlin und Prof. Susanne Mertens danken, mir die Gelegenheit geschaffen zu haben, bei KATRIN beziehungsweise TRISTAN meine Masterarbeit abzulegen. Ich möchte mich natürlich auch bei der gesamten Kollaboration bedanken, ohne die dieses Experiment nicht existieren würde. Außerdem möchte ich Prof. Weber für das Übernehmen der Rolle als Korreferent danken. Beim Korrekturlesen waren mir Prof. Weber, als auch Stephanie und Marc eine riesige Hilfe. Danke für die Mühe, die ihr euch für mich gemacht habt.

Besonderen Dank will ich meinem Betreuer Marc aussprechen. In enger Zusammenarbeit haben wir viel Zeit im Büro, als auch im Labor verbracht. Dabei hat Marc immer einen Scherz auf den Lippen, was für eine lockere Atmosphäre sorgt. Bei vielen wichtigen Entscheidungen über meine Vorgehensweise und den Inhalt meiner Arbeit war Marc sehr wegweisend und hatte oft hilfreiche Tipps. Und dennoch hat er mir die nötige Freiheit gegeben, eigene Ideen zu verfolgen. Des Weiteren geht mein Dank auch an Leo und Anton, sowie an alle weiteren Mitglieder der TRISTAN-Gruppe. Es hat mir Freude bereitet mit euch zusammenzuarbeiten und bei solch einem R&D-Projekt dabei zu sein.

Die KATRIN-Gruppe hier in Karlsruhe ist großartig. Ich will mich bei allen für die tolle Gemeinschaft bedanken. Besonders bei denjenigen die dafür sorgen, dass immer wieder soziale Aktionen wie Stammtisch, Weihnachtsfeier, gemeinsames Grillen, Pizza-Mittwoch, etc. stattfinden. Auch bei unserer ständig wachsenden Klettergemeinschaft möchte ich mich bedanken für die gemeinsame Zeit in der Kletterhalle, als auch beim Klettern am Hauptspektrometer. Special thanks goes to Leo, Stephanie, Sanshiro, Momo, and Tobi for the amazing rock climbing trips we had!

Durch das Studium hat mich stets ein enger und treuer Freundeskreis begleitet. Sowohl beim Lernen, als auch in der Freizeit. Vielen Dank Sebastian, Julian, Wieland und Mirko! Ohne euch wäre das Studium bestimmt viel schwieriger gewesen und hätte auf jeden Fall weniger Spaß gemacht. An dieser Stelle möchte ich noch einigen weiteren engen Freunden danken, die ich durch das Studium hier in Karlsruhe kennenlernen durfte. Danke an Wolle, Laura, Maria, Isi, Felix, Josh, Jan, León, Johanna und die kleine Luise.

Zu guter Letzt möchte ich mich bei meiner Familie bedanken, die mir immer den Rücken gestärkt hat, und mich bei allem was ich mache unterstützt. Vielen Dank!