



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Event detection in Project 8

**Detection efficiency and its implications for the
first tritium β -decay spectrum recorded with
Cyclotron Radiation Emission Spectroscopy**

Dissertation zur Erlangung des Grades
„Doktor der Naturwissenschaften“

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität Mainz

vorgelegt von

Christine Claessens

geboren am 5. September 1988 in Gelnhausen

Mainz, den 10. September 2020

Christine Claessens
Experimentelle Teilchen- und Astroteilchen Physik
Institut für Physik
Staudingerweg 7
Johannes Gutenberg-Universität Mainz
D-55128 Mainz, Germany
claessen@uni-mainz.de

Abstract

The Project 8 collaboration aims to measure the absolute antineutrino mass scale from the shape of the tritium β -decay spectrum near its endpoint. To this end, the collaboration has successfully established Cyclotron Radiation Emission Spectroscopy (CRES), a new frequency-based approach to detect electrons and determine their kinetic energy. The CRES event signal consists of a sudden onset of narrow-band power persisting for a short amount of time. In this thesis, an event detection system consisting of real-time triggering and offline event reconstruction has been developed. This system searches for the characteristic signal features in the continuously digitized and processed radio-frequency signal. Reconstructing the initial frequency of each electron event makes it possible to record a decay spectrum with high precision. Since the mass of the electron antineutrino is determined from the shape distortion it induces in the tritium spectrum near the high-energy endpoint, it is essential to quantify any dependence of the electron detection efficiency on energy or, equivalently, frequency. In this work, the detection efficiency has been measured and the importance of including it in the analysis for an accurate determination of the endpoint and the extraction of the antineutrino mass has been demonstrated. In addition, the efficiency uncertainty has been propagated in the analysis of the first tritium spectrum recorded with CRES and a precision requirement for a future CRES experiment with a 40 meV target sensitivity to the antineutrino mass has been determined.

Abstract (deutsch)

Das Ziel der Project 8 Kollaboration ist die Messung der absoluten Antineutrino-Massenskala anhand der Form des Tritium-Zerfallsspektrums nahe seines Endpunkts. Zu diesem Zweck hat die Kollaboration erfolgreich die Zyklotron-Strahlungsemissionsspektroskopie (CRES) etabliert, einen neuen frequenzbasierten Ansatz zum Nachweis von Elektronen und zur gleichzeitigen Bestimmung ihrer kinetischen Energie. Das Signal eines CRES-Ereignisses besteht aus dem Auftreten eines schmalbandigen Peaks im Frequenzspektrum, das für kurze Zeit andauert. Im Rahmen dieser Doktorarbeit wurde ein Ereignis-Detektionssystem, bestehend aus einem Echtzeit-Trigger und einer Offline-Ereignisrekonstruktion, entwickelt. Dieses System sucht in dem kontinuierlich digitalisierten und verarbeiteten Hochfrequenzsignal nach den charakteristischen CRES-Merkmalen. Die Rekonstruktion der Startfrequenz jedes Elektronenereignisses ermöglicht die hochpräzise Aufnahme eines Zerfallsspektrums. Da die Masse des Elektron-Antineutrinos aus der Formverzerrung bestimmt wird, die es im Tritiumspektrum in der Nähe des Endpunkts induziert, ist es unerlässlich, jede Abhängigkeit der Effizienz des Elektronennachweises von der Energie bzw. der Frequenz zu quantifizieren. In dieser Arbeit wurde die Detektionseffizienz gemessen und die Wichtigkeit ihrer Einbeziehung in die Analyse für eine genaue Messung des Endpunkts und die Extraktion der Antineutrinomasse demonstriert. Des Weiteren wurde ihre Messunsicherheit als systematischer Fehler in der Analyse des ersten mit CRES aufgenommenen Tritiumspektrums fortgepflanzt und eine Präzisionsanforderung bestimmt, um mit einem zukünftigen CRES-basierten Experiment eine Sensitivität auf die Antineutrinomasse von 40 meV erreichen zu können.

Contents

| | | |
|----------|---|------------|
| 1 | Introduction | 1 |
| 2 | Neutrinos and neutrino mass | 3 |
| 2.1 | Discovery of the neutrino | 3 |
| 2.2 | Neutrinos in the Standard Model | 4 |
| 2.3 | Neutrino mixing and oscillation | 5 |
| 2.4 | Massive neutrinos | 8 |
| 2.5 | Neutrino mass measurements | 12 |
| 2.6 | Neutrino mass in tritium β -decay | 17 |
| 3 | Project 8 | 24 |
| 3.1 | Cyclotron Radiation Emission Spectroscopy | 24 |
| 3.2 | A phased approach to measuring the neutrino mass | 36 |
| 4 | CRES in the Phase II apparatus | 45 |
| 4.1 | Signal characteristics in Phase II | 45 |
| 4.2 | Magnetic trap configuration | 59 |
| 5 | Data acquisition in Phase II | 65 |
| 5.1 | Analog signal processing | 66 |
| 5.2 | The ROACH2 | 70 |
| 6 | Triggering on CRES events | 78 |
| 6.1 | Psyllid: a data acquisition package for Project 8 | 78 |
| 6.2 | Trigger logic | 81 |
| 6.3 | Trigger implementation in Psyllid | 84 |
| 6.4 | Trigger performance | 92 |
| 6.5 | Ongoing development | 100 |
| 7 | Track and event reconstruction | 102 |
| 7.1 | Time series & short time Fourier transform | 103 |
| 7.2 | Background shape removal | 104 |
| 7.3 | Track identification | 105 |
| 7.4 | Event reconstruction | 111 |
| 7.5 | False events and power cuts | 113 |
| 7.6 | Data quality plots | 116 |

CONTENTS

| | | |
|-----------|---|------------|
| 8 | Performance studies with fake events | 118 |
| 8.1 | Locust | 118 |
| 8.2 | Simulation validation | 121 |
| 8.3 | Fake event studies | 130 |
| 9 | Detection efficiency in Phase II | 142 |
| 9.1 | Detection efficiency impact on the tritium endpoint | 143 |
| 9.2 | Studying frequency-dependent systematics | 148 |
| 9.3 | Constructing an efficiency curve for tritium | 165 |
| 10 | Detection efficiency as source of systematic uncertainty | 177 |
| 10.1 | Efficiency uncertainty in the Phase II tritium model | 178 |
| 10.2 | Analysis of the Phase II tritium spectrum | 185 |
| 10.3 | Detection efficiency in future Phases | 196 |
| 11 | Summary and outlook | 200 |
| | Appendices | 203 |
| A | Track and event properties | 204 |
| B | Simplified lineshape | 209 |

List of Figures

| | | |
|------|--|----|
| 2.1 | β -decay spectrum by Ellis and Wooster | 3 |
| 2.2 | Particles in the Standard Model | 4 |
| 2.3 | Neutrino mixing | 5 |
| 2.4 | Mass of the electron neutrino vs. summed mass of all mass eigenstates | 9 |
| 2.5 | Fermion masses in the Standard Model | 9 |
| 2.6 | Neutrino mass limits set in the last decades | 16 |
| 2.7 | Effect of the electron antineutrino mass on the tritium spectral shape | 20 |
| 2.8 | Molecular final state distribution of HeT ⁺ and HeD ⁺ | 21 |
| 2.9 | The KATRIN experiment | 23 |
| | | |
| 3.1 | Power spectrogram of CRES Event Zero | 26 |
| 3.2 | CRES sensitivity vs. exposure with 1% resolution uncertainties | 29 |
| 3.3 | CRES sensitivity vs. exposure with a 10% B resolution uncertainty | 30 |
| 3.4 | Experimental target sensitivities in comparison | 31 |
| 3.5 | Detection efficiency impact on the tritium spectrum | 36 |
| 3.6 | Experimental setup of CRES in a waveguide | 37 |
| 3.7 | Phase I CRES prototype detector insert and receiver electronics | 38 |
| 3.8 | Energy spectra of ^{83m} Kr lines recorded with the Phase I experiment | 39 |
| 3.9 | Free space CRES design with patch antenna array | 40 |
| 3.10 | Beamforming in the Phase III antenna array | 41 |
| 3.11 | Comparison of final state broadening in atomic and molecular tritium | 42 |
| 3.12 | Velocity selection | 44 |
| 3.13 | Phase IV schematic overview | 44 |
| | | |
| 4.1 | Phase II waveguide cell | 45 |
| 4.2 | Axial motion in a magnetic trap | 47 |
| 4.3 | Minimum trapped angle vs. trap depth | 47 |
| 4.4 | Sidebands comb structure | 49 |
| 4.5 | Harmonic trap approximation | 50 |
| 4.6 | Electron trajectory in a Penning trap | 52 |
| 4.7 | Mode impedance and coupling in the TE ₁₁ and TM ₀₁ modes | 53 |
| 4.8 | Transmitted power in the TE ₁₁ mode vs. kinetic energy | 54 |
| 4.9 | Power dependence on pitch angle in a harmonic trap | 54 |
| 4.10 | Power vs. radius and pitch angle in the TE ₁₁ mode | 55 |
| 4.11 | Inelastic scattering energy loss function | 56 |
| 4.12 | CRES lineshape of the ^{83m} Kr K-line in a shallow magnetic trap | 57 |

LIST OF FIGURES

| | | |
|------|---|-----|
| 4.13 | Start frequency histograms in harmonic traps of different depths | 60 |
| 4.14 | Start frequency histograms in harmonic traps of different depths fitted with the lineshape model | 61 |
| 4.15 | Lineshape parameter dependence on trap depth | 62 |
| 4.16 | K-line frequency positions in different traps with different depths | 63 |
| 4.17 | Q300 trap configuration for tritium data taking | 63 |
| 4.18 | Number of events vs. trap coil current in four individual harmonic traps | 64 |
| 4.19 | Q300 calibration data set | 64 |
| | | |
| 5.1 | Signal processing in Phase II | 66 |
| 5.2 | Signal processing in the conditioning box | 68 |
| 5.3 | Noise power profile and transmission function of the conditioning box . . | 68 |
| 5.4 | Image noise resulting from the low-pass filter in the conditioning box . . | 69 |
| 5.5 | ADC calibration noise power spectrum with test tone | 71 |
| 5.6 | Signal processing in a single ROACH2 channel | 72 |
| 5.7 | ROACH2 channel SNR vs. frequency | 75 |
| 5.8 | Central frequency scan procedure | 75 |
| 5.9 | Detection efficiency vs. frequency in the ROACH2 channels | 76 |
| 5.10 | Efficiency drop in adjacent channel configurations | 77 |
| 5.11 | ROACH2 channel configuration for tritium data taking | 77 |
| | | |
| 6.1 | Psyllid instances and interactions | 80 |
| 6.2 | Frequency mask example | 84 |
| 6.3 | Trigger illustration on CRES event | 84 |
| 6.4 | Psyllid node configuration in triggering mode | 86 |
| 6.5 | Example for trigger-flag processing in the event builder node | 90 |
| 6.6 | Psyllid node configuration in streaming mode | 91 |
| 6.7 | Noise trigger acquisition rate and live fraction in a threshold scan | 93 |
| 6.8 | Fraction of recorded event starts from trigger simulation | 96 |
| 6.9 | Simulated trigger rate reduction by requiring a 2 nd FMT-trigger | 97 |
| 6.10 | Simulated noise trigger rate vs. trigger threshold | 97 |
| 6.11 | Reference count rate in trigger optimization study | 98 |
| 6.12 | Count rate and data reduction in trigger test runs | 99 |
| 6.13 | Event rate vs. data volume in trigger test runs | 99 |
| | | |
| 7.1 | CRES event example | 103 |
| 7.2 | RF background shape | 105 |
| 7.3 | Example event in sparse spectrogram | 106 |
| 7.4 | Sequential lines in a sparse spectrogram | 108 |
| 7.5 | Sequential line and track rate in example Egg-file | 110 |

| | | |
|------|---|-----|
| 7.6 | Processed tracks in a sparse spectrogram | 110 |
| 7.7 | Reconstructed event in a sparse spectrogram | 112 |
| 7.8 | Noise track power distribution | 115 |
| 7.9 | Data quality plots | 117 |
| 8.1 | Locust simulation flow chart | 119 |
| 8.2 | Raw spectrogram of a simulated fake event | 121 |
| 8.3 | Event properties in un-triggered reference data | 122 |
| 8.4 | Number of tracks per event optimization in fake events | 123 |
| 8.5 | Phase II cell geometry in Kassiopeia simulation | 124 |
| 8.6 | Power distribution from particle tracking simulations | 125 |
| 8.7 | Distribution of analytically calculated coupled powers | 125 |
| 8.8 | Start pitch angle correction for longitudinal position | 126 |
| 8.9 | Simulated powers corrected for electron start conditions | 127 |
| 8.10 | Fake and real reconstructed event properties in comparison | 128 |
| 8.11 | Bin SNR comparison of real and simulated data | 129 |
| 8.12 | SNR standard deviation vs. average track SNR in real and simulated data | 130 |
| 8.13 | Fake event start frequency reconstruction error | 131 |
| 8.14 | Fake event start time error | 132 |
| 8.15 | Start times in triggered fake event acquisitions | 133 |
| 8.16 | Trigger efficiency studies with fake events | 134 |
| 8.17 | Comparison of three trigger efficiency definitions and their dependence on SNR | 135 |
| 8.18 | Trigger efficiency dependence on event length | 136 |
| 8.19 | Trigger efficiency dependence on SNR and event length | 136 |
| 8.20 | Event reconstruction efficiency dependence on SNR and event length . . . | 137 |
| 8.21 | Dependence of the trigger and reconstruction efficiency on SNR | 137 |
| 8.22 | SNR response matrix from simulation | 138 |
| 8.23 | Unfolding a reconstructed fake event SNR distribution | 139 |
| 8.24 | Trigger and reconstruction efficiency dependence on track slope | 140 |
| 8.25 | Fake event slope reconstruction error | 140 |
| 8.26 | Reconstructed average track SNR vs. track slope | 141 |
| 9.1 | Fake tritium data generated from an undistorted model | 144 |
| 9.2 | Endpoint vs. background contours in a pseudo-experiment | 145 |
| 9.3 | Fit result distributions for fake undistorted tritium data | 145 |
| 9.4 | Toy model efficiency | 146 |
| 9.5 | Fake tritium data generated from a distorted model | 147 |
| 9.6 | Biased fit result distributions for fake distorted tritium data | 147 |
| 9.7 | Unbiased fit result distributions for fake distorted tritium data | 148 |

LIST OF FIGURES

| | | |
|------|--|-----|
| 9.8 | Field-shifting solenoid | 150 |
| 9.9 | FSS scan current sequence | 152 |
| 9.10 | Predicted FSS scan temperature variation | 152 |
| 9.11 | K-line frequency position in FSS scans | 153 |
| 9.12 | Merging start frequency histograms per FSS current | 153 |
| 9.13 | Fitted start frequency histograms in the trap 4 FSS scan | 155 |
| 9.14 | Linearity of frequency shifts in FSS scans | 156 |
| 9.15 | Count rates vs. frequency in the trap 3 and 4 FSS scans | 157 |
| 9.16 | SNR vs. frequency in the trap 3 and 4 FSS scans | 157 |
| 9.17 | Count rate vs. SNR means in all single-trap FSS scans | 158 |
| 9.18 | Slopes vs. frequency in all single-trap FSS scans | 159 |
| 9.19 | Count rates vs. frequency in the trap 2 FSS scan | 159 |
| 9.20 | Lineshape parameters in the trap 3 FSS scan | 160 |
| 9.21 | Track and event lengths in the trap 3 FSS scan | 161 |
| 9.22 | Count rates and track slopes in the Q300 FSS scan | 162 |
| 9.23 | Comparison of event rates in single and multi traps | 163 |
| 9.24 | Count rate vs. SNR in the Q300 trap configuration | 164 |
| 9.25 | Lineshape parameters in Q300 FSS scan | 164 |
| 9.26 | Effect of a start frequency cut on the event rate to frequency relation . . | 165 |
| 9.27 | Comparison of received power vs. frequency in the TE_{11} mode | 166 |
| 9.28 | 2 nd unfolding matrix applied to reconstructed fake events | 167 |
| 9.29 | Example of an unfolded real SNR distribution | 168 |
| 9.30 | SNR analysis based on unfolding reconstructed SNR distributions | 169 |
| 9.31 | Single-trap field map | 170 |
| 9.32 | Simulated start frequency cut | 170 |
| 9.33 | SNR analysis based on forward folding a simulated SNR distribution . . . | 171 |
| 9.34 | SNR scaling factors in all single-trap FSS scans | 172 |
| 9.35 | Count rate vs. SNR scaling factor in all single-trap FSS scans | 173 |
| 9.36 | Scaled count rate vs. SNR scaling factor in all single-trap FSS scans . . | 174 |
| 9.37 | Energy-corrected single-trap FSS count rates | 175 |
| 9.38 | Q300 energy-corrected detection efficiency curve construction | 176 |
| 9.39 | Q300 slope-corrected detection efficiency curve construction | 176 |
| 10.1 | Tritium model multiplied with the Phase II detection efficiency curve . . . | 177 |
| 10.2 | Fitting fake data with the distorted model | 178 |
| 10.3 | Relative detection efficiency uncertainty vs. frequency | 180 |
| 10.4 | Initial distorted model fit to a fake data histogram | 181 |
| 10.5 | Statistical, systematic and combined endpoint uncertainty obtained in pseudo-experiments with a randomized efficiency curve | 182 |

| | | |
|-------|--|-----|
| 10.6 | Endpoint fit errors vs. true model endpoints in fake data experiments with randomized efficiencies | 183 |
| 10.7 | Example for a tilted efficiency curve | 184 |
| 10.8 | Effect of a tilted efficiency in fake data studies | 184 |
| 10.9 | Raw Phase II tritium data | 185 |
| 10.10 | Phase II tritium data after removing duplicate events | 186 |
| 10.11 | Molecular T ₂ final state spectrum | 187 |
| 10.12 | Validation of the approximate tritium model in fake data experiments . . . | 187 |
| 10.13 | Simplified lineshape model | 188 |
| 10.14 | All systematic endpoint uncertainties in a fake data analysis | 190 |
| 10.15 | Total systematic endpoint uncertainty in a fake data analysis | 191 |
| 10.16 | Full model validation in fake data experiments | 192 |
| 10.17 | Tritium data converted to energy and fitted with the approximate Phase II tritium model | 193 |
| 10.18 | Endpoint uncertainties in real Phase II data fit | 193 |
| 10.19 | Phase II neutrino mass fit | 195 |
| 10.20 | Neutrino mass vs. endpoint contour plot | 196 |
| 10.21 | Phase IV scenario simulation based on the Phase II efficiency curve | 197 |
| 10.22 | 90 % interval width vs. efficiency uncertainty in Phase IV simulations . . . | 197 |

List of Tables

| | | |
|-----|---|-----|
| 2.1 | Neutrino mass-squared differences | 8 |
| 3.1 | Input parameters for the analytical sensitivity calculation | 29 |
| 3.2 | Project 8 phases and milestones | 36 |
| 3.3 | Phase IV atomic tritium parameters | 44 |
| 5.1 | Specifications of the quad-core ADC. | 71 |
| 5.2 | ROACH2 packet content. | 73 |
| 6.1 | Best trigger configurations | 101 |
| 7.1 | Point properties | 107 |
| 7.2 | Event cut groups | 114 |
| A.1 | Track properties | 204 |
| A.2 | Event properties | 205 |
| B.1 | Model parameters | 209 |

Neutrinos are the most abundant particles of matter in the universe. In the Standard Model of particle physics they are described as three massless, electrically neutral leptons with spin $1/2$ and left-handed chirality. However, the observation of neutrino flavor oscillation proves that the neutrino flavor states mix and that at least two of the three mass eigenstates are massive. Up to now, cosmology and laboratory experiments could only set upper limits on the sum of the three neutrino masses $\sum m_\nu \lesssim 0.15 \text{ eV}/c^2$ and the effective mass of the antielectron flavor state $m_\beta < 1.1 \text{ eV}/c^2$ respectively. The mass limits from cosmology are stricter than the current laboratory limit, but they depend on cosmological models. Even if cosmological observations make it possible to distinguish neutrino masses from zero in the near future, an independent competitive laboratory measurement must serve as confirmation and cross-check.

The Project 8 experiment aims at extracting the antineutrino mass from the shape of the tritium spectrum. For this purpose, the Project 8 collaboration has developed Cyclotron Radiation Emission Spectroscopy (CRES) to detect electrons in a gas chamber and reconstruct their kinetic energy with high precision. Electrons in a strong magnetic field undergo cyclotron motion and emit radiation with a frequency narrowly peaked at the cyclotron frequency. This frequency is inversely proportional to the Lorentz factor γ of the electron. As a result, the electron's kinetic energy can be reconstructed with the excellent precision of frequency measurements.

Ultimately, the sensitivity of decay measurements is limited by the final state broadening of the daughter molecule and its uncertainty. To date, the most sensitive neutrino mass experiments have selected tritium as β -emitter, because the T_2 molecule is relatively simple and its final states are well understood. In order to achieve a sensitivity below $0.1 \text{ eV}/c^2$, the use of atomic instead of molecular tritium is required so that the molecular final state uncertainties are removed. In gaseous atomic tritium, recombination can only be prevented if the tritium atoms are isolated from all surfaces of the gas container. A major advantage of CRES over previous detection techniques is that the emitted electrons can be detected without extracting them from the source gas. This opens up the possibility to observe β -decays of tritium atoms that are trapped magnetically through the magnetic moment of the unpaired electron in their atomic shells.

The Project 8 collaboration has successfully shown that the direct detection of cyclotron radiation from a single electron is possible, although the emitted power is on the order of 1 fW. In fact, the $^{83\text{m}}\text{Kr}$ internal conversion spectrum has been reconstructed by recording the radiation of the emitted electrons and determining their cyclotron frequency. Furthermore, the excellent energy resolution and linearity that CRES provides has been demonstrated. The next milestone was the first molecular tritium spectrum recorded with CRES. The main difficulties that arose were the significantly longer half-life of tritium, the need for efficient data reduction, and the required accurate reconstruction of the spectral shape.

All these challenges are addressed as part of this work. In chapter 2 a short introduction to the theoretical background of massive neutrinos and the field of neutrino mass measurements is given. The CRES technique and its potential sensitivity to the effective mass of the electron antineutrino are presented in chapter 3, followed by an overview of the strategy employed by the Project 8 collaboration to achieve the 40 meV sensitivity target. In chapter 4 the properties of CRES signals originating from electrons trapped magnetically in the Phase II detector are analyzed, resulting in an optimized trap design for tritium data taking. One focus of this work is the data acquisition system, which was developed to enable continuous event detection over several months while ensuring sufficient data reduction. It is described in detail in chapters 5 and 6 and consists of analog and digital signal pre-processing in combination with a software trigger, which is designed and optimized to identify CRES events under tritium running conditions. Once recorded, the electron events are analyzed by the offline event reconstruction algorithm presented in chapter 7. The response of the event detection process to various event and RF system characteristics is one of the main sources of systematic errors in CRES. The precision and accuracy to which this response is understood is key to a successful spectrum reconstruction. Therefore, the performance of the event detection is examined in simulation in chapter 8. In chapter 9 this response is used in the analysis of magnetic field scan data to extract the dependence of the event detection efficiency on frequency and energy. Finally, in chapter 10 the impact of this efficiency and its uncertainty on the Phase II tritium spectrum analysis is investigated and an extrapolation to future CRES neutrino mass experiments is made. This allows a prediction of how well the detection efficiency must be known in order to meet the collaboration's goal to measure the antineutrino mass with a sensitivity of 40 meV.

2

Neutrinos and neutrino mass

Neutrinos are the only particles in the Standard Model that interact exclusively via the weak force (and gravity), which makes them inherently difficult to detect. Large neutrino telescopes have contributed to enormous progress in our understanding of neutrinos and their sources in recent decades. Neutrinos were long thought to be massless until neutrino flavor oscillation was discovered in the late 1990s. This chapter provides a very brief introduction to the history of the field, the mechanism of neutrino mixing, the theoretical background of possible neutrino mass models, and a summary of the approaches used to measure the neutrino mass scale.

2.1 Discovery of the neutrino

The discovery of the neutrino begins with the investigation of β -decay in the early 20th century. In 1914, James Chadwick [1] observed that the β -spectrum of "radium B and C" (^{214}Pb and ^{214}Bi) follows a continuous energy distribution which apparently violated energy conservation under the assumption that only a β -particle is emitted (figure 2.1). This observation was confirmed by Charles Ellis [2] and Lise Meitner [3], who disproved the hypothesis that the missing energy is dissipated by heat or emitted as γ -particles.

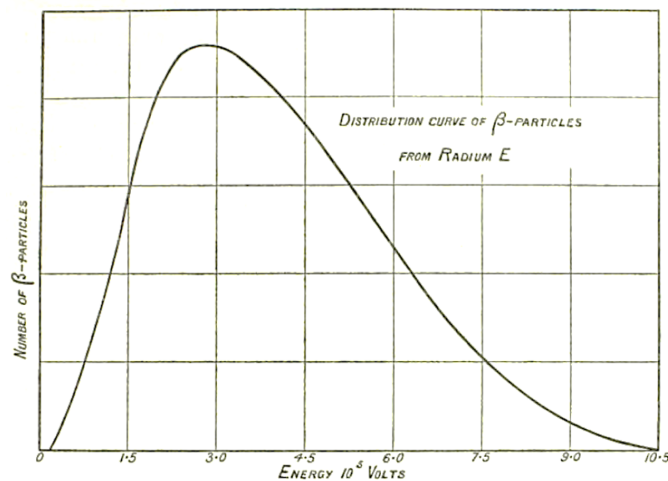


Figure 2.1: Continuous β -decay spectrum presented by Ellis and Wooster [2].

2.2. NEUTRINOS IN THE STANDARD MODEL

In a letter to Meitner in 1930, Wolfgang Pauli proposed the participation of a neutral, very light elementary particle in the decay, which he called "neutron", as an explanation for the apparent energy and momentum nonconservation [4]. In 1931 Enrico Fermi changed the name to neutrino to distinguish the newly postulated particle from the much heavier neutron. In 1933 Fermi developed a theory for β -decay inspired by electrodynamics [5]. Extended to all processes of the weak interaction, Fermi's theory is still valid in the low momentum transfer approximation. Only in 1948, the experiment of Maurice Goldhaber and Gertrude Scharff-Goldhaber proved that β -particles are electrons [6]. The first direct experimental detection of the neutrino via inverse β -decay (2.1) was achieved in 1956 by the Cowan-Reines neutrino experiment [7]

$$p^+ + \bar{\nu}_e \rightarrow n + e^+. \quad (2.1)$$

2.2 Neutrinos in the Standard Model

The Standard Model of particle physics was developed in stages throughout the 20th century and has been a highly successful theory for predicting and explaining phenomena [8–13]. It structures all known particles as shown in figure 2.2.

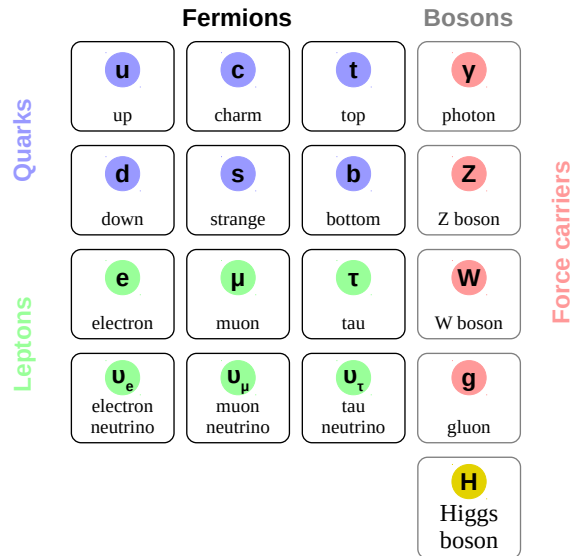


Figure 2.2: Particles in the Standard Model.

Since there was no evidence of nonzero neutrino masses at the time the Standard Model was formulated, neutrinos were introduced as massless fermions. The fermion group consists of quarks and leptons, each divided into three generations. In the leptonic sector

each generation consists of two leptons of the same flavor:

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}.$$

Neutrinos participate only in the charged-current (CC) and neutral-current (NC) weak interaction, with the lepton number being conserved separately for each flavor. In the Standard Model neutrinos and antineutrinos are distinct particles (they are Dirac fermions) with left-handed chirality. Since they are described as massless, the helicity is always negative for neutrinos and positive for antineutrinos.

2.3 Neutrino mixing and oscillation

The flavor states of neutrinos ν_e, ν_μ, ν_τ are eigenstates of the weak force and orthogonal linear combinations of the mass eigenstates ν_1, ν_2, ν_3 and vice versa. Figure 2.3 shows a graphical representation of (three-)neutrino mixing in vacuum. It can be seen that e.g. ν_2 is composed of almost equal parts of ν_e, ν_μ and ν_τ . The probabilities of detecting ν_2 as each of the three neutrino flavor states are therefore almost equal, in contrast to the composition of ν_3 , which consists of similar amounts of ν_μ and ν_τ , but has almost no ν_e contribution.

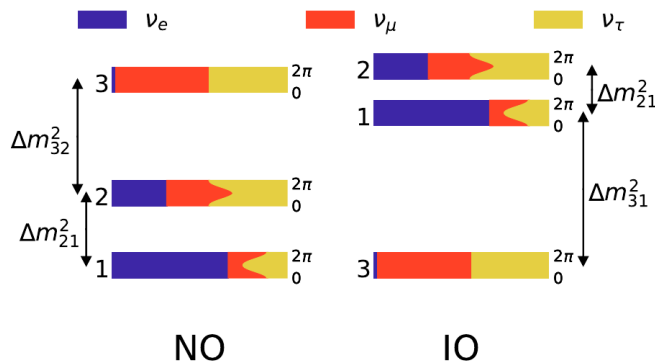


Figure 2.3: Probability of finding neutrino flavor l in the i^{th} neutrino mass eigenstate [14]. The colors reflect the mixing between the mass and the flavor basis and indicate the uncertainty of the CP-violating phase δ_{CP} . The mass-squared differences Δm_{21}^2 and $|\Delta m_{31}^2|$ are known from neutrino oscillation measurements. It is still unclear whether m_3 is the heaviest (left: normal ordering) or lightest (right: inverted ordering) of the three neutrino mass states.

2.3. NEUTRINO MIXING AND OSCILLATION

The state mixing (in the two-neutrino picture with flavors μ, τ and mass states 2, 3) can be expressed as an orthogonal transformation

$$\begin{aligned}\nu_\mu &= \cos \theta_{23} \cdot \nu_2 + \sin \theta_{23} \cdot \nu_3 \\ \nu_\tau &= -\sin \theta_{23} \cdot \nu_2 + \cos \theta_{23} \cdot \nu_3,\end{aligned}\tag{2.2}$$

where θ_{23} is the mixing angle. A neutrino state produced in a weak decay (for example a μ -neutrino in pion decay) is a coherent sum of the mass states weighted by $\cos \theta_{23}$ and $\sin \theta_{23}$. For the three-neutrino picture, equation (2.2) generalizes to

$$\nu_l = \sum_{i=1}^3 U_{li} \nu_i\tag{2.3}$$

with $l = e, \mu, \tau$. U_{li} is the 3×3 unitary mixing matrix (Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix)

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\text{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\text{CP}}} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha_{21}/2} & 0 \\ 0 & 0 & e^{i\alpha_{31}/2} \end{pmatrix}\tag{2.4}$$

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$. The phase factors α_{21} and α_{31} are only nonzero if neutrinos are Majorana particles [15] and hence their own antiparticles. If $\delta_{\text{CP}} \neq 0$, neutrino oscillation intrinsically violates CP-symmetry¹.

Neutrino oscillation

Neutrino oscillations are a consequence of the propagation of mixed neutrino states. Their discovery [16–18] provides compelling evidence that neutrinos cannot be massless and was recognized with the Nobel Prize in 2015².

During the propagation of mixed neutrino states, the relative phases of the states change with distance and time, which results in the occurrence of interference effects. The time dependence of the neutrino states is introduced by replacing the mass states in equation (2.2) by the propagating wave equation, which can be parameterized by the

¹Charge conjugation parity symmetry: If conserved, the laws of physics remain the same if a particle is replaced by its antiparticle while its spatial coordinates are inverted.

²The Nobel Prize 2015 was awarded to Takaaki Kajita and Arthur B. McDonald "for the discovery of neutrino oscillation, which shows that neutrinos have mass".

product of the shape and phase factors [19]

$$\Psi_k = g_k(x - v_k t) e^{i\Phi_k}. \quad (2.5)$$

Here, g_k is the envelope of the wave packet for the eigenstate k and depends on the group velocity v_k . The phase factor $e^{i\Phi_k}$ produces an oscillatory pattern inscribed in the envelope of the wave packet with the complex phase factor given by

$$\Phi_k(x, t) = p_k x - E_k t, \quad (2.6)$$

where p_k and E_k are the mean momentum and energy in the packet. Because the mass states ν_k have different masses, they also have different phase velocities $v_k^{\text{ph}} = \frac{E_k}{p_k}$. As a result, an additional phase difference between the mass eigenstates appears during the propagation of the wave packet, which leads to the oscillation phase

$$\Phi_{\text{osc}} \equiv \Phi_3 - \Phi_2 \approx \frac{\Delta m_{32}^2 L}{2E}, \quad (2.7)$$

where $\Delta m_{32}^2 = m_3^2 - m_2^2$ is the difference between the squared neutrino masses and E is the energy of the neutrino mass eigenstate. The increase of the phase difference with distance and time leads to a change of the oscillation pattern. For example, a ν_μ in the two-neutrino image above is composed of the two relevant mass states ν_2 and ν_3 . If the starting phase of this neutrino was $\Phi_{\text{osc}} = 0$, then with increasing time and distance Φ_{osc} becomes nonzero, leading to the appearance of the ν_τ component. If $\Phi_{\text{osc}} = \pi$, the deviation from the initial state and thus the contribution of ν_τ is maximal. If $\Phi_{\text{osc}} = 2\pi$, the initial state is restored. Therefore the probability of detecting ν_μ as ν_μ is

$$P_{\mu\mu} = 1 - P_{\mu\tau} = 1 - \sin^2(2\theta_{23}) \sin^2\left(\frac{\Delta m_{32}^2 L}{4E}\right). \quad (2.8)$$

Because of the presence of Δm_{ij} in the oscillation probabilities, the observation of neutrino oscillation proves that the masses of the neutrino mass eigenstates are different and therefore at least two of them (in the three-neutrino picture) cannot be zero. However, neutrino oscillation is neither sensitive to the absolute scale of the masses nor to the sign of Δm_{ij}^2 (in vacuum). To date, only the sign of Δm_{12}^2 has been determined from matter effects but the ordering of ν_3 relative to ν_1 and ν_2 remains an open question. The current global best fit [20] of the mass-squared differences is given in table 2.1. Here, as in the rest of this work, the speed of light c is defined as 1. A thorough review of neutrino mixing, oscillation, and mass ordering can be found for example in [14, 21, 22].

2.4. MASSIVE NEUTRINOS

| | Normal ordering | Inverted ordering |
|---------------------|--|---|
| Δm_{21}^2 | $7.42_{-0.20}^{+0.21} \cdot 10^{-5} \text{ eV}^2$ | $7.42_{-0.20}^{+0.21} \cdot 10^{-5} \text{ eV}^2$ |
| $ \Delta m_{3l}^2 $ | $2.517_{-0.028}^{+0.026} \cdot 10^{-3} \text{ eV}^2$ | $-2.498_{-0.028}^{+0.028} \cdot 10^{-3} \text{ eV}^2$ |

Table 2.1: Neutrino mass-squared differences from [20] with $\Delta m_{3l}^2 = \Delta m_{31}^2 > 0 \text{ eV}^2$ for normal ordering and $\Delta m_{3l}^2 = \Delta m_{32}^2 < 0 \text{ eV}^2$ for inverted ordering. Results include data on atmospheric neutrinos provided by the Super-Kamiokande collaboration.

2.4 Massive neutrinos

Oscillation experiments have shown that neutrino flavor states are superpositions of neutrino mass eigenstates. In addition, neutrino oscillation is sensitive to the difference in masses squared. Consequently, at least two of the three known mass eigenstates cannot be massless. Equation (2.3) yields a lower effective mass limit for the electron-neutrino mass

$$m_\beta = \sqrt{\sum_i |U_{ei}^2| m_{\nu_i}^2} \quad (2.9)$$

of $> 0.01 \text{ eV}$ for normal ordering and $> 0.05 \text{ eV}$ for inverted ordering. Since all neutrino masses are linked via the PMNS matrix, it is sufficient to measure the mass of any flavor or mass eigenstate to know all the neutrino masses (depending on the ordering). In addition, as long as the CPT³ theorem holds, antineutrinos carry the same rest mass as neutrinos. The most accessible of the neutrino masses is the effective mass of the electron antineutrino flavor state because it participates in β -decays. In this work m_β is therefore often referred to as *the* neutrino mass. Figure 2.4 shows m_β as a function of the summed mass of the mass eigenstates for normal and inverted ordering together with the current best limits set by cosmology and the KATRIN experiment (see sections 2.5.1 and 2.6.3).

2.4.1 Implications for the Standard Model

The minimal Standard Model (SM) does not allow for nonzero neutrino masses. Accordingly, neutrino oscillation implies that the Standard Model is incomplete and that new components of the quantum field theory of particle physics are required.

In the Standard Model, neutrinos are described as chiral left-handed fields ν_{lL} ($l = e, \mu, \tau$) with helicity $-\frac{1}{2}$ and zero mass. Their flavor and helicity are Lorentz-invariant and there is no inertial frame of reference in which their helicity is positive. If however, ν_l is not massless, the helicity is a linear superposition of the helicity states $-\frac{1}{2}$ and $+\frac{1}{2}$. Together

³Charge, parity, and time reversal symmetry.

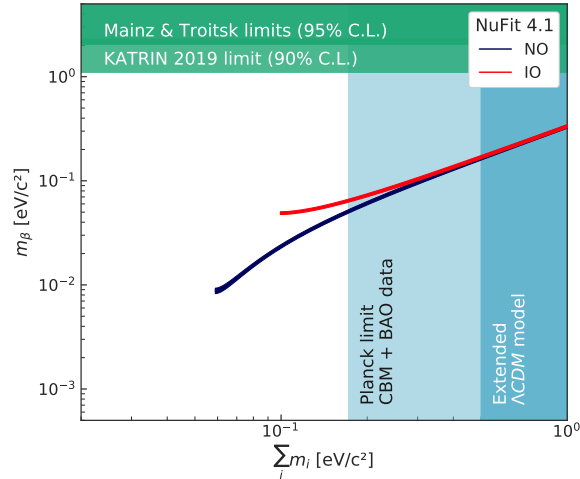


Figure 2.4: Effective mass of the electron antineutrino m_β vs. the summed mass of all mass eigenstates $\sum m_i$ based on [23, 24]. Limits on the absolute scale of the masses were established by precision β -spectroscopy [25–27] and cosmic surveys [28, 29]. Adapted image from private communication with Sebastian Böser.

with the left-handed lepton field l_L , ν_{lL} forms an $SU(2)_L$ doublet. So far, there is no evidence of the existence of right-handed neutrinos ν_R .

The current mass limits for the neutrino masses are six orders of magnitude smaller than the scale of all other fermion masses, as can be seen in figure 2.5. This in itself is not problematic, since the charged fermion masses also cover a range of six orders of magnitude. But while their masses homogeneously populate the charged-fermion mass range, there is no known mass between the neutrino masses and the electron mass at 511 keV. Ideally, an extension of the Standard Model would therefore not only provide nonzero neutrino masses but also explain why they are so small [30].

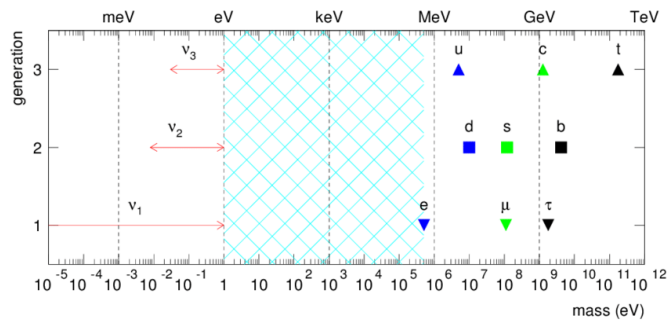


Figure 2.5: Fermion masses in the Standard Model [30]. There are no known fermions in the blue-hatched region. The red arrows indicate the allowed mass ranges for the neutrino mass eigenstates, assuming normal ordering.

2.4. MASSIVE NEUTRINOS

2.4.2 Neutrino mass models

There is a great variety of neutrino mass models, which are qualitatively very different. They contain different components and make divergent predictions about the nature of neutrinos. According to [31], the neutrino mass models can be grouped into three broad categories that complement the Standard Model:

1. New chiral fermion fields n^c : These allow for neutrino-Higgs-Yukawa type interactions that, after electroweak symmetry breaking, translate into Dirac masses for the neutrino states.
2. Other sources of electroweak symmetry breaking: A new Higgs boson field with a nonzero vacuum expectation value could contribute to nonzero neutrino masses under the condition that it does not affect any other fermion.
3. A new source of mass: The neutrino mass could be of dual nature as the consequence of two independent mass scales, the electroweak symmetry breaking scale and an additional new mass scale.

Most neutrino mass models are based on the hypothesis that neutrinos are Majorana fermions. In these models, neutrino masses are a high-energy phenomenon with heavy degrees of freedom and small neutrino masses as a consequence. The only nonrenormalizable operator of dimension five that can be constructed from SM fields and that is consistent with gauge invariance is [32]

$$\mathcal{L}_5 \rightarrow -\frac{\lambda^{\alpha\beta}}{2\Lambda}(L_\alpha H)(L_\beta H) + \text{H.c.}, \quad (2.10)$$

where Λ is the effective new physics scale, $\lambda^{\alpha\beta}$ are dimensionless coefficients, L_α is a lepton field with $\alpha = e, \mu$ or τ , H is a fundamental scalar field and H.c. is the shorthand for the Hermitian conjugates of all preceding terms. After electroweak symmetry breaking, equation (2.10) translates to

$$\mathcal{L}_5 \rightarrow -\frac{m_{\alpha\beta}}{2}\nu^\alpha\nu^\beta + \text{H.c.}, \quad (2.11)$$

where $m = \frac{\lambda v^2}{\Lambda}$. This nonzero Majorana mass is therefore a direct consequence of the presence of new physics with a new effective scale Λ . Furthermore, the neutrino Majorana masses are reduced by a factor $\frac{v}{\Lambda}$ compared to the weak scale v under the condition that $\Lambda \gg v$. Equation (2.11) has two important consequences:

1. Neutrinos are Majorana fermions.
2. Lepton number is not exactly conserved.

2.4.3 Seesaw mechanism

While many models can generate the Weinberg operator (2.10), only three of them are capable of doing so at the tree level [33–38]. All other models require that the connection between the new lepton number breaking physics and the Weinberg operator is more indirect and occurs e.g. in higher-order perturbation theory.

The seesaw I mechanism assumes that lepton number conservation is broken and both the Yukawa coupling γ and the entries of the right-handed neutrino Majorana mass matrix M_{ij} are nonzero. In this way, the Lagrangian for the combination of both right-handed neutrino potentials becomes

$$\mathcal{L}_{\text{Type I}} = \gamma^{\alpha i} L_{\alpha} n_i^c H - \frac{M^{ij}}{2} n_i^c n_j^c + \text{H.c.} \quad (2.12)$$

This equation is referred to as the type I seesaw Lagrangian. After electroweak symmetry breaking, equation (2.12) describes $(3 + N_n)$ potentially massive Majorana fermions that are all linear combinations of the ν_{α} and n_i^c fields. In this basis the mass matrix is

$$m_{\nu} = \begin{pmatrix} 0 & m_D \\ m_D^{\top} & M \end{pmatrix}, \quad (2.13)$$

where m_D are the Dirac neutrino mass matrices. The relation of M and m_D falls into three categories:

- $M \ll m_D$: The neutrino mass eigenvalues are quasi-degenerate and of order m_D . There is no natural explanation why the neutrino masses are so much smaller than the masses of the other fermions. For sufficiently large M , the mass splittings could be observed via neutrino oscillations between the active states ν_{α} and the new chiral fermion fields n^c , which are referred to as sterile neutrinos. Solar neutrino oscillations have ruled out long-wavelength, large-mixing angle oscillations and limit the right-handed neutrino masses to be smaller than $\sim 10^{-9}$ eV [39].
- $M \simeq m_D$: There are six neutrino mass eigenstates and the Majorana neutrino masses are approximately of the order of m_D . This scenario is excluded because there is no evidence for the existence of other oscillation frequencies of the same order of magnitude as the known atmospheric and solar oscillation frequencies.
- $M \gg m_D$: The neutrino masses split into two subsets. Heavy neutrinos of mass M , which consist mainly of the sterile neutrino states n^c , and light neutrinos, which consist almost exclusively of active states ν_{α} . The active neutrino masses would be lighter by a factor $\frac{m_D}{M}$ than the weak scale. This would explain why the active masses are so small compared to all other fermion masses. Since the neutrino

2.5. NEUTRINO MASS MEASUREMENTS

masses are reduced by the large values in M , the Yukawa coupling γ can be of the same order of magnitude as those of the other charged fermions. In this limit, the right-handed neutrino fields can be integrated out to obtain equation (2.10).

In equation (2.12) the singlet fermion n^c could be replaced with $SU(2)_L$ triplet fields with zero hypercharge. This is referred to as type III seesaw mechanism. Another possibility to generate the Weinberg operator is the type II seesaw mechanism. This approach adds a new scalar field ξ to the SM Lagrangian. ξ is a $SU(2)_L$ triplet with hypercharge -1 and couples only to the lepton doublet fields L^α via a Yukawa interaction

$$\mathcal{L}_{\text{triplet}} = \frac{f_{\alpha\beta}}{2} L^\alpha L^\beta \xi + \text{H.c.}, \quad (2.14)$$

where $f_{\alpha\beta}$ are dimensionless couplings. If the neutral component had a vacuum expectation value $\langle \xi \rangle$, equation (2.14) would lead to Majorana neutrino masses $m_\nu = f \langle \xi \rangle$. In this scenario, the neutrino masses are fundamentally different from the masses of charged fermions since they would be proportional to a different source of electroweak symmetry breaking. The small observed neutrino masses would require $\langle \xi \rangle \ll 100 \text{ GeV}$, which is much smaller than the expectation value of the Higgs doublet.

2.5 Neutrino mass measurements

The absolute scale of the neutrino mass and the mechanism that generates it remain open questions that cannot be answered by oscillation experiments. Measuring the mass will inform our understanding of the possible extension of the Standard Model of particle physics. Three different approaches being pursued to solve this puzzle have set more or less stringent upper limits on their respective observable:

1. Cosmological observations: Because of the abundance of relic neutrinos, the neutrino mass had an impact on structure formation in the early universe. Cosmological models are sensitive to the sum of the neutrino masses.
2. Neutrinoless double β -decay ($0\nu\beta\beta$): If neutrinos are Majorana particles, lepton number conservation is violated and a nuclear transition $(A, Z) \rightarrow (A, Z + 2)$ emitting two electrons but no neutrino is allowed. $0\nu\beta\beta$ -decay is sensitive to the effective mass $\langle m_{\beta\beta} \rangle$.
3. Direct neutrino mass measurements: The effective mass of flavor eigenstates could be observed in weak decay spectroscopy and time-of-flight measurements via the relativistic energy-momentum relation. This is the most model-independent approach, which in principle is based exclusively on kinematic considerations.

All three methods will be shortly summarized here. A more detailed overview can be found for example in [40].

2.5.1 Cosmology

During structure formation in the early universe, neutrinos acted as free streaming hot dark matter that smears out small-scale fluctuations. The free streaming length of relic neutrinos depends on their mass. If neutrinos became nonrelativistic at the time of recombination, they would cause a clear signal in the CMB (Cosmic Microwave Background) power spectra. If their mass is small enough ($m_\nu \ll 1 \text{ eV}$) so that they remained relativistic until after recombination, their impact on the CMB would mainly consist of a change of the angular diameter distance, which is degenerate with decreasing H_0 (Hubble constant). A review of the influence of neutrinos on cosmology can be found in [41].

The lower limit on the summed masses is given by the oscillation parameters Δm_{21}^2 and Δm_{3l}^2 . Depending on the ordering, the minimal summed mass is $\sum m_i \gtrsim 0.06 \text{ eV}$ (NO) or $\sum m_i \gtrsim 0.1 \text{ eV}$ (IO) (figure 2.4). The most recent upper limit from CMB observations was set by the Planck collaboration [28] based on CMB and BAO (baryon acoustic oscillations) data using the $\Lambda\text{CDM} + \sum m_\nu$ (Lambda Cold Dark Matter plus neutrino mass) model. These limits (95% C.L.) are $\sum m_\nu \leq 0.121 \text{ eV}$ (assuming degenerate masses), $\sum m_\nu \leq 0.146 \text{ eV}$ (assuming normal ordering) and $\sum m_\nu \leq 0.172 \text{ eV}$ (assuming inverted ordering).

The constraints on the neutrino mass from cosmology depend to some extent on the model that is used to describe structure formation. While the ΛCDM model has been very successful in describing most cosmological observations, it is only an approximation of a realistic structure formation scenario. In recent years, the model has been challenged by an emerging tension between its prediction for the Hubble constant and new independent local measurements of H_0 from luminosity distances of supernovae [42]. Depending on the extension and the combination of data sets to which the extended model is fitted, the sum of the neutrino masses could be as large as 0.5 eV [29].

Their model dependence notwithstanding, the cosmological neutrino mass limits are the most stringent yet. In combination with oscillation parameters, they indicate that laboratory experiments aiming to measure the effective mass of the electron neutrino flavor state require a sensitivity of $\ll 0.2 \text{ eV}$. A definite measurement of the neutrino mass scale by existing β -decay measurements would in turn hint at the need to extend the ΛCDM model.

2.5. NEUTRINO MASS MEASUREMENTS

2.5.2 Neutrinoless double β -decay ($0\nu\beta\beta$)

Double β -decay is an isobaric transition from a parent nucleus (A, Z) to a daughter nucleus $(A, Z+2)$ with a charge difference of $+2$. In the Standard Model this transition is only allowed under the emission of two electrons and two antineutrinos [43]

$$(A, Z) \rightarrow (A, Z + 2) + 2e^- + 2\bar{\nu}_e + Q_{\beta\beta}, \quad (2.15)$$

where $Q_{\beta\beta}$ is the energy released in the decay. The typical half-life of $2\nu\beta\beta$ -decay is $> 10^{19}$ years. The decay was first deduced in 1950 in ^{130}Te [44] and only directly observed in 1987 in ^{82}Se [45, 46]. In contrast to $2\nu\beta\beta$ -decay no neutrinos are emitted in the so-called neutrinoless double β -decay ($0\nu\beta\beta$)

$$(A, Z) \rightarrow (A, Z + 2) + 2e^- + Q_{\beta\beta}. \quad (2.16)$$

This decay violates lepton number conservation by two units and is thus forbidden in the Standard Model. Its observation would directly confirm the Majorana nature [15] of neutrinos and demonstrate the breaking of a global conservation law. This is one of the main motivations for $0\nu\beta\beta$ -decay, since the fact that there is a large abundance of matter over antimatter in the universe requires a mechanism beyond the Standard Model that resulted in this imbalance. $0\nu\beta\beta$ -decay is sensitive to the effective mass

$$\langle m_{\beta\beta} \rangle = \left| \sum_i U_{ei}^2 m_i \right|. \quad (2.17)$$

The expected decay rate depends on the neutrino mass model. The most commonly given relation, corresponding to the light neutrino exchange model, is

$$\Gamma_{1/2}^{0\nu} = \left(G |\mathcal{M}|^2 \langle m_{\beta\beta} \rangle^2 \right)^{-1} \approx 10^{27 \rightarrow 28} \left(\frac{0.01 \text{eV}}{\langle m_{\beta\beta} \rangle} \right)^2 \text{ y}, \quad (2.18)$$

where G is a phase space factor and \mathcal{M} is the nuclear transition matrix element. Via equation (2.18), experimentally measured half-lives are translated to limits on $\langle m_{\beta\beta} \rangle$. The current limits mostly range between 0.1 eV and 0.5 eV [47–51]. All limits have large uncertainties that are dominated by the uncertainty of the nuclear transition matrix element calculation. A number of experiments are continuing the quest for a possible observation of $0\nu\beta\beta$ to push the limit further down and decide the question of the neutrino nature. An overview over the recent development in the field and information about the individual $0\nu\beta\beta$ -experiments can be found for example in [52–54].

2.5.3 Direct neutrino mass measurements

As opposed to the cosmological approach and the search for $0\nu\beta\beta$ -decay, the neutrino mass can in principle be determined via the relativistic energy-momentum relation

$$E^2 = p^2 + m_0^2. \quad (2.19)$$

Methods making use of equation (2.19) are often referred to as direct neutrino mass measurements because, compared to cosmological analyses and $0\nu\beta\beta$ -decay, they are less affected by strong model dependencies. The approaches currently pursued are time-of-flight measurements and β -decay or electron capture spectroscopy.

Time-of-flight measurements

Because the neutrino mass is so small, very long-baseline time-of-flight measurements are required, which only astrophysical events such as core-collapse supernovae can provide, in order to set a competitive limit on the mass. On February 23rd 1987, the supernova SN1987 in the Large Magellanic Cloud (LMC) was observed both optically and by neutrino detection. Neutrinos originating from this event were detected by Kamioka [55], IMB (Irvine-Michigan-Brookhaven) [56] and Baksan [57]. The distance to the supernova location was (52 ± 5) kpc, which corresponds to a light travel time of $t_0 = (5.3 \pm 0.5) \cdot 10^{12}$ s. The travel time of a neutrino with mass m and energy E is

$$t_{\text{observed}} - t_{\text{emitted}} = t_0 \cdot \left(1 + \frac{m^2}{2E^2} \right). \quad (2.20)$$

Under the assumption that the emission time of the neutrinos was sharply peaked, the relation between arrival time and energy should follow a hyperbolic law. However, this was not the case for SN1987 neutrinos and consequently only upper limits could be obtained. The first limits found were in the range of 27 eV [58], but they could be tightened over the years to limits as low as 5.7 eV (95% C.L.) [59] and 5.8 eV (95% C.L.) [60] by improving the analysis and the underlying models.

The expected supernovae observation rate is 2 to 3 per century. If a galactic supernova occurs in the near future, the neutrinos will be detected by several neutrino telescopes that are interconnected by the SuperNova Early Warning System (SNEWS) [61]. But even if a core-collapse supernova could be observed soon, it would be difficult to achieve a more competitive neutrino mass limit due to a lack of understanding of the time and energy spectra of the emitted neutrinos. However, if the core-collapse process were to end in the formation of a black hole, the neutrino emission would end abruptly and more information about the detected neutrinos could be obtained [62].

2.5. NEUTRINO MASS MEASUREMENTS

Weak decay experiment

The majority of laboratory results originate from β -decay experiments. But there have also been electron capture experiments for example with ^{163}Ho that have obtained upper limits on the effective electron neutrino mass $m_\beta < 255\text{ eV}$ [63] and $m_\beta < 490\text{ eV}$ [64]. As stated before, the effective mass of the neutrino flavor state ν_e in β -spectroscopy is often defined as

$$m_\beta = \sqrt{\sum_i |U_{ei}^2| m_{\nu_i}^2}. \quad (2.21)$$

In contrast to $0\nu\beta\beta$ -decay, which is sensitive to $m_{\beta\beta}$, equation (2.21) sums over all neutrino mass states and is independent of whether the electron neutrino is a Majorana or a Dirac particle. Also m_β does not depend on any phases in the mixing matrix U . A more detailed description of the impact of neutrino masses on β -decay spectra will follow in section 2.6. Most β -decay experiments have used tritium as β -emitter. The earliest reported result from tritium spectroscopy came from the Curran experiment with an upper limit on m_β^2 of 1 keV [65]. Until September 2019, the lowest neutrino mass limits were obtained by the Mainz and Troitsk experiments [25, 26]. On September 16th, 2019, the KATRIN collaboration published its first neutrino mass limit of 1.1 eV (90% C.L.) [27]. The expected final sensitivity of KATRIN after three years of data taking time (five years of full run time) is 200 meV. Figure 2.6 shows results of previous β -decay experiments.

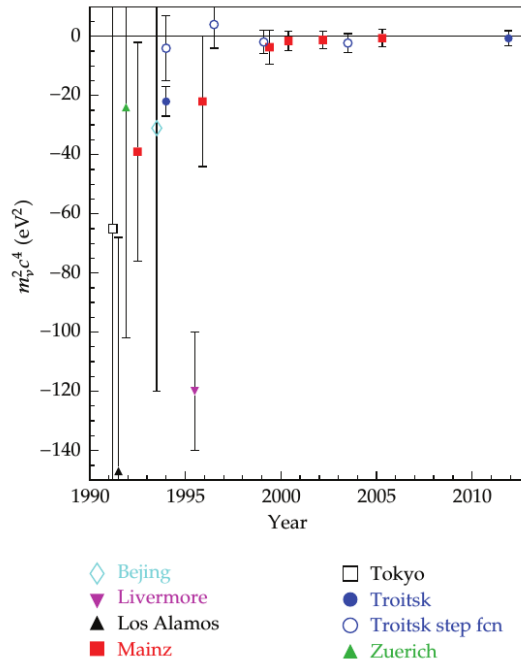


Figure 2.6: Results from β -decay experiments for the squared electron antineutrino mass since 1990 [40].

2.6 Neutrino mass in tritium β -decay

Tritium decays to helium via β -decay:



The decay is super-allowed and has a half-life of 12.3 years. The endpoint energy is 18.6 keV, which is relatively low compared to most β -decays and enhances the impact of the neutrino mass on the spectrum. In this section a theoretical description of tritium β -decay is presented following the derivations in [66, 67] and adopting the shorthand notation $m_\beta = m_{\nu_e} = m_{\bar{\nu}_e}$. In β -decay the surplus energy

$$Q = m_{\text{T}} - m_{{}^3\text{He}^+} - m_e \quad (2.23)$$

is shared between the kinetic energy of the electron, the total energy of the neutrino and the recoil energy E_{rec} of the daughter nucleus. If the neutrino was massless the difference between Q and the recoil energy E_{rec} would determine the endpoint energy of the decay spectrum: $E_0 = Q - E_{\text{rec}}$. Thus, the neutrino mass could be measured by determining the shift of E_0 by m_ν . However, Q is often subject to systematic shifts and the event rate in the last 1 eV of the spectrum is very low. Therefore, the mass extraction via a shape analysis of the spectrum near the endpoint is more feasible, treating E_0 as a fitted nuisance parameter.

According to Fermi's Golden rule, the decay probability is proportional to the summed integral of the transition matrix element M over all possible final states f

$$\Gamma = 2\pi \sum \int |M|^2 df \quad (2.24)$$

with $\hbar = c = 1$. The transition matrix element M can be divided into a leptonic and a nuclear part M_{lep} and M_{nuc} respectively. Usually it is written in terms of Fermi's coupling constant $G_F = 1.1663787(6) \cdot 10^{-5} \text{GeV}^{-2}$ [21] and the Cabibbo angle θ_C

$$M = G_F \cdot \cos \theta_C \cdot M_{\text{lep}} \cdot M_{\text{nuc}} \quad (2.25)$$

with $\cos(\theta_C) = 0.97425 \pm 0.00022$ [68]. The number of final states dn of outgoing particles within a normalization volume V into the solid angle Ω and the momentum interval $p \rightarrow p + dp$ can be calculated as

$$dn = \frac{V \cdot p^2 \cdot dp \cdot d\Omega}{(2\pi)^3} = \frac{V \cdot p \cdot E_{\text{tot}} \cdot dE_{\text{tot}} \cdot d\Omega}{(2\pi)^3}. \quad (2.26)$$

2.6. NEUTRINO MASS IN TRITIUM β -DECAY

Hence, the energy state density is

$$\frac{dn}{dE_{\text{tot}}d\Omega} = \frac{V \cdot p \cdot E_{\text{tot}}}{(2\pi)^3}. \quad (2.27)$$

Since the nucleus is much heavier than the two emitted leptons, it is a good approximation to assume that it absorbs almost no energy but balances the entire momentum. The total state density can therefore be reduced to the electron and neutrino phase space density only, which is given by

$$\rho(E_e, E_\nu, d\Omega_e, d\Omega_\nu) = \frac{dn_e}{dE_{e,\text{tot}}d\Omega_e} \cdot \frac{dn_\nu}{dE_{\nu,\text{tot}}d\Omega_\nu} = \frac{V^2 \cdot p_e \cdot E_{e,\text{tot}} \cdot p_\nu \cdot E_{\nu,\text{tot}}}{(2\pi)^6}. \quad (2.28)$$

Tritium decay is a super-allowed transition in which the nuclear transition matrix element M_{nuc} is independent of the electron energy and none of the leptons need to carry away angular momentum. However, there remains an angular correlation of the outgoing leptons. In β -decay parity is maximally violated and negative helicity particles and positive helicity antiparticles are preferred depending on their velocity. This translates to a correlation of the lepton momenta with their spins. The resulting angular correction factor is

$$1 + a \cdot (\vec{\beta} \vec{\beta}_\nu), \quad (2.29)$$

where $\vec{\beta}_\nu = \frac{\vec{v}_\nu}{c}$. The angular coefficient a is 1 for pure Fermi transitions and $-1/3$ for pure Gamov-Teller transitions. Recently, it has been measured as $a = -0.10430(84)$ for $\beta_e - \bar{\nu}_e$ transitions [69].

The squared leptonic transition matrix element $|M_{\text{lep}}|^2$ translates to the probability of the two leptons to be found at the nucleus. In allowed or super-allowed transitions, this probability is equal to $\frac{1}{V}$ for a plane neutrino wave emitted into the normalization volume V . For the emitted electron, the probability is modified by the Fermi function $F(E, Z)$, which accounts for the Coulomb interaction between the outgoing electron and the daughter nucleus. Hence, the total leptonic mixing matrix element is

$$|M_{\text{lep}}|^2 = \frac{1}{V^2} \cdot F(E, Z). \quad (2.30)$$

In the classical approximation the Fermi function is given by

$$F(E, Z) = \frac{2\pi\eta}{1 - \exp(-2\pi\eta)}, \quad (2.31)$$

where $\eta = \alpha Z/\beta$ is the Sommerfeld parameter, α is the fine structure constant and β the electron velocity relative to the speed of light. The surface of the two-particle phase space is defined by a δ -function ensuring conservation of energy. Integrating equation (2.24) over the continuum states using equation (2.28) yields the partial decay rate into a single

channel. Equation (2.32) shows the resulting atomic partial decay rate to the ground state with probability P_0 :

$$\begin{aligned} \Gamma_0 = P_0 \cdot & \int_{E_{e,\text{tot}}, E_{\nu,\text{tot}}, \Omega_e, \Omega_\nu} \frac{G_F^2 \cdot \cos^2 \theta_C \cdot |M_{\text{nucl}}^2|}{(2\pi)^5} \cdot F(E, Z) \\ & \cdot p_e \cdot E_{e,\text{tot}} \cdot p_\nu \cdot E_{\nu,\text{tot}} \\ & \cdot (1 + a \cdot (\vec{\beta}_e \vec{\beta}_\nu)) \\ & \cdot \delta(Q + m_e - E_{e,\text{tot}} - E_{\nu,\text{tot}} - E_{\text{rec}}) dE_{e,\text{tot}} dE_{\nu,\text{tot}} d\Omega_e d\Omega_\nu. \end{aligned} \quad (2.32)$$

To obtain the differential decay rate in terms of electron kinetic energy E , one has to integrate over the neutrino energy and all angles. Integration over the angles has to respect the angular correlation (2.29), which enters the recoil energy E_{rec} . The variation of E_{rec} is small and can be replaced by a constant offset $E_{\text{rec}} = 1.72$ eV. Integrating over E_ν is achieved by fixing it via the δ function to $\epsilon = E_0 - E$. In addition to integrating over the (β, ν) continuum, equation (2.32) has to be summed over all three neutrino mass eigenstates m_i with probability $|U_{ei}|^2$. The differential decay rate $\frac{d\Gamma}{dE}$ becomes

$$\begin{aligned} \frac{d\Gamma_0}{dE} = P_0 \frac{G_F^2 \cdot \cos^2 \theta_C \cdot |M_{\text{nucl}}|^2}{(2\pi)^3} \cdot F(E, Z) p_e E_{e,\text{tot}} \\ \cdot \sum_i |U_{ei}|^2 \epsilon \sqrt{\epsilon^2 - m_i^2} \cdot \Theta(\epsilon - m_i). \end{aligned} \quad (2.33)$$

The nuclear matrix element can be divided into a vector and an axial part

$$|M_{\text{nucl}}|^2 = g_V^2 + 3g_A^2 \quad (2.34)$$

with the vector coupling constant $g_V = 1$ and the axial vector coupling constant $\lambda = \frac{g_A}{g_V} = -1.2646 \pm 0.0035$ [70]. The phase space factor of the emitted neutrino

$$\epsilon \cdot (\sqrt{\epsilon^2 - m_i^2}) \quad (2.35)$$

determines the spectral shape near the endpoint. The Heaviside function Θ prevents that the electron's kinetic energy becomes negative. The theoretical spectrum and the neutrino mass impact are displayed in figure 2.7. For tritium spectroscopy with a sensitivity larger than the sum of the neutrino mass eigenstates, the differential decay rate can be simplified by using a quasi-degenerate model and considering the effective neutrino mass (2.21)

$$\begin{aligned} \frac{d\Gamma_0}{dE} = P_0 \frac{G_F^2 \cdot \cos^2 \theta_C \cdot |M_{\text{nucl}}|^2}{(2\pi)^3} \cdot F(E, Z) p_e E_{e,\text{tot}} \\ \cdot \epsilon \sqrt{\epsilon^2 - m_\beta^2} \cdot \Theta(\epsilon - m_\beta). \end{aligned} \quad (2.36)$$

2.6. NEUTRINO MASS IN TRITIUM β -DECAY

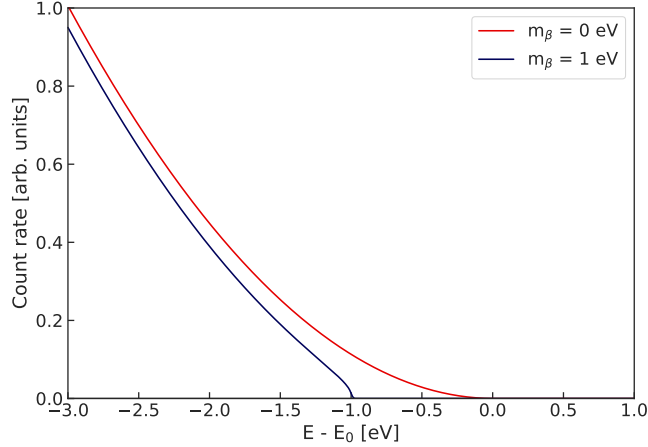
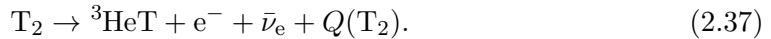


Figure 2.7: Effect of $m_\beta = 1$ eV on the shape of the tritium spectrum near the endpoint.

2.6.1 Molecular tritium decay

Molecular tritium T_2 decays via β -decay to ${}^3\text{HeT}$:



The release energy $Q(T_2)$ has to be corrected for the difference in binding energy between the mother and daughter system. In addition, the nuclear recoil can excite a spectrum of rotational and vibrational final states. The neutrino energy ϵ has to be replaced by the corrected energy $\epsilon_f = E_{0,T_2} - V_f - E$, with $E_{0,T_2} = Q(T_2) - E_{\text{rec}} = (18573.2 \pm 0.1)$ eV [71, 72]. The differential decay rate (2.36) has to be summed over all final states of the daughter molecule with probability P_f and excitation energy V_f :

$$\begin{aligned} \frac{d\Gamma}{dE} = & \frac{G_F^2 \cdot \cos^2 \theta_C \cdot |M_{\text{nucl}}^2|}{(2\pi)^3} \cdot F(E, Z) \cdot p_e \cdot (E + m_e) \\ & \cdot \sum_f P_f \epsilon_f \sqrt{\epsilon_f^2 - m_\beta^2} \cdot \Theta(\epsilon_f - m_\beta). \end{aligned} \quad (2.38)$$

According to theoretical calculations, 57% of all decays leave the daughter molecule in a rovibrational (rotational and vibrational) electronic ground state [73]. The remaining 43% decay to an excited electronic state. During the last decades there had been a discrepancy on the branching ratios to the different final states between theory and experiments conducted in the 50s [74–76]. While theory predicts a branching ratio of decaying T_2 and HT to the molecular bound states The^+ and HHe^+ respectively of at most 57%, mass spectrometer experiments found branching ratios of 95% and 90%. A new measurement of the branching ratios to the bound states by the TRIMS experiment [77] found experimental values of 56.5(6)% for the decay of HT and 50.3(15)% for the decay of T_2 , which is in agreement with the theoretical predictions. The accuracy of the

neutrino mass determination depends critically on the precision to which the binding energies and branching ratios of all states are known. Precise calculations for the state distribution in the endpoint region were performed by [78, 79]. A review of the theory of tritium final states and current validation efforts can be found in [71]. The result of the numerical calculation of final states in molecular tritium decay is shown in figure 2.8. The first group of excited electronic states show a broad peak above 25 eV. The decay to the ground state has a mean of 1.7 eV and a width of 0.4 eV. The final state spectra depend on the mass of the molecule and are consequently sensitive to the isotopes in the daughter and mother molecule. The exact composition of the source gas must therefore be known so that the relative contributions can be added to the full spectrum in a weighted sum. The width of the final states and their uncertainty are a fundamental limit in tritium spectroscopy with molecular tritium.

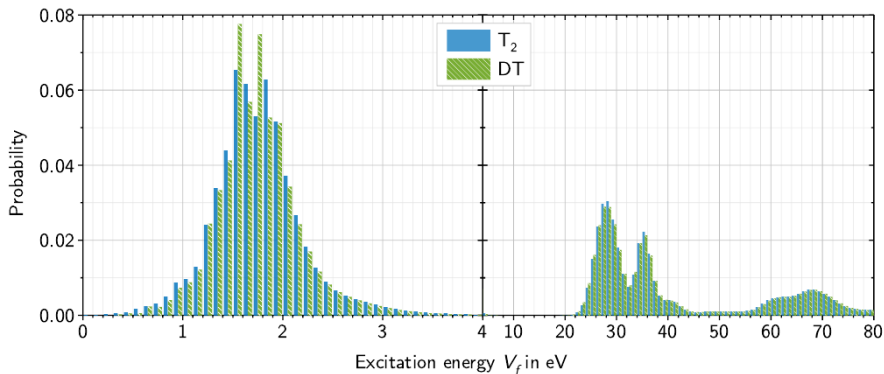


Figure 2.8: Molecular final state distribution of HeT^+ and HeD^+ at a temperature of 30 K [66].

2.6.2 Corrections

The β -spectrum in equation (2.33) is an approximation to the exact relativistic three body decay. The dependence of the recoil energy on the neutrino mass and the final state spectrum was neglected in the derivation. The approximation causes a shift of the maximum electron energy on the order of 0.1 meV, which can be compensated for by constraining the endpoint energy E_0 . Several other corrections need to be applied to the Fermi function $F(Z, E)$:

- Radiative corrections: Contributions from virtual and real photons can be accounted for by a correction factor $G(E, E_0)$. The impact on a neutrino mass determination is in the range of a few percent.

2.6. NEUTRINO MASS IN TRITIUM β -DECAY

- Screening: The Fermi function describes the Coulomb interaction between the emitted electron and the daughter molecule. Screening effects by the Coulomb field of the 1s-orbital electrons must be corrected by a screening factor $S(Z, E)$.
- Recoil effects: Energy-dependent recoil effects arising from the relativistic three-body phase space are typically combined into a single correction factor $R(E, E_0, M)$.
- Finite nucleus structure: Because the daughter molecule is not a point-like object, the Coulomb field does not scale with $\frac{1}{r^2}$. This is accounted for by a correction factor $L(Z, E)$. An additional factor $C(Z, E)$ enters from the convolution of the nuclear wave functions of the electron and the neutrino with the wave function of the nucleus.
- Recoiling Coulomb field: The emitted electron propagates in the recoiled Coulomb field. The associated correction factor is $Q(Z, E, E_0, M)$.
- Orbital interactions: The correction factor $I(Z, E)$ accounts for the quantum mechanical interactions between the emitted electron and the 1s-orbital electrons.

All correction factors are listed in the appendix of [66]. Including them yields the corrected differential decay rate

$$\left(\frac{d\Gamma}{dE}\right)_C = \frac{G_F^2 \cdot \cos^2 \theta_C}{(2\pi)^3} \cdot (g_V^2 + 3g_A^2) \cdot F_{\text{rel}}(E, Z) \cdot p_e \cdot (E + m_e) \cdot S \cdot L \cdot C \cdot I \cdot \sum_f G \cdot R \cdot Q \cdot P_f \epsilon_f \sqrt{\epsilon_f^2 - m_\beta^2} \cdot \Theta(\epsilon_f - m_\beta). \quad (2.39)$$

2.6.3 MAC-E filter spectroscopy

Since the discovery in 1939 that tritium is a β -decaying element with a small Q-value [80], it has been the isotope of choice for many neutrino mass experiments. In the recent decades, the most stringent laboratory limits for the neutrino mass were set by molecular tritium spectroscopy experiments employing the MAC-E filter (magnetic adiabatic collimation with electrostatic filter) technique. The most recent limit of $m_\beta < 1.1$ eV was set by the KATRIN experiment in 2019 [27].

In KATRIN, T_2 source gas molecules decay in a windowless gaseous tritium source. Strong magnetic solenoidal fields guide the emitted electrons to an electrostatic pre- and a main spectrometer [81, 82]. To prevent the spectrometers from being tritiated, the transport section reduces the gas pressure by 14 orders of magnitude via differential and cryogenic pumping. The pre-spectrometer serves as pre-filter and rejects all electrons with a kinetic energy smaller than about 8 keV below the endpoint energy. In the main spectrometer the remaining electrons are then further filtered to a narrow energy range

that is scalable by adjusting the electrostatic potential. The magnetic field strength changes in the transport section and in the spectrometers are slow compared to the cyclotron frequency of the guided electrons ensuring adiabatic invariance of the magnetic moment related to their spatial angular momenta. Along their path to the central plane of the spectrometers, the magnetic field drops by several orders of magnitude. Due to the conservation of the electrons' magnetic moments in the slowly varying fields, the momentum vector aligns with the field lines. This way, the electron trajectories are all parallel when they encounter the retarding potential, which therefore acts as high-pass energy filter. Electrons that surpass the potential are re-accelerated and counted by a segmented monolithic silicon detector. By step-wise scanning of the electric potential across an energy range from 40 eV below to 50 eV above the spectrum endpoint, an integrated β -spectrum is recorded. The energy resolution that can be achieved by the MAC-E filtering technique is limited by the ratio of minimum (B_A) to maximum (B_{\max}) magnetic field strength

$$\Delta E = \frac{B_A}{B_{\max}} \cdot E \cdot \frac{\gamma + 1}{2}, \quad (2.40)$$

where B_A is the field in the analyzing plane of the spectrometer, γ is the Lorentz factor, and E is the kinetic energy of the electrons. A schematic drawing of the KATRIN setup is depicted in figure 2.9.

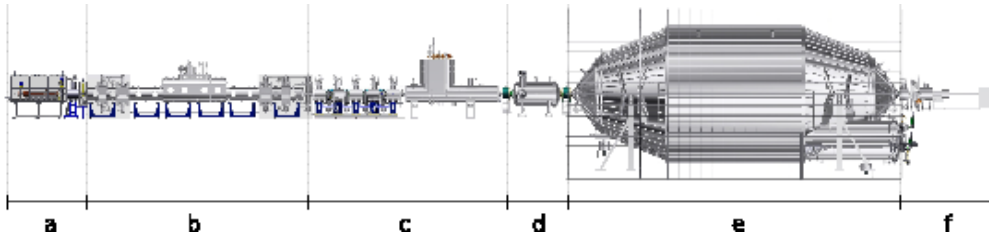


Figure 2.9: The 70 m long KATRIN experiment consists of a monitoring and calibration section (a), a windowless source (b), a magnetic transport section (c), a pre-spectrometer (d), a main spectrometer (e), and an electron counting detector (f). Image from [66].

By using the MAC-E filter technique, the Mainz and Troitsk experiments were able to set an upper limit on the effective electron antineutrino mass of $m_\beta \leq 2 \text{ eV}$ (95 % C.L.) [25, 26]. The KATRIN experiment is the successor of Mainz and Troitsk and aims for an unprecedented sensitivity of 0.2 eV after three years of live measurement time.

With KATRIN, the MAC-E filter technique reaches its ultimate limitations in terms of size and sensitivity. The technique does not allow the use of atomic tritium, which is crucial for pushing the sensitivity of a tritium experiment below 0.1 eV unless the uncertainty on the final states of molecular tritium can be further reduced. The Project 8 experiment is the first neutrino mass experiment that is going to attempt tritium spectroscopy with atomic tritium, using the CRES technique for recording a differential tritium spectrum.

Cyclotron Radiation Emission Spectroscopy was proposed in 2009 by Monreal and Formaggio [83] as a novel precision spectroscopy technique with the goal to measure the neutrino mass from tritium decay with a sub-eV sensitivity. The Project 8 collaboration was formed to pursue the development of the CRES technique, aiming to reach a neutrino mass sensitivity of $m_\beta \lesssim 40$ meV at 90% confidence. This requires the use of atomic tritium as β -emitting source, a very homogeneous magnetic field $\frac{\sigma(B)}{B} \leq 10^{-7}$, and a precise determination of the kinetic energy of emitted electrons $\sigma_e \sim 0.1$ eV.

3.1 Cyclotron Radiation Emission Spectroscopy

Since 2009, the CRES technique has been developed and demonstrated by the Project 8 collaboration. A vessel containing a source gas consisting of unstable β -emitters is placed in a uniform magnetic field. In this field, electrons will follow a cyclotron trajectory with a frequency inversely proportional to the Lorentz factor $\gamma = \left(1 + \frac{E_{\text{kin}}}{m_e c^2}\right)$

$$f_\gamma = \frac{f_e}{\gamma} = \frac{1}{2\pi} \cdot \frac{qB}{m_e + E_{\text{kin}}/c^2}. \quad (3.1)$$

The Lorentz factor of electrons with a kinetic energy of 18.6 keV (energy of the tritium spectrum endpoint) is $\gamma \approx 1.035$. As long as the electron motion along the magnetic field lines is nonrelativistic, the electrons emit coherent electromagnetic radiation with a power spectrum sharply peaked at f_γ . This radiation can be detected by a receiving antenna and allows for a reconstruction of the kinetic energy via the determination of the cyclotron frequency. The CRES technique offers several major benefits:

- First of all, it is relatively easy to achieve the required resolution in frequency measurements as long as the observation time is long enough. In a 1 T magnetic field, the cyclotron frequency of an 18.6 keV electron is 27.01 GHz. To resolve 0.1 eV at 18.6 keV, the uncertainty in determining the start frequency must be ≤ 5 kHz. The frequency resolution in signal processing is limited by the observation time

$$\sigma_f \cdot \sigma_t \geq \frac{1}{4\pi}. \quad (3.2)$$

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

Therefore, the decay electrons must be observed over a period of at least 0.016 ms. If the pitch angle θ , which is defined as the angle between the momentum of the electron and the magnetic field vector, is 89° , the electron's guiding center moves parallel to the field lines at $\sim 0.5\%$ of the speed of light. Within 0.016 ms it travels a distance of 21 m. Therefore, electrons emitted at a pitch angle $\theta < 90^\circ$ must be trapped magnetically to achieve the required observation time without changing their kinetic energy. This is accomplished by adding a magnetic bottle field to the homogeneous background field. The minimum pitch angle that remains trapped is determined by the trap depth ΔB

$$\theta \geq \sin^{-1} \left(\sqrt{1 - \frac{\Delta B}{B_{\max}}} \right). \quad (3.3)$$

- The second major advantage of the CRES technique compared to conventional particle detection methods is that the emitted electrons do not have to be extracted from the source gas to be detected. In principle, they can be observed directly after the individual decays. This allows the use of atomic rather than molecular tritium as a source gas. As a result, the minimum achievable sensitivity can be significantly improved, since the fundamental limitation of molecular tritium spectroscopy is removed: The uncertainty of the final state spectrum.

3.1.1 Basic signal characteristics

A CRES signal originating from a trapped electron has several features that can be used to distinguish it from RF noise. Figure 3.1 shows the first detected CRES event in a power spectrogram that was published in [84]. Narrow-band power with a frequency corresponding to the electron's cyclotron frequency abruptly emerges from the moment the electron is born in the magnetic trap. It persists in time until the electron escapes the trap, for example after changing its pitch angle θ due to a collision with a gas molecule. Not every collision leads to a large enough change in pitch angle to enable the electron to escape, but each collision will result in an energy loss and therefore in a sudden upward change in frequency. Between collisions, the electron steadily emits radiation and thus loses power, causing a slow upward drift in frequency with a slope of typically 350 MHz/s (in a 1 T field). Each event corresponding to one trapped electron therefore consists of one or more linear chirps which are named electron tracks. The sum of the track time durations (track lengths) amounts to the total event length during which the electron remained trapped. Trapping field geometries, Doppler and detection effects add additional features to an event, which are often correlated with the electron's pitch angle. These effects will be described in more detail in chapter 4.

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

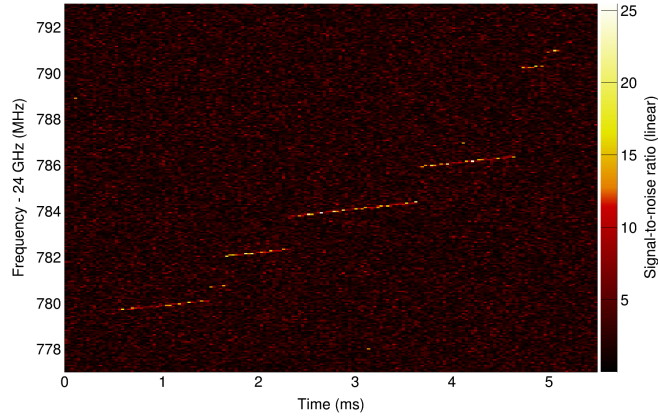


Figure 3.1: Power spectrogram of the first observed CRES event [84].

3.1.2 Sensitivity

To get a quick sense of the sensitivity to the neutrino mass of a CRES experiment with tritium, it is not necessary to perform a full spectrum analysis. A simplified analysis [85] already provides a result that is consistent with the Bayesian sensitivity studies performed by Talia Weiss following her analysis approach in [86]. The analytical estimation is based on a simplified description of the tritium spectrum (3.4), assuming that m_β is large and therefore the individual neutrino mass states need not be treated individually

$$\frac{dN_s}{dE} = 3rt \cdot (E_0 - E) \cdot \sqrt{(E_0 - E)^2 - m_\beta^2}. \quad (3.4)$$

Here, r is the event rate in the last 1 eV of the spectrum in the absence of neutrino mass, t is the total data taking duration and E_0 is the endpoint energy. Naively, the neutrino mass can be determined from counting the number of events in an interval $\Delta E = E_0 - E_1$, with $E_1 < E_0$, as long as the rate, time, endpoint energy, and background are known well enough. Additional information, such as the distribution of events in ΔE , could be used, but this simplification provides a robust statistical baseline for estimating the sensitivity of a differential spectrum measurement. The total number of signal events is obtained by integrating and expanding equation (3.4)

$$N_s = rt \cdot (\Delta E)^3 \left[1 - \frac{3}{2} \frac{m_\beta^2}{(\Delta E)^2} \right]. \quad (3.5)$$

For a flat background rate b , the total number of events is the sum of the number of signal events N_s and the number of background events $N_b = bt \cdot \Delta E$

$$N_{\text{tot}} = rt \cdot (\Delta E)^3 \left[1 - \frac{3}{2} \frac{m_\beta^2}{(\Delta E)^2} \right] + bt \cdot \Delta E. \quad (3.6)$$

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

The statistical neutrino mass uncertainty $\sigma_{m_\beta^2}$ is related to the variance of the total number of events

$$\sigma_N^2 = \left(\frac{\partial N_{tot}}{\partial m_\beta^2} \right)^2 \sigma_{m_\beta^2}^2 = \left(\sqrt{N_{tot}} \right)^2. \quad (3.7)$$

Differentiating equation (3.6) yields

$$\frac{\partial N_{tot}}{\partial m_\beta^2} = -\frac{3rt \cdot \Delta E}{2}, \quad (3.8)$$

which can be solved for $\sigma_{m_\beta^2}$

$$\begin{aligned} \sigma_{m_\beta^2} &= \frac{2}{3rt \cdot \Delta E} \sqrt{N_{tot}} \\ &= \frac{2}{3rt} \cdot \sqrt{rt \cdot \Delta E - \frac{3}{2} \frac{m_\beta^2}{\Delta E^4} + \frac{bt}{\Delta E}} \\ &\approx \frac{2}{3rt} \sqrt{rt \cdot \Delta E + \frac{bt}{\Delta E}}. \end{aligned} \quad (3.9)$$

The uncertainty has a minimum if ΔE is chosen to be $\Delta E = \sqrt{\frac{b}{r}}$. Assuming an experiment where r is high and b is low, the optimum analysis window is small and probably limited by experimental constraints. The energy window should not be smaller than the largest contribution to the energy resolution from either the width of the final state distribution or the instrumental energy resolution, since the contribution of the neutrino mass will be smeared over this range. It is conceivable that the smallest reasonable analysis window ΔE is the full width half maximum (FWHM) of the total energy smearing after adding the individual components in quadrature

$$\Delta E = \sqrt{8 \ln 2} \cdot \sqrt{\frac{b}{r \cdot 8 \ln 2} + \sigma_{\text{FSS}}^2 + \sigma_{\text{trans}}^2 + \sigma_{\text{instr}}^2}. \quad (3.10)$$

Here, σ_{FSS} is the final state width, σ_{trans} is the energy smearing from Doppler broadening and σ_{instr} is the instrumental energy uncertainty, which in turn is given by the frequency precision and the magnetic field inhomogeneity ($\sigma_{\text{instr}}^2 = \sigma_\omega^2 + \sigma_B^2$). The derivative of the cyclotron orbital frequency ω_c with respect to the kinetic energy E_{kin} is

$$\begin{aligned} \frac{d\omega_c}{dE_{\text{kin}}} &= -\frac{qB}{(m_0c^2 + E_{\text{kin}})^2} \\ \frac{dE_{\text{kin}}}{E_{\text{kin}}} &= -\frac{m_0c^2 + E_{\text{kin}}}{E_{\text{kin}}} \frac{d\omega_c}{\omega_c} \\ &= -\frac{\gamma}{\gamma - 1} \frac{d\omega_c}{\omega_c}. \end{aligned} \quad (3.11)$$

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

Depending on the gas pressure, an electron will scatter after a random amount of time. The average time before scattering is the mean free path of the electron $\lambda = 1/(\sigma_0 n)$ divided by the velocity $\tau = \lambda/\beta c$, where n is the number density of molecules or atoms and σ_0 is the scattering cross section. The FWHM of the angular frequency $\Delta\omega_c$ is related to τ via

$$\tau = \frac{1}{\Delta\omega_c}, \quad (3.12)$$

yielding

$$\frac{\Delta\omega_c}{\omega_c} = \frac{\beta c \sigma_0 n}{\omega_c}. \quad (3.13)$$

For convenience, it is assumed here that the contribution to the energy uncertainty due to scattering is bell-shaped. The magnetic field inhomogeneity seen by an orbiting electron can be parameterized by the rms field variation $\sigma(B)$. Collecting all terms yields

$$\Delta E = \sqrt{8 \ln 2} \cdot \sqrt{\frac{b}{r \cdot 8 \ln 2} + \sigma_{\text{FSS}}^2 + \sigma_{\text{trans}}^2 + \left(\frac{\gamma E_{\text{kin}}}{\gamma - 1} \cdot \frac{\sigma(B)}{B}\right)^2 + \left(\frac{\gamma E_{\text{kin}}}{\gamma - 1} \cdot \frac{\beta c \sigma_0 n}{\omega_c}\right)^2}. \quad (3.14)$$

The event rate r in a physical volume V is related to the number density n , the source gas lifetime τ_s , the branching ratio η to the last eV, and the probability that a decay is detected ϵ

$$r = \epsilon \eta \frac{nV}{\tau_s}. \quad (3.15)$$

Each energy uncertainty term also contributes to $\sigma_{m_\beta^2}$

$$\sigma_{m_\beta^2} = \sqrt{\sum_i \left(\frac{\partial m_\nu^2}{\partial \sigma_i} \sigma_{\text{res}(i)}\right)^2} \quad (3.16)$$

with σ_i being the Gaussian standard deviation of the resolution contribution of the component i . According to [87] the relationship between the error σ_{res} from an instrumental resolution contribution and the neutrino mass error it introduces is

$$\frac{\partial m_\beta^2}{\partial \sigma_i} = -2 \frac{\partial \sigma_i^2}{\partial \sigma_i}. \quad (3.17)$$

Equation (3.16) then becomes

$$\sigma_{m_\beta^2} = 4 \sqrt{\sum_i \sigma_i^4 \left(\frac{\sigma_{\text{res}(i)}}{\sigma_i}\right)^2}. \quad (3.18)$$

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

Combining equation (3.9) with (3.18) gives the final uncertainty on the squared neutrino mass

$$\sigma_{m_\beta^2} = \sqrt{\sigma_N^2 + \sigma_{\text{FSS}}^4 \cdot f_{\text{FSS}}^2 + \sigma_{\text{trans}}^4 \cdot f_{\text{trans}}^2 + \left(\frac{\gamma E_{\text{kin}}}{\gamma - 1} \cdot \frac{\sigma(B)}{B}\right)^4 \cdot f_B^2 + \left(\frac{\gamma E_{\text{kin}}}{\gamma - 1} \cdot \frac{\beta c \sigma_0 n}{\omega_c}\right)^4 \cdot f_{\omega_c}^2} \quad (3.19)$$

where f_i is the resolution factor $\left(\frac{\sigma_{\text{res}(i)}}{\sigma_i}\right)$ that describes the knowledge of the different energy widths. For simplicity, it is assumed to be 1 % for all contributing terms. Using the values from table 3.1, the result of $\sigma_{m_\beta^2}$ can be calculated as a function of the total number of events for molecular or atomic tritium. For an expected neutrino mass of 0 eV the sensitivity is taken to be the width of the one-sided 90 % confidence interval, calculated as the square root of $1.28 \cdot \sigma_{m_\beta^2}$. The result is shown in figure 3.2.

| Parameter | Molecular tritium | Atomic tritium |
|----------------------------|--|--------------------------------------|
| Inelastic cross section | $3.4 \cdot 10^{-18} \text{ cm}^2$ [88] | $9 \cdot 10^{-19} \text{ cm}^2$ [89] |
| Branching ratio | 0.57 [90] | 0.702 [91] |
| Final state spectrum width | 0.436 eV (30 K) [90] | 0.024 eV (1 K) [90] |

Table 3.1: Input parameters for the analytical sensitivity calculation of a CRES tritium endpoint measurement using molecular or atomic tritium as source gas.

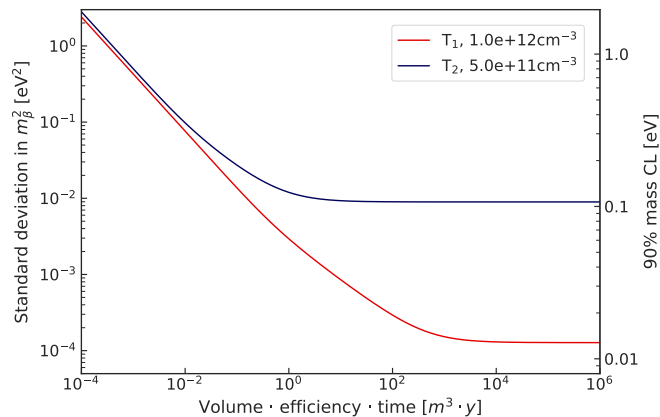


Figure 3.2: $\sigma_{m_\beta^2}$ and the corresponding 90 % neutrino mass interval assuming a best-fit mass of $m_\beta = 0 \text{ eV}$ for a molecular (blue) and an atomic (red) tritium CRES experiment. All resolution factors are set to 1 %. Adapted image from private communication with Lucie Tvrznikova.

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

From figure 3.2 it can be seen that the ultimate neutrino mass limit in a molecular tritium experiment converges to about 0.1 eV due to the final state spectrum width. In an atomic tritium experiment, the achievable sensitivity is in principle about 10 meV. The limiting contribution here is the uncertainty of the magnetic field and its resolution factor. For example, if f_B is 10 % instead of 1 %, the mass limit increases to 40 meV, while the sensitivity in a molecular tritium experiment remains mostly unaffected (figure 3.3).

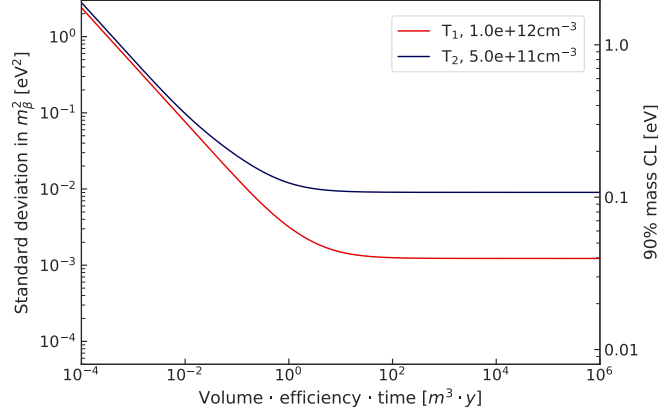


Figure 3.3: $\sigma_{m_\beta^2}$ and the corresponding 90 % neutrino mass interval as in figure 3.2, this time with a magnetic field resolution factor of 10 %.

The total number of recorded events is given by the product of effective volume, gas density, event detection efficiency, and net data acquisition live time. The curves in figures 3.2 and 3.3 are all based on a gas number density of $1 \cdot 10^{12} \frac{\text{atoms}}{\text{cm}^3}$. Convergence to the sensitivity limit given by the experimental systematic energy uncertainties in an atomic tritium experiment, as shown in figure 3.3, requires $V \cdot \epsilon \cdot t > 10 \text{ m}^3 \text{ y}$. Assuming an efficiency of 10 % and a data recording time of one year leads to the conclusion that a CRES neutrino mass experiment needs to contain a source gas volume of about 10^2 m^3 to 10^3 m^3 , unless the efficiency or the gas density can be increased. The goal of the Project 8 collaboration is to achieve a final sensitivity of 40 meV, which is smaller than the minimum effective neutrino mass for inverted ordering and therefore either gives a positive result or eliminates inverted ordering as an option. A comparison of the target sensitivities of Project 8 and KARTIN is shown in figure 3.4.

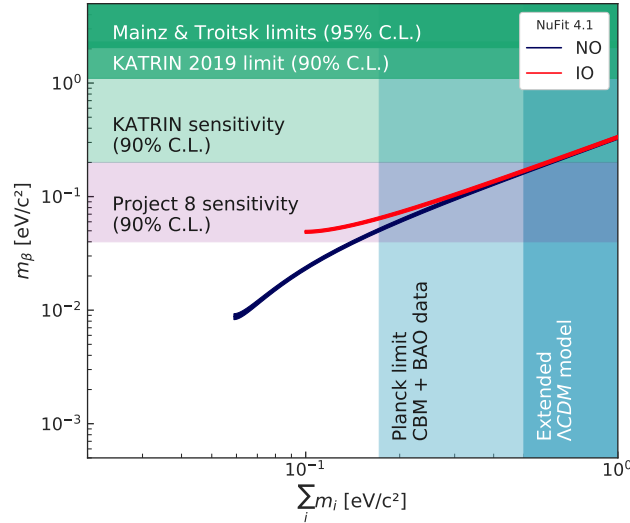


Figure 3.4: Effective mass of the electron antineutrino as a function of the summed mass eigenstates as shown in figure 2.4. The target sensitivity of the Project 8 experiment (purple) would exclude the entire mass range allowed by inverted ordering.

3.1.3 Contributions to the neutrino mass sensitivity of CRES

According to the derivation in section 3.1.2, the contributions to the neutrino mass sensitivity can be decomposed into three components:

1. Energy resolution σ_e consisting of the sum of several resolution components σ_i (for example the achievable frequency resolution σ_f)

$$\sigma_e = \sum_i \sigma_i. \quad (3.20)$$

2. Exposure N_{detected} defined as the number of detected events

$$N_{\text{detected}} = V \cdot n \cdot \frac{t}{2\tau_{1/2}} \cdot \epsilon, \quad (3.21)$$

where V is the total gas volume, n is the molecule number density, $\tau_{1/2}$ is the half-life of the source gas isotope, t is the measuring duration and ϵ is the detection efficiency. The detection efficiency is defined as the proportion of detected decays over all decays in the total gas volume.

3. Background rate b , which can be divided into two categories: Noise fluctuations and trapped electrons originating from sources other than the design gas e.g. cosmic rays.

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

These three contributions to the sensitivity depend on various technical parameters specific to the experiment. In order to achieve the sensitivity goal of 40 meV, both the value of each contribution and its uncertainty must be known.

Contributions to the energy resolution

The lower limit of the energy resolution in a CRES experiment results from the sum of the uncertainty on the start frequency σ_f and the magnetic field inhomogeneity σ_B . When measuring decays of tritium atoms, the Doppler effect leads to an additional uncertainty of the energy resolution σ_{trans} due to the motion of the parent atom.

- σ_f : For a signal with constant frequency, simply averaging the frequency spectrum over a long time would theoretically result in an arbitrarily precise frequency (and therefore energy) determination. However, electron tracks only last for a short duration (the time between collisions with gas molecules) and their frequency is not constant due to the radiative energy loss. The frequency precision from a discrete Fourier transform of the signal in figure 3.1 is limited by the duration Δt , during which the electron radiates at approximately the same frequency $\Delta f < \frac{1}{\Delta t}$. Yet, the start frequency determination is not limited to the width of such a frequency bin. It can be improved by taking advantage of the fact that the slope of the tracks is smaller than $\frac{\Delta f}{\Delta t}$. If the start frequency is calculated from the slope and the start time, the limitation to Δf can be overcome and is replaced by the limitations of the start time resolution and the slope uncertainty, which decrease with track length.
- σ_B : The cyclotron frequency is proportional to B . Any variation of the magnetic field strength along the electron's trajectory will lead to a broadening or a modulation of the received frequency. As long as this variation is known, the energy can still be precisely reconstructed. Therefore, accurate field mapping and field monitoring over time is required. Even with a field magnitude uncertainty $\leq 10^{-6}$ the magnetic field homogeneity is expected to be the main limitation for the energy resolution in an atomic CRES experiment.
- σ_{trans} : For a thermalized source gas the Doppler broadening of the spectrum is given by

$$\sigma_{\text{trans}} = \sqrt{\frac{p_{\text{rec}}^2}{2m_T} 2k_B T}, \quad (3.22)$$

where p_{rec} is the recoil momentum given by the sum of the momentum vectors of the emitted electron and neutrino, m_T is the mass of a tritium atom, k_B is the Boltzmann constant and T is the gas temperature. To keep the contribution of the

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

Doppler broadening small, the source gas must be cold. For an atomic tritium gas at 30 mK the expected Doppler broadening is ≤ 5 meV.

Background sources

Background events originating from the misidentification of noise fluctuations are an event reconstruction challenge and will be addressed in section 7.5. The following effects can lead to unwanted trapped electrons, that would appear as additional background in a recorded spectrum:

- Cosmic rays and radioactivity interactions with the source gas.
- β -decays of atoms adsorbed on the wall of the vessel containing the source gas: Electrons born outside of the magnetic trap can only be trapped if they lose energy or change their pitch angle, for example in scattering interactions. However, most of the electrons emitted from the vessel wall will be magnetically reflected back into the wall on the time scale of the inverse cyclotron frequency (40 ps).
- Electrons emitted in T_2 decay: If the source gas consisted of atomic tritium, decays from recombined molecules could be a significant background for the atomic tritium spectrum. The molecular tritium endpoint is approximately 8 eV above the atomic endpoint. This makes the purity of the gas composition a major design constraint (see section 3.2.2).

Only trapped charged particles that emit radiation in the frequency region of interest are relevant background for the neutrino mass determination with CRES. Electrons produced by cosmic ray interactions and decays occurring on the vessel walls may end up in the magnetic trap without necessarily contributing to the background.

Impact of experimental parameters on the sensitivity

The main effects that impact the sensitivity of the neutrino mass determination with CRES are:

- **Gas pressure:** For any given isotope the gas density determines the mean free path of electrons in the gas and therefore the average time between electron-gas collisions. The time between scatters is equal to the electron track length. Event detection techniques have to make use of the event characteristics to distinguish them from RF noise. The main features are the narrow-band power excess and its persistence in time. Therefore the detection probability is higher for longer tracks. In addition, the start frequency accuracy depends on the accuracy of the slope determination, which is also better for long tracks. There exist two kinds of optimum pressure values:

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

1. To maximize the event rate, a trade-off between detection probability and the number of decays per time in a unit volume can be found. The number of decays increases linearly with pressure. At the same time, longer tracks are easier to detect. The product of the efficiency and the event rate dependence on pressure will possess a maximum. This is however not necessarily the same pressure that maximizes the sensitivity of the experiment to the neutrino mass.
 2. Instead of maximizing the count rate, the pressure can be chosen such that it minimizes the achievable upper limit on the neutrino mass. This is accomplished by finding the optimum combination of exposure and energy resolution. In this process, the optimum pressure depends on many experimental parameters like the fraction of the emitted power that is collected and the performance of the event detection techniques.
- **Emitted and collected power:** According to Larmor's equation (3.23) the free space radiative power loss of a single electron in a magnetic field is

$$P(\gamma, \theta) = \frac{1}{4\pi\epsilon_0} \frac{2}{3} \frac{e^4}{m_e^2 c} B^2 (\gamma^2 - 1) \sin^2 \theta, \quad (3.23)$$

where ϵ_0 is the permittivity of free space and θ is the pitch angle of the electron. In a 1 T field, the Larmor power of 18 keV electrons amounts to approximately 1 fW. Due to the low radiative power, the signal-to-noise ratio (SNR) in CRES is relatively low, therefore requiring a high-performance receiver and signal processing setup with a very low noise temperature. The fraction of the emitted power that can be collected and the noise temperature that is added to the signal determine the detection efficiency of the system. If the detection efficiency is low, it has to be compensated for by a larger source gas volume, longer measurement time, or higher gas densities. Therefore, the fraction of collected power impacts the exposure and the energy resolution.

- **Frequency- or energy-dependent SNR:** It can easily be seen from equation (3.23) that the radiated power depends on γ . As γ is directly related to the cyclotron frequency, the emitted power is frequency-dependent. In addition, noise temperatures and noise added by the signal processing system also vary with frequency. All this induces a frequency and energy dependence of the signal-to-noise ratio and thus the event detection probability. An uncertainty of this dependence limits the precision to which the recorded tritium spectral shape can be known. As the neutrino mass is extracted from the spectral shape near the endpoint, an uncertainty on the spectral shape directly translates to a limitation of the sensitivity and has to be well understood. It is one of the goals of this work to study this

3.1. CYCLOTRON RADIATION EMISSION SPECTROSCOPY

systematic effect in detail and to make a prediction for how well the efficiency must be known in order to perform a competitive neutrino mass measurement with CRES. Figure 3.5 shows an example for the impact an energy-dependent detection efficiency could have on the tritium spectrum.

- **Missed first tracks:** The length (duration) of the electron tracks is random and exponentially distributed. It is thus unavoidable that a significant fraction of tracks is too short to be distinguishable from noise fluctuations. Events consisting of just a single short track will thus remain undetected. However, if an event has several tracks and only the first one is too short or faint to be detected, the start-frequency associated with the event will be the start frequency of the second (or n^{th}) track and consequently incorrect. As the frequency jumps between tracks are usually positive¹, the mis-reconstructed start frequencies are almost always too high. The minimum energy loss for an electron scattering from a hydrogen isotope is about 12 eV [88]. This sets a minimum for the size of the upward frequency jumps between tracks (ignoring pitch angle changes). If the analysis energy range for extracting the neutrino mass is e.g. limited to the last 10 eV, the mis-reconstructed start frequencies do not interfere with the neutrino mass analysis. However, depending on the magnetic trap and the antenna configuration, additional frequency shifts from pitch angle effects may lead to scattered tracks ending up in the same frequency region as un-scattered tracks. For a neutrino mass experiment using CRES, these effects have to be well understood during the design of the experiment. In the Project 8 prototype experiment that this work focuses on, the analyzed energy range extends down to almost 2.5 keV below the tritium endpoint, and events with missed first tracks are accounted for in the analysis.

¹A consecutive track has a higher start frequency than its predecessor. Only in rare cases can the change of pitch angle lead to a frequency decrease.

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

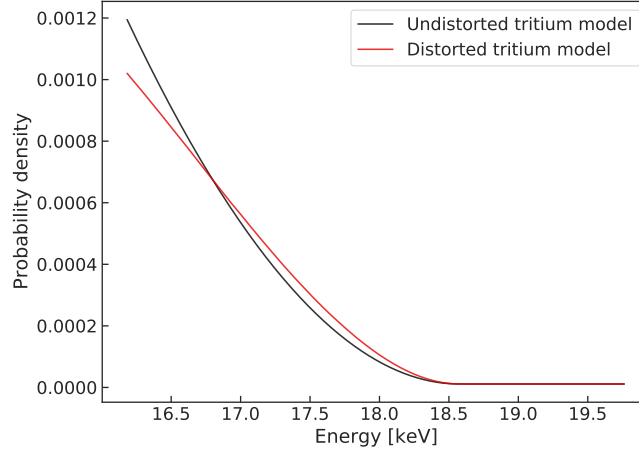


Figure 3.5: The result of a detection probability varying with energy is a distortion of the measured spectrum. The detection efficiency assumed for this plot was generated in a toy model simulation (see section 9.1.1).

3.2 A phased approach to measuring the neutrino mass

The Project 8 collaboration follows a phased approach to achieve its goal of measuring the electron antineutrino mass from tritium endpoint spectroscopy. An overview of the phases and R&D plan is given in [92]. The phases and their respective goals and milestones are listed in table 3.2.

| Phase | Timeline | Milestones | Science goals |
|-------|--------------------|---|---|
| I | 2010 - 2016 | Proof of concept: Single electron detection | Conversion electron spectrum from ^{83m}Kr |
| II | 2015 - 2020 | Systematic studies and first recorded CRES T_2 spectrum | Final state spectrum test: determination of tritium endpoint with $\sigma_{E_0} \sim 30 \text{ eV}$ |
| III | 2016 - 2026 | Demonstration of high exposure and atomic tritium feasibility | $m_\beta \lesssim 2 \text{ eV}$ |
| IV | 2022 - ≥ 2030 | Large scale atomic tritium experiment | $m_\beta \lesssim 40 \text{ meV}$ |

Table 3.2: Project 8 phases and milestones

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

The experiment that is supposed to measure the effective electron antineutrino mass, or at least set an upper limit of 40 meV, is Phase IV. Phase I and II are prototype experiments demonstrating the CRES technique for discrete and continuous spectra. In Phase II, the first CRES tritium spectrum was recorded between December 2019 and March 2020. Phase III summarizes the demonstrations of the technologies needed in order to build and conduct the large scale experiment in Phase IV.

3.2.1 CRES in a waveguide: Phase I and II

In Phase I and II a waveguide is placed inside a 1 T superconducting NMR magnet. The goal of Phase I was to show that the radiation of single electrons in a magnetic field can be detected and used to reconstruct the discrete $^{83\text{m}}\text{Kr}$ energy spectrum. This goal was reached [84] and Phase I was completed in 2016. A schematic diagram of the Phase I setup is shown in figure 3.6. The setup consisted of a standard rectangular

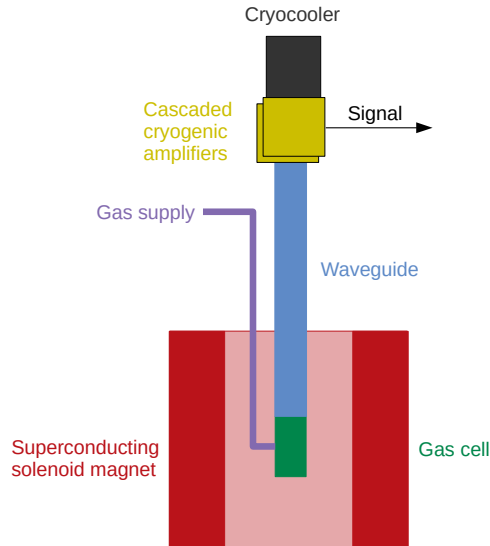


Figure 3.6: Schematic setup of the CRES experiments Phase I and II. A waveguide containing a gas cell is inserted into the bore of a superconducting solenoid magnet. The cell is filled with an electron-emitting source gas. Electrons in the cell follow a cyclotron trajectory and emit radiation that is transmitted via the waveguide to a receiver.

WR42 waveguide cell sealed by two Kapton windows. Gaseous $^{83\text{m}}\text{Kr}$ was continuously produced by a ^{83}Rb source adsorbed on zeolite beads. The cell was kept at low pressure while the krypton gas diffused through the gas system and the cell. $^{83\text{m}}\text{Kr}$ is an excited state of krypton and decays via internal conversion with a half-life of 1.83 hours. The $^{83\text{m}}\text{Kr}$ spectrum has a K-line near 17.8 keV, two L-lines near 30.4 keV, and two M-lines near 31.9 keV. The K-line is close to the tritium decay spectrum endpoint and the L-

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

and M-lines allow a calibration over a large energy range. This makes $^{83\text{m}}\text{Kr}$ an ideal isotope for the energy calibration of an experimental setup targeting the determination of the neutrino mass via tritium decay.

In Phase I, the 1 T magnetic field was provided by a warm-bore superconducting solenoid with an aperture of 52 mm. The bore diameter limited the size of the detector insert and therefore the gas cell volume to only a few cubic centimeters. The waveguide insert was placed coaxially inside the bore and the magnetic field caused the decay electrons to gyrate around the field lines. In this configuration, the electrons coupled strongly to the fundamental TE_{10} mode of the rectangular waveguide and a large fraction of the emitted power was transmitted to the input of a low noise radio-frequency receiver. The signal was mixed-down² at a high-frequency receiver stage to center the 2 GHz wide frequency region of interest at 1.8 GHz. The Phase I insert and receiver chain are shown in figure 3.7. The longitudinal confinement (along the solenoid's z -axis) was achieved by

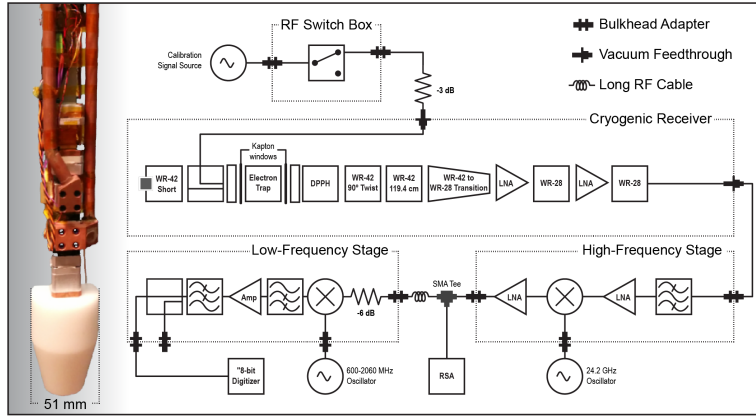


Figure 3.7: Phase I CRES prototype detector insert and receiver electronics [84]. The receiver chain consists of cascaded cryogenic amplifiers and a high- and low-frequency stage. The frequency band of interest (26 ± 1) GHz is mixed down to be centered at 1.8 GHz and is then subject to a second amplification and mixing stage.

adding a weak magnetic bottle field. This allowed for a sufficiently long observation time to detect the cyclotron radiation. The trapping field was provided by a combination of three copper coils that were wound around the gas cell. Each of them could generate a field perturbation with a depth of up to 8 mT. Depending on which coils were powered, different trapping field geometries could be generated to trap electrons with pitch angles $85^\circ \leq \theta \leq 90^\circ$. Figure 3.1 shows the spectrogram of a typical CRES event recorded with the Phase I detector. To extract the decay energy of such an event the first track has to be identified, as all other tracks are shifted in frequency due to the energy loss from scattering. The starting points of the first tracks are used to compose the decay

²In RF mixing, a signal is multiplied with a local oscillator signal. The product is a signal at the difference and the sum of the mixed frequencies.

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

frequency spectrum. Using the frequency to energy relation from equation (3.1), the start frequency spectrum is translated to energy. Figure 3.8 shows $^{83\text{m}}\text{Kr}$ spectra that were recorded in Phase I. Depending on the depth $\Delta B = B_{\text{max}} - B_{\text{min}}$ and the trap geometry, a FWHM of the two peaks in the 30.4 keV doublet of 3.3 eV (figure 3.8a) to 15 eV was achieved.

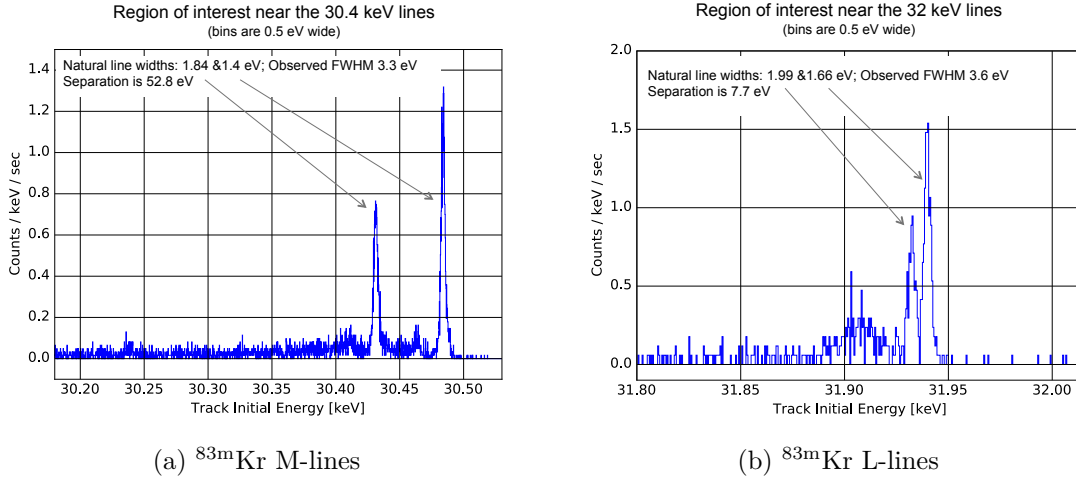


Figure 3.8: Energy spectra of $^{83\text{m}}\text{Kr}$ lines recorded with the Phase I experiment [92].

The Phase II setup is mostly identical to the Phase I experiment. The major difference is the replacement of the rectangular waveguide by a circular one with a diameter of 1 cm and a redesign of the waveguide cell. The goal of Phase II is to record the first molecular tritium spectrum with CRES and to study the systematic effects of the CRES technique. To this end, the gas system is designed to be able to supply the cell with krypton or tritium gas. The cell is equipped with five trapping coils and the cell windows are made of calcium fluoride, which prevents that tritium escapes from the cell via diffusion. The energy region of interest (ROI) of the Phase II tritium analysis almost covers the uppermost 2.5 keV of the spectrum. In fact, as the CRES technique records differential spectra, a large ROI does not require compromising on statistical power in the endpoint region. In addition, this choice of ROI places the 17.8 keV $^{83\text{m}}\text{Kr}$ line almost in the center of the analyzed energy range, which makes it ideal for studying systematic effects. These effects include (but are not limited to) magnetic trapping and waveguide mode coupling, both of which are described in detail in [93] and will be summarized in chapter 4. The Phase II signal processing and event detection was developed as part of this work and is presented in chapters 5 to 7, followed by the investigation of its performance in chapters 8 and 9. The recorded tritium spectrum is shown and analyzed in section 10.2.

3.2.2 Future Phases III and IV

Free space CRES demonstrator

The Phase I and II gas volume amounts to only a few cubic millimeters. For a CRES neutrino mass experiment with a sub-eV sensitivity, a volume of several cubic meters is required, which no longer fits into a waveguide. Therefore the future phases of Project 8 will be realized by the use of a patch-antenna or slotted-waveguide array surrounding the decay volume. In this array, the antennas (or waveguides) will be arranged cylindrically around a large gas cell as is depicted in figure 3.9. Trapped electrons will travel parallel to the longitudinal axis z of the cylinder while undergoing cyclotron motion. With a magnetic flux density of 1 T, the wavelength of the cyclotron radiation is on the order of 1 cm. The radius of the array is planned to be 5 cm in Phase III and significantly more in Phase IV. An electron that travels along z at a radius r emits radiation that is received by an antenna at the same z but different (r, ϕ) position³ with a phase shift $\Phi = |(\vec{r}_{\text{antenna}} - \vec{r}_{\text{electron}})| \cdot \frac{1}{\lambda_c}$. Unless the electron gyration center is on the array axis, all antennas will receive the signal at a different phase. To obtain the highest possible signal-to-noise ratio, the signal from each antenna must be phase-shifted before summation to compensate for the phase differences. As the position of the electron is unknown prior to detection, the identification of the combination of phase shifts that maximize the summed signal amplitude is an optimization task. Once the optimized phase shifts are found, the position of the radiating electron is known in r and ϕ .

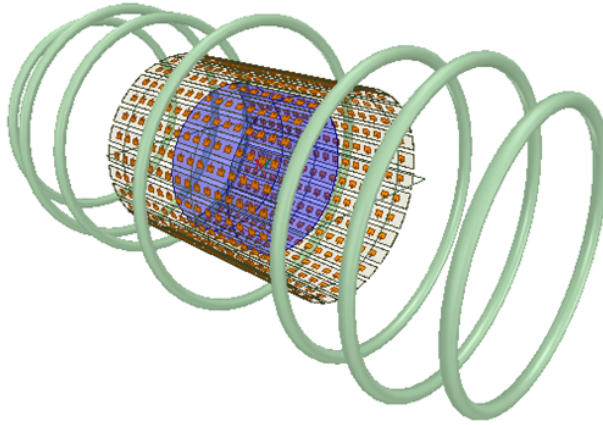


Figure 3.9: Patch antenna array design for Phase III from private communication with Timothy Wendler. The source gas is contained in an RF transparent vessel (blue). A cylindrical array of antennas encompasses the vessel and transmits the emitted cyclotron radiation. A range of coils (bright green) provide the trapping field.

Figure 3.10 shows the localization and signal power in the transverse plane of the antenna

³in cylindrical coordinates

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

array from figure 3.9. The color of each pixel indicates the summed signal power from all antennas after applying phase shifts according to the pixel location. One goal of Phase III is to demonstrate that free-space CRES with a phased antenna array can achieve the required efficiency to reach the exposure needed for the Phase IV neutrino mass experiment.

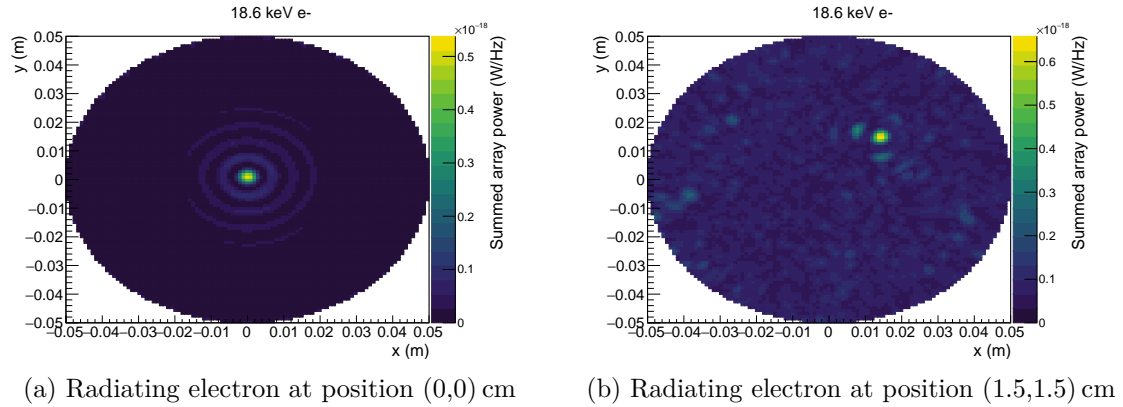


Figure 3.10: Simulated summed phase-shifted power in the transverse plane of an array with 30 antennas per ring. The position of the radiating electron can be reconstructed from the phase shifts that maximize the summed power. Images from private communication with Penny Slocum.

Atomic tritium

The shape of the β -spectrum of molecular tritium near the endpoint is affected by the broadening of the final states due to excitations of the daughter molecule. The width of these electronic, rotational or vibrational excitations has been calculated from theory, most recently by Saenz, Jonsell, and Froelich [94, 95]. Figure 3.11 shows the theoretical relative final bound state excitation spectrum of ${}^3\text{He}^+\text{T}$ and atomic ${}^3\text{He}^+$ [96] in comparison. The atomic final state spectrum is narrower by an order of magnitude and more than 8 eV lower in energy. In principle, any final state broadening can be included in the neutrino mass analysis, as long as the uncertainty on the broadening width and shape is small and precisely known. Based on the current theoretical calculations and their uncertainties, the remaining limitation of a neutrino mass measurement with molecular tritium is around 0.1 eV (depending on the size of the ROI).

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

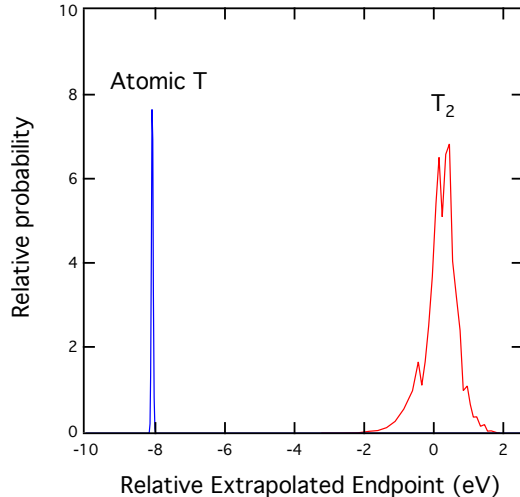


Figure 3.11: Final state broadening in atomic and molecular tritium decay [97]. The molecular spectrum has been calculated by [95].

To reach sensitivities beyond this limit, the Project 8 collaboration decided that it will use atomic tritium as source gas. This poses several challenges:

- **Doppler effect:** The effect of the thermal translational motion of the source atoms was described in section 3.1.3. Especially when using atomic tritium, the effect becomes nonnegligible. To minimize it, the tritium gas must be as cold as feasible. For the final Project 8 - Phase IV experiment, it is planned to cool the gas down to 30 mK.
- **Recombination:** Below a temperature of 30 K, tritium atoms hitting a surface are likely to recombine with available H atoms and reemerge as HT. To limit the number of contacts, the atoms are directed to an 8 K cold nozzle where they ideally scatter off the surface only once. This way, the gas is cooled by contact while still keeping the recombination rate at an acceptable level. For further cooling and trapping, a contact-less method is needed. Otherwise, it is not possible to store the atomic tritium gas in a vessel without losing the required T/T₂ purity.
- **Magnetic trapping:** As soon as the tritium is cooled to a temperature $T \leq 30$ K, it has to be isolated from all surfaces. Atomic tritium has an unpaired electron and therefore a magnetic moment $\vec{\mu}$ that allows for trapping in a magnetic potential

$$U = -\vec{\mu} \cdot \vec{B}. \quad (3.24)$$

The magnetic trapping of atomic antihydrogen has been demonstrated by the ALPHA experiment [98]. The experiment uses a Ioffe-trap field geometry with large alternating gradients near the trap walls and a negligible field in the trap

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

center. Trapping an atom with a velocity $v \leq 20$ m/s requires a magnetic field potential with $\Delta B = 2$ T. The velocities of thermal atoms follow a Maxwell-Boltzmann distribution and there is always a fraction of atoms with velocities that allow them to escape the trap. Depending on the loss rate, the trap must be constantly resupplied with new cold tritium atoms in order to maintain the gas pressure. A benefit of trapping atomic tritium magnetically is that recombined T_2 or HT molecules do not possess a large magnetic moment and will therefore escape the trap. This helps to maintain a high purity of the atomic tritium gas.

- **Atomic tritium production:** To compensate for the loss rate in the magnetic trap, tritium atoms need to be constantly resupplied. This requires an atomic tritium production of $\geq 10^{17}$ atoms/s. Atoms are produced by a thermal cracker via dissociation of molecular tritium. Tritium atoms and molecules that escape from the magnetic trap are recirculated via the thermal cracking and re-enter the cooling and selection processes.
- **Velocity selection:** To reach temperatures below 8 K, the fraction of atoms with a sufficiently low velocity can be selected from the Maxwell-Boltzmann distribution (see figure 3.12). The velocity-selected gas rethermalizes to a Maxwell-Boltzmann distribution with a temperature resulting from the average selected velocity. Velocity selection can be accomplished by the use of magnetic guiding fields, which direct the atoms in the desired velocity range into an acceptance aperture, while all other atoms are diverted from the allowed trajectory. However, this process suffers from low efficiency. Since CRES detection requires a high magnetic background field level, typically about 1 T, the selected velocity range is determined by the potential energy that the low-field seeking⁴ atoms lose when they enter the magnetic trap over a 1 T field step. To overcome this field step, they need a minimum velocity of $v = 60$ m/s and the velocity range to be selected is therefore 60 m/s to 80 m/s.

Like the phased antenna array, the atomic tritium concept must be demonstrated in Phase III. Only the combination of free space CRES with an atomic tritium source gas will allow Project 8 to achieve its target sensitivity of 40 meV. Figure 3.13 shows the schematic concept of the Phase IV experiment. The design parameters according to current knowledge are given in table 3.3.

⁴Low-field seeking: Spin state in which the atoms lose energy when climbing *up* in a magnetic potential.

3.2. A PHASED APPROACH TO MEASURING THE NEUTRINO MASS

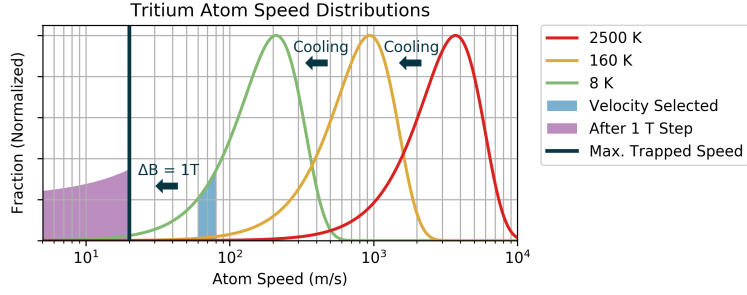


Figure 3.12: Velocity selection as shown in [97]. The velocity distribution of the tritium atoms is shifted downwards as a result of various cooling stages. By guiding the atoms magnetically through a chicane, only atoms within a defined velocity range reach the trap.

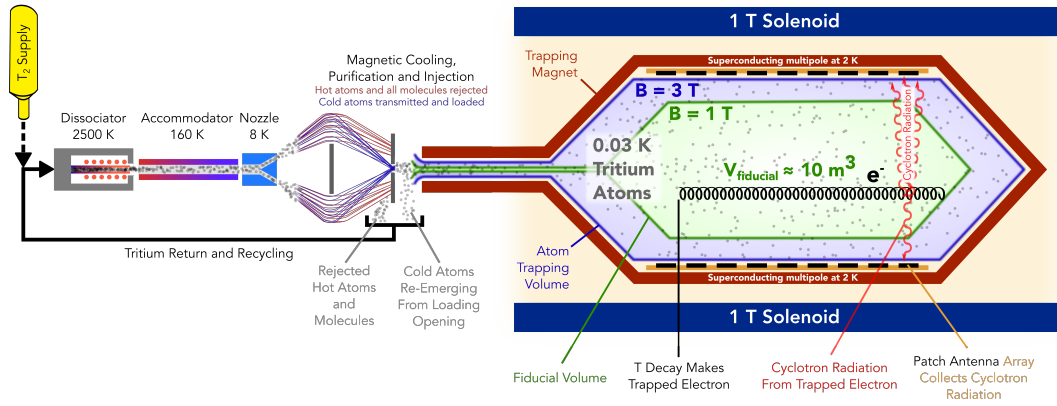


Figure 3.13: Design of the Phase IV experiment [97]. Tritium atoms are split by a thermal cracker and cooled via surface contact and velocity selection while being guided to a large gas vessel. The vessel is equipped with an array of antennas that detect the cyclotron radiation. An Ioffe-trap provides the trapping field for the atomic tritium gas.

| Parameter | Value | Comments |
|-----------------------|---------------------------|--|
| T density | 10^{12} cm^{-3} | Limited by detectable electron track length and frequency resolution requirement |
| T temperature | 30 mK | Keeps evaporation rate low |
| T_2/T concentration | $\leq 10^{-6}$ | Limits background from T_2 spectrum |
| $\sigma(B)/B$ | 10^{-7} | Required to achieve target frequency resolution |
| Ioffe trap depth | 2 T | Confines atomic tritium |
| Smallest dimension | 50 - 100 cm | 1 - 2 mean free path lengths for thermalization |

Table 3.3: Phase IV atomic tritium experimental requirements. All parameters are still subject to optimization studies.

4

CRES in the Phase II apparatus

One focus of this work is the CRES tritium spectrum and the systematic effects in the Phase II detector. To this end, a detailed understanding of the signal characteristics in the Phase II experimental setup is required. The signal signature of trapped electrons in a waveguide is described in detail in [93]. In this chapter, the effects that are most relevant to the Phase II tritium spectrum are summarized and the analytical predictions are compared to observations made in real krypton data taken during the calibration of the Phase II apparatus. Finally, the choice of trapping field geometry for the tritium data collection campaign is presented.

4.1 Signal characteristics in Phase II

The Phase II detector consists of a circular waveguide that is inserted into a superconducting NMR solenoid magnet so that their longitudinal axes are aligned (see section 3.2.1). The waveguide section containing the gas cell is depicted in figure 4.1. It is equipped with five copper coils that serve to provide the trapping field geometry. The trapping field optimization is addressed in section 4.2.2.

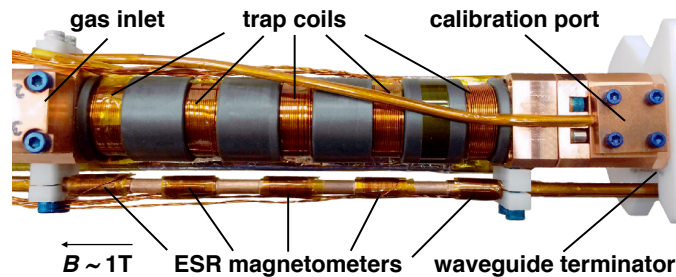


Figure 4.1: Phase II waveguide cell (as shown in [92]) with five copper coil windings that allow for adaptable trapping field configurations. The waveguide short from Phase I was replaced by a terminator to suppress reflections from the waveguide end.

The cell can be filled via the gas inlet with either $^{83\text{m}}\text{Kr}$ or T_2 gas. Electrons emitted in decays will start to follow a cyclotron trajectory. In addition, depending on their pitch angle¹ and the depth of the magnetic trap, they will oscillate inside the magnetic trapping field, along the longitudinal waveguide axis.

¹Pitch angle is the angle between \vec{B} and \vec{p}_e .

4.1.1 Effects from magnetic trapping

Trapping is required to achieve an observation time that is sufficiently long to distinguish the electron tracks from noise and to precisely determine their start frequency. Due to their oscillatory motion in the trapping field, electrons experience a modulated magnetic field strength. This leads to a range of trapping and pitch angle effects.

Trapped angles

Following the derivation from [93], the trajectory of the trapped electrons is decomposed into two components, parallel and perpendicular to the solenoid field axis

$$E_{\text{kin}} = E_{\text{kin}\parallel} + E_{\text{kin}\perp} = \frac{1}{2} \frac{p_0^2}{m_e} \cos(\theta(t))^2 + \mu(t)B(t). \quad (4.1)$$

Here, E_{kin} is the electron's kinetic energy, θ is the pitch angle, p_0 is the initial magnitude of the electron's momentum, and μ is the magnetic moment of the electron given by

$$\mu(t) = \frac{1}{2} \frac{p_0^2}{m_e} \frac{\sin(\theta(t))^2}{B(t)}. \quad (4.2)$$

In the adiabatic regime, the change of $B(t)$ experienced by the electron is slow compared to the duration of one cyclotron orbit. As a result, μ is a constant of the electron motion and any change of $B(t)$ causes $\sin(\theta(t))^2$ to change accordingly. An electron moving away from the field minimum B_{min} will possess an increasing magnetic potential energy $U(t) = \mu B(t)$. Energy conservation requires that $E_{\text{kin}}(t)$ and hence $\cos(\theta(t))^2$ decreases. If the trap depth $\Delta B = B_{\text{max}} - B_{\text{min}}$ is large enough such that $E_{\text{kin}} \rightarrow E_{\text{kin}\perp}$ before the electron reaches B_{max} , $E_{\text{kin}\parallel}$ vanishes as $\theta(t)$ approaches 90° and the electron is reflected. Thus, the condition for an electron to remain trapped is

$$\begin{aligned} \frac{1}{2} \frac{p_0^2}{m_e} \cos(\theta_0)^2 &\leq \mu \cdot (B_{\text{max}} - B_{\text{min}}) \\ \cos(\theta_0)^2 &\leq \frac{1}{B_{\text{max}}} \cdot (B_{\text{max}} - B_{\text{min}}) \\ \theta_0 &\leq \cos^{-1} \left(\sqrt{\frac{\Delta B}{B_{\text{max}}}} \right) = \sin^{-1} \left(\sqrt{1 - \frac{\Delta B}{B_{\text{max}}}} \right), \end{aligned} \quad (4.3)$$

where θ_0 is the pitch angle at the bottom of the trap, B_{max} is the maximum magnetic field magnitude and ΔB is the difference between the maximum and minimum B -field in the magnetic bottle. The trapping principle is depicted in figure 4.2. For the rest of this work, *pitch angle* will no longer refer to the instantaneous pitch angle $\theta(t)$, but instead to θ_0 , the pitch angle of any given electron at B_{min} . A trap coil current in the Phase II trapping coils of 1 A generates a field of approximately 4 mT. The relation of θ_0 to the trapping field depth and the trap coil current is shown in figure 4.3.

4.1. SIGNAL CHARACTERISTICS IN PHASE II

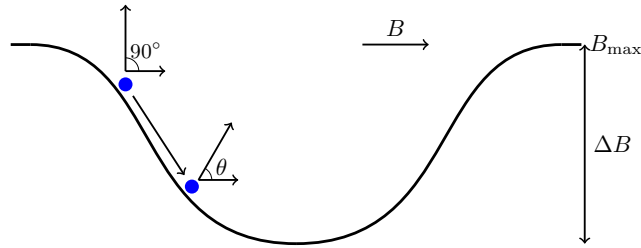


Figure 4.2: Axial motion in a magnetic trap. The trap depth ΔB is defined as the difference between minimum and maximum field magnitude. If the angle between momentum and field vector (pitch angle) satisfies equation (4.3) the electron remains trapped. At the turning points the pitch angle is 90° . Image similar to [93].

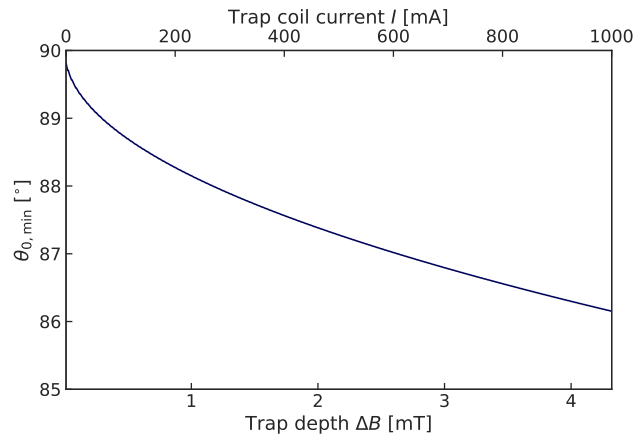


Figure 4.3: Minimum trapped pitch angle vs. trap depth. Electrons with pitch angles below the blue curve can escape the trap, while angles above are trapped. The trap depth is determined by the set trap coil current. A single trap coil powered by 300 mA causes a trapping field depth of 1.4 mT.

Frequency response to pitch angle changes

As the observation axis (waveguide axis) is aligned with the electron's oscillatory motion in the trapping field, the observed frequency $\omega_r(t)$ is shifted by the Doppler effect

$$\omega_r(t) = \omega_c(t_{\text{ret}}) \cdot \left(1 - \frac{v_z(t_{\text{ret}})}{v_p}\right)^{-1}, \quad (4.4)$$

where t_{ret} is the time of signal emission, v_z is the axial velocity, and v_p is the phase velocity. Because only large angles are trapped $\frac{v_z(t_{\text{ret}})}{v_p}$ is small compared to 1 and equation (4.4) can be approximated by

$$\omega_r(t) \simeq \frac{eB(t_{\text{ret}})}{m_e + E_{\text{kin}}/c^2} \cdot \left(1 + \frac{v_z(t_{\text{ret}})}{v_p}\right). \quad (4.5)$$

The dependency of equation (4.5) on the electron's pitch angle is twofold:

1. An electron with a smaller pitch angle θ_0 will travel further up in the trapping potential and therefore see a higher average magnetic field strength, which leads to a higher average cyclotron frequency. Signals from electrons with the same kinetic energy appear at different frequencies depending on their pitch angle. The width of the frequency distribution of a hypothetical population of trapped monoenergetic electrons is thus given by the range of trapped angles and therefore depends on the trap depth.
2. Both the magnitude of the magnetic field experienced by an electron and the Doppler shift vary periodically and induce frequency modulations [99] on the cyclotron signal, with the Doppler shift being the dominant effect. As a result, the received signal consists of a frequency comb structure with a main carrier signal at the average cyclotron frequency and evenly spaced sidebands (figure 4.4). The magnitude of the sidebands relative to the main carrier depends on the frequency modulation index

$$h = \frac{\Delta\omega}{\omega_a}, \quad (4.6)$$

where $\Delta\omega$ is the difference between maximum and minimum received frequency and ω_a is the axial trapping oscillation frequency. The smaller the pitch angle of an electron, the larger is its axial velocity and thus its maximum Doppler shift, which leads to a large modulation index h . In data recorded with the Phase II apparatus, the sidebands for any trapped pitch angle do usually not carry enough power to be detected and only main carriers are observed (depending on the trapping field geometry). The signal power in the main carrier has a steep dependence on the pitch angle (figure 4.9), which limits the pitch angles that can be detected to a small range. To reach the necessary exposure for a neutrino mass experiment, a

wide range of angles must be detectable. Therefore, the future Phases III and IV are designed with the aim to reduce the frequency modulation index by changing the observation axis to be perpendicular to the trapping motion.

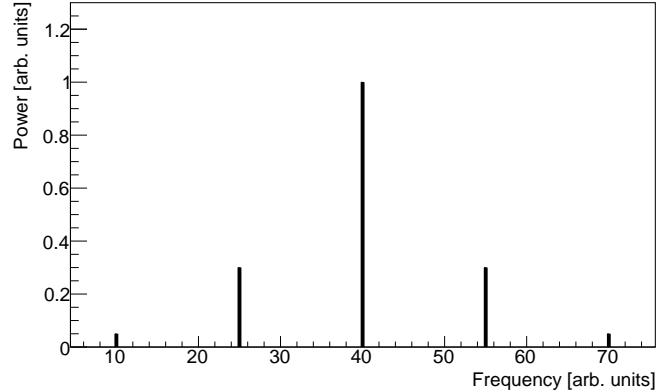


Figure 4.4: Sidebands comb structure [93]. The central peak is located at the average cyclotron frequency. The sideband spacing corresponds to the axial frequency in the trap.

Harmonic trap

Most of the data that is analyzed in this work was taken in a combination of multiple harmonic traps. A harmonic trap is the simplest magnetic bottle realization and is generated by a current in a single coil. To create a dip in the background field, the polarity of the trapping field has to be opposite to the polarity of the NMR magnet. The resulting axial field component $B_z(z)$ can be approximately described by a parabola with $z = 0$ m at the trapping field center

$$B_z(z) = B_0 \left(1 + \frac{z^2}{L_0^2} \right), \quad (4.7)$$

where L_0 is the trap length that adjusts the slope of the trap walls. B_0 is not the field strength of the NMR magnet, but instead the minimum field strength in the trap $B_0 = B_{\text{NMR}}(z = 0) + B_{\text{trap}}(z = 0)$. Since the field polarity of the harmonic trap is opposite to the NMR field, B_0 decreases with increasing trap coil current. Figure 4.5 shows the field profile along z together with the harmonic approximation from equation (4.7).

4.1. SIGNAL CHARACTERISTICS IN PHASE II

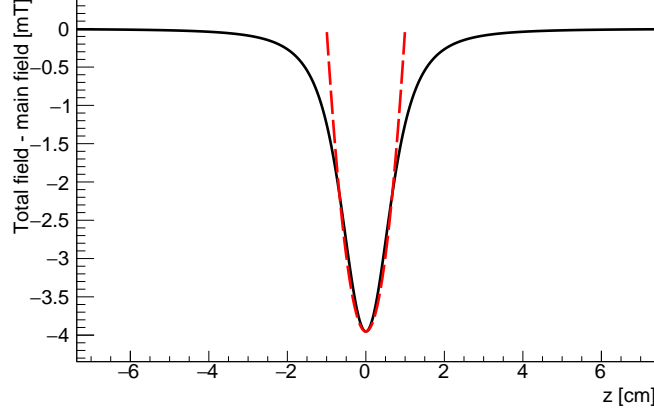


Figure 4.5: The axial field profile of a magnetic trap, as generated by a single coil, can be approximated by a parabola (4.7). Image from [93].

With the field approximation from equation (4.7) the electron's equation of motion (4.1) can be solved. For an electron with any given pitch angle θ_0 , the axial frequency follows from

$$\frac{1}{\omega_a} = \frac{2}{\pi} \int_0^{z_{\max}} \frac{dz}{\sqrt{\frac{2}{m}(E_{\text{kin}} - \mu B(z))}}. \quad (4.8)$$

The reflection position z_{\max} , at which $\theta(t) = 90^\circ$, can be calculated from $z_{\max} = L_0 \cot(\theta_0)$. The z -component of the magnetic field, seen by the moving electron, as a function of time is

$$B_z(t) = B_0 \cdot \left(1 + \frac{z_{\max}^2}{2L_0^2} - \frac{z_{\max}^2}{2L_0^2} \cdot \cos(2\omega_a t) \right). \quad (4.9)$$

Inserting $B_z(t)$ into equation (3.1) yields

$$\omega_c(t) = \frac{eB_0}{\gamma m_e} \cdot \left(1 + \frac{z_{\max}^2}{2L_0^2} - \frac{z_{\max}^2}{2L_0^2} \cdot \cos(2\omega_a t) \right). \quad (4.10)$$

Hence, $\omega_c(t)$ has a time-dependent term and a term that is given by the trapping field geometry. The varying term describes the frequency modulation, giving rise to the comb structure from figure 4.4. The time-independent term corresponds to the main carrier frequency ω_0 , averaged over the full electron trajectory

$$\omega_0 = \frac{eB_0}{\gamma m_e} \cdot \left(1 + \frac{z_{\max}^2}{2L_0^2} \right). \quad (4.11)$$

Via z_{\max} , ω_0 depends on the pitch angle of the electron. The signal from each pitch angle is thus observed at a different frequency. If this is not accounted for, a deeper trap (with larger ΔB) leads to a larger spread of the main carrier frequencies even for electrons with the same kinetic energy. There are two options in dealing with this frequency broadening:

1. If the pitch angle can be determined from signal characteristics other than the frequency, it can be used to correct the observed main carrier start frequency. This way, the true kinetic energy can still be extracted.
2. Without any frequency correction, the frequency broadening will result in an additional uncertainty in the energy determination. This can be treated as energy resolution in the spectrum analysis.

While the first option is more favorable, it could not be employed for the tritium data presented in this work. This has to do with the selected trapping field geometry (see section 4.2.2) that complicates the extraction of pitch angle information from the data. The effect of pitch angles on the spectrum and the tools needed for their correction are described and discussed in [93, 100].

Magnetic drift motions

In addition to the cyclotron orbit and axial trapping oscillation, the electron undergoes two drift motions induced by the nonuniformity of the magnetic trapping field. Local field gradients cause a drift perpendicular to both the magnetic field and its gradient (grad-B motion). This drift is analogous to the magnetron motion in Penning traps (figure 4.6). It describes a circle with the normal coinciding with the trapping axis z

$$\vec{v}_{\text{grad-B}} = \frac{\mu}{m\omega_c B} \cdot \vec{B} \times \nabla B. \quad (4.12)$$

With a field gradient on the order of 10 mT/m in a 1 T magnetic field, the grad-B drift is slow relative to the cyclotron and the axial trapping motion. In fact, the duration of one orbit is of the same magnitude as the average total track length and is therefore not relevant for the determination of the main carrier frequency. However, the grad-B drift can introduce a time dependence of the axial frequency that can help to distinguish main carriers from sidebands. Sideband classification was demonstrated based on Phase I data in [100].

The second drift is induced by the curvature of the magnetic field lines (curvature drift). Under Phase II conditions it is one order of magnitude smaller than the grad-B motion and therefore not observable on the trapping time scales.

4.1. SIGNAL CHARACTERISTICS IN PHASE II

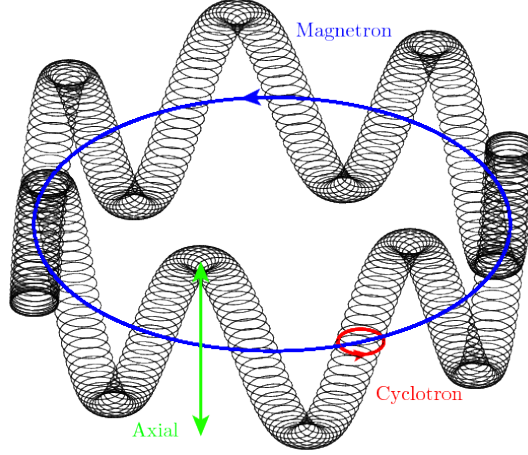


Figure 4.6: The trajectory of an electron in a Penning trap [101] is the combination of its cyclotron motion and axial trapping oscillation as well as a comparably slow magnetron drift. In a fully magnetic trap, the magnetron drift is replaced by the grad-B motion.

4.1.2 Received power

As the evanescent waves don't reach up to position of the amplifiers, the received power does not follow equation (3.23). Only power coupled to a propagating waveguide mode can be observed at the amplifier. In a circular waveguide the fundamental traveling mode is the transverse electric TE_{11} mode, with the power spectral density

$$P_{0,TE_{11}} = \frac{Z_{11}e^2v_0^2}{8\pi\alpha} \cdot \left(J_1'^2(k_c\rho_c) + \frac{1}{k_c^2\rho_c^2} \cdot J_1^2(k_c\rho_c) \right). \quad (4.13)$$

Here, Z_{11} is the TE_{11} mode impedance, J_n is the n^{th} Bessel function, ρ_c is the radial position of the electron's gyro-center in cylindrical coordinates, v_0 is the electron's velocity, k_c is the wavenumber for the cut-off frequency of the mode, and α is a normalization factor

$$\alpha = \int_0^R \left[J_1'^2(k_c\rho) + \frac{J_1^2(k_c\rho)}{k_c^2\rho^2} \right] \rho d\rho. \quad (4.14)$$

The mode impedance can be calculated as

$$Z_{11} = \frac{Z_0}{\sqrt{1 - \frac{f_k^2}{k^2}}} \quad (4.15)$$

with the free space plane wave impedance $Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$. Equation (4.13) is an approximation for the condition that the cyclotron radius $r_e = \frac{v}{2\pi c}\lambda_c$ is small compared to λ_c , the wavelength of the cyclotron radiation, which is obviously justified. The mode impedance depends on frequency as shown in figure 4.7a. The change of the coupling amplitude with radius is depicted in 4.7b. It can be seen that the amplitude decline with increasing radius

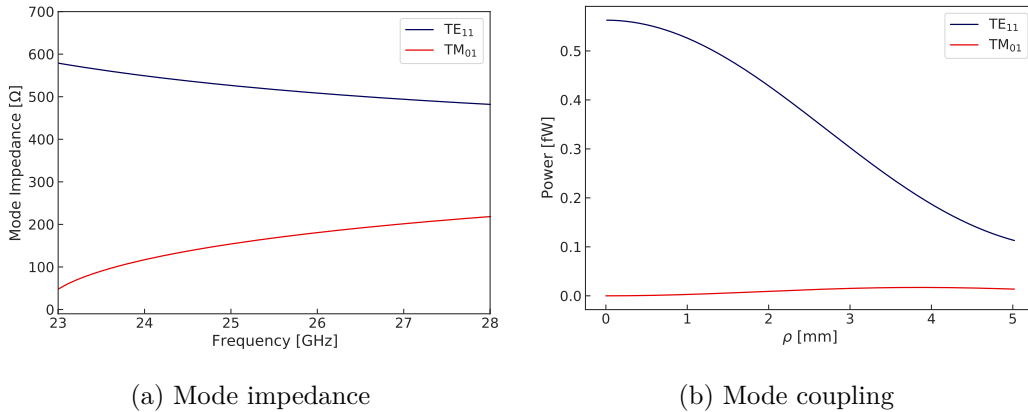


Figure 4.7: Mode impedance and coupling in the TE_{11} and TM_{01} modes. The radial dependence of the coupled power in the TE_{11} mode reduces the volume in which electrons can be detected.

is steep. The probability for an electron to be born at a radial position grows linearly with the radius. As a result, the fraction of detectable electrons is small, since the radial coupling dependence limits the volume in which electrons can be observed. In addition, the distribution of the received power is expected to be wide even for monoenergetic electrons.

The presence of v_0^2 in equation (4.13) reveals the dependence of the coupled power on the electron's kinetic energy. Classically, this dependence would be linear, albeit in special relativity the velocity increases with γ as

$$\begin{aligned} v &= c \cdot \sqrt{\left(1 - \frac{1}{\gamma^2}\right)} \\ &= c \cdot \sqrt{1 - \frac{1}{\left(1 + \frac{E_{\text{kin}}}{E_0}\right)^2}}. \end{aligned} \quad (4.16)$$

In the energy region of interest (< 20 keV) the 2nd order term in a Taylor series expansion of equation (4.16) only contributes about 10% compared to the 0th and 1st order terms. The energy dependence is thus almost linear (figure 4.8). The spectral power for each band P_n (main frequency or sideband) can be found by determining the Fourier coefficients a_n

$$P_n = P_{0,\lambda} |a_n|^2. \quad (4.17)$$

In the approximation of shallow harmonic magnetic traps, where the magnetic field modulation is small compared to the Doppler-induced frequency modulation, the power in the n^{th} band simplifies to

$$P_n = P_{0,\lambda} J_n(k_\lambda z_{\text{max}})^2. \quad (4.18)$$

4.1. SIGNAL CHARACTERISTICS IN PHASE II

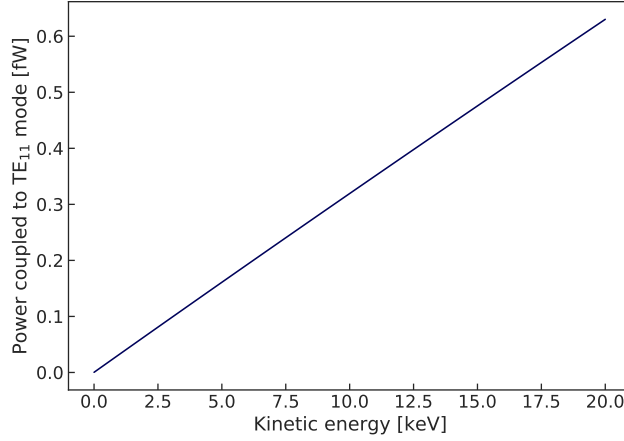


Figure 4.8: Power transmitted in the TE₁₁ mode at $\rho = 0$ cm vs. kinetic energy of a trapped electron for $\theta_0 = 90^\circ$ and $B \approx 0.96$ T.

Figure 4.9 shows the power of the main carrier signal and the first and second order sideband relative to the maximum power in the main carrier, as a function of θ_0 . The trap geometry underlying figure 4.9 is a 1.3 mT deep harmonic trap (trap coil 3 powered with 300 mA). It can be seen that the maximum relative sideband power amounts to approximately 35%. The fact that no sidebands were observed in Phase II, in combination with the maximum coupled power being approximately 0.57 fW (figure 4.7b), indicates that the detection threshold is ≥ 0.2 fW.

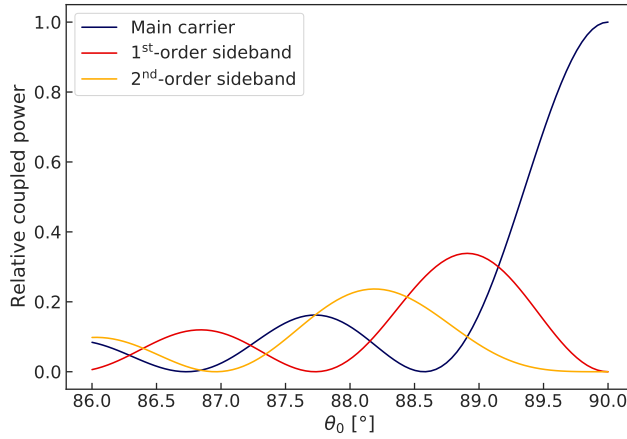


Figure 4.9: Power in the main carrier and sidebands in a harmonic trap ($L_0 = 0.178$ m). The power strongly depends on the frequency modulation index h . If the detection threshold is $> 0.35 \cdot P_{\max}$ no sidebands can be observed.

With equation (4.13) and (4.18), the power dependence on pitch angle θ_0 and radial position in the waveguide ρ are determined. The two-dimensional relation is shown in figure 4.10. Every trapped electron is born at some initial position in the ρ - θ_0 plane and

is observed with the according coupled power. The minimum received power required to detect an electron affects the minimum observable pitch angle and the maximum observable radial position. Since the minimum observed angle determines the width of the frequency broadening, a higher detection efficiency increases the exposure at the cost of leading to a larger energy uncertainty (if no frequency corrections are applied).

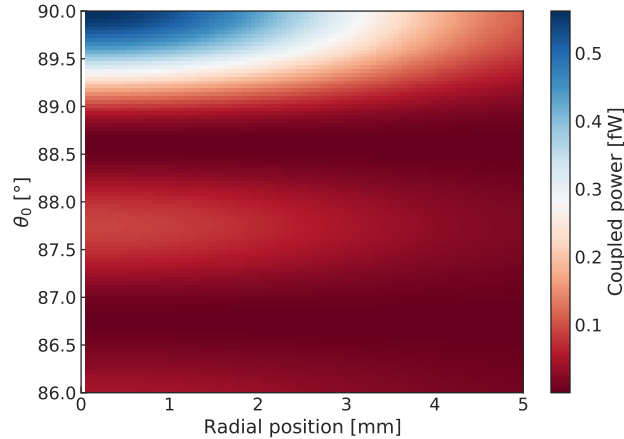


Figure 4.10: Combining equation (4.13) and (4.18) yields the two-dimensional power dependence on pitch angle and radial position in the TE_{11} mode.

4.1.3 Reflections and additional modes

The waveguide has a finite length. On one end the signal is transmitted via a receiver. The other end is closed off by a waveguide terminator. In this configuration, the receiver will only detect half of the electron's emitted power and the signal will ideally be the same as that of an infinitely long waveguide. In reality, reflections from the opposite waveguide end are small but not absent. This causes interference effects that enhance or suppress the received signal for certain combinations of electron position along the waveguide and frequency. As the reflections are small, it is not expected that the power from any given frequency and trap location is entirely suppressed. However, the signal power shows a frequency dependence that is distinct for each trap, as each trap has a different distance to the waveguide terminator. In addition, the presence of additional waveguide modes can lead to increased power losses to these modes. All these effects need to be measured in order to understand the shape of a spectrum that is recorded with a waveguide.

4.1.4 Energy loss

Slope

The total energy loss via radiation exhibits itself in the slope s of the electron tracks

$$s(t) = \frac{dE_{\text{kin}}}{dt} = -P(t), \quad (4.19)$$

where $P(t)$ is the total emitted power. Under the assumption that the power loss is constant over a short period of time and the loss is slow compared to the cyclotron and axial trapping motion, equation (3.1) becomes approximately

$$\omega_c(t) \simeq \frac{qB(t)}{m_e + E_{\text{kin},0}/c^2} \cdot \left(1 + \frac{Pt}{m_e c^2 + E_{\text{kin},0}} \right). \quad (4.20)$$

The constant frequency drift with time due to the energy loss is clearly visible in figure 3.1. The slope can be used to identify the frequency positions of additional waveguide modes, as well as the sum of the power from all frequency bands.

Scattering

When electrons scatter off residual gas molecules, they lose energy and their pitch angles change. The energy loss function (figure 4.11) of electrons scattering from hydrogen molecules was measured by [88].

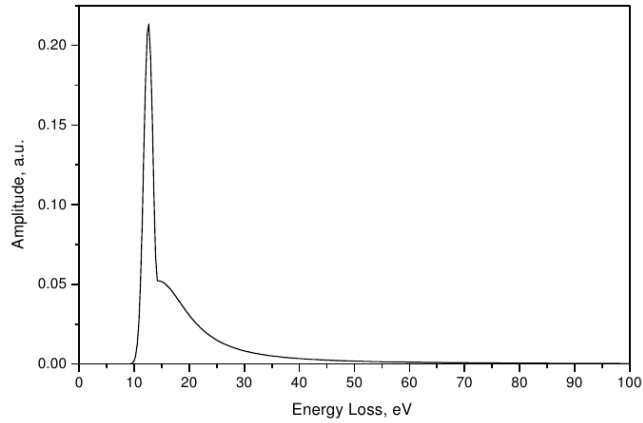


Figure 4.11: Fitted electron energy loss spectrum for single inelastic scattering off hydrogen molecules [88].

4.1. SIGNAL CHARACTERISTICS IN PHASE II

In CRES, electrons can be observed from the moment of the decay, which in principle could prevent the need to consider energy losses due to scattering when analyzing a spectrum recorded with CRES. However, a track can only be distinguished from noise if it is sufficiently long and powerful. The track lengths follow an exponential distribution and short tracks can remain unobserved before the first, or even the n^{th} , scatter. In addition, if the magnetic trap depth is such that electrons from a wider range of pitch angles are trapped than can be detected, it is possible that previously undetectable electrons only become visible to the event reconstruction after scattering. In Phase II, the energy region of interest is chosen to be large, since the event rate is low due to the small volume in the waveguide. Therefore, taking the energy loss from scattered events into account is crucial for the analysis of a spectrum. The effect of missed tracks can be seen from the $^{83\text{m}}\text{Kr}$ K-line peak in figure 4.12. The events that were only detected after the n^{th} scatter (with $n > 0$), appear as a high frequency (low energy) tail.

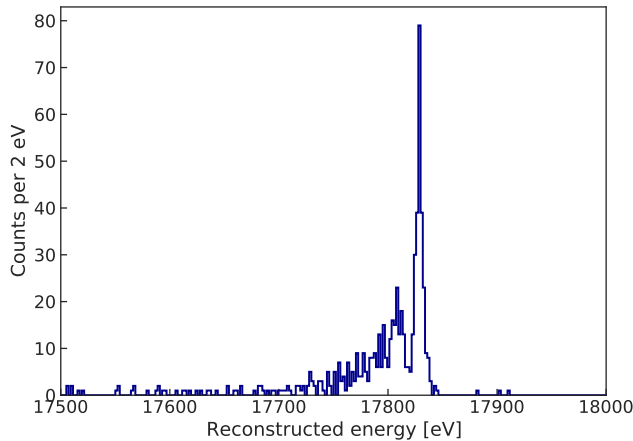


Figure 4.12: Energy spectrum of $^{83\text{m}}\text{Kr}$ K-line electrons recorded with CRES. The tail on the left-hand side of the main peak consist of events that were only detected after scattering.

In the Phase II cell, the gas consists not only of hydrogen. Especially when taking krypton data, a significant fraction of the gas is krypton. The krypton scattering cross section is larger than the one for hydrogen. Therefore, a multi-gas-species scattering model was developed for Phase II. This model consists of two components: a 0^{th} order peak describing the shape of the un-scattered start frequency distribution, and a tail that contains all events that were detected after scattering. The 0^{th} order peak shape F_0 consists of the intrinsic shape of the K-line [102] convolved with a normal distribution of standard deviation σ_0 , which represents the energy resolution function. The frequency position of the 0^{th} order peak is $f_0 = f(\bar{B}, 17.83 \text{ keV})$, where \bar{B} is the average magnetic field sampled by the electron trajectories in the trap. The scattering tail is composed of the sum of N scattering peaks. The shape of the 1^{st} scatter is the weighted sum of

4.1. SIGNAL CHARACTERISTICS IN PHASE II

the convolution of the gas-species-dependent energy loss function with the 0th order peak. The weights $\gamma_{\text{H}_2, \text{Kr}, \text{He}}$ represent the probability for an electron to scatter of one of the three most abundant gas species in Phase II (krypton, helium, and hydrogen). The ratio of the area under the 1st scattering peak over the area under the 0th order peak is given by the probability to detect the (n+1)th instead of the nth track as first track. This probability parameter is referred to as survival probability $p_{\text{scattering}}$. The nth scattering peak $F_{n, (k_1, k_2, k_3)}$ is composed of all possible combinations of scattering sequences (e.g. an electron in the second scattering peak could have scattered off once from hydrogen and once from krypton before it was detected), where each scattering event is incorporated by another convolution with the respective energy loss function. The indices $k_{i\text{-gas}}$ indicate how often the electron scattered from a molecule of the gas species i . The full lineshape (limited to three gas species) is described by

$$\text{PDF}_{\text{Lineshape}} = \sum_{n=0}^N \left[\sum_{k_1+k_2+k_3=n} \binom{n}{k_1, k_2, k_3} \cdot p_{\text{scattering}}^n \cdot \gamma_{\text{H}_2}^{k_1} \gamma_{\text{Kr}}^{k_2} \gamma_{\text{He}}^{k_3} \cdot F_{n, (k_1, k_2, k_3)} \right]. \quad (4.21)$$

This model is used throughout this work in the analysis of the presented krypton (e.g. figure 4.14) and tritium data. The shape of the tritium spectrum recorded in Phase II (section 10.2) corresponds to the intrinsic spectral shape (section 2.6) convolved with the lineshape model (based on a Dirac peak instead of the intrinsic K-lineshape) and multiplied with the detection efficiency dependence on energy (as in figure 3.5).

4.2 Magnetic trap configuration

4.2.1 Harmonic trap studies

Figure 4.13 shows the start frequency histograms of 17.8 keV (K-line) electrons from $^{83\text{m}}\text{Kr}$ decay recorded in a harmonic trap at different trap coil currents: 12 mA, 50 mA, 100 mA, 300 mA. Each current results in a different trap depth (figure 4.3). For every tested trap depth events were accumulated over a duration of 3 hours². As expected, the number of events in the recorded peak and the peak width grow with increasing trap depth. This is due to the correlation of the range of trapped pitch angles with ΔB . The impact on these two quantities is twofold:

- As discussed in the section 4.1, a wider range of trapped angles results in a higher rate of trapped electrons. The average magnetic field seen by an electron depends on its pitch angle. This slightly changes the frequency at which each electron is observed and leads to a broader start frequency peak.
- The larger ΔB , the more likely it is that an electron remains trapped after scattering of a residual gas molecule. As a consequence, the average number of tracks grows with trap depth (figure 4.15c). In addition, events with more tracks are easier to detect, because it is more likely that any of the tracks will be long enough and visible above the detection threshold. At the same time, this increases the risk of the first detected track not being the first actual track in the event. Therefore, a larger fraction of events appears in the tail of the start frequency histogram, leading to an even wider peak.

The start frequency histograms in figure 4.13 are fitted with the multi-gas-species lineshape model (4.21). The results are shown in figure 4.14. The gas weights γ_{H_2} , γ_{He} , and γ_{Kr} are fixed in each fit and were calculated from the product of the measured partial pressures in the gas system and the electron scattering cross section with a gas species. The fit results give a quantitative insight into how the peak is affected by increasing the trap depth. Figure 4.15b shows the linear shift of the frequency position f_0 with the trap coil current, resulting from the change of \bar{B} . The energy resolution σ_0 also increases linearly with current (figure 4.15a). The scatter peak ratio (figure 4.15d) shows a similar dependence on the trap depth as the average number of tracks per event (figure 4.15c).

²When this data was taken, the Rubidium source had aged by several half-lives and the source gas activity was low.

4.2. MAGNETIC TRAP CONFIGURATION

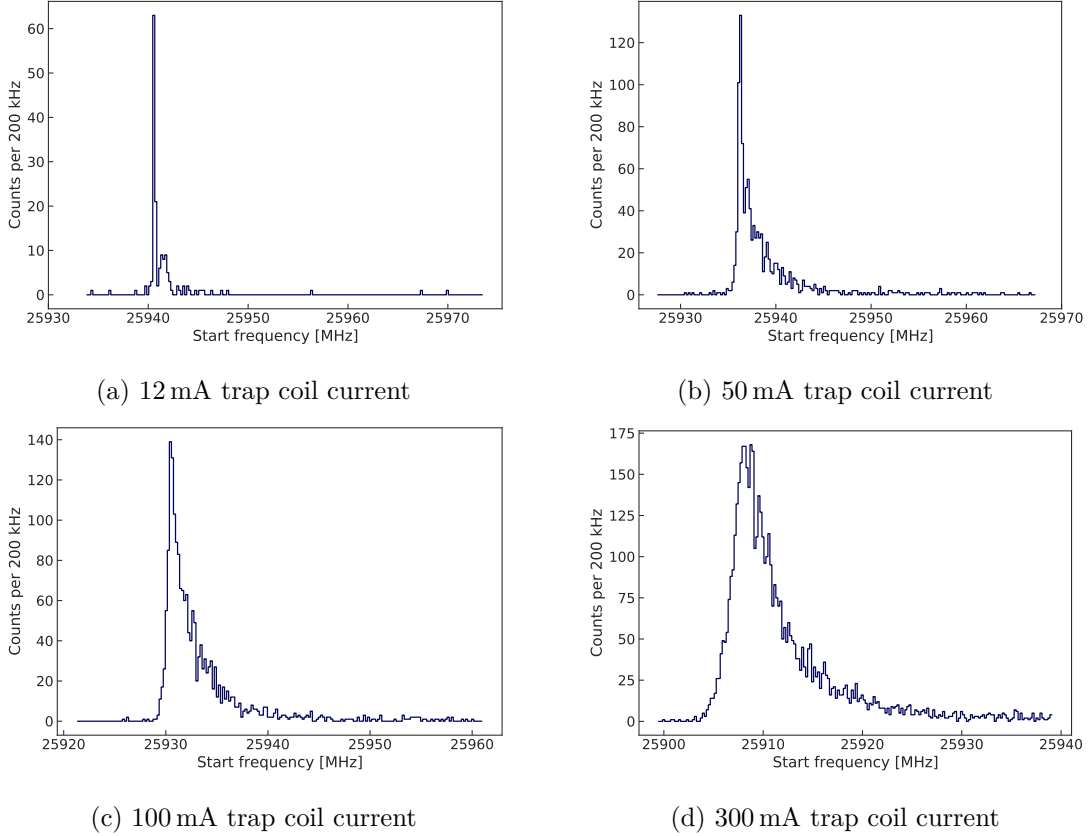


Figure 4.13: Start frequency histograms of K-line electrons recorded in harmonic magnetic trapping fields of different depths. With no start frequency corrections employed, the width of the peak broadens with increasing trap depth. This leads to a blurring of the main peak and the tail shape.

4.2.2 Trap optimization for tritium

With trapping, pitch angle, and power coupling effects in mind, the trap configuration for tritium data taking needed to be optimized. In Phase II, the main challenge in recording a tritium spectrum was the low count rate due to the combination of the much lower half-life of tritium (compared to $^{83\text{m}}\text{Kr}$) with the small volume in the waveguide gas cell of only a few cubic centimeters. Therefore, the goal of the trap optimization was the maximization of the observed event rate, even if that required giving up on the outstanding energy resolution that the CRES technique can offer. This was achieved by using a deep magnetic trap. In theory, the energy resolution can be recovered by identifying the pitch angle of each event, but in this work, no frequency corrections are employed. Due to the steep decrease of coupled power towards smaller pitch angles, a trap coil current of 300 mA already traps a larger range of angles than can be detected. To further increase the exposure, four trap coils were powered simultaneously to generate four serial harmonic trapping fields. The use of such a multi-harmonic trapping field

4.2. MAGNETIC TRAP CONFIGURATION

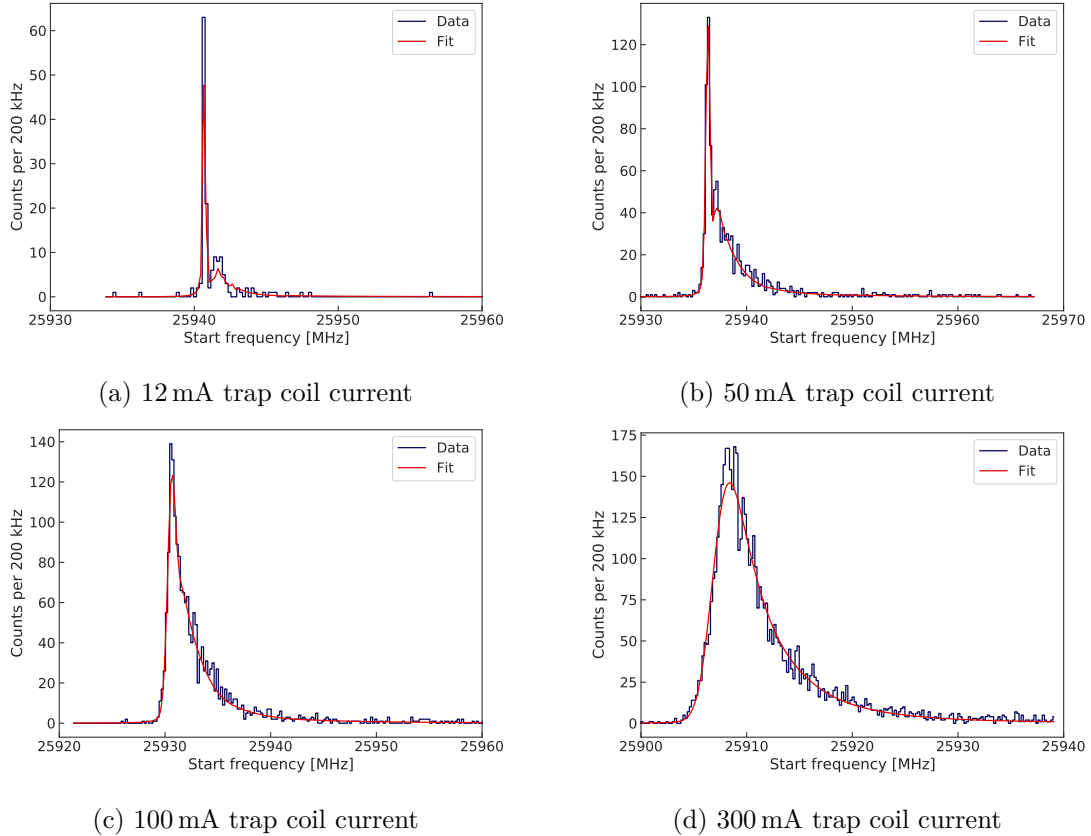


Figure 4.14: Start frequency histograms from figure 4.13 fitted with the multi-gas-species lineshape model (4.21). The remaining discrepancy between data and model likely originates from the assumption that the energy resolution function is bell-shaped.

configuration does not increase the width of a start frequency peak, as long as the bottom fields $B_{0,i}$ of all individual harmonic traps are aligned. To this end, the trap current scan was repeated for four of the five trap coils³. Figure 4.16 shows the peak frequency position $f_{0,i}$ for all four traps at several trap depths in comparison. While in all traps $f_{0,i}$ is a linear function of the coil current I_i , the intercept with $I_i = 0$ A differs from trap to trap. The reason for this behavior is the non-uniformity of the background magnetic field along z . From the linear fits of $f_{0,i}(I_i)$, the currents required to align $B_{0,i}$ were calculated. In doing this calculation, the effect of each trap coil field on the neighboring traps had to be taken into account. The currents were chosen such that $B_{0,i}$ in the multi-harmonic configuration is equal to $B_{0,3}$ in the single-harmonic trap 3 at 300 mA. Therefore, the trap configuration was named Q300 (quadruple-harmonic trap equivalent to 300 mA). The resulting field profile as a function of the longitudinal position z in the waveguide is shown in figure 4.17.

³The fifth coil is at an unfavorable location, where the background field strength makes an alignment with the other four coils impossible.

4.2. MAGNETIC TRAP CONFIGURATION

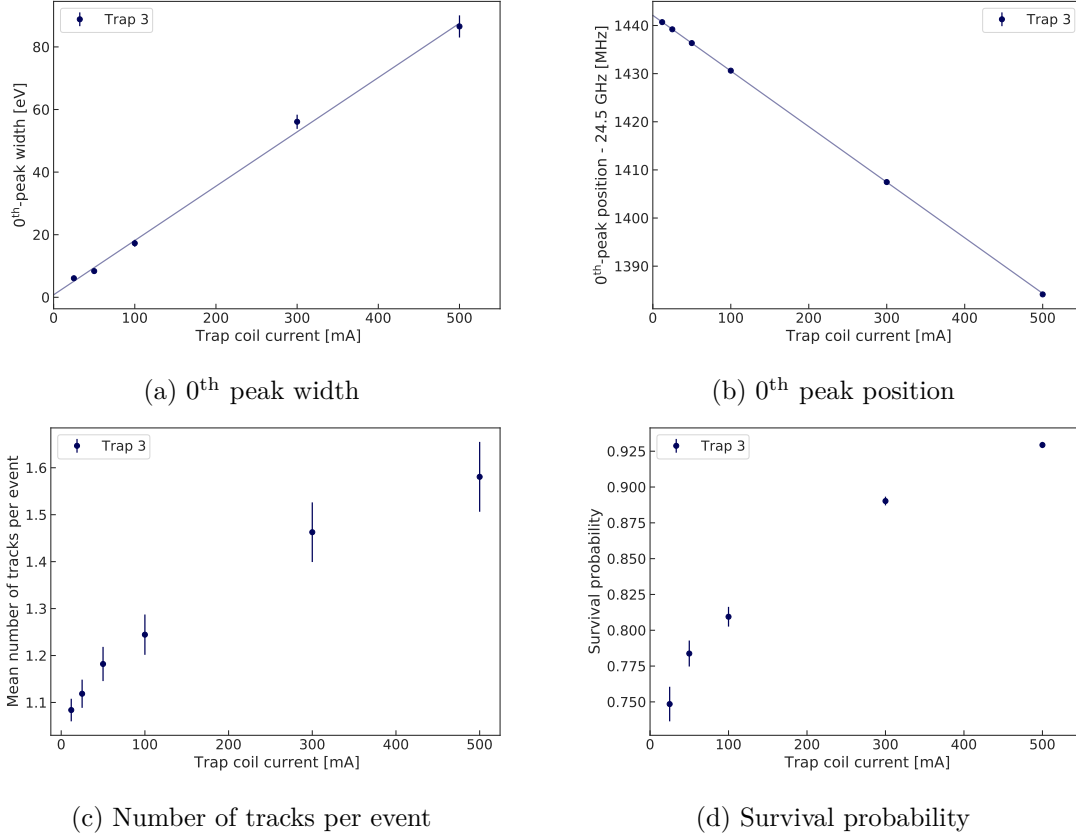


Figure 4.15: The lineshape model fit results give insight on various trapping effects. The range of trapped pitch angles and therefore the width of the un-scattered peak (a) increases with trap depth. As the field vector of the trap is antiparallel to that of the background field, the average magnetic field experienced by the trapped electrons decreases linearly with trap coil current and so does the frequency position of the K-line peak (b). Like the number of tracks per event (c), the survival probability (d) grows with trap depth, since a larger fraction of events is detected after scattering.

Because of the background field shape, the trap depths ΔB_i of the four harmonic traps in the Q300 configuration are different and the lineshapes from the individual traps are therefore not identical. The survival probability and the energy resolution in data recorded in this configuration adopt effective values. The detected event rate contributed from each trap, also differs as a result of the differences in ΔB_i . However, the dominant reason for count rate variations from trap to trap is induced by reflections and modes in the waveguides. In fact, using four traps instead of one does not increase the total count rate by a factor 4. Instead, the count rate in the Q300 trap is only about 1.8 times larger than the rate in the single-harmonic trap with the highest count rate per se (trap 3). Figure 4.18 shows the count rate dependence on the trap coil current in all four traps (powered individually). The systematic effects arising from the count rate behavior in each trap will be studied in chapter 9. The lineshape parameters originating

4.2. MAGNETIC TRAP CONFIGURATION

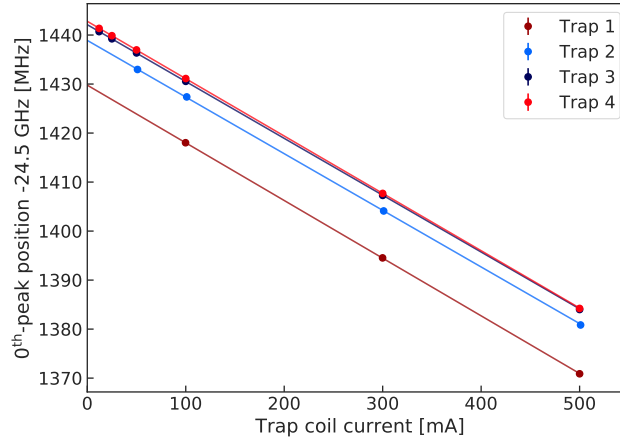


Figure 4.16: K-line frequency position vs. trap coil current in different traps. Since the background field profile is curved, the K-line frequency positions in trap 1 and 2 are offset compared to trap 3 and 4.

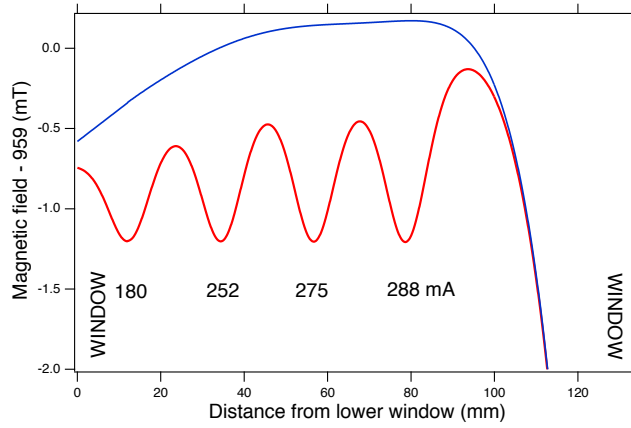


Figure 4.17: The Q300 trap configuration. To increase the number of trapped electrons four traps are used simultaneously generating a quadruple-harmonic trapping field geometry (red). The trap coil currents are set such that the curvature of the background magnetic field (blue) is compensated and the bottom field magnitudes are the same in all four traps. Image from private communication with R.G. Hamish Robertson.

from the Q300 trap configuration will impact the shape of the recorded tritium spectrum. To determine the lineshape parameters accurately, K-line electrons were recorded over 24 hours in this trap. Figure 4.19 shows the resulting start frequency histogram, fitted with the lineshape model. The fit yields the model parameters $\sigma_0 = (15.8 \pm 0.2)$ eV and $p = 0.869 \pm 0.001$. The magnetic field is calibrated to $\bar{B}_{Q300} = (957.80 \pm 0.003)$ mT.

4.2. MAGNETIC TRAP CONFIGURATION

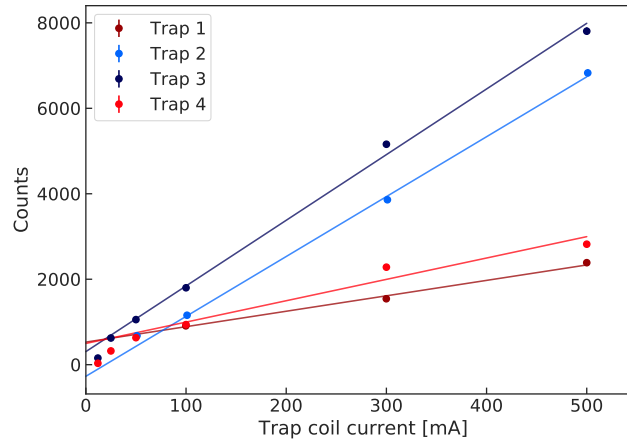


Figure 4.18: Detected number of events in the individual traps 1 to 4 as a function of the trap coil current. Because of interference effects in the waveguide, the received power from trapped electrons, and thus the fraction of detectable events, is frequency and trap dependent.

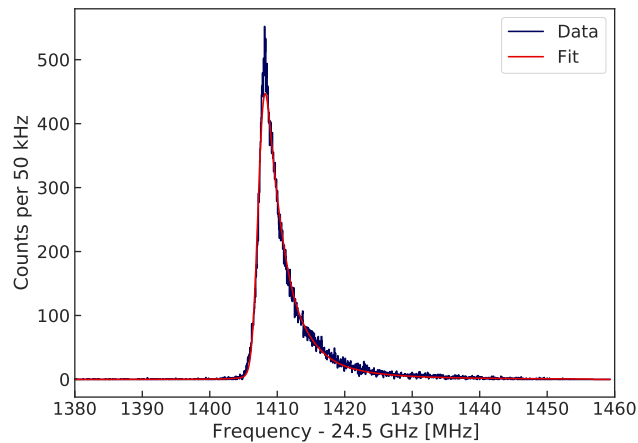


Figure 4.19: Start frequency histogram of $^{83\text{m}}\text{Kr}$ K-line events recorded in a 12 hours long run in the Q300 trap configuration. This data is used as magnetic field and lineshape calibration data for the Phase II tritium analysis in section 10.2.

5

Data acquisition in Phase II

In the previous section, the Phase II detector insert was described and an introduction to the detector response to various magnetic trapping field configurations was given. In order to produce start frequency histograms as shown for example in figure 4.19, the narrow-band power emitted by an electron has to be processed, digitized and its signal features have to be extracted. This is achieved in four steps:

1. Analog signal processing
2. Digitization and digital signal processing
3. Real-time triggering
4. Offline event reconstruction

Steps 2 to 4 are the core of this work. In this chapter, the analog and digital signal processing are described. The real-time and offline event detection techniques employed in Phase II are treated in chapters 6 and 7. Finally, the end-to-end performance of the event detection chain is studied in section 8.3 and section 9.2, followed by the analysis of its impact on the Phase II tritium spectrum as a source of systematic uncertainty in chapter 10.

The goal of the Phase II analog signal processing is to maintain the visibility of the CRES signal above the RF noise floor level. The lower the cell and terminator temperature the lower the thermal noise power and the better the signal-to-noise ratio (SNR). In Phase II, the RF system performance was studied using $^{83\text{m}}\text{Kr}$ data. To avoid that differences in temperature modify the behavior of systematic effects, it was decided to record tritium data at the same cell temperature as was set for the systematic studies performed with krypton gas. At cell temperatures below 115 K krypton begins to freeze on the cell walls. The trade-off temperature, which maximizes the detected $^{83\text{m}}\text{Kr}$ event rate by compensating the freeze-out rate with an improved SNR, is 85 ± 1 K. All data presented in this work were recorded at this cell temperature.

Figure 5.1 shows a schematic picture of the Phase II analog signal processing chain. The circular waveguide connects to a rectangular WR-42 waveguide in which only the TE_{10} mode propagates. Between the waveguide cell and the first low noise amplifier (LNA), the signal passes a waveguide isolator (a waveguide circulator with a terminator on its second port) that prevents the amplifier noise from leaking back into the waveguide. The

5.1. ANALOG SIGNAL PROCESSING

first stage of the receiver chain is formed by two cascaded LNAs (with a gain of 26.5 dB for $22 \text{ GHz} < f < 40 \text{ GHz}$). The system noise temperature is calculated from the sum of the terminator and isolator temperatures, weighted by the isolator's transparency and emissivity, and the temperature of the first stage amplifiers. Together they result in a noise temperature of about 132 K for a cell temperature of 85 K. All downstream stages only insignificantly add to the total noise temperature. Following the cold amplifiers, the signal is coupled to a coaxial cable and enters the ambient temperature part of the RF system, which consists of a high- and low-frequency stage. The combined system gain amounts to 76 dB.

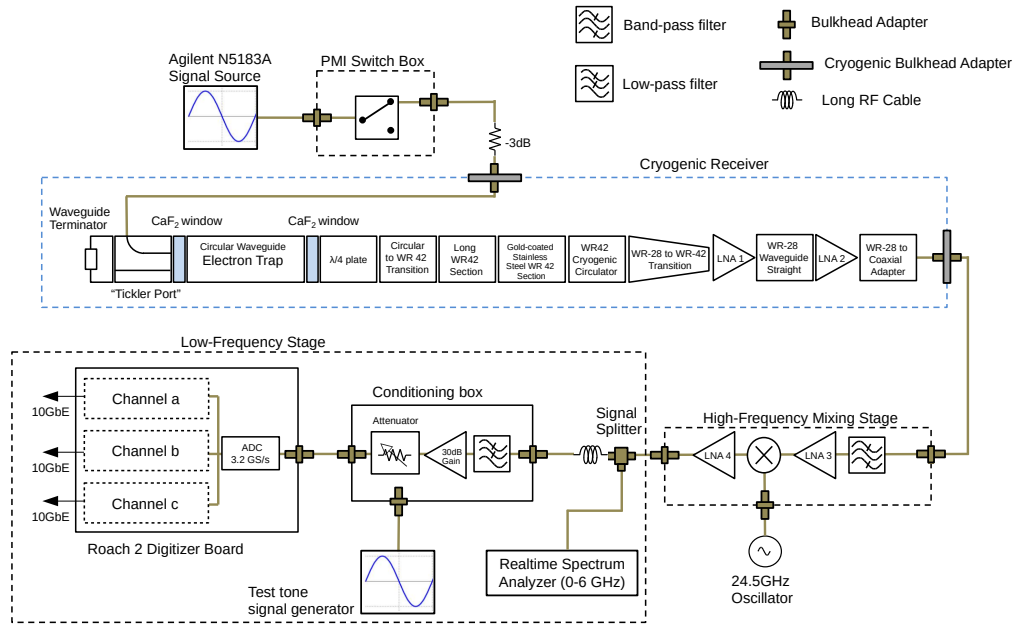


Figure 5.1: Schematic signal processing in the Phase II experiment. Power emitted by trapped electrons couples to the TE_{11} mode of a circular waveguide. The transmitted signal is amplified by cascaded cryogenic LNAs, followed by analog down-conversion with 24.5 GHz and digitization using a ROACH2 board. Image adapted from [92].

5.1 Analog signal processing

The data acquisition system should allow parallel acquisition of the quasi-monoenergetic K-, M- and L-lines (M and L are line doublets) of ^{83m}Kr decay and must also be capable of recording a wide frequency range corresponding to the energy region of interest for tritium data taking. At the same time, the amount of transferred data has to remain manageable. Before the analog signal is digitized, it undergoes two stages of mixing, filtering and amplification.

5.1.1 High-frequency mixing

In the digitization process, the signal has to be sampled with a sampling rate at least twice as high as the maximum frequency of interest (Nyquist theorem). The highest frequency analyzed in Phase II corresponds to the lower edge of the energy region of interest (ROI) in the tritium spectrum analysis. Due to the low expected count rate, the ROI is chosen to be as large as possible, while still allowing to control all systematic parameters. The resulting trade-off minimum kinetic energy is approximately 16.2 keV, which is 2.4 keV below the endpoint. The equivalent frequency range in the Q300 trap configuration is 25.81 GHz (~ 1.2 keV above the T_2 spectral endpoint) to 25.99 GHz (~ 2.4 keV below the T_2 spectral endpoint). It is not practical to sample the analog signal with ≥ 52 GS/s, especially as the entire frequency range below 25.2 GHz (frequency of the $^{83\text{m}}\text{Kr}$ M-lines) is not of interest. Therefore the signal is bandpass-filtered to keep the ROI clear from image noise and mixed down with a local oscillator signal (LO) with a frequency of 24.5 GHz. In this process, the two signals inputted to the mixer (waveguide and local oscillator signal) are multiplied. The result is a new signal at the sum and difference of the two original frequencies. This shifts the tritium frequency ROI (f_{ROI}) down to the range $1313 \text{ MHz} \leq f_{\text{ROI}} \leq 1488 \text{ MHz}$ (the signal at the summed frequency is discarded). In this work, CRES signal frequencies are often denoted as intermediate frequencies (IF) after analog down-conversion.

5.1.2 Matching to the digitizer

Before the mixed signal can be digitized, its total power has to be adjusted such that the signal amplitudes fit in the input voltage range of the 8-bit digitizer (0.5 Vpp) but also exceed the least significant bit (≥ 2 mV in the 8-bit system used here). In addition, a low pass filter needs to be applied to prevent noise aliasing in the digitization process. Without a filter, noise at frequencies higher than half the sampling rate would appear at $f_{\text{alias}} = |N \cdot f_{\text{sampling}} - f_{\text{noise}}|$, where N is the closest integer multiple of the sampling rate f_{sampling} . The required signal preparation is implemented in a *signal conditioning box* as shown in figure 5.2.

The conditioning box has a total of three inputs, which can be connected to the internal circuit via two switches. On one of the inputs, a test tone can be superimposed with the noise of a calibration noise source. RF band 1 and RF band 2 serve as inputs for the down-converted waveguide signal. Each of them could be fed the signal, mixed with a different local oscillator frequency. However, in Phase II only RF band 1 with the signal mixed-down by 24.5 GHz is used. The test tone input allows to easily generate a flexible test signal for performance measurements of both the conditioning box and the digitizer. The conditioning box contains a noise source that is added to the test tone via a coupler. The noise power profile as measured at the box output is shown in figure 5.3a.

5.1. ANALOG SIGNAL PROCESSING

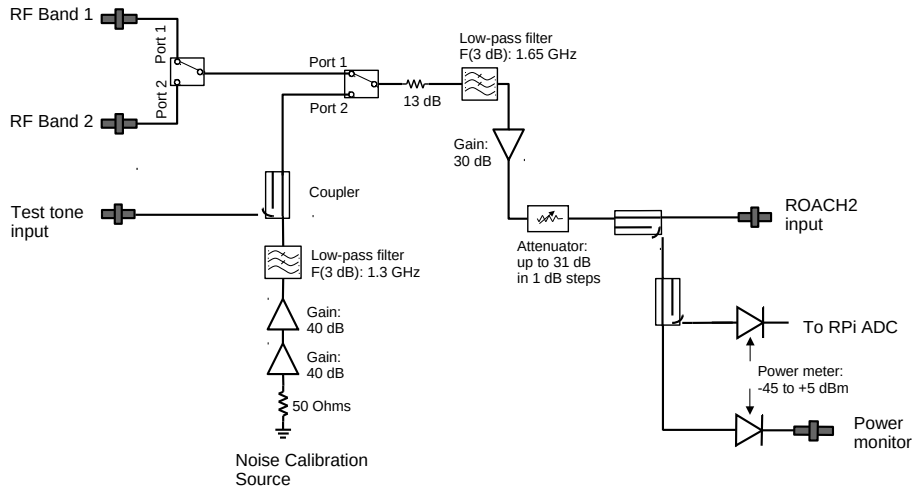
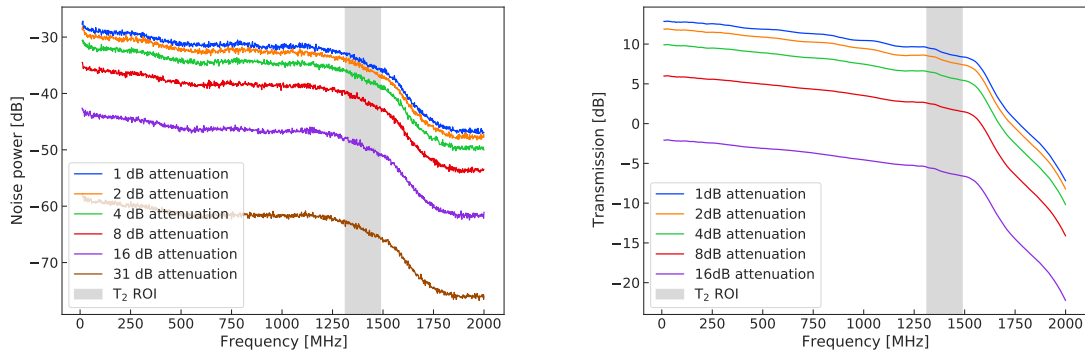


Figure 5.2: The conditioning box filters the signal to suppress image noise in the digitizer bandwidth. It further matches the signal power to the digitizer input requirements. The box is controlled via a RaspberryPi [103] and all switches as well as the step attenuator can be set remotely. Adapted image from private communication with Jonathan Tedeschi.



(a) Frequency profile of the calibration noise source. The profiles were recorded at different attenuator settings by a 100 kHz RBW spectrum analyzer.

(b) Transmission function of the RF 1 input as measured using a network analyzer (Keysight Technologies PNA) for different attenuator settings.

Figure 5.3: Calibration noise power spectrum (a) and transmission function from RF band 1 (b) in the signal conditioning box. Measurements performed by Jonathan Tedeschi.

All inputs are connected to a 30-dB amplifier followed by a 1-dB step attenuator. This allows for a flexible amplification over a 30-dB range. The transmission function of the low-pass filter and the amplification stage is shown in figure 5.3b. The low-pass filter cut-off frequency is 1.65 GHz. Noise above this frequency is suppressed, while the signal below is mostly unaffected. The filter shape is of course not a perfect Heaviside function with respect to frequency and all transmitted noise power above the cut-off frequency

will leak into the frequency region of interest via aliasing. Figure 5.4 shows the resulting image noise that one would expect from sampling a uniform noise background that was processed by the conditioning box, with a sampling rate of 3.2 GS/s.

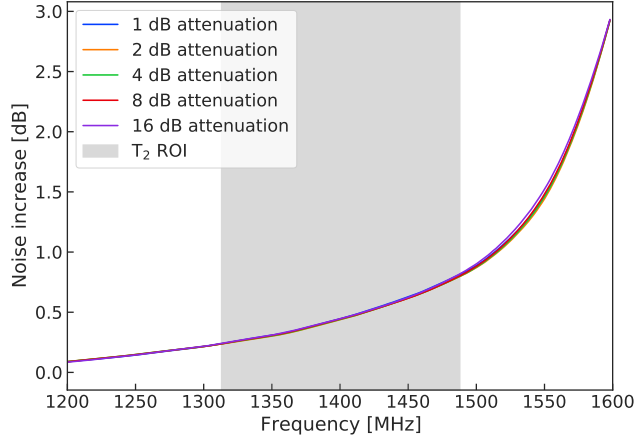


Figure 5.4: Image noise after 3.2 GS/s digitization calculated via equation (5.1) using the shape of the low-pass filter transmission (figure 5.3b) and assuming white noise on the RF band 1 input. The image noise increase is independent of the attenuator level.

The SNR will decrease inversely to the rise of noise power towards high frequencies

$$\text{Image noise}(f) = \text{Gain}(f_{\text{alias}}) \cdot \text{Noise}(f_{\text{alias}}). \quad (5.1)$$

The ROI of the Phase II tritium spectrum in the Q300 trap falls between 1313 MHz and 1488 MHz. Over this range, the SNR drops by 15% due to image noise. This has a significant impact on the frequency dependence of the detection efficiency and consequently on the shape of the tritium spectrum. The RF system design of a future CRES experiment should avoid the presence of image noise in the analysis range. For the prototype experiment Phase II, however, the goal was to develop a system capable of recording events from the T_2 spectrum as well as the K-, M- and L- lines of the $^{83\text{m}}\text{Kr}$ spectrum.

5.2 The ROACH2

The digitization and first stage digital signal processing of the intermediate frequency in Phase II is done using a ROACH2 (Reconfigurable Open-Architecture Computing Hardware version 2 [104]) platform. The platform was developed by CASPER (Collaboration for Astronomical Signal Processing and Electronics Research) and hosts a Xilinx Virtex 6 Field Programmable Gate Array (FPGA) and a PowerPC 440EPx processor. The PowerPC runs a Linux operating system and provides an interface over 1-Gbit Ethernet through which the board and its FPGA can be controlled. The ROACH2 is equipped with two analog-to-digital converter boards (ADC) [105] built around the EV8AQ160 ADC by *e2V*¹. Each ADC board has the capability of sampling the IF with up to 5 GS/s at a bit depth of 8 bit. The ROACH2 streams the processed data via a high-speed network board that features four SFP+ 10-Gbit Ethernet ports.

In the Phase II application, the purpose of the digital signal processing on the ROACH2 is to further down-convert the IF to a 100 MHz wide band centered on a configurable central frequency. As the CRES signal is only visible in the frequency-domain, the signal processing is split into a time-domain and a frequency-domain path. The Fourier transform of the sampled, down-converted signal is computed and both the time- and frequency-domain data are streamed via the 10-Gbit Ethernet connections to a server, where a software trigger searches the frequency-domain data for CRES-like signals. A description of the gateway can be found in [106]. It serves three main purposes: Global status and control, analog-to-digital conversion, and digital signal processing. Field registers with 32-bit depths provide monitoring and control possibilities. All registers are accessible within Python through the installed libraries on the PowerPC.

5.2.1 Analog to digital conversion

The ASIAA ADC1x5000-8 board [107] consists of four cores that each sample the IF with a resolution depth of 8 bit at a quarter of the target sampling rate. The samples from each core are interleaved to compose the full rate sampled signal. A misalignment of the cores in gain, phase, or offset would result in a spurious spectrum with peaks at multiples of the single-core sampling rates. The ADC-cores can be aligned via control registers. The properties and performance of the ASIAA ADC1 are described in [108]. The main specifications are listed in table 5.1.

¹<http://www.e2v.com>

| | |
|---------------------------------|----------------|
| Vertical resolution | 8 bit |
| Maximum sampling rate | 5 GS/s |
| Full power input bandwidth | 2 GHz |
| Analog input V_{pp} | 500 mV |
| Effective number of bits (ENOB) | 7.1 at 620 MHz |

Table 5.1: Specifications of the quad-core ADC.

In the configuration used in Phase II, only one of the two boards samples the IF signal with a rate of 3.2 GS/s (the other ADC card is unused). The power spectrum of the calibration noise source, after sampling with the ROACH2, is depicted in figure 5.5. It shows that the total gain of the conditioning box and the ADC drops by 12 dB over the 1.6 GHz bandwidth.

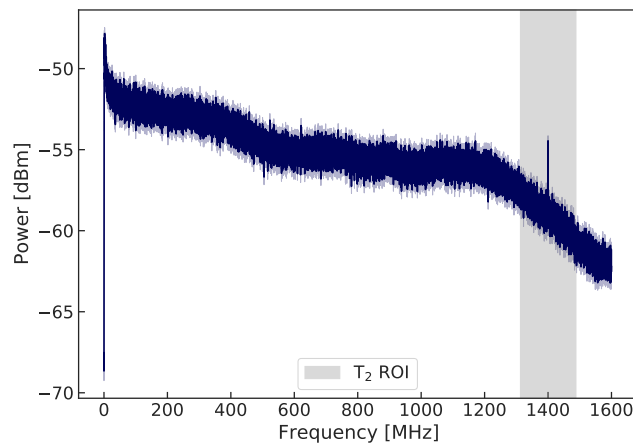


Figure 5.5: Power spectrum of the calibration noise (figure 5.3a) superimposed with a test tone at 1.4 GHz. The spectrum is averaged over 100 spectra that were recorded with the ROACH2 ADC. At an attenuation setting of 18 dB, the spectrum integrates to a total power of -6 dBm.

5.2.2 Multi-channel implementation

The FPGA is programmed by the execution of a BORPH file [109]. An overview of the digital processing stages that are implemented on the FPGA is shown in figure 5.6. The digitized samples from the ADC are de-multiplexed by a factor of 16: With every clock cycle of the FPGA board clock (200 MHz), 16 samples are transferred to the FPGA. To be able to record all three krypton lines simultaneously, the digital processing is implemented in three channels, labeled a , b , c . Each channel processes the same data

5.2. THE ROACH2

but can be configured to down-convert a different frequency band. The signal is mixed with a tunable complex local oscillator signal that is completely independent of the other channels. The oscillator frequency can be set in steps of 3.125 MHz. After the mixing stage, a 127th order finite impulse response (FIR) bandpass filter is applied. Only every 16th filter output is computed which down-samples the signal by a factor of 16 resulting in a bandwidth after mixing of 100 MHz (200 MS/s). At this point the signal path divides into parallel time- and frequency-domain processing. On the frequency-domain path, the 8192-point complex-valued discrete Fourier transform (DFT) is computed using a biplex FFT architecture [110], resulting in a frequency resolution of approximately 24.4 kHz. Only the positive half of the frequency spectrum is kept as no additional information is contained in the negative half. At the same time, the time-domain signal is subject to a second down-conversion stage in which it is mixed, filtered, and down-sampled. The results are IQ (in-phase & quadrature) samples of the same data volume and bandwidth as the frequency-domain data. The frequency central to the final 100 MHz wide frequency band (central frequency) is now shifted to zero. During processing, the bit resolution was increased to allow for bit growth. In the end, the data in both domains is re-quantized to 8 bit. The data is written to the payload of UDP packets. Each channel streams its time- and frequency-packets via its own 10-Gbit Ethernet connection in alternating order.

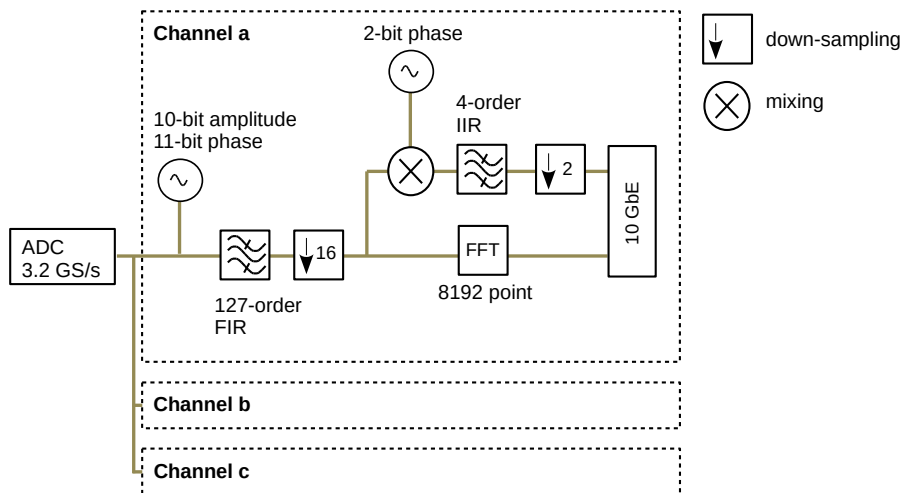


Figure 5.6: Block schematic of digital signal processing in a single ROACH2 channel. The channels *b* and *c* are identical to channel *a*. Adapted image from private communication with André Young.

5.2.3 ROACH2 Packets

The ROACH2 creates UDP packets and stores 4096 (complex) samples from either the time- or frequency-domain in one packet. The size of one packet amounts to 32 header bytes plus 8192 payload bytes. Hence, the total volume per domain in one channel is 200 MB/s (200781250 Bytes/s). The packet definition is given in table 5.2. The time- and frequency-domain packets can be distinguished by the *freq_not_time* variable in the packet header and matched via the *pkt_in_batch* counter id.

| Field name | Type | Description |
|---------------|--------|--|
| unix_time | uint32 | Unix time at packet creation (seconds since January 1 st 1970) |
| pkt_in_batch | uint32 | A counter that increments with each pair of time/frequency-packets. It counts to 390625 before wrapping to zero (in total every 16 seconds). |
| digital_id | uint32 | Identifies the digital channel (one of three). In the bitcode and control software, the channels are labeled <i>a</i> , <i>b</i> , and <i>c</i> , which maps to digital_id 0, 1, and 2 respectively. |
| if_id | uint32 | Identifies the IF signal input on the ROACH2, either 0 or 1. In the Phase II configuration, only one ADC (if0) is used. |
| freq_not_time | uint64 | If 1, the packet contains frequency-domain data, if 0, it contains time-domain data. |
| payload | uint8 | 8192 time- or frequency-domain samples. |

Table 5.2: ROACH2 packet content.

5.2.4 Channel configurations

The multi-channel implementation enables the ROACH2 to record data from different frequency ranges simultaneously. This offers many possible applications:

- In a trap configuration with a high axial frequency, the separation between main carrier frequency and sidebands could be larger than the (half-)width of a single channel (100 MHz). Using one channel to record the main carriers and the other two channels to record the sidebands would allow to record all visible bands and use their simultaneous appearance to reconstruct characteristic event properties. Although this application was the intended use case for the multi-channel implementation, it is not of interest here because no sidebands were detected in the deep harmonic trap configurations ($\Delta B \sim 1$ mT).
- The energies of the peaks with the highest branching ratios in the $^{83\text{m}}\text{Kr}$ spectrum are 17.8 keV (K-line), 30.4 keV (L-lines) and 31.9 keV (M-lines). In the Q300 trap configuration, the intermediate frequency positions of electrons originating from these lines are 1408 MHz, 806 MHz and 737 MHz. All three lines fall into the 1.6 GHz bandwidth of the ROACH2 and the three channels can be used to record all lines simultaneously.
- In a CRES experiment with a high exposure, observing a small energy range around the tritium endpoint would be sufficient to look for the neutrino mass and only a single channel would be required in a waveguide experiment. However, in Phase II, the electron count rate is low and the frequency ROI is much larger than the bandwidth of a single channel. Therefore, the channels are positioned adjacently such that they cover the full tritium ROI while maintaining a flat detection efficiency across all channels (figure 5.11).
- It can also be useful to record the same frequency range with separate channels, for example, to compare the trigger efficiency for different register settings on the ROACH2 or different software trigger configurations.

5.2.5 Channel configuration for tritium data taking

Before the Fourier transform is computed by the FPGA, the signal was band-pass filtered during the first down-conversion stage. The image noise rejection in filtering is never perfect and the signal-to-noise ratio therefore decreases towards the edges of a channel's frequency band as shown in figure 5.7. The event detection efficiency is expected to follow a qualitatively similar profile across a channel as the SNR. Its exact dependence on frequency relative to the central frequency is relevant for selecting the channel positioning for tritium data taking. To investigate the relative detection efficiency in a channel,

K-line electrons were recorded using all three channels simultaneously. The central frequencies of two channels were scanned from 50 MHz below the 17.8 keV peak position to 50 MHz above. The third channel was kept centered on the peak and served as a reference (figure 5.8).

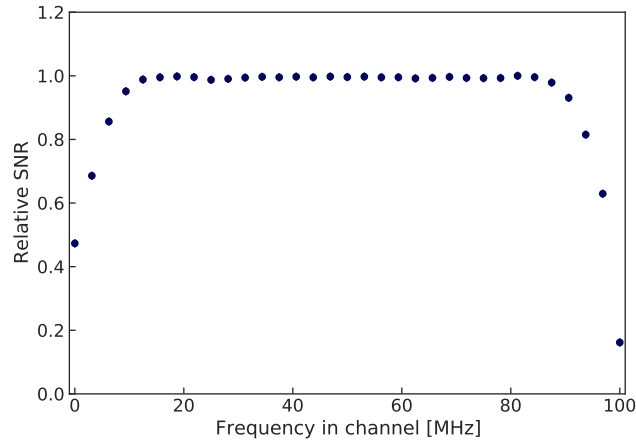


Figure 5.7: Measured relative SNR of an injected RF tone vs. tone position in a ROACH2 channel. Here, SNR is defined by the ratio of the height of an injected tone to the averaged noise background at the tone frequency.

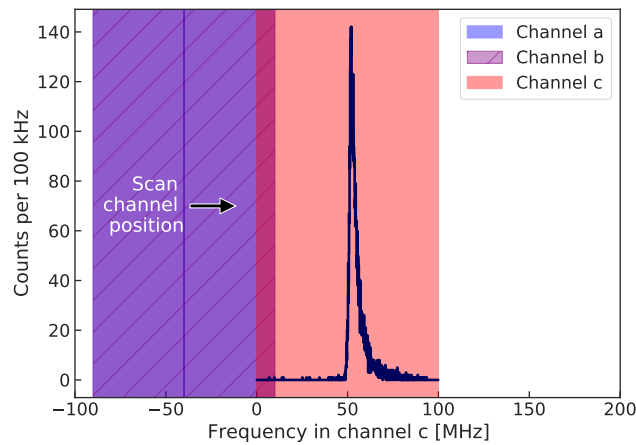


Figure 5.8: To record the trigger efficiency as a function of frequency in a channel, the central frequencies of two channels are scanned in 3.125 MHz steps from 50 MHz below to 50 MHz above the frequency position (IF) of the $^{83\text{m}}\text{Kr}$ K-line. Dividing the rate of detected events in a moving channel by the rate in the stationary reference channel (c) yields the relative detection efficiency (5.2).

5.2. THE ROACH2

For the count rate to be indicative of the efficiency at the peak position in a channel, only events close to the fitted frequency position f_0 of the 0th order peak (using the scattering model (4.21)) are counted: $|f - f_0| \leq 1$ MHz. Dividing the event rates in the two moving channels by the rate in the reference channel yields the detection efficiency dependence on the frequency position relative to the central frequency

$$\epsilon_{\text{channel}}(f) = \frac{N_{\text{scanned channel}}(f)}{N_{\text{reference channel}}(f)}, \quad (5.2)$$

where $\epsilon_{\text{channel}}$ denotes the relative channel efficiency and f ranges from 0 MHz to 100 MHz. The result is shown in figure 5.9. The channel efficiency can be fitted with a two-sided Butterworth [111] filter function

$$\text{Gain}(f) = A \cdot \frac{1}{\sqrt{1 + \left(\frac{f}{f_{\text{cut},1}}\right)^{2p_1}}} \cdot \frac{1}{\sqrt{1 + \left(\frac{f}{f_{\text{cut},2}}\right)^{2p_2}}}. \quad (5.3)$$

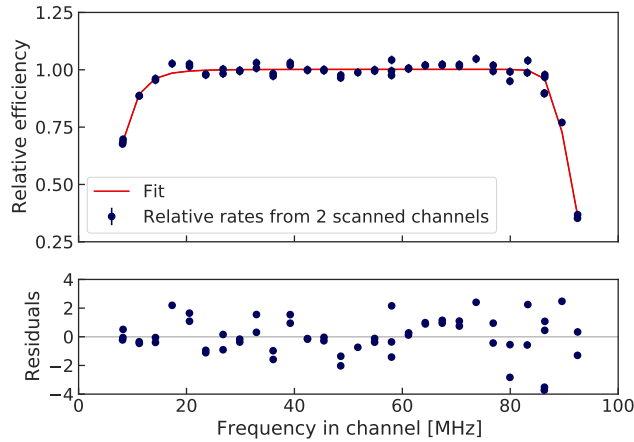


Figure 5.9: Detection efficiency vs. frequency in the ROACH2 channels. The recorded relative rates are fitted with a filter function. The fitted parameters are $A = 1.001 \pm 0.002$, $f_{\text{cut},1} = (89.8 \pm 0.1)$ MHz, $f_{\text{cut},2} = (8.4 \pm 0.1)$ MHz, $p_1 = 31.4 \pm 1.3$ and $p_2 = -2.4 \pm 0.2$. The asymmetric roll-offs originate from events being artificially shortened at the upper edge of the frequency band.

When multiple channels are used to record a spectrum, an efficiency drop between adjacent channels can in principle be included in the spectrum analysis. But if the central frequency configuration is chosen such that the channels overlap (figure 5.10a) by a margin that keeps the maximum efficiency decrease insignificant compared to the statistical error in the recorded spectrum, no correction of the count rates across a channel boundary is required. Events recorded beyond the intercept frequency are simply ignored. With the filter function from figure 5.9, the maximum efficiency loss for any given frequency overlap can be calculated (figure 5.10). Note that the size of

the frequency overlap takes discrete values, as the central frequencies can only be set in steps of 3.125 MHz.

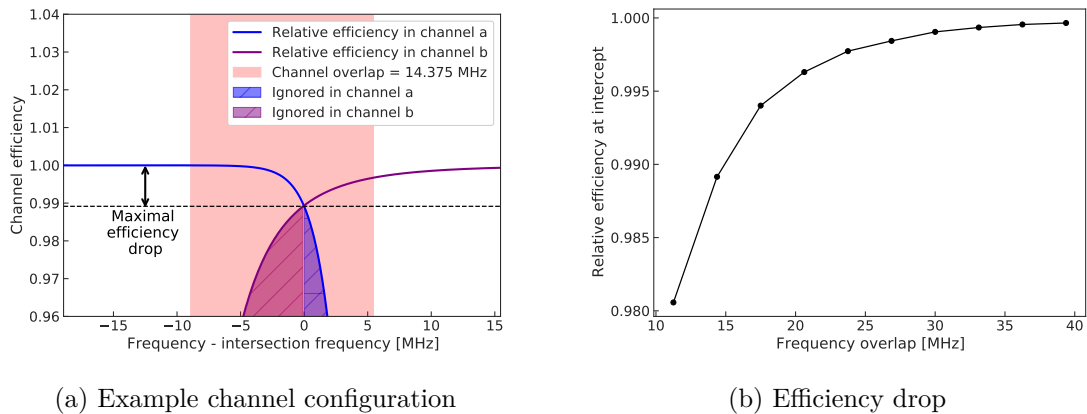


Figure 5.10: Efficiency drop between two channels. When channels are placed adjacently in frequency such that their frequency bands overlap, the channel efficiencies intersect. The magnitude of the efficiency at the intersection point depends on the size of the overlapping frequency region.

The initially expected total number of tritium events in Phase II was ≤ 10000 counts after 100 days of data taking. In a binned spectrum analysis, the statistical error would amount to at least 1 % if all events were contained in a single wide bin. If the efficiency drop in the frequency overlap is similarly small, it is negligible in the analysis of the tritium data. Figure 5.11 shows the channel configuration that the Phase II tritium data was recorded with.

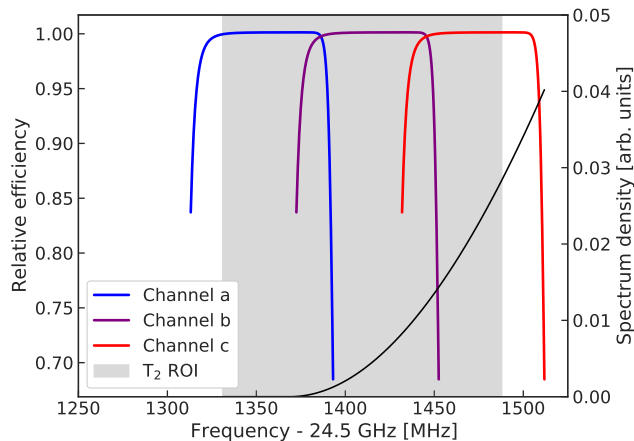


Figure 5.11: Channel configuration used for tritium data taking. The three ROACH2 channels are placed adjacently with a frequency overlap of 20.625 MHz resulting in a maximum efficiency drop between channels of 0.4 %. Together, they span across the full tritium ROI.

6

Triggering on CRES events

The ROACH2 provides real-time streams of time- and frequency-domain data. Via 10-Gbit SFP+ cables, 100 MS/s are transmitted per channel and per domain to a server. A triggering system has to inspect the incoming data and only save the content of the time-domain packets if they potentially contain CRES events. As events can occur at any time, the data acquisition should not suffer from any dead time during which the system would miss a new event start. To this end, the Project 8 collaboration has developed the real-time triggering package Psyllid [112]. Psyllid extracts the content of each ROACH2 packet and inspects the frequency-domain records (packet payloads) for CRES event candidates. Once a candidate is found, all records containing part of the presumed event are collected to form one acquisition and are written to disk. After fully introducing Psyllid in section 6.1, the trigger algorithm is described on a high level in section 6.2. In section 6.3 the implementation of the trigger logic in Psyllid is laid out in detail and in section 6.4 its performance as CRES trigger is demonstrated, followed by the optimization of the trigger parameters for Phase II.

6.1 Psyllid: a data acquisition package for Project 8

Psyllid is a data acquisition package written in C++. It was developed for the purpose of receiving UDP packets via a port, analyzing them in real-time and writing data to Egg-files [113], a file format that is based on the industry standard Hierarchical Data Format, version 5 [114]. The Psyllid architecture was designed to:

1. be modular and configurable at runtime,
2. safely run the signal processing steps in parallel, and
3. integrate into an existing Dripline [115] controlled system.

6.1.1 Midge

Psyllid is based on the Midge [116] framework. The working principle behind Midge are data streams that are published, read, and processed by different kinds of operators:

- **Publishers:** Operators that only publish data streams
- **Transformers:** Operators that consume and publish data streams
- **Consumers:** Operators that only consume a data stream, but do not publish an output stream

In principle, operators can possess an arbitrary number of input and output streams. The only limitation is that there can only be one publisher per data stream. All operators can be connected via data streams in almost arbitrary ways, where each operator takes action on the incoming data as soon as it appears on the output of the upstream operator. Upon initialization of a Midge instance, the network of operators is established and connected according to a pre-configured structure.

6.1.2 Psyllid control

Psyllid consists of two main layers: a control and a DAQ layer. The DAQ layer is a Midge object holding a set of nodes (implementations of Midge operators) that are defined in Psyllid. The control layer serves the purpose of managing the DAQ layer and of allowing interaction with a user by the exchange of Dripline [115, 117] requests via a rabbitMQ broker [118, 119]. It consists (among others) of the following key components (also see figure 6.1):

- **Run-server:** An instance of the run-server class is created by the Psyllid executable. After initialization the *execute* method of the run-server instance is called and the Psyllid executable waits for its return before exiting. The run-server creates an instance of the DAQ-control, the stream-manager and the request-receiver classes and calls their respective *execute* methods on separate threads. Before returning to the Psyllid executable, the run-server waits for the started threads to join.
- **Request-receiver:** The request-receiver's purpose is to receive Dripline requests and forward them to their destination (usually DAQ-control or one of the DAQ nodes). It therefore holds forwarding routes for all available set, command, and get requests.
- **Stream-manager:** With its initialization, the stream-manager is given the DAQ node configuration. The stream-manager creates a Midge instance for every set of connected nodes, followed by a node-building and a node-binding instance for every node in the set. A node consists of an instance of one of the DAQ classes and it is

6.1. PSYLLID: A DATA ACQUISITION PACKAGE FOR PROJECT 8

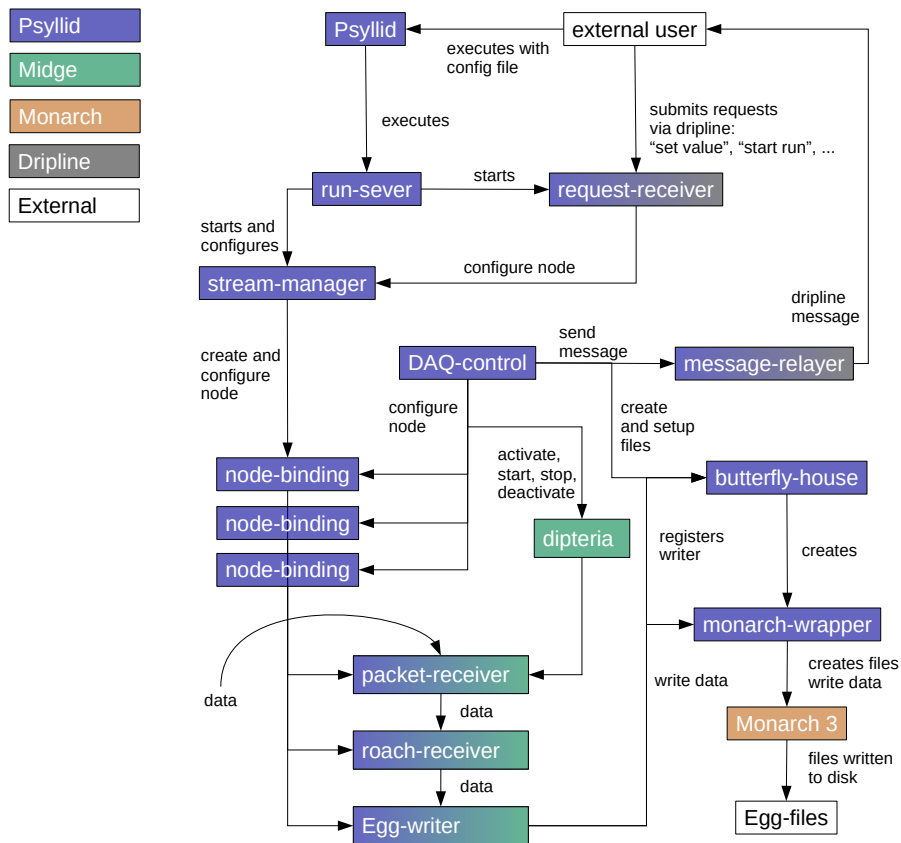


Figure 6.1: Setup and interactions of instances in Psyllid [112].

created and configured by its respective node-building instance. The node-binding instances allow to apply and dump node configurations while the nodes are actively running. The stream-manager further configures the data streams between the Midge operators.

- **DAQ-control:** DAQ-control is the key manager of data taking and controls the states of the DAQ layer. In each state, the DAQ nodes take different actions. The available states are:
 1. Deactivated (ground state): Midge is not running.
 2. Activating: A request to activate the DAQ has been received. The stream-manager will start a Midge instance and equip it with the configured DAQ nodes. The DAQ nodes prepare for running (data taking).

3. Activated: Midge is running but remains in a paused state. The DAQ nodes are processing data packets without actually taking data. DAQ-control is now ready to receive a start-run request.
4. Running: The payload of a start-run request contains the run duration and the filename of the Egg-file that will be created and filled with data once the run starts. A run is defined as a period of continuous (triggered) data taking. Upon receiving a start-run request, DAQ-control un-pauses Midge, which tells the DAQ nodes to start taking data. In the meantime, DAQ-control sleeps for the duration of the run. Once the sleep time limit has expired, Midge is re-paused but remains in the activated state, ready for more start-run requests.
5. Deactivating: A request to deactivate the DAQ has been received. Midge is told to stop. All DAQ nodes receive a stop command on their in-streams, which makes them clean up remaining buffered packets and then exit.
6. Canceled: Superior stopping instruction. When received, all activities in the DAQ nodes are stopped immediately and Psyllid will exit.
7. Do restart: A nonfatal error was experienced by Midge. The control moves to this state and restarts the DAQ. It will then go to the deactivated state.
8. Done: Operation has been completed and Psyllid is ready to exit.
9. Error: Psyllid is in a nonrecoverable error state and has to be restarted before any other action can be taken.

6.2 Trigger logic

The trigger logic employed in Phase II is implemented as a Midge stream consisting of multiple operators. Its components and configuration will be described in section 6.3. Here the functional principle of the trigger logic as a whole is described. Like any trigger, it serves two main goals:

1. Reduce the amount of recorded data.
2. Record as many events as possible.

In addition, it has to be able to (re-)start writing at any time. For each CRES event candidate, the trigger algorithm writes out the full 100 MHz-bandwidth time-domain data for the duration of the event. These acquisitions should contain the start points as well as the end points of the events (figure 3.1), which requires the trigger algorithm to recognize whether an event could be ongoing. While the event start is crucial, missing

6.2. TRIGGER LOGIC

an event end is less critical. Still, recording the full event lengths is favorable as track end event lengths yield important information about pressure, trap depth, and even gas composition. To record most event ends one could simply set a fixed acquisition length to be recorded after an event start was detected. But as the track and event lengths follow exponential distributions, the acquisition length would need to be set so that it is too long for most events while at the same time still being too short to record the full length of very long events. Therefore, the trigger algorithm was designed to provide the capability to produce acquisitions of flexible length.

Start-writing

The basis of the trigger algorithm is frequency mask triggering. The in-streaming frequency-domain data is compared to a pre-recorded frequency mask. As soon as the mask level is exceeded at any frequency, the algorithm checks if the start-writing condition is fulfilled. This condition can be dynamically configured at runtime to one of the following cases:

1. Detection of a single frequency bin with an SNR greater than a set threshold.
2. Detection of a frequency bin with an SNR greater than a set threshold followed by k additional high-SNR bins exceeding the same threshold within N frequency packets. Note that if $N > k$, the k bins don't have to be in consecutive packets.
3. Detection of a frequency bin with an SNR greater than a set threshold followed by k other high-SNR bins within N frequency packets, exceeding a less stringent threshold. It is less likely for an event to have two bins of very high SNR, rather than one very high and several relatively high-SNR bin. Therefore a lower second threshold increases the trigger efficiency compared to using the same threshold for all packets.

If the start-writing condition is fulfilled, Psyllid starts writing an acquisition. To prevent the event start from being missed if only a later part of the event matches the start-writing condition, Psyllid continuously buffers multiple records. The records in this pre-trigger buffer are all added to the triggered acquisition.

Continue-writing

Once Psyllid started writing, it will look out for new occurrences of high-SNR bins. To this end, it continues to compare incoming frequency spectra to a frequency mask. As there has already been at least one high-SNR bin before, the chance that a real event is ongoing is high and a more lenient threshold can be applied. Therefore the threshold and consequently the frequency mask used in the continue-writing conditions is lower than the mask that was used for the start-writing condition. Psyllid will end an acquisition as soon as the continue-writing condition is no longer fulfilled.

Threshold type: SNR or σ trigger mask

Before Psyllid can be used to trigger on incoming frequency packets, a frequency mask has to be defined. This is done by calculating the mean and the variance over all incoming frequency spectra for a set duration of time. The power that is added in this process by the presence of CRES events is insignificant because the event rate in Phase II is low and events are short. The frequency mask $m(f)$ is then calculated either by simply fitting a Spline curve to the averaged spectrum $\bar{p}(f)$ multiplied with an SNR threshold t_{SNR} (6.1), or by adding multiples of the square root of the noise power variance to the mean in order to produce a mask that is a defined number of standard deviations t_σ above the average power level in the frequency-domain:

$$\begin{aligned} m(f) &= t_{\text{SNR}} \cdot \bar{p}(f) \\ m(f) &= \bar{p}(f) + t_\sigma \cdot \sigma_p(f) \end{aligned} \tag{6.1}$$

In case of exponentially distributed noise power, an SNR threshold t_{SNR} is equivalent to a σ threshold of $t_\sigma = t_{\text{SNR}} - 1$. The noise power distribution in Phase II is exponential and the choice of threshold type is therefore irrelevant (figure 6.2).

An example of how the trigger logic applies to a typical CRES event is shown in figure 6.3. Once the start-writing condition is fulfilled, the pre-trigger buffer is written out, followed by the data from all records until the continue-writing condition is no longer matched by the incoming frequency spectra. All color-shaded areas are written to disk.

6.3. TRIGGER IMPLEMENTATION IN PSYLLID

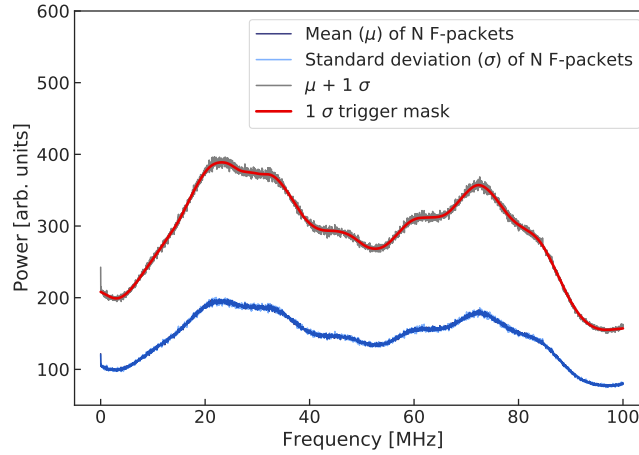


Figure 6.2: In σ -threshold mode, the frequency mask (red) is a third order Spline fit to the sum of a multiple of the standard deviation (bright blue) and the mean (dark blue) of pre-recorded frequency-domain data.

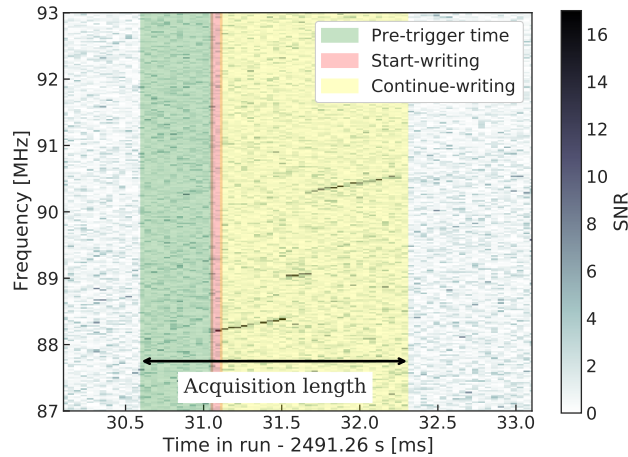


Figure 6.3: A triggered acquisition should extend over the entire length of an event. For this purpose, the content of in-streaming ROACH2 packets is constantly buffered. Once an event is detected (red), a pre-trigger time (green) is added to the acquisition. The acquisition continues (yellow) until no new activity is detected within a set time limit.

6.3 Trigger implementation in Psyllid

The trigger logic described in section 6.2 is implemented as a Midge data stream that is defined and controlled within Psyllid. To this end, Psyllid's stream-manager starts and configures the requested DAQ nodes and hands them over to the initialized Midge object. The nodes run on separate threads and the stream in- and outputs are connected by dipteria according to the configuration provided by the stream-manager.

Nodes are equipped with at least one in-stream and/or one out-stream buffer. The out-stream buffer of a producer is at the same time the in-stream buffer of its downstream consumer or transformer. The sizes of these buffers are defined at initialization and are part of the stream configuration. After initialization, each node starts a while loop. The first action in each consumer or transformer node's while loop is a check of the status (start, run, stop, exit) set in the next stream object in its in-stream buffer. Depending on this status, the node takes different actions. Each node passes the status on to the downstream nodes by setting the same status in the stream objects that it places into its out-stream buffer. As a node cannot proceed in its while loop before receiving a new in-stream object, the nodes are synchronized by the speed at which the first producer in the stream sequence produces its out-stream objects. This way it is ensured that the data is processed in sequence and every node waits until the upstream nodes have finished their tasks. The data content of a stream object can be of different types. The two main implemented types are roach-packets and trigger-flags:

- **Roach-packet:** The content of the ROACH2 UDP packets is stored in instances of Psyllid's freq-data and time-data classes. They both inherit from the roach-packet class that contains the functionality to read the header information as well as the payload content (record) from a UDP packet. The payload is written to a two-component 8-bit signed integer array with 4096 entries in each component, representing the real and imaginary part of the sampled voltages (IQ sampling).
- **Trigger-flag:** A trigger-flag object has three member variables:
 - *flag* (boolean): This variable is set by the node that produces the trigger-flag instance. For example, a simple frequency mask trigger node searches a frequency spectrum and finds one bin that exceeds the trigger mask. The node will create a trigger-flag, set the *flag* to *true* and put it in its out-stream buffer.
 - *id* (unsigned int64): The *id* serves as a unique (unique within 16 seconds) identifier of a packet during the whole analysis process in Psyllid. It corresponds to the ID of the UDP packet (`packet_in_batch`).
 - *high-threshold* (boolean): A frequency mask trigger node comparing the data to more than one frequency mask sets this variable to *true* if the highest threshold mask was exceeded.

6.3.1 Node configuration for triggered data taking

To record CRES events by applying the trigger logic from section 6.2, Psyllid has to complete five basic tasks:

1. Receive UDP packets from the ROACH2 and read their content.
2. Distribute the packet content depending on whether it contains frequency- or time-domain data.
3. Apply a frequency mask trigger to the frequency-domain data.
4. Decide which packets should be written to disk.
5. Write the time-domain data to Egg-files.

To accomplish these tasks the Midge stream is configured to use five nodes: A packet receiver (publisher), a tf-roach receiver (transformer for receiving time- and frequency-domain ROACH2 packets), a frequency mask trigger (transformer), an event builder (transformer), and an Egg-file writer (consumer). The stream connections of the five nodes are shown in figure 6.4.

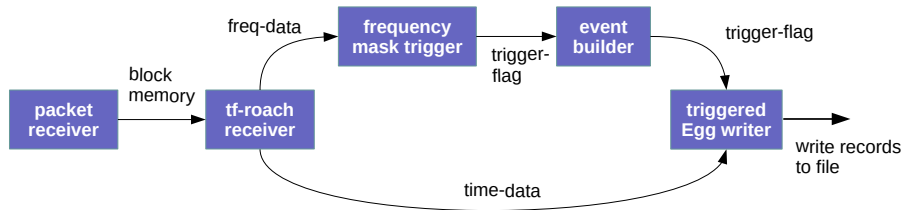


Figure 6.4: Node configuration in triggering mode. Records from the time-domain packets are buffered while their frequency-domain equivalent is inspected for high-power bins by the frequency mask trigger node (FMT). The event builder node uses the information forwarded by the FMT in the form of trigger-flags to identify the record IDs composing the next acquisition. The writer node then writes the time-domain records sorted as acquisitions to an Egg-file using the Monarch3 [113] library.

The tf-roach receiver has two out-streams, one with time-data and one with freq-data objects. The time-data out-stream is connected to one of the two in-streams of the Egg-file writer. The received stream object will remain in the buffer until the Egg-file writer receives a trigger-flag on its other in-stream with a matching *id*. Depending on whether the boolean *flag* variable of this trigger-flag is *true* or *false*, the content of the time-data object will be written to a file.

Data unpacking

Psyllid has two types of packet receivers: the socket receiver and the fast packet receiver. The socket receiver gets the UDP packets via the standard socket interface. Since this is a simple high-level process, it is very useful for testing purposes or for receiving packets from a packet simulator that produces UDP packets at a slower rate. However, receiving packets over a port is too slow for real-time triggering, because under Linux one system call is required to capture one packet. The fast packet receiver on the other hand gets the packets via the fast packet-acquisition interface that uses `PACKET_MMAP` [120]. `PACKET_MMAP` provides a circular buffer of configurable size that is shared between the kernel and the user. Instead of issuing a system call, the user can read packets from the shared buffer (potentially even multiple packets at once). Both receivers are Midge producers with a single out-stream and they both write the packet content as raw blocks of memory to their out-stream.

The `tf-roach` receiver is a transformer with one in-stream and two out-streams. It expects to receive memory-block data as produced by a packet receiver via its in-stream and processes the packet header information. It creates `freq-data` or `time-data` objects and places them in the two out-stream buffers sorted by domain.

Frequency mask trigger

The frequency mask trigger (FMT) is a transformer node with one in-stream and one out-stream. The in-streaming data type is `freq-data` and the output data are `trigger-flags`. The FMT can run in two different modes.

- **Add-to-mask:** In the add-to-mask mode, the FMT calculates the mean and the variance of the incoming power spectra over time. In order to do so, it takes the incoming complex data array from each packet and sums the absolute square and the square of the absolute square separately for each frequency bin. Once a pre-configured number of records N has been processed, the sums are divided by N and the trigger mask is calculated (see section 6.2). Depending on which start-writing condition will be applied, the FMT either calculates one or two trigger masks of different levels.
- **Apply-threshold:** In this mode the FMT compares the power in the in-streaming frequency packets to the saved frequency masks. It loops over the `freq-data` array, converts it to power by calculating the absolute square of each bin and compares it to its stored trigger masks. If the FMT is using two masks with different thresholds, it first checks whether the power in a bin exceeds the higher-threshold mask. If this comparison is true for any bin the FMT sets both the *flag* and the *high-threshold* variable in the out-stream `trigger-flag` to *true* and breaks from the for loop over

6.3. TRIGGER IMPLEMENTATION IN PSYLLID

the remaining bins. Else, it compares the power in each bin to the lower-threshold mask. If only the lower mask level is exceeded, the *flag* is also set to *true* but the *high-threshold* remains *false* and the FMT does not exit from the loop over the remaining bins. This way, the downstream operator is able to tell from the trigger-flag whether the frequency packet resulted in a trigger at all, and if it did, whether the power exceeded the higher or only the lower of the two mask levels. However, there is no information passed on about which frequency bin exceeded which threshold.

Event building

The event builder is a transformer node with trigger-flags in both its in-stream and out-stream buffer. The event builder's task is to decide which records should be written to acquisitions. The start-writing, continue-writing, and the pre-trigger sizes are all determined here and are implemented as a finite-state machine. In every iteration of its while loop, it will get the next trigger-flag from the in-stream buffer and depending on the value of the *flag* variable and the *high-threshold* decide to switch to a new state or remain in its current state. The four states of the state machine are un-triggered, collecting-triggers, triggered, and skipping and the action taken in each state is described below.

The event builder has two internal buffers in which it can temporarily store trigger-flags. In the pre-triggered state, it moves every incoming trigger-flag to the pre-trigger buffer. If the pre-trigger buffer is full, the oldest trigger-flag is moved from the pre-trigger buffer to the out-stream buffer and its *flag* variable is set to *false*. The second buffer is the skip buffer. This buffer is filled with trigger-flags that arrive when the event builder is in the collecting-triggers or skipping state. Whenever the event builder leaves either of these states, the skip buffer is emptied and all trigger-flags are moved either to the pre-trigger buffer or to the out-stream buffer. The pre-trigger and skip buffer sizes are configurable dynamically at run time via the pre-trigger time and the skip-tolerance parameters respectively. Figure 6.5 shows an example of a sequence of incoming trigger-flags and indicates into which buffers they would be placed by the event builder. This depend on the event builder's internal states:

- **Un-triggered:** The event builder starts in the un-triggered state. In this state it moves all trigger-flags to the pre-trigger buffer (and to the out-stream buffer if the pre-trigger buffer is full) until one of them contains a *flag* and a *high-threshold* that are both *true*. Once such a trigger-flag arrives, the event builder checks if the start-writing condition requires one or more FMT-triggers¹. If the number of required FMT-triggers is 1 it switches to the triggered state right away. Otherwise,

¹An FMT-trigger is a trigger-flag produced by the FMT that contains a *flag* that is set to *true*

if a higher number of FMT-triggers is required, it switches to the collecting-triggers state instead.

- **Collecting-triggers:** In this state the pre-trigger buffer is untouched and all new trigger-flags are saved in the skip buffer. If an incoming *flag* is *true*, a counter variable is incremented. As soon as the counter exceeds the required number of FMT-triggers, the event builder switches to the triggered state. If not enough FMT-triggers arrive and the skip buffer is full, the event builder returns to the un-triggered state. Trigger-flags are moved from the pre-trigger buffer to the out-stream, turning their *flags* to *false*, until the pre-trigger buffer has enough space to fit all trigger-flags from the skip buffer. The skip buffer is then emptied and all trigger-flags are moved to the pre-trigger buffer. This way, the event builder is ready for a new trigger to arrive, with no packet *ids* missing in the out-stream sequence.
- **Triggered:** Whenever the event builder moves to this state, all buffers are emptied and the trigger-flags are moved to the out-stream with all *flags* set to *true*. A newly arriving trigger-flag that is flagged as *true* is written to the out-stream right away. If the *flag* of the next trigger-flag is *false*, the trigger-flag is moved to the skip buffer and the event builder switches to the skipping state.
- **Skipping:** Similar to the collecting-triggers state, the event builder fills the skip buffer with the incoming trigger-flags. As soon as a new FMT-trigger arrives, the event builder returns to the triggered state. Otherwise it switches back to the un-triggered state once the skip buffer is full. Before doing so, the skip buffer has to be emptied in a way that avoids causing any dead time, even if an acquisition has just been completed. In addition, the event builder should make sure that the number of records written before the detected event start is always as least as large as the size of the pre-trigger buffer. This can be achieved either by moving all trigger-flags in the skip buffer to the out-stream (labeled *true*), or by moving the content of the skip buffer to the pre-trigger buffer. If the skip buffer size is larger than the pre-trigger buffer, trigger-flags are moved to the out-stream until their remaining number no longer exceeds the size of the pre-trigger buffer. If the size of the skip buffer is smaller or equal to the size of the pre-trigger buffer, the trigger-flags could all fit into the pre-trigger buffer directly. However, to prevent that event ends are missed, all trigger-flags are written to the out-stream as *true* instead, leaving the pre-trigger buffer empty. This way, a renewed match of the start-writing condition would lead to an uninterrupted continuation of the acquisition.

6.3. TRIGGER IMPLEMENTATION IN PSYLLID

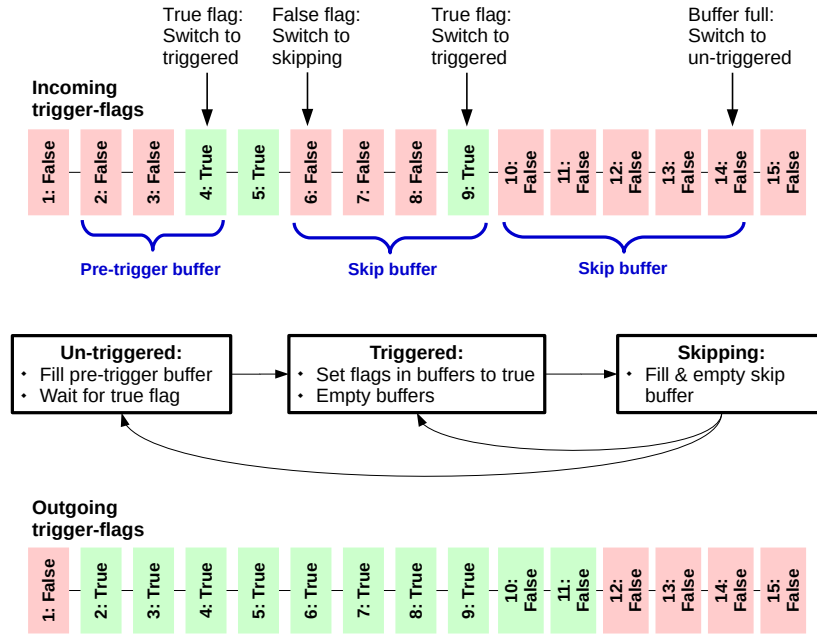


Figure 6.5: Example for trigger-flag processing in the event builder (without collecting-triggers state). While in the un-triggered state, in-streaming trigger-flags are constantly placed into the pre-trigger buffer. When the buffer is full and a new trigger-flag arrives, the first flag in the buffer is moved to the out-stream. Once a trigger-flag with the *flag-boolean* being *true* arrives (trigger-flag 4), the event builder switches to the triggered state. In this state all trigger-flags in the pre-trigger buffer (2 to 4) are labeled as *true* and are moved to the event builder’s out-stream. When the next incoming trigger-flag is labeled *false* (6), the event builder switches to the skipping state and starts filling the skip buffer. If a new *true flag* arrives (9) before the skip buffer is full, the event builder switches back to the triggered state. Once the skip buffer is full, trigger-flags are written out as *true* (10 and 11) until the remaining trigger-flags (12 to 14) fit into the pre-trigger buffer. All records corresponding to the trigger-flags that are moved to the out-stream as *true* (2 to 11) will be written to an acquisition by the triggered writer. The logic implemented in the event builder ensures that acquisitions always consist of a sequence of consecutive records.

Writers

Currently, Psyllid has two different writer nodes: a streaming and a triggered writer. Both are Midge consumers. The streaming writer only has one in-stream on which it receives the records that were contained in the time-domain data packets. All records are written to Egg-files. The triggered writer has two in-streams via which it receives time-data and trigger-flags. Both contain an *id* that originates from the ROACH2 time- and frequency-domain packets. The triggered writer writes the record content of the time-data objects to the Egg-file when it receives a trigger-flag with an *id* that has the

same value and a *flag* that is *true*. If the *flag* is *false* the time-data object is deleted without its content being written to disk. Both writers use the Monarch [113] library (via a Monarch wrapper) to create and write acquisitions to Egg-files.

6.3.2 Streaming mode

With Midge, highly flexible node configurations are possible. By combining different nodes in streams, different functionalities can be configured at run-time. The simplest data taking method is to not use triggering and write all packets to disk instead. For this acquisition mode the time-data out-stream of the tf-roach receiver is directly connected to the in-stream of the streaming writer, while the freq-data out-stream is connected to a terminator node where the data is deleted. This is necessary to prevent that the out-stream buffer of the tf-roach receiver is blocked. The streaming mode configuration is shown in figure 6.6.

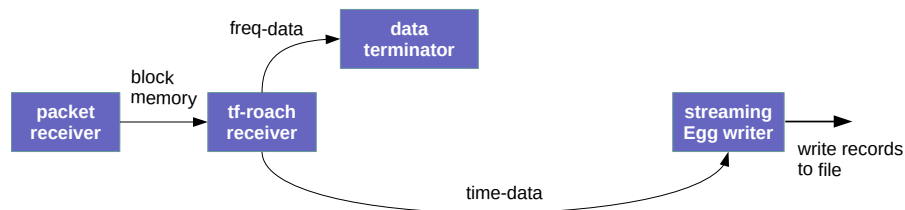


Figure 6.6: Node configuration in streaming mode. In this configuration, all time-domain records are written to Egg-files.

6.4 Trigger performance

Psyllid and Midge were designed to meet the following three performance goals as a CRES software trigger:

1. Reduce the amount of recorded data to a manageable level.
2. Record as many events as possible that can be reconstructed by the offline analysis.
3. Avoid suffering from any dead time during data taking.

In this section, it will first be shown that Psyllid fulfills the requirement of not suffering from dead time. In section 6.4.2 a trigger parameter optimization is presented with the goal of finding the best trade-off between data reduction and trigger efficiency.

6.4.1 Dead time

To verify that the Psyllid trigger algorithm really has no dead time, a threshold scan was conducted. To this end, a white noise² signal was processed by a ROACH2 to provide Psyllid with packets that contain only noise in both domains at the use-case rate (one packet per domain per $40.96 \mu\text{s}$). The DAQ nodes were configured as in figure 6.4. Psyllid's FMT was set to only use a single σ -threshold mask. The pre-trigger and skip buffer sizes were set to zero so that for each single FMT-trigger an acquisition containing one record would be written out. Then the FMT-trigger threshold was scanned from 5 to 17σ . The acquisition rate and the live fraction (fraction of records that were written to an Egg-file) is shown in figure 6.7. It can be seen that the live fraction converges to 1 for low thresholds, because the lower the threshold, the more likely it becomes that each record contains at least one noise fluctuation above the trigger mask. For high thresholds, the live fraction decreases exponentially to 0. This behavior is expected as the noise power follows an exponential distribution. The acquisition rate on the other hand first increases with rising threshold, as fewer consecutive records are merged into combined acquisitions, before decreasing again.

²White noise: A random signal having equal intensity at different frequencies, giving it a constant power spectral density.

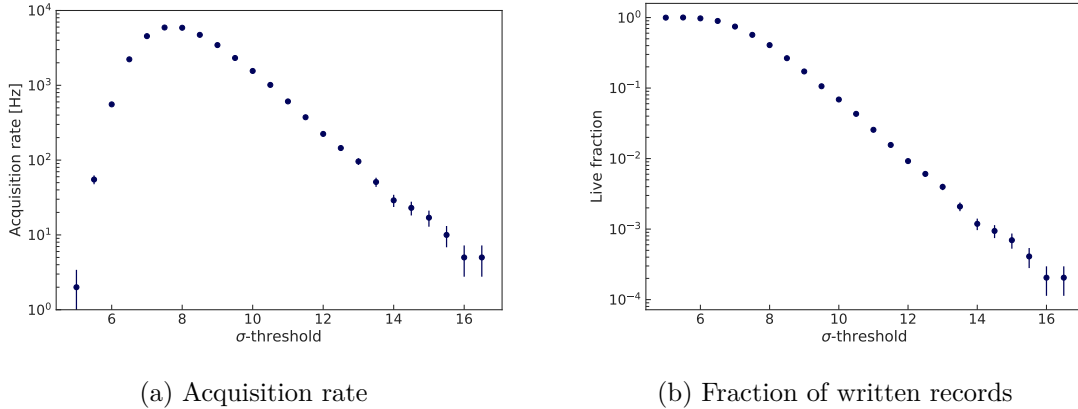


Figure 6.7: Acquisition rate and live fraction in a trigger threshold scan. White noise was injected into the ROACH2 input. Psyllid’s event builder node was configured with a pre-trigger and a skip buffer size of zero trigger-flags. This way, each single FMT-trigger (FMT trigger-flag with a *flag* equal to *true*) is directly forwarded to the triggered writer. The live fraction converges to 1 for low thresholds, which proves that Psyllid can trigger on each record without suffering from dead time. For high thresholds it converges to 0. The acquisition rate decreases towards low and high thresholds because consecutively written records are combined to single acquisitions.

6.4.2 Trigger optimization

From figure 6.7 it can be concluded that Psyllid is capable of both recording data continuously as well as reducing the recorded volume by 100 %. The goal is now to tune the configurable trigger parameters (two thresholds, two buffer sizes, and the number of required FMT-triggers) to find the best compromise between data reduction and detection efficiency.

Detection efficiency and run conditions

Detection efficiency is the combination of trigger and event reconstruction efficiency. The detected event rate is the product of the recorded event rate and the fraction of recorded events that can be reconstructed. It is not always the case that recording more events automatically results in a higher detection rate. If the trigger algorithm records more low SNR events than the reconstruction algorithm can find, further lowering the trigger threshold will not lead to more detected events. Therefore, the trigger optimization aims to provide the highest possible data reduction while minimizing the fraction of events that could have been reconstructed but were discarded by the trigger. This was partly an iterative process during which both triggering and reconstruction approached the achievable optimum for Phase II.

6.4. TRIGGER PERFORMANCE

To discriminate between CRES signals and noise fluctuations, the trigger has to make use of various event properties like event length and signal power. The distribution of these properties strongly depends on the run conditions. The most relevant conditions for the trigger are the SNR and thus the noise temperature, the gas pressure (track length), and the trap depth (number of tracks per event). When these conditions change, the optimal trigger configuration changes too, e.g.:

- Using a lower gas pressure leads to longer events. Long events are exposed to a higher risk of being split into several acquisitions, especially if their SNR is low. Milder time tolerances (buffer sizes) are therefore required, which increases the amount of data recorded per trigger. On the other hand, longer events are more likely to exceed the trigger threshold many times in their lifetime. This allows to configure the trigger logic to require a high number of FMT-triggers in the start-writing condition, which reduces the noise trigger rate.
- At high pressure, the events are very short and the chance of them overcoming the trigger threshold during their lifetime is lower. Thus, less strict trigger conditions must be set, which in turn increases the rate of noise triggers. This is compensated by setting a small pre-trigger time and skip-tolerance leading to very little recorded time per acquisition.

In the turn of the Phase II experiment, the run conditions were modified several times as the event reconstruction was continuously improving. A better reconstruction efficiency for short events changes the optimum pressure that maximizes the detected tritium event rate. Whenever the pressure is changed by an order of magnitude, a re-optimization of the trigger parameters is required.

Parallel trigger tests with $^{83\text{m}}\text{Kr}$ data

The performance of the Phase II trigger and the event reconstruction with respect to various event properties will be analyzed in section 8.3 by comparing recorded data to simulations. The goal here is the optimization of the detected event rate after triggering for the tritium running conditions in Phase II. The relative efficiency was studied by recording $^{83\text{m}}\text{Kr}$ K-line electrons with different trigger settings and comparing the event rate after reconstruction. Similarly to the channel efficiency test from section 5.2.4, all three ROACH2 channels were used to record krypton data in parallel. Again, one channel served as reference channel that was always configured with the same trigger settings, while the two other channels were used to test a new configuration in each run. This way, two configurations could be tested per parallel three-channel run. The reference channel is needed to ensure the result is not falsified by fluctuations of the source activity over time. To accumulate a sufficient number of events to achieve a statistical uncertainty

of less than 5 %, the duration of each run was set to one hour. Even after reducing the number of parameter combinations by carefully choosing the range over which each parameter is tested, the optimization still took three days in total.

Choosing trigger parameter settings

The trigger logic has a total of five configurable parameters:

1. Pre-trigger time: Sets the size of the pre-trigger buffer.
2. Skip-tolerance: Sets the size of the skip buffer.
3. N-triggers: The number of FMT-triggers required by the event builder to switch to triggered state.
4. High-threshold: Threshold of the trigger mask that has to be exceeded before the event builder leaves the un-triggered state and switches to the skipping or the triggered state.
5. Threshold: Threshold of the lower trigger mask that is sufficient to increment the counter in the event builder's collecting-triggers state and go back to the triggered state when in skipping state.

The pre-trigger time does not need to be scanned. It can simply be set according to the average event length, which depends on the gas pressure and the gas composition in the cell. When taking tritium data, the majority of the gas molecules are T₂, HT, ³He, and H₂. During ^{83m}Kr operation, the dominant gas species are (mostly) H₂, ³He, and Kr. The inelastic scattering cross section for electrons is much higher in krypton than in hydrogen gas. In order to match the track and event lengths in krypton and tritium operation, a different gas pressure has to be set depending on the gas composition: $\sim 1.6 \cdot 10^{-6}$ torr in krypton operation and $\sim 2.5 \cdot 10^{-6}$ torr in tritium operation (pressures measured by ion gauge, calibrated to hydrogen equivalent gases). The expected average event length in both operation modes is 0.65 ms (in the Q300 trap). The pre-trigger time is set such that 90 % of all event starts would still be contained in the acquisition window even if only the very last bin in an event triggered the acquisition. According to figure 6.8 this requires a pre-trigger time of 2 ms, taking into account that the event reconstruction applies a 0.5 ms start time cut.

6.4. TRIGGER PERFORMANCE

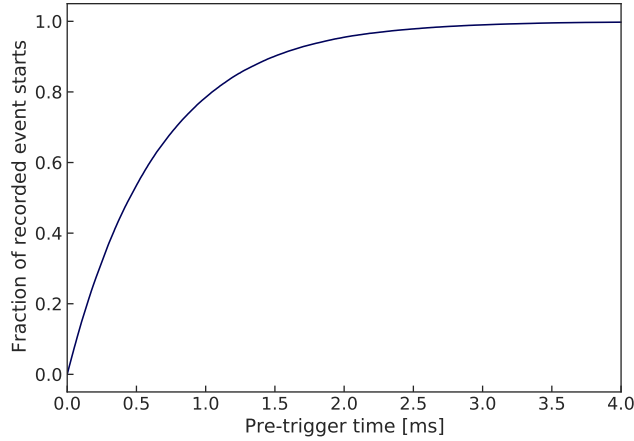


Figure 6.8: Simulated fraction of recorded events starts for an exponential event length distribution with a mean of 0.65 ms, assuming only the last time bins of the events trigger the acquisitions.

The skip-tolerance sets the skip buffer size, which determines how long the event builder waits for FMT-triggers to accumulate before starting an acquisition and how long the event power can stay below the FMT threshold before the acquisition is ended. It further determines how much time is written to the Egg-file after the last FMT-trigger occurred. As its optimum has so many dependencies, it is one of the most important parameters to be optimized in the scan. The tested range was limited to $0.5 \text{ ms} \leq \Delta t_{\text{skip}} \leq 2.5 \text{ ms}$, which roughly corresponds to the range of detected track lengths.

For the n-triggers parameter, only two settings were tested. Since most events are very short at the target pressures and are therefore not expected to exceed the trigger threshold very often in their lifetime, n-triggers was either set to 1 or 2. When the n-triggers parameter is 1, a single FMT-trigger is sufficient to start an acquisition. The additional data reduction achievable by using a second trigger (n-trigger = 2) can be seen in figure 6.9. It is based on a simulation of exponential white noise and shows the fraction of noise power fluctuations above 13σ (an example threshold that is arbitrarily chosen) that were not followed by a second fluctuation exceeding the threshold on the y-axis within the time limit on the x-axis. It therefore simulates the effect of combining a certain skip-tolerance with a lower FMT threshold on the noise trigger rate. Based on this plot, the decision was made to test a trigger threshold range between 10σ and 13σ .

To pre-select a range for the high-threshold parameter, one can take a look at the rate of occurrences of noise fluctuations that would exceed it. Figure 6.10 shows the FMT-trigger rate as a function of a set σ threshold for simulated noise. Based on the available disk space in Phase II and the target to take tritium data for 100 days, the goal was set to

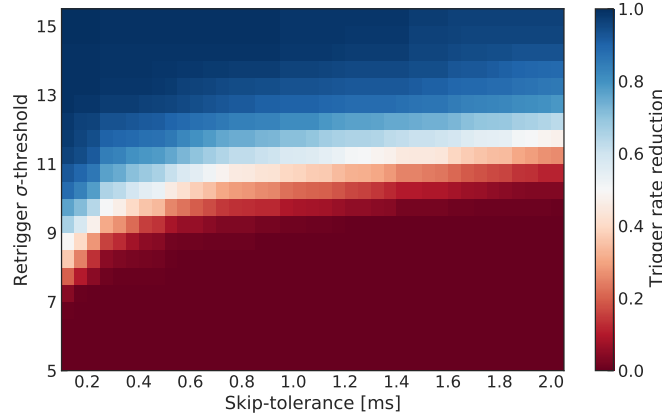


Figure 6.9: Simulated noise trigger rate reduction by requiring a 2nd FMT-trigger before starting an acquisition. The color encodes the fraction of acquisitions that are not recorded by requiring that a second trigger-flag from a re-trigger threshold σ occurs within the time limit set by the skip-tolerance.

achieve a recording rate that does not exceed 0.5 TB/day to 0.7 TB/day per channel. This corresponds to a data reduction of at least 95.9% to 97.1%. As the task is not to reduce the noise trigger rate, but instead to limit the filled disk volume per time, the threshold choices depend on the acquisition lengths. The assumption can be made that the majority of triggers are noise triggers with no re-triggers. Their acquisition lengths will be equal to the sum of the pre-trigger time plus the minimum time that is recorded after a trigger (skip buffer length). For a skip-tolerance of 0.5 ms, the resulting total acquisition length is 2.5 ms. To reach the data reduction goal of $\geq 96\%$, the trigger rate has to be smaller than 16 Hz. This leads to a lower limit on the high-threshold of about 14σ . For the trigger configurations with n-triggers = 2, even lower thresholds were tested as the re-trigger requirement provides additional data reduction.

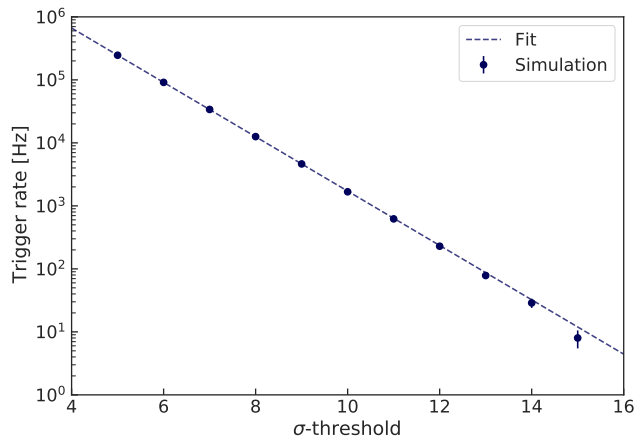


Figure 6.10: Simulated noise trigger rate vs. FMT threshold for n-triggers = 1.

6.4. TRIGGER PERFORMANCE

Data set

All data was taken in the Q300 trap configuration. The set gas pressure was $1.6 \cdot 10^{-6}$ torr. In total, 80 different configurations were tested. The reference trigger configuration originates from a previous optimization for a much lower pressure ($3 \cdot 10^{-7}$ torr). The data is split into two sets: n-triggers = 1 and n-triggers = 2. In both sets the high-threshold, re-trigger threshold and skip-tolerance were varied in all possible combinations, covering the previously described parameter space.

Analysis

Figure 6.11 shows the event rate over all runs in the reference channel. As the krypton source was several lifetimes old when this optimization study was performed, the total number of events accumulated to only ~ 1500 per one hour long run. The x-axis in figure 6.11 marks the run ID. Because all runs were equally long, it is equivalent to a time axis. The run IDs serve as identification of the tested configurations.

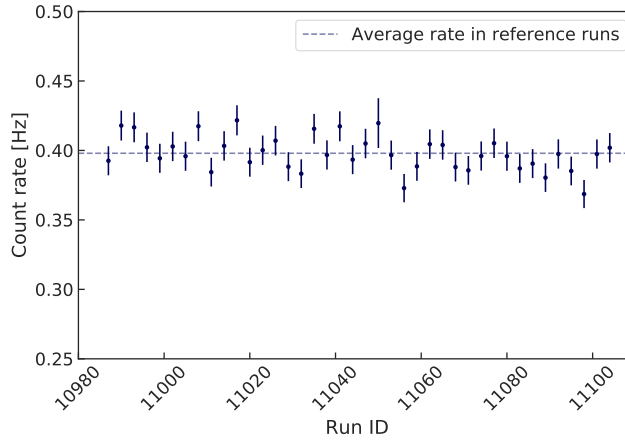
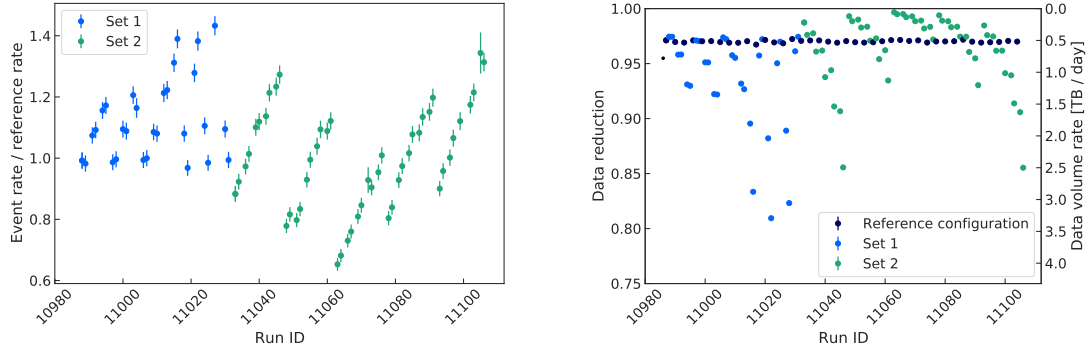


Figure 6.11: Detected event rate in the reference channel. The absolute detected event rates in this channel were mostly stable. To reduce the impact of statistical fluctuations on the result of the trigger optimization, the detected event rates in all trigger test runs are divided by the rates in the reference channel.

In figure 6.12a the event rate of all tested configurations divided by the rate in the reference channel is shown. It can be seen that the configurations yielded event rates between 0.6 and 1.5 times the reference rate. The data reduction compared to streaming data is shown in figure 6.12b. A data reduction of 1 would mean that no data was recorded at all. The reference configuration achieved a reduction of 0.96, which already fulfills the minimum reduction requirement. However, many configurations could increase the data reduction significantly.



(a) Detected event rate in trigger test runs relative to the event rate in the reference channel.

(b) Data reduction in trigger test runs. The reduction in the reference channel (dark blue) is mostly constant.

Figure 6.12: Relative count rate (a) and data reduction (b) in trigger test runs. Two types of trigger configurations were tested: In Set 1, a single FMT-trigger was required to start an acquisition ($n\text{-triggers} = 1$). In Set 2, acquisitions were only started after two FMT-triggers arrived within a short amount of time ($n\text{-triggers} = 2$).

To find the optimum trigger configuration, the two quantities from figure 6.12 are combined. In figure 6.13 the relative event rate (relative to the rate in the reference channel) is plotted against the data volume in TB/day. Each point in figure 6.13 represents a different test run with a different trigger configuration. The color (gray-shading) encodes the run ID. The crossing point of the dashed lines marks the performance of the reference configuration.

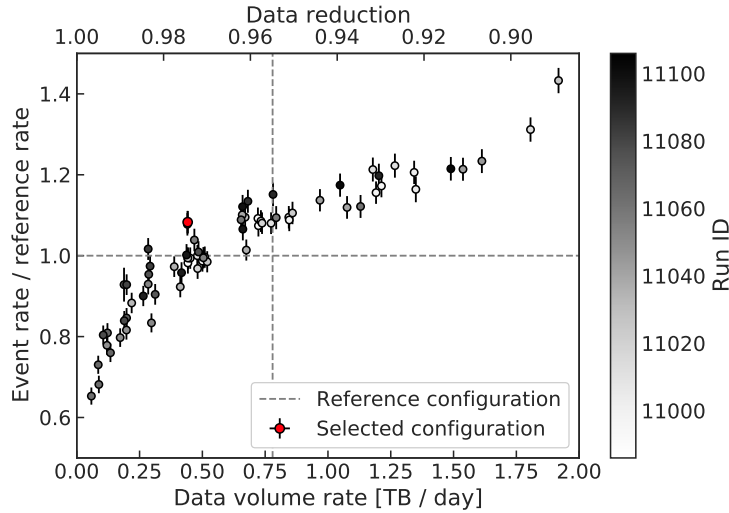


Figure 6.13: Event rate vs. data volume in the trigger optimization test runs. The run marked in red (ID 11087) was taken with the configuration that was chosen for recording tritium data.

6.5. ONGOING DEVELOPMENT

From all tested configurations, three were selected as most promising candidates (see table 6.1). They all have an event rate higher than the reference rate while at the same time improving the data reduction. In the end, the configuration from run 11087 was selected for recording the tritium spectrum in Phase II as well as most of the $^{83\text{m}}\text{Kr}$ data sets that were taken to study sources of systematic errors. Although it is not the configuration with the highest event rate, the additional data reduction (38 % less data volume than the reference configuration) outweighs its 10 % event rate loss compared to the allowed (in terms of recorded volume per day) configuration with the highest event rate.

6.5 Ongoing development

The Phase II trigger implementation achieves a data reduction of 97.4 % while removing only a small fraction of events that could have been reconstructed. The response of the triggering and reconstruction performance to different event properties and their distributions will be studied in more detail in section 8.3. It would be straight forward to improve the data reduction significantly, for example by keeping the information of which frequency bin resulted in an FMT-trigger. Additional nodes could be added that use that information to do online track reconstruction. Originally, a simplified track reconstruction implemented as a trapezoidal filter was planned to be developed for Phase II. Unfortunately, time constraints prevented this development from taking place before tritium data collection started. All real-time event reconstruction projects were postponed to Phase III.

In the future phases of Project 8, the data acquisition will need to process the signals from not just a single channel, but 30 channels or more depending on the antenna array design. The required data reduction is thus at least one order of magnitude higher than what was achieved in Phase II. In addition, the track reconstruction will have to handle additional dimension like transverse position in the antenna array and longitudinal extension of the trapping motion. The Psyllid framework offers the flexibility needed to extend the online data processing to tackle all real-time processing challenges of future CRES experiments. Its development in preparation for Phase III is actively ongoing.

| Run ID | Configuration | $\frac{N}{N_{ref}}$ / $\frac{\text{TB}}{\text{day}}$ | Description |
|--------|---|--|--|
| 11088 | <ul style="list-style-type: none"> • n-triggers = 2 • high-threshold = 13.5σ • threshold = 11σ • skip-tolerance = 0.75 ms | 1.13 ± 0.03 / 0.68 ± 0.01 | Medium level initial FMT-trigger followed by a very low threshold re-trigger within a skip-tolerance that is on the scale of the average event lengths. Data reduction is 96.1%. |
| 11039 | <ul style="list-style-type: none"> • n-triggers = 2 • high-threshold = 14σ • threshold = 11σ • skip-tolerance = 1.0 ms | 1.1 ± 0.03 / 0.66 ± 0.01 | More lenient skip-tolerance (compared to 11088) with higher high-threshold. Data reduction is 96.2%. |
| 11087 | <ul style="list-style-type: none"> • n-triggers = 2 • high-threshold = 13.5σ • threshold = 11σ • skip-tolerance = 0.5 ms | 1.08 ± 0.03 / 0.44 ± 0.01 | Similar to 11088 but shorter skip-tolerance. This reduces the recorded volume drastically on the cost of only a small reduction of the event rate. Data reduction is 97.4%. |

Table 6.1: Best trigger configurations in optimization study.

7

Track and event reconstruction

In the previous sections, the process of digitizing and detecting CRES events was described. The missing link to a reconstructed decay spectrum is the precise determination of the start frequencies of all events that were recorded by Psyllid. This requires that all tracks and events that are contained in the raw time-domain data are identified and analyzed. In general, the event reconstruction should fulfill the highest possible precision standards, as any mis-reconstruction of the start frequency would directly affect the energy spectrum. In addition, it should be robust against mistakenly identifying noise fluctuations as events, since all false events will contribute to the spectrum background. The data analysis is performed using the Katydid analysis software framework [121] developed by the Project 8 collaboration. It will not be described in depth here, instead only the event reconstruction process for Phase II is presented. The Phase II event reconstruction is a multi-step procedure that can be summarized to four main elements:

1. Short-time Fourier transform,
2. Background shape removal,
3. Track identification and
4. Event reconstruction.

The first two steps convert the time-domain data to a raw spectrogram (figure 7.1). The background shape removal ensures that the event detection is not impacted by a frequency dependence of the transmission functions and gains of the various components in the RF and DAQ systems. The track identification searches for the appearance of straight lines in the normalized raw spectrogram. The event reconstruction combines these tracks if they meet certain criteria, for example, that their start and end points are very close in time. Each step is described in detail in the following sections.

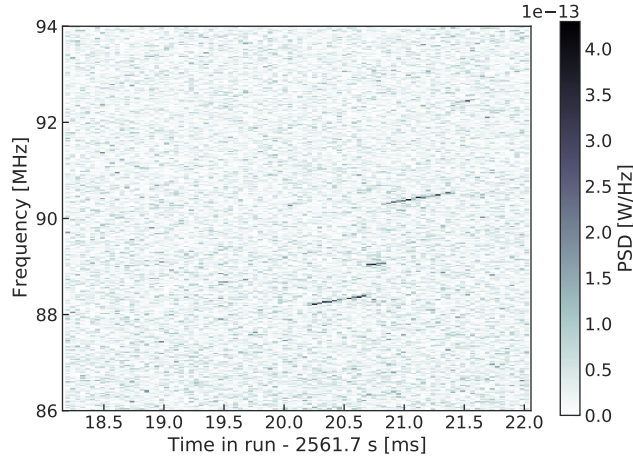


Figure 7.1: Raw spectrogram of a typical CRES event. In each DFT, 4096 time-domain samples (100 MHz sampling rate) are transformed to frequency-domain. Consequently, the frequency and time bin widths are 24.41 kHz and 40.96 μ s respectively. The task of the track and event identification is to detect the four electron tracks and combine them to a single event.

7.1 Time series & short time Fourier transform

As described in section 6.1, Psyllid processes the data in the ROACH2 UDP packets and writes the time-domain data that may contain CRES events as continuous acquisitions to Egg-files. Acquisitions contain records each consisting of 4096 IQ samples. Katydid uses the Monarch library [113] to read in the records and convert the 8-bit samples to voltages using the ADC characteristics (bit depth, voltage range and offset) that are noted in the file headers. During the reading process, the time series is split into *time slices*. The size of one slice does not have to coincide with the size of one record. Katydid could just continue reading samples from the next records until it has reached the pre-configured slice size. Each slice is transformed into frequency-domain using the Fast Fourier Transform (FFT) algorithm implemented in the FFTW3 software library [122]. The resulting frequency spectra are converted to power spectral density (PSD)

$$\text{PSD}_i = \frac{(I_i^2 + Q_i^2)}{50\Omega N \Delta f}, \quad (7.1)$$

where (I_i, Q_i) is the IQ voltage in the i^{th} bin, N is the size of the slice, Δf is the frequency bin width, and PSD_i is the power spectral density value of the i^{th} bin. The frequency bin width Δf is determined by the ratio of the sampling rate f_{sampling} and the number of samples in a slice N :

$$\Delta f = \frac{f_{\text{sampling}}}{N}. \quad (7.2)$$

7.2. BACKGROUND SHAPE REMOVAL

The time bin width corresponds to the length of a time slice:

$$\Delta t = \frac{N}{f_{\text{sampling}}} = \frac{1}{\Delta f}. \quad (7.3)$$

The bin sizes can be adjusted by the choice of the slice size. The choice will affect the signal to noise ratio for a track crossing a frequency bin. Tracks that radiate at 1 fW show a frequency drift (slope) of approximately 300 MHz/s. The slopes observed in Phase II mostly range from 350 MHz/s to 450 MHz/s. The SNR in a bin will be maximal if the track crosses it diagonally. The bin dimensions selected for the Phase II analysis are $\Delta t = 40.96 \mu\text{s}$ / $\Delta f = 24.41 \text{ kHz}$. The diagonal of these bin dimensions has a slope of 596 MHz/s and comes closest to the track slopes for all slice sizes with 2^N samples.

7.2 Background shape removal

The shape of the power spectrum background is the product of the frequency-dependent noise power, originating from the waveguide and the LNAs, and the transmission function of the RF system. Analog and digital filters applied within the data-acquisition system add image noise and basically all components have a frequency-dependent gain. A CRES signal with sufficiently high SNR will appear as a peak above the noise and remain visible for several time bins. Due to the track slope, the signal will cross a certain frequency range during its lifetime. Even if the track's SNR remained constant over time, the absolute power with which it is observed in the raw spectrogram varies with frequency according to the system gain. This introduces a frequency dependence of any track and event reconstruction method that relies on absolute power. To prevent a frequency dependence of the reconstruction efficiency and quality, the power in each spectrum is normalized by the amplitude and the variance of the noise power. This ensures that a signal with constant SNR has a constant power over the entire frequency range. The variation of the gain with frequency is slow as can be seen in figure 7.2. The average and variance of the frequency spectrum is calculated based on all time slices in the processed Egg-file. A 3rd order spline is fitted to the average and the variance to describe the background shape. The power in each time slice is normalized according to

$$P_{\text{norm},i} = \frac{P_i - \mu_i}{\sigma_i}, \quad (7.4)$$

where P_i is the un-normalized power in bin i , μ_i and σ_i are the average amplitude and standard deviation in that bin and $P_{\text{norm},i}$ is the normalized power.

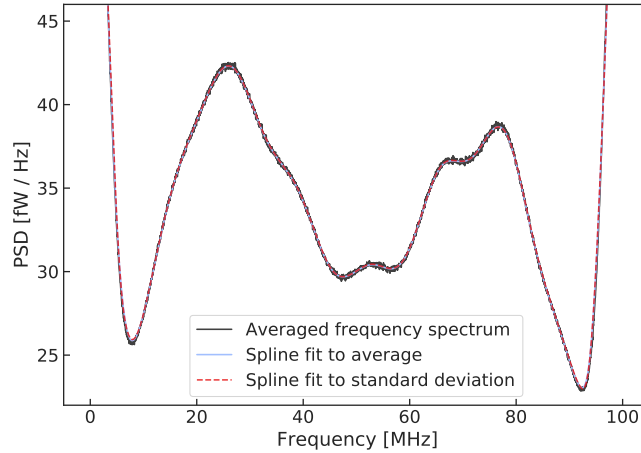


Figure 7.2: Demonstration of the background fit to an averaged power spectrum. As the noise power in each frequency bin is exponentially distributed the mean and standard deviation are identical.

7.3 Track identification

The result of the Fourier transform and background shape removal is a spectrogram that shows the time evolution of the normalized PSD. Tracks appear as lines with a positive slope. The track lines are typically straight, but can have curves due to changing power coupling in the waveguide with frequency. The task of identifying these lines is a classical image processing challenge and numerous algorithms have been developed for this exact purpose. Several feature detection techniques such as Canny-Edge detection or Hough-Transform were tested on Phase II data [123] but so far fell short of producing sufficiently good results. Typically they struggle with the combination of short and faint features in the exponential noise background. Therefore the Phase II track search is based on reducing the raw spectrogram to a sparse spectrogram where only bins with a power above a set threshold are kept (figure 7.3). Instead of searching for features in an image, a point cluster search can now be applied to the remaining bins in the sparse spectrogram.

In Phase I the sparse spectrogram was further cleaned by removing all points that did not have a certain number of neighboring points within a radius. The remaining points were merged to clusters by a density spaced spatial clustering (DBSCAN) algorithm [124]. To fit tracks to a point cluster a Hough transform was used. Unfortunately, this detection technique did not work well enough to achieve a sufficiently high detection efficiency for Phase II data. It was therefore replaced by the Sequential Track Finder (STF), an algorithm that was first proposed by [125] and further developed and adapted for Phase II as part of this work. The STF is not the only algorithm that has been shown to give good results in the analysis of Phase II data. For example, the Maximum

7.3. TRACK IDENTIFICATION

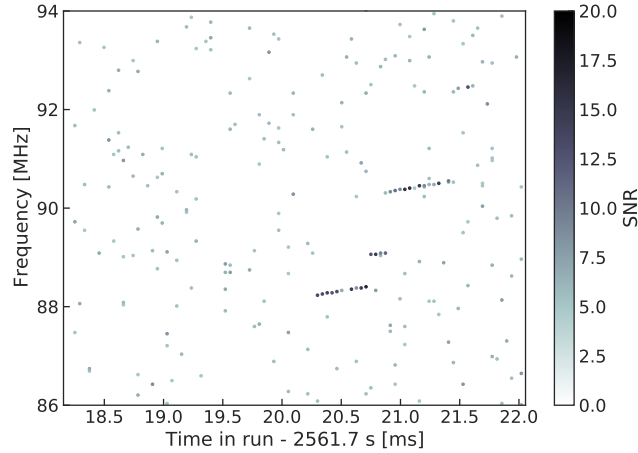


Figure 7.3: Sparse representation of the spectrogram from figure 7.1. The applied cut uses a threshold defined either as SNR or as standard-deviation (σ) threshold. In this example only bins with $P_i \geq \mu_i + 4 \cdot \sigma_i$ were kept.

Likelihood Clustering [126], developed by Evan Zayas, has performed similarly well. Nonetheless, in this chapter only the Sequential Track Finding is presented, since it is the reconstruction method used for the official Phase II analysis.

7.3.1 Sequential Track Finder

The sequential track finding algorithm is a track clustering method operating on sparse spectrograms. It was initially designed as a triggering method that could in principle be implemented for example on an FPGA. The fact that it processes the data time slice by time slice, without using the full 2D image, makes it ideal for fast real-time processing at the cost of not being able to look forward or backward in time and therefore giving up some information. However, as all electron signals slowly move upwards in frequency, the loss of information is compensated for by the algorithm using this knowledge to predict where the signal will appear in the next time slice.

For the track clustering, the algorithm distinguishes track points from noise points by making use of a key feature of the electron signal signatures: their quasi-linearity. It searches the sparse spectrogram for points that fall on a more or less straight line. To this end, it requires that the separation of the points in time and frequency is smaller than a set maximum distance and that the start and the end point of the electron tracks have an SNR above a certain threshold.

In its Katydid implementation, the algorithm receives a list of points for every time slice. The points in this list represent all bins from the normalized spectrum that were accepted by the SNR- or σ -thresholding. Every point contains all relevant information about the bin it originates from (table 7.1).

| Point property | Symbol |
|------------------------------------|---|
| Frequency position | f_i |
| Time in run (time slice center) | t_i |
| Power | P_i |
| Background power mean at f_i | μ_i |
| Background power variance at f_i | σ_i^2 |
| Threshold power | $P_{r,i} = \mu_i + r_\sigma \cdot \sigma_i$ |
| Signal to noise ratio | $\text{SNR}_i = \frac{P_i}{\mu_i}$ |
| Normalized unitless power (NUP) | $P_{\text{norm},i} = \frac{P_i - \mu_i}{\sigma_i}$ |
| Local power residual excess | $A_i = \frac{\sum_{j=i-d}^{i+d} (P_j) - d \cdot \mu_i}{\sigma_i}$ |

Table 7.1: Point properties that are passed on to the STF. i and j refer to the indices of the frequency bins in a slice. d and r_σ are configurable parameters defining the size of the area over which bin powers are summed and the threshold that was applied to create the sparse spectrogram.

In the very first time slice given to the STF, it will consider every point in the list as a potential new track start and create *sequential line* objects for each of them. A sequential line is an instance of a class with a start and end point in time-frequency space, a slope (in Hz/s), a summed SNR, and a summed NUP that are calculated from the properties of the points contained in the line object. At initialization the line consists of just a single point. The line's start and end point are both set to the frequency and time properties of this point and the slope is set to a configurable initial slope parameter. The choice of this initial slope has been shown to not be critical to the success of the track clustering as long as it is of a similar magnitude as the typical electron track slope.

In the consecutive time slices, the STF compares the extrapolation of the existing active lines in the direction of their slope to the newly incoming point lists. As longer lines are more likely to correspond to a real electron signal and points of higher SNR are more likely to originate from a radiating electron than from noise, the point-to-line comparison is done in ascending order of line start time and descending order of normalized power. While it is possible for a line to pick up more than one point per time slice (as long as

7.3. TRACK IDENTIFICATION

the point distances are within the frequency acceptance), the sorting of the line-point comparison ensures that lines are not deviated by points of lower SNR. In general any point is appended to an existing line if the condition

$$|f_i - (f_{\text{end},l} + s_l \cdot (t_i - t_{\text{end},l}))| < \Delta f_{\text{STF}} \quad (7.5)$$

is fulfilled, where $f_{\text{end},l}$ and $t_{\text{end},l}$ are the end frequency and time of a sequential line, s_l is the line slope, and Δf_{STF} is the frequency-acceptance. If a point cannot be matched to a pre-existing line, a new line instance is started with this point as start point. After a point has been added to a line, the line's slope is re-evaluated, taking the power residuals as weights

$$s_l = \frac{\sum_{i=1}^N P_{\text{norm},i} \cdot \frac{f_i - f_{\text{start},l}}{t_i - t_{\text{start},l}}}{\sum_{i=1}^N P_{\text{norm},i}}. \quad (7.6)$$

This ensures that the slope of a line is not deflected by the occasional addition of noise points. A line object is set inactive once the distance in time between the last point that was added and the current time slice is larger than the time-gap tolerance Δt_{STF} , which determines the maximally allowed time distance between the points in a line. Once inactive, a line undergoes a trimming process that repeatedly cuts off its start and/or its end point until the SNR of the start and end exceed the configurable trimming threshold. Eventually, if a line still contains more than a required minimum number of points after the trimming, and its slope exceeds a set minimum slope (typically > 0 Hz/s), it is passed on to downstream processing. The result of the sparse spectrogram from figure 7.3 processed by the STF is shown in figure 7.4a. The acquisition that is used as example contained one event consisting of four tracks, three of which were recognized as lines by the STF.

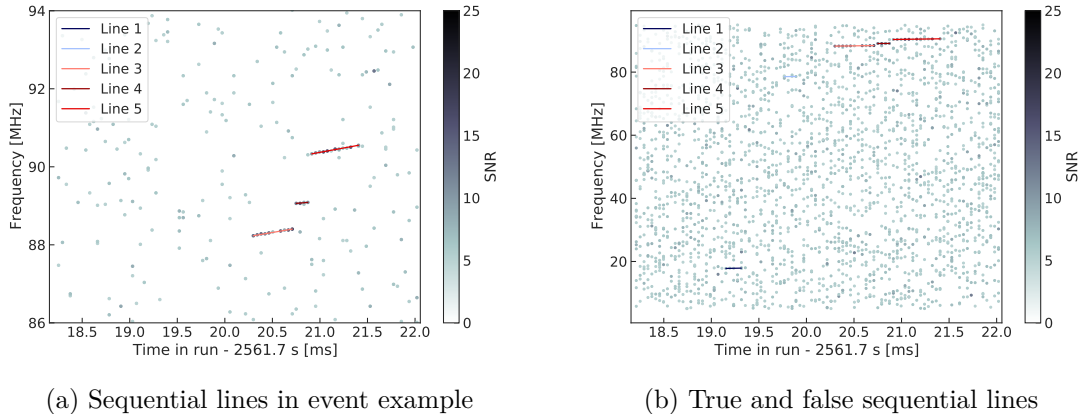


Figure 7.4: Two sparse spectrograms of the same event displaying different frequency ranges. The STF detects 3 out of the 4 tracks contained in the event from figure 7.1 but in (b) it can be seen that it further falsely identifies some noise fluctuations as track candidates.

The end product of the STF are line objects that still contain the full list of clustered points and all their properties. In addition, lines also already possess track like properties like start and end points, length, slope, summed SNR, and summed NUP. While the STF has proven to be very successful at finding lines that directly correspond to electron tracks, it also collects noise points that are accidentally aligned (figure 7.4b). In addition, power fluctuations along a track can cause the STF to identify a single track as two or more separate lines. Therefore, further processing steps are required that combine lines, which belong to a single track and that remove noise point clusters from the set of track candidates.

7.3.2 Track segment clustering

A varying SNR in a track leads to a probability greater than zero of breaking tracks during point clustering into several line segments. If these broken tracks are not repaired, the consequence could be the false detection of additional events. Because the electron signal shifts upwards in frequency with time, these additional false events have a higher start frequency and therefore cause high frequency tails in a start frequency spectrum. To prevent this, additional track segment clustering steps recombine broken track segments and prevent misidentification of events in the downstream Katydid processing. In the following cases, the segment clustering will combine two track segments to one single track:

- The start point of one track overlaps with the end point of another track in time. Their distance in frequency is small and their slopes are similar.
- The start point of one track is separated by the end point of another track by not more than a certain limit. Extrapolating both tracks in time brings them closer in frequency than the frequency-acceptance Δf_{tc} . In addition their slopes are similar.

$$\begin{aligned} |t_{1,\text{start}} - t_{2,\text{end}}| &< \Delta t_{tc} \\ |f_{1,\text{start}} - (f_{2,\text{end}} + s_2 \cdot (t_{1,\text{start}} - t_{2,\text{end}}))| &< \Delta f_{tc} \\ \left| \frac{s_1}{s_2} - 1 \right| &< \frac{\Delta s_{tc}}{s_2} \end{aligned} \quad (7.7)$$

At this point the track search is completed and the STF lines are processed to calculate the final track properties from the point properties contained in the line.

7.3. TRACK IDENTIFICATION

7.3.3 Track processing

After the point clustering and segment merging has been completed, the final track properties (see table A.1) are calculated. To this end, a straight line is fitted to the point cluster contained in the sequential line objects. The start frequency of a track is defined as the intersection of the fitted line with the time coordinate of the earliest point in the cluster. This way, a better start frequency precision is achieved than if the frequency bin center of the earliest point was used. Once processed, a power cut is applied and all tracks that don't exceed a set average normalized bin power are removed. This step alone removes almost all noise fluctuation from the event candidate list as can be seen in figure 7.5. Figure 7.6 shows that the true tracks in figure 7.4b survive the cut.

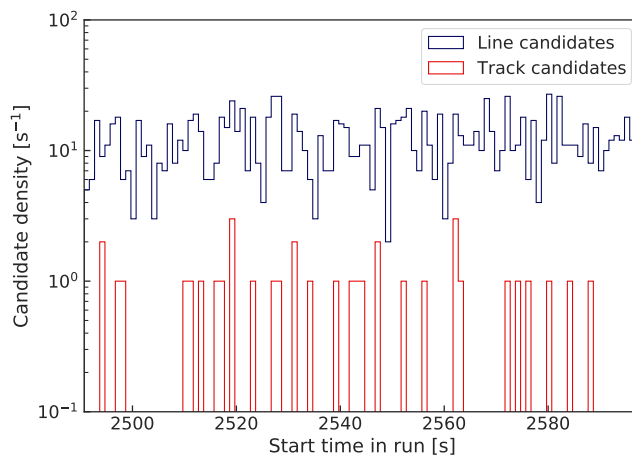


Figure 7.5: Sequential line and track rate in the Egg-file containing the example event from figure 7.1. The segment clustering and the average-power cut remove the majority of the false lines detected by the STF.

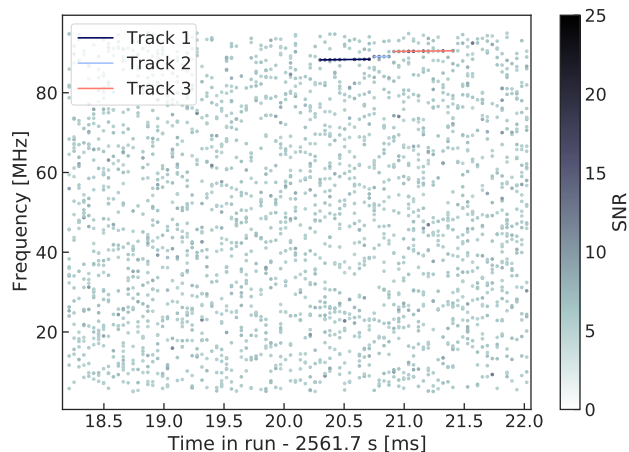


Figure 7.6: After the track processing is completed a power cut is applied and the false candidates found by the STF in figure 7.4b are removed.

7.4 Event reconstruction

After the track reconstruction, all tracks originating from the same electron have to be grouped to a single event. This is done in two stages:

1. Combination of parallel tracks corresponding to sidebands.
2. Combination of consecutive tracks that are offset in frequency and time but belong to the same electron.

7.4.1 Sideband clustering

First, the main carriers and sidebands have to be combined to multi-peak tracks. This is achieved by doing a head-to-head and tail-to-tail comparison. If tracks start or end approximately at the same time (within a time-gap tolerance $\Delta t_{\text{sidebands}}$), they are considered parallel appearances of the same electron in different bands. The grouping condition is

$$\begin{aligned} |t_{1,\text{start}} - t_{2,\text{start}}| &< \Delta t_{\text{sidebands}} \\ |t_{1,\text{end}} - t_{2,\text{end}}| &< \Delta t_{\text{sidebands}}, \end{aligned} \tag{7.8}$$

where t_1 and t_2 are either the start times or the end times of two reconstructed tracks. Every track is assigned to a multi-peak track object. Each of these objects consists of one or more tracks that are grouped together by the multi-peak track clustering in Katydid. The multi-peak track is assigned a combined start time and start frequency, corresponding to the start time and frequency of the track with the earliest start time in the group. The length of the multi-peak event is defined as the time in run of the latest end-point minus the time-in-run of the earliest start-point of a track in the object. In theory, all tracks would start and end at the same time, since their common source of existence is the radiating electron between two scatters. In reality, however, the start and end time of the tracks correspond to the earliest and latest bin that is assigned to the tracks by the STF, which is associated with an uncertainty. Choosing the earliest and latest bin for defining the duration of the multi-peak track is motivated by the assumption that the track finding is more likely to miss a bin that should have been assigned than to add a bin that does not belong to the track. The selection of the start frequency from the track with the earliest start time, on the other hand, has no physical motivation. Ideally, the multi-peak track clustering would understand the topology of the event and be able to distinguish between main carriers and sidebands. This was demonstrated by the means of machine learning track classification in [100]. If it is known which track is the main carrier and which are sidebands, the multi-peak track builder could either use the start frequency of the main carrier as true start frequency, or it could calculate the main carrier frequency, for example from the average of the first

7.4. EVENT RECONSTRUCTION

order sideband frequencies (if both are visible). In most Phase II data, no sidebands are detectable and the multi-peak tracks typically only contain the main carrier.

7.4.2 Event clustering

In the second stage of the event clustering multi-peak tracks are grouped together to multi-peak events, this time, by head-to-tail time comparisons:

$$\begin{aligned} |t_{1,\text{start}} - t_{2,\text{end}}| &< \Delta t_{\text{event}} \\ |t_{1,\text{end}} - t_{2,\text{start}}| &< \Delta t_{\text{event}} \end{aligned} \quad (7.9)$$

Multi-peak events contain all parallel and consecutive tracks originating from one electron. At the end of the multi-peak clustering, the event properties are calculated from the properties of the tracks they contain. The time and frequency related properties are defined equivalent to the track properties in table A.1. In addition, the SNR and power related properties of the first track are also available member variables in the event object. The event start frequency is equal to the start frequency of the earliest multi-peak track (figure 7.7). The precision and accuracy of the start frequency reconstruction will be evaluated in chapter 8. All multi-peak event properties are listed in table A.2.

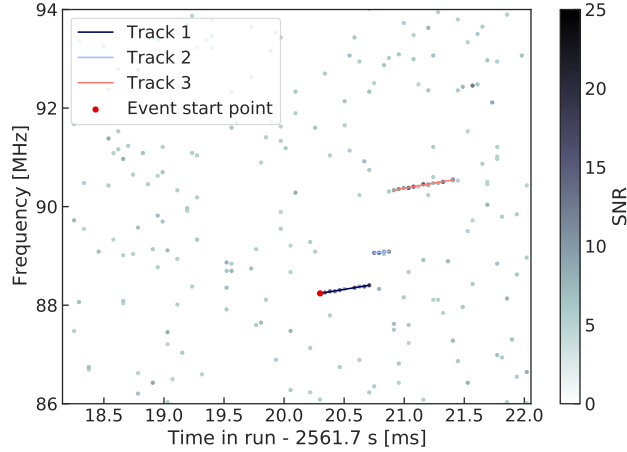


Figure 7.7: Final reconstructed event in the sparse spectrogram from figure 7.3.

7.5 False events and power cuts

The STF falsely identified two additional lines in the acquisition containing the event from figure 7.4. On average, it finds 12 lines per second in data that contains only noise. More than 95% of the noise fluctuations that fulfill the STF criteria are removed by the average power cut that is applied to the processed tracks. Around 0.3 false tracks per second still survive this cut (figure 7.5). No additional requirements have to be fulfilled for tracks to be converted to multi-peak events. Therefore, at the end of the track and event reconstruction more cuts are applied to the completed events. In this process, the maximum allowed false event rate determines the cut thresholds. The more stringent the cut thresholds are, the more real events are removed. The cut choice is therefore a trade-off between true event efficiency and false event rate. The total tritium run duration in Phase II was initially planned to amount to 100 days. It was the declared goal to not find a single false event in the tritium spectrum. To this end, the cuts were set to reduce the chance of having a single false event per ROACH2 channel in 100 days of data taking to below 10%.

There are several SNR related track and event properties that the cuts could be applied to (see table A.1 and table A.2). In the final Phase II reconstruction configuration, all cuts target the average P_{norm} (NUP) of the first track in an event. The average P_{norm} is calculated by dividing the summed normalized power residuals (TotalNUP) by the number of points contained in the STF point cluster (NTrackBins). Only the first track of an event is evaluated by the cut because it is obviously the most critical element and must be a plausible event start. To set the thresholds for the false event removal, there are two options:

- Choosing thresholds to remove all false events from 100 days of triggered data: The Phase II trigger reduces the recorded run duration by about 97.4%. The assumption could be made that the fraction of false events that are already removed by the trigger is of a similar order of magnitude. Only the remaining false events have to be removed by the SNR cuts.
- Choosing thresholds for 100 days of un-triggered data: Assuming that all false events that meet the STF line criteria would also result in a trigger leads to the conclusion that the trigger does not remove any false events. This is at least partly justified as the STF and the trigger algorithm search for similar event features (several high-SNR bins within a short amount of time).

The truth is probably somewhere in between. To be conservative, the thresholds are set following the second assumption.

7.5.1 Event grouping

As mentioned before, the event cuts applied are first track average NUP (P_{norm}) cuts. A global¹ SNR cut with just a single threshold would drastically reduce the event reconstruction efficiency. To limit the removal of too many real events, the events are sorted into groups and a separate cut is applied to each group. The properties that decide which group the event is assigned to are the number of tracks in the event and the number of bins in the first track. The idea behind this is that it is a lot less likely for a false event to contain long or numerous tracks than it is for real events. For example, the group of events composed of three tracks and a first track that contains six bins requires a less stringent cut than events that consist of only one single six-bin-long track. If however, the first track of the three tracks is constituted of only three bins, it is more questionable that the first track is not a noise fluctuation unless the average power of these three bins is very high. Table 7.2 shows the event groups that have specific thresholds applied to them. All events that do not match any of these groups undergo a global cut.

| First track points - Number of tracks | | | | | |
|---------------------------------------|-------|-------|-------|-------|-------|
| 3 - 1 | 3 - 2 | 3 - 3 | 3 - 4 | 3 - 5 | 3 - 6 |
| 4 - 1 | 4 - 2 | 4 - 3 | 4 - 4 | 3 - 5 | |
| 5 - 1 | 5 - 2 | 5 - 3 | 5 - 4 | | |
| 6 - 1 | 6 - 2 | 6 - 3 | 6 - 4 | | |
| 7 - 1 | 7 - 2 | 7 - 3 | | | |
| 8 - 1 | 8 - 2 | | | | |
| 9 - 1 | 9 - 2 | | | | |
| 10 - 1 | | | | | |
| 11 - 1 | | | | | |

Table 7.2: Event cut groups. Each combination of number of points in the first track and number of tracks in the event is subject to a separate power cut. The power thresholds are set to reduce the total number of false events to less than 1/100 days per ROACH2 channel.

¹Global cut: the same cut criteria are applied to all events

7.5.2 Choosing thresholds

To determine the required threshold level, one hour of streaming (un-triggered) data was recorded, while the cell was pumped out and the trap coils were switched off. This data is completely free of real events. Hence, anything reconstructed as an event has to be a false event. The track and event clustering is performed and the detected false events are categorized into groups according to table 7.2. For each group a histogram of the first track average NUPs is fitted with a Gamma distribution

$$f(x; \alpha, \beta) = \frac{\beta^\alpha \cdot x^{\alpha-1} \cdot e^{-\beta x}}{\Gamma(\alpha)}. \quad (7.10)$$

Figure 7.8 shows the first track average NUP of events in the 3 - 1 group (single-track events with three track bins). From the Gamma-fit, a cut threshold, which would remove all false events occurring at an arbitrary rate, can simply be read off the average NUP axis. This way, the threshold for each group is chosen so that the total rate of false events from all groups together would be smaller than 1 per 100 days (with 90 % confidence). The statistical power in groups with rare event topologies (many bins and many tracks) is naturally limited. All groups that did not contain a sufficient number of false events in order to extract a reliable cut threshold are subject to a global cut. Its threshold is set so that false events from the group with the lowest rate that still contained enough events for a targeted threshold are excluded. The final Katydid configuration with all configurable parameters for the track search, the event clustering, and the cuts can be found in listing A.1.

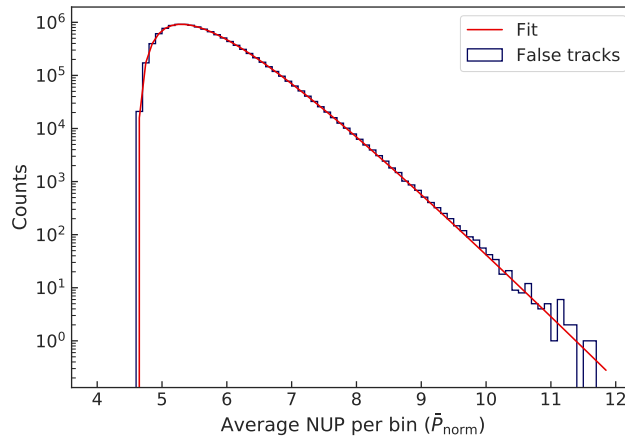


Figure 7.8: Gamma fit to the power distribution of false events. For the false event reconstruction in this data set, more lenient clustering parameters were used to increase the statistical power in the analysis.

7.6 Data quality plots

In CRES spectroscopy the variable of highest interest is the start frequency as it can be directly linked to the electrons kinetic energy right after the decay. However, many other track and event characteristics are extracted by the reconstruction algorithm. Some of these characteristics, like power and slope, can be used to improve the energy reconstruction [100]. Others yield information about the intrinsic quality of the recorded data as well as the performance of the reconstruction algorithm. Here, a few examples of quantities that help to evaluate the data and reconstruction quality are presented. The official reconstruction analysis in Phase II automatically produces a series of plots for each data run that consist of one- or two-dimensional histograms of some variables from table A.2. Some fits are applied to extract useful quantities like the mean of the track and event length distributions. Comparing these quantities between runs helps to track a change of pressure or gas composition in the cell, as these physical parameters affect the shape of the event property distributions. Figure 7.9 shows a sub-set of the quality plots for the calibration run that was already used in section 4.2.2 to calibrate the magnetic field and energy resolution for the Phase II run conditions.

In addition to pressure and trap depth information, the track and event length distributions in figure 7.9a also demonstrate the effect of the minimum track length that can be detected (3 bins) and the group-specific event cuts that are applied. Intrinsically, the track and event lengths follow an exponential distribution. The roll-off towards short tracks (and events) is caused by the performance limitations of the track and event reconstruction. Harsher power cuts on short tracks lead to a roll-off at higher minimal track lengths. The reconstruction algorithm and its parameters were optimized to push the roll-off in the lengths distributions as much as possible towards short lengths, as this allowed to set higher pressures in the gas cell and thus increases the tritium event rate.

The distribution of the event start times relative to the acquisition start (figure 7.9b) is maximum at 2 ms, which corresponds to the ROACH2 pre-trigger time. In fact, most events start before the pre-trigger time and are only detected by the trigger algorithm later in their lifetime. By design, almost no events start before 0.5 ms, but a small fraction starts after the trigger occurred. While some of these may have been recorded after a noise trigger, it is plausible that the vast majority consists of events with a missed first track. This indicates that the trigger efficiency is much better than the reconstruction efficiency which will be confirmed in chapter 8.

7.6. DATA QUALITY PLOTS

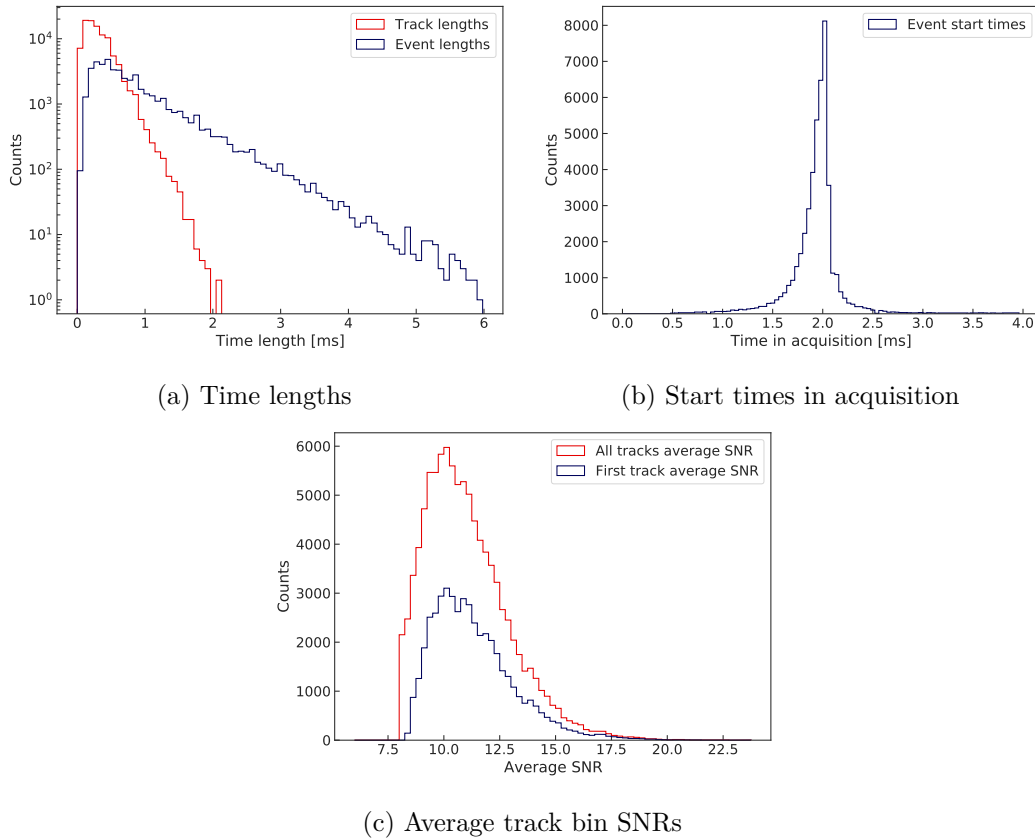


Figure 7.9: Data quality plots from a krypton data run. The goal of data quality plots is to provide information about the run conditions and/or the reconstruction performance. The track length distribution (a) yields information about the gas pressure in the cell. The distribution of event start times (b) serves as check for successful triggering and the distribution of track SNR (c) tells about the signal power and power thresholds that were used.

Figure 7.9c shows the distribution of the average SNR of the first tracks and of all tracks in the events that survived the event cuts. The first tracks (blue) are a subset of all tracks (red). The sharp cut-off in the all-tracks SNR distribution is caused by the power threshold applied to the tracks prior to the event clustering. The majority of electrons only couple weakly to the transporting waveguide modes. As a result, the underlying power distribution is expected to decrease monotonically towards higher powers. The reason why the maximum of the reconstructed SNR distribution is not at the cut-off SNR is that, depending on the event topology, a higher average SNR is required for the first track.

8

Performance studies with fake events

In the previous chapters, the reconstruction and triggering methods were optimized based on real Phase II data. Here, a thorough qualitative and quantitative understanding of the detection methods' performance and their dependence on various event properties will be achieved by the use of fake events. Fake events are simulated signals resembling the signal of CRES events. They are not based on actual electron tracking simulations but instead are generated as complex sinusoidal signals with event-like properties added to white noise. Their features are randomly drawn from probability density functions. The generated fake signals undergo a processing sequence (mixing, filtering, digitizing) that is similar to the processing in the experiment. The advantage of fake events over actual electron tracking simulations is that the computation costs per simulation are low and that instead of the electron trajectories, the true signal features are known. This allows the mass production of CRES-like events in an affordable amount of time ($\sim 10^5$ per day) and therefore makes it possible to generate sufficient statistical significance to understand the detection performance and its dependence on signal features in detail. The tool used to generate these fake events is Locust [127], a simulation package developed by the Project 8 collaboration (section 8.1). Because fake events are not simulated tracked electrons, their similarity to real recorded electron events must first be proven. This is accomplished in section 8.2. In section 8.3 the fake events are used to study the trigger and reconstruction efficiency dependence on event properties.

8.1 Locust

The Locust software package [128] is a simulation tool that can model the response of an RF receiver chain to time-varying electromagnetic fields. Its modular implementation allows for highly flexible extensions and algorithmic implementations of various receiver configurations. Locust accepts calculated electromagnetic signals as input in the form of sinusoidal waveforms or arbitrary externally defined signals. A signal can be processed in stages by various RF components that are implemented as configurable signal processing classes. Figure 8.1 shows a flow diagram of the components that provide the central functions of Locust.

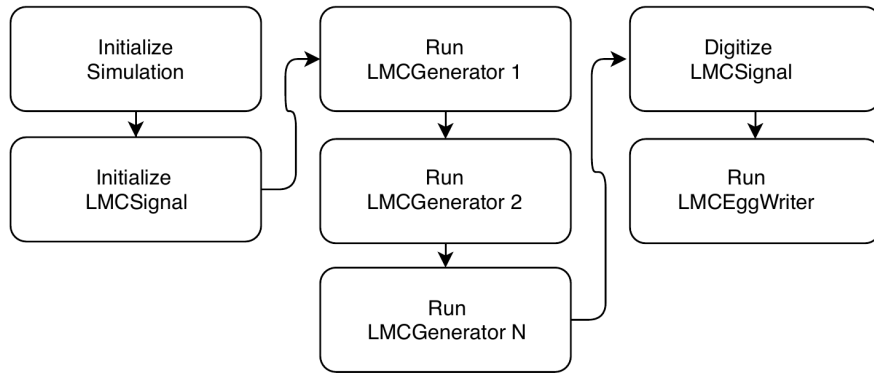


Figure 8.1: Locust simulation flow chart from [127]. After initialization, the signal is operated on by a series of generator and processing instances.

Locust can currently operate in three modes: It can independently generate ideal signals (sinusoid, noise, ...), generate fake CRES-like signals, or wait periodically for an electromagnetic signal that is generated externally by another software package.

8.1.1 Signal flow

The signal generation begins with the definition of all simulation properties. These include the sampling frequency, the record size, and the output file paths. During initialization, a signal object (LMCSignal) is created and configured with all defined simulation settings. The object contains an array of complex voltages following the provided definitions. For fake event generation, the signal is created internally and the array is first populated in LMCGenerator 1. All following blocks until LMCGenerator N, perform consecutive operations on the same signal object. An arbitrary number of blocks can be configured. The last two steps of the simulation are the digitization and writing of the processed signal to an Egg-file [113].

8.1.2 Fake event generator

In this work, the customized generator used to produce the voltage signal is the LMC-FakeTrackSignalGenerator (FTG). During a simulation it is configured with the track and event parameters of interest. In addition, noise is added as background to the signal by the LMCGaussianNoiseGenerator that generates random voltages following a normal distribution. In a single simulation, track and event properties are set by defining a collection of probability density functions (PDF) which describe the distributions of the CRES signal properties. A random seed from which the Monte Carlo-driven signal simulation is initialized can be specified or randomly drawn. The following signal properties are randomized in the FTG:

8.1. LOCUST

- **Start frequency (Hz):** The first track start frequency is sampled from a uniform PDF that is bounded by minimum and maximum values. The start frequencies of later tracks in the event are calculated by adding a frequency jump size to the end frequency of the previous track. Since Locust version v1.15.11 the start frequency can also be drawn from a PDF describing the 17.8 keV $^{83\text{m}}\text{Kr}$ lineshape. However, this feature is not used in this work.
- **Slope (MHz/ms):** The track slopes are drawn from a normal distribution with configurable mean and standard deviation. Alternatively, a uniform distribution can be selected. The slopes from consecutive tracks are not correlated.
- **Start time (s):** The event start times are sampled from a uniform PDF bounded by minimum and maximum values.
- **Track length (s):** The track length follows an exponential distribution defined by the mean track length.
- **Number of tracks per event:** The number of tracks per event can be configured either by defining a geometric distribution with a mean number of tracks, or by setting a minimum trapped pitch angle. To prevent the number of tracks drawn from the geometric distribution from being zero, the configured number is incremented by one. The actual mean number of tracks is therefore one greater than the value set in the configuration. This is not required when the number of tracks is determined via the pitch angle.
- **Start pitch angle ($^\circ$):** Each track in the FTG has an associated pitch angle that is used to calculate the received power in a harmonic trap relative to a 90° electron (figure 4.9). The configured track power is corrected by multiplying it with this relative power. The cosine of the start pitch angle follows a uniform distribution with a configurable minimum and maximum θ_0 .
- **Pitch angle changes:** If the number of tracks per event is not defined by a geometric distribution, a new pitch angle is calculated for each track. The distribution of the pitch angle changes follows an approximate model [129]

$$f(d\theta) = 1 / \left(1 + \sin^2 d\theta / \alpha^2 \right). \quad (8.1)$$

- **Frequency jumps (Hz):** The frequency jumps between tracks are randomly drawn from an energy loss function obtained by interpolating oscillator strength data [130] and correcting the resulting frequency increase by the frequency dependence on pitch angle in a harmonic trap (4.11).

- **Start z -position:** A cosine uniform distribution does not describe the distribution of pitch angles at the bottom of a trap correctly. To correct the start pitch angle, the longitudinal start position on the z -axis can be set to a uniform distribution. The start pitch angle is then adjusted by the angular shift an electron would experience when moving from its start z -position to the bottom of the harmonic trap ($z = 0$ m). This way a more realistic pitch angle distribution is obtained.

In addition to the event parameters and distributions listed above, the magnetic field value and the signal power must be specified for each simulation of a fake event. The value of the magnetic field is used, for example, for the calculation of the frequency jump size. Figure 8.2 shows an example fake event.

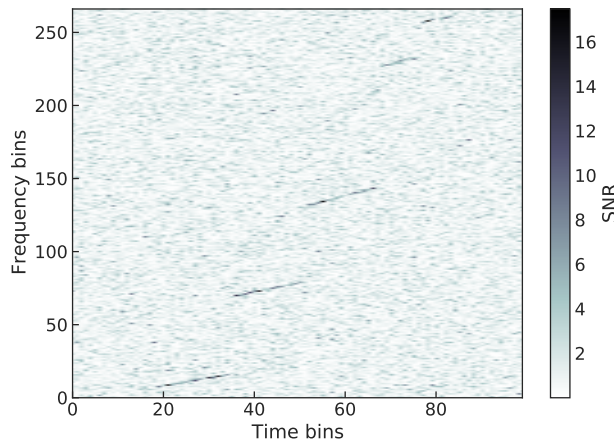


Figure 8.2: Raw spectrogram of a simulated signal generated by Locust to imitate the characteristics of a CRES event. As in all previous spectrograms (e.g. figure 7.1), the bin dimensions are $40.96 \mu\text{s}$ (time) and 24.414 kHz (frequency).

8.2 Simulation validation

There are two levels of agreement between simulated and experimentally observed signals that have to be confirmed before the use of fake events for reconstruction and triggering performance studies is justifiable:

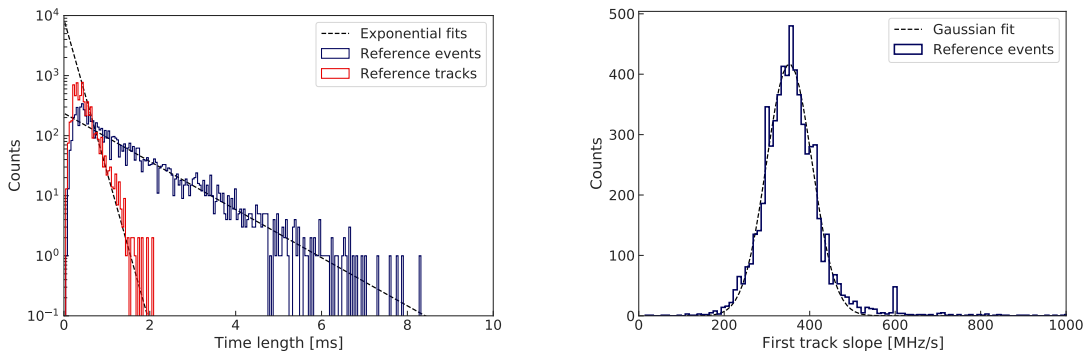
1. The event property distributions that are relevant for the reconstruction quality need to match their respective distributions in real events. These properties are track length, number of tracks per event and event length, maximum SNR of any bin in a track, and average SNR of a track.

8.2. SIMULATION VALIDATION

2. Because the track reconstruction is sensitive to the SNR in tracks on a bin-by-bin basis, the fluctuation of SNR along a track has to be as similar as possible to the fluctuations in a real electron track. At the same time, this requires that the noise fluctuations are identical too.

8.2.1 Reference data and event property distributions

In order to avoid any bias that may be introduced by the trigger, the simulated fake events are compared to un-triggered data. The comparison data set was taken in the Q300 trap at $1.6 \cdot 10^{-6}$ torr and contains 6257 reconstructed events. The track and event length and the track slope distributions are shown in figure 8.3. The exponential track and event length constants are (0.18 ± 0.02) ms and (1.09 ± 0.02) ms respectively. The slope distribution is approximated by a Gaussian distribution with a mean of (352.3 ± 0.8) MHz/s and a standard deviation of (54.5 ± 0.8) MHz/s.



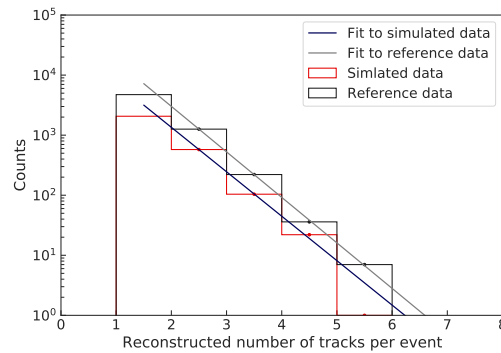
(a) Track and event lengths in reference data. Via a fit with an exponential distribution, the underlying mean track (or event) length is extracted. The fit is limited to lengths ≥ 0.5 ms to not include the roll-off caused by the reduced efficiency for short tracks and events.

(b) Slopes in reference data. The peaks at 298 MHz/s and 596 MHz/s correspond to the (half-)diagonal of a raw spectrogram bin. Otherwise, the slope distribution is approximately normal.

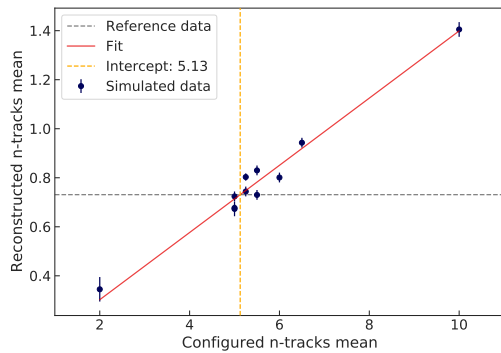
Figure 8.3: Event properties in un-triggered reference data.

The event length is defined as the difference between the end time of the last track and the start time of the first track in the event. It is therefore a quantity that cannot be directly set in the configuration of the simulations. The mean number of tracks per event in the reference data is 1.3. However, this number does not reflect the true number of tracks per event because individual tracks can be too short or too faint to be detected but still contribute to the event length if they are not the first or last track in the event. In addition, the track segment clustering and the multi-peak track builder regularly combine tracks if their start and end points are too close together in time or frequency, which further falsifies the reconstructed number of tracks per event. Therefore, the number of tracks per event in the fake events has to be matched to the reference data by finding

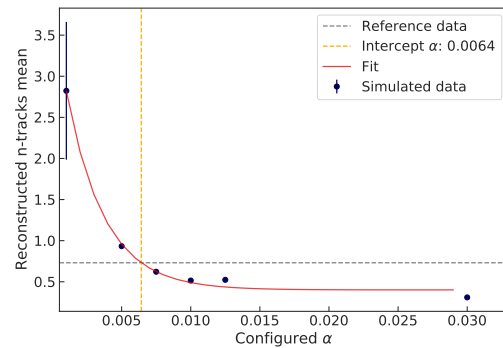
the parameter setting of the FTG that provides the best agreement after reconstruction (figure 8.4). To this end, sets of fake events with an average track length of 0.18 ms were generated. Each set was simulated with a different average number of tracks, either by choosing a different n-tracks mean parameter for the geometric distribution, or by setting a different α for the pitch angle changes in equation (8.1). The fake events were processed by Katydid and the resulting distribution of reconstructed number of tracks per event compared to the reference data. An exponential function was fitted to the distribution of $N_{\text{tracks}}/\text{event}$ as shown in figure 8.4. The fit result for each simulated set was plotted against the input parameters to predict the best value via interpolation.



(a) Example comparison of $N_{\text{tracks}}/\text{event}$ in simulated and reference data after reconstruction.



(b) Fitted exponential $N_{\text{tracks}}/\text{event}$ parameter vs. configured mean number of tracks.



(c) Fitted exponential $N_{\text{tracks}}/\text{event}$ parameter vs. configured α (8.1).

Figure 8.4: The optimum underlying number of tracks per event was found by comparing simulated to real data in parameter scans.

8.2. SIMULATION VALIDATION

The event start time and start frequency are irrelevant for the reconstruction and were set to a 100 bin wide range in each dimension. While the start frequency distribution of 17.8 keV (K-line) electrons in the Q300 trap configuration is not uniform, it is wide enough that no significant bin position effects are expected to bias the result.

SNR distribution

Like the number of tracks per event, the SNR distribution can also not be read directly from the reconstructed reference data. The reconstructed track SNR is always affected by track position in a time slice relative to the closest frequency bin center as well as the slope and the noise power fluctuations. In addition, the reconstruction efficiency has a big impact on the shape of the SNR distribution after reconstruction. There are two approaches to find the shape of the underlying power distribution in the Phase II apparatus:

- **Kassiopeia [131] simulations:** Similar to the simulations done in [127], the field geometry and waveguide of the Phase II setup were implemented in a Kassiopeia simulation. To find the power distribution of trapped electrons, particle tracking simulations of electrons with a kinetic energy of 17.8 keV, a random pitch angle $\theta_0 \geq 89^\circ$ and a random position in the cell were conducted. The magnetic field of a single coil with 300 mA was simulated and the generated electrons were tracked for the simulated time duration of 40.096 μs , which is the exact length of one spectrogram time bin. The incident electromagnetic fields calculated by Kassiopeia were mixed and sampled in Locust. The noiseless signal power spectrum was calculated from the Locust samples. Both the main band and first order sidebands are visible in the Fourier transformed spectra (figure 8.6a). From integrating each band, the main carrier and sideband power for each tracked electron was found. Combining the band powers from all generated electrons results in the expected Phase II power distribution of 17.8 keV electrons (figure 8.6b).

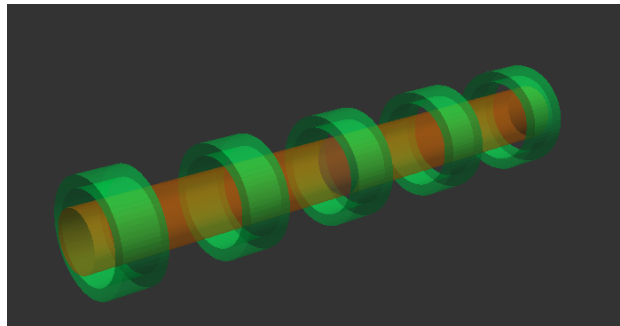


Figure 8.5: Phase cell II geometry as implemented in Kassiopeia [131]. Image from private communication with Penny Slocum.

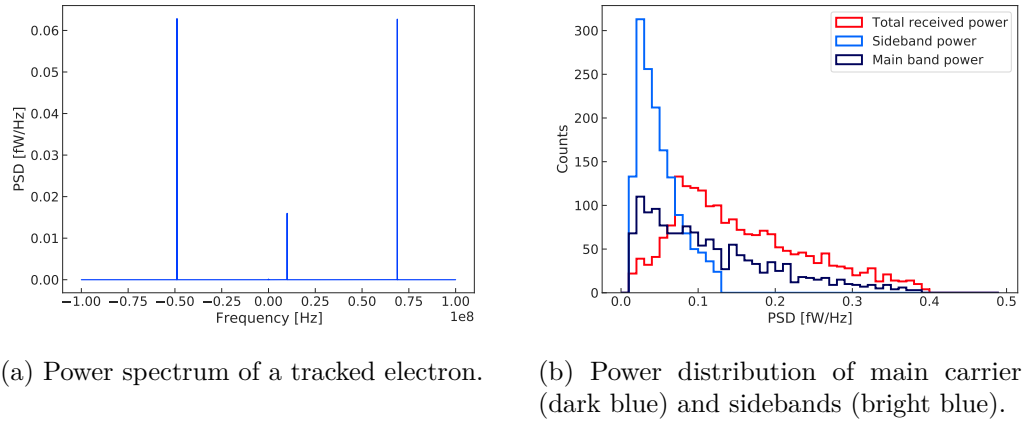


Figure 8.6: Power distributions (b) of simulated 17.8 keV electrons in the Phase II waveguide with a single 300 mA trap. The distributions are composed of main carrier and sideband peak heights in power spectra as in (a) from 2500 electron tracking simulations performed with Kassiopeia.

- Analytical calculation:** With the field solutions for a cylindrical waveguide from [93], the coupled power to the TE_{11} mode is calculated for electrons at random radial positions and pitch angles (using equations from section 4.1.2). Because relative power corrections for pitch angles are implemented in Locust, it is sufficient to generate a power distribution for 90° electrons (figure 8.7) and use it to configure the signal power in the fake event simulations. Generating the signal power distribution with the help of analytical formulas is of course much faster compared to generating and tracking electrons in Kassiopeia and yields significantly superior statistical power.

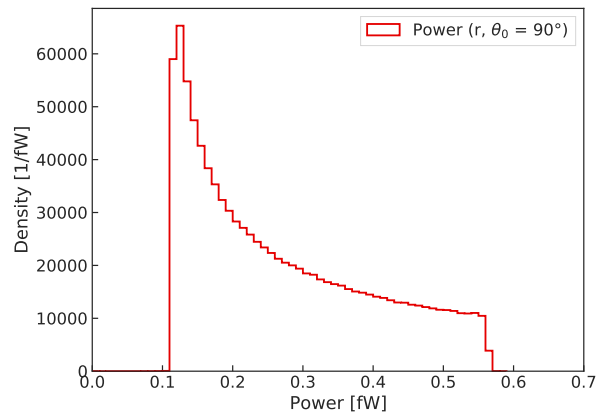


Figure 8.7: Main carrier power distribution from analytical calculation of the power in the TE_{11} mode for 17.8 keV electrons with a pitch angle of 90° at random radial positions $r \leq 5$ mm.

8.2. SIMULATION VALIDATION

In both approaches, the signal power has to be scaled by a gain to result in a realistic SNR distribution relative to the noise floor power. This was achieved by means of another parameter scan, which aimed at finding the best agreement of the reconstructed average track SNR distributions in simulated and reference data.

Electron start conditions

To use the semi-analytic power distribution from figure 8.7, the start conditions affecting pitch angles have to be chosen. Locust applies the main band power correction from figure 4.9 to the fake event signal power according to its associated pitch angle. The largest trap current in the Q300 trap configuration is 288 mA and the resulting trap depth keeps electrons with pitch angles $88^\circ \leq \theta_0 \leq 90^\circ$ trapped. As no sidebands are observed in this trap, it is safe to assume that main band powers below the maximum sideband power are not detectable anyway and don't need to be simulated. The pitch angle range in the simulations is therefore limited to $89^\circ \leq \theta_0 \leq 90^\circ$. The cosine of generated pitch angles follows a uniform distribution. In order to obtain a more realistic distribution, the longitudinal start position is set and the instantaneous start pitch angle θ_z corrected for the magnetic field increase between the trap center and the start z -position. Because electrons with a pitch angle θ_0 below 89° cannot be detected, the maximum z -position that needs to be simulated is the position that would cause an electron with $\theta_z = 90^\circ$ to be shifted to $\theta_0 = 89^\circ$. Figure 8.8 shows θ_0 as a function of the start z -position for electrons with $\theta_z = 90^\circ$ and $\theta_z = 89^\circ$. It was found that z can be limited to 3.1 mm.

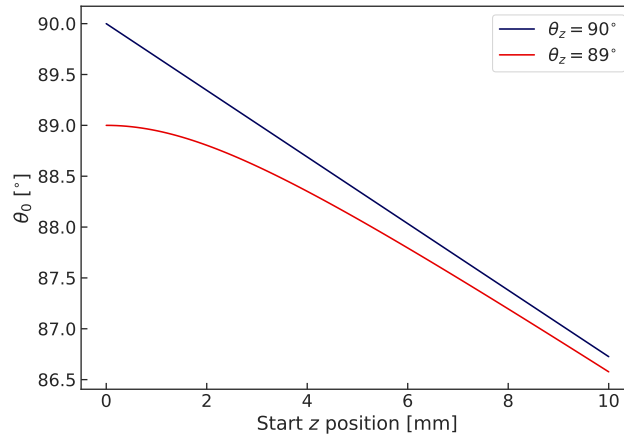
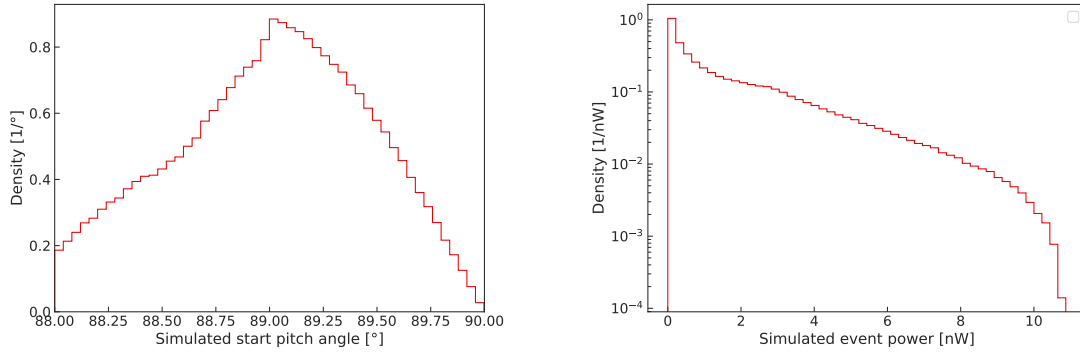


Figure 8.8: Pitch angle θ_0 of electrons with an initial pitch angle at $B(z) > B_0$ of $\theta_z = 90^\circ$ (dark blue) or $\theta_z = 89^\circ$ (red).

The shape of the pitch angle distribution resulting from the start z corrections in figure 8.9a is now almost realistic above 89° , where electrons are potentially detectable. The only missing contribution in the start pitch angle correction is the radial dependence of the magnetic field in a harmonic trap. In comparison to the longitudinal correction, the radial effect is expected to be small and it is therefore ignored here. The power distribution after pitch angle correction and multiplied system gain is shown in figure 8.9b.



(a) Pitch angle distribution θ_0 after correcting for start z -position ($z \leq 3.1$ mm).

(b) Power distribution from figure 8.7 after correction for start pitch angle θ_0 and gain.

Figure 8.9: Pitch angle and power distribution of 17.8 keV electrons after correcting the start pitch angles for the longitudinal z -position.

8.2.2 Comparison of fake and real events

To validate the fake event generator, 92,400 events were processed. The number of tracks per event in this validation study was set by a minimum trapped angle of 88° and a scattering α of 0.0064. The semi-analytic power distribution from figure 8.7 was used, together with a configured start pitch angle range of $89^\circ < \theta_z < 90^\circ$ and a longitudinal start position between 0 mm and 3.1 mm. The simulated events were processed with the same Katydid version (v2.17.0) and the same reconstruction parameter configuration (listing A.1) as the reference data. Only the number of fit points used for the spline fit to the averaged background power was set to a smaller number, because the background shape in simulated data is flat and requires a less detailed curve for its description. Figure 8.10 shows a comparison of the most relevant event properties:

- Track and event lengths,
- Average SNR of the first tracks and of all tracks,
- Maximum SNR of the first tracks and of all tracks.

8.2. SIMULATION VALIDATION

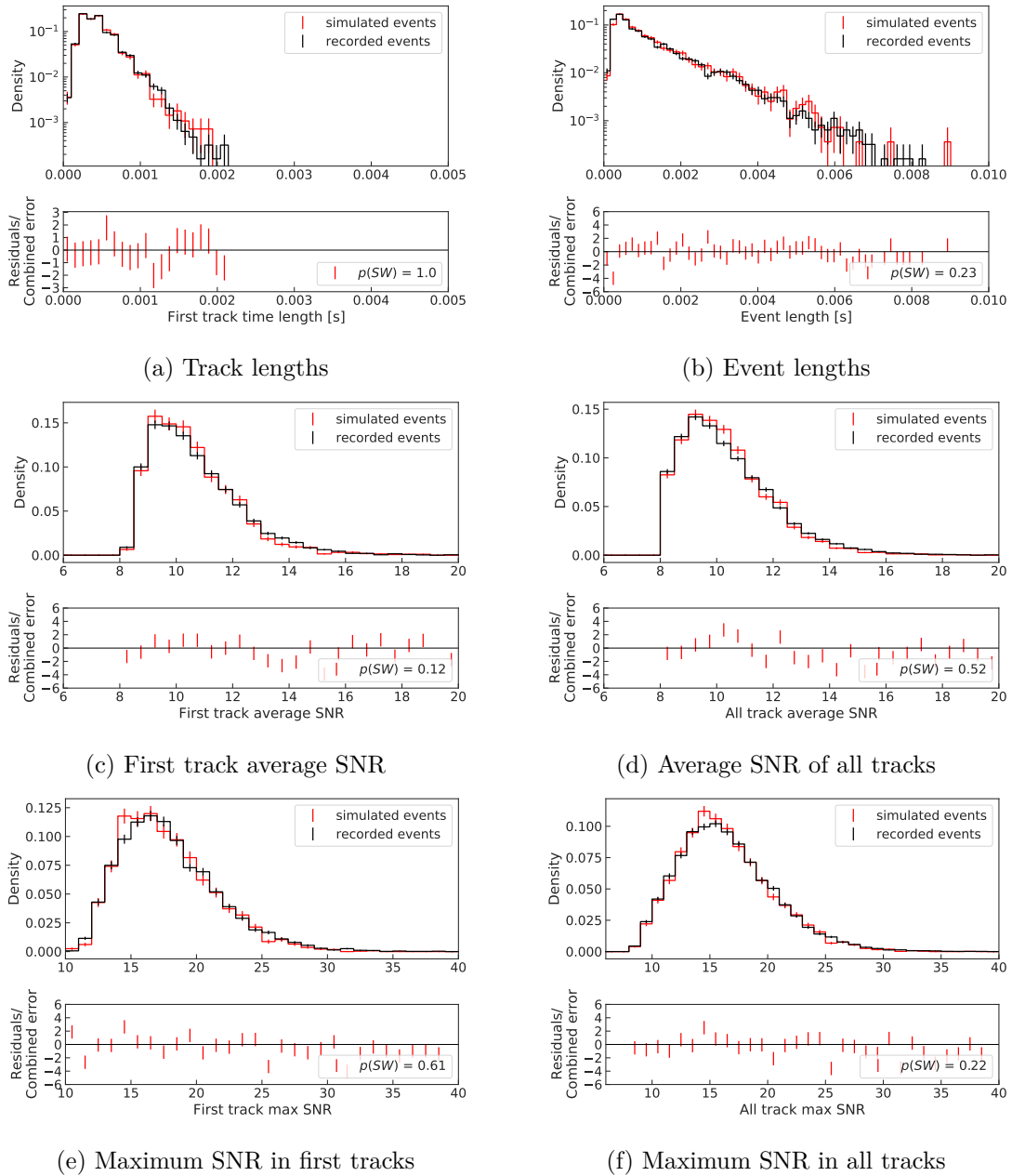


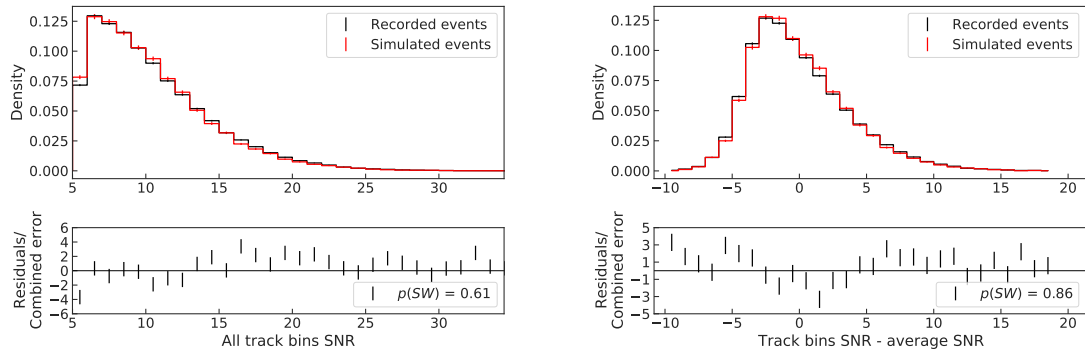
Figure 8.10: Comparison of reconstructed event properties in real and fake CRES signals.

The agreement of reconstructed event properties from simulated and recorded data is very good. The p-values of the Shapiro-Wilk tests¹ [132, 133], testing the hypothesis that the residuals are normally distributed, cannot rule out the equivalence of simulation and real data. At least with the statistical significance achieved here, it is not possible to tell from the distributions of reconstructed event properties whether a data set was recorded with the Phase II apparatus or simulated by Locust.

¹The Shapiro-Wilk test is a powerful test for normality [132].

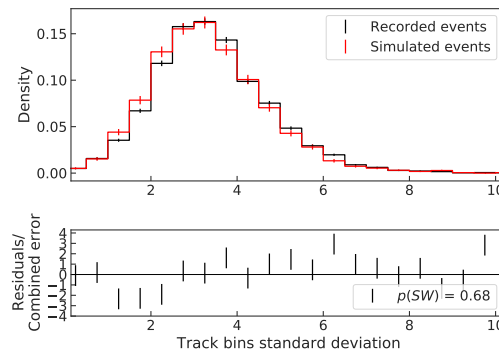
8.2.3 Power fluctuations in tracks

From section 8.2.2 it can be concluded that fake events can mimic real data on the track and event property level. To analyze the similarity of simulated and real data on a bin-by-bin level, the SNR distribution of the points clustered by the STF is analyzed here. Figure 8.11 shows the distribution of all track bin SNRs, the standard deviation of bin SNRs within single tracks, and the differences between a track's bin SNRs to its average SNR.



(a) SNR in all spectrogram bins identified as belonging to a track.

(b) Differences of all bin SNRs to the average SNR of the track they belong to.



(c) Standard deviations of bin SNRs in all tracks.

Figure 8.11: Comparison of SNR in fake and real events on the bin level. Since SNR fluctuations along tracks originate from noise power fluctuations and the DFT frequency response to the track positions, real and simulated tracks are indistinguishable on the bin-level.

The agreement of simulated and real data in figure 8.11 is excellent, which indicates that the main source of power fluctuations is common to real and fake events. Apparently, their dominant sources are the angle and the relative position at which tracks cross the frequency bins in combination with the fluctuation of the noise power that is added to all bins. As the shape of the noise power and the track slope distribution were carefully chosen to agree between simulation and real data it is not surprising that the resulting

8.3. FAKE EVENT STUDIES

power fluctuations are similar. Even the SNR standard deviation spread sorted by average track SNR (figure 8.12) doesn't show a difference beyond statistical fluctuations.

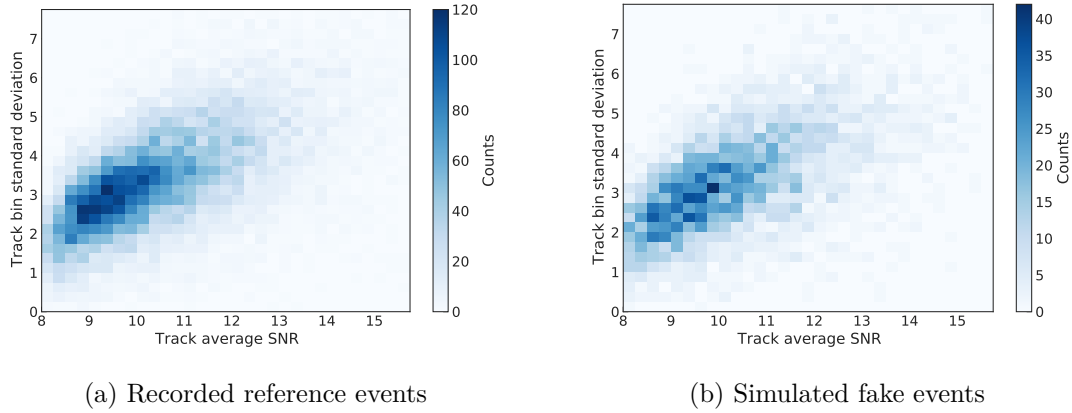


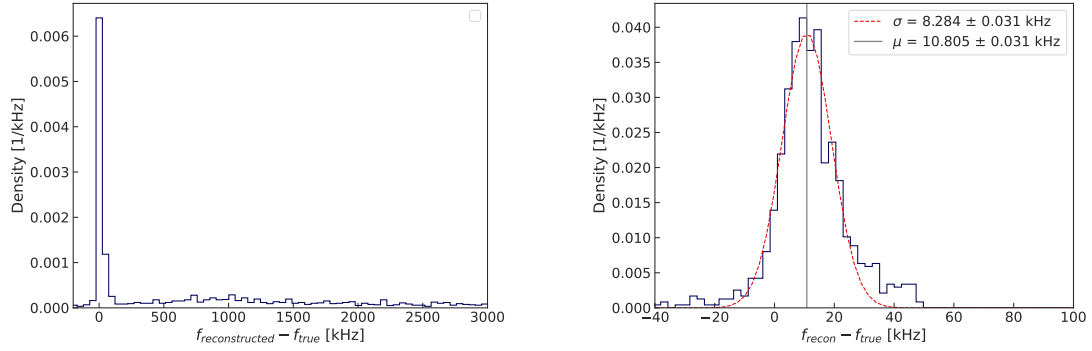
Figure 8.12: SNR standard deviation in tracks vs. average track SNR.

8.3 Fake event studies

For each fake event simulation, Locust writes out two files. One contains the time-domain samples, the other the properties of the signal that was generated (start frequency, power, slope and more). After the similarity between fake and real events has been established the fake events can be used for studying the detection and reconstruction performance. The goal is to determine the absolute detection efficiency and its dependence on various event properties, such as SNR, length, and slope, as well as to quantify the accuracy of reconstructed properties. The most important property to reconstruct precisely and accurately is, of course, the event start frequency.

8.3.1 Start frequency reconstruction

The accuracy of the reconstructed start frequency presumably depends on other event properties. For example, if the signal power is very high, it is less likely that the STF misses the first bin of a track. The typical slope of a CRES event in Phase II is 350 MHz/s. Tracks with this slope cross approximately 0.6 frequency bins per time slice (for a time slice width of $40.96 \mu\text{s}$ and a frequency bin width of 24.414 kHz). For a track starting in a bin center, the expected start frequency error from a missed first bin is therefore 14 kHz, which corresponds to an energy error of 0.3 eV. To get a sense of the distribution of start frequency errors under Phase II conditions, a realistic fake event data set like the one used in section 8.2 was generated and a comparison of the reconstructed and the true start frequencies was made. The result is shown in figure 8.13.



(a) Fake event start frequency reconstruction error. Events in the wide floor were only detected after the first track had ended. (b) Start frequency reconstruction error of fake events with a correctly identified first track.

Figure 8.13: Fake event start frequency reconstruction error.

From figure 8.13a it can be seen that the start frequency error exhibits a sharp peak at 0 kHz and a wide tail above 100 kHz. The events contained in the sharp peak are events with a correctly identified first track, while the events in the tail were only reconstructed starting from their second, third, or even later track. $(61 \pm 1)\%$ of the detected events were detected after the first track had ended. The start frequency error of the correctly identified first tracks can be roughly approximated by a Gaussian distribution with a mean of $\mu = (10.81 \pm 0.03)$ kHz and a standard deviation of $\sigma = (8.28 \pm 0.03)$ kHz (figure 8.13b). The systematic offset from 0 kHz proves that the first bin of a track is often missed. This is confirmed by the asymmetric start time error distribution in figure 8.14. Since the magnetic field used in the energy-to-frequency conversion in Phase II is calibrated from the frequency position of the $^{83\text{m}}\text{Kr}$ K-line, a systematic shift of the reconstructed start frequency is incorporated in the calibration and will not affect the tritium spectrum. Only the width of the start frequency error is relevant. The standard deviation of the error is approximately equal to only one third of the frequency bin width. This can be explained by the fact that the start frequency of a track is calculated from the fit of a linear line and its intercept with the reconstructed start time. As the average track crosses 0.3 frequency bins per half time bin, the width of the error is consistent with expectations and results in a decay energy uncertainty of 0.17 eV. This uncertainty is sub-dominant to the pitch angle effects described in section 4.1.1. However, in Phase IV the energy resolution requirement is much smaller and the start frequency determination must therefore be more precise. This will be achieved in a phase-space based analysis, in which the currently unused phase information will help to determine the start times of the tracks more precisely.

8.3. FAKE EVENT STUDIES

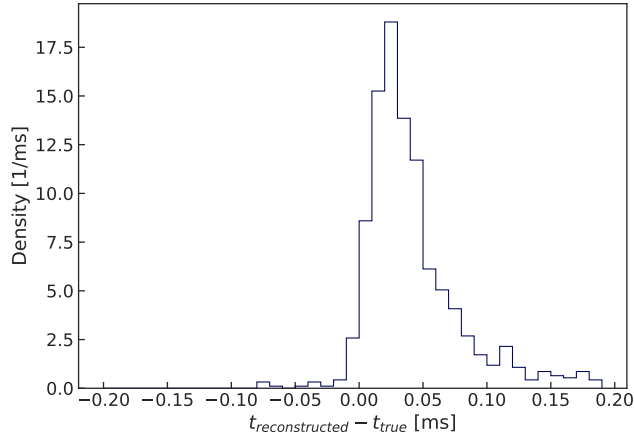


Figure 8.14: Start time reconstruction error for fake events with a correctly identified first track.

8.3.2 Absolute reconstruction efficiency

Of the 92,400 events with $\theta_0 \geq 89^\circ$ that were generated in section 8.2, only 2683 ($(2.9 \pm 0.06)\%$) were reconstructed, which shows that the Phase II experiment fails to achieve the required detection efficiency ($\geq 5\%$ of all decays) that would enable the accumulation of the target exposure for the Phase IV neutrino mass experiment. Several factors contribute to this - expected - insufficiency of the Phase II prototype apparatus:

- The power coupled to the waveguide steeply decreases with radius and pitch angles $\theta_0 < 90^\circ$. Only electrons on a trajectory with a radial position of $r \leq 3$ mm can be detected. In addition, the observation axis is aligned with the longitudinal trapping oscillation, leading to a large frequency modulation due to Doppler shifts, even for small angular deviations from 90° . In the future phases of Project 8, the emitted power will be detected by antennas that are arranged cylindrically around the CRES cell, to mitigate these effects.
- The power cuts applied by Katydid (section 7.5) are harsh to keep the background rate below one false event per 100 days of tritium operation. Lifting the power cuts to allow less than one false event per day (90% confidence) would increase the detection efficiency for real events by about 50%. The balance between background level and event efficiency is an optimization task for future phases and the optimum will depend on the fraction of the emitted power that is collected by the antenna array.

- The pressure was optimized to maximize the tritium count rate. Because the decay rate per volume increases linearly with pressure, the optimum lies at high pressures at the cost of the detection efficiency. In future phases, the pressure will be set to optimize the neutrino mass sensitivity, taking the impact of track lengths on the start frequency precision into account (see section 3.1.3).

8.3.3 Detection efficiency dependence on different event properties

To study the detection efficiency dependence on SNR, a fake event data set was generated with a uniform SNR distribution ranging from 0 to 25. This requires switching pitch angle corrections off and instead setting the number of tracks per event to follow a geometric distribution. As pitch angle changes are small and event power cuts are applied to the first tracks prior to any scattering, it is not expected that neglecting pitch angle corrections affects the efficiency dependence on SNR. All other event property distributions were configured as in section 8.2. Like Psyllid, Locust writes data to Egg-files. These files can be processed by Psyllid’s egg-reader node. A frequency-transform node produces the frequency-domain data, allowing the fake events to be processed with the same trigger implementation that is used to record real data (see section 6.3). This way, the data originally generated by a Locust simulation is written to a second, triggered Egg-file. The acquisition start and end times are logged in the file header, which allows for analysis of the true event start times relative to the time of triggering. Figure 8.15 shows the true start times in acquisition for the trigger configuration that was optimized in section 6.4.2.

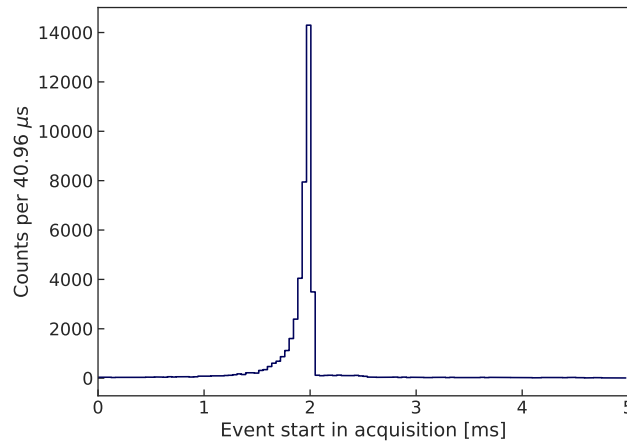
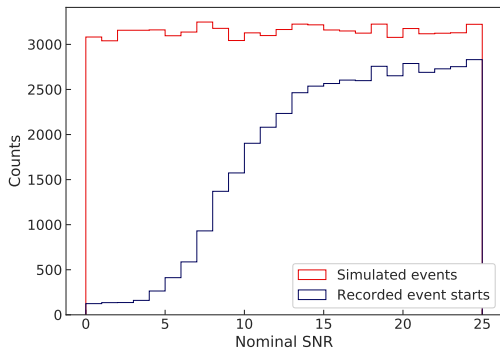


Figure 8.15: True start times of fake events in triggered acquisitions written by Psyllid. The trigger configuration used to process this fake event data set is identical to the configuration optimized to record tritium data (section 6.4.2).

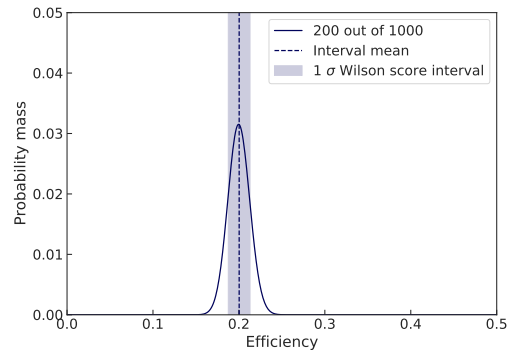
8.3. FAKE EVENT STUDIES

Figure 8.16 illustrates how the SNR dependence of the efficiency is extracted from the reconstruction of generated fake events. In this process, trigger efficiency can be defined in different ways:

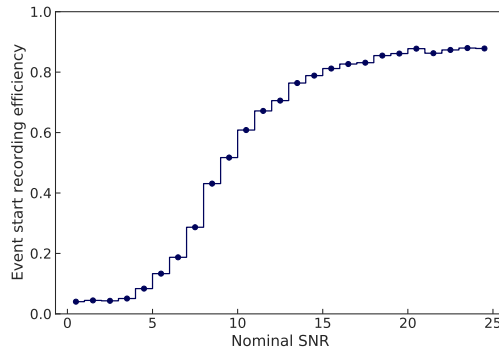
1. Because the event reconstruction cuts events that start earlier than 0.5 ms into an acquisition, events will only be counted as detected by the trigger in this work if their start point was recorded and the start time in the acquisition was > 0.5 ms.
2. A stricter definition could be to require that an event starts before and ends after the time of triggering. In this case, it would be very likely, though not guaranteed, that it was the event and not noise that triggered the acquisition.
3. Another interesting definition is the first track trigger efficiency, which only counts acquisitions in which the first track crosses the time of triggering.



(a) Nominal SNR distribution of all simulated fake events (red) and simulated events that were recorded (blue) by the Psyllid trigger.



(b) Detection efficiency for the example of 200 detected out of 1000 simulated events. The indicated uncertainty corresponds to the 1σ Wilson score interval [134, 135].



(c) Recording efficiency with uncertainty vs. nominal fake event SNR.

Figure 8.16: SNR dependence of the trigger (recording) efficiency for fake events. The nominal simulated SNR is given by the ratio of configured signal power to noise power. The SNR efficiency is taken to be the mean of the Wilson score interval for a binomial distribution with $p = \frac{N_{\text{detected}}}{N_{\text{simulated}}}$.

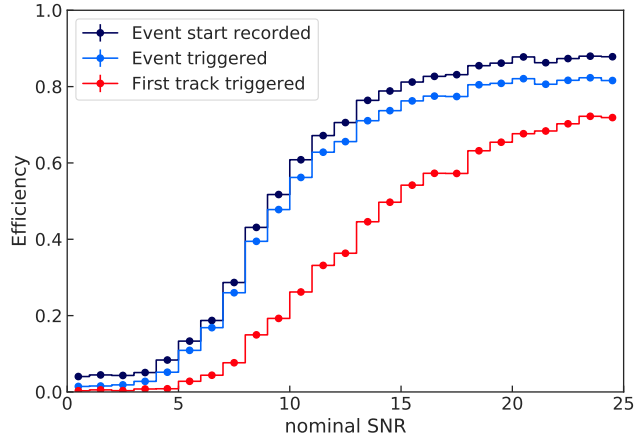


Figure 8.17: Event start recording (blue), event trigger (bright blue) and first track trigger efficiency (bright red) vs. nominal SNR.

Eventually, it is only relevant that the event starts were recorded. Nonetheless the efficiencies of all three definitions are shown in comparison in figure 8.17. As expected, the recording efficiency converges to an offset from zero that corresponds to the fraction of recorded time from noise triggers ($1 - \text{data reduction}$). In the low-SNR limit, the probability for events to cross the time of triggering corresponds to the probability that an event is ongoing at any random point in time. The efficiency only starts to increase significantly for an $\text{SNR} \geq 5$. The reason why the efficiency curves do not converge to 1 even at very high SNRs is that a significant fraction of events is too short to be detected. The trigger configuration (section 6.4.2) is based on the assumption that events result in at least two frequency mask triggers (FMT-triggers). Events that are shorter than two time bins are only recorded if a noise fluctuation causes the second FMT-trigger. Therefore, the trigger efficiency dependence on event length (figure 8.18) exhibits a sharp roll-off towards short lengths. In general, the trigger efficiency has a dual dependence on event length and SNR: Long events are more likely to cause a trigger even when their average SNR is relatively low and high-SNR events can still go undetected if they are short. Figure 8.19 shows the two-dimensional dependence of the recording rate (fraction of recorded event starts) on simulated SNR and event length. It can be seen that the recording efficiency quickly converges to 1 for simulated SNRs ≥ 10 and event lengths ≥ 0.5 ms. This behavior indicates that no other event properties in the simulated data adopt values that significantly diminish the trigger efficiency.

8.3. FAKE EVENT STUDIES

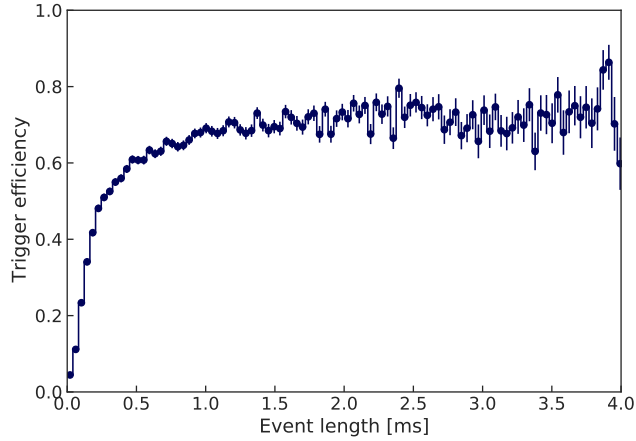


Figure 8.18: Dependence of the event start recording efficiency on the simulated event length.

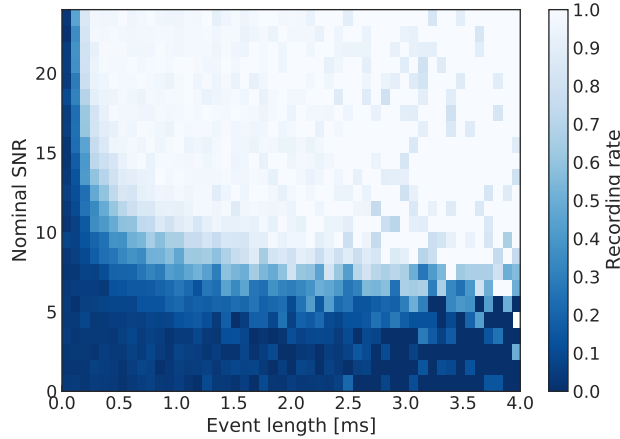


Figure 8.19: The fraction of recorded event starts strongly depends on the nominal SNR and the full event length. For SNRs > 10 and event lengths > 0.5 ms the trigger efficiency quickly converges to 1.

A similar behavior is expected for the event reconstruction efficiency. Here, an event counts as reconstructed if any of its tracks was found by Katydid and if it was not removed by the power cuts. Because the cuts apply different power thresholds to events depending on their length, the reconstruction efficiency, like the trigger efficiency, depends on the combination of the two properties (figure 8.20).

Unlike the Phase II trigger, the event reconstruction does not tolerate the misidentification of noise fluctuations as triggers. It is therefore unsurprising that the fraction of reconstructed events is significantly lower than the fraction of triggered events. Figure 8.21 compares the trigger and reconstruction efficiency. While in section 6.4.2 the optimized trigger configuration removed about 20% of events that could have been reconstructed,

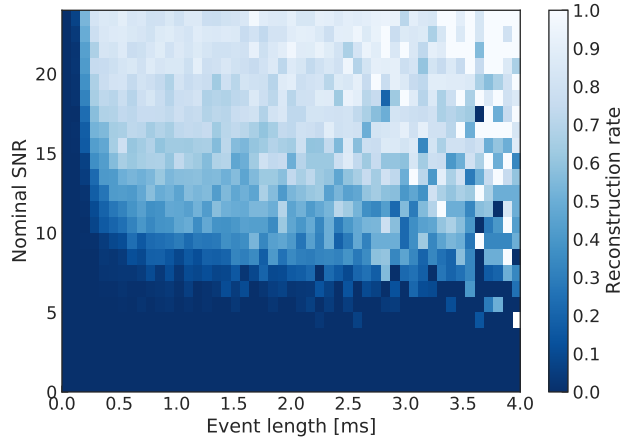


Figure 8.20: The dependence of the event reconstruction efficiency is qualitatively similar to figure 8.19. Since the event reconstruction process only accepts event candidates with signatures that are very unlikely to originate from noise, the reconstruction efficiency is generally lower than the trigger efficiency.

the detection efficiency in figure 8.21 is almost exactly equal to the reconstruction efficiency. The reason for this apparent discrepancy is that the power cuts in the event reconstruction were tightened after the trigger optimization studies. The time-domain data recorded in Phase II thus contains many more events than those reconstructed so far. If in the future a better algorithm is found, or the background requirements are loosened, the number of events in the tritium spectrum can be significantly increased ($\sim 50\%$).

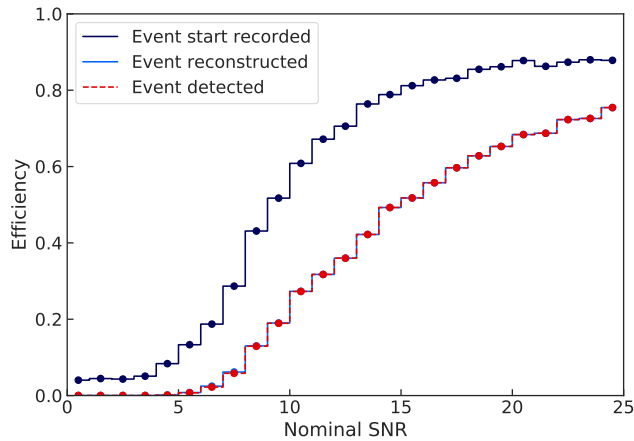


Figure 8.21: Comparison of the trigger and the reconstruction efficiency. The detection efficiency (fraction of events that is recorded and reconstructed (dashed bright red)) closely follows the reconstruction efficiency (bright blue). This shows that the trigger is not removing events that could have been reconstructed.

8.3.4 Mapping reconstructed to simulated event SNR

The fake event data set from section 8.3.3 can furthermore be used to understand how simulated SNR translates to reconstructed average first track SNR (figure 8.22). Because the SNR in track bins depends on the angle and the position of the track in the bin, the variation of the reconstructed SNR is expected to be large. In addition, bins can be identified as belonging to a track even if they only contain a fraction of the total track power from one time slice. This biases the average SNR towards low values. On the other hand, bins can only be considered in the clustering process if their SNR is larger than the threshold used to generate the sparse spectrograms, which introduces a bias in the opposite direction. Which effect predominates can vary from track to track.

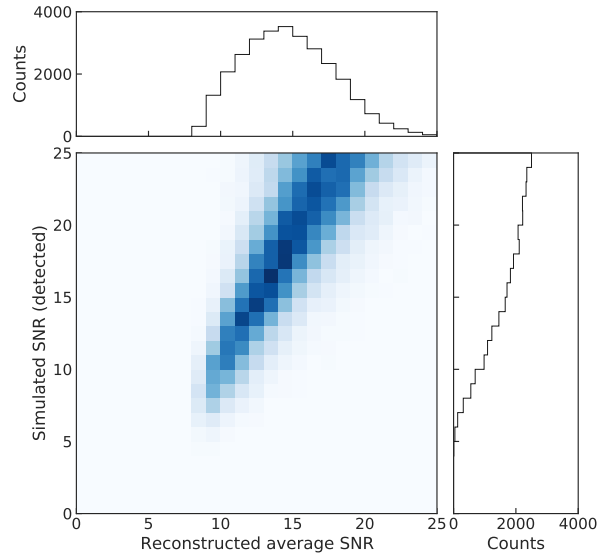


Figure 8.22: 2D histogram of the nominal (right) and reconstructed (top) first track SNR of detected fake events. After normalization along the simulated SNR axis, the 2D matrix becomes a response matrix, mapping simulated to reconstructed SNR.

The two-dimensional matrix in figure 8.22 can serve as response matrix to unfold a reconstructed power distribution and draw conclusions about the underlying true signal powers. To this end the Python package PyUnfold [136] is used to deconvolve signal smearing and efficiency via iterative Bayesian unfolding. As shown in figure 8.23, this method is able to conclude from the shape of the reconstructed SNR distribution that the underlying true distribution was uniform. Only below a nominal SNR of around 10 the uncertainties on the unfolding result start to grow. Below an SNR of 5, virtually no information is present in the population of events that were detected by Katydid and no conclusion about the simulated truth can be drawn.

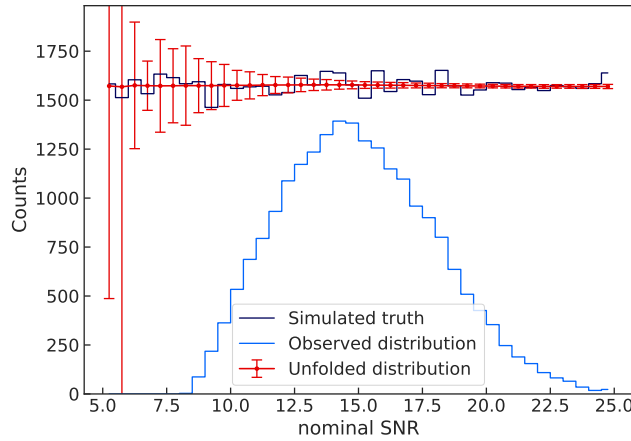


Figure 8.23: The shape of the simulated SNR distribution (dark blue) is correctly recovered by unfolding the reconstructed SNR distribution of detected first tracks (bright blue). The unfolding result with uncertainty (red) is calculated from multiplying the reconstructed SNR distribution with an unfolding matrix obtained using the PyUnfold package [136].

8.3.5 Detection efficiency dependence on slope

With the optimized trigger and reconstruction parameters, the detection efficiency should be independent of the track slopes, as long as they are in a realistic range (100 \rightarrow 1000 MHz/s). For normally distributed track slopes of (352 ± 54) MHz/s (figure 8.3b) this appears to be true, as both the trigger and reconstruction efficiencies converge to 1 for long events with a high SNR. However, it is worth testing whether this holds over a larger range of slopes, because high slope values may appear in frequency ranges in which electrons lose additional power to waveguide modes other than the TE₁₁ mode. To study the efficiency dependence on track slopes a fake event data set was generated with a uniform SNR distribution between 0 and 25 and a uniform slope distribution between 0 MHz/s and 2000 MHz/s. All other event property distributions were identical to those in section 8.2. The resulting first track detection efficiency dependence on slope can be seen in figure 8.24. The efficiency sensitivity to slope is clearly less pronounced than its dependence on SNR or event length. The trigger and reconstruction efficiency decrease slowly with increasing slope. In addition, the reconstruction efficiency exhibits a drop towards very low slopes, which can be explained by the minimum track slope cut of $s \geq 0$ MHz/s applied by the STF. If the slope is wrongly reconstructed to be negative, the track is discarded as event candidate. The difference between reconstructed and simulated slope is shown in figure 8.25. The sizes of reconstruction errors show that a fraction of slopes below 50 MHz/s may well be removed by the ≥ 0 MHz/s slope cut.

8.3. FAKE EVENT STUDIES

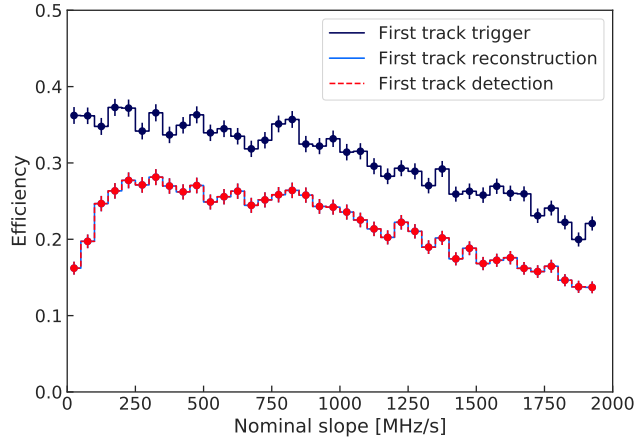


Figure 8.24: Slope dependence of the trigger (blue), the event reconstruction (bright blue) and the detection (bright red) efficiency. The roll-off in the reconstruction (and detection) efficiency towards small slopes is caused by a minimum-slope cut in the reconstruction process.

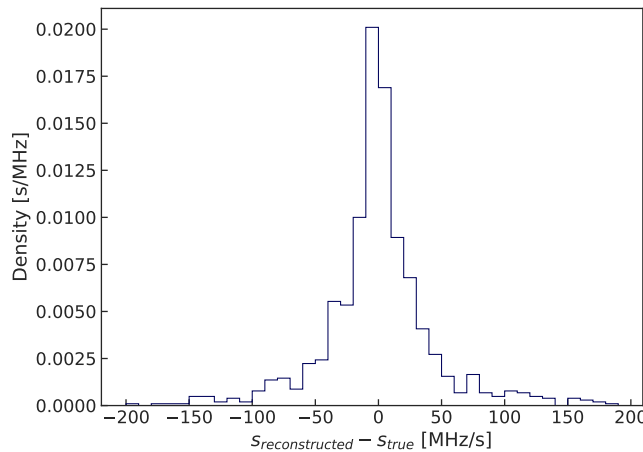


Figure 8.25: Slope reconstruction error for fake events.

The reason for the decrease of the detection efficiency for very high slopes can be found in figure 8.26. It shows a 2D histogram of the reconstructed track SNR and the simulated slope. Although SNR and slope were both generated from uniform distributions, a shape in the reconstructed SNR can be seen. The average reconstructed SNR decreases with increasing slopes, because the tracks cross multiple frequency bins per time slice. As a result, the power in a time slice is distributed among several bins. On the low slope side, tracks with a start frequency between bin centers do not have a high enough slope to cross the bin center within one time slice. Therefore, their maximum amplitude in the discrete Fourier transform remains low. All in all, it can be concluded, that the efficiency dependence on slope, in reality, is a dependence on the average SNR in track

bins (and a slope cut for very low slopes). For most of the tritium region of interest, this effect is irrelevant. However, section 9.2.2 will show that in some frequency regions mode resonances do cause some tracks to have very high slopes. As expected, the detection efficiency drops in these regions.

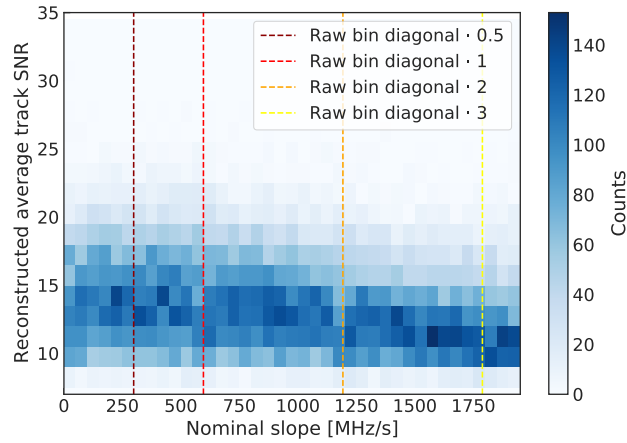


Figure 8.26: Reconstructed average track SNR vs. simulated slope. Towards very high slopes, the average reconstructed SNR visibly decreases, as track power is distributed over multiple frequency bins in each time slice.

9

Detection efficiency in Phase II

The goal of the Project 8 collaboration is to establish the CRES technique as a tool for recording a continuous tritium spectrum and extracting the neutrino mass. This not only requires a precise energy reconstruction for individual decay electrons, but also an exact determination of the relative detected event rates as a function of energy. The count rate at any given energy depends on the amplitude of the probability density function (PDF) describing the spectral shape and the detection efficiency for events at this energy. If there were no pitch-angle effects and magnetic field variations, there would be a one-to-one mapping of frequencies to energies. In this case, the spectral shape would be the product of the underlying PDF and the curve describing the detection efficiency as a function of frequency, or equivalently, energy.

The detection efficiency does not necessarily depend on the frequency itself, but rather on event properties that possess a frequency dependence. For example, the power of an electron signal varies with frequency since both the signal frequency and the emitted power scale with the kinetic energy of the electron. Furthermore, the noise temperature in the RF system is not perfectly flat with respect to frequency, which results in a frequency dependence of the SNR even for signals of constant power. Consequently, any CRES experiment with an SNR-dependent detection efficiency will always show a frequency-dependent distortion of the recorded spectrum. This distortion must either be eliminated, corrected for, or included in the analysis. The efficiency distortion is not the only frequency- (and energy-) dependent systematic effect. As will be shown in section 9.2.2, the parameters of the asymmetric scattering lineshape described in section 4.1.4 could also be frequency-dependent and lead to additional distortions or shifts of the spectrum. In contrast to the detection efficiency, the lineshape cannot be incorporated by simple multiplication, but instead has to be convolved with the tritium spectrum. While this chapter focuses on the investigation of the frequency and energy dependence of the detection efficiency in Phase II, the lineshape distortion will also be included in the analysis of the tritium data in section 10.2.

9.1 Detection efficiency impact on the tritium endpoint

The error introduced by a systematic uncertainty can be investigated by the means of Monte Carlo pseudo-experiments. To this end, random fake data¹ is generated from a model that describes the expected shape of the tritium spectrum (see section 2.6). Depending on the goal of a fake data study, systematic effects like efficiency or lineshape are included in the data generation model. The generated fake data is then analyzed using an accurate or erroneous model. Erroneous here means that either a systematic effect is completely neglected in the analysis or that the model parameters describing the effect are varied according to their uncertainty.

The four free parameters in the analysis are the number of signal events N_{T_2} and background events N_B , the spectrum endpoint energy E_0 , and the squared neutrino mass m_β^2 . Due to the low statistical power, it is expected that the tritium spectrum recorded with the Phase II apparatus is not sensitive to the neutrino mass with respect to current limits on m_β . Therefore, the main fit parameter of interest in this work is the endpoint energy, and the neutrino mass is fixed to $m_\beta^2 = 0 \text{ eV}^2$ unless otherwise stated. The data is binned and fitted by minimizing the negative log-likelihood

$$\mathcal{L} = - \sum_i (n_i \cdot \log(\mu_i) - \mu_i - \log(n_i!)). \quad (9.1)$$

The number of events in a bin n_i is assumed to follow a Poisson distribution with a mean μ_i equal to the tritium model amplitude integrated over the energy range covered by the bin. For minimization, the Python package `iMinuit` [137] is used, which provides a Python interface to the `Minuit2` [138] minimizer.

Repeating the pseudo-experiments many times gives insight into the robustness of the fits, reveals possible biases, and provides a distribution of fit results for each parameter. If the fake data is fitted with an accurate model (all systematic parameters are fixed and identical to the model used for data generation), the width of this distribution reflects the statistical uncertainty. For a large number of pseudo-experiments, the mean of the fit results converges to the model truth. If, on the other hand, the model used to fit the fake data differs from the model used to generate it, the results could be biased. Consequently, the coverage indicating the proportion of the time that the 1σ interval (defined by $\Delta\mathcal{L} = 0.5$) and the 90% interval (1.645σ) contain the true value could be reduced.

¹Unlike the fake time-domain signals studied in chapter 8, fake data here simply consists of a list of kinetic energies randomly drawn from a probability density function describing the tritium spectrum.

9.1.1 Undistorted and distorted tritium spectrum

To obtain a sense of the scale of the detection efficiency impact on the tritium endpoint fit result, fake data is generated from an undistorted and a distorted tritium spectrum model, and both fake data sets are fitted with the undistorted model. This reveals the magnitude of the introduced bias of the endpoint result in case efficiency is neglected in the analysis. Figure 9.1 shows fake data generated from the probability density function of an undistorted spectrum model with a true endpoint of 18.573 keV. To mimic Phase II conditions, the considered energy ROI covers the range from 2.39 keV below to 1.19 keV above the endpoint. The spectrum is convolved with a Gaussian with a standard deviation of 15.5 eV (energy resolution in the Q300 trap). Approximately 5000 events are generated from this model, about 100 of which are background events² that are evenly distributed over the entire energy range.

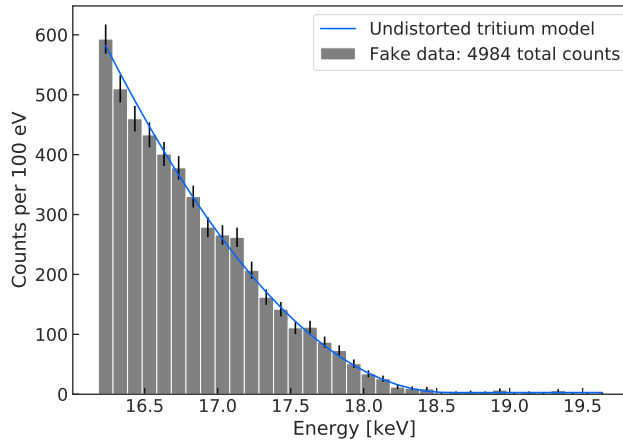


Figure 9.1: Smeared undistorted spectral shape (blue) and generated fake data (grey). The errorbars indicate the square root of the number of counts in each bin.

First, the fake data is fitted with the same model that was used to generate it. Figure 9.2 shows the 1, 2, and 3 σ contours of E_0 and N_B for one pseudo-experiment. The true values used for generating the data fall within the 1 σ contour. The parameter intervals that are returned by iMinuit correspond to the statistical uncertainties. In order to test the coverage of these intervals, the process of generating random data and fitting it with the undistorted model is repeated 1000 times. The statistical uncertainty is extracted from the spread of these fit results (figure 9.3a). The distribution of the fitted endpoint energies has a standard deviation of $\sigma_{E_0} = 23.8$ eV and a mean of $\mu_{E_0} = 18573.5$ eV. (90.4 \pm 0.9) % of the endpoint fit results contain the true (injected) value in their 90 % interval. The coverage of the 90 % background interval is (89.0 \pm 1.0) %.

²The background level in this study is an arbitrary choice and does not reflect the expected number of background events in the real tritium spectrum.

9.1. DETECTION EFFICIENCY IMPACT ON THE TRITIUM ENDPOINT

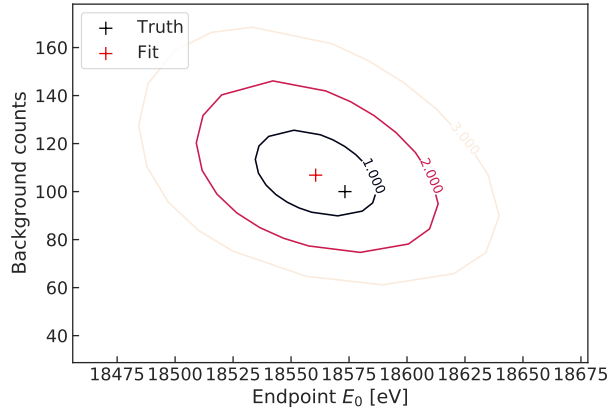


Figure 9.2: Contour plot of endpoint vs. number of background events for one pseudo-experiment. The generated data from figure 9.1 is fitted with the same model it was generated from. In this example, the results of both parameters are within 1σ of the model truth.

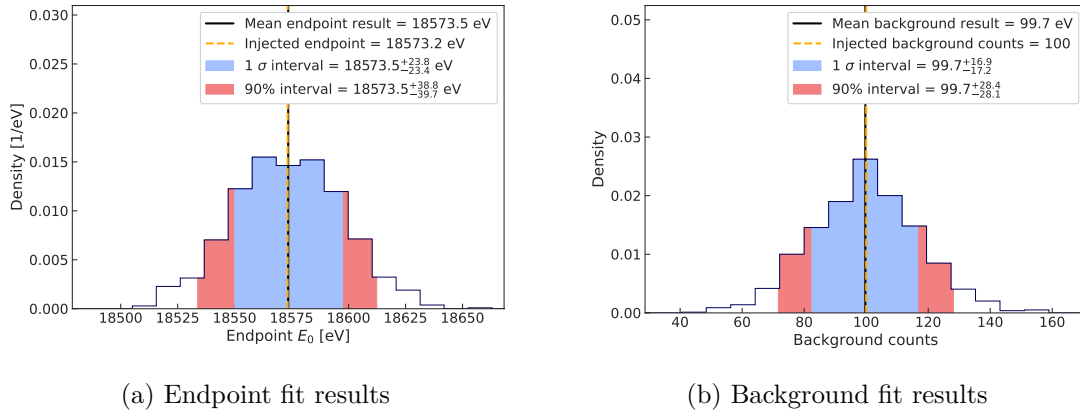
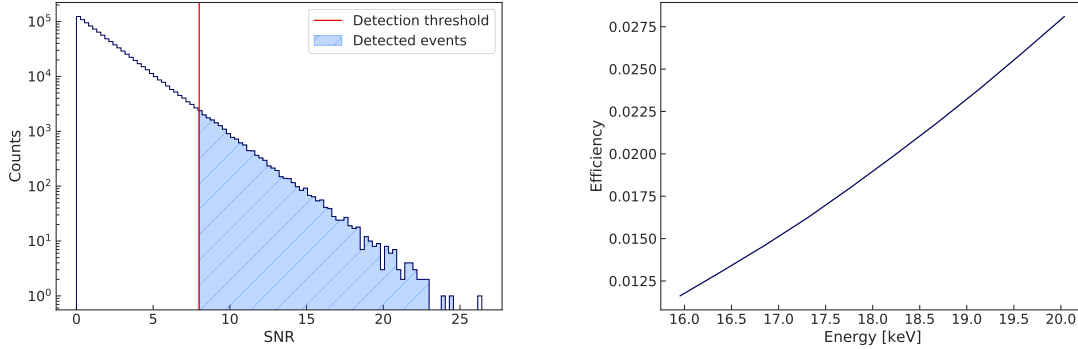


Figure 9.3: Endpoint and background fit results in 1000 pseudo-experiments. Data as in figure 9.1 was repeatedly generated and fitted with the undistorted model. The energy resolution was fixed to 15.5 eV. The intervals are given with respect to the mean. The 1σ and 90% intervals contain 68.3% and 90.0% of all results respectively, with the remaining results evenly split below and above the interval boundaries.

In a second study, a distorted model is created by multiplying the spectral shape with a hypothetical detection efficiency curve. The aim is to simulate the effect of neglecting the efficiency dependence on kinetic energy in the analysis. Since the background is expected to be caused by false events, the background shape is modeled as uniform and is only added to the spectrum after the spectrum has been multiplied by the efficiency. The hypothetical detection efficiency curve is produced by assuming an exponential event SNR distribution with a mean SNR of 5 and a detection threshold of $t_{\text{SNR}} = 8$. The SNR distribution is scaled by the received power dependence on kinetic energy

9.1. DETECTION EFFICIENCY IMPACT ON THE TRITIUM ENDPOINT

(figure 4.8) relative to the power emitted by 17.83 keV electrons. The detection efficiency then corresponds to the fraction of the distribution that is above the fixed detection threshold (figure 9.4).



(a) 10^5 samples from an exponential distribution with $\lambda = 5$. This distribution serves as toy model for a hypothetical SNR distribution. Only events with an $\text{SNR} \geq 8$ count as detected.

(b) The distribution in (a) is scaled by the received power (figure 4.8) relative to 17.8 keV. The resulting efficiency vs. energy curve rises by approximately 14 % per 1 keV.

Figure 9.4: A toy-model efficiency curve (b) is constructed by taking the ratio of samples above a hypothetical detection threshold to all samples from an exponential SNR distribution (a).

Just as before, fake data is generated, this time from the distorted model (figure 9.5) and fitted with the undistorted model. The 90 % interval of the endpoint results (figure 9.6a) is $E_0 = 18725.0^{+22.3}_{-23.6}$ eV. The mean of this interval is offset from the true value by 6.5 standard deviations (151.8 eV) and its coverage is reduced to 0 %. The large bias of the endpoint result is a consequence of the shape distortion induced by the efficiency curve. Since the efficiency increases with energy, the steepness of the tritium spectrum is reduced. A higher endpoint value also reduces the slope of the spectrum and the fit makes use of this to compensate the shape mismatch between model and data. The size of the endpoint bias is therefore expected to be correlated with the steepness of the efficiency curve. The endpoint offset in turn has a small effect on the result of the background fit. The coverage of its 90 % interval is reduced to (81.2 ± 1.2) %.

9.1. DETECTION EFFICIENCY IMPACT ON THE TRITIUM ENDPOINT

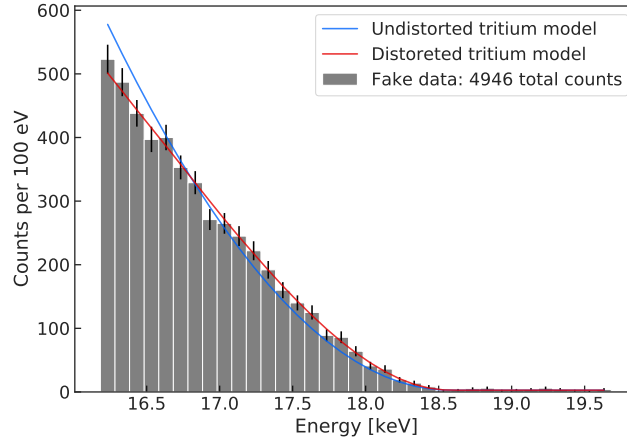
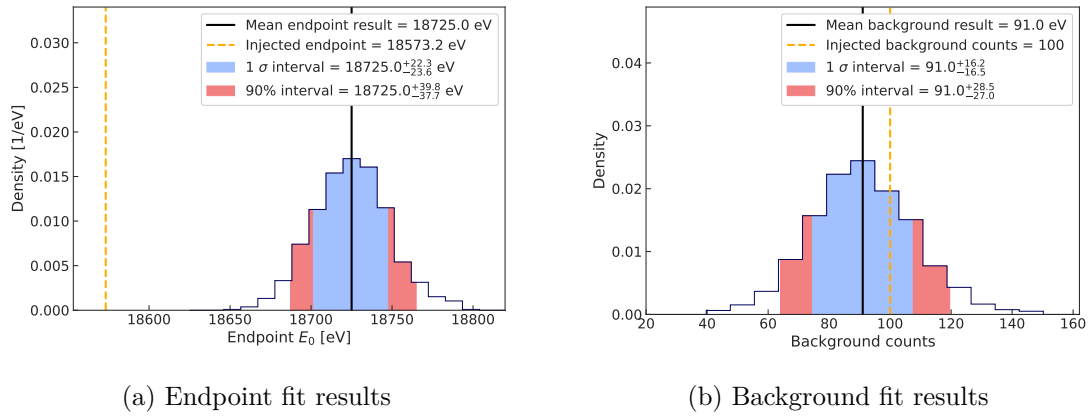


Figure 9.5: Undistorted (blue) and distorted tritium model (red). The distorted model is created by multiplying the tritium spectrum shape with the hypothetical detection efficiency curve in figure 9.4b. The fake data (grey) was generated from the distorted model.



(a) Endpoint fit results

(b) Background fit results

Figure 9.6: Distributions of fitted parameters in pseudo-experiments with fake data generated from the distorted tritium model and fitted with the undistorted model. The endpoint fit results show a strong bias and are shifted by 152 eV compared to the true model endpoint. The fitted number of background events is shifted downwards by less than one standard deviation.

9.1.2 Incorporating the detection efficiency in the spectrum fit

There are several approaches to incorporate the detection efficiency in the analysis. The section above has shown that neglecting the efficiency leads to a systematic shift of the fitted endpoint result. The simplest correction would be to subtract this shift from all fit results. A better approach is to use a distorted model to fit the data, as this allows for the propagation of systematic uncertainties on model parameters by randomly sampling them in Monte Carlo studies. In figure 9.7 it can be seen that fitting distorted data with the distorted model centers the distribution of fit results back on the true model values

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

and recovers an endpoint coverage of $(91.4 \pm 0.9)\%$. The third option to correct for the spectrum distortion is to weigh the events in the binning process and fit the resulting histogram with the undistorted model.

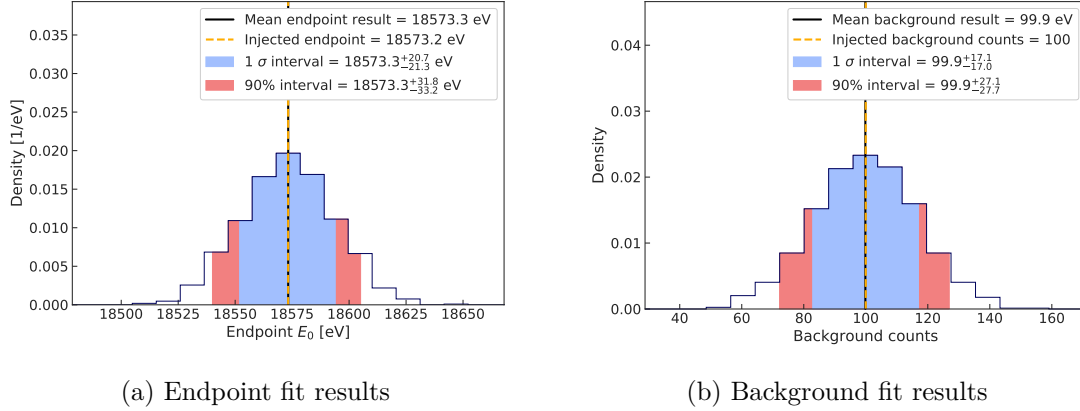


Figure 9.7: Results of pseudo-experiments with data generated and fitted with the same distorted tritium model. The coverages of the 90% intervals of the endpoint and the background results are restored to $(91.4 \pm 0.9)\%$ and $(89.9 \pm 1.0)\%$ respectively.

In this work the second approach is pursued, where distorted models are used to fit the recorded tritium data. But first, the true shape of the detection efficiency in the Phase II apparatus as a function of frequency has to be measured.

9.2 Studying frequency-dependent systematics

The examples in section 9.1 show the importance of a detailed knowledge of the variation of the detection efficiency to a level that keeps the resulting fit error within acceptable limits. The goal is the minimization of the systematic error induced on the endpoint fit result. In section 8.3, the efficiency dependence of the trigger and the reconstruction algorithm on various event properties was studied in simulations. In the analysis of a tritium spectrum, only the relative change of detection probability with frequency and energy are relevant. The detection efficiency is extremely sensitive to many experimental parameters and can therefore not be studied in simulation alone. Instead, it has to be measured directly, which can be accomplished in several ways:

- Using an electron gun as calibration source: An electron gun with a very narrow emission profile (in energy) would be the ideal calibration tool. Unfortunately, it is not possible to trap electrons that were emitted outside the magnetic trap unless they lose energy or change their angle relative to the magnetic field when crossing the trap. An energy loss would lead to an increased uncertainty of the kinetic energy. Only electrons that scatter elastically and not inelastically could preserve

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

their kinetic energy while changing their pitch angle. Whether this is feasible and produces a sufficient rate of trapped electrons with known energy will have to be tested in Phase III. In Phase II no electron gun is available and the trapping efficiency would be too small. Therefore, it is ruled out as calibration candidate.

- Using a β -decay calibration source: A β -emitting source gas with a well known spectral shape would be the ideal efficiency calibration source. For this source, similar properties as tritium would be required (for example, it has to be gaseous at the temperatures in the CRES cell) and the decay spectrum would have to cover the full tritium endpoint region (the discrete $^{83\text{m}}\text{Kr}$ spectrum does not fulfill this requirement). Such a source (other than tritium itself) is currently not known to the collaboration.
- Decreasing the magnetic field: In principle, the efficiency could be studied by decreasing the magnetic field and shifting the tritium spectrum down in frequency. In doing so, an energy range in the spectrum with a high event rate would now be detected in the same frequency range as the tritium endpoint before the shift. Because in Phase II the tritium event rate is low and the analysis energy range is large, this method would require a frequency shift of several 100s of MHz (corresponding to several keV). However, the power emitted by an electron depends on the kinetic energy and the magnetic field magnitude itself. As a result, a large frequency shift would change the SNR, which again affects the detection efficiency. Therefore, the use of this method is not possible without having to make significant corrections. In addition, the means for changing the magnetic field this drastically are not available in the Phase II apparatus.
- Using $^{83\text{m}}\text{Kr}$ electrons in combination with magnetic field scans: The Phase II apparatus is krypton-compatible³ and the $^{83\text{m}}\text{Kr}$ K-line is quasi-monoenergetic⁴. Because of its proximity to 18.6 keV it is ideal for calibration. By sweeping the background magnetic field up and down, the frequency of the 17.8 keV electrons can be scanned across the region of interest. The half-life of $^{83\text{m}}\text{Kr}$ is significantly shorter than that of tritium and the count rate correspondingly higher for similar gas densities. Therefore, a magnetic field scan of ≤ 100 MHz makes it possible to record an event rate vs. frequency profile with small enough statistical uncertainty to use it for the construction of a curve describing the relative efficiency changes with frequency.

³The cell temperature is high enough that gaseous krypton can be kept in the cell.

⁴The K-line is ~ 2 eV wide and therefore narrow compared to the energy resolution in the Q300 trap.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

While the latter approach also requires corrections for the energy dependence of the SNR, the magnetic field changes needed here to shift the frequency position of the K-line across the tritium region of interest are small compared to the absolute field. From all listed methods it was the most feasible for systematic error studies in Phase II. In the following sections, the field-shifting scans and their results are presented. In section 9.3 it will be shown, how fake event simulations can be used to extract the relative changes of SNR with frequency and how this knowledge is used to translate the K-line count rates to a tritium efficiency prediction.

9.2.1 Field-shifting solenoid scans

In a magnetic field scan, only the field magnitude should be changed while the field geometry has to remain unaffected. This way, the K-line position is shifted in frequency without modifying the trapping conditions. The field is scanned in steps, where the step size dB is kept small to ensure a good frequency resolution of the region of interest. Because it is not practical to scan the current of the superconducting NMR magnet, an additional copper coil winding was added to the Phase II apparatus (figure 9.8). This field-shifting solenoid (FSS) was inserted into the NMR magnet bore such that it encompassed the waveguide cell. Running current through the FSS coil slightly increases or decreases the total longitudinal field component depending on the current polarity. A FSS current of 1 A causes a magnetic field change of ~ 3 mT (0.3%), which shifts the frequency position of the K-line by approximately 85 MHz.

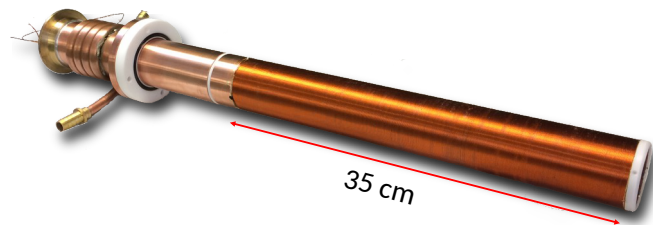


Figure 9.8: The field-shifting solenoid is a copper coil winding that was inserted into the NMR bore to surround the waveguide and serve as tool for scanning the background field over a few mT.

Scan procedure

The trap configuration used to record the final Phase II tritium data is the Q300 trap (using coils 1 to 4) at a depth that has the same effective average sampled magnetic field as the 300 mA single harmonic trap 3. To obtain a complete picture of the frequency dependencies in all traps, five FSS scans were taken in total: One in the Q300 configuration and one in each of the four individual harmonic traps.

The uncorrected FWHM of the 0th order (unscattered) ^{83m}Kr K-line peak in a 300 mA trap is approximately 2 MHz. Therefore, the FSS frequency step size was chosen to be exactly 2 MHz, which corresponds to FSS current steps of 0.023 A. It was decided to achieve a total count rate of at least 4000 events in each trap for $I_{\text{FSS}} = 0$ A, which yields a statistical uncertainty of 1.5%. This required a minimum run duration of 30 to 50 minutes, depending on the detection efficiency at the trap location in the waveguide. The same total run duration was then taken at all set currents.

Although the FSS is water-cooled, a current of several 10s of mA leads to a significant temperature rise within a few minutes. The FSS has a residual thermal coupling to the waveguide cell and terminator. Though the waveguide is passively cooled, a fast permanent change of the FSS temperature leads to a loss of the temperature stability of the detector insert. A temperature fluctuation of the cell changes the krypton freezing rate and the noise temperature, which in turn affects the event rate and the SNR of the CRES signals. Both effects have a strong impact on the event detection rate. To keep the temperature variation of the FSS to a minimum, the average squared FSS scan current was set sufficiently long (several hours) before the start of the scan to allow the cell temperature to re-stabilize at 85 K. After stabilization, the FSS scan was started with run durations at each current limited to 60 s. The FSS currents followed an interleaved sequence as shown in figure 9.9 to stabilize the average dissipated power, which kept the temperatures of the terminator and the cell stable to within 1 K. The current sequences were repeated many times until the accumulated total run time reached the target duration. Figure 9.10 shows the predicted temperature changes of the FSS during the first 10 iterations of a scan. Unfortunately, the thermal coupling of the FSS to the insert increased over time and eventually completely degraded temperature stability in the apparatus. In the end, the FSS had to be removed from the NMR bore. To be able to complete all scans before the FSS removal, the current step sizes in the trap 1 and 4 scans were increased to 20 MHz. Consequently, the analysis results in these traps have to be interpolated to the frequency positions that were left out. As these two traps together only contribute less than 30% to the total count rate in the Q300 trap (for $I_{\text{FSS}} = 0$ A), the additional uncertainty introduced by interpolation is not significant.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

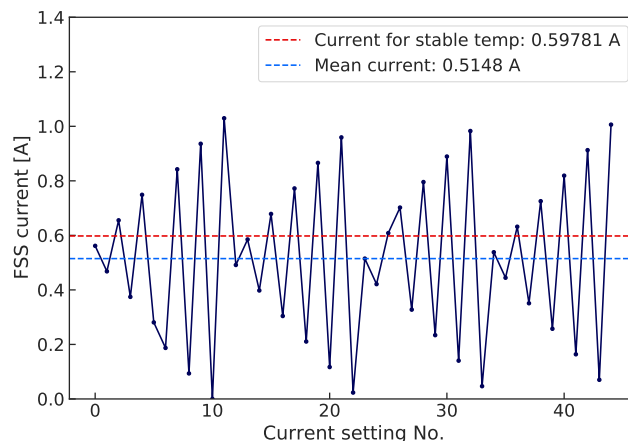


Figure 9.9: The alternating current sequence in a single FSS scan balances temperature fluctuations in the water cooled FSS coil and keeps the cell temperature within ± 1 K. Before each FSS scan, a current corresponding to the mean of all squared sequence currents was set to heat up the FSS to the average scan temperature. This way, large temperature fluctuations could be avoided during the scans.

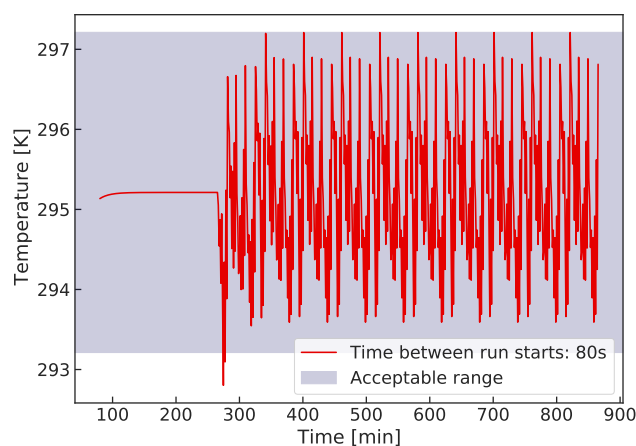


Figure 9.10: Predicted FSS temperature variation in the first 10 iterations of a single scan following the current sequence from figure 9.9.

Within one 60s-long run, the current and frequency position of the K-line is stable. Because the maximum frequency shifts (± 87 MHz) are larger than half the bandwidth of one ROACH2 channel, the central frequency of the recording channel is adjusted for each run such that it follows the K-line position. Figure 9.11 shows the expected peak position and central frequency settings during the FSS scan.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

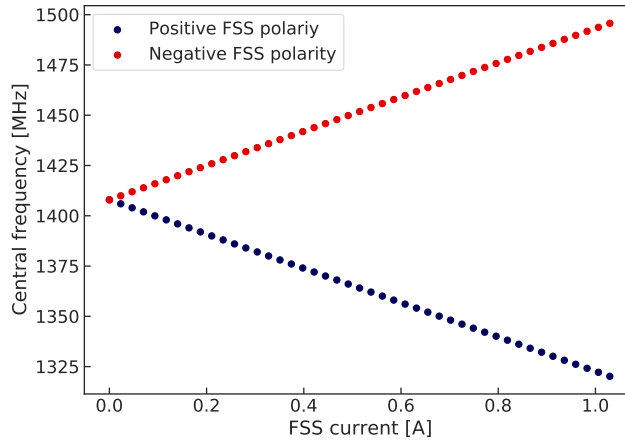
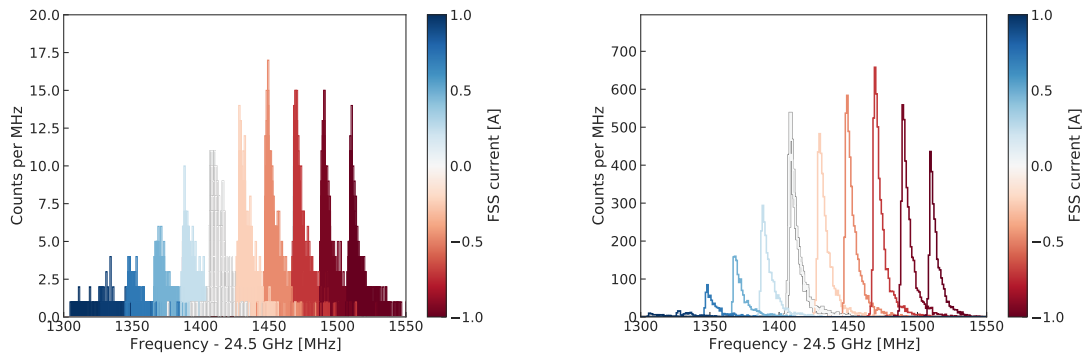


Figure 9.11: Position of the K-line in intermediate frequency (after mixing with 24.5 GHz) as a function of the current powering the FSS field.

All events that were recorded at the same current but during different iterations of a scan have to be assembled for the analysis. In this process, only events recorded when the cell temperature was within (85 ± 1) K are considered, in order to prevent temperature effects from impacting the analysis results. Figure 9.12 shows the start frequency histogram from the harmonic trap 4 scan before and after merging all sub-runs from all FSS currents.



(a) Start frequency histograms of all events recorded in the trap 4 FSS scan before merging by current.

(b) Start frequency histograms after merging runs by FSS current setting.

Figure 9.12: Start frequency histograms in the single harmonic trap 4 FSS scan before (a) and after (b) merging runs by FSS current. The FSS scan with trap 4 was taken with a frequency step size of 20 MHz.

9.2.2 Single harmonic trap scan results

The behavior of the following event properties are analyzed in this section:

- **Peak position:** Over the energy region of interest, the frequency-to-energy relation for 90° electrons is mostly linear. However, the reconstructed start frequency has various dependencies, like pitch angle, SNR, and slope. Because the start frequencies in Phase II are not corrected, their reconstruction error could vary with frequency. If the mis-reconstruction is constant with frequency, it would be equivalent to a mis-calibration of the magnetic field, leading to the same shift of the tritium start frequencies. This does not introduce a systematic error on the tritium endpoint result. If, on the other hand, the shift is not constant with frequency, the tritium spectrum could be stretched or compressed as a function of energy.
- **Count rates:** Under the assumption that the FSS does not affect the trapping efficiency or the SNR, the event count rates directly indicate the relative change of detection efficiency with frequency.
- **SNR:** The ratio of track power to noise power is the most decisive quantity in terms of detection efficiency. The event properties reflecting the SNR are the maximum SNR in any bin belonging to a track and the average SNR of all track bins.
- **Slope:** The slope of a track determines the angle at which the time-frequency bins are crossed by the CRES signal. A variation in slope may cause a variation of the distribution of track power fractions over different bins. In section 8.3.5 it was shown that this has an impact on the SNR and thus on the detection efficiency. The slopes are a result of the energy loss of the electrons in the waveguide. The presence of additional waveguide or cavity resonant electromagnetic modes can therefore diminish or enhance the detection efficiency.
- **Track and event length:** The longer a weakly radiating electron remains trapped, the more likely it is that its signal combined with a noise fluctuation exceeds the trigger threshold and is recorded. Similarly, the longer a track, the easier it is for the reconstruction methods to distinguish it from noise. The lengths of tracks and events are exponentially distributed. If the mean of this distribution depends on frequency and/or energy, the detection efficiency could be heavily affected. However, reconstructed track and event lengths have to be interpreted carefully. Because the event power cut in Katydid (section 7.5) applies different thresholds depending on the number of bins in the first track and the number of tracks in the event, the average reconstructed lengths can change as a result of a changing SNR.

- **Lineshape:** The tritium spectrum recorded with CRES is a convolution of the real spectrum with the lineshape (section 4.1.4). If the lineshape parameters are frequency-dependent, they have to be included in the tritium spectrum analysis.

As mentioned previously, FSS scans were performed in the Q300 multi-trap configuration as well as in each individual trap. The goal is to measure the frequency dependence of the efficiency and other systematics in the Q300 trap, but because this multi trap is a composition of four individual harmonic traps, analyzing the single-trap FSS scans helps in understanding its behavior. Therefore, the results of the single-trap FSS scans are presented first, before addressing the Q300 scan in section 9.2.3.

Start frequency peak position

As in section 4.2.1, all merged start frequency histograms are fitted with the lineshape model (4.21). This allows the determination of the un-scattered peak position as well as its width and the survival probability. Figure 9.13 shows the merged start frequency histograms including the fitted peak model for the harmonic trap 4 scan. In figure 9.14 the fitted peak positions in the trap 3 and trap 4 scans are plotted against the set FSS currents. The residuals of the linear fit to the relation of these two quantities are mostly contained within ± 0.5 MHz. Since the event properties are analyzed as a function of the peak position, nonlinearities in the frequency dependence on FSS current settings are irrelevant.

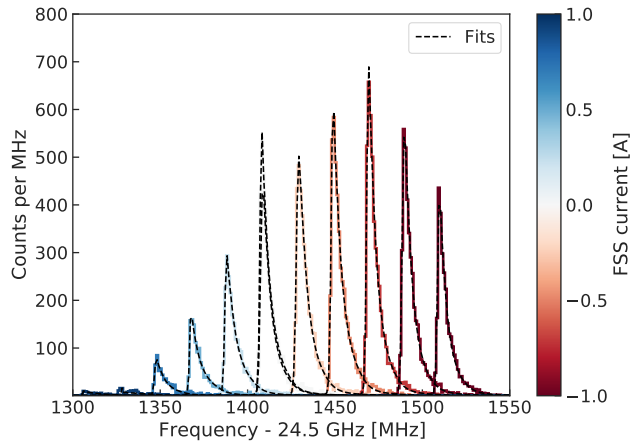


Figure 9.13: Fitted start frequency histograms in the trap 4 FSS scan. The K-line peaks in all FSS scans were fitted with a lineshape model (4.21) including 20 scatters. For simplicity the hydrogen fraction in the multi-gas model was approximated as 1.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

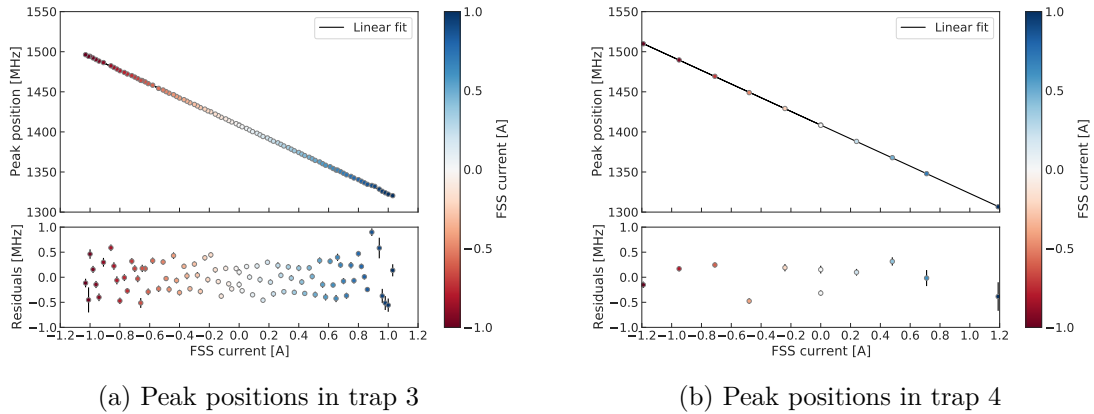


Figure 9.14: Linear fits to peak position (IF) vs. set FSS current in the trap 3 (a) and trap 4 (b) FSS scans. The trap 3 scan was taken with step sizes of 2 MHz while the trap 4 scan step sizes were 20 MHz. The large residuals can mostly be explained by the imprecision of the FSS current settings. At $I_{\text{FSS}} = 1$ A the residuals are increased due to waveguide resonances (see figure 9.18).

Count rates

The recorded event rate in the FSS scans is expected to be a result of the frequency-dependent detection efficiency. As the effect of SNR on the efficiency is predominant, the counted event rates should follow the SNR variation with frequency. The scanned frequency range is close to the upper edge of the ROACH2 bandwidth. Therefore, it was initially expected that the count rates would decrease with the rise of image noise that is not entirely rejected by the low-pass filter (section 5.1.2). In figure 9.15 the count rate is shown as a function of the fitted peak position in mixed-down frequency (IF). In contrast to these naive expectations, the count rates decrease towards low and high frequencies. It is furthermore apparent that the count rates in trap 3 and trap 4 exhibit distinct frequency dependencies. The reason for this behavior is the presence of modes and interference effects from reflections in the waveguide. Because the individual traps are located at different longitudinal positions in the cell, the interference effects lead to a different frequency dependence of the transmitted signal powers for each trap and therefore to a different average SNR. This also explains why even at $I_{\text{FSS}} = 0$ A the count rates in the traps are not the same (see section 4.2.2).

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

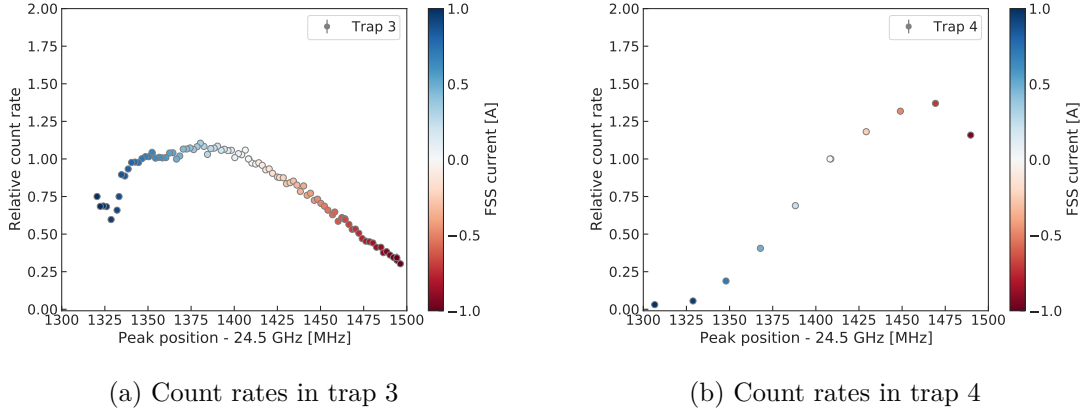


Figure 9.15: Detected event rate vs. frequency position in the trap 3 (a) and 4 (b) FSS scans. Count rates are given relative to the rate at the frequency position of the K-line at $I_{\text{FSS}} = 0$ A (1408 MHz in IF).

Signal-to-noise ratio

Figure 9.16 shows the mean of the first track average SNRs of all 17.8 keV events for each frequency position in trap 3 and 4. The shape of the SNR-mean vs. frequency correlates clearly with the count rates in figure 9.15. The SNR, however, does not follow the inverse shape of the image noise power from figure 5.4, which supports the conclusion that modes and interference are the main sources of SNR changes.

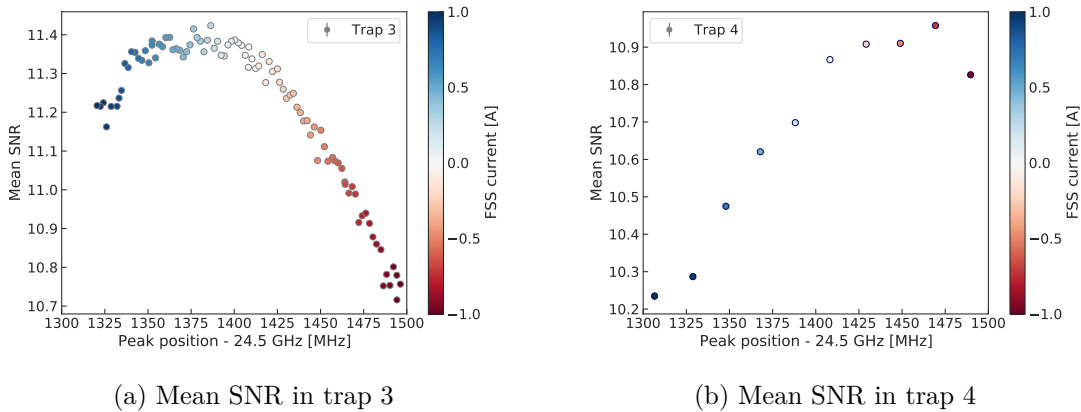
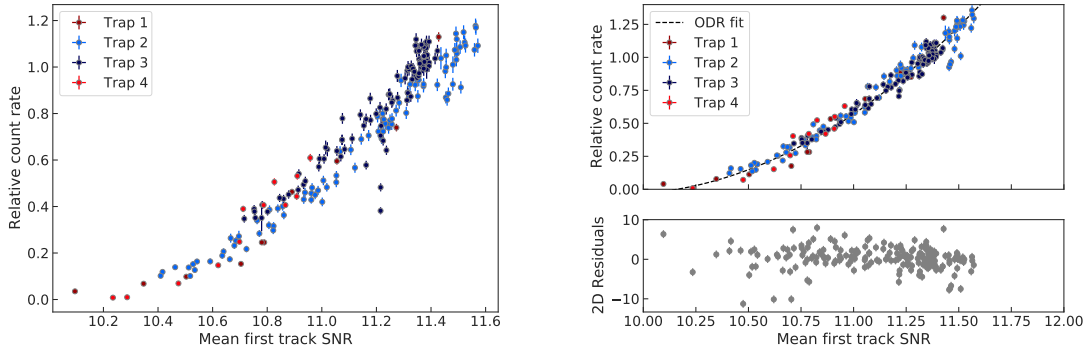


Figure 9.16: Means of the reconstructed first track average SNR distributions in the trap 3 (a) and 4 (b) FSS scans.

The hypothesis that count rates are a direct consequence of SNR is confirmed by figure 9.17. It shows the count rate to SNR relation in all four individual single-trap FSS scans. The absolute scales of the count rates are calculated by normalizing the measured rates in the scans to the rates in each trap at $I_{\text{FSS}} = 0$ A and then multiplying them by the fitted count rate dependence on trap depth from figure 4.18. While the count rates are close to falling onto the same curve, the absolute rates in the individual

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

traps differ slightly from each other. This is not surprising, as the frequency dependence of the efficiency was neglected in section 4.2.1. Figure 9.17a still strongly hints at a causality between SNR and count rates. Assuming causality, the relative trap rates can be calculated by minimizing the difference between the count rate to SNR relation in the four traps (figure 9.17b). The count rates retrieved this way are deemed more accurate as they do not rely on an interpolation that is not taking major frequency dependencies into account.



(a) Count rates vs. SNR in single-trap scans

(b) Polynomial fit to scaled count rates vs. SNR in single-trap scans

Figure 9.17: Count rate vs. mean first track SNR in all single-trap FSS scans (a). Because the scans were taken at different times with different krypton source activities, the absolute event rates were not comparable and were calculated from figure 4.18 instead. In (b), the summed squared differences between the four traps' count rate vs. SNR curves were minimized, yielding relative rates (relative to trap 3) at $I_{\text{FSS}} = 0$ A of 0.28 (trap 1), 0.86 (trap 2) and 0.42 (trap 4).

Slopes

The track slope reflects the total power emitted by a radiating electron in the waveguide. As described in section 4.1.2, it depends on the electron's kinetic energy, the radial position of the electron in the waveguide, and the presence of waveguide modes near the cyclotron frequency. The mean slope profile for all four traps is shown in figure 9.18. In all traps the slopes peak at the same frequency positions but with different magnitudes. Since high slopes lead to a decrease of the event reconstruction efficiency (see 8.3.5), it is expected that the count rates at high-slope frequency positions drop. This explains the count rate deficit in the trap 3 FSS scan (figure 9.15a) at 1328 MHz. Similarly, the count rate vs. frequency curve in trap 2 should exhibit pronounced local minima at the frequency positions where the slopes are maximal (at 1328 MHz, 1361 MHz and 1426 MHz). This is confirmed in figure 9.19.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

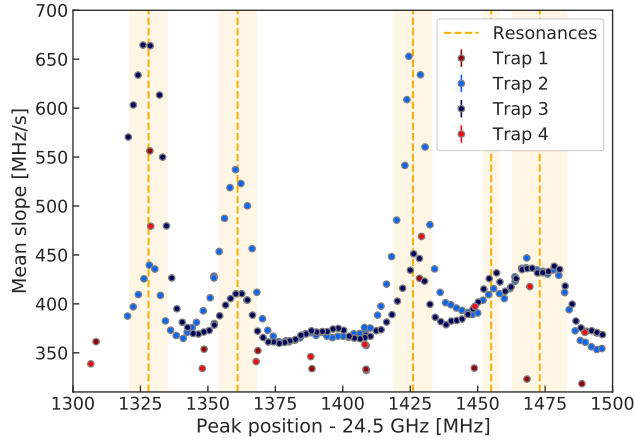


Figure 9.18: Mean of reconstructed slope distributions in all four single-trap FSS scans as a function of the fitted peak position. High slopes indicate the presence of resonances in the waveguide.

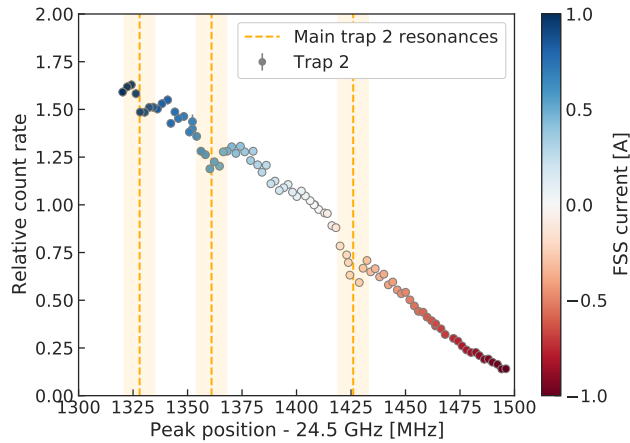
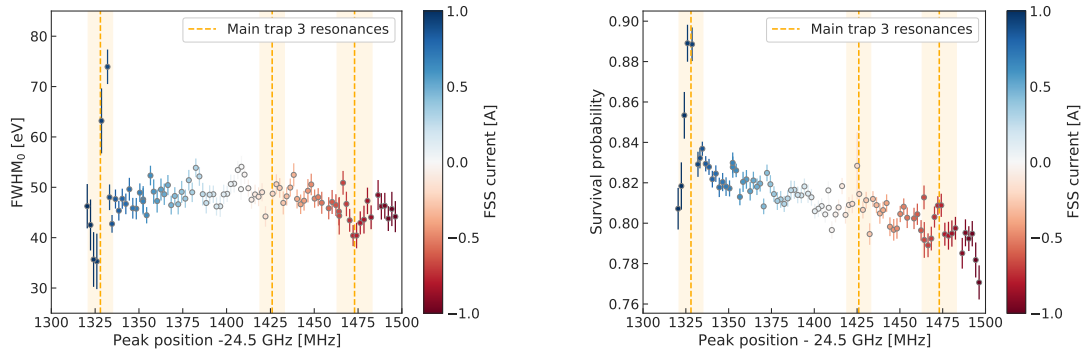


Figure 9.19: Count rates vs. frequency in the trap 2 FSS scan. While the count rate dependence on frequency is smooth in most regions, there are three distinct positions with count rate drops. These positions correspond to the frequencies of waveguide-mode-induced peaks in the slopes in figure 9.18.

Because tracks with higher slopes have a larger start frequency error if the first track bin is missed by the event reconstruction, slopes have an impact on the fitted peak position. This is why, in figure 9.14, the absolute residuals of the peak position fits are largest on the low frequency (positive FSS current polarity) side in trap 3 (at 1328 MHz).

Lineshape

The lineshape is composed of the underlying K-line spectrum smeared by a Gaussian with width σ_0 and N energy loss functions from scattered electrons (section 4.1.4). Without pitch angle correction, the width of the un-scattered peak corresponds to the frequency spread of the detectable pitch angle range. The power coupling to the transporting waveguide mode is weaker for trapped electrons with smaller pitch angles (figure 4.9). A change of SNR would affect the size of the detectable pitch angle range and thus lead to a change of the frequency resolution. In addition, the chances for the first track of an event to be detected are higher if its SNR is high. As a result, the size of the scattering tail relative to the main peak is expected to also show an SNR dependence. In figure 9.20a and figure 9.20b the FWHM of the un-scattered main peak (FWHM_0) and the survival probability (ratio of the number of events in the $(n + 1)^{\text{th}}$ and the n^{th} scatter) in the trap 3 FSS scan can be seen respectively. The frequency dependence of the lineshape parameters is visible but small compared to the variations of other event or data properties (like slope or SNR). Only at the frequencies where the mean slopes are large, the width and the scattering tail both differ drastically from their average. This is indicative for reconstruction errors resulting in broken tracks and events, which show up at higher frequencies and broaden the start frequency distribution.



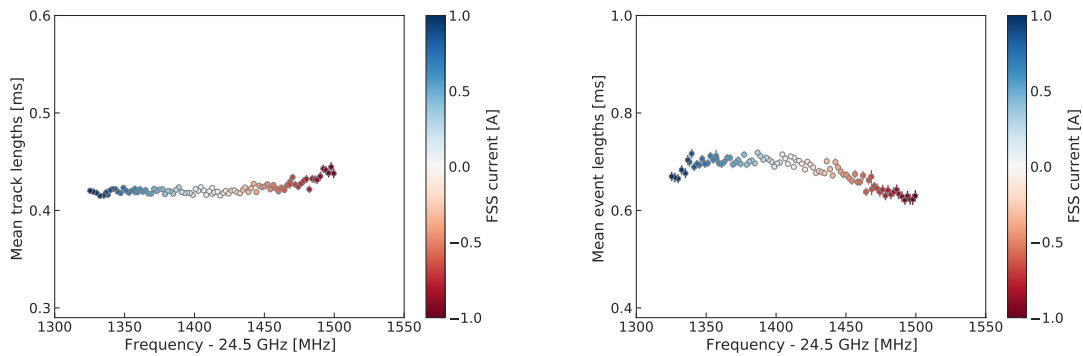
(a) FWHM of the un-scattered peak in trap 3

(b) Survival probability in trap 3

Figure 9.20: Lineshape parameters in the trap 3 FSS scan as a function of frequency. Especially at 1328 MHz (IF), the presence of waveguide mode results in tracks with very high slopes which impacts the event reconstruction and thus the lineshape fits.

Track and event lengths

The track and event lengths follow an exponential distribution. The track length is determined by the time between scatters and its mean is therefore proportional to the scattering cross section. The scattering cross section has no dependence on the magnetic field or the frequency and is expected to be constant for (quasi-)monoenergetic electrons during an FSS scan. The event length is defined by the difference between event end and event start and is theoretically equal to the summed track lengths of the tracks belonging to the event. All in all, the track and event length distributions should not be directly affected by a field scan. Yet, figure 9.21 shows that in the example of the trap 3 scan, the mean of first track lengths and the mean of the event lengths change with frequency in opposite directions. This is due to the power cuts in the reconstruction process (section 7.5). Harsher cuts are applied when the first track is short, which leads to an increase of the mean length with decreasing SNR. When the SNR is low, fewer tracks are detected and the number of tracks in an event, and therefore the average event length decreases.



(a) Reconstructed track lengths in trap 3

(b) Reconstructed event lengths in trap 3

Figure 9.21: Track and event lengths in the trap 3 FSS scan. Both quantities are expected to be constant in an FSS scan. The dependence on frequency is a result of the power cuts applied in the event reconstruction process.

9.2.3 Multi-trap scan results

The frequency dependence of all event properties in the Q300 trap is a composition of the dependencies in the four individual harmonic traps. Figure 9.22a and figure 9.22b show the count rates and average slopes of the Q300 FSS scan.

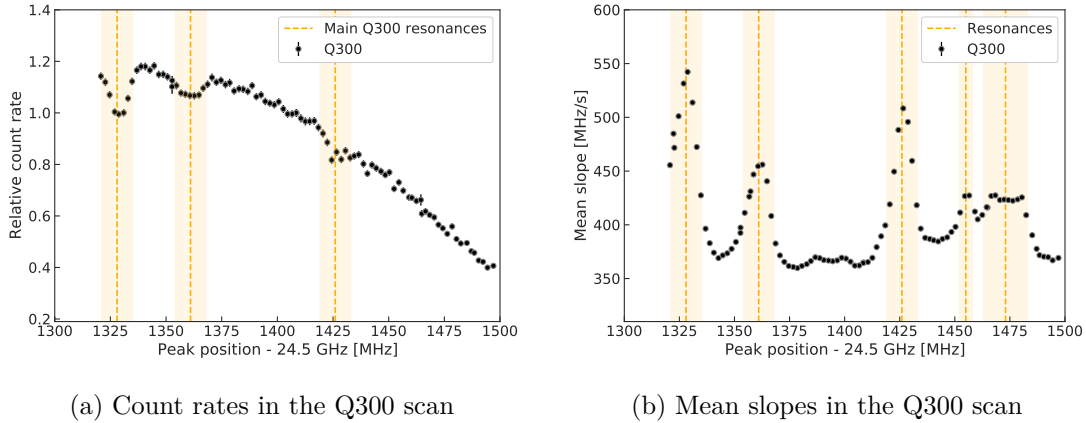


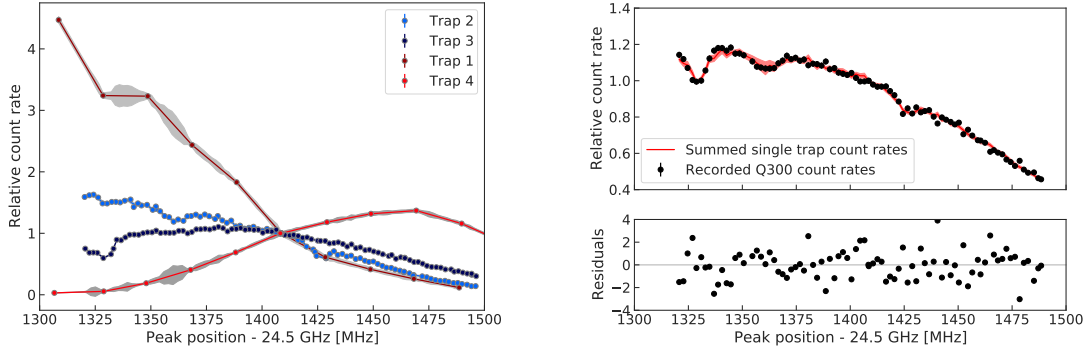
Figure 9.22: Count rates and track slopes vs. frequency in the Q300 trap FSS scan.

It can be seen that all peaks in the slope vs. frequency curves from all four single harmonic traps are present in the Q300 data, but the peak heights are weighted averages. Similarly, the count rate in the Q300 trap is the sum of the detected event rates in the individual traps. Therefore, all count rate drops due to slope resonances are present in the Q300 count rate curve. However, the relative contributions to the rate coming from each trap cannot be taken directly from the individual FSS scans, because the trap depths in the single-trap scans and in the Q300 trap configuration are not the same. In the single-trap FSS scans, the trap depths were set so that the 17.83 keV peak is located at the same frequency in each trap. The same is done in the Q300 configuration (see section 4.2.2), but since the fields from the sub-traps affect each other, the trap coil currents needed for the alignment are not identical. As a result the trap depths in single-trap and multi-trap operation are different.

The first order approximations for the count rate contributions from all sub-traps are again taken from the fitted rate-to-current relation in figure 4.18 and serve as initial guesses. The count rate vs. frequency curve in the Q300 data (figure 9.22a) is compared to the weighted sum of the rates in the individual scans. The weights are left floating in a least-squares minimization of the difference between the summed single-trap and the Q300 count rates. Because the scans in the traps 1 and 4 were only taken with large current steps, their count rates and count rate errors are interpolated to the frequency positions in the Q300 FSS scan (figure 9.23a). The count rate uncertainties in the interpolated frequency ranges are increased by the deviation from a linear interpolation in the fine scan of the trap (2 or 3) with the most similar mean slope at this frequency.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

The good agreement between summed single-trap rates and the Q300 rate (figure 9.23b) enables to make count rate predictions for the Q300 trap based on the count rates in the single traps.



(a) Interpolated relative count rates in all single harmonic traps. The rate in each trap is set to 1 at $I_{\text{FSS}} = 0$ A. The grey shaded bands indicate the rate uncertainties.

(b) Comparison of the count rates in Q300 to the weighted sum of the single-trap rates. The weights were optimized to minimize the residuals.

Figure 9.23: Comparison of event rates in single and multi traps. The relative contribution to the Q300 count rates at $I_{\text{FSS}} = 0$ A were fitted by minimizing the least squared residuals. The fit results for the traps' respective contributions to the Q300 count rates are: $p_{\text{opt}} = \{0.076 \pm 0.003, 0.341 \pm 0.013, 0.381 \pm 0.014, 0.203 \pm 0.02\}$.

The mean SNR from all traps can also be summed to the Q300 mean

$$\text{SNR}_{\text{Q300}} = \frac{\sum_i^{N_{\text{Q300}}} \text{SNR}_i}{N_{\text{Q300}}} = \sum_{j=1}^4 \left(r_j \cdot \frac{\sum_i^{N_j} \text{SNR}_i}{N_j} \right) \cdot \frac{1}{4}, \quad (9.2)$$

where i refers to the i^{th} event in a trap and j indicates the trap number. r_j is the relative number of events from each trap for any given frequency position. The relation of summed count rate to averaged SNR does not correspond to the relation observed in the single harmonic traps. This can be explained by a thought experiment: Assume 3 of the 4 traps have a constant count rate and a constant mean SNR as a function of frequency. The SNR in the 4th trap decreases at some frequency such that no electrons in this trap could be detected. The count rate in the Q300 trap would decrease at this frequency, while the mean SNR would increase. The strict monotony of count rate vs. SNR can therefore be broken in a multi-trap configuration. Similarly, there can be multiple single-trap SNR combinations that lead to the same total count rate in a multi trap. It is therefore not surprising, that the shape of the count rates vs. mean SNR curve in the Q300 trap configuration (figure 9.24) does not agree with the curves in figure 9.17.

9.2. STUDYING FREQUENCY-DEPENDENT SYSTEMATICS

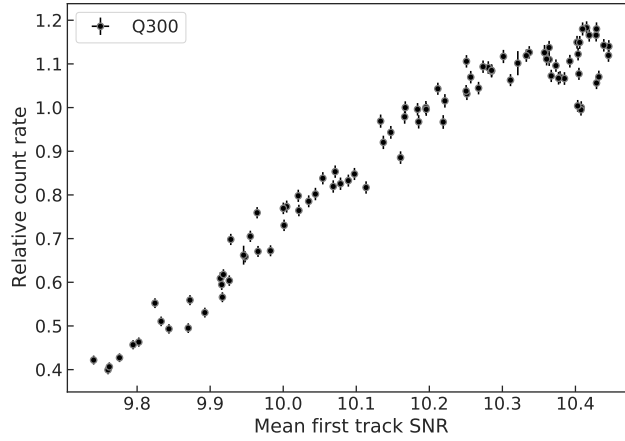
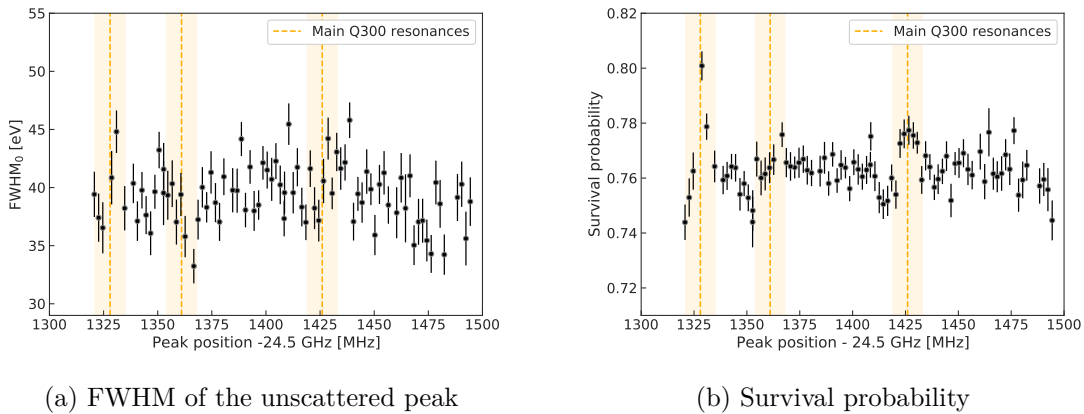


Figure 9.24: Count rate vs. SNR in the Q300 trap configuration. Because this trap is a multi-trap configuration, the relation between the two quantities is no longer deterministic.

In contrast to the summed count rate and averaged SNR, the composition of the lineshape in the Q300 trap is more complicated because the lineshape parameters are not additive. Instead, an effective lineshape can be fitted to the K-line peaks in the Q300 data. Figure 9.25 shows the frequency dependence of the effective lineshape parameters on frequency. The parameter variations with frequency appear to be purely statistical except for the mode resonance frequencies (where slopes are large).



(a) FWHM of the unscattered peak

(b) Survival probability

Figure 9.25: Lineshape parameters vs. frequency in the Q300 FSS scan. Except for the frequency positions where the track slopes are high, the lineshape parameters appear to be independent of frequency.

9.3 Constructing an efficiency curve for tritium

In figure 9.22a and figure 9.15 all events in each recorded peak were counted as a function of the peak's frequency position. Because the peak in traps powered by a 300 mA trap coil current is about 10 MHz wide (including scattering tail), the full event rate curves reflect the frequency-dependent detection efficiency smeared over 10 MHz. To obtain a finer picture of the efficiency, a start frequency cut is applied and only events from a 2 MHz⁵ wide range around the fitted peak position are counted (figure 9.26b). As expected, the regions that are affected by sharp changes of the mean slope exhibit an even more pronounced count rate deficit after the frequency cut.

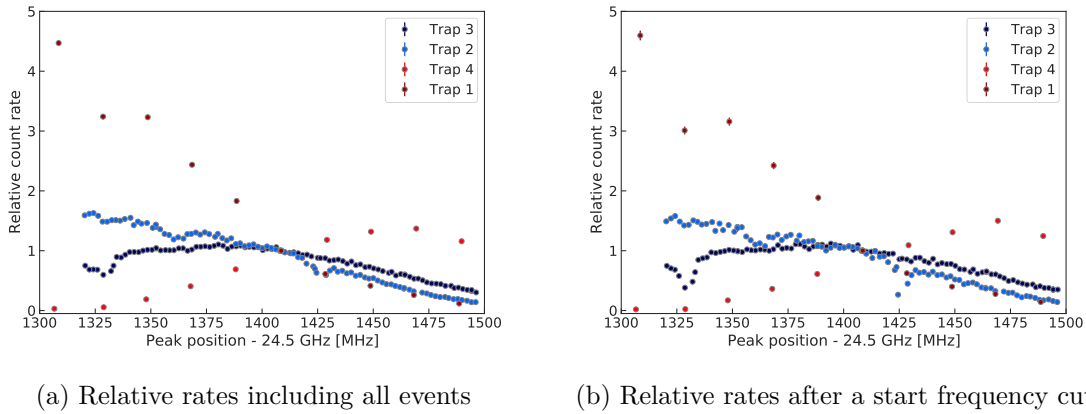


Figure 9.26: Relative count rates in all single harmonic traps before (a) and after (b) a start frequency cut of $|f - f_0| \leq 1$ MHz is applied to each recorded K-line peak. After the cut, the count rate deficits at the high-slope frequency positions are more pronounced.

Because the efficiency enters via multiplication into a distorted tritium spectrum model and the model is normalized, the absolute scale of the efficiency is irrelevant. Relative count rate differences can be directly translated to relative efficiency changes. Therefore, summing the cut count rates from figure 9.26 with their relative contributions from figure 9.23b results in the full frequency-dependent detection efficiency curve for the Phase II system in the Q300 trap. Effects like power coupling, reflections, analog and digital signal processing, and event reconstruction are all included. However, this efficiency curve is only valid for the quasi-monoenergetic K-line electrons with a kinetic energy of 17.83 keV. For a tritium efficiency curve, the absent power dependence on energy in the FSS scans has to be corrected for. The coupled power in the TE₁₁ mode as a function of frequency is shown in figure 9.27.

⁵A range narrower than 2 MHz would reduce the number of events too much and prevent a sound efficiency analysis.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

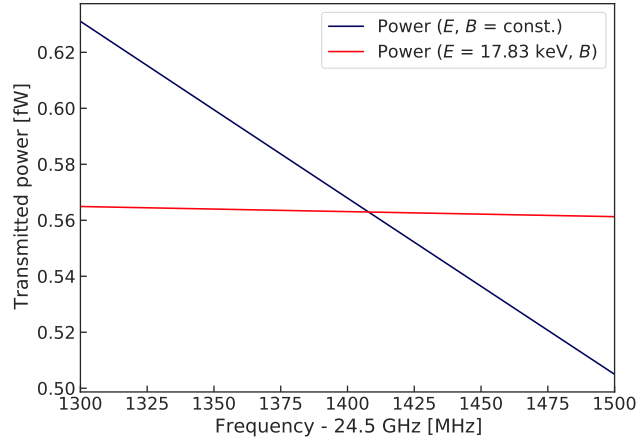


Figure 9.27: Comparison of the expected received power dependence on frequency in FSS scans compared to the expected dependence on kinetic energy. The ratio of the two curves is used to correct the event SNRs in FSS scans for the missing energy dependence.

The main event property that is affected by a change of the emitted power is of course the SNR. If the true SNR of all events (even of the undetected ones) in the FSS scans was known, it could simply be multiplied by the coupled power for any given energy relative to the power of 17.83 keV electrons. As a result, events would appear or disappear above the detection threshold, which would change the number of detected events and thus the detection efficiency. Therefore, the steps needed for the energy correction of the frequency-dependent efficiency curve are:

1. Knowing the SNR distributions in all traps at all frequencies in the FSS scans.
2. Multiplying the SNR of each event by the relative power from figure 9.27 to obtain energy-corrected SNR distributions.
3. Predicting how count rates change as a result of the energy-corrected SNRs.

The main challenge lies in the fact that the reconstructed track SNRs do not translate directly to an underlying signal power but are smeared by the reconstruction as was shown in section 8.3.4. In addition, it is difficult to predict how the count rates would change if all SNRs were multiplied by a factor because the shape of the SNR distribution below the detection threshold cannot be measured. In the following sections, an understanding of the relation between SNR and count rate changes will be developed, resulting in the construction of a curve describing the energy-dependent detection efficiency in Phase II.

9.3.1 SNR analysis using unfolding

While unfolding reconstructed data yields an underlying SNR distribution even in the SNR region below the detection threshold (section 8.3.4), there is no information present in the reconstructed data that would justify a prediction of its shape. Instead, the unfolding result resembles the shape of the simulated SNR distribution that was used to generate the response matrix in the first place. To test in which SNR ranges the unfolding result is reliable, two response matrices based on different SNR distributions were generated. The first response matrix based on a uniform SNR distribution was already shown in figure 8.22. The second response matrix is generated from a fake event data set in which the uniform signal powers were corrected for random start pitch angles and random start z -positions (figure 9.28). In figure 9.29 the reconstructed SNR distribution of events recorded in trap 3 at $I_{\text{FSS}} = 0$ A is unfolded.

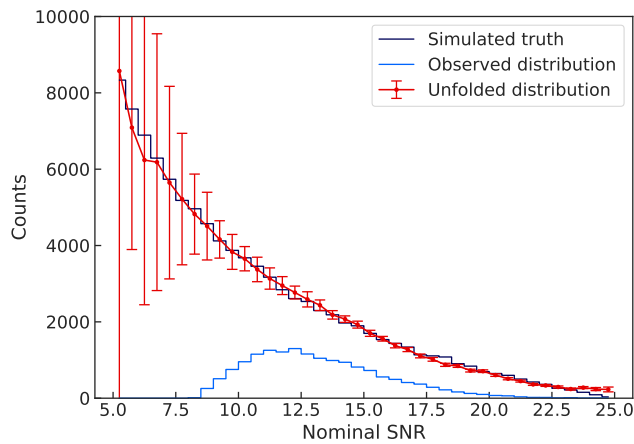


Figure 9.28: Unfolding result (red) for a reconstructed fake event distribution (bright blue). The true simulated event distribution (dark blue) is well retrieved.

As expected, the unfolding result is distorted by the original fake event distribution in the SNR regions where no or little information is available in the recorded data. Using the unfolding matrix based on the pitch-angle-corrected power distribution is therefore expected to give the most accurate result as it is closer to the expected true SNR distributions in Phase II data. Nonetheless, only the SNR region between 9 and 13 will be considered in the unfolding analysis, because in this region a sufficient number of events is present in the recorded data and both unfolding matrices produce a similar result.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

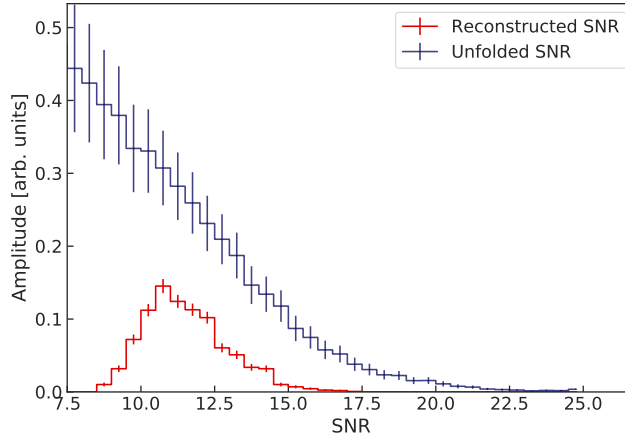
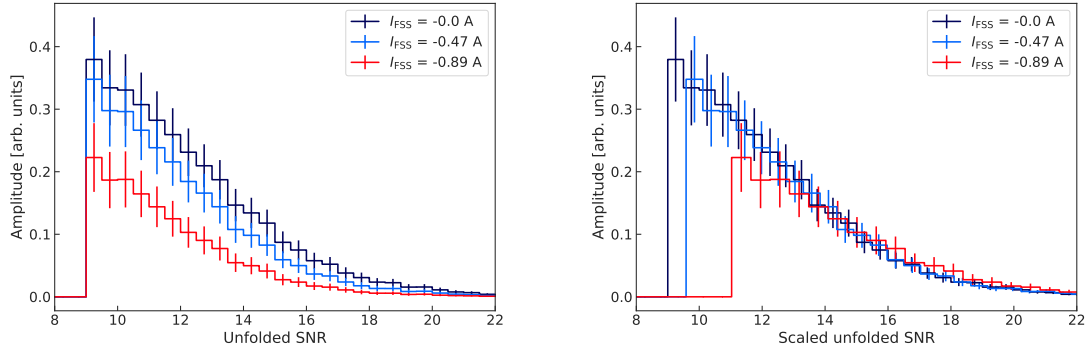


Figure 9.29: Unfolding of a first track average SNR distribution from the trap 3 FSS scan using the same unfolding matrix as in figure 9.28.

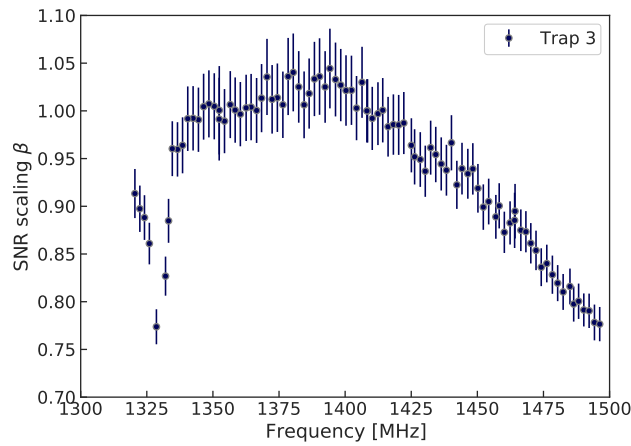
It can be assumed that the population of trapped electrons is the same with respect to the distribution of pitch angles and radial positions in all four traps at all FSS currents. This assumption is justified, because the starting conditions are random and determined by the boundary conditions of the trapping field. Without interference effects, the power coupling to the traveling TE_{11} mode would be independent of the longitudinal positions of the radiating electrons in the waveguide. But as stated previously, the power transmission is modified by interference and additional modes in the waveguide, which causes a frequency and trap position dependence. Like the power dependence on kinetic energy, the interference effects effectively scale the distribution of received power up and down the power axis. To measure this scaling effect in the FSS scans, a relative scaling factor β is extracted by which the power distributions are shifted from one FSS current to another and equivalently from one frequency position to another. For this purpose, the SNR distribution at each frequency (after the frequency cut) is unfolded as in figure 9.29 and the individual unfolding results are shifted relative to each other until the sum of the least squared differences in the reliable SNR region (9 to 13) is minimal. Figures 9.30a and 9.30b show the effect of scaling unfolded SNR distributions and in figure 9.30c the extracted scaling factors for the example of the trap 3 FSS scan can be seen.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM



(a) Examples for unfolded SNR distributions in the trap 3 FSS scan.

(b) Scaled unfolded distributions from (a) after optimizing the relative scaling factors β in a least square minimization.



(c) SNR scaling factors β extracted from shifting the unfolding results of the trap 3 FSS scan data relative to $I_{\text{FSS}} = 0$ A.

Figure 9.30: SNR analysis based on unfolding reconstructed first track SNR distributions.

9.3.2 SNR analysis using forward folding

The unfolding and scaling result in figure 9.30 is promising. Like the mean of the reconstructed SNR distribution in figure 9.16a, the fitted scaling factor exhibits the same qualitative frequency dependence as the count rates. However, the uncertainties are large and the bias of the unfolded distributions towards the fake event distribution limits the extent to which the result can be trusted. The opposite approach to extract a scaling factor is to use the (first track) response matrix from fake events and forward fold a simulated power distribution. This folded distribution should match the reconstructed first track SNRs. To this end, a power distribution even more precise than in section 8.2 is needed, which requires to take both the longitudinal and the radial magnetic field geometry of a harmonic trap into account. Figure 9.31 shows the field geometry used to correct the uniform start pitch angle distribution.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

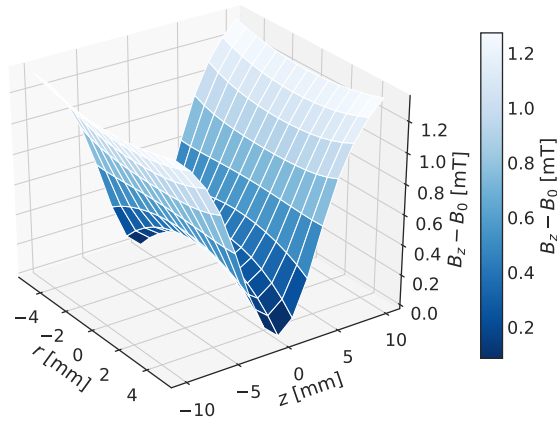
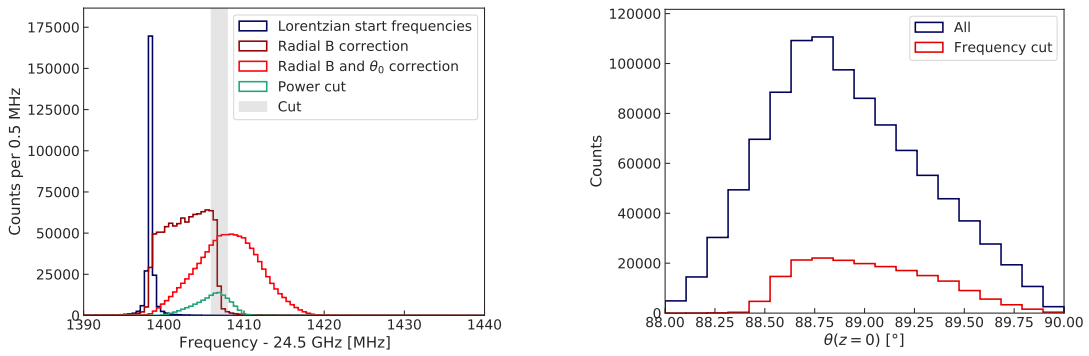


Figure 9.31: Single-trap field map used for the generation of a simulated power distribution.

In contrast to the power distribution in section 8.2, a distribution that is to be compared to FSS data has to be subjected to the same start frequency cut. For this purpose, the cyclotron frequency of each simulated event is calculated according to its radial position and pitch angle (figure 9.32a). All generated events with a start frequency outside a 2 MHz wide range centered on the frequency of the peak's maximum (after a power cut) are removed. The power distribution is then composed of the powers calculated from the starting positions in r and θ_0 of the remaining events.



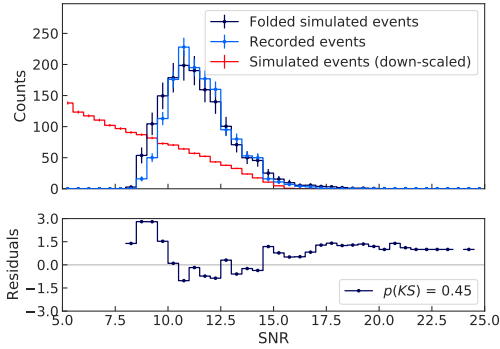
(a) Analytically calculated frequencies of simulated events. The grey shaded frequency range indicates the frequency cut.

(b) Simulated pitch angle distribution before (blue) and after (red) the frequency cut was applied.

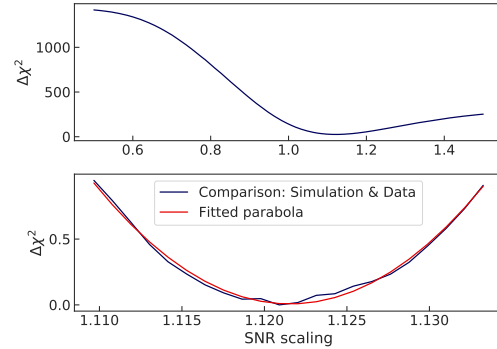
Figure 9.32: To generate a realistic power distribution, the frequency cut that is applied in real FSS data must also be applied in simulated data.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

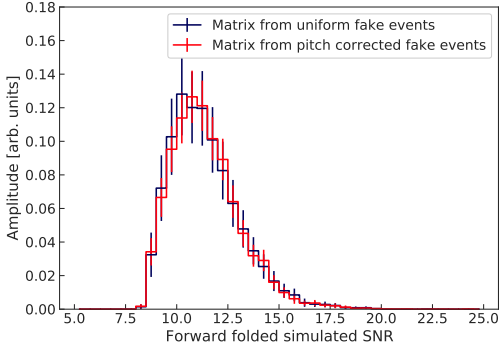
To compare the simulated power distribution with the reconstructed FSS data, it has to be scaled to a comparable SNR and multiplied with the response matrix. This process is repeated iteratively until the least squared difference to the reconstructed SNR distribution at any given FSS current is minimal. Figure 9.33a shows the comparison of the forward folded generated power distribution using the same response matrix as in figure 9.28 with the reconstructed SNR distribution from trap 3 at $I_{\text{FSS}} = 0$ A. While the agreement is imperfect, forward folding has the advantage that the result is not biased by the fake event distribution used to generate the response matrix (figure 9.33c).



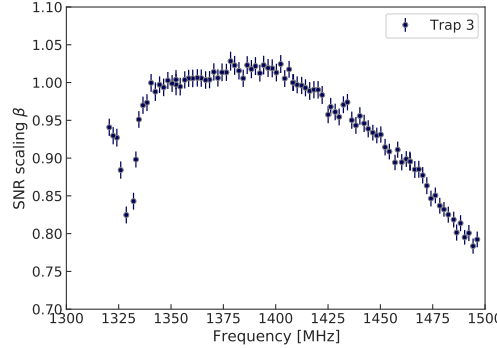
(a) Scaled, forward folded, simulated power distribution (dark blue) compared to real data (bright blue). The residuals are given relative to the combined uncertainty.



(b) Forward folding profile of the SNR scaling factor. The value that minimizes χ^2 is used to generate the folded simulated SNR distribution in (a).



(c) Comparison of forward folding results using the response matrices extracted from figure 8.23 and figure 9.28. The results are independent of the fake event set used to generate the response matrix.



(d) Extracted scaling factors from forward folding a simulated SNR distribution to reproduce the first track SNRs recorded in the trap 3 FSS scan.

Figure 9.33: SNR analysis based on forward folding a simulated SNR distribution. The forward folding result is independent of the fake events that were used to generate the response matrix. It is therefore considered more reliable than the unfolding result from 9.30.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

Similar to the unfolding analysis, the scaling required to minimize the difference to each recorded event set at each FSS current in all traps can now be compared. The scaling and folding of the generated power distribution has to reproduce both the shape and the amplitude of the reconstructed SNR distributions. Figure 9.33d shows the scaling factors extracted via forward folding relative to the scaling factor at $I_{\text{FSS}} = 0$ A.

Comparing the resulting scaling factors extracted via the two different approaches (figures 9.30c and 9.33d), it can be seen that although the forward folding approach suffers from different dominant uncertainties (input power distribution) than the unfolding approach (unfolding matrix), the resulting dependence of the relative SNR scaling on frequency is remarkably similar. Only the dip at the frequency position where the track slopes are highest (1328 MHz) is more pronounced in the scaling factors retrieved via unfolding. During the unfolding process the SNR distribution is divided by the SNR-dependent efficiency. The propagation of the efficiency uncertainties results in larger error bars on the SNR scaling curve than if generated data is forward folded. Because the forward folding approach gives a more precise result and does not suffer from a bias in the unfolding process, it is the preferred method and will be used for the energy correction of the FSS data. Figure 9.34 shows the SNR scaling result from all four traps relative to the scaling factor in trap 3 at 1408 MHz IF (K-line position at $I_{\text{FSS}} = 0$ A after HF-mixing).

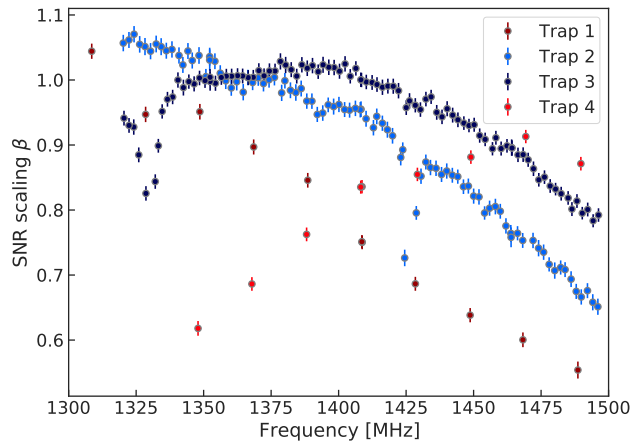


Figure 9.34: SNR scaling factors in all single-trap FSS scans relative to the scaling factor at $I_{\text{FSS}} = 0$ A in trap 3.

9.3.3 Count rate vs. SNR

The extracted SNR scaling factors from figure 9.34 can be combined with the count rates from all traps to obtain the relation of the two quantities (figure 9.35). The absolute count rates were again calculated by multiplying the relative detected event rates in all FSS scans (relative to the K-line frequency at $I_{\text{FSS}} = 0$ A) by the event rates in the trap depth scans (figure 4.18) interpolated to the FSS scan trap depths.

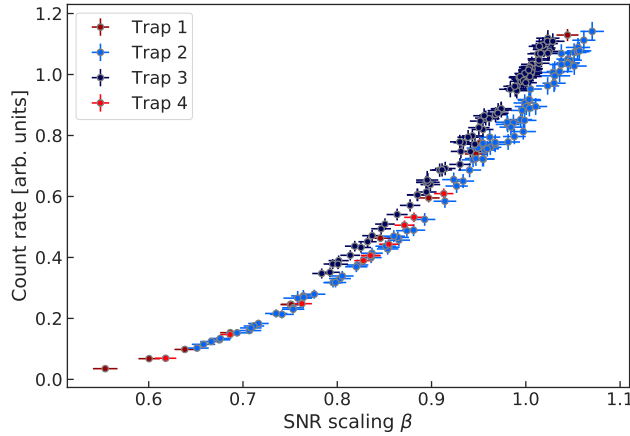


Figure 9.35: Count rate vs. SNR scaling factor in all single-trap FSS scans. The data from high-slope frequency positions are excluded.

The shape of the count rates as a function of the SNR scaling factors looks identical in all four traps. Only the absolute scale of the count rates is different. As the spread of scaling factors is significantly smaller than that of the mean SNR in figure 9.17a, the difference is more visible here. While in the further analysis all traps could be treated individually, it is safe to assume that the rate difference originates from the missing frequency dependence in the interpolation of the trap optimization data from section 4.2.2 and the differences in trap depths. As the energy correction in all traps is applied relative to the K-line, the absolute count rates in all traps can be scaled arbitrarily without affecting the resulting correction. The relation of count rate to scaling factor is fitted by a 4th order polynomial. In order to improve the precision on the fit parameters, the absolute count rates in trap 1, 2, and 4 are scaled to minimize the least squared error of a combined fit. This way the information from all 4 traps is used simultaneously (figure 9.36). Data from frequency positions with high mean slopes is excluded ((1328 ± 5) MHz in trap 3 and 4 and (1426 ± 5) MHz in trap 2) because the scaling factors extracted via unfolding and forward folding differ in these regions and the count rate vs. scaling relation also slightly deviates from the curves in figure 9.35.

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

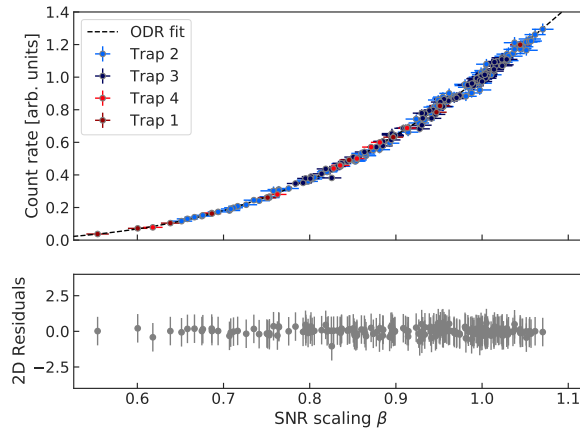


Figure 9.36: Counts vs. SNR scaling factors in all single-trap FSS scans. Here, the count rates were scaled to minimize the difference between the single-trap curves. The resulting count rates at $I_{\text{FSS}} = 0$ A relative to trap 3 are 0.26 (trap 1), 0.82 (trap 2), and 0.46 (trap 4).

The fit in figure 9.36 describes the relation so accurately that count rates can be taken as the better measure for the scaling factor extraction than the scaled SNR distributions. This way, even with count rate and polynomial fit parameter uncertainties propagated, the resulting uncertainties are smaller. This provides all the tools required to correct the count rates for the energy dependence: The count rates (after frequency cuts) at all FSS currents in all traps directly give rise to the SNR scaling parameter with the exception of the high-slope frequency regions. This parameter is multiplied by the power dependence on energy relative to the K-line power, which yields the energy-corrected SNR scaling parameter. Via the fitted count-rate-to-scaling-parameter relation, the corrected scaling parameter at the same time returns a corrected count rate. The energy-corrected relative count rates are shown in figure 9.37.

9.3.4 Slope error propagation

As in section 9.2.3, the energy-corrected count rates are summed to the expected Q300 relative detection efficiency (figure 9.38b). The corrected count rates from traps 1 and 4 are again interpolated to the frequency positions of the fine FSS scans (figure 9.38a). In this process, the interpolation uncertainty is calculated from the difference from a linear interpolation in the fine trap scans 2 and 3.

In section 8.3.5 it has been shown that the detection efficiency decreases with slope as events are reconstructed with lower average SNR than they truly possess. However, in the very high slope regions the unfolding and forward folding approaches gave slightly different results. To make sure that the efficiency curve is correct even in those regions,

9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM

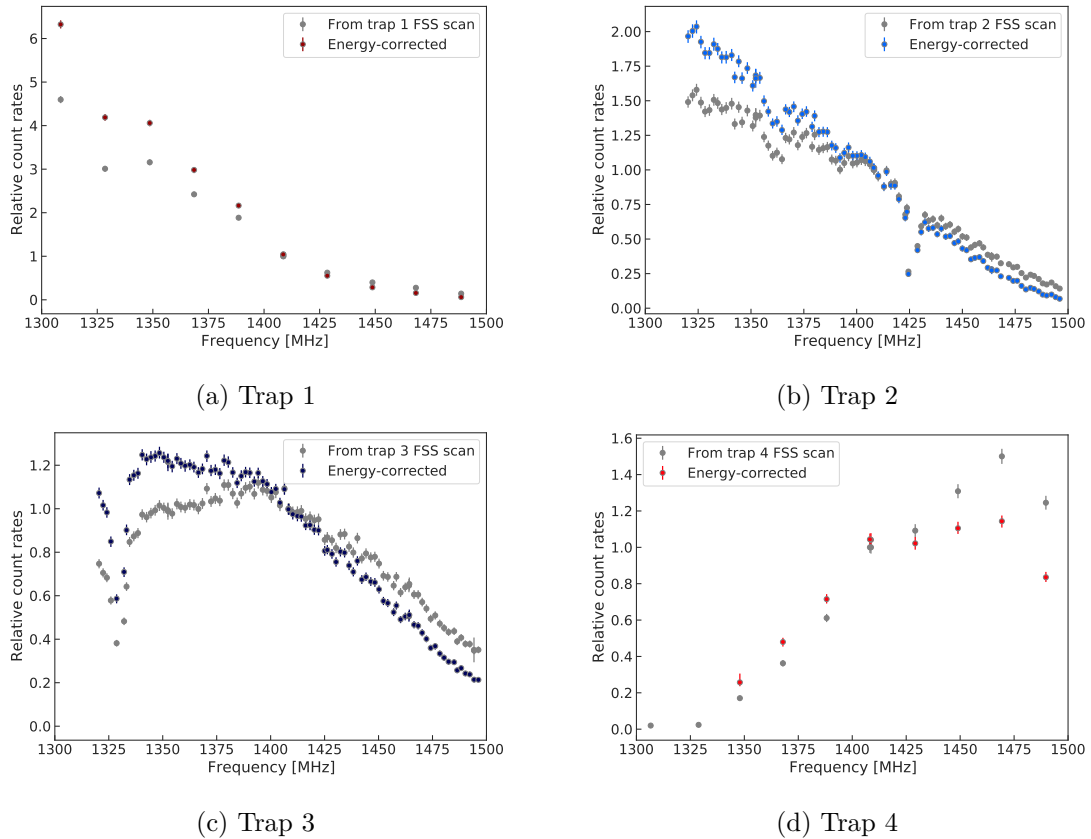
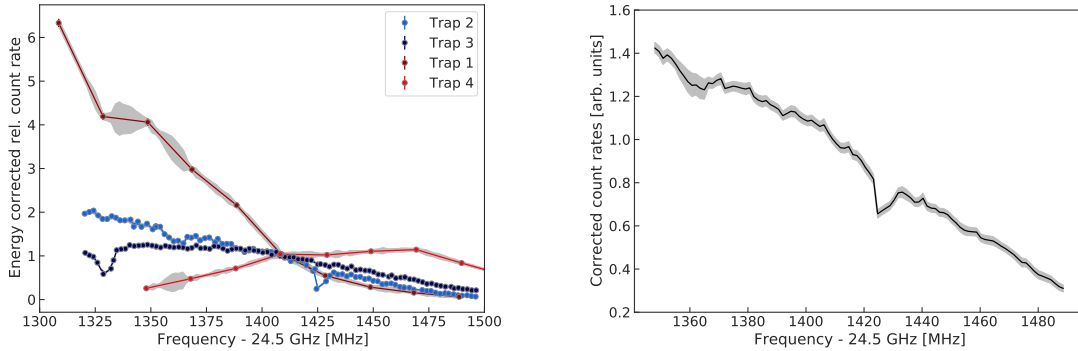


Figure 9.37: Relative count rates in FSS scans (grey) and energy-corrected count rates (color) from multiplying the SNR scaling factors with the relative transmitted power and reading the corrected count rates from figure 9.36.

the process of generating the curve from summing the interpolated and corrected count rates is repeated, this time with the high slope regions excluded (figure 9.39b). This curve is then multiplied with the ratio of the summed FSS-count-rate curves (prior to energy correction) without slope cuts and with slope cuts (figure 9.39a). This way, the slope effect is treated as relative deviation of the otherwise linearly interpolated count rates originating from (non-slope induced) SNR variations.

The predicted efficiency curve drops by only 16 % when it is corrected by the relative slope correction as opposed to 18 % when the high slope regions are not excluded and the efficiency is retrieved solely via the SNR. This difference is not statistically significant. The slope-corrected curve naturally has larger uncertainties in the region where the correction was applied. While there is no obvious reason to prefer one curve over the other, the slope-corrected curve will be used for the analysis of the tritium spectrum in chapter 10, as its error bars are larger in the high-slope regions, which makes it the conservative choice.

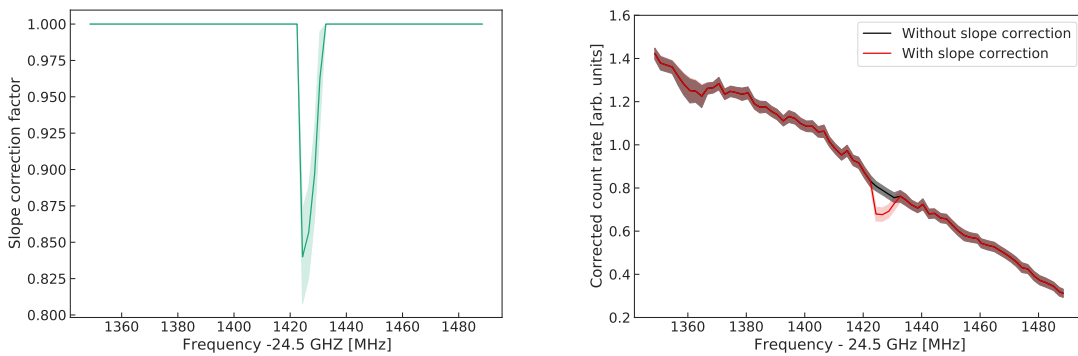
9.3. CONSTRUCTING AN EFFICIENCY CURVE FOR TRITIUM



(a) Interpolated energy-corrected single-trap count rates

(b) Summed energy-corrected count rates

Figure 9.38: Construction of an energy-corrected detection efficiency curve for the Q300 trap. The energy-corrected single-trap rates (a) are interpolated to the frequency positions in the Q300 FSS scan. Then they are summed by using the weights from figure 9.23. The uncertainties (grey) include the statistical and the systematic uncertainties from the polynomial fit in figure 9.36, the single-trap rate summation, and the interpolation between the 20 MHz FSS scan steps in trap 1 and 4.



(a) Ratio of the summed count rates vs. frequency curve from all single traps and the same curve with data points from high-slope frequency regions excluded.

(b) Comparison of the energy-corrected Q300 count rate curve excluding high-slope frequency regions (grey) to the same curve multiplied with the slope correction factor (red).

Figure 9.39: Q300 efficiency curve construction with high-slope regions excluded and then reintroduced as a correction factor. The slope- and energy-corrected count rate curve in (b) is used as detection efficiency curve in the analysis of the Phase II tritium spectrum.

10

Detection efficiency as source of systematic uncertainty

With the detection efficiency curve from figure 9.39b, the distorted model for the expected Phase II tritium spectrum shape can be created by multiplication (figure 10.1). In this process, the magnetic field strength used to translate the frequency-dependent efficiency curve to an energy-dependent one is extracted from the fit of the lineshape model to the start frequencies in the calibration data set in figure 4.19.

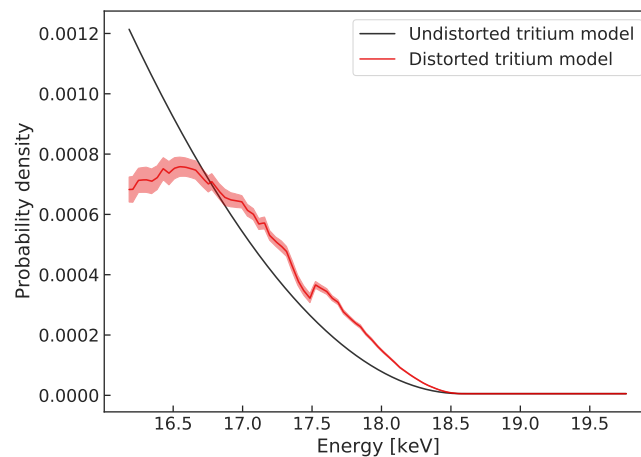
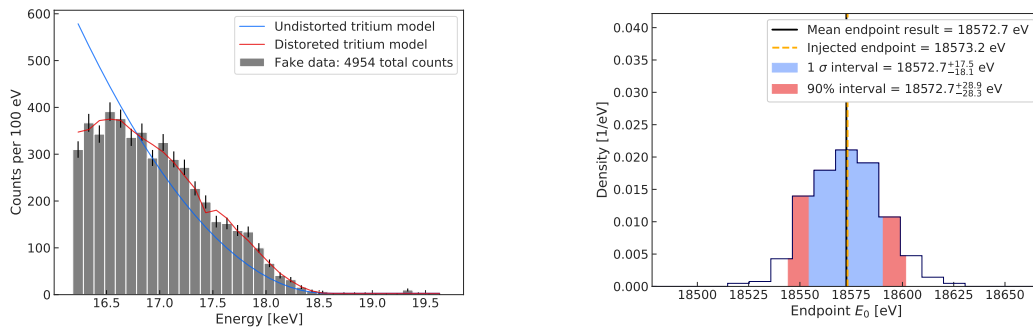


Figure 10.1: Comparison of the undistorted (black) and the distorted tritium model (red). The distortion results from multiplication with the efficiency curve from figure 9.39b. The red-shaded area reflects the 1σ uncertainty on the expected event rate originating from the efficiency uncertainty.

The efficiency curve comes with an uncertainty that propagates to an uncertainty on the expected tritium event rates as a function of frequency (or energy). In this chapter the impact of the detection efficiency and its uncertainty on the analysis of a tritium spectrum is investigated. To this end, the efficiency uncertainty is first propagated to an endpoint uncertainty in fake data experiments (section 10.1). In section 10.2 the real tritium spectrum recorded with the Phase II apparatus is presented and analyzed, including all systematic uncertainties that are present in the experiment. Finally, an extrapolation to a future CRES neutrino mass experiment is made in section 10.3 and a precision requirement for Phase IV is obtained.

10.1 Efficiency uncertainty in the Phase II tritium model

The difference between the undistorted spectrum shape and the distorted model (figure 10.1) is larger than the toy model in figure 9.5, especially towards lower energies. But even a distortion this large can be accounted for in the analysis model without leading to a systematic bias. In figure 10.2 data generated from the expected Phase II tritium model is fitted with the exact same model. The endpoint results in 1000 pseudo-experiments average to the true model endpoint and yield a 1σ interval of $18572.7^{+17.5}_{-18.1}$ eV. The coverage of the 90% interval is $(89.6 \pm 1.0)\%$. However, these results do not yet contain the endpoint uncertainty introduced by the uncertainty on the efficiency curve.



(a) Fake data generated from the distorted model. Both the tritium model and the efficiency curve are integrated over each bin width.

(b) Distribution of endpoint fit results. The mean of all results converges to the true (injected) model endpoint. The 1σ and 90% interval are defined as in section 9.1.

Figure 10.2: Fitting fake data with the distorted model. As long as the distortion is known, the fit results are unbiased.

10.1.1 Efficiency uncertainty

In the construction of the efficiency curve the following uncertainties were included:

- Statistical uncertainty of the count rates in the single trap FSS scans: Because the frequency cut of ± 1 MHz reduced the number of events significantly (by about 70%), the statistical uncertainty is larger than the 1.5% error targeted by the total number of accumulated events. It amounts to approximately 3% at $I_{\text{FSS}} = 0$ A in each trap scan. The relative uncertainty grows with lower count rates and it is therefore especially large in frequency regions with low efficiency (up to 10%). Due to time constraints and difficulties in maintaining a stable cell temperature, taking more FSS data was unfortunately not possible.

10.1. EFFICIENCY UNCERTAINTY IN THE PHASE II TRITIUM MODEL

- Uncertainty from summing single trap rates to the Q300 efficiency curve: The spread of residuals in summing four count rate curves and comparing the result to the rates in the Q300 FSS scan (figure 9.23b) is larger than expected from pure statistical fluctuations. The square root of the additional width of the residual distribution is propagated as a systematic error and added in quadrature to the uncertainty of the Q300 efficiency curve.
- Uncertainties of the fit parameters in the count rate vs. SNR scaling fit: When translating an SNR scaling factor to a relative count rate and vice versa, the uncertainty on the 4th order polynomial parameters is propagated to an uncertainty of the translated quantity by sampling the polynomial parameters from a multivariate normal distribution. As the fit describes the count rates to SNR scaling relation so well (figure 9.36) the error induced in addition to the statistical error propagated via the polynomial curve is small.
- Uncertainty from interpolation: As the FSS scans in trap 1 and 4 were taken with a larger step size, the count rates had to be interpolated to the frequency positions of the fine-step scans in trap 2, 3, and Q300. The rates were interpolated linearly, taking the deviation from a straight line between the same frequency positions in one of the other trap scans as interpolation uncertainty. In doing so, the trap with the most similar track slopes in this region was selected for comparison. In most frequency regions the interpolation uncertainty contributes < 10 % to the full uncertainty on the efficiency curve. Only at (1360 ± 10) MHz the contribution rises to about 50 %. This frequency region is beyond the expected tritium endpoint and therefore not relevant for the tritium analysis.
- Slope correction: While it may not even be necessary to do a slope correction, as most of the slope variation is incorporated via SNR variation already, a slope cut is applied in the construction process of the efficiency curve. The efficiency deficit in the high-slope frequency region at (1426 ± 7) MHz is reintroduced by multiplying the ratio of summed FSS rates without and with slope cut. This enhances the uncertainty in this region by up to 30 % since the correction factor is obtained by dividing two interpolated count rate curves with their respective uncertainties.

The resulting relative uncertainty of the interpolated efficiency curve can be seen in figure 10.3. The efficiency curve is lowest at the high frequency (low energy) end. The relative error at this end rises to 7 % from 2 % on the low-frequency side.

10.1. EFFICIENCY UNCERTAINTY IN THE PHASE II TRITIUM MODEL

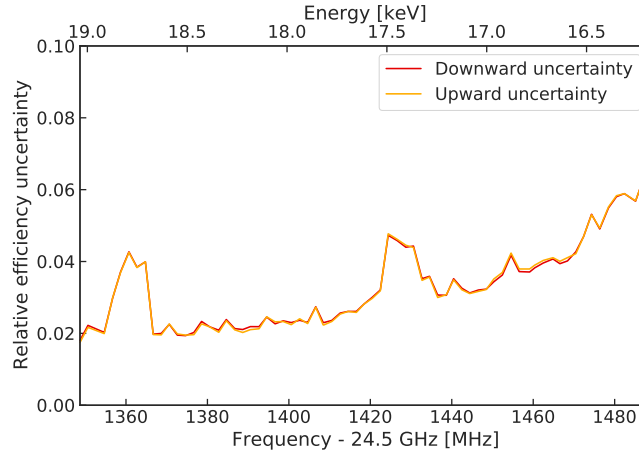


Figure 10.3: Relative 1σ efficiency uncertainty as a function of frequency or energy. The dominant source of uncertainty is the statistical uncertainty in the FSS scans, followed by the uncertainty from the interpolation of the trap 1 and 4 scans and the slope correction uncertainty. Since these three are all symmetric, the overall efficiency uncertainty is also nearly symmetric.

10.1.2 Uncertainty propagation to the endpoint fit result

In this section, the efficiency uncertainty is propagated as systematic error in the spectrum fits. A similar analysis strategy as in [27] is adopted. The systematic impact could in principle be investigated directly on real tritium data. However, the actual tritium data recorded in Phase II is the result of several systematic error sources, and the first goal here is to study the efficiency impact in isolation. To that end, a fake data set is created based on a tritium model that only includes a Gaussian energy resolution and the efficiency curve from figure 9.39b, but leaves out other systematic effects like energy loss due to scattering. The energy resolution is assumed to be independent of energy and frequency and its width is thus constant over the full spectrum. The generated data is fitted with the distorted (and smeared) tritium model. The free parameters in this fit are the number of tritium (signal) events, the number of background events, and the spectrum endpoint. The neutrino mass is fixed to 0 eV. Based on this initial-fit result (figure 10.4), the statistical and systematic errors are propagated in pseudo-experiments via Monte Carlo sampling.

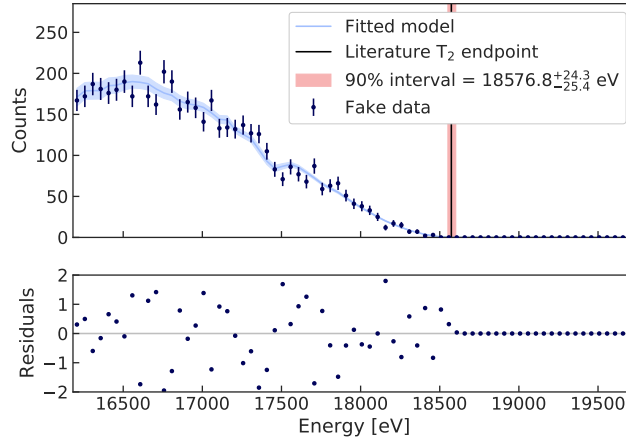


Figure 10.4: Fake data set with fitted model and 90% endpoint confidence interval (statistical uncertainty only). To propagate the statistical and systematic uncertainty, the number of events and the efficiency in each bin are varied within their respective uncertainties.

Because the efficiency curve used to generate and fit the data in figure 10.4 was fixed, the interval only reflects the statistical uncertainty. The statistical and systematic efficiency uncertainty as well as the combined uncertainty are propagated in three steps:

1. **Statistical uncertainty:** The purely statistical uncertainty is extracted via sampling similar to figure 10.2. Fake data is repeatedly generated from the initial-fit model in figure 10.4 and re-fitted. The spread of endpoint results yields the uncertainty interval shown in figure 10.5a (and figure 10.4).
2. **Systematic uncertainty:** All performed fits are binned fits. The bin size is set to 50 eV, which corresponds to almost exactly 2.5 MHz and is similar to the frequency step size in the FSS scans. For each bin, the efficiency and associated uncertainty are calculated by integrating the interpolated efficiency curve over the bin widths. To study the impact of the uncertainty in isolation, an Asimov [139] data set with no statistical fluctuations is created from the fitted model in figure 10.4. This fixed data set is repeatedly fitted with a recalculated tritium model. Each model is the product of the tritium spectral shape and an efficiency curve that was randomly sampled in each bin. The efficiency uncertainties are approximately symmetric and the efficiency samples are drawn from a Gaussian distribution with a mean equal to the assumed best efficiency curve and a standard deviation corresponding to the efficiency uncertainty in a bin. The spread of the endpoint fit results obtained in this way can be seen in figure 10.5b. Its width amounts to approximately 4 eV and reflects the systematic endpoint uncertainty originating from the uncertainty on the efficiency curve.

3. **Combined uncertainty:** The statistical and systematic uncertainty are usually added in quadrature. Instead, here the tritium counts and the efficiency curve are sampled simultaneously in pseudo-experiments. In each iteration, new fake data is generated from the initial-fit model and fitted with a model based on randomized bin-efficiencies. The resulting combined uncertainty is shown in figure 10.5c.

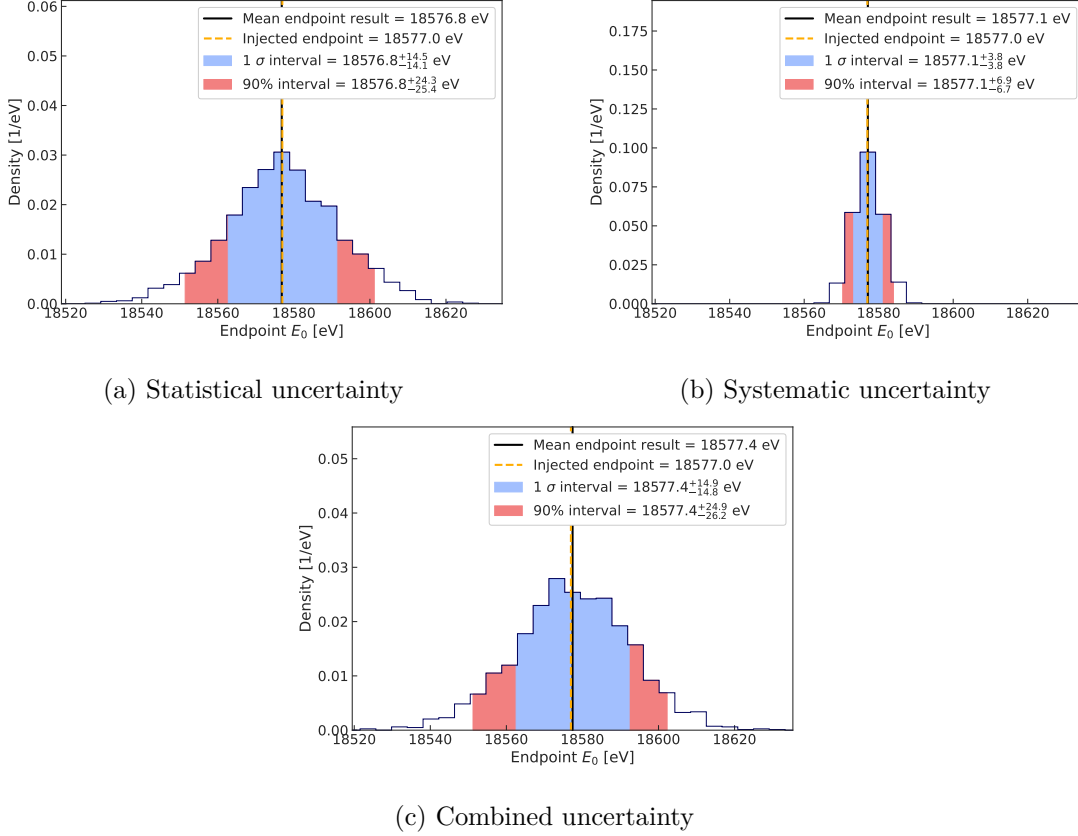


Figure 10.5: Endpoint fit results with propagated statistical (a), systematic (efficiency) (b) and combined uncertainties (c). The injected endpoint corresponds to the initial-fit value. The efficiency uncertainty does not contribute significantly to the total endpoint uncertainty.

In order to test whether the efficiency uncertainty introduces a bias in the analysis, the process of generating fake data and propagating the combined uncertainty via Monte Carlo sampling is repeated 200 times (figure 10.6). The true endpoint is sampled from a normal distribution with a mean value of 18573 eV and a width of 50 eV. The average offset of the initial-fit (injected) endpoint results is (-0.9 ± 1.1) eV and is thus unbiased. The analysis of the real Phase II tritium data in section 10.2 is performed in the same way as presented in this section, except that the uncertainties of all systematic model parameters are propagated. For this purpose, the tritium model used for the analysis of the real data includes not only the efficiency distortion, but also the effects of molecular final states and energy loss due to scattering.

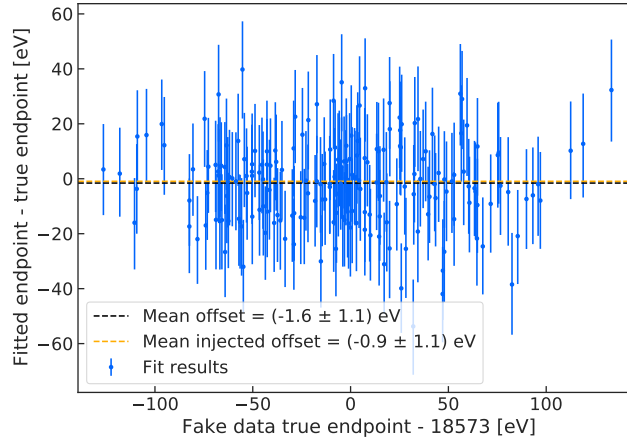


Figure 10.6: Endpoint fit errors vs. true model endpoints in 200 fake data experiments with sampled efficiency curves. The errorbars represent the combined uncertainty intervals (1σ). The coverage of the 90 % intervals is $(93.3 \pm 1.8) \%$.

10.1.3 Effect of a tilted efficiency curve

In the Monte Carlo studies in section 10.1.2 the uncertainty of the efficiency in each bin was regarded as independent of the neighboring bins. As most of the efficiency uncertainty originates from statistical fluctuations in the FSS scans, this approach is not unjustified. However, if the count rate vs. SNR relation (figure 9.35) or the size of the energy correction were under- or overestimated, the result could be an overall tilt of the efficiency curve (figure 10.7). To study the effect of such a tilt, fake data is generated from a model in which the efficiency curve is multiplied with a constant slope (in units of keV^{-1}). The fake data is analyzed by the same model as in section 10.1.2. This process is repeated with random relative efficiency tilts between -0.2 keV^{-1} and $+0.2 \text{ keV}^{-1}$. In figure 10.8a it can be seen that a tilted efficiency causes a systematic endpoint shift of $(3.4 \pm 0.2) \text{ eV}$ per 1 % tilt (0.01 keV^{-1}). A tilt in the efficiency not only affects the endpoint but also the shape of the entire tritium spectrum. When the tilt is added as a free parameter in the model that is used to analyze the data, both the endpoint and the tilt can be retrieved in a fit. Figure 10.8b shows the tilt fit result as a function of the true fake data tilt. The correlation between fitted and true tilt is 0.95 ± 0.05 which is compatible with 1. It can therefore be concluded that an unexpected tilt of the efficiency could be detected by adding it as a free parameter in the tritium fit. Note that the uncertainty of the tilt fit in figure 10.8b is larger for large negative tilts than for large positive tilts. The reason for this effect is that in the convention adopted here, negative tilts are equivalent to a lower efficiency far from the endpoint. As the tilt manifests itself most notably by the position of the roll-off of the spectrum shape far from the endpoint, fewer counts in this region lead to a larger statistical uncertainty on the fit result.

10.1. EFFICIENCY UNCERTAINTY IN THE PHASE II TRITIUM MODEL

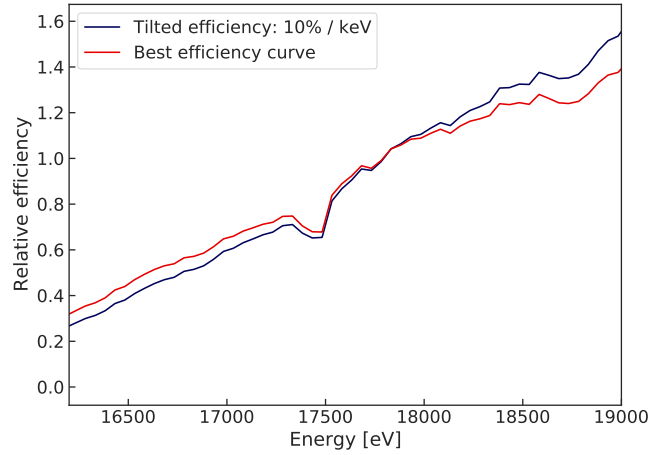
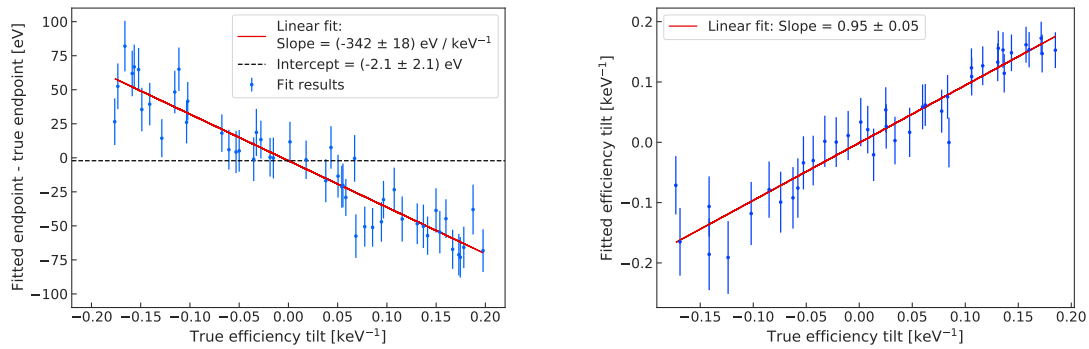


Figure 10.7: Effect of tilting the efficiency curve by 10% per 1 keV.



(a) Effect of a systematic tilt in the efficiency on the endpoint offset. If the tilt is not included in the analysis, the endpoint result is biased. The tested tilts are drawn from a uniform distribution between -0.2 keV^{-1} and $+0.2 \text{ keV}^{-1}$.

(b) Offset of the fitted tilt as a function of the true tilt in generated data. By treating the efficiency tilt as a free parameter in the analysis, the correct tilt can be extracted (on average).

Figure 10.8: An unknown tilt in the efficiency curve would lead to a systematic endpoint shift (a) that scales linearly with the size of the tilt. Such a tilt could be detected by adding it as an unknown free parameter in the fit (b).

10.2 Analysis of the Phase II tritium spectrum

After completing all krypton based efficiency, lineshape, and field calibration studies the cell of the Phase II setup was filled with molecular tritium gas (T_2 and HT). Tritium events were recorded over a three-month period from mid-December 2019 to mid-March 2020, with several weeks of downtime due to laboratory maintenance work in February. The total net run duration accumulated to ~ 82 days. Figure 10.9 shows the raw start frequency histograms of the recorded tritium data from all three ROACH2 channels (channels were arranged as in figure 5.11).

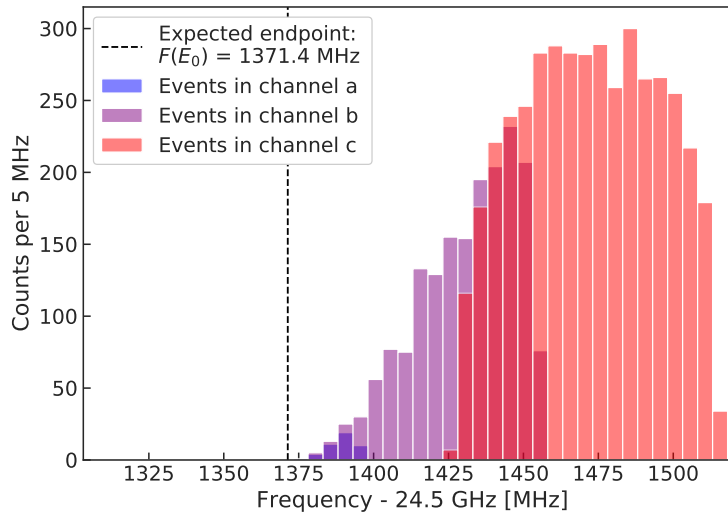


Figure 10.9: Raw Phase II tritium data with 6017 total counts recorded by three overlapping ROACH2 channels (a , b and c). The absence of background events in the frequency region below the endpoint allows to set an upper limit on the background event rate of $\leq 3 \cdot 10^{-10} \text{ eV}^{-1}\text{s}^{-1}$ (90% C.L.).

10.2.1 Data preparations

As expected, the recorded spectrum is strongly distorted. The roll-off at frequencies > 1500 MHz (IF) originates from the sharp efficiency decrease at the upper bound of channel c (section 5.2.4) in combination with the increase of image noise (figure 5.4). Before the spectrum is analyzed, events are removed from the overlapping frequency ranges between channels to avoid double counting. The frequencies that define the two channel transitions correspond to the frequencies at which the efficiency curves of adjacent channels intersect (figure 5.11). Hence, events are removed from channel a if their start frequency is greater than the frequency at which the efficiency drop between channel a and b is minimal $f_{\text{trans},ab} = 1386.2$ MHz (IF), while events recorded by channel b are cut if their start frequency is below $f_{\text{trans},ab}$ or above $f_{\text{trans},bc} = 1445.6$ MHz (IF). In channel c events starting below $f_{\text{trans},bc}$ are removed. Figure 10.10 shows the tritium data

10.2. ANALYSIS OF THE PHASE II TRITIUM SPECTRUM

from figure 10.9 after removing events beyond the channel transition frequencies. The grey shaded area indicates the frequency range covered by the efficiency curve derived in section 9.3.

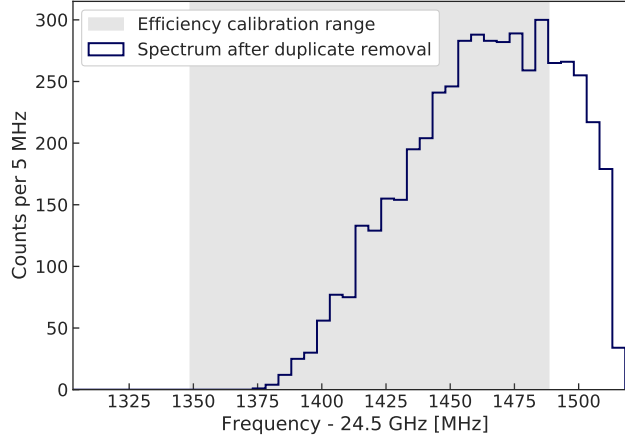


Figure 10.10: 4937 events remain after removing duplicate events in overlapping frequency ranges. The grey shaded area indicates the FSS calibration range: 25.849 GHz - 25.988 GHz. It contains 3742 tritium events. Since the background shape is assumed to be flat, no efficiency curve is required for the frequency region below the endpoint frequency. The lower frequency bound is therefore given by the lowest frequency recorded by channel *a* (plus a safety margin of 10 MHz): 25.813 GHz.

10.2.2 Tritium model tests

Prior to the analysis of the tritium data, the model used for the analysis is validated by testing it with fake data. To ensure that the validation studies are informative, the fake data generation model includes all correction factors to the tritium spectrum shape (section 2.6.2) and the full molecular final state spectrum of HeT^+ (figure 10.11). The fact that the final state spectrum depends on whether the parent molecule is HT, T_2 or DT is ignored here, since the differences between the binding energies of the daughter molecules are small (figure 2.8) compared to the energy resolution and sensitivity in Phase II. For the same reason the uncertainty on the final states is also not considered.

The analysis model, on the other hand, neglects all corrections of the electron phase space but the Fermi function and uses a sparse representation of the final state spectrum. This way a 5 times faster model evaluation is achieved. While in the following sections both the data generation and the data analysis model will be convolved with the scattering lineshape and multiplied with the efficiency curve from figure 9.39b, here all distortions originating from experimental systematic effects are omitted. In order to validate that the usage of the approximate model does not introduce any significant biases, fake data sets are generated from the detailed model and analyzed with the approximate model.

10.2. ANALYSIS OF THE PHASE II TRITIUM SPECTRUM

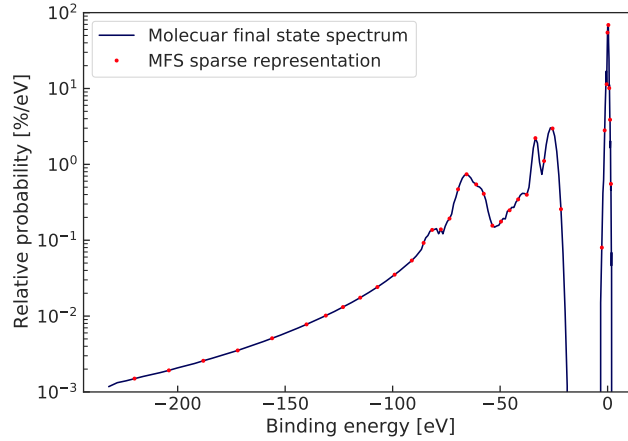


Figure 10.11: Molecular final state spectrum of T_2 decay [90]. In the analysis model only a sparse representation (red dots) is included.

Figure 10.12 shows the endpoint fit results of 200 such fake experiments with random true endpoints of (18573 ± 50) eV. Each fake data set contained on average 5000 events in the Phase II ROI. The fit results are unbiased and (88.8 ± 2.2) % of all fake experiments contained the true endpoint within their 90 % confidence interval.

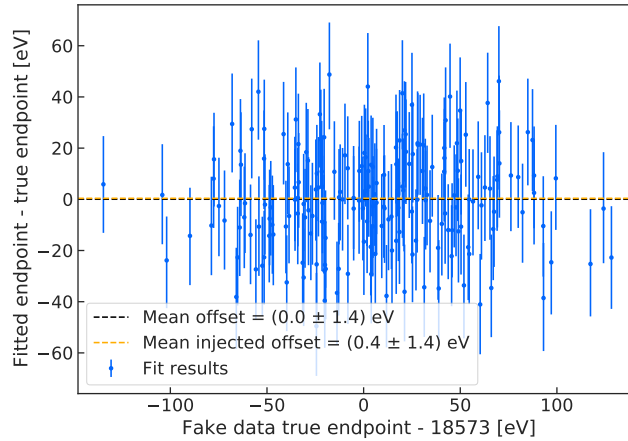


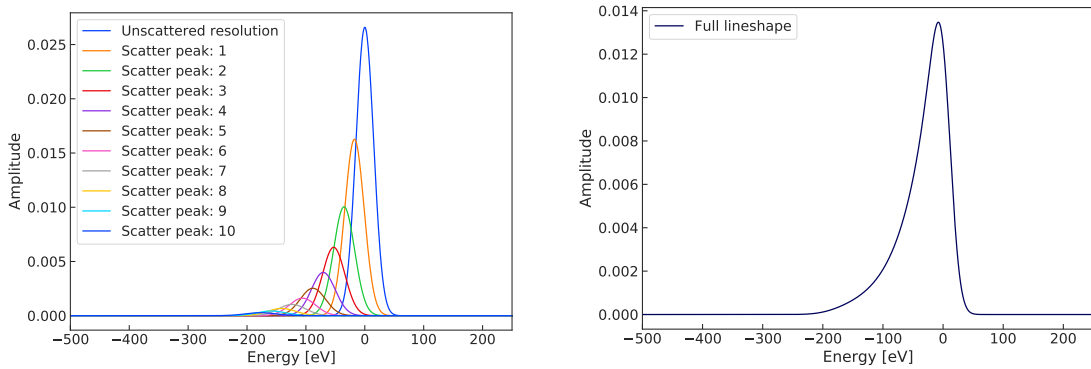
Figure 10.12: To validate the approximate tritium model, 200 fake data sets generated from a detailed tritium model (as derived in section 2.6) are analyzed. The true spectrum endpoint of the fake data was sampled from a normal distribution with $\mu = 18573$ eV and $\sigma = 50$ eV. The average endpoint fit result is unbiased. The coverages for the 1σ and 90 % interval are (66.9 ± 3.3) % and (88.8 ± 2.2) % respectively.

10.2.3 Including the lineshape

In section 10.1.2 the detection efficiency curve and its uncertainty have been studied. Of course, efficiency is not the only major systematic effect that affects the endpoint fit result. The asymmetric scattering lineshape (section 4.1.4) also leads to a systematic shift of the endpoint and must be accounted for in the analysis of the tritium data. Therefore, a model describing the lineshape has to be convolved with the tritium models. In the spirit of using the most detailed models for data generation, the detailed tritium model is convolved with the detailed lineshape model from equation (4.21). This makes it possible to specify the gas composition encountered by the emitted electrons and results in the most accurate description of the scattering tail. For the convolution of the tritium analysis model, a simplified version of the lineshape was developed. In this model, each scattering peak is approximated by a Gaussian shape, which is well justified in deep magnetic traps with a low energy resolution. The mean and standard deviation of each scattering peak were found to follow a linear and a logarithmic law respectively

$$\begin{aligned}\mu_i &= p_0(i) + p_1(i) \cdot \text{FWHM}_0 \\ \sigma_i &= p_3(1) + p_4(i) \cdot \log(\text{FWHM}_0).\end{aligned}\tag{10.1}$$

Here, FWHM_0 is the full width half maximum of the Gaussian describing the energy resolution of unscattered electrons, and the parameters p_1 to p_4 (table B.1) are derived from fits in a toy model analysis in which this Gaussian was convolved with the energy loss function from [88]. To obtain the complete lineshape, the individual scattering peaks are summed, with each successive peak contributing less than the previous scatter. The ratio between the scattering peaks corresponds to the survival probability from equation (4.21). Figure 10.13 shows how the full lineshape is composed by summing the weighted probability density functions of N Gaussian distributions.



(a) Each scattering peak is approximated by a Gaussian with μ_i and σ_i (10.1). (b) The full lineshape consists of the normalized sum of $N = 20$ individual scattering peaks.

Figure 10.13: Simplified lineshape model.

FWHM₀ and the survival probability $p_{\text{scattering}}$ are extracted from fits of the detailed lineshape model to calibration $^{83\text{m}}\text{Kr}$ data that was taken in the Q300 trap (figure 4.19). Because in deep trap data, the lineshape fits are not sensitive to the gas composition, the proportion of scatters from each gas species must be derived from the cross section and the partial pressure measurements of an RGA (Residual Gas Analyzer) in the gas system. Unfortunately, the partial pressures measured are not very reliable due to the temperature difference and the distance of the gas system section in which the RGA is located from the CRES cell. To account for the uncertainty of the gas composition, the $^{83\text{m}}\text{Kr}$ calibration data was fitted several times assuming a different composition in each fit. The spread of fit results for the energy resolution

$$\sigma_0 = \text{FWHM}_0 \cdot \frac{1}{2 \cdot \sqrt{2 \cdot \log(2)}} \quad (10.2)$$

and the survival probability $p_{\text{scattering}}$ is taken as systematic uncertainty on these parameters. The resulting uncertainty and mean value for the survival probability, energy resolution, and magnetic field strength that are used for the analysis in this work are identical to the input parameters to the Bayesian tritium analysis result presented at the Neutrino 2020 conference [140]:

$$\begin{aligned} p_{\text{scattering}} &= 0.67 \pm 0.11 \\ \sigma_0 &= (15.5 \pm 2.2) \text{ eV} \\ B &= (0.9578095 \pm 8.5 \cdot 10^{-6}) \text{ T} \end{aligned} \quad (10.3)$$

10.2.4 Propagation of systematic uncertainties

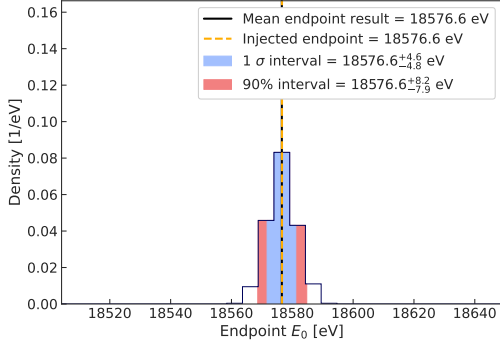
The sources of systematic uncertainty that are propagated in the Phase II tritium analysis are the detection efficiency, the lineshape parameters, and the magnetic field magnitude (10.3). Of these, only the shape of the uncertainty of the survival probability $p_{\text{scattering}}$ cannot be taken as Gaussian. Its uncertainty is so large that, if a bell-shaped representation was used, a random sample might fall outside of the allowed range $0 \leq p_{\text{scattering}} < 1$. To ensure it is contained between 0 and 1, the uncertainty on the survival probability is modeled as a Beta distribution

$$\text{PDF}_{\text{Beta}} = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{B}(\alpha, \beta)}, \quad (10.4)$$

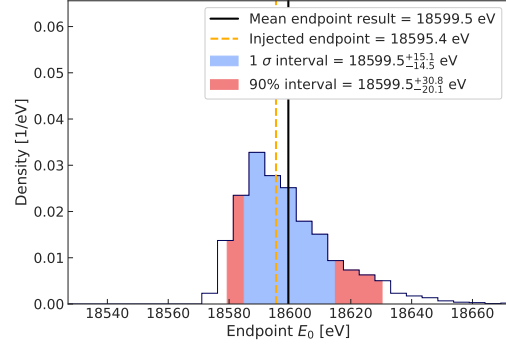
where $\text{B}(\alpha, \beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha+\beta)}$ with the Gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$. The parameters α and β are chosen such that the mean and standard deviation of the Beta distribution describe the uncertainty on $p_{\text{scattering}}$.

10.2. ANALYSIS OF THE PHASE II TRITIUM SPECTRUM

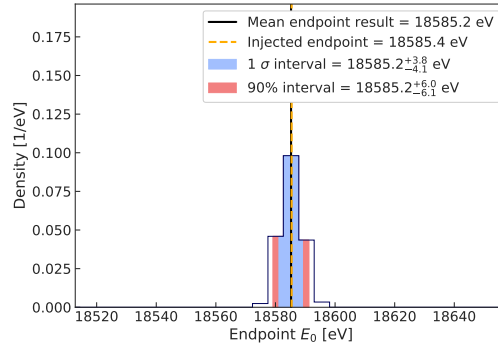
To obtain the induced systematic endpoint uncertainties, each systematic contribution is separately varied in a Monte Carlo study, in which an Asimov data set including all distortions (efficiency and lineshape) is repeatedly fitted and statistical fluctuations are consequently eliminated. Figure 10.14 shows the spread of endpoint fit results for fake data consisting of approximately 5000 events in a spectrum with the same ROI as the real Phase II data.



(a) Endpoint results from sampling B



(b) Endpoint results from sampling the lineshape parameters σ_0 and $p_{\text{scattering}}$



(c) Endpoint results from sampling the efficiency as in figure 10.5b

Figure 10.14: Systematic endpoint uncertainties in fake data experiments. An Asimov data set is fitted repeatedly using a model with randomly sampled systematic parameters. The parameters belonging to each systematic effect are sampled separately to obtain the distribution of endpoint fits originating from the uncertainties on the magnetic field (a), the lineshape parameters (b) and the efficiency (c). All intervals are given relative to the mean.

The largest systematic uncertainty on the endpoint result is caused by the scattering lineshape. The energy resolution alone has no considerable impact on the endpoint result. The entire width of endpoint uncertainty in figure 10.14b originates from sampling $p_{\text{scattering}}$. The asymmetry of the uncertainty of $p_{\text{scattering}}$ and its nonlinear correlation with the endpoint result causes the mean of the endpoint fit results to be offset from the injected (initial-fit) value. Adding all systematic uncertainties from figure 10.14

in quadrature yields a total uncertainty of approximately ± 16 eV which is just slightly smaller than the total uncertainty from combined sampling in figure 10.15.

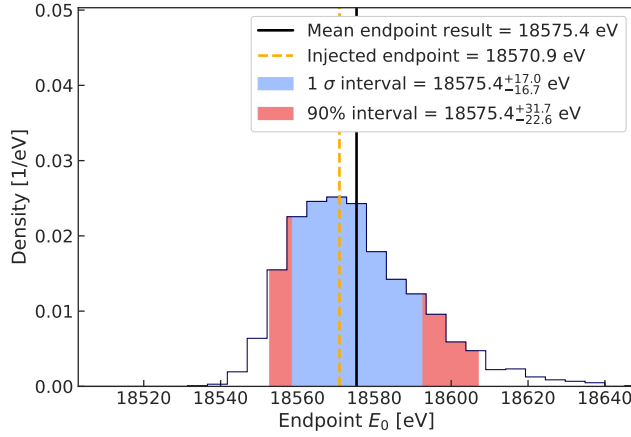


Figure 10.15: Full systematic endpoint uncertainty in fake data experiment with 5000 events. All uncertainties from figure 10.14 are propagated simultaneously.

10.2.5 Full model validation

Before the real tritium data is analyzed, the full model including all systematic effects is tested in fake data experiments. To this end, 200 sets of fake events are generated from the convolved and distorted detailed tritium model. In the data generation process, the true lineshape parameters, bin efficiencies and magnetic field strengths used to generate the fake data are all randomly drawn from their uncertainty distributions (Beta for survival probability, Gaussian for all other parameters). In addition, each fake data set has a different, random true endpoint. As in section 10.2.2, the values of the true endpoints are sampled from a Gaussian distribution with a standard deviation of 50 eV centered on the literature endpoint [90]. This ensures that the model is tested for biases over a range that is larger than the expected total endpoint uncertainty. The number of events in each fake data set is drawn from a Poisson distribution with a mean of 5000 events.

Just as before, the fake data is fitted with the approximate analysis model, whereby all systematic parameters are fixed at first to their presumed best value. This initial-fit result serves as starting point for the uncertainty propagation. To obtain the total endpoint uncertainty for each fake data experiment, the data and the systematic parameters are randomly sampled simultaneously. This is done by repeatedly generating new fake data from the initial-fit model and re-fitting them with a model defined by the randomized systematic parameters. Figure 10.16 shows the endpoint fit offsets of the 200 fake data experiments as a function of the true endpoint. The average 1σ interval is 22 eV wide and $(70.4 \pm 3.2)\%$ of the intervals contain the true endpoint (the 90% interval coverage

10.2. ANALYSIS OF THE PHASE II TRITIUM SPECTRUM

is $(92.3 \pm 1.8)\%$. The mean of the fit results is offset from the truths by (-1.7 ± 1.6) eV, which is probably a coincidence and not significant for Phase II. It can be seen that the mean of the initial-fit endpoints (injected endpoints) no longer converges to the mean of the true endpoint values. This is due to the impact of $p_{\text{scattering}}$ (figure 10.14b). Since the mean of the endpoint fits is unbiased, it is considered the best-fit result.

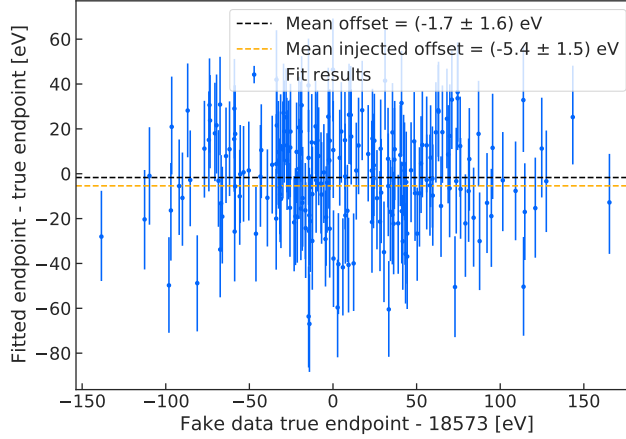


Figure 10.16: Full model test in fake data studies. Fake data is generated from the detailed model and fitted with the approximate analysis model. Both models contain the lineshape and efficiency distortion of the spectrum. The true endpoint of each fake data set is randomly drawn from a Gaussian distribution centered on 18573.24 eV with a standard deviation of 50 eV. In the fitting process the statistical and the systematic uncertainties are propagated simultaneously via Monte Carlo sampling.

10.2.6 Full model fit to real data

After the tritium model has been thoroughly tested in fake data experiments, the real tritium data shown in figure 10.10 is analyzed here. Again, the free parameters in the maximum likelihood fit are the number of background and signal events as well as the endpoint position. As in section 10.2.5, all uncertainties are propagated via Monte Carlo sampling. Figure 10.17 shows the analysis model fitted to the real tritium data in the ROI (energy range that is covered by the efficiency curve plus the background range above the endpoint). The model describes the data very well, which is confirmed by a KS-test, testing the residuals in the signal region (below the endpoint) for normality, resulting in a p -value of 0.95. The endpoint result including statistical and systematic uncertainties is $18555.6^{+24.0}_{-23.9}$ eV and it contains the literature endpoint (figure 10.18c).

10.2. ANALYSIS OF THE PHASE II TRITIUM SPECTRUM

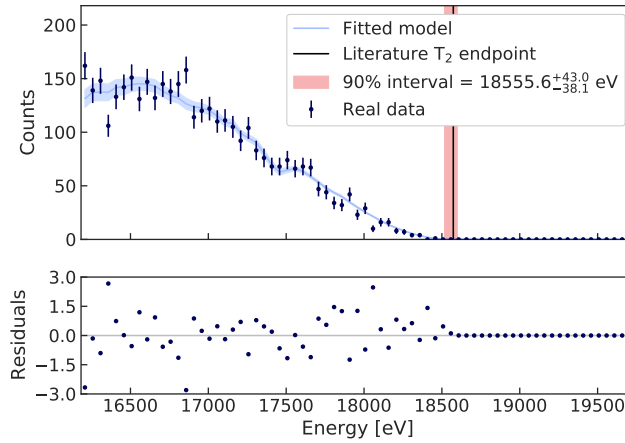


Figure 10.17: Real tritium data converted to energy and fitted with the approximate analysis model.

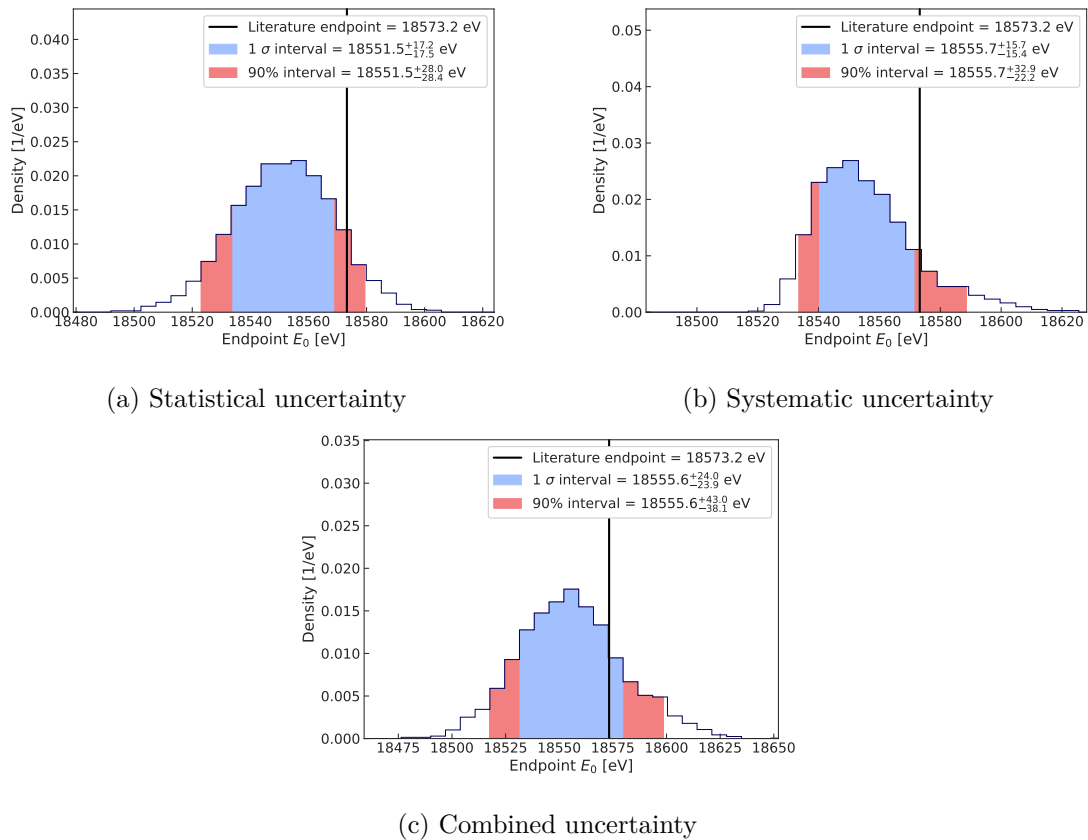


Figure 10.18: Endpoint fit results obtained in the analysis of real tritium data with statistical (a), systematic (b) and all (c) uncertainties propagated via Monte Carlo sampling. The literature endpoint of 18573.24 eV [90] is contained in the 1 σ interval $E_0 = 18555.6^{+24.0}_{-23.9}$ eV when all uncertainties are propagated.

10.2.7 Remaining uncertainties

The fake data studies in section 10.2.5 have demonstrated that the analysis is unbiased under the assumption that the propagated systematic uncertainties reflect the true uncertainty ranges. The offset in the analysis of the real data from the literature endpoint in figure 10.18 is insignificant, which confirms that the systematic contributions to the spectrum in Phase II are well understood and their uncertainties are adequately small. In section 10.1.3 the effect of a systematic tilt of the efficiency curve was a shift of the endpoint by several eV. Although the distorted tritium model already matches the recorded data very well (figure 10.17), the fit to the real data is repeated with the tilt added as a free parameter. The fit result for the tilt is $(0.0002 \pm 0.0052) \text{ keV}^{-1}$, which is compatible with 0 keV^{-1} and does not hint at a missed (or mis-scaled) effect in the construction of the efficiency curve. Thus, the result from section 10.2.6 stands.

Currently, the largest uncertainty in the Phase II analysis originates from the lineshape model. The multi-gas scattering model can describe the shape of the measured K-line peak well, even if the fitted contributions of each gas take unrealistic values, simply by adjusting the other model parameters accordingly. While the work on improving the model is ongoing, it is not guaranteed that a more reliable result can be obtained, since in the deep trap ($\Delta B \geq 1 \text{ mT}$) the shapes of the main peak and the scattering tail merge into a single shape. This makes it inherently difficult to distinguish scatters from individual gases and their respective contributions to the tail. In addition, the current model does not reflect the fact that the lineshape is a superposition of four lineshapes, one from each sub-trap. It has further been shown that describing the energy resolution function as Gaussian is insufficient in deep traps. A correct lineshape description is all the more important, since the magnetic field strength, the crucial parameter in converting the recorded frequency into energy, is extracted from a krypton line fit in the Q300 trap. Moreover, the lineshape convolved with the tritium spectrum is not identical to the krypton lineshape, because the spectral shape underlying the K-line and the scattering of krypton gas molecules only affect krypton data and are absent in tritium data. The situation could be improved if the gas composition in the cell were monitored more accurately. This way, one set of important unknown parameters in the lineshape model, namely the scattering contribution from individual gas species, could be eliminated from the uncertainty budget. But as mentioned before, the gas composition monitoring in Phase II is limited to the supplying gas system in a location relatively far away from the cell.

10.2.8 Neutrino mass

In the previous sections, m_β was always fixed to 0 eV. The fact that the endpoint best-fit result is offset from the literature value by 17.6 eV could already be taken as an upper limit on the neutrino mass. In theory, however, the endpoint result could be subject to a systematic bias and the mass is therefore usually extracted in a shape analysis (see section 2.6). The neutrino mass modifies the tritium spectrum by shifting its endpoint and distorting its shape near the endpoint by the shape factor

$$\frac{dN}{dE} \propto (E_0 - E) \cdot \sqrt{(E_0 - E)^2 - m_\beta^2}. \quad (10.5)$$

Since the neutrino mass enters the shape of the spectrum as m_β^2 , the actual parameter used in the analysis model is the squared mass. The endpoint of the spectrum is shifted by the absolute square root of m_β^2 .

Although the number of events near the endpoint in the Phase II tritium data is too small to allow a restriction of the neutrino mass on the eV-scale, a proper fit is of course still of interest. To obtain a neutrino mass fit result, the analysis from section 10.2.6 is repeated, this time treating m_β^2 as an additional free parameter. Figure 10.20 shows the contour plot of m_β^2 and E_0 . A strong correlation between the two parameters can be seen. As a result, the endpoint fit result is shifted upwards by 15 eV compared to section 10.2.6 and its uncertainty increased to approximately ± 29 eV (figure 10.19a). The 90% interval of m_β^2 is $0.009^{+0.031}_{-0.018}$ keV². The fit result is compatible with $m_\beta = 0$ eV and allows to set an upper limit on the mass via the sensitivity method by [141] of $m_\beta < 221$ eV (90% C.L.). As expected, the tritium data recorded with the Phase II apparatus does not allow to set a strict limit via the shape analysis. Propagating the statistical uncertainty alone yields an upper limit on m_β of 214 eV.

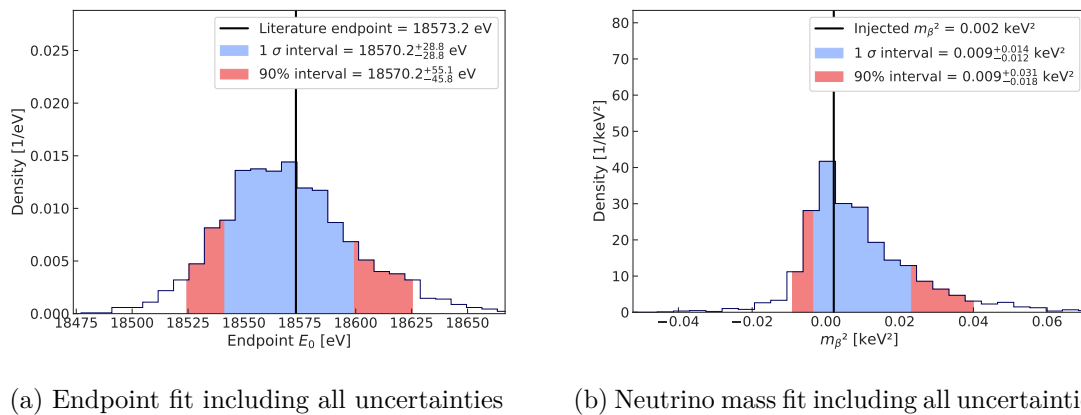


Figure 10.19: Endpoint (a) and neutrino mass squared (b) fit results when m_β^2 is treated as free parameter in the Phase II tritium analysis.

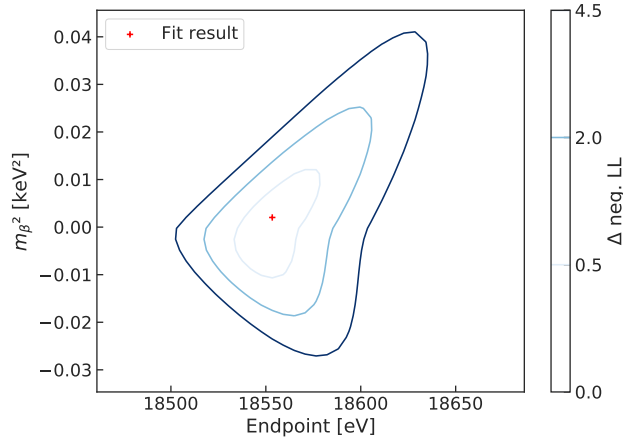


Figure 10.20: Squared neutrino mass vs. endpoint contours. The correlations between the neutrino mass and other model parameters is more than a factor 10 smaller than the correlation with the endpoint value. In this plot, all parameters but the endpoint and the squared neutrino mass are fixed to their best-fit values.

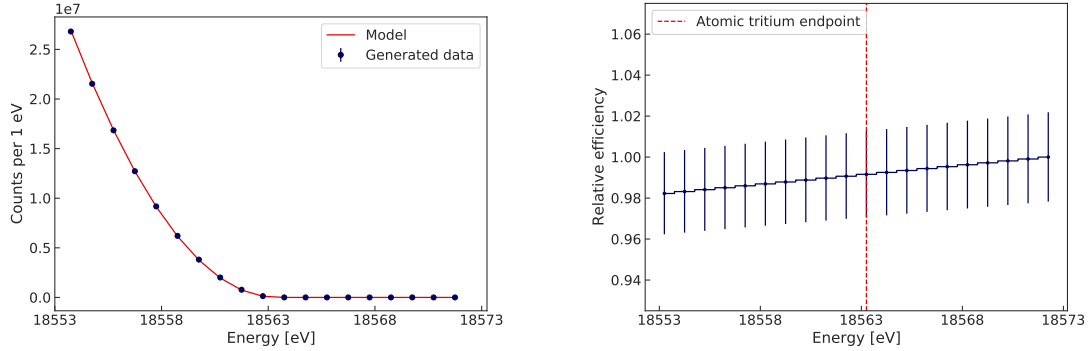
10.3 Detection efficiency in future Phases

10.3.1 Efficiency uncertainty in a Phase IV scenario

In Phase IV the goal is to achieve an upper limit on the neutrino mass of 40 meV. This requires a precise knowledge of the event rate especially in the last eV of the spectrum (see section 3.1.2). In Phase II the efficiency uncertainty close to the endpoint is about 2%, which is insufficient for a neutrino mass determination. Nonetheless, a Phase IV scenario is tested here using the efficiency curve from Phase II interpolated to 1 eV-wide bins (figure 10.21b). The true neutrino mass is set to 0 eV and 10^8 fake events are generated from the distorted model in a ROI ranging from 10 eV below to 10 eV above the atomic tritium endpoint (figure 10.21a).

The systematic uncertainty of the squared neutrino mass is obtained by repeatedly fitting an Asimov data set, while randomly sampling the efficiency independently in each bin. To test how small the uncorrelated efficiency uncertainty has to be for Phase IV, the experiment is repeated four times while the efficiency uncertainty is scaled down by a factor of 0.25^i in each experiment i . The full width of the 90% systematic uncertainty interval as a function of the relative efficiency uncertainty is shown in figure 10.22.

10.3. DETECTION EFFICIENCY IN FUTURE PHASES



(a) Fake data generated from a distorted tritium model in a 20 eV energy range centered on the atomic tritium endpoint (18563.25 eV). The energy resolution is set to 0.1 eV.

(b) Phase II efficiency interpolated to 1 eV bins in the Phase IV ROI.

Figure 10.21: Data and efficiency in a Phase IV scenario based on the Phase II detection efficiency studies.

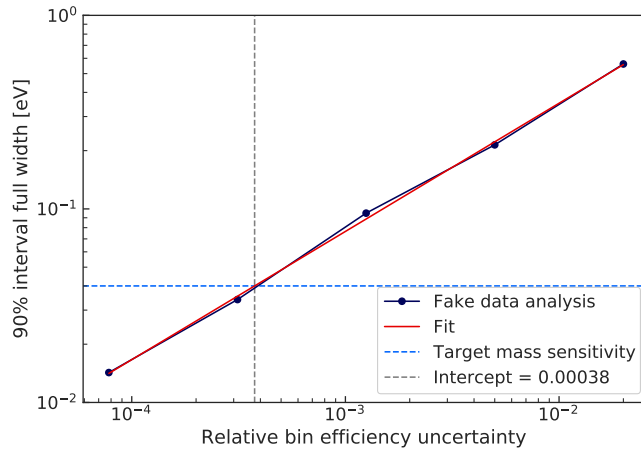


Figure 10.22: 90 % systematic neutrino mass interval (full width) in a simulated Phase IV scenario as a function of the relative efficiency uncertainty in 1 eV bins. The blue dashed line marks the 40 meV sensitivity goal for Phase IV.

The relation of the systematic neutrino mass interval to the efficiency uncertainty follows a double logarithmic law. The fit function can be solved for the relative efficiency uncertainty

$$\frac{\sigma_\epsilon}{\epsilon} = \exp\left(\frac{\log(\Delta m_\beta) - \log(c) - a}{b}\right), \quad (10.6)$$

where $a = 4.743$, $b = 0.662$ and $c = 0.065$ eV. A 40 meV full width of the 90 % interval is reached for an uncorrelated efficiency uncertainty $< 3.7 \cdot 10^{-4}$. Assuming that the neutrino mass uncertainty from other sources (3.1.2) amounts to 40 meV in Phase IV (90 % interval width with a best fit of $m_\beta = 0$ eV) and requiring that the increase of this

10.3. DETECTION EFFICIENCY IN FUTURE PHASES

interval width must be ≤ 1 meV, limits the allowed systematic uncertainty caused by the detection efficiency to 9 meV. Hence, the efficiency has to be known with a precision of $\leq 1.1 \cdot 10^{-4}$ on a 1 eV scale (corresponding to a full width 90% interval of 18 meV).

10.3.2 Calibration in future phases

In section 9.2 various methods to measure the detection efficiency were discussed. In light of the maximum allowed uncertainty on the efficiency of $1.1 \cdot 10^{-4}$ for Phase IV, their feasibility is reconsidered here. First of all, it can be concluded from section 10.3.1 that in order to measure the efficiency with the required precision, $8.3 \cdot 10^7$ events per eV would have to be recorded in a detection efficiency calibration measurement. If this had to be achieved in an FSS scan, in addition to a much smaller frequency stepsize (49 kHz), an exposure $3.3 \cdot 10^4$ times higher than in the Phase II FSS scans would be required. The design exposure in Phase IV is 10^8 in the last 10 eV below the atomic spectrum endpoint in one year of data taking. The fraction of events in this energy range compared to the fraction in the last 2.4 keV of the molecular spectrum is $5.3 \cdot 10^{-8}$, leading to an expected event rate $\sim 1.1 \cdot 10^{11}$ times the rate in Phase II in the same energy range. If the number of krypton events scales in a similar way, the required number of events in an FSS scan could easily be accumulated in a short amount of time.

However, an FSS scan in Phase IV seems unfeasible for other reasons: The Phase III and Phase IV setups are complex and don't have much space available for an additional solenoid between the antenna array and the superconducting solenoid. In addition, a field scan is not allowed to modify the field geometry, which is challenging on the large scales and field precision requirements of the future Phases. It is questionable whether this can be achieved by the usage of an FSS or even a magnetic field shift of the superconducting background field solenoid. But most importantly, the CRES cells of the future experiments will be colder (30 mK in Phase IV) which prevents the performance of systematic krypton studies at identical run conditions in the first place. Therefore using an electron gun or a highly active β -emitter that is gaseous at low temperatures would be most appealing.

One such candidate is tritium itself. Instead of a field scan, a single large frequency shift (without modifying the trapping field geometry) could move the frequency endpoint at the nominal magnetic field magnitude far into the atomic tritium spectrum. In a 10 eV range at 2 keV below the endpoint, the decay rate is $\sim 1.6 \cdot 10^5$ times higher than in the uppermost 10 eV of the spectrum. Accumulating the required counts per eV for the efficiency calibration would take less than one hour. This, however, assumes that the detection efficiency is the same at 2 keV below the endpoint, which is unlikely to be true, as the emitted power decreases with energy. Even if the decrease in detection efficiency could be compensated for by a longer calibration duration, the observed ranges of event

properties would not be the same. For example, the range of radial positions at which electrons could be detected would differ at energies far away from the endpoint. Using tritium for efficiency studies therefore seems unfeasible.

On the other hand, depending on the exact combinations of signal processing components in the Phase IV detector, unknown small scale efficiency fluctuations are not necessarily expected if all transmission functions are assumed to be smooth. All fluctuations introduced by digital signal processing and event detection analysis could then easily be studied in simulation. It is important to note that a systematic tilt of the efficiency curve affects the global shape of the spectrum and thus the endpoint fit. It could therefore be detected from the spectral shape itself. Consequently, it could be sufficient to rule out the presence of small scale efficiency fluctuations from resonances or filter shapes near the endpoint. Still, relying on an analytical calculation of the efficiency, even in a very well-known system, carries the major risk of an unknown error in the neutrino mass and endpoint analysis. It is therefore highly recommended to measure the efficiency in some way. Developing a feasible method to analyze the detection efficiency of the Phase IV experiment is one of the major goals of the Phase III free space CRES demonstrator.

The absolute scale of neutrino masses remains an open question in the Standard Model of particle physics. In recent decades, a number of experiments have set increasingly strict limits on the effective mass of the electron antineutrino by recording the spectral shape of tritium decay near the high-energy endpoint. The Project 8 collaboration has developed Cyclotron Radiation Emission Spectroscopy (CRES) with the aim of recording a tritium spectrum with unprecedented precision, thus achieving a sensitivity to m_β of 40 meV. In CRES, electrons emitted in decays are trapped in a magnetic field where they undergo cyclotron motion and consequently emit radiation that is narrowly peaked at their cyclotron frequency. This frequency is inversely proportional to the Lorentz factor γ , which allows to reconstruct the kinetic energy with high precision.

To achieve its goal of a 40 meV sensitivity, Project 8 is taking a phased approach. The current Phase II apparatus served the purpose of recording the first tritium spectrum with CRES. The setup consists of a small gas volume of a few cm^3 inside a waveguide that is inserted into a superconducting NMR magnet. The cyclotron radiation emitted by trapped electrons couples to a traveling waveguide mode and is transported to the receiver. Once the signal leaves the waveguide, it is amplified by a series of cryogenic and room-temperature amplifiers. The analog signal is digitized by a ROACH2 board, which computes its Fourier transform and streams the time- and frequency-domain samples to a server where they are processed by a software trigger. The trigger continuously inspects the frequency-domain data for the appearance of CRES-like signal characteristics. Whenever an event candidate is found, the corresponding time-domain data is written to disk for offline analysis. In the offline event reconstruction process, a normalized power threshold is applied to the Fourier-transformed power spectrogram and all bins that are not expected to be part of an electron signal are removed. The resulting sparse spectrogram is searched for electron tracks consisting of a series of bins that fall onto a straight line. With the help of simulated fake electron events, the end-to-end performance of the event detection process with respect to various event properties was studied. In addition, the frequency dependence of the event detection efficiency was investigated by the means of magnetic field scans. Combined with a response matrix obtained via simulation, an efficiency curve describing the dependence of the detection efficiency on energy was constructed. This curve is a key ingredient to the analysis of a spectrum recorded with CRES since the shape of any recorded spectrum corresponds to the product of the intrinsic spectral shape and the efficiency curve.

Using the Phase II apparatus, a tritium spectrum was recorded over the duration of 82 days. For the purpose of its analysis, a model including all systematic effects was developed and extensively tested in fake data experiments. The uncertainties of the detection efficiency and other systematic effects were propagated in Monte Carlo pseudo-experiments. It was shown that the analysis model is unbiased and the total systematic endpoint uncertainty amounts to 16 eV. The full frequentist analysis of the real tritium data including statistical and systematic uncertainties gave an endpoint result of $18555.6^{+24.0}_{-23.9}$ eV, which is in agreement with the literature value, and an upper limit on the neutrino mass of 221 eV.

In future phases, the CRES signals will be received by a cylindrical array of antennas, adding additional dimensions to the challenge of analysis. The antenna signals will be beamformed and thus sensitive to the spatial locations of the electrons in the gas cell volume. Because of the larger number of channels, the longer run duration, and the significantly higher event rate, the data reduction requirement is at least one order of magnitude higher than the reduction achieved in Phase II. It can only be met by establishing real-time event reconstruction. In such an online analysis system, a pre-selected region in location (ρ, ϕ) , frequency and time could be searched by one or several algorithms for the exact electron track positions. The frequency and energy dependence of this system will require careful calibration as it was shown in this work that the efficiency precision needed in Phase IV is $\leq 1.1 \cdot 10^{-4}$. To some extent, the slope of the detection efficiency curve with respect to energy and frequency can be extracted from the shape of the tritium spectrum over the entire ROI. But small scale efficiency fluctuations near the endpoint must be either ruled-out or measured. In principle, the efficiency could be calibrated in field scans similar to the Phase II FSS scan. However, a much higher statistical power will be required and many technical challenges will have to be tackled. For example, the source gas that is used to calibrate the system will need to be stored in the same container at the same temperature as the tritium gas. Which calibration method is feasible in Phase IV is one of the open challenges to be investigated in Phase III.

Appendices

A

Track and event properties

Table A.1: Track properties

| Property name | Unit | Description |
|----------------------|-------------|---|
| StartFrequency | MHz | Frequency coordinate of track start |
| EndFrequency | MHz | Frequency coordinate of track end |
| StartTimeInAcq | s | Time coordinate of track start since start of the acquisition |
| StartTimeInRun | s | Time coordinate of track start since start of the run |
| EndTimeInRun | s | Time coordinate of track end since start of the run |
| TimeLength | s | EndTimeInRun – StartTimeInRun |
| Slope | Hz/s | Track frequency drift |
| TotalPower | MHz | Sum of the track bin powers |
| NTrackBins | unit-less | Number of points that constitute a track |
| TotalTrackSNR | unit-less | Sum of the SNR of all track points |
| MaxTrackSNR | unit-less | SNR of the track point with the maximum SNR |
| TotalTrackNUP | unit-less | Like TotalTrackSNR but summing the normalized power residuals |
| MaxTrackNUP | unit-less | Like MaxTrackSNR but using NUP instead of SNR |

Table A.2: Event properties

| Property name | Unit | Description |
|--------------------------|-------------|--|
| AcquisitionID | unitless | Number of acquisition in the run |
| EventID | unitless | Unique identifier to an event from an egg file; As runs can have multiple files, EventID is not unique in a run. |
| TotalEventSequences | unitless | Number of tracks in this event |
| StartTimeInAcquisition | s | StartTimInAcquisition of the earliest track in the event |
| StartTimeInRun | s | StartTimeInRun of the earliest track in the event |
| EndTimeInRun | s | EndTimeInRun of track with the latest end time |
| TimeLength | s | Full duration of the event: EndTimeInRun – StartTimeInRun |
| StartFrequency | MHz | StartFrequency of the earliest track in the event |
| EndFrequency | MHz | EndFrequency of the latest ending track |
| FirstTrackID | unitless | TrackID of the earliest track |
| FirstTrackTimeLength | s | TimeLength of earliest track |
| FirstTrackFrequencyWidth | Hz | FrequencyWidth of earliest track |
| FirstTrackSlope | Hz/s | Slope of earlist track |
| FirstTrackIntercept | Hz | Intercept of earliest track |
| FirstTrackTotalPower | W/Hz | TotalPower of earliest track |
| FirstTrackNTrackBins | unitless | NTrackBins of earliest track |
| FirstTrackTotalSNR | unitless | TotalSNR of earliest track |
| FirstTrackMaxSNR | unitless | MaxSNR of earliest track |
| FirstTrackTotalNUP | unitless | TotalNUP of earliest track |
| FirstTrackMaxNUP | unitless | MaxNUP of earliest track |

Listing A.1: Katydid configuration

```

processor-toolbox:
  processors:
    - type: egg-processor
      name: egg1
    - type: forward-fftw
      name: fft1
    - type: convert-to-power
      name: to-ps1

    - type: egg-processor
      name: egg2
    - type: forward-fftw
      name: fft2
    - type: convert-to-power
      name: to-ps2

    - type: data-accumulator
      name: acc
    - type: gain-variation
      name: gainvar

    - type: variable-spectrum-discriminator
      name: discrim
    - type: sequential-track-finder
      name: seq-tr-clust
    - type: overlapping-track-clustering
      name: otc
    - type: iterative-track-clustering
      name: itc
    - type: track-proc-ws
      name: tr-proc-ws
    - type: multi-peak-track-builder
      name: mptb
    - type: multi-peak-event-builder
      name: mpeb

    - type: apply-cut
      name: ac1b
    - type: apply-cut
      name: event-nup-cut
    - type: apply-cut
      name: event-time-cut

    - type: basic-root-writer
      name: brw
    - type: root-tree-writer
      name: trw

  connections:
    # Header processing
    - signal: "egg1:header"
      slot: "fft1:header"

    - signal: "egg1:ts"
      slot: "fft1:ts-fftw"

    # First egg processing
    - signal: "fft1:fft"
      slot: "to-ps1:fs-fftw-to-psd"

    - signal: "to-ps1:psd"
      slot: "acc:ps"

    - signal: "acc:ps-finished"
      slot: "gainvar:ps-var"

    - signal: "acc:ps-finished"
      slot: "brw:psd"          # Write accumulated PSD
                              histogram

    - signal: "gainvar:gain-var"
      slot: "brw:gain-var"    # Write gain-var histogram

    # Second egg processing
    - signal: "egg2:header"
      slot: "fft2:header"

    - signal: "egg2:ts"
      slot: "fft2:ts-fftw"

    - signal: "fft2:fft"
      slot: "to-ps2:fs-fftw-to-psd"

    - signal: "to-ps2:psd"
      slot: "discrim:ps-pre"

    - signal: "discrim:disc-1d"
      slot: "seq-tr-clust:disc-1d"

    #
    #
    - signal: "discrim:disc-1d"
      slot: "trw:disc-1d"    # Writing discriminated
                              points to ROOT tree

    - signal: "seq-tr-clust:seq-cand"
      slot: "otc:seq-cand"

    - signal: "otc:seq-cand"
      slot: "itc:seq-cand"

    - signal: "itc:seq-cand"
      slot: "ac1b:apply"

    - signal: "ac1b:pass"
      slot: "tr-proc-ws:seq-cand"

    - signal: "tr-proc-ws:track"
      slot: "mptb:track"

    - signal: "tr-proc-ws:track"
      slot: "trw:proc-track"  # Write processed tracks
                              to a ROOT tree

    - signal: "egg2:egg-done"
      slot: "seq-tr-clust:done"

    - signal: "seq-tr-clust:clustering-done"
      slot: "otc:do-clustering"

    - signal: "otc:clustering-done"
      slot: "itc:do-clustering"

    - signal: "itc:clustering-done"
      slot: "mptb:do-clustering"

    - signal: "mptb:mpt"
      slot: "mpeb:mpt"

    - signal: "mptb:mpt-done"
      slot: "mpeb:do-clustering"

    - signal: "mpeb:event"
      slot: "event-nup-cut:apply"

    - signal: "event-nup-cut:pass"
      slot: "event-time-cut:apply"

    - signal: "event-time-cut:pass"
      slot: "trw:mt-event"

  run-queue:
    - egg1

```

```

- egg2
min-total-nup: 0
min-average-nup: 7.0
time-or-bin-average: bin
wide-or-narrow: wide

egg1:
filename: "foo.egg"
egg-reader: egg3
number-of-slices: 0
start-record: 1
slice-size: 4096
progress-report-interval: 10000

event-nup-cut:
ntracks-npoints-nup-cut:
wide-or-narrow: wide
time-or-bin-average: bin

fft1:
transform-flag: ESTIMATE

default-parameters: #thresholds for event first track
nup cut
min-total-nup: 0
min-average-nup: 7.454357052489394
min-max-nup: 0

acc:
number-to-average: 0
signal-interval: 0

parameters:
- ft-npoints: 3
ntracks: 1
min-total-nup: 0
min-average-nup: 15.74668644848628
min-max-nup: 0

gainvar:
normalize: false
min-frequency: 0
max-frequency: 100e6
fit-points: 50

egg2:
filename: "foo.egg"
egg-reader: egg3
number-of-slices: 0
start-record: 1
slice-size: 4096
progress-report-interval: 10000

- ft-npoints: 3
ntracks: 2
min-total-nup: 0
min-average-nup: 14.396403491431705
min-max-nup: 0

fft2:
transform-flag: ESTIMATE

- ft-npoints: 3
ntracks: 3
min-total-nup: 0
min-average-nup: 13.032932650345844
min-max-nup: 0

discrim:
min-frequency: 5e6
max-frequency: 95e6
#snr-threshold-power: 6.0
sigma-threshold: 4.0
normalize: true
neighborhood-radius: 2

- ft-npoints: 3
ntracks: 4
min-total-nup: 0
min-average-nup: 11.651179906121854
min-max-nup: 0

seq-tr-clust:
min-frequency: 5e6
max-frequency: 95e6
initial-slope : 800e6
slope-method: weighted-first-point-ref
n-slope-points: 10
time-gap-tolerance: 0.1e-3
frequency-acceptance: 56e3
trimming-threshold: 6.0
min-points: 3
min-slope: 0

- ft-npoints: 3
ntracks: 5
min-total-nup: 0
min-average-nup: 10.242207772240628
min-max-nup: 0

otc:
max-track-width: 150e3

- ft-npoints: 3
ntracks: 6
min-total-nup: 0
min-average-nup: 8.78782196327628
min-max-nup: 0

itc:
time-gap-tolerance: 1.0e-3
frequency-acceptance: 100.0e3
max-track-width: 100.0e3

- ft-npoints: 4
ntracks: 1
min-total-nup: 0
min-average-nup: 12.735418234003228
min-max-nup: 0

tr-proc-ws:
min-slope: 0

- ft-npoints: 4
ntracks: 2
min-total-nup: 0
min-average-nup: 11.681651640175646
min-max-nup: 0

mptb:
sideband-time-tol: 1e-3

- ft-npoints: 4
ntracks: 3
min-total-nup: 0
min-average-nup: 10.610759932504259
min-max-nup: 0

mpeb:
jump-time-tol: 3.0e-3

- ft-npoints: 4
ntracks: 4
min-total-nup: 0
min-average-nup: 9.515056968514685

ac1b:
seq-line-nup-cut:

```

```

min-max-nup: 0

- ft-npoints: 4
  ntracks: 5
  min-total-nup: 0
  min-average-nup: 8.379691045166915
  min-max-nup: 0

- ft-npoints: 5
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 10.869025304065605
  min-max-nup: 0

- ft-npoints: 5
  ntracks: 2
  min-total-nup: 0
  min-average-nup: 9.987041877860296
  min-max-nup: 0

- ft-npoints: 5
  ntracks: 3
  min-total-nup: 0
  min-average-nup: 9.083012737896947
  min-max-nup: 0

- ft-npoints: 5
  ntracks: 4
  min-total-nup: 0
  min-average-nup: 8.144983909889586
  min-max-nup: 0

- ft-npoints: 6
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 9.831510394306793
  min-max-nup: 0

- ft-npoints: 6
  ntracks: 2
  min-total-nup: 0
  min-average-nup: 9.148730213081482
  min-max-nup: 0

- ft-npoints: 6
  ntracks: 3
  min-total-nup: 0
  min-average-nup: 8.436488780835134
  min-max-nup: 0

- ft-npoints: 6
  ntracks: 4
  min-total-nup: 0
  min-average-nup: 7.67794591793375
  min-max-nup: 0

- ft-npoints: 7
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 8.893879615160312
  min-max-nup: 0

- ft-npoints: 7
  ntracks: 2
  min-total-nup: 0
  min-average-nup: 8.294285562892895
  min-max-nup: 0

- ft-npoints: 7
  ntracks: 3
  min-total-nup: 0
  min-average-nup: 7.658861173418163
  min-max-nup: 0

- ft-npoints: 8
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 8.277213651025455
  min-max-nup: 0

- ft-npoints: 8
  ntracks: 2
  min-total-nup: 0
  min-average-nup: 7.722303890626092
  min-max-nup: 0

- ft-npoints: 9
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 7.9962037692806796
  min-max-nup: 0

- ft-npoints: 9
  ntracks: 2
  min-total-nup: 0
  min-average-nup: 7.46339335340889
  min-max-nup: 0

- ft-npoints: 10
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 7.725369787485035
  min-max-nup: 0

- ft-npoints: 11
  ntracks: 1
  min-total-nup: 0
  min-average-nup: 7.454357052489394
  min-max-nup: 0

event-time-cut:
  event-time-in-acq-cut:
    min-time: 0.5e-3
    max-time: 5

brw:
  output-file: "GainVariation.root"
  file-flag: recreate

trw:
  output-file: "TracksAndEvents.root"
  file-flag: recreate

```

B

Simplified lineshape

Table B.1: Model parameters

| Scatter i | p_1 | σ_{p_1} | p_2 | σ_{p_2} | p_3 | σ_{p_3} | p_4 | σ_{p_4} |
|-------------|---------|----------------|---------|----------------|---------|----------------|----------|----------------|
| 1 | 1.94 | 3.83e-3 | 4.1e-1 | 5.79e-5 | 1.76e+1 | 7.24e-3 | 2.55e-2 | 2.1e-3 |
| 2 | 3.83 | 2.54e-3 | 3.96e-1 | 4.36e-5 | 3.51e+1 | 3.54e-3 | 5.62e-2 | 1.18e-3 |
| 3 | 5.42 | 2.86e-3 | 3.85e-1 | 5.12e-5 | 5.28e+1 | 4.33e-3 | 6.87e-2 | 1.44e-3 |
| 4 | 6.75 | 2.76e-3 | 3.79e-1 | 4.69e-5 | 7.05e+1 | 4.29e-3 | 6.38e-2 | 1.39e-3 |
| 5 | 8.25 | 2.89e-3 | 3.68e-1 | 4.88e-5 | 8.81e+1 | 4.68e-3 | 6.95e-2 | 1.5e-3 |
| 6 | 9.58 | 2.69e-3 | 3.61e-1 | 4.68e-5 | 1.06e+2 | 4.08e-3 | 6.46e-2 | 1.35e-3 |
| 7 | 1.07e+1 | 3.17e-3 | 3.56e-1 | 5.23e-5 | 1.24e+2 | 5.87e-3 | 6.15e-2 | 1.83e-3 |
| 8 | 1.21e+1 | 3.08e-3 | 3.46e-1 | 5.54e-5 | 1.41e+2 | 5.05e-3 | 5.38e-2 | 1.67e-3 |
| 9 | 1.3e+1 | 3.27e-3 | 3.44e-1 | 5.2e-5 | 1.59e+2 | 6.4e-3 | 4.94e-2 | 1.95e-3 |
| 10 | 1.45e+1 | 3.2e-3 | 3.3e-1 | 6.01e-5 | 1.77e+2 | 2.67e-3 | 1.58e-2 | 8.83e-4 |
| 11 | 1.54e+1 | 3.72e-3 | 3.3e-1 | 5.78e-5 | 1.94e+2 | 6.49e-3 | 2.81e-2 | 1.93e-3 |
| 12 | 1.66e+1 | 3.21e-3 | 3.21e-1 | 5.61e-5 | 2.12e+2 | 3.58e-3 | 1.11e-2 | 1.15e-3 |
| 13 | 1.78e+1 | 3.01e-3 | 3.16e-1 | 5.15e-5 | 2.3e+2 | 4.08e-3 | 1.23e-2 | 1.33e-3 |
| 14 | 1.85e+1 | 3.69e-3 | 3.14e-1 | 6.13e-5 | 2.47e+2 | 5.83e-3 | 9.78e-3 | 1.81e-3 |
| 15 | 1.94e+1 | 3.86e-3 | 3.1e-1 | 6.13e-5 | 2.65e+2 | 6.99e-3 | 6.22e-3 | 2.11e-3 |
| 16 | 2.06e+1 | 3.51e-3 | 3.02e-1 | 5.81e-5 | 2.83e+2 | 5.63e-3 | 3.54e-3 | 1.78e-3 |
| 17 | 2.16e+1 | 3.37e-3 | 2.97e-1 | 5.78e-5 | 3.01e+2 | 4.42e-3 | -2.53e-3 | 1.44e-3 |
| 18 | 2.26e+1 | 3.45e-3 | 2.91e-1 | 5.83e-5 | 3.18e+2 | 4.85e-3 | -3.08e-3 | 1.57e-3 |
| 19 | 2.33e+1 | 3.8e-3 | 2.92e-1 | 5.94e-5 | 3.36e+2 | 6.8e-3 | -2.22e-3 | 2.07e-3 |
| 20 | 2.4e+1 | 4.32e-3 | 2.88e-1 | 7.03e-5 | 3.54e+2 | 7.69e-3 | -2.44e-3 | 2.35e-3 |

Bibliography

- [1] J. Chadwick, *Intensitätsverteilung im magnetischen Spectrum der β -Strahlen von radium B + C*, Verhandl. Dtsch. Phys. Ges. **16**, 383 (1914).
- [2] C. D. Ellis and W. A. Wooster, *The average energy of disintegration of radium E*, Proc. Roy. Soc. Lond. A **117** (776), 109 (1927).
- [3] L. Meitner and W. Orthmann, *Über eine absolute Bestimmung der Energie der primären β -Strahlen von Radium E*, Z. Phys. **60** (3), 143 (1930).
- [4] W. Pauli (1930), *Pauli letter collection: letter to Lise Meitner*, typed copy.
- [5] F. L. Wilson, *Fermi's Theory of Beta Decay*, Am. J. Phys. **36** (12), 1150 (1968).
- [6] M. Goldhaber and G. Scharff-Goldhaber, *Identification of Beta-Rays with Atomic Electrons*, Phys. Rev. **73**, 1472 (1948).
- [7] C. L. Cowan, F. Reines, et al., *Detection of the free neutrino: A Confirmation*, Science **124**, 103 (1956).
- [8] S. Novaes (1999), *Standard model: An Introduction*, in *10th Jorge Andre Swieca Summer School: Particle and Fields*, pp. 5–102, arXiv:hep-ph/0001283.
- [9] D. Griffiths (2008), *Introduction to elementary particles*.
- [10] M. D. Schwartz (2014), *Quantum Field Theory and the Standard Model* (Cambridge University Press).
- [11] E. S. Abers and B. W. Lee, *Gauge theories*, Phys. Rep. **9** (1), 1 (1973).
- [12] T. Cheng and L. Li (1984), *Gauge theory of elementary particle physics*.
- [13] M. Baak, M. Goebel, et al., *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, Eur. Phys. J. C **72**, 2205, arXiv:1209.2716 [hep-ph] (2012).
- [14] P. F. de Salas, S. Gariazzo, et al., *Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects*, Front. Astron. Space Sci. **5**, 36, arXiv:1806.11051 [hep-ph] (2018).
- [15] E. Majorana and L. Maiani (2006), *A symmetric theory of electrons and positrons*, in *Ettore Majorana Scientific Papers: On occasion of the centenary of his birth*, edited by G. F. Bassani (Springer, Berlin, Heidelberg) pp. 201–233.

- [16] Y. Fukuda et al. (Super-Kamiokande), *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81**, 1562, arXiv:hep-ex/9807003 [hep-ex] (1998).
- [17] Q. R. Ahmad et al. (SNO), *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89**, 011301, arXiv:nucl-ex/0204008 [nucl-ex] (2002).
- [18] Q. R. Ahmad et al. (SNO), *Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87**, 071301, arXiv:nucl-ex/0106015 [nucl-ex] (2001).
- [19] A. Yu. Smirnov, *Solar neutrinos: Oscillations or No-oscillations?*, arXiv:1609.02386 [hep-ph] (2016).
- [20] I. Esteban, M. C. Gonzalez-Garcia, et al., *The fate of hints: updated global analysis of three-flavor neutrino oscillations*, arXiv:2007.14792 [hep-ph] (2020).
- [21] M. Tanabashi et al. (Particle Data Group), *Review of Particle Physics*, Phys. Rev. D **98** (3), 030001 (2018).
- [22] M. Gonzalez-Garcia and Y. Nir, *Neutrino Masses and Mixing: Evidence and Implications*, Rev. Mod. Phys. **75**, 345, arXiv:hep-ph/0202058 (2003).
- [23] I. Esteban, M. C. Gonzalez-Garcia, et al., *Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ_{23} , δ_{CP} , and the mass ordering*, J. High Energy Phys. **01**, 106, arXiv:1811.05487 [hep-ph] (2019).
- [24] I. Esteban, M. C. Gonzalez-Garcia, et al. (2019), *NuFIT 4.1*, accessed: 2020-08-24.
- [25] C. Kraus et al., *Final results from phase II of the Mainz neutrino mass search in tritium beta decay*, Eur. Phys. J. C **40**, 447, arXiv:hep-ex/0412056 (2005).
- [26] V. N. Aseev, A. I. Belesev, et al., *Upper limit on the electron antineutrino mass from the Troitsk experiment*, Phys. Rev. C **84**, 112003 (2011).
- [27] M. Aker et al. (KATRIN), *Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN*, Phys. Rev. Lett. **123** (22), 221802, arXiv:1909.06048 [hep-ex] (2019).
- [28] N. Aghanim et al. (Planck), *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209 [astro-ph.CO] (2018).
- [29] E. Di Valentino, A. Melchiorri, and J. Silk, *Cosmic Discordance: Planck and luminosity distance data exclude Λ CDM*, arXiv:2003.04935 [astro-ph.CO] (2020).
- [30] A. de Gouvêa (2004), *TASI lectures on neutrino physics*, in *Theoretical Advanced*

BIBLIOGRAPHY

- Study Institute in Elementary Particle Physics: Physics in $D \geq 4$* , pp. 197–258, arXiv:hep-ph/0411274.
- [31] A. de Gouvêa, *Neutrino Mass Models*, Ann. Rev. Nucl. Part. Sci. **66**, 197 (2016).
- [32] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, Phys. Rev. Lett. **43**, 1566 (1979).
- [33] M. Gell-Mann, P. Ramond, and R. Slansky, *Complex Spinors and Unified Theories*, Conf. Proc. C **790927**, 315, arXiv:1306.4669 [hep-th] (1979).
- [34] P. Minkowski, $\mu \rightarrow e\gamma$ at a Rate of One Out of 10^9 Muon Decays?, Phys. Lett. B **67**, 421 (1977).
- [35] T. Yanagida, *Horizontal gauge symmetry and masses of neutrinos*, Conf. Proc. C **7902131**, 95 (1979).
- [36] S. Glashow, *The Future of Elementary Particle Physics*, NATO Sci. Ser. B **61**, 687 (1980).
- [37] R. N. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Non-conservation*, Phys. Rev. Lett. **44**, 912 (1980).
- [38] J. Schechter and J. Valle, *Neutrino Masses in $SU(2) \times U(1)$ Theories*, Phys. Rev. C **22**, 2227 (1980).
- [39] A. Donini, P. Hernandez, et al., *Minimal models with light sterile neutrinos*, J. High Energy Phys. **07**, 105, arXiv:1106.0064 [hep-ph] (2011).
- [40] G. Drexlin, V. Hannen, et al., *Current direct neutrino mass experiments*, Adv. High Energy Phys. **2013**, 293986, arXiv:1307.0101 [physics.ins-det] (2013).
- [41] J. Lesgourgues and S. Pastor, *Massive neutrinos and cosmology*, Phys. Rept. **429**, 307, arXiv:astro-ph/0603494 (2006).
- [42] A. G. Riess, *The Expansion of the Universe is Faster than Expected*, Nature Rev. Phys. **2** (1), 10, arXiv:2001.03624 [astro-ph.CO] (2019).
- [43] M. Goeppert-Mayer, *Double beta-disintegration*, Phys. Rev. **48**, 512 (1935).
- [44] M. G. Inghram and J. H. Reynolds, *Double Beta-Decay of Te^{130}* , Phys. Rev. **78**, 822 (1950).
- [45] S. R. Elliott, A. A. Hahn, and M. K. Moe, *Direct Evidence for Two Neutrino Double Beta Decay in ^{82}Se* , Phys. Rev. Lett. **59**, 2020 (1987).

- [46] M. K. Moe, *The First Direct Observation of Double-Beta Decay*, *Ann. Rev. Nucl. Part. Sci.* **64**, 247 (2014).
- [47] M. Agostini, A. M. Bakalyarov, et al. (GERDA Collaboration), *Improved Limit on Neutrinoless Double- β Decay of ^{76}Ge from GERDA Phase II*, *Phys. Rev. Lett.* **120**, 132503 (2018).
- [48] C. E. Aalseth, N. Abgrall, et al. (Majorana Collaboration), *Search for Neutrinoless Double- β Decay in ^{76}Ge with the Majorana Demonstrator*, *Phys. Rev. Lett.* **120**, 132502 (2018).
- [49] C. Alduino, F. Alessandria, et al. (CUORE Collaboration), *First Results from CUORE: A Search for Lepton Number Violation via $0\nu\beta\beta$ Decay of ^{130}Te* , *Phys. Rev. Lett.* **120**, 132501 (2018).
- [50] J. B. Albert, G. Anton, et al. (EXO-200 Collaboration), *Search for Neutrinoless Double-Beta Decay with the Upgraded EXO-200 Detector*, *Phys. Rev. Lett.* **120**, 072701 (2018).
- [51] A. Gando, Y. Gando, et al. (KamLAND-Zen Collaboration), *Search for Majorana Neutrinos Near the Inverted Mass Hierarchy Region with KamLAND-Zen*, *Phys. Rev. Lett.* **117**, 082503 (2016).
- [52] M. J. Dolinski, A. W. Poon, and W. Rodejohann, *Neutrinoless Double-Beta Decay: Status and Prospects*, *Ann. Rev. Nucl. Part. Sci.* **69**, 219, arXiv:1902.04097 [nucl-ex] (2019).
- [53] M. K. Moe and P. Vogel, *Double beta decay*, *Ann. Rev. Nucl. Part. Sci.* **44**, 247 (1994).
- [54] S. Dell’Oro, S. Marcocci, et al., *Neutrinoless double beta decay: 2015 review*, *Adv. High Energy Phys.* **2016**, 2162659, arXiv:1601.07512 [hep-ph] (2016).
- [55] K. S. Hirata et al., *Observation in the Kamiokande-II Detector of the Neutrino Burst from Supernova SN 1987a*, *Phys. Rev. D* **38**, 448 (1988).
- [56] C. B. Bratton et al. (IMB), *Angular Distribution of Events From Sn1987a*, *Phys. Rev. D* **37**, 3361 (1988).
- [57] E. Alekseev, L. Alekseeva, et al., *Possible Detection of a Neutrino Signal on 23 February 1987 at the Baksan Underground Scintillation Telescope of the Institute of Nuclear Research*, *JETP Lett.* **45**, 589 (1987).
- [58] W. D. Arnett and J. L. Rosner., *Neutrino Mass Limits From Sn1987a*, *Phys. Rev. Lett.* **58**, 1906 (1987).

BIBLIOGRAPHY

- [59] T. J. Loredo and D. Q. Lamb, *Bayesian analysis of neutrinos observed from supernova SN-1987A*, Phys. Rev. D **65**, 063002, arXiv:astro-ph/0107260 (2002).
- [60] G. Pagliaroli, F. Rossi-Torres, and F. Vissani, *Neutrino mass bound in the standard scenario for supernova electronic antineutrino emission*, Astropart. Phys. **33**, 287, arXiv:1002.3349 [hep-ph] (2010).
- [61] P. Antonioli, R. T. Fienberg, et al., *SNEWS: the SuperNova Early Warning System*, New J. Phys. **6**, 114 (2004).
- [62] J. F. Beacom, R. N. Boyd, and A. Mezzacappa, *Technique for direct eV scale measurements of the mu and tau neutrino masses using supernova neutrinos*, Phys. Rev. Lett. **85**, 3568, arXiv:hep-ph/0006015 [hep-ph] (2000).
- [63] P. T. Springer, C. L. Bennett, and P. A. Baisden, *Measurement of the Neutrino Mass Using the Inner Bremsstrahlung Emitted in the Electron-Capture Decay of ^{163}Ho* , Phys. Rev. A **35**, 679 (1987).
- [64] S. Yasumi et al., *The Mass of the electron-neutrino from electron capture in Ho-163*, Phys. Lett. B **334**, 229 (1994).
- [65] S. C. Curran, J. Angus, and A. L. Cockroft, *Investigation of soft radiations—II. The beta spectrum of tritium*, Phil. Mag. **40** (300), 53, <https://doi.org/10.1080/14786444908561210> (1949).
- [66] M. Kleesiek et al., *β -Decay Spectrum, Response Function and Statistical Model for Neutrino Mass Measurements with the KATRIN Experiment*, Eur. Phys. J. C **79** (3), 204, arXiv:1806.00369 [physics.data-an] (2019).
- [67] E. W. Otten and C. Weinheimer, *Neutrino mass limit from tritium beta decay*, Rept. Prog. Phys. **71**, 086201, arXiv:0909.2104 [hep-ex] (2008).
- [68] K. A. Olive, *Review of Particle Physics*, Chin. Phys. C **38** (9), 090001 (2014).
- [69] M. Beck et al., *Improved determination of the β - $\bar{\nu}_e$ angular correlation coefficient a in free neutron decay with the aSPECT spectrometer*, Phys. Rev. C **101** (5), 055506, arXiv:1908.04785 [nucl-ex] (2020).
- [70] Y. A. Akulov and B. A. Mamyrin, *Half-life and $fT_{1/2}$ value for the bare triton*, Phys. Lett. B **610**, 45 (2005).
- [71] L. I. Bodine, D. S. Parno, and R. G. H. Robertson, *Assessment of molecular effects on neutrino mass measurements from tritium β decay*, Phys. Rev. C **91**, 035505 (2015).

- [72] E. G. Myers, A. Wagner, et al., *Atomic Masses of Tritium and Helium-3*, Phys. Rev. Lett. **114**, 013003 (2015).
- [73] S. Jonsell, A. Saenz, and P. Froelich, *Neutrino-mass determination from tritium β decay: Corrections to and prospects of experimental verification of the final-state spectrum*, Phys. Rev. C **60**, 034601 (1999).
- [74] A. H. Snell, F. Pleasonton, and H. Leming, *Molecular dissociation following radioactive decay: Tritium hydride*, J. Inorg. Nucl. Chem. **5** (2), 112 (1957).
- [75] S. Wexler, *Dissociation of TH and T2 by β -decay*, J. Inorg. Nucl. Chem. **10** (1), 8 (1959).
- [76] S. Wexler and F. T. Porter, *Probability of Formation of $[(^3\text{He})^+]_{As}$ following Beta Decay of Tritium*, J. Chem. Phys. **50** (12), 5428 (1969).
- [77] Y. T. Lin, T. H. Burritt, et al. (TRIMS Collaboration), *Beta Decay of Molecular Tritium*, Phys. Rev. Lett. **124**, 222502 (2020).
- [78] N. Doss, J. Tennyson, et al., *Molecular effects in investigations of tritium molecule β decay endpoint experiments*, Phys. Rev. C **73**, 025502 (2006).
- [79] N. Doss and J. Tennyson, *Excitations to the electronic continuum of $^3\text{HeT}^+$ in investigations of T2 β -decay experiments*, J. Phys. B **41** (12), 125701 (2008).
- [80] L. W. Alvarez and R. Cornog, *Helium and Hydrogen of Mass 3*, Phys. Rev. **56**, 613 (1939).
- [81] V. M. Lobashev and P. E. Spivak, *A method for measuring the electron antineutrino rest mass*, Nucl. Inst. Meth. A **240** (2), 305 (1985).
- [82] A. Picard, H. Backe, et al., *A solenoid retarding spectrometer with high resolution and transmission for keV electrons*, Nucl. Inst. Meth. B **63** (3), 345 (1992).
- [83] B. Monreal and J. A. Formaggio, *Relativistic Cyclotron Radiation Detection of Tritium Decay Electrons as a New Technique for Measuring the Neutrino Mass*, Phys. Rev. D **80**, 051301, arXiv:0904.2860 [nucl-ex] (2009).
- [84] D. M. Asner et al. (Project 8), *Single electron detection and spectroscopy via relativistic cyclotron radiation*, Phys. Rev. Lett. **114** (16), 162501, arXiv:1408.5362 [physics.ins-det] (2015).
- [85] P. J. Doe et al. (Project 8) (2013), *Project 8: Determining neutrino mass from tritium beta decay using a frequency-based method*, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi*

BIBLIOGRAPHY

- (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013, arXiv:1309.7093 [nucl-ex].
- [86] T. E. Weiss (2018), *Modeling Beta Decay Spectra to Analyze the Sensitivity of a Neutrino Mass Experiment*, Bachelor Thesis (MIT).
- [87] R. G. H. Robertson and D. A. Knapp, *Direct Measurements of Neutrino Mass*, Ann. Rev. Nucl. Part. Sci. **38** (1), 185 (1988).
- [88] V. N. Aseev et al., *Energy loss of 18 keV electrons in gaseous T and quench condensed D films*, Eur. Phys. J. D **10** (1), 39 (2000).
- [89] M. B. Shah, D. S. Elliott, and H. B. Gilbody, *Pulsed crossed-beam study of the ionisation of atomic hydrogen by electron impact*, J. Phys. B **20** (14), 3501 (1987).
- [90] L. I. Bodine, D. S. Parno, and R. G. H. Robertson, *Assessment of molecular effects on neutrino mass measurements from tritium β decay*, Phys. Rev. C **91**, 035505 (2015).
- [91] R. D. Williams and S. E. Koonin, *Atomic final-state interactions in tritium decay*, Phys. Rev. C **27**, 1815 (1983).
- [92] A. Ashtari Esfahani et al. (Project 8), *Determining the neutrino mass with cyclotron radiation emission spectroscopy—Project 8*, J. Phys. **G44** (5), 054004, arXiv:1703.02037 [physics.ins-det] (2017).
- [93] A. Ashtari Esfahani et al., *Electron Radiated Power in Cyclotron Radiation Emission Spectroscopy Experiments*, Phys. Rev. C **99** (5), 055501, arXiv:1901.02844 [physics.ins-det] (2019).
- [94] S. Jonsell, A. Saenz, and P. Froelich, *Neutrino-mass determination from tritium β decay: Corrections to and prospects of experimental verification of the final-state spectrum*, Phys. Rev. C **60**, 034601 (1999).
- [95] A. Saenz, S. Jonsell, and P. Froelich, *Improved Molecular Final-State Distribution of HeT⁺ for the β -Decay Process of T₂*, Phys. Rev. Lett. **84**, 242 (2000).
- [96] A. Stark and A. Saenz, *Final-State Spectrum of ³He after β^- Decay of Tritium Anions T⁻*, Phys. Rev. A **81**, 032501 (2010).
- [97] A. Lindman (2018), *Concept for an Atomic Tritium Experiment: Phase IV of Project 8*, accessed: 2020-04-7.
- [98] C. Amole, G. B. Andresen, et al., *The ALPHA antihydrogen trapping apparatus*, Nucl. Inst. Meth. A **735**, 319 (2014).

- [99] T. G. Thomas and S. C. Sekhar (2005), *Communication Theory* (Tata McGraw Hill).
- [100] A. Ashtari Esfahani et al., *Cyclotron Radiation Emission Spectroscopy Signal Classification with Machine Learning in Project 8*, New J. Phys. **22** (3), 033004, arXiv:1909.08115 [nucl-ex] (2020).
- [101] A. West, *Lepton Dipole Moments*, arXiv:1607.00925 [hep-ph] (2016).
- [102] R. G. H. Robertson and V. Venkatapathy, *Shakeup and shakeoff satellite structure in the electron spectrum of $^{83}\text{Kr}^m$* , arXiv:2003.12952 [nucl-ex] (2020).
- [103] Raspberry Pi Foundation (2012), <https://www.raspberrypi.org/documentation>, accessed: 2020-09-02.
- [104] CASPER (2012), *ROACH-2 Revision 2*, accessed: 2019-08-28.
- [105] CASPER (2012), *ADC1x5000-8*, accessed: 2019-08-27.
- [106] A. Young (2016), *ROACH2 Digital Acquisition System for Project8 Phase II*, Github package, accessed: 2019-08-27.
- [107] H. Jiang, H. Liu, et al. (2012), *Digitizing The Yuan Tseh Lee Array for Microwave Background Anisotropy by 5Gbps ADC boards*, in *2012 19th IEEE International Conference on Electronics, Circuits, and Systems (ICECS 2012)*, pp. 304–307.
- [108] N. A. Patel, R. W. Wilson, et al., *Characterizing the performance of a high-speed ADC for the SMA digital backend*, J. Astron. Instrum. **03** (01), 1450001, <https://doi.org/10.1142/S2251171714500019> (2014).
- [109] CASPER (2009), <https://casper.ssl.berkeley.edu/wiki/BORPH>, accessed: 2020-09-02.
- [110] CASPER (2013), <https://casper.ssl.berkeley.edu/wiki/Fft>, accessed: 2020-09-02.
- [111] S. Butterworth, *On the Theory of Filter Amplifiers*, Experimental Wireless and the Wireless Engineer **7**, 536 (1930).
- [112] C. Claessens, B. H. LaRoque, and N. S. Oblath (2016), *Psyllid*, Github package, accessed: 2019-09-16.
- [113] N. S. Oblath, J. Kofron, et al. (2018), *Monarch*, Github package, accessed: 2020-5-13.
- [114] The HDF Group (2000-2010), *Hierarchical data format version 5*, accessed: 2019-10-4.
- [115] B. H. LaRoque (2017), *Cyclotron Radiation Emission Spectroscopy: first demon-*

BIBLIOGRAPHY

- stration and performance benchmarks from the Project 8 experiment*, PhD Thesis (UCSB).
- [116] D. L. Furse (2013), *Midge*, Github package, accessed: 2019-09-16.
- [117] B. H. LaRoque (2014), *Dripline*, Github package, accessed: 2019-09-16.
- [118] Pivotal Software (2013), <https://www.rabbitmq.com>, accessed: 2019-09-25.
- [119] ISO/IEC 19464 (2014), *Information technology - advanced message queuing protocol (amqp) v1.0 specification (2014)*.
- [120] The Linux Kernel Archives (2009), https://www.kernel.org/doc/Documentation/networking/packet_mmap.txt, accessed: 2020-02-21.
- [121] N. S. Oblath, C. Claessens, et al. (2019), *Katydid*, Github package, accessed: 2019-10-4.
- [122] M. Frigo and S. G. Johnson, *The Design and Implementation of FFTW3*, Proceedings of the IEEE **93** (2), 216, special issue on “Program Generation, Optimization, and Platform Adaptation” (2005).
- [123] M. Gödel (2019), *Track reconstruction in Project 8*, Master Thesis (JGU Mainz).
- [124] M. Ester, H.-P. Kriegel, et al. (1996), *A density-based algorithm for discovering clusters in large spatial databases with noise* (AAAI Press) pp. 226–231.
- [125] D. L. Furse (2015), *Techniques for Direct Neutrino Mass Measurement Utilizing Tritium β -Decay*, PhD Thesis (MIT).
- [126] E. M. Zayas (2019), *New Analysis Developements for Reconstruction and Classification of CRES Signals in Project 8*, Master Thesis (MIT).
- [127] A. Ashtari Esfahani et al. (Project 8), *Locust: C++ software for simulation of RF detection*, New J. Phys. **21**, 113051, arXiv:1907.11124 [physics.comp-ph] (2019).
- [128] N. Buzinsky, P. L. Slocum, et al. (2013), *Locust*, Github package, accessed: 2020-03-10.
- [129] M. E. Rudd, *Differential and total cross sections for ionization of helium and hydrogen by electrons*, Phys. Rev. C **44**, 1644 (1991).
- [130] UBC database (1994), *Hydrogen high resolution (48meV fwhm) 1994 data set (T.N. Olney et al.)*, retrieved: 2019-06-27.
- [131] D. L. Furse et al., *Kassiopeia: A Modern, Extensible C++ Particle Tracking Package*, New J. Phys. **19** (5), 053012, arXiv:1612.00262 [physics.comp-ph] (2017).

- [132] N. Mohd Razali and B. Yap, *Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests*, J. Stat. Model. Analytics **2** (2011).
- [133] S. S. Shapiro and M. B. Wilk, *An analysis of variance test for normality (complete samples)*, Biometrika **52** (3-4), 591 (1965).
- [134] E. B. Wilson, *Probable Inference, the Law of Succession, and Statistical Inference*, J. Am. Stat. Assoc. **22** (158), 209 (1927).
- [135] S. Wallis, *Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods*, J. Quant. Linguist. **20** (3), 178 (2013).
- [136] J. Bourbeau and Z. Hampel-Arias, *PyUnfold: A Python Package for Iterative Unfolding*, 10.5281/zenodo.1258211, arXiv:1806.03350 [physics.data-an] (2018).
- [137] iminuit team (2018), *iminuit – A Python interface to Minuit*, Github package, accessed: 2020-04-13.
- [138] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10**, 343 (1975).
- [139] G. Cowan, K. Cranmer, et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71**, 1554, arXiv:1007.1727 [physics.data-an] (2011).
- [140] N. S. Oblath, *Project 8*, June, 2020.
- [141] A. Lokhov and F. V. Tkachov, *Confidence intervals with a priori parameter bounds*, Phys. Part. Nucl. **46** (3), 347, arXiv:1403.5429 [physics.data-an] (2015).