DIPLOMARBEIT

Convention versus Convolution: Two Methods of Low-Energy Event Classification in CRESST

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Masterstudium Technische Physik

eingereicht von

Damir Rizvanović, BSc

Matrikelnummer: 01525388

ausgeführt am Atominstitut
der Fakultät für Physik der Technischen Universität Wien
in Zusammenarbeit mit dem Institut für Hochenergiephysik
der Österreichischen Akademie der Wissenschaften

Betreuer: Univ.-Prof. Dr. Jochen Schieck
Mitwirkung: Dr. Florian Reindl, DI Felix Wagner

Unterschrift Verfasser:

Unterschrift Betreuer:

Wien, Mai 2022

Abstract

The introduction of new low-threshold detectors in CRESST-III has brought a new phenomenon in the low-energy region to surface. The origin of the Low-Energy Excess has to date not been identified, with new detectors being crafted in hopes of determining the source of the excess. This work will aim to provide a raw data analysis using conventional CRESST techniques, so-called quality cuts, of event discrimination on one hand, but also build upon the foundation laid in previous works and utilise Machine Learning as a tool for event selection. By employing Convolutional Neural Networks, one of today's most popular types of Machine Learning algorithms, this work's analysis will not only be able to reproduce the results achieved with conventional methods, but also allow for a more granular discrimination of event types for low energies.

Abstract

Durch die Einführung neuer extrem sensibler Detektoren in CRESST-III kommt es zur Messung eines neuen Phänomens für das CRESST-Experiment, dem sogenannten Low-Energy Excess. Die Einführung wiederum neuer Detektorentypen in den letzten Runs versucht den Ursprung dieses Überschusses an gemessenen Teilchen bei niedrigen Energien zu ermitteln. Das Ziel dieser Arbeit ist es einerseits eine Rohdatenanalyse mittels konventioneller CRESST-Analysemethoden, den Quality Cuts, zur Pulsunterscheidung bereitzustellen, andererseits aber auch Machine Learning, aufbauend auf der Analyse vorhergehender Arbeiten, als Werkzeug zu nutzen um Events voneinander zu unterscheiden. Durch die Nutzung von Convolutional Neural Networks, einem der heutzutage weitverbreitetsten Machine Learning Algorithmen, ist es in dieser Arbeit nicht nur möglich die Resultate altbekannter Methoden zu reproduzieren, sondern auch niederenergetische Events feiner und granularer zu klassifizieren.

Contents

Contents

Introduction

Dark matter (DM) and its inherent nature has eluded physicists for almost 90 years now, with earliest evidence found by Fritz Zwicky in 1933 [1]. The main issue is already rooted in its name: even though dark matter is about five times more abundant than baryonic matter in the universe, its incapability of interacting electromagnetically leads to it being invisible to any type of optical measurement device.

In order to measure dark matter one will therefore need to resort to different types of measurement methods, with Cryogenic Rare Event Search with Superconducting Thermometers (CRESST) being such an experiment. Located within the Abruzzi mountains, CRESST operates at extremely cold temperatures, with the aim of measuring the recoil signal of dark matter particle interactions with precisely grown target crystals. This setup allows for high precision measurements in its most recent iteration, CRESST-III. Results presented in [2] showcase not only measurements in the sub-Gev/c^2 dark matter mass regime, but also one of the lowest exclusion limits for light dark matter particles measured today.

Analysis of CRESST measurements necessitates, in a first step, the discrimination of event types into valid particle interactions in the target crystal and different artifacts. The rate and energy information for DM candidate events, i.e. nuclear recoils, is then used to calculate limits on the DM-nucleon interaction cross section. Until now the standard procedure within the CRESST collaboration has for the most part been to apply a large variety of parametric cuts in order to achieve this goal. In this work, we provide a possible way of presenting the process of event discrimination as a typical classification problem utilising Machine Learning (ML). We build upon the progress and knowledge already presented in [4–8]. Here we choose a slightly different path from the above mentioned works and attempt to frame the classification of events as a semi-supervised problem employing Convolutional Neural Networks (CNN) as our classifiers. That is to say that the training of the Neural Network is based of simulated events exclusively.

The work is structured as follows. The first chapter provides a short introduction into the history of dark matter and its observational evidence. From this we also discuss different candidates for dark matter and explore options for detection of one of the most popular candidates, Weakly Interacting Massive Particles (WIMPs). In chapter 2 the history of the CRESST experiment and in particular its experimental setup, working principle and detector modules are discussed in great detail. There

we specifically focus on the setup, schematics and behaviour of a novel beaker module, called "Gode3", as the data from this module is the basis of the analysis in the later chapters. Chapter 1 and 2 will closely follow the explanations given in the author's project work [7].

The following chapter 3 is divided into two, distinct parts. In the first part of the raw data analysis, the exploration of the hardware triggered data is showcased with a focus lying on the generation and extraction of the Gode3-specific characteristic pulse shapes (standard events) and the determination of the trigger thresholds. This section will also provide an introduction to concepts like the pulse shape model [9], standard events themselves, as well as the optimum filter [10]. The second part of the chapter then utilises the triggered stream data set for the main part of the dark matter analysis. In this section, quality cuts are applied to exclude artifacts from the data set and additionally distinguish between the two different types of events present within the phonon channel data. Following a general discussion concerning the conversion of amplitudes to recoil energies and the calculation thereof, we provide energy spectra and light yield plots showcasing the results of this chapter's analysis. Especially the results in the region below 1keV recoil energy will constitute a focal point of this and the final chapter's analysis.

The final chapter explores at last the application of CNNs as classifiers in a semi-supervised manner. In this chapter the theoretical background of CNNs is briefly outlined, followed by the description of the data preparation and the simulation of pulses which are correlated between the two phonon channels. The main part of the chapter then explores the training of the CNN and subsequent attempts at improving our results. Here we introduce the concepts of prediction probabilities and establish a model which incorporates predictions of both the Gode3's carrier and ring channel. The chapter concludes with a discussion of the model's performance in the low-energy region.

Chapter 1

The Search for Dark Matter

As the human species evolves, we become more aware of our own insignificance when it comes to scales larger than our solar system. Not only are we living on a metaphorical speck of dust in relation to the rest of the interstellar objects, the matter that constitutes us also plays only a subdominant role in the total energy density of the universe.

According to the Lambda cold dark matter (Λ-CDM) model [11], often termed as the standard model of Big Bang cosmology, the largest contribution to the total energy density of our universe of 69.2%, as measured by the Planck mission [14], belongs to a constant energy of empty space which accelerates the expansion of the universe, the so-called *dark energy*. On the other end of the spectrum *radiation* today only contributes 0.01% to the energy density of the universe after being dominant in its early stages. The remaining energy density is divided into 4.8% for *baryonic matter*, i.e. the type of matter that constitutes all matter described within the standard model (SM) of particle physics and for the biggest part large clouds of helium, and 25.8% *dark* matter [15]. Dark matter is a type of matter that interacts with baryonic matter via the gravitational force and is postulated to possibly interact weakly - either by the weak interaction or some other, to date unknown, type of interaction - as discussed in Section 1.2. Hypothetical dark matter constituents are not included in the SM of particle physics, and all SM particles have all been excluded as possible DM candidates for one reason or another [11].

This chapter will deal with this ominous type of matter and what it might be made out of. First, in Section 1.1, one needs to establish the evidence for this peculiar type of matter. Then Section 1.2 deals with models, primarily in form of particle candidates, attempting to describe dark matter. Lastly, in Section 1.3, the various methods of dark matter detection are mentioned.

1.1 Evidence for Dark Matter

First evidence for the existence of non-visible matter was found in 1932 by Dutch physicist Jan Oort who found proof for *extra hidden matter in our galaxy* [11]. The following year, Fritz Zwicky was able to infer - by usage of the Virial theorem (1.1) - that the mean mass density of the Coma galaxy cluster was much larger than the visible matter density would indicate [1, 12]:

$$\bar{T} = -\frac{1}{2}\bar{U} \quad .$$

(1.1)

The Virial theorem relates the kinetic and potential energy of a closed multi-body system and therefore allows through the measurement of the radial velocity a calculation of the total mass of the system. Zwicky notes that for average redshifts observed in the cluster, velocities of 1000 km/s or more are necessary which implicates an average density 400 times larger than suggested by the luminous matter in the system [12]. Modern estimates lead to a different factor of about 100 times more mass density than is present in luminous matter [13]. Zwicky's calculations was by far not the only piece of evidence leading towards the postulation of dark matter, as presented in this section, but was only the starting point of many landmark observations in the history of DM physics. An immediate consequence of the much higher velocities within galaxy clusters can be seen in the following example.

1.1.1 Rotation Curves

The observation of average velocities within galaxy clusters also lead to the peculiar behaviour of galaxy rotation curves. Application of simple classical mechanics and equating the centripetal and gravitational force acting on a body of mass m a distance r away from the center of the galaxy gives the following expression for its velocity $v(r)$

$$v(r) = \sqrt{\frac{GM(r)}{r}} \quad ,$$

(1.2)

where the mass M that attracts the object is given by,

$$M(r) = 4\pi \int_0^r \rho(r')r'^2 dr' \quad .$$

(1.3)

The object will feel the total mass within its orbit. Once beyond the borders of the galaxy the total mass of the cluster will interact with the object as its distance increases and therefore the velocity at the edge of the galaxy should scale as $v \propto \sqrt{1/r}$. This behaviour is not what one actually observes when looking at galaxy rotation curves as can be seen in Fig. 1.1. Contrary to our calculations the velocity at the edge of the visible galaxy is still increasing which can be explained by a dark matter halo surrounding the galaxy with a much larger radius than the visible matter would indicate.
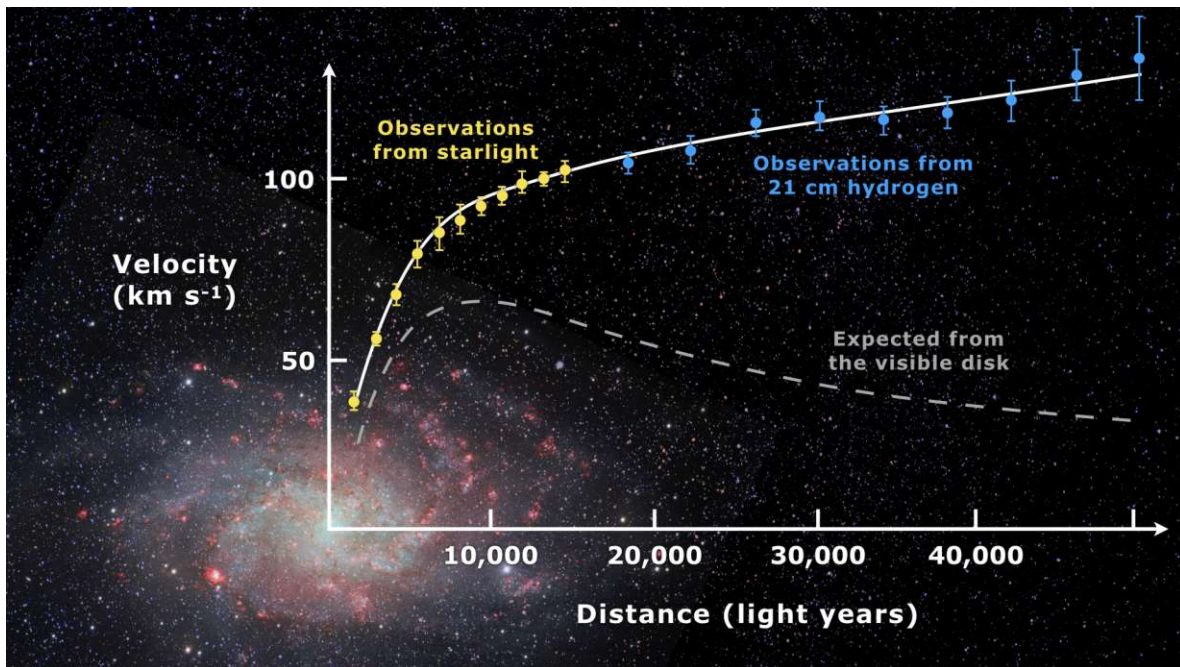
Figure 1.1: Rotation curve of the Messier 33 spiral galaxy (M33). The yellow and blue data points show the actual velocity distribution of the galaxy compared to the dashed line which shows the expected velocity for no dark matter for a distance R from the center of the galaxy (x-axis). Figure taken from [16].

1.1.2 Cosmic Microwave Background

Further evidence, first provided by the COBE mission operated from 1989 to 1993, for the existence of dark matter lies within the fact that the Cosmic Microwave Background (CMB) displays anisotropies of the order of $\mathcal{O}(10^{-5})$ in its black body temperature spectrum. Newer satellite missions such as WMAP and Planck were able to further improve the precision of the COBE measurements as shown in Fig. 1.2 and are able to estimate the dark and baryonic matter contents as well as infer the dark energy density and the universe's age [14].

The CMB was first observed by Arno Penzias and Robert Wilson in 1964 as a faint signal in the microwave region which originated from everywhere in the sky. This result is exactly what was expected of a universe that started in a very hot, dense state and whose fast expansion gave rise to photons which - now red-shifted - were being detected on earth. What we now call CMB was formed in the early stages of the universe when its temperature, through the process of expansion, dipped below 13.6 eV and produced photons that were now unable to produce proton-electron pairs by breaking the hydrogen bonds and propagated freely through a now transparent universe [15]. This process, called decoupling, gave us an incredibly homogeneous black body spectrum with a temperature of $T = (2.72548 \pm 0.00057)\mathrm{K}$ [15]. But what needs
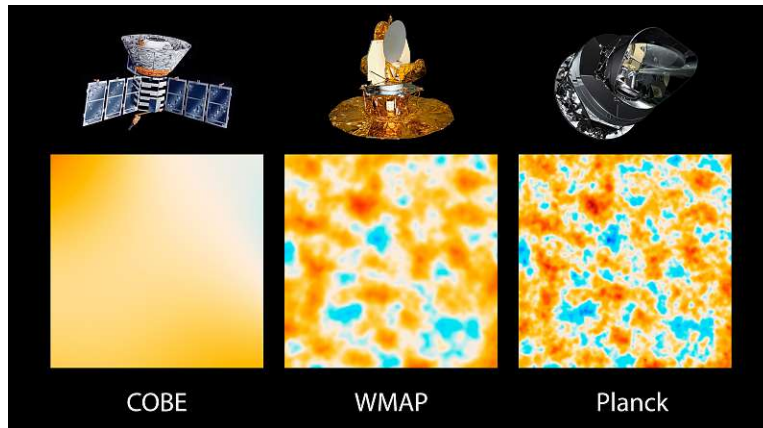
Figure 1.2: 10-square-degree all-sky patches of the CMB measured by different satellite missions ordered by starting date from left to right. The oldest satellite mission COBE started in 1989 exhibits the anisotropies in the CMB much less clearly than the more modern successors WMAP (2001, middle) and Planck (2009, right). Figure taken from [17].

to be acknowledged is that the CMB in fact exhibits inhomogenities which cannot be explained only by the existence of baryonic matter. Before the decoupling, baryonic matter and radiation were in equilibrium steadily producing proton-electron pairs and hydrogen-photon pairs

$$p + e^- \leftrightharpoons H + \gamma \quad , \tag{1.4}$$

any fluctuation in the baryonic matter soup would be immediately suppressed and would render a perfectly homogeneous black body spectrum in the CMB. For the production of anisotropies we see in the CMB today there would have needed to exist some matter which did not interact with this primordial radiation and only interacted with baryonic matter through gravitation. The clumping of dark and thus baryonic matter can then explain the structure of the CMB we observe today. The necessity of interaction between baryonic and dark matter is also evidenced by the following example.

1.1.3 Bullet Cluster

In the early 2000's another piece of evidence for the existence of dark matter was found in the so called *Bullet Cluster* (1E 0657-56) a collision of two galaxy clusters with vastly different sizes [18]. This process is shown in Fig. 1.3 where one sees the overlay of the optical image of the visible spectrum in the background, a pink clowd depicting the X-ray spectrum and the through gravitational lensing measured mass distribution of the cluster in blue.
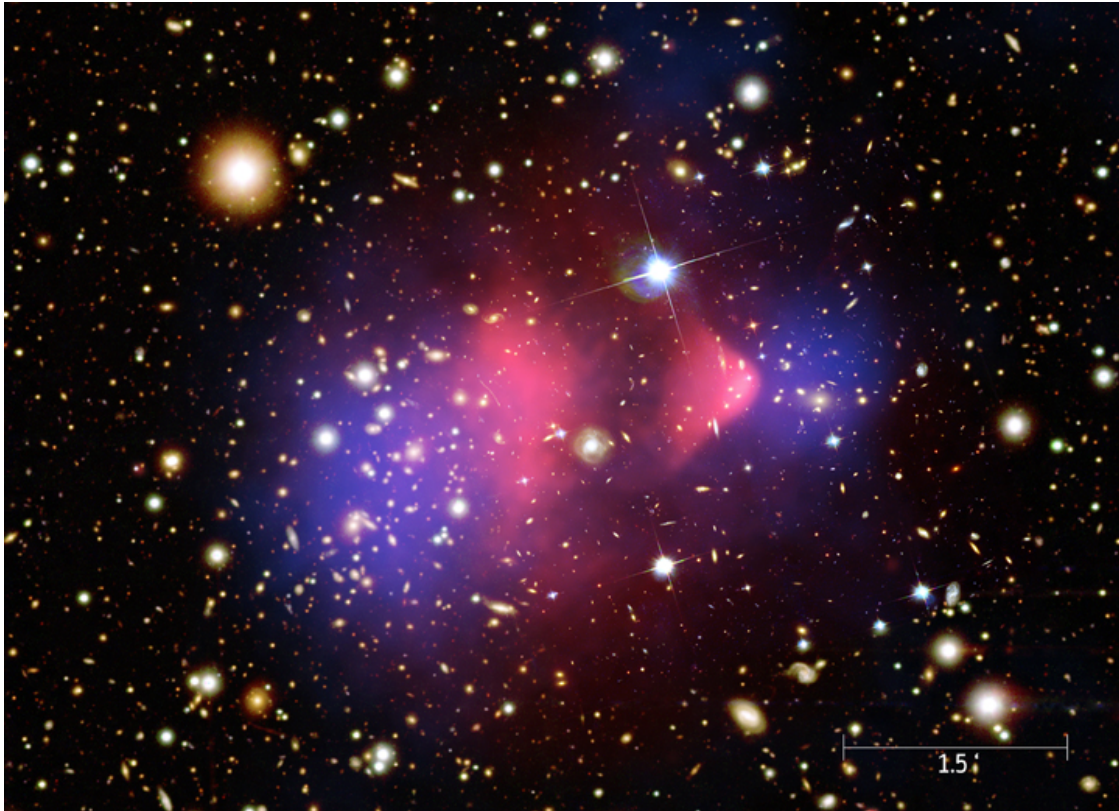
Figure 1.3: The Bullet Cluster or 1E 0657-56 shown as an overlay of the visible spectrum in the background, the X-ray spectrum in pink and the mass distribution in blue after the collision of two galaxy clusters. Figure taken from [19].

What becomes apparent in Fig. 1.3 is that the clouds corresponding to the mass distribution and the X-ray spectrum do barely overlap. This behaviour would not be expected for a collision where only baryonic matter was involved since the bulk of the galaxy - mainly galactic gas - would, through friction, slow down and the stars would simply pass through. But what we observe is that most of the mass distribution is located around the stars, who have much less mass than the gas present in the galaxy, which leads to the conclusion that next to the present baryonic matter there must also exist other matter, e.g. dark matter, which only weakly interacts with itself or not at all.

The bullet cluster plays an important role in disproving an alternate theory, the so-called Modified Newtonian Dynamics or MOND. Such theories impose flatness upon the galaxy rotation curves not by the introduction of dark matter, but by introducing a lower limit of acceleration a_0 below which Newton's second law is adapted to be $\vec{F} = m\vec{a}\frac{|\vec{a}|}{|\vec{a_0}|}$. The problem MOND theories face for the bullet cluster, other than violating momentum conservation, Lorentz invariance and the equivalence principle, is that they are incapable of describing the distinct mass and baryonic matter centers

observed in the Bullet Cluster and need 2 eV neutrinos to describe the non-baryonic matter density [20].

1.2   Dark Matter Candidates

From the previous section one is now able to infer conditions that dark matter needs to satisfy in an attempt of narrowing down the number of possible dark matter candidates. A first natural assumption, due its ability to interact gravitationally, is that dark matter consists of particles, in close analogy to the occurrence of baryonic matter in exclusively particle form. From here, let us recapitulate the above mentioned evidence that constrains dark matter to particles which, for one, only interact with themselves and baryonic matter via gravitational force and possibly via the weak interaction or some interaction of similar strength. Further, we must assume that dark matter was and still is cold i.e. not moving at relativistic speeds. Hot dark matter in the early universe would have produced structures far too small to be consistent with our observations of the CMB and in general with the structures of the universe today. Lastly, as mentioned before, dark matter must have existed in the earliest stages of the universe and therefore needs to be stable with lifetimes comparable to the age of the universe, or needs to exhibit a fitting reproduction mechanism.

With these constraints we are already able to exclude some possibilities within the standard model of particle physics such as neutrinos. Even though electrically neutral and therefore not interacting electromagnetically, neutrinos are far too *hot* to adhere to the structure of the CMB we measure today. Another option would be so-called *MACHOs* - MAssive Compact Halo Objects - such as black holes, neutron stars, white and brown dwarfs which have been observed but make up far too little of the total energy density as to be considered candidates for dark matter.

Outside the standard model, we encounter hypothetical particles like the *massive sterile neutrinos*, right-handed massive neutrinos which only interact by way of gravitation, and *axions*, a hypothetical particle postulated by Peccei and Quinn [21]. Latter ones would also explain the lack of CP-violation in the strong interaction and are able to decay into and be produced by photons in the presence of strong magnetic fields via the so-called Primakoff effect [22]. Both types of particles are possible to detect, the latter by applying this effect, the former by a mechanism called active-sterile neutrino mixing [23].

This leads us to the next candidate called *Weakly Interacting Massive Particle* (WIMP) which, as its name suggests, also interacts weakly with baryonic matter along with the requirements mentioned above, due to being massive and thus heavy and cold, and was favored by the physics community. WIMPs are postulated to have originated in the early universe and were then in thermal equilibrium with the surrounding plasma. But as the temperature dropped the production processes were *frozen out*, thus decreasing the WIMP density as the universe kept on expanding. Since this in turn prohibited the further annihilation of WIMPs with themselves, we obtain an ex-

planation for the abundance of relic dark matter in our universe today. With the relic dark matter density as a fixed parameter, we can calculate the dark matter cross section with the Boltzmann equation (1.5):

$$\frac{dn}{dt} + 3Hn = -\langle \sigma v \rangle_{\text{ann}}(n^2 - n_{eq}^2) \quad , \tag{1.5}$$

where n is the particle density, n_{eq} is the particle density in equilibrium, H is the Hubble constant, and the term in the angled brackets depicts the thermally averaged annihilation cross section of the WIMPs [24]. When assuming typical masses for the weak-scale of the order of $\mathcal{O}(10 \text{ GeV})$, as for these masses the cross section is proportional to m_{WIMP}^2, we find that the prevalence for such particles is inversely proportional to its interaction cross section. Thus, we are able to calculate the interaction cross section of the WIMPs at the time of the freeze-out from the dark matter abundance today. The resulting cross section lies in the order of the weak interaction, with this result being called the *WIMP miracle* for its surprisingly high agreement with known physics [24].

Promising candidates for what WIMPs might be, are found in the most popular extension of the standard model called *super symmetry* (SUSY). The SUSY model appoints a supersymmetric partner or superpartner to every particle in the standard model. Additionally, a new type of symmetry is introduced called R-Parity which states that the lightest supersymmetric partner (LSP) is guaranteed to be stable, a property ideal for a dark matter candidate. Similar candidates are also obtained by universal extra dimension (lightest Kaluza-Klein particle) [25] and little Higgs (lightest T-odd particle) [26] theories. A problem with SUSY theories at the moment is that no experimental evidence has been found to date and that some models are already excluded [33]. In conclusion, WIMPs provide viable candidates for dark matter as they can be detected directly and indirectly and offer possibilities for production at colliders, as will be discussed in the upcoming Section 1.3. There were already observations of suitable WIMP-like signals, most mysterious and promising the DAMA modulation signal [28]. However, almost all of these observations so far turned out to be explainable within known physics. Also the reliability of the DAMA signal is questioned by contradictory exclusion limits from other experiments, most recently from the ANAIS collaboration [27]. Under the most standard assumptions the DAMA/LIBRA result has been excluded by a multitude of different experiments, with ANAIS now also providing a model independent approach. Furthermore, the HEPHY and TU Wien are currently also part of the COSINUS collaboration which is attempting to cross-check the DAMA/LIBRA result in a model-independent way, applying a unique detector setup in which NaI crystals act as scintillating cryogenic calorimeters [29]. This thesis will still, as it is standard within the community, consider a heavy-mediator model, as is done in a WIMP scenario. Especially low-mass candidates in the sub-GeV region, as searched by the CRESST experiment, are considered as a viable dark matter particle due to their mentioned merits. We point out here that low-mass candidates (sub-Gev) as examined by CRESST are typically excluded by the so-called Lee-Weinberg

bound [30] limiting the lowest WIMP-masses to $\sim$2GeV, due to their behaviour in the freeze-out mechanism. The CRESST experiment is thus searching for particles constituting a type of WIMP, e.g. predicted by the asymmetric DM model [31], but not the classical WIMP introduced in literature.

1.3 Detecting Dark Matter

As we learned previously, dark matter particles need to interact with standard model particles via a force roughly about as strong as the weak force to be detected successfully. Fig. 1.4 shows the Feynman diagram of the 3 orthogonal detection channels [32]: direct and indirect detection and production at a collider.
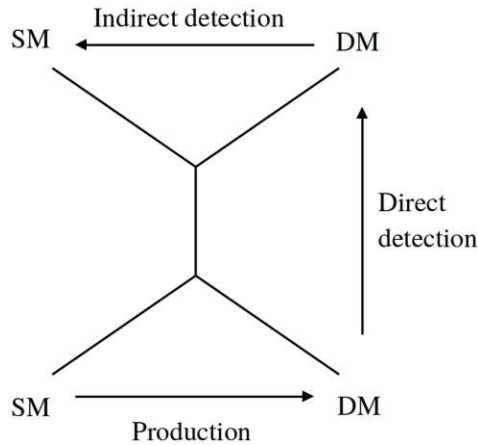

Figure 1.4: Feynman diagram of the 3 possible *orthogonal* detection routes for dark matter particles [32].

1.3.1 Direct Detection

Firstly, we cover the detection method of the CRESST experiment, the main subject of this work: direct detection. The experiment itself is discussed closely in chapter 2. This method focusses on the detection of dark matter particles which interact directly with standard model particles in target materials. These interactions take form as spin-independent coherent elastic scattering of particles. For Majorana particles and the LSP one also expects spin-dependent interactions [24] for which experiments have less sensitivity [32]. Since earth itself is located inside the dark matter halo of the Milky Way one expects to measure dark matter particles with detectors located on earth. Eq. (1.6) visualizes the expected detection rates for WIMPs in a simplified way [32],

$$R = \Phi_\chi \sigma_{\chi N} \frac{N_T}{m} \quad . \tag{1.6}$$

Here the rate of recoiling nuclei per target mass m is given by R where the number of nuclei in the target is given by N_T. Further, the flux of dark matter particles is given by $\Phi_\chi = n_\chi v_{\chi N}$ which depends on the velocity in relation to the target $v_{\chi N}$ and the density of particles in the dark matter halo n_χ. Especially the velocity of the particles can be inferred by measuring the velocity distribution of the galaxy. Lastly, the interaction cross section of WIMP-nucleus-scattering is given by $\sigma_{\chi N}$. For a closer discussion of anticipated dark matter recoil spectra, velocity distributions and form factor of the target material refer to [34]. As the work mentions one obtains low expected scattering rates and typical recoil energies of order $\mathcal{O}(\text{keV})$. The spectrum itself can be described as featureless and exponentially decaying. Therefore, highly sensitive, low threshold detectors with minimal background, large exposure times and sensitivity to low recoil energies are needed [32].
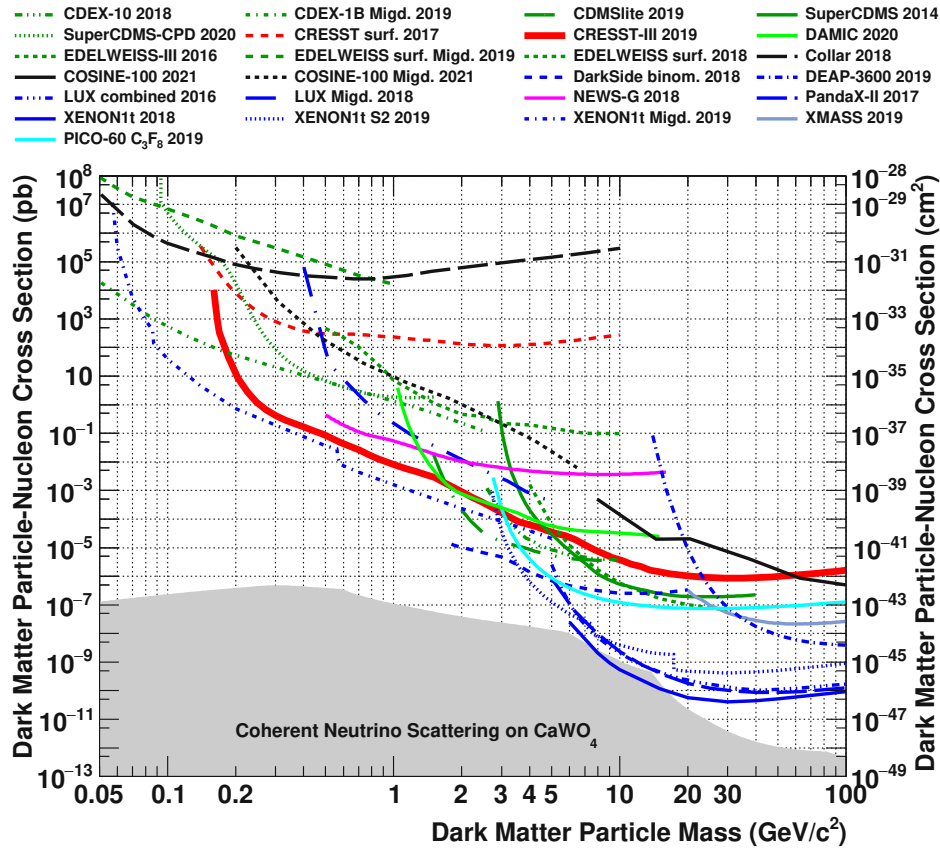

Figure 1.5: Exclusion limits for dark matter particles. Shown are limit lines for the CRESST experiment (red), liquid noble gas experiments (blue) and other cryogenic experiments (green) above which dark matter particles are excluded with 90% confidence. The limits are bounded from below by an irreducibe background from coherent neutrino-nucleus scattering [2].

To fulfill these requirements different experimental approaches can be taken. Generally subdivided into Phonon-, Scintillation-, Ionization- and Superheated Liquids detectors [32], here the different experiments are divided into the following categories. Firstly, we have liquid noble gas experiments e.g. XENON100 [35] and XENON1T [36] which currently hold the lowest exclusion limits. Experiments concerning low-mass dark matter particles achieve extremely low exclusion limits by accounting for the Migdal effect [36]. The exclusion limits are visualized via two dimensional plots i.e. Fig 1.5, where the dark matter mass is located on the x-axis, and the DM-nucleon cross section lies on the y-axis. Lines then correspond to the corresponding exclusion limits. Secondly, we have solid state experiments, where we want to single out cryogenic detectors such as the CRESST experiment which will be discussed in detail in the following chapter. The 2019 run by the CRESST-III experiment currently holds the second lowest exclusion limits behind the XENON1T experiment. On the spectrum of direct detection experiments we further find NaI-experiments such as the already mentioned DAMA/LIBRA- [28], COSINUS- [29] and ANAIS-experiments [27] which utilise scintillating NaI-crystals to detect annual modulation in their dark matter signal. This modulation is expected since the experiment's laboratory orbits around the sun and therefore passes different densities in the dark matter halo, thus measuring a modulating dark matter signal. We conclude by also mentioning so-called *Bubble Chambers* and charge-coupled devices (CCD) as possible direct detection detectors [34].

1.3.2 Indirect Detection

In contrast to direct methods we have indirect methods which aim to measure the standard model decay products of dark matter annihilation. This approach shows higher sensitivity towards WIMPs with high masses as opposed to the detection of dark matter produced at colliders. All standard model particles can be created in the annihilation, even γ-radiation by way of further decay into cosmic radiation. The three main types of cosmic radiation research focusses on are γ-radiation, charged leptons and neutrinos. Indirect searches mainly treat the GeV to TeV mass region as expected from the freeze-out model [32].

1.3.3 Production

As the name suggests, the hope with production at large particle colliders is to create low-mass dark matter particles - since the sensitivity for this kind of dark matter is much higher - from standard model particles, see [32]. The main obstacle here is that the produced particles are not detectable by the sensors as they usually measure the electromagnetic signal. In this scenario the experiment measures the lack of energy in the signal [34]. Furthermore, detectors at particle colliders such as the LHC (ATLAS [38], CMS [39]) or the SuperKEKB (Belle-II [40]) are sensitive towards the lifetime of the produced particle for only about 100 ns which is much shorter than the

postulated lifetime of WIMPs on the scale of the age of the universe. So far no trace of dark matter particles was found at the ATLAS and CMS experiments which operate at center of mass energies of 8 TeV for the former and 13 TeV for the latter [34]. This chapter provides only a small snippet of all possible DM detection experiments, with far more extensive introductions provided in Ref. [32] for instance. The following chapter will now completely focus on the one experiment whose data was used in the main part of this thesis.

Chapter 2

The CRESST Experiment

The *Cryogenic Rare Events Search with Superconducting Thermometers* (CRESST) is a dark matter direct detection experiment located 1400 m below the Italian Abruzzi mountains at the *Laboratori Nazionali del Gran Sasso* (LNGS) whose aim is to detect Dark Matter particles by measuring the particle-nucleus collisions in extremely cold, scintillating detector modules.

The history of the CRESST experiment is divided into three separate stages by their differing usage of detection crystals and sensitivity. This work uses the data taken from both CRESST-II (2013-2015) and the most recent stage CRESST-III (2016 - ongoing), where the setup has been updated for measurements of light dark matter particles with masses as low as 160 MeV/c^2 [2].

This chapter will first deal in Section 2.1 with the general experimental setup of CRESST and then, in Section 2.2, discuss detector modules used in the experiment, especially the detector whose data was used for this thesis.

2.1   Experimental Setup

One of the main obstacles of Rare Event Searches in a cryogenic high precision setup is the eradication of background. The first step in combating the problems of background radiation will therefore be to install such experiments in environments with natural shielding, e.g. below mountains. The already mentioned location of the CRESST experiment is then able to reduce the influx of muons by a factor of 10^{-6} compared to the sea level equivalent [34]. To reduce the remaining pollution from natural radioactivity of the surviving muons, surrounding rock and lead shielding, further precaution measures need to be taken. Further, any type of vibration from the outside, i.e. passing vehicles, needs to be accounted for, by decoupling the experiment from its surroundings.

The setup of the CRESST experiment is shown in Fig. 2.1.

liquid nitrogen

liquid helium

SQUIDs

cryostat

water

internal lead

cold finger

radon box

internal polyethylene

detector modules
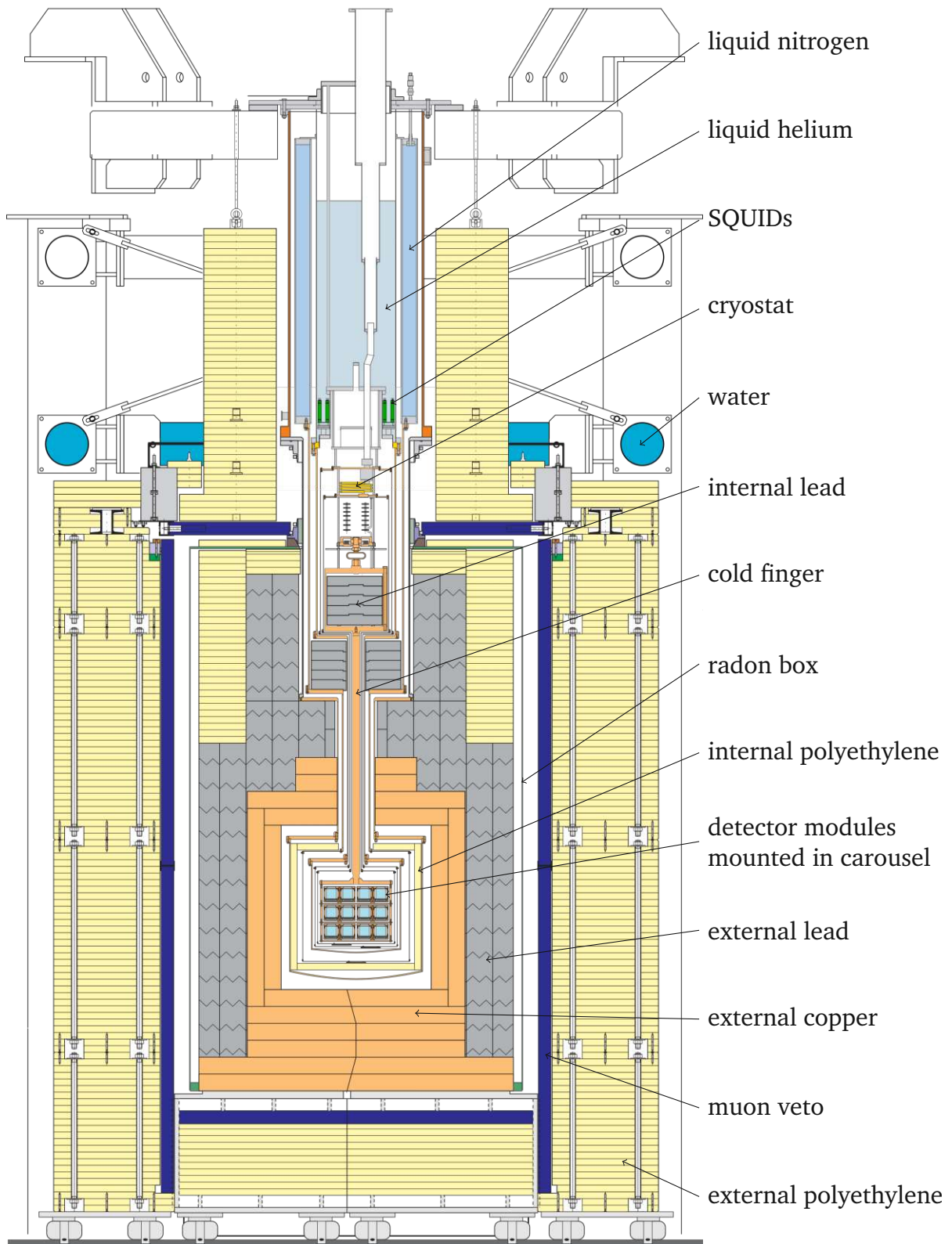mounted in carousel

external lead

external copper

muon veto

external polyethylene

Figure 2.1: Schematic drawing of the CRESST experimental setup. Figure taken from [37].

In order to tag the rest of the surviving muons plastic scintillator panels are installed and cover almost 100% of the solid angle around the cryostat. The panels then act as an active muon veto and allow for filtering of coinciding muon events.

Inside the plastic panels the thick lead and copper shieldings are applied and serve mainly to absorb the radiation from the surrounding rock, which is home to radioactive materials such as ^{40}K, ^{232}Th and ^{238}U. These elements and their decay products radiate difficult to block γ-radiation which is the reason for the necessity of thick shielding with materials that have high mass attenuation coefficients like lead. α- and β-particles on the other hand are easily shielded by thin metal sheets around the experiment. The radio-pure copper shielding is necessary to block radiation from the lead, which cannot be produced as radio-purely as copper.

The large polyethylene mantle around and inside the 6m high experiment as well as the internal smaller mantle of polyethylene is essential for the functioning of the CRESST experiment as the materials high content of hydrogen molecules is able to efficiently attenuate incoming neutrons. These are produced outside the laboratory or in the lead and copper shieldings via muon interactions or $(\alpha; n)$ reactions. This is essential because signals of the electrically neutral particles are almost indistinguishable from WIMP signals.

Finally, the existence of the gaseous ^{238}U decay product ^{222}Rn needs to be addressed. As ^{222}Rn itself is radioactive, continuous flushing of the surrounding gas-tight radon box with radioactively clean nitrogen gas is necessary. This is indicated in Fig. 2.1 by the colorless portion between scintillating panels and lead shielding. Additionally, to ensure gas-tightness, the radon box is kept under slight overpressure.

The whole experiment is cooled to extremely cold temperatures below 7 mK, as to mitigate thermal noise and to ensure ideal temperature rise in the thermometer ($\Delta T \propto T^{-3}$) [34]. This is achieved by employing a commercial dilution refrigerator circulating a ^{3}He /^{4}He mixture which is connected to the carousel, inside which the detectors are mounted. The carousel and mixing chamber are connected via a copper rod called cold finger. The purpose of the cold finger is twofold: Not only does it keep the carousel in place, but it also realizes the thermal coupling between the detectors, and the cryostat, which is mounted above the lead and copper shielding. The detector modules found in the interior of the carousel are explained in the following section.

2.2 Detector Modules

Since the CRESST-II stage scintillating $CaWO_4$ crystals have been used for most detector modules. Predominantly realized in a 2-Channel setup with one phonon and one scintillation light detector, one can also find modules in 3-Channel setups such as the Gode3-module which will be discussed in Section 2.2.1. Recently, also other materials such as lithium, sapphire and silicon have seen use. The purpose for this setup is the above mentioned need to distinguish between electrically neutral particles e.g. WIMPs and neutrons, and electrically charged particles and photons. This is achieved

through the light detector which measures the scintillation light generated by electro-magnetically interacting particles mostly interact with the atomic shells and produce therefore much more scintillation light then nuclear recoils would. This fact can be illustrated via the so-called Light Yield (LY), a quantity which relates the energy of the scintillation light E_l to the energy deposited in the phonon detector E_p:

$$\mathrm{LY} = \frac{E_l}{E_p} \quad . \tag{2.1}$$

The energies here are measured in the respective electron equivalents. Therefore, electron recoils as induced by scattering of β- and γ-particles have a Light Yield of 1 [34]. Events stemming from other particles are *quenched* and therefore exhibit lower Light Yield, see Fig. 2.2.
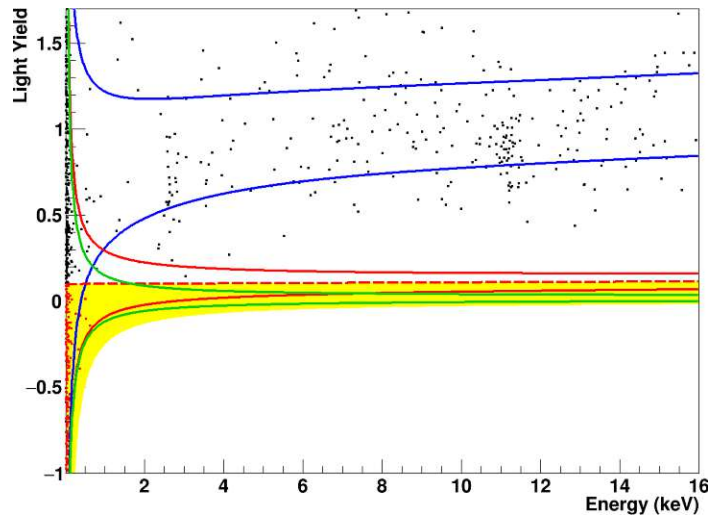

Figure 2.2: Light Yield for different scattering partners as a function of deposited energy. Within the coloured *bands* 80% of the expected event classes, i.e. β/γ (blue) or nuclear recoils off oxygen (red) and tungsten (green), can be found. The acceptance region, i.e. the region in which dark matter candidates are expected, is highlighted in yellow. Figure taken from [2].

The reason for the introduction of the electron equivalent is necessary because we calibrate the light detector with scintillation light, produced by electromagnetic recoils within the phonon detector from a γ-source. But the ratio of how much light is produced in such an interaction is highly dependent on the scattering partner. For electron recoils in $\mathrm{CaWO_4}$ about 6-7% of the energy goes into light production whereas for the scattering off nuclei only about 0.12%[1] of the energy produces light. The electron equivalent recoil energy is therefore an estimator for the recoil energy, in case

---

[1]This is calculated by assuming that the quenching factor for tungsten scattering ($\mathrm{QF_W}$) is given by, $\mathrm{QF_W} \sim 1/50$.

the scintillation phonon that scattered within the light detector, was produced by an electron recoil. The phonon signal itself is practically independent of the type of particle. Only a small difference is obtained due to light production. Finally, since both phonon and light signal are being measured, one can compute the total deposited energy by adding the energy deposited in the light and phonon channel [41].

Finally, the temperature of the target material is measured with a transition edge sensor (TES). As its name suggests such a sensor makes use of the fact that for superconducting materials operated in the vicinity of their critical temperature (*on the edge*) the resistance is highly dependent of the temperature changes, see Fig. 2.3. The CRESST-III-experiment utilises transition edge sensors made out of tungsten which are commonly evaporated directly onto the crystal. The crystals are equipped with a heater, needed for stabilizing the temperature of the superconductor. Together, crystal and thermometer constitute the phonon detector [8]. For the light detector a thin layer of tungsten - acting as a TES - is evaporated on the light absorber which is a silicon-on-sapphire (SOS) disk. The SOS disk now absorbs the scintillation light and converts the incoming photons into phonons. The following temperature increase can then be measured via the TES [8]. A typical transition curve of such a sensor is depicted in Fig. 2.3.
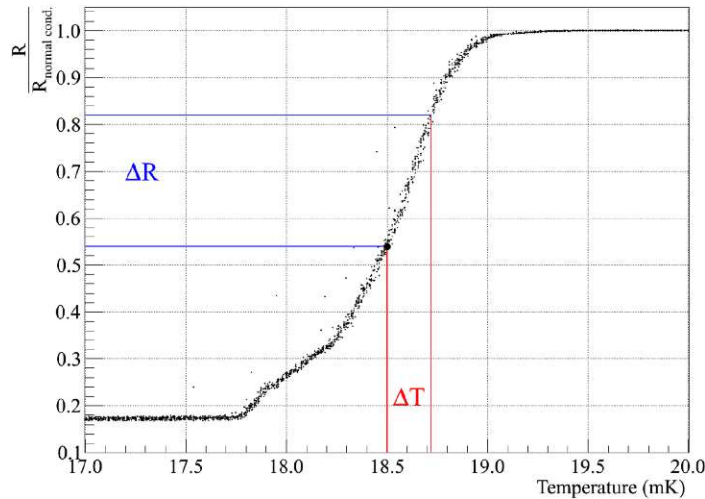

Figure 2.3: Transition curve of a Transition Edge Sensor used in the TUM40 detector module. For superconductors, as the one used in the TUM40 module's TES, one observes a typical drastic increase in resistance ΔR (blue) depending on the change in temperature ΔT (red) when operated around their critical temperature. Figure taken from [34].

In general, most of the detection modules used in CRESST are built in this way, only differing slightly in their nuances. For the purposes of this work only the single detector, whose data was used for later chapters, will be discussed in greater detail in the following.

2.2.1 Gode3 Detector

The Gode3 module is a 3-Channel module operated during the latest CRESST-III run, Run-36. The module has three separate detectors, two phonon and one light detector, with the goal of this setup and Run-36 in general being to explore the origins of the *Low-Energy Excess* (LEE) [2]. The excess, pictured in Fig. 2.4, is a exponential decay in the region below 1keV which is only visible in detectors with low enough thresholds.
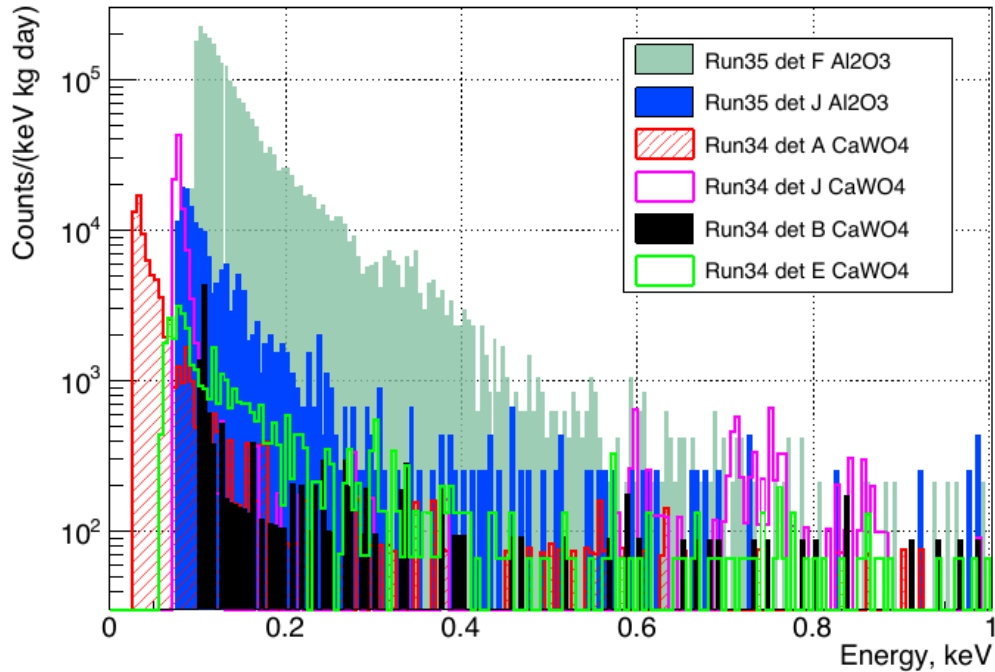

Figure 2.4: Visible here is the Low-Energy Excess for multiple detector modules of various CRESST-III runs. The signal seems to exhibit an exponential decay with respect to energy in the region below 1 keV. Figure taken from [42].

It is especially noteworthy that the excess is not a side effect measured exclusively within the CRESST experiment. As elaborated during the 2021 EXCESS workshop [46], we see excesses below the 1keV mark in multiple experiments, independent of *detector materials, sensors and holding structures, below and above ground, at different temperatures, and background levels* [46]. The authors find that the source of the LEE, since the signals appear with varying shape and rate between the experiments, seems to not stem from DM interactions. The excess furthermore appears to originate from multiple sources, due to its specific characteristics in the relevant experiments. The source might potentially coincide between experiments, but is unlikely shared between all of them. Details regarding measurement differences between experiment as well as an extensive comparative analysis of the data is presented in Ref. [46].

Building from the concepts introduced here, we will specifically delve into possible holder-related origins in the analysis of chapter 4. The idea of the Gode3-design is to specifically test if the LEE originates due to either shear strain in the holders of the phonon detector or events introduced by the surroundings [42]. This is achieved by implementing a second phonon detector, the so-called *ring* detector, which is being held by copper holders and additional implementation of the light detector as a silicon beaker providing an active 4π-surface veto. Therefore no line-of-sight is given to the absorber or other surfaces. A slightly adapted schematic of the Gode3 module is shown in Fig. 2.5.
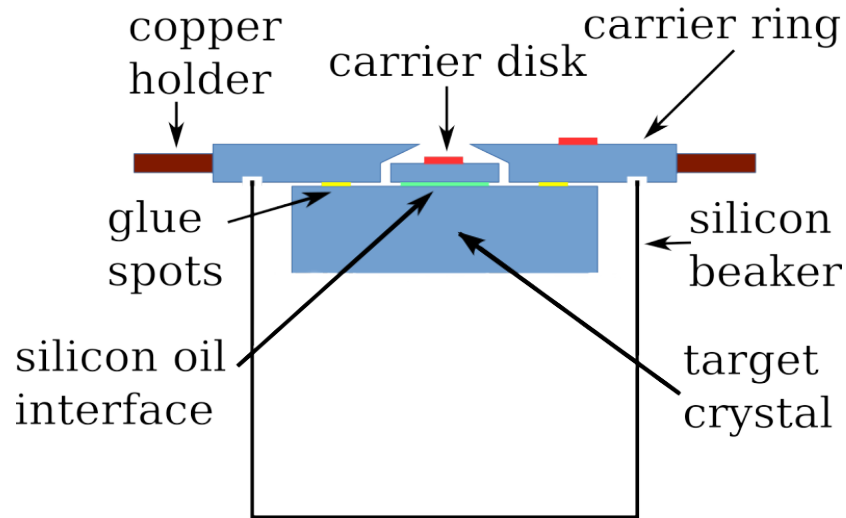

Figure 2.5: Schematic drawing of the Gode3 module. Pictured are: a) ring detector, labelled carrier ring (two parts held by the brown copper holders, blue). b) carrier detector, labelled carrier disk (thin rectangle, blue). c) light detector, a thin silicon beaker (black) with its own TES surrounds the phonon detector(s). The ring is attached to the absorber or target crystal (rectangular, blue) via glue spots (yellow) while the carrier detector is held by a silicon oil interface (green). Both the carrier and ring detector are equipped with a TES (red) and are therefore able to measure incoming particles. Figure taken and adapted from [42].

The Gode3 module revises a few aspects of the Gode1 detector, whose general setup is extremely similar to Gode3, operated in Run-35. Firstly, the thermal links between the detectors were weakend with the goal of reducing cross-heating between detectors [42]. The reduction amounts to a factor of 2 for the ring, while the strength of the thermal link in the carrier was weakened by a factor 10. What becomes obvious though when analysing the events of the Gode3 module is that the ring detector is extremely unstable even for only moderately high pulses which is most probably caused by an incredibly high sensitivity of both the TES and the SQUID. This has implications for the pulse shapes, turning them into artifacts as depicted in the following Fig. 2.6.
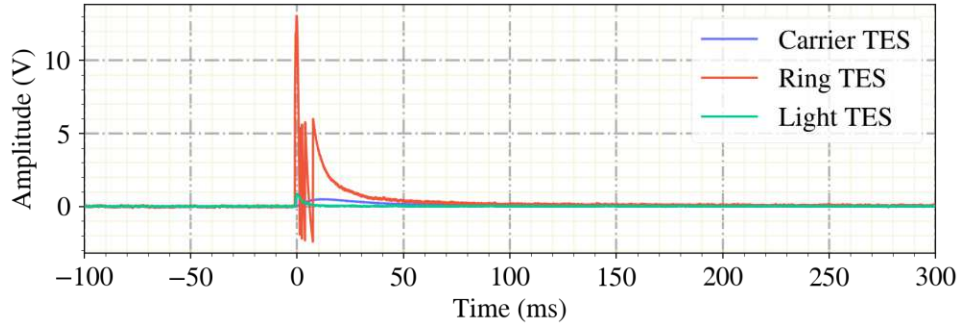
Figure 2.6: Overly saturated pulse in the ring channel while shape of carrier and right pulse show no signs of strong saturation.

In addition to that the TES size of the carrier was reduced by 75% and copper parts and screws were now etched. These measures lead to the fact that direct hits into the carrier are not visible in the bulk data contrary to the Gode1 data [42]. The discrimination of such event types and their implications for the analysis as well as the dynamic range for pulse detection will be discussed in the following chapter 3.1.1.
As was the case for the Gode1 detector, much lighter crystals enable the detection of low-energy interactions and therefore the ability of exploring the origins of the Low-Energy Excess [5, 7, 42]. In order to compare the results of the analysis in chapter 3, where the following terms are introduced, we will provide here the CPE factors, explained in great detail in Sec. 3.2.3, which converts test pulse equivalent pulse height to energy, baseline resolutions and threshold energies found by the author of [42] in the following table 2.1.

	CPE (keV/V)	Resolution OF fit (eV)	Threshold (eV)
Carrier	7.79	12.2	87.8
Ring	11.7	134	858
Light	7.2	133	944

Table 2.1: CPE factors, baseline resolutions and energy thresholds for ring, carrier and light detector of the Gode3 module found in [42].

Chapter 3

Raw Data Analysis for the Gode3 Module

The CRESST experiment, as all experiments attempting to detect light dark matter, is extremely dependent on the sensitivity of the detectors with noise from outside sources, described in the previous chapter, naturally being one of the largest limiting factors. Eradication of as much noise background noise as possible within the detectors, which in turn leads to lower detection thresholds, is only relevant if one is also able to preserve and extract the incoming particle pulses with appropriate tools for gathering and analysing the data. This chapter will thus aim to summarize and shine light on the analysis techniques developed during the CRESST-III phase by using the raw data gathered by the Gode3 detector showcased in the previous section 2.2.1. Gode3 presents itself as an especially interesting module due its previously explained, unique setup as a 3-channel detector which gives one the ability for research into the origin of the Low-Energy Excess. In addition to that, the module has seen only little analysis to date - the results so far were presented in [42] - and therefore promises highly fascinating results within the scope of this thesis. Here we will use the hardware-triggered training background data sets in a .rdt file format: `bck003`, `bck004`, `bck005`, `bck006`, `bck007`, `bck008`, `bck018`, `bck028`, `bck039`, `bck048`, `bck058` which amount to a total measurement time of 2850.832 hours. The measurements of the training data set took place between the period of October 2020 - April 2021.

In this chapter, the way from evaluation of the raw data up until the high level analysis, e.g. the generation of Light Yield plots and Dark Matter exclusion limits, will be described. The analysis process itself can generally be split into two parts portrayed in the following two sections. First, in Sec. 3.1, the hardware triggered data set is analysed, though the process of data acquisition itself will not be described. For an extensive description of the setup and procedure, especially hardware triggering, refer to either of [8, 34, 44]. Using the results of 3.1, the following section 3.2 will then outline the high level analysis of the triggered data stream. This, as well as all results in chapter 4, are accomplished with the help of the Python library *Cait* (Cryogenic

Artificial Intelligence Tools) [43]. The library was used for data handling, simulation and preparation and provides the CRESST-specific methods for all calculations done in this work.

3.1 Analysis of the Hardware-triggered Data

At first look it might seem superfluous to gather not only the data which was triggered by the hardware but also store the raw data stream. This technique has been implemented only recently during the CRESST-III phase for the ability of adjusting the threshold after datataking [44]. Since the hardware triggering necessitates the definition of a threshold beforehand, one is left with the following modus operandi, discussed in detail in [45]: First, one needs to analyse the hardware triggered data and reconstruct a typical pulse shape measured in the detector, a so called *standard event* (SEV). In addition to standard events the notion of the *Pulse Shape Model* as an analytic method of reconstructing an event signal is introduced. Using the information obtained in Sec. 3.1.1 we calculate the optimum filter and use the SEV to simulate a data set of 3000 events in 3.1.3. The standard deviation of these roughly Gauss distributed pulse height reconstructions will then represent the baseline resolution. Finally, Sec. 3.1.4 will deal with the determination of the trigger thresholds via an estimation of the noise trigger rates. With knowledge of the correct trigger threshold for the data at hand one is now able to again apply the triggering procedure which would now also account for low-energy pulses below the hardware trigger threshold. This data set will then constitute the basis for the high level analysis of section 3.2.

3.1.1 Standard Events and Pulse Shape Model

In this section we will first introduce the methods of reconstructing general pulse shapes found in a detector in CRESST. This will provide the base for taking the data generated by the hardware-triggering and extracting the standard events present in the data set, showcased in the following section 3.1.2. One can find for reference typical "good" pulse shapes of the Gode3 module in Fig. 3.3.
The SEVs are created by averaging multiple - usually the number is a few tens - pulse signals with similar pulse height. The large number of events has the effect that any noise present in the events is averaged out while one needs events of similar pulse height in order to avoid effects of the so-called *Trigger Walk*. The phenomenon has its origin in the differing rise times of pulses with different pulse heights. Pulses with larger amplitudes reach the threshold value much faster than low-amplitude signals. The standard events play a big part in the analysis of CRESST data due to their ability to reconstruct pulse shapes even for pulses with small amplitudes - as lower energy signals might often be drowned out by noise - and thus extracting important information about the energy of the incident particle [5]. The procedure with which one finds the onset time, pulse height, and other parameters of a certain pulse can be

the *standard event fit* [5] for instance. Another more commonly used method, the so-called *main parameters*, are utilised later on. The fit builds either upon the standard event template or the so-called *Pulse Shape Model* (see following explanations), first introduced in [9], by fixing most parameters of the model and fitting the pulse height, onset time, baseline offset and the drift [5]. The fit function that needs to be minimised is governed by either the SEV or the pulse shape model s superposed with some baseline offset d and a linear drift k of the form,

$$\arg\min_{h,t_0,k,d} \left(\sum_x (e[x] - hs(t - t_0)[x] + kt[x] + d)^2 \right) , \tag{3.1}$$

where the fitted pulse height h and onset time t_0 depict the true values of the event $e[x]$, with $t[x]$ acting as an array of time values. The variable x represents here the array indices of each value, as introduced in [5]. It is also important to point out that in principle one is not restricted to a linear model for the baseline. In addition to the four fit parameters h, t_0, k, and d the Root-Mean-Square (RMS) error of the fit is often also stored for analysis, since large fit errors on some event indicate the presence of an artifact in the data as we are able to extract deviations of the fitted pulse from the nominal pulse shape.

The Pulse Shape model [9] used in the fit constitutes an analytical function which tries to explain the response of a detector without considering detector cross-heating. One has to keep in mind here that this plays a large role in the Gode3 data, see Sec. 2.2.1. It assumes a thermal heat circuit which links power inputs on the TES with those on the absorber. For the short detour in this section explaining the model, the simplified circuit shown in [5] is used as well as the simplified considerations in Sec. 3.3 of [9].

The reconstruction of the pulse shape assumes a successive, rapid ($\mathcal{O}(\text{ns})$) relaxation of non-thermal optical phonons generated by the particle interaction into acoustic thermal phonons with frequencies of 300 GHz [9,44]. In our simplified case, depicted in Fig. 3.1, the absorber crystal and thermometer are coupled to each other and to a heat bath operating at temperature T_b via thermal conductivities G. The heat bath will in the case of the CRESST experiment correspond to the detector housing which in turn is connected to the carousel and cold finger via a copper wire [5].

The heat capacity of the thermometer, when neglecting the thermal conductance of the thermometer phonons [9], is given mainly by the electron temperature T_e and their heat capacity C_e with an additional external and time-dependent power input $P_e(t)$ arising due to the absorption of high-frequency phonons by free electrons of the thermometer's metal film. The situation for the absorber with temperature T_c and heat capacity C_c is similar where the power input $P_c(t)$ will in this case be given by the thermalisation of high-frequency phonons on the crystal surface [9]. P_e and P_c can then be defined to be of the following form

$$P_c(t) = \Theta(t)\varepsilon P_0 e^{-t/\tau_n} \tag{3.2}$$

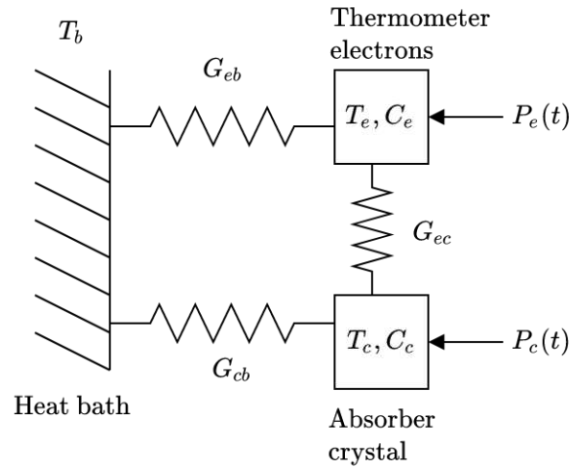$$P_e(t) = \Theta(t)(1 - \varepsilon)P_0 e^{-t/\tau_n} \quad , \tag{3.3}$$

Figure 3.1: Thermal model of CRESST detector with temperatures for the heat bath T_b, absorber crystal T_c, and TES electrons T_e. All three components are connected via heat conductivities G_{cb} (crystal - heat bath), G_{ce} (crystal - electrons), and G_{eb} (electrons - heat bath) with external time-dependent power inputs $P_c(t)$ and $P_e(t)$ acting upon the absorber and thermometer respectively. Figure adapted from [9].

when assuming that the rate of thermalisation is independent of the phonons' frequency and that a fraction ε is thermalised in the thermometer $(1-\varepsilon)$ in the absorber. The normalisation constant $P_0 = \Delta E/\tau_n$ is determined by the quotient of deposited energy ΔE by relaxation time of the non-thermal phonons τ_n. Finally, we assume that the population of non-thermal phonons is present immediately after the the particle interaction at time t which is given by the Heavyside function Θ.

Using the previously defined expressions (3.2) for P_e and P_c one can now solve the differential equations decribing the circuit in Fig. 3.1 and finds the following solution for the thermometer signal $\Delta T_e(t) = T_e(t) - T_b$ [9],

$$\Delta T_e(t) = \Theta(t) \left[A_n(e^{-t/\tau_n} - e^{-t/\tau_{in}}) + A_t(e^{-t/\tau_t} - e^{-t/\tau_n}) \right] \quad . \tag{3.4}$$

The temperature signal described by above equation has already been used successfully to model in [5]. What is to be noted in (3.4) is that the reconstructed pulse has two distinct components corresponding to the rapid non-thermal phonons with amplitude A_n and the slower thermal component A_t. The time constants τ_{in} and τ_t are found by solving the homogenous differential equation. All constants are defined by the heat capacities, resistances and the deposited energy [5].

The Pulse Shape model is especially dependent upon the mode of operation of the detector which are governed by the ratio of τ_{in} and τ_n and thus controlling the sign of the non-thermal component A_n. The operation modes are either that of the *bolometric* or *calorimetric mode*. The former, also known as energy flux mode, is defined

by a much faster intrinsic relaxation time of the thermometer compared to the non-thermal phonons ($\tau_{in} \ll \tau_n$) which basically leads to measurement of the energy flux. For the reversed case ($\tau_{in} \gg \tau_n$), the so-called energy integral mode, where the non-thermal phonons decay much faster, the detector functions as a calorimeter and measures the deposited energy. The size of τ_{in} can be adjusted by modifying the strength of the thermometer-heat bath coupling [34, 44]. The calorimetric mode is the predominantly used mode in light CRESST-III detectors [5, 44].

3.1.2 Gode3-specific Standard Events

With the technical description of standard events completed we now move towards the practical determination of the Gode3 standard events. As explained earlier, the generation of SEVs depends upon the overlay of several similar measured pulses in the hardware triggered data set.

The extraction of these events is usually done by applying *Quality Cuts* on the data and keying in on the sought-after pulse shape. Here it is important to note that one detector may contain more than one single SEV as will become clear soon. Therefore, the applied cuts need to not only remove any artifacts within the data set at hand but also uniquely define the present standard events. Such artifacts in need of removal are often one of the following anomalous events originating directly in the detector - more precisely in the electronic readout of the pulses - or stemming from environmental sources [34, 44]:

- **Voltage spikes**: single voltage outliers (can be either positive or negative) in an otherwise empty baseline.

- **Flux quantum losses**: the SQUID within the readout scheme is not able to follow the fast rising pulse leading to a change in baseline level.

- **SQUID reset**: reset of the SQUID by the electronics within a record window where the baseline jumps suddenly.

- **Decaying baseline**: the tail of a high-energy pulse in the previous record which has not relaxed in the current window.

- **Pile-Up**: Pile-up of two or more events within the record window.

- **Noise triggers**: The hardware trigger was activated by a noise fluctuations of an otherwise empty baseline.

- **Stick event**: Slowly rising and decaying pulse following an energy deposition in the $CaWO_4$-sticks holding the crystal. This type of event is not relevant for Gode3 since no sticks are present.

The choice of quality cut is most often determined by inspection of feature plots and subsequent selection of constraints on the features if possible. The analysis in this work made use of the parameters calculated within *Cait*, e.g. the main parameters [2, 43], found by taking a 50-sample moving average. and other constants derived from these values. For the data at hand it proved to yield the most information about the types of pulses measured by the detector when plotting the rise or decay times of the pulse versus its pulse height. The rise time is defined as time between the onset of the pulse up until the reaching of 80% of the pulse's maximum. The decay time on the other hand displays the time in which the pulse has decayed from 90% to 30% of its maximum value. For the analysis in this work the plot for rise time versus pulse height depicting all three channels, Fig. 3.2, acts as a stepping stone in finding the correct standard events.

(a) Carrier channel
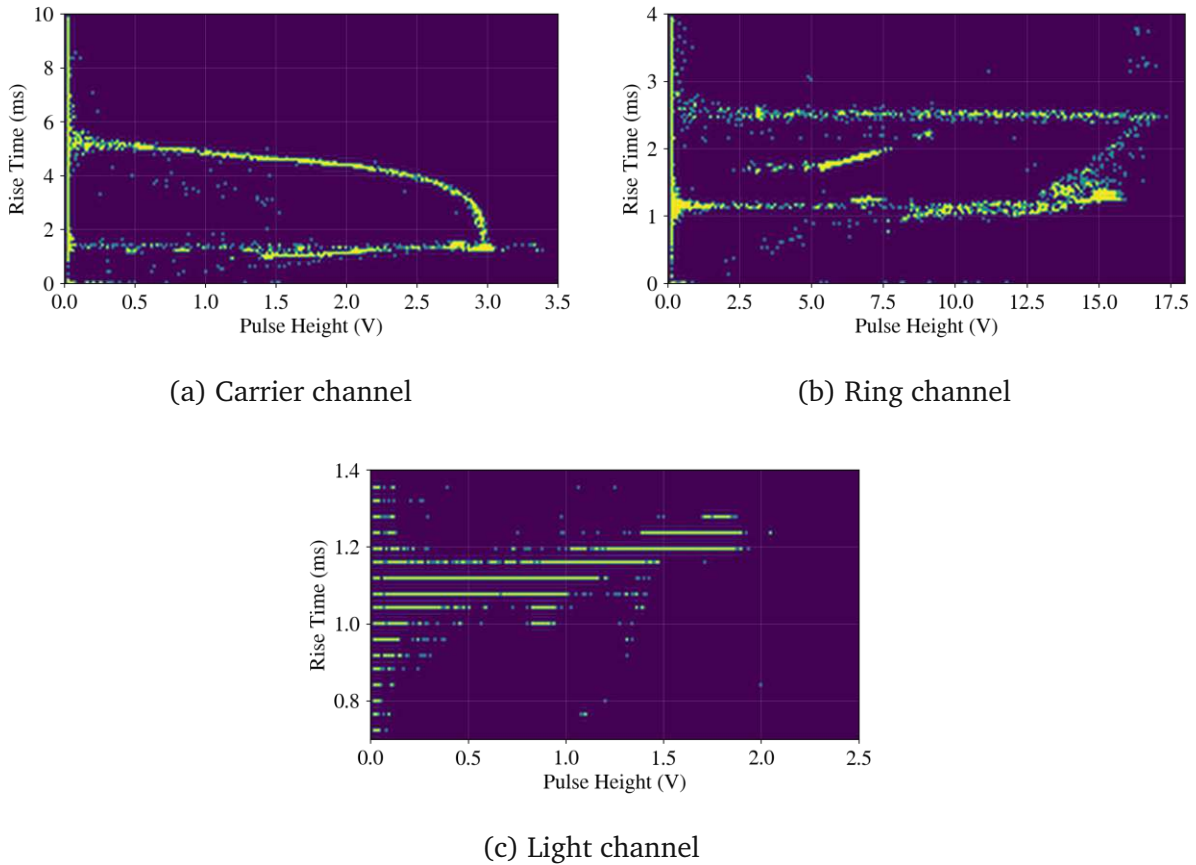

(b) Ring channel

(c) Light channel

Figure 3.2: Rise time vs. pulse height plots for all three Gode3 channels. The peculiar split into two bands for the phonon channels - channels 0 (a) and 1 (b) - indicates that there must exist at least two standard events in the measured data set. The light channel - channel 2 (c) - does not exhibit these characteristic hinting at separate phonon and light SEVs. The splitting into discrete values in the light channel is due to the precision of the digitizer.

For the phonon channels we ascribe the two bands to two different types of impact destinations for the incident particles. On one hand, we have hits into the target or absorber crystal and on the other hand direct detector ring hits. The different rise times of the two events stem from the fact that the distance phonon excitations travel until detected play a large role in the shape of a pulse. Therefore an impact in the ring of the detector will first need to travel through the absorber crystal before being detected in the carrier, rendering a much slower pulse with lower amplitude when compared to the same pulse detection in the ring channel. Conversely, an impact in the absorber crystal gives also two slightly different pulse shapes - though much less extreme than for direct detector hits - between phonon channels due to the differing detector geometries and setups between carrier and ring channel.

In the following figure Fig. 3.3 typical members of the two bands are presented.

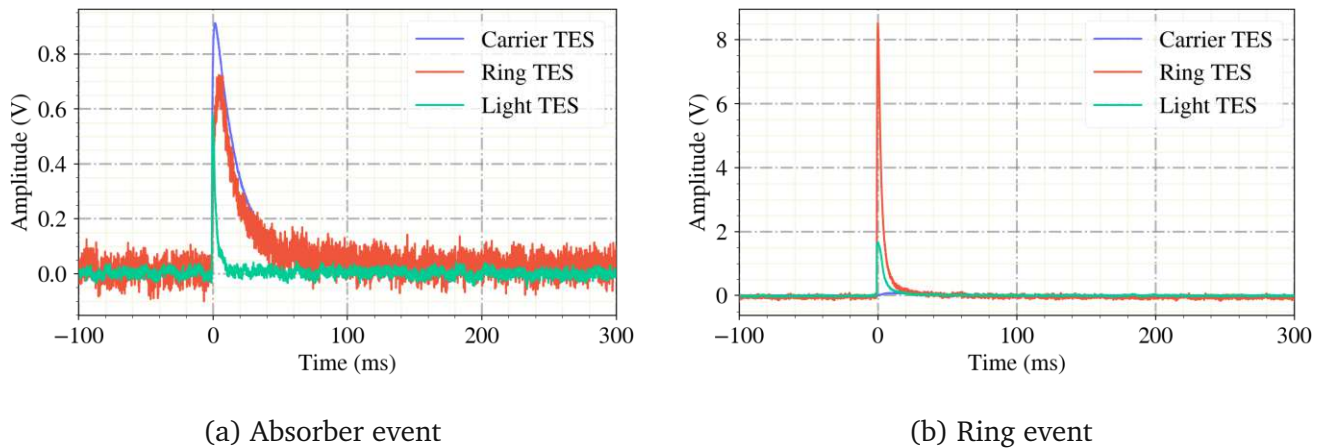(a) Absorber event   (b) Ring event

Figure 3.3: Shown here are examples of a typical a) absorber and b) ring event present in the Gode3 data. The absorber event showcases a visibly slower rise and decay time in the ring TES when compared with the carrier TES and an additionally only slightly smaller amplitude measured in the ring channel. The picture for the ring event is much more extreme where the signal amplitude in the carrier TES is barely visible and features a slow rise to maximum. In contrast we see a sharp peak with high amplitude in the ring TES.

Visible are especially the differences in pulse height and rise/decay time between detection channels and especially event types. For absorber events the differing rise times between channels (lower band in 3.2a at 1.8ms versus upper band in Fig. 3.2b at 2.8ms) in Fig. 3.2 are visible as well as the immense difference in measured pulse height between channels and the rise time difference for ring events. The discrepancy in pulse height also allows for interesting applications with regards to the dynamic range of the detectors. Consideration of the other channel when analysing helps differentiate between real events and artifacts such as over-saturation and others, e.g. Fig. 2.6.

What is most striking in the rise time plots Fig. 3.2 is that a third band corresponding to direct carrier crystal hits is missing. These types of events are detected in most of the CRESST module across all runs (see extensive analysis of carrier events featured in [5, 8, 34]) and also similar modules such as the Gode1 but are seemingly absent in the Gode3 data. The matter of the fact is that these events are not missing in the data at hand but rather incredibly rare when compared to the previously discussed event forms. Plotting the events in the short rise time region of the carrier channel yields only a hand full of direct carrier hits. The rarity of the carrier pulses can be explained by the specific geometry of the Gode3 module: the carrier channel is encircled by the ring which gives a smaller probability of impact for incident particles. Additionally the thermal link of the carrier TES was weakened as mentioned in 2.2.1 leading to smaller amplitudes measured by the detector. Analysis in [42] reflects the hypothesis that only two types of events are relevant for the subsequent analysis.

In order to achieve adequate separation of the data set into enough representative events for the standard event generation quality cuts need to be chosen which eliminate artifacts and distinguish between the two types of detector hits. The analysis proved that focus on the two upper bands of ring and carrier channel lead to better SEVs with noticeably less influence from artifacts. The chosen rise time intervals in milliseconds were for absorber events $[2.2, 2.8]$ and for direct ring hits $[4, 5.9]$. Additionally, the pulse height in Volts was restricted to $\mathrm{PH_{Abs}} \in [5.2, 6]$ for absorber events and to $\mathrm{PH_{Ring}} \in [0.9, 1]$ for ring events.

For the until now unmentioned light channel it is also necessary to find the SEV belonging to absorber hits. The issue here is that from earlier analysis in Run-35, it was known that the light channel also saw direct hits leading to differing pulse shapes in the data. The rise time plot (Fig. 3.2c) on the other hand shows no direct structure, as was the case for the phonon channels. Here another approach needed to be taken, namely that direct light hits have to show no pulse in the phonon channels, since the sensors were not linked as is the case for carrier and ring detector. Therefore, the direct light detector hits could be found in a data set where the pulse heights in ring and carrier were below 0.2 V while the scintillation light from the crystals has pulse heights above 0.2 V in the two channels. Further, a rise time cut on the light channel of $[0.9, 1.3]$ ms and an Onset cut of $[-20, 20]$ ms was applied for generation of the scintillation SEV, while for direct light hits a pulse height cut on the light channel with an interval of $[0.1, 1.6]$ V was applied to cut artifacts from the data set. During the generation of the different SEVs also a Left-Right cut of 0.5V was applied to filter out decaying baselines etc.

The above mentioned cuts produce SEVs which seem to mirror the types of events present in the data closely. The standard events which are generally normalised to an amplitude of 1V are depicted in the following plots Fig. 3.4. All applied cuts are summarized in Tab. 3.1.
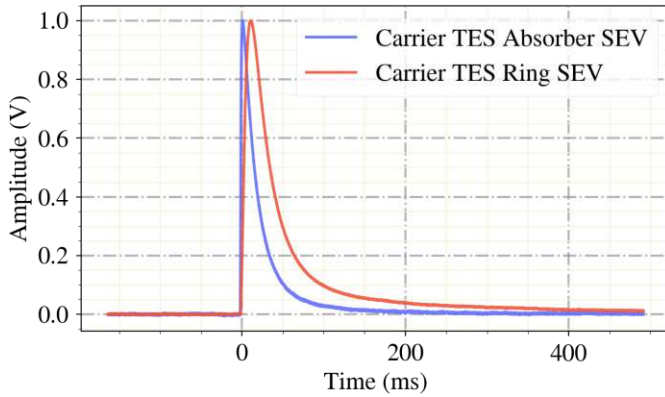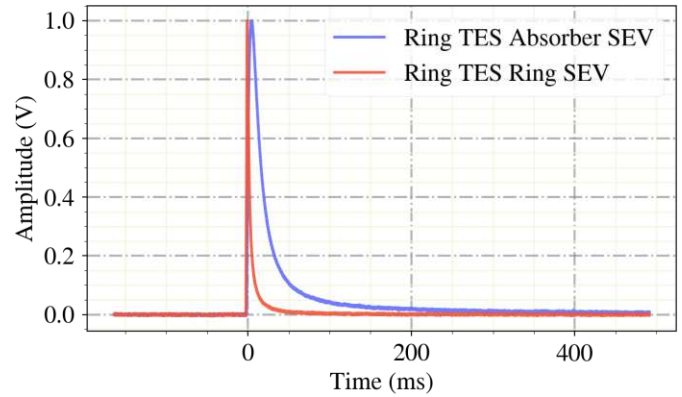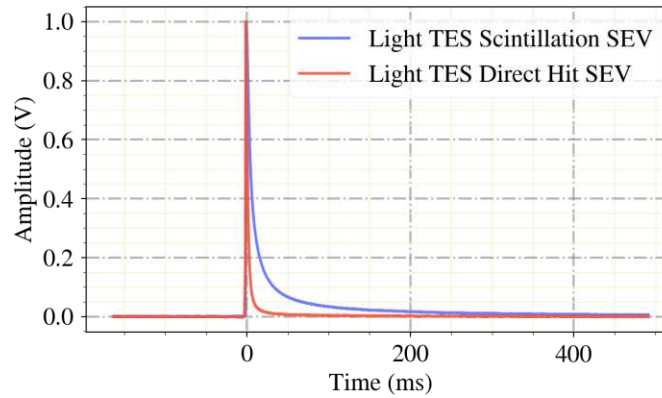
(a) carrier channel

(b) Ring channel

(c) Light channel

Figure 3.4: Pictured are the possible standard events observed in the Run-36 Gode3 data. In Fig. 3.4a we see the absorber (blue) and ring (red) standard events in the carrier TES. Fig. 3.4b displays the same two SEVs only in the ring TES. The expectation formulated in Fig. 3.2 rings true with the absorber SEV having a shorter rise time in the carrier TES than the ring SEV but the situation being reversed in the ring channel. Finally, Fig. 3.4c shows the possible standard events - the scintillation (blue) and the direct hit (red) SEV - in the light channel.

SEV	PH_C [V]	PH_R [V]	PH_L [V]	RT_C [ms]	RT_R [ms]	RT_L [ms]	OT_L [ms]
Absorber		$[2.2, 2.8]$			$[5.2, 6]$		
Ring	$[4, 5.9]$			$[0.9, 1]$			
Scintillation	> 0.2	> 0.2				$[0.9, 1.3]$	$[-20, 20]$
Direct light	< 0.2	< 0.2	$[0.1, 1.6]$				

Table 3.1: Applied cuts for SEV generation. The following shorthand was used for pulse height (PH), rise time (RT) and Onset (OT). The subcripts indicate on which channel the cut was applied: Carrier (C), Ring (R), or Light (L).

3.1.3 Optimum Filter and Baseline Resolution

After creation of the absorber SEV one is now able to move on to the calculation of the baseline resolution. The knowledge of the absorber SEV is especially important since we want to measure particles with the absorber crystal, ergo model the detector threshold to this type of event. The SEV in question will now first be used to determine the so-called *optimum filter* (OF) with which we calculate the baseline resolution. The OF will later also be utilised for the triggering process.

The method of employing the OF for the calculation of the resolution will later prove to lead to the best resolutions. First, it will be necessary to define the optimum filter. The method, introduced by Gatti and Manfredi in 1986 [10], used for pulse height estimation has already been succesfully applied in CRESST and other cryogenic rare event searches, e.g. [5, 44, 47]. The goal of filters is to minimise the influence of a noise signal on the true signal [48]. This is achieved by employing a filter which is matched to the target signal, the matched filter, and whose main property is the maximisation of the signal-to-noise-ratio (SNR) in frequency space [48, 49]. The output $g(t)$ of filtering is given by following equation where a filter $H(\omega)$ is applied to the input signal $S(\omega)$ via convolution [10]

$$g(t) = \frac{Q}{2\pi} \int_{-\infty}^{\infty} H(\omega)S(\omega)e^{i\omega t} \mathrm{d}\omega \quad , \tag{3.5}$$

where Q is a normalisation constant. The authors of [10] then calculate a filter which maximises and estimates the signal's amplitude at the time of measurement t_m to be

$$H(\omega) = K \frac{S^*(\omega)}{N(\omega)} e^{(i\omega t_m)} \quad , \tag{3.6}$$

with the star * denoting complex conjugation and K again a normalisation constant. The symbol $N(\omega)$ denotes the noise-power-spectrum (NPS) of the baseline signal. The NPS constitutes the squared absolute values of the average Fourier transformed empty noise baselines [5, 44].

Since the real input signal $S(t)$ of the event pulses is not known generally, it is here where the standard event comes into play. The incoming signal is assumed to be a scaled version of the present standard event $S_0(t)$ multiplied by some factor a, acting as the amplitude of the pulse,

$$S(t) = aS_0(t). \tag{3.7}$$

The value of the above menioned constants Q and K is then set by requiring the amplitude of the SEV to be $a = 1$ [44]. The Fourier transform of the standard event is inserted into the equation for the filter (3.6) which in turn is then used for the convolution (3.5) with the pulse to create the filtered signal. Thus, one finds the optimum filter pulse height to be equal to the value of the scaling factor a.

The OF calculated for the previously defined absorber SEV is pictured for all three channels in the figure 3.5 below.
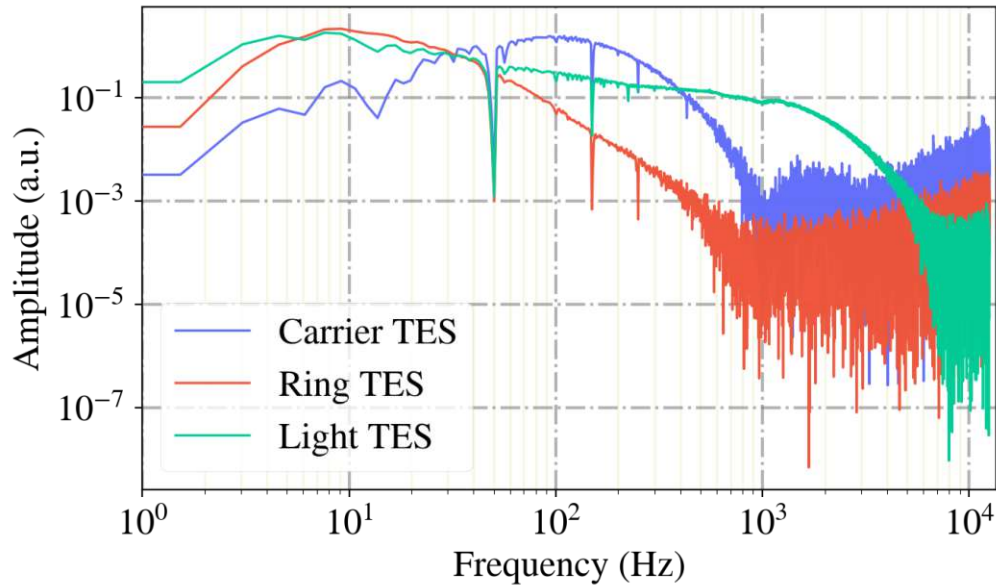
Figure 3.5: Optimum filter transfer function for all three channels of the Gode3 module. The light channel (green) experiences less supression by the filter over a larger frequency interval compared to the carrier (blue) and ring (red). This might be due to less noise in the detector or pulses of higher frequencies. The prominent voltage noise frequency at 50 Hz and its harmonics are are heavily suppressed by the filter.

With knowledge of how one can determine the optimum filter pulse height, we are now able to calculate the baseline resolution of the Gode3 module which in turn can then be used to extract the trigger threshold. The baseline resolution used in this work is determined by finding the standard deviation of the approximately Gauss distributed noise baselines measured during the run. In order to achieve this goal, 3000 simulated events are generated by superposing the absorber SEV with the empty noise baselines. Then, a histogram over the pulse height is generated which should be centered around $1V$, since the SEV is normalised exactly at this value. A Gauss distibution is finally fitted to the histogram via a Maximum Likelihood fit.

From the previous discussions it should be clear that finding the pulse height of events with noise is not a simple task. Therefore, we utilise the previously presented methods of pulse height reconstruction with the aim of generating the sought after histograms and thus finding the baseline resolution. Here we will employ the optimum filter pulse height, which generally produces the best resolution, the pulse height of the standard event fit as well as Cait's pulse height mainparameter. The baseline resolutions found by the three different methods are documented in Tab. 3.2.

	Carrier channel		Ring channel		Light channel	
	Mean [V]	σ [mV]	Mean [V]	σ [mV]	Mean [V]	σ [mV]
Optimum filter	0.9779	1.3	1.02	8.96	0.9722	5.05
SEV fit	0.99992	2.76	1.00007	8.95	0.9995	6.32
Pulse height	0.99	2.87	0.9988	15.2	0.931	6.69

Table 3.2: Mean and standard deviation σ or baseline resolution of the simulated events. The optimum filter pulse height produces the lowest resolution out of all three methods though it shows a negative bias in the means in all three channels. The SEV fit pulse height and the pulse height mainparameter perform significantly worse across all channels.

The table above shows clearly that the pulse height determined by the optimum filter produces the best baseline resolutions, called σ above, out of all three methods though at the cost of implementing a bias of about 2-3% into the mean of the distributions. This bias might originate from noise present in the SEV itself. The bias can be remedied by multiplying the baseline resolution with the inverse of the mean. We also see a slight bias in the pulse height mainparameter due to the fact that the parameter is determined as the maximum of the time series with implementing a 50 sample moving average. This inherently produces a negative bias as the peak is basically rounded off by a smoothing function [5]. All three channels show that the resolution in the ring channel is by far the worst which is again most probably due to the instability of the ring channel. Here the SEV fit pulse height also shows a slightly better, by $10\mu V$, baseline resolution than the optimum filter.

With the baseline resolution calculated for all three channel one is now able to find the trigger threshold of the detector module at hand. In CRESST the threshold is typically defined to measure 1 noise trigger per kg and day [44], a method presented below in the following section. Alternatively, one could choose an appropriate value for Gaussian noise, above which only little noise can be expected in the data. In CRESST this value is commonly chosen to be between 5 and 6.5 times of the standard deviation of the distribution. Thus, all that is left to find the preliminary trigger threshold is to multiply the unbiased OF baseline resolution of all three channels with the here chosen value 6.5 leading to the following set of thresholds.

	Carrier	Ring	Light
Threshold [mV]	8.64	57.1	33.77

Table 3.3: Trigger thresholds calculated using baseline resolution results.

This method of determining the trigger threshold is not the only one used within CRESST data analysis. In the following section we present the more widely used method, e.g. in [44], for finding the energy above which only one noise trigger per kg and day is measured and with which we later trigger the stream data set.

3.1.4 Trigger Threshold via Noise Trigger Rates

Lastly, this section will now deal with the way the trigger threshold was calculated for this thesis by utilisation of *noise trigger rates*. This method can thus be summarized as way of using the standard events extracted from the hardware-triggered data set to then apply the found trigger thresholds on the stream data set. We will see that the results found with the noise trigger method produces similar results to those in Tab. 3.3 though we will use the thresholds found in this section for the triggering of the datastream. The method will be described briefly in the following.

In a first step, the optimum filter is applied to the empty noise baselines, thus creating baselines which resemble those "seen" by the trigger during the triggering process. This is because the datastream is too filtered before it passes the software trigger. Following the filtering of the noise baselines a histogram for the "pulse height" or maxima of those noise triggers per kg, day and mV is generated. If Gaussian noise is assumed, one can again fit a Gauss distribution to the resulting histogram. This model produces in some cases, like channels with high rates, unsatisfying results at the tails of the distribution due to pollution with artifacts. Within Cait a novel method of modelling deviations from Gaussian noise has been implemented. There, this pollution of the noise can be filtered out by a second *pollution component* - modelled by an exponential decay - which leaves the undistorted *Gauss component*. Finally, one finds the trigger threshold by finding the amplitude in Volts at which the cummulative counts from the right tail equal exactly 1 noise trigger per kg and day, plotted in Fig. 3.6 for the light channel. From the plots one can extract the trigger thresholds subsequently proceed with the triggering of the datastream. The thresholds are listed in Tab. 3.4. Though the thresholds listed in Tab. 3.3 are generally lower than the ones found through the noise trigger rates, the following analysis will build upon these more conservative values as they reflect the true thresholds more accurately.

	Carrier	Ring	Light
Threshold [mV]	9.943	55.659	38.733

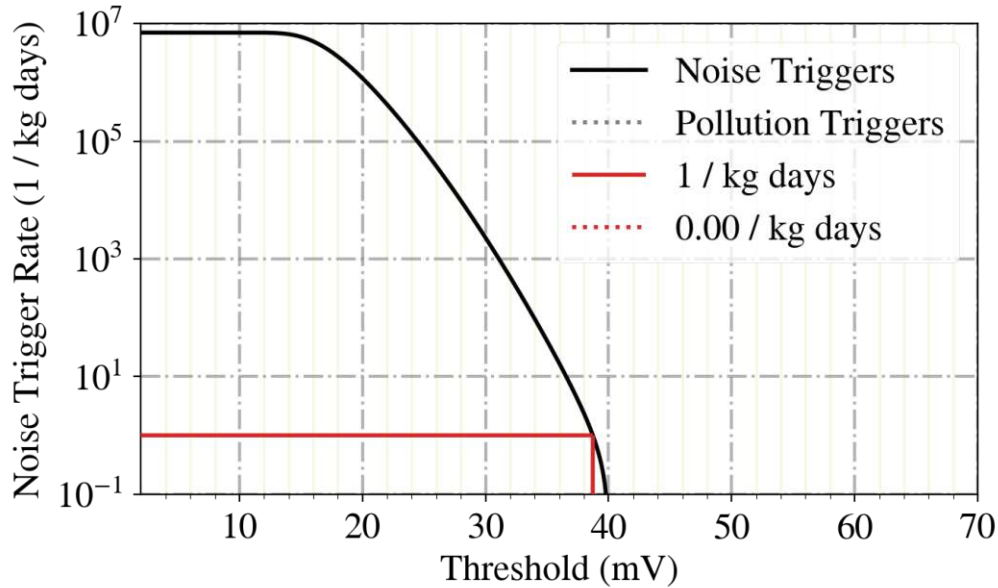Table 3.4: Trigger thresholds calculated using noise trigger rates.

Figure 3.6: Cumulative noise trigger rate per kg and day of the empty noise baselines for the light channel. The threshold is set at the amplitude at which 1 count per kg and day is measured. In this specific case of the light channel one can see that the pollution component of the model is not visible, ergo the data is mostly pollution-free.

3.2 Analysis of Data Stream

With all the necessary information obtained in the previous section one is now able to begin with the analysis of the continuous data stream. The data set, stored in three *.csmpl-files (one per channel), is triggered with the thresholds for each detector noted in Tab. 3.4. By applying the OF to the stream and extracting those record windows with OF pulse heights above the trigger threshold the events are generated. We define record windows as the voltage values which are stored in a window of 16384 samples, where the trigger is placed at $1/4$ of the record window. The triggering produced in sum 11622 events measured over a period of 2851 hours. The exact definition of the triggering algorithm is described in great detail in [45].

Following the export of the events, the first step, in Sec. 3.2.1 is to cut possible artifacts with two of the major cut approaches called the rate and stability cuts. Later, in Sec. 3.2.3, the data set is subjected to further quality cuts ending with the calculation of the cut efficiencies in 3.2.2. After determination of the conversion factor between TPA equivalent pulse heights and particle energy - the *CPE factor* mentioned in Chap. 2 - the section concludes with a presentation of the energy spectra and Light Yield plots in 3.2.4.

3.2.1 Rate and Stability Cut

The triggering procedure does not guarantee an artifact-free data set readily available for analysis. Therefore, the energy calibration of the underlying data is often preceded by a cut procedure which is split into two distinct parts. For one, it is possible to cut events from the data set by excluding time intervals of unusual behaviour for the detector. On the other hand the analyst then cuts the bulk of the artifacts from the data by quality cuts as done for the generation of standard events.

In CRESST, intervals with this "unusual behaviour" are defined by time periods of either increased event rates or increased control pulse heights deviating from the mean. The first cut is referred to as the *rate cut*, described in the following paragraphs, while the second - the *stability cut* - is dealt with afterwards.

Periods of heightened event rates are often caused by outside disturbances, e.g. by electromagnetic interference and similar phenomena [44], leading to unnaturally high rates of triggered events. The reason why the analysis of CRESST data expects an approximately constant event rate is due to the fact that a constant background data plus possible dark matter interactions are gathered during data taking. A segment of the first 300h of Gode3 measurement is displayed in Fig. 3.7 featuring such a phase of increased event rates at ca. 65h for instance.
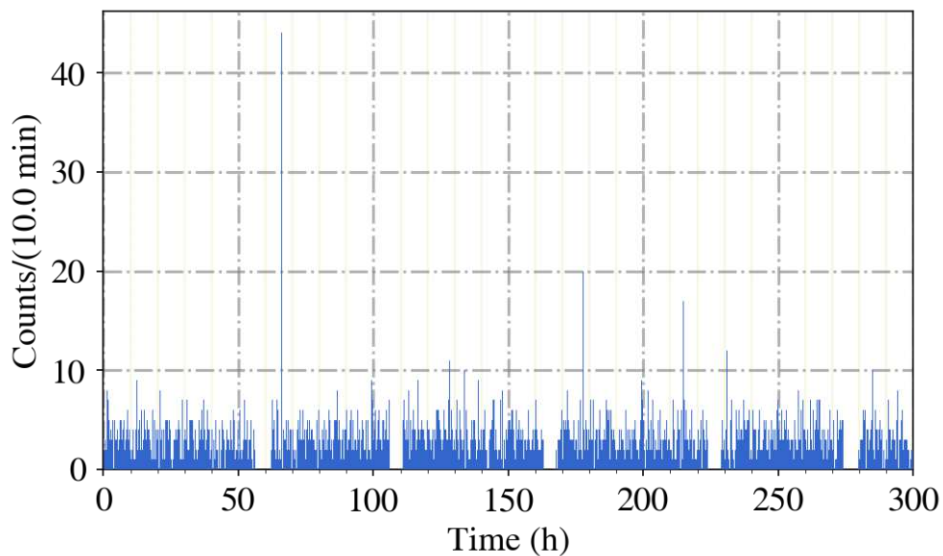

Figure 3.7: Event rate for the Gode3 detector in the first 300h of measurement. The events are combined into bins corresponding to a measurement time of 10 minutes. Phases with no detections correspond to helium refill times. Clearly visible are also bins with heightened event rates, e.g. at 65h, 180h, or 215h, which will be discarded by the rate cut as intervals with unstable behaviour.

The rate cut will thus exclude all intervals of certain length, generally 10 minutes, which deviate from the median by 3 or more standard deviations. As noted in [44], the period of 10min provides a sufficient trade-off between measured counts and bin size. Usually, one also excludes the neighbours of bins exceeding the maximally allowed rate, since instabilities might have also been visible within thos adjacent bins. For the analysis here we have chosen the bin sizes small enough to remedy this leeking of unstable periods into neighbouring bins. For the data taking of Gode3 the rate median amounts to 3 counts per 10 minutes which means that 10 minute intervals with 9 or more counts are excluded. Within this good event rate interval one finds 96,963% of the events measured which amounts to a total of 11269 events.

In contrast to intervals of high event rates one also aims to extract periods of unusually high or low pulse heights for periodically injected control pulses. Since these pulses monitor the operating point of the detector by heating the TES to its normal conducting regime [34] one wishes to exclude time periods where subsequent control pulses have amplitudes surpassing or falling short of the mean of the - as Gaussian modelled - distribution by 3σ. The stability cut is, contrary to the rate cut, applied to all channels of the observed detector. Mean μ and standard deviation σ of the control pulse heights found for the stability cut by modelling Gaussian noise within a reasonable pulse height interval are listed in Tab. 3.5. The interval is chosen such that the calculation excludes pulse heights which constitute definite outliers, i.e. close to zero amplitude etc., since these pulses would enlarge the standard deviation by quiet a bit.

	μ [V]	σ [mV]	Bounds [V]	Good Events	Good Event Ratio
Carrier	2.846	14.445	2-3	7714	66.374%
Ring	9.114	16.886	9-10	11622	100%
Light	1.33	22.718	1-2	10565	90.905%

Table 3.5: Stability cuts on all three channels of the Gode3 detector. Listed are the mean (μ) and standard deviation (σ) of the assumed Gauss distributed control pulse heights set within specific bounds of stable operation. Further the surviving "good" events as well as the ratio compared to all events in the triggered stream data set are listed. In total only 45.81% of all initial triggered events survive this cut.

The sum of both the rate and stability cuts leaves us with a total of 5324 particle events within "good" measurement periods. This corresponds to 45.81% of all the triggered events. The results of these cuts are clearly dominated by large periods of time where the control pulses in the carrier channel show unstable behaviour, necessitating the definition of the bounds above. Closer investigation shows that the control pulses display extremely low amplitudes ($\sim 40\mu V$) during periods of the first 300h of measurement. This might easily correspond to noise which would mean that during these periods no control pulses where injected, while events and test pulses were measured. In Fig. 3.8 the pulse heights of the control pulses in the carrier channel are plotted for the first 300h.
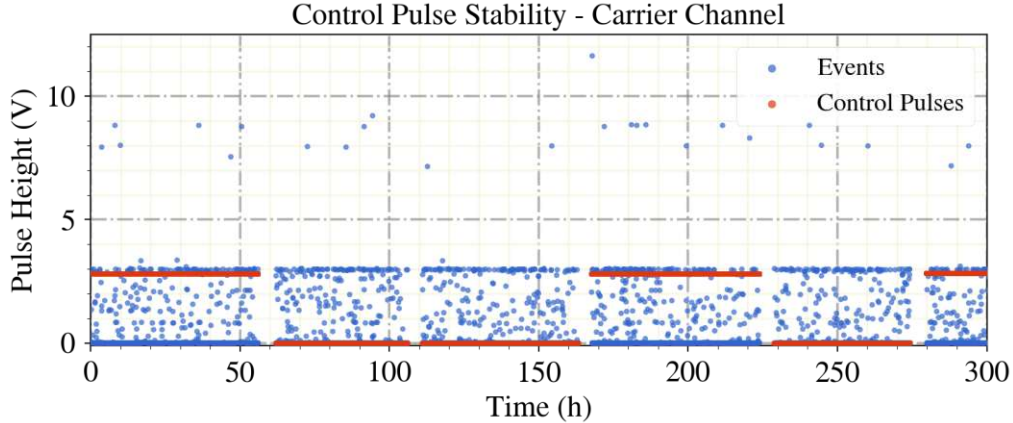
Figure 3.8: Pulse heights of the control pulses measured during the first 300h of measurement in Run-36. Visible are extended periods during which practically no amplitude is measured, while events are simultaneously detected, hinting at problems in the control pulse injection.

What is interesting though is that this phenomenon is either not visible, like in the ring channel, or does not line up temporally with the carrier channel as observed in the light channel. This must all lead to the conclusion that during these periods there must have been issues with the control pulse injection and not the detector module. Therefore, one must not exclude the events during this period and thus ignore the stability cut, since "good" events might fall victim to these detector-independent issues. A cross-check with triggered events and test pulses shows that during these periods the detector was not offline and that indeed events were measured with pulse heights not equal to zero. This analysis will hence proceed only with events surviving the rate cut.

3.2.2 Quality Cuts and Cut Efficiency

Keeping in mind that the rate cut above only cuts a fraction of all artifacts, further analysis of the raw data is essential for the creation of energy spectra and Light Yield plots. The previous section 3.1.2 already introduced the most common artifacts within the data set as well as the notion of quality cuts within Cait. Upon this foundation - and the knowledge of two distinct event classes in the phonon detector - this work's analysis will further expand and build upon.

Slope Cut

As a first measure for removing artifacts, or more particularly decaying or rising baselines and Squid Jumps, a slope cut is applied to the data set. The slope is defined

as the difference between the average of the first and last 500 sample amplitudes of the record window. Clean events without influence from thermal decay processes etc. will have a slope close to zero, given that the record window was defined as large enough. Plotting a histogram of the slope of all events reveals a nearly Gaussian distribution. The cuts are set symmetrically around zero outside of the bulk of events, making sure that as little valid events as possible are cut. For the phonon channels the slope was restricted to the interval of $[-0.2, 0.2]$ V, while for the light channel $[-0.1, 0.1]$ V was picked.

Optimum Filter Pulse Height Cuts

After the slope cut another reasonable strategy for removal of artifacts - especially any noise baselines which might still be left - is the threshold cut. Here any baselines which have not been excluded by the triggering will be extracted by demanding that their OF pulse height surpasses the defined trigger threshold of only the carrier channel, as this has the best sensitivity for the target, the absorber crystal.
Even though this cut might seem superfluous since it in principal acts extremely similarly as the triggering procedure, it alone cuts from 11269 events surviving the rate cut almost 8000 events from the data set. This is a consequence of the way the data stream was triggered: even though all channels were triggered, we here exclude ring and light only triggers. After rate and threshold cut there are only 3297 events left for the analysis, when taking the above mentioned Slope cut into account the number shrinks to 2480 events.
Another cut on the OF pulse height is given by Fig. 3.2a. After a pulse height of about 2.5V it becomes apparent that saturation starts to set in for the carrier channel. Therefore, events with OF pulse heights above 2.5V are counted as oversaturated pulses and thus excluded from the analysis, leaving 1837 baselines for the remaining analysis (15.81% of triggered baselines).

Onset Cut

Pulses saved via the triggering script will usually have onset times of approximately 0ms, since the maximum of the pulse will lie at 0ms. In some cases where event pile-ups and similar artifacts are measured, the onset of the event will not line up with above stated condition. Here, the maximum of the matched filter technique utilised during the triggering will be set to 0ms, while the onset calculated with the Cait library might show stark discrepancies between the onset time of a pulse. In Cait the onset of a pulse is simply defined as the last sample before 20% of the maximum amplitude of the record window are reached. By setting a tight window around 0ms, any artifacts like late triggers and pile-ups will be removed from the data set. For the purposes of the stream analysis again the same Onset cut as used for the generation of the SEVs was used ($t_0 \in [-20, 20]$ ms), leaving 1162 events for analysis.

Region Cut

In some cases, examinations of the main parameter plots yield information about the data beyond knowledge of multiple standard events, as was the case in Sec. 3.1.2. Clear detector issues, such as the instability of the ring detector towards higher amplitudes also manifest themselves in main parameter plots. For instance, one is able to observe such a cluster of unstable pulses in the rise time plot 3.2b of the carrier channel clearly deviating from the rest of the pulses.

The described cluster visible in the hardware triggered data's rise time plots actually turns out to be purely constituted of artifacts, since the triggering procedure was able to filter out those events. Here our first intuition seemed fruitful which would lead to the conclusion that - if any such clusters still exist - it would appear to be wise to cut these events from the data, ensuring clean baselines for the spectral analysis.

Further investigation of the stream data showed two interesting cases in which such clustering of potential artifacts did take place. The first is a type of plot already discussed previously, namely the decay time versus pulse height plot pictured in Fig. 3.9. Starting at about 7.5V there appears a third band below 2.5ms of decay time. This "extra" band seems to originate mainly from the ring channels inability to measure higher-energy pulses, leading to either distortion, see Fig. 2.6, or oversaturation effects common in CRESST detectors. This phenomenon can also be explored from another angle, pictured in Fig. 3.10, by plotting the pulse heights measured in the carrier and ring channels simultaneously.
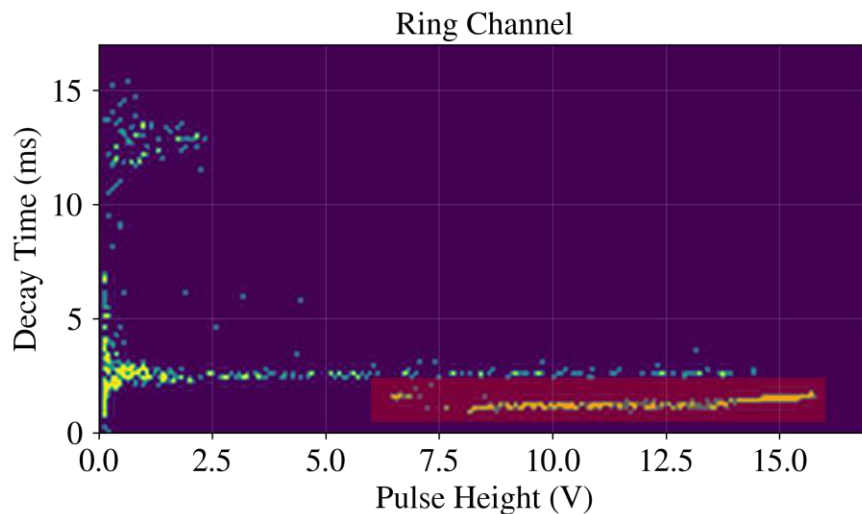

Figure 3.9: Decay time versus pulse height of the already reduced event data set in the ring channel. The red-shaded area shows clear clustering which might be attributed to instabilities in the ring channel, see Fig. 2.6, further manifesting in even shorter decay times, as typical of such artifacts.
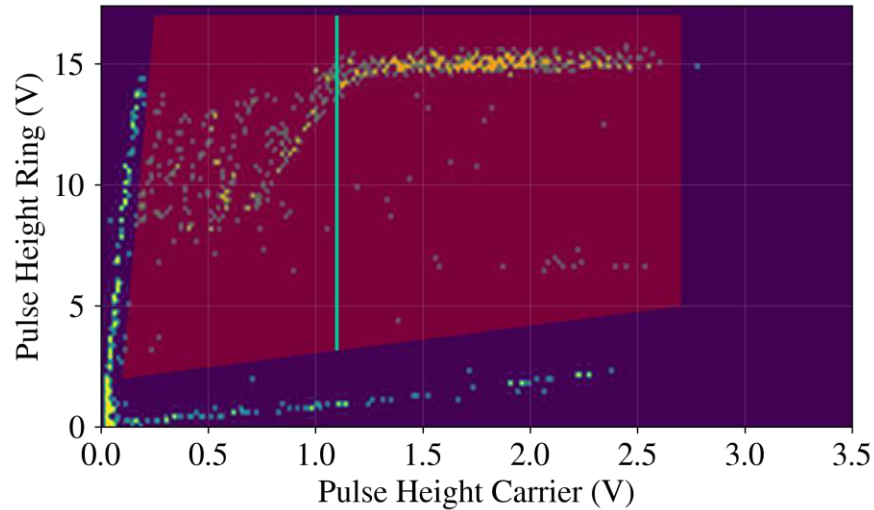
Figure 3.10: Ratio of carrier versus ring pulse heights. The shaded area represents for one the above mentioned artifacts due to the ring instability below 1.1V carrier pulse height (turqoise vertical line in plot) - also presented as such in Ref. [42] - but also shows an area clearly depicting "normal" oversaturation of events above 1.1V.

The graph shows a projection of the extra band visible in Fig. 3.9 onto the carrier channel pulse heights. Here one sees that this band is actually split into two regimes, one which shows erratic jumps in the ring pulse heights left of the turqoise line at 1.1V and another right of this vertical threshold. As mentioned in [42], these events below 1.1V carrier pulse height actually correspond for a large part to those unstable events briefly discussed in Sec. 2.2.1, a fact made plausible by the erratic jumps in ring pulse height similar to the pulses shape. Above 1.1V the picture becomes much more stable where one can assume that oversaturation starts to kick in, capping the maximum pulse height measured in the detector. Given this knowledge we might now make steps into excluding all events within the shaded regions, here called region cut, in Fig. 3.9, and Fig. 3.10.

The cut removing the events given by the third band in Fig. 3.9 is therefore given by excluding all events within the shaded region, i.e. $[6, 16]$ V and additionally $[0, 2.4]$ ms. Applying this cut removes over half of all events left and leaving 471 baselines from the stream data set (4.02% of intial stream data set).

Applying the second region cut though leads to a myriad of issues in further analysis. For one, utilising the cut bounded by the shaded region in Fig. 3.10 removes no further events, thus having only limited impact. Secondly, excluding events in above mentioned region leads to inconsistent cut-efficiencies for high energies in the later stages of the analysis, see Sec. 6.9 in [34]. These issues lead to the fact the region for the pulse height ratio plot will not be implemented as a cut.

Absorber vs. Ring Event Discrimination

Finally, the suriving events will now be split into two sets of event types, i.e. separated into absorber and ring events. In accordance with the previously performed analysis for the SEV generation, the event types will mainly be discriminated by the velocity of the pulse's rise to maximum or its decay.

As evidenced by Fig. 3.9, the two-band stucture observed in the hardware data is also visible in the already cut stream data set. Hence, we will perform rise and decay time cuts based on the corresponding plots. For the rise time, all events below 3ms in the carrier and above 2ms in the ring are classified as absorber hits while all events with the features inversed are labelled ring hits. Additionally, the decay time of pulses is restrained to an interval of 12-20ms in the carrier and above 10ms in the ring for absorber hits with events larger than 20ms in the carrier and smaller than 10ms in the ring being ring events. Following these cuts, there remain 65 absorber and 114 ring events (in sum 179 events, i.e. 1.76% of the initial data set) for the analysis of spectra and the Light Yield plots. Most of the remaining events exhibit the correct properties relating to its event label. It is evident though that pulses in the region of low amplitudes the discrimination of events becomes practically impossible. Therefore, we will turn to Machine Learning techniques in the following chapter 4 to circumvent these issues with regard to the labelling of the low-energy region, i.e. the region especially interesting for the investigation of possible sources of the LEE. The resulting classification of events by the quality cuts and the assoicated energy spectra will be presented following the energy calibration in Sec. 3.2.3.

Cut Efficiencies

All cuts applied in this analysis will involuntarily also eliminate good events from the data due to their intrinsic "crude" nature of simply extracting regions and intervals. A simple tool to verify one's choice of quality cuts is to apply them to a simulated data set and counting the number of events surviving the cut within an energy bin.

To simulate such events one superimposes the standard event of incident potential Dark Matter pulses with empty noise baselines. The simulated set of events is then free from most artifacts and pollution, allowing for precise determination of the number of removed valid events [34, 44]. To also measure the energy dependence of the cuts, pulse shapes of varying energies are injected into the data set by scaling the pulse height of the used standard event to match the desired energy. The method of converting pulse height amplitudes of pulses to their respective energies is described in the following section, Sec. 3.2.3. For the purposes of this work 10000 artificial events were generated on the basis of the absorber crystal SEV with pulse height scaling below the carrier channel's maximum measured pulse height of 3V. The resulting survival probabilities, calculated by counting surviving events within an energy bin (in sum 2000 bins over the whole energy spectrum), are pictured in Fig. 3.11.

Here, the absorber SEV was used, since the crystal represents the actual target in the
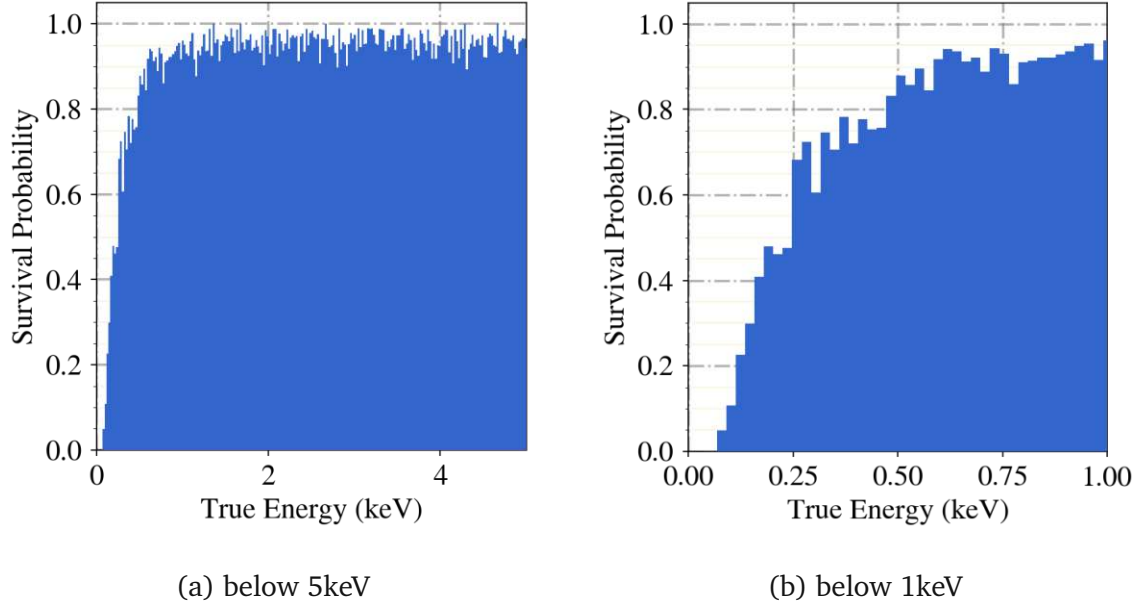
(a) below 5keV   (b) below 1keV

Figure 3.11: Absorber cut efficiencies or survival probabilities below 5keV (left) and 1keV (right) for the carrier channel simulation data set. The data set extends to energies of ca. 45keV and is divided into 2000 bins.

experiment with the ring detector acting as a veto detector, decoupling the absorber from any external influence, see Sec. 2.2. Thus, we will only consider the carrier detector as the primary detector for analysis.

The cuts give a similar picture seen in the analysis of [34, 44] with high survival rates near 100% above 750eV. Energy-dependence stems from the fact that for low energies the bands in the rise and decay time versus pulse height plots widen as lower pulse amplitudes are reached. This is due to growing influence from the noise in the signal and hence also being prominently present in Fig. 3.11b.

Moreover, we introduce here what is called the *analysis threshold*. This threshold is given by the energy at which the survival probabilities begin to exceed a value of 0.5, in this case this would correspond to an energy of approximately 200eV. Below this energy even valid events could be extracted from the data by the applied quality cuts. Thus, yielding an in-depth analysis of events at lower recoil energies practically impossible via quality cuts, since they could be removed.

3.2.3 Converting Amplitudes to Recoil Energies

In order to obtain meaningful results it is of utmost importance to convert the incident pulse heights into recoil energies. This goal is achieved by a two-step process where first, the reconstructed pulse heights in the triggered data set are converted into test pulse equivalent amplitudes (TPAs), with a subsequent determination of the conversion factor between pulse heights and energies, called the *CPE factor*. The usage of TPAs is necessary since they account for any temporary deviations within the TES' response due to heating or other mechanisms [44]. Most importantly, the amplitudes of test pulses are linear with respect to their energies, while this is not necessarily the case for event pulse heights. The explanations in this thesis are kept short and only outline the main ideas of the energy calibration and conversion process. A more in-depth overview of a larger variety of approaches can be found in either of [34, 44]. Test pulses are generally injected at known, constant time intervals at fixed amplitudes. For Gode3 the time between two test pulses of the same amplitudes amounts to roughly 0.5h, where in sum twelve distinct pulse heights - in Volts: (0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) - were used. The injected amplitudes can then be reconstructed via one of the methods presented in Sec. 3.1 as long as the test pulse heights scale in the same way as the reconstructed pulse heights [44]. This would not be the case for oversaturated events which would necessitate the utilisation of a truncated fit, presented for instance in [5, 44].

Before any analysis on the test pulses can be done, previous cleaning of any artifacts within the test pulse data set is required. Here, only the previously presented stability cut is applied with bounds of 0.25V around the mean of the event distribution, cutting any possible artifacts. The stability cut removes 0.12% of all test pulses from the Gode3 data.

The reduced data set is then split into smaller sets separated by half an hour of idle operation where the helium tanks are refilled and finally interpolated, in this work a linear interpolation is used, for each discreet test pulse amplitude. Thus, every injected test pulse is directly and time-dependently mapped to a reconstructed pulse height measured by the detector. The conversion of test pulse amplitudes to reconstructed pulse heights can again be accomplished by any of the above mentioned procedures - here the OF pulse heights were employed.

With information of the conversion factor between the pulse height of a test pulse and its true injected amplitude at any time of data acquisition, it is now possible to select the exact time of event detection and convert the events pulse height into its TPA. The so-called *transfer function* [44], i.e. the function converting the pulse's reconstructed height to TPA, is obtained again via an interpolation procedure. The choice of a sensible interpolation method is discussed in detail in [44], with a polynomial fit of third degree proving to provide the best results in the context of this work. The time-interpolation of the test pulses and the polynomial fit of the transfer function at a specific time stamp are plotted in Fig. 3.12.

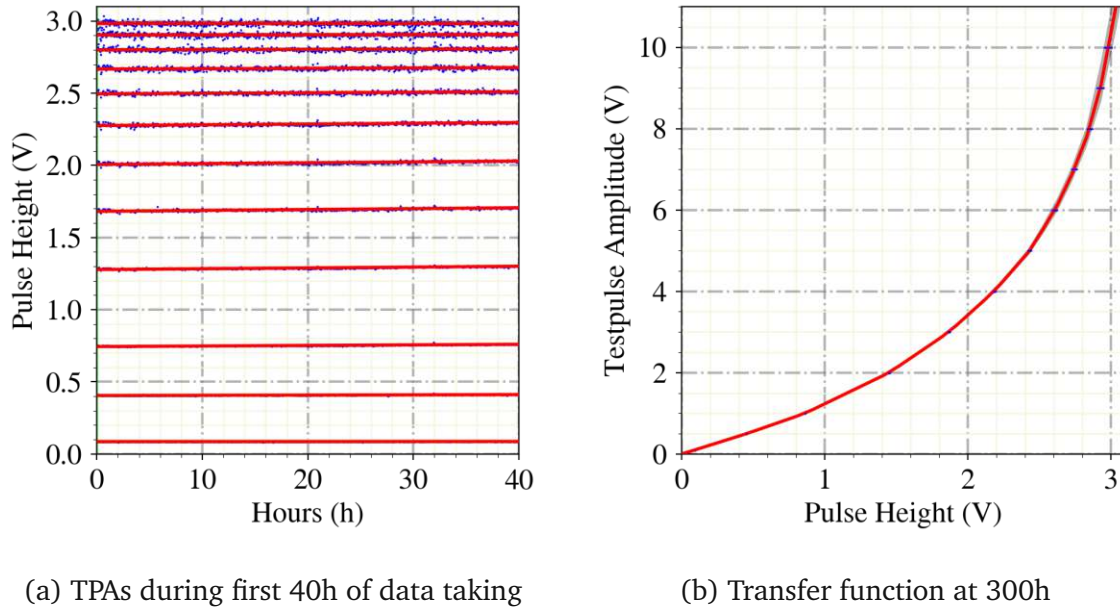(a) TPAs during first 40h of data taking   (b) Transfer function at 300h

Figure 3.12: Fig. 3.12a displays the linear interpolation of the measured TPAs during the first 40h of data acquisition with some visible deviations from the linear model for higher energies. These differences are also observable in the final transfer function from pulse heights to test pulse equivalent amplitudes in Fig. 3.12b obtained at the 300h mark.

Every CRESST run is preceded by calibration measurements using the 122keV γ-rays from a Co^{57} source in close proximity to the detector. Produced by Co^{57} decaying into Fe^{57}, the resulting γ-rays generate observable peaks in the pulse height spectrum of the detector in question. For one, the peak of the Co^{57}-γ but also the $K_{\alpha,1}$ and $K_{\beta,1}$ transitions, induced by the Co-decay in the present tungsten, are visible [44]. These peaks give insight into the detector's response in amplitude for the already known energy of the γ-radiation, and thus allowing the conversion of test pulse amplitudes to recoil energy, i.e. the CPE factor. The pulse heights of the peaks are reconstructed via the standard event fit [44] and subsequently converted into test pulse equivalent amplitudes, again accounting for any drifts in the detecor's response. Since this procedure was already done for the Gode3 detector, the values for the CPE-factor were taken from [42], see Tab. 2.1, and applied to the test pulse equivalent amplitudes acquired by the previously described fitting procedures. The resulting values for V-to-keV conversion factor (C), baseline resolution and trigger threshold are listed in the following table, Tab. 3.6.

The conversion to recoil energies for baseline resolution and trigger threshold is an approximation by necessity, since the detector's response is completely dependent upon the time of interaction within the detector, ergo information which is inaccessible to us beforehand. Therefore, the conversion factor C is calculated by scaling the

	C (keV/V)	Resolution OF fit (eV)	Threshold (eV)
Carrier	9.09	11.81	90.3
Ring	16.83	150.76	935.2
Light	20.12	101.6	779.3

Table 3.6: V-to-keV conversion factor (C), Baseline resolutions and energy thresholds for ring, carrier and light detector.

CPE factor with the ratio of the mean of one specific test pulse amplitude TPA^{j} by its corresponding reconstructed pulse height,

$$C = \mathrm{CPE} \cdot \sum_{i} \frac{1}{\mathrm{N}} \cdot \frac{\mathrm{TPA}_{i}^{j}}{\mathrm{PH}(\mathrm{TPA}_{i}^{j})}. \tag{3.8}$$

For this analysis, the lowest injected pulse height of 0.1V was chosen ensuring that the detector is within its linear region. Compared to the results shown in [42] here similar values are obtained, differing by a maximum of 12% in the light channel, affirming the validity of this work's analysis.

As a final note, a second possibility of obtaining the CPE factor is mentioned here. Many detector modules, with Gode3 not belonging to this list, have an iron-source inbuilt in the detector's casing. This enables the extraction of the CPE factor via analysis of the stream data by again transforming all pulse heights in the data into test pulse equivalent amplitudes, and finally finding the CPE factor from the peak present in the spectrum.

3.2.4 Spectrum and Light Yield Plots

We now conclude this chapter by producing the detector's Light Yield plots, Fig. 3.13, and energy spectra, Fig. 3.14. It is important to mention that all spectra are calculated as *absorber equivalent* and thus only valid for recoils within the absorber crystal, denoted by the unit [keV$_{\mathrm{ae}}$].

The energy spectra, Fig. 3.14, make evident that ring events are the dominant type of interaction for energies below 1keV$_{\mathrm{ae}}$. A fact which is to be expected for the carrier channel, since events in the ring naturally have lower amplitudes in the absorber, it is even more interesting that it is also clearly visible in the ring channel. Not only do ring events dominate high energies in the ring channel, but our quality cuts also leave only ring events, though extremely few, below 1keV$_{\mathrm{ae}}$ in this channel. This is further evidenced by the data in the light channel, leading to the assumption that only the detector ring experiences low-energy events. Possible reasons for this specific data distribution will be explored in the following and also chapter 4.

The Light Yield plots 3.13 show a similar situation as the energy spectra, namely that

the low-energy region is completely dominated by ring events. The Light Yield, as explained in Sec. 2.2, is simply the ratio of the deposited energy in the Light and observed phonon channel. Thus, we would expect electromagnetic interactions to have a Light Yield of around one, highlighted in the plots by a yellow line, and possible dark matter interaction to be close to a Light Yield of zero. Such events do exist in our cut data set especially in the carrier channel.
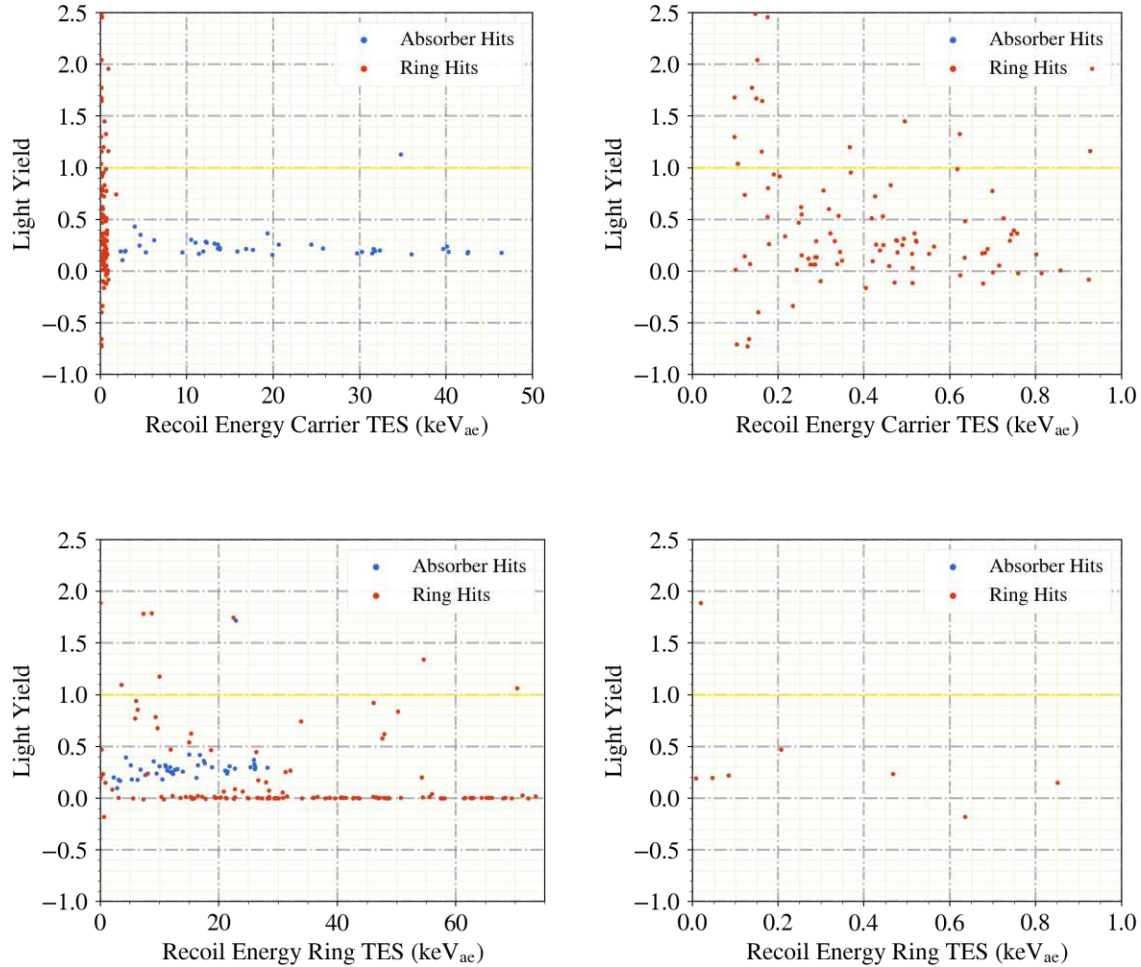
Figure 3.13: Light Yield plots for the measured absorber equivalent energies in carrier (top row) and ring channel (bottom row) split into absorber (blue) and ring hits (red). The left column displays the whole recoil energy range of the detector, while the right is a zoomed-in picture below $1\mathrm{keV_{ae}}$. Moreover, the yellow line indicates the e^-/γ-band. Here, the prevalence of ring hits in the lower energy range is again present. All energies are listed as absorber equivalent recoils, as mentioned above.
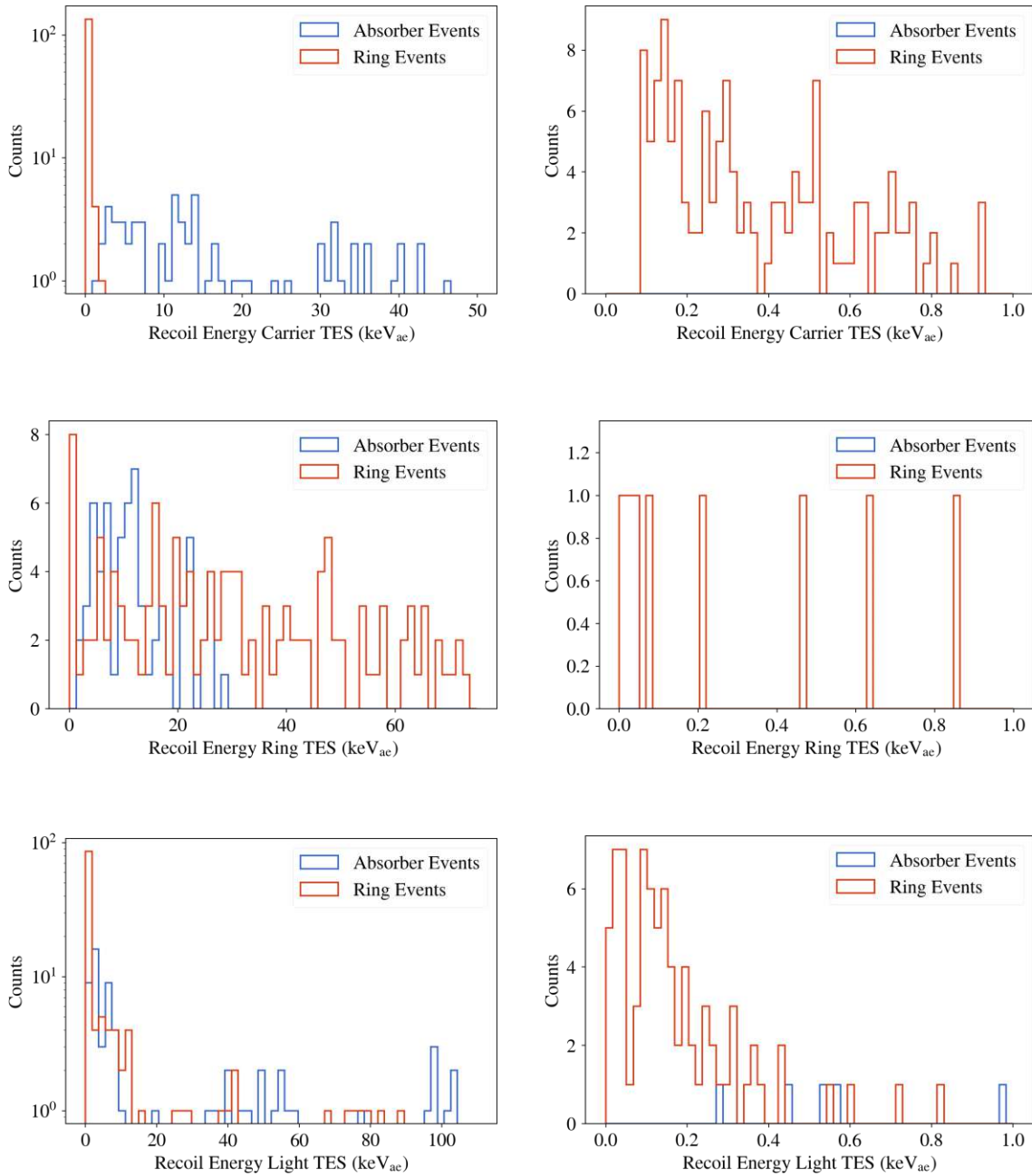
Figure 3.14: Recoil energy spectra of carrier (top row), ring (middle row), and light channels (bottom row) measured in keV$_\mathrm{ae}$. Shown are plots over the whole energy spectrum (left column) and a zoom into the recoil energy region below 1keV$_\mathrm{ae}$ (right column). All graphs are binned into 60 bins. Further, the counts are seprated into absorber (blue) and ring hits (red) displaying far more recorded ring hits in lower energies according to the quality cuts performed in Sec. 3.2.2.

The results presented in this chapter show hints of the origin of the Low-Energy Excess. In the following we will try our hand at another technique of event selection which may give an even better insight into the events present at lower energies, since - as the plots of rise/decay time show - quality cuts are viable only upto a specific energy where the absorber and ring bands are clearly differentiable.

Chapter 4

Low-Energy Event Classification via CNNs

The field of Machine Learning, though not novel per se, as first mentions of the term arise as early as 1959 by A. L. Samuel [50], has solidified itself as one of the most popular and quintessential sub-fields of computer science in the first decades of the 21st century. The overcoming of limitations in CPU power of the early days and an increase in available data has moved Machine Learning to the front of modern computational methods. Be it voice or image recognition [51–53] or purely mathematical applications such as solvers for partial differential equations [54, 55], more and more firms and research institutes are finding practical usages for Machine Learning.

Especially Neural Networks and its many different implementations have risen in popularity in recent years. The term *Neural Network* has its roots in human biology and was first used to describe the way synapses in our brains are interconnected, leading to our ability to learn. Hence, in literature one often finds the differentiation between *Biological Neural Networks* (BNN), i.e. the networks inside brains of living beings, and *Artificial Neural Networks* (ANN), special computational systems connecting neurons or *nodes*. With early mentions of Neural Networks as machines for computation by W. McCulloch and W. Pitts in 1943 [56], the first and most basic form of Neural Network was developed by F. Rosenblatt in 1958 [57] in the form of a *Perceptron*, a binary classification network of only one layer of input nodes and one single output [58].

This chapter will utilise one of the more common and nowadays popular variants of Neural Nets called *Convolutional Neural Networks* (CNN), which is especially capable and most often applied in image processing problems [59].

The following analysis in this chapter is mainly motivated by one thing: Finding an efficient and most importantly precise way to classify events in the low-energy region. More specifically, the knowledge of which events are measured in the Low-Energy excess interval is of great interest to the CRESST experiment. The conventional analysis, though being a competent tool for most other detectors and over a large part of the energy spectrum, faces obstacles when it comes to the discrimination of events with low energies. This is true in particular for Gode3, as noted in [42], where ab-

sorber and ring bands begin to blend and the to date "standard" method of quality cuts begins to show its weaknesses. One can only differentiate between absorber and ring hits with "rough" rise/decay time quality cuts, with "rough" being used here to indicate that they do not provide a clear separation of the events for lower energies. This implicates inevitably that one cannot probe the origin of the LEE with any certainty, and that, since all applied quality cuts were either energy-independent or in the high-energy region, low energies most interesting for Dark Matter analysis are not analysed. This inability to analyse the low-energy region via quality cuts is only further amplified by the analysis threshold of 200eV determined in Sec. 3.2.2, which fixes the lower limits of the analysis. Quality cuts, futhermore, ignore one of the most interesting facts about the Gode3: its inherently large dynamic range. In the previous analysis, Sec. 3.2.2, we have removed all events with any semblance of being artifacts, mostly ignoring the structure of the event in the parallel phonon detector, e.g. via the Region Cut 3.2.2. This for instance removes all events which are clearly artifacts in the ring channel, due to the detector's high instability, but are in fact completely valid pulses in the carrier channel.

This work will thus build upon previous work done using Machine Learning with CRESST data, in [4–7, 60], but additionally build upon the work done in event discrimination, e.g. discrimination between carrier and absorber hits in [5], and combine both in hopes of approaching a valid discrimination of events in the low-energy region. This chapter is structured as follows: First, a short introductory overview for Machine Learning and Convolutional Neural Networks is given in section 4.1. Then, the data are prepared, in Sec. 4.2, and trained using the CNN, Sec. 4.3. The final section of this chapter will focus on improvement strategies, enhancing the predictive power of the Machine Learning model at hand. Note that most plots depict the results in the carrier channel with any exceptions to this highlighted separately.

4.1 Convolutional Neural Networks

The field of Machine Learning can in general be divided broadly into three subgroups. These groups are defined by the way the models are trained and in which way they learn: The three different methods are: Reinforcement, Unsupervised, and Supervised Learning [59]. Supervised Learning trains via pre-labelled data sets and is a valid option when one has labelled data at one's disposal. Unsupervised Learning on the other hand is used when trying to split and separate data sets and learn patterns in the data without the need for labeling. This includes methods for clustering, dimensionality reduction or generative modelling [59]. Reinforcement learning is applied by demanding of an intelligent agent to change its behaviour in such a way which maximises some given reward, e.g. winning in a game. Both of the latter Machine Learning methods were not employed in this thesis. An introduction to Reinforcement Learning can be found in [62], while one can find a general discussion of Unsupervised Learning in chapter three of [63]. Unsupervised Learning techniques have also

been applied in the context of CRESST data analysis in [5, 6]. The model trained in this work made use of a labelled data set with which the Convolutional Neural Network was tained, i.e. learning under supervision. In the following only a rough overview of the way CNNs function will be given.

CNNs differ from other standard deep learning networks, such as a Multilayer Perceptron for example, by the fact that the hidden layers within the CNN are so-called *convolutional layers*. Convolutions of some input allow the network to perform accurate classifications of two-dimensional objects while simultaneously also being invariant towards translations, scaling and other forms of distortion [58, 59]. The method suits itself especially for grid-like data with sampling taking place at regular intervals. This would include time-series data - like CRESST event windows - in a 1D setting, but also pictures or volumetric data - like CT scans - in two or three dimensions [64]. CNNs will thus prove to be a valuable tool for CRESST event classification in the following sections.

An examplary schematic of the way a CNN would look like is depicted in Fig. 4.1. One can find an interactive visualization of how a CNN, which classifies hand written digits, functions provided in [65].
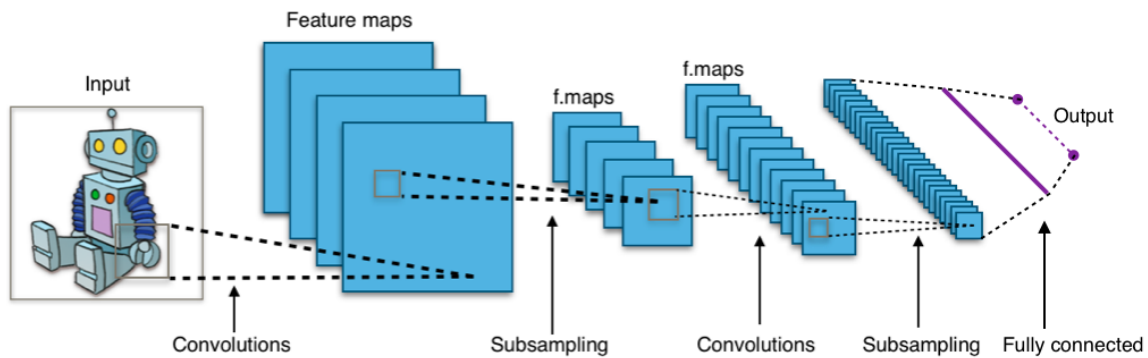
Figure 4.1: Schematic of a typical CNN architecture. Visible are multiple convolutional layers which transform the input, here the picture of a robot, into a fully connected layer at the end of the network before it is finally classified via the output layer. Figure taken from [66].

As is visible in the figure above, we see that CNNs are in general made up of alternating *convolution* and *subsampling* operations before finally adding the 2D-slices to a number of *fully connected layers* providing the output of the classification problem for instance. Convolutions are achieved by shifting multiple different *kernels* or *filters* of small size over the picture matrix with a given stride length S, the number by how many pixels the shifts are performed. Subsequently, one performs matrix multiplications of the regions values with the kernel at each step [67]. Different kernels might serve different detection methods, i.e. detecting edges, faces, etc. within pictures. The detection of these features is controlled by the weights within the kernel matrix,

which are shared between every neuron of this kernel's feature map. This stands in contrast to the way weights function in typical ANNs, where every neuron, in case of a 2D picture this corresponds to a pixel, has its own associated weight. This so-called *weight sharing* leads to positive effects within the network. Not only is the number of free parameters lowered, decreasing the necessary memory space, but the model's capacity for generalisation is also increased [58].

Typical kernel sizes W of 3x3, 5x5 or 7x7 are used in image recognition for convolving single parts of the sample matrix with the filter [5]. In some cases the definition of zero-padding pixels P around the output is convenient in order to control the size of the result [67]. For given input size W of some quadratic matrix for instance, one finds the output at the subsequent layer $l + 1$ to be,

$$W_{l+1} = \frac{W_l - F_l + 2P_l}{S_l} + 1 \quad .$$ (4.1)

Following each convolution step one most often finds a non-linearization of each pixel's value via an activation function, most often a ReLU-function, i.e. $\max(0, x)$ [67], concluding the operations of the convolutional layer.

The resulting feature maps are then often downsampled or subsampled via a pooling operation, leading to a reduced sensitivity in shifts and other distortions in the input. This means, that the exact location of the robot in Fig. 4.1 would be less important for an accurate classification of the object [64]. Subsampling is often accomplished by again scanning a matrix of small size, i.e. 2x2, across the feature maps and aggregating the values within the matrix to one pixel, so-called *pooling*. Pooling can be done in many different ways, with max pooling - the largest value of the pooling matrix is chosen as the new neuron - being one of the most popular choices [64].

After a few convolution and subsampling layers all neurons are connected to a fully connected layer, akin to a typical hidden layer of a standard Neural Network. The fully connected layer can be the output layer of the CNN, in which case the number of neurons has to match the number of classification classes. For the purposes of this work, we are only interested in binary classification, i.e. two output neurons. The network then learns, like most other Neural Networks, via backpropagation, an algorithm which adjusts the weights of the neurons in a backwards manner, minimizing the error compared to the labelled input. A more detailed description of the algorithm can be found in any of the Refs. [5, 7, 58, 59, 67, 68]. We now turn to the next section which will deal with the preparation and generation of a labelled data set fit for training a CNN.

4.2 Generation of the Training Data Set

The first step in any supervised Machine Learning technique is the preparation and potential cleaning of the data set. The data in question is then be labelled by hand, giving the Machine Learning model a data set to train on. In our case data cleaning

and preparation has already be done in the previous chapter 3, with an already established discrimination of absorber and ring hits. The process of labelling the data, though necessary for the training of the CNN, is quite impracticable for the data at hand. Not only is labelling incredibly time intensive, the number of surviving "good" events is quite small, which makes extracting a training set of adequate size impossible and putting in question the model's ability to generalise on the rest of the data set. Therefore, it proves wise to opt to pursue another variant of Supervised Learning called *semi-supervised learning*. The data used for the CNNs training is not labelled but instead synthesised via the pulse simulation method described for cut efficiencies in Sec. 3.2.2.

Since one knows which standard event was used as the template for the generation of the simulated pulse, the data are automatically labelled as either absorber or ring event. The method is not only much more time efficient then standard labelling by hand, but also generates arbitrarily many pulses, for the training of the CNN in this work 10000 absorber and ring events each were considered. The only downside is that the CNN is trained with artificially generated pulses which are entirely dependent upon choices made by the analyst, therefore possibly inducing errors in the pulse simulation, e.g. the extraction of SEVs or the generation of channel-correlated pulse heights. In order to replicate real events the pulse heights of simulated events will need to be correlated between the two phonon channels in a specific manner for both the absorber and ring events, as visible in Fig. 3.10. This enables us to train the CNN on not only the carrier channel, but also the ring channel, hence expanding the possibilities of cross-channel analysis and overlapping the results of the models for carrier and ring channel. The confidence in the predictions are improved if both carrier and ring model predict an event to be, for instance, an absorber hit.

Thus, before any events are generated, it is benefitial to first investigate the way absorber and ring hits behave in all three detection channels. This is accomplished by looking at the ratios of pulse heights between the channels and attempting to extract any possible law for the way the pulse heights are correlated between channels. From Fig. 3.10 one can easily observe that the pulse heights outside of the red region, which are eventually cut from the data, behave in an approximately linear manner. When applying the absorber/ring cuts it will become apparent that the left branch with a high slope will correspond to ring events, whereas the lower branch with slope of approximately one coincides with absorber events. This is easily visualized by the way particle hits in the crystals are measured in the detection channels, as explained for the generation of SEVs, Sec. 3.1.2. A hit in the target crystal produces roughly equal amplitudes in the carrier and the ring, while a particle hitting the ring produces a vastly dampened signal in the carrier channel. In the next step we make use of these approximately linear correlations between pulse heights in the phonon channels and fit linear functions to the data of the absorber and ring branches via linear regression, with an explanation of linear regression found in [69]. The resulting linear functions, depicted together with the original data points in Fig. 4.2, will act as a basis for correlated random pulse heights of the simulated events.
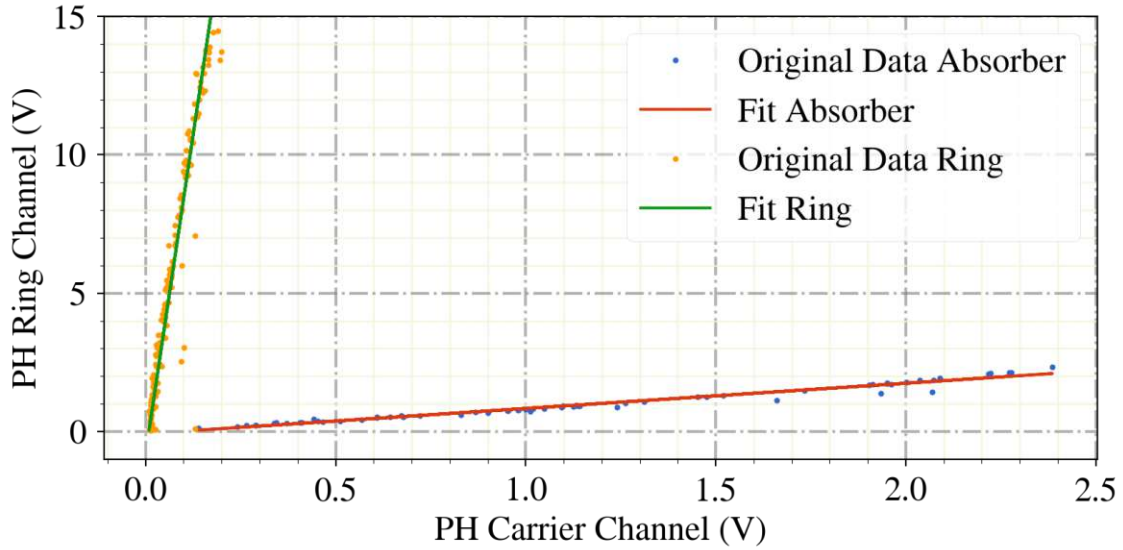
Figure 4.2: Results of the pulse height regression for both absorber and ring events. pulse heights for ring hits are extremely limited in the carrier channel while the regression of the absorber pulses exhibits almost a slope of one.

The correlated pulse heights are now generated by first randomly selecting 10000 pulse heights for the carrier channel from a uniform distribution U. For the random selection the following pulse height intervals for the carrier $\mathrm{PH_C} \in (0, 2.5)$, and light channel $\mathrm{PH_L} \in (0, 3)$ were chosen. The pulse heights for the ring channel are dispensable here, since we choose to generate via the linear model described above. The light channel, until now not discussed, displays no apparent correlation between pulse heights in the phonon channels and the light channel which is to be expected. This means that when one looks at the data of the light channel, there exists no way to ascribe the distribution of absorber or ring hits to any law, e.g. linear functions like in the phonon channels. Hence, pulse heights for the light channel are for the sake of completeness simply picked from the uniform distribution mentioned above. The pulse heights of the light channel are less important anyway, because only the phonon detectors' data was used for discrimination between absorber and ring hits. Additionally, noise ε on the correlated pulse height is generated by random selection from a Gauss distribution, assuming that noise around the linear functions can be modelled as Gaussian, with a standard deviation equal to the σ found in the linear regression. Thus, the generation of correlated pulse heights in the phonon channels is estimated by the following equations,

$$\begin{pmatrix} \mathrm{PH_C} \\ \mathrm{PH_R} \\ \mathrm{PH_L} \end{pmatrix} = \begin{pmatrix} \mathrm{rand}(U(0, 2.5)) \\ k \cdot \mathrm{rand}(U(0, 2.5)) + d + \varepsilon \\ \mathrm{rand}(U(0, 3)) \end{pmatrix} . \tag{4.2}$$

57

The parameters k and d represent the slope and the intercept of the linear regression for either the absorber or ring SEV respectively. For absorber hits one finds $k = 0.80255$, $d = 0.12$ and an R-value of 91.78%, while ring hits display the following values for the linear regression: $k = 0.011$, $d = 0.01$ and an R-value of 94.745%. The pulse heights for the light channel were generated for completeness' sake but are not used in the subequent analysis. Using the relations in Eq. (4.2) 10000 absorber and ring pulse heights and consequently artificial events were generated. These events are used to train a CNN in the following section.

4.3 Training of the Initial Model

This section will demonstrate a time-efficient method of training and applying the CNN presented in the beginning of this chapter to not only differentiate between artifacts and true events but also between two different types of events, the absorber and ring events.

First, we turn to the training of the CNN using the simulated event data set. As mentioned in Sec. 4.2, 10000 absorber and 10000 ring recoils are simulated with pulse heights correlated between the carrier and ring channels. The trained model, as it is defined as a binary classifier, will then be able to discriminate between those absorber and ring events. Using the Python library *Cait*, which itself employs a *PyTorch* backend, a one-dimensional CNN with two convolutional layers is defined, with a kernel of size eight and a stride length of four. Following the convolutional layers two fully-connected layers are inserted, one intermediary and one output layer. The output of the Neural Network is internally assigned to values of either zero for absorber (0) or one for ring (1) predictions. The predictions of the CNN are usually between zero and one which are assigned to an event by counting all predictions below 0.5 as absorber hits and all above 0.5 as ring events. The concrete values of the CNN's output, which are called in this work *prediction probabilities*, are saved alongside the binary classifications of the event types and will prove to also exhibit utility with regards to result improvement in Sec. 4.4.1. In order to save storage space the data are downsampled by a factor of 32, which must not to be confused with the subsampling used in CNNs. Downsampling in Cait is defined by splitting the record window into intervals of the downsampling parameter's size, in our case this would be 32, and averaging the resulting set of data points, thus reducing the size of the used dataset by said parameter. The model further uses 30 epochs during the learning process, which is to mean that the model passes the complete training data set a maximum of 30 times. The solver for weight propagation in the CNN model implemented within Cait utilises the *Adam* optimisation algorithm, an adapted stochastic gradient descent solver by Kingma and Ba (2014) [70]. The learning rate of the Adam solver was chosen to be 10^{-3} in order to reduce overshooting during the weight adjustments. All parameters during the training were chosen via a trial-and-error process and were heavily inspired by the Cait tutorial notebooks. The trained model performs with

an accuracy score of 99.75% on the labelled training data, a high score which also provides confidence in the model's ability to generalise to the actual Gode3 data set. Applying the CNN model on the streamed data set gives the following predictions for the carrier channel depicted in Fig. 4.3.
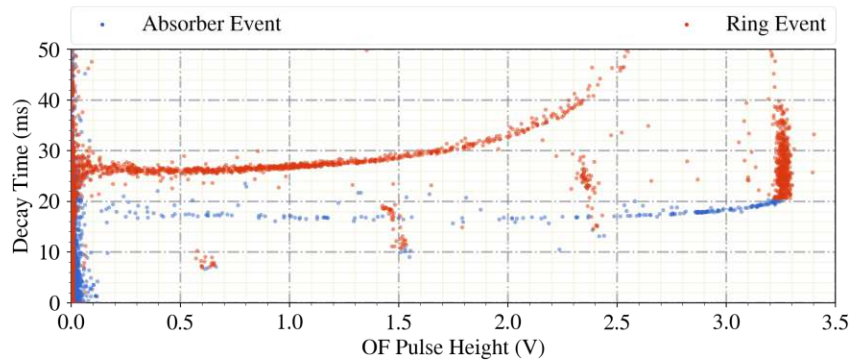

Figure 4.3: Resulting classification by the CNN of the carrier channel data. The CNN model correctly mirrors the separation of absorber (blue) and ring (red) events into two distinct bands depending on the decay time of the pulse. Interestingly, the model predicts all saturated events - indicated by the vertical line at 3.25 OF pulse height - as strictly ring events. This might imply that the model at this stage heavily relies on the decay time of the event which is naturally longer for saturated events.

The predictions by the CNN already display characteristics we might have expected by the previous raw data analysis. The only issue the model shows is that the model only classifies events as either absorber or ring, without accounting for artifacts. Recall here that the three structures in the center of the plot or the saturated events on the upper end of the OF pulse height spectrum represent such artifacts. A first intuition might be to now generate a data set of simulated artifacts and add a third label to the model. In order to introduce these artifacts into the Machine Learning model the classification is approached by using a *pre-trained model*. This other model was trained on a universal training set using many different simulated pulse shapes and is pre-shipped within the *Cait* library [71]. Subsequently, several types of artifacts were extracted and simulated in [60] to discriminate between valid events and artifacts. The model uses the same hyperparameters and posesses the same architecture as defined for the absorber/ring classifier. The results for the artifact/event classifier are depicted in Fig. 4.4. The combined model depicted in Fig. 4.5 is obtained by combining the predictions from both the absorber/ring and additionally the artifact/event classification models. The combined predictions are derived by labelling every event which is both absorber and event as an "absorber event" and a signal which is ring and event as a "ring event". Every pulse which is not labelled as an event, independent of its absorber/ring label, will be called an "artifact" from here on out.
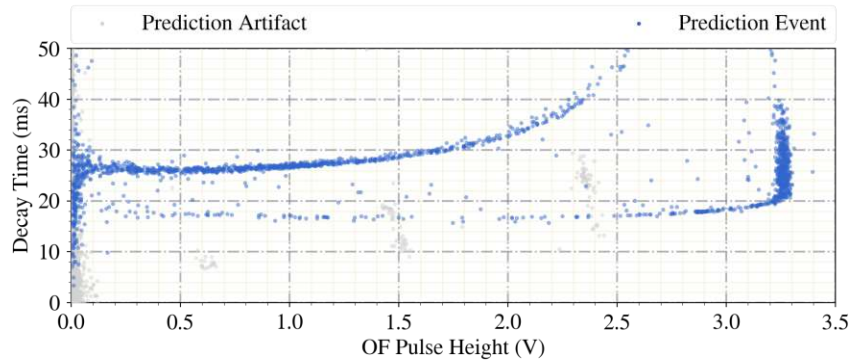
Figure 4.4: Classification of artifacts (grey) and events (blue) by the pre-trained model on the carrier channel.

The predictions of our initial combined model, Fig. 4.5, show parallels between the event discrimination made by the quality cuts in Sec. 3.2.2. That is to say that the model correctly translates the input of the two SEVs into predictions mirroring the two branches in the decay time plots. Interestingly, the model features more absorber events in the low-energy region below 1keV, while the quality cuts leave no absorber hits in the energy spectrum. Also notable is the interplay between predicted absorber and ring events below 1V OF pulse height, with event types leaking into each other's cluster, which is not the case in the quality cuts, since both events were separated at 20ms decay time. This feature of Machine Learning thus allows for a more granular way of event classification unattainable via standard quality cuts.

In opposition to these satisfying results one can also see that the combined model mislabels a few events, mainly as ring hits, below the trigger threshold as valid events. Due to this fact and since the CNN displays different behaviour for weaker particle interactions - though keeping similar structure - we try to check the overall validity of the model. In order to quantify the performance in the low-energy region we go back to the simulation of artificial pulses and again generate 20000 events - 10000 absorber and ring events respectively - below OF pulse heights of 0.2V, corresponding to 200eV recoil energy. This verifies if the model correctly predicts events in the interest region of this work. Applying the CNN to this new set of simulated pulses returns a satisfactory score of 96.96% correct predictions. The evaluated efficiency of the model on this data set is presented in Fig. 4.6. Even though the results by the combination of two different CNN models show fascinating results we are interested in potentially improving the performance of the Machine Learning model. The methods in the following section will attempt and remedy these shortcomings of the CNN applied here.
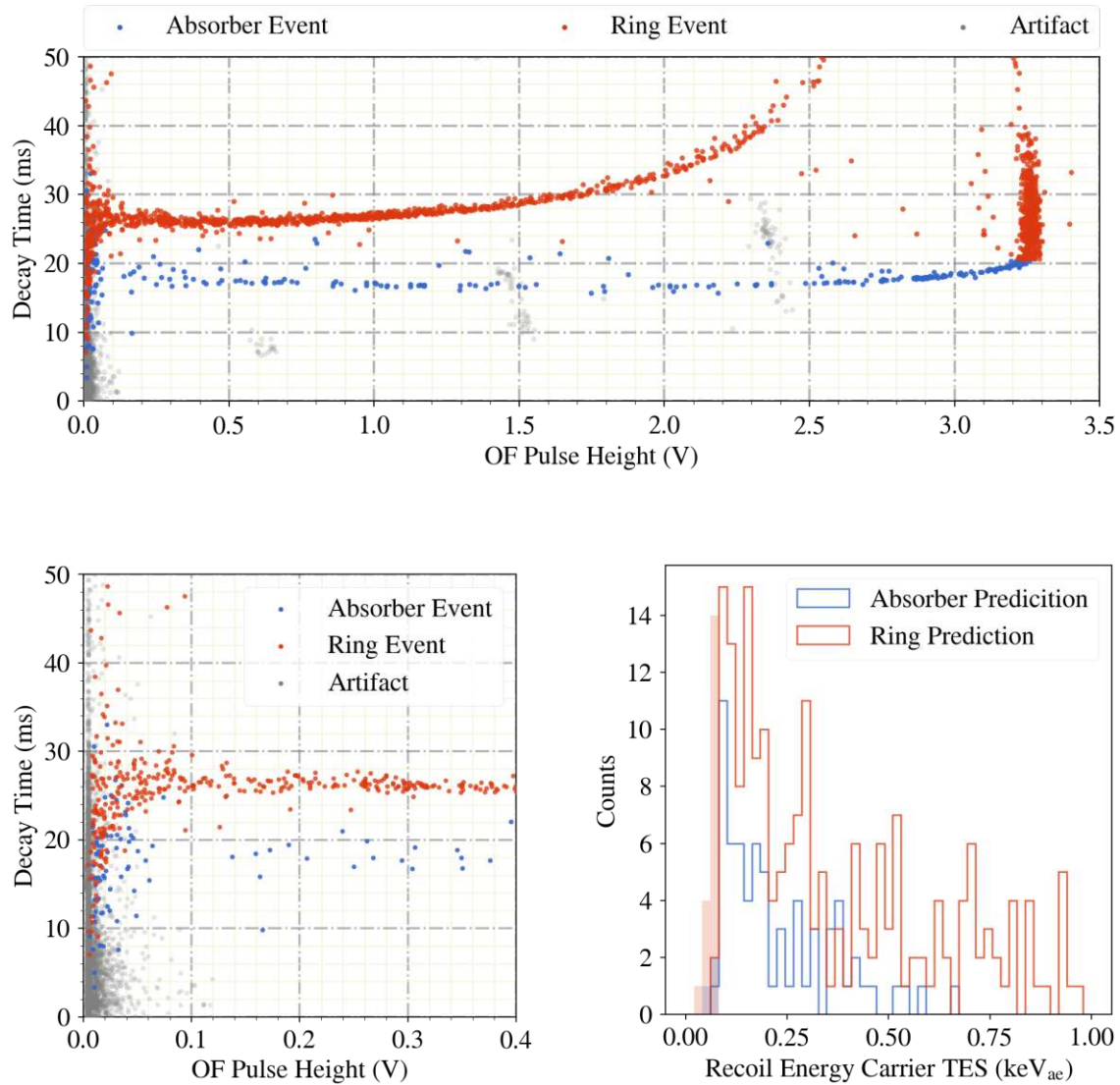
Figure 4.5: The top plot shows the results of the initial, combined CNN model which differentiates between absorber events (blue), ring events (red), and artifacts (grey). The model interestingly predicts most of the upper band as valid events even though the analysis shows that the pulses are distorted. The quality cuts remove these events in 3.2.2. The bottom pictures show the results of the first CNN model in the low-energy region (bottom left) and the energy spectrum below 1keV (bottom right) with light shaded areas indicating events below the trigger threshold which we would assume to be artifacts, given the analysis of Sec. 3.1.4. The plots measured in Volts and the energy spectrum are related via the V-to-keV conversion factor $C \sim 10\text{keV}/\text{V}$ presented in Sec. 3.2.3. The model predicts much more absorber events below 1keV compared to the quality cuts in Fig. 3.14. The spectrum for ring events on the other hand looks quite similar between the two plots.

Figure 4.6: Efficiency of initial model on low energy data set overlaid with the error function (red). Depicted is the average prediction within a pulse height bin (blue curve). For the plot 200 bins were used. Above true pulse heights of 0.045V one observes that the model is able to accurately predict, with an accuracy of more than 99.9%, all simulated pulses. Below these pulse heights the impact of noise on the simulated event is so large that the initial model starts to mislabel the events. The analysis threshold (grey), as introduced in Sec. 3.2.2, is calculated to be 2.2mV.

4.4 Result Enhancement of the CNN

Even though promising results mirroring the event classification of quality cuts in the low-energy region were achieved, it might be possible to extract higher precision classifications. This goal might be attained by re-using the outcomes of the initial model in an alternate way or implementing a model which discriminates between events based on ring channel data, thus exploiting cross-channel correlations in the pulse height data mentioned above, in Sec. 4.2. Finally, the outcome of a combination of both enhancement strategies will be explored, ending with a discussion of the obtained energy spectra below 1keV and their significance with regards to the Low-Energy Excess.

4.4.1 Improvement Strategy I

The general analysis of the event discrimination between ring and absorber hits in the previous section was performed by completely ignoring the separation of artifacts and valid events. Even though this was remedied by taking the cross-section of two CNN models in a later step, we will now attempt to extract all pulses which intro-

duced uncertainty in the absorber/ring model. Here we assume that any event whose prediction was in any way ambiguous for the network, i.e. the pulse shape was in both parts absorber and ring event, must lead to the conclusion that the event in question has to be discarded. Thus, the process of artifact/event classification by the previously trained model will be augmented by the concrete output of the CNN. This, as mentioned above, corresponds directly to the prediction probabilities implemented in the *PyTorch* library used in *Cait*. The results are and pictured in Fig. 4.7.
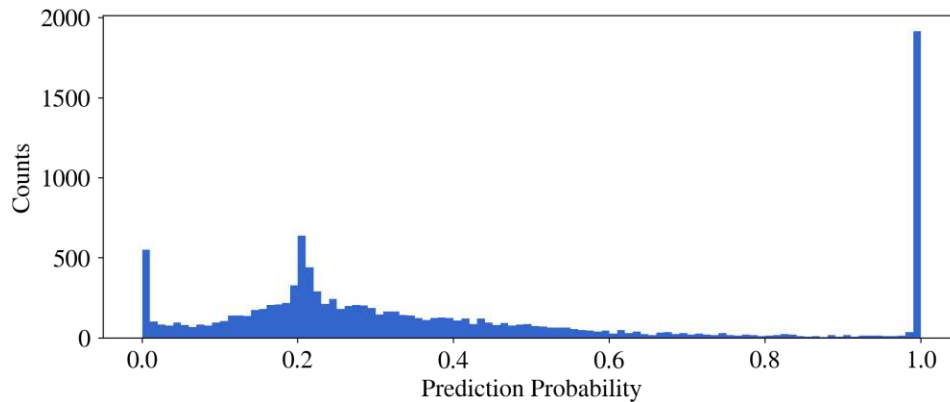
Figure 4.7: Histogram of prediction probabilities for all events in the stream data set calculated by the initial CNN. events below a probability of 0.5 are labelled as absorber, while events above the threshold are labelled as ring events. Particularly striking is the high number of events labelled as absorber hits but with probabilities around 0.2, hinting at the fact that the model is more "unsure" in its absorber predictions compared to the ring predictions. This fact is made use of in the first improvement strategy.

The plot of prection probabilities showcases the fact that there do exist a large number of events, especially predicted absorber events, which are quite far from being a "certain" label. A vast number of events, which one might assume are mainly artifacts, lies around the region of about 0.2. All events within this region of "uncertain" predictions will hence be labelled as artifacts. This is ensured by choice of a strict value at the prediction probability of the corresponding event type. For this analysis only events which differ by an amount of at most 0.01 from the maximum prediction probability of the related event type. The first improvement strategy, accepting only absorber predictions with probabilities of 0.01 or less and ring predictions with probabilities of 0.99 or more, yields the spectrum and event classifications depicted in Fig. 4.9.

The updated model displays, first of all, less valid events, i.e. "non-artifacts", overall and most obviously in the low-energy region below 1keV recoil energy. As expected by the distribution of prediction probabilities and their subsequent cuts, one can observe

a stark decrease in absorber predictions in this region with the number of events being cut from 71 to 31 in total. Further, many of the events below the trigger threshold have now been labelled as artifacts, with only 6 mislabels left. It can therefore be concluded that with the more drastic cut utilising prediction probabilities we were able to extract artifacts, especially those present at extremely low amplitudes, which were not picked up by the previously trained universal model. Even though the model applies a vast number of different pulse shapes it still seems to struggle with the Gode3 at lower energies. Hence, reinforcing the validity of utilising prediction probabilities for stricter event discrimination in the context of CRESST data.

Apart from the obvious mislabelling of events in the high-energy region which feature oversaturated and distorted events and the remaining below-threshold events which are still left, closer analysis of events below 0.1V OF pulse height leads to the fact that a lot less absorber hits are present in this region than predicted by the model. This conclusion is reached by successively checking each "non-artifact" pulse below these amplitudes. An examplary faux-absorber event is shown in Fig. 4.8. The predictions following the first improvement strategy therefore still exhibit inconsistencies in the way the model labels pulses and is consequently still in need of optimisation which will be attempted presently.
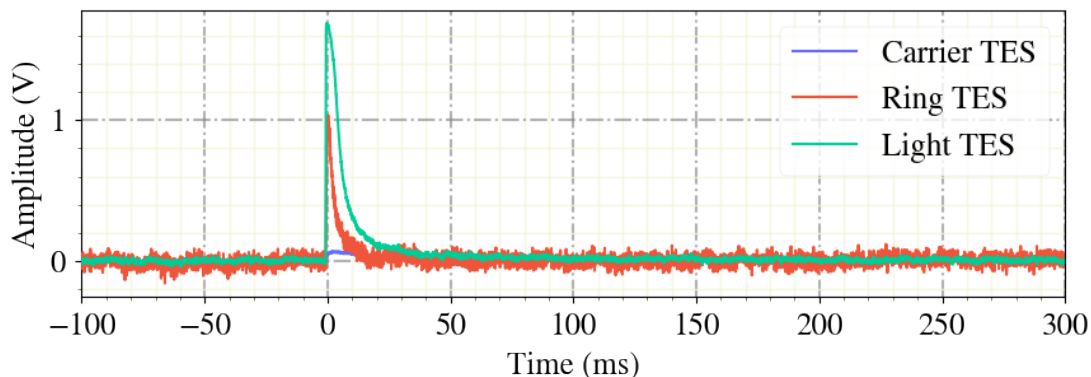
Figure 4.8: The event depicted here, with event index 2250, is labelled by the CNN as an absorber event but displays clear characteristics of a ring event, i.e higher pulse height and short rise/decay times in the ring channel.
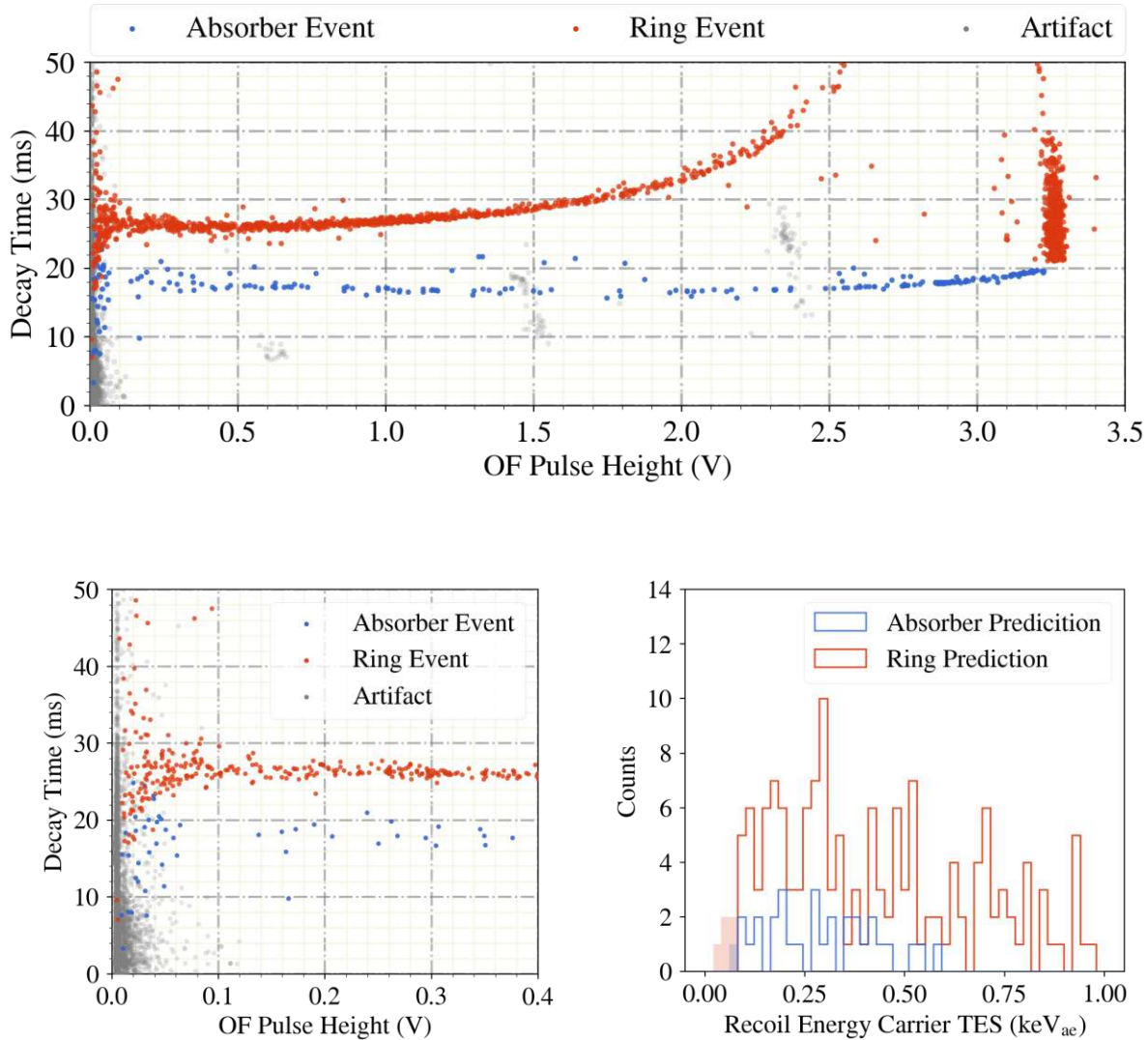
Figure 4.9: Shown are the results after first improvement strategy. The top plot shows the classifications by restricting the interval of prediction probabilities of the inital model to 0.01. The model differentiates between absorber events (blue), ring events (red), and artifacts (grey). Again the upper band of events are labelled as valid events even though analysis shows that many of these events are distorted or saturated. In the bottom pictures the focus lies solely on the low-energy region (bottom left) and the energy spectrum below 1keV (bottom right). Compared to the initial model, Fig. 4.5, the overall number of absorber and ring events is reduced here. This leads to the conclusion that at least some events which were previously labelled as ring or absorber events were associated with some amount of ambiguity in the intial model.

4.4.2 Improvement Strategy II

All previous analysis and subsequent predictions made by the various combinations of CNN models have not made use of the fact that pulse heights between phonon detectors are correlated. The generation of pulse heights described in 4.2 accounts for exactly this type of relation between carrier and ring channel amplitudes. The simulated events thus attempt to map these correlations and provide additional information which aim to improve the quality of the model's predictions.

The second improvement strategy is modelled by using the simulated events in the ring channel and training another separate CNN, which in essence aims to act as a veto on the predictions of the initial, combined model. The sum of the models is accomplished by imposing that a valid particle interaction is not only one which has not been extracted by the pre-trained model as an artifact but an interaction which has no contradicting label between the models trained on the carrier and ring channel data. This means that any event is discarded as being an artifact where the initial model predicts the event to be an absorber hit for instance, while the model using the simulated ring data predicts a ring hit. The results stemming from this procedure of labelling events are pictured in Fig. 4.10.

Fig. 4.10 shows that the new model now removes a large amount of distortions in the ring hit band and filters out oversaturated events in the cluster at approximately 3.25V OF pulse height. The number of absorber predictions is additionally diminished in the low-energy region which now approaches the results of the quality cuts and the analysis of Sec. 3.2.2. The only shortcoming of the model is that there is only little change in the incorrect predictions of events below the trigger threshold which constitute invalid events. In an effort to mend defficiencies present in the first and second improvement approaches we will attempt to combine both strategies into one combined discrimination model. This aims to remedy the prediction of distorted pulses and absorber hits below 1keV for the first strategy, while attempting to also improve upon the extraction of events below the trigger threshold for the second strategy.
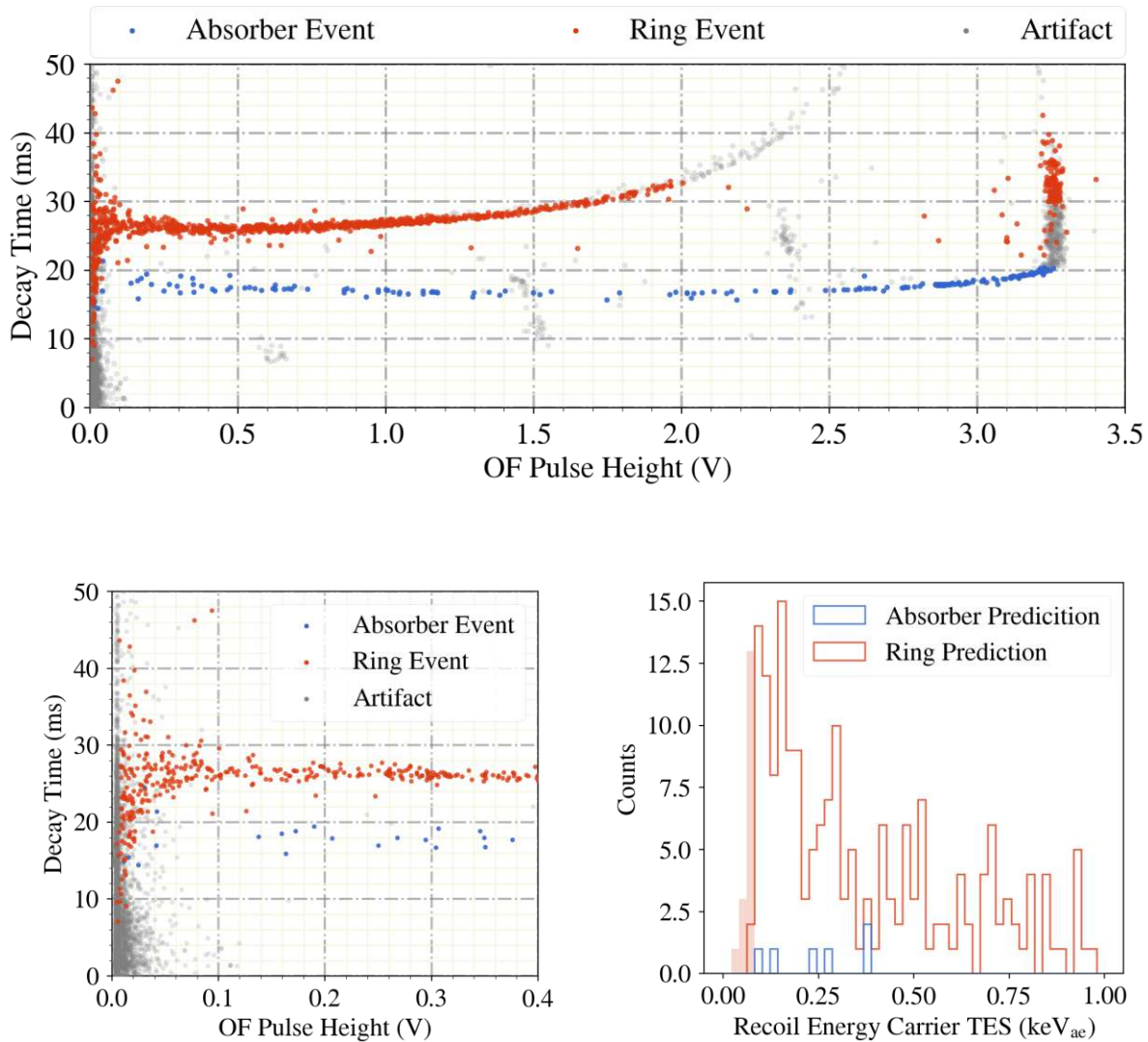
Figure 4.10: The results of the second strategy induced by including the veto by the ring channel predictions are depicted over the whole range of OF pulse heights in the top plots. The model, now incorporating the correlations between the pulse heights in the phonon detectors, differentiates between absorber events (blue), ring events (red), and artifacts (grey). The information provided by the data from the second channel now removes a big part of the distorted events in the ring hit band, as well as some of the over saturated events at 3.25V OF pulse height. The plots at the bottom provide a closer look at the low-energy region (bottom left) and the energy spectrum below 1keV (bottom right), with the light shaded area corresponding to events below the trigger threshold. The number of absorber predictions has decreased highly while only few of the ring predictions have been removed by the second strategy.

4.4.3 Combining Strategies I and II

To conclude this chapter's analysis and provide a model whose performance in labelling real world CRESST data is sufficiently high, the strategies presented in the last two sections are combined and presented in the following.

As explained in Sec. 4.4.1, prediction probabilities, which intrinsically arise in the way CNN's predict data, can be employed to drastically improve the model's ability of extracting artifacts in the data. The choice of prediction tolerance as 0.01 supresses predictions for which the model is not clearly decided. Combining this strategy additionally with a model trained solely on the simulated data set in the ring channel, will be able to discriminate between events extremely well. The merger of these strategies is modelled in the following way:

- **Absorber prediction**: An absorber prediction is labelled as an "event" by the pre-trained model and simultaneously has prediction probabilites in both CNNs trained on the carrier and the ring channel between $[0, 0.01]$.

- **Ring prediction**: Ring predictions are also necessarily "events" in the pre-trained model. The difference here is that the binary label for ring hits is "1" and thus both CNNs need to have prediction probabilities between $[0.99, 1]$.

- **Artifact**: An artifact will then simply be everything which does not fall in the absorber or ring category.

As this is an even stricter condition imposed upon the predictive ability of the final model less events will remain in the spectrum. This would then be consistent with the results of the quality cuts and strengthen the confidence in the model's ability to correctly discriminate events. The results of the final combined model implementing both improvement strategies is displayed in Fig. 4.11.

As postulated at the end of the last section one is able to observe that the combination of both result enhancement strategies indeed improves upon the issues which have been noted. The energy spectrum of the final model now shows that only 2 events which lie below the trigger threshold are being misclassified as ring hits which is a huge improvement to the inital model. Moreover, the implementation of the final model now also solves the issue that a large number of events in the high energy region were misidentified as ring hits. These artifacts above 2V OF pulse height are now succesfully filtered out mainly by the implementation of a model which trains on the simulated ring channel events. The CNN model we arrive at finally now also displays a really small number of absorber hits in the energy spectrum below 1keV which is a akin to the result of the event discrimination by the quality cuts. Measuring the model's performance on real Gode3 data on the basis of only consistency checks, as done above, leads to the conclusion that the CNN as employed in its final version presented in this section provides a suitable event discriminator.
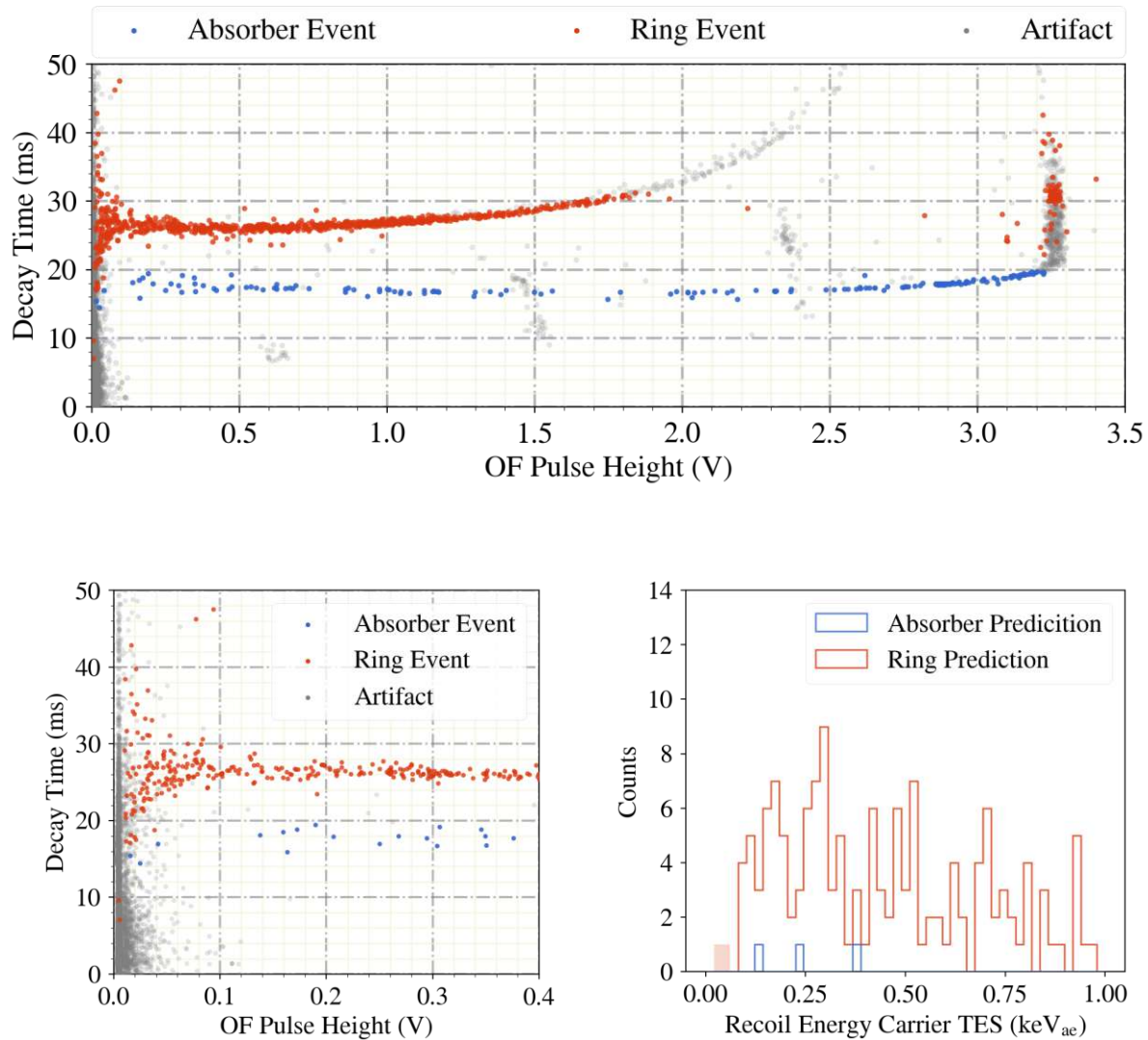
Figure 4.11: Shown are the results of the final CNN model for the whole range of OF pulse heights in the top plot and the low-energy region in the bottom plots. The model differentiates between absorber events (blue), ring events (red), and artifacts (grey). On the bottom right the energy spectrum below 1keV using 50 bins is displayed, with events below the trigger threshold displayed by the shaded areas. The predictions by the model implementing both the first and second improvement strategy showcase positive results in both higher energies but also in the low-energy region. It is visible that distorted and oversaturated events above 2V OF pulse heights are now correctly labelled as artifacts as well as reducing the number of non-artifacts below the trigger threshold to only 2.

Given that all means of assessing the quality of the final model without validating the predictions themselves have been exhausted, we will move to quantifying the performance of the model in the region below 1keV directly. Thus, each event and its corresponding prediction below this threshold is examined, resembling a sort of "reverse supervision".

Below 1keV one finds 148 non-artifact events in total, split into 145 ring predictions and 3 absorber predictions. Examination of all events one by one yields the following distribution:

- **Prediction correct**: All three absorber events have been correctly identified by the model with 126 further correct predictions of ring hits.

- **Noise baselines**: A total of four record windows are completely filled with noise but still predicted as ring hits. All four of those events are below 250eV in the spectrum.

- **Distorted ring hits**: From the ring predictions we see that 13 events are in fact hits directly into the ring detector but feature a distorted pulse shape due to the detector's instability. We count these events as mislabelled since they are defined as artifacts, though they are in fact direct ring hits.

- **Other artifacts**: Two unusual pulse shapes which do not constitute ring hits are mislabelled as such.

This prediction distribution gives a score of 87.16% for the final model. An impressively high score for a model which has not been trained on a single real world data set but only on simulated data. Neither the absorber/ring classifier nor the pre-trained artifact/event discriminator utilize real CRESST data but learn on completely artificial pulses. What is additionaly more worth mentioning is that the main classifier for artifacts was trained on exclusively artificial data. As a final consistency check, as done presented in Fig. 4.6, the model's efficiency on the simulated data set is once again calculated. This ensures that the combined CNN does not reach the high scores listed above by eliminating most of the data set as artifacts. The efficiency curve of the combined model is depicted in Fig. 4.12 and shows that even at very low pulse heights, as low as 42mV, the model is able to accurately differentiate absorber and ring events as well as artifacts.

In light of these facts one may conclude that the final, combined CNN model constitues a discriminator on par with manual quality cuts. This conclusion rings true especially due to the fact that the Neural Network was able to identify the only three absorber events in the region below 1keV recoil energy which were discarded by the quality cuts. The analysis here thus suggests that Neural Networks are well suited in discriminating events in the region affected by the Low-Energy Excess.

As mentioned at the end of the last chapter 3 one faces the same apparent situation. The first analysis via quality cuts hints at the fact the Low-Energy Excess might be generated solely in the ring of the detector, for instance due to shear stress. When

employing Machine Learning for exactly the same task one observes exactly the same phenomenon, with ring events dominating the region of lower recoil energies in both carrier and ring channel, the former pictured in Fig. 4.11 and the latter in Fig. 4.13. Both of these facts might hint at the origin of the Low-Energy Excess stemming from shear stress in the holders of the crystal, with more analysis and new data on this topic still proving to be necessary.
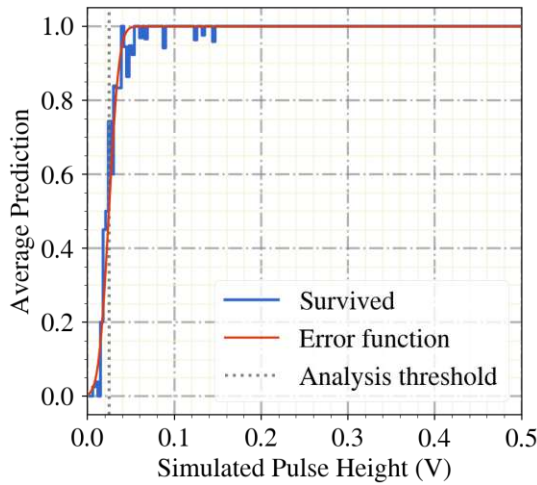

Figure 4.12: Efficiency curve of the combined model below 500eV. The average predictions in an energy bin (blue curve) and its corresponding error function (red curve) show that the model is capable of reaching high scores >99.9% at up to pulse heights of about 42mV. For the calculation of the efficiency 1000 bins in the interval [0,3] V were implemented. At even higher pulse heights >500eV the average prediction per bin remains steadily at ∼1. The threshold of the final model (grey) lies at 2.48mV which is in good agreement with the value found for Fig. 4.6.
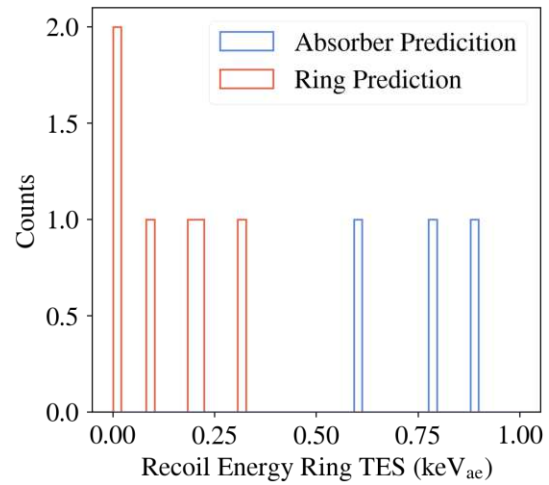
Figure 4.13: Energy spectrum of the ring channel predictions below 1keV. Though only few events are present in this region, only ring predictions are present with energies relevant for the Low-Energy excess. Note that no events with recoil energies below the trigger threshold are present in this spectrum compared to its counterpart in the carrier in Fig. 4.11.

Conclusion

This thesis first and foremost explores event classification of CRESST event data and utilises two vastly different concepts in this regard. Here, the standard way of CRESST event discrimination is juxtaposed with a more novel approach which employs Machine Learning for achieving this task. Based on the findings and progress already made in works such as [4–8] this thesis makes use of Machine Learning techniques explored there and expands upon these concepts.

Before any analysis via Machine Learning was done, we first provided a blueprint for raw data analysis typical for the most recent CRESST-III runs on the basis of Gode3 data within Run36. This includes the preliminary analysis of hardware triggered data necessary for the extraction of the trigger threshold which in turn is employed in the triggering of the captured data stream. In this section a large focus was laid on the types of events present in the Gode3 data set which - as it turned out - provided two distinct sets of detector interactions visible in the phonon channels. This fact, most clearly visible in decay and rise time plots of the channels in question, has the consequence that the analysis had to differentiate between two different types of Standard Events, absorber and ring events. The analysis was additionally able to designate two separate types of events for the light channel as well which were called scintillation events and direct light hits, but which are not important for the subsequent analysis. Finally, this section concluded with the definition of the baseline resolution and trigger threshold which are both in line with previous analysis on this detector in [42].

Following the triggering of the streamed data set we commenced with the stream analysis which is split into mainly two parts. On one hand Sec. 3.2 deals with the definition of different quality cuts in order to extract valid particle interactions in the data, while the second part on the other hand mainly concerns itself with the description and conversion of amplitudes into corresponding recoil energies. It was once again necessary, by appropriate choice of quality cuts, to not only remove artifacts from the data set but also differentiate between absorber and ring events. The subsequent simulation of events and calculation of cut efficiencies proved to coincide with cut efficiency spectra found in previous CRESST works, thus confirming the validity of the cuts performed here. The section finally also discussed the energy spectra and Light Yield plots of the Gode3 detector, leading to a novel observation in the low-energy region. The energy spectrum of the carrier channel and the designated veto channel, the ring, both exhibit a large amount of ring events with recoil energies

below 1keV which largely outnumber those of absorber hits. This fact can also be observed in the light channel and might hint at the fact that the source of the Low-Energy Excess might truly stem from shear stress in the holders of the detector.

In order to further analyse this circumstance we moved to training and applying a Machine Learning algorithm to the Gode3 data. By initially training one Convolutional Neural Network model which differentiates between absorber and ring hits and reusing a pre-trained model from Cait for Artifact/Event discrimination we were already able to reproduce some of the results achieved via quality cuts. The major step forward here was that both models were trained in semi-supervised fashion, only on the basis of simulated pulse shapes. This step completely circumvents the necessity of labelling upwards of 10000 events for standard Supervised learning, and thus also provides time efficiency next to satisfactory predictive capabilities.

In the aftermath of the intial model new methods of enhancing the results were explored. The use of prediction probabilities and the training of a third CNN using the simulated and pulse height correlated data on the ring channel in tandem, showcased a significant jump in performance of the combined, final model. The predictions show self-consistent behaviour for higher, removal of distortions and oversaturated events, and in lower energy regions, only two events below the trigger threshold of 90eV on the carrier channel are mislabelled as ring events. All in all, the final model displays high precision scores of 87.16% in the interest region below 1keV, with most mislabels arising due to distortions and instabilities of direct ring hits, see Fig. 2.6. These events could also be considered as correctly labelled ring events, since they in fact do represent the correct crystal-particle interaction. In this analysis though we removed them from the data set as artifacts. It should additionally be noted that the final model was able to correctly extract absorber events in the low-energy region where quality cuts failed to do so. Final considerations of the low-energy region verify the results from quality cuts and show that energies relevant for the Low-Energy Excess are completely dominated by ring hits in both phonon channels.

In conclusion, this thesis proves once more that Machine Learning within CRESST is able to take on event classification tasks and in some specific cases exceed the possibilities of standard techniques such as quality cuts. As quality cuts are designed to extract any types of artifacts from the data, one is necessarily commited to choosing cuts in a strict fashion. Machine Learning on the other hand can, if the model is sophisticated enough, provide a much more granular method of event discrimination. Future attempts at utilising semi-supervised learning in CRESST or similar experiments might achieve further improvements by, first of all, analysing pulse height relations more closely which in turn yield much more accurate representations of correlated pulse heights. The usage of linear correlations, as done in this work, faces problems in both extremely low energies, where noise dominates, and high energies, where pulses become oversaturated. The usage of Machine Learning in Dark Matter search has not nearly been completely exhausted and thus this thesis can be used as a stepping stone for more analysis in the future.

List of Figures

List of Tables

Bibliography

[1] F. Zwicky. "On the Masses of Nebulae and of Clusters of Nebulae", The Astrophysical Journal pp. 86-217 (1937). 10.1086/143864

[2] CRESST Collaboration et al. "First results from the CRESST-III low-mass dark matter program".Physical Review D 100(10) (2019).
DOI:10.1103/PhysRevD.100.102002.

[3] J. Schmidhuber. "Deep learning in neural networks: An overview". Neural Networks 61 (2015), pp. 85–117. DOI: 10.1016/j.neunet.2014.09.003

[4] C. Mühlmann. "Pulse-shape discrimination with deep learning in CRESST". Master's Thesis, TU Wien (2019).

[5] F. Wagner. "Machine Learning Methods for the Raw Data Analysis of Cryogenic Dark Matter Experiments ", Master's Thesis, TU Wien (2020)

[6] D. Bartolot. "Supervised and Unsupervised Classification of Data from Cryogenic Dark Matter Experiments". Project Work, TU Wien (2021).

[7] D. Rizvanović. "Supervised Classification and Feature Analysis for CRESST Event Data", Project Work, TU Wien (2021)

[8] A.J. Zöller. "Artificial Neural Network Based Pulse-Shape Analysis for Cryogenic Detectors Operated in CRESST-II". PhD Thesis, TU München (2016).

[9] F. Pröbst, M. Frank, S. Cooper, P. Colling, D. Dummer, P. Ferger, G. Forster, A. Nucciotti, W. Seidel, L. Stodolsky. "Model for cryogenic particle detectors with superconducting phase transition thermometers". Journal of Low Temperature Physics 100 (1995), pp. 69-104. DOI: 10.1007/BF00753837

[10] E. Gatti & P. Manfredi. "Processing the signals from solid-state detectors in elementary-particle physics". Riv. Nuovo Cim. 9, pp. 1–146 (1986).
DOI: 10.1007/BF028221567

[11] A. Liddle. "An Introduction to Modern Cosmology". Wiley, 2nd ed. (2003). ISBN: 9780470848357.

[12] F. Zwicky. "An Introduction to Modern Cosmology". Helvetica Physica Acta 6 (1933), pp. 110-127. DOI: 10.5169/seals-110267.

[13] A. Blanchard. "Clusters of galaxies". New Astronomy Reviews 45(4-5) (2001), pp. 401-417. DOI: https://doi.org/10.1016/S1387-6473(00)00162-7.

[14] Planck Collaboration et al. "Planck 2015 results. XIII. Cosmological parameters", Astronomy & Astrophysics, 594 (2016) A13. DOI: 10.1051/0004-6361/201525830.

[15] T. Wrase. Lecture notes on "Cosmology and particle physics". TU Wien, Summer Semester (2019).

[16] "Galaxy rotation curve", Wikipedia, published September 18, 2021. Visited December 29, 2021. Link: en.wikipedia.org/wiki/Galaxy_rotation_curve.

[17] "Catalog Page for PIA16874.", NASA. Visited December 29, 2021. Link: photojournal.jpl.nasa.gov/catalog/PIA16874.

[18] D. Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". The Astrophysical Journal Letters 648(2) (2006), p. L109. DOI: 10.1086/508162.

[19] "More Images of 1E 0657-56.", Chandra. Visited December 29, 2021. Link: chandra.harvard.edu/photo/2006/1e0657/more.html.

[20] V. Zacek. "Dark Matter". Fundamental Interactions (2007). DOI: 10.1142/9789812776105_0007.

[21] R.D. Peccei & H.R. Quinn. "CP Conservation in the Presence of Pseudoparticles". Physical Review Letters 38(25) (1977), pp. 1440-1443. DOI: 10.1103/PhysRevLett.38.1440.

[22] H. Primakoff. "Photo-Production of Neutral Mesons in Nuclear Electric Fields and the Mean Life of the Neutral Meson". Physical Review 81, p. 899. DOI: 10.1103/PhysRev.81.899.

[23] D: Gorbunov & A. Panin. "Minimal active-sterile neutrino mixing in seesaw type I mechanism with sterile neutrinos at GeV scale". Physical Review D 89. DOI: 10.1103/PhysRevD.89.017302.

[24] H.V. Klapdor-Kleingrothaus & K. Zuber. "Teilchenastrophysik", Teubner (1997). DOI: 10.1007/978-3-322-90548-2.

[25] G. Servant & T. Tait. "Is the lightest Kaluza–Klein particle a viable dark matter candidate?", Nuclear Physics B 650(1–2) (2003), pp. 391-419. DOI:10.1016/S0550-3213(02)01012-X

[26] H. Cheng & I Low. "Little Hierarchy, Little Higgses, and a Little Symmetry." Journal of High Energy Physics 2004(08) (2004), pp. 061–061. DOI: 10.1088/1126-6708/2004/08/061.

[27] ANAIS Collaboration et al. "Annual Modulation Results from Three Years Exposure of ANAIS-112". Phys. Rev. D 103, 102005 (2021). arXiv: 10.1103/PhysRevD.103.102005.

[28] R. Bernabei, R, P. Belli, F. Cappella et al. . "Final model independent result of DAMA/LIBRA-phase1". The European Physical Journal C 73, 2648 (2013). DOI: 10.1140/epjc/s10052-013-2648-7.

[29] COSINUS Collaboration et al. "Simulation-based design study for the passive shielding of the COSINUS dark matter experiment ". The European Physical Journal C 82, 248 (2022). DOI:10.1140/epjc/s10052-022-10184-5.

[30] B. Lee & S. Weinberg. "Cosmological Lower Bound on Heavy-Neutrino Masses". Phys. Rev. Lett. 39, 165 (1977). DOI: 10.1103/PhysRevLett.39.165.

[31] K. Petraki & R. Volkas. "Review of asymmetric dark matter". International Journal of Modern Physics A 28, 19 (2013). DOI:10.1142/S0217751X13300287.

[32] J. Schieck. Lecture notes on "Search for Dark Matter", TU Wien, Fall Semester (2021).

[33] P.A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020) and 2021 update. DOI: 10.1093/ptep/ptaa104.

[34] F. Reindl. "Exploring light dark matter with CRESST-II low-threshold detectors". PhD Thesis, TU München (2016). URL: mediatum.ub.tum.de/?id=1294132

[35] XENON Collaboration et al. "Dark Matter Results from 225 Live Days of XENON100 Data". Physical Review Letters 109(18) (2012), p. 181301. DOI: 10.1103/PhysRevLett.109.181301.

[36] XENON Collaboration et al. "Search for Light Dark Matter Interactions Enhanced by the Migdal Effect or Bremsstrahlung in XENON1T ". Physical Review Letters, 123(24) (2019). DOI: 10.1103/PhysRevLett.123.241803.

[37] Michael Kiefer. "Improving the Light Channel of the CRESST-II-Dark Matter Detectors". PhD Thesis, Technische Universität München (2012). URL: mediatum.ub.tum.de/?id=1097360

[38] G. Aad et al. "Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 8$TeV with the ATLAS detector". The European Physical Journal C 75.7 (2015), pp. 1–43. DOI: 10.1140/epjc/s10052-015-3517-3.

[39] V. Khachatryan et al. "Search for dark matter, extra dimensions, and unparticles in monojet events in proton–proton collisions at $\sqrt{s} = 8\text{TeV}$". The European Physical Journal C 75.5 (2015), pp. 1–25. DOI: 10.1140/epjc/s10052-015-3451-4.

[40] I. Adachi et al. "Search for an Invisibly Decaying Z' Boson at Belle II in $e^+e^- \rightarrow \mu^+\mu^-(e^\pm\mu^\mp)$ Plus Missing Energy Final States". Phys. Rev. Lett. 124, 141801 (2020). DOI: 10.1103/PhysRevLett.124.141801.

[41] CRESST Collaboration et al. "Results on low mass WIMPs using an upgraded CRESST-II detector". European Physics Journal C 74, 3184 (2014). DOI:10.1140/epjc/s10052-014-3184-9.

[42] M. Kaznacheeva. "Gode-modules in Runs 35-36". Slides presented at CRESST Collaboration Meeting (2021).

[43] D. Bartolot, J. Burkhart, D. Rizvanovic, F. Wagner. "Cait", PyPi. Visited January 2, 2022. PyPi link: Cait

[44] M. Stahlberg. "Probing low-mass dark matter with CRESST-III: data analysis and first results". PhD Thesis, TU Wien (2020). DOI: 10.34726/hss.2021.45935

[45] N. Ferreiro Iachellini. "Increasing the sensitivity to low mass dark matter in CRESST-III with a new DAQ and signal processing". PhD Thesis, LMU München (2019). DOI: 10.5282/edoc.23762

[46] P. Adari et al. "EXCESS workshop: Descriptions of rising low-energy spectra". arXiv:2202.05097 [astro-ph.IM] (2022). arXiv: astro-ph.IM/2202.05097.

[47] S. Di Domizio, F. Orio and M. Vignati. "Lowering the energy threshold of large-mass bolometric detectors". JINST 6 (2011) P02007. DOI: 10.1088/1748-0221/6/02/P02007

[48] J.-R. Ohm & H. D. Lüke. "Signalübertragung". Springer-Verlag Berlin Heidelberg, 12 ed. (2014). DOI: 10.1007/978-3-642-53901-5

[49] A. Oppenheim, G. Verghese. "Signals, Systems and Inference" Pearson, Global ed. (2017). ISBN: 9781292156201

[50] A. L. Samuel. "Some Studies in Machine Learning Using the Game of Checkers," in IBM Journal of Research and Development 3, no. 3 (1959), pp. 210-229. DOI: 10.1147/rd.33.0210.

[51] D. Ciregan, U. Meier, J. Schmidhuber. "Multi-column deep neural networks for image classification". 2012 IEEE Conference on Computer Vision and Pattern Recognition (2012), pp. 3642-3649. DOI: 10.1109/CVPR.2012.6248110.

[52] J. Hu, H. Niu, J. Carrasco, B. Lennox, F. Arvin. "Voronoi-Based Multi-Robot Autonomous Exploration in Unknown Environments via Deep Reinforcement Learning". IEEE Transactions on Vehicular Technology 69(12) (2020), pp. 14413-14423. DOI: 10.1109/TVT.2020.3034800.

[53] A. Krizhevsky, I. Sutskever, G.E. Hinton. "ImageNet classification with deep convolutional neural networks". Commun. ACM 60(6) (2017), pp. 84–90. DOI:10.1145/3065386.

[54] Lu Lu, P. Jin, G. Karniadakis. "DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators". arXiv:1910.03193 [cs.LG] (2019). arXiv: 1910.03193.

[55] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar. "Fourier Neural Operator for Parametric Partial Differential Equations". arXiv:2010.08895 [cs.LG] (2020). arXiv: 2010.08895.

[56] W. McCulloch & W. Pitts. "A logical calculus of the ideas immanent in nervous activity". Bulletin of Mathematical Biophysics 5 (1943), pp. 115–133. DOI:10.1007/BF02478259.

[57] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain". Psychological Review 65(6) (1958), pp. 386-408. DOI: 10.1037/h0042519.

[58] S.S. Haykin. "Neural Networks and Learning Machines". Pearson, 3rd ed. (2016). ISBN: 9780131471399.

[59] P. Mehta, M. Bukov, C. Wang, A.G.R. Day, C. Richardson, C.K. Fisher, D.J. Schwab. "A high-bias, low-variance introduction to Machine Learning for physicists". Physics Reports 810 (2019), pp. 1-124.
DOI: 10.1016/j.physrep.2019.03.001.

[60] F. Wagner, D. Bartolot, D. Rizvanović, F. Reindl, J. Schieck, W. Waltenberger. "A Python package with novel raw data analysis methods for cryogenic particle detectors". Unpublished paper (2022).

[61] K. Fukushima. "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position". Biol. Cybernetics 36 (1980), pp. 193–202. DOI: 10.1007/BF00344251.

[62] R.S. Sutton & A.G. Barto. "Reinforcement learning: An Introduction". MIT Press, 2nd ed. (2018). ISBN: 9780262039246.

[63] D. Greene, P. Cunningham, R. Mayer. "Machine Learning Techniques for Multimedia". Cognitive Technologies, Springer (2008), pp. 51-90.
DOI: 10.1007/978-3-540-75171-7.

[64] I. Goodfellow, Y. Bengio, A. Courville. "Deep Learning". MIT Press (2016). DOI: 10.1007/s10710-017-9314-z.

[65] "3D Visualization of a Convolutional Neural Network", visited February 10, 2022. Link: cs.ryerson.ca/ aharley/vis/conv/

[66] "Typical CNN architecture", Wikipedia, published December 15, 2015. Visited February 10, 2022.
Link: en.wikipedia.org/wiki/Convolutional_neural_network.

[67] F. Li, J. Johnson, S. Yeung. Lecture Notes on "CS231n: Convolutional Neural Networks for Visual Recognition". Stanford University, Fall Semester (2017). Link: cs231n.stanford.edu/2017/.

[68] D.E. Rumelhart, G.E. Hinton, R.J. Williams. "Learning representations by back-propagating errors". Nature 323, pp. 533-536 (1986). DOI: 10.1038/323533a0.

[69] W. Waltenberger. Lecture notes on "Statistische Methoden der Datenanalyse". TU Wien, Fall Semester 2020. Link: Lecture notes.

[70] D. Kingma & J. Ba. "Adam: A Method for Stochastic Optimization". International Conference on Learning Representations (2014). arXiv: 1412.6980.

[71] "Universal Training Data Augmentation". Cait documentation. Visited April 10, 2022. Link: 13universaltrainingset.html

Danksagung

Diese Arbeit ist das Produkt eines langen Prozesses mit vielen unvorhersehbaren Wendungen, Hindernissen und Rückschlagen, welche ich ohne die tatkräftige Unterstützung außerordentlich engagierter und rücksichtsvoller Personen nie alleine geschultert hätte. Vor allem in den letzten zwei Jahre, die für uns alle eine besonders große Belastung darstellten, hat mich mein tolles Umfeld aus jeder scheinbar unüberwindbaren Lage herausgeholt.

Zunächst gilt mein Dank meinem herausragendem Betreuerteam.
Ich danke Jochen, dem Institusleiter des HEPHY und Leiter der Forschungsgruppe an der ich mitwirken durfte, für die Möglichkeit in so einem besonderem Team und einem faszinierendem Experiment sowohl meine Projekt- als auch diese Masterarbeit schreiben zu dürfen.
Ich danke Florian für seine tatkräftige Unterstützung bei allen Dark Matter spezifischen Themen und für seinen stets lehrreichen Input zu dieser Arbeit. Vor allem möchte ich mich hier auch dafür bedanken, dass mir die Chance und das Vertrauen gegeben wurde bei einem Collaboration Meeting etwas zu präsentieren und ich so zum ersten mal ein wenig "Konferenzluft" schnuppern konnte, auch wenn ich diese nur über Zoom wahrnehmen konnte.
Zuletzt möchte ich mich noch sehr bei Felix bedanken, der mich nicht nur als Studienkollege, Mittutor und schließlich als toller Betreuer bei meinen Master- und Projektarbeiten durch das Studium begleitet hat, sondern auch in das Developerteam für sein Projekt Cait dazugeholt hat. Hier wurde eine extrem coole Software gemeinsam mit Daniel aufgebaut, welche diese Arbeit auch erst möglich gemacht, und auf die du wirklich stolz sein kannst. Danke dir, dass du stets und wirklich zu jeder Zeit erreichbar warst und mich so unnachgiebig bei der Fertigstellung dieses Projekts unterstützt hast.

Das größte Dankschön gilt meinen Freunden und meiner wundervollen Familie! Ich möchte euch allen von ganzem Herzen dafür danken, dass ihr mich das ganze Studium, durch alle Höhen und Tiefen, und durch jede noch so schwierige Prüfungs- und Schreibphase emotional unterstützt und jede meiner Launen ausgehalten habt. Diese Arbeit wäre ohne euch nie zustande gekommen.

Gora, ich danke dir für jedes Wort der Unterstützung, jede Sekunde des Trostes, und jeden Satz des Anspornes, den du mir in den letzten Jahren gegeben hast. Auch wenn du selbst noch so viel zu tun hattest, du hattest immer ein offenes Ohr für mich und mein Physikgebrabbel. Ich könnte meinen Dank dir gegenüber hier noch seitenweise fortführen und ich wäre noch immer nicht auch nur halb fertig. Ich hoffe du weißt wie glücklich ich bin dich in meinem Leben und an meiner Seite zu haben!

Hvala ti majko, da si iz najtežih situacija našla snagu i volju da nama trojima omogućiš sve ovo što sad imamo. Nemam riječi za to koliko sam ja tebi zahvalan, pa mi ostaje samo da ovaj rad tebi posvjetim. Volim tebe i moje male seke puno!