TECHNISCHE UNIVERSITÄT WIEN

ATOMINSTITUT

HEPHY INSTITUT FÜR HOCHENERGIEPHYSIK

DIPLOMARBEIT

Calculation of Dark-Matter Exclusion-Limits using a Maximum Likelihood Approach

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Masterstudium Technische Physik

eingereicht von

Daniel Schmiedmayer, BSc.
Matrikelnummer 01027796

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

Betreuung
Betreuer: Univ.Prof. Dr. Jochen Schieck
Mitwirkung: Univ.Ass. Dr. Florian Reindl

Wien, 12.02.2019

(Unterschrift Verfasser) (Unterschrift Betreuer)

In this study a maximum likelihood formalism is used to calculate dark matter particle nucleus scattering cross section exclusion limits from data obtained in the CRESST-II and CRESST-III experiment. For this purpose a suitable description for modeling of data is chosen. A comparison to previous maximum likelihood implementations as well as Yellin's optimum interval method is made.

Contents

1 Abstract

The search for dark matter (DM) is one of the most fundamental questions in modern physics. Evidence for matter beyond the Standard Model (SM) of particle physics is found in various astrophysical observations. These observations also provide the basis for theories about the type and properties of potential DM particles. One of the most promising explanations involves weakly interacting massive particles (WIMPs) and similar WIMP-like particles, which are also referred to as WIMPs in the course of this work. Numerous experiments attempt to find proof of their existence and measure characteristic properties. So far no conclusive proof has been found, it was however possible to put limits on the expected cross section for the interaction of potential particles with SM particles. One of these experiments is the Cryogenic Rare Event Search with Superconducting Thermometers (CRESST) which is the currently leading experiment for establishing cross section exclusion limits for WIMPs in the low-mass region.

In this work an extended maximum likelihood method capable of calculating exclusion limits from the data measured with CRESST detectors is presented and tested. At the beginning an overview of the evidence for dark matter, its possible particle models and detection methods with focus on WIMPs and direct detection is given. Then, a short introduction into the CRESST experiment, its setup and capabilities is presented. The performance of the extended maximum likelihood method is compared against another method for limit calculation, namely Yellin's optimal interval method [1]. Therefore, Yellin's method is introduced alongside the maximum likelihood formalism. To evaluate the measured data in the maximum likelihood framework a model for the properties of the detector as well as the known backgrounds is established. Finally, the maximum likelihood method is used to calculate WIMP exclusion limits using data from two CRESST experimental data taking campaigns with the purpose of achieving stronger limits.

2 Dark Matter

To a large extend this chapter follows the formalism of the lecture notes of the "Search for Dark Matter" lecture [2] taught at the TU Wien in 2017/2018.

2.1 Evidence for Dark Matter

2.1.1 Velocity Distributions of Galaxies

The first indications for the existence of dark matter were discovered by F. Zwicky in 1933. He observed that the velocity distribution of galaxies in the Coma cluster is greater than expected. Through the viral theorem, which states that in a closed, bound system in equilibrium the average kinetic energy is half the potential energy $-\langle E_{pot}\rangle = 2\langle E_{kin}\rangle$, it was possible to estimate the total mass of the clusters:

$$M = \frac{5R}{3G}\overline{v^2} \tag{1}$$

Through this the ratio of gravitationally interacting to luminous matter in terms of solar mass $M_\odot$ and luminosity $L_\odot$ was determined for numerous clusters and it was found that $\frac{M}{L} \approx 300\frac{M_\odot}{L_\odot}$. It was therefore established that most of the matter in galaxy clusters is dark.

Similar observations were made for individual galaxies in the 1970s where the velocity distribution of stars outside the visible halo was studied. For these stars a Keplerian behavior where the velocity drops with $\frac{1}{\sqrt{r}}$ was expected. For a large number of measured galaxies the curve stays flat well beyond the galactic body. This also holds true for the milky way. To explain those rotational curves non luminous matter is required. The rotational curve of the M33 spiral galaxy is presented in figure 1 as an example for the discrepancy between visible and total matter.

2.1.2 Gravitational Lensing

Another hint for the existence of dark matter comes in the form of gravitational lenses. Gravitational lenses are massive objects which are capable of bending the light trajectory, thus acting similar to optical lenses. The mass of the object bending the light can be determined using the deflection angle α and the impact parameter ξ:

$$M = \frac{\alpha\xi c^2}{4G} \tag{2}$$

In case of strong visible distortion the effect is called strong gravitational lensing. Without clearly visible distortion it is called weak gravitational lensing. A highly located strong gravitational lens is called microlens. Through the observation of distortions in the image of far away objects the gravitational mass between the object and the observer can be determined. This allows to map the distribution of mass in the universe, even if it is non-luminous.

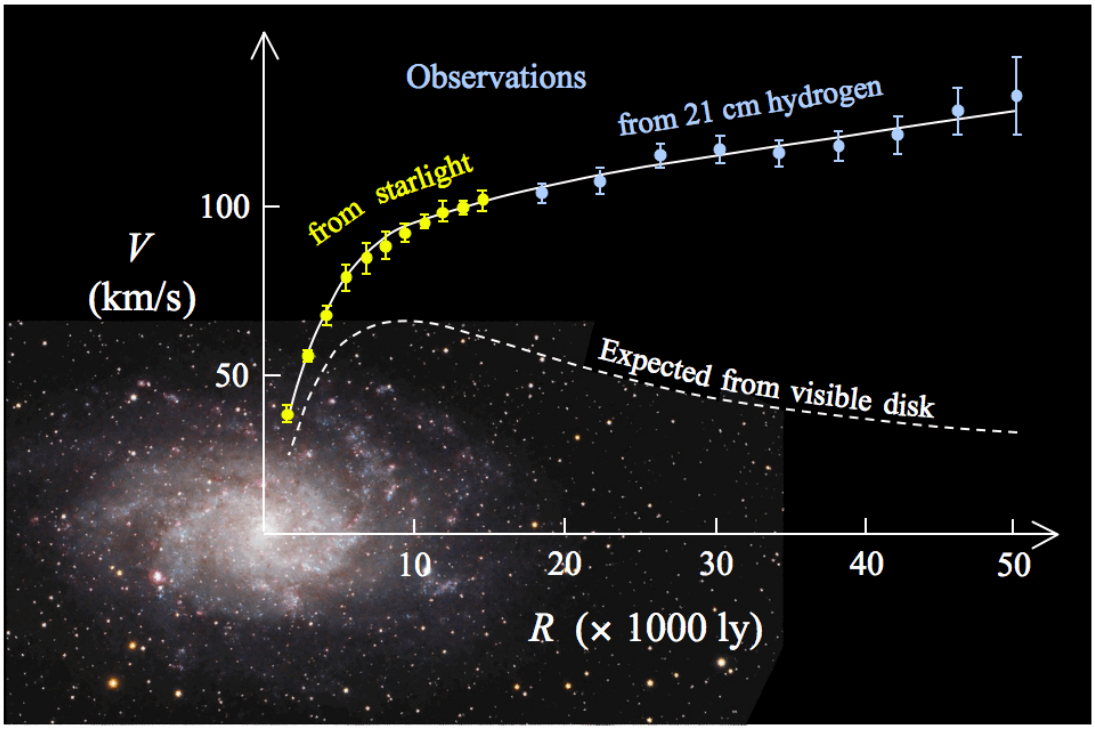


Figure 1: The measured rotation curve of the M33 spiral galaxy with the rotation curve expected from the visible matter as comparison [3].

One especially good example for non-luminous gravitational matter which was discovered this way is in the so-called bullet cluster, shown in figure 2. The bullet cluster consists of two well separated merging sub-clusters. Through measurement of the gravitational lensing it was found that the distribution of mass mostly coincides with the individual sub-clusters, but that the visible matter only contributes roughly 2% of the total mass. A standard baryonic dark matter candidate would be the intergalactic medium (hot gas mostly consisting of H and He ions). This gas was indeed found using X-ray observations. This matter only contributes to around 15% of the total mass and is located in-between the two subclusters with a visible bow shock in the gas. This is because the gas separated from the subclusters through friction when they were interpenetrating each other. Due to the large distance between individual stars they were mostly unaffected in this process. Since the distribution of remaining dark matter coincides with the subclusters it must have a low self-interaction cross section because it would otherwise have separated from the clusters through friction.

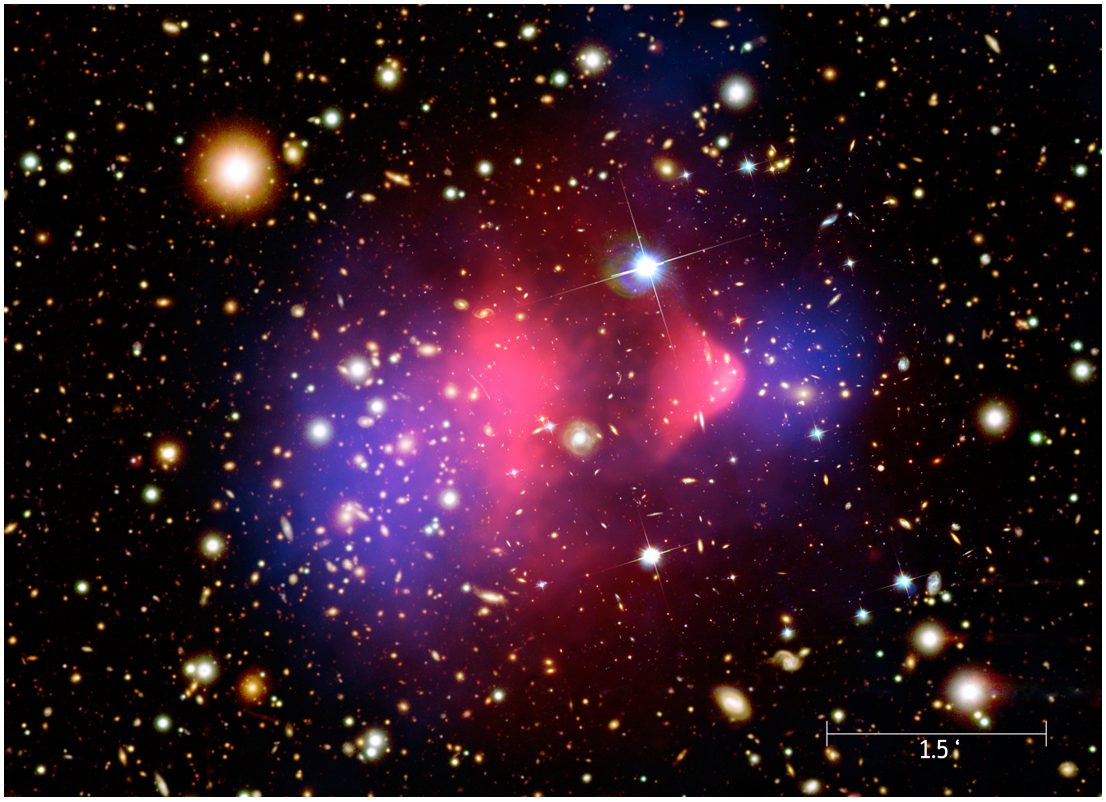


Figure 2: Bullet cluster; The x-ray emissions from the intergalactic medium are shown in pink and the blue overlay represents the mass distribution deduced from gravitational lensing [4].

2.1.3 Cosmic Microwave Background

In the early universe the temperatures were above the ionization energy of hydrogen and radiation and matter were in thermal equilibrium. Due to expansion of the universe the plasma cooled. At a temperature of roughly 3000 K the radiation decoupled and the universe became transparent. Today this radiation is redshifted to 2.7255 K and called cosmic microwave background (CMB). It is omnipresent and isotropic to an order of 10^{-4}. However, on the mK-scale the CMB is anisotropic, which is the result of density fluctuations of the matter at the point of decoupling, roughly 380000 years after the big bang. Therefore, the CMB represents a snapshot of the universe at that point. Small overdensities in the primordial plasma lead to a gravitational collapse, where the matter falls into the gravitational well. This causes the temperature to increase and with it the radiation pressure rises which counteracts the collapse. Due to this there were acoustic oscillations and higher harmonics in the plasma which stopped at the point of decoupling. An analysis of the angular power spectrum of the CMB allows to draw conclusions about these oscillations and therefore the distribution of matter. Non-electromagnetic interacting massive matter amplifies the gravitational collapse while it

does not contribute to the radiation pressure. It therefore leaves a different signature in the power spectrum than baryonic matter. Through this it was possible to estimate the content of baryonic and non-baryonic matter. The most common model states that baryonic matter only contributes around 15% of the total gravitational mass in the universe [5].

2.1.4 Structure Formation

The size of regions with over-densities depends on the speed of sound in the medium. For a large speed of sound small over-densities are almost instantly washed out by pressure changes. Therefore, the speed of sound in the universe before and after decoupling defines the size of the structures which are able to be formed. Without dark matter the very high speed of sound in the primordial plasma prevents any structures from forming before decoupling. This is, however, not consistent with observations. For a baryon-radiation-plasma the speed of sound is determined from radiation. Thus, any non-radiating matter reduces the speed of sound. Structure formation in the universe can therefore be explained through the existence of dark matter. Additionally, the speed of the dark matter particles also influences the speed of sound and therefore determines the size of structures which are able to form first. For relativistic moving (hot) dark matter the larger speed of sound allows very large structures, with the size of superclusters, to form first (top down scenario). For slow moving (cold) dark matter smaller structures, like galaxies, form first (bottom up scenario). Since the forming of superclusters is observed only now the cold dark matter model is heavily favored.

2.2 Dark Matter Candidates

There are many models attempting to explain the particle character of dark matter. They can be classified into baryonic and non-baryonic matter models. Baryonic matter is part of the Standard Model of particle physics while non-baryonic dark matter requires an extension of the Standard Model. Another distinction can be made between cold and hot dark matter. Cold dark matter refers to particles which move slow compared to the speed of light while hot dark matter moves at (ultra) relativistic speeds. Cold dark matter models are favored because of the better agreement of structure formation predictions with observations.

2.2.1 Baryonic Dark Matter

One explanation for the excess in non-luminous mass in the universe is the existence of massive compact halo objects (MACHOs). These massive objects consist of normal baryonic matter and emit little to no light. Common examples include black holes, neutron stars and brown dwarfs. However, observations show that such objects only contribute roughly 8% to the galactic mass. Furthermore, large amounts of non-baryonic mass are needed to explain various measurements, among them CMB observations.

2.2.2 Sterile Neutrinos

The weak interaction only couples to left-handed particles and right-handed antiparticles. Neutrinos are leptons without an electric charge and do not partake in the strong and electromagnetic interaction. Right-handed Neutrinos do not interact weakly and are called sterile neutrinos. In the Standard Model neutrinos are massless. However, observed neutrino oscillations are only possible if they have mass. The mass term can then propagate with a different frequency than the flavor eigenstate and the neutrino can oscillate between states. Massive sterile neutrinos could also explain the extremely low mass of active (left-handed) neutrinos, compared to other leptons, through an effect called seesaw mechanism. This effect states that through mass mixing of the Majorana- and Dirac-massterms the light active neutrino gets lighter as the sterile neutrino gets heavier. Massive sterile neutrinos are therefore a candidate for dark matter. Conventionally sterile neutrinos are considered hot dark matter. Recently however models have been found in which they can also act as cold dark matter [6].

2.2.3 Axions

Axions are hypothetical particles which arise from an additional symmetry that was introduced as a possible solution to the strong CP problem. Axion radiation transports energy out of stars and supernovae which limits the coupling strength of axions. Axions are non-thermal produced particles. They can be converted into photons and vice versa in the presence of a magnetic field, via the so called Primakoff effect. This effect is used to search for axions, however none were found so far.

2.2.4 Weakly Interacting Massive Particles

Weakly interacting massive particles (WIMPs) is the name for a group of hypothetical particles which, as their name suggests, only interact through the weak and the gravitational interaction. Depending on the model WIMPs can act as either cold or hot dark matter. They are currently the most popular dark matter candidate, which is in part the result of an effect called the WIMP miracle and their prediction through various extensions of the Standard Model.

The WIMP Miracle

Using the Boltzmann equation it is possible to determine the relic density of a weakly coupling particle with mass m. The Boltzmann equation in this case is

$$\frac{dn_\chi}{dt} = -3Hn_\chi - \langle\sigma_{ann}v\rangle\left(n_\chi^2 - n_{eq}^2\right) \tag{3}$$

with the Hubble constant H, the equilibrium density n_{eq}, the WIMP density n_χ and the average thermal WIMP annihilation cross section $\langle\sigma_{ann}v\rangle$. The Hubble constant is needed to account for the expanding universe, which reduces the density over time. For a relic density, which matches the observed dark matter density today, the Boltzmann

equation leads to a particle with a mass and cross section in the scale of the known weak interaction. This alignment is whats referred to as the WIMP miracle.

Prediction Through Supersymmetry

Another argument in favor of WIMPs is the prediction of WIMP-like particles through supersymmetric extensions of the Standard Model. The minimal supersymmetric model (MSSM) is an extension of the Standard Model which allows to explain many of the open questions in modern physics. In MSSM every Standard Model particle has a supersymmetric partner. In this extension baryon and lepton numbers are no longer conserved and the supersymmetric partners would have a mass identical to their Standard Model counterparts. However, baryon and lepton number conservation is tested very precisely, for example through the lifetime of the proton, which is at least 10^{33} years. Without baryon and lepton number conservation the proton would decay fast. Therefore, a new conserved parity, called R-parity, is introduced, which breaks this symmetry and strongly suppresses proton decay. As another consequence of this the lightest supersymmetric particle, the neutralino, cannot decay into Standard Model particles, which would make it a prime dark matter candidate. Due to their high mass and therefore non-relativistic behavior neutralinos are cold dark matter candidates. However, supersymmetric theories are under pressure because of recent particle accelerator experiments, in particular at the LHC, which have already excluded some supersymmetric models.

2.3 Search for Dark Matter

Due to various astrophysical observations most dark matter models assume particles that do not partake in the strong or electromagnetic interaction. However, to be successfully detected dark matter has to interact more than just through gravitation. Therefore, most dark matter searches focus on detection through interaction of dark matter particles with Standard Model particles through a weak interaction. The Feynman diagram of this interaction, shown in figure 3, allows to illustrate three different detection approaches.

2.3.1 Indirect Detection

In indirect searches dark matter is detected through its signature when annihilating into Standard Model particles. The type and strength of the expected signal is highly dependent on the properties of the particle as well as astrophysical assumptions. Indirect search is mostly used to detect WIMP-like particles in the high mass region, where this detection method is most sensitive. The model for the WIMP freeze-out provides estimates for the expected cross section and mass. Possible decay channels include all Standard Model particles, even photons through secondary decays. All regions with sufficiently high DM density are possible sources for decay signals, for example the center of the milky way, dwarf spheroidal galaxies outside of the galactic plane or even the sun. A few interesting signals have been found this way, however due to very large

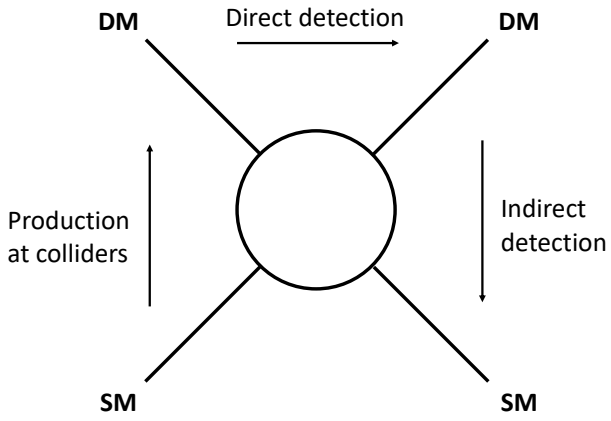


Figure 3: Feynman diagram showcasing the possible approaches for dark matter detection.

uncertainties in the astrophysical models or weak signals clear interpretation as dark matter annihilation was not possible.

2.3.2 Production

Dark matter search through production in particle colliders utilizes the same axis in the Feynman diagram as indirect detection, but with inverted time. This method is particularly sensitive for low-mass dark matter particles. Detectors in particle physics utilize electromagnetic interaction for detection. Since DM does not interact electromagnetically it is only detected indirectly through missing energy. It is also possible to search for dedicated dark matter decay chains. However, it is not possible to prove the lifetime of the produced particles and since the lifetime of DM-particles should be in the order of the age of the universe it is only possible to find DM-particle candidates with this method. So far no signal has been found using this search strategy.

2.3.3 Direct Detection

If dark matter interacts weakly it can transfer energy and momentum to Standard Model particles via scattering. According to the models of the dark matter distribution in our galaxy the earth is located inside a dark matter halo. Therefore, it should be possible to observe dark matter interactions with earth-based detectors. Since the interaction of dark matter with Standard Model particles is expected to be very minimal the detectors

need to be sufficiently sensitive. A model for the expected signal as well as the background is required to quantify the measurements. The expected signal consists of two contributions:

- The velocity distribution and density of the dark matter particles at the point of the detector, which depends on the mass of the particle as well as the distribution of dark matter
- The strength and type of interaction of the dark matter particles with the detector, depending on the cross section of the interaction, the mass of the dark matter particle as well as the detector material

These contributions are explained in more detail in chapter 4.1. The expected dark matter signal is very small, therefore an effective shielding is required to reduce the background, which consists of Standard Model particles interacting with the detector. To distinguish dark matter from Standard Model events many detectors utilize two different detection channels. This allows to distinguish between electromagnetic interacting Standard Model particles and non electromagnetic interacting WIMPs. However neutrons cannot be distinguished this way, therefore special care has to be taken to reduce and estimate the neutron background. The CRESST experiment is presented in section 3 and serves as an example for a direct dark matter detection technique.

3 The CRESST Experiment

The Cryogenic Rare Event Search with Superconducting Thermometers experiment is located at the Gran Sasso underground laboratory (LNGS) in Italy. The aim is the direct detection of WIMPs though elastic scattering off nuclei. An introduction into the experimental setup as well as its capabilities are presented here. A schematic of the experimental setup as well as the shielding structure and the cooling mechanism is presented in figure 4.

3.1 Experimental Setup

As a reduction of the background is very important various shielding techniques are employed to prevent or at least reduce certain backgrounds. The experiment is located at an underground laboratory inside the Gran Sasso massif where a minimal overburden of over 1400 m of rock provides shielding against high energy cosmic radiation and muons. While muons cannot be shielded entirely the muon rate at the detector is only one muon per hour per square metre, which is equivalent to a reduction in the order of 10^{-6} compared to sea level [7]. The remaining muons are taged through plastic scintillators covering almost 99% of the solid angle around the detector. If a detector event coincides with a muon veto event it is rejected. Most of the remaining radiation comes from natural radioactivity present in the rock, mainly through decays of ^{40}K, ^{232}Th, ^{238}U and members of their decay chains. α- and β-radiation are already blocked by thin layers of metal. However, γ-radiation and bremsstrahlung x-rays, which are emitted when β-radiation is absorbed, have large penetrating power. Therefore a thick lead shielding surrounds the detectors, due to its high Z lead is especially capable of stopping γ-radiation. Unfortunately, radio-pure lead cannot be produced. Therefore, a second shield consisting of radio-pure copper is installed in order to block radiation from the lead shielding.

Gaseous ^{222}Rn, which is an intermediate member of the ^{238}U decay chain, is present in the air of the laboratory and could penetrate to the insides of the shields. To prevent this, the shields are encased in a gas-tight container, which is under slight overpressure and constantly flushed with radioactively clean nitrogen gas.

As already mentioned in chapter 2.3.3, neutrons can imitate the signal of WIMPs. Therefore, the reduction of the neutron background is especially crucial. Because of its high capability of moderating neutrons hydrogen is used in the neutron shielding to reduce the energy of free neutrons below the detection threshold. Polyethylene ($[–CH_2–CH_2–]_n$) with its high hydrogen content is used for most of the neutron shielding. Only in small parts, where installation of polyethylene was not possible, water is used. A primary shielding of around 50 cm polyethylene is installed around the previously mentioned lead shielding to moderate external neutron radiation. However, since neutrons can also be produced further inside of the shielding layers, for example through the decay of light, through fission of heavy nuclei or through muon interactions, another polyethylene shielding is installed inside of the experimental volume.

The experimental volume is cooled to ≤ 10 mK using a commercial ^{3}He/^{4}He dilution refrigerator. The cryostat is not made out of radio pure materials and is therefore located outside the shielding layers. The cooling power is transferred to the experimental volume through a copper rod called "cold finger", which also holds the experimental volume in place.

3.2 Detectors

The detectors used in the CRESST experiment use scintillating $CaWO_4$ crystals as their target material. As mentioned in section 2.3.3, most modern dark matter detectors utilize two detection channels to separate particles based on whether they interact electromagnetically or not. The CRESST detectors are also designed this way and measure energy deposition using:

- Heating of the crystal through energy deposition caused by elastic scattering of particles on the atoms of the target material
- Scintillation light emitted when (through scattering) excited electrons relax into a lower energetic state

Electrically charged particles and gammas predominantly interact with the shell of the atoms and therefore cause much more scintillation light, compared to the thermal energy they deposit, than neutral particles, which practically only interact with the atomic nuclei. Using the ratio of emitted scintillation light to deposited phonon energy it is possible to determine the type of particle on an event-by-event basis. It has to be noted that even for gammas, which relatively produce the most scintillation light, only around 1-2% of the total deposited energy is detected as scintillation light. Therefore, the phonon energy is used to measure the total energy deposition.

The temperature change caused by a change in phonon energy is approximated as

$$\Delta T \simeq \frac{\Delta E}{C} \tag{4}$$

Since the expected energy deposition is extremely small a very high resolution in the thermometer as well as a low heat capacity C is required. For dielectric materials at low temperatures the heat capacity is proportional to T^3. As mentioned earlier the detectors are operated at 10 mK, which reduces the heat capacity of the crystal by around 13 orders of magnitude compared to room temperature. Nonetheless, the change in temperature is still very small in the order of μK/keV. The thermometer used in the CRESST detectors to detect this minimal change is a so-called transition edge sensor (TES). It utilizes the relatively large change in resistance of a material at its transition point between superconducting and normal state. A transition curve for a TES used in the CRESST experiment is shown in figure 5. The TES used for the detectors consist of a thin layer of tungsten which is evaporated onto the $CaWO_4$ crystal and a small heater used to keep the TES in a stable operating point.

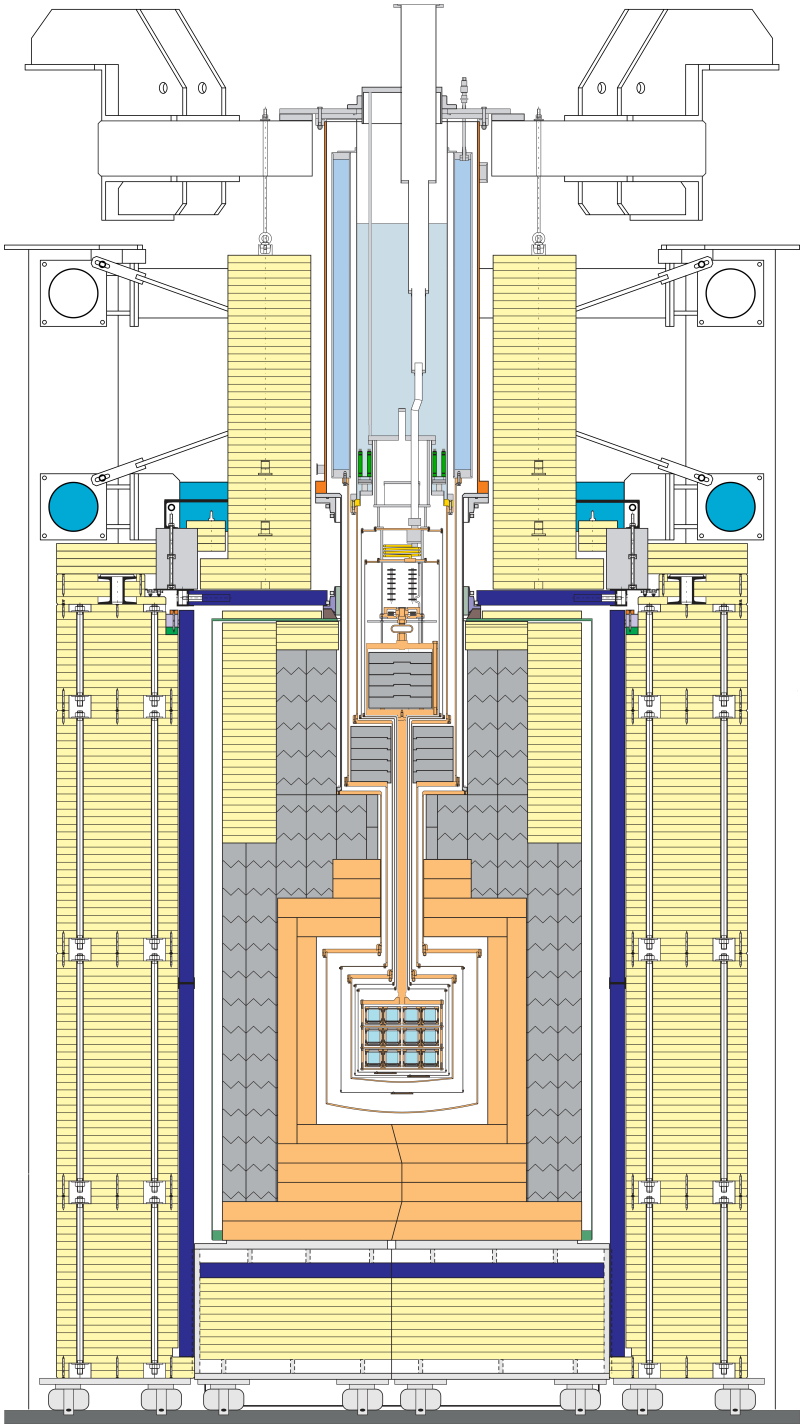


Figure 4: A drawing of the CRESST setup including the onion-like shielding structure and the dilution refrigerator.

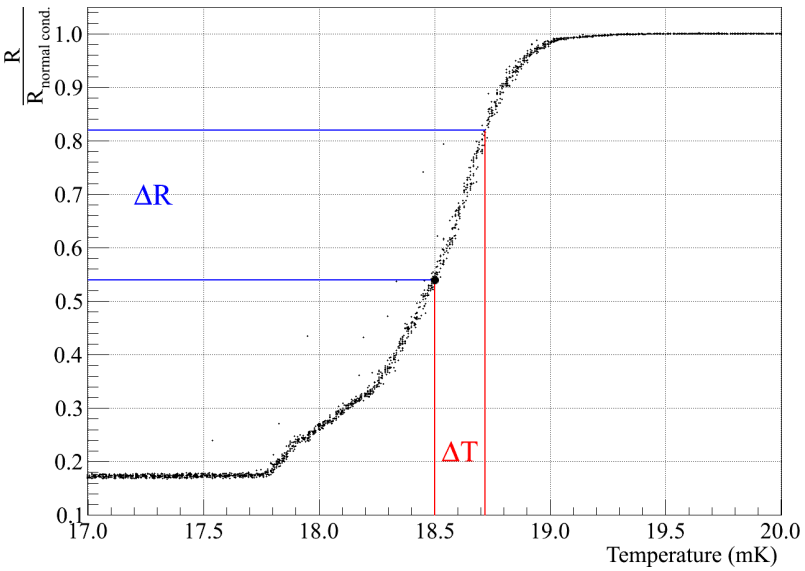


Figure 5: The transition curve of a TES used in detector TUM40 for the CRESST experiment [8].

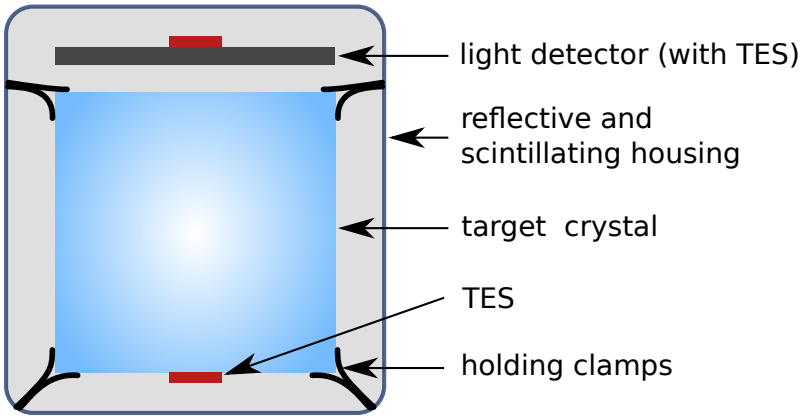
The scintillation light is detected through a light absorber with another tungsten transition edge sensor evaporated onto it. The absorbers are made out of a sapphire waver with a thin silicon layer. Sapphire is chosen because it absorbs light with wavelengths similar to those of the scintillation light especially well. If a photon which is emitted as scintillation light in the detector crystal is absorbed in the silicon-on-sapphire it slightly heats the absorber, which is measured using the TES. The measured resistance change is proportional to the total emitted scintillation light. Simplified drawings of the detector setup for different generations is shown in figure 6.

While there are a large variety of detectors in the CRESST experiment this work focuses on three detectors chosen specifically because of their characteristic features and properties. These detectors along with their peculiarities are presented in the following.

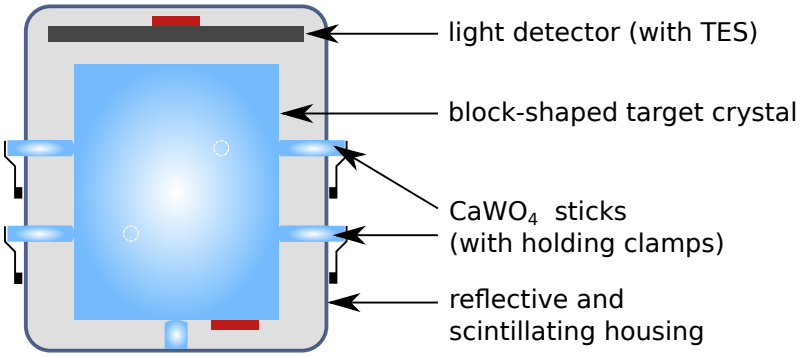
3.2.1 Detector TUM40

Detector TUM40 consists of a block-shaped crystal with a height of 40 mm, an edge length of 32 mm and a mass of 248 g which was grown at Technische Universität München. It was used in CRESST-II Phase 2. This detector is chosen because of its overall good properties, namely low background level ($3.51\frac{\text{counts}}{\text{kgd}}$), low trigger threshold (603 eV) and good energy resolution for phonon and light detector [9][8].

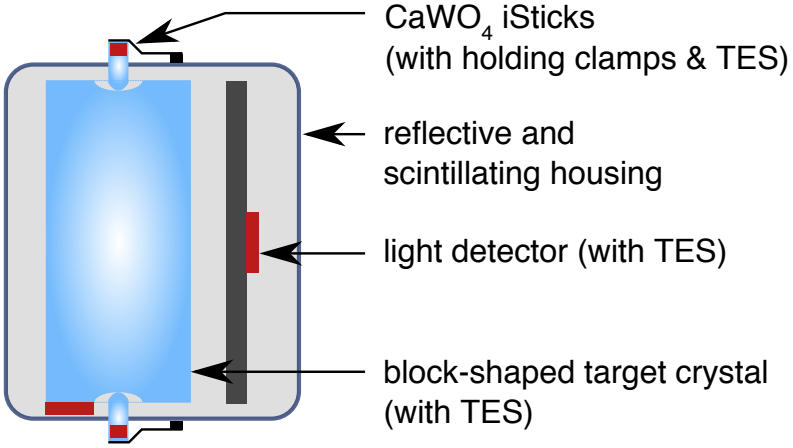
The measurement data on which the calculations are performed was collected in 2013 with a total exposure of 29.35 kgd (before cuts).



(a) A schematic representation of a CRESST-II detector with holding clamps, an example for such a detector would be LISE.



(b) A schematic representation of a CRESST-II detector with holding sticks, an example for such a detector would be TUM40.



(c) A schematic representation of a CRESST-III detector.

Figure 6

3.2.2 Detector Lise

The Lise detector was also used in CRESST-II Phase 2. It is a commercial crystal with a mass of 306 g in cylindrical shape with a height and radius of 40 mm. Unlike TUM40 this detector displays relatively large background levels and the resolution of the light detector is low. However, the trigger threshold of 307 eV is the lowest for all phase 2 detectors [10]. Overall this detector behaves quite differently than the TUM40 module and together these two detectors are good benchmark cases for describing the measurement data.

The data used in this work was collected during CRESST-II Phase 2 with a total exposure of 52 kgd.

3.2.3 Detector A

The detector A module is a new model of CRESST-III Phase 1. The detector design was optimized to achieve lower thresholds, below 100 eV, and, therefore, a sensitivity for even lower dark matter particle masses. Of all new detector modules detector A reaches the lowest threshold of only 30.1 eV and therefore is the new benchmark for low-mass dark matter search [11]. Additionally this detector module displays a very low level of background and has good energy resolution. The $CaWO_4$ crystal of detector A weighs 23.6 g and has a size of 20x20x10 mm..

The data used for the calculations in this work was collected between 10/2016 and 01/2018 with a total exposure of 5.689 kgd.

4 Calculating an Upper Limit on the Dark-Matter-Particle-Nucleus Scattering Cross-Section

If no dark matter signal is measured or too weak for a positive discovery one calculates an upper limit on the cross section on the potential dark matter particle for a given mass. This is based on the idea that a DM particle with a larger cross section has a higher chance of interacting with the detector and would therefore lead to a stronger signal, which at some point is no longer compatible with the measured data. For this approach a precise estimate of the density and velocity distribution of the potential DM particles at the point of the detector as well as knowledge of the interaction with the detector is required, except for the cross-section.

4.1 Calculating the Expected Dark Matter Signal

4.1.1 Dark Matter Distribution

The expected dark matter density and velocity distribution at the point of the detector can be estimated through astrophysical observations. For DM search experiments the conventionally used model assumes thermalized DM particles, which follow a Maxwell-Boltzmann velocity distribution in the galactic rest frame. For a spherical and isothermal galactic halo the density takes the form:

$$\rho\left(r\right) = \frac{\rho_c}{1 + \left(\frac{r}{r_c}\right)^2} \tag{5}$$

Using astrophysical observations the expected DM density at Earths position for this parametrization is estimated as:

$$\rho_\chi := 0.3 \frac{\mathrm{GeV/c^2}}{\mathrm{cm^3}} \tag{6}$$

This value is used in practically all DM searches to allow for easier comparison between experiments

The velocity distribution of the DM particles compared to the galactic rest frame is given as:

$$f_g\left(\vec{v}_G\right) = \frac{1}{N}\left(\frac{3}{2\pi w^2}\right)^{\frac{3}{2}} \exp\left(-\frac{3v_G^2}{2w^2}\right) \tag{7}$$

where $\vec{v}_G$ is the velocity of the particle compared to the rest frame, w is the root mean square velocity of the WIMP and N is the normalization factor. To determine the expected DM signal the integral of f_g has to be evaluated. The lower integration limit is defined through the minimal velocity v_{min} for which a DM particle can induce a nuclear recoil with a recoil energy of E_R on a nucleus N:

$$v_{min} = \sqrt{\frac{E_R m_N}{2\mu_N^2}} \tag{8}$$

The upper limit is determined through the galactic escape velocity v_{esc} over which nothing is expected to remain in the galactic halo. Therefore, no DM-particles with a speed higher than the escape velocity are expected to be measured. A value of 544 km/s for the escape velocity is used in this work. Using this, the normalization factor can be determined:

$$N = \operatorname{erf}(z) - \frac{2}{\sqrt{\pi}} z \exp\left(-z^2\right) \quad \text{with} \quad z^2 := \frac{3v_{esc}^2}{2w^2} \tag{9}$$

The velocity of the earth relative to the galactic rest frame $v_\oplus$ has to be taken into account. This is an overlap of earth's annual travel around the sun and the suns movement around the galactic centre.

$$v_\oplus(t) = 220 \left[1.05 + 0.07 \cos \frac{2\pi (t - t_0)}{1\text{yr}}\right] \frac{\text{km}}{\text{s}} \tag{10}$$

With this the integral can be determined for the masses expected in the CRESST experiment. Since measurements are conducted over a timespan in the order of a year the annual modulation can be neglected. Only the final result is presented here, a derivation can be found in [12].

$$I(v_{min}) = \frac{1}{N\eta} \left(\frac{3}{2\pi w^2}\right)^{\frac{1}{2}} \left(\frac{\sqrt{\pi}}{2} \left[\operatorname{erf}(x_{min} - \eta) - \operatorname{erf}(x_{min} + \eta)\right] - 2\eta \exp\left(-z^2\right)\right) \tag{11}$$

with

$$\eta^2 := \frac{3v_\oplus^2}{2w^2} \tag{12}$$

$$x_{min}^2 := \frac{3v_{min}^2}{2w^2} = \frac{3m_N E_R}{4\mu_N^2 w^2} \tag{13}$$

For w the standard value of 270 km/s [13] is used.

4.1.2 Interaction of Dark Matter Particles with the Detector

WIMPs only interact with regular matter via weak and gravitational interaction and are immune to electromagnetic interaction. Therefore, they only interact with the nuclei of the detector material. This interaction is described using a form factor which corresponds to the Fourier transformation of the nucleus density. For DM experiments the most common parametrization for the nucleus density is the one proposed by Helm [14], which assumes a constant core density convolved with a Gaussian function that models the

decreasing density at the edge of the core. This model however is too simple, especially for heavy nuclei. For this reason, a more sophisticated model was used.

For Calcium and Oxygen so-called model-independent form factors are used. They have been proposed by G. Duda et al. [15] and postulate the density as a sum of N Bessel functions. This leads to a form factor of:

$$F(q) = \frac{sin(qR)}{qr} \frac{\sum_{v=1}^{N} \frac{(-1)^v a_v}{v^2\pi^2 - q^2R^2}}{\sum_{v=1}^{N} \frac{(-1)^v a_v}{v^2\pi^2}} \tag{14}$$

A suitable choice for N as well as the parameters a_v and radius R have been determined for ^{16}O and ^{40}Ca in [15], however for tungsten no such parameters exist.

Therefore, a different form factor is chosen for tungsten, namely the Woods-Saxon form factor:

$$\rho(q) = \frac{\rho_c}{e^{\frac{r-c}{a}} + 1} \tag{15}$$

for which dedicated parameters exist for ^{184}W:

$$a = 0.535\text{fm}, \quad c = 6.51\text{fm} \tag{16}$$

For this parametrization no analytical Fourier transformation is possible, which would require a slow numerical convolution for every recoil energy q. To avoid this the form factor is precalculated for recoil energies between 0.1 and 300 keV and then interpolated using splines of order 1. This interpolation is then used to estimate the form factor for any recoil energy. The impact of form factors is much stronger for higher recoil energies, therefore form factors play only a minor role in low-mass DM search.

4.1.3 Differential Recoil Rate

Using these properties the differential recoil rate can be calculated:

$$\frac{dN}{dE} = \frac{\rho_\chi}{2m_\chi \mu_N^2} \sigma_\chi F^2(E_R) I(v_{min}) \tag{17}$$

This can be compared to the measured data to calculate an exclusion limit using e.g. the following methods.

4.2 Yellin's Optimal Interval Method

This method was proposed by Yellin [1] [16] as a way to calculate exclusion limits despite the presence of an unknown background by utilizing the difference in distribution of the expected signal to the measured background to find an exclusion limit.

In the following it is assumed that measured events are distributed along an 1D-intervall. Furthermore, the distribution consists of the signal for which the shape of the expected signal but not the total intensity is known, a known background for which

the shape and size is known, and an unknown background. In the case of the CRESST experiment the 1D measured data is the number of events along the energy axis. The dark matter signal is distributed according to the differential recoil rate in equation 17 and the number of expected events is determined by the DM-particle-nucleus cross section as well as the exposure. The background consists of electron, gamma, alpha and neutron events. In this case these contributions are not included as known background in Yellin's optimal interval mode.

The "maximum gap" method utilizes the gap between events along the horizontal axis to calculate a limit. The size of the gap between two measurement points E_i and E_{i+1} is defined by the expected number of events in that interval:

$$x_i = \int_{E_i}^{E_{i+1}} \frac{dN}{dE} dE \tag{18}$$

The maximum gap method now increases the intensity of $\frac{dN}{dE}$ until the largest gap reaches a size where, with a given probability, it is no longer compatible with the observation. Vividly said the number of expected events in that gap is increased until, under a certain probability, it is no longer possible to receive the given measurement result. The probability for a gap smaller than the observed one only depends on the size of the gap x and the expected number of events in that region μ:

$$C_0(x, \mu) = \sum_{k=0}^{\lfloor \frac{\mu}{x} \rfloor} \frac{(kx - \mu)^k e^{-kx}}{k!} \left(1 + \frac{k}{\mu - kx}\right) \tag{19}$$

In this case the sought after confidence for the upper limit is 90% and $\frac{dN}{dE}$ is increased by increasing the DM cross section σ_χ. Since $\frac{dN}{dE}$ is directly proportional to σ_χ at any point E, σ_χ is also directly proportional to x and μ.

For a high density of events the intervals between data points get relatively small, which requires a larger $\frac{dN}{dE}$ to reach an exclusion limit. In this case it is useful to use intervals over which there are especially few events instead of restricting it to none. $C_n(x, \mu)$ now corresponds to the probability that all intervals with $\leq n$ events have $\leq x$ expected events for a given $\frac{dN}{dE}$ and no background. For $n > 0$ it is not analytically determinable, but can be tabulated using Monte Carlo simulations. C_n can then be used the same way as C_0. To avoid an arbitrary choice of n, which could lead to skewed results, Yellin [1] proposed an automatic choice for n.

The largest possible intervals are those with an event or the bounds of the experimental range right before its starting-point and right after its endpoint. C_n is now calculated for all of these intervals and the largest C_n of these intervals is called C_{Max}. For an experiment with N events this requires an evaluation of $\frac{(N+1)(N+2)}{2}$ intervals. The function $\bar{C}_{Max}(C, \mu)$ is now defined in a way that for a fraction C of random experiments with no unknown background $C_{Max} < \bar{C}_{Max}(C, \mu)$. This guarantees that $\bar{C}_{Max}(0.9, \mu)$ always equals at least the 90% confidence level upper limit.

4.3 Extended Maximum Likelihood Method

The optimum interval method is one dimensional and utilizes the different shape of the energy spectrum in the measured data compared to the expected spectrum $\frac{dN}{dE}$. In CRESST, however, the data are available in two dimensions, namely the energy and the light energy. While this property is also used to cut background events from the data used in Yellin's method it enables the use of a maximum likelihood method. It should be noted that, it is also possible to use the optimal interval method in two dimensions [7] or apply a one dimensional likelihood formalism.

4.3.1 The Likelihood Function

The likelihood of a set of parameters for a given model describes the plausibility that these parameters describe the given data. It is therefore strongly connected to the probability which describes the plausibility of an outcome for a given set of parameters. Simply put, the likelihood function is the inverse of the probability function in the sense that:

$$\mathcal{L}(\theta|O) = \mathcal{P}(O|\theta) \tag{20}$$

where O is the set of observed data and θ is the set of parameters for the model. $\mathcal{P}$ is the discrete probability of observing O for the parameters θ. The likelihood function $\mathcal{L}$ is the plausibility of θ fo the given dataset O. This is also applicable for a (non-normalized) density function $\rho(O|\theta)$.

In this case the likelihood function can be seen as the product of the evaluation of the density function ρ at each individual datapoint O_i over all N datapoints:

$$\mathcal{L}(\theta|O) = \prod_{i=1}^{N} \rho(O_i|\theta) \tag{21}$$

4.3.2 Handling Non-Normalized Distributions - The Extended Likelihood

The maximum likelihood method maximizes the likelihood function $\mathcal{L}$ by finding those parameter values θ for which the model best describes the observed data. If the density function ρ is not normalized, but the normalization depends on the parameters θ, it is possible for the likelihood to increase indefinitely. To avoid this, the likelihood formalism can be expanded to include the number of expected events. This is called extended likelihood method. If the number of observed events N is, according to the model, a Poisson distributed random number with mean ν the extended likelihood function can be written as [17]:

$$\mathcal{L}(\theta|O) = e^{-\nu} \prod_{i=1}^{N} \nu f(O_i|\theta) = e^{-\nu} \prod_{i=1}^{N} \rho(O_i|\theta) \tag{22}$$

In this case $f(O_i|\theta)$ is a normalized density function and $\rho(O_i|\theta)$ is a distribution with norm ν.

$$\nu = \int_A \rho(X|\theta)\, dX \tag{23}$$

This introduces a penalty if the number of expected events deviates from the number of actual events in the data.

4.3.3 Considering Machine Precision - The Logarithmic Likelihood

Due to the finite precision in the representation of real numbers on computers a multiplication over many values with varying size might lead to a quite significant error. The accumulation of error can be considerably reduced by using the logarithm of the likelihood and adding up the individual contributions.

$$\ln[\mathcal{L}(\theta|O)] = \ln\left(e^{-\nu}\prod_{i=1}^{N}\rho(O_i|\theta)\right) = -\nu + \sum_{i=1}^{N}\ln[\rho(O_i|\theta)] \tag{24}$$

4.3.4 Calculating an Exclusion Limit

The likelihood can now be used to calculate an exclusion limit on the dark matter cross section depending on its mass. The number of expected events is directly proportional to the DM cross section σ_χ. For a given measurement it is possible that the true dark matter signal fluctuates up- or downwards, resulting in the observed number of events. In the context of this work the 90% upper exclusion limit describes the value of $\sigma_{\chi,excl}$ for which, under a given measurement, the true value of σ_χ is smaller with a probability of 90%. Another way to interpret this is that for 90% of experiments the calculated upper exclusion limit is above the real cross-section.

This is also known as the p-value in statistics which describes the probability that, assuming the model is true, the measurement is equal to or greater than a certain value. For a general pdf $\rho(x)$ p is the integral of ρ over the region from n to ∞:

$$p = \int_n^\infty \rho(x)\, dx \tag{25}$$

For Gaussian distributed random events this corresponds to:

$$p = \int_n^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{(n-\mu)^2}{2\sigma^2}\right)\right] \tag{26}$$

Where erf is the error function. The next step is to calculate the p value from the likelihood. The p value thus needs to be converted to the significance Z:

$$Z = \Phi^{-1}(1-p) \tag{27}$$

Φ is the cumulative of the distribution and Φ^{-1} is known as the quantile of the distribution. For the standard Gaussian distribution Φ^{-1} takes the form:

$$\Phi^{-1}(p) = \sqrt{2}\,\mathrm{erf}^{-1}(2p-1) \tag{28}$$

A visualization of Z and the p value for a Gaussian distribution can be seen in figure 7

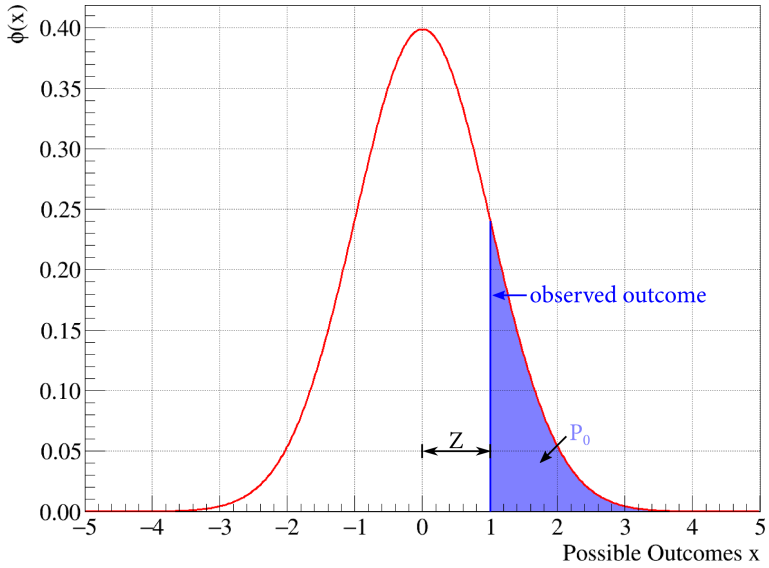


Figure 7: A visualization of the relation between the Z and the p value for a Gaussian distribution. A Z value of 1 corresponds to a (right-sided) p value of 0.1587 which is the integral of the (normalized) distribution from $\mu + (Z \cdot \sigma)$ to infinity, where μ is the mean and σ the standard deviation of the distribution (0 and 1 respectively for this plot). The figure is taken from [8] and follows the formalism used by Cowan et al. [18].

To calculate the pdf ρ the likelihood ratio λ is established, using $\theta = \left(m_\chi, \sigma_\chi, \vec{\vartheta}\right)$. This ratio is also referred to as profile likelihood:

$$\lambda(m_\chi) = \frac{\mathcal{L}\left(m_\chi, \sigma_{\chi,excl}, \vec{\vartheta}_{excl}|O\right)}{\mathcal{L}\left(m_\chi, \sigma_{\chi,best}\vec{\vartheta}_{best}|O\right)} = \frac{\mathcal{L}_{excl}}{\mathcal{L}_{best}} \tag{29}$$

$\sigma_{\chi,excl}$ corresponds to the exclusion and $\sigma_{\chi,best}$ to the optimal value for the DM cross section. Using λ the parameter q can be calculated:

$$q = \begin{cases} -2\ln(\lambda(m_\chi)) & \lambda > 0 \\ 0 & \lambda < 0 \end{cases} \tag{30}$$

which is a measure for the compatibility of the $\sigma_{\chi,excl}$ model with the measurement data with the best model $\sigma_{\chi,best}$ as reference. The higher q the higher the incompatibility between the two models. q is usually distributed according to a χ^2-distribution (note that in this case χ is not the one used to reference a potential DM particle). Z can be

calculated directly from q if q follows a χ^2-distribution. This cannot be assumed a priori and is examined in section 7.2.

$$Z \simeq \sqrt{q} \tag{31}$$

Using equations 29, 30 and 31 the likelihood value for the exclusion limit can be calculated:

$$Z = \sqrt{-2 \ln \left(\frac{\mathcal{L}_{excl}}{\mathcal{L}_{best}} \right)} \tag{32}$$

for the logarithmic likelihood this equates to:

$$Z = \sqrt{-2 \left[\ln \left(\mathcal{L}_{excl} \right) - \ln \left(\mathcal{L}_{best} \right) \right]} \tag{33}$$

This is solved for $\ln \left(\mathcal{L}_{excl} \right)$:

$$\ln \left(\mathcal{L}_{excl} \right) = \ln \left(\mathcal{L}_{best} \right) - \frac{Z^2}{2} \tag{34}$$

The Z value for a probability of 90% is 1.282.

4.3.5 Discovery Calculation

Another advantage of the likelihood formalism is the capability of a discovery calculation. This is done the same way as for an exclusion, except the best fit is compared to the null hypothesis where no dark matter signal is expected at all. In this case q corresponds to the compatibility of the null hypothesis with the data compared to the best fit. Z corresponds to the probability that the null hypothesis cannot explain the data compared to the best fit model with dark matter and p corresponds to the discovery probability (assuming a correct model!).

$$Z = \sqrt{-2 \ln \left(\frac{\mathcal{L} \left(0, 0, \vec{\vartheta}_{null} | O \right)}{\mathcal{L} \left(m_\chi, \sigma_{\chi,best}, \vec{\vartheta}_{best} | O \right)} \right)} \tag{35}$$

In addition to the much higher probability needed for a successful discovery unknown backgrounds might mimic the signal and make the discovery more difficult. This could lead to a false discovery or a discovery with a higher significance than in reality. In comparison for an exclusion limit an unknown (or not considered) background similar to the signal results in a weaker limit which is a conservative approach.

5 Parametrization of the Measurement Data

As already mentioned in section 3.2 the CRESST detectors measure the signal in two observables, the phonon energy and the scintillation light. A 2D plot of datapoints measured by a CRESST detector is shown in figure 8. To apply the likelihood formalism to the measurement data a suitable model for the distribution of events is required. Such a parametrization is presented here. It is based on an already existing parametrization [8][19] with added improvements for better agreement of the model with the data. Additionally, earlier implementations of a likelihood fit required an acceptance region while this implementation does not. Another improvement is the possibility to include neutron calibration data into the fit, which allows to better constrain the model and therefore put sharper limits on the cross section.

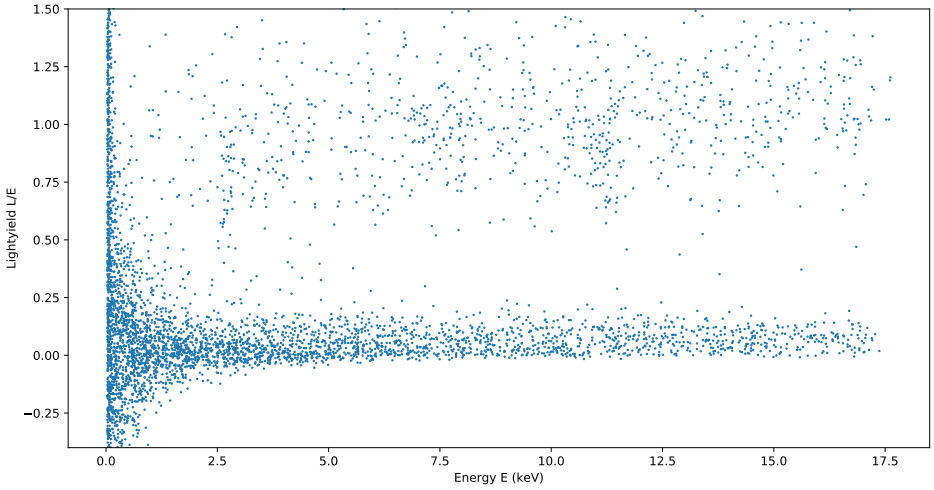


Figure 8: A 2D scatter plot of the detector A measurement and neutron calibration data.

5.1 The Electron-Gamma-Band

5.1.1 Parametrization of the Mean Line

Even with all the shielding described in section 3.1 most of the measured data still originates from gamma or beta background mainly from radioactive decays. The detectors are calibrated in a way that the light-yield (LY), $LY = \frac{L}{E}$, of a 122.1 keV gamma peak from a ^{57}Co source is $LY = 1$. Due to this calibration the mean of the electron-gamma-band is roughly at a light yield of 1. This is described using a linear function, this and the following equations all use energy (E) and light (L):

$$L_e(E) = L_0 E \tag{36}$$

However, since the scintillator does not behave perfectly linear a quadratic term is introduced to better describe this behavior, especially at higher energies.

$$L_e(E) = L_0 E + L_1 E^2 \tag{37}$$

At lower energies a saturation effect known as scintillator non-proportionality bends the electron band towards lower light yields [20][21]. This is taken into account using an exponentially decaying factor. Therefore, the complete description of the mean of the electron band is:

$$L_e(E) = \left(L_0 E + L_1 E^2\right)\left[1 - L_2 \exp\left(-\frac{E}{L_3}\right)\right] \tag{38}$$

This function is shown together with the measurement data in figure 9.

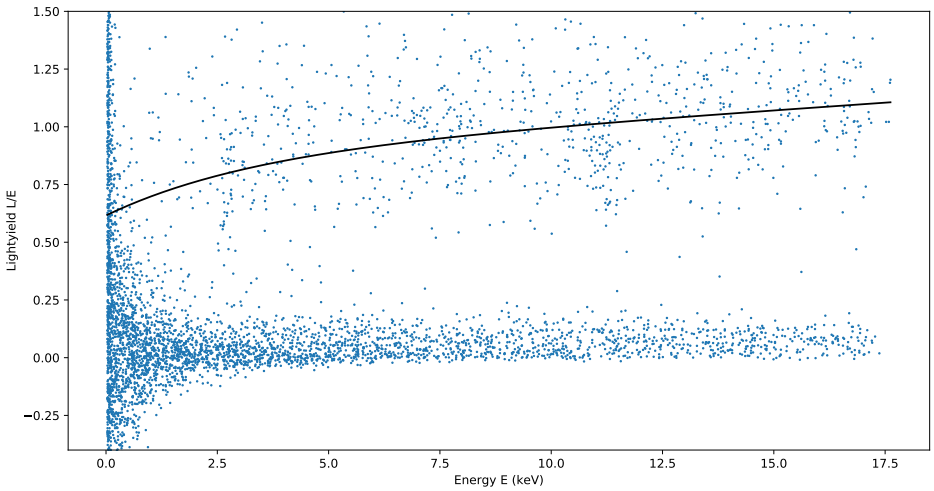


Figure 9: The 2D plot of figure 8 overlaid with the electron band mean light function.

In a first approximation this parametrization can also be used to describe the mean of the gamma-band, as was done in earlier versions of the likelihood fit. This, however, is only a rough approximation since a γ-photon, which interacts with the detector material, produces multiple secondary electrons, each with energies lower than the initial photon. Since the scintillation does not behave perfectly linear the sum of the emitted light from the individual electrons is not identical to the light a single electron with the same total energy would emit [21]. Therefore, the mean of the gamma-band is slightly quenched compared to the electron band. To describe this behavior two different phenomenological approaches have been implemented for the first time.

The first one is a simple linear quenching factor multiplied with the mean of the electron band:

$$L_\gamma(E) = L_e(E)\,[Q_{\gamma,1} + E\,Q_{\gamma,2}] \tag{39}$$

The second method attempts to stay more true to the underlying mechanism by evaluating the electron band mean function at a lower energy. This is again accomplished using a linear function:

$$L_\gamma(E) = L_e\,(E\,[Q_{\gamma,1} + E\,Q_{\gamma,2}]) \tag{40}$$

The idea behind this second approach is that multiple secondary electrons, which are emitted by an initial phonon, have lower individual energies and therefore a lower light yield due to the stronger impact of the non-proportionality effect, which can be approximated by evaluating the electron band at a lower energy. Both of these approaches have been implemented and both performed well. A comparison of the two can be found in section 7.1.

5.1.2 Width of the Bands

The phonon and light detectors have a finite resolution, therefore, the measured datapoints deviate from the mean line. Since the phonon signal is stronger compared to the scintillation light signal the resolution of the phonon detector usually is much better. Therefore, the width of the bands is described in terms of light. This is accomplished using a Gaussian function. The mean is defined by the functions 38 and 39 or 40. The width σ_L is influenced by 3 effects [7]:

1. Uncertainties due to baseline noise in the light detector. This manifests as an energy-independent width $\sigma_{L,0}$.

2. The number of scintillation photons produced can also fluctuate due to the finite energy required to produce a single photon. This effect is expected to be proportional to $\sqrt{L}$ and described with S_1.

3. Other light energy dependent uncertainties, for example a position dependance in the crystal, are considered using a parameter S_2

Together these three factors describe the total resolution of the light detector:

$$\sigma_L\,(L) = \sqrt{\sigma_{L,0}^2 + S_1 L + S_2 L^2} \tag{41}$$

Even if the influence of the phonon detector resolution σ_P is small compared to the light detectors it cannot be neglected. Therefore, a parametrization for the phonon detector resolution is implemented. Just like the light detector resolution it consists of an energy-independent resolution $\sigma_{P,0}$ caused by the baseline noise and a term $\sigma_{P,1}$ which scales with $\sqrt{E}$. This parametrization describes the behavior of the phonon detector well. Therefore, the total phonon detector resolution is:

$$\sigma_P(E) = \sqrt{\sigma_{P,0}^2 + \sigma_{P,1} E} \tag{42}$$

Since the width of the band is described in terms of light, the phonon resolution does not directly affect the width but has to be scaled with the slope of the corresponding band $\frac{dL^x}{dE}$ at that point. In previous iterations of likelihood fits this slope was approximated with 1 for the electron-gamma band and for quenched bands the quenching factor was used as slope. While this is a fairly good approximation for higher energies, where the band behaves very linear, it is not valid at lower energies, where the non-proportionality effect bends the electron- and gamma-bands. For energy-dependent quenching factors this is also not an ideal approximation. For that reason the analytical derivation of the electron band mean line was calculated and used as slope for that band:

$$\begin{aligned}\frac{dL_e}{dE}(E) = &\left[\left(L_0 E + L_1 E^2\right) \frac{L_2 \exp\left(-\frac{E}{L_3}\right)}{L_3}\right] \\ &+ \left[\left(L_0 + 2L_1 E\right)\left(1 - L_2 \exp\left(-\frac{E}{L_3}\right)\right)\right]\end{aligned} \tag{43}$$

Since the gamma-band is expected to be very similar to the electron band the same derivation is chosen, but with the quenching factor applied in the same way as it is done for the mean line. Meaning that for the linear light quenching the slope is just scaled while for the energy quenching the slope is evaluated at the same quenched energy as the mean line. For these implementations the slope of the quenched band is:

$$\frac{dL_\gamma}{dE}(E) \simeq \left[Q_{\gamma,1} + E\, Q_{\gamma,2}\right] \frac{dL_e}{dE}(E) \tag{44}$$

or

$$\frac{dL_\gamma}{dE}(E) \simeq \frac{dL_e}{dE}\left(E\left[Q_{\gamma,1} + E\, Q_{\gamma,2}\right]\right) \tag{45}$$

With these slopes the total width of the bands in terms of light can be written as:

$$\sigma_x(E) = \sqrt{\sigma_L\left(L_x(E)\right) + \frac{dL_x}{dE}(E)\,\sigma_P(E)} \tag{46}$$

With $L_x(E)$ being the mean line function of the corresponding band. The electron band with the upper and lower 90% line for the detector A data is shown in figure 10.

5.1.3 Excess Light Events

In the measured data there is an abundance of events with a higher light yield than expected for electron- or gamma-events. These events are called excess light events and might arise from electrons which produce additional scintillation light in the foil, which surrounds the detector, before they are absorbed in the detector crystal [8].

An empirical model for these events was developed by J. Schmaler [7]. Excess light events are more frequent at lower energies E and decrease in frequency with higher energy. This is modeled by an exponential decrease with an amplitude of El_{amp} at origin

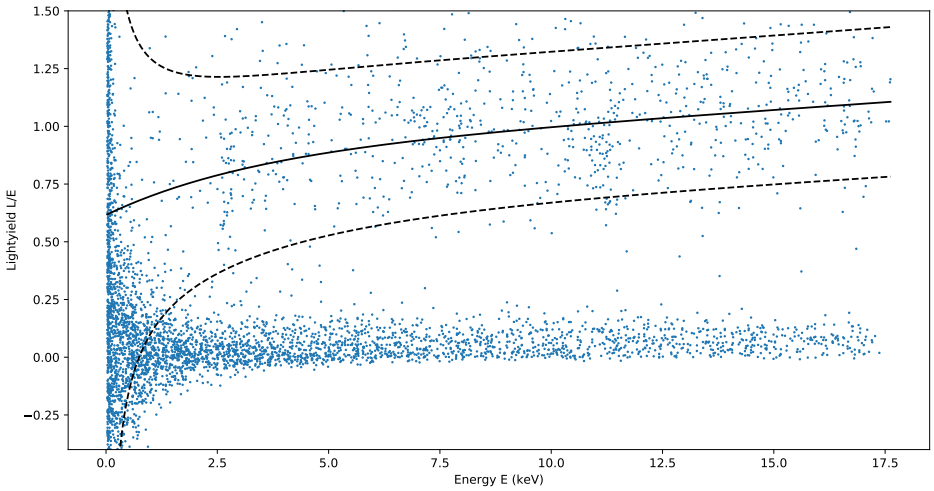


Figure 10: The 2D plot of figure 8 overlaid with the electron band mean light function and the upper and lower 90% lines, 80% of all electron events are expected to occur between the two dashed lines.

and a decay length of El_{dec}. In the original implementation of Schmaler a fraction of the electron-gamma-events X_{el} was used instead of the independent amplitude El_{amp}. However, since the gamma-peaks do not contribute to the abundance of excess light events an independent definition is more sensible, in particular for the current crystals. For the distribution of events along the light axis another exponential decay, with a decay length of El_{width}, was used. This time, however, the decay is not starting at origin but at the mean line of the electron band. This does not account for the resolution of the detector. For this reason J. Schmaler convolved this distribution with the Gaussian resolution of the detector. With this the total parametrization of the excess light distribution is:

$$\begin{aligned}\rho_{excess}\left(E,L\right) &= El_{amp}\exp\left(-\frac{E}{El_{dec}}\right)\\ &\cdot\left(\frac{1}{2El_{width}}\exp\left(-\frac{L}{El_{width}}+\frac{\left(\sigma_{L,e}\right)^2}{2El_{width}^2}\right)\cdot\left[1+\operatorname{erf}\left(\frac{L}{\sqrt{2}\sigma_{L,e}}-\frac{\sigma_{L,e}}{\sqrt{2}El_{width}}\right)\right]\right)\end{aligned} \tag{47}$$

Where $\sigma_{L,e}$ is short for the resolution of the detector at the mean line of the electron band $\sigma_L\left[L_e\left(E\right)\right]$

5.2 Nuclear Recoil Bands

Nuclear recoils, which originate from neutrons or potentially dark matter particles interacting with the nuclei of the detector atoms, emit far less scintillation light than

electromagnetic interactions of the same energy. In addition, the amount of scintillation light produced also depends on the type of atom which partakes in the interaction. The standard CRESST detector crystals are made out of $CaWO_4$, which contains three different chemical elements. Therefore, for the $CaWO_4$ detectors there is a total of three nuclear recoil bands, one corresponding to each chemical element. Since the nuclear recoil bands do not display the non-proportionality effect, the parametrization of the electron band mean line without the non-proportionality correction, see equation 37, is used as basis for the description of the nuclear recoil bands. An energy dependent quenching-factor is applied to this parametrization to produce the mean line of the nuclear recoil band. This description is based on the work of R. Strauss et al. [22], however, it was converted to use a quenching factor as basis of the energy dependent description instead of the light yield at infinite energy. The mean line can therefore be written as:

$$L_x(E) = \left(L_0 E + L_1 E^2\right) \cdot QF_x \cdot \left[1 + f_x \exp\left(-\frac{E}{\lambda_x}\right)\right] \tag{48}$$

Where x represents the chemical element to which the band belongs.

One key discovery by Strauss et al. was that the energy dependent quenching factor is the same for each detector, but scaled with a factor ϵ which is individual for each detector. This means that QF_x, f_x and λ_x can be generally determined for all detectors. This was done by Strauss et al. and the values they determined for the oxygen, calcium and tungsten bands are:

	QF_x	f_x	λ_x
O	0.0739 ± 0.000019	0.7088 ± 0.0008	567.1 ± 0.9
Ca	0.0556 ± 0.00073	0.1887 ± 0.0022	801.3 ± 18.8
W	0.0196 ± 0.0022	0	∞

Therefore, the complete description for the mean line of the nuclear recoil band is:

$$L_x(E) = \left(L_0 E + L_1 E^2\right) \cdot \epsilon \cdot QF_x \cdot \left[1 + f_x \exp\left(-\frac{E}{\lambda_x}\right)\right] \tag{49}$$

With the QF_x, f_x and λ_x values from the earlier table.

For these bands the resolution of the detectors is applied in the same way as for the electron band, again with an analytical derivation $\frac{dL_x}{dE}$ as slope. The nuclear recoil bands along with the measurement and neutron calibration data of detector A is shown in figure 11.

5.3 Description of the Energy Spectra

Now that the band structure is introduced a description for the energy spectra $\frac{dN_x}{dE}$ of the bands is required. For excess light events the differential recoil rate is already

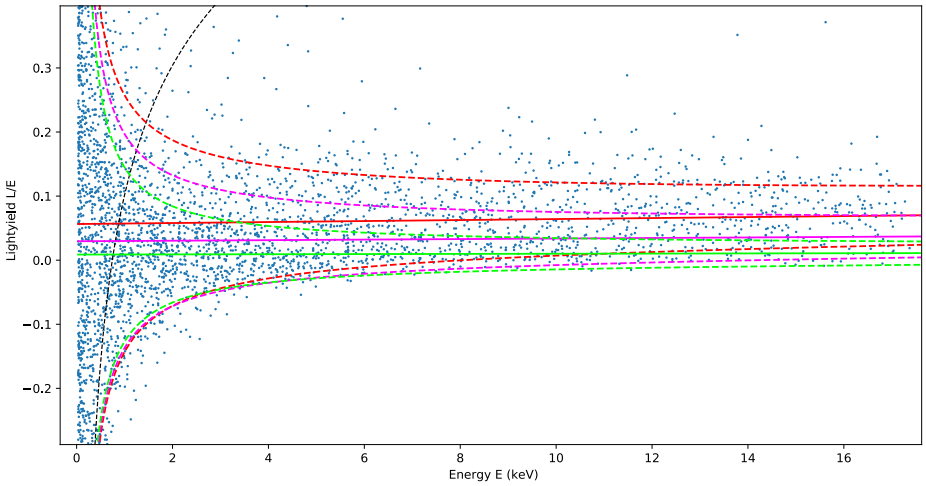


Figure 11: The 2D plot of the measurement and neutron calibration data of detector A overlaid with the nuclear recoil band description. The solid lines are the mean lines of the bands and the dashed lines represent the upper and lower 90% lines. The oxygen band is displayed in red, the calcium band in purple and the tungsten band in green. The lower 90% line of the electron band is shown in black and demonstrates the overlap of the bands at low energies.

included in equation 47. This leaves the spectra of the electron-, gamma- and neutron-backgrounds as well as a description for the dark matter signal and the signal of the neutron calibration.

5.3.1 Dark Matter Rate

The expected energy spectrum for the potential dark matter signal is already established in equation 17. However, this does not account for the phonon detector resolution. The resolution of the detector blurs the dark matter spectrum and also allows true energies which lie below the threshold to get detected. This is taken into account by numerically convolving the dark matter energy spectrum with a Gaussian function with the width σ_P. This is done for recoil energies down to six times the resolution σ_P away from the threshold. Since calculating the numerical convolution for every data and integration point is very time consuming, the dark matter spectrums are pre-calculated for a number of points and then a spline interpolation of order 4 is used to approximate the spectrum. In total 1000 points are used in this interpolation and the density of points is higher for lower energies since the spectrum displays sharper features in that region.

This method is not entirely correct. Therefore, for the newer detector A of run 34 a simulation for the resolution and cut-efficiency[1] replaces this convolution. The simulated data consists of the induced true energy, the measured energy and whether the event survives the cuts. With this data and the calculated spectrum $\frac{dN_\chi}{dE}$ the expected measured spectrum can be constructed. This is done by evaluating the dark matter spectrum at the induced energy and then adding this amplitude to bins according to the measured energy, if the event survives. The bins are then renormalized to account for the number of simulated events and a spline interpolation is used to construct the spectrum from the bins. The impact of the simulated cut-efficiency in comparison to the convolution is studied in more detail in chapter 7.7.

5.3.2 Electron Rate

For the energy spectrum of the electron band a linear spectrum in the form $\frac{dN_e}{dE} = P_0 + EP_1$ was chosen initially. Many detectors exhibit a sharp increase in events at the lower end of their detection range, which as long as there are detectors with a similar or lower threshold without an increase with similar energy dependence, cannot originate from a dark matter signal, which has to have the same energy dependance in all detectors. Therefore, an exponentially decaying term was introduced to the energy spectrum to describe this behavior. Thus, the complete parametrization of this spectrum is:

$$\frac{dN_e}{dE} = P_0 + EP_1 + F_e \exp\left(-\frac{E}{D_e}\right) \tag{50}$$

Due to the high overlap of the electron band with the nuclear recoil bands at lower energies this exponential function can, to some degree, mimic the shape of the dark matter signal. Therefore, special care has to be taken when using this parametrization. This will be further discussed in chapter 7.5.

5.3.3 Gamma Rate

Practically all measured gamma-events originate from radioactive decays inside of the shielding, most from the detector crystal itself. Therefore, the energy of the measured gamma-events should correspond to the characteristic energies of the associated radioactive decays. These measured energies slightly deviate from the real energy due to the resolution of the phonon detector. This is modeled using Gaussian functions with the peak position M_x as mean and the phonon resolution $\sigma_p(M_x)$ as width. The intensity of the peak is incorporated using an amplitude C_x. Therefore, the parametrization for each gamma peak is:

[1] The goal of the raw data selection is to filter false events from the data. However, this also filters some potential signal events. The chance for a signal events to survive this cuts is called cut-efficiency. The cut-efficiency is energy-dependent and can be measured by injecting artificial signal events (simulated events) and measuring fraction of events surviving the cut [8].

$$\frac{dN_{\gamma,x}}{dE} = \frac{1}{\sqrt{2\pi}\sigma_p\left(M_x\right)} \exp\left(-\frac{\left(E - M_x\right)^2}{2\sigma_p^2\left(M_x\right)}\right) \tag{51}$$

5.3.4 Neutron Rate

One special feature of the program developed in this work is its capability to include neutron calibration data into the likelihood function which allows a better fit and agreement of the model with the actual detector behavior. For this reason a description of the differential recoil rate for these type of events is required. The rate of the neutron calibration signal as well as of the neutron background are modeled as an exponentially decreasing function:

$$\frac{dN_{n,x}}{dE} = A_{n,x} \exp\left(-\frac{E}{N_{decay,x}}\right) \tag{52}$$

This implementation is based on previous implementations of a likelihood fit [19].

Since the number of expected neutron background events in the measured data is very small, usually less than 0.1 events per kgd of measurement data, it is not sensible to fit the energy dependence to the background data. Instead, the same energy dependance is used for the nuclear recoil spectrum in the neutron calibration data as well as for the spectrum of the neutron background. Only the amplitudes $A_{n,x}$ are determined separately. Since the neutron calibration is performed with an AmBe source and the background neutrons are a mix of source-like and muon-induced neutrons the energy spectrum of the neutron background can vary a lot from the spectrum in the calibration, but due to the very low number of neutron background events this approximation has to be made. In the future Monte Carlo simulations could replace this implementation and lead to a further improvement of the description. To avoid an overestimation of the neutron background, which would reduce the number of events identified as dark matter interactions and therefore lower the exclusion-limit, the number of neutron events determined by the fit is brought into relation to the expected number of neutron events via a Poisson distribution. The Poisson distribution is included in the likelihood function.

5.4 Construction of the Density Functions

The band structure as well as the energy spectra are now established. With this the distribution functions ρ_x of the individual contributions to the total event distribution in the measurement data is constructed. Except for the excess light distribution, which was already introduced in equation 47, all event distributions consist of a normalized Gaussian function around the corresponding mean line L_x with the width σ_x. This is then multiplied with the related energy spectrum $\frac{dN_x}{dE}$ to describe the expected number of events associated with that band. With this the distribution functions can be written as:

$$\rho_x(E,L) = \frac{\frac{dN_x}{dE}(E)}{\sqrt{2\pi}\sigma_x(E)} \exp\left(-\frac{[L - L_x(E)]^2}{2\sigma_x^2(E)}\right) \tag{53}$$

There are a total of five density functions: ρ_e for the electron background, ρ_γ fo the gamma background, ρ_{nb} for the neutron background, ρ_{ns} for the neutron signal in the neutron calibration and ρ_χ for the dark matter signal. The density functions describing the nuclear recoil bands, namely ρ_{nb}, ρ_{ns} and ρ_χ, actually consist of three separate density functions, one for each chemical element in the detector material, which are summarized into one density function. With this the two total density functions, which are used in the likelihood fit, can be constructed. Two density functions are needed since there are two data-files, one from the neutron calibration and one from the actual measurement. The total density function for the measurement is:

$$\rho_m = \rho_e + \rho_\gamma + \rho_{nb} + \rho_\chi \tag{54}$$

And the one for the neutron calibration:

$$\rho_{ncal} = \rho_e + \rho_\gamma + \rho_{ns} \tag{55}$$

In theory the density functions ρ_{nb} and ρ_χ should also be included in the neutron calibration, but since the exposure of the neutron calibration is extremely low they can be neglected. The description of the band structure is the same in both density functions, only the existence of a dark matter spectrum and the amplitude of the neutron spectrum varies between the two. Therefore, using both datasets and thus both density functions the quality of the fit should improve significantly. Since there are very few events in the nuclear recoil bands in the measurement the inclusion of the neutron calibration should be especially beneficial to the description of these bands. This is validated in chapter 7.3. Newer detector modules, e.g. detector A of run 34, display very few electron background events in the measurement data, which makes it harder to find the correct parameters. Therefore, for these detectors the neutron calibration with more events also improve the fit.

6 Minimizer Algorithms

The parametrization of the bands and the energy spectra require a large number of parameters. Currently, there are 12 detector-independent parameters, for example the ones describing the energy-dependent quenching factors. In the current model another 41 parameters are required for the description of each detector. Although not all parameters are free, for example the exposure is known for each detector, the parameter-space for the fit is large. Furthermore, due to the complexity of the problem the function is expected to have multiple local minima. To cope with this, 5 minimizer algorithms are included in the fit program. The optimization algorithms are part of two packages available for the Julia programming language. The first is called Optim.jl package [23], which is written entirely in Julia and includes a variety of algorithms for uni- and multivariate optimization. A total of four algorithms from this package are implemented into the program: Nelder-Mead, particle swarm, conjugate gradient and Brent's method. The second package is called BlackBoxOptim.jl [24]. It focuses on global optimization of not-differentiable functions. For this it provides (meta-)heuristic/stochastic algorithms. The differential evolution method provided by this package is used in the program. Each algorithm has different strengths and weaknesses and is therefore best used in specific situations. In the following a short overview of the methods is presented along with a guideline when they are best applied. For a more detailed description see the individual references.

6.1 Nelder-Mead

The Nelder-Mead method, also known as downhill simplex method, is a direct search method, which was first proposed by Nelder and Mead [25]. It does not require a gradient, instead it evaluates the function at various points in the search space, this forms a simplex. The algorithm then iteratively replaces the worst point with a better one. To do this, the algorithm can either reflect, expand, contract, or shrink the simplex. This is governed by four parameters, one for each type of operation: α for the reflection, β for the expansion, γ for the contraction and δ for shrinking. The parameters proposed by Nelder and Mead are:

$$\alpha = 1, \beta = 2, \gamma = \frac{1}{2}, \delta = \frac{1}{2} \tag{56}$$

However, a more sophisticated parameter choice, which takes the dimensionality n of the problem into account, was proposed by Gao and Han [26]:

$$\alpha = 1, \beta = 1 + \frac{2}{n}, \gamma = 0.75 + \frac{1}{2n}, \delta = 1 - \frac{1}{n} \tag{57}$$

this parametrization is used for the algorithm. The initial simplex is constructed as a $(n+1)$ dimensional vector filled with n dimensional vectors which are constructed using the start values $\vec{x}$.

This algorithm is very stable and converges to a very exact minimum, however, it requires a starting guess and there is the risk of only finding a local minimum. Therefore,

this method is best suited as a final step when the initial starting point is already pretty close to the expected global minimum.

6.2 Particle Swarm

Particle swarm optimization refers to a group of optimization algorithms which utilize a population of candidate solutions. This population can move around in the search space to find the optimal solution. It is a metaheuristic method and, like the Nelder-Mead method, requires no gradient. This method takes inspiration from flocks of bird searching for the optimal feeding or sleeping location. Each individual particle moves around the search space with a certain velocity. This velocity is influenced by the current position of the particle in relation to the optimal solution found by this particle, the optimal solutions found by the nearest particles as well as the global best positions.

Various implementations of this concept exist. The one used in this work is called "Adaptive Particle Swarm Optimization" and was developed by Zhan et al. [27]. In this implementation each particle can undergo one of four evolutionary states: exploration, exploitation, convergence or jumping out. The jumping out state allows the best particle to move away from its current state which is assumed to be a local optimum. This improves the ability to find a global optimum at the cost of convergence speed.

Due to the nature of this method finding an exact minimum is very slow, however, it is very efficient at finding solutions close to the optimal value. Additionally, it is possible to provide the algorithm with an initial guess for the optimal solution, which can significantly speed up the convergence. Therefore, this method is best suited to find a good estimate of the global minimum when a very rough estimate already exists. Usually, the minimum found by this method is only larger by 0.0001% compared to the best minimum. The convergence speed is very sensitive to the size of the search space. Therefore, it is very beneficial to provide a good choice for the region in which the optimal parameter values are expected.

6.3 Differential Evolution

The differential evolution (DE) method has similarities to the particle swarm optimization, both are population-based search strategies, however, it is not inspired by a biological model. An advantage of the DE method over other evolutionary algorithm is that the distance and direction information from the current population is taken into account. Between generations the population can undergo mutation and crossover operations. In DE algorithms mutation is always applied first to generate a trial vector which is then used in the crossover to create the offspring. Also mutation step sizes are not based on any pre-known assumption.

Simply said, the mutation process allows the candidate solutions to vary and the crossover operator then attempts to combine these candidate solutions in order to find a better solution.

The DE method used in this work is called DE/rand/1/bin. "rand" refers to the random selection of the target vector from the current population. The target vector is

used in the mutation operation to produce trial vectors. This leads to a good diversity in the candidate solutions, which improves the global convergence. However, this also yields a slower convergence speed compared to a methodical selection of the target vector. The number, in this case 1, gives the total number of difference vectors which are used for the determination of the magnitude and direction of the step size. Finally, "bin" refers to the binomial crossover, where the crossover points are randomly selected from the set of possible crossover points. An acceptance probability p_r governs the probability that a crossover point will be included. For a higher p_r more points from the trial vector will be used in the construction of the new candidate solution, while for a lower p_r more points from the previous generation remain. The algorithm uses radius limited sampling. For a more detailed explanation of this method see [28].

This algorithm does only require a search region instead of a starting point. While it initially converges very fast the precise convergence is pretty slow. The method also often converges only to a local minimum. It is therefore best used as a first step to create an initial estimate for the minimum from which a second algorithm takes over.

6.4 Conjugate Gradient

The conjugate gradient method is an iterative method for solving large systems of linear equations. It can be seen as an upgraded version of gradient descent which, for exact computation, converges in n steps, where n is the dimensionality of the problem. As the name already suggests it requires the gradient of the function to be minimized. However, it is not possible to calculate an analytical gradient of the likelihood-function since it depends on the measured data-points. Therefore, to use this method the gradient of the function has to be calculated numerically, which is slower and less precise. This further slows down the convergence. Additionally, this method also only converges to local minima. Because of these drawbacks the method is hardly used in this work, but due to its popularity and robustness it is included in the available minimizers. For a complete description of the method see [29].

The used conjugate gradient method uses the implementation of Hager and Zhang [30][31].

6.5 Brent's Method

Brent's Method combines inverse quadratic interpolation, bisection and secant methods to solve $f(x) = 0$ [32]. The method starts with an interval in which the function changes sign. The method then reduces the interval on one side in each iteration. This means that the method requires only one evaluation of the function in each step and requires no gradient. For this method convergence is guaranteed and usually fast. But Brent's Method can only be applied in one dimension. This means that it has limited use for a high dimensional likelihood function. However, due to its properties of fast guaranteed convergence it is perfectly suited to find an estimate for the exclusion cross section $\sigma_{exclusion}$ by finding the value of $\sigma_{exclusion}$, for which the likelihood is as implausible, compared to the optimal fit, as required.

7 Results

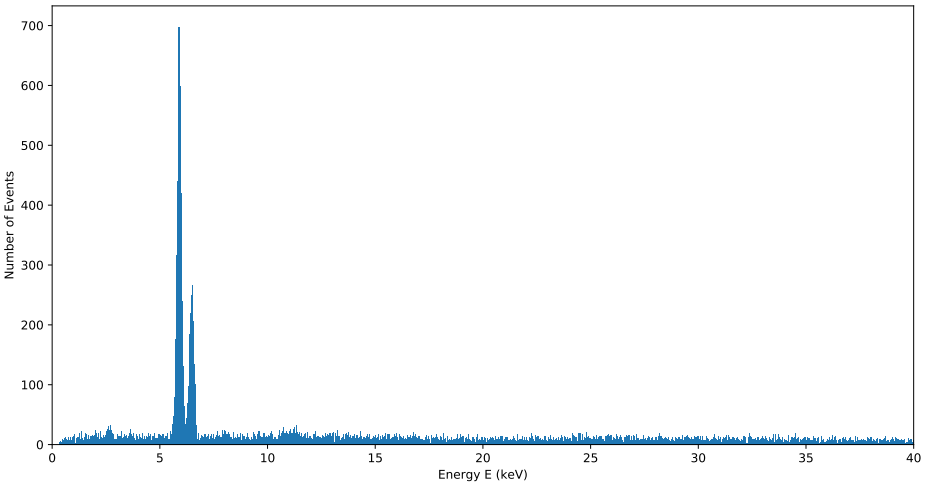
7.1 Comparing the Gamma Quenching Methods

The two gamma-band quenching methods were compared using the CRESST-II Phase 2 (run33) datasets of the TUM40 and Lise detectors. These two detectors were chosen as test cases because of their very different behavior. The Lise detector displays almost no non-proportionality in the electron band, however its background levels are very high and the bands are very wide, which causes a large overlap in the bands. In contrast, the TUM40 detector features well separated electron- and nuclear recoil bands down to very low energies and has a low background level, but displays a very pronounced scintillator non-proportionality. The features in the γ-spectrum of the two detectors are vastly different. While for the Lise detector the gamma spectrum between the threshold and 40 keV is almost entirely dominated by two large peaks at 5.9 (Mn K_α) and 6.48 keV (Mn K_β), the TUM40 detector displays four gamma peaks at 2.59 (^{179}Hf(M1)), 8.04 (Cu), 10.78 (^{179}Hf(L2)) and 11.27 keV (^{179}Hf(L1)). This means that for TUM40 the gamma peaks are distributed over a broader range, which better tests the description of the energy dependence, and for Lise the two peaks close together test the general capability of the quenching method. The energy spectrum for these two detector modules is displayed in figure 12.

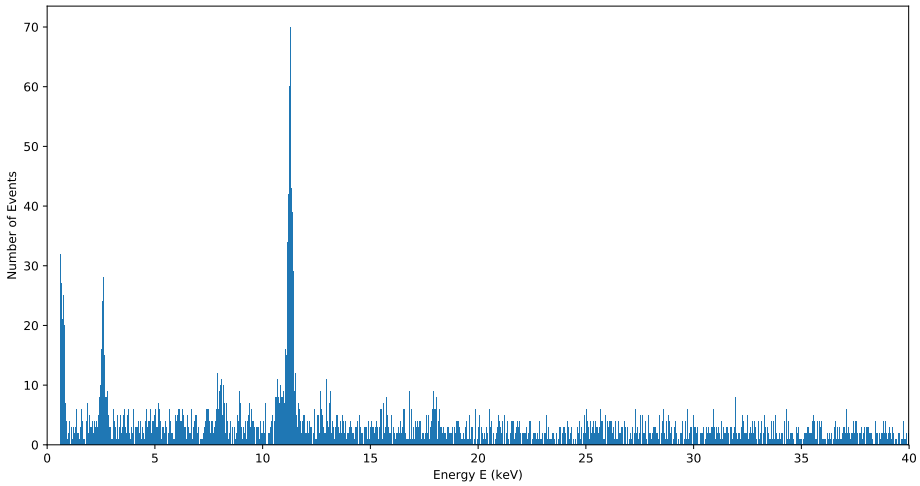
For this test the two detectors were each fitted with both implementations to compare the best possible likelihood obtained through the use of both methods. Since the likelihood is a measure of how good the model describes the data, the implementation which achieves higher likelihoods for the fit is preferable. However, due to the nature of the optimizer algorithms, which are minimizer, the negative likelihood is used in the program. Therefore a lower value is better. To avoid convergence to a local minimum on one of the results the fit is repeated multiple times from scratch. First, a fit without any starting values is applied using the Differential Evolution algorithm. Then the result of that is used as starting point for a fit using the Particle Swarm algorithm and afterwards a final optimization is done with the Nelder Mead algorithm. This procedure is repeated at least two times for every combination of method and dataset to avoid comparing a local minimum with a global one. The results of these tests are shown in the following table:

	TUM40	Lise	TUM40+Lise
L-quenching	-3948.884	-51878.34	-55827.22
E-quenching	-3949.257	-51878.32	-55827.58

Overall, the energy quenching slightly outperforms the light quenching. Considering the fact that the Lise gamma spectrum is completely dominated by two peaks and that a polynomial of second degree can perfectly position these two peaks it is easy to explain why the light quenching is marginally better for this dataset. However, for TUM40, where the peaks are distributed over a larger energy range, the energy quenching method is better. The total likelihood is larger for Lise since there are much more events in the



(a) Lise energy spectrum using 1000 bins for 52 kgd of measurement data.



(b) TUM40 energy spectrum using 1000 bins for 29.35 kgd of measurement data.

Figure 12

dataset due to the higher background levels. Thus, if the relative difference is taken into account the advantage of the energy quenching becomes even more pronounced.

As a result of these comparisons the method which evaluates the electron band at a lower energy is chosen over the light quenching method and is used in all further calculations.

7.2 Verifying the χ^2-Distribution of q

Simulated datasets are used to test whether q follows a χ^2 distribution for this case. The simulation is done with a different implementation of the measurement data parametrization. Therefore, this test also serves as a simple cross check between the two programs. In total 10000 datasets are generated using a Monte Carlo method. The simulated dark matter signal uses particles with a mass of 25.3 GeV/c^2 and a cross section of $1.6 \cdot 10^{-6}$ pb and is simulated for the TUM40 detector with an exposure of 28.548 kgd.

As a first step all datasets are combined into one dataset with a hypothetical exposure of 285480 kgd. With this a maximum likelihood fit is performed to calculate the probable dark matter particle properties. For this case the maximum likelihood is for a DM mass of 25.5 GeV/c^2 and a cross section of $1.627 \cdot 10^{-6}$ pb, which is close to the simulated properties. This small discrepancy can be explained by some minor differences between the two implementations.

The program used for the simulation does not use the derivations of the mean lines as slope. Furthermore, this program uses a spline interpolation of order two to handle the measured cut efficiency while the program used for the simulation uses splines of order one. Additionally, the convolution of the dark matter spectrum with the Gaussians is also implemented differently. As mentioned in chapter 5.3.1, the program developed in this thesis uses a spline interpolation of order 4 to approximate the dark matter spectrum, while the program used for the simulation uses binning for the pre-calculation. Together these three differences are most likely the cause of the small discrepancy.

Finally, an exclusion calculation is performed for a dark matter mass of 25.3 GeV/c^2 for all 10000 datasets individually. This is done to determine whether q is χ^2 distributed and to test if the cross section exclusion limit calculated using these data is larger than the simulated cross section in 90% of the cases. The result of these calculations is presented in figure 13.

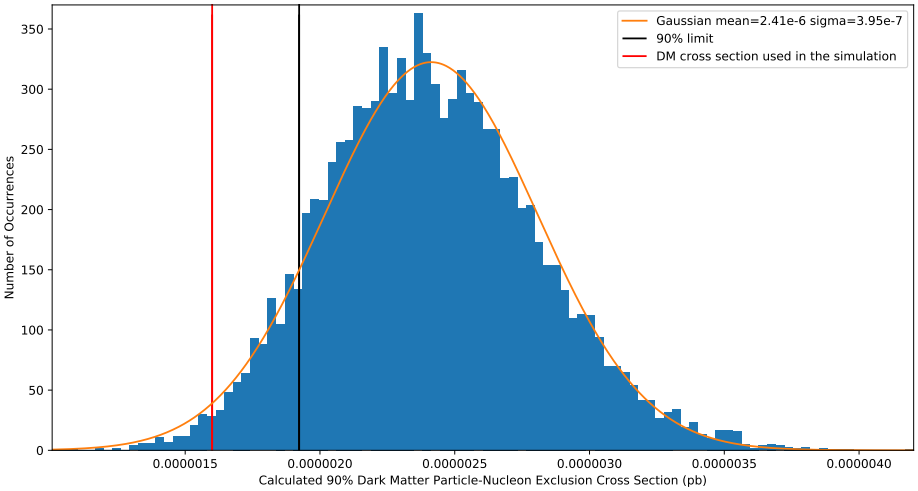
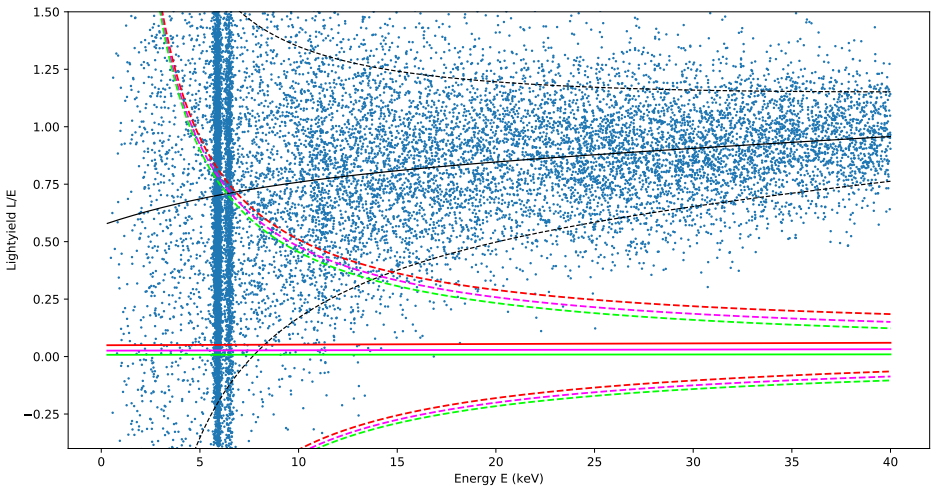


Figure 13: Histogram of the calculated exclusion limit for 10000 simulated datasets overlaid with a Gaussian distribution constructed with the mean and σof the exclusion limits. The black vertical lines represent the limit above which 90% of all exclusion limits lie and the red vertical line represents the simulated cross section.

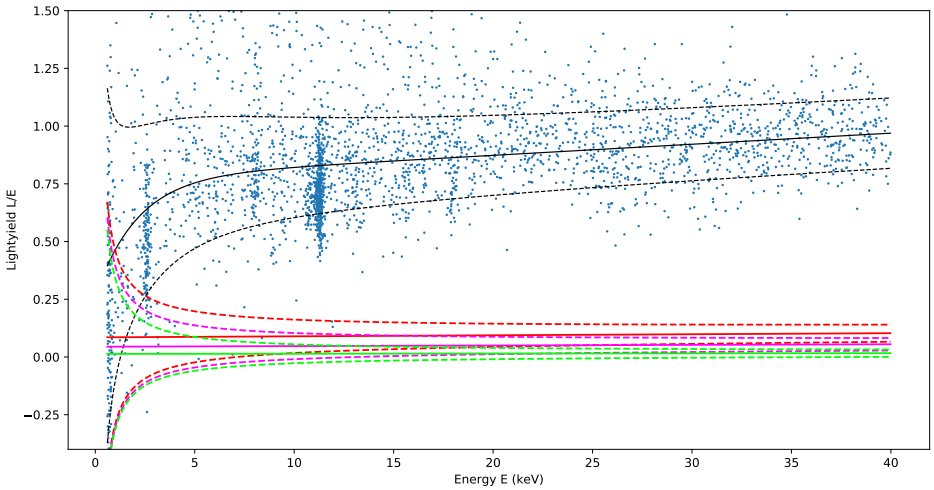
It can be seen that the gaussian and the histogram of the exclusion limits show a very good agreement. Additionally over 90% of all calculated exclusion limits lie above the cross section of the simulated dark matter.

7.3 Impact of Including the Neutron Calibration on the Calculated Exlusion-Limits

As already mentioned in chapter 5.4 the program has the ability to include the neutron calibration data into the likelihood fit to better constrain the parameters, especially the ones describing the nuclear recoil band. To test the impact of this inclusion on the fit and, therefore, on the calculated exclusion limits, the limits calculated using the TUM40 and Lise datasets with and without neutron calibration are compared. These two detectors are chosen again for the same reasons as in chapter 7.1. The parameters obtained in the previous fit are used as starting values for all four cases. Then a global optimization using the particle swarm algorithm and afterwards a local optimization with the Nelder Mead algorithm is performed. The band fit for both detectors is shown in figures 14 and 15.

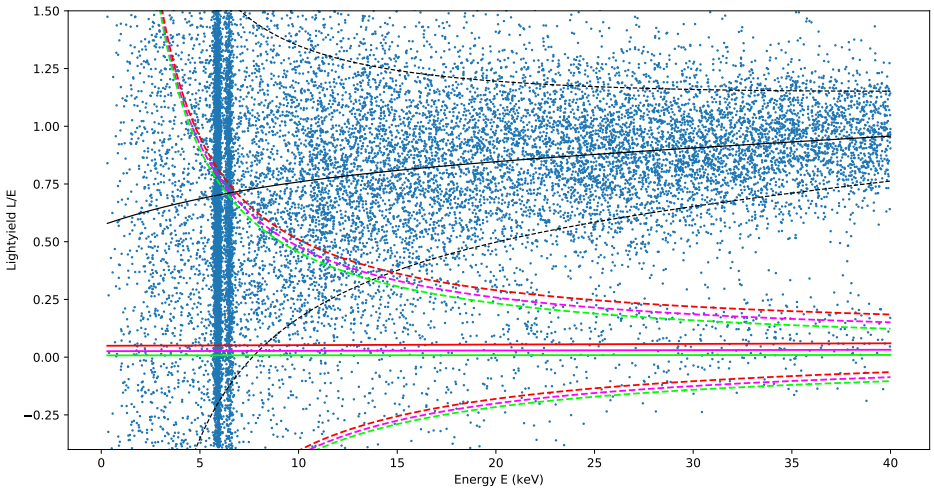


(a) 2D scatter plot of 51.1284 kgd detector Lise measurement data overlaid with the band fit functions introduced in chapter 5.1.

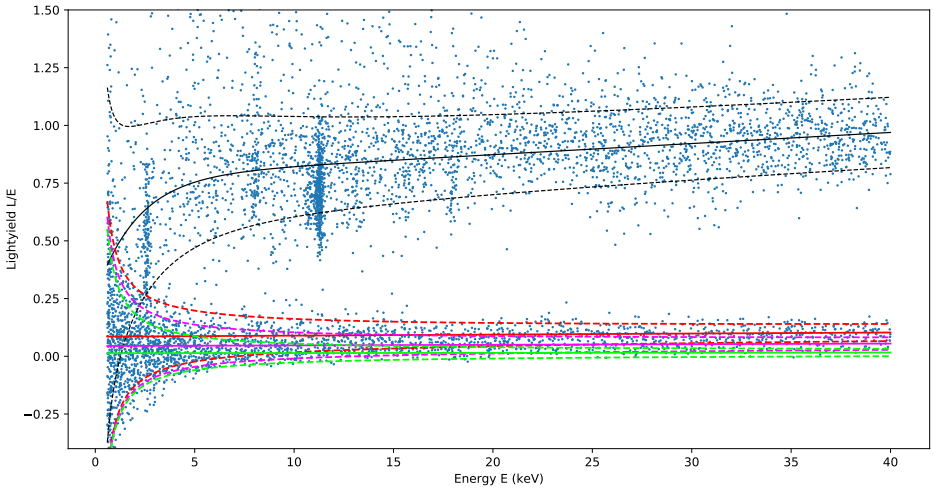


(b) 2D scatter plot of 28.548 kgd detector TUM40 measurement data overlaid with the band fit functions introduced in chapter 5.1.

Figure 14



(a) 2D scatter plot of the detector Lise measurement and neutron calibration data overlaid with the band fit functions introduced in chapter 5.1.



(b) 2D scatter plot of the detector TUM40 measurement and neutron calibration data overlaid with the band fit functions introduced in chapter 5.1.

Figure 15

From the parameters obtained through this fit, limit calculations with Yellin's method as well as with the profile likelihood method are carried out. The Yellin exclusion limits are calculated using the mean of the oxygen band as upper limit on the region of interest and the lower limit is chosen to include the upper 99.5 % of the tungsten band. The calculated limits are shown in figure 16.

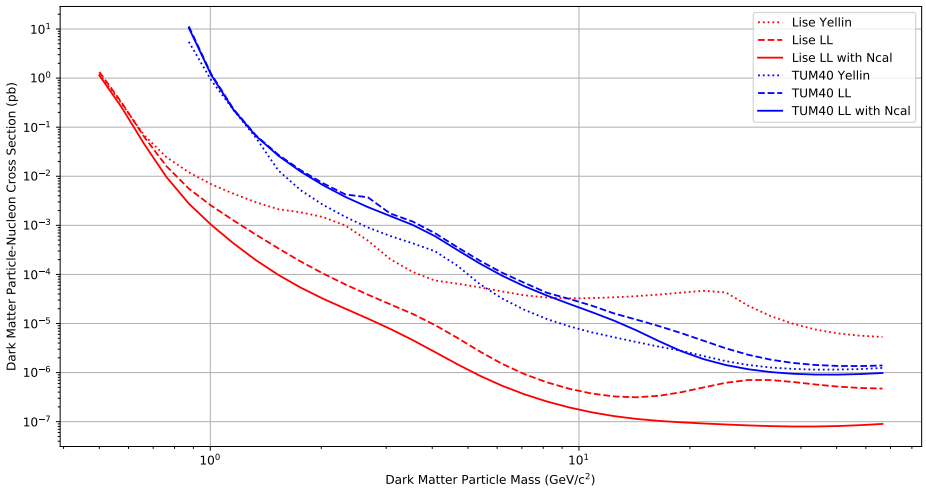


Figure 16: Dark matter particle exclusion limits for the Lise and TUM40 datasets with and without the inclusion of the neutron calibration (Ncal)

It can be seen that the profile likelihood method vastly outperforms Yellin's optimal interval method for the Lise detector data, especially for higher masses. This is expected since the large overlap of the bands leads to a significant leakage of electron-gamma-background to the nuclear recoil bands. The profile likelihood approach can account for this background while the optimal interval method cannot. For the TUM40 data Yellin's method actually outperforms the profile likelihood approach. This is most likely caused by two different effects. First, a large increase in event frequency near the threshold increases the exclusion limit for low dark matter particle masses and second, possible neutron background events increase the limit in the medium dark matter mass region. More on these contributions later in sections 7.5 and 7.6. For both detectors the inclusion of the neutron calibration data results in stronger exclusion limits. Especially the Lise limit benefits from the inclusion of the neutron calibration and is up to one order of magnitude lower than without. Since there are more data-points to evaluate and more degrees of freedom when the neutron calibration is included in the fit it is expected that the calculation requires significantly more time. However, the opposite is the case. The calculations without the neutron calibration data are mostly slower. This is especially apparent for the exclusion calculations, which often require roughly ten times longer without neutron calibration. This indicates that it is easier for the

optimizer algorithms to converge when the neutron calibration data are included. This effect is more prominent in exclusion calculation, where the largest changes between the fit happen in the nuclear recoil bands.

7.4 Combining Detectors

Another innovation of the fit program presented here is its capability to combine the measurement data of multiple detectors into the likelihood fit and limit calculation. The idea behind this is that a combination of detectors results in a higher total exposure and therefore better statistics which then potentially leads to a stricter exclusion limit. Another advantage might be that this method takes advantage of the individual strengths of the detectors. For example the lower threshold of the Lise detector module in combination with the lower background levels and better separated bands of TUM40 could result in a better combined limit. To test this hypothesis the Lise and TUM40 measurement data are both included in a profile likelihood exclusion calculation. The result is then compared to the exclusion limits of the individual detectors. The limits calculated with the inclusion of the neutron calibration are used as benchmark. This means that also for the combined limit calculation the neutron calibration of both detectors is used. The comparison of the individual and combined limits is presented in figure 17.

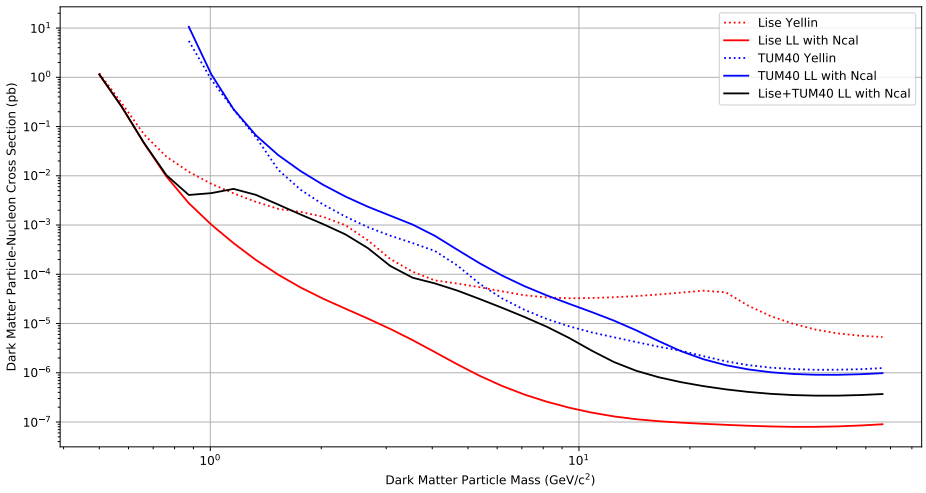


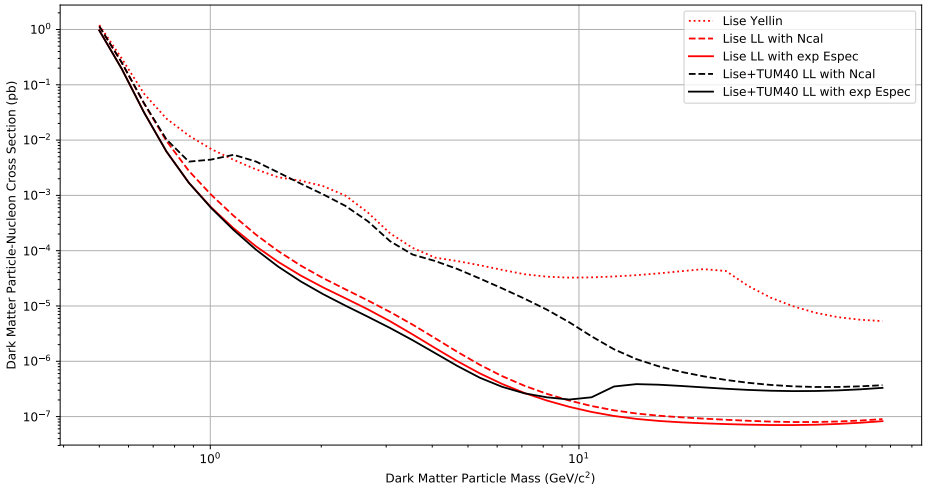
Figure 17: Dark matter particle exclusion limits for the Lise and TUM40 datasets separately and combined each including the neutron calibration (Ncal)

Below a dark matter mass of roughly 0.75 $\mathrm{GeV/c^2}$ the combined limit is almost identical to the Lise limit. The reason for this is that due to its higher threshold of 0.603 keV the TUM40 detector has nearly no sensitivity in that mass region. However, for DM masses above 0.75 $\mathrm{GeV/c^2}$ the combined limit is actually worse than the Lise

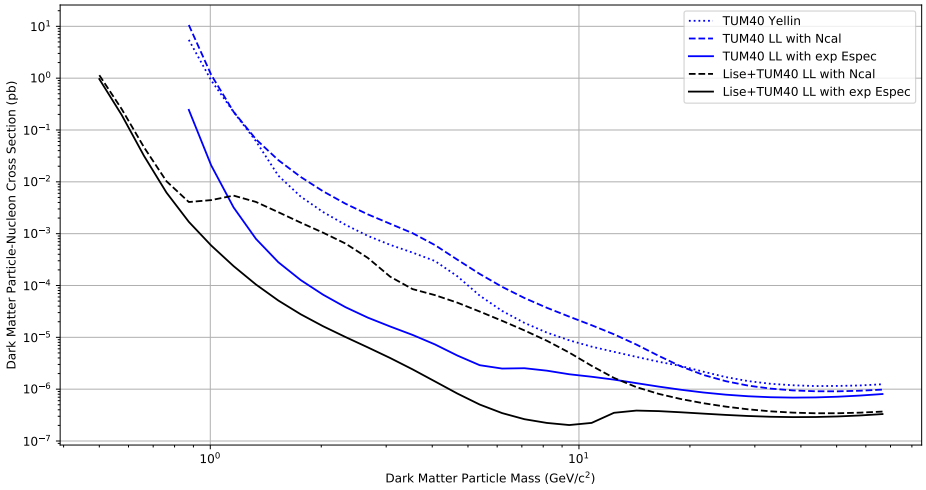
only limit. This is caused by the exponential increase in events near the threshold for the TUM40 dataset which mimics a dark matter signal and therefore increases the limit. However, since the Lise detector does not feature such an increase in event frequency, at least not at similar energies, it can be excluded as a dark matter signal. For this reason an exponential energy spectrum is introduced to the electron band to describe this behavior.

7.5 Testing the Exponential Electron Energy Spectrum

To test the impact of the exponential term in the electron energy spectrum the same procedure as in the last section is performed. As long as there are detectors with a lower threshold which do not display a similar exponential increase as the detectors with a higher threshold it should be possible to enable the exponential term in the electron energy spectrum. Therefore the exponential electron energy spectrum is enabled for both detectors, Lise and TUM40. It has to be noted here that the sharp increase in event frequency near the threshold does not happen in the electron recoil band. Instead it is centered around a light-yield of 0. However due to the very large overlap of the bands in that region and to avoid an exponential term in the nuclear recoil bands, which could be too similar to the possible dark matter spectrum, this spectrum is included in the spectrum of the electron band. Figure 18 shows the results of the exclusion calculations with that additional term.



(a) Lise and combined limits



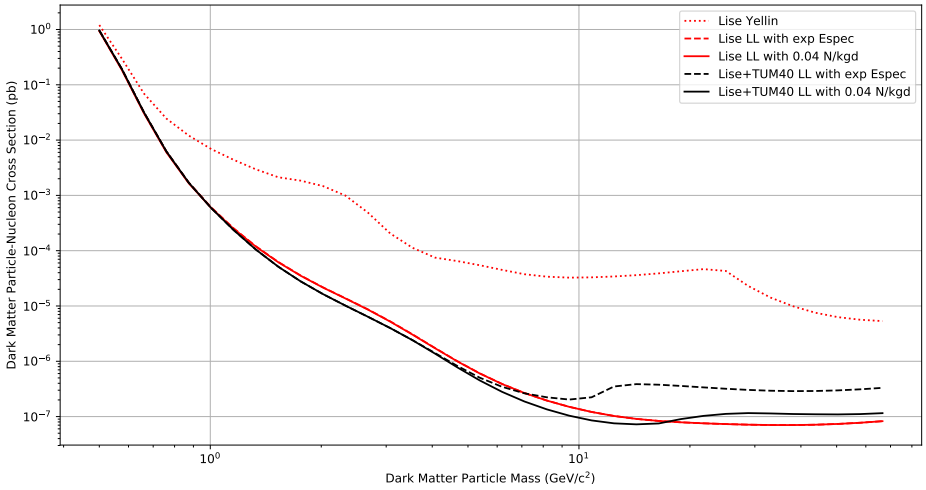
(b) TUM40 and combined limits

Figure 18: Dark matter particle exclusion limits for the Lise (a), TUM40 (b) and combined (a and b) datasets including the neutron calibration, with and without the exponential energy spectrum.

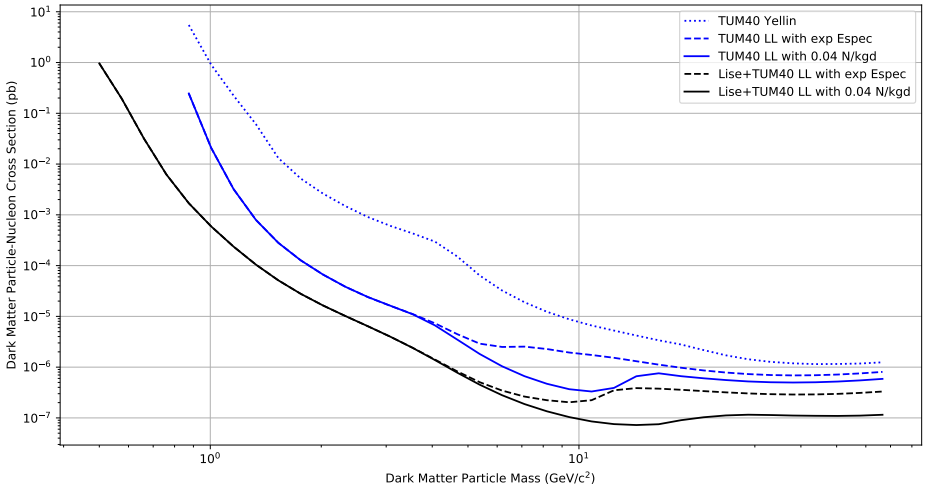
With this added term in the spectrum the limit calculated with the profile likelihood approach outperforms the optimal interval method for both detectors. The combined limit shows a slight improvement over the Lise limit for DM masses below 7 $\mathrm{GeV/c^2}$. For higher masses there is a sudden rise in the combined limit. This most likely arises from a neutron background event in the TUM40 dataset which is misidentified as a dark matter interaction. For the TUM40 dataset used in this calculation the existence of an event, which is not compatible with any other type of background or signal other than the neutron background, is known [8].

7.6 Impact of a Neutron Background Estimation

For earlier runs an estimate for the number of expected neutron background events exists [19]. However, this was before the inner polyethylene shielding was installed. Due to this the neutron background significantly changed and no measurement of its size exists. There are, however, simulations which estimate a reduction of the neutron background by a factor of 10, due to the polyethylene shielding [8]. Since this is just a very weak hint for the real neutron background the limits calculated with the inclusion of this background spectrum hold a very limited validity. Therefore the main aim of this test is just to demonstrate the potential impact of a background reduction on the individual as well as the combined limits. For this purpose two different expected neutron background rates are tested. First, a rate of 0.04 neutron background events per kgd of exposure is used. In this case there is roughly one expected neutron event in the TUM40 dataset, which has an exposure of 28.548 kgd. The limits calculated with this background rate are shown in figure 19. A second test, shown in figure 20, is performed with a rate of 0.07 events per kgd, which leads to two expected neutron events in the TUM40 dataset. These two rates are also used for the Lise and the combined limits. Since the poisson distribution, which restricts the number of neutron events in the fit, uses the total exposure of all included detectors the combined fit can attribute the neutron events to whichever dataset they match best.

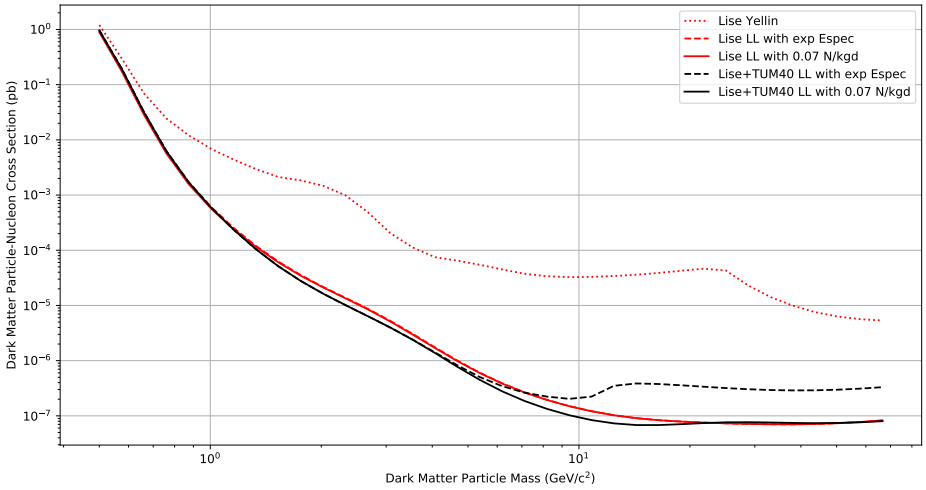


(a) Lise and combined limits

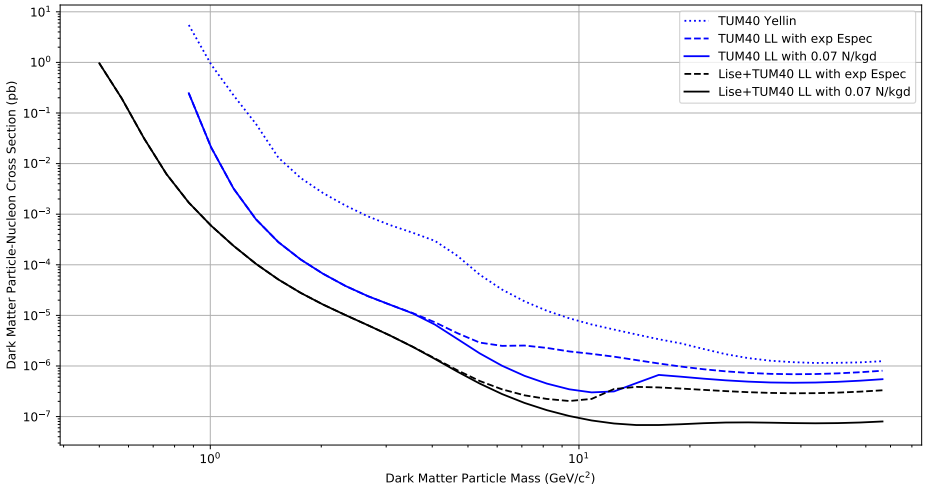


(b) TUM40 and combined limits

Figure 19: Dark matter particle exclusion limits for the Lise (a), TUM40 (b) and combined (a and b) datasets including the neutron calibration, exponential energy spectrum, with and without a neutron background of 0.04 events per kgd.



(a) Lise and combined limits



(b) TUM40 and combined limits

Figure 20: Dark matter particle exclusion limits for the Lise (a), TUM40 (b) and combined (a and b) datasets including the neutron calibration, exponential energy spectrum, with and without a neutron background of 0.07 events per kgd.

The additional background spectrum of 0.04 neutron events per kgd does not affect the exclusion limit calculated with the Lise dataset. However, for the TUM40 limit the inclusion of a neutron background significantly lowers the exclusion limit for dark matter particle masses between 5 and 13 GeV/c^2. For higher masses up to 60 GeV/c^2 the limit calculated with a neutron background spectrum is still slightly better than without one. The biggest change occurs in the combined limit which is clearly better with the neutron background for dark matter masses above 5 GeV/c^2. Additionally, the combined limit is lower than the limit calculated with only the Lise data up until roughly 15 GeV/c^2. However, for higher masses the individual limit is still slightly better. It is interesting to note that when looking at the parameters of the fit for the combined datasets it is apparent that the fit attributes all of the neutron background events to the TUM40 data.

For a rate of 0.07 neutron background events per kgd the exclusion limits of the individual detectors barely change compared to a rate of 0.04 events per kgd. A larger change occurs in the combined limit which is now entirely lower or at least on par with the Lise only limit. Unfortunately, it is only lower up until roughly 20 GeV/c^2. The combined fit once again attributes all neutron background events to the TUM40 dataset. Therefore, the fit most likely overestimates the actual neutron background, but the combined limit improves only marginally. This indicates that the model used to describe the neutron background spectrum is not appropriate. However, due to the very low rate of events fitting a separate spectrum is not practicable. Therefore a simulation of the expected neutron background could significantly improve the model and therefore the limit calculation.

7.7 Impact of the Simulated Efficiency on the Exclusion Limits

As already mentioned in section 5.3.1 the convolution of the dark matter spectrum with a Gaussian distribution to describe the measured dark matter spectrum is not an entirely correct approach. For this reason the convolution shall be replaced with a simulation of the detector behavior. Such a simulation exists for detector A, a new detector module in CRESST-III Phase 1 (run 34). The impact of the simulation compared to the convolution and measured cut efficiency is tested using the measurement data of this module. The measurement data for detector A along with the band fit is shown in figure 21 and the energy spectrum for this detector module is displayed in figures 22 and 24.

Unfortunately, this detector displays a large increase in event frequency near the threshold, even more extreme than for the TUM40 detector, while also being the module with the lowest threshold of 30 eV. Thus it is not justified to use the exponential spectrum in this case. As a result of this the profile likelihood limit performs worse than the limit calculated with the optimal interval method for higher dark matter particle masses. This can be seen in figure 23.

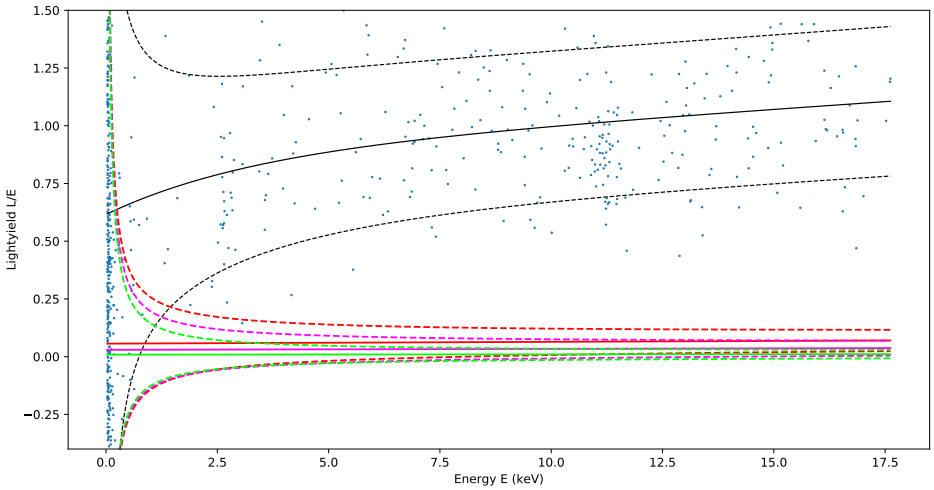


Figure 21: 2D scatter plot of 5.689 kgd of detector A measurement data overlaid with the band fit functions introduced in chapter 5.1.

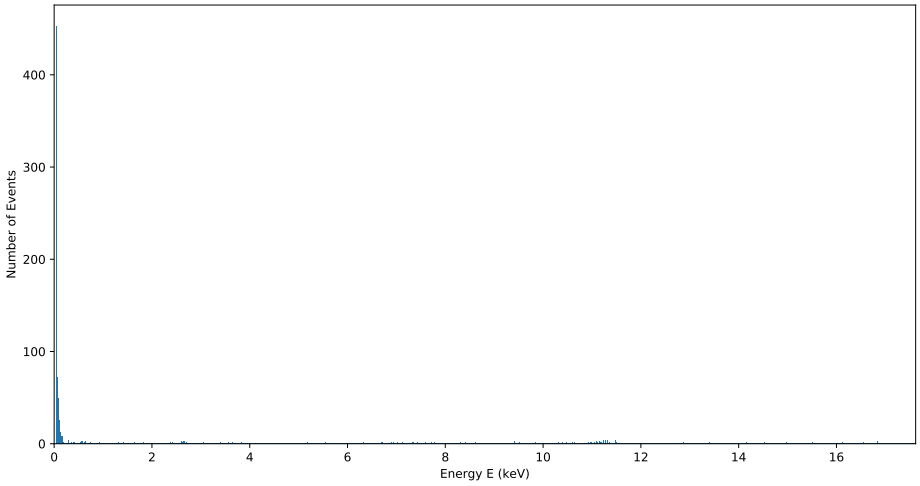


Figure 22: Nuclear recoil energy spectrum using 1000 bins for 5.689 kgd of detector A measurement data.

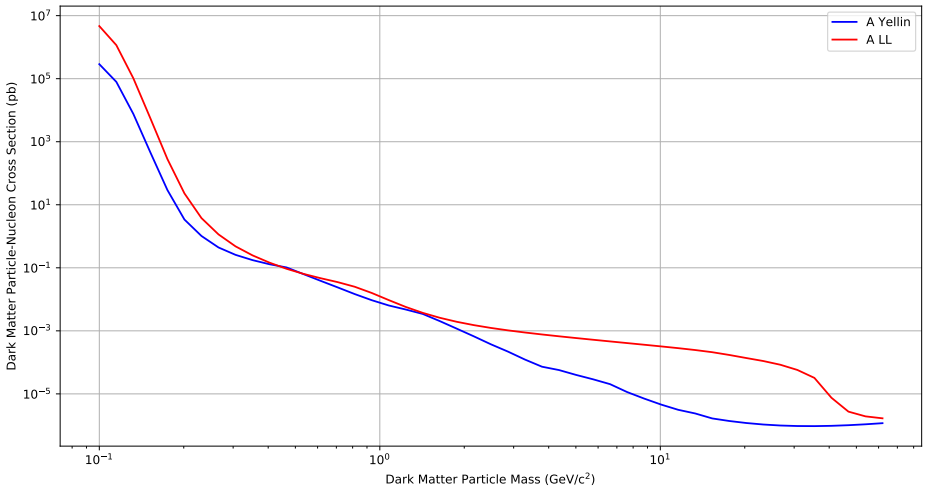


Figure 23: Dark matter particle exclusion limits calculated with Yellin's optimal interval and profile likelihood methods for the detector A dataset with neutron calibration, using the measured cut effieciency.

For higher dark matter masses above 1.5 GeV/c^2, where the dark matter spectrum reaches higher recoil energies, Yellin's optimal interval method simply ignores the low recoil energy region with a large background and utilizes the higher energy region with few events. The profile likelihood does not have the ability to ignore a region and while the shape of the measured spectrum does not match the shape of the dark matter spectrum it is still beneficial for the fit to attribute some of the events to the dark matter signal and thus increasing the exclusion limit. For very low dark matter particle masses from 0.1 to 0.4 GeV/c^2 the optimal interval method also outperforms the profile likelihood method significantly. This indicates that the shape of the dark matter spectrum for such masses differs from the spectrum in the measurement data. However, for masses between 0.4 and 1.5 GeV/c^2 the two exclusion limits are fairly similar. This indicates that the shape of the measured spectrum is similar to the recoil spectrum of a dark matter particle in that mass range. A maximum likelihood fit including the dark matter particle parameters as free fit parameters is done to test this hypothesis. The optimization converges to a minimum for a dark matter spectrum of a particle with a mass of 1.34 GeV/c^2 and a cross section of $4.25 \cdot 10^{-3}$ pb.

Since Yellin's optimal interval method is preforming better in this case it is used for the comparison between the convolution and simulation method, which is shown in figure 25. While the limit calculated with the convolution suggests a sensitivity down to dark matter masses of 0.1 GeV/c^2 the limit for the simulation only reaches down to roughly 0.17 GeV/c^2. This is somewhat expected since for the convolution the sensitivity theoretically reaches towards DM masses of almost 0 GeV/c^2. The convolution pushes some events above the threshold, no matter how low the actual recoil energy is, as

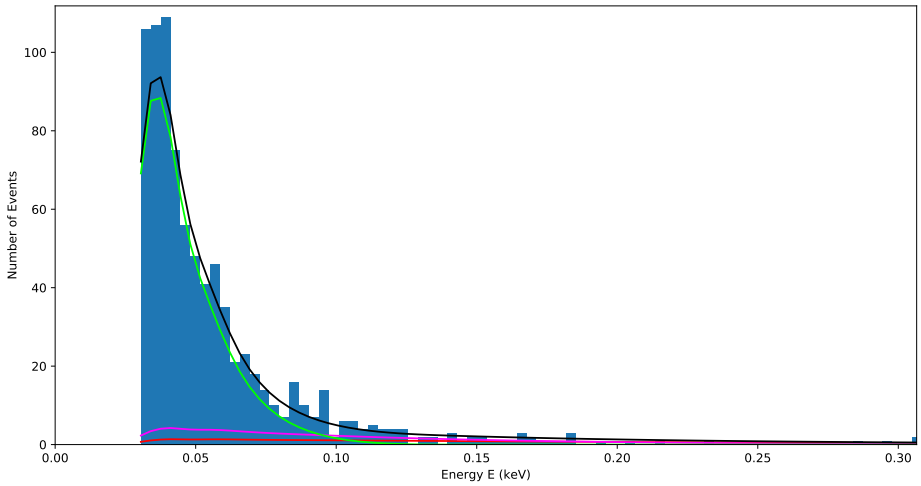


Figure 24: Lower part of the nuclear recoil energy spectrum for 5.689 kgd of detector A measurement data. This time in 5000 bins and showing the spectrum of a dark matter particle with a mass of 1.34 GeV/c^2 and a cross section of $4.25 \cdot 10^{-3}$ pb. The black line is the total spectrum of all three bands and the red, purple and green lines represent the oxygen, calcium and tungsten bands respectively.

long as the cross section is large and therefore the signal is strong enough. However, this behavior is wrong and cannot happen with the simulation. Therefore, the limit calculated with the simulation method starts at higher dark matter masses since the detector has no sensitivity towards lower masses. For the complete mass range the exclusion limit calculated with the simulation stays above the limit for the convolution method. This is again caused by the convolution which pushes more low recoil energy events above the threshold. The difference between the methods is therefore most prominent in the low recoil energy region, which also explains why the difference gets smaller for larger dark matter particle masses.

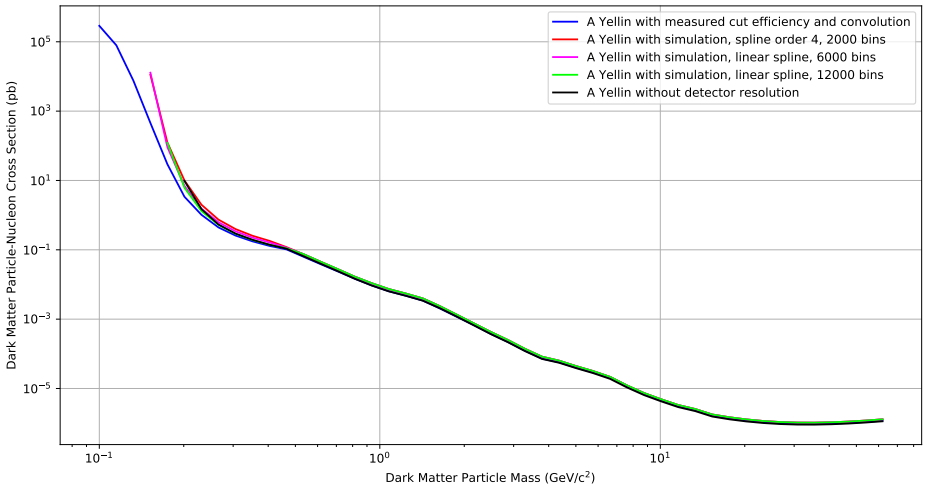


Figure 25: Dark matter particle exclusion limits calculated with Yellin's optimal interval method for the detector A dataset with neutron calibration using the measured cut efficiency (blue) and the simulation with different bin sizes (red, purple and green). Additionally a limit calculation without the convolution with the detector resolution is performed as comparison, which is shown in black.

8 Conclusion and Outlook

The improvements to the parametrization of the model for the measurement data as well as new implementations in the maximum likelihood fit introduced in this thesis have the capability to enhance the exclusion limit calculation for the CRESST experiment. Especially the inclusion of the neutron calibration data into the maximum likelihood fit significantly increases its accuracy and convergence speed. But also smaller changes like the γ-band quenching method and the more accurate implementation of the slopes improve the model and bring it closer to the true detector behavior.

The change from the convolution to the simulation to calculate the measured dark matter recoil spectrum also enhances the correctness of the exclusion limits. This method stays truer to the real detector behavior which leads to slightly less sensitivity towards low mass dark matter particles.

With the new model a profile likelihood exclusion limit calculation has the capability to vastly outperform a calculation with Yellin's optimal interval method. However, this is very dependent on the knowledge of the background spectra in the detector data. Especially the increase in event frequency towards the detector threshold in many datasets has a larger negative impact on the profile likelihood limits than on the optimal interval limits. Therefore, a better knowledge of this behavior and possibly a better overall description of these events, which in reality do not happen in the electron recoil band,

but are centered around a light-yield of 0, could further enhance the limit calculation.

Unfortunately, the combination of detectors in the maximum likelihood formalism does not yield the desired beneficial impact on the exclusion limits. Instead the limits are mostly weaker with this combination. However, tests have shown that this is most likely caused by one or more neutron background events. The current model for the neutron background energy spectrum is not suited to describe the behavior correctly. Therefore, a simulation of the expected neutron background could possibly improve the combined limits significantly.

The improved maximum likelihood implementation presented in this thesis has the capability to calculate stronger exclusion limits for the dark matter particle - nucleon interaction cross section for the data obtained in the CRESST experiment. There is, however, still room for improvement. Especially a better knowledge of the expected backgrounds and the detector behavior can further better the limit calculation with this method.

References

[1] Steven Yellin. "Finding an upper limit in the presence of an unknown background". In: *Physical Review D* 66.3 (2002), p. 032005.

[2] Jochen Schieck and Holger Kluck. *Dark Matter*. Vorlesungsskriptum TU Wien. 2017.

[3] *Extended rotation curve of M33*. URL: `https://commons.wikimedia.org/wiki/File:M33_rotation_curve_HI.gif`.

[4] NASA/CXC/M. Weiss. *File:1e0657 scale.jpg*. URL: `https://commons.wikimedia.org/wiki/File:1e0657_scale.jpg`.

[5] Peter AR Ade et al. "Planck 2015 results-xiii. cosmological parameters". In: *Astronomy & Astrophysics* 594 (2016), A13.

[6] Graciela B Gelmini, Efunwande Osoba, and Sergio Palomares-Ruiz. "Inert-sterile neutrino: cold or warm dark matter candidate". In: *Physical Review D* 81.6 (2010), p. 063529.

[7] Jens Michael Schmaler. "The CRESST Dark Matter Search-New Analysis Methods and Recent Results". PhD thesis. Technische Universität München, 2010.

[8] Florian Reindl. "Exploring Light Dark Matter With CRESST-II Low-Threshold Detectors". PhD thesis. Technische Universität München, 2016.

[9] G Angloher et al. "Results on low mass WIMPs using an upgraded CRESST-II detector". In: *The European Physical Journal C* 74.12 (2014), p. 3184.

[10] G Angloher et al. "Results on light dark matter particles with a low-threshold CRESST-II detector". In: *The European Physical Journal C* 76.1 (2016), p. 25.

[11] M Mancuso et al. "A Low Nuclear Recoil Energy Threshold for Dark Matter Search with CRESST-III Detectors". In: *Journal of Low Temperature Physics* (2018), pp. 1–8.

[12] F Donato, N Fornengo, and S Scopel. "Effects of galactic dark halo rotation on WIMP direct detection". In: *arXiv preprint hep-ph/9803295* (1998).

[13] Frank J Kerr and Donald Lynden-Bell. "Review of galactic constants". In: *Monthly Notices of the Royal Astronomical Society* 221.4 (1986), pp. 1023–1038.

[14] Richard H Helm. "Inelastic and elastic scattering of 187-MeV electrons from selected even-even nuclei". In: *Physical Review* 104.5 (1956), p. 1466.

[15] Gintaras Dūda, Ann Kemper, and Paolo Gondolo. "Model-independent form factors for spin-independent neutralino–nucleon scattering from elastic electron scattering data". In: *Journal of Cosmology and Astroparticle Physics* 2007.04 (2007), p. 012.

[16] Steven Yellin. Feb. 2011. URL: `http://titus.stanford.edu/Upperlimit/`.

[17] Stephanie Hansmann-Menzemer. *Modern Methods of Data Analysis - Lecture VII (26.11.07)*. URL: https://www.physi.uni-heidelberg.de/~menzemer/Stat0708/statistik_vorlesung_7.pdf.

[18] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011), p. 1554. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: https://doi.org/10.1140/epjc/s10052-011-1554-0.

[19] Godehard Angloher et al. "Results from 730 kg days of the CRESST-II Dark Matter Search". In: *The European Physical Journal C* 72.4 (2012), p. 1971.

[20] Patrick Huff. "The Detector Parameters Determining the Sensitivity of the CRESST-II Experiment". PhD thesis. Technische Universität München, 2010.

[21] RF Lang et al. "Scintillator non-proportionality and gamma quenching in CaWO4". In: *arXiv preprint arXiv:0910.4414* (2009).

[22] R Strauss et al. "Energy-dependent light quenching in CaWO4 crystals at mK temperatures". In: *The European Physical Journal C* 74.7 (2014), p. 2957.

[23] Patrick Kofod Mogensen and Asbjørn Nilsen Riseth. "Optim: A mathematical optimization package for Julia". In: *Journal of Open Source Software* 3.24 (2018), p. 615. DOI: 10.21105/joss.00615.

[24] *BlackBoxOptim.jl*. URL: https://github.com/robertfeldt/BlackBoxOptim.jl.

[25] John A Nelder and Roger Mead. "A simplex method for function minimization". In: *The computer journal* 7.4 (1965), pp. 308–313.

[26] Fuchang Gao and Lixing Han. "Implementing the Nelder-Mead simplex algorithm with adaptive parameters". In: *Computational Optimization and Applications* 51.1 (2012), pp. 259–277.

[27] Zhi-Hui Zhan et al. "Adaptive particle swarm optimization". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39.6 (2009), pp. 1362–1381.

[28] *Differential evolution*. URL: https://ucilnica.fri.uni-lj.si/pluginfile.php/7950/mod_resource/content/5/Engelbrecht07-DE.pdf.

[29] Jonathan Richard Shewchuk. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*. URL: http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf.

[30] William W Hager and Hongchao Zhang. "Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent". In: *ACM Transactions on Mathematical Software (TOMS)* 32.1 (2006), pp. 113–137.

[31] William W Hager and Hongchao Zhang. "The limited memory conjugate gradient method". In: *SIAM Journal on Optimization* 23.4 (2013), pp. 2150–2168.

[32] Richard P. Brent. "An algorithm with guaranteed convergence for finding a zero of a function". In: *The Computer Journal* 14.4 (1971), pp. 422–425.