Master Thesis

XENON100 Dark Matter Search with the PAX Raw Data Processor for XENON1T

vorgelegt von
Christian Wittweg

Institut für Kernphysik

- Dezember 2016 -

Erster Gutachter: Prof. Dr. C. Weinheimer
Zweiter Gutachter: Prof. Dr. A. Kappes

EIGENSTÄNDIGKEITSERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel **XENON100 Dark Matter Search with the PAX Raw Data Processor for XENON1T** selbstständig verfasst habe, und dass ich keine anderen Quellen und Hilfsmittel als die angegebenen benutzt habe und dass die Stellen der Arbeit, die anderen Werken – auch elektronischen Medien – dem Wortlaut oder Sinn nach entnommen wurden, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht worden sind.

Münster, den 19.12.2016

Christian Wittweg

CONTENTS

INTRODUCTION

A plethora of observations from cosmology and astronomy constitute an overwhelming body of evidence that a new form of non-luminous and non-baryonic matter makes up about a quarter of the mass-energy content of the universe. The constituents of this *dark matter* are commonly assumed to be new elementary particles arising from theories beyond the Standard Model of particle physics. One well-motivated class of candidates are weakly interacting massive particles (WIMPs) which would have been created as thermal relics in the Big Bang and subsequently played a crucial role in the evolution of the cosmos.
As the whole concept of dark matter stems from indirect observations, there has been an ongoing and diverse experimental effort to directly detect the interactions of WIMPs with ordinary matter. The identification and characterization of the quanta of dark matter would solve one of modern physics' greatest puzzles and provide new insights into the evoultion of the universe.
The *XENON Dark Matter Project* is aimed at the direct detection of dark matter utilizing xenon dual-phase time projection chambers in the legacy experiments XENON10 and XENON100 as well as the operational next generation detector XENON1T. Interactions of WIMPs with the xenon target are expected to produce scintillation light and ionization charges. Both signal channels are read out with photomultiplier tubes (PMTs). The raw data has to be processed, i. e. converted from PMT channel-wise ADC counts into physically meaningful information, for a dark matter search. The *Processor for Analyzing Xenon* (PAX) has been developed for this task in XENON1T.
This work will validate the novel data processor by carrying out a full dark matter search on PAX-processed XENON100 data. The dataset at hand contains of 225 live days worth of data recorded in 2011 and 2012. It has been subjected to numerous analyses utilizing the XENON100 raw data processor. This makes the final results on dark matter exclusion limits and remaining events in the signal region directly comparable. In addition analysis code and data cuts applicable for XENON1T are developed.
The first chapter provides the theoretical background of dark matter evidence, candidate particles and their direct detection. The *XENON Dark Matter Project* is introduced briefly. Chapter two introduces the tools and methods that are used in a dark matter search and gives an outline of the analysis steps that have to be performed in order to get from light and charge signals to a so called exclusion plot that puts constraints on the mass and interaction cross sections of WIMPs. The third chapter is devoted to the preparation of the dataset making use of calibration data from various sources. The focus here lies on two crucial analysis steps: the translation and development of cuts that are

used to separate physical and non-physical backgrounds from hypothetical dark matter events, and an energy calibration of the detector. After the preparation of the dataset the actual dark matter analysis is performed in chapter four. A dark matter search region is defined by introducing a discrimination between signal candidates and background events. Considering the signal acceptance of the cuts a mass-dependent limit on the WIMPs spin-independent interaction cross section is calculated applying a Feldman & Cousins approach. The result is then compared with the original XENON100 result. The last chapter presents a conclusion and outlook.

CHAPTER

ONE

THEORY OF DIRECT DARK MATTER DETECTION

This chapter provides a short overview of the wealth of evidence for dark matter, both from astronomy and cosmology. The most common dark matter candidates are introduced with the focus on weakly interacting massive particles (WIMPs). General considerations for their direct detection are presented including their experimental signature, interaction cross sections, background sources, common detector designs, measured signals and a generic direct detection result. In the last section of this chapter the *XENON Dark Matter Project* is presented.

1.1 Observational Evidence for Dark Matter

In the 1920s and 1930s astronomers pioneered in adopting the concept of *dark matter* as a means of explaining the unexpected motions of astronomical objects [1, 2]. Fritz Zwicky noted that the high velocities of nebulae in the Coma cluster could not be explained by taking into account the gravitational interactions of visible matter only [3]. The observation by Vera C. Rubin that the orbital velocities of stars in galaxies are approximately constant with increasing distance r to the galactic center falls into the same category of observations [4]. According to the distribution of visible mass these stars should exhibit a Newtonian $v \propto 1/\sqrt{r}$. The introduction of a uniformly distributed dark matter halo can explain this deviation as shown in Fig. 1.1 [5, 6].

These observations on relatively small galactic scales are augmented by measurements of gravitational lensing [7]. Light rays are deflected due to the deformation of spacetime by a massive object in the line of sight between the observer and the object under study. This results in multiple images (strong lensing) or a deformation (weak lensing) of the observed object's original image. The degree of deformation allows for the reconstruction of the gravitational potential of the deflecting object. In general reconstructed masses are up to a few hunded times greater than the amount of luminous matter. A promiment example for the use of this method is the "Bullet" cluster [8]. The mass distribution of two colliding galaxy clusters is inferred from an image in visible light taken by the Magellan probe. This image is shown on the left hand side of Fig. 1.2. The baryonic matter in both clusters is concentrated in gas clouds that are heated by the collision and emit X-rays that can be made visible by the Chandra satellite. This is shown on the right hand side of the figure. The offset of the mass distribution and the gas clouds is clearly visible. This implies

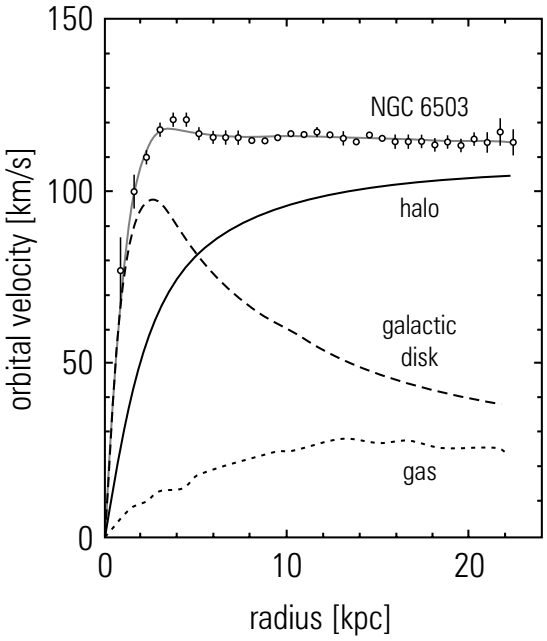

Figure 1.1: Orbital velocities of stars in NGC 6503 as a function of their distance from the galactic center. The expected orbits from the mass distribution of the galactic disk are displayed together with contributions from a dark matter halo and gas. [6].

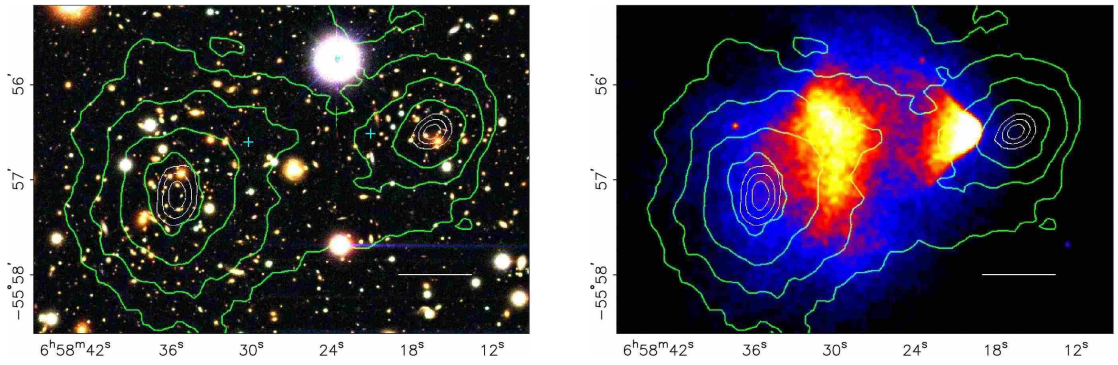

Figure 1.2: Left: color image from the Magellan images of the merging cluster 1E0657-558 (Bullet cluster) with the white bar indicating 200 kpc at the distance of the cluster. The green contours in both panels show the weak lensing reconstruction of the mass distribution of the clusters. The white contours show the errors on the positions of the lensing peaks and correspond to 68.3 %, 95.5 %, and 99.7 % confidence levels. The cyan crosses mark the the center of gravity of the heated gas clouds that can be seen in the right image. Right: Chandra image of the cluster. The offset of the mass distribution and the visible matter distribution (hot gas emitting X-rays in Chandra image) is clearly visible [8].

that the mass of the colliding clusters is dominated by non-luminous matter that barely interacts with baryonic matter and itself allowing it to penetrate the colliding clusters unperturbed.

Although numerous observations of cluster collisions [7] widely discourage alternative approaches to dark matter such as modified Newtonian dynamics (MOND) and baryonic massive halo compact objects (MACHOs), the most compelling case for particle dark matter comes from cosmology.

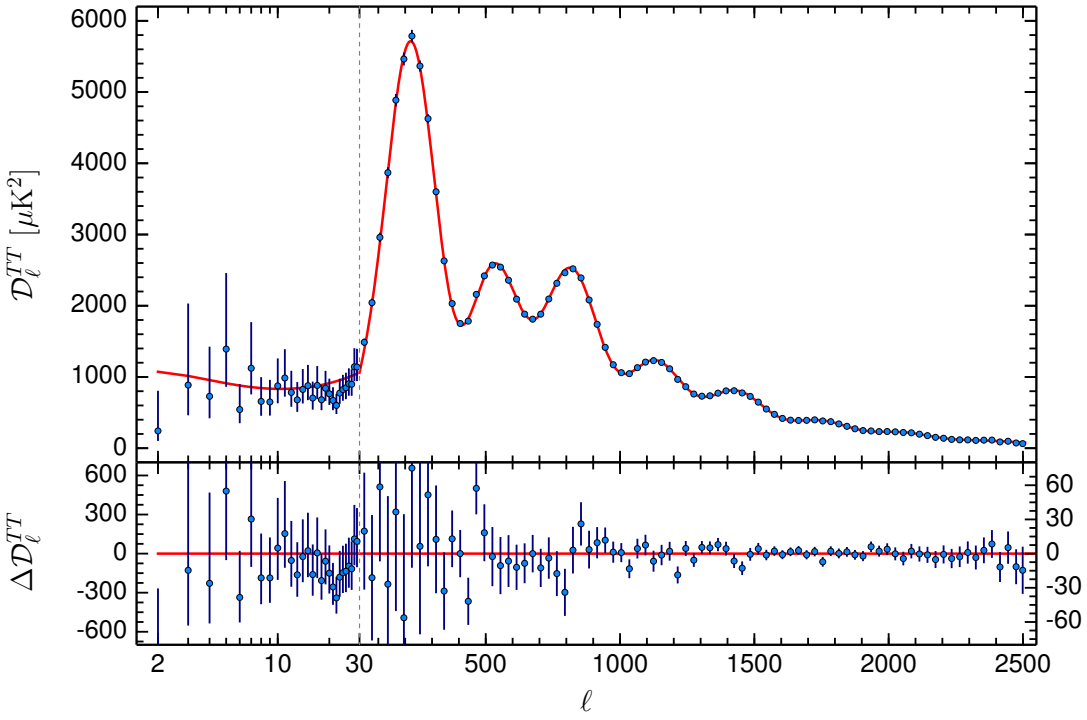

Figure 1.3: Planck 2015 temperature power spectrum by multipole moment ℓ with fitted ΛCDM model and fit residuals. The error bars represent 1σ-uncertainties [10].

The satellite missions WMAP [9] and PLANCK [10] have measured temperature anisotropies in the cosmic microwave background (CMB) with unprecedented precision and have opened a window into the universe at an age of about 380.000 years. The power spectrum of the CMB, shown in Fig. 1.3, exhibits a characteristic pattern arising from the oscillations of baryons and photons in the gravitational potential that is dominated by dark matter density perturbations [7]. A fit of this spectrum with the ΛCDM (Λ cold dark matter) cosmological model gives, among others, the baryonic matter, dark matter and dark energy contents of the universe. The current values for the densities of these components are $\Omega_{\mathrm{DM}} = 0.265$, $\Omega_{\mathrm{baryon}} = 0.049$[1] and $\Omega_\Lambda = 0.686$ [10].

It is commonly assumed that the temperature anisotropies in the CMB reflect inhomogeneities in the matter and dark matter distributions in the early universe. These have their origin in quantum fluctuations during the period of inflation. The further evolution of the universe in the form of structure formation from about 10 kpc on to the largest scales is usually studied with N-body simulations of dark matter particles. A baryonic component can be included in these simulations that can then be compared to galaxy surveys and other observations [11, 12]. This leads to the postulate of a dominantly cold, i.e non-relativistic, dark matter component in order to explain the clustering scale of galaxies [7].

Thus, from the full set of observations it can be concluded that dark matter is composed of massive, neutral and non-relativistic particles which exhibit weak self-interaction. Several candidates with these characteristics are presented in the next section.

[1] A baryonic nature of dark matter is further ruled out by Big Bang nucleosynthesis (BBN), as the predicted abundances of light elements depend on the baryon density. Measurements constrain the value $\Omega_{\mathrm{baryon}} \approx 0.04$ [7].

1.2 Dark Matter Candidates

Postulated dark matter particles usually belong to one of three theoretical classes [7, 13]:

1. **Neutrinos** are the only Standard Model (SM) particles with fitting properties, but too light to constitute cold dark matter. A fourth **sterile neutrino** can be introduced in order to explain the smallness of the neutrino masses and observed anomalies in reactor neutrino fluxes. This ν_{sterile} could take part in neutrino oscillations, would have an allowed mass range from a few to tens of keV$/c^2$ and could constitute cold or warm dark matter depending on the production mechanism. Recent measurements by the IceCube neutrino telescope disfavour sterile neutrinos [14].

2. **Axions** and axion-like particles arise from the spontaneous breaking of a postulated symmetry that resolves the smallness of CP violation in Quantum Chromodynamics (QCD) [7] . They are cold dark matter candidates, strongly constrained by astrophysical and cosmological observations, and could account for the total Ω_{DM}.

3. Hypothetical stable, neutral and massive (usually $m \sim \text{GeV}/c^2 - \text{TeV}/c^2$) **weakly interacting massive particles (WIMPs)** are suggested by numerous models beyond the SM [7]. As thermal relic particles WIMPs were in thermal equilibrium with the plasma in the early universe until the temperature of the plasma dropped below the WIMP mass causing the WIMPs' freeze-out. The dark matter relic density was subsequenty reached when the Hubble expansion rate of the universe exceeded the WIMP annihilation rate governed by its self-interaction cross section. The necessary cross section for the correct relic abundance of $0.3\,\text{GeV}/c^2$ is of the order of the electroweak scale.

However, this is only a tiny selection of the dark sector. A comprehensive review can be found in [15]. The next section will focus on aspects of direct WIMP detection.

1.3 Direct Detection Experiments

A triad of processes can be exploited for testing the particle dark matter hypothesis. These are the production at colliders, the indirect search for annihilation signals and scattering processes with ordinary matter in direct detection. The corresponding couplings to ordinary matter are depicted in Fig. 1.4.
Direct detection relies on measuring the small energy depositions of dark matter particles colliding with target nuclei within ultra-low-background detectors. This is in the following referred to as a nuclear recoil (NR). Elastic scattering of WIMPs with target nuclei produce NRs in the range of $(1 - 100)$ keV [16]. The identification of these events and their discrimination from abundant backgrounds requires considerations with regard to the experimental signature as well as particle and nuclear physics aspects. Furthermore, assumptions regarding the local dark matter density and velocity distribution in the Milky Way are necessary for the interpretation of experimental results.

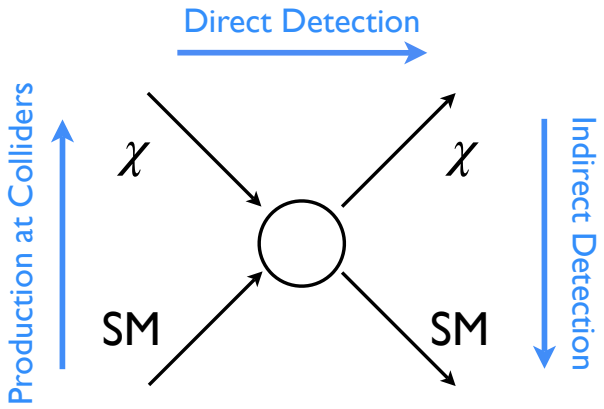

Figure 1.4: Schematic couplings of dark matter (χ) to standard model (SM) particles with their detection channels. Sketch based on [7].

1.3.1 Experimental Signature in Direct Detection

Given the minuscule interaction strength of dark matter particles the expected signature consists of a recoil spectrum of single scatter events at an extremely low rate. The differential WIMP recoil spectrum $\frac{\mathrm{d}R}{\mathrm{d}E}$ can be expressed as

$$\frac{\mathrm{d}R}{\mathrm{d}E}(E,t) = \frac{\rho_0}{m_\chi \cdot m_A} \cdot \int v \cdot f(\mathbf{v},t) \cdot \frac{\mathrm{d}\sigma}{\mathrm{d}E}(E,v)\,\mathrm{d}^3 v, \tag{1.1}$$

where ρ_0 is the local dark matter density, m_χ the dark matter mass, $\frac{\mathrm{d}\sigma}{\mathrm{d}E}(E,v)$ its differential cross section and m_A the mass of the target nucleus. The distribution $f(\mathbf{v},t)$ of the dark matter velocity v is time-dependent due to the motion of the earth around the sun within the dark matter halo. Consequently, the interaction rate is expected to show an annual modulation with a one-year period: A higher relative velocity translates into a greater number of dark matter particles that are able to produce nuclear recoils which are sufficiently energetic to produce detectable signals [7]. This minimum energy required for the detection of a physical signal is in the following referred to as the energy threshold of a dark matter detector.
At low energies the differential interaction rate from equation 1.1 can be described by a falling exponential [16]:

$$\frac{\mathrm{d}R}{\mathrm{d}E}(E) \approx \left(\frac{\mathrm{d}R}{\mathrm{d}E}\right)_0 \cdot F^2(E) \cdot \exp\left(-\frac{E}{E_c}\right). \tag{1.2}$$

Here $\left(\frac{\mathrm{d}R}{\mathrm{d}E}\right)_0$ is the differential rate for zero momentum transfer, E_c an energy scale constant depending on m_χ and m_A, and $F^2(E)$ the nuclear form-factor correction which is described in the next section.

1.3.2 Cross Sections

In direct detection a distinction is usually made between interactions that are independent of spin (SI) and interactions that depend on spin (SD) [7, 16]. For the former it is assumed that neutrons and protons contribute equally to the scattering. At a sufficiently low momentum transfer g the scattering amplitudes of the nucleons add in phase leading

to coherent WIMP-nucleus scattering. For the latter spin dependent interactions only unpaired nucleons contribute to the interaction making only nuclei with odd numbers of protons and neutrons viable targets. The total differential WIMP-nucleus cross section can be written as a sum of the SI and SD components:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E} = \frac{m_A}{2\mu_A^2 v^2} \cdot (\sigma_0^{\mathrm{SI}} \cdot F_{\mathrm{SI}}^2(E) + \sigma_0^{\mathrm{SD}} \cdot F_{\mathrm{SD}}^2(E)). \tag{1.3}$$

Here μ_A is the reduced WIMP-nucleus mass, and σ_0^{SI} and σ_0^{SD} are the cross sections at zero momentum transfer for the two interaction types. The form factors $F_{\mathrm{SI,\,SD}}^2$ account for low momentum transfers q where the de Broglie wavelength of the particle is no longer large compared to the dimension of the nucleus. In this case the cross section decreases with increasing q. Hence, the cross section for an interaction of a WIMP with the whole nucleus can be expressed as

$$\sigma \propto \sigma_0 \cdot F^2(q^2). \tag{1.4}$$

From now on we focus on SI interactions[2]. One can write the WIMP-nucleus cross section at zero momentum transfer as

$$\sigma_0^{\mathrm{SI}} = \sigma_p \cdot \frac{\mu_A^2}{\mu_p^2} \cdot [Z \cdot f_p + (A - Z) \cdot f_n]^2 \tag{1.5}$$

with the WIMP-nucleon reduced mass μ_p, the WIMP-nucleon cross section at zero momentum transfer σ_p and the contributions $f_{p,n}$ of protons and neutrons to the coupling. Under the assumption of $f_p = f_n$ one can express the cross section for WIMP-nucleon interactions as

$$\sigma_{\mathrm{nucleon}}^{\mathrm{SI}} \propto A^2 \cdot F^2(E). \tag{1.6}$$

This relation highlights the influence of the chosen target isotope on the result of direct detection experiments. The event rate corresponding to Eq. 1.1 is visualized in Fig. 1.5 for different target materials. It is given in number of events per keV, day and kg for spin-independent interactions. A fixed $\sigma = 10^{-45}\,\mathrm{cm}^2$ and $m_\chi = 100\,\mathrm{GeV}/c^2$ are assumed. On the left of Fig. 1.5 one can see that heavier targets such as xenon and tungsten profit from an enhancement of σ at low energies while a loss of coherence at higher energies suppresses the event rate. The drop in the rate for tungsten originates in a pole of $F^2(E)$. This translates to the fact that a lower energy threshold is crucial for experiments with heavy targets such as xenon while it is of less importance for light targets such as argon. In general the effect of the form factor correction is stronger for heavier targets, as can be seen in the right plot where the rate with and without the form factor correction is depicted. One can also see that lighter targets are sensitive to lower m_χ over a broader energy range resulting in a higher integrated total rate. Heavier targets, however, are more sensitive to high mass WIMPs at low recoil energies given a sufficiently low energy threshold.

[2] Additional information on SD interactions and their cross sections is compiled in [16].

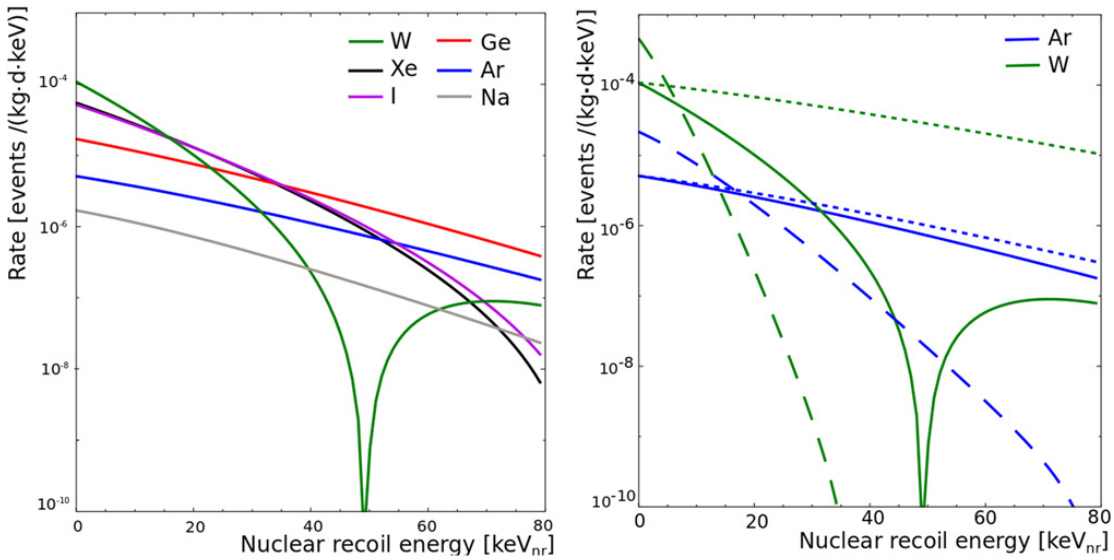

Figure 1.5: **Left:** Differential event rate for the direct detection of a 100 GeV/c^2 WIMP with a cross section of 10^{-45} cm^2 for tungsten (green), xenon (black), iodine (magenta), germanium (red), argon (blue) and sodium (grey) as target materials. **Right:** Event rate for a heavy (tungsten, green) and a light target (argon, blue). The effect of neglecting the form factor correction is depicted as a dotted line. The rate for a lighter WIMP of 25 GeV/c^2 is shown by the dashed line [7].

1.3.3 Background Sources and Reduction

Due to the low expected event rates, ultra-low background experimental conditions are required as a starting point for any dark matter direct detection experiment. The most common background contributions are [7]:

- Environmental γ-radiation originates from natural decay chains and common isotopes such as ^{137}Cs, ^{40}Ka and ^{60}Co. As a dominant process in the γ-energy regime, Compton scattering produces a broad background spectrum.
- Cosmic muons can deposit energies in the range of tens of MeV in a detector and cause the activation of target and detector materials.
- Free neutrons are generated in spallation reactions with cosmic muons, (α,n)-reactions and spontaneous fission.
- Internal backgrounds are radioactive isotopes of or within the target such as ^{85}Kr for xenon. Additionally, radon is contained in xenon and continuously emanated from surfaces.
- Neutrinos and rare physical processes such as double β-decay start to play a role at ever growing target masses in the ton-range.

Tab. 1.1 summarizes the relevant background components in xenon dark matter detectors with their common sources, interactions and mitigation strategies. Most background components such as γ and β radiation rather scatter off the electronic shell of the target than the nucleus. These electronic recoils (ER) can be discriminated from dark matter signals in data analysis.

Table 1.1: Common background sources in xenon dark matter detectors with energy ranges and mitigation strategies.

Background component	Source	Interaction type	Mitigation strategy
Environmental	^{137}Cs	661.2 keV ER	Material selection,
γ-radiation	^{40}K	1460.8 keV ER	passive shielding
	^{60}Co	1173.2 keV ER	
		1332.5 keV ER	
	U/Th decay chains	$<$ 2.6 MeV ER	
Muons	Cosmic rays	ER	Underground detectors,
		Activation of target	active veto
Cosmogenic neutrons	Spallation	NR (multiple)	Active veto
Radiogenic neutrons	(α,n)-reactions	NR (multiple)	(Water Cherenkov, LXe)
	Spontaneous fission	NR (multiple)	
Internal backgrounds	^{85}Kr	173.4 keV ER (β),	Cryogenic distillation
		514 keV ER (γ)	
	^{222}Rn (and daughters)	$<$ 6 MeV (α)	Material selection,
		$<$ 1.6 MeV ER (β, γ)	radon-free air, distillation
Neutrinos	Solar	ER, NR	
Double β-decay	^{136}Xe	2458 keV ER	

1.3.4 Background Discrimination

After an initial reduction of backgrounds in the design and construction of a direct detection experiment the discrimination of backgrounds from signals forms a central part of data analysis. Physical and non-physical (e.g. electronic noise) backgrounds are removed by selection criteria based on specific signal properties usually referred to as cuts. As these may also affect a fraction of true signal events both their background rejection as well as their signal acceptance have to be quantified.

ER interactions can be distinguished from NR interactions by combining information from multiple signal channels [7], as the response of media to an interaction depends both on the deposited energy and the depositing particle. In consequence, most dark matter experiments read out at least two of the following three signal channels: charge, light, phonons/heat. These channels can be used to sort dark matter detectors into different experiment classes depicted in Fig. 1.6.

The XENON100 experiment [17], which is introduced in detail in section 1.4, is a xenon

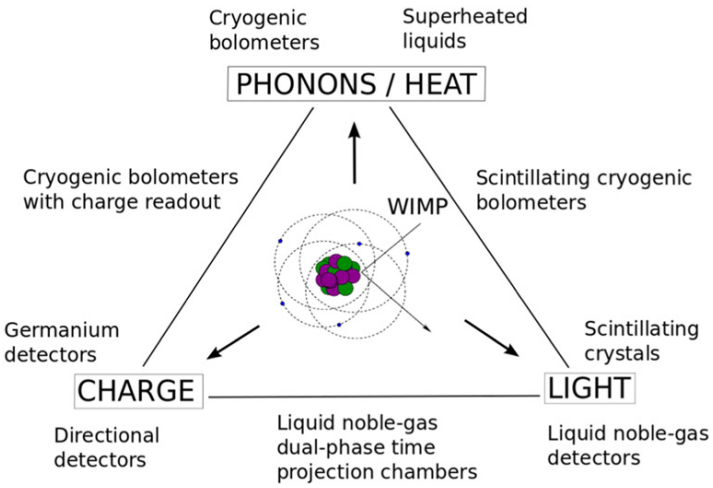

Figure 1.6: Schematic of signal channels that are exploited by dark matter direct detection experiments [7].

dual-phase time projection chamber (TPC) that measures light and charge in order to discriminate electronic from nuclear recoils. Here, the two signal channels also provide information on the interaction site in the detector which can be used to define a sensitive volume within the full target. This so-called fiducialization takes advantage of the target's self-shielding capability. It is particularly effective for targets with high Z, as these exhibit a large stopping power for charged α and β particles as well as γ-radiation.
After the discrimination of ER interactions NR interactions from neutrons, α-particles and neutrinos remain. Neutrons tend to scatter multiple times within the detector and can therefore be removed by cutting on signals which exhibit more than one physical interaction. The energy depositions from α particles within the TPC usually exceed the expected ones from elastic WIMP scattering and can easily be distinguished. Coherent nuclear scattering of solar neutrinos perfectly mimics the dark matter event signature and presents the ultimate background. Yet, due to the low interaction cross section neutrinos only start to play a role in newly emerging multi-ton detectors like XENON1T.

1.3.5 Generic Result of a Direct Detection Experiment and its Interpretation

As can be seen from Eq. 1.1 the mass m_χ and WIMP-nucleus interaction cross section σ_0 are the observables of direct detection experiments. Statistical methods are applied to claim a detection or, as in the case of this work, set an upper limit on σ_0 as a function of m_χ [7]. The most simple ansatz for deriving a mass dependent cross section is the reduction of the analysis result to a counting experiment. In this work this is done via a Feldman & Cousins approach [18].
First, a signal to background discrimination variable based on the experiment's signal channels needs to be defined. Then a region of interest with a high signal to background ratio is established in the recoil energy spectrum of this variable. The measured events in the signal region can contain signal and background events or background only. The principle is illustrated in Fig. 1.7. In the last step a background expectation from Monte Carlo and calibration data is applied to the events in the signal region. Now one can derive

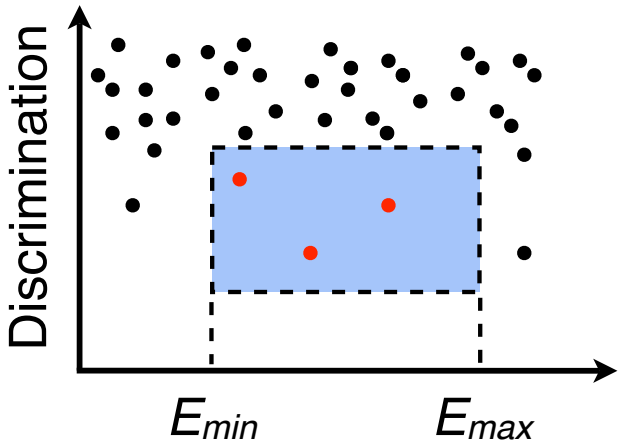

Figure 1.7: Illustration of a dark matter counting experiment based on [7].

an upper limit for the event rate with tabulated values at given confidence levels for signal counts over background [18].
For SI interactions the upper limit of the event rate $\frac{dR}{dE}$ can be directly translated into an upper limit of σ_0 by using a simplified version of Eq. 1.1:

$$\frac{\mathrm{d}R}{\mathrm{d}E}(E,t) = \frac{\rho_0}{2\mu_A^2 \cdot m_\chi} \cdot \sigma_0 \cdot A^2 \cdot F^2 \int_{v_{\min}}^{v_{\mathrm{esc}}} \frac{f(\mathbf{v},t)}{v}\, \mathrm{d}^3 v. \tag{1.7}$$

Here v_{esc} is the escape velocity while $\rho_0 = 0.3$ GeV/c^2 and $f(\mathbf{v},t)$ mark the choice of a specific dark matter halo model[3]. One commonly describes the dark matter velocity profile by an isotropic Maxwell-Boltzmann distribution:

$$f(\mathbf{v}) = \frac{1}{\sqrt{2\pi\sigma_v}} \cdot \exp\left(-\frac{|\mathbf{v}|^2}{2\sigma_v^2}\right). \tag{1.8}$$

The dispersion velocity σ_v is related to the circular velocity $v_c = 220$ km/s of the sun around the galactic center with

$$\sigma_v = \sqrt{\frac{3}{2}} \cdot v_c. \tag{1.9}$$

The lowest velocity at which dark matter particles can be detected depends on the detector threshold E_{thr} with

$$v_{\min} = \sqrt{\frac{2m_A \cdot E_{\mathrm{thr}}}{2\mu_A^2}}. \tag{1.10}$$

The left panel in Fig. 1.8 shows a generic exclusion plot with the black curve as a benchmark and the colored curves as deviation from the benchmark due to the variation of experimental variables [7]. These are a different target, lower E_{thr} and a longer exposure, i.e. a longer data taking period and/or increased target mass. The form of these curves

[3]The choice of a specific halo model affects the comparability of results from different experiments because they might probe different parts of the velocity distribution. This can be addressed by introducing a halo-independent parameter space. Refer to [7] and the cited references for a more detailed discussion.

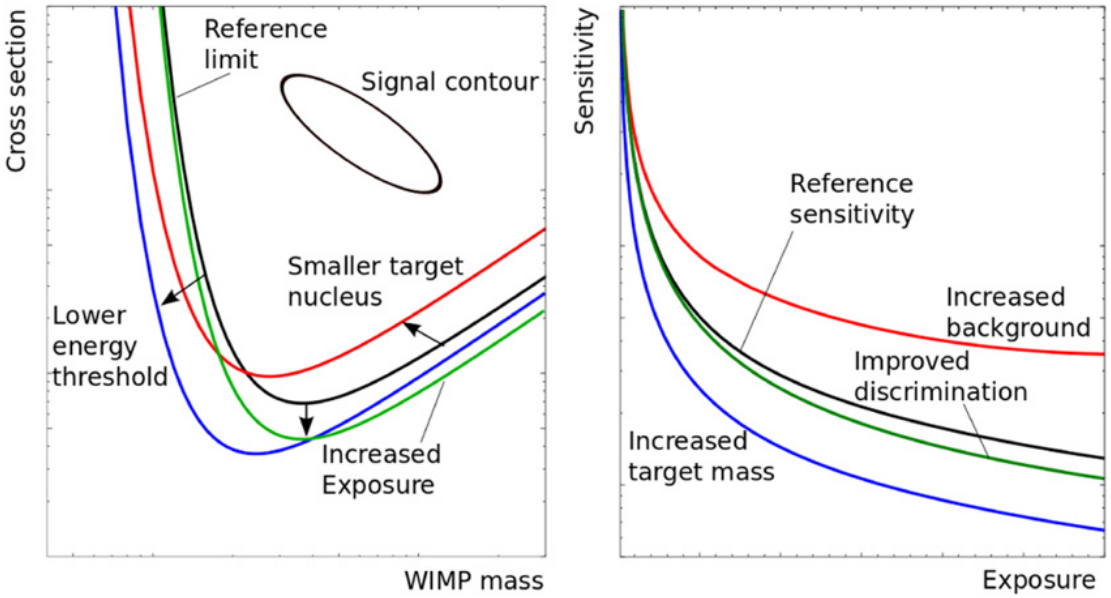

Figure 1.8: **Left:** Illustration of a generic direct detection result in the form of an exclusion plot with the upper limit for σ as a function of m_χ. The black line is the reference for an upper limit or a signal contour. **Right:** Evolution of the experimental sensitivity with increasing exposure [7].

can be understood when reconsidering Fig. 1.5. Due to the kinematics of elastic scattering the choice of a lighter target increases the sensitivity at lower WIMP masses, but decreases the overall sensitivity by shifting the curve upwards. The energy threshold dominates the low energy part of the curve. An increased exposure shifts the curve downwards as a whole without a qualitative change.

The right plot illustrates the detector sensitivity as a function of the exposure. In this case a higher sensitivity translates to a higher rate at the same cross section, WIMP mass and energy. An increased target mass provides more scattering centers and offers better self-shielding, as the volume to surface ratio of e.g. a cylindric detector increases. An improved discrimination results in fewer background events in the region of interest while an increased background has the opposite effect.

Although more advanced frequentist and Bayesian statistical methods have become a standard in the field [7, 19], counting experiments are commonly used as cross-checks and usually yield results within the same order of magnitude. Advanced methods like profile likelihood can, for example, use knowledge about the spectral shape of the signal and the probability density functions of signal and background [20, 21]. Uncertainties in the background predictions can also be addressed in order to achieve stronger limits.

1.4 The XENON100 Experiment

The XENON100 experiment [17] is located deep underground at Laboratori Nazionali del Gran Sasso (LNGS). It uses liquid and gaseous xenon in a dual-phase time projection chamber for the direct detection of WIMPs. By a combination of three data taking campaigns resulting in a combined exposure of 48 kg $\cdot$ years the XENON collaboration could set an upper limit of $\sigma_{\mathrm{SI}} < 1.1 \cdot 10^{-45}\,\mathrm{cm}^2$ for a $50\,\mathrm{GeV}/c^2$ WIMP at 90 % confidence

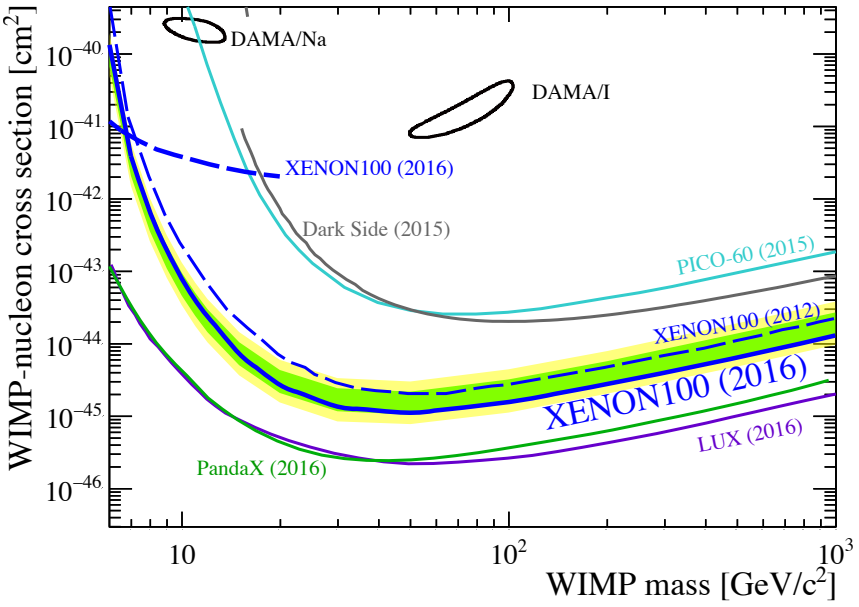

Figure 1.9: Upper limit for the SI cross section (blue line) from the combined analysis of three XENON100 science runs. The 1σ (green band) and 2σ (yellow band) expected sensitivity regions at 90 % CL are also drawn. A selection of other experimental limits at the same confidence level is included. [22].

level [22]. Here, exposure is defined as the product of the target mass and the duration of taking data. The XENON100 limit is superseded by the most recent results from the LUX and PandaX-II experiments. These employ a similar technology, but harness the advantages of larger target masses achieving 92 kg · years for LUX and 90 kg · years for PandaX-II [23] [24]. The limit plots for spin-independent interactions are delineated in Fig. 1.9. The now operational XENON1T experiment is projected to achieve a sensitivity that is two orders of magnitude above XENON100 with 2 ton · years of data.

1.4.1 Liquid Xenon as a Dark Matter Target

The advantages of liquid xenon for radiation detection are manifold and directly applicable to dark matter detection [25]. It has the highest stopping power of all noble liquids which is owed to its high atomic number (54) and density (3 g/cm^3). This means that self-shielding is highly effective when applying fiducialization. Furthermore, liquid xenon detectors are easily scalable compared to crystalline detectors and can achieve a high purity by constant purification. Internal backgrounds such as the anthropogenic ^{85}Kr can be removed by cryogenic distillation[4] [26].

Due to its isotopic composition with similar amounts of even and odd isotopes, xenon detectors are able to detect both SI and SD interactions. The high mean atomic weight $A = 131.3$ makes them most sensitive to WIMPs in the mass range of about 10 GeV/c^2 to a few hundred GeV/c^2.

Natural xenon does not contain radioactive xenon isotopes that constitute a background. Rare decay processes such as double β-decay and double electron capture with half-lives on

[4]The removal of ^{222}Rn by cryogenic distillation is also considered as an option of reducing internal backgrounds.

the order of 10^{21} years only present a significant background contribution at large target masses. Yet, they are also of experimental interest themselves [27, 28].

1.4.2 Scintillation and Ionization Process

An energy deposition on the order of a few keV produces excited and ionized xenon atoms in the liquid. Both direct excitation and the recombination of electrons and ions produce scintillation light by the formation of excited Xe_2^* dimers. The process for excited atoms Xe^* is

$$\begin{aligned} Xe^* + Xe + Xe &\rightarrow Xe_2^* + Xe, \\ Xe_2^* &\rightarrow 2Xe + h\nu. \end{aligned} \tag{1.11}$$

For Xe^+ ions there is an additional recombination and de-excitation step:

$$\begin{aligned} Xe^+ + Xe &\rightarrow Xe_2^+, \\ Xe_2^+ + e^- &\rightarrow Xe^{**} + Xe, \\ Xe^{**} &\rightarrow Xe^* + \text{heat}, \\ Xe^* + Xe + Xe &\rightarrow Xe_2^* + Xe, \\ Xe_2^* &\rightarrow 2Xe + h\nu. \end{aligned} \tag{1.12}$$

As the light is emitted by dimers whose energy states are different from the ones of individual atoms, xenon is transparent to its own scintillation light with a wavelength of 178 nm. The scintillation yield is comparable to that of NaI(Tl) with a faster time response [25].
The de-excitation of the singlet and triplet states of the excited dimers are both on the order of a few to tens of nanoseconds for different ionizing particles. This makes pulse-shape discrimination for different particle types non-feasible [25]. Still, discrimination between ER and NR interactions can be achieved by treating scintillation and ionziation as separate signal channels. For this purpose recombination can be partly suppressed by separating the ionization charges from the ions with an electric field. In this case light and charge at a given energy deposition become anti-correlated which can be seen in Fig. 1.10 [29].
At low energies E and a drift field of the strength $\mathcal{E}$ the combined probability P for generating N_γ photons and N_e escaping electrons can be approximated by two independent Poisson processes [30]:

$$P(N_\gamma, N_e | E_u, \mathcal{E}) \approx \text{Poi}(N_\gamma | n_\gamma) \cdot \text{Poi}(N_e | n_e). \tag{1.13}$$

Here, the expectation values $n_\gamma(E_u, \mathcal{E})$ and $n_e(E_u, \mathcal{E})$ depend on the deposited energies, the interaction type with $u = \text{nr}$ for NR and $u = \text{ee}$ for ER, and the electric field. The average energies for the creation of information carriers are influenced by all three aforementioned parameters and are usually expressed by effective W-values. The field dependence is factorized with functions S_u for the loss of light yield from field quenching effects and T_u for the loss of charge in recombination. One assumes $S_u(\mathcal{E} = 0) = 1$ and $T_u(\mathcal{E} \rightarrow \infty) = 1$.

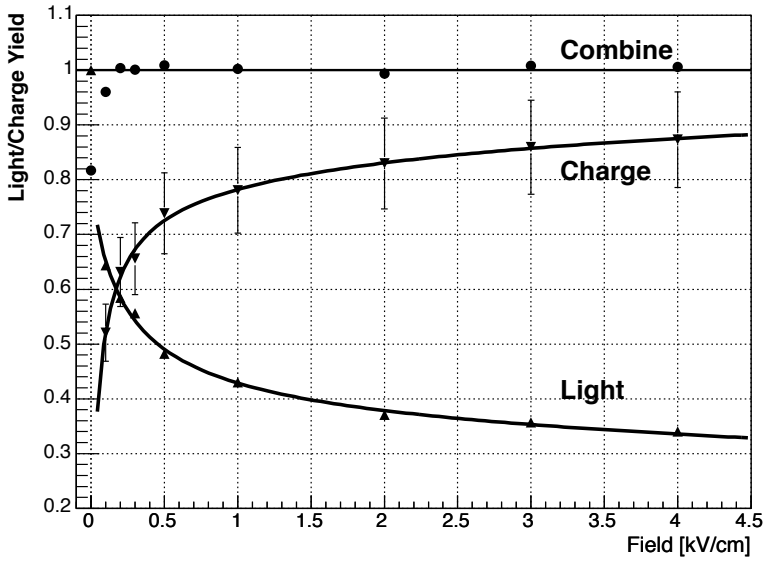

Figure 1.10: Light and charge yield as a function of the applied electric drift field from 661.2 keV γ-rays of a ^{137}Cs source. The light yield is relative to that at zero field. The anticorrelation can be seen: Higher light yields that correspond to proportionally lower charge yields at a given energy [29].

This allows to write the expectation values as

$$n_\gamma(E_u, \mathcal{E}) \approx \frac{E_u}{W_\gamma(E_u, \mathcal{E}=0)}\, S_u(\mathcal{E}), \tag{1.14}$$

$$n_e(E_u, \mathcal{E}) \approx \frac{E_u}{W_e(E_u, \mathcal{E}\to\infty)}\, T_u(\mathcal{E}). \tag{1.15}$$

In the data analysis of the XENON100 experiment an energy calibration of nuclear recoils is performed using 122 keV$_{\mathrm{ee}}$ as a reference energy, [30], where keV$_{\mathrm{ee}}$ denotes the electronic recoil-equivalent energy. The energy of 122 keV$_{\mathrm{ee}}$ corresponds to a γ-line of ^{57}Co which was historically used as a calibration source. Due to its low penetration depth it is not practical anymore given the growing dimensions of current detectors. The experimental light and charge yields at the reference energy can be used as fixed points in combination with dedicated measurements of the functions $W_{\gamma,e}$, S_u and T_u to build an energy scale for lower γ-ray energies and nuclear recoils. The light yield L_y at a reference energy E_{ref}, electric field $\mathcal{E}$ and position $\vec{r}$ in the TPC is given as

$$L_y(E_{\mathrm{ee}} = E_{\mathrm{ref}}, \mathcal{E}, \vec{r}) = \frac{S_{\mathrm{ee}}(\mathcal{E}) \cdot \mu(\vec{r})}{W_\gamma(E_{\mathrm{ee}} = E_{\mathrm{ref}}, \mathcal{E}=0)}. \tag{1.16}$$

The position dependent light detection efficiency $\mu(\vec{r})$ combines the quantum efficiency of the photodetectors with geometric parameters and absorption. It has to be measured with calibration sources.

The scintillation yields of nuclear and electronic recoils are different at the same fields and energy depositions. The relative scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$ of nuclear recoils with respect

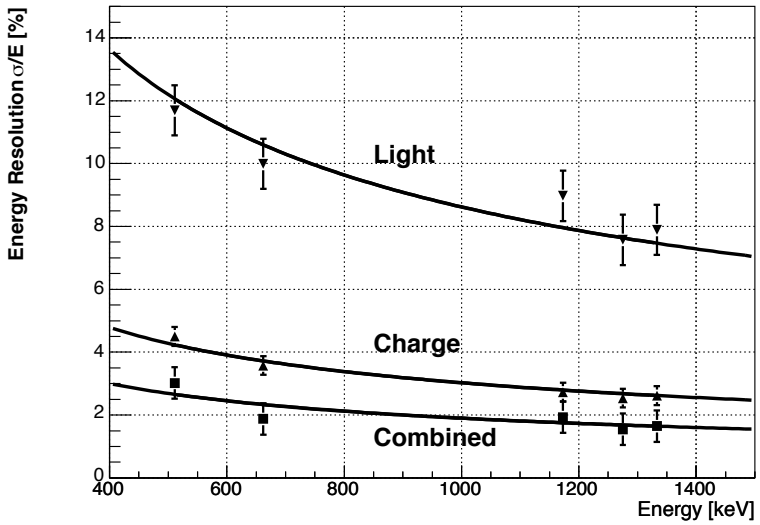

Figure 1.11: Energy dependent resolution of a liquid xenon detector measured from ^{22}Na, ^{137}Cs and ^{60}Co for three calibration methods: light only, charge only and combined light and charge. The applied electric drift field is 3 kV/cm [29].

to the reference γ-line at zero field is needed to convert E_{ee} to E_{nr}. It can be expressed as:

$$\mathcal{L}_{\mathrm{eff}}(E_{\mathrm{nr}}, \mathcal{E} = 0) = \frac{W_\gamma(E_{\mathrm{ee}} = E_{\mathrm{ref}}, \mathcal{E} = 0)}{W_\gamma(E_{\mathrm{nr}}, \mathcal{E} = 0)}. \tag{1.17}$$

At the time the XENON100 results were published $W_\gamma(E_{\mathrm{nr}})$ had been measured to lower energies than $W_e(E_{\mathrm{nr}})$. Hence, the energy scale for nuclear recoils was established from scintillation light only [30]. Although the W-values are different for the generation of photons and electron-ion pairs [25], simulation packages like the *Noble Element Simulation Technique* (NEST) assume $W_\gamma(E_{\mathrm{ee}}) = W_e(E_{\mathrm{ee}}) = 13.7$ eV for the sake of simplicity [31].
If only one of the signal channels, i.e. light or charge, is used for energy reconstruction the resolution is usually orders of magnitude worse than expected from the Fano factor. The reason for this are recombination fluctuations. The combination of both signal channels reduces the influence of fluctuations and provides a better energy resolution [29]. This is illustrated in Fig. 1.11.

1.4.3 Dual-Phase Time Projection Chamber

Xenon dual-phase time projection chambers (TPCs) make use of the principles illustrated in the section above. The typically cylindrical detectors contain liquid xenon as a target with a thin gaseous xenon layer at the top. Fig. 1.12 is a schematic representation of such a detector.
An interaction of a WIMP in the liquid xenon target produces both prompt scintillation light and ionization charges. These constitute two available signal channels and are measured in the form of photoelectrons by two PMT arrays at the bottom and top of the TPC. The working principle of photomultiplier tubes is explained in detail in [33]. The prompt scintillation signal (**S1**) is registered mostly by the bottom PMT array due to total reflection at the liquid-gas interface. The electrons from ionization that have not recombined are drifted towards the gas phase by an electric drift field. There they are

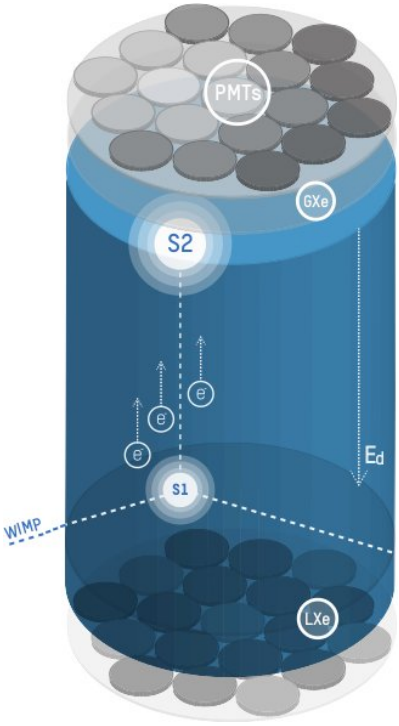

Figure 1.12: Schematic of a xenon dual-phase TPC. Used under Creative Commons license from [32].

extracted and accelerated by a stronger field producing a delayed electroluminescence signal (**S2**).
The product of the time delay t_d between S1 and S2 and the constant drift velocity v_d of the electrons in the liquid gives the depth (z-coordinate) of the interaction. The (x,y) coordinate can be derived from the hit-pattern of the S2 on the top PMT array by means of neural networks and fitting algorithms.

1.4.4 Measurement of the S1 Signal

The PMTs in the TPC measure the primary scintillation signal in the form of electrical pulses. These pulses are digitized and converted into units of photoelectrons (pe) by the raw data processor using derived values for the individual PMT gain factors including an electronic amplification factor of 10 [30]. The contribution $s1_i$ of each PMT to the full S1[5] signal can be written as:

$$s1_i(\vec{r}) \approx n_\gamma(E_u,\mathcal{E})\mu_i(\vec{r}). \tag{1.18}$$

The total light detection efficiency in L_y from Eq. 1.16 is the sum of the individual PMT light detection efficiencies $\mu_i(\vec{r})$. Hence, the expectation value for the area $s1$ of the total

[5] In the chosen notation S1 denotes the prompt scintillation signal without regard for its measurement. The lower case and italic $s1$ stands for the measured property which is an area of PMT pulses in photoelectrons. The signal processing from pulses to signals is explained in detail in the next chapter. An analogous notation is used for the S2 signal.

S1 signal for a nuclear recoil is

$$s1(\vec{r}) = \sum_{i=1}^{M} s1_i(\vec{r}) \approx n_\gamma(E_{\mathrm{nr}}, \mathcal{E})\mu(\vec{r}). \tag{1.19}$$

The data analysis is usually performed with an $s1$ that is corrected with the spatial average $\langle\mu\rangle$ of $\mu(\vec{r})$ in order to achieve a uniform signal in the whole TPC:

$$cS1 = s1(\vec{r})\frac{\langle\mu\rangle}{\mu(\vec{r})}. \tag{1.20}$$

Using equations 1.14, 1.16 and 1.17 the $cS1$ expectation value can be related to E_{nr} such that

$$cS1 = E_{\mathrm{nr}}\langle L_y(E_{\mathrm{ref}}, \mathcal{E})\rangle \cdot \mathcal{L}_{\mathrm{eff}}(E_{\mathrm{nr}}, \mathcal{E}=0)\frac{S_{\mathrm{nr}}(\mathcal{E})}{S_{\mathrm{ee}}(\mathcal{E})}, \tag{1.21}$$

where $\langle L_y\rangle$ is the detector-averaged light yield. Rearranging this equation gives the rule for the S1-based NR energy calibration for the dark matter search in this work:

$$E_{\mathrm{nr}} = \frac{cS1}{\langle L_y(E_{\mathrm{ref}}, \mathcal{E})\rangle \cdot \mathcal{L}_{\mathrm{eff}}(E_{\mathrm{nr}}, \mathcal{E}=0)} \cdot \frac{S_{\mathrm{ee}}(\mathcal{E})}{S_{\mathrm{nr}}(\mathcal{E})}. \tag{1.22}$$

1.4.5 Measurement of the S2 Signal

Several factors contribute to the magnitude of the measured S2 signals in a TPC. For instance, charge losses occur during the drifting of the electrons. This is due to an attachment to electronegative impurities in the liquid and results in a characteristic electron-lifetime τ_e. Continuous purification of the xenon increases the lifetime.
The extraction yield κ at the liquid gas interface depends on the extraction field $\mathcal{E}_{\mathrm{gas}}$. This field is also responsible for the amplification factor Y of the proportional scintillation. The expectation value for the contribution $s2_i$ of an individual PMT to the S2 signal in units of pe is described by [30]:

$$s2_i(\vec{r}) \approx n_e(E_u, \mathcal{E}) \cdot e^{\frac{-t_d}{\tau_e}} \cdot \kappa(\mathcal{E}_{\mathrm{gas}})Y\left(\frac{\mathcal{E}_{\mathrm{gas}}}{\rho}, h_g\right) \cdot \beta_i(x,y)\eta_i. \tag{1.23}$$

The gas density ρ and size of the gas gap h_g also influence the amplification. Both can be position-dependent due to a warping of the meshes for applying an electric field or an inclined liquid level. $\beta_i(x,y)$ is the probability for a photon created at (x,y) to reach the photocathode of the PMT and η_i is the product of the quantum and collection efficiencies for the PMT. The sum $\delta(x,y)$ over the PMT-wise products of $\delta_i = \kappa Y \beta_i \eta_i$ can be measured with calibration sources. The analysis is then performed with a spatially uniform corrected signal

$$cS2 = s2(\vec{r}) \cdot e^{\frac{t_d}{\tau_e}}\frac{\langle\delta\rangle}{\delta(x,y)}. \tag{1.24}$$

The expectation value for the area $s2$ of the total NR S2 can be expressed as

$$s2(\vec{r}) = \sum_{i=1}^{M} s2_i(\vec{r}) = E_{\mathrm{nr}} \cdot Q_y(E_{\mathrm{nr}}) \cdot e^{\frac{-t_d}{\tau_e}} \delta(x,y). \tag{1.25}$$

The charge yield Q_y for a nuclear recoil in units of e^- per $\mathrm{keV_{nr}}$ according to Eq. 1.14 is

$$Q_y = \frac{T_{\mathrm{nr}}(\mathcal{E})}{W_e(E_{\mathrm{nr}}, \mathcal{E}_{\mathrm{ref}})}. \tag{1.26}$$

A detailed discussion of the Poisson processes in the generation and measurement of the signals with the corresponding probability distribution functions for $s1$ as well as $s2$ can be found in [30].

1.4.6 The XENON100 Detector

The XENON100 detector [17] [30] is divided into two concentrical volumes which contain a total of 161 kg of ultra-pure liquid xenon. The inner volume is a dual-phase TPC of 30.5 cm height and 15.3 cm radius. A diving bell design is chosen to maintain a stable liquid level of the 62 kg liquid xenon target. In terms of photodetectors the TPC is instrumented with 178 Hamamatsu R8520 PMTs selected for a high quantum efficiency of up to 32 %. The 80 PMTs at the bottom of the TPC are arranged for maximum light collection efficiency while the 98 at the top form concentric rings for (x,y)-position reconstruction. The PMT arrays are depicted in Fig. 1.13. A schematic of the detector is also shown.
The inner volume of the detector is surrounded to all sides by a liquid xenon veto that is instrumented with 64 PMTs of the same type. Optical separation of the veto and the TPC is achieved by polytetrafluorethylene (PTFE) panels. The material has been chosen for its high reflectivity at the 178 nm wavelength of the xenon VUV scintillation photons. The white holding structure of the top PMT array that can be seen on the right of Fig. 1.13 is also made of PTFE.
A drift field of 530 V/cm is applied between the cathode mesh at the bottom of the TPC and the gate mesh close to the liquid gas interface. The drift velocity is $v_d \approx 1.73$ mm/μs. The extraction field of 12 kV/cm is applied between the gate and anode mesh. For all practical purposes the meshes can be assumed to be transparent. A detailed description of the full experiment and its subsystems can be found in [17].

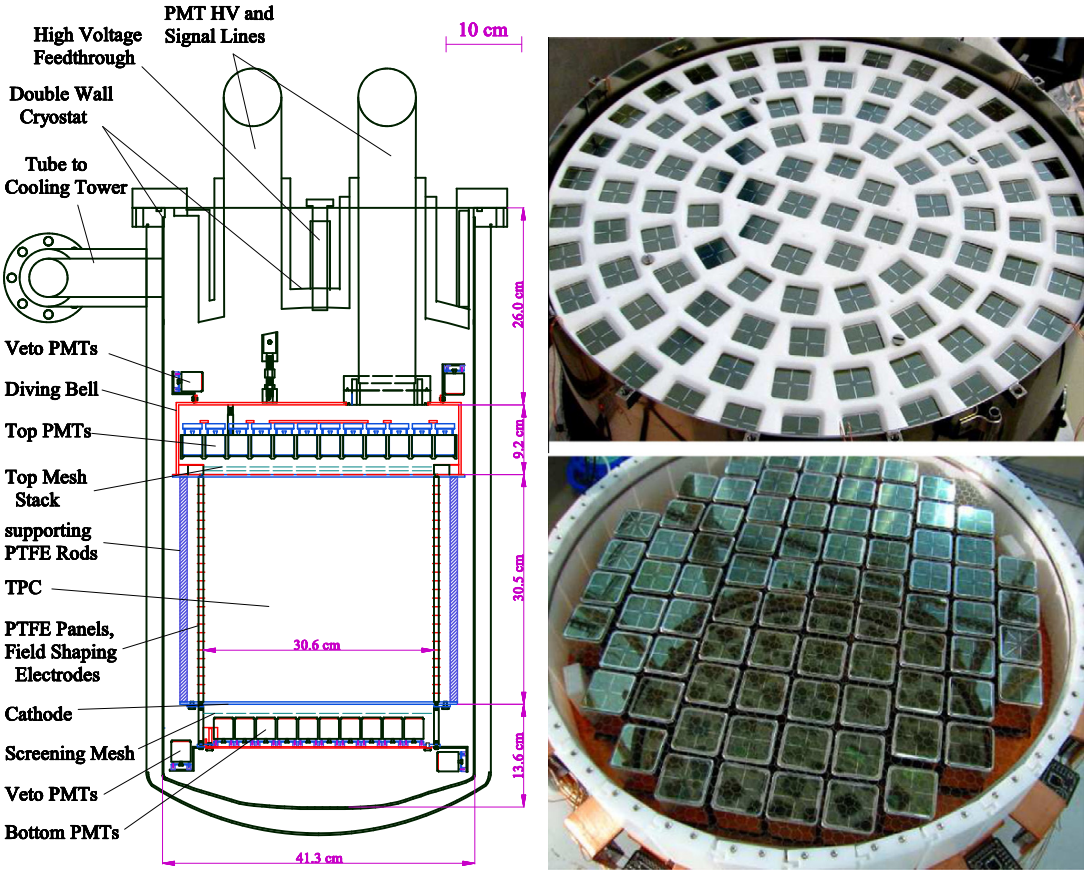

Figure 1.13: **Left:** Technical drawing of the XENON100 dark matter detector. **Right:** Top and bottom PMT arrays of the XENON100 detector before final assembly [17].

CHAPTER

TWO

MATERIALS AND METHODS

There are two realms of direct dark matter detection. The preceding chapter was centered around aspects rooted in the physical world describing particles interacting with highly specialized detectors. Now, this chapter focuses on the immaterial tools that translate measured quantities into scientific results. The first section introduces the *Processor for Analyzing Xenon* (PAX) which was developed for XENON1T. An agenda for the dark matter search with the processed data is set in the second section. It follows the example of the original analysis of the dataset at hand from [34]. The respective published result is introduced in the last section of this chapter.

2.1 The PAX Raw Data Processor

PAX has been developed in Python and utilizes its growing scientific ecosystem which provides packages for mathematics, data handling and manifold other purposes [35]. It is an open source program and customizable to handle data from other TPCs [36]. Physical events have to pass through a series of plugins which share a common structure. This series can be altered to use PAX for other tasks than data processing, for instance, conversion of file formats, event building and triggering.
The PAX processing chain is sketched in Fig. 2.1 and can be best understood following the developers' nomenclature. The blue tiles in the figure represent the PAX objects that are

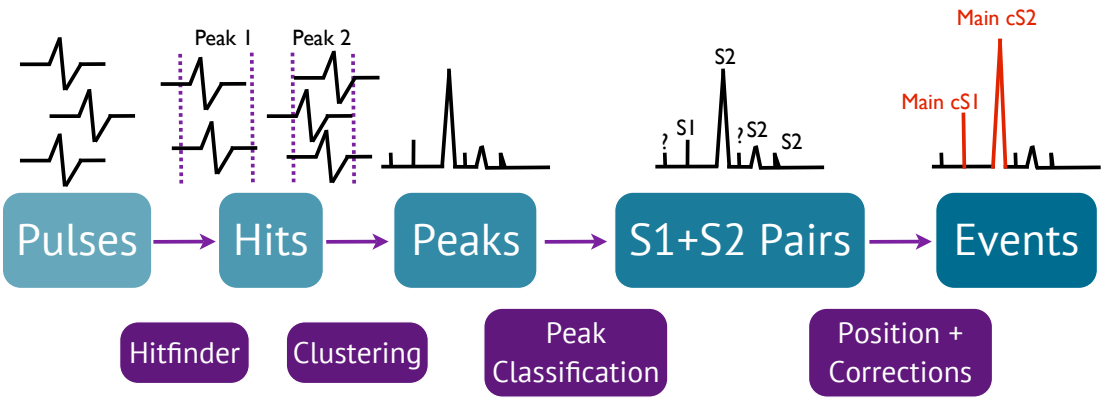

Figure 2.1: Schematic of the PAX processing chain. The blue tiles represent the PAX objects saved in the output file while the purple tiles represent the plugins.

stored in the output file of the processor. Here, the increasing opacity of the tiles represents the fact that every object incorporates information from the preceding processing step. These steps are identified by the purple tiles which also correspond to processor plugins. In PAX a physical *event* contains pairs of S1 and S2 *peaks* resulting from *interactions* inside the TPC. These *peaks* are the summed coincident *hits* of photons in multiple PMT channels. A *hit* is a *pulse* that shows a sufficient excursion above the baseline of the waveform while *pulses* are regions of raw digitized data of the individual PMT waveforms. Fig. 2.2 shows an example waveform consisting of 10 ns-samples recorded by the ADCs which is treated as a *pulse* by PAX. The baseline of the waveform is calculated from samples at the beginning and end of each *pulse* and the higher of both values is used. Two *hitfinder* thresholds are set dynamically using the standard deviation σ of the baseline. If a sample exceeds the lower threshold at 1σ over baseline the *hitfinder* starts to sum the consecutive samples until falling below threshold. These are stored as a *hit* if one of the contained samples surmounts the high threshold at 5σ. One *pulse* can contain multiple *hits* whose unit is photoelectrons. These are calculated from the ADC counts applying the PMT gain values. Coincident *hits* are clustered into *peaks* in two steps. A rough clustering is based on the gaps between hits. Then a finer clustering is achieved with a recursive natural breaks algorithm. If a channel exhibits a lone *hit* that is not coincident with one recorded by a

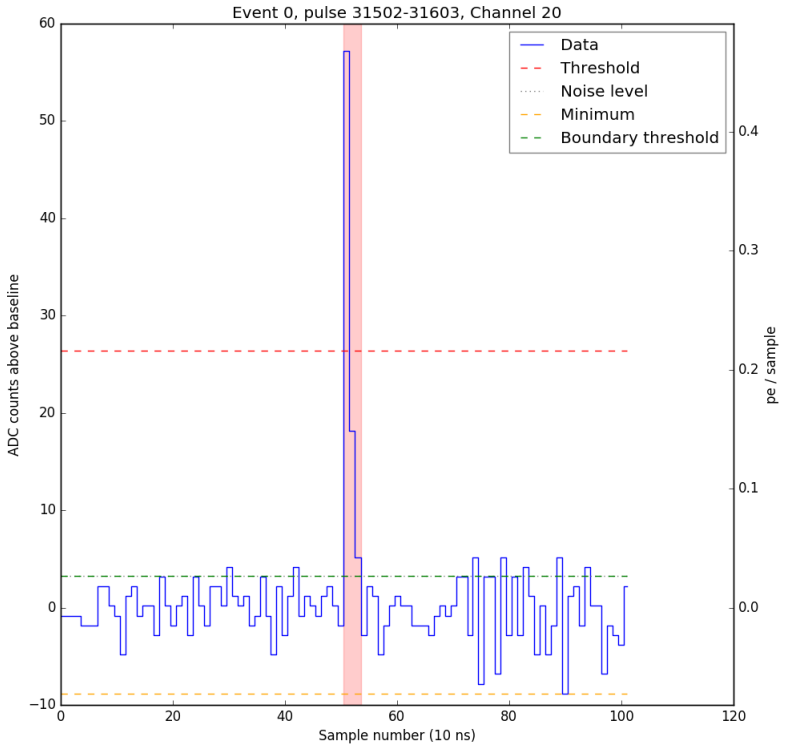

Figure 2.2: Example waveform from the first event of a XENON100 raw data file for PMT channel 20. The y-axis on the left shows ADC counts above baseline which are converted into photoelectrons per sample using the PMT gain calibration. The x-axis states the sample number of the waveform with a sample size of 10 ns given by the sampling rate of the ADCs. The dashed lines are the sample minimum and the thresholds calculated from the standard deviation of the baseline. In this case they are at 1σ and 5σ. The gray dotted line is the noise level coinciding with the boundary threshold. The area shaded in red is a *hit*.

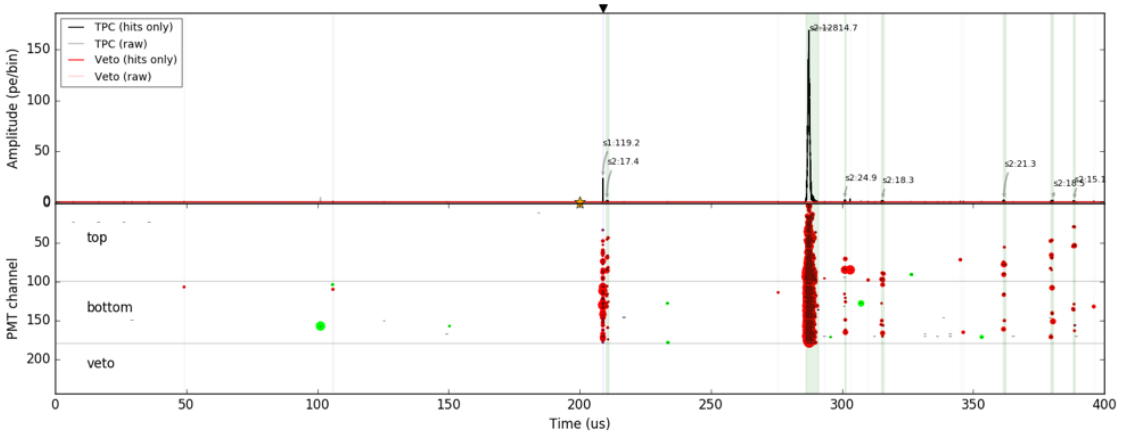

Figure 2.3: Summed waveform of a PAX example event. The top panel shows the classified *peaks* summed from clustered *hits*. The bottom panel shows the *hits* with contributing ones shaded in red and those from suspicious channels in green.

different PMT, it is marked as suspicious. This two-fold coincidence criterion makes PAX robust against noise. The uncorrected *peak* area A in pe is then calculated as the sum of all contributing *hits*.

In the next step *peaks* are classified according to their area and width. In this case, the width is classified as the time window around the *peak* center that contains 50 % of the *peak* area. *Peaks* with an area $A > 50$ pe are classifed as an S1 if their width is smaller than 100 ns. Otherwise they are labeled S2. For $A < 50$ pe an S1 must have a width smaller than 75 ns. Special cases are present for *peaks* that cannot be classified.

Pairing the S1s and S2s in order to form *interactions* is the penultimate processing step. All physical pairings where the S1 precedes the S2 are saved into an array. The first pair in this array is the main *interaction* comprised of the largest area S1 and S2 in the event[1]. As an *interaction* contains the full information of S1 and S2, a position can be reconstructed. The reconstructed positions and the position dependent signal corrections for obtaining $cS1$ and $cS2$ are saved for all pairings. The final *event* is built from all *interactions* occuring in a time window of 400 μs and contains the full set of processed information down to the level of *pulses*. The summed waveform of an *event* with the classified peaks and *hits* in all PMT channels is depicted in Fig. 2.3.

2.2 XENON100 Dark Matter Search Agenda

The data analyzed in this work has been subject to a number of publications dedicated to different types and interactions of dark matter [22, 27, 34, 37–41]. Its analysis will follow the example of the WIMP search carried out in [34] which is detailed in [30]. The total livetime of the data recorded over the course of 13 months from 28th February 2011 to 31st March 2012 is 224.6 days. The trigger of the data acquisition system (DAQ) was set to save data if the summed signal of all PMT channels exceeded 150 pe. Key parameters of the science run can be found in Tab. 2.1. Now, the full dark matter search requires the completion of several analysis steps in a specific order. Starting after the completion of

[1] It should be kept in mind as a caveat that the old data processor for XENON100 defined the main *interaction* as the pairing of the most coincident S1 with the largest area S2.

data acquisition these are:

1. Blinding
2. Selection and processing of recorded datasets
3. Detector characterization and calibration
4. Definition, acceptance and application of cuts
5. Background expectation
6. Unblinding and statistical evaluation of the data

2.2.1 Blinding

In order to avoid the introduction of bias, a blind analysis is standard in low count experiments and is, consequently, applied by the XENON collaboration. The data in a pre-defined dark matter search region are made inaccessible during the steps two to five [30, 34]. No blinding is performed in this work for two reasons. First, an already unblinded dataset with published analysis results is analyzed. Second, this work is not aimed at competing with other direct dark matter searches, but at understanding the difference PAX makes in the XENON100 analysis compared to the old raw data processor.

2.2.2 Data Selection and Processing

Analysis data is selected on the grounds of stable detector operating conditions which are permanently monitored with the experiment's slow control system. Relevant operation parameters include xenon pressure and temperature, cathode and anode voltages, purification flow, cryostat vacuum pressure and the liquid xenon level. Data showing irregularities, e.g. a 5σ deviation from the mean value of the parameter of interest, is excluded from the analysis. Periods with increased electronic pickup noise are also removed from the dataset. Consequently, the exposure is calculated from the remaining data. This work uses the same data selection as the original analysis for a direct comparison of the remaining events after cuts.

Table 2.1: Key parameters of the dark matter dataset analyzed in this work and [30].

Parameter	Value
Livetime [days]	224.6
Anode voltage [kV]	4.4
Cathode voltage [kV]	-16.0
DAQ deadtime	$< 1\,\%$
DAQ trigger threshold [pe]	150

2.2.3 Detector Characterization and Calibration

The position-dependent light collection efficiency (LCE) and electron lifetime are required by the data processor for the position corrections of the S1 and S2 signals. Both are derived from weekly calibrations with a ^{137}Cs source. The γ-line at 662 keV$_{\mathrm{ee}}$ allows to monitor the evolution of light yield, charge yield and liquid xenon purity, as well as the width of S2 pulses. The electron lifetime during the science run continuously increased and ranged between $374 - 661\,\mu$s. The electron lifetime is saved in the raw data files while the LCE map and PMT gains from weekly LED calibrations are integrated into PAX via the XENON100 configuration file. Therefore, this calibration step will not be repeated in this work. An overview of the calibration parameters from the original analysis can be found in Tab. 2.2.

The detector calibration with respect to ER and NR interactions is carried out with radioactive sources [30]. The response to ER interactions is characterized by the low energy Compton continuum of a ^{60}Co γ-source and the broad ER spectrum of a ^{232}Th source. The nuclear recoil region is defined with elastic neutron interactions from an (α,n) ^{241}AmBe source. Moreover, inelastic scattering of the AmBe neutrons leads to an activation of the liquid xenon target itself and gives γ-lines in an energy range between 39.6 keV and 319.9 keV. These prompt deexcitation lines from ^{129}Xe and ^{131}Xe as well as the delayed deexcitations from ^{129m}Xe and ^{131m}Xe are used for the energy calibration.

As outlined in section 1.4.4, the energy calibration for nuclear recoils is based on the S1 signal. $L_y(122\,\mathrm{keV_{ee}}) = (2.28 \pm 0.04)\,\mathrm{pe/keV_{ee}}$ is determined from interpolating the light yields of the neutron-activated xenon and ^{137}Cs γ-lines with the NEST model [31]. The relative scintillation efficiency of nuclear recoils $\mathcal{L}_{\mathrm{eff}}(E_{\mathrm{nr}})$ is parametrized by overlaying all existing direct measurements with a cubic spline [42]. This is shown in Fig. 2.4. Below 3 keV$_{\mathrm{nr}}$ a logarithmic extrapolation is applied, so that $\mathcal{L}_{\mathrm{eff}} = 0$ at 1 keV$_{\mathrm{nr}}$ [42]. The field parametrization is given by $S_{\mathrm{nr}} = 0.95$ and $S_{\mathrm{ee}} = 0.58$.

Both L_y and the S1 energy calibration are re-evaluated in this work, as these are influenced by the way PAX builds S1 and S2 peaks. In addition, a combined energy scale incorporating the information from both the S1 and the S2 signal is determined.

Table 2.2: Calibration parameters from the analysis in [30].

Parameter	Value
Livetime ER calibration [days]	48.0
Livetime ^{241}AmBe [days]	2.7
Electron lifetime AmBe [μs]	360
Electron lifetime science data [μs]	$374 - 611$
$L_y(122\,\mathrm{keV_{ee}})$ [pe/keV$_{\mathrm{ee}}$]	2.28 ± 0.04
$\frac{S_{\mathrm{ee}}}{S_{\mathrm{nr}}}(\mathcal{E} = 530\,\mathrm{V/cm})$	$\frac{0.58}{0.95}$

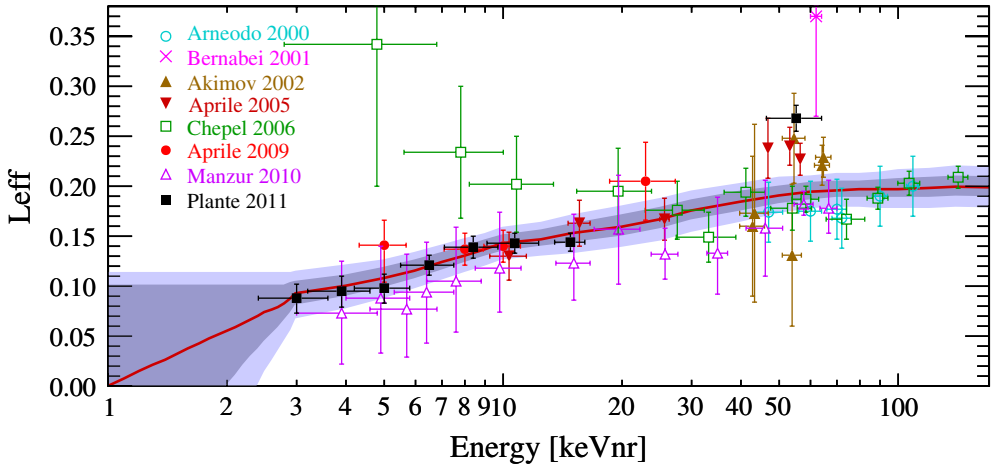

Figure 2.4: All direct measurements of $\mathcal{L}_{\text{eff}}$ described by a Gaussian distribution to obtain the mean (solid red line) and the uncertainty band (1σ and 2σ). Below $3\,\text{keV}_{\text{nr}}$ the trend is logarithmically extrapolated to $\mathcal{L}_{\text{eff}} = 0$ at $1\,\text{keV}_{\text{nr}}$ [42].

2.2.4 Cuts

Cuts separate potential signal events from background, electronic noise and other experiment-specific pathologies that are picked up by the highly sensitive detectors. In general, cuts are selection criteria that are applied on the science data in order to select dark matter candidate events in a region of interest (ROI) [30]. These criteria span from data quality, energy selection and signal threshold, over single scatter event selection, data consistency and an event location in the fiducial volume to a final signal region selection. A few examples for these categories are presented here for a general overview while the full set of cuts is presented in detail in section 3.1.

Data quality cuts remove noisy waveforms and those that unlikely result from a physical interaction inside the TPC. For example, small widths of S1 peaks can result from electronic noise pulses with samples fluctuating around the baseline. These effectively shorten the extent of the identified peak due to negative excursions. They can, thus, be removed by requiring a minimal S1 width. Additionally, sanity checks like the existence of at least one S1-S2 pair in a waveform fall into this category. For instance, a lone S2 can occur if the S1 is too faint or to close to the S2 signal to be resolved by the detector.

The second category of selection criteria are **energy selection and threshold cuts**. Dark matter interactions are expected to occur in an energy interval where the bounds are given by a number of parameters. The upper bound is optimized for sensitivity by cutting off at recoil energies where an increase of the energy window does not enhance sensitivity further due to increasing background. The detector and trigger thresholds set the lower bound.

Cosmogenic neutrons produce low energy nuclear recoils inside the TPC that perfectly resemble a WIMP interaction. However, in contrast to WIMPs neutrons scatter multiple times in the detector producing a cumulated S1 and multiple S2 signals. Thus, a **double scatter cut**[2] removes events with multiple S2 signals. Additionally, waveforms with a coincident trigger of the liquid xenon veto are removed as these originate from particles

[2]The double scatter cut is in the following also referred to as a single scatter selection, as it preserves single scatter events by cutting double scatters out of the data.

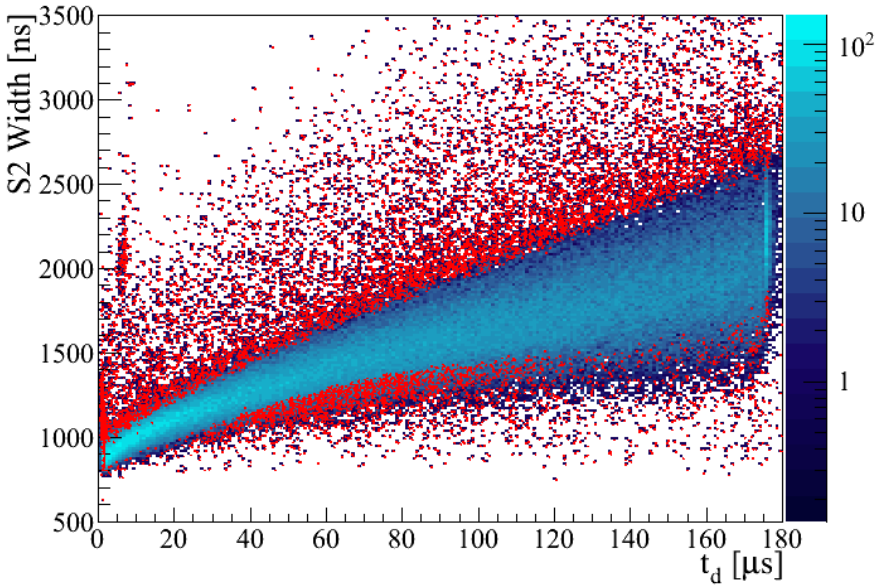

Figure 2.5: S2 width distribution (defined at 10 % of the peak height) versus drift time for nuclear recoils with energies up to 30 pe in S1 (blue histogram). The red dots are the events removed by the S2 width selection [30].

with multiple energy depositions such as muons and Compton-scattering γ-rays.

Consistency cuts distinguish valid events from experiment-specific pathologies introduced by detector effects or the data processor itself. They include comparing the difference of the reconstructed positions provided by several algorithms and evaluating the likelihood of PMT hit patterns. A well-motivated consistency check is the width of the S2 signal as a function of the drift time t_d. Due to diffusion of the electron cloud the S2 signal widens the deeper an interaction occurs inside the TPC. Events that exhibit a large deviation of the S2 width from the mean at a given t_d are excluded from the analysis as they likely result from interactions in the gas phase or unphysical pairings of S1 and S2 signals. The cut from the original XENON100 analysis is shown in Fig. 2.5.

The **fiducial volume** is chosen by maximizing the rejection of external backgrounds while maintaining an as large as possible mass of the liquid xenon target. The result is an ellipsoidal fiducial volume containing 34 kg of liquid xenon. Fig. 2.6 shows the fiducial volume cut from the original publication [34] that analyzed the dataset from this work with the original XENON100 data processor. The volume's shape is indicated in red and the borders of the TPC are marked by the yellow dashed line. Events reconstructed outside the TPC are due to the uncertainty in the position reconstruction. The event distribution shows the full dark matter science data with a discrimination of nuclear and electronic recoils applied to it. Gray points are rejected as falling above an ER discrimination line while black points are below it and, consequently, kept.

The discrimination is derived from the bands of ER and NR datapoints of the ^{60}Co and ^{232}Th sources, and the elastic neutron scatters from the AmBe source, respectively. The key to discrimination are the different ionization and scintillation yields of electronic and nuclear recoils that translate to a different value of the discrimination variable $\log_{10}\left(\frac{cS2_b}{cS1}\right)$ at a given $cS1$ or energy. Here $cS2_b$ is the fraction of the corrected S2 peak area seen by the bottom PMT array which is less prone to saturation than the top array. The discrimination

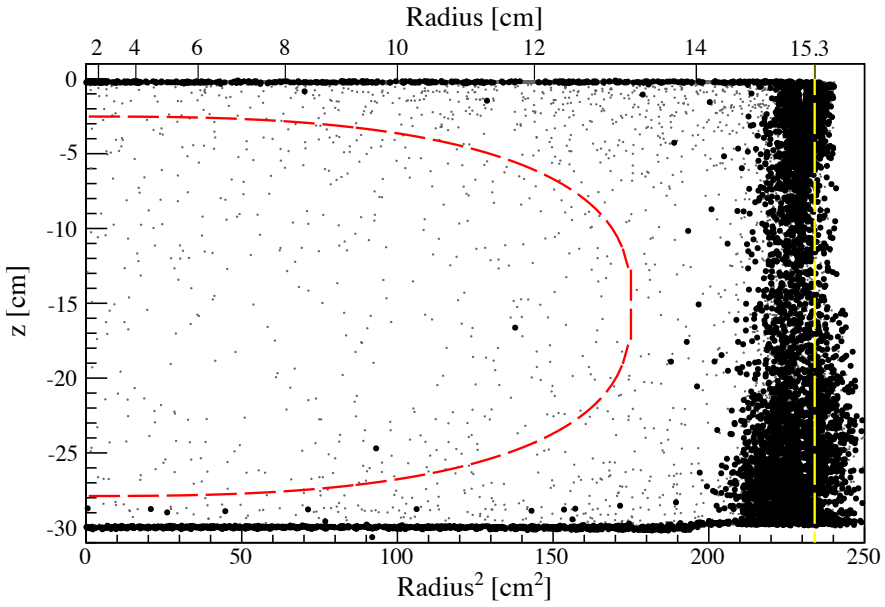

Figure 2.6: Spatial event distribution inside the TPC using a $6.6 - 43.3\,\mathrm{keV_{nr}}$ energy window. The 34 kg fiducial volume is indicated by the red dashed line. Gray points are above the electronic recoil rejection line, black circles fall below [34].

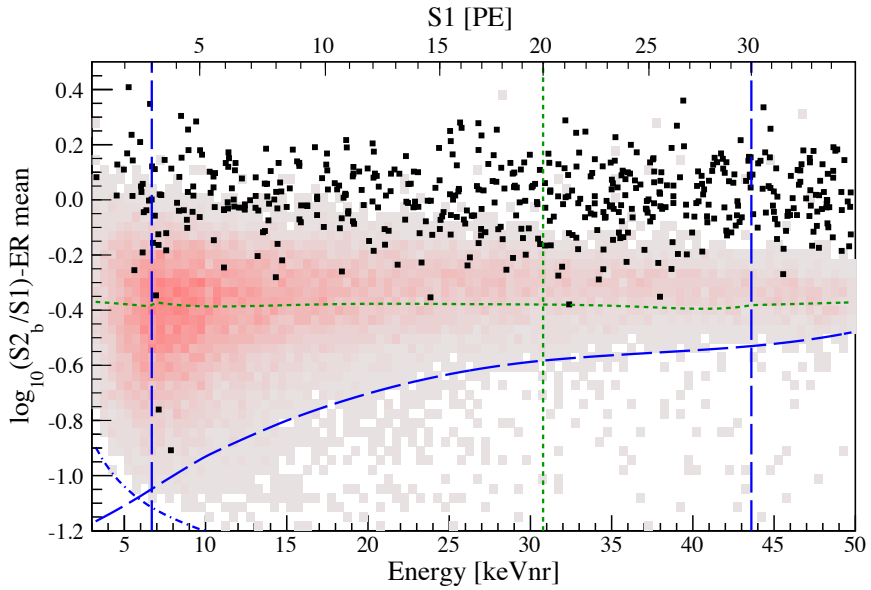

Figure 2.7: Dark matter data after cuts from [34] in flattened discrimination space as a function of E_{nr} and S1 area. The black datapoints are the remaining events after all cuts. The gray points show the nuclear recoil band from AmBe data. The dashed blue line following its lower contour is the 97 % quantile of the band which is the lower bound for accepting events. The dashed blue lines left and right mark the energy window between $6.6 - 43.3\,\mathrm{keV_{nr}}$, corresponding to an S1 area between $3 - 30$ pe, used for the dark matter search with a Profile Likelihood analysis. The vertical, dotted and green line marks the upper S1 boundary at 20 pe for a cut-based analysis with a discrimination cut. This discrimination cut is marked by the horizontal, dotted and green line which is at 3.09σ from the mean of the Gaussian ER band from ^{60}Co and ^{232}Th data.

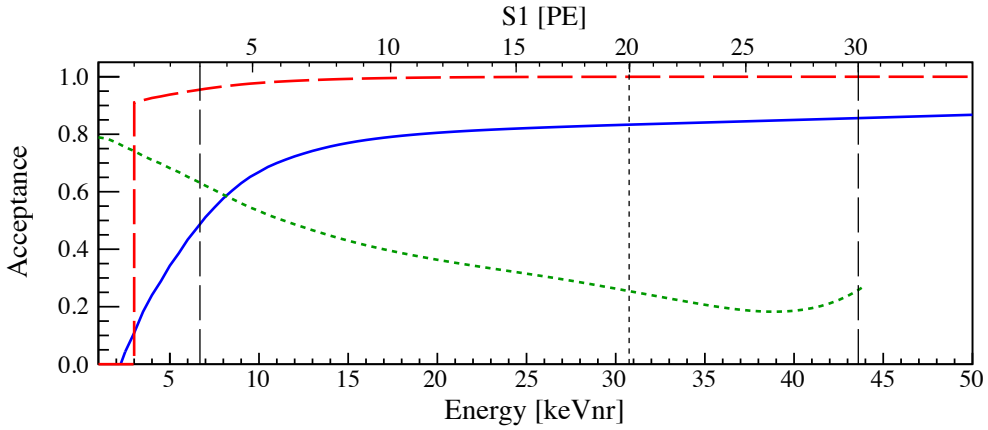

Figure 2.8: Combined cut acceptance from [34] (blue line). The S2 threshold cut requiring $S2 > 150\,\mathrm{pe}$ (dashed red line) is independent of possible fluctuations in S1 and has to be applied to the S1 spectrum before taking into account the S1 resolution. It is conservatively set to zero below 1 pe. The hard discrimination cut's acceptance to NRs is shown by the dotted green line. The lower analysis energy thresholds for the Profile Likelihood are $6.6 - 43.3\,\mathrm{keV_{nr}}$, whereas the cross-check is restricted to $30.5\,\mathrm{keV_{nr}}$ (dashed and dotted black lines).

plot from the published result in [34] is depicted in Fig. 2.7. The discrimination variable is given as a function of E_{nr} in flattened discrimination space. Band flattening removes the energy dependence of the ER band from the plot. Here, the mean value of the discrimination variable for the ER band is subtracted from the discrimination variable values of the individual datapoints:

$$\text{Flattened discrimination variable} = \log_{10}\left(\frac{cS2_b}{cS1}\right) - \text{ER mean}. \tag{2.1}$$

The Gaussian distribution of the ER band around the mean is used for defining an ER and NR discrimination cut. It is given by a 3.09σ deviation from the ER band mean which translates to a 99.75 % rejection of ER events inside this interval. The lower acceptance line for nuclear recoils is defined by the 97 % quantile of the NR band.

Apart from removing background, noise and other detector pathologies, cuts also discard signal events. The fraction of surviving signal events is denominated the cut acceptance. It is evaluated on NR calibration data for each cut using the fraction of events dismissed by this selection alone while still applying all other cuts. This approach addresses the fact that an acceptance calculation with calibration data cannot rely on a pure event sample without contaminations from non-signal events. The application of $n - 1$ cuts minimizes this contamination when testing the acceptance of cut n. Therefore, the resulting acceptance estimate for each cut is rather conservative. In addition to the NR calibration data the non-blinded part of the science data is used for checking the acceptance of data quality cuts, as they could exhibit a time-dependence with changing detector conditions. The combined acceptance of the full cut set is the product of the individual acceptances. The result for the combined cut acceptance from [34] is shown as a function of E_{nr} and S1 in Fig. 2.8 as a solid blue line. The dotted green line marks the nuclear recoil acceptance of the discrimination cut introduced in Fig. 2.7. The threshold cut on the minimum uncorrected

S2 area is treated separately and shown as a dashed red line. It is treated separately due to its independence of the S1 resolution [34]. Its acceptance is conservatively set to zero at energies below the analysis threshold.
PAX introduced a new datastructure, so all XENON100 cuts are translated for it in the scope of this work. The background discrimination in flattened discrimination space is remade while retaining the original definition of the ROI in pe. This is necessary because the shape of the ER and NR bands changes with cuts that have a different effect due to the way PAX calculates the properties of *peaks* and *interactions*. This also necessitates a recalculation of the cut acceptance.

2.2.5 Background Expectation

After cuts and background discrimination a leakage of background events into the region of interest is still expected, as statistically only 99.75 % of ER events are rejected. Ergo, a background prediction in the ROI has to be formulated. This uses both calibration data scaled to the dark matter data as well as Monte Carlo simulations of the background components introduced in Tab. 1.1. While the majority of ER background events follows a Gaussian distribution anomalous leakage into the NR band can be present [34]. Such events are explained by double scatters with only one energy deposition inside the TPC and a second scatter in a charge-insensitive region. Consequently the combined S1 of both interactions is paired with the single S2. Taking this into account, the total background expectation in the ROI between $(3-20)$ pe is (1.0 ± 0.2) events [34]. This is also adopted in this work.

2.2.6 Unblinding and Statistical Evaluation

The last step in the analysis after unblinding is the evaluation of the remaining events in the ROI against the background expectation. The original analysis of the dataset examined in this work was performed using a Profile Likelihood method which does not employ a fixed discrimination of electronic and nuclear recoils. The application in the ROI of $6.6-43.3\,\mathrm{keV_{nr}}$ is described in detail in [30] and [34]. A cuts-based maximum

Table 2.3: Parameters for the statistical evaluation of events after cuts [30].

Parameter	Value
Profile Likelihood ROI [pe]	$3-30$
Profile Likelihood ROI [$\mathrm{kev_{nr}}$]	$6.6-43.3$
Maximum Gap ROI [pe]	$3-20$
Maximum Gap ROI [$\mathrm{kev_{nr}}$]	$6.6-30.5$
Background Expectation [events]	1.0 ± 0.2
ER discrimination	99.75 %
Lower NR contour	97 %
Fiducial mass [kg]	34

gap [19] analysis with a fixed 99.75 % ER rejection was used as a cross-check in a smaller energy region from $6.6 - 30.5\,\mathrm{keV_{nr}}$. A summary of the parameters for the statistical analysis is given in Tab. 2.3. In this work, a Feldman & Cousins approach [18] is applied on the events after cuts situated in the original Maximum Gap ROI. An upper limit on the WIMP-nucleon interaction cross section as a function of m_χ is derived taking the background expectation into account.

2.3 Original Dark Matter Search Result

After unblinding two events remained in the ROI for the Maximum Gap analysis [34]. It can be seen in Fig. 2.7 that both are close to the energy threshold with $7.1\,\mathrm{kev_{nr}}$ (3.3 pe) and $7.8\,\mathrm{kev_{nr}}$ (3.8 pe), respectively. No events were observed below 3 pe. Both statistical analysis methods were in agreement with a background-only hypothesis and provided exclusion limits compatible in terms of the known systematic differences. The exclusion limits for σ_{SI} were calculated assuming an isothermal WIMP halo with the key parameters given in Tab. 2.4. The limit calculated with the Profile Likelihood approach is given in Fig. 2.9 as a blue line. It is situated within the expected sensitivity of the experiment in the absence of a dark matter signal indicated as green (1σ) and yellow (2σ) bands. At the time of publication it represented the most stringent limit for σ_{SI} at WIMP masses m_χ above $8\,\mathrm{GeV}/c^2$ with a minimum of $\sigma_{\mathrm{SI}} = 2 \cdot 10^{-45}\,\mathrm{cm}^2$ at $m_\chi = 55\,\mathrm{GeV}/c^2$. It is now superseded by more recent limits from the *XENON* collaboration, as well as the *PandaX-II* and *LUX* collaborations which are shown in Fig. 1.9.

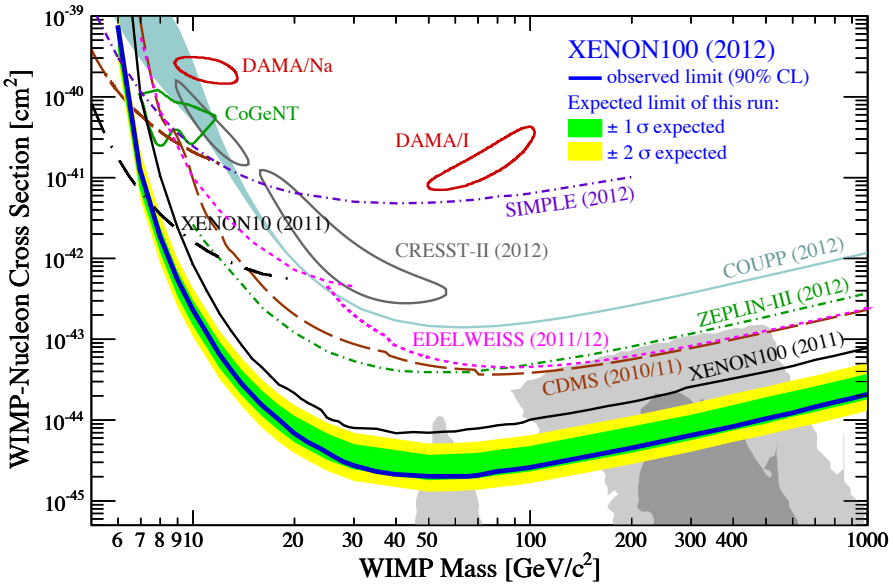

Figure 2.9: Result for σ_{SI} as a function of m_χ from XENON100. Expected sensitivity of the 225 live day run is shown by green and yellow bands ($1\sigma/2\sigma$). The resulting exclusion limit (90 % CL) is depicted in blue. Other experimental limits (90 % CL) and detection claims (2σ) from the time of publication and regions ($1\sigma/2\sigma$) preferred by supersymmetric (CMSSM) models are also shown [34].

Table 2.4: Parameters for the limit calculation with Eq. 1.7 [34].

Parameter	Value
Local dark matter density ρ_0 [GeV/cm^3]	0.3
Minimum relative velocity v_0 [km/s]	220
Escape velocity v_{esc} [km/s]	544
Acceptance-corrected exposure [kg $\cdot$ days]	2323.7

CHAPTER

THREE

CUTS AND CALIBRATION

With an agenda for the dark matter search fixed it is time to set the stage by preparing the dataset through cuts and calibration. First, the cuts used in the original analysis, from now on referred to as legacy cuts, are introduced and their application is motivated. Then they are translated to conform to the variables provided by PAX. Incompatible cuts are highlighted and an S2 width consistency cut is newly developed in order to replace the respective legacy cut. The second section is dedicated to the calibration of the detector with γ-rays. These are provided by an external ^{137}Cs source as well as neutron activated isotopes inside the detector. Three methods for a combined ER energy calibration incorporating the information from both S1 and S2 are introduced. The following nuclear recoil energy calibration is based on the S1 signal alone in analogy to the original analysis.

3.1 XENON100 Cuts and PAX Variables

The use of the legacy cuts is motivated by the reduction of systematics challenging the comparability of the old and new analysis results. If cuts are tuned to the advantages and pitfalls of PAX in a non-blind analysis, a bias may be hard to distinguish from an effect that is actually induced by the processor[1]. Moreover, bearing in mind that the XENON100 analysis effort has been wrapped up, the development of new XENON100-exclusive cuts is not economical. An emphasis should be put on developing code applicable in forthcoming XENON1T analyses. Still, the XENON100 cuts are highly adapted to the characteristics and data structure of the old data processor, so a few points have to be addressed regarding their use.

First of all, the cuts have to be translated owed to the different data structures of the two processors. Tab. 3.1 gives a schematic overview of the legacy cuts and their function. A formal definition of the translated cuts with a detailed description of the variables can be found in the appendix, section 6.2. Not all variables given by PAX are directly comparable to the ones provided by its predecessor. In some cases the units of a variable are different, some have slightly different values and others have been dropped entirely. Therefore, three

[1] Cut development and refinement also constitute an extensive analysis effort which entertains dozens of researchers over the course of years. Introducing a completely new set of cuts with reasonable rejection powers and acceptances would have overstretched the time resources for this master thesis. Therefore, an existing set of cuts presents a good stepping stone.

Table 3.1: Overview of XENON100 legacy cuts. Cuts that are not directly translatable are marked with a *.

Cut	Selection criteria
S1 coincidence	S1 is seen by more than one (non-noisy) PMT.
S1 width	Width is larger than 26 ns.
S2 trigger threshold	S2 area is above trigger threshold 150 pe.
S2 seen by top PMTs	Fraction of the S2 area seen by the top PMT array is larger than zero.
Fit quality of position reconstruction	Checks $\chi^2_{(x,y)}$ of position reconstruction.
Consistency of position reconstruction	Deviation of reconstructed (x,y)-positions from different algorithms is small.
Veto	No signal observed in liquid xenon veto.
Avoid early S2s	Main S2 can only occur after 178 μs.
S2 asymmetry	Remove events with a high S2 signal asymmetry between top and bottom PMT array which is characteristic for gas phase events.
γ-like events in AmBe neutron band	Cut in discrimination space: $\log_{10}\left(\frac{cS2_b}{cS1}\right) < 3.1$
Signal to noise	The difference between the area of all peaks in the waveform and the summed area of the main interaction S1 and S2 is close to zero.
Single S1	Remove events that have more than one S1 of either a physical origin or caused by noise.
S2 single scatter	Remove double scatter events with secondary S2 signals above a given area ratio between secondary and main S2.
Signal entropy*	Entropy variable is not implemented in PAX.
S2 width consistency*	Cut events showing unphysical S2 widths as a function of t_d. The cut is performed for different ranges of $s2$ with a flattened and corrected width variable which is not implemented in PAX.
S1 hit pattern*	Check if S1 hit pattern on bottom PMT array matches the one expected from the reconstructed position. Pattern likelihood variable not implemented in PAX.
34 kg fiducial volume	Ellipsoidal inner detector volume.

cuts, namely the entropy, the S1 PMT pattern likelihood and the S2 width consistency cut, could not be translated. While the two former ones are dropped in the analysis, a simple S2 consistency cut is newly defined in section 3.1.1.

Disparities in variable values between processors can be attributed to the respective event builders. The old processor searched for, identified and quantified S1s and S2s on the

summed waveform of all PMT channels. This made it susceptible to noise. PAX on the other hand does not run a peakfinder on the summed waveform, but clusters coincident *hits* from individual channels first. Due to the dynamic threshold setting in each channel and the exclusion of suspicious channels, PAX shows an inherent robustness against a certain level of noise. Therefore, noise discrimination variables like the signal entropy are entirely dropped by PAX. However, this causes the calculated peak areas, widths and PMT contributions to differ slightly which will become evident when comparing the properties of the dark matter candidate events from [34] in both processors. In addition, the peak classification is different in PAX. The main S1 in an event is assigned on the basis that it has the largest area of all S1 peaks in the waveform. In the old processor the main S1 was defined as S1 seen by most PMTs. At last, the old data processor did not continue searching for peaks after the main S2 in the waveform while PAX does so. Therefore, the cumulative area of all peaks in a waveform, which is used in one of the noise cuts, should be larger in PAX.
Such discrepancies can cause the rejection of PAX-processed events that have been regarded as valid events before. The opposite case of accepting an originally rejected event is also possible. Still, these tuning effects are not direct counterarguments against applying the legacy cuts in this analysis. A higher event rejection can be accounted for by the cut acceptance which is calculated in section 3.2. Remaining pathological events in the dark matter search region post cuts are evaluated on an event by event basis. This way PAX is tested on real dark matter data with a published result as the benchmark.

3.1.1 S2 Width Consistency Cut

A newly developed S2 width consistency cut replaces the respective legacy cut which could not be translated. Fig. 3.1 shows the S2 width of ER events from ^{232}Th and ^{60}Co sources at 90 % of the peak area against t_d. Typical signal widths in the energy region of $cS1 < 150\,\mathrm{pe}$ range between $(1 - 1.5)\,\mu$s. The band of S2 widths is divided into slices of $\Delta t_d = 10\,\mu$s and each slice is fitted with a Gaussian distribution. The fit gives the mean and the $\pm 5\sigma$ exclusion intervals for the cut.
In order to obtain continuous functions of the bounds and the mean, the points for all slices are fitted with a square root function motivated by diffusion theory [25]:

$$\text{S2 width}_{90} = \sqrt{a \cdot t_d + b}. \tag{3.1}$$

In the original analysis the width cut was defined taking into account the energy dependency of the S2 width, as S2s from low-energy events show a larger spread due to low statistics. Additionally there is a small (x,y)-dependence of the S2 width which was accounted for in the original analysis by a corrected width-variable provided by the old processor. This allowed for stricter cuts in the respective energy intervals compared to the high 5σ acceptance in the newly defined cut. The new cut is defined as:

$$\begin{aligned} &\text{S2 width}_{90} \leq \sqrt{16.058 \cdot 10^{-9}\,\mathrm{s} \cdot t_d + 1.156077 \cdot 10^{-12}\,\mathrm{s}^2} \\ \wedge \quad &\text{S2 width}_{90} \geq \sqrt{8.609 \cdot 10^{-9}\,\mathrm{s} \cdot t_d + 0.492219 \cdot 10^{-12}\,\mathrm{s}^2}. \end{aligned} \tag{3.2}$$

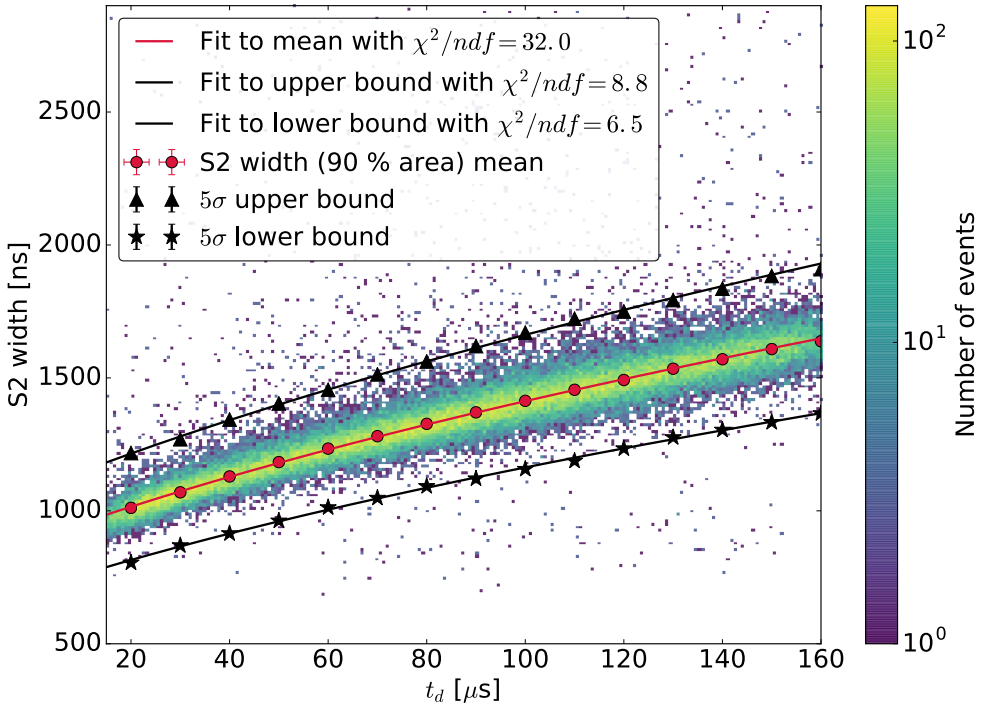

Figure 3.1: S2 width of ER events from ^{232}Th and ^{60}Co sources at 90 % of the peak area against t_d in the energy region of $cS1 < 150$ pe. Mean and 5σ intervals are determined by Gaussian fits of $\Delta t_d = 10\,\mu$s-slices of the S2 width band. The points for the slices are overlaid with square root functions to obtain cut functions for the upper and lower bound. The uncertainties of the points are too small to be visible in the plot.

A complementary S2 width consistency cut is defined on NR data in the same fashion with 3σ intervals. The definition can be found in section 6.3 of the appendix.

3.1.2 Sanity Cuts

Sanity checks request that an *event* contains at least one *interaction* with a reconstructed position inside the TPC. Such events are usually removed by other cuts such as the double scatter and fiducial volume cuts. However, sanity checks are useful as a pre-selection for saving memory because they only request a single variable, do not require much computing time, and sort out up to 60 % of invalid events within a processed file. The sanity checks are:

$$cS1 > 0\,\text{pe}, \tag{3.3}$$

$$cS2 > 0\,\text{pe}, \tag{3.4}$$

$$-z < 30.5\,\text{cm}. \tag{3.5}$$

3.1.3 Transitioning from ROOT to Python

Besides translating processor-specific variables in the cut definitions, the application of the cuts themselves has to be reworked. The XENON100 analysis relied heavily on the ROOT data analysis framework [43], but this is subject to change. While PAX provides ROOT output, the XENON1T analysis is centered around the Python-based PAX environment

which integrates ROOT support via modules. The transition manifests itself in the fact that most new analysis tools are already written in Python making heavy use of interactive, web-based *Jupyter* notebooks. Following the rationale of developing code applicable in XENON1T analyses all code for this thesis has been developed from scratch.
The application of cuts is illustrated in a short example with minimal code. Below, the sanity cut on $cS1$ is defined.

```
def Xs1custom0(data):
        cut_xs10 = data['cs1'] > 0
        return(cut_xs10)
```

Its input is a `data` object containing all relevant event properties of up to millions of events. The function returns arrays of `true` or `false` values which can be used to index the `data` object. The cut is performed by passing the data selection `Xs1custom0` and the `data` object to a cut function which creates a `new_data` object only containing the events that passed the cut.

```
def cut(data, selection):
        new_data = data[selection]
        return(new_data)
```

All events in `data` where the index returned by the selection is `true` are kept, while the others are discarded. Now the selection `Xs1custom0` is applied.

```
data = cut(data,Xs1custom0(data), 'cS1 > 0 selection')
```

In this case the old `data` object is overwritten with the new one after the cut. One can also create a new object and keep the original one.

3.2 Cut Acceptance

Counting the events in a sample of WIMP scatters before and after cuts would be a straightforward approach to determine the combined cut acceptance. Such an ideal sample could be created using a waveform simulator. However such data is neither available here, nor was it in the original analysis [34]. So, ER and NR calibration data had to stand in for real dark matter events. A clean event sample was prepared by applying all cuts except the one the acceptance was tested for[2]. In the end, the combined acceptance was calculated as the product of the individual ones. This follows the assumption that an event being removed by two independent cuts is more likely to be an invalid event than one removed by a single selection alone. The result is a conservative assessment which translates to an overestimation of the actual event rate given the measured number of events in the dark matter search region. As a consequence a weaker limit on the WIMP-nucleon interaction cross section can be set than with a higher acceptance and the same number of events.
In [34] a part of the cuts is tested on ER data[3] while the other is tested on NR data.

[2] In some cases correlated cuts were removed in addition to the tested cut. Correlations occur if two cuts work with the same variables.

[3] Electronic recoil calibration data is used to evaluate cuts addressing noise. As only 2.7 days of AmBe data are available they cannot account for noise conditions changing with time. ER calibration data is available over the course of the whole dark matter search campaign.

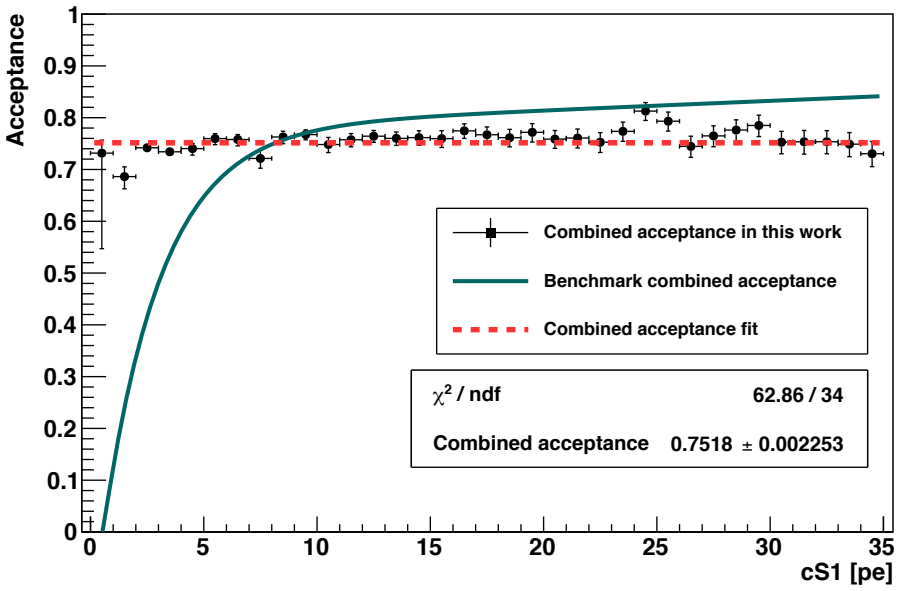

Figure 3.2: Combined cut acceptance in $cS1$ from 0 pe to 35 pe. A constant fit of the datapoints (dashed red line) and the combined cut acceptance from [34] (solid teal line) are overlaid over the datapoints.

The cuts tested on the different datasets are:

- NR data: S1 width, S2 seen by top PMTs, fit quality and consistency of position reconstruction, removal of γ-like events in the neutron band, gas event removal, and the S2 width consistency cut.
- ER data: single S1, S2 single scatter, and signal to noise cut.

The 34 kg fiducial volume was excluded from the acceptance calculation while the S2 trigger threshold was treated independently of the other cuts because it does not depend on S1 fluctuations. The acceptance of the veto cut was estimated in an extensive campaign of investigating waveforms of vetoed events and identifying wrongly vetoed valid events. In turn, the S1 coincidence criterion has been evaluated by selecting physical events as single coincidence events that also triggered the veto.
This work relies on a simplified approach. First, the fiducial volume, S2 trigger threshold and veto cuts[4] are applied to the ER and NR data before testing the other cuts assuming an acceptance of 100 %. Sanity cuts are applied in addition to further purify the sample. By virtue of the two different datasets all cuts that are tested on NR data can without reservations be applied on the ER data before evaluating the acceptance of the remaining cuts and vice versa. This should provide clean ER and NR datasets where the acceptances of the remaining cuts are derived by applying them on the respective datasets at once. Building the ratio of event numbers before (N_{before}) and after (N_{after}) cuts yields two acceptances: one from the cuts evaluated on NR data ($\mathcal{A}_{\text{NR}}$) and one from those evaluated on ER data ($\mathcal{A}_{\text{ER}}$). Both are calculated as a function of $cS1$ by sorting the cut and uncut

[4]The acceptance of the veto cut in the original work was assumed to be 99.4 % over the whole energy range.

data into $cS1$ histograms with 1 pe bins.

$$\mathcal{A}_{\mathrm{NR}}(cS1) = \frac{N_{\text{after, NR}}(cS1)}{N_{\text{before, NR}}(cS1)},$$
$$\mathcal{A}_{\mathrm{ER}}(cS1) = \frac{N_{\text{after, ER}}(cS1)}{N_{\text{before, ER}}(cS1)} \tag{3.6}$$

The combined acceptance $\mathcal{A}_{\mathrm{comb}}$ is then calculated as the product of both.

$$\mathcal{A}_{\mathrm{comb}} = \mathcal{A}_{\mathrm{NR}} \cdot \mathcal{A}_{\mathrm{ER}} \tag{3.7}$$

Computing the combined acceptance the Poissonian errors of the $cS1$ histograms have to be considered. As the acceptance cannot exceed unity, asymmetric errors must be used and preserved when dividing and multiplying histograms. This is taken care of by the `TGraphAsymmErrors` class in ROOT [43].
More events are expected to be removed by the tested cuts because less cuts have been applied beforehand. A lower combined acceptance should be the result. This can be observed in Fig. 3.2. The combined acceptance has been fitted with a constant function (dashed red line) yielding $\mathcal{A}_{\mathrm{comb}} = (75.2 \pm 0.3)\,\%$. Due to a few outliers one finds $\chi^2/ndf \approx 1.85$.
A divergence from the combined acceptance in [34] which is shown as the solid teal line is clearly visible. Above 10 pe the overall lower acceptance can be explained by lacking two non-translatable cuts in the pre-selection, namely the entropy and PMT pattern likelihood cuts. In addition, the custom definition of the S2 width consistency cut which has not been optimized for a high acceptance[5]. The steep rise below 10 pe featured in the original is introduced by the S1 coincidence criterion and reflects that the smaller an S1 signal is, the lower is its chance to be detected by more than one PMT. In this work the acceptance of the S1 coincidence cut is unity over the whole energy range. This is due to PAX's internal coincidence criterion which does not allow S1 peaks seen by less than two PMTs. Therefore, an assessment of the true acceptance for actual physical events with a singular coincidence is not possible.

3.3 Energy Calibration of the Detector

Information on the energy of nuclear recoil interactions is essential in the case of a dark matter discovery as it allows to calculate the WIMP mass. Moreover, the calibration provides insights on detector parameters such as, for example, the energy resolution $\Delta E/E$ or light yield L_y. It is of particular interest if these are altered by the data processor, so the calibration also serves the validation process. Two calibration methods are introduced in this work. An energy calibration for electronic recoils uses the combined information of S1 and S2. The nuclear recoil energy calibration is exclusively S1-based following the example of [34].

[5] Omitting this cut shifts the acceptance to a flat 80 %, closer to the values from the original work.

3.3.1 Calibration sources

The γ-lines of neutron-activated xenon and fluorine, as well as ^{137}Cs are utilized to determine the light and charge yields at the given energies by fitting their full absorption signals inside the detector. These are signals where an incident γ deposits its full energy in a single resolved interaction. Tab. 3.2 provides an overview of the relevant γ-lines.
Xenon and fluorine are activated by inelastic scattering of neutrons with the liquid xenon target and PTFE walls of the TPC. The neutrons originate in (α,n)-reactions of the ^{241}AmBe source. The capture of an α-particle from the decaying ^{241}Am frees a neutron from the ^{9}Be target:

$$\begin{aligned} {}^{241}_{95}\mathrm{Am} &\xrightarrow{432.2y} {}^{237}_{93}\mathrm{Np} + {}^{4}_{2}\,\alpha + \gamma\ (59.5\ \mathrm{keV}), \\ {}^{9}_{4}\mathrm{Be} + {}^{4}_{2}\,\alpha &\xrightarrow{(\alpha,n)} {}^{12}_{6}\mathrm{C} + n + \gamma. \end{aligned} \tag{3.8}$$

Except for the neutrons all decay products are contained in the source and cannot enter the TPC. Their inelastic scatters populate excited nuclear states of ^{19}F, ^{129}Xe and ^{131}Xe which deexcite via γ-decay [44]:

$$n + {}^{129}\mathrm{Xe} \rightarrow {}^{129}\mathrm{Xe}^{*} + n,$$

$$^{129}\mathrm{Xe}^{*} \rightarrow {}^{129}\mathrm{Xe} + \gamma. \tag{3.9}$$

$$n + {}^{131}\mathrm{Xe} \rightarrow {}^{131m}\mathrm{Xe} + n,$$

$$^{131m}\mathrm{Xe} \rightarrow {}^{131}\mathrm{Xe} + \gamma. \tag{3.10}$$

The lifetime of the low energy lines up to 109.9 keV is on the order of 1 ns. As the neutrons

Table 3.2: Overview of γ-lines used for the energy calibration of the XENON100 experiment sorted by their energies in ascending order [45].

Isotope	Energy	Half-life	Notes
^{129}Xe	39.6 keV	0.97 ns	Coincident neutron scatters
^{131}Xe	80.2 keV	0.48 ns	Coincident neutron scatters
^{19}F	109.9 keV	0.6 ns	Coincident neutron scatters
^{131m}Xe	163.9 keV	11.8 d	
^{19}F	197.1 keV	89 ns	
^{129m}Xe	236.2 keV	8.9 d	196.6 keV+39.6 keV (fast)
^{129}Xe	319.9 keV	$\sim$ ps	Weighted mean of neighbouring levels (318 keV, 321 keV)
^{137}Cs	661.7 keV	30.07 y	External

activating these lines need about 10 ns to travel through the whole TPC[6], they are likely to scatter again within the decay time of the excited states and the subsequent detector response to the emitted γ. Therefore, these lines can be altered in shape and energy due to coincident neutron scatters. This has to be taken into account in the fits[7].
The decay scheme of ^{137}Cs is shown in Fig. 3.3. The β-decay of ^{137}Cs preferably decays into a metastable state of ^{137}Ba which deexcites under γ-emission at the characteristic energy of 661.7 keV. Being an external source, energy loss via Compton scattering can occur before entering the sensitive volume of the TPC. This directly translates to features in the S1 and S2 spectra of this source.

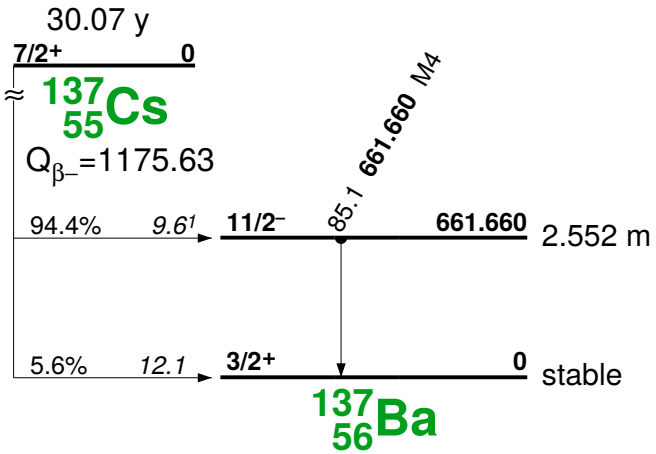

Figure 3.3: Decay scheme of ^{137}Cs [44].

3.3.2 Features of γ-lines in S1-S2-space

The energy calibration of the detector requires that the full energy of the γ is deposited inside the sensitive region of the detector and that it is read out as a single S1 and S2 signal. For incident photons within the energy range of 39.6 keV to 661.7 keV Compton scattering and photoelectric absorption are the principal mechanisms of energy loss in xenon [25, 46]. Fig. 3.4 shows the attenuation coefficients in xenon for 1 keV to 10 MeV γ-rays. At low energies up to around 300 keV photoelectric absorption is dominant while Compton scattering prevails above.
The interplay of multiple energy losses and detector resolution causes characteristic features in the γ-spectra of the calibration sources. These can be best observed in S1-S2-space like in Fig. 3.5. The plot shows $cS1$ vs. $cS2_b/100$ for data from an external ^{137}Cs source. From now on, the y-axis in such plots is normalized by 100 in order to have S1 and S2 in the same order of magnitude. Only the S2 contribution seen by the bottom PMT array $cS2_b$ is treated to avoid saturation effects of the top array PMTs. The color scale of the

[6] The derivation can be found in the appendix in section 6.4.
[7] There are other inelastic reactions which are possible given the typical incident neutron energies of a few MeV that an ^{241}AmBe source [45] provides. However, such interactions have not been observed in the detector next to backgrounds and the other lines.

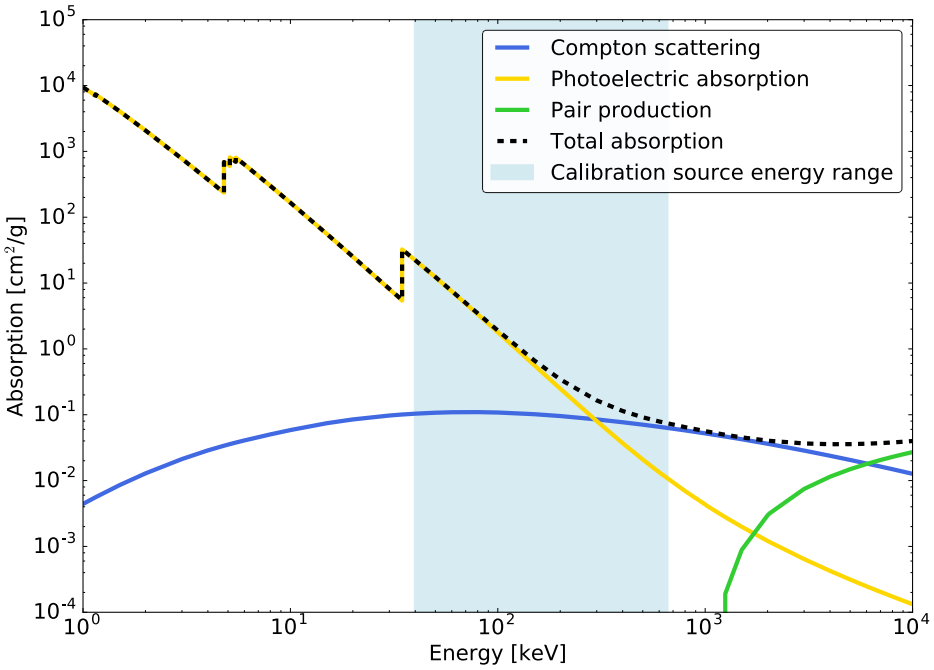

Figure 3.4: Attenuation coefficients for γ-rays in xenon as a function of energy for photoelectric absorption, Compton scattering and pair production. Nuclear and electron field contributions to pair production have been summed to obtain a total attenuation coefficient. All coefficients are normalized by density, so the unit is cm^2/g. Values taken from [46].

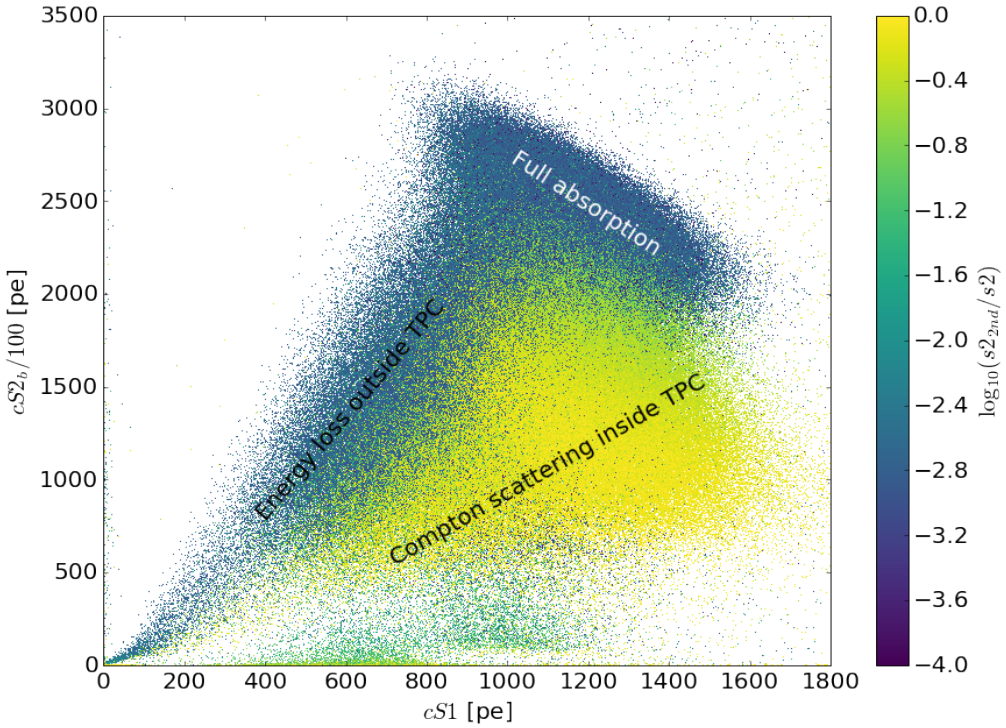

Figure 3.5: Scatter plot of $cS2_b/100$ against $cS1$ for an external ^{137}Cs source before the application of a double scatter cut. The color scale is normalized by $\log_{10}\left(\frac{s2_{2\mathrm{nd}}}{s2}\right)$ with the range restricted to $[-4,0]$.

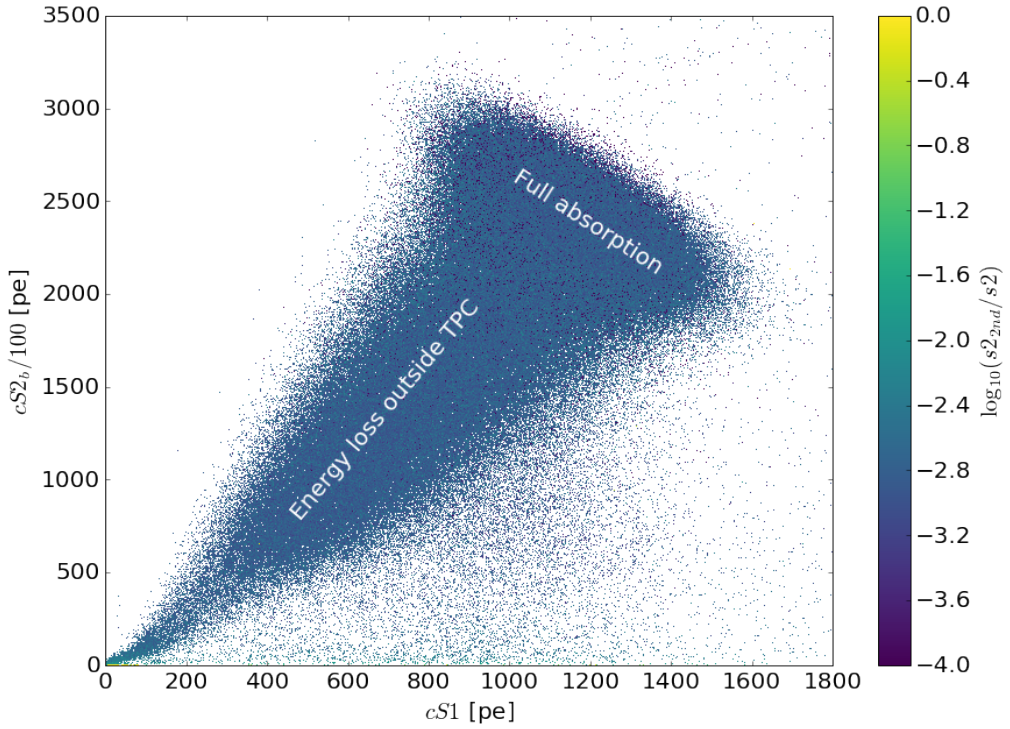

Figure 3.6: Scatter plot of $cS2_b/100$ against $cS1$ for an external ^{137}Cs source after the application of a double scatter cut. The color scale is normalized by $\log_{10}\left(\frac{s2_{2\text{nd}}}{s2}\right)$ with the range restricted to $[-4,0]$.

plot is normalized by

$$\text{color} \sim \log_{10}\left(\frac{s2_{2\text{nd}}}{s2}\right). \tag{3.11}$$

Here, $s2$ and $s2_{2\text{nd}}$ are the areas of the main and second largest S2 in the event, so the yellow points have a larger secondary S2 than the blue ones.

The anti-correlation of both signal channels is clearly visible in the plot as the tilt of the blue and yellow event populations. The full energy peak of a given γ-line is a Gaussian peak if observed in $cS1$ as well as in $cS2_b$ only. In the combined space it represents a tilted ellipse, so more light means less ionization charge and vice versa.

At $E_\gamma = 661.7\,\text{keV}_\text{ee}$ Compton scattering is expected to be the predominant mode of energy loss in liquid xenon. The blue population consists of events where individual Compton scatters of photons in the TPC could not be resolved. Thus, both S1 and S2 include the full energy of the event. If no energy loss occured outside the sensitive volume of the TPC, events contribute to the full absorption line. The diagonal so-called *Compton continuum* is populated by events with external energy loss where the γ has either left the sensitive volume before full absorption or scattered before entering the TPC. The second case is expected for external sources where a γ has to traverse the cryostat, liquid xenon veto and copper-ringed PTFE shell of the TPC first.

The yellow event population consists of Compton-scatter events inside the TPC. Here the S2 does not represent the full energy of the incident photon while the S1 does. The S2s from multiple scatters can be resolved if the individual S2s are not clustered into a single peak. In such events the main S2 lacks energy from multiple interactions. This is different

for the S1s. As the prompt scintillation photons travel with the speed of light, individual Compton scatters cannot be resolved in light and a summed S1 is measured. Therefore, the S1 spectrum exhibits a Compton-continuum with a full energy peak just like the blue population. The yellow population can be removed by the S2 single scatter selection. The spectrum of the ^{137}Cs source after the removal of double scatters is shown in Fig. 3.6.

3.3.3 Energy Calibration from Full Absorption Ellipses

The tilted ellipses of the full absorption lines in S1-S2-space are mathematically described by 2D Gaussian functions in a rotated coordinate system [47, 48]:

$$f(x) = A \cdot e^{\frac{-((x-\mu_x)\cdot\cos\theta-(y-\mu_y)\cdot\sin\theta)^2}{2a^2} - \frac{((x-\mu_x)\cdot\sin\theta+(y-\mu_y)\cdot\cos\theta)^2}{2b^2}} \tag{3.12}$$

A is the amplitude at the center, $\mu_{x,y}$ are the respective means, and the scaling parameters a and b of the main axes incorporate the rotated standard deviations $\sigma_{x,y}$ of a standard 2D Gaussian. After substitutions $x = cS1$ and $y = cS2_b/100$ one obtains the fit function. The meaning of the parameters is illustrated in Fig. 3.7. The counterclockwise rotation of the coordinate system by the angle θ is described by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cdot\cos\theta - y\cdot\sin\theta \\ x\cdot\sin\theta + y\cdot\cos\theta \end{pmatrix}. \tag{3.13}$$

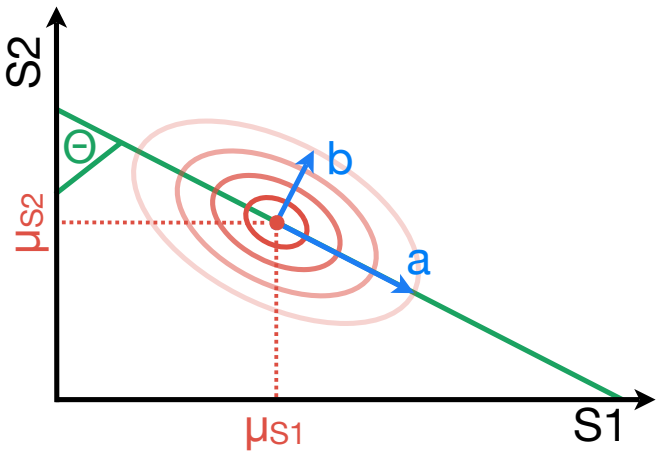

Figure 3.7: Illustration of the elliptical fit function for the fits of the γ-lines in S1-S2-space. The amplitude A is indicated by the fading contour lines.

The light yield L_y and charge yield Q_{ee} of electronic recoils[8] at the respective energies and the applied drift field of 530 V/cm can be directly calculated from the means $\mu_{\text{S1,S2}}$ by simple divisions:

$$\begin{aligned} L_y(E, \mathcal{E} = 530\,\text{V/cm}) &= \frac{\mu_{\text{S1}}}{E}, \\ Q_{\text{ee}}(E, \mathcal{E} = 530\,\text{V/cm}) &= \frac{\mu_{\text{S2}}}{E}. \end{aligned} \tag{3.14}$$

[8] These are not equal to the light and charge yields for nuclear recoils. The conversion of L_y by $\mathcal{L}_{\text{eff}}$ is outlined in section 1.4.2 and will be performed in section 3.3.9.

These as well as the rotation angle θ can be used for an energy calibration [47, 48]. The combined light and charge signals contain the full energy information of a particle interaction regardless of recombination (c.f. Fig. 1.10). Thus, one can write

$$E = \alpha(E) \cdot cS1 + \beta(E) \cdot \frac{cS2_b}{100} \tag{3.15}$$

with energy-dependent scaling parameters $\alpha(E)$ and $\beta(E)$. These are related to the effective energies W_γ and W_e which are required to generate information carriers in liquid xenon and which have been introduced in Eq. 1.14. By assuming a mean effective energy

$$W_\gamma = W_e = W = 13.7\,\mathrm{eV} \tag{3.16}$$

and introducing the photon detection efficiency g_1 as well as the charge amplification factor g_2 one obtains

$$E = W \cdot \left(\frac{cS1}{g_1} + \frac{cS2_b}{g_2} \right) \tag{3.17}$$

The detector-specific variable g_1 is an effective value related to the photon detection efficiency $\mu(\vec{r})$ introduced in Eq. 1.16. It still incorporates the optical properties of the TPC and the quantum efficiency of the PMTs, but the position dependence has been removed by $cS1$. It is assumed to be constant over the energy range of interest. If no saturation sets in, g_2, which bears resemblance to $\delta(x,y)$ from Eq. 1.24, can also be assumed as constant. This approach with a mean effective W-value resembles the one taken by the Monte Carlo simulation package NEST [31] which provides the raw numbers of photons and electrons created in liquid xenon by an incident particle[9]. In order to compare the raw numbers with measured S1 and S2 signals they have to be converted by using g_1 and g_2:

$$\begin{aligned} cS1 &= n_\gamma \cdot g_1, \\ cS2_b &= n_e \cdot g_2. \end{aligned} \tag{3.18}$$

This work requires the knowledge of at least g_1 for interpolating the measured light yields with NEST to ascertain $L_y(122\,\mathrm{keV_{ee}})$ for the S1-based energy calibration. Comparing Eq. 3.15 and Eq. 3.17 and neglecting the energy dependence of α and β gives

$$\begin{aligned} g_1 &= \frac{W}{\alpha}, \\ g_2 &= \frac{100 \cdot W}{\beta}. \end{aligned} \tag{3.19}$$

Now α and β can be determined by fitting the measured light and charge yields at all energies with a linear function [49]

$$Q_{\mathrm{ee}} = -\alpha \cdot L_y + \frac{1}{\beta} \tag{3.20}$$

from rearranging Eq. 3.15 and inserting Eq. 3.14. At this point the information from the

[9] NEST is introduced in more detail in section 3.3.9.

rotation angle has not been used. One can derive geometrically that

$$\tan\theta = \frac{cS2_b/100}{cS1}. \tag{3.21}$$

This allows to construct an energy calibration from a single line if α and β are assumed constant[10]. As the main axis of each ellipse constitutes a line of constant energy E, one can generalize Eq. 3.21 for all points on it:

$$\mu_{\mathrm{S2}} + m \cdot \mu_{\mathrm{S1}} = cS2_b/100 + m \cdot cS1 = k \cdot E \tag{3.22}$$

The scaling constants m and k include $\theta(E)$ with

$$\begin{aligned} m &= \frac{cS2_b/100 - \mu_{\mathrm{S2}}}{cS1 - \mu_{\mathrm{S1}}} = \tan\theta(E), \\ k &= \frac{m \cdot \mu_{\mathrm{S1}} + \mu_{\mathrm{S2}}}{E}, \end{aligned} \tag{3.23}$$

where (μ_{S1},μ_{S2}) is the center point of the ellipse. It will be investigated in section 3.3.8 if this parametrization improves the accuracy and resolution of the energy calibration. The accuracy is from now on defined as

$$\Delta E = E_{\mathrm{true}} - E_{\mathrm{calibrated}} \tag{3.24}$$

with the actual energy E_{true} of a given γ-line and the energy $E_{\mathrm{calibrated}}$ from the calibration. The resolution is calculated as

$$\frac{\sigma_E}{E_{\mathrm{calibrated}}} \tag{3.25}$$

The standard deviation σ_E and mean $E_{\mathrm{calibrated}}$ are extracted from Gaussian fits of to the 1D energy histograms after the conversion from $cS1$ and $cS2_b$ to E.

3.3.4 Combined Energy Scale

A combined energy scale (CES) of S1 and S2 should provide a better energy resolution than one from a single signal channel because recombination fluctuations in S1 are accounted for in S2. However, it has not been used in the benchmark dark matter search [34] because the charge yield Q_y for nuclear recoils was not known to sufficiently low energies at that time. Thus, it will not be applied on the dark matter data in this work either.
Still, there is a case for the definition of a CES. Chiefly, it provides $g1$ which is needed for matching the measured light yields of the individual lines with NEST. This lays the groundwork for constructing the S1-based energy scale applied in the dark matter search. Moreover, the use of both light and charge for energy reconstruction has become state of the art, as Q_y has now been measured to similarly low energies as $\mathcal{L}_{\mathrm{eff}}$ [50]. In turn, the calibration method applied in this work can be directly used for XENON1T after slight modifications for the calibration lines in the new detector.
Three methods for the definition of a CES are introduced, evaluated and compared in terms

[10] Yet, $\theta(E)$ is treated as an independent parameter in the fit function and can be used to parametrize the energy-dependence of the energy scaling parameters.

of accuracy and resolution: a CES from a single γ-line, in this case the one from ^{137}Cs, a global linear fit of the light and charge yields for all lines, and an angular parametrization that accounts for the energy dependence of the scaling parameters.
In the following, all fits to data are performed with ROOT using the MINUIT minimizer in χ^2-minimization mode. For the low statistics fits with bin contents of less than 100 events a likelihood maximization is performed. The errors for all fit parameters are calculated in a second fitting procedure with the MINOS minimizer which takes the correlation of fit parameters into account when calculating their 1σ uncertainties. The bin errors in all histograms are $\sqrt{N}$ Poisson errors. Unless specified otherwise, a single rotated 2D Gaussian is fitted to the γ-lines in $cS1$ versus $cS2_b/100$.

3.3.5 Single Line CES from ^{137}Cs

A single line CES can in principle be constructed from any γ-line by rearranging Eq. 3.22 to

$$E = \frac{1}{k} \cdot \left(\frac{cS2_b}{100} + m \cdot cS1 \right). \tag{3.26}$$

and using μ_{S1}, μ_{S2} as well as θ to determine k and m. Preferably one should choose a clean line with sufficient statistics, little contamination from other lines or coincident neutron scatters, and a low background. Due to the fact that ^{137}Cs calibrations were carried out weekly plentiful statistics are available with little background from other lines. Fig. 3.8 shows the event distribution from the external ^{137}Cs source at three different (x,y)-positions around the TPC. The fiducial volume spanning over 80 % of the TPC radius and the central 70 % of the TPC length is indicated by the dashed red lines. One can see that most events are close to the walls owed to the self-shielding of liquid xenon: Both the calibration source and external backgrounds are attenuated. The vertical streaks in the left plot are a result of the position reconstruction in PAX.
The reconstruction of the full absorption line is not trivial due to the remaining Compton spectrum from external energy loss that is still present after the basic double scatter cut.

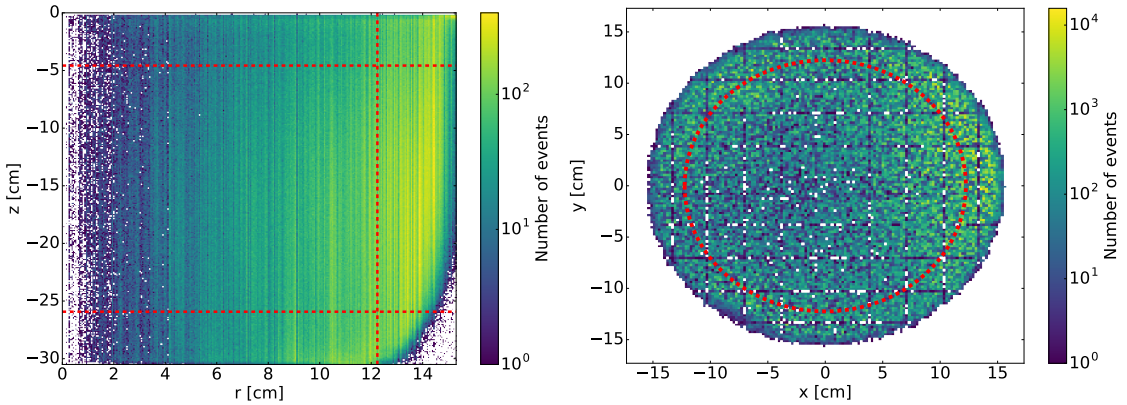

Figure 3.8: **Left:** Event distribution of ^{137}Cs in (r,z) before cuts. **Right:** Event distribution in (x,y) of the same dataset. The bin width in both plots is 0.5 mm. The dashed red lines mark the fiducial volume. The source positions are slightly visible in the right plot at 3, 7 and 11 o'clock.

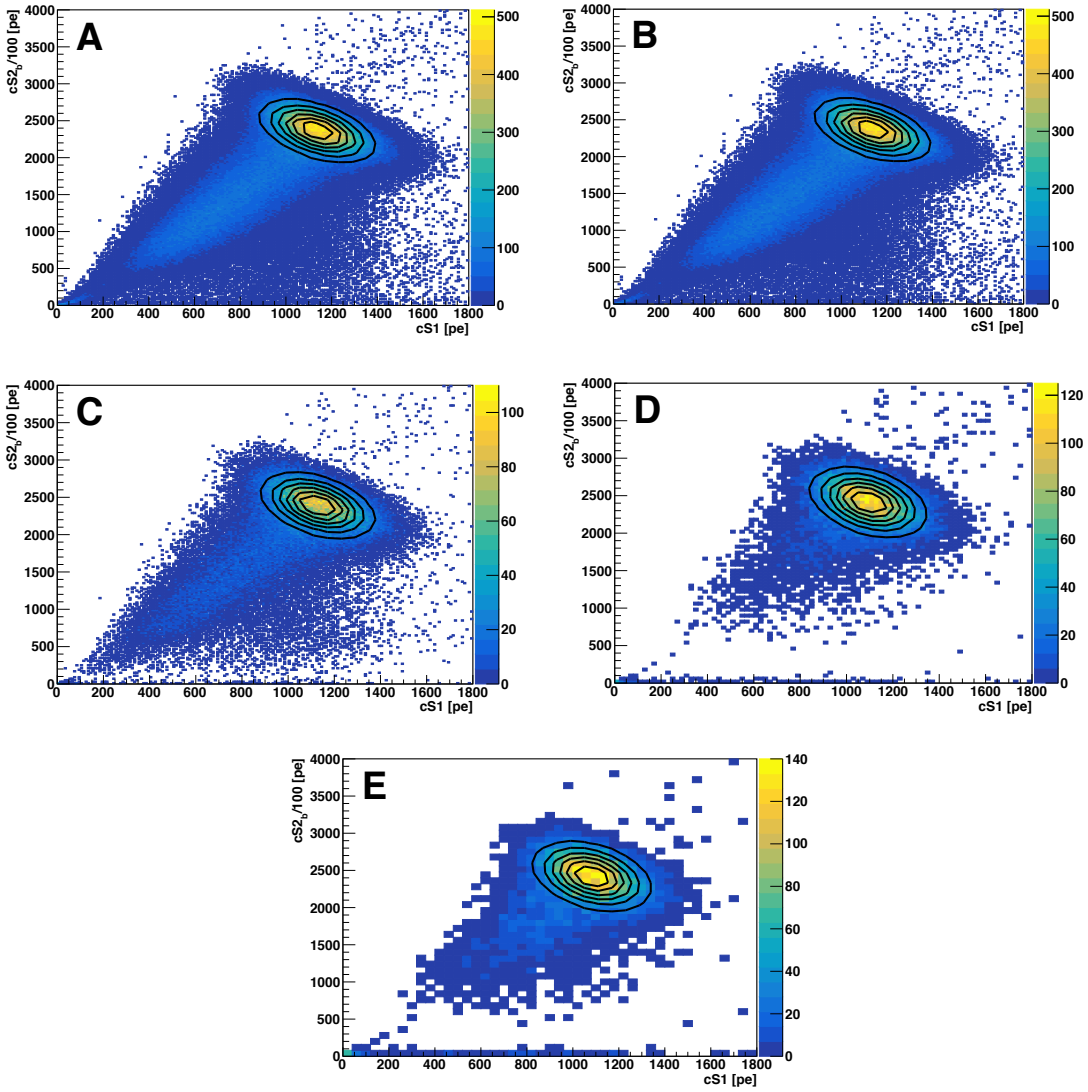

Figure 3.9: Elliptical fits of the ^{137}Cs data in S1-S2-space with increasingly strict double scatter cuts.

Table 3.3: Fit parameters for the unconstrained ^{137}Cs fits on the same dataset with different double scatter cuts. Uncertainties are taken from MINOS errors.

	Selection	μ_{S1} [pe]	μ_{S2} [pe]	θ [rad]	χ^2/ndf
A	Legacy cut	1136 ± 1	2361 ± 1	1.21 ± 0.01	14.55
B	$\log_{10}\left(\frac{s2_{2nd}}{s2}\right) < -2$	1136 ± 1	2361 ± 1	1.21 ± 0.01	14.56
C	$\log_{10}\left(\frac{s2_{2nd}}{s2}\right) < -3$	1128 ± 1	2390 ± 2	1.27 ± 0.01	1.90
D	$\log_{10}\left(\frac{s2_{2nd}}{s2}\right) < -4$	1094 ± 1	2415 ± 2	1.32 ± 0.01	1.31
E	$\log_{10}\left(\frac{s2_{2nd}}{s2}\right) < -5$	1083 ± 2	2422 ± 4	1.32 ± 0.01	2.21

This can be solved by stricter double scatter cuts at the cost of statistics or by constraining fit ranges and parameters. As the second approach introduces bias, stricter cuts are tried first.

Fig. 3.9 shows the elliptical fits to the event distributions in S1-S2-space after increasingly

strict double scatter cuts. The fits are neither constrained by fit ranges nor by parameter limits, except for θ. Only positive angles are allowed to fix the rotation direction. A constant for modeling background has been added to the fit function. Initial guesses for the fit parameters are taken from a preliminary and strongly constrained fit around the ellipse center. The binning is adapted to the decline in statistics for stricter cuts. Starting with bin sizes of 10 pe in $cS1$ and 20 pe in $cS2_b/100$ in the first plot and an initial number of 637047 events, the binning is 40 pe in $cS1$ and 80 pe in $cS2_b/100$ in the last plot with 10282 remaining events. The S2 double scatter cuts compare $s2_{\mathrm{2nd}}$, the area of the second largest S2, to the uncorrected area $s2$ of the main interaction requiring

$$\log_{10}\left(\frac{s2_{\mathrm{2nd}}}{s2}\right) < -2\ (,\ -3,\ -4,\ -5). \tag{3.27}$$

The definition of the legacy cut for the basic single scatter selection can be found in section 6.2 in the appendix. After the maximum requirement no events with a secondary S2 are left in the data. This means that the particles in these events lost their full energy in a single resolved interaction. The effect of the reduced Compton background can be observed in the evolution of the fit parameters and quality.

Table 3.3 gives an overview of the fit parameters and the reduced χ^2 of the fits for the different double scatter cuts. The overall trend is that μ_{S1} shifts left to lower values while μ_{S2} moves to higher values. This also causes an increased rotation. The reduced χ^2 improves from selections A to D due to the removal of events with energy losses outside the TPC. These events moved and tilted the ellipses to the lower left of the frame. For less events outside the full energy ellipse the fit function describes the data better. However, this is not true for selection E where the reduced χ^2 and also the fitted background are higher than before. This can be explained by the removal of actual full energy events to a larger amount than extra-TPC Compton scatters. In the removed events a secondary S2 could be present due to single electrons from photoionization of impurities inside the TPC. As this does not impact the full energy S2 such an event constitutes a valid full absorption event. Still, it is removed by the extreme single scatter requirement. Therefore, selection D is chosen for constructing the single line CES. The fit parameters yield:

$$\begin{aligned} L_y &= (1.65 \pm 0.01)\,\frac{\mathrm{pe}}{\mathrm{keV}}, \\ Q_{\mathrm{ee}} &= (364.97 \pm 0.31)\,\frac{\mathrm{pe}}{\mathrm{keV}}, \\ m &= 3.91 \pm 0.17, \\ k &= (10.11 \pm 0.29)\,\frac{\mathrm{pe}}{\mathrm{keV}}. \end{aligned} \tag{3.28}$$

The CES is applied to the remaining data after the selection. The full energy peak is fitted with a Gaussian function in order to determine the resolution. A second Gaussian function is included to model the background from Compton scatters. This is shown in Fig. 3.10. Modeling the Compton background as a Gaussian is not strictly the right mathematical model, so one gets $\chi^2 = 1.31$ in the selected fit range from 400 keV to 800 keV[11]. A constant

[11] The expected energy of the Compton edge has been estimated to 476.9 keV from the maximum energy transfer.

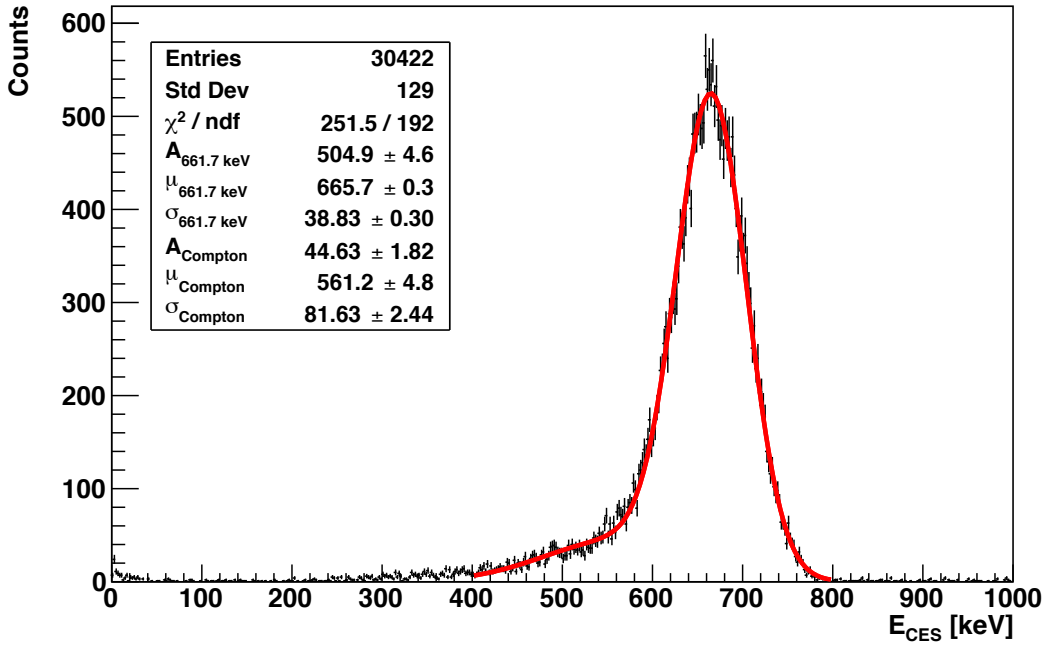

Figure 3.10: Energy spectrum of the ^{137}Cs data with with $\log_{10}\left(\frac{s2_{2\mathrm{nd}}}{s2}\right) < -4$. The CES uses the parameters from the unconstrained elliptical fit.

background has been added in early fit attempts, but assumed negative values due to the absence of background on the right side of the peak. The fit parameters were not constrained. The mean and standard deviation of the full energy peak are

$$\begin{aligned} \mu_{661.7\,\mathrm{keV}} &= (665.7 \pm 0.3)\,\mathrm{keV}, \\ \sigma_{661.7\,\mathrm{keV}} &= (38.83 \pm 0.30)\,\mathrm{keV}. \end{aligned} \tag{3.29}$$

Therefore, the mean energy does not agree within the error, but lies within the standard deviation of the peak. The energy resolution is

$$\frac{(38.83 \pm 0.30)\,\mathrm{keV}}{(665.7 \pm 0.3)\,\mathrm{keV}} = (5.83 \pm 0.05)\,\% \tag{3.30}$$

which is inferior to the resolution of 2.3 % at 661.7 keV quoted in the original analysis [17]. This is an indication that the fit of the full energy ellipse can be improved further. For instance, the rotation angle of the ellipse and the center point of the ellipse do not obey the expected relation

$$\tan(1.32) = 3.90 \neq \frac{2415\,\mathrm{pe}}{1094\,\mathrm{pe}} = 2.207. \tag{3.31}$$

In an attempt to improve the fit θ is constrained by Eq. 3.21 and fit ranges (950-1250) pe in $cS1$ and (2250-2750) pe in $cS2_b/100$ are added. In addition, a smaller 34 kg ellipsoidal fiducial volume is chosen. Now that Compton background can be suppressed by the selection of the fit range the dataset with selection C is chosen because it provides more statistics. The constrained elliptical fit is displayed in Fig. 3.11. $\chi^2/ndf \approx 2.95$ is higher than previously. This does not necessarily mean that the previous fits described

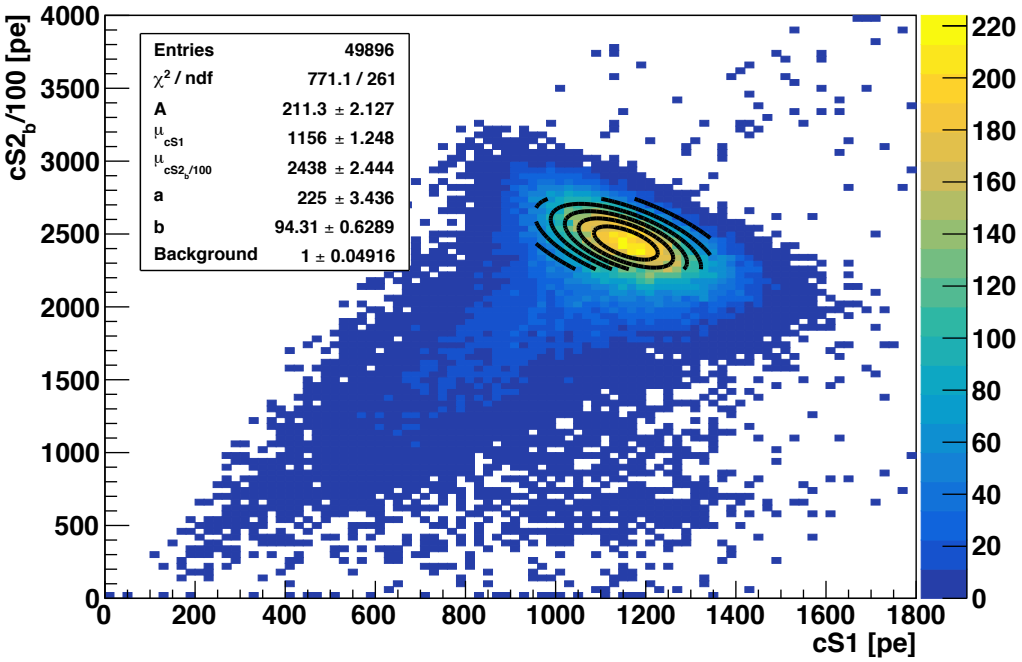

Figure 3.11: Constrained elliptical fit of the ^{137}Cs full absorption line in the 34 kg fiducial volume with $\log_{10}\left(\frac{s2_{2\mathrm{nd}}}{s2}\right) < -3$.

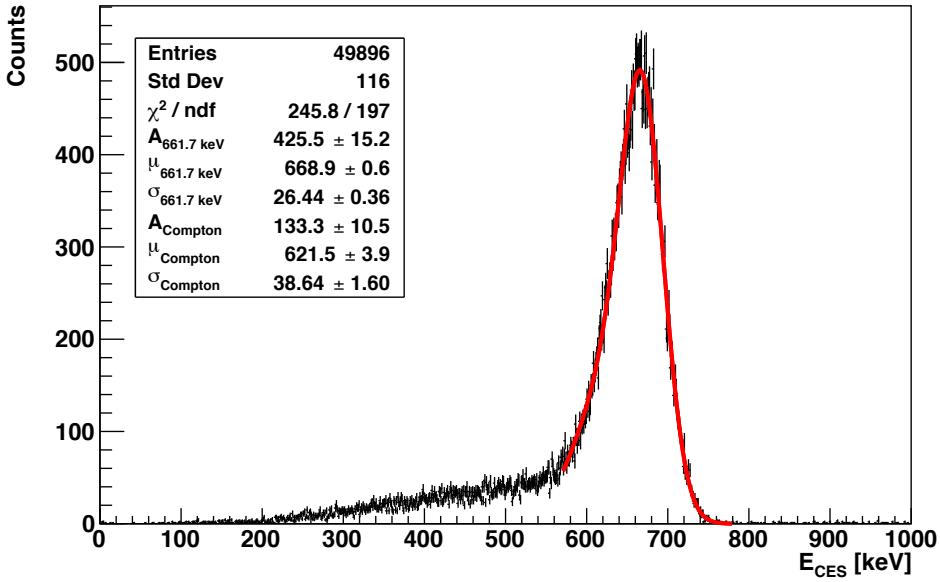

Figure 3.12: Energy spectrum of the ^{137}Cs data calculated with the CES parameters from the constrained elliptical fit.

the data better, but that bins with low statistics could have artificially suppressed χ^2. The ellipse itself is narrower on the small main axis with $a = (94.31 \pm 0.63)$ pe. While $\mu_{S2} = (2438 \pm 3)$ pe has been shifted further up again, $\mu_{S1} = (1156 \pm 2)$ pe is higher than

in all previous fits. The fit parameters yield

$$\begin{aligned} L_y &= (1.74 \pm 0.01)\,\frac{\text{pe}}{\text{keV}}, \\ Q_{\text{ee}} &= (368.44 \pm 0.45)\,\frac{\text{pe}}{\text{keV}}, \\ m &= 2.11 \pm 0.01, \\ k &= (7.37 \pm 0.02)\,\frac{\text{pe}}{\text{keV}}. \end{aligned} \tag{3.32}$$

The CES fit displayed in Fig. 3.12 shows a lower $\chi^2/ndf \approx 1.27$ than before, but the fit range had to be decreased to an interval of (570-780) keV due to the more prominent Compton background that cannot be described by a single Gaussian over the whole range. Decreasing the fit range further kept the fit result stationary within the parameter uncertainties. An improvement of the energy resolution is apparent with

$$\frac{(26.44 \pm 0.36)\,\text{keV}}{(668.9 \pm 0.6)\,\text{keV}} = (3.95 \pm 0.60)\,\%. \tag{3.33}$$

The mean energy is shifted further away from the true energy, but agrees within the resolution. Even though further optimization could be viable, it is not pursued to give way to a CES from several γ-lines over a broad energy range.

3.3.6 Global CES from Xenon, Fluorine and Caesium Ellipse Means

Accompanying the single point from the ^{137}Cs line, the neutron activation lines from the ^{241}AmBe calibration are fitted. Fig. 3.13. shows the entire $cS1$-$cS2_b$-spectrum with the source still present. With the exception of the fiducial volume the data selection is based on all sanity and a selection of legacy cuts. The applied legacy cuts are: S1 coincidence, S1 width, veto, S2 trigger threshold, S2 single scatter, S2 seen by top PMTs, avoid early S2s and consistency of position reconstruction. Owed to the lower energies no pronounced Compton spectrum can be discerned.

All events are situated within 95 % of the TPC radius and the central 80 % of the z-coordinate. The large radius of 14.54 cm has been chosen because the ^{19}F lines originate in the TPC's PTFE shell. The lower energy 109.9 keV γ has a mean free path of less than 3 mm in liquid xenon [51], which is only half of the radius that has already been cut. Due to the low lifetime and high activity the line is visible nonetheless. A larger fiducial volume is not viable, as data close to the walls suffers from high background and incomplete charge collection. The data is displayed on logarithmic scales to account for the large range of energies and the different interaction types it encompasses.

Three major populations can be identified. The lower left is populated by elastic scatters of Ambe neutrons. These are later used to define the nuclear recoil band, but, apart from them activating the detector walls and target, they are not relevant for the energy calibration.

The small event population at the upper right are α-lines from decays of intrinsic ^{222}Rn and the daughter isotope ^{218}Po [48]. Signals from such decays are marked by an extremely high S1 yield paired with a comparatively small S2 signal. Both signals can cause saturation and the correction maps for $cS1$ and $cS2$ are not valid.

The relevant cohort of events for this analysis is at the center of the figure. These are the γ-lines of activated xenon and fluorine. These are fitted with the 2D Gaussian functions, so the modalities outlined in section 3.3.4 also apply here. In some cases the fit approach has to be modified due to an overlap of lines or due to coincident neutron scatters. Depending

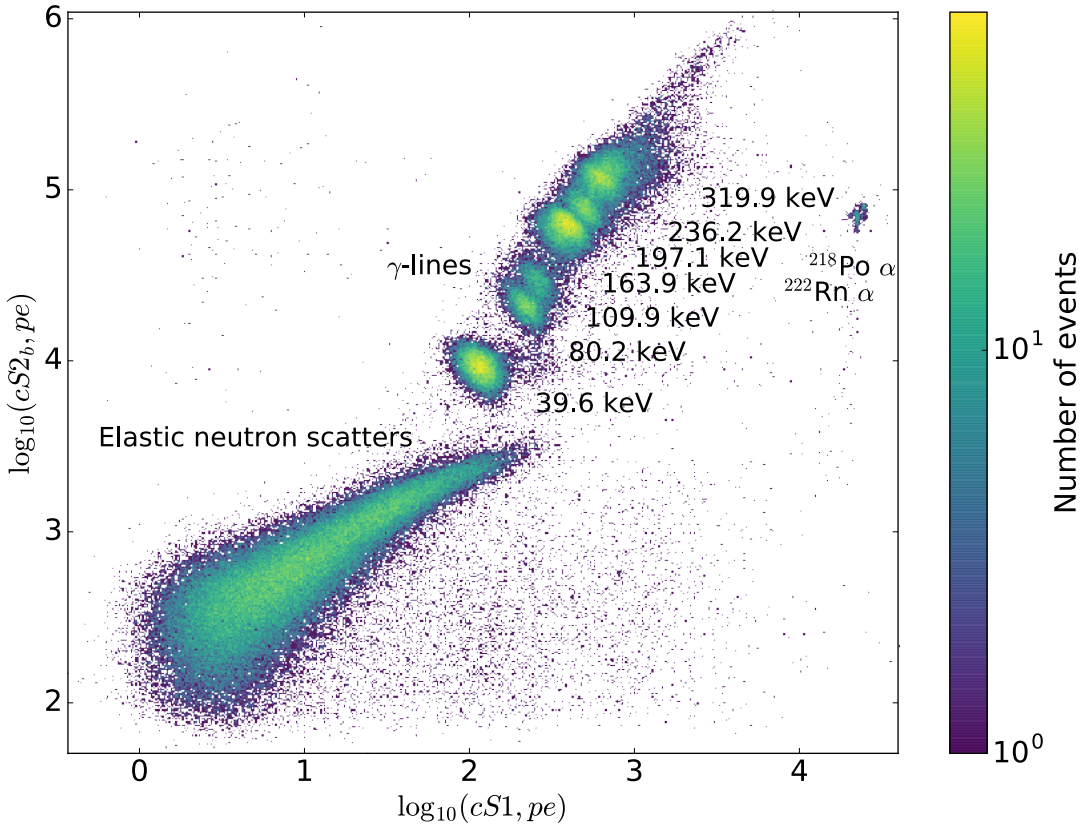

Figure 3.13: Spectrum in $cS1$ and $cS2_b$ on logarithmic scales during ^{241}AmBe calibration (2.7 days livetime) after cuts. The cylindrical fiducial volume spans 95 % of the r-coordinate and the central 80 % of the z-coordinate.

Table 3.4: Fiducial volumes and fit methods for the γ-lines of activated Xe and F.

E [keV]	Fiducial Volume	Fit Method
39.6, 80.2	67 % TPC radius, central 80 % TPC length	Ellipse with constrained fit range, additional 1D exponentially modified Gaussians in $cS1$ and $cS2_b/100$ modeling coincident neutron scatters.
109.9	95 % TPC radius, central 80 % TPC length	Combined fit of two ellipses together with 80.2 keV line.
163.9, 197.1, 236.2	67 % TPC radius + 85 % to 95 % TPC radius + 95 % TPC radius, central 80 % TPC length, and 34 kg fiducial volume for background data	Combined and constrained elliptical fit of the 163.9 keV, 197.1 keV and 236.2 keV lines from AmBe (three different fiducial volumes) and background data. The combined χ^2 for five fit functions is minimized.

Table 3.5: Fit ranges and parameter constraints for all γ-lines.

E [keV]	$cS1$ range [pe]	$cS2_b/100$ range [pe]	Constraints
39.6 (2D)	$70-160$	$60-130$	$0<\theta_{39.6\text{ keV}}<0.8$
39.6 (1D)	-	$50-160$	-
80.2 (2D)	$150-280$	$150-280$	$0.5<\theta_{80.2\text{ keV}}<1.0$
80.2 (1D)	$140-350$	$140-320$	-
109.9	$130-350$	$130-380$	$0.6<\theta_{80.2\text{ keV}}<0.9$, $0.6<\theta_{109.9\text{ keV}}<1.5$
163.9	$220-480$	$400-630$	by combined fit
197.1	$250-620$	$400-980$	by combined fit
236.2	$380-640$	$600-980$	by combined fit

on the line different fiducial volumes, datasets and fit methods are chosen which are outlined in Tab. 3.4. Parameter constraints and fit ranges are summarized in Tab. 3.5. The assessment of systematic errors is outlined in the specific sections. Using the 319.9 keV line was not possible due to low statistics.

39.6 keV and 80.2 keV

The elliptical fits of the 39.6 keV and 80.2 keV lines are shown in Fig. 3.14. In both $\chi^2/ndf \approx 1$ suggests a good description of the data by the model, but this is artificially lowered by the large $\sqrt{N}$-errors from bins with low event counts[12]. Larger bins in both fits do not alter the parameter values, but an increase of $\chi^2/ndf \approx 2$ is visible. The current binning has been chosen because the shapes of the event distributions are more visible this way.

An apparent offset between the ellipse center and the event distribution's center of gravity is visible for the 39.6 keV line. This is due to the fact that the line does not represent a perfect ellipse due to coincident neutron scatters. These, similarly to Compton scatters, cannot be separated in $cS1$ which increases the peak area. Strictly, the Gaussian $cS1$ distribution has to be folded with the coincident neutrons' $cS1$ distribution which in turn depends on the energy spectrum. An empirical exponentially modified Gaussian function assumes that the NR energy spectrum follows an exponential decay while the γ component is treated as a Gaussian:

$$\begin{aligned} f(x) &= \int_{\mu}^{\infty} \tfrac{1}{\tau} e^{-\frac{t-\mu}{\tau}} \cdot \tfrac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-t)^2}{2\sigma^2}} dt \\ &= \tfrac{e^{\frac{\sigma^2}{2\tau^2}}}{2\tau} \cdot e^{-\frac{x-\mu}{\tau}} \cdot \operatorname{erfc}\left(\tfrac{\mu - x + \frac{\sigma^2}{\tau}}{\sqrt{2}\sigma}\right). \end{aligned} \tag{3.34}$$

In addition to the Gaussian mean μ and standard deviation σ a third free parameter, the exponential constant τ of the NR spectrum is introduced. The model is also applied

[12] In later data-taking campaigns the amount of recorded AmBe data was doubled in order to improve statistics. However this data cannot be used in this work due to different detector conditions.

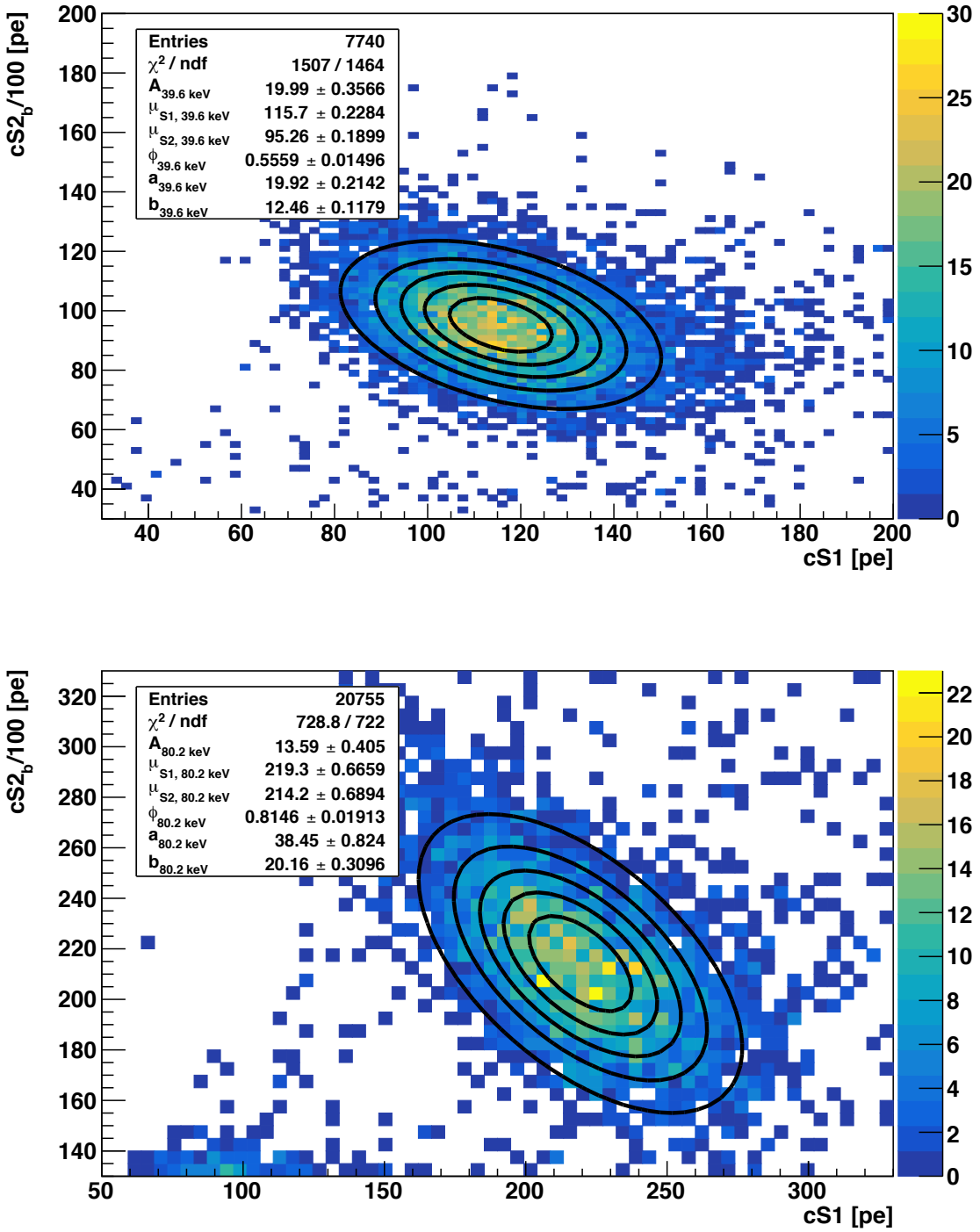

Figure 3.14: Elliptical fits of the 39.6 keV (**top**) and 80.2 keV (**bottom**) lines in $cS1$ and $cS2_b/100$. The cylindrical fiducial volume spans 67 % of the r-coordinate and the central 80 % of the z-coordinate.

to to the $cS2_b/100$-distribution, as events where the coincident neutron scatters can be discriminated in S2 should have been removed by the single scatter selection while those with merged S2s are still present.

The results for the two lines are shown in figures 3.15 and 3.16. Values for the reduced χ^2 do not increase significantly when the binning is reduced which is an indication for a better description of the data. It is evident that both means are shifted to lower values by this method. This implies that L_y and Q_{ee} would have been overestimated with the 2D

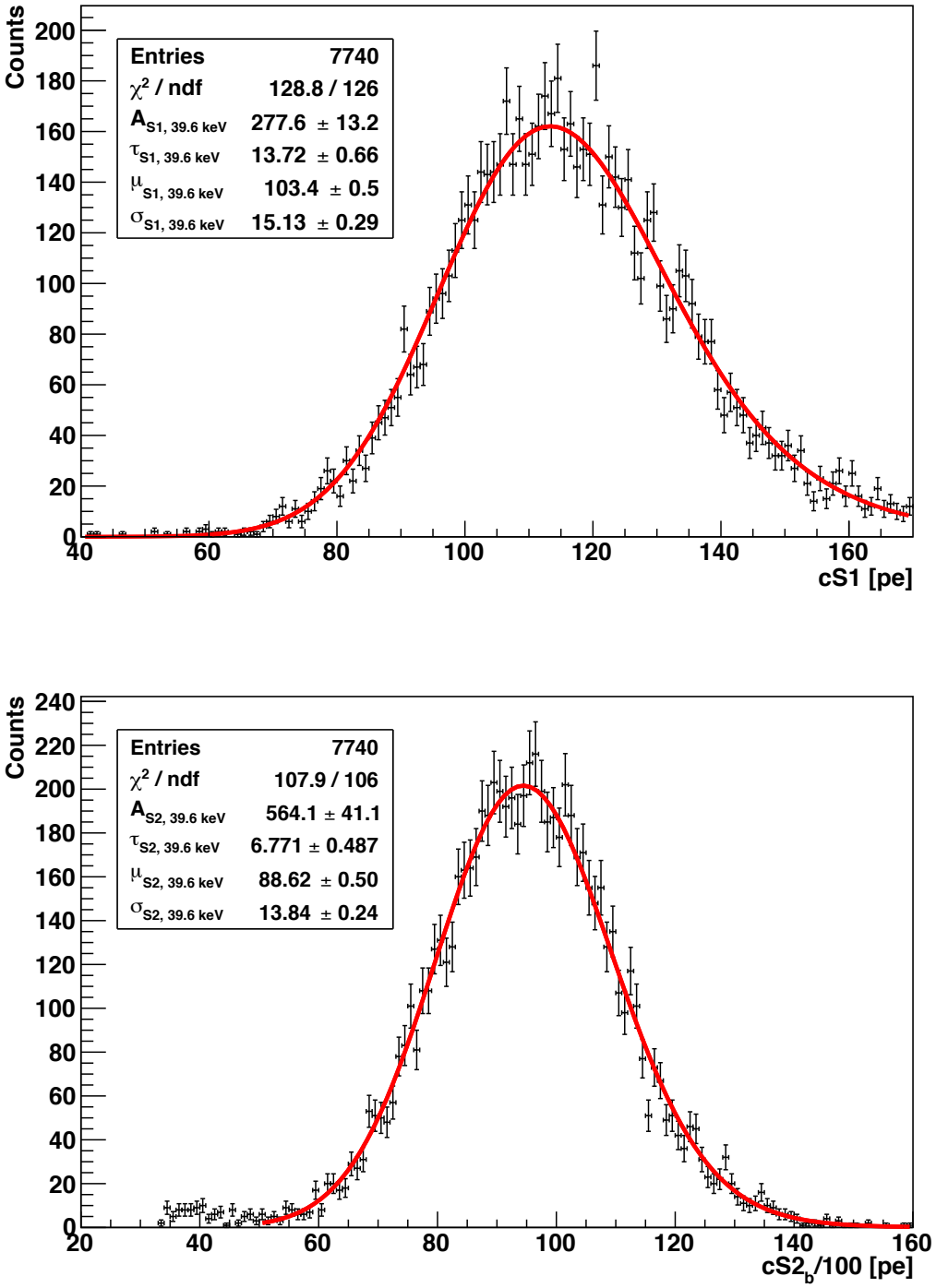

Figure 3.15: **Top:** Exponentially modified Gaussian fit of the 39.6 keV-line in $cS1$. **Bottom:** Fit of the same line in $cS2_b/100$.

model. Therefore, only the rotation angle of the ellipse is taken from the elliptical fit. Its deviation from the expected value from the means has to be accounted for by a systematic uncertainty. Comparing the means from the 1D fits with the angle at 39.6 keV yields

$$\arctan\left(\frac{(88.6 \pm 0.5)\ \text{pe}}{(103.4 \pm 0.5)\ \text{pe}}\right) = 0.71 \pm 0.01 \neq \theta = 0.56 \pm 0.02, \tag{3.35}$$

so $\Delta\theta_{\text{syst, 39.6 keV}} = 0.15$ is assumed. In the following all systematic uncertainties for θ are

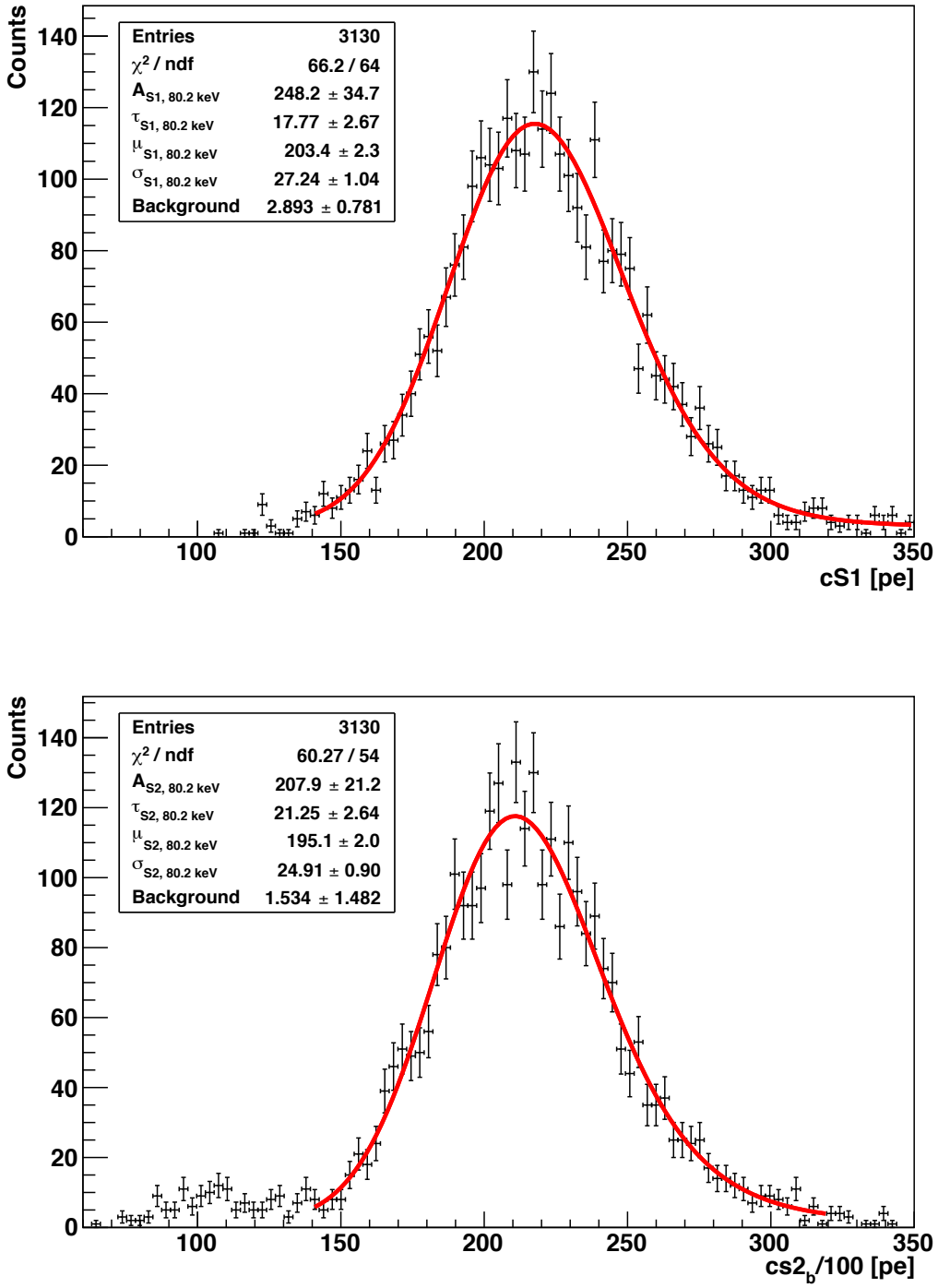

Figure 3.16: **Top:** Exponentially modified Gaussian fit of the 80.2 keV-line in $cS1$. **Bottom:** Fit of the same line in $cS2_b/100$.

estimated with this method. The systematic uncertainties of the means are estimated by studying the change of the values when varying the fit range.

It should be noted that the decay constants τ disagree between the two lines for $cS1$ and $cS2_b/100$ respectively. The decay constants should be the same for both lines if the spectra of the coincident scatters do not change. However, due to the different energy losses from the excitation of the xenon nuclei this could be the case.

Approaching the outer shell of the TPC the 109.9 keV-line activated in the PTFE starts to

appear in the S1-S2-spectrum. It is evident in Fig. 3.13 that it overlaps with the 80.2 keV, so a combined 2D fit, i.e. a fit with a sum of two tilted Gaussian ellipses, is applied. 1D fits with exponentially modified functions have been tested, but did not yield a usable result as both peaks merged into one when projecting to $cS1$. The fit results from the previous 2D fit of the 80.2 keV line have been used as initial guesses for the fit parameters of the respective ellipse. A constant background was added to both fits. The result is shown in Fig. 3.17.

109.9 keV

The low statistics caveats from before also apply here, albeit with a $\chi^2/ndf \approx 1.26$. All fit parameters for 80.2 keV agree with the previous isolated fit within the given uncertainties except for the μ_{S2}. At (206.7 ± 0.5) pe it is (3.50 ± 0.41) % lower than the value from the isolated 2D fit. This can be explained by incomplete charge collection close to the walls that is not sufficiently accounted for by the S2 correction map.
The shape of both ellipses given by a and b is similar, but the rotation is larger at $\theta_{109.9\ \mathrm{keV}} = 1.00 \pm 0.02$ compared to $\theta_{80.2\ \mathrm{keV}} = 0.84 \pm 0.02$. As a neutron-contribution to $\mu_{\mathrm{S1,\ 109.9\,keV}}$ and $\mu_{\mathrm{S2,\ 109.9\,keV}}$ could not be modeled by 1D fits, the centerpoint cannot be directly compared to the ellipse rotation angle. A conservative 10 % systematic uncertainty is assumed for both mean values in order to account for the fit method and loss of charge close to the TPC wall. In the spectrum from Fig. 3.13 the 197.1 keV line overlaps with both the 163.9 keV and 236.2 keV lines. The overlap is still present when reducing the radius for the data selection and it shifts the center positions of the 163.9 keV and 236.2 keV ellipses when fitting them individually on AmBe data. As the lines belong to metastable states of

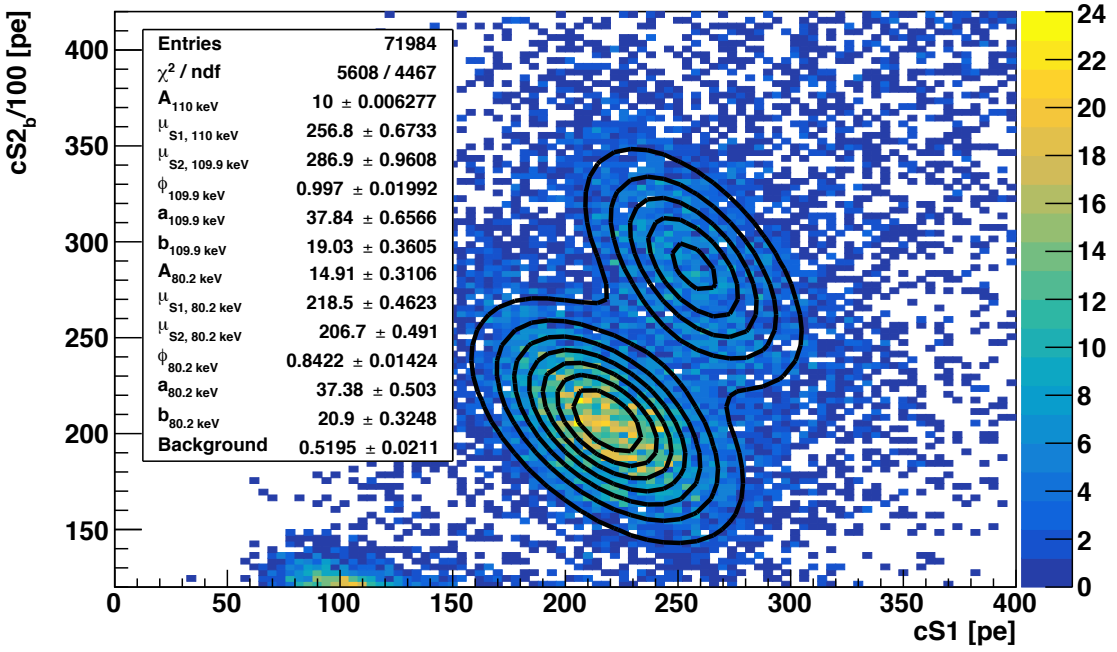

Figure 3.17: Combined elliptical fit of the 80.2 keV-line and the 109.9 keV-line. The cylindrical fiducial volume spans 95 % of the r-coordinate and the central 80 % of the z-coordinate.

^{129m}Xe and ^{131m}Xe with half-lives on the order of 10 days, they can also be found in data taken after the removal of the source. Here the short-lived ^{19}F lines are not present.

163.9 keV, 197.1 keV and 236.2 keV

The S1-S2-spectrum from such data taken over the course of one month is shown in Fig. 3.18. The same cuts have been applied as on the AmBe data except for a stricter ellipsoidal fiducial volume. This is necessary due to the long exposure where external backgrounds aggravate. It can bee seen in the plot that most low energy backgrounds are absent while high energy γ-lines from the ^{222}Rn decay chain and other internal sources are present above the metastable xenon lines and the aforementioned α-lines are also more visible. Only few interactions are present in the region of the elastic neutron recoils.
At this point it would be straightforward to fit the metastable lines individually. Afterwards a combined fit of three ellipses to the overlapping lines could be performed using the values from the individual fits on background as starting parameters. However, there is a more elegant solution founded in the fact that there are now two independent datasets containing the xenon lines. It is possible to perform a constrained fit, i.e. a fit where the χ^2 for multiple functions is minimized simultaneously.
In this case the multiple functions are single ellipses for metastable lines in background data and a sum of three ellipses for the overlapping lines in the AmBe data. This nets three functions. Yet, instead of fitting the overlapping lines in a single fiducial volume,

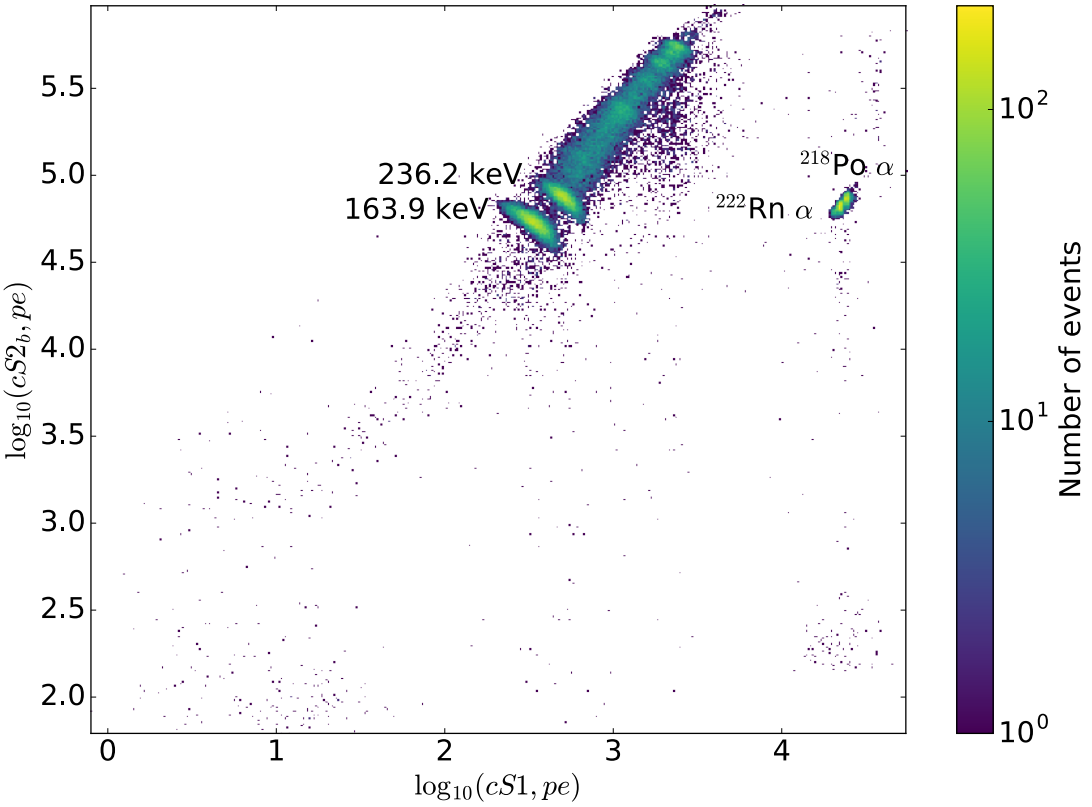

Figure 3.18: Background spectrum after AmBe calibration in $cS1$ and $cS2_b$ on logarithmic scales after cuts. The ellipsoidal 34 kg fiducial volume is used. The data has been taken from 11th February 2011 to 11th April 2011.

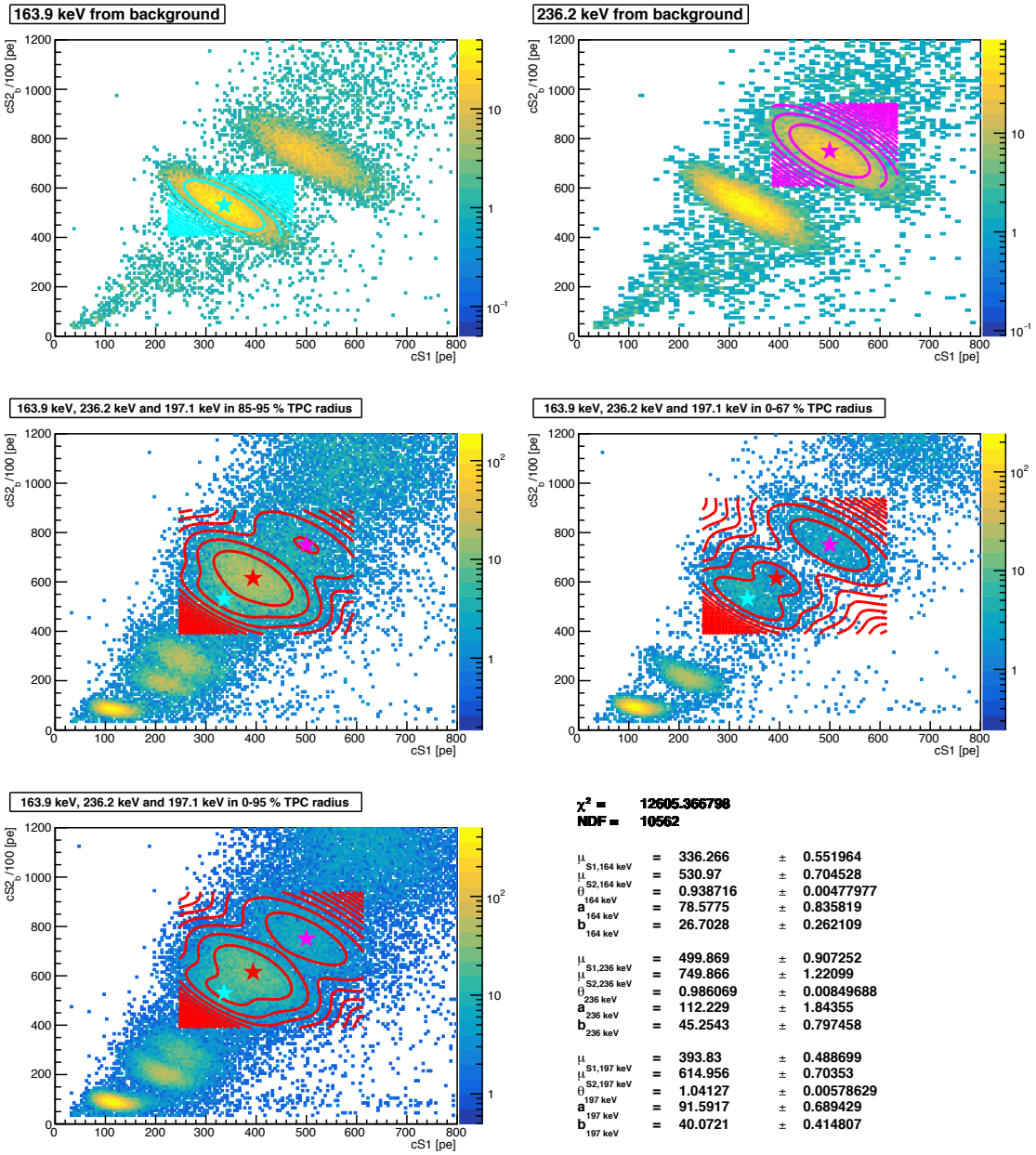

Figure 3.19: Combined fit of the γ-lines at 163.9 keV, 197.1 keV and 236.2 keV. Both top panels show the fits of the metastable xenon lines from background data. Starting from the left panel in the center row, the three other panels show the fits of the overlapping lines in three different fiducial volumes, namely an outer volume from 85 % to 95 % of the TPC radius, an inner volume up to 67 % of the TPC radius, and the full volume from the center of the TPC to 95 % of the radius. The color scales are logarithmic due to the large variation of event counts between the lines. The stars mark the ellipse centers.

three different fiducial volumes are chosen for the AmBe data: an outer volume from 85 % to 95 % of the TPC radius where the overlap is dominated by the fluorine line, an inner volume up to 67 % of the TPC radius where the metastable lines are more prominent, and the full volume from the center of the TPC to 95 % of the radius. Together with the background fits this makes five functions where two are single rotated 2D Gaussians and three are sums of three rotated 2D Gaussians. Apart from the amplitudes, the fit parameters for the respective ellipses are shared by all five functions. This amounts to a

total of 26 fit parameters that are optimized simultaneously. Fit ranges have to be set for all five fit functions in order to ensure fit convergence. The initial parameters are partly taken from individual fits of the metastable lines in background data. The result of the combined fit is shown in Fig. 3.19. The combined $\chi^2/ndf \approx 1.19$ is comparable to the individual fits of the other lines at similar statistics. The mean values as well as the rotation angles increase with energy. The systematic errors of the rotation angle and means are determined by a variation of the fit ranges.

Constructing the Linear CES

Tab. 3.6 summarizes the estimated parameters for all γ-lines with MINOS and systematic uncertainties. The values for ^{137}Cs have been taken from Fig. 3.11. Systematic uncertainties have been conservatively estimated from the differences in $\mu_{\mathrm{S1,\ 661.7\ kev}}$ and $\mu_{\mathrm{S2,\ 661.7\ kev}}$ between different double scatter cuts and fit ranges. These values are used to determine the light and charge yields of the datapoints for constructing the linear CES. Fig. 3.20 shows Q_{ee} against L_y for all lines. The error bars of the datapoints contain the parameter uncertainties from the fits as well as systematic uncertainties from the estimations outlined before. The uncertainties for the 109.9 keV line are large compared to the other points and result from the conservative estimation of systematics.

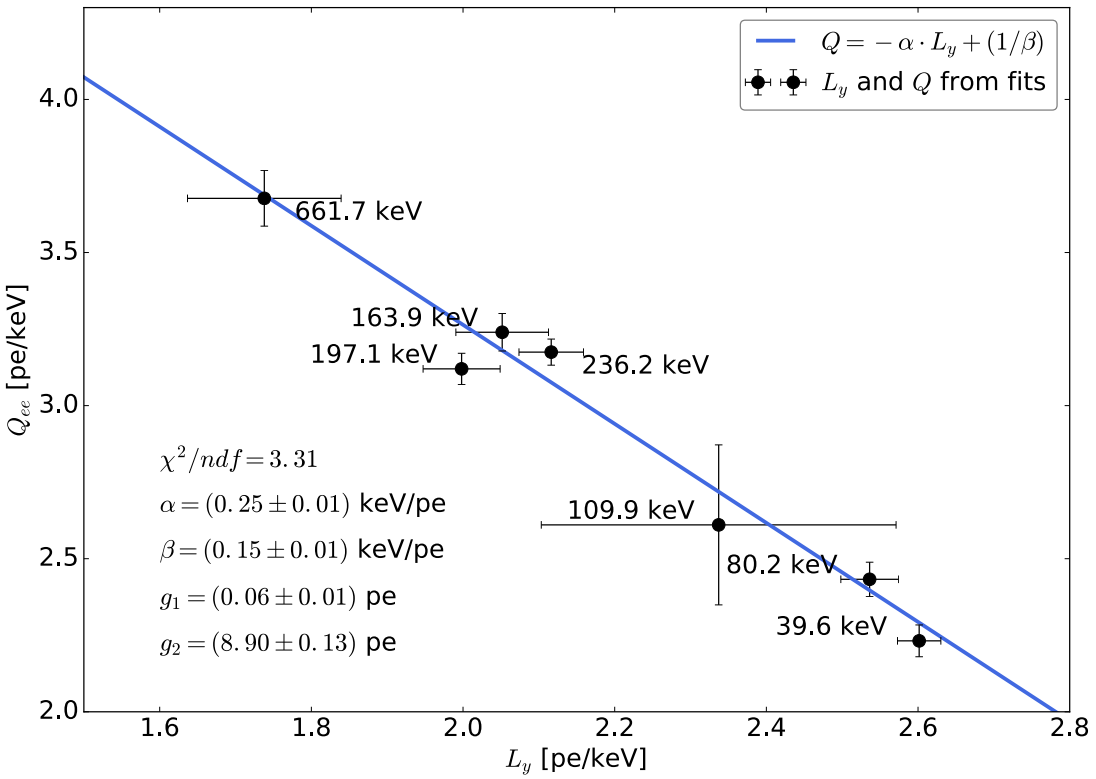

Figure 3.20: Derivation of the linear CES with constant α and β from fits of L_y and Q_{ee} calculated from the means of all fitted γ-lines.

Table 3.6: Fit results for all γ-lines. Both MINOS and systematic uncertainties are given in this order. For values not used in later analyses no systematic uncertainties have been estimated.

E [keV]	$cS1$ [pe]	$cS2_b/100$ [pe]	θ	χ^2/ndf
39.6 (2D)	115.7 ± 0.3	95.3 ± 0.2	$0.56 \pm 0.02 \pm 0.15$	1.03
39.6 (1D)	$103.4 \pm 0.5 \pm 1$	$88.6 \pm 0.5 \pm 2$	-	1.02, 1.02
80.2 (2D)	219.3 ± 0.7	214.2 ± 0.7	$0.81 \pm 0.02 \pm 0.05$	1.01
80.2 (1D)	$203.4 \pm 2.3 \pm 2$	$195.1 \pm 2.0 \pm 4$	-	1.03, 1.12
109.9	$256.8 \pm 0.7 \pm 2.7$	$286.9 \pm 1.0 \pm 28.7$	$1.00 \pm 0.02 \pm 0.09$	1.26
163.9	$336.3 \pm 0.6 \pm 10$	$531.0 \pm 0.8 \pm 10$	$0.94 \pm 0.01 \pm 0.07$	1.19
197.1	$393.8 \pm 0.5 \pm 10$	$615.0 \pm 0.8 \pm 10$	$1.04 \pm 0.01 \pm 0.04$	1.19
236.2	$499.9 \pm 1.0 \pm 10$	$749.9 \pm 1.3 \pm 10$	0.99 ± 0.01	1.19
661.7	$1156 \pm 2 \pm 67$	$2438 \pm 3 \pm 60$	$1.17 \pm 0.01 \pm 0.05$	2.95

The CES parameters α and β are determined by a linear fit with the function from Eq. 3.20 [47]. The related parameters g_1 and g_2 are calculated using Eq. 3.19. The result is

$$\begin{aligned} \alpha &= (0.25 \pm 0.01)\,\text{keV/pe}, \\ \beta &= (0.15 \pm 0.01)\,\text{keV/pe}, \\ g_1 &= (0.06 \pm 0.01)\,\text{pe}, \\ g_2 &= (8.90 \pm 0.13)\,\text{pe}. \end{aligned} \tag{3.36}$$

The fit function does not cross the error bars of all points and yields $\chi^2/ndf \approx 3.31$. The general trend of the model indicates that that the higher a γ energy is, the higher Q_{ee} is while L_y decreases. This trend is broken by the overlapping lines from the combined fit. The position of the 236.2 keV line can be explained by its twofold decay. The first decay with an energy of 196.6 keV should have a light yield of about 2 pe/keV while the second 39.6 keV decay corresponds to the lowest energy line in the plot with $L_y \approx 2.6$. Thus, the overall light yield is higher while the charge yield is reduced. Still, this only explains the shift on the line to the right, but not the deviation from the line itself. The combined energy information from light and charge should correspond to a point on the line. One explanation is that a non-negligible portion of the energy is lost into a channel invisible for the detector and from now on referred to as heat. The low charge yields of the 109.9 keV and 197.1 keV fluorine lines could result from incomplete charge collection close to the TPC walls.

The linear CES with constant α and β is applied to the ^{137}Cs, ^{241}AmBe and background datasets in the respective fiducial volumes. The results are shown in figures 3.21, 3.22

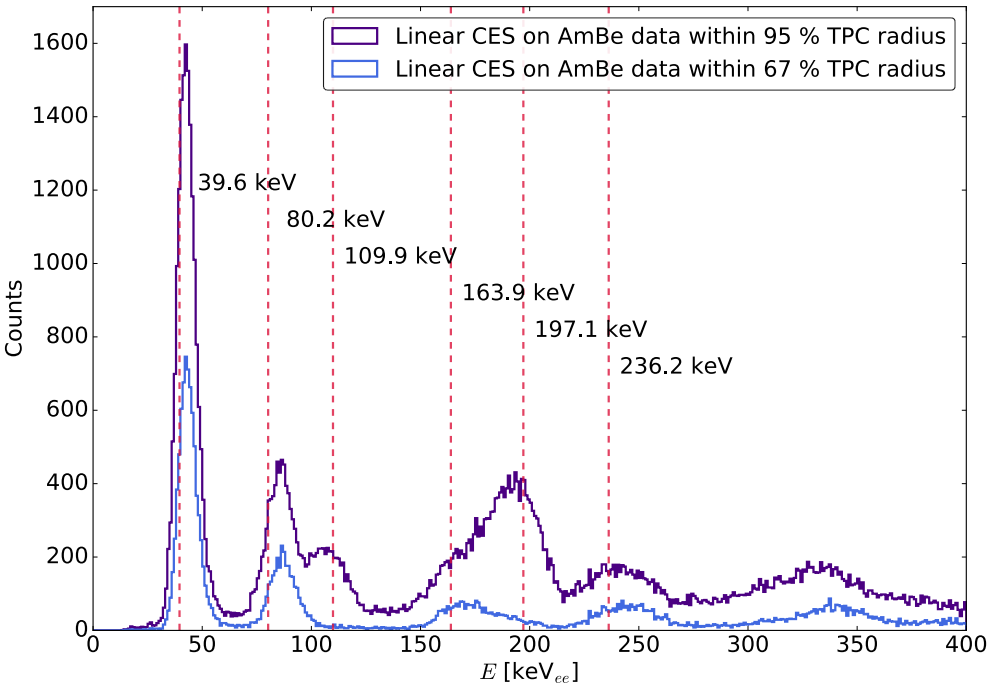

Figure 3.21: Linear CES with constant α and β applied to ^{241}AmBe data in the two cylindrical fiducial volumes.

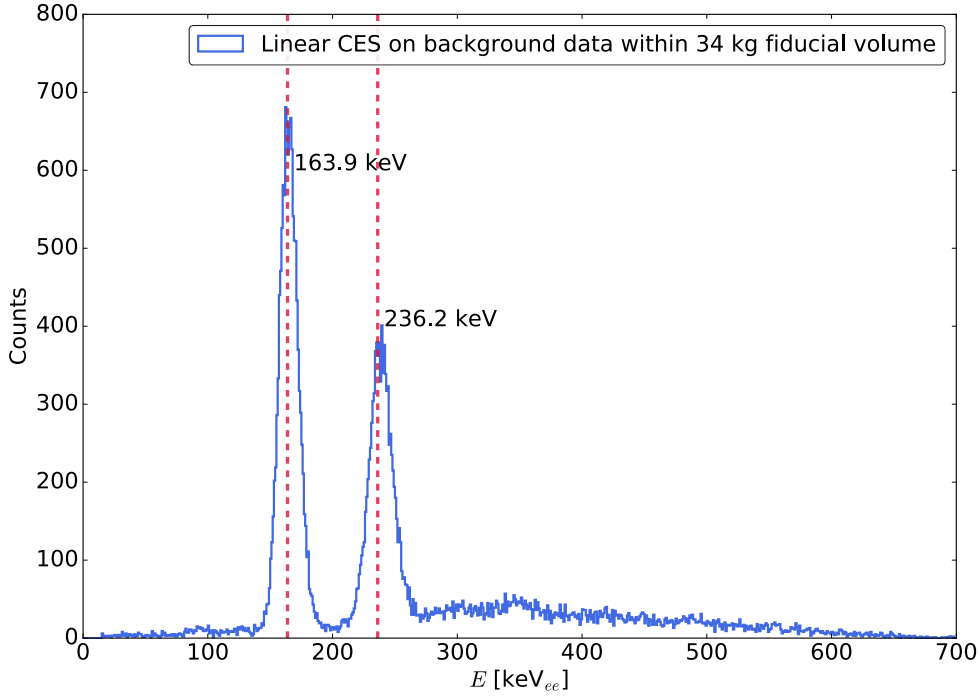

Figure 3.22: Linear CES with constant α and β applied to background data in the 34 kg fiducial volume.

and 3.23. All lines show a reasonable agreement with the expected values. The slight offset of the low energy xenon lines to the right in the AmBe datasets is due to the altered shape of the peaks by coincident neutron scatters. The higher energy peaks exhibit low statistics in the more restrictive volume and are immersed in background in the larger one. As expected the 109.9 keV line is not present in the smaller volume while there is a small contribution from fluorine at 197.1 keV. In the background and ^{137}Cs datasets the

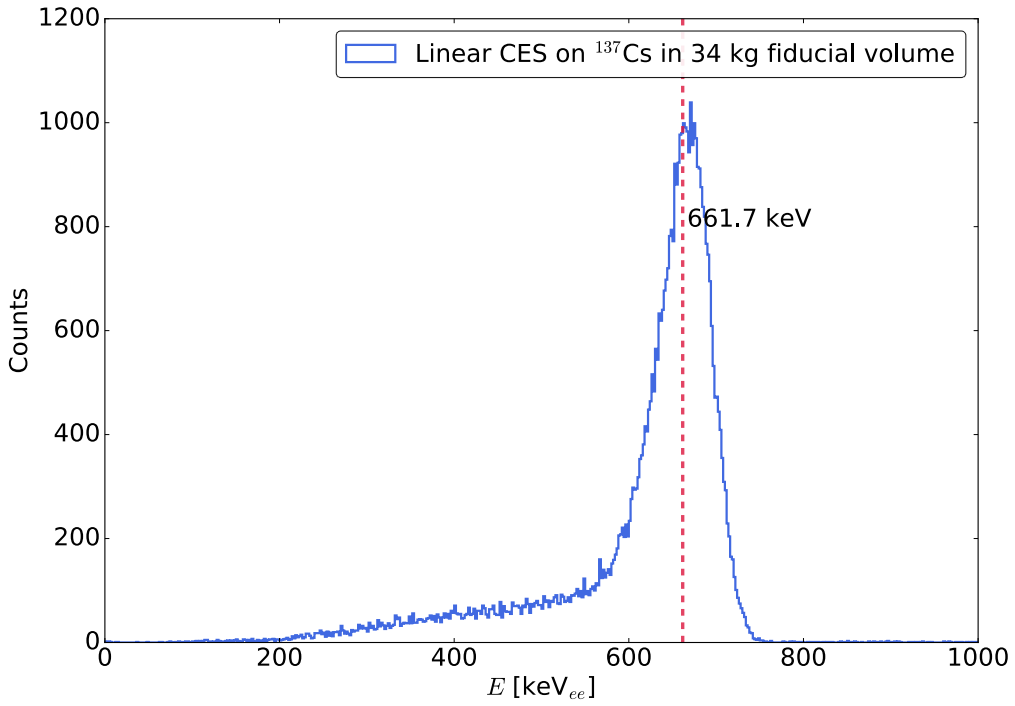

Figure 3.23: Linear CES with constant α and β applied to ^{137}Cs data in the 34 kg fiducial volume.

lines show good agreement with the true energies. The spectrum from the background data exhibits a mostly featureless background distribution with a peak around 350 keV. This can be attributed to a γ decay of ^{214}Pb at 352 keV [44] which is a daughter isotope of ^{222}Rn. Using a longer exposure dataset more background peaks could be identified, but this is not part of this thesis. A detailed discussion of the resolution and accuracy of this and the other two CES will be provided in section 3.3.8.

3.3.7 Global CES by Angular Parametrization of α and β

The linear CES assumes that α and β are constant over the whole energy range. However, in the general case both are energy-dependent. The energy dependence can be parametrized by transitioning from α and β to $k(E)$ and $\theta(E)$ from Eq. 3.23. The CES can then be constructed by fitting both parameters at different energies with continuous functions. These are then inserted into Eq. 3.22. This approach has one pitfall. In the end each event's value pair $(cS1, cS2_b/100)$ has to be converted into an energy with $k(E)$ and $\theta(E)$. If these are functions of energy, they require the energy of each event to be given as an input before providing it as an output. Consequently, another measure of the energy has to be found avoiding the recursive definition. Viable candidates are $cS1$, $cS2_b$ or a sum of both. This sum can additionally be weighted, for example with α and β from the linear CES. In this analysis no weighting is applied in order to compare the linear and the parametrized CES independent of each other.
The fits of $k(x)$ and $\theta(x)$ with $x = cS1 + cS2_b/100$ are shown in Fig. 3.24. The datapoint for the 236.2 keV line has been excluded from the fit due to the twofold decay. The angle $\theta_{cS137} = 1.17 \pm 0.01 \pm 0.05$ has been determined from an ellipse fit to the dataset in Fig.

3.11 with the same fit range, but an unconstrained rotation angle. As for the other angles the systematic error has been assessed by a comparison with the means and a variation of

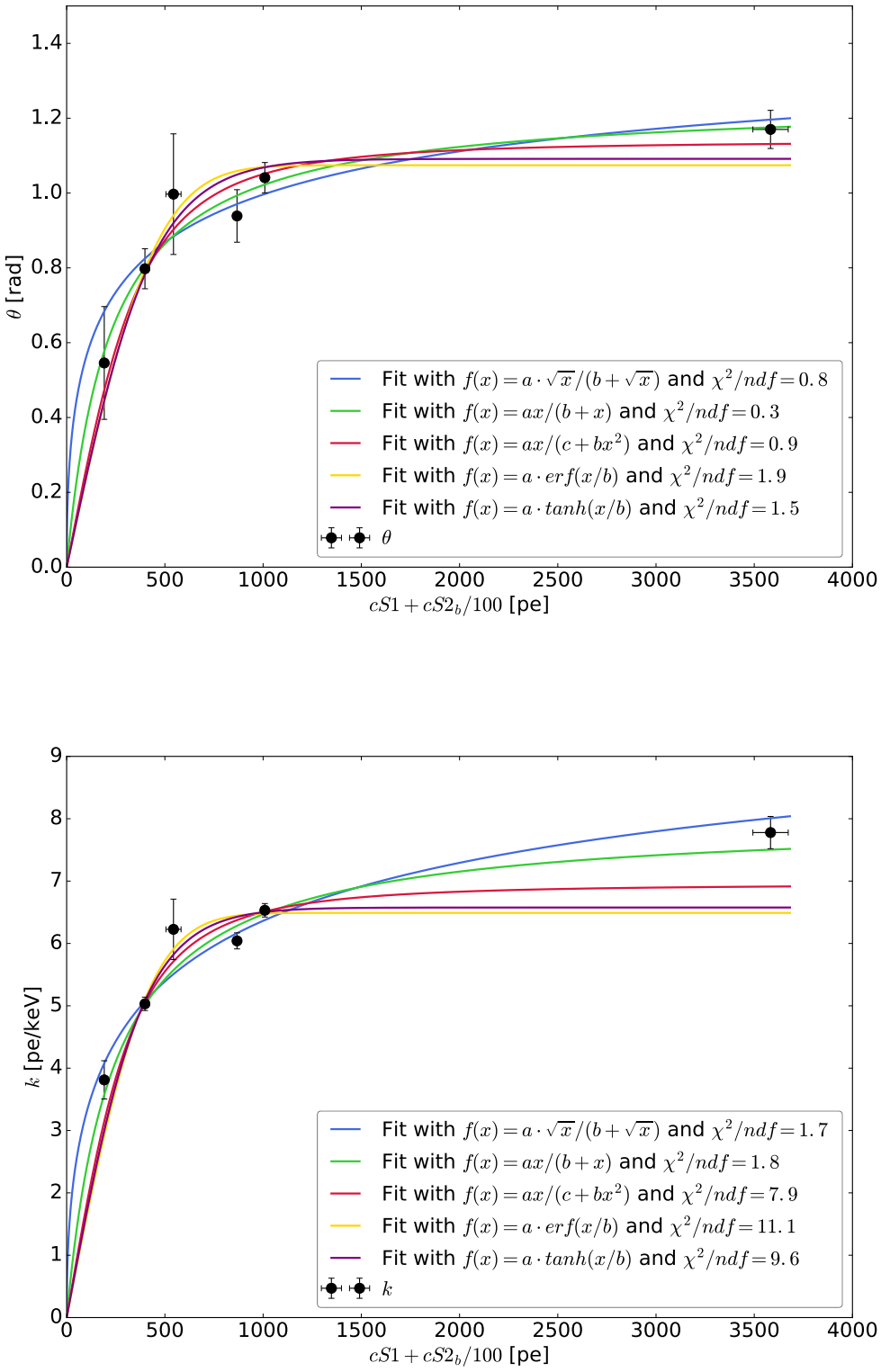

Figure 3.24: Fits of k and θ in $(cS1 + cS2_b/100)$ for all points except 236.2 keV.

the fit range. Assuming an asymptotic behaviour

$$\begin{aligned} \lim_{x\to 0} \theta(x) &= 0, \\ \lim_{x\to 0} k(x) &= 0. \end{aligned} \tag{3.37}$$

both parameters show a steep rise at low energies and saturation at higher energies. Sigmoid functions are evaluated as phenomenological models describing this behaviour with as few fit parameters as possible. The candidate functions are:

$$\begin{aligned} f_1(x) &= \frac{a \cdot \sqrt{x}}{b + \sqrt{x}}, \\ f_2(x) &= \frac{a \cdot x}{b + x}, \\ f_3(x) &= \frac{a \cdot x}{b + cx^2}, \\ f_4(x) &= a \cdot \mathrm{erf}\left(\frac{x}{b}\right), \\ f_5(x) &= a \cdot \tanh\left(\frac{x}{b}\right). \end{aligned} \tag{3.38}$$

The first three functions yield $\chi^2/ndf < 1$ in the fits of θ which can be explained by an overly conservative estimation of the systematic errors. As the right model is not known, the errors cannot be normed by requiring $\chi^2/ndf = 1$. Decreasing the errors for testing shows that f_1 still achieves the best reduced χ^2 followed by f_2. However, the conservative errors are kept to be on the safe side. In the future an actual model could be derived for the evolution of the angle instead of using a phenomenological function. Based on the low quality of the fits in both parameters f_4 and f_5 are discarded. So is f_3 given the performance in the k fit. Besides it requires a third fit parameter other than f_1 and f_2. Providing the best fits in each case f_2 is adopted for θ while f_1 is used for k with

$$\begin{aligned} \theta(x) &= \frac{a_\theta \cdot x}{b_\theta + x}, \\ k(x) &= \frac{a_k \cdot \sqrt{x}}{b_k + \sqrt{x}}, \end{aligned} \tag{3.39}$$

where the fit parameters are

$$\begin{aligned} a_\theta = (1.25 \pm 0.01) \quad &\text{and} \quad b_\theta = (222.59 \pm 9.64), \\ a_k = (11.27 \pm 0.86)\,\mathrm{pe/keV} \quad &\text{and} \quad b_k = (24.49 \pm 1.90)\,\mathrm{pe}. \end{aligned} \tag{3.40}$$

The application of the CES to the ^{241}AmBe and background data is shown in Fig. 3.25. As with the linear CES there is an apparent offset of the 39.6 keV and 80.2 keV peaks to the right of the true energy. The matching appears good for the 109.9 keV and 163.9 keV peaks, but the mean of the 236.2 keV line is shifted to a higher than the true energy. This anomaly could be due to its double decay property. A detailed discussion of the accuracy and resolution is presented in the next section.

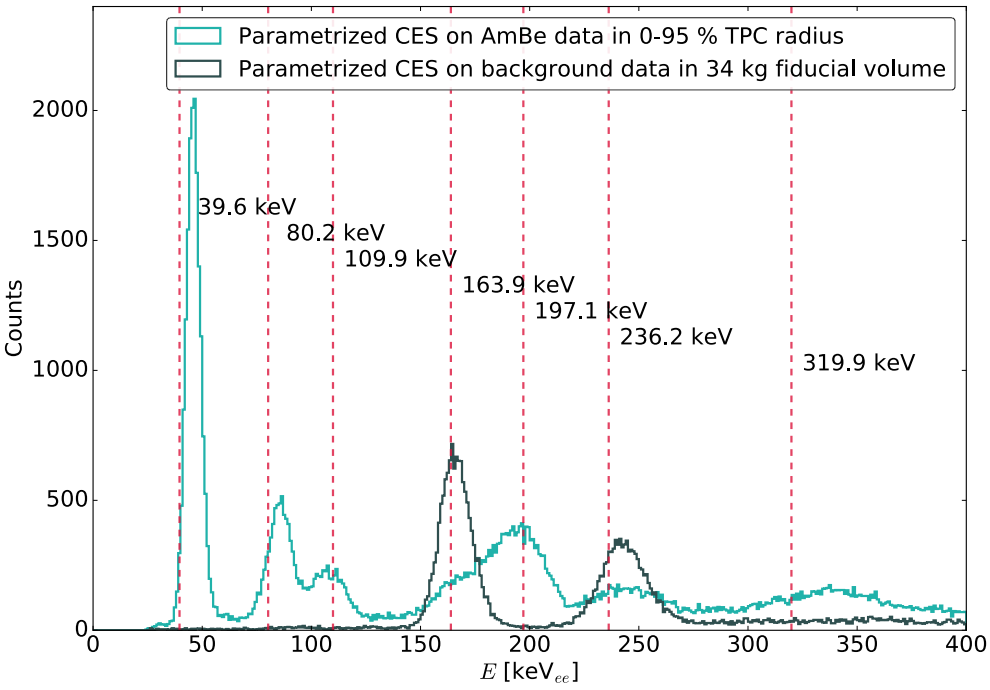

Figure 3.25: Energy spectrum from the parametrized CES on ^{241}AmBe and background data.

3.3.8 Comparison of the Three CES Methods

All three CES methods are compared with regard to the accuracy to the true energy and the energy resolution. The resolution of all three energy calibrations is depicted in Fig. 3.26. Only the $^{129(m)}$Xe, $^{131(m)}$Xe and ^{137}Xe lines in the 34 kg fiducial volume have been used. The ^{19}F lines are excluded because of their proximity to the TPC wall. The mean energies E and standard deviations σ have been extracted from Gaussian fits of the data after the application of the different CES. The metastable lines have been fitted on the background data while the AmBe data has been used for the other two. The fits can be found in the appendix in section 6.6. In addition to the parameter uncertainties of the fits an additional uncertainty resulting from the uncertainties of the CES parameters themselves has been added to the points. The assessment of this uncertainty can be found in section 6.6. The points have been fitted with the function

$$\frac{\sigma}{E} = \frac{c_1}{\sqrt{E}} + c_2 \tag{3.41}$$

which has been used in [30] to describe energy resolution as a function of energy. It is evident that both the linear CES as well as the CES parametrized with the ellipse rotation angle offer a better energy resolution than the CES from the single ^{137}Cs line. Although larger errors have been assumed for the parametrized CES, $\chi^2/ndf \approx 9.7$ of the fit to the linear CES is smaller. Still, the line does not cross the error bars of all points. Additional systematics can result from the fit ranges as well as from parameter correlations from the combined fits of several lines although these should be addressed by MINOS. The fit

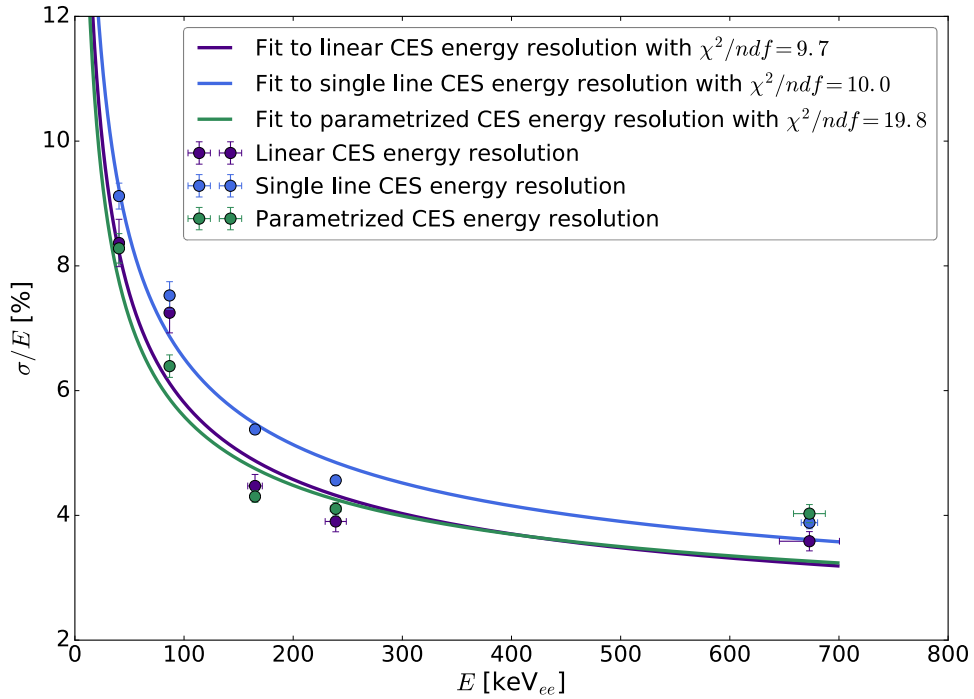

Figure 3.26: Energy resolution of the linear CES, single line CES from ^{137}Cs and the CES parametrized by the ellipse rotation angle.

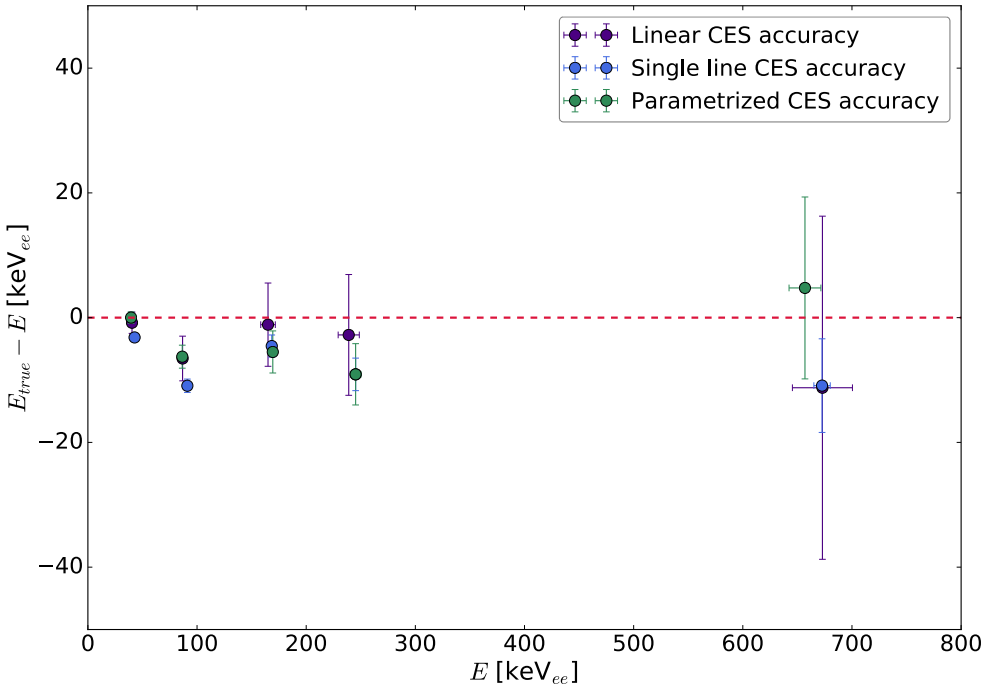

Figure 3.27: Accuracy of the linear CES, single line CES from ^{137}Cs and the CES parametrized by the ellipse rotation angle.

parameters for the linear CES are

$$\begin{aligned} c_1 &= (42.50 \pm 3.17)\,\sqrt{\text{keV}}, \\ c_2 &= (1.58 \pm 0.23)\,. \end{aligned} \tag{3.42}$$

This yields a resolution of $(3.1 \pm 0.3)\,\%$ at $1\,\text{MeV}_{\text{ee}}$ which is a factor of 0.5 larger than the benchmark value of $1.9\,\%$ in [30]. However, the energy resolution could be further improved

with a better tuning of cuts for electronic recoils in the 100 keV range. All applied cuts are optimized for low energy nuclear recoils of only few keV.
The accuracy of all three fits is shown in Fig. 3.27. Only the error bars of the linear CES datapoints cross the line of zero deviation from the true energy. As mentioned before, systematics could have been underestimated. Even then the linear CES offers several benefits over the other two methods.
First of all, it incorporates the energy information of more than one γ-line. The single line CES is highly susceptible to fit modalities which is in this case the difficulty of the Compton spectrum for the ^{137}Cs line. In addition, a single line is blind to specific effects for certain energy ranges, for example saturation. Even though a linear model cannot account for saturation either, the effect becomes visible if more than one line is used and can be accounted for by changing the detector settings or model. Compared to the parametrized CES the linear model is simpler, as it requires less fits and no decision for a model of θ and k. The impact of twofold decays or energy going into other channels than light and charge on θ is also not clear. In this work no benefit of the chosen parametrization, which would justify the higher complexity of this approach, over the other methods of constructing a CES could be shown. Still, a further investigation of the ellipse rotation and anti-correlation might be of interest because it provides a deeper insight into the mechanisms generating signals in liquid xenon. It is also relevant for other experiments exploiting it. One example is *EXO-200* which observes minimal change of the ellipse rotation angles at the relevant energies exceeding 1 MeV [52].

3.3.9 Nuclear Recoil Energy Scale

The energy scale for nuclear recoils is determined using the S1 signal. This is done by interpolating L_y of all datapoints with the NEST model [31]. NEST explains the energy dependence of light and charge yields in liquid noble gases by modeling recombination of ionization charges in Monte Carlo simulations. The main mechanisms governing recombination stem from incident particle type, linear energy transfer $\frac{dE}{dx}$, and the applied electric field. The model is continuously extended and also allows to analytically calculate the absolute light and charge yields for α and β particles as well as neutrons at a given electric field. In this work the NEST spreadsheet [31] is used which provides tabulated values for absolute light yields of γ-rays from the software's version v0.99. Values are given in the energy range from 1 $\mathrm{keV_{ee}}$ to 5 $\mathrm{MeV_{ee}}$ at six electric fields. The values at $\mathcal{E} = 500\,\mathrm{V/cm}$ are taken insofar as they are closest to the actual field of $\mathcal{E} = 530\,\mathrm{V/cm}$. The tabulated values will provide a slightly higher number of photons than expected from the original field, but the effect is expected to be on the order of 1 % which is smaller than the quoted systematic uncertainty of the model at 4 %. Figures 3.28 and 3.29 show the datapoints for all L_y and Q_{ee} datapoints overlaid with the NEST predicition.
The absolute numbers of generated quanta per unit energy have been converted to measured S1 and S2 signals by applying g_1 and g_2 from the linear CES. These have been calculated with Eq. 3.19 and $W = 13.7\,\mathrm{eV}$ assumed by NEST:

$$L_y = g_1 \cdot L_{\mathrm{abs}} \tag{3.43}$$

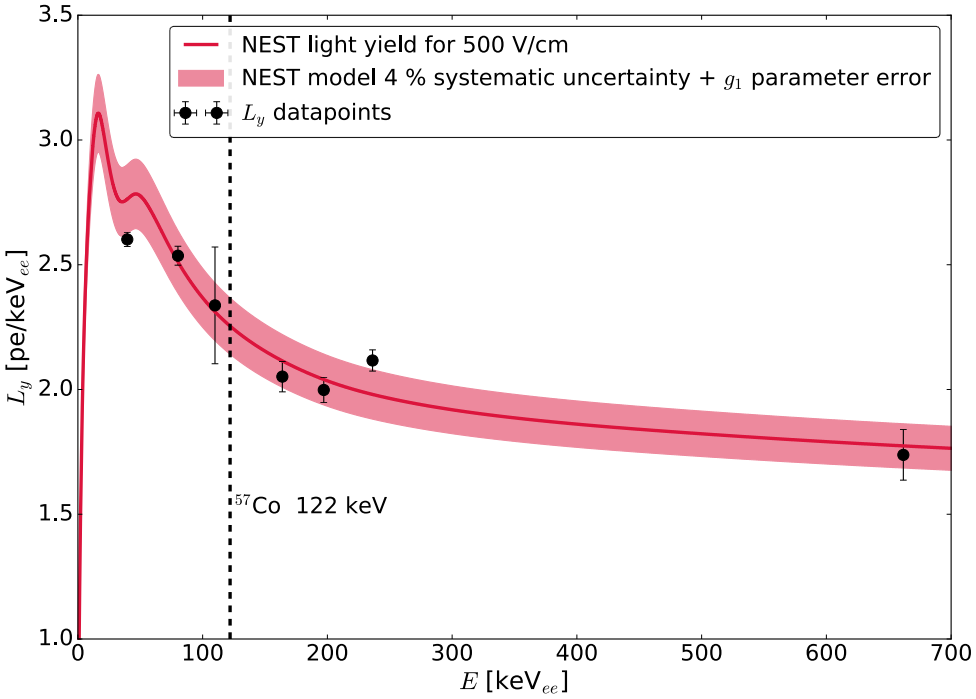

Figure 3.28: L_y datapoints overlaid with the NEST model.

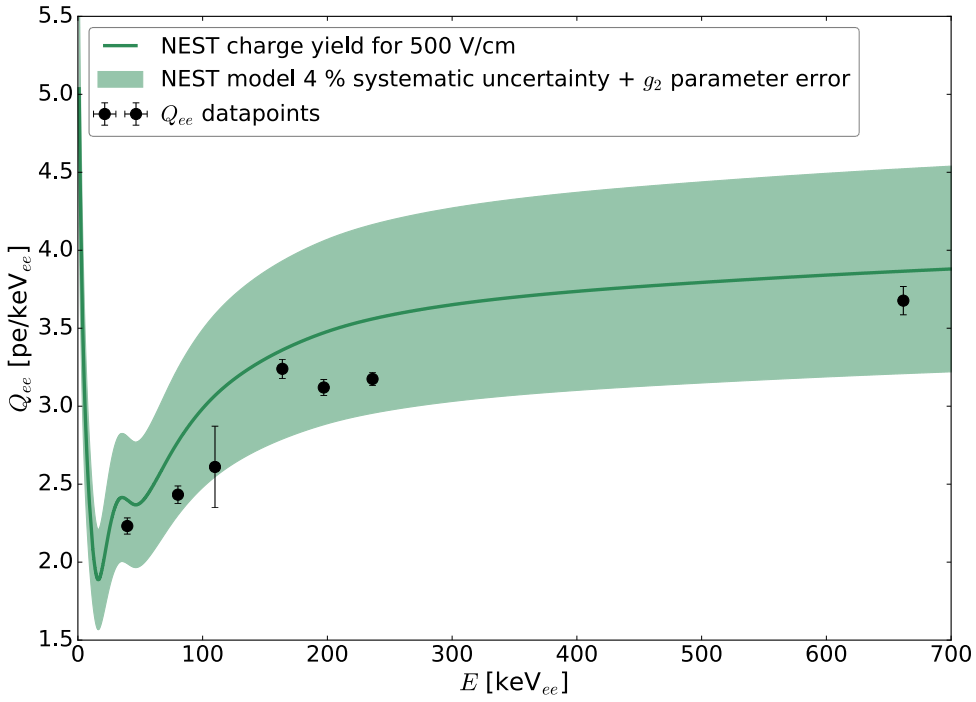

Figure 3.29: Q_{ee} datapoints overlaid with the NEST model.

The effective W-value signifies the generation of 73 quanta/keV regardless of the total energy of the incident particle. Thus, the charge yield can be calculated by

$$Q_{\mathrm{ee}} = g_2 \cdot (73\,\mathrm{keV}^{-1} - L_{\mathrm{abs}}). \tag{3.44}$$

The parameter uncertainties of g_1 and g_2 as well as the 4 % systematic uncertainty of the model are taken into account in the uncertainty bands of both L_y and Q_{ee}. It is evident

that all points are in agreement with the prediction for Q_{ee} where the model has a wide uncertainty band.
The L_y prediction is in good agreement with all datapoints with the exception of 39.6 keV and 236.2 keV. The latter exhibits a higher light yield than a monoenergetic γ of this energy due to its two-fold decay where both individual light yields are expected to be higher. Therefore an upward shift from the prediction is expected and rather confirms its validity than challenge it.
The datapoint at 39.6 keV is situated in a prominent dip which is present because of a xenon K-edge resonance for the production of low-energy Auger electrons. These have a different linear energy transfer than γ-rays which requires the use of another recombination model which outputs a decrease in scintillation. The developers note that the modeled decrease is too shallow, though, due to an inedequacy of the utilized simulation framework NEST has been embedded in [31]. Consequently, the downward deviation of this datapoint can be explained.
After assuring the accordance of model and data NEST can be used to determine the light yield at the reference electronic recoil equivalent energy. The interpolation of the datapoints provides an absolute light yield of $L_{\mathrm{abs}}(122\,\mathrm{keV_{ee}}) = 37.57\,\mathrm{keV}^{-1}$ which yields

$$L_y(122\,\mathrm{keV_{ee}}) = (2.25 \pm 0.38_{g_1} \pm 0.09_{\mathrm{model}})\,\mathrm{pe/keV_{ee}}. \tag{3.45}$$

This agrees with the original value of (2.28 ± 0.04) pe/keVee within the error [34]. However, the uncertainty of the value in this work is larger from the model uncertainty alone. Neglecting this, the parameter uncertainty is still almost a factor of ten larger than in the original work. One contributing factor is the conservative estimation of systematic errors in the linear CES that g_1 is derived from.
Knowing L_y, NR energies can be directly determined from $cS1$ using Eq. 1.22 with $S_{\mathrm{nr}} = 0.95$, $S_{\mathrm{ee}} = 0.58$, and the cubic spline interpolation of $\mathcal{L}_{\mathrm{eff}}$ from [42] shown in Fig. 2.4. The boundaries of the dark matter search in $cS1$ are translated to E_{nr}:

$$\begin{aligned} 3\,\mathrm{pe} &\rightarrow 6.7\,\mathrm{keV_{nr}}, \\ 20\,\mathrm{pe} &\rightarrow 30.9\,\mathrm{keV_{nr}} \end{aligned} \tag{3.46}$$

Due to the lower light yield in this work it is shifted upward compared to the original bounds from 6.6 $\mathrm{keV_{nr}}$ to 30.5 $\mathrm{keV_{nr}}$ [34].

3.3.10 Preliminary summary

The definition of a nuclear recoil energy scale marks the completion of dataset preparation. The steps introduced in this chapter marked the migration from the old analysis framework centered around ROOT to the pythonic PAX environment. This included the translation and application of the legacy cuts which have been developed for and tuned to the old data processor. A new set of cuts has been developed to augment and replace them.
Moreover, calibrations for ER and NR interactions have been introduced. Three methods for the construction of a combined energy calibration using both the S1 and the S2 signal of various γ-lines have been compared with each other. The linear CES constructed from

the means of the full absorption lines offered the best resolution and accuracy to the true energy. The method for its construction is directly applicable to XENON1T data after slight modifications and can be extended with Q_y, L_y and $\mathcal{L}_{\text{eff}}$ to describe nuclear recoils as well. The photon detection efficiency as well as the charge amplification factor of the XENON100 detector have been found with

$$\begin{aligned} g_1 &= (0.06 \pm 0.01)\,\text{pe}, \\ g_2 &= (8.90 \pm 0.13)\,\text{pe}. \end{aligned} \tag{3.47}$$

Subsequently, the light yield at the reference energy has been determined from matching measured light yields with the NEST model. A good agreement with the model yielded:

$$L_y(122\,\text{keV}_{\text{ee}}) = (2.25 \pm 0.38_{g_1} \pm 0.09_{\text{model}})\,\text{pe/keV}_{\text{ee}}. \tag{3.48}$$

which is in agreement with the value from the benchmark analysis [34]. The NR energy calibration was then defined applying a parametrization of the relative scintillation efficiency of nuclear recoils $\mathcal{L}_{\text{eff}}$ using all existing measurements. With the energy calibration finished, the search for dark matter events in 224.6 live days of data commences.

CHAPTER

FOUR

DARK MATTER SEARCH

At this point only three steps are missing before an upper limit σ_{SI} for spin-independent WIMP nucleon interactions can be set. First the nuclear and electronic recoil bands in discrimination space $\log_{10}\left(\frac{cS2_b}{cS1}\right)$ have to be defined. These bands define a dark matter search region with an optimized ratio of signal acceptance vs background rejection. The datapoints left in this region after cuts are the dark matter signal candidates and have to be scrutinized for pathologies that have not been removed by the cuts[1]. At this point the remaining signal events in the benchmark analyis and this work can be directly compared. The limit on σ_{SI} is then set by comparing the number of candidate events with the background expectation.

4.1 Definition of Electronic and Nuclear Recoil Bands

The electronic recoil band is defined from the Compton continuum of an external ^{60}Co source and the decay radiations from a ^{232}Th source. The band is characterized by its mean and 3.09σ interval. Both are taken from fits of Gaussian functions to 1 pe slices of the band in discrimination space. The 3.09σ excursion from the Gaussian mean marks the upper boundary of the dark matter search region. Only events below this discrimination line are regarded as signal candidates which translates to an electronic recoil event rejection of 99.75 %. Continuous functions for the ER mean and discrimination line are obtained by overlaying the means and standard deviations with the sum of an exponential function and a fourth order polynomial:

$$f_{\mathrm{ER,\ NR}}(x) = e^{a-b\cdot x} + c\cdot x^4 + d\cdot x^3 + e\cdot x^2 + f\cdot x + g. \tag{4.1}$$

This phenomenological model[2] is inspired by the band flattening procedure in [34]. The ER band result is displayed in Fig. 4.1. The band is defined in the $cS1$ interval between 3 pe and 60 pe[3]. This conforms to the lower threshold of the dark matter ROI in [34]. The band is dominated by the exponential decrease up until ~ 30 pe and flat thereafter.

[1] This procedure is colloquially referred to as *waveform watching*.

[2] More recent developments mark the incremental replacement of phenomenological approaches by sophisticated ER band models that are constructed from Monte Carlo simulations [53].

[3] The fits for the ER and NR bands can be found in the appendix in section 6.8.

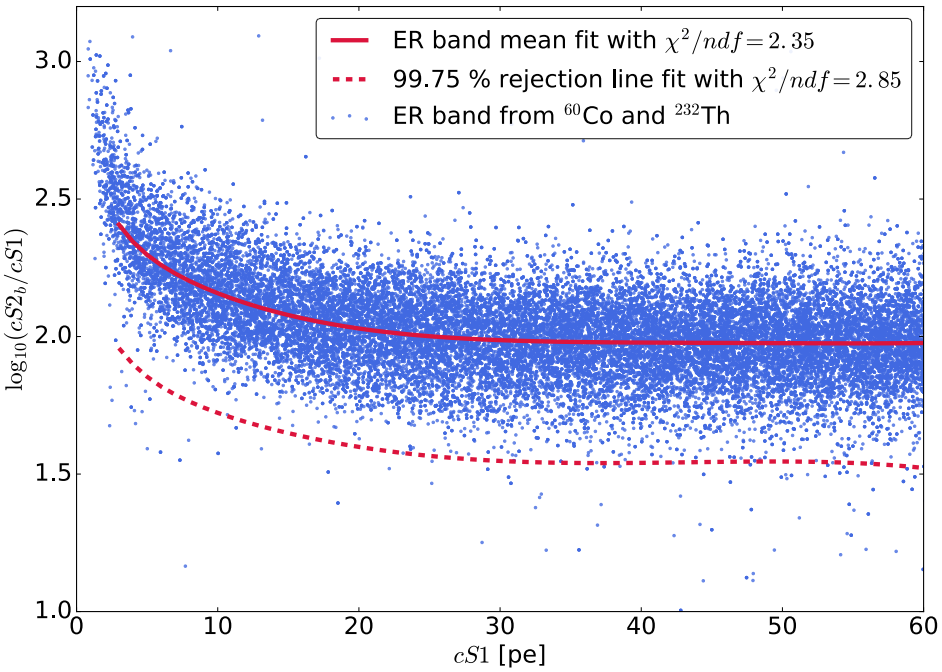

Figure 4.1: ER band from ^{60}Co and ^{232}Th data up to 60 pe in $cS1$. All legacy cuts applied on the dark matter data are also applied on this dataset. The band mean and 99.75 % exclusion line are constructed from fitting 1 pe slices with 1D Gaussian distributions. The means and 3.09σ intervals are in turn fitted with a phenomenological function starting at the lower dark matter search region threshold of $cS1 = 3$ pe.

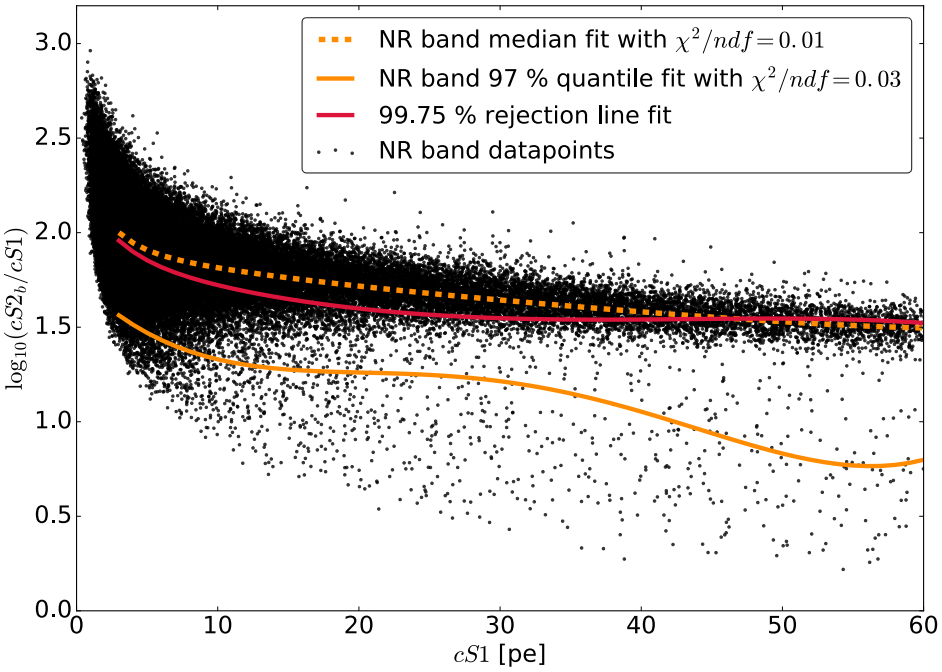

Figure 4.2: NR band from ^{241}AmBe data up to 60 pe in $cS1$. All legacy cuts applied on the dark matter data are applied. The band median and 97 % quantile are constructed from 1 pe slices. The datapoints from the individual slices are overlaid with a phenomenological function starting at the lower dark matter search region threshold of $cS1 = 3$ pe.

A similar approach is applied to the NR band from elastic ^{241}AmBe neutron scatters shown in Fig. 4.2. The same range and cuts are used and the band is again divided into slices of 1 pe. As the shape of the band is different a statistical approach is pursued rather than fits of the slices. Due to lower statistics outliers strongly influence the mean of the band. Thus, the NR band median is used to define the band instead of the mean. The lower boundary of the dark matter ROI is defined as the 97 % quantile of the NR band. Only points above this line fall into the category of signal candidates. The points for the individual slices are joined by the same phenomenological function as above.
The reduced χ^2 of the NR band fits is on the order of 0.01. This is partly due to the fact that no errors have been assumed for the datapoints from the individual slices. Looking at the fits in section 6.8 it is evident that the median datapoints are well described by the model. In contrast, the 97 % quantile exhibits large fluctuations above 15 pe causing a poor fit quality. These fluctuations originate from low statistics within the 1 pe slices. Above 30 pe a single slice of the band contains less than 200 events. This means that a lone outlier point can shift the position of the 97 % quantile significantly. The presence of outliers can be attributed to an insufficient tuning of cuts in this energy region, so more effective cuts would alleviate the problem. A larger AmBe dataset could also mitigate the effect by delivering more statistics close to the band median.
Yet, the approach is maintained for the time being in order to maintain comparability of this analysis with [34]. Additionaly, the effect is small in the benchmark dark matter search region from the original analysis stretching from 3 pe to 20 pe. It should be considered when comparing the original NR band with one in this work though. A direct comparison of the bands requires the transition to flattened discrimination space. This is treated in the next section.

4.1.1 Band Flattening

Flattened discrimination space is defined by the mean of the electronic recoil band which is subtracted from all values of the discrimination parameter. Thus, the ER band in flattened discrimination space is centered around zero:

$$\text{Flattened discrimination variable} = \log_{10}\left(\frac{cS2_b}{cS1}\right) - \text{ER mean} \tag{4.2}$$

The flattening is, in this case, purely cosmetic because the 99.75 % ER discrimination line and the lower NR acceptance threshold at the 97 % NR quantile have been determined on the non-flattened discrimination variable. The only purpose is achieving an illustration of the ER and NR bands resembling the one in [34][4]. It is performed in two steps and utilizes the datapoints for the means from the 1D Gaussian fits of 1 pe ER band slices. In the first step the means are fitted with the sum of an exponential plus a second order polynomial

$$g_{\text{flat},1}(x) = e^{a+b\cdot x} + c\cdot x^2 + d\cdot x + e \tag{4.3}$$

The resulting fit function is subtracted from the $\log_{10}(cS2_b/cS1)$-values of all ER mean datapoints. The points after the first correction are then centered around zero by fitting

[4]The original flattening function has been tested, but did not yield satisfactory results.

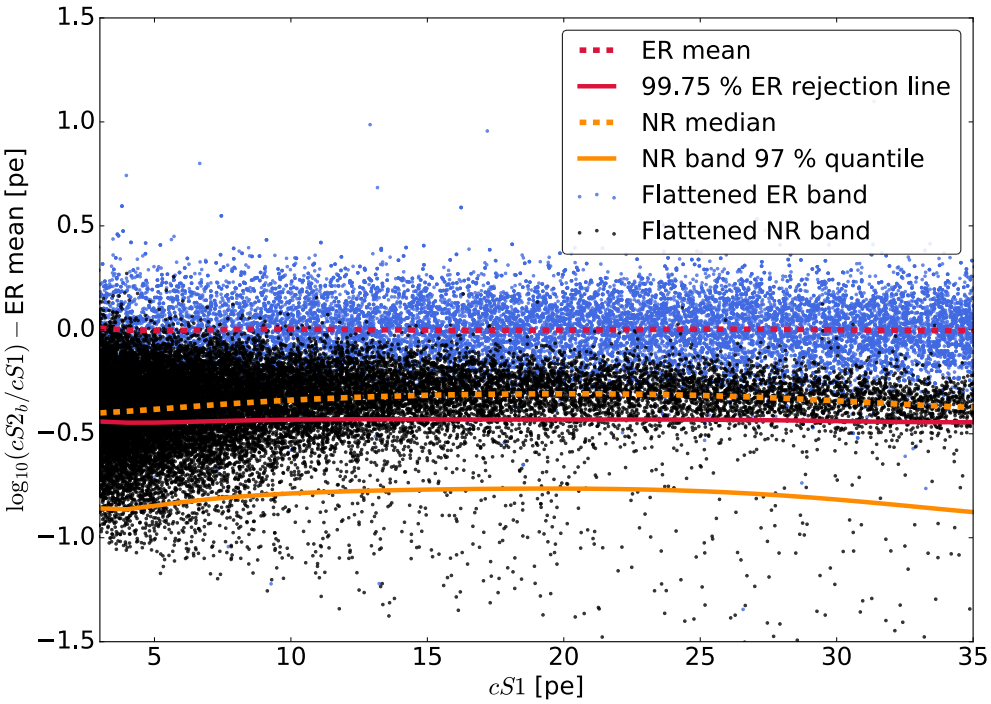

Figure 4.3: ER (blue) and NR (black) band datapoints in flattened discrimination space from $cS1 = 3\,\text{pe}$ to $35\,\text{pe}$. The ER band mean and 3.09σ line (red) are sketched as well as the NR band median and 97 % quantile (orange).

and subtracting a sixth order polynomial

$$g_{\text{flat},2}(x) = f \cdot x^6 + g \cdot x^5 + h \cdot x^4 + j \cdot x^3 + k \cdot x^2 + m \cdot x + n. \tag{4.4}$$

Plots of the two flattening steps can be found in section 6.9 in the appendix. The flattened bands in the energy region of interest are displayed in Fig. 4.3. As expected, the ER band is centered around zero. The ER discrimination line lies flat above the curved NR median at ≈ -0.45 in the flattened discrimination variable. The lower NR acceptance quantile exhibits a similar curve. Starting at ≈ -0.85 it reaches a maximum around $cS1 \approx 20\,\text{pe}$ at ≈ -0.75 followed by a gradual decrease.

4.1.2 Comparison of the Flattened Bands with the Benchmark Analysis

Several disparities can be found between the flattened ER and NR bands in this work and those from [34] which are illustrated in Fig. 2.7. Although they nominally constitute the same NR acceptance and ER rejection, the shape and location of the bands differ in the two analyses. The NR 97 % quantile in the benchmark analysis starts with ≈ -1.1 at 3 pe and rises up to ≈ -0.6 at 20 pe continuing in its upward tendency thereafter with an inflection point at around 30 pe. Above 15 pe the difference can be attributed to the problematic description of the line by the fit function mentioned earlier. Differently, the overall higher values of the line until $cS1 \approx 15\,\text{pe}$ indicate either a worse NR acceptance at low energies, a higher rejection of pathological events in the NR band, or a shift of points to higher values. This would have to be investigated by an evaluation of events on the level of waveforms between the two processors. Such a comparison could also show if events are missing from the band due to a low cut acceptance or if they are still present, but shifted

to higher values. An upward shift could for example be caused by smaller values of $cS1$ or larger values of $cS2_b$ attributed by PAX. However such an analysis would overstretch the scope of this work and is not pursued.
In the original analysis the 99.75 % ER rejection line lay flat at ≈ -0.4 in discrimination space. The lower position of the line in this work indicates a larger width of the ER band compared to the benchmark. There, the smaller width could be achieved with more suited cuts that reject ER events far away from the band. As only a small number of events can shift the position of the line significantly at the available statistics this is a likely explanation. If a 1σ band was available for the data in [34], this assumption could be investigated by checking the difference for intervals closer to the mean.
The wider band could also be a processor effect. However, one would expect a general shift of the band rather than a widening. If, for instance, $cS2_b$ is generally larger due to the clustering algorithm, the band is moved upward which is corrected by the band flattening. A widening would on the other hand signify an inconsistent clustering that generates peaks with both larger and smaller areas compared to the old data processor. Thus, points are shifted away from the band mean. A third explanation comes from the fact that only a fraction of the total ^{60}Co and ^{232}Th data has been used owed to memory constraints at the analysis facility. It would have to be investigated if the ratio of datapoints from both sources changes the shape of the ER band.

4.1.3 Nuclear Recoil Acceptance of the 99.75 % Discrimination Line

The NR acceptance of the 99.75 % ER discrimination cut $\mathcal{A}_{\mathrm{nr}}$ is computed as the ratio between NR events below the discrimination line and all NR events. The resulting NR acceptance is displayed in Fig. 4.4 together with the one from the benchmark analysis [34]. The acceptance of the discrimination cut is lower over the whole range. Above 5 pe it reaches values between 60 % and 80 % of the benchmark NR acceptance and resembles its decrease with larger $cS1$, albeit shallower. This can be explained by the downward shift of the discrimination line from ≈ 0.4 in discrimination space down to ≈ 0.45. If the NR band is largely unchanged between the two analyses more events in the are rejected.
Below 5 pe where most NR events are situated (cf. Fig. 4.2) a steep rise is apparent. This can be seen in context with the higher combined cut acceptance in this energy region. If more events have been removed from the NR band by cuts in the benchmark analysis, there are less events left to be removed by the discrimination. This is an indicator that PAX's capability for finding S1 peaks with $cS1 < 5$ pe is actually higher than with the old data processor. As the events containing these peaks pass all cuts one can also deduce that the higher combined cut acceptance at low energies cannot be solely attributed to an artificial acceptance boost induced by PAX's twofold PMT coincidence requirement.
The statistical approach applied 4.3 does not require a bin-wise NR acceptance in $cS1$, but the total one in the dark matter search interval between 3 pe and 20 pe. This is found to be $\mathcal{A}_{\mathrm{nr}} = (31.1 \pm 0.3)\,\%$ in the same fashion as above. It should be noted that the profile likelihood analysis in the original work did not use such a fixed discrimination, but signal and background models. The fixed discrimination was only applied in a maximum gap cross-check [34].

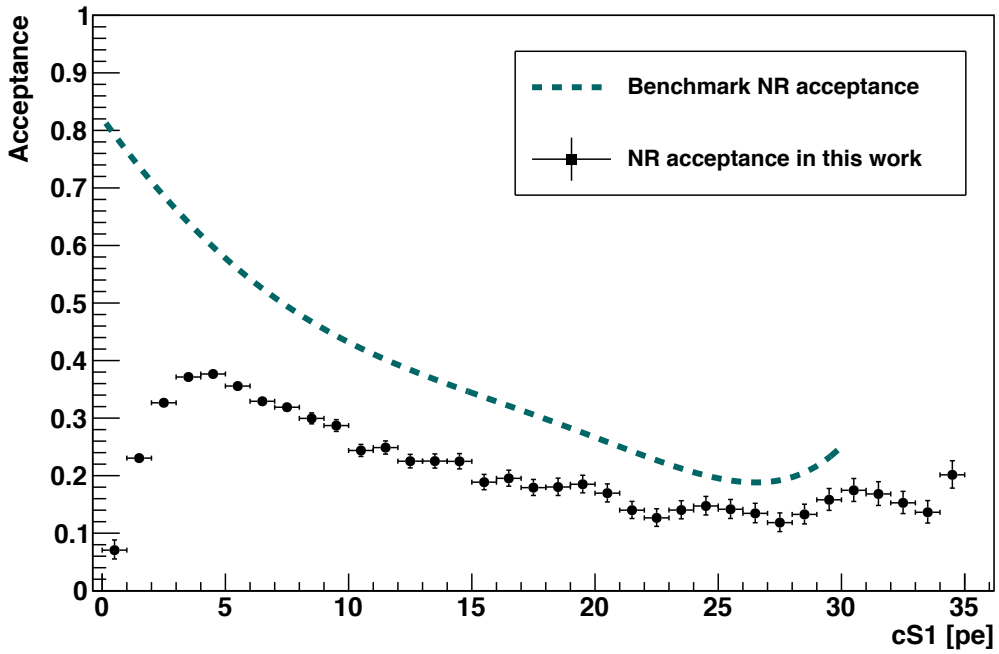

Figure 4.4: NR acceptance of the 99.75 % ER discrimination cut vs $cS1$ from 3 pe to 35 pe. The acceptance from the benchmark analysis is given as the dashed teal line up to 30 pe [34].

4.2 Analysis of Signal Candidate Events

The boundaries of the dark matter search region are defined by the upper and lower $cS1$ thresholds of the benchmark ROI from [34], namely 3 pe and 20 pe, the 97 % quantile of the NR band, and the 99.75 % ER rejection line. All legacy cuts are applied on the original dataset containing 224.6 live days of dark matter data. The remaining data points are depicted in flattened discrimination space in Fig. 4.5 with the ROI indicated in yellow.
In total 26 events appear below the ER discrimination line for $cS1 < 30$ pe and 11 of these events are situated in the ROI. All of these events either appear below 10 pe or above 25 pe. 8 events are below the lower $cS1$ threshold while 2 are situated above the higher ROI threshold. The waveforms of all 26 events are examined in order to discover if they constitute valid signals or not.

Removal of Noise Events

Eleven events below the discrimination line originate from the same data file recorded on 13th April 2011 at 11.17 a.m. Of these 11 events 8 are located inside the ROI. The file has not been marked as noisy and consequently not been excluded from the analysis. Yet, all remaining events originating from it exhibit high peak counts. An example waveform highlighting the common features is shown in Fig. 4.6.
In this so-called event display the panels in the top row show, from left to right, the main S1 amplitude vs time, the main S2 amplitude vs time, the S1 hit pattern on the bottom PMT array, and the S2 hit pattern on the top PMT array. The contributions of the individual PMTs to the peak areas are color-coded. The middle panel illustrates the summed waveform. Peaks that are shaded grey originate from hits in suspicious PMT

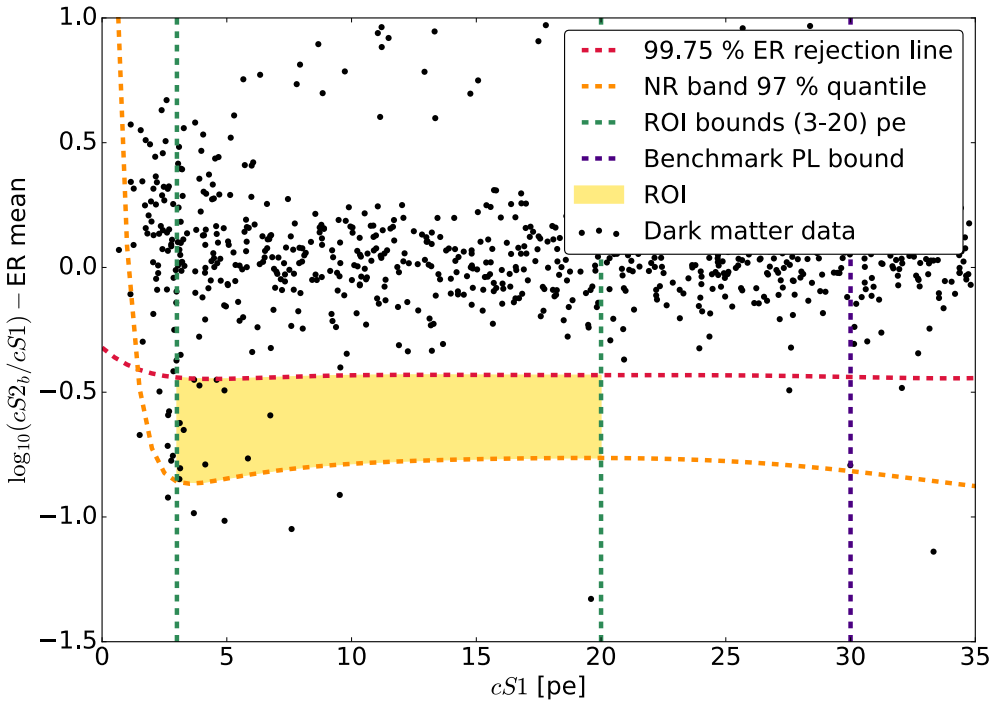

Figure 4.5: Dark matter data from 224.6 live days in flattened discrimination space vs $cS1$ from 0 pe to 30 pe. The dark matter search region between 3 pe and 20 pe is indicated in yellow. The 97 % quantile of the NR band, and the 99.75 % ER rejection are displayed as orange and red lines. Below the lower ROI threshold (dashed green line) the band flattening is not valid. The upper bound of the profile-likelihood analysis from [34] is indicated as the dashed purple line.

channels. The hits are shown channel-wise for top, bottom and veto PMTs in the base panel. Red points are hits contributing to peaks while green points are isolated hits causing the respective channels to be marked as suspicious. The size of the points is scaled to the area of the hit.

All these events feature a valid S2 signal, visible in the middle of the summed waveform, which is paired with one of multiple S1 signals. In addition several S2s and unknowns are present in the summed waveform, so the pairing of the main interaction is not reliable. In addition multiple channels exhibit lone hits and are marked as suspicious. It is, therefore, likely that such events fall into a period of increased noise. Thus, they should have been removed by the signal to noise cut which compares the total area of all peaks with the summed area of the main S1+S2 pair. As the cut has been optimized for the old processor's peak definitions, insufficient tuning explains their keeping. All events from this file and two further events from other files[5] are removed manually on this basis. This leaves a total of 13 signal candidate events below the discrimination line with 2 of these in the ROI. Candidates from now on denote events with $cS1 < 30$ pe which are located below the discrimination line and inside the fiducial volume.

The event distribution in the TPC after the removal of noisy events is shown in Fig. 4.7. Black events are located below the ER discrimination line while gray points are located above. The blue points are candidate events outside the ROI and the red stars are the signal candidates inside. One can see that they are spread throughout the fiducial volume

[5]The waveforms of the removed noisy events can be found in the appendix, section 6.10.

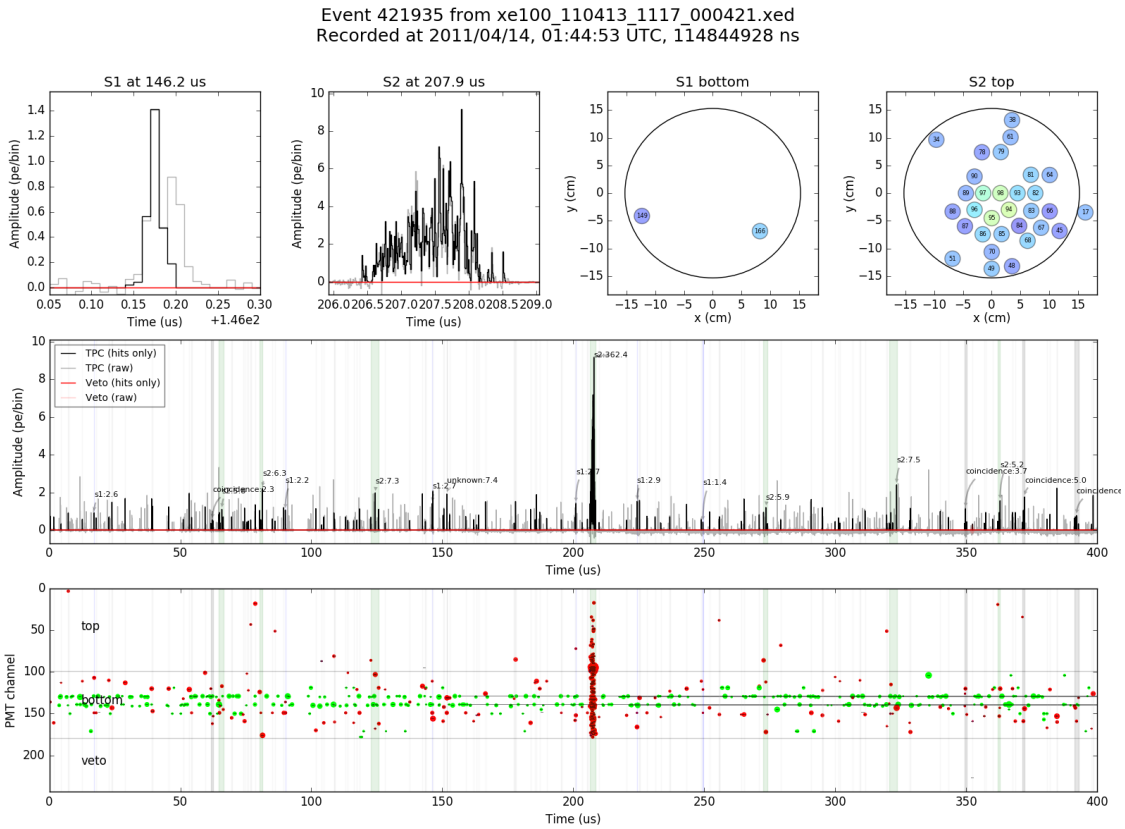

Figure 4.6: Waveform of noise event 421935 from the data file recorded on 13th April 2011 at 11.17 a.m. showing the common features of all remaining events from this file after cuts and discrimination.

with one non-ROI event being located at the very edge. This event could have been removed in the original analysis due to the uncertainty of the position reconstruction. Several other features are visible such as leakage outside the TPC borders to the right and bottom as well as the light gray band of the gas gap close to $z = 0$. At the bottom outer edge of the TPC events are shifted inwards. This is a feature which has not been observed in [34]. It is unclear wether this is a position reconstruction artifact or due to charge accumulation on the walls which would alter the drift path of electrons [23].

Events Below the ROI Threshold

Events below the lower ROI threshold of 3 pe are not considered as valid dark matter events. However, they are examined for common features because no events have been observed here in the original analysis [34]. In summary these events show a mixed picture. Representative waveforms can be found in section 6.10 of the appendix. Three events have multiple S2 peaks and were not removed because the secondary and tertiary S2s are too small, i.e. $s2 < 70$ pe, to be considered by the double scatter cut. One event has three S1s and is still passed by the single S1 cut. This can be understood taking a closer look at the cut. Events are removed if they have more than S1 signal that is likely to be of physical origin. This means, for example, that the secondary S1 must have at least a twofold coincidence before the cut removes the event. If a noisy PMT channel, which is a PMT channel that is known to produce noise pulses regardless of being marked as

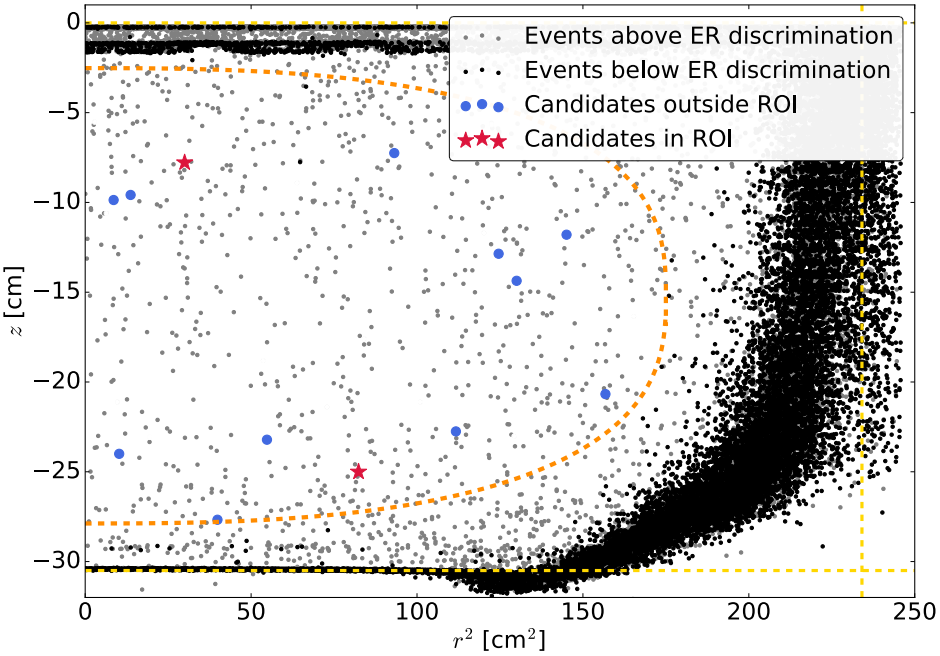

Figure 4.7: Spatial distribution of dark matter data from 224.6 live days in the TPC after the exclusion of noisy events based on their waveforms. Gray points constitute events above the ER discrimination line, black events are situated below. Blue events are candidate events, i.e. events below the discrimination line and inside the fiducial volume, but outside the ROI. The red events are signal candidates inside the dark matter search ROI. The dashed yellow lines indicates the TPC borders while the fiducial volume is outlined in orange.

suspicious by PAX, contributes to the S1, the required coincidence is bumped up to three. This is the case here. One secondary S1 has been seen by two PMTs, one of them being a known noisy PMT. Accordingly, it is not removed by the cut. The other additional S1 precedes the S2 by more than 185 μs which also prompts the single S1 cut to ignore it. In addition, the S1 has not been seen by the bottom PMT array which would have prompted its removal by the S1 PMT pattern likelihood cut. This is also true for the last event below the ROI threshold leaving no valid events here.

Remaining Events Outside the ROI

Of the two events above the upper bound of the ROI one has a small secondary S2[6]. The other event exhibits a single S1 and S2 pair in its waveform shown in Fig. 4.8. It has not been present in the original XENON100 analysis although it is likely to be a valid signal with a triple coincidence S1. Still, it might have been removed by the PMT pattern likelihood because the S1 with an uncorrected area of 37 pe has not been seen by any of the top PMTs.

4 events remain inside the $cS1$ bounds of the ROI, but below the NR acceptance threshold. One of these shows periodic noise and a second small S2. Apparent pathologies in the three other events are periodic noise in the veto channels, large isolated pulses succeeding the main S2, and S1 signals not registered by the bottom PMTs. Event displays can be found in section 6.10 of the appendix.

[6]Waveform in section 6.10 of the appendix.

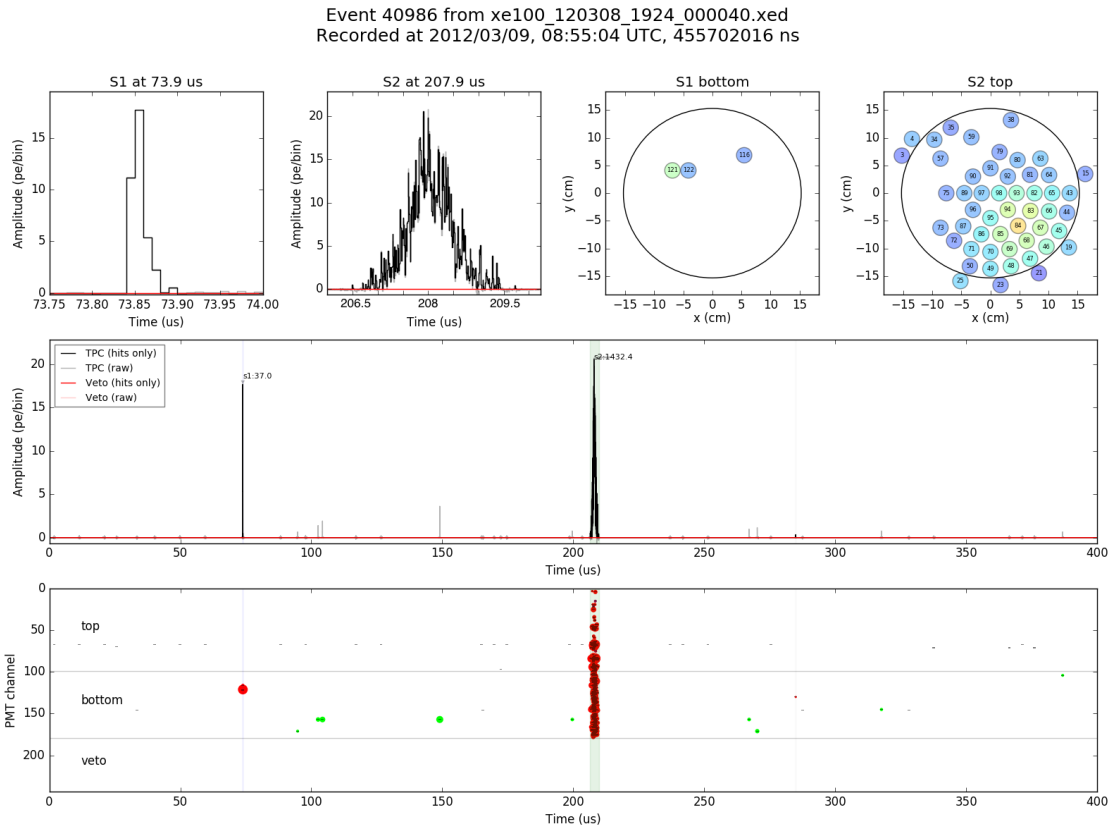

Figure 4.8: Waveform of event 40986 from the data file recorded on 8th March 2012 at 7.24 p.m. The event is above the upper ROI threshold, but inside the $cS1$ region of the profile-likelihood analysis in [34]. It has not been present in that analysis although it has all features of a valid signal event [34]. One caveat is the missing registration of the S1 by the top PMT array.

Signal Candidates in the ROI

After the removal of noise events two signal candidates remain in the dark matter search region. Their waveforms are presented in Fig. 4.9 and Fig. 4.10. The first candidate exhibits periodic noise and three unidentified peaks. Still, these do not warrant the rejection of the event, as they can be isolated from the dual coincidence S1. An additional caveat is that the S1 has not been seen by the bottom PMT array. Looking at the drift time of $\approx 50\,\mu$s one can deduce that the event occured in the upper third of the TPC which makes a non-detection of the small S1 by the bottom PMT array plausible. The candidate is marked as a signal event accordingly.

The second remaining event in the ROI is one of the two original signal events in [34], namely event 23859 from the file recorded on 23rd October 2011 at 11.01 a.m. Its properties are now directly compared between the PAX reconstruction in this work and the original data processor. The original PMT hit patterns as well as the summed waveforms are displayed in Fig. 4.11.

Regarding the S1 signal, the hit patterns of the S1 signal on the bottom PMT array are equal and no PMT hits have been registered on the top array. The uncorrected peak areas are $s1 = 4.6\,$pe for PAX and $s1 = 4.5\,$pe for the old data processor. The larger S1 area in PAX can be attributed to the dynamic threshold setting in PAX as well as negative area contributions in the old data processor. A double peak structure of the S1 is clearly visible

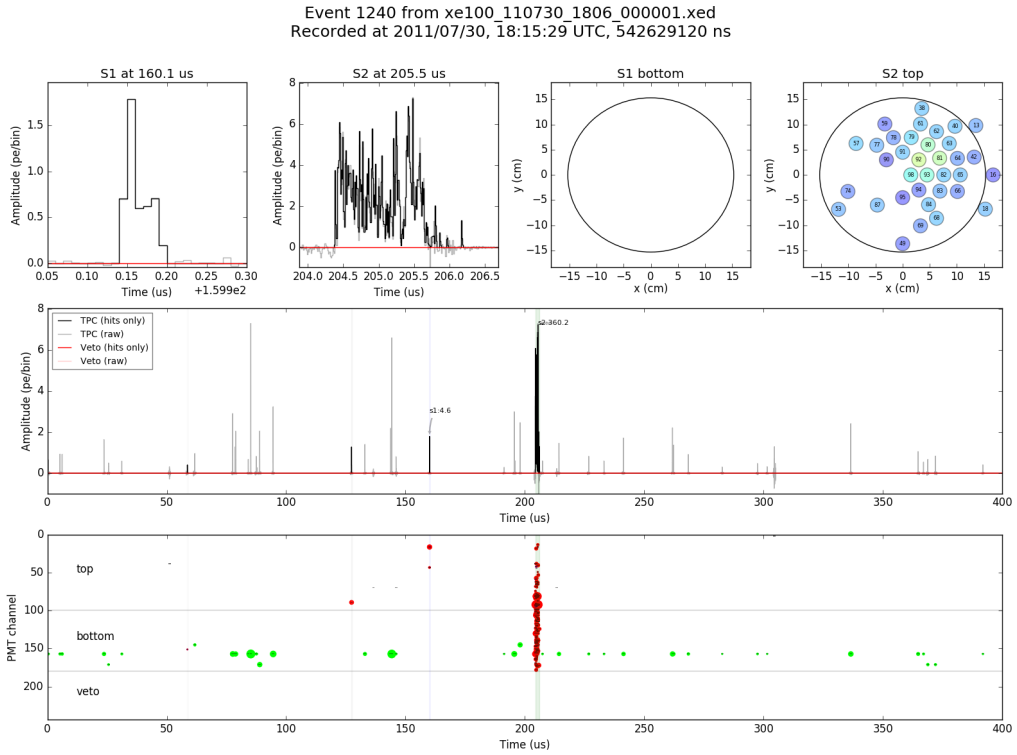

Figure 4.9: First candidate event in the ROI (event 1240 from the data file recorded on 30th July 2011 at 6.06 p.m.). The event shows large lone hits and the main S1 was not seen by the bottom PMT array. It has not been present in [34].

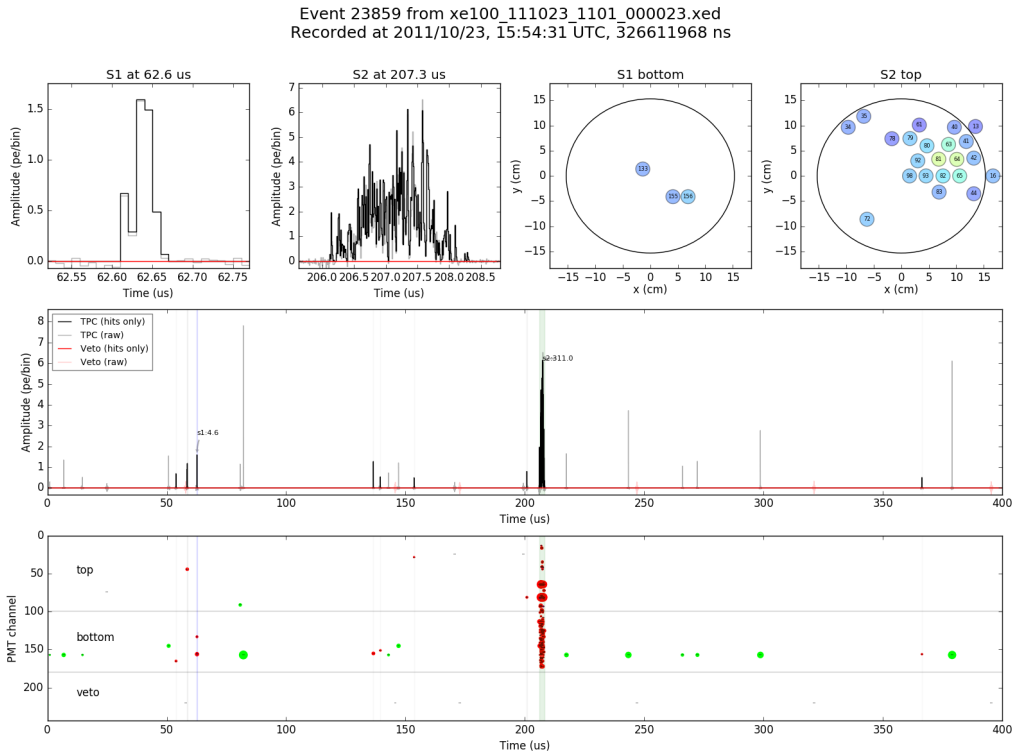

Figure 4.10: Second candidate event in the ROI (event 23859 from the data file recorded on 23rd October 2011 at 11.01 a.m.). The event exhibits lone pulses resulting in three unclassified peaks, but the main S1 features a triple coincidence on the bottom PMT array. This is one of the two signal events in [34].

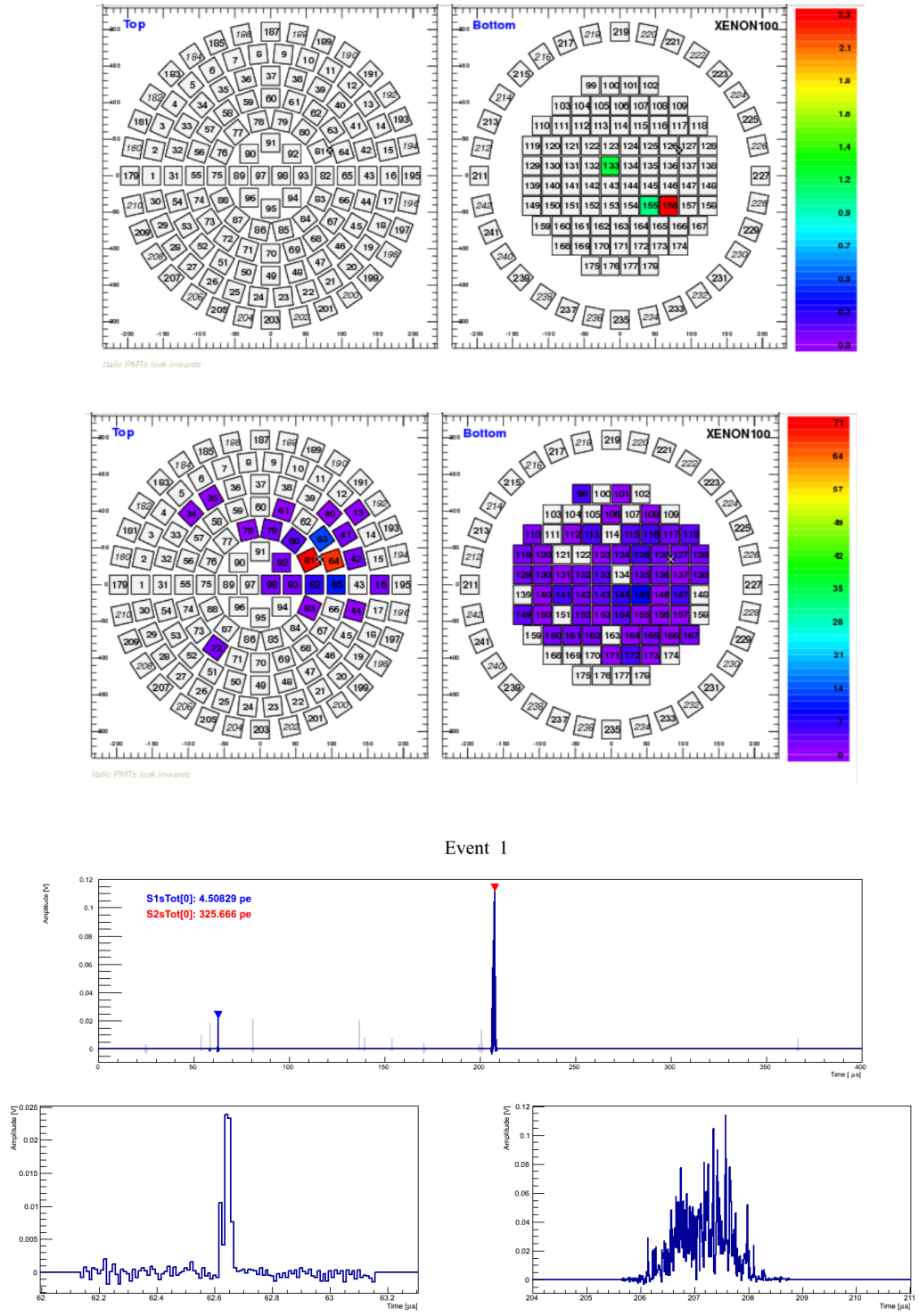

Figure 4.11: Overview of event 23859 processed with the old data processor. From top to bottom: S1 hit patterns on both PMT arrays, S2 hit patterns, summed waveform, magnified S1 waveform (left) and magnified S2 waveform (right) [34, 48].

in both waveforms.

Comparing the S2 waveforms in both processors the hit patterns on the top PMT array are indistinguishable. Yet, the uncorrected area of the S2 signal in the old processor is larger than the one from PAX with $s2 = 325.7\,\text{pe}$ vs $s2 = 311.0\,\text{pe}$. This can be explained by the smaller peak width in PAX. Here contributions to the peak area emerge shortly after

206.0 μs and cease before 208.5 μs. In the old processor hits are present before and after this window illustrating the effect of the PAX peak finder and clustering[7]. The resulting peak area deviation is $\approx 4\,\%$ between processors.
The reconstructed position of the event with the old data processor is $x^2 + y^2 = r^2 = 93.16\,\mathrm{cm}^2$ and $z = -24.7\,\mathrm{cm}$ which disagrees with the reconstructed position from PAX. Here one finds $r^2 = 82.38\,\mathrm{cm}^2$ and $z = -25.0\,\mathrm{cm}$. The disagreement exceeds the position resolution with $\Delta z \approx 0.03\,\mathrm{cm}$ and $\Delta r^2 < 0.09\,\mathrm{cm}^2$. The radial inward shift corresponds to the general tendency for events at the bottom of the TPC visible in Fig. 4.7. Addressing the deviation in the longitudinal coordinate, one has to consider that the value of the electron drift velocity is hard-coded into the PAX configuration file and not determined from data. If the drift velocity is assumed too high, events are reconstructed deeper in z than they actually are. This can also explain the leakage of events below $z = -30.5\,\mathrm{cm}$ visible in Fig. 4.7. This should be further investigated as PAX's position reconstruction evolves with XENON1T.
Concluding the discussion of the one remaining signal event in the ROI it has to be acknowledged that event 23859 has been discarded in more recent analyses. It has been found that isolated bursts of S1 signals occur repeatedly within the dataset which can mimic dark matter signals if paired with a lone random S2. The remaining signal from this analysis falls close to one such period [22]. Taking the knowledge at the time of the original publication as a basis, the event is kept as a signal candidate though.

Removal of an Original Signal Candidate

One of the original signal events that were left after cuts, discrimination and waveform examination in [34] is missing in this analysis. Its original waveform from the old data processor is detailed in Fig. 4.12 while the PAX event display is shown in Fig. 4.13.
Both events are in agreement with respect to the PMT hit patterns. The areas $s1$ as well as $s2$ are $\approx 11\,\%$ and $\approx 4\,\%$ smaller respectively than their old processor counterparts. The original zoomed waveform of the S1 exhibits large fluctuations around the baseline which are not present in PAX. The original S2 appears elongated compared to the PAX one. Again, a disparity of the reconstructed positions is visible with $r^2 = 137.8\,\mathrm{cm}^2$ and $z = -16.6\,\mathrm{cm}$ for the old processor and $r^2 = 131.4\,\mathrm{cm}^2$ and $z = -16.8\,\mathrm{cm}$ for PAX. The downward and inward shift exceeding the uncertainty of the position reconstruction recur, albeit to a smaller degree.
The cut that removed this original signal candidate here is the S2 asymmetry cut which compares the area contributions of the top and bottom PMT arrays to the S2 signal. Its main purpose is the removal of events with interactions in the gas phase. Fig. 4.14 shows the AmBe neutron event distribution the asymmetry cut was defined on with the cut boundaries indicated as black dashed lines. The discarded dark matter candidate is located close to the upper boundary of the cut. Given that the hit patterns are similar, one can deduce that in the original analysis this event lay just inside the cut boundary where a lone aberrant PMT contribution could lead to a deviating result. As the point is shifted upward, either one bottom PMT channel contributed less or one top PMT channel

[7] Of course no general trend with respect to peak areas can be derived from a single event. These rather manifest themselves in the definition of the ER and NR bands.

contributed more. Due to the shorter duration of the S2 in PAX this explanation is the most likely one. Therefore, the threshold setting and peak clustering of PAX caused the exclusion of this event.

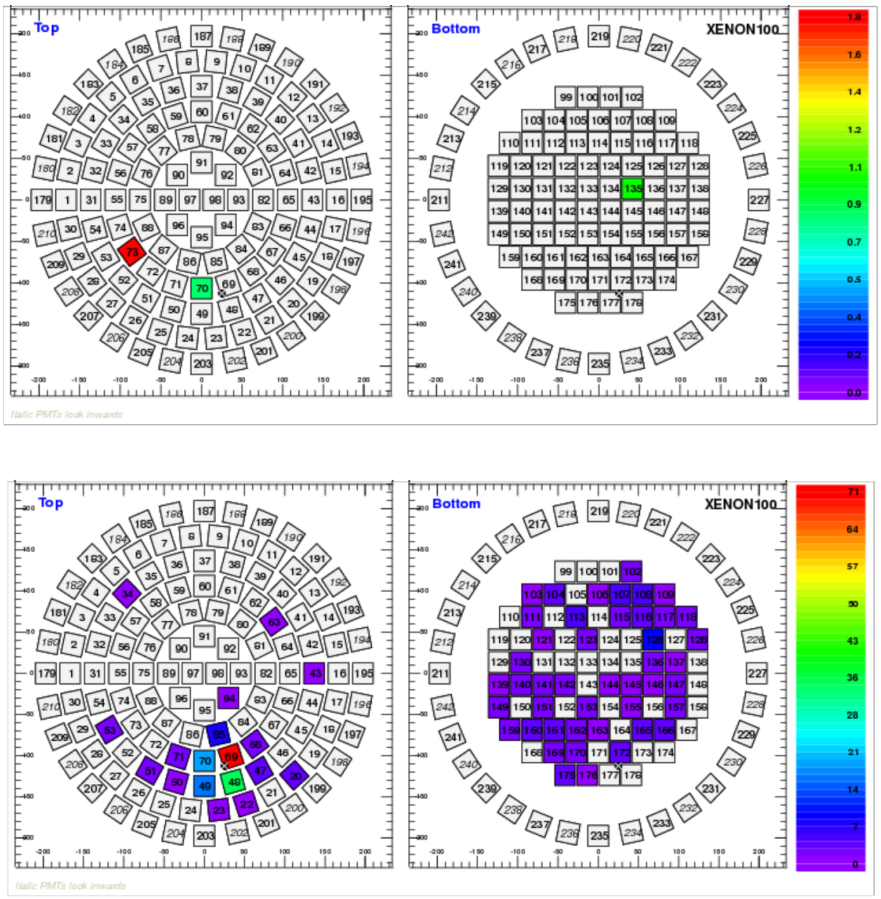

Figure 4.12: Overview of event 40253 processed with the old data processor. From top to bottom: S1 hit patterns on both PMT arrays, S2 hit patterns, summed waveform, magnified S1 waveform (left) and magnified S2 waveform (right) [34, 48].

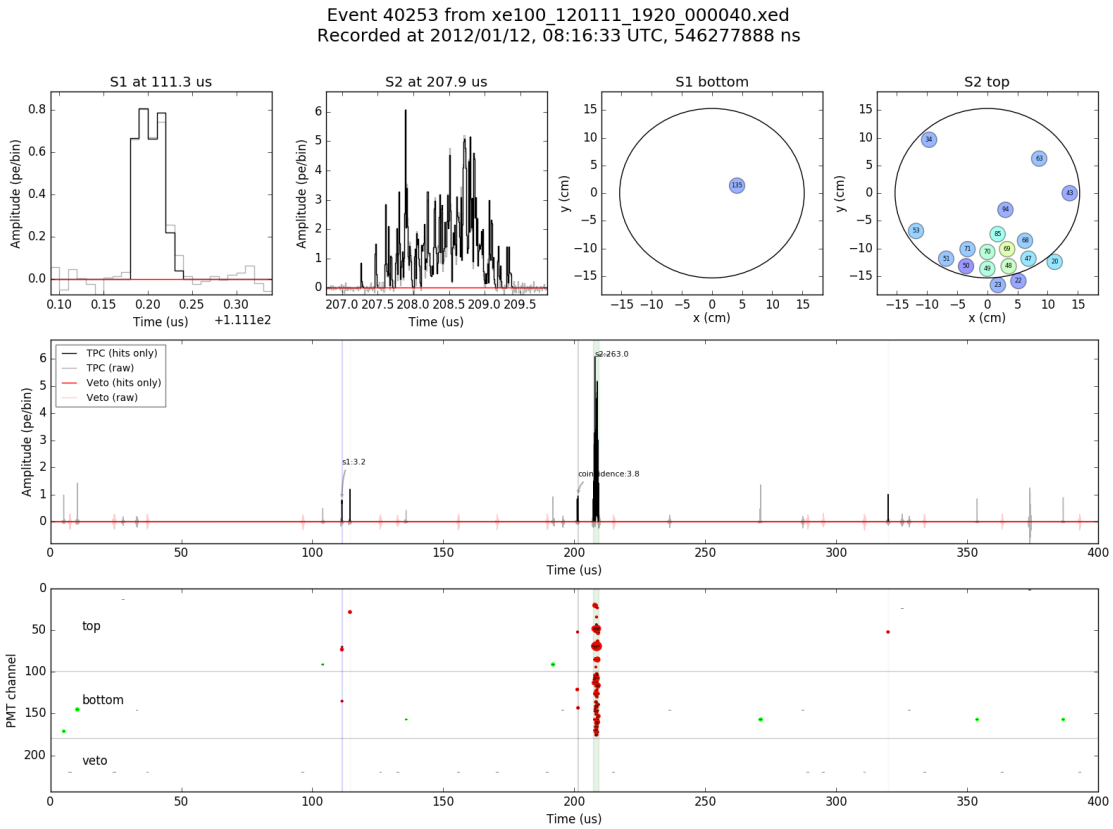

Figure 4.13: PAX event display of the second candidate event from the original analysis (event 40253 from the file recorded on 11th January 2012 at 7.20 p.m.) [34].

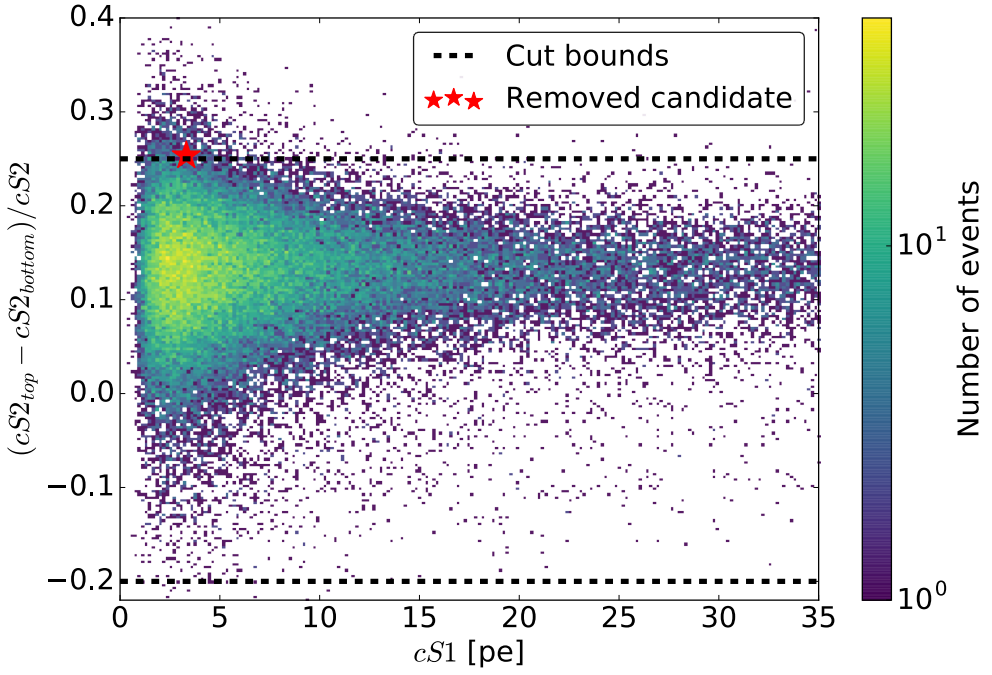

Figure 4.14: Location of the discarded original dark matter candidate (red star) in the parameter space of the S2 asymmetry cut. The dashed black lines signify the upper and lower cut boundaries. The event distribution in the background are AmBe neutron events the cut was defined on.

4.3 Upper Limit on the WIMP-Nucleon Cross Section

Having established the existence of two signal events inside the dark matter search region an upper limit on the WIMP nucleon cross section $\sigma^{\mathrm{SI}}_{\mathrm{nucleon}}$ can be calculated using Eq. 1.7 [7, 54]. While the most important steps are outlined here, detailed calculations can be found in section 6.11 of the appendix. After the calculation of the velocity integral one obtains the differential dark matter interaction rate per unit mass inside the detector:

$$\frac{\mathrm{d}R}{\mathrm{d}E} = \frac{\rho_0 \cdot \sigma_0 \cdot F^2(q^2)}{2 m_\chi \mu_A^2} \cdot \left\langle \frac{1}{v} \right\rangle \tag{4.5}$$

with

$$\left\langle \frac{1}{v} \right\rangle = \frac{2}{\sqrt{\pi} v_c} \cdot \left(e^{-\frac{E \cdot m_A}{2\mu_A^2 v_c^2}} - e^{-\frac{v_{\mathrm{esc}}^2}{v_c^2}} \right). \tag{4.6}$$

In these equations σ_0 is the WIMP-nucleus interaction cross section, $\rho_0 = 0.3\,\mathrm{GeV/c^2}$ the local dark matter density, σ_0 the spin-independent interaction cross section for WIMPs with the entire nucleus, $F^2(q^2)$ the nuclear form factor, $m_{\chi,A}$ the masses of the WIMP and target nucleus, and μ_A their reduced mass. The velocities are given by astrophysical constraints, namely the escape velocity of our galaxy v_{esc} and the velocity of the sun around the galactic center v_c [7, 54]:

$$\begin{aligned} v_{\mathrm{esc}} &\approx 544\,\mathrm{km/s}, \\ v_c &\approx 220\,\mathrm{km/s}. \end{aligned} \tag{4.7}$$

The motion of the earth around the sun is neglected in the following because its influence on the limit is minuscule. Up to this point all equations assumed natural units with $\hbar = c = 1$, but from now on these factors are included to facilitate the recalculation of limits[8].
In the next step the differential rate is integrated over the dark matter ROI boundaries from $E_{\mathrm{min}} = 6.7\,\mathrm{keV}$ to $E_{\mathrm{max}} = 30.9\,\mathrm{keV}$. The nuclear form factor $F^2(q^2)$ is energy-dependent and approximated by a simple spherical model [16]:

$$F^2(q^2) = e^{-\frac{r^2 q^2}{6\hbar^2 c^2}}. \tag{4.8}$$

The nuclear radius r and the non-relativistic momentum transfer q^2 are given by:

$$\begin{aligned} r &= 1.2\,\mathrm{fm} \cdot A^{\frac{1}{3}}, \\ q^2 &= 2 m_A \cdot E \end{aligned} \tag{4.9}$$

After inserting the factors into Eq. 4.5 an integration and rearrangement yields

$$\sigma_0 = \frac{\sqrt{\pi} R \cdot \mu_A^2 m_\chi c^2 v_c}{\rho_0} \cdot \left(\frac{e^{-a \cdot E_{\mathrm{min}}} - e^{-a \cdot E_{\mathrm{max}}}}{a} + \frac{e^{-b \cdot E_{\mathrm{max}} - c} - e^{-b \cdot E_{\mathrm{min}} - c}}{b} \right)^{-1} \tag{4.10}$$

[8] First and foremost this choice is intended to help readers with their own limit calculations.

with the mass-dependent factor a and the constants b and c:

$$\begin{aligned} a = & \frac{m_A \cdot r^2}{3\hbar^2 c^2} + \frac{m_A c^2}{2\mu_A^2 v_c^2}, \\ b = & \frac{m_A \cdot r^2}{3\hbar^2 c^2}, \\ c = & \frac{v_{\mathrm{esc}}^2}{v_c^2}. \end{aligned} \tag{4.11}$$

The event rate R and number of events in the search window are connected by the acceptance-corrected exposure $(T \cdot M)_{\mathrm{corr}}$ which gives

$$R = \frac{N}{(T \cdot M)_{\mathrm{corr}}} = \frac{N}{T \cdot M \cdot \mathcal{A}_{\mathrm{comb}} \cdot \mathcal{A}_{\mathrm{nr}}} \tag{4.12}$$

with the live time, target mass and acceptances

$$\begin{aligned} T = & \quad 224.6\,\mathrm{d}, \\ M = & \quad 34\,\mathrm{kg}, \\ \mathcal{A}_{\mathrm{comb}} = & \quad (75.2 \pm 0.3)\,\%, \\ \mathcal{A}_{\mathrm{nr}} = & \quad (31.1 \pm 0.3)\,\%. \end{aligned} \tag{4.13}$$

N is derived utilizing frequentist confidence intervals at 90 % confidence level. The statistical approach in this work is based on the method devised by Feldman and Cousins [18]. The number of observed events

$$N_{\mathrm{obs}} = 2\,\mathrm{events} \tag{4.14}$$

and the background expectation in the ROI are the inputs here. One retrieves the values of N at the confidence interval boundaries either from a lookup table or the ROOT `TFeldmanCousins` class [43]. Based on the inputs single-sided (upper limit) or double-sided (detection claim) intervals are chosen automatically. Taking the original background expectation [34] of

$$N_{\mathrm{bg}} = (1.0 \pm 0.2)\,\mathrm{events} \tag{4.15}$$

a single sided interval with an upper limit of

$$N = 4.91\,\mathrm{events} \tag{4.16}$$

is the result. The acceptance-corrected event rate then reads

$$R = \frac{4.91\,\mathrm{events}}{224.6\,\mathrm{d} \cdot 34\,\mathrm{kg} \cdot 0.752 \cdot 0.311} \approx 2.7 \cdot 10^{-3}\,\frac{\mathrm{events}}{\mathrm{kg} \cdot \mathrm{d}}. \tag{4.17}$$

The quantity which is usually given in the spin-independent limit plots is the WIMP-nucleon cross section because it allows to compare results from detectors with different target nuclei. If the couplings of protons and neutrons to WIMPs are equal, with the

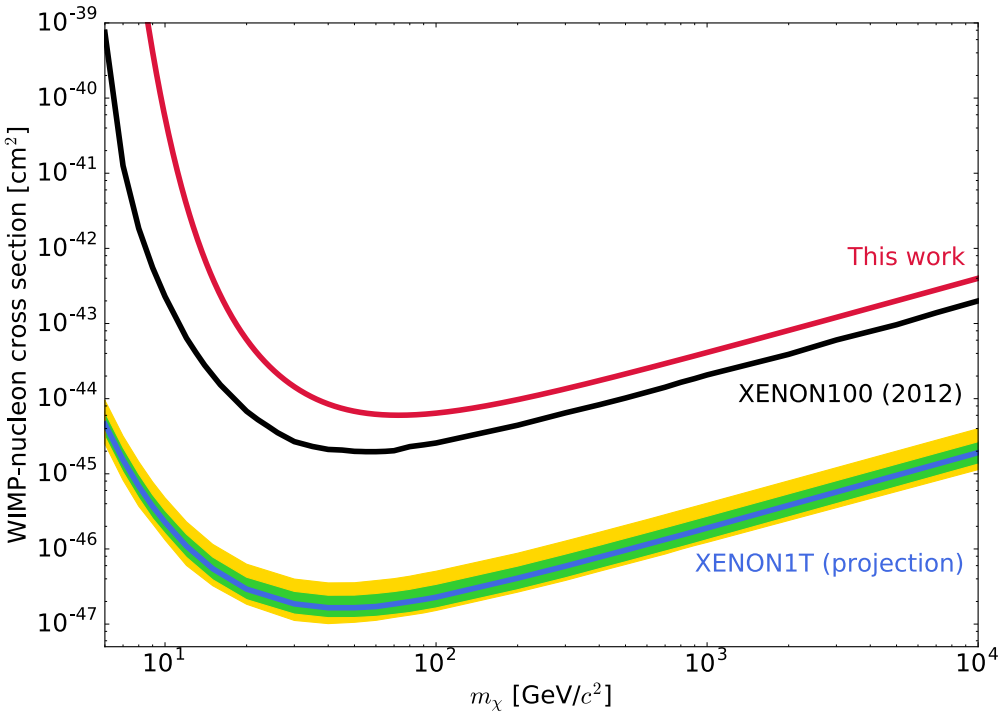

Figure 4.15: Upper limits of the WIMP-nucleon interaction cross section from this work (red), the benchmark profile-likelihood analysis (black) [34] and the XENON1T sensitivity projection after an exposure of $2\,\mathrm{t}\cdot\mathrm{y}$ (blue) [55]. 1σ (green) and 2σ (gold) uncertainty bands are given for the XENON1T sensitivity prediction.

nucleon mass m_N^2 and the mass number A of the target isotope [54] one can write

$$\sigma^{\mathrm{SI}}_{\mathrm{nucleon}} = \frac{\sigma_0 \cdot m_N^2}{\mu_A^2 \cdot A^2}. \tag{4.18}$$

Fig. 4.15 shows the resulting exclusion curve for $\sigma^{\mathrm{SI}}_{\mathrm{nucleon}}$ as a solid red line. The original limit from [34] and the XENON1T sensitivity prediction with 1σ und 2σ uncertainty bands are displayed as black and blue lines respectively. It is evident that the result in this work is on the same order of magnitude as the original exclusion curve, but its sensitivity cannot be matched. This can be attributed to two factors. First of all the combined cut and NR discrimination acceptances are lower in total because the cuts have not been tuned for PAX. The result is that the exposure is corrected to lower values which shifts the limit upwards (c.f. Fig. 1.8). In the original analysis the product of NR discrimination (c.f. Fig. 3.2) and combined cut acceptance (c.f. Fig. 4.4) was roughly

$$0.45 \cdot 0.80 \approx 0.36 \tag{4.19}$$

whereas in this work it was

$$0.31 \cdot 0.75 \approx 0.23 \tag{4.20}$$

which amounts to a factor of ≈ 1.5 reduced exposure compared to the original analysis. At WIMP masses above $100\,\mathrm{GeV/c^2}$ where the behaviour of the curve is dominated by the local dark matter density ρ_0 one finds a factor ≈ 2 deviation. The additional offset in this energy range can be explained by the different statistical methods applied for setting the

limit [30, 34, 56]. The original profile-likelihood analysis reaches a higher sensitivity, as it extracts more information from the data applying sophisticated signal and background models. The event counting in a single pre-defined box applied in this work does not use information on the signal shape. This could be addressed by a bin-wise analysis or the application of a maximum-gap approach [19]. However, the effect of the chosen statistical method is minor compared to cut optimization for an enhanced acceptance.
It should also be noted that the effect of the nuclear form factor $F^2(q^2)$ is largest at high WIMP masses. Even though only a simple spherical approximation has been put to use in the calculation of the limit the curves match remarkably, albeit with the aforementioned deviation in sensitivity.
At lower masses there is an apparent difference in behaviour of the exclusion curves. The whole curve in this work is shifted to the right compared to the original. As a result, the best upper limit on the WIMP-nucleon cross section achieved in this work is

$$\sigma^{\mathrm{SI}}_{\mathrm{nucleon}} = 6 \cdot 10^{-45}\,\mathrm{cm}^2 \text{ for } m_\chi = 70\,\mathrm{GeV} \tag{4.21}$$

whereas it is $\sigma^{\mathrm{SI}}_{\mathrm{nucleon}} = 2 \cdot 10^{-45}\,\mathrm{cm}^2$ for $m_\chi = 55\,\mathrm{GeV}$ in the benchmark analysis [34]. Going to lower masses the relative deviation of the two curves increases. The steep rise around the energy threshold is reached faster exceeding $10^{-40}\,\mathrm{cm}^2$ at $m_\chi = 10\,\mathrm{GeV}$ which is three orders of magnitude above the original curve.
A different threshold behaviour is expected due to the slightly lower light yield found in this work. Yet, the energy threshold is moved upwards by only 0.1 keV which does not warrant the large disparity. This can be explained by the fact that the large Poisson fluctuations

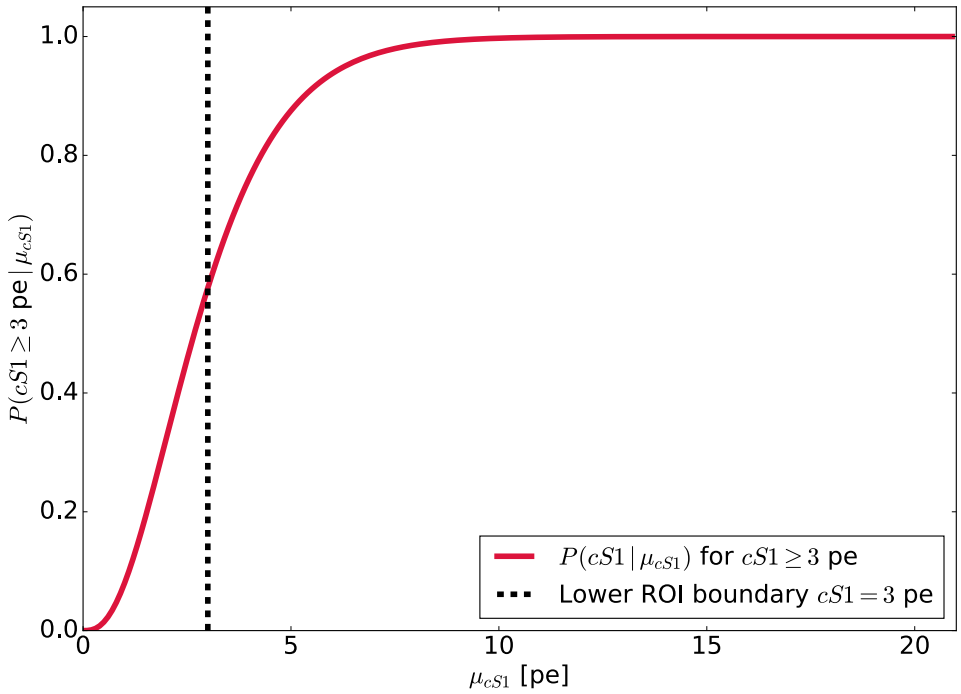

Figure 4.16: Illustration of Poisson smearing. The probability for $cS1 \geq 3$ at an expectation value of μ_{cS1} can be calculated using a Poisson distribution and has both a non-zero value below the $cS1$ threshold of the ROI and approaches unity above.

the S1 signal is subject to have not been taken into account. After the incorporation of NR and cut acceptances into the interaction rate model, the analysis in this work assumes that the lower ROI threshold is a step function in $cS1$. However this is not true. At an expectation value of μ_{cS1} the Poisson probability of generating an S1 signal above threshold $P(cS1 \geq 3\,\mathrm{pe}\,|\,\mu_{cS1})$ is greater than zero [30]. This is illustrated in Fig. 4.16. It indicates that even at a fixed lower ROI threshold the acceptance for S1 signals with $cS1 \leq 3\,\mathrm{pe}$ does not vanish. This changes the threshold which has to be smeared by a function. In the benchmark analysis Poisson fluctuations in S1 are considered by the so-called Poisson smearing [30, 34]. Pursuit of the same approach would require the recalculation of the limit, so only a rough assessment of the effect is performed. Looking at Fig. 4.16 one finds that $P(cS1 \geq 3\,\mathrm{pe}\,|\,\mu_{cS1})$ has fallen by half at $\mu_{cS1} = 1.5\,\mathrm{pe}$ compared to the value at $\mu_{cS1} = 3\,\mathrm{pe}$. This is taken as an approximation of the actual ROI threshold and translates to $E_{\mathrm{min}} = 4.1\,\mathrm{keV_{nr}}$.

The resulting upper limit is shown as the green curve in Fig. 4.17. Due to the larger search window in $cS1$ the curve is slightly shifted downwards as a whole and the shapes of both curves are in better agreement with similar slopes around the threshold. The best upper limit has been moved to lower masses with

$$\sigma^{\mathrm{SI}}_{\mathrm{nucleon}} = 4.6 \cdot 10^{-45}\,\mathrm{cm}^2 \text{ for } m_\chi = 55\,\mathrm{GeV}. \tag{4.22}$$

It can be concluded that the benchmark result is well-matched by the PAX-based limit after initial deviations at low WIMP masses. The shape of the curves becomes similar

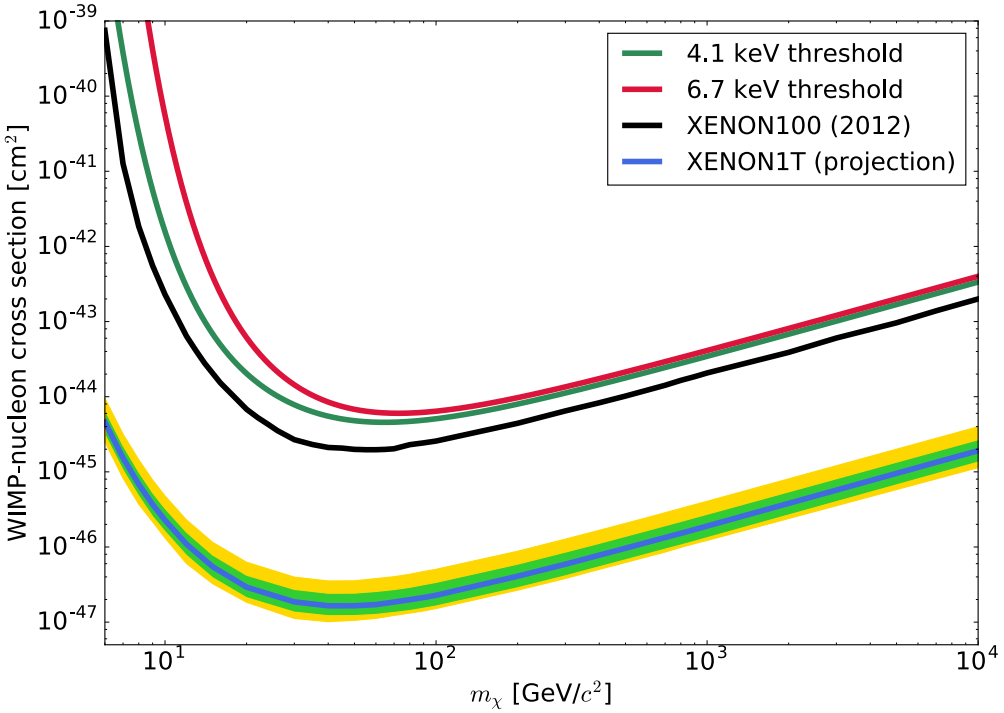

Figure 4.17: Upper limits of the WIMP-nucleon interaction cross section with a rough correction for Poisson fluctuations at the S1 signal threshold (green line). The plot also shows the previous limit from this work (red), the benchmark profile-likelihood analysis (black) [34] and the XENON1T sensitivity projection after an exposure of $2\,\mathrm{t} \cdot \mathrm{y}$ (blue) [55]. 1σ (green) and 2σ (gold) uncertainty bands are given for the XENON1T sensitivity prediction.

when the effect of Poisson smearing is approximated. A similarity at high masses has been achieved even though a simplified spherical model for the nuclear form Factor F^2 has been chosen. The overall upper limit on the spin-independent WIMP-nucleon interaction cross section is a factor of ≈ 2.3 above the original work. This can be attributed to both the lower signal acceptance resulting from the cuts not being optimized for PAX as well as the simpler statistical analysis method.

The result could be improved in several ways. A more competitive lower limit could be achieved by applying likelihood methods that utilize signal and background models rather than a fixed discrimination. Different form factors could be tested for their effect. In addition, a smearing function could be applied to the whole dark matter search region for a thorough treatment of Poisson fluctuations. Uncertainties of L_y, $\mathcal{L}_{\text{eff}}$ as well as those of the astronomical parameters could also be addressed. However, all these approaches would overstretch the scope of this master thesis which was aimed at examining the capabilities of PAX for dark matter search.

CHAPTER

FIVE

CONCLUSION AND OUTLOOK

This work set about to test the impact of the novel PAX raw data processor for XENON1T on a dark matter search with XENON100. To this end, the introduction highlighted the quest for a direct detection of dark matter, the *XENON Dark Matter Project*, and outlined the structure of this work.

Chapter one focused on observational evidence for dark matter, candidate particles and general aspects of direct detection. Evidence from galaxy rotation curves, CMB, large scale structure and cluster collisions have been detailed. Weakly interacting massive particles with masses on the electroweak scale have then been introduced as the particles of interest for direct detection experiments. The leading experiments in the field are based on dual-phase TPCs filled with noble liquids. These detectors look for elastic scatters off target nuclei at extremely low event rates. Therefore, the reduction and discrimination of external and intrinsic backgrounds are vital and paired with sophisticated statistical analysis methods. Liquid xenon TPCs are highly competitive owed to their scalability, the self shielding and purity of liquid xenon, as well as the dual signal of light and charge that is read out.

The mechanisms of generating light and charge in elastic scatters of WIMPs with target nuclei have been put forward. In liquid xenon these are the deexcitation of dimers as well as ionization followed by recombination. In this respect the relative scintillation efficiency of nuclear recoils compared to electronic recoils $\mathcal{L}_{\text{eff}}$ has been introduced. Measurement aspects of the prompt scintillation signal S1 and the delayed scintillation from drifted amplified charges S2 include position-dependent light collection efficiencies as well as charge losses in the liquid and amplification in the gas phase. The last section of the first chapter was dedicated to the XENON100 detector which now makes way for the first operational multi-ton dual-phase TPC for direct dark matter searches: XENON1T.

Chapter two introduced the namesake tool of this work, the PAX raw data processor, and formulated an agenda for the dark matter search to be performed. PAX's working principle is centered around a PMT channel-wise search for photon hits and their subsequent clustering to peaks. This approach is different from the original XENON100 processor, which ran a peakfinder on the summed waveform, and makes PAX more robust against noise. As a means of validation, PAX has been directly compared to the old processor in a full dark matter search on 224.6 live days of XENON100 data. All processor-independent aspects of the benchmark analysis [34] have been adopted in this work including the

background expectation, parametrization of $\mathcal{L}_{\text{eff}}$ and the cuts from the original analysis. However, the new data structure of PAX necessitated the translation of all cuts to PAX variables. This was done in chapter three which introduced the legacy cuts, outlined their translation and marked possible challenges for the comparability of variables. Cases where a translation was not possible where also highlighted and a new S2 width consistency cut has been introduced. The fundamental transition from the C++ ROOT framework to the Python-based PAX environment has also been illustrated in a small code example about the application of cuts. Concluding this chapter the combined cut acceptance has been determined yielding a flat 75.2 % acceptance over the full range of the later-defined dark matter search region.

The second part of chapter three focused on detector calibration with respect to electronic and nuclear recoils. Both calibrations put to use γ-lines from external and internal sources. The mono-energetic lines have been extracted from the data by means of cuts and, due to anti-correlation of light and charge, appear as tilted ellipses in combined S1+S2 space. The means and rotation angles of these ellipses can be used to determine the light and charge yields at the given energies which are used to derive an energy calibration. An electronic recoil energy calibration has been carried out using the combined information of S1 and S2 (CES). Three methods for constructing the CES have been compared using either the mean and rotation angle of a single line, the means of several lines, or their rotation angles. The second method emerged as the simplest and offered an energy resolution $\sigma_E/E \leq 4\,\%$ paired with high accuracy. With minor modifications it is directly applicable to the energy calibration of XENON1T. Moreover, it provided the photon detection efficiency g_1 and charge amplification factor g_2 for the definition of the nuclear recoil energy scale.

These have been used to interpolate measured light yields with the modeled ones from the NEST simulation package. The result is the light yield at the reference energy for the nuclear recoil energy scale $L_y(122\,\text{keV}_{\text{ee}}) \approx 2.25 \pm 0.38_{g1} \pm 0.09_{\text{model}}$ which is lower than in the original analysis, but agrees within the uncertainties. With the original parametrization of $\mathcal{L}_{\text{eff}}$ the energy energy interval of the dark matter search region has been defined as $(6.7 - 30.9)\,\text{keV}_{\text{nr}}$. This assumed the original S1 boundaries of $(3 - 20)\,\text{pe}$ which have been tuned for a high signal to background ratio.

Chapter four was dedicated to the dark matter search. The electronic and nuclear recoil bands have been defined on data from ^{232}Th and ^{60}Co sources for ER and ^{241}AmBe for NR. They have then set the upper and lower boundaries of the dark matter search region in discrimination space $\log_{10}\left(\frac{cS2_b}{cS1}\right)$ with the 99.75 % interval of the Gaussian ER band and the 97 % quantile of the NR band. For comparability with the original analysis the bands have been flattened by subtracting the electronic recoil band mean from all values of the discrimination variable. Disparities in the shape and location of the bands have been identified and explained with processor-induced effects and untuned cuts.

After cuts 26 events remained below the electronic recoil discrimination line of which two events could be identified as genuine signal events in the dark matter search region. One of these events had also been observed in the original analysis. The properties of the event provided by PAX have been compared to their old processor counterparts. Disparities emerged in the S1 and S2 areas and the reconstructed position. The former have been attributed to differences in clustering while the origin of the latter is unclear. A downward-

shifted reconstructed z-coordinate could originate from a false value of the assumed electron drift velocity. Yet, it could also be related to the radial inward shift observed in the (x,y) coordinate at larger depths. If this shift is due to field distortions the effective drift velocity is altered due to the bent path of the electrons.
The second remaining event in the ROI had not been found in the original analysis. It is assumed that it might have been removed by the PMT pattern likelihood cut that could not be translated to PAX variables.
One of the original signal candidates from the benchmark analysis has been discarded by the S2 asymmetry cut in this work. It could be shown that the event is located at the parameter edge of the cut boundary. A comparison of the old processor and PAX waveforms showed different peak areas which translate to different contributions from the top and bottom PMT arrays. These caused the event to be rejected by the cut. With regard to position reconstruction an inward shift in (x,y) and a downward shift in z could be found again. Being located at a smaller z-coordinate the effect is smaller. This indicates that field distortions could be present at the bottom TPC edges. With regard to XENON1T it would have to be investigated if this is physical or a processor artifact.
In a final step, an upper limit on the spin-independent WIMP-nucleon interaction cross section has been set. The two remaining events and the background expectation in the ROI have been inserted into a statistical approach devised by Feldman and Cousins. This yielded a single sided interval for the maximum number of observed events at 90 % confidence level. With the acceptance-corrected exposure an upper bound for the event rate of $2.7 \cdot 10^{-3}$ events/(kg $\cdot$ d) has been found. Applying simplified models for the nuclear form factor F^2 and the Poisson fluctuations of the S1 signal $\sigma^{\mathrm{SI}}_{\mathrm{nucleon}} = 4.6 \cdot 10^{-45}\,\mathrm{cm}^2$ for $m_\chi = 55\,\mathrm{GeV}$ has been obtained. This is reasonable compared to the original result. A lower sensitivity by a factor of ≈ 2.3 has been explained by a less sophisticated statistical approach to the data as well as a lower acceptance of the unoptimized legacy cuts.
Consequently this work illustrates that the PAX-based analysis can reproduce the original XENON100 result from 2012. This marks a validation of PAX under real data conditions in a full analysis. Still, the result could have been improved at several points. For instance, likelihood-based methods for statistical analysis have become a standard in direct dark matter detection. As they use signal and background models they don't require hard discrimination cuts which amounted to an acceptance loss of $\approx 70\,\%$ in this work. A higher acceptance could also have been achieved by an optimization of the cuts instead of a mere translation. However, these points have not been pursued because this work was rather aimed at gaining insights on a broad range of analysis aspects related to PAX.
In addition to a performance test for the new data processor this work can be seen as a guide to XENON analysis directed at prospective undergraduate and graduate students. As all analysis code for this work had to be newly developed, it provides a full analysis suite covering aspects of calibration, cut development and dark matter analysis. All programs are provided as accessible, interactive and browser-based *Jupyter* notebooks. This provides newcomers to XENON with cues for their own analyses. By virtue of the shared data structure it does not matter if the prospective data comes from XENON100, a small institute TPC or XENON1T.

CHAPTER

SIX

APPENDIX

6.1 List of Abbreviations

AFB	Peak area fraction on bottom PMT array
AFT	Peak area fraction on top PMT array
BBN	Big Bang nucleosynthesis
CES	Combined energy scale
CMB	Cosmic microwave background
DAQ	Data acquisition system
DM	Dark matter
EFT	Effective field theory
ER	Electronic recoil
LNGS	Laboratori Nazionali del Gran Sasso
MACHO	Massive compact halo object
MOND	Modified Newtonian Dynamics
NEST	Noble Element Simulation Technique
NR	Nuclear recoil
PMT	Photomultiplier tube
PTFE	Polytetrafluorethylene
QCD	Quantum Choromodynamics
QED	Quantum Electrodynamics
QFT	Quantum field theory
ROI	Region of interest

S1	Prompt scintillation signal in a dual-phase TPC
S2	Secondary scintillation signal from electroluminescence of drifted, extracted and accelerated ionization charges in a dual-phase TPC
SD	Spin-dependent (interaction)
SI	Spin-independent (interaction)
SM	Standard Model (of particle physics)
TPC	Time projection chamber
WIMP	Weakly interacting massive particle

6.2 Formal Cut Definitions

These are the formal Boolean definitions of the XENON100 legacy cuts. If the condition formulated in the cut is false, the event is rejected by the cut. The definitions of the variables are given in Tab. 6.1.

Data Quality Cuts

S1 coincidence:

$$n_{\mathrm{S1}} > 1$$

S1 width:

$$W_{\mathrm{S1,90}} > 26\,\mathrm{ns}$$

S2 seen by top PMT array:

$$\mathrm{S2_{aft}} > 0$$

Fit quality of position reconstruction:

$$\frac{\chi^2_{(x,y)}}{n_{\mathrm{S2}} - 1} < 7$$

Difference of positions reconstructed by neural network and robust weighted mean position reconstruction algorithms from the final reconstructed position:

$$\sqrt{(x_{\mathrm{nn}} - x)^2 + (y_{\mathrm{nn}} - y)^2} + \sqrt{(x_{\mathrm{rwm}} - x)^2 + (y_{\mathrm{rwm}} - y)^2} < 7$$

No signal in LXe veto:

$$\mathrm{veto_{max}} < 0.35\,\mathrm{pe}$$

Avoid early S2s:

$$\frac{t_{\mathrm{S2}}}{1000} > 178\,\mu\mathrm{s}$$

S2 asymmetry cut against gas events:

$$-0.2 < \mathrm{S2_{aft}} < 0.25$$

Cut γ-like events from AmBe neutron band:

$$\log_{10}\left(\frac{cS2_b}{cS1}\right) < 3.1$$

Signal to noise cut against high peak counts in events:

$$\log_{10}\left(\frac{s1+s2}{A_{\text{tot}}}\right) > 0.284 \cdot (\log_{10}(s2))^3 - 0.281 \cdot (\log_{10}(s2))^2 + 10.1 \cdot \log_{10}(s2) - 12.3$$

Energy and Threshold Cuts

DAQ threshold:

$$s2 > 150\,\text{pe}$$

Single Scatter Cuts

Single physical S1 in an event:

$$n_{\text{S1,j}} < 2 \vee t_{\text{S2}} - t_{\text{S1,j}} < 18500\,\text{ns} \vee \log_{10}\left(\frac{cS2}{cS1_{\text{j}}}\right) > 3.2\ \forall j \in \text{2nd, 3rd, 4th, ...}$$

S2 single scatter:

$$s2_{\text{2nd}} < \left(70\,\text{pe} + \frac{s2 - 300\,\text{pe}}{100}\right) \vee n_{\text{interactions}} = 1$$

Fiducial Volume Cut

Ellipsoidal 34 kg fiducial volume:

$$\left(\frac{|-z-15.2\,\text{cm}|}{12.68\,\text{cm}}\right)^{2.7} + \left(\frac{x^2+y^2}{175.0\,\text{cm}^2}\right)^{2.7} < 1$$

Table 6.1: Description of variables used in the translated cuts.

Variable	Unit	Definition	Notes
$cS1_{\mathrm{j}}$	pe	S1 area corrected for the interaction position of interaction j $\in$ 2nd, 3rd, 4th, ... No index indicates the main interaction.	Position correction with S1 LCE map.
$cS2_{\mathrm{j}}$	pe	S2 counterpart of the above variable.	Position correction with S2 LCE map and electron lifetime.
$cS2_b$	pe	S2 area seen by the bottom PMT array: $cS2_b = cS2 \cdot (1 - \mathrm{S2_{aft}})$.	Higher robustness against saturation than $cS2$.
$s1_{\mathrm{ij}}/s2_{\mathrm{ij}}$	pe	Uncorrected area of the S1/S2 peak. The index i refers to one of the 178 PMTs inside the TPC. A missing index translates to the summed signal over all PMTs. The index j gives the interaction.	Dynamic threshold setting and clustering change area with respect to old processor.
x, y	$(0 - 15.3)\,\mathrm{cm}$	(x,y)-position of the main interaction. Origin at the center of the top PMT array. Subscript indices nn and rwm refer to the neural network and robust weighted mean position reconstruction algorithms.	Given in mm by the old processor.
z	$(-30.5 - 0)\,\mathrm{cm}$	z-position of the main interaction.	Given in mm by the old processor.
$n_{\mathrm{S1,j}}$	-	Number of PMT channels with a contribution $s1_i > 0.35\,\mathrm{pe}$ to $s1$. Noisy PMTs are not counted. The index j refers to the interaction.	Noisy PMT channels bumping up the coincidence criterion are 152, 170, 167.
$n_{\mathrm{S2,j}}$	-	S2 counterpart to the above variable	-
$\mathrm{S2_{aft}}$	-	Fraction of $s2$ seen by the top PMT array.	-
$t_{\mathrm{peak,\ j}}$	ns	Center time for an S1/S2 peak from interaction j.	-
$W_{\mathrm{S1,90}}$	ns	S1 peak width at 90 % peak area.	Given in samples by the old processor: 1 sample = 10 ns.
$\mathrm{veto_{max}}$	pe	Largest hit registered in liquid xenon veto.	-
$\chi^2_{(x,y)}$	-	Goodness of fit for the reconstructed (x,y)-position of an event.	-
A_{tot}	pe	Total summed and uncorrected area of all peaks in an event.	The old data processor does not find peaks after the main S2 while PAX does.

6.3 S2 Width Consistency Cut on Nuclear Recoil Data

An S2 width consistency cut defined on NR data replaces the respective legacy cut which could not be translated. The definition on ER data is not sufficient because it has been found that its NR acceptance is at just 50 % in the dark matter search region.
Fig. 6.1 shows the S2 width of NR events from elastic scatters of ^{241}AmBe neutrons at 90 % of the peak area against t_d. Typical signal widths in the energy region of $cS1 < 60\,\mathrm{pe}$ range between $(0.7 - 1.7)\,\mu\mathrm{s}$. The band of S2 widths is divided into slices of $\Delta t_d = 10\,\mu\mathrm{s}$ and each slice is fitted with a Gaussian distribution. The fit gives the mean and the $\pm 3\sigma$ exclusion intervals for the cut.

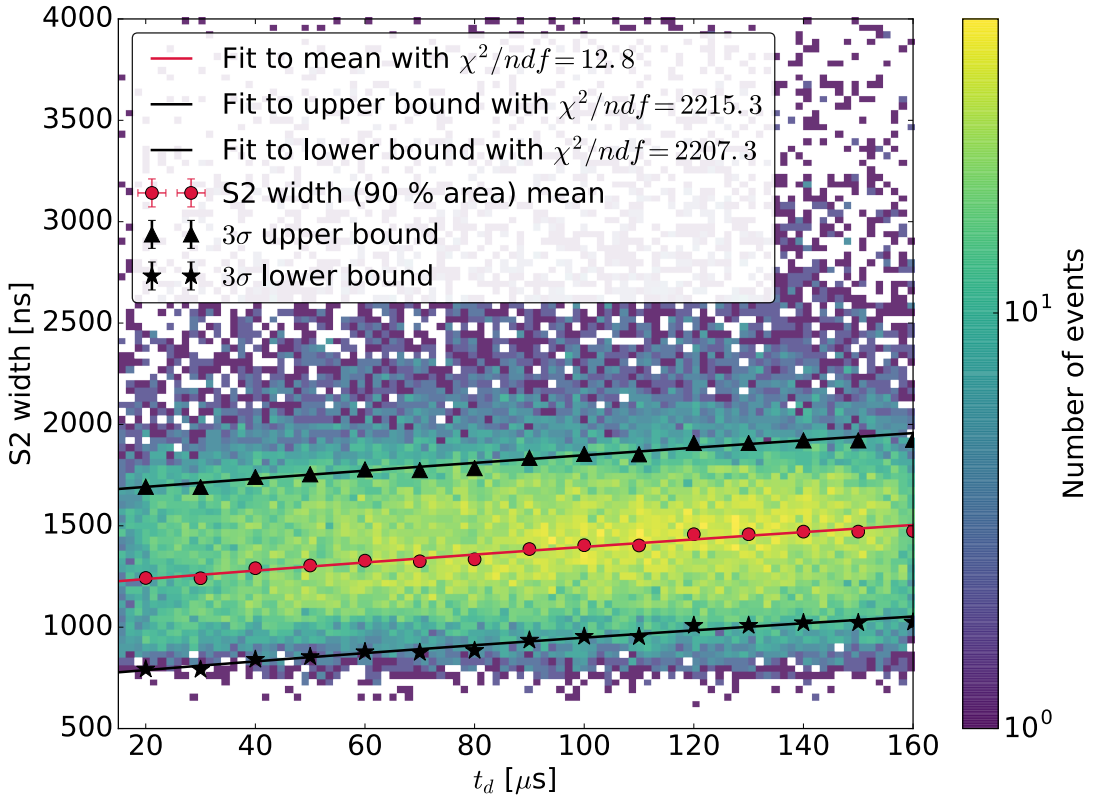

Figure 6.1: S2 width of NR events from ^{241}AmBe at 90 % of the peak area against t_d in the energy region of $cS1 < 150\,\mathrm{pe}$. Mean and 3σ intervals are determined by Gaussian fits of $\Delta t_d = 10\,\mu\mathrm{s}$-slices of the S2 width band. The points for the slices are overlaid with square root functions to obtain cut functions for the upper and lower bound. The uncertainties of the points are too small to be visible in the plot.

The cut is then defined as:

$$\begin{aligned} &\text{S2 width}_{90} \leq \sqrt{6.910 \cdot 10^{-9}\,\mathrm{s} \cdot t_d + 2.724928 \cdot 10^{-12}\,\mathrm{s}^2} \\ \wedge \quad &\text{S2 width}_{90} \geq \sqrt{3.474 \cdot 10^{-9}\,\mathrm{s} \cdot t_d + 0.552186 \cdot 10^{-12}\,\mathrm{s}^2}. \end{aligned} \tag{6.1}$$

It is applied on all NR and dark matter search data where it has an almost flat acceptance of $\approx 80\,\%$ for $cS1 \leq 60\,\mathrm{pe}$.

6.4 Neutron Velocity Calculation

The velocity of a 2 MeV neutron emitted by the ^{241}AmBe source can be calculated by:

$$E_{\text{kin}} = (\gamma - 1)mc^2 \tag{6.2}$$

where E is the kinetic energy of the neutron, c is the speed of light, $m = 939.565\,\text{MeV/c}^2$, and

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{6.3}$$

This yields

$$v = c \cdot \sqrt{1 - \left(\frac{mc^2}{E + mc^2}\right)^2} \approx 1.95 \cdot 10^7\,\text{m/s} \tag{6.4}$$

Therefore a neutron with this kinetic energy needs

$$\frac{0.306\,\text{m}}{1.95 \cdot 10^7\,\text{m/s}} \approx 15.67 \cdot 10^{-9}\,\text{s} \tag{6.5}$$

to traverse the 30.6 cm radius of the TPC. Although scattering prolongates the path the total travel time should stay on the order of 10 ns.

6.5 Uncertainties for single line CES parameters

The uncertainties for the parameters of the single line CES are calculated by Gaussian error propagation:

$$\Delta m = \frac{2\Delta\theta}{\cos(2\theta) + 1}, \tag{6.6}$$

$$\Delta k = \frac{1}{E} \cdot \sqrt{(\mu_{\text{S1}} \cdot \Delta m)^2 + (m \cdot \Delta\mu_{\text{S1}})^2 + (\Delta\mu_{\text{S2}})^2}. \tag{6.7}$$

Resolution:

$$\Delta R = \sqrt{\left(\frac{\Delta\sigma}{E}\right)^2 + \left(\frac{\sigma \cdot \Delta E}{E^2}\right)^2} \tag{6.8}$$

6.6 Fits to γ-lines in CES

The γ-lines in all three CES are fitted with the respective Gaussian and exponentially-modified Gaussian functions. A constant or linear background is added in order to improve fit convergence and quality. Errors of the χ^2-minimization are taken from MINOS.

Linear CES

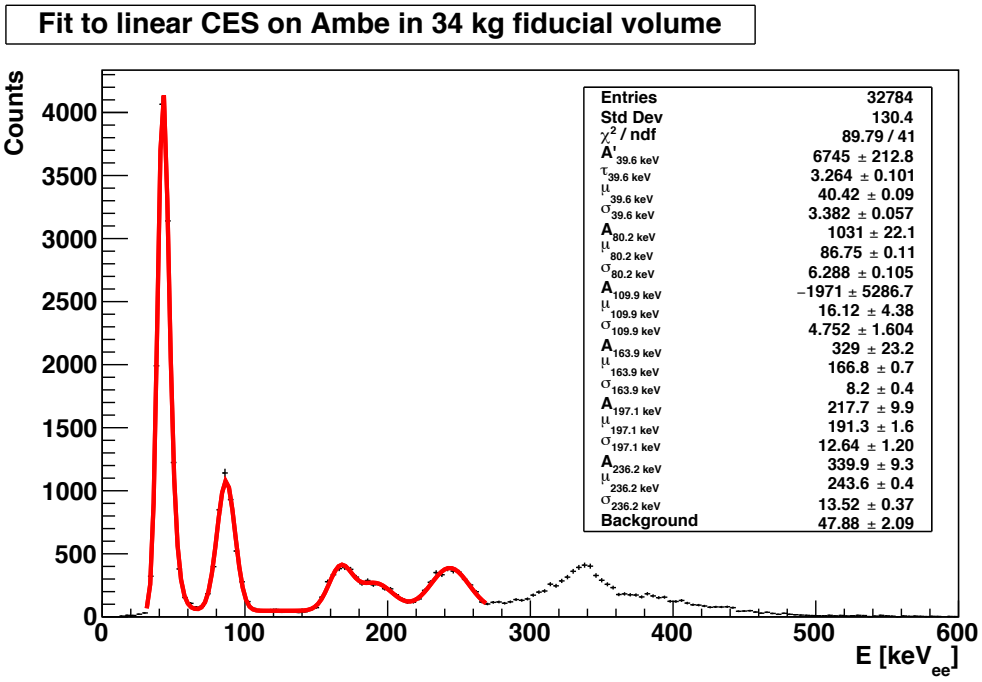

Figure 6.2: Fit of the AmBe spectrum with the linear CES in the 34 kg fiducial volume.

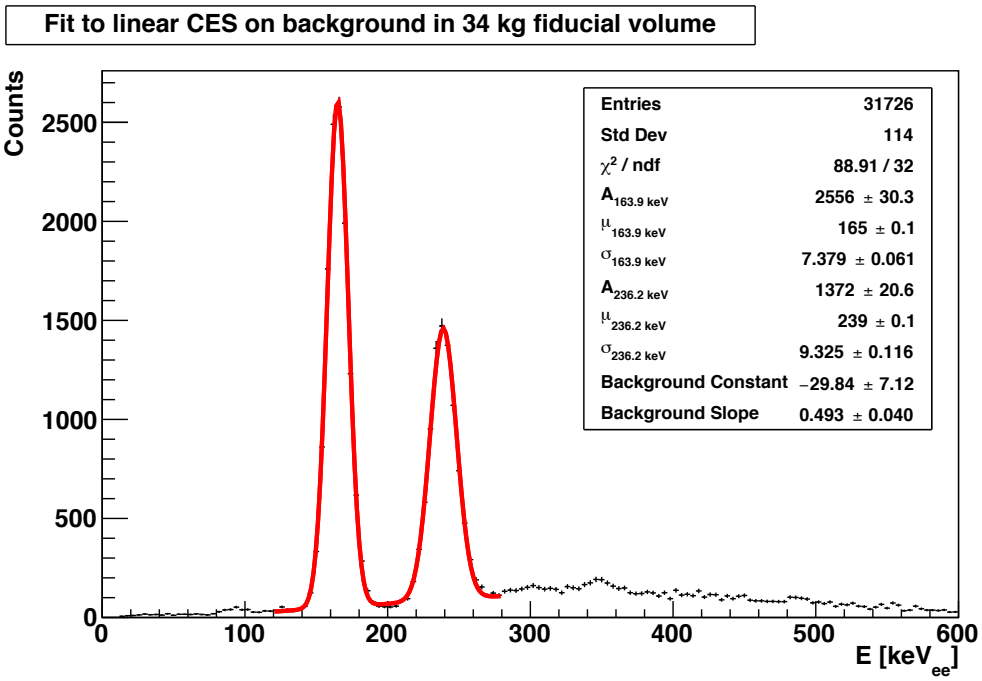

Figure 6.3: Fit of the background spectrum with the linear CES in the 34 kg fiducial volume.

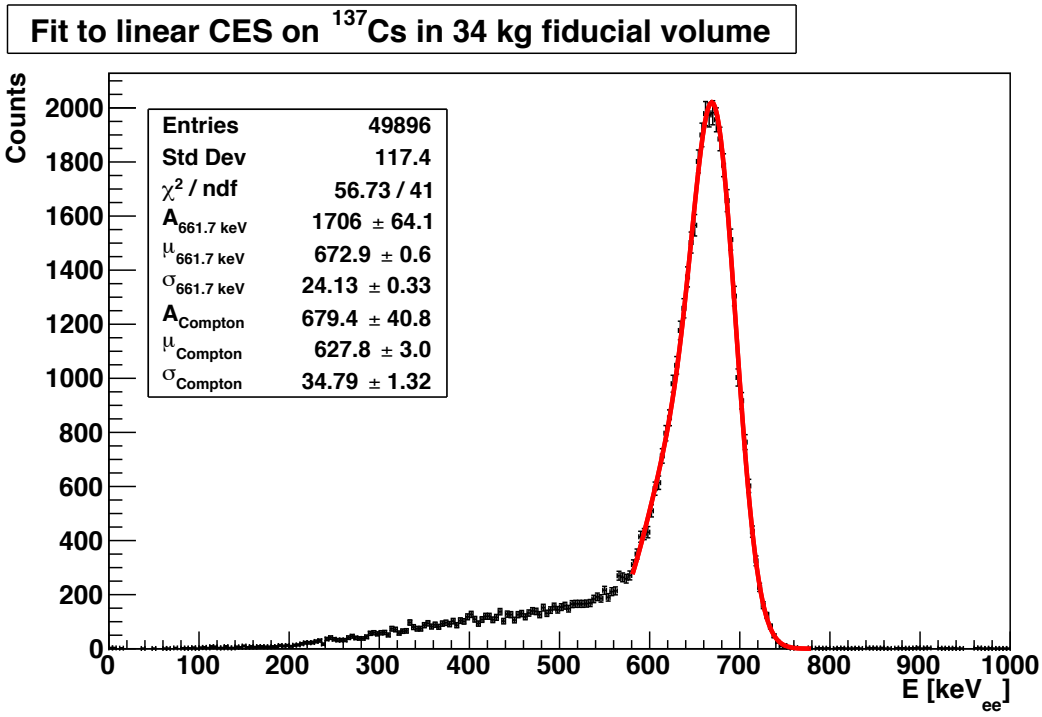

Figure 6.4: Fit of the ^{137}Cs spectrum with the linear CES in the 34 kg fiducial volume.

Single Line CES from ^{137}Cs

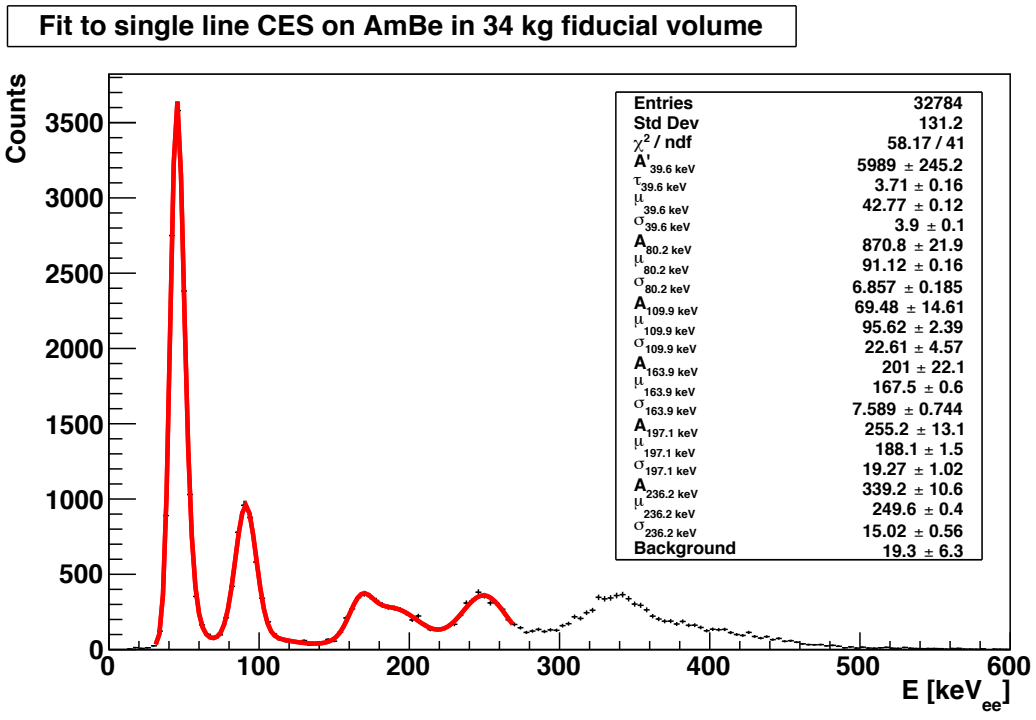

Figure 6.5: Fit of the AmBe spectrum with the single line CES in the 34 kg fiducial volume.

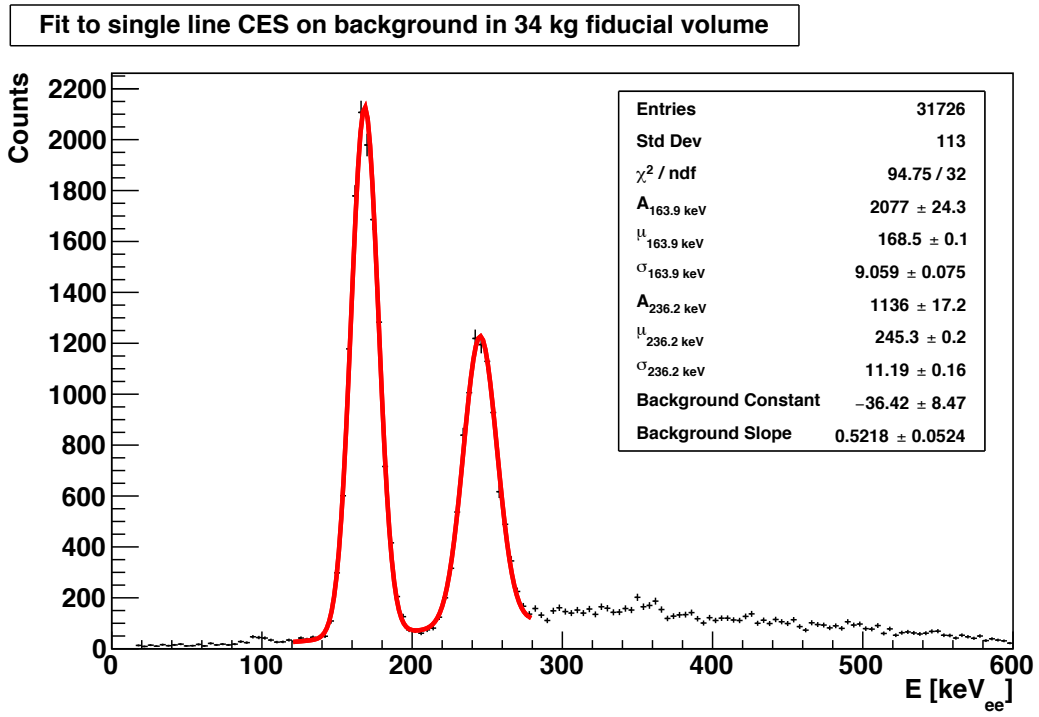

Figure 6.6: Fit of the background spectrum with the single line CES in the 34 kg fiducial volume.

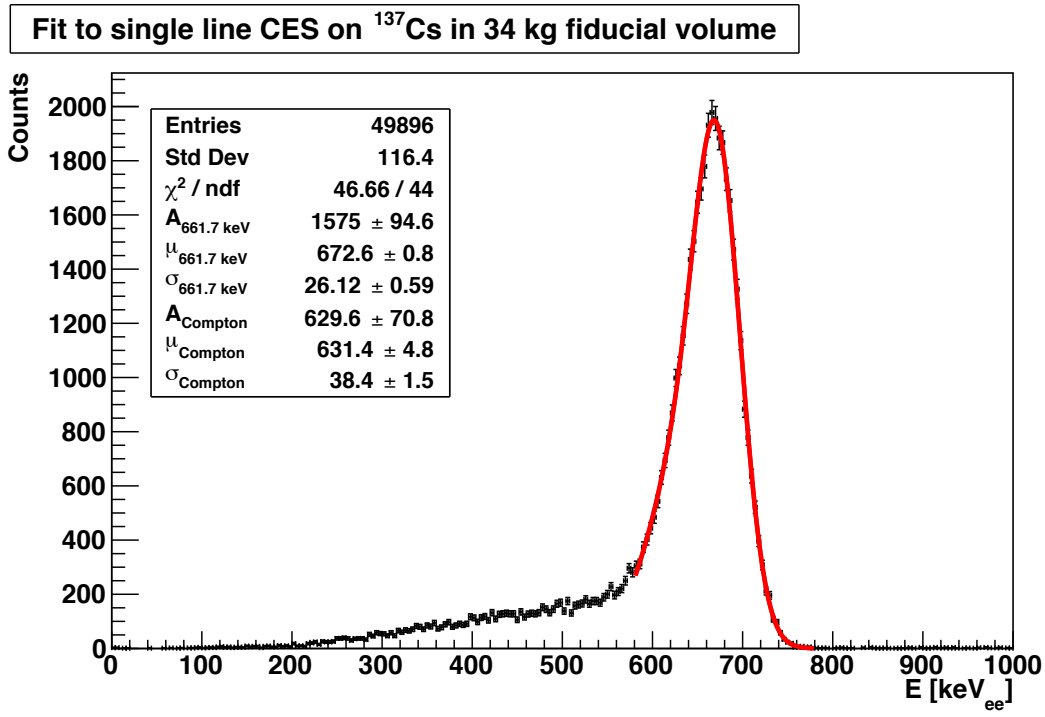

Figure 6.7: Fit of the ^{137}Cs spectrum with the single line CES in the 34 kg fiducial volume.

CES Parametrized with $\theta(E)$ and $k(E)$

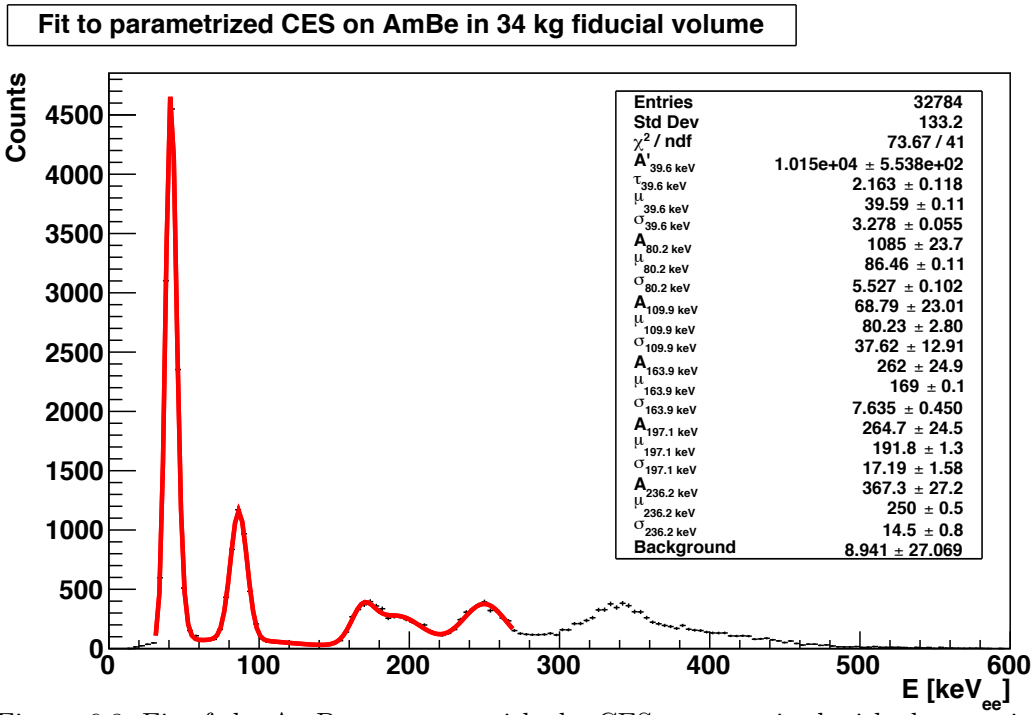

Figure 6.8: Fit of the AmBe spectrum with the CES parametrized with the rotation angle in the 34 kg fiducial volume.

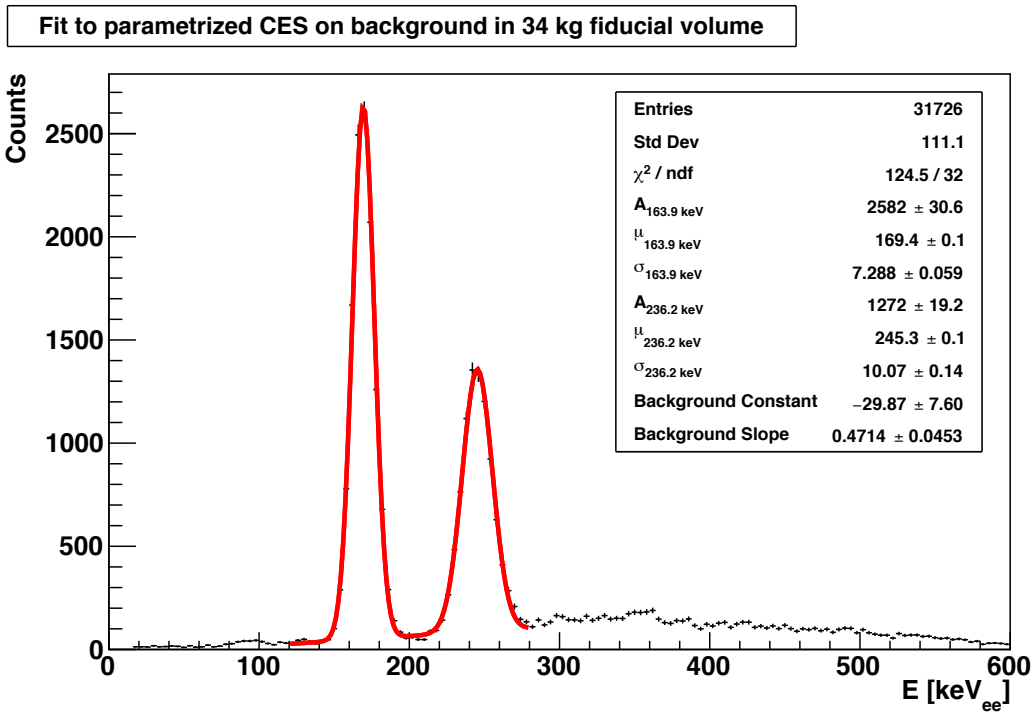

Figure 6.9: Fit of the background spectrum with the CES parametrized with the rotation angle in the 34 kg fiducial volume.

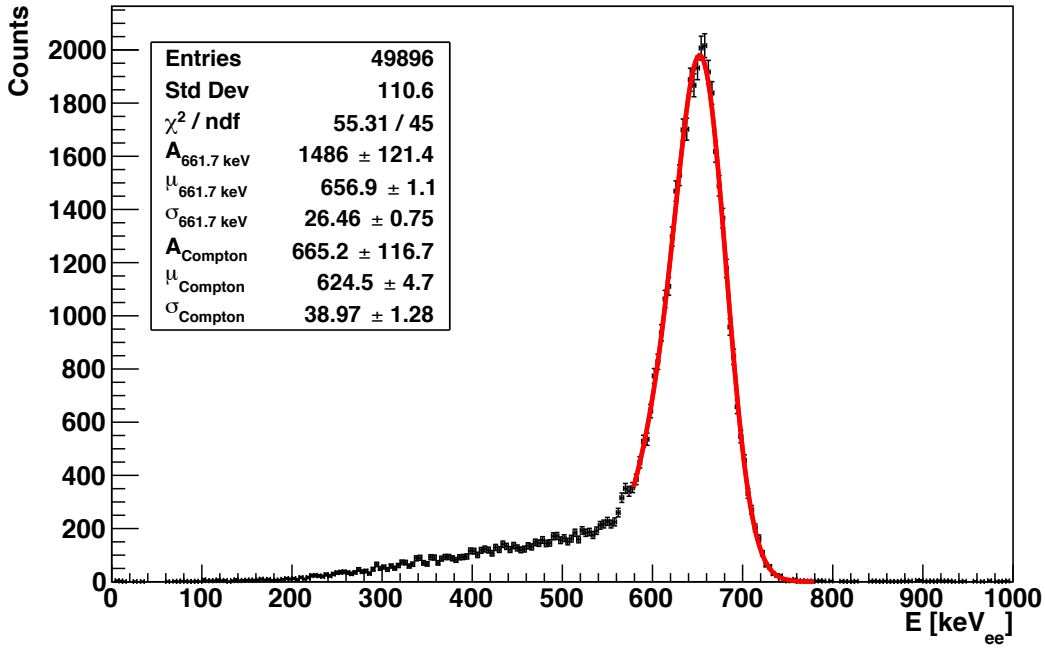

Figure 6.10: Fit of the ^{137}Cs spectrum with the CES parametrized with the rotation angle in the 34 kg fiducial volume.

Estimation of CES Parameter Error Effects

The effect of the parameter errors on the mean values derived from the fits to the γ-lines in each CES is investigated in this section. This is done by calculating the propagated errors ΔE_{CES} for the reconstructed energy of each datapoint:

$$\Delta E_{\mathrm{CES,\ linear}} = \sqrt{(\Delta\alpha \cdot cS1)^2 + (\Delta\beta \cdot cS2_b/100)^2} \tag{6.9}$$

for the linear CES and

$$\Delta E_{\mathrm{CES,\ single\ line/parametrized}} = \sqrt{\left(\frac{\Delta k}{k^2} \cdot \left(\frac{cS2_b + m \cdot cS1}{100}\right)\right)^2 + \left(\frac{\Delta m \cdot cS1}{k}\right)^2} \tag{6.10}$$

for the single-line and parametrized CES. Now the fractions

$$\Psi = \frac{\Delta E_{\mathrm{CES}} - E_{\mathrm{CES}}}{E_{\mathrm{CES}}} \tag{6.11}$$

are computed for each point in the respective CES. The result is shown in Fig 6.6. Additional errors

$$\begin{aligned} (1 - \Psi_{\text{linear CES}}) \cdot E_{\text{linear CES}} &= 0.04 \cdot E_{\text{linear CES}}, \\ (1 - \Psi_{\text{single line CES}}) \cdot E_{\text{single line CES}} &= 0.01 \cdot E_{\text{single line CES}}, \\ (1 - \Psi_{\text{parametrized CES}}) \cdot E_{\text{parametrized CES}} &= 0.05 \cdot E_{\text{parametrized CES}} \end{aligned} \tag{6.12}$$

are added to the mean energies for each line in the respective CES when determining accuracy and resolution. The error for the parametrized CES is not calculated, but estimated conservatively.

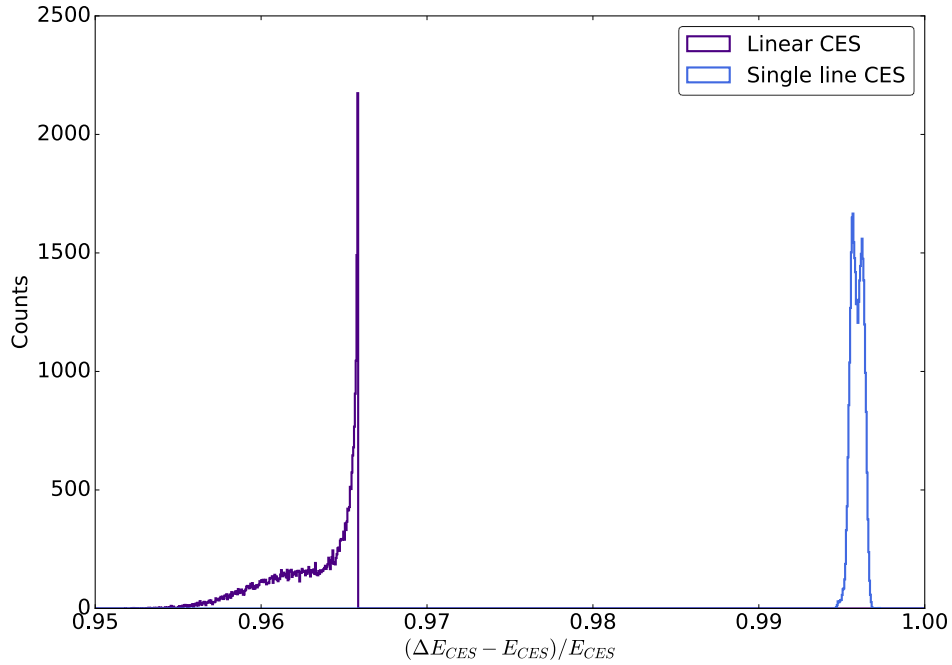

Figure 6.11: Effect of CES parameter errors on linear and single line CES.

6.7 Nuclear Recoil Energy Spectrum from ^{241}AmBe

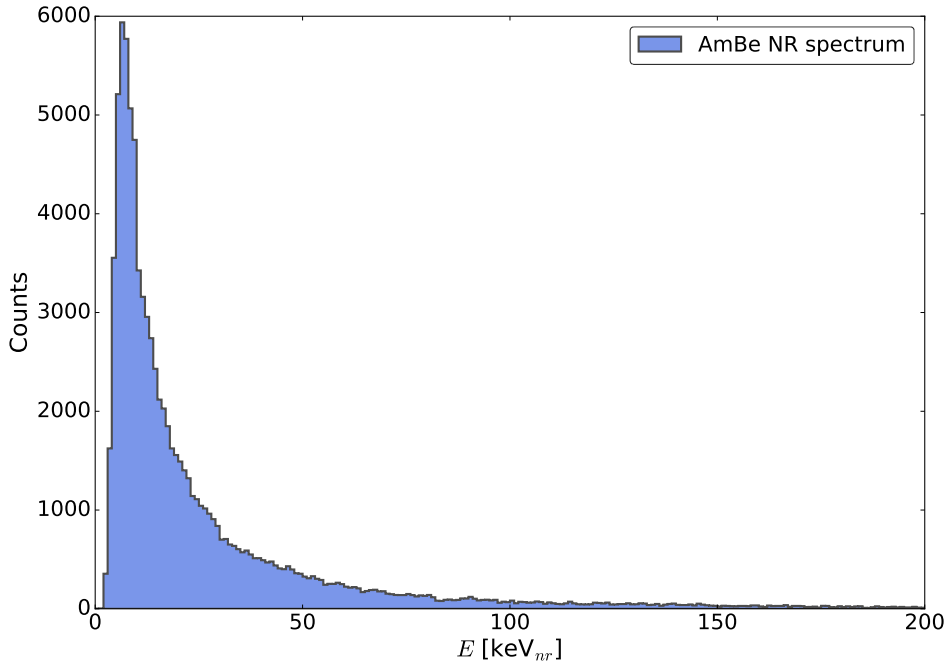

Figure 6.12: E_{nr} spectrum of neutrons from the ^{241}AmBe source up to 200 keV$_{\mathrm{nr}}$. The conversion of $cS1$ to E_{nr} has been carried out with the parametrization of $\mathcal{L}_{\mathrm{eff}}$, $L_y(122\,\mathrm{keV_{ee}}) = (2.25 \pm 0.38_{g_1} \pm 0.09_{\mathrm{model}})\,\mathrm{pe/keV_{ee}}$, $S_{\mathrm{ee}} = 0.58$ and $S_{\mathrm{nr}} = 0.95$.

Fig. 6.12 shows the NR energy spectrum of AmBe neutrons up to 200 keV$_{\mathrm{nr}}$. No entries are present at 0 keV$_{\mathrm{nr}}$ because all value pairs or E_{nr} and $\mathcal{L}_{\mathrm{eff}}$ below 1 keV$_{\mathrm{nr}}$ are set to zero in the parametrization. Additionally, the detector acceptance has to be considered. From 10 keV$_{\mathrm{nr}}$ on the spectrum shows a steep decrease with rising energy.

6.8 Fits for ER Band and NR Band

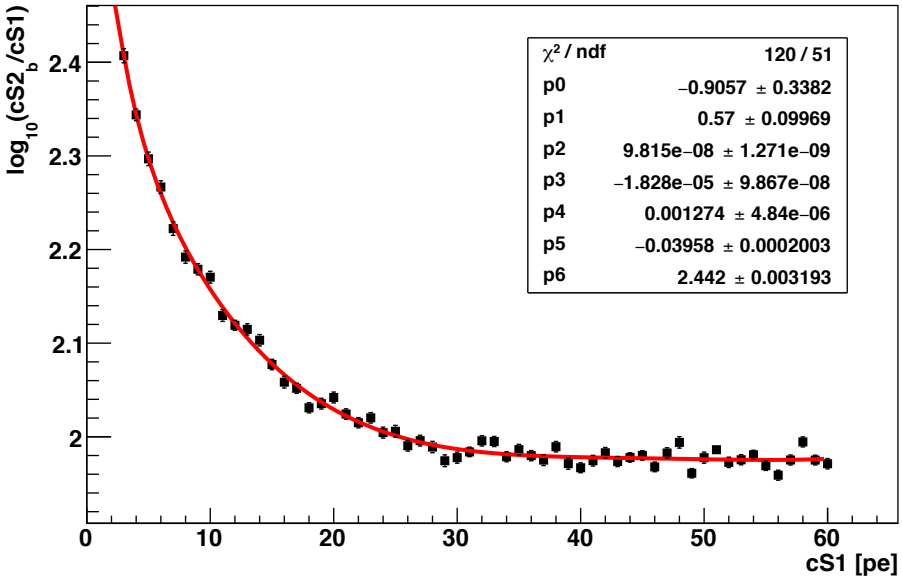

Figure 6.13: ER band mean from ^{60}Co and ^{232}Th data up to 60 pe in $cS1$. All legacy cuts applied on the dark matter data are also applied on this dataset. The band mean is constructed from fitting 1 pe slices with 1D Gaussian distributions. The mean points are in turn fitted with a phenomenological function starting at the lower dark matter search region threshold of $cS1 = 3$ pe.

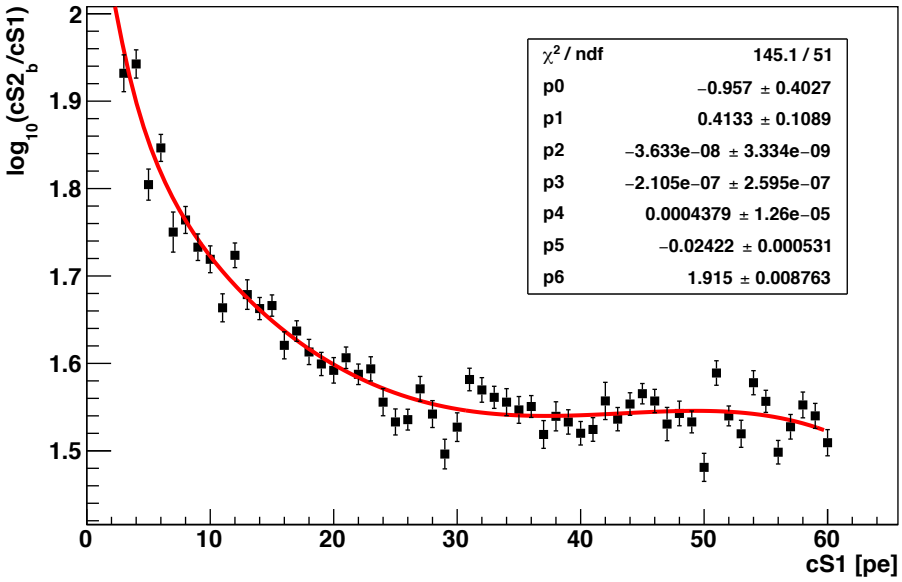

Figure 6.14: ER band 3.09σ exclusion line from ^{60}Co and ^{232}Th data up to 60 pe in $cS1$. All legacy cuts applied on the dark matter data are also applied on this dataset. The 99.75 % exclusion line is constructed from fitting 1 pe slices with 1D Gaussian distributions. The 3.09σ intervals are in turn fitted with a phenomenological function starting at the lower dark matter search region threshold of $cS1 = 3$ pe.

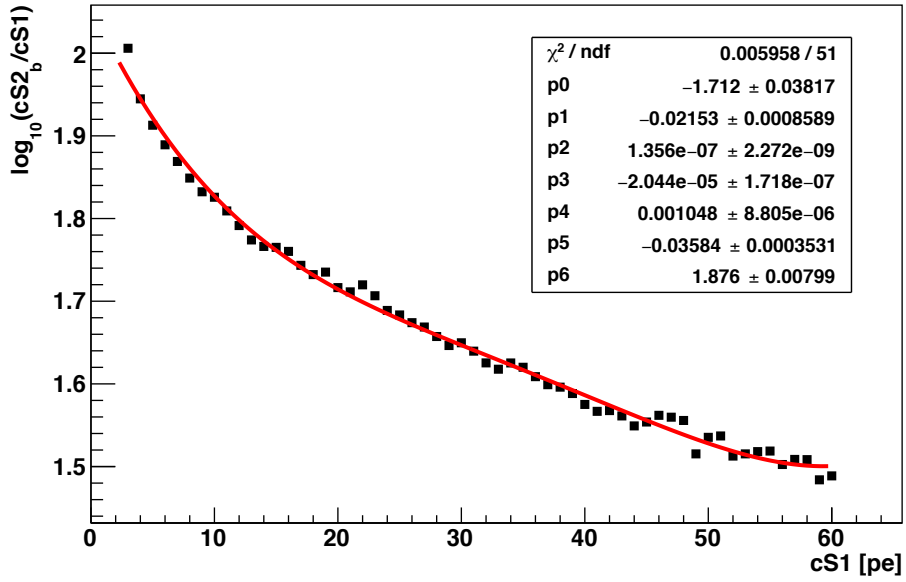

Figure 6.15: NR band median from ^{241}AmBe data up to 60 pe in $cS1$. All legacy cuts are applied. The band median is constructed from 1 pe slices. The median points are fitted with a phenomenological function starting at the lower dark matter search region threshold of $cS1 = 3$ pe.

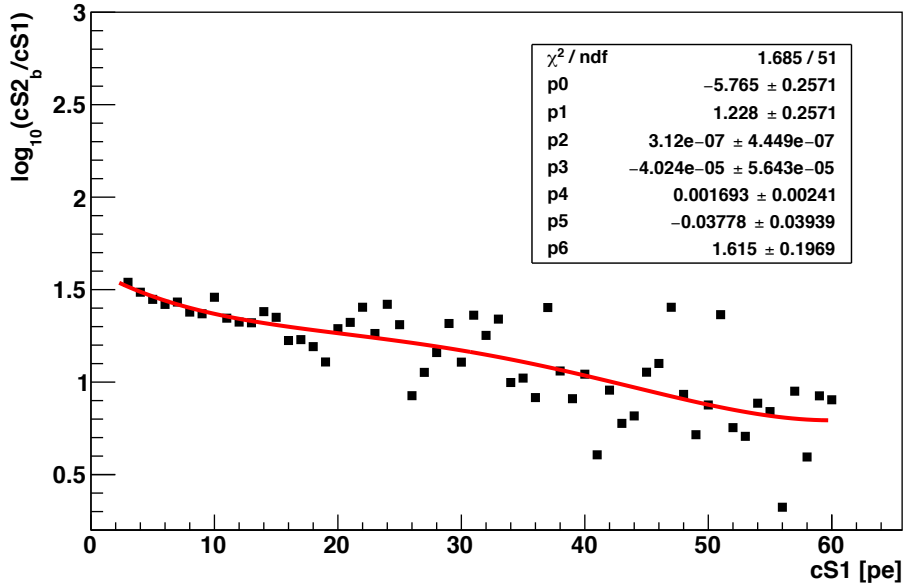

Figure 6.16: NR band 97 % quantile from ^{241}AmBe data up to 60 pe in $cS1$. All legacy cuts are applied. The lower NR acceptance quantile is constructed from 1 pe slices. The datapoints for the individual slices are fitted with a phenomenological function starting at the lower dark matter search region threshold of $cS1 = 3$ pe.

6.9 Band Flattening

The two band flattening steps outlined in section 4.1 are executed between 2 pe and 35 pe. Accordingly, the flattened bands are not valid outside of this $cS1$-range. This is especially evident in Fig. 6.19 where the exponential fit from the first flattening step causes the mean to rise exponentially above the range of flattening.

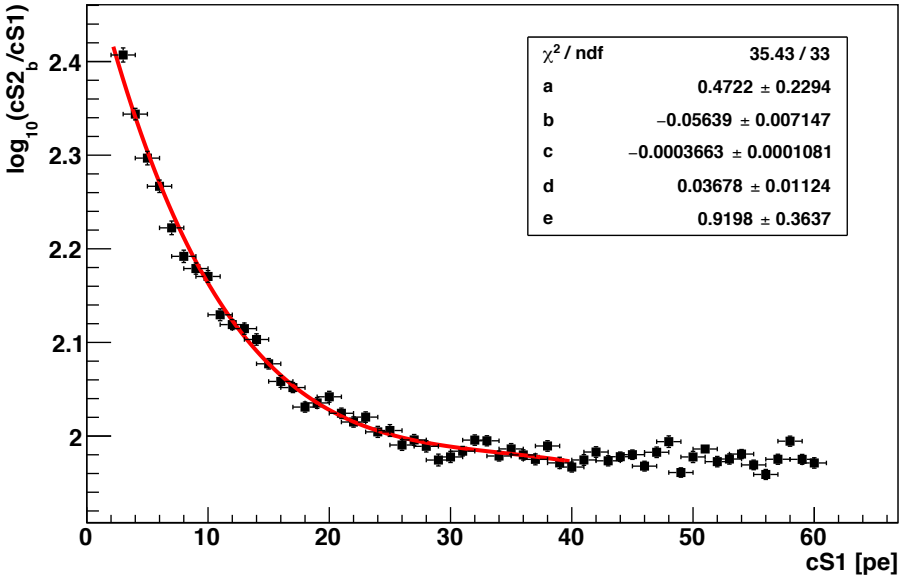

Figure 6.17: ER band means before flattening overlaid with a sum of an exponential function and a second order polynomial between 2 pe and 35 pe.

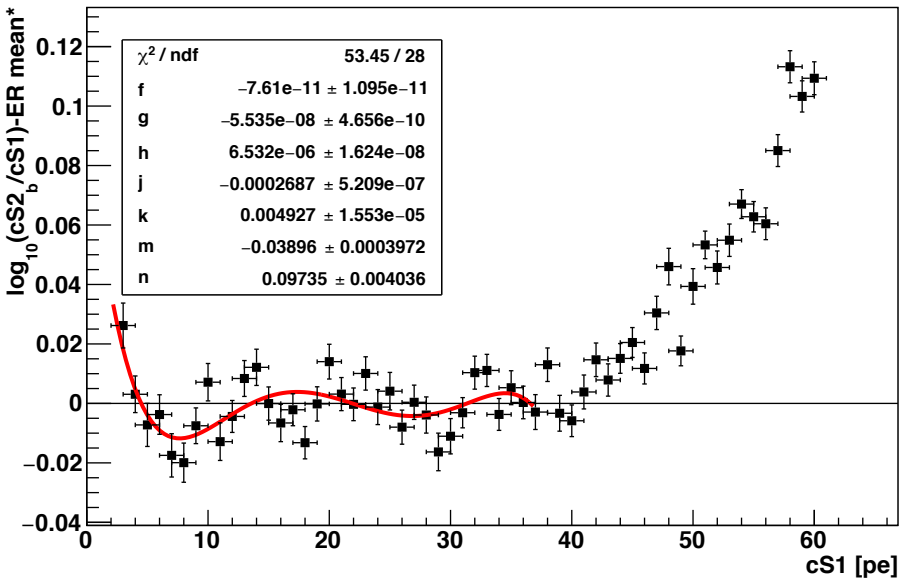

Figure 6.18: ER band means after the first band flattening step overlaid with a sixth order polynomial. The exponential rise at higher $cS1$ values is due to the flattening range between 2 pe and 35 pe.

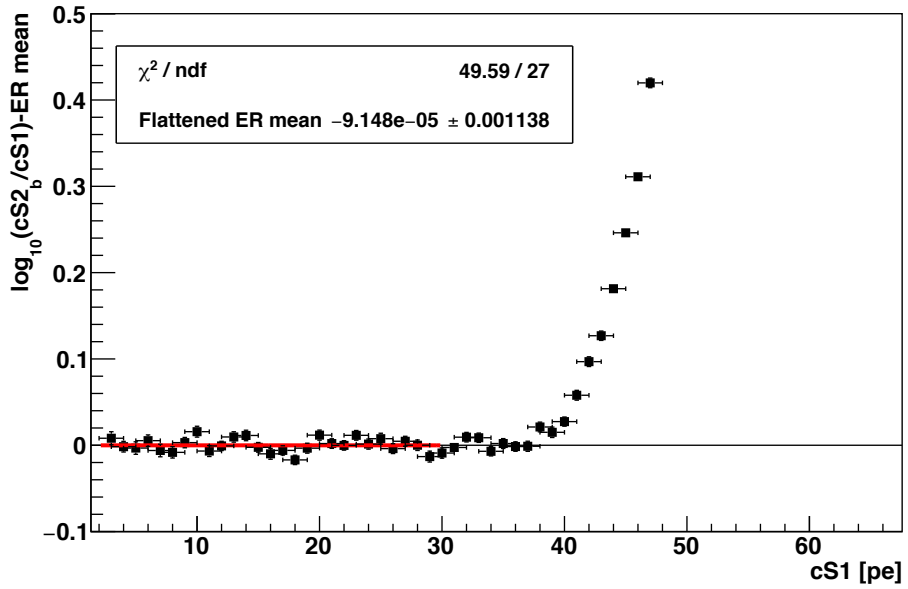

Figure 6.19: ER band means after the second band flattening step overlaid with a constant function. The exponential rise at higher $cS1$ values is due to the flattening range between 2 pe and 35 pe.

6.10 Plots of Events Below the Discrimination Line

Noisy Waveforms

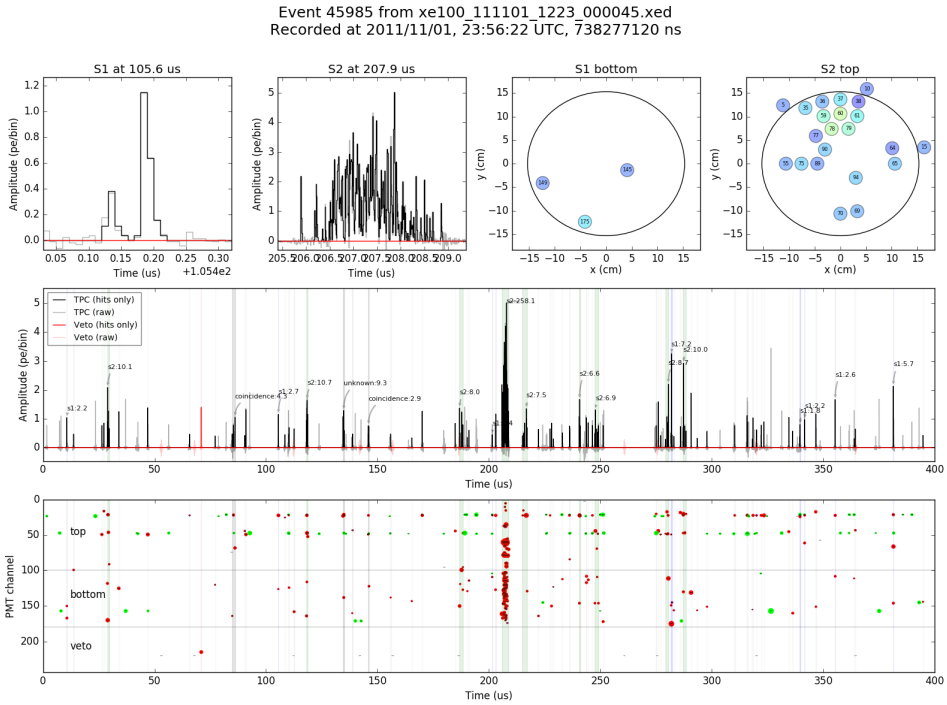

Figure 6.20: Waveform of event 45985 from the data file recorded on 1st November 2011 at 12.23 p.m.

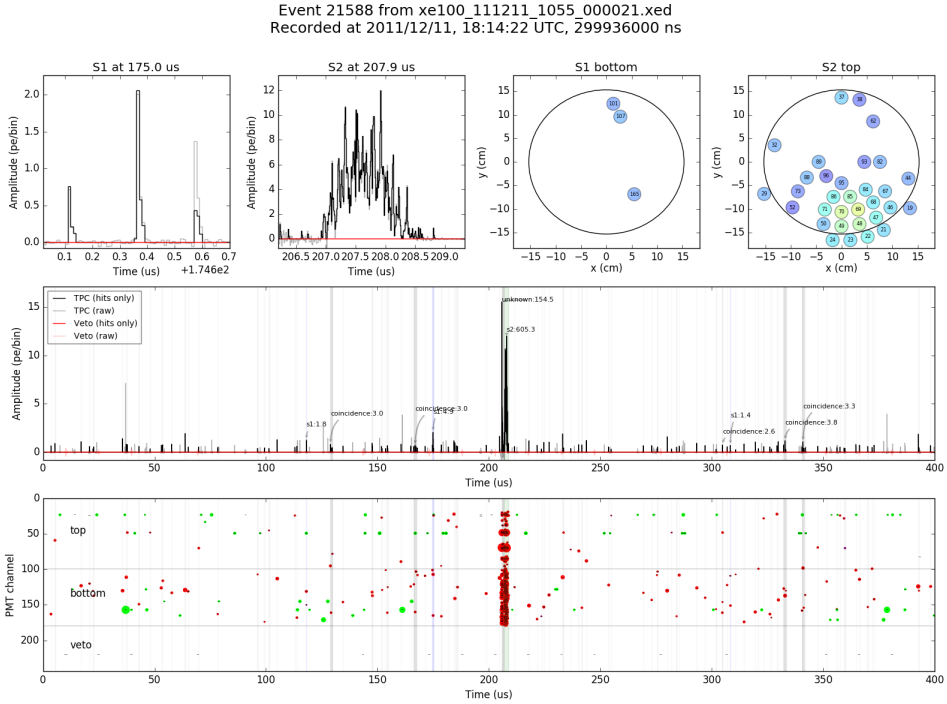

Figure 6.21: Waveform of event 21588 from the data file recorded on 11th December 2011 at 10.55 a.m. This event was present within the dark matter search ROI, but has been excluded due to its noisy waveform with numerous peaks and a large unidentified peak directly preceding the main S2.

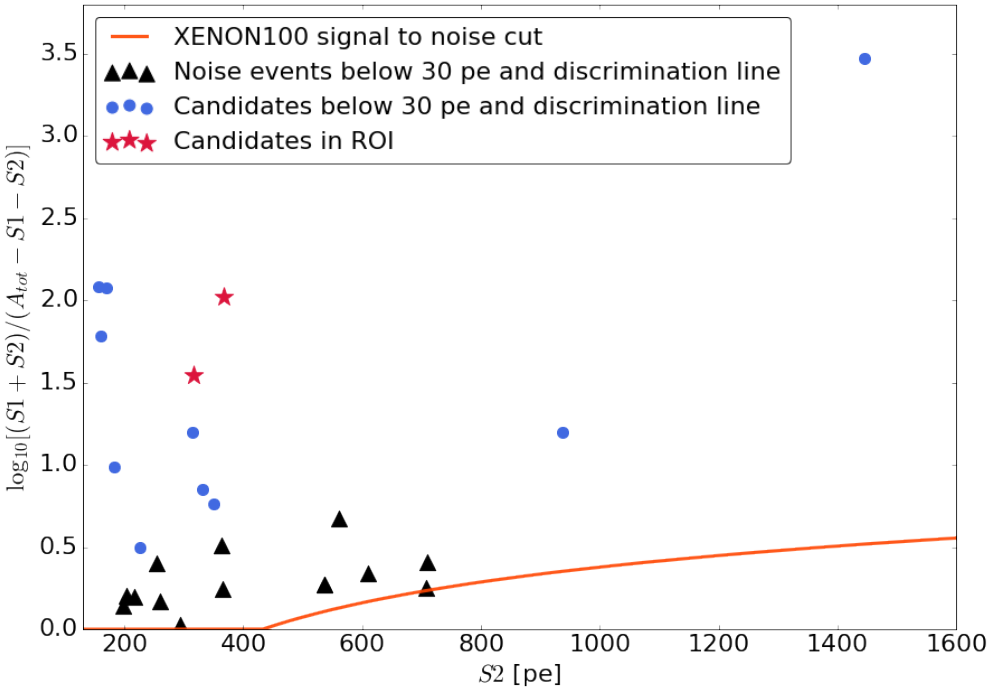

Figure 6.22: Location of the noisy events and event candidates in the parameter space of the signal to noise cut. The cut is delineated by the solid orange line. Noisy events are signified by black triangles, candidates outside the ROI by blue circles and candidates inside the ROI by red stars.

Points Below Lower ROI Threshold

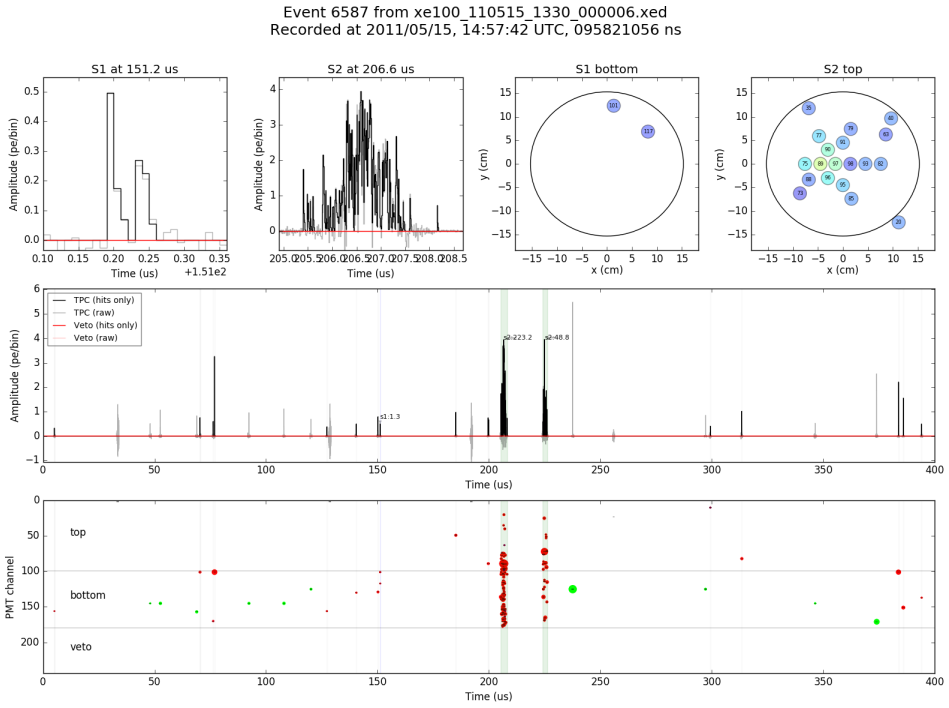

Figure 6.23: Waveform of event 6587 from the data file recorded on 15th May 2011 at 1.30 p.m. It exhibits dual S2 peaks too small to be removed by the S2 double scatter cut.

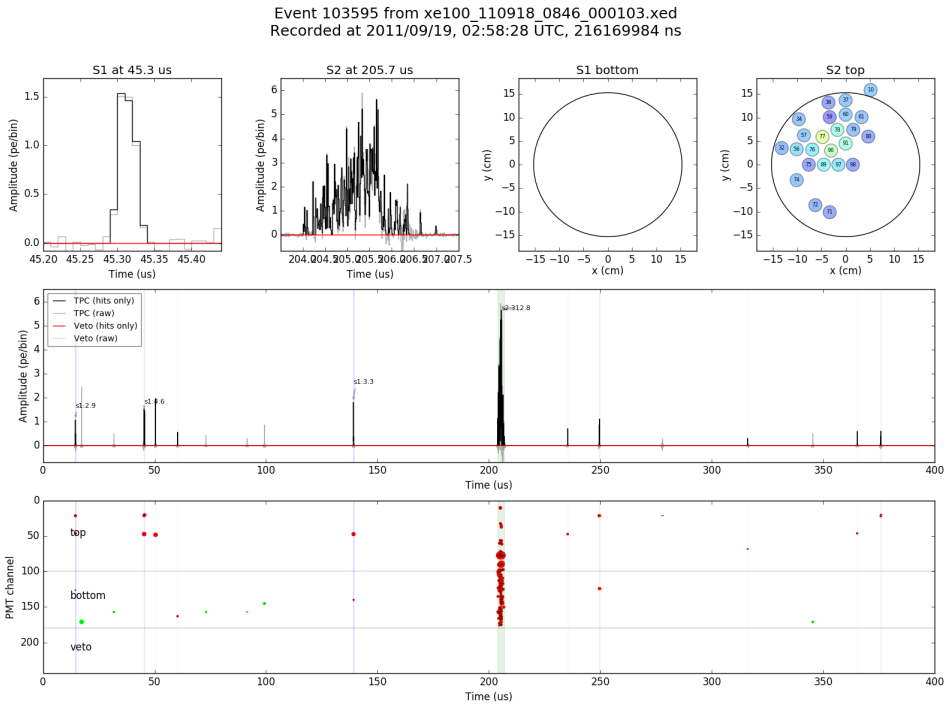

Figure 6.24: Waveform of event 103595 from the data file recorded on 18th August 2011 at 8.46 a.m. A secondary and tertiary S1 are present which did not satisfy the conditions for this event to be removed by the single S1 selection.

Points Above Upper ROI Threshold

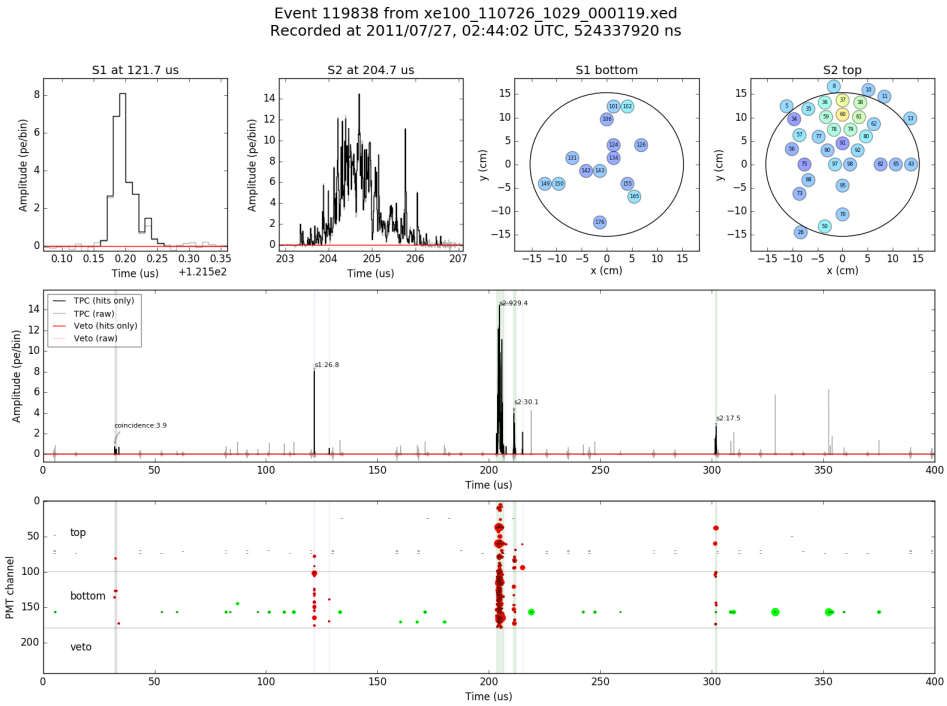

Figure 6.25: Waveform of event 119838 from the data file recorded on 26th July 2011 at 10.29 a.m. A secondary and tertiary S2 are present which were too small for this event to be removed by the S2 double scatter cut.

Points with $3\,\mathrm{pe} < cS1 < 20\,\mathrm{pe}$ Below NR Acceptance

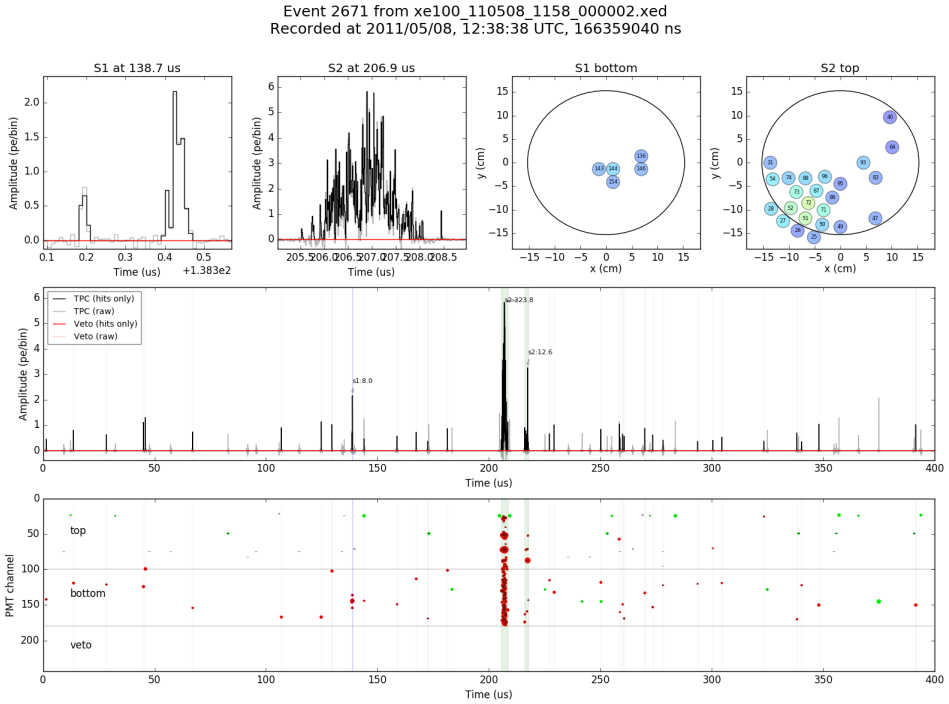

Figure 6.26: Event 119838 from the data file recorded on 8th May 2011 at 11.58 a.m. A small secondary S2 and periodic noise are present on the summed waveform.

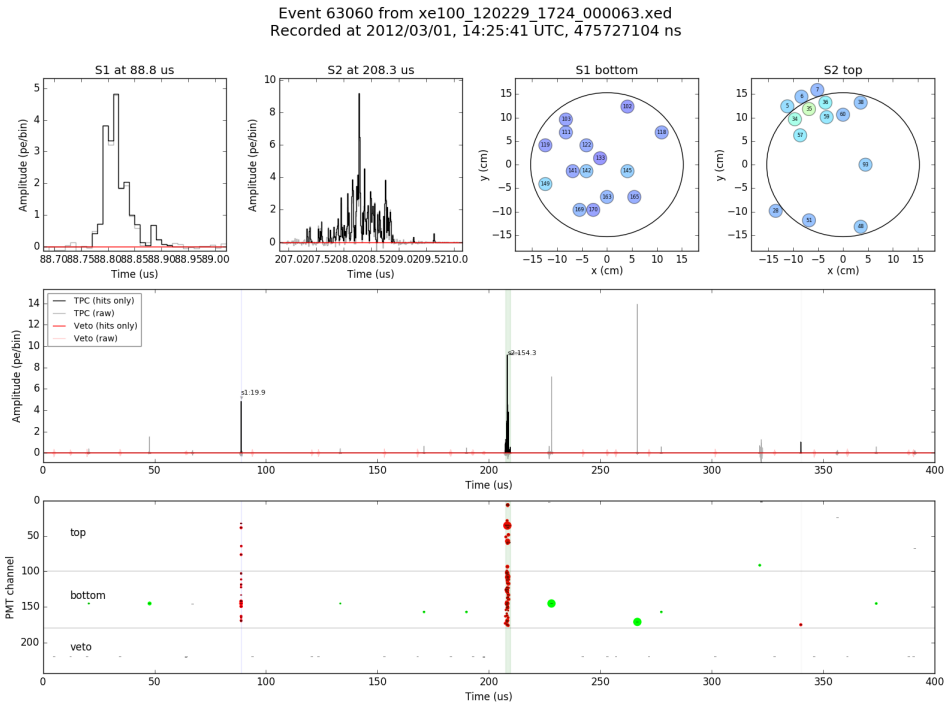

Figure 6.27: Event 63060 from the data file recorded on 29th February 2012 at 1.52 p.m. The event displays a clean waveform resembling a signal, but periodic noise in the veto and large isolated hits after the main S2 are present.

6.11 Calculations for Limit Setting

This section is based on the calculations which were demonstrated in the lecture on *Neutrino Physics and Low Energy Astroparticle Physics* by C. Weinheimer [54]. Formulae do not use natural units and include factors $\hbar$ and c. In order to get sensible results it is necessary to match units between all quantities. In this work all values have been converted to cm, GeV and seconds. Consequently one has:

$$\hbar \cdot c = 197.3 \cdot 10^{-16}\,\mathrm{GeV} \cdot \mathrm{cm}. \tag{6.13}$$

One starts with the solution of the integral over the dark matter velocity distribution. This can best be obtained in spherical coordinates [54]:

$$\begin{aligned}\left\langle \frac{1}{v} \right\rangle =& \frac{4}{\sqrt{\pi} v_c^3} \cdot \int_{v_{\min}}^{v_{\mathrm{esc}}} v \cdot e^{-\frac{v^2}{v_c^2}}\, dv, \\ =& \frac{-2}{\sqrt{\pi} v_c} \cdot \left[e^{-\frac{v^2}{v_c^2}} \right]_{v_{\min}}^{v_{\mathrm{esc}}}, \\ =& \frac{2}{\sqrt{\pi} v_c} \cdot \left(e^{-\frac{E \cdot m_A}{2\mu_A^2 v_c^2}} - e^{-\frac{v_{\max}^2}{v_c^2}} \right).\end{aligned} \tag{6.14}$$

In these equations $\rho_0 = 0.3\,\mathrm{GeV/c^2}$ is the local dark matter density, σ_0 the spin-independent WIMP-nucleon interaction cross section, $m_{\chi,A}$ the masses of the WIMP and target nucleus, and μ_A their reduced mass. The velocities are given by astrophysical constraints, namely the escape velocity of our galaxy $v_{\mathrm{esc}} \approx 544\,\mathrm{km/s}$ and the velocity of the sun around the galactic center $v_c \approx 220\,\mathrm{km/s}$ [7, 54]. The minimum velocity of WIMPs above the detection threshold is given as:

$$v_{\min} = \sqrt{\frac{2 m_A \cdot E}{2\mu_A^2}}. \tag{6.15}$$

In a next step one has to integrate the differential dark matter rate per unit mass and time. This integral includes the nuclear form factor $F^2(q^2)$ which describes the loss of coherence at higher momentum transfers. In the picture of particle waves a higher momentum transfer translates to a lower De Broglie wavelength and, accordingly, to a higher resolution. The particle starts to "see" individual nucleons rather than a uniform nucleus. Staying in the optics analogy the nuclear form factor is the Fourier transform of the nucleon distribution similar to a diffracted image in the far field being the Fourier transform of the diffracting structure. Assuming a spherical nucleus one can approximate the form factor by

$$F^2(q^2) = e^{-\frac{1}{6\hbar^2 c^2} \cdot r^2 q^2} \tag{6.16}$$

where r is the nuclear radius

$$r = 1.2\,\mathrm{fm} \cdot A^{\frac{1}{3}} \tag{6.17}$$

and q^2 is the maximum non-relativistic momentum transfer

$$q^2 = 2m_A \cdot E \tag{6.18}$$

which can be derived from the kinematics of elastic scattering. Inserting these factors yields the differential rate

$$\frac{\mathrm{d}R}{\mathrm{d}E} = \frac{\rho_0 \sigma_0}{\mu_A^2 m_\chi c^2 \cdot \sqrt{\pi} v_c} \cdot e^{-\frac{r^2 m_A \cdot E}{3\hbar^2 c^2}} \cdot (e^{-\frac{E \cdot m_A}{2\mu_A^2 v_c^2}} - e^{-\frac{v_{\mathrm{esc}}^2}{v_c^2}}). \tag{6.19}$$

One makes the substitutions:

$$\begin{aligned} a = &\ \tfrac{m_A \cdot r^2}{3\hbar^2 c^2} + \tfrac{m_A c^2}{2\mu_A^2 v_c^2}, \\ b = &\ \tfrac{m_A \cdot r^2}{3\hbar^2 c^2}, \\ c = &\ \tfrac{v_{\mathrm{esc}}^2}{v_c^2}. \end{aligned} \tag{6.20}$$

This gives a simple integral over the rate in the energy bounds of the dark matter search region:

$$\begin{aligned} R &= \textstyle\int_{E_{\min}}^{E_{\max}} \frac{\sigma_0 \rho_0}{\mu_A^2 m_\chi c^2 \cdot \sqrt{\pi} v_c} \cdot (e^{-aE} - e^{-bE-c}), \\ &= \tfrac{\sigma_0 \rho_0}{\mu_A^2 m_\chi c^2 \cdot \sqrt{\pi} v_c} \cdot \left[-\tfrac{e^{-aE}}{a} + \tfrac{e^{-bE-c}}{b} \right]_{E_{\min}}^{E_{\max}}. \end{aligned} \tag{6.21}$$

The equation is rearranged for the interaction cross-section of a WIMP with the entire nucleus

$$\sigma_0 = \frac{\sqrt{\pi} R \cdot \mu_A^2 m_\chi c^2 v_c}{\rho_0} \cdot \left(\frac{e^{-aE_{\min}} - e^{-aE_{\max}}}{a} + \frac{e^{-bE_{\max}-c} - e^{-bE_{\min}-c}}{b} \right)^{-1}. \tag{6.22}$$

In order to compare different experiments it is necessary to convert this into the WIMP-nucleon cross-section by

$$\sigma_{\mathrm{nucleon}}^{\mathrm{SI}} = \frac{\sigma_0 \cdot m_N^2}{\mu_A^2 \cdot A^2}. \tag{6.23}$$

It should be kept in mind that the above is an approximation that is only true for $m_\chi \gg m_N$ where m_N is the nucleon mass.

CHAPTER

SEVEN

BIBLIOGRAPHY

[1] J. C. Kapteyn. First Attempt at a Theory of the Arrangement and Motion of the Sidereal System. *Astrophys. J.*, 55:302–328, 1922.

[2] J. H. Oort. The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems. *Bulletin of the Astronomical Institutes of the Netherlands*, 6:249, August 1932.

[3] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta*, 6:110–127, 1933.

[4] V.C. Rubin, N. Thonnard, and J. Ford. Extended rotation curves of high-luminosity spiral galaxies. *Astrophys. J*, 225:L107, 1978.

[5] Emily E. Richards et al. Baryonic distributions in the dark matter halo of NGC 5005. *Mon. Not. Roy. Astron. Soc.*, 449(4):3981–3996, 2015, 1503.05981.

[6] C. Grupen. *Astroparticle physics*. Springer-Verlag GmbH, 2005.

[7] Teresa Marrodán Undagoitia and Ludwig Rauch. Dark matter direct-detection experiments. *J. Phys.*, G43(1):013001, 2016, 1509.08767.

[8] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *Astrophys. J.*, 648:L109–L113, 2006, astro-ph/0608407.

[9] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *Astrophys. J. Suppl.*, 208:19, October 2013, 1212.5226.

[10] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016, 1502.01589.

[11] Mark Vogelsberger, Shy Genel, Volker Springel, Paul Torrey, Debora Sijacki, Dandan Xu, Gregory F. Snyder, Dylan Nelson, and Lars Hernquist. Introducing the Illustris

Project: Simulating the coevolution of dark and visible matter in the Universe. *Mon. Not. Roy. Astron. Soc.*, 444(2):1518–1547, 2014, 1405.2921.

[12] Lauren Anderson et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples. *Mon. Not. Roy. Astron. Soc.*, 441(1):24–62, 2014, 1312.4877.

[13] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005, hep-ph/0404175.

[14] M. G. Aartsen et al. Searches for Sterile Neutrinos with the IceCube Detector. *Phys. Rev. Lett.*, 117(7):071801, 2016, 1605.01990.

[15] J. Silk et al. *Particle Dark Matter: Observations, Models and Searches*. Cambridge, UK: Univ. Pr., 2010.

[16] J. D. Lewin and P. F. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astropart. Phys.*, 6:87–112, 1996.

[17] E. Aprile et al. The XENON100 Dark Matter Experiment. *Astropart. Phys.*, 35:573–590, 2012, 1107.2155.

[18] Gary J. Feldman and Robert D. Cousins. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev.*, D57:3873–3889, 1998, physics/9711021.

[19] S. Yellin. Finding an upper limit in the presence of unknown background. *Phys. Rev.*, D66:032005, 2002, physics/0203002.

[20] E. Aprile et al. Likelihood Approach to the First Dark Matter Results from XENON100. *Phys. Rev.*, D84:052003, 2011, 1103.0303.

[21] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011, 1007.1727. [Erratum: Eur. Phys. J.C73,2501(2013)].

[22] E. Aprile et al. XENON100 Dark Matter Results from a Combination of 477 Live Days. 2016, 1609.06154.

[23] D. S. Akerib et al. Results from a search for dark matter in the complete LUX exposure. 2016, 1608.07648.

[24] Andi Tan et al. Dark Matter Results from First 98.7 Days of Data from the PandaX-II Experiment. *Phys. Rev. Lett.*, 117(12):121303, 2016, 1607.07400.

[25] E. Aprile and T. Doke. Liquid Xenon Detectors for Particle Physics and Astrophysics. *Rev. Mod. Phys.*, 82:2053–2097, 2010, 0910.4956.

[26] E. Aprile et al. Removing krypton from xenon by cryogenic distillation to the ppq level. 2016, 1612.04284.

[27] E. Aprile et al. Search for Two-Neutrino Double Electron Capture of ^{124}Xe with XENON100. 2016, 1609.03354.

[28] N. Ackerman et al. Observation of Two-Neutrino Double-Beta Decay in ^{136}Xe with EXO-200. *Phys. Rev. Lett.*, 107:212501, 2011, 1108.4193.

[29] E. Aprile, K. L. Giboni, P. Majewski, K. Ni, and M. Yamashita. Observation of Anti-correlation between Scintillation and Ionization for MeV Gamma-Rays in Liquid Xenon. *Phys. Rev.*, B76:014115, 2007, 0704.1118.

[30] E. Aprile et al. Analysis of the XENON100 Dark Matter Search Data. *Astropart. Phys.*, 54:11–24, 2014, 1207.3458.

[31] M. Szydagis, N. Barry, K. Kazkaz, J. Mock, D. Stolp, M. Sweany, M. Tripathi, S. Uvarov, N. Walsh, and M. Woods. NEST: A Comprehensive Model for Scintillation Yield in Liquid Xenon. *JINST*, 6:P10002, 2011, 1106.1613.

[32] Gaudiman. Skizze zur Funktionsweise einer Xenon dual-phase TPC, 2016, "`https://commons.wikimedia.org/wiki/File:2phaseTPC_b.jpg`".

[33] Hamamatsu Photonics. *Photomultiplier Tubes. Basics and Applications*. Hamamatsu Photonics K. K., 2007.

[34] E. Aprile et al. Dark Matter Results from 225 Live Days of XENON100 Data. *Phys. Rev. Lett.*, 109:181301, 2012, 1207.5988.

[35] J. Aalbers, C. Tunnell, and B. Pelssers. PAX: The XENON1T raw data processor. XENON1T inauguration poster, November 2015.

[36] J. Aalbers, C. Tunnell, and B. Pelssers. PAX documentation, 2016.

[37] E. Aprile et al. Low-mass dark matter search using ionization signals in XENON100. *Phys. Rev.*, D94(9):092001, 2016, 1605.06262.

[38] E. Aprile et al. Search for Event Rate Modulation in XENON100 Electronic Recoil Data. *Phys. Rev. Lett.*, 115(9):091302, 2015, 1507.07748.

[39] E. Aprile et al. Exclusion of Leptophilic Dark Matter Models using XENON100 Electronic Recoil Data. *Science*, 349(6250):851–854, 2015, 1507.07747.

[40] E. Aprile et al. First Axion Results from the XENON100 Experiment. *Phys. Rev.*, D90(6):062009, 2014, 1404.1455.

[41] E. Aprile et al. Limits on spin-dependent WIMP-nucleon cross sections from 225 live days of XENON100 data. *Phys. Rev. Lett.*, 111(2):021301, 2013, 1301.6620.

[42] E. Aprile et al. Dark Matter Results from 100 Live Days of XENON100 Data. *Phys. Rev. Lett.*, 107:131302, 2011, 1104.2549.

[43] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth.*, A389:81–86, 1997.

[44] Richard B. Firestone, Coral M. Baglin, and S.Y.F. Chu. *Table of isotopes*. Number Bd. 4. Wiley, 1999.

[45] K. Ni, R. Hasty, T. M. Wongjirad, L. Kastens, A. Manzur, and D. N. McKinsey. Preparation of Neutron-activated Xenon for Liquid Xenon Detector Calibration. *Nucl. Instrum. Meth.*, A582:569–574, 2007, 0708.1976.

[46] M.J. Berger, J.H. Hubbell, S.M. Seltzer, J. Chang, J.S. Coursey, R. Sukumar, D.S. Zucker, and PML Radiation Physics Division) Olsen, K. (NIST. XCOM: Photon Cross Sections Database. NIST Standard Reference Database 8 (XGAM), 2016.

[47] Alexander Kish. *Dark Matter Search with the XENON100 Experiment*. PhD thesis, Zurich U., 2011.

[48] XENON Collaboration. Internal communication. 2011-2016.

[49] Tadayoshi Doke, Akira Hitachi, Jun Kikuchi, Kimiaki Masuda, Hiroyuki Okada, and Eido Shibamura. Absolute scintillation yields in liquid argon and xenon for various particles. *Japanese Journal of Applied Physics*, 41(3R):1538, 2002.

[50] D. S. Akerib et al. Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics. 2016, 1608.05381.

[51] Kaixuan Ni. *Development of a Liquid Xenon Time Projection Chamber for the XENON Dark Matter Search*. PhD thesis, Columbia U. (main), 2006.

[52] M. Auger et al. Search for Neutrinoless Double-Beta Decay in ^{136}Xe with EXO-200. *Phys. Rev. Lett.*, 109:032505, 2012, 1205.5608.

[53] D. S. Akerib et al. Signal yields, energy resolution, and recombination fluctuations in liquid xenon. 2016, 1610.02076.

[54] Christian Weinheimer. Neutrino Physics and Low Energy Astroparticle Physics. Lecture Notes.

[55] E. Aprile et al. Physics reach of the XENON1T dark matter experiment. *JCAP*, 1604(04):027, 2016, 1512.07501.

[56] E. Aprile et al. Likelihood Approach to the First Dark Matter Results from XENON100. *Phys. Rev.*, D84:052003, 2011, 1103.0303.

DANKSAGUNG

Ein gewaltiges "Danke!" an alle, die beim Gelingen dieser Arbeit geholfen haben!

Zuallererst möchte ich mich bei Prof. Dr. Christian Weinheimer bedanken, der es mir ermöglicht hat an einem spannenden Thema zu arbeiten, bei dem ich nicht nur viel über die Suche nach dunkler Materie lernen, sondern an dem ich wachsen konnte. Nicht zuletzt ist das seiner immer gut dosierten Kritik und Ermutigung zu verdanken, mit denen am Ende noch für jede auftretende Schwierigkeit eine Lösung gefunden werden konnte. Herzlichen Dank möchte ich auch für die stete Unterstützung aussprechen, die mir die Teilnahme an einer Sommerschule in Obertrubach, ein Sommerstudienprogramm am GSI Helmholtzzentrum für Schwerionenforschung und einen USA-Aufenthalt über den DAAD ermöglicht hat, um nur einige von vielen tollen Erfahrungen zu nennen. Ich hatte eine sehr gute Zeit in der Arbeitsgruppe und freue mich auf die weitere Zusammenarbeit in den nächsten Jahren.

Herrn Prof. Dr. Alexander Kappes danke ich für die Übernahme der Zweitkorrektur dieser Arbeit.

Professor Rafael Lang gebührt ein enormes Dankeschön, denn ich bin nicht sicher, ob ich ohne die Arbeit in seinem Labor zu diesem wissenschaftlichen Abenteuer gekommen wäre, das wir die Suche nach der dunklen Materie nennen. Vielen Dank für das Willkommen sein, den direkten Draht und die mir in bester Erinnerung bleibende Zeit an der Purdue University! Cheers!

Alexander Fieguth hatte es nicht immer leicht in den letzten Monaten, denn er musste nicht nur meine Auseinandersetzungen mit ROOT, Python und Linux aushalten, sondern auch meinen Humor. Für die Ruhe, Besonnenheit und den eigenen Witz, mit denen er das gemeistert hat, möchte ich ihm genau so danken wie für seine Expertise. Ein dickes Danke muss natürlich auch für das gewissenhafte und schonungslose Korrekturlesen ausgesprochen werden. Merci!
Für die Anmerkungen und Kommentare bedanke ich mich ebenfalls herzlich bei Michael Murra, Lutz Althüser und Miguel Ángel Vargas Jara. Sie haben diese Arbeit lesbarer

gemacht, als sie sonst hätte werden können. Cheers aber auch für den alltäglichen Wahnsinn im Büro.
An Dr. Stephan Rosendahl und Christian Huhmann, von denen ich so einige Handgriffe und Kniffe im Labor lernen durfte, geht ein besonderes Dankeschön.

Dennis Niemeier darf ich an dieser Stelle auf keinen Fall vergessen. Seit Anfang des Studiums haben wir Mathevorlesungen, Praktika und Klausuren zusammen durchgestanden und, wenn es drauf ankam, dann konnte man sich immer auf dich verlassen. Danke dafür! Jetzt wünsche ich dir alles erdenklich Gute für die Promotion in Dresden. Wir sehen uns! Das gilt auch für Benedikt und Alina. Wie schnell fünf Jahre vorbeigehen, wenn man gut beschäftigt ist. Die Kaffeepausen, Mensaessen und gemeinsamen Unternehmungen waren da immer eine gelegene Ablenkung, wenn im Labor oder am Bildschirm gerade nichts mehr ging. Danke, dass ihr da wart!

Von Herzen möchte ich meinen Eltern, Beate und Volker Wittweg, und meiner Großmutter, Ursula Wehrmann, danken, die mir mit ihrem Vertrauen, ihrem Zuspruch und ihrer Unterstützung ein erfolgreiches Studium ermöglicht haben.

Der Schlusssatz dieser Dankesworte gebührt Juliane. Du bist da und das ist alles, was ich brauch. Danke!