Department of Physics and Astronomy
Heidelberg University

Bachelor Thesis in Physics
submitted by

Julian Kusch

born in Homburg (Germany)

April 2024

Position reconstruction of scattering events in the TPC of the DARWIN observatory to improve the radiogenic neutron background separation

This Bachelor Thesis has been carried out by Julian Kusch at the
Physikalisches Institut in Heidelberg
under the supervision of
Prof. Dr. Stephanie Hansmann-Menzemer

Abstract

The existence of Weakly Interacting Massive Particles (WIMPs), which are promising candidates for Dark Matter, remains an unresolved question in modern physics. The DARWIN observatory is a future experiment planned to detect the scattering of WIMPs in liquid xenon. This work deals with the reconstruction of the horizontal position of such a scattering process within the time projection chamber of the DARWIN detector. Based on Monte Carlo simulation, the position can be reconstructed with an accuracy of 2 mm. Furthermore, the reconstruction algorithm can be used to identify 89% of the radiogenic neutron background with a signal acceptance of 90%. In comparison to the previous identification, 29% of the radiogenic neutrons can be additionally rejected.

Zusammenfassung

Die Existenz von Weakly Interacting Massive Particles (WIMPs), welche vielversprechende Kandidaten für Dunkle Materie sind, ist weiterhin eine ungelöste Frage der modernen Physik. Das DARWIN Observatorium ist ein zukünftiges Experiment, mit dem die Streuung von WIMPs in flüssigem Xenon detektiert werden soll. Diese Arbeit beschäftigt sich mit der Rekonstruktion der horizontalen Position eines solchen Streuprozesses innerhalb der Zeitprojektionskammer des DARWIN Detektors. Basierend auf Monte-Carlo-Simulationen kann die Position mit einer Auflösung von 2 mm rekonstruiert werden. Darüber hinaus kann der Rekonstruktionsalgorithmus genutzt werden, um 89% des radiogenen Neutronenhintergrunds bei einer Signalakzeptanz von 90% zu identifizieren. Verglichen mit der vorherigen Identifikation kann so der radiogene Neutronenhintergrund zusätzlich um 29% reduziert werden.

Contents

Chapter 1

Introduction

Numerous observations of the universe provide evidence for the existence of Dark Matter, indicating that physics beyond the Standard Model of Particle Physics exists. Indeed, current investigations by the Planck collaboration [1] show, that Baryonic matter makes up only 5% of the energy content in the universe, whereas 26% is made up of Dark Matter and the remaining fraction belongs to Dark Energy. One promising model for the microscopic nature of Dark Matter are Weakly Interacting Massive Particles (WIMPs). Therefore, liquid-xenon detectors are built to detect the possible scattering processes of WIMPs with xenon atoms. Current experiments like XENONnT in Gran-Sasso or the LZ experiment at the Sanford Lab are operating and setting limits for the direct detection cross section of WIMPs. To increase the sensitivity for direct Dark Matter detection further, the construction of a detector with around six times more liquid xenon is planned, namely the DARWIN observatory, which will operate with a prospective total amount of 50 tons of liquid xenon.

This thesis presents the implementation and performance of algorithms that are used to reconstruct the x-y-position of a scattering process in the DARWIN time projection chamber. The main reconstruction algorithm is based on χ^2 minimization. Its performance is tested on Monte Carlo data sets. The reconstructed position is required for the corrections of the scattering signals. These corrections take the spatial-dependent detector efficiencies into account and therefore correct the corresponding signal strength. The χ^2 algorithm of the position reconstruction can also be used to reduce the neutron background. The identification of neutron scatters plays a significant role in background reduction since neutrons can interact in the same way as WIMPs with xenon atoms.

The algorithms implemented in this thesis are based on the ones developed by Yuan Mei [2] for the XENON100 experiment and on the analysis of the position reconstruction in XENON100 and XENON1T by Bert Pelssers [3].

Chapters 2 and 3 present evidence and models for Dark Matter and explain the detection principle of the DARWIN observatory. The position reconstruction algorithms are explained in Chapter 4 and their performances are presented in Chapter 5. The reduction of the neutron background is studied in Chapter 6. Chapter 7 summarizes this work and gives an outlook for further studies.

Chapter 2

Dark Matter

Although the Standard Model of Particle Physics is a well-tested theory that can predict the properties of particles with impressive precision, there are still phenomena it cannot describe. One of these phenomena is the nature and composition of Dark Matter. Measurements by the Planck collaboration in 2018 [1] show that only around 5% of the energy content in the universe is made up of Baryonic, visible matter as described in the Standard Model of Particle Physics. Actually, around 26% is Dark Matter and 69% is made up of Dark Energy. While Dark Matter has a huge gravitational impact on the structure of galaxies and clusters, Dark Energy – the even more unknown quantity – is needed to explain the accelerated expansion of the universe. The fact that Dark Matter does not or only very weakly interacts electromagnetically is the reason for the generic name Dark. The study of the behavior of galaxies and clusters provides some evidence for the existence of something which can be at the moment primarily observed through macroscopic gravitational effects. The microscopic nature of Dark Matter is one big unanswered question in physics, which opens the door for several different theories and experiments.

2.1 Evidence for Dark Matter

The first evidence for Dark Matter was found in 1932 by Jan Hendrik Oort [4] who observed the rotation velocity of stars in the Milky Way from which he derived with the virial theorem a mass density that was higher than expected from the visible matter. One year later, Fritz Zwicky [5], who also was the first person to call it "Dunkle Materie" (German for Dark Matter), measured the velocity distribution of stars in the Coma galaxy cluster and made similar observations to Oort. The scientific community paid little attention to these observations until the 70s, when Vera Rubin and Kent Ford [6] again observed an unexpectedly high velocity distribution in the Andromeda Galaxy. From there on, more evidence for the existence of Dark Matter was found with different observation methods. This section, however, focuses only on two of them.

2.1.1 Rotation curves of galaxies

Using Newtons' classical mechanics, one can derive the velocity $v(r)$ of a body with mass m in a gravitational bound orbit to be

$$G\frac{M(r)m}{r^2} = \frac{mv^2}{r} \quad \Leftrightarrow \quad v(r) = \sqrt{\frac{GM(r)}{r}}, \tag{2.1}$$

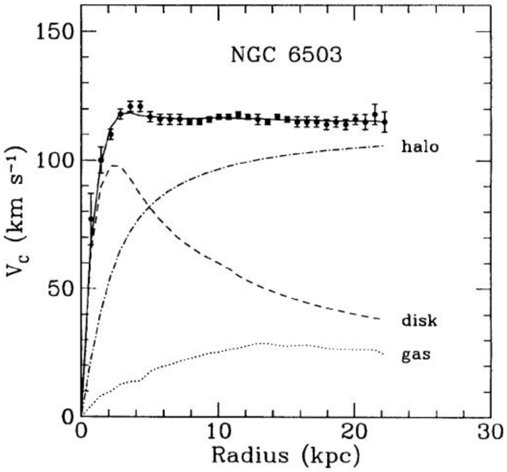


Figure 2.1: The solid curve shows the fit to the rotation curve of the dwarf spiral galaxy NGC 6503 [7]. The contributions of visible matter (disk, dashed), the gas in the galaxy (gas, dotted) and also the contribution of Dark Matter halo (halo, dashed-dotted) can be seen.

where $M(r)$ is the mass distribution in the galaxy dependent on the radial distance r to the center and G being the gravitational constant. In the absence of Dark Matter, most of the mass in the galaxy should be concentrated in the center, so that for larger radii $M(r)$ stays constant, which would lead to the proportionality $v \sim r^{-1/2}$.

However, multiple observations show a different behavior for the velocity distribution of stars. Figure 2.1 shows the rotation curve of the dwarf spiral galaxy NGC 6503 measured by Begeman et al. [7]. It can be seen that the velocity does not fall proportional to $r^{-1/2}$ but rather remains constant. This would require a mass distribution of the shape $M(r) \sim r$ to get a constant behavior of the rotation curve. Thus, in addition to visible matter, further mass is needed, which is called Dark Matter. The plot also shows the individual components of the rotation curve: the contribution of visible matter, of gas in the galaxy and the contribution of the Dark Matter halo of the galaxy.

2.1.2 Cosmic microwave background

The cosmic microwave background (CMB) was accidentally discovered by Arno Penzias and Robert W. Wilson in 1964. The CMB is a relic from a time shortly after the Big Bang and almost perfectly follows the spectrum of a black body with the temperature of $(2.725\,48 \pm 0.000\,57)\,\mathrm{K}$ [8]. The spatial dependent temperature fluctuations of the CMB, which can be seen in Figure 2.2, can be decomposed into spherical harmonics. The angular power spectrum of the fluctuations describes the contributions of the different multipole moments on the temperature fluctuations. By performing a fit to the angular power spectrum, one can use the multipole moments of the peaks in the spectrum to calculate the amount of Baryonic matter, Dark Matter and Dark Energy in the universe.

Table 2.1: Energy and matter composition of the universe [1].

Quantity	**Value**	**Explanation**
Ω_Λ	0.6847 ± 0.0073	Energy fraction
Ω_m	0.3153 ± 0.0073	Matter fraction
$\Omega_b h^2$	0.02237 ± 0.00015	Physical baryon density parameter
$\Omega_c h^2$	0.1200 ± 0.0012	Physical Dark Matter density parameter

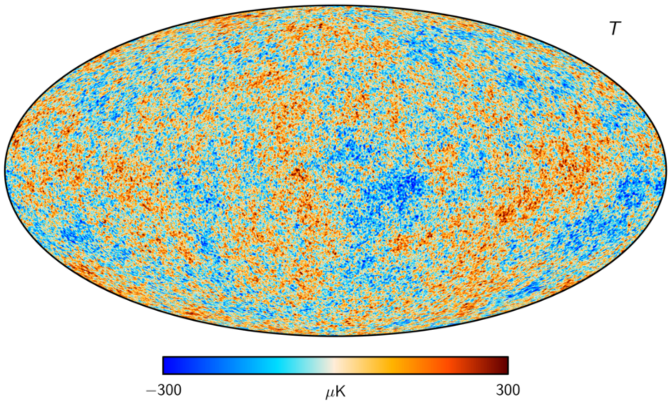


Figure 2.2: Temperature fluctuations of the CMB [9].

The first two values in Table 2.1, taken from the Planck collaboration from 2018, show that around 68% of the energy content in the universe is made up of Dark Energy, while the remaining 32% are Dark Matter and Baryonic matter. The matter and energy densities in the universe are defined such that $\Omega_i = \frac{\rho_i}{\rho_c}$, with the critical energy density ρ_c and $i \in \{\Lambda, m\}$. Here, ρ_c is set to the value at which the universe has a flat geometry. The actual measured value (from the fit of the power spectrum) is consistent with the theoretical value, yielding the relation $\Omega_\Lambda + \Omega_m = 1 \equiv \Omega_{tot}$. The fraction of Dark Matter, Ω_c, out of the total amount of matter, Ω_m, is scaled by the square of the reduced Hubble constant h, which is defined by

$$h = \frac{H_0}{100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}} = 0.674, \tag{2.2}$$

with the Hubble constant $H_0 = 67.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Therefore, the ratio of Dark Matter out of the whole matter in the universe amounts to 26% and only 5% is made up of Baryonic matter.

2.2 The WIMP model

As it has so far been ruled out [10] that Dark Matter solely exists in the form of brown dwarfs or black holes, research into the particle properties of Dark Matter has become increasingly important. In order to describe the nature of Dark Matter, a massive particle would have to exist, which until now could only be observed through gravitational interaction. Since the only possible particle of the Standard Model is the neutrino, which could only account for a small fraction of the total mass of Dark Matter, it stands to reason that there are particles (and possibly also forces) outside the Standard Model. An overview of the different possible Dark Matter particles can be seen in Figure 2.3. One promising candidate are Weakly Interacting Massive Particles (WIMPs). WIMPs are promising because they have not been proposed to explain Dark Matter, but rather arise naturally from theories beyond the Standard Model. The following overview of WIMPs is based on [10].

WIMPs are electrically neutral with masses in the 1 to 10^5 GeV range and were thermally created in the early phase of the universe with an interaction scale of 10^{-41} to 10^{-51} cm^2. They only interact weakly, gravitational or maybe through a force that is not yet included in the Standard Model. Their relatively low self-annihilation cross section leads to the property that WIMPs are now "frozen-out", making them candidates for Cold Dark Matter. Due to their high mass, they are moving at non-relativistic velocities, which leads to the assumption that they tend to clump together since their low

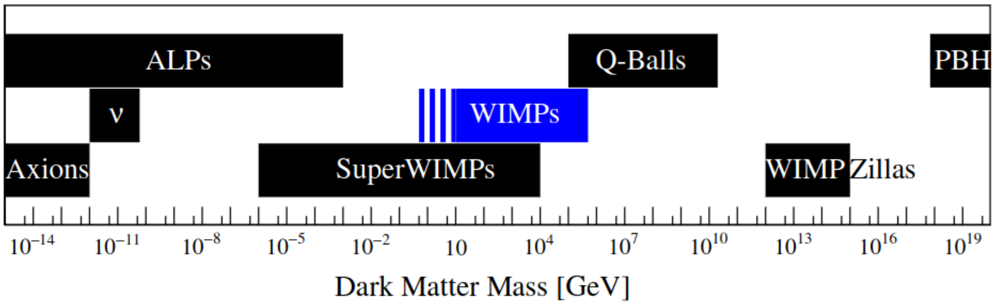


Figure 2.3: Mass ranges of the different Dark Matter particle candidates, modified from [10].

velocity can not overcome the gravitational forces. A favorable property of WIMP theory is, that the current observed amount of Dark Matter in the universe can be derived with the properties described above. To detect these weakly interacting particles, very sensitive experiments have to be built.

2.3 Search for Dark Matter particles

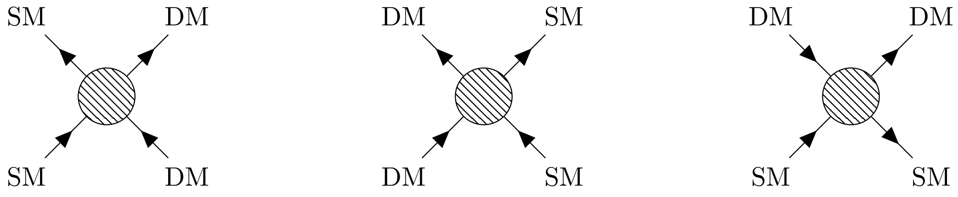
Since the mass of possible Dark Matter particles has a vast range (as seen in Figure 2.3) and different types of interaction with Baryonic matter may exist, there are numerous experiments to study these different mass ranges and interactions. In general, there are three possible interaction types of Dark Matter with the matter known from the Standard Model. The Feynman diagrams of these processes can be seen in Figure 2.4. While exploiting the production and annihilation of Dark Matter are called indirect detection methods, the search for Dark Matter particles via scattering with Standard Model matter is called direct detection.

For the production of Dark Matter, particle colliders like the Large Hadron Collider (LHC) at CERN can be used to create Dark Matter out of ordinary matter. Doing this, the presence of Dark Matter would be indicated by missing momentum in the event because the momentum carried by the Dark Matter particles can not be detected. These kinds of searches have been performed for example at the LHCb [11], ATLAS and CMS experiments [12].

The annihilation of Dark Matter can be detected by looking at gamma rays originating from galaxies, which is performed for example by the Fermi Gamma-ray Space Telescope (FGST) using the Large Area Telescope (LAT). Recently, in 2023, data from 12 years of observation by Fermi-LAT has been analyzed to search for potential Dark Matter annihilation [13].

Since the earth is traversing the Dark Matter halo of our galaxy, Dark Matter could be detected by a scattering process with ordinary matter. In the Cresst (Cryogenic Rare Event Search with Superconducting Thermometers) experiment [14], the temperature increase resulting from such a scattering in a scintillating crystal is to be measured. Another direct way of detecting Dark Matter is exploited by the future DELight experiment [15]. Here, the scattering is detected using superfluid helium and the search is designed for WIMPs in the sub-GeV range.

The experiment in the focus of this thesis is the future DARWIN observatory [16]. DARWIN is planned to be a dual-phase noble gas time projection chamber in which electrons and photons caused by a scattering of WIMPs with xenon atoms can be detected. There are already smaller experiments based on this principle such as XENONnT and LZ which are now operating. The limits of the cross section for the direct search of Dark Matter using such type of detectors can be seen in Figure 2.5.



(a) Production. **(b)** Annihilation. **(c)** Scattering.

Figure 2.4: Interaction types of Dark Matter (DM) with Standard Model matter (SM). The first two diagrams describe the interaction behind indirect searches of DM candidates and the last one shows the mechanism used for the direct search.

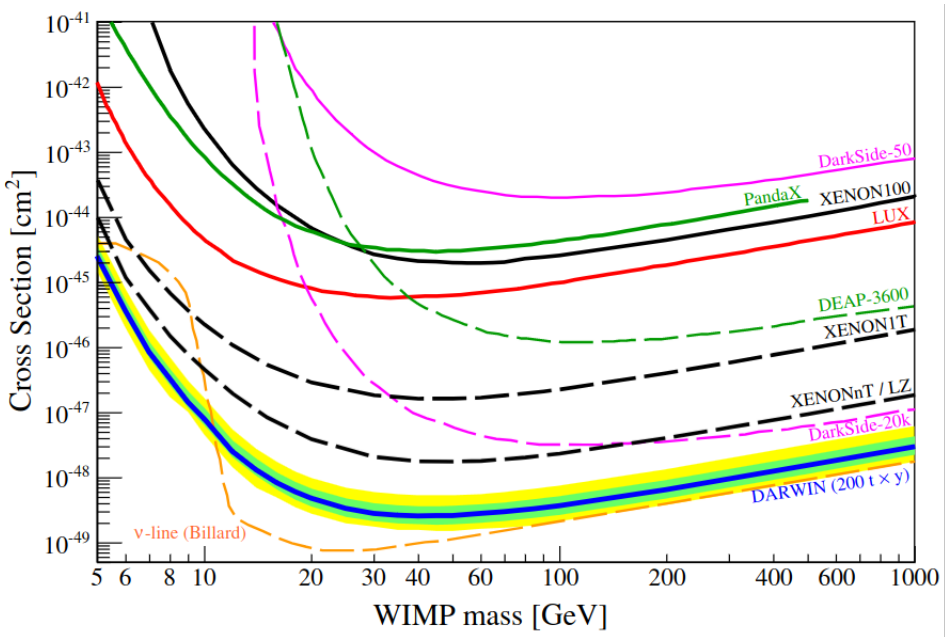


Figure 2.5: Sensitivity curves of finished (solid lines) and planned or still operating (dashed lines) experiments, which search for Dark Matter. The DARWIN observatory is around 10 times more sensitive at higher WIMP masses than the current XENONnT or LZ experiments. The sensitivity is restricted by the neutrino background (ν-line, dashed orange) since neutrinos make up the largest irreducible background. Modified from [16].

Chapter 3

The DARWIN observatory

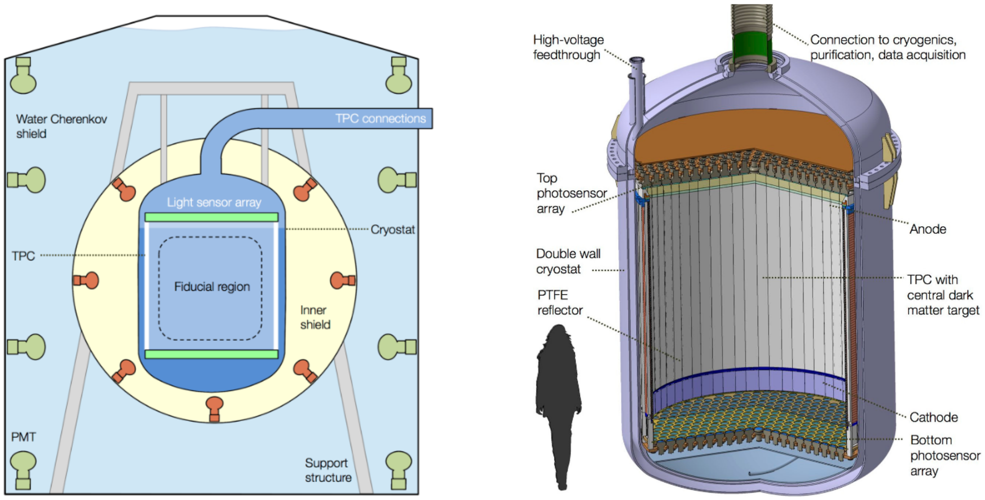
The DARWIN (DARk matter WImp search with liquid xenoN) observatory is not the first xenon detector for Dark Matter. At the moment, two experiments are taking data, namely the XENONnT Experiment in Gran-Sasso and the LUX-ZEPLIN (LZ) Experiment at the Sanford Lab. The detection principle is similar for all detectors: The WIMPs scatter elastically with the xenon nucleus in a cylindrical liquid-xenon tank. The recoil of the xenon atoms leads to the emission of photons and electrons (in detail explained in section 3.2). By analyzing these signals, information on the scattering process is provided. Since particles from the Standard Model can also scatter with xenon, the detectors are built deep underground to minimize the external background. Due to the low interaction cross section of WIMPs, the detection probability is increased by a large target and a long exposure time. One of the major improvements of the DARWIN detector, compared to the other two, will be the higher amount of xenon in the TPC. While the XENONnT and LZ experiments have eight and seven tons of xenon at their disposal, respectively, the DARWIN detector contains around 50 tons of xenon. One reason why the construction of DARWIN is on a timescale of about 10 to 20 years is that the xenon from the current experiments is reused. The significant increase in size compared to previous experiments also presents further technical hurdles. Therefore, the detector design is still under optimization. The current status of planning is given in the Whitepaper from 2016 by Aalbers et al. [16].

There are also further studies that can be performed with the DARWIN detector besides the search for WIMPs. Due to the required high sensitivity, axions (further candidates for Dark Matter), solar neutrinos (which make up a large part of the background) and also the neutrinoless double beta decay ($0\nu\beta\beta$) can be investigated.

3.1 Experimental setup

A schematic of the whole detection system can be seen in Figure 3.1a. The outermost part is a water Cherenkov muon veto system to detect cosmic muons. The design of the neutron veto (in Figure 3.1a denoted as Inner shield) is under further investigation.

The core of the DARWIN detector is the time projection chamber (TPC) which can be seen in Figure 3.1b. The TPC has a cylindrical shape with a radius of 1.3 m and a height of around 2.6 m. The DARWIN TPC is a dual-phase xenon TPC meaning that the major part of the TPC is filled with liquid xenon (LXe) and a thin layer at the top of the TPC is filled with gaseous xenon. To operate the TPC with LXe, the xenon has to be cooled down to 165 K inside a large cryostat. Since there is xenon pumped through the cryostat system, the active target mass amounts to 40 tons, while the



(a) Sketch of the whole detection system.　**(b)** Sketch of the TPC.

Figure 3.1: The DARWIN detector [16].

whole detector operates with 50 tons of xenon.

At the bottom and the top of the LXe are a cathode and a gate electrode, generating an electric drift field of $0.2\,\mathrm{keV\,cm^{-1}}$ (baseline value). Electrons originating from the recoil process of a particle with the xenon in the TPC (in detail explained in section 3.2) drift towards the top of the TPC, passing the gate and entering the xenon gas phase of the TPC. The gas phase is a few mm thick and the electric field is significantly higher than the drift field with $10\,\mathrm{kV\,cm^{-1}}$. At the bottom and the top of the detector are the photomultiplier tubes (PMTs). Each array has a radius of 1300 mm and contains 955 PMTs. The PMTs also have a circular shape with a radius of 38 mm. The walls of the TPC and the space between the PMTs are coated with PTFE (also known as Teflon) to reflect the scattering light.

3.2 Detection principle of DARWIN

Not only WIMPs but also particles like neutrons, neutrinos and Gammas can scatter with the xenon atoms and can generate a scintillation signal and an ionization signal, each with a wavelength of 175 nm. This section explains the basic origin of the signals based on [2].

The processes behind these two signals can be seen in Figure 3.2. There are two states formed after the elastic scattering process of a particle with a xenon atom: one excited Xe^* state and one ionized $\mathrm{Xe}^+ + \mathrm{e}^-$ state. While the excited state recombines to an excimer ${\mathrm{Xe}_2}^*$ with a Xe atom, the ionized Xe^+ recombines to an ionized dimer ${\mathrm{Xe}_2}^+$. The excimer ${\mathrm{Xe}_2}^*$ decays into two ground-state xenon atoms and emits UV light. The previously formed ionized dimer ${\mathrm{Xe}_2}^+$ captures an electron from the earlier ionization process and forms an excimer ${\mathrm{Xe}_2}^*$ which decays in the same way as the excimer before. Additionally, this time also phonons (heat) are produced which, however, can not be detected by the DARWIN detector. These two processes emit the so-called scintillation light, $S1$. The capture of the previously ionized electrons is suppressed by the drift field in the TPC. This field ensures that a part of the ionized electrons drift towards the gas phase. After entering the gas phase, the electrons excite the gaseous xenon atoms to excimers because of the high extraction field. These subsequently decay and emit photons. These photons form the ionization signal, $S2$. Both signals are detected by the PMTs.

There are two important reasons for using xenon. Firstly, xenon is transparent to its own scintillation

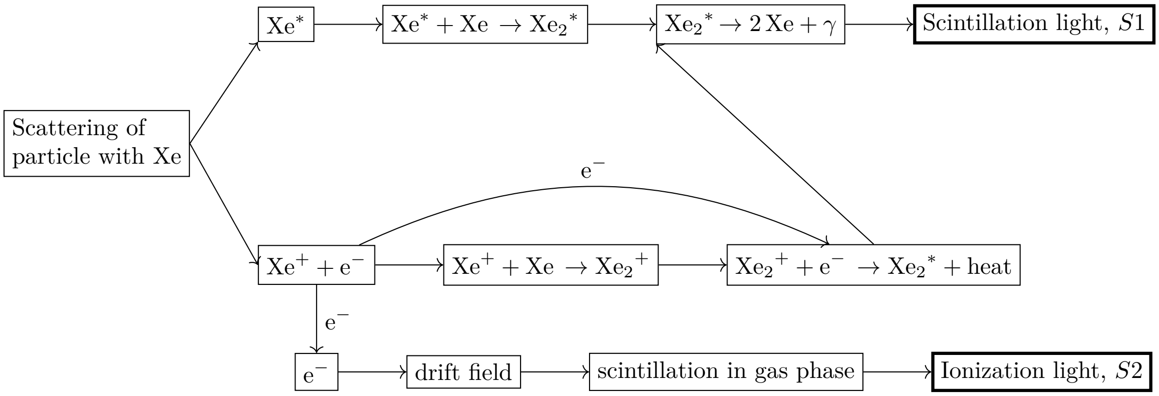


Figure 3.2: Sources of scintillation and ionization light. Both states Xe^* and $\mathrm{Xe}^+ + \mathrm{e}^-$ are formed after a scattering, whereby the respective number of states depends on the type of recoil, which is explained in Section 3.2.1.

light because photons are created through the decay of an excimer and their energy does not match any difference of energy levels in a Xe atom. Secondly, the high density of $3\,\mathrm{g\,cm^{-3}}$ leads to small interaction lengths of particles like photons and neutrons. With this property, background identification is easier because one expects these particles to scatter multiple times, preferred at the edges of the detector due to the self-shielding of LXe. This is a crucial point in background reduction because, despite all efforts, components of the detector, like the PMTs or the cryostat vessel contain radioactive trace elements due to limited radio purity. Neutrons can be emitted by $(\alpha,\, n)$ reactions from the decay of trace elements and spontaneous fission and can scatter (even multiple times) in the TPC.

The $S1$ and $S2$ signals arrive at different times at the PMTs. This time difference is caused by the time the electrons need to get from the scattering position to the gas phase, the so-called drift time. Due to the nearly homogeneous field, the drift velocity is constant and the z position of the scattering process can be extracted from the drift time and the drift velocity. For the position reconstruction in the x-y-plane, the $S2$ pattern is used. The location at which the electrons arrive in the gas phase and start to scintillate is strongly correlated to the horizontal scattering position due to the vertical drift of the ionized electrons. $S1$ photons are emitted isotropically and scatter several times before they reach the PMTs. Thus, this signal only contains less information on the x-y position.

3.2.1 Electronic and nuclear recoil

Particles can in general interact through two different types of scattering processes: electronic recoil (ER) and nuclear recoil (NR). Gamma rays, β particles and also neutrinos mainly interact with the electrons of the xenon atom. Therefore, more electrons are ionized and drift toward the gas phase, which leads to a large $S2$ signal for electronic recoils. Nuclear recoil can be induced by WIMPs, neutrinos and also by neutrons. They mainly interact with the nucleus and thus fewer electrons are ionized, leading to a smaller $S2$ signal. Neutron-induced recoils emit the same signal as WIMPs and therefore their identification is important. WIMPs, neutrons and neutrinos can also scatter via electronic recoil. However, the energy range of these processes is not in the region of interest.

By calculating the ratio $S1/S2$, ER can be distinguished from NR. Figure 3.3 shows the location of NR and ER events in the $(cS1,\, cS2_b)$-space. The "c" indicates that these are the corrected $S1$ and $S2$ values and the "b" indicates that the $S2$ signal is only measured in the bottom array. The corrections are required to consider the detector's spatial-dependent properties. For example, there are fewer photons collected for an event that takes place at the edge or at the top of the detector,

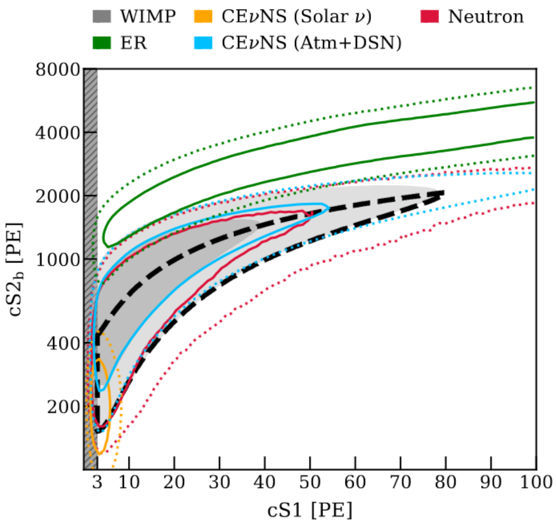


Figure 3.3: The signals of NR and ER events in the ($cS1$, $cS2_b$)-space while $S2$ is only measured in the bottom array, taken from [17]. The 1σ (solid) and 2σ (dashed) contours are shown for the background. The green ER band is higher than the gray NR 50 GeV WIMP band. It can be seen that neutrons also contribute to the NR signal as their band overlaps with the WIMP band. The bands of Coherent Elastic Neutrino-Nucleus Scattering (CEνNS) from solar neutrinos (solar ν, orange) and Atmospheric and Diffuse Supernova Neutrinos (Atm+DSN, blue) are also shown.

since the photons are scattered at the edges and at the liquid surface. More information about the associated light collection efficiency can be found in Section 4.2.1. As it can be seen in Figure 3.3, ER events (green curve) have a larger $S2$ signal than WIMP-induced NR events (gray area). The signal is measured in photoelectrons (PEs) of the PMTs. It should be noted that this is a schematic from XENONnT.

3.3 Simulation framework

The DARWIN simulation framework consists of three main parts: Geant4 [18], Tray [19] and Alea [20]. Simulating NR or ER events starts with Tray, which is based on the IceTray software from the IceCube experiment. Tray itself consists of many different modules and every module executes one task, like the computation of the energy deposit of the WIMP and the resulting number of electrons and $S1$ photons. After that, Tray simulates the drift of the electrons through the liquid xenon (LXe) until they enter the gas phase. The propagation of the $S1$ photons and the $S2$ photons, which are created by scintillation of the electrons in the gas phase, is performed with Geant4. Geant4 exploits detector properties like the reflection of the PTFE and the absorbance and the scattering in the LXe to simulate realistic photon propagation. Since this is a very time-consuming process, so-called light collection efficiency (LCE) maps are used, which are in detail explained in Section 4.2.1. In short, the LCE maps provide probabilities for emitted photons to hit certain PMTs. These probabilities are then used in Tray to simulate the actual photon hits in the PMTs, taking the detection efficiencies of the PMTs into account. After the following simulated readout process, the whole data is stored. The statistical data analysis such as the sensitivity calculation is done with Alea.

Chapter 4

Position reconstruction algorithms in the x-y-plane

The goal of this chapter is the presentation of the two different algorithms used for the reconstruction of the x-y-position of a scattering event, which are based on the $S2$-photon distribution in the top PMT array. Two different algorithms are exploited: a weighted sum algorithm and a χ^2 minimization. The χ^2 algorithm is very time-consuming, while the weighted sum is much faster. However, the χ^2 minimization is – as described in the next chapter – more accurate and also gives the possibility to get a reliable uncertainty estimate of the reconstructed position.

The position reconstruction in this thesis is based on the $S2$ signal and the algorithms need the $S2$ pattern as a measured quantity, i.e. the number of photons n_i detected in each PMT i. However, the quantity provided by the simulation is not the number of photons, but the number of photoelectrons (PEs). These are connected through the PMT efficiency via

$$n_i = \frac{n_{pe,i}}{\varepsilon_i}, \tag{4.1}$$

where $n_{pe,i}$ is the number of measured PEs in PMT i. The measured quantity in PMTs is the so-called area, which can be transferred to the number of PEs. The number of PEs is usually not an integer. The efficiency of PMT i is denoted as ε_i. Since the simulation assumes the same efficiency for every PMT, this value is constant for all PMTs with $\varepsilon_i \equiv \varepsilon$. If the algorithm is applied on real data, the PMTs could have different efficiencies. The efficiency ε is calculated with the quantum efficiency, QE, which is the efficiency of one photon triggering one photoelectron. Since one photon could also trigger two PEs, the double-photon-efficiency, DPE, has to be taken into account. Inserting the values of the simulation for QE and DPE leads to a total PMT efficiency of

$$\varepsilon = \frac{QE}{1+DPE} = \frac{0.35}{1+0.22} = 0.2869. \tag{4.2}$$

The number of photoelectrons follows a Poisson distribution with an uncertainty

$$\sigma_{n_{pe,i}} = \sqrt{n_{pe,i}}. \tag{4.3}$$

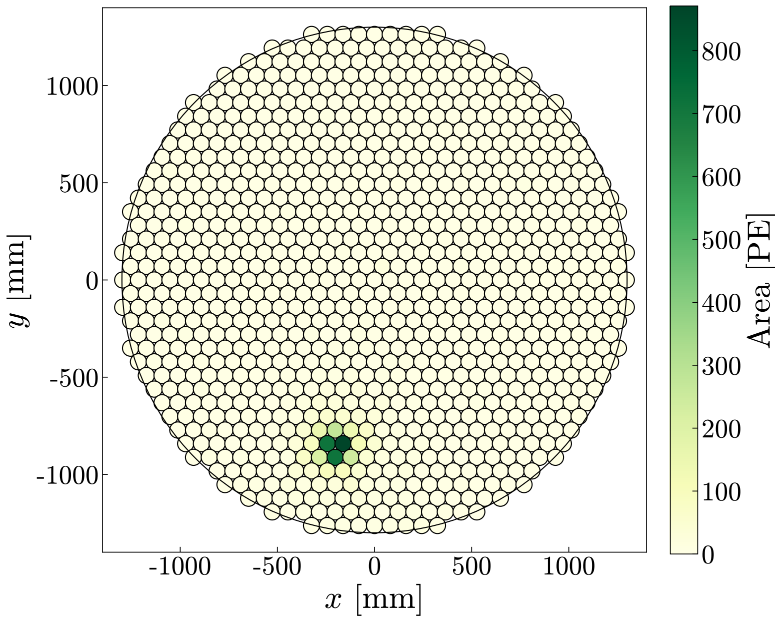


Figure 4.1: The $S2$ pattern for a 80 keV NR event in the TPC, the colormap indicates the number of measured PEs.

This leads to the following uncertainty for the number of photons:

$$\sigma_{n,i}^2 = \left(\frac{\partial n_i}{\partial n_{pe,i}} \cdot \sigma_{n_{pe,i}} \right)^2 = \frac{n_{pe,i}}{\varepsilon^2} = \frac{n_i}{\varepsilon}. \tag{4.4}$$

A typical $S2$ pattern of an NR event in the top array can be seen in Figure 4.1, where the colormap indicates the $S2$ area (in PEs). Most of the photons are distributed over a few dozen PMTs, clustered directly above the scattering position. Hits in other PMTs are due to scattered photons in the detector, for example at the LXe surface or at the PTFE on the walls and between the PMTs.

4.1 Weighted sum algorithm

The weighted sum algorithm weights the x and y positions of the PMTs with the respective photon counts, comparable to calculating the center of mass with

$$\vec{x}_{rec} = \frac{1}{N_{ph}^{top}} \sum_{\{PMT \geq t\}} n_i \cdot \vec{x}_{PMT,i}. \tag{4.5}$$

N_{ph}^{top} is the total number of $S2$ photons collected in the top array, n_i is the number of photons collected by PMT i and $\vec{x}_{PMT,i}$ is the position of PMT i. If n_{max} is the highest number of photons measured in a PMT, then the threshold, t, is set such that only PMTs are taken into account for which $n_i \geq 0.1 \cdot n_{max}$ holds. The reason for this is to avoid a bias towards the center of the TPC, which is related to random noise explained in Section 5.1.

4.2 χ^2 minimization algorithm

The χ^2 minimization is based on the regular χ^2 sum but adapted to take into account that most PMTs have very low counts. Kenneth J. Mighell did this adaption [21] who proposes a modified χ^2 distribution

$$\chi^2_\gamma = \sum_{i=1}^{N} \frac{(n_i + \min(n_i, 1) - m_i)^2}{n_i + 1}, \tag{4.6}$$

where N is the total number of observations. For observation i, n_i is the measured value and m_i is the theoretically expected value based on a certain model. The $+1$ takes into account that a common χ^2 distribution underestimates the true mean by 1 for Poissonian distributed signals (i.e. low counts) and $\sigma_i^2 \equiv n_i + 1$.

Adapting Equation 4.6 to the position reconstruction at DARWIN based on [2], leads to the so-called test statistic, $T(x, y)$,

$$T(x, y) = \sum_{\{PMT \geq 50\}} \frac{(n_i + 1 - N_{ph}^{top} p_i(x, y))^2}{\sigma_i^2}, \tag{4.7}$$

with

$$\sigma_i^2 = \frac{n_i}{\varepsilon} + \frac{N_{ph}^{top}}{\varepsilon} p_i(x, y) + 1. \tag{4.8}$$

Here, n_i is again the number of photons in PMT i, $N_{ph}^{top} = \sum n_i$ is the total number of $S2$ photons collected in the top array and $p_i(x, y)$ is the probability distribution function of PMT i, which gives the probability for PMT i to receive a photon from position (x, y) out of all received photons in the top array. The calculation of $p_i(x, y)$ is performed with light collection efficiency simulations, which are explained in Section 4.2.1.

The uncertainty, σ_i, in Equation 4.8 is calculated by first considering the uncertainty on the measured value, as in Equation 4.4. The model uncertainty is also assumed to follow Poissonian statistics and is given by

$$\sigma_{m_i} = \sigma_{Np_i} = \sqrt{\frac{N_{ph}^{top}}{\varepsilon} p_i(x, y)}. \tag{4.9}$$

The sum in Equation 4.7 only considers those PMTs above a threshold of 50 photons, similar to the weighted sum. This threshold is introduced so that the test statistic $T(x, y)$ in Equation 4.7 follows a χ^2 distribution. If PMTs with low or no counts were also included, certain properties of a χ^2 distribution, such as the calculation of confidence regions (explained in Section 5.3), could no longer be applied to the test statistic. The analysis of this threshold is provided in Section 5.4.

The reconstructed position is found by calculating the test statistic $T(x, y)$ for one event and finding the minimum of the so-obtained two-dimensional test function.

4.2.1 Light collection efficiency

The number of $S2$ photons of an event, which are actually captured by the PMTs, is obtained by calculating the light collection efficiency (LCE). This is done with a Monte Carlo (MC) simulation since the LCE can not be calculated analytically. In this simulation, 10^9 photons are equally distributed in the gas phase and are propagated with Geant4. This simulation takes the geometrical properties of the detector, the reflectivity of the detector materials, the absorbance of the photons and their scattering in the xenon into account. From this propagation, one obtains the number of photons received by each PMT. For calculations and mapping, the photons are binned by a quadratic grid size of around 8.6 mm. To get the LCE map, the received photons in the PMTs are divided by the total number of simulated photons, which leads to an x-y dependent LCE map, which can be seen in Figure 4.2. The LCE map indicates that around 35% of the produced photons are detected by any PMT. The LCE is lower at the edge of the detector and therefore the $S2$ signals of events at the edges of the detector have to be corrected, using the reconstructed position. The $S1$ signal must also be corrected, however, the LCE used for this is only weakly dependent on the x-y-position and has a strong z dependency.

Calculating the $p_i(x, y)$, which are also called Probability Distribution Functions (PDFs), is an essential part of the position reconstruction. The difference between the LCE map and the PDF is that for $p_i(x, y)$ the detected photons are not divided by the total number of simulated photons, but by the number of received photons in all PMTs, which can be denoted as

$$p_i(x, y) = \frac{\nu_i(x, y)}{\sum\limits_{j=1}^{955} \nu_j(x, y)},$$

where $\nu_i(x, y)$ is the number of received photons for PMT i as a function of the x-y-position. These PDFs are normalized, so that

$$\sum_{i=1}^{955} p_i(x, y) = 1 \tag{4.10}$$

holds. The PDFs have a very sharp peak at their center position and are decreasing rapidly. The PDF of the center PMT $i = 477$ can be seen in Figure 4.3. The PDF of a PMT at the edge of the detector is cut off, which plays a role for the reconstruction of events at the edges, as described in Section 5.2. Additional plots of the PDFs can be found in Appendix A.

4.3 Implementation

The implementation of the χ^2 minimization is a significant fraction of the work of this thesis and therefore this section gives a short overview of the technical aspects.

While reconstructing one event, the calculated test statistic in Equation 4.7 of one event is only defined on a grid, due to the fact, that the PDFs of the PMTs are calculated on an 8.6×8.6 mm grid. To get a smooth function, one has to interpolate between the grid points. The interpolation can be done in two different ways: Every single PDF of one PMT can be interpolated and afterward the test statistic can be calculated. This leads to artifacts based on the LCE simulation grid, which is a serious problem for the minimization. The expedient way of interpolating is to calculate the whole test statistic $T(x, y)$ on a grid and interpolate afterward. For the interpolation, a rectangular bivariate spline is used without smoothing, to avoid losing the information of the PDFs' sharp peaking maxima. Another important aspect of the interpolation is to avoid the algorithm extrapolating outside the detector due to the few PMTs that are placed at the edges of the detector and the grid (e.g. at $(1300, 0)$ mm). This extrapolation occurs because their peaks are cut off and the PDFs are not approaching zero. Therefore, the LCE simulation grid is artificially increased at the edges by adding zeros. To find the minimum after the interpolation, the reconstructed position of the weighted sum algorithm is used as a first guess. The position reconstruction algorithm is implemented as a Tray module into the simulation framework and the computation time for one event is at around 0.03 s.

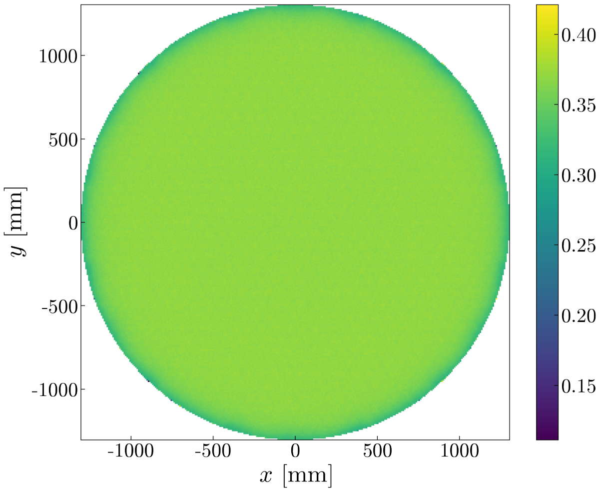


Figure 4.2: LCE map of the DARWIN detector. Around 35% of the photons are detected by any PMT. Due to the absorbance and reflection at the detector walls, the LCE is lower at the edge of the detector.

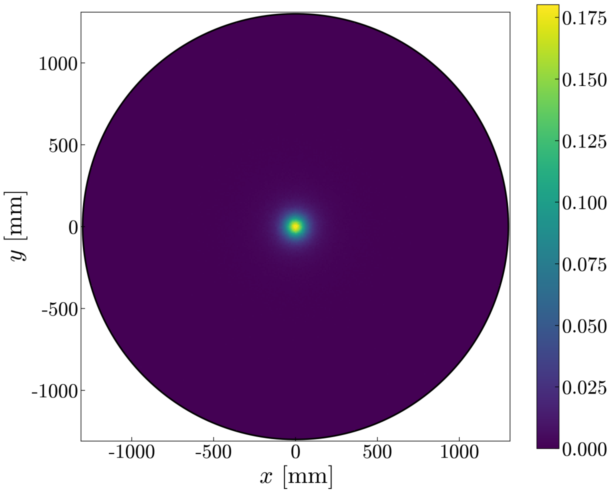


Figure 4.3: Colormap of the PDF of the center PMT. Around 17.5% of the detected photons originating from $(0, 0)$ are detected by the center PMT.

Chapter 5

Performance of the position reconstruction

In this chapter, the position reconstruction algorithms described in chapter 4 are tested on a Monte Carlo (MC) data set. This data set contains 10^5 NR events with random positions and energies. The energies correspond to the deposited energy of the scattering and are uniformly distributed in the range from 0 to 100 keV. The only data selection criterion in this chapter is to pass the N4TW200 (baseline) threshold: $S1$ photons of one event have to hit at least 4 PMTs (four-fold coincidence, denoted as N4) in the timespan of 200 ns (denoted as TW200) to get a clear $S1$ peak.

Since the true MC positions are known, the first question to be asked is how much the reconstructed position deviates from the true position. The radial residual, Δr, is defined as

$$\Delta r = \sqrt{(x_{rec} - x_{MC})^2 + (y_{rec} - y_{MC})^2}, \tag{5.1}$$

and is calculated for the weighted sum and the χ^2-minimization algorithm in the following. Furthermore, studies on the thresholds of the two algorithms and the features of the χ^2 algorithm are presented in this section.

5.1 Weighted sum algorithm

Since the reconstructed position of the weighted sum is used as the starting point for the χ^2 minimization, the radial residual of the weighted sum plays an important role in the χ^2 algorithm. Calculating the sum over all 955 PMTs leads to a mean radial residual of about 16 cm. Introducing a threshold for the considered PMTs reduces this residual significantly. This can be seen in Figure 5.1. The threshold is chosen in the way, that only PMTs are considered which are above a certain percentage of the maximum photon counts in one PMT. At a relative photon threshold of 10%, the mean residual is minimized to 4.0 mm. Without a threshold, the reconstructed positions are biased towards the center ($x = 0, y = 0$) because the noise is equally distributed and would result in an average reconstructed position at ($x = 0, y = 0$). Thus, any measurement is biased toward the center of the TPC if PMTs below the threshold are taken into account as well. An additional visualization of this bias can be seen in Appendix B. A histogram of the radial residuals of the weighted sum together with those of the χ^2 algorithm can be seen in Figure 5.2.

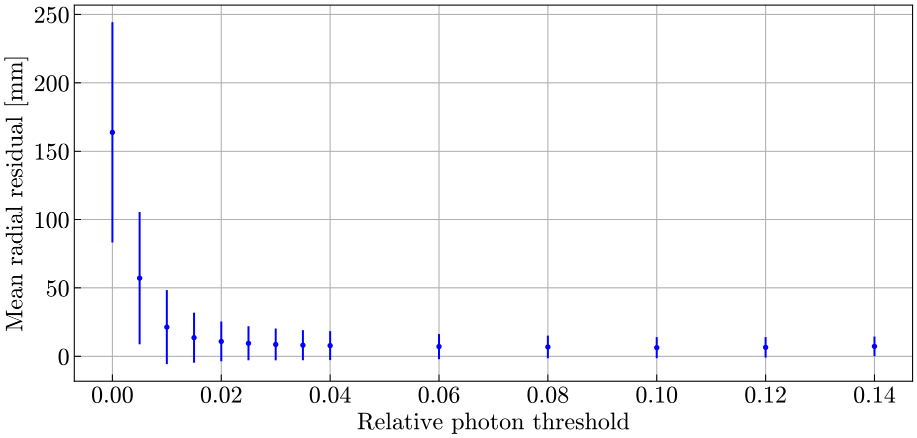


Figure 5.1: Mean radial residuals of the weighted sum in dependence of the chosen relative photon threshold.

5.2 Resolution of the χ^2 algorithm

Calculating the radial residual is also the first check for the χ^2 algorithm. Figure 5.3 shows that the mean radial residual increases for events with $1200\,\mathrm{mm} \leq r < 1300\,\mathrm{mm}$. The reasons for that are edge effects which can be seen in the LCE map in Figure 4.2. Also, the PDFs of the edge PMTs are cut off, which can be seen in Appendix A. For further analysis, only events with $r < 1200\,\mathrm{mm}$ are selected to stay in the so-called fiducial volume. The fiducial volume describes the part of the detector in which events are used for data analysis, all external events are rejected. It is defined in such a way that 10 cm from the edges and 10 cm from the top and bottom of the detector are cut away. Further aspects for the choice of the fiducial volume can be found in Section 6.1. A histogram of radial residuals inside the fiducial volume can be seen in Figure 5.2. It clearly shows that the χ^2 minimization is – as expected – more accurate than the weighted sum algorithm. The mean radial residual is at about 2 mm, which is noticeably smaller than the PMT radius of 38 mm.

It would be interesting to study if the resolution could be increased by simulating the PDFs on a finer grid. Currently, as explained in Section 4.2.1, the 10^9 photons are binned on 8.6×8.6 mm grid, equivalent to 301×301 bins. Increasing the number of bins with the current Geant4 data would not increase the resolution because the new bins would not contain enough photons. This means that the Geant4 simulation has to be repeated with a larger number of photons, which is a very time-intensive calculation beyond the scope of this work.

5.3 χ^2 distribution and confidence regions

As explained in Section 4.2, the test statistic follows a χ^2 distribution with a threshold. This can be used to calculate a goodness-of-fit parameter and confidence regions, providing an uncertainty estimate on the reconstructed position.

The minimum χ^2 value, χ^2_{min}, at the reconstructed position provides information on how well the reconstruction worked. However, the absolute χ^2 value is ambiguous, since the number of considered PMTs for the sum in Equation 4.7 is not constant. To get a meaningful goodness-of-fit parameter, the

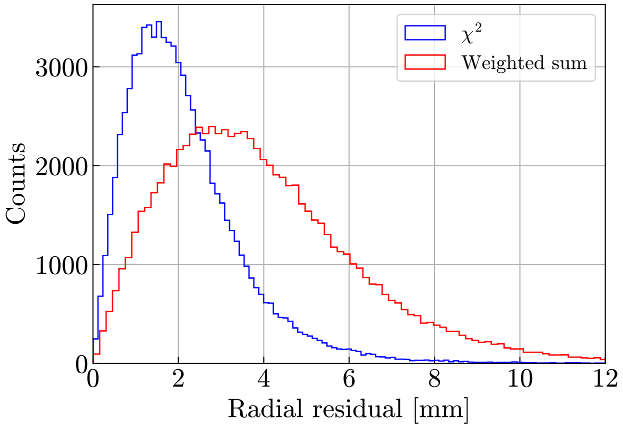


Figure 5.2: Histogram of radial residuals for around 80000 events in the fiducial volume.

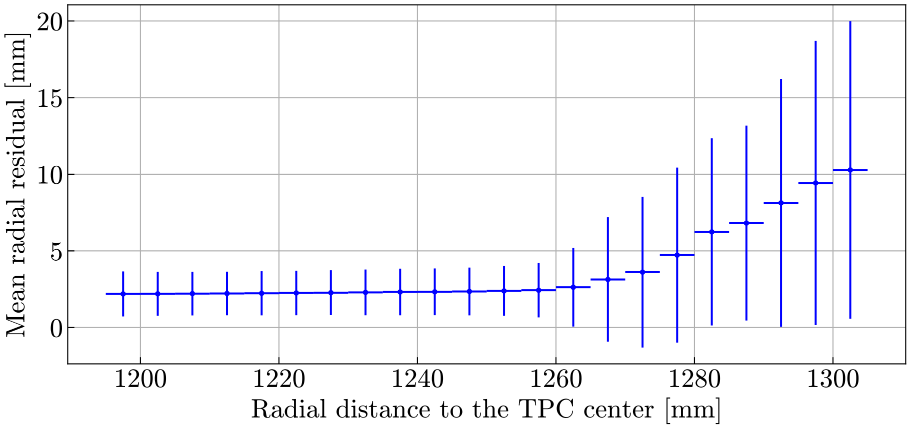


Figure 5.3: Mean radial residuals of the χ^2 algorithm against the radial distance to the TPC center, binned by reconstructed radius. Edge effects from the PDFs and the LCE occur for large radii. Although the uncertainties are large, the tendency is recognizable that there are larger radial residuals for larger radii.

reduced χ^2 value is calculated with

$$\chi^2_{red} = \frac{\chi^2}{\mathrm{ndof}} \quad \text{and} \quad \mathrm{ndof} = \#(\text{Considered PMTs}) - 1 - 2, \tag{5.2}$$

where ndof is the number of degrees of freedom which corresponds to the number of considered PMTs reduced by one because one degree of freedom is lost due to the normalization and again reduced by two because there are further two free parameters (reconstructed x and y position). A histogram of χ^2_{red} is displayed in Figure 5.4 on the left. On the right in Figure 5.4, the relation between χ^2_{red} and N_{ph}^{top}, which corresponds to the recoil energy, can be seen. There is no statistically significant energy dependence on the χ^2_{red} value.

Another advantage over the weighted sum is the possibility of extracting an uncertainty estimate on the reconstructed position with confidence regions. To find these confidence regions, one has to consider the absolute minimum χ^2 value. Table 5.1 shows the definitions of the confidence regions of a two-dimensional χ^2 distribution with two parameters of interest, taken from [22]. With that, a

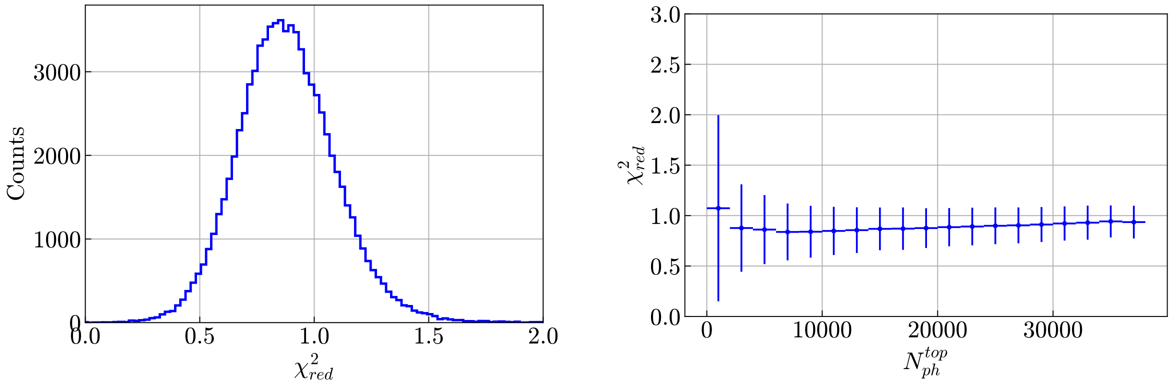


Figure 5.4: Left: Histogram of χ^2_{red} values of the reconstructed position. Right: Dependency of the χ^2_{red} value of the reconstructed position on the total photon number of $S2$ photons, which corresponds to the initial energy deposit. The energy dependence is less than 10% and therefore negligible.

Table 5.1: Definition of confidence regions for a test function with two parameters of interest [22].

Confidence level	1σ	2σ	3σ
Increase	$\chi^2_{min} + 2.3$	$\chi^2_{min} + 6.18$	$\chi^2_{min} + 9.21$

confidence contour can be drawn such that for all points in the 1σ confidence contour $\chi^2 \leq \chi^2_{min} + 2.3$ holds. The 1σ and 2σ confidence regions in the so-called χ^2 landscape can be seen in Figure 5.5. While the orange triangle indicates the starting position of the minimization given by the weighted sum, the yellow cross marks the found minimum and the red circle indicates the true position.

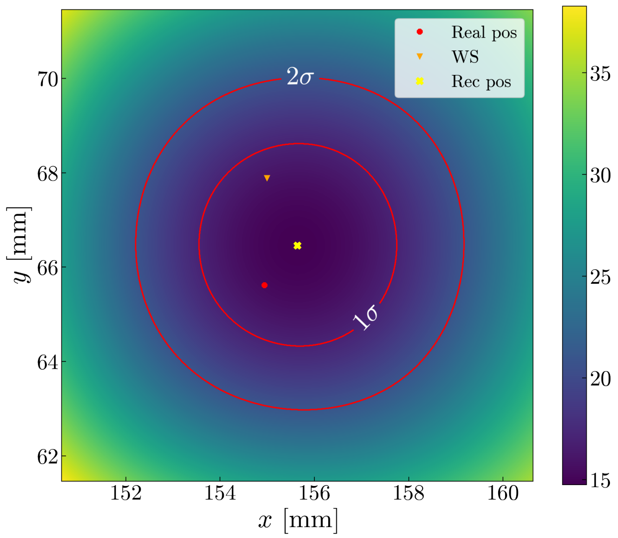


Figure 5.5: χ^2 landscape of a reconstructed NR scattering position and the corresponding confidence regions as they are defined in Table 5.1. The orange triangle marks the starting position of the minimization given by the weighted sum and the yellow cross marks the reconstructed position at the minimum. The red dot indicates the true MC position.

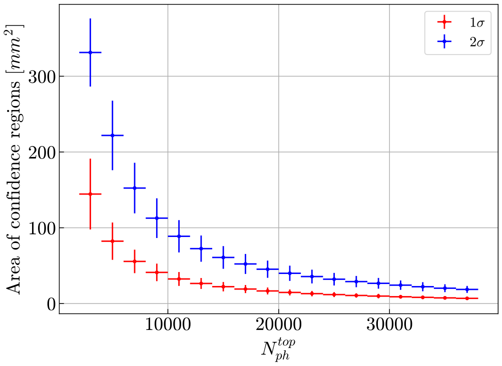


Figure 5.6: The uncertainty of the reconstructed position measured in the area of the confidence regions. The uncertainty decreases for a larger number of photons, following a $\left(N_{ph}^{top}\right)^{-1/2}$ behavior.

Since the χ^2 algorithm principally fits a mean to the $S2$ photon cloud, the uncertainty, $\overline{\sigma}$, of this fitted mean should decrease with the number of photons N_{ph}^{top} since

$$\overline{\sigma} \sim \frac{1}{\sqrt{N_{ph}^{top}}}. \tag{5.3}$$

Figure 5.6 shows this correlation of the area of the 1σ and 2σ confidence regions as in Equation 5.3.

5.3.1 Eccentricity of confidence regions

The confidence regions in Figure 5.5 have a very circular shape, which is given for Gaussian distributed photon clouds. By fitting ellipses to the confidence regions, the shape of those can be analyzed by calculating the eccentricity, ε, of the ellipses. A histogram of the eccentricities obtained from single scatter events can be seen in Figure 5.7 on the left. The eccentricities are centered at around $\varepsilon = 0.6$. A fitted ellipse of the 2σ confidence region with an eccentricity of $\varepsilon = 0.58$ can be seen in Figure 5.7 on the right. The eccentricities of multiple scatter events are expected to be larger. This will be studied in Section 6.2.1

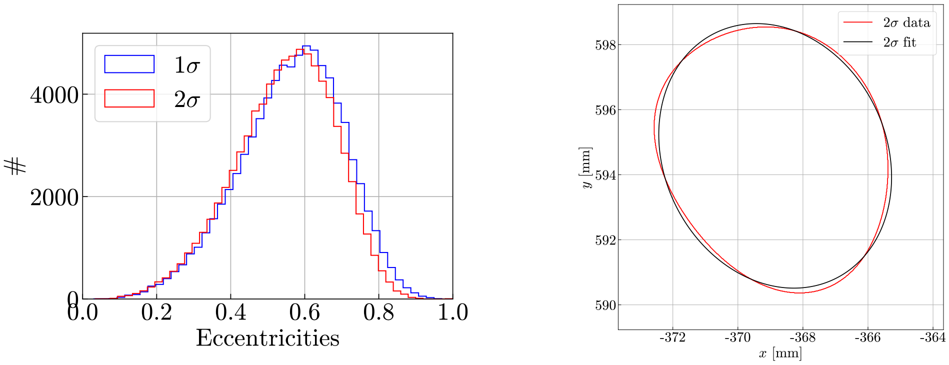


Figure 5.7: Left: Eccentricities of the 1σ and 2σ confidence regions. Right: Fitted ellipse to the 2σ confidence region with an eccentricity of $\varepsilon = 0.58$.

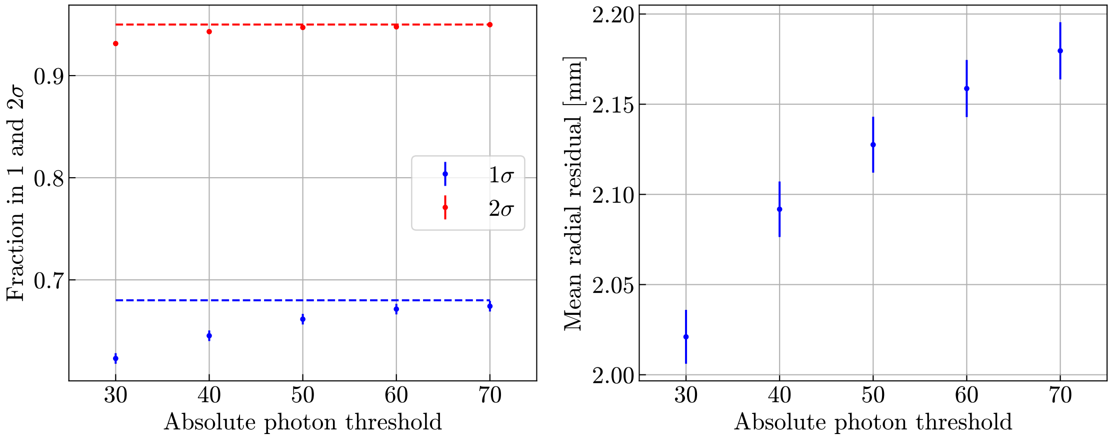


Figure 5.8: Left: Fraction of true events in 1σ and 2σ confidence contours in dependence of the chosen absolute photon threshold. The dashed lines indicate the theoretical expected fractions of 68% and 95%. Right: Radial residual in dependence on the chosen threshold.

5.4 Choice of threshold

The test statistic does not follow a χ^2 distribution without the introduction of a threshold. This can be seen by calculating the fraction of events for which the true position is within the 1σ and 2σ confidence regions of the reconstructed positions. The left plot in Figure 5.8 shows that the fractions differ with the chosen photon threshold and approach the desired levels for larger thresholds. The absolute photon threshold of 50 photons is chosen by considering the following aspects.

- As it can be seen in the right plot of Figure 5.8, the mean radial residual is increasing for larger thresholds. The threshold should therefore be as low as possible. However, it must be noted that this decrease is in the order of 0.2 mm, which is remarkably small compared to the size of the PMTs.
- Almost all events that passed the N4TW200 threshold have at least one PMT above the threshold. If other thresholds like N3TW200 or N2TW200 are applied, the absolute photon threshold could cut off more low energetic events, which forms another reason why the threshold should be as low as possible.
- By considering the two aspects above, the absolute photon threshold is set to 50 photons. The 2σ fraction is following the required value, as seen in the left plot of Figure 5.8. The 1σ value deviates slightly from the required 68% level, but the 68% level is still in the 3σ range of the value.

A general investigation of this behavior and the influence of different thresholds has been performed using a toy model, which is described in the following section.

5.4.1 Toy model

To investigate whether the effects described above are of statistical nature or are caused by effects in the simulation of events, a toy model is created. To stay in the context of the experiment, the arriving photon cloud is simulated by sampling with a two-dimensional Gaussian distribution, while the mean of the Gaussian represents the initial position, which has to be reconstructed. The sample size corresponds to the total number of photons. This number is energy-dependent but is generally in the order of 10^4 counts. The detection of photons with the PMTs can be transferred to the toy model by using a 30×30 binning of the Gaussian sample, similar to the 955 PMTs. After the binning, χ^2 sums are calculated based on two different statistics, as in Equation 5.4. Neyman statistic is used for Poissonian distributed signals (i.e. low counts), where the division with zero is avoided by using $\max(n_i, 1)$ as an uncertainty in the denominator. Since this approach still underestimates the true mean by one [21], the Mighell statistic is considered.

$$\textbf{Mighell:} \quad \chi^2 = \sum \frac{(n_i + \min(n,1) - m_i^2)}{n_i + 1}, \quad \textbf{Neyman:} \quad \chi^2 = \sum \frac{(n_i - m_i)^2}{\max(n_i, 1)}, \tag{5.4}$$

In Equation 5.4, one sums over all 900 bins, where n_i is the number of simulated counts in bin i. The corresponding expected number of counts in bin i is denoted as m_i. The expected number of counts in bin i is calculated by using an ordinary two-dimensional Gaussian distribution, whereby this must still be scaled with the sample size and the bin size. The x and y positions of the respective bin i are then used as an input for the $2D$ Gaussian and the mean values in x and y of this Gaussian represent the position to be reconstructed and are therefore free parameters for the minimization. By minimizing the difference between n_i and m_i, the position can be reconstructed. Compared with the χ^2 algorithm, m_i corresponds to $N_{ph}^{top} p_i(x, y)$ in Equation 4.7.

The corresponding χ^2 value at the minimum of the reconstructed mean can be seen in Figure 5.9 on the left. The dashed line indicates $\chi^2 = 897$. This value corresponds to the expected minimum χ^2 value since there are three degrees of freedom less: one for the normalization of the model (sample size) and two for the fitted means in x and y. For both statistics, the best fit χ^2 value is too low. The reason for that are firstly bins with zero counts, which do not contribute to the χ^2 sums in Equation 5.4. This also holds for the simulation data of the DARWIN experiment, because a lot of the PMTs count zero photons. Therefore, empty bins artificially increase the degrees of freedom without containing any information. Including also bins with low counts still does not generate a proper χ^2 distribution, although the Mighell statistic tries to correct this. This issue is prevented, by restricting the range of the binning: The binning is only applied in a small region around the true mean, which corresponds to the introduced photon threshold. To compare this method to the unrestricted method, the number of bins is kept the same. This stands in contrast to the experiment where, of course, the number of PMTs stays the same and χ^2_{red} has to be calculated for the comparison. This method is chosen, although it does not exactly match the procedure in the experiment, as this avoids the choice of a threshold. These methods with a restricted binning range are denoted as "MighellRange" and "NeymanRange", since they use the same χ^2 sums as in Equation 5.4 for the Mighell and Neyman statistics. The restricted Mighell statistic reaches the required χ^2 value already at a sample size of 10^4, which is the region of interest. The restricted Neyman statistics approaches the required χ^2 only for larger sample sizes. On the right-hand side of Figure 5.9, the radial residual is shown, which follows the inverse square-root relation given in Equation 5.3.

5.5 Leakage of the fiducial volume

The leakage of the fiducial volume is defined as the amount of events that are reconstructed in or outside the fiducial volume, but originate from the other side. Leak-in events are those events whose true position is outside the fiducial volume and the reconstructed position is inside the fiducial volume. Leak-out events are defined the other way around. From the same MC data set as before, one obtains

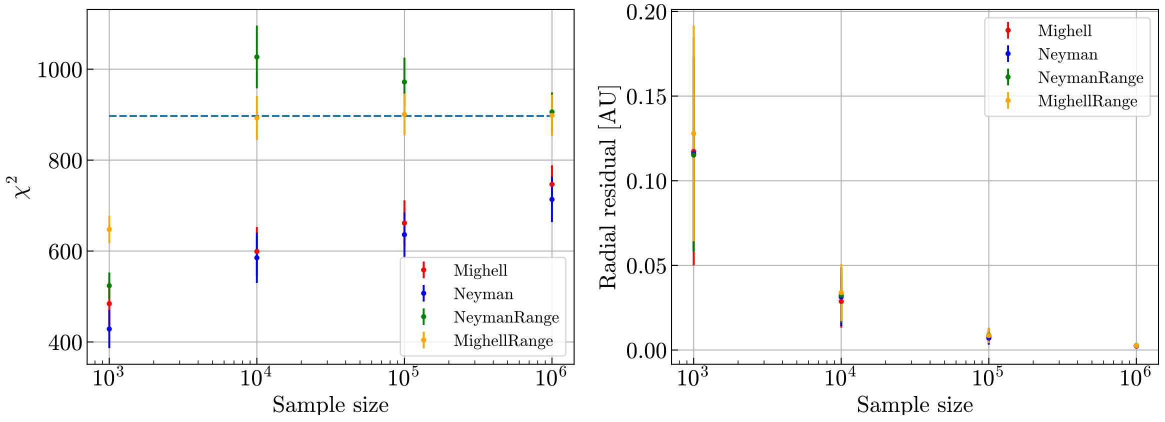


Figure 5.9: Left: χ^2 value of the found minimum for the different methods. Right: Radial residual of the reconstructed means. The region of interest is at about 10^4.

the values

$$\textbf{Leak-in:}\ (0.131 \pm 0.011)\,\% \quad \text{and} \quad \textbf{Leak-out:}\ (0.075 \pm 0.009)\,\% \tag{5.5}$$

for the χ^2 algorithm. This result makes sense since the position reconstruction works better for smaller radii. A histogram of the true radii of the leakage events can be seen in Figure 5.10.

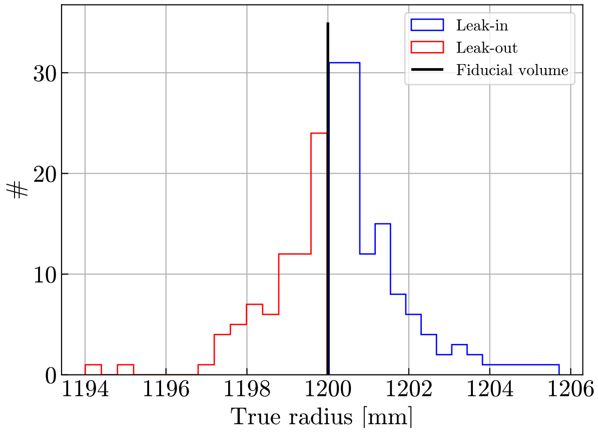


Figure 5.10: Radii of the leak-in and leak-out events, binned by the true radius.

Chapter 6

Radiogenic neutron background separation

Since it is expected that only very few WIMPs may scatter in the TPC, it is essential to reject signals of other scattering processes. Firstly, the background is minimized by building the experiment deep underground. This shields almost all particles, except for muons and neutrinos. Using the ratio $S1/S2$, as described in Section 3.2.1, and the muon veto system, the scattering of muons can be identified, whereas the scattering of neutrinos cannot be identified and remains a constant background. Another background, independent of the location of the experiment, is the radiogenic neutron background, which is caused by radioactively contaminated detector materials. (α, n) reactions from the decay of trace elements and spontaneous fission result in the emission of neutrons, which can scatter with the xenon atoms in the TPC of the detector. As neutrons have a much larger cross section than WIMPs and neutrinos, they can scatter several times in the xenon. Identifying these multiple scatters (MS) makes it possible to separate signals from neutrons, which is an advantage over the irreducible neutrino background.

To study the radiogenic neutron background at DARWIN, Geant4 is used to simulate the emitted neutrons of the different materials and the contained trace elements. The PMTs consist of ceramic, Cirlex, Kovar and quartz. The detector consists of copper (used for the driftfield), PTFE (the TPC is coated with PTFE for a better reflection) and titanium, while titanium is used both in the electrodes and in the cryostat vessel. The radioactive trace elements ^{226}Ra, ^{228}Th, ^{232}Th, ^{235}U and ^{238}U are contained in the different materials. Considering the neutron production rate of every isotope and their amount in the materials, the expected rate of emitted neutrons per year amounts to 1991 neutrons/year [23]. By considering the activities of the isotopes and the locations of the materials in the TPC, Geant4 simulates the propagation of the neutrons and their energy deposits.

Since there are 40 different material-isotope pairs and the processing of the simulation data is very time-consuming, the following analysis has only been done for seven different material-isotope pairs, which make up 73% of the total emitted neutrons which then amounts to 1453 neutrons/year. The considered material-isotope pairs can be seen in Table 6.1.

Table 6.1: Overview of the considered material-isotope pairs which make up 73% of the total emitted neutrons.

Material	**Isotopes**
Cryostat (titanium)	^{228}Th, ^{232}Th, ^{238}U
PTFE	^{226}Ra, ^{235}U, ^{238}U
Ceramic	^{226}Ra

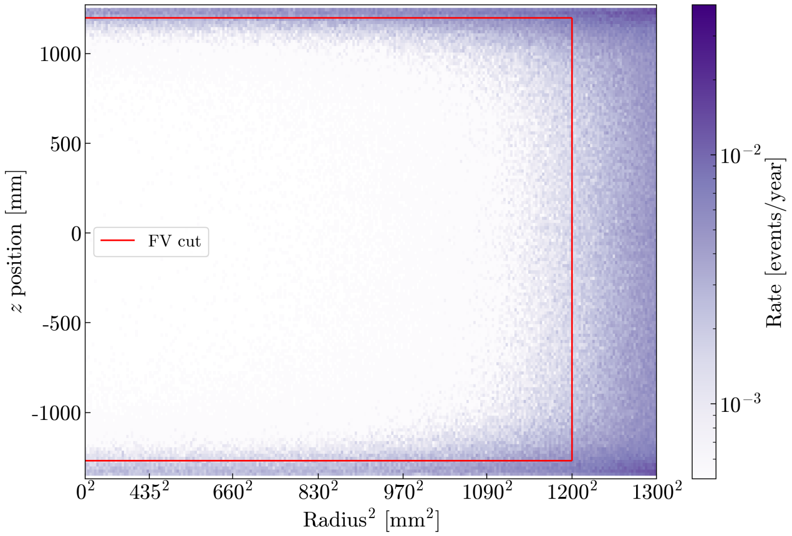


Figure 6.1: Spatial distribution of the radiogenic neutron scatters. Due to the self-shielding of xenon, the neutrons mostly scatter at the edges of the detector. By only considering events inside the so-called fiducial volume, which is indicated by the red lines, the neutron event rate can be reduced by 73%.

6.1 Data selection and previous MS identification

Whereas in the previous chapter only the N4TW200 threshold is used as a data-selection criterion, further criteria are now applied to the neutron data. In the following, only scatters are called events if they pass the first three criteria in the list below, the so-called event cuts. The last cut is called the fiducial-volume cut.

- **N4TW200:** The $S1$ photons have to hit at least four PMTs in a timespan of 200 ns to ensure that every event has a clear $S1$ peak.
- **S2** : There must be at least 73 $S2$ photoelectrons in one event to cut off one-electron events.
- **cS1** : All events with more than 150 $S1$ photons after the corrections are rejected to stay in desired the search window for WIMPs.
- **Fiducial volume:** The spatial distribution of the radiogenic neutron events in Figure 6.1 shows that most of the neutrons scatter at the edges of the detector due to the self-shielding of xenon. The fiducial volume (FV) is obtained by cutting off the outer 10 cm of the detector horizontally and vertically. Therefore, the active xenon target mass of 40 tons is reduced to the fiducial mass of 31.4863 tons and only events inside the fiducial volume are considered, which already reduces the rate of neutron events by 73% as seen in Table 6.2. The weighted-sum algorithm is used for

Table 6.2: Reduction of the neutron background using the different cuts. The FV cut reduces the neutron event rate already by 73%. The total reduction of neutron events after the z cut amounts to 97.8%.

Cut	**Neutron event rate [events/year]**
All events	41.46
FV	11.19
FV and z	0.902

the position reconstruction, which should reconstruct a center position of the different scatters and which is also faster than the χ^2 algorithm.

Since a lot of neutrons scatter multiple times in the TPC, the neutron background can be reduced by identifying those MS events. In addition to the identification using the χ^2 algorithm explained in the next section, there is already an MS identification in the DARWIN framework that uses the timing distribution of the $S2$ photons, which typically has a Gaussian shape. The associated width follows the relation

$$\text{width} \sim \sqrt{\text{drift time}}, \tag{6.1}$$

with the drift time of the electrons, which is directly connected to the z position of the scatter through the constant drift velocity in the TPC. Therefore, this rejection method is also called z cut. If a neutron scatters multiple times in the detector, the $S2$-timing-distributions of the individual scatters overlap and the width is bigger than $\sqrt{\text{drift time}}$. By parameterizing Equation 6.1 with single scatter MC data, unusual events as MS can be identified and rejected. Since the widths spread around the square root function, a 90% acceptance level is chosen for the MS identification. Since the spread is larger for lower energies due to statistical fluctuations, the width of the acceptance band has an energy dependence. This is considered by splitting up into different energy ranges, corresponding to the number of measured photoelectrons.

Table 6.2 shows that the rate of neutron events in the TPC amounts to 41.46 events/year. This rate is significantly smaller than the activity of all isotopes with 1453 neutrons/year since most of the neutrons scatter outside the detector and only scatters are considered as an event if they pass the event cuts. The neutron event rate can be reduced by 73% with the FV cut, using the self-shielding property of xenon. The z cut reduces the rate again to 0.902 events/year and a total reduction of 97.8% is achieved.

The following section explains how the χ^2 algorithm can be used for a further reduction of the neutron event rate. For this study, the z and the χ^2 cuts are always applied to those events inside the FV.

6.2 MS identification by using the χ^2 algorithm

If the χ^2 algorithm gets the $S2$ hit pattern from an MS event as an input, it fails to reconstruct a "correct" position, since it is only implemented for the reconstruction of single scatters. This leads to a high χ^2_{red} value. It is therefore necessary to choose a suitable cut value for the MS identification.

By additionally applying the cuts in Section 6.1 to the histogram in Figure 5.4, NR acceptance levels can be defined by considering different χ^2_{red} cut values. The correlation between different cut values and the corresponding NR acceptance can be seen in Figure 6.2 on the left. A low NR acceptance level leads on the one hand to a high rejection of MS but at the same time also to a WIMP-signal loss.

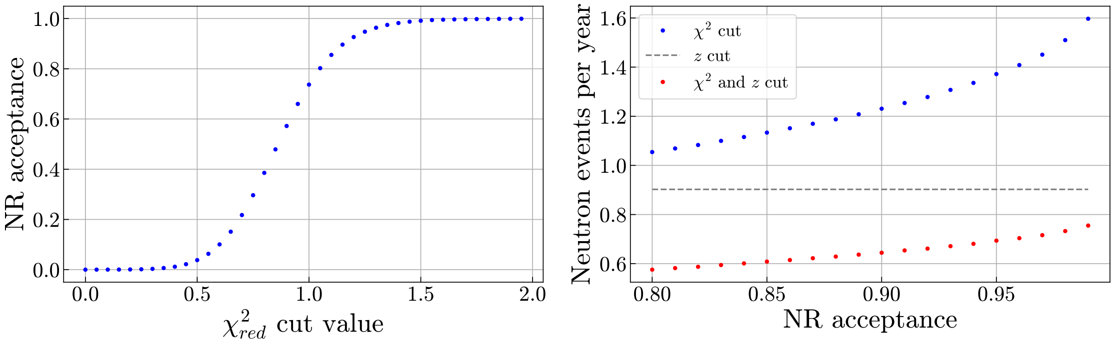


Figure 6.2: Left: Correlation between the χ^2_{red} cut value and the corresponding NR acceptance. Right: Neutron rates for different NR acceptances, while the different cuts are applied to the events that already passed the event cuts and the FV cut. The blue curve shows the rate after applying the χ^2 cut. The dashed line indicates the rate after the z cut as shown in Table 6.2 and the red curve shows the rate dependent on the NR acceptance after applying all cuts.

For the selection of a suitable acceptance level, the right plot in Figure 6.2 is considered, in which the neutron event rate is plotted against the NR acceptance. The blue curve is calculated by applying the MS identification with the χ^2 algorithm, the so-called χ^2 cut, to the events inside the FV. The dashed line indicates the neutron event rate after the z cut, as shown in Table 6.2. The red curve indicates the neutron event rate after applying all cuts to the data. The blue curve follows a linear dependency for NR acceptances smaller than 90%. For this reason, and to agree with the z cut, an NR acceptance of 90% is chosen, corresponding to a cut value of $\chi^2_{red,cut} = 1.155$.

Figure 6.3 shows the energy spectrum of the neutron events. The black curve shows all events which passed the event cuts and the yellow curve indicates the events which are inside the FV. The χ^2 cut (green curve) and the z cut (red curve) are then applied to the events inside the FV. With the z cut, more MS can be rejected than with the χ^2 cut. The reason for that is that the z cut uses the detection times of the $S2$ electrons which can be measured very precisely. The χ^2 algorithm is on the other hand restricted by the PMT size. Both the z cut and the χ^2 cut are more efficient at higher energies since the acceptance band of the z cut is smaller for higher energies as explained in Section 6.1 and the MS identification with the χ^2 algorithm works better for higher energies due to smaller statistical fluctuations as seen in Section 6.4. The blue curve shows the remaining events after applying all cuts to the data. The improvement due to the χ^2 cut compared to the previous cuts is represented by the difference between the red and the blue curve, which is calculated explicitly in Section 6.3. The neutrons can also scatter only one time in the TPC. Therefore, the gray curve shows the amount of true single scatters out of all events that passed all previous cuts, which is an irreducible background.

6.2.1 MS identification by using eccentricities of confidence regions

Another way of identifying MS is by using the eccentricities, ε, of the fitted ellipses of the confidence regions, as explained in Section 5.3.1. The confidence regions of an MS have higher eccentricities since the photon clouds are no longer Gaussian distributed or even have more indistinct shapes. By considering the histogram with the additional cuts from Section 6.1 of the 1σ and 2σ NR single-scatter eccentricities in Figure 5.7, ε cut values can be defined corresponding to a certain NR acceptance,

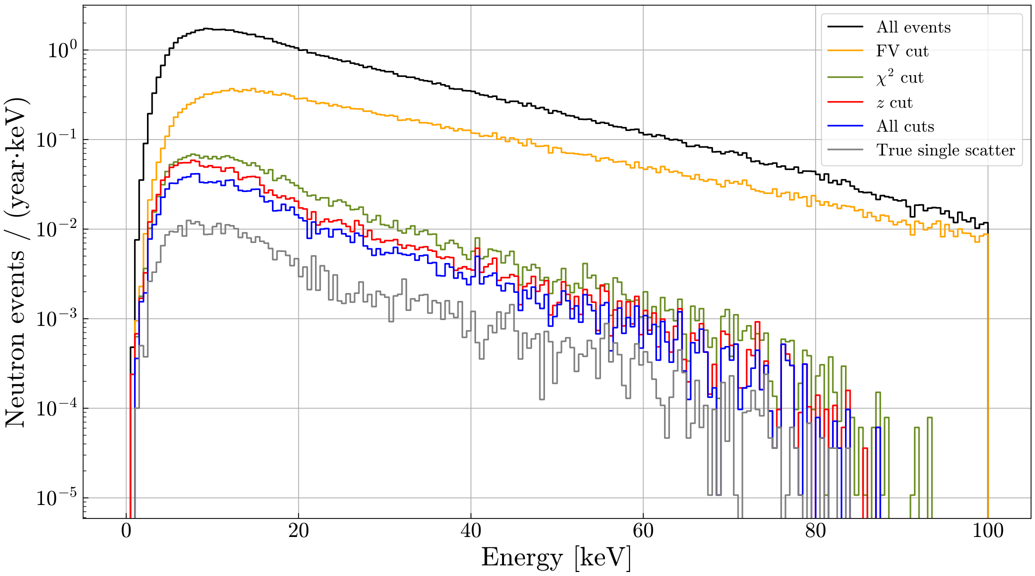


Figure 6.3: Energy spectrum of the neutron events after different cuts. The black curve shows those events which passed the event cuts and the yellow curve shows the events inside the FV. The χ^2 cut (green) and the z cut (red) are applied to the events inside the FV. Both cuts were applied with an NR acceptance of 90%. The z and χ^2 cuts are more efficient for higher energies due to smaller statistical fluctuations. The gray curve shows the true single scatters of the neutrons.

similar to the χ^2 algorithm. With a 90% acceptance, one obtains the cut values

$$\varepsilon_{cut}^{1\sigma} = 0.739 \quad \text{and} \quad \varepsilon_{cut}^{2\sigma} = 0.711. \tag{6.2}$$

The cuts are applied to the 1σ and 2σ eccentricities, the so-called 1σ-ε and 2σ-ε cuts. The ε cut describes the combination of both cuts. Figure 6.4 shows that the ε cut is less efficient than the χ^2 cut. The improvement compared to the previous cuts corresponds to the difference between the blue curve and the yellow curve. The fitting of the ellipses and the calculation of the eccentricities is around ten times slower than the χ^2 algorithm. For further investigations, the ε cut should only be applied to the events that passed the χ^2 cut. Beyond that, the χ^2 cut already rejects most of the MS events with larger distances between the individual scatters, for which the confidence regions have more indistinct shapes which can be a problem for the fit.

6.3 Expected number of neutron events at DARWIN

The rate of neutron events after a certain cut corresponds to the integral of the energy spectra in Figures 6.3 and 6.4. This rate corresponds to the number of expected neutron events per year at the DARWIN observatory, which can not be distinguished from possible WIMP signals since they passed all cuts. In this section, the amount of neutron events for the whole operational time of the DARWIN observatory is calculated.

Since only events inside the fiducial volume are considered, the active volume is fiducialised from 40 to 31.4863 tons, leading to an operational time of 6.352 years to stay at the baseline exposure of 200 t y. Using the neutron event rates obtained from the previous section, the expected number of neutron events in the TPC during the operational time can be calculated. Since the previous analysis was only

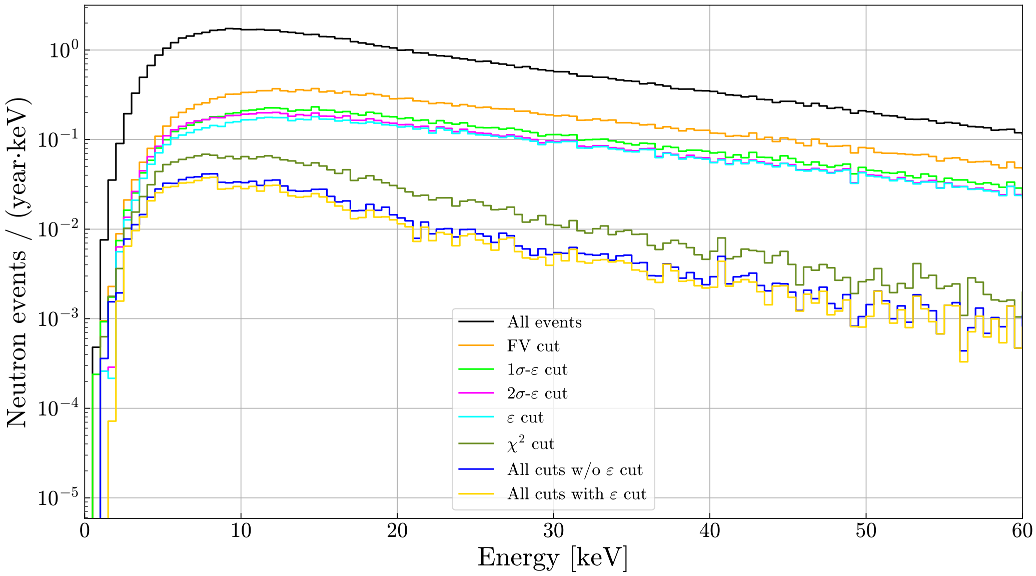


Figure 6.4: Energy spectrum showing the influence of the eccentricity cuts. The 1σ-ε (lime) and the 2σ-ε (magenta) curves describe the cuts on the 1σ and 2σ eccentricities, while the cyan curve shows the combination of both cuts. The yellow curve shows the remaining rate per energy after applying the ε cut in addition to the z and χ^2 cut.

Table 6.3: Expected number of neutron events with different cuts, where the χ^2, z and ε cuts are always applied on the events inside the FV.

Cuts	**Neutron events per year**	**Number of neutron events for** 200 t y	**WIMP acceptance** [%]
All events	41.459	360.75	-
FV	11.187	97.35	-
χ^2	1.230	10.71	90
z	0.902	7.85	90
z and χ^2	0.644	5.61	81
z, χ^2 and ε	0.562	4.89	72

done with 73% of the neutron data, the rates have to be scaled up. It is assumed that the scaling with a simple scaling factor is reasonable since the χ^2 cut has approximately the same efficiency for all different material-isotope pairs. The expected number of neutron events, n_{neutrons}, during the whole exposure time is then calculated with

$$n_{\text{neutrons}} = \frac{\text{neutron events}}{\text{year}} \cdot 6.352\,\text{y} \cdot \frac{100}{73}. \tag{6.3}$$

The numbers of expected events for the different cuts can be seen in Table 6.3. With the χ^2 cut alone, 89% of the neutron events inside the FV can be rejected. The χ^2 cut can improve the neutron background separation, additionally to the z cut, by 29% with a WIMP-signal loss of only 9%. Applying the ε cut additionally improves the background rejection by 13% with a WIMP signal loss of 9%. This shows that the ε cut is not as efficient as the χ^2 cut. The total improvement of the background rejection compared to the z cut, after additionally applying the χ^2 and ε cuts, amounts to 38% with a total WIMP-signal loss of 18%.

6.4 Double scatter separation

The identification of MS with the χ^2 algorithm works best for scatters that are far apart in the x-y-plane compared to the PMT size. It therefore makes sense to ask what the minimum distance between the two scatters in the x-y-plane must be so that the event is identified as a double scatter.

Double scatters (DS) have multiple free parameters: the horizontal and vertical distance between the scatters and the recoil energies of the two deposits. Since the separation in z, by using the timing of the $S2$ photoelectrons, is a very accurate method, the following analysis is performed by considering only DS with the same z position in each scatter as an extreme case. While the two individual scatters of one DS have the same z position, the different DS have been simulated at random z positions in the detector with increasing horizontal distances between the two scatters. This has been processed for energies of 10, 40 and 90 keV, where both scatters have the same recoil energy.

The identification with the χ^2 algorithm is again performed for an NR acceptance of 90%, corresponding to a χ^2_{red}-cut value of $\chi^2_{red,cut} = 1.155$, which is indicated by the dashed line in Figure 6.5. It can be seen that the DS separation works slightly better for high energetic DS, since the statistical fluctuations and the uncertainties of the reconstructed positions are smaller for larger energies, as seen in Figure 5.6. The minimal separation distance therefore has a slight energy dependence and is of the order of 4 to 7 cm. This distance is in the same order of magnitude as the size of the PMTs, which is the limiting factor for the spatial separation resolution.

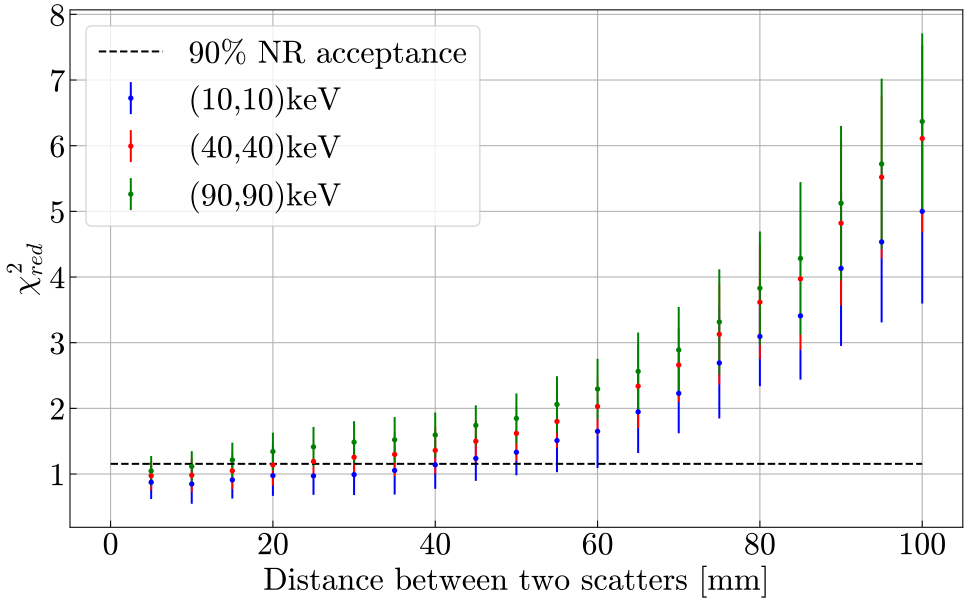


Figure 6.5: Separation of double scatter events with different energies. The error bars represent the width of the χ^2_{red} distribution at a certain distance. The black dashed line indicates the 90% NR acceptance. The separation distance is slightly energy dependent and is of the same order of magnitude as the PMT size, with a few cm.

Chapter 7

Summary and outlook

This thesis explains the reconstruction of the horizontal position of a scattering process within the time projection chamber of the DARWIN observatory. The main method used is a χ^2 minimization, for which a weighted-sum algorithm (WS) is used as the starting point for the minimization. A photon threshold must be introduced for both the WS and the χ^2 algorithm. The thresholds avoid a bias towards the center of the TPC for the WS and generate a proper χ^2 distribution for the χ^2 minimization. Based on Monte Carlo simulation, the WS reconstructs the horizontal position with an accuracy of about 4 mm, whereas the χ^2 algorithm achieves an accuracy of 2 mm. This is a very accurate reconstruction, compared to the radius of the PMTs of 38 mm. With the χ^2 algorithm, the confidence regions can be used to estimate an uncertainty for the reconstructed position that is of the order of a few millimeters. At the same time, the algorithm offers with the reduced χ^2 value, χ^2_{red}, a goodness-of-fit parameter to restrict the analysis only to proper reconstructed events.

With the analysis of the χ^2_{red} value and a 90% NR acceptance level, 89% of the neutron events inside the fiducial volume can be identified, using the fact that most of the neutrons scatter multiple times. Compared to the previous identification by using only the z cut, the χ^2 cut improves the neutron background separation additionally by 29%. The identification of neutron scatters using the eccentricities of the confidence regions is in turn not as efficient as the χ^2 cut. Only 12% can additionally be rejected with a WIMP-signal loss of 9%. The number of neutron events for an exposure of 200 t y amounts to 5.61 neutron events with the z and χ^2 cuts.

A possible next study could be to check whether the resolution of the χ^2 algorithm can be improved by simulating the PDFs on a finer grid, which is however a very time-consuming process.

In addition to the two algorithms used in this thesis, the horizontal position can also be reconstructed by using a Neural Network, which can be faster and more accurate than the χ^2 algorithm. Since Neural Networks usually do not provide a goodness-of-fit parameter, the χ^2 algorithm can still be used to check if the reconstructed position of the Neural Network is reasonable.

The radial separation of double scatter is limited by the size of the PMTs. A finer granular optical readout, e.g. by silicon photomultipliers, could result in further improvements which is another interesting topic to be studied.

Appendix A

Probability distribution functions

This appendix shows some additional plots for the PDFs of the PMTs, as described in Section 4.2.1.

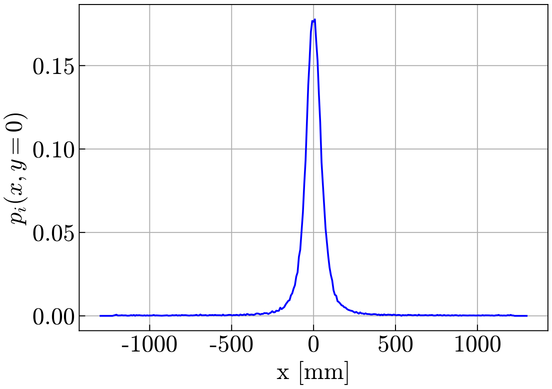


Figure A.1: Profile of the PDF of the center PMT at $y = 0$. The PDF has a sharp peak at the PMT position and is decreasing rapidly.

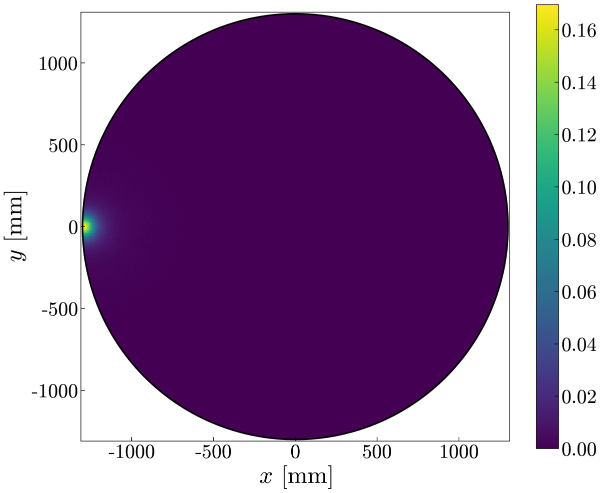


Figure A.2: PDF of a PMT at the edge of the detector. The PDF is asymmetric.

Appendix B

Bias of the weighted sum

As described in Section 5.1, the weighted sum is biased towards the center of the TPC without a threshold, since there are more PMTs considered in the direction of the center of the TPC.

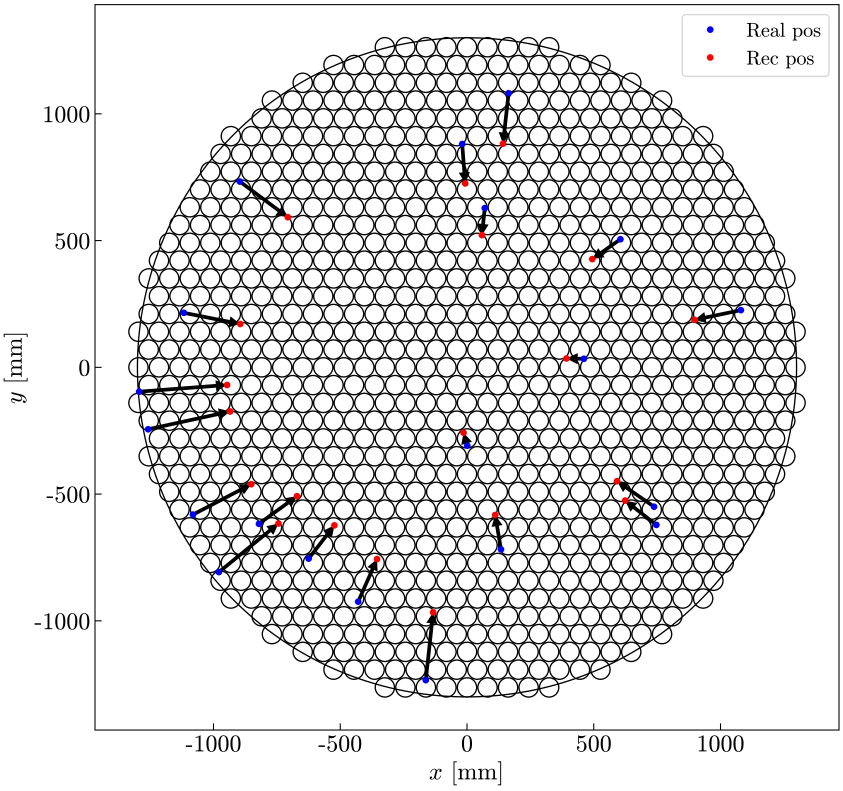


Figure B.1: Red dots indicate the reconstructed positions of the weighted sum without a threshold, and the blue dots indicate the true positions. The reconstructed positions of the weighted sum are biased towards the center of the TPC.

Bibliography

[1] Planck Collaboration. “Planck 2018 results - I. Overview and the cosmological legacy of Planck”. In: *A&A* 641 (2020). DOI: 10.1051/0004-6361/201833880. URL: https://doi.org/10.1051/0004-6361/201833880 (cit. on pp. 1–3).

[2] Yuan Mei. “Direct Dark Matter Search with the XENON100 Experiment”. PhD thesis. RICE UNIVERSITY, Apr. 2011 (cit. on pp. 1, 8, 13).

[3] Bart Pelssers. “Position Reconstruction and Data Quality in XENON”. MA thesis. University of Utrecht, 2015 (cit. on p. 1).

[4] J. H. Oort. “The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems”. In: 6 (Aug. 1932), p. 249. URL: https://ui.adsabs.harvard.edu/abs/1932BAN.....6..249O (cit. on p. 2).

[5] F. Zwicky. “Die Rotverschiebung von extragalaktischen Nebeln”. In: *Helvetica Physica Acta* 6 (Jan. 1933), pp. 110–127. URL: https://ui.adsabs.harvard.edu/abs/1933AcHPh...6..110Z (cit. on p. 2).

[6] Vera C. Rubin and Jr. Ford W. Kent. “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”. In: 159 (Feb. 1970), p. 379. DOI: 10.1086/150317 (cit. on p. 2).

[7] K. G. Begeman, A. H. Broeils, and R. H. Sanders. “Extended rotation curves of spiral galaxies : dark haloes and modified dynamics.” In: 249 (Apr. 1991), p. 523. DOI: 10.1093/mnras/249.3.523 (cit. on p. 3).

[8] D. J. Fixsen. “The temperature of the cosmic microwave background”. In: *The Astrophysical Journal* 707.2 (Nov. 2009), p. 916. DOI: 10.1088/0004-637X/707/2/916. URL: https://dx.doi.org/10.1088/0004-637X/707/2/916 (cit. on p. 3).

[9] Planck Collaboration. “Planck 2018 results - IV. Diffuse component separation”. In: *A&A* 641 (2020). DOI: 10.1051/0004-6361/201833881. URL: https://doi.org/10.1051/0004-6361/201833881 (cit. on p. 4).

[10] Marc Schumann. “Direct detection of WIMP dark matter: concepts and status”. In: *Journal of Physics G: Nuclear and Particle Physics* 46.10 (Aug. 2019), p. 103003. ISSN: 1361-6471. DOI: 10.1088/1361-6471/ab2ea5. URL: http://dx.doi.org/10.1088/1361-6471/ab2ea5 (cit. on pp. 4, 5).

[11] Titus Mombächer. *Dark matter searches at LHCb*. 2021. arXiv: 2111.00306 [hep-ex] (cit. on p. 5).

[12] Yiming Abulaiti. *Status of searches for dark matter at the LHC*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2799299 (cit. on p. 5).

[13] Mattia Di Mauro et al. “Constraining the dark matter contribution of γ rays in clusters of galaxies using Fermi-LAT data”. In: *Phys. Rev. D* 107 (8 Apr. 2023), p. 083030. DOI: 10.1103/PhysRevD.107.083030. URL: https://link.aps.org/doi/10.1103/PhysRevD.107.083030 (cit. on p. 5).

[14] G. Angloher et al. “Detector development for the CRESST experiment”. In: (2023). arXiv: 2311.07318 [physics.ins-det] (cit. on p. 5).

[15] Belina von Krosigk et al. “DELight: a Direct search Experiment for Light dark matter with superfluid helium”. In: (2022). arXiv: 2209.10950 [hep-ex] (cit. on p. 5).

[16] J. Aalbers et al. “DARWIN: towards the ultimate dark matter detector”. In: *Journal of Cosmology and Astroparticle Physics* 2016.11 (Nov. 2016), pp. 017–017. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2016/11/017. URL: http://dx.doi.org/10.1088/1475-7516/2016/11/017 (cit. on pp. 5–8).

[17] E. Aprile et al. “Projected WIMP sensitivity of the XENONnT dark matter experiment”. In: *Journal of Cosmology and Astroparticle Physics* 2020.11 (Nov. 2020), pp. 031–031. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2020/11/031. URL: http://dx.doi.org/10.1088/1475-7516/2020/11/031 (cit. on p. 10).

[18] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688 (cit. on p. 10).

[19] R. Abbasi et al. “Framework and tools for the simulation and analysis of the radio emission from air showers at IceCube”. In: *Journal of Instrumentation* 17.06 (June 2022), P06026. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/06/p06026. URL: http://dx.doi.org/10.1088/1748-0221/17/06/P06026 (cit. on p. 10).

[20] Dacheng Xu et al. *XENONnT/alea: v0.2.4*. Version v0.2.4. Mar. 2024. DOI: 10.5281/zenodo.10829030. URL: https://doi.org/10.5281/zenodo.10829030 (cit. on p. 10).

[21] Kenneth J. Mighell. “Parameter Estimation in Astronomy with Poisson-distributed Data. I. The χ^2_γ Statistic”. In: *The Astrophysical Journal* 518.1 (June 1999), p. 380. DOI: 10.1086/307253. URL: https://dx.doi.org/10.1086/307253 (cit. on pp. 13, 22).

[22] Y. Avni. “Energy spectra of X-ray clusters of galaxies.” In: 210 (Dec. 1976), pp. 642–646. DOI: 10.1086/154870 (cit. on pp. 18, 19).

[23] Sebastian Vetter. “Simulations of the radiogenic neutron background”. Internal DARWIN wiki. 2022 (cit. on p. 24).

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 01. April 2024