DISS. ETH NO. 29205

Direct WIMP Search with the Dual-Phase Data of the Argon Dark Matter Experiment

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES
(Dr. sc. ETH Zürich)

presented by
Alexander Martin Stauffer
MSc. ETH Zürich

born on *03.01.1995*

accepted on the recommendation of
Prof. Dr. André Rubbia
Prof. Dr. Günther Dissertori

2023

Abstract

The **Ar**gon **D**ark **M**atter (ArDM) experiment is the first tonne-scale dual-phase argon-based direct dark matter detection experiment. It searches for nuclear recoils of "Weakly interacting massive particles" (WIMPs), a promising dark matter candidate, with the argon nuclei. The recoils produce primary scintillation light signals (S1s) within the liquid phase, and free electrons, which drift towards the gaseous phase inducing the secondary scintillation light signals (S2s) therein. The data reconstruction extracts fundamental characteristics of the S1 and S2 signals and matches a number of S2 signals to the original S1 signal. Subsequent data analysis allows the calibration of experimental parameters and the confinement of a region of interest, wherein a particularly high likelihood of nuclear recoils is expected, based on a dedicated neutron calibration data set. In this region of interest the ArDM WIMP search is performed within a fiducial volume of $m_{\mathrm{Ar,fid}} = 574.57$ kg and during a measurement period of 8.429 live-days. The corresponding ArDM exposure of 4'843 kg $\cdot$ days results in an upper exclusion limit on the spin-independent WIMP-nucleon interaction cross section of $\sigma_p^{\mathrm{SI}} = (2.28 \pm 0.02) \cdot 10^{-41}$ cm^2 for a WIMP candidate of mass $m_\chi = 100$ GeV and a minimal upper exclusion limit of $\sigma_p^{\mathrm{SI}} = (2.16 \pm 0.02) \cdot 10^{-41}$ cm^2 for a WIMP candidate of mass $m_\chi = 120.2$ GeV at the 90% confidence level. With the first successful completion of a tonne-scale dual-phase argon-based dark matter experiment ArDM represents a major milestone in the search for dark matter.

This thesis introduces the reader to the history of the search for dark matter and the current status of the field. The theoretical framework of the WIMP interaction in a liquid noble gas detector is explored, and adapted to the presented ArDM detector. The reader is accompanied through the data selection and reconstruction process based on the data collected by ArDM. A detailed overview of the subsequent analysis of the ArDM data is provided with a focus on the discrimination of signal-like nuclear recoils versus a background of electronic recoils, and the estimation of the detector's resulting efficiency for WIMP-like nuclear recoils. The thesis concludes with a thorough explanation of the statistical procedure involved in setting an upper limit on the expected WIMP signal. The results of the ArDM WIMP search are compared to the current world-leading experimental results, and an outlook to the future of argon-based direct dark matter detection experiments is presented.

Kurzfassung

Das **A**rgon **D**ark **M**atter (ArDM) Experiment ist das erste zwei-phasige, Argon-basierte, direkte Detektionsexperiment für Dunkle Materie auf der Tonnen Skala. Es sucht nach nuklearen Rückstössen zwischen den Argon Kernen und den "Weakly interacting massive particles" (WIMPs), einem vielversprechenden Dunkle-Materie-Kandidaten. Die Rückstösse produzieren primäres Szintillationslicht (S1) in der flüssigen Phase, und freie Elektronen, welche bis zur gasförmigen Phase driften, um darin das sekundäre Szintillationslicht (S2) zu induzieren. Die Datenrekonstruktion extrahiert grundlegende Eigenschaften der S1 und S2 Signale und ordnet eine freie Zahl an S2 Signalen den originalen S1 Signalen zu. Die darauffolgende Datenanalyse erlaubt die Kalibration der experimentellen Parameter und die Einschränkung einer Interessensregion, in der eine hohe Wahrscheinlichkeit der nuklearen Rückstösse erwartert wird, basierend auf einem separaten Neutronen Kalibrationsdatensatz. In dieser Interessensregion wird die ArDM WIMP Suche in einem Referenzvolumen von $m_{\mathrm{Ar,fid}} = 574.57$ kg und während einer Messzeit von 8.429 Live-Tagen durchgeführt. Die entsprechende ArDM Datengrösse von 4'843 kg $\cdot$ Tage führt zu einem oberen Exklusionslimit des spin-unabhängigen WIMP-Nukleonen Interaktionswirkungsquerschnittes von $\sigma_p^{\mathrm{SI}} = (2.28 \pm 0.02) \cdot 10^{-41}$ cm^2 für einen WIMP Kandidaten der Masse $m_\chi = 100$ GeV und einem minimalen oberen Exklusionslimit von $\sigma_p^{\mathrm{SI}} = (2.16 \pm 0.02) \cdot 10^{-41}$ cm^2 für einen WIMP Kandidaten der Masse $m_\chi = 120.2$ GeV bei einem Vertrauenslevel von 90%. ArDM repräsentiert einen fundamentalen Meilenstein in der Suche nach dunkler Materie durch die erste erfolgreiche Durchführung eines zwei-phasigen, Argon-basierten, Dunkle-Materie-Experiment der Tonnen Skala.

Diese These führt den Leser in die Geschichte der Suche nach dunkler Materie und den aktuellen Status des Feldes ein. Das theoretische Prinzip der WIMP Interaktion in Detektoren mit flüssigen Edelgasen wird vorgezeigt und adaptiert an den vorgestellten ArDM Detektor. Der Leser wird anhand der ArDM Daten durch den Datenselektions- und -rekonstruktionsprozess begleitet. Es wird eine detaillierte Übersicht der nachfolgenden Analyse der ArDM Daten gegeben mit einem Fokus auf der Unterscheidung der signalähnlichen nuklearen Rückstösse von dem Hintergrund der elektronischen Rückstösse, und der resultierenden Effizienz des Detektors für WIMP-ähnliche nukleare Rückstösse. Die These endet mit einer rigorosen Erklärung der statistischen Auswertung, welche notwendig ist, um dem erwarteten WIMP Signal ein oberes Limit zuzuordnen. Das Resultat der ArDM WIMP-Suche wird mit den aktuellen weltführenden experimentellen Resultaten verglichen und es wird ein Ausblick in die Zukunft der Argon-basierten direkten Detektionsexperimente für Dunkle Materie gegeben.

Acknowledgements

First and foremost, I am deeply grateful to my supervisor, Professor Dr. André Rubbia, for his invaluable advice, continuous support, and extraordinary patience during my PhD study. Without his continuous belief in my abilities, mentoring of my progress and outstanding emotional investment in this journey, the personal and professional growth experienced during my PhD would never have been possible.

My sincere gratitude extends to my co-supervisor, Professor Dr. Günther Dissertori, for his treasured support and fantastic mentoring. He has been a pillar of advice and emotional support, which significantly impacted the progression of my PhD.

Additionally, I would like to express my gratitude to Professor Dr. Paolo Crivelli, who, with his honesty, professional experience and crucial emotional support, helped me during some rougher patches of my PhD and is fundamentally responsible that I managed to keep going.

Furthermore, I would like to sincerely thank Dr. Balint Radics as well as my colleagues Botao Li and Johannes Wüthrich, who helped me tremendously on the PhD journey, professionally as well as personally. This PhD would have turned out significantly worse without the countless times I was able to consult the outstanding knowledge of each of them and the fundamental collaboration with Botao Li specifically.

I would like to extend my special thanks to all my PhD colleagues who supported me emotionally and accompanied me on this journey. It was a pleasure to get to know all of them and also experience each of their journeys as well. Shared suffering is much more bearable.

I am also deeply grateful for the emotional support received from my dearest friends and family. Without their invaluable continuous support and endurance of listening to my complaints I would not have been able to carry through and remain mentally healthy. Special thanks to my parents, Kurt and Barbara Stauffer, who have always supported and enabled my academic ambitions.

Contents

Introduction

The history of the discovery of dark matter is very interesting and closely intertwined with the history of the discovery of the universe. In the end of the 19^{th} century early astronomical photographs had revealed completely dark regions in the sky. Naturally, the question arose whether these regions are entirely empty of stars. In 1904, the Irish mathematical physicist and engineer William Thomson, 1^{st} Baron Kelvin treated the Milky Way galaxy like a thermodynamic gas of stars around the sun in the center. Based on the observed velocities of nearby stars, he estimated the size of the Milky Way galaxy and the number of sun-like stars in it. He stated: "Many of our supposed thousand billion stars, perhaps a great majority of them, may be dark bodies;" [1] (p. 274). In 1906, in an answer to Lord Kelvin, the French mathematician Henri Poincaré first used the term "matière obscure" or "dark matter" in a perception of matter, that shone out or still has to become bright. But he also argued, that Lord Kelvin's projected number of stars already agreed with observations and therefore dark stars must not be significant in number [2].

In 1922, the Dutch astronomer Jacobus Kapteyn published one of the first attempts to model the arrangement of the Milky Way galaxy, using again the kinetic theory of gases. He proposed rotational motion in the Milky Way galaxy and that the sun could indeed be off center. However, he estimated the distance of the sun to the galactic center to be less than or equal to 700 parsec (1 pc $\approx 3.0857 \cdot 10^{16}$ m), and consequently assumed the sun to be at the center and the intermediate galactic neighborhood to be motionless in his calculations. Assuming a Maxwellian distribution of the velocities of nearby stars and including the observations thereof, he calculated the average star mass in ellipsoidal shells of the Milky Way galaxy. On the possibility of dark matter he stated: "We therefore have the means of estimating the mass of dark matter in the universe. As matters stand at present it appears at once that this mass cannot be excessive." [3].

In 1927, the Dutch astronomer Jan Hendrik Oort, a former student of Kapteyn, heavily favored a rotating Milky Way galaxy. Based on the assumption of the observed velocities of globular clusters being entirely rotational, he estimated the distance from the sun to the center to be 5900 pc (later corrected to 5100 pc). He also stated that a majority of clusters would escape the system due to their high velocity, and proposed an attracting and non-observable mass of $8 \cdot 10^{10}$ solar masses ($M_\odot$) at the center of the galaxy, in order to provide enough gravitational potential to contain these high velocity clusters. He remarked: "There are indications that enough dark matter exists to blot out all galactic starclouds beyond the limits of the Kapteyn-system." [4]. In 1928, the Swiss physicist Fritz Zwicky published his first article studying the thermodynamic equilibrium in the universe, wherein he treated "dust" as an intermediate reaction step between atoms and stars. He concluded that matter in the universe should exist either concentrated in stars or in an evaporated gaseous form. But he also stated: "This of course cannot mean that local conditions like the vicinity of stars cannot make the existence of dust particles more probable (dark clouds in interstellar space?)." [5]. In 1929, the American astronomer Edwin Hubble made the first estimation of the Hubble constant [6], the velocity (per distance) with which all galactic-scale-objects are moving away from each other (as a function of the distance). This inspired the German theoretical physicist Albert Einstein and the Dutch mathematical physicist and astronomer Willem de Sitter to propose the Einstein-de Sitter universe which corresponds to a flat universe filled only with matter [7]. In 1932, Oort used the acceleration of stellar bodies towards the galactic Milky Way plane to estimate the local total matter density and

therefore also the local dark matter density and wrote: "We may conclude that the total mass of nebulous or meteoric matter near the sun is less than 0.05 $M_\odot/\mathrm{pc}^3$ or $3{\cdot}10^{-24}$ g/cm^3; it is probably less than the total mass of visible stars, possibly much less." [8]. In 1933, Zwicky researched the redshift of intergalactic nebula and stated (translated): "To result in an average Doppler effect of 1000 km/sec, as observed, the average density in the Coma cluster would have to be at least 400 times larger than the shining mass derived from observations." [9].

The realization that dark matter might in fact occupy a fraction of the total matter much larger than the one of ordinary matter concludes the first part of the history of dark matter. As technology evolved and measurements of galactic and intergalactic objects became much more precise and numerous, so evolved the understanding of the universe. Modern science has resulted in some prolific knowledge regarding the previously speculated concepts. Today the distance from the sun to the center of the Milky Way galaxy is estimated to be $7.86 \pm 0.14 \pm 0.04$ kpc [10]. Oort's attracting and non-observable mass at the center of the galaxy is accounted to the black hole named Sagittarius A* with a mass of $(4.03 \pm 0.16 \pm 0.04) \cdot 10^6\ M_\odot$ [10]. The existence of cold dark matter (CDM) is fundamental in the currently favored ΛCDM model of the universe. The cosmological constant Λ accounts for the unknown dark energy which acts akin to negative pressure. Therefore, dark energy pushes space away from itself and fuels the expansion of space which explains Hubble's observation of all galactic objects moving away from each other. In the ΛCDM model, the energy content of the universe is estimated to be divided into ∼71.9% of dark energy, ∼23.5% of dark matter and ∼4.6% of ordinary matter (from the "Wilkinson Microwave Anisotropy Probe", WMAP spacecraft [11], and similarly from the Planck space observatory [12]). The Milky Way galaxy is seen as a barred spiral galaxy with a spherical dark matter halo [13], and local dark matter densities of 0.005 to 0.01 $M_\odot/\mathrm{pc}^3$ are proposed [14]. Zwicky's revolutionary claim of an abundance of dark matter over ordinary matter proved to be true, albeit at a different magnitude, and today dark matter is estimated to be roughly five times more abundant in the universe than ordinary matter. While quite a lot is known about ordinary matter and its composition, modern science still is in the dark on the actual composition of dark matter and its characteristics despite its abundance in the universe. The list of potential dark matter candidates is extensive and includes concepts such as primordial black holes [15], sterile neutrinos [16], axion-like particles [17, 18], supersymmetry [19, 20], extra dimensions [21], massive compact halo objects (MACHOs) [22], weakly interacting massive particles (WIMPs) [23,24] and modified gravity [25]. While numerous models for the origin of dark matter exist, no conclusive evidence has been found yet to identify one or multiple models as plausible or correct. MACHOs specifically, a past dark matter favorite, have not been excluded but found to be only abundant enough to account for roughly 20% [26] or even lower than 8% [27] of the dark matter content of the Milky Way galaxy. While traces of the gravitational interaction of dark matter have been measured, there is no evidence of any other interaction with ordinary matter, which would allow to probe its nature with known techniques.

To test hypotheses about dark matter candidates is the goal of ArDM, the "**Ar**gon **D**ark **M**atter" experiment. The research of ArDM is exploring the possibility of the WIMP as a viable dark matter candidate and the gain of insights on its nature via direct detection. Dark matter particles are expected to constantly hit the Earth, as it travels through the dark matter halo. The WIMPs are hypothesized to scatter with the argon atoms in the tonne-scale liquid argon detector of ArDM through very weak interaction via either a

known non-gravitational interaction or an unknown new interaction. Consequently, this would lead to scintillation light and ionization charges which could then be detected and result in insights on the nature, abundance and interaction probability of the dark matter particles. ArDM looks to contribute to the global dark matter search by either finding the WIMP or by excluding a part of the parameter space. This doctoral thesis conducts the data analysis of the final ArDM data and the resulting WIMP search. In Chapter 1 the history of dark matter is continued into modern times to give the reader an overview of previous efforts to find dark matter and the current status of the field. In Chapter 2 the ArDM detector is introduced as well as the specific detection technique which should provide the reader with a physical understanding of the working principle of the ArDM dark matter search. Chapter 3 summarizes the collected data and describes the tentative data selection and pre-processing and Chapter 4 concludes the processing chain to guide the reader through the challenges of data cleansing and processing. Chapter 5 introduces the expected experimental backgrounds and explains what information can be gathered about them from the available data. This provides the reader with an adequate understanding of the challenges in the discrimination of backgrounds in experiments such as ArDM. In Chapter 6 the actual WIMP search is conducted and an exclusion limit is set in the parameter space of the spin-independent WIMP-nucleon interaction cross section and the unknown WIMP mass. Finally, a preview to the future of the ArDM dark matter search is given in the Outlook to conclude this doctoral thesis. Additional material to complement the understanding as well as a summary of the chapter-specific contributions of the author can be found in the Appendices A and B respectively.

1 Dark matter, WIMPs and their direct detection on Earth

This introductory chapter is divided into three parts. Section 1.1 continues the story of the discovery of dark matter in modern days. From there on, this doctoral thesis focuses on the WIMP as the dark matter candidate of interest. While the presence of dark matter is evident, its nature is unclear. Section 1.2 presents a brief overview of attempts to detect WIMPs in the present days in order to possibly conclude on its nature, and provides the reader with an assessment of the status of the field. Lastly, Section 1.3 derives the theoretically expected WIMP interaction spectrum on Earth, as it constantly travels through the WIMP halo. This chapter aims to complement the reader's phenomenological understanding of dark matter in the universe with a theoretical perspective on the behavior of WIMPs. All three parts are summaries of the present knowledge of the scientific community which are assembled by the author and do not claim to be all-encompassing.

1.1 The recent history of dark matter

This section will extend the story of the discovery of dark matter, started in the Introduction, into modern times. Some of the most cited scientific experiments showing the existence of dark matter are presented. These include the galaxy rotation curves (Section 1.1.1) which point to missing mass in the Milky Way and other galaxies. With finding MACHOs to not be abundant enough to explain dark matter, particle physics eventually entered the search with the WIMP candidate. The rise of precision technology significantly defined the view of dark matter (Section 1.1.2) and lastly, gravitational lensing calculations based on an X-ray image of the bullet cluster showed, for the very first time, a reconstructed image of the presence of dark matter (Section 1.1.3).

1.1.1 Galaxy rotation curves

In classical Newtonian physics the gravitational force exerted from one point-like object on another is determined by their masses m and M, the distance r separating them and the gravitational constant G. In a stable system in equilibrium, where one lighter mass m is rotating with velocity $v(r) = |\vec{v}(r)|$ around a much heavier mass M approximately at rest, the gravitational force should be compensated by the centripetal force. This simple physical equilibrium is, in good approximation, not only applicable to the Moon-Earth and the Earth-Sun system, but also to the rotation of stellar objects around the galactic center. In order to properly describe the rotation of stellar objects, the point-like center of mass M has to be replaced by the mass $M(r)$ of the galaxy which is encompassed up until the radius r. Hence, the mass $M(r)$ is the integral of the spherically symmetric density $\rho(r)$ over radial shells $\mathrm{d}V$ as

$$G\frac{mM(r)}{r^2} = m\frac{v^2(r)}{r}, \qquad \text{where} \qquad M(r) = \int_0^r \rho(r^*)\,\mathrm{d}V(r^*)\,. \tag{1.1}$$

The dependence of the rotational velocity $v(r)$ on the distance r to the center can easily be gathered as

$$v(r) = \sqrt{G\frac{M(r)}{r}} = \sqrt{G\frac{\int_0^r \rho(r^*)\,\mathrm{d}V(r^*)}{r}}, \tag{1.2}$$

where the question remains what to assume for the density $\rho(r)$ within a galaxy. The most simplistic assumption for a spherical model galaxy of size R (and hence spherical volume $\frac{4\pi}{3}R^3$) and total mass M_{Tot} would be a constant density ($\rho = \text{constant}$) within the galaxy and zero outside the galaxy. Hence the mass $M(r)$ encompassed in the spherical volume of radius r and the rotational velocity at the radius r are

$$M(r) = \begin{cases} \frac{4\pi}{3}r^3\rho, & \text{if } r < R. \\ \frac{4\pi}{3}R^3\rho = M_{\text{Tot}}, & \text{otherwise.} \end{cases} \quad \rightarrow \quad v(r) = \begin{cases} \sqrt{\frac{4\pi G\rho}{3}r^2} \propto r, & \text{if } r < R. \\ \sqrt{\frac{4\pi G\rho}{3}\frac{R^3}{r}} \propto r^{-\frac{1}{2}}, & \text{otherwise.} \end{cases} \tag{1.3}$$

Consequently, one would expect the rotational velocity $v(r)$ in the perfect spherical galaxy to increase proportionally to r within the radius R, where the majority of stellar objects reside. The rotational velocities of the few scattered stellar objects at larger radii than R are expected to decrease proportionally to $r^{-\frac{1}{2}}$. While this model is simplistic and only serves the purpose of presenting the general phenomenon, more sophisticated models of the radial dependence of the rotational velocity always include a linear increase below R and then a drop-off beyond R. The radius R is usually assigned to the size of the central bulge of the galaxy which encompasses the majority of the visible mass.

Since Oort drew conclusions on the matter density in the Milky Way galaxy based on the observation of velocities of stellar bodies in 1932 [8], there have been many efforts to measure the radial dependence of the rotational velocity in galaxies. Even though it is an injustice to all other significant contributions before and after, the work of the American astronomers Vera Rubin and Kent Ford is often singled out due to being the first precision measurement of galaxy rotation curves. They used the Doppler redshift and blueshift of the hydrogen emission lines to measure the rotational velocities of stellar objects within galaxies, first of the nearby Andromeda galaxy in 1969 [28] and later, together with the American physicist Norbert Thonnard, of multiple other galaxies in 1978 [29] and 1980 [30].

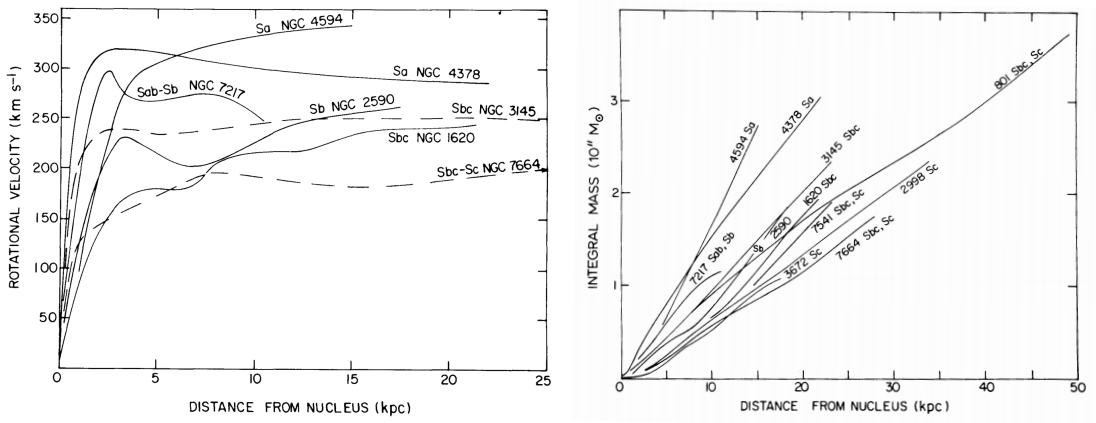

Figure 1.1: *Left*: The smoothened rotational velocity curves as a function of the distance to the core for seven galaxies from Rubin, Ford and Thonnard (1978) [29]. While the sharp increase at low radii is as expected, the approximately flat continuation at higher radii showed the inaccuracy of galaxy models containing only ordinary luminous matter. *Right*: The derived mass encompassed within a disk of radius r as a function of r; also from Rubin, Ford and Thonnard (1978) [29].

The resulting galaxy rotation curves from Rubin, Ford and Thonnard's paper in 1978 can be found in Fig. 1.1 (left). They clearly show a sharp increase in the rotational velocity at low radii. However, instead of a $\propto r^{-\frac{1}{2}}$ or a similar drop-off at higher radii, the rotational velocity stays approximately constant. This implies the presence of much more matter than just the luminous one. Going back to the simplistic model of a perfect spherical galaxy, a constant rotational velocity at radii larger than R would motivate an encompassed mass $M(r) \propto r$ above R, which also is Rubin and Ford's conclusion (Fig. 1.1, right). These observations strongly motivate the presence of massive halos outside of the central bulge, filled, to a large degree, with non-luminous matter. Or in the words of Vera Rubin, Kent Ford and Norbert Thonnard in 1980: "The conclusion is inescapable that non-luminous matter exists beyond the optical galaxy." and "Is the luminous matter only a minor component of the total galaxy mass?" [30].

1.1.2 MACHOs, WIMPs and structure formation

In the 1980s two dominant ideas about the nature of dark matter were becoming more concrete. The first and more traditional one is still that dark matter could be just some form of non-luminous matter, but not necessarily a new type of constituent. This idea materialized in the concept of MACHOs or massive compact halo objects. Black holes, neutron stars, brown dwarfs, old white dwarfs or even single planets may constitute MACHOs. Generally, MACHOs are considered massive astronomical bodies in the galaxy halo which emit little to no radiation but which could explain the missing mass described in the previous paragraph. Naturally, MACHOs are difficult to detect. In 1986, the Polish astronomer Bohdan Paczyński published the idea of MACHOs and their possible detection via mircolensing events [22]. He argued that the massive dark matter candidates would act as a gravitational lense to a light source and change the apparent brightness of the source over time. Subsequently, research groups began looking for microlensed light sources in the galactic neighborhood. In 2000, the MACHO project [26] estimated a MACHO contribution of roughly 20% of the Milky Way galactic halo dark matter density by searching for areas in the Milky Way galaxy that act as micro-lenses for the brightness of objects in the Large Magellanic Cloud. In 2007, the EROS-2 project [27] estimated the contribution to be less than 8%. These findings effectively ruled out the MACHOs as the only source of dark matter. If at all, MACHOs are likely to only contribute a minimal fraction of the Milky Way galaxy dark matter content.

Simultaneously with the emergence of the MACHO candidate in the 1980s many other candidates such as neutrinos and light axions were summarized in the term WIMP, the Weakly Interacting Massive Particles [23]. WIMPs remain a promising candidate until today but have yet to be detected convincingly. The common property of WIMP candidates apart from gravitational interaction, is a very rare interaction with ordinary matter via either the weak or a completely new interaction. A further commonality often is the cold and hence relatively slow nature of the WIMP particle. This consideration originates from the astrophysical subject of structure formation. To fully understand it, one must go back to the history of the very early, hot and dense universe. Therein, one expects the particle species to be in thermal equilibrium with each other via non-gravitational interactions. Some very small anisotropies in density and hence temperature (originating e.g. from quantum fluctuations of the inflaton field) are present, which would later develop into gravitationally bound structures. Upon the expansion of the universe, the density of particles is eventually small

enough that the respective particle species cannot sustain the required interaction rate to maintain the thermal equilibrium between each other. This moment is usually referred to as a "freeze-out" and effectively determines the abundance of particle species which predominantly self-interact after freezing out. The same applies to the WIMP which, at an early stage of the universe, interacted with ordinary matter non-gravitationally. Consequently, the strength of the non-gravitational interaction with ordinary matter and the early dark matter density determined the beginning of the freeze-out and hence the starting conditions of the self-interacting evolution of the dark matter. After freezing out the WIMPs' main interaction with other particle species as well as with themselves remains gravitation. The small density anisotropies grow and become the empty or mass filled regions of the universe. The structure in the universe observed today must be a relic of the gravitational interaction of the ordinary matter with the much more abundant dark matter. In other words, the filaments of ordinary matter in the universe today must only be a trace of the underlying dark matter structure, which in turn is a consequence of the temperature, mass, density and interaction probability of dark matter. Hot dark matter is usually disregarded as its relativistic nature would have smeared out most primordial fluctuations and hence would have erased any seeds of any eventual structure. Consequently, detailed knowledge of the anisotropies in the early universe allows conclusions on the nature of dark matter.

In 1965, the American astronomers and later Nobel laureates Arno Allan Penzias and Robert Woodrow Wilson accidentally discovered the Cosmic Microwave Background (CMB) as a noise while trying to detect radio waves reflected by satellites. The CMB consists of electromagnetic radiation coming from all directions of the sky at almost the exact same temperature. As these are the reminiscent photons that cooled off during the expansion of the universe, they originate from the freeze-out (of the photons) and hence are a photograph of the universe when it was only approximately $372'000 \pm 14'000$ years old compared to its 13.7 ± 0.2 billion years now [31].

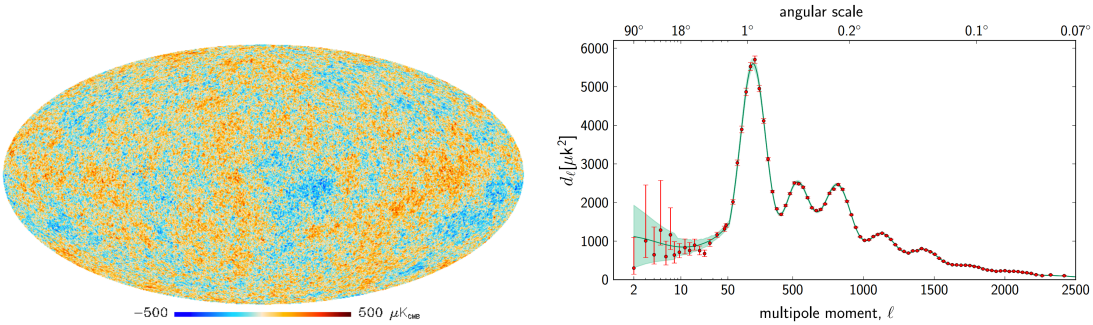

Figure 1.2: *Left*: The temperature map of the CMB anisotropies as measured by Planck. It shows the deviation from 2.7255 K after the subtraction of the dipole anisotropy and galactic plane. Taken from the Planck 2013 results (2014) [12]. *Right*: The derived power-spectrum (red dots) showing the relation of the anisotropy amplitudes to their sizes. The heights and positions of the acoustic peaks determines the cosmological parameters in a fit to the ΛCDM model (green). Also taken from the Planck 2013 results (2014) [12]. Planck reported the energy content of the universe to be ∼69.6% dark energy, ∼25.6% dark matter and ∼4.8% ordinary matter.

The first detailed and precise measurement of the CMB was done by the Cosmic Background Explorer (COBE) in 1992 [32] and 1996 [33] which led to the Nobel Prize being

awarded to the American astrophysicists George Smoot and John Mather and which heralded the age of precision cosmology. COBE found a perfect black body radiation spectrum with $T_0 = 2.73$ K and a temperature variation $\Delta T/T$ in the CMB of only $11 \cdot 10^{-6}$ which is in good agreement with the theory of an expanding universe and especially the ΛCDM model (with a non-zero cosmological constant Λ and cold dark matter). The map of anisotropies is commonly transformed into a power-spectrum to unfold the dependence of the anisotropy amplitude to its size on the map. The angular wavenumber or multipole l is a measure of the size of the anisotropy (higher l correspond to smaller sizes and $l = 100$ is approximately 1°). The heights and positions of the power-spectrum can be predicted from the cosmological model and can hence constrain the parameters of the model. In 2013, the Wilkinson Microwave Anisotropy Probe (WMAP) reported the energy content of the universe to be ∼71.9% dark energy, ∼23.5% dark matter and ∼4.6% ordinary matter according to the 9-year results. In 2014, the Planck surveyor reported ∼69.6% dark energy, ∼25.6% dark matter and ∼4.8% ordinary matter [12]. Fig. 1.2 shows Planck's CMB map (left) and the resulting power-spectrum (right). These phenomenal efforts of precision cosmology showcasing the agreement with the ΛCDM model cemented the notion of the abundance of dark matter in the universe.

1.1.3 The Bullet cluster

In 2004, Clowe, Markevitch et al. were observing the cluster 1E 0657-56, in which they found a "bullet-like" gaseous subcluster emerging [34, 35]. They had observed a collision of two clusters and potentially the beginning of a subsequent merger. In 2006, they refined their observations and published a *Direct empirical proof of existence of dark matter* [36].

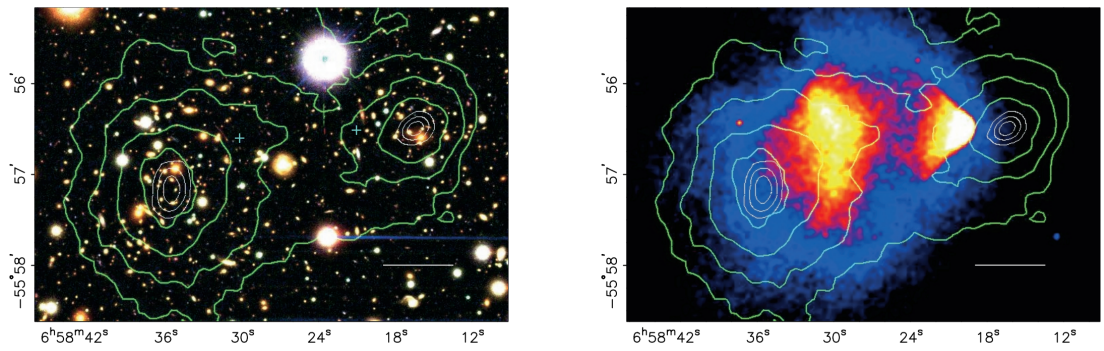

Figure 1.3: *Left*: Telescopic image of the visible light observed from the ongoing collision of the 1E 0657-56 or "bullet" cluster. *Right*: X-ray image showing the distribution of the plasma components of both merging clusters. The dark matter distribution can be reconstructed via weak gravitational lensing (green contours). The centers of the dark matter components (white contours) are eight standard deviations away from the centers of the plasma components (blue crosses) and imply that the largest parts of the cluster masses are indeed in the form of dark matter and that it is only rarely interacting non-gravitationally if at all. The scale is given in the white bar which corresponds to 200 kpc. Directly taken from Clowe, Markevitch et al. (2006) [36].

What makes the subsequently named bullet cluster so interesting, is the behavior of its main constituents. The stellar components, which can be detected by their emitted

visible light (see Fig. 1.3, left), have traversed each other without significant electromagnetic interaction. Their motion is only affected by the gravitational effect of the other cluster. The plasma (hot gas) components, which are self-interacting electromagnetically and can be detected via X-ray imaging, are focused at the center of the two clusters (see Fig. 1.3, right). Finally, the dark matter components can be reconstructed via weak gravitational lensing. The distortion of the image of galaxies behind the bullet cluster can be converted to a two-dimensional reconstruction of the mass distribution (green contours in Fig. 1.3). The centers of the reconstructed mass distributions (white contours) are eight standard deviations distant from the reconstructed mass centers of the plasma components (blue crosses). This was and is seen as maybe the most compelling evidence of the existence of dark matter. Next to its existence, this observation also implied the gravitational dominance in the interaction of dark matter. If it were to interact via any other form, it must be a very weak interaction. These observations also allowed Clowe, Markevitch et al. to make an order of magnitude estimate on the upper limit of the self-interaction cross section $\sigma_{\chi\chi}$ of dark matter (of mass m_χ) as $\sigma_{\chi\chi}/m_\chi < 1\ \mathrm{cm}^2\mathrm{g}^{-1}$ [35].

1.2 Detection of WIMPs in the present day

Section 1.1 has shown a major part of the history of the dark matter search. While there exists a compelling amount of evidence for the existence of dark matter, it has yet to be detected. Neither has a detailed theory yet been convincingly confirmed by experiment. According to the four interactions known today, such a detection would have to be made via a non-gravitational, very rare but non-vanishing, interaction. The weak interaction or an unknown new interaction are prime candidates and make the WIMP (see Section 1.1.2) the principal hypothesis to be tested by detection experiments.

In general, the detection attempts of WIMPs can be divided into the two groups of direct and indirect detection. While both types of experiments aim to detect traces, annihilation or decay products of dark matter, indirect detection experiments search for traces of WIMPs in outer space and direct detection experiments are located on Earth. In general, indirect detection experiments usually focus their efforts towards regions of the universe where a large dark matter density is expected or hypothesized, like the center of the Milky Way galaxy, near satellite galaxies or even the sun[1].

Contrarily, direct detection experiments search for traces of WIMPs on Earth and as such, already have a long history. In 1984, Drukier and Stodolsky presented their idea to detect MeV-range neutrino elastic scatterers on nuclei with recoil energies of 10 eV to 1 keV via phonons or minimal temperature changes in superconductors [38]. In 1985, Goodman and Witten remarked that the detector principle proposed by Drukier and Stodolsky could also be used to search for eV to keV nuclear recoils (NR) from elastic scatterings of GeV-range dark matter particles [39]. Goodman and Witten proposed three groups of possible nuclear recoil interactions: coherent or spin-independent weak interactions, spin-dependent interactions of roughly weak scale and strong interactions. Coherent and spin-dependent scattering are the two methods in which results of modern dark matter searches are reported. For spin-dependent interactions the scattering amplitude changes sign with the spin direction, such that only nuclei with an odd number of neutrons or an odd number

[1] Were the dark matter particle to have an anti-particle, those two could annihilate via an s-channel into a particle-anti-particle pair of a known species. Via a t-channel, the WIMP and its anti-particle could annihilate into two neutral bosons. E.g. the Large Area Telescope on the Fermi Gamma-Ray Space Telescope is looking for γ-rays in the galactic halo that could originate from such dark matter annihilations [37].

of protons can be used as a detection material. While the theory of WIMP elastic scattering will be discussed in detail in Section 1.3, there are two independent unknown dark matter parameters involved. These are usually chosen to be the WIMP mass m_χ and the spin-independent (or spin-dependent) WIMP-nucleon cross section σ_p^{SI} (or σ_p^{SD}, commonly referring to the proton p as the nucleon) as a measure of the interaction probability. Experiments searching for dark matter try to exclude as much of this two-dimensional parameter space as possible when no dark matter signal is observed. The very first exclusion limit was set by Ahlen et al. at the Homestake mine in South Dakota in 1987 using a germanium ionization detector [40]. Their upper exclusion limit (at the 68% confidence level) for the cross section was at the order of 10^{-40} cm^2 for a WIMP mass of 60 GeV (see Fig. 1.5, left). The challenges of WIMP search experiments lie, among others, in the necessary sensitivity to nuclear recoils in the low-energy range of eV to keV and the extreme scarcity of signal events and hence the rejection of background. The main backgrounds are scattering with electrons of the target nuclei or electronic recoils (ER) for short and non-WIMP originated nuclear recoils (NR) that can be induced by e.g. neutrons or muons. To minimize the flux of cosmic muons, dark matter searches are often carried out in former underground mines.

While Goodman and Witten's original idea of a superconductor detector never materialized on a large scale or sensitive to signal, some of the modern dark matter experiments are looking for phonons or heat signatures of the elastic scattering of WIMPs with nuclei. Experiments such as PICO use superheated liquids as target materials in bubble chambers, where the superheated liquid is kept close to its boiling point, such that a minimal energy deposition can lead to a phase transition and the formation of bubbles which can then be recorded. PICO set the most stringent bubble chamber direct detection limits to the spin-dependent cross section at $2.5 \cdot 10^{-41}$ cm^2 for 25 GeV WIMPs using C_3F_8 in 2019 [41]. Germanium and silicon semiconductors only require energies in the eV range to create electron-hole pairs and can hence be used to detect ionizations resulting from recoils. For example DAMIC set an exclusion limit of $\sigma_p^{\mathrm{SI}} = 3 \cdot 10^{-41}$ cm^2 for m_χ from 7 to 10 GeV in 2020 [42]. Next, scintillator crystals can be used to detect the scintillation light resulting from recoils. With the scintillation light alone, detectors could only confirm the observation of WIMP events via the annual modulation of the overall count rate but could not identify single WIMP events. DAMA (using NaI crystals) claims to have observed an annual modulation in their signal that is according to the theory of dark matter at the 8.2σ confidence level (e.g. in 2008 [43]), but the same region of the parameter space has been excluded multiple times by several other detection materials, and an independent experiment has yet to confirm the same finding using also NaI crystals.

Almost all modern direct detection WIMP search experiments are looking for one or two of the three previously measured recoil signatures: heat, ionization charge or scintillation light. Fig. 1.4 groups a selection of modern dark matter direct detection experiments according to the detection signatures used. The advantage of using two different signatures is the improved discrimination of background via the ratio of the two signatures. For example the EDELWEISS experiment used the ratio of the ionization energy and the recoil energy (a linear combination of the heat and ionization energy) to discriminate ERs from NRs [44]. EDELWEISS and also CRESST aim to search for dark matter in the MeV [45] or even sub-MeV [46] mass range. Both belong to the group of cryogenic bolometers which try to detect tiny interaction-induced temperature changes (typically ≤ 50 mK). They have excellent energy resolution and are usually paired with either an ionization (EDELWEISS) or a scintillation channel (CRESST) for the background rejection.

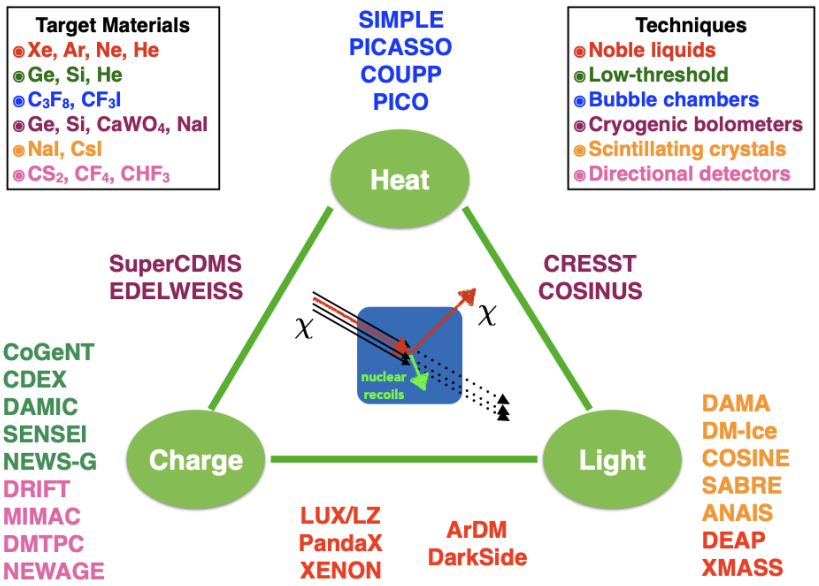

Figure 1.4: Schematic overview of WIMP search direct detection experiments. The majority of WIMP direct detection experiments aim to either detect heat or phonons, ionization charge or scintillation light, or a combination thereof. Examples are presented in the text. Schematic provided by the ArDM collaboration.

In recent years, detectors using liquid noble gases have produced the most stringent exclusion limits on σ_p^{SI}. Specifically, XENON1T excluded a σ_p^{SI} of $4.1 \cdot 10^{-47}$ cm^2 for a WIMP mass of 30 GeV in 2018 [47] while PandaX-4T excluded a σ_p^{SI} of $3.8 \cdot 10^{-47}$ cm^2 for a WIMP mass of 40 GeV in 2021 [48]; both at a 90% confidence level using 1 tonne and 3.7 tonnes of xenon as the detection material respectively. The commonly used noble gases are xenon or argon, which are excellent scintillators and easily ionized. While xenon is naturally cleaner of radioactive isotopes than the naturally abundant form of argon, the latter has two very different lifetimes of the excited states which lead to additional ER background discrimination power via the light signal pulse shape. Dual-phase detectors aim to detect the scintillation light as well as the ionization charge of recoils and the ratio of the two can be used to discriminate the ERs from the NRs. ArDM is the first tonne-scale liquid argon dual-phase detector and will be discussed in more detail in Chapter 2. DEAP used a 3.6 tonne single-phase liquid argon detector to exclude a σ_p^{SI} of $3.9 \cdot 10^{-45}$ cm^2 for a WIMP mass of 100 GeV at a 90% confidence level in 2020 [49]. And DarkSide-50, a dual-phase argon detector used 46.4 ± 0.7 kg of depleted argon to set the 90% confidence level exclusion limit of σ_p^{SI} at $1.14 \cdot 10^{-44}$ cm^2 for a WIMP mass of 100 GeV in 2018 [50]. The main goal of DarkSide-50 was to show the usability of underground argon which is depleted in ^{39}Ar, the dominant source of ER background and the main disadvantage of argon over xenon. DarkSide-50 used a fiducial volume of roughly 50 kg of underground argon, which had a reduced ^{39}Ar activity of 0.73 ± 0.11 mBq/kg as opposed to $\sim$1 Bq/kg for ordinary atmospheric argon [51]. The main reason why argon experiments are currently trailing behind the xenon experiments in terms of the attained cross section exclusion level is the lacking of a tonne-scale dual-phase argon detector using depleted argon. While the prospects for the near future will be discussed in the Outlook, ArDM's successful operation proves the feasibility of a tonne-scale dual-phase argon detector, albeit using atmospheric

argon.

A graphical overview of the time evolution of the cross section sensitivity for 60 GeV WIMPs is given in Fig. 1.5 (left). The two-dimensional 90% confidence level exclusion limits in the σ_p^{SI} versus m_χ plane can be seen in Fig. 1.5 (right). It also shows the so-called neutrino-floor, which is the region where interactions with neutrinos will become non-negligible and constitute an additional background which will be difficult to discriminate. One possibility would be to turn to directional experiments which also measure the direction of nuclear recoils. As the motion of the sun dominates the motion of the Earth in the Milky Way galactic frame (see Section 1.3.1), a peak in the WIMP interaction rate would be expected in the opposite direction of the sun's motion due to the WIMPs present in the galactic halo. DRIFT has reported a 90% confidence level exclusion limit of σ_p^{SI} at $2.8 \cdot 10^{-37}$ cm^2 for 100 GeV WIMPs in 2017 [52], using a multi-wire proportional chamber filled with a low-pressure gas to reconstruct tracks. Reaching the neutrino floor will be a significant milestone and continuing below (in the case of no discovery) will be an even more challenging endeavor.

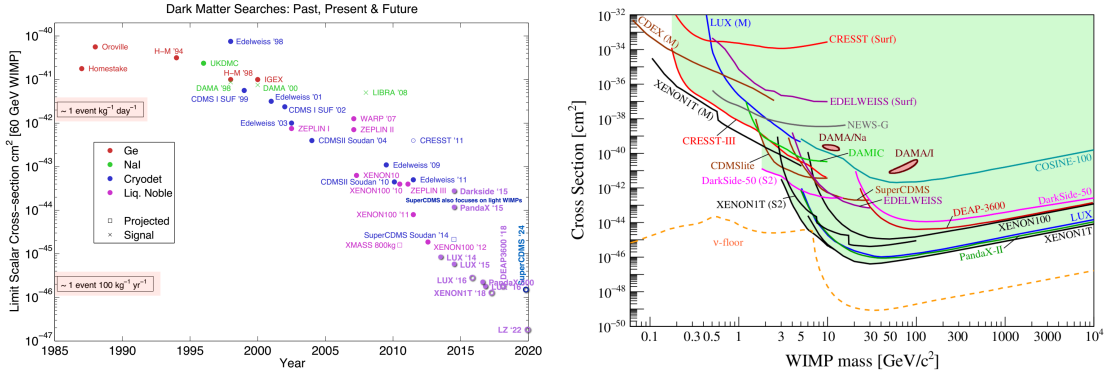

Figure 1.5: *Left*: The progression of the sensitivity to the WIMP-nucleon cross section σ_p^{SI} for 60 GeV WIMPs. Taken from Gaitskell (2020) [53]. *Right*: The current status of the field (green) of the WIMP search 90% exclusion upper limits in the two-dimensional parameter space of σ_p^{SI} versus m_χ. Taken from Billard et al. (2021) [54]. The neutrino-floor (orange) will be the scale at which neutrino interactions become a non-negligible background. More details and examples can be found in the text.

1.3 Theory of WIMPs on Earth

Having an understanding of the observational evidence of dark matter and the current experimental efforts to detect WIMPs it is time to study the theory of WIMPs. This section will give an overview of the expected WIMP spectra and differential WIMP interaction rates (on Ar and Xe) traversing the Earth as a function of the nuclear recoil energy E_{NR} (per ton and day). The expected rate of events R is proportional to the WIMP number density n_χ, their average velocity relative to the target $\langle v_\chi \rangle$ and the interaction cross section with the target σ as $R \propto n_\chi \sigma \langle v_\chi \rangle$. Accordingly, one needs to study all three of those quantities theoretically.

Section 1.3.1 will start with a summary of the Earth's velocity in the frame of the Milky Way galaxy. The importance lies in the influence of the date of the experiment on the resulting Earth's velocity, and consequently the WIMP rate. Section 1.3.2 will treat the

kinematics of nuclear recoils which is the theorized main non-gravitational interaction of WIMPs with baryonic matter. Therein one will see that the De Broglie wavelength of the nuclear recoils will be at the order of the size of Ar atoms, which necessitate a discussion of the form factor to account for the decoherence of scattering. Accordingly, Section 1.3.3 will introduce the Helm form factor to modify the cross section of the interaction. Lastly in Section 1.3.4, all components will be combined in the isothermal WIMP halo model which treats WIMPs like particles in an isothermal non-rotating gas. The particle density and velocity distribution of the WIMPs will also be discussed. The three commonly referred values of the isothermal WIMP halo model are the local WIMP density ρ_χ, the rotational velocity of the solar system through the galaxy v_0 and the local escape velocity v_{esc}, which are chosen as $\rho_\chi = 0.3$ GeVcm^{-3}, $v_0 = 230$ kms^{-1} and $v_{\text{esc}} = 600$ kms^{-1} for this work respectively. The resulting differential WIMP interaction rate $\mathrm{d}R/\mathrm{d}E_{\text{NR}}$ and the total rate above a threshold R_{aT} for Ar and Xe will be shown for the specific values of the spin-independent WIMP-nucleon cross section $\sigma_p^{\text{SI}} = 10^{-45}$ cm^2 and the WIMP mass $m_\chi = 100$ GeV. This section is targeted to readers interested in the theoretical assumptions underlying the calculations of WIMP exclusion limits. It serves as the first part of the theoretical WIMP expectation in ArDM, and as such is universally applicable. The second part in Section 2.2 is shaped after the principles of WIMP detection in dual-phase noble gas detectors specifically.

1.3.1 Earth velocity

The isothermal WIMP halo model essentially models the Earth to be flying through a cloud of WIMPs. Therefore the exact value of the Earth's velocity is of importance. The dependence of the WIMP velocity distribution on the Earth's velocity will become clear in Section 1.3.4. This section is based on the models of [55, 56] which are paraphrased here. For a more detailed description the reader should turn to the cited publications. In general, the Earth velocity $\vec{v}_{E,\text{MW}}$ with respect to the Milky Way (MW) galaxy center is composed of the galactic rotation of the solar system $\vec{u}_r$, the Sun's peculiar motion $\vec{u}_s$ and the Earth's orbital velocity relative to the Sun $\vec{u}_E$ as

$$\vec{v}_E \equiv \vec{v}_{E,\text{MW}} = \vec{u}_r + \vec{u}_s + \vec{u}_E, \tag{1.4}$$

omitting the Earth's rotation around its own axis for practical purposes and also omitting the reference frame MW superscript for brevity. The galactic rotational velocity $\vec{u}_r$ is chosen to be 230 kms^{-1}, with similar values reported in [57–59]. The Sun's peculiar motion $\vec{u}_s$ is given by its mean motion relative to nearby stars [60]. This results in

$$\vec{u}_r = \begin{pmatrix} 0 \\ 230 \\ 0 \end{pmatrix} \text{kms}^{-1}, \quad \vec{u}_s = \begin{pmatrix} 11.1 \\ 12.24 \\ 7.25 \end{pmatrix} \text{kms}^{-1} \quad \text{and} \quad \vec{u}_E = u_E(\lambda) \begin{pmatrix} \cos(\beta_x)\sin(\lambda - \lambda_x) \\ \cos(\beta_y)\sin(\lambda - \lambda_y) \\ \cos(\beta_z)\sin(\lambda - \lambda_z) \end{pmatrix}, \tag{1.5}$$

where the Earth's orbital velocity $\vec{u}_E$ relative to the Sun is taken from [55,56] and evaluated as a function of the ecliptic longitude λ (which is $\sim 0^\circ$ at the vernal equinox and increases by roughly 1° per day) and parametrized by the ecliptic latitudes $\beta_{x,y,z}$ and longitudes $\lambda_{x,y,z}$ of the x, y, z axes in galactic coordinates. The latter are summarized as

$$\beta_x = -5.5303^\circ, \qquad \beta_y = 59.575^\circ, \qquad \beta_z = 29.812^\circ$$
$$\lambda_x = 266.141^\circ, \qquad \lambda_y = -13.3485^\circ, \qquad \lambda_z = 179.3212^\circ$$

The amplitude $u_E(\lambda)$ of the Earth's orbital velocity as a function of the ecliptic longitude λ is given by

$$u_E(\lambda) = \langle u_E \rangle \left[1 - e \sin(\lambda - \lambda_0)\right], \tag{1.6}$$

where the average orbital velocity is $\langle u_E \rangle = 29.79\ \mathrm{kms}^{-1}$, the ellipticity of the Earth's orbit $e = 0.016722$ and the longitude of the orbit's minor axis $\lambda_0 = 13^\circ$. The ecliptic longitude λ can be estimated by

$$\lambda = L + 1.915^\circ \sin(g) + 0.020^\circ \sin(2g) \tag{1.7}$$

with

$$L = 280.460^\circ + 0.9856474^\circ \cdot n \quad \text{and} \quad g = 357.528^\circ + 0.9856003^\circ \cdot n \tag{1.8}$$

modulo 360°, where n is the n-th day of the year. For practical purposes, the amplitude of the Earth's velocity can also be approximated by [55]

$$\|\vec{v}_E\| \simeq 244\ \mathrm{kms}^{-1} + 15 \sin\left(2\pi \frac{(n-60)}{365.25}\right) \mathrm{kms}^{-1} \tag{1.9}$$

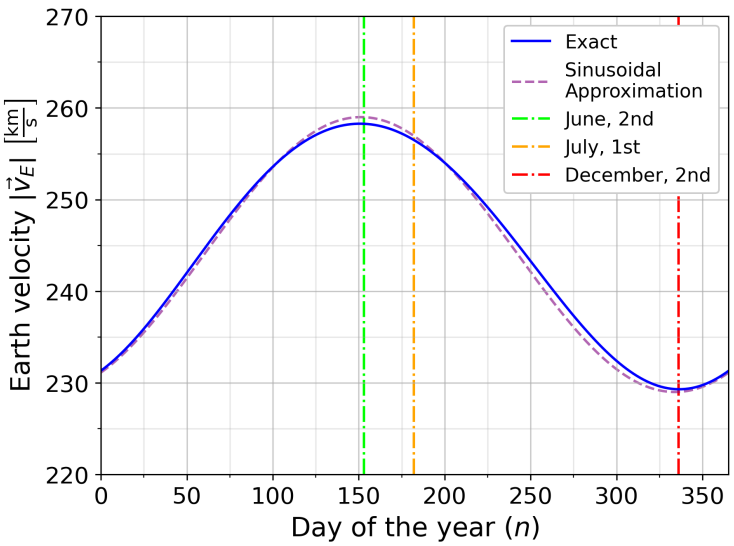

Figure 1.6: The absolute value of the Earth velocity $\vec{v}_E$ in the frame of the Milky Way galaxy in exact form (Eqs. 1.4 to 1.8) and in sinusoidal approximation (Eq. 1.9) as a function of the day of the year based on the models [55, 56]. January, 1st corresponds to the 1st; July, 1st to the 182nd and December, 31st to the 365th day of the year.

Both, the exact Earth velocity (blue, Eqs. 1.4 to 1.8) as well as the sinusoidal approximation (violet, Eq. 1.9) are shown in Fig. 1.6 as a function of n. Consequently, the Earth's

absolute velocity varies between 229.30 kms^{-1} (December, 2nd) and 258.27 kms^{-1} (June, 2nd) during the year. A higher Earth velocity generally implies higher relative WIMP velocities and increases the expected WIMP interaction rate by up to ∼5% [55]. For the purposes of the theoretical expectation in this chapter, the exact formula has been used as well as an assumed average date of data-taking of July, 1st ($n = 182$) which corresponds to an Earth velocity of $v_E = 256.51$ kms^{-1}.

1.3.2 Kinematics of nuclear recoils

The kinematics of a WIMP-nucleus interaction are those of the elastic two-body scattering. The nuclear recoil energy is the energy transferred from one particle to the other which, in the case of elastic scattering, can be written as

$$E_{\mathrm{NR}} = \frac{-Q^2}{2m_N}, \tag{1.10}$$

where m_N is the mass of the target nucleus and Q the exchanged four-momentum. Since Q^2 is a Lorentz invariant quantity, so is E_{NR} and thus one can choose any frame of reference to evaluate Q^2. The center of mass (CM) frame is the most advantageous as therein $-Q^2 = \|\vec{q}_{\mathrm{CM}}\|^2$ with the exchanged three-momentum $\vec{q}_{\mathrm{CM}}$. However, one would also prefer to describe the recoil energy as a function of variables in the fixed target (Lab) frame (e.g. the WIMP velocity relative to a target on Earth $v_\chi \equiv v_{\chi,\mathrm{Lab}} \equiv v_{\chi,\mathrm{E}}$ in Section 1.3.4) for practical purposes. In the CM frame, the momentum of the incoming WIMP $\vec{p}_{\chi,\mathrm{CM}}$ and outgoing WIMP $\vec{k}_{\chi,\mathrm{CM}}$ can be chosen to be

$$\vec{p}_{\chi,\mathrm{CM}} = \mu_N \vec{v}_{\chi,\mathrm{Lab}} = \begin{pmatrix} \mu_N v_{\chi,\mathrm{Lab}} \\ 0 \end{pmatrix} \equiv \begin{pmatrix} \mu_N v_\chi \\ 0 \end{pmatrix} \quad \text{and} \quad \vec{k}_{\chi,\mathrm{CM}} = \begin{pmatrix} \mu_N v_\chi \cos(\theta_{\mathrm{CM}}) \\ \mu_N v_\chi \sin(\theta_{\mathrm{CM}}) \end{pmatrix}, \tag{1.11}$$

where

$$\mu_N = \frac{m_\chi m_N}{m_\chi + m_N} \tag{1.12}$$

is the reduced mass of the WIMP-nucleus system and θ_{CM} is the scattering angle of the WIMP in the CM frame as illustrated in Fig. 1.7 (left). Then $\vec{q}_{\mathrm{CM}}$ can simply be evaluated as the loss of three-momentum of the WIMP as

$$\vec{q}_{CM} = \vec{k}_{\chi,\mathrm{CM}} - \vec{p}_{\chi,\mathrm{CM}} = \begin{pmatrix} \mu_N v_\chi \cos(\theta_{\mathrm{CM}}) \\ \mu_N v_\chi \sin(\theta_{\mathrm{CM}}) \end{pmatrix} - \begin{pmatrix} \mu_N v_\chi \\ 0 \end{pmatrix} \tag{1.13}$$

and hence

$$-Q^2 = \|\vec{q}_{\mathrm{CM}}\|^2 = 2\mu_N^2 v_\chi^2 \left(1 - \cos(\theta_{\mathrm{CM}})\right). \tag{1.14}$$

Consequently, the nuclear recoil energy can then be written as

$$E_{\mathrm{NR}} = \frac{\mu_N^2 v_\chi^2}{m_N} \left(1 - \cos(\theta_{\mathrm{CM}})\right) = E_\chi r \frac{\left(1 - \cos(\theta_{\mathrm{CM}})\right)}{2}, \tag{1.15}$$

where

$$E_\chi \equiv E_{\chi,\mathrm{Lab}} = \frac{m_\chi}{2} v_\chi^2 \quad \text{and} \quad r = \frac{4 m_\chi m_N}{(m_\chi + m_N)^2} \tag{1.16}$$

specify the kinetic energy E_χ of the WIMP in the Lab frame and the kinematic factor r respectively. Under the assumption of isotropic scattering (the interaction is spin-independent for noble gases) the nuclear recoil energy E_{NR} must hence be evenly distributed within $0 \leq E_{\mathrm{NR}} \leq E_\chi r$ for a specific E_χ. As will be explained in Section 1.3.4, the most probable WIMP velocity v_0 relative to Earth is commonly approximated by the local galactic rotational velocity $u_r = 230\ \mathrm{kms}^{-1}$ (see Section 1.3.1). Considering a WIMP of mass $m_\chi = 100$ GeV of velocity $v_0 \simeq 230\ \mathrm{kms}^{-1}$ and an Ar target ($m_N^{\mathrm{Ar}} = 37.2158$ GeV) one would have an upper limit of $E_\chi r \approx 25$ keV. Accordingly, the upper limit of the momentum exchanged with the Ar nucleus would be $q \approx 40$ MeV, which corresponds to a De Broglie wavelength of ≈ 4.74 fm. As this is comparable to the nuclear charge radius of the Ar atom ($r_N^{\mathrm{Ar}} \approx 3.897$ fm, see Eq. 1.20) the consideration of the influence of form factors is important and will be discussed in Section 1.3.3. The possible range for the nuclear recoil energy $0 \leq E_{\mathrm{NR}} \leq E_\chi r$ allows to write the differential WIMP event rate $\frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}}$ as a function of $\mathrm{d}R\,(E_\chi)$ as motivated visually in Fig. 1.7 (right). Then

$$\mathrm{d}R\,(E_\chi) = \mathrm{d}\left(\frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}}\right)\cdot(E_\chi r) \quad \rightarrow \quad \frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}} = \int_{E_{\mathrm{min}}}^{E_{\mathrm{max}}} \frac{1}{E_\chi r}\mathrm{d}R\,(E_\chi) = \frac{1}{E_0 r}\int_{v_{\mathrm{min}}}^{v_{\mathrm{max}}} \frac{v_0^2}{v_\chi^2}\mathrm{d}R\,(v_\chi)\,, \tag{1.17}$$

which depends on the minimal and maximal WIMP kinetic energies E_{min} and E_{max} or velocities v_{min} and v_{max} respectively. While v_{max} will be given by v_{esc} (see Section 1.3.4), v_{min} is determined by $E_{\mathrm{min}} = \frac{E_{\mathrm{NR}}}{r}$, which is the smallest particle energy which can result in a recoil of energy E_{NR}. Hence

$$v_{\mathrm{min}} = \sqrt{\frac{2E_{\mathrm{min}}}{m_\chi}} = \sqrt{\frac{E_{\mathrm{NR}}}{E_0 r}}v_0 \quad \text{with} \quad E_0 = \frac{m_\chi}{2}v_0^2 \tag{1.18}$$

being the most probable WIMP kinetic energy.

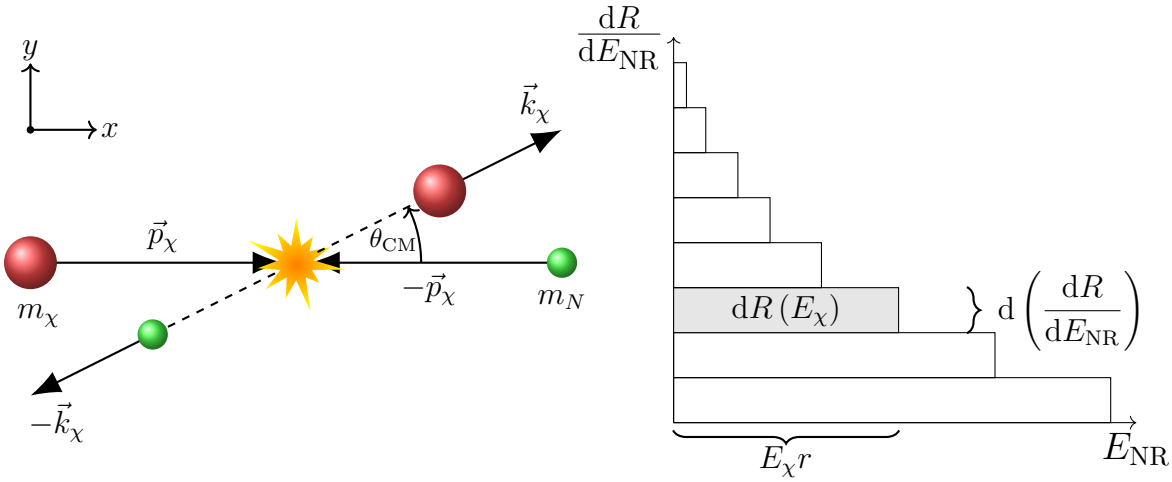

Figure 1.7: *Left*: The kinematics of the WIMP nuclear recoil on a target nucleus in the CM frame. The indicated variables are explained and used in the text. *Right*: Visualization of the differential WIMP interaction rate as a function of the WIMP energy E_χ.

1.3.3 Nuclear form factor and cross section

As already motivated in Section 1.3.2, the De Broglie wavelength of typical WIMP nuclear recoil energies is of the same order as the nuclear charge radius. This invokes the inclusion of the form factor to account for the decoherence of scattering. In practice the WIMP-nucleus cross section σ_N depends on the exchanged momentum q (here in the unit fm^{-1}, related to GeV via $\hbar c = 197.326$ MeVfm) as well as the atomic radius r_N as

$$\sigma_N\left(qr_N\right) = \sigma_{N,0} F^2\left(qr_N\right), \tag{1.19}$$

where $\sigma_{N,0}$ is the WIMP-nucleus cross section in the limit of zero exchanged momentum. Therefore $\sigma_N\left(0\right) = \sigma_{N,0}$ holds and $F\left(0\right) = 1$. The exchanged momentum q can directly be related to the nuclear recoil energy E_{NR} by using Eq. 1.10 as

$$q = \sqrt{-Q^2} = \sqrt{\|q_{\mathrm{CM}}\|^2} = \sqrt{2 m_N E_{\mathrm{NR}}}, \quad \text{and} \quad r_N \simeq 1.14 \cdot A^{\frac{1}{3}}\ \mathrm{fm} \tag{1.20}$$

allows to approximately predict the nuclear charge radius for an atom of atomic weight A [55,61]. For $A^{\mathrm{Ar}} = 40$ this results in $r_N^{\mathrm{Ar}} \approx 3.897$ fm which is comparable to the estimated De Broglie wavelength of ≈ 4.74 fm for a 100 GeV WIMP with a velocity of $v_0 \simeq 230\ \mathrm{kms}^{-1}$ impinging on Ar (see Section 1.3.2). In the first Born approximation, the form factor is the Fourier transform of the charge density of the scattering centers. As such, there are many possible choices of the form factor, but a common one is the combination of a solid sphere and a Gaussian scatterer resulting in the Helm form factor [62] given by

$$F\left(qr_N\right) = \underbrace{\frac{3 j_1\left(qr_N\right)}{qr_N}}_{\text{Solid sphere}} \underbrace{e^{-\frac{(qs)^2}{2}}}_{\text{Gaussian}} = 3\frac{\left[\sin\left(qr_N\right) - qr_N\cos\left(qr_N\right)\right]}{\left(qr_N\right)^3} e^{-\frac{(qs)^2}{2}}, \tag{1.21}$$

where j_1 is the spherical Bessel function of the first kind and $s = 0.9$ fm [55] is the nuclear skin thickness. The WIMP-nucleus cross section for spin-independent scalar interactions [19] (e.g. for noble gas targets) at zero exchanged momentum can be evaluated as

$$\sigma_{N,0} = \frac{4\mu_N^2}{\pi}\left[Z f_p + \left(A - Z\right) f_n\right]^2, \tag{1.22}$$

where Z is the atomic number of the target ($Z^{\mathrm{Ar}} = 18$) and f_p and f_n are the effective couplings to a proton and neutron respectively (in the units mkg^{-1}, or E^{-2} in natural units). Assuming $\|f_p\| \approx \|f_n\|$, one can approximate

$$\sigma_{N,0} \approx \frac{4\mu_N^2}{\pi} A^2 f_p^2 = \sigma_p^{\mathrm{SI}} \frac{\mu_N^2}{\mu_p^2} A^2, \tag{1.23}$$

where

$$\sigma_p^{\mathrm{SI}} = \frac{4\mu_p^2}{\pi} f_p^2 \quad \text{and} \quad \mu_p = \frac{m_\chi m_p}{m_\chi + m_p} \tag{1.24}$$

are the WIMP-proton (or WIMP-nucleon) cross section and the WIMP-proton reduced mass respectively. Commonly, WIMP exclusion plots are shown as a function of the two unknowns m_χ and σ_p^{SI} to remove the dependence on the target element. The Helm form factor squared as function of the nuclear recoil energy is shown for Ar and Xe in Fig. 1.8. As expected, the form factor starts at unity and then decreases with increasing recoil energy.

Eventually, the form factor starts dipping to very low values and recovering in an oscillatory manner. At which energies this motion starts highly depends on the nuclear radius. The larger the atom, the lower the required energy to resolve the nucleus (and consequently the earlier starts this oscillatory motion). In Fig. 1.8 this oscillatory behavior can be seen for Xe, while Ar will assume the same behavior but only at higher recoil energies.

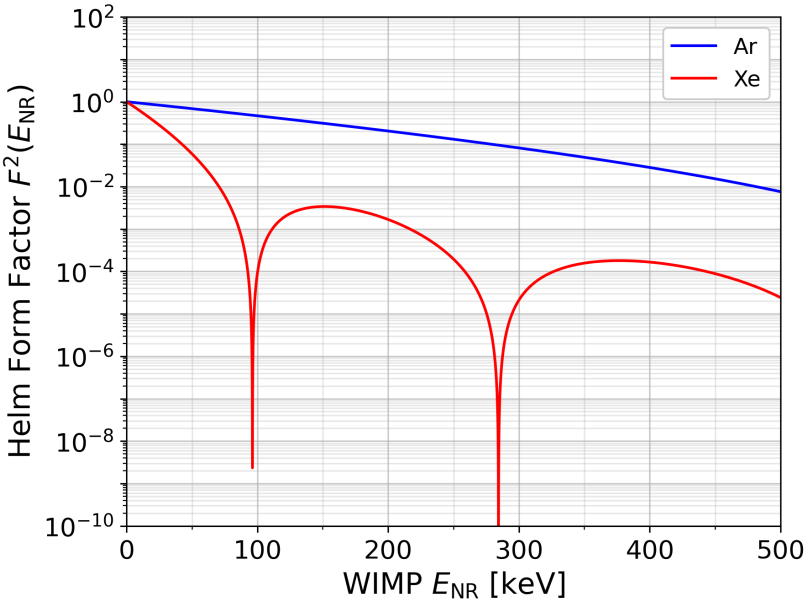

Figure 1.8: The Helm form factor (Eq. 1.21) squared of Ar and Xe as a function of the recoil energy E_{NR}. The form factor accounts for the decoherence of the scattering at higher transferred momenta and is highly dependent on the atomic radius.

1.3.4 The isothermal WIMP halo model

In 1985, when Goodman and Witten proposed the first direct detection technique (see Section 1.1), they approximated the velocity of the WIMPs on Earth as roughly 200 kms^{-1} [39]. In 1986, Drukier, Freese and Spergel then introduced what is now known as the (standard) isothermal WIMP halo model [63]. A more refined discussion is presented by Lewin and Smith in 1996 [55] and constitutes the model which is still used today in the calculation of WIMP exclusion limits. The standard isothermal WIMP halo model postulates a non-rotating WIMP halo in the Milky Way galaxy which is held up by the inherent pressure of the random velocities of the WIMPs. The WIMPs themselves behave like the atoms in an ideal gas in thermodynamic equilibrium. As such, the WIMP velocities $v_{\chi,\mathrm{MW}}$ in the Milky Way galaxy frame follow the Maxwell-Boltzmann distribution. While the velocity dispersion parameter of the Maxwell-Boltzmann distribution is not inherently clear, the related most probable velocity of the WIMPs is chosen as the local galactic rotational velocity as $v_0 = 230$ kms^{-1} (see Section 1.3.1). However, for an estimated WIMP interaction rate on Earth the decisive quantity of interest is the kinetic energy (and hence the velocity) distribution of WIMPs in the reference frame of the Earth. Therein one can simply define the WIMP velocity $\vec{v}_{\chi,\mathrm{E}} \equiv \vec{v}_{\chi}$ as

$$\vec{v}_{\chi} \equiv \vec{v}_{\chi,E} = \vec{v}_{\chi,\mathrm{MW}} - \vec{v}_{E,\mathrm{MW}} = \vec{v}_{\chi,\mathrm{MW}} - \vec{v}_{E} \tag{1.25}$$

with the Earth's velocity in the reference frame of the Milky Way $\vec{v}_{E,\mathrm{MW}} = \vec{v}_E$ given in Eq. 1.4. Consequently, in the isothermal WIMP halo model

$$\vec{v}_{\chi,\mathrm{MW}} = \vec{v}_\chi + \vec{v}_E \sim f\left(\vec{v}_E, \vec{v}_\chi\right) \propto e^{-\frac{\left(\vec{v}_E+\vec{v}_\chi\right)^2}{v_0^2}}. \tag{1.26}$$

follows the Maxwell-Boltzmann distribution f as well. Any proportionality factor in f can be set to unity for the purpose of this treatment as it will be canceled out in all theoretical estimations. Having chosen a velocity distribution one can now examine the particle density of WIMPs $\mathrm{d}n_\chi$ in the WIMP velocity space. For any general velocity distribution

$$\mathrm{d}n_\chi \left[\frac{1}{\mathrm{m}^3}\right] = n_{\chi,0}\frac{f\left(\vec{v}_E, \vec{v}_\chi\right)\mathrm{d}^3\vec{v}_\chi}{k} \tag{1.27}$$

holds. Therein $f\left(\vec{v}_\chi, \vec{v}_E\right)\mathrm{d}^3\vec{v}_\chi$ is the non-normalized fraction of particles in the infinitesimal volume $\mathrm{d}^3\vec{v}_\chi$ in WIMP velocity space. The constant k normalizes the velocity distribution and hence corresponds to the integral of $f\left(\vec{v}_\chi, \vec{v}_E\right)\mathrm{d}^3\vec{v}_\chi$ over all possible velocities. It is given by

$$k\left[\frac{\mathrm{km}^3}{\mathrm{s}^3}\right] = \int_0^{2\pi}\mathrm{d}\phi\int_{-1}^{1}\mathrm{dcos}\left(\theta\right)\int_0^{\|\vec{v}_E+\vec{v}_\chi\|=v_{\mathrm{esc}}} f\left(\vec{v}_E, \vec{v}_\chi\right)v_\chi^2\mathrm{d}v_\chi, \quad \text{such that} \quad \int \mathrm{d}n_\chi \equiv n_{\chi,0}, \tag{1.28}$$

where v_{esc} is an upper limit on the local WIMP velocities in the Milky Way frame of reference. In practice, it is reasonable to truncate all velocities above the local galactic escape velocity of $v_{\mathrm{esc}} \approx 600\ \mathrm{kms}^{-1}$ [63], but it will be left unfixed for now. As motivated earlier, the interaction rate on Earth is directly proportional to the average WIMP velocity $\langle v_\chi \rangle$ relative to Earth, which corresponds to

$$\langle v_\chi \rangle = \frac{\int v_\chi \mathrm{d}n_\chi}{n_{\chi,0}} = \frac{\int v_\chi f\left(\vec{v}_E, \vec{v}_\chi\right)\mathrm{d}^3\vec{v}_\chi}{k}. \tag{1.29}$$

For the simplified case $v_{\mathrm{esc}} = \infty$, the normalization k is commonly labeled k_0 and can be directly integrated as

$$k_0 = k\left(v_{\mathrm{esc}} = \infty\right) = \left(\pi v_0^2\right)^{\frac{3}{2}} = \pi^{\frac{3}{2}}v_0^3. \tag{1.30}$$

The resulting probability density for $v_{\mathrm{esc}} = \infty$ then corresponds to

$$\frac{1}{n_{\chi,0}}\frac{\mathrm{d}n_\chi}{\mathrm{d}\left(\|\vec{v}_E + \vec{v}_\chi\|\right)} = \frac{4\pi\|\vec{v}_E + \vec{v}_\chi\|^2 f\left(\vec{v}_E, \vec{v}_\chi\right)}{k_0} = \frac{4\pi\|\vec{v}_E + \vec{v}_\chi\|^2}{k_0}e^{-\frac{\|\vec{v}_E+\vec{v}_\chi\|^2}{v_0^2}} \tag{1.31}$$

which can be seen in Fig. 1.9 as a function of the WIMP velocity in the Milky Way galaxy frame. The most probable velocity is $v_0 = 230\ \mathrm{kms}^{-1}$ by construction.

More generally, for a Maxwell-Boltzmann distribution truncated at $\|\vec{v}_E + \vec{v}_\chi\| = v_{\mathrm{esc}}$, the resulting normalization is commonly labeled k_1 and can be evaluated as

$$k_1 = k\left(v_{\mathrm{esc}}\right) = k_0\left[\mathrm{erf}\left(\frac{v_{\mathrm{esc}}}{v_0}\right) - \frac{2}{\sqrt{\pi}}\frac{v_{\mathrm{esc}}}{v_0}e^{-\frac{v_{\mathrm{esc}}^2}{v_0^2}}\right] \tag{1.32}$$

involving the error function. One can reassuringly see that $\lim_{v_{\text{esc}}\to\infty} k_1 = k_0$ as expected. The ratio of k_0 and k_1 will enter the differential WIMP interaction rate as a correction for a finite escape velocity v_{esc}. As already mentioned, it is reasonable to assign the local galactic escape velocity of ≈ 600 kms^{-1} to v_{esc}, which only cuts away a small fraction of the tail of the velocity distribution (see Fig. 1.9). In that case k_0/k_1 equals 1.0035.

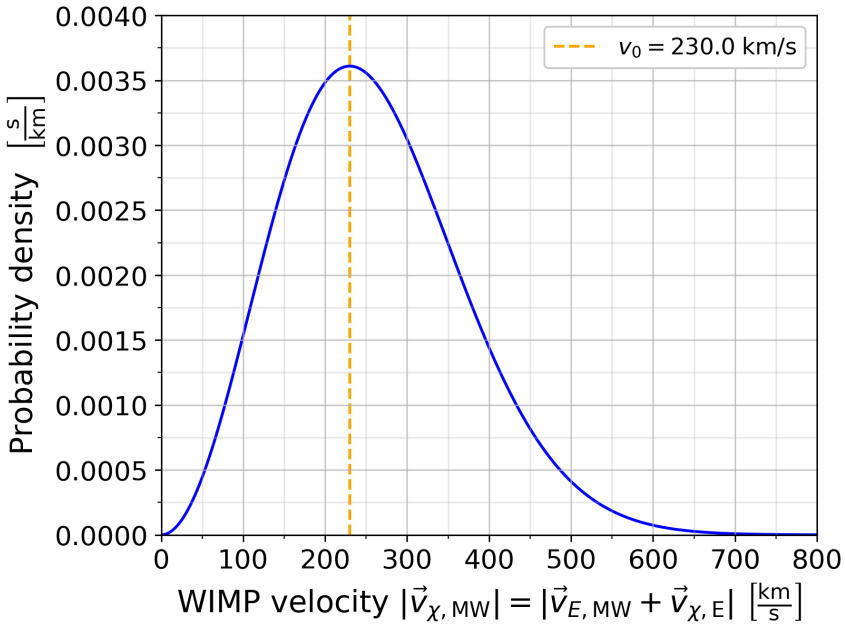

Figure 1.9: The probability density distribution (Eq. 1.31) of the absolute WIMP velocity in the reference frame of the Milky Way galaxy. The most probable velocity is $v_0 = 230$ kms^{-1} by construction.

Under the hypothesis of WIMPs being abundant in the Milky Way galaxy halo, the Earth is constantly flying through a cloud of WIMPs, or equivalently, WIMPs are constantly flying through the Earth. As a consequence, the flux Φ_χ of WIMPs on Earth can be estimated via the average velocity $\langle v_\chi \rangle$ relative to the Earth (given by Eq. 1.29) and the total WIMP particle density $n_{\chi,0}$ as

$$\Phi_\chi \left[\frac{1}{\text{m}^2\text{s}}\right] = n_{\chi,0}\langle v_\chi \rangle, \quad \text{where} \quad n_{\chi,0} = \frac{\rho_\chi}{m_\chi}, \tag{1.33}$$

with ρ_χ being the local WIMP mass density (in GeVcm^{-3}) and m_χ the unknown WIMP mass (in GeV). For the purpose of this analysis, the common choice of $\rho_\chi \simeq 0.3$ GeVcm^{-3} is adopted [19] (or [64] for a recent summary). For elastic collisions of the WIMPs with the target nuclei N the total WIMP event rate R on Earth can then be written as

$$R \left[\frac{1}{\text{kgs}}\right] = \frac{N_A}{m_a}\Phi_\chi \sigma_N = \frac{N_A}{m_a}\sigma_N n_{\chi,0}\langle v_\chi \rangle, \tag{1.34}$$

where $N_A = 6.022 \cdot 10^{23}$ mol^{-1} is Avogadro's number, m_a the atomic mass of the target medium ($m_a^{\text{Ar}} = 39.948$ gmol^{-1}) and σ_N is the unknown WIMP-nucleus interaction cross

section (discussed in Section 1.3.3). Consequently, using Eq. 1.29, the differential event rate $\mathrm{d}R$ can simply be written as a function of the WIMP velocity v_χ relative to the target as

$$\mathrm{d}R = \frac{N_A}{m_a}\sigma_N v_\chi \mathrm{d}n_\chi, \tag{1.35}$$

where the differential WIMP particle density $\mathrm{d}n_\chi$ is given by Eq. 1.27. One can again compute for the simplified case of $v_E = 0$ and $v_{\mathrm{esc}} = \infty$, that

$$\langle v_\chi \rangle (v_E = 0, v_{\mathrm{esc}} = \infty) = \frac{2v_0}{\sqrt{\pi}} \quad \text{and} \quad R_0 = R(v_E = 0, v_{\mathrm{esc}} = \infty) = \frac{2}{\sqrt{\pi}}\frac{N_A}{m_a} n_{\chi,0}\sigma_N v_0 \tag{1.36}$$

where the labeling convention of Section 1.3.4 has been kept in labeling R_0 as the rate in the case of $v_{\mathrm{esc}} = \infty$ (and here also $v_E = 0$). Including also the form factor for incoherent scattering from Section 1.3.3, Eq. 1.19 into R_0 results in

$$R_0 = \frac{2}{\sqrt{\pi}}\frac{N_A}{m_a}\frac{\rho_\chi}{m_\chi} v_0 \sigma_p^{\mathrm{SI}} \frac{\mu_N^2}{\mu_p^2} A^2 F^2 (q r_N). \tag{1.37}$$

Now the total rate can be written more simply as

$$R = R_0 \frac{\sqrt{\pi}}{2}\frac{\langle v_\chi \rangle}{v_0} = R_0 \frac{k_0}{k_1}\frac{1}{2\pi v_0^4}\int v_\chi f(\vec{v}_E, \vec{v}_\chi)\, \mathrm{d}^3\vec{v}_\chi \tag{1.38}$$

and hence the differential event rate $\mathrm{d}R(v_\chi)$ as a function of the WIMP velocity follows immediately as

$$\mathrm{d}R(v_\chi) = R_0 \frac{k_0}{k_1}\frac{1}{2\pi v_0^4} v_\chi f(\vec{v}_E, \vec{v}_\chi)\, \mathrm{d}^3\vec{v}_\chi \tag{1.39}$$

Inserting back into Eq. 1.17 (from the discussion of nuclear recoil kinematics in Section 1.3.2) and identifying $v_{\max}$ with v_{esc} finally yields

$$\frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}} = \frac{\mathrm{d}R(v_E, v_{\mathrm{esc}})}{\mathrm{d}E_{\mathrm{NR}}} = \frac{R_0}{E_0 r}\frac{1}{2\pi v_0^2}\int_{v_{\min}}^{v_{\mathrm{esc}}} \frac{1}{v_\chi} f(\vec{v}_E, \vec{v}_\chi)\, \mathrm{d}^3\vec{v}_\chi \tag{1.40}$$

as the general solution for the differential WIMP rate as a function of the nuclear recoil energy E_{NR}. The integration has solutions of increasing complexity when choosing values for v_E and v_{esc}:

$$\frac{\mathrm{d}R(0,\infty)}{\mathrm{d}E_{\mathrm{NR}}} = \frac{R_0}{E_0 r} e^{-\frac{E_{\mathrm{NR}}}{E_0 r}} \tag{1.41a}$$

$$\frac{\mathrm{d}R(0, v_{\mathrm{esc}})}{\mathrm{d}E_{\mathrm{NR}}} = \frac{k_0}{k_1}\frac{R_0}{E_0 r}\left[e^{-\frac{E_{\mathrm{NR}}}{E_0 r}} - e^{-\frac{v_{\mathrm{esc}}^2}{v_0^2}}\right] = \frac{k_0}{k_1}\left[\frac{\mathrm{d}R(0,\infty)}{\mathrm{d}E_{\mathrm{NR}}} - \frac{R_0}{E_0 r} e^{-\frac{v_{\mathrm{esc}}^2}{v_0^2}}\right] \tag{1.41b}$$

$$\frac{\mathrm{d}R(v_E,\infty)}{\mathrm{d}E_{\mathrm{NR}}} = \frac{R_0}{E_0 r}\frac{\sqrt{\pi}}{4}\frac{v_0}{v_E}\left[\mathrm{erf}\left(\frac{v_{\min}+v_E}{v_0}\right) - \mathrm{erf}\left(\frac{v_{\min}-v_E}{v_0}\right)\right] \tag{1.41c}$$

$$\frac{\mathrm{d}R(v_E, v_{\mathrm{esc}})}{\mathrm{d}E_{\mathrm{NR}}} = \frac{k_0}{k_1}\left[\frac{\mathrm{d}R(v_E,\infty)}{\mathrm{d}E_{\mathrm{NR}}} - \frac{R_0}{E_0 r} e^{-\frac{v_{\mathrm{esc}}^2}{v_0^2}}\right] \tag{1.41d}$$

The most simple case (a) concerns a stationary Earth and no local WIMP escape velocity, whereas case (b) corrects for a finite escape velocity. Case (c) includes the Earth velocity, but with infinite escape velocity. Finally, case (d) incorporates the Earth velocity and the escape velocity of the WIMPs to conclude the general derivation of the differential event rate. In its full form the differential WIMP event rate, which is used to set modern exclusion limits, is

$$\begin{aligned}\frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}} =& \frac{2}{\sqrt{\pi}} \frac{N_A}{m_a} \frac{\rho_\chi}{m_\chi} \frac{v_0}{E_0 r} \sigma_p^{\mathrm{SI}} \frac{\mu_N^2}{\mu_p^2} A^2 F^2 \left(q r_N\right) \\ & \times \left[\mathrm{erf}\left(\frac{v_{\mathrm{esc}}}{v_0}\right) - \frac{2}{\sqrt{\pi}} \frac{v_{\mathrm{esc}}}{v_0} e^{-\frac{v_{\mathrm{esc}}^2}{v_0^2}} \right]^{-1} \\ & \times \left\{ \frac{\sqrt{\pi}}{4} \frac{v_0}{v_E} \left[\mathrm{erf}\left(\frac{v_{\mathrm{min}} + v_E}{v_0}\right) - \mathrm{erf}\left(\frac{v_{\mathrm{min}} - v_E}{v_0}\right) \right] - e^{-\frac{v_{\mathrm{esc}}^2}{v_0^2}} \right\}. \end{aligned} \tag{1.42}$$

The resulting differential event rate for the commonly referred WIMP mass $m_\chi = 100$ GeV and the reference cross section $\sigma_p^{\mathrm{SI}} = 10^{-45}$ cm^2 is shown in Fig. 1.10 (left) for an average date of July, 1st. The sharp fall-off at ≈ 242 keV (for Ar) is mainly a consequence of the chosen escape velocity v_{esc}, but its position also depends on the chosen WIMP mass m_χ, the most probable WIMP velocity v_0 and the chosen average date (which determines v_E, see Section 1.3.1). Beyond this fall-off the differential rate changes sign and hence the physically meaningful energy range of the differential rate is from zero up to the fall-off only. As the differential rate scales linearly with the WIMP-proton cross section, the chosen value of σ_p^{SI} influences the y-scale proportionally. For larger target nuclei (e.g. Xe), the oscillatory behavior of the form factor mentioned in Section 1.3.3 will be directly visible in the differential rate. This is not the case for Ar as the cut-off is at a lower energy than the first dip in the form factor squared, which can be seen in Fig. 1.8.

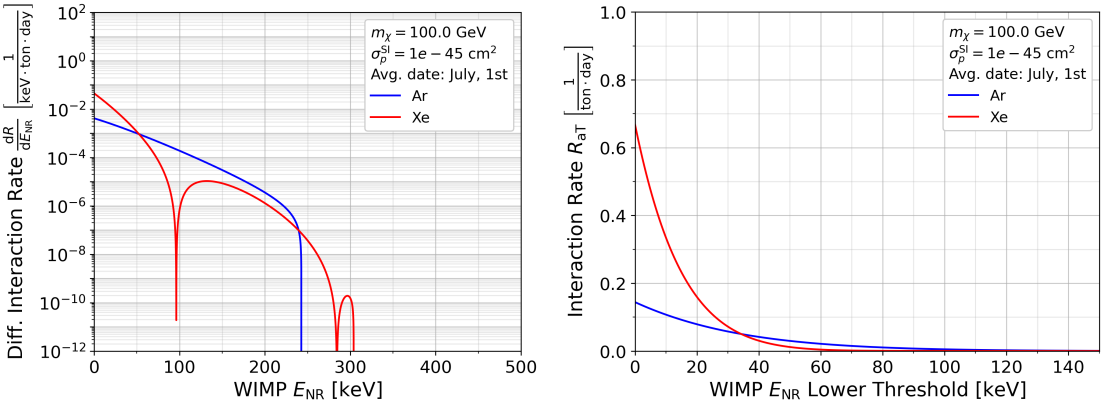

Figure 1.10: *Left*: The differential WIMP interaction rate (Eq. 1.42) on Ar and Xe targets as a function of the recoil energy E_{NR} for a chosen set of the WIMP mass m_χ, the spin-independent WIMP-nucleon cross section σ_p^{SI} and an average date of data-taking. *Right*: The total WIMP rate above a threshold of E_{NR} (Eq. 1.43) for Ar and Xe. The total expected rate of WIMP events $R_{\mathrm{aT}}(0)$ is 0.144 per ton and day for Ar and 0.668 per ton and day for Xe.

The total WIMP rate R_{aT} above a threshold T of E_{NR} can be easily gathered via numerical integration as

$$R_{\mathrm{aT}}\left(\mathrm{T}\right) = \int_{\mathrm{T}}^{\infty} \frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}} \mathrm{d}E_{\mathrm{NR}} = \int_{\mathrm{Treshold}}^{\mathrm{Fall-off}} \frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}} \mathrm{d}E_{\mathrm{NR}} \tag{1.43}$$

and can be seen in Fig. 1.10 (right). The total expected rate of WIMP events $R_{\mathrm{aT}}\left(0\right)$ is 0.144 per ton and day for Ar and 0.668 per ton and day for Xe. However, the nuclear recoil energy E_{NR} is not a directly detectable quantity. Generally, detectors have a lower energy detection threshold and background discrimination becomes increasingly difficult at lower energies, as will be thoroughly discussed and presented in Chapter 5 for the case of ArDM. The second part of the theoretical treatment (in Section 2.2) will convert the energy variable from the nuclear recoil energy E_{NR} into the ArDM observable of the primary scintillation light S1L.

2 ArDM - The Argon Dark Matter Experiment

Liquid noble gas direct detection experiments have set the most stringent limits to the spin-independent WIMP-nucleon interaction cross section σ_p^{SI} (see Section 1.2). Liquid noble gases, especially argon and xenon, bring with them a unique combination of scintillation properties, chemical inertness, scalability and the advantage that free electrons from ionization can remain free. The liquid phase is much more dense than the gaseous phase and hence has an increased probability of WIMP interactions. Dual-phase detectors use a large liquid target volume with an applied electric field to allow the free electrons to drift upwards in a straight trajectory into a small gaseous layer. Therein the ionization signal is amplified via another electric field and detected via secondary scintillation. This allows to reconstruct the event vertex position in three dimensions, essentially promoting the detector to a time projection chamber. This chapter will focus on **ArDM**, the **Ar**gon **D**ark **M**atter experiment, which implements e tonne-scale dual phase detector. Section 2.1 will discuss the physical processes involved when WIMPs or other particles recoil with the liquid noble gas and briefly describe the possible experimental backgrounds and the underlying discrimination variables of ArDM. Knowing the physical challenges that a liquid noble gas direct detection experiment has to fulfill, the section will conclude with a technical overview of the ArDM detector including its location, design and data acquisition system. Section 2.2 will continue the theoretical study of the interaction properties of WIMPs started in Section 1.3. Therein the total rate of WIMPs inside ArDM will be derived as a function of the ArDM observable S1L, the light of the primary scintillation. The conclusion will be a projected WIMP sensitivity for the ArDM WIMP search. This chapter aims to give the reader an overview of the physical processes involved as well as the requirements and technical challenges a modern liquid noble gas direct detection experiment has to fulfill. The content of this chapter consists of general knowledge of the ArDM collaboration and existing research on liquid noble gases which are summarized by the author.

2.1 The ArDM detection principle

This section will discuss the particle detection principle of ArDM in detail. Upon interaction of a WIMP inside the active liquid argon volume of ArDM, the induced nuclear recoil will lead to scintillation light as well as free ionization charge. Similarly, β particles interact predominantly with the electrons of the target nuclei in electron recoils constituting the largest experimental background. These physical processes will be discussed in Section 2.1.1 in the context of liquid argon and the imposed technical challenges that any liquid noble gas direct detection experiment has to manage. A brief technical overview of the detector and an explanation of the experimental setup will be given in Section 2.1.2.

2.1.1 Recoil types, scintillation and ionization

Particles traversing the detector either interact predominantly with the electrons or the nuclei of the argon atoms depending on their particle type[2]. Heavy charged particles, such as α particles, predominantly interact with the electrons of the argon atoms via Coulomb scattering. They undergo various interactions along their tracks in quick succession while continuously loosing energy. Similarly, β particles also continuously Coulomb-interact with

[2] An excellent review of the interactions of different particle types can be found in Knoll's second chapter of "Radiation Detection and Measurement" [65]. This also served as the basis for this section.

the electrons in the target medium while gradually slowing down. Fast β particles also might loose energy via radiative processes such as bremsstrahlung, but this is negligible for the energy range of interest in WIMP searches. Additionally, γ rays interact with the electrons predominantly via Compton scattering or the photo-electric effect. All these particle types interacting with the electrons either excite or ionize the argon atoms and are summarized as **electron recoils (ERs)** of energy E_{ER}. In this context one defines also the electron equivalent energy E_{ee} as the energy that is transferred to the electrons of the target medium. Undergoing continuous interactions, ERs quickly deploy all their energy to the electrons and hence $E_{\mathrm{ee,ER}} = E_{\mathrm{ER}}$.

Conversely, neutrons and WIMPs undergo elastic scattering with the nuclei of the argon atoms and are collectively called **nuclear recoils (NRs)**. Neutrons often only transfer very little of their energy in each elastic scattering. On the scale of the ArDM detector, multiply scattering neutrons are likely. Hence, since it is extremely unlikely for the so far undetected WIMP to scatter multiple times in the detector, the **scattering multiplicity** can be used to eliminate some of the neutrons among the NRs. In general, as neutron-induced NRs are not discriminable from WIMP-induced NRs, the rate of neutron events has to be minimized. Independent of the origin either being a WIMP or a neutron, the nuclear recoil energy E_{NR} in each scattering can not be transferred fully to the electrons of the target medium. In practice, a NR will result in a recoiling atomic nucleus moving through the medium while interacting with other nuclei forming a collision cascade. This will produce excited and ionized atoms, free electrons but also heat in the form of kinetic energy of the nuclei. Hence, one defines an energy-dependent nuclear scintillation or transferring efficiency $\mathcal{L}_{\mathrm{eff}}(E_{\mathrm{NR}})$ as $E_{\mathrm{ee,NR}} = \mathcal{L}_{\mathrm{eff}}(E_{\mathrm{NR}})\, E_{\mathrm{NR}}$, which will be further discussed in Section 2.2.1. Summarizing, the energy of both recoil types transferred to the electrons of the argon atoms are

$$E_{\mathrm{ee,ER}} = E_{\mathrm{ER}} \qquad \text{and} \qquad E_{\mathrm{ee,NR}} = \mathcal{L}_{\mathrm{eff}}(E_{\mathrm{NR}})\, E_{\mathrm{NR}}. \tag{2.1}$$

The main interest lies in the electron equivalent energy E_{ee} as this is the energy that is available in the form of excited atoms, ions and free electrons, which will in turn lead to the observable scintillation light and ionization charge. Already in 1961, Platzman described the energy distribution of E_{ee} [66] as

$$E_{\mathrm{ee}} = N_{\mathrm{i}} E_{\mathrm{i}} + N_{\mathrm{ex}} E_{\mathrm{ex}} + N_{\mathrm{i}} \varepsilon. \tag{2.2}$$

Therein the energy E_{ee} is distributed among N_{ex} number of excited atoms Ar^* which are excited by an average energy of E_{ex}, and N_{i} number of pairs of ionized atoms Ar^+ with average expenditure E_{i} and liberated electrons e^- of average kinetic energy ε. Dividing by N_{i} yields the so-called W value as

$$W = \frac{E_{\mathrm{ee}}}{N_{\mathrm{i}}} = E_{\mathrm{i}} + \frac{N_{\mathrm{ex}}}{N_{\mathrm{i}}} E_{\mathrm{ex}} + \varepsilon. \tag{2.3}$$

The W value corresponds to the average energy required for the production of a pair of a free electron and an ion and has been estimated by Doke et al. to be 23.6 ± 0.3 eV [67,68]. The ratio of excited and ionized atoms $N_{\mathrm{ex}}/N_{\mathrm{i}}$ is often denoted as α and Doke et al. reported α_{ER} to be in the range 0.19-0.29 for ERs in argon (measured with 1 MeV electrons specifically). E_{i} is reported at 15.4 eV, E_{ex} at 12.7 eV and ε at 5.26-6.0 eV [67]. Consequently, the

number of pairs of ionized atoms and free electrons can be written as

$$N_{\mathrm{i}} = \frac{E_{\mathrm{ee}}}{W} = \frac{\mathcal{L}_{\mathrm{eff}} E_{\mathrm{NR}}}{W} \qquad \text{(NRs)} \tag{2.4}$$

for NRs.

In the absence of an electric field, the ions Ar^+ and free electrons e^- will ideally all recombine and lead to further excitons Ar^*. The entirety of all excitons Ar^* will lead to excimers Ar_2^* which relax via the emission of scintillation light $\mathrm{h}\nu$. Phenomenologically, the excimer formation [69] can occur either through direct impact excitation via vibrationally excited states $\mathrm{Ar}_2^{*,\nu}$ as

$$\begin{aligned} &\mathrm{e}^- + \mathrm{Ar} \rightarrow \mathrm{Ar}^* + \mathrm{e}^- && \text{impact excitation (e.g. ERs)} \\ &\mathrm{Ar}^* + \mathrm{Ar} \rightarrow \mathrm{Ar}_2^{*,\nu} && \text{excimer formation} \\ &\mathrm{Ar}_2^{*,\nu} + \mathrm{Ar} \rightarrow \mathrm{Ar}_2^* + \mathrm{Ar} && \text{relaxation} \end{aligned} \tag{2.5}$$

or via recombination of the Ar_2^+ with the free e^- as

$$\begin{aligned} &\mathrm{e}^- + \mathrm{Ar} \rightarrow \mathrm{Ar}^+ + 2\mathrm{e}^- && \text{ionization} \\ &\mathrm{Ar}^+ + \mathrm{Ar} + \mathrm{Ar} \rightarrow \mathrm{Ar}_2^+ + \mathrm{Ar} && \\ &\mathrm{e}^- + \mathrm{Ar}_2^+ \rightarrow \mathrm{Ar}^{**} + \mathrm{Ar} && \text{recombination} \\ &\mathrm{Ar}^{**} + \mathrm{Ar} \rightarrow \mathrm{Ar}^* + \mathrm{Ar} + \text{heat} && \\ &\mathrm{Ar}^* + \mathrm{Ar} + \mathrm{Ar} \rightarrow \mathrm{Ar}_2^* + \mathrm{Ar} + \text{heat}. && \end{aligned} \tag{2.6}$$

The excimers then relax via the scintillation of vacuum ultraviolet (VUV) photons $h\nu$ as

$$\mathrm{Ar}_2^* \rightarrow \mathrm{Ar} + \mathrm{Ar} + \mathrm{h}\nu \qquad \text{VUV emission,} \tag{2.7}$$

which is called **prompt or primary scintillation (S1)** in the liquid phase. The scintillation wavelength of argon is 128 nm and argon is transparent to its own scintillation light. Henceforth, the scintillation light travels freely through the detector medium and can be observed directly, constituting the ArDM observable S1L. The excimers come in two different molecular states, a singlet and a triplet, with different lifetimes. For argon, the triplet state has a lifetime of 1600 ± 100 ns, whereas the singlet's lifetime is only 7.0 ± 1.0 ns [70]. The observed S1L signal hence is a combination of photons coming from decays of short-lived singlet and long-lived triplet states. In practice, it is observed that the corresponding ratios of triplet and singlet state excimers is drastically different for NRs and ERs [70]. The underlying reasons are not known exactly, but it is believed that the high concentration of the ionization, caused by NRs, boosts triplet-destructive interactions such as Penning ionization [71]. Hence, the pulse shape of the S1 signal can be used to discriminate the ERs and the NRs in argon. In practice the **prompt fraction** or **F90** is used as the pulse shape discriminator. The S1F90 corresponds to the fraction of the S1L observed in the first 90 ns and is an excellent indicator of the fraction of singlet states in the original recoil. In Fig. 2.1 (left) a plot of ArDM data taken with (bottom) and without (top) ^{252}Cf neutron source can be seen in the S1F90 versus S1L plane. The dominant ER branch at an S1F90 of $\sim$0.3 originating mostly from ^{39}Ar beta decays is clearly visible in both data sets. The

additional NR branch at an S1F90 of ∼0.7 becomes apparent in the ^{252}Cf neutron data set (bottom). The ability of pulse shape discrimination due to very different lifetimes is a main advantage of argon over xenon, as the lifetimes of the xenon triplet at 22.0 ± 2.0 ns and the singlet at 4.3 ± 0.6 ns are of similar order [70]. The pulse shape discrimination of the ArDM data will be studied in Chapter 5.

In the past single-phase argon detectors, operated with no electric drift field or gaseous phase, used the pulse shape discrimination of S1s alone to discriminate the recoil type. Ideally, in the absence of an electric field, the number of emitted VUV photons N_γ in the S1 would correspond to the sum of the created excitons N_{ex} (via Eq. 2.7) and free electron and ion pairs N_{i} (via Eqs. 2.5 to 2.7) as $N_\gamma \approx N_{\text{ex}} + N_{\text{i}}$. Consequently, one denotes the W_{ph} value as the average energy needed to produce a single scintillation photon at zero field and perfect scintillation efficiency. The value W_{ph} is also often referred to as the effective work function or the scintillation yield. Therefore, with W being the average energy required to produce a single pair of a free electron and an ion (see Eq. 2.3), W and W_{ph} are related via the conservation of energy eluding to

$$W \cdot N_{\text{i}} = W_{\text{ph}} \cdot (N_{\text{i}} + N_{\text{ex}}) \iff W_{\text{ph}} = \frac{W}{1 + \frac{N_{\text{ex}}}{N_{\text{i}}}} = \frac{W}{1+\alpha}. \tag{2.8}$$

In 2002, Doke et al. reported a value of 19.5 eV for W_{ph} in accordance with their previously mentioned W value of 23.6 ± 0.3 eV and a calculated value of $\alpha_{\text{ER}} = (N_{\text{ex}}/N_{\text{i}})_{\text{ER}} = 0.21$ (for 1 MeV electrons in argon) [68]. Unfortunately, W_{ph} is often abbreviated as W in modern literature thereby confusing the historical definition. In practice however, the assumption of perfect scintillation efficiency is not valid as some electrons always escape recombination and $N_\gamma < N_{\text{ex}} + N_{\text{i}}$ holds even in the absence of an electric drift field.

In a dual-phase detector, an electric field (of the order of 100-300 V/cm) is applied across the liquid phase in order to drift the freed electrons towards the gaseous phase. For an electric field of 230 V/cm and a temperature of 87 K (close to the boiling point of 87.303 K of argon) the estimated electron drift velocity amounts to 1.25 mm/μs [72]. During the drift, the free electrons can recombine, which has to be corrected for in the data analysis (see drift time correction in Section 4.1.2). The electrons which reach the boundary to the gaseous phase are extracted thereinto via an additional electric extraction field (on the order of ∼4 kV/cm) which is necessary to overcome the potential barrier between the liquid and the gaseous phase. In the gaseous phase the electrons induce the same excitation-and-scintillation processes (as in Eqs. 2.5 to 2.7) which result in the **delayed or secondary scintillation (S2)**, with the S2L being the corresponding light observed by ArDM. Consequently, the S2L is expected to be proportional to the free electrons drifting upwards and hence the ionization charge. Naturally, a higher electric drift field strength will increase the number of electrons which escape recombination during the drift and hence also increases the S2L. However, it will consequently also decrease the S1L production through recombination via Eq. 2.6 (the recombination quenching of the S1 due to the electric drift field will be discussed in Section 2.2.1). The strength of the extraction field can also be used to amplify the S2 signals as more free electrons can overcome the potential barrier and higher energy electrons lead to more excitations and ionizations via ERs in the gaseous phase. Accordingly, the average ratio of the ionization charge variable S2L and the prompt scintillation variable S1L can to some extent be controlled with the applied drift and extraction field strengths.

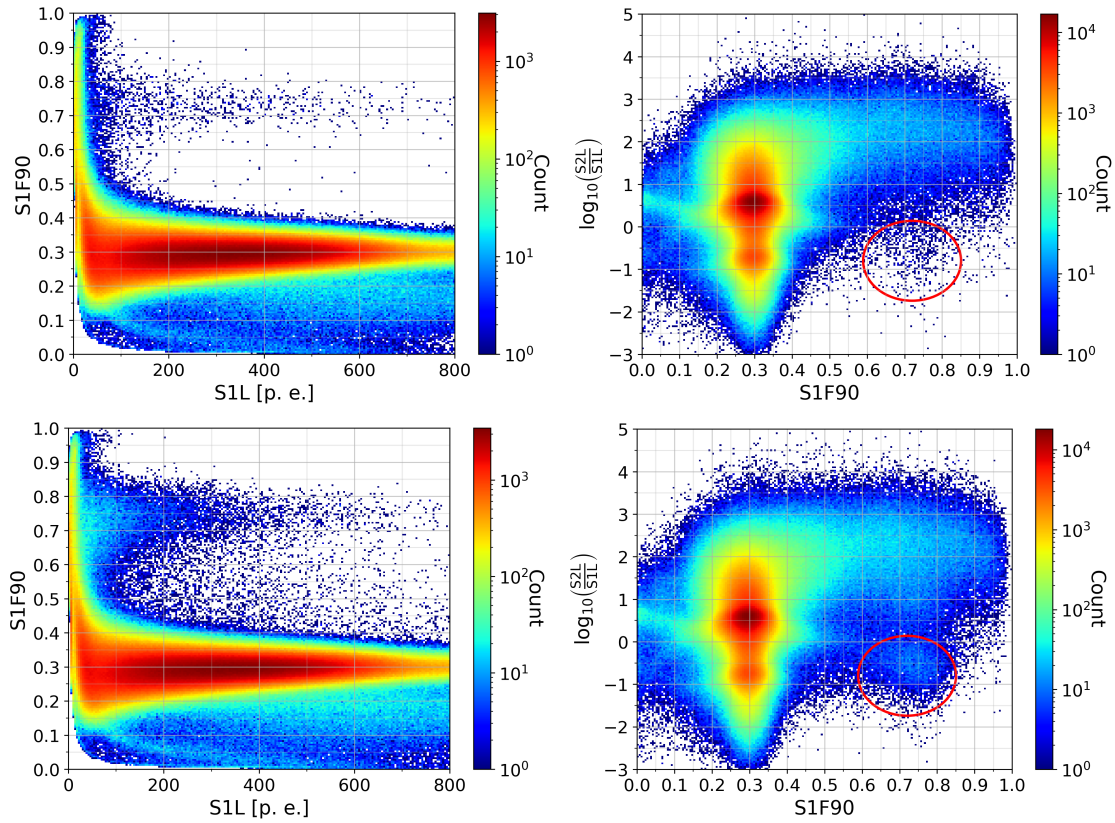

Figure 2.1: *Left*: ArDM data in the plane of the pulse shape discriminator S1F90 versus the prompt scintillation light S1L. *Right*: ArDM data in the plane of the charge to light ratio log(S2L/S1L) versus the pulse shape discriminator S1F90. The data sets are shown with (*bottom*) and without (*top*) a ^{252}Cf neutron source present. The dominant ERs originating mostly from ^{39}Ar beta decays are centered around an S1F90 of ∼0.3 and a log(S2L/S1L) of ∼0.7. The NR region becomes apparent in the Neutron data sets at an S1F90 of ∼0.7 and a log(S2L/S1L) of ∼-0.8 (red circle). See text for details.

The **charge-to-light-ratio S2L/S1L** is also instrumental in the background rejection of ERs versus NRs. As discussed, in electronic recoils the recoil energy is fully transferred to the electrons of the argon atoms, resulting in a high number of free electrons and hence a large S2. Conversely, in nuclear recoils the gained kinetic energy of the recoiling argon atoms after elastic scattering is of the order of 1-100 keV, which allows the electrons of the argon atoms to keep up with the nuclei. Therefore, the argon atoms conserve most of their electrons and transfer their excess energy into other nuclei in collision cascades, essentially creating heat. While there exists no all-encompassing model of the energy transfer mechanisms involved in NRs, it is clear that fewer free electrons are emitted in NRs than in ERs of comparable recoil energy. Consequently, one expects the S2L/S1L to be smaller for NRs than for ERs. While this is known to be an excellent recoil type discriminator for xenon detectors[3] [74], it has also been shown to add discrimination power to argon experiments [75, 76]. Therefore, xenon-based detectors use the S2L/S1L ratio as their main

[3]Also owed to the fact that $\alpha_{\mathrm{ER,Xe}} = (N_{\mathrm{ex}}/N_{\mathrm{i}})_{\mathrm{ER,Xe}} \approx 0.06\text{-}0.20$ [68] and $\alpha_{\mathrm{NR,Xe}} = (N_{\mathrm{ex}}/N_{\mathrm{i}})_{\mathrm{NR,Xe}} \approx 1$ [69, 73] are vastly different in liquid xenon.

discriminator of ERs versus NRs, while argon-based experiments can use the pulse shape discriminator S1F90 as well as the S2L/S1L ratio. The distribution of the log(S2L/S1L) versus the S1F90 of ArDM data can be seen in Fig. 2.1 (right) with (bottom) and without (top) a ^{252}Cf neutron source present. The dominant ER events originating mostly from ^{39}Ar beta decays are focused around an S1F90 of ∼0.3 and a log(S2L/S1L) of ∼0.7. The accumulation of NR events becomes visible at an S1F90 of ∼0.7 and a log(S2L/S1L) of ∼-0.8 (red circle) in the Neutron data. The two-dimensional recoil type discrimination will be discussed in detail in Chapter 5.

2.1.2 Detector overview

This section provides a technical overview of the ArDM detector. While the author has not been present himself during the construction and upgrading of the ArDM detector, a detailed review of the earlier ArDM setup can be found in Bárbara-Rosario Montes Núñez' PhD thesis submitted in 2016 [77] and a summary of the dual-phase upgrade of ArDM can be found in Wei Mu's PhD thesis submitted in 2019 [78]. The summary presented here is heavily based on these PhD theses as well as the submitted ArDM status reports.

The main challenges in a modern competitive liquid noble gas dual phase dark matter detector lie in the scarcity of potential signal events, the abundance of background events, the necessary experimental precision and the required longevity of the high quality measurement conditions. At the current level of sensitivity to the WIMP-nucleon cross section σ_p^{SI} above the neutrino floor, the non-discriminable background in a WIMP search consists solely of low energy single-scattering neutrons. The most common neutron sources are the decay chains of the radioactive ^{232}Th and ^{238}U isotopes, of which trace amounts are naturally present in common rock as well as detector components. The second external source of neutrons is cosmic muons that scatter inelastically with nuclei or produce neutrons in secondary reactions in electromagnetic or hadronic cascades. Therefore, liquid noble gas WIMP searches are carried out underground as the muon flux decreases significantly with depth in the Earth's crust (which is usually measured in *m.w.e.* or "meters water equivalent"). ArDM is located at the Laboratorio Subterráneo de Canfranc (LSC) in Canfranc, Spain. The depth of the LSC experimental halls corresponds to 2450 m.w.e. and the muon flux therein has been measured to be (2-4)$\cdot 10^{-3}$ m^{-2}s^{-1} and the neutron flux to be $3.47 \pm 0.35 \cdot 10^{-2}$ m^{-2}s^{-1} [79]. In addition, the whole detector setup is surrounded by a high-density polyethylene shield to passively suppress external neutrons. It is arranged in an equilateral octagon around the main liquid argon vessel with a total height of 505 cm and being at least 50 cm thick in each dimension around the detector. While external neutron sources can be shielded to a very high degree, the internally emitted neutrons through the natural radioactivity of the detector components constitute the instrumental background and WIMPs can only be measured as a significant fluctuation thereof.

The main liquid argon (LAr) vessel is a cylindrical stainless steel dewar of 100 cm in diameter and 200 cm in height. It hosts a 80 cm diameter sensitive LAr volume with a maximal drift length of 112 cm, a gaseous argon (GAr) layer on top, an inner polyethylene shield, the electric field cages as well as the light detection system. It can hold approximately 850 kg of LAr. In order to maintain the quality of the LAr in the target volume, it has to be continuously cooled and purified. A separate 15 mm thick cooling bath of LAr surrounds the internal walls of the main dewar to stabilize the temperature of the LAr target volume slightly below its boiling point of 87 K (the density of argon at its boiling

point of 87.303 K is 1.396 g/cm^3 [4]). The cooling system consists of three 300-W Gifford-MacMahon cryocoolers that re-condense boiled off argon gas in the separate LAr cooling bath. A high purity of the LAr is of the utmost importance, as impurities can significantly quench the production of scintillation light from excimers (Eq. 2.7). It has been shown that for example a contamination with N_2 or O_2 of 1 ppm reduces the scintillation light by ∼20% [80, 81] or ∼38% [82] respectively. Therefore, the LAr is continuously circulated through purification cartridges filled with copper powder to filter electro-negative impurities. The LAr is circulated via a cryogenic membrane pump which has a flow rate at the order of 170 L/h. Similarly, ArDM also hosts a separate purification cycle for the GAr layer on top of the LAr. During cryogenic operation the pumps, pressure, temperature, oxygen and liquid argon levels are remotely monitored.

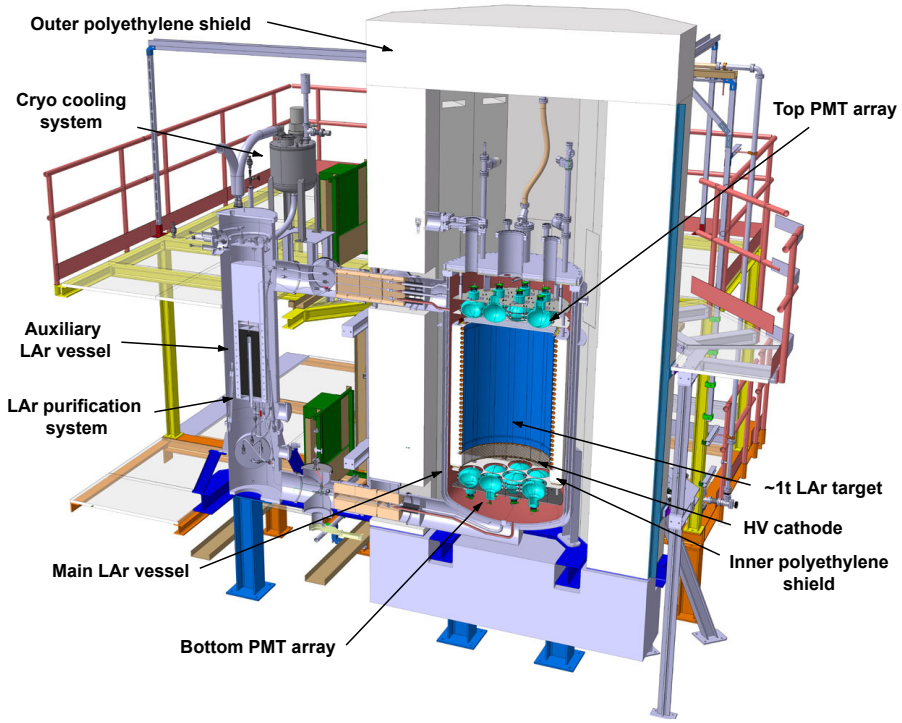

Figure 2.2: ArDM detector overview with the main components. See the text for a detailed description. Courtesy of the ArDM collaboration.

The innermost part of the detector, the active LAr target inside the inner polyethylene shield, contains the electric field cages used to extract the ionization charge, as well as the optics to observe the resulting scintillation lights S1L and S2L. The drift field cage is of cylindric form of height 1112 mm and inner diameter 735 mm (see Fig. 2.3, left). It is immersed inside the LAr and GAr volume and encloses a subvolume of 660 kg of LAr (corresponding to the maximal possible active target). With an electron drift velocity of 1.25 mm/μs (for a drift field strength of F_{Drift} = 230 V/cm and a temperature of 87 K (close to the boiling point of 87.303 K of argon) [72]) and an ^{39}Ar activity of ∼1 Bq/kg of LAr [83], the maximal drift length avoids significant pile-up of S1 and S2 events. The top and bottom of the drift field cage are composed of polymethylmethacrylate (PMMA) plates

[4] See `https://lar.bnl.gov/properties/`

which are covered with indium tin oxide (ITO) layers on both sides. ITO is a conductive material which is also transparent to visible light. Therefore, the top and bottom PMMA plates act as the anode and the cathode respectively. The wall of the drift field cage contains 28 field-shaper rings made of 2.4 mm diameter copper tubes which act as voltage dividers to guide the electric drift field. The central volume is completed by two 10 mm copper plates on the top and bottom to hold the polyethylene shield and the copper rings in place. During operation, a high voltage of the order of -30 kV is applied to the cathode ITO layer resulting in the desired electric field strength of $\sim$240 V/cm within the liquid phase. Increasing voltages from bottom to top are applied to the 28 field-shapers to guide the electric field uniformly. The extraction grid is immersed in the LAr and located below the anode ITO layer which is inside the GAr. There-between an electric field of the order of 4 kV/cm is applied to aid the free electrons to overcome the potential barrier at the LAr-GAr boundary (which is adjusted to be in the middle of the extraction grid and the anode). Finally, the electronics at the bottom of the vessel are protected by a protection grid with an applied voltage of ~ -0.5 kV.

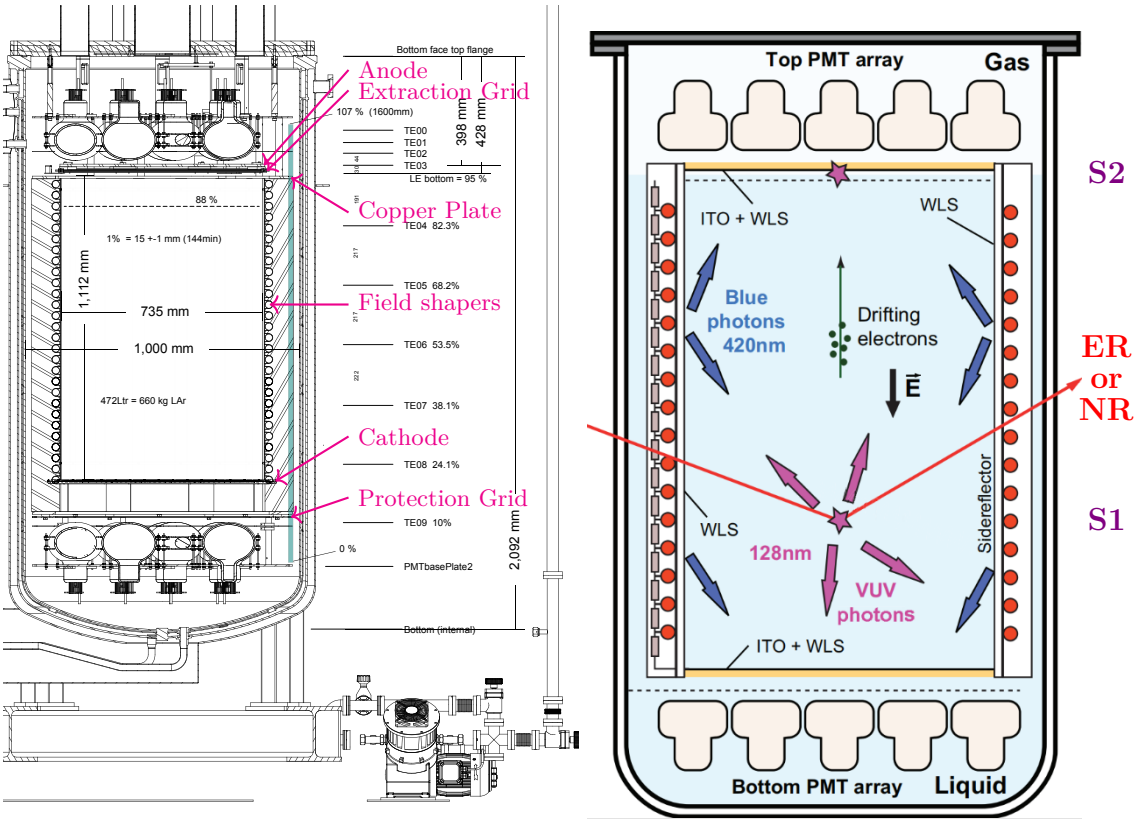

Figure 2.3: *Left*: Technical cross section of the ArDM main vessel containing the drift field cage as well as the extraction grid and the PMTs to detect the scintillation light from S1 and S2 signals. *Right:* Schematic of the ArDM detector main vessel in operation. An incoming particle recoils in the LAr volume leading to the prompt scintillation S1 and free electrons. These are drifting upwards in the applied electric field towards the gaseous layer, wherein they are extracted via the applied extraction field. Upon entering the gaseous phase, the electrons induce further recoils leading to the delayed scintillation light S2L. Both scintillation signals are wavelength-shifted and reflected along all boundaries of the drift cage until they are collected by the top and bottom arrays of PMTs. Courtesy of the ArDM collaboration.

The light detection system of ArDM consists of reflectors, wavelength shifters and 24 photomultiplier tubes (PMTs). The 24 8" cryogenic Hamamatsu borosilicate PMTs are divided into two symmetric arrays of twelve PMTs (see Fig. 2.4, left), one at the top and one at the bottom of the drift field cage (see Fig. 2.3). They are operated at bias voltages of 946-1345 V avoiding electric discharges and resulting in a nominal gain of $5{\cdot}10^7$ [78]. As their optimal light detection wavelength is in the range of 360-436 nm, the PMT glass windows are covered with tetraphenyl butadiene (TPB). TPB is an organic wavelength shifter that converts the VUV scintillation signals (128 nm) into blue light (∼420 nm). The lateral areas of the drift cage are covered with high-reflectivity foils made of 3M Vikuiti^{TM} Enhanced Specular Reflector Film to maximize the solid angle coverage of the light detection system. The reflecting foils in the drift cage are also covered with TPB to again act as wavelength shifters.

Figure 2.4: *Left*: The (bottom) PMT array used inside ArDM to detect the scintillation signals S1 and S2. *Right:* The whole ArDM detector inside the neutron polyethylene shield located at the LSC in Canfranc, Spain. Courtesy of the ArDM collaboration.

The data acquisition (DAQ) system of ArDM consists of four 8-channel, 250 MHz, 12-bit FADC (flash analog-to-digital converter) digitizer modules (CAEN V1720). The 24 PMTs are divided among the four boards where six channels on each board handle six PMTs. The signal detection can be triggered by any of the four boards. Each is connected to a DAQ PC via an optical link with a data rate of up to 80 MB/s (∼320 MB/s for all four optical links combined). The sampling time window during operation was set to either 4 or 8 μs in order to detect single S1 and S2 events separately. Events are recorded with a 250 MHz sampling rate resulting in 1024 or 2048 samples per event and PMT channel. The recorded partial events of the four boards are stored separately, together with an absolute time stamp (the trigger time tag or TTT) which is reset and synchronized at every restart of the DAQ. The full event is then reconstructed offline from the four partial events based on the TTT. This system can handle the ∼2 kHz of S1 and S2 events which are expected in one tonne of atmospheric argon with an ^{39}Ar activity of ∼1 Bq/kg. The first step in the data reconstruction, a software called LRAnalysis, calculates the pedestal of triggers, calibrates the single-photon-response and evaluates the light signals of each of the 24 PMT channels. All relevant reconstructed quantities are stored in numbered run-based ROOT files containing ∼100k triggers. During operation, a MySQL database is continuously updated with the electronics settings of the field cage as well as the DAQ via a slow control system. This database also bookkeeps the general trigger rate and which events

have been triggered by how many of the four boards. This information can be used in a first selection of quality runs which have to be fully reconstructed. The run selection criteria and the reconstruction steps will be discussed in Chapter 3 while the data processing steps and necessary data corrections will be covered in detail in Chapter 4.

2.2 Theory of WIMPs inside ArDM

This section serves as a continuation of the WIMP theory discussed in Section 1.3, where the theoretical WIMP spectrum on Earth has been derived as a function of the nuclear recoil energy E_{NR}. As discussed in Section 2.1.1, the actual ArDM observable is the S1L, the number of detected photoelectrons of the primary scintillation, which is expected to be proportional to the energy E_{ee} transferred to the electrons of the argon atoms in the recoils. The full conversion from the nuclear recoil energy E_{NR} to the primary scintillation light S1L incorporates multiple effects as

$$\mathrm{S1L}\left(F_{\mathrm{Drift}}\right) = \overbrace{\mathrm{LY}\left(F_{\mathrm{Drift}} = 0\ \mathrm{V/cm}\right)}^{\text{Light Yield}} \cdot \overbrace{\underbrace{\mathcal{L}_{\mathrm{eff}}\left(E_{\mathrm{NR}}\right)}_{\substack{\text{Nuclear and}\\ \text{bi-excitonic}\\ \text{quenching}}} \cdot E_{\mathrm{NR}}}^{E_{\mathrm{ee}}} \cdot \overbrace{q_{\mathrm{rec}}\left(E_{\mathrm{NR}}, F_{\mathrm{Drift}}\right)}^{\substack{\text{Recombination}\\ \text{quenching}}}. \tag{2.9}$$

The nuclear scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$ from the nuclear recoil energy E_{NR} to the electron equivalent energy E_{ee}, as well as the quenching of recombination due to the presence of the electric drift field will be treated in Section 2.2.1 in the context of the underlying theory. The light yield LY representing the conversion from keV of quenched recoil energy to the detected photoelectrons constituting the observed S1L of a signal event is incorporated within Section 2.2.2. Therein, the resulting total WIMP rate above a detection threshold of S1L will be presented and converted into the projected sensitivity of ArDM to the spin-independent WIMP-proton interaction cross section σ_p^{SI} versus the WIMP mass m_χ for reasonable experimental conditions.

2.2.1 Nuclear scintillation efficiency and drift field dependence

In contrast to ERs, the recoil energy of NRs is not fully transferred to the electrons of the atoms in the target medium (see discussion in Section 2.1.1). In practice, a nuclear recoil will result in a recoiling atomic nucleus moving through the medium while interacting with other nuclei forming a collision cascade. This will produce excited and ionized atoms, free electrons but also heat in the form of kinetic energy of the involved nuclei. The energy converted to heat is inadvertently lost for the production of scintillation light. This is accounted for via the nuclear scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$ from nuclear recoil energy E_{NR} to the electron equivalent energy E_{ee} (as in Eq. 2.1).

Generally, the nuclear scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$ accounts for two types of losses of electron equivalent energy as

$$\mathcal{L}_{\mathrm{eff}}\left(E_{\mathrm{NR}}\right) = q_{\mathrm{nc}}\left(E_{\mathrm{NR}}\right) \cdot q_{\mathrm{bi}}\left(E_{\mathrm{NR}}\right) \tag{2.10}$$

The first term, the nuclear quenching q_{nc} summarizes the previously mentioned energy lost in collision cascades in the form of heat. The underlying theory on atomic collisions was

published by Lindhard in 1963 [84] and equates the nuclear quenching to

$$q_{\mathrm{nc}}\left(E_{\mathrm{NR}}\right) = \frac{k \cdot g\left(\varepsilon\right)}{1 + k \cdot g\left(\varepsilon\right)} \quad \text{with} \quad k = 0.133 \cdot Z^{\frac{2}{3}} \cdot A^{-\frac{1}{2}}. \tag{2.11}$$

In the Lindhard theory [76,84] of nuclear quenching, k is the proportional constant between the electronic stopping power $\left(\mathrm{d}E/\mathrm{d}x\right)_{\mathrm{el}}$ and the velocity of the recoiling atom. It depends on the atomic weight A and number Z (where $A^{\mathrm{Ar}} = 40$ and $Z^{\mathrm{Ar}} = 18$). In 2008, in what is often called the Mei model, Mei et al. have reiterated the Lindhard theory and put it on a modern footing [85]. They extracted the function $g\left(\varepsilon\right)$ by a fit to data and found it to be well described by

$$g\left(\varepsilon\right) = 3 \cdot \varepsilon^{0.15} + 0.7 \cdot \varepsilon^{0.6} + \varepsilon, \qquad \text{where} \qquad \varepsilon\left(E_{\mathrm{NR}}\right) = 11.5 \cdot Z^{-\frac{7}{3}} \cdot \left(\frac{E_{\mathrm{NR}}}{\mathrm{keV}}\right) \tag{2.12}$$

relates to the nuclear recoil energy E_{NR}. The second type of quenching is called bi-excitonic quenching q_{bi}. Therein two excited atoms (which would each lead to scintillation light via Eqs. 2.5 to 2.7) collide with each other via

$$\mathrm{Ar}^{*} + \mathrm{Ar}^{*} \quad \rightarrow \quad \mathrm{Ar} + \mathrm{Ar}^{+} + \mathrm{e}^{-}, \tag{2.13}$$

effectively reducing the number of excited atoms (or excitons) which could lead to scintillation light. Additionally, the Penning process can further reduce the number of excited atoms in bi-excitonic collisions in a high excitation density environment [71]. While it is known that the nuclear quenching q_{nc} alone can not explain the observed nuclear scintillation efficiency $\mathcal{L}_{\mathrm{eff}}$, the theory of the bi-excitonic quenching q_{bi} is not fully established yet. Mei et al. proposed to use Birk's saturation law, which is commonly done in the community.

$$q_{\mathrm{bi}}\left(E_{\mathrm{NR}}\right) \overset{\text{Mei}}{=} \frac{1}{1 + k_{\mathrm{B}} \cdot \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\mathrm{el}}} \quad \text{or} \quad q_{\mathrm{bi}}\left(E_{\mathrm{NR}}\right) \overset{\text{Hitachi}}{=} 0.60 = \text{constant}. \tag{2.14}$$

Here, $\left(\mathrm{d}E/\mathrm{d}x\right)_{\mathrm{el}}$ is the electronic stopping power of an argon atom inside argon which is dependent on the recoiling atom's kinetic energy and hence on E_{NR}. For the free parameter k_B values in the range of $(3.12\text{-}7.4) \cdot 10^{-4}$ $\mathrm{MeV}^{-1}\mathrm{gcm}^{-2}$ can be found [76, 85–87]. In 2018, the ARIS (Argon Response to Ionization and Scintillation) experiment proposed to include a quadratic term in the denominator [88]. In 2019, Hitachi et al. argued that a correct theory of bi-excitonic quenching should include the linear energy transfer instead of the stopping power, and used an approximation to evaluate a constant value for q_{bi} showing good agreement with the available argon data [89]. While the theoretical treatment of bi-excitonic quenching still seems to be an open debate, the constant value $q_{\mathrm{bi}} = 0.60$ from Hitachi et al. will be used for the purposes of this WIMP search due to its good agreement with the available argon data as shown by Hitachi et al.. The resulting nuclear scintillation efficiency $\mathcal{L}_{\mathrm{eff}}\left(E_{\mathrm{NR}}\right)$ combining the nuclear quenching from the Lindhard theory and Hitachi's constant bi-excitonic quenching can be seen in Fig. 2.5 (right).

The final type of S1 quenching is the recombination or drift field quenching q_{rec}. At vanishing electric fields, all free electrons remain in the original recoil area and ideally contribute to the S1 via the recombination channel (Eqs. 2.6 and 2.7). In a dual-phase detector, the applied electric field forces the free electrons to drift away from the original

recoil center towards the gaseous layer. Hence, a stronger drift field will amplify the S2 while quenching the S1 as less free electrons are available for the recombination channel production. Formally, the recombination quenching of the S1 is evaluated in the Thomas-Imel box model [90] and modified for argon [76, 87] as

$$q_{\mathrm{rec}}\left(E_{\mathrm{NR}}, F_{\mathrm{Drift}}\right) = \frac{\alpha_{\mathrm{NR}}\left(F_{\mathrm{Drift}}\right) + R^{*}\left(E_{\mathrm{NR}}, F_{\mathrm{Drift}}\right)}{\alpha_{\mathrm{NR}}\left(F_{\mathrm{Drift}}\right) + 1}, \qquad \text{where} \qquad \alpha_{\mathrm{NR}} = \left(\frac{N_{\mathrm{ex}}}{N_{\mathrm{i}}}\right)_{\mathrm{NR}} \tag{2.15}$$

is the ratio of N_{ex} excited argon atoms and N_{i} produced pairs of ionized atoms and free electrons (in argon), which was already introduced in Section 2.1.1. The value of α_{NR} for NRs in argon is not exactly known. While DarkSide-50 [86] and ARIS [88] have assumed α_{NR} to be constant and equal to 1 without scientific evidence, the SCENE collaboration has reported values increasing from $\alpha_{\mathrm{NR}} = 0.6 \pm 0.4$ for $E_{\mathrm{NR}} = 16.9$ keV up to $\alpha_{\mathrm{NR}} = 2.1 \pm 1.2$ for $E_{\mathrm{NR}} = 57.3$ keV at drift field strengths below 1 kV/cm [91]. In 2019, Kimura et al. [87] have introduced a model of α_{NR} which includes its dependence on the strength of the electric drift field F_{Drift} which follows an exponential as

$$\alpha_{\mathrm{NR}} = \alpha_0 \cdot \exp\left(-D_{\alpha} F_{\mathrm{Drift}}\right), \tag{2.16}$$

where α_0 is fixed to 1. Kimura et al. reported a best fit value of $D_{\alpha} = (8.9 \pm 0.5) \cdot 10^{-4}$ cm/V [87]. The model of Kimura et al. is based on a similar model as used for xenon [92] where scientific evidence exists for $\alpha_{\mathrm{NR,Xe}} \approx 1$ [69, 73]. Kimura's model is adopted for this WIMP search as there currently exists no perfect model which would also explain the dependence on E_{NR} as observed by SCENE [91]. The recombination probability R^{*} is given by

$$R^{*} = 1 - \frac{\ln\left(1 + N_i \zeta\right)}{N_i \zeta}, \quad \text{where} \quad \zeta = \gamma \cdot \left(F_{\mathrm{Drift}}\right)^{-\delta} \tag{2.17}$$

follows a power law. The best fit values of Kimura et al. to ^{252}Cf neutron data are $\gamma = 1.15 \pm 0.02$ $(\mathrm{V/cm})^{\delta}$ and $\delta = (5.76 \pm 0.03) \cdot 10^{-1}$ [87], which are used for this WIMP search. Finally, the number of produced pairs N_i of ionized atoms and free electrons can be retraced from Eq. 2.4 (Section 2.1.1) as

$$N_i = \frac{E_{\mathrm{NR}}}{W_{\mathrm{ph}}} \cdot \frac{1}{\alpha_{\mathrm{NR}} + 1} \cdot \mathcal{L}_{\mathrm{eff}}, \tag{2.18}$$

where W_{ph} denotes the scintillation yield or effective work function introduced in Eq. 2.8. The value of Doke et al. of $W_{\mathrm{ph}} = 19.5$ eV [68] is assumed to be recoil-type-independent (as done by the community, e.g. [76, 86–88, 91]) and adopted for the purposes of this WIMP search. The resulting drift field strength dependence of the recombination quenching q_{rec} can be seen in Fig. 2.5 (left) for various fixed nuclear recoil energies E_{NR}. The total combined S1 quenching, including the nuclear and bi-excitonic quenching, can be seen in Fig. 2.5 (right) for the drift field strength value $F_{\mathrm{Drift}} = 230$ V/cm as used during the WIMP search in ArDM. The combined quenching effect reduces the overall S1 to lower than 30% in the energy region of interest for WIMP searches. The S1 reduction due to the quenching effects is strongest at the lowest energies which is also where most WIMPs are expected to reside (see Fig. 1.10, Section 1.3.4) constituting a further physical difficulty of WIMP searches. The recombination quenching due to the drift field can be controlled by the experimental setup and does not inherently pose a problem as lost S1 scintillation light is transformed to S2 scintillation light.

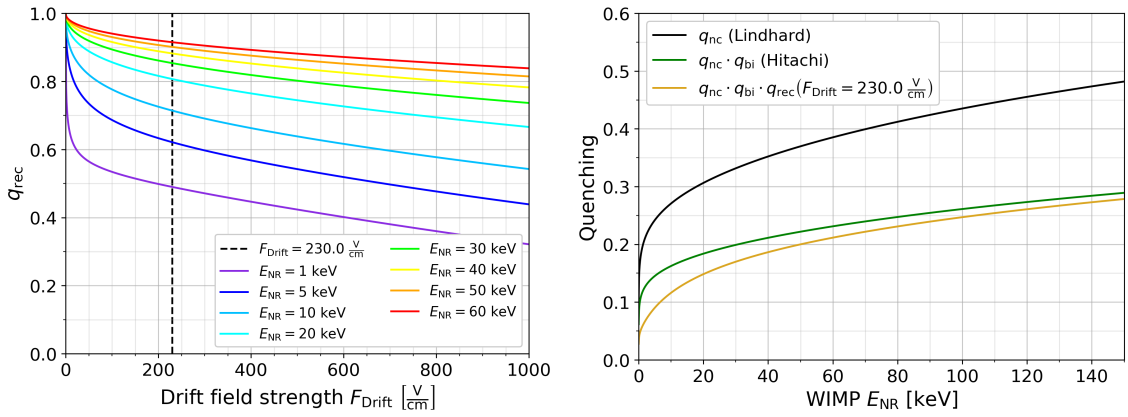

Figure 2.5: *Left*: The recombination quenching q_{rec} of the S1 signals due to the presence of an electric drift field for various nuclear recoil energies E_{NR} (see Eqs. 2.15 to 2.18). The drift field strength in ArDM was set to $F_{\text{Drift}} = 230$ V/cm during the WIMP search. *Right*: Nuclear (q_{nc}, black, Eqs. 2.11 and 2.12) and bi-excitonic quenching (q_{bi}, green, Eq. 2.14) combined for argon as a function of the nuclear recoil energy E_{NR}. The combination with the drift-field-dependent recombination quenching (q_{rec}, gold, Eqs. 2.15 to 2.18) reduces the overall S1 signal to less than 30% of its original value in the energy region of interest for WIMP searches. Further discussion and details can be found in the text.

2.2.2 Projected WIMP sensitivity inside ArDM

Before an actual WIMP search, the projected sensitivity to the WIMPs is commonly estimated. It is based on reasonable assumptions of the experimental conditions, as well as the theoretical concepts and the analysis. In order to do that for ArDM, the WIMP spectrum on Earth as a function of the recoil energy E_{NR} (see Fig. 1.10, Section 1.3.4) needs to be converted into a WIMP spectrum inside ArDM as a function of the observable S1L. Reiterating the full conversion from Eq. 2.9 as

$$\text{S1L} = \text{LY} \cdot \mathcal{L}_{\text{eff}} \cdot E_{\text{NR}} \cdot q_{\text{rec}}, \tag{2.19}$$

where the nuclear scintillation efficiency $\mathcal{L}_{\text{eff}}$ and the drift-field-dependent recombination quenching q_{rec} have been discussed in the previous Section 2.2.1 and visualized in Fig. 2.5. The remaining factor, the light yield LY, accounts for the overall conversion efficiency from the quenched recoil energy available to produce scintillation light to the actual detected and reconstructed photoelectrons. As such it incorporates the influence of the purity level and dimensions of the target medium, the detection efficiency and resolution of the PMTs, the data acquisition, as well as the reconstruction of the detected photoelectrons. It is commonly presented in units of photoelectrons per keV of electron equivalent energy (after all quenching effects) and is strictly measured at zero drift field, as the field dependence is already incorporated in the recombination quenching q_{rec}. As the dominant background of ArDM is the naturally abundant radioactive isotope ^{39}Ar, its well defined energy-dependent β spectrum is the natural candidate to gauge the LY of ArDM. The results of the fit of the theoretical ^{39}Ar β spectrum are detailed in Section 5.1 and the good agreement shows that the LY_{ER} is independent of the ER recoil energy E_{ER} up to 565 keV (the corresponding Q value or maximal possible β kinetic energy). For the experimental phase of the WIMP search the ArDM LY_{ER} is estimated to be 1.33 p.e./keV (see Section 5.1). An essential

assumption is then, that the conversion from the quenched recoil energy to photoelectrons is independent of the recoil type and correspondingly $\mathrm{LY} = \mathrm{LY_{ER}} = \mathrm{LY_{NR}}$. This completes the discussion of all variables included in the conversion from the WIMP nuclear recoil energy E_{NR} to the ArDM observable S1L as in Eq. 2.19. One can then apply the chain rule of derivatives as

$$\frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}} = \frac{\mathrm{d}R}{\mathrm{dS1L}}\frac{\mathrm{dS1L}}{\mathrm{d}E_{\mathrm{NR}}} \quad \rightarrow \quad \frac{\mathrm{d}R}{\mathrm{dS1L}} = \frac{\mathrm{d}R}{\mathrm{d}E_{\mathrm{NR}}}\left(\frac{\mathrm{dS1L}}{\mathrm{d}E_{\mathrm{NR}}}\right)^{-1} \tag{2.20}$$

to convert the WIMP spectrum on Earth as a function of E_{NR} (Eq. 1.42, Section 1.3.4) to a WIMP spectrum inside ArDM as a function of S1L. The required derivative of S1L with respect to E_{NR}

$$\frac{\mathrm{dS1L}}{\mathrm{d}E_{\mathrm{NR}}} = \mathrm{LY}\cdot\mathcal{L}_{\mathrm{eff}}\cdot E_{\mathrm{NR}}\cdot q_{\mathrm{rec}}\cdot\left(\frac{1}{\mathcal{L}_{\mathrm{eff}}}\frac{\mathrm{d}\mathcal{L}_{\mathrm{eff}}}{\mathrm{d}E_{\mathrm{NR}}} + \frac{1}{E_{\mathrm{NR}}} + \frac{1}{q_{\mathrm{rec}}}\frac{\mathrm{d}q_{\mathrm{rec}}}{\mathrm{d}E_{\mathrm{NR}}}\right) \tag{2.21}$$

is quite involved but straightforward and not quoted here. The resulting differential WIMP spectrum as a function of the ArDM observable S1L for the same chosen parameters $m_\chi = 100$ GeV, $\sigma_p^{\mathrm{SI}} = 10^{-45}$ cm^2 and an average date of July, 1st, as well as the light yield $\mathrm{LY} = 1.33$ p.e./keV and the drift field strength $F_{\mathrm{Drift}} = 230$ V/cm can be seen in Fig. 2.6 (left). While the general features of the WIMP spectrum have not changed drastically due to the conversion from E_{NR} to S1L, the sharp fall-off in the spectrum transitioned from ∼242 keV of E_{NR} to ∼103 p.e. of S1L and the amplitude increased slightly due to the smaller energy scale.

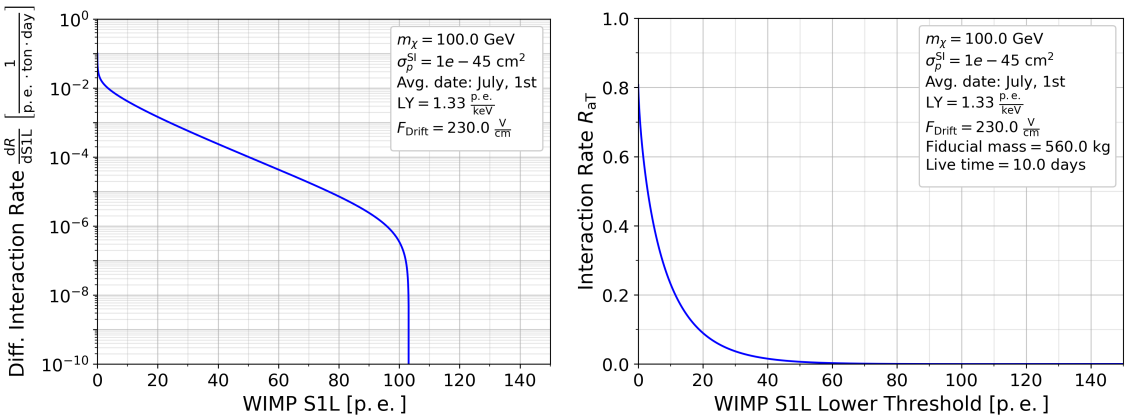

Figure 2.6: *Left*: The differential WIMP interaction rate (Eqs. 2.20 and 2.21) as a function of the ArDM observable S1L for a chosen set of the WIMP mass m_χ, the spin-independent WIMP-nucleon cross section σ_p^{SI} and a presumed set of data-taking conditions in the ArDM WIMP search. *Right*: The total expected WIMP rate inside the assumed ArDM fiducial volume during a measurement period of 10 live days. The total expected WIMP rate $R_{\mathrm{aT}}\,(0)$ is 0.807 (8.07) events for $m_\chi = 100$ GeV at $\sigma_p^{\mathrm{SI}} = 10^{-45}$ (10^{-44}) cm^2.

The total WIMP rate R_{aT} above a threshold T of S1L can then be gathered via numerical integration as

$$R_{\mathrm{aT}}\,(\mathrm{T}) = \int_{\mathrm{T}}^{\infty}\frac{\mathrm{d}R}{\mathrm{dS1L}}\mathrm{dS1L} = \int_{\mathrm{Threshold}}^{\mathrm{Fall-off}}\frac{\mathrm{d}R}{\mathrm{dS1L}}\mathrm{dS1L}. \tag{2.22}$$

The total WIMP rate above a threshold of S1L is the same integration as the total WIMP rate above a threshold of E_{NR} (Eq. 1.43, Section 1.3.4) with a substitution according to the conversion in Eq. 2.19. Hence, for practical purposes, the total WIMP rate above a threshold can be calculated with the less complex WIMP spectrum as a function of E_{NR} (Eq. 1.43) as long as the threshold is set according to the corresponding S1L threshold (Eq. 2.19). Consequently, the total expected WIMP rate $R_{\mathrm{aT}}(0)$ for the same set of parameters is still 0.144 events per day and tonne of argon. Finally, the target mass and the experimental live time of the ArDM WIMP search need to be incorporated. In practice, there is always a possibility of boundary effects and accumulated background events at the edges of the detector vessel. Therefore, rare event search experiments commonly "fiducialize" or reduce the detector volume to an inner fiducial volume, which constitutes the inherently smaller target volume. While the exact size of the fiducial volume and the total live time of the quality data can only be fixed after the data analysis, a presumable fiducial volume of 560 kg (∼92% of the 660 kg drift field cage volume) and an experimental live time of 10 days follow the expectations. The total WIMP rate R_{aT} above a given S1L threshold can be seen in Fig. 2.6 (right) for this set of ArDM experimental parameters. As the WIMP rate scales linearly with the target size (in tonnes) and the experimental live time (in days), the expected total WIMP rate $R_{\mathrm{aT}}(0)$ in ArDM is 0.807 events for $m_\chi = 100$ GeV at $\sigma_p^{\mathrm{SI}} = 10^{-45}$ cm^2, and 0.232 (0.090) WIMP events are expected above a S1L threshold of 20 (10) photoelectrons. As the WIMP rate also scales linearly with σ_p^{SI} (see Eq. 1.42, Section 1.3.4) the total expected WIMP rate $R_{\mathrm{aT}}(0)$ inside ArDM would be 8.07 (80.7) events for $m_\chi = 100$ GeV at $\sigma_p^{\mathrm{SI}} = 10^{-44}$ (10^{-43}) cm^2.

The theoretically expected rate of WIMPs inside ArDM can not only be used to constrain the WIMP exclusion limit in the final WIMP search, but also to estimate the projected sensitivity before the actual search. Therefore, assumptions have to be made on the data, which can only be confirmed or adapted after the full data analysis. In order to estimate the projected sensitivity, a counting experiment in a region of interest of the variable space will be assumed. The WIMP cut efficiency $\varepsilon_{\mathrm{Cut},\chi}$ combines the efficiency of the data analysis cuts as well as the efficiency into the region of interest, which is bound by a lower S1L threshold. The majority of WIMP searches end in a region of interest which has an expected combined background of $N_{\mathrm{Bkg}} = 0$, and likewise $N_{\mathrm{Obs}} = 0$ for the observed events surviving all cuts. The number of signal events N_{Sig} can then be constrained by the maximal compatible value with $N_{\mathrm{Obs}} = N_{\mathrm{Bkg}} = 0$ at the desired confidence level. While the procedure depends on the exact statistical analysis, for the simplified case of Poisson distributed signal and background event counts in a region of interest with $N_{\mathrm{Obs}} = N_{\mathrm{Bkg}} = 0$, the 90% confidence level intervals on N_{Sig} are 0-2.30 in the Classical [93] and Bayesian [94] approaches and 0-2.44 in the Feldman-Cousins [95] approach. Fig. 2.7 shows the projected ArDM 90% confidence level sensitivity on the upper limit of the spin-independent WIMP-proton interaction cross section σ_p^{SI} versus the WIMP mass m_χ for the simplified analysis scenario of an overall WIMP cut efficiency of $\varepsilon_{\mathrm{Cut},\chi} = 10\%$ in the region of interest for three different lower S1L thresholds. A smaller S1L threshold does not only enable higher sensitivity but also allows to resolve smaller potential WIMP masses. The projected minima are at $\sigma_p^{\mathrm{SI}} = 1.286{\cdot}10^{-43}$ cm^2 for $m_\chi = 95.50$ GeV, $\sigma_p^{\mathrm{SI}} = 3.156{\cdot}10^{-43}$ cm^2 for $m_\chi = 131.83$ GeV and $\sigma_p^{\mathrm{SI}} = 6.694 \cdot 10^{-43}$ cm^2 for $m_\chi = 165.96$ GeV for lower S1L thresholds of 10, 20 and 30 p.e. respectively. While this is not competitive with modern xenon (e.g. Xenon1T [47]) experiments, it is at the order of the currently leading argon experiments (DEAP-3600 [49] and DarkSide-50 [50]); see Fig. 1.5, Section 1.2.

While the simplified scenario of $N_{\mathrm{Obs}} = N_{\mathrm{Bkg}} = 0$ is realistic, the true region of interest will likely not consist of a simple lower S1L threshold with a constant WIMP cut efficiency $\varepsilon_{\mathrm{Cut},\chi}$. In practice, the region of interest is not expected to be defined by a hard cut on the S1L but rather a more complicated cut in a multi-dimensional plane of all discrimination variables S2L/S1L, S1F90 and S1L (see Section 2.1.1). Consequently, the WIMP cut efficiency $\varepsilon_{\mathrm{Cut},\chi}$ into the region of interest is expected to be energy-dependent. However, the choice of $\varepsilon_{\mathrm{Cut},\chi} = 0.10$ is conservative and renders the projected sensitivity presented here a justifiable expectation. With the expected energy dependence of the WIMP cut efficiency, the theoretically expected total WIMP rate is better evaluated in ranges of energy or S1L. Specifically, the expected total WIMP rate $R_{1,2}$ in a bin of S1L with lower edge T_1 and upper edge T_2 is given by

$$R_{1,2} = \int_{\mathrm{T}_1}^{\mathrm{T}_2} \frac{\mathrm{d}R}{\mathrm{dS1L}} \mathrm{dS1L} = R_{\mathrm{aT}}\left(\mathrm{T}_1\right) - R_{\mathrm{aT}}\left(\mathrm{T}_2\right). \tag{2.23}$$

This concludes the necessary tools regarding the theoretical expectations of WIMPs inside ArDM, which will be used to constrain the upper limits on σ_p^{SI} in the ArDM WIMP search in Chapter 6.

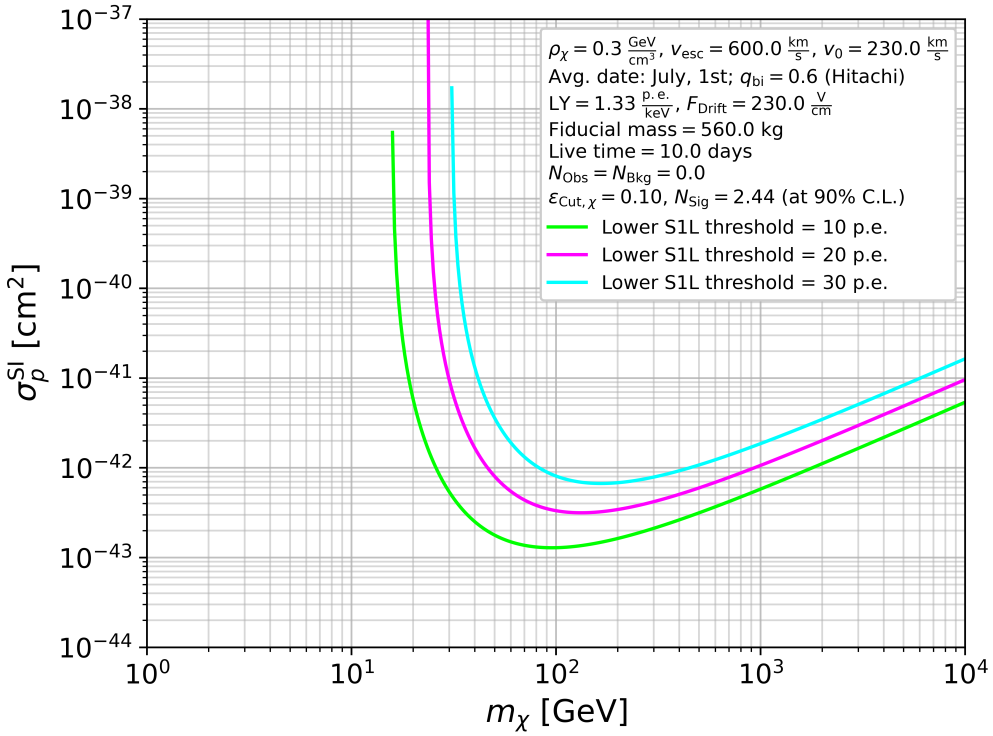

Figure 2.7: Projected sensitivity of the ArDM WIMP search based on reasonable experimental conditions and a simplified region of interest, consisting of a lower S1L threshold and an overall WIMP cut efficiency of 10%. The projected minima are at $\sigma_p^{\mathrm{SI}} = 1.286 \cdot 10^{-43}$ cm^2 for $m_\chi = 95.50$ GeV, $\sigma_p^{\mathrm{SI}} = 3.156 \cdot 10^{-43}$ cm^2 for $m_\chi = 131.83$ GeV and $\sigma_p^{\mathrm{SI}} = 6.694 \cdot 10^{-43}$ cm^2 for $m_\chi = 165.96$ GeV for lower S1L thresholds of 10, 20 and 30 p.e. respectively, and are comparable to the most recent argon experiment results. More details can be found in the text.

3 Data selection and signal reconstruction

Search experiments of the magnitude of ArDM collect an enormous amount of data. Between Feb. 12th, and Aug. 29th, 2019 ArDM collected ∼260 Tb of data divided into 130'367 experimental runs. The subset of dual-phase (DP) data dedicated to the WIMP search was taken between June 19th, and Aug. 25th, 2019, corresponding to ∼127 Tb of data divided into 63'877 experimental runs. A major challenge during the data acquisition is the proper bookkeeping of the experimental runs and the selection of runs which should be processed for further analysis. In ArDM a slow control system continuously monitored the hardware settings and recorded any change (see Section 2.1.2). The resulting information, as well as the feedback from the trigger system of the DAQ were stored in a MySQL database. Additionally, an electronic logbook (ELOG, Appendix A) was kept by the ArDM collaboration, wherein information like the active placement of the neutron source, the exchange of an Ar gas bottle, or general DAQ and detector mode changes were recorded manually. The combination of all these sources of information was used to select the experimental runs of good quality and label them according to the respective detector mode.

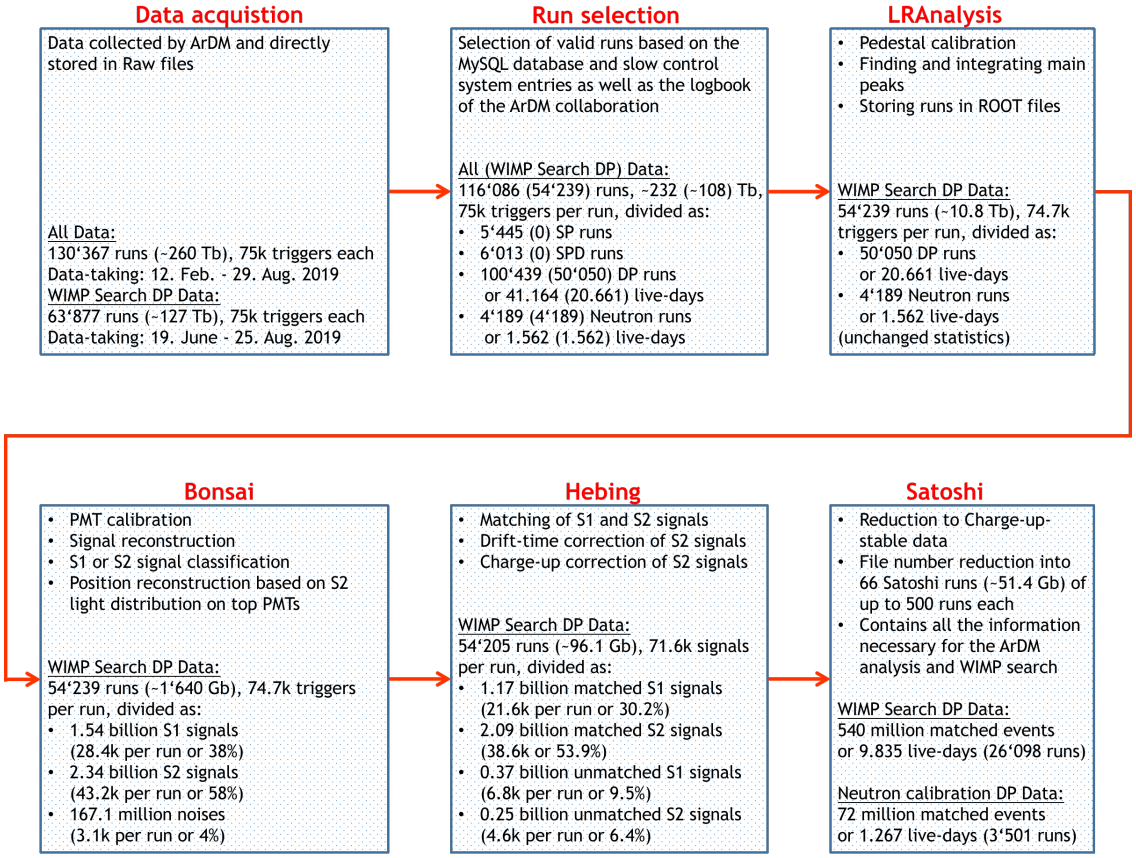

Figure 3.1: Schematic overview and general statistics of the ArDM data processing chain. The data acquisition system is described in Section 2.1.2. The run selection criteria are discussed in detail in Section 3.1, while the trigger and signal reconstructions in the form of LRAnalysis and Bonsai are summarized in Section 3.2 and Section 3.3 respectively. Finally, the event reconstruction is completed within the last step called Hebing; while Satoshi merges chunks of data to reasonable sizes for the WIMP search. The event reconstruction as well as the data selection to Charge-up-stable data are discussed in Chapter 4.

The quality selection criteria were extracted by the author in collaboration with Prof. Dr. André Rubbia, and are the main subject of this chapter treated in Section 3.1. The run selection resulted in ∼232 Tb (or 116'086 runs) and ∼108 Tb (or 54'239 runs) for the total and the WIMP search quality data respectively, as noted in Fig. 3.1. The selected runs are divided into four modes of data-taking. DP runs are dual-phase mode runs with the detector operating with electric drift and extraction fields turned on as described in Section 2.1.2. Neutron runs are DP runs with an active ^{252}Cf source, aimed to investigate the detector's response to nuclear recoils. SPD runs are single-phase drift runs where the extraction field is turned off and the ionization charge is able to drift towards the gaseous phase but without extraction thereinto. Finally, SP runs are single-phase runs where both the drift and extraction fields are turned off, resulting in the primordial detector mode wherein only the primary scintillation light S1L can be detected. As seen in Section 2.2, the conversion of nuclear recoil energy E_{NR} to the detected primary scintillation light S1L is proportional to the drift-field-dependent recombination quenching q_{rec} as well as the light yield LY at zero drift field. While SPD runs could be used to examine the drift-field-dependent recombination quenching q_{rec}, by scanning through various drift field strengths, the SP runs are the ideal mode to gauge the LY of the detector (see Section 5.1).

The signal reconstruction consists of two separate software packages. Firstly, LRAnalysis (Section 3.2) calibrates the pedestal (or baseline) of each individual PMT, finds and integrates the main peaks, and stores the condensed data in ROOT files. This reduces the WIMP search DP data-size to ∼10.8 Tb, divided into 50'050 DP runs (or equivalently 20.661 live days) and 4189 Neutron calibration runs (or 1.562 live days). Secondly, Bonsai (Section 3.3) calibrates the individual PMT single-photoelectron-responses, thereby reconstructing the signals in terms of the number of detected photoelectrons. This results in the ArDM variables of the total light, F90 and TTR, which are then used to classify the signals into S1s and S2s. The distribution of the total S2 light seen by the top PMTs allows to reconstruct the approximate horizontal position of the corresponding signal. Therewith, Bonsai further condenses the WIMP search DP data-size to ∼1'640 Gb divided into 1.54 billion S1 signals, 2.34 billion S2 signals and 167.1 million noise events. Both software packages, LRAnalysis and Bonsai, were already existing before the arrival of the author as part of the ArDM collaboration, and will only be briefly described for the purpose of completeness.

The data processing continues via Hebing which will be treated in detail in Chapter 4. Hebing includes an algorithm that matches the S2 signals to the corresponding S1 signals. A complete event consists of one S1 signal and a number of matched S2 signals. This is due to the fact that multiple S1 signals originating from the same recoiling particle can not be resolved in time, in contrast to the corresponding S2 signals. Hebing is further reducing the WIMP search DP data-size to ∼96.1 Gb, divided into 1.17 billion matched S1, 2.09 billion matched S2, 0.37 billion unmatched S1 and 0.25 billion unmatched S2 signals. Finally, Satoshi merges multiple small runs into reasonably sized chunks of data, resulting in 1.083 billion matched events or 20.661 live days of WIMP search DP data, and 87 million matched events or 1.562 live-days of Neutron calibration DP data. A further selection of Charge-up-stable runs (explained in detail in Section 4.2) will reduce the data to 540 million matched events or 9.835 live-days of WIMP search DP data, and 72 million matched events or 1.267 live-days of Neutron calibration DP data.

3.1 The quality run selection

In a first step of data reduction, the 130'367 experimental runs collected by ArDM between Feb. 12th, and Aug. 29th, 2019 were divided into runs of quality and runs to disregard. The division is based on the information from the ArDM slow control system, the trigger information from the DAQ system and the electronic logbook (ELOG) of the collaboration. The combined information from the DAQ trigger system and the slow control system is stored in a MySQL database[5] on a dedicated server located at CERN (pcnometh29). Among others, the entries from the DAQ trigger system contain the run number, the start- and end-time of the run, the number of triggers collected by one, two, three or all four of the DAQ boards (see Section 2.1.2), and the total DAQ trigger rate. This is complemented by the information from the slow control system. It monitored any update in the voltages applied to the extraction grid, the protection grid, the anode and the cathode (see Fig. 2.3, left), allowing to deduce the respective drift and extraction fields. Runs are disregarded in case of invalid or unreasonable entries, in case of a quality flag set by hand by Prof. Dr. André Rubbia, or if a non-negligible percentage of the signals was not captured by all four DAQ boards (see Section 3.1.1). The ELOG adds further information about manual changes to the DAQ, potential issues with the equipment, and periods where the ^{252}Cf source was positioned next to the detector. A digest of the ELOG can be found in Appendix A. The reconstructed drift and extraction field strengths from the slow control system, and the ^{252}Cf source placements from the ELOG allowed the labeling of the runs into the subtypes SP, SPD, DP and Neutron, and to exclude further runs with unreasonable values (see Section 3.1.2). In order to store the information of which selection criteria a run failed, a second quality flag was introduced by the author. This quality flag is summarized in Section 3.1.3, together with a final summary of the quality data statistics. The combination of all selection criteria reduced the total number of runs from 130'367 to 116'086 (∼89.05%). A more detailed technical documentation of the entire quality run selection, provided by the author, can be found on the CERN GitLab page [96].

Later in the data analysis process, the decision was made to only include the data taken between June 19th, and Aug. 29th, 2019 in the WIMP search, as the DAQ settings were occasionally changed before then (see Appendix A, ELOG entries #25 to #27). Specifically for this subset of the data dedicated to the WIMP search, the quality run selection reduced the total number of runs from 63'877 to 54'239 (∼84.91%). However, the quality selection criteria were fine-tuned based on the complete set of 130'367 runs under the justified assumption that the relevant variables would not change due to the different DAQ settings.

3.1.1 Data cleansing cuts

The aim of the data cleansing cuts is to remove runs which, from the viewpoint of the monitoring database, have unreasonable or invalid characteristics. In the following, the set of data cleansing cuts applied and the resulting individual cut efficiencies are explained based on the entirety of the data. This reflects the manner in which these cuts were defined; independently and based on physical or technical expectations and with the entirety of the

[5] The MySQL database can be accessed directly on an LXPlus terminal via "*mysql -h pcnometh29 -u ardm -p*" followed by a query of the password. It can also be accessed remotely from a local machine by tunnel-forwarding. Therefore, one should use "*ssh -L 12345:pcnometh29.cern.ch:3306 [username]@lxplus.cern.ch*" in a shell to open the port 12345 for access. Then, one should be able to run "*mysql -h 127.0.0.1 -P 12345 -u ardm -p*" in another shell to connect to the MySQL database.

available data. A summary of the individual and combined cut efficiencies on the entirety of the data as well as the reduced subset of WIMP search data can be found in Section 3.1.3.

To start, a negligible amount of 131 (∼0.10%) out of 130'367 runs have either their **starting or ending time set to "*0000-00-00 00:00:00*"** and consequently also have **no events recorded** and an **invalid run duration**. Next, since the number of triggers per run is set to 75'000, the number of total triggers per run should be around 300'000, which corresponds to 75'000 triggers collected on each of the four DAQ boards. The distribution of the total triggers per run (Fig. 3.2, left) shows a dominant peak at 300'000 triggers, as expected, as well as a few trigger-poor runs and some SP runs with up to ∼460'000 triggers (collected with different trigger settings). Consequently, **a lower limit of 295'000 total triggers per run** has been chosen which excludes 283 (∼0.22%) out of 130'367 runs.

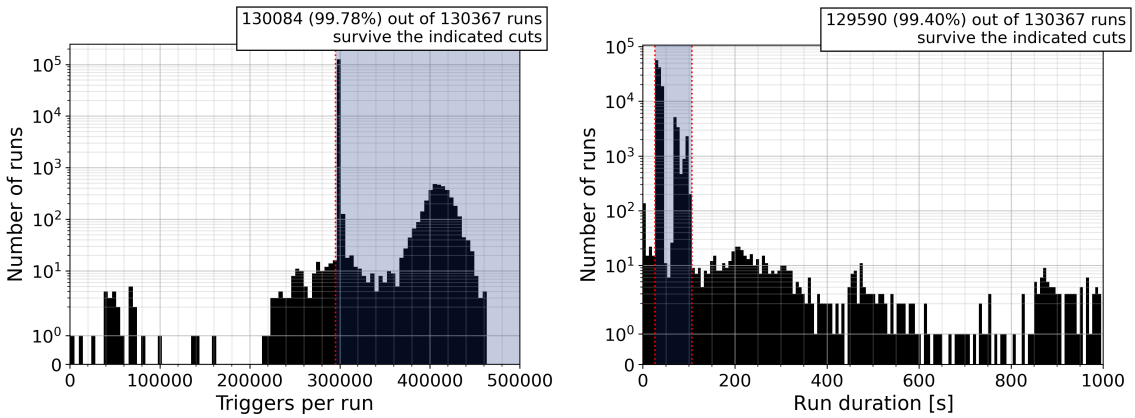

Figure 3.2: The distribution of the total triggers combined for all four DAQ boards per run (left) and the respective run durations (right) for the considered ArDM experimental runs. In the data cleansing process the total number of triggers has an imposed lower limit of 295'000 triggers per run and the run duration is required to be between 26.0 s and 106.0 s.

The trigger rate of the detector is expected to be dominated by electronic recoils originating from the β^- decays of the radioactive isotope ^{39}Ar. As the ^{39}Ar β^- activity of atmospheric LAr is ∼1 Bq/kg$_{\mathrm{LAr}}$ [83, 97] and the total volume of LAr is of the order of one tonne, the expected signal rate is around 1 kHz for single-phase (SP) and single-phase drift (SPD) runs. As electronic recoils are predominantly single-scatterers, the expected signal rate in dual-phase (DP) runs is expected to be at the order of 2 kHz. The distribution of the trigger rate (Fig. 3.3, left) does reveal a narrow peak around 1 kHz and a broader peak around 2 kHz. The evolution of the trigger rates at the order of 2 kHz over time (Fig. 3.3, right) shows a distinct behavior of increasing and resetting, which was later found to be a problem dubbed "Charge-up" and attributed to the extraction grid. It will be extensively discussed in Section 4.2. Based on the distribution of the data **the trigger rate was requested to either be in the range of 950 Hz to 1125 Hz for SP or SPD runs or 1575 Hz to 2700 Hz for DP runs**, thereby excluding 1'281 (∼0.98%) out of 130'367 runs.

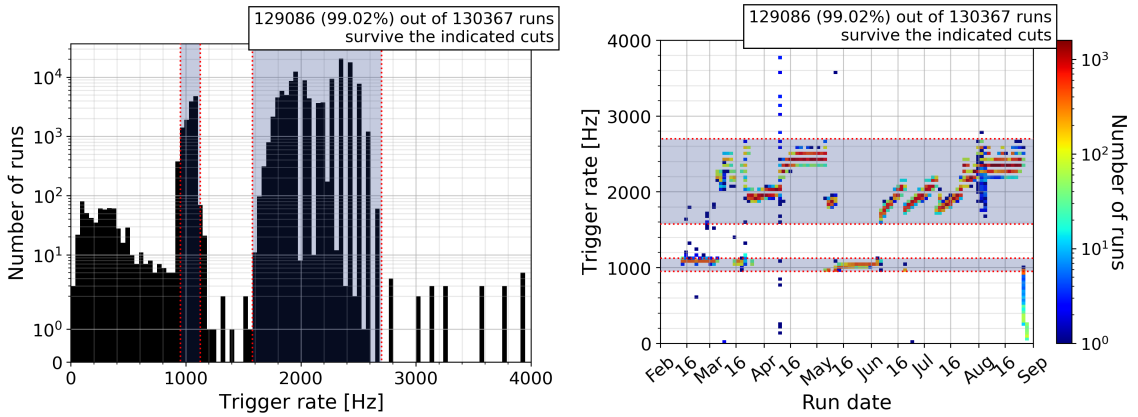

Figure 3.3: The distribution of the trigger rates during the runs (left) and their time evolution (right) for the considered ArDM experimental runs. In the data cleansing process the trigger rate is required to be between 950 Hz and 1125 Hz for SP and SPD runs and between 1575 Hz and 2700 Hz for DP runs. The time evolution shows a repeating increase and reset of the trigger rate caused by the Charge-up effect discussed in Section 4.2.

Furthermore, combining the expected trigger rates of approximately 1 kHz (SP/SPD) or 2 kHz (DP) and the set number of 75'000 triggers per run leads to run durations which are expected to be around 75 s and 37.5 s respectively. The distribution of the run durations (Fig. 3.2, right) indeed reflects this and a cut is imposed on **the run duration to be between 26.0 s and 106.0 s**, which disqualifies 777 (∼0.60%) out of 130'367 runs. The scatter diagram of the total triggers per run versus the trigger rates (Fig. 3.4, left) reveals that almost all DP runs oblige the hard-set trigger number of 300'000 per run (75'000 triggers on all four DAQ boards) and the subgroup of good runs with more than 300'000 triggers are almost exclusively SP or SPD runs. Naturally, the scatter diagram of the trigger rates versus the run durations (Fig. 3.4, right) follows a hyperbolic trend as the product is set to be constant.

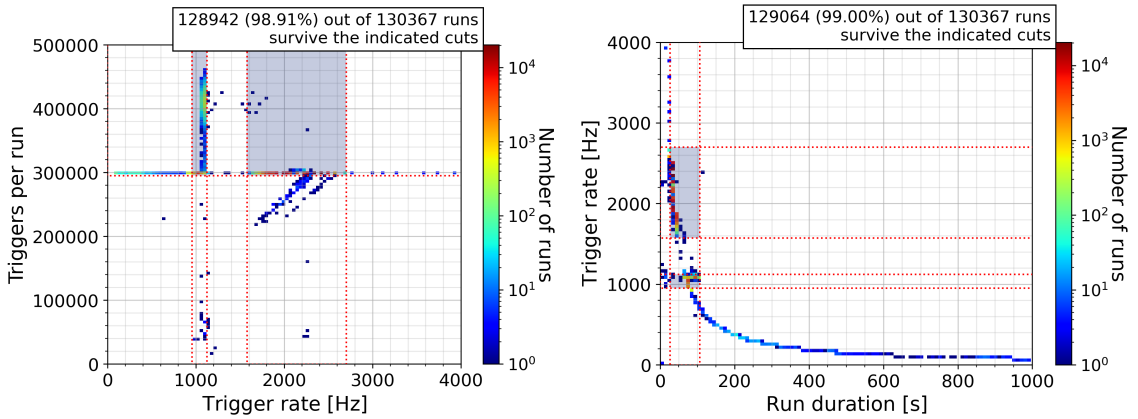

Figure 3.4: *Left*: The scatter diagram of the total trigger numbers versus the trigger rates of the ArDM runs reveals that DP runs are concentrated at the set 300'000 total triggers (75'000 triggers on all four DAQ boards), while SP and SPD runs were also collected with more triggers. *Right*: The scatter diagram of the trigger rates versus the run durations shows a hyperbolic curve for DP runs as a consequence of the preset number of 300'000 triggers.

However, the ideal case of all four DAQ boards registering all 75'000 triggers is not always met in reality. Consequently, the monitored information from all four separate DAQ boards was combined to form the percentage of triggers that was collected by all four DAQ boards. While, in principle, it is feasible that a low energy signal located closely to the bottom or top of the detector deviates this percentage from 100%, this is expected to be a very rare scenario, and in general, a lower percentage rather points to problems in the DAQ. The distribution and time evolution of this percentage (Fig. 3.5) are deliberately shown with an additional 10'967 runs taken immediately prior to the so far considered 130'367 runs. These additional runs resulted in a data-driven **lower limit of 92.0% enforced on the percentage of triggers collected by all four DAQ boards**. This disqualifies 19'731 (∼13.96%) out of the combined 141'334 runs including all 10'967 runs taken before Feb. 12th, 2019. These results imply problems with the DAQ before that date and justify why these early 10'967 runs were not included in the fine-tuning of all other aforementioned cuts. While this percentage also drops below the required 92.0% for a subset of runs taken past Aug. 1st, 2019, this is not alarming as the distribution of all 130'367 runs is strongly peaked at 100% and rapidly decreasing towards lower percentages. However, there is also a considerable amount of runs where this percentage is 0.0%, indicating that at least one DAQ board was not functioning properly. The majority of the runs disqualified in this cut are of this nature. This cut individually removes 8'764 (∼6.72%) out of the 130'367 runs, making it the most exclusive cut.

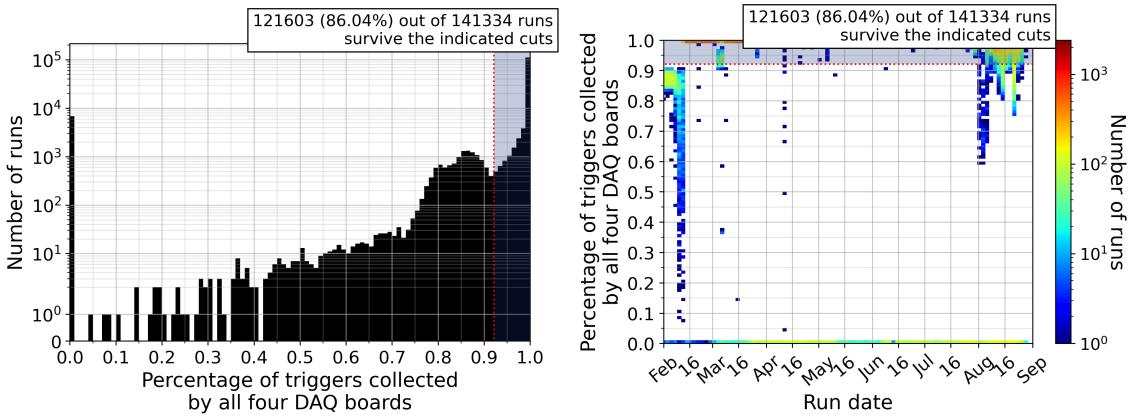

Figure 3.5: The distribution of the percentage of triggers collected by all four DAQ boards for the respective runs (left) and its time evolution (right) for all available ArDM data. All runs taken before Feb. 12th, 2019 resulted in a percentage below 92.0% necessitating a cut at this lower limit in the data cleansing process. The runs taken before Feb. 12th, 2019 were not considered in the fine-tuning of all other data cleansing cuts.

3.1.2 Run-type labeling

During the data-taking period of Feb. 12th, to Aug. 29th, 2019, ArDM was run either in single-phase (SP), single-phase drift (SPD) or dual-phase (DP) mode, and the experimental runs have to be labeled accordingly. While the ArDM collaboration book-kept any changes to the detector settings and the placement of a neutron source (see the ELOG digest, Appendix A), the general distinction into SP, SPD and DP run-types is performed accurately by the monitored field cage settings. The neutron source was only introduced

during DP runs and the relabeled Neutron runs have to be distinguished manually based on the logbook and internal communications.

The electric drift field strength is controlled via the voltage difference applied between the cathode and the extraction grid (see Fig. 2.3, left). The primary applied drift field strength, with which all WIMP search DP data was taken, is 230 V/cm. Additionally, a small number of DP runs were taken at a drift field strength of 340 V/cm and SPD runs were taken at a drift field strength of 270 V/cm. These three run-types are clearly visible in the distribution of the drift field strengths and its temporal progression in Fig. 3.6. By design, the drift field is turned off during SP runs resulting in the fourth major peak around 0 V/cm. Apart from a minor amount of transitional field strengths, there is only 1 (∼0.000008%) out of 130'367 runs which is excluded by a cleansing cut on the drift field strengths. The cleansing cut is applied on the set cathode voltage to be either between 0.0 kV and 0.03 kV, or between 29.5 kV and 43.0 kV. This corresponds to **drift field strengths below 0.27 V/cm or between 226.5 V/cm and 348.4 V/cm** encompassing SP and SPD/DP runs respectively.

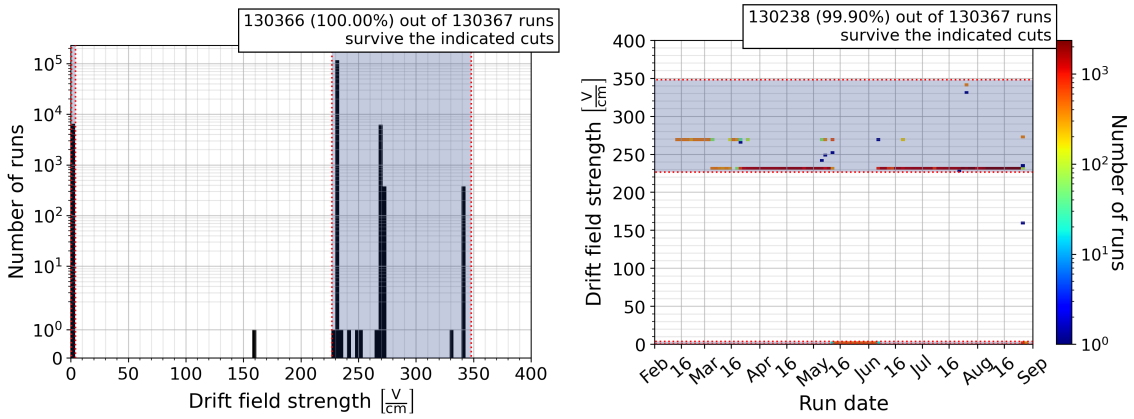

Figure 3.6: The distribution of the set drift field strengths during the runs (left) and its time evolution (right) for the considered ArDM experimental runs. In the data cleansing process the drift field strength is required to be smaller than 0.27 V/cm or between 226.5 V/cm and 348.4 V/cm. SP runs correspond to a vanishing drift field strength, SPD runs were taken at a drift field strength of 270.0 V/cm, while DP runs were primarily performed at 230.0 V/cm (with a small sample of experimental runs at 340.0 V/cm).

The electric extraction field strength is controlled via the voltage difference applied between the extraction grid and the anode (see Fig. 2.3, left). With a field strength of the order of 4 kV/cm it enables the drifting electrons to overcome the potential barrier between the liquid and the gaseous argon phase above. Additionally, it amplifies the kinetic energy of the electrons reaching the gaseous argon, resulting in the delayed scintillation signals S2. As the exact extraction field strength depends on the fill level of the liquid argon (and is further amplified by the unwanted Charge-up effect, as described in Section 4.2), the run settings are reported in terms of the set voltages. The extraction field voltage is either set close to 0 V or 4400 V, for the detector to be in SPD or DP mode respectively, as can be seen in Fig. 3.7. On July 24th, 2019 the ArDM collaboration scanned through the extraction grid voltages from 4.2 kV to 4.6 kV in steps of 0.1 kV for experimental purposes, and subsequently maintained the set extraction voltage of 4.6 kV. Consequently, from the

data cleansing perspective, **the extraction voltage is required to be either lower than 0.35 V or between 4150.0 V and 4650.0 V**, which excludes 10 (∼0.00008%) out 130'367 runs.

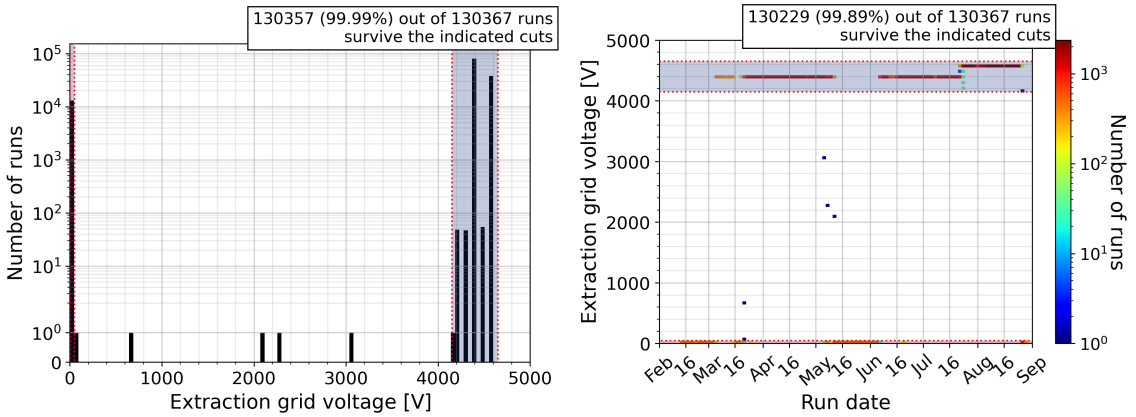

Figure 3.7: The distribution of the set extraction grid voltages during the runs (left) and its time evolution (right) for the considered ArDM experimental runs. In the data cleansing process the extraction grid voltage is required to be smaller than 0.35 V or between 4150.0 V and 4650.0 V. SP and SPD runs correspond to vanishing extraction grid voltages. In DP runs, the extraction field allows the drifting electrons to overcome the potential barrier from the liquid to the gaseous phase in order to produce the delayed scintillation signals S2 therein.

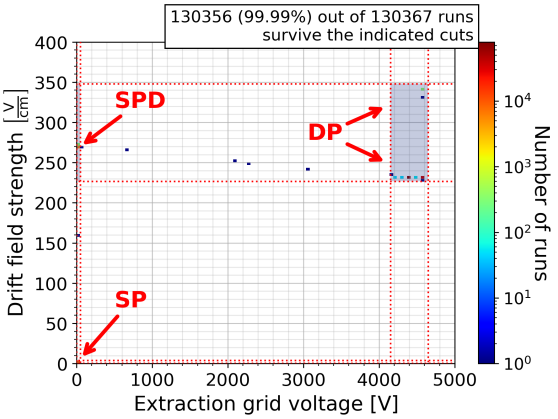

Figure 3.8: The scatter diagram of the drift field strengths versus the extraction grid voltages for the considered ArDM runs reveals the different detector modes SP, SPD and DP.

The three detector modes are revealed when combining the drift field strengths and extraction voltages graphically (see Fig. 3.8). In SP mode, the extraction and drift fields are turned off; in SPD mode, the drift field is turned on allowing the free electrons to drift towards the gaseous phase; in DP mode, the extraction field is also turned on, extracting the electrons into the gaseous phase and producing the S2 signals. A summary of the statistics of the different run-types can be found in Table 2 in Section 3.1.3.

3.1.3 The quality data statistics

With the application of all data cleansing cuts, the 130'367 runs taken between Feb. 12th, and Aug. 29th, 2019 reduce to 116'086 (∼89.05%). As the DAQ settings were also tampered with (see e.g. ELOG entries #25 to #27, Appendix A), the WIMP search data is comprised of a smaller time period of constant final DAQ conditions between June 19th, and Aug. 29th, 2019. Applying the same data cleansing cuts to the WIMP search data results in a reduction from 63'877 to 54'239 (∼84.91%) runs, which are green-lighted to undergo the full reconstruction. After the run-type labeling, the WIMP search data is comprised of 50'050 runs (20.661 live-days) of DP data and 4'189 runs (1.562 live-days) of Neutron calibration data. A summary of all the applied cuts and their individual and combined cut efficiencies (when applied in the listed order) for both data sets is given in Table 1. The final quality data statistics can be found in Table 2.

Cut		**All data**			**WIMP search data**		
#	Cut name	Number of runs	Combined efficiency	Individual efficiency	Number of runs	Combined efficiency	Individual efficiency
	Total	130'367	100.000%	100.000%	63'877	100.000%	100.000%
0	No triggers	130'236	99.900%	99.900%	63'770	99.832%	99.832%
1	All DAQ boards	121'603	93.277%	93.277%	58'651	91.819%	91.819%
2	Trigger rate	120'535	92.458%	99.017%	58'649	91.816%	99.829%
3	Drift field	120'534	92.457%	99.999%	58'649	91.816%	100.00%
4	Extraction field	120'527	92.452%	99.992%	58'649	91.816%	100.00%
5	Too short	120'507	92.437%	99.855%	58'649	91.816%	99.831%
6	Too long	120'507	92.437%	99.549%	58'649	91.816%	100.00%
7	Too few triggers	120'502	92.433%	99.783%	58'647	91.812%	99.646%
8	Weak source	120'221	92.217%	99.774%	58'374	91.385%	99.551%
9	Neutron veto	119'860	91.940%	99.704%	58'013	90.820%	99.396%
10	DAQ problems	116'086	89.046%	96.820%	54'239	84.912%	93.509%

Table 1: A summary of all the quality data selection cuts explained in Section 3.1 with their respective run statistics (after the cut), the individual and combined cut efficiencies (when applied in the listed order). One run corresponds to ∼75'000 signal triggers (combined from all four DAQ boards). Cuts #0 to #2 and #5 to #7 are explained in detail in Section 3.1.1, cuts #3 and #4 are discussed in Section 3.1.2 and cuts #8 to #10 are detailed here in Section 3.1.3. A continuation of this table past the reconstruction can be found in Table 4, in Section 4.2.1.

The cuts #0 to #2 and #5 to #7 are explained in detail in the data cleansing in Section 3.1.1 while cuts #3 and #4 are detailed in the run-type labeling in Section 3.1.2. Thereof, only cut #1 "All DAQ boards", which requires the percentage of triggers that were detected by all four DAQ boards to be higher than 92.0%, removes a non-negligible part of the data. Additionally, cuts #8, #9 and #10 were introduced during the data analysis process. The first part of the neutron calibration data was taken with a ^{252}Cf source with a neutron activity of ∼100 Bq (see Appendix A, ELOG entries #24 and #29), which was soon replaced by a stronger ^{252}Cf source with a neutron activity of ∼500 Bq (see Appendix A, ELOG entries #30, #32, #35, #36, #41 and #43 to #45). The weaker ^{252}Cf source

led to a small number of detected neutrons, which was below the threshold of aiding the calibration of the detector's response to nuclear recoils. Hence, the weaker ^{252}Cf source runs were discarded with cut #8, which removes 295 (∼0.226%) out of 130'367 runs of the entire data and 287 (∼0.449%) out of 63'877 runs of the WIMP search data respectively. In terms of the neutron calibration data, cut #8 removes 295 (∼6.58%) out of 4'484 Neutron runs in the entire data and 287 (∼6.41%) out of 4'476 Neutron runs in the WIMP search data respectively.

The cut #9 "Neutron veto" refers to time periods wherein other experiments in the LSC experimental hall used a neutron source for their own purposes and therewith increased the instrumental neutron background flux (see Appendix A, ELOG entries #34 and #40). The neutron veto, via cut #9, removes 386 (∼0.296%) out of 130'367 runs of the entire data and 386 (∼0.604%) out of 63'877 runs of the WIMP search data. Finally, the cut #10 "DAQ problems" refers to a subset of 3'616 runs with unreasonably large fractions of S2 signals, which were discovered during the data analysis process (see Section 4.2), as well as 530 runs which led to crashes of the reconstruction process. The DAQ problems could potentially be related to an exchange of the argon gas bottle or a subsequent crash of the DAQ (see Appendix A, ELOG entries #39 and #42), and cover the period from Aug. 1st, until Aug. 5th, 2019. The cut #10 removes a combined 4'146 (∼3.18%) out of 130'367 runs and 4'146 (∼6.49%) out of 68'877 runs for the entire data and the WIMP search data respectively; concluding the run selection cuts. A continuation of this table past the reconstruction can be found in Table 4, in Section 4.2.1.

A quality flag was added to the ArDM run database in order to record the quality information of each run. For each failed cut the respective run's quality flag (initialized at zero) was subtracted by 10 to the power of the cut number. Thus, quality runs that survived all cuts would maintain a quality flag of "0", whereas a run with the example quality flag of "−10000110" has less than 92.0% of triggers collected by all four DAQ boards, an unreasonable trigger rate, and less than 295'000 total triggers.

Finally, the distribution of the surviving runs classified into the different run-types can be found in Table 2. The 4.611 live-days of SP data are dedicated to gauge the light yield of the detector (see Section 5.1); the 1.562 live-days of Neutron data are assigned to gauge the detector's response to nuclear recoils (see Section 5.2); and finally, the 20.661 live-days of DP data are dedicated to conduct the WIMP search (see Chapter 6). An updated version of this table can be found in Table 5 in Chapter 5, following the entire reconstruction and subsequent data analysis.

	All data			**WIMP search data**		
Label	Runs	Live-days	Percentage	Runs	Live-days	Percentage
SP	5'445	4.611	8.66%	0	0.000	0.00%
SPD	6'013	5.917	11.11%	0	0.000	0.00%
DP	100'439	41.164	77.30%	50'050	20.661	92.97%
Neutron	4'189	1.562	2.93%	4'189	1.562	7.03%
Total	116'086	53.254	100.00%	54'239	22.223	100.00%

Table 2: A summary of the statistics of the different run-types of the entire data and the WIMP search data surviving all quality selection cuts listed in Table 1. One run corresponds to ∼75'000 signal triggers. An updated version of this table can be found in Table 5 in Chapter 5, following the entire reconstruction and subsequent data analysis.

3.2 Trigger reconstruction - LRAnalysis

LRAnalysis is the first step in the data reconstruction chain (see Fig. 3.1), which transforms the raw data from the PMTs into ROOT files containing 75'000 triggers each. Therefore, the raw waveforms from the DAQ boards, each reading six PMTs, are combined according to the universally synchronized time tags. The first part of LRAnalysis consists of the pedestal calculation or baseline finding for each PMT. Therefore, the 800 ns before the trigger time are evaluated using an iterative pedestal finding algorithm which aims to remove dark counts and pile-up.

An example S1 (left) and S2 (right) waveform can be seen in Fig. 3.9 after the pedestal subtraction, in Analog-to-Digital-Converter (ADC) units (per time samples of 4 ns). The trigger time is located at ∼468 time samples or ∼1872 ns and the duration of the whole signal accounts to 4096 time samples or 16.384 μs. The example S1 waveform shows the expected dominant fast scintillation component, that attains the F90, or "the fraction of signal detected in the first 90 ns (22.5 time samples) after the trigger time", at a value of ∼0.3 for ERs and ∼0.7 for NRs. Conversely, the S2 trigger consists of the slower scintillation process in the gaseous phase resulting in a more continuous waveform and consequently very low values of the F90.

The second part of LRAnalysis consists of summarizing the timestamps and waveform integrals of the individual PMT hits. These individual PMT hits are potentially combined to so-called clusters, wherein adjacent time samples with counts above the threshold of 10σ of pedestal fluctuations are summed. The integrated sizes and timestamps of these clusters are then stored in the continuously numbered run-based ROOT files of 75'000 triggers each. These files can then undergo the signal reconstruction step next (Bonsai, Section 3.3).

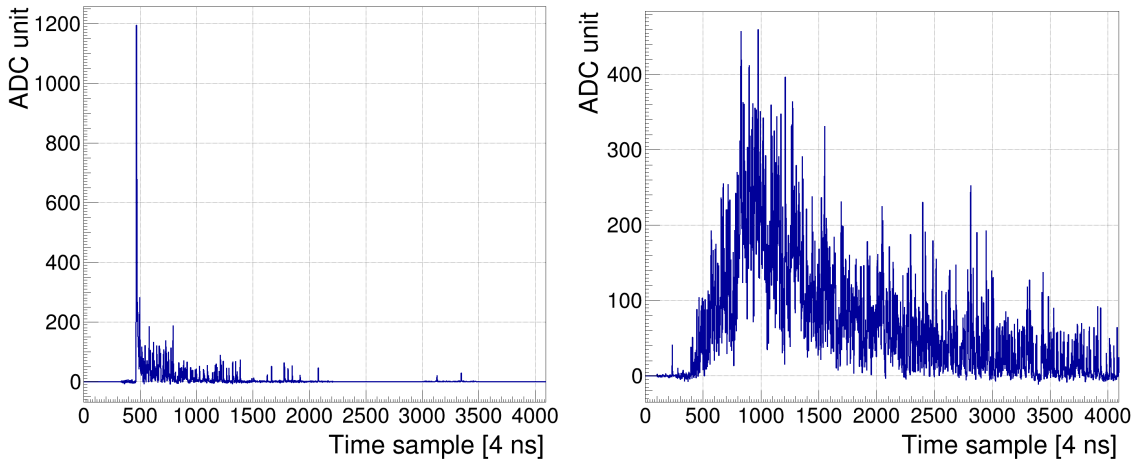

Figure 3.9: A sample S1 (left) and S2 (right) trigger of an ER event resulting from the first part of the ArDM data reconstruction chain. The pedestal of each PMT is calculated individually from the 800 ns (or 200 time samples) before the trigger time and subtracted. LRAnalysis then stores the sizes and timestamps of all the clusters of single and adjacent PMT hits.

3.3 Signal reconstruction - Bonsai

Bonsai is the second step in the data reconstruction chain (see Fig. 3.1), which transforms the still unclassified reconstructed triggers (from LRAnalysis, Section 3.2) in units of ADC counts into classified S1 and S2 signals in units of detected and reconstructed photoelectrons (p.e.). While the main features and results of Bonsai are summarized here, a more detailed description and study of Bonsai can be found in the PhD thesis of Wei Mu [78].

The conversion factor from cluster integrals of ADC counts to actual photoelectrons is given by the center of the single-photon-peak in the distribution of the peak integrals for each individual PMT and run. Fig. 3.10 (left) shows an example distribution of the peak integrals of a PMT and the resulting single-photoelectron response (SPE, red line). Bonsai uses a fast and iterative algorithm, which estimates the SPE based on the averaged maxima of the peak integral distributions in various slices of time samples. In principle, a best-fit result of a linear combination of a single-, double- and triple-photon peak would be the most physically motivated approach. However, past efforts of the collaboration have shown that the current Bonsai method leads to very comparable results, which agree within 5% with the physically motivated approach [78]. In practice, Bonsai's automated algorithm to extract the SPE is preferred due to its robustness, which is key when handling the large amounts of data collected by ArDM.

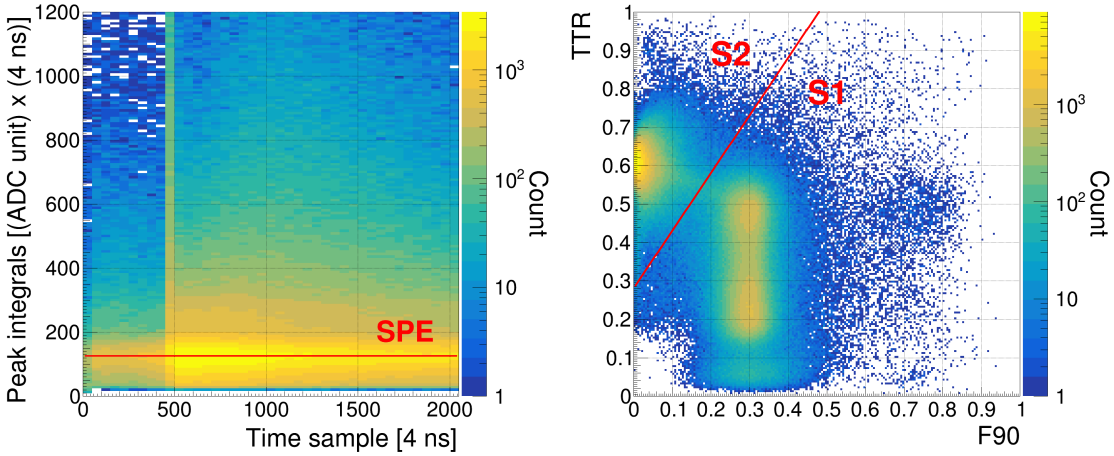

Figure 3.10: *Left*: An example distribution of the peak integrals from a single PMT. The extracted single-photoelectron response (SPE, red) allows the calibration of each PMT (ADC counts to detected photoelectrons). *Right*: An example distribution of the reconstructed TTR versus F90 from a small number of DP runs. The accumulations of S1 and S2 signals are clearly visible and are classified by the distinction of $\mathrm{TTR} = 1.5 \cdot \mathrm{F90} + 0.28$ (red line).

Having calibrated the individual PMT photoelectron responses, the signal summary variables of the total light, the TTR and the F90 can be calculated. With the **total light** corresponding to the total signal integral in photoelectrons, and the **F90**, or the percentage of the light detected in the first 90 ns of the signal, following directly. However, the longer lifetimes of the triplet states of 1.6 μs and 3.0-3.2 μs [98] in LAr and GAr can potentially lead to multiple consecutive DAQ triggers for single signals (see the incomplete S2 signal in Fig. 3.9, right). Therefore, consecutive sets of triggers with a time difference of less than 1000 ns to the previous trigger, a total light of less than 8 p.e., and an F90 of less than 0.2

are merged with the original trigger; and the total light, F90 and TTR are updated for the complete signal.

The **TTR** corresponds to the "top-to-total ratio", and hence the fraction of the total light which was detected by the twelve PMTs located at the top of the detector. Consequently, the TTR correlates with the vertical position of the original scatter centers within the detector. Fig. 3.10 (right) shows the distribution of the TTR versus the F90 of a small number of example DP runs. Therein, a clear distinction between S2 signals with high TTRs of ∼0.6 and F90s ≤ 0.1, and S1 signals with TTRs distributed throughout the whole detector length and F90s around ∼0.3 for ERs and ∼0.7 for NRs is visible. Consequently, Bonsai classifies the signals into S1 and S2 according to the separation of these two distributions (red line, $\mathrm{TTR} = 1.5 \cdot \mathrm{F90} + 0.28$).

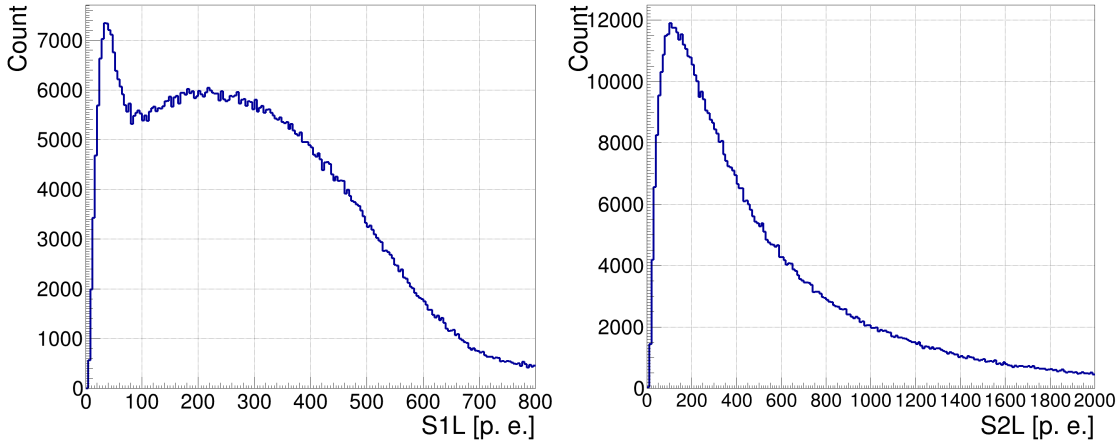

Figure 3.11: Example distributions of the S1 (left) and S2 (right) total light in photoelectrons from a small number of DP runs. The distribution of the S1 total light (S1L) consists of an ^{39}Ar β^- spectrum on top of an exponential background, convoluted with a Gaussian detector resolution. It can be used to gauge the energy scale from photoelectrons to keV (of ER) in Section 5.1. The S2 total light (S2L) spectrum needs to be corrected for the absorption of electrons along the drift path (Section 4.1.2) as well as the later discovered Charge-up (Section 4.2).

Having classified the signals into S1s and S2s, the general reconstructed signal variables can be displayed separately. Fig. 3.11 shows the distributions of the total lights for S1s and S2s. The **S1 total light (or S1L)** distribution consists of an ^{39}Ar β^- spectrum on top of an exponential background convoluted with a Gaussian detector resolution. The corresponding function will be fitted to extract the detector's characteristic light yield in photoelectrons (p.e.) per keV of ER via the comparison to the known ^{39}Ar β^- spectrum energy scale (see Section 5.1). Ideally, the distribution of the **S2 total light (or S2L)** would be correlated with the S1L distribution. However, this is not the case for two specific reasons. Firstly, the S2L is attenuated due to the absorption of the ionization charges along their drift tracks. This attenuation follows a Lambert-Beer exponential absorption law and is corrected for in Section 4.1.2. Secondly, the Charge-up effect, discovered at a later stage, has significantly altered the applied extraction field in some parts of the detector, influencing resulting S2L distribution. The Charge-up is explained further and corrected for in Section 4.2. Therefore, the S2L distribution can not be interpreted directly

at this stage, as there are height-dependent (drift time attenuation) and position- and live-time-dependent (Charge-up) effects involved.

The distributions of the F90 for S1s and S2s can be seen in Fig. 3.12, which are essentially one-dimensional projections of the two distinct regions in the TTR versus the F90 distribution (Fig. 3.10, right). The **S1F90** distribution is strongly peaked at 0.3 corresponding to the ERs from the ^{39}Ar β^- spectrum and has a small NR contribution at ∼0.7. Conversely, the **S2F90** distribution is strongly peaked around 0.0 as forced by construction and as expected due to the long lifetime of the triplet state of 3.0-3.2 μs [98].

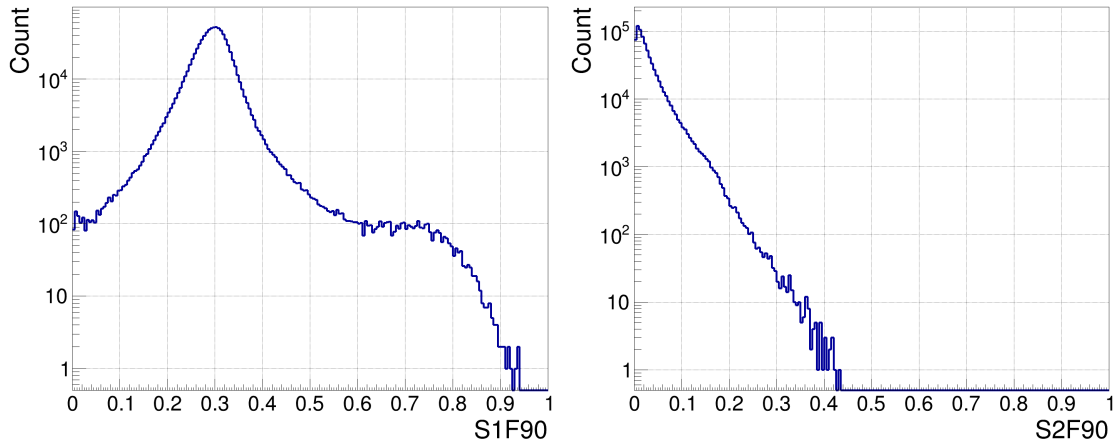

Figure 3.12: Example distributions of the S1F90 (left) and S2F90 (right) from a small number of DP runs. The F90 distributions are projections of the subsets of S1 and S2 signals in the combined TTR versus F90 plane (Fig. 3.10, right). The ER events are the dominant contribution to the S1F90 spectrum at an F90 value of ∼0.3 and the NR events represent a WIMP-like signal at an F90 value of ∼0.7. The S2F90 distribution is strongly peaked at 0.0 due to the S2 events being of ER nature and dominated by the longer triplet lifetime of 3.0-3.2 μs [98] in the gaseous phase.

The distributions of the TTR for S1s and S2s (Fig. 3.13) represent the distributions of the vertical positions of the signals. The individual features of the **S1TTR** distribution are not completely understood. The two dominant peaks at S1TTR values of ∼0.22 and ∼0.48 were reproduced to an extent by past simulation efforts [78] of the ArDM collaboration, incorporating the ^{39}Ar β^- spectrum as well as γ backgrounds from environmental ^{238}U, ^{232}Th and ^{40}K. The smaller peak at an S1TTR of ∼0.06 is currently not understood. The S1TTR variable will be introduced into the analysis as a measure of the matching quality of S2-S1 pairs (Section 4.1.1). In the case of single-scatters with only one S2 signal per S1 signal, the S1TTR is expected to be well correlated with the time difference between the detections of the S1 and the corresponding S2 signal. Apart from being used as a matching quality variable, the S1TTR and S2TTR values are not used in the actual WIMP search. However, the **S2TTR** distribution allows to gauge the vertical position of the LAr-to-GAr boundary to a TTR value of roughly 0.55 (as S2 signals originate from within the gaseous phase). The fact that the TTR is not extending all the way up to unity is attributed to an increased absorption of photons by the PMMA plate (with the ITO and TPB coating, see Section 2.1.2) at the top of the detector [78].

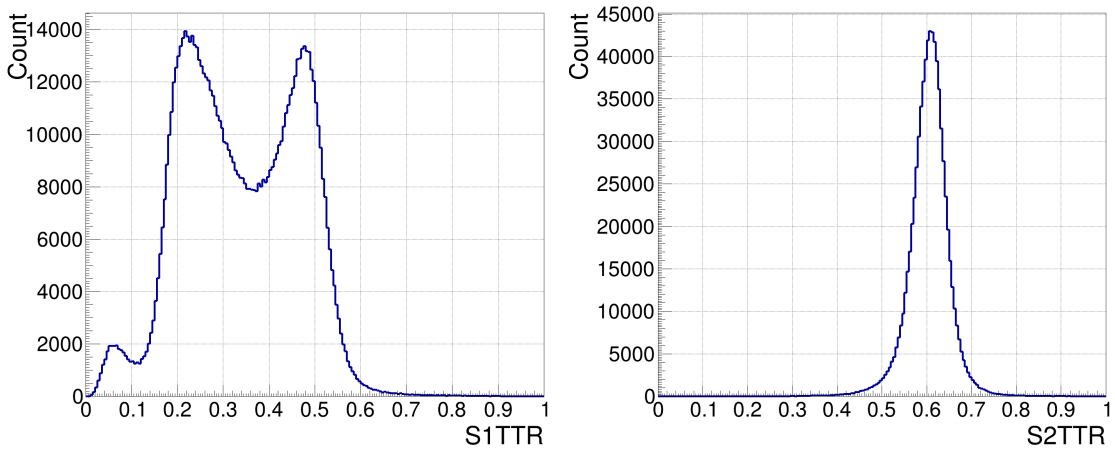

Figure 3.13: Example distributions of the S1TTR (left) and S2TTR (right) from a small number of DP runs. The TTR distributions are projections of the subsets of S1 and S2 signals in the combined TTR versus F90 plane (Fig. 3.10, right). While the two dominant S1TTR peaks at TTR values of ∼0.22 and ∼0.48 were, to an extent, reproduced by simulation efforts [78], the detailed features of the S1TTR distribution are not completely understood. In terms of the WIMP search analysis, the S1TTR is only used as a measure of the matching quality of S1-S2 pairs (Section 4.1.1). The S2TTR distribution is strongly peaked at a TTR value of ∼0.62 and allows to gauge the LAr-to-GAr boundary to a TTR of roughly 0.55.

Another interesting reconstructed variable, which is part of the WIMP search, is the **S1maxFracPMT**, or the maximal fraction of the total S1L which was recorded by a single PMT. By construction this variable is limited to values between $1/24 = 0.041\overline{6}$, where all the S1L is evenly distributed among all 24 PMTs, and 1.0, where all the S1L is captured by a single PMT. The distribution of the S1maxFracPMT (see Fig. 3.14, left) is strongly peaked at a maxFracPMT value of ∼0.093 and sharply decreases towards higher values. This indicates a quite even light distribution for the majority of the S1 signals, which certainly corresponds to the preferred state of a signal. Whereas a one-sided light distribution can be a consequence of background radiation or dark counts. Consequently, the S1maxFracPMT variable will be cut to an upper limit during the WIMP search (Chapter 6).

The last part of the Bonsai signal reconstruction software concerns the horizontal position reconstruction of the classified S2 events. After having drifted upwards in the liquid phase inside the drift field cage, the surviving electrons should still be centered around the horizontal position of the original scatter. Consequently, the S2 light pattern on the twelve top PMTs should clearly hint towards the original horizontal position. Generally, there exist many methods to reconstruct the horizontal position based on the light pattern, which include the center-of-gravity method [99], the least squares test [100] and neural networks [77]. Bonsai uses the **least squares (LS) test to reconstruct the S2 horizontal position** from a least squares fit result of the detected light pattern to precomputed PMT hit-pattern maps [78]. The PMT hit-pattern maps are generated from the ArDM Monte Carlo simulation and store the normalized probability for each individual PMT to capture photons as a function of the horizontal x and y position. An example resulting reconstructed position distribution (using the LS algorithm) can be seen in Fig. 3.14 (right). In general, the reconstructed horizontal distribution of S2 events is expected to be uniform within the

detector's central volume. The data suggests that this is not the case with the applied LS algorithm, as there seems to exist a bias due to the PMT positions. Additionally, the shadows of some detector components are visible at larger radii of the detector's horizontal cross section. For the purpose of the ArDM WIMP search however, this is not a problematic obstacle. The main purpose of the reconstruction of the horizontal position is to reduce the WIMP search region within the detector to a fiducial volume at the center to avoid boundary effects and backgrounds originating from detector components (see the fiducial volume cut in Section 5.3). For this purpose, the precision of several centimeters in the central volume (radii smaller than 35 cm) given by the LS algorithm [78] is fully sufficient. Therewith, the signal reconstruction is concluded and the data is ready to undergo the last reconstruction step, which consists of recombining S1 and S2 signals to complete events within Hebing (Chapter 4).

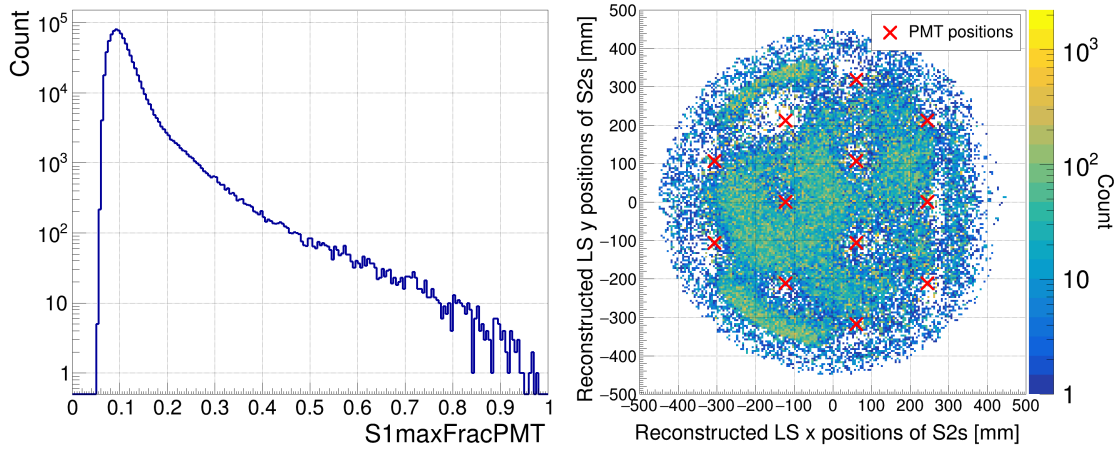

Figure 3.14: Example distributions of the S1maxFracPMT (left) and S2 LS reconstructed horizontal positions (right) from a small number of DP runs. The S1maxFracPMT distribution is strongly peaked at a maxFracPMT value of ∼0.093 indicating that the S1 light is evenly distributed among the 24 PMTs in the majority of the S1 signals. The distribution of the S2 LS reconstructed horizontal positions (right) shows features of PMTs as well as other detector components. The S2 LS reconstructed positions are used to confine a central fiducial volume for the WIMP search region (Section 5.3), and therefore require a precision of several centimeters in the central volume of the detector.

4 Event reconstruction - Hebing

This chapter covers the continuation of the data processing chain (Fig. 3.1), after the trigger reconstruction in the form of LRAnalysis (Section 3.2) and the subsequent signal reconstruction in the form of Bonsai (Section 3.3). Hebing completes the data reconstruction by correcting the S2 signals for physical effects and by matching them to their respective S1 signals, resulting in the fully reconstructed events.

Section 4.1 describes the functionality of the matching algorithm and presents the additional variables which result from the combination of S1 signals with the corresponding S2 signals. Furthermore, a matching quality variable is introduced (Section 4.1.1), which measures the physical accuracy of single-scatter matches by comparing the observed time difference between the S2 and S1 signals and the approximate vertical position of the S1 signal. During the upward drift a fraction of the free ionization charges recombines, which therefore reduces the observed S2 light (S2L) as a function of the drift time. The corresponding drift-time correction is explained in Section 4.1.2. During the data analysis procedure a phenomenon, designated "Charge-up", was discovered (Section 4.2). Therein, electric charges have accumulated in the upper part of the extraction grid which led to a continuously increasing extraction field strength further altering the S2L. While the Charge-up led to a series of unwanted effects, a thorough analysis procedure allowed to understand their consequences, accordingly adapt the data selection and processing, and correct the S2L to the extent possible (Section 4.2.1). The Charge-up correction completes the event reconstruction of Hebing. The matching algorithm was primarily developed by Dr. Balint Radics of the ArDM collaboration. The author's work revealed the Charge-up effect, and contributed the drift time and Charge-up corrections.

Hebing was a major milestone in the ArDM analysis process, and was developed and iteratively improved over the course of multiple years. While it is not possible to cover this entire process in detail, this chapter aims to explain the latest version of Hebing in a concise but complete manner. However, the discovery of the Charge-up necessitated an adaptation of the matching criterion. As such, parts of the discovery (Section 4.2) and the correction (Section 4.2.1) of the Charge-up are based on data using a previous matching criterion and thus an earlier version of Hebing. The corresponding figures are labeled accordingly to avoid confusion.

Furthermore, the data presented in this chapter originates from many different stages of the analysis process. While the final cuts used in the WIMP search are explained in detail in Chapters 5 and 6, the most commonly used ones are listed in Table 3. Emphasis is invested on presenting all variables before presenting their respective cuts. If not stated otherwise, the listed cuts are applied in order and only the ultimate cut is used to label the respective data.

Cut name	Explanation	Variable(s) before cut
General	S1L $\leq$ 800 p.e.	Fig. 3.11 (left)
Low energy	S1L $\leq$ 100 p.e.	Fig. 3.11 (left)
Preselection	nS2 $\geq$ 1 and S1maxFracPMT $\leq$ 0.3	Fig. 3.14 (left) and Fig. 4.2
Single-scatterers	nS2 = 1	Fig. 4.2 (left)
Quality matches	0.75 $\leq$ S2maxQMatch $\leq$ 1.25	Fig. 4.4 and Fig. 4.5
Pile-up cut	nS2perS1 = 1 and nS1perS2max = 1	Figs. 4.23 and 4.24

Table 3: A list of the most commonly applied cuts to the data presented in this chapter. If not stated otherwise, the listed cuts are applied in order and only the ultimate cut is used to label the respective data.

After having completed the data reconstruction, Satoshi (see Fig. 3.1) merges together multiple experimental runs into reasonably sized chunks of data. The final data statistics result in 540 million matched events (26'098 runs) or 9.835 live-days of WIMP search DP data, and 72 million matched events (3'501 runs) or 1.267 live-days of Neutron calibration DP data (see the data selection in Table 4 in Section 4.2.1 and the detailed statistics in Table 5 in Chapter 5). The fully reconstructed data then is ready to be studied for the experimental backgrounds (Chapter 5) and to be analyzed in the final WIMP search (Chapter 6).

4.1 Matching of S2s to S1s

One of the major challenges on the software side of dual-phase experiments, such as ArDM, is to reconstruct the proper pairing of S1 and S2 signals. After the successful signal reconstruction, the data consists of a sequence of S1 and S2 signals, each with a timestamp, and corresponding signal variables such as the total light, the TTR and others. The required result is a set of complete events consisting of one S1 signal and a number of matched S2 signals. In contrast to the corresponding S2 signals, multiple S1 signals originating from the same recoiling particle are generally not resolvable in time. The boundary conditions of the matching are given by the allowed time differences between pairs of S1s and S2s, which must be in the range of zero to the maximal possible drift time (maxDT). This problem, excluding the boundary conditions, is akin to the "College Admissions" problem discussed by Gale and Shapley [101]. Therein, a set of colleges with quotas for students to be admitted and a set of potential students each have a preference list for the respective other entity. A matching algorithm then finds a mapping of students and colleges which is considered stable, in the sense that there exists no other mapping with the same or a higher number of pairings with higher preferences. ArDM applies a similar algorithm to the matching problem of S1s and S2s [102], which was primarily the work of Dr. Balint Radics. Therein, the S2s can be considered the students and the S1s with unknown scattering multiplicity are akin to the colleges with their quotas. The preference of S1s and S2s for the respective other signal type can be expressed by a likelihood function based on the available signal variables. The advantage of the dual-phase signal matching over the college admission problem is the fact, that this likelihood function, or explicitly the preference for the other entity, is symmetric. Hence the likelihood that an S2 signal will assign to an S1 signal is the same as vice versa, as the full available information is used for each potential pair. The temporal boundary conditions dictate the potential candidates. Specifically, the ArDM matching algorithm loops through small groups of consecutive S2 signals. It then includes all potential S1 signals as per the boundary conditions and starts the matching procedure. A margin for errors is given by allowing maximal time differences of $1.2 \cdot \text{maxDT}$ as visualized in Fig. 4.1. Then, the S1s and S2s construct ranking tables for the respective other signal candidates and the matching starts to undergo an iterative process. The students, or S2s, apply to the preferred colleges, or S1s. The S1s then puts the preferred S2 among the potential applicants on a waiting list. Specifically, in the schematic example in Fig. 4.1, the S2s #2 and #3 both prefer the S1 #2. However, the S1 #2 prefers the S2 #3 over the S2 #2 and hence puts the S2 #3 on the waiting list. Then, in an iterative process, the S2s apply to their next highest priority and the matching repeats itself. If the colleges, or S1s, were to have a maximal quota of places they would eventually start to compare preferred new students, or S2s, to the ones which are already on the waiting list

and potentially dump some of them in favor of the new, more preferred S2s. However, in the case of ArDM the maximal scattering multiplicity is not limited and the quota is put to an arbitrarily high number.

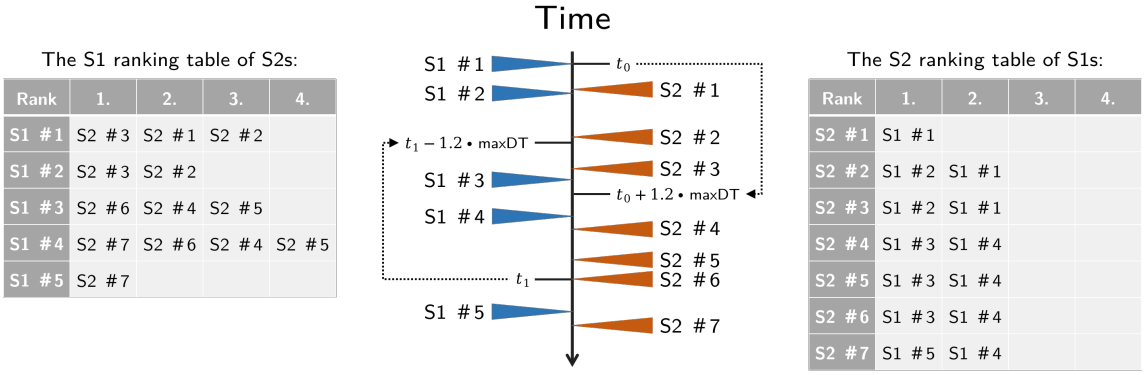

Figure 4.1: A visualization of the working principle of the matching algorithm. The S1 and S2 signals are distributed on a timeline and rank every potential candidate of the respective other entity according to a supplied ranking criterion. The matching algorithm then builds a collection of stable S1-S2 pairings of the highest possible compatibility. One S1 signal combined with all S2 signals matched to itself constrain one fully reconstructed event.

The visualization in Fig. 4.1 is an exaggeration of a realistic matching scenario. As one will be able to see in Fig. 4.10 (right), the rate of S1 signals is approximately 800 Hz after the reconstruction. Combined with the maximal observed drift time of 1.044 ms (shown in Fig. 4.11, right) and the margin-for-error factor of 1.2, this results in an expectation of $\lambda = 1.00224$ S1 signals per matching window. Hence, for a Poisson process, the probability to find 0, 1, 2, 3, 4 or 5 S1 signals within a matching window accounts to 36.7%, 36.8%, 18.4%, 6.2%, 1.5% or 0.3%. This simplified estimation gives an idea of the likelihood of the number of S1s in a matching scenario. Conversely, the situation is different from the S2 perspective. The average S2 signal frequency starts at 700-800 Hz (Fig. 4.10, right) and increases up to 1400 Hz over the data-taking period due to the Charge-up (explained in Section 4.2). As the final WIMP search data will concern the stable data period with an S2 signal frequency of 1400 Hz, the Poisson parameter equates to $\lambda = 1.2 \cdot 1.044\text{ms} \cdot 1.4\text{ms}^{-1} = 1.75392$ expected S2 signals for one matching window. Naturally, the time distribution of the S2 signals is dependent on the time distribution of the S1 signals and their energy, and hence is not exactly a Poisson process. Nevertheless, the approximation of a Poisson process serves as an estimate of the lower bound of S2s per matching window. If the S2 signals were a true Poisson process, the probability to find 0, 1, 2, 3, 4 or 5 S2 signals within a matching window would account to 17.3%, 30.4%, 26.6%, 15.6%, 6.8% or 2.4%. Based on that, one would naively estimate the probability for an S1 signal to see more than one S2 signal within a matching window to be ∼50%. However, as discussed, this probability can only serve as an estimated lower bound, as the S2 signal time distribution is not independent of the distribution of S1 signals. In reality, the occurrence of an S1 signal will positively influence the probability of S2 signals within the following matching window. Conversely, the occurrence of an S2 signal without any S1 signal in the matching window before can only be a consequence of an S1 signal not being reconstructed properly. The actual distributions of the number of S1 and S2 signals in the data, from the perspective

of the other entity, are shown later in Fig. 4.23 and are discussed there. The presented statistical thought experiment should serve to illustrate why the question of matching is so important for ArDM. Since it is expected to be more probable for a pile-up of signals to occur than not, the choice of the correct matching is crucial.

Fig. 4.2 shows an example matching result, by preferring the best correlation of the time difference of the S2-S1 signal pair (drift time) and the vertical position information of the S1 signal (S1TTR). The resulting scattering multiplicity (left) reveals that this matching criterion results in ∼20% of S1 signals remaining unmatched. The matching efficiency as a function of the energy S1L (right) indicates that most of these unmatched S1 signals reside at low energies (below 100 p.e.). This is generally undesirable, as the low S1L region is where the majority of WIMPs is hypothesized, but it is also where the signal parameters S1F90 and S1TTR become less reliable due to the low number of photoelectrons comprising the signal.

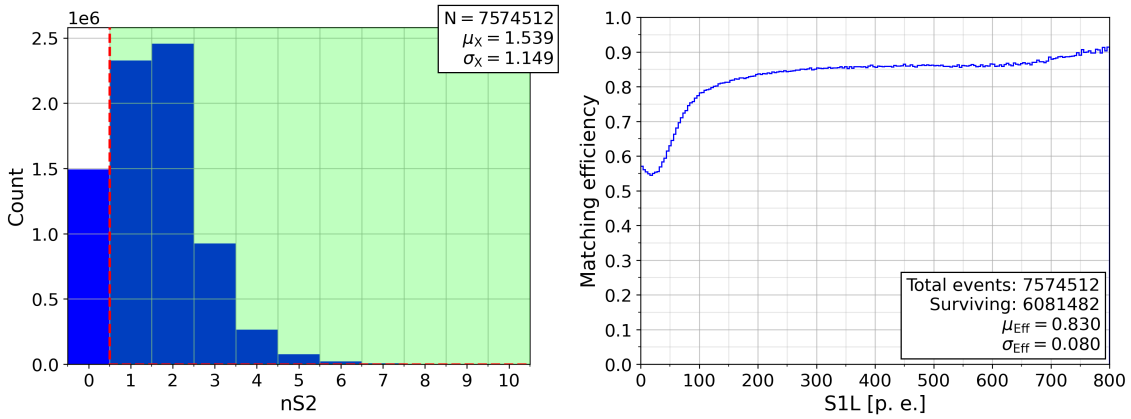

Figure 4.2: Example matching result by preferring the S2-S1 pair with the best correlation of the S2-S1 time difference and the S1 vertical position (S1TTR). While this matching criterion was not employed in the final analysis, it helped drive the process of finding the right matching and remained part of the analysis in a different form (see Section 4.1.1). *Left*: The scattering multiplicity of this matching. It results in ∼20% of S1 signals remaining unmatched. *Right*: The matching efficiency as a function of the energy S1L pinpoints the majority of unmatched S1 signals to very low energies (below 100 p.e.).

The corresponding positional distribution of the matched events is shown in Fig. 4.3. The matching results in a linear distribution of events along the vertical axis of the detector (Fig. 4.3, left). As expected, the number of events is increasing in the radial projection until the boundary of the inner LAr vessel at 36.75 cm (Fig. 4.3, right). In the case of multiple S2 signals being matched to the same S1 signal, the corresponding S2 signal with the maximal fraction of the total light represents the corresponding event (e.g. S2maxL, S2maxDT, ...).

While this choice of ranking of the potential S1-S2 pairs did not end up being the final matching, it was a vital first step in the iterative process of finding the best matching and helped to uncover potential problems. The criterion of a good correlation of the S2 drift time and the S1TTR was later disfavored, but it is physically well motivated and remains a part of the later analysis (see Section 4.1.1). Nevertheless, it is essential to understand that the choice of a matching is similar to the choice of any other data analysis cut. It also has to be fine-tuned and its performance needs to be optimized on the signal-like events.

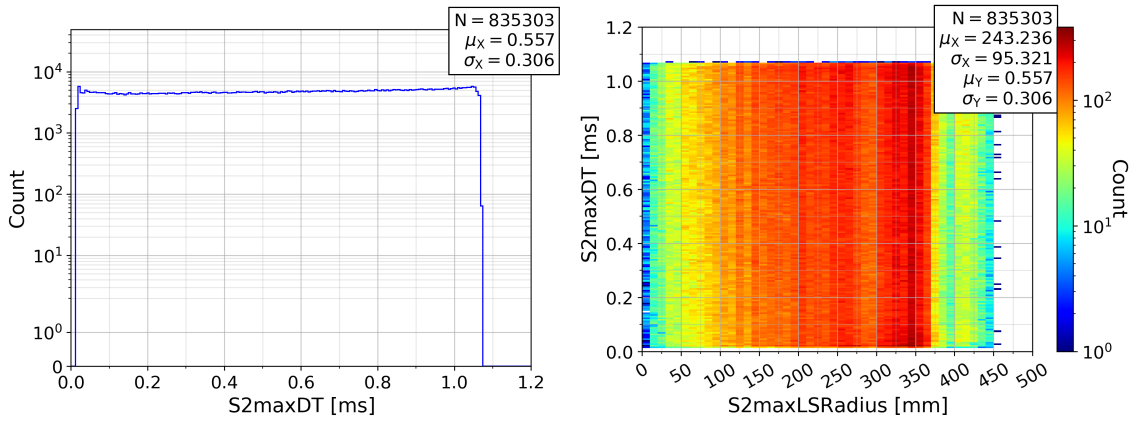

Figure 4.3: Example matching result by preferring the S2-S1 pair with the best correlation of the S2-S1 time difference and the S1 vertical position (S1TTR). The information of the matched S2 signal with maximal light allows to analyze into the positional distribution of the reconstructed events. The vertical distribution (left) along the drift time (S2maxDT) of the S2 signals appears linear. As expected, the number of events increases radially until the boundary of the inner LAr vessel at 36.75 cm.

Taking a step back: the previously presented matching algorithm (Fig. 4.1) is a tool to carry out the matching procedure. It essentially reduces the question of the matching to a choice of the likelihood function that assigns the rankings. During the analysis process of the ArDM data, multiple likelihoods have been investigated. The first likelihood was based solely on the correlation of the TTR of the S1 signal and the drift time of the potential S1-S2 pair (as in Figs. 4.2 and 4.3). Due to the detection of the Charge-up (see Section 4.2) the likelihood had to be adapted to maximizing the S2/S1. However, the change of the matching algorithm and the applied Charge-up correction (see Section 4.2.1) led to a group of pile-up events which disturbed the signal region. Consequently, the ArDM collaboration decided to carry out the baseline WIMP search with only the cleanest matches. Specifically, for the baseline WIMP search, only S1-S2 pairings with no competition are allowed, at the cost of a significant loss in efficiency. This will be presented and discussed further in Section 4.2.1. One should note, that accepting only competition-less S1-S2 pairs essentially removes the necessity of any matching likelihood, as the competition-less S1-S2 pairs are those which are matched together independently of the chosen likelihood.

4.1.1 Matching quality variable

As discussed in the matching working principle (Section 4.1), the correlation of the TTR of the S1 signal and the drift time of the ionization charges resulting in the S2 signal is physically motivated. In the case of a true single-scattering event, the TTR of the S1 signal indicates the vertical location of the scattering. The resulting free electrons drift upwards with a constant drift velocity v_{Drift}. After drifting for the time S2DT, the electrons reach the liquid-gas interface, and, upon entering the gaseous phase produce the S2 signal. Therefore, the product of the constant drift velocity v_{Drift} and the "S2 drift time" S2DT should correlate strongly with the vertical position of the initial scattering, and thus the S1TTR. In the case of multiple-scatterers, this correlation is weakened. As the different S1 signals originating from the same initial particle are not resolvable in time, the measured

S1 signal is the sum of the underlying individual S1 signals. This renders the S1TTR a weighted average in the case of multiple-scatterers. The distribution of the S1TTR versus the S2maxDT can be seen in Fig. 4.4 (left). As expected, a strong correlation is visible. As mentioned earlier, the S2 signal with maximal light is chosen to represent the events' S2 variable values in the case of multiple-scatterers, hence the sub-term "max" in the S2 drift time S2maxDT.

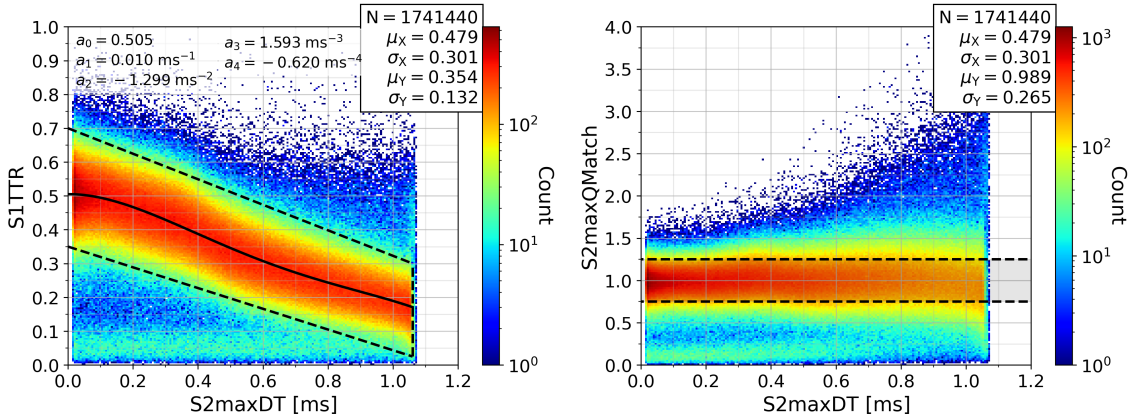

Figure 4.4: *Left*: Example distribution of the S1TTR versus the matched maximum S2 signal drift time. A strong correlation is visible as physically motivated by the constant drift velocity v_{Drift}. *Right*: A linearization of this correlation (via Eq. 4.1) allows to transform the S1TTR and the S2maxDT into a measure of the strength of this correlation via the new event variable S2maxQMatch. For single-scattering events with strongly correlated S1TTR and S2maxDT, the S2maxQMatch is expected to be distributed around unity.

As the correlation of the S1TTR and the S2maxDT is especially good for single-scatterers but worse for multiple-scatterers, it is desirable to transform the S1TTR into an event variable that is approximately independent of the drift time, such that it can serve as a measure of the correlation strength in the data analysis. This is implemented with the new matching quality event variable, called S2maxQMatch, as

$$\mathrm{S2maxQMatch} = \frac{\mathrm{S1TTR}}{\sum_{i=0}^{4} a_i \cdot (\mathrm{S2maxDT})^i} \ , \tag{4.1}$$

where the correlation of the S1TTR and the S2maxDT is fitted well with a polynomial of fourth order. The resulting distribution of the S2maxQMatch event variable can be seen in Fig. 4.4 (right) and in Fig. 4.5. By design, it is heavily centered around unity and larger than zero everywhere. The small bump around 0.1 stems from a small background at very low S1TTR which is also visible in Fig. 4.4. A sensible cut to remove events with bad correlation is at the points of half maximum. In this case a symmetrical cut of $0.75 \leq \mathrm{S2maxQMatch} \leq 1.25$ would cut away $\sim$20% of events. Adding the S2maxQMatch as an event variable allows to incorporate the matching criterion of the S1TTR-S2maxDT correlation in a later data analysis cut even though the initial matching criterion is changed.

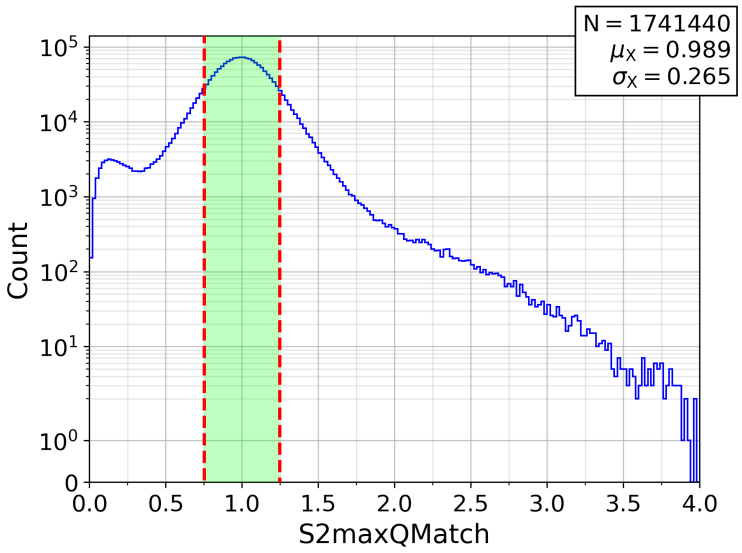

Figure 4.5: The distribution of the event variable S2maxQMatch. By design, the S2maxQMatch serves as a measure of the S1TTR-S2maxDT correlation with the strongest correlation at unity. The indicated symmetrical cut $0.75 \leq$ S2maxQMatch ≤ 1.25 results in a cut efficiency of $\sim$80%.

4.1.2 Light attenuation of S2s due to their drift time

While the free electrons drift upwards in the LAr due to the applied drift field, there is a probability of them recombining in the LAr. This reduces the number of drifting electrons that reach the gaseous phase, hence reducing the resulting S2L. Consequently, the S2L has to be corrected for this attenuation. As the drift velocity of the free electrons is constant ($v_{\text{Drift}} = 1.065$ mm/μs, see Fig. 4.12), the probability of absorption is also expected to be constant. Consequently, the number of drifting electrons is expected to decrease exponentially as a function of the drifted distance or the drifting time. This is in essence equivalent to the Lambert-Beer law of physical optics, even though it isn't the light that is absorbed, but rather the electrons. However, as the emitted S2L is proportional to the number of drifting electrons reaching the gaseous phase, one can write the attenuation of the S2L as a function of the drift time (S2DT) as

$$\text{S2L}_{\text{measured}} = \text{S2L}_0 \cdot e^{-\frac{\text{S2DT}}{\tau_{\text{Drift}}}} \qquad \rightarrow \qquad \text{S2L}_0 = \text{S2L}_{\text{measured}} \cdot e^{\frac{\text{S2DT}}{\tau_{\text{Drift}}}}. \tag{4.2}$$

In order to properly compare the S2L of events originating from all vertical regions of the detector, one seeks to recreate the non-attenuated value and restore S2L_0 (which will be referred to just as S2L). Naturally, the drift time S2DT is dependent on a matching and as such, the evaluation of the drift time attenuation coefficient τ_{Drift} has to be carried out on the cleanest matches. Fig. 4.6 (left) shows a calibration fit on the competition-less matches resulting in $\tau_{\text{Drift}} = 0.585$ ms or $l_{\text{Drift}} = 623.02$ mm respectively; signifying, that after a drift time of 585 μs or a drift length of 623.025 mm an average of 1/e or 36.8% of the drifting electrons are absorbed. Fig. 4.6 (right) shows the same distribution with the drift-time correction applied.

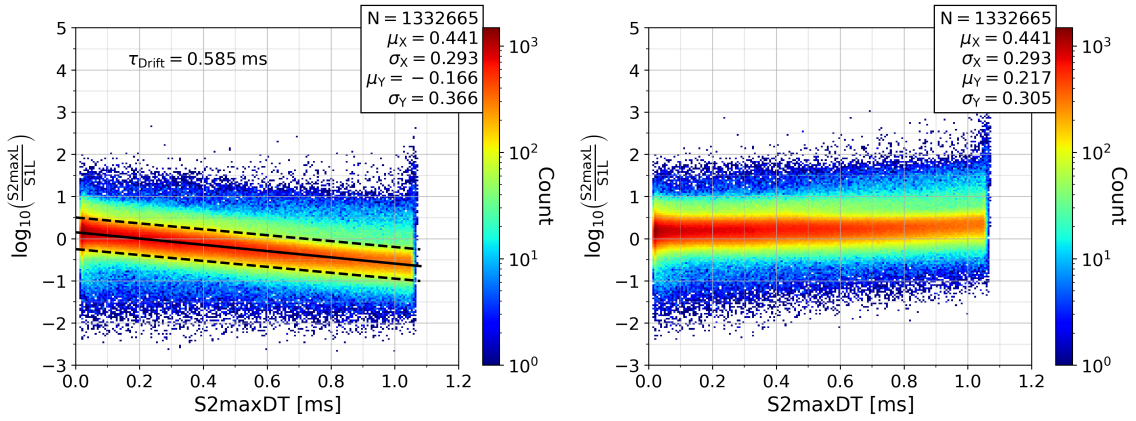

Figure 4.6: The absorption of the drifting electrons inside the LAr follows an exponential law and needs to be corrected for. As the emitted S2L is proportional to the number of drifting electrons reaching the gaseous phase, the drift time attenuation coefficient τ_{Drift} can be calibrated based on the S2L of the competition-less matches (left). After applying the drift-time correction, the distribution of the S2L is approximately independent of the vertical position (right).

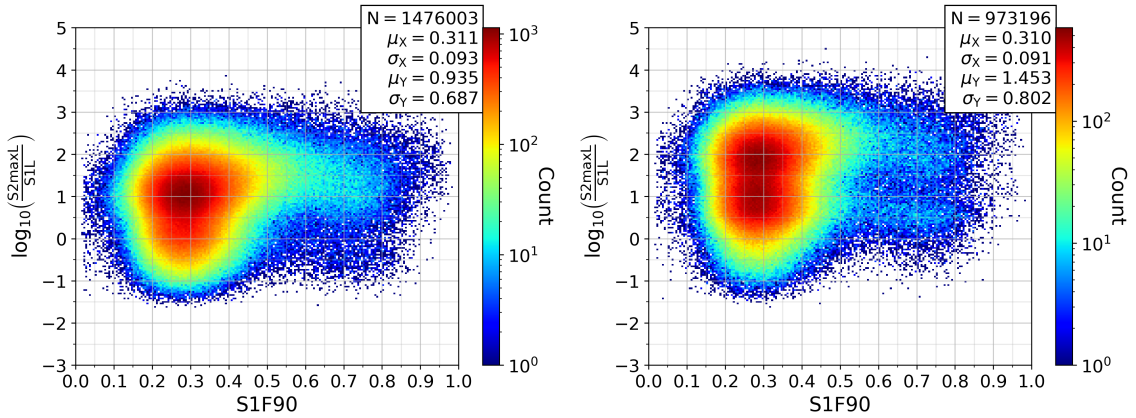

Figure 4.7: Two sample Neutron calibration data sets (cut to S1L $\leq$ 800 p.e.) collected at different times. While the S1F90 perspective is behaving as expected, the distribution of the log(S2maxL/S1L) is drastically changing over time. The left-hand data set was taken around July, 10th 2019, while the right-hand data set was collected around August, 13th 2019.

Having corrected the S2L for the drift time attenuation allows a first look at the distribution of the main ER/NR discrimination variables: the log(S2maxL/S1L) versus the S1F90. In general, ERs are expected to be distributed around an S1F90 of ∼0.30, whereas the NRs are expected to be distributed around an S1F90 of ∼0.70. At lower energies, the S1F90 becomes a less reliable estimator and the ER and NR distributions start to merge. While the distribution of the log(S2maxL/S1L) is dependent on the energy, the detector and the experimental conditions, the general expectation is for the NRs to have a lower average log(S2maxL/S1L) than the ERs. Fig. 4.7 shows the distribution of the log(S2maxL/S1L) versus the S1F90 of two sample Neutron calibration data sets (cut to

S1L $\leq$ 800 p.e.) collected at different times. It becomes immediately apparent that, while the S1F90 perspective is behaving as expected, the log(S2maxL/S1L) distribution is not clearly distinguishing between an ER and an NR population, and also changes over time (left to right). In the process of the ArDM data analysis, this observation led to the discovery of the Charge-up process. The effects of this Charge-up process and its correction are discussed in detail in the upcoming Section 4.2.

4.2 Charge-up of S2s - Discovery

The correction of the S2L attenuation allowed a first look at the log(S2maxL/S1L) versus the S1F90; the main ER/NR discrimination variables (see Section 4.1.2, Fig. 4.7). It became apparent that the log(S2maxL/S1L) distribution is drastically changing over time. Fig. 4.8 shows the same two sample Neutron data sets (as Fig. 4.7) also with events cut to S1L below 800 p.e. and additionally cut to reconstructed single-scatterers only. While one data set was collected around July, 10th 2019 (left), the other data set was taken around August, 13th 2019 (right). While the later Neutron data set approaches more the expectation of two distinct populations with a visible NR accumulation at an S1F90 of $\sim$0.70 and a log(S2maxL/S1L) of $\sim$0.5, the earlier Neutron data set doesn't have such a visibly clear NR population which should be at lower average log(S2maxL/S1L) compared to the ER population. In fact, there is quite a populated region with an S1F90 of $\sim$0.70 and a log(S2maxL/S1L) of $\sim$1.0 and the ER region of an S1F90 of $\sim$0.30 appears stretched in terms of the log(S2maxL/S1L). Also noteworthy is the change of the average scale from μ(log(S2maxL/S1L)) $\sim$0.444 to μ(log(S2maxL/S1L)) $\sim$0.788.

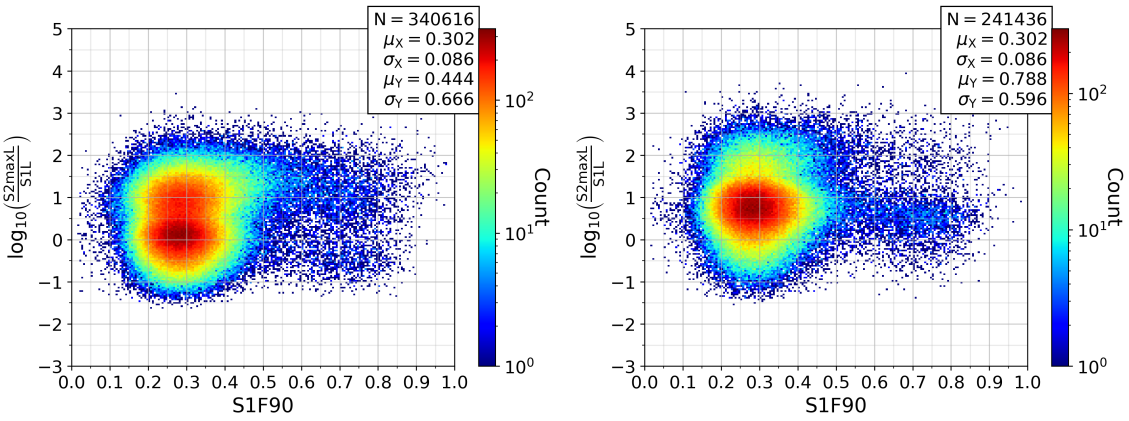

Figure 4.8: The same two sample Neutron calibration data sets as in Fig. 4.7 (cut to reconstructed single-scatterers with S1L $\leq$ 800 p.e.) collected at different times. The left-hand-side data set was taken around July, 10th 2019, while the right-hand-side data set was collected around August, 13th 2019. The log(S2maxL/S1L) distribution changes drastically over time in scale and range.

This discovery necessitates a detailed study of the time dependence of this phenomenon. And indeed, the progression of the average log(S2maxL/S1L) versus time (Fig. 4.9, left) reveals a repeated increase, which resembles the charging of a capacitor, followed by sudden resets. The resets coincide with an SP period (green), an SPD period (blue) and an HV failure (purple) in time. This motivates a connection of this "Charging" behavior to

the status of the extraction field. Explicitly, every time the extraction field was turned off in-between DP periods, the subsequent average log(S2maxL/S1L) reset to a starting value. The strength of the extraction field directly relates to the energy of the drifting electrons entering the gaseous phase and hence the amplitude of the average S2L. The time progression of the average S1L, S2L and S2maxL in Fig. 4.9 (right) supports this hypothesis. Any change of the average S1L is insignificant on the scale that one can observe in terms of the progression of the average S2L and S2maxL. E.g. the average S2L initially started at ∼1500 p.e. and increased up to ∼9000-10000 p.e. which is an increase by ∼500-567%.

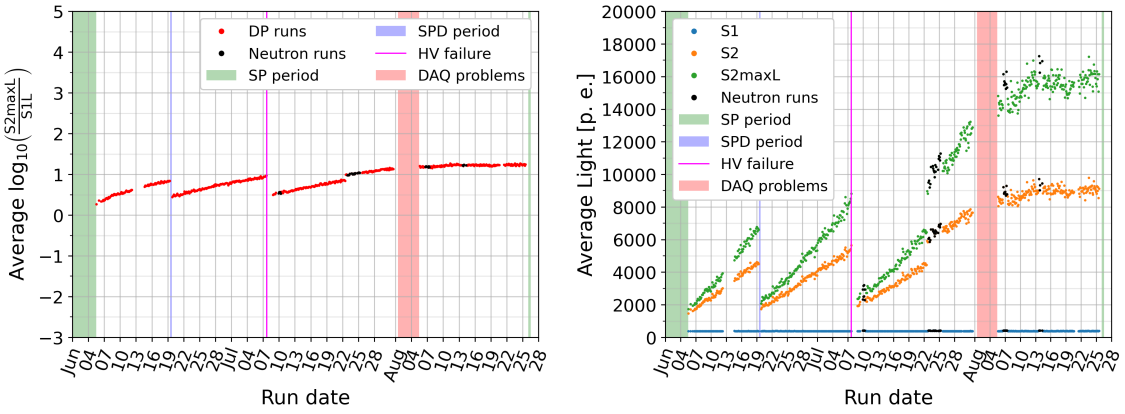

Figure 4.9: The time progression of the average log(S2maxL/S1L) (left) reveals a repeated reset-and-recharging behavior with the resets coinciding with points in time where the extraction field was turned off. The progression of the individual S1L, S2L and S2maxL (right) confirms that only the S2 signals are significantly affected which strengthens the hypothesized connection to the extraction field.

An increase in the strength of the applied extraction field would not only amplify the average S2L but also allow more potential S2 signals to cross the detection threshold. Again, an aspect which is supported by the data as the progression of the fraction of signal types is shown in Fig. 4.10 (left). The fraction of reconstructed signals which are S2s (orange) is increasing in similar recharge-and-reset cycles which is compensated by the fraction of S1 signals (blue). Naturally, the fractions of matched and unmatched S1 signals both decrease as a consequence. However, the fraction of matched S1 signals (green) approaches more and more the fraction of S1 signals (blue) with recharging. This concludes that the percentage of S1 signals which are matched increases within a charging cycle simply due to the presence of more S2 signals as matching candidates. To, again, untie the influence on S1 and S2 signals separately, Fig. 4.10 (right) shows the progression of the signal rates. The S1 signal rate decreases constantly by a total of ∼10% over the span of nearly three months. This is deemed unconnected to the Charge-up behavior as there are no resets and is likely caused by the degradation of the PMTs over time. Conversely, the rate of S2 signals intensifies from ∼650 s^{-1} to ∼1300-1400 s^{-1} which corresponds to an increase of ∼100-115% again supporting the extraction field hypothesis. In principle, this repeated increase and reset of the S2 signal rate was already visible in the trigger rate of the DAQ (see Section 3.1.1, Fig. 3.3, right) which showed a similar behavior. However, only the detailed study of the reconstructed events could reveal the underlying aspects.

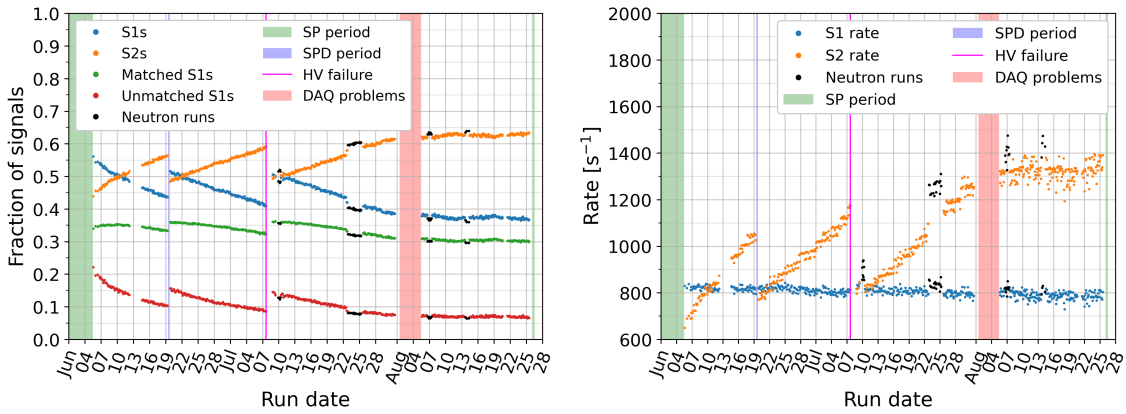

Figure 4.10: The time progression of the fractions of signal types (left) shows again a repeated reset-and-recharging behavior for the S2s which is compensated by the fraction of S1s. With recharging, the increased number of S2s leads to a smaller percentage of S1s ending up unmatched. The progression of the signal rates (right) emphasizes that only the S2 rate is affected by the Charge-up behavior, which further supports the extraction field hypothesis.

The immediate follow-up concern is whether the influence of this Charge-up phenomenon is positionally dependent or uniform within the detector. Fig. 4.11 (left) shows the effect on the horizontal positions. While the average x and y positions (blue and orange) are stable over time, the S2L-weighted average x and y positions (green and red) are affected by the Charge-up. This eludes to the fact that the positional distribution of events is not changing on average, but the effect of the Charge-up behavior on the S2L is positionally dependent. The vertical perspective (Fig. 4.11, right) reveals an effect on the average drift time (yellow). This can be understood in the manner that the detection efficiency of S2 signals is in principle smaller, the longer the free electrons have to drift to reach the gaseous phase, due to the constant absorption probability of the electrons covered in Section 4.1.2. Independently, as previously discussed, the stronger extraction field shifts more of the S2L distribution over the detection threshold. The combination of these two aspects results in this relatively lower threshold being more effective for events of smaller drift time as these events, that survive the attenuation, profit more from the shift of the S2L distribution relative to the detection threshold.

While it now is evident that the Charge-up effect is caused by an uncontrolled change of the extraction field strength, it is pivotal to check any potential influence on the drift field strength. The stability of the drift field strength is rectified by the consistency of the maximal drift time (Fig. 4.11, right, dark blue). This signifies that the drift velocity of the free electrons remained constant during the data-taking and consequently also the drift field strength. The observed maximal drift time of 1044.2 ± 3.4 μs combined with the length of the drift field chamber of 1112.0 mm results in a drift velocity of $v_{\text{Drift}} = 1.065 \pm 0.003$ mm/μs. A comparison to the published literature on the electron drift velocity in LAr [72] allows to gauge the effective drift field that is felt by the free electrons to $F_{\text{Drift}} = 152.1 \pm 1.4$ V/cm. This is, as a matter of fact, significantly less than the initially desired 230 V/cm. This reduces the effect of the recombination quenching (see Section 2.2.1; Eqs. 2.15 to 2.18 and Fig. 2.5, left), which has to be incorporated into the expected WIMP signal in Chapter 6. Additionally, it also reduced the drift velocity to the deduced 1.065 ± 0.003 mm/μs (instead

of $v_{\text{Drift}} = 1.25$ mm/μs at $F_{\text{Drift}} = 230$ V/cm). This in turn led to the observed lower drift attenuation length coefficient of $l_{\text{Drift}} = 623.02$ mm which results in a larger S2 signal loss, especially from the bottom part of the detector. However, this difference of the desired and the effective drift field is stable and not caused by the Charge-up effect.

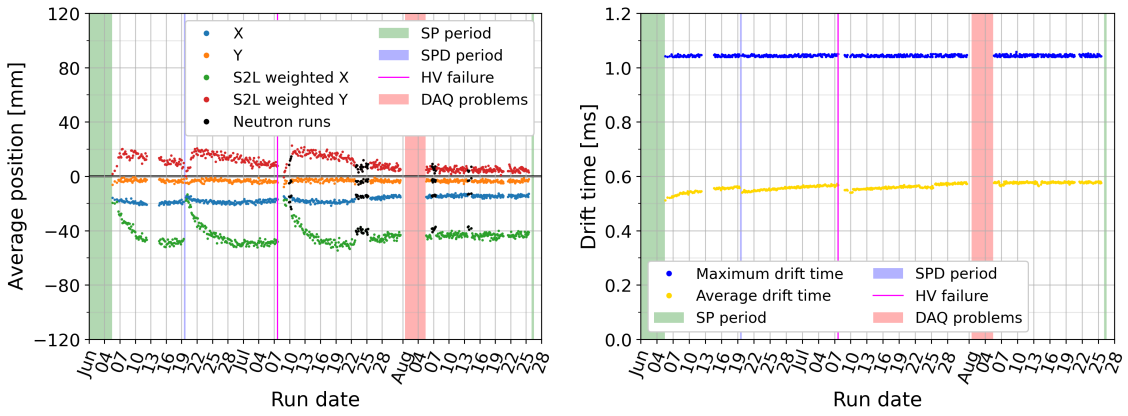

Figure 4.11: *Left*: The time progression of the average x and y as well as the average S2L-weighted x and y positions imply that the Charge-up effect had no influence on the average positional distributions but its effect on the average S2L is significantly dependent on the position. *Right*: The influence on the average drift time is a combination of the free electrons with shorter drift time having a higher probability to survive attenuation and the shift of more of the S2L distribution over the detection threshold due to the Charge-up. The stability of the maximal drift time ensures the consistency of the drift field strength and allows to gauge the electron drift velocity.

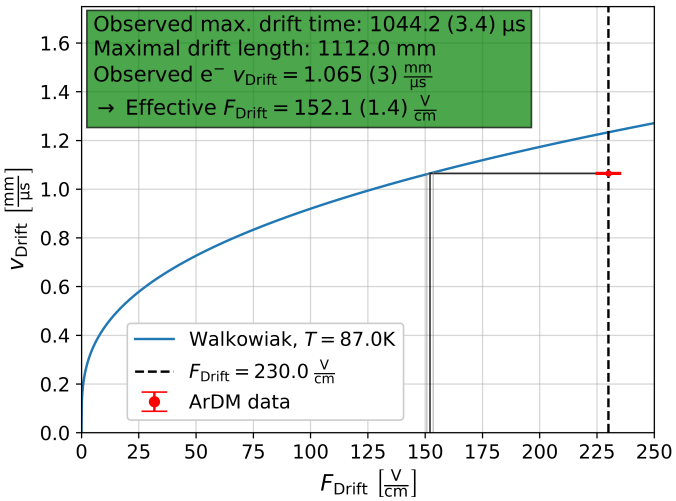

Figure 4.12: The observed maximal drift time (Fig. 4.11, right, dark blue) allows to deduce the electron drift velocity of 1.065 ± 0.003 m/μs. A comparison to the literature (Walkowiak [72]) gauges the effective drift field strength felt by the electrons to $F_{\text{Drift}} = 152.1 \pm 1.4$ V/cm instead of the desired 230 V/cm.

While it is now established that the Charge-up effect is caused by an increase in the extraction field strength, which is reset by turning off the extraction grid voltage, the horizontal dependence remains to be analyzed. Fig. 4.13 shows the horizontal positional distribution inside the inner LAr target volume ($R \leq 36.75$ cm) of the average S2maxL of two sample DP data sets taken at different times. The earlier data set (left) was taken on June, 20th 2019, or the start of the second Charge-up cycle, and the later data set (right) was collected on July, 8th 2019, or the end of the second Charge-up cycle. In the earlier data, the average S2maxL is distributed relatively uniform with the exception of a lower S2maxL region at higher reconstructed S2maxLSX which is a signature of the geometry of the detector top. Conversely, in the later data, the average S2maxL has increased drastically towards lower S2maxLSX.

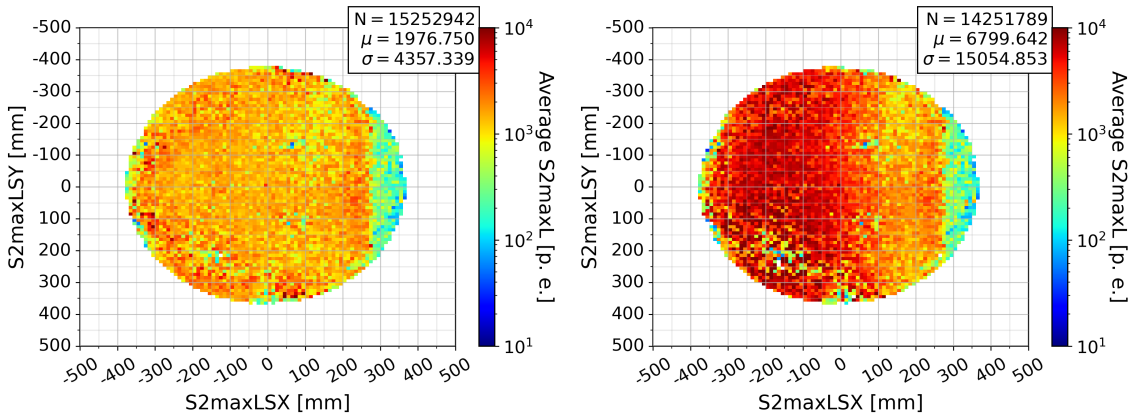

Figure 4.13: The horizontal positional distribution inside the inner LAr target volume ($R \leq 36.75$ cm) of the average S2maxL of two sample DP data sets taken at the start (June, 20th 2019) and at the end (July, 8th 2019) of the second Charge-up cycle. While in the earlier data set (left) the average S2maxL is distributed relatively uniform, the later data set (right) shows a drastic increase towards lower S2maxLSX. The relative differences can be seen in Fig. 4.14.

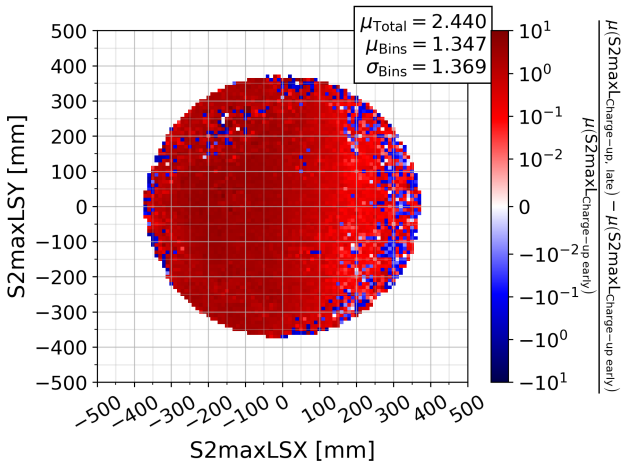

Figure 4.14: The relative differences of the average S2maxL of the two data sets (shown in Fig. 4.13) normalized to the earlier data set. The average of the S2maxL of the entire data set has increased by 244% and it becomes apparent that the Charge-up effect is stronger towards lower S2maxLSX.

Fig. 4.14 shows the respective differences of the average S2maxL from the two data sets normalized to the earlier data set. The statistics box is to be understood in the sense that the average S2maxL of the entire data set has increased by 244% (μ_{Total}) while the bin-wise changes are characterized by their mean μ_{Bins} and standard deviation σ_{Bins}. The relative differences emphasize again the increase of the Charge-up effect with lower S2maxLSX. The likeliest cause is an accumulation of electric charges over time at the top of the extraction grid. This increases the extraction field strength over time until the accumulated charges reach a saturation level. Hence the similarity to a capacitor and the name "Charge-up".

The first visual impression of the positional dependence of the Charge-up effect on the average S2maxL (Fig. 4.13, right) implies a y-symmetry. To test this claim, one can investigate the standard deviations of the S2L-weighted position centers of all runs for various rotation angles as shown in Fig. 4.15. Therein, the x and y coordinate system is rotated by a specific angle before the S2L-weighted centers are calculated. The standard deviations of the x and y centers' time progressions then result in a quantitative measure of the accuracy of the symmetry. E.g. the standard deviations the S2L-weighted x and y centers' time progressions for a rotation angle of 0° (corresponding to the green and red data points in Fig. 4.11, left) result in σ_x and σ_y for 0°. The scan over various angles confirms the visual claim. The standard deviation of the x centers σ_x is greatest and the standard deviations of the y centers σ_y is smallest for the Charge-up symmetry angle of ∼2° relative to the standard detector axes. Reducing the symmetry to 0° is a reasonable assumption which simplifies the correction of the charged-up S2L. The calibration of the correction of the Charge-up influence on the S2L as well as the consequences for the data analysis will be discussed in the upcoming Section 4.2.1.

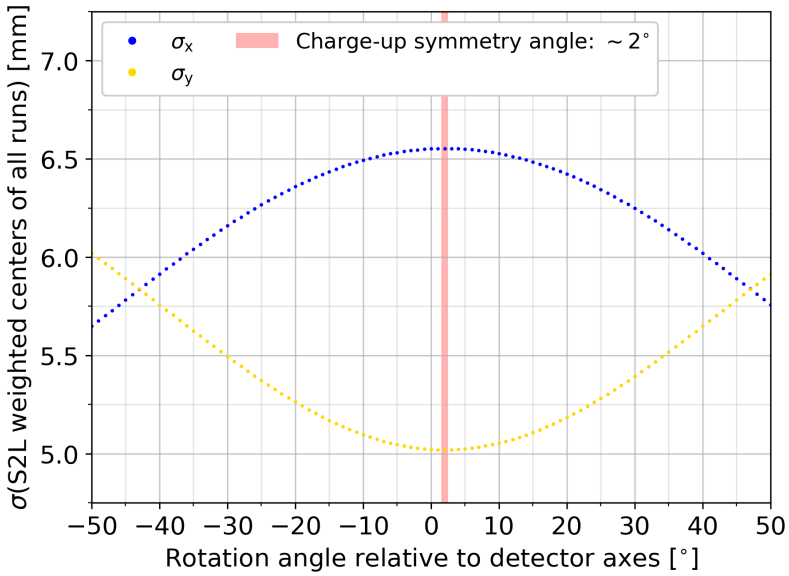

Figure 4.15: The standard deviations of the S2L-weighted x and y centers of all runs for various rotation angles of the coordinate system. E.g. the standard deviations for 0° correspond to the standard deviations of the green and red data points in Fig. 4.11 (left). The maximal σ_x and minimal σ_y at ∼2° quantitatively confirm the almost perfect y-symmetry.

4.2.1 Correction of the Charge-up effect and its implications

Having thoroughly researched the effects of the Charge-up effect the ArDM collaboration remained confident that the S2L could be restored by a calibrated correction in the data analysis process. Such a correction essentially has to cure the positional dependence (see Fig. 4.13) as well as the time dependence (see Fig. 4.9, left). Based on the quantitatively supported assumption of a y-symmetry (see Fig. 4.15), the x-y plane is divided into numbered bins of 10 by 10 cm as visualized in Fig. 4.16.

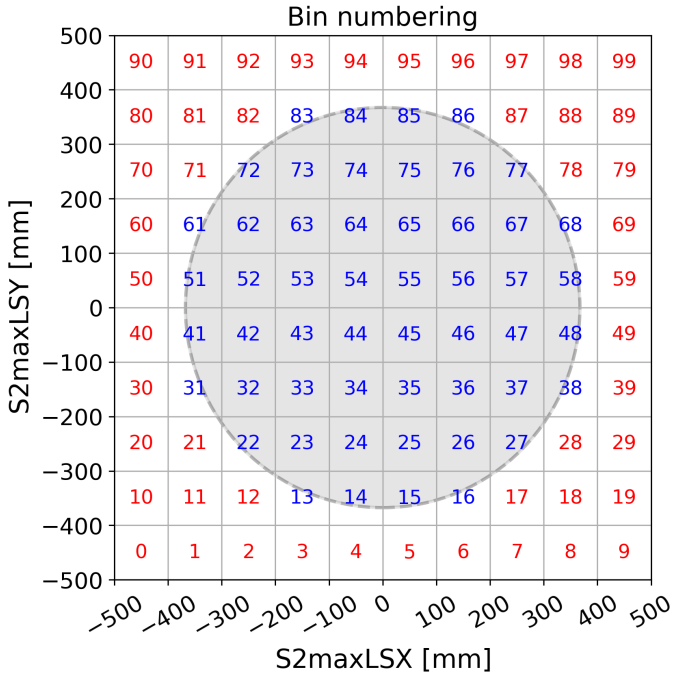

Figure 4.16: The bin numbering for the Charge-up correction to treat the positional dependence. In each 10 by 10 cm bin individually, the S2maxL is cured of its time dependence by a fit of the progression of the log(S2maxL/S1L) versus time according to the Charge-up function in Eq. 4.3. For bins at the edge of the inner LAr target volume (black shaded, $R = 36.75$ cm) only events within the target area are considered.

The bin-wise correction serves to restore the positional independence to a reasonable degree. A too large bin size doesn't catch the positional dependence while a too small bin size results in correcting events based on increasingly low statistics. Multiple bin sizes have been tested and the bin width of 10 cm strikes a good balance between the two effects as the statistics are numerous enough for a stable calibration but also the positional dependence is treated adequately. For every bin, the time-wise correction is then based on a calibration fit of the time progression of the log(S2maxL/S1L) following the Charge-up of a capacitor in an RC circuit as

$$\begin{aligned} \log_{10}\left(\frac{\text{S2maxL}}{\text{S1L}}\right)(t) &= A_\infty\left[1 - e^{-\frac{t-t_{\text{Offset}}}{\tau_{\text{Charge-up}}}}\right], \\ \text{where} \quad A_\infty &= \log_{10}\left(\frac{\text{S2maxL}}{\text{S1L}}\right)(t=\infty)\,. \end{aligned} \tag{4.3}$$

For bins at the edge of the inner LAr target volume (Fig. 4.16, black shaded, $R = 36.75$ cm) only events within the target area are considered. An example fit result can be found in Fig. 4.17 (left) for the bin with number 53. The great agreement of the fit and the data again emphasizes the analogy to a capacitor charging up inside an RC circuit. Therein, the total charge stored at infinite time would equal to $Q_{\text{max}} = CV_{\text{b}}$ or correspondingly to A_∞ inside ArDM as the log(S2maxL/S1L) which is attained after infinite time. The time offset t_{Offset} corresponds to the start time at which the Charge-up process is started. And finally, the time constant (or sometimes called the transient response time) $\tau_{\text{Charge-up}}$ would correspond to RC in an RC circuit and quantify the charging time of the capacitor. Respectively, one, two, three and five $\tau_{\text{Charge-up}}$ periods would correspond to ∼63%, ∼86%, ∼95% and ∼99% of the maximal charge accumulated.

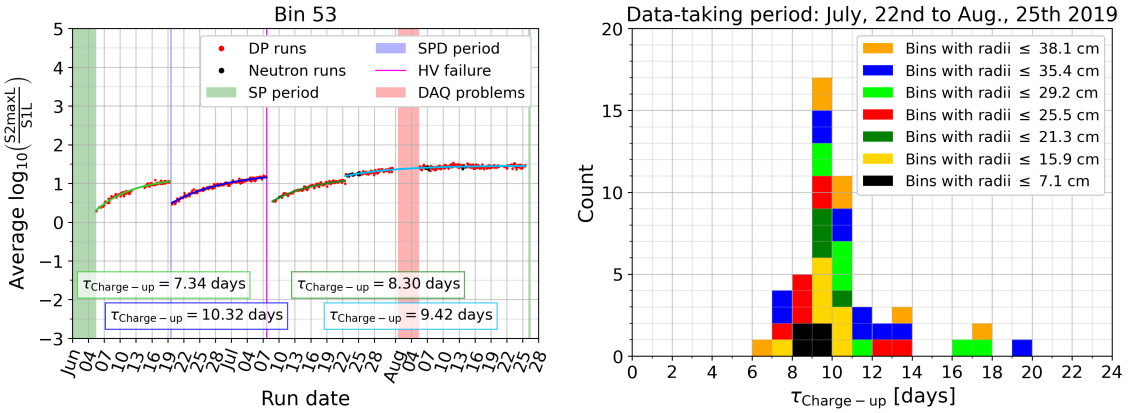

Figure 4.17: *Left*: An example fit result of the Charge-up function (Eq. 4.3) to the time progression of the average log(S2maxL/S1L) for the bin with the number 53 as in Fig. 4.16. *Right*: Resulting distribution of the best fit Charge-up time constant $\tau_{\text{Charge-up}}$ for the last of the four fitting regions. Central bins are centered around $\tau_{\text{Charge-up}} \approx 10$ days meaning that roughly 10 days result in a Charge-up effect of an "effectivity" of ∼63% and a time of roughly a month results in an "effectivity" of ∼95%.

As established, turning off the extraction grid voltage (or altering it by hand) resets (or shifts) the Charge-up effect. The total time period of the data-taking is divided into four separate fitting regions (as showcased in the example bin result in Fig. 4.17, left). The resulting distribution of the best fit $\tau_{\text{Charge-up}}$ is shown in Fig. 4.17 (right) for the fourth and last time period ranging from July, 22nd to August, 25th 2019. Therein, the distribution is colored according to increasing radii of the centers of the correction bins. This emphasizes that the correction fits lead to consistent results in almost all bins with only a minority of non-central bins having inconsistent results. The majority of the best fit results are distributed around a center of $\tau_{\text{Charge-up}} \approx 10$ days implying that the Charge-up effect reaches ∼63% and ∼95% effectivity after roughly 10 days and a month respectively. The corresponding distributions of the amplitude parameter A_∞ and the starting time t_{Offset} are shown in Fig. 4.18 (left and right) for the fourth fitting time period. Again, the majority of the bins are consistent in terms of the best fit parameters with only a handful of outliers. These outliers aren't considered problematic. Specifically, there is also an area in the horizontal plane at the edge of the inner LAr volume at high S2maxLSX (see Fig. 4.14) where it appears that the Charge-up has little effect. Consequently, the fits in those bins

do not necessarily have to be consistent. The amplitude A_∞ as such is difficult to interpret or judge given the context that the Charge-up effect has messed significantly with the S2L scale and the gauge to a specific voltage or field strength is uncertain. However, the exact gauge of the log(S2maxL/S1L) scale is not important as long as it is not compared to the scales of other experiments. The crucial point is to conserve the relative differences of the ER and NR distributions as well as possible. Finally, the distribution of the starting times t_{Offset} acts as another confirmation of the link of Charge-up resets to points in time where the extraction field was turned off. Specifically, in this case, there is a strong trend of the starting times to be accumulated around the indicated HV failure which also enforced a turned off extraction field. This concludes the detailed look into the calibration of the Charge-up correction.

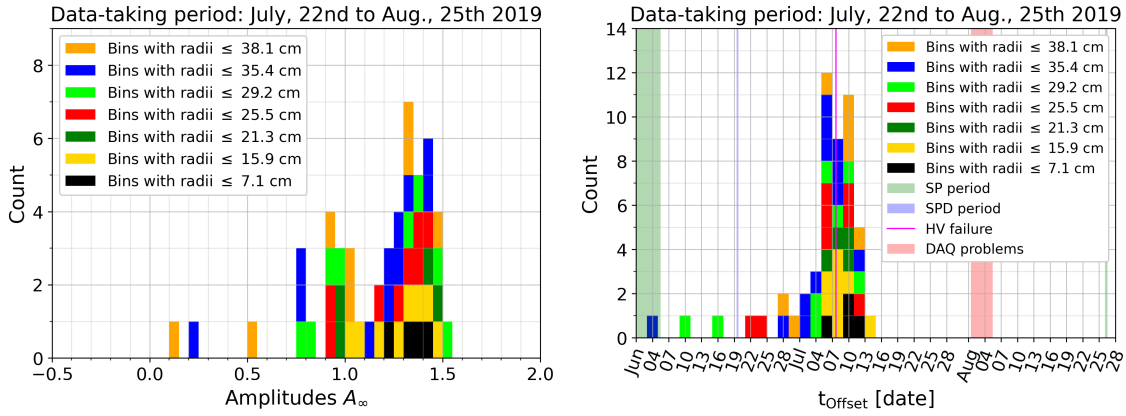

Figure 4.18: Resulting distributions of the best fit amplitudes A_∞ (left) and the starting times t_{Offset} (right) for the last of the four fitting time periods. The amplitudes A_∞ as such, shouldn't be interpreted intensively as the strength of the Charge-up effect is not uniform within the horizontal plane. The starting times t_{Offset} are consistent with the point in time where an HV failure occurred. Similar consistency for the other three fitting periods with the HV failure, the SP and the SPD period again support the causality of a turned off extraction field with the reset of the Charge-up.

Naturally, it is vital to check the outcome of the applied correction on the S2L variable over time and position. Firstly, Fig. 4.19 (left) shows the time progression of the corrected average light (compare to the uncorrected light in Fig. 4.9, right). As discussed, the Charge-up effect renders the average log(S2maxL/S1L) scale somewhat arbitrary. For this correction, the average log(S2maxL/S1L) has been rescaled to a value of 0.3. As designed, the average S2L is now stable within 10% over the duration of the entire data-taking time. Secondly, Fig. 4.19 (right) presents the progression of the corrected S2L-weighted x and y centers (compare to the uncorrected S2L-weighted x and y centers in Fig. 4.11, left). As desired, the S2L-weighted x and y centers are, to a high degree, congruent with the average x and y positions meaning that the bin-wise correction successfully removed the majority of the positional dependence of the Charge-up effect on the S2L.

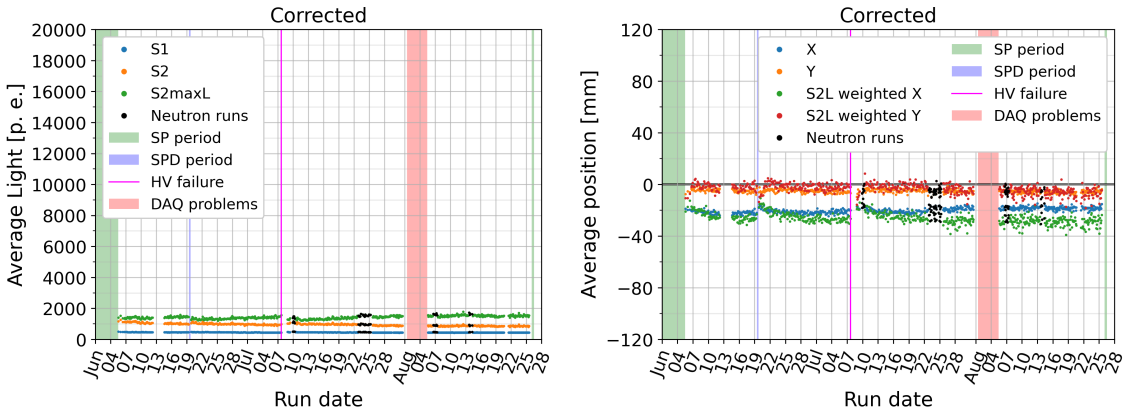

Figure 4.19: *Left*: The progression of the average light is stable within 10% after the Charge-up correction (compare to the uncorrected light in Fig. 4.9, right). *Right*: The progression of the S2L-weighted x and y centers and the x and y averages are, to a high degree, congruent (compare to the uncorrected S2L-weighted x and y centers in Fig. 4.11, left).

Thirdly, Fig. 4.20 showcases the averages of the corrected S2maxL versus the horizontal position of two sample DP data sets collected at different times. The earlier data set (left) corresponds to the same data, taken on June, 20th 2019, as shown uncorrected in Fig. 4.13 (left). The later data set (right) corresponds to the same data, collected on July, 8th 2019, as shown uncorrected in Fig. 4.13 (right). Visually, the two data sets now are much more conform with each other and show no significant positional dependence anymore. Quantitatively, while the uncorrected data sets incorporated an increase of the total average S2maxL of 244%, the corrected data sets result in a relative increase of only 5.2%. Moreover, the standard deviation of the positional averages is well conserved over both data sets.

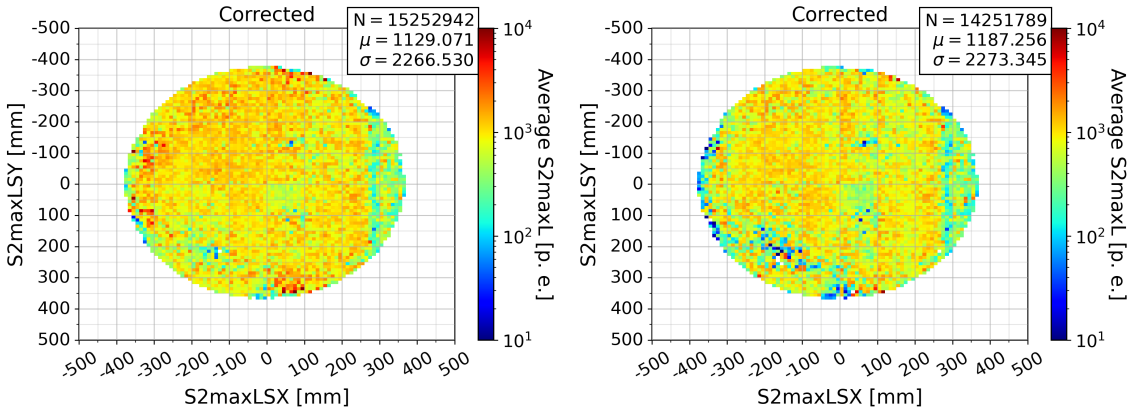

Figure 4.20: The horizontal positional distribution of the corrected average S2maxL of two sample DP data sets taken at the start (June, 20th 2019) and at the end (July, 8th 2019) of the second Charge-up cycle. The uncorrected data sets (see Fig. 4.13) incorporated an increase of 244% in the total average of the S2maxL between the two data sets. The corrected data sets now show no significant sign of positional dependence. The total average S2maxL only increases by 5.2% in-between the two data sets and the standard deviation of the averages is well conserved.

Fourthly, the relative differences of the positional averages of the log(S2maxL/S1L) of both data sets (normalized to the earlier data set) are shown in Fig. 4.21 before (left) and after (right) the application of the Charge-up correction. While, in the uncorrected data, the total average of the log(S2maxL/S1L) increased by 256% with a strong correlation to lower S2maxLSX, the increase accounts to only 7.6% after the application of the Charge-up correction. The bin-wise average change is a decrease of 1.2% further supporting the successful removal of the positional dependence. This concludes the positive test of the immediate outcome of the Charge-up correction.

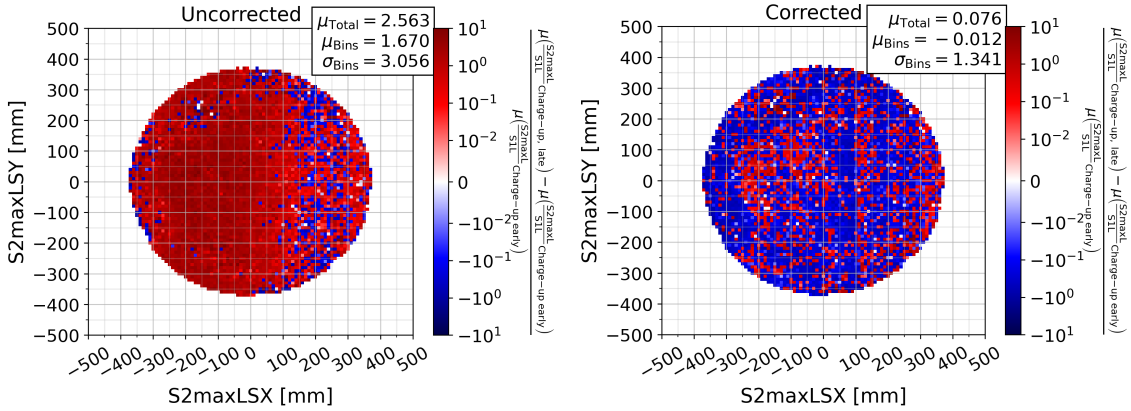

Figure 4.21: The relative differences of the average log(S2maxL/S1L) of the two data sets (collected on June, 20th and July, 8th 2019) normalized to the earlier data set. In the uncorrected data (left), the total average log(S2maxL/S1L) is increasing by 256% in-between the two data sets with a strong trend towards lower S2maxLSX. In the corrected data (right), the total average log(S2maxL/S1L) is increasing by only 7.6% in-between the two data sets, thereby underlining the successful Charge-up correction.

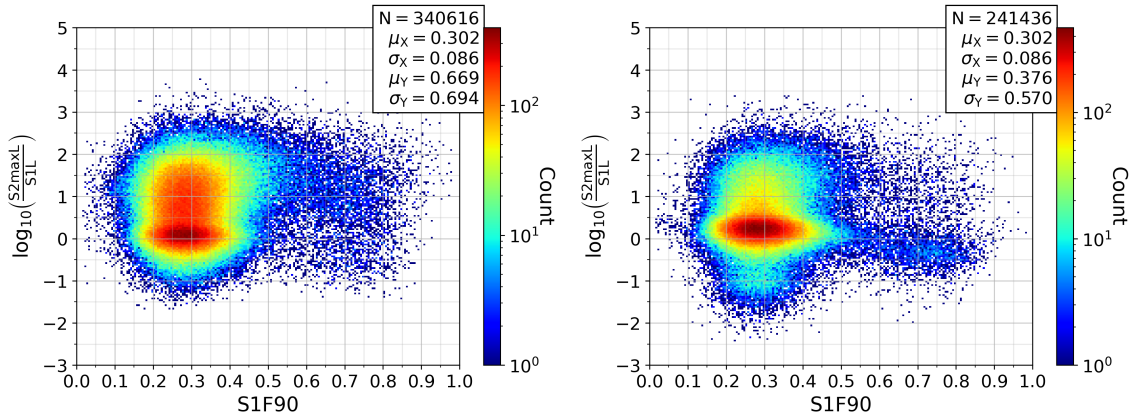

Figure 4.22: The same two sample Neutron calibration data sets as in Fig. 4.8 (cut to reconstructed single-scatterers with S1L $\leq$ 800 p.e.) collected around July, 10th (left) and August, 13th (right) 2019. There still is a large spread on the log(S2maxL/S1L) scale and a group of events at high log(S2maxL/S1L) values which are dubbed "pile-up" events.

Finally, having applied the Charge-up correction, it is time to revisit the main variables

of interest. The distribution of the corrected log(S2maxL/S1L) versus the S1F90 is shown in Fig. 4.22. This corresponds to the same two Neutron data sets presented earlier in Fig. 4.8 collected on July, 10th and August, 13th 2019. While the later data set (right) now is much closer to the expectation of two separate ER and NR populations, the spread on the log(S2maxL/S1L) scale is still very large. This feature is even stronger in the earlier data set (left) where the ER population at an S1F90 of ∼0.3 is dramatically spread over a large range of the log(S2maxL/S1L). Most of the potential NR events at an S1F90 of ∼0.7 appear at larger log(S2maxL/S1L) than the maxima of the ER population at a log(S2maxL/S1L) of ∼0.1.

At that point in time, the ArDM collaboration changed the matching likelihood (see the matching algorithm in Section 4.1) to favoring S1-S2 pairs of maximal log(S2maxL/S1L) due to the discovery of the Charge-up effect. This decision aimed at avoiding potential false signals at low log(S2maxL/S1L) as, any competing S2 of higher light would always be preferred. In this context, it became clear that those events at unreasonably high log(S2maxL/S1L) could very well be "pile-up" events. Or explicitly, matching scenarios where the competing S2 signals indeed drive up the resulting log(S2maxL/S1L) due to the matching likelihood preferring the higher S2L. Essentially, a solution had to be found which distinguishes these pile-up events but keeps the risk of false signals at a minimum. Eventually, this meant falling back to the safe option of competition-less matches (as mentioned in Section 4.1), where the S1 and the S2 signals only have one potential partner of the respective other species within the allowed time window. Fig. 4.23 shows the distributions of the matching scenarios of low energy (S1L $\leq$ 100 p.e.) events. Explicitly, the nS2perS1 variable (left) is simply the count of the number of S2 signals which are available in the allowed matching time window. For these low-energy DP events, this accounts to 33.6% of S1 signals having no potential S2 candidates within the allowed matching time window. Inevitably, these events become unmatched and do lower the data analysis cut efficiency. Next, 22.3% of S1 signals have exactly one potential S2 signal within the allotted matching window making those the unequivocal pairings from the perspective of the S1 signals. The remaining 44.1% concern cases where the S1 signal sees at least two potential S2 signals to be matched to. The naive Poisson process expectation for nS2perS1 given in Section 4.1 of 17.3% (nS2perS2 = 0), 30.4% (nS2perS2 = 1) and 52.3% (nS2perS2 $\geq$ 2)) is quite a bit off as expected. However, it has to be taken into account, that the detection efficiency for these lower-energy S2 signals is also lower and longer drift times might completely attenuate them. Consequently, a shift of the nS2perS1 distribution to lower values is to be expected for low-energy events.

Fig. 4.23 (right) shows the S2 signal perspective of all S2maxL signals matched to the low-energy S1 signals. Shown are only the corresponding S2 signals of maximal light to be conform with the S1 perspective (left) and avoid double-counting. This has to be understood in the sense that, out of all events with an S1 signal matched to at least one S2 signal, 23.7% of the matched S2max signals saw only one potential S1 candidate during the matching process. The remaining 76.3% of the matched S2max signals saw multiple S1 candidates within the time window and where matched to the S1 signal of lowest light according to the likelihood. Naturally, matching scenarios where at least one signal species sees multiple candidates of the respective other species are prone to pile-up. Fig. 4.24 shows the correlation of both counting variables, the nS2perS1 and the nS1perS2max, for the DP data set of low-energy S1 events. Only ∼15.3% of the low-energy events correspond to the competition-less matches with nS2perS1 = nS1perS1max = 1.

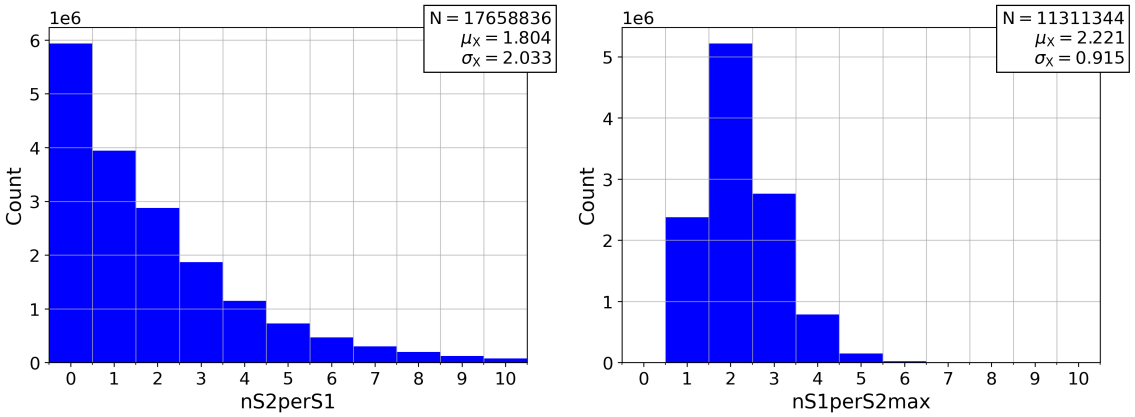

Figure 4.23: The distributions of the counts of the respective other signal species within the allowed matching window (for low-energy events with S1L $\leq$ 100 p.e.). The nS2perS1 distribution (left) shows that 33.6% of S1 signals are unaccompanied by any S2 signal, 22.3% have exactly one potential S2 candidate, while the remaining 44.1% see multiple S2 signals. The nS1perS2max distribution (right) reveals that, out of all matched events, 23.7% of the matched S2max signals saw only one potential S1 candidate, while the other 76.3% saw multiple S1 candidates during the matching process. The cleanest matches are those where both perspectives contain only one candidate of the respective other signal species (see Fig. 4.24).

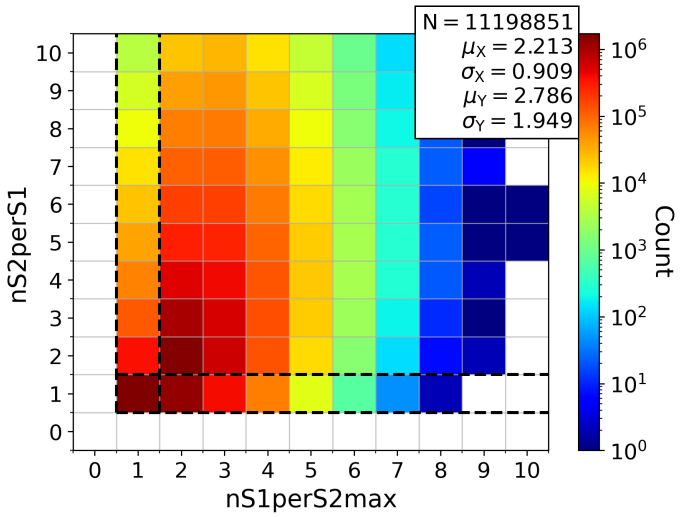

Figure 4.24: The correlation of the variables (shown in Fig. 4.23) counting the respective other signal species' candidates within the matching time window (for low-energy events with S1L $\leq$ 100 p.e.). Only the events where both signal species see only one candidate of the respective other signal species are the unequivocal matches. A "pile-up" cut to these events has an efficiency of $\sim$15.3% for these low-energy DP events.

Evidently, due to the large range of the log(S2maxL/S1L) distribution observed in Fig. 4.22, a cut to the pile-up-free competition-less matches is necessary; albeit at a large cost in terms of the cut efficiency. Fig. 4.25 shows the distributions of the log(S2maxL/S1L)

versus the S1F90 of the same two Neutron data sets collected on July, 10th and August, 13th 2019 after the application of the pile-up cut. Finally, the later data set (right) follows the expectation of a concentrated NR population around an S1F90 of ∼0.70 and at a lower log(S2maxL/S1L) than the similarly concentrated ER population which is distributed around an S1F90 of ∼0.30. This is a great success after the discovery of the Charge-up! Unfortunately, the earlier data set (left) does not obey the same standards and still has a large log(S2maxL/S1L) spread and consequently no similarly concentrated NR population. At this point, the only reasonable explanation is a consequence of the Charge-up effect increasing the detection efficiency of S2 signals, especially for low-energy S1 events (as discussed with respect to Fig. 4.10 and Fig. 4.11, right). Consequently, the large log(S2maxL/S1L) range can still be explained by wrong matches due to "pile-up", but with the twist, that the true low-energy S2 signal evaded the detection. Moreover, potentially clean events with exactly one true S1 and one true S2 signal are more likely to end up as only one detected S1 signal, and hence an unmatched event, in the earlier data. This is, again, strongly supported by the progression of the fraction of unmatched S1 signals in Fig. 4.10 (left) which decreases with the Charge-up effect.

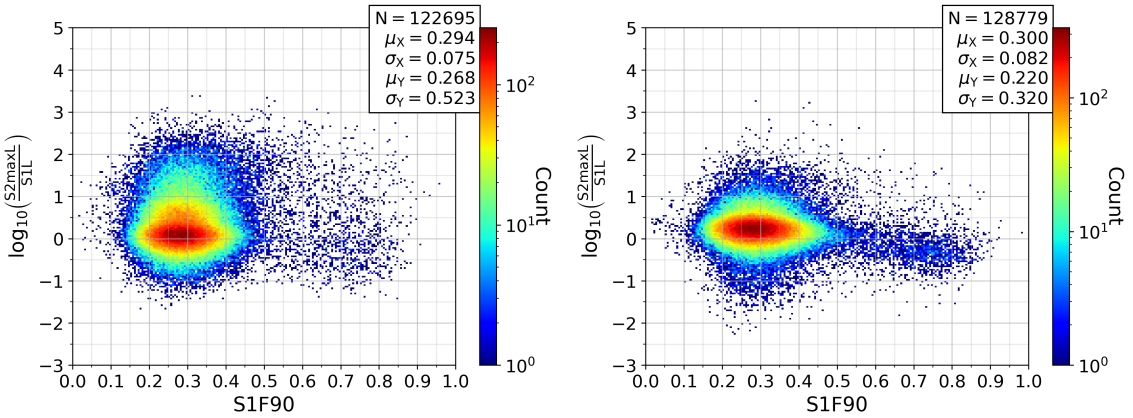

Figure 4.25: The same two sample Neutron calibration data sets as in Fig. 4.22 (pile-up-cut to nS2perS1 = nS1perS2max = 1) collected around July, 10th (left) and August, 13th (right) 2019. While the later data set finally follows the expectation of a concentrated ER and a corresponding NR population, the earlier data set lacks any focused NR region on the log(S2maxL/S1L) scale. This led to the exclusion of the data collected before the aproximate saturation of the Charge-up effect on July, 22nd 2019 (see Fig. 4.26).

A quantitative way of characterizing the evolution of the log(S2maxL/S1L) distribution is simply by its mean and standard deviation. Fig. 4.26 shows the time progression of the mean and the standard deviation of the pile-up-cut data. Therein, the data has been divided into two groups labeled "DP1" and "DP2", collected before and after July, 22nd 2019. In the DP1 data, the spread of the log(S2maxL/S1L) distribution is converging with the progression of the Charge-up effect while the mean is still shifting to a lower value. Only in the DP2 data, the Charge-up effect has reached a saturation level where both, the mean and the standard deviation of the log(S2maxL/S1L) distribution, have stabilized. Therefore, the decision was made to focus the baseline WIMP search entirely on the stable DP2 data and exclude the entire DP1 data. This comes with a detrimental loss of statistics as the WIMP search DP data is reduced from 54'205 runs (or 22.224 live-days) to 29'599 runs

(or 11.102 live-days) which accounts to an efficiency of 54.606% in runs (or 49.952% in live-time). Fortunately, the subset of the Neutron calibration DP data remains relatively unaffected as it is reduced from 4'156 runs (or 1.550 live-days) to 3'501 runs (or 1.267 live-days), which accounts to an efficiency of 84.240% in runs (or 81.737% in live-time). A minor loss of data (34 runs or 0.063%) within the reconstruction chain and the cut to Charge-up-stable data are summarized as cut #11 and #12 respectively in Table 4 which serves as a continuation of the data selection statistics summarized in Table 1, Section 3.1.3. This concludes the detailed explanation of the consequences of the Charge-up effect. Chapter 5 will continue with the study of the experimental ER and NR backgrounds based on the SP and the DP2 data.

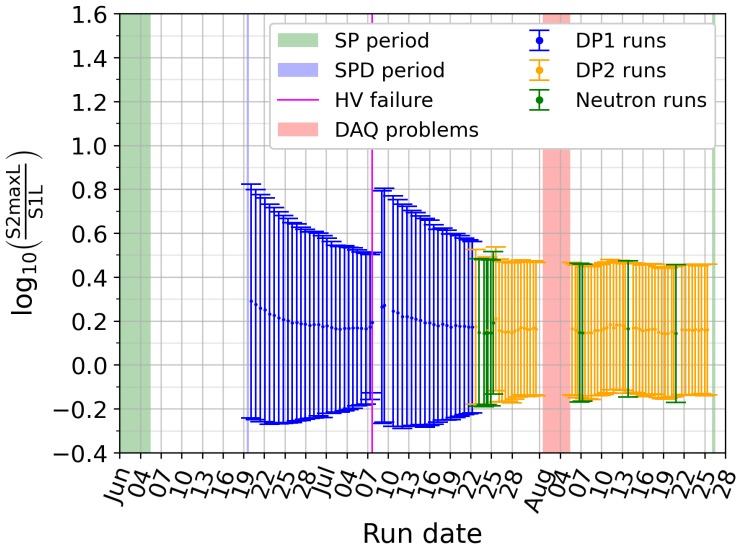

Figure 4.26: The progression of the mean and the standard deviation of the pile-up-cut log(S2maxL/S1L) distribution. While the characteristics of the log(S2maxL/S1L) distribution are converging under the increasing Charge-up effect in the DP1 data, the DP2 data shows a stable behavior resulting in clear ER and NR regions as in Fig. 4.25 (right). This led to the exclusion of the unstable DP1 data from the WIMP search thereby reducing the WIMP search DP data from 54'205 runs (or 22.224 live-days) to 29'599 runs (or 11.102 live-days), see Table 4.

Cut		**WIMP search DP data**		
#	Cut name	Number of runs	Collective efficiency	Individual efficiency
	Total	63'877	100.000%	100.000%
0-10	Data selection	54'239	84.912%	84.912%
11	Reconstruction chain	54'205	84.858%	99.937%
12	Charge-up-stable data	29'599	46.337%	54.606%

Table 4: A continuation of the data selection statistics summary given in Table 1, Section 3.1.3. The cuts #0 through #10 are summarized as #0-10 and explained in detail in Section 3.1. Cut #11 corresponds to the application of the entire reconstruction chain which results in a minor loss of statistics. Cut #12 corresponds to the exclusion of the Charge-up-unstable DP1 data as shown in Fig. 4.26 and explained in detail here in Section 4.2.1.

5 Understanding the experimental backgrounds

Searching for a WIMP signal in a current direct detection experiment is in some sense more difficult than looking for a needle in a haystack. Apart from the abundant background of electronic recoils (ERs), the hay, the sought-after signal is indifferent in signature to another background in form of the nuclear recoils (NRs), the needle imitations. Therefore, a thorough understanding of both types of backgrounds is fundamental. While the ERs and NRs have been introduced conceptually in detail in Section 2.1.1, the completed event reconstruction (Chapter 4) allows the thorough study of both recoil types from a data perspective.

Firstly, the dominant background of the ERs facilitates the gauge of the energy scale of the detector. As the vast majority of ER events origin from the radioactive β^--decay of the ^{39}Ar isotopes, a fit of the resulting β spectrum consolidates the conversion from keV of recoil energy to photoelectrons of S1L via the light yield (LY) parameter as originally introduced in Section 2.2. The LY is assumed to be independent of the recoil type and as such relates the WIMP's nuclear recoil energy to the corresponding S1L. The extraction of the LY via the ER background is presented in Section 5.1.

Secondly, the WIMP-like NR background is treated in Section 5.2. While any direct detection experiment undergoes enormous efforts to minimize the flux of NR-creating neutrons (as detailed in Section 2.1.2), a dedicated source of neutrons allows to gauge the detector's response to WIMP-like events. Therefore, the ArDM collaboration uses a dedicated "Neutron calibration" data set where a ^{252}Cf source emitted a large number of neutrons next to the detector. The study of the Neutron calibration data set allows to confine an energy-dependent "Region of Interest" (RoI), in terms of the recoil type discrimination variables, where the majority of clean NRs is expected and the contribution of ERs is minimal. The cut efficiency of NRs into this RoI is then estimated with the help of Monte Carlo simulations.

Thereafter, in Section 5.3, the last cut in form of the fiducial volume is defined based on a subset of the available DP data. This subset, labeled the "DP calibration" data, is assumed to be WIMP-free and serves as the control experiment to study the backgrounds in ArDM. Finally, having introduced all cuts relevant to the data analysis, the respective general cut efficiencies are evaluated on the DP calibration data set in Section 5.4. This will conclude the chapter and the preparation for the WIMP search, which will be conducted in Chapter 6.

While the analysis and results presented in this chapter are produced by the author, the ^{252}Cf Monte Carlo simulation is a result of multiple previous efforts of the ArDM collaboration. The PhD student Botao Li made a major contribution via his work on improving the ^{252}Cf simulation and the master student Duy Lai provided great help during his master thesis on the quantification of the ER and NR backgrounds [103].

A summary of the available data after the completed event reconstruction and the applied Charge-up stability cut (Chapter 4) is given in Table 5. The 332.7 million events or 4.365 live-days of SP data are used to extract the LY at zero drift field in Section 5.1. The 72.0 million events or 1.267 live-days of Neutron calibration data are used to understand the detector's response to NRs and confine the RoI in Section 5.2. Finally, the available DP data is split into the DP calibration data set of 77.7 million events or 1.406 live-days and the WIMP search data set of 462.8 million events or 8.429 live-days. The DP calibration data set is used to confine the RoI and fiducial volume cuts in Sections 5.2 and 5.3 and

evaluate the respective cut efficiencies in Section 5.4, whereas the WIMP search data will only be un-blinded in Chapter 6. The combined available DP data therefore accounts to 540.4 million events or 9.834 live-days. It also has to be emphasized, that in the statistics of the SP data (in Table 5), every S1 signal is counted as a fully reconstructed event as there are no S2 signals in the absence of the drift and extraction fields. Conversely, the statistics of the DP and the Neutron data only count reconstructed events with one S1 signal and at least one matched S2 signal.

	Data involved in the WIMP search				
Label	Runs	Events	Live-days	Percentage	Purpose
Total	34'755	-	15.467	100.00%	
SP	5'156	$332.7 \cdot 10^6$	4.365	28.22%	LY estimation in Section 5.1
Neutron calibration	3'501	$72.0 \cdot 10^6$	1.267	8.19%	Detector response to NRs in Section 5.2
DP calibration	3'836	$77.7 \cdot 10^6$	1.406	9.09%	Confinement of RoI and fiducial volume in Sections 5.2 and 5.3; Evaluation of cut efficiencies in Section 5.4
WIMP search DP	22'262	$462.8 \cdot 10^6$	8.429	54.50%	WIMP search in Chapter 6

Table 5: A summary of the available data statistics involved in the WIMP search. This table is a progression of the data statistics in Table 2 (Data selection, Section 3.1.3) to include the completed event reconstruction and the applied Charge-up stability cut (Chapter 4 and Section 4.2.1 specifically). While for the SP data all S1 signals are counted as events, only matched S1 signals are counted in the DP and Neutron data sets. The available DP data (9.834 live-days) is divided into a DP calibration set (14.29%) and the WIMP search data set (85.71%). The latter will only be un-blinded in Chapter 6.

Cut name	**Explanation**	**Variable(s) before cut**
EoI (Energy of Interest)	S1L $\leq$ 100 p.e.	Fig. 3.11 (left)
Event quality	S1maxFracPMT ≤ 0.3	Fig. 3.14 (left)
Matching	nS2perS1 = 1 and nS1perS2max = 1	Figs. 4.23 and 4.24
Match quality	$0.75 \leq$ S2maxQMatch ≤ 1.25	Fig. 4.4 and Fig. 4.5
Fiducial volume cut	0.02 ms $\leq t_{\text{Drift}} \leq$ 1.0242 ms and $R \leq$ 35.0 cm	Fig. 5.14
RoI (Region of Interest)	Energy-dependent S1F90 and S2maxL/S1L limits; see Section 5.2.2, Table 7	E.g. Fig. 5.7 (left)

Table 6: A list of the final cuts, as applied to the data presented in Chapters 5 and 6. If not stated otherwise, the listed cuts are applied in order and only the ultimate cut is named to label the respective data.

With the progressing analysis and the completed event reconstruction (Chapter 4) the data cuts have also evolved. Table 6 shows a summary of the updated and final cuts which are used in Chapters 5 and 6. Compared to Table 3 (from Chapter 4), the low-energy cut is

still the same but relabeled as "Energy of Interest". The "Event quality" cut is simply a part of the previous "Preselection" cut. While the "Match quality" is still assessed equally, the "Matching" itself consists of events with only one possible S1 and S2 signal respectively. The "fiducial volume" cut removes the edge volume of the detector to avoid unwanted sources of backgrounds and is presented in Section 5.3. Finally, the "Region of Interest" (RoI) cut is the aforementioned reduction to a region in the recoil type discrimination variable space, where the ratio of NRs over ERs is expected to be high. The detailed cut values of the energy-dependent RoI cut are given in Table 7, Section 5.2.2. Additionally, all the respective and combined cut efficiencies evaluated on the DP calibration data will be presented in Section 5.4. If not stated otherwise, the listed cuts are applied in order and only the ultimate cut is named to label the respective data.

5.1 The ER background - Gauging the Light yield

ArDM uses atmospheric argon (AAr) as the detection material. The main abundance in AAr is the ^{40}Ar isotope. In the atmosphere, the ^{39}Ar isotope is produced cosmogenicly via the neutron capture of the ^{40}Ar isotope. The production rate of ^{39}Ar from cosmic ray neutrons at sea-level is expected to be 759 ± 128 atoms/kg$_{\mathrm{Ar}}$/day [104]. In AAr, this production is eventually saturated and results in an abundance of ^{39}Ar of $8.2 \cdot 10^{-16}$ (compared to an abundance of ^{40}Ar of 0.9960, [104]). ^{39}Ar itself is a β^- emitter with a half-life of 268 ± 8 years [104]. In the AAr used in ArDM, the ^{39}Ar activity has been measured to be $\sim$1 Bq/kg$_{\mathrm{LAr}}$ [83]. This renders the β^--rays, or electrons, from the ^{39}Ar the by-far dominant source of events inside ArDM. While this is an, in principle, unwanted background, its dominating presence and the distinct form of the β^- spectrum allow the gauge of the energy scale via a fit. The theoretical conversion from keV of recoil energy to photoelectrons of S1L is discussed in detail in Section 2.2 and equates to

$$\mathrm{S1L_{NR}}\left(F_{\mathrm{Drift}}\right) = \overbrace{\mathrm{LY}\left(F_{\mathrm{Drift}} = 0\ \frac{\mathrm{V}}{\mathrm{cm}}\right)}^{\text{Light Yield}} \cdot \overbrace{\underbrace{\mathcal{L}_{\mathrm{eff}}\left(E_{\mathrm{NR}}\right)}_{\substack{\text{Nuclear and}\\ \text{bi-excitonic}\\ \text{quenching}}} \cdot E_{\mathrm{NR}}}^{E_{\mathrm{ee}}} \cdot \overbrace{q_{\mathrm{rec}}\left(E_{\mathrm{NR}}, F_{\mathrm{Drift}}\right)}^{\substack{\text{Recombination}\\ \text{quenching}}} \tag{5.1}$$

for nuclear recoils (repetition of Eq. 2.9). For electronic recoils, this conversion simplifies significantly, as the electron-equivalent energy equals the electronic recoil energy as $E_{\mathrm{ee}} = E_{\mathrm{ER}}$. To simplify even further, the recombination quenching q_{rec} is only present in the case of a non-zero drift field strength. Hence, at zero drift field strength, the conversion reduces to

$$\mathrm{S1L_{ER}}\left(F_{\mathrm{Drift}}\right) = \mathrm{LY}\left(F_{\mathrm{Drift}} = 0\ \frac{\mathrm{V}}{\mathrm{cm}}\right) \cdot E_{\mathrm{ER}} \cdot q_{\mathrm{rec}}\left(E_{\mathrm{ER}}, F_{\mathrm{Drift}}\right) \overset{F_{\mathrm{Drift}}=0\ \frac{\mathrm{V}}{\mathrm{cm}}}{\Rightarrow} \mathrm{S1L_{ER}} = \mathrm{LY} \cdot E_{\mathrm{ER}} \tag{5.2}$$

for electronic recoils. Eqs. 5.1 and 5.2 inherently include the assumptions that the light yield LY is independent of the recoil energy $\frac{\partial \mathrm{LY}}{\partial E_{\mathrm{ee}}} = 0$ and the recoil type $\mathrm{LY} = \mathrm{LY_{ER}} = \mathrm{LY_{NR}}$. As such, the SP data with deactivated drift field is the ideal candidate to directly gauge the conversion factor from keV of recoil energy to photoelectrons of S1L. This LY can be understood as a detector property which, in turn, determines down to which recoil energies a reasonable statistic of photoelectrons can be collected.

In order to evaluate the LY and its consistency over time, the available SP data is split into eleven time-ordered subsets of data. Fig. 5.1 (left) shows the distribution of the S1TTR versus the S1L for one of those SP data subsets. As becomes apparent, the S1L spectrum is not completely uniform along the vertical TTR axis of the detector. This can be caused by non-uniform optical behavior of the LAr vessel coating as well as different sensibilities of the 24 PMTs. Consequently, the SP data subsets are further divided into S1TTR bins of 0.05 ranging from 0.15 up to 0.55. Below an S1TTR of 0.15 the SP data appears inconsistent and has no significant statistics above an S1L of 400 photoelectrons. Above an S1TTR of 0.55 is very close to the top of the LAr volume (see Fig. 3.13, Section 3.3) and also has low statistics. Fig. 5.1 (right) shows the S1L spectrum of the S1TTR bin from 0.25 to 0.30 of the same SP data subset.

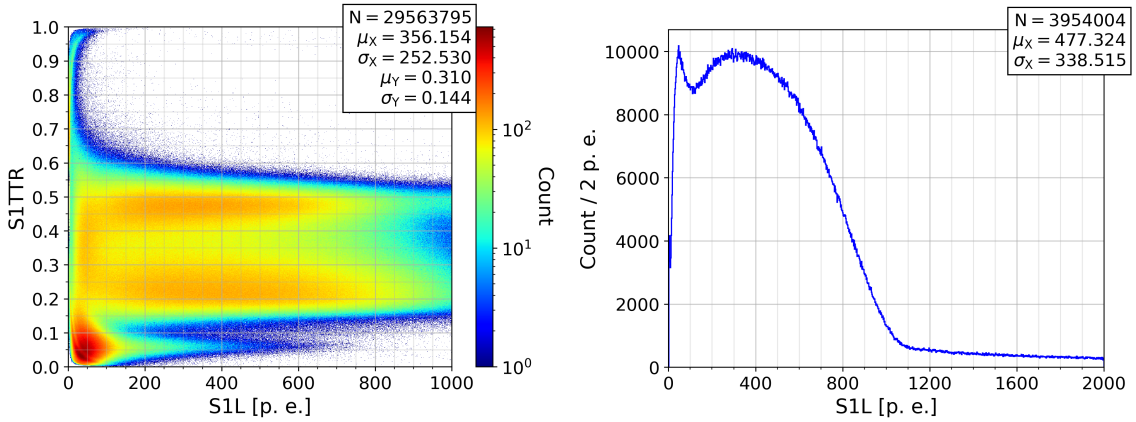

Figure 5.1: *Left*: The distribution of the S1TTR versus the S1L of one subset of the available SP data. The S1L spectrum is not uniform along the vertical TTR axis. This can be a consequence of different sensibilities of the 24 PMTs or non-uniform optical behavior of the coating on the LAr vessel. *Right*: The S1L spectrum from that SP data subset in the S1TTR bin from 0.25 to 0.30.

In order to extract the LY estimate, the data is fitted to the well-known β-spectrum. Concretely, the detector resolution is taken into account by convolving the β-spectrum with a Gaussian distribution. The corresponding standard deviation σ of the detector resolution is a second order polynomial as a function of the energy. Additionally, an exponential background is added to the convolved β-spectrum to account for other event sources, such as nuclear recoils, or electronic recoils which do not originate from the ^{39}Ar decay. A more detailed description of the fitting function with all parameters used for ArDM can be found in Khoi Nguyen's PhD thesis submitted in 2016 [105]. Fig. 5.2 (left) visualizes an example β-spectrum fit result for the same SP data subset in the same S1TTR bin of 0.25 to 0.30. The fit results in a good agreement of the data with the right shoulder of the β-spectrum as well as the exponential tail. Specifically, the exponential tail is fitted first and then used as a guessing guide in the combined fit. The resulting LY is then slightly varied around the best fit value in a χ^2-scan (see Fig. 5.2, right). The subtraction of the minimal χ^2 value allows the quantification of the final best fit value with a measure of its uncertainty.

Repeating this fitting procedure for all eight S1TTR bins and all eleven SP data subsets subsequently results in 88 fit results as showcased in Fig. 5.3. While there is some visible time variation of the resulting LY in all S1TTR bins, there is no clear trend, and the

overall time variation within an S1TTR bin is smaller than ∼10%. The larger variation of up to ∼25% appears between the different S1TTR bins. In general, with the exception of the [0.15, 0.20] S1TTR bin, there is a trend of higher LY towards lower S1TTR; and hence lower position inside the detector. This can point to more sensible PMTs or better optical conditions at the bottom of the detector. Nevertheless, these acceptable variations can be averaged out as a reasonable assumption as the WIMP signal shouldn't have a time or positional preference. As such, the resulting averaged LY of the 88 fit results accounts to LY = 1.82±0.13 p.e./keV. This is significantly higher than the initially expected ∼1.33 p.e./keV (in Section 2.2.2) and will be incorporated into the expected WIMP S1L spectrum in Section 6.1.

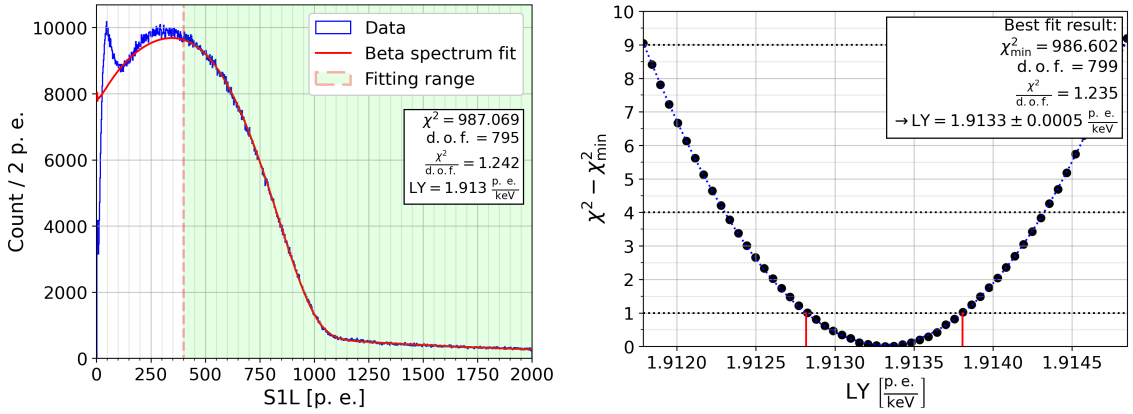

Figure 5.2: *Left*: An example β-spectrum fit on the same subset of data as in Fig. 5.1 (right). The fitting function is a sum of an exponential background, which accounts for events not originating from the ^{39}Ar decays, as well as the ^{39}Ar β-spectrum convolved with a Gaussian detector resolution. *Right*: After the β-spectrum fit, the initially resulting LY guides a χ^2-scan, where the final LY best fit value is fixed with a measure of its uncertainty.

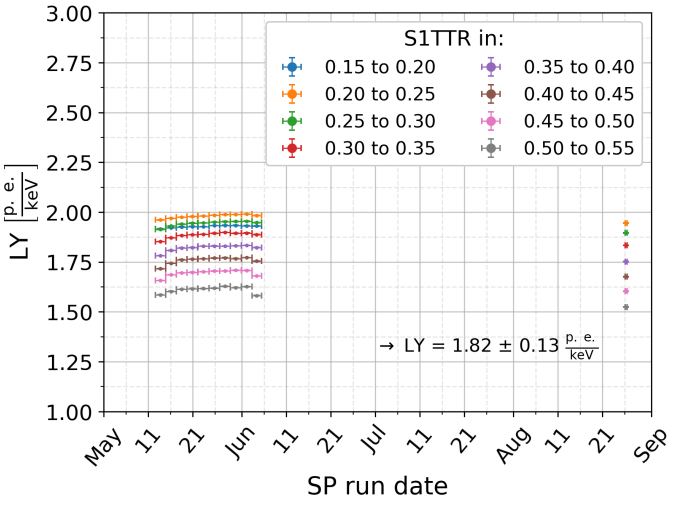

Figure 5.3: The combined β-spectrum fit results of all eight S1TTR bins in all eleven time-ordered subsets of SP data. The time variation of up to ∼10% and the S1TTR-related variation of up to ∼25% can be averaged out as the WIMP signal shouldn't have a time or positional preference. The resulting averaged LY accounts to LY = 1.82 ± 0.13 p.e./keV and will be incorporated into the expected WIMP S1L spectrum in Section 6.1.

5.2 The NR signal - Discriminating the WIMP-like ^{252}Cf events

The neutron-induced nuclear recoils (NRs) are the closest possible events to a hypothesized WIMP-induced signal. As such, they are a blessing and a curse. While they allow to gauge the detector's response to WIMP-like events, they also constitute a non-discriminable background. To combat the neutron-induced NR background, direct detection experiments either make enormous efforts to reduce neutron-induced NR events, use veto detectors around the main target volume to identify them, or both. As described in Section 2.1.2, ArDM is located at the Laboratorio Subterráneo de Canfranc (LSC), in Canfranc, Spain. The depth of the LSC experimental halls corresponds to 2450 m.w.e. (meters water equivalent), which significantly reduces the flux of potentially neutron-inducing cosmic muons. Additionally, ArDM surrounded its main LAr vessel with a minimally 50 cm thick polyethylene shield to minimize external sources of neutrons. Lastly, the natural radioactivity of the detector compartments has been screened in previous efforts of the ArDM collaboration [78]. The remaining neutron flux constitutes the instrumental NR background, which can only be quantitatively monitored via a control experiment in ArDM (the DP calibration data, as will be evaluated in Section 6.2).

On the positive side, the NR-like nature of the hypothesized WIMPs allow to imitate its event signature with a dedicated neutron source. Therefore, ArDM used a ^{252}Cf source with a neutron activity of 529 ± 6 Bq at the time of use. The DP data collected with this source positioned at the top of the detector constitutes the 1.267 live-days of "Neutron calibration" data as listed in Table 5 at the beginning of the chapter. The goal of the Neutron calibration data is twofold. Firstly, it should lead to the confinement of an energy-dependent "Region of Interest" (RoI) in terms of the recoil type discrimination variables log(S2maxL/S1L) and S1F90. Therein, one expects a significant amount of NR events and a statistically small amount of ER events. Secondly, it should then help quantifying the NR-specific cut efficiencies into this RoI, as these are expected to be the same for WIMPs. The second goal is achieved with the help of a Monte Carlo simulation of the ^{252}Cf source inside ArDM. As presented in Section 5.2.1, the ^{252}Cf simulation results in an expectation of reconstructed single-scattering NR events inside the WIMP search volume; before any data analysis cuts. This can then be compared to the extracted ^{252}Cf source events from the Neutron calibration data, after the applied data analysis cuts, to constrain the NR-specific efficiencies. In summary, Section 5.2.2 first constrains an RoI in bins of energy and then evaluates the NR-specific efficiencies of the data analysis cuts into this RoI via the comparison to the simulated expectation. This then completes the necessary preparation to evaluate the theoretical WIMP signal (in Section 6.1) based on the expected cut efficiencies.

5.2.1 The simulation perspective: MC expectation

The Monte Carlo simulation of the ^{252}Cf source events inside ArDM has been fine-tuned in multiple previous efforts of the ArDM collaboration and is explained in detail in Wei Mu's PhD thesis submitted in 2019 [78]. More recently, the PhD student Botao Li invested significant effort into the ^{252}Cf simulation and produced the results presented here in collaboration with the author. The actual ^{252}Cf source, used for the Neutron calibration data, had a known neutron activity of 529 ± 6 Bq during the 1.267 live-days of data-taking. This resulted in 57.9 ± 0.6 million neutron emissions. In order to study the expected ^{252}Cf neutron events, 60.0 million neutron emissions have been simulated. Thereof, only 2.6 million simulated neutron emissions led to interactions within the ArDM detector, resulting in a

geometric efficiency of $\varepsilon_{\text{Geo}} = 0.04344\ (3)$. The resulting energy spectrum (blue) in terms of nuclear recoil energy E_{NR} and reconstructed light S1L (related via Eq. 5.1) can be seen in Fig. 5.4.

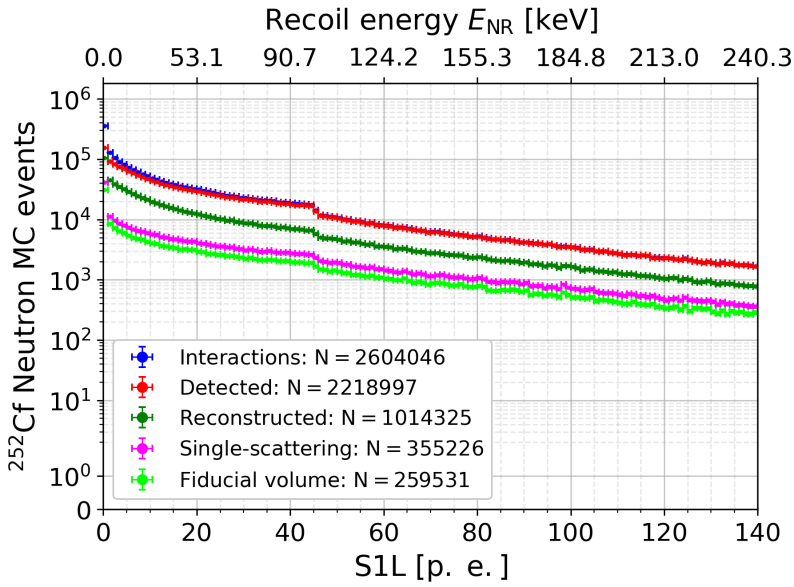

Figure 5.4: The energy spectra resulting from 60 million simulated ^{252}Cf neutron emissions. Thereof, only 2.6 million simulated neutron emissions led to interactions within the ArDM detector, insinuating a geometric efficiency of $\varepsilon_{\text{Geo}} = 0.04344\ (3)$. The resulting efficiencies of the detection, reconstruction, single-scattering and fiducial volume are presented in Fig. 5.5.

In general, the spectrum follows an approximately exponential decrease with increasing energy, which is the nature of the elastic scattering cross section (that dominates at low energies). There is a small recognizable dip at ∼100 keV of recoil energy E_{NR}, which is attributed either to a decrease in the cross section of the elastic neutron scattering on ^{40}Ar nuclei[6], or the manner in which the simulation software GEANT4[7] handles the cross section at different energies.

Subsequently, the spectra of the simulated ^{252}Cf neutron interactions undergo several cuts. The resulting spectra can be seen in Fig. 5.4, while the respective cut efficiencies are presented in Fig. 5.5 in the same respective color. Firstly, the detection efficiency (red) shows a sharply increasing behavior at very low energies (S1L $\leq$ 10 p.e.), akin to a fluid detection threshold, and subsequently saturates with increasing energy resulting in a total efficiency of $\varepsilon_{\text{Det}} = 0.8521\ (2)$. Secondly, the reconstruction efficiency (dark green) shows a slight decrease at low energies (S1L $\leq$ 40 p.e.) but stabilizes around a total efficiency of $\varepsilon_{\text{Rec}} = 0.4571\ (3)$. These NR-specific detection and reconstruction cut efficiencies will

[6]Compare, for example, with the data available on `https://www-nds.iaea.org/exfor/servlet/E4sSearch2`, and specifically the TENDL nuclear data library [106]. This specific decrease in the cross section occurs at kinetic neutron energies of ∼1 MeV, where the elastic scattering is still dominant. The kinematics of elastic nuclear recoils (Eq. 1.15), relate the maximally possible nuclear recoil energy for a given neutron kinetic energy as $E_{\text{NR,max}} \approx rE_{\text{n,kin}}$. The kinematic factor r (Eq. 1.16), with the replacement of the WIMP m_χ with an ^{40}Ar nuclei, evaluates to ∼1/10, therefore reasoning the observed dip in the spectra.

[7]See `https://geant4.web.cern.ch/`.

subsequently also be assumed for the theoretical WIMP signal (in Section 6.1). Thirdly, the efficiency of events which are reconstructed as single-scatterers (purple), increases at low energies (S1L $\leq$ 40 p.e.) and eventually saturates resulting in a total efficiency of $\varepsilon_{\mathrm{SSc}} = 0.3502\ (5)$. Fourthly, the efficiency of single-scattering events which are reconstructed inside the fiducial volume (light green), is approximately constant and equates to a total efficiency of $\varepsilon_{\mathrm{Fid}} = 0.7306\ (7)$. The cut values of the fiducial volume cut are listed in Table 6 and will be presented in detail in Section 5.3. In summary, the efficiency of ^{252}Cf neutron interactions to reconstructed single-scatterers inside the fiducial volume (black) accounts to $\varepsilon_{\mathrm{Tot}} = 0.0997\ (2)$ or roughly $\sim$10%.

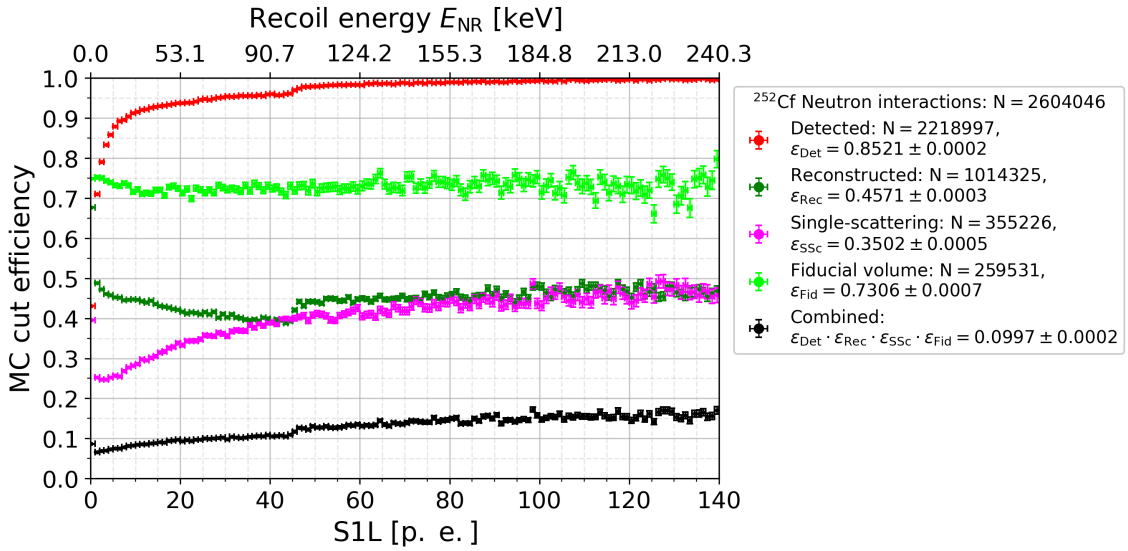

Figure 5.5: The estimated cut efficiencies resulting from 60 million simulated ^{252}Cf neutron emissions. Thereof, only 2.6 million simulated neutron emissions led to interactions within the ArDM detector, insinuating a geometric efficiency of $\varepsilon_{\mathrm{Geo}} = 0.04344\ (3)$. The resulting energy spectra after the interaction, detection, reconstruction, single-scattering and fiducial volume cuts are presented in Fig. 5.4. The detection and the reconstruction efficiency of the simulated ^{252}Cf neutron events are henceforth assumed to be the respective efficiencies applying to the theoretical WIMP signal (as implemented in Section 6.1).

Consequently, these efficiencies of the simulated ^{252}Cf neutron emissions allow an estimated expectation of the ^{252}Cf neutron events in the Neutron calibration data. This expectation can be seen Fig. 5.6 in the same respective colors. As mentioned previously, the known neutron activity of the used ^{252}Cf source of 529 ± 6 Bq during the 1.267 live-days of data-taking results in 57.9 ± 0.6 million expected neutron emissions. After the consideration of the geometric efficiency of $\varepsilon_{\mathrm{Geo}} = 0.04344\ (3)$, 2.52 ± 0.03 million neutron interactions (blue) are expected inside the detector. Finally, the simulation-based estimation of reconstructed single-scatterers inside the fiducial volume (light green) accounts to 0.251 ± 0.003 million events. These expected spectra are shown in S1L bins of ten photoelectrons, as the WIMP search will also be conducted in these bins. In the upcoming Section 5.2.2, this expectation of the ^{252}Cf neutron events before the data analysis cuts will be compared to the extracted ^{252}Cf neutron events after the application of the data analysis cuts to estimate the NR-specific cut efficiencies thereof.

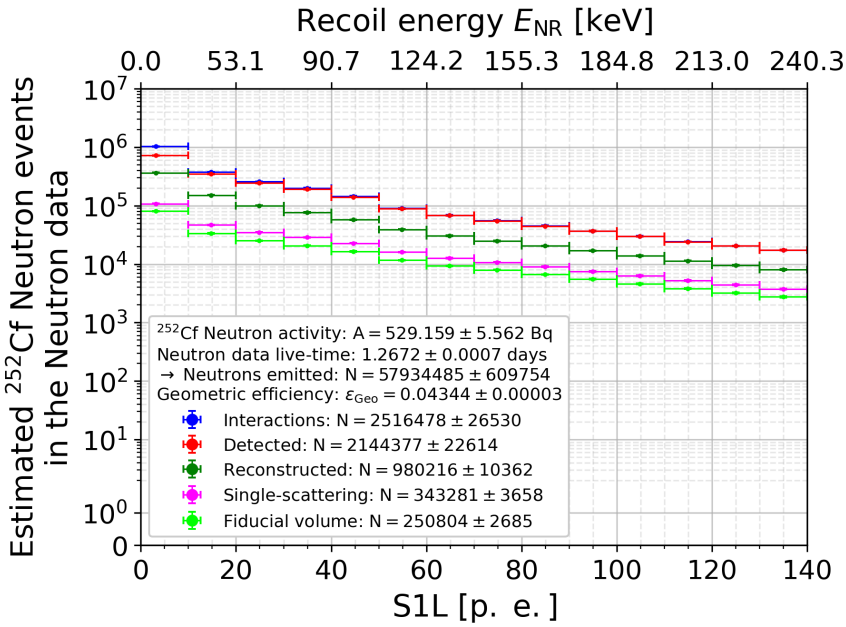

Figure 5.6: The expected spectra (in bins of ten photoelectrons) of the ^{252}Cf neutron events in the 1.267 live-days of the Neutron calibration data. The data analysis perspective will result in a spectrum of extracted ^{252}Cf NR events after all cuts inside the RoI (see Section 5.2.2). The comparison of the expected spectrum before the data analysis cuts (light green) and the extracted spectrum after the application of the cuts will facilitate the estimation of the NR-specific cut efficiencies into the RoI.

5.2.2 The data perspective: Confining and evaluating an RoI

On the data side, the comparison of the available Neutron and DP calibration data allows to confine the "Region of Interest" (RoI), where the ratio of the estimated likelihoods of NR events over ER events is greatest. As summarized in Table 5, the designated DP calibration data corresponds to 77.7 million events or 1.406 live-days of DP data. The Neutron calibration data consists of 72.0 million events or 1.267 live-days of DP data taken with a ^{252}Cf source active on top of the detector. As elaborated in Section 5.2.1, the simulation-guided expectation for the Neutron calibration data is 0.980 ± 0.010 million reconstructed ^{252}Cf neutron events, whereof 0.251 ± 0.003 million ^{252}Cf neutron events are single-scatterers inside the fiducial volume. In order to preserve the energy dependence of the RoI in terms of the recoil type discrimination variables (the log(S2maxL/S1L) and the S1F90), the data is split into S1L bins of ten photoelectrons each, up to 100 photoelectrons. Above 100 photoelectrons of S1L, or 184.8 keV of nuclear recoil energy E_{NR}, the expected WIMP signal becomes negligible (see Section 6.1). Fig. 5.7 shows the available DP (left) and Neutron (right) calibration data for the example S1L bin of 20 to 30 photoelectrons in the two recoil type discrimination variables of the log(S2maxL/S1L) versus the S1F90. All cuts, according to Table 6, up to and including the fiducial volume cut are already applied to both calibration data sets. The surplus of ^{252}Cf neutron events at an S1F90 of ∼0.70-0.90 and a log(S2maxL/S1L) of ∼-0.30 is clearly visible.

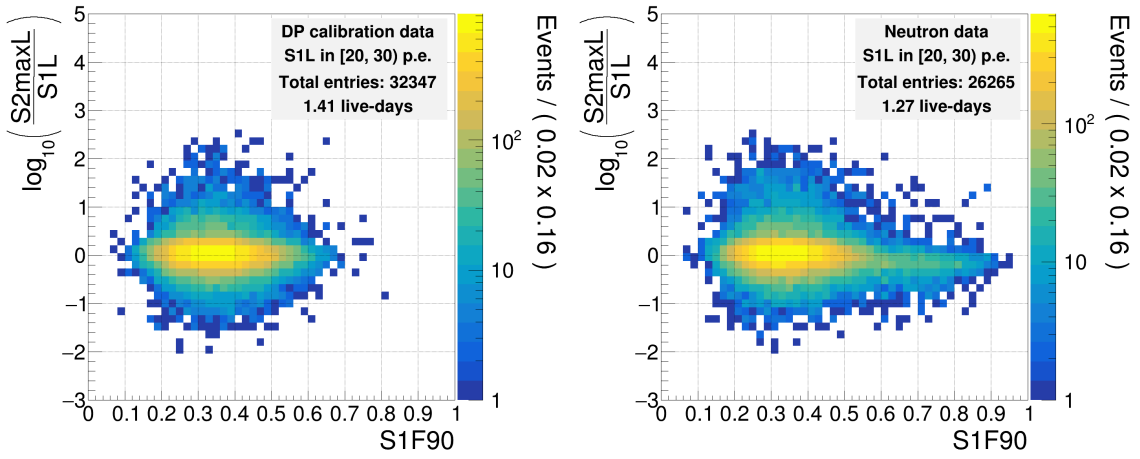

Figure 5.7: The available DP (left) and Neutron (right) calibration data for the S1L bin of 20 to 30 photoelectrons in the two recoil type discrimination variables of the log(S2maxL/S1L) versus the S1F90. All cuts, according to Table 6, up to and including the fiducial volume cut are already applied to both calibration data sets. The surplus of ^{252}Cf neutron events at an S1F90 of ∼0.70-0.90 and a log(S2maxL/S1L) of ∼-0.30 is clearly visible.

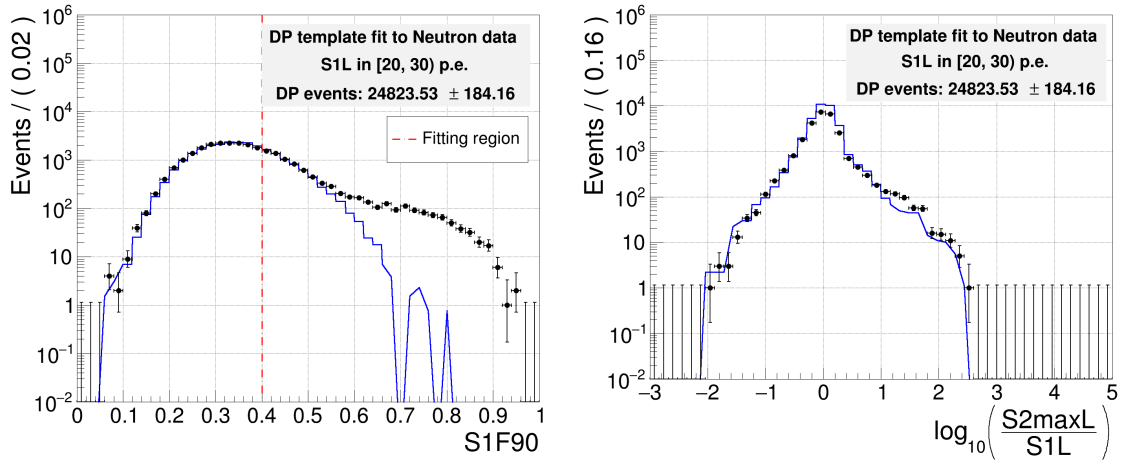

Figure 5.8: The one-dimensional perspectives of the fit of the extended DP probability density function (extracted from Fig. 5.7, left) to the Neutron calibration data (Fig. 5.7, right) in the S1L bin of 20 to 30 photoelectrons. While the surplus of ^{252}Cf neutron events is clearly visible in the S1F90 perspective (left) due the discrepancy of the Neutron calibration data (black) and the DP template (blue) at S1F90 $\geq$ 0.50, the log(S2maxL/S1L) perspective (right) alone reveals no significant discrimination.

In order to constrain the RoI for a specific bin of S1L, the DP calibration data is first normalized to a two-dimensional probability density function. This DP probability density function or DP template is then extended with an amplitude which labels the number of DP-like events. Thereafter, this extended DP template is fitted to the Neutron calibration data to gauge the DP-like events. Fig. 5.8 shows the one-dimensional perspectives of the fit result in the same S1L bin of 20 to 30 photoelectrons. In this case, the fitting region

was limited to the ER-dominated data with S1F90 $\leq$ 0.40. The fit result estimates, that 24'823.52 $\pm$ 184.16 events out of the 26'265 entries (see Fig. 5.7, right) are DP-like. The S1F90 perspective of the fit result (Fig. 5.8, left) visualizes the strength of the S1F90 as a recoil type discriminator. While the agreement of the extended DP template (blue) and the Neutron calibration data (black) is extremely good at S1F90 values up to ~0.50, the surplus of ^{252}Cf neutron events becomes apparent at higher S1F90 values. The small number of DP-like events with an S1F90 $\geq$ 0.70 is either due to a high fluctuation of ER events, or the instrumental NR background in the DP calibration data. Contrarily, the log(S2maxL/S1L) perspective (Fig. 5.8, right) emphasizes, that the log(S2maxL/S1L) is not a sufficiently discriminating recoil type variable on its own in the case of the ArDM data. However, it is impossible to judge, how much this fact is a consequence of the discovered S2L-affecting Charge-up effect and its necessary correction (see Section 4.2).

Alternatively, Fig. 5.9 (left) shows the two-dimensional fit result, in terms of the discrepancy, for the same S1L bin of 20 to 30 photoelectrons. Therein, a newly defined variable, called the "Neutron discriminator", accounts for the subtraction of the fit result from the data, normalized by the data. Consequently, blueish histogram bins show overfitting, mint greenish bins signify great agreement or the absence of events and reddish bins show underfitting or the surplus of events in the Neutron calibration data. While the largest region of dark red bins at an S1F90 $\geq$ 0.64 and a log(S2maxL/S1L) $\leq$ 0.20 corresponds to the accumulation of ^{252}Cf neutron events, there are also other dark red bins caused by fluctuations of the ER events or the pile-up of signals. Therefore, the RoI has to be set restrictively, as to allow a large part of the NR region but reject pile-up and ER events effectively. For this specific S1L bin of 20 to 30 photoelectrons, the RoI has been chosen to be limited by S1F90 $\geq$ 0.70 and log(S2maxL/S1L) $\leq$ 0.0 in a simple "box approach". This RoI serves as a cut in the WIMP search in Chapter 6 and confines wherein the expected background and signal are evaluated and the observed events are counted.

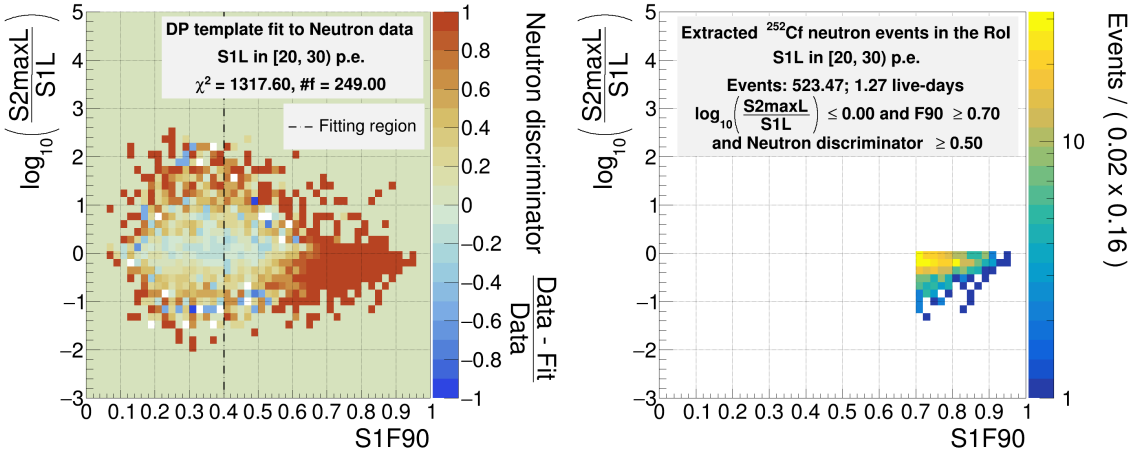

Figure 5.9: *Left*: The two-dimensional fit result of the extended DP probability density function (extracted from Fig. 5.7, left) to the Neutron calibration data (Fig. 5.7, right) in the S1L bin of 20 to 30 photoelectrons. The bin values in the histogram label the discrepancy of the data and the fit normalized to the data. *Right*: The discrepancy of events inside the RoI corresponds to the extracted ^{252}Cf neutron events in this S1L bin.

Finally, Fig. 5.9 (right) presents the extracted ^{252}Cf neutron events inside the RoI

resulting from the fitting routine for the same S1L bin of 20 to 30 photoelectrons. Therein, the additional RoI criterion of a Neutron discriminator value larger than 0.50 is listed, but this is not in effect in any S1L bin due to the restrictive nature of the chosen RoI limits in terms of the S1F90 and the log(S2maxL/S1L). The extracted ^{252}Cf neutron events are given by the Neutron calibration data minus the fit result, which accounts to 523.45 estimated ^{252}Cf neutron events inside the RoI for the S1L bin of 20 to 30 photoelectrons.

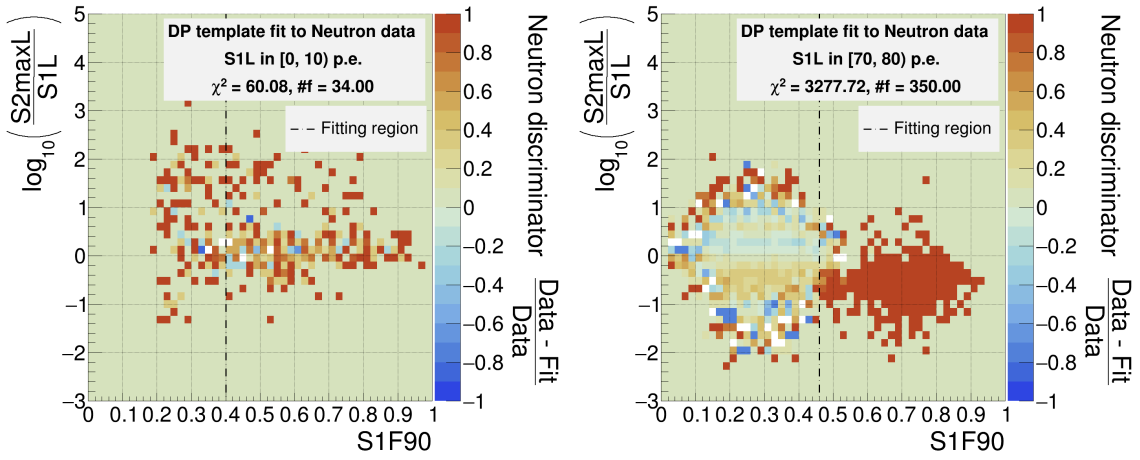

Figure 5.10: Further examples of the fit result of the respective extended DP templates to the Neutron calibration data. *Left*: In the extreme case of zero to ten photoelectrons, the NR and ER regions completely overlap and the event statistics are very low. The only reasonable choice for an RoI is to include the entire plane of the log(S2maxL/S1L) versus the S1F90. With bins of increasing S1L up to 50 photoelectrons, the NR and ER regions start to separate, as can be seen in Fig. 5.9 (left). *Right*: In even higher S1L bins, the NR and ER regions are eventually fully separated. The chosen RoI limits for all S1L bins are summarized in Table 7.

Fig. 5.10 illustrates examples of other S1L bins to facilitate the discussion of the RoI choices for all S1L bins. As expected, the separation of the NR and ER regions grows with increasing energy and larger statistics of photoelectrons. This evolution starts at the extreme case of zero to ten photoelectrons (Fig. 5.10, left). There, the event statistics are very low (with 413 and 497 total events in the DP and Neutron calibration data respectively), and the NR and ER regions seem to completely overlap. Consequently, the only possible RoI choice for the lowest S1L bin is to include the entire plane of the log(S2maxL/S1L) versus the S1F90 and exclude no events. Therefore, the extracted ^{252}Cf neutron events up to ten photoelectrons simply are the Neutron calibration data events minus the DP calibration events (rescaled according to the respective live-times). Continuing towards higher energies, the next S1L bins up to 50 photoelectrons all behave similar as the previously discussed bin of 20 to 30 photoelectrons (see Fig. 5.9, left), with a partial overlap of the NR and ER regions. Therein, the S1F90 lower limits of the RoI decrease with higher S1L bins. However, the upper log(S2maxL/S1L) limits are always set to 0.0 due to the consistency of the log(S2maxL/S1L) distribution's upper end as a function of energy, and the danger of pile-up events at higher values. A summary of the chosen RoI limits can be found in Table 7. Finally, for the remaining S1L bins above 50 photoelectrons (e.g. Fig. 5.10, right), the overlap of the NR and ER regions becomes smaller and smaller and eventually non-

existent. Consequently, the chosen S1F90 lower limits decrease further down to 0.50 at 90 to 100 photoelectrons. However, the log(S2maxL/S1L) upper limit remains at 0.0 for the same, previously mentioned reasons.

S1L bin	**S1F90**	**log(S2maxL/S1L)**
0-10 p.e.	≥ 0.0	≤ 5.0
10-20 p.e.	≥ 0.80	≤ 0.0
20-30 p.e.	≥ 0.70	≤ 0.0
30-40 p.e.	≥ 0.66	≤ 0.0
40-50 p.e.	≥ 0.62	≤ 0.0
50-60 p.e.	≥ 0.56	≤ 0.0
60-70 p.e.	≥ 0.54	≤ 0.0
70-80 p.e.	≥ 0.54	≤ 0.0
80-90 p.e.	≥ 0.50	≤ 0.0
90-100 p.e.	≥ 0.50	≤ 0.0

Table 7: The chosen RoI limits for all S1L bins which are considered in the WIMP search in Chapter 6. This defines the energy-dependent RoI cut, which aims to reduce the data to a region of high ratios of the likelihoods of NR over ER events.

Apart from the confinement of the RoI, the main purpose of the Neutron calibration data is to extract the NR-specific cut efficiencies into the RoI based on the abundance of the ^{252}Cf neutron events. Fig. 5.11 shows the spectrum of extracted ^{252}Cf neutron events for all S1L bins up to 100 photoelectrons according to the previously explained procedure (compare with Fig. 5.9 for the S1L bin of 20 to 30 photoelectrons). This accounts to a total of 7'208 $\pm$ 85 extracted ^{252}Cf neutron events below 100 photoelectrons for the 1.267 live-days of Neutron calibration data. The spectrum can be understood as a combination of the simulation-guided expectation before the cuts (see Section 5.2.1) and the specific cut efficiencies of the respective cuts (which will be evaluated on the DP calibration data in Section 5.4). Firstly, the simulation-guided expectation of the ^{252}Cf NR single-scatterers inside the fiducial volume is a roughly exponential decrease of events with increasing energy (see Fig. 5.6 specifically). The remaining cuts which go into the observed spectrum of extracted ^{252}Cf neutron events are the event quality, the matching, the match quality and the RoI cuts (as summarized in Table 6). The NR-specific individual cut efficiencies can not be extracted with absolute certainty, but the respective cut efficiencies on the DP calibration data are expected to be very similar as a reasonable assumption (except the RoI cut). As will be seen in Section 5.4, the efficiencies of the event and match quality cuts are sharply increasing with energy. This can be understood, in considering, that the underlying respective variables of the S1maxFracPMT and the S1TTR are both variables, which become less accurate or favorable with lower photoelectron statistics. Additionally, the matching cut increases roughly linearly with increasing energy. Lastly, the RoI cut (Table 7), with the exception of the very first S1L bin, increases the NR acceptance of the RoI with increasing energy and better separation of the two recoil regions (compare with the progression shown in Figs. 5.9 and 5.10). Consequently, the expected spectrum of surviving ^{252}Cf NR events is the multiplication of an exponential decrease (simulation-guided expectation) with a strictly increasing function (cuts). This also is, what is observed in the extracted ^{252}Cf spectrum in Fig. 5.11, as the increasing cut efficiency is dominating

up to an S1L of roughly 50 photoelectrons. Above that threshold, the cut efficiency is high enough that the exponentially decreasing spectrum becomes visible. The exception of a slightly higher number of extracted ^{252}Cf events in the very first S1L bin of zero to ten photoelectrons, is given by the fact, that the RoI cut exceptionally excludes no events in that energy range (see Table 7).

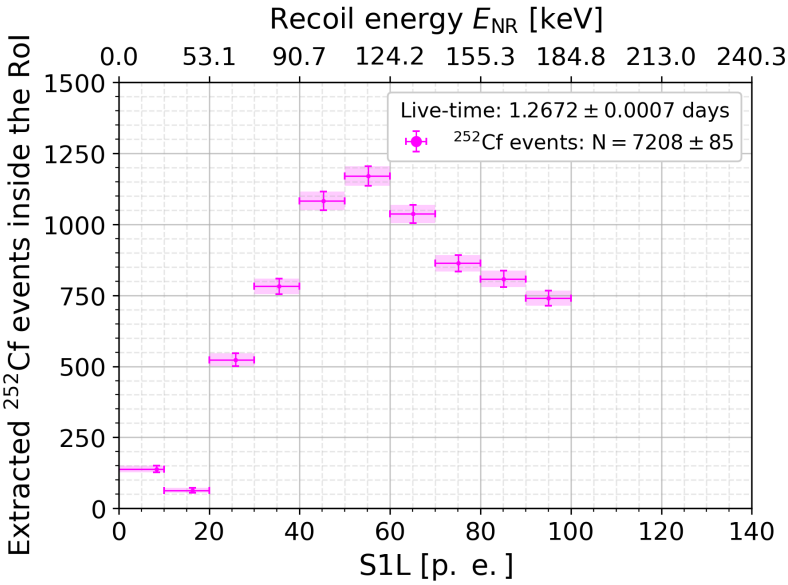

Figure 5.11: The spectrum of extracted ^{252}Cf NR events up to 100 photoelectrons after all cuts applied (as listed in Tables 6 and 7). The observed spectrum is the combination of the exponentially decreasing simulation-guided expectation (see Fig. 5.6) and the strictly increasing cut efficiency with increasing energy (see Section 5.4). The comparison of the observed spectrum after all cuts with the simulation-guided expectation before the cuts (Fig. 5.6) will allow the estimation of the NR-specific cut efficiencies (Fig. 5.12).

Finally, the observed spectrum of ^{252}Cf NR events, after the data analysis cuts (Fig. 5.11), can be compared to the simulation-guided expectation of the ^{252}Cf NR events, before the application of the data analysis cuts (Fig. 5.6). Therefore, Fig. 5.12 shows the resulting NR-specific cut efficiencies into the RoI up to 100 photoelectrons of S1L. These cut efficiencies specifically include the event quality, the matching, the match quality and the RoI cuts as listed in Tables 6 and 7. As expected, the NR-specific cut efficiency is an increasing function of the energy. As discussed previously, the event and match quality cuts sharply increase with increasing energy, therefore explaining the low efficiency below 20 photoelectrons. The main culprit for the generally low efficiency is the matching cut. As will be seen in Section 5.4, the matching cut efficiency in the DP calibration data is at the order of ∼10-20%, which reasons why the NR-specific cut efficiency doesn't exceed ∼13-14% at an S1L of 100 photoelectrons. The overall NR-specific cut efficiency of $\varepsilon_{\text{Cut}} = 0.03312 \pm 0.00038$ however, is dominated by the large simulation-guided expectation of events below ten photoelectrons before the cuts (see Fig. 5.6), and should not be understood as a defining value. This completes the study of the experimental NR background in ArDM. The NR-specific cut efficiencies are also expected to apply to the WIMPs and will be used in the theoretical WIMP expectation in Section 6.1.

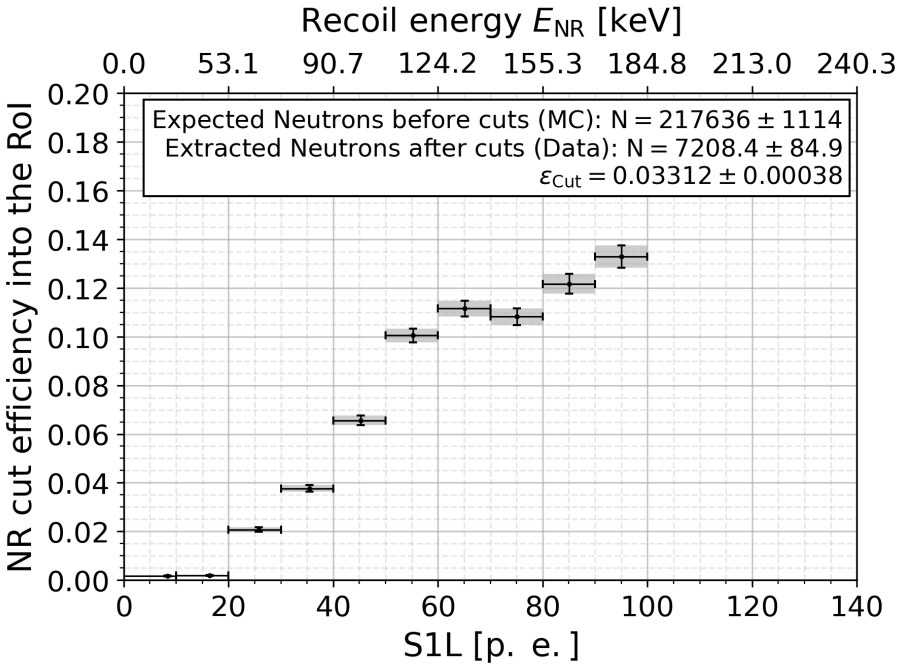

Figure 5.12: The resulting NR-specific cut efficiencies by comparing the simulation-guided expectation before the cuts (Fig. 5.6) and the extracted ^{252}Cf NR spectrum after the application of the cuts (Fig. 5.11). The herein contained NR-specific cut efficiencies include the event quality, the matching, the match quality as well as the RoI cuts as listed in Tables 6 and 7. While the event and match quality cuts exclude the majority of events below 20 photoelectrons, the matching cut is the main limiting factor at higher energies up to 100 photoelectrons (compare with the matching cut efficiency at the order of ∼10-20% for the DP calibration data in Section 5.4). These NR-specific cut efficiencies are also expected to apply to the WIMPs and will be used in the theoretical WIMP expectation in Section 6.1.

5.3 The fiducial volume cut

The last remaining cut (in the complete list in Table 6), which is yet to be discussed, concerns the confinement of a fiducial volume. The fundamental purpose of a fiducial volume lies in removing any excess background or inconsistent events. Alternatively, the fiducial volume can be seen as the region, where one is expecting to be sensitive to the signal. In the case of ArDM, that view would translate to a volume, where one is certain, that the boundary of the LAr vessel is not interfering with the desired signal. In liquid noble gas direct detection experiments, it is common to remove the outermost centimeters from the detection volume for the above reasons. E.g., in the comparable argon experiment DarkSide-50, the collaboration decided to cut away 1-2 cm radially and ∼4 cm from the top and bottom of the LAr detection volume [50]. While this serves as a guideline for the fiducial volume confinement of ArDM, the DP calibration data has to be checked for potential accumulations of events near the edges of the LAr detection volume. Therefore, Fig. 5.13 shows the vertical and horizontal perspectives of the DP calibration data in the S1L range from 10 to 100 photoelectrons after the application of all cuts, as listed in Table 6, with the exception of the fiducial volume cut. The lowest S1L bin of zero to ten

photoelectrons is deliberately excluded in Fig. 5.13, as the corresponding 740 events would dominate the appearance due to the relaxed RoI cut in said bin. The indicated volume (black, shaded) corresponds to the LAr vessel with a radius of 36.75 cm and a height of 111.2 cm or 1.0442 ms of drift time respectively. Due to the constant drift velocity of $v_{\text{Drift}} = 1.065$ mm/μs (see Fig. 4.12, Section 4.2), 10 μs of drift time correspond to 10.65 mm of drifted height. In the DP calibration data, only 53 of the 61 total events are reconstructed inside the LAr vessel. Out of these 53 events, only one single event is located within the top and bottom 2 cm. Radially speaking, the volume at radii between 35 cm and 36.75 cm corresponds to the outermost 9.30% of the LAr detection volume and, given a uniform distribution of events, one would expect 4.93 events therein. The actually observed 8 events correspond to a positive 1.38 sigma Poisson-fluctuation. Therefore, the number of events observed in the outermost part of the LAr detection volume is a positive fluctuation, but not a concerning one.

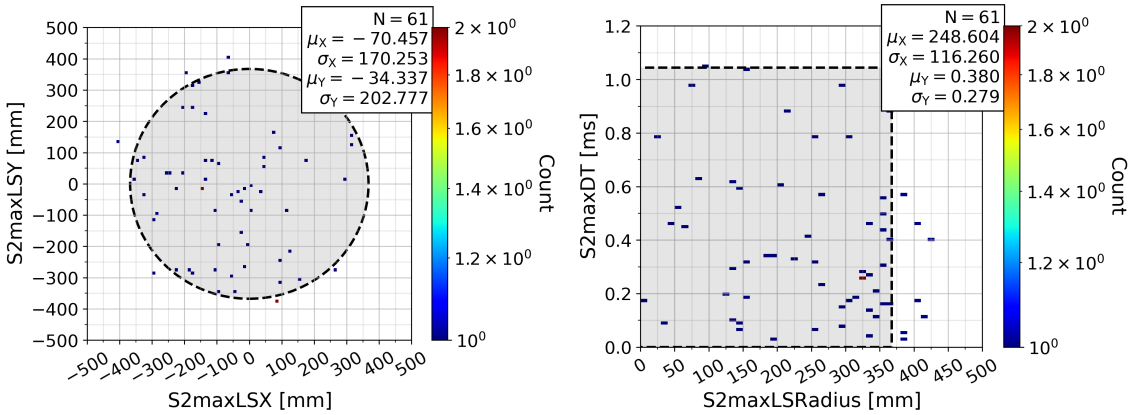

Figure 5.13: The positional distribution of all events from the DP calibration data set with an S1L between 10 and 100 photoelectrons; after all cuts, except the fiducial volume, are applied according to Table 6. The indicated volume (black, shaded) corresponds to the LAr vessel with a radius of 36.75 cm and a height of 111.2 cm or 1.0442 ms of drift time.

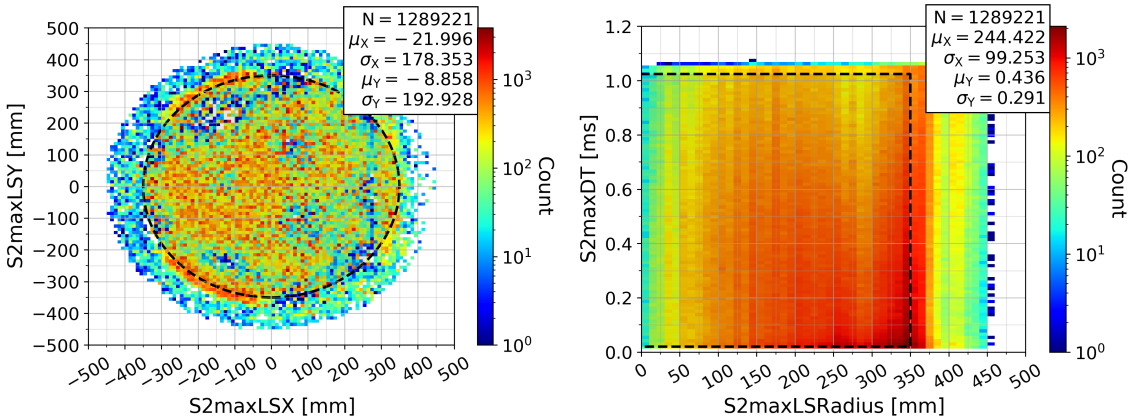

Figure 5.14: The positional distribution of all events in the DP calibration data, after the application of the match quality cut according to Table 6. The indicated fiducial volume (black, shaded) has a total acceptance of 81.7% in the case of the DP calibration data (see Section 5.4).

Consequently, a fiducial volume cut similar to the one confined by DarkSide-50 is deemed safe, and the ArDM fiducial volume cuts away the top and bottom 0.02 ms or 2.13 cm and the outermost 1.75 cm radially. Therefore, the ArDM fiducial volume restricts the signal region radially to obey $r \leq 35$ cm, and confines the vertical component to obey 0.02 ms $\leq t_{\text{Drift}} \leq$ 1.0242 ms. Fig. 5.14 illustrates the confinement of the fiducial volume with the match-quality-cut DP calibration data as in the order listed in Table 6. In the DP calibration data, the fiducial volume cut has a total efficiency of 81.7%. The detailed cut efficiency as a function of the energy and the recoil type discrimination variables of the S1F90 and log(S2maxL/S1L) will be presented in Section 5.4.

Having confined a fiducial volume, the relevant measure for the theoretical WIMP signal (Section 6.1) is the LAr target mass within. The fiducial mass $m_{\text{Ar,fid}}$ of ArDM can be evaluated as:

$$\begin{aligned} m_{\text{Ar,fid}} &= \rho_{\text{Ar,b.p.}} \cdot \pi r_{\text{max}}^2 \cdot v_{\text{Drift}} \Delta t_{\text{Drift}} \\ &= 0.4670731462 \cdot \left(\frac{r_{\text{max}}}{\text{cm}}\right)^2 \cdot \left(\frac{\Delta t_{\text{Drift}}}{\text{ms}}\right) \text{ kg}, \\ &= 574.57 \text{ kg, for } r_{\text{max}} = 35 \text{ cm and } \Delta t_{\text{Drift}} = 1.0042 \text{ ms}. \end{aligned} \tag{5.3}$$

Therein, the argon density at the boiling point $\rho_{\text{Ar,b.p.}} = 1.396$ g/cm^3 is considered, as well as the known drift velocity $v_{\text{Drift}} = 1.065$ mm/μs. With the fixed fiducial volume and the corresponding radial limit of $r_{\text{max}} = 35$ cm and the allowed drift time window of $\Delta t_{\text{Drift}} = 1.0042$ ms, the fiducial mass accounts to $m_{\text{Ar,fid}} = 574.57$ kg. This therefore concludes the study of the fiducial volume.

5.4 All cut efficiencies evaluated on the DP calibration data

This section serves as a reference, presenting all respective and combined cut efficiencies, as evaluated on the DP calibration data. All cuts are applied as listed in Table 6 and the corresponding references to the plots of the cut variables (before the respective cuts), can be found therein. For the purpose of presenting, every cut efficiency is shown in order of the application and as a function of the main variables of interest; the energy in the form of the S1L and the recoil type discrimination variables, the log(S2maxL/S1L) versus the S1F90. It has to be emphasized, that this section only serves to improve the understanding of the general behavior of the DP data. Conversely, the NR-specific cut efficiencies, which will be assumed for the WIMPs in Section 6.1, where already evaluated in Section 5.2.

The very first cut to the energies of interest (EoI) or to S1L $\leq$ 100 p.e. is straight-forward and not presented here. Next, the event quality cut aims to remove events which where dominantly detected by a single PMT. This could be caused by strongly located background radiation or dark counts. Therefore, the applied cut is S1maxFracPMT ≤ 0.30, meaning that the PMT with the largest percentage of the collected light can not have collected more than 30% of the total light (see Section 3.3, Fig. 3.14, left). Fig. 5.15 shows the resulting cut efficiencies in the DP calibration data as a function of the main variables of interest. Naturally, the event quality cut removes the majority of events with very few detected photoelectrons (S1L $\leq$ 15 p.e., left). There are no reconstructed events with S1L $\leq$ 4 p.e. and the respective cut efficiency defaults to 0.0. With increasing energy above 20 photoelectrons, the event quality cut becomes very efficient resulting in a total cut efficiency of $\varepsilon_{\text{Total}} = 0.817$ in terms of the S1L. The total cut efficiency in terms of the recoil type discrimination variables (right) is slightly higher with $\varepsilon_{\text{Total}} = 0.847$. This difference is

caused by the fact, that the DP calibration, at this stage, still consists of all possible events. Consequently, the cut efficiency as a function of the S1L (left) includes all 17.7 million events with a valid S1 signal. Contrarily, the discrimination variable log(S2maxL/S1L) is only defined for the 11.3 million events with one S1 signal and at least one matched S2 signal. Interestingly, there appears a correlation of high cut efficiency with lower S1F90. Consequently, one expects the majority of the high-S1maxFracPMT events to be of very low energy (S1L $\leq$ 15 p.e.) and have an NR-like S1F90. While this does mimic the characteristics of the WIMP signal, it remains a necessary cut due to the aforementioned dangers of background radiation and dark counts.

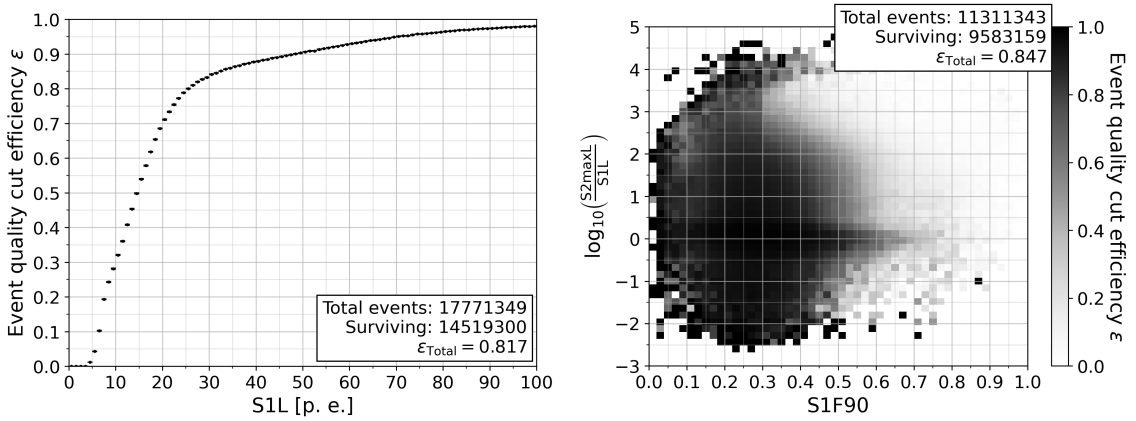

Figure 5.15: The event quality cut efficiency as a function of the energy (left) and the recoil type discrimination variables (right) evaluated on the DP calibration data. The cut is applied as listed in Table 6 and in the respective order therein. The specific behavior of the cut efficiency is discussed in the text.

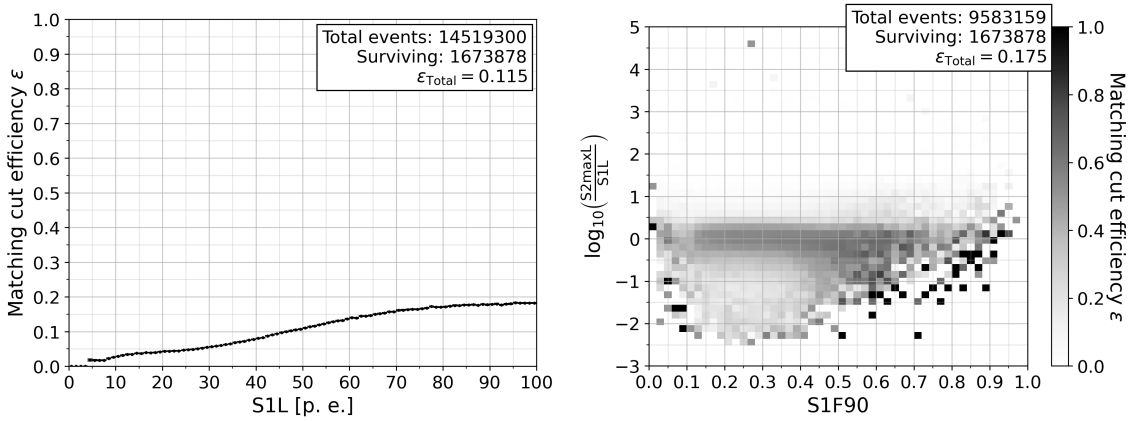

Figure 5.16: The matching cut efficiency as a function of the energy (left) and the recoil type discrimination variables (right) evaluated on the DP calibration data. The cut is applied as listed in Table 6 and in the respective order therein. The specific behavior of the cut efficiency is discussed in the text.

The next cut, applied in order, is the matching in Fig. 5.16. Specifically, the matching requires nS2perS1 = 1 and nS1perS2max = 1 and consequently excludes all events with

any competing signals, and hence, any ambiguous pairing possibilities. As expected, this matching cut is very harsh, but necessitated by the discovery of the Charge-up (see Section 4.2) to ascertain the conformity of events. In terms of the 14.5 million events with an S1 signal (left), the total cut efficiency results in $\varepsilon_{\text{Total}} = 0.115$. The roughly linear increase of the cut efficiency with the energy implies that a pile-up of signals is likelier the lower the energy. As designed, the matching removes the vast majority of pile-up events at high log(S2maxL/S1L) (right). Again, the cut efficiency in terms of the recoil type discrimination variables is higher, with $\varepsilon_{\text{Total}} = 0.175$, as it only pertains to the 9.6 million events which consist of one S1 and at least one S2 signal. This discrepancy of the cut efficiencies is subsequently lifted, as the matching reduces the data to the 1.67 million events which consist of exactly one S1 and one S2 signal.

The subsequently applied cut is the match quality cut in Fig. 5.17. It aims to reduce the data to quality single-scatterers, which obey the correlation of the S2 drift time (S2maxDT) and the S1 positional height in the form of the S1TTR (see Section 4.1.1). This is enforced by requiring $0.75 \leq \text{S2maxQMatch} \leq 1.25$. Naturally, the match quality cut efficiency is increasing as a function of the energy (left), as the S1TTR becomes a more reliable variable with more reconstructed photoelectrons. The electronic recoils are expected to be dominated by single-scatterers, which is visible in the high cut efficiency in the center of the ER region at an S1F90 below 0.50 and a log(S2maxL/S1L) of ∼0.20. The total efficiency of the match quality cut accounts to $\varepsilon_{\text{Total}} = 0.770$.

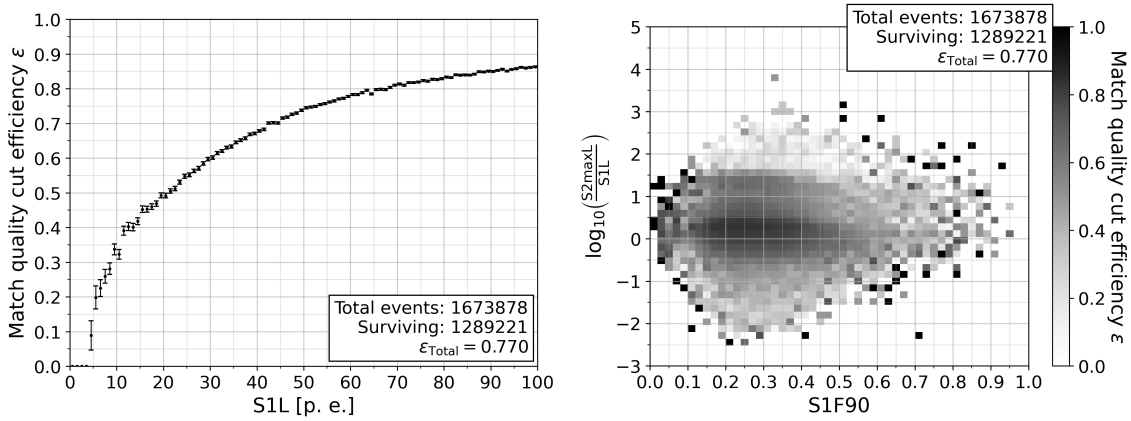

Figure 5.17: The match quality cut efficiency as a function of the energy (left) and the recoil type discrimination variables (right) evaluated on the DP calibration data. The cut is applied as listed in Table 6 and in the respective order therein. The specific behavior of the cut efficiency is discussed in the text.

The next cut, applied in order, is the fiducial volume cut in Fig. 5.18. The fiducial volume cut efficiency shows an increase as a function of the energy (left). This emphasizes the importance of the fiducial volume confinement, as the efficiency supports the increased appearance of low-energy events near the boundary of the LAr vessel. The fiducial volume cut efficiency accounts to $\varepsilon_{\text{Total}} = 0.817$. Consequently, the fiducial volume cut removes 9.3% of the LAr detection volume (see Section 5.3) and 18.3% of the events, further signifying the importance of the fiducial volume. In terms of the recoil type discrimination variables (right), the fiducial volume appears to remove more pile-up-like events at higher log(S2maxL/S1L). This could pertain to specific pile-up events which survived the previous

cuts. It is generally possible, that a matched event consists of two true underlying events, where the respective other signal species was not reconstructed, and the correlation of the drift time and the S1TTR is good by chance.

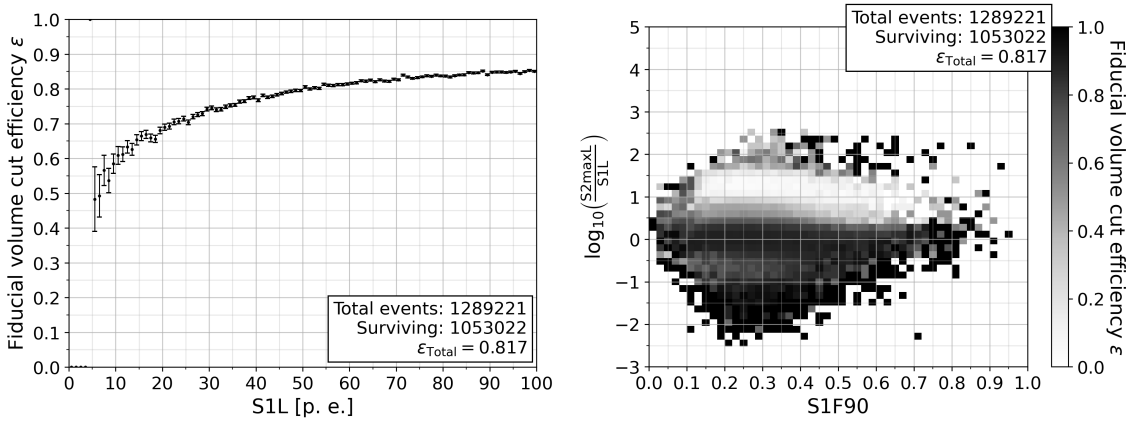

Figure 5.18: The fiducial volume cut efficiency as a function of the energy (left) and the recoil type discrimination variables (right) evaluated on the DP calibration data. The cut is applied as listed in Table 6 and in the respective order therein. The specific behavior of the cut efficiency is discussed in the text.

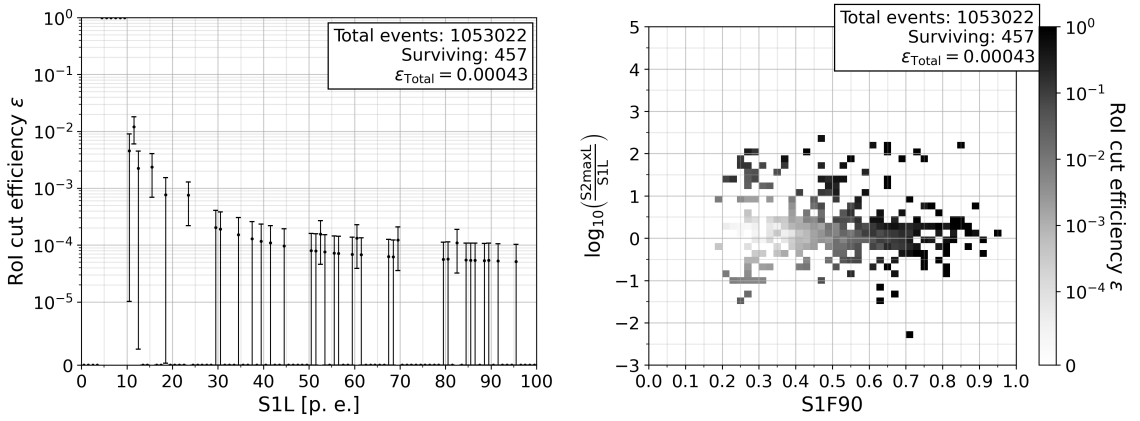

Figure 5.19: The RoI cut efficiency as a function of the energy (left) and the recoil type discrimination variables (right) evaluated on the DP calibration data. The cut is applied as listed in Table 7 and in the respective order of Table 6. The specific behavior of the cut efficiency is discussed in the text.

Finally, the RoI cut is applied as listed in Table 7, with the resulting efficiencies shown in Fig. 5.19. Exceptionally for this cut, due to the low values, the efficiency is shown on logarithmic scales. Due to the increasing separation of NRs and ERs with increasing energy (as discussed in Section 5.2.2), the RoI cut efficiency on the ER-dominated DP calibration data decreases strongly at very low energies (S1L $\leq$ 40 p.e., left), and then eventually saturates at $\approx 5 \cdot 10^{-5}$. The majority of bins of one photoelectron have an efficiency of zero, as no events survive the RoI cut. Contrarily, the events with an S1L between five and ten photoelectrons have an efficiency of one, as the RoI cut excludes no events until

up to ten photoelectrons. The total RoI cut efficiency results in $\varepsilon_{\text{Total}} = 4.3 \cdot 10^{-4}$ with just 457 events surviving all cuts. The RoI cut efficiency as a function of the recoil type discrimination variables (right) is skewed due to the dominating presence of the events below ten photoelectrons. As listed in detail in Table 7, above ten photoelectrons, no events with a log(S2maxL/S1L) ≥ 0.0 or an S1F90 ≤ 0.50 survive the RoI cut. As designed, the RoI cut reduces the data to strongly NR-like events.

As a summary, Fig. 5.20 shows the combined cut efficiency of all cuts, applied in order, versus the energy on a logarithmic scale. All respective total efficiencies $\varepsilon_{\text{Total}}$ refer to the combined total efficiency up to, and including the respective cut. The efficiencies emphasize again the dominance of the matching cut. Thereafter, the efficiency of the events in the DP calibration data is $\varepsilon_{\text{Total}} = 0.094$ or 9.4%. After the fiducial volume cut, 5.9% of the DP calibration data events are still surviving. By design, the RoI cut removes the vast majority of the remaining events and accounts to a total combined efficiency of $\varepsilon_{\text{Total}} = 2.6 \cdot 10^{-5}$, resulting in 457 events of instrumental background. These remaining background events will serve as the estimation of the expected background in the WIMP search data in Section 6.2. The cut efficiency shown versus the recoil type discrimination variables in Fig. 5.21 (on a logarithmic scale), refers to the total combined efficiency of all cuts. Again, by construction, the combination of all cuts leads to strongly NR-like events. The total cut efficiency $\varepsilon_{\text{Total}} = 4.0 \cdot 10^{-5}$ is slightly higher, as this efficiency only refers to those events, where one S1 signal was matched with at least one S2 signal (to result in a valid log(S2maxL/S1L) value). This concludes the detailed study of the behavior of the DP calibration data under all applied cuts.

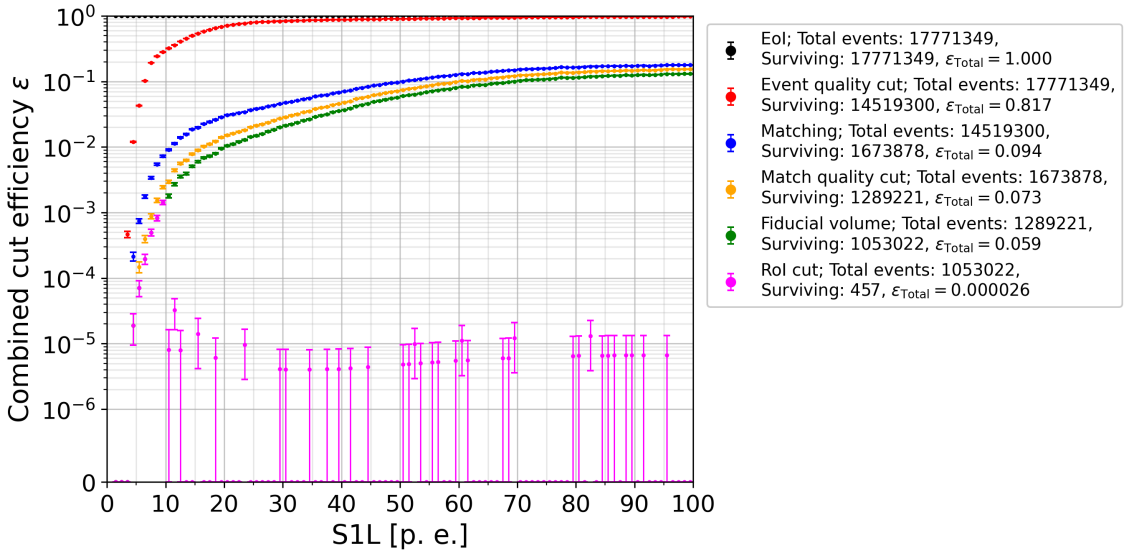

Figure 5.20: The combined cut efficiency of all cuts as a function of the energy evaluated on the DP calibration data. All cuts are applied as listed in Table 6 and in the respective order therein. The respective total cut efficiencies $\varepsilon_{\text{Total}}$ refer to the efficiency up to, and including the respective cut. The specific behavior of all cut efficiencies is discussed in the text.

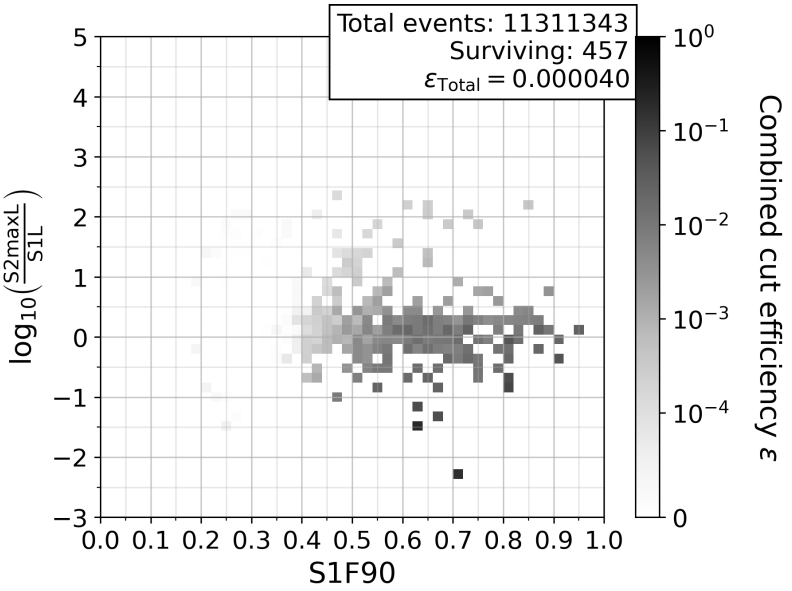

Figure 5.21: The combined cut efficiency of all cuts as a function of the recoil type discrimination variables evaluated on the DP calibration data. All cuts are applied as listed in Table 6. The specific behavior of all cut efficiencies is discussed in the text.

6 The WIMP search

Finally, after a long journey, all pieces are ready to be assembled in order to conduct the actual WIMP search. The theory of the hypothesized WIMPs has been discussed in detail in terms of their flux on Earth (Chapter 1, Section 1.3) and inside ArDM (Chapter 2, Section 2.2). While this served well to facilitate a general understanding, the actual values of the experimental parameters used therein, were based on assumptions and predictions. Having selected the data (Chapter 3), performed the event reconstruction (Chapter 4) and studied the experimental backgrounds (Chapter 5), these parameters need to be updated. The data analysis involved therein led to the knowledge of the exact experimental live-time, the effective drift field strength (Section 4.2) and the actual light yield value (Section 5.1). These exact values are incorporated into the theoretical WIMP spectrum in Section 6.1. In combination with the extracted WIMP-affecting cut efficiencies (Section 5.2), these ingredients result in an expected WIMP spectrum after the application of all cuts. As neither the WIMP mass m_χ (or potentially its distribution), nor the spin-independent WIMP-nucleon cross section σ_p^{SI} are known, this expected WIMP spectrum has to be re-evaluated for every different WIMP mass m_χ and scaled linearly with the spin-independent WIMP-nucleon cross section σ_p^{SI}. The expected WIMP spectrum corresponds to the first ingredient for the statistical analysis performed in the WIMP search.

The second and third ingredients are the expected background and the actually observed events, which are evaluated in Section 6.2. The therein used data sets correspond to the exact same data sets as listed in Table 5 from the introduction to Chapter 5. The 1.406 live-days of "DP calibration" data are hypothesized to be WIMP-free and used to gauge the expected background of events after all applied cuts. The 8.429 live-days of "WIMP Search" data are utilized to count the observed events after all applied cuts.

Finally, all ingredients are combined in a statistical analysis, which constrains the signal strength μ based on the expected background and the observed events. As the expected signal spectrum has to be calculated for every tested WIMP mass m_χ individually, the statistical analysis has to be repeated respectively. For each tested WIMP mass m_χ, the constrained value of the signal strength μ will correspond to an upper limit of the spin-independent WIMP-nucleon cross section σ_p^{SI} at the 90% confidence level. Section 6.3 accompanies the reader through the statistical testing process for one specific WIMP mass m_χ (Section 6.3.1), and then presents and discusses the resulting WIMP exclusion limit after scanning through a wide range of potential masses (Section 6.3.2). This will conclude the baseline WIMP search conducted with the principal ArDM DP data set. A preview to potential short- and long-term improvement plans will be given in the Outlook. All the results presented in this chapter are the work of the author.

6.1 The theory perspective: Expected WIMP signal spectrum

The theoretical expectation of the WIMP signal spectrum has already been discussed in detail in Chapters 1 and 2 and implemented into a projected ArDM sensitivity in Section 2.2.2. However, while these discussions served well to create a general understanding, they were based on specific assumptions about the experimental process. The completion of the data selection (Chapter 3), the event reconstruction (Chapter 4) and the study of the DP and Neutron calibration data (Chapter 5) results in necessary adjustments of these assumptions. Consequently, this section incorporates the updated prerequisites in order to align the theoretical WIMP expectation with the available data, the experimental conditions of

ArDM and the analysis procedure.

Firstly, the available WIMP search data has been overestimated. The discovery of the Charge-up effect (Section 4.2) necessitated a reduction to Charge-up-stable data resulting in an efficiency of 54.61% (see Table 4 in Section 4.2.1). The remaining data was collected between July, 22nd and August, 21st 2019. The consequence of the date range lies in the resulting velocity $\vec{v}_E$ of the Earth in the frame of the Milky Way galaxy, which, in turn, determines the velocity v_χ of the WIMPs relative to the Earth (see Eq. 1.25, Section 1.3.4). Fig. 6.1 shows the distribution of the data (green) over the days of the year. The theory (blue, from Eqs. 1.4 to 1.8) translates each run of data into the resulting Earth velocity $|\vec{v}_E|$ in the frame of the Milky Way galaxy. The resulting average Earth velocity $\overline{v_E}$ (orange) is shown with the respective standard deviation as $\overline{v_E} = 250.3 \pm 1.8$ km/s. In a justified simplification, this value is taken as the Earth's velocity during the entirety of the available data. This average velocity $\overline{v_E}$ translates to an "average" date of data-taking in 8th, August, 16:00 (rounded to the hour). Naturally, a higher velocity v_χ of the WIMPs relative to the Earth results in higher nuclear recoil energies E_{NR}. Consequently, the average date of July, 1st, previously assumed in the projected sensitivity (Section 2.2.2), which corresponds to an average Earth velocity of $\overline{v_E} = 256.5$ km/s, has overestimated the resulting WIMP flux at higher recoil energies. Incorporating the WIMP-affecting cut efficiencies (which will be shown in Figs. 6.3 and 6.4), the correction of the average date of data-taking leads to a reduction of the total WIMP flux after all cuts at the order of ∼5%. Additionally, the total duration of the WIMP search data has been estimated to be 10.0 live-days in the projected sensitivity of ArDM. The final value of 8.429 live-days is lower, which decreases the expected WIMP flux after all cuts by the relative factor (15.71%). Likewise, the fiducial mass is another linear multiplication factor. The final fiducial mass of 574.57 kg (see Section 5.3) is slightly higher than the initially expected 560.0 kg and consequently results in an increase of 2.60% of the expected WIMP flux after all cuts.

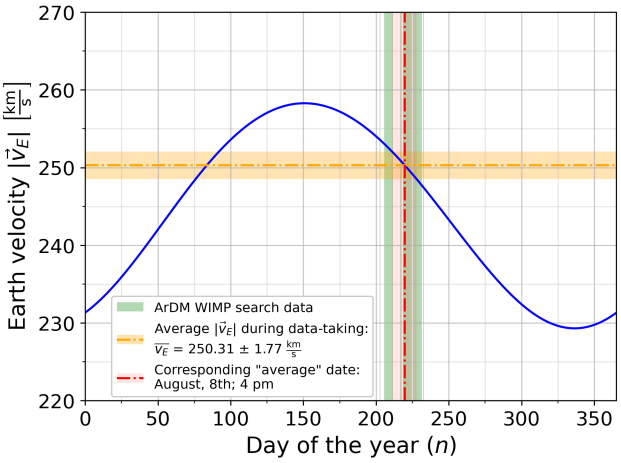

Figure 6.1: The absolute value $|\vec{v}_E|$ of the Earth velocity in the frame of the Milky Way galaxy (blue, from Eqs. 1.4 to 1.8) as a function of the day of the year based on specific models [55, 56]. January, 1st corresponds to the 1st; July, 1st to the 182nd and December, 31st to the 365th day of the year. The average Earth velocity during the collection of the ArDM WIMP search data (green) accounts to a value of $\overline{v_E} = 250.3 \pm 1.8$ km/s (orange). In a justified simplification, this average is taken as the constant value of the Earth velocity during the entirety of the WIMP search data. It corresponds to an "average" date of data-taking of August, 8th, 16:00 (rounded to the hour).

Next, the effective electric drift field strength F_{Drift}, felt by the drifting electrons, has been overestimated to be 230.0 V/cm in the projected sensitivity (Section 2.2.2). This planned drift field strength would have resulted in an electron drift velocity of $v_{\text{Drift}} = 1.25$ mm/μs. However, the observed reduced drift velocity of $v_{\text{Drift}} = 1.065 \pm 0.003$ mm/μs gauges the effective drift field strength, felt by the electrons, to $F_{\text{Drift}} = 152.1 \pm 1.4$ V/cm (see Figs. 4.11 and 4.12, Section 4.2). The direct implications of this have already been discussed in detail in Section 4.2. The slower drift of the electrons increases the percentage of electrons, which are absorbed along the way, and consequently the S2L attenuation (see Section 4.1.2). This results in a larger loss of S2 signals; especially those, that origin from the lower parts of the detector. The exact loss of events is not easily quantifiable, but implicitly incorporated in the matching and the resulting WIMP-affecting cut efficiency (which will be shown in Fig. 6.3). However, the different drift field strength F_{Drift} also affects the recombination quenching q_{rec} (see Eqs. 2.15 to 2.18, Section 2.2.1), which is a linear factor in the conversion of the nuclear recoil energy E_{NR} into the observed primary scintillation light S1L (see Eq. 2.9, Section 2.2). Fig. 6.2 shows the recombination quenching q_{rec} (left) and the resulting total quenching (right) based on the updated value of the effective drift field strength as $F_{\text{Drift}} = 152.1$ V/cm. The effect is largest at low nuclear recoil energies ($E_{\text{NR}} \leq 50$ keV), however, it is mitigated by the very small WIMP-affecting cut efficiencies at these nuclear recoil energies. At higher nuclear recoil energies of $E_{\text{NR}} \geq 50$ keV, the change in the effective drift field strength only affects the S1L scale at the order of $\sim$1%. Consequently, the explicit implications on the WIMP spectrum, caused by the different effective drift field strength, are negligible.

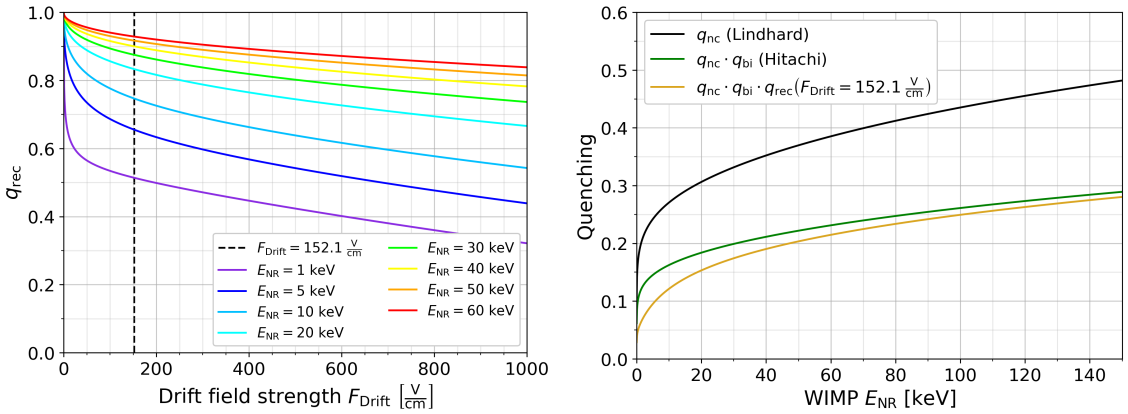

Figure 6.2: *Left*: The recombination quenching q_{rec} of the S1 signals due to the presence of an electric drift field for various nuclear recoil energies E_{NR} (see Eqs. 2.15 to 2.18, Section 2.2.1). The effective drift field strength felt by the electrons was extracted to be $F_{\text{Drift}} = 152.1 \pm 1.4$ V/cm (see Figs. 4.11 and 4.12, Section 4.2) instead of the assumed 230.0 V/cm in the projected sensitivity (Section 2.2.2). *Right*: The combination of the nuclear (q_{nc}, black, Eqs. 2.11 and 2.12, Section 2.2.1), the bi-excitonic (q_{bi}, green, Eq. 2.14, Section 2.2.1) and the drift-field-dependent recombination quenching (q_{rec}, gold, Eqs. 2.15 to 2.18, Section 2.2.1) as a function of the nuclear recoil energy E_{NR}. The change in the effective drift field strength F_{Drift} leads to a negligible effect on the S1L scale (via Eq. 2.9, Section 2.2) of the order of $\sim$1%.

Finally, the light yield has been underestimated as LY = 1.33 p.e./keV in the projected

sensitivity (Section 2.2.2). The LY is a linear factor in the conversion from the nuclear recoil energy E_{NR} to the primary scintillation light S1L (Eq. 2.9, Section 2.2). The gauging of the S1L scale via the β^- spectrum of the ^{39}Ar isotope (Section 5.1) extracted the light yield at a significantly higher value of LY $= 1.82 \pm 0.13$ p.e./keV. This consequently rescales the E_{NR} scale relative to the S1L scale and allows to resolve lower nuclear recoil energies more efficiently. Effectively, this shifts events with certain nuclear recoil energies to higher bins of the S1L. As the combined WIMP-affecting cut efficiency (see Fig. 6.3) is increasing with the S1L, this increases the expected WIMP spectrum. After all applied cuts, the expected WIMP flux is increased by 81.9% in the energy range of interest up to 100 photoelectrons of S1L due to the increase of the LY.

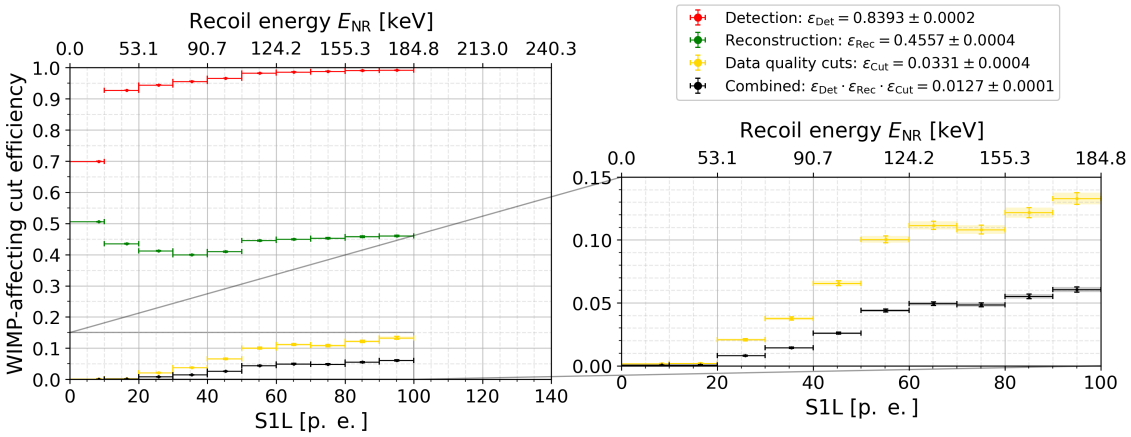

Figure 6.3: The WIMP-affecting cut efficiencies as a function of the energy, shown in S1L bins of ten photoelectrons. While the detection (red) and reconstruction (green) efficiencies are extracted from the ^{252}Cf Monte Carlo simulation directly (Section 5.2.1), the efficiency of the data analysis cuts (yellow) is inferred by the comparison to the extracted ^{252}Cf events in the Neutron calibration data after the application of all cuts (Section 5.2.2). Between 0 and 10, and 10 and 20 photoelectrons, the combined cut efficiency is close to $\varepsilon = 0.0$ with $\varepsilon = 6.08 \cdot 10^{-4}$ and $\varepsilon = 7.63 \cdot 10^{-4}$ respectively. It increases roughly linearly with the primary scintillation light S1L, until finally, in the S1L bin of 90 to 100 photoelectrons, the combined efficiency reaches $\varepsilon = 0.061$.

Fig. 6.3 shows the WIMP-affecting cut efficiencies that have been concluded based on the study of the Neutron calibration data (Section 5.2). In the simplified assumption for the projected ArDM sensitivity (Section 2.2.2), the total cut efficiency had been estimated to be $\varepsilon = 0.0$ below a threshold of 10, 20 or 30 photoelectrons, and $\varepsilon = 0.1$ above that threshold. Naturally, the actual WIMP-affecting cut efficiencies, inferred from the data analysis of the Neutron calibration data and the ^{252}Cf Monte Carlo simulation, are more complicated. The detection (red) and reconstruction (green) efficiencies are extracted from the ^{252}Cf Monte Carlo simulation directly (Section 5.2.1). The efficiency of the data analysis cuts (yellow) are inferred by the comparison of the extracted ^{252}Cf events in the Neutron calibration data, after the application of all cuts, with the expectation of ^{252}Cf events before any cuts from the Monte Carlo simulation (Section 5.2.2). The combined cut efficiency (black) increases roughly linearly as a function of the primary scintillation light S1L. Between 0 and 10, and 10 and 20 photoelectrons, the combined cut efficiency is close to $\varepsilon = 0.0$ with $\varepsilon = 6.08 \cdot 10^{-4}$

and $\varepsilon = 7.63 \cdot 10^{-4}$ respectively. Finally, in the S1L bin of 90 to 100 photoelectrons, the combined efficiency reaches $\varepsilon = 0.061$. As discussed in the evaluation of the cut efficiencies on the DP calibration data (Section 5.4, and specifically Fig. 5.20), the main reduction factor in the data analysis cuts comes from the matching, with an individual efficiency of $\varepsilon = 0.115$ (see also Fig. 5.16). While the matching efficiency, evaluated on the DP calibration data, can not be translated directly to the WIMP-affecting matching efficiency, it serves as a reasonable estimation. Consequently, the assumed combined efficiency ε of 10% (above a threshold) in the projected sensitivity (Section 2.2.2) was an overestimation.

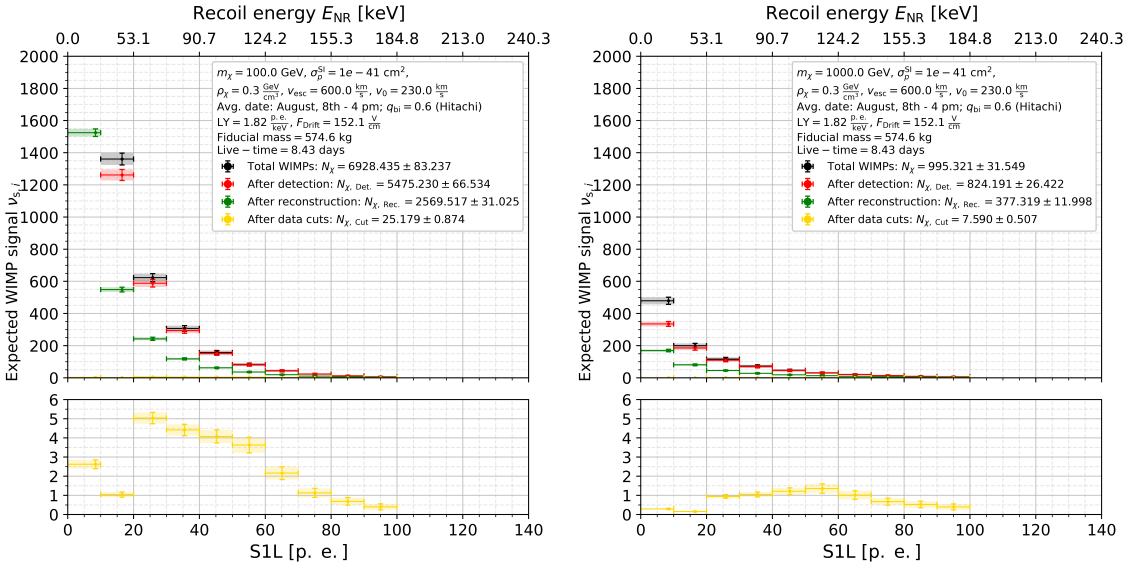

Figure 6.4: The expected signal for a hypothesized 100 GeV (left) and 1 TeV (right) WIMP mass m_χ at a reference spin-independent WIMP-nucleon cross section of $\sigma_p^{SI} = 10^{-41}$ cm^2. In the S1L bin of zero to ten photoelectrons and the case of the 100 GeV WIMP, the shown scale omits the expected 4'313 ± 66 total WIMPs, and the expected 3'015 ± 46 detected WIMPs. The resulting expected WIMP spectra, after all cuts (yellow), are a combination of the exponential WIMP spectrum (black) and the ArDM cut efficiency (Fig. 6.3, black). The upper limit on the signal strength will be constrained in Section 6.3 for a range of WIMP masses from 1 GeV to 10 TeV.

Having incorporated all correct values of the experimental parameters, Fig. 6.4 finally shows the expected WIMP spectrum for a hypothesized 100 GeV (left) and 1 TeV (right) WIMP mass m_χ at a reference spin-independent WIMP-nucleon cross section of $\sigma_p^{SI} = 10^{-41}$ cm^2. In favor of the presentation of the spectra of the 1 TeV WIMP, the y-axis has been clipped at 2000 events. In the S1L bin of zero to ten photoelectrons and the case of the 100 GeV WIMP, this omits the expected 4'313 ± 66 total WIMPs, and the expected 3'015±46 detected WIMPs from the shown scale. Similarly to the spectrum of the extracted ^{252}Cf events from the Neutron calibration data (see Fig. 5.11, Section 5.2.2), the expected WIMP spectrum after all cuts (yellow) is a combination of an exponential WIMP spectrum (black) with a roughly linear cut efficiency (Fig. 6.3, black). While the harsh cut efficiencies below 20 photoelectrons diminish the expected WIMP signal significantly, the spectrum follows a roughly linearly decreasing function above 20 photoelectrons. In comparison to the lighter 100 GeV WIMP (left), the expected spectrum of the heavier

1 TeV WIMP (right) is shifted to higher values of the energy and diminishes less with increasing energy. While the 100 GeV WIMP spectrum is covered significantly by the incorporated S1L bins up to 100 photoelectrons, the 1 TeV WIMP spectrum motivates the future inclusion of even higher S1L bins into the analysis. The choice of the reference cross section of $\sigma_p^{\mathrm{SI}} = 10^{-41}$ cm^2 is justified retrospectively, as the upper limit will approximately be constrained where the expected WIMP signal is at the order of three events plus the expected background. This concludes the discussion of the theoretical expectation of the WIMP spectrum. The expected background and the observed events will be extracted in Section 6.2, while the upper limit on the signal strength will be constrained in Section 6.3 for a range of WIMP masses from 1 GeV to 10 TeV.

6.2 The data perspective: Expected background and observed events

The confinement of the RoI (Section 5.2.2) completed the final list of all cuts (see Tables 6 and 7, Chapter 5). As such, the WIMP search simplifies to a counting experiment. The DP calibration data serves as the control experiment, where the absence of signal is assumed. The observed events therein, after all applied cuts, gauge the expected instrumental background of the WIMP search. The WIMP search data itself, hypothetically consists of the instrumental background, and a potential signal. The observed events within the WIMP search data, after all applied cuts, will then be used in a statistical analysis, together with the expected instrumental background, to constrain the strength of the potential signal in Section 6.3.

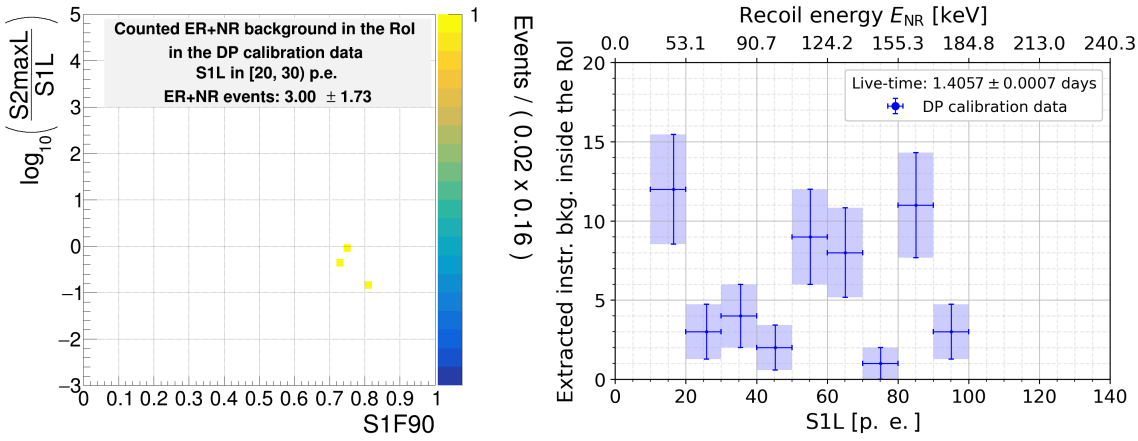

Figure 6.5: The counted estimation of the instrumental DP background in the 1.406 live-days of the hypothetically signal-free DP calibration data. The specific evaluation in the example S1L bin of 20 to 30 photoelectrons (left) and the resulting spectrum up to an S1L of 100 photoelectrons (right). Due to the relaxed nature of the RoI cut below ten photoelectrons (see Tables 6 and 7, Chapter 5), the resulting counted estimation of 413 events of instrumental background is far outside the plotting scale.

Firstly, Fig. 6.5 (left) shows the counted expectation of the instrumental background within the S1L bin of 20 to 30 photoelectrons, based on the DP calibration data and the respective 1.406 live-days. In this S1L bin, the expected instrumental background amounts to three events with a Poissonian error. Fig. 6.5 (right) shows the resulting spectrum of the expected instrumental background in all S1L bins. As discussed in Section 5.2.2, the

NR and ER regions completely overlap at an S1L below ten photoelectrons. Consequently, the RoI cut excludes no events below ten photoelectrons (see Tables 6 and 7, Chapter 5) and the resulting counted background estimation of 413 events is far outside the plotting scale. Above ten photoelectrons, the counted background estimation is reasonably small with a maximum of twelve events per S1L bin. Naturally, there is a trade-off involved in the exact confinement of the cuts. A more restrictive RoI cut would remove more instrumental background, but also more of the WIMP signal region. As detailed in Section 5.2.2, the RoI cut, with the help of the Neutron calibration data, has been optimized to regions of very high ratios of the estimated likelihoods of NRs over ERs.

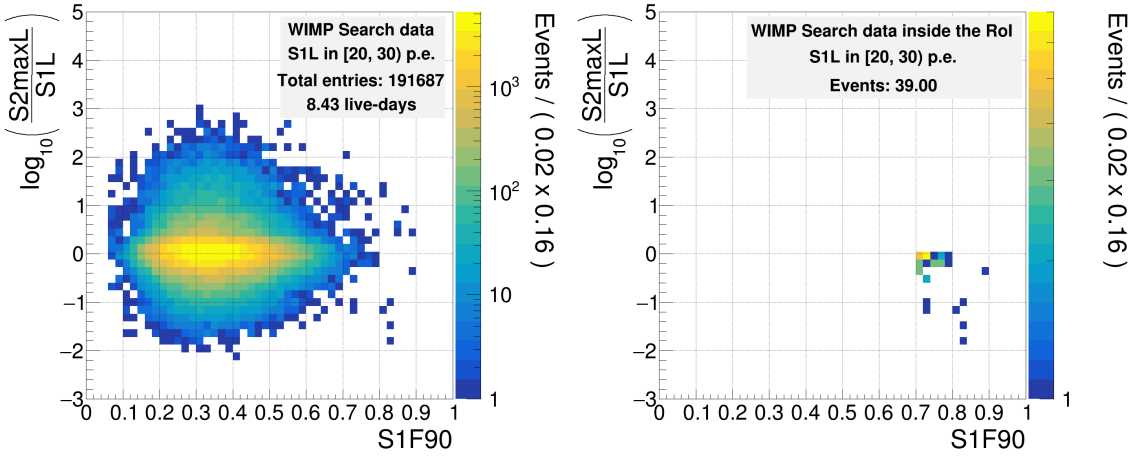

Figure 6.6: The 8.429 live-days of the un-blinded WIMP search data, in the example S1L bin of 20 to 30 photoelectrons, after the fiducial volume cut (left), and after the subsequent RoI cut (right) applied according to Tables 6 and 7 (in Chapter 5). The number of observed events in the RoI is compared to the scaled expectation of the instrumental DP background in Fig. 6.7.

Secondly, Fig. 6.6 shows the 8.429 live-days of the un-blinded WIMP search data after the fiducial volume cut (left), and after the subsequent RoI cut (right), applied according to Tables 6 and 7 (in Chapter 5). In the example bin of 20 to 30 photoelectrons, the fiducial volume cut results in 191'687 events, while the RoI cut is survived by just 39 events. In this specific bin, the tail of the instrumental background distribution appears to leak into the RoI. This is a potential danger of the RoI "box approach" (see Section 5.2.2) in S1L bins with a partial overlap of the NR and ER regions. Essentially, the instrumental background in the WIMP search data (8.429 live-days) is the same distribution as in the DP calibration data (1.406 live-days), just sampled more often. Naturally, the tail of the distribution is expected to be larger. However, based on the underlying Neutron and DP calibration data, which guided the cut confinement, there was no implication to constrain the respective RoI further (see Figs. 5.7 and 5.9, Section 5.2.2).

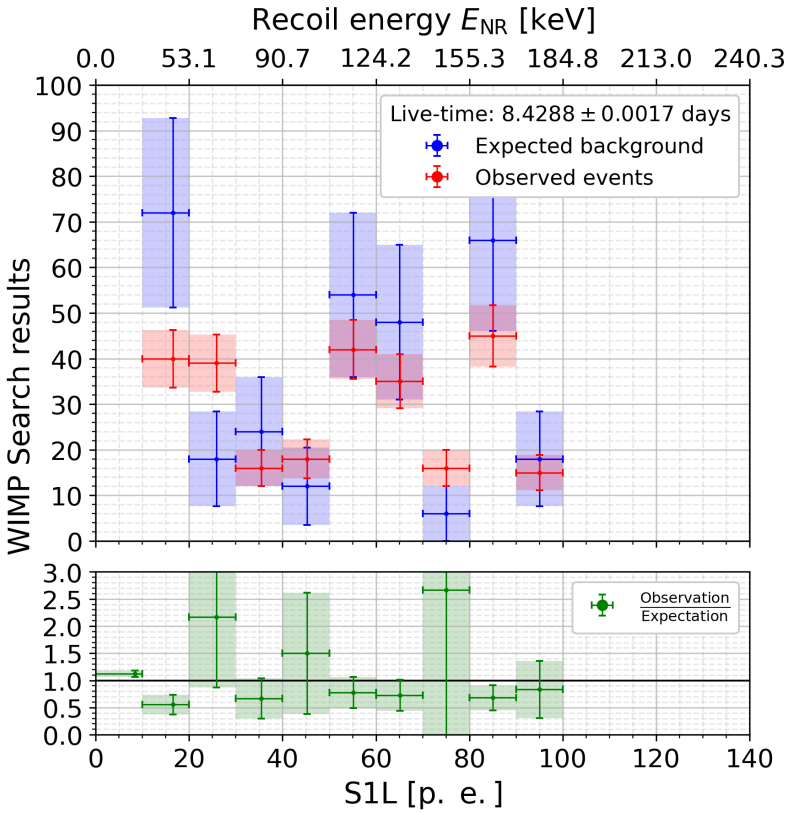

Figure 6.7: The WIMP search results up to 100 photoelectrons of S1L based on the up-scaled expected background from the DP calibration data of 1.406 live-days (see Fig. 6.5); and the actually observed events in the WIMP search data of 8.406 live-days (see the example S1L bin of 20 to 30 photoelectrons in Fig. 6.6). Due to the relaxed nature of the RoI cut below ten photoelectrons, the respective expected background of 2'476 ± 122 events, and the respective 2'780 ± 53 observed events in the lowest S1L bin are far outside the plotting range. Generally, the ratios of the observation versus the expectation (green) are well centered around unity. The specific cases, where the errors (green, shaded) do not include unity, can be explained by the lack of the incorporation of the systematic error of the RoI cut into the estimation of the expected background. These results, together with the theoretical estimation of the WIMP spectrum (Section 6.1), are subsequently used to constrain the signal strength statistically and set an upper limit in Section 6.3.

Finally, Fig. 6.7 presents the results of the WIMP search as a counting experiment, evaluated up to 100 photoelectrons. The expected background (blue) stems from the observed instrumental background in the DP calibration data of 1.406 live-days (see Fig. 6.5), scaled up to the WIMP search data of 8.429 live-days. The observed events (red) are directly counted in the WIMP search data (see the example S1L bin of 20 to 30 photoelectrons in Fig. 6.6). Due to the relaxed nature of the RoI cut below ten photoelectrons, the respective expected background of 2'476 ± 122 events and the respective 2'780 ± 53 observed events in the lowest S1L bin are far outside the plotting range. Subsequently, this leads to a small observation excess (green) in the first S1L bin. However, in the second and second-to-last S1L bins, the expectation exceeds the observation. Overall, the ratios of the observation

versus the expectation are well centered around unity. The fact, that the statistical error (green, shaded) doesn't always include unity shouldn't be frowned upon. This is caused by the fact, that the error on the expected background (blue, shaded) only takes the statistical uncertainty into account. Strictly speaking, a thorough study of the systematic error would be necessary to judge any significant observation excess. In the WIMP search, the main driver of the systematic error would be the uncertainty connected to the box-like RoI cut, as the specific cut values can significantly influence the result. However, in the interest of brevity, and the fact that ArDM only aims to constrain an upper limit and not claim a discovery, a very detailed systematic uncertainty estimation based on the RoI cut is left out. The results of the WIMP search, together with the theoretical estimation of the WIMP spectrum (Section 6.1), are subsequently used to constrain the signal strength statistically and set an upper limit to it in Section 6.3.

6.3 The WIMP exclusion limit of ArDM - A statistical analysis

With the fine-tuning of the theoretically expected WIMP spectrum (Section 6.1) and the extraction of the expected background as well as the observed events in the WIMP search data (Section 6.2), all ingredients are ready to be converted into a statistical statement on the WIMP exclusion limit of ArDM. The general statistical procedure and underlying WIMP model largely follows the recent (2021) reporting recommendations of Baxter et al. [107]. As discussed previously, the WIMP hypothesis has two remaining unknowns; the WIMP mass m_χ, and the spin-independent WIMP-nucleon cross section σ_p^{SI}. The rate of WIMP interactions scales linearly with the spin-independent WIMP-nucleon cross section σ_p^{SI} and depends on the WIMP mass m_χ in a much more complex manner (see Eq. 1.42, Section 1.3). Consequently, the expected WIMP spectrum (Section 6.1) also has to be re-evaluated for every specific value of m_χ and σ_p^{SI}. To minimize the required computational effort, the linear dependence of the WIMP spectrum on σ_p^{SI} is exploited in the so-called "raster scan". Therein, the theoretical WIMP spectrum is evaluated for a specific value of m_χ and a reference cross section of $\mu \cdot \sigma_{\mathrm{Ref}}$. The floating parameter μ refers to the signal strength and defaults to unity. Incorporating the expected background and the observed events, the statistical analysis procedure then constrains the signal strength $\mu \rightarrow \mu_{\mathrm{Constrained}}$ to the value corresponding to the upper limit at the desired confidence level. Therefore the upper limit on the cross section then corresponds to $\sigma_p^{\mathrm{SI}} = \mu_{\mathrm{Constrained}} \cdot \sigma_{\mathrm{Ref}}$ for the specific WIMP mass m_χ. Subsequently, this process is repeated for a large number of WIMP masses m_χ, in the mass range of interest, to complete the raster scan. Section 6.3.1 accompanies the reader through the statistical testing process for a 100 GeV WIMP candidate. Thereafter, Section 6.3.2 presents and discusses the resulting WIMP exclusion limit of ArDM after the raster scan in the plane of the σ_p^{SI} versus the m_χ. This will complete the ArDM baseline WIMP search. After the conclusion in Section 6.3.2, a preview to the potential short- and long-term improvement plans will be given in the Outlook.

6.3.1 The statistical testing process of a single WIMP mass

This section walks the reader through the statistical process of confining an upper limit on the spin-independent WIMP-nucleon cross section σ_p^{SI} for a 100 GeV WIMP candidate. The required ingredients are the expected background and the observed events in the WIMP search data (Section 6.2), as well as the to-be-constrained WIMP spectrum. The respective theoretical 100 GeV WIMP spectrum ν_{s} (Section 6.1, Fig. 6.4, left, yellow) is based on

a reference cross section of $\mu \cdot \sigma_{\mathrm{Ref}}$, with σ_{Ref} fixed at 10^{-41} cm^2 and μ defaulting to unity. The statistical process consists of two main parts. Firstly, the discovery test, which studies the statistical significance of potential high event fluctuations above the expected background level. According to the recommended reporting standards [107], a statistical signal significance of at least 3σ is required to claim evidence (and at least 5σ for a claim of discovery following the convention of particle physics). Secondly, upon refusal of the discovery claim, inverted hypothesis testing allows to constrain a maximal value of μ, which is consistent with the observed fluctuations at the desired confidence level. This constrained value then directly gauges the upper limit on the cross section as $\sigma_p^{\mathrm{SI}} = \mu_{\mathrm{Constrained}} \cdot \sigma_{\mathrm{Ref}}$ for this respective WIMP mass m_χ. The repetition of this statistical testing procedure for various WIMP masses then results in an exclusion limit in the plane of σ_p^{SI} versus m_χ, which is showcased and discussed in Section 6.3.2.

The presented likelihood approach follows the recommended statistical testing procedure for direct WIMP searches [107]. Therein, a statistical model, in the form of a likelihood, describes the statistical behavior of the experimental results, the observables, as a function of the underlying parameters. A likelihood is a, not necessarily normalized, combination of statistical distributions. Each distribution describes the statistical behavior of one or more observables as a function of the underlying parameters. The entire model then measures the likelihood of a specific set of parameters resulting in a specific set of observables or vice versa. Let's consider a simplified case of the ArDM experiment, where all ten S1L bins are combined to a single bin. The observables then are the measured values of the total observed events N_{Obs} and the total expected background N_{Bkg}. As ArDM is fundamentally a counting experiment, both observables follow Poisson distributions. The underlying rate of the background Poisson process is labeled ν_{b}, whereas the rate of the "WIMP signal and background" Poisson process is given by $\mu\nu_{\mathrm{s}}+\nu_{\mathrm{b}}$, therewith incorporating a signal source ν_{s} with signal strength μ, as desired. In the WIMP search, the parameter ν_{s} is fixed through the theoretical WIMP spectrum, while the true background rate ν_{b} and the signal strength μ are unknown and have to be constrained. While the signal strength μ is the primary parameter of interest, the background rate ν_{b} is considered a nuisance parameter, which has to be accounted for, but is not a parameter one is fundamentally interested in. The resulting likelihood $\mathcal{L}\left(\mu, \nu_{\mathrm{b}} | \nu_{\mathrm{s}}, N_{\mathrm{Obs}}, N_{\mathrm{Bkg}}\right)$ of the parameters μ and ν_{b} given a fixed signal reference rate ν_{s} and a set of observables N_{Obs} and N_{Bkg} accounts to

$$\mathcal{L}\left(\mu, \nu_{\mathrm{b}} | \nu_{\mathrm{s}}, N_{\mathrm{Obs}}, N_{\mathrm{Bkg}}\right) = \frac{\left(\mu\nu_{\mathrm{s}} + \nu_{\mathrm{b}}\right)^{N_{\mathrm{Obs}}}}{N_{\mathrm{Obs}}!} e^{-(\mu\nu_{\mathrm{s}}+\nu_{\mathrm{b}})} \cdot \frac{\left(\nu_{\mathrm{b}}\right)^{N_{\mathrm{Bkg}}}}{N_{\mathrm{Bkg}}!} e^{-\nu_{\mathrm{b}}}. \tag{6.1}$$

This directly translates into ArDM, as the first Poisson term corresponds to the WIMP search data and the second Poisson term describes the background estimation with the control experiment in form of the DP calibration data. One has to emphasize, that this model does not fully incorporate the uncertainty of the background measurement in the control experiment. As described in Section 6.2, the estimated background N_{Bkg} in the WIMP search is a value extracted from a smaller data set and scaled up according to the live-times. However, the above second Poisson term describes the estimated background N_{Bkg} as a value extracted from a data set of the same live-time as the WIMP search data. As a consequence, the uncertainty of the actually estimated background value is larger by a factor of the square root of the ratio of the live-times (which accounts to $\approx \sqrt{6}$) than what is reflected in the above model. Additionally, the systematic uncertainty, driven by the RoI confinement (as discussed in Section 6.2, with respect to Fig. 6.7), is not accounted

for either. Both sources of additional uncertainty on the background measurement could be incorporated with the extension of ν_b to $\tau\nu_\mathrm{b}$ and a multiplicative Gaussian distribution. The respective Gaussian parameters μ_τ and σ_τ would reflect the degree of belief in the background measurement, with μ_τ fixed to unity and σ_τ set to 0.1, 0.2 or the respective estimated accuracy. However, this would only affect the discovery significance, and since ArDM's expected sensitivity should not lead to a discovery, the discussed extension is left out for the time being.

After the simplified model, the extension to the ten S1L bins of the ArDM experiment is straightforward. Each S1L bin i is incorporated with a multiplicative factor as

$$\mathcal{L}\left(\mu, \nu_{\mathrm{b},i} | \nu_{\mathrm{s},i}, N_{\mathrm{Obs},i}, N_{\mathrm{Bkg},i}\right) = \Pi_i \frac{\left(\mu\nu_{\mathrm{s},i} + \nu_{\mathrm{b},i}\right)^{N_{\mathrm{Obs},i}}}{N_{\mathrm{Obs},i}!} e^{-\left(\mu\nu_{\mathrm{s},i} + \nu_{\mathrm{b},i}\right)} \cdot \frac{\left(\nu_{\mathrm{b},i}\right)^{N_{\mathrm{Bkg},i}}}{N_{\mathrm{Bkg},i}!} e^{-\nu_{\mathrm{b},i}}. \quad (6.2)$$

The corresponding observables as well as the signal and background rates naturally extend through indexing as $\nu_\mathrm{b}, \nu_\mathrm{s}, N_\mathrm{Obs}, N_\mathrm{Bkg} \rightarrow \nu_{\mathrm{b},i}, \nu_{\mathrm{s},i}, N_{\mathrm{Obs},i}, N_{\mathrm{Bkg},i}$. Only the main parameter of interest, the signal strength μ, remains the same in all S1L bins. This can be understood as combining the statistics of ten experiments trying to measure the same parameter of interest. Naturally, the respective values of $N_{\mathrm{Obs},i}$ and $N_{\mathrm{Bkg},i}$ are given by the resulting experimental counts (Fig. 6.7, blue and red respectively). Moreover, the signal rates $\nu_{\mathrm{s},i}$ correspond to the theoretical WIMP spectrum for the studied WIMP mass; as Fig. 6.4 (left, yellow) for the 100 GeV WIMP. Consequently, the binned model results in eleven parameters; one parameter of interest in the form of the signal strength μ and ten nuisance parameters in the separate background rates $\nu_{\mathrm{b},i}$. Analogous to the simplified one-bin model (Eq. 6.1), the additional uncertainty of the background measurements could be reflected by extending the background rates $\nu_{\mathrm{b},i}$ to $\tau_i\nu_{\mathrm{b},i}$ and multiplying with respective Gaussian distributions with the parameters $\mu_{\tau,i}$ and $\sigma_{\tau,i}$. Again, this is currently left out with the main priority being to set a WIMP exclusion limit.

Having defined the ArDM model (Eq. 6.2) appropriately, the likelihood can be studied for its behavior as a function of the parameters. The maximum likelihood estimator (MLE), often denoted with a hat $\hat{\cdot}$, refers to the value of a variable, which maximizes the value of the likelihood $\mathcal{L}$. Consequently, the likelihood $\mathcal{L}\left(\hat{\mu}, \hat{\nu}_{\mathrm{b},i}\right)$, evaluated at the MLE values $\hat{\mu}$ and $\hat{\nu}_{\mathrm{b},i}$, refers to the maximal possible likelihood value with an optimized set of parameters. Or, in other words, the best possible agreement of the parameter values with the data given the underlying model. Being interested in a specific parameter μ, one can "profile" the likelihood as $\mathcal{L}\left(\mu, \hat{\hat{\nu}}_{\mathrm{b},i}\left(\mu\right)\right)$, where the parameter of interest μ is left free, while the nuisance parameters $\hat{\hat{\nu}}_{\mathrm{b},i}\left(\mu\right)$ are optimized as a function of the floating μ. Subsequently, one can calculate the profile likelihood ratio $\lambda\left(\mu\right)$ as a function of the parameter of interest μ as

$$\lambda\left(\mu\right) = \frac{\mathcal{L}\left(\mu, \hat{\hat{\nu}}_{\mathrm{b},i}\left(\mu\right)\right)}{\mathcal{L}\left(\hat{\mu}, \hat{\nu}_{\mathrm{b},i}\right)}, \quad \text{with} \quad t_\mu = -2\ln\left(\lambda\left(\mu\right)\right) \quad (6.3)$$

being the corresponding test statistic t_μ. As the likelihood $\mathcal{L}\left(\hat{\mu}, \hat{\nu}_{\mathrm{b},i}\right)$, evaluated at all MLEs, is the maximal possible likelihood value, the profile likelihood ratio $\lambda\left(\mu\right)$ must be limited between zero and one. Consequently, the test statistic t_μ is a helpful transformation. Its values range from infinity to zero representing an increasing agreement of the parameter value and the data given the underlying model.

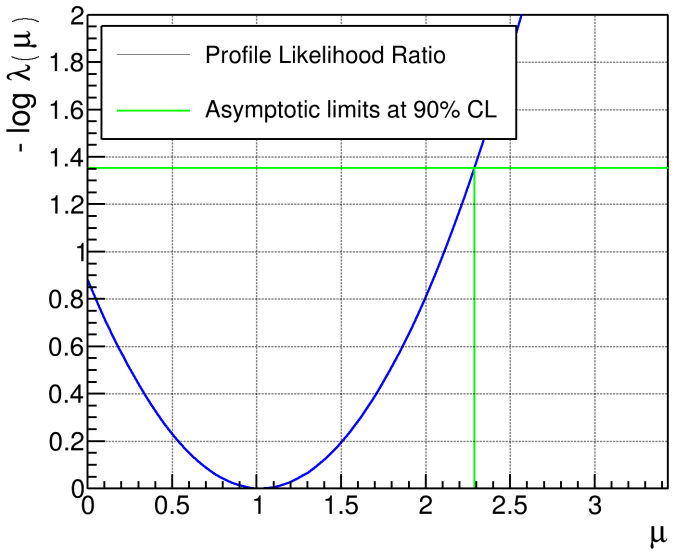

Figure 6.8: The profile likelihood ratio $\lambda(\mu)$ (Eq. 6.3), or the test statistic t_μ (blue) respectively, as a function of the signal strength μ evaluated for the ArDM experiment for a 100 GeV WIMP at the reference cross section $\mu \cdot 10^{-41}$ cm^2. The asymptotic properties of the profile likelihood ratio $\lambda(\mu)$ (Eq. 6.4, Wilk's theorem [108]) guide an estimate of the upper limit at a value of $\mu = 2.29$ at the desired confidence level of 90%. See the text for details.

Fig. 6.8 shows the profile likelihood ratio $\lambda(\mu)$, or the test statistic t_μ (blue) respectively, as a function of the signal strength μ evaluated for the ArDM experiment for a 100 GeV WIMP at the reference cross section $\mu \cdot 10^{-41}$ cm^2. The test statistic t_μ, evaluated at the MLE $\hat{\mu}$, equals zero and is situated around a value of $\mu \approx 1$ in this case. Therefore, a value of $\mu \approx 1$ results in the best agreement of the data and the model. In order to consolidate any statement on a potential upper limit (or a discovery significance), one would need to know the probability density function of the test statistic t_μ. Therewith, one could elaborate which lower range of t_μ would be covered by 90% of repeated experiments given a true value μ'. Naturally, the probability density function $f(t_\mu|\mu')$ of t_μ depends on what the true value μ' corresponds to. For the special case of the tested value μ equaling the true value μ', Wilk's theorem [108] provides a simple approximation as

$$-2\ln(\lambda(\mu)) = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \quad \overset{N\to\infty}{\Rightarrow} \quad f\left(-2\ln(\lambda(\mu))\,|\mu'=\mu\right) = \chi^2_1. \tag{6.4}$$

Here, the MLE $\hat{\mu}$ follows a Gaussian distribution with a mean μ' and standard deviation σ, given by $\sigma^2 = \mathrm{cov}[\hat{\mu},\hat{\mu}]$. N represents the data sample size, and, in the large sample size limit, the distribution $f(t_\mu|\mu'=\mu)$ consequently approaches a χ^2 distribution with one degree of freedom. Consequently, the well-known cumulative distribution function of χ^2_1 allows to make an educated guess of the upper limit on μ at a 90% confidence level; as indicated in Fig. 6.8 (lime). However, it has to be emphasized again, that this simplistic approach is based on the assumptions that the tested value μ corresponds to the true value μ', as well as the data consisting of a large sample size. Both of these assumptions are not appropriate for the WIMP search. Indeed, one would actually like to test the hypothesis of a specific signal strength versus a background-only case. Therefore, one also

needs to know the distribution $f\left(t_\mu|\mu'\right)$ for cases where $\mu' \neq \mu$. While corresponding asymptotic distributions exist, and will be discussed in the following paragraphs, they naturally rely on the large sample size of the data. The actual solution lies in the "Toy Monte Carlo" approach, where a large number of pseudo-experiments are sampled based on a model and a set of parameters. Therein, the background measurements $N_{\mathrm{Bkg},i}$ are promoted to global variables, which are kept constant, whereas the observed events $N_{\mathrm{Obs},i}$ are resampled in every toy experiment. The distribution of the resulting test statistics t_μ consequently approximates the distribution $f\left(t_\mu|\mu'\right)$. What follows in the next paragraphs, is a discussion of the asymptotic behavior of the distributions of the test statistics relevant to the WIMP search. All corresponding distributions will then be shown next to the Monte Carlo toy-sampled distributions. After an introduction to the Asimov data set, the discovery significance will be evaluated, followed by the setting of an upper limit for the 100 GeV WIMP through inverted hypothesis testing. The majority of the upcoming paragraphs is treated in great detail in various sources [109,110], and is only rephrased here based thereon. The general underlying Wald theorem [111] allows the calculation of the distributions $f\left(t_\mu|\mu'\right)$ in the asymptotic limit and is not covered here.

The Asimov data set and the signal profile likelihood ratio: The asymptotic behavior of the distributions $f\left(t_\mu|\mu'\right)$ relies on an artificial data set called the Asimov data set. The Asimov data set, first justified by Cowan et al. in 2010 [109], corresponds to the artificial data of the median experiment, where all observables equal their expected values based on a given set of parameters. In the WIMP search, this would specifically result in the Asimov data of

$$N_{\mathrm{Bkg},i,\mathrm{A}} = \nu_{\mathrm{b},i} \quad \text{and} \quad N_{\mathrm{Obs},i,\mathrm{A}} = \mu\nu_{\mathrm{s},i} + \nu_{\mathrm{b},i}. \tag{6.5}$$

Subsequently, the corresponding "Asimov likelihood" $\mathcal{L}_\mathrm{A}$ and the resulting profile likelihood ratio λ_A evaluate to

$$\lambda_\mathrm{A}\left(\mu\right) = \frac{\mathcal{L}_\mathrm{A}\left(\mu, \hat{\hat{\nu}}_{\mathrm{b},i}\left(\mu\right)\right)}{\mathcal{L}_\mathrm{A}\left(\mu', \nu_{\mathrm{b},i}\right)}. \tag{6.6}$$

Independently, Wald's theorem [111] generalizes the first part of Wilk's theorem (Eq. 6.4) for any μ unrelated to μ' as

$$-2\ln\left(\lambda\left(\mu\right)\right) = \frac{\left(\mu - \hat{\mu}\right)^2}{\sigma^2} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right). \tag{6.7}$$

Consequently, for large sample sizes, and the Asimov data set, where $\hat{\mu} = \mu'$ per definition, the corresponding Asimov data test statistic $t_{\mu,\mathrm{A}}$ simplifies to

$$t_{\mu,\mathrm{A}} = -2\ln\left(\lambda_\mathrm{A}\left(\mu\right)\right) \approx \frac{\left(\mu - \mu'\right)^2}{\sigma^2} \quad \Rightarrow \quad \sigma^2 \approx \frac{\left(\mu - \mu'\right)^2}{t_{\mu,\mathrm{A}}}. \tag{6.8}$$

Hence, the true power of the Asimov data lies in the simplicity of estimating the standard deviation σ of the assumed Gaussian distribution of $\hat{\mu}$. The estimation of σ through the Asimov data set will be used subsequently in the asymptotic approximations of the distributions $f\left(t_\mu|\mu'\right)$ for a tested value μ and a true value μ'. Therefore, in the mathematical

context, the general profile likelihood ratio (Eq. 6.3) is often adapted for the signal search as

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{\mathcal{L}\left(\mu, \hat{\hat{\nu}}_{\mathrm{b},i}(\mu)\right)}{\mathcal{L}\left(\hat{\mu}, \hat{\nu}_{\mathrm{b},i}\right)} & \text{for } \hat{\mu} \geq 0, \\ \frac{\mathcal{L}\left(\mu, \hat{\hat{\nu}}_{\mathrm{b},i}(\mu)\right)}{\mathcal{L}\left(0, \hat{\nu}_{\mathrm{b},i}(0)\right)} & \text{for } \hat{\mu} < 0. \end{cases} \tag{6.9}$$

This signal profile likelihood ratio $\tilde{\lambda}(\mu)$ specifically does not extend to $\hat{\mu} < 0$. This is only important in the context of the asymptotic approximations, as the toy Monte Carlo approach naturally prohibits a negative signal strength.

The discovery significance: The first of two test cases concerns the potential discovery. Therefore, the tested value of μ is strictly zero. The respective null hypothesis H_0 reflects the background-only scenario with the true signal strength $\mu' = 0$. If it were rejected, the alternative hypothesis $\mathrm{H}_{\mathrm{Alt}}$ would be a non-zero signal strength μ'.

The concerning test statistic t_0 for $\mu = 0$ is commonly adapted and relabeled as the one-sided test statistic $\tilde{q}_0$, which specifically equates to

$$\tilde{q}_0 = \begin{cases} -2 \ln\left(\tilde{\lambda}(0)\right) & \text{for } \hat{\mu} \geq 0, \\ 0 & \text{for } \hat{\mu} < 0. \end{cases} \tag{6.10}$$

The resulting probability density function $f(\tilde{q}_0 | \mu' = 0)$ of H_0, in the asymptotic limit, approximates to

$$f\left(\tilde{q}_0 | \mu' = 0, \hat{\hat{\nu}}_{\mathrm{b},i}(\mu' = 0)\right) \approx \frac{1}{2}\delta(\tilde{q}_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{q}_0}}e^{-\frac{1}{2}\tilde{q}_0}, \tag{6.11}$$

consisting of a delta function $\delta(\cdot)$ and a half chi-square distribution $\frac{1}{2}\chi^2_1$ with one degree of freedom. Fig. 6.9 (left) shows a histogram (black) of 2'000 toy Monte Carlo pseudo-experiments (with resampled $N_{\mathrm{Obs},i}$) and their resulting test statistic values $\tilde{q}_0$, given that $\mu' = 0$. Additionally, the asymptotic approximation (Eq. 6.11) is plotted on top (red), resulting in an acceptable agreement with the toy-sampled distribution. However, for the purpose of the ArDM data analysis, the distribution $f(\tilde{q}_0 | \mu' = 0)$ is approximated by the histogram of the Monte Carlo toy experiments. The values of the test statistic $\tilde{q}_{0,\mathrm{Obs}}$ of the actually observed values $N_{\mathrm{Obs},i}$ is also indicated (blue). Naturally, one can evaluate the corresponding p-value p_0 of H_0, or the probability of a more extreme observation than $\tilde{q}_{0,\mathrm{Obs}}$, as

$$p_0 = P(\tilde{q}_0 \geq \tilde{q}_{0,\mathrm{Obs}} | \mu' = 0) = \int_{\tilde{q}_{0,\mathrm{Obs}}}^{\infty} f\left(\tilde{q}_0 | \mu' = 0, \hat{\hat{\nu}}_{\mathrm{b},i}(\mu' = 0)\right) \mathrm{d}\tilde{q}_0. \tag{6.12}$$

In the asymptotic approximation (Eq. 6.11), the p-value p_0 of H_0 can be evaluated as

$$p_0 = 1 - F(\tilde{q}_{0,\mathrm{Obs}} | \mu' = 0) \approx 1 - \Phi\left(\sqrt{\tilde{q}_{0,\mathrm{Obs}}}\right), \tag{6.13}$$

where $F(\tilde{q}_{0,\mathrm{Obs}} | \mu')$ is the cumulative distribution function of $f(\tilde{q}_0 | \mu')$. In the case of $\mu' = 0$, the cumulative distribution function $F(\tilde{q}_{0,\mathrm{Obs}} | \mu' = 0)$ simply corresponds to the cumulative distribution function $\Phi(\cdot)$ of the Normal distribution. In the ArDM experiment, for the WIMP mass of 100 GeV, the p-value evaluates to $p_0 = 0.094$ (Fig. 6.9, left, grey shaded) based on the 2'000 Monte Carlo toy experiments. This signifies, that only an expected 9.4%

of repeated experiments would result in more extreme values of the observed $N_{\mathrm{Obs},i}$, given that $\mu = \mu' = 0$.

The respective probability density function $f\left(\tilde{q}_0|\mu'\right)$ for the alternative hypothesis $\mathrm{H_{Alt}}$ of $\mu' \neq 0$ can be asymptotically approximated as

$$f\left(\tilde{q}_0|\mu'\right) \approx \left(1 - \Phi\left(\frac{\mu'}{\sigma}\right)\right)\delta\left(\tilde{q}_0\right) + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{q}_0}}e^{-\frac{1}{2}\left(\sqrt{\tilde{q}_0}-\frac{\mu'}{\sigma}\right)}, \quad \text{where} \quad \frac{\mu'}{\sigma} \overset{\mu=0}{=} \sqrt{\tilde{q}_{0,\mathrm{A}}}\,, \tag{6.14}$$

as the standard deviation σ can be evaluated with the corresponding test statistic $\tilde{q}_{0,\mathrm{A}}$ of the Asimov data set (see Eq. 6.8). In fact, Eq. 6.14 accounts for the general case of any μ' and the asymptotic approximation of $\mathrm{H_0}$ (Eq. 6.11) can be retrieved by setting $\mu' = 0$. Consequently, the p-value p_{Alt} of the alternative hypothesis $\mathrm{H_{Alt}}$ is evaluated by

$$p_{\mathrm{Alt}} = P\left(\tilde{q}_0 \geq \tilde{q}_{0,\mathrm{Obs}}|\mu' \neq 0\right) = \int_{\tilde{q}_{0,\mathrm{Obs}}}^{\infty} f\left(\tilde{q}_0|\mu', \hat{\hat{\nu}}_{\mathrm{b},i}\left(\mu'\right)\right)\mathrm{d}\tilde{q}_0. \tag{6.15}$$

In the asymptotic approximation (Eq. 6.14), this accounts to

$$p_{\mathrm{Alt}} = 1 - F\left(\tilde{q}_{0,\mathrm{Obs}}|\mu' \neq 0\right) \approx 1 - \Phi\left(\sqrt{\tilde{q}_{0,\mathrm{Obs}}} - \frac{\mu'}{\sigma}\right). \tag{6.16}$$

Again, this approximation actually holds for the general case of any μ' and the special case of p_0 (Eq. 6.13) can be retrieved with setting $\mu' = 0$. Fig. 6.9 (right) shows the Monte Carlo toy-sampled distributions of $f\left(\tilde{q}_0|\mu' = 0\right)$ ($\mathrm{H_0}$, red) and $f\left(\tilde{q}_0|\mu' = 2.29\right)$ (a specific $\mathrm{H_{Alt}}$, blue) and the respective asymptotic approximations (Eq. 6.11 and Eq. 6.14 respectively). The specific value of $\mu' = 2.29$ has been chosen as the upper limit guess based on the asymptotic approximation of the signal profile likelihood ratio (Fig. 6.8). The respective distributions, relative to the ArDM experiment (lime), result in $p_0 = 0.082$ (red, shaded) and $p_{\mathrm{Alt}} = 0.955$ (blue, shaded) based on 2'000 toy experiments each. This signifies, that under the background-only hypothesis $\mathrm{H_0}$, an estimated 8.2% of repeated experiments would result in a more extreme observation than ArDM. The small difference to the previously estimated $p_0 = 0.094$ (in Fig. 6.9, left, grey shaded) motivates, that more than 2'000 toy experiments are necessary to attain a reasonable precision in the test of $\mathrm{H_0}$. In the case of the alternative hypothesis $\mathrm{H_{Alt}}$ and $\mu' = 2.29$, an estimated 95.5% of repeated experiments would result in a more extreme observation than ArDM. Explicitly, the ArDM results favor $\mathrm{H_{Alt}}$ over $\mathrm{H_0}$, but not enough to reject $\mathrm{H_0}$. A rejection of $\mathrm{H_0}$, or the claim of signal, would require a minimal significance of 3σ ($p_0 = 1.4 \cdot 10^{-3}$) or 5σ ($p_0 = 2.9 \cdot 10^{-7}$), depending on the convention. Nevertheless, the low p-value p_0, observed in ArDM, is a natural consequence of the observed over-fluctuation in some S1L bins (see Fig. 6.7). Additionally, as previously discussed, the ArDM model (Eq. 6.2) does not include a complete treatment of the systematic uncertainty of the RoI box approach; and the statistical error on the background measurement is underestimated by a factor $\approx \sqrt{6}$ due to the live-time rescaling of the data. Such an inclusion would increase the fluctuation of the expected background, and hence render the ArDM observation less extreme. This concludes the study of the discovery significance of the ArDM analysis result, as the background-only hypothesis $\mathrm{H_0}$ can not be rejected.

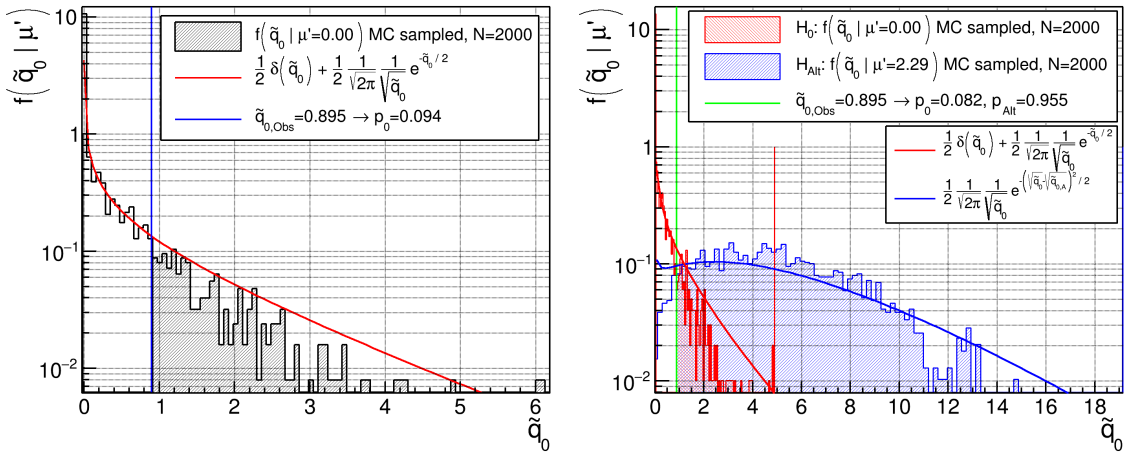

Figure 6.9: The resulting distributions $f\left(\tilde{q}_0|\mu'\right)$ from 2'000 Monte Carlo toy experiments for a 100 GeV WIMP and the tested value $\mu = 0$, given a true signal strength μ'. Shown are the background-only hypothesis H_0 of $\mu' = 0$ (left, black and right, red), as well as the alternative hypothesis $\mathrm{H}_{\mathrm{Alt}}$ of $\mu' = 2.29$ (right, blue) with respective asymptotic approximations (Eq. 6.11 and Eq. 6.14). The test statistic $\tilde{q}_{0,\mathrm{Obs}}$, corresponding to the ArDM observations $N_{\mathrm{Obs},i}$, allows the estimation of the corresponding p-values p_0 and p_{Alt}. The resulting low p-value p_0 of $\sim$0.09 is not significant enough to reject H_0. It is a natural consequence of the over-fluctuation observed in ArDM (see Fig. 6.7). It would be mitigated with an incorporation of the systematic uncertainty of the RoI box approach and the proper statistical error of the live-time rescaling of the background measurement into the ArDM model (Eq. 6.2).

Setting an upper limit through inverted hypothesis testing: The second and last case concerns the setting of an upper limit on the signal strength μ after not being able to reject the background-only hypothesis. Now, for the upper limit, the two hypotheses are interchanged. In the new null hypothesis H_0, the true signal strength μ' is strictly non-zero, while the alternative $\mathrm{H}_{\mathrm{Alt}}$ hypothesis consists of a background-only scenario with $\mu' = 0$.

The general two-sided test statistic $\tilde{t}_\mu$ for any tested value μ accounts to

$$\tilde{t}_\mu = -2\ln\left(\tilde{\lambda}\left(\mu\right)\right) = \begin{cases} -2\ln\left(\dfrac{\mathcal{L}\left(\mu,\hat{\nu}_{\mathrm{b},i}(\mu)\right)}{\mathcal{L}\left(\hat{\mu},\hat{\nu}_{\mathrm{b},i}\right)}\right) & \text{for } \hat{\mu} \geq 0, \\ -2\ln\left(\dfrac{\mathcal{L}\left(\mu,\hat{\nu}_{\mathrm{b},i}(\mu)\right)}{\mathcal{L}\left(0,\hat{\nu}_{\mathrm{b},i}(0)\right)}\right) & \text{for } \hat{\mu} < 0. \end{cases} \tag{6.17}$$

The corresponding probability density function $f\left(\tilde{t}_\mu|\mu'\right)$ is asymptotically approximated by

$$f\left(\tilde{t}_\mu|\mu'\right) \approx \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\frac{1}{2}\left(\sqrt{\tilde{t}_\mu}+\frac{\mu-\mu'}{\sigma}\right)^2} + \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\frac{1}{2}\left(\sqrt{\tilde{t}_\mu}-\frac{\mu-\mu'}{\sigma}\right)^2} & \tilde{t}_\mu \leq \mu^2/\sigma^2, \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}e^{-\frac{1}{2}\frac{\left(\tilde{t}_\mu-\frac{\mu^2-2\mu\mu'}{\sigma^2}\right)^2}{(2\mu/\sigma)^2}} & \tilde{t}_\mu > \mu^2/\sigma^2. \end{cases} \tag{6.18}$$

The standard deviation σ can once again be evaluated with the test statistic $\tilde{t}_{\mu,\mathrm{A}}$ of the

corresponding Asimov data set (see Eq. 6.8). For the null hypothesis H_0, and the case of the tested signal strength μ being equal to the true signal strength μ', this simplifies to

$$f\left(\tilde{t}_\mu|\mu'=\mu\right) \approx \begin{cases} \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\frac{1}{2}\tilde{t}_\mu} & \tilde{t}_\mu \leq \mu^2/\sigma^2, \\ \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\frac{1}{2}\tilde{t}_\mu} + \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}e^{-\frac{1}{2}\frac{\left(\tilde{t}_\mu+\mu^2/\sigma^2\right)^2}{(2\mu/\sigma)^2}} & \tilde{t}_\mu > \mu^2/\sigma^2. \end{cases} \tag{6.19}$$

Fig. 6.10 (left) shows a histogram (black) of 2'000 Monte Carlo toy experiments (with resampled $N_{\mathrm{Obs},i}$) and their resulting distribution for a specific signal strength of $\mu = \mu' = 2.29$ (which again corresponds to the initial guess, motivated by the simple signal profile likelihood ratio estimation in Fig. 6.8). Additionally, the asymptotic approximation (Eq. 6.19) is plotted on top (red), resulting in an acceptable agreement with the toy-sampled distribution. Again, for the purpose of the ArDM analysis, all probability density functions of test statistics are approximated by the distribution of the Monte Carlo toy experiments and the corresponding p-values are evaluated accordingly. The test statistic $\tilde{t}_{\mu,\mathrm{Obs}}$ of the actually observed values $N_{\mathrm{Obs},i}$, and given that $\mu = 2.29$, is also indicated (blue). The corresponding p-value p_μ of H_0, or the probability of a more extreme observation than $\tilde{t}_{\mu,\mathrm{Obs}}$, given $\mu = \mu'$, is calculated naturally as

$$p_\mu = P\left(\tilde{t}_\mu \geq \tilde{t}_{\mu,\mathrm{Obs}}|\mu'=\mu\right) = \int_{\tilde{t}_{\mu,\mathrm{Obs}}}^{\infty} f\left(\tilde{t}_\mu|\mu'=\mu,\hat{\nu}_{\mathrm{b},i}\left(\mu'=\mu\right)\right)\mathrm{d}\tilde{t}_\mu. \tag{6.20}$$

In the asymptotic approximation (Eq. 6.19), the p-value p_μ of H_0 can be evaluated as

$$p_\mu = 1 - F\left(\tilde{t}_{\mu,\mathrm{Obs}}|\mu'=\mu\right) \approx \begin{cases} 2\Phi\left(\sqrt{\tilde{t}_{\mu,\mathrm{Obs}}}\right) & \tilde{t}_\mu \leq \mu^2/\sigma^2, \\ \Phi\left(\sqrt{\tilde{t}_{\mu,\mathrm{Obs}}}\right) + \Phi\left(\frac{\tilde{t}_{\mu,\mathrm{Obs}}+\mu^2/\sigma^2}{2\mu/\sigma}\right) & \tilde{t}_\mu > \mu^2/\sigma^2. \end{cases} \tag{6.21}$$

In the ArDM experiment, for the WIMP mass of 100 GeV and $\mu = \mu' = 2.29$, the p-value evaluates to $p_\mu = 0.101$ (Fig. 6.10, left, grey shaded) based on the 2'000 Monte Carlo toy experiments. This signifies, that only an expected 10.1% of repeated experiments would result in more extreme values of the observed $N_{\mathrm{Obs},i}$, given that $\mu = \mu' = 2.29$.

The respective asymptotic approximation of the probability density function $f\left(\tilde{t}_\mu|\mu'=0\right)$ for the alternative hypothesis $\mathrm{H}_{\mathrm{Alt}}$ of $\mu \neq \mu' = 0$ can be directly synthesized from the general case (Eq. 6.18) as

$$f\left(\tilde{t}_\mu|\mu'=0\right) \approx \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\frac{1}{2}\left(\sqrt{\tilde{t}_\mu}+\frac{\mu}{\sigma}\right)^2} + \begin{cases} \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{t}_\mu}}e^{-\frac{1}{2}\left(\sqrt{\tilde{t}_\mu}-\frac{\mu}{\sigma}\right)^2} & \tilde{t}_\mu \leq \mu^2/\sigma^2, \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)}e^{-\frac{1}{2}\frac{\left(\tilde{t}_\mu-\frac{\mu^2}{\sigma^2}\right)^2}{(2\mu/\sigma)^2}} & \tilde{t}_\mu > \mu^2/\sigma^2. \end{cases} \tag{6.22}$$

Therein, the standard deviation σ can once again be evaluated from the test statistic $\sqrt{\tilde{t}_{\mu,\mathrm{A}}}$ of the corresponding Asimov data (see Eq. 6.8) as

$$\sqrt{\tilde{t}_{\mu,\mathrm{A}}} \overset{\mu'=0}{=} \frac{\mu}{\sigma} \quad \Rightarrow \sigma \overset{\mu'=0}{=} \frac{\mu}{\sqrt{\tilde{t}_{\mu,\mathrm{A}}}}. \tag{6.23}$$

Subsequently, the p-value $p_{\mu,\mathrm{Alt}}$ of the alternative hypothesis $\mathrm{H_{Alt}}$ is calculated as

$$p_{\mu,\mathrm{Alt}} = P\left(\tilde{t}_\mu \geq \tilde{t}_{\mu,\mathrm{Obs}} | \mu' = 0\right) = \int_{\tilde{t}_{\mu,\mathrm{Obs}}}^{\infty} f\left(\tilde{t}_\mu | \mu' = 0, \hat{\hat{\nu}}_{\mathrm{b},i}\left(\mu' = 0\right)\right) \mathrm{d}\tilde{t}_\mu \tag{6.24}$$

In the asymptotic approximation (Eq. 6.22), this accounts to

$$p_{\mu,\mathrm{Alt}} = 1 - F\left(\tilde{t}_{\mu,\mathrm{Obs}} | \mu' = 0\right) \approx \Phi\left(\sqrt{\tilde{t}_{\mu,\mathrm{Obs}}} + \frac{\mu}{\sigma}\right) + \begin{cases} \Phi\left(\sqrt{\tilde{t}_{\mu,\mathrm{Obs}}} - \frac{\mu}{\sigma}\right) & \tilde{t}_\mu \leq \mu^2/\sigma^2, \\ \Phi\left(\frac{\tilde{t}_{\mu,\mathrm{Obs}} - \mu^2/\sigma^2}{2\mu/\sigma}\right) & \tilde{t}_\mu > \mu^2/\sigma^2. \end{cases} \tag{6.25}$$

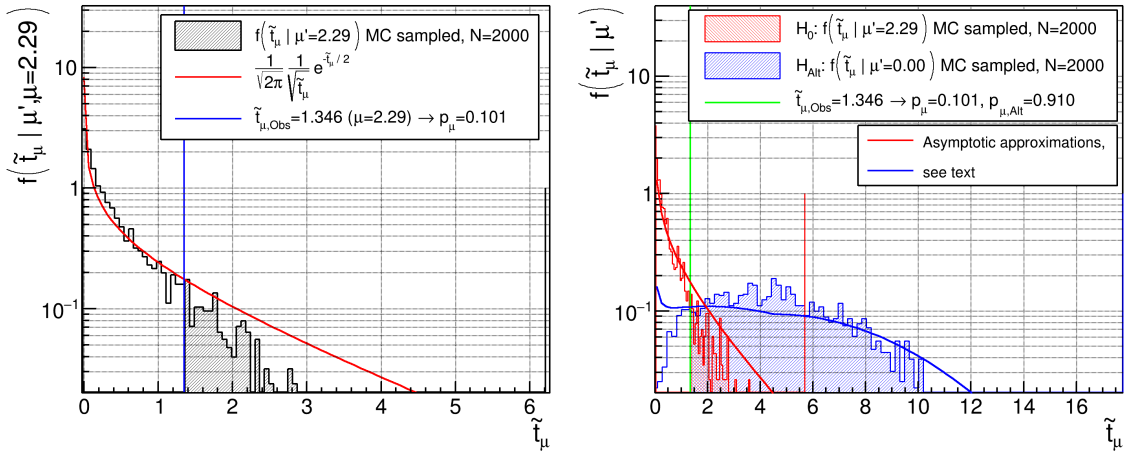

Figure 6.10: The resulting distributions $f\left(\tilde{t}_\mu|\mu'\right)$ from 2'000 Monte Carlo toy experiments for a 100 GeV WIMP and the tested value $\mu = 2.29$, given a true signal strength μ'. Shown are the $\mathrm{H_0}$ hypothesis of $\mu' = 2.29$ (left, black and right, red), as well as the alternative hypothesis $\mathrm{H_{Alt}}$ of $\mu' = 0$ (right, blue) with respective asymptotic approximations (Eq. 6.19 and Eq. 6.22). The test statistic $\tilde{t}_{\mu,\mathrm{Obs}}$, corresponding to the ArDM observations $N_{\mathrm{Obs},i}$, allows the estimation of the corresponding p-values p_μ and $p_{\mu,\mathrm{Alt}}$. While the ArDM result supports $\mathrm{H_{Alt}}$ ($\mu' = 0$, $p_{\mu,\mathrm{Alt}} = 0.910$) over $\mathrm{H_0}$ ($\mu' = 2.29$, $p_\mu = 0.101$), the importance for the WIMP exlusion limit lies in the value of $p_\mu = 0.101$. It translates to the statistical statement, that, at an estimated confidence level of 89.9%, the tested signal strength of $\mu = \mu' = 2.29$ doesn't result in a more extreme observation than experienced in the ArDM experiment.

The corresponding hypothesis test for the signal strength $\mu = 2.29$ is showcased in Fig. 6.10 (right). It contains the Monte Carlo toy-sampled distributions of $f\left(\tilde{t}_\mu|\mu' = 2.29\right)$ (a specific $\mathrm{H_0}$, red) and $f\left(\tilde{t}_\mu|\mu' = 0\right)$ ($\mathrm{H_{Alt}}$, blue) and the respective asymptotic approximations (Eq. 6.19 and Eq. 6.22 respectively). The tested value of $\mu = 2.29$ once again corresponds to the guess of the upper limit through the asymptotic approximation of the simple signal profile likelihood ratio (Fig. 6.8). The respective distributions, relative to the ArDM experiment (lime), result in $p_\mu = 0.101$ (red, shaded) and $p_{\mu,\mathrm{Alt}} = 0.910$ (blue, shaded) based on 2'000 toy experiments each. This signifies, that under the null hypothesis $\mathrm{H_0}$ of the true signal strength $\mu' = \mu = 2.29$, an estimated 10.1% of repeated experiments

would result in a more extreme observation than ArDM. The fact that this is congruent with the previously estimated $p_\mu = 0.101$ (in Fig. 6.10, left, grey shaded), motivates, that 2'000 toy experiments are enough to attain a reasonable precision in the test of H_0. In the case of the alternative background-only hypothesis H_{Alt} and $\mu' = 0 \neq \mu$, an estimated 91.0% of repeated experiments would result in a more extreme observation than ArDM. Consequently, the ArDM results favor H_{Alt} ($\mu' = 0$) over H_0 ($\mu' = 2.29$). But the fundamental result here is the H_0 p-value $p_\mu = 0.101$, which supports, in other words, that, at an estimated confidence level of 89.9%, the tested signal strength of $\mu = \mu' = 2.29$ doesn't result in a more extreme observation than experienced in the ArDM experiment.

This directly leads to the principle of the inverted hypothesis testing to set an upper limit on the signal strength μ at a desired confidence level. Therein, the above testing routine (Fig. 6.10, right) with $f(\tilde{t}_\mu|\mu')$ is repeated over a lattice of tested values μ. Essentially, this results in an estimation of the p-values p_μ as a function of the tested signal strengths μ. The intersection of this estimated function with the desired confidence level (corresponding to $1 - p_\mu$) then allows the inversion of relating a specific value of the signal strength μ to the respective confidence level. This is commonly done in the frame of the "CL_s-method" [112, 113]. Therein, the p-values p_μ and $p_{\mu,\,Alt}$ are relabeled as

$$CL_{s+b} = p_\mu \quad \text{and} \quad CL_b = p_{\mu,\,Alt} \quad \text{respectively, and} \quad CL_s = \frac{CL_{s+b}}{CL_b} \tag{6.26}$$

is an attempt to reduce the influence of the background measurement on the estimated signal strength confidence level. The corresponding CL_{s+b}, CL_b and CL_s values for the 100 GeV WIMP, resulting from the inverted hypothesis testing over a lattice of the signal strength μ, are shown in Fig. 6.11. Therein, the observed values are compared to the expected values in the absence of signal ($\mu' = 0$), which are showcased in form of the "Brazil band". This expectation is based on the toy-sampled distribution $f(p_\mu|\mu' = 0)$ of p-values p_μ given that the alternative hypothesis H_{Alt} ($\mu' = 0$) were true. Explicitly, if there were no signal, the expected observed test statistic would be distributed as $f(\tilde{t}_\mu|\mu' = 0)$ (Fig. 6.10, right, blue, for the tested value $\mu = 2.29$). And, if one were to then test the null hypothesis H_0 ($\mu' = \mu = 2.29$), one would evaluate the p-values p_μ as the quantiles of the distribution $f(\tilde{t}_\mu|\mu' = 2.29)$ (Fig. 6.10, right, red). This results in the expected distribution $f(p_\mu|\mu' = 0)$ of the p-values p_μ for a specific tested μ, given the absence of any signal. Consequently, the "Expected CL_{s+b}" values in Fig. 6.11 (black, dotted) correspond to the median $CL_{s+b} = p_\mu$ values of the distributions $f(p_\mu|\mu' = 0)$ for a range of tested μ values. The respective 1σ (lime) and 2σ (yellow) bands correspond to the boundaries of the central ~68.27% and ~95.45% of the distributions $f(p_\mu|\mu' = 0)$ of the tested μ values. The actually observed CL_{s+b} values (black errorbars) are situated in the upper 2σ band. Naturally, lower observed CL_{s+b} values lead to a more stringent upper limit on the signal strength μ. Consequently, Fig. 6.11 shows how "unlucky" the over-fluctuation observed in ArDM is in terms of the statistical expectation. The observed $CL_b = p_{\mu,\,Alt}$ values (turquoise errorbars) are consistently close to one. Consequently, the resulting CL_s values are very similar to the CL_{s+b} values. Due to the minimal difference, the ArDM upper limit is set based on the more traditional CL_{s+b} values. The resulting constrained upper limit $\mu_{Constrained}(CL_{s+b}) = 2.28 \pm 0.02$ at a 90% confidence level corresponds to the intersection of the progression of the observed CL_{s+b} values (black) and the desired p-value of 0.10 (1 − "confidence level", red). This concludes the statistical testing process for the specific WIMP of $m_\chi = 100$ GeV with setting the corresponding upper limit on the

spin-independent WIMP-nucleon cross section σ_p^{SI} at $(2.28 \pm 0.02) \cdot 10^{-41}$ cm^2 at the 90% confidence level. The resulting exclusion limit in the plane of m_χ and σ_p^{SI}, after testing a large range of WIMP masses m_χ, is shown and discussed in Section 6.3.2.

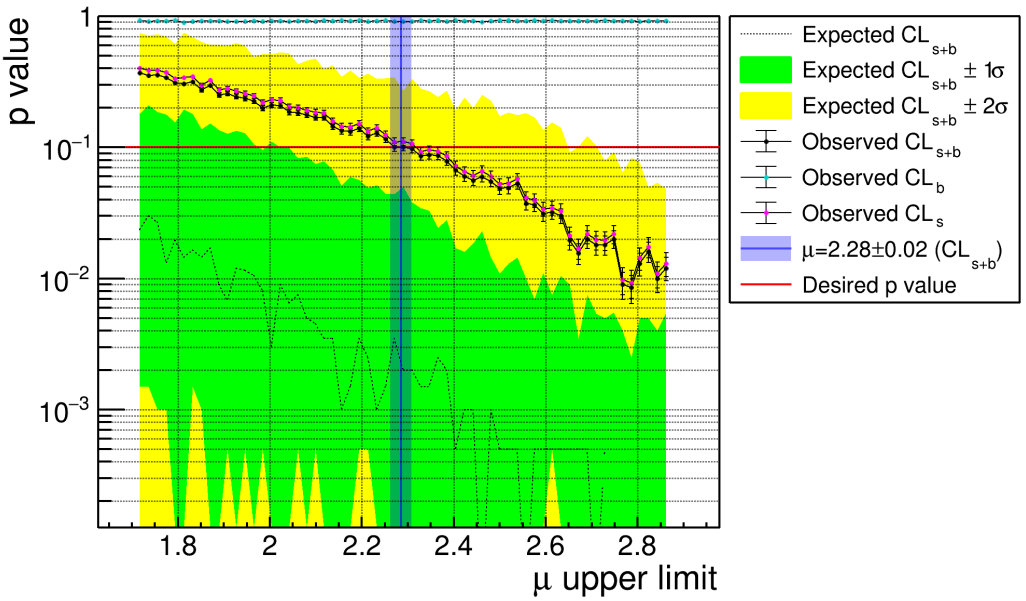

Figure 6.11: The inverse hypothesis testing result for a 100 GeV WIMP after repeating the hypothesis test on $f\left(\tilde{t}_\mu|\mu'\right)$ (see Fig. 6.10) for a lattice of μ values. Each evaluated p-value (or confidence level via Eq. 6.26) is based on 2'000 toy-sampled experiments and the resulting distributions $f\left(\tilde{t}_\mu|\mu'\right)$. The observed $\mathrm{CL_{s+b}}$ values (black errorbars) correspond to a high over-fluctuation compared to the toy-sampled expected distributions $f\left(p_\mu|\mu'=0\right)$ in the absence of any true signal (black dotted, lime and yellow). The intersection of the progression of the $\mathrm{CL_{s+b}}$ values with the desired p-value of 0.10 (red) results in a constrained upper limit of $\mu_{\mathrm{Constrained}} = 2.28 \pm 0.02$. Consequently, the ArDM experiment results in an upper limit on the spin-independent WIMP-nucleon cross section σ_p^{SI} of $(2.28 \pm 0.02) \cdot 10^{-41}$ cm^2 for a 100 GeV WIMP at the 90% confidence level.

6.3.2 The WIMP exclusion limit of ArDM - Conclusion

The previously introduced "raster scan" consists of the repeated evaluation of the theoretical WIMP spectrum ν_{s} (Section 6.1), and the corresponding statistical testing process (Section 6.3.1) to constrain the respective signal strength μ, for a wide range of WIMP masses m_χ. As detailed in Section 6.3.1, the inverted hypothesis testing results in an upper limit on the spin-independent WIMP-nucleon cross section $\sigma_p^{\mathrm{SI}} = \mu_{\mathrm{Constrained}} \cdot 10^{-41}$ cm^2 for the respective WIMP mass m_χ at the 90% confidence level. The evaluation of the ArDM WIMP search data in the m_χ range from 1 GeV to 10 TeV is shown in Fig. 6.12 (red) and compared to other WIMP search experiments. The ArDM exposure of 4'843 kg $\cdot$ days (with the fiducial mass $m_{\mathrm{Ar,fid}} = 574.57$ kg and a live-time of 8.429 live-days) results in an upper limit of $\sigma_p^{\mathrm{SI}}\left(m_\chi = 100 \text{ GeV}\right) = (2.28 \pm 0.02) \cdot 10^{-41}$ cm^2 and a minimal upper limit of $\sigma_p^{\mathrm{SI}}\left(m_\chi = 120.2 \text{ GeV}\right) = (2.16 \pm 0.02) \cdot 10^{-41}$ cm^2 at the 90% confidence level. In general, the theoretical WIMP spectrum (Section 6.1) shifts to higher nuclear recoil energies, and consequently higher primary scintillation light S1L, for higher WIMP masses m_χ. Conse-

quently, ArDM becomes sensitive to an increasing number of the S1L bins with increasing WIMP mass m_χ. Explicitly, below a WIMP mass of m_χ ∼15 GeV, the ArDM experiment is only sensitive to signal below ten photoelectrons of S1L. Towards higher WIMP masses above ∼15 GeV, 22 GeV and 28 GeV, the ArDM experiment hypothetically could see signal in the second, third and fourth S1L bin up to 20, 30 and 40 photoelectrons respectively. Finally, at WIMP masses above ∼63 GeV, ArDM is sensible in all ten S1L bins up to 100 photoelectrons. Consequently, the statistical model of ArDM (Eq. 6.2) below m_χ of ∼63 GeV simplifies to a lower number of S1L bins, but the remaining statistical testing process remains the same. While the estimated WIMP-affecting signal cut efficiency of $6.08 \cdot 10^{-4}$ in the lowest S1L bin of up to ten photoelectrons (Section 6.1, Fig. 6.3, black) is very low, it not being zero allows ArDM to constrain an upper limit below WIMP masses of ∼15 GeV. The respective statistical uncertainty (blue, shaded) on the upper limit, originating from the inverted hypothesis testing (see Fig. 6.11, Section 6.3.1), is lower than 1% above WIMP masses of 1.5 GeV. This statistical uncertainty is primarily driven by the number of toy experiments sampled per hypothesis (2'000 toy experiments in this analysis).

Compared to other leading WIMP direct detection experiments, the presented ArDM result is not competitive. The world-leading WIMP direct detection noble gas results belong to the xenon-based experiments XENON1T [47] (pink) and PandaX-4T [48] (black). Their corresponding exclusion limits for a 100 GeV WIMP are approximately 5.5 orders of magnitude better than ArDM's corresponding result. Moreover, the CMS collaboration [114] (blue) is responsible for the lowest exclusion limit for WIMPs with masses below ∼12 GeV. The CMS collaboration searched for invisible decays of the Higgs boson, produced via vector boson fusion. However, the comparison of ArDM with xenon-based experiments, or the search for invisible decays, is not straightforward and is forgone here; in favor of a comparison with the corresponding argon-based experiment.

The world-leading direct detection result, using argon as the detection material, stems from DarkSide-50. The corresponding DarkSide-50 exclusion limit above a m_χ of ∼20 GeV (2018 [50], dark green) is better by up to three orders of magnitude, and the dedicated DarkSide-50 low-mass search below ∼10 GeV (2022 [115], with quenching fluctuations, lime green) is better by up to four orders of magnitude. This difference is a consequence of multiple factors, with the first being the total exposure. While DarkSide-50 was working with a much smaller fiducial mass of ∼31.3 kg (and ∼20 kg in the case of the low mass search), the respective live-time of 532.4 live-days (and 653.1 live-days respectively) results in an exposure of 16'660 kg · days (and 12'306 kg · days respectively). Consequently, the exposure of DarkSide-50 is a factor ∼3.44 (and ∼2.54 in the case of the low-mass search) higher than the exposure of ArDM. If ArDM were taking data for the same live-time as DarkSide-50 (532.4 live-days) the resulting exposure would be larger by a factor of ∼63.16 compared to the current ArDM exposure (305'901 kg·days instead of 4'843 kg·days). With the resulting increased exposure, the hypothetically prolonged ArDM experiment would expect proportionally more signal by a factor of ∼63.16. Depending on whether the expected background also scales linearly or less, the resulting hypothetical sensitivity would be better by an estimated factor in the range of $\sim \sqrt{63.16}$ to ∼63.16. The resulting projected sensitivity of the ArDM experiment operated over the live-time of 532.4 live-days is also shown in Fig. 6.12 (red, shaded). Secondly, the data cut efficiencies are generally favorable for the DarkSide-50 experiments. The NR acceptance of DarkSide-50 is approximately constant at ∼70% above ∼60-70 keV of nuclear recoil energy. The respective WIMP-affecting cut efficiency in the case of ArDM (see Fig. 6.3, black, Section 6.1) is increasing approximately

linearly from ∼1% at ∼60-70 keV up to ∼6% at ∼185 keV of nuclear recoil energy (or 20 to 100 photoelectrons of S1L). One of the underlying reasons lies in the more efficient event reconstruction in DarkSide-50. As will be discussed in the Outlook, DarkSide-50 used underground argon (UAr), which is depleted in the ^{39}Ar isotope, and consequently the resulting β^- ER spectrum, by an estimated factor of ∼1400 relative to the atmospheric argon (AAr) used in ArDM. Additionally, the target volume of DarkSide-50 was surrounded with 30 tonnes of a liquid scintillator, which could veto neutron-induced NRs highly efficiently [116]. Consequently, the DarkSide-50 collaboration managed to reject much more ER and NR background without loosing the hypothetical signal strength. Additionally, the drastically lower event frequency due to the use of UAr and the smaller maximal drift length of 376 μs result in practically no pile-up events. ArDM, through the use of AAr and its sheer dimensional prowess, resulting in a considerable pile-up of events (showcased in Figs. 4.23 and 4.24, Section 4.2.1), has a clear disadvantage, which is reflected in the low matching efficiency of an average 11.5% below 100 photoelectrons of S1L (see Fig. 5.16, Section 5.4). Additionally, the discovered Charge-up effect (see Section 4.2) led to an increase of the S2 frequency over time, further complicating the matching inside ArDM. Moreover, as shown in the statistical analysis of the 100 GeV WIMP (see Fig. 6.11, Section 6.3.1), the chosen split of the WIMP search data and the respective DP calibration data resulted in a statistically unfavorable over-fluctuation of background events in the WIMP search data, which further weakens the constrainable exclusion limit.

The last fundamental difference is the very high LY of 8.0 ± 0.2 p.e./keV in DarkSide-50 compared to the LY of 1.82 ± 0.13 p.e./keV in ArDM (see Section 5.1). While Darkside-50 generally had smaller holes in the packing of their 19 3" PMTs compared to ArDM's 12 8" PMTs in the bottom and top arrays, the main reason of the vast difference in the LYs simply is the scalability of the entire optical system and its PMT coverage from ∼50 kg to the tonne-scale. An estimation of how the sensitivity of ArDM with a LY of 8.0 p.e./keV would improve, is not straight-forward, as the respective cut efficiencies might also change. However, it can be insinuated via the change in the expected number of WIMPs before all cuts in ArDM under the same experimental conditions (except the LY). Therefore, the respective energy range has to be adapted. In a reasonable assumption, the statistical behavior of the discrimination variable S1F90 will not change drastically as a function of the number of photoelectrons. Consequently, one can assume, that also with a LY of 8.0 p.e./keV, ArDM would not achieve significant efficiency below 20 photoelectrons. Above the threshold of 20 photoelectrons, ArDM, with a hypothetical LY of 8.0 p.e./keV and the same other experimental conditions, would expect a factor of 3.253 more WIMPs before any cuts given a WIMP mass of $m_\chi = 100$ GeV. Additionally, the relative LY improvement by a factor of ∼4.4 (from a LY of 1.82 p.e./keV to a LY of 8.0 p.e./keV) would directly translate to the relative scale of the S1L, resulting in the same statistical recoil type discrimination power at the same S1L but correspondingly lower nuclear recoil energies. The exact behavior of the ArDM WIMP-affecting cut efficiencies (Fig. 6.3) is difficult to predict for a hypothetical LY of 8.0 p.e./keV. Generally, the additional ∼225% of WIMPs above 20 photoelectrons of S1L (compared to the LY of 1.82 p.e./keV) shift to higher S1L bins with respectively higher cut efficiencies. Consequently, an additional improvement factor for the number of estimated WIMPs after all cuts, that could far exceed the previously discussed factor of 3.253, is expected for a hypothetical ArDM LY of 8.0 p.e./keV.

While these considerations predominantly serve to estimate the statistical influence and are purely hypothetical, they reason the observed difference of approximately three

orders of magnitude in terms of the exclusion limit of DarkSide-50 and ArDM. Some of the fundamental differences, the lower LY and the drastically lower matching efficiency, are primarily a consequence of the scale difference of the two experiments and the use of AAr inside ArDM instead of UAr inside DarkSide-50. Increasing the detector size from 50 kg to the tonne-scale simply comes accompanied with technical difficulties. ArDM still is the first tonne-scale dual-phase argon-based WIMP direct detection experiment, which was successfully performed; and therewith fulfills its fundamental purpose. ArDM proves the general feasibility of tonne-scale dual-phase argon-based detectors for the WIMP search. The inauguration of this new size scale remains a fundamental step in the progression of the search for dark matter. A preview on the future perspective of the WIMP search with argon-based direct detection experiments is given in the Outlook.

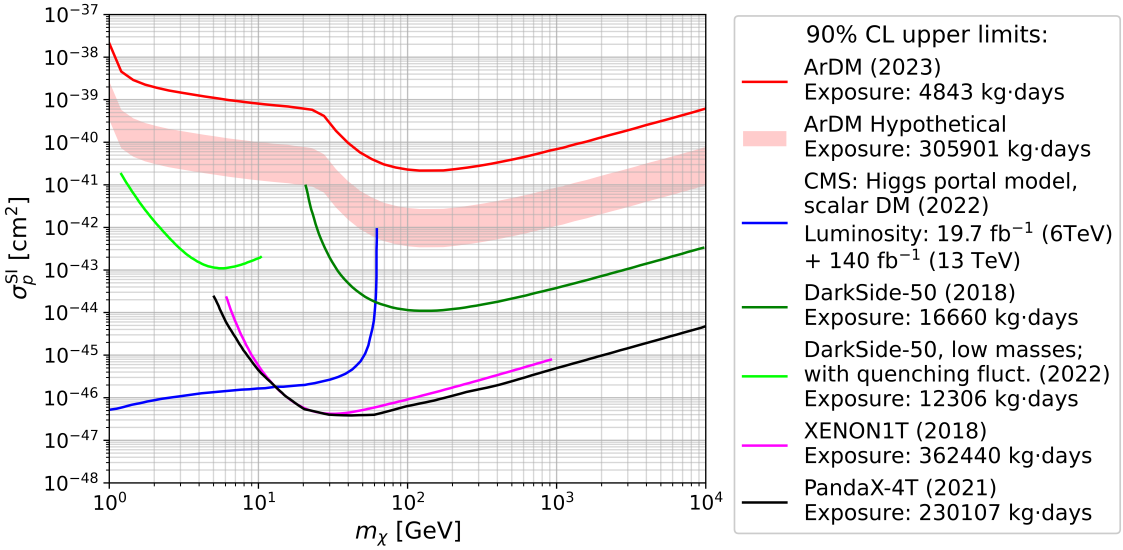

Figure 6.12: The resulting ArDM exclusion limit (red) of the spin-independent WIMP-nucleon cross section σ_p^{SI} as a function of the respective WIMP mass m_χ. The ArDM experiment, with an exposure of 4'843 kg $\cdot$ days results in an upper limit of $\sigma_p^{\mathrm{SI}}\,(m_\chi = 100\ \mathrm{GeV}) = (2.28 \pm 0.02) \cdot 10^{-41}\ \mathrm{cm}^2$ and a minimal upper limit of $\sigma_p^{\mathrm{SI}}\,(m_\chi = 120.2\ \mathrm{GeV}) = (2.16 \pm 0.02) \cdot 10^{-41}\ \mathrm{cm}^2$ at the 90% confidence level. The results of the world-leading direct detection experiments XENON1T [47] (pink), PandaX-4T [48] (black) and the CMS collaboration [114] (blue) are shown with their respective exposures. A detailed comparison with the world-leading argon-based direct detection experiment DarkSide-50 [50, 115] (green) is presented in the text. The hypothetical ArDM sensitivity (red, shaded) shows the projected exclusion limit curve if ArDM were taking data for the same 532.4 live-days as DarkSide-50.

Outlook

The ArDM experiment is only one milestone in a long line of dark-matter-searching experiments. Naturally, there are short- and long-term improvement plans to further progress the development of direct detection WIMP searches and breach new boundaries in terms of the testable WIMP-nucleon interaction cross section.

On the shorter term, the baseline result of the ArDM experiment itself can be improved potentially. The current analysis, presented herein, is focused on the accuracy of the results. The matching cut (see Section 5.4, Fig. 5.16) results in an estimated efficiency of $\varepsilon_{\text{Total}} = 0.115$ below 100 photoelectrons of S1L, with the aim of removing any matches with multiple competing candidates of S1 or S2 signals. This restrictive matching cut was necessitated by the discovery of the Charge-up effect (see Section 4.2) and the consequently observed enhanced pile-up of S2 signals, resulting in unreasonably high values of the log(S2maxL/S1L) (see Section 4.2.1). While restricting the matching to those perfect competition-less events did wonders in terms of minimizing the percentage of pile-up events and emphasizing the expected ER and NR distributions (compare Fig. 4.22 to Fig. 4.25 specifically, before and after the matching/pile-up cut), it also reduced the cut efficiency significantly. The aforementioned matching efficiency of $\varepsilon_{\text{Total}} = 0.115$ below 100 photoelectrons of S1L was evaluated on the DP calibration data (see Section 5.4, Fig. 5.16) and only serves as an estimation. However, the actual NR- or WIMP-affecting cut efficiency is not expected to be greater than the respective DP efficiency. A less restrictive matching based on the correlations of other event variables (such as the S2 drift time and the S1TTR) could potentially result in a very high degree of belief of correctness, and a significantly higher cut efficiency. Therefore, experimenting with different matching criteria and comparing to the outcome of the "perfect" restrictive matching could salvage a significant portion of additional data, which would consequently increase the expected number of WIMPs and lower the final ArDM exclusion limit.

An additional consequence of the discovery of the Charge-up effect was the eventual dismissal of any data, which was not dubbed "Charge-up stable"; resulting in an individual data selection efficiency of 54.606% (see Section 4.2.1, Fig. 4.26 and Table 4). An adapted matching method might also be able to relax the criterion of the Charge-up stability and restore more of the remaining 45.394% of the potential WIMP search DP data. Consequently, the resulting increased exposure would allow ArDM to probe smaller WIMP-nucleon cross sections and set a lower exclusion limit. As discussed previously, the observed ArDM results (Section 6.2, Fig. 6.7) correspond to a statistically unlikely and unfavorable over-fluctuation compared to the expectation from the DP calibration data (see Section 6.3.1, Fig. 6.11). While speculative, the added data from these potential improvements is statistically expected to guide the ArDM observation more towards the average experiment. The correspondingly less over-fluctuating result should also naturally lower the constrained ArDM exclusion limit. These main points, all related to a less restrictive matching method, conclude the principal short-term improvement options.

On the longer term, ArDM joined the **Global Argon Dark Matter Collaboration (GADMC)** in 2017 together with the experiments of DarkSide-50 [50], DEAP-3600 [49] and MiniCLEAN [117, 118] with the aim of combining the efforts in order to reach new magnitudes in the direct WIMP search with argon. As discussed in Section 5.1, the atmospheric argon (AAr) used in ArDM incorporates an ^{39}Ar activity of $\sim$1 Bq/kg$_{\text{LAr}}$ due to the cosmogenic production of the ^{39}Ar isotope via the neutron capture of the ^{40}Ar isotope. The

electronic recoils, caused by the electrons from the β^- decay of the ^{39}Ar atoms, subsequently account for the vast majority of events and produce the complicated matching scenarios inside ArDM. It was first shown in 2015, that underground argon (UAr), enriched from a CO_2 well in Cortez, Colorado, has a significantly smaller ^{39}Ar activity of 6.6 mBq/kg$_{\mathrm{LAr}}$ (corresponding to a depletion factor of $\sim$153.8 relative to AAr) [119]. As a major upgrade in 2017, DarkSide-50 incorporated 153 ± 1 kg of UAr enriched from the same CO_2 well, and further purified by Fermilab [120], resulting in a measurement of the ^{39}Ar depletion factor of $1'400 \pm 200$ relative to AAr [121].

Naturally, the use of UAr over AAr is a fundamental improvement in future direct detection experiments, as it results in significantly less ER background and also simplifies the matching scenarios tremendously. On the drift length of ArDM, practically no pile-up of S1 signals would be expected with UAr. Therefore, the next phase planned by the GADMC, called the **DarkSide-20k** experiment, incorporates a fiducial volume of 20 tonnes of liquid UAr [122]. The required UAr will be extracted by the **Urania** plant from the same CO_2 well in Cortez, Colorado with an estimated production rate of up to 330 kg/day [123, 124]. For the purpose of further chemical purification, a 350 m cryogenic isotopic distillation column, called the **Aria** plant, is currently being installed in a mine shaft in southern Sardinia, Italy [124, 125]. The primary purpose of Aria will be to remove all chemical impurities (e.g. traces of N_2, O_2 and Kr) with a separation power better than 10^3 per pass. Therefore, Aria will manage to process a rate of several tonnes per day of UAr. As a secondary purpose, Aria will allow to further deplete the concentration of the ^{39}Ar isotope in the UAr by an estimated depletion factor of 10 per pass. The models of the DarkSide collaboration indicate, that the Aria plant could produce up to 10 kg/day of depleted UAr (DAr). After successful testing and operation, Aria will receive a wider second column which is estimated to increase the processing rate of UAr to DAr to 150 kg/day [122,124]. Over multiple passes through Aria, the total combined ^{39}Ar depletion factor could be improved even further. The final ^{39}Ar depletion factor will be measured for specific detector-grade UAr or DAr batches inside the single-phase chamber called **DArT**; which will be located inside ArDM, with ArDM acting as a single-phase veto [123, 125].

Another technical improvement is the use of silicon photomultipliers (SiPM) instead of the currently used borosilicate PMTs. SiPMs are a promising advancement with the potential for a higher photon detection efficiency and a better single-photon resolution, while operating at lower bias voltages [126]. Additionally, the tile shape of SiPMs and their relatively small size at the order of mm^2 allows a close arrangement, resulting in an improved possible coverage compared to the traditional PMTs. DArT will use two 1 cm^2 SiPMs [123, 125], while the LAr time projection chamber of DarkSide-20k is planned to incorporate SiPMs covering a total area of 14 m^2 [122].

Upon successful operation, DarkSide-20k is expected to be sensitive to a spin-independent WIMP-nucleon cross section σ_p^{SI} of $1.2 \cdot 10^{-47}$ cm^2 for a 1 TeV WIMP and a live-time of five years [122] (see Fig. O.1). In the even longer term, the GADMC plans to eventually build the **Argo** detector with an even larger fiducial mass of 200 tonnes [122]. Argo, after five years of operation, would hypothetically even be sensitive to the "neutrino floor" of coherent neutrino-nucleus scattering [127–131]. This would result in an additional source of non-discriminable background in the form of neutrino-induced nuclear recoils. With the status of today's knowledge, one would then have to turn to other WIMP search alternatives (e.g. directional experiments, based on the annual modulation of the WIMP signal and the neutrino background, as proposed in [132, 133]). However, until the neutrino floor

is reached, there are certainly exciting times ahead in the direct detection search for dark matter candidates.

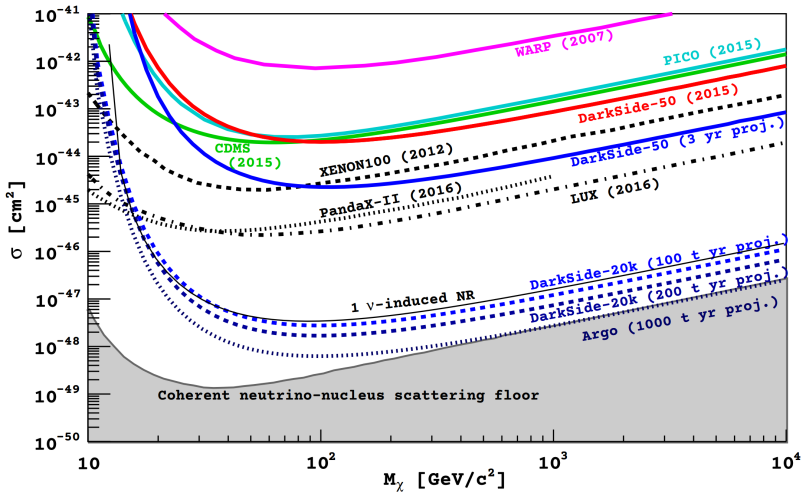

Figure O.1: The results of current direct dark matter search experiments and the projected sensitivities of the planned DarkSide-20k and Argo experiments with fiducial masses of 20 and 200 tonnes respectively. This figure is taken directly from the publication of the DarkSide-20k experiment [122].

A ELOG digest

The ArDM collaboration kept an electronic logbook (ELOG), wherein changes to the DAQ system or the experimental setup were updated manually. The following table is a summary of the ELOG entries, the ArDM data-processing MySQL database, and the mail traffic among the collaboration. The summary starts with the experimental run 288'967 taken on Feb. 12th, 2019 and ends with run 419'350 taken on Aug. 29th, 2019, which simultaneously concluded the data-taking period of ArDM. The ELOG entries were primarily used to reconstruct the time periods, where data was taken with the ^{252}Cf neutron calibration source (see Section 3.1.2), as well as finding periods of DAQ problems and constraining the valid data-taking range for the WIMP search (see Section 3.1).

#	**Run ID and date**	**Category**	**Event**
1	Run 288967; 12th Feburary 2019	First Data	First good SPD run
2	Run 292558; 02nd March 2019	DAQ	Arduino starts controlling Buffer level
3	Run 292967; 04th March 2019	DAQ	All Pedestals set to 4000
4	Run 293029; 04th March 2019	Data	First good DP run; Bias copper ring set to 4300V
5	Run 293474; 06th March 2019	DAQ	Arduino veto levels adjusted → DAQ running with 100% matching boards
6	Run 293847; 08th March 2019	DAQ	ZLE mode (on); followed by DAQ problems (Board #1 unreachable)
7	Run 293998; 09th March 2019	DAQ	DAQ problems again; Board #1 unreachable
8	Run 294683; 13th March 2019	DAQ	ZLE mode (off); positively tested connection of ZLE mode to DAQ problems
9	Run 295193; 15th March 2019	Data	Back to SPD
10	Runs 296018-296046; 19th March 2019	DAQ	Board #1 replaced; ZLE mode (on); Pedestals of new board set to 4000
11	Run 296064; 19th March 2019	Data	Back to DP
12	Runs 296354-296493; 20th March 2019	DAQ, Data	ZLE mode (off), quick SPD period; PMT HVs changed to reach 140 intADCcnts per SPE; back to DP in ZLE mode (on)
13	Runs 296641-296665; 21st March 2019	DAQ	ZLE mode (off); Trigger tests
14	Run 296742; 21st March 2019	DAQ	ZLE mode (on)
15	Run 297021; 22nd March 2019	DAQ	DAQ start at 16 μs and 1 min waiting time
16	Run 299482; 25th March 2019	DAQ, Data	ZLE mode (off), back to SPD

#	**Run ID and date**	**Category**	**Event**
17	Run 299575; 25th March 2019	DAQ, Data	Back to DP, ZLE mode (on)
18	Runs 312685-312771; 10th April 2019	DAQ	Trigger adjustments
19	Run 335408; 06th May 2019	Data	Back to SPD
20	Run 335821; 07th May 2019	Data	Back to DP
21	Runs 338959-339530; 10th-12th May 2019	DAQ, Data	Back to SPD; Ar refill bottle exchanged; DAQ waiting time set to 300 s; Switch to SP
22	Runs 344769-345023; 06th June 2019	Data	Runs with open cover (still SP)
23	Run 345025; 06th June 2019	Data	Back to DP
24	Runs 345027-345034; 06th June 2019	Calibration	Weak (∼100 Bq) ^{252}Cf neutron source placed
25	Run 345318; 06th June 2019	DAQ	DAQ waiting time set to 60 s
26	Run 353951; 19th June 2019	Data	Back to SPD
27	Run 354142; 19th June 2019	Data	Back to DP
28	Run 369039; 07th-08th July 2019	DAQ	HV failure → Power module in HV main frame replaced
29	Runs 369614-369900; 09th July 2019	Calibration	Weak (∼100 Bq) ^{252}Cf neutron source placed
30	Runs 369901-370392; 09th-10th July 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
31	Runs 380300-380309; 22nd July 2019	DAQ	Extraction voltage increased from 4.4 kV to 4.6 kV
32	Runs 380611-381097; 22nd-23rd July 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
33	Run 381014; 23rd July 2019	DAQ	DAQ waiting time changed to 35 s
34	Run 381098-381402; 23rd July 2019	Data veto	External neutron source placed in experimental hall for other experiment (CLYC)
35	Run 381403-382267; 23rd-24th July 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
36	Run 382340-383473; 24th-25th July 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
37	Run 382389-382630; 24th July 2019	Calibration	Extraction voltage scan from 4.2 kV to 4.6 kV (in steps of 0.1 kV; with ^{252}Cf neutron source)
38	Run 383332-383708; 25th July 2019	DAQ	Temporary switch to higher drift field of ∼350 V/cm

#	**Run ID and date**	**Category**	**Event**
39	∼Run 390545; 31th July 2019	DAQ	Argon gas bottle replaced
40	Run 390704-390784; 01st August 2019	Data veto	External neutron source placed in experimental hall for other experiment (TREX)
41	Run 390785-391551; 01st-02nd Aug. 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
42	Run 391551; 02nd August 2019	DAQ	DAQ crash
43	Run 395521-396486; 06th-07th Aug. 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
44	Run 403811-404770; 13th-14th Aug. 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
45	Run 412108-413080; 20th-21th Aug. 2019	Calibration	Strong (∼500 Bq) ^{252}Cf neutron source placed
46	Run 418019; 25th August 2019	Data	Last good DP run; then back to SPD (bad runs) and later to SP
47	Runs 418395-418732; 26th August 2019	Data	Last good SP runs
48	Run 419350; 29th August 2019	Last Data	Last ArDM run

B Contributions of the author

This appended section contains a summary of the contributions of the author in each chapter. The description also states where other collaborators made specific contributions and where the author summarized existing knowledge of the scientific community. Naturally, this doctoral thesis is built on many previous efforts of the ArDM collaboration which are not included in this summary.

Chapter 1 contains a short summary of the discovery of dark matter (Section 1.1), a brief overview of direct detection attempts of WIMPs (Section 1.2) and the derivation of the theoretical expectation of the WIMP's interaction rate on Earth (Section 1.3). All three parts are summaries of the present knowledge of the scientific community which are only assembled by the author.

Chapter 2 introduces the reader to the ArDM experiment including its backgrounds and discrimination variables as well as the ArDM detector (Section 2.1). This is followed by an extension of the theory from Chapter 1 to incorporate the physical effects in the interaction of WIMPs with liquid argon (Section 2.2). The content of this chapter consists of general knowledge of the ArDM collaboration and existing research on liquid noble gases which are only summarized by the author.

Chapter 3 contains a summary of the data statistics collected by ArDM, the selection of quality runs of ArDM data (Section 3.1) and a brief description of the first two data processing steps in the form of LRAnalysis (Section 3.2) and Bonsai (Section 3.3). The quality selection criteria for the runs of ArDM data were extracted by the author in collaboration with Prof. Dr. André Rubbia. Both software packages, LRAnalysis and Bonsai, were already existing as part of the ArDM data reconstruction before the arrival of the author. The author contributed minimal changes to the Bonsai software.

Chapter 4 describes the completion of the data processing chain in the form of Hebing. This includes the matching of primary and secondary scintillations to each other as well as the reconstruction of resulting combined event variables (Section 4.1). The chapter is concluded with a detailed description of the discovery of the Charge-up effect and the consequently necessitated correction thereof (Section 4.2). The ArDM matching algorithm in Hebing was primarily developed by Dr. Balint Radics of the ArDM collaboration. The contribution of the author lies in the evaluation and fine-tuning of event variables as well as the discovery and subsequent correction of the Charge-up effect. This correction also involved the improvement and adaptation of the matching algorithm according to the given circumstances.

Chapter 5 focuses on the analysis of the experimental backgrounds. This includes the extraction of the ArDM light yield (Section 5.1), the gauge of the detector's response to signal-like events (Section 5.2), the confinement of a fiducial volume cut (Section 5.3) and the evaluation of all data analysis cut efficiencies based on a subset of the ArDM data (Section 5.4). The extraction of the light yield has previously been performed by the ArDM collaboration and was updated and repeated by the author on ArDM's data under the latest data-taking conditions. The analysis of the response of the detector to signal-like events, as

well as the fine-tuning of all data analysis cuts and the evaluation of their efficiencies was also performed by the author. The PhD student Botao Li provided a major contribution via his work on improving the Monte Carlo simulation of the neutron emissions from the ^{252}Cf source (leading directly to Figs. 5.4 to 5.6 in Section 5.2.1) and the author received support in the data analysis from the master student Duy Lai during his master thesis.

Chapter 6 concludes with the WIMP search of the ArDM data. This includes an update of the theoretically expected WIMP spectrum according to the experimental conditions (Section 6.1), the evaluation of the expected background and the observed events in the WIMP search data (Section 6.2) and the thoroughly described statistical testing process involved in the WIMP search (Section 6.3). All results presented in this chapter are the work of the author.

Bibliography

[1] W. T. Kelvin. *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, (Cambridge University Press Warehouse1904). URL https://archive.org/details/baltimorelecture00kelviala.

[2] H. Poincaré. Bulletin de la société astronomique de france (January 1906).

[3] J. C. Kapteyn. First attempt at a theory of the arrangement and motion of the sidereal system. *The Astrophysical Journal*, volume 55 (1922): page 302. URL http://dx.doi.org/10.1086/142670.

[4] J. H. Oort. Observational evidence confirming Lindblad's hypothesis of a rotation of the galactic system. volume III (1927) (120). URL https://ui.adsabs.harvard.edu/abs/1927BAN.....3..275O/abstract.

[5] F. Zwicky. On the thermodynamic equilibrium in the universe. *Proceedings of the National Academy of Sciences*, volume 14 (1928) (7): pages 592–597. URL http://dx.doi.org/10.1073/pnas.14.7.592.

[6] E. Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, volume 15 (1929) (3): pages 168–173. URL http://dx.doi.org/10.1073/pnas.15.3.168.

[7] A. Einstein and W. de Sitter. On the relation between the expansion and the mean density of the universe. *Proceedings of the National Academy of Sciences*, volume 18 (1932) (3): pages 213–214. URL http://dx.doi.org/10.1073/pnas.18.3.213.

[8] J. H. Oort. The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems. volume VI (1932) (238). URL https://ui.adsabs.harvard.edu/abs/1932BAN.....6..249O/abstract.

[9] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, volume 6 (1933): pages 110–127. URL https://ui.adsabs.harvard.edu/abs/1933AcHPh...6..110Z/abstract.

[10] A. Boehle, et al. An improved distance and mass estimate for sgr a* from a multistar orbit analysis. *The Astrophysical Journal*, volume 830 (2016) (1): page 17. URL http://dx.doi.org/10.3847/0004-637x/830/1/17.

[11] G. Hinshaw, et al. Nine–year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results. *The Astrophysical Journal Supplement Series*, volume 208 (2013) (2): page 19. URL http://dx.doi.org/10.1088/0067-0049/208/2/19.

[12] P. A. R. Ade, et al. Planck2013 results. i. overview of products and scientific results. *Astronomy & Astrophysics*, volume 571 (2014): page A1. URL http://dx.doi.org/10.1051/0004-6361/201321529.

[13] P. J. McMillan. The mass distribution and gravitational potential of the milky way. *Monthly Notices of the Royal Astronomical Society*, volume 465 (2016) (1): pages 76–94. URL http://dx.doi.org/10.1093/mnras/stw2759.

[14] M. Weber and W. de Boer. Determination of the local dark matter density in our galaxy. *Astronomy and Astrophysics*, volume 509 (2010): page A25. URL http://dx.doi.org/10.1051/0004-6361/200913381.

[15] S. Bird, et al. Did LIGO detect dark matter? *Physical Review Letters*, volume 116 (2016) (20). URL http://dx.doi.org/10.1103/physrevlett.116.201301.

[16] X. Shi and G. M. Fuller. New dark matter candidate: Nonthermal sterile neutrinos. *Physical Review Letters*, volume 82 (1999) (14): pages 2832–2835. URL http://dx.doi.org/10.1103/physrevlett.82.2832.

[17] S. Weinberg. A new light boson? *Physical Review Letters*, volume 40 (1978) (4): pages 223–226. URL http://dx.doi.org/10.1103/physrevlett.40.223.

[18] N. Du, et al. Search for invisible axion dark matter with the axion dark matter experiment. *Physical Review Letters*, volume 120 (2018) (15). URL http://dx.doi.org/10.1103/physrevlett.120.151301.

[19] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Physics Reports*, volume 267 (1996) (5-6): pages 4, 99. URL http://dx.doi.org/10.1016/0370-1573(95)00058-5.

[20] M. van Beekveld, W. Beenakker, S. Caron, R. Peeters, and R. R. de Austri. Supersymmetry with dark matter is still natural. *Physical Review D*, volume 96 (2017) (3). URL http://dx.doi.org/10.1103/physrevd.96.035015.

[21] M. Pospelov. Secluded u(1) below the weak scale. *Physical Review D*, volume 80 (2009) (9). URL http://dx.doi.org/10.1103/physrevd.80.095002.

[22] B. Paczynski. Gravitational microlensing by the galactic halo. *The Astrophysical Journal*, volume 304 (1986): page 1. URL http://dx.doi.org/10.1086/164140.

[23] G. Steigman and M. S. Turner. Cosmological constraints on the properties of weakly interacting massive particles. *Nuclear Physics B*, volume 253 (1985): pages 375–386. URL http://dx.doi.org/10.1016/0550-3213(85)90537-1.

[24] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, volume 405 (2005) (5-6): pages 279–390. URL http://dx.doi.org/10.1016/j.physrep.2004.08.031.

[25] M. Milgrom. A modification of the newtonian dynamics as a possible alternative to the hidden mass hypothesis. *The Astrophysical Journal*, volume 270 (1983): page 365. URL http://dx.doi.org/10.1086/161130.

[26] C. Alcock, et al. The MACHO project: Microlensing results from 5.7 years of large magellanic cloud observations. *The Astrophysical Journal*, volume 542 (2000) (1): pages 281–307. URL http://dx.doi.org/10.1086/309512.

[27] P. Tisserand, et al. Limits on the macho content of the galactic halo from the EROS-2 survey of the magellanic clouds. *Astronomy & Astrophysics*, volume 469 (2007) (2): pages 387–404. URL http://dx.doi.org/10.1051/0004-6361:20066017.

[28] V. C. Rubin and J. F. W. Kent. Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, volume 159 (1969): page 379. URL http://dx.doi.org/10.1086/150317.

[29] V. C. Rubin, N. Thonnard, and J. F. W. K. Extended rotation curves of high-luminosity spiral galaxies. IV - systematic dynamical properties, SA through SC. *The Astrophysical Journal*, volume 225 (1978): page L107. URL http://dx.doi.org/10.1086/182804.

[30] V. C. Rubin, N. Thonnard, and J. F. W. K. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /r = 4kpc/ to UGC 2885 /r = 122 kpc/. *The Astrophysical Journal*, volume 238 (1980): page 471. URL http://dx.doi.org/10.1086/158003.

[31] D. N. Spergel, et al. First-year wilkinson microwave anisotropy probe (WMAP) observations: Determination of cosmological parameters. *The Astrophysical Journal Supplement Series*, volume 148 (2003) (1): pages 175–194. URL http://dx.doi.org/10.1086/377226.

[32] G. F. Smoot, et al. Structure in the COBE differential microwave radiometer first-year maps. *The Astrophysical Journal*, volume 396 (1992): page L1. URL http://dx.doi.org/10.1086/186504.

[33] C. L. Bennett, et al. Four-year [ITAL]COBE[/ITAL] DMR cosmic microwave background observations: Maps and basic results. *The Astrophysical Journal*, volume 464 (1996) (1): pages L1–L4. URL http://dx.doi.org/10.1086/310075.

[34] D. Clowe, A. Gonzalez, and M. Markevitch. Weak-lensing mass reconstruction of the interacting cluster 1e 0657-558: Direct evidence for the existence of dark matter. *The Astrophysical Journal*, volume 604 (2004) (2): pages 596–603. URL http://dx.doi.org/10.1086/381970.

[35] M. Markevitch, et al. Direct constraints on the dark matter self-interaction cross section from the merging galaxy cluster 1e 0657-56. *The Astrophysical Journal*, volume 606 (2004) (2): pages 819–824. URL http://dx.doi.org/10.1086/383178.

[36] D. Clowe, et al. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal*, volume 648 (2006) (2): pages L109–L113. URL http://dx.doi.org/10.1086/508162.

[37] W. B. Atwood, et al. THE LARGE AREA TELESCOPE ON THEFERMI GAMMA-RAY SPACE TELESCOPEMISSION. *The Astrophysical Journal*, volume 697 (2009) (2): pages 1071–1102. URL http://dx.doi.org/10.1088/0004-637x/697/2/1071.

[38] A. Drukier and L. Stodolsky. Principles and applications of a neutral-current detector for neutrino physics and astronomy. *Physical Review D*, volume 30 (1984) (11): pages 2295–2309. URL http://dx.doi.org/10.1103/physrevd.30.2295.

[39] M. W. Goodman and E. Witten. Detectability of certain dark-matter candidates. *Physical Review D*, volume 31 (1985) (12): pages 3059–3063. URL http://dx.doi.org/10.1103/physrevd.31.3059.

[40] S. Ahlen, F. Avignone, R. Brodzinski, A. Drukier, G. Gelmini, and D. Spergel. Limits on cold dark matter candidates from an ultralow background germanium spectrometer. *Physics Letters B*, volume 195 (1987) (4): pages 603–608. URL http://dx.doi.org/10.1016/0370-2693(87)91581-4.

[41] C. Amole, et al. Dark matter search results from the complete exposure of the PICO-60 c3f8 bubble chamber. *Physical Review D*, volume 100 (2019) (2). URL http://dx.doi.org/10.1103/physrevd.100.022001.

[42] A. Aguilar-Arevalo, et al. Results on low-mass weakly interacting massive particles from an 11 kg d target exposure of DAMIC at SNOLAB. *Physical Review Letters*, volume 125 (2020) (24). URL http://dx.doi.org/10.1103/physrevlett.125.241803.

[43] R. Bernabei, et al. First results from DAMA/LIBRA and the combined results with DAMA/NaI. *The European Physical Journal C*, volume 56 (2008) (3): pages 333–355. URL http://dx.doi.org/10.1140/epjc/s10052-008-0662-y.

[44] E. Armengaud, et al. Performance of the EDELWEISS-III experiment for direct dark matter searches. *Journal of Instrumentation*, volume 12 (2017) (08): pages P08010–P08010. URL http://dx.doi.org/10.1088/1748-0221/12/08/p08010.

[45] A. Abdelhameed, et al. First results from the CRESST-III low-mass dark matter program. *Physical Review D*, volume 100 (2019) (10). URL http://dx.doi.org/10.1103/physrevd.100.102002.

[46] J. Gascon, et al. Low-mass dark matter searches with edelweiss. http://arxiv.org/abs/2112.05467v1.

[47] E. Aprile, et al. Dark matter search results from a one ton-year exposure of XENON1t. *Physical Review Letters*, volume 121 (2018) (11). URL http://dx.doi.org/10.1103/physrevlett.121.111302.

[48] Y. Meng, et al. Dark matter search results from the pandax-4t commissioning run. (2021). URL http://dx.doi.org/10.48550/ARXIV.2107.13438.

[49] S. Westerdale. Analysis and dark matter search results from DEAP-3600 with 231 live days at SNOLAB. *Journal of Physics: Conference Series*, volume 1468 (2020) (1): page 012031. URL http://dx.doi.org/10.1088/1742-6596/1468/1/012031.

[50] P. Agnes, et al. DarkSide-50 532-day dark matter search with low-radioactivity argon. *Physical Review D*, volume 98 (2018) (10). URL http://dx.doi.org/10.1103/physrevd.98.102006.

[51] P. Agnes, et al. Results from the first use of low radioactivity argon in a dark matter search. *Physical Review D*, volume 93 (2016) (8). URL http://dx.doi.org/10.1103/physrevd.93.081101.

[52] J. Battat, et al. Low threshold results and limits from the DRIFT directional dark matter detector. *Astroparticle Physics*, volume 91 (2017): pages 65–74. URL http://dx.doi.org/10.1016/j.astropartphys.2017.03.007.

[53] R. Gaitskell. Dark Matter Direct Detection Experiment Review. In *Identification of Dark Matter - IDM 2020* (2020). URL https://indico.cern.ch/event/766367/contributions/3889100/attachments/2076463/3487902/200720_Dark_Matter_Direct_Detection_Gaitskell_IDM2020_July_2020_v4.pdf.

[54] J. Billard, et al. Direct Detection of Dark Matter – APPEC Committee Report. http://arxiv.org/abs/2104.07634v1.

[55] J. Lewin and P. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics*, volume 6 (1996) (1): pages 87–112. URL http://dx.doi.org/10.1016/s0927-6505(96)00047-3.

[56] C. McCabe. The earth's velocity for direct detection experiments. *Journal of Cosmology and Astroparticle Physics*, volume 2014 (2014) (02): pages 027–027. URL http://dx.doi.org/10.1088/1475-7516/2014/02/027.

[57] G. R. Knapp, S. D. Tremaine, and J. E. Gunn. The global properties of the galaxy. i - the h i distribution outside the solar circle. *The Astronomical Journal*, volume 83 (1978): page 1585. URL http://dx.doi.org/10.1086/112367.

[58] F. J. Kerr and D. Lynden-Bell. Review of galactic constants. *Monthly Notices of the Royal Astronomical Society*, volume 221 (1986) (4): pages 1023–1038. URL http://dx.doi.org/10.1093/mnras/221.4.1023.

[59] J. A. R. Caldwell and I. M. Coulson. Milky way rotation and the distance to the galactic center from cepheid variables. *The Astronomical Journal*, volume 93 (1987): page 1090. URL http://dx.doi.org/10.1086/114393.

[60] R. Schönrich, J. Binney, and W. Dehnen. Local kinematics and the local standard of rest. *Monthly Notices of the Royal Astronomical Society*, volume 403 (2010) (4): pages 1829–1833. URL http://dx.doi.org/10.1111/j.1365-2966.2010.16253.x.

[61] A. Bohr and B. R. Mottelson. *Nuclear Structure*, (World Scientific Publishing Company1998). URL http://dx.doi.org/10.1142/3530.

[62] R. H. Helm. Inelastic and elastic scattering of 187-mev electrons from selected even-even nuclei. *Phys. Rev.*, volume 104 (1956): pages 1466–1475. URL http://dx.doi.org/10.1103/PhysRev.104.1466.

[63] A. K. Drukier, K. Freese, and D. N. Spergel. Detecting cold dark-matter candidates. *Phys. Rev. D*, volume 33 (1986): pages 3495–3508. URL http://dx.doi.org/10.1103/PhysRevD.33.3495.

[64] S. Garbari, J. I. Read, and G. Lake. Limits on the local dark matter density. *Monthly Notices of the Royal Astronomical Society*, volume 416 (2011) (3): pages 2318–2340. URL http://dx.doi.org/10.1111/j.1365-2966.2011.19206.x.

[65] G. F. Knoll. *Radiation Detection and Measurement*, (WILEY2010). ISBN 0470131489. URL https://www.ebook.de/de/product/12251381/glenn_f_knoll_radiation_detection_and_measurement.html.

[66] R. Platzman. Total ionization in gases by high-energy particles: An appraisal of our understanding. *The International Journal of Applied Radiation and Isotopes*, volume 10 (1961) (2-3): pages 116–127. URL http://dx.doi.org/10.1016/0020-708x(61)90108-9.

[67] T. Doke, A. Hitachi, S. Kubota, A. Nakamoto, and T. Takahashi. Estimation of fano factors in liquid argon, krypton, xenon and xenon-doped liquid argon. *Nuclear Instruments and Methods*, volume 134 (1976) (2): pages 353–357. URL http://dx.doi.org/10.1016/0029-554x(76)90292-5.

[68] T. Doke, A. Hitachi, J. Kikuchi, K. Masuda, H. Okada, and E. Shibamura. Absolute scintillation yields in liquid argon and xenon for various particles. *Japanese Journal of Applied Physics*, volume 41 (2002) (Part 1, No. 3A): pages 1538–1545. URL http://dx.doi.org/10.1143/jjap.41.1538.

[69] V. Chepel and H. Araújo. Liquid noble gas detectors for low energy particle physics. *Journal of Instrumentation*, volume 8 (2013) (04): pages R04001–R04001. URL http://dx.doi.org/10.1088/1748-0221/8/04/r04001.

[70] W. H. Lippincott, et al. Erratum: Scintillation time dependence and pulse shape discrimination in liquid argon [phys. rev. c78, 035801 (2008)]. *Physical Review C*, volume 81 (2010) (3). URL http://dx.doi.org/10.1103/physrevc.81.039901.

[71] T. M. Ito, et al. Effect of an electric field on superfluid helium scintillation produced by α-particle sources. *Physical Review A*, volume 85 (2012) (4). URL http://dx.doi.org/10.1103/physreva.85.042718.

[72] W. Walkowiak. Drift velocity of free electrons in liquid argon. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, volume 449 (2000) (1-2): pages 288–294. URL http://dx.doi.org/10.1016/s0168-9002(99)01301-7.

[73] P. Sorensen and C. E. Dahl. Nuclear recoil energy scale in liquid xenon with application to the direct detection of dark matter. *Physical Review D*, volume 83 (2011) (6). URL http://dx.doi.org/10.1103/physrevd.83.063501.

[74] E. Aprile, et al. Simultaneous measurement of ionization and scintillation from nuclear recoils in liquid xenon for a dark matter experiment. *Physical Review Letters*, volume 97 (2006) (8). URL http://dx.doi.org/10.1103/physrevlett.97.081302.

[75] P. Benetti, et al. First results from a dark matter search with liquid argon at 87k in the gran sasso underground laboratory. *Astroparticle Physics*, volume 28 (2008) (6): pages 495–507. URL http://dx.doi.org/10.1016/j.astropartphys.2007.08.002.

[76] T. Washimi, M. Kimura, M. Tanaka, and K. Yorita. Scintillation and ionization ratio of liquid argon for electronic and nuclear recoils at drift-fields up to 3 kV/cm. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, volume 910 (2018): pages 22–25. URL http://dx.doi.org/10.1016/j.nima.2018.09.019.

[77] B.-R. Montes Núñez. *Analysis of the first underground run and background studies of the Argon Dark Matter experiment*. Ph.D. thesis, UCM, Madrid, Dept. Phys., Madrid U. (April 2016). URL https://eprints.ucm.es/id/eprint/44127/.

[78] W. Mu. *Direct WIMPs Search with the ArDM Experiment and Preliminary Results from 1.43 Days of Physics Data*. Ph.D. thesis (March 2019). URL http://dx.doi.org/10.3929/ETHZ-B-000335024.

[79] A. Bettini. The canfranc underground laboratory (LSC). *The European Physical Journal Plus*, volume 127 (2012) (9). URL http://dx.doi.org/10.1140/epjp/i2012-12112-1.

[80] S. Himi, T. Takahashi, J. zhi Ruan(Gen), and S. Kubota. Liquid and solid argon, and nitrogen-doped liquid and solid argon scintillators. *Nuclear Instruments and Methods in Physics Research*, volume 203 (1982) (1-3): pages 153–157. URL http://dx.doi.org/10.1016/0167-5087(82)90623-8.

[81] R. Acciarri, et al. Effects of nitrogen contamination in liquid argon. *Journal of Instrumentation*, volume 5 (2010) (06): pages P06003–P06003. URL http://dx.doi.org/10.1088/1748-0221/5/06/p06003.

[82] R. Acciarri, et al. Oxygen contamination in liquid Argon: combined effects on ionization electron charge and scintillation light. *Journal of Instrumentation*, volume 5 (2010) (05): pages P05003–P05003. URL http://dx.doi.org/10.1088/1748-0221/5/05/p05003.

[83] J. Calvo, et al. Measurement of the attenuation length of argon scintillation light in the ArDM LAr TPC. *Astroparticle Physics*, volume 97 (2018): pages 186–196. URL http://dx.doi.org/10.1016/j.astropartphys.2017.11.009.

[84] J. Lindhard, V. Nielsen, M. Scharff, and P. Thomsen. Integral equations governing radiation effects. *Mat. Fys. Medd. Dan. Vid. Selsk*, volume 33 (1963) (10): pages 1–42.

[85] D.-M. Mei, Z.-B. Yin, L. Stonehill, and A. Hime. A model of nuclear recoil scintillation efficiency in noble liquids. *Astroparticle Physics*, volume 30 (2008) (1): pages 12–17. URL http://dx.doi.org/10.1016/j.astropartphys.2008.06.001.

[86] P. Agnes, et al. Simulation of argon response and light detection in the DarkSide-50 dual phase TPC. *Journal of Instrumentation*, volume 12 (2017) (10): pages P10015–P10015. URL http://dx.doi.org/10.1088/1748-0221/12/10/p10015.

[87] M. Kimura, M. Tanaka, T. Washimi, and K. Yorita. Measurement of the scintillation efficiency for nuclear recoils in liquid argon under electric fields up to 3 kV/cm. *Physical Review D*, volume 100 (2019) (3). URL http://dx.doi.org/10.1103/physrevd.100.032002.

[88] P. Agnes, et al. Measurement of the liquid argon energy response to nuclear and electronic recoils. *Physical Review D*, volume 97 (2018) (11). URL http://dx.doi.org/10.1103/physrevd.97.112005.

[89] A. Hitachi and A. Mozumder. Properties for liquid argon scintillation for dark matter searches . http://arxiv.org/abs/1903.05815v1.

[90] J. Thomas and D. A. Imel. Recombination of electron-ion pairs in liquid argon and liquid xenon. *Physical Review A*, volume 36 (1987) (2): pages 614–616. URL http://dx.doi.org/10.1103/physreva.36.614.

[91] H. Cao, et al. Measurement of scintillation and ionization yield and scintillation pulse shape from nuclear recoils in liquid argon. *Physical Review D*, volume 91 (2015) (9). URL http://dx.doi.org/10.1103/physrevd.91.092007.

[92] B. Lenardo, K. Kazkaz, A. Manalaysay, J. Mock, M. Szydagis, and M. Tripathi. A global analysis of light and charge yields in liquid xenon. *IEEE Transactions on Nuclear Science*, volume 62 (2015) (6): pages 3387–3396. URL http://dx.doi.org/10.1109/tns.2015.2481322.

[93] N. Gehrels. Confidence limits for small numbers of events in astrophysical data. *The Astrophysical Journal*, volume 303 (1986): page 336. URL http://dx.doi.org/10.1086/164079.

[94] R. P. Kraft, D. N. Burrows, and J. A. Nousek. Determination of confidence limits for experiments with low numbers of counts. *The Astrophysical Journal*, volume 374 (1991): page 344. URL http://dx.doi.org/10.1086/170124.

[95] G. J. Feldman and R. D. Cousins. Unified approach to the classical statistical analysis of small signals. *Physical Review D*, volume 57 (1998) (7): pages 3873–3889. URL http://dx.doi.org/10.1103/physrevd.57.3873.

[96] A. Stauffer. ArDM - Run Selection and Cuts - October 2019. Technical Report, ETH Zürich (2021). URL https://gitlab.cern.ch/ArDM/Analysis/-/blob/master/daqtrf/run_selection_and_cuts.pdf.

[97] P. Benetti, et al. Measurement of the specific activity of 39ar in natural argon. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, volume 574 (2007) (1): pages 83–88. URL http://dx.doi.org/10.1016/j.nima.2007.01.106.

[98] A. Buzulutskov. Electroluminescence and electron avalanching in two-phase detectors. (2020). URL http://dx.doi.org/10.48550/ARXIV.2006.11017.

[99] G. Landi. Properties of the center of gravity as an algorithm for position measurements. (2019). URL http://dx.doi.org/10.48550/ARXIV.1908.04447.

[100] Y. Mei. *Direct Dark Matter Search with the XENON100 Experiment*. Ph.D. thesis, Rice University (2012). URL https://hdl.handle.net/1911/70350.

[101] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, volume 69 (1962) (1): page 9. URL http://dx.doi.org/10.2307/2312726.

[102] B. Radics, E. Burjons, and A. Rubbia. Matching problem for primary and secondary signals in dual-phase tpc detectors. (2018). URL http://dx.doi.org/10.48550/ARXIV.1809.01006.

[103] Q. D. Lai. *The ArDM Experiment - Quantification of Electronic and Nuclear Recoils at Low Energies*. Master thesis, ETH Zürich (2022).

[104] R. Saldanha, et al. Cosmogenic production of ^{39}Ar and ^{37}Ar in argon. *Physical Review C*, volume 100 (2019) (2): page 024608. URL http://dx.doi.org/10.1103/physrevc.100.024608.

[105] Nguyen, Khoi Nguyen. *Proof of concept for the direct WIMP search with ArDM and analysis of its first data*. Ph.D. thesis (2016). URL http://dx.doi.org/10.3929/ETHZ-A-010616022.

[106] A. Koning, D. Rochman, J.-C. Sublet, N. Dzysiuk, M. Fleming, and S. van der Marck. TENDL: Complete nuclear data library for innovative nuclear science and technology. *Nuclear Data Sheets*, volume 155 (2019): pages 1–55. URL http://dx.doi.org/10.1016/j.nds.2019.01.002.

[107] D. Baxter, et al. Recommended conventions for reporting results from direct dark matter searches. *The European Physical Journal C*, volume 81 (2021) (10). URL http://dx.doi.org/10.1140/epjc/s10052-021-09655-y.

[108] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, volume 9 (1938) (1): pages 60–62. URL http://dx.doi.org/10.1214/aoms/1177732360.

[109] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. (2010). URL http://dx.doi.org/10.48550/ARXIV.1007.1727.

[110] K. Cranmer. Practical statistics for the lhc. (2015). URL http://dx.doi.org/10.48550/ARXIV.1503.07622.

[111] A. Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, volume 54 (1943) (3): pages 426–482. URL http://dx.doi.org/10.1090/s0002-9947-1943-0012401-3.

[112] A. L. Read. Modified frequentist analysis of search results (the CL_s method) (2000). URL http://dx.doi.org/10.5170/CERN-2000-005.81.

[113] A. L. Read. Presentation of search results: the CL(s) technique. *Journal of Physics G: Nuclear and Particle Physics*, volume 28 (2002) (10): pages 2693–2704. URL http://dx.doi.org/10.1088/0954-3899/28/10/313.

[114] A. Tumasyan, et al. Search for invisible decays of the higgs boson produced via vector boson fusion in proton-proton collisions at $\sqrt{s} = 13$TeV. *Physical Review D*, volume 105 (2022) (9): page 092007. URL http://dx.doi.org/10.1103/physrevd.105.092007.

[115] 50 Collaboration, et al. Search for low-mass dark matter wimps with 12 ton-day exposure of darkside-50 (2022). URL http://dx.doi.org/10.48550/ARXIV.2207.11966.

[116] P. Agnes, et al. CALIS—a CALibration insertion system for the DarkSide-50 dark matter search experiment. *Journal of Instrumentation*, volume 12 (2017) (12): pages T12004–T12004. URL http://dx.doi.org/10.1088/1748-0221/12/12/t12004.

[117] A. Hime. The miniclean dark matter experiment (2011). URL http://dx.doi.org/10.48550/ARXIV.1110.1005.

[118] K. Rielage, et al. Update on the MiniCLEAN dark matter experiment. *Physics Procedia*, volume 61 (2015): pages 144–152. URL http://dx.doi.org/10.1016/j.phpro.2014.12.024.

[119] J. Xu, et al. A study of the trace ^{39}Ar content in argon from deep underground sources. *Astroparticle Physics*, volume 66 (2015): pages 53–60. URL http://dx.doi.org/10.1016/j.astropartphys.2015.01.002.

[120] H. O. Back, et al. First large scale production of low radioactivity argon from underground sources (2012). URL http://dx.doi.org/10.48550/ARXIV.1204.6024.

[121] M. Wada. Wimp search with underground argon in darkside-50. Technical Report (2017). URL http://dx.doi.org/10.3204/DESY-PROC-2009-03/MASAYUKI_WADA.

[122] C. E. Aalseth, et al. DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS. *The European Physical Journal Plus*, volume 133 (2018) (3). URL http://dx.doi.org/10.1140/epjp/i2018-11973-4.

[123] E. S. Garcia. DArT, a detector for measuring the ^{39}Ar depletion factor. *Journal of Instrumentation*, volume 15 (2020) (02): pages C02044–C02044. URL http://dx.doi.org/10.1088/1748-0221/15/02/c02044.

[124] P. Agnes, et al. Separating ^{39}Ar from ^{40}Ar by cryogenic distillation with aria for dark-matter searches. *The European Physical Journal C*, volume 81 (2021) (4). URL http://dx.doi.org/10.1140/epjc/s10052-021-09121-9.

[125] 20k Collaboration. Measurement of isotopic separation of argon with the prototype of the cryogenic distillation plant aria for dark matter searches (2023). URL http://dx.doi.org/10.48550/ARXIV.2301.09639.

[126] C. E. Aalseth, et al. Cryogenic characterization of FBK RGB-HD SiPMs. *Journal of Instrumentation*, volume 12 (2017) (09): pages P09030–P09030. URL http://dx.doi.org/10.1088/1748-0221/12/09/p09030.

[127] J. Monroe and P. Fisher. Neutrino backgrounds to dark matter searches. (2007). URL http://dx.doi.org/10.48550/ARXIV.0706.3019.

[128] J. D. Vergados and H. Ejiri. Can solar neutrinos be a serious background in direct dark matter searches? (2008). URL http://dx.doi.org/10.48550/ARXIV.0805.2583.

[129] L. E. Strigari. Neutrino coherent scattering rates at direct dark matter detectors. (2009). URL http://dx.doi.org/10.48550/ARXIV.0903.3630.

[130] A. Gütlein, et al. Solar and atmospheric neutrinos: Background sources for the direct dark matter searches. (2010). URL http://dx.doi.org/10.48550/ARXIV.1003.5530.

[131] J. Billard, L. Strigari, and E. Figueroa-Feliciano. Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments. (2013). URL http://dx.doi.org/10.48550/ARXIV.1307.5458.

[132] P. Grothaus, M. Fairbairn, and J. Monroe. Directional dark matter detection beyond the neutrino bound. *Physical Review D*, volume 90 (2014) (5): page 055018. URL http://dx.doi.org/10.1103/physrevd.90.055018.

[133] C. A. J. O'Hare, A. M. Green, J. Billard, E. Figueroa-Feliciano, and L. E. Strigari. Readout strategies for directional dark matter detection beyond the neutrino background. (2015). URL http://dx.doi.org/10.48550/ARXIV.1505.08061.