The Search for Low Mass Dark Matter with the NEWS-G and SBC Experiments

by

Daniel J. Durnford

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics
University of Alberta

Abstract

The search for particle dark matter has been ongoing for several decades to date, without any conclusive evidence of detection events. While large, tonne-scale experiments accumulate exposure (larger masses and longer data taking) in hopes of finding this elusive signal, other experiments are expanding the search to include novel detector technologies and non-traditional dark matter candidates. Even more so than usual, it is critical to have a robust understanding of the physics of these detectors to make a credible claim of discovery, or to exclude the existence of dark matter.

Among these are the NEWS-G and SBC collaborations, which utilize spherical proportional counters (SPCs) and scintillating bubble chambers (SBCs) respectively. This thesis presents the progress made in the characterization and understanding of the physics relevant to both technologies. This includes measurements of the ionization yield of SPCs, the development of novel radon removal systems, and the measurement of the bubble creation efficiency of nuclear recoils in both more traditional C_3F_8 bubble chambers, as well as a liquid xenon SBC. Also presented is the full analysis and physics results of a NEWS-G dark matter search campaign in pure methane gas, establishing world-leading exclusions limits on low-mass dark matter with a spin-dependent coupling with protons.

Preface

The work presented in this thesis was completed by myself, Daniel Durnford, as a member of the NEWS-G and SBC collaborations, and under the supervision of Dr. Marie-Cécile Piro at the University of Alberta. The nature of this work was collaborative, building on the efforts of others who designed and constructed the detectors used, and benefiting from past analysis efforts.

The W-value measurements presented in sections 3.2 and 3.3 were carried out by myself, in collaboration with Philippe Gros and Alexis Brossard, as well as other NEWS-G collaborators at Queen's University who assisted in data taking. A publication based on this work is currently in preparation. The radon trap measurements described in section 3.4 built upon previous measurements by Patrick O'Brien, Carter Garrah, Yuqi Deng, Julieta Hernandez, and Marie-Cécile Piro. The analyses presented in this section were performed by myself. A publication based on these results is currently in preparation. The fabrication of the U of A SPC detector used for this study was aided by Philippe Gros.

The S140 detector was designed and fabricated by the NEWS-G collaboration and other technicians, detailed in ref. [1] (principally authored by Alexis Brossard). The ^{210}Pb assay for the S140 detector described in section 4.1 was a joint effort of the NEWS-G and XMASS collaborations (the latter collecting the data), with the analysis performed by myself. This result was included in a broader publication [2], which was principally authored by Patrick Knights.

I assisted in the installation and operation of the detector at the Laboratoire Souterrain de Modane, including the collection of the physics data presented in Chap-

ter 4. I was involved in all aspects of the subsequent dark matter analysis, with large portions being solely my own work. Specific collaborators are acknowledged throughout this chapter where relevant (the work was solely my own unless otherwise stated). Contributions were made by Francisco Vazquez de Sola, Jean-Marie Coquillat, Carter Garrah, George Savvidis, Aviv Padawer-Blatt, and Yuqi Deng. Publications based on this study are currently in preparation.

Chapter 5 – detailing nucleation efficiency calibrations for both the PICO and SBC collaborations – built on the previous analysis work and experimental efforts of others. Specifically, the data acquired for the PICO study was gathered by collaborators at the Université de Montreal and other sites. A preliminary analysis was performed by Eric Dahl and Miaotianzi Jin. The final results, including my analyses presented in sections 5.3 – 5.4 and 5.7, were published in ref. [3]. The composition of this manuscript was done in collaboration with Marie-Cécile Piro, with whom I share primary authorship. The data used for the xenon bubble chamber analysis presented in sections 5.5 – 5.6 and 5.7 was gathered by Eric Dahl, Matthew Bressler, and other collaborators at Northwestern University. The simulations needed for the analysis were performed by Ariel Zuñiga. The subsequent analysis was my own, with suggestions and contributions by Eric Dahl and Russell Neilson. A publication featuring these results is currently in preparation.

My doctoral studies were supported by funding from the Natural Sciences and Engineering Research Council of Canada (the Alexander Graham Bell CGS-D award), the Alberta Graduate Excellence Scholarship, and the Arthur B. McDonald Canadian Astroparticle Physics Research Institute.

Acknowledgements

It is an odd experience to look at this finished document, and think of it as the cumulative summary of the last 5 years of life. It seems rather long, but at the same time it leaves out uncountable hours of work. In between the lines of these pages there was also my wedding, a global pandemic, multiple moves, multiple comings and goings of loved ones, and the arrival of a particularly crazy cat. But here we are, 5 years – and a little shy of 400 pages – later.

This moment would certainly would not have been reached without the help and guidance of my supervisor, Professor Marie-Cécile Piro, who gifted to me her unwavering support from day one – I cannot thank you enough for that! As I have been a member of the NEWS-G collaboration for more than 7 years, the list of people I have to thank extends back in time long before my PhD studies, beginning with Gilles Gerbier for opening this chapter of my life at Queen's University. I extend many thanks to all my collaborators there, at the University of Alberta, and all over the world. There are more of you than I can name here, but I certainly would not have accomplished much without Alexis Brossard, Philippe Gros, Maurice Chapellier, Anouchka Ronceray, Julie McDonald, Marie Vidal, George Savvidis, Carter Garrah, Yuqi Deng, Patrick O'Brien; you are all superb! A special thanks also goes to Francisco Vazquez de Sola and Jean-Marie Coquillat, my companions through the long, dark forest of the LSM dark matter analysis; I can say with great confidence that the paper will only take two more weeks! Without any doubt in my mind, I can ascribe any success I might achieve in physics to Quentin Arnaud, whose friendship, mentor-ship, and confidence in me at a critical time set me on my current path.

During my PhD studies, I was also welcomed into another collaboration (and also sort of a third) with enthusiasm, and lots of support. I would like to thank Eric Dahl for affording me the opportunity to work on two very rewarding analyses for PICO and SBC, as well as my coauthors on both papers. Their efforts (spanning back 10 years!) allowed me to be there for both projects to cross the finish line. I would also like to thank Matthew Bressler for patiently answering my many questions when I started with SBC.

Of course, I also owe a great deal to my family for supporting me in my studies, especially my mother Carol. She first sparked my love of science at a young age, by creating endless enrichment opportunities for me. Although I'm sure you didn't always understand what I was working on, or why, or why I had to stay up working so late, or why I had to subject you to constant fear by going hiking in the mountains alone, I always knew I could count on your support.

Finally – and with all the sincerity in my heart – my biggest thank you is for my wife, Felicia. I am in constant awe of you, the grace with which you've adapted your life to enable me to pursue this pipe-dream career path, and the tenacity with which you still pursue your own lofty goals. I could not in my wildest dreams imagine a better companion for all of our crazy adventures, and to build a life with.

Table of Contents

List of Tables

List of Figures

Abbreviations

ADU Arbitrary digital units.

CAD Computer-aided design.

CEvNS Coherent elastic neutrino nucleus scattering.

CL Confidence level.

CMB Cosmic microwave background.

DAQ Data acquisition.

FC Feldman-Cousins.

FN False negative.

FNV False negative veto.

FP False positive.

HDPE High density polyethylene.

KDE Kernel density estimation.

LAr Liquid argon.

LAS Laser absorption spectroscopy.

LCL Lower confidence limit.

LSM Laboratoire Souterrain de Modane.

LXe Liquid xenon.

MC Monte Carlo.

MCMC Markov chain Monte Carlo.

MH Metropolis Hastings.

NBT Negative baseline transient.

NEWS-G New Experiments with Spheres – Gas.

PD Photodiode.

PLR Profile likelihood ratio.

PNNL Pacific Northwest National Lab.

PSD Pulse shape discrimination.

SBC Scintillating bubble chamber.

SD Spin dependent.

SI Spin independent.

SPC Spherical proportional counter.

SZ Silver zeolite.

UCL Upper confidence limit.

UV Ultraviolet.

WIMP Weakly interacting massive particle.

Chapter 1

Introduction

For nearly a century, the existence and nature of dark matter has been one of the most intriguing problems of astrophysics, and by extension, particle physics [4]. The persistent and nagging lack of discovery of postulated dark matter candidates has prompted consideration of alternative, theoretically motivated solutions for the astrophysical quandary, such as low-mass particle dark matter [5–7]. Numerous experimental efforts are now underway to stake out this largely uncharted region of parameter space in the hopes of either discovering a positive dark matter signal or at least ruling out another plausible class of candidates [8, 9]. The New Experiments with Spheres – Gas (NEWS-G) and Scintillating Bubble Chamber (SBC) collaborations are two such groups. The former utilizes proportional counters with light-atom gas mixtures, while the latter employs liquid noble bubble chambers. Like all experiments in this field, both types of detectors must be constructed with a high degree of radio-purity, operated in shielded environments, and are subject to the stringent requirements of calibration and characterization.

Chapter 2 of this work introduces the long-running search for dark matter, and the field of direct detection experiments more specifically. The context of the NEWS-G and SBC experiments within this active area of study is presented, as well as the overall status of the field. Chapter 3 introduces the NEWS-G experiment and the physics of Spherical Proportional Counters (SPCs). The mainstay calibration techniques of

the experiment are presented, as well as their application to dedicated measurements of the energy response properties of these detectors with a small, lab-based SPC experiment. A similar prototype-scale detector was installed at the University of Alberta, which has been used to develop and characterize a novel gas-handling apparatus to remove radon from the target gas, and monitor the composition of the gas in real-time. Future applications of this SPC are explored.

Over the course of my M.Sc. (during which time I worked with NEWS-G) and Ph.D. studies, I have taken part in the entire development cycle of the collaboration's dark matter search experiment, from fabrication to analysis and WIMP results. This included the construction of a next-generation detector, subsequent installation and data taking at the Laboratoire Souterrain de Modane, which will soon culminate in the publication of world-leading constraints on low-mass dark matter interactions on a hydrogen-rich target. The details of this detector, and the methane target gas dark matter analysis, are presented in Chapter 4.

The SBC collaboration is currently designing twin liquid argon detectors intended to search for low-mass dark matter and coherent elastic neutrino scattering [10]. In the meantime, existing calibration data taken with C_3F_8 (cooperatively with the PICO collaboration) and liquid xenon bubble chambers was analyzed to measure the response of both media to low-energy nuclear recoils. These efforts are given in Chapter 5, which represents the most robust measurements of the nucleation efficiency of bubble chambers to date. The C_3F_8 result has been published [3], with the liquid xenon result to follow shortly. This quantity is critical for the calculation of these experiments' sensitivity to dark matter.

Chapter 2

Particle Dark Matter

In the early 1920's, the prevailing concept of the Milky Way was that of an island Universe. Measurements of the local stellar density by Jeans and Kapteyn [11], and later Oort in 1932 [12], lead them to conclude that more mass than was visible must be present, speculated to be un-observable stars or gas. In 1933, Fritz Zwicky's measurement of the velocities of galaxies in the Coma Cluster allowed him to calculate that the mass of the cluster was 400 times greater than its luminous contents could account for [11, 13]. In the 1970's, Vera Rubin *et al* carried out measurements of the rotational velocities of objects within galaxies, starting with the nearby M31 [14–17]. These measurements – which bore out the same conclusion in many separate observations – demonstrated that the mass of the galaxies observed far out-weighed their luminous content, reinforcing Zwicky's proposed "dunkle Materie" – or dark matter – but on a galactic scale. Following these ground-breaking observations, the theory-based and experimental work to explain these phenomena began in earnest.

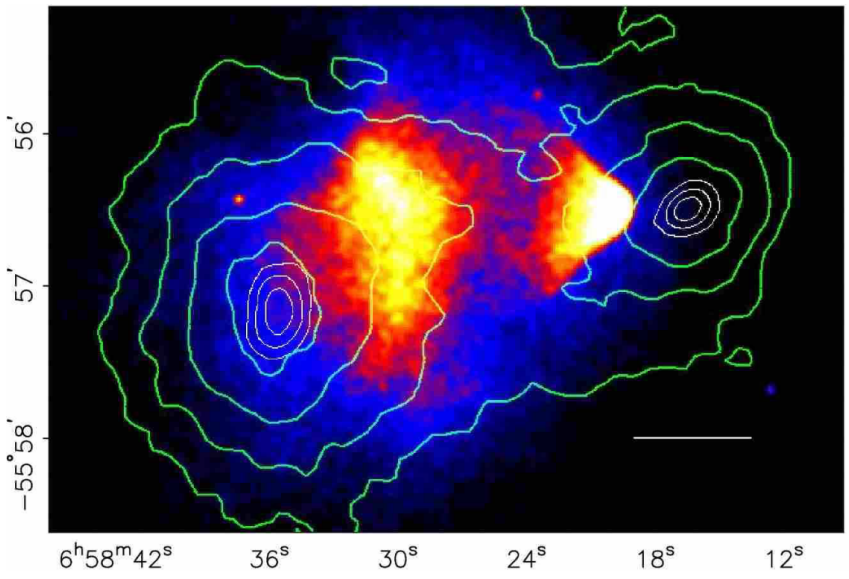


Figure 2.1: The bullet cluster imaged in X-ray (colour-scale), with the contour lines indicating the mass distribution reconstructed from gravitational lensing data. The white bar is shown for scale (200 kPc at the distance of the cluster). This demonstrates the separation between intergalactic gas/dust (prominent in X-ray) and the strongest gravitational potential of the clusters. Taken from figure 1 of ref. [20].

2.1 Theories and evidence for Dark Matter

2.1.1 Modified theories of gravity

An early class of hypotheses to explain the apparent extra mass in the Universe (or more accurately, unexplained gravitational force) posited that our understanding of gravity on cosmic scales is incorrect. Thus these theories primarily involve corrections to Newtonian gravity – MOND (MOdified Newtonian Dynamics) [18]. MOND theories are capable of explaining the anomalous galaxy rotation curves measured by Rubin and in other subsequent measurements [16, 18, 19]. Their central conjecture is that the phenomenon of "dark matter" caused by an undiscovered scale dependence in the strength of gravity, leading to stronger attractive forces between objects spanning galactic or intergalactic distances.

However, some observed phenomena currently cannot be explained by MOND theories, such as the Bullet Cluster [20, 21], pictured in Fig. 2.1. In this object, two

galaxy clusters are observed post-collision, the act of which stripped away much of the intergalactic dust and gas present in both clusters originally. This is observable in X-ray imaging due to the heat imbued by the collision. Although this gas and dust is known to make up the majority of the baryonic matter in galaxy clusters, a gravitational micro-lensing observation of the object indicates that the majority of their mass still resides in the clusters themselves [20, 21]. This is evidence that (in this case) the majority of the gravitational potential of the clusters does not coincide with the majority of their baryonic matter, refuting the MOND-picture of this event which would posit no additional source of gravity, just an increased potential caused by the luminous content of the clusters.

Recently, observations of galaxies with little or no dark matter (i.e. no anomalous galactic rotation curve) have been made [22–24]. Somewhat counter-intuitively, this also refutes MOND-like theories, which would predict the same "anomalous gravity" behaviour for all galaxies in the Universe with no exception. Thus, galaxies lacking dark matter suggest that the cause of this phenomenon is some substance exerting additional gravitational attraction that is not uniformly distributed throughout the Universe.

2.1.2 Non-luminous astrophysical objects

Another set of hypotheses proposes that the observed dark matter consists of normal, baryonic matter that simply is not visible to telescopes (optically or in any other wavelength band). Asteroids and comets, planets or planetesimals, and black or brown dwarf stars are all credible examples of baryonic matter that would evade detection by telescopes. This hypothesis is appealing because it does not require the existence of undiscovered physics to explain the mystery of dark matter.

However, efforts to find such objects with gravitational lensing surveys (such as EROS and OGLE) have not observed enough objects for them to account for all (or even a significant fraction of) dark matter [25–28]. These searches monitor a wide

field of stars for small fluctuations in their brightness caused by gravitational lensing of astrophysical objects passing in front of the background stars; the resultant gravitational lensing can distort or multiply the image of the star, causing an increase in brightness. While these surveys are only sensitive to objects in the mass range of 30 to 10^7 solar masses [27, 28], if dark matter was dominated by smaller objects, the number density required to achieve the gravitational effect observed would imply the existence of many such objects even locally within our own solar system. Astrophysical objects on larger mass scales have also been constrained by surveys looking for stellar binary systems that would be disrupted by encounters with such objects [29]. Moreover, specific survey results that do suggest the existence of populations of compact objects indicate that they inhabit the disk of the galaxy, rather than the spherical halo where dark matter is expected to reside [26, 30].

One specific case of a compact astrophysical object that remains a viable and interesting dark matter candidate is primordial black holes (PBHs). These are hypothesized black holes – with a wide possible mass range of 10^{-2} to 10^5 solar masses – formed before recombination (the moment electrons became bound to nuclei) in the early Universe from extreme density fluctuations in the primordial soup [31]. The existence of PBHs could also help explain the prevalence of supermassive black holes in the Universe at later times, with PBHs serving as possible progenitors [32]. Pulsar timing array experiments (e.g. NANOGrav) and gravitational wave interferometers (e.g. LIGO, LISA) have the potential to provide evidence of PBHs, rekindling interest in this dark matter candidate [31, 33–35].

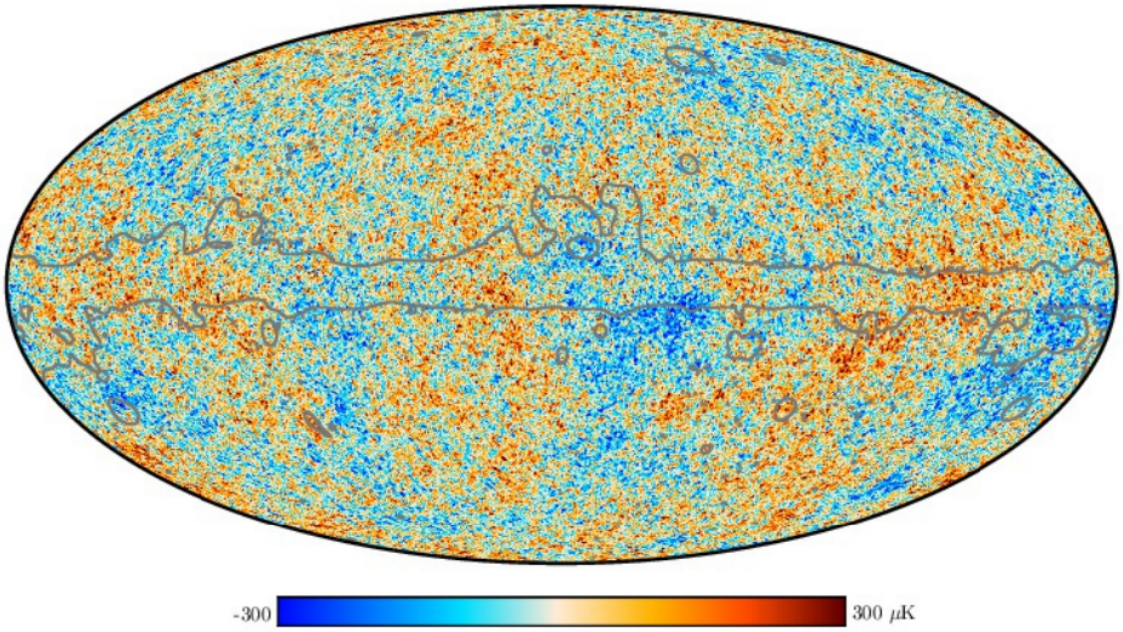
2.1.3 Non-baryonic particles and ΛCDM

The remaining category of hypotheses is that dark matter consists of one or more undiscovered, fundamental particles. If present in sufficiently high density, electrically-neutral (non-luminous) particles could provide the anomalous gravitational potential known since the 20^{th} century, with the Milky Way and other galaxies residing

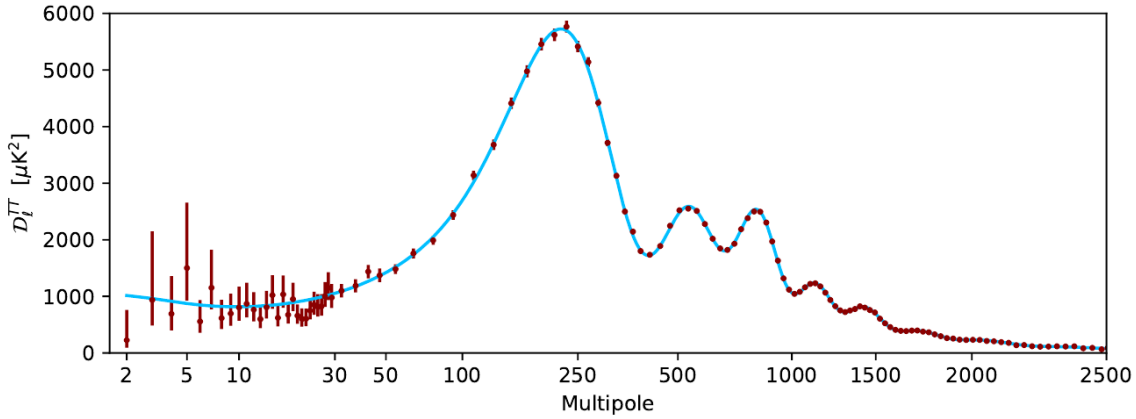
in fuzzy clouds of these particles extending far beyond their luminous bodies. Although the Standard Model (SM) of particle physics does include some seemingly viable candidates, none can satisfy all the requirements of dark matter. Free neutrons decay too quickly to provide the pervasive, stable influence of dark matter, and Standard Model neutrinos are relativistic, such that they cannot remain gravitationally bound to structures such as galaxies. Thus, dark matter must come from beyond the Standard Model. Undiscovered particle dark matter can be seen as a deductive implication of the challenges faced by other hypotheses, described above, but the general concept is supported independently as well. Perhaps the most compelling evidence for this paradigm comes from the Cosmic Microwave Background (CMB); the virtually uniform wall of photons arriving from the early Universe, red-shifted to micro-wavelengths over the course of their 13.8 billion year journey.

This light, released during recombination, contains information about the interactions of the constituents of the early Universe [37]. Despite the temperature of these photons fluctuating by less than a milli-Kelvin, and the obscuring collage of astronomical foregrounds to contend with, the temperature map of the CMB has been measured to high precision by several experiments, notably WMAP and Planck [36, 38]. The angular size of the fluctuations are quantified in the power spectrum of the CMB fit with spherical harmonic functions [36], shown in Fig. 2.2.

Evidence for dark matter comes from interpreting measurements of the CMB in the context of the standard cosmological model, ΛCDM (Λ-Cold Dark Matter) [37]. This paradigm assumes a Universe containing some component of Baryonic matter (with density parameter Ω_B), a non-relativistic (cold) non-baryonic dark matter component Ω_χ, radiation, and dark energy represented by the cosmological constant Λ. Fits of the CMB power spectrum assuming a particle dark matter component that does not experience radiation pressure (the relic density of which is one of six free parameters in the model) provide strong support for the cold dark matter paradigm. The successful application of this model to understand baryon acoustic oscillations and to predict



(a) Figure 6 from ref. [36].



(b) Figure 9 from ref. [36].

Figure 2.2: Recent results from the Planck experiment [36]. Temperature anisotropies in the foreground-subtracted sky map (top panel) are shown with respect to the average of 2.7 K, with the galactic plane, outlined in gray. The temperature angular power spectrum as a function of multipole moment l is compared to the best-fit ΛCDM model (bottom panel).

structure formation further reinforces this interpretation [39]. From the ΛCDM fit of the CMB, there is approximately 5 times more dark matter than baryonic matter in the Universe [40]. MOND-like theories are often incompatible with the CMB temperature spectrum or else would lead to structure formation not consistent with the Universe we live in [37, 41]. The requirement that dark matter exists in some form before recombination rules out all compact astrophysical objects (such as dwarf stars) except primordial black holes.

Despite the ΛCDM paradigm's apparent agreement with observations on cosmological and galactic scales, unresolved issues remain, especially on smaller scales. For example, the density profile of most galaxies does not seem to rise sharply at their centres as predicted for straightforward cold dark matter (the "core-cusp problem") [42]. There is also a conspicuously small number of dark matter sub-halos observed orbiting around or within our own Milky Way galaxy [42]. Tensions such as these may be alleviated within the ΛCDM paradigm, however, by including more second-order dark matter behaviours such as radiation pressure and interactions with SM matter into structure formation N-body simulations [43].

2.2 Particle Dark Matter

From the observations described in section 2.1, we can construct a list of properties that particle dark matter must satisfy [4]:

- Particle dark matter must be stable on cosmological timescales [4].
- Dark matter must not couple strongly to SM particles (particularly electromagnetically), and thus does not likely have an electric charge, but may interact with other particles through the weak nuclear force (and gravity) [40].
- The ΛCDM paradigm requires that a large component – if not all dark matter – is non-baryonic, i.e. that it does not experience radiation pressure. Further, this matter must be non-relativistic (cold) in order to be gravitationally bound to cosmological objects, and therefore participate in structure formation.

As there are no SM particles in compliance with all of these constraints, we are forced to appeal to new physics for suitable candidates.

2.2.1 Candidate particles

Weakly Interacting Massive Particles (WIMPs) represent a general class of candidates that follow this paradigm; relatively massive (10 - 1000 $\mathrm{GeV/c^2}$), stable particles that interact gravitationally and very weakly through other forces with the SM (note that "weak" does not imply that it interacts via the weak nuclear force specifically). This class of dark matter candidate has been historically favoured because of the "WIMP miracle", the seeming coincidence that the predicted freeze-out relic density of thermally-produced WIMPs – with a weak-scale mass and annihilation cross-section – is close to the observed relic density [44]. One theoretical extension to the SM that predicts a specific WIMP candidate is Supersymmetry, which calls for boson/fermion partner particles for all SM fermions/bosons respectively [44]. The

lightest supersymmetric particle – the neutralino – is a viable dark matter candidate [4, 40]. However, searches for evidence of Supersymmetry at the Large Hadron Collider have not revealed any sign of new physics at this time [45].

The lack of evidence for classic WIMPs thus far has led to increasing interest in other cold dark matter particle candidates [4, 8, 19, 46, 47], such as:

- Axions; meV/c^2 - scale neutral particles proposed to solve the Strong CP problem of the Standard Model, which could also serve as dark matter [40, 48].

- Kaluza-Klein states; higher-dimension excitations of lighter particles (such as axions) predicted by various quantum gravity theories [49–51].

- Sterile neutrinos; hypothesized additional mass eigenstates (possibly $\mathcal{O}$(keV/c^2)) of neutrinos which could explain the very low masses and left-handed chirality of known neutrinos. These heavy neutrinos would not interact with other matter directly, only when rarely oscillating to other neutrino flavours [52].

- Asymmetric dark matter; models that predict pairings of dark matter particles with $\sim 1\,\mathrm{GeV}$ and $\sim 5\,\mathrm{GeV}$ masses to balance the matter-antimatter asymmetry of the Universe [6, 53].

In general, there has been renewed experimental activity to search for low-mass (on the order of $1\,\mathrm{GeV/c^2}$ or less), WIMP-like dark matter of various theoretical motivations in recent years [5–8, 53]. It is also possible that dark matter includes a whole dark sector beyond the SM, of which we may be able to detect one or more constituents [54]. Despite the myriad theoretical models, experimental searches can exclude the existence of many different dark matter candidates by exploiting their similar masses and couplings, such as $\sim$ GeV $-$ TeV particles that couple to nucleons (like the classic WIMP).

2.2.2 Experimental observation

Assuming that there is some interaction between dark matter particles and the SM, all dark matter search experiments can be categorized as follows:

- **Collider production** If dark matter can interact with SM particles at all, then it should be possible to produce it during particle collisions of sufficient energy. If dark matter was produced in a collider such as the LHC, it would not be apparent directly in detectors like ATLAS or CMS, but the rapid decay of an unstable dark matter state or missing momentum/energy carried out of the detector by a dark matter particle could be observable. [4, 55].
- **Indirect detection** If particle dark matter can annihilate (either because it is a Majorana particle, or part of a dark sector of particles that can interact), its annihilation products may be visible in the form of x-rays/gamma rays, or high energy neutrinos. Experimental searches may involve neutrino observatories like IceCube [56] or space-based x-ray telescopes [57]. An x-ray signal from the galactic core has been detected, but the many astronomical foregrounds make it difficult to draw conclusions [57].
- **Direct detection** Particle dark matter may be able to scatter off of SM particles, through couplings with nucleons, nuclear spin, or electrons [58, 59]. A large, low-background particle detector may be able to detect the recoils of SM particles from such interactions. It is also possible that "dark photon" particles could be detected via a photoabsorption-like process [5], or for axions to be observed in "light shining through a wall" experiments (photons momentarily transforming into axions in the presence of a magnetic field) or with resonant cavity searches [47, 48].

The focus of the rest of this work is on direct detection experiments, searching for WIMP-like dark matter specifically. This sub-field itself involves dozens of international collaborations using many different detector technologies (more on these in

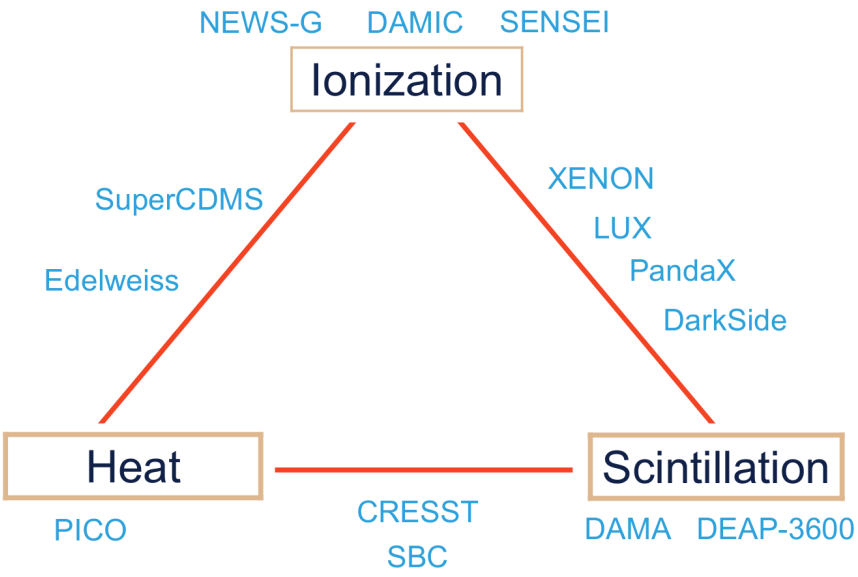


Figure 2.3: Direct detection experiments categorized by their energy observeables. Note that "heat" may refer to nucleations in bubble chambers, or phonon production in bolometers. Adapted from figure 4 of ref. [60].

section 2.4). These experiments can be further categorized by the observables that they measure (typically as estimates of particle interaction energy), which broadly speaking may be ionization, light (i.e. scintillation light), or heat. Fig. 2.3 shows a 3-lobed Venn diagram of many current dark matter direct detection experiments based on their respective observables. In particular, the NEWS-G (New Experiments with Spheres - Gas) experiment detects only ionization, while the Scintillating Bubble Chambers (SBC) detect both scintillation light and heat (in the form of bubble production).

2.3 Direct detection searches for WIMP-like Dark Matter

Generally speaking, all direct detection searches consist of some particle detector capable of observing WIMP-induced events. To have a chance of making an observation of these rare interactions, direct detection experiments go to great lengths to reduce background event rates by:

- using radio-pure construction materials,
- applying rigorous cleaning procedures to prevent contamination,
- shielding detectors from external radiation with water, lead or other shielding materials,
- placing experiments deep underground (in facilities such as SNOLAB [61] or the Laboratoire Souterrain de Modane [62]) to attenuate the flux of muons produced in the atmosphere by cosmic rays. These can produce background events directly in a detector, or through the production of spallation neutrons from interactions in nearby material.

To observe a statistically-significant signal, direct detection experiments also typically maximize their sensitivity to dark matter by having large target masses and/or taking data for long periods of time.

Another consideration for direct detection experiments is particle-species identification; that is, the ability to differentiate between interactions of lightly ionizing particles (electron or photon events), nuclear recoils, or highly-ionizing interactions (from alpha particles). This can greatly aid in suppressing non-nuclear recoil background events. Many detectors with multiple observables (see Fig. 2.3) are capable of this, exploiting the relative yields of different observables from different interaction types. For experiments like NEWS-G, it is also possible to discriminate against

track-like events (from lightly-ionizing background events) or to suppress background events coming from the impurity-laden surface of the detector with spatial reconstruction [63]. In the case of bubble chambers, lightly-ionizing particle events are naturally suppressed, and alpha particle recoils are identifiable acoustically [64, 65].

2.3.1 WIMP recoil energy spectrum

The WIMP nuclear recoil signal that would be hypothetically observed by a direct detection experiment can be calculated following the treatments presented in refs. [58, 66]. Recently, parameters associated with this paradigm were formalized by the PHYSTAT collaboration (see section 2.4) in ref. [67], which has been adopted by many collaborations.

The energy deposited by an elastic nuclear recoil interaction in the lab frame (E_R) by an incident dark matter particle with kinetic energy E_i and mass M_χ, assuming isotropic scattering in the centre-of-momentum frame, is given as [66]:

$$E_R = \frac{1}{2} r E_i (1 - \cos\theta) = \frac{1}{4} r M_\chi v^2 (1 - \cos\theta) \tag{2.1}$$

Here r is a kinematic factor similar to the reduced mass of the system:

$$r = \frac{4 M_\chi M_\mathrm{A}}{(M_\chi + M_\mathrm{A})^2} \tag{2.2}$$

where M_A is the mass of the target atom. This factor is maximized when $M_\chi = M_\mathrm{A}$, which means that direct detection experiments are generally most sensitive to (due to the relatively larger recoil energies produced by) dark matter with a similar mass to the target atoms employed.

To determine the differential rate of dark matter interactions, it is necessary to know the velocity distribution of such particles. This is given by the standard halo model of dark matter in the Milky Way, which assumes that our galaxy resides in the centre of a large, isothermal sphere of dark matter particles [66–68]. Given this, the distribution of dark matter particle velocities $\vec{v}$ on the earth is Maxwellian, given by:

$$f(\vec{v}, \vec{v}_{\rm E}) = e^{-(\vec{v}+\vec{v}_{\rm E})^2}/v_0^2 \times \Theta(v_{\rm esc} - ||\vec{v}||) \tag{2.3}$$

where $\vec{v}_{\rm E} = 232\,{\rm km/s}$ is the average of the time-varying velocity of the Earth relative to the galactic frame, and $\vec{v}_0 = 220\,{\rm km/s}$ is the local circular velocity of the dark matter [67, 68]. The Heaviside function serves to truncate the distribution at the escape velocity of the galaxy, $v_{\rm esc} = 544\,{\rm km/s}$ [69]. Next, we can write the differential number density of dark matter particles locally, expressed in terms of the local mass density $\rho_\chi = 0.3\,{\rm GeV/c^2cm^3}$ [67] :

$$dn = \frac{\rho_\chi}{kM_\chi} f(\vec{v}, \vec{v}_{\rm E}) d^3\vec{v} \tag{2.4}$$

where k is a normalization factor given in [58]. Using this, the differential event rate dR/dE_R is given as:

$$\frac{dR}{dE_R} = \frac{\rho_\chi}{kM_\chi M_{\rm A}} \int_{v_{\rm min}}^{v_{\rm esc}} \frac{d\sigma_{\chi-N}}{dE_R} v f(v) d^3\vec{v} \tag{2.5}$$

with $d\sigma_{\chi-N}/dE_R$ being the differential WIMP-nucleon cross-section. This equation integrates over all dark matter velocities that lead to a given recoil energy E_R, from the minimum such velocity $v_{\rm min} = \sqrt{2E_R/rM_\chi}$ up to the galactic escape velocity $v_{\rm esc}$.

The energy dependence of the differential cross-section can then be separated from a WIMP-nucleon cross-section $\sigma_{\chi-N}$ that is energy-independent [66]:

$$\frac{d\sigma_{\chi-N}}{dE_R} = \frac{\sigma_{\chi-N} F^2(q)}{4\mu_A^2 v^2} \tag{2.6}$$

where $\mu_A = (M_\chi M_A)/(M_\chi + M_A)$ is the reduced mass of the WIMP-nucleus system, and $F^2(q)$ is a structure form factor of the momentum transferred during a collision $q = \sqrt{2M_A E_R}$. The form factor is essentially the Fourier transform of the spatial density of the nucleons (and therefore of scattering sites) inside the nucleus. The standard choice for $F^2(q)$ for the spin-independent interactions is the Helm form

factor [58, 66], based on the Woods-Saxon potential as a choice of density distribution for the nucleus [70]:

$$F(q) = \left(\frac{3j_1(qR_N)}{qR_N}\right)^2 e^{-q^2s^2/2} \tag{2.7}$$

where j_1 is the first spherical Bessel function, $s \approx 0.9\,\text{fm}$ is the nuclear "skin thickness" [58], and R_N is the effective nuclear radius [58]:

$$R_N^2 = \left(1.23A^{1/2} - 0.6\,\text{fm}\right)^2 + \frac{7}{3}\left(0.52\pi\,\text{fm}\right)^2 - 5s^2 \tag{2.8}$$

In the case of lighter target atoms such as hydrogen, helium, carbon, and neon, $F(q)$ is very nearly equal to 1 (in both the spin-independent and dependent cases) up to interaction energies of $\approx 100\,\text{keV}$, far above the energy range of interest for experiments such as NEWS-G and SBC. Higher-A target atoms loose coherence at lower scattering energies, elevating the importance of form factor choices for experiments employing these targets, which also tend to have higher energy regions of interest [58].

The energy-independent cross-section of eq. 2.6 can be further divided into terms representing spin-independent (SI) coupling with nucleons $\sigma^{\text{SI}}_{\chi-N}$, and spin-dependent (SD) coupling $\sigma^{\text{SD}}_{\chi-N}$ [66]:

$$\sigma_{\chi-N} = \sigma^{\text{SI}}_{\chi-N} + \sigma^{\text{SD}}_{\chi-N} = \frac{4\mu_A^2}{\pi}(Zf_p + (A-Z)f_n)^2 \tag{2.9}$$

$$+ \frac{32G_F^2\mu_A^2}{\pi}\frac{J+1}{J}\left(a_p\langle S_p\rangle + a_n\langle S_n\rangle\right)^2 \tag{2.10}$$

where G_F is the Fermi constant, Z is the number of protons in the nucleus, and A the total number of nucleons. f_p/f_n and a_p/a_n represent effective couplings to proton and neutrons in the SI and SD cases respectively, all usually taken to be ≈ 1.

In the SD coupling term, J is the total nuclear spin of the nucleus, and $\langle S_p\rangle$ and $\langle S_n\rangle$ are the expectation values of the spin of protons and neutrons within the nucleus

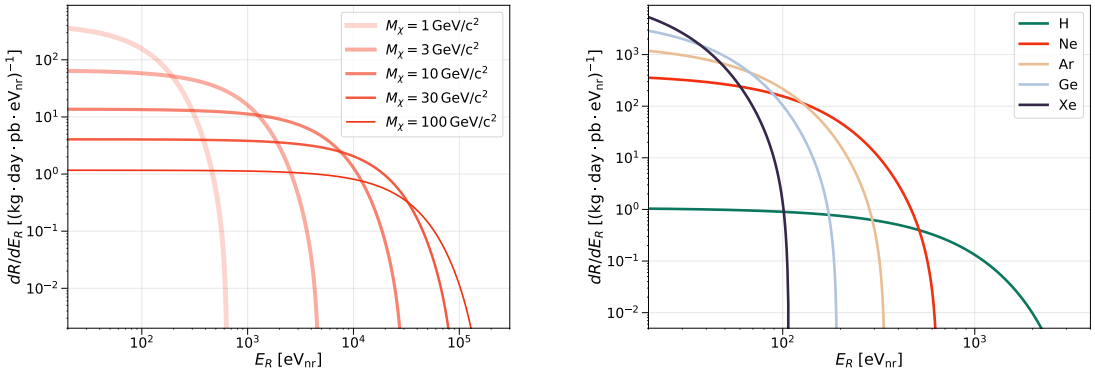


Figure 2.4: Examples of WIMP-nucleon recoil energy spectrum (assuming a 1 pb SI cross-section). Left: recoil spectra for different masses of WIMP-like dark matter scattering off of a neon target nuclei. Right: recoil spectra for a $1\,\mathrm{GeV/c^2}$ WIMP scattering off of various target nuclei.

respectively (which would be 1/2 for free nucleons) [66]. A table of values for $\langle S_p \rangle$ and $\langle S_n \rangle$ of different target atoms is given in Ref. [71]. Experiments using C_3F_8 as a target (e.g. the PICO experiment) benefit from having an odd proton in fluorine, giving a relatively high SD-proton coupling of $\langle S_p \rangle = 0.477$, but only $\langle S_n \rangle = -0.004$ [71].

Examples of WIMP recoil spectra (SI interactions only) are shown in Fig. 2.4, which show them to be approximately exponential but with a cut-off at a certain recoil energy. From this one can see that the recoil spectrum terminates at lower recoil energies for lower WIMP masses, and a given target atom. On the other hand, the spectrum extends to higher recoil energies for a given WIMP mass but lower-A target atoms. This highlights two fundamental realities of low-mass dark matter experiments in particular:

1. low-mass dark matter experiments need very low energy thresholds to be sensitive to candidate events, and

2. lighter target atoms increase sensitivity to low-mass dark matter.

The exponential nature of the recoil spectra also means that the majority of the

sensitivity of any direct detection experiment comes from low-energy recoils and that for low WIMP masses in general, the signal observed by an experiment comes only from recoils at/below $\mathcal{O}(1\,\mathrm{keV})$.

2.3.2 Calculation of exclusion limits

The statistical analysis of direct detection searches hinges on the fundamental question of whether or not the experiment detects dark matter-like events with sufficient confidence to claim a discovery, or to cautiously assume that all observed events are actually from background sources. In the latter case, an exclusion limit on the existence of dark matter can be calculated in the parameter space of dark matter cross-section vs. particle mass. The decision to claim discovery or publish an exclusion limit is somewhat subjective; modern dark matter experiments are often confronted with the difficult situation of observing events consistent with dark matter interactions that exceed the expected background rate, but a lack of confidence in background modelling and detector characterization often leads to exclusion limits being published instead. This is especially the case for current low-mass dark matter experiments (see section 2.4). In some such cases – with potentially compelling WIMP-like observations – experimental groups may publish both interpretations of their data, as a possible discovery and exclusion limit [72]. The dark matter analyses discussed in this work all lead to the publication of exclusion limits, and so a brief outline of their calculation is given below.

Exclusion limit calculations are performed for a given particle mass (and then repeated for a range of WIMP masses to define a continuous curve in the cross-section vs. mass parameter space). First, the WIMP recoil spectrum for the applicable WIMP mass and target atom is calculated, scaled by the exposure of the experiment (see eq. 2.5). In real analyses, this is then multiplied by the WIMP signal acceptance due to all data quality cuts, hardware and analysis trigger efficiencies etc. to obtain the recoil spectrum actually observable by the experiment. For a given cross-section σ_0,

the total expected number of events is the integral of this recoil spectrum.

In the simplest scenario possible, an experiment observes 0 events. In this case, the background-free exclusion limit can be calculated with the following steps:

1. Calculate the expected number of WIMP events N_0 (as described above) for an arbitrary reference cross-section σ_0:

$$N_0 = \int_{E_{\text{thresh.}}} \frac{dR}{dE_R}(\sigma_0) dE_R \tag{2.11}$$

2. Given the observed 0 events, the upper 90 % confidence limit on the expected number of events $N_{\text{excl.}}$ (the standard choice of the field) is calculated using

$$N_{\text{excl.}} = \frac{1}{2} F^{-1}_{\chi^2}(0.9, 2) \approx 2.303 \tag{2.12}$$

where $F^{-1}_{\chi^2}$ is the inverse cumulative distribution of the χ^2 distribution, and 0.9 corresponds to the desired confidence level [73].

3. The excluded cross-section can then be computed as:

$$\sigma_{\text{excl.}} = N_{\text{excl.}} \times \frac{\sigma_0}{N_0} \tag{2.13}$$

In the case where the experiment observes $N > 0$ events, the subsequent treatment depends on whether or not a subtraction of modelled background rates is performed. Without well-understood background models (which is the case for many low-mass dark matter experiments currently [74]), an exclusion limit may still be calculated by conservatively treating all observed events as WIMP candidates (which is admittedly somewhat counter-intuitive for an analysis claiming no detection of WIMPs). A simple "Poisson statistics" exclusion limit can be calculated in the same manner as a background-free experiment using the 90 % limit on the expected number of events given N observed [73]:

$$N_{\text{excl.}} = \frac{1}{2} F^{-1}_{\chi^2}(0.9, 2N + 2) \tag{2.14}$$

For example, if 1 (resp. 2, 3) events are observed, then the 90 % limit on the mean number of events is $\approx$ 3.890 (resp. 5.322, 6.681). The overall exclusion limit as a function of dark matter particle mass (see ahead to Figs. 2.5, 2.6 for example) typically have a "check-mark" shape when plotted on a log–log plot. The most strongly excluded cross-section is approximately at a WIMP mass equal to the target atom mass. At low masses the exclusion limit weakens asymptotically as the WIMP recoil spectrum cuts off at lower and lower energies, nearing and then reaching the limit of detection by the experiment (see Fig. 2.4). The limit rises at higher masses due to the $1/M_\chi$ dependence of the WIMP recoil spectrum (see eq. 2.5). For background-free experiments, the exclusion limit typically improves proportionally to increasing exposure time (i.e. a ten-fold increase in exposure will result in an excluded cross-section ten times lower). In the presence of background events, the excluded cross-section only improves as approximately $1/\sqrt{\Delta\text{exposure}}$.

More advanced statistical methods can produce stronger exclusion limits even if an experiment's background is not modelled, typically by defining targeted energy regions of interest (ROIs) with an improved signal-to-noise ratio compared to the whole energy range detectable by the experiment. One approach is to use machine learning algorithms such as Boosted Decision Trees (trained on simulated data) to find the ideal ROI for a given WIMP mass [63, 75, 76]. The Optimum Interval (OI) Method [77] is another approach that is similar in spirit, but which uses the real data of the experiment to define optimal ROIs constructed out of one or more consecutive intervals between the data points. Using the actual dataset of the experiment for this procedure incurs a "statistical penalty", calculated to ensure that obtained limits have the desired 90% coverage [77]. An example of a limit produced for a hypothetical neon experiment using the OI method is shown in Fig. 2.5. This example demonstrates how in some cases, the OI method can achieve an exclusion limit comparable to the background-free expectation of an experiment even in the presence of a background signal if there is some ROI with a favourable signal-to-noise ratio.

A third method that strikes a balance between trained machine learning algorithms and the OI method is an as-yet un-named technique employed by the Edelweiss and XENON1T experiments [78, 79]. In this approach, a fraction of the experimental data is "sacrificed" as training data, being used to manually find the optimal energy ROI for each WIMP mass. Then, the exclusion limit is calculated using the untouched remainder of the data, in the tuned ROIs. While this method does throw away some of the collected physics data (and thereby reduces the exposure of the experiment), this may be an acceptable loss for experiments faced with particularly high background rates (i.e. hundreds or thousands of observed events), for which the loss of some exposure likely will have a negligible impact on the resulting exclusion limit. Computation of OI limits for large datasets such as this is also often prohibitively slow, making this data-tuned ROI method an attractive alternative. Compared to typical machine-learning approaches, this method also has the advantage of not needing to rely on the accuracy of simulated data to tune ROIs. It is worth noting that all three of these limit-setting methods may be applied to any variable – not just energy-estimators as is usually the case – with possible extensions to multivariate analyses as well [63, 80].

The methods described above for calculating exclusion limits assume that all observed events are WIMP-like, including no information/contribution from background signals. If experimental backgrounds are known however (perhaps up to a normalization factor), then a "background subtraction" can be performed as a part of a likelihood fit of the data. In general, this involves a likelihood function for the observed data (often as a function of energy) given a WIMP signal with cross-section σ, and a set of nuisance parameters θ (often scaling parameters) describing the background contributions, $\mathcal{L}(\sigma, \theta)$. The likelihood function is then used to form a profile likelihood ratio $\lambda(\sigma)$ for the hypothesis that there is a WIMP signal with a specific cross-section σ vs. that there is not. "Profiling" refers to the process of optimizing the nuisance parameters (usually denoted with a double-hat) for a fixed value of the pa-

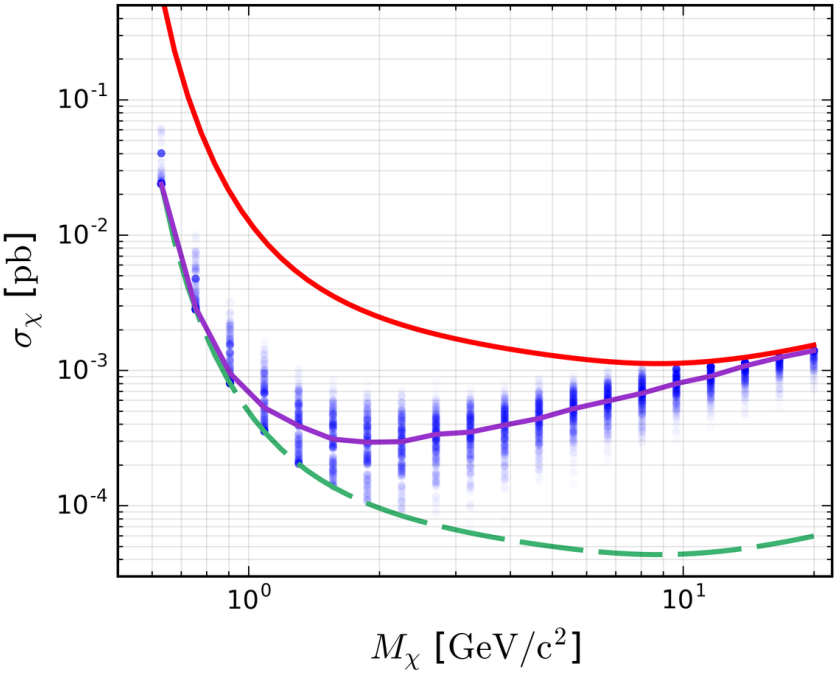


Figure 2.5: Demonstration of the OI method [77] applied to a hypothetical neon dark matter experiment, for SI coupling. This assumes an energy-independent background rate of 5/kg/day/keV$_{\rm nr}$, no ionization quenching factor or resolution effects, a 1 kg day exposure, and an ROI of 180 eV$_{\rm nr}$ to 1 keV$_{\rm nr}$. Shown are the results of 500 Monte Carlo trials of the OI method at each WIMP mass (blue markers) resulting in a median OI exclusion limit (purple curve), compared to the simple Poisson-statistics limit (red curve) and background-free limit (green dashed curve) for the same experiment.

rameter of interest, as opposed to simultaneously optimizing all parameters (denoted with single hats) [81]:

$$\lambda(\sigma) = \frac{\mathcal{L}\left(\sigma, \hat{\hat{\theta}}\right)}{\mathcal{L}\left(\hat{\sigma}, \hat{\theta}\right)} \tag{2.15}$$

This is then used to form the common test-statistic $q = -2\log\lambda(\sigma)$, which in ideal circumstances is known to follow the χ^2 distribution with 1 degree of freedom according to Wilk's theorem [82]. To calculate WIMP exclusion limits, this test-statistic can be modified as follows [81]:

$$q_\sigma = \begin{cases} -2\log\lambda(\sigma), & \hat{\sigma} \leq \sigma \\ 0, & \hat{\sigma} > \sigma \end{cases} \tag{2.16}$$

This definition inherently produces a one-sided confidence interval, considering only $\hat{\sigma} < \sigma$ to be at odds with the null hypothesis [67]. For this specific test-value of σ and the null hypothesis that this signal is present, the distribution of q is then calculated: $f(q_\sigma|H_\sigma)$. This can be done using Wilk's theorem, or more prudently with Monte Carlo simulations. The distribution $f(q_\sigma|H_\sigma)$ is then compared to the observed value $q_\sigma^{\text{obs.}}$, to calculate a p-value [83]:

$$p = \int_{q_\sigma^{\text{obs.}}}^{\infty} f(q_\sigma|H_\sigma)dq_\sigma \tag{2.17}$$

If this p-value is higher than the desired 0.1 (corresponding to a 90 % upper confidence limit), then the test value of σ is raised, or vice versa until σ is found such that the null hypothesis is rejected exactly with $p = 0.1$ to obtain the cross-section that can be excluded at 90 % [81, 83], all the while including optimized contributions from background signals.

A slightly modified, more robust version of the test statistic given in eq. 2.16 can be constructed to allow for two-sided confidence intervals to be formed [67]:

$$q_\sigma = \begin{cases} -2\log\lambda(\sigma), & \hat{\sigma} \geq 0 \\ -2\log\frac{\mathcal{L}\left(\sigma,\hat{\hat{\theta}}\right)}{\mathcal{L}\left(\sigma=0,\hat{\theta}\right)}, & \hat{\sigma} < 0 \end{cases} \tag{2.18}$$

In this case, $\hat{\sigma} > \sigma$ is allowed (corresponding to an excess dark matter signal). This construction also allows for negative values of σ (despite this being non-physical), but restricts their likelihood ratio to the physical lower bound of $\sigma = 0$. The advantage of this construction is that the same test statistic can be used for both exclusion limit setting, and discovery claims in the case that a non-zero WIMP signal is favoured by the likelihood ratio, at which point the one-sided limit automatically switches to a two-sided limit with a lower bound on the WIMP cross-section [67].

2.4 Status of the field

Direct detection searches have been ongoing for approximately 40 years to date, producing no conclusive/widely accepted evidence of particle dark matter despite the ever-increasing sensitivity of experiments [8, 9]. The field encompasses many different detector technologies (roughly categorized in Fig. 2.3); a non-exhaustive summary of the current status of experimental collaborations is given below.

Gaseous detectors Other than NEWS-G (the subject of chapters 3 and 4), gas-phase dark matter experiments include the CYGNUS/CYGNO, DRIFT, and TREX-DM projects, which typically employ proportional counter designs with MicroMegas-type devices for readout [84–87]. These finely-pixelated gas chamber anodes offer excellent spatial resolution, allowing for discrimination between point-like and track-like events on the basis of event reconstruction. Gaseous detectors also benefit from flexibility in target material choice, and the possibility of single-ionization energy thresholds (although this is not yet achievable by experiments other than NEWS-G).

Proportional counters such as these (sometimes referred to as gaseous time projection chambers) have been used to search for traditional low-mass WIMP-like dark matter nuclear recoils [87, 88]. However, ultra-low pressure proportional chambers are also being considered for directional detection of WIMPs; in rarefied target materials, the usually point-like interactions of nuclear recoils may be stretched into macroscopic tracks (i.e. of an observable scale). Typically, slow-drift gases such as helium mixtures of SF_6 are used to reduce diffusion of the particle tracks, aiding in reconstruction [85, 86]. The prevailing travel of the Earth in the direction of the constellation Cygnus means that WIMP recoils would originate primarily from that bearing in the sky, as opposed to isotropically distributed (i.e. detector radioactive backgrounds) or localized background sources (like neutrinos from the Sun) [85]. Directional detection would greatly boost the signal-to-noise ratio of any direct detection experiment and

is the primary hope for overcoming the "neutrino floor" background from coherent elastic neutrino-nucleus scattering [89–92].

Bubble Chambers Fluorocarbon bubble chamber experiments (i.e. PICO) have already produced world-leading spin-dependent WIMP exclusion limits [93], benefiting from intrinsic suppression of electronic recoil events. The first full-scale liquid noble bubble chamber for dark matter detector (by the SBC collaboration) – additionally sensitive to scintillation light, with possibly greater electronic recoil suppression – is currently in the fabrication stage [94]. A full discussion of bubble chambers as dark matter detectors is given in chapter 5.

Noble liquid noble detectors The search for traditional WIMP dark matter (with masses from 10 to 1000 $\mathrm{GeV/c^2}$) is currently dominated by liquid noble (cryogenic), dual phase time projection chambers (TPCs) sensitive to ionization and scintillation signals [9, 83]. The target liquid is held in a cylindrical vessel with an electric field applied along its axis. Particle interactions can produce scintillation light directly at the interaction site, detected by 2D arrays of photo-multiplier tubes (PMTs) on the ends of the cylinder. Additionally, ionization electrons may be produced in the liquid and pulled by the drift field toward a liquid/gas interface near the top of the detector, at which points the electrons are converted into more scintillation light. The 2D PMT arrays offer X-Y position reconstruction, while the delay between prompt (S1) and delayed (S2) scintillation light gives the Z position of an event. Fortunately, nuclear and electronic recoils produce different relative yields of prompt and delayed scintillation light, allowing for background discrimination [83].

The XENON [95], LZ [96], and PandaX [97] collaborations employ liquid xenon (LXe) dual-phase TPCs. The 5.5-tonne fiducial mass LZ experiment currently leads the field, with the lowest cross-section excluded to date; $9.2 \times 10^{-48}\,\mathrm{cm^2}$ at a WIMP mass of 36 $\mathrm{GeV/c^2}$ [96]. The 5.9-tonne XENONnT [98] and 3.7-tonne PandaX-4T

[97] devices have similar physics potential, and are expected to publish competitive results in the coming years. Despite their large fiducial masses, these experiments have also achieved the lowest background rates of all direct detection experiments, rendering them sensitive to coherent neutrino scattering in the coming years [99].

In contrast, the DarkSide-50 and DEAP-3600 collaborations use liquid argon (LAr) detectors [100, 101]. Although xenon is a higher-A target atom better suited to search for high-mass WIMPs, argon benefits from scintillation pulse shape differences between electronic and nuclear recoil events, with a discrimination factor against electronic recoils as high as 3×10^9 [102]. However, this pulse-shape discrimination is crucial for LAr detectors, which are afflicted by high activities of the beta emitter ^{39}Ar ($\sim$ Bq/kg). To mitigate this problem, future LAr experiments plan to use ^{39}Ar-depleted gas from underground reservoirs [103]. DarkSide currently operates a 46 kg dual-phase TPC at the Gran Sasso lab in Italy [100, 104], with plans to construct a 20-tonne detector [102]. DEAP-3600 is the sole example of a single phase noble liquid detector for dark matter detection, detecting only S1 (prompt) scintillating light [101]. The spherical geometry of the vessel allows for greater PMT coverage than typical dual-phase TPCs, at the expense of S1/S2 electronic recoil discrimination. However, discrimination against electronic recoils is still possible, exploiting the characteristically longer-lived triplet excimer state (compared to the singlet state) of LAr. This is disproportionately produced by electronic recoil events, particularly by the beta emitter ^{39}Ar, motivating the ratio of prompt to late scintillation light as a PSD variable [105]. Currently, a 3.3-tonne detector is operated at SNOLAB [106].

Both the LAr and LXe experimental collaborations are undergoing mergers (into the GADMC and XLZD groups respectively) due to the similar designs of their detectors, and the advantage of pooling resources to achieve larger combined fiducial masses of their costly target materials [103, 107].

Cryogenic solid-state detectors The SuperCDMS and Edelweiss collaborations [72, 108] employ germanium (and silicon in the former case) semiconductor detectors to search for dark matter. These consist of crystals (typically $\mathcal{O}(100\,\mathrm{g})$ sized pucks) cooled to $\mathcal{O}(\mathrm{mK})$ temperatures. Particle interactions may produce athermal phonons in the crystal, detected by transition edge sensors sensitive to the associated minute temperature fluctuations. Electron-hole pairs may also be produced, and drifted by an electric field to anodes/cathodes for readout on the surface of the crystal block [9]. As is often the case with two-channel detectors, the different relative energy partitioning of electronic and nuclear recoils into ionization and phonon production allows for background discrimination. Sensor arrays on the surface of the crystal can also give some X-Y position reconstruction, and the spread of the signal in the X-Y plane can be used to infer the Z-position, allowing for the rejection of surface background events [72]. These detectors can achieve energy thresholds as low as a few eV, benefiting from the low band-gap energies of these materials (2.9 eV and 3.6 eV in Ge and Si respectively [9]).

SuperCDMS has operated various Ge and Si crystal detectors with total masses on the order of 100s of grams (using an array of smaller crystals) [72, 109, 110], with ongoing efforts to construct a kg-array of detectors at SNOLAB [111]. The Edelweiss collaboration [108] has operated Ge cryogenic detectors at the Laboratoire Souterrain de Modane in France [62, 76]. By applying stronger drift fields, ionization may be converted into additional heat energy through the Neganov-Trofimov-Luke (NTL) effect [112], allowing for single quanta energy thresholds. Both collaborations have produced small, high-voltage devices designed to maximize this effect [78, 113, 114]. The SELENDIS project aims to tune the NTL effect (and exploit narrow energy resolution) to discriminate between electronic and nuclear recoils even for single-quanta events [115].

In addition to Si and Ge, polyatomic crystalline materials may be used for cryogenic detectors, such as calcium tungstate and sapphire. The former was chosen for

its high scintillation light yield, and inclusion of low-A target nuclei (oxygen) [116]. The CRESST collaboration operates their $LiAlO_2$ and $CaWO_4$ crystals as bolometers, collecting only the phonon/heat signal produced by particle interactions [117]. The low-A target materials (particularly the inclusion of lithium and oxygen) give CRESST excellent sensitivity to low-mass dark matter, both through SI and SD couplings [118]. The CRESST-III experiment has set the currently-strongest constraints on WIMP-like dark matter below masses of $1\,\mathrm{GeV/c^2}$ [117].

CCD experiments The DAMIC [119], SENSEI [120], and OSCURA collaborations [121] re-purpose traditional silicon CCDs into dark matter detectors, reading out the charge collected on each pixel over days-long exposures. These achieve the best spatial resolution of any dark matter detector, while also retaining single ionization energy thresholds, and sub-electron energy resolution (e.g. $\sigma = 0.07\,\mathrm{e^-}$ [122]). The Skipper-CCD readout technique of SENSEI further reduces readout noise [120]. With these capabilities, CCD experiments are sensitive to low-mass WIMP-like dark matter, as well as non-traditional dark matter candidates such as absorption of dark photons (requiring excellent energy resolution) [123]. The main drawback of this technology is scalability; the small silicon wafers used have masses on the order of grams, although there are plans to scale up to kg year exposures [121].

Crystal scintillators Sensitive only to scintillation light, these detectors comprise a scintillating crystal (usually NaI) instrumented with PMTs. This technology is relatively lower-cost at full scale, with $\mathcal{O}$(keV) energy thresholds. The most prominent example of such an experiment is DAMA/LIBRA, which has long claimed detection of an annually modulating signal consistent with dark matter nucleon scattering from a particle with a mass around 10 or $50-100\,\mathrm{GeV/c^2}$ [124]. Although such a signature is expected from particle dark matter due to the revolution of the Earth around the sun, this region of parameter space has been thoroughly excluded by other results, as

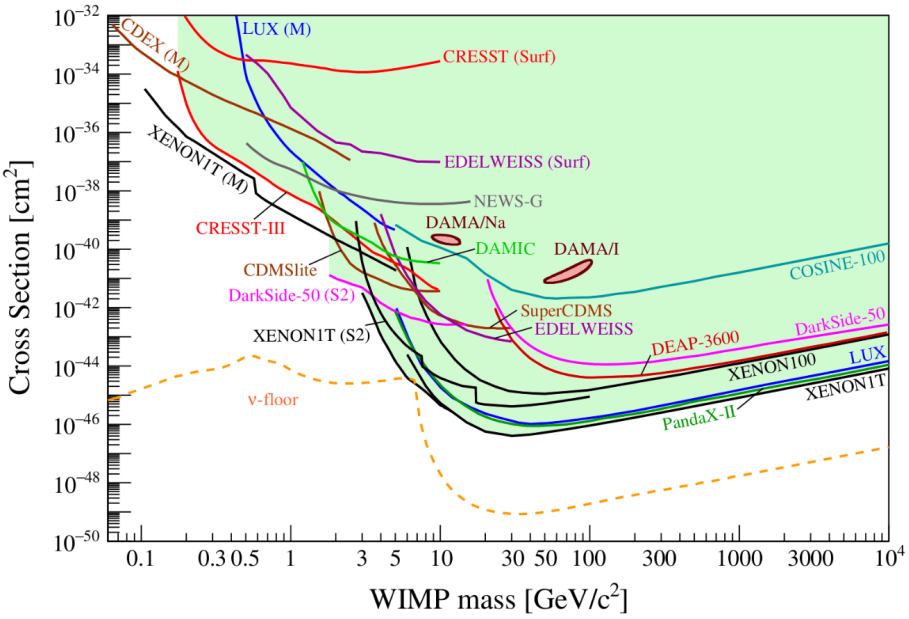


Figure 2.6: Spin-independent WIMP-nucleon interaction exclusion limits from many recent publications. The parameter space collectively ruled out is highlighted in green. A version of the "neutrino floor" is shown in orange [9]. Taken from figure 3 of ref. [9].

shown in Fig. 2.6.

There are dedicated experiments underway (i.e. COSINE-100, ANAIS-112) aiming to directly refute DAMA/LIBRA's claim with similar NaI detectors [125, 126], or to explain the undeniably strange signal observed. Thus far these experiments cannot strongly reject – but do not support – the existence of the modulation signal claimed by DAMA/LIBRA, due to their limited statistics [127, 128]. Postulated explanations for the modulation signal (other than dark matter) include diffusion of helium into the glass bulbs of the PMTs [129], or that a false modulation signal could be created by DAMA/LIBRA's practice of subtracting annually-averaged background rates [130, 131]. The latter hypothesis could be easily refuted/proven given access to DAMA/LIBRA's full dataset over time, something the collaboration has historically (and recently) refused to make public [132].

With this controversy brewing quietly within the field, dozens of experimental groups (many more than the sampling listed above) continue to publish increasingly

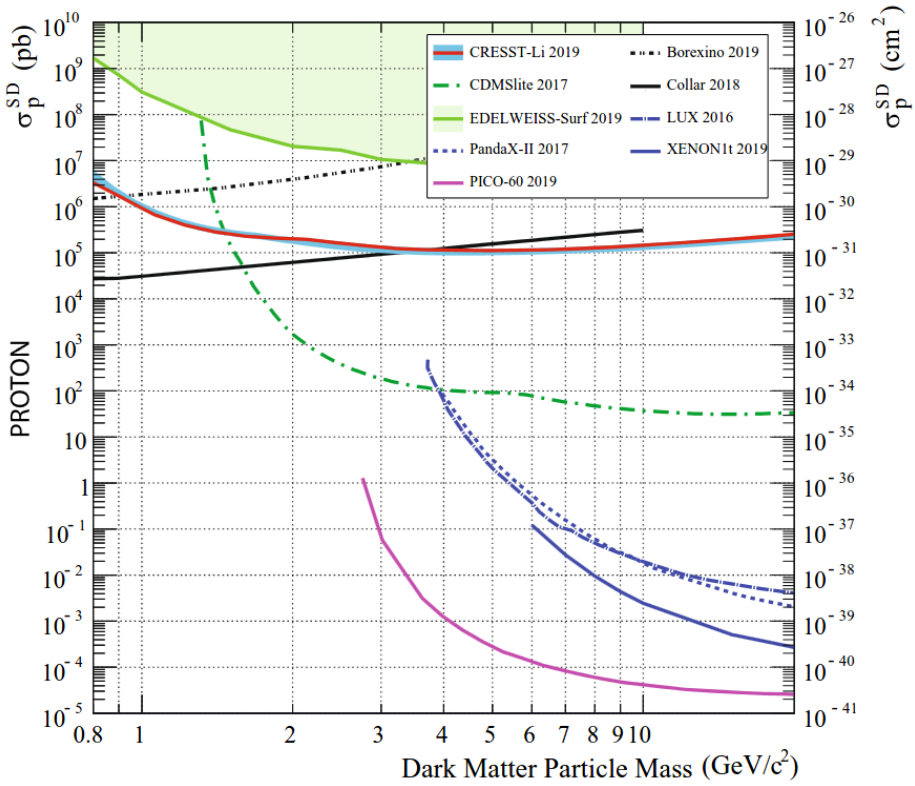


Figure 2.7: Spin-dependent (proton) WIMP-nucleon interaction exclusion limits from several recent publications. Taken from figure 4 of ref. [117].

stringent exclusion limits on the existence of WIMP-like dark matter. Examples of many currently leading exclusion limits are given in Figs. 2.6 (for SI coupling) and 2.7 (SD-p coupling). The search for classic WIMPs is currently led by LXe dual-phase TPCs. As mentioned in sec. 2.2.1, many experiments are now looking for low-mass dark matter particles that couple with nucleons [8]. The search in this parameter space (the primary targets of the NEWS-G and SBC experiments) is currently dominated by the CRESST-III experiment for spin-independent interactions [133], and PICO-60 for spin-dependent proton interactions [93].

Departing even more from the classic paradigm, it has been postulated that dark matter nucleon interactions may be enhanced by additional channels, namely the Migdal effect [134] and the nuclear Bremsstrahlung effect [7]. Both signals extend to higher recoil energies than elastic scattering for the same collisions, allowing existing

experiments to probe sub-GeV dark matter particle masses [79]. Neither effect has been demonstrated experimentally as of yet, although efforts to that end are ongoing [135, 136]. Additionally, there have been experimental results exploring (and excluding for now) dark matter coupling with electrons [59, 79], the existence of "dark photon" dark matter which would undergo a photoelectric-analogous process [78, 109, 113], and other non-traditional particle dark matter candidates [5].

2.4.1 Community efforts

There have been some hints of possible dark matter signals in recent years, or at least a growing number of experiments reporting data that is not fully accounted for by their background models. Perhaps not by chance, this has coincided with the rise of low-mass dark matter experiments, which may point to poorly-understood detector effects as an explanation. One example of an unexplained signal comes from the XENON experiment, which observed a strong excess of electronic recoils in the $2-3\,\mathrm{keV}$ range, which is consistent with solar axions as a dark matter particle [137]. Many other low-mass dark matter experiments using different technologies have also reported excess signals in recent years, including NEWS-G [63], SENSEI [123], DAMIC [138], SuperCDMS [113, 114], CRESST-III [133], and others [74]. While dark matter interpretations of these excess signals have been proposed [139], others contend that detector pathologies are to blame, such as stress-fractures in crystalline detectors [74]. Both interpretations struggle to reconcile the observations coming from many detector technologies. A multi-part special workshop (EXCESS 2021, EXCESS 2022) was convened to collaboratively address this controversial issue, resulting in a community white paper [74]. Although no firm conclusions have been reached, this workshop led to greater transparency and sharing of data from the collaborations involved.

Recognition of the challenges faced by low-threshold/low-mass dark matter experiments is also growing; modelling new detector physics, addressing unknown back-

grounds, and choosing appropriate statistical analysis protocols are all problems that need to be addressed. To that end, the PHYSTAT low-threshold working group was created to summarize the status of this field in a white paper (currently in preparation). While solutions do not exist yet for all of these questions, a thorough description of the problems and potential solutions may help guide future work; my involvement has included different statistical techniques to calculate dark matter exclusion limits including the optimal interval method [77], the use of machine learning techniques [63], or training-based approaches [78, 79]. This white paper also presents an opportunity to standardize certain modelling choices to describe the physics of low-threshold detectors, such as my previous work on modelling ionization statistics (the Fano factor) with the COM-Poisson distribution [140, 141]. This effort is a follow-up to the PHYSTAT-DM conference [142] and subsequent white paper, aimed at codifying statistical best-practices for traditional dark matter experiments. This set of standards has community endorsement from the NEWS-G, SBC, DAMIC, DarkSide, DEAP-3600, LZ, PandaX, PICO, SENSEI, SuperCDMS, and XENON collaborations [67].

Chapter 3

The NEWS-G experiment: physics and detector development

Gaseous particle detectors are among the earliest technologies developed for measuring ionizing radiation, dating back to the beginning of the 20^{th} century. Rutherford and Geiger created a cylindrical detector in 1908 with a single wire anode, with a tunable yield for incident radiation (by changing the voltage on the wire) [143, 144]. This device was the progenitor of both proportional counters and Geiger counters, which are still in prolific use today. Since then, gaseous particle detectors have found countless uses, and myriad forms to fit those needs. Prominent applications include radiation dosimetry (measuring the amount of energy imparted per unit mass) for oncological therapy [145], detection of high energy cosmic rays, particle identification at colliders, and even muon tomography experiments used to map the interior of the pyramids of Giza [146].

All gaseous detectors, regardless of their geometry, must include an anode structure to which voltage is applied to collect charges produced when incident particles ionize some of the gas in the detector. If the applied voltage is large enough, the incoming charges will be multiplied in a process called the Townsend avalanche, wherein electrons are accelerated to the point of further ionizing the gas, creating an exponential cascade [143]. Therefore, different detector configurations can be broadly categorized according to how the observed signal changes with anode voltage. The general trend

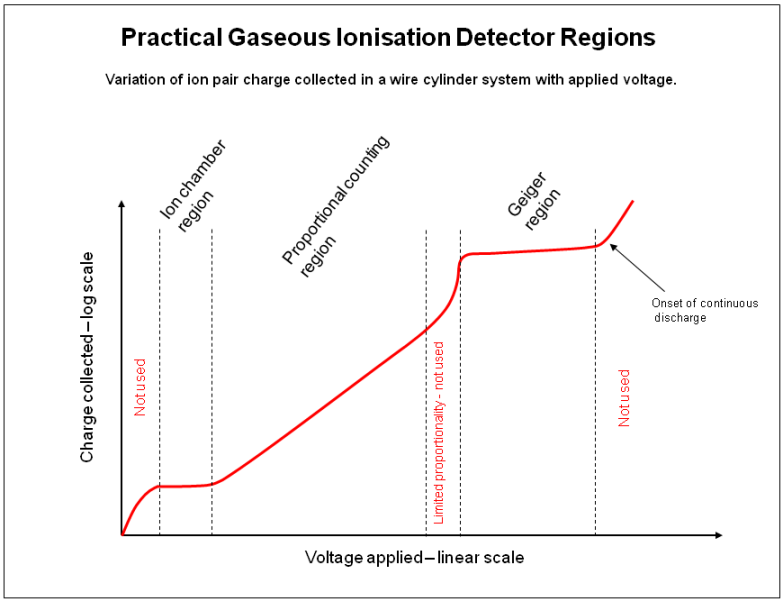


Figure 3.1: Ionization/charge response of gaseous detectors as a function of applied voltage, indicating the ion chamber, proportional counter, and Geiger counter regimes. Taken from ref. [147].

of this behaviour is shown in Fig. 3.1, highlighting the three regimes that are often used [143]:

1. Ion chamber region: at the lowest voltages, no avalanche amplification occurs, and the current measured on the anode is equal to the amount of charge produced. The chambers are often used in radiation dosimetry and can be made very compactly.

2. Proportional region: at intermediate voltages, a Townsend avalanche does occur, with a tunable yield that increases proportionally with increasing anode voltage. The current collected is also proportional to the amount of initial ionization (hence "proportional counter").

3. Geiger region: at the high voltages, a large Townsend avalanche always occurs (involving a large fraction of the gas medium available), leading to a large, clear signal that no longer depends on the energy of the incident radiation.

Therefore, detectors operating in this range can only function as counters (i.e. Geiger counters [148]).

Due to their tunable avalanche yield, and typical energy-linearity, proportional counters are most often used in physics research applications. A major milestone in their development was the Multiwire Proportional Counter (MWPC), a Nobel prize–winning achievement of George Charpak [143, 149]. Instead of a single wire anode, a grid of anode and cathode wires is used (often several stacked grids) to collect charges from multiple wires. With separate readout for each anode wire, the planar position of a particle interaction may be deduced [149]. This can be taken a step further by using the timing information from different anode wires in a 3D grid to reconstruct the perpendicular trajectory of the incident particle. This attribute of excellent spatial reconstruction was carried forward by MicroMegas and GEM style proportional chambers with sub–mm spaced anodes, which are used for some dark matter direct detection experiments currently (see section 2.4) among many other applications.

Spherical Proportional Counters (SPCs) are another example of gaseous detectors, originating in 2008 for the purpose of rare event detection (i.e. dark matter, neutrino searches) and neutron spectroscopy [150, 151]. Consisting of a single spherical anode inside a spherical vessel, these detectors are designed to achieve high avalanche amplification while still having a linear energy response, low electronic noise (due to the low capacitance of this spherical geometry), with a high volume to surface area ratio (where most radioactive contaminants reside). The target gases used are usually noble gases with small admixtures of methane, operated at pressures ranging from $\approx 50\,\text{mbar}$ up to several bars. The ability to use light atomic gases such as helium and neon (and hydrogen from methane), as well as the low energy threshold made possible by the large avalanche yield, make SPCs particularly well suited to search for low-mass dark matter. Examples of an SPC vessel and the anode within are shown

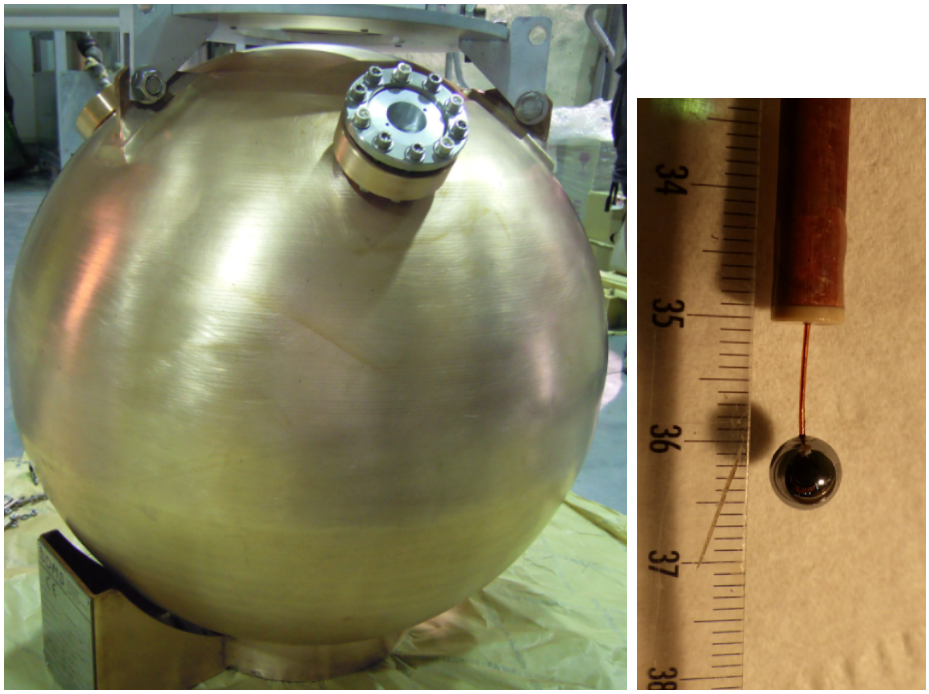


Figure 3.2: The SEDINE detector (a 60 cm copper SPC), with a 6.3 mm anode on a grounded rod [63]. Taken from figure 1 of ref. [63] (left) and figure 3.22 from ref. [152] (right).

in Fig. 3.2.

The rest of this chapter details the use of SPCs by the NEWS-G collaboration to search for light dark matter, beginning with an overview of the physics of the experiment and past results. Calibration techniques are described in section 3.2, which are then applied in measuring the ionization yield of SPCs with methane gas in section 3.3. A description of the device currently operated by the collaboration is given in section 4.1, after which the analysis and physics results of a dark matter search with pure methane gas are presented in section 4.2-4.5. The design of a small-scale prototype SPC to develop radon trapping filters is discussed in section 3.4.

3.1 SPC Physics

The first step in understanding the physics of SPCs is to examine the electric field structure of the detector volume. A first-order approximation may be obtained by assuming that the anode is a sphere floating in the middle of the grounded SPC vessel, with no support rod. In this simple case, the electric field inside an SPC of radius r_1 with an anode of radius r_2 is given by [153]:

$$E(r) = \frac{V_0}{r^2}\left(\frac{1}{r_2} - \frac{1}{r_1}\right)^{-1} \tag{3.1}$$

where V_0 is the voltage applied to the anode. In reality, the grounded support rod significantly distorts the electric field in volume near it, necessitating a finite element simulation of the electric field. An example of such a simulation – performed with the COMSOL software [154] – is given in Fig. 3.3 for a 60 cm SPC, with an applied voltage of 2520 V [63]. These simulations also require detailed modelling of the sensor structure; typically these are either silicon or stainless steel balls attached to insulated HV wires, fed through grounded metal rods. At the end of the rod, an "umbrella" structure may be attached. These can include a cylinder of resistive material such as Bakelite or glass, possibly containing a secondary electrode designed to make the electric field more isotropic throughout the SPC volume [155, 156].

Another significant consideration for SPCs is the gas mixture used. Typically, noble gases are chosen as they are monoatomic (simplifying the physics of ionization), chemically inert, and have relatively low ionization thresholds. For NEWS-G, this includes helium, neon, and argon gas. It is also possible to operate SPCs with xenon gas, as is done by the R2D2 collaboration to search for neutrinoless double β^- decay [157]. In addition to the main gas constituent, a small admixture of a molecular "quenching gas" is added [143], often chosen to be methane for NEWS-G experiments. This gas (usually in concentrations of $\sim$ 1 to 7% by volume) serves to suppress run-away ionization caused by the recombination of ions produced in the

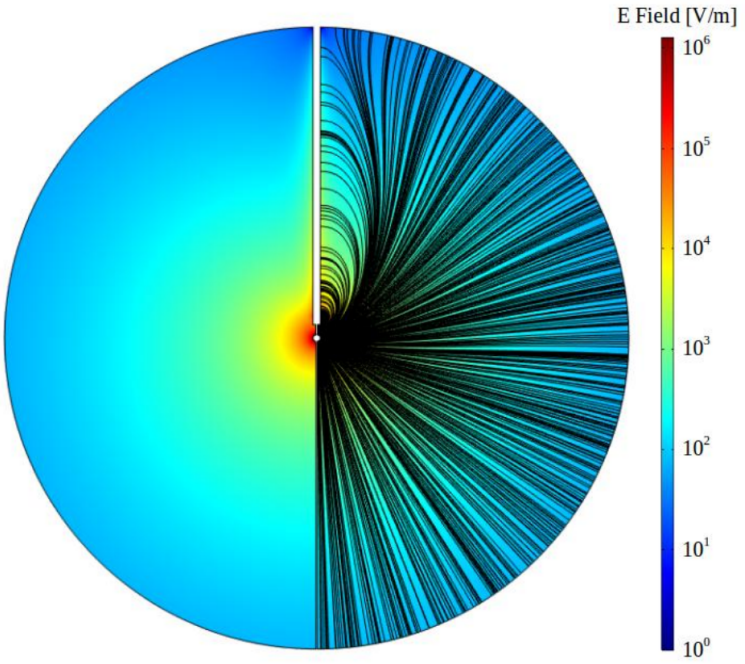


Figure 3.3: The electric potential (colour-scale) and electric field lines (right-hand side, black curves) of a 60 cm diameter SPC with a 6.3 mm anode at the centre, with an applied voltage of 2520 V. This finite element simulation was performed with the software COMSOL [154]. Taken from figure 1 of ref. [63].

Townsend avalanches of past interactions. Molecular gases achieve this by introducing additional, non-ionizing channels of energy deposition, i.e. through rotation and vibration of the molecule [143]. The addition of such gases can increase the gain of the experiment through the Penning effect [158], by which excited noble gas atoms may collide with a gas molecule with a lower ionization threshold, increasing the yield of charges [159]. The strength of this effect depends on the composition and pressure of the gas mixture.

The general principle of detection for SPCs is given below, corresponding to steps 1–5 in Fig. 3.4:

1. When ionizing radiation (be it dark matter or any type of background radiation) interacts in the gas volume of the SPC, a stochastic number of primary electron/ion pairs will be produced. For electronic recoil events (from photons/electrons) this involves a sequence of photon scatters, photo-ionizations, and

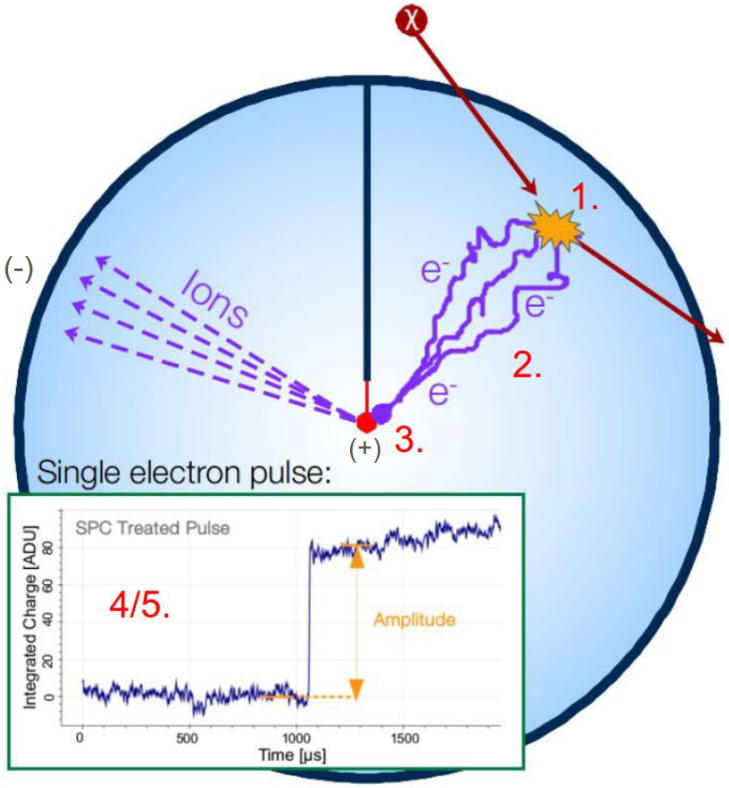


Figure 3.4: Cartoon depiction of an SPC, showing the steps of particle detection from the primary ionization produced by the incident radiation (step 1) up to signal formation and treatment (step 4/5).

ionizations by electrons. For neutral particles, the recoiling nucleus after a collision ionizes the gas. The average number of primary pairs for a given deposited energy is calculated using the W-value, and their dispersion is dictated by the Fano factor, both of which are described in more detail in sections 3.1.1 and 3.1.2.

2. The HV anode attracts the primary electrons produced in the gas, which diffuse as they travel through repeated elastic collisions with other gas atoms/-molecules. The drift of electrons can take anywhere from 10s of microseconds to more than 1 millisecond depending on the gas, size of the SPC, anode voltage, and starting position of the electrons. More details on this process are given in section 3.1.3.

3. When reaching within several millimeters of the sensor (again, this depends on the geometry, gas and voltage of the SPC), primary electrons are accelerated sufficiently to start a Townsend avalanche [143], producing 1000s of secondary electron/ion pairs for every primary electron. These electrons travel the final distance to the sensor almost instantaneously, while the ions can take as many as several seconds to reach the SPC wall (the cathode). Empirical modelling of this process is described in section 3.1.4.

4. At this point, the large number of secondary ions drifting away from the anode region (rapidly at first, but quickly decelerating) induce a current on the anode that is large enough to be observed. The current induced by q_{ion} ions as a function of time is given as [153]:

$$Q_{\text{ind}}(t) = -q_{\text{ion}}\rho\left(\frac{1}{r_2} - \frac{1}{\left(r_2^3 + 3\alpha t\right)^{1/3}}\right) \tag{3.2}$$

where $\alpha = \mu^0_{\text{ion}}\frac{V_0}{P} \times \left(\frac{1}{r_2} - \frac{1}{r_1}\right)^{-1}$, and μ^0_{ion} is the mobility of the gas ions at atmospheric pressure (typically $\mathcal{O}\left(10^{-6}\right)$ $\text{cm}^2\text{V}^{-1}\mu\text{s}^{-1}$ [160]). Examples of the signal

shape induced by the ions are shown in Fig. 3.5, for various gas/voltage assumptions. At this stage, the signal passes through a charge-sensitive preamplifier with an exponential response function (the manufacturer-given decay time is usually $\tau \approx 100\,\mu$s). This integrates the induced current given by eq. 3.2 – and multiplied by the preamplifier response function – to give the following voltage signal over time [153]:

$$S(t) = -q_{\text{ion}}\alpha\rho e^{-t/\tau} \int_0^t e^{t'/\tau} \left(r_2^3 + 3\alpha t'\right)^{4/3} dt' \tag{3.3}$$

This voltage signal is then digitized by a DAQ (data acquisition) board which also applies a trigger algorithm to store sections of the continuous signal trace containing the pulses caused by particle interactions.

5. The raw signal recorded by the digitizer is treated (as described in section 3.1.5) to produce an output such as the example shown in Fig. 3.4 step 5, which was produced by a single primary electron in the gas.

The DAQ boards used by NEWS-G are either custom FPGA-based devices nicknamed "caliboxes" [63, 161], or commercially available Red Pitaya FPGA boards [162]. In either case, the data is recorded at a sampling frequency of $1 - 2\,$MHz, with each event consisting of a $1 - 8\,$ms long signal trace. The DAQ software is a custom application called "SAMBA" [161], previously employed (along with the calibox) by experiments such as Edelweiss and CUPID-Mo [163, 164]. When the DAQ algorithm is triggered, Samba records an event window centred around the maximum of the observed raw pulse.

3.1.1 Primary ionization

The stochastic process of primary ionization is traditionally characterized by two quantities – the W-value and the Fano factor. The former is an energy-dependent, gas-specific quantity that represents the average energy required to ionize a single gas

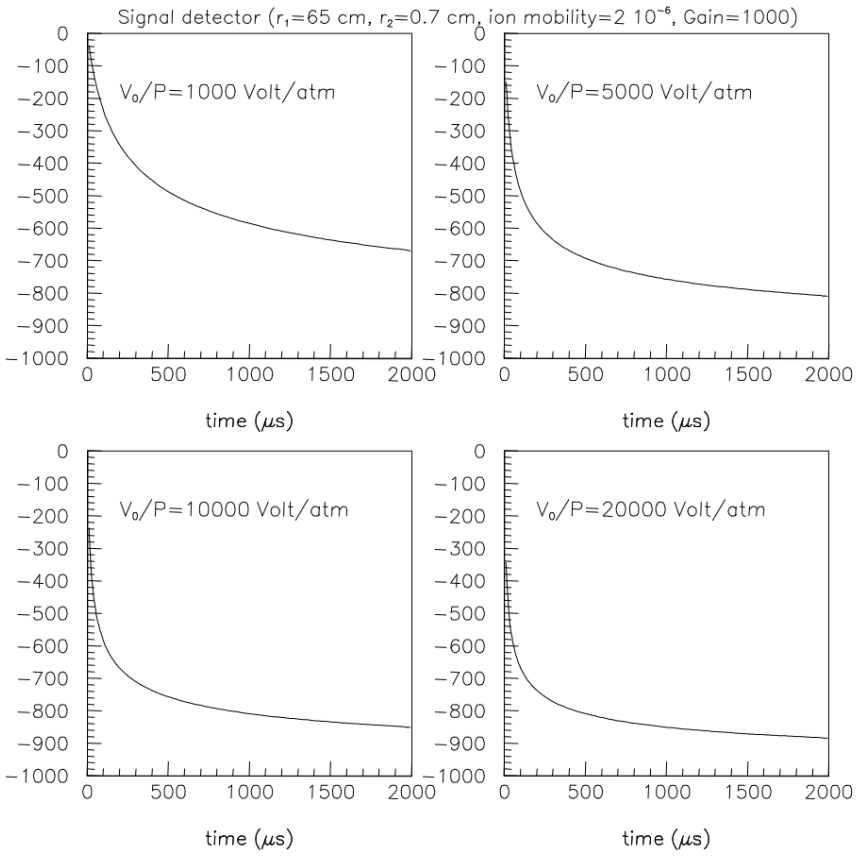


Figure 3.5: The ion-induced current (according to eq. 3.2) in a 65 cm radius SPC, with a 7 mm radius anode, with typical values of the gain and ion mobility, under various voltage/gas pressure conditions [153]. Credit to J. Derré.

atom. The parameterization used by NEWS-G to describe its evolution with energy is [165]:

$$W(E) = W_0 \times \frac{E}{E - U} \tag{3.4}$$

where W_0 is the high energy limit to which $W(E)$ converges (often quoted by itself as the W-value), with values typically around $W_0 = 30\,\text{eV}$ and $U\mathcal{O}(10\,\text{eV})$. U is meant to roughly equal the average energy of sub-ionization excitation levels, but in practice is an empirically derived value [165]. This model has been shown to be in good agreement with measurements in various proportional counter gases, including for sub-keV energies [166, 167]. Efforts to measure this quantity for NEWS-G operating conditions are presented throughout this chapter. The average number of primary ionizations μ produced by a particle interaction depositing energy E is therefore:

$$\mu = \frac{E}{W(E)} \tag{3.5}$$

The Fano factor gives the index of dispersion of primary ionization [168], that is:

$$F = \frac{\sigma_N^2}{\mu} \tag{3.6}$$

where σ is its standard deviation. While a Poisson process would have $F = 1$, it is known that for gaseous detectors F is typically around 0.2. Indeed the Fano factor has been measured to be $\lesssim 0.2$ for a variety of ionization detector media including argon, xenon, germanium and silicon [169–172].

However, representation of the entire probability distribution of primary ionization (rather than just the Fano factor as a descriptive quantity) remains something of an unresolved issue in low-energy particle physics. While this can be simulated from first principles using electron scattering cross-section data [173], this approach is impractical for modelling an experiment's energy response over the entire range of its sensitivity and is known to produce results that do not agree with experiments

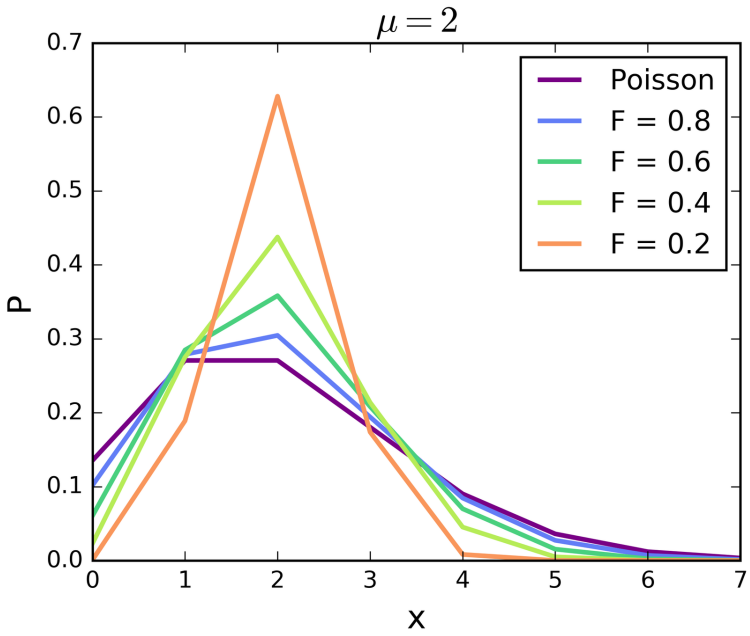


Figure 3.6: Examples of the COM-Poisson probability distribution with a mean of 2 and various Fano factors.

[140]. Therefore, the COM-Poisson distribution was adapted for use as a flexible statistical model of primary ionization in refs. [90, 140]. This probability distribution – originally developed for modelling queuing behaviour [141, 174] – is a discrete probability distribution with two shape parameters [175]:

$$P(X = x|\lambda, \nu) = \frac{\lambda^x}{(x!)^\nu Z(\lambda, \nu)}, \qquad \text{for } x \in \mathbb{N}_0,\ \lambda > 0,\ \nu \geq 0 \tag{3.7}$$

where Z is a normalization factor:

$$Z(\lambda, \nu) = \sum_{s=0}^{\infty} \frac{\lambda^s}{(s!)^\nu} \tag{3.8}$$

While λ largely controls the mean of the distribution, ν mostly changes its dispersion (in the special case $\nu = 1$, it reduces to the regular Poisson distribution). With the ability to represent ionization distributions with any mean and any F (that is mathematically possible [140]), the COM-Poisson PDF is a flexible choice that allows NEWS-G to fit and model data with $F < 1$. Examples of this distribution are shown in Fig. 3.6.

The difficulty in using the COM-Poisson distribution comes from the fact that λ and ν do not correspond directly to μ and F, and no closed-form expression relating them exists. Thus, ref. [140] includes the calculation of lookup tables for λ and ν to give the desired values of μ and F, as well as asymptotic approximations that are valid for $\mu \gtrsim 20$. The strategies used to employ the COM-Poisson distribution are shown in Fig. 3.7, which correspond to the publicly available COM-Poisson code used by NEWS-G [176].

3.1.2 Ionization quenching factor

The primary ionization yield of nuclear recoil events is reduced compared to electronic recoil interactions, due to the predilection for nuclear recoils to deposit energy into non-ionization channels, specifically heat and radiation from excited states [177]. This effect is described by the so-called "quenching factor" Q, a scaling factor on the W-value for electronic recoil events W_{ee} (between 0 and 1) that gives the primary ionization yield for nuclear recoil events using eq. 3.5):

$$W_{\mathrm{nr}} = \frac{W_{\mathrm{ee}}}{Q(E)} \tag{3.9}$$

The quenching factor is an energy-dependent, gas-dependent quantity. The main model for predicting the quenching factor is due to Lindhard [177]. For some ionization detector media (i.e. liquid argon and xenon), Lindhard theory has been shown to be in rough agreement with experimental results [178, 179], but not in other instances including measurements in various gas mixtures [180–182].

Another approach to predicting this quantity uses software such as SRIM [183], which simulates ions slowing down in matter. Previous NEWS-G experiments used SRIM to calculate the quenching factor for neon [63], which was parameterized as

$$Q(E) = \alpha \times E^{\beta} \tag{3.10}$$

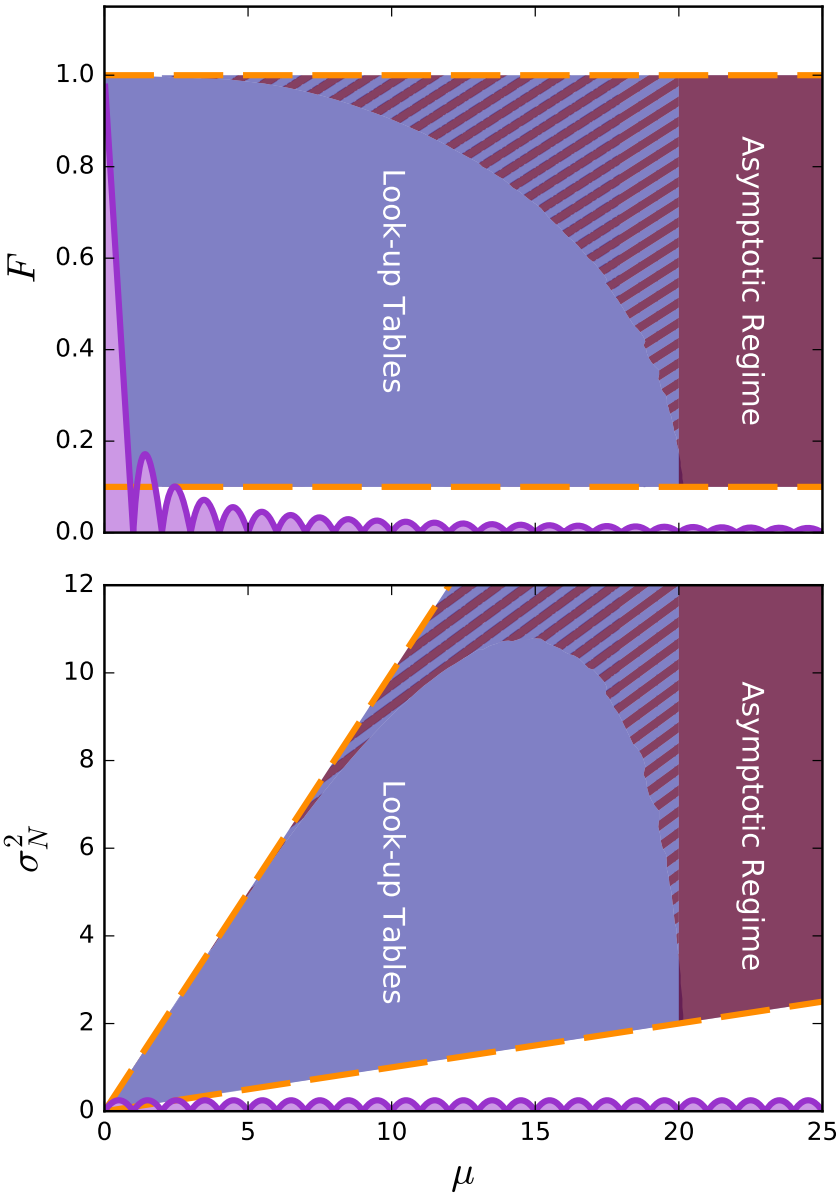


Figure 3.7: Strategies for using the COM-Poisson distribution for ionization modelling in different regions of μ/F (upper panel) and μ/σ_N^2 (lower panel) parameter space. This includes the use of pre-computed lookup tables (blue-shaded region), and asymptotic approximations (red-shaded region), bounded by the minimum possible dispersion as a function of μ (purple-shaded region). In the implementation of the code used by NEWS-G, these strategies are verified to be accurate to within 0.1% from $F = 0.1$ to $F = 1$ (orange dashed lines). Taken from figure 1 of ref. [140].

with $\alpha = 0.216$ and $\beta = 0.163$. However, this approach using SRIM is also not thoroughly tested against measurements. Therefore, it is critical to obtain empirical results for this SPC property. Quenching factor measurements are notoriously hard to perform, in part because a source of nuclear recoil events is needed (e.g. from neutrons), but also because of the many other correlated factors associated with a detector's energy response, such as the W-value and avalanche yield.

Despite these difficulties, the NEWS-G collaboration has performed measurements in neon/methane gas mixtures using a neutron beam produced at the TUNL facility [184]. In this case, the empirical results were found to disagree with both the SRIM prediction and Lindhard theory, motivating future measurements. A semi-empirical treatment for the quenching factor in methane is described in section 4.3.4.

3.1.3 Electron transport

Once primary electrons are created in the gas, they begin the process of drifting toward the central anode, and are subject to diffusion along the way. In the absence of an electric field, the thermal motion of the electrons would lead them to diffuse outwards (with time-evolving spatial density ρ) through repeated elastic collisions with other gas atoms according to Fick's law [143, 185]:

$$\frac{d\rho}{dt} = D\nabla^2\rho \tag{3.11}$$

where D is a coefficient of diffusion, with dimension $\mathrm{m}^2\,\mathrm{s}^{-1}$. Solving this equation gives an essentially Gaussian distribution with $\sigma = \sqrt{2Dt}$ for the spatial density over time [185]:

$$\rho(r,t) = \left(\frac{\rho_0}{4D\pi t}\right)^{3/2} e^{-r^2/4Dt} \tag{3.12}$$

Determining the impact of the electric field on this diffusion is non-trivial. The CERN simulation package Magboltz gives the coefficient of diffusion of electrons in gas (with the pressure and composition specified) as a function of electric field

strength [186]. The diffusion coefficient is different for parallel and perpendicular motion relative to the electric field. Using this information, a custom MC is used to model the diffusion and drift time of electrons starting at different positions in the SPC. Note that analogous simulations have been performed by other members of the NEWS-G collaboration (i.e. refs. [152, 160, 187]). In the version developed for the work shown in this document, a simulated electron is propagated through the following steps:

1. At the initial position of the electron, the electric field strength is determined (either using the idealized electric field given in eq. 3.1 or using a finite element simulation of the SPC).

2. Parallel and perpendicular diffusion coefficients – as well as electron drift velocity at this electric field strength – are obtained from Magboltz [186].

3. The electron is stepped towards the anode of the SPC according to its drift velocity, for a fixed time step ($\mathcal{O}(\mu s)$).

4. In addition to this drift step, random parallel and perpendicular diffusion displacements are applied to the electron (drawn from a Gaussian with $\sigma = \sqrt{2D_{\parallel}t}$ or $\sigma = \sqrt{2D_{\perp}t}$ respectively). The parallel diffusion moves the electron forward or backward along its drift trajectory, while perpendicular diffusion moves it along a spherical shell at the same radius, i.e. while maintaining the same distance from the anode along the electric field line the electron follows. Therefore, if the idealized electric field of eq. 3.1 is used, perpendicular diffusion is irrelevant.

5. Steps 1 to 4 are repeated until the electron reaches within a preset distance of the anode.

Note that in this MC scheme with fixed time intervals, the last few centimetres before reaching the anode may occur in only a few MC steps (as the electron rapidly

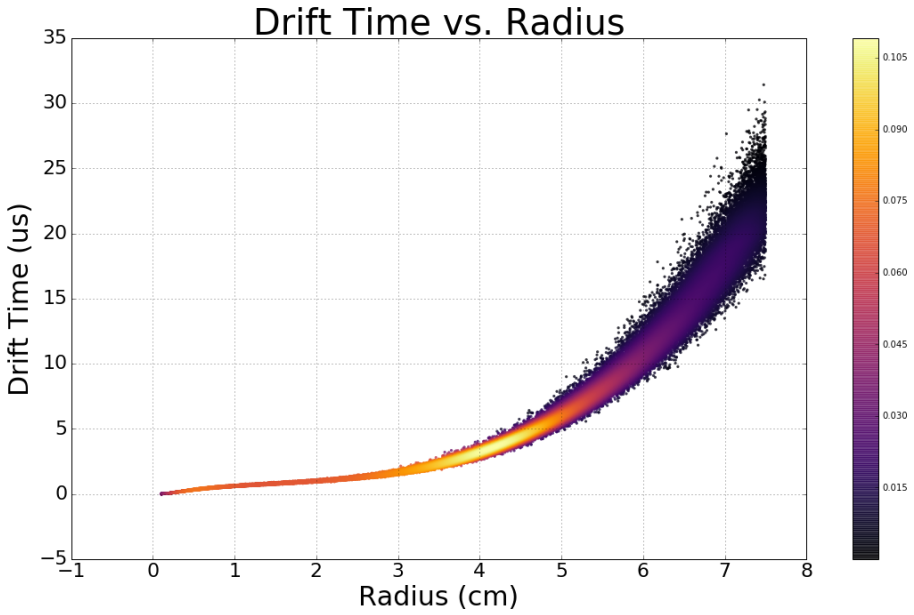


Figure 3.8: Electron drift time vs. initial radius simulated for a 15 cm diameter SPC, with 800 mbar, Ar + 2% CH_4 gas. The spread in drift time represents the amount of diffusion experienced by electrons. The colour scale is a 2D Gaussian KDE score (arbitrary units) representing the density of points on the plot. Figure 3.5 from ref. [90].

accelerates). This coarse stepping in the region is not problematic, however, as the electron spends the majority of its drift time (and experiences the most diffusion) far from the sensor, in low electric-field regions.

An example of the results of such a simulation – for many electrons at many different starting radii in an SPC – is shown in Fig. 3.8. The spread in electron arrival times, due to stochastic diffusion, leads to varying pulse shapes depending on how much diffusion is experienced. Generally, diffusion is greater for events starting at larger radii, allowing for limited pulse-shape discrimination between surface (large initial radii) and volume events in an SPC. Several variables designed to characterize this behaviour are described in section 3.1.5.

The other process to consider while primary electrons drift is charge trapping, or "attachment". This occurs when free electrons drifting through the gas collide with electronegative gas particles (primarily molecular oxygen and water), which trap the

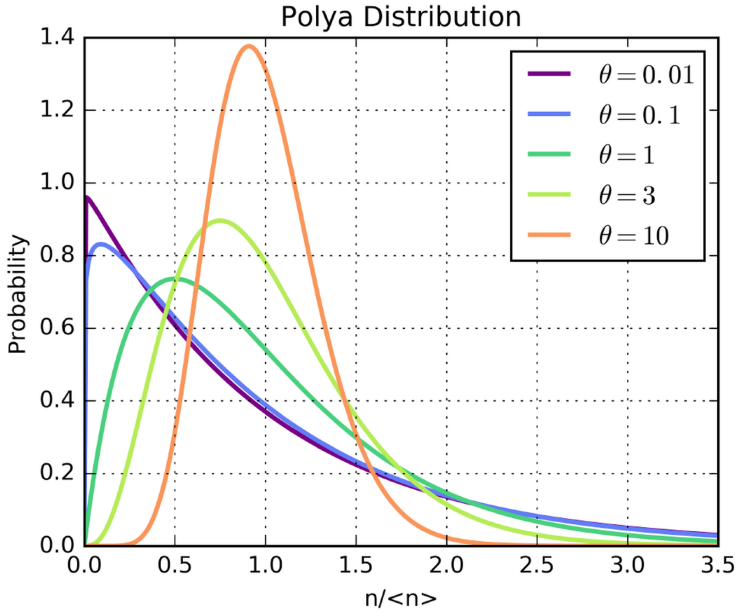


Figure 3.9: The Polya probability distribution with different values of θ, plotted in units of the mean gain in number of secondary electrons $\langle n \rangle$ such that the mean of the distribution is always 1.

free electrons [63, 152]. In this way, charge trapping reduces the signal observed from particle interactions, with events starting further from the sensor more likely to be affected (due to their greater chances of encountering contaminants). This can produce a noticeable impact with oxygen and water contamination as low as $\mathcal{O}$(ppm), necessitating great efforts to keep these contaminants out of SPCs (more on the apparatus involved in section 4.1).

3.1.4 Townsend avalanche

If a primary electron has sufficient kinetic energy, it can ionize more gas atoms/-molecules in the medium as it drifts toward the sensor. Typically this only happens when the electric field is strong enough to accelerate electrons to this degree in between elastic collisions. If the mean free path length between ionizing interactions is λ, then the inverse quantity $\alpha = 1/\lambda$ – the first Townsend coefficient – gives the average number of secondary ionizations produced per unit length [143]. This quantity

depends on the gas composition and pressure – dictating the ionization cross section and number density of atoms – and electric field strength, and can be determined with the Magboltz software [186]. Each of the secondary electrons produced will go on to ionize more atoms, creating an exponentially growing cascade (the Townsend avalanche) [143]. The number of secondary electrons n produced by a single primary electron grows with the travel distance x of the primary electron as [152]:

$$n = e^{\alpha x} \tag{3.13}$$

In an SPC, this occurs within a millimetre (or less) of the anode, with typically thousands of secondary pairs produced for every primary electron. The mean gain of the detector is therefore defined as the yield of secondary electrons n produced by N primary electrons:

$$\langle G \rangle = \frac{n}{N} \tag{3.14}$$

As with primary ionization, this is also a stochastic process. For large numbers of primary electrons, the avalanche yield E (the observed energy, various units) will be necessarily Gaussian in nature (with a mean of $\langle G \rangle \times N$) due to the Central Limit Theorem [73], but for low energy events, this is not necessarily the case. The statistics of the avalanche process are historically well-described by the Polya distribution [188–192], which has a single shape parameter θ that allows the distribution to vary between exponential ($\theta = 0$) and Gaussian ($\theta \to \infty$):

$$\begin{aligned} P_{\text{Polya}}\left(E|\theta, \langle G \rangle\right) = \frac{1}{\langle G \rangle} \frac{(1+\theta)^{1+\theta}}{\Gamma\left(1+\theta\right)} \left(\frac{E}{\langle G \rangle}\right)^{\theta} \\ \times \exp\left(-\left(1+\theta\right)\frac{E}{\langle G \rangle}\right) \end{aligned} \tag{3.15}$$

Typically θ is close to 0 for SPCs [192]. This distribution is depicted with different values of θ in Fig. 3.9. The avalanche response for N primary electrons is the N^{th}

convolution of the Polya distribution (each avalanche is assumed to be independent), and is given by [192]:

$$P_{\text{Polya}}^{(N)}\left(E|N,\theta,\langle G\rangle\right) = \frac{1}{\langle G\rangle}\left(\frac{(1+\theta)^{1+\theta}}{\Gamma\left(1+\theta\right)}\right)^{N} \times \left(\frac{E}{\langle G\rangle}\right)^{N(1+\theta)-1} \times exp\left(-\left(1+\theta\right)\frac{E}{\langle G\rangle}\right) \times \prod_{i=1}^{N-1} B\left(\left(i+i\theta\right),\left(1+\theta\right)\right) \tag{3.16}$$

where $B(x,y)$ is the beta function. Equation 3.16 can then be combined with the COM-Poisson distribution section 3.1.1) to give the full detector response E to a particle event depositing energy E_0. This is achieved by multiplying the probability of having N primary electrons with the corresponding avalanche response, and summing over the possible numbers of primary electrons:

$$\mathcal{P}(E|E_0) = \sum_{N=1}^{N_{\max}} P_{\text{COM}}(N|\mu(E_0),F(E_0)) \times P_{\text{Polya}}^{(N)}\left(E|\theta,\langle G\rangle\right) \tag{3.17}$$

where the mean number of primary electrons is $\mu = E_0/W(E_0)$, and the Fano factor may also be a function of the particle energy. $N_{\max}$ is the maximum number of primary electrons possible to produce given the ionization threshold of the atom in question:

$$N_{\max} = \left\lfloor \frac{E_0}{E_{\text{ionization}}} \right\rfloor \tag{3.18}$$

The detector energy response E can be thought of in units of the number of secondary electrons, or the pulse amplitude observed by the detector as an estimate of the energy response (more details on pulse treatment are given in section 3.1.5). In the latter case, the mean gain $\langle G\rangle$ is in units of ADU (arbitrary digital units) per primary electron. This response function is shown for different values of E_0 in Fig. 3.10, for $\theta = 0$ and $F = 0.2$. This demonstrates the relatively broad energy resolutions of

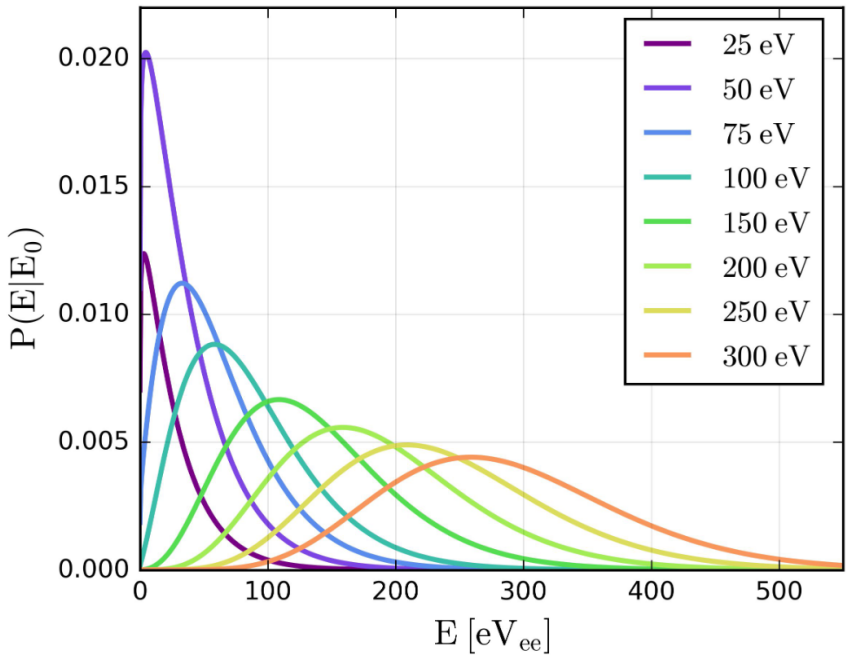


Figure 3.10: The total energy response of a hypothetical SPC, for different deposited energies, obtained using eq. 3.17. This includes the dispersion in primary ionization with $F = 0.2$ and Townsend avalanche dispersion, with $\theta = 0$ and $\langle G \rangle = W^{-1}$ such that the resultant scale is equivalent to the deposited energy scale.

SPCs, which is nearly exponential at low energy due to both primary and secondary ionization statistics.

Note that even though eq. 3.17 can be cast in terms of recorded pulse amplitudes, it does not include the impact of electronic noise in the recorded data, or other distortions due to pulse processing (in particular pulse smoothing, more on this in section 3.1.5). Typically these effects are collectively taken into account by convolving eq. 3.17 with a Gaussian distribution, with a standard deviation representing the extra dispersion added to the amplitude of events by these processes.

3.1.5 Signal processing and Pulse-shape discrimination

The raw signal of a particle event in an SPC is a combination of the total energy response in the gas (described in the previous subsection), and the induced signal on the anode by the avalanche charges drifting away (given by eq. 3.3). Qualitatively,

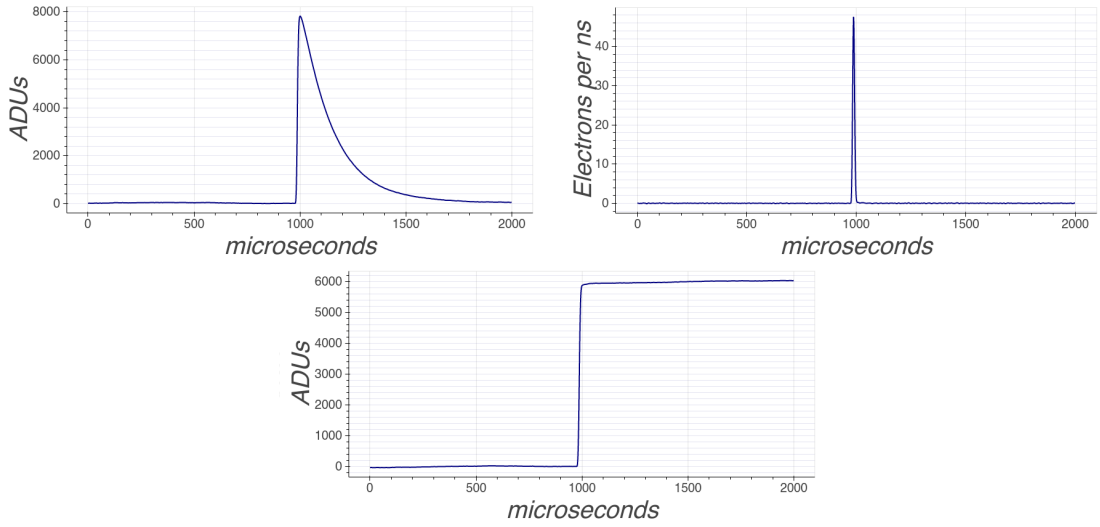


Figure 3.11: Example of a real SPC event (a 2.8 keV ^{37}Ar event) in various stages of treatment. First panel: the raw pulse, baseline-subtracted (y-axis in arbitrary digital units). Second panel: the double-deconvolved, smoothed pulse. Third panel: the cumulative integral of the double-deconvolved pulse (y-axis in arbitrary digital units).

this appears as a sharply rising pulse, rounded slightly near its maximum, followed by a long decaying tail due to the exponential response of the pre-amplifier. One uses the height of this pulse as an estimate of the energy of the event. However, if primary electrons arrive at the sensor spread out in time, the decaying preamplifier response will reduce the signal height from one electron before the next arrives. Therefore, the maximum amplitude of the raw pulse will be less than the sum of the raw amplitudes of all the participating electrons, making this a non-linear energy estimator that behaves differently depending on where the event originated from, gas conditions, etc.

To solve the issue, the raw pulse is "corrected" for the preamplifier response function – as well as a similar pulse shape due to the changing induced current over time – by deconvolving the raw pulse from both of these effects. The resulting treated pulse is essentially a sequence of Dirac delta functions with amplitudes and times corresponding to the avalanches from individual primary electrons. Pulse smoothing and electronic baseline noise smear these into narrow Gaussian-like peaks. The pulse

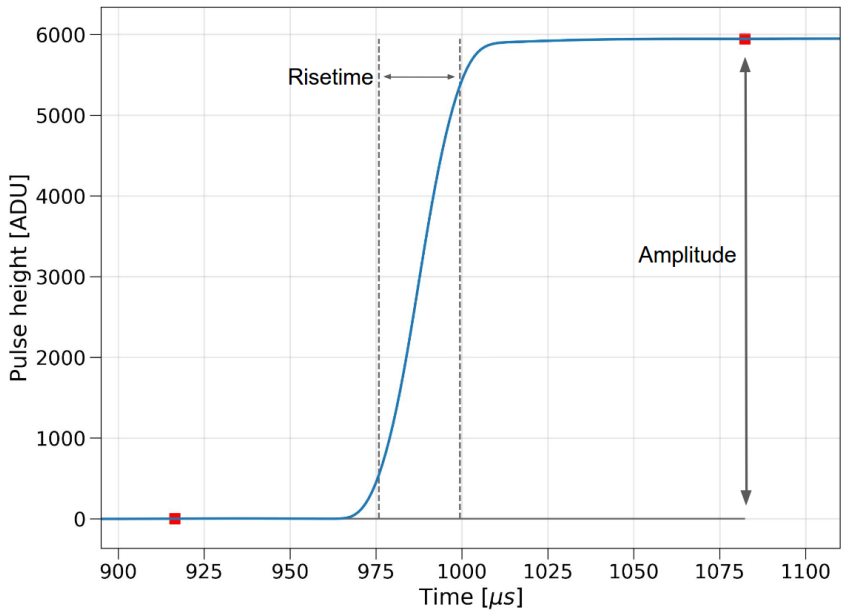


Figure 3.12: A double-deconvolved, integrated pulse (an 2.8 keV ^{37}Ar event). The red markers indicate the start and end of the pulse. The amplitude and risetime (10 – 90%) of the pulse are also labelled; in this case, the amplitude is ≈ 5940 ADU and the risetime is $\approx 11.3\,\mu$s.

treatment algorithm used by NEWS-G is largely the work of F. Vazquez [160], and is summarized below:

1. A trapezoidal filter applied to the raw pulse (the derivative calculated over a span of samples) is used to determine the starting point of the event, i.e. when the derivative exceeds a set threshold.

2. A span of samples before the start of the pulse is used to calculate the average baseline, which is then subtracted from the pulse.

3. The pulse is then deconvolved from the preamplifier (exponential) response function.

4. The trace is smoothed with a rolling average smoothing algorithm.

5. Next the pulse is deconvolved (in frequency space) from the ion-induced current response.

6. The double-deconvolved pulse is smoothed again with a rolling average algorithm.

7. Finally, the cumulative integral of the double-deconvolved pulse is calculated.

An example of a pulse throughout various stages of this process is shown in Fig. 3.11. The main variables used in analyses are defined using the integrated, double-deconvolved pulse, which are depicted in Fig. 3.12. The amplitude of the pulse (an estimate of the event energy) is simply the difference between the pulse height after and before the end and start of the pulse, respectively. This is sometimes denoted by the variable name `DD_AmplADU` throughout this work. Risetime is defined as the time between the pulse reaching 10 % and 90 % (or sometimes 75 %) of its maximum height.

Risetime is useful as a pulse-shape discrimination (PSD) variable, as explained in section 3.1.3. Events originating from larger radii within the SPC tend to experience more diffusion of primary electrons as they travel to the sensor, and therefore will have large risetimes on average compared to events uniformly distributed throughout the volume (as dark matter would be). An example of this behaviour is shown in Fig. 3.13. As many sources of background events are surface contaminants, this limited fiducialization of the detector can be used to reject many background events. This approach does break down at low energy events consisting of just a few electrons; single electron events have the same risetime (which is non-zero due to pulse smoothing and baseline noise) regardless of where they originate in the SPC.

Some useful PSD variables are also extracted from the baseline-subtracted raw pulse, namely the full-width half-maximum (FWHM) and risetime of the raw pulse. These are employed to reject non-physical, discharge-like events thought to originate from microscopic electrical shorts inside the sensor structure. Since these events do not experience a Townsend avalanche in the gas, their pulse shape is an instantaneous rise with a purely exponential decay due to the pre-amplifier. Therefore, one expects

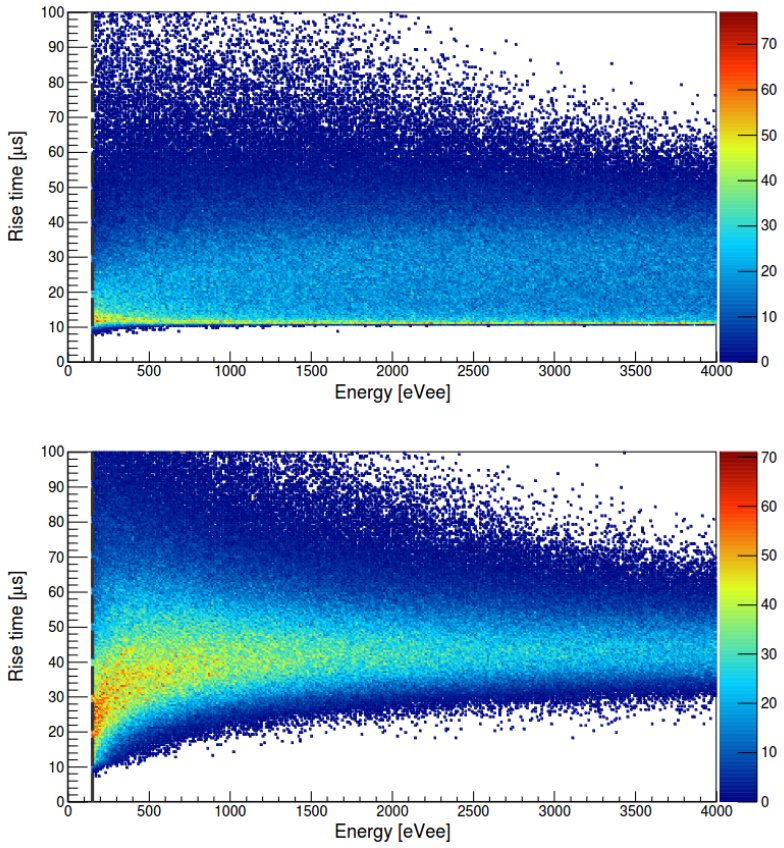


Figure 3.13: 2D histograms of risetime vs. amplitude data from simulated volume (top panel) and surface (lower panel) events, with a uniform deposited energy distribution. This simulation is specific to the SEDINE experiment at the LSM [63], and demonstrates the clear difference in risetime between these populations, except where it breaks down at low energy. Taken from figure 3 of ref. [63].

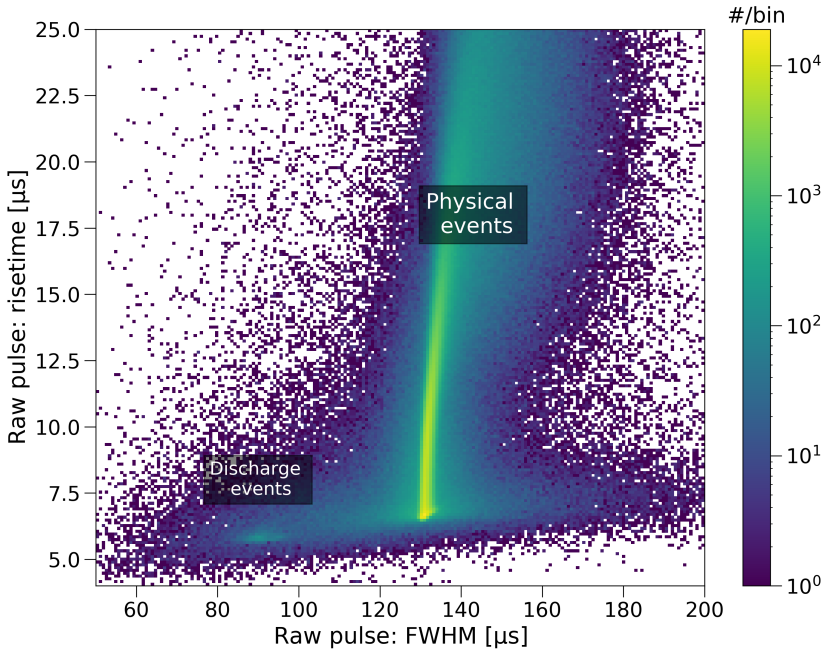


Figure 3.14: 2D histogram of the raw pulse risetime vs. FWHM from ^{37}Ar data taken in a 200 mbar argon gas mixture. The majority of events are physical events, with a small sub-population of discharge events at lower risetimes and FWHM values.

these pulses to have characteristically shorter FWHM and risetime values. Note that the risetime calculated using the double-deconvolved pulse is not useful in this case, as that algorithm deconvolves the discharge pulses for a shape component they do not have, severely distorting them. An example of these different PSD populations is shown in Fig. 3.14. More specifics on PSD cuts for specific analyses are discussed later on in section 4.4.

When the diffusion experienced by primary electrons is very large, it is possible to distinguish individual primary electrons in the double-deconvolved pulse signal, as opposed to more typical examples like Fig. 3.11 where electrons arrive in rapid succession and appear as one peak. The conditions required for this to happen include having a slow-drifting gas, and a large SPC. To date, this has only been the case with the recent LSM physics campaign (discussed in sections 4.2 to 4.5) in a 140 cm SPC with methane gas. More details of this special pulse-processing are given in section 4.2.2.

3.1.6 Previous NEWS-G results

The first dark matter physics result of the NEWS-G collaboration was derived from physics data taken with the "SEDINE" detector at the LSM in 2015 [63]. This device was a 60 cm copper SPC with a 6.3 mm single anode (pictured in Fig. 3.2). A target gas of 3.1 bar, Ne + 0.7% CH_4 was used, with an applied voltage of 2520 V. Almost 42 kg · days of data was collected, leading to the publication of the WIMP exclusion limit shown in Fig. 3.15. At the time, this was world-leading at a WIMP mass of 0.5 GeV/c^2 [63].

Although this was a notable achievement (being the first dark matter result from any gaseous detector collaboration at the time), the experiment was affected by several shortcomings. The only calibration source used was ^{37}Ar, which showed evidence for strong anisotropies in the electric field produced by this early sensor with no corrective umbrella anode. Because of the degeneracy between the W-value and $\langle G \rangle$, the lack of independent UV laser calibration (see section 3.2.1) forced the assumption of the W-value found in the literature for neon – 36 eV [63, 193]. No realistic model for primary ionization statistics (like the COM-Poisson distribution) was known at that time, so it was taken to be Poissonian. No neon quenching factor measurements had been performed [184], so the SRIM simulation software was used despite its known shortcomings [63]. The detector also observed a large, unexplained excess of low energy events far exceeding any modelled background sources (similarly to other low-mass dark matter experiments, see section 2.4.1).

Moving forward from this first result, the collaboration has put a great deal of effort into correcting hardware deficiencies, improving calibration techniques, and better understanding the physics of SPCs. Some of these improvements have already been presented earlier in this section, with many more to be presented throughout the rest of this chapter as part of the upcoming second NEWS-G dark matter result.

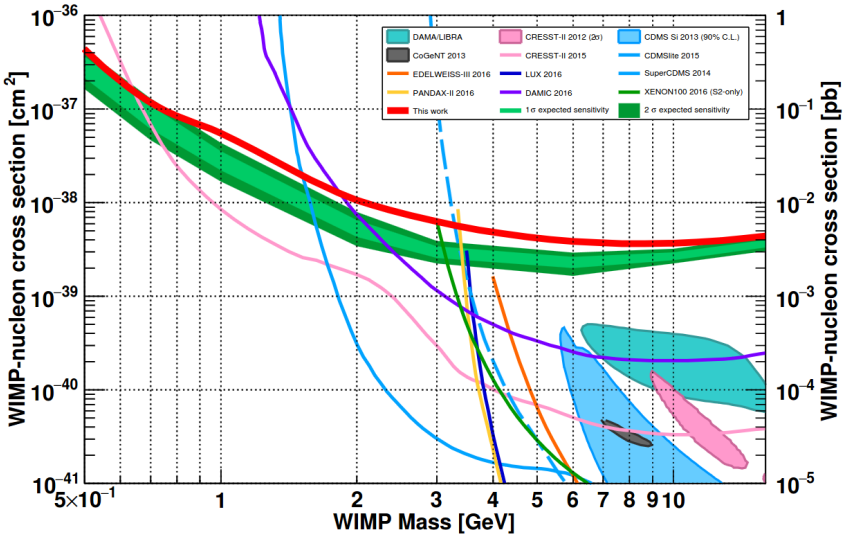


Figure 3.15: Spin-independent WIMP-nucleon scattering cross-section exclusion limits from a variety of experiments, including the NEWS-G exclusion limit from the LSM obtained with the SEDINE detector (red curve), along with its sensitivity band (green shaded bands). Also shown are several possible dark matter signal regions from the CDMS, CoGeNT, DAMA/LIBRA and CRESST-II experiments (blue, grey, turquoise, and pink-shaded regions respectively). Figure 10 from ref. [63].

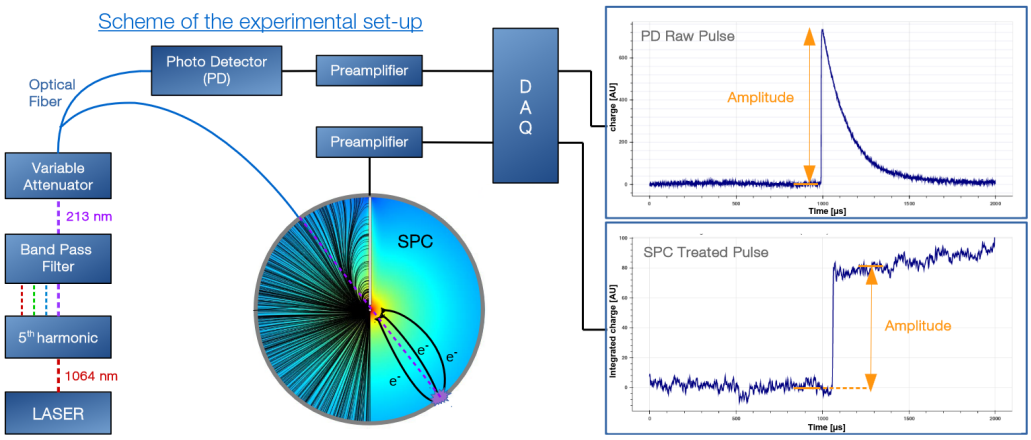


Figure 3.16: UV laser calibration setup for NEWS-G, including the internal components of the laser, optical paths into an SPC and photodiode, and subsequent signal acquisition. Examples of a photodiode (PD) pulse and treated SPC laser pulse are also shown. Figure 1 from ref. [192].

3.2 Calibration strategies

3.2.1 UV laser calibration

A principle calibration tool used by NEWS-G is a UV laser system, used to produce photoelectrons from the inner surface of the SPC via the photoelectric effect [192]. The standard setup employed in multiple labs uses a 213 nm pulsed laser, shone into the sphere through a fibre optic cable and an optical fibre vacuum feed-through. The lasers used are produced by Quantum Light Instruments [194]. A 10 mJ/pulse infrared (1064 nm) YAG laser is coupled to a (internal) 5^{th} harmonic waveform generator to ultimately produce the 213 nm light. The device also includes a variable transmission neutral density attenuator to control the laser intensity. A split optical fibre is used to simultaneously shine the light inside the sphere as well as on a ThorLabs silicon photodiode for event tagging. The photodiode signal is captured by the same DAQ system (optionally with external signal amplification) as the sphere to ensure proper timing between these two channels. This setup is depicted in Fig. 3.16, with more details given in ref. [192].

At peak intensity, the laser produces 0.5 mJ pulses of 213 nm light, which corre-

sponds to roughly 5.4×10^{14} photons. However, the transmission of this light through the optical fibre and feed-through is very low (as they are not optimized for UV light). Additionally, there is an unknown (but evidently very low) efficiency for these photons to extract electrons that arrive at the sensor, rather than immediately recombining with the SPC surface. Collectively, these factors can reduce the large initial number of photons produced to $\mathcal{O}(100)$ photoelectrons per pulse. The laser is often operated with a relatively low pump current and internal transmission, further reducing this yield to $\mathcal{O}(1)$ photoelectron.

The primary application of the UV laser is to measure the avalanche response of an SPC, specifically the parameters θ and $\langle G \rangle$. While other calibration sources can also be used to measure these parameters (i.e. ^{37}Ar data), the signal from particle interactions in the gas are also dictated by the W-value and Fano factor, which are degenerate with $\langle G \rangle$ and θ respectively. As the number of primary electrons produced by the laser only depends on the laser intensity and work function of the metal (and is subject to Poisson statistics), it provides a means to decouple these energy response parameters with an independent measurement [192]. Typically for such calibrations, the DAQ is set to trigger on the photodiode channel so that only laser events are recorded. The intensity of the laser is tuned using the variable attenuator to produce few-electron events in the SPC.

At a given laser intensity, the number of primary electrons reaching the sensor is assumed to follow a Poisson distribution with mean μ. Each UV photon striking the surface has a given probability to extract an electron (i.e. following a Bernoulli distribution), so the probability for obtaining N primary electrons given K incident photons follows a binomial distribution when K is small. In this case, the number of incident photons in each pulse is so large that this tends towards a Poisson distribution [73] with a mean number μ. Therefore it is impossible to collect laser calibration data with only single primary electrons – if $mu \lesssim 1$, there will necessarily be contributions of null (zero electron) events, and others with two, three, etc. electrons. It

is worth noting that this assumption still holds in the face of charge trapping, which may be present. In this case, the original Poisson distribution of primary electrons reaching the sensor with mean μ would be convoluted with a Binomial distribution with trapping probability p, resulting in a Poisson distribution with mean $\mu \times (1-p)$ (see Appendix of ref. [192]). As the laser intensity does fluctuate throughout data taking (due to instabilities in the laser itself), the recorded photodiode data is used to select events with as close to a constant μ as possible, typically taking a subsection of the data with a photodiode amplitude within $\pm\,5\,\%$ of a given value [192].

Laser calibration data is processed largely as described in section 3.1.5. The only difference is that the pulses are integrated within a pre-defined window, rather than relying on a trapezoidal filter to determine the start of the pulse. This is necessary to properly treat null events in the laser data, with no evident pulse. The fixed window can be easily defined, as the DAQ software automatically attempts to center all pulses in the middle of each event window. This also ensures that every event integrates the same amount of baseline noise, which is not the case if the pulse integration window is automatically determined. Therefore this noise can be represented as a Gaussian signal with the same standard deviation for all events [192].

The amplitude (E) spectrum for laser events can be calculated analytically in a similar way as eq. 3.17, but with the number of primary electrons being Poisson-distributed rather than using the COM-Poisson distribution, and with an additional term for null events:

$$\begin{aligned}\mathcal{P}(E) = \Bigg(& P_{\text{Poisson}}(0|\mu) \times \delta(E-\omega_b) + \\ & \sum_{N=1}^{\infty} P_{\text{Poisson}}(N|\mu) \times P^{(N)}_{\text{Polya}}(E|\theta, \langle G \rangle) \Bigg) \\ & \otimes P_{\text{Norm}}(E|\omega_b, \sigma_b)\end{aligned} \tag{3.19}$$

where δ is the Kronecker delta function. $P_{\text{Norm}}(E|\omega_b, \sigma_b)$ represents the electronic baseline noise with mean and standard deviation ω_b and σ_b. An example of typical

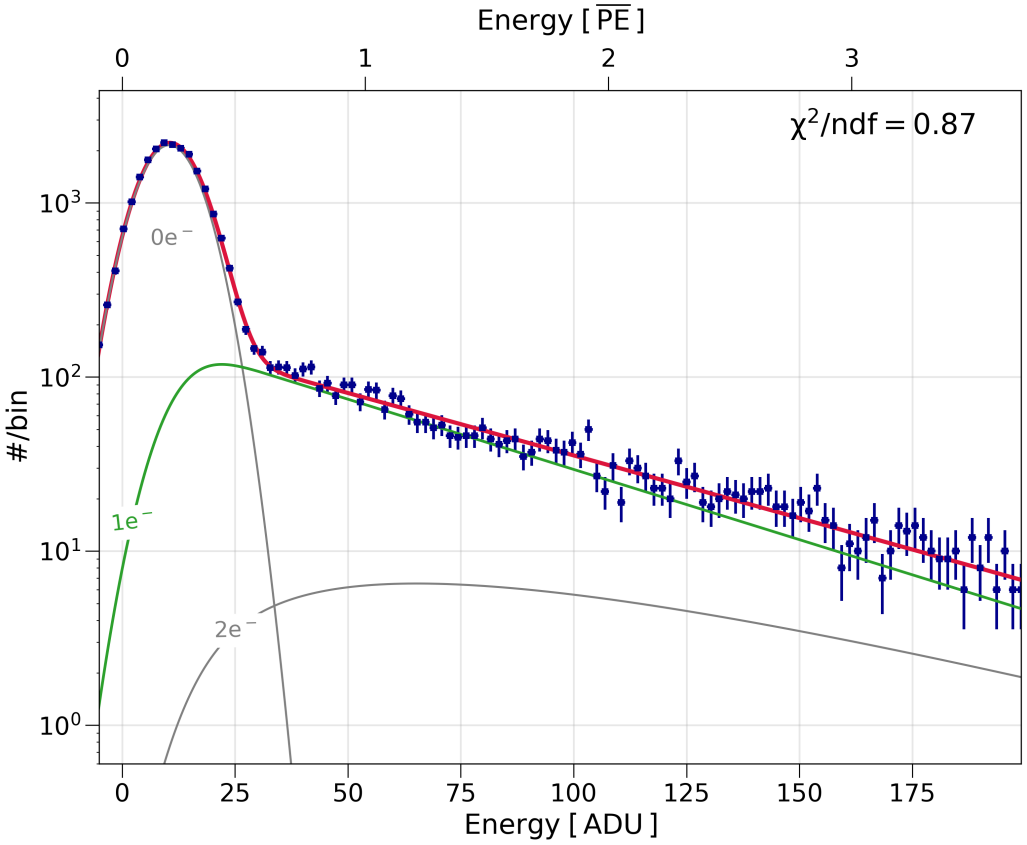


Figure 3.17: An example of a laser calibration amplitude spectrum (blue histogram), taken in a 30 cm SPC with 50 mbar of pure methane gas. The total fit of the spectrum is shown (red curve), as well as the contributions from zero, one, and two-electron events (grey and green curves). In this example, the best-fit parameters were $\theta = 0.0012^{+0.03}_{-0.0012}$, $\langle G \rangle = 53.8^{+1.3}_{-1.2}$ ADU, and $\mu = 0.23 \pm 0.01$. Using these results, the amplitude scale can be converted to units of the average amplitude for a single primary electron (top scale). Note that the null electron peak is shifted above 0 ADU ($\omega_b > 0$); this is an occasional artifact of pulse processing, wherein the double-deconvolved pulse has a slightly rising baseline.

laser calibration data is shown in Fig. 3.17. This spectrum can be fit with eq. 3.19 (which is also shown in Fig. 3.17) to extract estimates for θ and $\langle G \rangle$. Although more sophisticated laser analyses are presented in sections 3.3 and 4.3, this basic approach developed in ref. [192] demonstrates the salient features of laser data. Specifically, one can see the Gaussian peak centred near 0 ADU from null events, and the approximately exponential tail of mostly single-electron events, corresponding to the Polya distribution with $\theta \approx 0$.

The UV laser may also be operated in a high-intensity mode, wherein potentially hundreds of photoelectrons are produced in each event. The main application of this type of data is real-time monitoring of the detector response [192]. The gain of the detector may change over time due to changing electric field conditions, as well as changing gas conditions (e.g. temperature/pressure fluctuations). The detector response may also be affected over time due to increasing gas contamination, and therefore electron attachment. All of these effects can cause measurable changes in the amplitude of high-intensity laser events. As is shown in Fig. 3.18, these fluctuations also apply to other physical events in the sphere and therefore can be accounted for in analysis. The laser data can also optionally be used to correct for these fluctuations over time if appropriate.

Another important output of laser calibration data is the measurement of electron drift times. Because the photodiode and SPC signals are synchronized, this can be directly calculated using the time difference between the start of the photodiode pulse and SPC pulse (see ahead to section 4.2.2) – an example of this time difference can be seen in Fig. 3.16. Electron drift times depend on the size of the SPC, the drift properties of the gas (see section 3.1.3), and also on the rate of other events in the detector due to space charge effects (see section 4.3). For the LSM methane campaign, typical drift times ranged from ~ 0.7 to 1.5 ms.

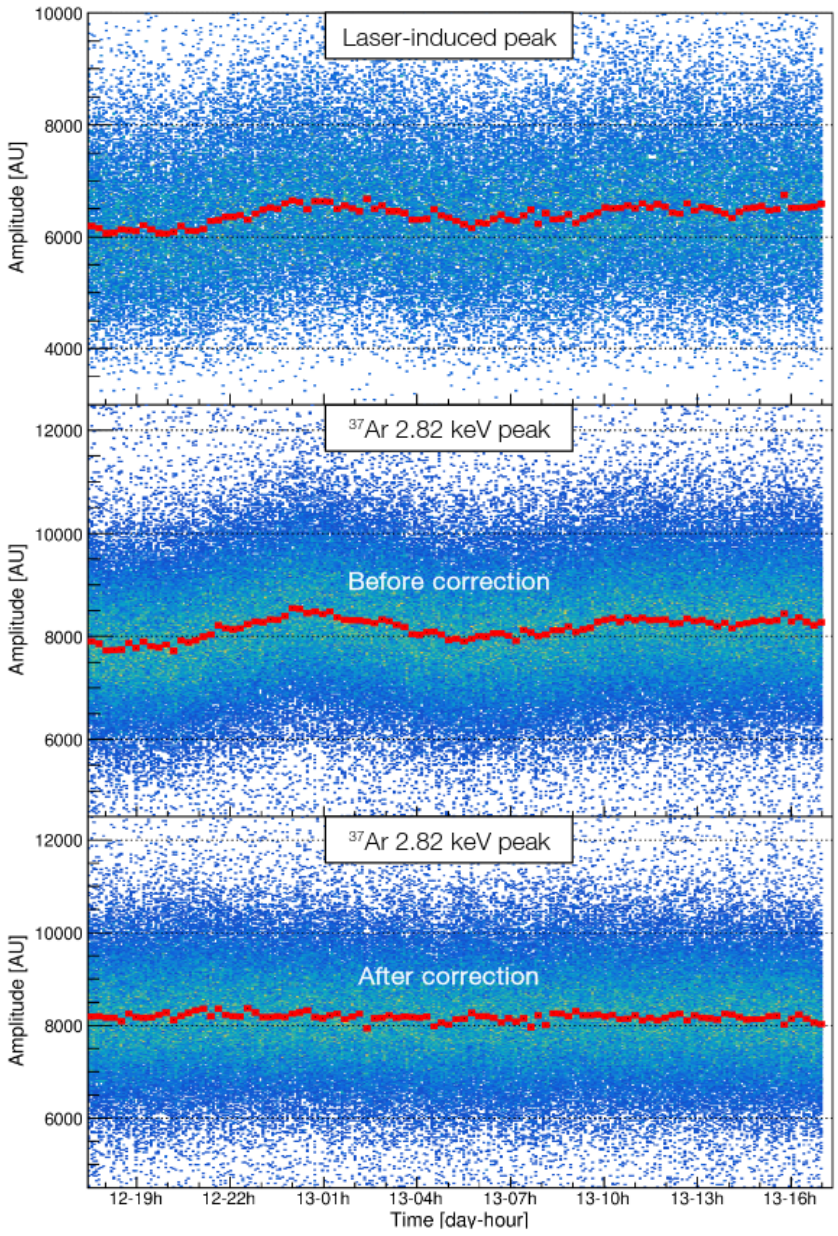


Figure 3.18: A 2D histogram of the amplitude of high-intensity laser events over time (top panel), as well as 2.8 keV ^{37}Ar events (middle panel). The red markers indicate the mean amplitude in different time segments. The bottom panel shows the same ^{37}Ar events but corrected for fluctuations in the detector response over time using the laser data. Taken from figure 6 of ref. [192].

3.2.2 ^{37}Ar Calibration

A frequent calibration source used by NEWS-G is ^{37}Ar. As a radioactive gas, ^{37}Ar provides calibration events from the entire SPC volume. Its decay also produces events with a range of energies, which can be used to verify the energy-linearity of a detector. Samples of this gas – with a 35 day half-life [195] – are produced regularly by collaborators at the Royal Military College of Canada [196]. Samples of CaO powder are inserted close to a SLOWPOKE-II research reactor. The neutron flux from the reactor produces ^{37}Ar via the reaction [196]:

$$^{40}\mathrm{Ca} + n \rightarrow {}^{37}\mathrm{Ar} + \alpha \tag{3.20}$$

The ^{37}Ar is then liberated from the CaO matrix by agitating and heating the vial, then passively transferred to an un-irradiated sample container by allowing it to diffuse through a HEPA filter (to ensure no CaO powder is carried through). The quantity of ^{37}Ar produced is calculated based on germanium counter measurements of ^{41}Ar and other radio-isotopes produced as by-products of irradiating the CaO and air trapped within it [196]. This technique provides a fast and efficient way to produce ^{37}Ar, requiring only ~ 15 minutes of irradiation to produce $\sim 1\,\mathrm{kBq}$ of ^{37}Ar in a $0.5\,\mathrm{g}$ sample of CaO powder [196]. Producing the sample under vacuum conditions ensures that negligible gaseous contaminants (especially O_2 and H_2O) are introduced into the SPC when the sample is allowed to diffuse into the detector [196].

^{37}Ar decays via electron capture [70] – primarily K-shell or L_1-shell capture, with branching ratios of 90.4 % and 8.4 % respectively [197]:

$$^{37}\mathrm{Ar} + \mathrm{e}^- \rightarrow {}^{37}\mathrm{Cl} + \nu_\mathrm{e} \tag{3.21}$$

The electron vacancy created leaves the atom in an excited state, which resolves through the emission of photons and/or Auger electrons. The total energy emitted per decay is 277 eV or 2.83 keV for L_1 and K shell electron capture, respectively [198]. To

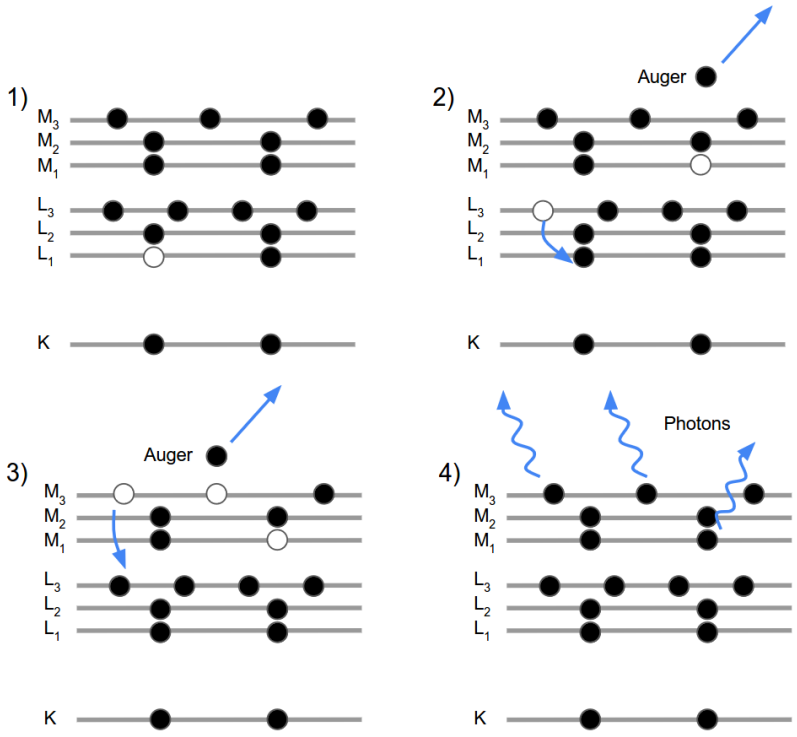


Figure 3.19: An example of one atomic relaxation path following an L_1-shell electron capture decay of ^{37}Ar. The level diagrams show the K, L and M shells, with electrons being represented as white circles and vacancies as black circles. The atom starts off with a vacancy in the L_1 shell (1), followed by two Auger electron transitions (2-3), and then photon emission (4).

use ^{37}Ar as a calibration source, the different decay paths must be modelled. Knowing the total energy deposited per event is not adequate due to the energy dependence of the W-value. That is, the average number of primary ionizations produced by two particles with energies E_A and E_B is not the same as the average yield of a single particle with energy $E_A + E_B$, especially in the sub-keV regime. Additionally, the different decay paths contribute different numbers of free electrons from the decay itself (Auger electrons) that contribute to the total ionization signal as well.

Therefore, a detailed simulation of the electron-capture decay of ^{37}Ar is necessary. The MC produced for NEWS-G described in this work follows the method described in ref. [199], as well as simulations of ^{37}Ar performed by the DarkSide collaboration [200–202]. Atomic transition data (including the types, energies, and probabilities of

different transitions) was obtained with code "RELAX", using the EADL2017 library [198, 203]. For Auger electron transitions, this data specifies the initial vacancy, the electron that fills the vacancy, and the electron that is ejected from the atom. These three shells are specified in order. For example, a KL_1L_3 Auger transition involves an electron falling from the L_1 to K shell, and an electron being ejected from the L_3 shell (in this case with kinetic energy 2.346 keV [198]), leaving two vacancies. Photon (X-ray) transitions simply involve an electron from a higher shell filling a vacancy. The atomic relaxation process terminates with low-energy photon transitions, with the atom accepting free electrons from the environment to vacancies in an M shell. An example of a complete decay path from an L_1-shell capture is depicted in Fig. 3.19.

As there are 45 Auger and 17 unique X-ray transitions possible following K-shell capture, there are thousands of possible decay paths to consider. This is done using an MC of the relaxation process, according to the following steps:

1. Either a K-shell or L_1-shell decay is randomly chosen according to their branching ratios [197].

2. The initial state of the atom is determined, with either a K or L_1 shell vacancy from the electron capture interaction. Note that there is a 2/3 chance of a second electron residing in the M2 shell, or a 1/3 chance that it is in the M3 shell, according to the RELAX transition data [198].

3. A possible decay path proceeding from the current state is chosen (i.e. one for which the initial vacancy is in the correction position, and other electrons involved are present).

4. Vacancies are noted accordingly, with the previous one moving, and possibly an additional vacancy being created by an Auger transition. The emitted particles are recorded.

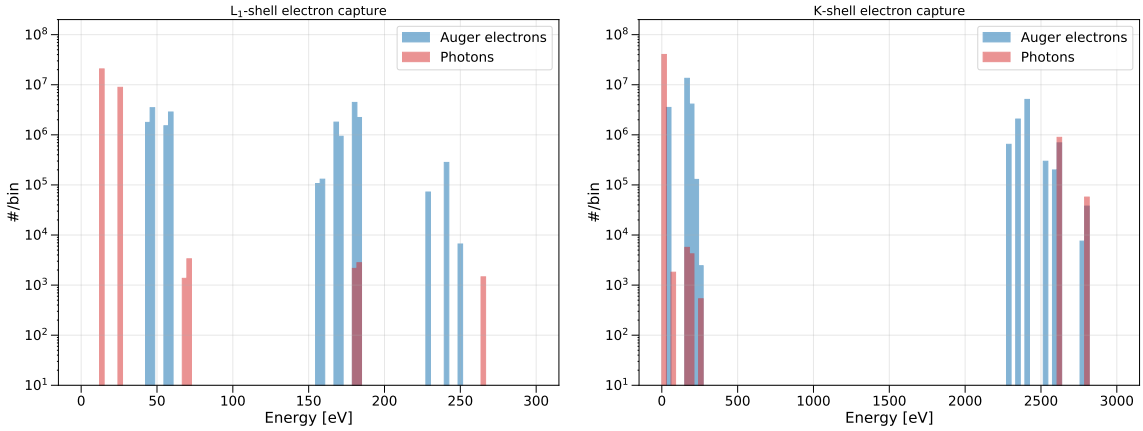


Figure 3.20: The spectrum of electrons (blue) and photons (red) emitted from the L_1 and K-shell electron capture decays of ^{37}Ar (left and right panels, respectively), obtained by simulating 10^7 decays.

5. Steps 3 and 4 are repeated until there are only vacancies in the M shells, at which point no Auger or X-ray transitions are possible.

6. The M-shell vacancies are resolved by photon emission, assuming that there is an infinite supply of free electrons in the environment to participate.

This simulation was performed for 10^7 decays, resulting in the emitted spectra of photons and electrons shown in Fig. 3.20. Typical decay paths involve one or more Auger electron transitions (which are generally more probable than initial X-ray transitions), ending with several low-energy M-shell photon transitions, as expected. The 5 most common L_1-shell decay cascades as given in Table 3.1, along with a list of the particles emitted and the relative probability of that decay path.

In total, 2251 unique decay paths were obtained through this simulation (for K-shell electron capture). Although this is less than the 5213 unique decays reported by the DarkSide collaboration [200], this is likely due to the limited statistics of this MC. To verify that this is the case, the simulation was re-run with the relative probability of every transition automatically set to 1, to strongly favour decays that would otherwise be so rare that even in 10^7 decays they may not occur. Using this approach, 8469 unique decay paths were found, suggesting that many rare paths exist

Decay transitions	Auger e^- [eV]	Photons [eV]	Relative Probability
$L_1L_3M_1$ Auger $\rightarrow L_3M_3M_3$ Auger $\rightarrow 2 \times M_3,\ M_1$ photon	46.46, 180.06	$2 \times 12.97,\ 24.54$	1
$L_1L_3M_1$ Auger $\rightarrow L_3M_2M_3$ Auger $\rightarrow M_3,\ M_2,\ M_1$ photon	46.46, 180.06	$2 \times 12.97,\ 24.54$	0.67
$L_1L_2M_1$ Auger $\rightarrow L_2M_2M_3$ Auger $\rightarrow M_3,\ M_2,\ M_1$ photon	44.46, 182.06	$2 \times 12.97,\ 24.54$	0.65
$L_1L_3M_1$ Auger $\rightarrow L_3M_1M_3$ Auger $\rightarrow M_3,\ 2 \times M_1$ photon	46.46, 168.49	$12.97,\ 2 \times 24.54$	0.64
$L_1L_3M_3$ Auger $\rightarrow L_3M_3M_3$ Auger $\rightarrow 3 \times M_3$ photon	58.03, 180.06	3×24.54	0.60

Table 3.1: The 5 most common decay paths following L_1-shell electron capture of ^{37}Ar, ordered by their relative probability

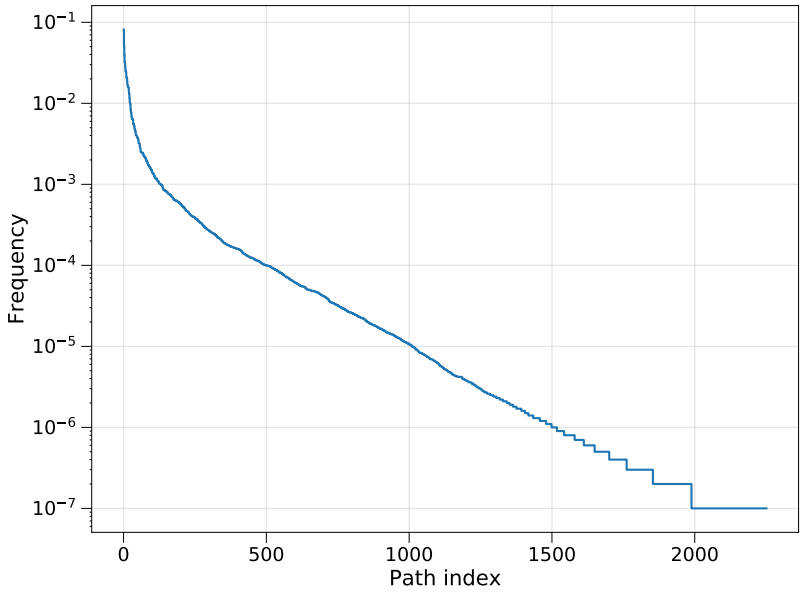


Figure 3.21: Frequency of occurrence vs. decay path index number for K-shell electron capture decays of ^{37}Ar, from an MC with 10^7 events.

that can be missed without extremely large simulations. One can see that this is the case by plotting the frequency of occurrence vs. a decay path index, using the original MC results, which is shown in Fig. 3.21. Of course, these rare decay cascades will not have an appreciable effect on the overall model of ^{37}Ar decay because of their scarcity.

One notable assumption made by this simulation pertains to the order in which vacancies are filled when multiple exist at the same step of the simulation. One approach would be to fill vacancies according to the order in which they are created in the simulation. However, lacking any information on the lifetimes of the various intermediate excited states from the RELAX code [198], it may be generally assumed that the highest-energy states will have the shortest lifetimes [70]. Therefore in this simulation, all existing vacancies in an event are sorted by their energy level following each step, and filled in that order. The "chronological" approach was also attempted for the sake of comparison. As shown in Fig. 3.22, this results in negligible differences that are attributable to statistical uncertainty in the MC.

The next step in modelling the decay of ^{37}Ar in an SPC is to consider the propaga-

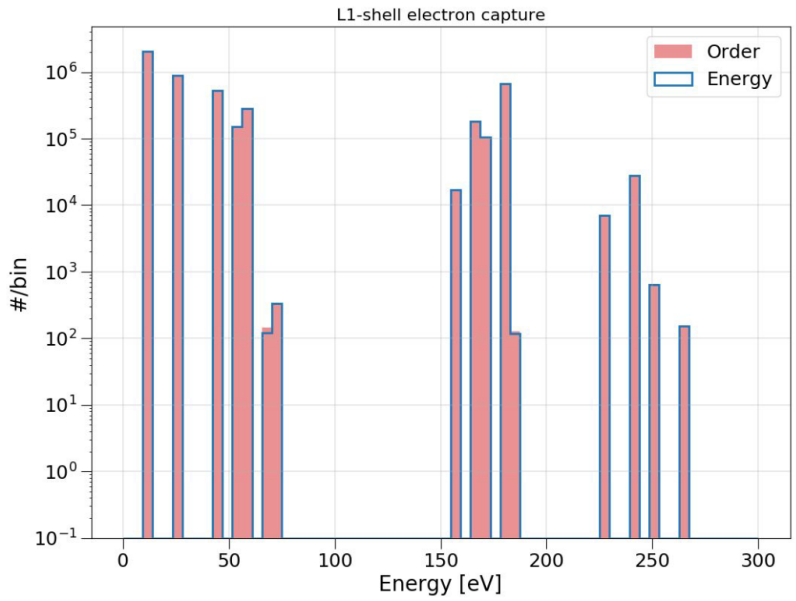


Figure 3.22: The energy spectra of emitted particles following L_1-shell electron capture of ^{37}Ar (Auger electrons and photons), simulated by filling vacancies in MC chronological order (red histogram) as well as by their energy level (blue empty histogram).

tion of the decay products in the low-density gas environment of these experiments. In particular, there may be a non-negligible (or indeed large) probability that high-energy X-rays will escape the sphere without depositing any energy, or that electrons may collide with the detector wall before depositing their full kinetic energy. These effects were modelled with Geant4 simulations [204], specifically using the "Shieldings" physics list [205]. All of the possible decay products of ^{37}Ar were simulated, distributed homogeneously throughout the SPC, with isotropic trajectories. The spectrum of energy deposited in the gas by each decay particle is obtained, giving a probability distribution function for the energy deposited by that specific radiation. The results of such simulations are dependent on the size of the SPC and target gas (both composition and pressure) in question; therefore the modelling of ^{37}Ar decays becomes experiment-specific at this stage. Examples of the deposited energy spectra of several ^{37}Ar decay electrons and X-rays are shown in Fig. 3.23, for the conditions of the pure methane gas physics campaign at the LSM (see section 4.1.3) as well as

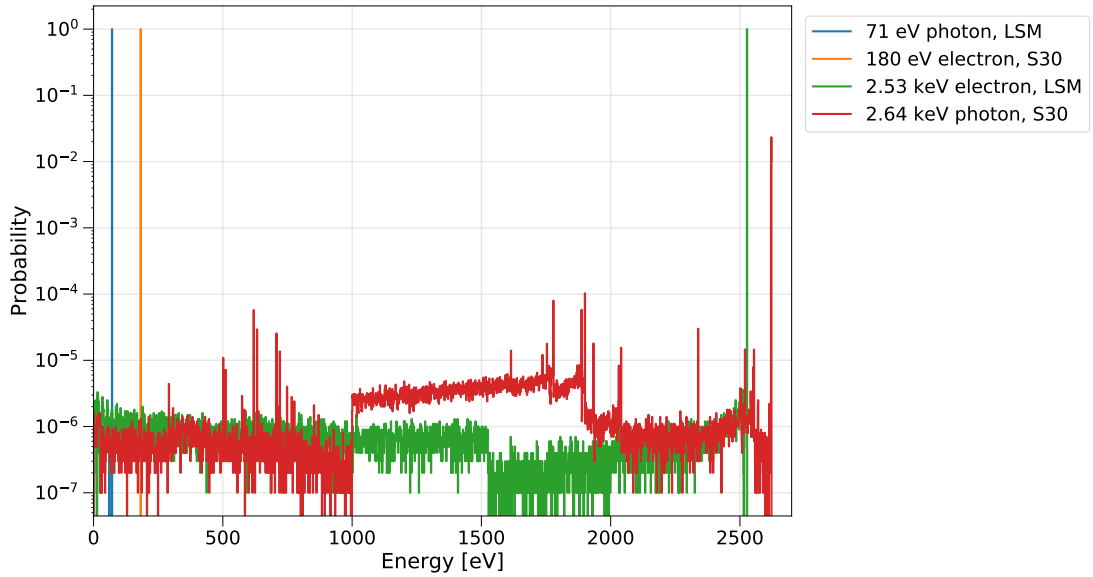


Figure 3.23: Deposited energy spectrum of several electrons and X-rays produced by ^{37}Ar decays, obtained from Geant4 simulations. These examples are given for the conditions of the pure methane gas physics campaign undertaken at the LSM ("LSM", see section 4.1.3), as well as methane gas experiments done in a 30 cm SPC at Queen's University ("S30", see section 3.3).

pure methane measurements in a 30 cm SPC (see section 3.3). One can see from these select examples that, while low-energy X-rays and electrons are likely to deposit their full kinetic energy in the SPC, high-energy particles often do not. The $\mathcal{O}$(keV) X-rays often Compton scatter before escaping, while the electrons may reach the vessel wall before depositing their full energy.

Combining the cascade simulation results and Geant4 propagation spectra described above, the total description of the decay of ^{37}Ar inside an SPC is achieved. An example of the spectrum of the number of decay particles vs. total energy deposited per decay is shown in Fig. 3.24. As many as 11 individual particles may be produced in a single decay, or as few as 1 that actually deposit energy in the SPC. In addition to the strong contributions at the full decay energies of the L_1 and K shell (277 and 2830 eV respectively [198]), there is a broad continuum of deposited energy in between those peaks, as well as a prominent line at 200 eV from partially escaped K-shell decays. A large number of 12.5 eV photons are produced from M-shell tran-

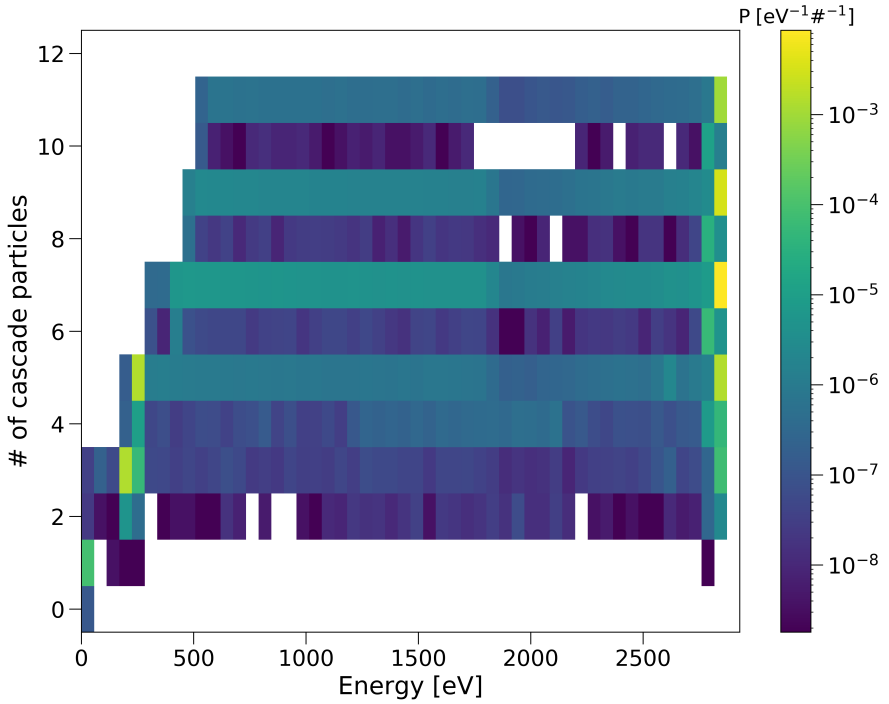


Figure 3.24: A 2D histogram of the number of decay cascade particles (Auger electrons and photons) vs. the total energy deposited in an SPC per event following the electron capture decay of ^{37}Ar. This result is simulated for a 30 cm SPC with 50 mbar of pure CH_4 gas.

sitions, but these are below the ionization threshold of the gases used by NEWS-G. A full study showing how these results are used to fit ^{37}Ar calibration data is given in section 3.3.

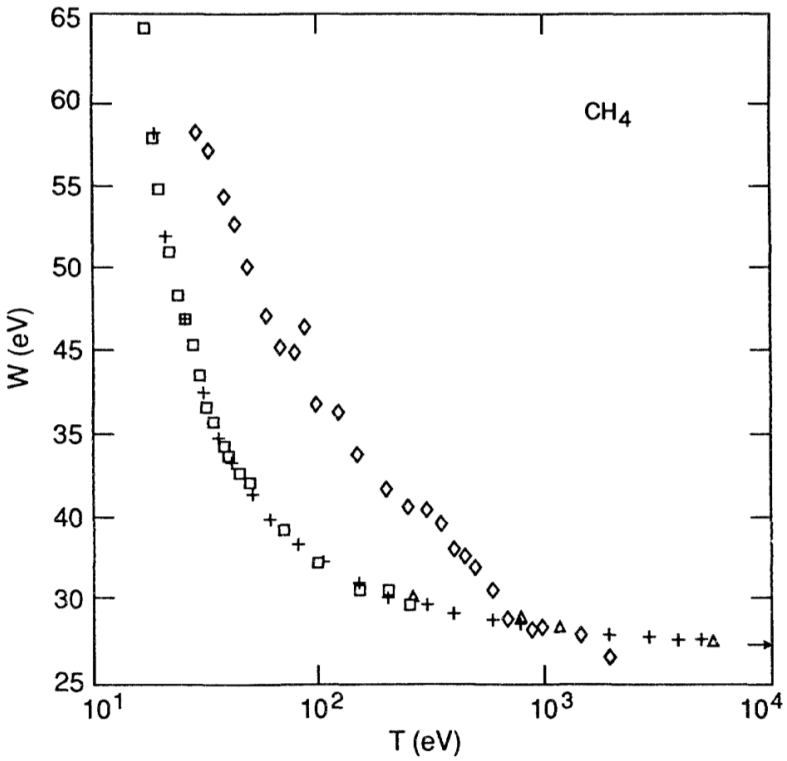


Figure 3.25: Various measurements of the W-value in pure methane gas as a function of energy. The measurements include those by Combecher [209] (square markers), Krajcar-Bronič *et al.* [207] (triangle markers), Waibel and Grosswendt [208] (plus-sign markers), and Smith and Booz [206] (diamond markers). The ICRU report 31 high-energy W-value is indicated by an arrow [167]. Taken from figure 8.7 from ref. [166].

3.3 W-value measurements in methane

As discussed in section 3.1.1, the W-value is central to describing the energy response of SPCs. While measurements of this quantity have been made in a variety of target gases and at varying energies [166], not all conditions relevant to NEWS-G have been studied. In particular, few measurements exist of the W-value in pure methane gas. Those that do exist are somewhat conflicting, as can be seen in Fig. 3.25 [166, 206–208].

As this was the target gas used in the recent NEWS-G physics campaign at the LSM (discussed for much of the rest of this chapter), dedicated measurements of the W-value in methane were deemed necessary. An *in situ* calibration was performed at

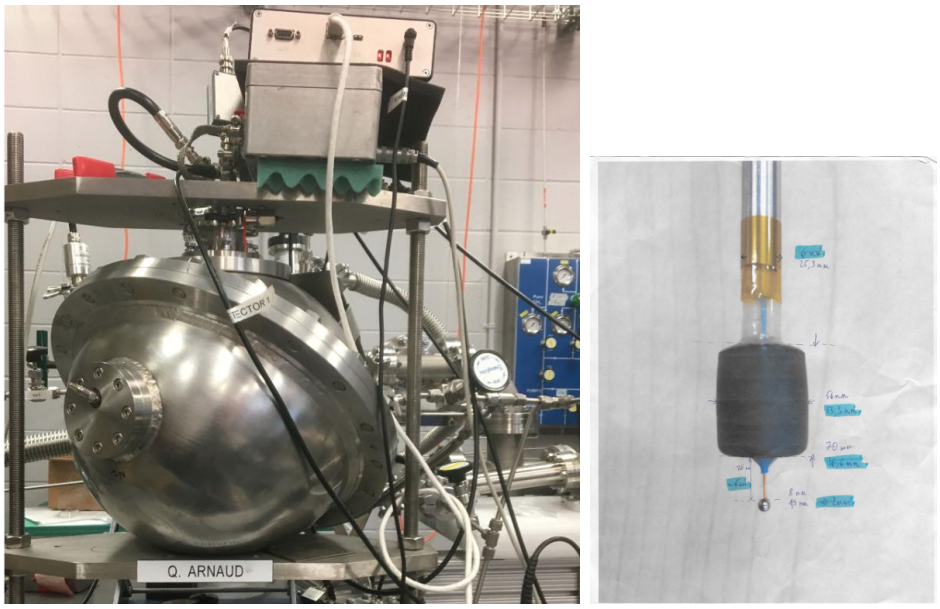


Figure 3.26: The S30 detector setup at Queen's University (left), and the 2 mm anode employed for W-value measurements (right). Taken from figure 1 of ref. [63] (left) and figure 3.22 of ref. [152] (right).

the LSM. However, the impossibility of performing repeat measurements in different conditions prompted further study using a lab-based SPC. The "S30" detector at Queen's University – pictured in Fig. 3.26 – was purposed for this task. The device is a 30 cm diameter stainless steel SPC [192]. The sensor used was a 2 mm stainless steel anode with a Bakelite secondary electrode (also shown in Fig. 3.26. Measurements were performed with pure methane gas, primarily at a pressure of 50 mbar. Other pressures were studied as well, but as will be discussed in later subsections, this did not yield data that is fully understood.

A Canberra Model 2006 charge-sensitive preamplifier was used (with a decay time of $\sim 46\,\mu$s) [210], and the data was digitized at a sampling frequency of 1041 kHz, recorded by the custom calibox DAQ board and Samba software (see section 3.1). ^{37}Ar produced by the RMCC was used as the primary calibration source for these measurements. The rate of decays in the sphere was approximately 40 to 60 Hz for each measurement. Additionally, an aluminum foil-wrapped ^{241}Am source was inserted into the sphere. This setup produced a few Hz of 1.486 keV aluminum flu-

orescence X-rays [198], induced by the 5.5 MeV α-decay of ^{241}Am [211]. Finally, the UV laser system described in section 3.2.1 was used to measure the detector gain $\langle G \rangle$ for all W-value measurements.

Data was collected at anode voltages of 1200, 1250, and 1300 V. This was done to rule out (to the largest extent possible) changing detector gain/electric field conditions having a systematic effect on the study results. Ideally a larger range of anode voltages would have been tested. However, this was not practically achievable due to the limited dynamic range of the DAQ, and the need to have a gain high enough to detect single electrons. The data was processed following the double-deconvolution algorithm described in section 3.1.5, with the only change being a fixed integration window 75 μs long in the middle of the overall 2 ms event window.

An example of the collected data is shown in Fig. 3.27, showing both the high and low energy peaks produced by ^{37}Ar decay (the low energy peak also includes partially escaped K-shell decays, see section 3.2.2). In between the ^{37}Ar peaks, the aluminum fluorescence events are also visible (at around 2800 ADU in this case), at a relatively lower rate. Because of the close overlap between the high energy ^{37}Ar and aluminum fluorescence peaks, these two calibration lines were analyzed together, with the low energy ^{37}Ar data fit separately. In addition to these calibration lines, a large diffuse background of muon events was present. These track-like events tend to have higher risetimes than point-like calibration events. Therefore this background can be largely removed by keeping only events with low risetimes (from approximately 7 to 11.7 μs).

3.3.1 Laser measurements

For each W-value measurement (at a given pressure, anode voltage) it is necessary to know the mean gain of the detector $\langle G \rangle$ to extract an estimate of the degenerate W-value. Further, knowing the avalanche response (modelled with the Polya distribution with shape parameter θ) allows for a disentangled measurement of primary ionization statistics in the SPC, described by the Fano factor. Therefore, UV laser calibrations

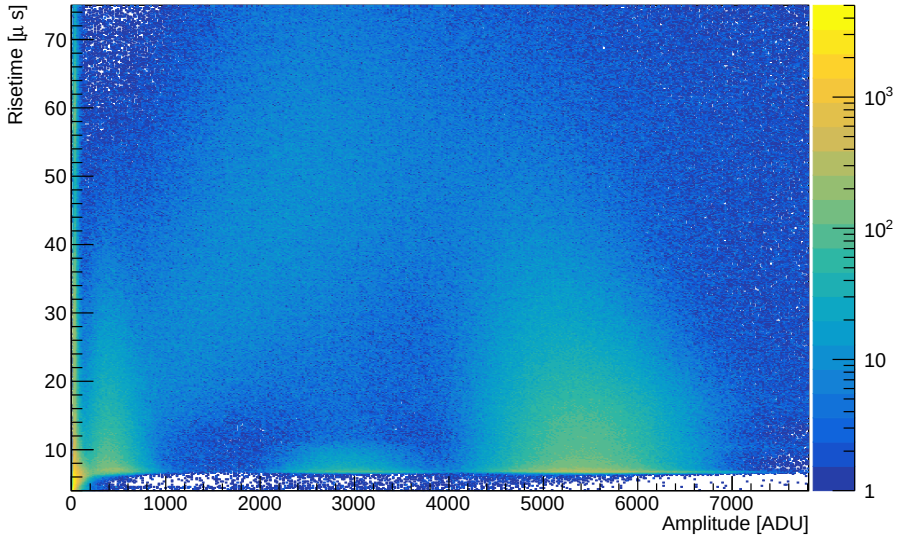


Figure 3.27: A 2D histogram (with the colour-scale giving counts per bin) of risetime vs. amplitude data from the 1250 V W-value measurement, showing ^{37}Ar decays (the high and low amplitude peaks), aluminum fluorescence (the middle peak), and muon events (the broad population at high risetime).

were performed corresponding to each W-value measurement (with the laser pulsing at 10 Hz), with this data being taken immediately before/after the main dataset, in the same gas and conditions.

As described in section 3.2.1 and ref. [192], the laser intensity is tuned to produce $\lesssim$ 1 photoelectron in the SPC on average, and acquisition is triggered by the photodiode coupled in parallel. In this way, even events where no photoelectrons reach the SPC anode are recorded. This data is processed in the exact same way as the W-value calibration data (described above), with a fixed integration window of 75 μs within the larger 2 ms event window. To account for fluctuations in laser intensity, the data is divided into segments based on the recorded photodiode amplitude of each event, with each segment only accepting events with $\pm$ 5% of a given photodiode amplitude [192].

Although the spectral shape of null (0 photoelectron) events is nominally expected to be Gaussian, this was found to not be the case in these measurements. Due to

the high rate of ^{37}Ar decays occurring at the same time, a non-negligible fraction of photodiode-triggered events were in near or total coincidence with other pulses. This results in some laser events having very large amplitudes compared to the expected few-electron events actually produced by the laser. This issue is easily addressed by limiting the range of amplitudes in which the laser data is actually fit. However, near-coincident events with parasitic pulses before or after the integration window defined for pulse processing can noticeably distort the amplitude of laser events, but not so much as to shift them outside of the analysis range completely. This effect can be directly measured by recording photodiode-triggered data with no laser photoelectrons but including the same rate/types of coincident events. This was achieved by simply removing the fibre-optic cable from the SPC, but still allowing the laser to shine on and trigger the photodiode. Note that this is similar to collecting data with a random trigger. An example of data with and without laser photoelectrons in the SPC is shown in Fig. 3.28, from which one can see the slightly non-Gaussian tails of the background amplitude spectrum. Although this effect is small (the spectrum is still approximately Gaussian), the measured background data can be used to inform an empirical background model that matches the data better than the nominal Gaussian approximation.

To that end, this type of no-photoelectron background data was recorded for every laser calibration measurement. To obtain a functional model of the background spectra (which has no clear parametric form), a Kernel Density Estimation (KDE) is a possible choice [212]. KDE algorithms assign a Kernel function (which may be a Gaussian, a "top-hat" function, etc.) for each data point, producing a continuous approximation of the dataset. The type and relative width of each kernel function (the "bandwidth") may be tuned to achieve the desired result, ideally without over or under-smoothing the data [212].

In this case, an additional complication comes from the fact that the background spectra (such as the one shown in Fig. 3.28) span a large range of densities. A KDE

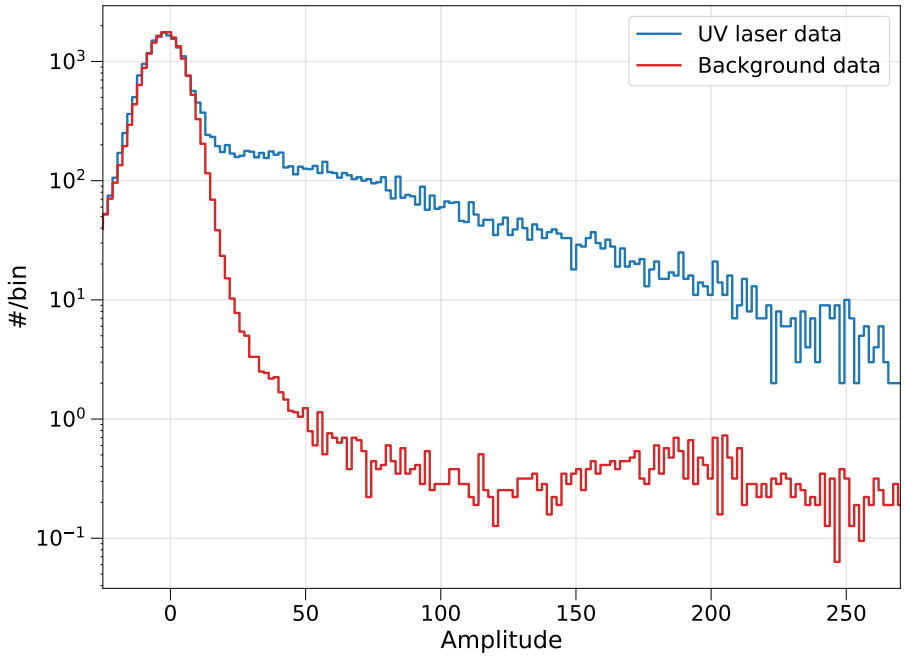


Figure 3.28: The amplitude spectrum of laser data taken with the fibre coupled to the SPC (blue histogram) vs. without (red histogram), showing the spectrum of background events, with non-Gaussian tails and a possible small peak due to coincidences with ^{37}Ar decays. Note that the background histogram is scaled to have the same maximum as the laser spectrum.

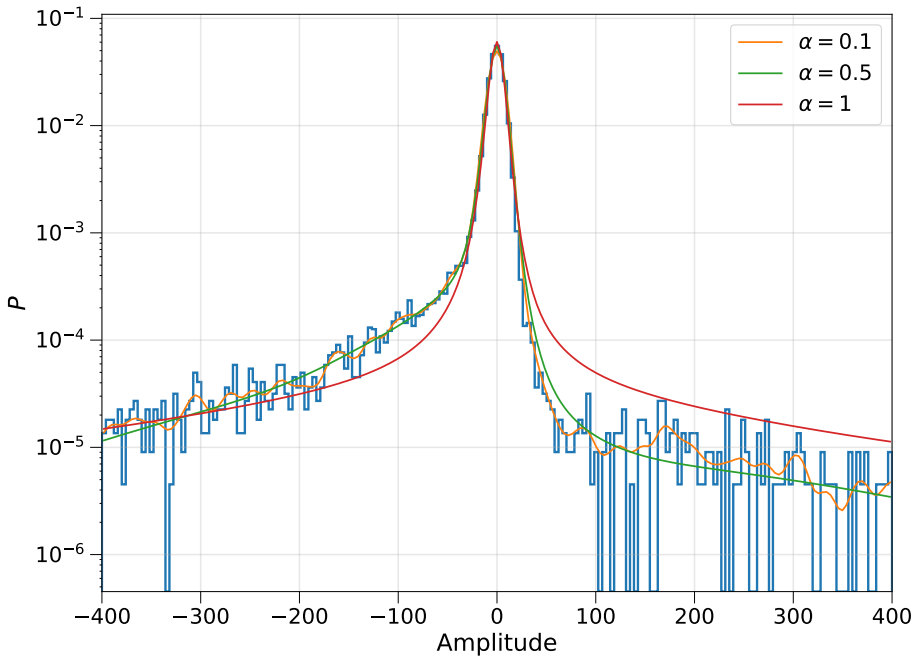


Figure 3.29: An example of a laser background amplitude spectrum (blue histogram), approximated by adaptive bandwidth KDEs with different values of α.

approximation using a single bandwidth value for such a dataset is likely to over-smooth the high probability part of the spectrum if it models the low-probability tails well, or conversely under-smooth the tails if it models the peak well.

A solution to this problem is to use an adaptive bandwidth KDE algorithm, where the bandwidth changes based on the local density of data points [213]. An implementation of such an algorithm for Python is the package `awkde` [214, 215]. In this approach, the kernel function is restricted to be a Gaussian. The bandwidth λ at low density (for a single data point) is given, and a normal KDE $f(X)$ with this single bandwidth is calculated for the data X. Then, the bandwidth is recalculated locally for each data point i by scaling the initial bandwidth according to the nominal density at that data point:

$$\lambda_i = \lambda \times (f(X_i))^{-\alpha} \tag{3.22}$$

where the parameter α is introduced to control the strength of this density-dependence

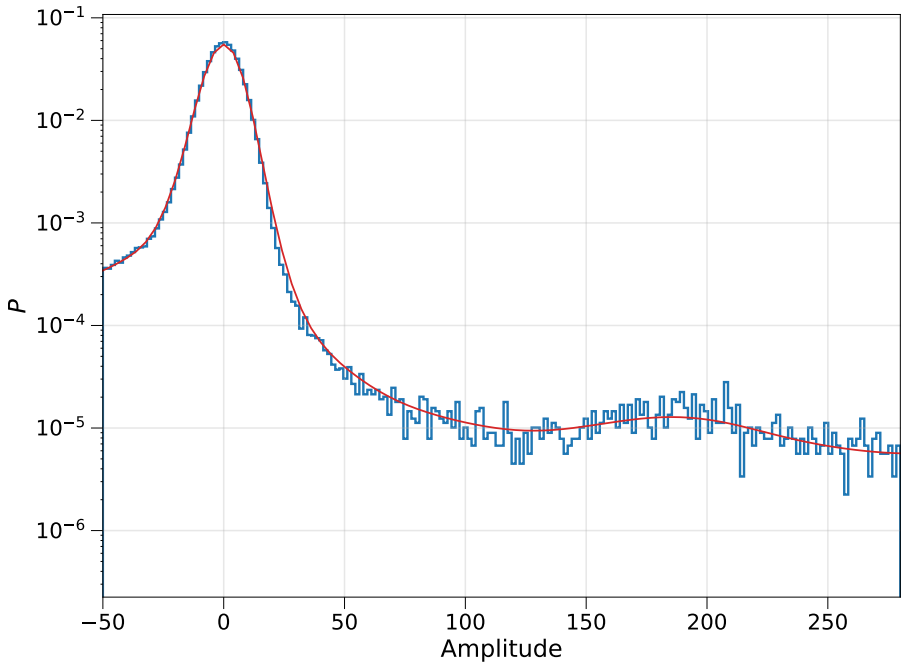


Figure 3.30: The laser background amplitude spectrum for the 1250 V W-value measurement (blue histogram) with the corresponding KDE model, calculated with a nominal bandwidth obtained using Silverman's rule [212], and $\alpha = 0.3$.

[214]. Examples of a laser background spectrum approximated with this adaptive bandwidth KDE, using various values of α, are shown in Fig. 3.29. It is still possible to over or under-smooth the data, but that with a judicious choice of α, it is possible to represent the data well even as it varies in density by many orders of magnitude. The parameter α of the KDE approximation was independently tuned for each laser calibration, and the nominal bandwidth was chosen using Silverman's rule [212], to give results such as the example shown in Fig. 3.30.

The overall laser amplitude spectrum (with E being amplitude) can be represented using eq. 3.19, substituting the normal Gaussian noise approximation with the new measured background model $P_{\mathrm{KDE}}(E)$:

$$\mathcal{P}(E) = \Bigg(P_{\text{Poisson}}(0|\mu) \times \delta(E - \omega_b) + \sum_{N=1}^{\infty} P_{\text{Poisson}}(N|\mu) \times P^{(N)}_{\text{Polya}}(E|\theta, \langle G \rangle) \Bigg) \otimes P_{\text{KDE}}(E) \tag{3.23}$$

Note that this (helpfully) reduces the number of free parameters in the model from 5 to 3. This can be used to construct a binned likelihood function for the histogrammed laser data, with c_i counts in bin i (also applying Stirling's approximation [73]):

$$\log \mathcal{L} = c_i \log\left(\frac{\nu_i}{c_i}\right) - \nu_i + c_i \tag{3.24}$$

where ν_i is the expected counts per bin given a model, $\nu_i = \int_{\text{bin}_i} \mathcal{P}(E) dE$. This likelihood function is maximized for each laser dataset using a Markov Chain Monte Carlo, implemented in Python with the `emcee` library [216] (see Appendix A for more on MCMC approaches for this body of work). The statistical errors reported in this section (or depicted in Fig. 3.32) assume Gaussian likelihood functions. However, the application of these results to subsequent W-value analyses uses the MCMC samples directly.

An example of the fit of a single laser data set (taken at an anode voltage of 1300 V, in the lowest photodiode amplitude segment considered) is shown in Fig. 3.31. The non-Gaussian nature of the modelled background is evident in this example, although its effect on the $> 0\,\text{e}^-$ components of the model is not easily discernible. However, the use of the KDE-modelled background does have an appreciable effect on the final results of the fit. For this particular example, the data was fit using a Gaussian background model for the sake of comparison. This resulted in estimates of $\langle G \rangle$ and θ that differ by approximately 6 % and 30 % respectively, as shown in Table 3.2. For the rest of this study, the only results used are those employing the KDE background models.

In addition to fitting each photodiode-amplitude segment of the laser data sep-

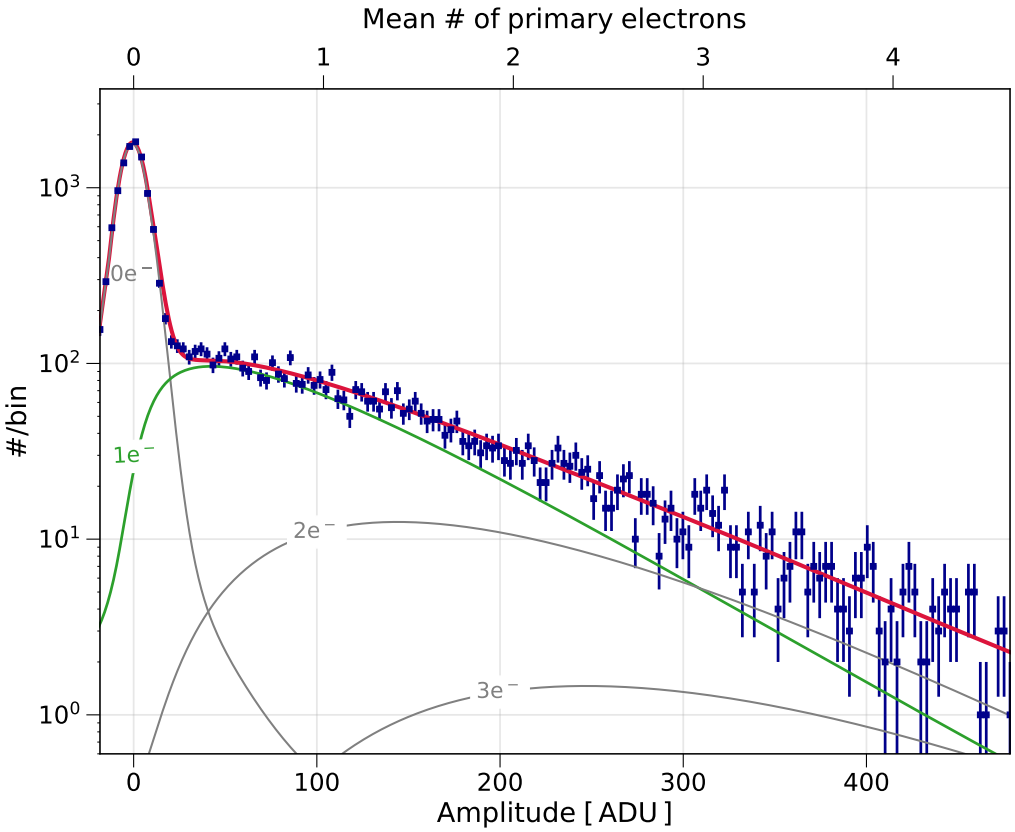


Figure 3.31: The laser amplitude spectrum from the 1300 V measurement (one particular photodiode amplitude segment, blue histogram) along with the best-fit model shown in red. Components of this best-fit model for 0 to 3 electrons are indicated. Using the fit result, the amplitude scale is converted to units of average amplitude for a single primary electron (top scale). The p-value of this particular fit is 0.29.

	Gaussian Bkgd.	KDE Bkgd.
$\langle G \rangle$ [ADU]	97.9 ± 1.7	103.6 ± 1.5
θ	$0.45^{+0.09}_{-0.07}$	$0.64^{+0.07}_{-0.06}$

Table 3.2: Fit results for the 1300 V laser data (one particular photodiode amplitude segment, corresponding to Fig. 3.31), assuming a Gaussian background vs. the KDE background model.

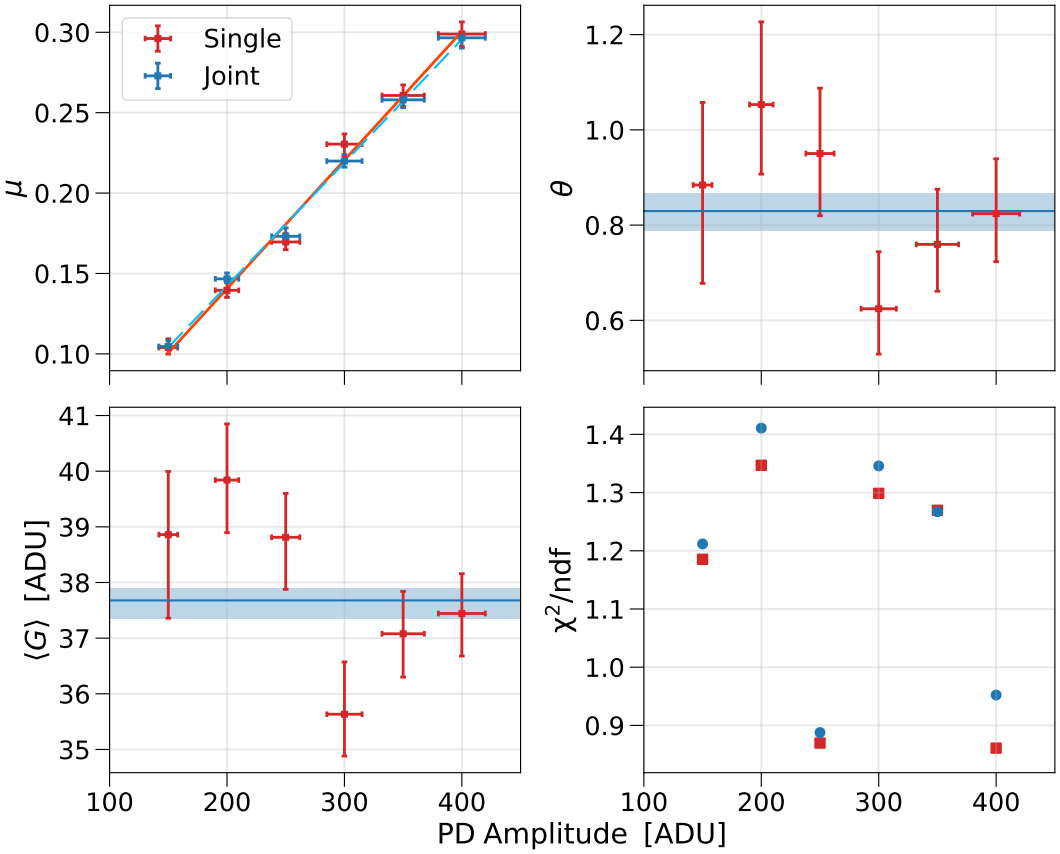


Figure 3.32: Laser calibration results for data taken at an anode voltage of 1200 V, with varying photodiode (PD) amplitude data selections. The single fit results for the mean number of electrons μ, $\langle G \rangle$, and θ are shown for each photodiode segment (red markers), as well as for the joint fit of all data segments (blue markers/bands). In all cases, the best-fit and 1 σ statistical uncertainties are shown. For the μ results, linear fits of the single and joint values are also shown. The lower right panel gives χ^2/ndf values for the single and joint fit of each data segment.

arately as in Fig. 3.31, a joint fit of all data segments is performed for each laser calibration [192]. For this joint fit, common values of $\langle G \rangle$ and θ are used to describe all the data, with independent values of μ for each segment. This is done to confirm that choosing laser data with different intensities does not bias the estimates of $\langle G \rangle$ and θ, only affecting the mean number of electrons μ [192].

The joint and single fit results for the 1200 V laser data are shown in Fig. 3.32 as an example. The general agreement between the single and joint results for $\langle G \rangle$ and θ (and the lack of any noticeable trend as a function of photodiode amplitude)

Anode voltage [V]	$\langle G \rangle$ [ADU]	θ
1200	$37.83^{+0.07}_{-0.47}$	$0.84^{+0.02}_{-0.07}$
1250	$60.44^{+0.45}_{-0.23}$	$0.63^{+0.03}_{-0.02}$
1300	$97.27^{+0.42}_{-0.38}$	$0.60^{+0.02}_{-0.02}$

Table 3.3: Laser calibration joint-fit results combining data at different laser pulse intensities. As expected, the avalanche gain increases proportionally with anode voltage.

confirms that there is no bias introduced by the varying laser intensity of each dataset. This was found to be the case for the other datasets as well. As expected, μ varies linearly with photodiode amplitude in both the joint and single fits, both exhibiting a nearly identical trend. Finally, the χ^2/ndf values of the joint and single fits of each data set are shown to be comparable, confirming at least that the joint fit represents the data as well as the more flexible single fits. This procedure – single fits of each photodiode amplitude segment, a joint fit, and the examination of the results shown in Fig. 3.32– is repeated for every laser calibration. The joint fit results (with statistical uncertainty) of each are used for the rest of this study and are given in Table 3.3.

3.3.2 Low energy analysis

To fit the low energy ^{37}Ar amplitude spectra (which includes L_1 and K-shell electron capture events), the decay cascade simulation results described in section 3.2.2 must be used. A Geant4 simulation of the SPC used for these measurements was performed to that end, namely of a 30 cm SPC with 50 mbar of CH_4. This can be used to produce the 2D histogram of the number of decay particles vs. total deposited energy per events shown in Fig. 3.24. However, since the primary electron yield depends on the energies of all the individual particles in the decay cascade, the full information of decay particles and their individual deposited energy was retained from the MC.

Using eq. 3.5, the total mean number of primary electrons μ produced by a single

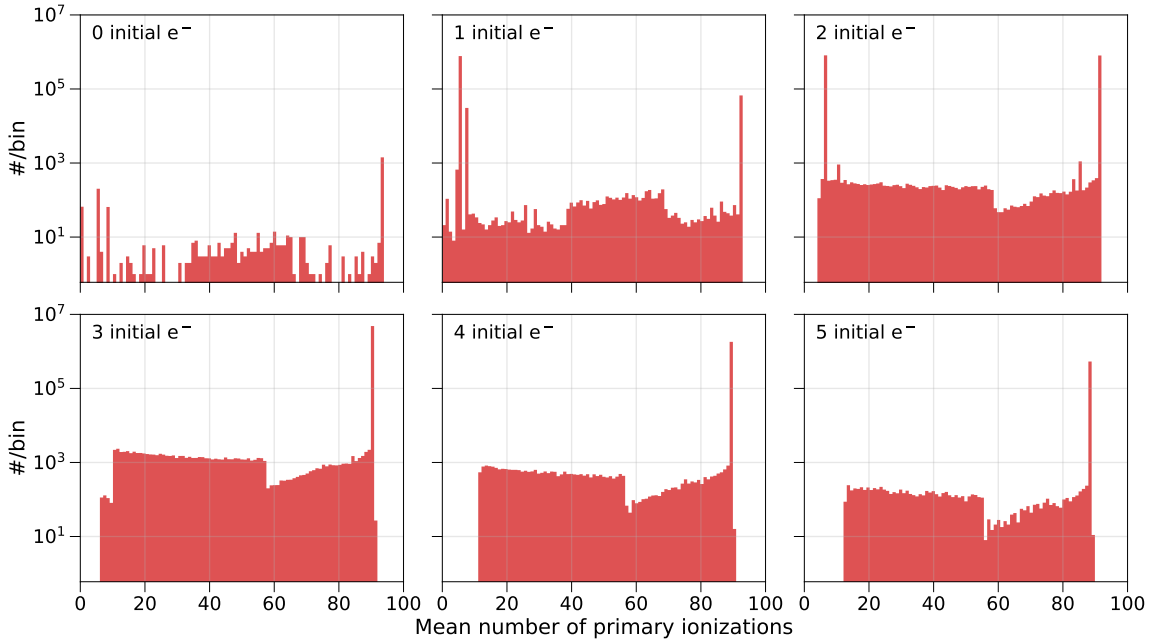


Figure 3.33: Distributions of the mean number of primary ionizations produced by ^{37}Ar decays in the S30 measurements, for different numbers of initial electrons. This is calculated assuming $W_0 = 30\,\text{eV}$ and $U = 15\,\text{eV}$.

event with N total decay products can be written as

$$\mu = \sum_{i}^{N} \frac{E_i}{W(E_i)} \tag{3.25}$$

which can then be expanded in terms of the W-value parameters W_0 and U using eq. 3.4:

$$\mu = \frac{1}{W_0}\left(\sum_{i}^{N} E_i - N \cdot U\right) \tag{3.26}$$

Calculating μ for all of the simulated ^{37}Ar events gives a probability distribution for the mean number of primary ionizations produced, $P(\mu)$. This can then be "smeared" with the COM-Poisson distribution (see section 3.1.1) to include dispersion in primary ionization, for a given Fano factor F. However, before this can be done, $P(\mu)$ must be separated into events with different numbers of initial, Auger electrons from the ^{37}Ar decays (from 0 to 5 in this case). These electrons contribute to the total primary ionization signal of the event, however, they are not subject to Fano factor

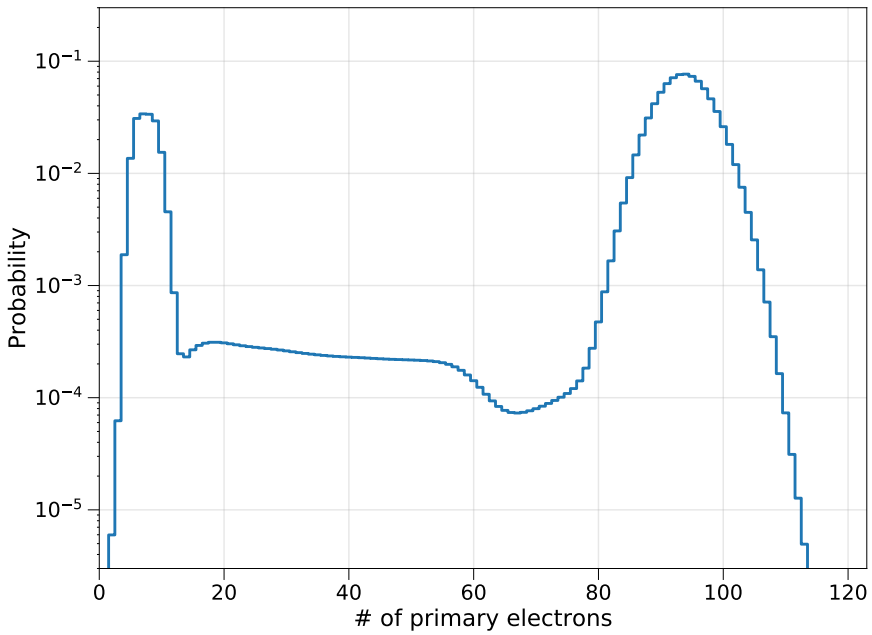


Figure 3.34: Distribution of primary electrons produced by ^{37}Ar, calculated for the S30 measurements with $W_0 = 30\,\text{eV}$, $U = 15\,\text{eV}$, and $F = 0.2$ as an example.

fluctuations.

The separate $P(\mu)$ for different numbers of initial electrons k is shown in Fig. 3.33 (with $W_0 = 30\,\text{eV}$ and $U = 15\,\text{eV}$ chosen arbitrarily for example). Each of these individual separate $P^{(k)}(\mu)$ distributions is first convolved with the COM-Poisson distribution, after which the appropriate number of initial electrons is added. Finally, all the separate distributions are added together to give the total probability distribution of i primary electrons from ^{37}Ar decays for given values of W_0, U, and F:

$$P(i) = \sum_{k=0}^{k=5} \left(\int P_{\text{COM}}(i-k|\mu, F) P^{(k)}(\mu) d\mu + k \right) \tag{3.27}$$

The total primary electron probability distribution is shown in Fig. 3.34. This is then combined with the avalanche response for i primary electrons, and summed over i to give the total amplitude E spectrum. This is also convolved with a baseline noise model $P_{\text{Noise}}(E)$. The KDE-modelled backgrounds used in section 3.3.1 are not applicable because, in the case of relatively large ^{37}Ar events, the DAQ software

will preferentially center the larger of two coincident pulses, changing the behaviour of coincident event reconstruction. However, the measured background spectrum (such as the example shown in Fig. 3.28) does demonstrate that the baseline noise is Gaussian to first order. Therefore, Gaussian distributions are fit to the peaks of these measured background spectra to define $P_{\mathrm{Noise}}(E)$.

In all of the W-value measurements, there exists a large background of low-energy events, including single-electron pulses as well as non-physical spike events. Rather than attempting to suppress these backgrounds and incur the issue of cut acceptance, they are included in the fit and modelled with a generic exponential distribution $B_1 e^{-B_2 E}$. This closely mimics the single electron distribution in an SPC if $\theta \approx 0$, as well as non-physical events. Additionally, a flat background rate B_{Flat} is included in the model. Taking all of the above into consideration, the overall model $\mathcal{P}(E)$ for the low energy ^{37}Ar calibration data is given as:

$$\begin{aligned}\mathcal{P}\left(E\right) = &\left[\sum_{i}\left(\sum_{k=0}^{k=5}\left(\int P_{\mathrm{COM}}(i-k|\mu,F)P^{(k)}(\mu)d\mu + k\right)\right) \times P^{(\mathrm{i})}_{\mathrm{Polya}}(E|\left\langle G\right\rangle,\theta)\right] \\ &\otimes P_{\mathrm{Noise}}\left(E\right) + B_{\mathrm{Flat}} + B_1 e^{-B_2 E}\end{aligned} \tag{3.28}$$

Note that this model is truncated at $i = 30$ primary electrons for the purpose of fitting the low energy ^{37}Ar events alone. The Polya distribution shape parameter θ was fixed to the best-fit value obtained from the laser calibration results. Leaving this as a free parameter proved to be very computationally expensive, requiring approximately 1 CPU $\cdot$ hour to compute the N^{th} convolution of the Polya distribution for all necessary values of N, for each new value of θ. While this may have increased statistical uncertainty on the Fano factor F in this study (which might be varied more to match the dispersion of the calibration data), F is not the main parameter of interest. Indeed, F is already known to incorporate other resolution broadening effects (i.e. from pulse-processing) beyond the true dispersion of primary ionization in SPCs [192]. Further, the $\sim 5\,\%$ statistical uncertainty on θ (see Table 3.3) translates

to an even smaller uncertainty on the dispersion of avalanche fluctuations f, with is given by [192]:

$$f = \frac{1}{1+\theta} \tag{3.29}$$

For example, using eq. 3.29, a 5 % change in $\theta = 0.6$ results in a 1.8 % change in f.

A binned likelihood function was constructed to fit this model to the W-value data (with c_i counts in bin i). The detector gain $\langle G \rangle$ was left as a free parameter in this model but with a constraint term added to the likelihood function ($\log \mathcal{L}_{\langle G \rangle}$), incorporating the laser calibration result and statistical uncertainty:

$$\log \mathcal{L} = c_i \log\left(\frac{\nu_i}{c_i}\right) - \nu_i + c_i + \log \mathcal{L}_{\langle G \rangle} + \log \mathcal{L}_{W_0-\text{High E.}} \tag{3.30}$$

where ν_i is the expected counts per bin obtained by integrating eq. 3.28, $\nu_i = \int_{\text{bin}_i} \mathcal{P}(E)dE$. The constraint for $\langle G \rangle$ was calculated using the MCMC samples of the laser calibration fit; by taking the boundary of all the MCMC samples, the profile likelihood is obtained, as shown in Fig. 3.35. An additional constraint term on W_0 was added to eq. 3.30 using the result of the high-energy W-value fit for each dataset, which was performed first (described in the next section).

For this study, the model was fit to the data using a custom MCMC approach – the "fast burn-in" method – which is described in Appendix A. An example of one fit result (for the data gathered at 1250 V) is shown in Fig. 3.36. The results for all three datasets are summarized in Table 3.4.

3.3.3 High energy analysis

The high-energy analysis was performed on both the (predominantly) 2.83 keV ^{37}Ar peak, as well as the somewhat overlapping 1.486 keV aluminum fluorescence peak [198]. The amplitude spectrum of the ^{37}Ar events was modelled in the same manner described in section 3.3.2, using the COM-Poisson distribution to characterize primary ionization statistics, applied to the decay cascade simulation presented in section

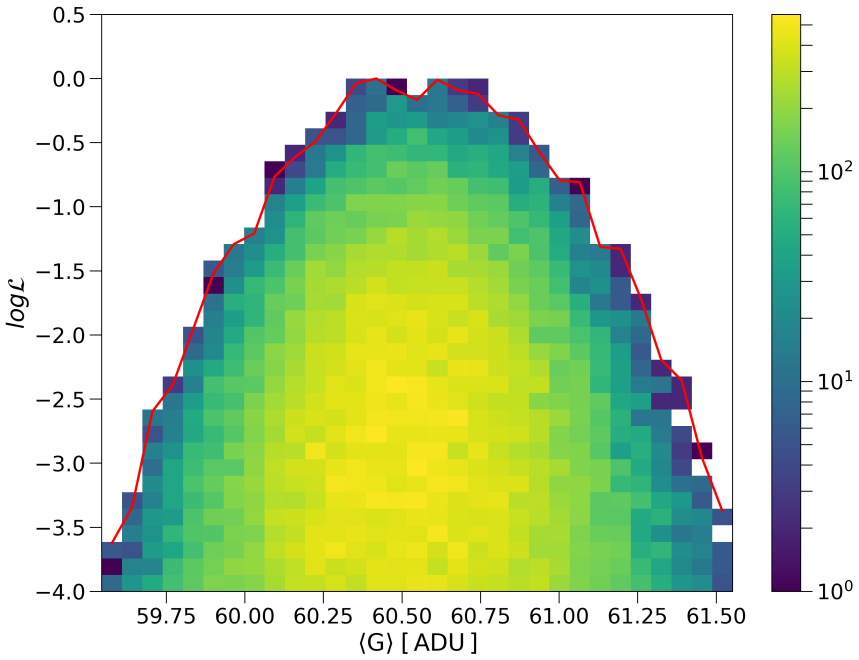


Figure 3.35: MCMC samples from the fit of the 1250 V laser calibration data, projected onto the parameter $\langle G \rangle$ vs. the log-likelihood (2D histogram with the colour-scale representing counts per bin). The red curve is therefore the profile log-likelihood function for $\langle G \rangle$.

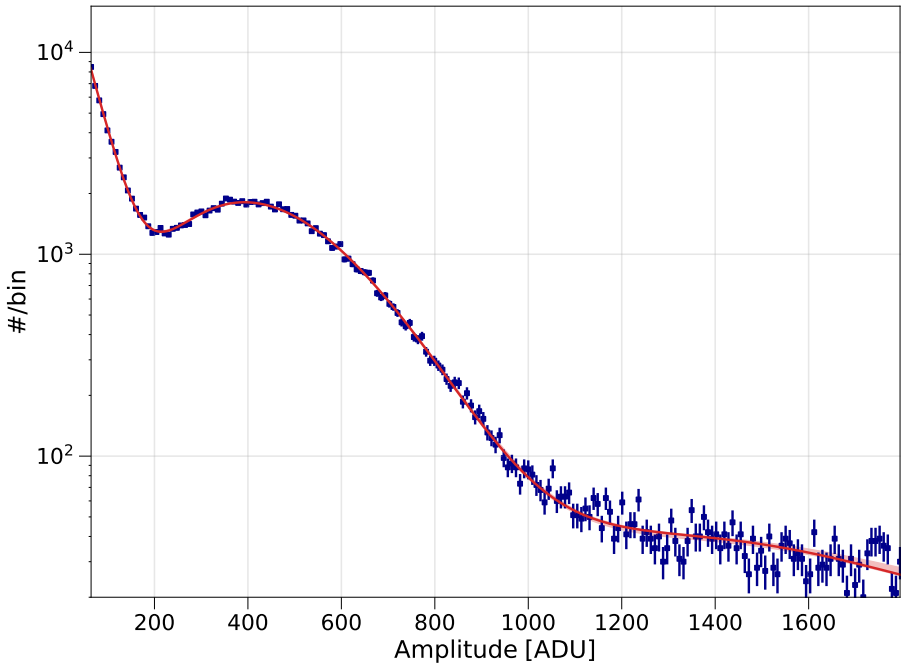


Figure 3.36: Low energy ^{37}Ar calibration data taken at 1250 V (blue histogram) compared to the best-fit model with $1\,\sigma$ statistical uncertainty band (red curve and shaded band). The p-value for this particular fit is 0.14.

Anode HV [V]	W_0 [eV]	U [eV]	F
1200	$32.1^{+0.7}_{-1.7}$	$11.7^{+2.7}_{-1.5}$	$0.51^{+0.03}_{-0.04}$
1250	$31.1^{+0.9}_{-0.5}$	$14.6^{+0.9}_{-1.4}$	$0.43^{+0.01}_{-0.02}$
1300	$33.1^{+0.1}_{-1.5}$	$11.7^{+2.4}_{-0.4}$	$0.61^{+0.01}_{-0.03}$

Table 3.4: Fit results of the low-energy spectrum at all anode voltages.

3.2.2. However, in the regime of this data, the evolution of the W-value with energy is negligible, so $W(E)$ was replaced with a single parameter W (equivalent to the high-energy limiting value W_0). Therefore the $N \cdot U$ term in eq. 3.26 was neglected.

Aluminum fluorescence produces monoenergetic, 1.486 keV X-rays. It was included in the model with an independent Fano factor $F_{\text{Alum.}}$ and W-value, giving a mean number of primary electrons:

$$\mu_{\text{Alum.}} = \frac{1486\,\text{eV}}{W_{\text{Alum.}}} \tag{3.31}$$

The relative branching ratio of this peak compared to the high-energy ^{37}Ar peak is denoted R. The overall model – building off of eq. 3.28 – is given as (without any background components):

$$\begin{aligned}\mathcal{P}(E) = \Bigg[\sum_i \Bigg(\sum_{k=0}^{k=5} \int P_{\text{COM}}(i-k|\mu,F) P^{(k)}(\mu) d\mu + k \\ + R \times P_{\text{COM}}\left(i|\mu_{\text{Alum.}}, F_{\text{Alum.}}\right)\Bigg) \times P^{(\text{i})}_{\text{Polya}}(E|\left\langle G\right\rangle, \theta)\Bigg] \\ \otimes P_{\text{Noise}}(E)\end{aligned} \tag{3.32}$$

Characteristically for $\mathcal{O}$(keV) peaks in SPC data, the flat background of the amplitude spectrum is slightly higher for amplitudes below the peaks than above. This is likely due to a combination of Compton scattering events depositing less than the full energy of the interacting photon [70] (if applicable), as well as misreconstructed

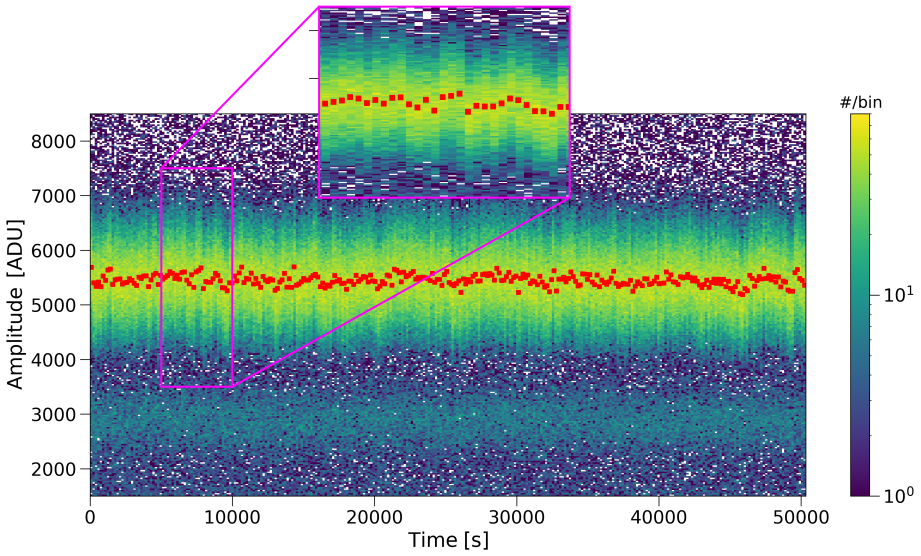


Figure 3.37: 2D histogram of amplitude data vs. time, specifically for the high-energy 1250 V dataset. The red markers represent the mode amplitude of the high-energy ^{37}Ar peak in segments of time.

events. To continuously connect the separate flat background components below and above the calibration peaks (B_1 and B_2 respectively), the function model $\mathcal{P}(E)$ itself is used. Specifically, the inverse CDF of the model eq. 3.32 connects the two flat background components to complete the background model, which is then added to eq. 3.32:

$$B(E) = B_1 \times \left[1 - \int^{E} \mathcal{P}(E')dE'\right] + B_2 \tag{3.33}$$

One potential systematic that can be addressed in the high-energy calibration data is sensor discharges when operating at relatively high voltages. This can lead to small, temporary changes in the gain of the detector. Indeed, plotting the amplitude data vs. time (such as in Fig. 3.37) reveals small fluctuations of a few percent, on time scales of thousands of seconds. A Lomb-Scargle periodogram of this time series data reveals that the highest relative magnitude of the fluctuations occurs for periods of approximately 500 to 5000 s [217, 218]. This effect causes extra dispersion in the amplitude data, leading to an overestimation of F.

To account for this as a systematic, these fluctuations were corrected in the data. This was done by dividing the data into $\sim 170\,\mathrm{s}$ time bins, and calculating the mode of the high-energy ^{37}Ar peak in each, as shown in Fig. 3.37. The amplitude in each bin was then translated so that the mode in that bin matched that of the first time bin (an arbitrary choice). In this way, the dispersion caused by gain fluctuations was removed. After this "corrected" amplitude data is fit, the maximum relative difference in mode amplitude between all of the time bins was propagated to give a range of corresponding values of W to report for the final result, using the proportionality between W and amplitude E:

$$W \propto \frac{\langle G \rangle}{E} \tag{3.34}$$

This is done to remain agnostic as to what the "true", un-distorted value of the gain is.

Examining the trend of the amplitude data over time is also useful to guard against another potential source of systematic uncertainty in this study – charge trapping. Increasing oxygen and water contamination in the gas can cause gradual decreases in the detector gain over time (see section 3.1.3). However, no such effect was observed in these datasets within the risetime range considered for analysis. Another way in which attachment may manifest is to cause an anti-correlation between risetime and amplitude. As shown in Fig. 3.38 for one dataset, this effect was also not observed. However, there is a slight apparent increase in amplitude at the lowest risetimes. This artifact is typical of all SPC data and may be caused by the pulse treatment algorithm, or possibly evidence of non-zero charge attachment for all events except those originating immediately next to the sensor. It has also been hypothesized that this is evidence of a type of "space-charge" effect; high-rate conditions in the sphere may produce a large enough flux of ions to significantly distort the electric field experienced by subsequent, incoming electrons. Such effects are more clearly observed (albeit not fully understood) when operating at higher voltage. Higher pressure/volt-

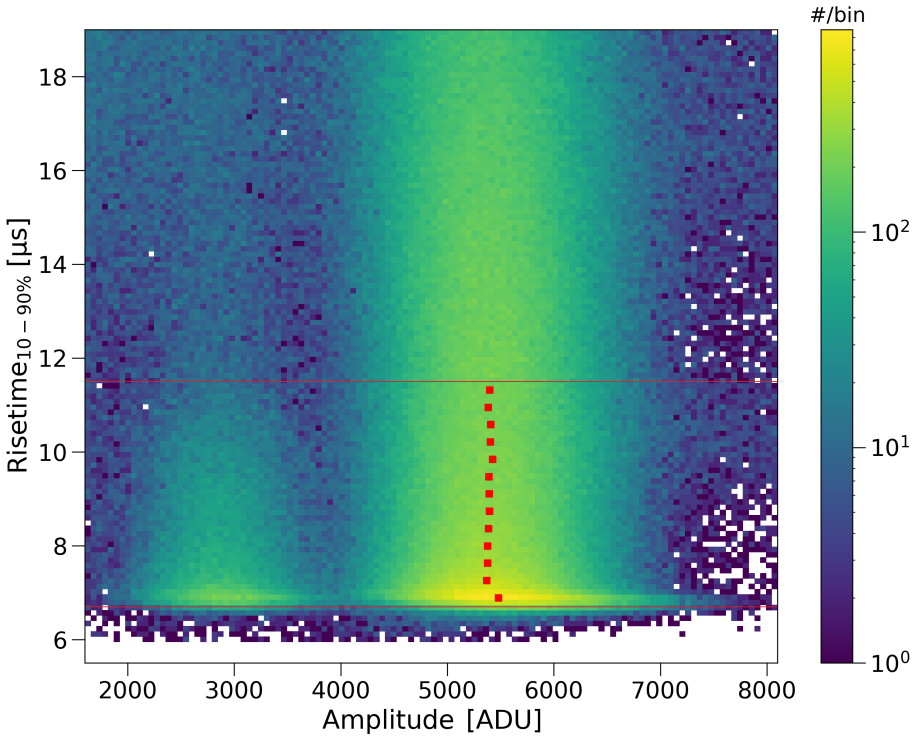


Figure 3.38: A 2D histogram of risetime vs. amplitude data, specifically for the high-energy 1250 V dataset. The red markers show the mode amplitude of the high-energy ^{37}Ar peak in different slices of risetime. The red lines indicate the minimum and maximum risetime cuts applied to select data for this analysis.

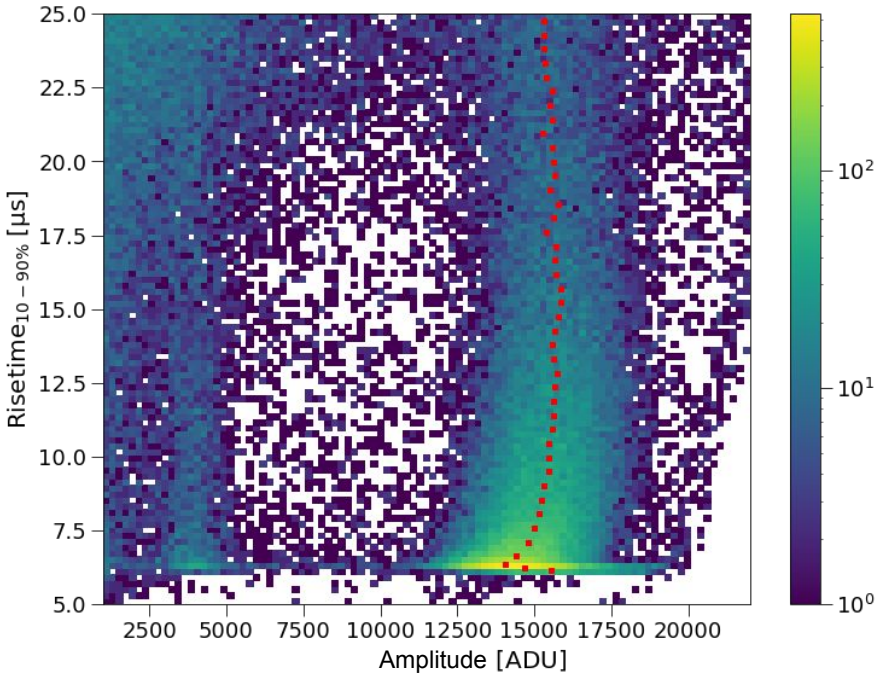


Figure 3.39: 2D histogram (with the colour scale giving counts per bin) of risetime vs. amplitude data, taken with 130 mbar of methane gas, with an anode voltage of 1800 V. The evident peak is 5.9 keV X-rays from ^{55}Fe. The red markers denote the mean amplitude in different slices of risetime.

age data was taken in methane gas for this study but was deemed unusable because of the severity of this space-charge distortion. An example of this data (130 mbar, 1800 V) is shown in Fig. 3.39.

As this effect is not fully understood, a further correction was applied to the data (after already applying the time-based correction described above). The mode amplitude of the high-energy ^{37}Ar data was found in slices of risetime as shown in Fig. 3.38, within the risetime range considered for this analysis. The amplitude data was then translated so that the mode in each slice matched the overall average of all slices. As before, these translations were then propagated to give a range of values of W for the final result after fitting the corrected data, so as to not make any assumptions about the nature of this artifact. The corrected amplitude data is shown compared to the original spectrum in Fig. 3.40; the operations described above did not amount to a significant change in amplitude.

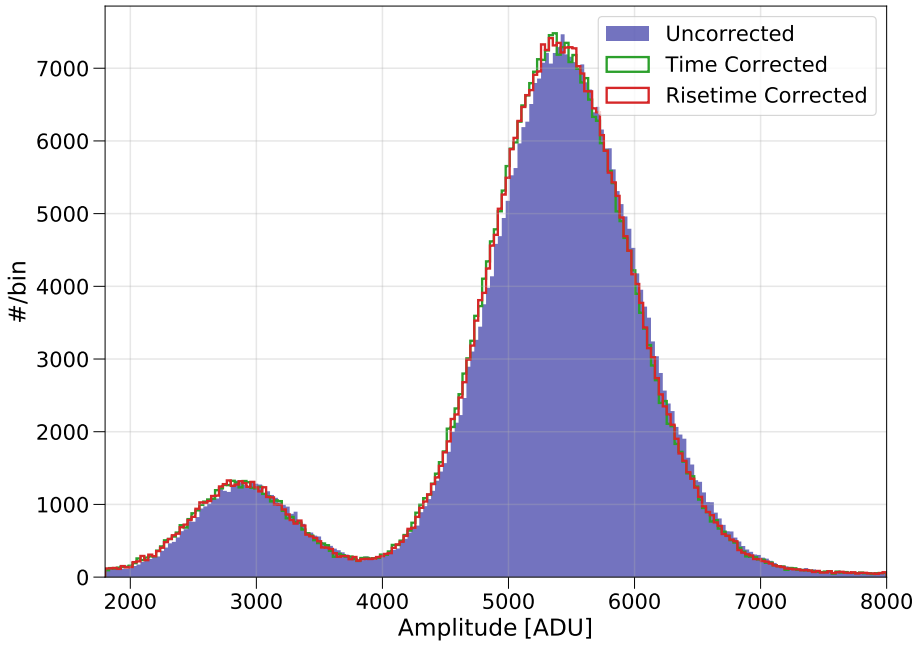


Figure 3.40: The amplitude spectrum of the original high-energy calibration data (blue histogram), compared to the data with time-based corrections applied (green histogram), as well as data with both time and risetime based corrections applied (red histogram).

As with the low-energy analysis, a log-likelihood function was constructed to fit the model to the data, including a constraint term for the detector $\langle G \rangle$ based on the UV laser calibrations (see eq. 3.30). This was optimized using the custom fast burn-in MCMC method (see Appendix A) to give fit results for the corrected data. An example of such a fit is shown in Fig. 3.41. The corrections described above were then propagated to give additional systematic uncertainty on the W-value results of the fits; note that these are reported as ranges with equal probability. The final results of this study, including systematic uncertainties, are given in Table 3.5.

3.3.4 Results and discussion

The key results of this study – $W(E)$ for both the high and low energy analyses – are presented in Fig. 3.42. By using multiple calibration energies, these results span an energy range of approximately 50 eV to 3 keV. There is no clear trend of $W(E)$ vs. anode voltage as expected, and the separate high and low energy results are consistent

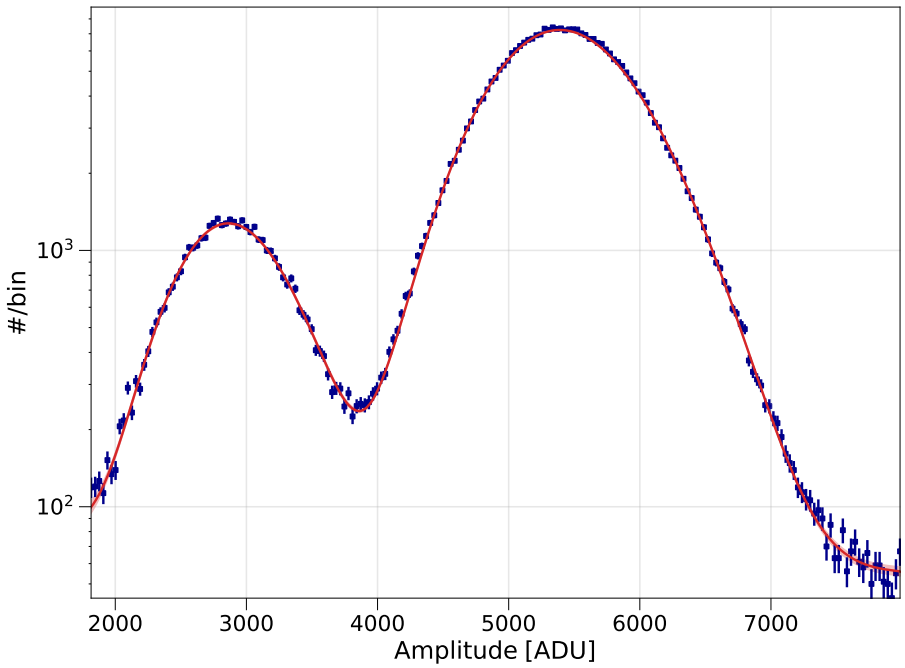


Figure 3.41: High energy ^{37}Ar and aluminum fluorescence calibration data taken at 1250 V (blue histogram) with the best-fit model with 1 σ statistical uncertainty band (red curve and shaded band). The p-value for this particular fit is 0.15.

Anode voltage [V]	Peak	W_0 [eV]	F
1200	Alum.	$(29.6 \rightarrow 35.4)^{+0.2}_{-0.3}$	$0.33^{+0.01}_{-0.01}$
1200	^{37}Ar	$(31.5 \rightarrow 35.1)^{+0.2}_{-0.3}$	$0.39^{+0.01}_{-0.01}$
1250	Alum.	$(28.9 \rightarrow 32.4)^{+0.2}_{-0.1}$	$0.33^{+0.02}_{-0.02}$
1250	^{37}Ar	$(29.4 \rightarrow 31.9)^{+0.2}_{-0.1}$	$0.30^{+0.01}_{-0.01}$
1300	Alum.	$(28.2 \rightarrow 33.3)^{+0.2}_{-0.1}$	$0.49^{+0.02}_{-0.02}$
1300	^{37}Ar	$(29.7 \rightarrow 35.3)^{+0.2}_{-0.2}$	$0.43^{+0.01}_{-0.01}$

Table 3.5: Fit results of the high-energy peaks at all anode voltages. The W-value results are shown given with ranges representing the systematic uncertainty of those results, with additional statistical uncertainty.

with each other.

When compared to results from the literature – particularly refs. [208, 209] – it is apparent that there is a $\sim$ 10% discrepancy (although not with all historical results [166, 206]). Specifically, the most probable W-value at 2.8 keV obtained in this study is 12% higher than the combined results of refs. [208, 209]. A more precise quantification of this disagreement is not possible because the results of refs. [208, 209] do not include any uncertainty.

Although several systematic effects were accounted for in this study (charge trapping, sensor discharges), it is possible that these effects are underestimated. Other phenomena that are not fully understood in SPCs (such as space charge effects) may have also impacted these results. Therefore these measurements may be considered "effective W-value" results, rather than measurements of the intrinsic W-value quantity of methane gas. Regardless, the results of this study do reflect the true ionization yield of methane gas in SPCs, which is what must be known to characterize the S140 experiment at the LSM. Future W-value measurements are planned with SPCs (with some currently ongoing at the University of Alberta) to further investigate this discrepancy. In particular, the dependence of the W-value on gas pressure (or lack thereof, theoretically [166]) is being pursued.

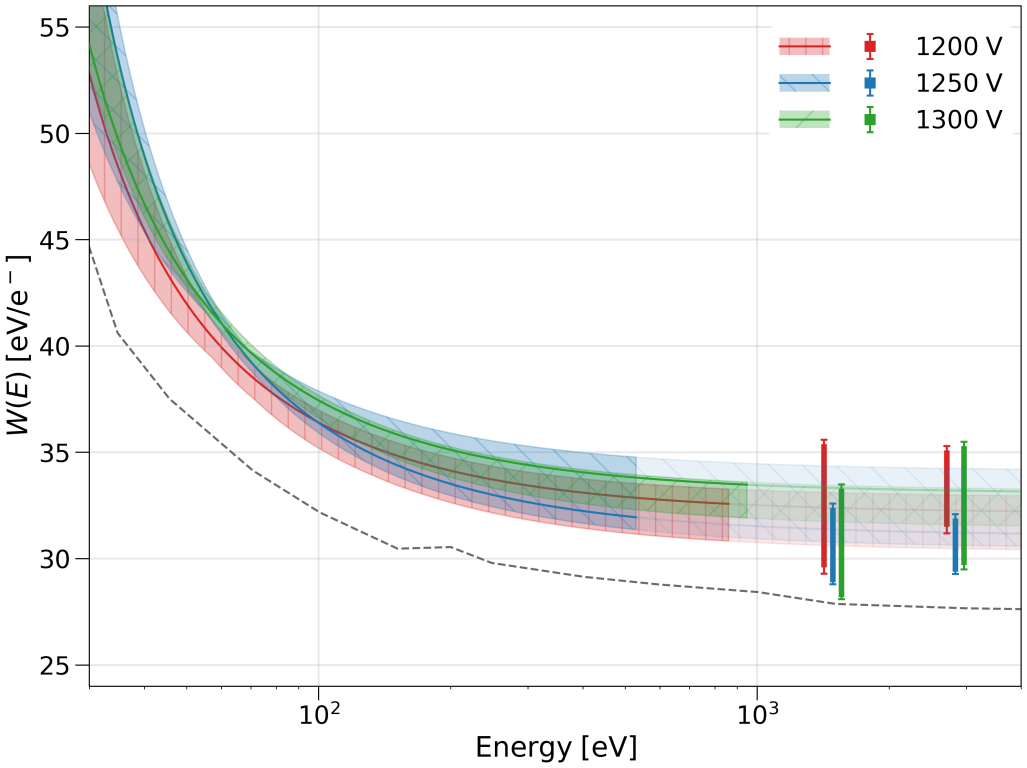


Figure 3.42: $W(E)$ results for all three calibration data sets (at different anode voltages). The low energy results are shown as best-fit curves and shaded bands representing the $1\,\sigma$ uncertainty bands, terminating at the end of their fitting ranges (and then continuing on more faintly). The high energy results (for ^{37}Ar and aluminum fluorescence) are represented with long markers for the systematic range of each measurement, with additional error bars for statistical uncertainty. The black dashed curve is the combined result of refs. [208, 209].

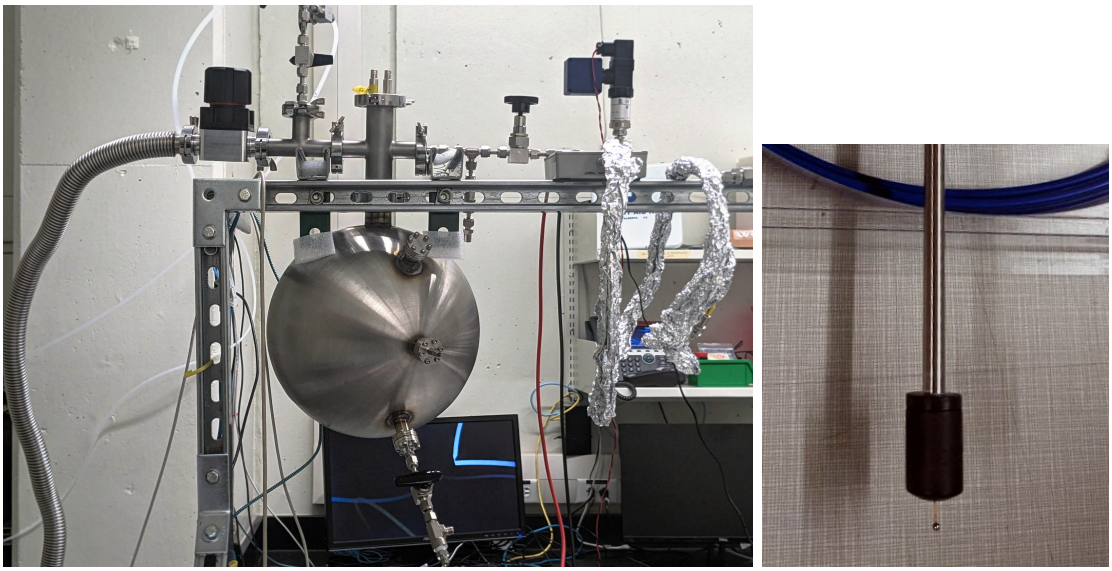


Figure 3.43: The U of A SPC (left) on its support structure, and the 2 mm sensor with Bakelite umbrella (right).

3.4 SPC at the University of Alberta

An R&D-scale, 30 cm SPC has been constructed and operated at the University of Alberta. Similarly to the SPC used for the measurements described in section 3.3, having a small-scale experiment in a local lab offers easy access, and the opportunity to develop hardware and analysis techniques in a way that is not possible with the full-scale S140 device. Since becoming fully operational in spring 2023, this SPC has already been used to perform W-value measurements (which are still ongoing), utilized in the development of a Laser Absorption Spectroscopy system (LAS) to measure the concentration of methane gas in the SPC in real-time [187]. Finally, the device has been used in ongoing experiments to quantify the performance of novel radon trap substrates and operating methods.

3.4.1 Detector setup

The U of A SPC – pictured in Fig. 3.43 – consists of a 30 cm diameter stainless steel vessel, which may be operated with pressures up to 1 bar (it is not high-pressure rated). Gas flow and the sensor insertion/support rod are both handled by a single

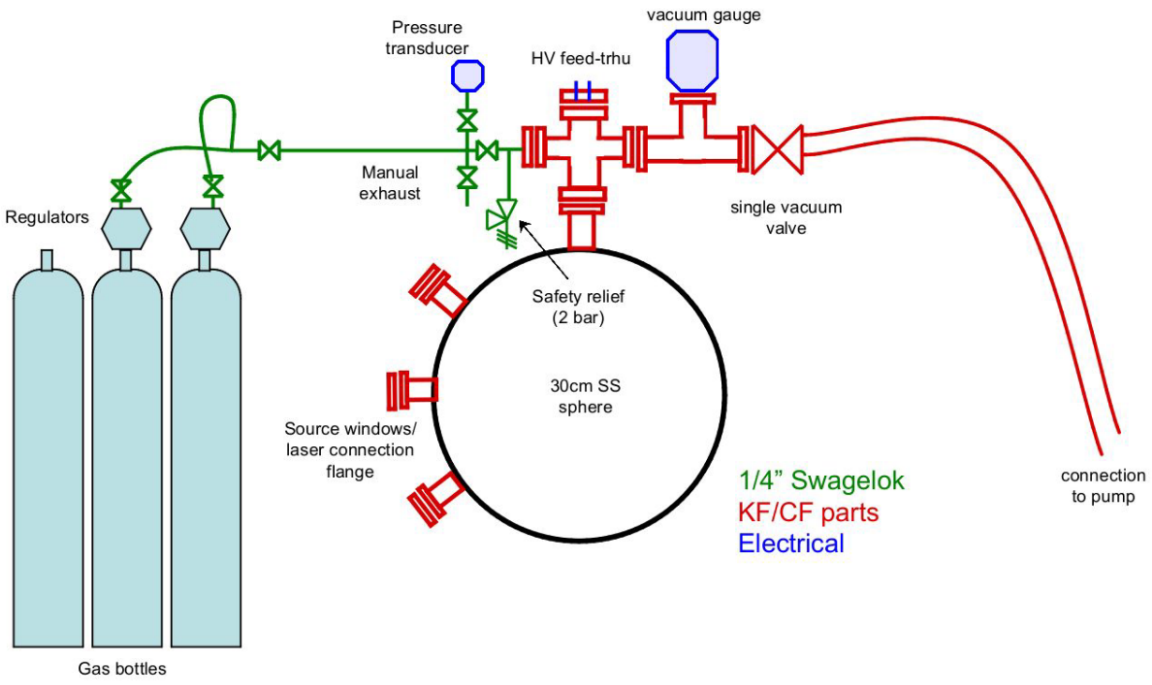


Figure 3.44: Schematic of the basic U of A SPC setup, with gas and vacuum plumbing identified.

40 mm port on the top of the sphere. A series of internal support disks with set screws allow for movement and alignment of the sensor support rod. Currently, a traditional 2 mm stainless steel anode is used, with a Bakelite umbrella (also pictured in Fig. 3.43). Three additional ports on the side of the SPC allow access for calibration, including one with a fiber optic cable feed-through for the use of a UV laser system (as described in section 3.2.1). An Agilent turbo-pump setup is used to evacuate the sphere. Currently argon and methane gas are available for use in the SPC (all 5N quality). The general plumbing P&ID diagram of the SPC by itself is shown in Fig. 3.44.

A CAEN desktop HV supply is used in this setup [219], with a CREMAT CR-110 charge-sensitive preamplifier [220]. This particular pre-amplifier has a time constant of 140 μs, and a gain of 1.4 V/pC. For a 1 keV event for example, with an avalanche gain of $\mathcal{O}(10^3)$, this would result in a signal on the order of 10 mV. The digitizer board used in this case is a Red Pitaya, with the custom NEWS-G DAQ software

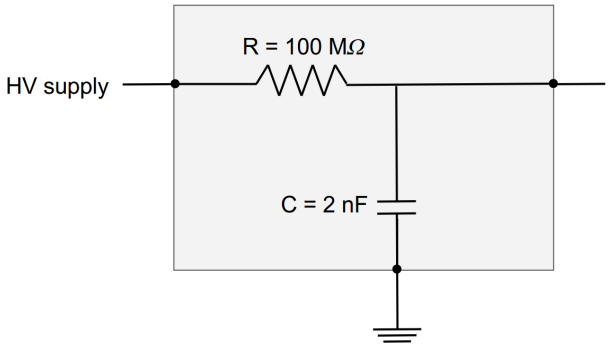


Figure 3.45: Schematic of a low-pass filter used for the U of A SPC.

Samba installed (see section 3.1). A homemade low-pass filter – depicted in Fig. 3.45 – was added after the HV supply to dampen noise from it. While there is no clear experimental evidence of it, it is possible that this filter causes a significant voltage drop on the anode that may take some time to recover. If these voltage drops following a particle event coincided with the avalanche of a subsequent event, this would reduce the avalanche gain. Further analysis of this circuit is warranted (and being undertaken) to understand this potential issue. The entire setup is connected to the electrical ground of the building (through the metal frame the sphere rests on). All of the coaxial cables are shielded with braided steel cable to reduce induced noise.

A radon trap setup – and radon source intended for testing the former – are described in the following subsection. Accurate, real-time measurement of the composition of gas inside the sphere (particularly the amount of methane) is crucial (specifically the concentration of methane); to that end, a Binary Gas Analyzer (BGA) is included in the setup [221]. This device exploits the speed of sound in two-gas mixtures to measure concentration [187].

However, this device is limited in that it can only perform in mixes of two gases, and is highly dependent on temperature and pressure. To obtain more robust measurements of methane concentration, a Laser Absorption Spectroscopy (LAS) system has been constructed [187, 222]. This device – pictured in Fig. 3.46 – consists of an infrared laser with a tunable wavelength, shone through a tube of gas. Wavelengths

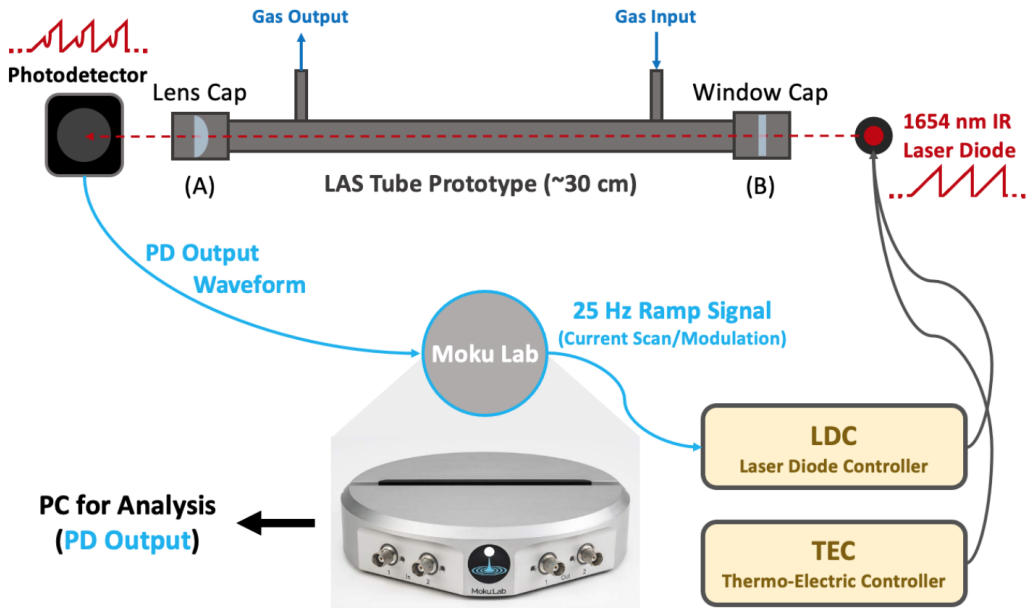


Figure 3.46: The laser absorption spectroscopy (LAS) system designed to measure methane concentrations in the U of A SPC. This includes the infrared laser, photodiode, and electronics. Taken from figure 6.11 of ref. [187].

corresponding to the absorption lines of the gas in question are scattered, reducing the intensity of light measured by a photodiode at the other end of the gas column [223]. Similar setups with reflecting/integrating spheres are sometimes used to measure very low concentrations of gases, however for % concentrations, a linear tube of approximately 10 cm is sufficient [222]. The amplitude of the absorption peak is proportional to the concentration. This allows for real-time measurements of methane concentrations in the U of A SPC [187]. A Moku:Lab device is used to drive the laser wavelength scan, as well as to readout the signal from the photodiode [224].

3.4.2 Radon trap measurements

^{222}Rn is a common gaseous radioisotope that is problematic for many rare-event searches, including NEWS-G. It decays via 5.49 MeV α emission, with a half-life of 3.82 days [225]. Additionally, its subsequent decay chain includes two other α emitters (^{218}Po and ^{214}Po [226, 227]) before terminating in the production of ^{210}Pb, itself a prominent background source in SPCs. This decay chain is shown in Fig. 3.47. Steps

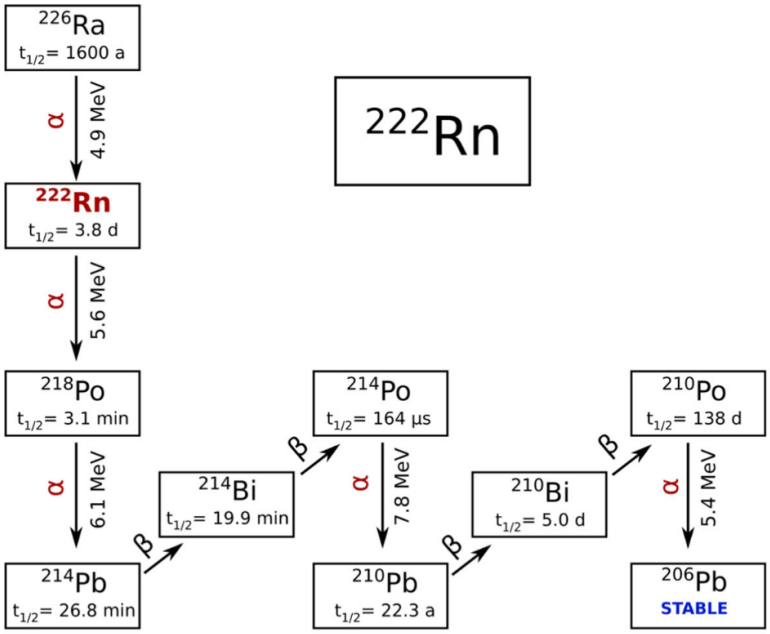


Figure 3.47: The (simplified) decay chain of ^{222}Rn, taken from figure 1 of ref. [229].

are taken for many rare-event searches to reduce the amount of radon introduced into the experiment, such as storing devices in sealed, pure gas environments, see section 4.1. Detector vessels may also be etched to remove radon deposition, such as with the DEAP-3600 experiment [228]. Still, radon persists as a major background source in many experiments. In the case of NEWS-G, many of the gas filters used to remove oxygen and water – such as the "getter" used for the S140 experiment – emit radon themselves.

To counter this problem, radon traps are often employed [230–232]. These consist of some form of substrate through which the detector medium is filtered to remove radon. As radon is chemically inert, these devices operate on the principle of adhesion of radon to the trap material through Van der Waals forces [233]. While this adsorption is not permanent, if it delays the passage of radon for long enough, it will decay inside the trap before being re-emitted [221]. Typical traps involved the use of carbon-based substrates (such as activated, coconut charcoal) in filters cooled down to approximately dry-ice temperatures (-78°C) [233, 234]. While effective at

cold operating temperatures, these traps can be inconvenient to operate in underground laboratories. A further problem for NEWS-G is the unfortunately similar size of Radon atoms and CH_4 molecules. This causes methane gas to be absorbed by the substrate as well. This can change the ionization properties of gas in the SPC, as well as the amount of hydrogen target material that is sensitive to dark matter interactions.

These drawbacks of typical carbon-based radon traps have spurred interest in alternative filter substrates. A particularly interesting prospect studied at the University of Alberta is Silver Zeolite (SZ), produced by the company Extraordinary Absorbents (based in Edmonton) [235]. This material (also known as Ag-ETS-10) consists of a porous, titanium-based mineral with incorporated silver ions. Such materials are known to function well as radon trap substrates [236]. Measurements of the properties of this specific substrate (i.e. performance at different operating temperatures) have been carried out with a Rad7 device [221], showing that at room temperature, SZ can remove orders of magnitude more radon than activated charcoal (more on this below).

However, these experiments were adversely affected by leaks in the Rad7 setup, leading to anomalous and unpredictable behaviour of radon activity over time in the experiment. SPCs can also serve as α counters, and so it was deemed beneficial to measure the performance of a room-temperature, SZ radon trap with the SPC at the U of A. The detector was connected in line with the radon source and trap for these measurements, with entirely Swagelok-type plumbing connections that leak less than the barb-fitting connecting of Rad7 devices. The setup for this experiment included an enclosed uranium ore source of radon (which diffuses out of the ore), and a SZ radon trap in a 1/2-inch, $\sim 20\,\text{cm}$ long stainless steel tube. Note that this radon trap can be operated at room temperature or dry-ice temperatures [221]. A schematic of the setup is shown in Fig. 3.48. Before the test, the radon trap was "purged" of any remaining radon/gases by heating the trap to $160 - 200°\text{C}$ while alternating

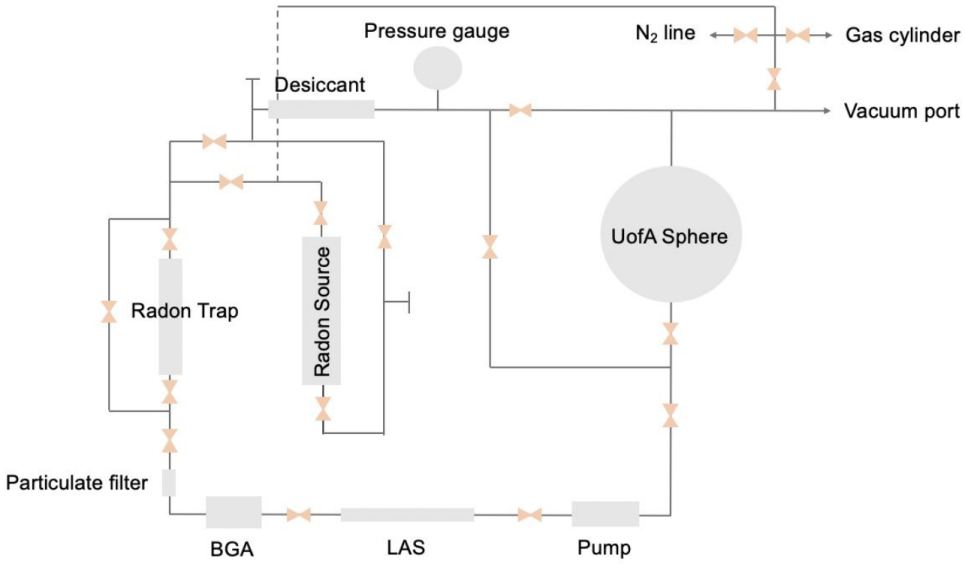


Figure 3.48: P&ID schematic of the U of A SPC setup for radon trap measurements. This includes the SPC itself, radon source and trap used, as well as the LAS system and binary gas analyzer (BGA) to measure gas composition. When circulating gas through the SPC and radon trap, the flow is in the counter clockwise direction.

between a nitrogen flush and vacuum. Following this, the entire system was evacuated for several days (to reach pressures $\mathcal{O}(10^{-6}\,\text{mbar})$), and then filled with 200 mbar of $\text{Ar} + 3\%\,\text{CH}_4$ gas. The SPC was operated at an anode voltage of 800 V.

First, background data was collected for approximately one day, consisting of primarily atmospheric muons, although these are at much lower energies than the α decays of interest in this study. Additionally, there was a constant 0.201 ± 0.002Hz background of ^{210}Po α decays (5.3 MeV) [237], originating from depositions of the radioisotope on the inner surface of the detector. Following the background measurement, radon was injected into the sphere until a total activity of $\sim 70 - 90$ Hz was achieved (the activity of radon itself was only ~ 60 Hz). This process took several hours to allow for diffusion into the SPC. The decay of radon was then observed for approximately two days to robustly measure the trend of activity. Finally, the radon trap was opened, and a small diaphragm pump was used to circulate gas through the system. Following the opening of the trap, the α rate was observed to drop precipi-

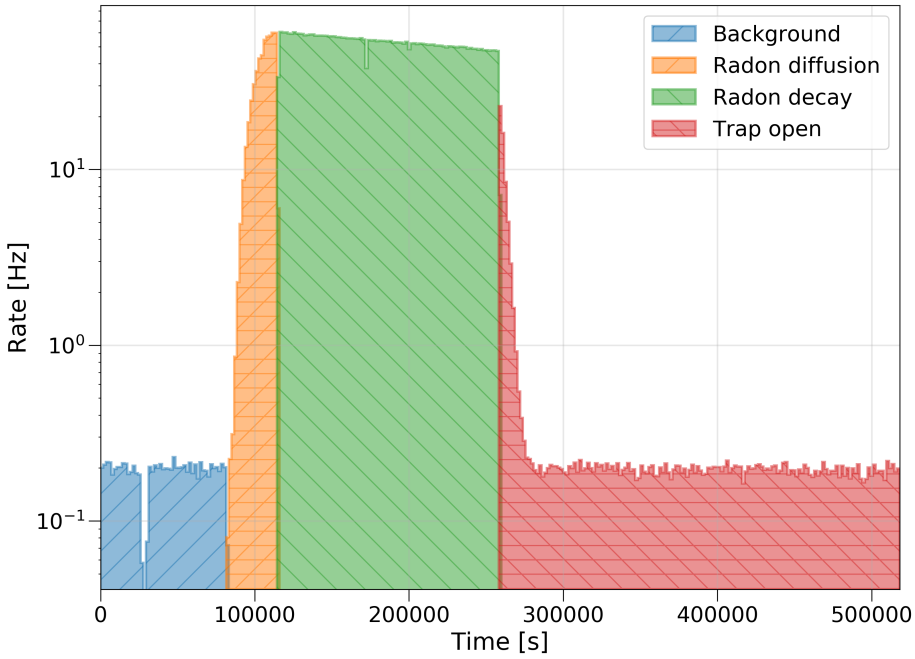


Figure 3.49: Event rate vs. time during the SZ radon trap test in the U of A SPC, showing the different phases of the experiment; background data, radon injection and decay monitoring, after which the trap is opened. PSD cuts and muon cuts are applied to this data. Note that there are some visible gaps in the data due to brief pauses in data collection.

tously over several hours, reaching the pre-radon background level again. Data was taken for 2.5 days after opening the trap to accurately measure the post-trap rate, and watch for potential re-emission of radon from the trap [221].

The data was treated using the usual double-deconvolution method (see section 3.1.5). Cuts were applied to remove muon events (specifically requiring Amplitude $>$ 800 ADU), as well as PSD cuts based on the raw pulse risetime and FWHM to remove non-physical background events (see section 3.1.5). With both cuts applied, the rate of events throughout the experiment is shown in Fig. 3.49, showing the low background rate of events, injection and decay of radon, and rapid effect of opening the trap. Note that there are some visible gaps in the data due to brief pauses in data collection.

The track lengths of α particles in this gas are relatively long, i.e. 21.5 cm for ^{210}Po and 37.9 cm for ^{214}Po (see ahead to Table 3.6 [183]). Because of this, many α particles

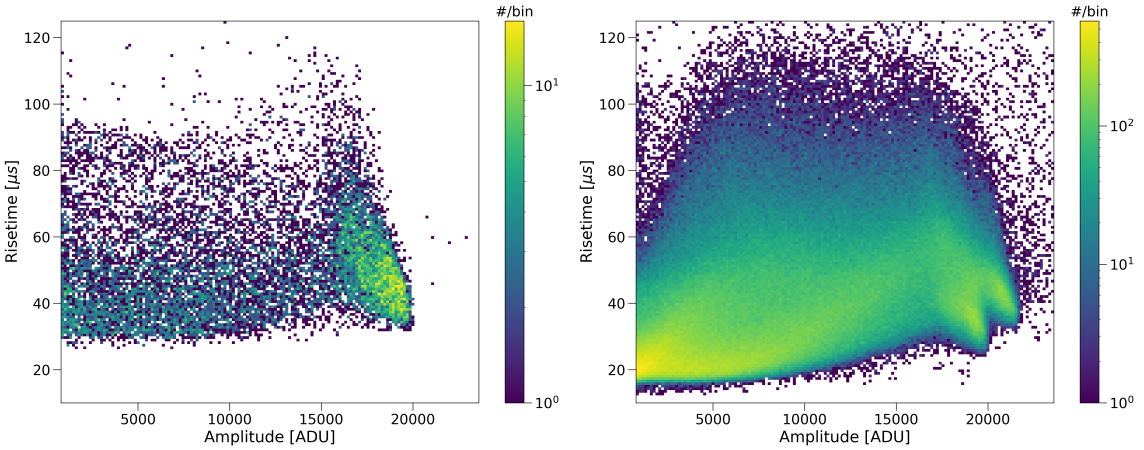


Figure 3.50: Radon trap test data taken in the U of A SPC with 200 mbar of $Ar + 3\% CH_4$ gas. Background-only data consisting of predominantly ^{210}Po α events is shown in the left panel, compared to radon data with α decay populations produced by ^{222}Rn, ^{218}Po, and ^{214}Po (right panel).

strike the inner wall of the SPC before depositing their full kinetic energy in the gas, leading to events recorded with a continuum of amplitudes below the full-energy peak. In the case of ^{214}Po, full energy deposition in the sphere was actually impossible. One can see the impact of this effect by examining the risetime vs. amplitude data, shown for both ^{210}Po and ^{222}Rn data in Fig. 3.50. Because of this, there was a great deal of overlap between the background ^{210}Po and radon decay events, making it impossible to disentangle the two signals. The distribution of risetimes for these events is due to the trajectory (and in the case of radon decays the position) of α events, with trajectories directed radially tending to have longer risetimes. Contrastingly, in an infinitely large sphere (i.e. with no wall effects for any αs directed away from the surface), the risetime vs. amplitude plot would exhibit a single amplitude peak for each decay energy, with a continuum of risetimes. Radon decays would access lower risetimes, since they occur throughout the volume instead of exclusively on the surface. The progeny of radon (^{218}Po and ^{214}Po) would exhibit a mix of surface/volume event behaviour, as they deposit on the surface of the detector over time.

The measurement was also subject to increasing charge trapping over time (due to out-gassing of oxygen and water in the system). This is evident from the changing

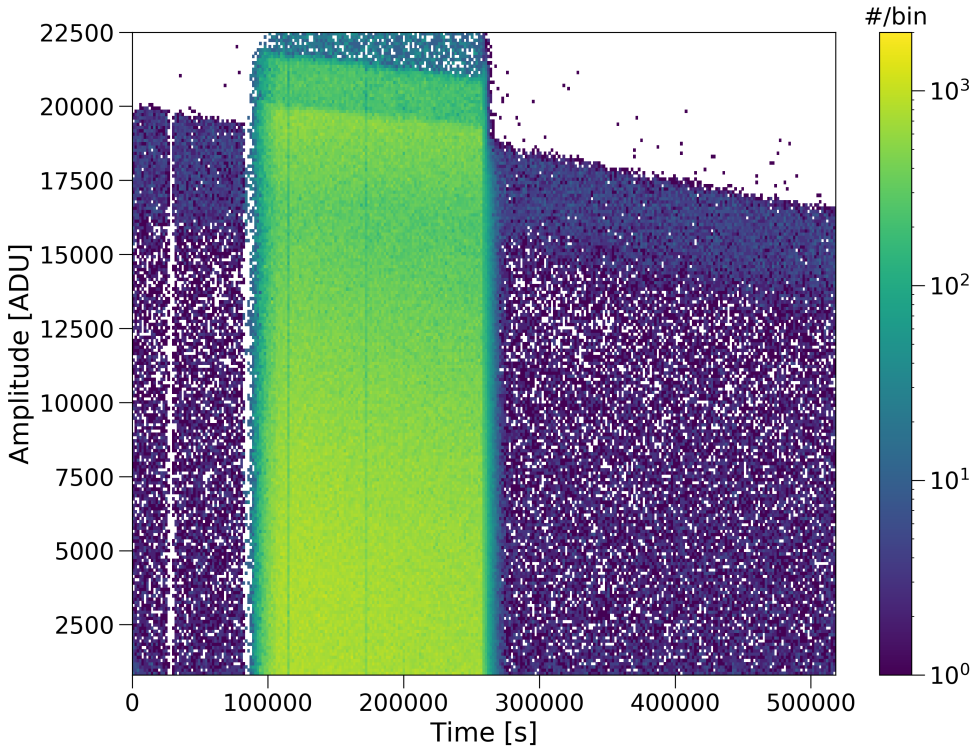


Figure 3.51: 2D histogram of amplitude vs. time for the SZ radon trap test in the U of A SPC, with PSD cuts applied.

amplitude spectrum over time shown in Fig. 3.51, which exhibits a linear decrease in the amplitude of α events, with a slight increase in slope at the time the trap is opened (the additional gas components increased the rate of out-gassing). Fig. 3.51 also makes clear how after the trap is opened, the radon events quickly disappear, leaving only the same peak/spectrum from ^{210}Po present in the background data.

To quantify the performance of the SZ trap, the radon reduction factor R (or R-value) is defined as the ratio between the expected radon activity (if the trap wasn't used) and the observed activity:

$$R(t) = \frac{A_{\text{exp.}}(t)}{A_{\text{obs.}}(t)} \tag{3.35}$$

Because of the natural decay of radon, this is necessarily a time-varying quantity. If the trap worked instantaneously to remove radon, the optimum R-value would occur at the moment of opening the trap (when the unadulterated activity would be the

highest). However, the time constant of the trap is not known, so it is not trivial to predict when the optimum measured R-value occurs. Therefore in this analysis, R was calculated as a function of time, and then the optimal value across the duration of the experimental period is reported for the final result.

To calculate R, it is first necessary to determine the expected activity of radon over time without the intervention of the trap. This was done by fitting the radon decay data (the green histogram of Fig. 3.49) with an exponential decay function:

$$A(t) = A_0 \left(\frac{1}{2}\right)^{t/t_{1/2}} + B \tag{3.36}$$

where B is a constant background component, and A_0 is the initial radon activity at $t = 0$. A binned likelihood function was constructed for the data with a prior probability for the background rate applied (using the measured background data), and maximized with an MH-step MCMC, implemented with the python package `emcee` [216]. 20 random walkers were runs for 5000 steps ($\sim 100\times$ the auto-correlation time as determined by `emcee`), after a burn-in period of 500 steps. The result was a measured half-life of 4.65 ± 0.08 days, an initial activity of 73.61 ± 0.41 Hz, and a background rate consistent with the prior measurement. Note that this is higher than the expected 3.8 days for ^{222}Rn [225]. While the contribution of ^{218}Po and ^{214}Po decays does cause a $\sim 0.1\%$ increase in half-life, the larger discrepancy is explained later in this subsection. The fit of the data over time is shown in Fig. 3.52, including statistical uncertainty from the MCMC. This result gives the extrapolated expected activity over time $A_{\text{exp.}}(t)$, if the trap were not opened.

The binned data over time (as shown in Fig. 3.52) is used to define the observed activity over time $A_{\text{obs.}}(t)$. Note that both $A_{\text{obs.}}(t)$ and $A_{\text{exp.}}(t)$ actually represent the total α activity of radon and its two daughter isotopes which remain in equilibrium with it, ^{214}Po and ^{218}Po. The activity of radon itself therefore actually only 1/3 of this, but this factor cancels out in eq. 3.35.

The following likelihood function is used to find the best-fit value of R^{-1} at time

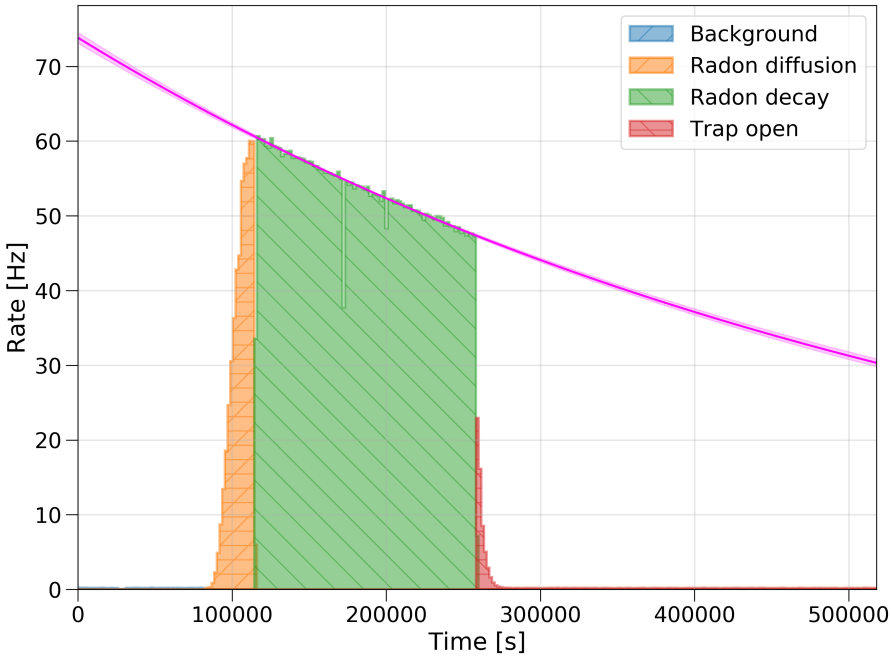


Figure 3.52: Event rate vs. time during the SZ radon trap test in the U of A SPC, showing the different phases of the experiment; background data, radon injection and decay monitoring, after which the trap is opened. PSD cuts and muon cuts are applied to this data. An exponential fit of the radon activity (green histogram) is shown as well (magenta curve with shaded $1\,\sigma$ uncertainty band).

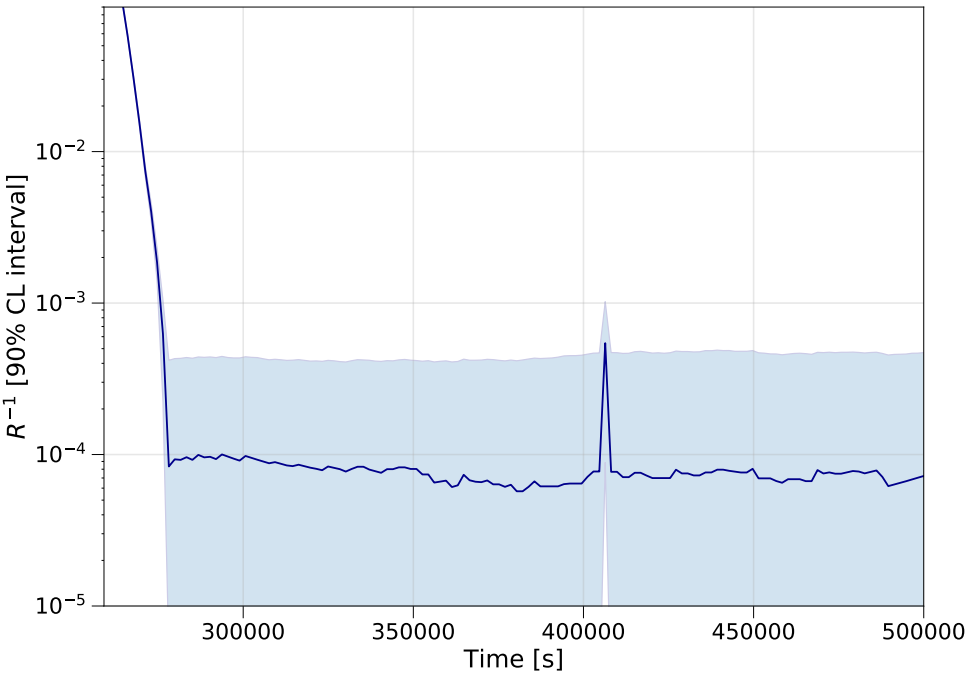


Figure 3.53: Best-fit (blue curve) and 90% CL interval (shaded region) for $R^{-1}(t)$ for the SZ radon trap test in the U of A SPC.

bin i (with counts c_i):

$$\mathcal{L}_R = P_{\text{Poisson}}(c_i | A^{(i)}_{\text{exp.}} \times R^{-1} + B) \times P_{\text{MCMC}}(A_0, t_{1/2}, B) \tag{3.37}$$

where $A^{(i)}_{\text{exp.}} = \int_{\text{bin}_i} A_{\text{exp.}}(t)dt$ is the expected counts in bin i given the radon decay model. Statistical uncertainty on the radon decay trend is incorporated by leaving these as free parameters in the fit, with a constraint term $P_{\text{MCMC}}(A_0, t_{1/2}, B)$ obtained using the MCMC samples from the fit of the radon decay data described above. The likelihood is expressed in terms of R^{-1} instead of R for convenience; having essentially zero expected counts per bin (on top of the background rate) after the trap is opened leads to a divergent range of values of R to consider. The python function `scipy.optimize.minimize` is used to optimize the likelihood function, with a Nelder-Mead algorithm [238, 239].

The Feldman-Cousins (FC) method was used to calculate a 90% CL uncertainty band for R^{-1} [240]. The advantage of this method is that the choice to report a

two-sided or one-sided confidence interval is determined automatically. Indeed, as will be shown below, early after opening the trap the result is a two-sided limit. This quickly transitions to a one-sided upper limit on R^{-1} once the rate returns down to the background level. Application of the FC method in this study involves calculating the profile likelihood ratio Λ, based on eq. 3.37:

$$\Lambda = \frac{\mathcal{L}_R\left(R^{-1}, \hat{\hat{A}}_0, \hat{\hat{t_{1/2}}}, \hat{\hat{B}}\right)}{\mathcal{L}_R\left(\hat{R^{-1}}, \hat{A}_0, \hat{t_{1/2}}, \hat{B}\right)} \tag{3.38}$$

where one circumflex symbol indicates parameters that are optimized globally (so $\hat{R^{-1}}$ is the global best-fit value of R^{-1}). A double circumflex indicates that the parameter has been profiled, i.e. optimized for a fixed value of the parameter of interest R^{-1} [73]. Λ is calculated for a (discrete, evenly spaced) range of values of R^{-1} relative to the best-fit value (for any given time bin i). Test values of R^{-1} are sorted by order of decreasing Λ, and added to the confidence interval until a total probability of 90% is reached [240].

A smoothing algorithm (a rolling average over 15 bins) is applied to the best-fit $R^{-1}(t)$ as well as the 90% CL interval, to ease statistical fluctuations in the data. Note that this is a conservative approach that cannot over-inflate $R(t)$ (including the 90% LCL), but can slightly reduce it. This procedure also guards against lower fluctuations in the post-trap data rate to some extent. The result for $R^{-1}(t)$ is shown in Fig. 3.53, from which the final result is extracted: the 90% LCL on R for the SZ room-temperature test is 2.69×10^3 (from a 90% UCL on R^{-1} of 3.72×10^{-4}). Note that even with the smoothing applied, one large fluctuation still exists (possibly due to transient, external noise events in the dataset). While a best-fit value of $R^{-1}(t)$ can be obtained, its relevance is questionable as the experiment is background limited. Therefore, for the rest of this analysis, only the 90% LCL on $R(t)$ (resp. the 90% UCL on $R^{-1}(t)$) is shown.

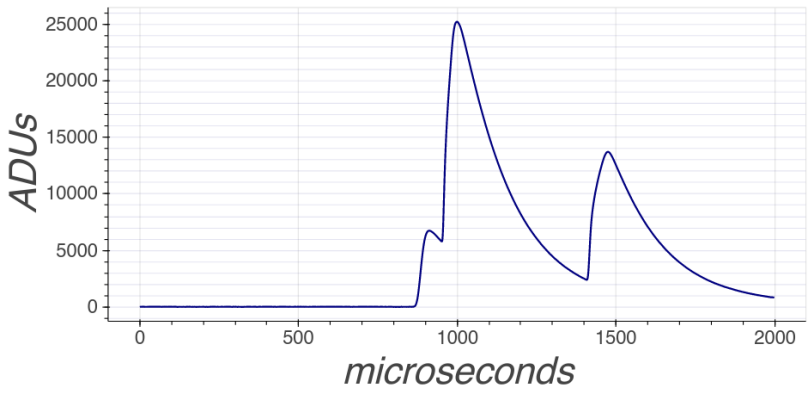


Figure 3.54: An event from the radon trap test data exhibiting potentially 3 α pulses in coincidence.

Coincident pulse and attachment systematics When operating the SPC at event rates approaching 100 Hz, the probability of having coincident peaks increases (as a function of rate) to the point of becoming common. The consequence of this is that during the radon trap experiment described above, there was a significant chance of "pileup". If two or more α decay pulses are coincident with each other, then they are under-counted – an example of such an event is shown in Fig. 3.54. Evidence of these pileup events can also be seen at amplitudes (Fig. 3.51) and risetimes (Fig. 3.50) higher than the main radon-decay populations. Another type of coincident event is one in which an α pulse falls within the 1 ms deadtime imposed by the DAQ after each event. These pulses are permanently lost, but their existence can be inferred. Both of these rate effects were more significant soon after the radon was injected (when the total event rate was highest), before decreasing over time, until becoming near-negligible again after the trap was opened. It is due to these effects that the measured half-life of radon in this analysis was greater than the expected 3.8 days.

To further demonstrate this, the radon data was re-analyzed using a peak-finding algorithm developed by F. Vazquez for the NEWS-G Kaluza-Klein axion search [49, 160]. The peak-counting results for the background ^{210}Po data and radon data are shown in Fig. 3.55. From this, one can see that although there are apparent pileup events in the background data (some of which are likely false-positive peaks found by

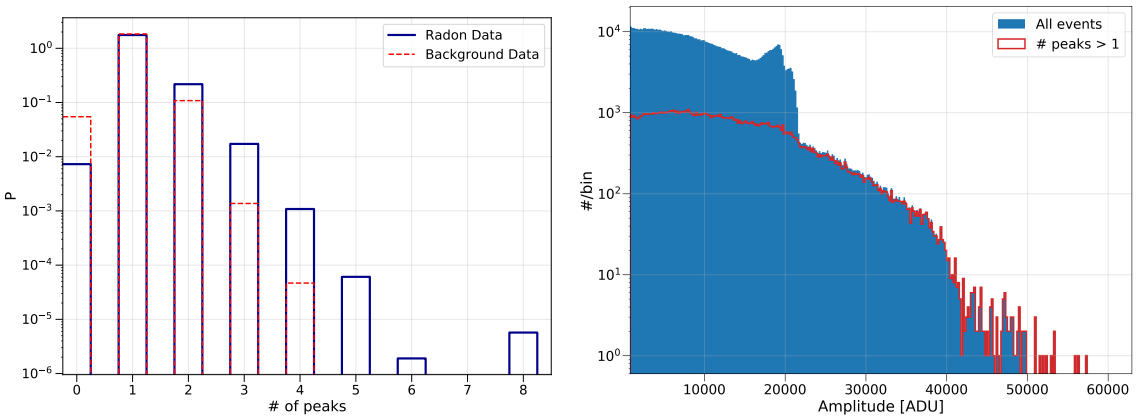


Figure 3.55: Left: peak-counting results for both background (^{210}Po) and radon data. Right: the amplitude spectrum of radon data, comparing all events to those with more than one peak identified.

the algorithm), there is a far greater proportion of such events in the radon data. Using these results (albeit with many probable false-negative and false-positive peaks), the effect of pileup can be *partially* negated by weighting the binned data in Fig. 3.52 according to the number of peaks found in each event, and re-fitting the radon decay trend. This resulted in a lower measured half-life of 4.19 ± 0.06 days. However, the peak-counting method likely misses many coincident α pulses, and does not at all account for events lost in the deadtime of previous events.

While the issue of coincident pulses is not inherently problematic, any distortion to the radon activity trend extrapolated from the fit shown in Fig. 3.52 would cause a systematic error in the result for R presented in Fig. 3.53. Investigating this potential systematic is complicated by the fact that pileup events themselves obscure the true rate of radon decays in the SPC. Therefore, it was first necessary to infer the true radon rate based on the observed rate affected by pileup. This was done by simulating radon datasets with varying rates, including the radon daughters ^{218}Po and ^{214}Po. Each MC dataset spanned the same amount of time as the real data, and was fit with the same MCMC procedure to extract the observed rate and half-life. Pileup and deadtime effects were emulated by considering any decays happening within 1 ms of each other (half the duration of a single event window) to be coincident, and any

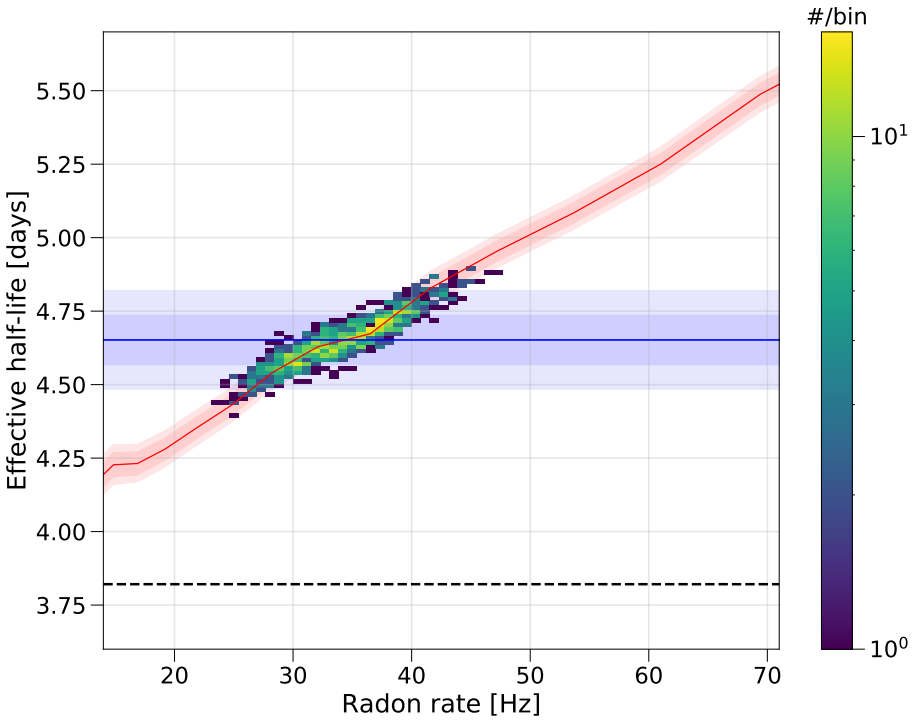


Figure 3.56: Effective half-life measured for MC datasets with varying initial radon decay rates (red curve and shaded $1\,\sigma$ uncertainty band) compared to the effective half-life of the real data (blue curve and shaded $1\,\sigma$ uncertainty band). The 2D histogram shows values randomly drawn jointly from both trends. The black dashed line indicates the true half-life of radon, 3.8 days [225].

falling within 1 to 2 ms of a preceding pulse to be lost to DAQ deadtime. This takes into account the fact that the Samba DAQ software centers the maximum amplitude in an event, so no two pulses can be more than half an event window apart and both be captured.

The results were used to construct a relationship between the measured "effective" half-life and true initial radon activity for this experiment, shown in Fig. 3.56. Accordingly, it was inferred that the true initial activity of radon (by itself, not including other α decays) was 34.2 ± 4.4 Hz. As expected, this is higher than the apparent initial radon activity (from the fit shown in Fig. 3.52) of 24.54 ± 0.14 Hz.

One MC radon dataset generated with a true radon rate corresponding to the real data is shown in Fig. 3.57. From this, it is apparent that while the exponential fit seems to match the data within the original fitting window, there are visible

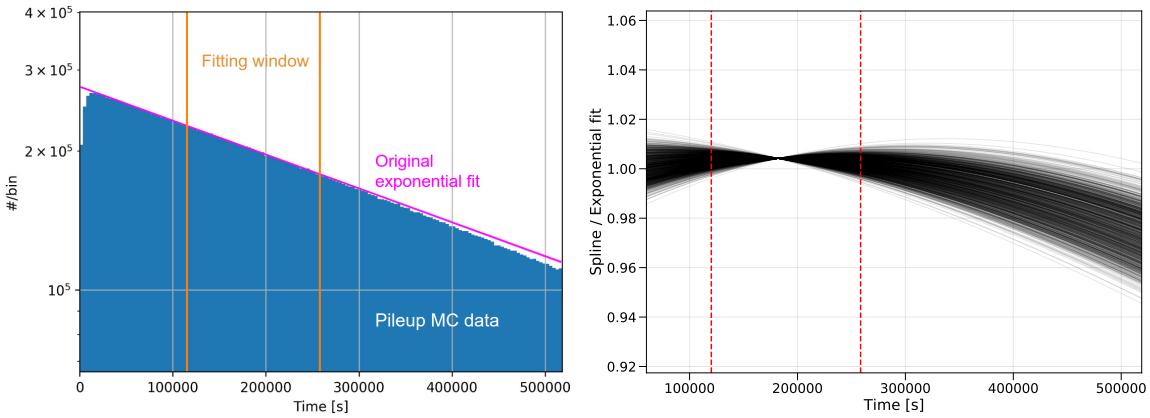


Figure 3.57: Left: the exponential fit of the real data (magenta curve), compared to an example of an MC radon dataset (with pileup accounted for) generated to match the real data. Right: the ratio of 1000 spline fits of MC datasets to the exponential fit of the real data, with the original fitting time window indicated by red dashed lines.

discrepancies with the extrapolated decay trend at laser times. To determine the systematic uncertainty on the radon decay trend, 1000 random values of the true radon activity were drawn jointly from both the uncertainty on the measured effective half-life for the real data and the MC trend of half-life vs. decay rate (also shown in Fig. 3.56). MC radon datasets were then calculated using these, and fit with a cubic spline [241] instead of an exponential decay function, to better represent the actual (slightly non-exponential) activity trends. The ratios of these spline fits compared to the exponential fit of the real data are shown in Fig. 3.57, which demonstrates that there is an appreciable discrepancy between the extrapolated radon trend and MC radon decay trends. Specifically, there is an average difference of $\sim 3\%$ by the end of the time period of the real data ($\sim 520000\,\mathrm{s}$).

Charge attachment – which results in a linearly decreasing amplitude for α events over time – gradually decreases the signal acceptance of the cut applied to remove muon/non-physical events at low amplitudes. Taking this effect into account as a systematic is complicated by the fact that the attachment trend discretely changes at the time the radon trap is opened, likely because the plumbing components associated with this system out-gas more than the rest of the system. This change in trend can

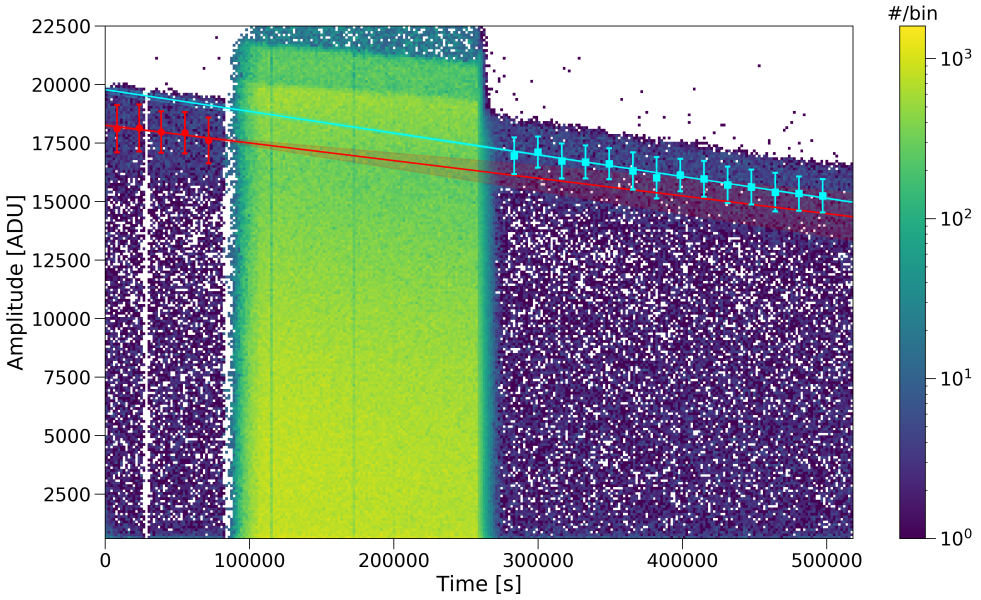


Figure 3.58: 2D histogram of amplitude vs. time for the SZ radon trap test in the U of A SPC, with PSD cuts applied. The ^{210}Po amplitude trend is fit before and after the radon injection, and fit with first order polynomials (red and cyan markers and curves, respectively).

be seen in Fig. 3.58, in which the ^{210}Po amplitude trends are fit before and after the trap is opened. The mean amplitude is determined in multiple time segments, which are then fit with first-order polynomials.

To determine the trend that radon α events would follow over time if the trap was not opened, the last 10000s of radon decay data was used to boot-strap simulated radon data [242], which was then scaled over time using the post-trap trend shown in Fig. 3.58. One such set of this boot-strapped data is shown in Fig. 3.59. This data is then directly used to calculate the cut acceptance of selecting only events with Amplitude > 600 ADU. This procedure was then repeated 1000 times, with the resulting cut-acceptance trends shown in Fig. 3.59.

Both of these systematic effects were then incorporated into the calculation of $R^{-1}(t)$. This was performed as described above, except using the spline fits of MC radon datasets in place of the exponential fit of the radon decay data, and with the time-varying cut acceptance applied. Ultimately, both of these systematics had very

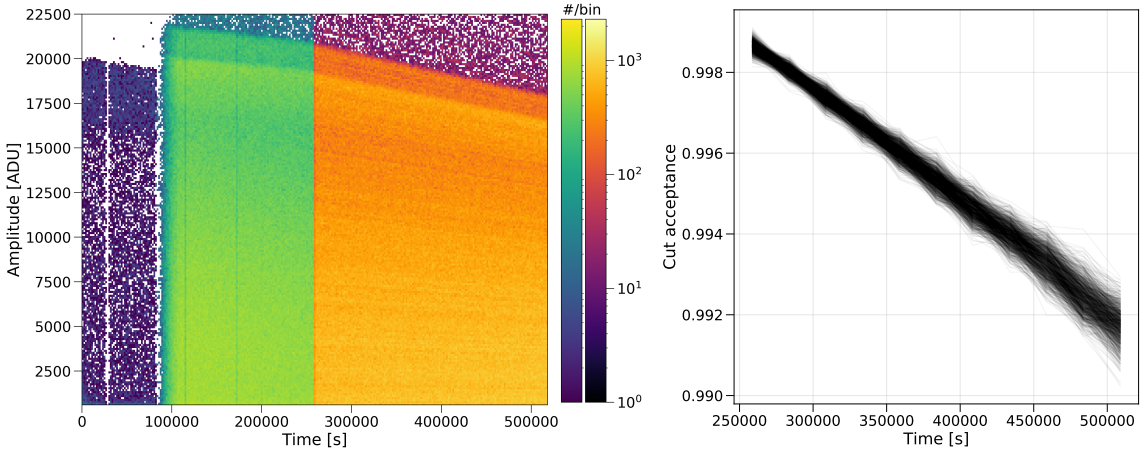


Figure 3.59: Left: the original background and radon data ("viridis" colour-map) and bootstrapped radon data ("inferno" colour-map) obtained using the last 10000s of radon data scaled by the post-trap ^{210}Po amplitude trend. Right: the resulting cut acceptances for radon events as a function of time (with 1000 trials).

small effects on the original result. The 90% UCL on R^{-1} increased by 4% compared to the original result shown in Fig. 3.53 at the end of the time period of the dataset. However, the final reported value of R was taken to be the optimum value of the 90% LCL, which occurred early enough in time that these systematic effects had a near-negligible impact: the 90% LCL on R changed from 2.69×10^3 to 2.66×10^3.

Test with the radon trap pre-saturated with methane As discussed earlier in this subsection, the tendency for radon traps to also remove CH_4 is a critical problem for the NEWS-G collaboration. One potential operating procedure that could mitigate this problem is to pre-saturate the radon trap with methane. To determine if the trap would still remove radon under these conditions, the experiment described above was repeated (with the SZ trap operated at room temperature). However, prior to the test, the radon trap was saturated with methane gas. CH_4 was injected in small bursts, after which the pressure was often observed to quickly drop as the gas became trapped in the porous substrate. This titration was repeated until the pressure no longer dropped after injection, at which point the trap was sealed. The experiment was then carried out following the same procedure described for the first

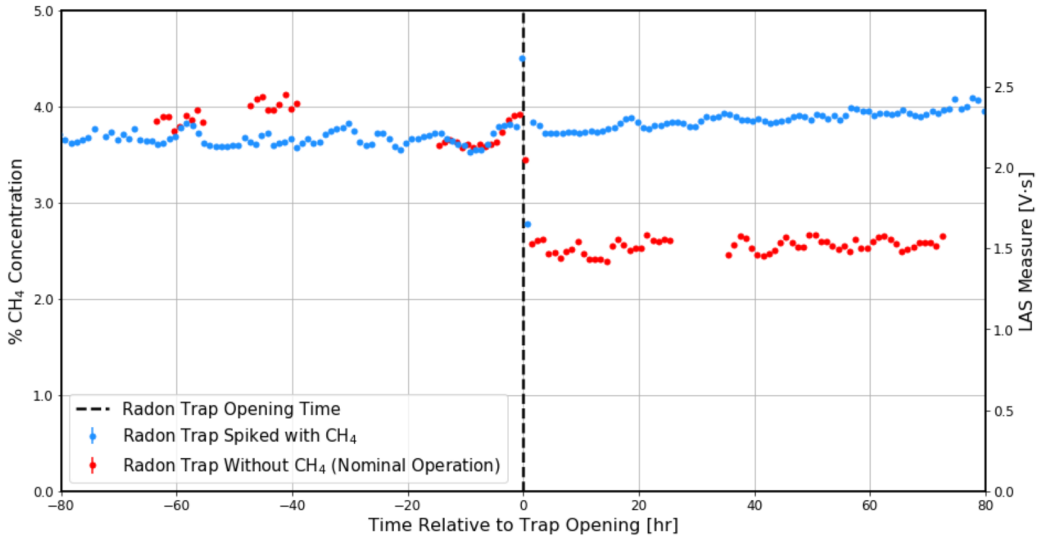


Figure 3.60: Preliminary measurements of methane concentration over time in the U of A SPC, measured with the LAS system (direct LAS output on right axis) [187], during a normal radon trap test (red) and one in which the trap was pre-saturated with CH_4. Credit to C. Garrah.

test. The analysis was also performed as described above, including the procedures to account for charge trapping and pileup/rate effects as systematics. The resulting 90% LCL on R obtained for this test was 3.70×10^3 – this is slightly higher than the original test due to a slightly higher initial radon activity, improving the sensitivity of the experiment. This demonstrates an equal ability to remove radon from the gas despite the pre-saturation of the trap.

The LAS system described in section 3.4.1 was also used to monitor the methane concentration of the gas in real-time for both experiments. The preliminary results (the LAS system still requires absolute calibration at this time [187]) are shown in Fig. 3.60. In the second experiment, no methane was lost to the trap, and in fact, the pre-saturated trap may have re-emitted CH_4 during the experiment. This demonstrates that the goal of removing radon while maintaining a constant CH_4 concentration in an SPC is feasible, given the correct operating procedure. The LAS system may be used in the future to more precisely pre-saturate the radon trap, as well as inform periodic methane concentration corrections for long-running experiments.

Comparison with Rad7 tests As discussed above, tests of SZ radon traps (in different operating conditions) were carried out by P. O'Brien using a setup with a Rad7 device in place of the SPC [221]. The general procedure of these tests was the same; radon was injected, and monitored for some time, following which the radon trap was opened. Full details of these tests are given in ref. [221], as this is beyond the scope of this work. However, the original analysis of this data used a slightly different definition of the R-value, prohibiting comparison with the SPC results. Specifically, R was defined to be the ratio of the radon activity just before the trap was opened, compared to the (averaged) activity observed afterwards. To facilitate a more accurate comparison, the Rad7 data was re-analyzed using the definition of R given by eq. 3.35.

Similarly to the SPC analysis, the trend of radon decay before opening the trap was fit with an exponential decay function to estimate the extrapolated radon activity if the trap was never opened ($A_{\text{exp.}}$). This was then compared to the observed activity after the trap was opened to calculate the best-fit values of $R^{-1}(t)$, with 90% CL intervals again calculated using the FC method [240]. Examples of this data and the resulting $R(t)$ trends are shown in Fig. 3.61. Note that in cases where the post-trap rate was not low, the analysis was done in terms of $R(t)$ instead of $R^{-1}(t)$. These tests often ended in noticeable, oscillating re-emission of radon from the trap after long latency periods [221]. Additionally, in cases where the radon activity was not reduced to background levels, there were oscillations in the activity on short time scales after the trap was opened [221]. To avoid these time periods, the calculation of $R^{-1}(t)$ was truncated to times after the oscillations dampened to $< 3\,\sigma$ of the post-trap trend.

The Rad7 data includes measurements of radon traps with SZ as well as activated charcoal – specifically the substrate Carboxen® 1000 (Sigma Aldrich). Data was taken at room temperature (and dry-ice temperatures in some instances). Various carrier gases were also used: N_2, pure argon, as well as $\text{Ar} + 3\%\,\text{CH}_4$. The latter surprisingly seemed to lead to generally worse outcomes with the Rad7, for unknown

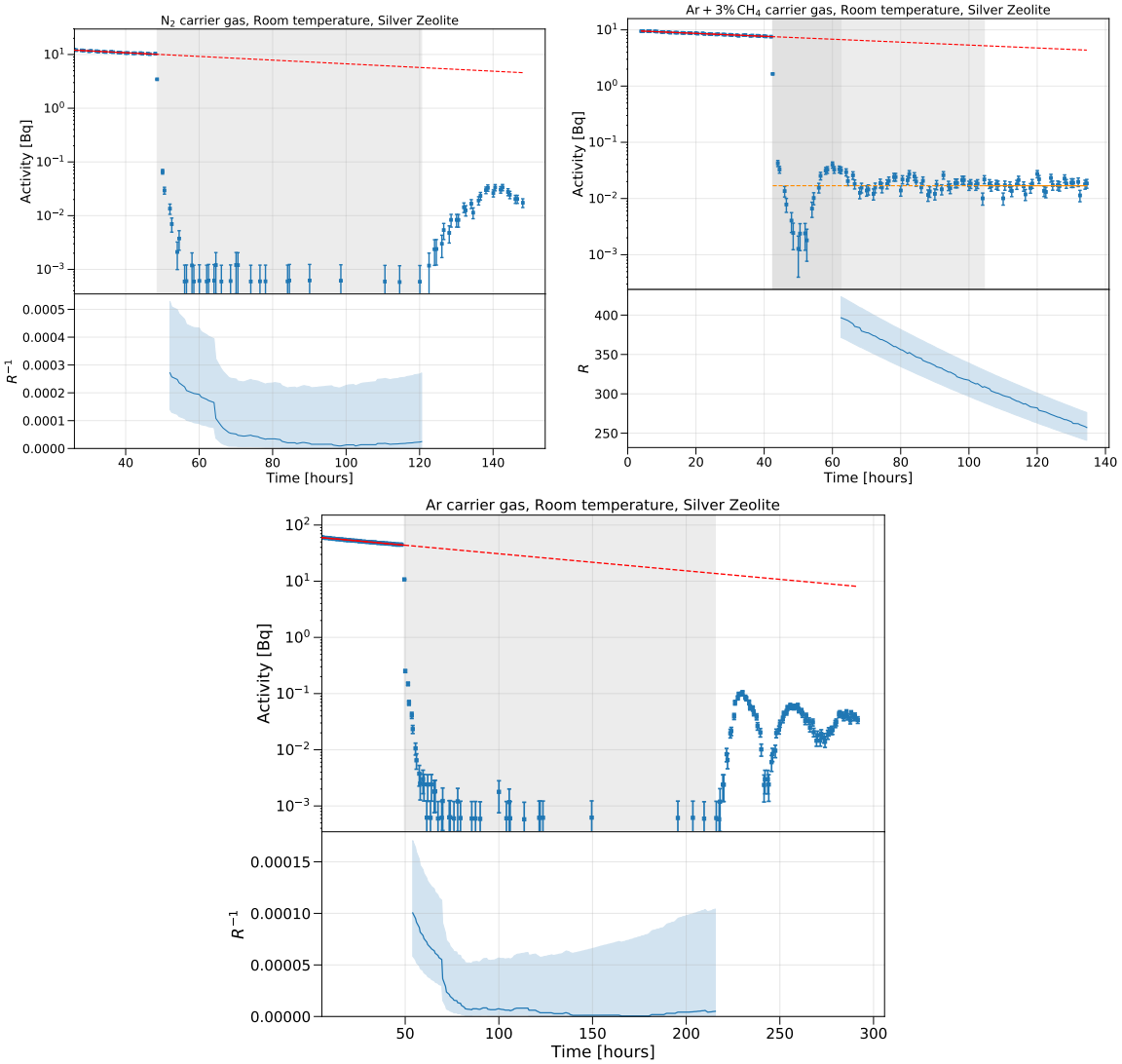


Figure 3.61: Select examples of Rad7 radon trap test data (blue histograms) with the resulting $R(t)$ (or $R^{-1}(t)$) trends, which best-fit and 90% CL intervals shown in blue. The pre-trap decay trend is shown in red, and the post-trap trend (where needed) is shown in orange. These experiments were performed with a room temperature SZ trap with N_2 carrier gas (top left), Ar + 3% CH_4 carrier gas (top right), and pure argon carrier gas (bottom).

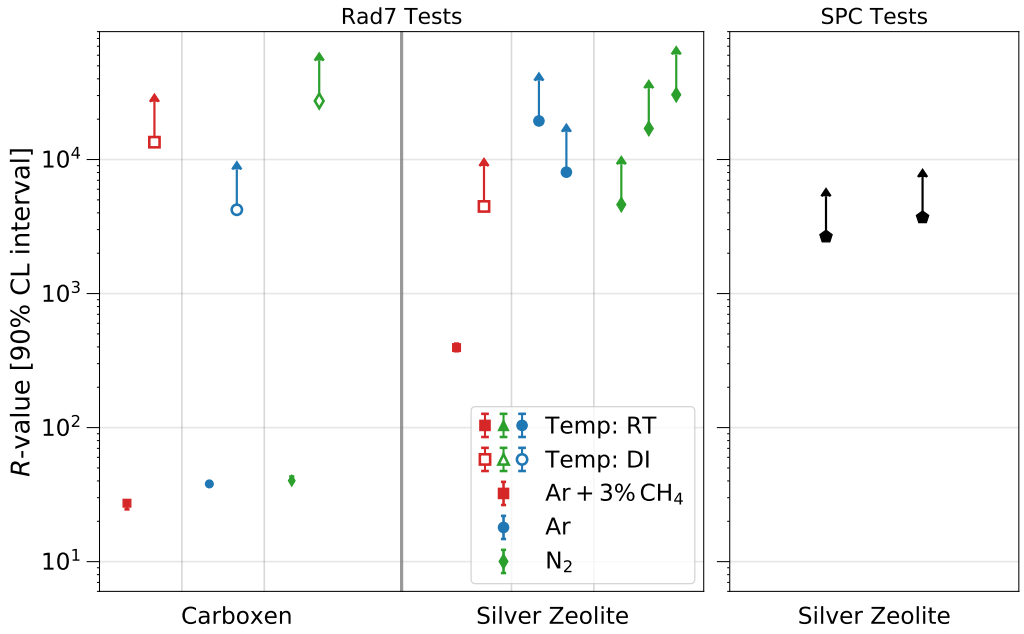


Figure 3.62: Left: preliminary R-value results from the Rad7 tests [221] with carboxen and SZ radon traps, with different carrier gases operating temperatures, re-analyzed for this work. Right: R-values for room temperature SZ obtained from the SPC tests. Data points with upper and lower error bars represent the two-sided 90% CL intervals, while the markers with arrows represent one-sided 90% LCLs for R.

reasons [221]. All of the re-calculated Rad7 R-values are presented in Fig. 3.62. These promising results confirm – as in ref. [221] – that SZ functions approximately as well at room temperature as activated charcoal does at dry-ice temperatures. On the other hand, the Carboxen traps performed orders of magnitude worse at room temperature. The different results for R with repeated measurements are due to different initial activities of radon in each test (a higher initial activity potentially results in a higher 90% LCL for R).

The room-temperature SZ tests performed with the Rad7 are compared to the SPC result in Fig. 3.62 as well. One can see that these results are generally compatible. The SPC tests yield a slightly lower 90% LCL for R in most cases, which is due to the limited sensitivity of the experiment. Although injecting higher amounts of radon into the SPC could lead to better results, the issue of coincidences would become significantly worse. Future improvements to the SPC experiment method-

Pressure [mbar]	^{210}Po	^{214}Po	^{218}Po	^{222}Rn
200	21.5 cm	37.9 cm	25.9 cm	22.7 cm
500	8.62 cm	15.1 cm	10.3 cm	9.09 cm

Table 3.6: Ranges of α particles produced by radon decay-chain isotopes in $Ar + 3\% CH_4$ gas (at 20°C), calculated using SRIM [183].

ology are discussed in section 3.4.3. While this re-analysis of the Rad7 data allows for better comparison with the SPC results, it does suffer from the shortcoming that the measured half-lives of radon decay in the Rad7 data vary significantly (with values both higher and lower than 3.8 days [225]), for unknown reasons. Unlike the SPC tests, rate effects alone cannot explain this phenomenon. Therefore, more study and/or repeated measurements are required.

3.4.3 Future work

Further radon trap tests are planned with the U of A SPC. This includes a full series of measurements of the performance of SZ under different operating conditions, as well as tests of charcoal and carboxen traps for direct comparison. Following the measurements presented in this work, the operating conditions of the radon trap procedure have been improved somewhat. By operating the SPC with a higher-pressure gas (500 mbar of $Ar + 3\% CH_4$ instead of 200 mbar), the range of the α particles in the gas is significantly reduced. The ranges of radon decay chain αs as well as ^{210}Po in these different conditions are given in Table 3.6, calculated using the SRIM simulation software [183]. Consequently, a smaller portion of events will collide with the SPC wall, depositing their full energy in the gas instead. Thus, a higher proportion of events from radon and its progeny will be distinguishable in amplitude from those of ^{210}Po, the main background of these measurements. Examples of background ^{210}Po and radon data in this higher pressure gas are shown in Fig. 3.63, showing the improved separation between these different populations compared to Fig. 3.50. Note that the lower risetimes of some radon events compared to ^{210}Po is partially due to

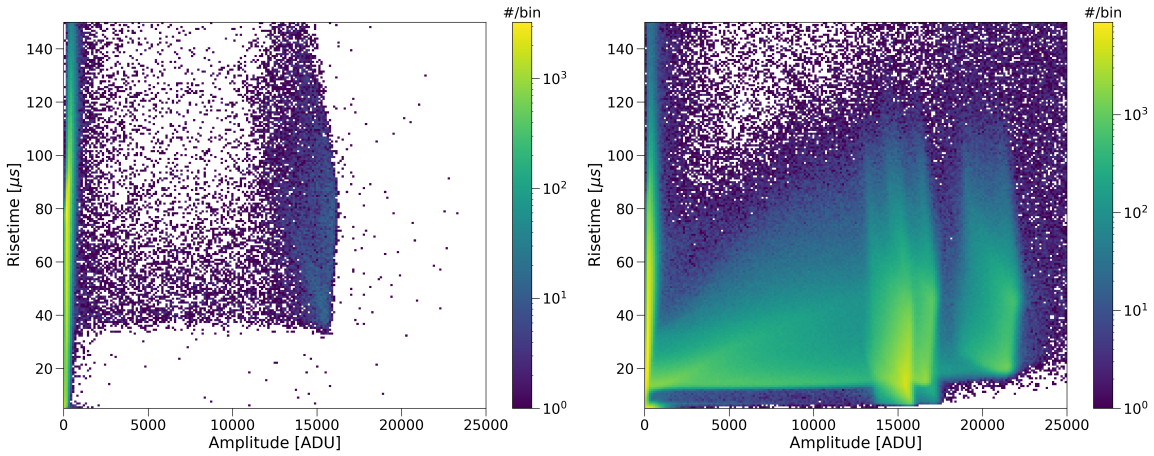


Figure 3.63: Radon trap test data taken in the U of A SPC filled with 500 mbar of $Ar + 3\% CH_4$ gas. Background-only data including ^{210}Po α events is shown in the left panel, compared to radon data with visible α decay lines from ^{222}Rn, ^{218}Po, and ^{214}Po (right panel).

the fact that radon decay isotopes may emit αs from within the volume of the SPC rather than on the surface. This is also likely affected by the different space-charge conditions caused by the high-rate conditions of the radon data.

By selecting only ^{214}Po events (above $\sim$ 18000 ADU in Fig. 3.63), the signal-to-background ratio of the experiment will be significantly improved. As ^{214}Po activity will be in equilibrium with ^{222}Rn, it can be used as a proxy for radon activity over time. The goal of these future measurements will be to definitively quantify the performance of SZ vs. traditional radon trap substrates and to identify optimal operating conditions for future NEWS-G experiments. Using the LAS system installed in this setup, the effect of radon traps on methane concentration can be measured, and procedures to correct/reduce this effect can be further improved.

The U of A SPC is also currently being used to carry out W-value measurements, following similar methods to those presented in section 3.3. Specifically, tests in $Ar + 3\% CH_4$ gas at various pressures are being performed to check for any dependence in the W-value. Future tests will expand to neon and helium gas mixtures, as well as varying methane concentrations (which again will be aided by the use of the LAS system). The work of other graduate students in the Piro group also includes

studies of space charge effects in SPCs; data taken while varying the intensity of UV laser events can be used to directly test simulations of electron/ion transport in the sphere.

Finally, a long-term goal of the NEWS-G collaboration is to perform quenching factor measurements via the neutron-capture method [90]. While such experiments have been carried out/are planned with other detector technologies [243–246], this method has not been applied to gaseous detectors or any NEWS-G target materials. In this approach, a neutron source is used to induce (n, γ) reactions in the gas. The recoiling nucleus provides a source of low-energy nuclear recoil events with which to measure the quenching factor, and the emitted high-energy γ-rays can be used to tag events/specific de-excitations. As with any quenching factor experiment, confounding energy response parameters such as the W-value and detector gain must be measured *in situ*. Therefore – with the proven capacity/equipment to do so – the U of A SPC is well-suited to this task. An additional advantage of this detector is its multiple calibration ports, which may be used for γ coincidence detectors in a neutron capture experiment. A potential schematic of such an experiment with an SPC is shown in Fig. 3.64.

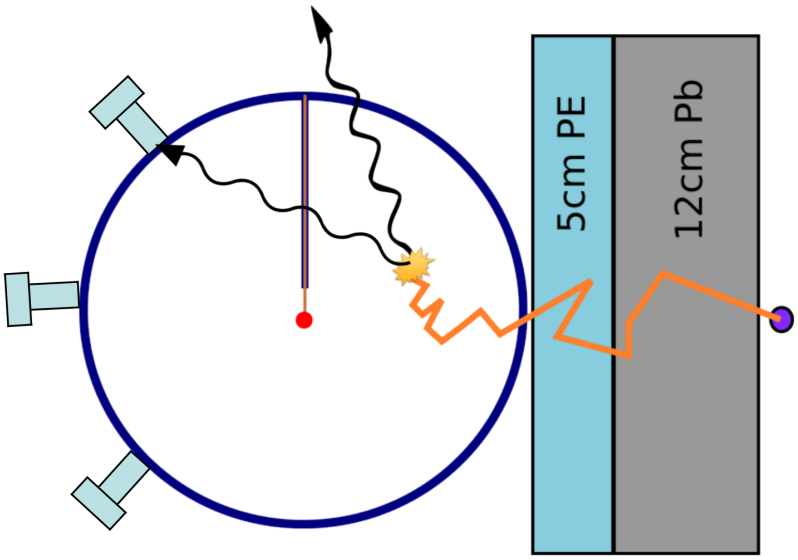


Figure 3.64: Schematic of a hypothetical neutron capture experiment with the U of A SPC, to measure quenching factors. A neutron source (such as an Am-Be source [247]) placed outside the sphere induces (n, γ) reaction inside the gas, with the emitted γs being observed by backing detectors outside the SPC. A polyethylene layer serves to thermalize the neutrons, while lead would help attenuate any non-neutron backgrounds produced by the source.

Chapter 4

The NEWS-G experiment: dark matter search results in methane gas

After the SEDINE experiment at the LSM concluded (see section 3.1.6), the main effort of the collaboration has been the construction and operation of a larger 140 cm diameter detector named S140 (sometimes "SNOGLOBE" unofficially) [1, 248]. While temporarily installed at the Laboratoire Souterrain de Modane, a dark matter search experiment was carried out with a methane gas target. The subsequent analysis – involving many novel techniques and improved SPC characterization – produced a world-leading exclusion limit on low mass WIMPs with a spin dependent proton coupling.

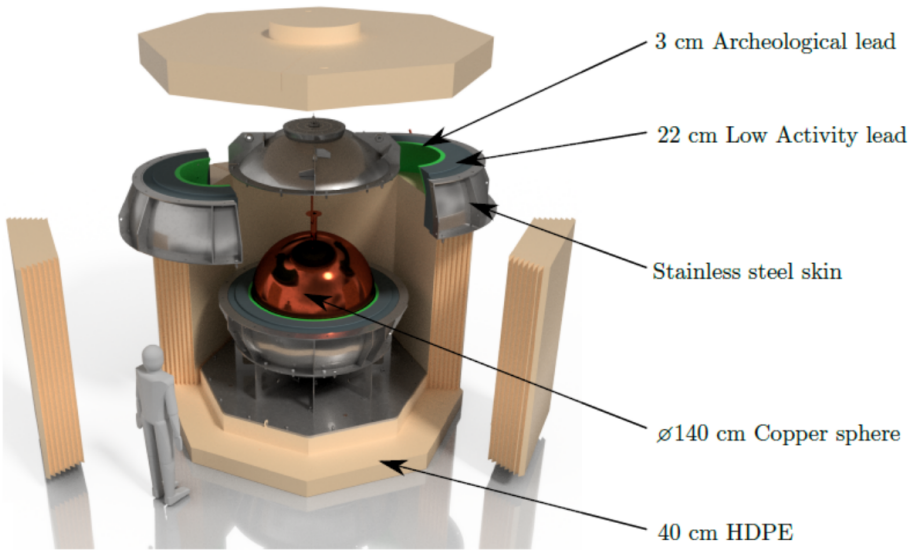


Figure 4.1: The S140 detector (inner copper sphere) is encased in shielding layers of archaeological lead, low-activity lead, and high-density polyethylene. Taken from figure 1 of ref. [248].

4.1 The S140 detector

A 3D model of the S140 detector is shown in Fig. 4.1. This SPC has a single port at the top of the sphere for vacuum/gas handling, through which the sensor is also inserted. The vessel is constructed of 99.99 % pure, C10100 oxygen-free copper [1]. Two disks of the commercially-available copper were spun/pressed into hemispheres, that were later electron-beam welded together. Copper is a common choice of material for rare-event search experiments, as there are relatively few long-loved radioisotopes that can be produced cosmogenically from copper [2].

The shielding setup for the experiment – also depicted in Fig. 4.1 – includes (starting with its outermost layer) a 40 cm thick octagonal castle of high-density polyethylene (HDPE) blocks, intended to shield against neutrons from the environment [1]. The blocks have saw-toothed, interlocking edges to avoid direct line-of-sight paths through the HDPE shielding. These components were manufactured at the University of Alberta. Inside this, a spherical assembly of commercial low-activity lead encloses the detector, forming a 22 cm thick layer [1]. These pieces are held in place

with a stainless steel "skin" structure and scaffolding. Finally, the innermost layer of shielding consists of 3 cm of Roman archaeological lead [249]. This material, taken from a Roman shipwreck in the Mediterranean, has been shielded from cosmic ray activation for thousands of years. Thus, it is expected to have exceptionally low radioactivity, making it a valuable shielding material. A plug made of Roman archaeological lead is designed to complete the shielding setup, with spiral channels running through it for gas flow, with additional channels for the UV laser optical fibre and electric connections [1]. In this way, there is no direct line-of-sight path through the detector shielding that would permit γ-rays into the SPC.

The spaces inside the HDPE are constantly flushed with boil-off nitrogen to reduce radon and dust levels in the vicinity of the detector [1]. The SPC itself is also always kept under vacuum or filled with inert gases to avoid the introduction of radioactive contaminants, as well as water and oxygen. The electronics of the detector (for specific hardware, see section 4.1.3) are housed inside a sealed glovebox, attached to the top flange of the sphere. A separate, air-tight tube houses the sensor and support rod and can be connected to the glovebox. Using this system, the top flange of the SPC can be opened, the sensor inserted via a winch at the top of the sensor storage tube, and electrical connections made without any exposure to air [1].

In such a large SPC volume, the electric field produced by a typical 2 mm anode is relatively weak at the outer reaches of the sphere, even with $\mathcal{O}$(kV) voltages applied. In these circumstances, there is a significant probability that primary electrons will diffuse backwards and collide with the vessel wall. To avoid this problem, a larger-radius sensor could be used (as large as a few cm), but this would lead to weaker electric field strengths near the anode, resulting in smaller Townsend avalanches. A new sensor design – the "ACHINOS", from the Greek word for sea urchin – was designed for S140 to solve this problem [250, 251]. The sensor consists of 11 individual 1.7 mm anodes arranged in an isotropic fashion relative to the centre of the overall sensor structure. A central electrode supports this arrangement, coated in a resistive

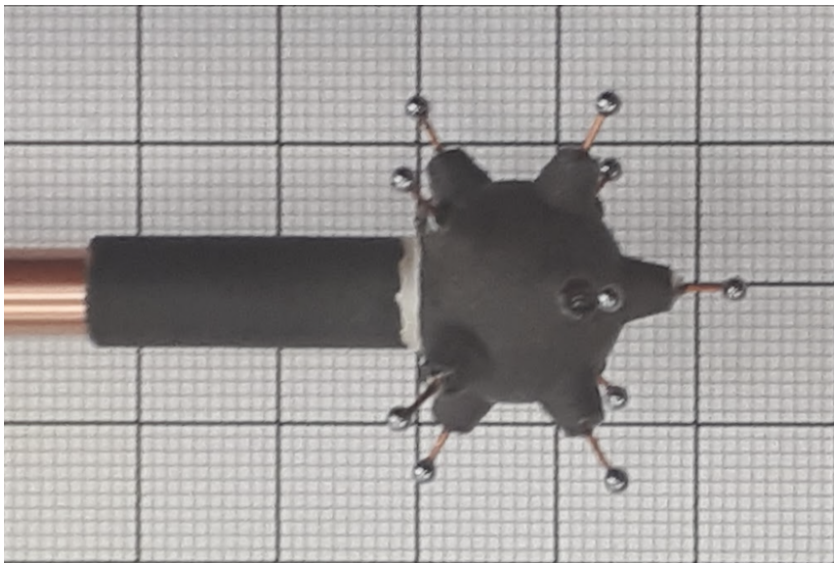


Figure 4.2: An example of a diamond-like carbon-coated ACHINOS sensor. Taken from figure 2 of ref. [1].

material such as diamond-like carbon to prevent discharges from the anodes [251]. Each anode is equally displaced from the centre by $\approx 1\,\text{cm}$. An example of this sensor design is shown in Fig. 4.2.

When used, this sensor provides a strong drift field at the surface of the sphere (due to the collective field of all 11 anodes), but locally each anode produces strong Townsend avalanches. Critically, the electric field remains roughly isotropic in the vicinity of each anode, providing (in principle) uniform avalanche yields from all anodes [251, 252]. Studies of the gain and resolution of ACHINOS sensors as a function of incoming particle trajectory have been carried out for some sensors [251], but notably not of the specific sensor used in S140 for the first physics campaign at the LSM.

The electric field produced by the S140 ACHINOS is shown in Fig. 4.3, which exhibits an oscillating pattern (rotating about the central axis) due to the arrangement of the anodes. Currently, ACHINOS sensors are fabricated so that all 6 anodes far from the support rod are grouped together into a single channel for readout (with connected wires), and all 5 near the support rod are grouped into a second channel.

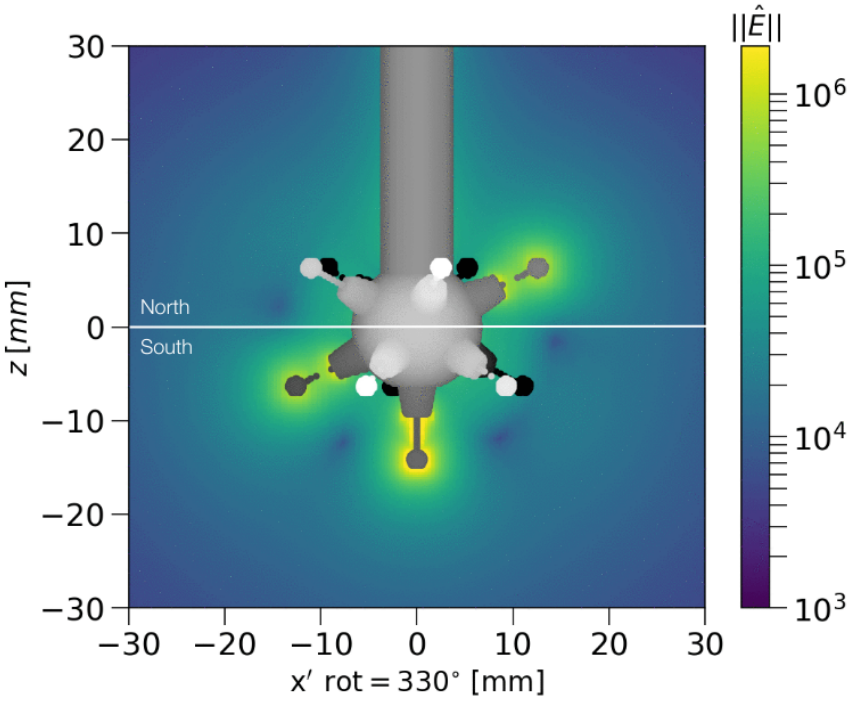


Figure 4.3: A finite element simulation of the electric field near an ACHINOS sensor (performed using COMSOL [154]), with the colour map showing the electric field strength (in V/m) in the $y = 0$ plane. Figure 4.69 shows electric field lines from the same simulation.

These are sometimes dubbed the "south" and "north" channels respectively, due to their orientation inside S140. The fiducial volume of these two channels is discussed in section 4.4.3. Future NEWS-G experiments may be able to support individual readout of all 11 ACHINOS anodes.

The single largest source of background radiation for the S140 experiment is from ^{210}Pb in the copper of the vessel itself. Measurements of this contaminant and mitigation strategies are discussed in the following two subsections. Another significant contribution of background events is from cosmogenic activation of the copper [1, 152]. Cosmic ray exposure on the surface of the Earth produces the radioisotopes ^{56}Co, ^{57}Co, ^{58}Co, ^{60}Co, and ^{64}Mn, which then produce backgrounds of electronic recoil events. While the detector materials were kept underground for storage and fabrication as much as possible, there was some inevitable exposure. Fortunately, these isotopes all have relatively short half-lives, with the longest being 5.27 years for

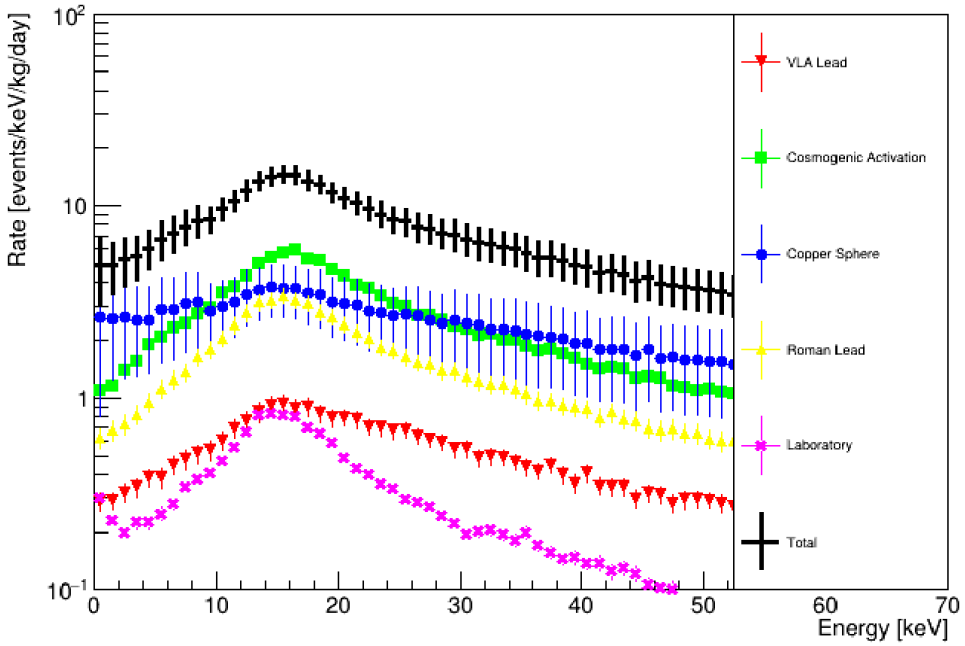


Figure 4.4: Simulated differential background rate in the S140 sphere from different components in the experiment/environment. This simulation was carried out for a target gas of 135 mbar, at the time of the LSM physics campaign (see section 4.1.3). Taken from figure 10 of ref. [1].

^{60}Co [253]. This means that the background rate from cosmogenics in the experiment will decrease over time, but also means that simulations of this background must take the age of the detector into account [1].

Naturally occurring ^{238}U, ^{232}Th, and ^{40}K is found in small concentrations in the SPC copper, lead shielding material, and lab environment. ^{210}Pb is also found in the lead shield. The concentrations of these contaminants was measured using the ICP-MS technique at PNNL [152]. Considering all of the above-listed sources of background events, a Geant4 simulation of the total expected background spectrum in S140 was carried out, which is shown in Fig. 4.4. More details of this background model and contamination assays are given in refs. [1] and [152].

4.1.1 ^{210}Pb Assay

As determined by Geant4 simulations [152, 204], ^{210}Pb is a dominant background signal in SPCs. In particular, Bremsstrahlung emission induced by the decay of ^{210}Pb

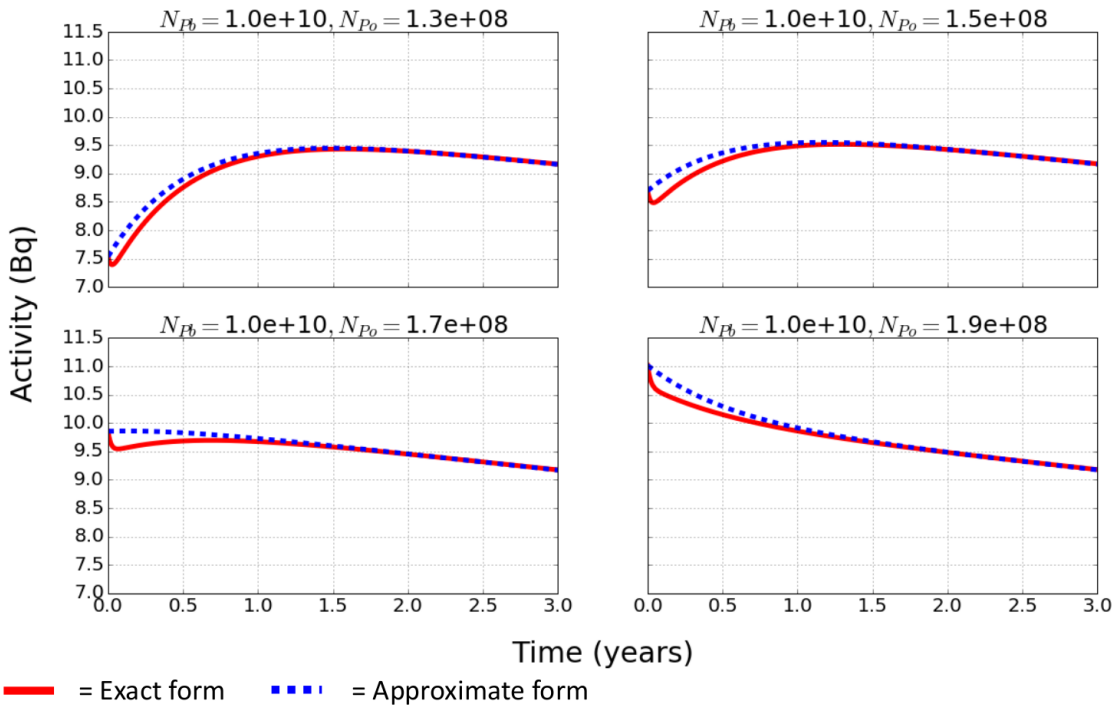


Figure 4.5: The evolution of measured ^{210}Po over time for different initial activities of ^{210}Po and ^{210}Pb, using the exact solution (eq. 4.1, red curve) and approximate formula (blue dashed curve).

(and its short-lived daughter ^{210}Bi) inside the copper of the S140 vessel leads to a homogeneous background of low energy events that cannot be discriminated against with PSD variables. Being a long-lived isotope with a 22 year half-life [254], it is impractical to plan to let it decay away. Therefore, effort was directed towards both reducing and precisely quantifying this background. To pursue the latter, NEWS-G worked with the XMASS collaboration to measure the ^{210}Pb content of the C10100 copper used for the S140 detector. The method employed by XMASS for similar previous studies [255] involves measuring the α-decay activity of ^{210}Po (the grand-daughter of ^{210}Pb) with an XIA UltraLo-1800 ionization chamber [256]. This chamber uses an active veto anode layer to reject α events not originating from the sample being assayed. The detector is also flushed with argon gas to reduce backgrounds from radon gas. For the measurements performed by XMASS, the assay sample tray of the detector was replaced with an electroformed copper tray to further reduce

Measurement date	Count rate [α/cm^2/hour]
July 2-25, 2018	$2.3 \pm 0.4 \times 10^{-4}$
October 5-17, 2018	$2.2 \pm 0.4 \times 10^{-4}$
December 28, 2018 - January 9, 2019	$1.4 \pm 0.3 \times 10^{-4}$
April 19 - May 7, 2019	$1.4 \pm 0.3 \times 10^{-4}$

Table 4.1: Raw ^{210}Po count rates measured from the S140 copper sample by the XMASS collaboration [2].

backgrounds [255].

The measured ^{210}Po activity originates from a combination of whatever initial quantity of ^{210}Po is present in the copper sample after casting, as well as that which builds up over time from decaying ^{210}Pb. As equilibrium between these isotopes is broken during the casting process [255], multiple measurements are required over time to disentangle the initial quantities of ^{210}Po and ^{210}Pb. To that end, the XMASS collaboration performed 4 measurements of a sample of NEWS-G copper over a period of $\sim$ 1.5 years since the manufacture of the copper sample [2]. This sample was treated and electro-polished at the Pacific Northwest National Lab (PNNL [257]) following the same procedure as the copper of S140 itself. The count rates of these four measurements are given in Table 4.1.

The activity of ^{210}Po over time can be calculated using the Bateman equation, which gives the quantity $N(t)$ of the n^{th} isotope in a decay chain over time:

$$N_n(t) = \sum_{i=1}^{n} \left(N_i(0) \times \left(\prod_{j=i}^{n-1} \lambda_j \right) \times \left(\sum_{j=i}^{n} \frac{e^{-\lambda_j t}}{\prod_{p=i,p\neq j}^{n} \lambda_p - \lambda_j} \right) \right) \tag{4.1}$$

where λ_i are the lifetimes of the isotopes involved, and $N_i(0)$ are their initial activities [258]. It is worth noting that one could take an approximate form of this result by neglecting the intermediate isotope ^{210}Bi, which has a short half-life of $\sim$ 5 days compared to its parent and daughter. However, this can cause a systematic over-prediction of ^{210}Po activity for measurements within about 1 year of production,

although the magnitude of this bias depends on the initial activities and is never more than $\sim 5\%$. Examples of the evolution of ^{210}Po activity over time following eq. 4.1 (and the approximate form used by [255]) are shown in Fig. 4.5.

As the primary focus of this study is measuring ^{210}Pb in the bulk volume of the copper sample, rather than on the surface, the bulk activity of ^{210}Po must be measured specifically. This activity is distinguished from the surface ^{210}Po activity by selecting α decays with partially degraded energies compared to the full 5.3 MeV kinetic energy of ^{210}Po αs [255]. An energy window of 2.5 to 4.8 MeV is used to identify decays at a depth of approximately 2 to 8 μm. A Geant4 simulation was performed by the XMASS collaboration to convert the count rate observed in this energy band (for a given sample surface area for counting, given in table 4.1) to the bulk activity of ^{210}Po; 2.7×10^{-2}(Bq/kg) / (α/cm^2/hour) [255]. The XMASS collaboration performed an experimental check of this conversion factor and found a systematic uncertainty of -10% to $+30\%$ [255].

After the raw observed count rates were converted into bulk ^{210}Po activity, the activity over time was fit as a function of the initial activity of bulk ^{210}Po and ^{210}Pb (A^0_{Po} and A^0_{Pb}). Eq. 4.1 is multiplied by the lifetime λ of ^{210}Po to give its activity over time:

$$A_{\mathrm{Po}}(t) = \lambda_{\mathrm{Po}} \times N_{\mathrm{Po}}(t) \tag{4.2}$$

For each measurement i between time t_{1_i} and t_{2_1}, the number of counts observed is Poissonian, with an expected rate given by:

$$\mu_i = \int_{t_{1_i}}^{t_{2_i}} A_{\mathrm{Po}}\left(t | A_{\mathrm{Po}}, A_{\mathrm{Pb}}\right) dt + N_B, \tag{4.3}$$

where N_B is a constant background contribution included in the fit model. This was then used to form a joint, unbinned likelihood function for the four measurements [73] each with c_i counts:

$$\mathcal{L}\left(\{c, t_1, t_2\} \,|A_{\mathrm{Pb}}, A_{\mathrm{Po}}, N_B\right) = \prod_{i=1} \frac{\mu_i^{c_i} e^{-\mu_i}}{c_i!} \times \mathcal{L}_B, \tag{4.4}$$

The term $\mathcal{L}_B$ is a Gaussian constraint on the background activity of using the rate of $5.6 \pm 5.6\,\alpha/\mathrm{cm}^2/\mathrm{hr}$ given by ref. [255]:

$$\mathcal{L}_B = \frac{1}{\sqrt{2\pi\sigma_B^2}} e^{-\frac{(N_B-\mu_B)^2}{2\sigma_B^2}}. \tag{4.5}$$

In this analysis, the likelihood function was maximized using the ROOT Minuit2 optimization package [259] to find best-fit values for the three parameters of this model. To estimate uncertainties (statistical), a profile-likelihood approach was used [73] to profile over the background activity as a nuisance parameter, and then a raster-scan was used to estimate the $1\,\sigma$ uncertainty contour for the initial Pb and Po activities (taking the region within 0.5 of the maximum log-likelihood) [73]. The 2D profiled likelihood functions are shown in Fig. 4.6 including the first two measurements (left panel) and all four measurements (right panel). This reconfirms the necessity of multiple measurements over time for this analysis: without data taken over a long period of time, the contributions of ^{210}Po and ^{210}Po to the observed activity are completely degenerate.

A marginal likelihood method was initially considered to give a more robust estimation of these uncertainties since the likelihood function was strongly bounded by physical constraints after the earlier measurements (i.e. not having an activity less than 0), as can be seen in Fig. 4.6. However, after all four measurements this was no longer an issue, and the profiled-likelihoods for the parameters of interest converged to a 2D Gaussian, so the two approaches are equivalent.

The resulting modelled ^{210}Pb and ^{210}Po activity over time are shown in Fig. 4.7, which shows that there is a significant contribution from ^{210}Pb. The result of this analysis after all four measurements was an initial Pb-210 activity of $28.5^{+8.3\,+9}_{-7.9\,-3}\,\mathrm{mBq/kg}$, with statistical and systematic uncertainty (from the surface-to-bulk activity conver-

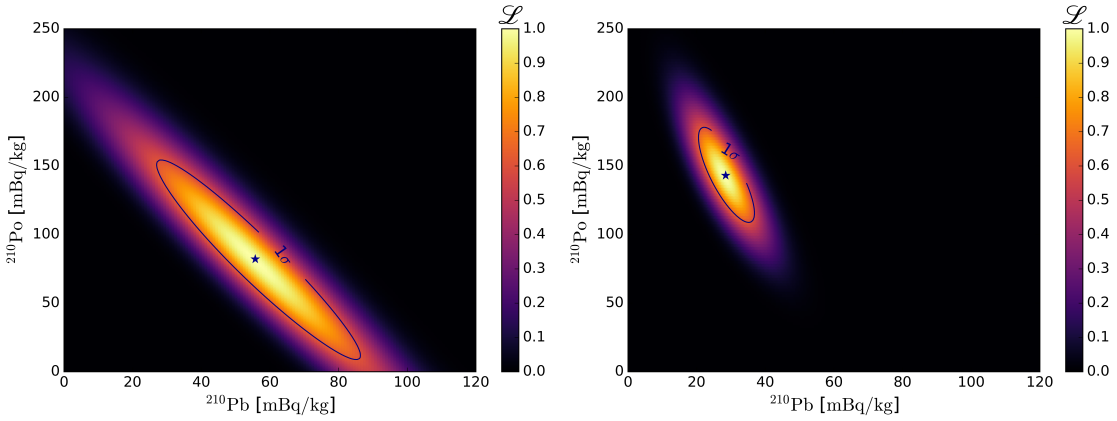


Figure 4.6: The likelihood function for initial ^{210}Po and ^{210}Pb activities after the second measurement of the C10100 copper sample (left) and after the fourth measurement (right panel), showing best-fit points and $1\,\sigma$ error contour. The likelihood values (colour scale) are normalized so that the maximum is $\mathscr{L} = 1$.

sion factor), respectively [2]. This is consistent with measurements of similar types of copper [255].

4.1.2 Copper electroplating

The level of ^{210}Pb contamination found in the copper of S140 leads to a significant background of low energy X-rays (from Bremmstrahlung interactions in the copper following β^- decays). The simulated energy spectrum of this background in S140 with a neon gas mixture – calculated using Geant4 [204] – is shown in Fig. 4.8 [1]. To mitigate this issue, the inside of the S140 vessel was electroplated with pure copper. In electroplating, an applied current supplies electrons to an anode causes oxidation reactions, with the liberated ions then drifting through an electrolyte towards the cathode, where they undergo a reduction reaction and become a part of the surface of the cathode [2].

Performed in collaboration with PNNL, the two hemispheres were plated (prior to being welded together) as described in ref. [2]. The setup – shown in Fig. 4.9 – included a hemispherical anode suspended inside the SPC hemisphere being plated (the cathode). The space in between was filled with a distilled water/sulphuric acid

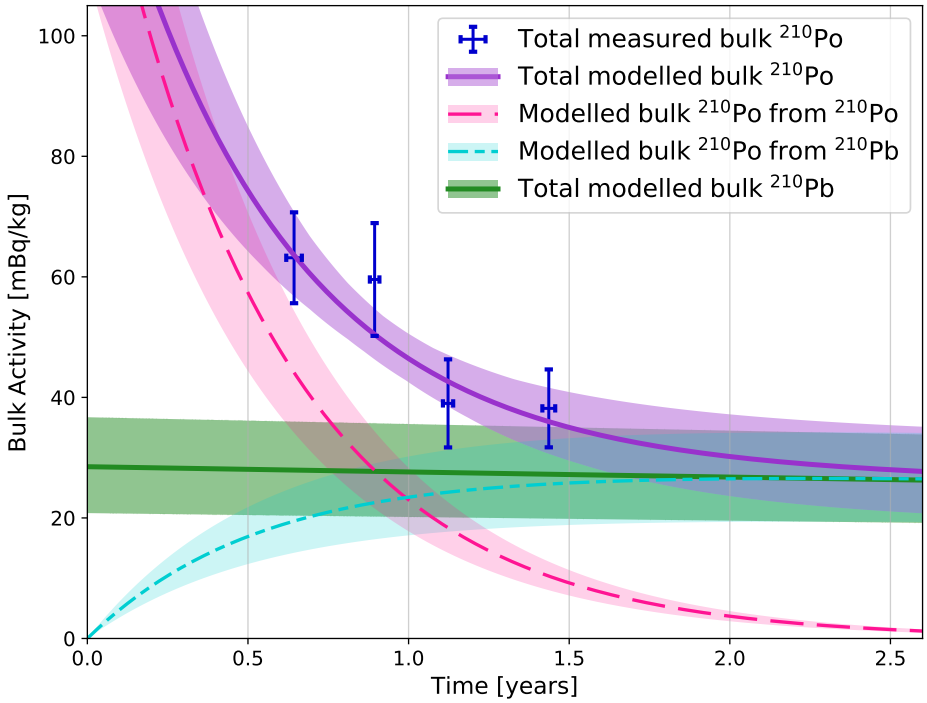


Figure 4.7: Measurements of bulk ^{210}Po over time with statistical error bars (blue data points), shown along with the modelled contributions of total ^{210}Po activity (purple), the ^{210}Po activity from initial ^{210}Po (pink dashed), and from initial ^{210}Pb (dot-dashed cyan). The modelled ^{210}Pb activity is shown in green. In all cases, the best-fit curves are shown as well as shaded $1\,\sigma$ error bands.

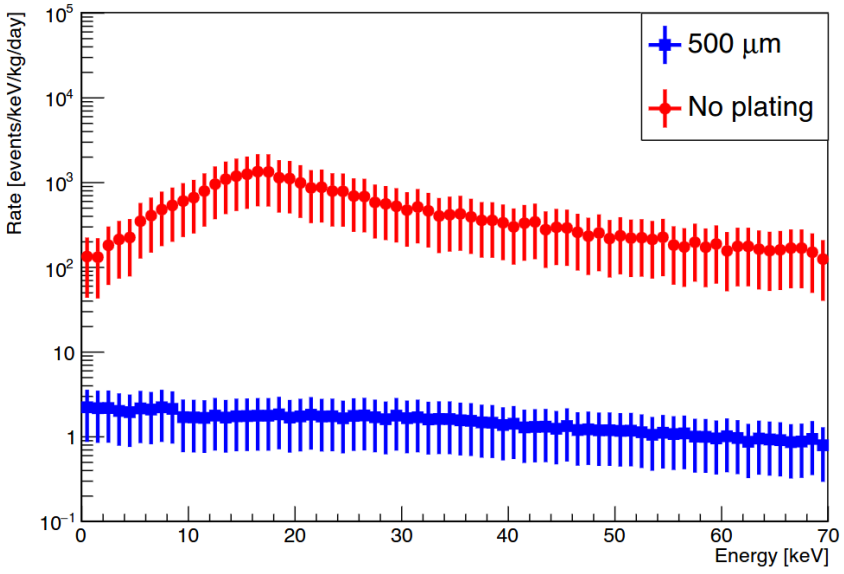


Figure 4.8: Simulated background rate in a neon gas mixture in S140 from ^{210}Pb copper contamination with and without the $500\,\mu$m of internal electroplating of the vessel. Plot taken from figure 7 of ref. [1].

electrolyte [260]. The voltage was set to be higher than the reduction potential of contaminants, such as lead ($-0.13\,\mathrm{V}$ [261]), uranium ($-1.8\,\mathrm{V}$ [262]), or thorium ($-1.9\,\mathrm{V}$ [263]). In this way, only pure copper (with a reduction potential of $+0.34\,\mathrm{V}$ [264]) was plated onto the sphere [2].

The mass of material plated m as a function of time and current I is [260]:

$$m = \frac{M \int I(t)dt}{zF} \tag{4.6}$$

where M is the molar mass of the material, and F is the Faraday constant, which is the charge of one mole of electrons. The number of electrons involved in the reduction interaction is z, which in the case of copper is 2, as the reaction is $\mathrm{Cu}^{2+} + 2\,\mathrm{e}^- \rightleftharpoons \mathrm{Cu}$ [2].

Electroplating of the S140 vessel took place at the LSM, lasting approximately 20 days per hemisphere. This deposited a layer of roughly $500\,\mu$m of pure copper [2]. Following this process, the S140 vessel was electron-beam welded together. The SPC was then etched with a hydrogen peroxide/sulphuric acid mix and pacified with citric acid twice, removing a few μm of copper each time. This was done to remove surface contamination from air exposure, and to improve the smoothness of the inner surface. The background reduction delivered by the copper plating is shown in Fig. 4.8, reducing the background rate by approximately 98 % below 1 keV [1].

4.1.3 Operation at the LSM

During spring to fall of 2019, the S140 detector was installed at the LSM for an initial commissioning phase and physics campaign. This was convenient following the electron-beam welding of the sphere in France, and took advantage of a time span where the space for the experiment at SNOLAB was not yet ready. For this temporary setup, the HDPE shield shown in Fig. 4.1 was replaced with a cylindrical-shell water tank with a thickness of approximately 20 cm, surrounding the SPC and lead shield [248].

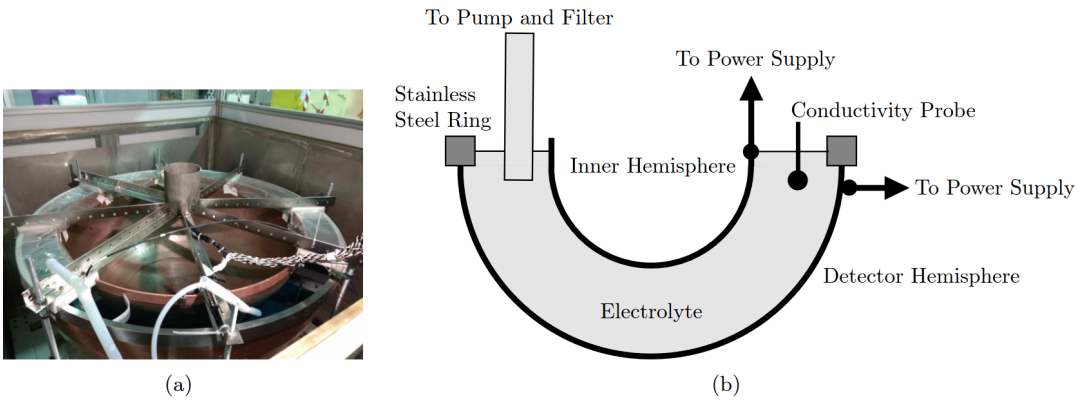


Figure 4.9: The electroplating setup for the S140 vessel with the inner hemispherical anode (a), and a schematic of the setup (b). Taken from figure 5 of ref. [260].

The 2-channel ACHINOS (as described above) was installed using the sealed sensor deployment and glove box apparatus. The gas-handling setup – depicted in Fig. 4.10 – included a liquid nitrogen cooled, carboxen radon trap [221], as well as a SAES MC700-902F "getter" [1]. The latter is intended to remove oxygen and water. The SPC was filled by passing gas through the getter and radon trap (in that order). Future procedures will likely involve re-circulation of the gas, with additional apparatus to monitor the loss of methane in the gas mixture to the radon trap [187].

The SPC high-voltage was supplied by an Iseg EHS 44 100x-K02 module [265] (up to 5 kV). A custom box with multiple CREMAT CR-110 [220] preamplifier chips were used; these are often used by NEWS-G [63], with a response decay time of $\tau = 140\,\mu$s and a signal gain of 1.4 V/pC. The calibox board was used for data acquisition (see section 3.1) [1]. A UV laser calibration system was also installed (as described in section 3.2.1), as well as a deployment port for ^{37}Ar samples.

After a commissioning phase – including data taken with argon and neon gas mixtures – a short physics campaign was undertaken. For this data, the unique opportunity to use pure methane gas (at 135 mbar) was exploited. Pure CH_4 as a target provides ample hydrogen target atoms (ideal for low-mass WIMP searches), while being relatively more transparent to some backgrounds – such as Compton

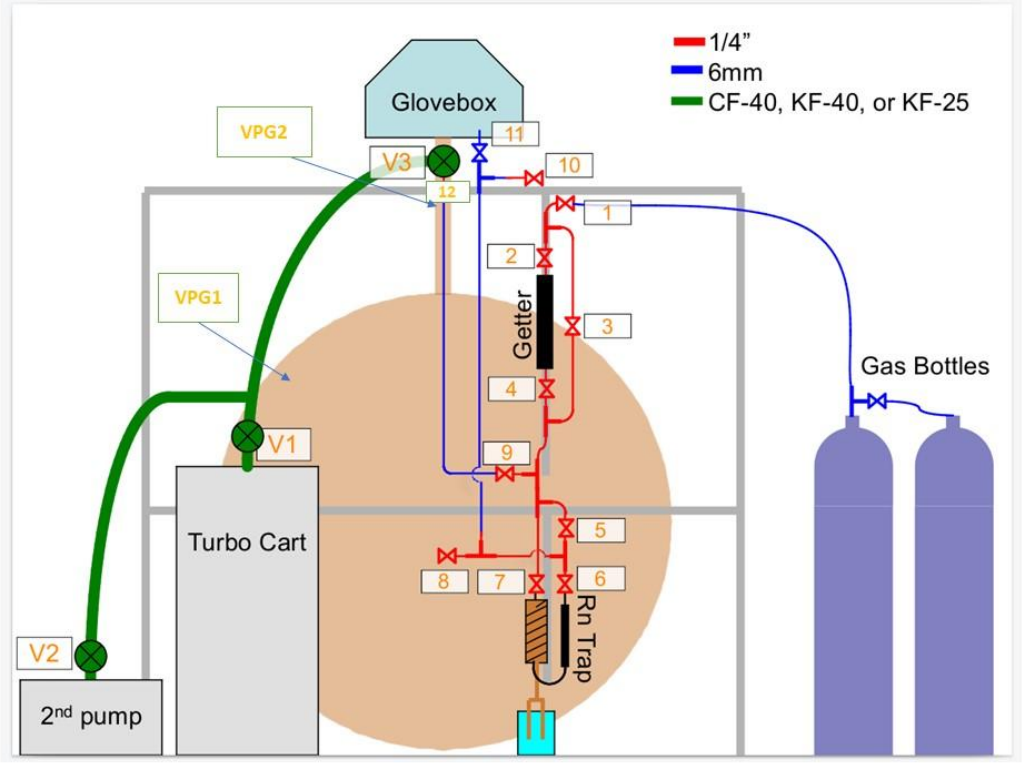


Figure 4.10: P&ID diagram of the S140 detector installation at the LSM, showing the sphere and surrounding frame structure, glove box, vacuum pumps, and gas cylinders. The plumbing included larger diameter vacuum tubing (green) and narrow, gas flow tubing (red and blue). Vacuum valves are labelled as V1 - V3, and gas-flow valves 4-12. The gas system also includes a liquid nitrogen-cooled carboxen radon trap and SAES getter filter, with bypasses for both.

scattering from γ-rays [70] – compared to more typical neon gas mixtures. The dark matter search data acquired during this campaign is described in section 4.2 and its analysis is described in sections 4.3 to 4.5.

4.1.4 Installation at SNOLAB

Following the 2019 physics campaign at the LSM, the S140 experiment was shipped (always with nitrogen cover gas) to SNOLAB, in Sudbury Ontario [61]. Compared to the LSM, SNOLAB has an additional $\sim 2\,\text{km}$ water-equivalent shielding, reducing muon fluxes even further [61, 266]. It is also a clean-lab facility, reducing the possibility of environmental contaminants being introduced into the SPC. Currently, the detector has been fully installed with the exception of some gas-handling components,

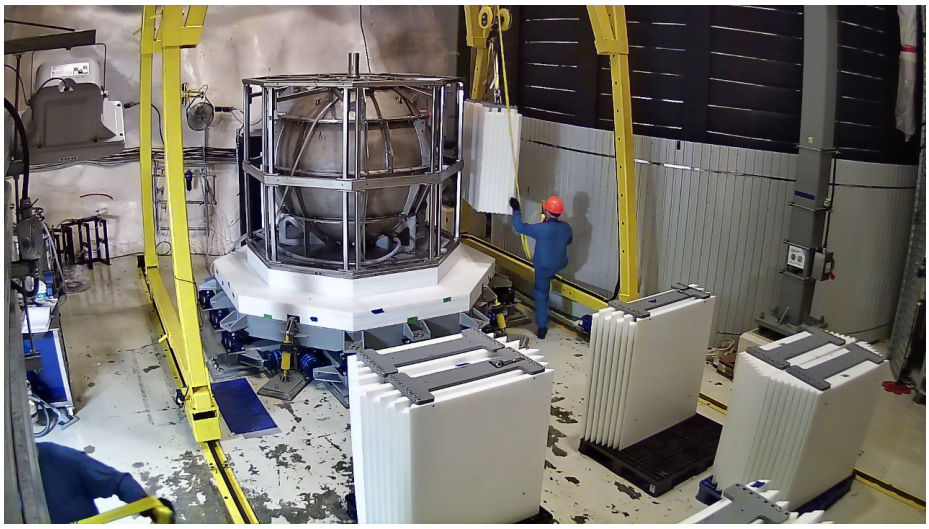


Figure 4.11: The S140 detector during installation at SNOLAB; the SPC and lead shield are in place, while the HDPE shield is being assembled.

and commissioning is underway [267]. A possible third etching of the inside of the sphere may be carried out in the near future. The setup at SNOLAB is shown in Fig. 4.11; the experiment is located in the "Cube hall", next to the DEAP-3600 and (future) PICO-500 experiments.

The first dark matter physics program at SNOLAB will be to take data with a 1 bar neon-methane gas mixture (6% methane). The total background rate expected from Geant4 simulations 1.67 ± 0.5 events/keV/day/kg (uniform in energy) [1, 152]. The expected energy threshold of the experiment – based on UV laser calibrations at Queen's University – is 0.5 primary electron equivalent energy [192]. Using these assumptions, as well as other nominal values describing the energy response of the detector, the projected SI WIMP exclusion limit of S140 with $20\,\mathrm{kd}\cdot\mathrm{days}$ of data is shown in Fig. 4.12 [268, 269]. This limit was calculated with the optimum interval method [77] (see section 2.3.2). Due to the large amount of hydrogen target material (and the typical low energy threshold of SPCs), it is possible that NEWS-G can exclude dark matter masses as low as $500\,\mathrm{MeV/c^2}$, which at the current time would be a world-leading result [8, 9].

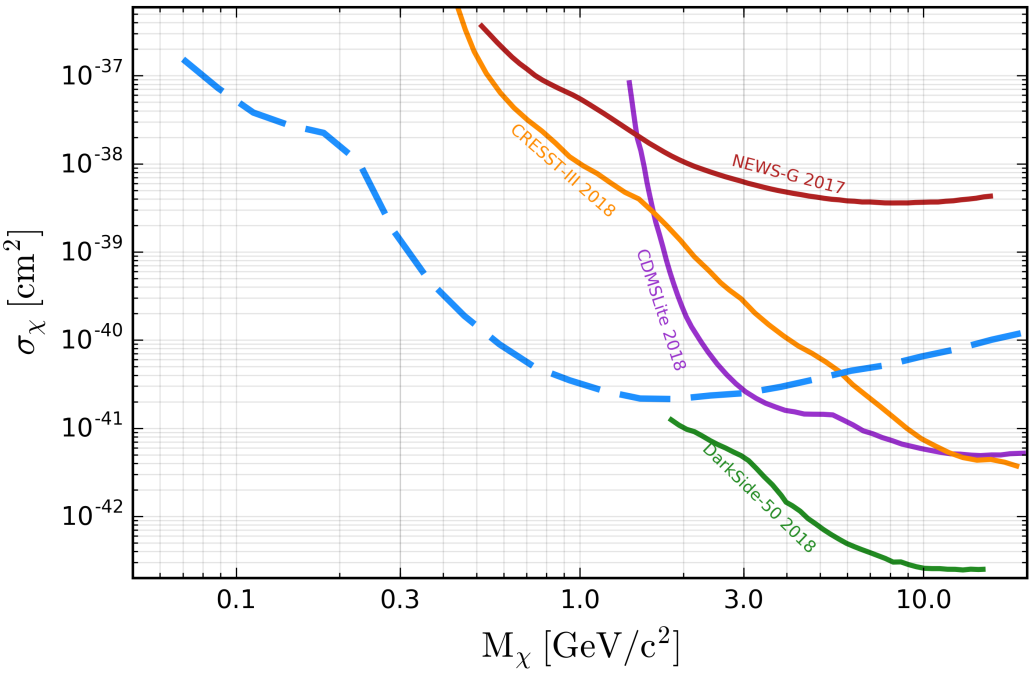


Figure 4.12: Projected sensitivity of the 140 cm SPC at SNOLAB [248]. Assumptions: a 20 kg · day exposure with Ne + 6% CH_4 gas, a 0.5 e^- threshold [192], and a flat background of 1.67 ± 0.5 events/keV/day/kg [152]. This limit is calculated using the optimum interval method [77]. Results from other experiments are shown for comparison [63, 104, 109, 270].

4.2 LSM campaign: Data

As previously introduced in section 4.1.3, a physics campaign with 135 ± 10 mbar of pure methane gas was undertaken at the LSM in Fall 2019. The S140 detector was used with a 2-channel ACHINOS sensor, which was divided into two channels for readout and dubbed "north" and "south" (with 5 and 6 anodes respectively). A voltage of 2030 V was applied to both channels, as this ratio was found to result in the lowest resolution for this sensor, thereby indicating their gains were closely matched. Unless otherwise stated, only the south channel was used to trigger physics data, and only south-channel data was analyzed. The signal acceptance for dark matter events with this fiducialization cut applied is discussed in section 4.4.3.

After inserting the sensor and support rod using the glovebox assembly (see section 4.1), the sphere was filled by passing the gas once through a getter filter to remove electronegative contaminants, and then through an activated charcoal radon trap to remove radon introduced by the getter. Data was recorded in 8 ms windows when the DAQ was triggered, with the traces being digitized at 1.042 MHz.

4.2.1 Collected data

Data was taken over the course of 10 days and alternated between 1-2 day long physics runs and 1 hour dedicated laser calibration runs each day. Additional calibration data was collected towards the end of the campaign with an Am-Be neutron source deployed close to the detector, as well as ^{37}Ar injected inside the SPC at the very end of the campaign. The DAQ was triggered with a trapezoidal filter (essentially based on the derivative of the raw signal trace). In initial physics runs, the trigger was defined by requiring an 80 ADU rise over 25 ms and after each pulse, a 4 ms deadtime was imposed. Part way through the campaign, the trigger was changed to 50 ADU over 15 ms (a slight improvement) with a longer deadtime of 10 ms.

For dedicated laser calibration runs, the SPC is triggered using the photodiode

signal (see section 3.2.1) in order to also record events where no laser photoelectrons reach the sensor. This trigger was set to 30 ADU over 50 ms with a 100% efficiency. For these laser calibration runs, the laser intensity was generally tuned such that on average $0 - 7$ photoelectrons were produced by each pulse, by changing the laser pump current. The pulse rate was 10 Hz for these runs. However, the laser was also operated during the physics run to monitor changes in detector response over time. In these cases, the pulse rate was reduced to 5 Hz, but the intensity increased (to a pump current of 140 A) to produce events with $\mathcal{O}(100)$ photoelectrons.

The campaign data (including run types, trigger conditions, etc.) is summarized in table 4.2. Other datasets were taken that are not listed in this table (such as ^{37}Ar at different anode voltages), but all data used in this work are included. The run naming scheme used by NEWS-G is as follows: 1) the first letter represents the year, 2) the second letter gives the month, 3) two numbers for the date, 4) a letter code identifying the detector ("s" is for the S140 device), and 5) three numbers as a run identifier for that day. For example, the third run in S140 on October 5$^{\text{th}}$ 2019 would be listed as tj05s002.

Before proceeding with any analyses, the LSM campaign physics data was partitioned into "test" and "blind" portions. The test data was used for all subsequent analyses, whereas the blind data was kept sequestered until the analysis is finalized. As of the date of writing (October 2023), the blind data remains untouched. The data was partitioned by first removing all laser events (using their photodiode amplitudes, with presumed 100% efficiency). Additionally, all α events and the subsequent 5s of data were removed (more on this cut in section 4.4.4). These events were removed from the physics runs and combined together in separate data files. The test data was then defined to include all of run tj04s002 (which had been used for preliminary analyses), and 20% (randomly selected events) of all other physics runs. In total, this represented about 27% of the total exposure of the physics runs, with the remainder being kept blind. Because the data was partitioned by randomly selecting events

Run	Duration [h]	Type	Trigger	Laser [A / Hz]
ti30s002	1	Laser	Photodiode	113 / 10
ti30s003	46.5	Physics	Cond. 1	140 / 5
tj02s000	1	Laser	Photodiode	115 / 10
tj02s001	21	Physics	Cond. 1	140 / 5
tj03s000	0.5	Laser	Photodiode	120 / 10
tj03s001	0.5	Laser	Photodiode	130 / 10
tj03s002	46.5	Physics	Cond. 1	140 / 5
tj04s000	1	Laser	Photodiode	120 / 10
tj04s002	21	Physics	Cond. 2	140 / 5
tj05s000	1	Laser	Photodiode	120 / 10
tj05s001	20.5	Physics	Cond. 2	140 / 5
tj06s000	1	Laser	Photodiode	120 / 10
tj06s001	23	Physics	Cond. 2	140 / 5
tj07s000	1	Laser	Photodiode	120 / 10
tj07s001	22.5	Physics	Cond. 2	140 / 5
tj08s000	1	Laser	Photodiode	120 / 10
tj08s001	23	Physics	Cond. 2	140 / 5
tj09s000	1	Laser	Photodiode	120 / 10
tj09s001	26.5	Physics	Cond. 2	140 / 5
tj10s000	1	Laser	Photodiode	120 / 10
tj10s001	19	Physics	Cond. 2	140 / 5
tj11s002	1	Laser & Am-Be	Photodiode	120 / 10
tj11s004	1	Laser & Am-Be	Photodiode	120 / 10
tj11s005	7	Am-Be	Cond. 2	140 / 5
tj12s003	3.5	^{37}Ar	Cond. 2	140 / 5
tj13s000	4	^{37}Ar	Cond. 2	140 / 5
tj14s001	4	^{37}Ar	Cond. 2	140 / 5

Table 4.2: Data collected during the LSM physics campaign. Trigger conditions are defined in the text. The laser was running at all times.

throughout each run, the test and blind datasets should have the same properties (i.e. the same proportion of different event populations). The procedure and checks for unblinding the physics data are discussed in section 4.5.4.

4.2.2 Pulse processing

The LSM data was processed using the DD2 method described in section 3.1.5 [160]. For the physics runs, the integration window was fixed instead of being defined using a trapezoidal filter. Within the 8 ms event window, the double deconvolved pulse was integrated from 3418 to 4656 μs (a duration of 1238 μs), roughly centered within the overall event window. This was necessary for the LSM analysis due to the significantly delayed arrival time of many electrons, slowing drifting through the large SPC and subject to greater-than-normal diffusion. If the integration window was determined automatically as usual, the algorithm would likely miss many delayed electrons, and result in integration window sizes varying significantly from event to event. This would make the amount of baseline noise integrated in each event inconsistent. Therefore this large window was chosen to (likely) contain all primary electrons. An example of an ^{37}Ar pulse (raw and treated) is shown in Fig. 4.13.

As mentioned in section 3.1.5, the LSM analysis was unique in that the diffusion experienced by primary electrons was large enough that individual primary electrons produced distinct, separable pulses. This was only the case when dealing with the treated pulses (specifically the double-deconvolved traces), as the pre-amplifier response still caused significant overlap in the raw pulses, as can be seen in Fig. 4.14.

The ROOT library TSpectrum (specifically `TSpectrum::SearchHighRes`) was used to identify potential peaks in the double-deconvolved trace [271–274]. All candidate pulses with heights exceeding an absolute threshold of 0.48 e^-/ns were kept. The double-deconvolved trace was then fit with a function including a flat baseline, and a Gaussian distribution for all peaks identified. In this analysis, the central times and heights of each Gaussian were fixed using the TSpectrum results. The baseline

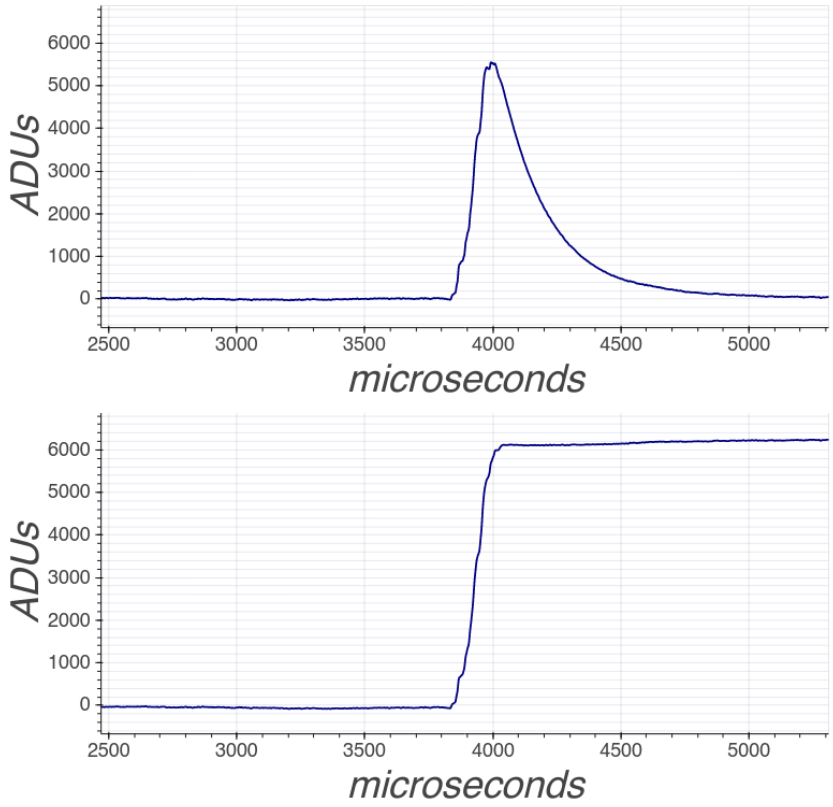


Figure 4.13: Example of an ^{37}Ar event from the LSM campaign, showing a raw pulse (top panel) and the subsequent double-deconvolved, integrated pulse (lower panel). Note that the NBT correction detailed ahead in section 4.2.3 has already been applied.

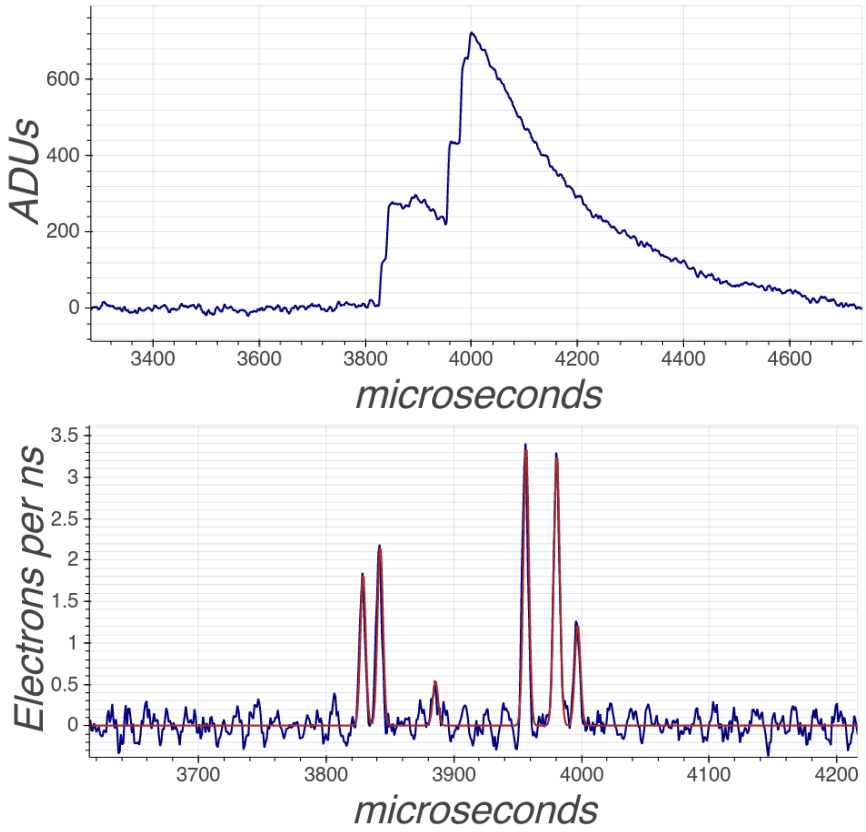


Figure 4.14: Example of a laser event from the LSM campaign (tj04s000 specifically), showing a raw pulse (top panel) and the subsequent double-deconvolved (lower panel). The PF algorithm result for this event is overlaid in red. Note that the NBT correction detailed ahead in section 4.2.3 has already been applied.

and width of each peak were left as free parameters in the fit, which was performed using ROOT TMinuit2 [259]. The peak widths were constrained to be no more than 6 samples (slightly less than 6 μs) to prevent the algorithm from occasionally fitting long, slow baseline fluctuations with very wide Gaussians. Typical, physical peaks have widths of approximately 2-3 samples. An example of an LSM low-intensity laser event (from tj04s000) is shown in Fig. 4.14, along with the resulting double deconvolved pulse and peak-finding result.

The threshold for the PF algorithm (0.48 e^-/ns) was chosen to balance between sensitivity to small primary electron avalanches (where the distribution of peak heights will be nearly exponential), vs over-sensitivity in misidentifying baseline noise as physical peaks. The latter was estimated by applying the PF algorithm with varying thresholds to pre-trace laser data (the first 1238 μs) of each event in tj04s000. In this way, the only peaks that are found are likely to be false positive (FP) noise peaks. This trend is shown in Fig. 4.15, which motivated the choice of 0.48 e^-/ns. The ability to directly count primary electron peaks offers better estimators of energy and the radial position of pulses; the time separation between the first and last peak (sometimes denoted as Δt_{peak}) has similar properties to the pulse risetime, but with better discrimination capability for few-electron events. These variables were used extensively for the LSM methane analysis presented later throughout this chapter.

Another variable calculated using the processed LSM data was the drift time of laser events. This was calculated as the time difference between the leading edge of each raw photodiode pulse, and the first primary electron peak identified in the double deconvolved SPC pulse. Note that other similar definitions are used in other works, e.g. in ref. [275]. An example of a laser event with the definition of its drift time is shown in Fig. 4.16. Typical drift times in the LSM data ranged from approximately 1200 μs to 1550 μs depending on prevailing space-charge conditions (see section 4.2.5), and as low as 700 μs following α-decay events (see section 4.4.4).

The DD2 processing code library (`quadis` [160]) saves the calculated variables in

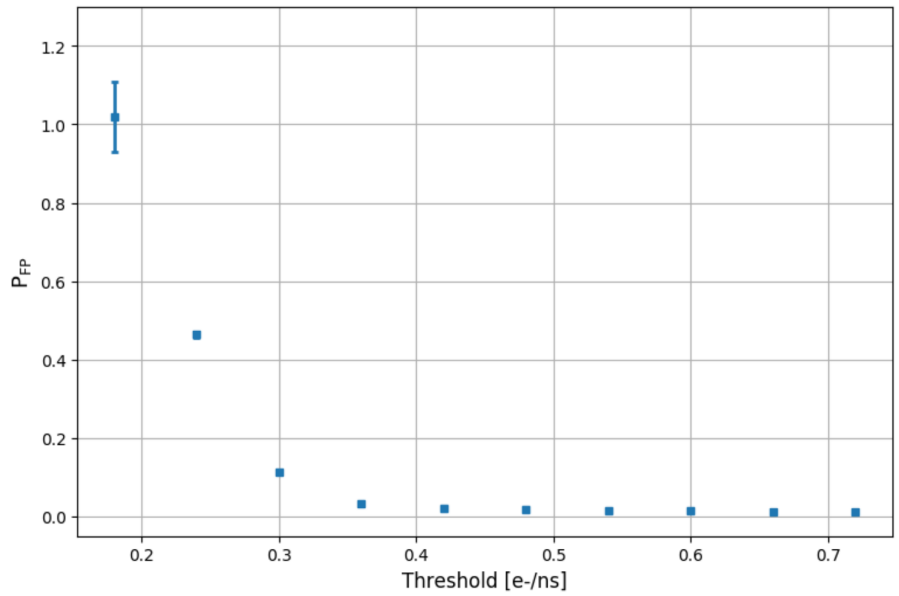


Figure 4.15: False positive peak identification probability (P_{FP}) as a function of PF algorithm threshold, as assessed with pre-trace data.

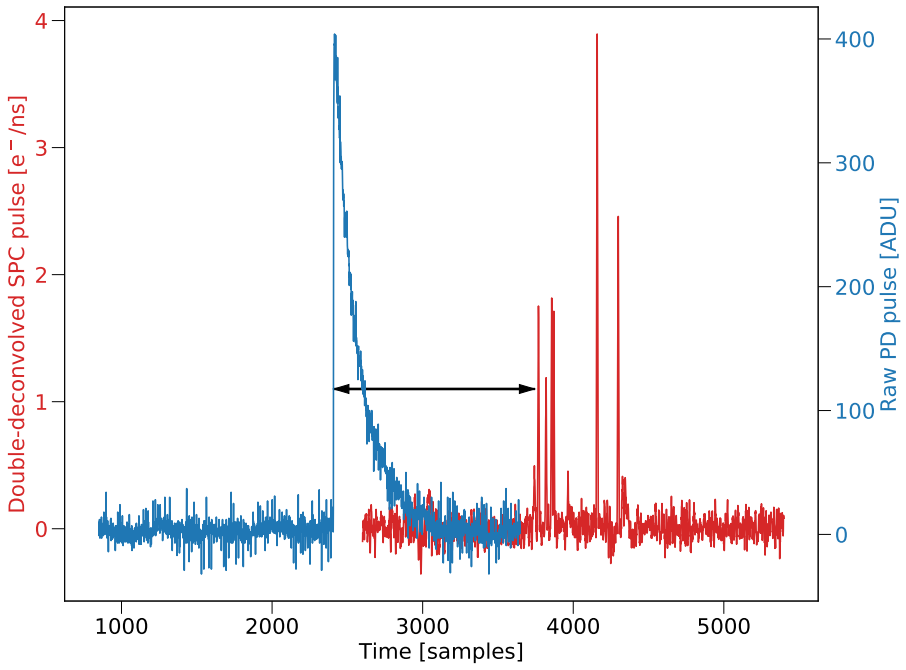


Figure 4.16: A low-intensity laser event (from tj04s000) showing the double-deconvolved south channel SPC pulse (red) and the raw photodiode (PD) pulse (blue). The arrow indicates the drift time for this event, which is 1332.5 μs in this case. Note that there is a time offset between the two traces because this corrects for an offset in the DAQ between the two channels.

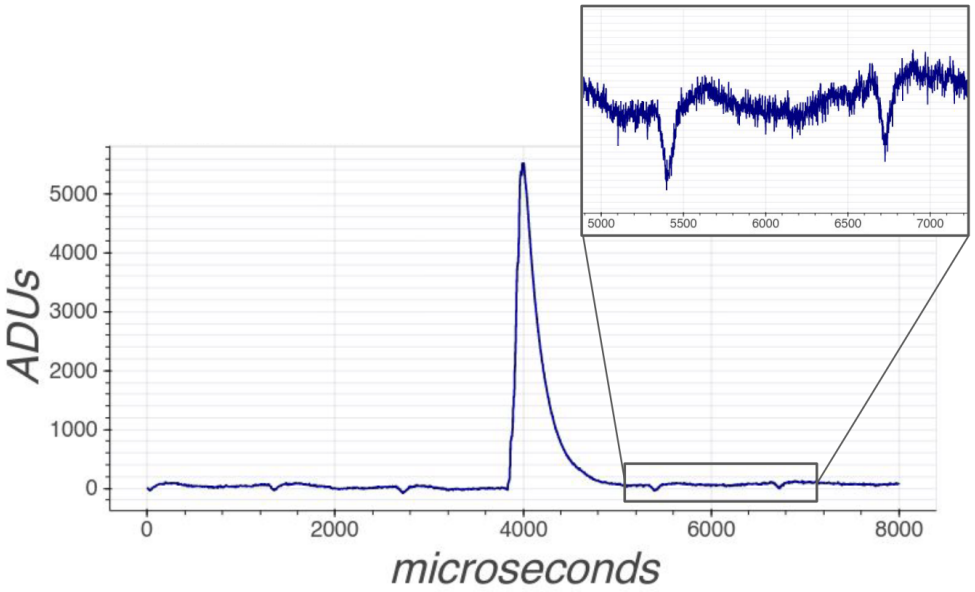


Figure 4.17: A raw ^{37}Ar pulse that has not been corrected for NBTs, with an inset to zoom in on two NBTs. The corrected raw pulse for this event is shown in Fig. 4.13.

ROOT TTree data files. All data files were also converted into Python dictionary files, which served as the primary data structure for analyses. This includes the regular DD2 method results, PF-calculated variables, and specially-calculated variables such as drift time. While both data structures are well-suited to multivariate analysis, the use of Python includes a host of useful code libraries for this analysis including `scipy`, `emcee`, and so on.

4.2.3 Negative baseline transient correction

The LSM campaign data was afflicted by an unfortunate, unusual type of electronic noise that was prominent only in the south channel. On top of typical baseline fluctuations (across a wide range of frequencies), there were periodic, large exponential downward fluctuations in the baseline. These occurred with a period of roughly 1.3 ms, so multiple were present in each 8 ms event window. This species of noise was dubbed "negative baseline transients" or NBTs, and were likely caused by the charging up and discharging of some detector component, although this is speculative. Further investigation into their cause is not possible, as they were only observed

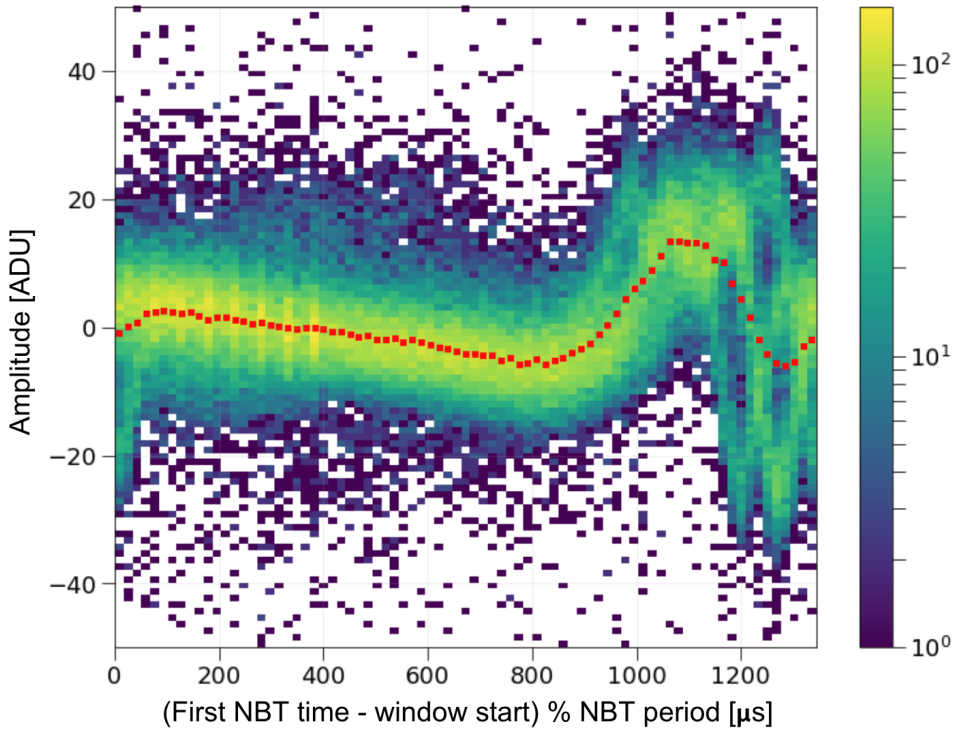


Figure 4.18: Pre-trace amplitude vs. time for all events in tj04s000, aligned by NBT phase (2D histogram with the colour scale giving counts per bin). The red markers represent the mean amplitude in individual slices.

during the LSM campaign. An example of this type of noise in an event is shown in Fig. 4.17. The amplitude of the NBTs was sufficiently large that they could distort the amplitudes calculated for events.

To demonstrate/characterize this effect, the LSM data was specially processed using an integration window at the beginning of each pulse, rather than at the centre. NBTs occurring in this region do not overlap with physical SPC signals, so the amplitude of the NBTs alone can be calculated. The time of the minimum value of the raw trace in this window was used to estimate the location of the first NBT within this pre-pulse window. The phase of the NBT can then be determined by subtracting the start time of the pre-trace window (slightly after 0 ms). Fig. 4.18 plots this quantity modulo the average period of the NBTs (approximately 1.3 ms) to approximately align the time scales of all events in laser run tj04s000. In this way, the amplitude distortion as a function of time due to NBTs can be seen. The maximum distortion

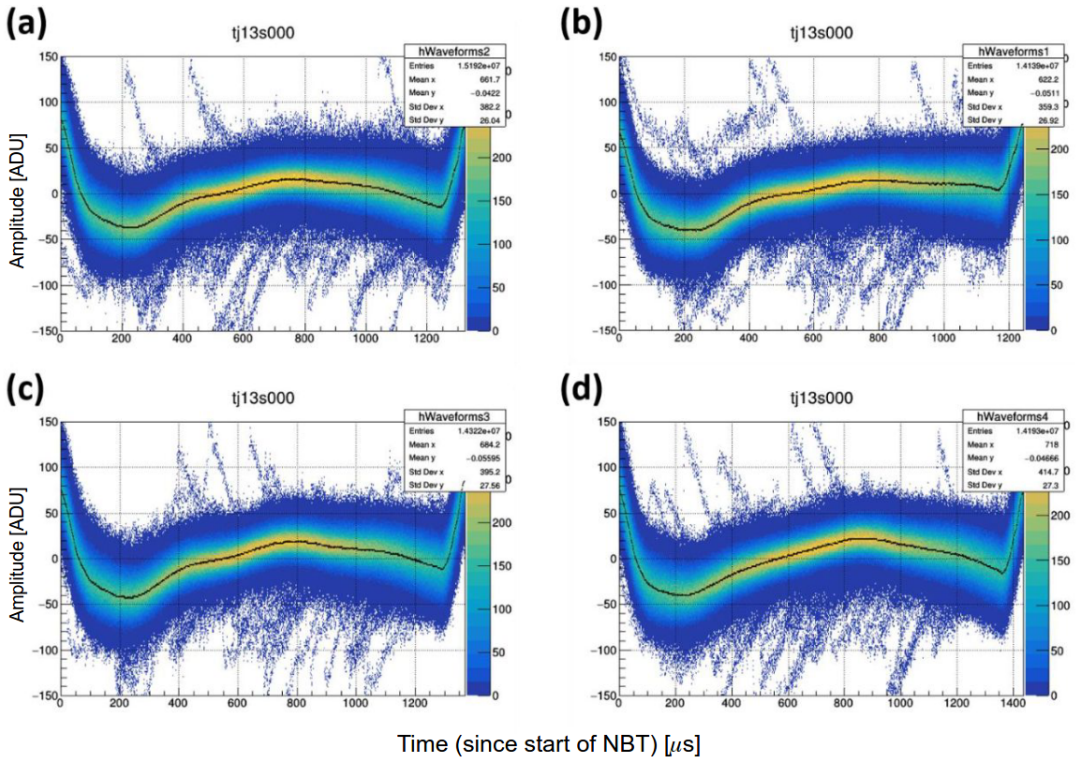


Figure 4.19: The four (inverted) NBT templates (black curves) based on tj13s000 data (2D histograms), separated by the cycling NBT period. Credit to A. Padawer-Blatt [276].

of approximately 30 ADU is small on the scale of some events (e.g. ^{37}Ar, with pulses typically around 8000 ADU), but it is significant compared to the amplitude of few-electron events. Fig. 4.18 also demonstrates that the NBTs cyclic noise causes (small) distortions at all times.

To mitigate the effect of NBTs, an algorithm was developed by A. Padawer-Blatt to precisely locate them within each event window, and then correct the raw traces for them [276]. First, the NBTs were identified by applying the `scipy.find_peaks` method to the inverted, smoothed raw pulses. The NBTs were found to cycle through four different periods of repetition, such that the overall period of the whole pattern was approximately 5.4 ms. Evidence of this behaviour can be glimpsed in Fig. 4.18. Four templates were then defined to match the four phases of the NBT cycle (including the time periods in between the visible excursions), by averaging all applicable time

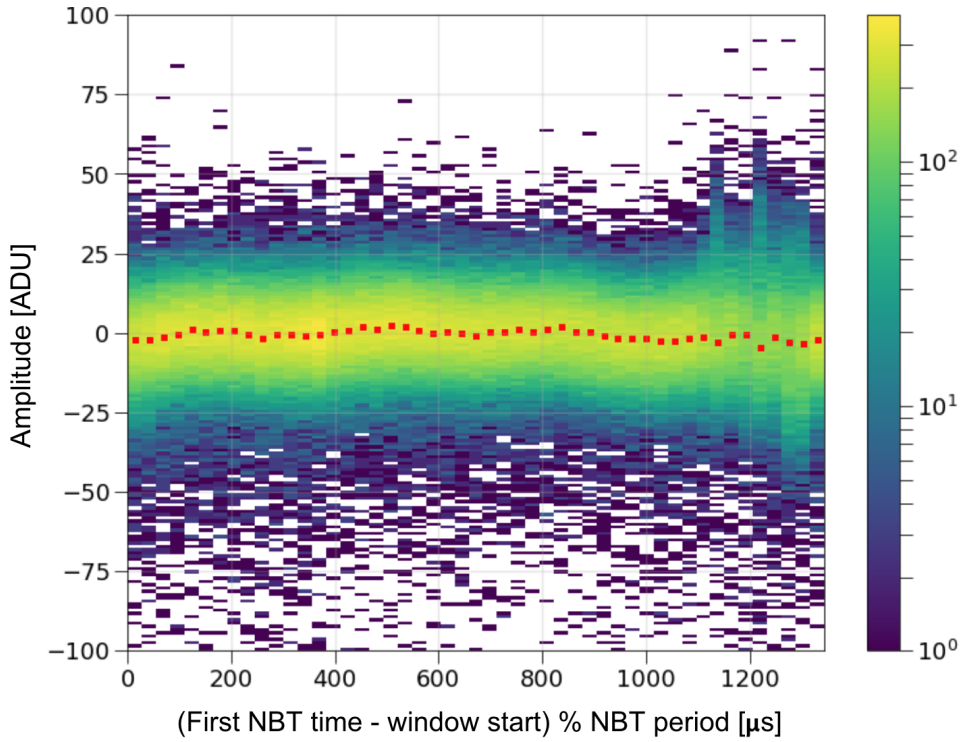


Figure 4.20: NBT-corrected, pre-trace amplitude vs. time for all events in tj04s000, aligned by NBT phase (2D histogram with the colour scale giving counts per bin). Th red markers represent the mean amplitude in individual slices. This data is the same as that shown in Fig. 4.18.

segments in the events of run tj13s000. These templates are shown in Fig. 4.19. The raw traces for all events in the complete LSM campaign were then corrected for NBTs by subtracting the appropriate templates (time-aligned) from each event. The corrected data was treated as described in section 4.2.2. The effect of this correction on the amplitude distortions shown in Fig. 4.18 is seen in Fig. 4.20.

Of course the NBT correction itself may not be perfect and is known in rare cases to produce false, artifact peaks in the double-deconvolved data. The frequency of such pathologies was estimated by comparing the number of peaks found in the original and NBT-corrected data, using laser run tj04s000 as an example, as shown in Fig. 4.21. In this case, approximately 3.9% of events gained one or more peaks after the NBT correction, while 3.7% lost peaks (and never more than two in either direction). Since these rates are roughly equal, the impact of NBT-induced peaks is

not of concern. Fig. 4.21 also shows the total peak distribution of tj04s000 with and without the NBT correction, which was largely unchanged. This demonstrates that the templates defined using the run tj13s000 are applicable to the NBT noise in other runs.

4.2.4 Signal "cross-talk"

A unique phenomenon observed in the LSM experiment was signal "cross talk" between the north and south channels of the ACHINOS sensor. While this term sometimes refers to the induction of a signal on a cable by another (or between any two electronic channels), in this case, it refers to signals induced on the opposite channel from where a Townsend avalanche occurs in the SPC. For example, if a primary electron produces an avalanche at one of the 5 north-channel anodes, a smaller, inverted signal is induced on the south channel by the drift of the avalanche ions. This is clearly observed in the double deconvolved traces of events, such as the example shown in Fig. 4.22. The inverted cross-talk signal was typically $\approx 20\%$ of the pulse height of the original signal and was observed for both north and south channel events. The strength of this cross-talk seemed to be invariant with the anode where the progenitor avalanche was taking place.

Because this phenomenon takes place in the gas volume of the SPC, it can serve as the basis for a useful PSD variable to reject non-physical events [275]. In particular, electronic discharge (or "spike") events are thought to originate from small electrical shorts within the sensor/rod assembly. Since these events do not produce a Townsend avalanche, they also do not produce a cross-talk signal. The application of this phenomenon as a PSD variable is described in detail in section 4.4.5.

4.2.5 Space charge effects

The term "space-charge effect" is often used to describe any changes in SPC conditions caused by large fluxes of avalanche ions in the sphere, and could include the screening

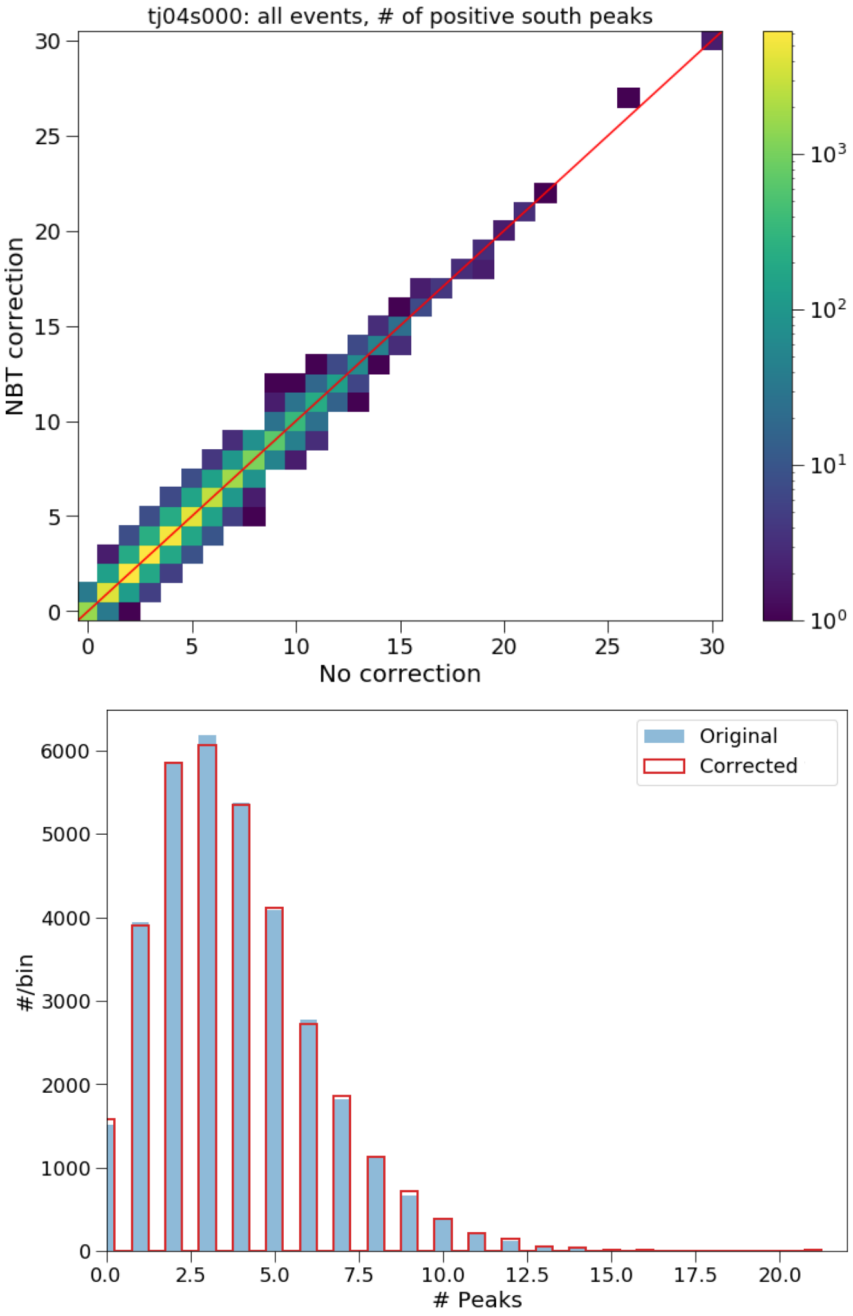


Figure 4.21: Top: A comparison of the number of positive, south-channel peaks found in laser run tj04s000 with and without the NBT correction. Bottom: the total distribution of the number of peaks in tj04s000, with and without the NBT correction.

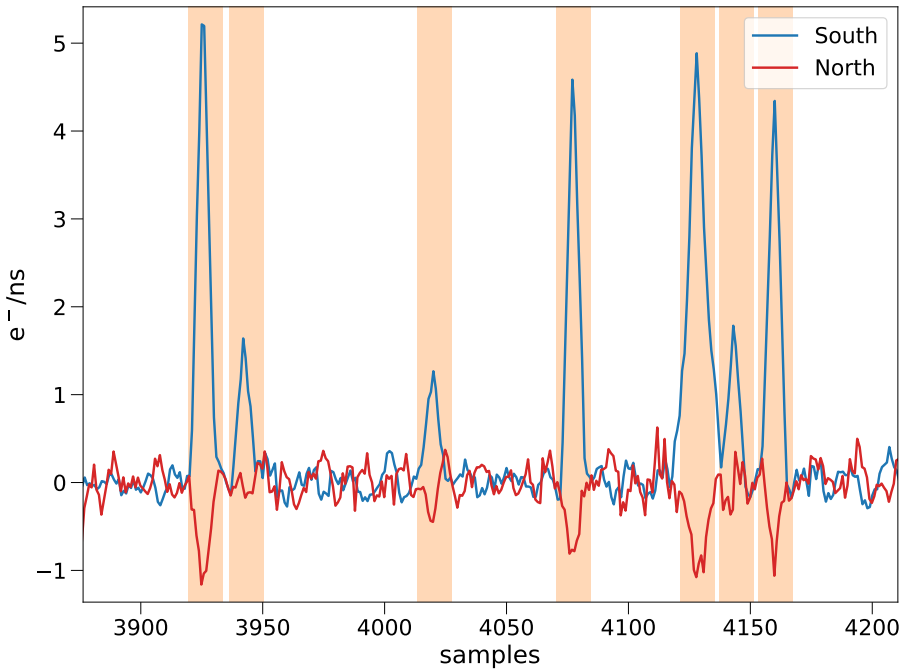


Figure 4.22: An example of a double deconvolved low-intensity laser event, showing the traces for both the north and south channels overlaid, demonstrating signal cross-talk. The highlighted bands show the regions identified as having peaks by the PF algorithm. The time-scale is in digitizer samples ($1\,\mu$s = 1.04167 samples). Taken from figure 2 of ref. [248].

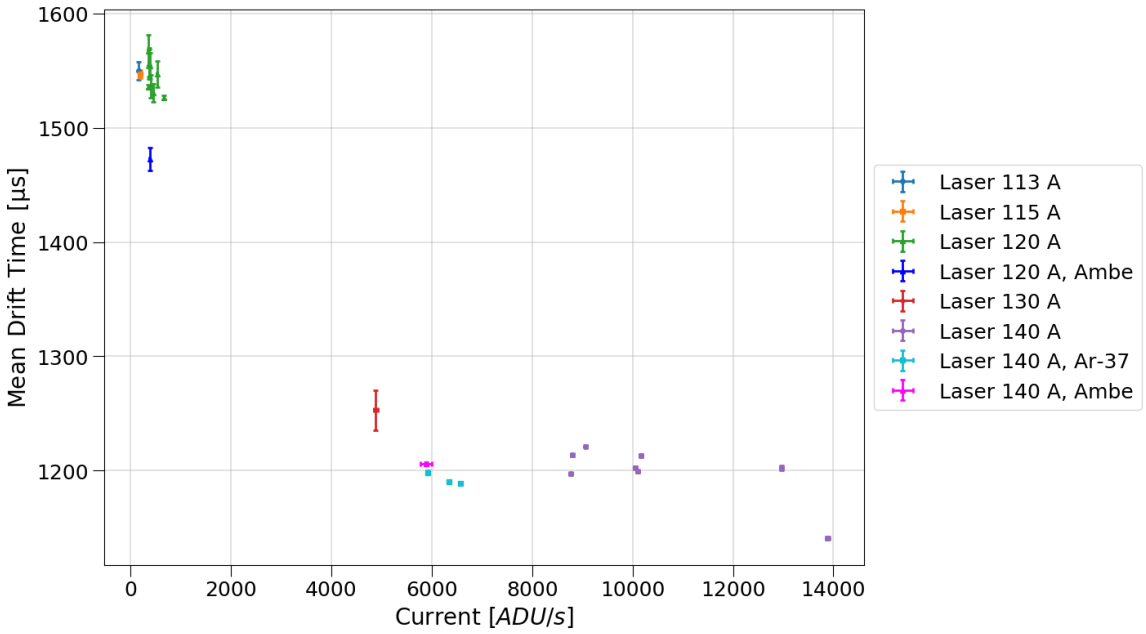


Figure 4.23: Mean drift time of all LSM campaign data sets listed by their laser intensity and other calibration sources, as a function of their average amplitude current (ADU/s).

of incoming primary electrons or changes in drift time. In this analyses, it refers specifically to the latter effect. The LSM campaign was particularly affected by these phenomena, possibly due to the large nominal drift times in the methane gas in the large-SPC environment of the experiment. The most dramatic impact of space charges on drift time is the sharp drop in drift time induced by ^{210}Po α decays [275], as discussed in section 4.4.4. The large flux of secondary ions produced ($\mathcal{O}(10^7)$) can cause a roughly 30% drop in drift time. However, lower energy events also had a large influence on the drift time in the SPC, especially when there was a constant source of events, such as high-intensity laser pulses. The presence of ^{37}Ar or Am–Be sources also had an impact.

To elucidate the effect of coherent space charge influences from all sources on the SPC, a quantity called "amplitude current" was defined as the average pulse amplitude for all events in second-long time windows. Drift time as a function of this "current" is shown in Fig. 4.23, which shows that drift time does decrease with increasing

current as expected, until plateauing for all high-intensity laser runs (physics runs), ^{37}Ar, and Am–Be data. A possible explanation for this plateau is that avalanche ions produced by laser events will entirely affect the path of future incoming laser electrons, and so will have an out-sized impact on drift time. Further work to understand and model these processes is being undertaken by other graduate students of the NEWS-g collaboration.

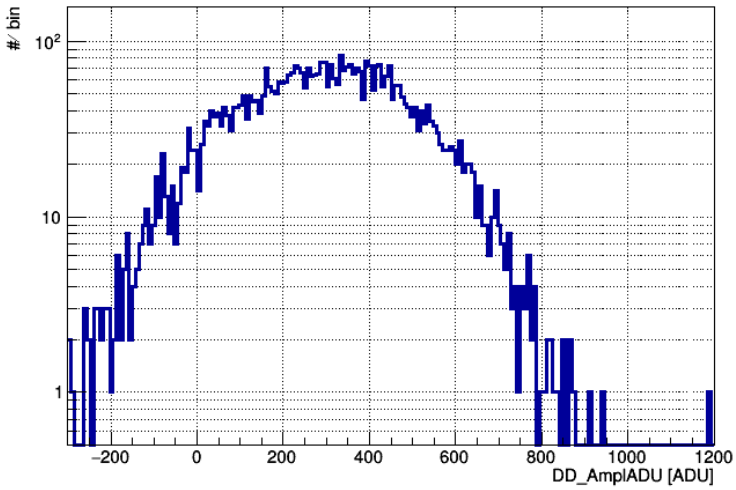


Figure 4.24: Amplitude spectrum of run ti30s002 processed with a fixed integration window of 1290 samples.

4.3 LSM campaign: SPC characterization

To determine the sensitivity of the S140 experiment at the LSM to a hypothetical dark matter signal, it is necessary to know how that signal would be observed by the SPC. This includes all physics pertaining to energy deposition and observation in the detector, as well as electron transport of volume events. These facets of the detector response are drawn from existing literature results in some cases, but for the most part, are the result of calibrations in the S140 detector.

4.3.1 Low-intensity laser calibration

As in previous analyses (such as in section 3.3), the primary application of the UV laser calibration data was to model the avalanche response of the detector, including the mean gain $\langle G \rangle$ and θ parameter of the Polya distribution [192]. This was done using the low-intensity calibration data listed in Table 4.2.

In typical laser analyses (such as that described in section 3.2.1), the pulses would be processed with a fixed integration window chosen to be wide enough to contain all photoelectrons, but not overly so, as this would integrate unnecessary baseline noise.

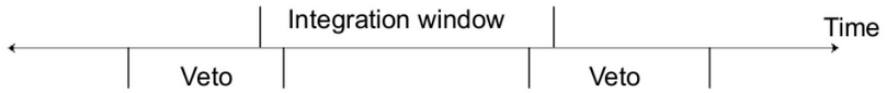


Figure 4.25: Cartoon depicting the post-processing of laser events; 60 sample integration windows are kept if they have any number of peaks fully inside the integration window, with no other peaks found inside the 25 sample veto windows before or after (these have a 10 sample overlap with the integration window).

In the case of the LSM analysis, the substantial diffusion of electrons in the gas would require an integration window more than 1 ms long to contain all photoelectrons. In doing so, a large amount of baseline noise affects the treated pulse, severely distorting the amplitude values calculated for laser pulses. An example of this is shown in Fig. 4.24. Rather than a spectrum with a resolved $0\,e^-$ peak and exponential tail such as in Fig. 3.31, the large amount of baseline noise washes out all features.

To circumvent this problem, the large 8 ms acquisition windows of each event were divided into many smaller sub-windows wide enough to only contain a few electrons, while integrating much less baseline noise. To choose acceptable sub-windows, the laser data was processed with the PF algorithm with a lower threshold than usual (0.24 instead of $0.48\,e^-$/ns) to detect any peaks on the border of the sub-window that might lead to a mis-reconstruction. A higher rate of FP peak identification is not a concern in this case. Specifically, the following procedure was used:

1. Each 8 ms event was partitioned into 60 sample sub-windows to avoid integrating too much baseline noise,

2. Sub-windows were rejected if there were any peaks found in 25 sample veto windows before and after the sub-window. These veto windows have 10 samples of overlap with the sub-window (see Fig. 4.25),

3. The data was reprocessed using the usual double-deconvolution processing method [160] repeatedly, with the integration window corresponding to all possible sub-windows (using the NBT-corrected traces (see section 4.2.3).

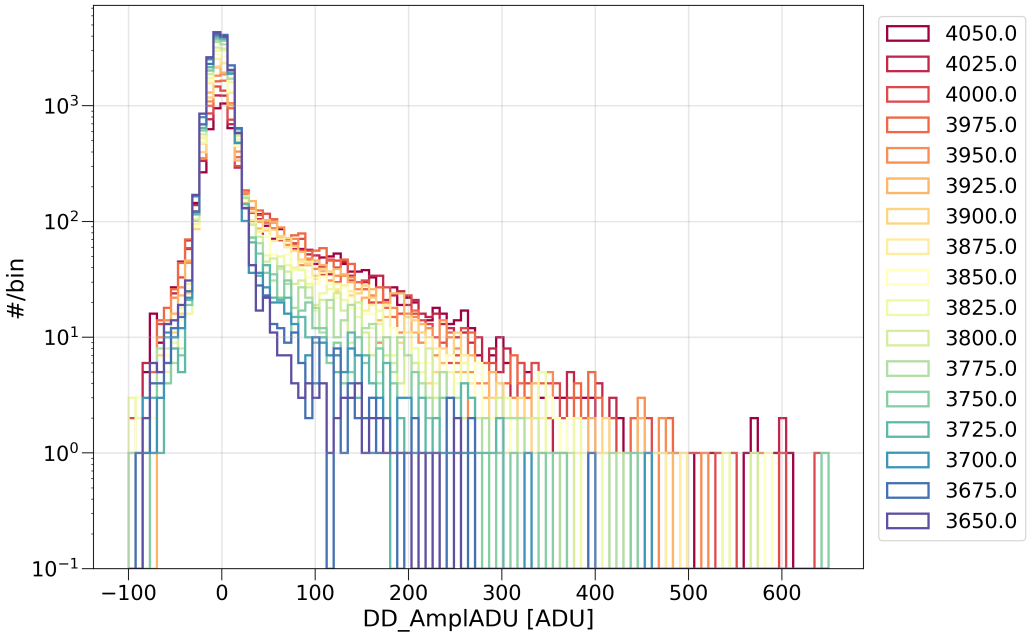


Figure 4.26: Amplitude (`DD_AmplADU`) spectra for sub-events with different window starting times (in samples), showing the changing fraction of null events to non-empty events.

With the PF-based vetting of sub-windows, a new dataset of usable laser events was extracted from each run. Despite using a lower PF threshold than usual, some problematic cases may still be missed. For example, if a peak exists in the early veto window of Fig. 4.25 that is too small to be detected by the PF algorithm, the baseline of the double-deconvolved pulse will be calculated incorrectly, leading to a distorted amplitude for the signal inside the sub-window. These pathologies – false negative veto (FNV) events – could be further reduced by still lowering the PF threshold used for this processing, but this would result in an unusably small fraction of laser sub-windows surviving the selection (most being not true FNV events). The only additional cleaning cut applied to further mitigate this problem is a cut on the slope of the pre-integration window double deconvolved trace. Sub-events with peaks in the pre-veto window (FNVs) will have unusually high or low slope values. Additionally, post α-decay time periods were removed from the datasets (see section 4.4.4).

The amplitude spectra of data with different sub-window starting times are shown

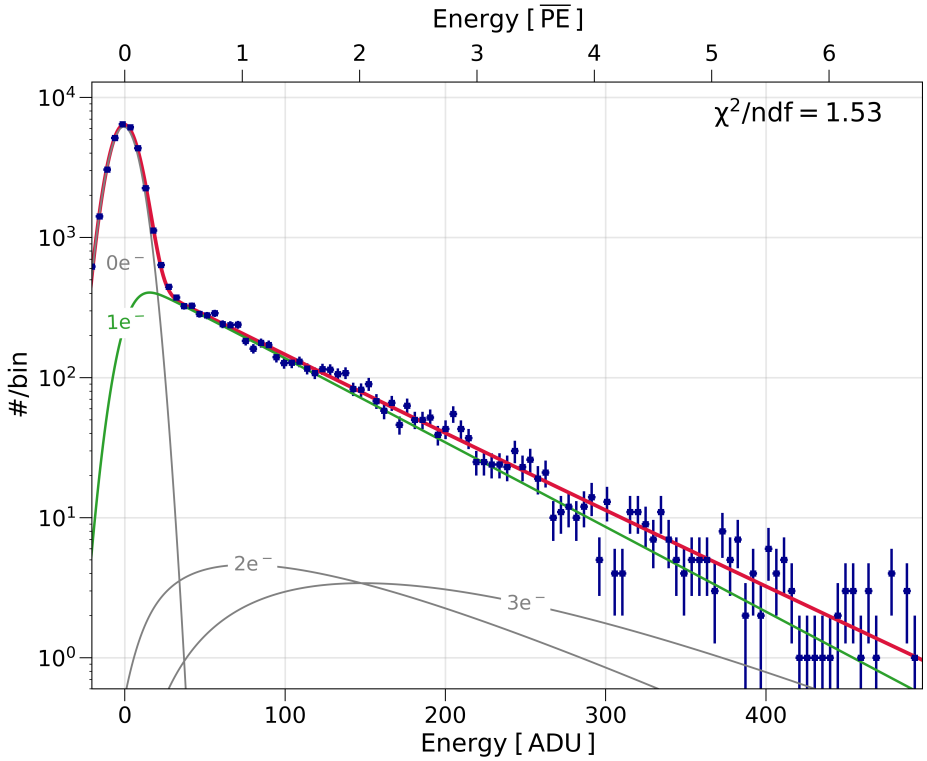


Figure 4.27: Example of the fit of a combined amplitude spectrum of low-intensity laser calibration data (blue histogram), tj04s000 specifically. The total model (red) as well as components for different numbers of electrons are shown (green and grey curves).

in Fig. 4.26. The changing fractions of null events to events with electrons is due to the mean drift time of laser electrons being centred near the middle of the 8 ms total event window - hence sub-windows starting near 4000 samples are likely to have more electrons. To make a combined amplitude spectrum, data with all sub-window starting times was included, except for some early/late data sets with high fractions of null events. This was done to achieve a fraction of non-empty events of approximately 0.1 to 0.5, which is typically best to give a strong constraint on the noise peak shape, but not overwhelm the spectrum of non-empty events. This same reasoning applies to the laser data from S30 and the classic analysis paradigm. The final result is an amplitude spectrum that closely resembles the "classic" laser data, like the example shown in Fig. 4.27.

Combined laser spectrum analysis One consequence of the special processing of laser data described above is that the statistics for the number of electrons in each sub-window are non-Poissonian. Therefore, instead of one fit parameter μ for the average number of electrons as in sec. 3.2.1, there are now 4 parameters for the probability of having 0 to 3 electrons, P_0 to P_3. It is not necessary to consider more than 3 electrons (4 would not fit in a sub-window of 60 samples), but there is a small proportion of cases with 3 electrons. With the exception of this change, the same paradigm described in section 3.2.1 was used to fit the data, with the following model:

$$\begin{aligned}\mathcal{P}(E) = &\left(P_0 \times \delta(E - \omega_b) + \right. \\ &\left. \sum_{N=1}^{\infty} P_N \times P^{(N)}_{\text{Polya}}(E|\theta, \langle G \rangle) \right) \\ &\otimes P_{\text{Norm}}(E|\omega_b, \sigma_b)\end{aligned} \tag{4.7}$$

where the P_N are all separate parameters, and $P^{(N)}_{\text{Polya}}$ is given by eq. 3.16. In this model, the P_N are constrained to sum to 1, and to be strictly decreasing, i.e. $P_N \geq P_{N+1}$.

The analysis included all low-intensity laser data sets with sufficient statistics, i.e. with a sufficient mean number of sub-windows selection. This included most data sets taken at a laser pump diode current of 120 A, which had μ of approximately 5–7 (see table 4.2). Laser runs with an even lower intensity typically had too low of a yield of usable sub-events for fitting. The model was optimized (based on a binned likelihood function) for the data using an MCMC [216], with 100 walkers and 10^4 steps. A burn-in time of 500 samples was used, exceeding the 220 sample auto-correlation time. An example of such a fit is shown in Fig. 4.27, and the key results from all analyzed data sets are given in table 4.3. Uncertainties for the fit parameters were determined assuming a multivariate Gaussian likelihood function [73].

Table 4.3: Fit results of the combined spectrum low-intensity laser calibration data. All datasets were taken with a laser pump current of 120 A (except ti30s002, which was at 113 A), and at a rate of 10 Hz.

Run	χ^2/ndf	$\langle$G$\rangle$ [ADU]	θ
* ti30s002	2.09	$71.3^{+6.3}_{-5}$	$0.001^{+0.026}_{-0}$
tj03s000	1.17	$73.4^{+4.2}_{-3.4}$	$0.001^{+0.006}_{-0}$
tj04s000	1.22	$69.7^{+6.1}_{-3.6}$	$0.0012^{+0.02}_{-0.0002}$
tj05s000	1.55	74^{+11}_{-11}	$0.008^{+0.02}_{-0.003}$
tj06s000	1.46	$74.9^{+8}_{-7.5}$	$0.001^{+0.05}_{-0}$
tj07s000	1.32	$63.5^{+4}_{-2.7}$	$0.001^{+0.02}_{-0}$
tj08s000	1.73	$69.5^{+2.4}_{-2.4}$	$0.001^{+0.01}_{-0.0001}$
tj09s000	1.24	$62.8^{+2.7}_{-2.3}$	$0.001^{+0.008}_{-0}$
tj10s000	1.08	$78.4^{+2}_{-7.9}$	$0.008^{+0.005}_{-0.08}$
tj11s002	1.64	$70.8^{+2.6}_{-2.3}$	$0.001^{+0.004}_{-0}$

Resolved laser spectra analysis The analysis approach described above suffers from a lack of constraint on the proportion of different electrons. A different paradigm that solves this problem, and also offers the opportunity to quantify the performance of the PF algorithm, involves resolving the laser amplitude spectrum into components according to the number of peaks found in each sub-event (in this case, 0, 1, or 2 peaks). This is done using the PF data obtained with a threshold of 0.48 e^-/ns (see section 4.2.2). The PF results are of course subject to various errors that cause the number of peaks found to differ from the number of actual electrons in the event. The pathologies considered were:

- False positives (FPs): when the peak-finding algorithm misidentifies noise as a peak,

- False negatives (FNs): individual electrons with lower avalanche yields will create a double-deconvolved pulse that falls below the threshold, and
- Coincidences: when neighbouring electron peaks coincide closely enough in time so as to be indistinguishable.

Any number of these pathologies may combine to transform the true underlying number of electrons m into the observed number of peaks n. These pathologies lead to several distorted spectral shapes compared to the basic one described in section 3.2.1 that contribute to the overall amplitude distribution of the data.

Accordingly, several different spectral shapes contribute to the laser spectrum, for true positive electrons, false negatives, and coincidences. The spectral shape $\mathcal{F}^{(\ell)}(E)$ of ℓ true positive electrons is the same as the "traditional" laser model, but the amplitude probability distribution for each electron is multiplied by an efficiency curve $f(E)$ representing the PF algorithm threshold, with free parameters ω_h and σ_h:

$$f\left(E|\omega_h,\sigma_h\right)=\frac{1}{2}\left(1+Erf\left(\frac{E-\omega_h}{\sqrt{2\sigma_h^2}}\right)\right) \tag{4.8}$$

and the spectral shape $\mathcal{F}$ itself is written as:

$$\mathcal{F}^{(\ell)}(E)=\begin{cases}\delta\left(E-\omega_b\right) & \ell=0\\ P_{\text{Polya}}\left(E|\theta,\langle G\rangle\right)\times f(E|\omega_h,\sigma_h) & \ell=1\\ \mathcal{F}^{(\ell-1)}(E)\otimes\mathcal{F}^{(1)}(E) & \ell>1\end{cases} \tag{4.9}$$

The amplitude probability distribution for false negatives $\mathcal{G}$ is derived similarly, except it is multiplied by the reciprocal of the PF efficiency curve:

$$\mathcal{G}^{(\ell)}(E)=\begin{cases}\delta\left(E-\omega_b\right) & \ell=0\\ P_{\text{Polya}}\left(E|\theta,\langle G\rangle\right)\times\left(1-f(E|\omega_h,\sigma_h)\right) & \ell=1\\ \mathcal{G}^{(\ell-1)}(E)\otimes\mathcal{G}^{(1)}(E) & \ell>1\end{cases} \tag{4.10}$$

For coincident events $\mathcal{H}$, a single contribution ($\ell=1$) is a combination of two or more electrons, only one of which need pass the peak-finding efficiency criteria:

$$\mathcal{H}^{(\ell)}(E) = \begin{cases} \delta\left(E - \omega_b\right) & \ell = 0 \\ \left[P_{\text{Polya}}\left(E|\theta, \langle G\rangle\right) \times f(E|\omega_h, \sigma_h)\right] \otimes P_{\text{Polya}}\left(E|\theta, \langle G\rangle\right) & \ell = 1 \\ \mathcal{H}^{(1)}(E) \otimes P_{\text{Polya}}^{(\ell-1)}\left(E|\theta, \langle G\rangle\right) & \ell > 1 \end{cases} \tag{4.11}$$

Only certain combinations of different contributions are permissible to build the amplitude spectrum for a given number of observed peaks n. In particular, n must be equal to the true number of electrons in an event m, plus contributions from i false positives, minus j false negatives and k coincidences. This requirement can be encoded as:

$$c_{ijk}^{nm} = \begin{cases} 1, & n = m + i - j - k \\ 0, & \text{otherwise} \end{cases} \tag{4.12}$$

The overall model for n peaks as a function of amplitude E is derived by summing over all possible values of m, i, j, and k, namely keeping only those terms where $c_{ijk}^{nm} = 1$, but also with the constraints that there cannot be more false-negatives or coincidences than true electrons to begin with. Additionally, false-positives cannot be considered to be in coincidence with each other or with true positive peaks. The statistics of false-positives is Poissonian with some average number of false-positives per window p_{FP}. False negatives and coincident contributions follow binomial statistics, with p_{CE} as the probability of a single coincidence, and p_{FN} as the probability of a single false negative:

$$p_{FN} = \int \mathcal{G}^{(1)}(E) dE \tag{4.13}$$

The probability of having m electrons is also a free parameter in the fit, as Poissonian statistics are not assumed. The appropriate contributions of $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$ are then convolved with each other, and finally, the whole model is convolved with baseline noise $P_{\text{Norm.}}(E|\omega_b, \sigma_b)$:

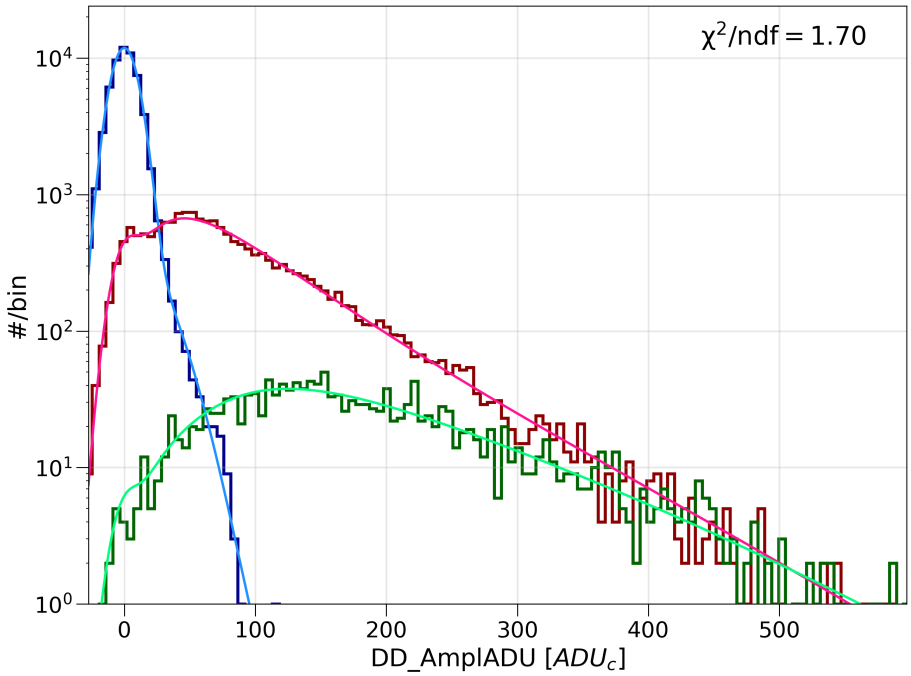


Figure 4.28: Histograms of peak-resolved low-intensity laser data (tj04s000) for different numbers of peaks (blue = 0, red = 1, green = 2) and best-fit model (matching curves).

$$\mathcal{P}\left(\{E,n\}\right) = \sum_{m=0}^{} \sum_{i=0}^{} \sum_{j=0}^{m} \sum_{k=0}^{m-1} \Bigg[c_{ijk}^{nm} \times p_m P_{\text{Poisson}}(i|p_{FP}) \binom{m}{j} p_{FN}^{j} \left(1-p_{FN}\right)^{m-j} \binom{m}{k} p_{CE}^{k} \left(1-p_{CE}\right)^{m-k} \times \left(\mathcal{F}^{(m-j-k)}(E) \otimes \mathcal{G}^{(j)}(E) \otimes \mathcal{H}^{(k)}(E)\right) \Bigg] \otimes P_{\text{Norm.}}(E|\omega_b, \sigma_b) \tag{4.14}$$

Practically, computation of this model requires some truncation condition for the nested sums, which is enforced as $c_{ijk}^{nm} = 0$ if $m > 4$, $i > 2$, $j > 4$, or $k > 3$. The only constraints on the model (other than basic physical limits of parameters) are that the probabilities for different numbers of electrons are strictly decreasing ($p_m \geq p_{m+1}$) and sum to 1, $\sum p_m = 1$. Fitting this model to the data was computationally challenging, as there are 13 free parameters in the model with m being allowed to go up to 4. Rather than a typical MCMC with MH stepping, the "fast burn-in" method described in Appendix A was used. An example of a fit to the data is shown in Fig.

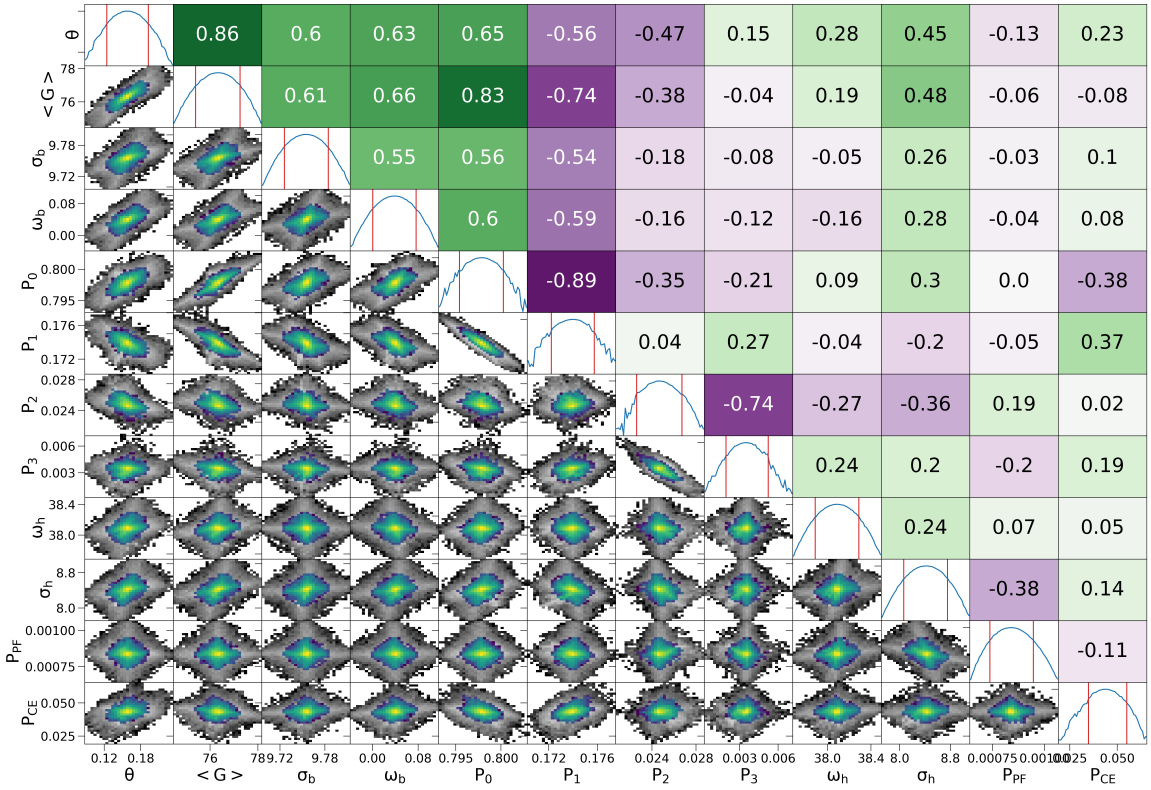


Figure 4.29: Example of a corner plot for the fit of a peak-resolved amplitude spectrum of low-intensity laser calibration data (tj04s000 specifically). The diagonal panels show the profile log-likelihood functions for each parameter, the lower panels show the 2D histograms of MCMC samples between each pair of parameters, and the corresponding upper panels give the Pearson correlation coefficient for each pair of parameters.

4.28. A corner-plot (for the same example fit) showing the profile log-likelihoods and correlations between all parameters is given in Fig. 4.29, which demonstrates that the likelihood is approximately Gaussian in this case. Table 4.4 gives the key fit results using this method, with uncertainties calculated assuming a Gaussian likelihood [73].

In Fig. 4.28, one can see the trade-off between the single peak spectrum and the non-Gaussian tail of the zero peak spectrum, indicating the peak-finding threshold. Some contribution of null events identified as one or two peak events can also be seen (FPs). Encouragingly, the estimates of $\langle G \rangle$ obtained through this independent analysis largely agree with the complete laser spectrum results, as can be seen in table 4.5. The results for θ tend to be slightly higher through this approach. However, the

Table 4.4: Key fit results of the resolved spectrum low-intensity laser calibration data. All datasets were taken with a laser pump current of 120 A, and at a rate of 10 Hz.

Run	$\langle$G$\rangle$ [ADU]	θ	ω_h [ADU]	σ_h [ADU]
tj03s000	$77.1^{+1.3}_{-1.4}$	$0.04^{+0.04}_{-0.04}$	$37.5^{+0.3}_{-0.3}$	$8.2^{+0.5}_{-0.5}$
tj04s000	$76.3^{+0.9}_{-1.0}$	$0.16^{+0.03}_{-0.03}$	$38.1^{+0.2}_{-0.2}$	$8.4^{+0.3}_{-0.4}$
tj05s000	$77.7^{+0.9}_{-0.9}$	$0.15^{+0.03}_{-0.03}$	$37.5^{+0.2}_{-0.2}$	$7.6^{+0.3}_{-0.3}$
tj06s000	$76.5^{+0.9}_{-0.9}$	$0.17^{+0.03}_{-0.03}$	$37.5^{+0.2}_{-0.2}$	$8.1^{+0.3}_{-0.3}$
tj07s000	$73.7^{+1.0}_{-1.0}$	$0.09^{+0.03}_{-0.04}$	$37.7^{+0.2}_{-0.2}$	$7.9^{+0.4}_{-0.4}$
tj08s000	$72.0^{+0.9}_{-1.0}$	$0.11^{+0.04}_{-0.04}$	$38.2^{+0.2}_{-0.2}$	$8.2^{+0.4}_{-0.3}$
tj09s000	$69.9^{+0.9}_{-1.0}$	$0.12^{+0.04}_{-0.04}$	$38.5^{+0.2}_{-0.2}$	$7.9^{+0.3}_{-0.4}$
tj10s000	$69.6^{+1.0}_{-1.0}$	$0.04^{+0.04}_{-0.04}$	$37.9^{+0.2}_{-0.2}$	$7.2^{+0.4}_{-0.4}$
tj11s002	$71.9^{+1.1}_{-1.0}$	$0.02^{+0.04}_{-0.02}$	$38.2^{+0.2}_{-0.2}$	$7.2^{+0.4}_{-0.4}$

resolved-spectrum analysis should be more sensitive to θ since correlations with the p_m parameters are reduced, and the split between the 0 and 1 peak spectra is where θ will have the strongest impact on the data. The average measured value of theta in this analysis is $\theta = 0.125^{+0.026}_{-0.023}$. It is these results (for θ and $\langle G \rangle$) that are used throughout the rest of the LSM analysis.

Full details of the characterization of the PF algorithm are presented in section 4.4.2. The PF threshold (described by ω_h and σ_h) obtained from this analysis is directly used for the overall PF model. However, the parameterization of FPs and coincident peaks from this analysis is not directly applicable to the rest of the LSM analysis. For the former, the estimates of P_{FP} in this study are specifically for sub-windows only 60 samples long; the probability of noise triggers in the 1290 sample window used to process the data (see section 4.2.2) is much greater. This quantity was also measured directly using pre-trace data. As for coincident peaks, the laser analysis

Table 4.5: Fit results for $\langle G \rangle$ and θ obtained through both laser analysis methods.

	Combined Spectra		Resolved Spectra	
Run	$\langle$G$\rangle$ [ADU]	θ	$\langle$G$\rangle$ [ADU]	θ
tj03s000	$73.4^{+4.2}_{-3.4}$	$0.001^{+0.006}_{-0}$	$77.1^{+1.3}_{-1.4}$	$0.04^{+0.04}_{-0.04}$
tj04s000	$69.7^{+6.1}_{-3.6}$	$0.0012^{+0.02}_{-0.0002}$	$76.3^{+0.9}_{-1.0}$	$0.16^{+0.03}_{-0.03}$
tj05s000	74^{+11}_{-11}	$0.008^{+0.02}_{-0.003}$	$77.7^{+0.9}_{-0.9}$	$0.15^{+0.03}_{-0.03}$
tj06s000	$74.9^{+8}_{-7.5}$	$0.001^{+0.05}_{-0}$	$76.5^{+0.9}_{-0.9}$	$0.17^{+0.03}_{-0.03}$
tj07s000	$63.5^{+4}_{-2.7}$	$0.001^{+0.02}_{-0}$	$73.7^{+1.0}_{-1.0}$	$0.09^{+0.03}_{-0.04}$
tj08s000	$69.5^{+2.4}_{-2.4}$	$0.001^{+0.01}_{-0.0001}$	$72.0^{+0.9}_{-1.0}$	$0.11^{+0.04}_{-0.04}$
tj09s000	$62.8^{+2.7}_{-2.3}$	$0.001^{+0.008}_{-0}$	$69.9^{+0.9}_{-1.0}$	$0.12^{+0.04}_{-0.04}$
tj10s000	$78.4^{+2}_{-7.9}$	$0.008^{+0.005}_{-0.08}$	$69.6^{+1.0}_{-1.0}$	$0.04^{+0.04}_{-0.04}$
tj11s002	$70.8^{+2.6}_{-2.3}$	$0.001^{+0.004}_{-0}$	$71.9^{+1.1}_{-1.0}$	$0.02^{+0.04}_{-0.02}$

only includes (and models) coincident events with primary electrons beginning at the surface of the SPC. Events originating closer to the sensor will experience less diffusion, therefore increasing the probability of having coincident pulses. The laser data cannot reveal the radial dependence of this quantity.

4.3.2 High intensity laser calibration

Another principal use of UV laser data is to monitor changes in the detector response over time, as increasing gas contamination may slowly reduce the gain of the SPC. Although the daily low-intensity calibration data described in the previous subsection does give information about the changing gain over time, the relatively large statistical uncertainty on these results obscures any obvious trends. Instead, the high-intensity laser data collected during all physics runs (see section 4.2.1) can be used for this purpose. The amplitude of these events (with approximately 50 to 130 photoelectrons)

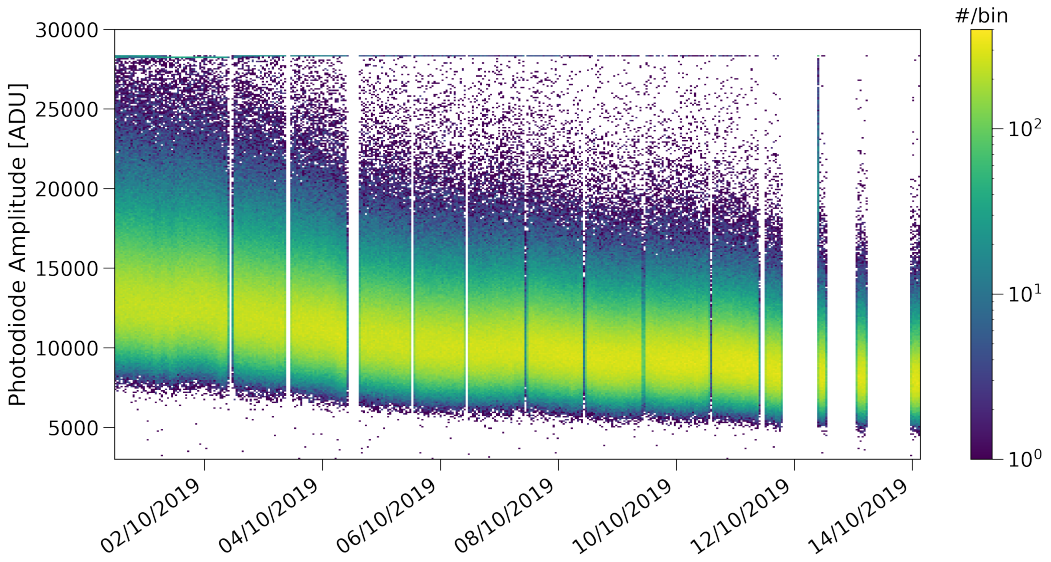


Figure 4.30: Photodiode amplitude data over time from high-intensity laser runs for the whole LSM campaign shown as a 2D histogram (colour-scale is counts per bin).

in the SPC was high enough that the triggering was 100% efficient on the SPC channel, and the tagging was 100% efficient using the photodiode amplitude.

A separate, confounding reason that the laser signal changes over time is due to the optical fibre aging. The fibres used were not intended for use at UV wavelengths, so over time laser light damages the fibre, gradually reducing their transmission. This effect can be seen in Fig. 4.30. To decouple the effect of fibre aging from changes in the SPC gain over time (both of which lead to decreasing laser signal amplitudes in the SPC), the ratio of SPC amplitude to photodiode amplitude was used. This is shown in Fig. 4.31. A three-segment piece-wise linear function with floating knots (the (x,y) points where the line segments connect) was fit to the time series data, excluding a short period where an Am-Be source increased the surface electric field (part way through tj10s001 to part way through tj12s003). The major drop in laser signal for the last three data periods was after the introduction of ^{37}Ar, which likely increased the contamination of oxygen and water in the sphere, hastening degradation of the gas. Between the beginning and end of the campaign, the gain dropped by 11% during the physics runs, and by 27% in total by the end of the ^{37}Ar data taking.

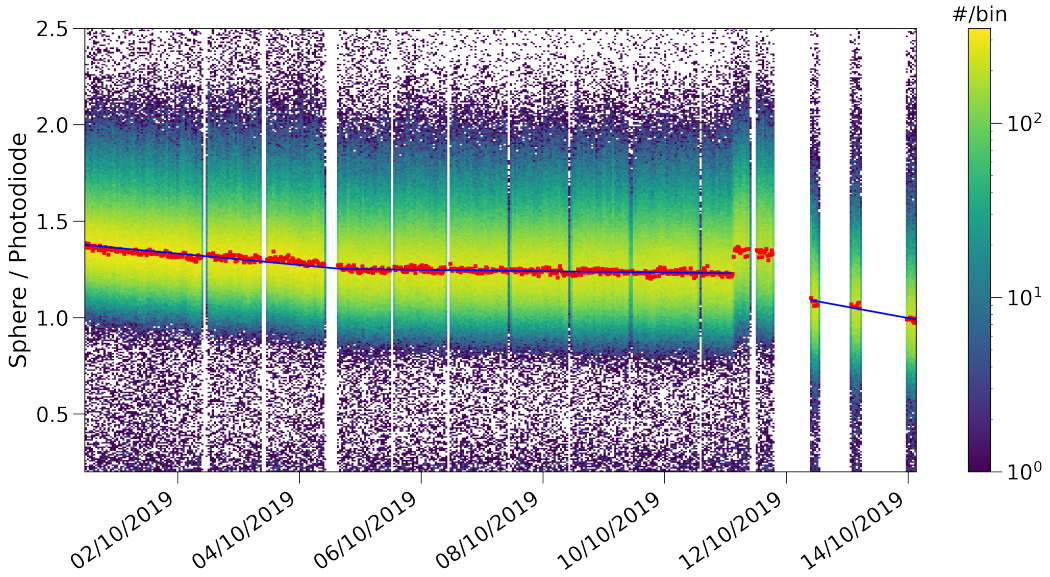


Figure 4.31: High-intensity laser data (sphere amplitude divided by photodiode amplitude) over time for the whole LSM campaign shown as a 2D histogram, with the mean value per slice (red) and piece-wise linear trend (blue). The period omitted from fitting is when the Am-Be source was present.

To assign an absolute scale to this trend in $\langle G \rangle$ over time, the low-intensity laser calibration results – specifically the resolved spectra results given in table 4.4 – were used to determine a single scaling factor for the piece-wise linear trend. This scaled result is shown in Fig. 4.32, which also depicts the averaged error bands for $\langle G \rangle$ from all low-intensity calibration runs. This result could be used to calculate a complete time-dependent energy response model for the SPC as a function of time, but for practical reasons, the dark matter analysis presented in this chapter uses the gain trend at the median time of the last physics run (tj10s001) unless stated otherwise.

4.3.3 ^{37}Ar analysis

As mentioned in section 4.2.1, ^{37}Ar was a primary calibration source used for the LSM campaign, which was injected at the end of the data-taking period. This volume-distributed source allows for calibration of the detector energy response throughout the whole detector volume, at low energies, with multiple prominent calibration peaks. Together with the results of UV laser calibrations described in the previous section,

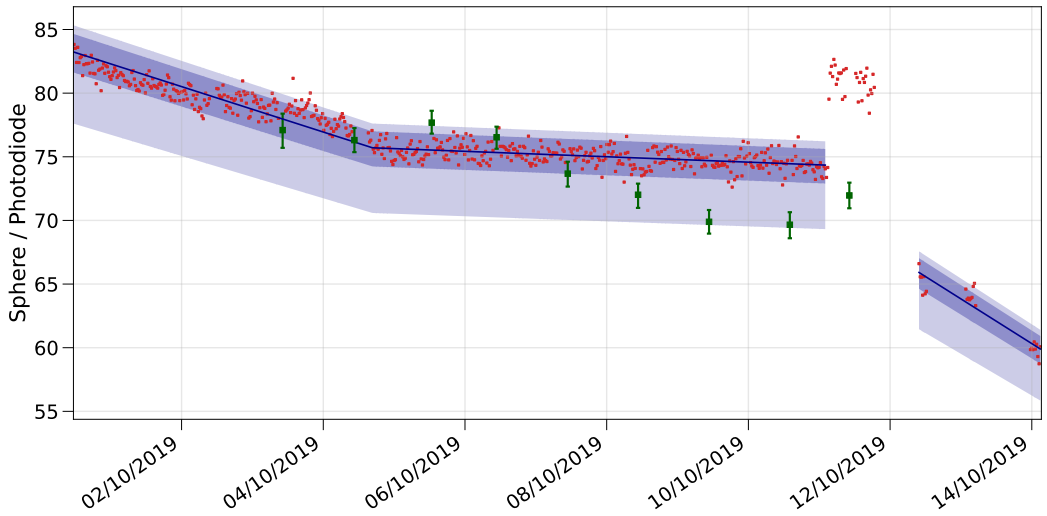


Figure 4.32: Detector mean gain over time for the entire LSM physics campaign, with individual low-intensity measurement results shown in green, the average of the measurements in blue (with 1 and $s\,\sigma$ uncertainty bands), and scaled high-intensity laser data (red). The period omitted from fitting is when the Am-Be source was present.

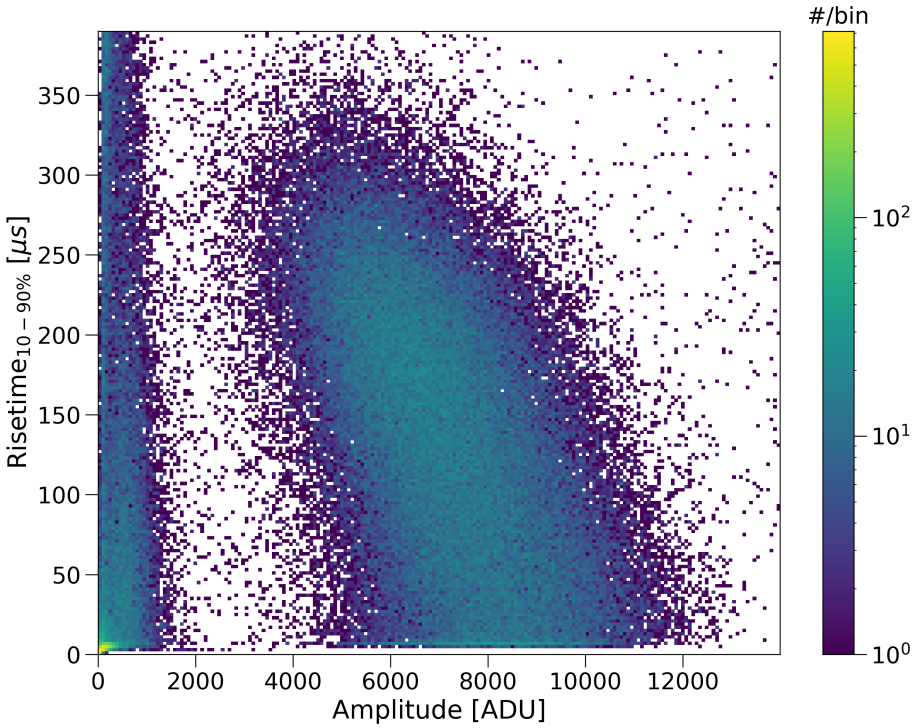


Figure 4.33: 2D histogram of risetime vs. amplitude data for tj13s000, with cuts described as in section 4.3.3. The 2.8 keV peak is seen centered around 8000 ADU. The lower energy peak sits from approximately 500 – 1500 ADU, overlapping with a population of non-physical events below 500 ADU.

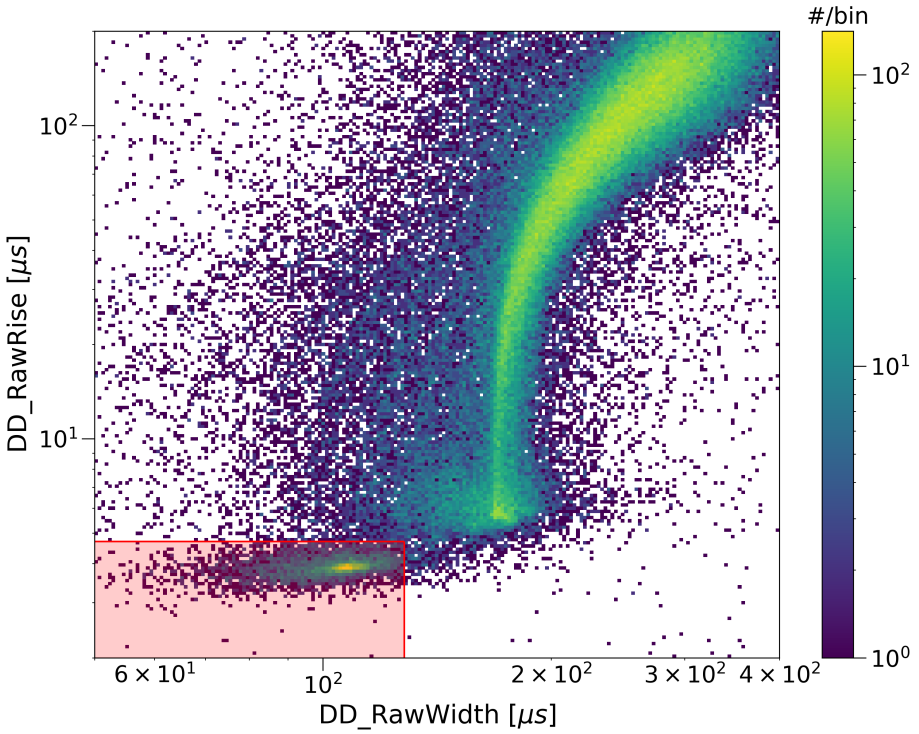


Figure 4.34: 2D histogram of `DD_RawRise` vs. `DD_RawWidth` for run tj13s000, showing the PSD cut made on these variables to remove non-physical events (red shaded region).

this ^{37}Ar analysis gives a complete model of the detector's energy response (to electronic recoils) by allowing for in-situ measurements of the W-value $W(E)$ and Fano factor F. The analysis was done to complement the independent measurements done in methane gas with the S30 detector, described in section 3.3.

The data used in this analysis comes from the run tj13s000, the first ^{37}Ar data collected after diffusion of the sample throughout the SPC had ceased (this was still visible in tj12s003). A later run was taken again at the very end of the campaign (tj14s001), but by this point, the gas quality had degraded somewhat due to the injection of the sample (see Fig. 4.31). The data was processed using the standard double deconvolution parameters described in section 4.2.2, albeit with a slightly smaller integration window size of 1090 samples. The risetime vs. amplitude plot of the data is shown in Fig. 4.33, which clearly shows the high and low energy peaks from ^{37}Ar, as well as a background of low energy events that slightly increases in risetime

with amplitude. The only cuts applied to the data were to remove post-alpha time periods (see section 4.4.4), laser pulses, and events with shared electrons between the north and south hemispheres based on the peak-finding results applied to the north channel. Finally, a loose PSD cut based on `DD_RawWidth` and `DD_RawRise` was made; the selection was of events with $> 128\,\mu$s and $> 4.7\mu$s respectively (see section 3.1.5). This cut/parameter space is shown in Fig. 4.34, and was intended to reject saturated events with negative widths, and obvious spike events with instantaneous risetimes – the efficiency of this loose cut is taken to be 100%.

Modelling of the ^{37}Ar starts with the same electron-capture decay simulation described in section 3.2.2, based on similar work in refs. [199, 200]. The Geant4 simulations of decay products in the S30 detector were repeated for S140 with 135 mbar of methane gas (see Fig. 3.23), the main effect of which was greater containment of the decay products, with more energy deposited in the gas on average. Using the subsequent MC electron capture events, the spectrum of total energy deposited in the gas as a function of the number of decay particles was obtained, which is shown in Fig. 4.35. Next, the probability distribution of the mean number of primary electrons produced in the gas was calculated according to eq. 3.26, for given values of W_0 and U. Practically, this was a discrete distribution of the mean $P(\mu)$ for linearly spaced values of μ from 0 to 160, spaced by 0.05. As in section 3.3, this distribution of the mean number of electrons is then "smeared" with the COM-Poisson distribution to represent dispersion in primary ionization statistics (with a Fano factor F). These spectra were resolved into components by the number of Auger (initial) electrons k in the simulated ^{37}Ar events (0 – 5), and then stacked together (see Fig. 3.33).

Avalanche statistics were modelled using the Polya distribution, as described in section 3.2.1, with the average result of $\theta = 0.125$ being used for the main fit in this analysis. A prior on the gain of one anode was derived using the average result of all low-intensity laser calibration data, interpolated at the time of tj13s000 (see Fig. 4.32). Finally, the (Gaussian) baseline noise of the data was measured directly using

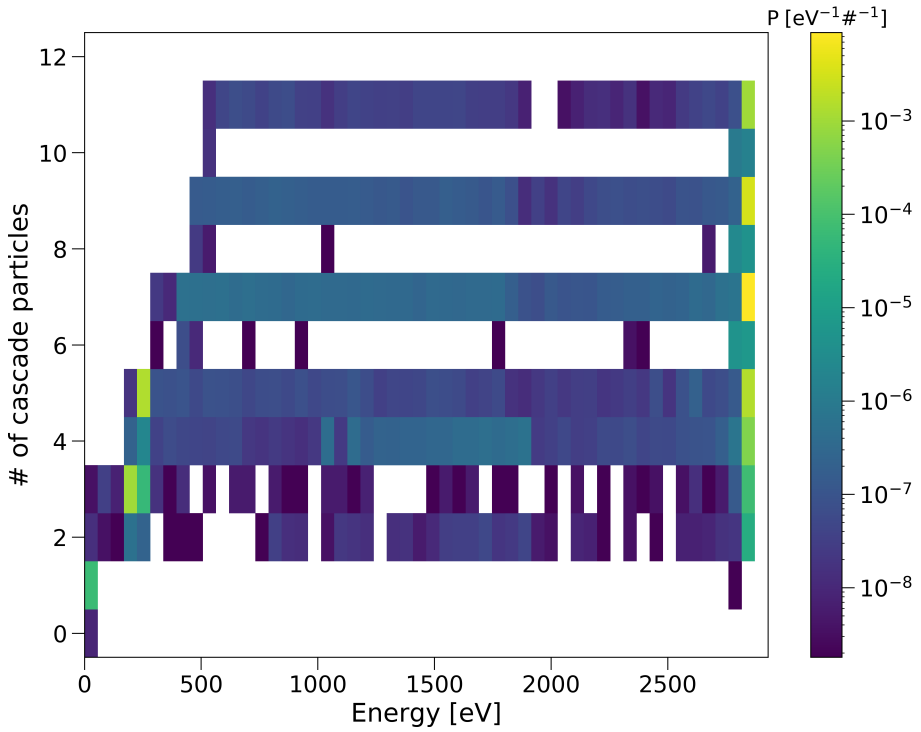


Figure 4.35: A 2D histogram of the number of decay cascade particles (Auger electrons and photons) vs. the total energy deposited in an SPC per event following the electron capture decay of ^{37}Ar. This result is simulated for S140 with 135 mbar of pure CH_4 gas.

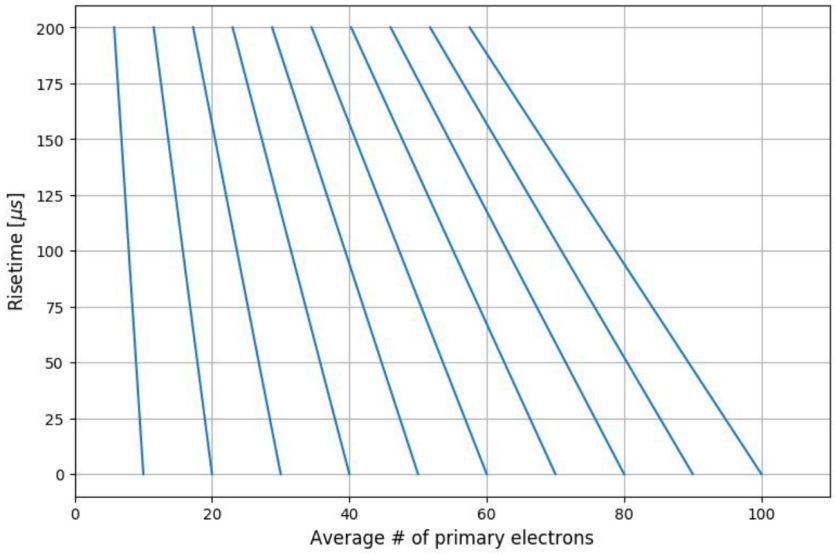


Figure 4.36: Risetime vs. amplitude plot depicting the attachment model employed in this analysis (arbitrary example with somewhat realistic parameter values). Each curve shows the effect of attachment for a given initial mean number of electrons as a function of risetime.

pre-trace data, which indicated noise with $\sigma_{\text{Noise}} = 87.0 \pm 0.5\,\text{ADU}$.

Another aspect of this ^{37}Ar analysis that differs from the S30 measurements is charge attachment, which can be seen clearly in Fig. 4.33 as an anti-correlation between the risetime and amplitude data of the high-energy peak. Modelling the impact of charge attachment on the detector's energy response can therefore be achieved if a two-dimensional (2D) fit of the data is carried out (using both risetime and amplitude variables), rather than just fitting the amplitude data at the lowest risetimes as is usually done. Attachment was modelled by assuming that the distribution of primary electrons that reach the sensor follows binomial statistics, with the success probability being a linear function of risetime. Thus, the model parameter to account for attachment is simply the scaling slope between risetime and attachment (charge-trapping) survival probability T_p (attachment is assumed to be negligible at a risetime of 0). A depiction of this parameterization is shown in Fig. 4.36 (arbitrary example with somewhat realistic parameter values).

Combining all of the above aspects of the physics of the ^{37}Ar data, the total detector

response can be written as a function of amplitude E, at a fixed risetime RT. By integrating over the distribution of the mean number of primary ionizations, summing over k auger electrons and $i-k$ produced primary electrons, and j surviving charge trapping, the model is given as can be written as (ignoring background events for now):

$$\begin{aligned}\mathcal{P}\left(E, RT\right) = \Bigg[\sum_{i}\Bigg(\sum_{j=0}^{j=i}\Bigg(\sum_{k=0}^{k=5}\Bigg(\int P_{\mathrm{COM}}(i-k|\mu, F)P^{(k)}(\mu)d\mu + k\Bigg) \\ \times P_{\mathrm{Binom.}}\left(j|i, p = T_p \times RT\right) \times P^{(\mathrm{j})}_{\mathrm{Polya}}(E|\left\langle G\right\rangle, \theta)\Bigg)\Bigg)\Bigg] \\ \otimes P_{\mathrm{Noise}}\left(E\right)\end{aligned} \tag{4.15}$$

Anomalous dispersion and modelling Unfortunately, the model described above is not adequate to match the data of tj13s000, which exhibits far more dispersion (broader energy resolution) than expected. To use eq. 4.15 as-is to fit the data would require $F \gg 1$, which conflicts with all previous measurements and physical expectations of the quantity. There are various possible explanations for this behaviour, such as the possibility that the electric field structure surrounding the southernmost anode differs considerably from the other south-hemisphere anodes. To test this, an electric field model of the SPC made using the COMSOL finite-element simulation software [154] (produced by F. Vazquez and Y. Deng) was used to examine the field strength near the anodes, and is shown in Fig. 4.37. This indicated that there is no significant difference in the electric field strength in the immediate vicinity of the anodes. Furthermore, the densities of field lines reaching the anodes as a function of polar angle also match for different south channel anodes. Large baseline noise cannot be responsible for this dispersion, as this was measured directly and already included in the model of the data (see above).

Another hypothesis that cannot be ruled out is that the mean gain $\langle G\rangle$ of each of the six southern anodes may vary, due to defects in each anode, or other struc-

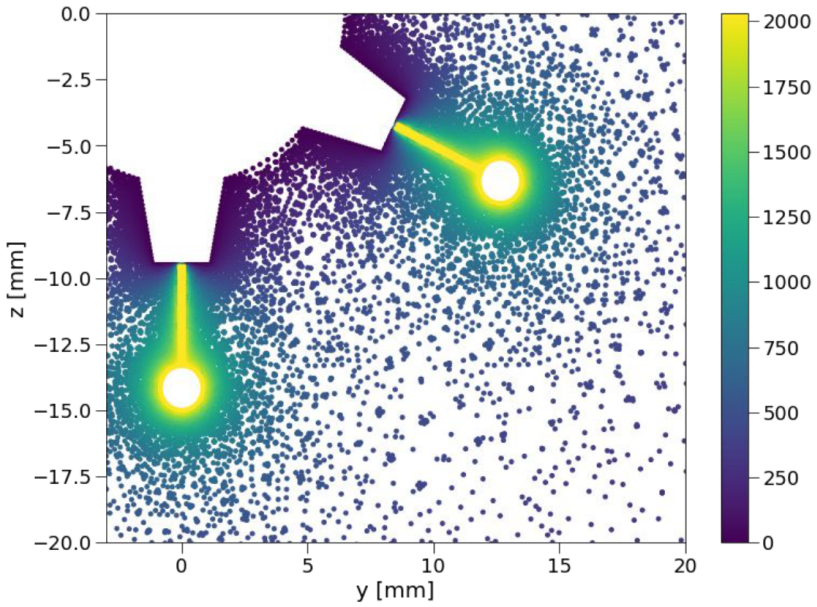


Figure 4.37: COMSOL model of the electric potential (in V) around the ACHINOS sensor, zoomed in on the southernmost anode and one neighbouring anode. The ACHINOS itself is an empty space devoid of points.

tural differences from the fabrication of the ACHINOS that are not included in the COMSOL model. While $\langle G \rangle$ is constrained by the laser data, there is no way to know which or how many south-channel anodes capture photoelectrons produced by the laser, only that at least one does. Based on approximations of the beam spread of the laser, indeed it seems probable that only one south-channel anode acquires laser electrons (in addition to one or more north-channel anodes, see section 4.4.6). Therefore, to account for this possibility and afford the model flexibility to match the data, the mean gain of each of the six south anodes was treated independently as a free parameter. Accordingly, the fit model given in eq. 4.15 was repeated six times for each anode (and summed together). The six parameters $\langle G \rangle_l$ are merely constrained to be monotonically increasing ($\langle G \rangle_l \leq \langle G \rangle_{l+1}$), and one of the gain parameters has a constraint term (prior) from the results of the laser analysis (see section 4.3.1). Other possible explanations do exist for the dispersion of the data, such as time-variations in the gain or detector response; unfortunately, the data collected does not have sufficient statistics to explore this idea. However, the choice to include six gain

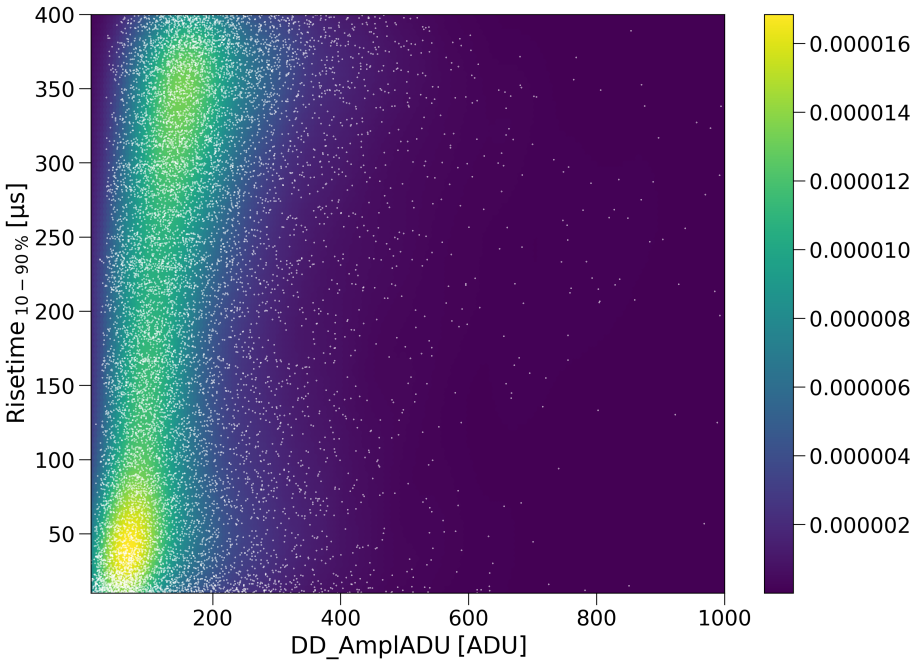


Figure 4.38: Risetime vs. amplitude data of run tj04s002 (white points) with the two-dimensional kernel density estimation model produced using the data (colour scale).

parameters in the model can also functionally account for any explanation involving variation of the gain, and was therefore a generic modelling choice to account for the anomalous dispersion of the amplitude data in tj13s000.

Background model In addition to the high and low-energy contributions of ^{37}Ar apparent in Fig. 4.33, there is also a prominent background of low amplitude events across all risetimes, with a slight correlation with risetime. This background likely consisted of many single electron events (at the minimum risetime), possibly α-induced despite the cut intended to remove them (see section 4.4.4). However, the correlation with risetime also suggests that some of these events were non-physical spike events. Rather than trying to remove most of these events and introducing the complication of the efficiency of such a cut as a function of risetime and amplitude, these non-^{37}Ar events were left in the dataset, and modelled as a background component to be included in the fit.

This was accomplished using a source of data with presumably matching contribu-

tions of low energy physics and spike events, namely the physics data itself. One of the non-blind physics runs (tj04s002) was used for this purpose, with identical cuts applied. To obtain an analytical representation, a 2D KDE of the data was calculated, with a Gaussian kernel, using the python library `scipy` [238]. The resulting KDE and data on which it is based are shown in Fig. 4.38, where one can see the expected population of low amplitude events slightly correlated with risetime.

In addition to this modelled background, a flat background component was included with separate rates B_1 and B_2 below and above the 2.8 keV respectively, smoothly connected by the reciprocal CDF of the ^{37}Ar model according to eq. 4.15. If the 2.8 keV peak is described as $P(E)$ (at a given risetime RT), then the overall background model $B(E)$ around this peak is given below, and appended to the rest of the model given in eq. 4.15:

$$B(E) = B_1 \times \left[1 - \int^{E} P(E')dE'\right] + B_2 \tag{4.16}$$

Data fitting and results To simplify the fit of the 2D ^{37}Ar data, events were divided into 8 risetime bins between 10 and 250 μs, which relies on the assumption that there is no significant difference in attachment between the highest and lowest risetimes in each bin. The KDE background model described previously was scaled independently for each risetime bin i by a factor R_i. This reduced the model to essentially 8 separate one-dimensional fits in amplitude. Overall, there are 10 parameters of interest in the model, and 10 nuisance parameters describing the background model, which are summarized in table 4.6. Prior constraints applied to this fit include one on $\langle G \rangle_{l=1}$ from the laser calibration results of section 4.3.1 scaled to the time of tj13s000, as well as 2D priors on W_0 and U from the three independent W-value measurements described in section 3.3.

Fitting this high-dimensional model was done using the "'fast burn-in" MCMC method described in Appendix A [3, 277]. A corner plot (for the main fit of the

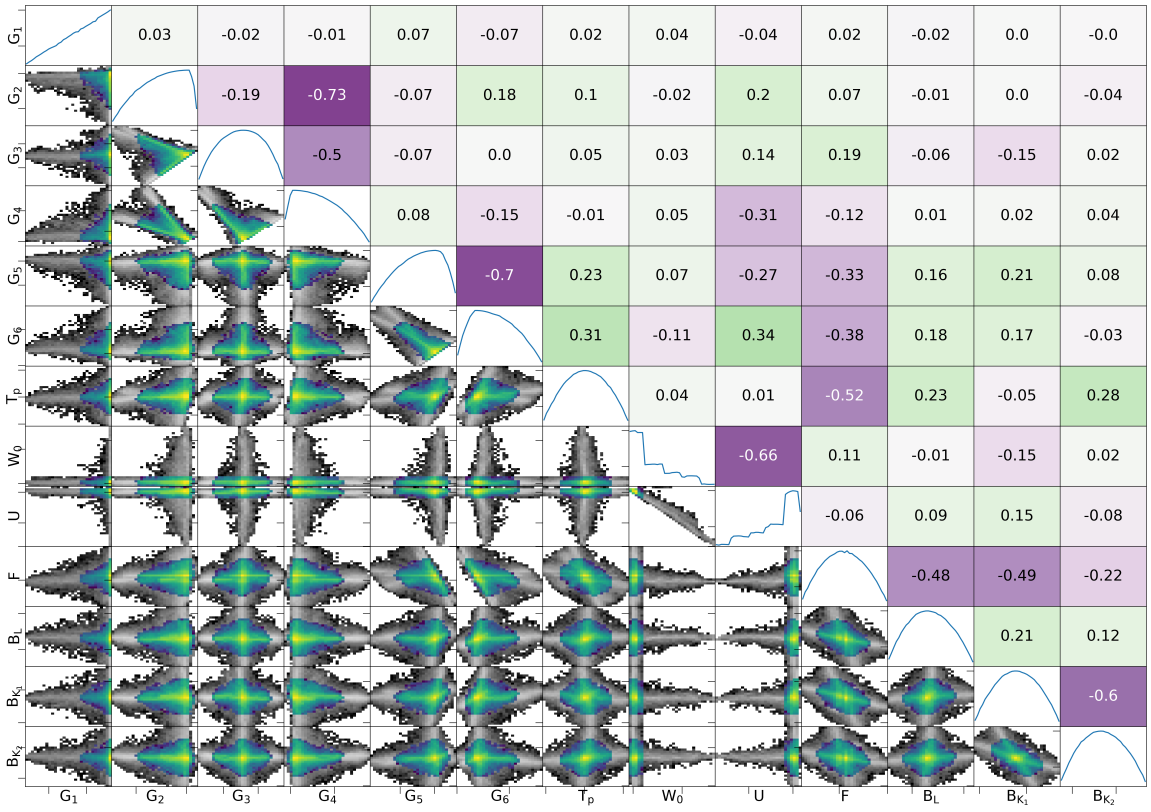


Figure 4.39: Corner plot of the ^{37}Ar MCMC fit with $\theta = 0.125$ and $\sigma_{\text{Noise}} = 87.0\,\text{ADU}$. The background KDE scaling parameters (R_i) are omitted for clarity, and because they do not have notable correlations with any other parameters. The diagonal panels show the profile log-likelihood function for each parameter. The lower left panels depict 2D histograms of the MCMC samples for every pairing of parameters, with the coloured (resp. grey) portion showing approximate $1\,\sigma$ (resp. $2\,\sigma$) uncertainty regions assuming a Gaussian likelihood [73]. The upper-right panels give the Pearson correlation coefficients between the pairings of parameters with a corresponding colour-scheme.

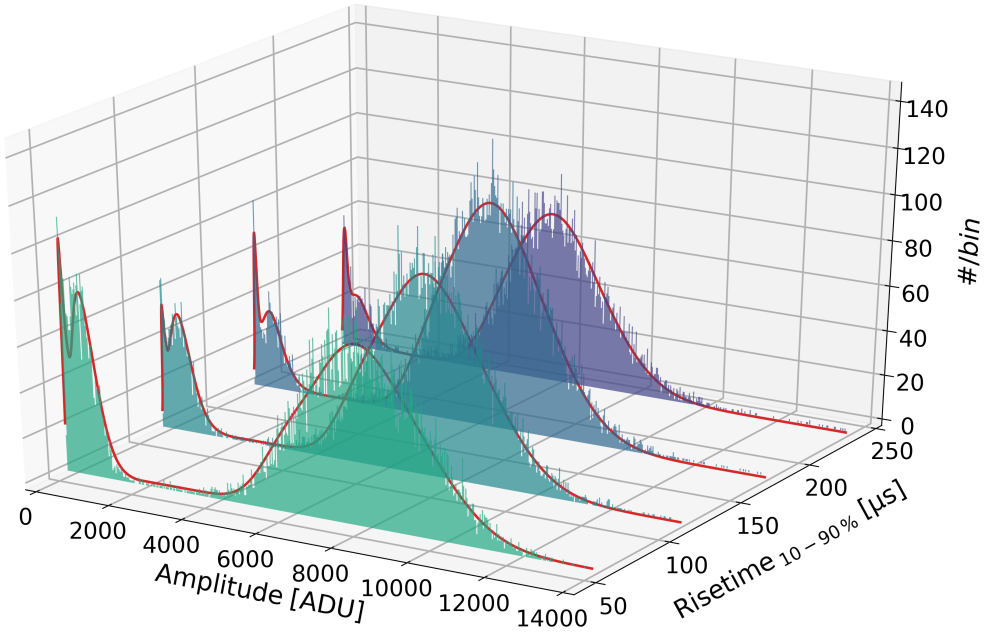


Figure 4.40: ^{37}Ar data from tj13s000 (histograms), shown for four out of eight risetime bins, with the best-fit model to the data (red curves).

data with $\theta = 0.125$ and $\sigma_{\text{Noise}} = 87.0\,\text{ADU}$) displaying the MCMC samples and 1D profile likelihood distributions for all of the parameters of interest (plus B_1 and B_2) is shown in Fig. 4.39. Approximate representations of uncertainties are obtained from the MCMC samples assuming a multivariate Gaussian likelihood [73]. For error propagation and other analyses, the MCMC samples were used directly unless otherwise stated. The resulting fit of the data is shown as a function of risetime and amplitude in Fig. 4.40. One can see that the model matches both ^{37}Ar peaks and the low-amplitude background well, including with changing risetime, which is depicted more clearly for the 2.8 keV peak in Fig. 4.41.

To propagate uncertainty on the parameter θ without incurring a severe cost in computing time (it took roughly 30 minutes to compute all the necessary distributions of $P^{(N)}_{\text{Polya}}$ for a single value of θ), the fit was redone using the upper and lower $1\,\sigma$ values of θ estimated in the laser analysis ($\theta = 0.125^{+0.26}_{-0.23}$). The value of the baseline noise $\sigma_{\text{Noise}} = 87.0 \pm 0.5\,\text{ADU}$ was also propagated in this way, in perfect

Parameter	Description	Fit result
$\langle G \rangle_1$	Gain of one anode	$88.5 \pm 0.11\,\mathrm{ADU}$
$\langle G \rangle_2$		$88.55 \pm 0.25\,\mathrm{ADU}$
$\langle G \rangle_3$		$88.62 \pm 0.26\,\mathrm{ADU}$
$\langle G \rangle_4$		$107.04 \pm 0.28\,\mathrm{ADU}$
$\langle G \rangle_5$		$107.05 \pm 0.21\,\mathrm{ADU}$
$\langle G \rangle_6$		$107.05 \pm 0.22\,\mathrm{ADU}$
$\mathrm{T_P}$	Trapping coefficient	$1.26 \pm 0.01 \times 10^{-4}$
W_a	W-value asymptotic value	$30.00^{+0.14}_{-0.15}\,\mathrm{eV}$
U	W-value low energy cut-off	$15.70^{+0.52}_{-0.30}\,\mathrm{eV}$
F	Fano factor	0.43 ± 0.05
B_1	Flat bkgd. below 2.8 keV	$2.3 \pm 0.3 \times 10^{-5}$
B_2	Flat bkgd. above 2.8 keV	$4.6 \pm 0.3 \times 10^{-6}$
R_1	Scaling of bkgd. KDE in risetime bin	0.427 ± 0.001
R_2		0.241 ± 0.002
R_3		0.173 ± 0.001
R_4		0.185 ± 0.001
R_5		0.238 ± 0.001
R_6		0.353 ± 0.002
R_7		0.416 ± 0.001
R_8		0.537 ± 0.001

Table 4.6: Description of model parameters of the LSM ^{37}Ar calibration, with fit results. The first portion of the table lists results for parameters of interest, while the lower lists the nuisance parameters describing the background model.

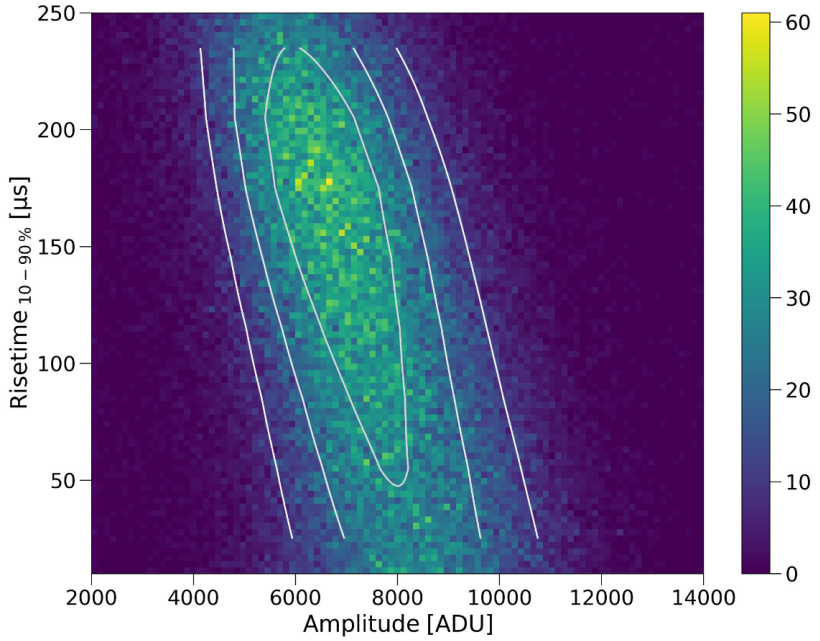


Figure 4.41: High-energy ^{37}Ar data of tj13s000 (two-dimensional histogram) with best-fit model (contours) as a function of risetime and amplitude.

positive correlation with θ, i.e. $\theta = 0.151$ and $\sigma_{\text{Noise}} = 86.5\,\text{ADU}$ vs. $\theta = 0.106$ and $\sigma_{\text{Noise}} = 87.5\,\text{ADU}$. As F is also somewhat degenerate with these two parameters, this procedure yields the maximum uncertainty on F. The three MCMCs performed were combined to produce the overall fit results (incorporating uncertainty on $\theta/\sigma_{\text{Noise}}$ as well as statistical uncertainty from each MCMC).

The six gain values measured ranged from 69.9 ADU (a close match to the laser analysis constraint) to 106.9 ADU, with most $\sim 88\,\text{ADU}$. For later use in the fiducial volume calculation (see section 4.4.3), a covariance matrix is estimated for the six gain parameters using the python package `numdifftools` [278], and is given below:

$$\begin{pmatrix} 0.0118 & 0.0220 & 0.0168 & 0.0092 & 0.0206 & 0.0212 \\ 0.0220 & 0.0639 & 0.0411 & -0.0071 & 0.0475 & 0.0520 \\ 0.0168 & 0.0411 & 0.0663 & -0.0378 & 0.0311 & 0.0367 \\ 0.0092 & -0.0071 & -0.0378 & 0.0799 & 0.0107 & 0.0050 \\ 0.0206 & 0.0475 & 0.0311 & 0.0107 & 0.0423 & 0.0044 \\ 0.0212 & 0.0520 & 0.0367 & 0.0050 & 0.0044 & 0.0494 \end{pmatrix} \tag{4.17}$$

The key parameter estimates obtained in this fit were $W_0 = 30.001 \pm 0.005\,\text{eV}$, $U = 15.700^{+0.007}_{-0.009}\,\text{eV}$, and $F = 0.43 \pm 0.05$. The combined MCMC results included nearly negligible statistical uncertainty for W_0 and U. However, because the three separate S30 measurements used as a prior in this analysis do not entirely agree within their respective statistical uncertainties for W_0 and U, an *ad hoc* systematic error for this analysis paradigm was added. This was calculated by fitting a double-sided Gaussian systematic error term for W_0 and U to all three S30 measurements (see section 3.3). After adding this in quadrature, the final results are $W_0 = 30.00^{+0.14}_{-0.15}\,\text{eV}$, $U = 15.70^{+0.52}_{-0.34}\,\text{eV}$.

This result for the W-value does differ from some (but not all [166, 206]) historical measurements [208, 209], as shown in Fig. 4.42. This $\sim 10\%$ discrepancy is consistent with the results of the independent S30 measurements (see section 3.3). Therefore, it is the interpretation of the NEWS-G collaboration that there are SPC-specific effects or systematics that are as-of-yet unknown. Further investigation will be needed to explain this difference between SPC-based and other results, and so this work does not claim this result to be a measurement of the true or intrinsic W-value for methane. Rather, this is taken to be an (accurate) measurement of the ionization yield of methane gas specifically in SPCs, a conclusion bolstered by the independent measurements described in section 3.3. Therefore, this result was used in the LSM dark matter analysis presented in this chapter despite its apparent discrepancy with the literature results.

The result for F from this ^{37}Ar calibration yields an estimate across a range of energies (200 eV to 2.8 keV) of $F = 0.43 \pm 0.05$. This value is significantly higher than typically expected for ionization in many different materials in which previous measurements have been done, including argon ($F \approx 0.23$), xenon ($F \approx 0.17$), silicon ($F \approx 0.16$), and germanium ($F \approx 0.12$) [169–172]. For methane gas in particular, previous measurements of F have been performed at a few energies ($\sim 300\,\text{eV}$ and $\sim 1.5\,\text{keV}$) [279, 280]. These compare favourably to numerical predictions [281],

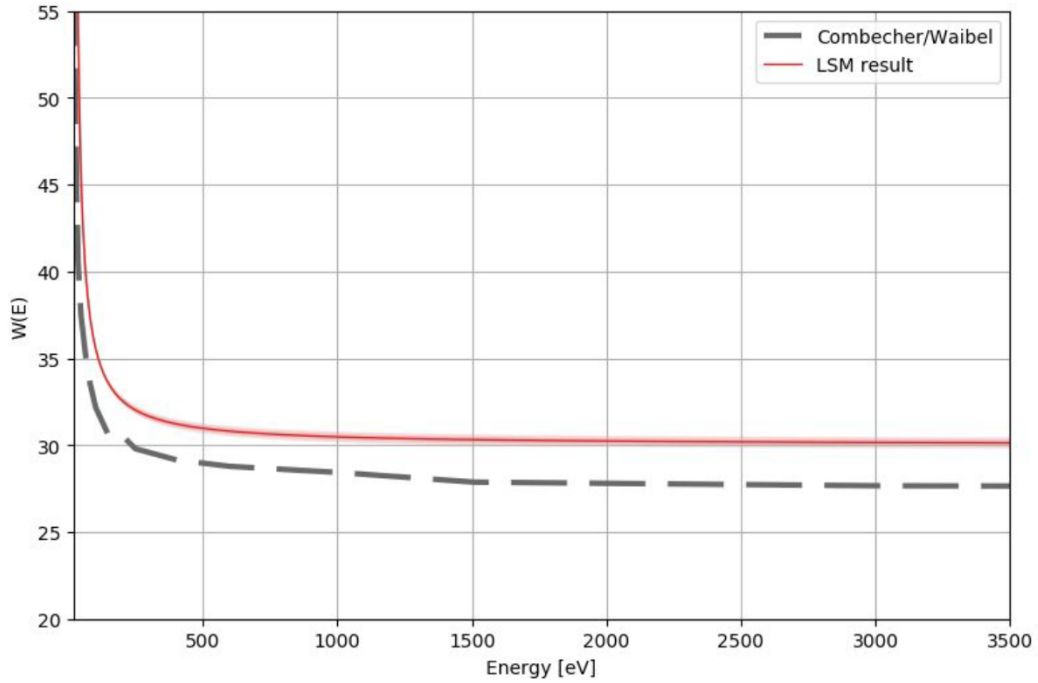


Figure 4.42: W-value as a function of energy obtained from the LSM ^{37}Ar calibration (red curve with $1\,\sigma$ uncertainty band), compared to the combined historical measurements of [208, 209].

which also predict a rise in F at low energies, avoiding the Bernoulli modes (the lowest mathematically possible value of F as a function of energy) [140]. All of these results are depicted in Fig. 4.43.

For the LSM dark matter analysis, while the measured calibration results accurately describe the energy resolution of the SPC, the value of $F = 0.43$ is anomalously high, and so it was taken to be an upper bound on F. To choose a lower bound on F, one may consider the agnostic choice of using the Bernoulli modes (the lowest possible $F(E)$), or a slightly more practical bounding curve of the Bernoulli modes also shown in Fig. 4.43. However, the COM-Poisson lookup tables used to compute the energy-response model of the detector is currently only defined down to $F = 0.1$ [140]. Given that practical constraint, and the agreement of existing measurements [279, 280] with numerical computations [281], the latter was used for this analysis.

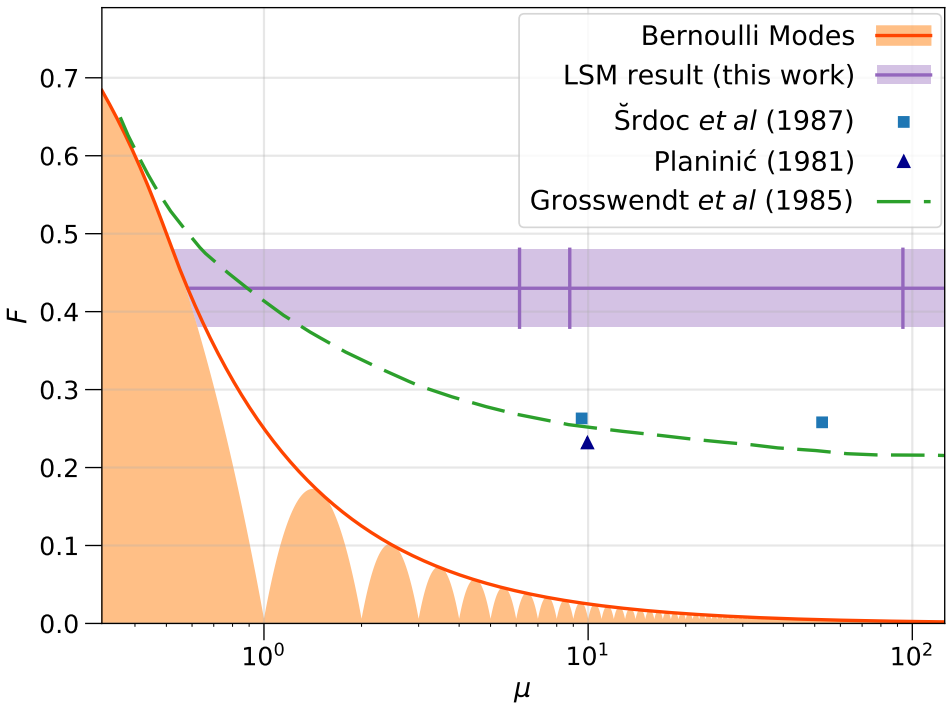


Figure 4.43: Fano factor F as a function of energy (in terms of mean number of primary electrons μ) for this work (purple), compared to existing results (blue points and green curve) [279–281]. The Bernoulli modes (with bounding curve shown, in orange) represent the minimum possible trend for F, a consequence of the discrete statistics of primary ionization.

4.3.4 Methane quenching factor

The final (and most critical) component of the SPCs energy response is the quenching factor of methane gas. As discussed in section 3.1.2, the quenching factor expresses the ratio of the primary ionization yield of nuclear recoil events (such as dark matter) compared to electronic recoil events, as given in eq. 3.9. Since the LSM analysis is performed only for hydrogen recoils in methane, it is the quenching factor for this interaction that must be known.

While the Lindhard theory [177] and simulation software SRIM [183] can be used to calculate the quenching factor, they frequently disagree with empirical results (see section 3.1.2). However, in the case of this study, there are limited empirical results to look to. One example is the COMIMAC collaboration in France, which has used their table-top ion accelerator to perform quenching factor measurements [180, 181]. In this case, the accelerator was used to fire protons into a methane-filled SPC through a microscopy hole (window) through a foil layer. These measurements spanned incident energies from 2 to 13 keV [282]. The other existing result available is W-value measurements of protons in methane, made in the context of dosimetry studies [283]. These results were re-interpreted as QFs by taking the ratio of W_{ion} to W for electronic recoils (from refs. [209] and [208]) [284]. This study yielded an estimate of the quenching factor for protons in methane from 0.51 to 375 keV [284].

Both of these existing results are less than ideal; the COMIMAC experiment faces challenges related to the deflection of ions as they pass through the foil window, while the W_{ion} reinterpretation uses measurements made not in an SPC. Neither takes into account the pressure of the gas used. Furthermore, neither result extends down to the very low-energy regime where the majority of dark matter sensitivity originates. Regardless, these are the only experimental results currently available. Future quenching factor measurements are planned by NEWS-G to improve this situation, including both the TUNL beam accelerator facility [184], or possibly neutron capture

experiments (see section 3.4.3).

When comparing the methane quenching factor results, we favour the COMIMAC measurement because it was made with an SPC, and because it is globally lower than the W_{ion} quenching factor, and is thus a conservative choice for this analysis. To extend the energy range in which the quenching factor is known, the W_{ion} trend is scaled down (by a factor of 0.874) to match the COMIMAC quenching factor where the two measurements overlap. Below the lowest-energy measured point, several extrapolation options were considered. First – and most optimistically – a Lindhard-like extrapolation was defined by fitting a bounded power-law approximation of the Lindhard theory:

$$Q(E) = \frac{mkE^{\alpha}}{(1 + kE^{\alpha})} \tag{4.18}$$

Fitting the (scaled) datapoints between 0.51 and 2 keV yielded the parameter values $m = 0.83$, $k = 1.02$, and $\alpha = 0.94$. This extrapolation implies that the quenching factor only reaches 0 at $E = 0\,\text{keV}$.

A more conservative (albeit not physically motivated) parameterization is a generic logarithmic function of the form:

$$Q(E) = a + b \log E \tag{4.19}$$

with $a = 0.44$ and $b = 0.23$ from the data below 2 keV. A final, extreme option is to assume that $Q(E) = 0$ below the lowest energy data point. While indisputably conservative, this approach is also almost certainly not physical. Ultimately, the logarithmic extrapolation was used as the main quenching factor parameterization below 0.51 keV for this analysis. This was continuously joined with a power law fit of the scaled data between 0.51 and 2 keV (with $m = 0.83$, $k = 1.02$, and $\alpha = 0.94$), and with a similar power law model fit of the COMIMAC data above 2 keV:

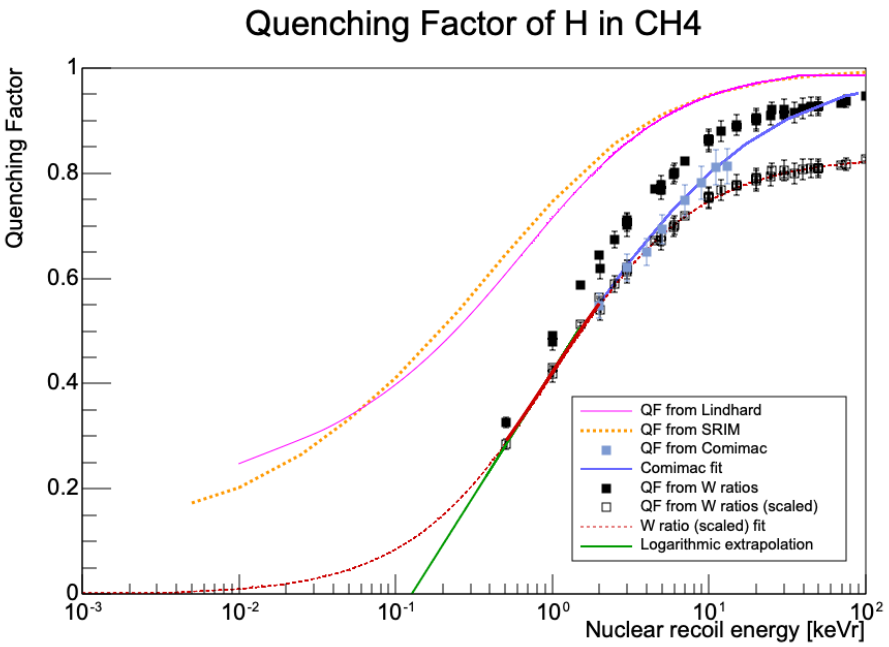


Figure 4.44: The quenching factor for protons in methane gas, including COMIMAC measurements (blue markers) [282], as well as the nominal W_{ion} literature results (black markers) [284] and scaled data points (empty square markers). The Lindhard theoretical model and quenching factor calculated with SRIM are shown as the pink and orange-dashed curves respectively. The logarithmic fit is shown as the green curve, and the power law fit in red. The extrapolations of both models are shown as dashed curves. Credit to F. Vazquez.

$$Q(E) = \frac{E^{0.70}}{1.37 + E^{0.70}} \tag{4.20}$$

The various quenching factor parameterizations and data are shown in Fig. 4.44. The model used for the dark matter analysis was derived by F. Vazquez.

4.3.5 Diffusion model

In order to use PSD variables such as risetime or Δt_{peak}, the electron transport behaviour of the SPC must be known. Specifically, a model for the diffusion of electrons as a function of their initial radial position in the sphere is required. The first attempts to do so for the S140 methane campaign involved simple first-principles electron transport simulations, as described in section 3.1.3. These were completed by Y. Deng using the COMSOL simulation of the S140 detector introduced in section 4.3.3.

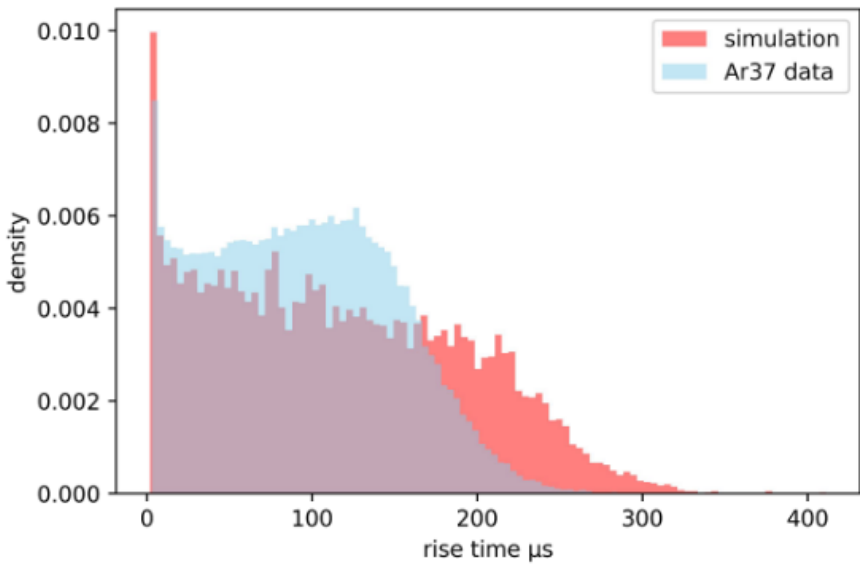


Figure 4.45: The $10-75\%$ risetime distribution of 2.8 keV ^{37}Ar events (from run tj13s000, blue histogram) compared to the first-principles electron transport simulation performed for the LSM methane campaign of ^{37}Ar events distributed uniformly throughout the detector volume (red histogram). Credit to Y. Deng.

Magboltz was used to provide the diffusion coefficients and drift velocity as a function of electric field strength [186], and a simulation following the steps listed in section 3.1.3 was carried out. Since dark matter candidate events would be homogeneously distributed throughout the SPC volume, ^{37}Ar (which is also a volume source) was used as a test to validate the simulation. A comparison of the first-principles simulation and ^{37}Ar risetime data is shown in Fig. 4.45 – while similar, the two distributions do not agree. This disagreement results in significantly erroneous diffusion/risetime predictions for the LSM dark matter signal, and is not reconcilable by modifying the electron transport parameters of the simulation, or electric field model.

The presumed cause of this discrepancy is the impact of space charge effects, which are not accounted for in the first-principles simulation. As shown in section 4.2.5, the coherent space-charge effects of high-intensity laser events, ^{37}Ar, α-decays, etc. can significantly alter electron drift times (and presumably diffusion as well). Further evidence that space-charge effects are the key physical process omitted is the fact that first-principles drift time simulation of low-intensity laser events (which are much less

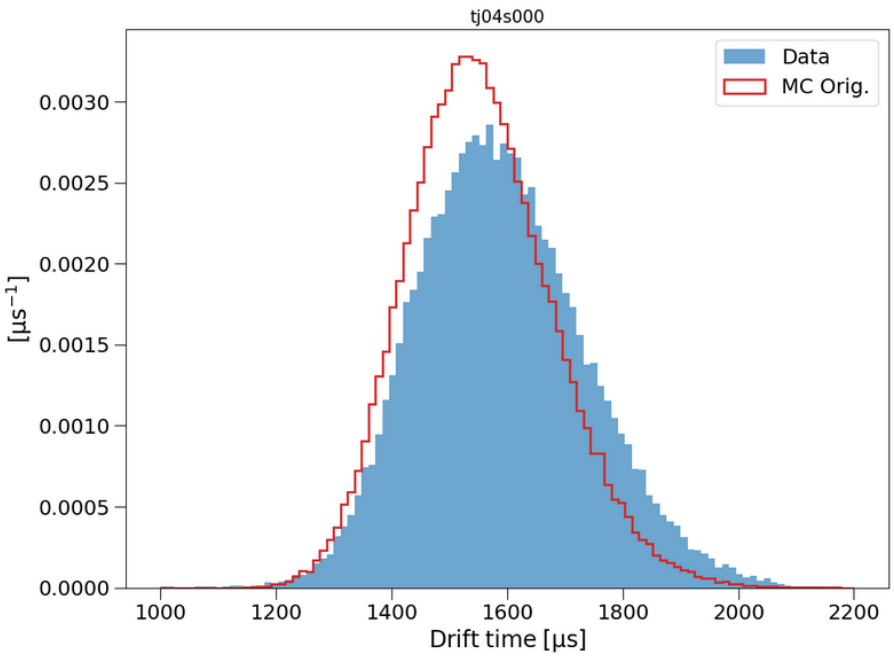


Figure 4.46: Drift time data from low-intensity laser run tj04s000 after applying post-alpha cuts (see ahead to section 4.4.4, blue histogram), compared to the first-principles electron transport MC prediction of drift time in S140 (red histogram).

affected by space charges) more closely matches the data, as can be seen in Fig. 4.46.

Although the average space-charge conditions during low-intensity laser runs is thought to be small (see Fig. 4.23), fluctuations in the SPCs electric field structure on smaller time-scales were suspected of playing a role (event-to-event space charge fluctuations). To demonstrate this, a risetime distribution for tj04s000 was built by drawing drift times from the total data distribution (Fig. 4.46). The simulated risetimes were then calculated by drawing random avalanche yields from the Polya distribution (with measured $\theta = 0.125$, see section 4.3.1) for each electron. Next, approximate integrated pulses were constructed as a series of step functions (defined by the randomly drawn drift times and avalanche amplitudes), from which the risetime can be extracted. The result of this simulation also does not match the risetime data, despite not relying on the electron transport MC in any way. This method assumes that the approximate method of calculating risetimes for simulated events is valid, so a more robust line of reasoning is presented below.

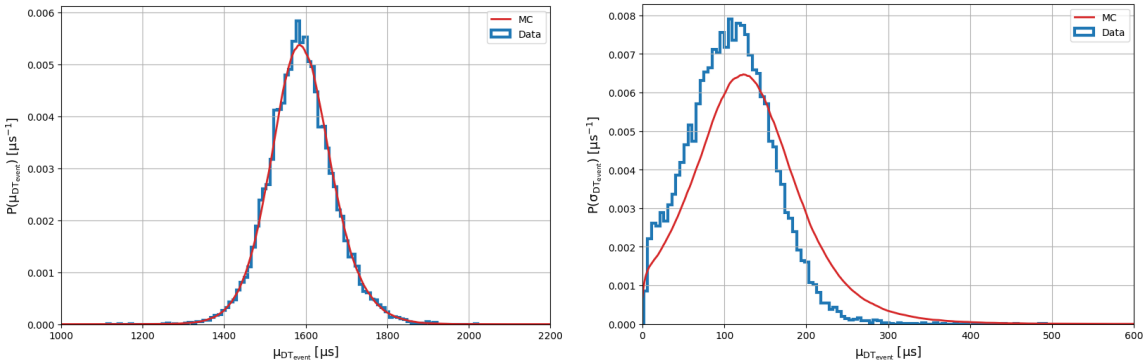


Figure 4.47: Left: mean drift time per event of tj04s000 data (blue histogram) compared to the MC reconstruction using the total drift time distribution (red histogram). Right: standard deviation of drift time per event of tj04s000 data (blue histogram) compared to MC reconstruction using the total drift time distribution (red histogram).

If it is assumed that the drift time distribution of the data shown in Fig. 4.46 is true for the whole dataset (space-charge and electric field conditions are static, or at least in equilibrium), then one should be able to reproduce the drift time characteristics of the data on an event-by-event basis. Specifically, the mean and standard deviation of the drift time in each event are calculated for tj04s000. To reproduce this with an MC based on the total drift time distribution, the following procedure is used for each simulated event:

1. Draw a random number of peaks from the distribution of peaks per event of the data,

2. For each peak, draw a drift time using the total data distribution (Fig. 4.46),

3. Calculate the mean and standard deviation of the drift time in that event.

The results of this boot-strapped simulation do not match the per-event drift time data, as can be seen in Fig. 4.47. In particular, the standard deviation of drift time per event is much smaller for the data than it is for simulation, suggesting the dispersion of the total distribution comes partially from a changing mean drift time. Although

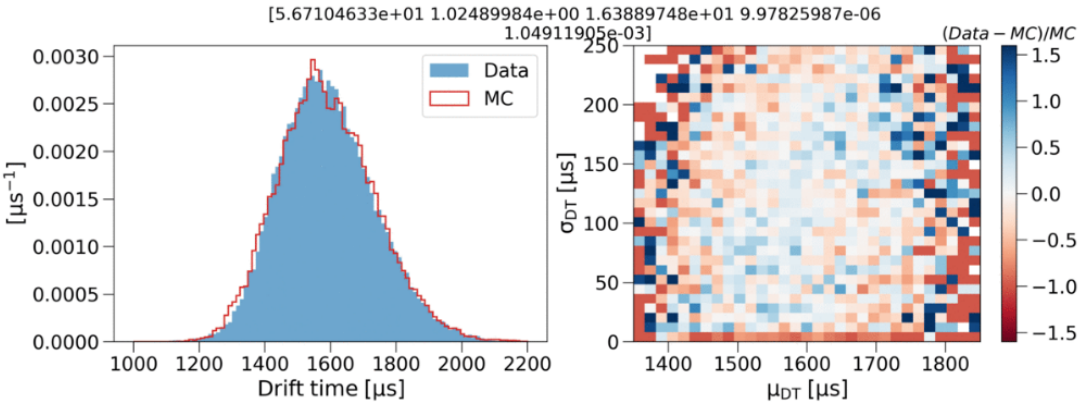


Figure 4.48: Comparison between the per-event (residuals of the mean and standard deviation drift time per event) and total drift time data of tj04s000 (right and left respectively) and semi-empirical drift time model, with parameters shown above (not necessarily the best-fit).

the discrepancy of the mean drift time per event is more subtle, a Komologrov-Smirnov test [218] performed indicates a p-value of only 0.1, so the MC model can be rejected at the 90% confidence level (but not higher). Altogether, this boot-strap MC refutes the idea that the electric field conditions in the sphere are static; non-negligible fluctuations occur on time-scales much shorter than the duration of the whole run, plausibly caused by changing space-charge densities.

Semi-empirical diffusion model At the time of writing, no *a priori* model of space charge effects and subsequent electron transport behaviour is available. Therefore, in order to accurately model electron transport in S140, the data itself (in which the space charge fluctuations are encoded) must inform the model in some way. This was done by applying a stochastic change of variables (to reflect the stochastic nature of the fluctuations) to the original electron transport simulation results. Specifically, the drift time distribution of the original MC $P_{\mathrm{MC}}(DT)$ was stretched and translated by parameters f and ω respectively:

$$P_{\mathrm{MC}}(DT) \rightarrow P'(DT) = P_{\mathrm{MC}}(f \cdot (DT - \omega)) \tag{4.21}$$

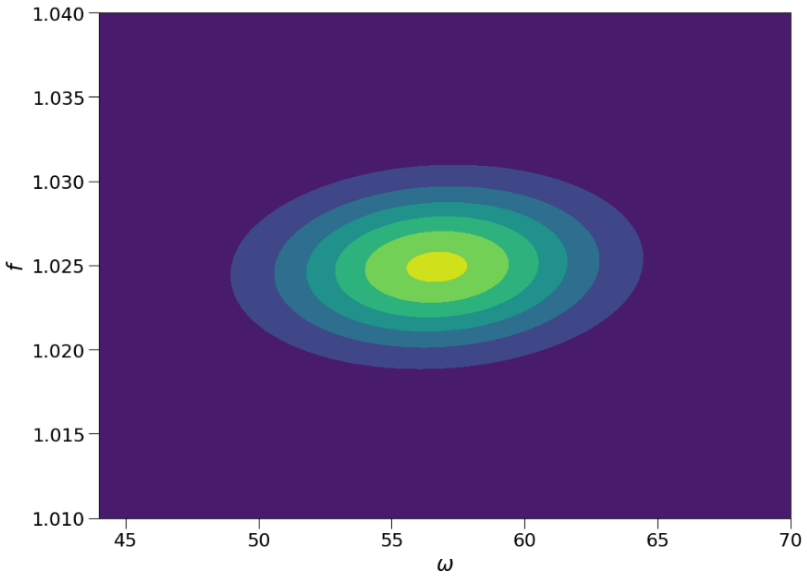


Figure 4.49: Best-fit correction of the original electron transport MC; a bivariate Gaussian of the stochastic stretch (f) and translation (ω) applied to the drift time data.

where f and ω are random variables drawn from a bivariate Gaussian. Thus the free parameters of the model are those which characterize the 2D Gaussian; μ_f, μ_ω, σ_f, σ_ω, and $cov(f, \omega)$.

The model was fit jointly to the event-level drift time data of tj04s000 (mean and standard deviation drift time per event), and total drift time distribution, since it must produce a result that matches both. A comparison of the model to both the one-dimensional and 2D data is shown in Fig. 4.48 (generated during the optimization process), showing how the drift time data is reproduced by reconstruction through this model. The best-fit bivariate Gaussian to correct the original drift time simulation is shown in Fig. 4.49; the average stretch and translation are around 1.025 and 57 μs respectively, with little covariance between the two parameters. This optimization was performed by using the `scipy.optimize.minimize` function with a Nelder-Mead algorithm [238, 239] to minimize the residuals shown in Fig. 4.48.

Using the corrected semi-empirical model, a new risetime simulation for S140 was performed. This was done for tj04s000 with the following procedure:

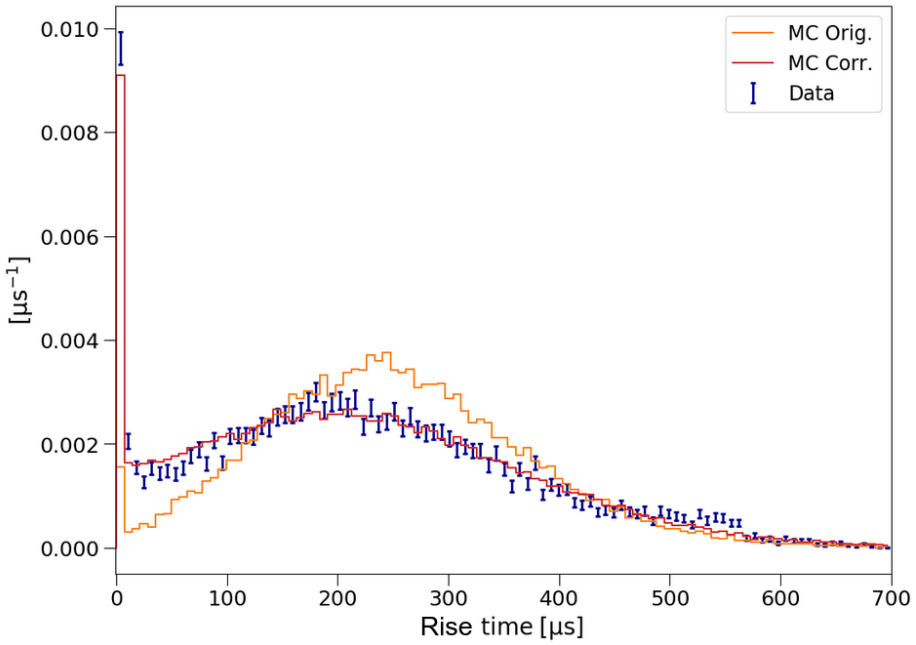


Figure 4.50: Risetime data from tj04s000 after applying post-alpha cuts (blue histogram), compared to the simulated risetime distribution of the original electron transport Monte Carlo (MC; orange) and semi-empirical MC (red).

1. For each simulated event, draw a random number of peaks from the distribution of peaks per event of the data,

2. For each simulated event, draw a random stretch and translation correction (f and ω) from the distribution shown in Fig. 4.49, which are then applied to the original electron transport MC drift time distribution P_{MC},

3. Draw an appropriate number of drift times for the event from the corrected distribution,

4. Calculate risetime as usual; create a series of step functions for each electron in the event (with randomly drawn amplitudes from the Polya distribution with $\theta = 0.125$) to emulate the integrated pulse, and calculate risetime.

A comparison of this result and the data is shown in Fig. 4.50, in which one can see there is much better agreement than with the original electron transport simulation. While this semi-empirical approach appears to successfully represent the fluctuating

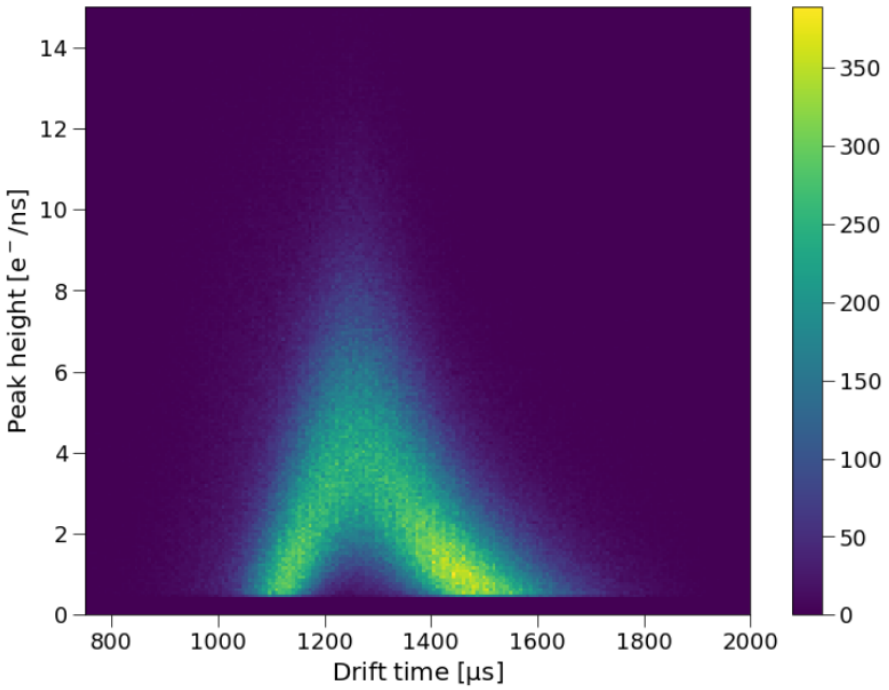


Figure 4.51: 2D histogram of peak height vs. drift time data for tj13s000, showing the effect of Samba's centering algorithm.

space-charge conditions of tj04s000 (low-intensity laser data), this same procedure fails to reproduce the risetime data of higher-intensity data (such as tj13s000). A possible culprit of this problem is the centering algorithm of Samba, which tends to center either high amplitude electrons, or in the case of runs like tj13s000, pileup of many primary electrons. This leads to the odd relationship between individual peak heights and drift time shown in Fig. 4.51, which is not accounted for in the simulation described here.

Generic functional diffusion model While the semi-empirical approach described above was useful to help understand the nature of space-charge effects in the SPC (and may provide the impetus for future simulations), a simpler approach was needed for the present dark matter analysis. The direct, fully empirical approach taken was to model diffusion as a function of radius with a completely generic parameterization. This analysis was performed by F. Vazquez. In this case, ^{37}Ar data (tj13s0000)

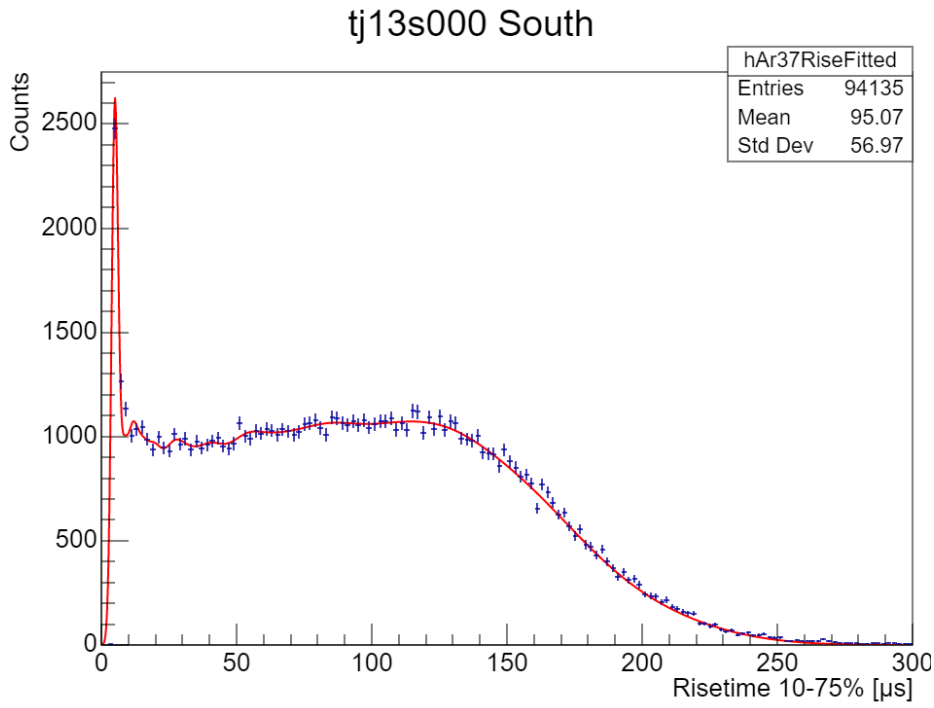


Figure 4.52: 10-75% risetime data of 2.8 keV ^{37}Ar events from run tj13s000 (blue histogram with statistical error bars). The sum-of-Gaussians fit (red curve) is the sum of 27 risetime distributions simulated for varying diffusion times. Credit to F. Vazquez.

provides the source of volume events used as a reference.

First, the relationship between reconstructed risetime and diffusion time $\sigma_{\text{Diff.}}$ – the standard deviation of electron arrival times – was determined by simulating 2.8 keV ^{37}Ar events with given diffusion times. The number of electrons simulated in each MC event was determined using the $\langle G \rangle$ and W-value results of sections 4.3.3, and the reconstructed risetimes were calculated using the step-function approach described in section 3.1.3. The distribution of $10-75\%$ risetime for a given diffusion time was found to be approximately Gaussian, with a standard deviation equal to $0.145\times$ the mean risetime.

The $10-75\%$ risetime data of tj13s000 (2.8 keV events only) was fit with the sum of such Gaussian distributions for risetime (with standard deviation equal to 0.145 times the mean risetime); 27 semi-evenly spaced Gaussians were used. This fit is

shown in Fig. 4.52. The scaling of each of these Gaussian contributions gives the relative probability of ^{37}Ar experiencing the corresponding amount of diffusion, and can then be mapped to the radial position of the events. This is done by matching the probability and assuming that diffusion increases monotonically with radial position, using the known radial probability distribution of these volume events:

$$\rho(r) = 3\frac{r^2}{r_{\text{max}}^3} \tag{4.22}$$

The resulting trend of diffusion time vs. radial position is shown in Fig. 4.53. Finally, this trend was parameterized using a generic functional form used by NEWS-G historically [63]:

$$\sigma_{\text{Diff.}} = \sigma_{\text{Diff.max}} \left(\frac{r}{r_{\text{max}}}\right)^{\alpha} \tag{4.23}$$

where $\sigma_{\text{Diff.max}}$ (maximum diffusion time) and α are free parameters in the model. The resulting diffusion model is described by $\sigma_{\text{Diff.max}} = 98.6\,\mu$s and $\alpha = 3.05$ with the following covariance matrix (for the parameters in that order):

$$\begin{pmatrix} 46.2 & 0.856 \\ 0.856 & 0.0225 \end{pmatrix} \tag{4.24}$$

An appropriate diffusion model is used to subtract background surface events. For this, the risetime spectra of both the high-intensity laser data taken during a physics run, as well as the physics data itself, was examined (from tj04s002 specifically). The risetime distributions of these two even populations are shown in Fig. 4.54. Curiously, there were slightly higher risetimes for laser events compared to the physics data (which is completely dominated by surface background events). The probable cause of this is that the laser was found to be partially shining on the portion of the inner SPC surface from which electrons drift to the north channel of the sensor, instead of the south. Evidence for this is presented in section 4.4.6. Therefore, the surface background data of tj04s002 was used to model diffusion for surface events,

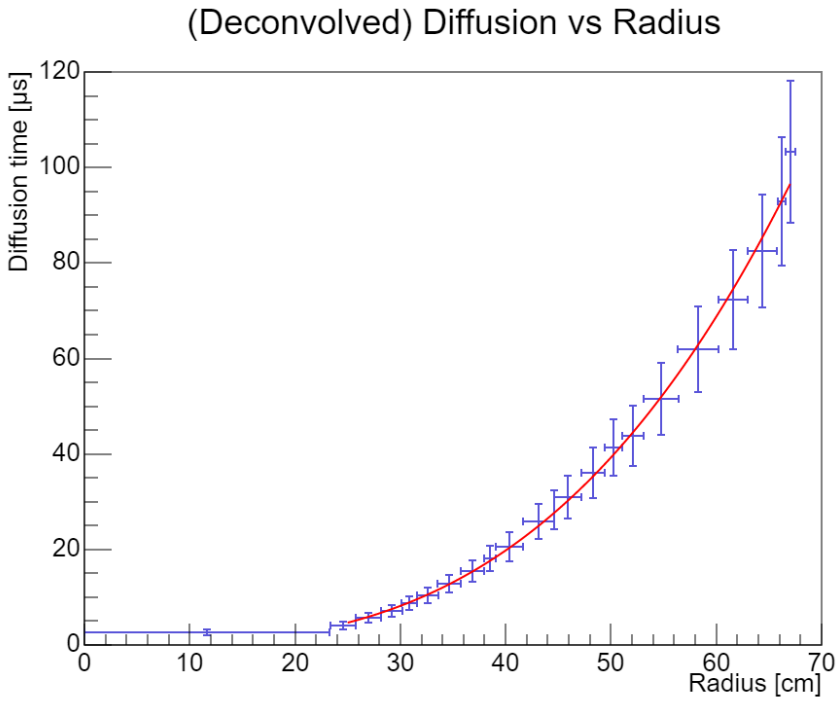


Figure 4.53: Diffusion time vs. radius trend for 2.8 keV ^{37}Ar events in run tj13s000 (markers with error bars), fit with eq. 4.23 with $\sigma_{\mathrm{Diff.max}} = 98.6\,\mu$s and $\alpha = 3.05$. Credit to F. Vazquez.

from which the average diffusion time was measured to be $123 \pm 1.1\,\mu$s. While the $10-90\%$ risetime is typically equal to $2.57\times$ the diffusion time (assuming a perfectly Gaussian step function integrated pulse), an extra smearing by a factor of 14.5% was needed to match the risetime distribution of the data in this case. This is again likely due to fluctuating space-charge effects.

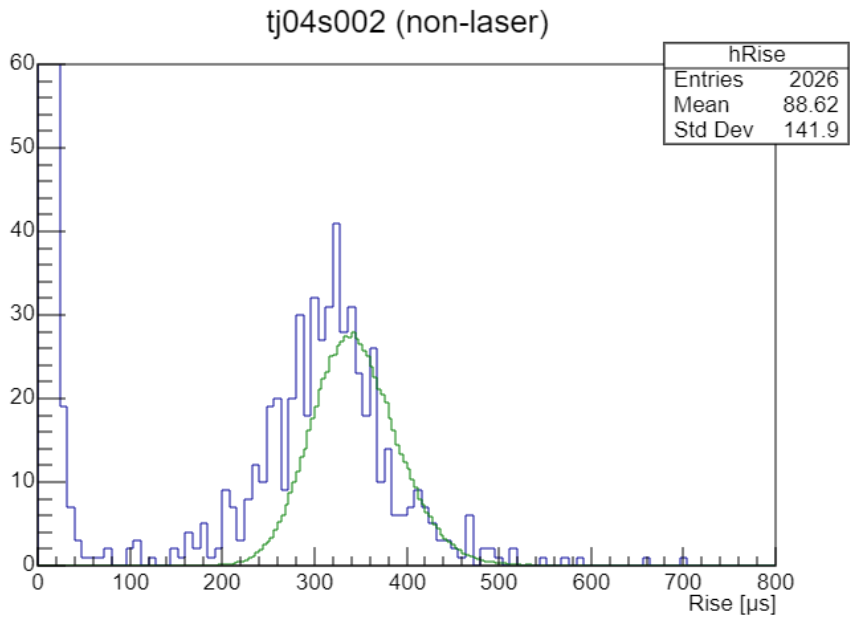


Figure 4.54: Risetime distribution of surface background events from run tj04s002 (blue histogram), compared to the spectrum for high-intensity laser events from the same run, normalized by maximum distribution height (green histogram). Standard PSD and α cuts are applied to select the surface background events, as well as the cut $3000 < \text{DD\_AmplADU} < 25000$, corresponding to an energy range of approximately 1.5 to 12.5 keV. Credit to F. Vazquez.

4.4 LSM campaign: Data cuts and signal acceptance

Once the physics of the hypothesized dark matter signal has been modelled (see the previous section), the signal acceptance must also be calculated. This includes the effect of cuts/data selection, the hardware trigger efficiency of the experiment, and the loss of livetime due to the DAQ algorithm.

4.4.1 Hardware trigger efficiency

The hardware trigger efficiency of the detector is the probability that Samba (the DAQ software, see section 3.1) was triggered by an event. Naively this is only a function of the amplitude of the event, with some Gaussian error function of pulse amplitude representing the trigger efficiency function. In prior NEWS-G experiments, this was approximately true. However, in the few-electron regime, and in circumstances with significant electron diffusion such as with the LSM data, it also depends on the separation of electrons over time. Consider for example an event with two primary electrons, each with an amplitude such that separately they have a 50% trigger probability. In an experiment with little diffusion in which the electrons arrive nearly simultaneously, the amplitudes of the two electrons would add, presumably leading to a $\sim 100\%$ trigger probability. However, if spaced apart their amplitudes will not add together, and the combined trigger probability for the event would only be 75%. Therefore, calculating the trigger efficiency of the experiment will depend not on the total energy of an event, but instead on the number of electrons, the diffusion model of the experiment, and the radial position of the events in question (i.e. surface vs. volume events).

The diffusion model is described in section 4.3.5, but the energy dependence of the trigger efficiency must be measured experimentally. Since laser data provided a source of tagged calibration events with few electrons, this data was used to calculate

the energy dependence of the hardware trigger for single electrons only. The laser run ti30s002 was used, as it was the lowest intensity laser run available. For this analysis, the data was processed in nearly the same way as the physics data – integrating the double deconvolved pulses over a 1090 sample-wide integration window, starting at 3660 samples. Note that this is not the same special processing done for the laser data as described in section 4.3.1, but integrating the entire pulse.

The first step after processing was to determine the mean number of primary electrons per event. To simplify fitting the very broad amplitude spectrum that is created with this processing, several parameter values are assumed from other analyses. The average value of $\theta = 0.125$ from all low-intensity laser datasets is used, and the mean gain $\langle G \rangle$ is obtained by extrapolating the gain trend shown in section 4.3.2 back to the time of ti30s002. The amplitude spectrum is then fit assuming a Poissonian distribution of primary electrons (see section 3.2.1), the average number of which was found to be $\mu = 1.215 \pm 0.033$. As desired, this data contained mostly null and single-electron events.

The next step was to apply the SPC-channel trigger to this photodiode-triggered data. The DAQ hardware trigger was emulated in software (in the `quadis` code library [160]), and applied to the ti30s002 data. This was done assuming both the early and late trigger settings (conditions 1 and 2 respectively, see section 4.2.1) of the campaign. Events with emulated DAQ triggers within the integration window were considered to have a successful SPC-trigger for this analysis. The effect of applying either trigger conditions (or both) to the amplitude spectrum of ti30s002 is shown in Fig. 4.55 and shows that the early trigger conditions were universally worse than the later trigger conditions. To simplify the overall analysis, the conservative decision was made to consider only the early trigger conditions going forward.

The amplitude-dependence of the trigger efficiency was modelled as a Gaussian error function of the double deconvolved amplitude (`DD_AmplADU`) for a single peak in isolation, with free parameters for the location (ω_{trigger}) and spread (σ_{sigma}) of the

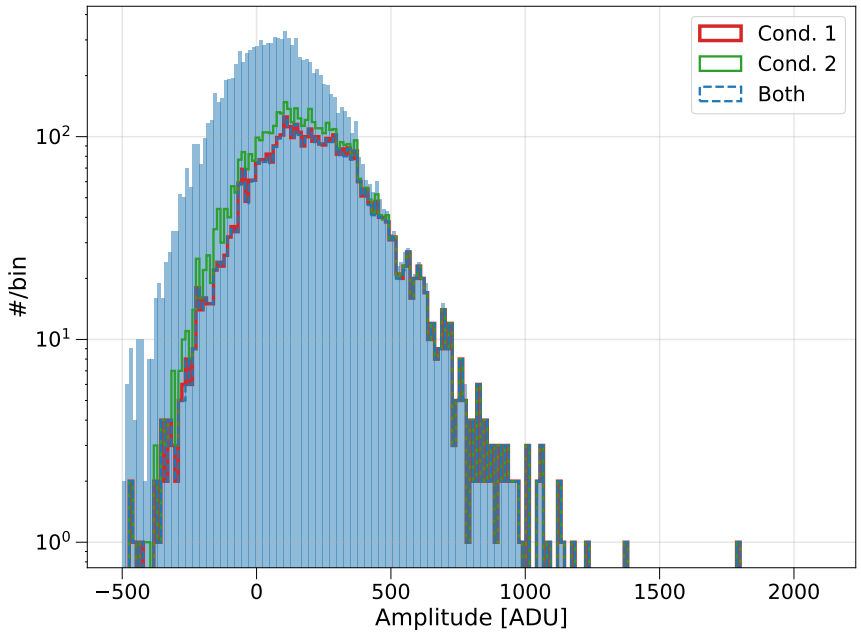


Figure 4.55: Amplitude spectrum of laser run ti30s002 (using a wide integration window) shown in blue. The effect of the emulated DAQ hardware trigger cuts applied with the early and late trigger conditions of the campaign are shown in orange and green respectively, as well as the effect of applying both conditions (red).

efficiency curve. Of course, there were many events with multiple electrons, which may or may not overlap. The trigger for overlapping peaks behaves the same as for one, larger peak. For peaks sufficiently far apart, the trigger efficiency curve applies to both separately, each with a chance to successfully trigger the DAQ. Therefore the model fit to the data was calculated using an MC with the following steps:

1. A random number of electrons was drawn for each event according to the fit of the ti30s002 data discussed above (a Poisson random number with $\mu = 1.215 \pm 0.033$). Peak times were drawn using the surface diffusion model (see section 4.3.5). Electrons close enough in time to overlap were considered to be clustered together. The cutoff time was chosen to be 25 μs – the derivative calculation window used for the early trigger settings.

2. For each cluster, random amplitudes were drawn using the interpolated gain for ti30s002, and the overall average value of $\theta = 0.125$. Clusters with multiple

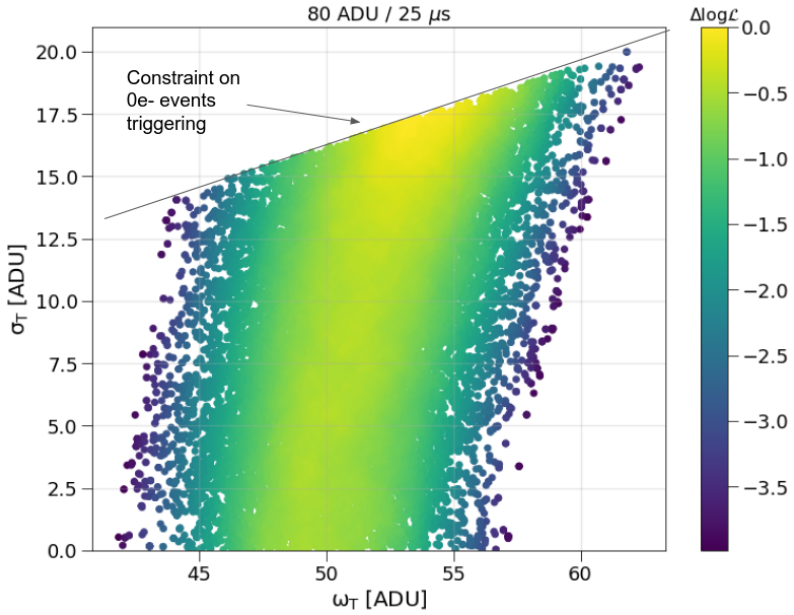


Figure 4.56: Posterior likelihood function (MCMC samples) of the fit of the spherical proportional counter-triggered ti30s002 amplitude spectrum as a function of ω_{trigger} and σ_{trigger}. A constraint based on the probability of 0-electron triggers is labelled as well.

electrons were assigned the sum of the individual amplitudes drawn.

3. For every cluster in an event, the trigger probability was determined for given values of $\omega_{\text{trigger}}/\sigma_{\text{trigger}}$, and a random success-draw was done. If one or more cluster successfully triggered, the whole event was considered to trigger. The amplitudes of the clusters were added to give the total amplitude, and convolved with baseline noise.

This model – with free parameters ω_{trigger} and σ_{trigger} – was then fit to the amplitude spectrum of ti30s002 with the early-conditions hardware trigger cut applied (see Fig. 4.55). The fit was performed based on a binned likelihood function, and optimized with an MCMC [216]. Specifically, 20 MH walkers were used, with a 500 step burn-in and then 2500 sample run (compared to the auto-correlation time of 35 and 42 samples for ω_{trigger} and σ_{trigger} respectively). The likelihood function for this model/data was extremely degenerate, and allowed very high values of both $\omega_{\text{trigger}}/\sigma_{\text{trigger}}$), which

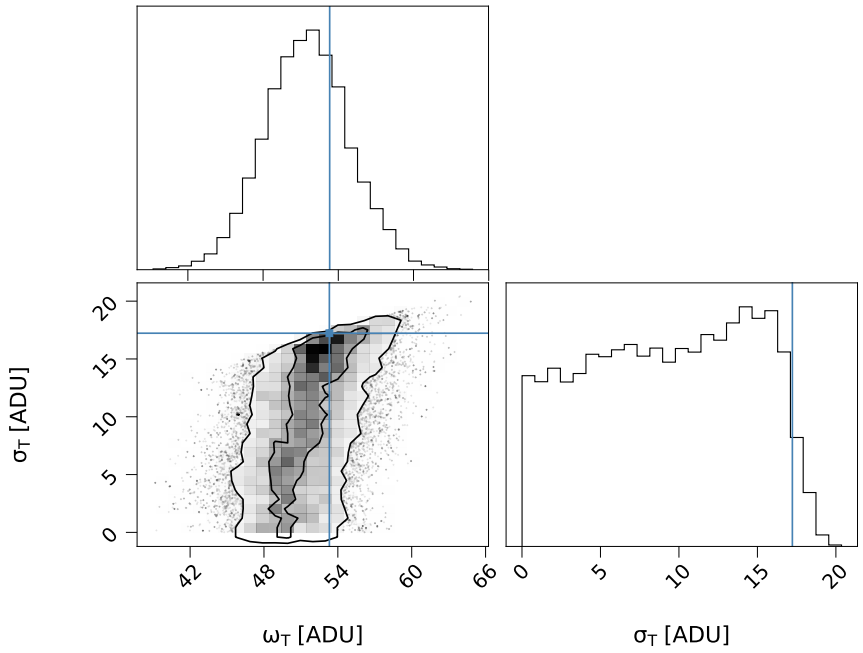


Figure 4.57: Corner plot of the MCMC-sampled hardware trigger efficiency likelihood function with ω_{trigger} and σ_{trigger}. The bottom left panel is a 2D histogram of the MCMC samples with contours for the "1 and 2 σ" levels (if the posterior was Gaussian), and a scatter plot of the samples beyond. The best-fit values are indicated in blue.

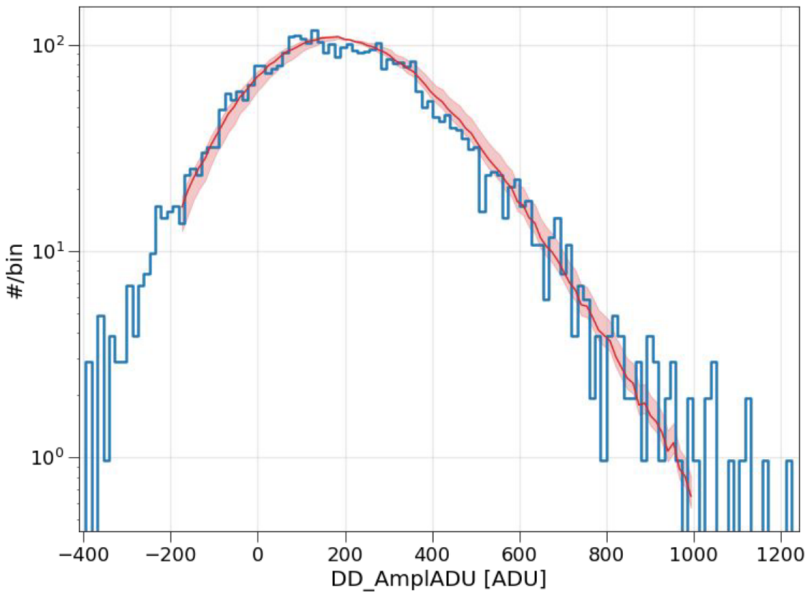


Figure 4.58: Amplitude spectrum of ti30s002 applying the early-condition hardware trigger cut (blue histogram), and resulting fit of the data with a hardware efficiency curve applied (red curve with $1\,\sigma$ uncertainty band).

would imply a large number of null-events triggering the data acquisition. Therefore a constraint was applied to the likelihood based on the probability of null-event triggers observed for a simulated data set with only null events, which was 10^{-6}. The MCMC sampling of the likelihood function with this constraint is shown in Fig. 4.56, and a corner plot of the likelihood function is shown in Fig. 4.57. The result of the MCMC is best-fit values of $\omega_{\text{trigger}} = 53.3\,\text{ADU}$, and $\sigma_{\text{trigger}} = 17.2\,\text{ADU}$, and the fit of the data is shown in Fig. 4.58. The unusual shape of the likelihood function precludes the calculation of a covariance matrix or uncertainties with any standard recipe, so the MCMC samples from the fit were used directly in subsequent error calculations. Note that this calculation incorporates uncertainties on $\langle G \rangle$ (one anode, for this laser data), θ, the baseline noise of the full-window integration, peak-finding performance parameters, and the surface diffusion model, as well as statistical uncertainties from the fit.

Applying this result, the hardware trigger efficiency for a given number of electrons was calculated with the same MC described above, depending on the initial radial

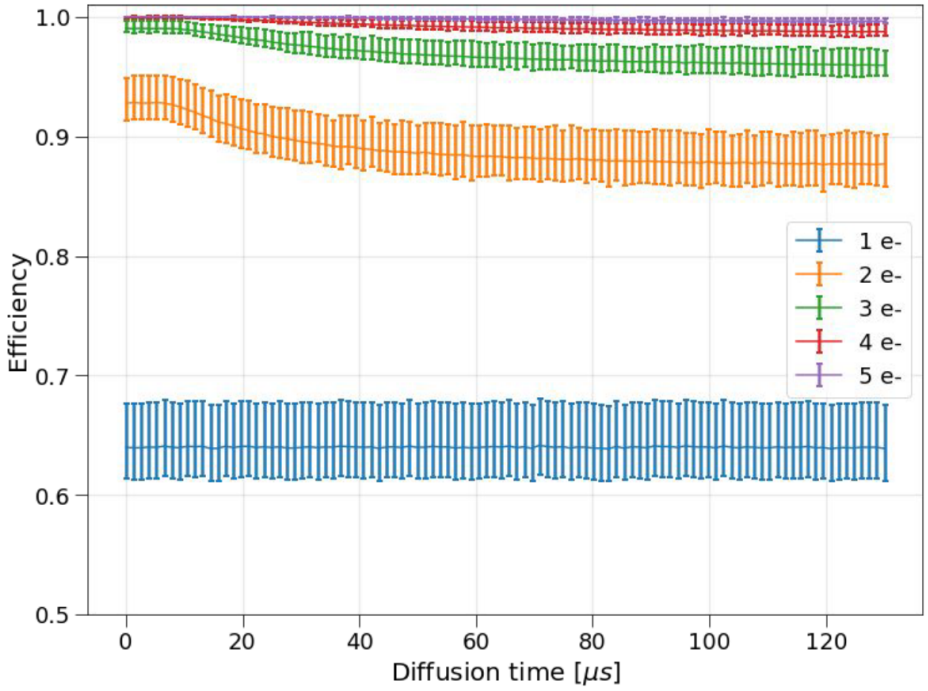


Figure 4.59: Hardware trigger efficiency as a function of number of primary electrons, in bins of diffusion time.

position of the electrons. To cast this result in a more useful fashion for the next steps of the analysis, the hardware trigger efficiency was calculated in diffusion-time bins (so that the result is applicable to surface or volume events), which is shown in Fig. 4.59. This final result includes uncertainty on the gain of all 6 anodes (using the results of section 4.3.3). As the gain decreased slowly over time, the trigger efficiency also decreased, and so the conservative decision was made to perform this calculation interpolated at the median time of tj10s001 – the last physics run. The efficiency for single electrons is of course fixed at any radius/diffusion time at 64^{+4}_{-3} %. For two electrons, the hardware trigger efficiency was as high as 93^{+2}_{-1} % (for events near the sensor).

4.4.2 Peak selection efficiency

Using the PF results described in section 4.2.1, the LSM physics data was partitioned in subsets with 2, 3, and 4 peaks for the final dark matter analysis presented in sec-

tion 4.5. Events with more peaks (corresponding to proportionally more electrons) are beyond the energy region of interest for the low-mass dark matter searches performed by NEWS-G. On the other hand, the single peak (mostly single electron) rate is anomalously higher than the expected rate from background events. This has historically been the case for NEWS-G experiments (see section 2.4.1), and in this case is thought to be at least partially caused by α events, which may produce some single electron events that leak past the cuts described in section 4.4.4. Therefore, these events are omitted from this analysis. The PF algorithm results are also used to remove any events with a peak in the north channel (see section 4.4.6).

Of course, in selecting data based on the number of peaks, the defects/failures of the PF algorithm itself must be taken into account to calculate the resulting signal acceptance for dark matter events. This characterization was partially informed by the peak-resolved low-energy laser analysis presented in section 4.3.1. Specifically, the PF efficiency threshold results described by σ_h and ω_h (see table 4.4) and eq. 4.8 were used to calculate the efficiency of the algorithm for identifying single electrons $P_{1\,\mathrm{e}^-}$:

$$P_{1\,\mathrm{e}^-} = \int P_{\mathrm{Polya}}(E|\theta, \langle G \rangle) \times f(E|\omega_h, \sigma_h) dE \tag{4.25}$$

using the values of $\langle G \rangle$ at the time of each run, and the global average value $\theta = 0.125$. This result was calculated for all runs, and is shown in Fig. 4.60, including the average value across all runs with statistical uncertainty. The generally-decreasing trend over time is expected, as the detector gain was decreasing relative to the fixed PF algorithm threshold (see section 4.3.2).

Another limitation of the PF method is the probability of FPs (noise peaks). The laser analysis provided an estimate of this quantity, but was only applicable to the small window sizes of the specially-processed laser data. For normally-processed events, P_{FP} was calculated directly by applying the PF algorithm to pre-trace data (see Fig. 4.15), which yielded a mean number of FPs per event of 3×10^{-4}.

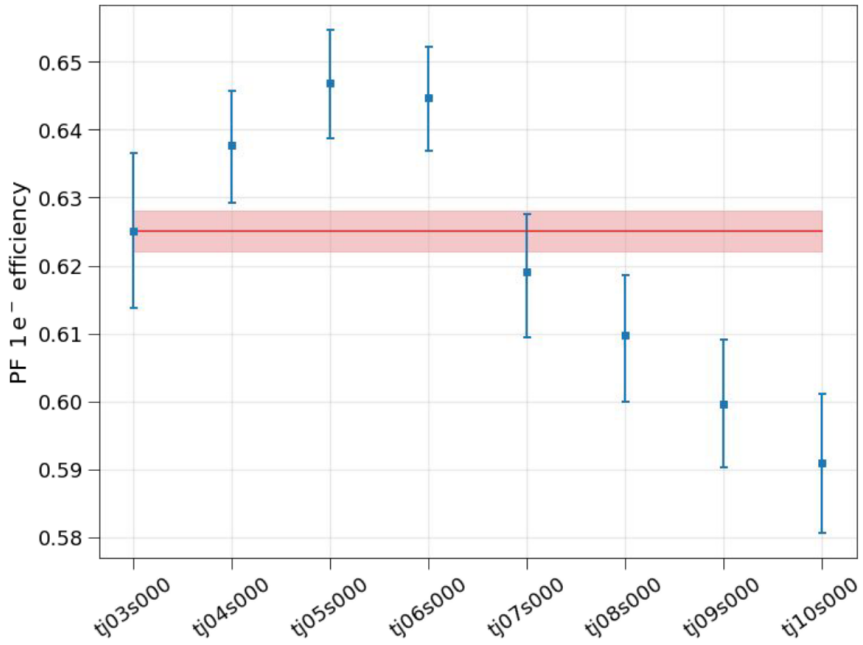


Figure 4.60: Single electron efficiency of the peak finding algorithm shown for all low-intensity laser calibration fits (blue points), as well as the average of all points (red with $1\,\sigma$ uncertainty band).

Finally, there is the issue of coincident peaks. Again, while the laser analysis does model this pathology, the result is only applicable for small window surface events. Volume events are far more likely to produce coincident peaks, but this is counter-balanced by charge attachment; the former reduces the number of observed peaks per event for small radii events, and the latter for higher radii events. Aside from the characteristics of the events being studied, this pathology is also fundamentally connected to the limits of the PF algorithm to separate peaks close in time.

To estimate the temporal resolution of the PF algorithm – i.e. the time separation at which peaks are not identified as being distinct – 2 peak laser data was used. The time separation for these events is modelled by assuming a Gaussian distribution for the arrival time of the electrons with standard deviation $\sigma_{\text{Diff.}}$ (see section 4.3.5). In this case, the distribution of time separation is a half-Gaussian beginning at $\Delta t_{\text{peak}} = 0\,\mu$s with a standard deviation of $\sigma_{\text{Diff.}}/\sqrt{2}$. To model the PF algorithms' loss of resolution at small time separations, this distribution is then multiplied by an error

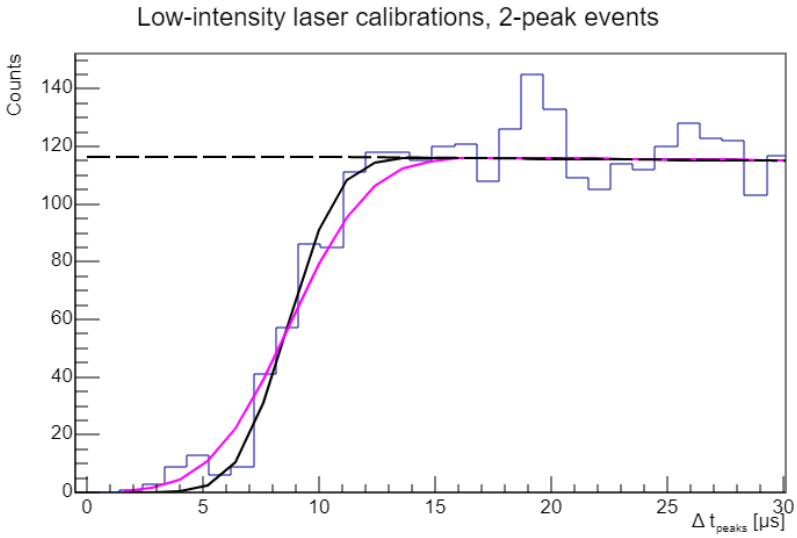


Figure 4.61: The (zoomed-in) time separation distribution for 2 peak laser events (histogram), fit with a half-Gaussian distribution at 0 (black dashed curve) multiplied by an error function to represent the temporal efficiency curve of the PF method. The black curve is an error function starting at $\Delta t_{\text{peak}} = 1\,\mu$s, and the pink curve is the model used in this analysis starting at $\Delta t_{\text{peak}} = 5.3\,\mu$s. Credit to F. Vazquez.

function defined by μ_{PF} and σ_{PF} – the PF method's temporal efficiency curve. To avoid the possibility of very close 2-peak events consisting of one electron and a small FP, this efficiency curve was restricted to be > 0 only at $\Delta t_{\text{peak}} = 5.3\,\mu$s. This model fits the time separation distribution well, as is shown in Fig. 4.61. The results for the temporal efficiency curve were $\mu_{\text{PF}} = 8.66 \pm 0.17\,\mu$s and $\sigma_{\text{PF}} = 2.41 \pm 0.26\,\mu$s, as determined by F. Vazquez.

All of these pathologies must be considered together in order to determine the PF selection efficiency for events with n peaks that actually have m electrons. This can be done analytically using the same PF formalism developed in section 4.3.1, specifically using eqs. 4.13 and 4.12 and parts of eq. 4.14:

$$P_{\text{PF}}\left(n\,\text{peaks}|r, m\,\text{e}^-\right) = \sum_{i=0}^{m}\sum_{j=0}^{m-1}\sum_{k=0}^{m-1} c_{ijk}^{nm} \times P_{\text{Poisson}}(i|P_{\text{FP}})\binom{m}{j} p_{\text{FN}}^{j}(1-p_{\text{FN}})^{(m-j)} \times \binom{m}{k} p_{\text{CE}}(r)^{k}(1-p_{\text{CE}}(r))^{(m-k)} \tag{4.26}$$

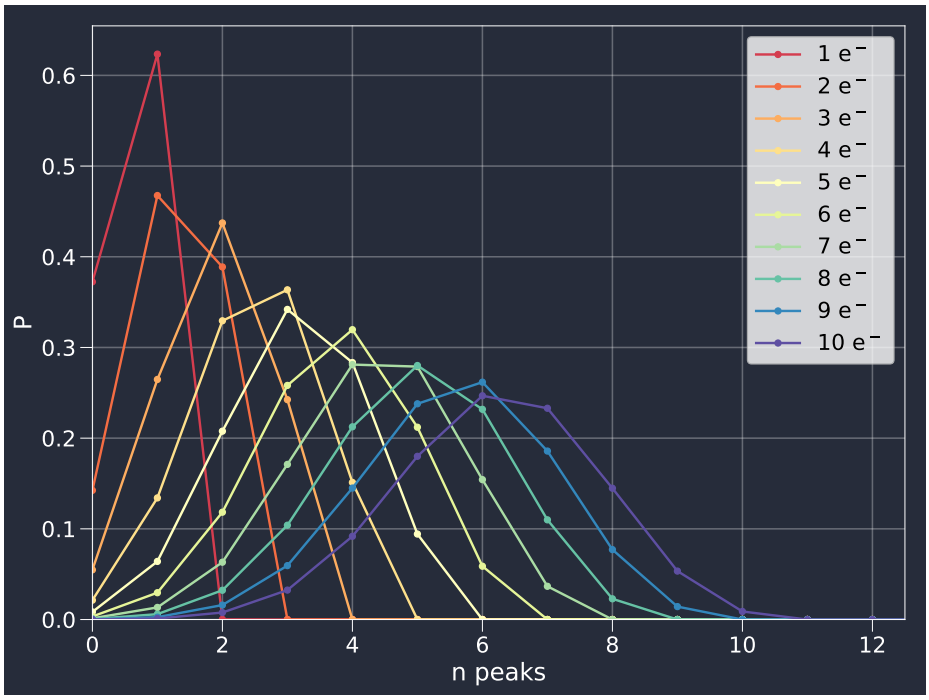


Figure 4.62: The PF efficiency for n peaks given m electrons, for surface events, using the average value of $P_{1\,\mathrm{e}^-} = 1 - p_{\mathrm{FN}}$ from all laser runs, $P_{\mathrm{FP}} = 3 \times 10^{-4}$, and the average coincidence probability of all low-intensity laser runs studied in section 4.3.1.

summing over i FPs, j FNs, and k coincidences, with the probability of the latter (p_{CE}) being a function of radial position r. For surface events, the PF efficiency for different numbers of electrons is shown in Fig. 4.62. WIMP candidate events (of a given mass) will, of course, produce a distribution of m primary electrons (denoted $P_{\mathrm{WIMP}}(m)$ here), and would be distributed throughout the detector volume. Therefore, to calculate the PF selection efficiency for n peaks from such events, eq. 4.26 must be summed for different contributions of m and integrated over r:

$$P(n|m) = \sum_m P_{\mathrm{WIMP}}(m) \int_0^{r_{\mathrm{max}}} \left(\frac{3r^2}{r_{\mathrm{max}}^3}\right) \times P_{\mathrm{PF}}\left(n\,\mathrm{peaks}|r, m\,\mathrm{e}^-\right) dr \tag{4.27}$$

However, in this case, there was no (validated) analytical expression for the probability of having coincident peaks as a function of radial position, $p_{\mathrm{CE}}(r)$. Work was done towards this goal, but was not completed. Instead, a toy MC of the PF process was designed by F. Vazquez, with inputs for the probabilities of different pathologies

described above. This was also used to simultaneously calculate the time separation distributions for surface and volume events with different numbers of peaks. As will be presented in section 4.5, time separation was the main variable used in the final dark matter analysis, the modelling of which is connected to the PF algorithms' performance. The MC proceeded according to the following (simplified) steps:

1. m primary electrons are generated for an event, and their arrival times at the sensor are randomly drawn according to the diffusion model described in section 4.3.5, accordingly for surface or volume events.

2. The number of electrons lost to attachment is drawn randomly according to the parameterization described in section 4.3.3 (using the approximate risetime of the un-affected MC event).

3. The probability of each peak alone being identified by the PF algorithm is assigned (using the results of Fig. 4.60), and all the surviving peaks are determined.

4. Starting from the first peak identified in the MC event, the probability of this peak being coincident with the next peak (chronologically) is randomly determined based on their time separation, and the temporal efficiency curve defined above. If the peaks are coincident, they are re-assigned as a single feature, and the above steps are repeated recursively until all features identified by the PF algorithm are found.

5. The time separation between the first and last peak, Δt_{peak} is calculated, and the number of identified peaks/features n is tracked, as the final results will be partitioned by n, not m.

In practice, all possible first/last peak combinations are considered for each MC event, with their corresponding probability calculated. In this way, each event contributes

multiple outcomes to the construction of the Δt_{peak} probability distribution for events with n peaks. The simulation was carried out for events with up to 15 electrons.

Comparison with data To validate the MC approach described above, the modelled time separation distributions for surface and volume events were compared to low-intensity laser and low energy ^{37}Ar data, respectively. For the former, the run tj04s000 was used. The MC time separation distributions for different numbers of electrons – *not peaks* – were added together, weighted by the fractions of events with m peaks. This quantity was determined for the laser run using eq. 4.26 and the PF characterization results from section 4.3.1. The summed MC contributions for 2 – 5 peaks are then compared to the data in Fig. 4.63 with no additional scaling applied, showing good agreement. Although this comparison relied on the laser PF characterization to achieve this agreement with the laser data, it confirms that the toy MC for PF efficiency and Δt_{peak} agrees with the empirically-driven PF results.

A similar comparison was then performed between the MC time separation distributions for volume events and low energy ($\lesssim 300\,\text{eV}$) ^{37}Ar data from tj13s000, which is shown in Fig. 4.64. In this case, there was not good agreement between the scaling of the different time separation distributions and the data, although the shapes of the distribution still matched the data well.

One possible cause of this discrepancy that was investigated was the contribution of background events in tj13s000, which certainly includes some physical events with electrons. The background model developed for the ^{37}Ar analysis was used for this study, the amplitude spectrum of which is shown in Fig. 4.65. Compared to the single electron amplitude spectrum also shown, one can see that this background data must include other non-physical events with a narrower amplitude distribution. To determine the maximal electron-only contribution of this data, the spectrum was fit assuming free contributions of different numbers of electrons (p_1, p_2, etc.), using energy response parameters determined in the laser and ^{37}Ar analyses (see sections

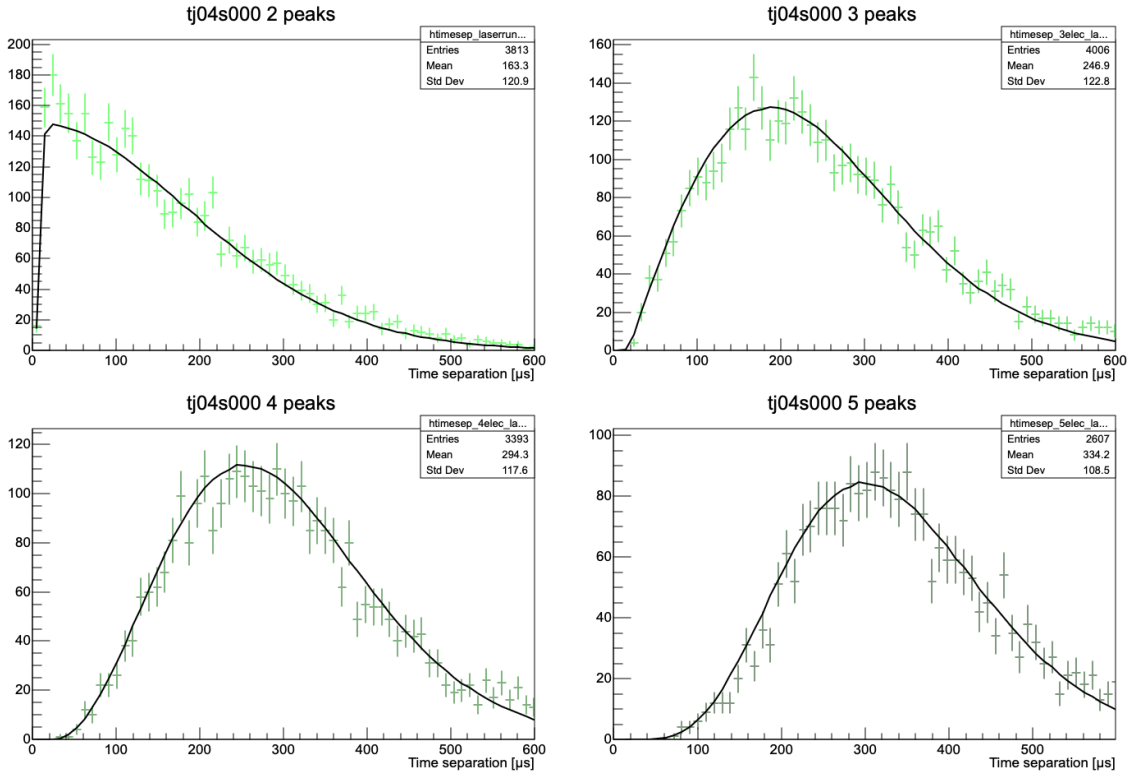


Figure 4.63: Comparison between the time separation spectra of laser run tj04s000 with different numbers of peaks (green histograms), and the toy MC simulated time separation distributions (black curves), summed/weighted using the PF characterization of the laser run from section 4.3.1. Credit to F. Vazquez.

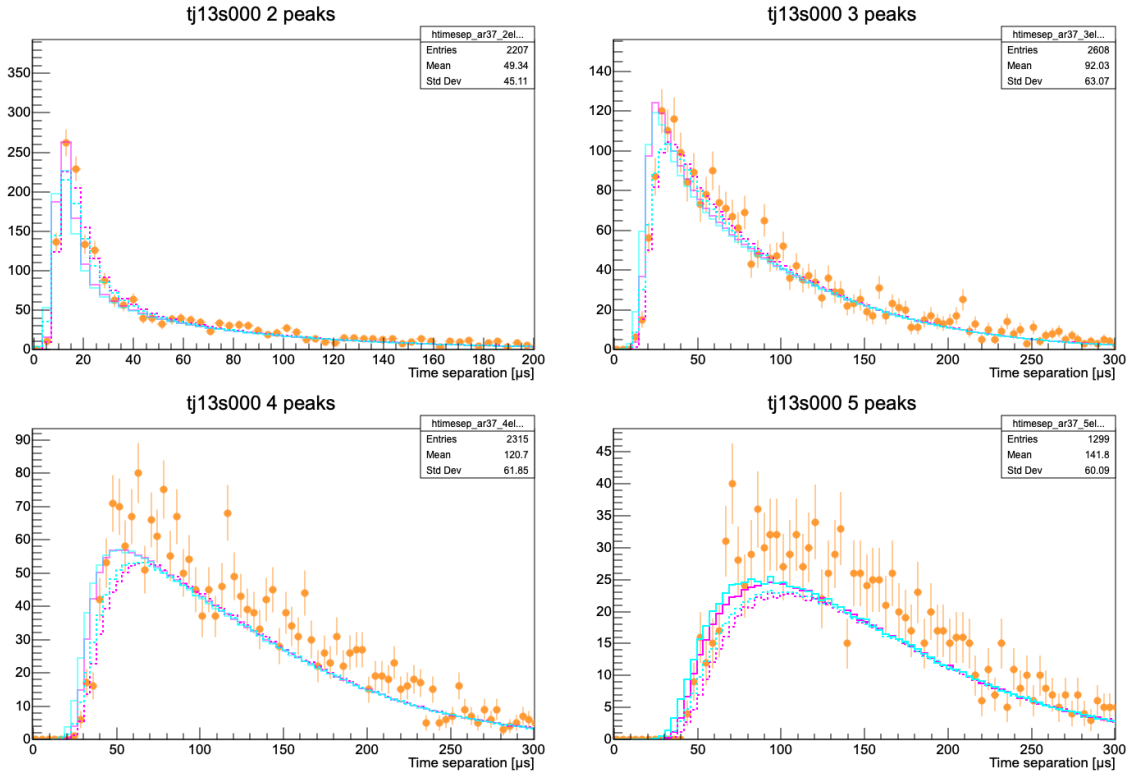


Figure 4.64: Comparison between the time separation spectra of low energy ^{37}Ar data (from run tj13s000) with different numbers of peaks (orange histograms), and the toy MC simulated time separation distributions scaled by 1.35, calculated with temporal resolution model constrained at $2.41 \pm 0.26\,\mu$s and $3.78 \pm 0.24\,\mu$s (pink and cyan curves respectively), and with/without a multi-coincident peak penalty factor (solid and dashed curves respectively). Credit to F. Vazquez.

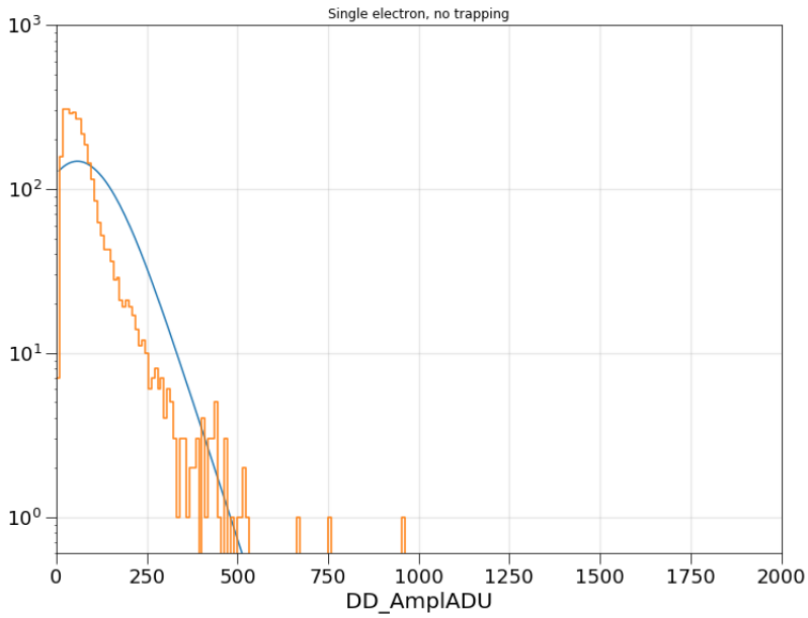


Figure 4.65: Low amplitude data of tj04s002 (with 1+ peaks and all other cuts applied, orange histogram), compared to the single electron amplitude spectrum using best-fit energy response parameters from sections 4.3.1 and 4.3.3 (blue curve). The scaling of the model is arbitrary.

4.3.1 and 4.3.3). The only free parameters are the contributions of different numbers of electrons P_n. The amplitude spectrum model $\mathcal{P}(E)$ is then calculated as (similar to eq. 4.15, summing over 8 risetime bins RT from 10 to 250 μs):

$$\begin{aligned}\mathcal{P}\left(E, RT\right) = \Bigg[\sum_n \left(\sum_{k=0}^{k=n} P_n P_{\text{Binom.}}\left(k|n, p = T_m \cdot RT\right)\right) \\ \times P^{(\text{k})}_{\text{Polya}}(E|\left\langle G\right\rangle, \theta)\Bigg] \otimes P_{\text{Noise}}\left(E\right).\end{aligned} \tag{4.28}$$

To account for the fact that only part of the amplitude spectrum shown in Fig. 4.65 can be modelled with physical event contributions, the data was fit using a modified binned likelihood function:

$$\mathcal{L} = \begin{cases} -\infty, & o_i < \nu_i - 3\sqrt{\nu_i} \\ \mathcal{L}_{\text{Poisson}}(\nu_i + 3\sqrt{\nu_i}, o_i), & o_i > \nu_i + 3\sqrt{\nu_i} \\ \mathcal{L}_{\text{Poisson}}(\nu_i, o_i), & \text{otherwise} \end{cases} \tag{4.29}$$

where o_i is the observed counts and ν_i is the expected counts given the model in

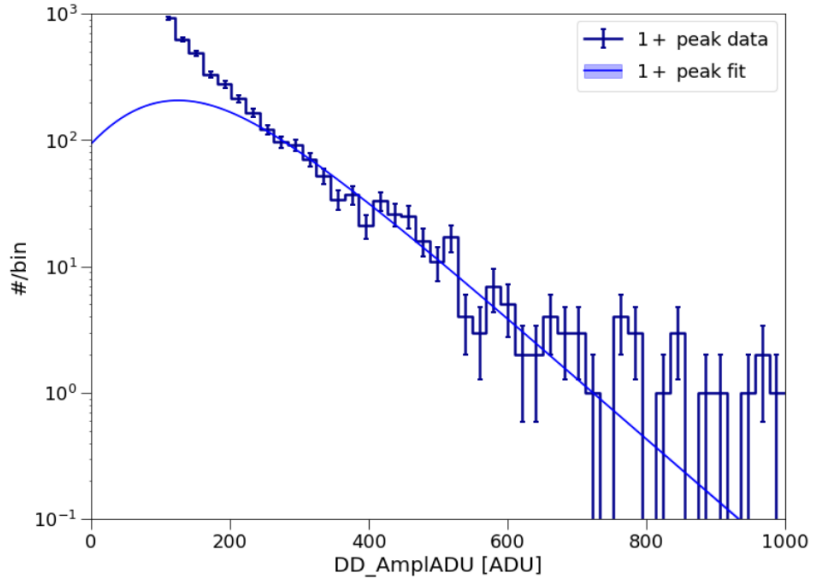


Figure 4.66: Low amplitude data of tj04s002 (with 1+ peaks and all other cuts applied, blue histogram), and physics model fit (blue curve with $1\,\sigma$ uncertainty band) obtained using eq. 4.29.

bin i. In bins where the observed number of counts was 3 standard deviations lower than the expected number, the likelihood function is assigned the value $-\infty$ to reject models that significantly overestimate the data in any bin. On the other hand, in bins where the observed counts are at least 3 standard deviations above the model, the likelihood was evaluated as though it were only 3 standard deviations low, to not over-penalize the model in bins where it cannot match the observed counts. In this way, the physics amplitude model is fit to match *but not exceed* the data in all bins possible.

The optimization was done with an MCMC, in this case [216], and the resulting fit is shown in Fig. 4.66. The relative proportions of different numbers of electrons (truncated at 4 for practical purposes) were found to be $P_1 = 0.995$, $P_2 = 0.002$, and $P_3 = P_4 = 0.0015$. The shape of the likelihood function is highly non-Gaussian, so the MCMC samples themselves were used for error propagation whenever needed. The scaled background result is shown alongside the distribution of primary electrons for tj13s000 (calculated using the results of section 4.3.3) in Fig. 4.67. Since

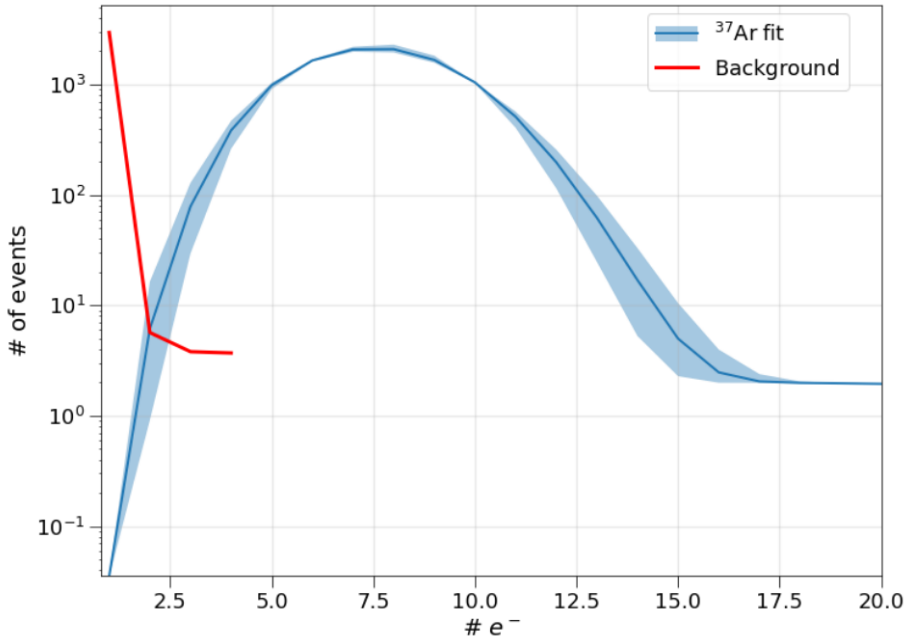


Figure 4.67: Primary electron spectrum of low-energy ^{37}Ar data, using the fit results for tj13s000 from section 4.3.3 (blue curve with $1\,\sigma$ uncertainty band) and scaled background derived with tj04s002 data (red curve).

the background overwhelmingly consists of single electron events (which in turn will produce overwhelmingly single peak events), its effect on the time separation data was found to be negligible. Ultimately, to match the ^{37}Ar time separation data and toy MC shown in Fig. 4.64, an arbitrary scaling factor of 1.35 was applied to the simulated spectra (based on the 2-peak data). Fig. 4.64 also depicts time separation results calculated using different temporal resolution definitions. Additionally, a 50% penalty factor was applied in the simulation on the probability of having events with multi-coincident electron features (cases with more than 2 electrons in coincidence). Despite this unresolved discrepancy in the scaling of the time separation MC for ^{37}Ar, the good agreement in the shape of the spectra validates the simulation insofar as it pertains to properly calculating the time separation distributions of volume events.

4.4.3 Fiducial volume

As stated in section 4.2.1, only south-channel events are kept in this analysis. This is because the electric field is known to be more isotropic in this region of the SPC, further from the grounded sensor support rod. Additionally, only the south channel of the ACHINOS could be characterized with the UV laser calibrations. As such, it is crucial to know what fraction of events occurring uniformly throughout the volume of the SPC (as dark matter events would) reach the south channel of the ACHINOS sensor vs. the north. Some fraction of events will also produce electrons that are shared between the two hemispheres. North-only and shared-hemisphere events are removed with a cut by removing events with any peak in the north channel (see section 4.4.6); this section is concerned with the signal acceptance of this data selection. In other words, this is the calculation of the fiducial volume of the experiment.

A simulation-based approach was used for this calculation, beginning with a COMSOL finite element simulation of the electric field in S140 was used [154] (created by F. Vazquez and Y. Deng). Using this, C. Garrah performed a "traditional" electron transport simulation of the SPC similar to the general MC described in section 3.1.3 [187]. This simulation is not expected to reproduce the electron drift/diffusion characteristics of the data (see section 4.3.5); it need only simulate which hemisphere of the ACHINOS the primary electrons drift to. Drift and diffusion parameters as a function of electric field strength were calculated for 125 mbar of methane gas with 0 ppm of O_2 contamination using the Magboltz software [186]. Note that these values for the gas pressure and oxygen level were the extremes of the uncertainty ranges for these parameters, conservatively chosen as they resulted in the lowest fiducial volume in this study [187].

Due to the stochastic nature of the electron transport, the south-channel selection efficiency for uniformly distributed events (or south-channel fiducial volume) varies significantly for events with different numbers of primary electrons. Therefore,

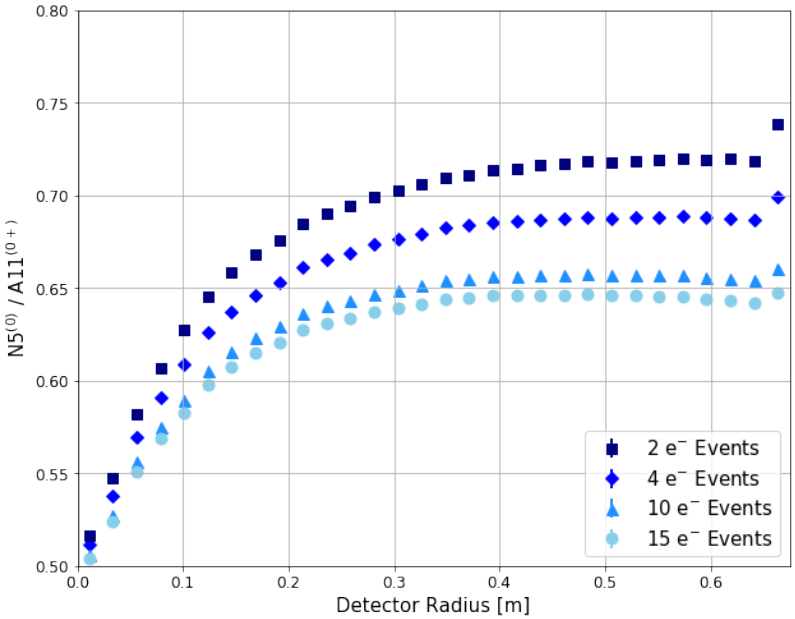


Figure 4.68: The south-channel fiducialization efficiency of events with different numbers of primary electrons, as a function of initial radial position. Taken from figure 5.13 of ref. [187].

thousands of events were simulated uniformly throughout the detector volume, with between 2 and 15 primary electrons. The fiducial volume was calculated as the ratio of the number of events in which zero electrons reach the north hemisphere (denoted $N5^{(0)}$) to the number in which at least one electron reaches any anode ($A11^{(0+)}$). This result, calculated for different numbers of primary electrons and as a function of initial radial position – is shown in Fig. 4.68 (for select numbers of electrons). This figure shows that even though events occurring at only 6 out of 11 anodes are kept in the LSM analysis, this channel subtends approximately 70% of the volume of the sphere (for events with a few primary electrons). This is due to the asymmetric nature of the electric field structure in S140, which is shown in Fig. 4.69, and depicts the electric field lines ending on north or south hemisphere anodes. Using this simulation result, the effective fiducial target mass of the experiment can be calculated by integrating the trend shown in Fig. 4.68 over radius, for a given number of primary electrons.

To validate this simulation approach, the LSM ^{37}Ar calibration data (tj13s000)

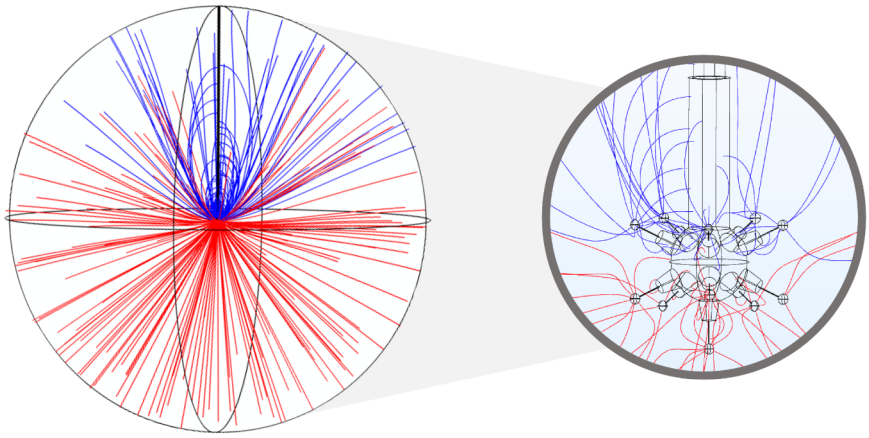


Figure 4.69: Depiction of a number of electric field lines in the COMSOL simulation of the S140 detector, coloured by which hemisphere of the sensor they terminate on (red for south-channel, blue for north-channel). The right inset shows a zoomed-in view around the sensor. Taken from figure 4.1c of ref. [187].

was used. For this comparison, the above simulation was repeated, but for uniformly distributed 2.8 keV events. This was done assuming the results for $\langle G \rangle$ of the six south anodes, θ, $W(E)$ and F given in sections 4.3.1 and 4.3.3. This simulation was repeated using drift/diffusion parameters for varying oxygen concentrations and total gas pressures according to the uncertainty on both of these quantities for the experiment [187].

To isolate events in the data with electrons reaching the different (or both) hemispheres, the fact that only the south channel triggered the DAQ presented an obstacle. Therefore the only north-channel ^{37}Ar events recorded were coincidental. To identify this population, G. Savvidis used a multi-pulse treatment algorithm, initially developed for NEWS-G Kaluza Klein axion searches [49], to find coincident north-channel ^{37}Ar pulses in post-trigger regions of each 8 ms event window [252]. The following "asymmetry" parameter was defined based on the double-deconvolved, integrated amplitudes of the separate south and north channel pulses identified:

$$\text{Asymmetry} = \frac{A_{\text{south}} - A_{\text{north}}}{A_{\text{south}} + A_{\text{north}}} \tag{4.30}$$

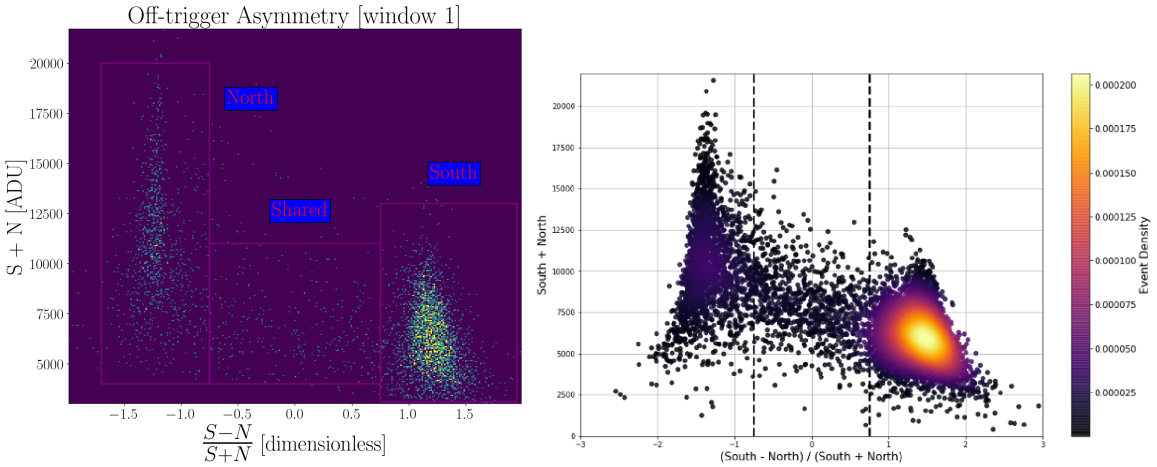


Figure 4.70: Left: the fiducialization asymmetry plot of 2.8 keV ^{37}Ar events in tj13s000, with the north channel pulses being identified in a post-trigger time window. Different event populations (purely north, south, or shared) are highlighted. Right: the corresponding plot for simulated ^{37}Ar events with cross-talk and baseline noise corrections applied. Cuts meant to mirror the data cuts are indicated by dashed lines. The colour scale is a 2D KDE score indicating the general density of data points. Taken from figure 6.8 of ref. [252] (left) and figure 5.7 of ref. [187] (right).

This variable – plotted in Fig. 4.70 (left) – was used to identify populations of events with purely south-channel, north-channel, and shared ^{37}Ar events [252]. Statistical uncertainties on these event population fractions were calculated assuming binomial statistics (for each population out of the total dataset). The final results were averaged using north-channel pulses identified with two different post-trigger time windows [252].

The ^{37}Ar simulation of C. Garrah was then extended to produce the same asymmetry plot, applying an artificial cross-talk correction (to include the small, inverted signals produced on the opposite channel as the ^{37}Ar signals), as well as convolution with a baseline noise distribution. An additional *ad hoc* scaling factor was needed in some cases to match the average amplitudes of different simulated event populations to the data. The resulting simulated asymmetry plot is shown in Fig. 4.70 (right). From this, analogous cuts were made to select populations of north, south, and shared events in the same manner as the analysis of tj13s000. The final comparison of the simulation and data results for the fractional populations of different event types is

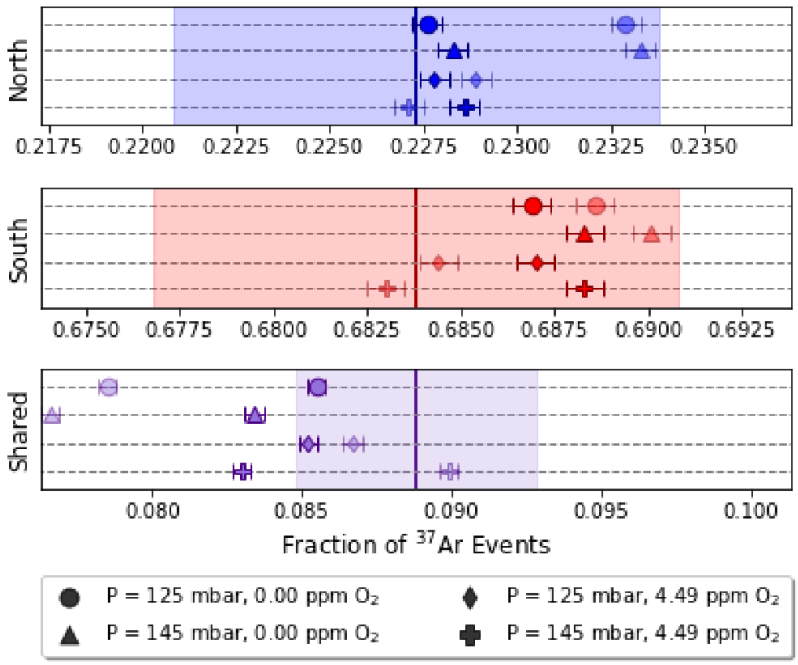


Figure 4.71: The north, south, and shared fraction of simulated 2.8 keV ^{37}Ar events, for different pressures and oxygen concentrations (markers with error bars), compared with the analogous fractions estimated from the data by G. Savvidis (shaded regions). Some simulation results had an additional *ad hoc* scaling applied to match average amplitudes with the data (faded points) [187]. Taken from figure 5.11 of ref. [187].

shown graphically in Fig. 4.71. This favourable comparison validates the simulation approach taken to calculate the south-channel fiducialization acceptance for the LSM analysis.

4.4.4 α-induced events

As mentioned in section 4.2.5, high energy events in the SPC caused significant space-charge disturbances in the sphere. The main source of such events in the LSM campaign was ^{210}Po α decays on the inner surface of the detector, with an activity of 28 mBq. Having a kinetic energy of 5.3 MeV [237], these particles produced $\mathcal{O}(10^7)$ avalanche ions. In addition to their prominent space charge impact, it was observed in this campaign – and in previous datasets [63] – that such high energy events can also induce a chain of delayed single electrons [275]. These occur on short timescales after their progenitor event (as can be seen in Fig. 4.72), but also carry on for

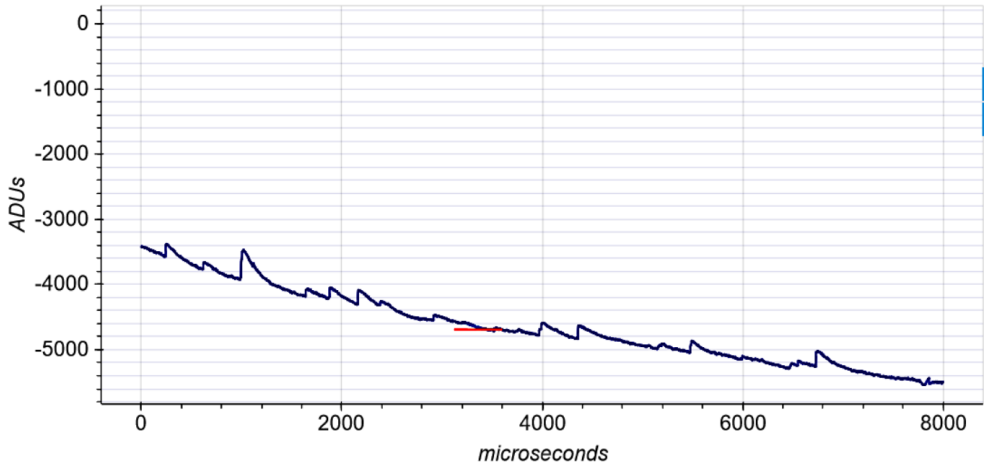


Figure 4.72: A raw pulse from run tj04s002 that occurred immediately after an α decay. In this case, the event captured a chain of single electron events on top of the falling exponential tail of the progenitor α event.

seconds afterwards. It is these long-delayed single electron events that are of great concern for this dark matter analysis. This phenomenon may be the (or a partial) explanation for the excess of low-energy events seen by NEWS-G experiments, as discussed in section 2.4.1. Currently, it is hypothesized within the collaboration that these induced-electrons are caused by the recombination of avalanche ions, and subsequent re-ionization of the gas. The purity of the gas (i.e. the levels of electronegative contaminants) also seems to play a role; electronegative ions may be the sources of recombination themselves, or possibly boost the ionization signal from other recombination events.

To characterize and remove these single electron background events, it is first necessary to identify the progenitor decays. Unfortunately, the ^{210}Po α pulses were saturated in the LSM campaign data. However, saturated pulses in this population shared a characteristic FWHM that was roughly proportional to the energy of the event, as the size of the pulse is correlated with the time taken for the pre-amplifier response tail to return back to baseline. The rate of events in this population was found to be consistent with the rate of alpha events measured at a lower anode voltage, for which these events were not saturated. However, there was imperfect

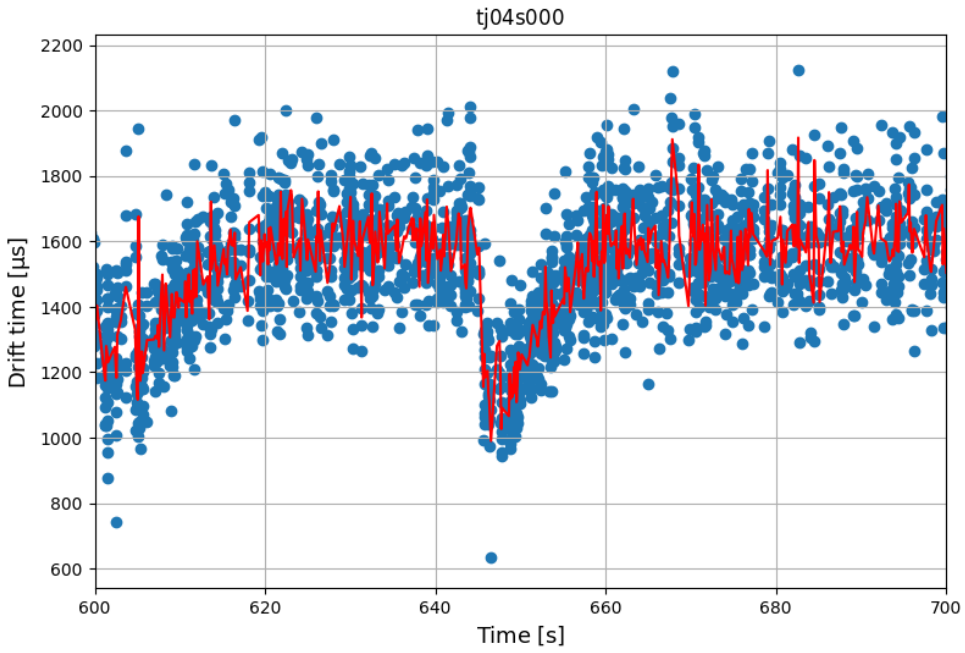


Figure 4.73: An example of the effect of an α particle on drift time: the drift time of all events in low intensity laser run tj04s000 (blue marker), with the average over time indicated in red.

separation from other saturated events not associated with induced electrons (such as some electronic noise pulses). Another way in which α events may be identified is through their space-charge impact, especially on the drift time of laser photoelectrons. Fig. 4.73 shows an example of the drop in drift time recorded after an α event, during a low-intensity laser run. Examples such as this also demonstrate the time-scale of the influence of α events. The average time constant for drift time disturbances was found to be 6.63 ± 0.10 s [275]. This was dictated by the time taken by the avalanche ions to slowly drift back to the surface of the SPC.

The final marker of this problem was the rate of single electron events (and the rate of all events) over time; this would temporarily (and dramatically) increase after α events, as can be seen in Fig. 4.74. The average time constant for the production of single electrons was measured to be 1.73 ± 0.02s [275]. Using a combination of these different techniques to identify α events, cuts were calculated for every LSM campaign run to remove a time period of 5s after each identified α event. Note that in the case

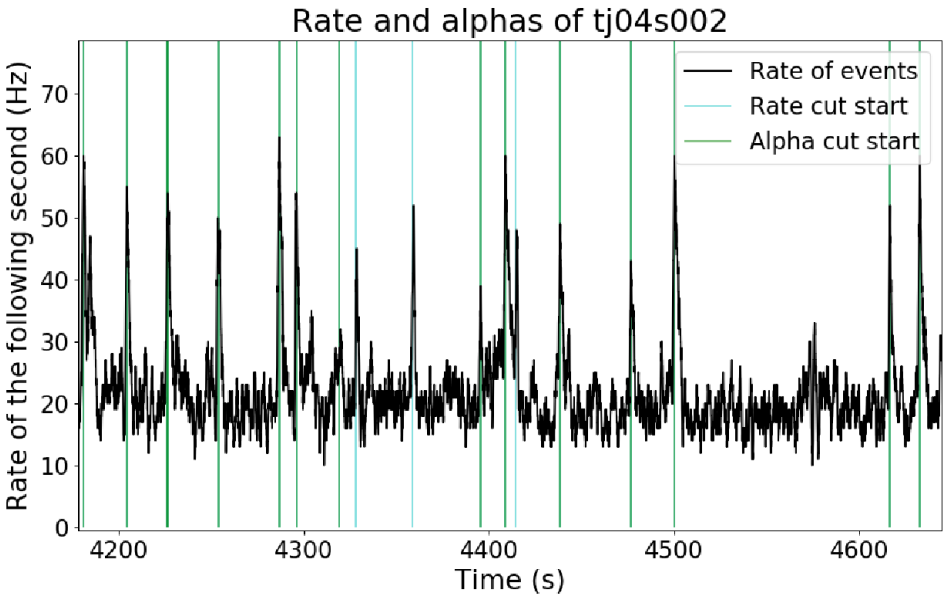


Figure 4.74: Examples of the rate increases caused by α events in run tj04s002. The rate-based identifications of α events are marked with cyan lines, and the independent FWHM identifications with green lines. Taken from figure 5.13 of ref. [275].

of photodiode-triggered laser calibration runs, only the drift time information could be used to identify α events, since the saturated α pulses and induced electrons were not recorded. This analysis was performed by J.-M. Coquillat.

The signal acceptance (for dark matter events of these cuts) is simply the remaining livetime after removing post-α periods, since these are uncorrelated with dark matter candidate events. For example, after applying this cut, 11.8% (respectively 11.1%) of the livetime of the physics run tj04s002 (respectively the ^{37}Ar run tj13s000) was removed. More details on these cuts and their signal acceptances can be found in ref. [275]. However, in the case of tj04s002, this cut removed 95.17% of single electron events passing other cuts, demonstrating that the vast majority of the single electron signal in this experiment was due to α decays. It is partially for this reason that – despite the efficacy of the cuts described here – single peak data is omitted from the dark matter analysis, due to the probably still-persistent contamination from α-induced single electrons (see section 4.5).

4.4.5 PSD cuts

A major source of background events in the LSM campaign data was non-physical "spike" events. As mentioned in section 4.2.4, these are thought to be caused by internal discharges in the sensor. Fortunately, these events demonstrate several characteristic traits that may be used to discriminate against them. One such property is the signal cross-talk that is experienced by physical events, but not by spike events. To exploit this phenomenon, the variable "N/S amplitude" (north/south amplitude) was defined as the ratio of the integral of the double-deconvolved pulse of the north channel vs. south channel. This variable was calculated on a peak-by-peak basis (rather than for the whole event), wherever positive south peaks are identified by the PF algorithm. Peaks closer than $\sim 13.4\,\mu$s in time are grouped together. A visual example of this variable can be seen in Fig. 4.22. The strength of the cross talk signal (where present) was typically around 20% relative to the progenitor signal, so N/S typically had values around -0.2 for physical, south channel events, and values centred around 0 for non-physical events.

To study the efficacy of this PSD variable and to construct a cut to remove spike events, it is necessary to have samples of both physical and spike events. The former is provided by the low-intensity laser calibration data (specifically tj04s000); the photodiode-triggered data will contain only physical events barring any coincidences with spikes, the rate of which is assumed to be negligible. A sample of spike events is provided using some of the test physics data (tj04s002), which is itself dominated by non-physical events after removing laser pulses and post-α time periods (see section 4.4.4). Further cuts are made based on the classical PSD variables used by NEWS-G; the FWHM and risetime of the raw pulse (`DD_RawWidth` and `DD_RawRise` respectively). The spike population selection in this parameter space is shown in Fig. 4.75. Although it is hard to quantify the efficiency of this PSD cut – and it must be acknowledged that using one imperfect PSD cut to define another is somewhat circular logic – since

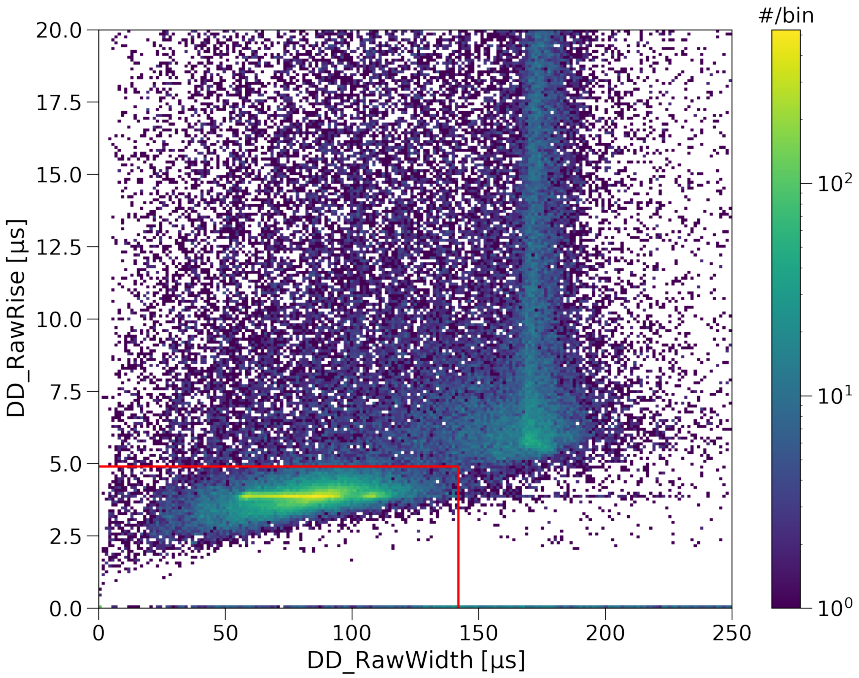


Figure 4.75: 2D histogram of `DD_RawRise` vs. `DD_RawWidth` for physics run tj04s002 (the colour scale gives counts per bin), with cuts to select spike events shown in red (selecting the lower left quadrant).

no background subtraction of spike events is performed, only the signal acceptance of the cut (based on the laser data) will matter for the LSM analysis. Imperfect selection of a spike population on which to base the subsequent PSD cuts will result in a less-than-optimal cut, but one whose signal acceptance for dark matter is properly calculated. The N/S distributions of the laser and spike event populations are shown as functions of the height of the peaks involved in Fig. 4.76, which demonstrates the separation between the two types of events. As expected, this discrimination breaks down for small peaks, for whom the cross-talk signal (if present) is lost in the noise of the double deconvolved pulses.

An additional PSD variable defined for this analysis was dubbed "spikiness", intended to be an improvement on the older raw risetime and FWHM PSD variables. Specifically, this pertains to the characteristically sharp leading edges of non-physical events; an example of a spike event in which this feature is evident is shown in Fig. 4.77. Spikiness is defined by taking the maximum 2-sample derivative of the raw

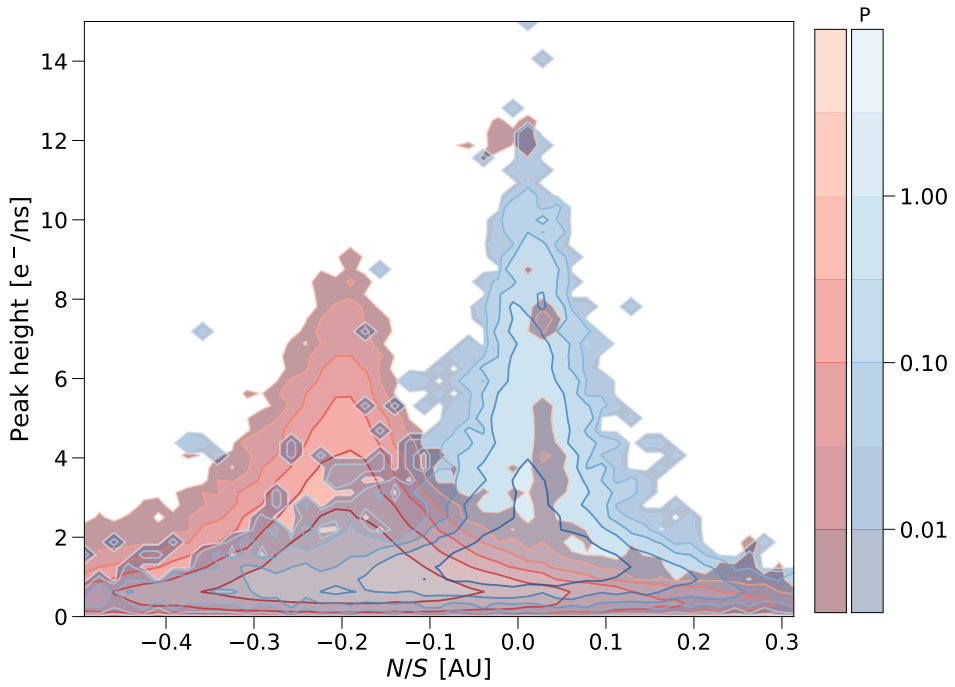


Figure 4.76: 2D distributions of peak height vs. N/S for the laser and spike event populations (red and blue respectively). The colour-scale and contours represent the density of the distributions.

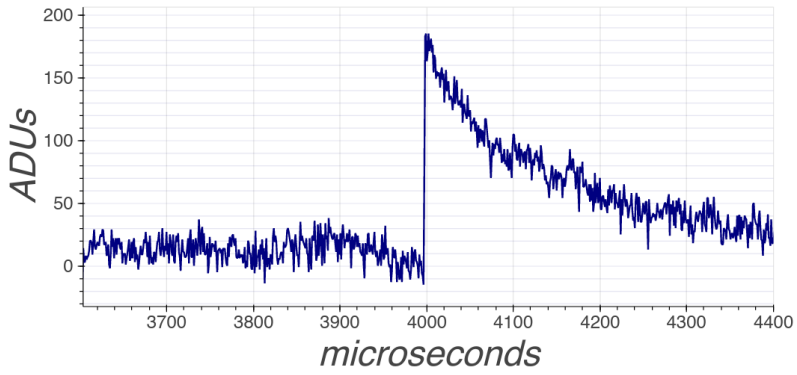


Figure 4.77: An example of a "spike" noise event in the physics run tj04s002 (raw pulse), in which the trace rises entirely in one time bin.

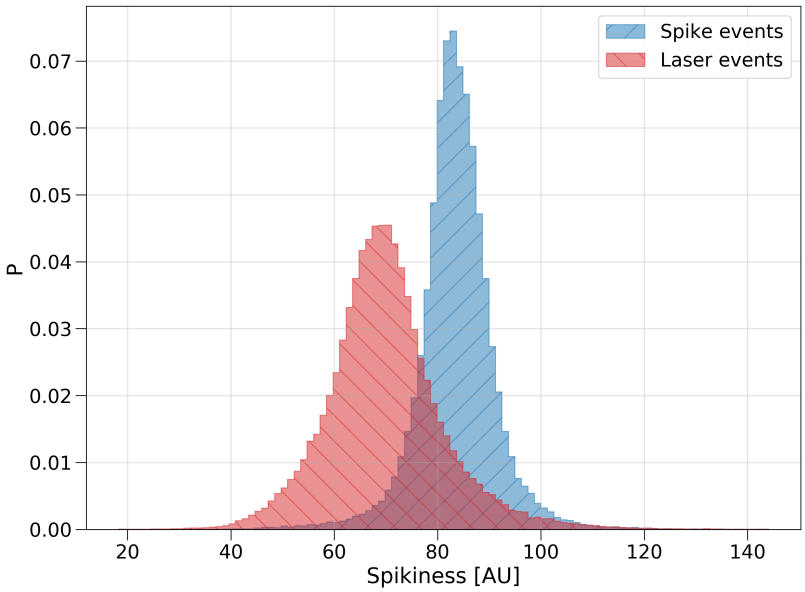


Figure 4.78: Histograms of spikiness for laser (red) and spike event populations (blue).

event trace within 10 μs of each peak identified by the PF algorithm, divided by the corresponding peak height. For spike events, this variable has a relatively high value, as these events rise more abruptly than physical events. The distributions of spikiness for laser and non-physical events are shown in Fig. 4.78.

While these two PSD variables would separately help remove spike events from the LSM physics datasets, the correlated nature of these two quantities means that a 2D cut in both would be even more effective. The 2D distributions of spikiness and N/S are shown in Fig. 4.79, from which it is evident that a diagonal cut would serve well to discriminate against spike events. The method used to determine the multivariate cut was to calculate the linear Fisher discriminant between the two variables. The Fisher discriminant determines a coordinate rotation for which the separation between the two populations is greatest when projected onto this rotated axis [285, 286].

To do this, the discriminant τ is defined as a linear combination of the variables in question $\vec{x}$ (this applies to more than two variables):

$$\tau = \vec{\alpha} \cdot \vec{x} \tag{4.31}$$

Next, the data samples of the two populations A and B are used to determine the mean parameter values $\vec{\mu}$ and covariance matrices V of each sample. The coefficients $\vec{\alpha}$ that lead to the best separation between the two populations (based on their relative means and covariances) are given by [275, 285]:

$$\vec{\alpha} \propto (\mathbf{V}_A + \mathbf{V}_B)^{-1} \cdot (\vec{\mu}_A - \vec{\mu}_B) \tag{4.32}$$

For the N/S and spikiness variables, the optimum coordinate rotation was defined by:

$$\tau = \vec{\alpha} \cdot \vec{x} = (-76.87,\ -0.6786) \cdot \begin{pmatrix} N/S \\ \text{Spikiness} \end{pmatrix} \tag{4.33}$$

The distributions of τ for the laser events and spike population events are shown in Fig. 4.80, which demonstrates the superior separation between these two populations compared to either variable alone.

Using this Fisher discriminant, J.-M. Coquillat determined that the optimum cut to separate the two populations (based on the ratio of the two distributions, weighted for the relative population sizes of the spike events and dark matter candidate events) was at $\tau = -40$ [275]. This cut is depicted in the Spikiness vs. N/S parameter space in Fig. 4.79. In addition to this cut, another was made to select events with $N/S < -0.1$ to remove another class of non-physical events. This species of background – dubbed "wide events" – had low Spikiness values, but $N/S \sim 0$, and so were easily removed with this additional cut [275]. A hint of these events can be seen in Fig. 4.79, but they were a more prominent background for 2-peak data specifically. The origin of these background events is not known.

Since the PSD variables described above were calculated on a peak-by-peak basis, in the case of events with multiple peaks, the average value of each variable per event

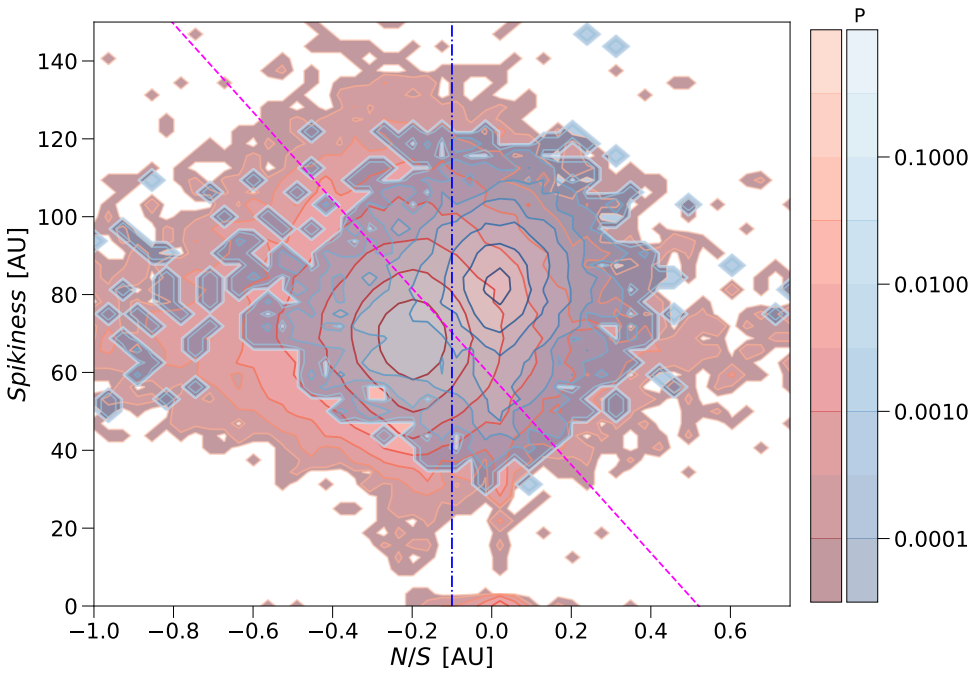


Figure 4.79: 2D distributions of spikiness vs. N/S for the laser and spike event populations (red and blue respectively). The colour-scale and contours represent the density of the distributions. The two PSD cuts defined in this analysis ($\tau > -40$ and $N/S < -0.1$) are denoted by magenta and blue dashed lines, respectively.

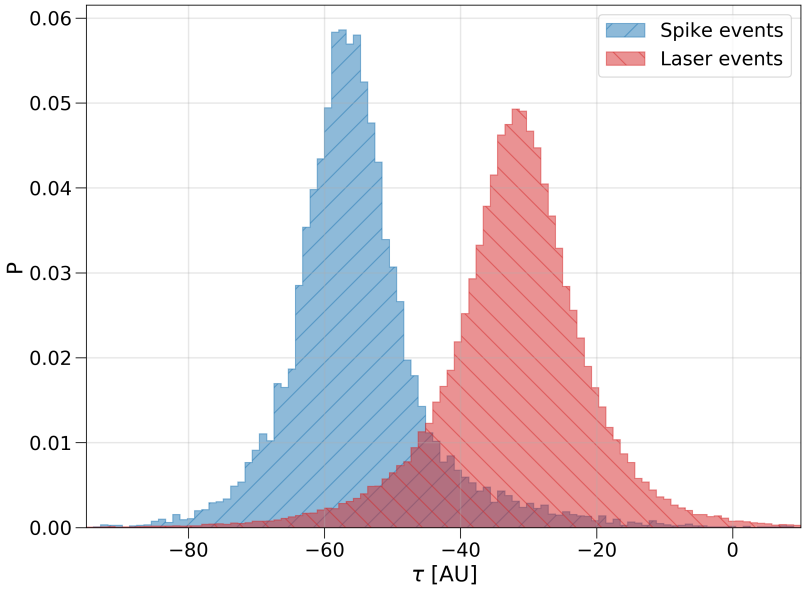


Figure 4.80: Distributions of the Fisher discriminant τ for laser (red) and spike event (blue) populations.

Table 4.7: Efficiencies of PSD cuts on N/S and τ for different numbers of peaks, with statistical uncertainties [275].

# of peaks	**Efficiency**
1	$77.22 \pm 0.11\%$
2	$83.30 \pm 0.60\%$
3	$85.35 \pm 0.56\%$
4	$86.98 \pm 0.59\%$

was used. Again, using the laser data from run tj04s000, J.-M. Coquillat estimated the signal acceptance of the cuts on τ and N/S for different numbers of peaks. These efficiencies are listed in Table 4.7. It is estimated that after applying these PSD cuts, 5% of the spike events in the populations studied would still remain in the dataset [275].

The results described above are entirely derived from 2 datasets (tj04s000 and tj04s002). Therefore, it is crucial to ensure that the behaviour of the PSD variables is consistent throughout the course of the campaign, especially for laser events. The N/S and Fisher discriminant populations were calculated for all low-intensity laser datasets (as representations of physical events), and analogous spike event populations from the unblinded physics data. These populations are plotted for all runs chronologically in Fig. 4.81. The trends are generally consistent, indicating that the PSD properties inferred from the few runs analyzed hold throughout the campaign.

4.4.6 North-channel coincidence cut

An unexpected phenomenon in the LSM data was the observation that the number of positive peaks found in the north channel traces of laser events is strongly correlated with the number of positive south channel peaks. This runs counter to the assumption that only particle interactions in the SPC volume may lead to events with shared

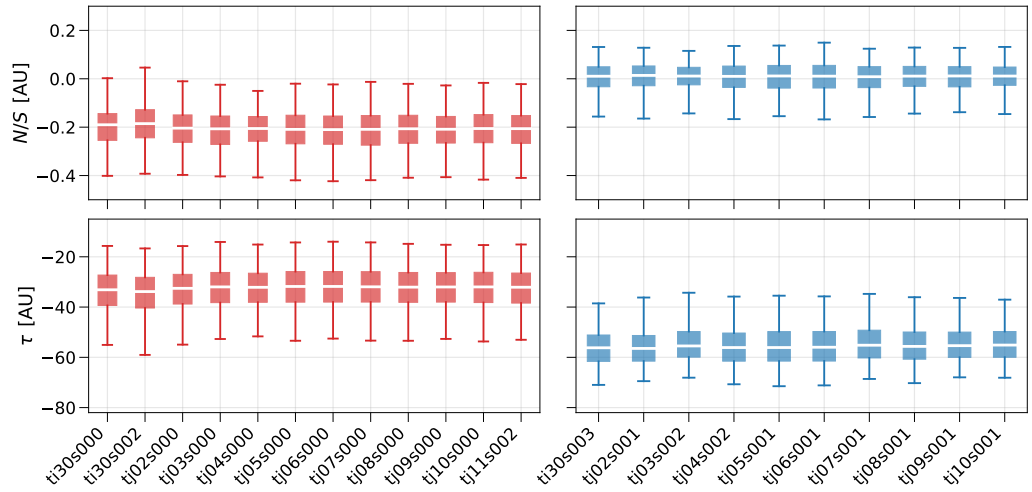


Figure 4.81: Box and whisker plots showing the trends of N/S and the Fisher variable τ (bottom) populations for all laser runs (left, red) and spike event populations from the test physics data (right, blue). The boxes represent the $25-75\%$ percentile range, the central dash gives the median of each distribution, and the whiskers are the $5-95\%$ percentile range.

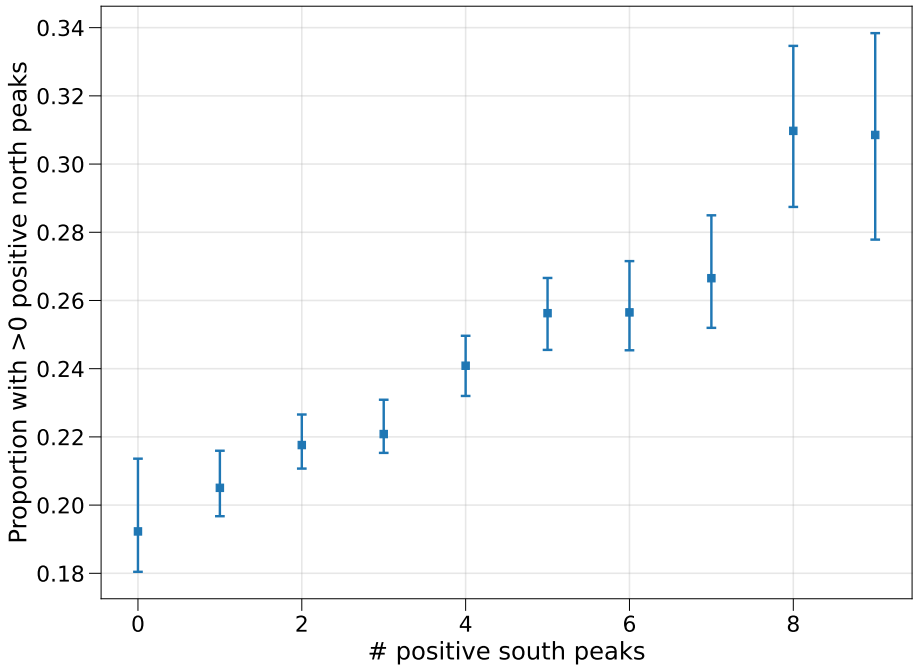


Figure 4.82: Proportion of laser events with more than zero positive north-channel peaks as a function of the number of positive south-channel peaks. This data is taken from the low-intensity laser run tj04s000, which cuts applied to remove post-α events (see section 4.4.4) saturated events.

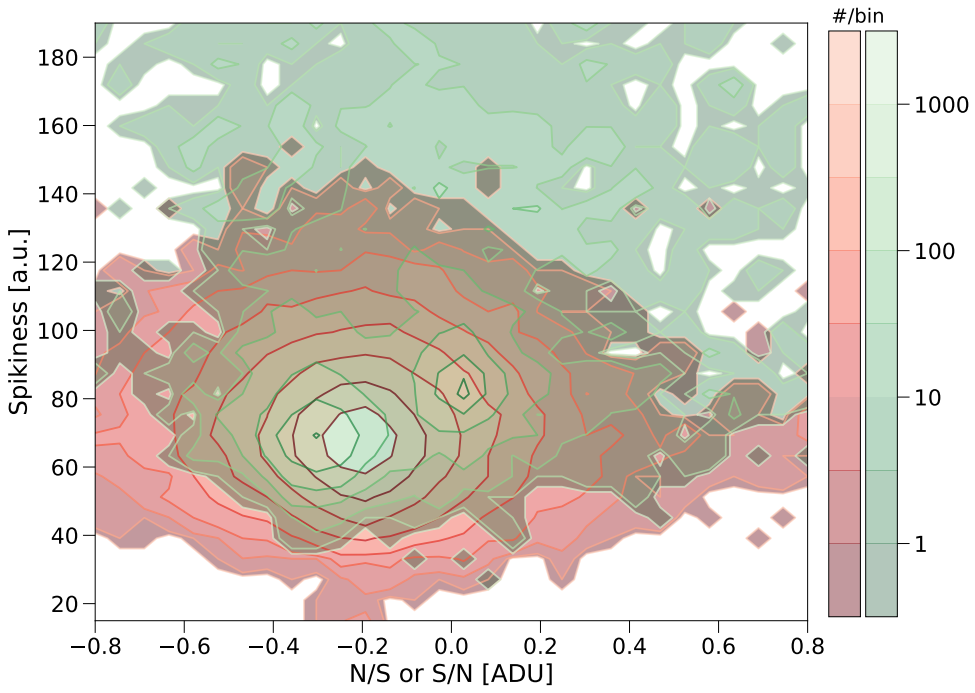


Figure 4.83: Pulse shape discrimination populations of positive north-channel peaks (shown in green in terms of S/N amplitude and spikiness, see section 4.4.5), compared to that of laser-only south channel peaks (N/S instead, red distribution), for tj04s000.

electrons (such as ^{37}Ar). This correlated behaviour is shown in Fig. 4.82. The source of these correlated north-channel peaks was thought to be (at least partially) the laser itself. It was initially assumed (naively) that the laser fibre was directed roughly due south in the sphere. However, there is no direct proof of this, nor was it even possible to observe the exact orientation of the fibre. Observed north-channel peaks also presumably include non-laser physical pulses (coincident with south-channel peaks), coincident spike events, as well as possible noise-triggers from the peak-finding algorithm applied to the north channel. Support for the north-channel events including a mix of physical and non-physical events comes from the bimodal PSD distribution of these north-channel peaks, shown in Fig. 4.83.

The impact of this phenomenon is due to the fact that a significant cut applied to the LSM data is to remove events with positive north-channel peaks, meant to reject events that may have electrons shared between the two channels. If a significant

fraction of north peaks removed are in fact coincident pulses and not truly shared events, then there is an additional loss of dark matter signal acceptance due to this cut, beyond the fiducial volume efficiency calculated in section 4.4.3. Furthermore, this cut would represent an unnecessary loss of exposure (although perhaps an unavoidable loss in the current analysis). Lastly, there was also the question of whether or not the larger-than-expected correlation in the number of peaks across both channels is due to benign causes such as coincidences and/or the laser shining partially on the north hemisphere of the SPC, or another unknown issue.

To quantify the fraction of north-channel peaks in run tj04s000 that are produced by the laser, the contribution coincident north-channel peaks – *not including laser electrons* – is estimated using pre and post-trace north channel data. The north peaks found in these data selections (with peak-finding windows starting at 1400, 1660, and 5660 samples) presumably consist of coincident physical peaks, spike events, noise peaks, and possibly α-induced electrons (that leak past the 5 s cut). This non-laser distribution is shown in Fig. 4.84, compared to the distribution of north-channel peaks in the main event window. The excess of the latter compared to the former is due to laser electrons. This contribution on top of the non-laser "background" is fit to the data assuming a linear relationship between the photodiode amplitude of each laser event and the average number of north peaks produced, the slope of which is the free parameter of the fit. This fit was done with the `scipy.optimize.curve_fit` least-squares function [238].

The resulting model shown in Fig. 4.84 – including a laser contribution – fits the data well, and indicates that 53.3% of all north-channel peaks in the selected tj04s000 data are indeed laser electrons. Conversely, 86.7% of events have no north-channel laser electrons (since most events have zero north-channel peaks). Multiplying the latter quantity by the proportion of events that have more than zero north-channel peaks but that pass all other analysis cuts gives the fraction of events removed due to north-channel peaks not produced by the laser: 20.7%. This represents the loss

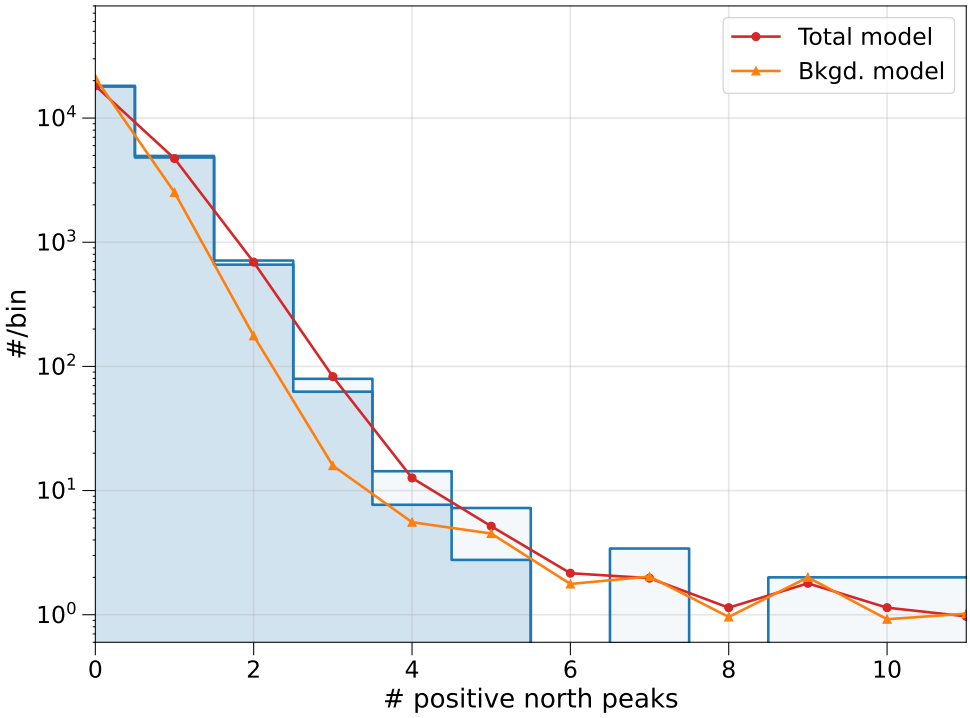


Figure 4.84: Distribution of the number of positive north-channel peaks in tj04s000 (blue histogram with the top empty bars giving the $1\,\sigma$ statistical uncertainty ranges). The non-laser background of peaks estimated using pre and post-trace data is shown in orange, and the fit to the data assuming an additional laser-induced component is shown in red.

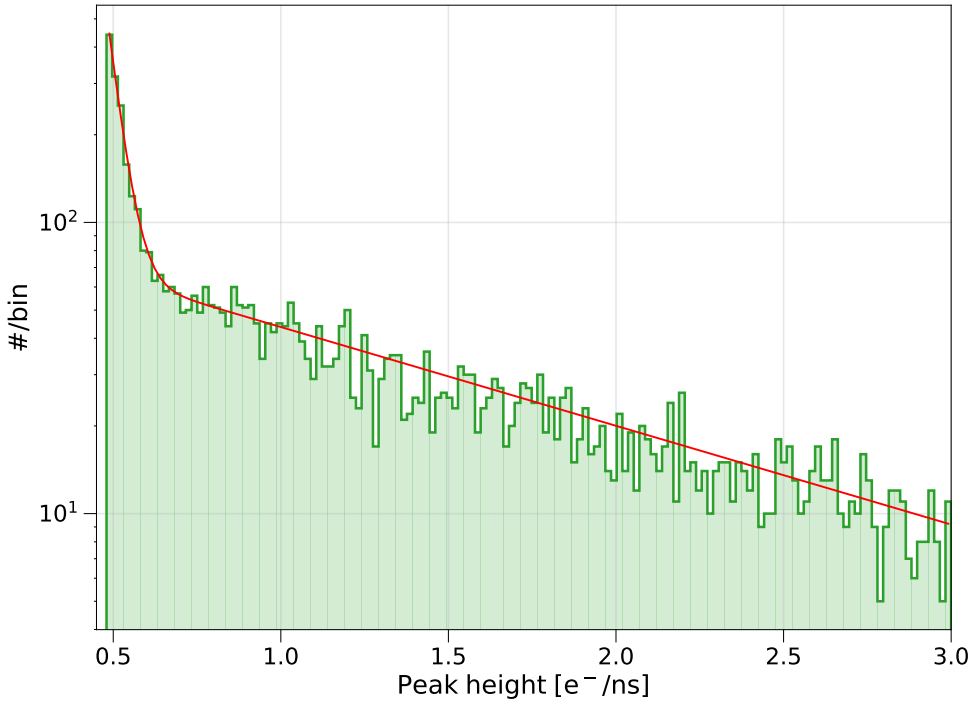


Figure 4.85: Peak height distribution of north-channel peaks in tj04s000 (green histogram), alongside a fit to the data including normally-distributed noise peaks and non-noise peaks modelled generically as an exponential distribution (red).

of exposure incurred by removing events with positive north-channel peaks in the physics data, conservatively assuming that any physical peaks in the north-channel are coincidental rather than shared-hemisphere events.

A large fraction of the events cut may have been due to noise-peaks found by the peak-finding algorithm, especially since it seems the north channel was subject to higher power, high-frequency noise. This was investigated by looking at the distribution of north-channel peak heights found by the PF algorithm, which is a combination of a Gaussian distribution of noise peaks (cut off at the peak-finding threshold of 0.48 e^-/ns), and an exponential distribution from non-noise peaks (physical, laser-induced, spikes etc.). This is shown in Fig. 4.85, the fit of which indicates that approximately 26% of all north-channel peaks may be noise peaks. Therefore, the exposure loss incurred by removing events with north-channel peaks could be improved in future analyses by applying a separate, optimized peak-finding threshold for the

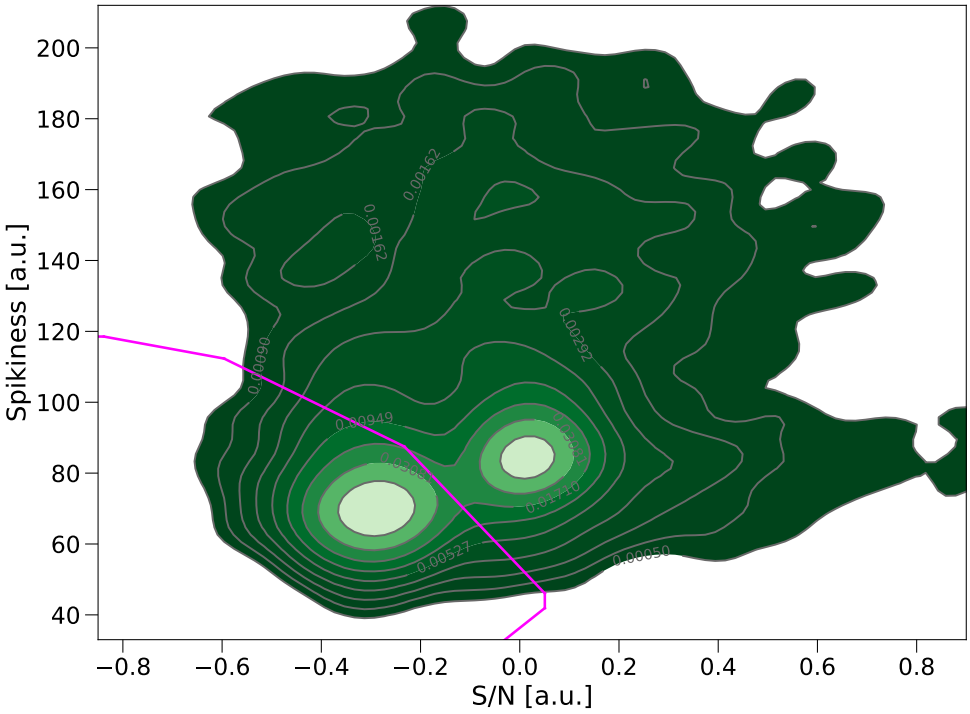


Figure 4.86: Two-dimensional kernel density estimation model of north-channel pulse shape discrimination data from tj04s000, with a visible bimodal nature. A pulse shape discrimination cut to select physical events with power $\alpha = 0.1$ is shown in magenta.

north channel.

Effect of north-channel PSD cuts In the future, it is possible that the loss of exposure due to north-channel peaks may be mitigated by applying a PSD cut to the north-channel data. This would improve the efficacy of the north-channel peak cut for removing only events with physical north-channel peaks, as noise triggers and spikes on the north channel are not necessarily cause to reject those events entirely. For the current analysis, this is not possible due to a lack of north-channel laser data with which to calculate the exact efficiency of such a cut. However, a north-channel PSD cut is considered here to inform future efforts.

To estimate the 2D PSD populations of physical and non-physical north channel peaks, a 2D KDE was calculated for the north-channel PSD data (using a Gaussian kernel [238]), which is shown in Fig. 4.86. The physical-peak population was then

approximated by taking the south-channel PSD population (again modelled with a 2D KDE) and translating it slightly to match the peak value of S/N amplitude for north-channel peaks (see Fig. 4.83), as evidently the north channel experiences slightly stronger cross-talk. These two KDEs are then used to calculate a 2D likelihood ratio for the hypotheses of physical events vs. physical plus non-physical. A cut in terms of S/N amplitude and spikiness is then defined using a contour of the likelihood ratio with a power of $\alpha = 0.1$ (shown in Fig. 4.86) and then $\alpha = 0.01$. The weaker cut is defined as `PSD cut 1`, and the stronger `PSD cut 2`.

After employing these cuts to remove physical north-channel peaks (presumably decreasing the fraction of north-channel peaks caused by laser events), the same procedure described above was used to quantify the fraction of laser-induced events remaining. With the weaker cut applied, the proportion of laser-induced north channel peaks is reduced to 32.5% (down from 53.3%), and reduced further to 27.4% by the stronger PSD cut. This proves that in the future, a north-channel PSD cut may effectively allow for the rejection of events with specifically physical north peaks.

Drift time evidence A separate line of reasoning that supports the conclusions presented in this subsection involves looking at the "drift time" of both north and south channel peaks, namely the time difference between the photodiode pulse and peaks in question, which is shown in Fig. 4.87. For non-laser events, this quantity should be uniformly distributed, since there is no temporal association with the laser pulse. This is the case for pre and post-trace drift time data for north channel peaks. Conversely, for laser events, a roughly Gaussian distribution is expected, which is true for south channel peaks. However, the north-channel peaks in the main event window demonstrated a mixture of both modalities, as shown in Fig. 4.87. This confirms that these events are a mixture of laser and non-laser events. If the PSD cuts described above are applied, the "drift time" distribution of north-channel peaks approaches a uniform distribution as expected (especially for the stronger cut `PSD`

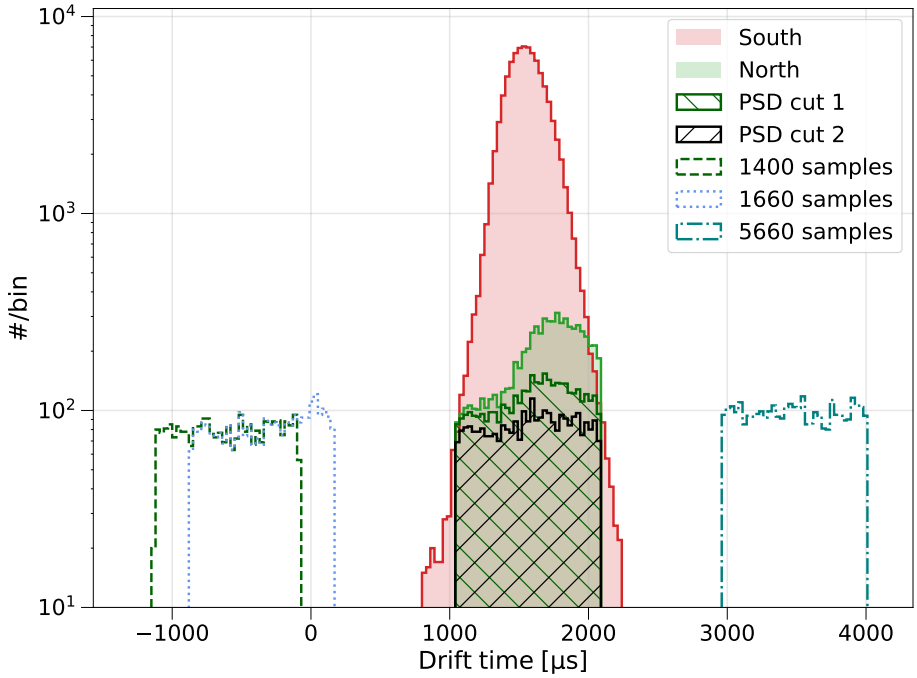


Figure 4.87: Drift time distributions of north (green) and south channel (red) peaks for tj04s000, including pre and post-trace north channel peaks. The cross-hatched histograms show the result of applying pulse shape discrimination cuts to the north-channel peaks.

`cut 2`), demonstrating the efficacy of these cuts.

4.4.7 Other data quality cuts

A significant cut made to the LSM data was to remove laser events from the physics datasets (which was done before the blind/test data partitioning). This was done using the photodiode amplitude, with a presumed 100% efficiency. Removing these events (which occurred at a rate of 5 Hz during physics runs) did lead to a reduced livetime for the physics data. Including the event windows containing laser events, and the subsequent DAQ deadtime, removing the laser events reduced the total exposure of the campaign by 6% (respectively 9%) when trigger condition 1 (respectively trigger condition 2) was employed, as defined in section 4.2.1.

Several additional data quality cuts were applied to remove some classes of pathological events. Several examples of such events are shown in Fig. 4.88. The additional

quality cuts were as follows:

- A cut to remove events with saturated pulses in the south channel. This was done based on the raw DAQ-measured amplitude, which has the same maximal value for all saturated events.

- A similar cut to events with saturated pulses in the north channel. Even though only the south channel data was used for the dark matter analysis, cross-talk from large north channel pulses sometimes caused abrupt baseline artifacts.

- Events occurring on the falling exponential tails of preceding high amplitude pulses (such as an un-triggered α event occurring immediately prior) sometimes resulted in artifacts in the baseline of the double deconvolved pulse. Therefore events with anomalous baselines (rising, falling, or significantly offset) were removed. Specifically, this applied to events where the baseline of the double-deconvolved pulse (the average value of the trace in the 20 samples preceding the integration window) differed by more than 0.20 e^-/ns between the north and south channels.

- Occasionally, primary electron peaks were observed outside the integration window defined in section 4.2.2. In these cases, the calculated value of Δt_{peak} would not represent the true diffusion of that event. To remove such events, the PF processing was extended 300 samples beyond the original integration window in both directions, and events with peaks in these windows were removed.

The cuts described above are all assumed to have no effect on the signal acceptance for dark matter candidate events beyond the associated reduction in livetime, as none of the pathologies listed are in any way correlated with specific properties of dark matter candidate events. The total loss of exposure due to the cuts listed here was 4.78%. This deadtime calculation was performed by J.-M. Coquillat.

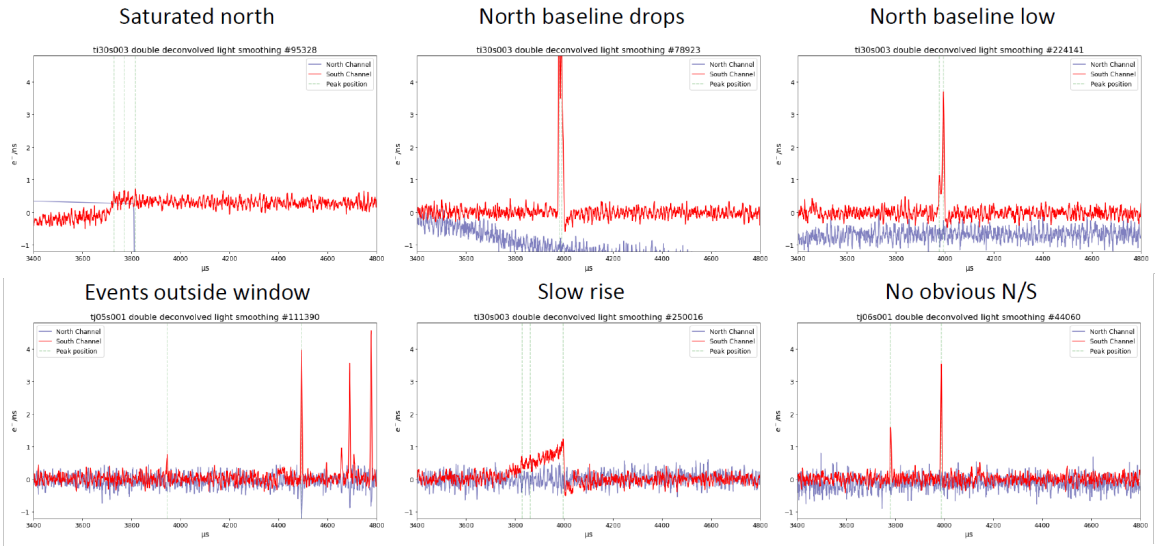


Figure 4.88: Examples of pathological events removed with specific data quality cuts listed in section 4.4.7. Credit to J.-M. Coquillat.

4.5 LSM campaign: Physics results

The test data of the dark matter search with methane gas did not reveal any excess of events believed to be WIMP-like. As was the case in the previous NEWS-G experiment at the LSM [63], the dark matter physics campaign experienced a significant rate of background events. Therefore, the final result of this analysis will be to produce an exclusion limit on the existence of WIMP dark matter, specifically for an SD-p coupling – to which the experiment is most sensitive. While this will ultimately be calculated with the blinded physics data, for the purposes of predicting the final result, the test data is studied here. The following subsection summarizes the work done by F. Vazquez to calculate an exclusion limit. Time separation is used as the main variable of interest, rather than an energy estimator as is often the case in dark matter studies. For events with few electrons, there is a great overlap in the amplitude/energy spectrum due to the physics of SPCs. Therefore, energy-related variables do not provide much discrimination power between WIMP-like and background events beyond what the PF results already provide. Time separation, on the other hand, is useful to disentangle volume, surface, and coincident-peak events, the latter two being the dominant remaining background event types of the experiment. To exploit this background discrimination capability, a profile likelihood ratio (PLR) analysis was performed for the limit calculation in order to subtract these background contributions [81].

The test data consisted of approximately 27% of the full physics data (see section 4.2.1), with 46.39 hours of livetime after all cuts. The data selection cuts applied were:

- A cut to remove post-α event time periods (see section 4.4.4),
- Cuts on the linear fisher discriminant and N/S PSD variables to remove spike events and "wide pulses" (see section 4.4.5),

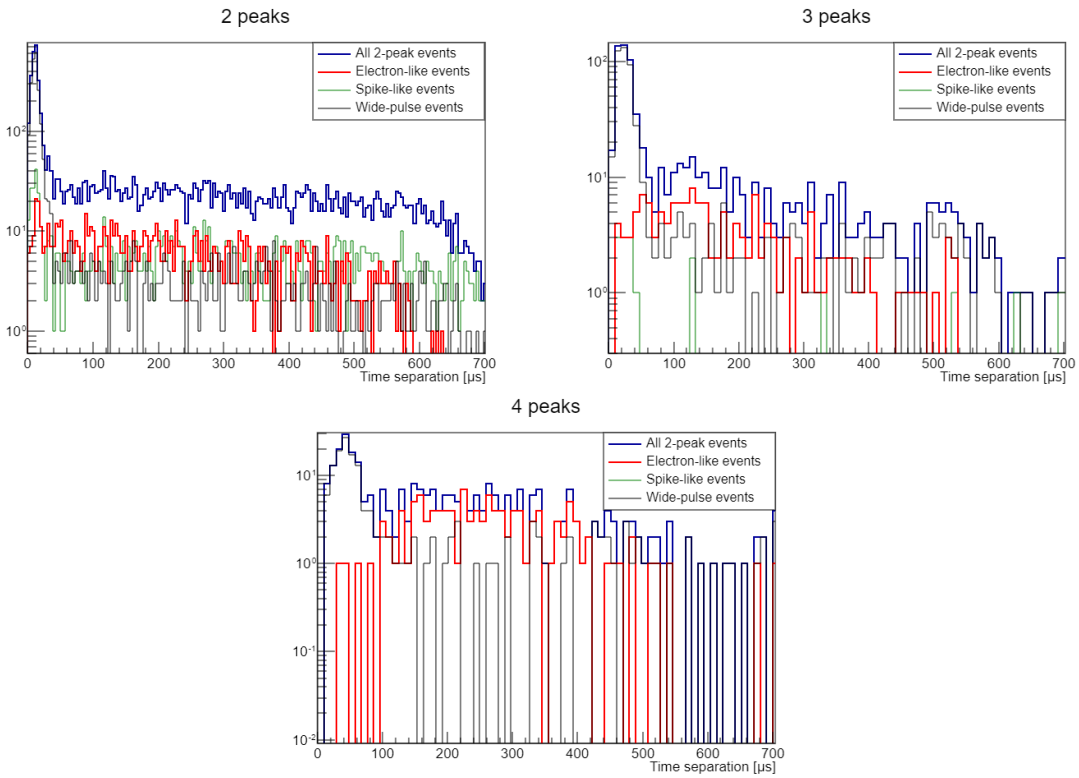


Figure 4.89: The Δt_{peak} distributions of the LSM test data separated into events with 2–4 peaks, from different PSD populations. The final physics data selection is shown in red ("electron-like" events). Credit to F. Vazquez.

- A cut to remove laser events from the physics data, as well as other quality cuts to remove pathological events listed in section 4.4.7,

- Events with positive peaks found in the north channel were cut, removing events with electrons shared between the two hemispheres (see sections 4.4.3 , 4.4.6).

Finally, the PF results were used to partition the data into events with 2, 3, and 4 peaks for the physics analysis, for which the number of events passing all cuts was 878, 146, and 122 respectively. The resulting time separation distributions are shown in Fig. 4.89. To avoid the excess of "wide peak" background events present at low Δt_{peak} (especially in the 2 peak data) – as well as high the Δt_{peak} regime where background modelling is less robust (see section 4.5.2) – the analysis range was restricted according to Table 4.8.

Table 4.8: Analysis ranges of Δt_{peak} for the WIMP analysis of the test data.

# of peaks	**Analysis range** [μs]
2	$19.2 \rightarrow 595.2$
3	$38.4 \rightarrow 748.8$
4	$76.8 \rightarrow 864$

4.5.1 WIMP time separation signal

To calculate the time separation signal from hypothetical WIMP events, the recoil energy spectrum given in section 2.3.1 must first be converted to the differential spectrum of the number of primary electrons produced in the SPC. This is done using the energy-response parameterization results of section 4.3, specifically $W(E)$, $F(E)$, and $Q(E)$, and using the COM-Poisson distribution to represent primary ionization statistics [140]. In this analysis, only WIMP scattering with the hydrogen in the methane gas molecules is considered. For a given WIMP recoil energy spectrum – defined by the particle mass, cross-section, and coupling – the total rate of events producing m electrons is given as:

$$R_{\text{WIMP}}(m) = \int_0^{E_{R_{\max}}} P_{\text{COM}}\left(m|\mu(E_R - E_{\text{Diss.}}), F(E_R - E_{\text{Diss.}})\right) \frac{dR}{dE_R} dE_R \qquad (4.34)$$

where the WIMP recoil spectrum dR/dE_R is taken from eq. 2.5, and $E_{R_{\max}}$ is the maximum possible recoil energy imparted by a WIMP (corresponding to the escape velocity of the galaxy) [58]. The mean number of primary electrons is calculated as follows:

$$\mu(E) = \frac{E}{W(E)} \times Q(E) \qquad (4.35)$$

Note that the W-value, Fano factor, and quenching factor are all evaluated at the

WIMP recoil energy less $E_{\text{Diss.}} = 4.55\,\text{eV}$, the energy required to disassociate a hydrogen atom from a methane molecule [287].

Next, the rate of events with m electrons can be used to calculate the expected time separation signal for different numbers of peaks using the PF model given in section 4.4.2, while at the same time taking into account all the aspects of the experiment's signal acceptance described in section 4.4:

$$\frac{dR\left(n, \Delta t_{\text{peak}}\right)}{d\Delta t_{\text{peak}}} = \sum_{m} \left[R_{\text{WIMP}}(m) \times \left(\int_{r_{\text{sensor}}}^{r_{\text{max}}} P_{\text{PF}}\left(n, \Delta t_{\text{peak}} | m, r\right) \times p_{\text{Fid.}}(m, r) \times p_{\text{Trig}}(m, \sigma_{\text{Diff.}}(r)) dr \right) \right] \times p_{\text{PSD}}(n) \times p_{\text{North coin.}} \times p_{\text{Livetime}} \tag{4.36}$$

where r_{sensor} is the outer radius of the ACHINOS sensor structure, the limit of particle detection in S140. Listed by their order in eq. 4.36, the aspects of the detector's signal acceptance are:

- $P_{\text{PF}}\left(n, \Delta t_{\text{peak}} | m, r\right)$ is the joint probability distribution function for the number of peaks observed and time separation, given the true number of electrons and radial position of the event. This was developed in section 4.4.2.

- The efficiency of the fiducial volume cut (selecting only south-channel events) is included as a function of the number of electrons and radial position of events as $p_{\text{Fid.}}(m, r)$ (see section 4.4.3).

- $p_{\text{Trig}}(m, \sigma_{\text{Diff.}}(r))$ is the hardware trigger efficiency of the S140 detector calculated in section 4.4.1, which is a function of both the number of electrons and the diffusion time of events (and therefore of their radial position).

- $p_{\text{PSD}}(n)$ is the signal acceptance for the PSD cuts placed on the N/S and Fisher discriminant variables, which is calculated for different numbers of peaks (see section 4.4.5).

- $p_{\text{North coin.}}$ – the loss of livetime associated with north coincident peaks – is applied in addition to the loss of signal acceptance due to the south-channel fiducialization. This is because the latter calculation only takes into account events with north-channel peaks originating from the same event producing south-channel peaks, not random coincidences with south-only events (see section 4.4.6).

- p_{Livetime} is the loss of livetime incurred by removing post-α time periods (section 4.4.4), laser events, and other quality cuts (section 4.4.7).

Note that there is an additional loss of exposure due to only utilizing the test data in this case. In practice, eq. 4.36 is calculated for up to $m = 15$ electrons, and truncated for $m < n$, neglecting the vanishingly small probability of FP noise peaks. Examples of the WIMP Δt_{peak} spectra for a dark matter particle mass of $0.76\,\text{GeV/c}^2$ and 2–4 peaks are shown in Fig. 4.90.

4.5.2 Background models

A large component of background events in the LSM data (especially at energies greater than one primary electron) was surface events – the source(s) of which remain unknown. This background may be due to ^{210}Pb and ^{210}Bi on the inner surface of the detector despite the electroplating and cleaning procedures described in section 4.1. Another hypothesis is that a thin layer of hydrocarbon material was deposited on the inner surface of the SPC, which would contain trace amounts of tritium. Regardless, this background component can be subtracted from the possible WIMP signal on the basis of its different time separation characteristics. This was done by generically assuming a differential background energy spectrum with a component that is constant in energy, as well as an exponentially rising component at low energy (to mimic the general low energy trends seen in the past [63]):

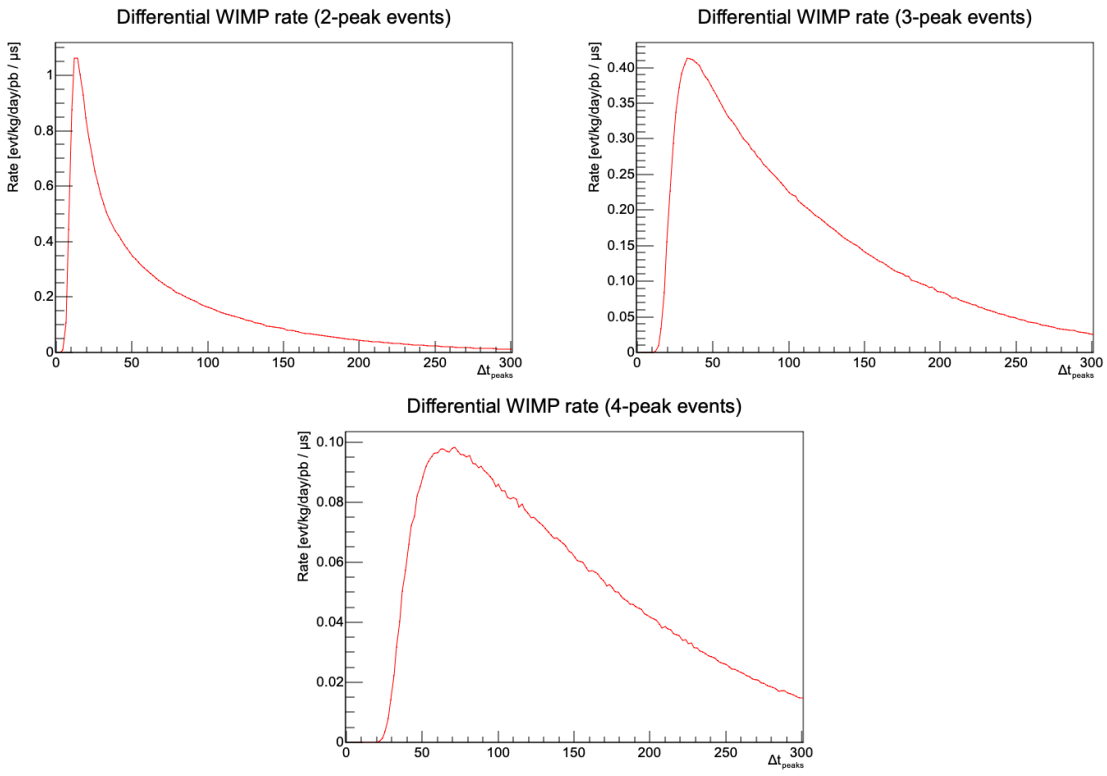


Figure 4.90: The Δt_{peak} signal (in μs) with 2–4 peaks, for a $0.76\,\text{GeV/c}^2$ WIMP scattering with atomic hydrogen. Credit to F. Vazquez.

$$\frac{d}{dE}R_{\text{Bkgd.}} = R_{\text{unif.}} + R_{\text{exp.}}e^{-E/\lambda_{\text{exp.}}} \tag{4.37}$$

Nominally, this distribution would be converted to a differential rate in terms of the number of primary electrons using eq. 3.4 and the COM-Poisson distribution for primary ionization. However, there is no closed form Jacobian for this change of variables, so to simplify, the Poisson distribution and Binomial distribution are used instead to represent the two extreme possible Fano factors for ionization. In both cases, the differential spectrum in terms of primary electrons m is of the form:

$$\frac{d}{dm}R_{\text{Bkgd.}} = A + B \times C^{-(m+1)} \tag{4.38}$$

Therefore, the surface background parameters are denoted as $A = R_{\text{unif.}}$, $B = R_{\text{expo.}}$, and $C = E_{\text{expo.}}$, all in terms of the number of electrons in an event m. The resulting time separation distributions for different numbers of peaks n is calculated using $P_{\text{PF}}\left(n, \Delta t_{\text{peak}} | m, r\right)$, the joint probability distribution function for n and time separation (see section 4.4.2), for $r = r_{\text{max}}$.

This parameterization – which connects the background levels for different numbers of electrons – was chosen specifically to enforce physically-possible relationships between the background rates in different numbers of peaks. If the background level were fit independently for different numbers of peaks, large fluctuations between them would be allowed by the fit, despite not being physically possible. Using this same generic form, a volume background contribution was considered for the final WIMP analysis as well, but ultimately was not used for the exclusion limit calculation, as the rates of volume and WIMP events are extremely degenerate.

Another significant source of background events is random coincidences – or "accidentals" (to distinguish these from PF algorithm coincidences) – containing two or more peaks by pure chance. Accidentals most likely involved post-α event single electrons (the largest source of single electrons by far) that evade cuts. These can pro-

mote single-peak events (which are omitted from the analysis) to 2-peak events which may be included in the physics data selection. Since the peaks in an accidental event are (largely) not correlated with each other, the distribution of their arrival time is necessarily uniform. While α-induced electrons do have an exponentially falling rate, the time constant for this process is estimated to be seconds-long, so on the time-scale of an 8 ms event window their rate is effectively constant.

Nominally, the distribution of Δt_{peak} for such events (at least for a 2 peak accidental) would also be uniform, truncating at the duration of the PF window. However, due to the DAQ software Samba's pulse centering algorithm, this is not the case. To account for this effect, a simple MC emulating the DAQ centering algorithm's effect on pulse treatment was carried out. In each MC event, the first peak was assigned an arrival time drawn from the distribution of peak times from single-electron data. This accounts for the effect of the raw pulse centering algorithm (which does not necessarily mean the double-deconvolved peak will be centered). Then, additional peaks were added to the MC event with a uniform arrival time distribution. The empirical impetus for this MC was α-correlated events, as their elevated rate of single electrons lead to many accidental events. As a background component of the physics data, the Δt_{peak} distributions for accidentals have an independent scaling factor for different numbers of peaks: $R_{2,\text{acc.}}$, $R_{3,\text{acc.}}$, and $R_{4,\text{acc.}}$.

4.5.3 WIMP exclusion limit

A WIMP 90% CL exclusion limit (for a SD-p coupling with atomic hydrogen) was calculated by F. Vazquez using the test data with the PLR method, as described in section 2.3.2 [81]. This was done to allow for a subtraction of the background components described in the previous subsection. This calculation was performed separately for logarithmically spaced values of WIMP mass M_χ – the results of which are then connected to form the exclusion limit.

First, a binned likelihood was constructed for the combined contributions of all

background components and WIMP signals in time separation, jointly for the 2, 3, and 4 peak spectra:

$$\mathcal{L}_{\text{WIMP}} = \mathcal{L}_{2p}(\sigma_{\text{SDp}}, M_\chi, \theta) \times \mathcal{L}_{3p}(\sigma_{\text{SDp}}, M_\chi, \theta) \times \mathcal{L}_{4p}(\sigma_{\text{SDp}}, M_\chi, \theta) \tag{4.39}$$

where each component of the likelihood function is the typical Poisson likelihood for the number of counts in each time separation bin. The parameter of interest in this case is the WIMP-hydrogen cross-section σ_{SDp}, which effectively scales the WIMP Δt_{peak} signal given by eq. 4.36. Next, θ represents the full set of nuisance parameters describing the background model, namely $\theta = \{R_{\text{unif.}}, R_{\text{expo.}}, E_{\text{expo.}}, R_{\text{2,acc.}}, R_{\text{3,acc.}}, R_{\text{4,acc.}}\}$. The expected number of counts in each time separation bin – for each number of peaks n – is given by adding contributions of the WIMP signal eq. 4.36, scaled accidentals background, and surface event background contributions. The PLR itself is calculated according to eq. 2.15:

$$\lambda(\sigma_{\text{SDp}}) = \frac{\mathcal{L}_{\text{WIMP}}(\sigma_{\text{SDp}}, \hat{\hat{\theta}})}{\mathcal{L}_{\text{WIMP}}(\hat{\sigma}_{\text{SDp}}, \hat{\theta})} \tag{4.40}$$

where $\hat{\hat{\theta}}$ are the profiled nuisance parameters (i.e. optimized for a given fixed value of σ_{SDp}), whereas the denominator of eq. 4.40 is the global maximum likelihood, optimized for all parameters. The optimization of all likelihoods in this case was done with ROOT's `TMinuit2` package [259]. Finally, the PLR test-statistic q is defined according to eq. 2.16.

In order to use this to calculate a limit, the value of σ_{SDp} is tuned to reach a certain threshold based on the distribution of the test statistic q [73]. Nominally, one could assume Wilk's theorem, which posits that this will be a χ^2 distribution with one degree of freedom [82]. In this case, the 90% CL exclusion limit would be set at the value of σ_{SDp} for which $q = 2.71$, the 90^{th} percentile of the χ^2 distribution. However, the general consensus in the field is that this should not be assumed [67]; instead one should directly compute the distribution of the test statistic with MC datasets. This

was done for a range of values of σ_{SDp} by:

1. Fitting the time separation spectra (of the test data) including the background contributions, as well as the WIMP signal for a given value of σ_{SDp}.

2. 2×10^5 MC datasets are generated using the best-fit.

3. The PLR test statistic (q) is calculated for each dataset, relative to the true underlying hypothesis for σ_{SDp}.

This procedure was carried out to generate the distribution of q assuming a background of surface + accidentals only, as well as with an added volume background following the same parameterization given in eq. 4.38, with independent parameters. The latter test was done to estimate the impact of volume background events if they are present, to ensure the PLR distribution is robust in the face of this possibility (note that the PLR calculation still assumes no volume background in this case). Additionally, the distribution of q was calculated with the input background scaled up and down by 30%, to investigate the impact of background mis-modelling. While the PLR distributions were found to deviate at times from Wilk's theorem (in different ways depending on the presence of a volume background or not), crucially these deviations were only present at values of σ_{SDp} far below the 90% CL exclusion cross-sections calculated for the data. For the relevant regime of σ_{SDp} at all WIMP masses, Wilk's theorem was shown to be true (within statistical fluctuations) in all cases, despite the various background mis-modelling scenarios. This can be explained by considering that for high enough WIMP cross-sections, the volume background contribution would be pushed to 0, negating its influence.

The WIMP exclusion limit on the test data was therefore calculated assuming Wilk's theorem, assuming only surface and accidentals background contributions. Examples of fits of the time separation spectra (for different numbers of peaks) are shown in Fig. 4.91 for a $0.66\,\mathrm{GeV/c^2}$ WIMP. This shows that the surface backgrounds

Table 4.9: Number of background events due to different contributions, for different numbers of peaks, corresponding to the profiled background model.

# of peaks	Surface	Accidentals
2	240	609
3	141	2.8
4	119	2.6

dominate, except for the 2 peak data which has an even greater component of accidentals. The WIMP contribution is most prominent in the 3 peak data, which is where most of the sensitivity of the experiment comes from. The profiled background model contributions are given in table 4.9. In addition to the exclusion limit, 1 and $2\,\sigma$ sensitivity bands were calculated for the experiment by generating datasets from the background-only, best-fit of the time separation spectra. The limits for these MC datasets are then determined in the same way. These results are shown in Fig. 4.92.

This calculation includes uncertainties on all the physics characterization/signal acceptance parameters described in the previous subsections. Uncertainties that affect the exclusion limit notably include U (from $W(E)$), and the pressure of gas in the SPC. The conservative $2\,\sigma$ error-limits of all such parameters were used to calculate the exclusion curve. However, by far the largest systematic uncertainty in this analysis is the quenching factor. As discussed in section 4.3.4, some extrapolation of $Q(E)$ beyond the lowest energy measurements must be assumed, with no choice being beyond reproach. Therefore, the exclusion limit was calculated using two choices for this extrapolation: the generic logarithmic extrapolation (meant to be more conservative than a Lindhard-like trend), and the null extrapolation, (i.e. assuming $Q(E) = 0$ below the lowest energy measurement). The limits calculated for both extremes are shown in Fig. 4.92, which shows the dramatic effect of this choice. However, even in the null-extrapolation case, the exclusion limit obtained is still a

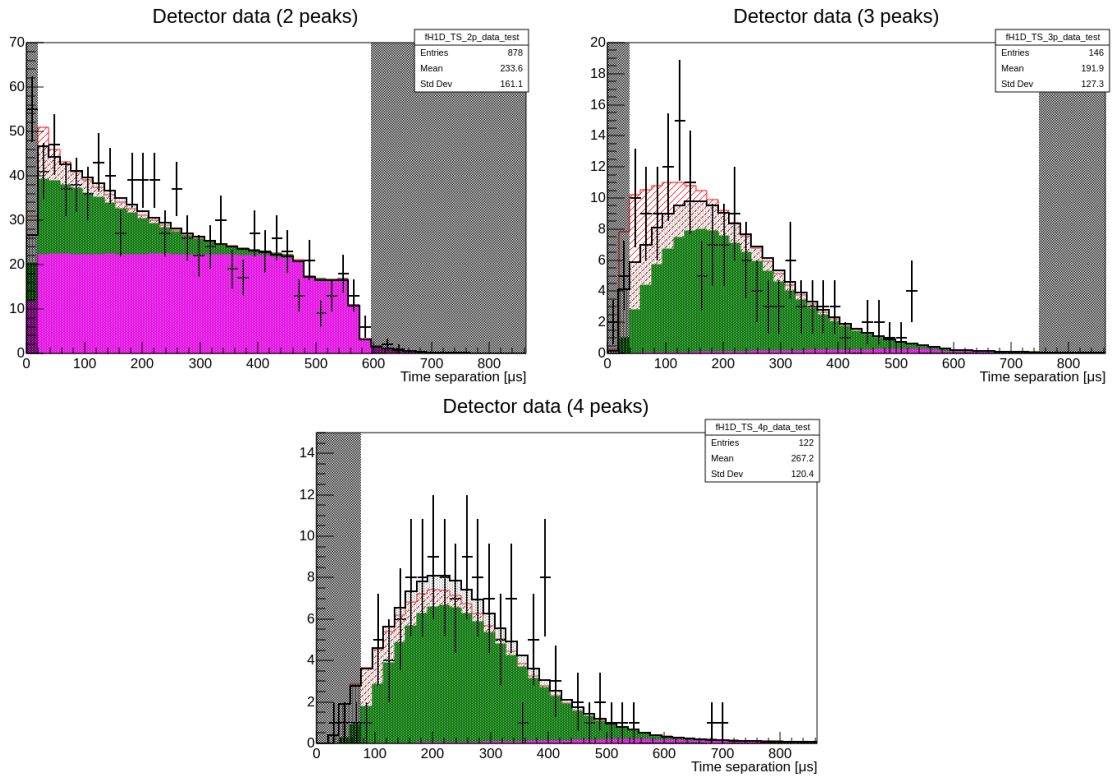


Figure 4.91: The Δt_{peak} physics data with 2–4 peaks (black histograms). The stacked contributions of accidentals (magenta), surface backgrounds (green), and the WIMP signal corresponding to the 90% CL exclusion limit for a $0.66\,\text{GeV/c}^2$ dark matter particle (with $\sigma_{\text{SDp}} = 32.7\,\text{pb}$ (red hashed histogram) are shown. For comparison, a backgrounds-only fit of the data including a volume background component is shown as the black curve. Credit to F. Vazquez.

notable improvement in the low-mass sector beyond existing results.

Based on the preliminary dark matter result presented in this section (using the test data), the NEWS-G collaboration can expect to set world-leading constraints on low-mass WIMPs with a SD-p coupling. Specifically, the exclusion limit calculated with the test data is currently the world-best for particle dark matter in the mass range 0.17 – $1.03\,\text{GeV/c}^2$ (or 0.26 – $0.68\,\text{GeV/c}^2$ if the zero-extrapolation quenching factor is assumed). The best sensitivity is achieved at a mass of $0.66\,\text{GeV/c}^2$, with $\sigma_{\text{SDp}} = 32.7\,\text{pb}$. When re-calculated with the full blind physics dataset, this limit is expected to improve by a factor of approximately 65%, based on the larger exposure of the blind data.

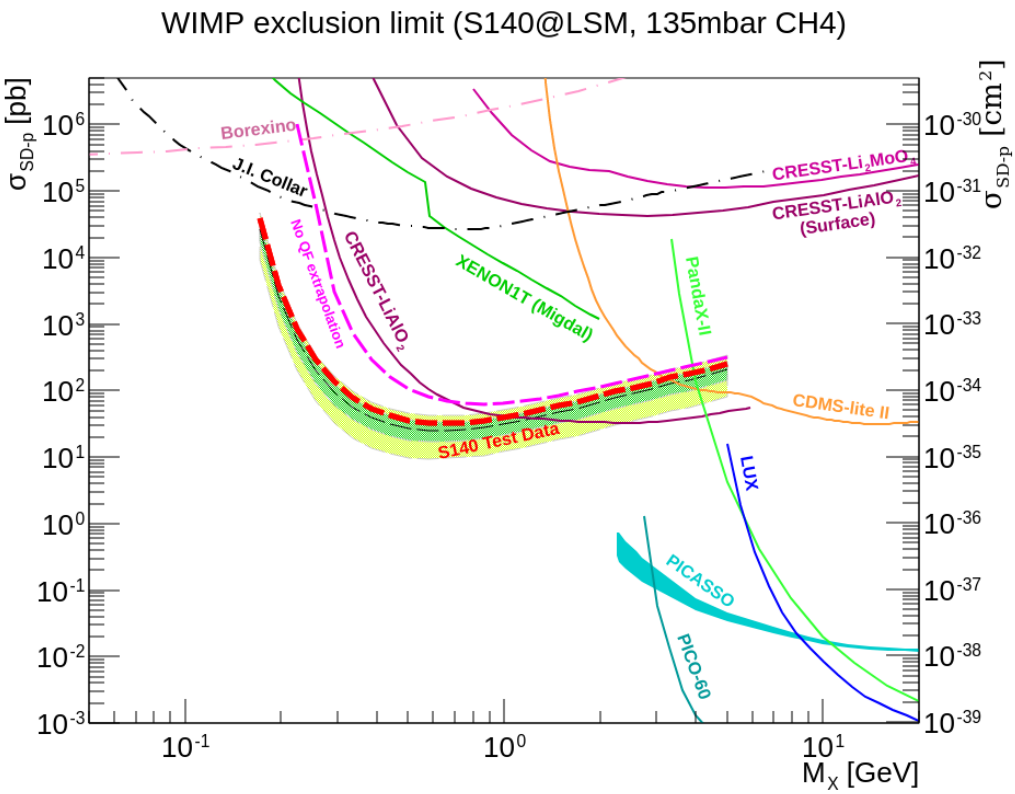


Figure 4.92: Preliminary WIMP 90% CL exclusion limit for SD-p coupling with atomic hydrogen, calculated with the LSM test data (red dashed curve), as well as the 1 and 2 σ sensitivity bands (green and yellow shaded regions). Limits from other existing results as shown as well, including from the PICO-60 experiment [93], PICASSO [288], LUX [289], PandaX-II [290], CDMS-lite II [110], CRESST-III Li_2MoO_4 device [117], CRESST-III $LiAlO_2$ detector [118, 291], XENON1T [292], J.I. Collar result [293], and Borexino experiment [294]. Credit to F. Vazquez.

4.5.4 Data unblinding

At the time of this writing (October 2023), the LSM analysis is in the final stages of collaboration review. Once the analysis performed on the non-blind test data is validated, the remaining $\sim 70\%$ of the physics data will be un-blinded. This data will be analyzed in the same manner described in this document, including cuts applied, background model fitting, and calculation of the WIMP exclusion limit. Since the blind data was chosen randomly from amongst the full datasets, there is no reason to suspect that it will have fundamental differences compared to the test data. However, immediately upon unblinding, there are several checks planned intended to verify that this is the case. They include:

- The evolution of PSD variable populations over time will be examined. Specifically, this will involve re-making Fig. 4.81, with the N/S and Fisher discriminant PSD variables for the blind data. This is intended only as a qualitative check. If in fact there are differences in the PSD populations of the blind data, the PSD cuts designed based on the test data will not be fully optimized. However, the signal acceptance for WIMP events – calculated only using the laser data – would not be called into doubt.

- The stability of the rate of single-electron events over time will be verified, as a proxy for the general rate stability of WIMP candidate events. While a change in the rate of single-electron events in the blind data compared to the test data would be deeply puzzling, as before, the WIMP exclusion limit calculation would still be robust.

- Background-only fits of the time separation spectra (with and without a volume contribution), and separate limits for 2–4 peak events will be generated with the blind data. This test is motivated by the tension between the 3 peak time separation data and the background-only model. This is especially evident

without a volume contribution, and still at a $2\,\sigma$ discrepancy level with a volume contribution (comparing the 3 peak exclusion limit and sensitivity bands). The blind data may reinforce – or erase – this tension. Therefore this comparison will be repeated, as well as χ^2 values calculated for the 3 time separation spectra. To account for the look elsewhere effect of these multiple tests, the adjusted significance for each test will be 1.3% to achieve a total significance level of 5% ($p = 0.05$).

Chapter 5

Bubble chamber nucleation efficiency studies

Bubble chambers have existed as a detector technology since the mid-twentieth century [295], with applications including high energy particle physics [296] and radiation dosimetry for neutrons [297]. The basic premise is the use of superheated fluids as a target medium. In this meta-stable state, particle interactions that deposit sufficient energy in a small enough volume can induce cavitation in the fluid, leading to a visible bubble. The sensitivity of these detectors to $\mathcal{O}$(keV) nuclear recoils – while remaining relatively insensitive to electronic recoil events – makes them natural choices as nuclear recoil, rare-event search experiments. This includes both direct detection searches for WIMP-like dark matter, as well as coherent neutrino scattering experiments, with a strong interest in large-scale experiments planned/currently under construction [298–301].

To conduct rare-event searches (resulting in either a claim of discovery or exclusion of a hypothesized interaction), it is crucial to precisely know the response of bubble chambers to nuclear recoils, i.e. the probability that recoils of a given energy will induce an observable nucleation. To that end, dedicated nuclear recoil calibrations have been carried out for bubble chambers from both the PICO and SBC collaborations (for fluorocarbon and liquid noble target fluids, respectively). These calibrations and subsequent analyses are presented here. The PICO analysis is already the subject

of a publication [3] – therefore the description of this work presented here will draw heavily from this paper. The SBC analysis will be featured in an upcoming publication, although some details have already been made public in a conference proceeding [302]. The results obtained through these analyses are of significant impact on the bubble chamber experiments they involve, and will be directly used for any future dark matter or coherent neutrino scattering results from either collaboration. Future nucleation efficiency calibrations will likely reuse these methodologies as well.

5.1 Bubble chambers for dark matter detection

Early bubble chamber experiments were typically large detectors employed in high energy physics experiments, such as the Big European Bubble Chamber or Gargamelle [303, 304]. Target media included Freon (a pragmatic choice due to its near room-temperature boiling point), liquid hydrogen, deuterium, and so on. Typically the chamber would be placed at the collision point of a particle accelerator beam to image the resulting cacophony of subatomic particles, with an applied magnetic field to bend the tracks of charged particles to aid in identification. In a similar fashion to a cloud chamber, this resulted in clear tracks from incident particles, revealing their travel, collisions, and decays. An example of a bubble chamber image with particle identification from the Big European Bubble Chamber is shown in Fig. 5.1. This type of particle-ID analysis led to major breakthroughs in the field, such as the discovery of neutral current interactions (implying the existence of the Z_0 boson) [305]. Such analyses did not (nor did they need to) consider the exact mechanics of bubble nucleation or its subtleties (i.e. interactions of different particle species). This is in stark contrast to the approach of rare-event searches, for which the exact threshold nature of bubble chambers must be known.

The PICASSO [307] and COUPP [308] experimental groups adopted bubble chamber technology to search for particle dark matter beginning in the late 1990s, using fluorocarbon target liquids including C_4F_{10}, CF_3I and C_3F_8. Fluorine-rich targets provide excellent sensitivity to spin-dependent coupling with protons, as they have an odd proton number [66, 71]. While the PICASSO experiment used a series of droplet suspension bubble chambers (similar to dosimeter bubble chambers), COUPP used a fully liquid target medium, compressing the fluid to arrest each nucleation.

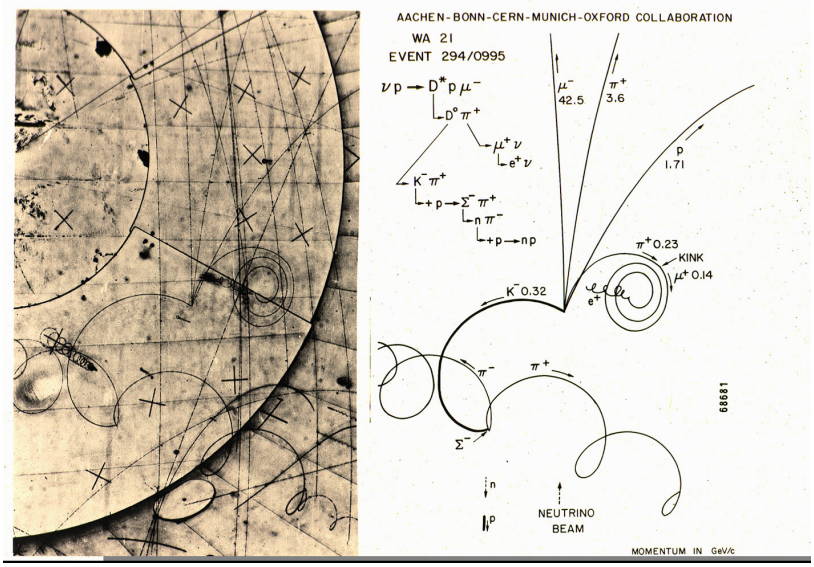


Figure 5.1: An example of an image of tracks produced by a neutrino beam in the Big European Bubble Chamber, with a labelled interpretation of the event. In this case, a neutrino-proton collision produced a D meson, which subsequently decayed into various charged leptons and hadrons. Image taken from ref. [306], copyright @ 2023 CERN.

5.1.1 The PICO collaboration

The PICO collaboration (the merger of PICASSO and COUPP) uses bubble chambers with fluorocarbon fluids. In most previous detector configurations, an artificial quartz glass vessel filled with target liquid is suspended from a pressure bellows system. A water buffer layer above the target fluid prevents it from coming into contact with the bellow's surface. The active volume is then chilled to its target temperature (e.g. 14 to 16° in the case of [93]), and the target pressure is controlled with the bellows system. An example of this design – the PICO-2L detector – is shown in Fig. 5.2 [309].

Detectors are instrumented with piezo-electric transducers to record audio signals from nucleation events, and a stereoscopic camera system. The recorded data allows for 3D position reconstruction of bubble positions (to reject surface background events). Further, discrimination against alpha-particle recoils is possible based on their characteristically different acoustic power [64]. With this technology, the PICO

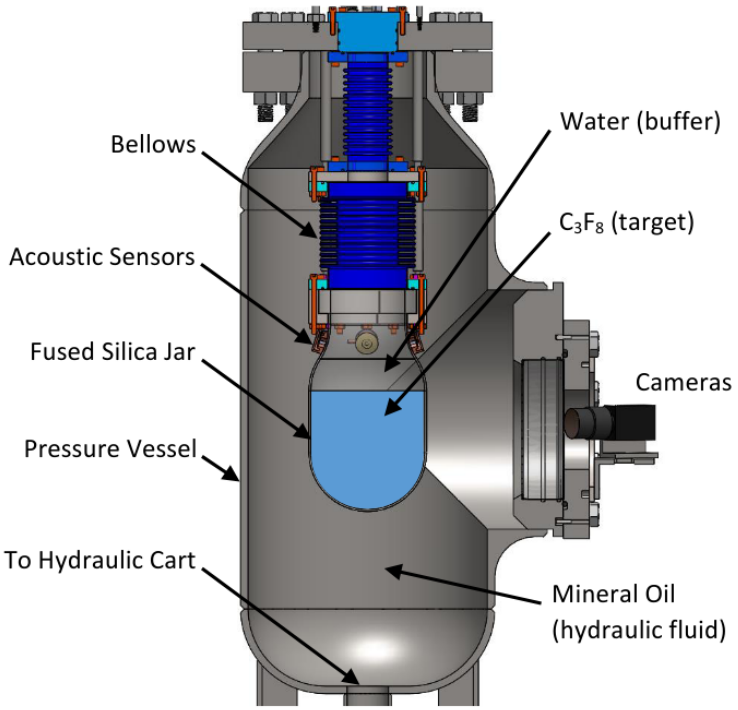


Figure 5.2: Schematic of the PICO-2L chamber with the target volume (in a silica jar) suspended below a water buffer layer, and then compression bellows. Taken from figure 1 of ref. [309].

collaboration recently published results of the PICO-60 detector, with a 52 kg C_3F_8 target [93]. With a 1400 kg day exposure, the WIMP-proton spin-dependent interaction is excluded down to $2 \times 10^{-41}\,\mathrm{cm}^2$ at a WIMP-mass of $20\,\mathrm{GeV/c}^2$, the best constraint to date in this parameter space. To calculate the sensitivity of the experiment, the nucleation efficiency of chamber/target fluid must be known – in this case, a preliminary version of the analysis presented in this chapter was used [277].

Detectors with the typical PICO design suffer from several contamination issues. In some iterations of the experiment, it was found that microscopic metal debris produced by the movement of the steel bellows could drift down to the active target volume and induce nucleation [310]. Additionally, mixing between the water buffer and target fluid can create induce nucleation near the liquid interface [93]. Newer chamber designs seek to solve these issues by having the compression bellows situated below the active volume, and applying pressure via a second nested jar. A strict

inverted temperature gradient is enforced, so that the fluid at the bottom of the vessel that is in contact with the bellows (a feature likely to induce bubbles) is kept in a cooler, non-superheated state. In this way, the active volume of liquid is only in contact with the silica jars, with no risk of particulate contamination from the bellows components or buffer liquid [311, 312].

5.1.2 The SBC collaboration

The Scintillating Bubble Chamber (SBC) collaboration employs liquid noble bubble chambers to search for low-mass dark matter. Unlike fluorocarbon bubble chambers, liquid nobles (i.e. xenon and argon) produce scintillation light – in addition to nucleation – while in a superheated state [313]. This extra channel offers the possibility of energy reconstruction for particle interactions. Specifically, 175 nm and 128 nm scintillation light is produced by LXe and LAr respectively. Photons of such wavelengths as LAr usually need to be wavelength shifted towards the optical spectrum to allow for detection with photomultiplier tubes (PMTs), but this can be accomplished through a small addition of LXe as a dopant [94]. SBCs made by the collaboration typically follow a similar "right side up" design to recent PICO experiments [311, 312], with either PMTs or silicon photomultipliers to record scintillation light. Since LXe and LAr are cryogenic fluids, the temperatures at which they reach superheat are significantly lower than fluorocarbon liquids (typically around $-40°$C for LXe), requiring cryogenic cooling systems to operate.

The first experiment conducted by the SBC collaboration was with a small 30 g LXe bubble chamber, constructed in the "right side up" configuration. It was instrumented with a single PMT to measure scintillation light produced (see section 5.5 for more details about the experimental setup) [314]. This experiment gave the first confirmed scintillation light in coincidence with bubble nucleation, an example of which is shown in Fig. 5.3, a critical proof-of-concept result for the collaboration [314].

Historical LXe bubble chambers revealed an interplay between scintillation light

yield and the probability of nucleation [315]. Specifically, the latter was found to increase when a scintillation quencher was added to the target fluid. This energy partitioning – possibly different for varying particle species – may explain the apparent suppression of electronic recoil-induced bubbles observed in LXe [314]. With an even greater electronic event suppression than fluorocarbon chambers, it may be possible to operate scintillating bubble chambers with energy thresholds as low as $\sim 100\,\mathrm{eV}$ (at which point fluorocarbon chambers would be overwhelmed by electronic recoils), greatly expanding their sensitivity to low-mass dark matter.

After the success of the LXe demonstration experiment, the SBC collaboration is currently fabricating two 10 kg LAr chambers for placement at SNOLAB and Fermilab, to search for dark matter and coherent elastic neutrino-nucleus scattering (CEνNS) respectively [10, 94, 298]. The design of these chambers is depicted in Fig. 5.4. With greater scintillation light collection efficiency (via the many silicon photo-multipliers they are instrumented with), energy reconstruction may be possible. Taking advantage of the expected strong electronic recoil suppression of LAr, these chambers may be operated successfully in a high-background environment next to a nuclear reactor to search for CEνNS. The ability to operate the chamber with a threshold of 1 keV or less will also allow for unprecedented sensitivity to WIMP-like dark matter of particle masses on the order of $1\,\mathrm{GeV/c^2}$ [298].

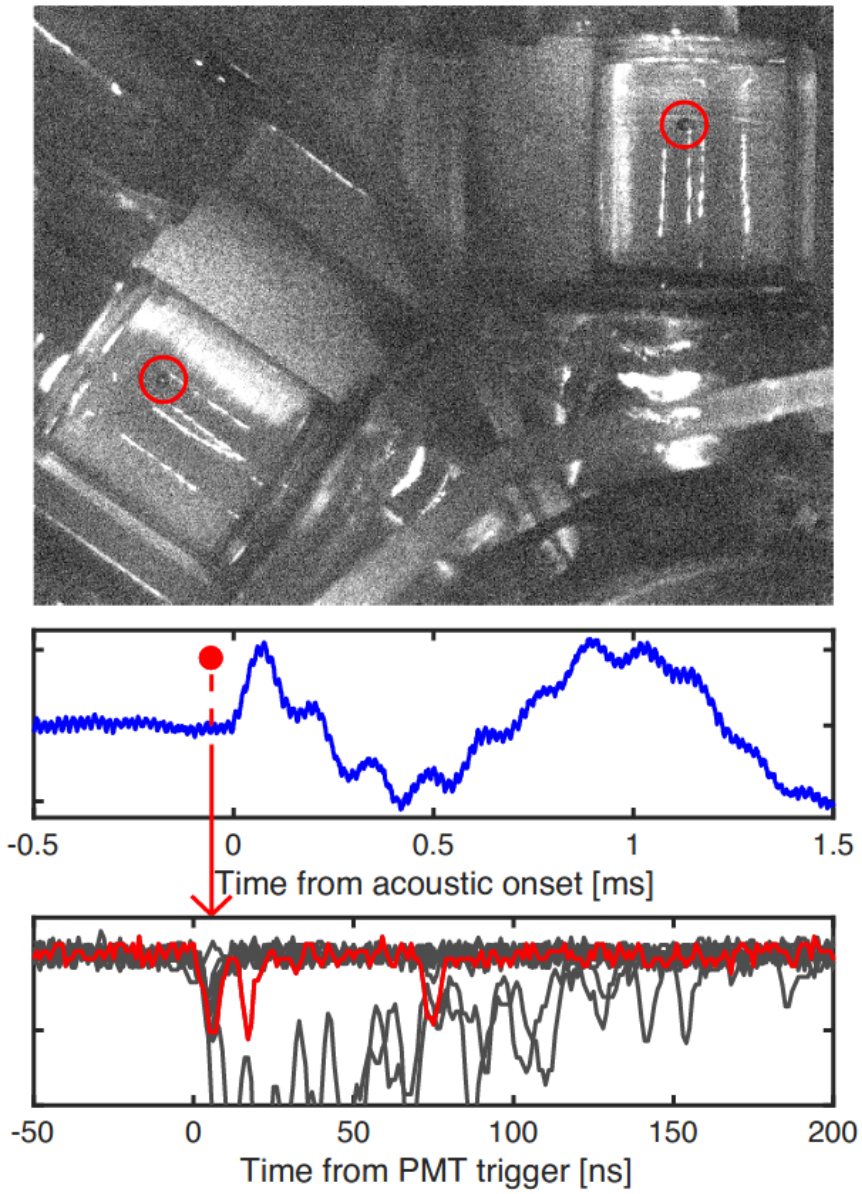


Figure 5.3: An event from the LXe bubble chamber with an optical image of the bubble shortly after formation (top) in coincidence with an acoustic signal recorded by a piezo-electric transducer (middle), and PMT signal from scintillation light (bottom). The gray PMT traces are other waveforms recorded throughout the acquisition window for the event, with the red trace being the coincident scintillation light. The slight delay between the onset of the acoustic pulse compared to the PMT and optical signal is indicative of the speed of sound in LXe. Taken from figure 2 from ref. [298].

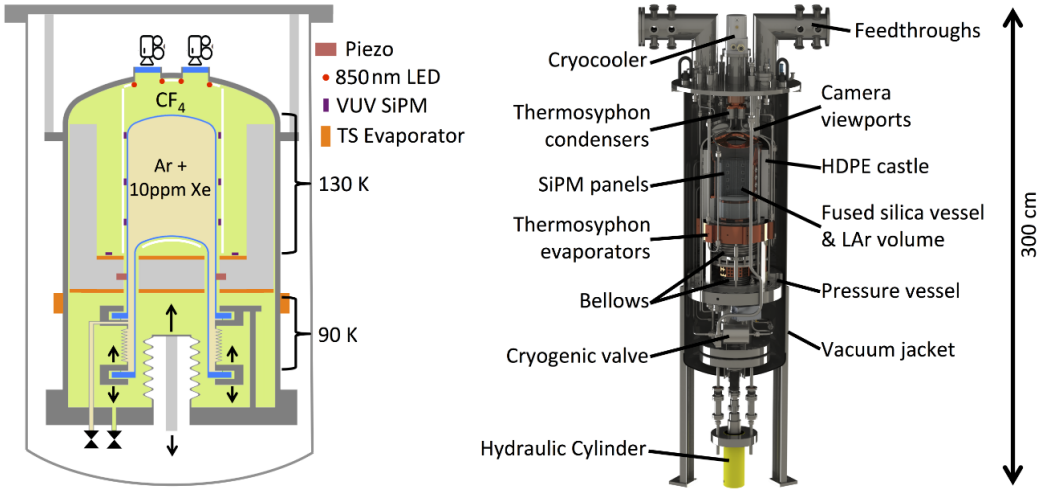


Figure 5.4: Schematic (left) and CAD model (right) of the 10 kg LAr scintillating bubble chamber design, instrumented with silicon photo-multipliers, cameras, and LEDs for photo-multiplier calibration. Taken from figure 6 of ref. [298].

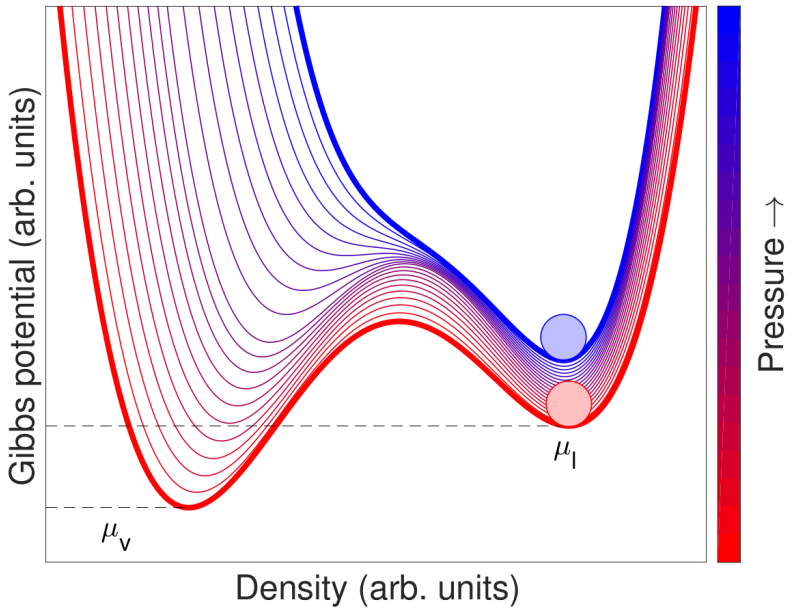


Figure 5.5: The Gibbs potential of a generic fluid as a function of density, for different pressures, revealing the metastable super-heated state μ_l at a higher potential than the vapour state μ_v. Taken from figure 2.1 of ref. [316].

5.2 Seitz model of nucleation

A bubble chamber may be operated with any target fluid in which it is practical to achieve the necessary thermodynamic conditions. The fluid begins as a liquid; pressure is slowly reduced while maintaining a constant temperature. As the target fluid's pressure is reduced below its vapour pressure, a metastable minimum develops in its Gibbs potential at a higher density than its stable minimum state as a vapour. This is generically depicted in Fig. 5.5 [316]. In this metastable super-heated state, the fluid remains a liquid. However, a small addition of energy can cause an amount of fluid to overcome this potential barrier and rapidly boil, causing a cavitation [317]. The nature of this process makes bubble chambers threshold detectors.

This potential barrier represents the energy required to overcome the surface tension of the liquid, for the vapour bubble to expand, etc. This energy can be provided by a particle interaction in the liquid that deposits sufficient energy in a small volume, by disrupting the potential of the fluid by reducing surface tension (i.e. from a piece

of dust, vessel-wall imperfections, etc.), or by any local injection of heat. Once a small cavitation bubble of vapour forms, boiling is triggered in the surrounding fluid, leading to the growth of a macroscopic bubble within nanoseconds [318]. Typically only nuclear recoil interactions can deposit sufficient energy densely enough in the fluid to cause boiling, hence bubble chamber's natural suppression of electronic recoil signals, which do not heat the liquid enough locally.

A theoretical treatment of the threshold for nucleation in bubble chambers is given by the Seitz "hot spike" model [317], with a modern treatment from [65]. In this paradigm, nuclear recoils are assumed to locally heat the target liquid, creating a small bubble of vapour known as a "proto-bubble". Proto-bubbles larger than a certain critical radius r_c will continue to grow into macroscopic bubbles of vapour, which is defined by the point at which the pressure of the vapour inside the bubble P_b overcomes the pressure from the surrounding liquid P_l and surface tension of the fluid σ [65]:

$$r_c = \frac{2\sigma}{P_b - P_l} \tag{5.1}$$

The heat energy required to reach this critical radius, therefore, defines the theoretical nucleation threshold. This is given by [65]:

$$Q_{\text{Seitz}} \approx 4\pi r_c^2 \left(\sigma - T\frac{\partial\sigma}{\partial T}\right) + \frac{4\pi}{3} r_c^3 \rho_b (h_b - h_l) - \frac{4\pi}{3} r_c^3 \left(P_b - P_l\right) \tag{5.2}$$

where T is the temperature of the fluid, ρ_b is the density of vapour in the proto-bubble, and h_b and h_l are the specific enthalpies of the vapour and liquid respectively. These three terms come from the work done to create the proto-bubble surface, to vapourize its' contents and the reversible work of the first two terms together. A unitless factor $\lambda \approx 1$ (the "Harper parameter") is used as an *ad hoc* scaling factor for the critical radius [65]. Thus in the Seitz model, any particle depositing energy E_{dep} greater than Q_{Seitz} within a bubble of distance less than λr_c (typically $\mathcal{O}(1\,\text{nm})$) will always cause

nucleation [3]:

$$E_{\text{dep}} = \int_0^{\lambda r_c} \frac{dE}{dx} dx \geq Q_{\text{Seitz}} \tag{5.3}$$

In practice, this theoretical threshold can be entirely controlled by the temperature and pressure maintained in the superheated fluid. The requirement of depositing energy in such a small radius usually precludes lightly-ionizing particles from creating a bubble, such as photons or electrons. This provides a strong natural background suppression for bubble chamber experiments [65]; indeed only nuclear recoils or alpha particle interactions typically induce nucleation.

5.2.1 Existing empirical results

While the Seitz model provides a useful – and indeed the only (currently) – first-principles framework to estimate the threshold for nucleation, it is well-known empirically that the Seitz model does not accurately describe the probability that a nuclear recoil will induce nucleation [3, 302, 309, 319–325]. Straggling of nuclear recoils, variability in recoil track structure, and surface tension effects [326] shift the true threshold to higher energies than the Seitz prediction, and broaden the efficiency curve, such that it is no longer a step function threshold. An example of disagreement between measurement and the Seitz model (independent from the results of the studies presented here [3, 302]) is a previous calibration by the COUPP experiment of a CF_3I chamber with Am-Be and Y-Be neutron sources, which indicate that the Sietz model underestimates the threshold for bubble nucleation, as shown in Fig. 5.6 [320].

Various parameterizations have been adopted to describe the nucleation threshold in previous analyses, as a function of recoil energy and thermodynamic state (Q_{Seitz}). For example, in ref. [322] the predicted Seitz step threshold is smeared by a Gaussian function to generically represent the stochastic processes that may lead to deviation from the predicted threshold. However, this approach is not physically motivated and

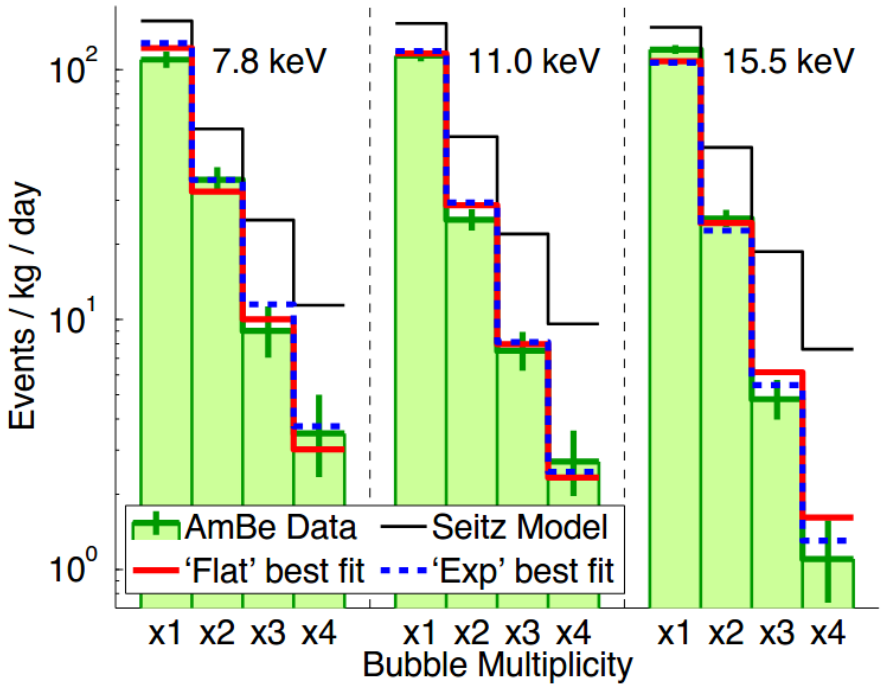


Figure 5.6: Observed count rates of different bubble multiplicities at various thermodynamic thresholds (values of Q_{Seitz}), from an Am-Be neutron calibration dataset. The red and blue dashed curves represent various best-fit parameterizations of the nucleation threshold derived from an MCNP simulation. The solid black curve shows the Seitz model prediction, which overestimates the observed count rates. Taken from figure 5 of ref. [320].

predicts nucleation from recoil energies below the Seitz threshold, for which there are no hypothesized mechanisms.

Another approach taken in ref. [320] was to either to use a scaled-down step function (one that steps at the Seitz threshold, but to some efficiency $< 100\,\%$), or a function of the form

$$P(E_r, Q_{\mathrm{Seitz}}) = 1 - exp\left(-\alpha \frac{E_r - Q_{\mathrm{Seitz}}}{Q_{\mathrm{Seitz}}}\right) \tag{5.4}$$

where α is a shape parameter estimated from calibration data ($\alpha = 0.15$ for carbon and fluorine) [320]. This model assumes that the probability of nucleation will never reach $100\,\%$ at any recoil energy, as well as that the efficiency rises above $0\,\%$ directly at the Seitz threshold, which is incompatible with subsequent measurements [3, 302,

319]. Therefore, a more flexible nucleation efficiency model was adopted for the work presented here (described in section 5.4.1), which can approximately accommodate both of the previous functional forms as well as many others, albeit at the expense of requiring many more model parameters.

5.3 PICO nuclear recoil calibrations

To obtain an empirical measurement of the nucleation efficiency in C_3F_8 bubble chambers, the PICO collaboration gathered nuclear recoil calibration data with various neutron sources, including compound neutron sources and accelerator-produced neutrons. The detectors used included small 2 L and 0.1 L bubble chambers [65, 309] (the former is depicted in Fig. 5.2). The PICO-0.1L vessel was designed to be a portable test chamber for detector characterization and was therefore suitable for temporary installation at an accelerator beam site. Both chambers were designed in the fashion of early PICO experiments, with a quantity of superheated C_3F_8 target fluid in quartz jars, with a water buffer above so that there was no contact between the steel bellows of the chamber re-compression system and the target fluid.

5.3.1 Calibration sources

Am-Be and Sb-Be compound sources were employed, which produce fast neutrons through (α, n) and (γ, n) reactions respectively [247]. Am-Be sources produce a broad spectrum of $\mathcal{O}$(MeV) neutrons, which can scatter multiple times even in a small chamber (producing up to seven or more bubbles in the PICO-2L detector). The many multi-scatter events provide a unique constraint on bubble nucleation, as the ratio of different scatter multiplicities is entirely due to the geometry of the experiment.

Sb-Be sources primarily produce 24 keV neutrons via the reaction $^9\mathrm{Be}(\gamma, \mathrm{n})^8\mathrm{Be}$, induced by a 1690 keV gamma-ray emitted by ^{124}Sb (branching ratio 48.4 %). A small number of 378 keV neutrons (5.7 %) are also produced [247, 327]. The downside of this calibration source is the large number of gamma rays emitted (at approximately 10^6 times the rate of neutrons produced). While bubble chambers are largely inherently blind to such radiation, the high rate of gammas and relatively low thermodynamic threshold of the test chambers did result in some rate of gamma-induced bubbles. Lead disks were placed in front of the source to partially address this issue

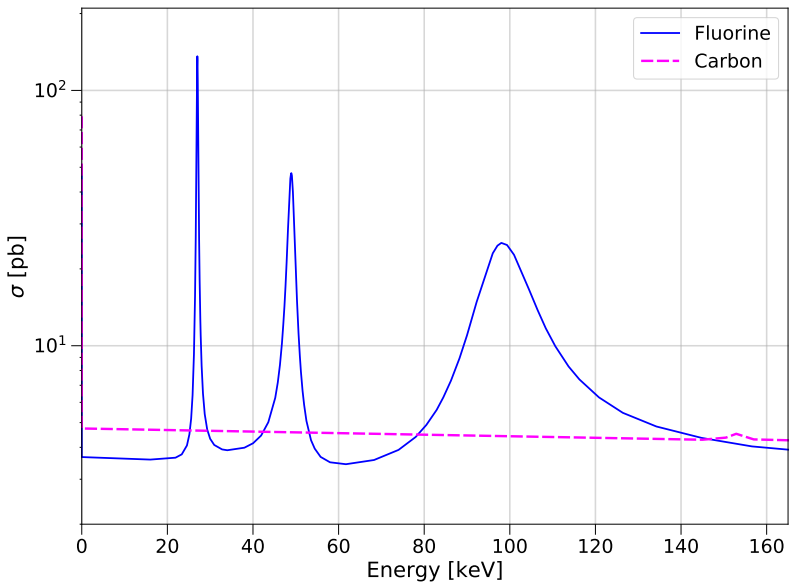


Figure 5.7: Total neutron interaction cross-sections for fluorine (^{19}F specifically) and elemental carbon, showing several resonances for fluorine [331].

by attenuating the gamma flux.

Other calibration experiments measured monoenergetic neutrons produced using the TANDEM accelerator at the Université de Montréal [328]. The proton beam was directed at a vanadium target, which produced nearly-monoenergetic neutrons via the reaction $^{51}\mathrm{V(p,n)}^{51}\mathrm{Cr}$; in a thin vanadium target, the neutron kinetic energy is equal to the proton energy (tunable from the accelerator) minus the Q-value of the reaction, 1564 keV. Neutron production was enhanced by matching the proton energy to several vanadium (p, n) cross-section resonances [329, 330]. Calibration data was taken with neutron kinetic energies of 50, 61, and 97 keV.

The varying kinetic energy of the accelerator neutrons allowed for the exploitation of resonances and anti-resonances in the neutron-fluorine cross-section, shown in Fig. 5.7 [331]. Compared to the roughly constant cross-section of carbon across the approximately 50 keV to 100 keV energy range of the beam-induced neutrons, the varying fluorine cross-section results in some calibration datasets were much more sensitive to fluorine-induced nucleation than others, disentangling the nucleation contributions of the two target atom species.

Dataset	Detector	Thresholds (keV)	Livetime (minutes)	Multiplicity
97 keV Beam	PICO-0.1	3.0, 3.2, 3.6	21, 9.9, 20	1,2,3+
61 keV Beam	PICO-0.1	2.9, 3.1, 3.6	16, 160, 18	1,2,3+
50 keV Beam	PICO-0.1	2.5, 3.0, 3.5	3.1, 7.7, 7.3	1,2,3+
Sb-Be	PICO-0.1	2.1, 2.6, 3.2	320, 310, 300	2,3+
Am-Be	PICO-2L	3.2	2200	1,2,3,4,5,6,7+

Table 5.1: PICO neutron calibration datasets used in this analysis, listing the detector and Seitz thresholds used, as well as bubble multiplicities (number of bubbles per event) included in the subsequent analysis.

5.3.2 Neutron scattering data

To thoroughly explore nucleation efficiency as a function of recoil energy and the thermodynamic state of the bubble chambers, data was collected using the neutron sources listed above, with varying temperature and pressure conditions. The thermodynamic state of each experiment is associated with its corresponding Seitz threshold Q_{Seitz} for the sake of simplicity, even though the inaccuracy of the Seitz model is the foundational argument for this work. The datasets used in the final analysis – the neutron sources, detectors used, thermodynamic states, and bubble multiplicities observed – are presented in Table 5.1. Data was collected at higher thermodynamic thresholds (see Fig. 5.9), but this was not used in the final analysis presented in this work.

In all cases, the data from every event consists of camera images captured when a pressure or acoustic spike caused by a bubble triggers the acquisition. The images are hand-scanned to verify bubble counts and positions in each event, which is assumed to be 100% efficient for bubble identification [3]. Time periods in which the pressure or temperature of the chambers deviated from the nominal conditions were removed from the datasets. Following every nucleation and recompression cycle, 10 (resp. 30) seconds were discarded in data taken with the PICO-0.1L (resp. PICO-2L) chamber to ensure the target fluid was fully restored to its intended superheated state. To remove

events possibly caused by wall imperfections or water-target interface bubbles (see section 5.1.1), single bubbles occurring near the vessel wall were discarded in PICO-2L datasets (as described in [309]). For PICO-0.1L datasets, a stronger fiducial volume cut was applied to remove single or multi-bubble events originating within a few cm of the water-target interface, which was found to become especially mixed through repeated recompression cycles [3]. After applying all the analysis cuts described above, the total livetime of the detector in each dataset was calculated, which is given in Table 5.1. The bubble event rates observed after analysis cuts are applied are shown in Fig. 5.8, with rate uncertainties calculated using the Feldman-Cousins method [240].

Background data was collected for all calibration setups (keeping the same geometry and conditions), in the absence of the source/beam, taken close in time to the dataset with the corresponding source. For all calibration sources except the Sb-Be experiment, the background was dominated by cosmogenic neutrons. In the case of Sb-Be, the high gamma flux induced a large rate of single bubble background events, so this data was deemed unusable. However, gamma rays are only likely to produce a single bubble, such that multi-bubble events from the Sb-Be source are dominated by neutrons and are therefore included in the analysis. This behaviour – the high rate (and lack) of gamma-induced single bubble (resp. multi-bubble) events was confirmed by taking data with the Sb source present but without the beryllium disk.

For the beam experiments, the background was found to be approximately 0.2 single bubbles/minute (with a negligible rate of coincident, multi-bubble background events) [3]. For the Sb-Be calibration, the background rate of 2-bubble events was measured as 0.004 bubbles per minute, and the background rate of the Am-Be experiment was consistent with 0. The background-subtracted data is shown in Fig. 5.9. This includes datasets taken at higher thermodynamic thresholds that were not included in the final analysis (or Table 5.1). However, these data points illustrate the expected trend of decreasing bubble rates as the thermodynamic threshold increases,

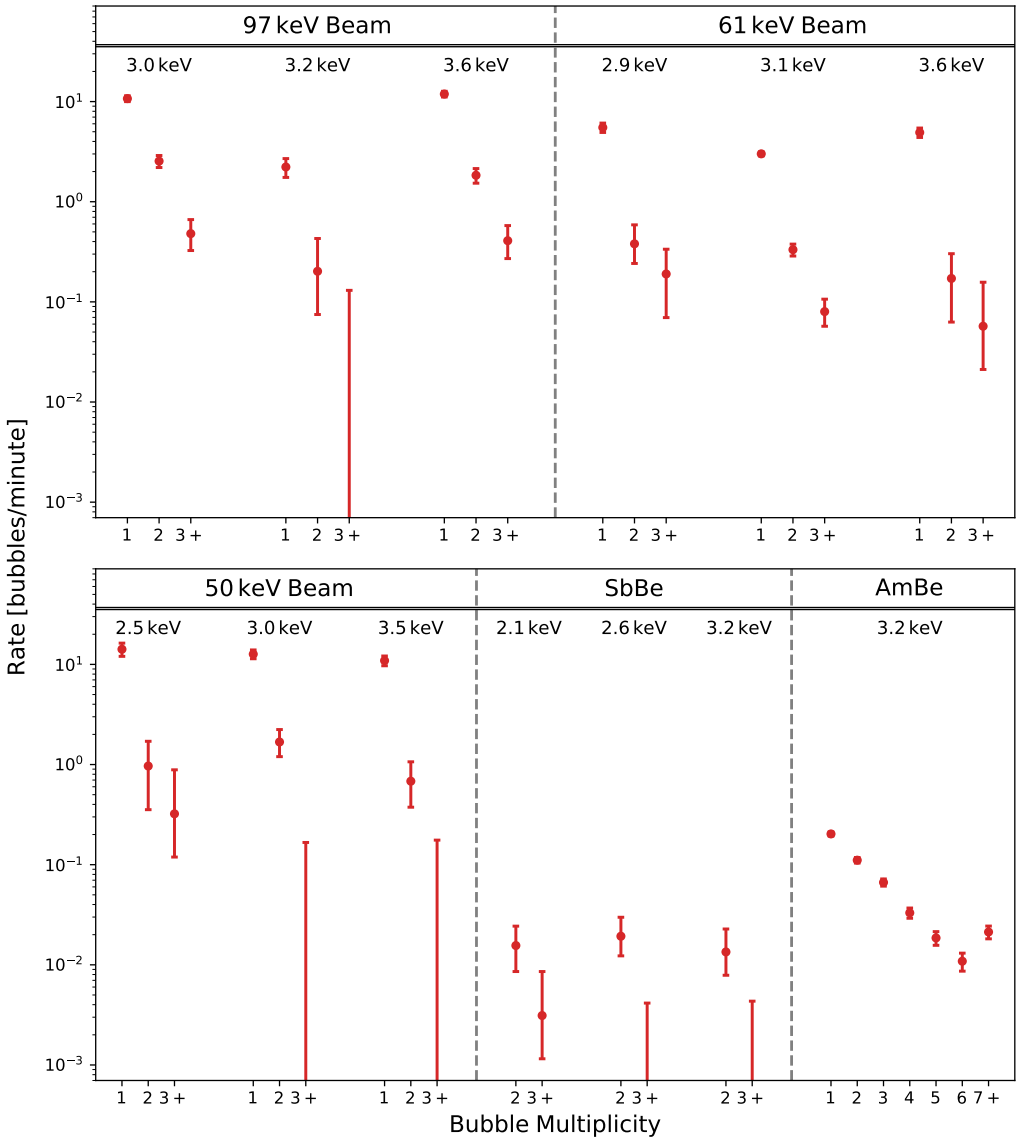


Figure 5.8: Bubble count rates from all PICO nuclear recoil calibrations (sources indicated at the top of each section), for all thermodynamic states and bubble multiplicities. The error bars shown are Feldman-Cousins $1\,\sigma$ confidence intervals [240].

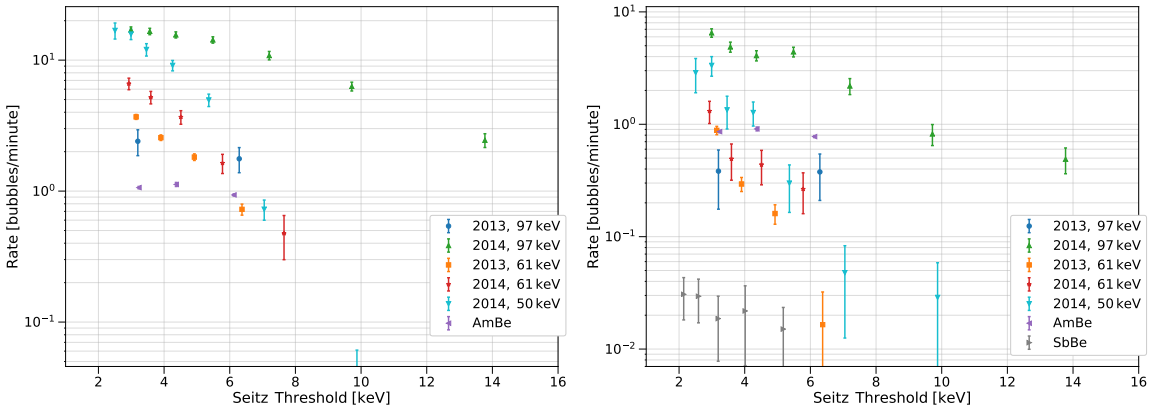


Figure 5.9: Background-subtracted bubble rates of single bubble (left) or multi-bubble (right) events. Each bubble in a multi-bubble event counts separately towards the rate. The error bars shown represent statistical uncertainty. Taken from figure 7 of ref. [3].

resulting in sensitivity to less of the neutron recoils from each source.

5.3.3 Calibration simulations

For each experimental setup, a detailed Monte Carlo simulation was created to calculate the incident neutron recoil spectrum in the corresponding bubble chamber. This was carried out with Geant4 for the Sb-Be experiment [204], and with MCNP-POLIMI for the beam experiments [332]. The Am-Be setup was simulated with both softwares, which were found to agree (the MCNP simulation was ultimately used in the final analysis [3]). As many materials can moderate or reflect neutrons, these simulations were necessarily very detailed, including all components surrounding each bubble chamber or associated with the neutron source. An example of a CAD model used to produce an MCNP simulation is shown in Fig. 5.10.

The cross-section data used by the simulation software is taken from the ENDF/B-VII library [331, 333], with corrections to the fluorine cross-sections from [334]. For the beam setup simulations, the differential rate of neutrons coming from the vanadium target ($dR/dEd\Omega$) was adopted from ref. [330] and fed into the MCNP simulation. For the Sb-Be simulation, a GEANT4 simulation of the gamma interactions

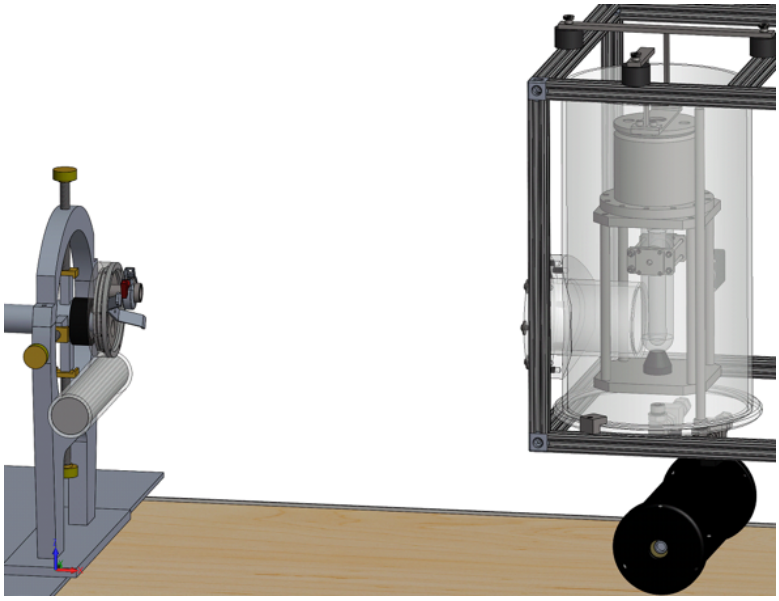


Figure 5.10: CAD diagram of the beam calibration setup at the Université de Montréal. On the left is the vanadium beam target in a calipered holder, with a helium-3 counter below it. On the right is the PICO test bubble chamber, with a second helium-3 counter below. Taken from figure 5 of [3], copyright Miaotianzi Jin, 2019 [277].

produced by the Sb source was first calculated, giving initial starting positions and energies for neutrons within the beryllium disk as an input for a separate GEANT4 simulation [3]. The neutron recoil spectra used in subsequent analyses are shown in Fig. 5.11. The final calculation of these rates includes the fiducial volume cuts applied to the data (see section 5.3.2).

5.3.4 Simulation flux normalization

Ancillary measurements of the neutron flux in each calibration setup were performed to reduce uncertainty on the neutron flux that would be adopted for the simulations described in section 5.3.3 alone. The exception is the Am-Be data, whose high-multiplicity events self-constrain the flux of neutrons.

The neutron flux of the Sb-Be source was directly measured with a ^{3}He counter placed at the location of the PICO-2L detector, which yielded an estimate of 209 ± 22 neutrons per second at the target location at the time of the measurement. The decay

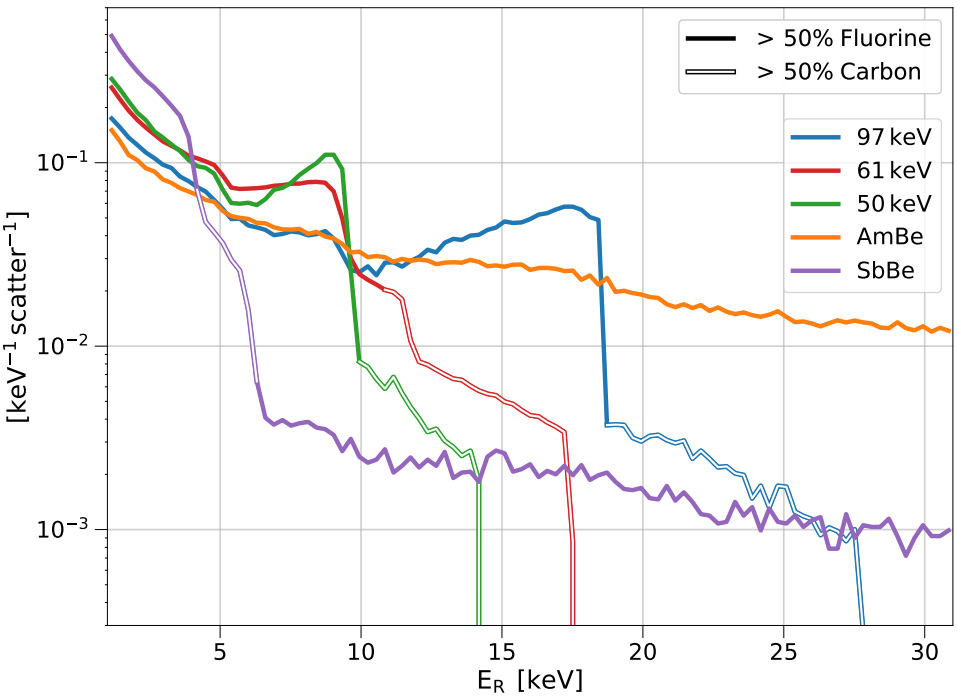


Figure 5.11: Simulated neutron recoil spectrum for all calibration setups, normalized per scatter. The spectra are labelled to indicate in what energy ranges scattering with fluorine or carbon target atoms is dominant. Taken from figure 8b of ref. [3].

of ^{124}Sb over time reduced the neutron yield, so the flux was corrected depending on the time any particular dataset was taken.

For the neutron beam experiments, the neutron flux was continuously measured using a ^{3}He counter placed below the vanadium target, as shown in Fig. 5.12. Data was recorded with a second ^{3}He counter placed near the PICO-0.1L chamber (as shown in Fig. 5.10), however comparison between the rate observed by the detector and the simulated rate at that location is affected by uncertainties in the geometry of the PICO-0.1L detector, its surrounding water bath, etc.

To better assess the neutron flux at the location of the target fluid of the PICO-0.1L chamber, a measurement was carried out in which the bubble chamber was removed, and a ^{3}He counter was placed at that location instead. The ratio of the neutron flux at the vanadium target compared to the target location was measured and is given for each beam energy in Table 5.2. The simulation-predicted flux ratio is also

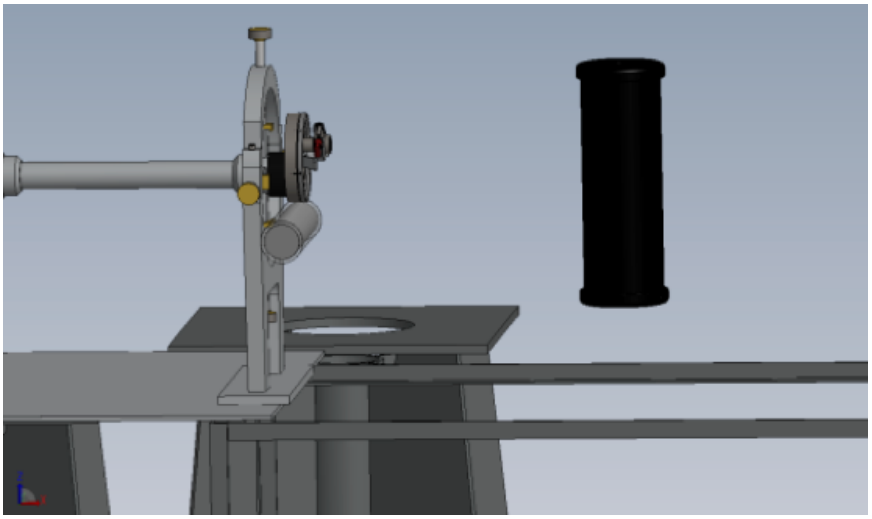


Figure 5.12: Experimental setup at the Université de Montréal for measuring the neutron flux. Two ^{3}He counters were used: one directly below the ^{51}V target and one suspended directly downstream of the target. Taken from figure 9 of ref. [3], copyright Miaotianzi Jin, 2019.

given, and reveals a tendency of the simulation to over-predict this flux ratio. Rather than correcting for this potential bias, the discrepancies are used to define systematic uncertainties on the neutron fluxes for those experiments. A second method of empirically estimating the neutron flux of the beam experiments involved measuring the decay of activated ^{51}Cr produced by the proton beam in the vanadium target [330]. This approach yielded agreeable results for the neutron flux of the beam experiments but was not used in the final analysis [3].

Energy	Measured ratio	Simulated ratio	Measurement / simulation
50 keV	2.28 ± 0.07	2.26 ± 0.08	1.01 ± 0.05
61 keV	2.02 ± 0.07	2.26 ± 0.08	0.89 ± 0.04
97 keV	2.07 ± 0.10	2.21 ± 0.07	0.93 ± 0.05

Table 5.2: Table II from ref. [3]. Flux ratio of the vanadium target site to PICO-0.1L fluid location for different beam experiments (listed by neutron energy), obtained by ^{3}He counter measurements and from MCNP simulations. The final column gives the ratio of the two results, indicating that the MCNP simulations more often over-predict this ratio compared to the measurements.

5.4 PICO nucleation efficiency analysis

To obtain a measurement of the nucleation efficiency of C_3F_8 from the neutron calibrations presented above, the expected recoil rate from the corresponding simulations ($R_{\text{Sim.}}$) is multiplied by a given nucleation efficiency model to calculate the expected count rate, which is then compared to the data to fit the model. The nucleation model – denoted as $\epsilon_s(E_R, Q)$ – may be a function of recoil energy E_R as well as thermodynamic state parameterized by Q_{Seitz}, and will depend on the atomic species s with which the nuclear recoil occurred. In this notation, the expected rate of nucleations for a given efficiency model is:

$$R_{\text{obs.}} = \sum_s \int_0^\infty \frac{dR_{\text{sim.}}}{dE_R} \times \epsilon_s(E_R, Q) dE_R \tag{5.5}$$

In this analysis, a joint fit of all calibration datasets was carried out (data summarized in Table 5.1), resulting in an estimation of the nucleation efficiency model for both carbon and fluorine [3].

5.4.1 Nucleation efficiency model

Due to the lack of a physically-motivated model for nucleation efficiency as a function of recoil energy (see section 5.2.1), the model chosen for this analysis was a generic piecewise function. It is defined by fixed efficiency knots at 0 %, 20 %, 50 %, 80 %, and 100 % with floating recoil energies at which each efficiency knot is reached (the free parameters of the model). These floating recoil energies are denoted as $\{x\}_{s,Q}$ (for atomic species s and thermodynamic state Q). This set of parameters, along with the efficiency knots, define the piecewise linear efficiency threshold function $\epsilon_s(E_R, Q)$ for a given Q, as a function of E_R. This scheme is depicted in Fig. 5.13.

To address the dependence of the model on thermodynamic state Q, as before there is no *a priori* functional form to be used. In the case of this analysis, most calibration datasets were clustered around $Q_{\text{Seitz}} = 2.45\,\text{keV}$ and $3.29\,\text{keV}$ (see Table 5.1), which

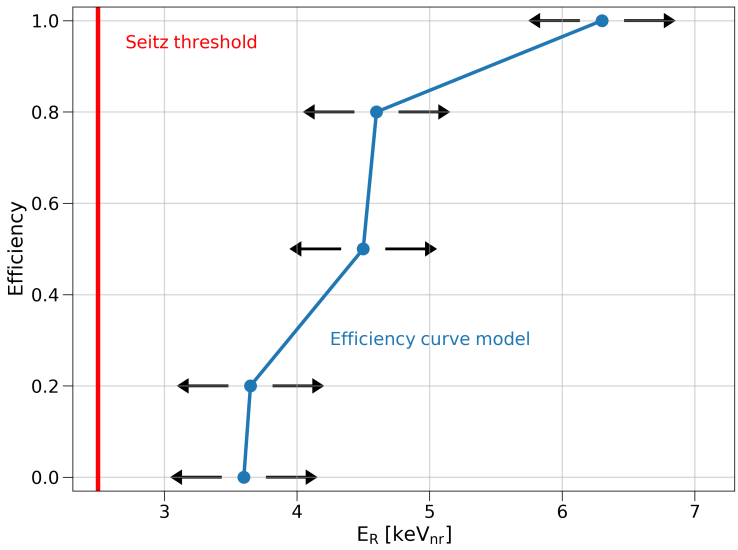


Figure 5.13: Depiction of the piecewise linear model used as a nucleation efficiency model in the present analysis, with fixed efficiency knots and floating recoil energies at which those efficiencies are reached.

were the main operating conditions of the PICO-60 physics campaign [93]. Therefore, the data would poorly constrain $\epsilon_s(E_R, Q)$ at values of Q not close to either cluster. In light of this, the simple solution adopted was to define $\epsilon_s(E_R, Q)$ at $Q = 2.45\,\text{keV}$ and $Q = 3.29\,\text{keV}$, and proportionally scale the efficiency curve for each calibration measurement at Q to the nearest of these two setpoint curves – a sort of "nearest neighbour scaling" to $\hat{Q}$, the closer of the two threshold setpoints:

$$\{x\}_{s,Q_i} = \frac{Q_i}{\hat{Q}}\{x\}_{s,\hat{Q}} \tag{5.6}$$

Overall, this model is described by 20 parameters of interest:

$$\{x\}_{C,Q=2.45\,\text{keV}}, \quad \{x\}_{F,Q=2.45\,\text{keV}}, \quad \{x\}_{C,Q=3.29\,\text{keV}}, \quad \{x\}_{F,Q=3.29\,\text{keV}} \tag{5.7}$$

where each set $\{x\}$ includes the recoil energies at which the five efficiency knots are reached. Several physical constraints apply, the first of which is that the efficiency curves must monotonically increase as a function of recoil energy:

$$\frac{\partial \epsilon_s}{\partial E_R} \geq 0 \tag{5.8}$$

Additionally, the efficiency curves are taken to monotonically increase as a function of Q:

$$\frac{\partial \epsilon_s(E_R, Q)}{\partial Q} \geq 0 \implies \{x\}_{s,Q=3.29} \geq \{x\}_{s,Q=2.45} \tag{5.9}$$

Because there is the expectation that the true nucleation threshold is at higher recoil energies than the predicted Seitz threshold, the model efficiency curves are constrained to always be greater than or equal to the corresponding value of Q:

$$\{x\}_{s,Q} \geq Q. \tag{5.10}$$

One final constraint comes from the consideration that the dE/dx of fluorine in C_3F_8 is higher than that of carbon [183], meaning that it is always more likely that a recoil on fluorine will produce nucleation than one on carbon. Therefore, the efficiency curves for carbon are constrained to be greater than or equal to the ones for fluorine at the same value of Q, and at the same efficiency knot:

$$\epsilon_C(E_R, Q) \geq \epsilon_F(E_R, Q) \implies \{x\}_{C,Q} \geq \{x\}_{F,Q} \tag{5.11}$$

In the code implementation of this model, several coordinate transformations are performed to the parameters of the model [3]. Specifically for each $\{x\}_{s,Q}$, the coded parameters $\{y\}_{s,Q}$ are the cumulative difference of the logarithm of the un-transformed parameter values. The efficiency curves in the nominal coordinate system can therefore be reconstructed from the code parameters as follows:

$$\{x_k\}_{s,Q} = \sum_{i=1}^{i=k} e^{\{y_i\}_{s,Q}} \tag{5.12}$$

The advantage of this coded representation is that it naturally enforces a physical constraint of the model, namely eq. 5.8. At the same time it also removes abrupt

parameter boundaries (by using the logarithm of the parameters), ideally resulting in a more Gaussian likelihood for those parameters.

5.4.2 Likelihood function

The likelihood function for the data given a nucleation efficiency model $\epsilon_s(E_R, Q)$ is formed by taking the Poisson probability for the observed number of counts k, given an expected number of counts ν. The log-likelihood for this term is then calculated, using the Stirling approximation for the denominator [335]:

$$P_{\text{Poisson}}(k|\nu) = \frac{\nu^k e^{-\nu}}{k!} \tag{5.13}$$

$$\implies \log P_{\text{Poisson}}(k|\nu) = k\log\nu - \nu - k\log k + k \tag{5.14}$$

$$= k - \nu + k\log\left(\frac{\nu}{k}\right) \tag{5.15}$$

This term is summed over all experiment/thermodynamic state combinations i and bubble multiplicities j (see Table 5.1) to form the global log-likelihood function:

$$\log\mathcal{L} = \sum_i \sum_j \left[-\nu_{i,j} + k_{i,j} + k_{i,j}\log\left(\frac{\nu_{i,j}}{k_{i,j}}\right)\right] \tag{5.16}$$

The expected counts $\nu_{i,j}$ are calculated by multiplying the expected count rate (see eq. 5.5) by the corresponding livetime for each dataset (see Table 5.1), adding the measured background rate (see section 5.3.2) to correspond to the observed counts (which include the same backgrounds):

$$\nu_{i,j} = \left(Ri, j + \mathrm{R}_{\mathrm{Bkgd}_{i,j}}\right) \times \mathrm{Livetime}_{i,j} \tag{5.17}$$

5.4.3 Systematic uncertainties

There are several systematic uncertainties associated with the nuclear recoil calibrations, some of which apply to all experiments equally, others to specific source setups. These systematic uncertainties relate to either the thermodynamic state of

each dataset (represented by Q_{Seitz}), or to the total exposure of the experiment to neutrons. The former derives from the stability of pressure and temperature control of the bubble chambers (approximately 1 psi and 0.1 °C) as well as long-term fluctuations in temperature. The latter includes uncertainties on the calibration geometry (such as the depth of the water bath of the PICO-0.1L detector), or the source flux of neutrons, incorporating neutron flux measurements described in section 5.3.4.

Systematic uncertainty on neutron exposure has a direct effect on the expected number of counts in a dataset (due to eq. 5.5). The impact of uncertainty on the thermodynamic state of each experiment comes in the form of a) uncertainty on the lower bound of the efficiency curve for that dataset (see eq. 5.10) and b) a systematic on the thermodynamic state-scaling of the efficiency curve for the dataset, based on eq. 5.6. These experimental systematic uncertainties are given in Table 5.3. A final, global systematic on the thermodynamic state of 3 % is included to capture uncertainty on the Seitz calculation itself, which contains *ad hoc* parameters such as the Tolman length [3, 326].

All systematic uncertainties are incorporated into the nucleation efficiency model as nuisance parameters. These are treated as multiplicative factors applied to their corresponding quantity in the simulation or model, and are taken to be log-normally distributed. Thus, the likelihood for the logarithm of each term is a Gaussian likelihood expressed in terms of η_l, the multiple of standard deviations each nuisance parameter l deviates from its expected value. Thus, the total log-likelihood for all nuisance parameter values η_l is a log-Gaussian centered at 0 with standard deviation 1,

$$\log \mathcal{L}_{\mathrm{Nuisance}} = -\sum_{l} \frac{\eta_l^2}{2}, \tag{5.18}$$

which is then appended to the global log-likelihood given in eq. 5.16. To slightly reduce the number of nuisance parameters in the model, all thermodynamic state

Table 5.3: Systematic uncertainties of the PICO nuclear recoil calibration data [3]. Uncertainty on Q_{Seitz} derives from uncertainty on the pressure/temperature regulation of the bubble chambers used, and both geometry and source systematics result in uncertainty on the total neutron exposure of the experiment. In instances where there are multiple values given per table cell, the uncertainty varies for the different thermodynamic states that the experiment was operated with.

Experiment	Q_{Seitz} [%]	Geometry-exposure [%]	Source-exposure [%]
Beam, 2013, 97 keV	8.0	7.5	8.3
61 keV	8.0		4.7
Beam, 2014, 97 keV	7.0 / 9.9	3.1	5.9
61 keV	1.7 / 2.5		5.0
50 keV	7.0 / 14		5.4
Sb-Be	6.5 / 7.0 / 7.5	10.3	
Am-Be	8.0	26	

systematics for each experimental setup (the 2013 and 2014 beam, Sb-Be source, and Am-Be source setups) are taken to be perfectly correlated. Thus, they can be grouped into four nuisance parameters in addition to the global Q_{Seitz} systematic due to the Tolman length. Together with the nine nuisance parameters describing the geometry and source strength systematics (see Table 5.3), there are a total of fourteen nuisance parameters in the model.

5.4.4 MCMC fitting approach

As discussed in appendix A, fitting high-dimensional models such as the one described above can present a severe computational challenge, as sampling of the likelihood function near the optimum can be very inefficient. Additionally, the likelihood presented in section 5.4.2 includes a relatively small number of data points, fit jointly from multiple separate measurements, with highly constrained model parameters. All

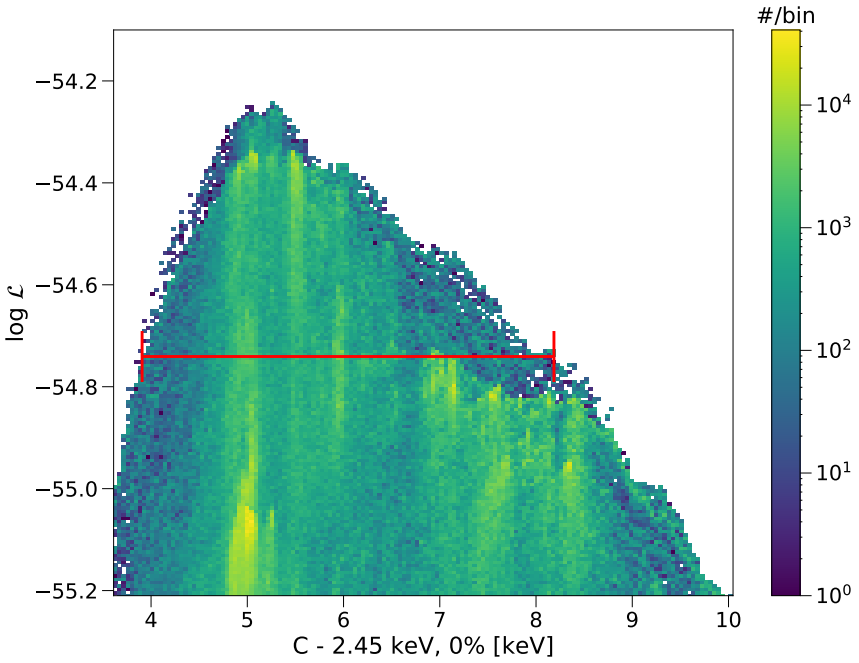


Figure 5.14: A depiction of how the "1 σ volume" is calculated for a single parameter, in this case the 0 % efficiency point for carbon with $Q_{\text{Seitz}} = 2.45\,\text{keV}$.

of these factors may lead to a severely non-Gaussian likelihood function, making the use of a typical gradient-descent optimization algorithm a dubious choice. Therefore, the fast burn-in MCMC approach [277] described in appendix A was utilized for this analysis, as a robust and efficient optimization tool.

A critical consideration in using this MCMC approach (or indeed any MCMC algorithm) to fit a model to data is ensuring/assessing convergence. One common heuristic used in MCMC analyses is the random walker auto-correlation time [216, 336], the number of steps needed until the MCMC walker's positions are independent of their starting guess. This ensures that the fit is not affected by a poor initial guess, but is not sufficient to guarantee that the MCMC has thoroughly explored the region of parameter space around the maximum likelihood point, meaning the results of the fit may still be subject to significant statistical fluctuations. Further, this common metric is not applicable to the fast burn-in method, since the walkers are frequently interrupted.

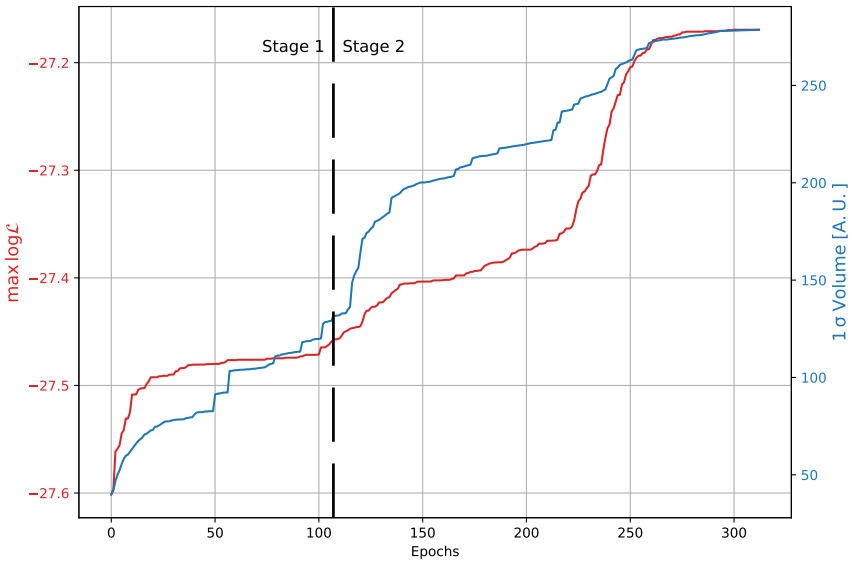


Figure 5.15: The maximum log-likelihood and "1 σ volume" as a function of epoch number for the fit of the PICO calibration data, as convergence metrics. Taken from figure 18 of ref. [3].

A logical alternative is to look at the progression of the maximum log-likelihood value found by the MCMC as a function of step number (or equivalently, the epoch number). Requiring that this value be stable over a large number of steps would serve as a good metric of the convergence of the best fit. However, it is also important that the boundary of the likelihood function as mapped by the MCMC becomes stable, as this is how parameter uncertainties are calculated. To assess this, a "1 σ volume" quantity is defined as the sum of the 1 σ uncertainty spans of all model parameters. For this purpose, the (intentionally simplistic) 1 σ uncertainty ranges are defined assuming an N-D Gaussian likelihood function, so that the 1 σ interval can be constructed by taking the span of parameter space at a value of $\log \mathcal{L} \geq \max(\log \mathcal{L}) - \frac{1}{2}$ [73]. A depiction of this calculation for a single parameter is shown in Fig. 5.14.

For its application to fitting the PICO nuclear recoil calibration data, the fast burn-in method was applied in two successive stages, first with relatively large MCMC proposal step sizes, then with finer steps. This was done to (in the first stage) ensure broad exploration of the parameter space and ensure that the MCMC was not stuck

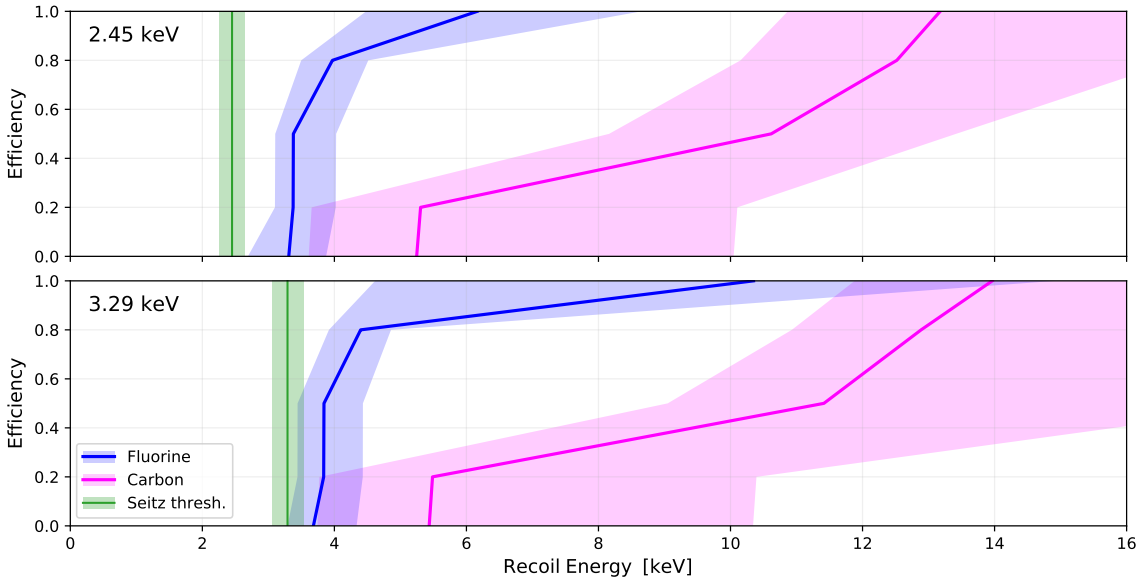


Figure 5.16: Nucleation efficiency results obtained for the PICO analysis, showing the best-fit model and $1\,\sigma$ uncertainty bands for carbon (magenta) and fluorine (blue) target atoms as a function of recoil energy, for the two thermodynamic states considered ($Q_{\text{Seitz}} = 2.45$ and $3.29\,\text{keV}$) in the upper and lower panels (also indicated by green vertical bands). Taken from figure 12 of ref. [3].

in a local optimum, and then finely sample the true global optimum. The step-size scaling parameter a – which in the Metropolis-Hastings algorithm implemented in `emcee` controls the relative scale of steps [216] – was tuned for these purposes, specifically $a = 2$ for the first stage and $a = 1$ for the second stage). The fast burn-in binning was also adapted; in the first stage, $B = 100$ bins after 10 steps per epoch were used to encourage rapid "travel" of the MCMC walkers, while in the second stage $B = 500$ bins after 5 steps per epoch resulted in thorough sampling.

Towards the end of the second stage, convergence was assessed using the metrics described above, requiring 25 successive epochs to have less than a 0.1 % change in maximum log-likelihood or $1\,\sigma$ volume. After this convergence criteria was satisfied, the MCMC was run for an additional 50 epochs. No larger changes in either convergence metric were observed in this last series of epochs, suggesting the convergence criteria were indeed sufficient [3]. The full progression of maximum log-likelihood and $1\,\sigma$ volume for the fit of the data is shown in Fig. 5.15.

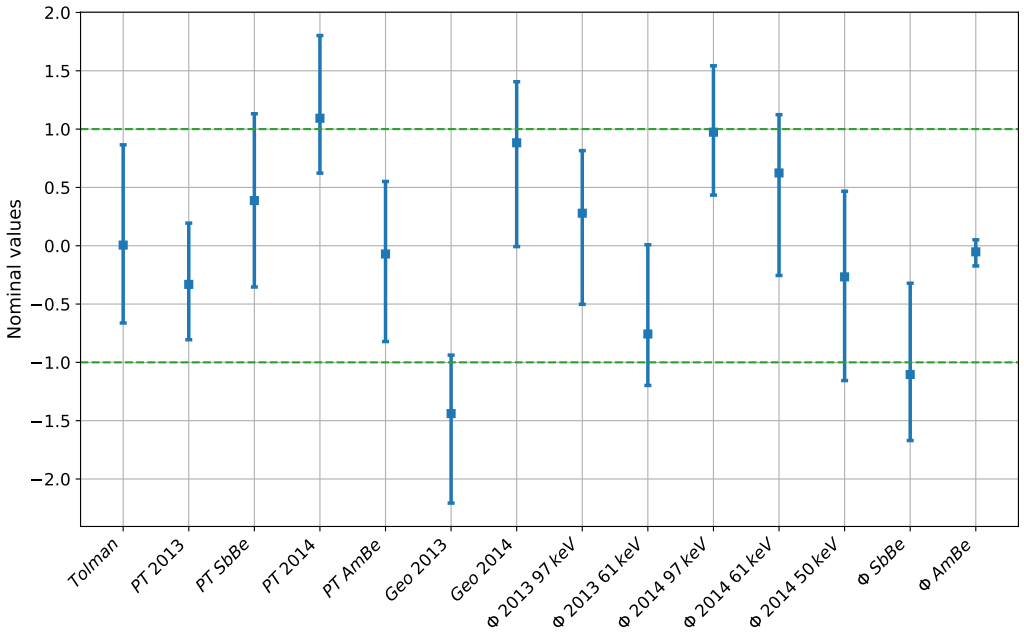


Figure 5.17: Fit results for the 14 nuisance parameters in the PICO analysis, expressed in terms of standard deviation multiples that each systematic differs from its nominal value (see section 5.4.3). The error bars shown are $1\,\sigma$ error bars derived assuming a Gaussian likelihood [73], indicating that all nuisance parameters deviated by one standard deviation or less (within $1\,\sigma$ statistical uncertainty). Taken from figure 14 of ref. [3].

5.4.5 Nucleation efficiency results

The resulting fit of the PICO neutron calibration data is shown in Fig. 5.16, which gives the best-fit nucleation efficiency curves for fluorine and carbon, for the two representative thermodynamic states considered. Also shown are $1\,\sigma$ uncertainty bands, derived approximating the likelihood as a multi-variate Gaussian, and determining the range in which $\log\mathcal{L} \geq \max(\log\mathcal{L}) - 0.5$ to give 68 % confidence intervals for each parameter of interest independently [73]. While this approach is a convenient recipe, the likelihood function in this analysis is often not Gaussian (see Figs. A.1 and 5.14 for example). Therefore, the uncertainties calculated in this way are intended only to provide an approximate visual representation of the statistical uncertainty of the fit, which is considerable. The fit results of the nuisance parameters in the model are shown in Fig. 5.17, which all ended up being within one standard deviation of

their nominal values. Another useful way of comparing the fit results to the data is to directly plot the data count rates vs. the rates predicted by the model, which is shown in Fig. 5.18. This shows good agreement between the model and data in all cases.

As expected from historical measurements of nucleation efficiency in bubble chambers (see section 5.2.1), the Seitz model underestimates the measured efficiency curves. While the fluorine threshold begins within 1 keV of Q_{Seitz} in both cases, the discrepancy for carbon is significant. None of the measured efficiency curves are step-thresholds in form, nor indeed do they resemble other generic functional forms such as a sigmoid curve (granted these forms could possibly be compatible with the large statistical uncertainties of this result). This result justifies the use of the flexible model designed for this analysis. The carbon efficiency curves naturally tend to higher energies than those of fluorine, meaning the hard constraint enforced by eq. 5.11 did not cause tension in the fit. The smaller statistical uncertainty of the fluorine efficiency curves can be understood by considering that the majority of nuclear recoils in all datasets were produced with fluorine target atoms, as can be seen in Fig. 5.11.

While the work included in this document represents the final, published analysis from ref. [3], an earlier preliminary version of this work was carried out in the interest of rapidly publishing the dark matter results of the PICO-60 detector with C_3F_8 [93]. Therefore, a chief question to be answered in this work is that of compatibility with the preliminary result used in ref. [93]. A comparison between the two nucleation efficiency models is shown in Fig. 5.19, which shows that the two results do indeed agree. The nucleation efficiency model is a direct and critical input used in calculating the WIMP sensitivity of the experiment, and therefore of the exclusion limits set on the existence of WIMP-like dark matter. The PICO-60 experiment established the most stringent (to-date) constraints on the WIMP-proton spin-dependent cross-section, as shown in Fig. 5.20.

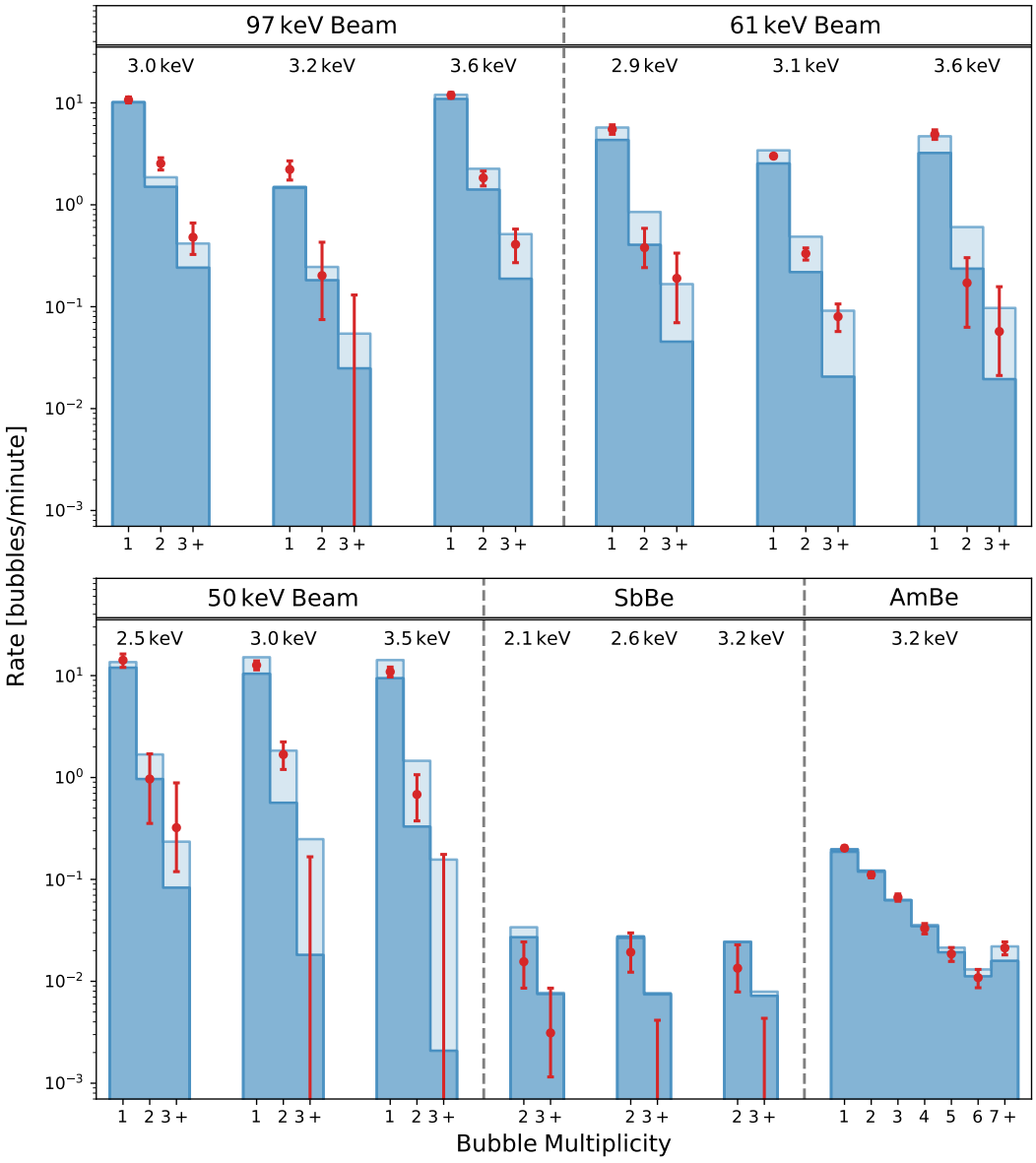


Figure 5.18: Bubble count rates (red points) listed by neutron source, Q_{Seitz}, and bubble multiplicity, with $1\,\sigma$ Feldman-Cousins error bars [240]. The blue bar charts are the bubble rates predicted by the efficiency model fit to the data, with the empty top portions of each bar representing the $1\,\sigma$ statistical uncertainty from the fit. Taken from figure 13 of ref. [3].

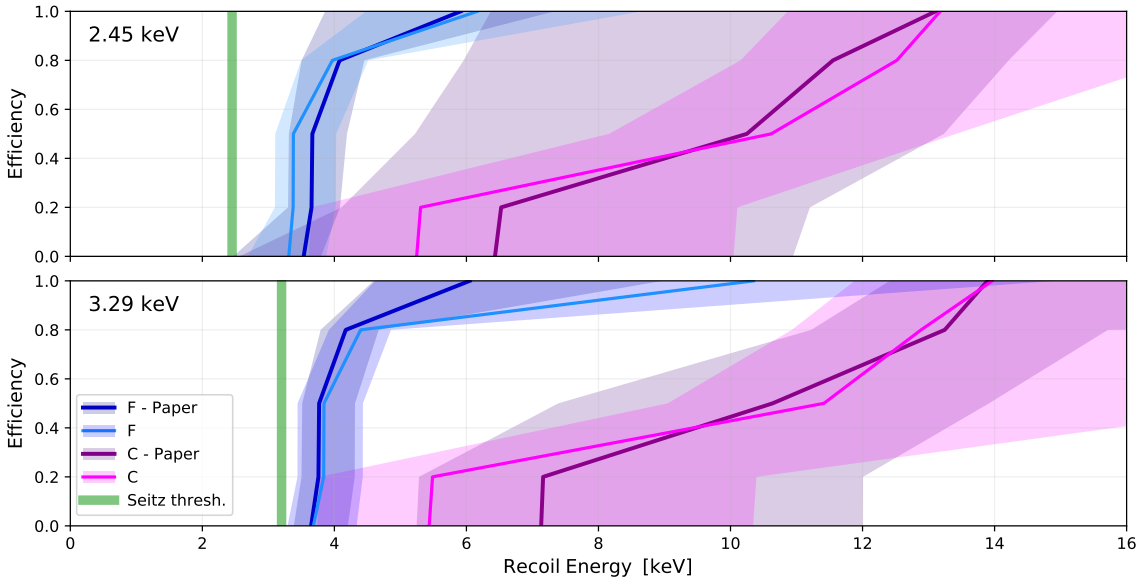


Figure 5.19: Comparison between the nucleation efficiency model obtained in this analysis (labelled "C" and "F", the same result as Fig. 5.16), and that used in the PICO-60 dark matter analysis [93] (labelled "paper"). Best-fit curves and $1\,\sigma$ uncertainty bands are shown for both results, for carbon and fluorine, at both values of Q_{Seitz} considered.

5.4.6 Application to WIMP sensitivity calculations

Past PICO dark matter analyses using C_3F_8 took conservative $1\,\sigma$ nucleation efficiency curves to produce WIMP exclusion limits [309, 343]. However, using the MCMC samples characterizing the likelihood function of the new calibration data, the full nucleation efficiency estimate can be directly incorporated into the calculation of WIMP sensitivity. This removes a redundant analysis step, and avoids relying on uncertainty limits calculated using generous assumptions (see section 5.4.5). This more robust approach was also used to calculate the dark matter result of ref. [93], which is partially reproduced here as another cross-check that the final PICO nucleation efficiency analysis agrees with that of ref. [93].

To calculate the sensitivity of the PICO-60 experiment to WIMPs (and therefore to calculate the exclusion limit claimed), the log-likelihood function for the nuclear recoil calibration data described in eq. 5.16 is recast as a likelihood function for the

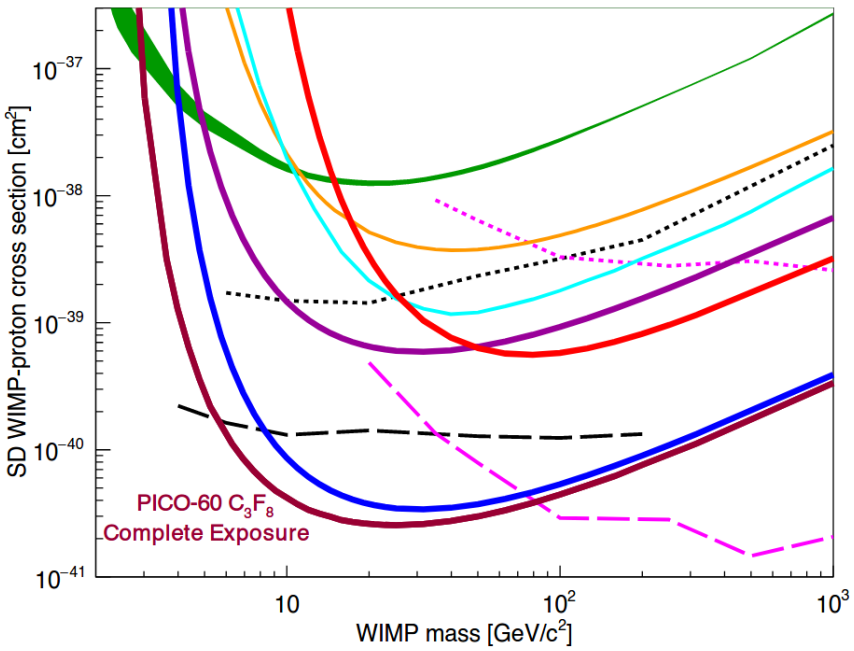


Figure 5.20: 90% CL exclusion limits on WIMP dark matter, with a spin-dependent coupling, with final the PICO-60 result with C_3F_8 indicated. Limits from PICO-60 with CF_3I (thick red), PICO-2L (thick purple), PICASSO (green), SIMPLE (orange), PandaX-II (cyan), as well as indirect detection limits from IceCube (dashed and dotted pink) and SuperK (dashed and dotted black) [288, 319, 337–342]. Taken from figure 7 of ref. [93].

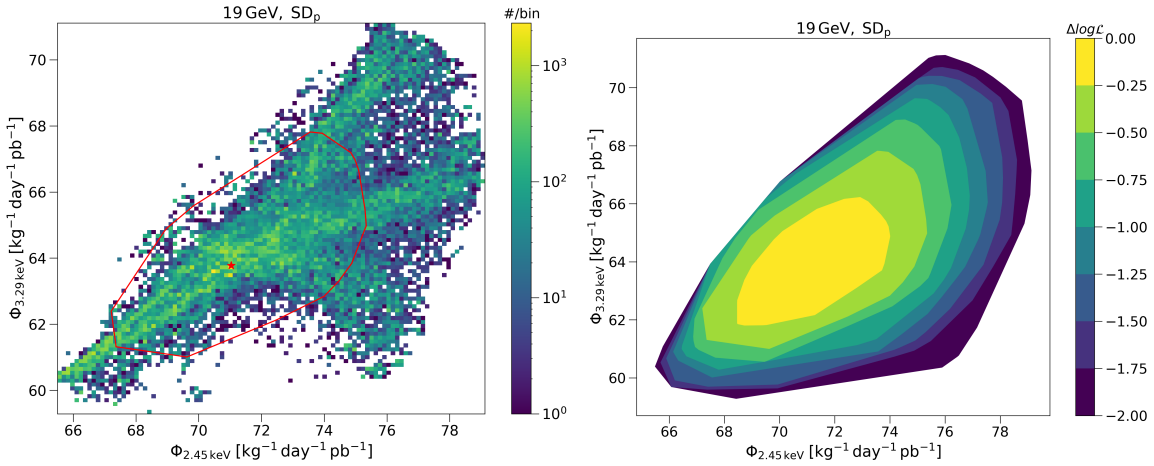


Figure 5.21: Left: 2D histogram of MCMC samples of WIMP interaction rates for a 19 GeV/c^2 WIMP with SDp coupling, with the interaction rate for $Q_{\mathrm{Seitz}} = 3.29\,\mathrm{keV}$ vs. $Q_{\mathrm{Seitz}} = 2.45\,\mathrm{keV}$. Right: the MCMC samples are used to construct closed contours at specified levels of $\Delta \log \mathcal{L}$ to represent the likelihood function.

expected rate of WIMP interactions Φ for a given dark matter particle mass M_χ and coupling $\mathcal{O}$, spin-independent (SI) or spin-dependent on proton (SDp). For the recent PICO-60 results [93], physics data taken at thermodynamic thresholds of 3.29 keV and 2.45 keV is combined to form a WIMP exclusion limit, so correlated uncertainties in the bubble nucleation efficiencies at these two thresholds must also be characterized. To achieve this, the fast burn-in MCMC procedure of section 5.4.4/appendix A is used to explore the likelihood function projected in terms of the WIMP event rate

$$\Phi_{M_\chi,Q,\mathcal{O}} = \sum_{s=\mathrm{C,F}} \int_0^\infty R_{s,M_\chi,\mathcal{O}}(E_R) \cdot \epsilon_s(E_R, Q)\, dE_r, \tag{5.19}$$

where $R_{s,M_\chi,\mathcal{O}}$ is the differential recoil spectrum (normalized by total WIMP-nucleon cross-section) for WIMPs with mass M_χ and coupling $\mathcal{O}$ on recoil species s. For the PICO-60 analysis, the WIMP masses that were considered were $M_\chi \in \{3\ \mathrm{GeV/c^2}, 10\ \mathrm{GeV/c^2}, 19\ \mathrm{GeV/c^2}, 50\ \mathrm{GeV/c^2}\}$, and $\mathcal{O} \in \{\mathrm{SI}, \mathrm{SDp}\}$, calculating recoil spectra for a standard halo model as described in [58] with $\rho = 0.3\,\mathrm{GeV/cm^3}$, $v_{\mathrm{esc}} = 544\,\mathrm{km/s}$, $v_{\mathrm{Earth}} = 232\,\mathrm{km/s}$, and $v_0 = 220\,\mathrm{km/s}$. Calculations are performed following [58, 66] — the python package *dmdd* was used specifically for fluorine/SDp coupling [344].

In addition to the set of 16 $\{\Phi_{M_\chi,Q,\mathcal{O}}\}$ parameters specified above, an additional 16 parameters were added to the MCMC sampling scheme, constructed from linear combinations of Φ's at different values of Q:

$$\Phi_{M_\chi,\pm,\mathcal{O}} = \pm\Phi_{M_\chi,2.45\mathrm{keV},\mathcal{O}} \pm \Phi_{M_\chi,3.29\mathrm{keV},\mathcal{O}}, \tag{5.20}$$

This was done to more efficiently explore correlated uncertainties at different thermodynamic states. The reseeding of new epochs is also changed slightly in this analysis compared to the algorithm described in appendix A, by binning the likelihood (40 bins from the best-fit to $\Delta \log \mathcal{L} = -1$) rather than in the projected dimensions, and taking the high and low extreme values of the projected parameter in each likelihood bin to seed the next epoch. After running the fast burn-in MCMC on this parameter construction, the results can then be used to plot the 2D posterior likelihood of WIMP

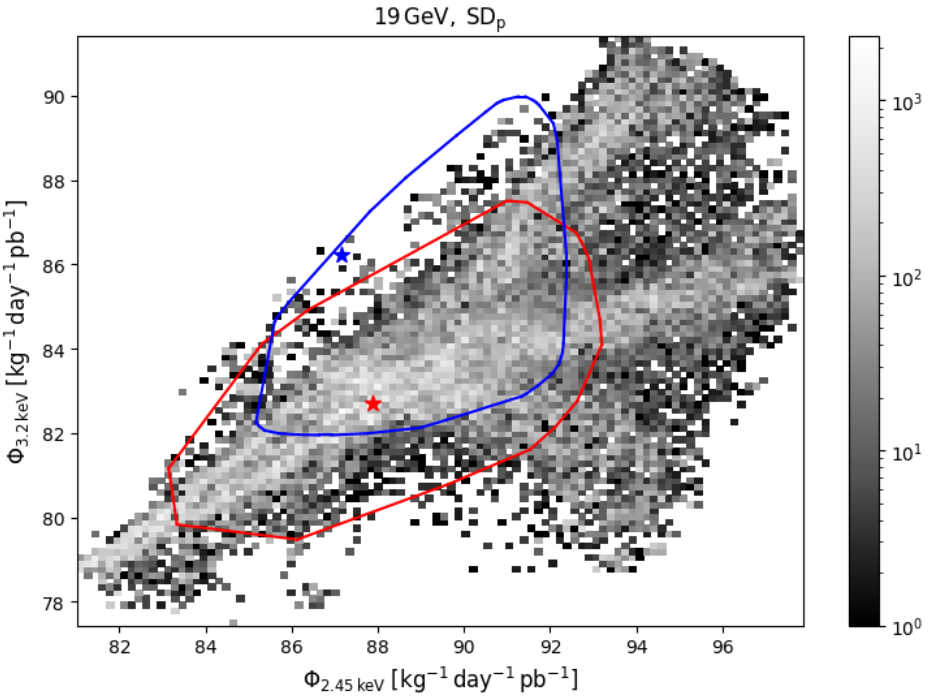


Figure 5.22: MCMC samples of WIMP interaction rates for a 19 GeV/c^2 WIMP with SDp coupling (grey 2D histogram) as well as the 1 σ contour and best-fit (red curve and star) obtained in this analysis, compared to that of ref. [277] (blue curve and star).

interaction rates for a given WIMP mass and coupling. An example of the MCMC samples for $M_\chi = 19\,\mathrm{GeV/c^2}$ and $\mathcal{O} = \mathrm{SDp}$ is shown in fig. 5.21 (left). Bounding curves at different levels of $\Delta \log \mathcal{L}$ can be created to form a more regular posterior likelihood (Fig. 5.21 right). The result obtained through this approach is compatible with the results of refs. [93, 277], as shown in Fig. 5.22.

5.4.7 Parametric Monte Carlo study

A further step taken to validate the results presented in section 5.4.5 was a study of simulated datasets. Using the best-fit model shown in Fig. 5.16, 25 Monte Carlo datasets were generated – hence this can be considered a "parametric MC study". Examples of this simulated data are shown in Fig. 5.23. Specifically, the expected number of counts for each experiment/bubble multiplicity was calculated using the best-fit model, from which a random Poisson number of counts was drawn. These

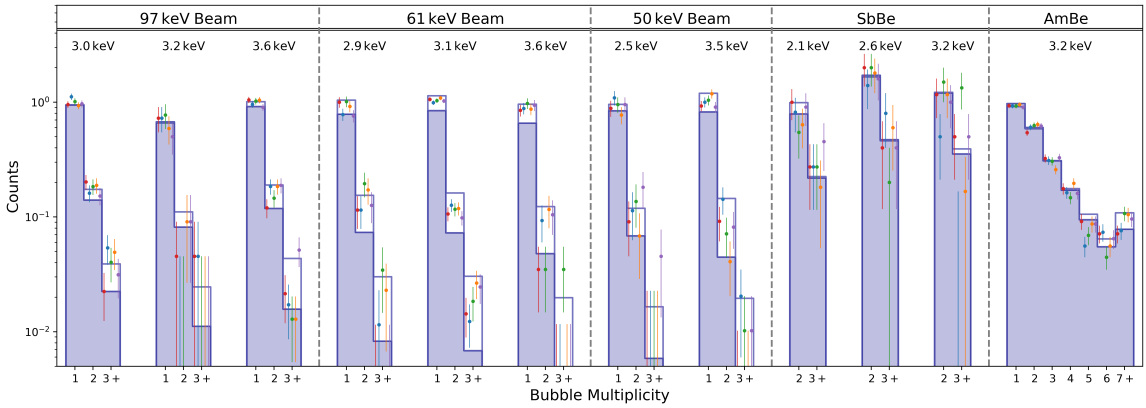


Figure 5.23: Best-fit nucleation efficiency model (blue bars), with 5 examples of simulated datasets (coloured points).

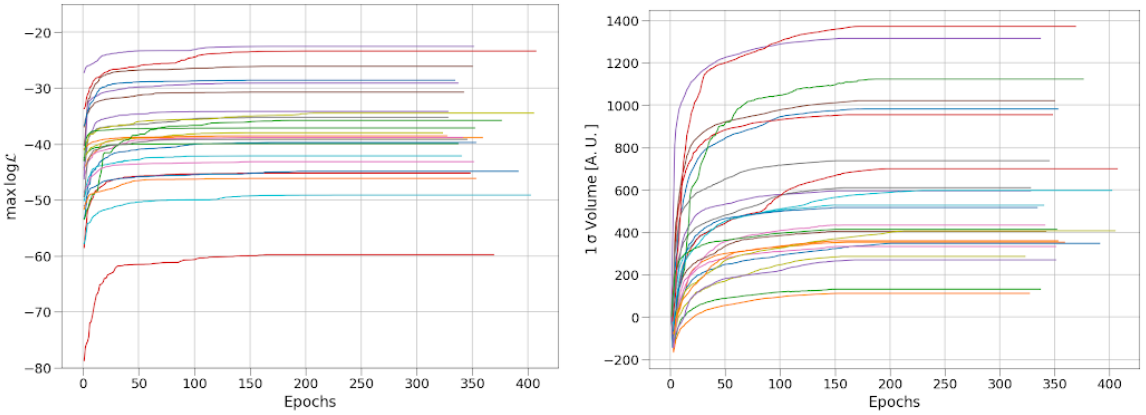


Figure 5.24: The maximum log-likelihood (left) and $1\,\sigma$ volume (right) of the fits of all 25 simulated PICO nuclear recoil calibration datasets, as a function of MCMC epoch.

MC datasets were fit using the same MCMC procedure described in section 5.4.4. One result of this study is confirmation that the convergence criteria used in this analysis were sufficient, as all fits of MC data converged to the true model with similar uncertainties in approximately the same number of epochs as the fit of the real data. The convergence metrics of all MC datasets as a function of MCMC epoch are shown in Fig. 5.24.

Another output of this study is an estimation of the goodness-of-fit of the result obtained in section 5.4.5. As a test statistic, the maximum log-likelihood value obtained in any given fit can be used to estimate a "χ^2" value, as $\chi^2 = -2\log\mathcal{L}$. Normally, one

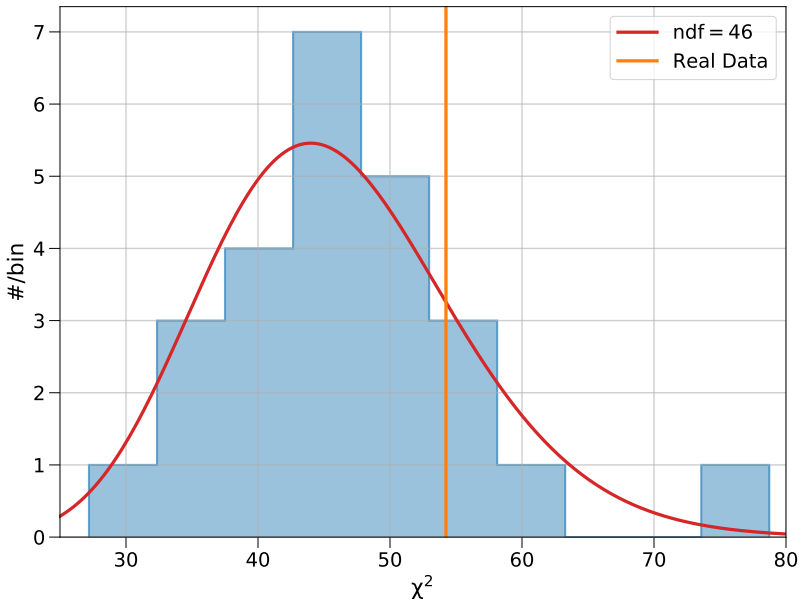


Figure 5.25: Distribution of χ^2 obtained from fits of MC datasets (blue histogram), as well as a χ^2 distribution fit to this data (orange curve) having 46 degrees of freedom. The value of χ^2 for the fit of the real data is indicated by a red line, corresponding to a p-value of 0.19. Taken from figure 16 of [3].

could calculate a p-value using this statistic given the number of degrees of freedom of the fit (the number of data points less the number of free parameters). However, in this case, there are many complicated correlations between the model parameters, resulting in a lower effective number of degrees of freedom. One can instead rely on the distribution of χ^2 obtained for the fits of the MC data – which necessarily represent a fit that agrees with the data (assuming the MCMC fits converged) – as their input model is known. Using this distribution of the test-statistic χ^2 shown in Fig. 5.25, the effective number of degrees of freedom is found to be 46. With this result, a p-value of 0.19 is obtained for the fit of the real data. This confirms that the model used in this analysis adequately represents the data and that they are in reasonable agreement. Further applications of this MC study are discussed ahead in section 5.7.

5.5 Xe-SBC nuclear recoil study

As a rare-search experiment hoping to detect nuclear recoils from WIMP-like dark matter and from coherent neutrino scattering, the SBC (Scintillating Bubble Chamber) collaboration requires nucleation efficiency calibrations for their bubble chambers. To date, the only liquid noble bubble chamber operated by the collaboration was a small liquid xenon (LXe) vessel (introduced in section 5.1.2), with initial results published in ref. [314]. This included the first demonstration of coincident scintillation and nucleation in a bubble chamber (measured with a ^{252}Cf neutron source), as well as a first estimate of the electronic recoil nucleation rate in LXe. With this same chamber, extensive neutron calibration data was collected from various sources, and in many thermodynamic operating conditions, allowing for an analysis of the chamber's nuclear recoil nucleation threshold. This analysis was carried out similarly to the PICO analysis presented in sections 5.3 and 5.4.

5.5.1 Neutron calibration data

The LXe bubble chamber (or Xe-SBC) – shown in Fig. 5.26 – held 30 g of pure liquid xenon. The "right side up" design was used (see section 5.1.1) albeit with a blank metal flange rather than an interior jar. A temperature gradient was created to maintain the upper volume in a superheated state (-55°C to -38°C), while keeping the lower extremities in a stable liquid state (at -100°C). This was achieved by placing the entire vessel in a vacuum cryostat, and cooling the entire system with a cold finger from a liquid nitrogen dewar. Separate heating coils maintained the two different regions at their intended temperatures to within 0.1°C [314]. Compression after nucleation was achieved via a hydraulically-driven bellows system.

The detector was instrumented with a single R6834 Hamamatsu photomultiplier tube (PMT) positioned above the LXe volume, two piezoelectric transducers to record acoustic signals, and a camera to image bubbles (through a sapphire window in the

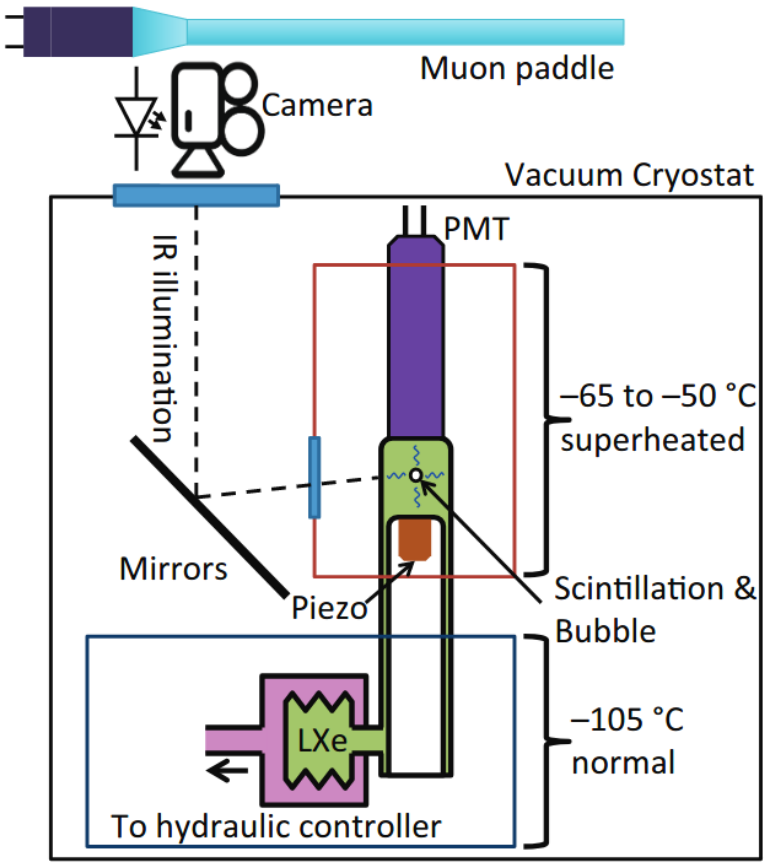


Figure 5.26: Schematic of the LXe scintillating bubble chamber, with a 30 g superheated target volume, housed in a vacuum cryostat. Instrumentation included a single PMT, piezoelectric transducer, and IR-illuminated camera. Taken from figure 1 of ref. [314].

cryostat). Infrared LEDs were flashed to provide illumination for the camera images, which were recorded every 10 ms. A series of mirrors produced a stereoscopic effect in the camera images for 3D bubble position reconstruction. Additionally, a plastic scintillator paddle was positioned above the entire experiment as a muon veto. The PMT was calibrated using a blue LED situated inside the cryostat, as well as exposure to a ^{57}Co source producing 122 keV gamma-rays [345]. The latter peaked at a PMT signal of approximately 4 photoelectrons, indicating a light collection efficiency of $\sim$ 0.05%. Assuming the Lindhard model [178] for the scintillation yield of nuclear recoils in LXe, this means a recoil energy of 94 keV$_{\mathrm{nr}}$ corresponds to a single photoelectron [345].

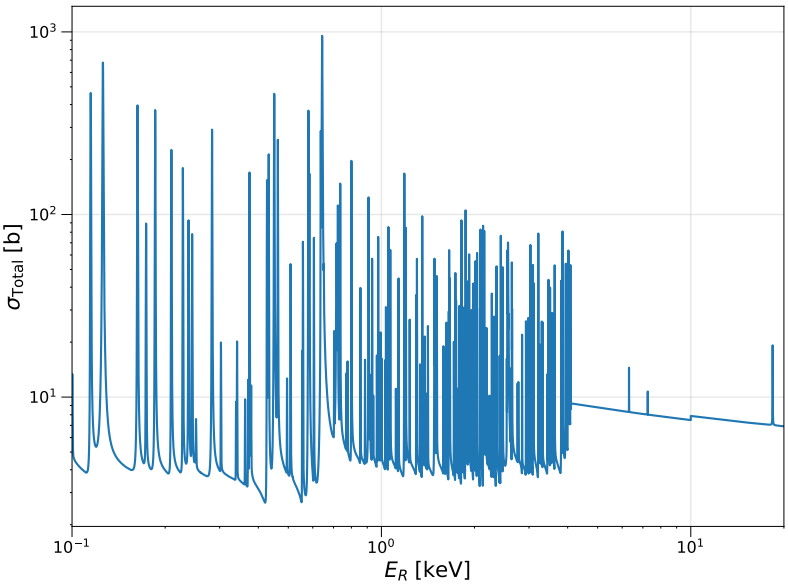


Figure 5.27: Total neutron-xenon cross-section, assuming natural isotopic abundance of xenon [331, 346].

Source	^{252}Cf	Bi-Be	Y-Be
Q_{Seitz}	0.9, 1.19, 1.89, 2.06	1.14	0.9, 0.97, 1.04, 1.12, 1.14, 1.33, 1.44, 1.56
Livetime	106, 1060, 71.2, 841	193	22.8, 32.0, 31.6, 34.2, 52.6, 67.9, 68.3, 76.5

Table 5.4: Xe-SBC neutron calibration datasets used in this analysis, listing the Seitz thresholds (in keV) and corresponding livetimes (in minutes) at which data was collected, for all calibration sources.

The neutron calibrations were performed using a ^{252}Cf source, as well as Y-Be and Bi-Be photoneutron sources [327]. The former is a spontaneous-fission source producing a broad, fast neutron spectrum, peaking around 2 MeV [347]. The Y-Be sources produces nearly monoenergetic 152 keV neutrons (corresponding to a maximum recoil energy of 4.9 keV), at a rate of approximately 800 neutrons per second [345]. The Bi-Be source, on the other hand, was estimated to only produce < 3 neutrons per second, and with a lower kinetic energy of 91 keV [345]. The total cross-section of neutrons with xenon (assuming natural isotopic abundance [346]) is shown in Fig. 5.27, which shows a dense clustering of resonances for $\mathcal{O}$(keV) neutrons.

With the calibration sources described above, data was collected in the Xe-SBC

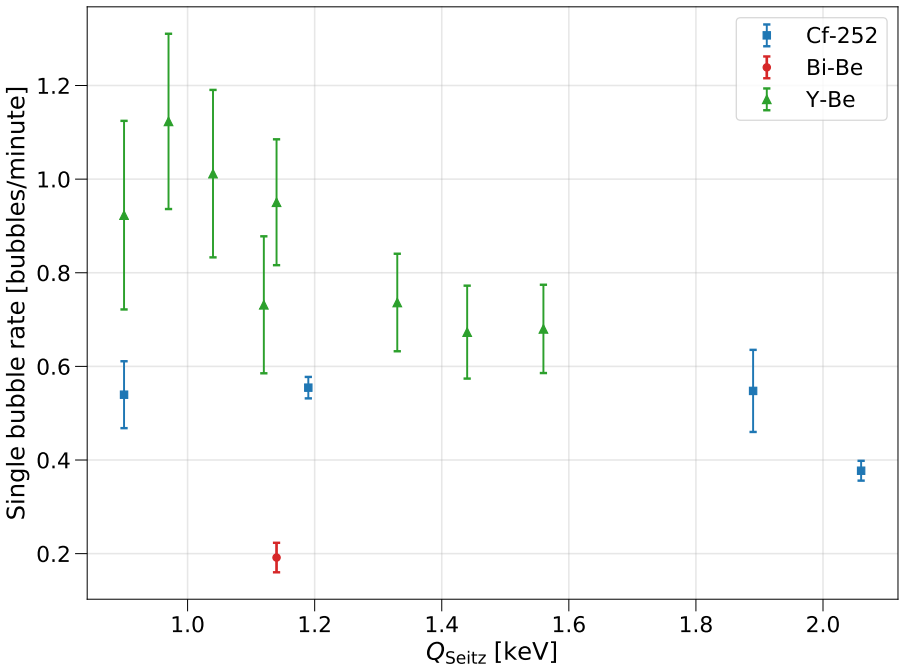


Figure 5.28: Single bubble event count rates (with Poisson error bars) as a function of Q_{Seitz}, for all Xe-SBC calibration sources.

chamber at a wide range of thermodynamic stats, ranging from $Q_{\mathrm{Seitz}} = 0.9\,\mathrm{keV}$ to 2.06 keV. This data is summarized in Table 5.4. Bubble chamber events were triggered by the camera system, based on a change in entropy from frame to frame (recorded at 100 fps). Coincident acoustic data and PMT waveforms corresponding to the bubble were also collected - note that the scintillation signal is recorded before the bubble is visible to the camera, but a coincident window was defined based on the speed of sound in LXe [314, 348]. The muon detector above the detector vetoed any triggers in coincidence with events in that channel. A camera trigger also starts the recompression cycle, returning the detector to operating condition; only when the vessel is restored to within 0.5 psi of its' target pressure is it considered live again. As with the PICO analysis, the bubble image processing was done by hand-scanning the events. The work of preparing and formatting the raw data was done by M. Bressler; more details can be found in ref. [345].

The collected data is grouped into events with a single bubble, or two or more

bubbles. The rate of single bubble events (number of counts divided by total livetime) is presented in Fig. 5.28 as a function of Q_{Seitz}. This reveals the expected trend of decreasing bubble rates as the thermodynamic threshold increases, for all calibration sources. Additionally, background data was taken in every experimental setup (source and thermodynamic configuration) near in time to signal dataset. In the case of the Y-Be source, an aluminum disk was inserted in the place of the Be target disk, so that the gamma background produced by the ^{88}Y source itself could still be measured (the aluminum similarly attenuates the gamma-rays).

5.5.2 Coincident scintillation spectrum

One unique ability of liquid noble bubble chambers compared to their fluorocarbon counterparts is the possibility of event-by-event energy reconstruction, using the scintillation signal produced. In practice, however, the scintillation signal from nuclear recoils is very small compared to that from electronic recoil backgrounds [314, 345]. Additionally, the single PMT configuration of the Xe-SBC chamber does not yield high light-collection efficiency. These issues rule out energy reconstruction for the data taken with this detector. Fortunately, the larger scintillation signal produced by electronic recoil background events is easily recorded by the PMT, meaning this channel can be used effectively for background suppression.

In the case of ^{252}Cf in particular – for which the majority of recoils induced fall below the 94 $\mathrm{keV_{nr}}$ light collection threshold – most events will produce none or single photoelectrons. However, the measured background data contains mainly events with greater than 1000 photoelectrons, as can be seen in Fig. 5.29. Therefore, a cut on the PMT signal was made (at 512.5 photoelectrons, corresponding to a bin boundary in Fig. 5.29) to suppress electronic recoil background events. The single-bubble background rate measured for all sources – including the PMT cut for ^{252}Cf – is shown in Fig. 5.30 as a function of Q_{Seitz}.

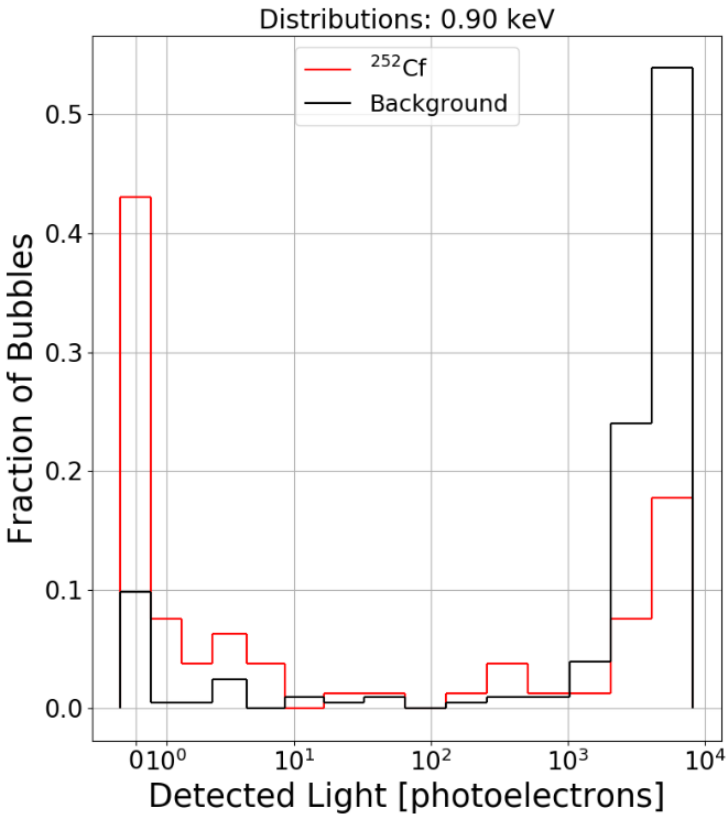


Figure 5.29: Fraction of bubble events distributed by their PMT signal (in photoelectrons, with irregular binning) collected with ^{252}Cf at $Q_{\text{Seitz}} = 0.90\,\text{keV}$, as well as background data taken in the same conditions afterwards. Taken from figure 5.25b of ref. [345].

5.5.3 Calibration simulations

As with the PICO nucleation efficiency analysis, full simulations of the neutron calibration experiments are needed to anticipate the expected recoil energy spectrum, and therefore to calculate the expected event rate for a given nucleation efficiency model. For the Xe-SBC analysis, this was done exclusively using GEANT4 [204]. The work of performing the simulation presented in this subsection was mostly done by A. Zuñiga of the SBC collaboration. The final simulation spectra produced are shown in Fig. 5.31.

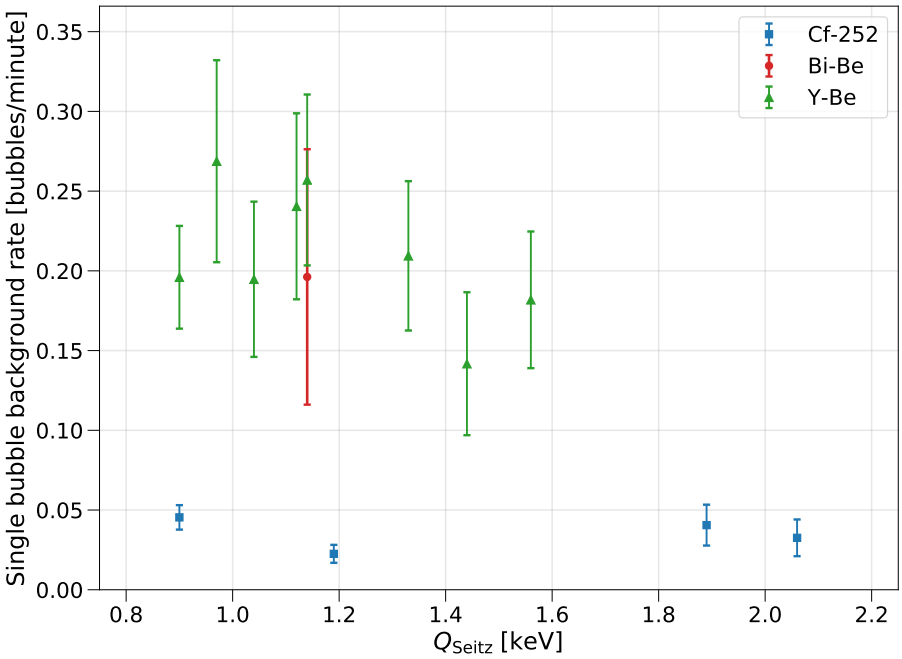


Figure 5.30: Rate of single bubble events (with Poisson error bars) from background data taken after neutron source data was collected, as a function of Q_{Seitz}.

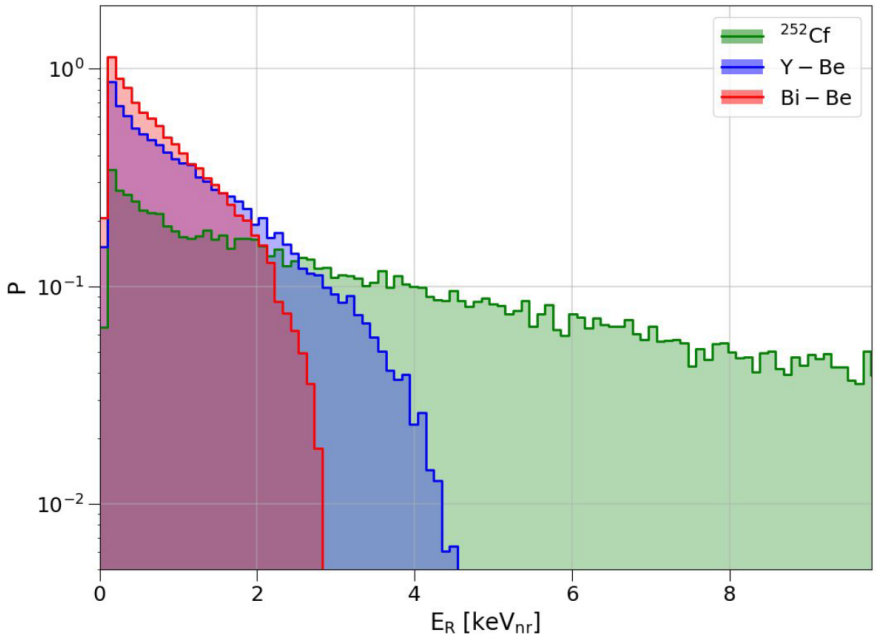


Figure 5.31: Simulated nuclear recoil spectra from different calibration sources with the Xe-SBC detector, produced using GEANT4 [204].

5.6 Xe-SBC nucleation efficiency analysis

Similarly to the PICO analysis, the nucleation efficiency of nuclear recoils in liquid xenon can be obtained from the neutron calibration described in the previous section by applying a given nucleation model to the simulated recoil spectrum and comparing the predicted bubble rate to the observed data (eq. 5.5). In this case, however, the efficiency model $\epsilon(E_R, Q_{\text{Seitz}})$ is slightly simpler due to the monatomic target material. In this analysis, a joint fit is performed using all calibration data listed in Table 5.4.

5.6.1 Nucleation efficiency model

The nucleation efficiency model used in this analysis was largely similar to that used in the PICO analysis (see section 5.4.1), save the differences described below. Efficiency curves for a given threshold Q_{Seitz} are defined as piecewise linear functions $\epsilon(E_R, Q)$, with the segments defined by fixed efficiency knots at $\{0, 0.2, 0.5, 0.8, 1\}$ and the floating nuclear recoil energies E_R at which those efficiency knots are reached, $\{x\}_Q$.

The data gathered in the Xe-SBC calibration campaign include a more continuous set of Q_{Seitz} values compared to the PICO campaign, as can be seen in Table 5.4. Therefore, a different Q interpolation scheme was developed for this analysis. The model includes efficiency curves at linearly-spaced values of Q, beginning (and ending) at the lowest (and highest) values of Q for all datasets, which were 0.9 keV and 2.06 keV. The number of curves in total varied from 2 (just the highest and lowest values of Q) to 5, which would include $Q = 0.9\,\text{keV}$, 1.19 keV, 1.48 keV, 1.77 keV, and 2.06 keV.

To interpolate in between these efficiency curves, a mesh of some kind is needed to define a bilinear interpolation. One option considered was a triangular mesh as a function of E_R and Q to define a planar interpolation on the mesh panels for efficiency. The $\{x\}_Q$ points would not form a regular grid, so an algorithm such as a Delaunay triangulation would be required, which seeks to maximize the smallest interior angle

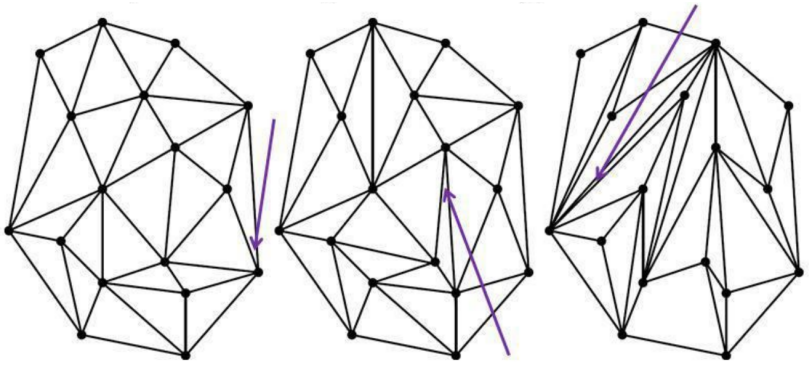


Figure 5.32: An example of various triangular meshes defined on irregularly-spaced points. A Delaunay triangulation (left) seeks to maximize the smallest internal angle of all triangles in a mesh, which is indicated by an arrow in each mesh example. Image taken from ref. [349].

of all triangles formed in the mesh, creating what is as close as possible to a regular mesh on irregular points [350]. A generic example of this is shown in Fig. 5.32.

However, if the mesh is instead defined on the parameters Q and efficiency ϵ, a regular grid of points can be recovered. The nuclear recoil energies for values of Q in between the explicitly defined ones are then calculated using the interpolation defined on this mesh. In this case, the mesh was defined to be all right-angle triangles (which corresponds to a Delaunay mesh for this regular grid), with each quadrant of points split into two triangles with their shared hypotenuse going from the lower Q/efficiency point to the high Q/efficiency point. This scheme is depicted in Fig. 5.33.

The nucleation efficiency model is also subject to the same constraints imposed on the PICO analysis, namely that the efficiency curves increase monotonically as a function of recoil energy (eq. 5.8), and are bounded by their corresponding value of Q_{Seitz} (eq. 5.10). The specific construction of the triangular mesh described above guarantees that this is the case. A proof of this is given in Appendix B. Note that this would not necessarily be the case for a mesh defined differently, i.e. if the hypotenuse of each cell was flipped by 90° in Q/ϵ.

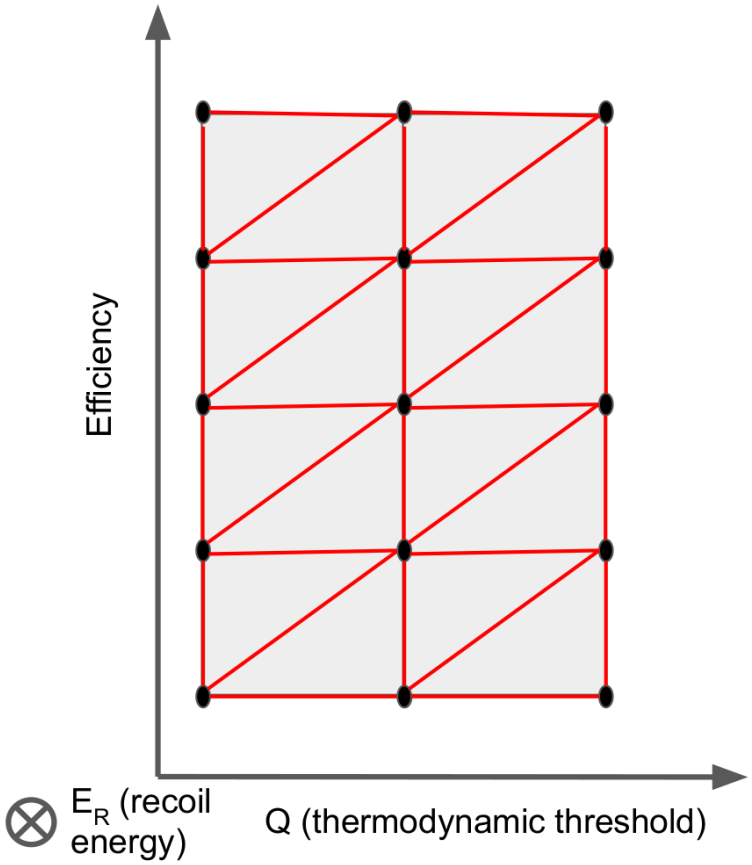


Figure 5.33: Cartoon depiction of the triangular mesh approach used in the Xe-SBC analysis, defined on the regular grid of efficiency and Q_{Seitz} points, with the nuclear recoil energy scale being the third dimension pointing out of the page.

5.6.2 Background treatment and systematic uncertainties

As discussed in section 5.5.1, background data was collected corresponding to every source by taking data in the same configuration as the calibration, near in time, simply without the source present (or by replacing the Be-O disk with an aluminum one in the case of the Y-Be calibration). Unlike with the PICO analysis – in which the best-fit background rate was subtracted from the background plus signal rate – the background rate is treated as an independent nuisance parameter for every experiment (source/Q_{Seitz}) and bubble multiplicity, resulting in 26 such nuisance parameters. This large addition of dimensions to the nucleation model would be a computational challenge to fit, so the background rate nuisance parameters are profiled-over analytically

to avoid needing to sample these parameters in the MCMC fit.

Note that the following derivation of this analytical approach to the background nuisance parameters was prompted by the work of C.E. Dahl on this topic (SBC Collaboration) [351]. In this scenario, there is one measurement of the background rate alone (which observes n_b counts), and another with the source measuring the signal and background rate together (observing n_s counts). The expected number of counts in the signal plus background data, μ_s, is a combination of the expected number of background events μ_b and the expected number of signal-only events ν, the parameter of interest in this case. Because the two datasets may not have the same exposure necessarily, the background rate is adjusted by the ratio of the exposures of the two datasets r:

$$\mu_s = \nu + r\mu_b \tag{5.21}$$

A joint Poisson likelihood can be formed for both measurements:

$$\mathcal{L}_{\text{Bkgd}} = \frac{e^{-\mu_s}\mu_s^{n_s}}{n_s!} \cdot \frac{e^{-\mu_b}\mu_b^{n_b}}{n_b!} \tag{5.22}$$

The log-likelihood is then simplified as (using Stirling's approximation $\log n! \approx n \log n - n$ [73]):

$$\mathcal{L}_{\text{Bkgd}} = -\mu_s + n_s \left(1 + \log \frac{\mu_s}{n_s}\right) - \mu_b + n_b \left(1 + \log \frac{\mu_b}{n_b}\right) \tag{5.23}$$

Note that in the case where $n_s = 0$ or $n_b = 0$, the terms in the above equation starting with the corresponding n_s or n_b evaluate to 0, to avoid the undefined evaluation of the logarithms. Next, we can substitute in eq. 5.21 so we are dealing with the parameter of interest:

$$\mathcal{L}_{\text{Bkgd}} = -\nu - r\mu_b + n_s \left(1 + \log \frac{\nu + r\mu_b}{n_s}\right) - \mu_b + n_b \left(1 + \log \frac{\mu_b}{n_b}\right) \tag{5.24}$$

and then differentiate the likelihood with respect to μ_b and set it equal to 0, to find the optimum value of μ_b for a given value of ν:

$$\begin{aligned} \frac{\partial \mathcal{L}_{\text{Bkgd}}}{\partial \mu_b} &= -r + \frac{r n_s}{\nu + r\mu_b} - 1 + \frac{n_b}{\mu_b} = 0 \\ \implies \mu_b(\nu + r\mu_b) &= -r\mu_b(\nu + r\mu_b) + r n_s \mu_b + n_b(\nu + r\mu_b) \\ \implies 0 &= (r+1)\mu_b^2 + \left(\frac{\nu(r+1)}{r} - n_s - n_b\right)\mu_b - \frac{n_b \nu}{r} \end{aligned} \tag{5.25}$$

Thus we have a quadratic equation for the optimum in terms of μ_b, for which we can obtain solutions for the profiled value of $\mu_b = \hat{\mu_b}$ using the quadratic equation (one of which will be positive):

$$\hat{\mu_b} = n_s + n_b - \frac{\nu(r+1)}{r} \pm \frac{\sqrt{\left(\frac{\nu(r+1)}{r} - n_s - n_b\right)^2 + 4\frac{(r+1)}{r} n_b \nu}}{2(r+1)} \tag{5.26}$$

Using this result, the profiled (optimized) value of μ_b can be obtained for a given predicted ν for every data point, greatly saving on computational resources when performing the fit of the data.

Other systematic uncertainties are also included in this analysis, however in this case the greater control over the pressure/temperature of the bubble chamber compared to the PICO campaign allowed uncertainty in Q_{Seitz} to be ignored. Further, because only one chamber was used, with only one calibration data-taking setup for each source, all systematic uncertainty on the source strengths and experiment geometries was combined into one systematic per calibration source, and one common systematic for all three calibrations. The latter is to account for uncertainty in the total mass of LXe in the bubble chamber. These uncertainties are listed in Table 5.5 [345]. As with the PICO nucleation efficiency study, these systematics are incorporated as log-normally distributed scaling factors in terms of multiples of the listed $1\,\sigma$ uncertainties.

Source	Uncertainty
^{252}Cf	1.2%
Bi-Be	50%
Y-Be	7%
All	10%

Table 5.5: Systematic uncertainties on the total neutron-LXe exposure for the Xe-SBC calibration experiments [345].

5.6.3 Model fitting

The likelihood function constructed for this model follows the same structure as that of the PICO analysis (see section 5.4.2), namely a Poisson log-likelihood summed over all experiments i and bubble multiplicities j, with expected counts $\nu_{i,j}$ and observed counts $k_{i,j}$.

$$\log \mathcal{L} = \sum_i \sum_j \left[-\nu_{i,j} + k_{i,j} + k_{i,j} \log \left(\frac{\nu_{i,j}}{k_{i,j}} \right) \right] \tag{5.27}$$

The expected counts for a given model are calculated by adding the profiled background rate (for that experiment/multiplicity as derived in section 5.6.2) to eq. 5.5, and multiplying by the corresponding livetime (see Table 5.4):

$$\nu_{i,j} = \left(R_{i,j} + \hat{\mu}_{b i,j} \right) \times \text{Livetime}_{i,j} \tag{5.28}$$

The nuisance parameters representing systematic uncertainties are included in the likelihood exactly as described in section 5.4.3, resulting in an extra term being added to the likelihood function given by eq. 5.18.

Fitting the model to the data was done with the same custom MCMC algorithm and procedure described in appendix A; the "fast burn-in" MCMC method. This again was done in two stages (see section 5.4.4), first with coarse binning (250 bins for each parameter) and large MCMC walker step sizes (with $a = 2$), and then with finer binning (500 bins) and smaller steps ($a = 1$) to map the likelihood function in

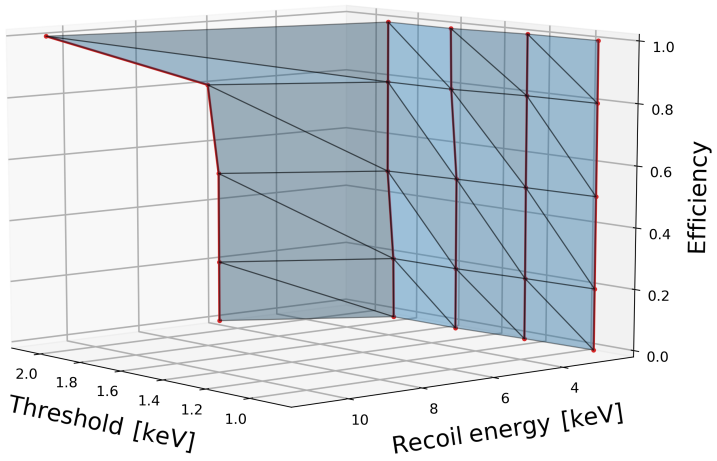


Figure 5.34: Best-fit LXe nucleation efficiency model with 4 threshold setpoints.

greater detail. The same convergence criteria based on the progression of $\log \mathcal{L}_{\max}$ and the "$1\,\sigma$ volume" of the likelihood function were used as well.

5.6.4 Model selection

As mentioned in section 5.6.1, the number of threshold setpoints in the nucleation efficiency model is potentially tunable (2-5 were considered for this analysis), with more threshold setpoints yielding a more flexible model. The trade-off with increasing the number of threshold setpoints (other than leading a to higher-dimensional, computationally-expensive model) is the potential for over-fitting. To determine the ideal number of threshold setpoints, models with 2, 3, 4, and 5 setpoints were fit to the data (as described in section 5.6.3).

A model-comparison analysis can then be used to determine the optimal trade-off between too many and too few threshold setpoints. Normally one could simply compute the likelihood ratio for two models (using the maximum log-likelihood found for each). However, in this case, the models with different numbers of threshold setpoints are not nested models, i.e. interpolations defined on meshes with 2-4 setpoints is not

# of setpoints	# of parameters	AIC
2	39	143.4
2	44	140.6
2	49	147.3
2	54	156.3

Table 5.6: Table of AIC values for LXe nucleation efficiency models with different numbers of threshold setpoints.

necessarily representable by a mesh with 5 threshold setpoints. A model selection criteria that does not require nested models for comparison is the Akaike Information Criterion (AIC) [352, 353]:

$$\mathrm{AIC} = 2k - 2\log\mathcal{L}_{\max} \tag{5.29}$$

where k is the number of free parameters in the model. In this case, k is the number of points in all of the piecewise linear functions $\{x\}_Q$ ($5 \times$ # of threshold setpoints), plus the 3 exposure systematic nuisance parameters, as well as the 26 background rate nuisance parameters. Minimizing this quantity optimizes the balance between a model that fits the data well, while also adding a penalty for having additional parameters.

Models with 2–5 threshold setpoints were fit to the data as described in section 5.6.3; an example with 4 threshold setpoints is shown in Fig. 5.34. The AIC values for the best-fit models are presented in Table 5.6, which indicates that 3 threshold setpoints are the ideal model definition. Therefore, this is used in the rest of the analysis going forward.

5.6.5 Nucleation efficiency results

The fit obtained using the optimally-selected model with 3 threshold setpoints, and the MCMC procedure described in section 5.6.3, is the main result of this analysis. The progress of maximum likelihood and $1\,\sigma$ volume used to confirm the convergence

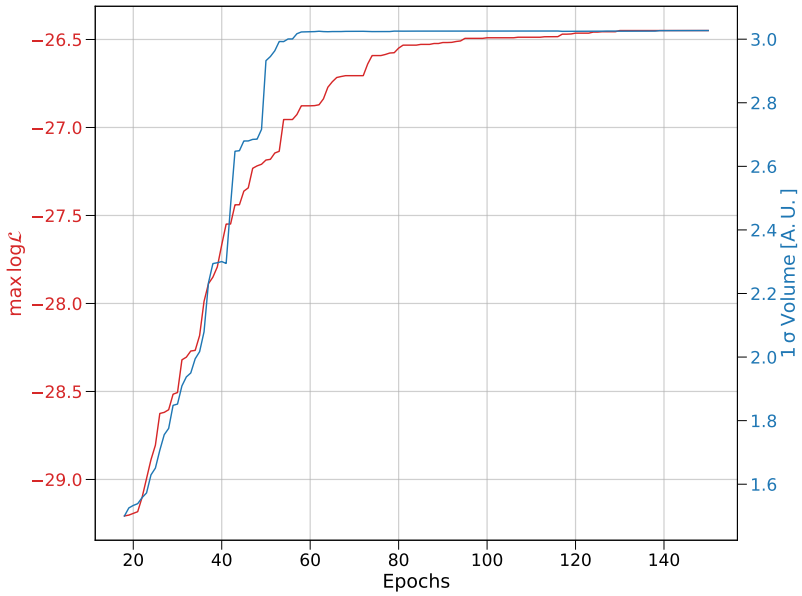


Figure 5.35: Maximum log-likelihood and $1\,\sigma$ volume vs. MCMC epoch for the fit of the Xe-SBC calibration data, using a model with 3 threshold setpoints.

of the MCMC is shown for this fit in Fig. 5.35. The resulting best-fit and $1\,\sigma$ uncertainty bands for the nucleation efficiency of LXe are shown in Fig. 5.36. As described in section 5.4.5 for the PICO analysis, the uncertainty bands presented in this plot are derived assuming a Gaussian likelihood function (which is not necessarily true) and are meant to provide an approximate visual representation of the statistical uncertainty of the fit only. Any subsequent analyses using this result make use of the actual MCMC samples instead. The fit results for all three nuisance parameters – representing systematic uncertainty on the neutron exposure for the different calibration experiments (see section 5.6.2) – all were found to be within one standard deviation of their nominal values.

As was found in the PICO analysis presented in section 5.4 and ref. [3], the measured nucleation threshold deviates from the predicted Seitz model thresholds for these experiments. For both lower threshold setpoints ($Q_{\mathrm{Seitz}} = 0.9$ and $1.48\,\mathrm{keV}$), the measured threshold is potentially compatible with a step-function threshold (the theoretical prediction), but displaced from Q_{Seitz} by approximately $1\,\mathrm{keV}$. The trend

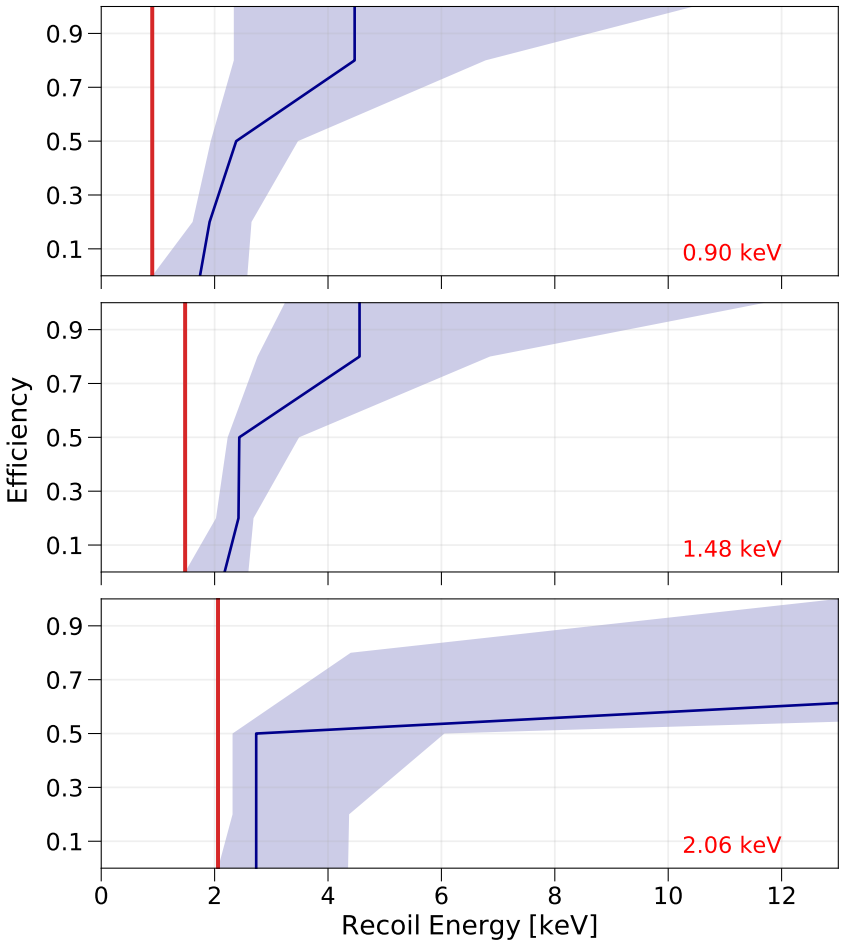


Figure 5.36: Best-fit and $1\,\sigma$ uncertainty bands (navy curve and shaded regions) of the nuclear recoil nucleation efficiency model obtained from the Xe-SBC calibrations. In this case, the model included piecewise linear efficiency curves at three different thermodynamic threshold setpoints (Q_{Seitz} values indicated in red).

of the $Q_{\text{Seitz}} = 2.06\,\text{keV}$ curve to high recoil energies at higher efficiencies is likely due to the poor constraining power of the data in this range. It is important to note the wide uncertainty band in this case, which permits compatibility with a threshold curve closer to a step function displaced by a few keV, matching the behaviour of the rest of the model. A major milestone demonstrated with these results is the first measured sub-keV nucleation from nuclear recoils, in a rare-event search bubble chamber (operated in conditions such that it is not overwhelmed by electronic recoil backgrounds). This gives credence to the feasibility of the SBC collaboration's goal of operating a liquid noble bubble chamber with $\mathcal{O}(100\,\text{eV})$ nucleation thresholds [94, 298].

To directly compare the nucleation model with the data, Fig. 5.37 shows the bubble rate predicted for each experiment by the best-fit model, compared to the data itself. This reveals that the Bi-Be data did not provide any useful constraint, as it was essentially compatible with the background rate. It is also apparent from this figure that there is good agreement between the model and almost all data points, with slightly more tension for the 2+ bubble events from the Y-Be source, for $Q_{\text{Seitz}} = 1.14$ and $1.33\,\text{keV}$ in particular. This small discrepancy was investigated further by examining the likelihood function decomposed for each experiment individually, to see if there was tension between the Y-Be calibration and others. This is shown for the $Q_{\text{Seitz}} = 0.9\,\text{keV}$ efficiency curve in Fig. 5.38. This plot shows that the global best fit, as well as the preferred parameter values of the ^{252}Cf and Y-Be likelihoods are almost identical, meaning there is no tension between these different calibration experiments. Unsurprisingly the Bi-Be data provides a roughly constant likelihood function in this plot since it is largely uninformative. Further investigation of this issue is ongoing, specifically by re-fitting the data using an independent simulation of the Y-Be experiment as a cross-check.

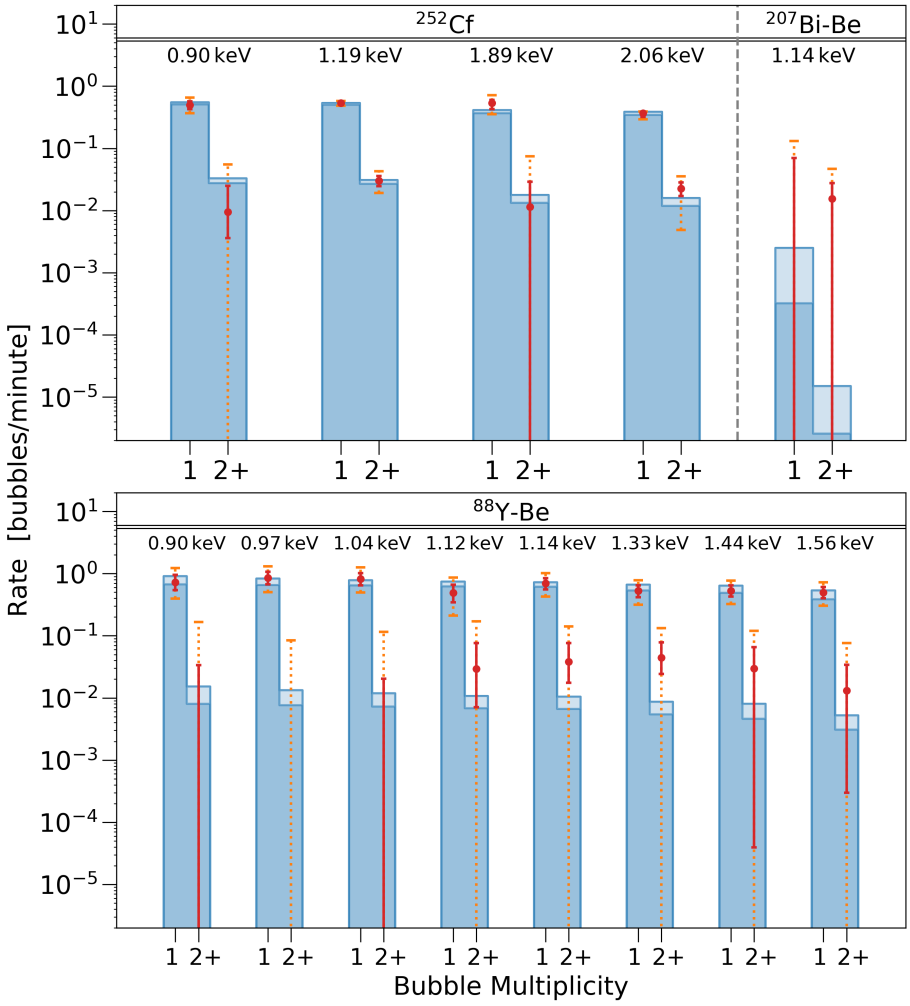


Figure 5.37: Background subtracted Xe-SBC calibration data (red points) listed by experiment, Q_{Seitz}, and bubble multiplicity. The red (respectively orange dashed) data error bars are the $1\,\sigma$ (respectively $2\,\sigma$) uncertainties on the measured bubble rats, including uncertainty on the background rate. The blue histogram represents the best-fit model to the data, with the empty portion at the top of each bar indicating the $1\,\sigma$ uncertainty band on the predicted signal rate.

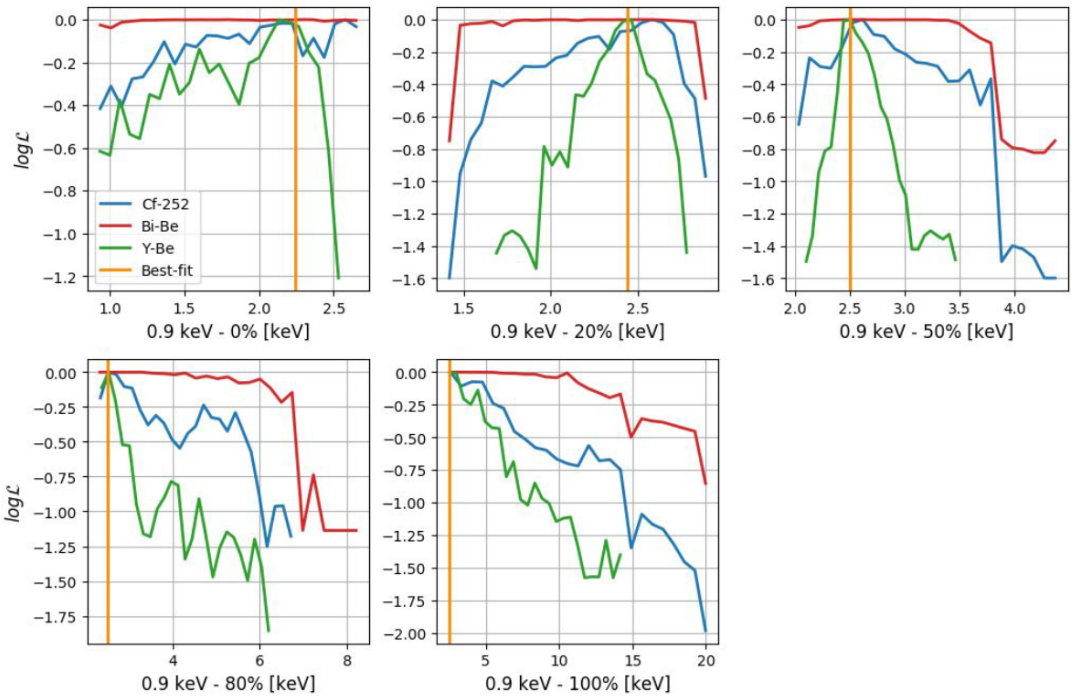


Figure 5.38: Profile log-likelihood functions for the $Q_{\mathrm{Seitz}} = 0.9\,\mathrm{keV}$ efficiency curve, decomposed into the likelihood functions for each of the three Xe-SBC calibration experiments. The best-fit parameter values obtained from the global fit are indicated by orange vertical lines. The likelihood curves are translated so that their maximum occurs at 0, for easier comparison.

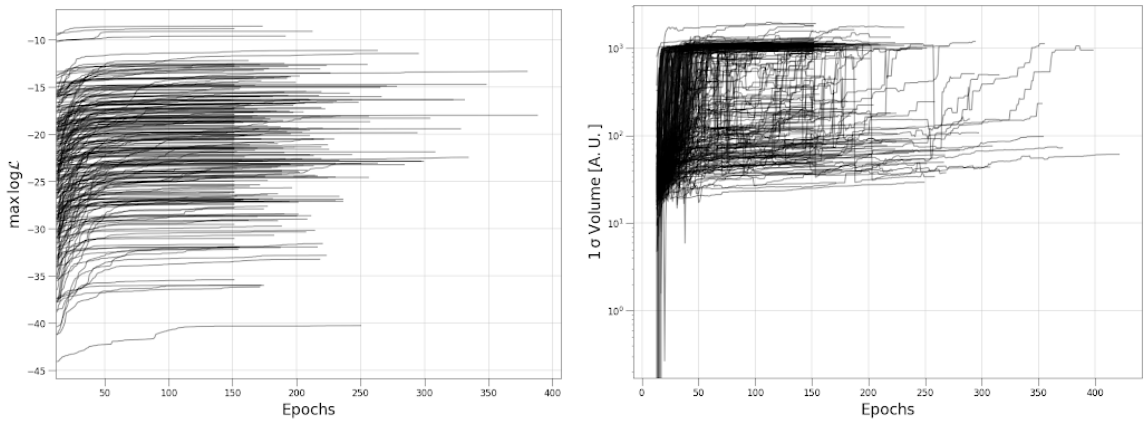


Figure 5.39: The maximum log-likelihood (left) and $1\,\sigma$ volume (right) of the fits of all 250 simulated Xe-SBC nuclear recoil calibration datasets, as a function of MCMC epoch.

5.6.6 Parametric Monte Carlo study

Following the steps laid out in the PICO nucleation analysis [3], a parametric Monte Carlo (see section 5.4.7) study was carried out for this LXe nucleation efficiency study. Specifically, 250 MC datasets were generated using the best-fit model, shown in Fig. 5.36. This includes randomly drawn background and source plus background counts for each experiment and bubble multiplicity. Note that in this case, the faster computation times compared to the PICO study allowed for $10\times$ the number of simulated datasets to be generated and studied. This MC data was fit using the same procedure as the real data (see section 5.6.3). The convergence metrics of these fits as a function of MCMC progress are shown in Fig. 5.39.

Exactly as was done for the PICO study, the main result of this additional step was to allow for interpretation of the $\chi^2 = -2\log\mathcal{L}_{\max}$ value for the actual best-fit model as a proper goodness-of-fit statistic. This was done by calculating χ^2 using the best-fit points of the MC datasets and examining their distribution, which is shown in Fig. 5.40. This can then directly be used to calculate a p-value for the fit of the real data, $p = 0.22$. This confirms that the model does adequately represent the data, and that the MCMC fit of the real data converged. Fitting a χ^2 distribution to this

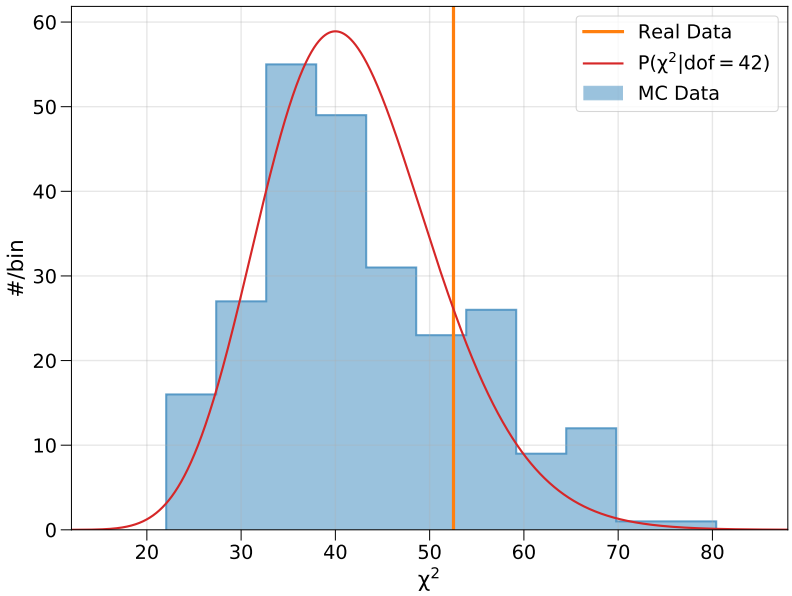


Figure 5.40: Distribution of χ^2 for 250 MC datasets (blue histogram), which is then fit with a χ^2 distribution with 42 degrees of freedom (red curve). The χ^2 value for the fit of the real data is indicated by a vertical orange line. The binning of this histogram was chosen using the Freedman-Diaconis rule [354].

data (with an unbinned likelihood fit) reveals that the effective number of degrees of freedom is 42 (also shown in Fig. 5.40).

5.7 Model bias studies

The nuclear recoil nucleation efficiency results obtained for both C_3F_8 and LXe seemingly fit their respective datasets well (based on their p-values), and are in line with expectations based on previous literature (see section 5.2.1). However, the use of an *ad hoc* model not rooted in first-principles physics and fit to data with a custom MCMC algorithm, fairly raises the question of how confident one can be in these results. To partially address these concerns, MC data studies were carried out to verify that the MCMC fits actually converged, and that the data is well represented by these models (see sections 5.4.7 and 5.6.6). Another concern that can be explored is whether or not this analysis paradigm yields biased results. That is, even if the data can be accurately approximated with the model used, does the optimization of the likelihood function return an unbiased estimate of the data. The following investigations are based on the appendix B of ref. [3] for the PICO study, as well as a similar study for the Xe-SBC nucleation efficiency study.

5.7.1 Naive PICO bias study

To provide some initial insight as to whether or not the PICO nucleation efficiency analysis is biased, one can examine the MC data already described in section 5.4.7. These 25 MC datasets are generated from the best-fit model of the real data, and thus can provide an immediate check on the bias of the analysis. The fit results of the MC data are compared to their true, input values in Fig. 5.41. One can see that while for most parameters there is no clear bias, in some cases there are persistent offsets, such as the efficiency curves for carbon at low efficiencies, as well as for fluorine at 100% efficiency. While these offsets are small (and hard to ascribe any significance to with a sample size of only 25), they may be emblematic of a systemic bias for the overall analysis.

The average residual of each parameter θ – denoted as $\Delta\theta$ – could be used to

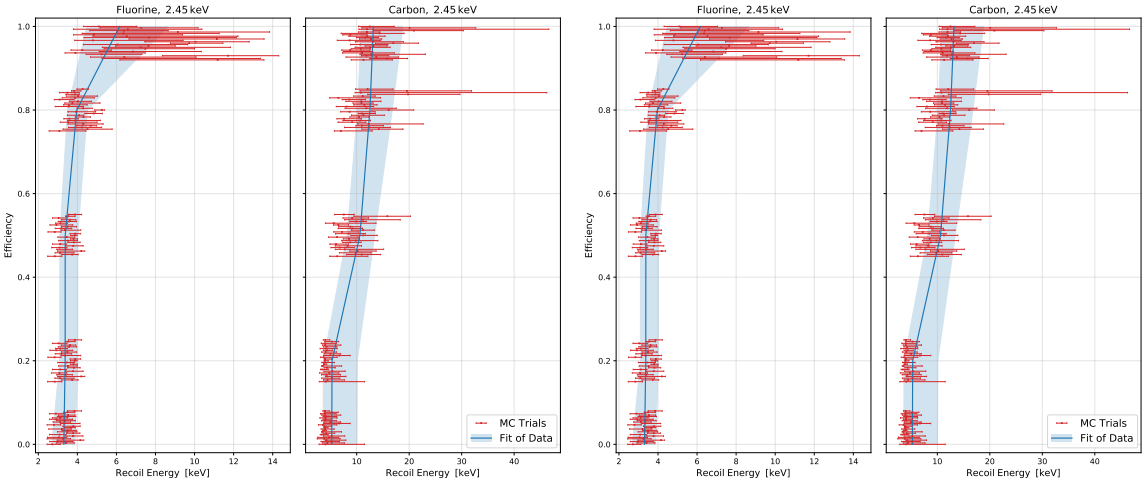


Figure 5.41: The fit results of 25 MC datasets (red markers) compared to the best-fit nucleation efficiency model used to generate the MC data (blue curve and band, the same as shown in Fig. 5.16). Taken from figure 15 of ref. [3].

calculate a "bias function" B for that parameter, which relates the true, bias-free value θ_{True} and the observed fit value $\hat{\theta}$ [355]:

$$\hat{\theta} = \theta_{\text{True}} + B_{\theta} \tag{5.30}$$

At this point in similar analyses, a critical assumption is often made, which is that the bias function is simply an offset, or a constant function of the parameter θ, allowing for simple calculation of the bias-free result:

$$B(\theta_{\text{True}}) = B\left(\hat{\theta}\right) \implies \theta_{\text{True}} = \hat{\theta} - B\left(\hat{\theta}\right) \tag{5.31}$$

A generic depiction of the bias correction implied in this scenario is shown in Fig. 5.42.

5.7.2 Robust bias correction for the PICO nucleation study

The key assumption made in the previous section is not known to be true in this case, prompting a more thorough investigation of the PICO nucleation efficiency results for bias. Specifically, it may be the case that the bias (if there is indeed any) of each parameter θ_j is a non-constant function of that parameter, but also potentially a

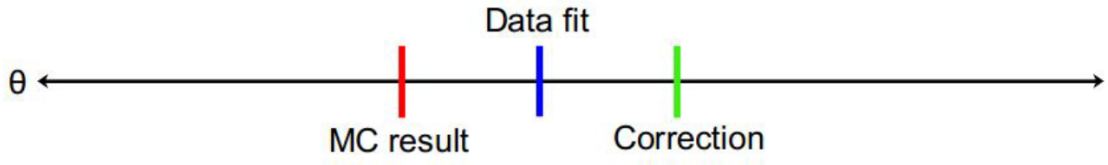


Figure 5.42: Generic example of an estimate of a parameter value ("data fit") compared to the average MC result of data generated using the best-fit value ("MC result"). Assuming the bias is a constant offset, the bias-corrected result is shown in green.

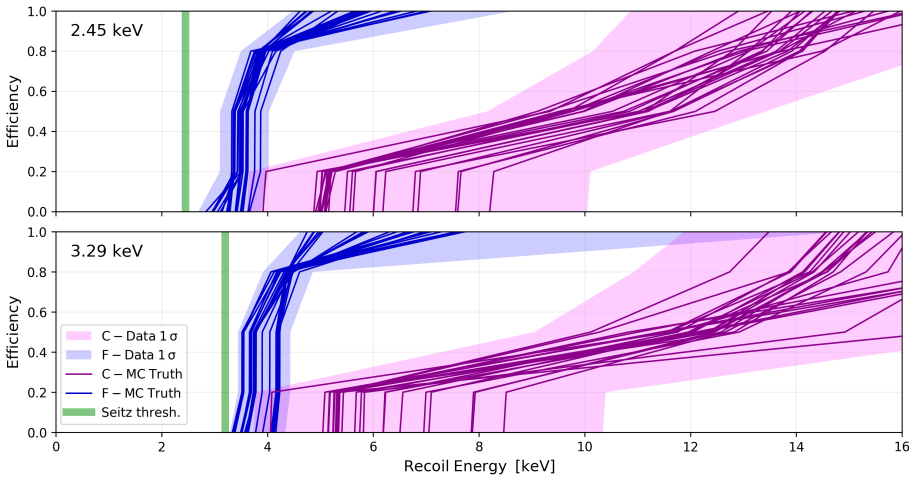


Figure 5.43: Nucleation efficiency models used to generate 25 MC datasets for the PICO nucleation analysis. Taken from figure 19 of ref. [3].

function of the true value of every other parameter θ_i. Thus, we can consider that there are 400 unique bias functions $B_{i,j}(\theta_i)$ for every combination of the 20 parameters of interest in the PICO analysis. The relation between the bias-free parameter values and the observed fit values $\hat{\theta}_j$ follows the same form as eq. 5.30 (here θ_j represents the true, bias-free value of that parameter):

$$\hat{\theta}_j = \theta_j + B_{i,j}(\theta_i) \tag{5.32}$$

To fully determine all bias functions $B_{i,j}$, ideally MC datasets would be generated by varying θ_i and fixing all other parameters, with multiple trials for every combination of i and $j \leq 20$, and then fitting all of the resulting datasets as described in section 5.4.4. However, this is not computationally practical. To more succinctly obtain an estimate of the bias functions with minimal computation, 25 MC datasets were generated using varying input models, drawn randomly from within the $1\,\sigma$ uncertainty band of the original PICO nucleation efficiency result of section 5.4.5. In this way, the bias functions can be tested for a range of parameter values locally around the original best-fit model. The input models used are shown in Fig. 5.43. The resulting MC datasets are fit as usual.

As a first step, these results can be compared to the average biases obtained in the naive study of section 5.7.1, which is shown in Fig. 5.44. One can see that the average biases are similar to those of the MC data fits generated using the original best-fit model, but not in all cases; this limited comparison does not fully answer the question of whether or not the bias functions are constant. To get a better idea, the bias of every parameter j can be plotted against the input value of every parameter i to see if there is a non-constant relationship (a constant bias function would result in no apparent correlation between the two). This information is summarized by the correlation coefficient of each pair of $\Delta\theta_j$ and θ_i. For this, the Spearman correlation coefficient ρ is used instead of the usual Pearson coefficient, since the former does not assume normality of the distributions [356]. This quantity is calculated for 2 random

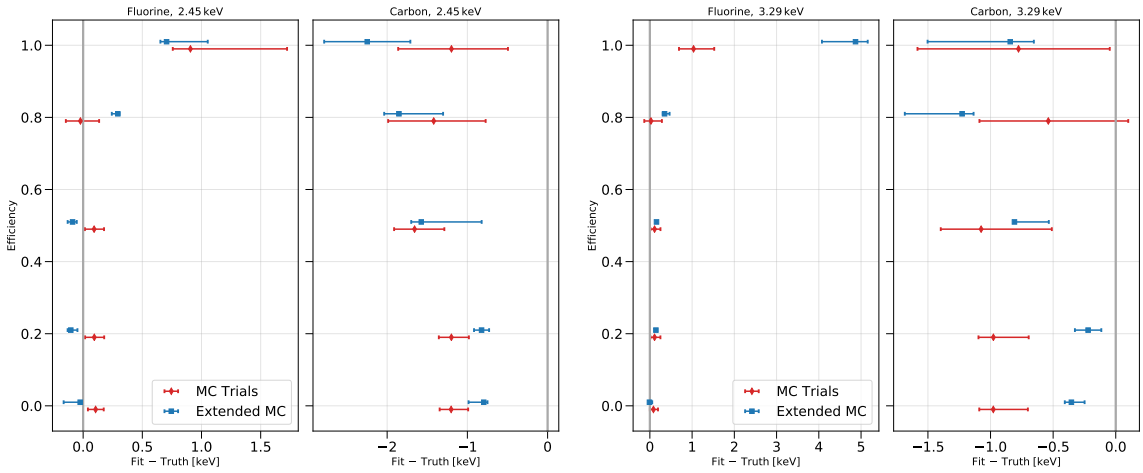


Figure 5.44: Average biases of 25 MC datasets generated using the best-fit PICO nucleation efficiency model (red) compared to MC data generated with varying input models (blue). The average values and error bars were calculated by stacking the profile likelihoods for all fits, for a given parameter, assuming that the resultant combined likelihoods are Gaussian (only for visualization purposes).

variables using their rank (their position out of N data points in ascending numerical order):

$$\rho_{i,j} = 1 - \frac{6}{N^3 - N}\left[\sum_{k=1}^{k=N}\left(rank(\Delta\theta_j) - rank(\theta_i)\right)^2\right] \tag{5.33}$$

Like the common Pearson correlation coefficient, $-1 \leq \rho \leq 1$, with larger absolute values indicating strong correlations/anti-correlations. All values of $\rho_{i,j}$ (for the 20 parameters of interest in the PICO analysis) are shown in Fig. 5.45.

While nearly all combinations exhibit some correlation, it is important to keep in mind the low statistics of this study (only 25 trials). For this sample size, a statistically significant correlation (i.e. the null hypothesis of no correlation can be rejected at the 90% confidence level) requires that $|\rho| > 0.324$ [356]. There are indeed many combinations with a statistically significant correlation, which proves that many of the bias functions in this analysis are not constant, and do depend on the input values of many different parameters in some cases. Thus, the simple bias-correction procedure presented in section 5.7.1 is not valid for this study.

To proceed, functional representations of all $B_{i.j}$ are needed. In this analysis, all

Figure 5.45: The Spearman correlation coefficients $\rho_{i,j}$ for every combination of bias function $B_{i,j}$ and input parameter value θ_i, as calculated in eq. 5.33. Abbreviated parameter notation is used indicating the atomic species (C/F), thermodynamic threshold fencepost (2.45 keV = L, 3.29 keV = H), and efficiency (0, 0.2, 0.5, 0.8, 1). The colour-scale also represents the indicated value in each grid space for easier visualization. Grid spaces with boxed values indicate cases of a statistically significant correlation (at the 90% confidence level). Taken from figure 21 of ref. [3].

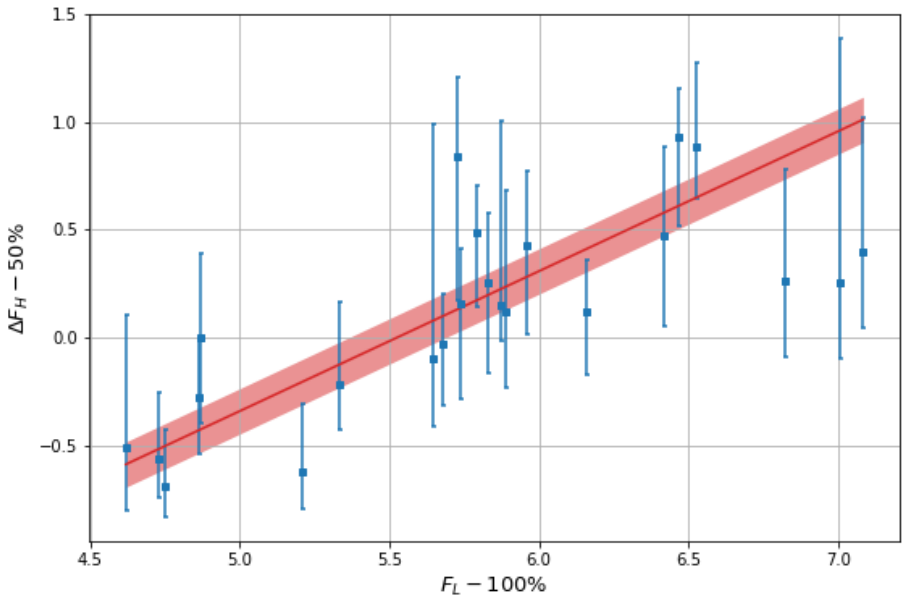


Figure 5.46: Example of a bias function fit with a first-order polynomial. In this specific case, the bias of the 50% efficient point of the fluorine nucleation threshold at $Q_{\text{Seitz}} = 3.29\,\text{keV}$ is plotted as a function of the fluorine 100% efficiency point with $Q_{\text{Seitz}} = 2.45\,\text{keV}$. No outliers are rejected in this case. Taken from figure 20 of ref. [3].

graphs of $\Delta\theta_j$ vs. θ_i are approximately constant or linear, which means the Spearman correlation coefficients presented in Fig. 5.45 can be conveniently used to delineate the two cases. Thus, all bias/input combinations are fit with a first-order polynomial when $|\rho_{i,j}| > 0.324$, or a zeroth-order polynomial otherwise (a constant function). The fitting in this case is done simply using least-squares [73]. An outlier rejection procedure is also applied to the bias data; all data points except one are fit with the appropriate function, and if the exempted point is a $> 2\,\sigma$ outlier, it is rejected. One example of bias data and the resulting first-order polynomial fit is shown in Fig. 5.46.

The resulting bias functions collectively represent 400 constraints (the polynomial fits of the bias data with statistical uncertainty) on the true values of the 20 parameters of interest in this analysis. Thus this problem can be thought of as an over-determined system of equations, except there's also the statistical uncertainty in the bias function fits to be included. To incorporate all of these constraints, the fol-

lowing likelihood function can be constructed for the set of bias-corrected parameter values $\{\theta\}$ given an observed set of parameters $\left\{\hat{\theta}\right\}$:

$$\log \mathcal{L}_{\text{bias}} \left(\left\{\hat{\theta}\right\} | \{\theta\} \right) = \sum_i \sum_j \log P \left(\hat{\theta}_j = \theta_j + B_{i,j}(\theta_i) \right) \tag{5.34}$$

where the probability associated with each term P comes from the fit of each bias function (with uncertainties determined by the covariance matrix of the fit). Maximizing this likelihood function using the fast burn-in MCMC method proved to be even more computationally intensive than the original fit of the data, requiring almost 10000 CPU hours. Therefore, the bias correction was only calculated for the original best-fit nucleation efficiency model (see Fig. 5.16). For the sake of demonstration, the obtained bias correction was applied to the $\pm 1\,\sigma$ uncertainty limits of the original result, to produce a bias-corrected PICO nucleation efficiency model. This result is shown in Fig. 5.47.

Fortunately, the bias-corrected result is nearly identical to the original fit result. Qualitatively, this suggests that although there are many individual cases of small systemic bias in the fit (as evidenced by Fig. 5.45), in aggregate they tend to cancel each other out and are merely representative of tensions in the fit. This close agreement lends validity to the original result presented in section 5.4.5 (which potentially will be used in future analyses), as well as the preliminary version of this analysis used for the already published PICO-60 dark matter result [93].

5.7.3 Bias study for the Xe-SBC experiment

A bias study was also carried out for the Xe-SBC nucleation efficiency analysis. Note that, unfortunately, this study used slightly outdated nuisance parameters compared to those given in table 5.5. These differences are not expected to change the conclusions of this study. At the same time, it was deemed not practical to repeat this study with this correction made. Regardless, as with the PICO bias study, this result should bolster confidence in the model/analysis paradigm for the main result.

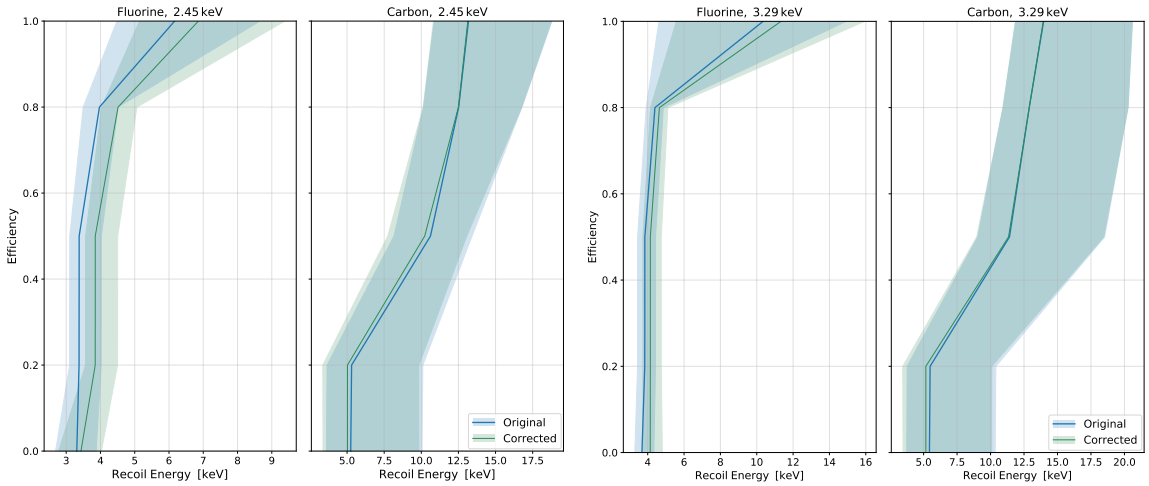


Figure 5.47: Original (blue) and bias-corrected (green) nucleation efficiency results for the PICO analysis, with best-fits and $1\,\sigma$ uncertainty bands. Taken from figure 22 of ref. [3].

250 MC datasets were generated using randomly-drawn input models from the $1\,\sigma$ uncertainty band of the original fit result (see Fig. 5.36). Note that this differs from the MC data described in section 5.6.6, which were generated from the best-fit nucleation efficiency model. In this way, possible fit biases can be explored in the local region of parameter space around the best fit of the real data, allowing for a bias correction that does not assume a constant offset bias (see section 5.7.1).

These 250 MC datasets were fit following the same procedure described in section 5.6.3. The fit results can then be used to determine the bias function $B_{i,j}$ of every parameter j as a function of every parameter i, as in section 5.7.2. To see if there are non-constant biases present, the Spearman correlation coefficient of every combination of bias $\Delta\theta_j = \hat{\theta}_j - \theta_j$ vs. input parameter values θ_i is calculated, and shown in Fig. 5.48. As with the PICO analysis, there are indeed cases where biases are present that are not constant, as implied where there are statistically-significant correlations. Note that in this study, model parameters are abbreviated indicating their efficiency setpoint and Q_{Seitz} set point ($1 \rightarrow Q_{\text{Seitz}} = 0.9\,\text{keV}$, $2 \rightarrow Q_{\text{Seitz}} = 1.48\,\text{keV}$, $3 \rightarrow Q_{\text{Seitz}} = 2.06\,\text{keV}$).

Unlike the PICO bias study, the bias relationships defined by these fit results do not all exhibit a first-order polynomial (or constant) behaviour. While some

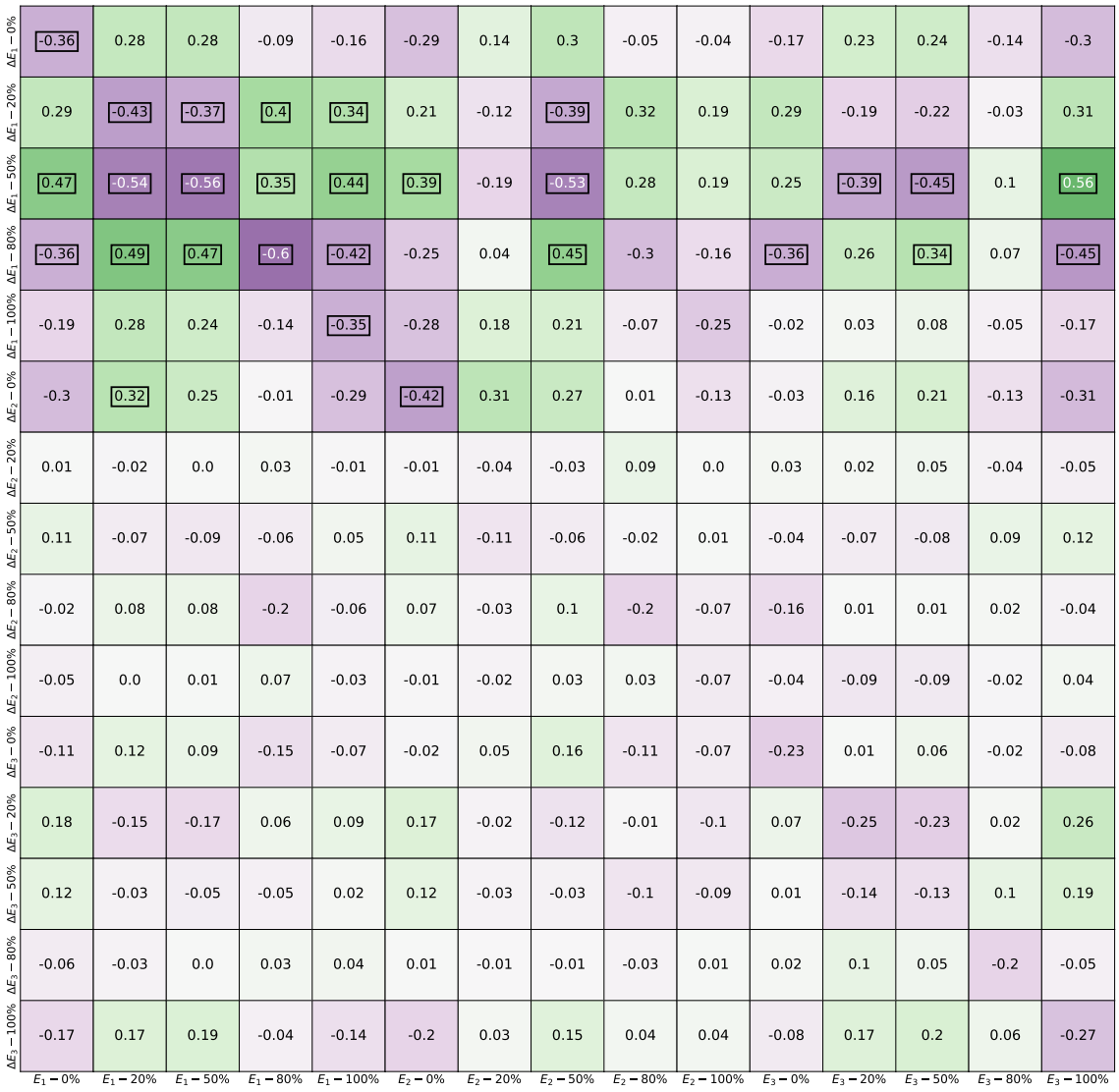


Figure 5.48: Spearman correlation coefficients between all parameter biases and true parameter values for the Xe-SBC nucleation efficiency study. The colour-scale also represents the listed values to aid in visualization. Boxed values are statistically significant correlations (at the 90% confidence level).

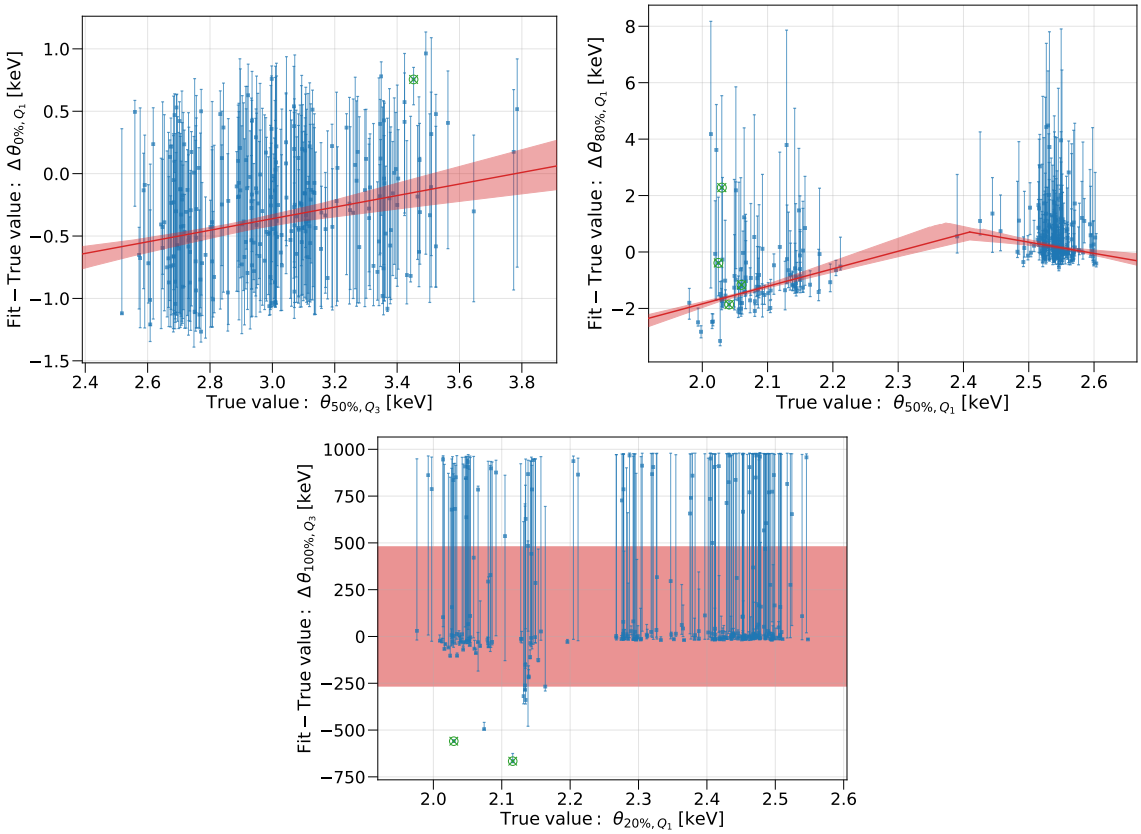


Figure 5.49: Examples of bias data for the Xe-SBC nucleation efficiency analysis (blue points), with rejected outliers labelled (green "x"). The three different parametrization options are shown; linear (left), piece-wise linear with a single knot (center), and an "equal probability" band (right) in red with $1\,\sigma$ uncertainty bands for the former two cases.

parameter/bias combinations do, others show a clear two-part linear relationship, which is parameterized with a piecewise linear function with a variable knot position. Yet other bias relationships are very poorly constrained, i.e. the bias values obtained have very large uncertainty ranges regardless of the input parameter value. This behaviour is characterized by a function with constant probability for the value of the bias, between two extreme values a and b which are variable:

$$P(\Delta\theta_j) = \begin{cases} \frac{1}{b-a}, & \text{when } a \leq \Delta\theta_j \leq b \\ 0, & \text{otherwise} \end{cases} \tag{5.35}$$

To obtain functional forms for all 225 bias relationships in the Xe-SBC analysis (from having 15 parameters of interest), it is necessary to algorithmically select between these three modalities. This was done by fitting every set of bias data with all three function types and using the Bayesian odds ratio to select the best parametrization in each case. Before this, the same outlier rejection method described in section 5.7.2 was applied. Examples of all three bias function modalities and the resulting parameterizations are shown in Fig. 5.49.

These 225 bias functions are then combined to create the bias-correction likelihood function for this analysis, using eq. 5.34. Optimizing this likelihood function gives an estimate for a bias-corrected set of parameters $\{\theta\}$ for a given observed set of parameters $\left\{\hat{\theta}\right\}$. While it is not computationally practical to determine the correction for all of the millions of MCMC samples constituting the fit of the real data, this procedure was carried out for the original best fit as well as 10000 other random MCMC samples. The resulting bias-corrected best fit and uncertainty band are shown in Fig. 5.50. This figure demonstrates that – as with the PICO bias analysis – the overall bias correction is very small. Perhaps the only significant change is the high efficiency, $Q_{\text{Seitz}} = 2.06\,\text{keV}$ parameters. In the original result, these parameters had very large uncertainty and were poorly constrained, which is somewhat curtailed in the bias-corrected result. The narrowness of the uncertainty bands of the other bias-

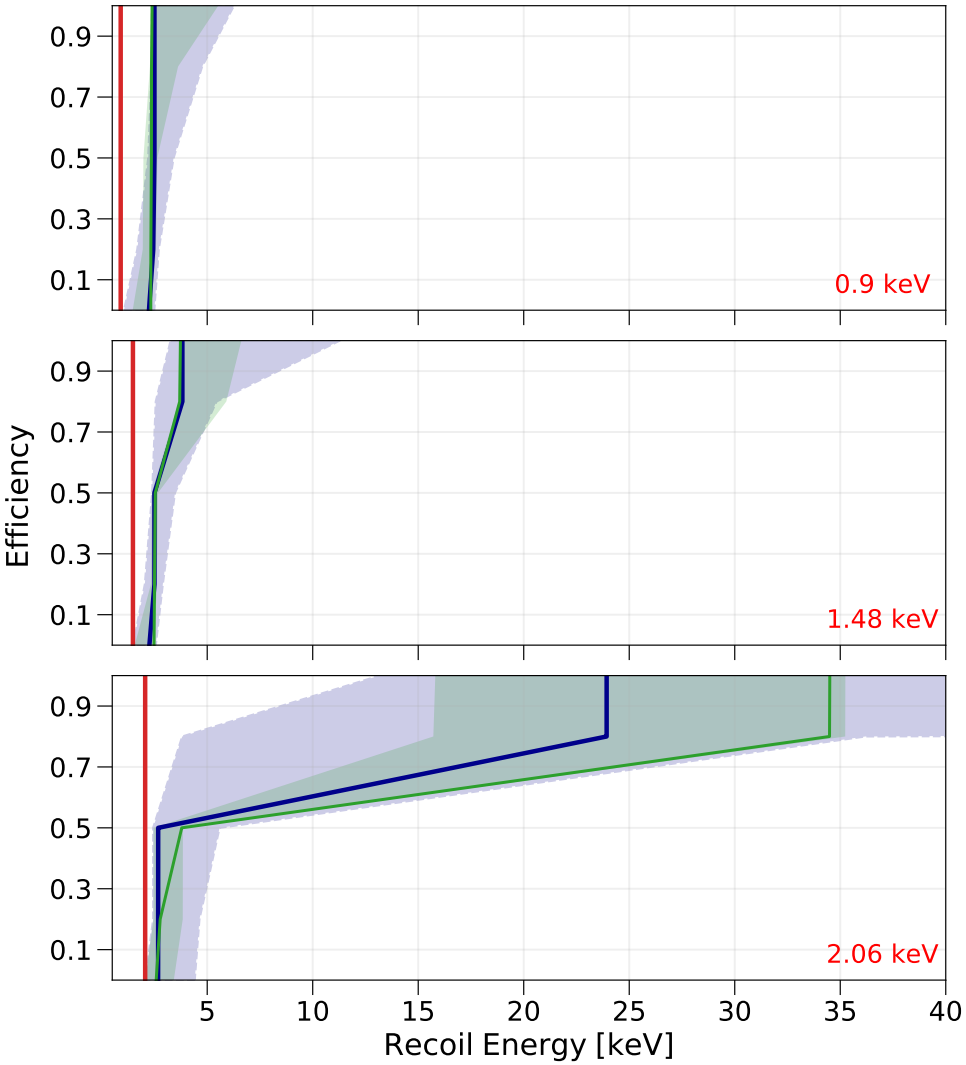


Figure 5.50: The original best-fit nucleation efficiency model as shown in Fig. 5.36 (blue) compared to the bias-corrected model with $1\,\sigma$ uncertainty bands (green).

corrected parameters is likely due to the poor statistics of the bias-corrected sampling (10000 samples vs. $\mathcal{O}(10^6)$ samples in the original fit). Overall, this analysis adds validity to the Xe-SBC nucleation efficiency study and the original fit result of Fig. 5.36.

Chapter 6

Conclusions

This body of work documents efforts to better understand the physics of proportional counter and bubble chamber-based dark matter search experiments for the NEWS-G and SBC collaborations. While the search for WIMP-like dark matter continues – and expands to lower mass candidate particles – this work is crucial to both continue to constrain the parameter space of particle dark matter, and to hopefully claim a discovery of dark matter one day. Particle dark matter remains a well-supported class of dark matter hypothesis – one of the most significant questions ever asked by the scientific community.

6.1 Dark matter searches with NEWS-G

In chapter 3, UV laser and ^{37}Ar calibration techniques for the NEWS-G experiment were presented. These methods – new to the collaboration since the time of its first dark matter publication [63] – allow for a full characterization of the detector's energy response, including primary ionization and the Townsend avalanche process. The physics of the former are now better characterized with the COM-Poisson distribution for primary ionization [140]. However, measurements of the W-value ($W(E)$) for different gases remains an ongoing pursuit of NEWS-G. The methane gas W-value study described in section 3.3 represents the most robust to-date measurement of this quantity by NEWS-G, with many systematic effects considered. This study

will provide the recipe for future measurements of $W(E)$ that are ongoing and/or planned at the U of A and at other SPC setups. These future measurements are all the more important given the small – but persistent – discrepancy between the W-value results of section 3.3 and some existing results in the literature [166, 208, 209]. This discrepancy is either due to as-of-yet undiscovered systematics in the NEWS-G approach, the lack of uncertainties reported for the historical results, or due to genuinely different SPC-specific physics. No matter the answer to this question, SPC measurements of the W-value will continue to be mandatory for all future NEWS-G experiments.

In section 3.4, the prototype SPC installed at the U of A was described, as well as ongoing efforts to develop and characterize a novel gas handling apparatus with it. In addition to an LAS system to measure methane gas concentrations in real-time, silver zeolite was explored as a new radon trap substrate. Based on the preliminary results given in this work and in ref. [221], this material shows great promise as a successor to the traditional activated charcoal substrates used by other experiments. Its competitive performance even while operated at room temperature makes silver zeolite an appealing choice, which will likely result in continued interest from NEWS-G and other rare-event search experiments. In the future, the U of A SPC will be used for W-value measurements in new gas conditions, among other physics objectives.

In chapter 4, the S140 detector was described, the culmination of years of effort from NEWS-G collaborators to improve upon the previous generation of SPCs, including improved material selection, internal shielding, and contaminant assaying [1, 2]. The dark matter physics campaign undertaken with this detector at the LSM was the first-ever to use a methane target – a hydrogen-rich material with desirable properties as a medium in which to search for low-mass dark matter. Furthermore, this was the first time that the novel ACHINOS sensor was used for a dark matter experiment – a necessary technological development for the operation of large SPCs. The work presented in Chapter 4 constitutes a complete overhaul of the analysis tech-

niques of the NEWS-G experiment. New pulse-treatment techniques were needed (section 4.2.3), a completely new approach to UV laser calibrations were developed (section 4.3.1), and existing analysis methods for ^{37}Ar were expanded to account for large backgrounds, and charge attachment (section 4.3.3). Due to its large size and significant electron diffusion, it was possible to count individual primary electron peaks – a reality that was as beneficial to the final result of the experiment as it was demanding of more sophisticated analysis techniques to characterize this information (sections 4.3.1, 4.4.2). The unique diffusion properties of this experiment also had consequences for the modelling of electron transport in the detector, calculation of the hardware trigger efficiency, and the observation of space-charge effects on a much more dramatic scale than seen before (sections 4.2.5, 4.3.5, 4.4.1). The two-channel ACHINOS readout resulted in the discovery of powerful PSD variables such as the N/S amplitude, but again necessitated extra efforts in characterization and calculation of the detector's fiducial volume (sections 4.4.3, 4.4.5).

The over-arching theme of the LSM dark matter analysis was of greatly improved knowledge of SPC physics and improved sensitivity to low-mass dark matter, coupled with the need to develop new analysis and characterization techniques to match. All of this culminated in the first PLR analysis of the NEWS-G collaboration, reflecting our improved knowledge of experimental backgrounds. The preliminary exclusion limit based on non-blind data for low-mass WIMPs (with a spin-dependent proton coupling) creates the expectation of a world-leading result when the full data is analyzed in the coming months. As the collaboration currently prepares for its next dark matter search with S140 at SNOLAB (with neon gas), it is hoped that the methods developed and outlined in this work will provide the foundation for future analyses. The improved gas-handling methods developed with the U of A SPC will also play a key role in this next phase of the experiment.

Still, there are outstanding questions for the NEWS-G collaboration to answer in the future. The large background of single electron events experienced at the LSM

(similar to the experiences of other low-mass dark matter experiments [74]) poses a major challenge. These may be caused by high-energy α-decay events, despite cuts applied to remove most of the single electrons induced by these (section 4.4.4). The physics of this phenomenon are not fully understood. Other than future W-value measurements, NEWS-G will continue to struggle with the deficiency of quenching factor measurements. While some data exists, more studies in the relevant target gases are needed, especially at low energies ($\mathcal{O}(100)\,\mathrm{eV}$), since that is where the majority of our sensitivity to dark matter derives from. The relatively conservative logarithmic quenching factor extrapolation used in the LSM analysis (see section 4.3.4) implies that the quenching factor drops to 0 around 100 eV, a behaviour that is not expected to reflect reality [177]. Therefore, any measurements of the quenching factor in this energy range – no matter how large the uncertainty or how low the measured value is – would dramatically improve our sensitivity to low-mass dark matter. Some future measurements are planned using accelerator neutrons at TUNL [184], as well as possible neutron capture experiments (section 3.4.3).

6.2 Bubble chamber nucleation efficiency studies

In chapter 5, measurements of the nucleation efficiency of bubble chambers – for both fluorocarbon and liquid xenon fluids – were presented. This quantity is analogous to the quenching factor in gaseous detectors, and is the principal physical quantity that determines these experiments' sensitivity to dark matter. The C_3F_8 study presented in sections 5.3 and 5.4 (resulting in a publication [3]) is the culmination of years of effort by the PICO collaboration – with data dating beginning in 2013/2014. The final nucleation efficiency result of this study validated the preliminary one used in recent PICO publications [93, 343]. Applying similar techniques, the study for the LXe scintillating bubble chamber (sections 5.5 and 5.6) is the first such measurement performed for any liquid noble bubble chamber. These novel detector fluids (LAr and LXe) show great promise for dark matter and neutrino detectors [94], due in part

to their large-A target atoms, and energy reconstruction/veto capabilities due to the scintillation light produced during events. Additionally – and perhaps more importantly for the first - dark matter SBC at SNOLAB – these target media appear to have even greater suppression of electronic recoils, and lower nuclear recoil thresholds than traditional fluorocarbon liquids. The latter was demonstrated in this LXe study, which yielded measured nucleation thresholds as low as $0.9\,\mathrm{keV_{nr}}$. This result, as well as the electronic recoil measurements carried out for the LXe bubble chamber [345], will be featured in an upcoming publication.

The bias analysis for both studies (section 5.7) reinforces the validity of the model and analysis techniques presented in this work. With that reassurance – and the now double-application of these methods – this approach will likely be used for future liquid noble and fluorocarbon nucleation efficiency measurements with both the PICO and SBC collaborations. In the future, more expansive neutron calibration data will allow for more detailed study of the threshold dependence on thermodynamic state and recoil energy. Simpler functional forms than the piece-wise linear model used in this work may be appropriate to use in the nucleation efficiency model, and will make data fitting more computationally practical.

However, underlying these studies is the fundamental question of why the measured nucleation thresholds differs from the first-principles thermodynamical theory – the Seitz model [3, 317]. Additional energy-loss channels must play a role – i.e. ways by which a nuclear recoil might deposit energy that do not contribute to bubble-forming heat – but the specific mechanisms are not known. There are ongoing efforts (including by U of A M.Sc. student Xiang Li) to elucidate these effects using molecular dynamics simulations, which model these microscopic processes. When such efforts bear predictive results, it will be of great interest to compare these to the empirical results presented here. The LXe calibrations in particular offer a reasonable expectation of agreement with simulation, due to the added simplicity of dealing with a monatomic target liquid. This provides a path to better understanding the physics

of bubble chambers, as the search for dark matter continues.

Bibliography

[1] L. Balogh *et al.*, "The NEWS-G detector at SNOLAB," *JINST*, vol. 18, no. 02, T02005, 2023. DOI: 10.1088/1748-0221/18/02/T02005.

[2] L. Balogh *et al.*, "Copper electroplating for background suppression in the news-g experiment," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 988, p. 164 844, 2021. DOI: https://doi.org/10.1016/j.nima.2020.164844.

[3] B. Ali *et al.*, "Determining the bubble nucleation efficiency of low-energy nuclear recoils in superheated C_3F_8 dark matter detectors," *Phys. Rev. D*, vol. 106, p. 122 003, 12 2022. DOI: 10.1103/PhysRevD.106.122003.

[4] J. Silk *et al.*, *Particle Dark Matter: Observations, Models and Searches*, G. Bertone, Ed. Cambridge Univ. Press, 2010, ISBN: 9781107653924.

[5] R. Essig *et al.*, "Working Group Report: New Light Weakly Coupled Particles," in *Community Summer Study 2013: Snowmass on the Mississippi*, 2013. arXiv: 1311.0029 [hep-ph].

[6] Zurek, K.M., "Asymmetric Dark Matter: Theories, signatures, and constraints," *Phys. Rep.*, vol. 537, no. 3, p. 91, 2014, Asymmetric Dark Matter: Theories, signatures, and constraints. DOI: https://doi.org/10.1016/j.physrep.2013.12.001.

[7] C. Kouvaris and J. Pradler, "Probing sub-gev dark matter with conventional detectors," *Phys. Rev. Lett.*, vol. 118, p. 031 803, 3 2017. DOI: 10.1103/PhysRevLett.118.031803.

[8] R. Essig *et al.*, "Snowmass2021 Cosmic Frontier: The landscape of low-threshold dark matter direct detection in the next decade," in *2022 Snowmass Summer Study*, 2022. arXiv: 2203.08297 [hep-ph].

[9] J. Billard *et al.*, "Direct Detection of Dark Matter – APPEC Committee Report," 2021. arXiv: 2104.07634 [hep-ex].

[10] L. Flores *et al.*, "Physics reach of a low threshold scintillating argon bubble chamber in coherent elastic neutrino-nucleus scattering reactor experiments," *Phys. Rev. D*, vol. 103, p. L091301, 9 2021. DOI: 10.1103/PhysRevD.103.L091301.

[11] V. Trimble, "The World Line of Dark Matter: Its Existence and Nature Through Time," in *Proceedings of the 1st International Symposium on Sources of Dark Matter in the Universe*, World Scientific, 1994, pp. 9–22.

[12] J. Oort, "The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems," *Bull. Astron. Inst. Netherlands*, vol. 6, p. 249, 1932.

[13] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helvetica Physica Acta*, vol. 6, pp. 110–127, 1933.

[14] V. Rubin and W. Ford, "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions," *Astrophys. J.*, vol. 159, pp. 379–403, 1970.

[15] V. Rubin, W. Ford, and N. Thonnard, "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," *Astrophys. J.*, vol. 238, pp. 471–487, 1980.

[16] K. Begeman, A. Broeils, and R. Sanders, "Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics," *Mon. Not. R. Astron. Soc.*, vol. 249, no. 3, pp. 523–537, 1991.

[17] E. Corbelli and P. Salucci, "The extended rotation curve and the dark matter halo of M33," *Mon. Not. R. Astron. Soc.*, vol. 311, no. 2, pp. 441–447, 2000.

[18] M. Milgrom, "A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis," *Astrophys. J.*, vol. 270, pp. 365–370, 1983.

[19] G. Bertone and D. Hooper, "History of dark matter," *Rev. Mod. Phys.*, vol. 90, p. 045 002, 4 2018. DOI: 10.1103/RevModPhys.90.045002.

[20] M. Bradač *et al.*, "Revealing the Properties of Dark Matter in the Merging Cluster MACS J0025.4–1222," *Astrophys. J.*, vol. 687, no. 2, p. 959, 2008.

[21] D. Clowe *et al.*, "A Direct Empirical Proof of the Existence of Dark Matter," *Astrophys. J. Lett.*, vol. 648, no. 2, p. L109, 2006.

[22] P. van Dokkum *et al.*, "A galaxy lacking dark matter," *Nature*, vol. 555, pp. 629–632, 2018.

[23] P. van Dokkum, S. Danieli, R. Abraham, C. Conroy, and A. Romanowsky, "A second galaxy missing dark matter in the ngc 1052 group," *Astrophys. J. Lett.*, vol. 874, no. 1, p. L5, 2019. DOI: 10.3847/2041-8213/ab0d92.

[24] E. Shin *et al.*, "Dark Matter Deficient Galaxies Produced via High-velocity Galaxy Collisions in High-resolution Numerical Simulations," *Astrophys. J.*, vol. 899, no. 1, p. 25, 2020. DOI: 10.3847/1538-4357/aba434.

[25] C. Alcock *et al.*, "The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations," *Astrophys. J.*, vol. 542, no. 1, p. 281, 2000.

[26] L. Strigari, "Galactic Searches for Dark Matter," *Phys. Rept.*, vol. 531, pp. 1–88, 2013.

[27] C. Alcock *et al.*, "EROS and MACHO Combined Limits on Planetary-Mass Dark Matter in the Galactic Halo," *Astrophys. J.*, vol. 499, no. 1, p. L9, 1998. DOI: 10.1086/311355.

[28] L. Wyrzykowski *et al.*, "The OGLE view of microlensing towards the Magellanic Clouds – III. Ruling out subsolar MACHOs with the OGLE-III LMC data*," *MNRAS*, vol. 413, no. 1, pp. 493–508, 2011. DOI: 10.1111/j.1365-2966.2010.18150.x.

[29] J. N. Bahcall, P. Hut, and S. Tremaine, "Maximum mass of objects that constitute unseen disk material," *Astrophys. J.*, vol. 290, pp. 15–20, 1985. DOI: 10.1086/162953.

[30] C. Flynn, J. Holopainen, and J. Holmberg, "White dwarfs and galactic dark matter," *Monthly Notices of the Royal Astronomical Society*, vol. 339, no. 3, pp. 817–824, 2003. DOI: 10.1046/j.1365-8711.2003.06223.x.

[31] J. García-Bellido, "Massive Primordial Black Holes as Dark Matter and their detection with Gravitational Waves," *J. Phys. Conf. Ser.*, vol. 840, no. 1, p. 012 032, 2017.

[32] M. Kawasaki, A. Kusenko, and T. Yanagida, "Primordial seeds of supermassive black holes," *Physics Letters B*, vol. 711, no. 1, pp. 1–5, 2012. DOI: https://doi.org/10.1016/j.physletb.2012.03.056.

[33] J. Dror, H. Ramani, T. Trickle, and K. Zurek, "Pulsar timing probes of primordial black holes and subhalos," *Phys. Rev. D*, vol. 100, p. 023 003, 2 2019. DOI: 10.1103/PhysRevD.100.023003.

[34] K. Kohri and T. Terada, "Solar-mass primordial black holes explain NANOGrav hint of gravitational waves," *Phys. Lett. B*, vol. 813, p. 136 040, 2021. DOI: https://doi.org/10.1016/j.physletb.2020.136040.

[35] S. Bird *et al.*, "Snowmass2021 Cosmic Frontier White Paper: Primordial black hole dark matter," *Phys. Dark Universe*, vol. 41, p. 101 231, 2023. DOI: 10.1016/j.dark.2023.101231.

[36] Y. Akrami *et al.*, "Planck 2018 results. I. Overview and the cosmological legacy of Planck," *arXiv:1807.06205*, 2018.

[37] B. Ryden, *Introduction to Cosmology*, 1^{st}. Addison-Wesley, 2002, ISBN: 978-0805389128.

[38] C. L. Bennett *et al.*, "Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results," *Astrophys. J. Suppl. S.*, vol. 208, no. 2, p. 20, 2013.

[39] M. Roos, "Astrophysical and Cosmological Probes of Dark Matter," *J. Mod. Phys.*, vol. 3, pp. 1152–1171, 2012.

[40] C. Patrignani *et al.*, "Review of Particle Physics," *Chin. Phys.*, vol. C40, no. 10, 2016. DOI: 10.1088/1674-1137/40/10/100001.

[41] S. Dodelson, "The Real Problem with MOND," *Int. J. Mod. Phys.*, vol. D20, pp. 2749–2753, 2011.

[42] R. Workman *et al.*, "Review of Particle Physics," *PTEP*, vol. 2022, p. 083C01, 2022. DOI: 10.1093/ptep/ptac097.

[43] J. Bullock and M. Boylan-Kolchin, "Small-Scale Challenges to the *Lambda* CDM Paradigm," *Ann. Rev. Astron. Astrophys.*, vol. 55, no. 1, pp. 343–387, 2017. DOI: 10.1146/annurev-astro-091916-055313.

[44] G. Jungman, M. Kamionkowski, and K. Griest, "Supersymmetric dark matter," *Phys. Rep.*, vol. 267, no. 5, pp. 195–373, 1996.

[45] C. Seitz, *Searches for strong production of supersymmetry at atlas and cms*, 2017. arXiv: 1710.05327 [hep-ex].

[46] M. Khlopov, "Particle Dark Matter Candidates," *Mod. Phys. Lett.*, vol. A32, p. 1 702 001, 2017.

[47] J. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection," *Ann. Rev. Astron. Astrophys.*, vol. 48, pp. 495–545, 2010.

[48] F. Steffen, "Dark Matter Candidates - Axions, Neutralinos, Gravitinos, and Axinos," *Eur. Phys. J.*, vol. C59, pp. 557–588, 2009.

[49] Q. Arnaud *et al.*, "Solar kaluza-klein axion search with news-g," *Phys. Rev. D*, vol. 105, p. 012 002, 1 2022. DOI: 10.1103/PhysRevD.105.012002.

[50] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "Phenomenology, astrophysics, and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity," *Phys. Rev. D*, vol. 59, p. 086 004, 8 1999. DOI: 10.1103/PhysRevD.59.086004.

[51] S. Chang, S. Tazawa, and M. Yamaguchi, "Axion model in extra dimensions with TeV scale gravity," *Phys. Rev. D*, vol. 61, p. 084 005, 8 2000. DOI: 10.1103/PhysRevD.61.084005.

[52] A. Boyarsky, M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy, "Sterile neutrino dark matter," *Prog. Part. Nucl. Phys.*, vol. 104, pp. 1–45, 2019. DOI: 10.1016/j.ppnp.2018.07.004.

[53] K. Petraki and R. Volkas, "Review of asymmetric dark matter," *Int. J. Mod. Phys. A*, vol. 28, no. 19, p. 1 330 028, 2013. DOI: 10.1142/S0217751X13300287.

[54] J. Alexander *et al.*, "Dark Sectors 2016 Workshop: Community Report," 2016. arXiv: 1608.08632.

[55] F. Queiroz, "Dark Matter Overview: Collider, Direct and Indirect Detection Searches," in *Proceedings, 51st Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 12-19, 2016*, 2016, pp. 427–436. arXiv: 1605.08788.

[56] R. Abbasi *et al.*, "Search for dark matter annihilation in the center of the Earth with 8 years of IceCube data," *PoS*, vol. ICRC2019, p. 541, 2020. DOI: 10.22323/1.358.0541. arXiv: 1908.07255 [astro-ph.HE].

[57] M. Cirelli, "Indirect Searches for Dark Matter: a status review," in *Proceedings, 25th International Symposium on Lepton Photon Interactions at High Energy (LP11): Mumbai, India, August 22-27, 2011*, vol. 79, 2012, pp. 1021–1043. DOI: 10.1007/s12043-012-0419-x.

[58] J. Lewin and P. Smith, "Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil," *Astropart. Phys.*, vol. 6, p. 87, 1996. DOI: 10.1016/S0927-6505(96)00047-3.

[59] R. Essig, T. Volansky, and T.-T. Yu, "New constraints and prospects for sub-gev dark matter scattering off electrons in xenon," *Phys. Rev. D*, vol. 96, p. 043 017, 4 2017. DOI: 10.1103/PhysRevD.96.043017.

[60] T. Undagoitia and L. Rauch, "Dark matter direct-detection experiments," *J. Phys. G: Nucl. Part. Phys.*, vol. 43, no. 1, p. 013 001, 2015. DOI: 10.1088/0954-3899/43/1/013001.

[61] F. Duncan, A. Noble, and D. Sinclair, "The construction and anticipated science of snolab," *Annu. Rev. Nucl. Part. S.*, vol. 60, no. 1, pp. 163–180, 2010.

[62] F. Piquemal, "Modane underground laboratory: Status and project," *Eur. Phys. J. Plus*, vol. 127, no. 1, pp. 110–114, 2012. DOI: 10.1140/epjp/i2012-12110-3.

[63] Q. Arnaud *et al.*, "First results from the NEWS-G direct dark matter search experiment at the LSM," *Astropart. Phys.*, vol. 97, p. 54, 2018. DOI: 10.1016/J.ASTROPARTPHYS.2017.10.009.

[64] F. Aubin *et al.*, "Discrimination of nuclear recoils from alpha particles with superheated liquids," *New J. Phys.*, vol. 10, no. 10, p. 103 017, 2008. DOI: 10.1088/1367-2630/10/10/103017.

[65] C. Amole *et al.*, "Data-driven modeling of electron recoil nucleation in PICO C_3F_8 bubble chambers," *Phys. Rev. D*, vol. 100, p. 082 006, 8 2019. DOI: 10.1103/PhysRevD.100.082006.

[66] R. Schnee, "Introduction to dark matter experiments," *Proc. TASI09*, pp. 775–829, 2011. DOI: 10.1142/9789814327183_0014.

[67] D. Baxter *et al.*, "Recommended conventions for reporting results from direct dark matter searches," *Eur. Phys. J. C*, vol. 81, no. 10, p. 907, 2021. DOI: 10.1140/epjc/s10052-021-09655-y. arXiv: 2105.00599.

[68] Cerdeño, D.G. and Green, A.M., "Direct detection of WIMPs," in *Particle Dark Matter: Observations, Models and Searches*, G. Bertone, Ed., Cambridge, UK: Cambridge University Press, 2010, p. 347, ISBN: 9780521763684.

[69] M. C. Smith *et al.*, "The RAVE survey: constraining the local Galactic escape speed," *Mon. Not. R. Astron. Soc.*, vol. 379, no. 2, p. 755, 2007. DOI: 10.1111/j.1365-2966.2007.11964.x.

[70] W. Cottingham and D. Greenwood, *An Introduction to Nuclear Physics*. Cambridge University Press, 2001, ISBN: 9780521651493.

[71] D. Tovey, R. Gaitskell, P. Gondolo, Y. Ramachers, and L. Roszkowski, "A new model-independent method for extracting spin-dependent cross section limits from dark matter searches," *Phys. Lett. B*, vol. 488, no. 1, pp. 17–26, 2000. DOI: https://doi.org/10.1016/S0370-2693(00)00846-7.

[72] R. Agnese *et al.*, "Silicon Detector Dark Matter Results from the Final Exposure of CDMS II," *Phys. Rev. Lett.*, vol. 111, p. 251 301, 25 2013. DOI: 10.1103/PhysRevLett.111.251301.

[73] G. Cowan, *Statistical Data Analysis*, 1st ed. Oxford: Clarendon Press, 1998, ISBN: 9780198501558.

[74] P. Adari *et al.*, "EXCESS workshop: Descriptions of rising low-energy spectra," 2022. arXiv: 2202.05097.

[75] [Online]. Available: https://root.cern/doc/v614/classTMVA_1_1MethodBDT.html.

[76] E. Armengaud *et al.*, "Constraints on low-mass WIMPs from the EDELWEISS-III dark matter search," *J. Cosmol. Astropart. Phys.*, vol. 2016, no. 05, pp. 019–019, 2016. DOI: 10.1088/1475-7516/2016/05/019.

[77] S. Yellin, "Finding an upper limit in the presence of an unknown background," *Phys. Rev. D*, vol. 66, p. 032 005, 3 2002.

[78] E. Armengaud *et al.*, "Searching for low-mass dark matter particles with a massive Ge bolometer operated above ground," *Phys. Rev. D*, vol. 99, p. 082 003, 8 2019. DOI: 10.1103/PhysRevD.99.082003.

[79] E. Aprile *et al.*, "Light Dark Matter Search with Ionization Signals in XENON1T," *Phys. Rev. Lett.*, vol. 123, no. 25, p. 251 801, 2019. DOI: 10.1103/PhysRevLett.123.251801. arXiv: 1907.11485 [hep-ex].

[80] S. Yellin, *Extending the optimum interval method*, 2007. arXiv: 0709.2701 [physics.data-an].

[81] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J.*, vol. C71, p. 1554, 2011, [Erratum: Eur. Phys. J.C73,2501(2013)]. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727.

[82] S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Ann. Math. Stat.*, vol. 9, no. 1, pp. 60 –62, 1938. DOI: 10.1214/aoms/1177732360.

[83] E. Aprile *et al.*, "Likelihood approach to the first dark matter results from XENON100," *Phys. Rev. D*, vol. 84, no. 5, 2011. DOI: 10.1103/physrevd.84.052003.

[84] E. Baracchini *et al.*, "CYGNO: A gaseous TPC with optical readout for dark matter directional search," *J. Inst.*, vol. 15, no. 07, pp. C07036–C07036, 2020. DOI: 10.1088/1748-0221/15/07/c07036.

[85] S. Vahsen *et al.*, *CYGNUS: Feasibility of a nuclear recoil observatory with directional sensitivity to dark matter and neutrinos*, 2020. arXiv: 2008.12587 [physics.ins-det].

[86] J. Battat *et al.*, "First measurement of nuclear recoil head-tail sense in a fiducialised WIMP dark matter detector," *J. Inst.*, vol. 11, no. 10, P10019–P10019, 2016. DOI: 10.1088/1748-0221/11/10/p10019.

[87] J. Castel *et al.*, "The TREX-DM experiment at the Canfranc Underground Laboratory," *J. Phys.: Conf. Ser.*, vol. 1468, no. 1, p. 012 063, 2020. DOI: 10.1088/1742-6596/1468/1/012063.

[88] J. Battat *et al.*, "Low threshold results and limits from the DRIFT directional dark matter detector," *Astropart. Phys.*, vol. 91, pp. 65–74, 2017. DOI: 10.1016/j.astropartphys.2017.03.007.

[89] J. Billard, E. Figueroa-Feliciano, and L. Strigari, "Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments," *Phys. Rev. D*, vol. 89, p. 023 524, 2 2014. DOI: 10.1103/PhysRevD.89.023524.

[90] D. Durnford, *Phenomenological studies and analysis techniques to search for light dark matter with NEWS-G*, 2018. [Online]. Available: https://qspace.library.queensu.ca/handle/1974/24878.

[91] C. O'Hare, "New definition of the neutrino floor for direct dark matter searches," *Phys. Rev. Lett.*, vol. 127, no. 25, 2021. DOI: 10.1103/physrevlett.127.251802.

[92] P. Grothaus, M. Fairbairn, and J. Monroe, "Directional dark matter detection beyond the neutrino bound," *Phys. Rev. D*, vol. 90, no. 5, 2014. DOI: 10.1103/physrevd.90.055018.

[93] C. Amole *et al.*, "Dark matter search results from the complete exposure of the PICO-60 C_3F_8 bubble chamber," *Phys. Rev. D*, vol. 100, p. 022 001, 2 2019. DOI: 10.1103/PhysRevD.100.022001.

[94] P. Giampa, "The Scintillating Bubble Chamber (SBC) Experiment for Dark Matter and Reactor CEvNS," *PoS*, vol. ICHEP2020, p. 632, 2021. DOI: 10.22323/1.390.0632.

[95] E. Aprile *et al.*, "Dark Matter Search Results from a One Ton-Year Exposure of XENON1T," *Phys. Rev. Lett.*, vol. 121, p. 111 302, 11 2018. DOI: 10.1103/PhysRevLett.121.111302.

[96] J. Aalbers *et al.*, "First dark matter search results from the lux-zeplin (lz) experiment," *Phys. Rev. Lett.*, vol. 131, p. 041 002, 4 2023. DOI: 10.1103/PhysRevLett.131.041002.

[97] Y. Meng *et al.*, "Dark matter search results from the PandaX-4t commissioning run," *Phys. Rev. Lett.*, vol. 127, no. 26, 2021. DOI: 10.1103/physrevlett.127.261802.

[98] E. Aprile *et al.*, "First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment," *Phys. Rev. Lett.*, vol. 131, p. 041 003, 4 2023. DOI: 10.1103/PhysRevLett.131.041003.

[99] M. Cadeddu and E. Picciau, "Impact of neutrino background prediction for next generation dark matter xenon detector," *J. Phys.: Conf. Ser.*, vol. 956, no. 1, p. 012 014, 2018. DOI: 10.1088/1742-6596/956/1/012014.

[100] P. Agnes *et al.*, "First results from the DarkSide-50 dark matter experiment at laboratori nazionali del gran sasso," *Phys. Lett. B*, vol. 743, pp. 456–466, 2015. DOI: 10.1016/j.physletb.2015.03.012.

[101] P.-A. Amaudruz *et al.*, "First results from the DEAP-3600 dark matter search with argon at SNOLAB," *Phys. Rev. Lett.*, vol. 121, no. 7, 2018. DOI: 10.1103/physrevlett.121.071801.

[102] C. Aalseth *et al.*, "DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS," *Eur. Phys. J. Plus*, vol. 133, no. 3, 2018. DOI: 10.1140/epjp/i2018-11973-4.

[103] C. Galbiata, *Global Argon Dark Matter collaboration, "Future dark matter searches with low-radioactivity argon*, 2018. [Online]. Available: https://indico.cern.ch/event/765096/contributions/3295671/attachments/1785196/2906164/DarkSide-Argo_ESPP_Dec_17_2017.pdf.

[104] P. Agnes *et al.*, "Low-Mass Dark Matter Search with the DarkSide-50 Experiment," *Phys. Rev. Lett.*, vol. 121, p. 081 307, 2018. DOI: 10.1103/PhysRevLett.121.081307.

[105] P. Adhikari *et al.*, "Pulse-shape discrimination against low-energy Ar-39 beta decays in liquid argon with 4.5 tonne-years of DEAP-3600 data," *Eur. Phys. J. C*, vol. 81, p. 823, 2021. DOI: https://doi.org/10.1140/epjc/s10052-021-09514-w.

[106] R. Ajaj *et al.*, "Search for dark matter with a 231-day exposure of liquid argon using DEAP-3600 at SNOLAB," *Phys. Rev. D*, vol. 100, p. 022 004, 2 2019. DOI: 10.1103/PhysRevD.100.022004.

[107] J. Aalbers *et al.*, "A next-generation liquid xenon observatory for dark matter and neutrino physics," *J. Phys. G: Nucl. Part. Phys.*, vol. 50, no. 1, p. 013 001, 2022. DOI: 10.1088/1361-6471/ac841a.

[108] Q. Arnaud *et al.*, "Optimizing edelweiss detectors for low-mass wimp searches," *Phys. Rev. D*, vol. 97, p. 022 003, 2 2018. DOI: 10.1103/PhysRevD.97.022003.

[109] R. Agnese *et al.*, "Search for low-mass dark matter with CDMSlite using a profile likelihood fit," *Phys. Rev. D*, vol. 99, p. 062 001, 6 2019. DOI: 10.1103/PhysRevD.99.062001.

[110] R. Agnese *et al.*, "Low-mass dark matter search with CDMSlite," *Phys. Rev. D*, vol. 97, p. 022 002, 2 2018. DOI: 10.1103/PhysRevD.97.022002.

[111] R. Agnese *et al.*, "Projected sensitivity of the SuperCDMS SNOLAB experiment," *Phys. Rev. D*, vol. 95, p. 082 002, 8 2017. DOI: 10.1103/PhysRevD.95.082002.

[112] C. Isaila *et al.*, "Low-temperature light detectors: Neganov–Luke amplification and calibration," *Phys. Lett. B*, vol. 716, no. 1, pp. 160–164, 2012. DOI: 10.1016/j.physletb.2012.08.003.

[113] R. Agnese *et al.*, "First Dark Matter Constraints from a SuperCDMS Single-Charge Sensitive Detector," *Phys. Rev. Lett.*, vol. 121, p. 051 301, 2018. DOI: 10.1103/PhysRevLett.121.051301.

[114] D. W. Amaral *et al.*, "Constraints on low-mass, relic dark matter candidates from a surface-operated supercdms single-charge sensitive detector," *Phys. Rev. D*, vol. 102, p. 091 101, 9 2020. DOI: 10.1103/PhysRevD.102.091101.

[115] H. Lattaud *et al.*, "Phonon and charge signals from IR and x excitation in the SELENDIS ge cryogenic detector," *J. Low Temp. Phys.*, vol. 209, no. 3-4, pp. 263–270, 2022. DOI: 10.1007/s10909-022-02826-5.

[116] [Online]. Available: https://cresst-experiment.org/the-cresst-experiment/cresst-detectors.

[117] A. Abdelhameed *et al.*, "First results on sub-GeV spin-dependent dark matter interactions with 7Li," *European Physical Journal C*, vol. 79, 7 2019. DOI: 10.1140/epjc/s10052-019-7126-4.

[118] G. Angloher *et al.*, "Probing spin-dependent dark matter interactions with 6Li: CRESST collaboration," *European Physical Journal C*, vol. 82, 3 2022, ISSN: 14346052. DOI: 10.1140/epjc/s10052-022-10140-3.

[119] J. de Mello Neto *et al.*, "The DAMIC dark matter experiment," *Proc. Sci.*, vol. ICRC2015, p. 1221, 2016. eprint: 1510.02126.

[120] J. Tiffenberg *et al.*, "Single-Electron and Single-Photon Sensitivity with a Silicon Skipper CCD," *Phys. Rev. Lett.*, vol. 119, p. 131 802, 2017. DOI: 10.1103/PhysRevLett.119.131802.

[121] A. Aguilar-Arevalo *et al.*, "The oscura experiment," 2022. arXiv: 2202.10518 [astro-ph.IM].

[122] I. Arnquist *et al.*, "The DAMIC-M Experiment: Status and First Results," 2022. arXiv: 2210.12070 [hep-ex].

[123] L. Barak *et al.*, "SENSEI: Direct-Detection Results on sub-GeV Dark Matter from a New Skipper CCD," *Phys. Rev. Lett.*, vol. 125, p. 171 802, 17 2020. DOI: 10.1103/PhysRevLett.125.171802.

[124] R. Bernabei *et al.*, "The dark matter: DAMA/LIBRA and its perspectives," 2021. arXiv: 2110.04734 [hep-ph].

[125] G. Adhikari *et al.*, "Strong constraints from COSINE-100 on the DAMA dark matter results using the same sodium iodide target," *Science Advances*, vol. 7, no. 46, eabk2699, 2021. DOI: 10.1126/sciadv.abk2699.

[126] M. Antonello *et al.*, "Monte Carlo simulation of the SABRE PoP background," *Atropart. Phys.*, vol. 106, pp. 1–9, 2019. DOI: https://doi.org/10.1016/j.astropartphys.2018.10.005.

[127] G. Adhikari *et al.*, "Three-year annual modulation search with COSINE-100," *Phys. Rev. D*, vol. 106, p. 052 005, 5 2022. DOI: 10.1103/PhysRevD.106.052005.

[128] J. Amaré *et al.*, "Annual modulation results from three-year exposure of ANAIS-112," *Phys. Rev. D*, vol. 103, no. 10, 2021. DOI: 10.1103/physrevd.103.102005.

[129] D. Ferenc, D. Šegedin, I. Šegedin, and M. Šegedin Ferenc, "Helium Migration through Photomultiplier Tubes – The Probable Cause of the DAMA Seasonal Variation Effect," 2019. arXiv: 1901.02139 [physics.ins-det].

[130] D. Buttazzo, P. Panci, N. Rossi, and A. Strumia, "Annual modulations from secular variations: Relaxing DAMA?" *JHEP*, vol. 2020, no. 4, 2020. DOI: 10.1007/jhep04(2020)137.

[131] G. Adhikari *et al.*, "An induced annual modulation signature in COSINE-100 data by DAMA/LIBRA's analysis method," *Sci. Rep.*, vol. 13, no. 1, 2023. DOI: 10.1038/s41598-023-31688-4.

[132] R. Maruyama, "Resolving DAMA," Presented at the 2023 UCLA DM Conference, Los Angeles, USA, 2023. [Online]. Available: https://indico.cern.ch/event/1188759/contributions/5238459/attachments/2622531/4534688/UCLA_DM2023_Maruyama_fin.pdf.

[133] A. Abdelhameed *et al.*, "First results from the CRESST-III low-mass dark matter program," *Phys. Rev. D*, vol. 100, p. 102 002, 10 2019. DOI: 10.1103/PhysRevD.100.102002.

[134] M. Ibe *et al.*, "Migdal Effect in Dark Matter Direct Detection Experiments," *JHEP*, vol. 03, p. 194, 2018. DOI: 10.1007/JHEP03(2018)194.

[135] T. Marley, "Current status of preparations for the MIGDAL experiment," Presented at the RD51 collaboration mini-week, 2021. [Online]. Available: https://indico.cern.ch/event/989298/contributions/4225239/.

[136] J. Xu *et al.*, *Search for the Migdal effect in liquid xenon with keV-level nuclear recoils*, 2023. arXiv: 2307.12952 [hep-ex].

[137] E. Aprile *et al.*, "Excess electronic recoil events in XENON1T," *Phys. Rev. D*, vol. 102, p. 072 004, 7 2020. DOI: 10.1103/PhysRevD.102.072004.

[138] A. Aguilar-Arevalo *et al.*, "Results on Low-Mass Weakly Interacting Massive Particles from an 11 kg d Target Exposure of DAMIC at SNOLAB," *Phys. Rev. Lett.*, vol. 125, p. 241 803, 24 2020. DOI: 10.1103/PhysRevLett.125.241803.

[139] N. Kurinsky, D. Baxter, Y. Kahn, and G. Krnjaic, "Dark matter interpretation of excesses in multiple direct detection experiments," *Phys. Rev. D*, vol. 102, p. 015 017, 1 2020. DOI: 10.1103/PhysRevD.102.015017.

[140] D. Durnford, Q. Arnaud, and G. Gerbier, "Novel approach to assess the impact of the fano factor on the sensitivity of low-mass dark matter experiments," *Phys. Rev. D*, vol. 98, no. 10, 2018. DOI: 10.1103/physrevd.98.103013.

[141] K. Sellers, S. Borle, and G. Shmueli, "The COM-Poisson model for count data: A survey of methods and applications," *Appl. Stoch. Models Bus. Ind.*, vol. 28, no. 2, p. 104, 2011. DOI: 10.1002/asmb.918.

[142] D. Durnford, "The COM-Poissondistribution for modellingionization statistics," Presented at the 2019 PHYSTAT-DM Conference, Stockholm, Sweden), 2019. [Online]. Available: https://indico.cern.ch/event/769726/contributions/3479756/attachments/1889004/3114803/DDurnford_PHYSTAT.pdf.

[143] F. Sauli, *Gaseous Radiation Detectors: Fundamentals and Applications*. Cambridge, England: Cambridge University Press, 2014. DOI: 10.1017/CBO9781107337701.

[144] E. Rutherford and H. Geiger, "An electrical method of counting the number of α–particles from radio–active sources," *Pro. Royal Soc. A*, vol. 81, 1908.

[145] C. Thomas, Ed., *The Physics of Radiology*, 4th ed. 1983.

[146] K. Morishima *et al.*, "Discovery of a big void in Khufu's Pyramid by observation of cosmic-ray muons," *Nature*, vol. 552, 2017. DOI: https://doi.org/10.1038/nature24647.

[147] D. Sim. (2012), [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/archive/f/f7/20120606214516%21Detector_regions.gif.

[148] H. Geiger and W. Müller, "Das Elektronenzählrohr," *Phys. Zetis.*, vol. 29, 1928.

[149] F. Sauli, "Principles of Operation of Multiwire Proportional and Drift Chambers," 1977. [Online]. Available: https://cds.cern.ch/record/117989.

[150] I. Giomataris *et al.*, "NOSTOS experiment and new trends in rare event detection," *Nucl. Phys. B Proc. Suppl.*, vol. 150, C. Bosio, P. S. Marrocchesi, F. L. Navarria, M. Paganoni, and P. G. Pelfer, Eds., pp. 208–213, 2006. DOI: 10.1016/j.nuclphysbps.2005.01.245. arXiv: hep-ex/0502033.

[151] I. Giomataris *et al.*, "A novel large-volume spherical detector with proportional amplification read-out," *J. Inst.*, vol. 3, no. 09, P09007–P09007, 2008. DOI: 10.1088/1748-0221/3/09/p09007.

[152] A. Brossard, "Optimization of spherical proportional counter backgrounds and response for low mass dark matter search," Ph.D. dissertation, Queen's University, 2020.

[153] J. Derre, "Pulse shape in the SPC prototype," Tech. Rep., 2007. [Online]. Available: https://newsgorg.files.wordpress.com/2018/07/pulse-shape-in-the-spc-prototype-by-j-derrc3a9.pdf.

[154] "COMSOL Multiphysics v. 6.0., COMSOL AB, Stockholm, Sweden." (), [Online]. Available: https://www.comsol.com/.

[155] A. Brossard, "Spherical proportional counters; development, improvement and understanding," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 936, pp. 412–415, 2019, Frontier Detectors for Frontier Physics: 14th Pisa Meeting on Advanced Detectors. DOI: https://doi.org/10.1016/j.nima.2018.11.037.

[156] I. Katsioulas *et al.*, "A sparkless resistive glass correction electrode for the spherical proportional counter," *J. Inst.*, vol. 13, no. 11, P11006–P11006, 2018. DOI: 10.1088/1748-0221/13/11/p11006.

[157] C. Jollet, "A new neutrinoless double beta decay experiment: R2D2," *J. Phys.: Conf. Ser.*, vol. 1468, no. 1, p. 012 108, 2020. DOI: 10.1088/1742-6596/1468/1/012108.

[158] M. Druyvesteyn and F. Penning, "The Mechanism of Electrical Discharges in Gases of Low Pressure," *Rev. Mod. Phys.*, vol. 12, pp. 87–174, 2 1940. DOI: 10.1103/RevModPhys.12.87.

[159] O. Şahin, I. Tapan, E. Özmutlu, and R. Veenhof, "Penning transfer in argon-based gas mixtures," *J. Inst.*, vol. 5, no. 05, P05002, 2010. DOI: 10.1088/1748-0221/5/05/P05002.

[160] F. A. Vazquez de Sola Fernandez, "Solar KK axion search with NEWS-G," Ph.D. dissertation, Queen's University, 2020.

[161] A. Dastgheibi Fard, "Study of a spherical gaseous detector for research of rare events at low energy threshold," Ph.D. dissertation, Universitè Parid Sud, 2014. [Online]. Available: https://theses.hal.science/tel-01134076.

[162] [Online]. Available: https://redpitaya.com/.

[163] E. Armengaud *et al.*, "Performance of the EDELWEISS-III experiment for direct dark matter searches," *J. Inst.*, vol. 12, no. 08, P08010, 2017. DOI: 10.1088/1748-0221/12/08/P08010.

[164] E. Armengaud *et al.*, "The CUPID-Mo experiment for neutrinoless double-beta decay: performance and prospects," *Eur. Phys. J. C*, vol. 80, no. 44, 2019. DOI: 10.1140/epjc/s10052-019-7578-6.

[165] M. Inokuti, "Ionization Yields in Gases under Electron Irradiation," *Radiat. Res.*, vol. 64, no. 1, p. 6, 1975. DOI: 10.2307/3574165.

[166] *Atomic and Molecular Data for Radiotherapy and Radiation Research* (TECDOC Series 799). Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 1995. [Online]. Available: https://www.iaea.org/publications/5444/atomic-and-molecular-data-for-radiotherapy-and-radiation-research.

[167] H. Bichsel *et al.*, "ICRU Report 31," *J. ICRU*, vol. os-16, no. 2, 1979. DOI: 10.1093/jicru_os16.2.1.

[168] U. Fano, "Ionization yield of radiations. II. The fluctuations of the number of ions," *Phys. Rev.*, vol. 72, no. 1, p. 26, 1947. DOI: 10.1103/PhysRev.72.26.

[169] A. Hashiba, K. Masuda, T. Doke, T. Takahashi, and Y. Fujita, "Fano factor in gaseous argon measured by the proportional scintillation method," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 227, no. 2, p. 305, 1984. DOI: https://doi.org/10.1016/0168-9002(84)90138-4.

[170] A. Policarpo, M. Alves, M. Salete, S. Leite, and M. dos Santos, "Detection of soft X-rays with a xenon proportional scintillation counter," *Nucl. Instrum. Methods*, vol. 118, no. 1, p. 221, 1974, ISSN: 0029-554X. DOI: https://doi.org/10.1016/0029-554X(74)90706-X.

[171] A. Owens, G. Fraser, and K. McCarthy, "On the experimental determination of the Fano factor in Si at soft X-ray wavelengths," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 491, no. 3, p. 437, 2002, ISSN: 0168-9002. DOI: 10.1016/S0168-9002(02)01178-6.

[172] B. Lowe, "Measurements of Fano factors in silicon and germanium in the low-energy X-ray region," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 399, no. 2, p. 354, 1997. DOI: https://doi.org/10.1016/S0168-9002(97)00965-0.

[173] B. Grosswendt, "Statistical fluctuations of the ionisation yield of low-energy electrons in He, Ne and Ar," *J. Phys. B*, vol. 17, no. 7, p. 1391, 1984.

[174] R. Conway and W. Maxwell, "A queuing model with state dependent service rates," *J. Ind. Eng.*, vol. 12, p. 132, 1962.

[175] G. Shmueli, T. Minka, J. Kadane, S. Borle, and P. Boatwright, "A useful distribution for fitting discrete data: Revival of the Conway-Maxwell-Poisson distribution," *J. R. Stat. Soc. C*, vol. 54, no. 1, p. 127, 2004. DOI: 10.1111/j.1467-9876.2005.00474.x.

[176] [Online]. Available: https://www.queensu.ca/physics/news-g/com-poisson-code.

[177] J. Lindhard, V. Nielsen, M. Scharff, and P. Thomsen, "Integral equations governing radiation effects. (notes on atomic collisions, iii)," *Kgl. Danske Videnskab., Selskab. Mat. Fys. Medd.*, vol. 33, no. 10, 1963.

[178] D. Akerib *et al.*, "Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics," 2016. arXiv: 1608.05381 [physics.ins-det].

[179] M. Szydagis *et al.*, *Noble Element Simulation Technique*, version v2.2.0beta, Nov. 2020. DOI: 10.5281/zenodo.4262416.

[180] O. Guillaudin *et al.*, "Quenching factor measurement in low pressure gas detector for directional dark matter search," in *EAS Publications Series*, vol. 53, 2012, p. 119. DOI: 10.1051/eas/1253015.

[181] D. Santos *et al.*, "Ionization Quenching Factor Measurement of Helium 4," 2008. arXiv: 0810.1137 [astro-ph].

[182] B. Tampon *et al.*, "Ionization Quenching Factor measurement of 1 keV to 25 keV protons in Isobutane gas mixture," *EPJ Web Conf.*, vol. 153, p. 01 014, 2017. DOI: 10.1051/epjconf/201715301014.

[183] J. Ziegler and J. Biersack, *SRIM (The Stopping and Range of Ions in Matter)*, 2008.

[184] L. Balogh *et al.*, "Quenching factor measurements of neon nuclei in neon gas," *Phys. Rev. D*, vol. 105, p. 052 004, 5 2022. DOI: 10.1103/PhysRevD.105.052004.

[185] B. Sitar, G. I. Merson, V. A. Chechin, and Y. A. Budagov, *Ionization Measurements in High Energy Physics*, 1st ed. Springer-Verlag Berlin Heidelberg, 1988, p. 338, ISBN: 978-3-662-14923-2.

[186] S. F. Biagi, "Monte Carlo simulation of electron drift and diffusion in counting gases under the influence of electric and magnetic fields," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 421, no. 1, p. 234, 1999.

[187] C. Garrah, "On Advances in LAS Instrumentation and Fiducial Volume Simulations of the S140 Detector for the NEWS-G Dark Matter Search Experiment," Ph.D. dissertation, University of Alberta, 2022. [Online]. Available: https://www.snolab.ca/news-projects/private/TWiki/bin/view/Main/NewsTheses.

[188] J. Derré *et al.*, "Fast signals and single electron detection with a MICROMEGAS photodetector," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 449, no. 1, p. 314, 2000. DOI: https://doi.org/10.1016/S0168-9002(99)01452-7.

[189] T. Zerguerras *et al.*, "Single-electron response and energy resolution of a Micromegas detector," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 608, no. 3, p. 397, 2009. DOI: https://doi.org/10.1016/j.nima.2009.07.015.

[190] M. Kobayashi *et al.*, "A novel technique for the measurement of the avalanche fluctuation of gaseous detectors," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 845, p. 236, 2017. DOI: https://doi.org/10.1016/j.nima.2016.06.073.

[191] R. Bellazzini *et al.*, "Imaging with the invisible light," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 581, no. 1, p. 246, 2007. DOI: https://doi.org/10.1016/j.nima.2007.07.098.

[192] Q. Arnaud *et al.*, "Precision laser-based measurements of the single electron response of spherical proportional counters for the news-g light dark matter search experiment," *Phys. Rev. D*, vol. 99, p. 102 003, 10 2019. DOI: 10.1103/PhysRevD.99.102003.

[193] J. Weiss and W. Bernstein, "The Current Status of W, the Energy to Produce One Ion Pair in a Gas," *Radiat. Res.*, vol. 6, no. 6, pp. 603–610, 1957. [Online]. Available: http://www.jstor.org/stable/3570413.

[194] [Online]. Available: https://www.qlinstruments.com/.

[195] S. Sangiorgio *et al.*, "First demonstration of a sub-keV electron recoil energy threshold in a liquid argon ionization chamber," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 728, p. 69, 2013. DOI: https://doi.org/10.1016/j.nima.2013.06.061.

[196] F. Kelly *et al.*, "The production of Ar-37 using a thermal neutron reactor flux," *J. Radioanal. Nucl. Ch.*, vol. 318, no. 1, p. 279, 2018. DOI: 10.1007/s10967-018-6130-8.

[197] X. Mougeot, "Towards high-precision calculation of electron capture decays," *Appl. Radiat. Isotopes*, vol. 154, p. 108 884, 2019. DOI: https://doi.org/10.1016/j.apradiso.2019.108884.

[198] D. Cullen, "Program RELAX: A code designed to calculate atomic relaxation spectra of x-rays and electrons," 1992. DOI: 10.2172/5360235.

[199] A. Ringbom, "Calculation of electron–photon coincidence decay of ^{131m}Xe and ^{133m}Xe including atomic relaxation," *Appl. Radiat. Isotopes*, vol. 70, no. 8, pp. 1499–1508, 2012. DOI: https://doi.org/10.1016/j.apradiso.2012.04.012.

[200] P. Agnes *et al.*, "Calibration of the liquid argon ionization response to low energy electronic and nuclear recoils with DarkSide-50," *Phys. Rev. D*, vol. 104, p. 082 005, 8 2021. DOI: 10.1103/PhysRevD.104.082005.

[201] J. Rode, "DarkSide-50: characterisation of the LAr ionization response in the keV regime," Presented at the 2021 GDR Deep Underground Physics plenary meeting, Paris France, 2021. [Online]. Available: https://indico.in2p3.fr/event/25051/contributions/101853/.

[202] Private exchange with D. Franco (Laboratoire Astroparticule et Cosmologie) of the DarkSide collaboration, about their ^{37}Ar electron capture simulation.

[203] D. Cullen, *A Survey of Atomic Binding Energies for use in EPICS2017*, 2018. [Online]. Available: https://www-nds.iaea.org/publications/nds/iaea-nds-0224/.

[204] S. Agostinelli *et al.*, "Geant4 - a simulation toolkit," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 506, no. 3, p. 250, 2003. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8.

[205] *Shielding Physics List Description*, 2012. [Online]. Available: http://www.slac.stanford.edu/comp/physics/geant4/slac_physics_lists/shielding/physlistdoc.html.

[206] B. Smith and J. Booz, "Experimental results on W-values and transmission of low-energy electrons in gases," Proc. Sixth Symp. Microdosimetry, 1978.

[207] D. Srdoč, B. Obelić, and I. Krajcar-Bronić, "Statistical fluctuations in the ionization yield for low-energy photons, absorbed in polyatomic gases," *J. Phys. B Atom Molec. Phys.*, vol. 20, pp. 4473–4484, 1987.

[208] E. Waibel and B. Grosswendt, "Spatial energy dissipation profiles, W values, backscatter coefficients, and ranges for low-energy electrons in methane," *Nucl. Instrumen. & Meth. Phys. Res.*, vol. 211, no. 2, pp. 487–498, 1983. DOI: https://doi.org/10.1016/0167-5087(83)90278-8.

[209] D. Combecher, "Measurement of w values of low-energy electrons in several gases," *Rad. Res.*, vol. 84, no. 2, pp. 189–218, 1980. [Online]. Available: https://www.jstor.org/stable/3575293.

[210] [Online]. Available: https://www.mirion.com/products/technologies/spectroscopy-scientific-analysis/gamma-spectroscopy/detector-electronics/2006-proportional-counter-preamplifier/.

[211] M. Basunia, "Nuclear data sheets for $A = 241$," *Nucl. Data Sheets*, vol. 107, p. 3323, 2006.

[212] B. Silverman, *Density Estimation for Statistics and Data Analysis*, 1st ed. 1998. DOI: 10.1201/9781315140919.

[213] H. Shimazaki and S. Shinomoto, "Kernel bandwidth optimization in spike rate estimation," *J. Comput. Neurosci.*, vol. 29, pp. 171–182, 2010. DOI: https://doi.org/10.1007/s10827-009-0180-4.

[214] M. Thorben and M. Walmsley, *Adaptive Width KDE with Gaussian Kernels*, 2022. [Online]. Available: https://github.com/mennthor/awkde.

[215] B. Wang and X. Wang, *Bandwidth selection for weighted kernel density estimation*, 2011. arXiv: 0709.1616 [stat.ME].

[216] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC Hammer," *Publ. Astron. Soc. Pac.*, vol. 125, pp. 306–312, 2013. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph.IM].

[217] J. Scargle, "Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data.," *Astrophys. J.*, vol. 263, pp. 835–853, Dec. 1982. DOI: 10.1086/160554.

[218] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*. Cambridge University Press, 2005. DOI: 10.1017/CBO9780511791277.

[219] [Online]. Available: https://www.caen.it/subfamilies/up-to-6-kv-dt55xxe/.

[220] [Online]. Available: https://www.cremat.com/.

[221] P. O'Brien, "Optimization of Processing Parameters and Development of a Radon Trapping System for the NEWS-G Dark Matter Detector," Ph.D. dissertation, University of Alberta, 2021. [Online]. Available: https://era.library.ualberta.ca/items/2a4b3045-5791-442d-9235-537f2f98e08d.

[222] W. Morrish, "Sensing with Optical Microresonators and Integrating Cavities," Ph.D. dissertation, University of Alberta, 2021. [Online]. Available: https://era.library.ualberta.ca/items/9bc5cb4d-8965-4809-98f8-5e2310b71621.

[223] A. Fried and D. Richter, "Infrared Absorption Spectroscopy," in *Analytical Techniques for Atmospheric Measurement*. John Wiley & Sons, Ltd, 2006, ch. 2, pp. 72–146, ISBN: 9780470988510. DOI: https://doi.org/10.1002/9780470988510.ch2.

[224] [Online]. Available: https://www.liquidinstruments.com/products/hardware-platforms/mokulab/.

[225] K. S. Toth, "Nuclear data sheets for A = 222," *Nucl. Data Sheets*, vol. 21, pp. 479–492, 1977. DOI: 10.1016/S0090-3752(77)80028-8.

[226] B. Singh *et al.*, "Nuclear Data Sheets for $A = 218$," *Nucl. Data Sheets*, vol. 160, pp. 405–471, 2019. DOI: 10.1016/j.nds.2019.100524.

[227] S. Zhu and E. A. McCutchan, "Nuclear Data Sheets for A=214," *Nucl. Data Sheets*, vol. 175, pp. 1–149, 2021. DOI: 10.1016/j.nds.2021.06.001.

[228] P. Giampa, "On the DEAP-3600 resurfacing," *AIP Conf. Proc.*, vol. 1921, no. 1, p. 070 005, 2018. DOI: 10.1063/1.5019008.

[229] S. Bruenner *et al.*, "Radon daughter removal from PTFE surfaces and its application in liquid xenon detectors," *Eur. Phys. J. C*, vol. 81, Apr. 2021. DOI: 10.1140/epjc/s10052-021-09047-2.

[230] E. Aprile *et al.*, "Online ^{222}Rn removal by cryogenic distillation in the XENON100 experiment," *Eur. Phys. J. C*, vol. 77, 6 2017. DOI: 10.1140/epjc/s10052-017-4902-x.

[231] K. Abe *et al.*, "Radon removal from gaseous xenon with activated charcoal," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 661, no. 1, pp. 50–57, 2012. DOI: https://doi.org/10.1016/j.nima.2011.09.051.

[232] B. Cai, M. Boulay, B. Cleveland, and T. Pollmann, "Surface backgrounds in the DEAP-3600 dark matter experiment," *AIP Conf. Proc.*, vol. 1338, no. 1, pp. 137–146, 2011. DOI: 10.1063/1.3579572.

[233] B. Cohen and E. Cohen, "Theory and practice of radon monitoring with charcoal adsorption," *Health Phys.*, vol. 45, no. 2, pp. 501–508, 1983.

[234] L. Guo *et al.*, "The temperature dependence of adsorption coefficients of ^{222}Rn on activated charcoal: an experimental study," *Appl Radiat Isot*, vol. 125, pp. 185–187, 2017.

[235] S. Kuznicki *et al.*, "Xenon adsorption on modified ETS-10," *J. Phys. Chem. C*, vol. 111, no. 4, pp. 1560–1562, 2007.

[236] I. Bikit *et al.*, "Radon adsorption by zeolite," *Radiat. Meas.*, vol. 72, pp. 70–74, 2015. DOI: https://doi.org/10.1016/j.radmeas.2014.12.001.

[237] M. Shamsuzzoha Basunia, "Nuclear Data Sheets for A = 210," *Nucl. Data Sheets*, vol. 121, pp. 561–694, 2014. DOI: 10.1016/j.nds.2014.09.004.

[238] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nat. Methods*, vol. 17, pp. 261–272, 2020. DOI: 10.1038/s41592-019-0686-2.

[239] F. Gao and L. Han, "Implementing the Nelder-Mead simplex algorithm with adaptive parameters," *Comput. Optim. App.*, vol. 51, no. 1, pp. 259–277, 2012.

[240] G. Feldman and R. Cousins, "Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D*, vol. 57, p. 3873, 1998. DOI: 10.1103/PhysRevD.57.3873.

[241] [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.UnivariateSpline.html.

[242] B. Efron, "Bootstrap Methods: Another Look at the Jackknife," *Ann. Stat.*, vol. 7, no. 1, pp. 1 –26, 1979. DOI: 10.1214/aos/1176344552.

[243] G. Angloher *et al.*, "Observation of a low energy nuclear recoil peak in the neutron calibration data of the CRESST-III experiment," *Phys. Rev. D*, vol. 108, no. 2, 2023. DOI: 10.1103/physrevd.108.022005.

[244] G. Soum-Sidikov *et al.*, "Study of collision and γ-cascade times following neutron-capture processes in cryogenic detectors," 2023. arXiv: 2305.10139 [physics.ins-det].

[245] H. Abele *et al.*, "Observation of a Nuclear Recoil Peak at the 100 eV Scale Induced by Neutron Capture," *Phys. Rev. Lett.*, vol. 130, p. 211 802, 21 2023. DOI: 10.1103/PhysRevLett.130.211802.

[246] K. Harris, A. Gevorgian, A. Biffl, and A. Villano, "Neutron capture-induced silicon nuclear recoils for dark matter and CEνNS," *Phys. Rev. D*, vol. 107, p. 076 026, 7 2023. DOI: 10.1103/PhysRevD.107.076026.

[247] J. I. Collar, "Applications of an $^{88}Y/Be$ photo-neutron calibration source to Dark Matter and Neutrino Experiments," *Phys. Rev. Lett.*, vol. 110, no. 21, p. 211 101, 2013. DOI: 10.1103/PhysRevLett.110.211101. arXiv: 1303.2686 [physics.ins-det].

[248] D. Durnford and M.-C. Piro, "The search for Light Dark Matter with NEWS-G," *J. Phys.: Conf. Ser.*, vol. 2156, no. 1, p. 012 059, 2021. DOI: 10.1088/1742-6596/2156/1/012059.

[249] N. Nosengo, "Roman ingots to shield particle detector," *Nature*, 2010. DOI: https://doi.org/10.1038/news.2010.186.

[250] A. Giganon *et al.*, "A multiball read-out for the spherical proportional counter," *J. Inst.*, vol. 12, no. 12, P12031, 2017. DOI: 10.1088/1748-0221/12/12/P12031.

[251] I. Giomataris *et al.*, "A resistive achinos multi-anode structure with dlc coating for spherical proportional counters," *JINST*, vol. 15, no. 11, P11023, 2020. DOI: 10.1088/1748-0221/15/11/P11023. [Online]. Available: https://dx.doi.org/10.1088/1748-0221/15/11/P11023.

[252] G. Savvidis, "Sensor Characterization and Fiducial Volume Studies for the NEWS-G Dark Matter Experiment," Ph.D. dissertation, Queen's University, 2023. [Online]. Available: https://qspace.library.queensu.ca/items/55d349eb-dc66-4eab-8ead-33941688eb45.

[253] J. Browne E. Tuli, "Nuclear Data Sheets for A = 60," *Nuclear Data Sheets*, vol. 114, no. 12, pp. 1849–2022, 2013, ISSN: 0090-3752. DOI: https://doi.org/10.1016/j.nds.2013.11.002.

[254] M. Basunia, "Nuclear data sheets for $A = 210$," *Nucl. Data Sheets*, vol. 121, p. 561, 2014. DOI: https://doi.org/10.1016/j.nds.2014.09.004.

[255] K. Abe *et al.*, "Identification of ^{210}Pb and ^{210}Po in the bulk of copper samples with a low-background alpha particle counter," *Nucl. Instrum. Methods Phys. Res., Sect. A*, vol. 884, p. 157, 2018.

[256] [Online]. Available: https://xia.com/products/ultralo-1800/.

[257] [Online]. Available: https://www.pnnl.gov/.

[258] H. Bateman, "Solution of a system of differential equations occurring in the theory of radioactive transformations," *Proceedings of the Cambridge Philosophical Society, Mathematical and physical sciences*, vol. 15, 1908.

[259] [Online]. Available: https://root.cern.ch/doc/master/classROOT_1_1Minuit2_1_1Minuit2Minimizer.html.

[260] P. Knights, "Gas and copper purity investigations for NEWS-G," *J. Phys.: Conf. Ser.*, vol. 1312, no. 1, p. 012 009, 2019. DOI: 10.1088/1742-6596/1312/1/012009.

[261] P. Atkins, *Physical Chemisty*, 6th ed. W.H. Freeman and Company, 1997.

[262] D. Lide, *CRC Handbook of Chemistry and Physics*, 87th ed. Taylor & Francis, 2006. [Online]. Available: https://books.google.co.uk/books?id=yTTUQgAACAAJ.

[263] W. Haynes, *CRC Handbook of Chemistry and Physics*, 92nd ed. CRC Press, 2011. [Online]. Available: :https://books.google.co.uk/books?id=pYPRBQAAQBAJ.

[264] A. Bard, R. Parsons, and J. Jordan, *Standard Potentials in Aqueous Solution, Monographs in Electroanalytical Chemistry and Electrochemistry*. Taylor & Francis, 1985. [Online]. Available: https://books.google.co.uk/books?id=fuJV1H18KtEC.

[265] [Online]. Available: https://iseg-hv.com/en/?pk_campaign=google_hv_en.

[266] A. Ianni, "Considerations on Underground Laboratories," *J. Phys.: Conf. Ser.*, vol. 1342, no. 1, p. 012 003, 2020. DOI: 10.1088/1742-6596/1342/1/012003.

[267] [Online]. Available: https://www.queensu.ca/physics/news-g/first-light.

[268] D. Durnford, "NEWS-G: Status and future prospects," Presented at SNOWMASS CF1 Meeting, 2020. [Online]. Available: https://indico.fnal.gov/event/44956/contributions/194187/attachments/133387/164505/snowmass_cf1_news-g.pdf.

[269] L. Balogh *et al.*, "DarkSPHERE: Exploring light dark matter with a spherical proportional counter electroformed underground at the Boulby Underground Laboratory," 2023. arXiv: 2301.05183 [hep-ex].

[270] G. Angloher *et al.*, "Results on light dark matter particles with a low-threshold CRESST-II detector," *Eur. Phys. J. C*, vol. 76, no. 1, p. 25, 2016. DOI: 10.1140/epjc/s10052-016-3877-3.

[271] M. Morhac, J. Kliman, V. Matousek, M. Veselsky, and I. Turzo, "Background elimination methods for multidimensional coincidence gamma-ray spectra," *Nucl. Instrum. Meth. A*, vol. 401, pp. 113–132, 1997. DOI: 10.1016/S0168-9002(97)01023-1.

[272] M. Morhac, J. Kliman, V. Matousek, M. Veselsky, and I. Turzo, "Efficient one- and two-dimensional gold deconvolution and its application to gamma-ray spectra decomposition," *Nucl. Instrum. Meth. A*, vol. 401, no. 2, pp. 385 –408, 1997. DOI: https://doi.org/10.1016/S0168-9002(97)01058-9.

[273] M. Morhac, J. Kliman, V. Matousek, M. Veselsky, and I. Turzo, "Identification of peaks in multidimensional coincidence gamma-ray spectra," *Nucl. Instrum. Meth. A*, vol. 443, pp. 108–125, 2000. DOI: 10.1016/S0168-9002(99)01005-0.

[274] M. Morhac, *TSpectrum class reference*. [Online]. Available: https://root.cern.ch/doc/master/classTSpectrum.html.

[275] J.-M. Coquillat, "Calibration and background data analysis in the scope of the NEWS-G dark matter experiment," M.S. thesis, Queen's University, 2021.

[276] A. Padawer-Blatt, *Internal note: Removing Microphonic Noise to Improve Light Dark Matter Resolution*, 2021.

[277] M. Jin, "Measurements and Analysis of the Sensitivity of Superheated C3F8 Bubble Chambers to Interactions from WIMP Dark Matter," Ph.D. dissertation, Northwestern University, 2019.

[278] [Online]. Available: https://pypi.org/project/numdifftools/.

[279] D. Srdoc *et al.*, "Statistical fluctuations in the ionisation yield of low- energy photons absorbed in polyatomic gases," *J. Phys. B: Atom. Mol. Phys.*, vol. 20, p. 4473, 1987.

[280] J. Planinić, "Statistical fluctuations in the proecss of primary ionization in methane," *Nucl. Instrumen. Methods*, vol. 187, pp. 473–475, 1981.

[281] B. Grosswendt and E. Waibel, "Statistical ionisation yield fluctuations and determination of spatial ionisation and energy absorption for low energy electrons," *Radiat. Prot. Dos.*, vol. 13, no. 1-4, pp. 95–102, 1985.

[282] L. Balogh *et al.*, "Measurements of the ionization efficiency of protons in methane," *Eur. Phys. J. C*, vol. 82, p. 1114, 2022. DOI: https://doi.org/10.1140/epjc/s10052-022-11063-9.

[283] M. Chemtob, N. Parmentier, and V. Nguyen, "Some experimental results on W-values for heavy particles," *Phys. Med. Biol.*, vol. 23, no. 6, p. 1197, 1978. DOI: 10.1088/0031-9155/23/6/018.

[284] I. Katsioulas, P. Knights, and K. Nikolopoulos, "Ionisation quenching factors from w-values in pure gases for rare event searches," *Astropart. Phys.*, vol. 141, p. 102 707, 2022. DOI: https://doi.org/10.1016/j.astropartphys.2022.102707.

[285] W. N. Venables and B. D. Ripley, "Classification," in *Modern Applied Statistics with S*. New York, NY: Springer New York, 2002, pp. 331–351, ISBN: 978-0-387-21706-2. DOI: 10.1007/978-0-387-21706-2_12.

[286] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. Wiley, 2000, ISBN: 978-0-471-05669-0.

[287] B. Ruscic, "Active thermochemical tables: Sequential bond dissociation enthalpies of methane, ethane, and methanol and the related thermochemistry," *J. Phys. Chem. A*, vol. 119, no. 28, pp. 7810–7837, 2015. DOI: 10.1021/acs.jpca.5b01346.

[288] E. Behnke *et al.*, "Final results of the PICASSO dark matter search experiment," *Astropart. Phys.*, vol. 90, pp. 85–92, 2017. DOI: https://doi.org/10.1016/j.astropartphys.2017.02.005.

[289] D. Akerib *et al.*, "Limits on spin-dependent WIMP-nucleon cross section obtained from the complete LUX exposure," *Phys. Rev. Lett.*, vol. 118, p. 251 302, 25 2017. DOI: 10.1103/PhysRevLett.118.251302.

[290] J. Xia *et al.*, "PandaX-II constraints on spin-dependent WIMP-nucleon effective interactions," *Physics Letters B*, vol. 792, pp. 193–198, 2019. DOI: https://doi.org/10.1016/j.physletb.2019.02.043.

[291] G. Angloher *et al.*, "Testing spin-dependent dark matter interactions with lithium aluminate targets in cresst-iii," *Phys. Rev. D*, vol. 106, p. 092 008, 9 2022. DOI: 10.1103/PhysRevD.106.092008.

[292] E. Aprile *et al.*, "Search for light dark matter interactions enhanced by the Migdal effect or Bremsstrahlung in XENON1T," *Phys. Rev. Lett.*, vol. 123, p. 241 803, 24 2019. DOI: 10.1103/PhysRevLett.123.241803.

[293] J. Collar, "Search for a nonrelativistic component in the spectrum of cosmic rays at Earth," *Phys. Rev. D*, vol. 98, p. 023 005, 2 2018. DOI: 10.1103/PhysRevD.98.023005.

[294] T. Bringmann and M. Pospelov, "Novel direct detection constraints on light dark matter," *Phys. Rev. Lett.*, vol. 122, p. 171 801, 17 2019. DOI: 10.1103/PhysRevLett.122.171801.

[295] D. Glaser, "Some effects of ionizing radiation on the formation of bubbles in liquids," *Phys. Rev.*, vol. 87, pp. 665–665, 4 1952. DOI: 10.1103/PhysRev.87.665.

[296] F. Hasert *et al.*, "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment," *Phys. Lett. B*, vol. 46, pp. 138–140, 1973. DOI: 10.1016/0370-2693(73)90499-1.

[297] F. Flakus, *Detection and measuring ionizing radiation - a short history*, 1981.

[298] E. Alfonso-Pita *et al.*, "Snowmass 2021 Scintillating Bubble Chambers: Liquid-noble Bubble Chambers for Dark Matter and CEνNS Detection," 2022. arXiv: 2207.12400 [physics.ins-det].

[299] S. Fallows *et al.*, "Letter of Interest: Multi-ton scale bubble chambers," *Snowmass 2021*, 2021. [Online]. Available: {https://www.snowmass21.org/docs/files/summaries/CF/SNOWMASS21-CF1_CF0-079.pdf}.

[300] E. Alfonso-Pita *et al.*, "Letter of Interest: Enabling the next generation of bubble-chamberexperiments for dark matter and neutrino physics," *Snowmass 2021*, 2021. [Online]. Available: {https://www.snowmass21.org/docs/files/summaries/IF/SNOWMASS21-IF8_IF0_Eric_Dahl-135.pdf}.

[301] J. Cooley *et al.*, "Report of the topical group on particle dark matter for snowmass 2021," 2022. arXiv: 2209.07426 [hep-ph].

[302] D. Durnford and M.-C. Piro, "Nucleation efficiency of nuclear recoils in bubble chambers," *J. Instrumen.*, vol. 17, no. 01, p. C01030, 2022. DOI: 10.1088/1748-0221/17/01/c01030.

[303] G. Harigel, "Bubble chambers, technology and impact on high energy physics," in *30 Years of Bubble Chamber Physics*, 2003.

[304] G. Harigel, "The big European bubble chamber at CERN Pt 1," *Physikalische Blaetter*, vol. 31, pp. 13–28, 1975. [Online]. Available: {http://inis.iaea.org/search/search.aspx?orig_q=RN:06206980}.

[305] D. Haidt, "The Discovery of Weak Neutral Currents," *Adv. Ser. Dir. High Energy Phys.*, vol. 23, pp. 165–183, 2015. DOI: 10.1142/9789814644150_0007.

[306] "Bubble chamber: D meson production and decay," 1978. [Online]. Available: https://cds.cern.ch/record/39469.

[307] V. Zacek, "Search for dark matter with moderately superheated liquids," *Il Nuovo Cimento*, vol. 107A, no. 2, p. 291, 1994. DOI: 10.1088/1742-6596/2156/1/012059.

[308] J. Collar *et al.*, "COUPP - a search for dark matter with a continuously sensitive bubble chamber," 2007. DOI: 10.2172/898400.

[309] C. Amole *et al.*, "Dark Matter Search Results from the PICO-2L C_3F_8 Bubble Chamber," *Phys. Rev. Lett.*, vol. 114, p. 231 302, 2015.

[310] P. Mitra, "PICO-60: A Dark Matter Search Experiment with C3F8 in a Bubble Chamber," Ph.D. dissertation, University of Alberta, 2018. DOI: https://doi.org/10.7939/R30G3HD1W.

[311] M. Bressler *et al.*, "A buffer-free concept bubble chamber for PICO dark matter searches," *J. Instrumen.*, vol. 14, no. 08, P08019–P08019, 2019. DOI: 10.1088/1748-0221/14/08/p08019.

[312] G. Giroux, "Searching for Dark Matter with the PICO Bubble Chambers," Presented at the 2017 CAP Congress, Kingston, Canada, 2017. [Online]. Available: https://indico.cern.ch/event/593812/contributions/2499749/attachments/1468988/2272242/Guillaume_Giroux_-_CAP_Congress_2017.pdf.

[313] E. Morikawa, R. Reininger, P. Gürtler, V. Saile, and P. Laporte, "Argon, krypton, and xenon excimer luminescence: From the dilute gas to the condensed phase," *J. Chem. Phys.*, vol. 91, no. 3, pp. 1469–1477, 1989. DOI: 10.1063/1.457108.

[314] D. Baxter *et al.*, "First demonstration of a scintillating xenon bubble chamber for detecting dark matter and coherent elastic neutrino-nucleus scattering," *Phys. Rev. Lett.*, vol. 118, p. 231 301, 23 2017. DOI: 10.1103/PhysRevLett.118.231301.

[315] J. L. Brown, D. A. Glaser, and M. L. Perl, "Liquid xenon bubble chamber," *Physical Review: Letters To The Editor*, vol. 102, 1956.

[316] D. Baxter, *Eliminating Backgrounds in the Search for Dark Matter with the PICO-60 Bubble Chamber*, 2018. DOI: https://doi.org/10.21985/N2TV1Z.

[317] F. Seitz, "On the theory of the bubble chamber," *The Physics of Fluids*, vol. 1, no. 1, pp. 2–13, 1958. DOI: 10.1063/1.1724333.

[318] R. Shutt, *Bubble and Spark Chambers: Principles and Use, Part 1*. Academic Press, 1967.

[319] C. Amole *et al.*, "Dark matter search results from the PICO-60 CF_3I bubble chamber," *Phys. Rev. D*, vol. 93, p. 052 014, 5 2016. DOI: 10.1103/PhysRevD.93.052014.

[320] E. Behnke *et al.*, "First dark matter search results from a 4-kg CF_3I bubble chamber operated in a deep underground site," *Phys. Rev. D*, vol. 86, no. 5, 2012. DOI: 10.1103/physrevd.86.052001.

[321] E. B. et. al. (COUPP Collaboration), "Direct measurement of the bubble-nucleation energy threshold in a CF_3I bubble chamber," *Physical Review D*, vol. 88, 021101(R), 2013, arXiv:1304.6001v2.

[322] F. Tardif, *Direct detection of dark matter with PICO experiment and the PICO-0.1 calibration chamber*, 2019. [Online]. Available: papyrus.bib.umontreal.ca/xmlui/handle/1866/22214.

[323] S. Archambault *et al.*, "New Insights into Particle Detection with Superheated Liquids," *New J. Phys.*, vol. 13, p. 043 006, 2011. DOI: 10.1088/1367-2630/13/4/043006. arXiv: 1011.4553 [physics.ins-det].

[324] F. d'Errico, "Fundamental Properties of Superheated Drop (Bubble) Detectors," *Radiat. Prot. Dosim.*, vol. 84, no. 1-4, pp. 55–62, 1999. DOI: 10.1093/oxfordjournals.rpd.a032796.

[325] F. d'Errico, "Radiation dosimetry and spectrometry with superheated emulsions," *Nucl. Instrum. Methods Phys. Res., Sect. B*, vol. 184, no. 1, pp. 229–254, 2001. DOI: https://doi.org/10.1016/S0168-583X(01)00730-3.

[326] R. Tolman, "The effect of droplet size on surface tension," *J. Chem. Phys.*, vol. 17, no. 3, pp. 333–337, 1949. DOI: 10.1063/1.1747247.

[327] A. E. Robinson, "Reanalysis of radioisotope measurements of the $^9\mathrm{Be}(\gamma, n)^8\mathrm{Be}$ cross-section," *Phys. Rev. C*, vol. 94, no. 2, p. 024 613, 2016. DOI: 10.1103/PhysRevC.94.024613. arXiv: 1602.05911 [nucl-ex].

[328] R. Van de Graaff, "Tandem Electrostatic Accelerators," *Nucl. Instrum. and Methods*, vol. 8, p. 195, 1960.

[329] M. Lafreniere, "Mesures d'étalonnage aux neutrons et caractérisation par étude Monte Carlo de la réponse des détecteurs à gouttelettes surchauffées conçus pour la recherche et la détection directe du neutralino (la matière sombre) menant aux résultats finaux de l'expérience PICASSO," Ph.D. dissertation, Université de Montréal, 2018. [Online]. Available: https://papyrus.bib.umontreal.ca/xmlui/handle/1866/20603.

[330] J. Gibbons, R. Macklin, and H. Schmitt, "^{51}V(p,n)^{51}Cr Reaction as a 5- to 120-keV Neutron Source," *Phys. Rev.*, vol. 100, p. 032 801, 2014.

[331] D. A. Brown *et al.*, "ENDF/B-VIII.0: The 8th Major Release of the Nuclear Reaction Data Library with CIELO-project Cross Sections, New Standards and Thermal Scattering Data," *Nucl. Data Sheets*, vol. 148, pp. 1–142, 2018. DOI: 10.1016/j.nds.2018.02.001.

[332] S. Pozzi, E. Padovani, and M. Marseguerra, "MCNP-PoliMi: a Monte-Carlo code for correlation measurements," *Nucl. Instrument. Meth.*, vol. 513, no. 3, p. 550, 2003.

[333] E. Mendoza and D. Cano-Ott, "Update of the evaluated neutron cross section libraries for the geant4 code," International Atomic Energy Agency (IAEA), Tech. Rep. INDC(NDS)–0758, 2018.

[334] A. E. Robinson, "New libraries for simulating neutron scattering in dark matter detector calibrations," *Phys. Rev. C*, vol. 89, no. 3, p. 032 801, 2014. DOI: 10.1103/PhysRevC.89.032801. arXiv: 1401.7900 [nucl-ex].

[335] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition*. Wiley, 1968, ISBN: 978-0471257080.

[336] J. Goodman and J. Weare, "Ensemble samplers with affine invariance," *Comm. Appl. Math. Comp. Sci.*, vol. 5, no. 1, pp. 65–80, 2010. DOI: 10.2140/camcos.2010.5.65.

[337] C. Amole *et al.*, "Improved dark matter search results from PICO-2L Run 2," *Phys. Rev. D*, vol. 93, p. 061 101, 6 2016. DOI: 10.1103/PhysRevD.93.061101.

[338] M. Felizardo *et al.*, "The SIMPLE Phase II dark matter search," *Phys. Rev. D*, vol. 89, p. 072 013, 7 2014. DOI: 10.1103/PhysRevD.89.072013.

[339] C. Fu *et al.*, "Spin-Dependent Weakly-Interacting-Massive-Particle–Nucleon Cross Section Limits from First Data of PandaX-II Experiment," *Phys. Rev. Lett.*, vol. 118, p. 071 301, 7 2017. DOI: 10.1103/PhysRevLett.118.071301.

[340] M. Aartsen *et al.*, "Search for annihilating dark matter in the Sun with 3 years of IceCube data," *Eur. Phys. J. C*, vol. 77, no. 3, 2017. DOI: 10.1140/epjc/s10052-017-4689-9.

[341] T. Tanaka *et al.*, "An Indirect Search for WIMPs in the Sun using 3109.6 days of upward-going muons in Super-Kamiokande," *Astrophys. J.*, vol. 742, no. 2, p. 78, 2011. DOI: 10.1088/0004-637x/742/2/78.

[342] K. Choi *et al.*, "Search for Neutrinos from Annihilation of Captured Low-Mass Dark Matter Particles in the Sun by Super-Kamiokande," *Phys. Rev. Lett.*, vol. 114, p. 141 301, 14 2015. DOI: 10.1103/PhysRevLett.114.141301.

[343] C. Amole *et al.*, "Dark Matter Search Results from the PICO-60 C_3F_8 Bubble Chamber," *Phys. Rev. Lett.*, vol. 118, no. 25, p. 251 301, 2017. DOI: 10.1103/PhysRevLett.118.251301. arXiv: 1702.07666 [astro-ph.CO].

[344] V. Gluscevic, M. I. Gresham, S. D. McDermott, A. H. G. Peter, and K. M. Zurek, "Identifying the Theory of Dark Matter with Direct Detection," *JCAP*, vol. 12, p. 057, 2015. DOI: 10.1088/1475-7516/2015/12/057. arXiv: 1506.04454 [hep-ph].

[345] M. Bressler, "Operation and Calibration of Right-Side-Up Bubble Chambers at keV Thresholds: Towards New Superheated Dark Matter Searches," Ph.D. dissertation, Drexel University, 2022. DOI: DOI:https://doi.org/10.17918/00001104.

[346] S. Mukhopadhyay, "Xenon Isotopes," in *Encyclopedia of Geochemistry: A Comprehensive Reference Source on the Chemistry of the Earth*, W. White, Ed. Springer International Publishing, 2018, pp. 1500–1508, ISBN: 978-3-319-39312-4. DOI: 10.1007/978-3-319-39312-4_203.

[347] "Cf-252 neutron spectrum," International Atomic Energy Agency (IAEA), Tech. Rep. IAEA-NDS–98(Rev1), 1987.

[348] M. Javed *et al.*, "Thermodynamic speed of sound of xenon," *J. Chem. Thermodyn.*, vol. 141, p. 105 933, 2020.

[349] M. van Kreveld, *Delaunay Triangulation and Tetrehedrilization*. [Online]. Available: {https://slideplayer.com/slide/2488745/}.

[350] A. Okabe, B. Boots, K. Sugihara, and S. Nok Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley, 2000, ISBN: 978-0471986355. DOI: 10.1017/CBO9780511791277.

[351] C. Dahl, *Internal note: Poisson background subtraction*, 2022.

[352] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004. DOI: 10.1109/MSP.2004.1311138.

[353] K. Burnham and D. Anderson, in *Model Selection and Multimodel Inference*. Springer New York, NY, 2002, pp. 49–97, ISBN: 978-0-387-22456-5.

[354] D. Freedman and P. Diaconis, "On the histogram as a density estimator: L2 theory," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, pp. 453–476, 1981. DOI: 10.1007/BF01025868.

[355] R. Godwin and D. Giles, "Analytic bias correction for maximum likelihood estimators when the bias function is non-constant," *Commun. Stat. Simulat.*, vol. 48, no. 1, pp. 15–26, 2019.

[356] W. Daniel, *Applied Nonparametric Statistics* (Duxbury advanced series in statistics and decision sciences). PWS-KENT Publishers, 1990, ISBN: 9780534919764. [Online]. Available: https://books.google.com/books?id=0hPvAAAAMAAJ.

[357] C. Taylor, *Applications Of Dynamic Programming To Agricultural Decision Problems*. CRC Press, 2019, ISBN: 9780429723094. [Online]. Available: https://books.google.ca/books?id=71SsDwAAQBAJ.

[358] R. Bellman, *Dynamic Programming* (Rand Corporation research study). Princeton University Press, 1957, ISBN: 9780691079516. [Online]. Available: https://books.google.ca/books?id=wdtoPwAACAAJ.

[359] D. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2004.

[360] H. Rosenbrock, "An Automatic Method for Finding the Greatest or Least Value of a Function," *Comput. J.*, vol. 3, no. 3, pp. 175–184, 1960. DOI: 10.1093/comjnl/3.3.175.

[361] *Optimization (scipy.optimize)*, 2014. [Online]. Available: ttps://docs.scipy.org/doc/scipy-0.14.0/reference/tutorial/optimize.html#unconstrained-minimization-of-multivariate-scalar-functions-minimize.

[362] J. U. Lange, *Nautilus: Boosting bayesian importance nested sampling with deep learning*, 2023. arXiv: 2306.16923 [astro-ph.IM].

[363] I. Emiola and R. Adem, "Comparison of Minimization Methods for Rosenbrock Functions," 2021. eprint: 2101.10546.

[364] Private exchange with the author of J.U. Lange, 2023 (arXiv:2306.16923).

Appendix A: Fast burn-in MCMC algorithm

Fitting a high-dimensional models such those described in chapters 3 – 5 often presents a critical computational challenge, due to the "curse of dimensionality", wherein sampling or optimization algorithms must probe exponentially rising parameter space for every additional parameter that is added to a model [357, 358]. Simply put, sampling near the optimum of a many-parameter likelihood function becomes increasingly inefficient, to the point where finding the true global optimum may be intractable for some algorithms.

A Markov Chain Monte Carlo (MCMC) is a natural tool for such a problem, of which there are many powerful implementations in common scripting languages, such as the `emcee` library for python [216]. This software package employs a variant of the typical Metropolis-Hastings (MH) walker step proposal algorithm [359] called the "stretch move" [216, 336], which was shown to out-perform the MH algorithm. However, this approach was deemed too inefficient for some of the analyses included in this work. For the PICO nucleation efficiency analysis, a custom MCMC approach was developed [277] called the "fast burn-in method". Rather than letting the MCMC walkers sample uninterrupted (with a given proposal function), this algorithm proceeds as follows:

1. Start M walkers with randomized initial guesses, using the standard `emcee` stretch move MH proposal.

2. After a small number of steps (usually 5 to 10), interrupt the walkers.

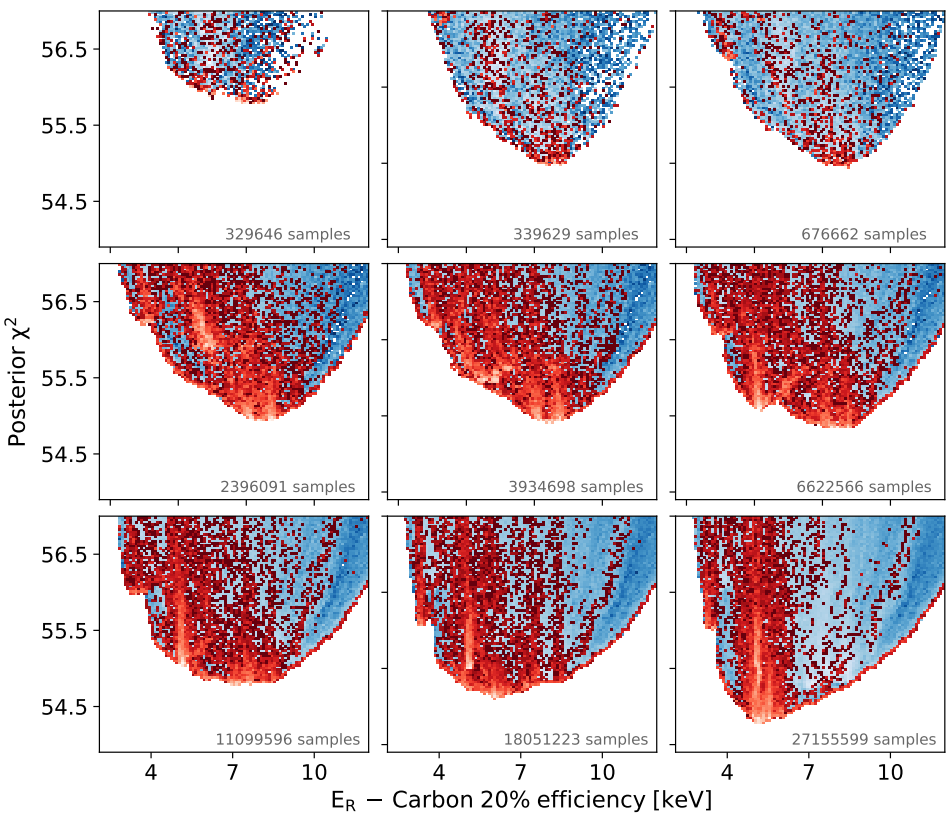


Figure A.1: An example of the progression of a fast burn-in MCMC fit of the PICO calibration data, for the parameter giving the recoil energy at which there is 20 % nucleation efficiency on carbon. In each frame, the cumulative MCMC samples are shown as a blue 2D histogram, while the starting points for the current epoch are shown in red. This emphasizes the algorithm's preferential sampling of the boundary of the likelihood function. Taken from figure 17 of [3].

3. Bin the samples explored by the walkers to that point in B bins of each parameter in the model. Determine the maximum likelihood set of parameters in each bin and append to a list. Remove non-unique sets of parameters from this list.

4. Start a new set of walkers from the list of best-fit points determined above.

5. Repeats steps 2-4 until convergence is reached. Each iteration of this process is dubbed an "epoch".

A depiction of this approach for one of the variables in the PICO nucleation efficiency model is shown in Fig. A.1. This approach forces the MCMC walkers to

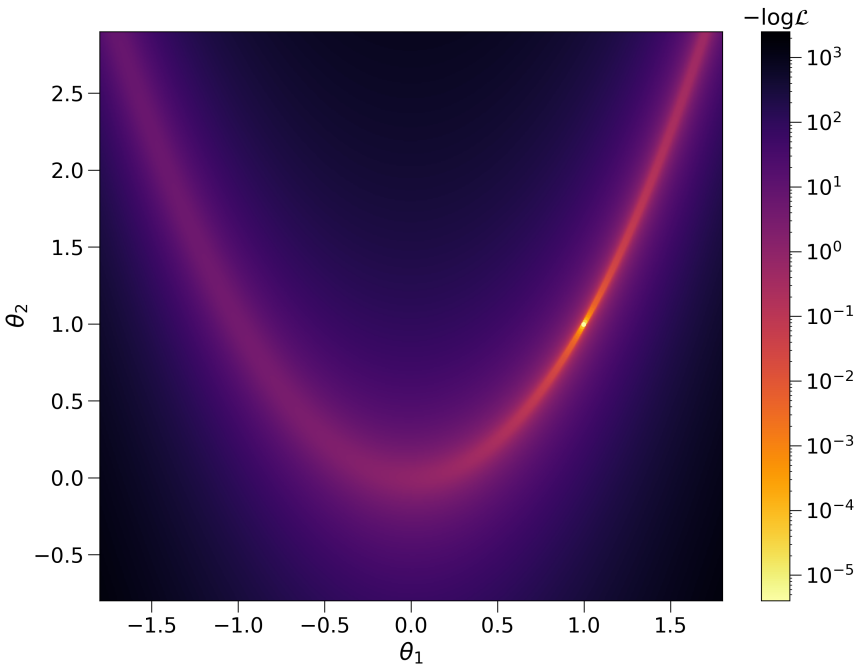


Figure A.2: Depiction of the 2-D Rosenbrock likelihood function [360].

preferentially sample the boundary of the likelihood function, which can be used (in some approaches) to estimate parameter confidence intervals as well as the best-fit model to the data. This does however come at the expense of no longer sampling the posterior proportionally, a desirable trait of typical MCMC algorithms.

To demonstrate the performance/efficacy of the fast burn-in method compared to a traditional MCMC, we can pit both against a standard test of optimization and sampling algorithms - the Rosenbrock function [360–363]. This challenging likelihood function, generalizable to N dimensions, is defined as [362]:

$$-\log \mathcal{L} = \sum_{i=0}^{N-2} \left[(1-\theta_i)^2 + 100 \left(\theta_{i+1} - \theta_i^2 \right)^2 \right] \tag{A.1}$$

The log-likelihood contains a parabolic valley with a narrow global optimum, as can be seen in Fig. A.2 for the 2-D function. The non-trivial correlations between parameters (and possibly large number of dimensions) poses a difficult task for all algorithms.

For this comparison, an MCMC was run on the 10-D version of this likelihood ($N = 10$ above), using the normal MH "stretch step" of the `emcee` package [216]

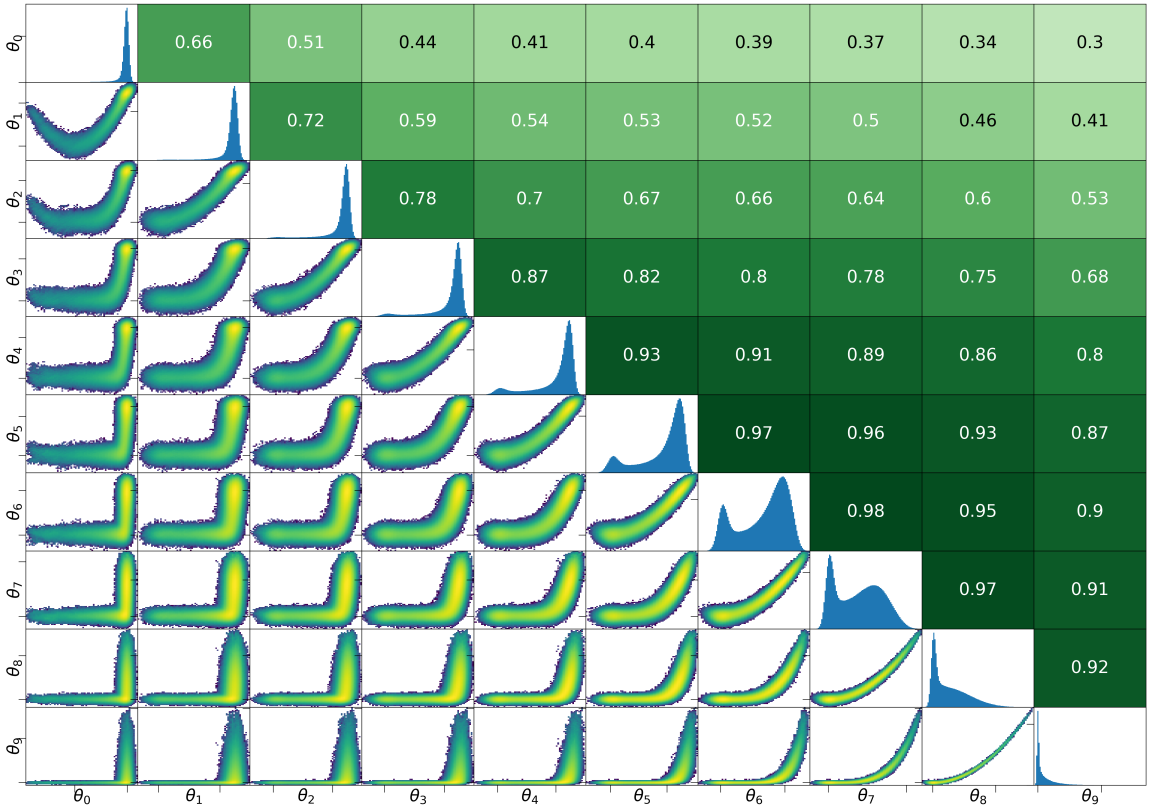


Figure A.3: Corner plot of the 10-D Rosenbrock function sampled with a MH "stretch move" MCMC (using `emcee` [216]). The lower-left panels show 2-D histograms of the MCMC samples of every possible parameter combination with a logarithmic colour-scale. The diagonal panels show the marginal posterior probability distributions of each individual parameter (a histogram of the MCMC samples). The upper-right panels display the Spearman correlation coefficient [356] of each pair of parameters, with a colour-scale to aid visualization.

with 100 walkers, a step size scaling of $a = 1$, running for 5×10^7 steps (a total of 5×10^9 samples), with a burn in period of 10^6 steps. This long MCMC run was needed to fully sample this complicated function, with a burn-in time at least as long as the auto-correlation time (approximately 10^6 samples [362]), and a main MCMC run many times longer than that. The overall shape of this 10-D likelihood function can be gleamed – including highly non-Gaussian, correlated features – from a corner plot of these samples, shown in Fig. A.3. The fast burn-in method was then used to sample the 10-D Rosenbrock function with 200 bins per parameter, 10 steps per epoch, a step size scaling of $a = 1$, for 3000 epochs (a total of $\sim 5 \times 10^7$ samples).

It's worth noting the Rosenbrock likelihood is a case where the marginalized likelihood for any given parameter does not necessarily agree with the profile likelihood. Specifically, the marginalized likelihood integrates over all other parameters Θ except the one in question, θ [73]:

$$\mathcal{L}_{\text{Marg.}}(\theta) = \int_{\Theta} \mathcal{L}(\Theta|\theta) d\Theta \tag{A.2}$$

While the profile likelihood is maximized over the nuisance parameters (indicated with a hat) for a given value of the parameter of interest [73]:

$$\mathcal{L}_{\text{Prof.}}(\theta) = \mathcal{L}(\hat{\Theta}|\theta) \tag{A.3}$$

The marginalized posterior probability distribution – the normalized marginal likelihood – is the natural output of an MCMC optimization, since for the standard MH algorithm, the posterior is estimated as simply the histogram of MCMC samples. To obtain the profile likelihood from MCMC samples, one must determine the maximum value of $\mathcal{L}$ possible for a given value (or given bin range) of the parameter of interest θ. This is easily done in python with functions such as `scipy.stats.binned_statistic` [238].

In this example study, there is substantial disagreement between the marginal and profile likelihood functions for some parameters, with the former often yielding biased best-fit values that differ significantly from the true optimal parameter values, which one can see is 1 for all θ_i from eq. A.1. A prominent example of this is shown in Fig. A.4, comparing the profile and marginal likelihoods of θ_7 estimated using the MH MCMC. In this case, similar case studies in the literature confirm that the marginalized distribution is calculated correctly (namely figure 8 of ref. [362]), but the best-fit value one would read off of this distribution is close to 0, whereas the profile likelihood is peaked around the true value of 1. To understand this, it bears remembering that the marginalized posterior probability at a given value is

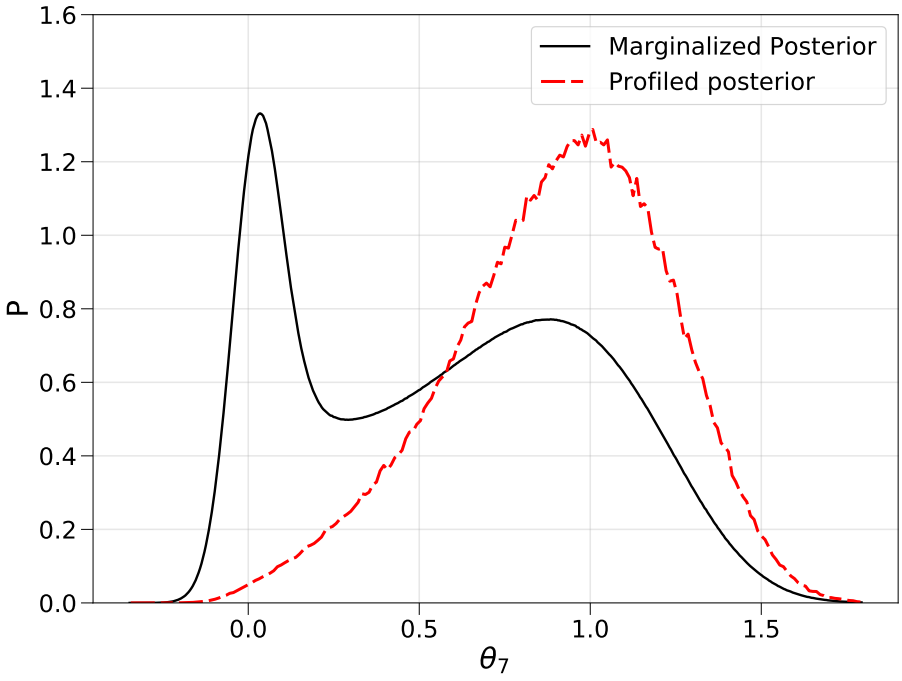


Figure A.4: The marginalized posterior probability distribution of Rosenbrock function parameter θ_7 (black curve) compared to the posterior calculated from the profile likelihood (red dashed curve), both estimated using the MH "stretch move" MCMC [359, 360].

determined by the likelihood times the volume of parameter space; in this case, the volume around the true optimum of 1 is very small, so the posterior probability is relatively lower [364].

The fast burn-in method can only produce the profile likelihood for specified parameters, not the marginalized posterior, since the sampling is not proportional to the posterior probability as is the case with most MCMC algorithms (admittedly a deficit of this method). Therefore, the output of both MCMC algorithms must be compared on the basis of profile likelihood functions, which are displayed for several parameters of the Rosenbrock function in Fig. A.5. From this, it is clear that both approaches produce nearly identical likelihood function results, although the MH estimate does seem to under-sample the likelihood slightly compared to the fast burn-in method.

As it is established that both the MH and fast burn-in MCMC approaches produce roughly equivalent results in the case of the 10-D Rosenbrock likelihood, the remaining

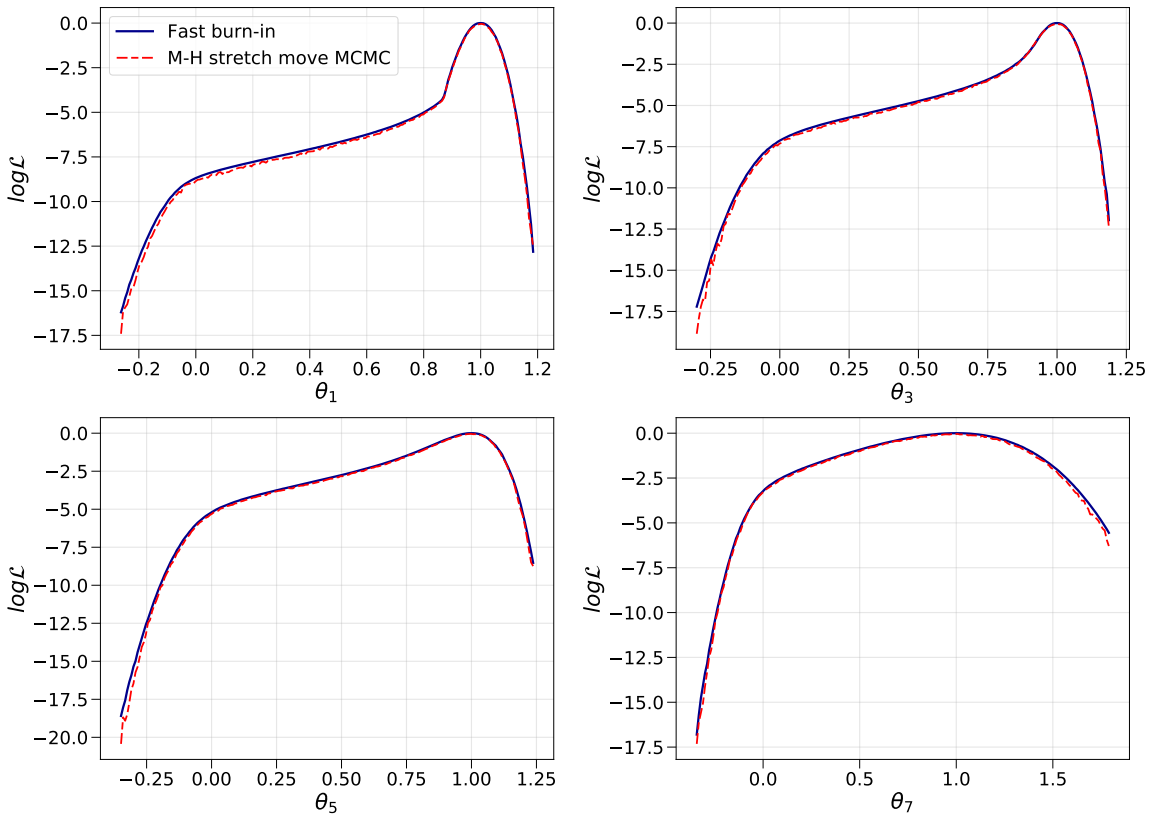


Figure A.5: The profile likelihood functions of several 10-D Rosenbrock function parameters, comparing the results obtained with the MH MCMC (red dashed curve) and fast burn-in MCMC methods (navy curve).

question is of the relative performance/efficiency of each method. This can be assessed by considering how many samples were required to reach the true global optimum (which occurs at $\log \mathcal{L} = 0$), and how far off the best-fit estimate was in each case from the true optimum parameter values (which occur at $\theta_i = 1$ for all i). The latter quantity is defined as:

$$||\text{best} - \text{fit}|| = \sqrt{\sum_{i=0}^{9} \left(\hat{\theta}_i - 1\right)^2} \tag{A.4}$$

The evolution of both of these quantities ($\max \log \mathcal{L}$ and $||\text{best} - \text{fit}||$) as a function of MCMC samples is shown in Fig. A.6. This clearly demonstrates that compared to MH, the fast burn-in algorithm obtains a higher maximum likelihood and closer best-fit, with orders-of-magnitude fewer samples. This improved performance for

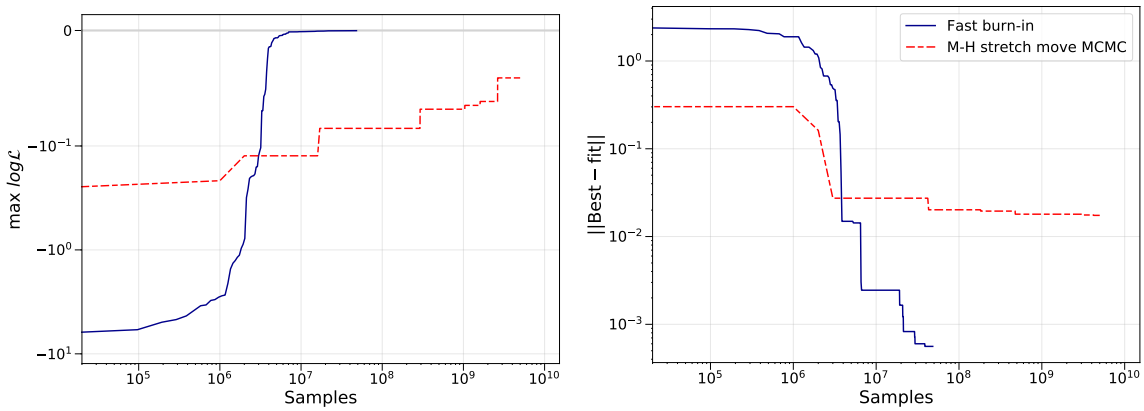


Figure A.6: The evolution of the maximum log-likelihood (left) and distance between the best-fit and true optimum parameters (right) of the 10-D Rosenbrock likelihood as a function of MCMC samples, for both the MH MCMC (red dashed curve) and fast burn-in MCMC (navy curve). Note that the highest possible maximum log-likelihood is 0, and the minimum possible best-fit distance $||\text{best} - \text{fit}||$ is 0.

high-dimensional models is the motivation for using this MCMC approach in several instances throughout this work.

Appendix B: LXe nucleation efficiency bilinear interpolation

The LXe nucleation efficiency model described in section 5.6.1 includes a triangular mesh defined on a grid of points of Q_{Seitz} and efficiency ϵ, to setup a planar interpolation on the mesh cells for ϵ as a function of Q_{Seitz} and recoil energy E_R. Efficiency curves extracted from this interpolation object for some Q_{Seitz} in between the fencepost values of Q must still obey the physical constraint of monotonicity in Q. A proof that this holds is given below.

Consider a single quadrilateral cell of the mesh (i.e. two triangles), defined by the points $(Q_1, \epsilon_1, x_{11})$, $(Q_1, \epsilon_2, x_{12})$, $(Q_2, \epsilon_2, x_{22})$, and $(Q_2, \epsilon_1, x_{21})$, as shown in Fig. B.1. Necessarily, these points must follow the constraints

$$x_{i2} \geq x_{i1}, \qquad x_{2j} \geq x_{1j}, \qquad x_{ij} \geq Q_i \tag{B.1}$$

to respect the usual physical constraints on the model. The shared hypotenuse of the two triangles we denote as $\overleftrightarrow{H}$, the two opposite triangle edges as $\overleftrightarrow{O_1}$ and $\overleftrightarrow{O_2}$ when $\epsilon = \epsilon_1$ and ϵ_2 respectively. These lines can be defined parametrically as follows:

$$\overleftrightarrow{H} = t \begin{pmatrix} Q_2 - Q_1 \\ \epsilon_2 - \epsilon_1 \\ x_{22} - x_{11} \end{pmatrix} + \begin{pmatrix} Q_1 \\ \epsilon_1 \\ x_{11} \end{pmatrix}, \quad t \in [0, 1] \tag{B.2}$$

$$\overleftrightarrow{O_1} : u \begin{pmatrix} Q_2 - Q_1 \\ 0 \\ x_{21} - x_{11} \end{pmatrix} + \begin{pmatrix} Q_1 \\ \epsilon_1 \\ x_{11} \end{pmatrix}, \quad u \in [0, 1] \tag{B.3}$$

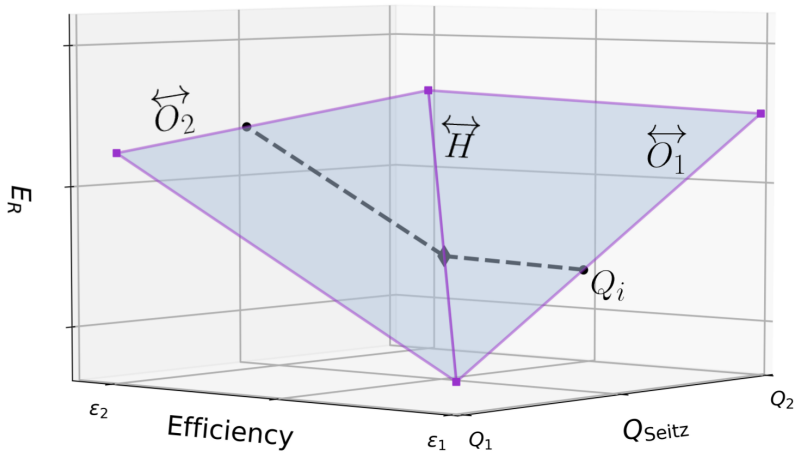


Figure B.1: Arbitrary single quadrilateral of the mesh scheme used for the Xe-SBC nucleation efficiency model, with labels, showing an efficiency curve interpolation for some Q_i.

$$\overleftrightarrow{O_2} : v \begin{pmatrix} Q_2 - Q_1 \\ 0 \\ x_{22} - x_{12} \end{pmatrix} + \begin{pmatrix} Q_1 \\ \epsilon_2 \\ x_{12} \end{pmatrix}, \quad v \in [0, 1] \tag{B.4}$$

An efficiency curve extracted from this model for a given Q_i is then defined by two line segments $\overleftrightarrow{X_1}$ and $\overleftrightarrow{X_2}$, intersecting along the hypotenuse of the triangle at a point X_*. That point of intersection can be found by finding $t/u/v$ such that $Q = Q_i$:

$$t_i(Q_2 - Q_1) + Q_1 = Q_i \implies t_i = \frac{Q_i - Q_1}{Q_2 - Q_1} \tag{B.5}$$

and indeed it's trivial then that $u_i = v_i = t_i$, defining the points of intersection along $\overleftrightarrow{O_1}$ and $\overleftrightarrow{O_2}$. Next, the slope and intercept of $\overleftrightarrow{X_1}$ and $\overleftrightarrow{X_2}$ (m_1/b_1 and m_2/b_2 respectively) can be found as follows:

$$\begin{aligned} m_1 &= \frac{t_i(x_{22} - x_{11}) + x_{11} - t_i(x_{21} - x_{11}) - x_{11}}{t_i(\epsilon_2 - \epsilon_1) + \epsilon_1 - \epsilon_1} \\ &= \frac{x_{22} - x_{21}}{\epsilon_2 - \epsilon_1} \end{aligned} \tag{B.6}$$

$$\begin{aligned} b_1 &= t_i(x_{21} - x_{11}) + x_{11} - m_1\epsilon_1 \\ &= t_i(x_{21} - x_{11}) + x_{11} - \frac{x_{22} - x_{21}}{\epsilon_2 - \epsilon_1}\epsilon_1 \end{aligned} \tag{B.7}$$

$$
\begin{aligned}
m_2 &= \frac{t_i(x_{22} - x_{12}) + x_{12} - t_i(x_{22} - x_{11}) - x_{11}}{\epsilon_2 - t_i(\epsilon_2 - \epsilon_1) - \epsilon_1} \\
&= \frac{t_i(x_{11} - x_{12}) + x_{12} - x_{11}}{(1 - t_i)(\epsilon_2 - \epsilon_1)} \\
&= \frac{x_{12} - x_{11}}{\epsilon_2 - \epsilon_1}
\end{aligned} \tag{B.8}
$$

$$
\begin{aligned}
b_2 &= t_i(x_{22} - x_{12}) + x_{12} - m_2\epsilon_2 \\
&= t_i(x_{22} - x_{12}) + x_{12} - \frac{x_{12} - x_{11}}{\epsilon_2 - \epsilon_1}\epsilon_2
\end{aligned} \tag{B.9}
$$

To prove that any efficiency curve model obtained from the mesh obeys all the physical constraints of the model, it must be the case that for some $Q_k > Q_i$, $\overleftrightarrow{X_{k1}}$ and $\overleftrightarrow{X_{k2}}$ are strictly greater than $\overleftrightarrow{X_{i1}}$ and $\overleftrightarrow{X_{i2}}$. This can be examined in three separate cases; where $\overleftrightarrow{X_{k1}}$ and $\overleftrightarrow{X_{i1}}$ overlap as a function of ϵ (case 1), where $\overleftrightarrow{X_{k1}}$ and $\overleftrightarrow{X_{i2}}$ overlap (case 2), and where $\overleftrightarrow{X_{i2}}$ and $\overleftrightarrow{X_{k2}}$ overlap (case 3). Case 1, where $\epsilon < t_i(\epsilon_2 - \epsilon_1) + \epsilon_1$, is proven as follows:

$$
\begin{aligned}
& t_k > t_i \\
\implies & t_k(x_{21} - x_{11}) + x_{11} - m_1\epsilon_1 > t_i(x_{21} - x_{11}) + x_{11} - m_1\epsilon_1 \\
\implies & m_1\epsilon + b_{1k} > m_1\epsilon + b_{1i} \\
\implies & \overleftrightarrow{X_{k1}} > \overleftrightarrow{X_{i1}}, \quad \text{when } \epsilon < \epsilon_{i*} \qquad \square
\end{aligned} \tag{B.10}
$$

Similarly , for case 3 we have $\epsilon > t_k(\epsilon_2 - \epsilon_1) + \epsilon_1$:

$$
\begin{aligned}
& t_k > t_i \\
\implies & t_k(x_{22} - x_{12}) + x_{12} - m_2\epsilon_2 > t_i(x_{22} - x_{12}) + x_{12} - m_2\epsilon_2 \\
\implies & m_2\epsilon + b_{2k} > m_2\epsilon + b_{2i} \\
\implies & \overleftrightarrow{X_{k2}} > \overleftrightarrow{X_{i2}}, \quad \text{when } \epsilon > \epsilon_{k*} \qquad \square
\end{aligned} \tag{B.11}
$$

Case 2 is somewhat more complicated, as m_2 is not necessarily always greater than m_1, but we can prove that the postulate holds by demonstrating that $\overleftrightarrow{X_{k1}}$ at point ϵ_{i*} is greater than x_{i*} (case 2a):

$$
\begin{aligned}
& t_i < t_k \\
\implies & t_i(x_{21} - x_{11}) < t_k(x_{21} - x_{11}) \\
\implies & t_i(x_{22} - x_{11} - x_{22} + x_{21}) < t_k(x_{21} - x_{11}) \\
\implies & t_i(x_{22} - x_{11}) < t_k(x_{21} - x_{11}) \\
\implies & t_i(x_{22} - x_{11}) < m_1 t_i(\epsilon_2 - \epsilon_1) + t_k(x_{21} - x_{11}) \\
\implies & t_i(x_{22} - x_{11}) < m_1 \left(t_i(\epsilon_2 - \epsilon_1) + \epsilon_1\right) + t_k(x_{21} - x_{11}) + x_{11} - m_1\epsilon_1 \\
\implies & x_{i*} < m_1\epsilon_{i*} + b_{k1}
\end{aligned}
\tag{B.12}
$$

and that x_{k*} is greater than $\overleftrightarrow{X_{k1}}$ at point ϵ_{k*}:

$$
\begin{aligned}
& t_k > t_i \\
\implies & t_k(x_{22} - x_{12}) > t_i(x_{22} - x_{12}) \\
\implies & t_k(x_{22} - x_{11} - x_{12} + x_{11}) > t_i(x_{22} - x_{12}) \\
\implies & t_k(x_{22} - x_{11}) + x_{11} > t_k(x_{12} - x_{11}) - (x_{12} - x_{11}) + t_i(x_{22} - x_{12}) + x_{12} \\
\implies & t_k(x_{22} - x_{11}) + x_{11} > m_2 \left(t_k(\epsilon_2 - \epsilon_1) + \epsilon_1\right) + t_i(x_{22} - x_{12}) + x_{12} - m_2\epsilon_2 \\
\implies & x_{k*} > m_2\epsilon_{k*} + b_{i2}
\end{aligned}
\tag{B.13}
$$

and therefore $\overleftrightarrow{X_{k1}} > \overleftrightarrow{X_{i2}}$ when $\epsilon_{i*} \leq \epsilon \leq \epsilon_{k*}$. Therefore the postulate holds in all cases, and any efficiency curve function interpolated using this specific triangular mesh formulation obeys the physical constraints of the nucleation model. Note that this would not necessarily be the case for a mesh defined differently, e.g. if the hypotenuse of each cell was flipped by 90° in Q/ϵ.