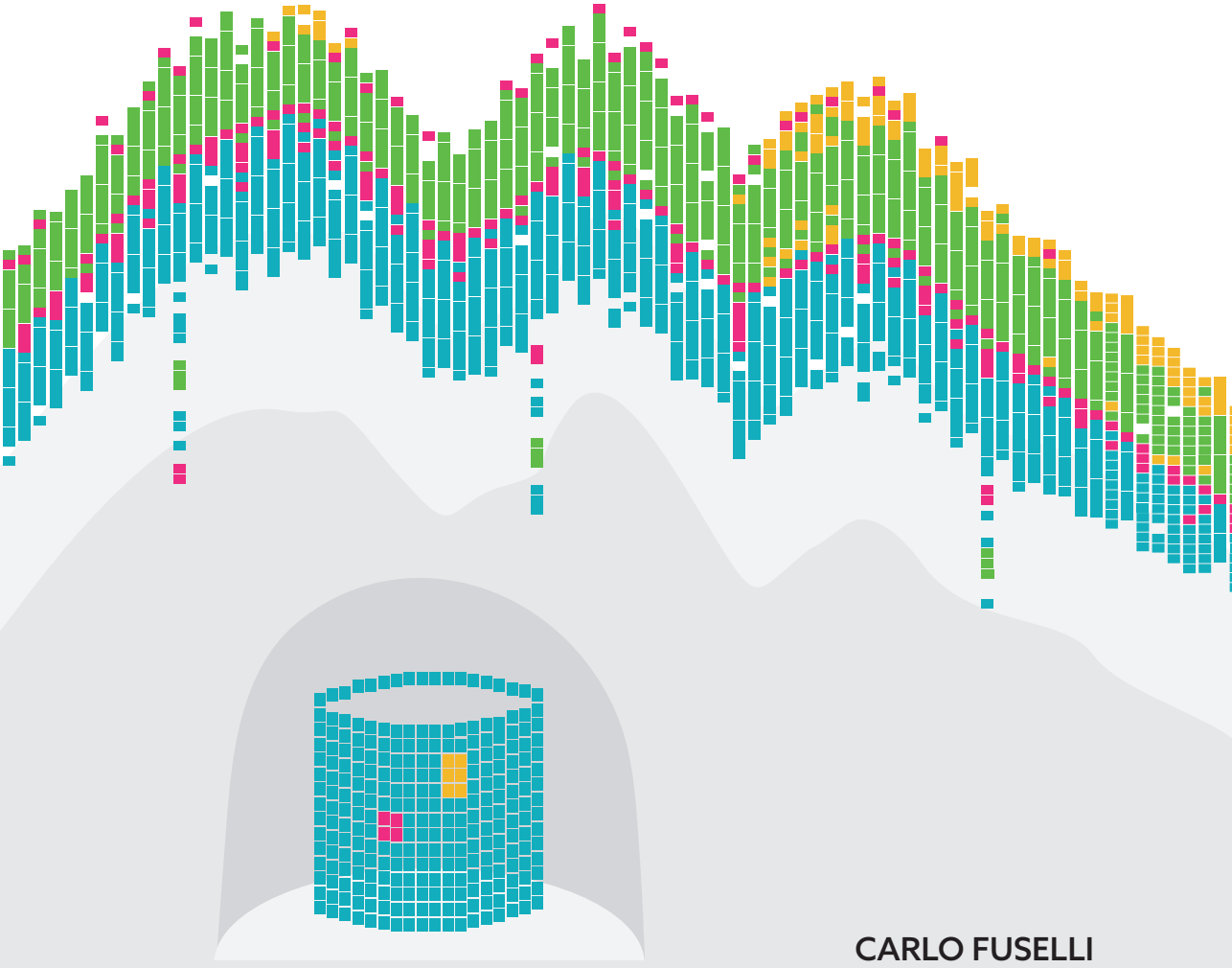
A MATTER OF COINCIDENCES

Electronic Recoils from Beta Decays to Dark Matter Absorption in XENONnT



CARLO FUSELLI

Propositions

belonging to the thesis

A Matter of Coincidences

Carlo Fuselli, Universiteit van Amsterdam

1. It is possible that we will never know whether dark matter is made of WIMPs.

2. Beta decay has been understood for over a century. When it happens inside a heavy nucleus, it is still quite difficult to model.

3. Seemingly innocent detector design choices can have a dramatic impact on the complexity of the data analysis. See: perpendicular wires.

4. There is a trade-off between making simulations reproduce the data and keeping enough modelling in them to learn something from the comparison.

5. The definition of a single scatter or multiple scatter interaction is not unique. It depends on how closely you look.

6. The XENONnT data acquisition system is well aware of when DAQ team members are on holiday, and reliably decides to break then.

7. AI as a tool is evil and addictive, but it can allow physicists to think more about physics and less about software details.

8. Measuring something unimportant is more satisfying than not measuring something that is.

A Matter of Coincidences

Electronic Recoils from Beta Decays to Dark Matter Absorption in XENONnT

Carlo Fuselli

© 2026 Carlo Fuselli. All rights reserved.

ISBN: 978-94-6534-694-6

Printed by ProefschriftMaken | www.proefschriftmaken.nl





This work has been performed at the National Institute for Subatomic Physics (Nikhef), which is funded by the Dutch Research Council (NWO).

A Matter of Coincidences

Electronic Recoils from Beta Decays to Dark Matter Absorption in XENONnT

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op donderdag 12 november 2026, te 14.00 uur

door Carlo Fuselli
geboren te Genova

Promotiecommissie

Promotores:	prof. dr. M.P. Decowski	Universiteit van Amsterdam
	prof. dr. A.P. Colijn	Universiteit van Amsterdam
Overige leden:	prof. dr. E.L.M.P. Laenen	Universiteit van Amsterdam
	prof. dr. A.J. Heijboer	Universiteit van Amsterdam
	prof. dr. ir. P.J. de Jong	Universiteit van Amsterdam
	dr. T.R. Pollmann	Universiteit van Amsterdam
	prof. dr. L. Grandi	University of Chicago
	dr. M. Senghi Soares	Vrije Universiteit Amsterdam
	dr. J. Aalbers	Rijksuniversiteit Groningen

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

PUBLICATIONS

Chapter 5 of this work is based on the following publication. The author provided essential contributions, as described below.

Chapter 5: The XENON Collaboration: E. Aprile, J. Aalbers, K. Abe, M. Adrover, S. Ahmed Maouloud, L. Althueser, B. Andrieu, E. Angelino, D. Antón Martin, S. R. Armbruster, F. Arneodo, L. Baudis, M. Bazyk, L. Bellagamba, R. Biondi, A. Bismark, K. Boese, R. M. Braun, A. Brown, G. Bruno, R. Budnik, C. Cai, C. Capelli, J. M. R. Cardoso, A. P. Cimental Chávez, A. P. Colijn, J. Conrad, J. J. Cuenca-García, V. D'Andrea, L. C. Daniel Garcia, M. P. Decowski, A. Deisting, C. Di Donato, P. Di Gangi, S. Diglio, K. Eitel, S. el Morabit, R. Elleboro, A. Elykov, A. D. Ferella, C. Ferrari, H. Fischer, T. Flehmke, M. Flierman, D. Fuchs, W. Fulgione, C. Fuselli, R. Gaior, F. Gao, R. Giacomobono, F. Girard, R. Glade-Beucke, L. Grandi, J. Grigat, H. Guan, M. Guida, P. Gyorgy, R. Hammann, A. Higuera, C. Hils, L. Hoetzsch, N. F. Hood, M. Iacovacci, Y. Itow, J. Jakob, F. Joerg, Y. Kaminaga, M. Kara, S. Kazama, P. Kharbanda, M. Kobayashi, D. Koke, K. Kooshkjalali, A. Kopec, H. Landsman, R. F. Lang, L. Levinson, I. Li, S. Li, S. Liang, Z. Liang, Y.-T. Lin, S. Lindemann, M. Lindner, K. Liu, M. Liu, J. Loizeau, F. Lombardi, J. A. M. Lopes, G. M. Lucchetti, T. Luce, Y. Ma, C. Macolino, J. Mahlstedt, F. Marignetti, T. Marrodán Undagoitia, K. Martens, J. Masbou, S. Mastroianni, V. Mazza, A. Melchiorre, J. Merz, M. Messina, K. Miuchi, A. Molinario, S. Moriyama, K. Morå, M. Murra, J. Müller, K. Ni, C. T. Oba Ishikawa, U. Oberlack, S. Ouahada, B. Paetsch, Y. Pan, Q. Pellegrini, R. Peres, J. Pienaar, M. Pierre, G. Plante, T. R. Pollmann, A. Prajapati, L. Principe, J. Qin, D. Ramírez García, M. Rajado, A. Ravindran, A. Razeto, R. Singh, L. Sanchez, J. M. F. dos Santos, I. Sarnoff, G. Sartorelli, J. Schreiner, P. Schulte, H. Schulze Eißing, M. Schumann, L. Scotto Lavina, M. Selvi, F. Semeria, P. Shagin, S. Shi, H. Simgen, A. Stevens, C. Szyszka, A. Takeda, Y. Takeuchi, P.-L. Tan, D. Thers, G. Trinchero, C. D. Tunnell, F. Tönnies, K. Valerius, S. Vecchi, S. Vetter, G. Volta, C. Weinheimer, M. Weiss, D. Wenz, C. Wittweg, V. H. S. Wu, Y. Xing, D. Xu, Z. Xu, M. Yamashita, J. Yang, L. Yang, J. Ye, M. Yoshida, L. Yuan, G. Zavattini, Y. Zhao, M. Zhong, T. Zhu, *Spectral Measurement of the* ^{214}*Bi* β*-Decay to the* ^{214}*Po Ground State with the XENONnT Experiment*, Phys. Rev. C **113**, 044303 (2026), reference [1].

C.F. designed and performed the analysis, comprising the ^{214}Bi identification algorithm, the spectral measurement, the electronic recoil yield extraction, and the comparison with nuclear structure models. C.F. and Y.-T. Lin formed the writing team for the collaboration. The manuscript was submitted to arXiv by C.F. and published in Phys. Rev. C. All authors of the XENON Collaboration contributed to the quality of the paper through the internal review process.

CONTENTS

SUMMARY 1

1 BETA DECAYS 7
- 1.1 Fermi Theory and the Beta Spectrum 8
- 1.2 Selection Rules and Classification of Beta Decays 11
- 1.3 The Shape Factor and Nuclear Matrix Elements 14
- 1.4 Conserved and Partially Conserved Currents 17
- 1.5 Nuclear Structure Models 19
- 1.6 The ^{214}Bi Decay as a Case Study 24

2 THE XENONNT DETECTOR 27
- 2.1 Energy Deposition in Liquid Xenon 29
- 2.2 Detection Principle of a Dual-Phase Xenon TPC 30
- 2.3 Detector Layout 32
- 2.4 Data Acquisition and Signal Reconstruction 36
- 2.5 Energy Scale and Detector Gains 37
- 2.6 Science Runs and Detector Conditions 39
- 2.7 Summary 41

3 DETECTOR SIMULATION FRAMEWORK 43
- 3.1 Microphysics and Quanta Generation 44
- 3.2 S1 Signal Simulation 45
- 3.3 S2 Signal Simulation 47
  - 3.3.1 Electron Drift and Diffusion 47
  - 3.3.2 Extraction and Secondary Scintillation 48
  - 3.3.3 Near-Wire Electric Field 53
  - 3.3.4 Delayed Electrons and Photoionisation 54
- 3.4 Detector Response 56
- 3.5 Truth Information 58
- 3.6 Applications of the Simulation 59
- 3.7 Summary 62

4 XENON YIELDS AT HIGH ENERGIES 63
- 4.1 Light and Charge Production in Liquid Xenon 63
  - 4.1.1 Recombination: The Light-Charge Partition 64
  - 4.1.2 Thomas-Imel and Doke-Birks Models 65
  - 4.1.3 Quanta and Recombination Fluctuations 69
  - 4.1.4 From Recombination to Yield Models 72
- 4.2 Yields and Clustering in Simulations 75
- 4.3 Detector Response Modelling with Single-Beta ^{214}Bi 79
- 4.4 Physics-Based Lineage Clustering 82

4.4.1 Principle: Clustering by Particle Lineage 83
4.4.2 Lineage Break Rules 83
4.4.3 Illustration and Impact 85
4.5 Validation with Multiple Sources 86
4.5.1 ^{214}Bi Ground-State Decays 86
4.5.2 ^{214}Bi with Excited-State Decays 86
4.5.3 ^{212}Pb Beta and Gamma Decays 88
4.5.4 ^{232}Th High-Energy Gamma Rays 88
4.6 Limitations and Possible Applications 89
4.6.1 The Recombination Fluctuation Model 89
4.6.2 Conversion-Electron Dominated Transitions 91
4.6.3 Implications for Double-Beta-Decay Studies 95
4.7 Summary and Implications 97
5 MEASUREMENT OF THE ^{214}BI ENERGY SPECTRUM 99
5.1 Introduction 99
5.2 Theoretical Framework 101
5.3 Experimental Method 103
5.3.1 XENONnT Detector and Datasets 103
5.3.2 Event Topology 105
5.3.3 Event Selection 107
5.3.4 Simulations and Signal Modeling 108
5.3.5 Energy Bias Correction 112
5.4 Results 114
5.4.1 Models to Data Comparison 114
5.4.2 Systematic Uncertainties 116
5.5 Conclusion and outlook 117
5.6 Appendix 117
5.6.1 Yields Parameters 117
5.6.2 Energy Bias Fit Results 118
6 FERMIONIC DARK MATTER ABSORPTION 119
6.1 Theory and Nuclear Inputs 121
6.1.1 Effective Interaction and Event Rate 121
6.1.2 ^{136}Xe Nuclear Structure 124
6.1.3 ^{136}Xe Signal Topology 126
6.2 Analysis Methods 128
6.2.1 ML Segmentation Model 128
6.2.2 Model Performance and Validation 130
6.2.3 Event Selection 131
6.2.4 Signal Template Construction 133
6.3 Background Model 134
6.3.1 Pile-up and Neutron Background 136
6.3.2 Muon-Spallation Products 136

6.3.3 Solar-Neutrino Interactions 138
6.4 Statistical Inference 141
6.4.1 Template Profile-Likelihood 142
6.5 Results 144
6.5.1 Observed Data 144
6.5.2 Observed Limits 144
6.5.3 Discussion 148
6.6 Conclusion 148
A XENONNT SOFTWARE: DESIGN AND EXPERIENCE 151
A.1 The Infrastructure 151
A.1.1 The Analysis Framework 151
A.1.2 Corrections and Reproducibility 154
A.1.3 Data Acquisition and Storage 155
A.2 Observations 157
A.2.1 What Worked 157
A.2.2 The Reprocessing Bottleneck 158
A.2.3 The Missing Analysis Layer 159
A.2.4 Shared Storage Constraints 160
A.2.5 Onboarding Scales with Proximity 161
A.2.6 Software Governance 161
A.3 Conclusions 162
SAMENVATTING 163
RIASSUNTO 169
ACKNOWLEDGEMENTS 177
BIBLIOGRAPHY 185

SUMMARY

of "A Matter of Coincidences"

"Come sta il barile?" is the question I have been asked at every family gathering for the past few years. The *barrel* is the funny name my family has for the detector I have been working with for my PhD: the XENONnT experiment. My usual answer is that the *barrel* is doing fine, sitting under a rock with nothing to do. The truth is not so far from that, but in this thesis I will try to give a more detailed answer.



The barrel itself is a tank filled with liquid xenon, buried under a mountain at the underground Laboratori Nazionali del Gran Sasso in central Italy, and sitting inside a larger tank of water for further shielding (Figure 1). It was built to catch a particle of dark matter: the invisible substance that makes up most of the matter in the universe, felt through its gravity but never directly detected. Dark matter is thought to fill our galaxy and drift through ordinary matter almost undisturbed. Because it barely interacts with anything, it should easily pass through a mountain of rock, and so we hope it bumps into one of the heavy xenon atoms in our detector.

At the same time, plenty of other things keep interacting with the xenon. Cosmic rays are one source, though the mountain above the laboratory blocks most of them before they arrive. Neutrinos from the Sun are another source: each one interacts only very rarely, but the Sun sends so many of

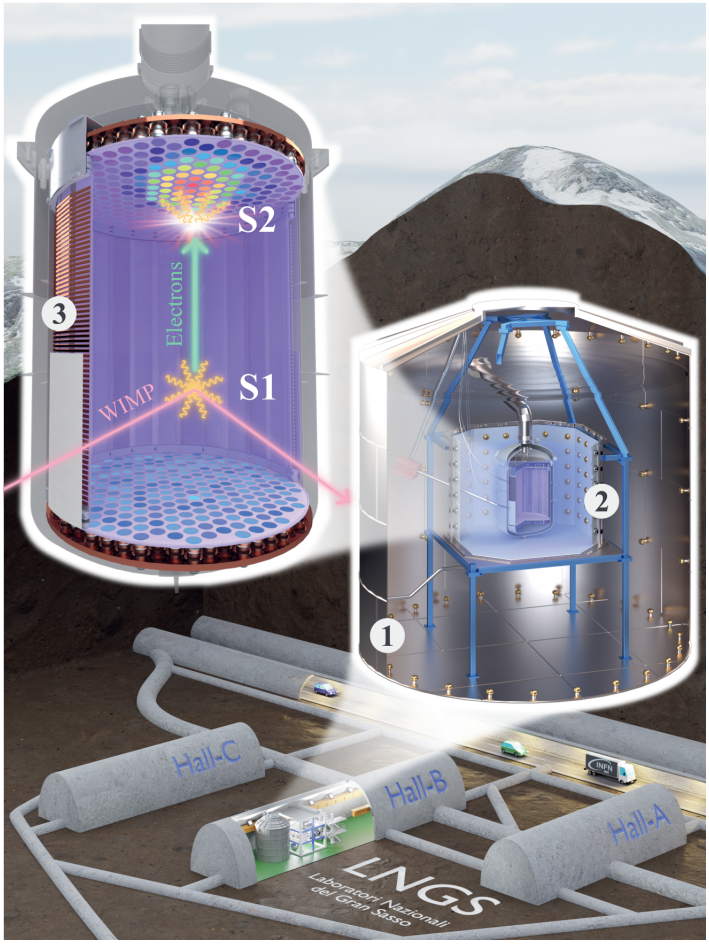


Figure 1 *The XENONnT detector at the Laboratori Nazionali del Gran Sasso (LNGS), Italy. The liquid xenon tank (3) is surrounded by light sensors above and below. A particle interaction produces the flash of light we call S1, and a second, larger flash, S2, once the freed electrons drift up and reach the gas layer. The detector sits inside a water tank (2) for shielding, inside Hall B of the underground laboratory (1). Figure taken from [2].*

them that some strike the xenon anyway, and when they do, they can look a lot like what we expect from dark matter.

But the biggest source is in the detector's own surroundings. Ordinary materials, rock, the detector's own components, are slightly radioactive, and even the liquid xenon itself is not perfectly pure: trace amounts of radon, a radioactive gas, are dissolved in it. Their atoms are unstable and decay one by one, emitting alpha particles, electrons, or gamma rays, and

occasionally neutrons. The water tank around the detector stops most of the neutrons, but the rest of this radiation reaches the xenon constantly. Keeping it under control took extreme care in choosing and cleaning every material in the detector, making XENONnT one of the cleanest, quietest places in the world.

When a particle interacts inside the xenon, two things happen: electrons are liberated from xenon atoms, and a flash of light is produced. Sensors around the tank see this flash right away, and we call this first signal S1. The freed electrons then drift up through the liquid under an electric field. A few microseconds later they reach the gas above the liquid and make a second, bigger flash, which we call S2. Together, S1 and S2 tell us where the interaction happened and how much energy it deposited. This process, and the details of the detector that make it possible, from the photomultiplier tubes to the electronics that reconstruct S1 and S2, are the subject of Chapter 2. Interpreting it precisely also means trusting a simulation of the detector. Chapter 3 describes `fuse`, the simulation framework that reproduces the entire process from a particle interaction to a digitised waveform, producing fake data nearly indistinguishable from the real thing.

XENONnT is the latest step in a hunt for dark matter that the XENON collaboration has pursued for almost 25 years, through a series of detectors, XENON10, XENON100, XENON1T, each one bigger and cleaner than the previous one. The current version holds 8.6 tonnes of xenon and has been running since 2020, without finding dark matter yet.

While waiting to discover dark matter, the same detector is also good at other things. Its very low background and its ability to measure S1 and S2 with high precision make it useful well beyond the search it was built for. In this thesis, we use data from the detector to do three things: to model how liquid xenon converts energy into light and charge, to measure a nuclear decay's energy spectrum precisely enough to test different theories of nuclear structure, and to search for dark matter through an alternative interaction, absorption rather than scattering.

All three analyses also rely on the same trick to fight background. In each case, the process we are after produces not one signal, but two: a first signal followed, a short and predictable time later, by a second one. Ordinary background essentially never does this, so requiring both signals, in the right order and within the right time window, removes almost all of it at once. This is usually called a *delayed coincidence*.

One of those decay chains is particularly useful. An atom of ^{214}Bi decays by emitting an electron, turning into ^{214}Po, which itself decays a fraction of a millisecond later by emitting an alpha particle. This is the signature we use: whenever we see this pair, we know we caught a ^{214}Bi

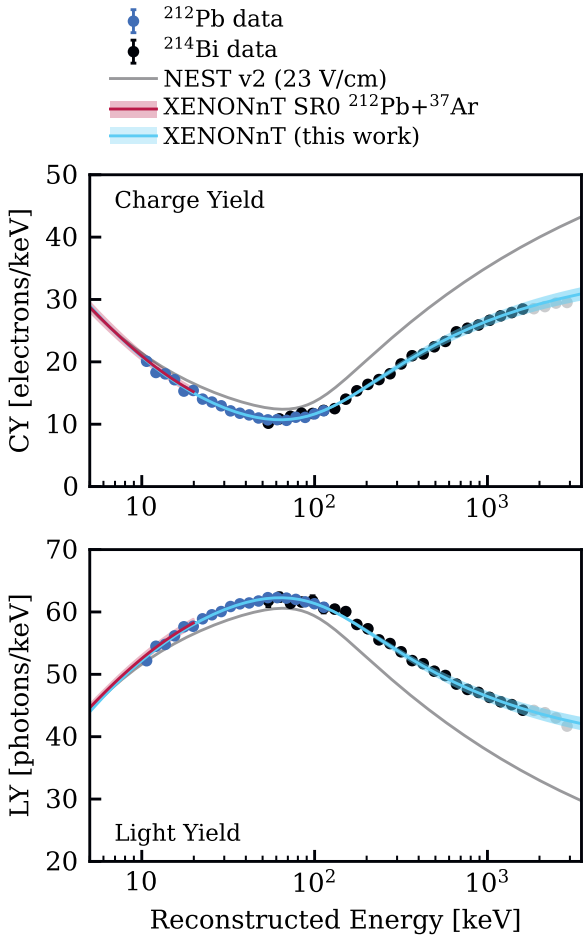


Figure 2 *How a given amount of deposited energy splits between light (LY, bottom) and charge (CY, top) in liquid xenon, extracted from the tagged ^{214}Bi decays described in the text. This measurement corrects a substantial disagreement in earlier models and is used throughout this thesis.*

decay, cleanly separated from everything else happening in the detector. Chapter 4 uses this tagged sample to work out how a given amount of deposited energy splits between the light and charge signals in liquid xenon, a relationship that previous models got wrong by a wide margin at these energies (Figure 2). Getting it right matters for every other measurement described in this thesis, since it sets the ruler everything else is measured against.

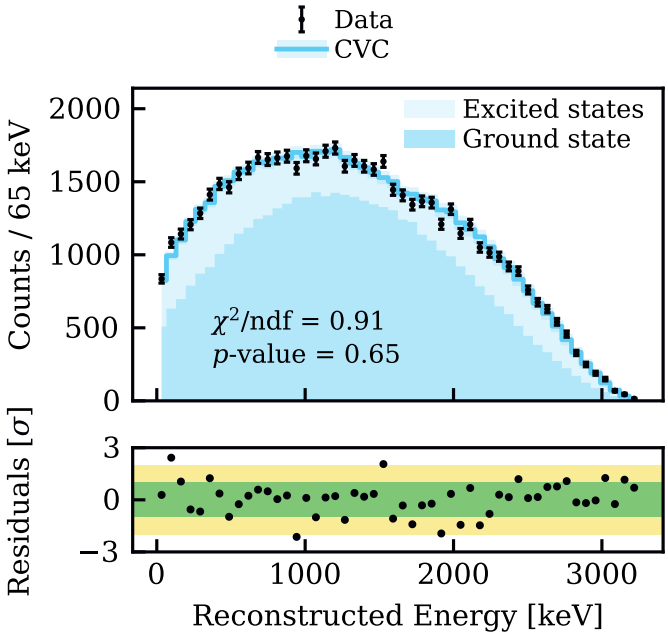


Figure 3 *Measured ^{214}Bi beta spectrum compared to a full-chain `fuse` simulation based on the general symmetry argument (CVC) described in the text, which matches the data best, with residuals shown below. Two more detailed nuclear-structure calculations, not shown here, disagree with the data by comparison.*

With that ruler calibrated, Chapter 5 goes a step further and asks what the beta decay itself looks like in detail: not just how much energy it releases on average, but the full spectrum of energies the electron can carry away. That spectrum is not a simple curve. Its exact shape depends on the internal structure of the atomic nucleus in a way that is hard to calculate, and different nuclear models disagree noticeably on what it should look like. Chapter 1 lays out the nuclear physics behind this: why the shape is hard to predict, and what a precise measurement of it can teach us about the nucleus. Using the same tagged sample of clean ^{214}Bi decays, we measured that spectrum with enough precision to tell the competing models apart (Figure 3).

Chapter 6 turns to dark matter again, but from a different angle. Most searches look for a dark matter particle bumping into a xenon nucleus and knocking it slightly, the way one billiard ball nudges another. For very light dark matter, that nudge is too gentle to see. An alternative is absorption: the dark matter particle is not deflected but destroyed, handing its entire

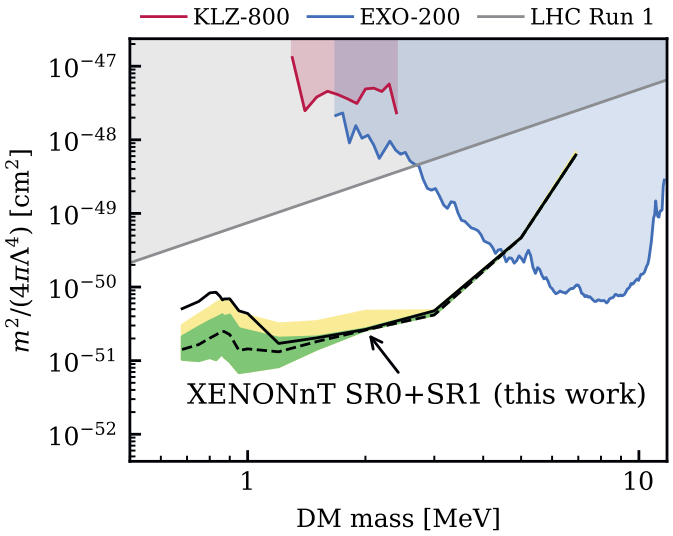


Figure 4 *Upper limits at 90% confidence on the strength of the dark matter absorption process searched for here, as a function of dark matter mass, from this work and from other experiments sensitive to the same process. In part of the mass range, this measurement improves on the previous best limit by up to four orders of magnitude.*

mass over to the nucleus in one go. In xenon, this can leave behind an excited nucleus that settles down shortly afterwards by emitting an extra, delayed flash of light of its own, on top of the original signal. Requiring that second, delayed flash strips away almost all of the ordinary backgrounds that would otherwise swamp the search. It is the same trick that made the beta-decay measurement possible, now applied to a different physics question.

No dark matter signal was found. But in part of the mass range, the limits this search places on how strongly it could interact are around ten thousand times tighter than what was known before (Figure 4).

The thesis closes with an appendix on the software: the infrastructure that processes, stores, and distributes XENONnT's data, built and kept running over several years by many people across the collaboration. It gives a short description of the main software and infrastructure, followed by some personal reflections on the practicalities of running and using it.

The barrel, it turns out, had more to do than sit there and wait.

1

BETA DECAYS

Nuclear beta decay has been studied for over a century [3, 4] and continues to yield new insights into nuclear structure and the weak interaction. In its most common form, a neutron converts to a proton, emitting an electron and an antineutrino ($n \to p + e^- + \bar{\nu}_e$). When this occurs inside a nucleus, the nuclear charge shifts by one unit: (Z, N) becomes $(Z+1, N-1)$. For most unstable nuclei, beta decay is the dominant decay mode.

Unlike alpha decay or gamma emission, which produce radiation at fixed energies, the beta electron shares the available energy with the antineutrino across a continuous distribution running from zero to the endpoint Q-value. Its shape is governed by a dimensionless function of electron energy, the shape factor $C(E_e)$. For the simplest decays it is constant and the spectral shape is set entirely by phase space. For others it depends on the nuclear transition in a way that can visibly distort the spectrum and encodes information about the structure of the nucleus.

The ^{214}Bi $\to$ ^{214}Po decay studied in this thesis is a case where $C(E_e)$ is not fixed by kinematics. As a first-forbidden non-unique transition (Section 1.2), its shape factor depends on several independent nuclear matrix elements that must be computed from a nuclear model. Different models give different matrix elements and therefore different spectral shapes. Measuring the spectrum precisely enough to distinguish these predictions is the subject of Chapter 5.

The nuclear-model question is relevant beyond ^{214}Bi. Neutrinoless double beta decay ($0\nu\beta\beta$), the process $(A, Z) \to (A, Z+2) + 2e^-$ with no neutrino emission, would establish that neutrinos are Majorana particles. The neutrino mass inferred from a measured half-life depends inversely on the squared nuclear matrix element $|M^{0\nu}|^2$ [5, 6]. Predictions from different models spread by roughly a factor of three in $|M^{0\nu}|$, producing an order-of-magnitude range in the extracted mass for any given experimental limit [6]. The same shell-model and QRPA methods used here are the main tools for these calculations. Testing them against a measured beta spectrum in the same mass region is one concrete check on their reliability.

A second application arises in Chapter 6, which searches for fermionic dark matter absorption on ^{136}Xe. The signal rate in the charged-current

channel $\chi + {}^{136}\mathrm{Xe} \to e^- + {}^{136}\mathrm{Cs}^{(*)}$ depends on the Gamow-Teller matrix element for the ${}^{136}\mathrm{Xe} \to {}^{136}\mathrm{Cs}$ transition [7]. The nuclear-structure framework of Sections 1.2–1.5 applies directly.

The common thread is the nuclear matrix element: understanding what determines its value, how models compute it, and how a measured beta spectrum tests those calculations. Behrens & Bühring [8] and Hayen et al. [9] give comprehensive formal treatments. The focus here is on the physical picture and the specific machinery needed to understand the shape factor and its model dependence. Section 1.1 derives the beta spectrum and introduces the shape factor. Section 1.2 covers the classification of transitions, with emphasis on the unique versus non-unique distinction. Section 1.3 decomposes the shape factor into nuclear matrix elements. Section 1.4 introduces the conserved vector current hypothesis, which constrains the vector NMEs and explains, via the PCAC relation, why g_A is not similarly protected. Section 1.5 covers the nuclear-structure approaches used to compute the axial NMEs and the phenomenology of g_A quenching. Section 1.6 applies this framework to the specific case of ${}^{214}\mathrm{Bi}$, including the precision corrections to the spectrum needed for a quantitative comparison with data.

1.1 Fermi Theory and the Beta Spectrum

Beta decay is a process governed by the weak interaction, mediated by the exchange of a massive $W^{\pm}$ boson between hadronic and leptonic currents. At the energies relevant for nuclear decays the momentum transfer is far below the W mass, and the interaction reduces to an effective four-fermion contact term. This is Fermi's original vector interaction [10]. An axial-vector component is also required, as evidenced by spin-changing transitions identified by Gamow and Teller [11]. Wu et al. [12] demonstrated parity violation in 1957, ruling out scalar and tensor contributions. The specific V$-$A combination was established by Sudarshan and Marshak [13, 14] and by Feynman and Gell-Mann [15]. The physical content is that the weak interaction couples exclusively to left-handed fermions. A pure vector interaction couples equally to both helicities. A pure axial-vector interaction does the same but with opposite sign. Their difference, V$-$A, projects out the left-handed component entirely. Wu et al. observed that the emitted electrons are preferentially left-handed, which is only consistent with a maximally left-handed current. The V$-$A structure determines which nuclear operators contribute to the transition amplitude, and thereby what decay channels are accessible, as discussed in Section 1.2.

The probability density $P(E_e)$ for the electron to be emitted with total energy E_e follows from Fermi's golden rule. Assuming the decaying nucleon

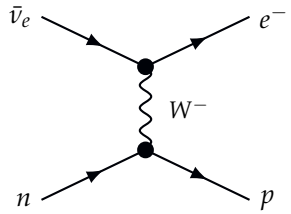


Figure 1.1 *Beta decay at the nucleon level, $n \to p + e^- + \bar{\nu}_e$, mediated by the W^- boson. At nuclear energies the W^- propagator reduces to Fermi's effective four-fermion contact interaction [10]. The hadronic vertex couples to both the vector ($\hat{\tau}_\pm$) and axial-vector ($\hat{\sigma}\hat{\tau}_\pm$) currents of the $V-A$ structure, as discussed in Section 1.2.*

does not interact with the remaining $A - 1$ nucleons, it takes the form

$$P(E_e)\,\mathrm{d}E_e = \frac{G_F^2}{(\hbar c)^6}\,\frac{1}{2\pi^3\hbar}\,C(E_e)\,p_e c\,E_e\,(E_0 - E_e)^2\,F_0(Z, E_e)\,\mathrm{d}E_e, \tag{1.1}$$

where p_e is the electron momentum, $E_0 = Q + m_e c^2$ the maximum total electron energy, Z the proton number of the daughter nucleus, and G_F the Fermi constant as determined from muon decay. In nuclear beta decay the effective coupling is $G_F|V_{ud}|$, where $|V_{ud}|$ is the CKM matrix element governing the $\mathrm{d} \to \mathrm{u}$ quark transition that underlies beta decay at the microscopic level. The factor $|V_{ud}|^2 \approx 0.948$ is absorbed into the normalisation here and discussed further in Section 1.4. Three factors do distinct jobs: the phase-space factor $p_e c\,E_e\,(E_0 - E_e)^2$, the Fermi function $F_0(Z, E_e)$, and the shape factor $C(E_e)$.

The factor $p_e c\,E_e\,(E_0 - E_e)^2$ is phase space. It counts the number of ways the electron and neutrino can share the available energy, and it gives the spectrum its characteristic shape: rising from zero at threshold, peaking broadly in the middle, and returning to zero at E_0. The continuous nature of the spectrum confused physicists in the 1920s and led Pauli to postulate the neutrino in 1930 [16]. It follows directly from the phase space: the electron can take any share of the available energy, with the neutrino carrying the rest undetected.

The Fermi function $F_0(Z, E_e)$ accounts for the Coulomb interaction between the outgoing electron and the daughter nucleus. For β^- emission the nuclear charge attracts the electron, enhancing the decay rate at low energies. For β^+ emission the sign reverses and the Coulomb barrier suppresses it. The effect scales with Z, and for a heavy nucleus like ^{214}Po it is a significant correction across the full spectrum. It is determined by the electron wavefunction evaluated near the nuclear surface, where the Coulomb field is strongest. For $Z = 84$, the Coulomb parameter $\alpha Z \approx 0.61$

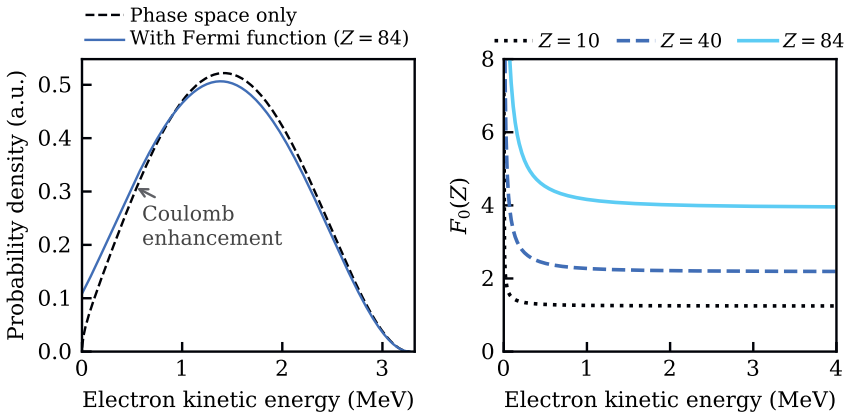


Figure 1.2 *Left: kinetic energy spectrum of the emitted electron for ^{214}Bi ($Q = 3.27$ MeV). The Fermi function uses $Z = 84$ for the daughter ^{214}Po. Phase space only (dashed) gives a symmetric shape. Including $F_0(Z)$ (solid) redistributes probability toward lower energies through Coulomb enhancement. Both curves are normalised to unit area. Right: the Fermi function $F_0(Z)$ for three values of Z. The low-energy enhancement grows strongly with Z.*

(where $\alpha \approx 1/137$ is the fine-structure constant) is not small, and it controls the strength of relativistic corrections to the wavefunction at short distances. The nuclear field pulls the electron to locally relativistic velocities close to the surface regardless of its asymptotic energy. The non-relativistic wavefunction gives the wrong electron density there, and the Dirac equation is required [17].

Figure 1.2 illustrates the roles of phase space and the Fermi function for ^{214}Bi, showing how Coulomb enhancement redistributes the spectrum toward lower energies and how strongly this effect grows with Z.

The shape factor $C(E_e)$ carries the nuclear structure information. Transitions where the lepton pair (electron and antineutrino) carries zero orbital angular momentum ($L = 0$) are called *allowed*. Those requiring $L > 0$ are *forbidden*. The full classification is in Section 1.2. For an allowed transition, $C(E_e)$ is constant, and the spectral shape reduces to phase space times the Fermi function. For forbidden transitions it acquires an energy dependence that can visibly distort the spectrum. The source of that dependence is the subject of Sections 1.2 and 1.3.

The Kurie plot [18] makes the distortion easy to identify. It is defined as

$$K(E_e) \equiv \sqrt{\frac{P(E_e)}{p_e c\, E_e\, F_0(Z, E_e)}} \propto \sqrt{C(E_e)}\,(E_0 - E_e), \tag{1.2}$$

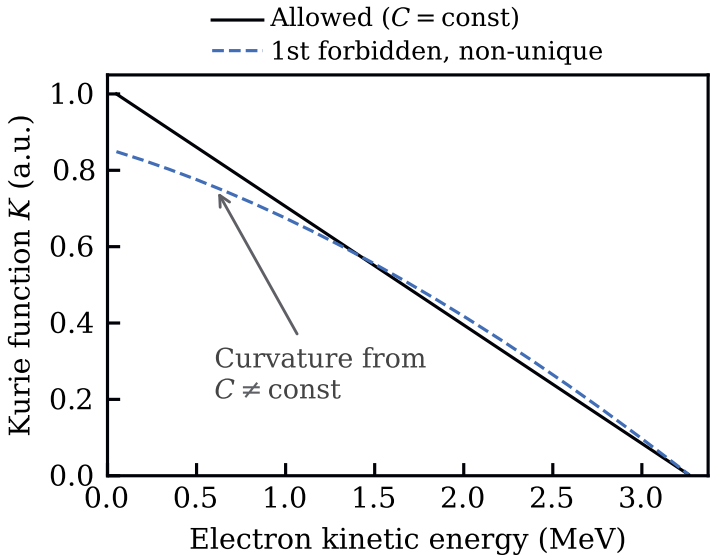


Figure 1.3 *Kurie plot $K \propto \sqrt{P(E_e)/(p_e E_e F_0)}$ as a function of electron kinetic energy, for an allowed decay (solid) and a schematic first-forbidden non-unique decay (dashed). For the allowed case $K \propto (E_0 - E_e)$, giving a straight line that intersects zero at the endpoint. The non-constant shape factor $C(E_e)$ causes the forbidden spectrum to deviate from this line across the full energy range. Both curves use $Z = 84$ and $Q = 3.27\,\mathrm{MeV}$. The shape factor is parametric and not specific to a particular nuclear model.*

so that an allowed decay ($C = \text{const}$) produces a straight line intersecting zero at E_0. Any curvature in the Kurie plot is a direct signature of $C(E_e) \neq \text{const}$. The Kurie plot was also the original tool used to bound the neutrino mass: a finite m_ν rounds the spectrum near E_0, a distortion that direct mass measurements search for. Figure 1.3 shows the contrast between an allowed decay and a first-forbidden non-unique one (as classified in Section 1.2): the distortion runs across the whole spectrum, not only near the endpoint.

1.2 Selection Rules and Classification of Beta Decays

The transition amplitude in Fermi's golden rule involves the overlap of the leptonic current with the nuclear transition operator. In nuclear beta decay, the outgoing leptons have momenta of at most a few MeV/c, corresponding to de Broglie wavelengths of $\lambda \sim \hbar c/p \gtrsim 200\,\mathrm{fm}$. For a nucleus with $A = 214$, the nuclear radius is $R \approx 1.2\,A^{1/3}\,\mathrm{fm} \approx 7\,\mathrm{fm}$ [19], so $p_e R/\hbar \ll 1$ across most of the spectrum. Under this *allowed approximation*, the lepton

wavefunctions are treated as constant over the nuclear volume and evaluated at the origin. The lepton field can be expanded in partial waves of orbital angular momentum L. Only the $L = 0$ component survives in this limit. This is the defining condition of an *allowed* decay.

Which nuclear operators can drive a given transition is governed by a general rule from angular momentum addition. A rank-k operator carries k units of angular momentum under rotations and can only connect states whose total angular momenta satisfy $|J_i - J_f| \leq k$; if $\Delta J > k$, the matrix element vanishes identically. The classification of transitions in this section follows from this rule.

Allowed transitions

Two classes of nuclear operators contribute to allowed decays. Here $\Delta J \equiv |J_f - J_i|$ denotes the change in total nuclear spin, and $\Delta\pi \equiv \pi_i \pi_f \in \{+1, -1\}$ is the parity product of initial and final states ($\Delta\pi = +1$: no parity change; $\Delta\pi = -1$: parity changes). The isospin raising and lowering operators $\hat{\tau}_\pm$ convert a neutron into a proton (or vice versa), acting on a single nucleon within the nucleus. Since they are rank-0 operators carrying no angular momentum, the total nuclear angular momentum cannot change: *Fermi transitions* require $\Delta J = 0$ and $\Delta\pi = +1$. The operator $\hat{\boldsymbol{\sigma}}\hat{\tau}_\pm$ additionally acts on the nucleon spin via the Pauli matrix $\hat{\boldsymbol{\sigma}}$, allowing a spin flip that can shift the total angular momentum by one unit. *Gamow-Teller transitions* have $\Delta J = 0$ or 1 (excluding $0 \to 0$) and $\Delta\pi = +1$. The $0 \to 0$ case is excluded because $\hat{\boldsymbol{\sigma}}$ is a rank-1 operator, carrying one unit of angular momentum under rotations. Coupling one unit of angular momentum onto a $J = 0$ state cannot produce another $J = 0$ state. Every allowed beta decay proceeds through one or both channels, weighted by the corresponding nuclear matrix elements M_F and M_{GT} [8, 20].

The shape factor for any allowed transition is constant: $C(E_e) = g_V^2 M_F^2 + 3 g_A^2 M_{GT}^2$, where the factor of 3 arises from summing over the three components of $\hat{\boldsymbol{\sigma}}$, and M_{GT} is the matrix element of a single spin component. The spectral shape is therefore the same for all allowed decays with the same Z and Q-value, determined entirely by phase space and the Fermi function.

Forbidden transitions

When the angular momentum and parity quantum numbers of the nuclear transition cannot be satisfied with $L = 0$ lepton emission, higher multipoles must contribute. Physically, this happens when the lepton wavefunction can no longer be treated as uniform over the nuclear volume. Each forbidden

order corresponds to one more power of $p_e R/\hbar$ in the multipole expansion of the lepton field. The nth-forbidden class corresponds to lepton partial waves with $L \geq n$. Each additional factor of the expansion parameter $\xi \equiv p_e R/\hbar$ suppresses the transition amplitude. The physical origin is straightforward: for $L = 0$, the lepton wavefunction is treated as constant across the nucleus, and only its value at the origin matters. For $L > 0$, the field must vary spatially across the nuclear volume. Each multipole of order L requires L spatial derivatives of the wavefunction within the nucleus, and each derivative introduces one factor of $p/\hbar$ evaluated at R. The expansion parameter ξ is exactly that factor. For first-forbidden decays in the $A \sim 200$ region, $\xi^2 \sim 10^{-3}$ to 10^{-2} across the spectrum [8, 20]. This makes the decay roughly two orders of magnitude slower than a comparable allowed transition with the same Q-value. Despite this suppression, they are experimentally accessible. For many heavy neutron-rich nuclei, including ^{214}Bi, all energetically allowed transitions to the daughter ground state involve a parity change, so first-forbidden decays are the only available channel.

The parity selection rule alternates with forbidden order: $\Delta\pi = (-1)^n$ [8]. First-forbidden transitions involve a parity change. Second-forbidden do not.

Unique versus non-unique

Within a given forbidden class, the transition is called *unique* if the nuclear spin change takes its maximum value, $\Delta J = n + 1$, where n is the order of forbiddenness ($n = 1$ for first-forbidden, $n = 2$ for second-forbidden, and so on). For unique forbidden transitions, only a single combination of nuclear matrix elements contributes to the shape factor, and its momentum dependence is fully fixed by kinematics. This occurs because only a single type of nuclear operator can drive a transition with $\Delta J = n + 1$: lower-rank operators cannot connect states with that large an angular momentum difference, as follows from the rank–ΔJ rule above. With only one operator contributing, its energy dependence is fully determined by kinematics, and no nuclear-model input is needed to predict the spectral shape. This property has been exploited in precision determinations of the effective g_A value (Section 1.5). Deviations of the measured spectrum from the kinematic prediction provide a model-independent observable without nuclear-model input [21, 22].

For *non-unique* forbidden transitions, $\Delta J < n + 1$ and multiple NME combinations enter $C(E_e)$ simultaneously. Their relative magnitudes and signs depend on nuclear structure and must be computed from a model. For non-unique decays, this model dependence is what makes the spectrum

a direct probe of nuclear matrix elements. The $1^- \to 0^+$ transition of ^{214}Bi is of this type, and its specific structure is discussed in Section 1.6.

Table 1.1 summarises the selection rules for the transitions most relevant in experimental beta spectroscopy.

The same classification applies in Chapter 6, which searches for fermionic dark matter absorption on ^{136}Xe. The dominant nuclear transition there is a Gamow-Teller $0^+ \to 1^+$ excitation of ^{136}Cs, directly analogous to the allowed Gamow-Teller transitions described above.

Table 1.1 *Classification of beta-decay transitions. The nth-forbidden class involves lepton partial waves with $L \geq n$. Unique transitions have $\Delta J = n + 1$ and a kinematically determined spectral shape. Non-unique transitions have $\Delta J < n + 1$ and a model-dependent $C(E_e)$. The symbol * indicates that $0 \to 0$ is excluded for GT transitions.*

Type	ΔJ	$\Delta\pi$	$C(E_e)$
Allowed (Fermi)	0	no	constant
Allowed (Gamow-Teller)	$0, 1^*$	no	constant
1st forbidden, non-unique	0, 1	yes	model-dep.
1st forbidden, unique	2	yes	kinematic
2nd forbidden, non-unique	1, 2	no	model-dep.
2nd forbidden, unique	3	no	kinematic

1.3 The Shape Factor and Nuclear Matrix Elements

The shape factor $C(E_e)$ in Eq. (1.1) can be decomposed into contributions from the vector, axial-vector, and mixed components of the V$-$A interaction [8, 20, 23]:

$$C(E_e) = g_V^2\, C_V(E_e) + g_A^2\, C_A(E_e) + g_V g_A\, C_{VA}(E_e), \tag{1.3}$$

where g_V and g_A are the vector and axial-vector coupling constants. For an allowed transition, all three components reduce to constants and $C(E_e)$ is energy-independent. For forbidden transitions, each component is a polynomial in the electron total energy E_e and the neutrino momentum $p_\nu = (E_0 - E_e)/c$. For first-forbidden non-unique decays, C_V, C_A, and C_{VA} contain terms proportional to p_e^2, p_ν^2, $p_e p_\nu$, and $E_e p_\nu$, each multiplied by a different nuclear matrix element. These combinations arise from bilinear products of the electron and neutrino wavefunctions evaluated at first order in the $pR/\hbar$ expansion.

The kinematic factors in $C(E_e)$ are fully determined by the leptonic wavefunctions and carry no nuclear information. The nuclear information enters through a set of numbers called nuclear matrix elements (NMEs). Each NME quantifies how strongly a specific nuclear operator $\hat{O}_k$ can drive the transition from the initial to the final nuclear state:

$$M_k = \langle J_f,\, \pi_f \mid \hat{O}_k \mid J_i,\, \pi_i \rangle, \tag{1.4}$$

where J_i, π_i and J_f, π_f are the total angular momentum and parity of the initial and final nuclei. For ^{214}Bi $\rightarrow$ ^{214}Po, the initial state has $J_i^\pi = 1^-$ and the final state $J_f^\pi = 0^+$. Each operator $\hat{O}_k$ acts on individual nucleons within the many-body nuclear state. M_k measures the overlap between the final nuclear wavefunction and the initial one after $\hat{O}_k$ has acted on it.

The shape factor is then a sum over all contributing operators:

$$C(E_e) = \sum_k g_k^2\, M_k^2\, f_k(p_e,\, p_\nu) + \sum_{k \neq l} g_k\, g_l\, M_k\, M_l\, h_{kl}(p_e,\, p_\nu), \tag{1.5}$$

where f_k and h_{kl} are kinematic polynomials in the electron and neutrino momenta, fully determined by the leptonic wavefunctions and independent of nuclear structure, and $g_k \in \{g_V,\, g_A\}$ is the coupling constant for operator k. The NMEs M_k are real-valued. Time-reversal invariance of the nuclear operators permits a phase convention in which all matrix elements entering $C(E_e)$ are real, so $M_k^2 = |M_k|^2$ and $M_k M_l = \mathrm{Re}(M_k M_l^*)$. The squared terms M_k^2 arise from squaring the decay amplitude. The cross terms $M_k M_l$ arise from interference between different operators. Because different terms carry different powers of p_e and p_ν, varying the NME values shifts spectral weight across the energy range in distinct ways. Figure 1.4 illustrates this. NMEs multiplying p_e^2 terms shift weight toward higher electron energies. NMEs multiplying p_ν^2 terms do the opposite, since p_ν increases as E_e decreases. Each combination produces a different distortion visible across the whole spectrum.

For a first-forbidden non-unique decay, the relevant operators are the rank-0 axial charge (the time component of the axial current), the rank-1 axial current (the space components), and a rank-1 vector operator constrained by CVC (Section 1.4). The rank determines which values of ΔJ the operator can connect. Writing out all combinations explicitly, including squared terms, cross terms, and small relativistic corrections, produces the lengthy shape-factor expression given in Refs. [8, 22].

NMEs are not directly observable. They are the output of nuclear structure calculations, and their values depend on the model used to describe the wavefunctions of ^{214}Bi and ^{214}Po. For $A = 214$, exact solutions

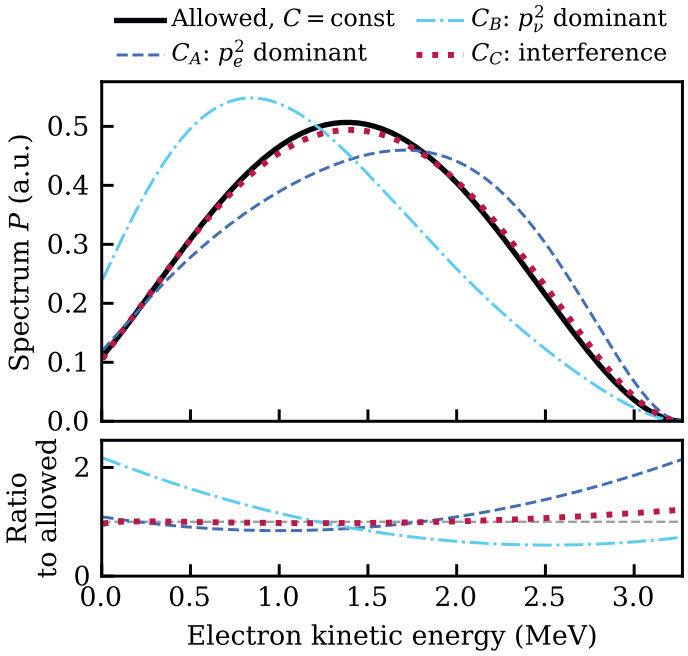


Figure 1.4 *Upper panel: schematic beta spectra illustrating the effect of a non-constant shape factor $C(E_e)$ for $Z = 84$, $Q = 3.27\,\text{MeV}$. The allowed case ($C = \text{const}$, solid) is shown alongside three parametric first-forbidden non-unique shape factors: one dominated by the p_e^2 NME term (dashed), one by the p_ν^2 term (dash-dot), and one with a mixed interference contribution (dotted). All curves are normalised to unit area. Lower panel: ratio of each spectrum to the allowed case. The deviations reach 30 to 50% and are present across the full energy range, not confined to the endpoint. The shape factors are schematic. Specific NME values for ^{214}Bi are discussed in Chapter 5.*

of the many-body Schrödinger equation are not available. Every practical approach makes approximations, as discussed in Section 1.5.

Because different NMEs multiply terms with different powers of p_e and p_ν, the shape factor varies across the spectrum. For ^{214}Bi, the effect is large: the momentum-dependent terms cause deviations from the pure phase-space shape of 10 to 30% across the full energy range from threshold to Q [24]. This is not a small correction confined to the endpoint region.

Not all NMEs entering $C(E_e)$ for first-forbidden transitions are of the same order. The small relativistic vector NMEs, or sNMEs, arise from relativistic corrections to the vector current and are suppressed relative to the leading contributions by a factor of order $v/c \approx 0.1$ for typical nucleon momenta. Despite this suppression, they can significantly affect

the predicted spectrum when cancellations occur among the larger terms. Their specific role in the ^{214}Bi decay is discussed in Section 1.6.

1.4 Conserved and partially conserved currents

Under the approximate isospin symmetry of the strong interaction, the proton and neutron are two states of the same particle, differing only in their isospin projection $I_3 = \pm\frac{1}{2}$. This symmetry extends to nuclei: for a given mass number A, the states $(A, Z-1)$, (A, Z), and $(A, Z+1)$ form an isospin triplet with $I = 1$ and $I_3 = -1, 0, +1$ respectively. The isovector electromagnetic current, the component of the EM current that distinguishes protons from neutrons, is the $I_3 = 0$ member of this same triplet.

The conserved vector current (CVC) hypothesis, proposed by Feynman and Gell-Mann in 1958 [15], identifies the vector component of the charged weak hadronic current as the $I_3 = \pm 1$ members of this triplet. Beta decay and the isovector electromagnetic transition are then different components of the same underlying current (Figure 1.5). The electromagnetic current is exactly conserved by gauge invariance in QED, so the weak vector current, as a member of the same multiplet, is also conserved [20].

CVC has two consequences relevant to beta spectroscopy. First, g_V is not renormalised by the nuclear medium: $g_V = 1.0$ exactly in any nucleus [25]. The reasoning parallels the conservation of electric charge: a conserved current carries a conserved charge that the surrounding medium cannot shift. The full beta-decay coupling is $G_F|V_{ud}|g_V$, where $|V_{ud}|$ is the CKM element governing the $d \to u$ quark transition (Section 1.1). CVC states that the nuclear medium leaves g_V at its free-nucleon value. Second, the vector form factors entering the forbidden shape factor are related, through the isospin-triplet structure, to isovector electromagnetic matrix elements. These are fixed by the same isospin symmetry that defines the triplet, without a free nuclear-model parameter. The two consequences are distinct. The first concerns the overall coupling strength and follows from current conservation alone. The second constrains the momentum-dependent form factors through the full triplet connection.

The second consequence is most transparent for the $1^- \to 0^+$ transition of ^{214}Bi. The Fermi operator $\hat{\tau}_\pm$ is a rank-0 isospin scalar and cannot connect states with $\Delta J = 1$, so M_F vanishes exactly for this transition. The remaining vector contribution enters through the *weak magnetism* form factor. The nucleon is a composite object with a magnetic moment, and this generates an induced correction to its weak vector current. Under CVC, this form factor is constrained by the isovector magnetic moment $\mu_p - \mu_n = 4.706$ nuclear magnetons, known from electromagnetic measurements [25, 26]. No free nuclear-model parameter enters this relation. The vector

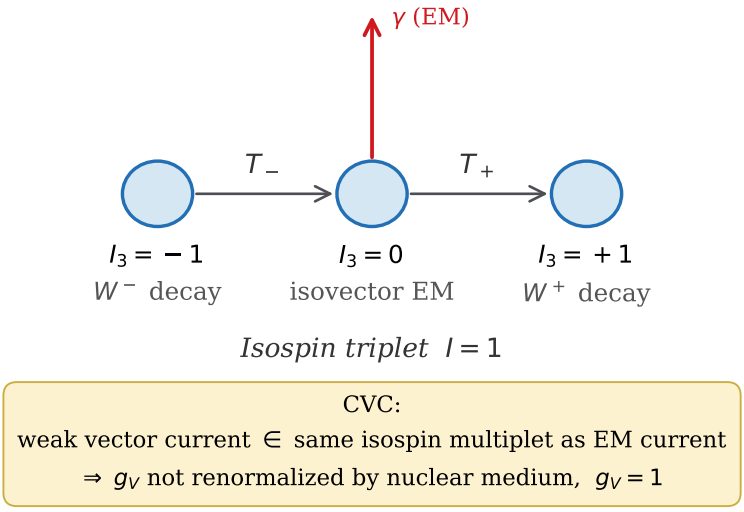


Figure 1.5 *The $I = 1$ isospin triplet underlying the CVC hypothesis. The three components correspond to W^- decay ($I_3 = -1$), the isovector electromagnetic current ($I_3 = 0$), and W^+ decay ($I_3 = +1$), connected by isospin raising and lowering operators $T_\pm$. Because the $I_3 = 0$ member is the conserved electromagnetic current, the whole triplet is conserved. The weak vector current, as an $I_3 = \pm 1$ member of the same multiplet, is therefore also conserved, and $g_V = 1$ exactly in any nuclear environment.*

contribution to the shape factor is therefore strongly constrained, reducing the model dependence of the spectrum largely to the axial NMEs and g_A^{eff}. Its application to forbidden non-unique transitions is less well tested experimentally. For superallowed Fermi decays ($0^+ \to 0^+$ transitions between isobaric analogue states), CVC is confirmed at the 10^{-4} level [25]. The present measurement probes CVC in a qualitatively different regime.

No analogous conservation law protects the axial current. The observed quenching of g_A relative to its free-nucleon value traces back to chiral symmetry. If all quark masses were zero, the strong interaction would be symmetric under independent rotations of left- and right-handed quarks, and both the vector and axial currents would be conserved. Chiral symmetry is spontaneously broken, however, and the pion emerges as the associated pseudo-Goldstone boson. Because the pion has a nonzero mass, the axial current is only partially conserved: $\partial_\mu J_A^\mu \propto m_\pi^2 \, \pi(x)$. This is the PCAC relation (partially conserved axial current). The nuclear medium can therefore modify g_A in a way it cannot modify g_V.

In a nucleus, nucleons exchange pions continuously. The weak axial current can couple to these pion-exchange processes: rather than acting on a single nucleon, the current acts on a correlated nucleon-pion system. Single-nucleon calculations miss these two-body contributions entirely and compensate by using a reduced g_A^{eff}. There is no analogous mechanism for the vector current. The pion is a pseudoscalar and couples to the axial current, not the vector current, so two-body corrections to g_V are forbidden at leading order by the same symmetry that protects it.

The prediction $g_V = 1$ is consistent with all superallowed Fermi data at the level of a few parts per thousand [25]. Small violations can arise from isospin-breaking corrections and, in principle, from non-nucleonic contributions to the vector current [22]. For the kinematic regime of ^{214}Bi ($Q = 3.27\,\text{MeV}$, $Z = 84$), such effects are expected to be small but not negligible for a precision measurement. A spectral shape consistent with the CVC prediction tests the symmetry in a new regime. Deviations would point to either isospin-breaking in the vector current or a failure of the nuclear model that happens to mimic CVC breaking.

1.5 Nuclear structure models

Computing the NMEs entering $C(E_e)$ requires solving the nuclear many-body problem for ^{214}Bi and ^{214}Po. With $Z = 83$ and $A = 214$, the full Hilbert space of antisymmetric A-nucleon states is vastly beyond any exact treatment. Every practical approach makes approximations. The two approaches used for computing NMEs in the ^{214}Bi decay are the nuclear shell model and the proton-neutron quasiparticle random-phase approximation (pnQRPA). Ab initio methods are discussed for completeness but are not yet applicable at this mass.

Nuclear shell model

In the shell model, nucleons fill discrete orbits defined by a mean-field potential [20, 27]. The Hilbert space is truncated to a *valence space*, the subset of orbits nearest the Fermi surface. A residual two-body interaction is then diagonalised exactly within this space.

For $A \sim 214$, the natural choice of inert core is ^{208}Pb. It is doubly magic: both the proton shell at $Z = 82$ and the neutron shell at $N = 126$ are completely filled. This gives ^{208}Pb exceptional stability and pushes its lowest excitations to high energy. It can therefore be treated as frozen, and only the nucleons above it need to be treated explicitly. The valence space then comprises the proton orbits above $Z = 82$ and the neutron orbits above $N = 126$. The truncation reduces the problem to a manageable size: ^{214}Bi has one

valence proton and five valence neutrons above the core, while ^{214}Po has two valence protons and four valence neutrons. Distributing six particles across a moderate set of single-particle orbits produces a Hamiltonian matrix of tractable dimension, rather than a 214-body Schrödinger equation.

The calculation requires two types of input. Single-particle energies specify the energy of each valence orbit relative to the ^{208}Pb core. They are taken from the experimental level spectra of ^{209}Bi and ^{209}Pb, the nuclei with one nucleon outside the closed shell. Two-body matrix elements encode the residual interaction between any pair of valence nucleons. These are provided by the Kuo-Herling effective interaction (khpe) [28], named after T. T. S. Kuo and G. H. Herling who originally derived it from the free nucleon-nucleon potential. It is fitted to spectroscopic data in the $A \approx 208$–220 region, starting from the free nucleon-nucleon potential and including corrections from configurations outside the valence space.

Diagonalising the Hamiltonian matrix yields the nuclear eigenstates: the ground state and low-lying excitations of ^{214}Bi and ^{214}Po. The calculation is validated by comparing the predicted level scheme to the experimental spectrum before any NMEs are computed. Calculations using KSHELL [29] and NuShellX@MSU [30] reproduce the low-lying levels of both nuclei to within tens of keV [24]. This is a prerequisite for trusting the NMEs: if the model cannot reproduce the observed energy levels, the eigenstates are wrong, and matrix elements computed from them have no quantitative meaning. The same eigenstates then enter the NME computation. The ^{214}Bi state plays the role of $|J_i, \pi_i\rangle$ and the ^{214}Po state plays $|J_f, \pi_f\rangle$ in Eq. (1.4). The operators $\hat{O}_k$ in that expression are not produced by the shell model: they are fixed by the structure of the weak interaction, as described in Section 1.3. The shell model contributes only the nuclear wavefunctions. The resulting NMEs are inserted into Eq. (1.5), giving a predicted shape factor $C(E_e)$ that enters the spectral formula, Eq. (1.1).

The shell model conserves all relevant quantum numbers and has access to the full set of transition operators required for the forbidden shape factor. Its main limitation is valence-space truncation: excitations of the ^{208}Pb core must be absorbed into effective operators rather than computed explicitly, introducing an approximation [31]. Figure 1.6 shows the valence space schematically. The resulting level predictions are compared to experiment in Figure 1.7, for both nuclear systems studied in this thesis: ^{214}Bi/^{214}Po and ^{136}Xe/^{136}Cs. In both cases the shell model reproduces the general ordering and energy scale of the low-lying levels, though the density of predicted states exceeds the experimental count.

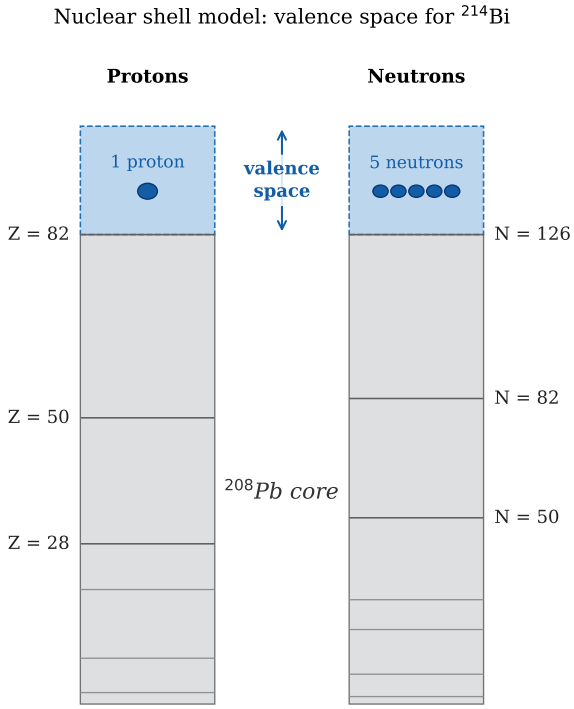


Figure 1.6 *Schematic nuclear shell structure for the $A \approx 214$ region. Filled shells (grey) up to the ^{208}Pb doubly-magic core ($Z = 82$, $N = 126$) are treated as an inert core in the shell-model calculation. The valence space (blue) spans the proton orbits above $Z = 82$ and the neutron orbits above $N = 126$. For ^{214}Bi, one valence proton and five valence neutrons occupy these orbits above the core. The Kuo-Herling (khpe) interaction provides the residual interaction within this space. Energy levels are schematic and not to scale.*

Proton-neutron QRPA

The proton-neutron quasiparticle random-phase approximation (pnQRPA) starts from a Bardeen-Cooper-Schrieffer (BCS) pairing vacuum [33]: a ground state in which proton-proton and neutron-neutron pairs form correlated $J = 0$ states analogous to Cooper pairs in a superconductor. Correlated proton-neutron excitations are built on top of this vacuum [27], with the Hilbert space restricted to two-quasiparticle configurations. This keeps the computational cost far lower than the full shell model and makes the approach applicable to any mass region. The pnQRPA has been applied in systematic surveys of g_A quenching across many isotopes [34, 35] and in calculations of forbidden beta spectra for background modelling in rare-event

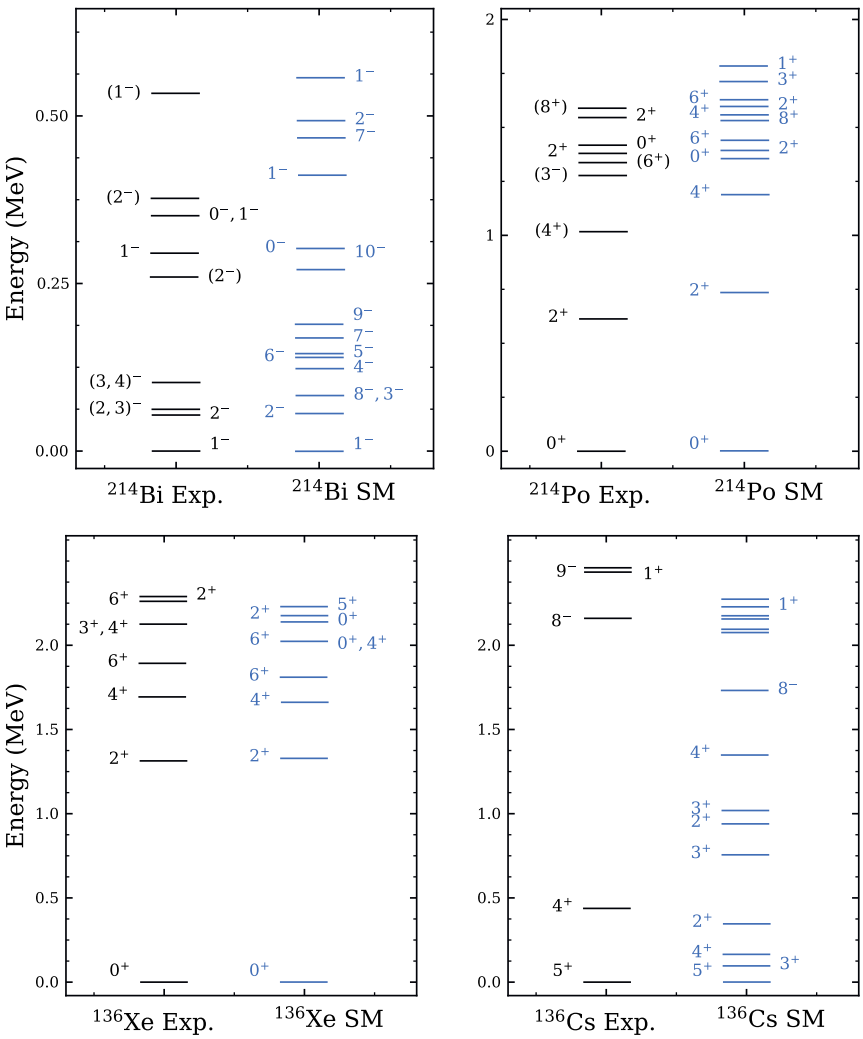


Figure 1.7 *Shell-model predictions (blue) compared to experimental energy levels (black) for the nuclear systems studied in this thesis. Top:* ^{214}*Bi (left,* $J^\pi = 1^-$ *ground state) and* ^{214}*Po (right,* $J^\pi = 0^+$*), the initial and final nuclei of the beta decay measured in Chapter 5. Bottom:* ^{136}*Xe (left,* 0^+*) and* ^{136}*Cs (right,* 5^+*), the initial and final nuclei for fermionic dark matter absorption (*$\chi + {}^{136}Xe \rightarrow e^- + {}^{136}Cs$*) studied in Chapter 6. Top reproduced from Ramalho and Suhonen [24]; bottom reproduced from Haselschwardt et al. [32].*

searches [36]. The effective g_A values extracted from such fits depend on g_{pp} and other model inputs, making the quenching itself model-dependent. The principal limitation is the restriction to two-quasiparticle states. Results are also sensitive to the particle-particle interaction strength g_{pp}: near the critical value where the pnQRPA ground state becomes unstable, NMEs can change rapidly and require careful tuning [27].

Ab initio methods

Methods such as coupled-cluster theory and the in-medium similarity renormalisation group now reach $A \sim 100$ starting from realistic nucleon-nucleon and three-nucleon interactions. Their uncertainties can in principle be systematically controlled. They are not computationally feasible for $A \sim 214$ at present, and none of the spectral predictions compared in Chapter 5 relies on an ab initio input. Their eventual extension to this mass region would provide a valuable cross-check on the shell-model and QRPA results.

Table 1.2 compares the three approaches.

Table 1.2 *Comparison of nuclear-structure approaches for computing NMEs relevant to* ^{214}*Bi. Ab initio methods are listed for completeness; they are not computationally feasible at* $A \sim 214$ *at present.*

Method	Applicable A	Key approximation
Shell model	up to ~ 230	valence-space truncation
pnQRPA	any	2-quasiparticle space
Ab initio	up to ~ 100	approx. many-body solver

The g_A quenching problem

The free-nucleon axial coupling constant is $g_A^{\text{free}} = 1.2756 \pm 0.0013$ [37]. In nuclear-structure calculations, the effective value g_A^{eff} is systematically smaller. Fitting transition data across many nuclei yields values in the range 0.7 to 1.0, depending on the nucleus, the model space, and the approximation used [21, 31, 38]. Ab initio calculations confirm that two-body electroweak currents account for essentially all the apparent quenching in light and medium-mass nuclei [39]. Whether this picture extends to the $A \sim 214$ region, where ab initio methods are not yet applicable, remains an open question.

For forbidden transitions, g_A enters the shape factor quadratically through C_A and linearly through C_{VA}. A 20% reduction in g_A relative to its free value propagates to a $\sim$ 36% shift in the axial contribution to $C(E_e)$. Different prescriptions for g_A^{eff} are one of the principal sources of spread among the theoretical spectral predictions for ^{214}Bi.

1.6 The ^{214}Bi Decay as a Case Study

The ground state of ^{214}Bi has $J^\pi = 1^-$ and the ground state of ^{214}Po has $J^\pi = 0^+$ [40]. The transition has $\Delta J = 1$ and a parity change, making it first-forbidden non-unique in the classification of Section 1.2. For this $1^- \to 0^+$ transition, $\Delta J = 1$ lies below the maximum $\Delta J = n + 1 = 2$. Both the rank-0 and rank-1 axial-vector operators are therefore compatible with the angular momentum selection rules and contribute to $C(E_e)$ (Section 1.2). Each couples to an independent NME. The vector operator contributes a further term, constrained by CVC but not kinematically fixed. Several independent NMEs are therefore needed to fully specify $C(E_e)$, and they cannot be determined from branching ratios alone. The explicit polynomial structure of $C(E_e)$ for this transition, with all contributing operators identified, is given in Eq. (1.5) and the surrounding discussion of Section 1.3.

Model predictions

The shell-model description of the ^{214}Bi and ^{214}Po nuclear structure is shown in Figure 1.7 (top). The variation among model predictions for the spectral shape is illustrated schematically in Figure 1.4. The specific shell-model spectral predictions are compared to the XENONnT data in Chapter 5. A key source of the spread is the sNMEs introduced in Section 1.3. The shell model computes them in principle, but because they are relativistic corrections their values carry large relative uncertainties. Ramalho and Suhonen [24] therefore treated them as free parameters and fitted them simultaneously to reproduce the experimental branching ratios for all transitions in the ^{214}Bi decay scheme. The fit has a quadratic dependence on the sNMEs, producing two distinct solutions (sNME1 and sNME2).

The degeneracy arises because branching ratios depend on integrated rates, which are proportional to $|M|^2$, the squared amplitude. Two sNME combinations with opposite signs give the same $|M|^2$ and therefore the same branching ratios. The shape factor $C(E_e)$, however, contains interference terms between different NMEs, with energy dependences proportional to p_e^2, p_ν^2, and $p_e p_\nu$. These interference terms change sign between the two

solutions, producing different spectral shapes even when the total rates are identical. Branching ratios alone cannot distinguish the solutions. The energy-resolved spectrum can.

As a baseline, the "Allowed" approximation simply sets $C(E_e) = \text{const}$, as if the transition were an allowed decay and the forbidden corrections were absent entirely. It is not a physical model of ^{214}Bi decay, but a benchmark that makes the size of the shape-factor effect visible. This prediction and the CVC prediction differ by up to 20% in the mid-spectrum region, a difference large enough to be resolved experimentally but requiring careful control of detector response and background.

Q-value, decay scheme, and precision corrections

The ground-state Q-value of 3.27 MeV places approximately 80 excited states of ^{214}Po within the kinematically accessible range [40]. Most of the ^{214}Bi decay rate populates these excited branches through the associated γ cascade. The ground-state-to-ground-state transition carries approximately 19% of the total rate [40]. Isolating this component requires either a γ coincidence or a careful statistical subtraction, as detailed in Chapter 5. On the experimental side, the large Q-value is an advantage: the momentum-dependent terms in $C(E_e)$ grow with p_e, so the spectral distortions are larger and easier to measure at higher energies than for lower-Q nuclei.

Precision comparison of the measured spectrum requires several corrections beyond Eq. (1.1). The nuclear charge is distributed over a finite volume, modifying the Fermi function relative to the point-Coulomb approximation. The atomic electrons screen the nuclear charge, suppressing the Fermi function near threshold. QED radiative corrections contribute at the 0.5–1% level and are concentrated near the endpoint. The induced nucleon form factors, in particular weak magnetism (constrained by CVC) and the mesonic enhancement of the weak axial charge [41], shift spectral predictions by several percent at forbidden order. All four corrections are implemented in `BetaShape` [42], which generates the theoretical spectra compared to data in Chapter 5. A comprehensive treatment is given in Hayen et al. [9].

The $^{214}\text{Bi} \rightarrow {}^{214}\text{Po}$ ground-state transition has a large Q-value and a first-forbidden non-unique $1^- \rightarrow 0^+$ spin-parity assignment, with several competing nuclear-model predictions. CVC provides the only model-independent constraint on the vector form factors. The axial sector remains the dominant source of theoretical uncertainty. Whether the measured spectral shape is consistent with CVC, and whether the shell-model description of the axial NMEs is correct, is the question addressed experimentally in Chapter 5.

2

THE XENONNT DETECTOR

The XENONnT experiment [43] is a rare-event search detector located at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy, beneath approximately 3600 meter-water-equivalent of rock overburden. The underground location reduces the cosmic muon flux by approximately six orders of magnitude [44], providing a low-radiation environment essential for rare-event searches. The detector consists of a dual-phase liquid xenon time projection chamber (TPC) surrounded by a gadolinium-loaded water Cherenkov neutron veto (NV) and an outer muon veto (MV), as shown in Figure 2.1. XENONnT is designed primarily for the direct detection of dark matter in the form of weakly interacting massive particles (WIMPs). The target signal is the nuclear recoil produced when a WIMP scatters off a xenon nucleus, depositing energy in the keV range. Despite decades of experimental effort, no conclusive detection has been made, and the search continues with ever-larger and ever-cleaner detectors [45, 46].

Liquid xenon offers a particularly favourable detection medium for this purpose [47, 48]. Its high atomic mass enhances sensitivity to WIMP–nucleus scattering, while its high density allows compact detector geometries with large target masses. Interactions in liquid xenon produce both scintillation photons and ionisation electrons. Nuclear recoils (NR) arise when an incident particle, such as a WIMP or a neutron, scatters elastically off a xenon nucleus. Electronic recoils (ER) arise when energy is deposited primarily via interactions with atomic electrons, as in gamma-ray Compton scattering or beta decay. The relative balance between scintillation and ionisation differs systematically between these two classes of interaction. This property provides the basis for discriminating potential WIMP signals from the dominant ER backgrounds, and underlies the dual-phase TPC detection concept described in the following sections.

XENONnT is the latest in a series of increasingly sensitive liquid xenon experiments operated by the XENON collaboration at LNGS, following XENON10, XENON100, and XENON1T [49–51]. Each generation brought a significant increase in target mass and a reduction in background rate. XENONnT contains approximately 5.9 tonnes of liquid xenon in its active volume, roughly three times that of its predecessor XENON1T.

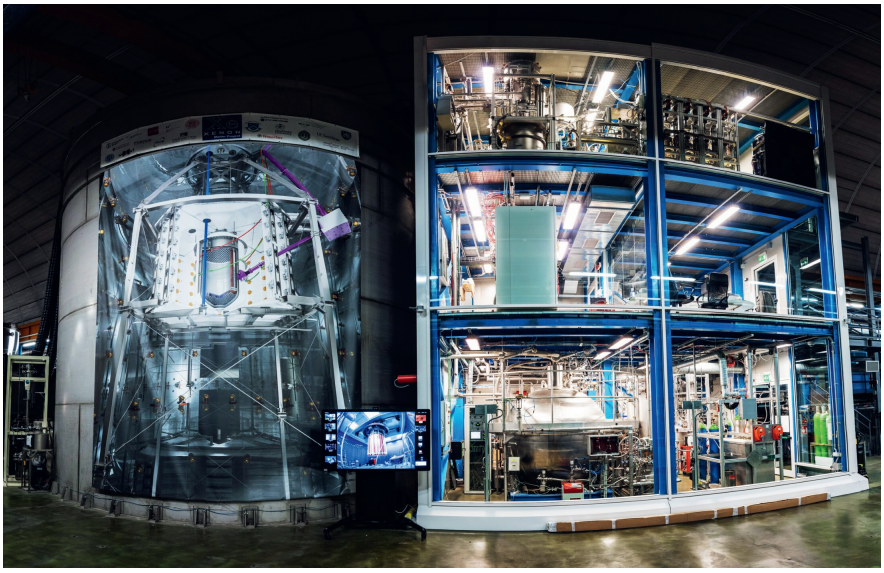


Figure 2.1 *Photo of the XENONnT detector in the LNGS underground laboratory. On the left stands the large water tank of the muon veto, which encloses the cryostat and the neutron veto. On the right, the three-storey service building houses the cryogenic systems, the xenon storage and recovery system, the liquid and gaseous xenon purification systems, and the online radon removal system. The data acquisition (DAQ) system is located on the first floor of the service building.*

While the primary physics goal of XENONnT is the detection of WIMP-induced nuclear recoils of a few keV, the same detector properties also enable a broader class of rare-event searches. At higher energies, extending into the few-MeV range, the large target mass and ultra-low background environment enable precision measurements of nuclear processes and searches for new physics involving electronic recoils. Processes such as the beta-decay of ^{214}Bi from radon progeny [1], double beta-decay of ^{136}Xe [5, 52], and potential signals from beyond-Standard-Model particles [53] all become accessible at these energies.

At higher energies the interaction topologies become increasingly complex, with multi-site energy depositions that challenge the standard signal modelling developed for low-energy searches. Understanding and modelling the detector response across this energy range is therefore both a scientific goal in itself and a prerequisite for the physics analyses it enables. The work presented in this thesis focuses on the higher-energy electronic recoil sector.

Data taking is organised in science runs: Science Runs 0, 1, and 2 (SR0, SR1, SR2) were conducted between 2021 and 2025 under similar detector

conditions, accumulating a total WIMP-search live time of 631.7 days. After completion of these science runs, XENONnT underwent a final upgrade implementing a new electrode configuration to increase the drift field and improve discrimination between electronic and nuclear recoils.

2.1 Energy Deposition in Liquid Xenon

When a particle traverses liquid xenon, it deposits energy through excitation and ionisation of xenon atoms along its track. Excited atoms (Xe^*) and electron–ion pairs ($Xe^+ + e^-$) are produced in proportions that depend on the type and energy of the incident particle. A fraction of the deposited energy is dissipated as heat and does not contribute to observable signals in the detector.

Excited xenon atoms rapidly associate with neighbouring ground-state atoms to form excited dimers, or excimers (Xe_2^*). These excimers are bound in the excited state but repulsive in the ground state, and decay by emitting a vacuum ultraviolet (VUV) photon at 175 nm and dissociating back into two ground-state xenon atoms, as illustrated schematically in Figure 2.2. Since ground-state atoms cannot reabsorb photons at this wavelength, liquid xenon is transparent to its own scintillation light, and emitted photons propagate freely to the detector walls.

Excimers are formed via two distinct molecular states, the singlet and triplet, which differ in their spin configuration and characteristic decay times. The singlet state decays promptly with a time constant of approximately 3 ns, while the triplet state is longer-lived with a decay time of approximately 24 ns [54]. Both states emit photons at the same wavelength, so the total scintillation pulse is a superposition of these two components. The relative population of singlet and triplet states depends on the local ionisation density along the track, which differs between electronic and nuclear recoils. In liquid xenon, the contrast between the singlet-to-triplet ratios for ER and NR is small, and the similarly short decay times of the two states make pulse shape discrimination less effective than in liquid argon [55]. Recent demonstrations in xenon detectors have nonetheless been reported [56].

Ionisation electrons produced along the track may either escape the interaction region or recombine with xenon ions. Recombination produces additional excited states, which subsequently form excimers and contribute to the scintillation signal, effectively converting ionisation charge back into light. Nuclear recoils deposit energy at higher ionisation density than electronic recoils of comparable energy. Combined with Lindhard quenching of the ionisation, this leads to stronger recombination and a larger scintillation-to-ionisation ratio [57]. This difference underlies the

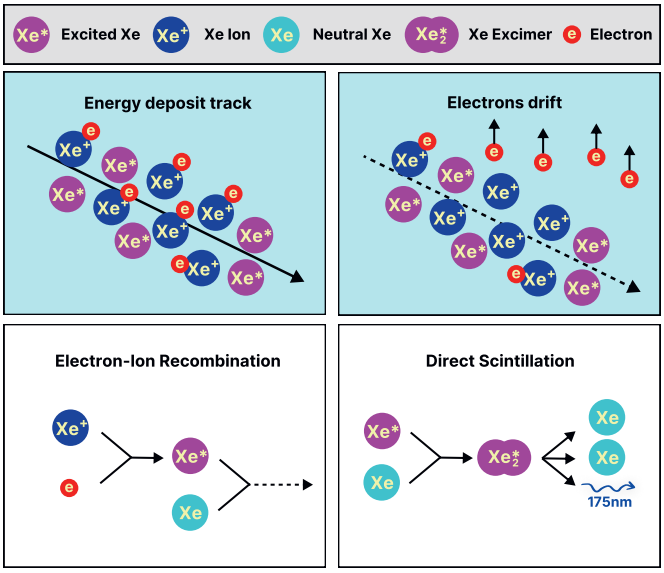


Figure 2.2 *Schematic illustration of microscopic energy deposition processes in liquid xenon. An incident particle produces excited xenon atoms (Xe^*) and ionised atoms (Xe^+) along its track. Excited atoms form excimers (Xe_2^*) that decay radiatively, emitting vacuum ultraviolet scintillation photons at 175 nm. Ionisation electrons may recombine with xenon ions, generating additional excited states, or remain free. The relative balance between excitation, ionisation, and recombination depends on the interaction type and local ionisation density.*

ER/NR discrimination exploited by dual-phase TPCs, as discussed further in Section 2.2 and Chapter 4.

2.2 Detection Principle of a Dual-Phase Xenon TPC

A dual-phase xenon TPC consists of a volume of liquid xenon with a layer of gaseous xenon maintained above the liquid surface. Particle interactions in the liquid produce both prompt scintillation light and ionisation electrons. A cathode at the bottom of the liquid volume and a gate electrode near the liquid–gas interface define the drift region. A uniform electric field in this region suppresses electron–ion recombination and drifts free electrons upward. Above the liquid surface, a stronger extraction field between the gate and the anode extracts the electrons into a thin gas layer, where they produce secondary scintillation.

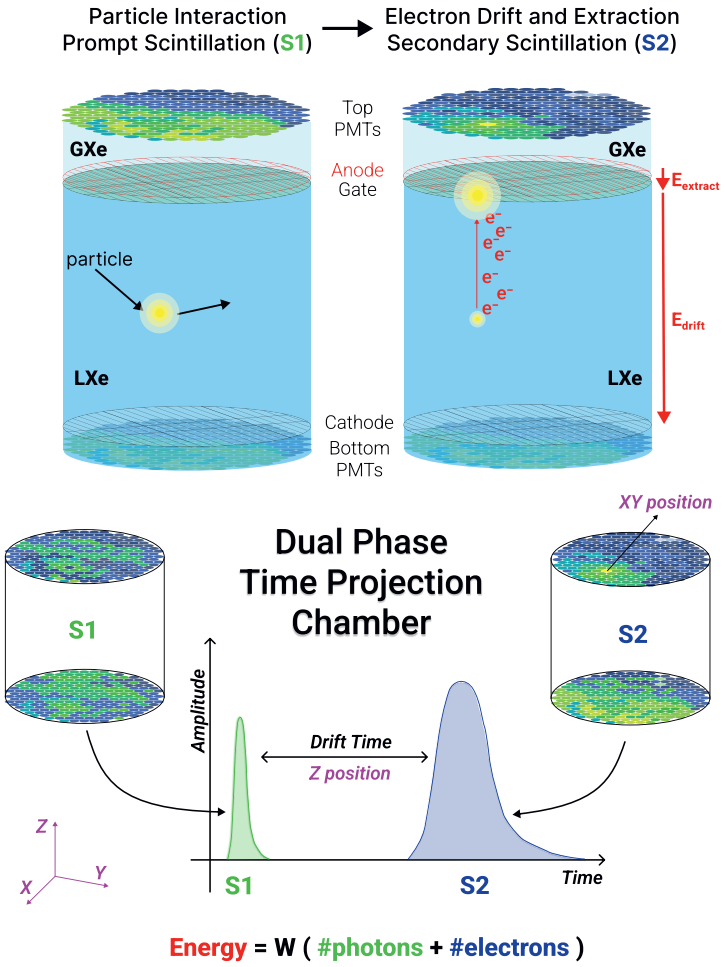


Figure 2.3 *Operating principle, signal structure, and reconstruction in the XENONnT dual-phase TPC. Top left: a particle interaction produces prompt scintillation light (S1), detected by both the top and bottom PMT arrays. Top right: ionisation electrons drift upward under the applied field, are extracted into the gas phase, and produce proportional electroluminescence (S2), detected predominantly by the top array. Bottom: example S1 and S2 waveforms with the corresponding PMT hit patterns on the top and bottom arrays. The time difference between S1 and S2 equals the electron drift time, giving the z position of the interaction. The S2 light distribution on the top array gives the horizontal (x–y) position (coordinate axes shown in purple, bottom left). The S1 and S2 signal areas determine the numbers of produced photons n_γ and ionisation electrons n_e, from which the deposited energy is reconstructed as $E = W(n_\gamma + n_e)$.*

The prompt scintillation signal, referred to as S1, is produced by the decay of xenon excimers as described in Section 2.1. S1 photons are emitted isotropically from the interaction site and detected simultaneously by both the top and bottom PMT arrays. The S1 pulse is narrow, typically extending over tens of nanoseconds, and provides a precise time reference for the interaction.

Ionisation electrons that escape recombination drift upward toward the liquid surface. Upon reaching the liquid–gas interface, they are extracted into the gas by the extraction field. In the gas, the electrons are accelerated and excite xenon atoms, producing secondary scintillation photons proportional to the number of extracted electrons. This proportionality makes S2 a direct measure of the ionisation charge deposited in the interaction. Because S2 is produced at the top of the detector, its photons are detected predominantly by the top PMT array. The S2 pulse is much broader than S1, extending over several microseconds, due to longitudinal diffusion during drift and the finite duration of the electroluminescence process.

The electrons drift at a velocity set by the drift field, typically of order 1 mm/μs in liquid xenon. The time elapsed between S1 and S2 therefore corresponds to the electron drift time and determines the depth (z coordinate) of the interaction. The distribution of S2 light across the top PMT array encodes the horizontal (x–y) position. Together, these measurements provide full three-dimensional event reconstruction. The S1 and S2 signal areas determine the numbers of produced photons and electrons, from which the deposited energy is reconstructed. The full signal chain is summarised in Figure 2.3.

The ratio of S2 to S1 encodes the relative partition between ionisation and scintillation, which differs systematically between ER and NR, and provides the primary handle for background discrimination.

2.3 Detector Layout

The XENONnT detector consists of a central TPC nested inside two concentric veto detectors, all housed within a water tank at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy. The overall detector layout is shown in Figure 2.4.

Surrounding the cryostat is the NV [58], a water Cherenkov detector that identifies neutrons which have scattered in the TPC and subsequently thermalized and captured in the surrounding water. Neutrons captured on gadolinium release a cascade of gamma-rays totaling approximately 8 MeV, detected as Cherenkov light by the NV PMTs. Coincident detection of this signal allows neutron-induced nuclear recoil events to be rejected as backgrounds. The outermost detector is the MV [59], a water Cherenkov

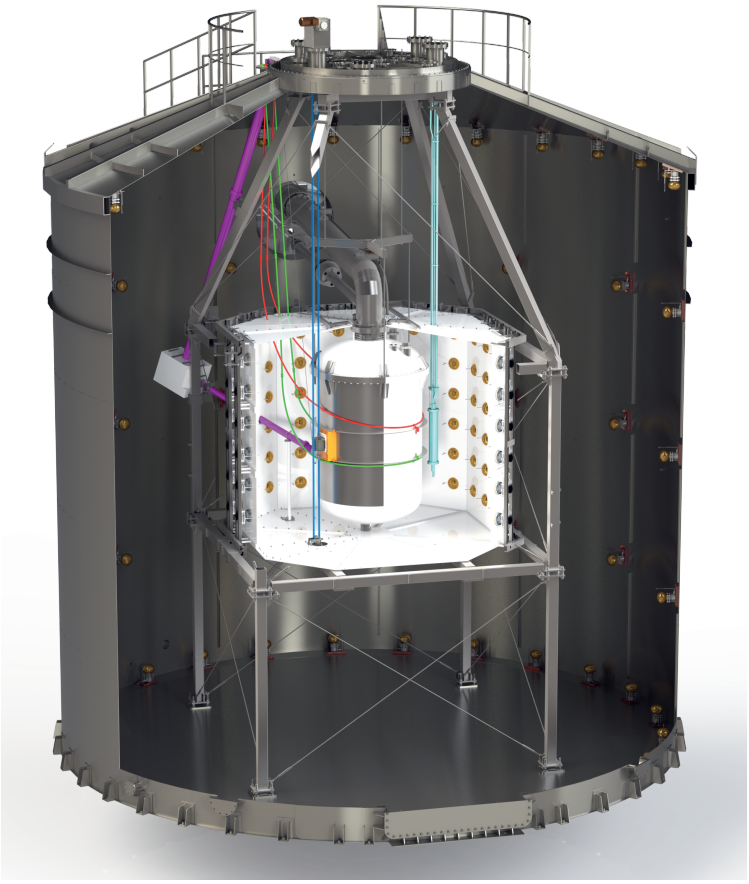


Figure 2.4 *CAD rendering of the XENONnT detector. The central TPC cryostat is surrounded by the neutron veto (NV) and the outermost muon veto (MV) water tank. Figure from Ref. [43].*

detector that tags residual cosmic muons and provides passive shielding against neutrons produced by cosmic-ray interactions in the surrounding rock.

The XENONnT TPC [43] is a cylindrical active volume 1.61 m high and 1.33 m in diameter, with a drift length of 1.49 m, enclosed in a cryostat containing approximately 5.9 tonnes of liquid xenon. The inner walls are lined with highly reflective polytetrafluoroethylene (PTFE) panels to maximise light collection. The TPC is instrumented with 494 Hamamatsu R11410-21 PMTs [43], low-background devices optimised for VUV sensitivity [60], arranged in a top array of 253 PMTs and a bottom array of 241 PMTs. The



Figure 2.5 *Left: CAD rendering of the XENONnT cryostat and TPC, showing the top and bottom PMT arrays, the field-shaping rings, and the PTFE reflector panels. Right: photograph of the inner cryostat during assembly.*

bottom array collects a larger fraction of the isotropically emitted S1 light, while the top array is primarily sensitive to the S2 signal produced in the gas phase. The field cage, consisting of a series of field-shaping rings, ensures uniformity of the drift field across the active volume [61]. The TPC layout is shown in Figure 2.5.

The field configuration described in Section 2.2 is established by three wire-grid electrodes: a cathode, a gate, and an anode. Taking the gate as the reference plane ($z = 0\,\mathrm{mm}$), the anode is located at $z = +8\,\mathrm{mm}$ and the cathode at $z = -1486\,\mathrm{mm}$, which defines the 1.49 m drift region. Two additional screening electrodes, placed 55 mm below the cathode and 28 mm below the top PMT array, shield the PMTs from the high drift and extraction fields. The gate and anode use stainless steel wires of 216 μm diameter with a pitch of 5 mm. The cathode uses thicker wires of 304 μm diameter with a 7.5 mm pitch [43]. To prevent sagging under electrostatic forces, two transverse wires are installed on the gate and four on the anode, oriented perpendicular to the main wire planes. These introduce local field inhomogeneities near the electrode planes that affect S2 signal formation and position reconstruction. Their treatment in the detector modelling is described in Section 3.3. The wire grid construction is shown in Figure 2.6.

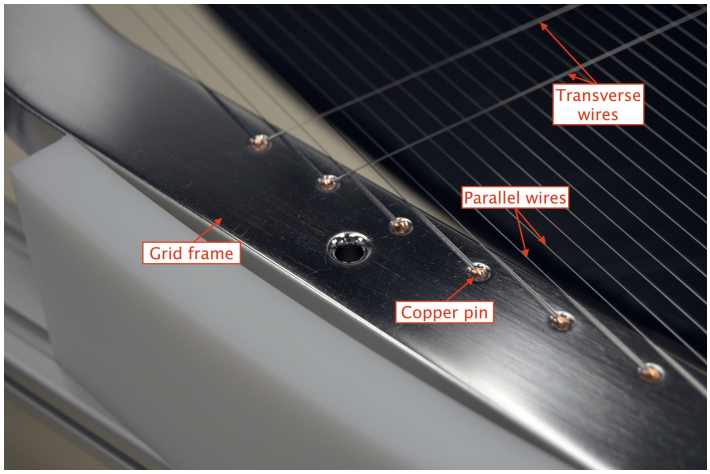


Figure 2.6 *Close-up photograph of the anode electrode during assembly. The parallel signal wires (216 µm diameter, 5 mm pitch) are fixed individually to the stainless steel frame via copper pins. Four transverse wires, installed perpendicular to the main wire planes to prevent sagging under electrostatic load, are visible crossing the array. The gate electrode uses the same wire dimensions but carries only two transverse wires. Figure from Ref. [43].*

The design drift field for XENONnT was approximately 200 V/cm. In practice, a short circuit between the bottom screening electrode and the cathode limited the cathode voltage. The resulting operating drift field was only approximately 23 V/cm throughout all three science runs. This limitation motivated the electrode upgrade carried out after SR2, described at the end of this chapter.

The TPC itself requires extremely high xenon purity for efficient charge transport. The xenon is continuously circulated through gas- and liquid-phase purification systems to remove electronegative impurities such as oxygen [62], achieving electron lifetimes exceeding 10 ms throughout SR0. The maximum electron drift time is approximately 2.2 ms across the full drift length. Stable operating conditions are maintained at approximately 175 K (-98 °C) and 1.9 bara by cryogenic systems [43].

Radioactive backgrounds from dissolved contaminants are controlled through several complementary strategies. Krypton is removed via cryogenic distillation exploiting the difference in vapor pressures between xenon and krypton [63], reducing the $^{\mathrm{nat}}$Kr/Xe concentration to the part-per-quadrillion level and suppressing the beta-emitting ^{85}Kr background. Radon, which emanates continuously from detector materials, is mitigated through careful material selection during construction [64] and an online radon removal system based on cryogenic distillation of the gas-phase xenon [2]. Progressive improvements to the radon removal strategy across

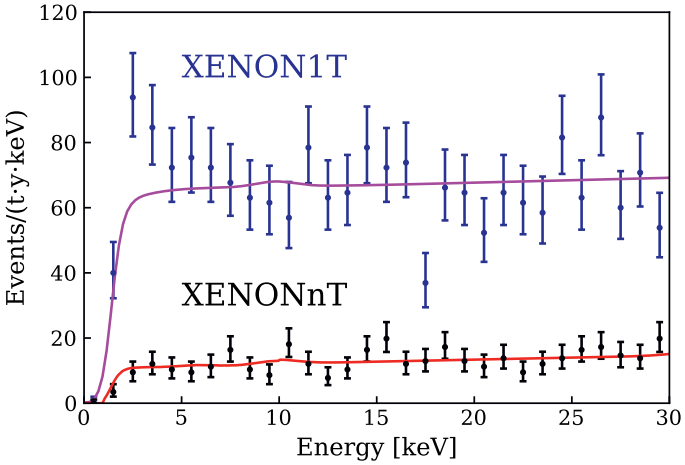


Figure 2.7 *Electronic recoil background rate as a function of energy in XENON1T and XENONnT. The reduction reflects the lower radon activity achieved through progressive improvements to the radon removal system [2], establishing XENONnT as the lowest-background liquid xenon detector to date, reaching a radon activity lower than 1 μBq/kg.*

the science runs ultimately achieved radon activity levels below 1 μBq/kg. The resulting reduction in the electronic recoil background rate, compared to XENON1T, is shown in Figure 2.7.

2.4 Data Acquisition and Signal Reconstruction

The XENONnT detector operates with a fully triggerless data acquisition (DAQ) system [65]. PMT signals are continuously digitized and stored as waveform segments exceeding a predefined threshold, using zero-length encoding to suppress baseline samples and reduce data volume. This approach avoids biases introduced by hardware triggers and allows event definitions to be applied entirely in software.

Raw waveform data are processed using the `strax` framework and its XENONnT-specific extension `straxen` [66]. The processing chain applies baseline subtraction, pulse integration, and time alignment to produce calibrated waveform segments. A peak-finding algorithm then groups coincident PMT signals into peaks. Peaks are classified as S1-like or S2-like based on their width, area, and temporal structure: S1 peaks are narrow (tens of nanoseconds) and distributed across both PMT arrays, while S2 peaks are broad (several microseconds) and concentrated on the top array. Events are constructed by pairing S1 and S2 candidates according to temporal and spatial criteria.

Position reconstruction uses the drift time to determine the z coordinate and the S2 light pattern on the top PMT array to reconstruct the horizontal position. The three-dimensional position reconstruction enables the definition of a fiducial volume, selecting only interactions in the inner part of the TPC. This reduces backgrounds from radioactive contaminants on the detector walls and from radon daughters plating out on the cathode. In WIMP-search analyses, a fiducial mass of approximately 4 tonnes is defined [45].

The raw S1 and S2 signals are subject to a series of position- and time-dependent corrections to ensure a uniform detector response across the active volume and over the duration of each science run. The S1 signal is corrected for spatial variations in light collection efficiency and drift field inhomogeneities using a three-dimensional correction map derived from ^{83m}Kr calibration data. The isotope ^{83m}Kr is a short-lived radioactive tracer injected into the active volume, decaying uniformly throughout and providing a high-statistics sample of monoenergetic deposits at all positions in the TPC. The S2 signal is subject to three corrections: for electron losses to electronegative impurities during drift, for spatial non-uniformities in the extraction and electroluminescence regions via a two-dimensional correction map, and for time-dependent variations in the single-electron gain and extraction efficiency. The resulting corrected observables are denoted $cS1$ and $cS2$ [66]. The same reconstruction and correction chain is applied to both experimental and simulated data, ensuring a consistent comparison throughout this thesis.

2.5 Energy Scale and Detector Gains

The deposited energy is reconstructed from the corrected light and charge signals as

$$E = W \left(\frac{cS1}{g_1} + \frac{cS2}{g_2} \right), \tag{2.1}$$

where $W \approx 13.7$ eV is the average energy required to produce one quantum [67] (see Chapter 4). The gain parameters g_1 and g_2 convert the detected signals into the number of produced photons and electrons, respectively: g_1 is the average number of detected photoelectrons per produced scintillation photon, and g_2 is the average number of detected photoelectrons per ionisation electron produced in the interaction.

The combined energy scale exploits the anti-correlation between scintillation and ionisation arising from recombination fluctuations. When recombination is high, more photons and fewer electrons are produced, while the sum $cS1/g_1 + cS2/g_2$ remains approximately constant. This anti-

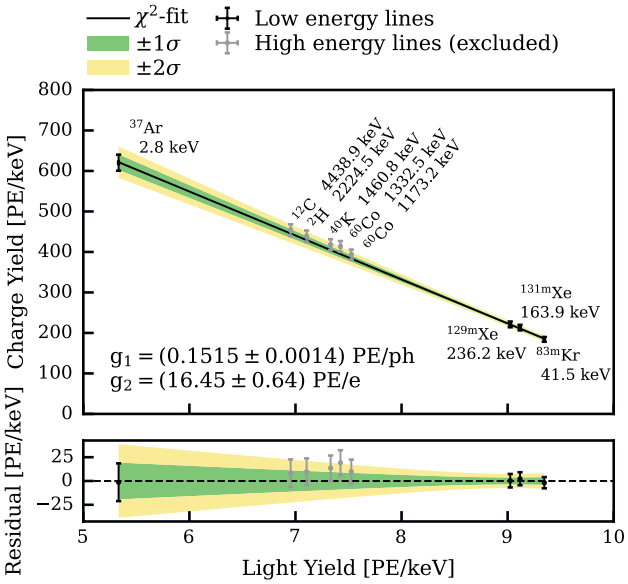


Figure 2.8 *Doke plot used to determine the detector gain parameters g_1 and g_2. Each point corresponds to the mean cS1 and cS2 signals measured for a calibration source or background line at a known energy. The linear relation of Eq.* (2.3) *is fitted simultaneously to all points, yielding g_1 and g_2. The calibration sources used include ^{37}Ar, ^{83m}Kr, ^{129m}Xe and ^{131m}Xe, ^{60}Co, ^{40}K, ^{2}H and ^{12}C. Figure from Ref. [66].*

correlation significantly improves the energy resolution compared to using either signal alone, and is discussed in detail in Chapter 4.

The gain parameters g_1 and g_2 are not fixed detector constants but analysis-dependent quantities. Their values depend on which calibration sources are included in the fit, on the corrections applied to cS1 and cS2, and on the assumed value of W. Consequently, different analyses within the collaboration may obtain slightly different values of g_1 and g_2 while remaining consistent with the same underlying detector response. The calibration strategy exploits the linear relationship between the mean cS1 and cS2 signals at fixed deposited energy E,

$$\langle cS1 \rangle = g_1 \frac{E}{W} - g_1 \langle N_e \rangle, \qquad \langle cS2 \rangle = g_2 \langle N_e \rangle, \tag{2.2}$$

where $\langle N_e \rangle$ is the mean number of free electrons. Eliminating $\langle N_e \rangle$ yields

$$\langle cS1 \rangle = -\frac{g_1}{g_2} \langle cS2 \rangle + g_1 \frac{E}{W}, \tag{2.3}$$

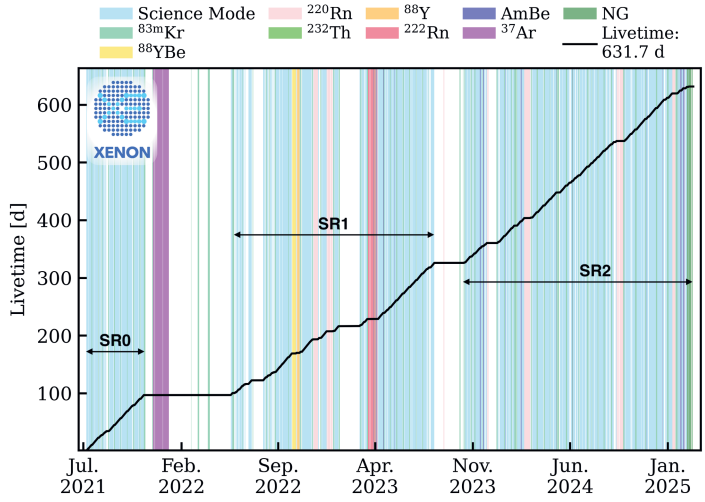


Figure 2.9 *Cumulative live time of the XENONnT science runs. Coloured bands indicate different data taking modes for the three science run periods (SR0, SR1, SR2) including calibration campaigns with different radioactive sources. The black line shows the cumulative WIMP-search live time, excluding calibration and non-standard operation periods. The total amounts to 631.7 days.*

which is linear in $\langle cS2 \rangle$ with slope $-g_1/g_2$ and intercept $g_1 E/W$. By measuring the mean cS1 and cS2 signals for multiple calibration sources at known energies, g_1 and g_2 can be extracted simultaneously from a combined fit. This approach, known as the Doke plot [66], is illustrated in Figure 2.8.

The resulting gain values are used throughout this thesis to convert corrected cS1 and cS2 signals into reconstructed energies via Eq. (2.1). The values of g_1 and g_2 are specific to each science run and analysis: they vary at the few-percent level across SR0, SR1, and SR2, reflecting changes in detector conditions and calibration choices. The numerical values used in the analyses presented here are quoted in Chapter 5.

2.6 Science Runs and Detector Conditions

XENONnT data taking is organised in science runs, denoted SR0, SR1, and SR2, separated by calibration campaigns. The total science live time used in WIMP-search analyses amounts to 631.7 days, distributed as 97.1 days (SR0), 224.6 days (SR1), and 310.0 days (SR2), as summarised in Table 2.1 and shown in Figure 2.9.

Table 2.1 *Summary of XENONnT science runs and corresponding live time used in WIMP-search analyses.*

Science Run	Period	Live time [days]
SR0	2021	97.1
SR1	2022–2023	224.6
SR2	2023–2025	310.0
Total	2021–2025	631.7

All three runs were conducted under broadly similar detector conditions. The drift field was maintained at approximately 23 V/cm throughout, and detector operating parameters such as electrode voltages, liquid level, pressure, temperature, and purification settings remained stable within small variations. The baseline detector response was therefore comparable across the three science runs.

The most significant evolution across the runs concerns the radioactive background conditions. The radon activity in the xenon was progressively reduced as the radon removal strategy matured. Early operation relied primarily on gas-phase radon reduction, while later running incorporated additional radon suppression in the liquid circulation loop, ultimately achieving the lowest radon concentrations ever recorded in a liquid xenon detector [2].

During SR1, an increasing rate of delayed electron emission was observed. Delayed electrons are attributed to several sources: electrons trapped on electronegative impurities and released at later times, photoionization of impurities by S1 and S2 photons, and field emission from detector materials. These contributions introduced additional low-level charge activity and affected the stability of the detector response. A purification filter installed during SR2 reduced the contamination and restored the detector stability.

The analyses in this thesis draw on different subsets of these data. Chapter 3 uses calibration data from SR2, the most recent science run, to establish the agreement between data and simulation and the detector modelling this requires. The analyses of the subsequent chapters, the ^{214}Bi beta-decay measurement in Chapter 5 and the fermionic dark matter search in Chapter 6, use data from SR0 and SR1, the datasets available at the time of those analyses.

2.7 summary

This chapter has introduced the XENONnT detector and the signal formation, reconstruction, and calibration framework that underpin the analyses in this thesis. The dual-phase TPC converts particle interactions into S1 scintillation and S2 electroluminescence signals, from which three-dimensional position and deposited energy are reconstructed. The operating drift field of 23 V/cm, constrained by a short circuit in the electrode system, characterises the detector response across all three science runs and motivated the post-SR2 upgrade. The gain parameters g_1 and g_2 determined from the Doke calibration enter directly into the energy reconstruction used throughout. The detector simulation and yield modelling are developed in Chapters 3 and 4.

3

DETECTOR SIMULATION FRAMEWORK

Interpreting XENONnT data requires a quantitative model of how particle interactions produce the S1 and S2 signals recorded by the detector. The `fuse` framework [68] provides this by simulating the full chain from Geant4 energy depositions [69–71] to digitised PMT waveforms, replacing the earlier `WFSim` [72] and `epix` [73] packages with a more modular implementation. The simulation proceeds through four sequential physical stages, summarised in Table 3.1 and illustrated in Figure 3.1. Particle transport through the detector volume is handled entirely by Geant4, which records energy depositions as a list of $\{\Delta E, x, y, z, t\}$ records together with the full particle ancestry. These deposits are the sole starting input to `fuse`; the two packages are otherwise independent.

A key design feature is that `fuse` produces output in the native `strax` data format used by the `straxen` [74] reconstruction pipeline. Simulated events therefore pass through exactly the same reconstruction chain as real data, and any efficiency or bias introduced by the reconstruction software affects simulation and data in a similar way.

Table 3.1 *The four physical stages of the `fuse` simulation chain for XENONnT. The microphysical model is described in Chapter 4; all remaining stages are detailed in this chapter.*

Stage	Physical process modelled	Reference
Microphysics	Clustering of Geant4 energy deposits by particle lineage; stochastic generation of scintillation photons n_γ and ionisation electrons n_e per cluster via the NEST yield model	Ch. 4
S1 simulation	Position-dependent detection of primary scintillation photons; optical propagation timing; scintillation decay components	Sec. 3.2
S2 simulation	Electron drift, diffusion, and attachment; liquid-to-gas extraction; proportional scintillation in the gas gap	Sec. 3.3
Detector & DAQ	Per-PMT single-photoelectron response; afterpulses; electronic noise; waveform digitisation	Sec. 3.4

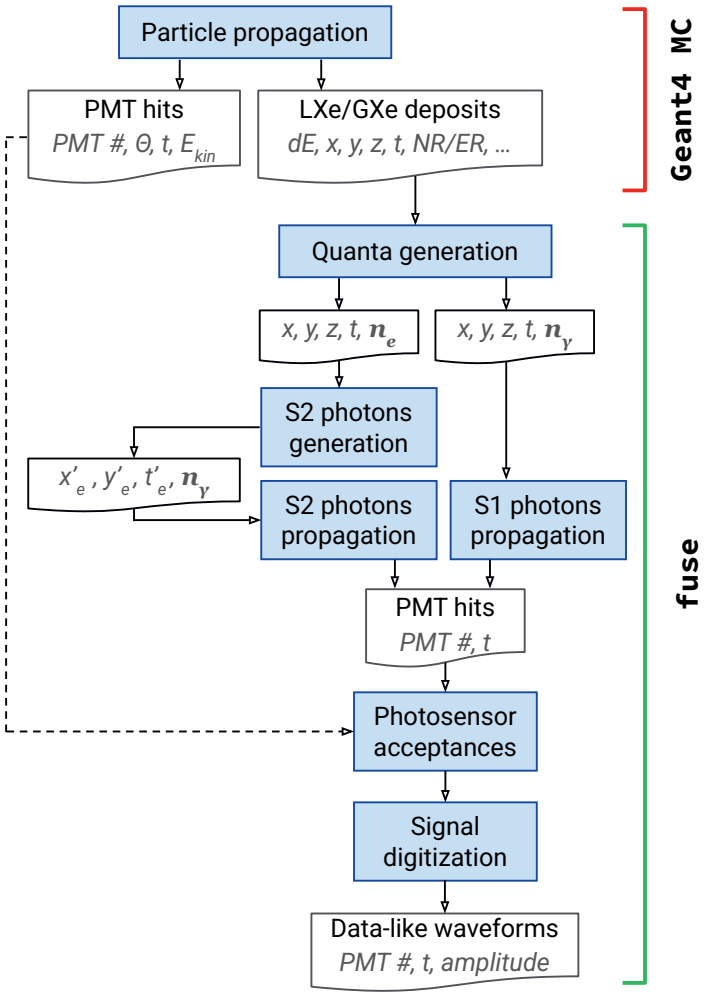


Figure 3.1 *Schematic overview of the* `fuse` *simulation chain. Blue boxes represent the main processing steps: particle propagation in Geant4, quanta generation, S1 and S2 photon propagation, photosensor acceptance, and signal digitisation. White boxes show the intermediate data products passed between steps, including energy deposits* $\{\Delta E, x, y, z, t\}$*, scintillation photon and ionisation electron counts* (n_γ, n_e)*, PMT hit times, and digitised waveform amplitudes. The final output enters* `straxen` *[74] identically to real detector data.*

3.1 Microphysics and Quanta Generation

Geant4 tracks particles through the detector geometry and records energy depositions, but does not model how that deposited energy converts to

scintillation photons and ionisation electrons in liquid xenon. This conversion requires a separate microphysical model. In `fuse`, this is not a first-principles calculation: rather than simulating ionisation and excitation at the atomic level, quanta are generated using semi-empirical yield models calibrated to experimental data. These models are defined at the level of a single, physically identifiable interaction site, at the spatial and temporal resolution at which individual interactions are resolved by the detector. Applying them directly to the hundreds of individual Geant4 tracking steps that a single event produces would therefore be physically inconsistent. The deposits must first be grouped into clusters, each representing a distinct scintillation site, before yields can be applied. The physical motivation for this clustering step, and its particular importance for high-energy multi-site topologies, is discussed in detail in Chapter 4.

`fuse` uses a lineage-based clustering algorithm that uses the full parent-child track hierarchy recorded by Geant4. Clusters are assembled by following the genealogy of each particle track. Particle-type changes (such as when a photon scatters into a Compton electron at a new site), large spatial separations between steps of the same track, and radioactive-decay transitions each start a new cluster. Each cluster is assigned a position and time equal to the energy-weighted mean of its constituent deposits, and is classified as either an electron recoil (ER) or a nuclear recoil (NR), the two interaction categories recognised by the yield model.

Each cluster is passed to NEST (Noble Element Simulation Technique) [75], a semi-empirical framework for modelling quanta production in liquid noble elements. Given the cluster energy, the local electric field from the COMSOL [76] map, and the ER or NR interaction type, NEST returns the mean numbers of scintillation photons n_γ and ionisation electrons n_e, together with recombination fluctuations. The free parameters of the yield model are not taken from the NEST defaults but are fitted to XENONnT calibration data. The yield physics and the parameter fit are described in Chapter 4.

Clusters in the TPC active volume undergo full quanta generation and contribute to both S1 and S2 signals. Clusters in the region below the cathode also generate quanta, but the drift field reverses below the cathode and pushes ionisation electrons toward the bottom PMT array, away from the active volume; they produce an S1 signal only, with no corresponding S2. Interactions in the gas phase are ignored by default.

3.2 S1 signal simulation

The S1 signal arises from primary scintillation: the excited xenon dimers (excimers) formed by direct excitation and by electron–ion recombination

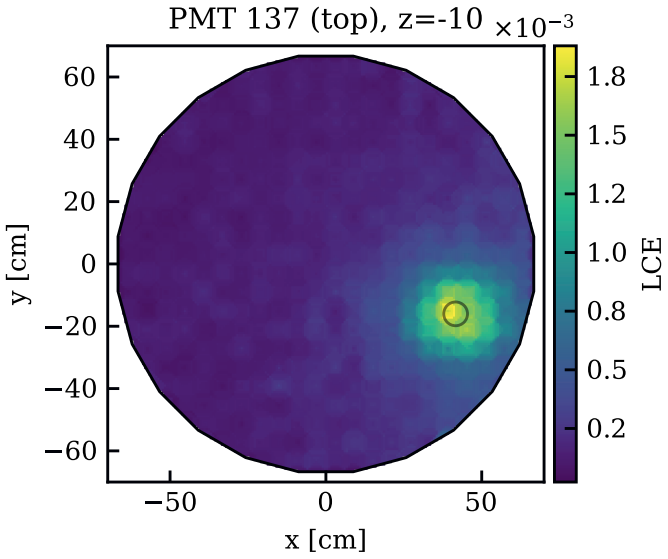


Figure 3.2 *Light collection efficiency (LCE) for PMT 137 as a function of (x, y) interaction position in a horizontal slice at $z = -10$ cm. The colour scale encodes the photon detection probability. The grey circle indicates the projected position of the PMT photocathode; the LCE peaks directly above the PMT and falls off with distance, reflecting the solid angle subtended by the photocathode.*

decay radiatively, emitting vacuum ultraviolet photons at 175 nm (Section 2.1) [47]. Excimers decay through two channels, the fast singlet state with lifetime $\tau_s = 3.1$ ns and the slow triplet state with lifetime $\tau_t = 24.0$ ns in liquid xenon [54], with the ratio of singlet to triplet emission depending on the interaction type. For nuclear recoils the singlet fraction is higher than for electron recoils, giving NR pulses a faster mean rise time. An additional time-distributed component arises from recombination luminescence, with photons arriving across an extended tail with a characteristic timescale of order 100 ns [77]. For nuclear recoils, the NEST timing model is called with fixed effective parameters (energy and exciton fraction) tuned to NR calibration data, with the local electric field retained event-by-event. Photon emission times beyond 1 μs are discarded, cutting a negligible fraction of the tail.

Tracking individual scintillation photons through the detector geometry at runtime would be computationally prohibitive. Instead, `fuse` uses a pre-computed three-dimensional light collection efficiency (LCE) map that encodes the probability of a photon produced at position (x, y, z) reaching each PMT photocathode. The map is derived from a Geant4

optical simulation of the TPC in which the reflectivity and absorption parameters of all surfaces are tuned to reproduce ^{83m}Kr calibration data, incorporating the full detector wire geometry. The number of detected S1 photons is drawn from a binomial distribution with probability equal to the LCE at the cluster position. Figure 3.2 shows an example for PMT 137, a bottom-array channel, in a horizontal slice at $z = -10$ cm.

Each detected photon is assigned to a specific PMT channel by sampling from the spatial LCE pattern. Its arrival time is drawn from the scintillation decay model convolved with a per-PMT optical propagation timing spline, derived from the same photon-transport simulation and encoding propagation delays and photocathode reflections. Double-photoelectron emission and the PMT transit time spread are applied as described in Section 3.4. The result is a set of per-PMT photon hit times and amplitudes constituting the simulated S1 signal.

3.3 S2 signal simulation

The S2 signal is produced by ionisation electrons that drift to the liquid–gas interface, are extracted into the gas phase, and undergo proportional scintillation as they are accelerated toward the anode. This multi-step process is sensitive to the electric field in the TPC volume, to the purity of the liquid xenon, and to the detailed geometry of the electrode region. The following subsections describe each stage in turn, including two physical effects specific to the XENONnT anode geometry that require explicit modelling: the non-uniform proportional scintillation induced by the wire structure (Section 3.3.2), and the strongly localised field distortion caused by perpendicular support wires on the anode grid (Section 3.3.3).

3.3.1 *Electron Drift and Diffusion*

Ionisation electrons produced in the TPC drift upward under the approximately uniform electric field of $\sim$23 V/cm established between the cathode (approximately -2.75 kV) and the gate electrode (approximately $+0.30$ kV). Two processes modify the electron distribution during drift. First, electrons attach to electronegative impurities with survival probability

$$P_{\mathrm{surv}}(t) = e^{-t/\tau_e}\,, \tag{3.1}$$

where τ_e is of order 10 ms. Second, the electron cloud spreads longitudinally and transversely, with standard deviations

$$\sigma_L(t) = \sqrt{2\,D_L\,t}\,, \qquad \sigma_T(t) = \sqrt{2\,D_T\,t}\,, \tag{3.2}$$

where D_L and D_T are the longitudinal and transverse diffusion coefficients. σ_L enters as an arrival-time spread σ_L/v_d; σ_T is applied directly as a transverse position shift. Longitudinal spreading directly determines the temporal width of the S2 signal; D_L is taken from a data-driven map extracted from ^{83m}Kr calibration data rather than a single constant, capturing position-dependent variations. The transverse coefficient $D_T \approx 4.8 \times 10^{-8}$ cm^2 ns^{-1} is obtained from [78] and has a small effect on S2 observables. For a point interaction at $z = -100$ cm the drift time is approximately 1.4 ms, giving $\sigma_L \approx 3.6$ mm ($\approx 5\,\mu$s in arrival time) and $\sigma_T \approx 3.7$ mm.

The electric field throughout the bulk of the active volume is shown in Figure 3.3. The drift velocity is not uniform across the TPC volume due to the non-uniformity of the field near the electrodes. A position-dependent drift velocity map derived from a COMSOL electrostatic simulation is used, normalised to reproduce the measured mean drift velocity of $v_d \approx 7 \times 10^{-5}$ cm ns^{-1} at the centre of the TPC. The simulation takes as input the full TPC geometry and the electrode potentials, and solves for the three-dimensional electrostatic field throughout the active volume. The same COMSOL simulation provides a survival probability map accounting for field lines that terminate on the PTFE reflector panels rather than reaching the gate electrode, causing a fraction of electrons near the TPC walls to be lost before completing their drift. This defines the charge-insensitive volume at the bottom corners of the TPC. The total drift-survival factor P_{drift} is the product of the attachment probability $P_{\text{surv}}(t)$ in Eq. (3.1) and the position-dependent field-line survival from this map.

Because the electric field is not perfectly uniform, electrons do not drift along vertical paths. A charge cloud originating at true position (x_0, y_0, z_0) exits the liquid surface at an apparent position $(x_{\text{obs}}, y_{\text{obs}})$ that can differ from the radial projection of the true position, particularly near the field-shaping rings and PTFE walls where field distortions are strongest. This effect is modelled using a three-dimensional field-distortion map derived from the same COMSOL simulation, which provides for each true position the corresponding observed surface-exit position. The position reconstruction algorithm reconstructs this surface-exit position rather than the true interaction site; a field distortion correction is applied in analysis to recover the true position.

3.3.2 *Extraction and Secondary Scintillation*

Electrons that survive the drift must cross the liquid–gas interface, after which they are accelerated through the gas gap toward the anode wires and emit electroluminescence light. The three stages of S2 formation are summarised in Figure 3.4.

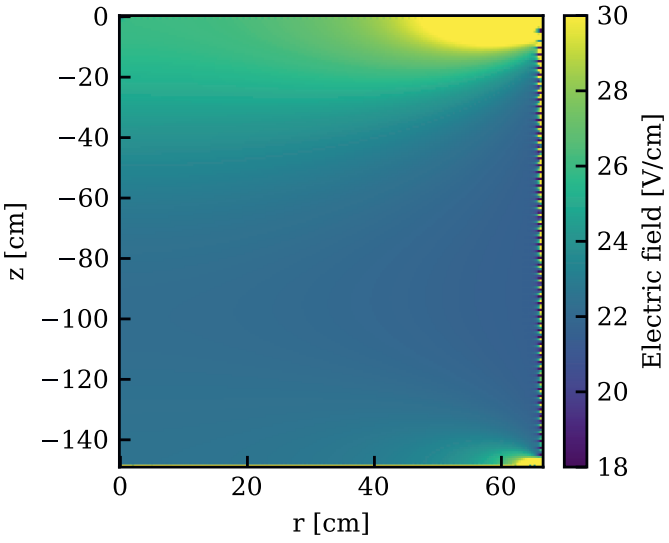


Figure 3.3 *COMSOL simulation of the electric field magnitude in the XENONnT TPC. The drift field is approximately uniform at ∼23 V/cm in the bulk active volume, with distortions near the gate and cathode electrodes. The same simulation provides the position-dependent drift velocity and electron survival probability maps used in* `fuse`.

An electron reaching the interface is extracted with probability $\varepsilon_{\text{ext}}(x, y)$, and each extracted electron produces on average $g_{\text{SE}}(x, y)$ photoelectrons in the gas gap. The mean S2 response therefore scales as

$$\langle S2_{\text{raw}}(x, y) \rangle \propto n_e^0 \, P_{\text{drift}} \, \varepsilon_{\text{ext}}(x, y) \, g_{\text{SE}}(x, y), \tag{3.3}$$

where n_e^0 is the number of ionisation electrons produced at the interaction site and P_{drift} is the combined drift-survival factor described in Section 3.3.1, accounting for both attachment losses and field-line termination on the PTFE walls.

In data, ε_{ext} and g_{SE} are partially degenerate: most S2 observables constrain their product more strongly than the two factors individually. `fuse` therefore parameterises the well-constrained combination. The spatial dependence is taken from the data-driven S2 correction map, while the absolute normalisation is anchored to the calibrated g_2 parameter of Section 2.5. With $g_{\text{SE}} \approx 25$ PE/e$^-$, the effective extraction efficiency is approximately 50%. Because the simulation is already normalised to the corrected S2 scale, the `straxen` single-electron-gain and extraction-efficiency corrections are not re-applied to simulation, to avoid double-counting.

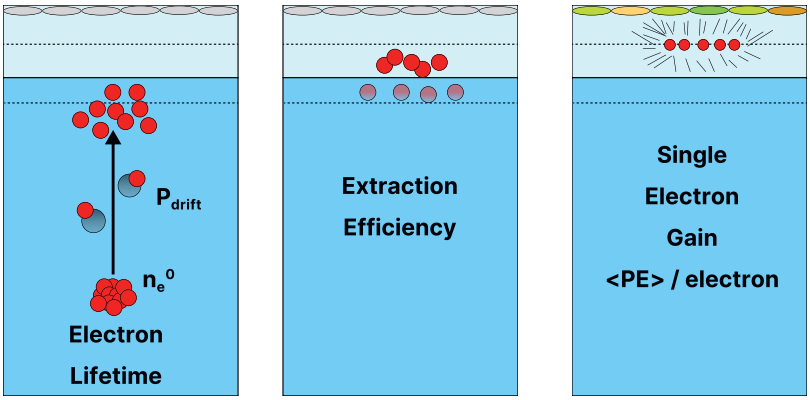


Figure 3.4 *Schematic of the three stages of S2 formation. Left: drift of the ionisation electron cloud from the interaction site to the liquid surface, broadened by longitudinal and transverse diffusion and attenuated by attachment on electronegative impurities (survival probability P_{drift}). Centre: extraction across the liquid–gas interface with probability $\varepsilon_{\mathrm{ext}}(x,y)$, modelled as a Bernoulli trial per electron. Right: proportional scintillation in the gas gap, where each extracted electron produces a Poisson-distributed number of photons with mean $g_{\mathrm{SE}}(x,y)$ set by the gas-gap field and the electron's lateral position relative to the anode wires.*

Extraction and emission are sampled per electron. The number of extracted electrons is

$$n_e^{\mathrm{ext}} \sim \mathrm{Binomial}(n_{e,\mathrm{if}}, \varepsilon_{\mathrm{ext}}) \,, \tag{3.4}$$

with $n_{e,\mathrm{if}}$ the number of electrons arriving at the interface after drift. Each extracted electron is then delayed by a trapping time of approximately 200 ns, representing the time spent at the surface before emission into the gas, and produces

$$N_\gamma \sim \mathrm{Poisson}(g_{\mathrm{SE}}(x,y)/(1+p_{\mathrm{DPE}})) \,, \tag{3.5}$$

photons. Since g_{SE} is measured in PE, dividing by $1+p_{\mathrm{DPE}}$ converts to photons, where $p_{\mathrm{DPE}} = 0.219$ [79] is the probability that a single VUV photon liberates two photoelectrons at the PMT photocathode. Double-photoelectron emission is then re-applied at the PMT level (Section 3.4).

As described in Chapter 2, the anode grid is a wire mesh with a 5 mm pitch and $\sim$216 μm wire diameter, reinforced by two perpendicular support wires oriented at 30° to the main wire direction. This wire structure makes the proportional scintillation strongly non-uniform on millimetre scales, for

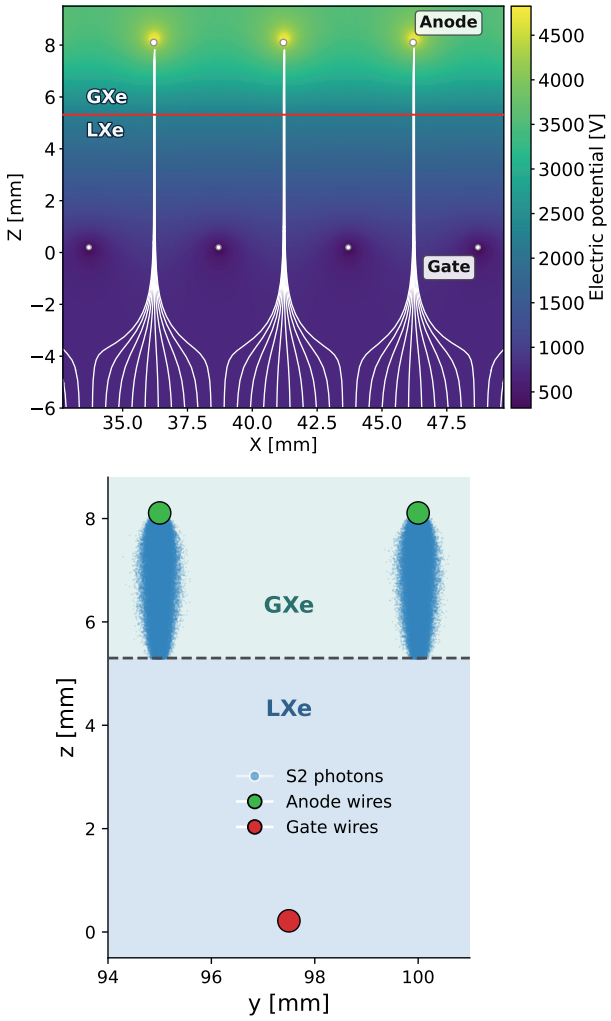


Figure 3.5 *Top: COMSOL simulation of the electric potential (colour scale, V) and field lines in the gas gap. Gate wires (white dots, $Z \approx 0$ mm) and anode wires (white dots, $Z \approx 8$ mm) are labelled; the red horizontal line marks the liquid-gas interface. Field lines from the bulk liquid converge tightly onto each anode wire, showing that extraction occurs preferentially below the wires. Bottom: S2 photon emission positions (blue dots) from Garfield++ electron transport simulations, in the YZ plane. Photon production is concentrated in a column directly below each anode wire (green circles), spanning the full gas gap height. The gate wire is shown as a red circle; the dashed line marks the liquid-gas interface.*

two reasons. First, electrons extracted at different lateral distances from an anode wire traverse different amplification paths in the gas gap: those

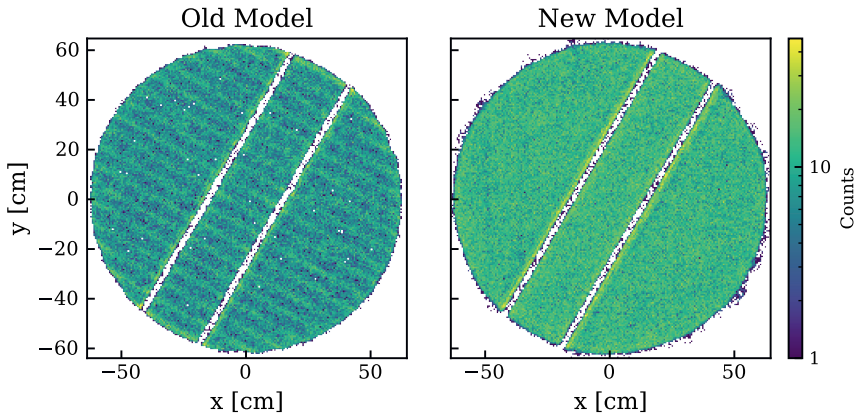


Figure 3.6 *Reconstructed XY distributions from simulations with uniform XY event positions, using two position reconstruction models: one trained without anode wire shadowing (left) and one trained with shadowing included in the optical map (right). The stripe artefacts visible on the left reflect the wire-pitch periodicity in the LCE map, which the model without shadowing cannot account for. Including shadowing in the training simulation recovers a flat distribution. A residual stripe at the perpendicular support wire position remains visible; this effect is discussed in Section 3.3.3.*

extracted directly below a wire experience a stronger field and a shorter path to the anode, producing more photons at greater heights within the gas gap, while electrons extracted midway between wires follow longer, lower-field trajectories. Second, the wire structures physically intercept a fraction of the photons produced nearby, with the fraction depending on the photon emission height. This is the anode shadowing effect. Together, these produce a position-dependent S2 light yield and top-to-bottom photon ratio that repeat with the 5 mm wire pitch across the anode plane.

To model the position-dependent emission correctly, pre-computed Garfield++ [80] maps of electron transport and electroluminescence in the gas gap are used to determine the three-dimensional photon emission profile as a function of electron extraction position relative to the wire. At the gas gap field of approximately 5700 V/cm (between the gate at approximately +0.30 kV and the anode at approximately +5 kV, separated by ∼0.8 cm), xenon scintillates proportionally. Figure 3.5 shows both the underlying field structure and the resulting emission profile. The COMSOL field lines (top) show that electrons extracted from the liquid are guided toward the anode wires from below; the Garfield++ emission profile (bottom) confirms that photon production is concentrated in a column directly below each anode wire, spanning the full gas gap height.

As with the S1 simulation, photons produced in the gas gap are distributed to individual PMT channels using a pre-computed S2 LCE map. This map is produced by a Geant4 optical simulation of the TPC, in which the position-dependent emission profile parameterised from the Garfield++ output is used as the photon source distribution. The simulation propagates photons through the full detector geometry, including absorption and reflection by the wire structures, so that the resulting LCE map encodes both the non-uniform emission profile and the geometric shadowing by the anode wires.

Figure 3.6 illustrates the effect of anode wire shadowing on position reconstruction. The simulated events are uniform in XY by construction, so any non-uniformity in the reconstructed positions is a reconstruction artefact. Two models are compared: one trained on simulations without shadowing, and one trained on simulations with shadowing included in the optical map. The first model has no knowledge of the wire-pitch periodicity in the LCE, and maps it into position space, producing visible stripe artefacts. The second model learns the correct light collection non-uniformity and recovers a flat distribution. Applied to data, this model correspondingly removes the stripe pattern observed in measured XY distributions. A residual stripe at the position of the perpendicular support wires is also visible; this effect requires separate treatment and is discussed in Section 3.3.3.

3.3.3 *Near-Wire Electric Field*

The macroscopic COMSOL electric field map captures the bulk drift field throughout the TPC volume but cannot resolve the strongly localised distortions near the anode wire structures. Figure 3.7 (top) shows the result of a finite-element simulation of this region. The perpendicular support wires, oriented at 30° to the main wire direction at $x \approx 13.1$ cm in the rotated frame, suppress the field to well below 10 V/cm directly beneath the wire, compared to the nominal $\sim$23 V/cm in the bulk. The perturbation extends ± 3 cm into the drift volume; electrons in this region are deflected laterally and arrive at the anode with a modified transverse position and a broader distribution of arrival times.

The same simulations are used to derive one-dimensional lookup tables giving the lateral displacement $\delta x(x_0)$ and additional arrival-time broadening $\sigma_t(x_0)$ as a function of initial transverse position x_0 relative to the wire centre. These corrections are applied to all electrons within ± 3 cm of the wire. Figure 3.7 (bottom) compares the S2 50% width between ^{83m}Kr data and simulation as a function of position in the rotated frame. Away from the wire the width sits at its nominal value of approximately 5 μs; it rises

to nearly 15 μs in the ± 3 cm region around $x \approx 13$ cm. The correction is derived purely from the COMSOL electrostatic simulation without tuning to data. The simulation reproduces the shape of the effect well but slightly underestimates its magnitude, which is visible as a small systematic offset in the normalised residuals near the wire position.

3.3.4 *Delayed Electrons and Photoionisation*

Both S1 and S2 VUV photons can photoionise trace impurities dissolved in the bulk liquid xenon, liberating electrons that drift upward and produce small delayed S2 pulses. Because the S2 signal contains far more photons than S1, only S2-induced photoionisation is modelled in `fuse`. Figure 3.8 illustrates the mechanism. S2 photons originate in the gas phase, so the photon flux decreases with depth into the liquid; impurities closer to the gas interface are more likely to be ionised, and short delays are more probable than long ones. This exponential delay-time distribution is observed in data and is motivated by a toy simulation of photon propagation in the TPC. Each delayed electron arrives within one drift time of the primary S2; its transverse position is drawn uniformly across the TPC cross-section, reflecting the uniform distribution of impurities in the bulk. Photoionisation contributes to the single-electron background and perturbs the inferred S2 area at the percent level.

The effect is parameterised in `fuse` by a photoionisation fraction PI, defined as

$$\mathrm{PI} \equiv \frac{\sum A_{\mathrm{delayed}}}{A_{S2}}\,, \tag{3.6}$$

where A_{S2} is the area of the primary S2 signal and $\sum A_{\mathrm{delayed}}$ is the total area of all delayed S2 pulses within one drift time. The number of delayed electrons generated per event is drawn from a Poisson distribution with mean proportional to PI and the primary S2 area. Each delayed electron is assigned a delay time drawn from an exponential distribution with time constant τ:

$$f(\Delta t) = \mathrm{PI} \cdot e^{-\Delta t/\tau}\,, \tag{3.7}$$

truncated at $t_{\max}$ corresponding to the full drift window, and placed at depth $z = -v_d\,\Delta t$, with a transverse position drawn uniformly within the TPC cross-section. The time constant is found to be stable across data-taking at $\tau \approx 1.3$ ms, while the photoionisation fraction varies across runs, tracking the evolution of xenon purity. Figure 3.9 shows the normalised delayed S2 area as a function of delay time for two runs at different purity conditions (PI $= 0.06$ and PI $= 0.15$), covering the full drift window of approximately

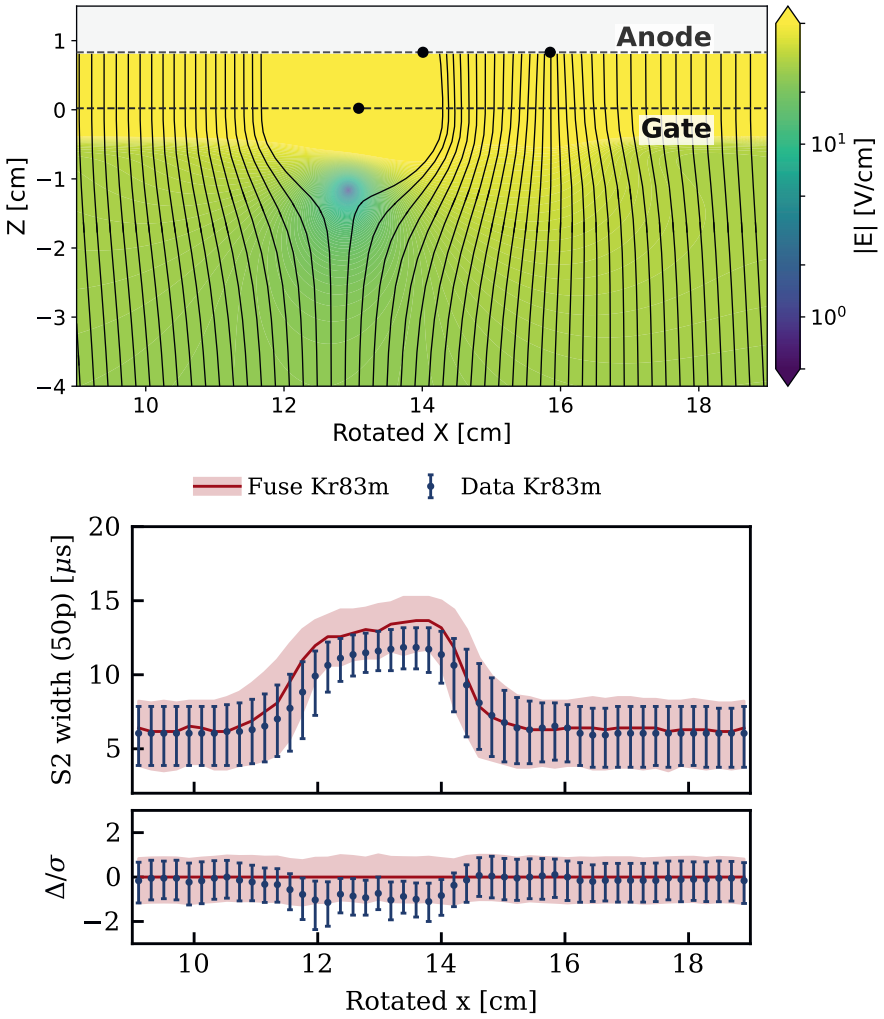


Figure 3.7 *Top: Electric field magnitude (log colour scale, V/cm) and field lines near the perpendicular support wires of the XENONnT anode grid. The gate and anode planes are shown as dashed lines; perpendicular wire positions are marked as black dots (two on the anode, one on the gate). Field lines converge onto the gate-level perpendicular wire, with the field dropping well below 10 V/cm directly beneath it compared to the ∼23 V/cm bulk value. Bottom: S2 50% width as a function of reconstructed position in the rotated frame, comparing data (dark blue) and simulation (red) for ^{83m}Kr. The pronounced peak near $x \approx 13$ cm is reproduced in shape by the simulation; the small residual offset reflects a slight underestimate of the effect magnitude.*

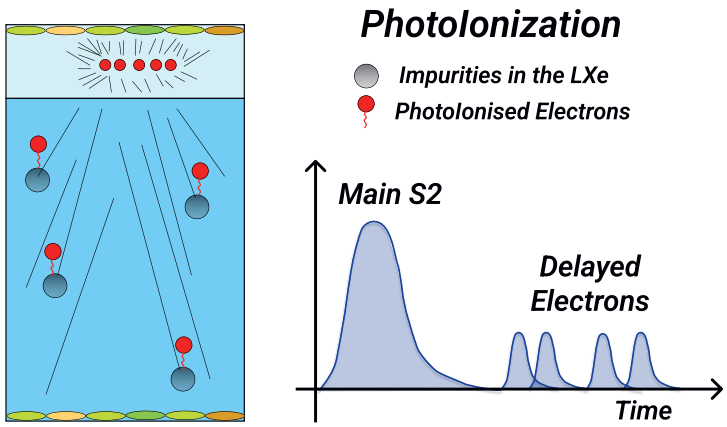


Figure 3.8 *Schematic of the photoionisation mechanism. VUV photons from the S2 signal (and, to a lesser extent, S1) ionise trace impurities distributed through the bulk liquid xenon, liberating electrons that drift toward the anode. Each liberated electron produces a delayed S2 pulse arriving after the main signal by a time determined by the depth at which it was produced.*

2.2 ms. Eq. (3.7) is overlaid on each distribution, and the parametrisation reproduces the data well in both cases.

The photoionisation model is implemented and parameterised for the science run conditions but is disabled in the default simulation configuration. Enabling it multiplies the electron count and increases runtime substantially, while the impact on primary-peak reconstruction is at the percent level and is handled as a separate correction in the analysis chain. The model is re-enabled for targeted studies of single-electron backgrounds and delayed emission, where the delayed-S2 spectrum is itself the observable of interest.

3.4 Detector Response

The per-PMT photon hit times produced by the S1 and S2 simulation stages are converted into digitised electrical waveforms by a model of the photosensor and readout response.

Each photon hit is converted into an electrical pulse using a per-PMT single-photoelectron (SPE) waveform template measured from calibration data, which has been noise-subtracted and normalised to unit mean area. The template spans a time window from -13 to $+195$ ns relative to the photon arrival time. A fraction p_{DPE} of photons produces two photoelec-

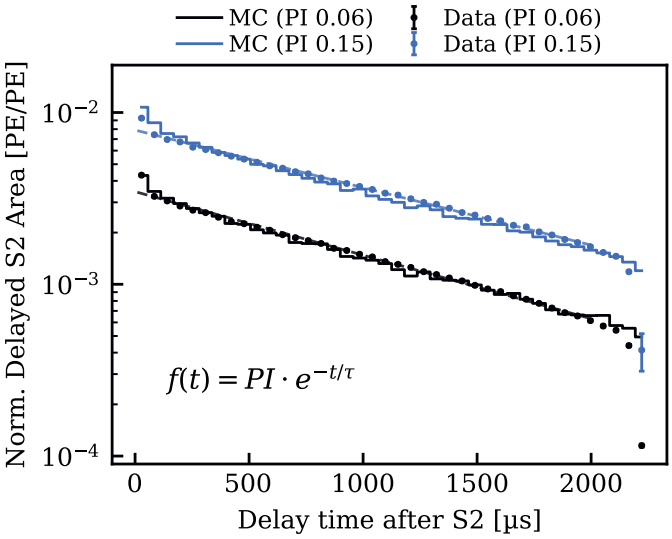


Figure 3.9 *Normalised delayed S2 area as a function of delay time for two runs at different xenon purity conditions (*PI $= 0.06$ *in black,* PI $= 0.15$ *in blue), covering the full drift window of approximately 2.2 ms. Data are shown as dots;* `fuse` *simulations as solid lines. The exponential model* $f(t) = \mathrm{PI} \cdot e^{-t/\tau}$ *is shown in the figure. The fitted parametrisation reproduces the measured distributions well in both cases, as expected from the fit.*

trons at the photocathode (as in the S1 simulation); the PMT transit time spread is modelled as a Gaussian with mean 46.0 ns and FWHM 9.0 ns, and per-channel gain values from calibrations are applied.

PMT afterpulses, caused by residual gas ions accelerating back to the photocathode after a primary pulse and producing secondary photoelectron emissions [60], are generated stochastically for each detected photon. Their time delays and amplitudes are sampled from per-channel empirical cumulative distribution functions (CDFs) measured during PMT calibrations. These afterpulses appear at delays of tens to hundreds of nanoseconds and contribute to the high-area tail of the signal response.

The overlapping SPE pulses on each channel are summed to form the analogue PMT output voltage, scaled via the circuit load resistance and an external pre-digitiser amplification of $\times 10$. Baseline electronic noise is added by superimposing a segment of a measured noise waveform recorded during noise calibration runs. The noise-plus-signal voltage trace is then quantised by a 14-bit analogue-to-digital converter with a dynamic range of 2.25 V and a reference baseline of 16000 counts, matching the XENONnT digitiser parameters [65]. A zero-length encoding (ZLE) algorithm identifies

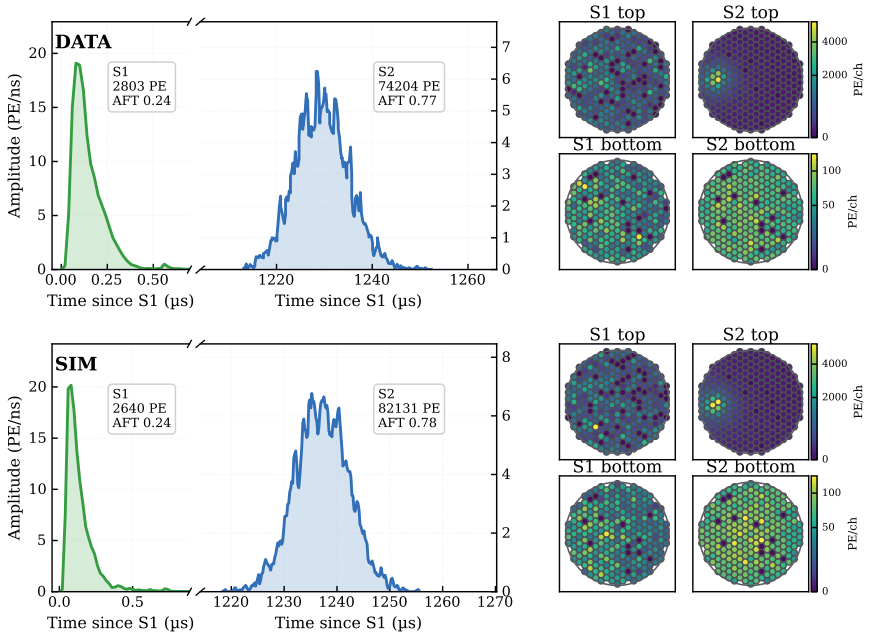


Figure 3.10 *Simulated (bottom) and data (top) TPC events shown at the reconstructed peaks level. Left: Summed PMT waveforms with a broken time axis, showing the S1 signal (green, 0–0.5 µs) and the S2 signal (blue, ∼1200–1260 µs) separately. Right: PMT hit patterns for the S1 and S2 signals, shown separately for the top and bottom arrays.*

contiguous samples exceeding a threshold of 15 ADC counts above baseline and retains only those segments, exactly replicating the data-reduction behaviour of the physical DAQ system. The resulting `raw_records` are format-identical to those from the physical DAQ (Figure 3.10) and are passed to `straxen` for reconstruction without modification, following the same chain described in Section 2.4.

3.5 Truth Information

`fuse` provides native access to truth-level information. Because both `fuse` and `straxen` are built on `strax` and share the same data model, the microphysics truth, comprising true interaction positions, numbers of quanta produced, particle types, and cluster identities, is stored alongside the simulated waveforms in the same format as any other `strax` data type. The data-processing chain itself is unaware of whether it is handling simulated or real data: it simply reads `raw_records` and writes processed data types. This means that any reconstructed quantity produced by `straxen` can be

compared directly to its true simulated counterpart on an event-by-event basis, without any additional bookkeeping.

Truth labels are propagated at multiple levels of the processing hierarchy: at the cluster level (which microphysics interaction contributed to each S1 or S2), at the peak level (which clusters are associated with each reconstructed peak), and at the raw-record level. This granularity enables a precise characterisation of reconstruction efficiencies, signal model biases, and the contributions of individual physical processes to observed signals. The bias characterisation described in Section 3.6 is one direct application. The systematic offset between reconstructed and true signal areas cannot be measured in real detector data, but is straightforward to evaluate in simulation because truth and reconstructed quantities share a common data representation.

3.6 Applications of the simulation

Three applications illustrate how `fuse` simulations enter the XENONnT analysis; detailed comparisons with calibration data are presented in the corresponding analysis chapters.

S2 morphology observables for event classifiers

A central application of `fuse` is the production of training samples for machine-learning-based event classifiers used in WIMP and low-energy analyses. These classifiers discriminate signal-like events from backgrounds on the basis of S2 morphology observables, and the resulting cut efficiencies are estimated directly from simulation. Accurate modelling of these observables is therefore a requirement for reliable efficiency corrections. The key morphology observables used as classifier inputs are the S2 50% width, the S2 rise time, and the corrected S2 area fraction top (cS2 AFT). The S2 width grows monotonically with drift time as the electron cloud spreads longitudinally; the AFT, the fraction of S2 light detected by the top PMT array, is sensitive to the depth of the interaction and to the S2 light collection model. A prerequisite for using these observables in simulation-based classifiers is that `fuse` reproduces them reliably across the full fiducial volume.

Figure 3.11 shows the comparison between simulation and SR2 calibration data for these observables. The S2 50% width as a function of drift time, checked against ^{37}Ar, ^{83m}Kr, and ^{220}Rn calibration sources, grows from near zero at short drift times to approximately 9 μs at the maximum drift time of $\sim$2 ms; Figure 3.11 (top) shows the ^{37}Ar comparison. The simulation reproduces this behaviour quantitatively across the full drift range. The AFT as a function of S2 area is evaluated for events near the

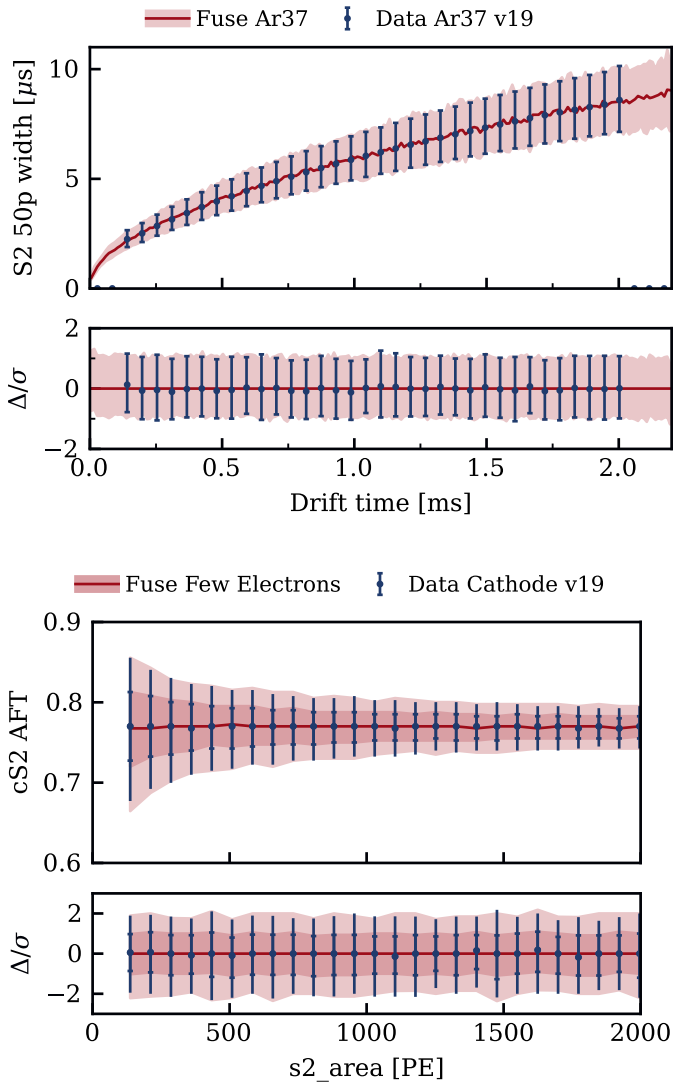


Figure 3.11 *Comparison of S2 morphology observables between SR2 data (blue) and simulation (red band). Top: S2 50% width as a function of drift time for the ^{37}Ar calibration source, growing from near zero to approximately 9 µs at maximum drift driven by longitudinal diffusion. Bottom: corrected S2 area fraction top (cS2 AFT) as a function of S2 area for cathode events, which probe the S2 light collection model most stringently. Error bars on the data points represent the 1σ statistical uncertainty in each bin. Normalised residuals are consistent with zero in both panels.*

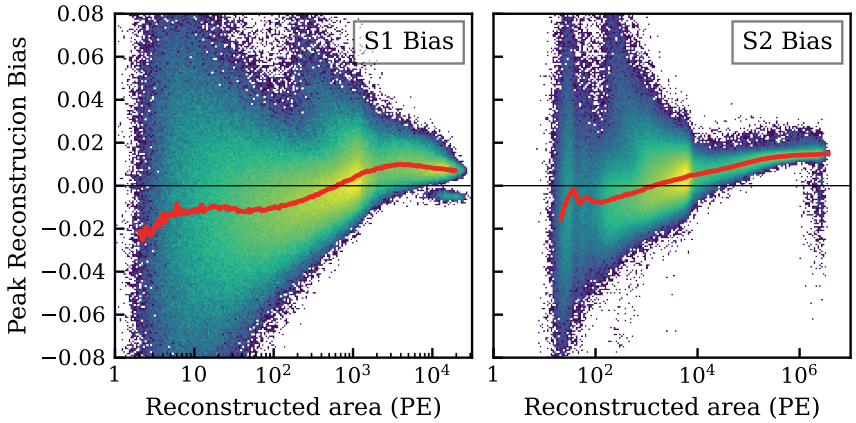


Figure 3.12 *Reconstruction bias for S1 (left) and S2 (right) signals as a function of reconstructed area, defined as* $(A_{\text{reco}} - A_{\text{true}})/A_{\text{true}}$*. The 2D histogram shows the distribution of individual events; the red line shows the median bias. At low areas the bias is negative, driven by the DAQ digitisation threshold and electronic noise; at high areas it turns positive due to PMT afterpulses contributing additional charge to the reconstructed signal.*

cathode. These events have drifted the full detector height and lost a large fraction of their electrons to attachment, producing small S2 signals in the range most relevant to low-energy analyses. The simulation reproduces the measured AFT ≈ 0.77 across 200–2000 PE.

Signal reconstruction bias

The access to truth information described in Section 3.5 enables a precise characterisation of the signal reconstruction bias, i.e. the systematic offset between the true signal area and the area reported by `straxen`. This bias arises from the combined effect of the DAQ digitisation threshold, electronic noise, PMT afterpulses, and any merging of delayed photoionisation electrons with their parent S2 signal. Because it requires comparing reconstructed and true quantities on an event-by-event basis, it is accessible only through simulation.

Figure 3.12 shows the S1 and S2 reconstruction bias as a function of signal area for SR2 simulations. At low areas the bias is negative, dominated by threshold and noise effects that cause small signals to be underestimated; at high areas it turns positive, driven by afterpulse contributions merging with the parent signal. The S1 bias is approximately -2% at low areas, crosses zero near 100 PE, and rises to $\sim +1\%$ at high areas. The S2 bias

follows a qualitatively similar trend, reaching $\sim +1.5\%$ at large signal sizes. These bias curves are used as corrections in the XENONnT analysis [66].

yield model validation and signal modelling

The yield model must reproduce the observed charge and light output across the energies and interaction types probed by calibration sources. The `fuse` framework is the primary tool for testing and validating the improved microphysical yield model developed in Chapter 4. Running the full simulation chain with the new yield parameterisation and comparing the output to calibration data shows whether the yield model correctly reproduces the observed charge and light signals across a wide range of energies and interaction types. In Chapter 5, the same full-chain simulation is used to produce the predicted detector response to ^{214}Bi beta decays, which is compared directly to the measured energy spectrum to extract information on the nuclear matrix elements governing the decay. The agreement between simulation and calibration data for each of these applications is established in the relevant analysis chapters.

3.7 summary

The four simulation stages described in this chapter, microphysics quanta generation, S1 photon propagation, S2 electron transport and proportional scintillation, and detector and DAQ digitisation, together produce output that is format-identical to real XENONnT data and passes through the same `straxen` reconstruction pipeline. Validations of the yield model against calibration data are presented in Chapter 4. The application to the ^{214}Bi spectrum analysis, including the bias corrections derived here, is described in Chapter 5.

4

XENON YIELDS AT HIGH ENERGIES

Reliable modelling of the detector response to electronic recoils is essential for a wide range of analyses in XENONnT. Prior to this work, simulations of ER above $\sim$100 keV showed significant discrepancies with data, particularly in the reconstructed S1 and S2 area distributions. These mismatches limited the reliability of Monte Carlo predictions.

Two main factors were identified as the dominant sources of these discrepancies. First, the light and charge yields at the XENONnT drift field were poorly constrained at high energies, requiring extrapolation of semi-empirical models beyond the region where experimental data were available. Second, the treatment of complex, multi-site interaction topologies, especially at high energies due to Compton scattering and secondary radiation, was insufficient in the standard clustering approach.

The improvements presented in this chapter address both limitations. The first ingredient is the availability of a clean sample of single-site ^{214}Bi beta-decays, selected via the BiPo delayed-coincidence technique and described in detail in Chapter 5. This dataset provides a direct, data-driven handle on ER light and charge yields in a topology that is effectively single-cluster. The second ingredient is a physics-motivated clustering strategy for Geant4 energy deposits, designed to treat multi-site high-energy topologies in a way that is consistent with how yields are defined and applied.

The combination of a data-driven ER yield model and physics-motivated clustering leads to a substantial improvement in the agreement between data and simulations over a wide energy range. The resulting yield model and clustering logic are used throughout the remainder of this thesis and are now the default settings of the simulation framework adopted by the XENONnT collaboration.

4.1 LIGHT AND CHARGE PRODUCTION IN LIQUID XENON

Electronic recoils in liquid xenon deposit energy predominantly through ionisation and excitation of xenon atoms. These processes lead to the

production of electron–ion pairs and excited states that subsequently form excimers. Not all deposited energy produces observable signals: about 20–30% is lost to sub-excitation electrons, which dissipate their energy as heat rather than creating additional quanta [47].

A key empirical observation is that this energy partition between observable quanta and heat remains approximately constant for electronic recoils over a wide energy range, up to the MeV scale. As a consequence, the total number of produced quanta is proportional to the deposited energy, motivating the definition:

$$N_q \equiv N_{\mathrm{ex}} + N_i = \frac{E}{W}, \qquad W \approx 13.7 \text{ eV/quanta}, \tag{4.1}$$

where E is the deposited energy, N_{ex} is the number of excitons, N_i the number of electron–ion pairs, and W denotes the average energy required to produce one observable quantum [67]. Here, W is an effective parameter that already includes the average energy lost to heat.

While electronic recoils lose a nearly constant fraction of their energy to heat, nuclear recoils lose an additional, strongly energy-dependent fraction of their energy to atomic motion through elastic collisions. As a result, only a fraction of the true nuclear recoil energy is converted into observable quanta. This energy-dependent quenching is commonly described by the Lindhard factor [81]. For electronic recoils, the approximately constant heat fraction can be absorbed into the effective W-value, leading to a linear relation between deposited energy and total quanta. For nuclear recoils, this is no longer the case: the observable energy depends non-linearly on the true recoil energy. This is why nuclear recoils are typically expressed in a separate energy scale ($\mathrm{keV_{nr}}$) distinct from the electronic-recoil scale ($\mathrm{keV_{ee}}$).

Recent measurements suggest a lower value of $W = 11.5$ eV/quanta [82]. In this work we adopt $W = 13.7$ eV/quanta for consistency with the XENONnT collaboration convention. This choice has negligible impact on reconstructed energies, since W is degenerate with the calibrated gain factors g_1 and g_2.

4.1.1 *Recombination: The Light-Charge Partition*

Immediately after the interaction, a fraction of the deposited energy is stored in excitons (N_{ex}), while the remaining fraction produces electron-ion pairs (N_i). The initial ratio between excitons and ion pairs is nearly constant. Early measurements in liquid xenon indicate a value of $N_{\mathrm{ex}}/N_i \simeq 0.06$ for ER interactions [83]. This implies that, to first order, the partition between N_{ex} and N_i is fixed once the total number of quanta N_q is determined by

Eq. (4.1). Excitons form excimers and de-excite on picosecond timescales, emitting scintillation photons [84]. The electron-ion pairs, on the other hand, can either be separated by the applied drift field or undergo recombination. Recombination leads to the formation of an excited xenon dimer, which subsequently de-excites and emits an additional scintillation photon. As a result, each recombined electron effectively adds one photon to the scintillation signal. Denoting by r the recombination probability for the ionisation electrons, the final number of scintillation photons (N_γ) and free electrons (N_e) produced after recombination can be written as:

$$N_\gamma = N_{\text{ex}} + r\, N_i, \tag{4.2}$$

$$N_e = (1 - r)\, N_i. \tag{4.3}$$

These relations illustrate the intrinsic anti-correlation between scintillation light and ionisation charge at fixed deposited energy: increased recombination leads to more photons and fewer free electrons, while reduced recombination has the opposite effect. In a dual-phase TPC, this microscopic anti-correlation directly manifests as an event-by-event anti-correlation between the measured S1 and S2 signals, as described in [47]. It is convenient to define the light yield and charge yield as:

$$L_y(E, \mathcal{E}) \equiv \frac{N_\gamma}{E}, \qquad Q_y(E, \mathcal{E}) \equiv \frac{N_e}{E}, \tag{4.4}$$

where $\mathcal{E}$ denotes the drift field. Combining Eqs. (4.1)–(4.3) shows that the sum of yields is constrained by energy conservation:

$$L_y(E, \mathcal{E}) + Q_y(E, \mathcal{E}) \simeq \frac{N_q}{E} \simeq \frac{1}{W}. \tag{4.5}$$

This relation provides a useful guiding principle: once one of the two yields (light or charge) is measured precisely, the other is strongly constrained. This point will be important later when we extract yields from data and enforce a consistent partition between light and charge.

4.1.2 *Thomas-Imel and Doke-Birks Models*

Historically, the recombination behaviour of ER in noble liquids has been described in terms of two limiting regimes, illustrated schematically in Figure 4.1.

thomas–imel box model (low-energy / short tracks)

At low energies, typically below ~ 15 keV, electronic recoil tracks are shorter

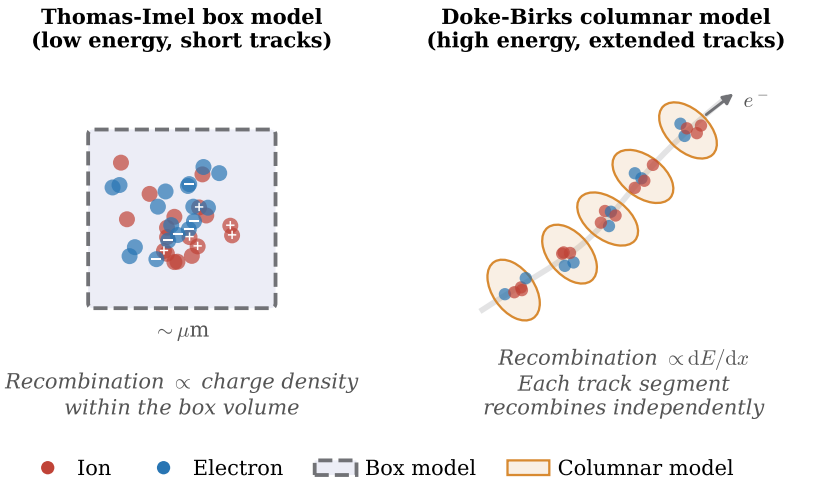


Figure 4.1 *Schematic illustration of the two limiting recombination regimes in liquid xenon. Left: at low energies, ionisation electrons thermalise within a few micrometers of their creation point, forming a compact charge cloud in which recombination is governed by the total electron-ion density, as described by the Thomas–Imel box model. Right: at high energies, the electron track extends over longer distances and the ionisation is distributed along a column. Each segment recombines independently, with a probability controlled by the local stopping power* $\mathrm{d}E/\mathrm{d}x$*, as described by the Doke–Birks model.*

than the characteristic electron thermalization length of a few micrometers in liquid xenon [67, 85]. In this regime, ionisation electrons diffuse and thermalize before they can spatially separate from the ion cloud, and the entire ionisation distribution can be treated as a compact "box" of charge. Recombination is then governed by the competition between the initial electron–ion density within this volume and the ability of the applied electric field to extract electrons before they recombine.

Under these assumptions, Thomas and Imel derived an analytic expression for the mean recombination probability [57, 67, 75]:

$$r = 1 - \frac{\ln(1+\xi)}{\xi}, \tag{4.6}$$

where, in the Thomas–Imel "box" model,

$$\xi = \frac{N_i\,\alpha}{4a^2 v_d}. \tag{4.7}$$

Here N_i is the number of ion pairs produced in the track segment, v_d is the electron drift velocity (with units of length per time), α is an effective

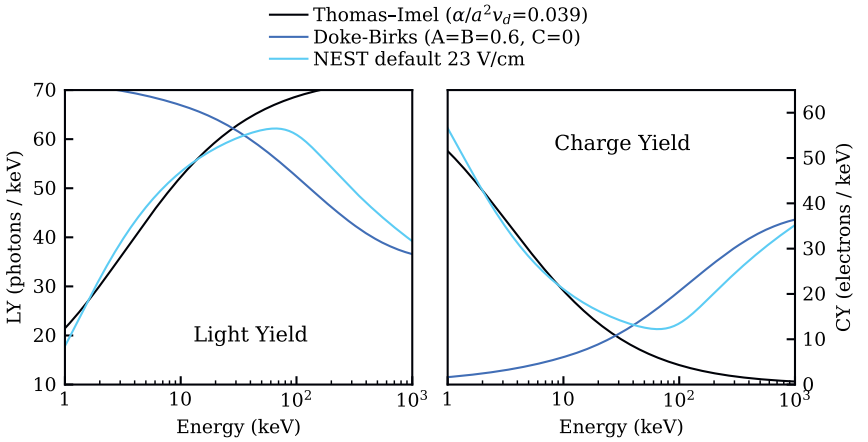


Figure 4.2 *Light yield L_y (left) and charge yield Q_y (right) evaluated from the Thomas–Imel and Doke–Birks models at the XENONnT drift field of* $\sim$23 V/cm*, compared to the default NEST β-ER prediction. For the Thomas–Imel model, $\alpha/(a^2 v_d) = 0.039$ is used; for Doke–Birks, $A = B = 0.6$ and $C = 0$. The stopping power $\mathrm{d}E/\mathrm{d}x$ is taken from the NIST ESTAR tables [86]. These parameter values are chosen for illustration only and are not the result of a fit to XENONnT data; a full fit is required to determine the field-dependent constants, as done in the NEST framework [75].*

volume recombination coefficient (with units of length^3/time), and a is a characteristic length scale entering the Thomas–Imel formalism.[1] The combination $N_i\,\alpha/(4a^2 v_d)$ is dimensionless: α carries units of volume per time, while $a^2 v_d$ has units of volume per time as well, so their ratio is unitless. The dimensionless parameter ξ therefore quantifies the competition between recombination and electron escape. It increases with ionisation density ($\xi \propto N_i$) and decreases when electrons are removed more efficiently by the applied drift field ($\xi \propto 1/v_d$).

Experimental data constrain only the grouped ratio $\alpha/(a^2 v_d)$ rather than the individual parameters separately. Fits to electronic recoil data yield $\alpha/(a^2 v_d) \simeq 0.03$–$0.04$ [67], corresponding to effective length scales of order 0.2–2 μm, comparable to the electron thermalization distance in liquid xenon [75, 85].

Doke–Birks model (high-energy / extended tracks)

At higher energies, typically above $\sim$ 50 keV, electronic recoil tracks extend

[1] In the original Thomas–Imel model, a was interpreted as the size of the ionisation box. Dahl [67] reformulated the model using detailed track simulations, where a acts as a local smoothing scale rather than the literal track size. The data constrain only $\alpha/(a^2 v_d)$, so the microscopic interpretation of a remains model-dependent.

over distances much larger than the electron thermalization length. The ionisation column is now spatially extended, and different segments along the track recombine independently. In this regime, recombination becomes a local process controlled by the ionisation density rather than the total number of ion pairs. The relevant variable is the Linear Energy Transfer (LET) or stopping power $\mathrm{d}E/\mathrm{d}x$, which quantifies the energy deposited per unit path length.

A phenomenological description for the recombination probability of this regime is the Doke–Birks model [87], inspired by Birks' law [88]:

$$r(\mathrm{d}E/\mathrm{d}x) = \frac{A \cdot (\mathrm{d}E/\mathrm{d}x)}{1 + B \cdot (\mathrm{d}E/\mathrm{d}x)} + C, \tag{4.8}$$

where A, B, and C are empirical parameters that depend on the applied drift field. At low and intermediate LET, increasing $\mathrm{d}E/\mathrm{d}x$ increases the local ion density and therefore enhances recombination. At sufficiently high LET, however, the recombination probability approaches a plateau. This saturation behaviour is experimentally observed as a "flat-top response," in which the scintillation yield remains approximately constant over a wide range of high ionisation densities [87]. Physically, in very dense tracks the local space charge partially screens the external electric field, and electron escape becomes increasingly suppressed. As a result, r cannot grow indefinitely and asymptotically approaches unity. At even higher ionisation densities, such as those produced by alpha particles with energies of a few MeV, the light yield can decrease due to bi-excitonic quenching processes, in which two excitons interact and produce only a single detectable photon. This effect lies beyond the simple Doke–Birks parameterisation.

An alternative perspective that has been explored in liquid noble detectors is to describe extended, high-energy tracks as a superposition of many localised recombination domains. In this picture, a long electron track is decomposed into a sequence of small segments, each treated approximately within a Thomas–Imel-like formalism. Such an approach was demonstrated in liquid argon by the ArgoNeuT collaboration [89], where highly ionising tracks were modelled as a collection of independent "boxes", avoiding the need for an explicit transition between distinct low- and high-energy formalisms. Conceptually, this multi-segment picture emphasizes that the difference between low- and high-energy regimes may arise from track structure rather than from fundamentally different recombination physics. In this view, the energy dependence of light and charge yields emerges from how energy is distributed along the track, rather than from switching between separate parametrizations.

Despite these different formulations, the Thomas–Imel and Doke–Birks expressions, Eqs. (4.6) and (4.8), should ultimately be viewed as effective descriptions of limiting regimes rather than exact microscopic laws. In real detectors and across wide energy ranges, ER interactions interpolate between these regimes and include additional complications such as secondary radiation (e.g., bremsstrahlung at high energies). This is one of the reasons why a direct, first-principles implementation of yields in full detector simulations is difficult.

To illustrate how these two limiting recombination pictures translate into different light and charge yields, we follow the approach presented in the Noble Element Simulation Technique (NEST) review paper [75], a semi-empirical framework that parameterises ER light and charge yields from a compilation of liquid xenon measurements, where the Thomas–Imel and Doke–Birks models are evaluated using fixed parameter values corresponding to a representative drift field.

Here, we adopt the same strategy, evaluating both models at the XENONnT drift field of $\sim$23 V/cm using illustrative parameter values ($\alpha/(a^2 v_d) = 0.039$ for Thomas–Imel; $A = B = 0.6$, $C = 0$ for Doke–Birks). The stopping power $\mathrm{d}E/\mathrm{d}x$ for electrons in xenon is taken from the NIST ESTAR tables [86]. The resulting recombination fractions are converted into light and charge yields, L_y and Q_y, using Eqs. (4.2) and (4.3).

Figure 4.2 shows the resulting yield curves alongside the default NEST beta yield prediction at the same drift field. At low energies, the Thomas–Imel box model dominates, while at high energies the Doke–Birks columnar description becomes appropriate. The NEST model smoothly interpolates between these limiting behaviours, demonstrating why semi-empirical parameterisations are necessary for practical detector simulations.

4.1.3 *Quanta and Recombination Fluctuations*

Having established how the mean recombination fraction depends on energy and track structure, we now turn to recombination fluctuations, which govern the event-by-event spread of the light and charge signals. At fixed deposited energy and field, the recombination fraction fluctuates due to variations in the initial ionisation pattern and in electron transport and thermalization. These recombination fluctuations are a dominant contribution to the intrinsic width of S1 and S2 distributions, and therefore to the achievable energy resolution.

For a fixed deposited energy E, neglecting fluctuations in the energy deposition itself, the total number of produced quanta fluctuates around the mean value $N_q = E/W$. In an ideal Poisson process, one would expect the variance to be $\mathrm{Var}(N_q) = N_q$. In reality, quanta production is correlated:

the creation of one excitation or ionisation reduces the energy available for subsequent quanta, resulting in suppressed fluctuations.

This behaviour is commonly described by introducing a Fano factor F, defined as:

$$\mathrm{Var}(N_q) \equiv F\, N_q, \tag{4.9}$$

with $F < 1$ indicating sub-Poissonian fluctuations. This concept was originally introduced by Fano in the context of gas ionisation [90] and later extended to condensed media. Early measurements and theoretical considerations for liquid xenon indicate an exceptionally small microscopic Fano factor, $F \simeq 0.04$–0.06, comparable to that of semiconductor detectors [91]. At higher energies, Fano fluctuations become less important relative to other sources of variance. The absolute uncertainty $\sigma_{N_q} \propto \sqrt{N_q}$ grows with energy, but the fractional uncertainty $\sigma_{N_q}/N_q \propto 1/\sqrt{E}$ decreases, making recombination fluctuations the dominant contribution to the observed signal widths.

Once N_q is realised, additional variance arises from the partition between photons and electrons through recombination. In a purely binomial picture, each of the N_i ionisation electrons would recombine independently with probability r, leading to a variance $r(1-r)N_i$ in the number of recombined electrons. Under this assumption, fluctuations in the charge signal N_e (and correspondingly in the light signal N_γ) would scale linearly with N_i.

Experimental data in liquid xenon show that this expectation is not sufficient. The observed widths of the electronic-recoil band in single-channel observables, such as N_e, N_γ, or their ratio, are much larger than predicted by binomial statistics [67, 92, 93]. It is important to emphasise that this discrepancy refers to the partition between light and charge, not to the total deposited energy.

This distinction became clear with the experimental demonstration of the intrinsic anti-correlation between scintillation and ionisation. Because the total number of quanta $N_q = N_\gamma + N_e$ is approximately fixed at given energy, recombination fluctuations shift events along an anti-correlated direction in the $(S1, S2)$ plane: an electron that recombines increases N_γ and decreases N_e, while an escaping electron has the opposite effect. This anti-correlation was first observed in liquid argon [94] and later in liquid xenon by the EXO collaboration [95]. By summing the scintillation and ionisation signals, a much improved energy resolution was obtained, demonstrating that the large single-channel fluctuations correspond to a redistribution of quanta between light and charge rather than to fluctuations of the total energy. In the combined energy estimator, recombination fluctuations largely cancel, and the variance is again governed by the Fano factor of the total quanta production.

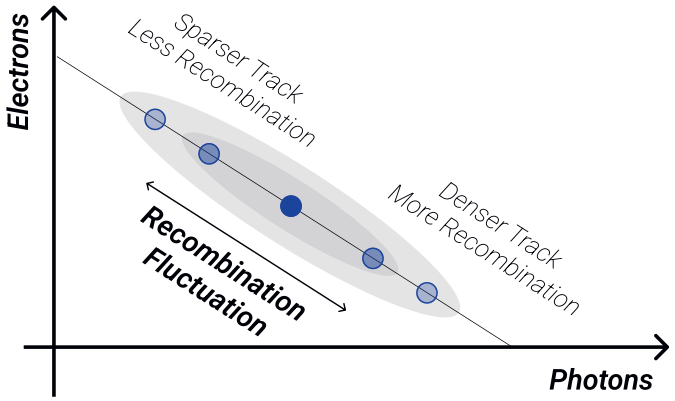


Figure 4.3 *Schematic illustration of recombination fluctuations at fixed deposited energy in liquid xenon. The horizontal axis represents the number of scintillation photons N_γ (proportional to S1), while the vertical axis shows the number of ionisation electrons N_e (proportional to S2). Each blue point corresponds to an interaction with the same total number of quanta $N_q = N_\gamma + N_e$, but with varying recombination. Differences in track structure lead to event-by-event fluctuations in the recombination fraction: increased recombination converts electrons into photons, shifting events along a line of constant N_q. This results in an intrinsic anti-correlation between light and charge.*

At low energies relevant for dark matter searches, Dahl showed using XENON10 data that the width of the electronic-recoil band in the discrimination parameter $\log_{10}(N_e/N_\gamma)$ does not follow the binomial expectation, being off by factors of $\mathcal{O}(100)$ in variance [67]. These observations established that recombination fluctuations are not independent electron-by-electron processes. Instead, the effective recombination probability itself fluctuates from event-to-event, reflecting variations in local track structure and ionisation density.

A quantitative description of this behaviour was later developed in the LUX experiment. Dobi observed that the excess variance scales approximately quadratically with the number of ion pairs, $\sigma^2 \propto N_i^2$ [96]. This quadratic scaling provides the empirical basis for the recombination fluctuation model introduced by the LUX collaboration [93] and later implemented in NEST [75].

The variance of the number of recombined electrons R is modelled in NEST by explicitly separating binomial and non-binomial contributions:

$$\mathrm{Var}(R) \simeq r(1-r)N_i + \omega^2 N_i^2, \tag{4.10}$$

where the first term represents binomial recombination statistics and the second term captures the empirically observed excess variance that scales approximately with N_i^2. The dimensionless parameter ω is energy dependent and controls the strength of these non-binomial recombination fluctuations.

A defining feature of the NEST recombination fluctuations model is that ω is not parameterised directly as a function of deposited energy. Instead, it is expressed as a function of the electron fraction,

$$f_e \equiv \frac{N_e}{N_e + N_\gamma} = \frac{N_e}{N_q}, \tag{4.11}$$

which characterises the light–charge partition of the event. Physically, f_e encodes how close the interaction is to the limits of full recombination ($f_e \to 0$) or full charge extraction ($f_e \to 1$). In both limits, fluctuations are expected to be suppressed, since the partition becomes effectively saturated. To reproduce the observed behaviour of the ER band widths, NEST parameterises $\omega(f_e)$ using a skewed Gaussian function:

$$\omega(f_e) = A(\mathcal{E}) \exp\left[-\frac{(f_e - \xi)^2}{2\omega_0^2}\right] \left[1 + \mathrm{erf}\left(\alpha_p \frac{f_e - \xi}{\omega_0\sqrt{2}}\right)\right], \tag{4.12}$$

where $A(\mathcal{E})$ is a field-dependent amplitude, ξ determines the centroid of the distribution, ω_0 its width, and α_p a skewness parameter. This form ensures that recombination fluctuations peak near intermediate electron fractions, where light and charge are most balanced, and are reduced toward the edges of the ER band. Once the mean ER yields are fixed, the mapping $E \mapsto f_e(E)$ uniquely determines an effective energy dependence $\omega(E) = \omega(f_e(E))$. In this way, the behaviour of ω as a function of energy is entirely inherited from the yield model and the resulting electron fraction.

At low energies, detector-related effects such as photon counting statistics and gain fluctuations dominate the observed S1 and S2 widths. At higher energies, these instrumental effects become sub-dominant and recombination fluctuations described by Eqs. (4.10-4.12) set the intrinsic width of the ER band. As discussed in Section 4.6, this modelling choice has important consequences for the description of high-energy ER data in XENONnT.

4.1.4 *From Recombination to Yield Models*

For detector simulation and analysis it is often impractical to work directly with a microscopic model of the recombination fraction r. Instead, light

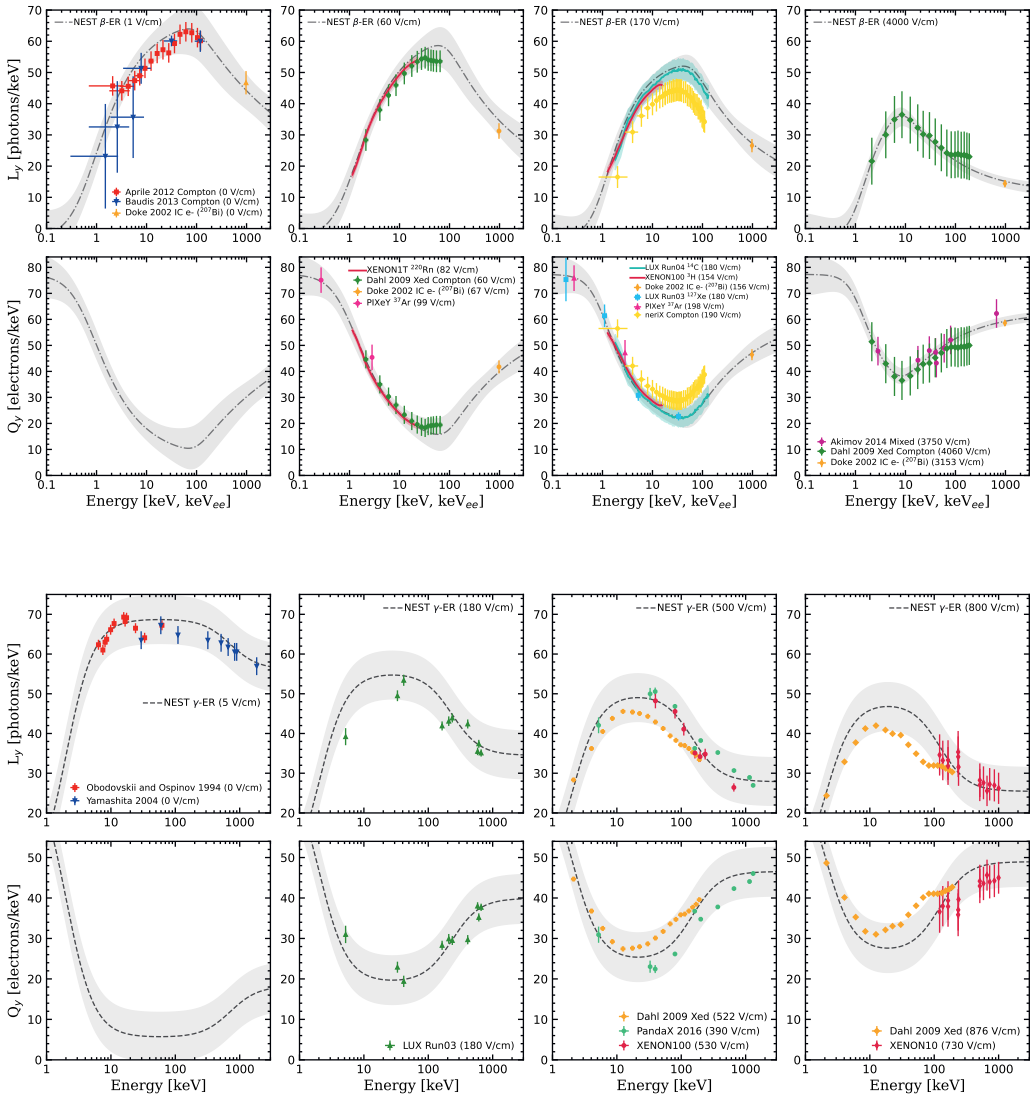


Figure 4.4 *Semi-empirical light and charge yield models for electronic recoils in liquid xenon. Figure from the NEST review [75]. The upper set of panels shows the β-ER model, while the lower set shows the γ-ER model. For each model, the top row displays the light yield L_y and the bottom row the charge yield Q_y as a function of deposited energy. Columns group measurements taken at similar drift fields, increasing from approximately zero field (left) to $\sim 800\,\mathrm{V/cm}$ (right). Note that even at vanishing external field, a non-zero charge yield is predicted: although the measured ionisation signal goes to zero, a finite fraction of electrons escape recombination and can be inferred from the total quanta minus the observed scintillation signal.*

and charge production are typically described in terms of energy- and field-dependent yield curves, $L_y(E, \mathcal{E})$ and $Q_y(E, \mathcal{E})$, which directly parameterise the mean detector response.

A widely used implementation of such yield models in liquid xenon is provided by NEST [75]. NEST compiles a large body of experimental data from different detectors, drift fields, and interaction types, and constructs semi-empirical parameterisations of the ER light and charge yields. NEST provides two distinct ER yield models: one intended for β-like interactions and one for γ-like interactions, as shown in Figure 4.4. The distinction reflects the fact that γ rays typically deposit energy through multiple Compton scatters, while beta-decays produce a single, continuous electron track.

Both models, however, share a common feature: a non-zero charge yield persists even at very low or zero drift field. This reflects the fact that electron-ion recombination in liquid xenon is not automatic: following ionisation, electrons thermalize and diffuse over micrometer-scale distances, and a significant fraction escape recombination even in the absence of an applied field [67, 75, 91]. The role of the drift field is therefore to suppress recombination further, rather than to enable electron escape in the first place.

A common feature of the NEST yield parameterisations is that they interpolate smoothly between the two limiting recombination regimes discussed in the previous section. For example, the charge yield is parameterised as:

$$Q_y(E, \mathcal{E}) = m_1(\mathcal{E}) + \frac{m_2 - m_1(\mathcal{E})}{\left[1 + \left(\frac{E}{m_3(\mathcal{E})}\right)^{m_4(\mathcal{E})}\right]^{m_9}} + m_5(\mathcal{E}) - \frac{m_5(\mathcal{E})}{\left[1 + \left(\frac{E}{m_7(\mathcal{E})}\right)^{m_8}\right]^{m_{10}(\mathcal{E})}}, \quad (4.13)$$

where the field-dependent parameters $m_i(\mathcal{E})$ control the behaviour in different energy regimes. The first term group $(m_1, m_2, m_3, m_4, m_9)$ governs the low-energy behaviour, where the Thomas–Imel box model dominates. The second term group (m_5, m_7, m_8, m_{10}) controls the high-energy saturation, corresponding to the Doke–Birks columnar regime. The parameter m_6 does not appear because it is set to zero in both the β- and γ-ER models, being degenerate with m_1. This functional form allows a smooth transition between the two limiting cases while maintaining consistency with experimental data across a wide range of energies and drift fields.

While NEST provides a well-established framework for many applications, two limitations are relevant for the present work. First, the experimen-

tal coverage at high energies is sparse, particularly for β-like interactions, where only a small number of measurements exist above $\sim$100 keV. Second, the coverage at low drift fields is limited, requiring extrapolation to the XENONnT operating field (23 V/cm). As a result, the default NEST predictions for high-energy ER interactions at the XENONnT field rely on extrapolations that are only weakly constrained by data.

These limitations motivate the two central questions addressed in this chapter. First, can the high-energy ER yields at the XENONnT drift field be determined directly from XENONnT data, rather than inferred from extrapolated global models? Second, to what extent is it necessary to treat β- and γ-induced ER interactions with distinct yield models, as opposed to describing their different detector responses through interaction topology?

We show that a single ER yield model extracted from single beta-decays of ^{214}Bi can accurately describe both β- and γ-dominated datasets when combined with a physically motivated clustering of energy deposits. The following section connects the microphysical picture introduced here to the full detector simulation chain and explains why clustering is a necessary ingredient when applying non-linear yield models to Geant4 energy deposits.

4.2 yields and clustering in simulations

The simulation chain in XENONnT usually starts from Geant4 simulations [69–71]. Geant4 provides a list of microscopic energy depositions, each associated with a position, time, deposited energy, particle identity, and parent particle information. To transform this microscopic interaction record into reconstructed peaks and events, XENONnT uses the waveform simulation framework `fuse` [68], followed by the standard processing software `straxen` [74]. Within `fuse`, one of the first steps after reading the Geant4 output is the conversion of energy deposits into scintillation photons and ionisation electrons. This conversion is performed using semi-empirical yield models, such as those provided by NEST [75], which return the expected number of photons and electrons for a given deposited energy and drift field. Yield models are not defined at the level of individual Geant4 steps. Rather, they are constrained by experimental data, where the detector observes integrated S1 and S2 signals arising from the combined effect of multiple microscopic energy deposits that cannot be resolved individually. As a result, Geant4 energy deposits must first be grouped into clusters before yields are applied. Each cluster represents an effective unit of energy deposition at the granularity with which the detector response is defined within `fuse`.

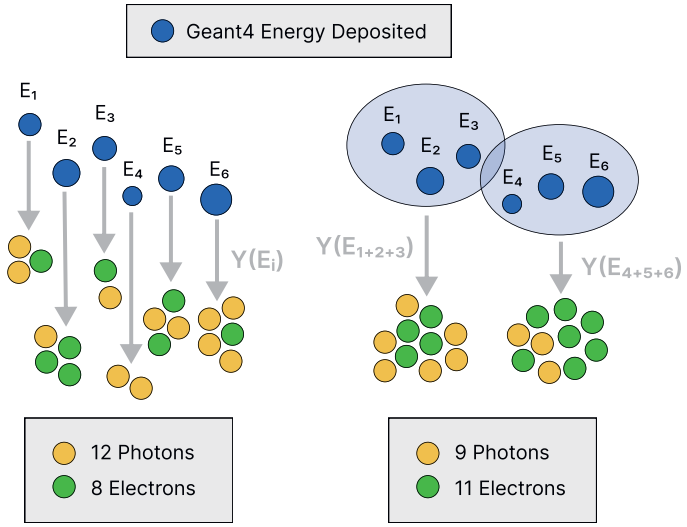


Figure 4.5 *Effect of clustering on the predicted number of scintillation photons and ionisation electrons. The same six Geant4 energy deposits are processed in two ways. Left: yields are applied to each deposit individually, producing 12 photons and 8 electrons in total. Right: the deposits are grouped into two clusters before yield application, producing 9 photons and 11 electrons. The total deposited energy is identical in both cases, but the S1/S2 partition differs because yield models are non-linear in energy.*

The need for clustering is not only technical. As discussed in Section 4.1, the partition of energy between light and charge is governed by recombination and depends on the density of deposited energy and on the drift field. Yield models are therefore non-linear in energy. Applying a yield model independently to many small deposits is not equivalent to applying it once to their sum. While the total reconstructed energy remains approximately conserved, the relative sizes of the S1 and S2 signals can change substantially depending on how energy is grouped. This is illustrated in Figure 4.5.

At low energies, below approximately 100 keV, this issue is largely irrelevant. Electronic recoils are typically compact and most events are effectively single-site at the scale relevant for yield assignment. In this regime, all deposited energy is naturally associated with a single physical cluster, and different reasonable clustering choices lead to nearly identical results.

Above $\sim$ 100 keV, the dominant interaction mechanism for γ rays changes. As shown in Figure 4.6, the Compton scattering cross section

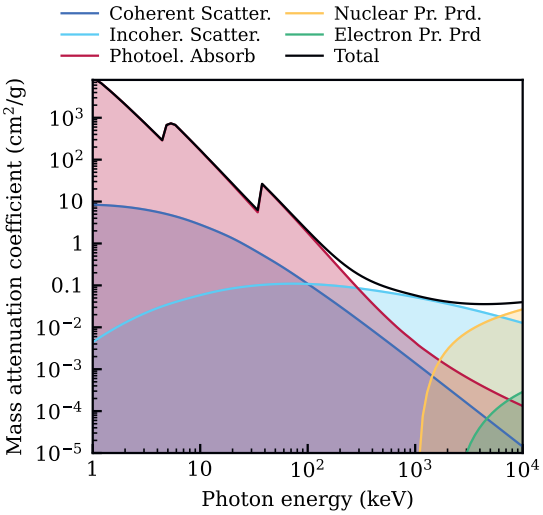


Figure 4.6 *Mass attenuation coefficients for γ rays in xenon as a function of photon energy, based on tabulated data from the NIST XCOM database [97]. Shown are the contributions from coherent (Rayleigh) scattering, incoherent (Compton) scattering, photoelectric absorption, pair production in the nuclear or electronic Coulomb field, and the total attenuation.*

becomes comparable to and then larger than the photoelectric cross section. Instead of depositing its full energy at a single site, a γ ray can undergo multiple Compton scatters before being fully absorbed. This produces spatially separated energy deposits that may still merge into a single S1 and S2 signal in the detector.

In addition, for energetic electrons, bremsstrahlung becomes relevant. Figure 4.7 shows the radiative energy-loss fraction from NIST ESTAR [86], and the corresponding expected number of bremsstrahlung photons above 50 keV emitted along the full electron track, approximated by the continuous slowing-down approximation (CSDA) range. The mean photon multiplicity shown in Fig.4.7 is estimated by integrating the standard bremsstrahlung spectrum in the complete-screening approximation [37] over an effective thickness $d/X_0 \simeq R_{\mathrm{CSDA}}/X_0$, where R_{CSDA} is the ESTAR continuous-slowing-down range.

This provides an order-of-magnitude estimate of the number of photons emitted along the full electron track. Above the MeV scale, bremsstrahlung photons can carry energy away from the primary electron track and create additional, spatially separated deposits.

In the previous simulation framework, Geant4 energy deposits are grouped using a clustering algorithm based on spatial and temporal prox-

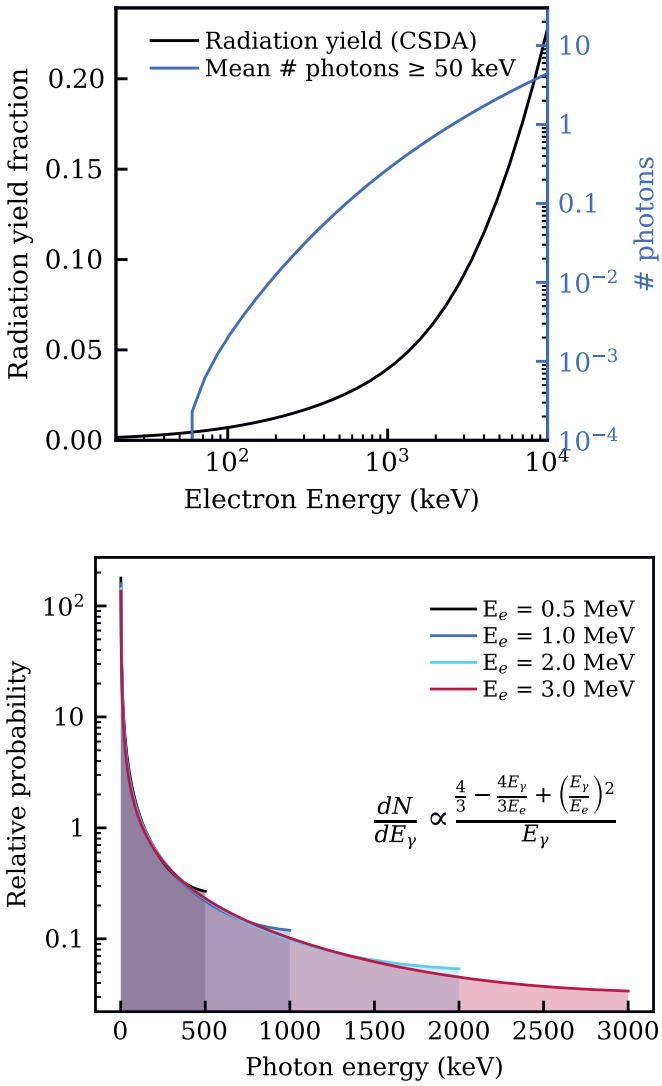


Figure 4.7 *Bremsstrahlung relevance for electrons in liquid xenon. Top: Radiation yield fraction (black) from the NIST ESTAR database [86] and the estimated mean number of bremsstrahlung photons with energy $k \geq 50\,\mathrm{keV}$ (blue) produced along the full CSDA track, as a function of the electron energy. Bottom: Bremsstrahlung photon energy spectrum for fixed electron energies of 0.5, 1, 2, and 3 MeV, computed using the Bethe–Heitler spectrum in the complete-screening approximation, $\mathrm{d}N/\mathrm{d}E_\gamma \propto (4/3 - 4E_\gamma/3E_e + (E_\gamma/E_e)^2)/E_\gamma$ [37], which describes the photon energy distribution emitted by a relativistic electron of energy E_e.*

imity. While adequate at low energies, this approach becomes ambiguous in the high-energy regime. Small changes in clustering parameters can alter the number of clusters and their assigned energies, and because yields are energy-dependent, this directly affects the predicted mean S1/S2 partition.

This inconsistency between how yields are extracted from data and how they are applied in simulations is the main source of the large data-simulation discrepancies observed at high energies.

We next determine ER yields from single-site ^{214}Bi beta-decays and then develop a clustering strategy that applies these yields consistently in multi-site high-energy events.

4.3 Detector response modelling with single-beta ^{214}Bi

To determine the light and charge yields directly from data, it is essential to use a dataset in which the deposited energy can be associated with a single physical cluster. Single beta-decays provide such a topology: the electron track is localised and is usually well described by a single cluster. This property makes beta-decays well suited for extracting intrinsic ER yields that can later be applied, cluster by cluster, in simulations of more complex interaction topologies.

For energies below approximately 20 keV, single beta-decays of ^{212}Pb from ^{220}Rn calibration data have been widely used in XENONnT and other experiments to extract ER yields [66]. Up to about 238 keV, the spectrum is dominated by single beta-decays. Above this energy, however, decays to excited states become significant and are accompanied by prompt gamma-ray emission, making the sample contaminated by multi-site topologies. As a result, there is no existing calibration dataset that provides clean single beta-decays at energies above $\sim$ 200 keV.

To overcome this limitation, in this work we exploit a high-purity sample of ^{214}Bi beta-decays selected via the BiPo delayed-coincidence technique, as described in detail in Chapter 5. The ^{214}Bi decay has a Q-value of 3.27 MeV, providing continuous ER coverage well into the high-energy regime. Approximately 19% of these decays proceed directly to the ground state of ^{214}Po and therefore emit no prompt gamma-ray radiation. In Chapter 5, we describe how these ground-state branches are isolated to obtain a clean sample of effectively single-site beta-decay interactions. In combination with ^{212}Pb beta-decays at lower energies, this dataset spans a wide ER energy range and enables a direct, data-driven determination of the charge and light yields at the XENONnT drift field.

The yield extraction proceeds in two steps. The number of extracted electrons is obtained from the corrected S2 signal as

$$N_e = \frac{\text{cS2}}{g_2}, \tag{4.14}$$

where g_2 is the detector gain parameter defined in Eq. (2.1). The charge yield $Q_y(E)$ is then obtained using the definition in Eq. (4.4). The empirical NEST model introduced in Section 4.1 is fitted to the resulting $Q_y(E)$ data points to determine the best-fit parameters for XENONnT.

Rather than fitting the light yield independently, we derive it from energy conservation as $L_y = 1/W - Q_y$. This enforces the constraint that the total number of quanta remains $N_q = E/W$, with a constant W, as assumed in the simulation framework. In the detector, light and charge are measured independently, and small mismodellings or calibration biases could in principle distort their sum if both yields were fitted separately. By fitting only Q_y and deriving L_y from the sum rule, we avoid introducing an artificial energy-dependent bias in the total quanta used in simulations. The light yield measured directly from data is nevertheless consistent with this constraint within uncertainties, confirming that no significant violation of the $1/W$ relation is observed.

A further complication is that the selected ^{214}Bi sample is not composed exclusively of pure single beta-decays. As detailed in Chapter 5, a controlled and well-modelled fraction of decays that proceed through excited states of ^{214}Po leak into the selection. These transitions are accompanied by prompt gamma-ray emission. In such cases the maximum beta energy is lower than in the ground-state transition, but the emitted gamma-ray can deposit part of its energy near the interaction site, causing the signals to merge. The reconstructed energy then corresponds to the combined beta and gamma deposition. Although such events satisfy the single-site selection criteria, they are therefore not true single-cluster beta-decays. Because the total deposited energy corresponds to the combined beta and gamma emission, these events predominantly populate the high-energy region of the spectrum.

The size of this effect is evaluated using simulations. With Geant4 truth information, simulated ^{214}Bi events passing the single beta-decay selection can be separated into genuine ground-state beta-decays and decays proceeding through excited states with accompanying gamma-ray emission. Figure 4.8 compares the reconstructed charge yield obtained for these two categories. While the yield extracted from true ground-state decays follows the input NEST model used in the simulation, the yield obtained from the full selected sample is biased at high energies due to unresolved gamma-ray contributions. This bias becomes relevant above approximately 1 MeV.

The difference between the two simulated selections provides a direct estimate of the bias introduced by excited-state contamination. We therefore

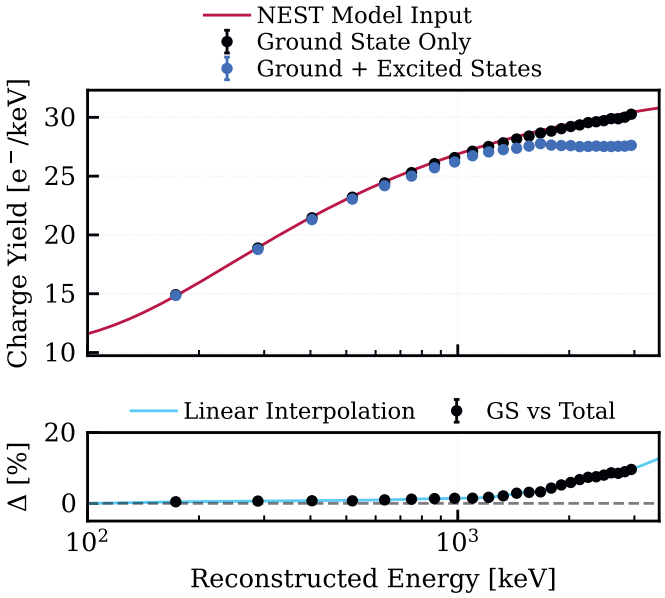


Figure 4.8 *Impact of excited-state contamination on the reconstructed charge yield in simulated ^{214}Bi events passing the single beta-decay selection. Top: Reconstructed charge yield as a function of energy. The black points show the yield obtained from true ground-state beta-decays only, while the blue points show the yield obtained from the full selected sample, including excited-state decays that leak into the selection. The red curve indicates the input NEST yield model used in the simulation, for reference. The deviation between the two simulated selections reflects the bias introduced by unresolved gamma-ray contributions from excited-state transitions. Bottom: Relative difference between the yields extracted from the full selection and from true ground-state decays. The light-blue curve shows the linear interpolation used to derive the energy-dependent correction applied to the data. The effect becomes significant only above approximately* 1 MeV.

derive an energy-dependent correction by comparing the yield extracted from all selected events with that obtained from true ground-state decays only. This correction removes the shift induced by unresolved gamma-ray energy deposits and is applied to the measured $Q_y(E)$ values before fitting the yield model.

The resulting charge-yield measurements are shown in Figure 4.9. The black points show the yields measured directly from the data, while the blue points show the values after applying the correction derived from simulations. The corrected data are then used to determine the parameters of the empirical NEST yield model introduced in Section 4.1. Specifically,

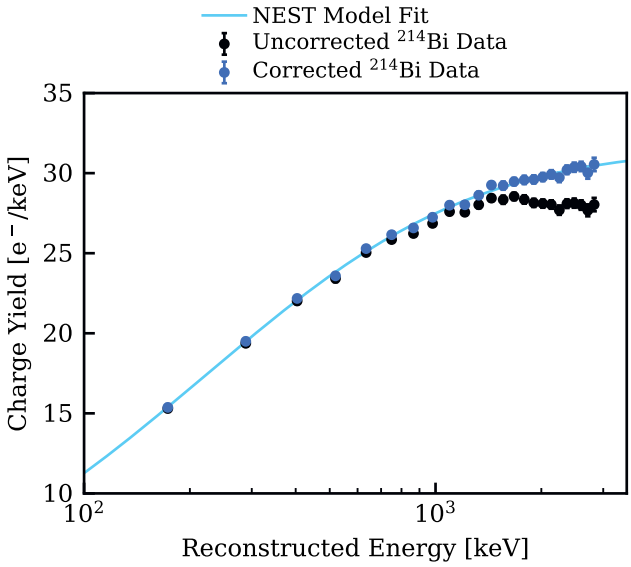


Figure 4.9 *Charge yield $Q_y(E)$ extracted from ^{214}Bi BiPo data as a function of reconstructed energy. Black points show the yield measured directly from the selected data sample, while blue points show the same data after applying the correction derived from simulations to account for contamination from excited-state decays. The corrected points are used to fit the NEST ER yield model. The resulting best-fit model is shown as the light-blue curve.*

the parametrisation given in Eq. (4.13) is fitted to the corrected $Q_y(E)$ data points to obtain the model describing the ER response in XENONnT.

The resulting best-fit model provides a consistent description of the charge yield over the full energy range of the ^{214}Bi spectrum. A detailed discussion of the fit quality, systematic uncertainties (including the treatment of energy bias), and comparison with previous measurements is provided in Chapter 5. In the remainder of this chapter, these yields are used as fixed inputs for the validation of the simulation framework and for the study of clustering effects in high-energy ER events.

4.4 Physics-Based Lineage Clustering

High-energy ER events are often multi-site, and yield models are non-linear in energy. Once the ER yields are fixed, the remaining question is how to define the clusters to which the yield model is applied. The predicted S1/S2 partition can change when the same total energy is distributed over multiple

clusters. We therefore introduce a topology-aware clustering method that uses the Geant4 particle history to define physically motivated clusters.

4.4.1 *Principle: Clustering by Particle Lineage*

We introduce a physics-based clustering method, referred to as lineage clustering. The idea is to group Geant4 energy deposits according to the particle lineage, rather than only according to geometric proximity.

For each energy deposit, Geant4 provides the particle identity (track ID), the parent track ID, the creation process, the energy-deposit process, and the space–time coordinates of the step. These quantities allow the reconstruction of the particle lineage. The lineage represents the sequence of interactions through which primary and secondary particles are produced and deposit energy in the detector. In lineage clustering, a lineage is defined as a connected sequence of energy deposits that is treated as a single unit for yield evaluation. Starting from the first interaction of a particle (or from a newly created secondary), subsequent deposits are assigned to the same lineage unless a lineage break condition is met.

A key design choice is that we do not attempt to construct separate yield models for beta decays and gamma-rays. Instead, the yields used in this work are derived from single-beta data and are treated as a unified electronic-recoil (ER) yield model, distinguishing only between ER and nuclear-recoil (NR) interactions. This choice is physically motivated: gamma-rays ultimately deposit their energy through the production of energetic electrons via processes such as Compton scattering and photoelectric absorption. At the microscopic level, both beta decays and gamma interactions therefore produce electron tracks, and the per-cluster detector response must be identical. The observed difference between gamma-ray and beta-decay detector response at high energy then arises naturally from event topology. Gamma rays typically produce several spatially separated electron tracks and are therefore split into multiple lineages (multiple yield evaluations at lower energies), whereas a beta-decay track usually remains within a single lineage and is treated as one energy deposit for yield evaluation.

4.4.2 *Lineage Break Rules*

The lineage break rules are constructed to reflect two physical observations: (i) β tracks are locally dense and, at the scale relevant for yield assignment, can be treated as a single cluster unless emitted radiation carries energy away; and (ii) γ interactions represent distinct scattering sites and should generally be treated as separate clusters.

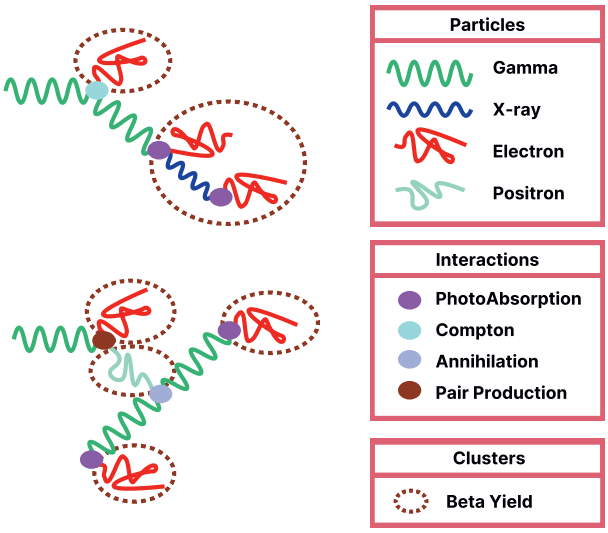


Figure 4.10 *Schematic illustration of the lineage clustering concept for two examples of high-energy γ interactions. Energy deposits are grouped into lineages following the particle lineage. In the first example, a Compton scatter followed by photoabsorption produces two lineages. In the second, pair production followed by two annihilation photons produces three lineages. Local secondary deposits associated with a single site remain grouped within the same lineage. This physics-based clustering stabilizes the application of energy-dependent yield models in multi-site events.*

Concretely, the implementation follows the rules below.

- **Start new lineages at decay vertices and for new particles.** Particles produced by radioactive decay start new lineages. In addition, if a particle appears that has not been assigned to any lineage yet, a new lineage is started unless it is explicitly merged with its parent according to the rules below.
- **Break at changes of nuclear parent (ions).** If the parent particle is an ion (nuclear track), the lineage is broken. This prevents deposits belonging to different nuclear products from being merged and mirrors the fact that different decay products correspond to distinct physical objects. Excited nuclear-state bookkeeping is treated separately to avoid artificial splitting of internal conversion chains.
- **Break γ rays into separate lineages at scattering sites.** For γ rays, the lineage is typically broken whenever the spatial separation between deposits exceeds a configurable threshold. The default configuration uses a zero threshold, corresponding to an always-break behaviour that treats successive γ interactions as separate clusters.

- **Keep photoabsorption + x-rays together.** The photoelectric interaction and the subsequent local de-excitation photons (x-rays/Auger-like deposits) are kept within the same lineage to avoid splitting what is effectively a single site.
- **Electrons make one track.** All deposits from a single electron emitted in beta-decays are kept within the same lineage, reflecting their localised energy loss.
- **Treat bremsstrahlung as a secondary.** When an electron produces a bremsstrahlung photon, the emitted photon can either be merged with the electron lineage if it is produced essentially locally, or split into a new lineage if it travels away. This behaviour is controlled with an independent distance threshold for bremsstrahlung.
- **Break on large time separations.** If the time separation between a deposit and its parent exceeds a configurable threshold, the lineage is broken. This protects against unphysical merging across delayed physical processes.

The thresholds used by lineage clustering are configurable and include: (i) a γ-distance threshold, (ii) a bremsstrahlung-distance threshold, and (iii) a time threshold. Every cluster will get assigned the same type of yield model, corresponding to the extracted parametrisation from single-beta-decays.

4.4.3 *Illustration and Impact*

Figure 4.10 illustrates the qualitative behaviour of lineage clustering. Compared to a purely geometrical clustering, the lineage-based approach produces clusters that follow the particle lineage of the interaction: distinct Compton scatters are separated, while local secondary deposits associated with a single scatter remain grouped.

The practical consequence is that a single, data-driven ER yield model derived from ^{214}Bi beta-decays can be used consistently for both β and γ sources. Differences in the observed S1/S2 partition between γ and β interactions are reproduced because γ events are naturally decomposed into multiple lower-energy lineages, shifting the effective yield sampling to lower energies. In the next section we validate this approach with multiple calibration sources and show that the combination of data-driven yields and lineage clustering substantially improves data–simulation agreement at high energy.

4.5 validation with multiple sources

This section tests the simulation framework against datasets with increasing topological complexity. The ER yields used in this chapter are constrained only by single-β events (Section 4.3), while the validation samples below are dominated by multi-site γ interactions and mixed $\beta + \gamma$ decays. Agreement in these cases therefore probes whether the clustering and yield application remain consistent beyond the extraction topology.

In the following, we present four validation cases of increasing complexity. In all figures, data are shown in black. The updated simulation framework using data-driven yields and lineage clustering is shown in blue (the New MC), and the previous simulation framework using default NEST yields and time-position clustering is shown in red (the Old MC).

For ease of comparison with other experiments, the detector response is expressed in terms of the reconstructed number of photons n_γ and electrons n_e, obtained from the corrected signals as $n_\gamma = cS1/g_1$ and $n_e = cS2/g_2$.

4.5.1 ^{214}*Bi Ground-State Decays*

We begin with the simplest case: the single-beta ^{214}Bi sample selected with the BiPo tag and used to extract the ER yields (Section 4.3). This dataset is dominated by localised electron tracks that are well described by a single cluster. It therefore provides an internal consistency check of the yield extraction and of the assumption that high-energy beta-decays can be treated as single-cluster interactions.

The top-left panel of Figure 4.11 shows the distribution of events in the (n_γ, n_e) plane. The New MC reproduces both the mean detector response and the observed spread of the data. This agreement confirms that the extracted yields correctly describe the single-site electronic recoil response and that the simulation accurately models the topology of high-energy beta-decays. In particular, it supports the assumption that the ^{214}Bi beta-decays used for the yield extraction behave as effectively single-cluster interactions.

4.5.2 ^{214}*Bi with Excited-State Decays*

We next consider ^{214}Bi decays proceeding through excited states, in which the beta-decay is accompanied by one or more prompt γ rays. These events exhibit more complex detector signatures: while S1 signals are typically merged, S2 signals can partially separate depending on the spatial distribution of the energy deposits. This leads to observed distortions in the (n_γ, n_e) distribution.

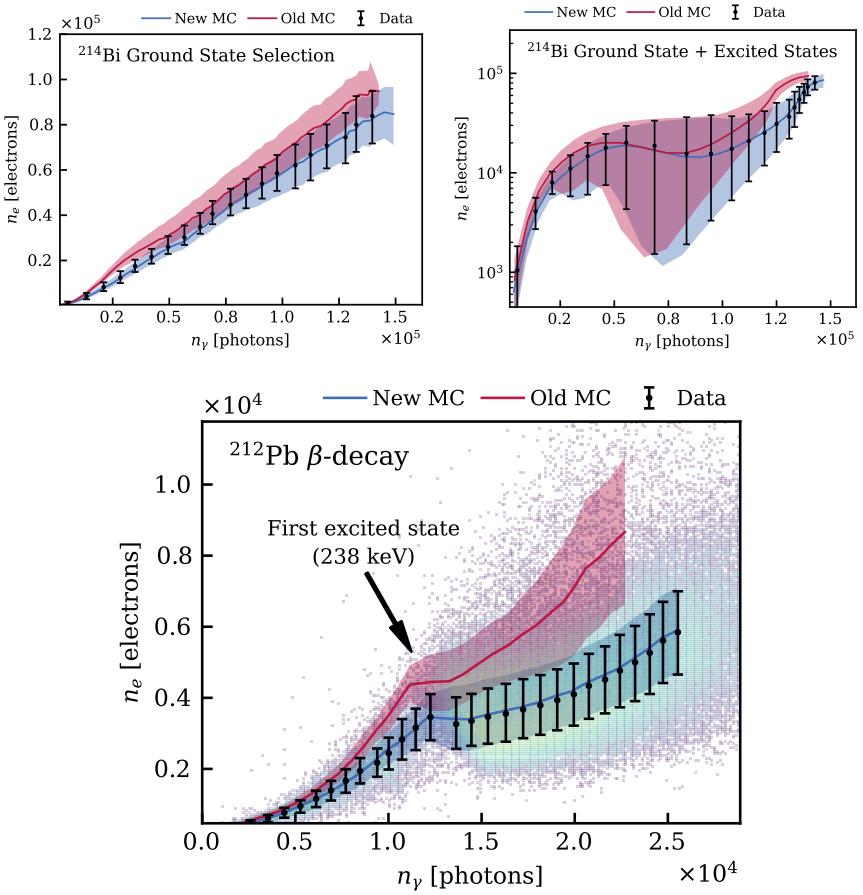


Figure 4.11 *Comparison of data and simulations in the (n_γ, n_e) plane for beta-decay-dominated samples. Black points show the mean response measured in data; the error bars indicate the 1σ width of the distributions. Blue bands show the New MC and red bands show the Old MC. Top-left: ^{214}Bi decays passing the ground state selection. Top-right: the full ^{214}Bi sample, including decays through excited states with accompanying gamma-ray emission, which produce visibly broader and distorted distributions. Bottom: ^{212}Pb beta-decays from ^{220}Rn calibration data. The kink at 238 keV marks the onset of the first excited state; below this energy the spectrum is dominated by single beta-decays, while above it mixed beta-decay and gamma-ray topologies contribute.*

The corresponding panel in Figure 4.11 shows that the New MC reproduces the characteristic shape observed in data, capturing both the mean response and the width of the distribution. In contrast, the Old MC exhibits clear systematic deviations. This agreement is non-trivial: the yield model is derived exclusively from pure single beta-decays, and the description of these mixed β+γ events relies entirely on the lineage clustering to decompose the topology into physically meaningful clusters.

4.5.3 ^{212}*Pb Beta and Gamma Decays*

A particularly instructive validation sample is provided by ^{212}Pb decays from ^{220}Rn calibration data. Up to approximately 238 keV, the spectrum is dominated by single beta-decays, while above this energy the contribution from excited states and associated γ emission becomes significant. This produces a clear transition in topology within a single dataset.

The third panel of Figure 4.11 shows the (n_γ, n_e) distribution for ^{212}Pb data, together with the mean and 1σ bands from simulations. The characteristic "kink" at the onset of gamma emission is clearly visible in the data and is reproduced by the New MC both in position and in width. This result demonstrates that the New MC correctly describes not only pure beta interactions, but also the smooth transition from single-site to mixed topologies within the same physical process.

4.5.4 ^{232}*Th High-Energy Gamma Rays*

The most stringent validation is provided by ^{232}Th calibration data. ^{232}Th is deployed as an external calibration source, and the recorded events are therefore dominated by high-energy gamma-rays entering the detector from outside. These gamma-rays have energies up to 2.6 MeV and typically undergo multiple Compton scatters before being fully absorbed, producing strongly multi-site topologies. Such events probe precisely the regime in which the Old MC was known to perform poorly.

We note that the simulation used here includes a modest retuning of the recombination fluctuation model at high energies. The latter adjustment was necessary to reproduce the observed width of the ER band and is discussed in detail in the next section.

Figure 4.12 shows a grid of energy slices from 200 keV up to 2.6 MeV. For each slice, the distribution of the charge-to-light ratio (n_e / n_γ) is shown. Across the full energy range, the New MC reproduces both the mean and the spread of the data distributions well. The Old MC captures the qualitative behaviour of the distributions but shows systematic deviations in both the mean response and the width.

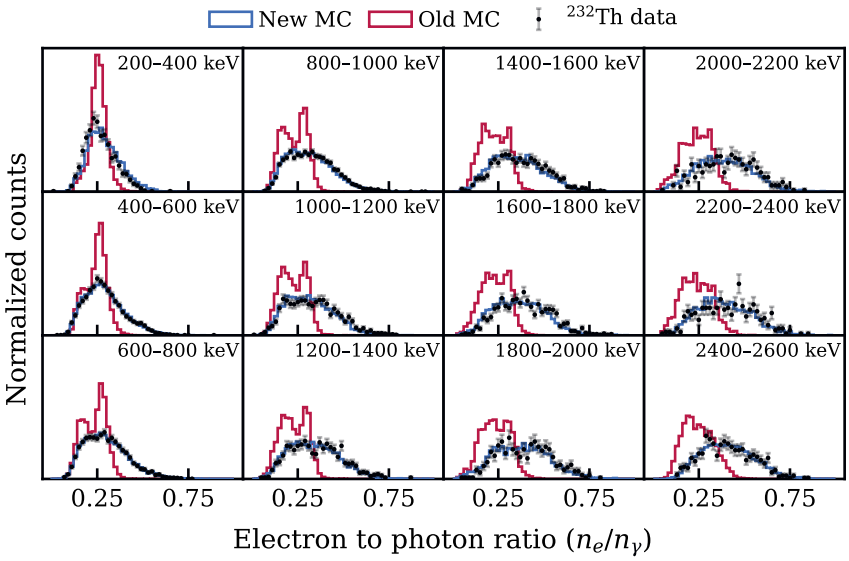


Figure 4.12 *External ^{232}Th calibration data dominated by high-energy gamma-ray interactions. For different energy slices between* 200 keV *and* 2.6 MeV*, the distribution of the charge-to-light ratio is shown. Black points correspond to data, blue curves to the New MC and red curves to the Old MC.*

The agreement obtained with the New MC is notable given the complexity of the underlying topologies. In these events, energy is distributed over multiple spatially separated Compton scatters, and the detector response arises from the combined effect of several yield evaluations at different energies. The results demonstrate that a single ER yield model derived from beta-decays, when combined with topology-aware clustering, can reproduce the response to high-energy gamma interactions without introducing a separate gamma yield parameterisation.

4.6 Limitations and Possible Applications

4.6.1 *The Recombination Fluctuation Model*

In Section 4.5 we showed that the combination of data-driven ER yields and lineage-based clustering reproduces both the mean response and the observed width of the S1 and S2 distributions across a broad energy range, including the external ^{232}Th calibration. Achieving this agreement, however,

required a modification of the recombination fluctuation parameters used in the NEST framework.

In this work, the functional form of the ER recombination fluctuation model was left unchanged. The model is described in Section 4.1 and by Eqs. (4.10–4.12). In brief, the variance is written as the sum of a binomial term and a non-binomial contribution proportional to N_i^2, with a width parameter ω that is parameterised as a function of the electron fraction f_e rather than directly as a function of energy.

Because $f_e(E)$ is fixed once the mean ER yields are specified, this prescription induces an implicit energy dependence $\omega(E) = \omega(f_e(E))$. The recombination fluctuations model therefore does not introduce an independent energy-dependent degree of freedom: its behaviour as a function of energy is fully determined by the yield model.

Figure 4.13 illustrates the impact of this modelling choice. The left panel shows the mean electron fraction $f_e(E)$ at the XENONnT drift field. Using the parameterisation $\omega(f_e)$ described in Eqs. (4.10–4.12), this electron fraction is mapped to an effective energy-dependent width parameter $\omega(E)$. The curves compare the default NEST parameter set with the modified parameters introduced in this work. The modification broadens and enhances ω at low values of electron fractions, which translates into larger recombination fluctuations at high energy. The right panel of Figure 4.13 compares the resulting variance of the charge-to-light ratio in data with simulations using the default and modified recombination fluctuations parameters. While the default NEST parameters describe the observed variance well at low energies, they systematically underestimate the spread at high energy. The modified parameter set reproduces the high-energy behaviour accurately, at the cost of an overestimation of the variance below approximately 80 keV.

Using the current NEST recombination fluctuation model, it was not possible to identify a single parameter set that simultaneously describes both regimes. This behaviour reflects an intrinsic limitation of the semi-empirical recombination fluctuations model, rather than of the yield extraction or clustering strategy presented in this chapter. In particular, it suggests that recombination fluctuations at high energy are influenced by additional physical effects, such as energy-dependent track structure, extended ionisation topology or correlations along long electron tracks. These effects are not fully captured by a model in which ω depends only on the electron fraction [75].

For the high-energy ER applications targeted in this work, the modified recombination fluctuations parameters provide a substantially improved description of the data and are therefore adopted in the simulations presented in this chapter. The observed tension between low- and high-energy

regimes points to the need for a more flexible treatment of recombination fluctuations. Recent ongoing work in the NEST collaboration is focused on implementing an explicit energy dependence in the ω parameter, which would allow a more accurate description of the full energy range. This feature is expected to be available in future NEST releases and will be adopted in XENONnT simulations as soon as it becomes available.

4.6.2 *Conversion-Electron Dominated Transitions*

Despite the substantial improvement achieved for high-energy beta and gamma interactions, the current combination of data-driven ER yields and lineage clustering does not fully reproduce the detector response for several mono-energetic sources. In particular, noticeable discrepancies remain for the ^{83m}Kr calibration and for the ^{129m}Xe and ^{131m}Xe signals from xenon activation, which are commonly used for ER calibration in the low-energy regime. In all these cases, simulations systematically underestimate the charge yield (or equivalently overestimate recombination) compared to data.

These transitions are characterised by a dominant contribution from internal conversion (IC), rather than from pure γ emission. Internal conversion is a nuclear de-excitation process in which an excited nucleus transfers its transition energy directly to one of the atomic electrons via electromagnetic coupling. The electron is ejected from the atom with a discrete kinetic energy equal to the nuclear transition energy minus the electron binding energy.

Internal conversion is possible whenever a gamma transition is allowed (except in fully ionised atoms) and competes directly with gamma-ray emission. The relative probability of IC versus gamma-ray emission is quantified by the internal conversion coefficient $\alpha = e/\gamma$, which increases rapidly for low transition energies and high atomic number [19]. As a result, many low-energy nuclear transitions in xenon and krypton decay predominantly via internal conversion rather than via gamma emission.

Following the emission of the conversion electron, the atom is left with a vacancy in an inner electron shell. The subsequent atomic relaxation proceeds through a cascade of characteristic X-rays and Auger electrons. From the detector perspective, the deposited energy therefore consists of a dense, highly localised cascade of low-energy electrons and photons originating from atomic de-excitation, rather than from a single gamma-ray interaction.

This feature distinguishes IC-dominated transitions from the datasets used to extract and validate the ER yield model in this chapter. Table 4.1 summarizes the dominant decay channels for several isotopes for which

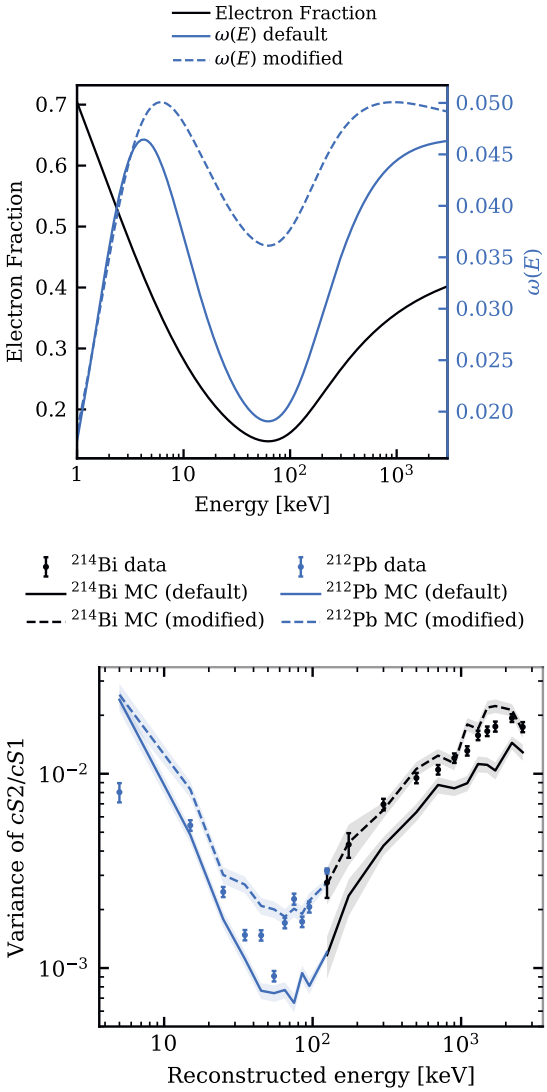


Figure 4.13 *Left: Electron fraction derived from the extracted ER yields (black) together with the corresponding $\omega(E)$ function used in the recombination fluctuation model. The solid blue curve shows the default NEST parametrisation, while the dashed blue curve shows the modified parametrisation adopted in this work. Right: Variance of the charge-to-light ratio as a function of energy measured in data and predicted by simulations. Black points correspond to ^{214}Bi data and blue points to ^{212}Pb data; the error bars represent the 1σ width of the distributions. Simulation predictions are shown as bands, with the solid curve corresponding to the default configuration and the dashed curve to the simulation using the modified $\omega(E)$ parameters.*

Table 4.1 *Summary of dominant decay channels for selected ER calibration and background isotopes. The total internal conversion coefficient α_T quantifies the ratio of emitted electrons to photons. Data from Nuclear Data Sheets [40, 98–101].*

Isotope	Transition Energy [keV]	α_T
^{83m}Kr	32	~ 2000
	9	~ 20
^{129m}Xe	196	~ 20
	39	~ 12
^{131m}Xe	164	~ 50
^{212}Pb	238	~ 0.87
^{214}Bi	609	~ 0.02

data are available in XENONnT. While ^{214}Bi decays are dominated by high-energy γ emission with negligible internal conversion, isotopes such as ^{83m}Kr, ^{129m}Xe, and ^{131m}Xe exhibit large internal conversion coefficients, implying that a substantial fraction of decays deposit their energy through conversion electrons and atomic cascades.

In the current simulation framework, conversion electrons are treated as ordinary ER tracks using the yield model derived from single beta-decays. While this is adequate for extended electron tracks, it may not hold for IC-dominated transitions, where the conversion electron is accompanied by a localised shower of Auger electrons and characteristic X-rays from atomic shell refilling, producing a local ionisation density that differs from a simple beta track.

This limitation is illustrated in Figure 4.14, which compares data and simulations for the emissions of neutron-activated xenon ^{129m}Xe and ^{131m}Xe, as well as for ^{83m}Kr from calibrations, in the $(\mathrm{cS1}, \mathrm{cS2})$ plane. In all cases, simulations slightly underestimate the charge signal (or equivalently overestimate recombination), resulting in a lower cS2 at fixed cS1 compared to data.

We verified that this discrepancy can be mitigated by introducing an ad hoc modification to the yield model, assigning a reduced charge yield specifically to lineages dominated by conversion electrons. With such a prescription, both the xenon activation lines and the ^{83m}Kr response can be reproduced more accurately. However, implementing a conversion-electron yield model would introduce additional empirical tuning without

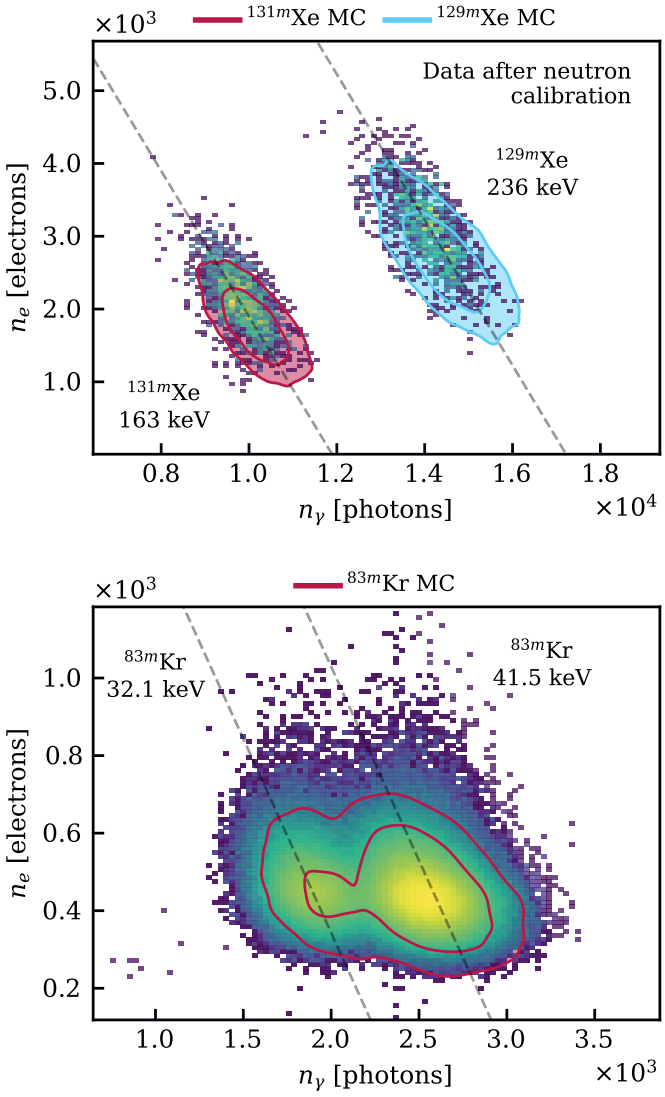


Figure 4.14 *Comparison of data and simulations for conversion-electron–dominated mono-energetic emissions. Left:* ^{129m}*Xe and* ^{131m}*Xe activation lines. Right:* ^{83m}*Kr, where the two visible clusters correspond to the 9 keV and 32 keV emission lines. Simulations reproduce the general topology but systematically underestimate the charge yield, indicating excess recombination relative to data.*

a clear microscopic justification and would weaken the goal of maintaining a unified and minimally tuned ER response description.

For this reason, no special treatment for internal conversion electrons is included in the final simulation framework presented in this chapter. Instead, the observed discrepancies are interpreted as an indication of residual limitations in the semi-empirical modelling of ER microphysics. Similar challenges have been reported for electron-capture and double-electron-capture processes in other experiments, including recent studies by XENONnT [45] and in particular by LZ [102], where deviations from standard ER yield models are observed.

In summary, while the framework developed in this work provides an accurate and predictive description of high-energy β and γ interactions, it does not fully capture the detector response for IC-dominated mono-energetic transitions. This reflects the intrinsic complexity of ER microphysics in liquid xenon and points to the need for future developments that explicitly account for atomic de-excitation cascades and track-structure effects.

4.6.3 *Implications for Double-Beta-Decay Studies*

An additional and important consequence of the detector-response modelling developed in this chapter is its applicability to double-beta-decay studies. While the primary focus of this work is to accurately simulate high-energy ER signals from single β and γ interactions, the same ingredients also enable a predictive description of events in which the deposited energy is shared between two electrons.

In liquid xenon, the detector response to two-electron emission is not, in general, identical to that of a single beta-decay depositing the same total energy. The partition between light and charge depends on how the energy is distributed among the electrons and on the resulting track structure, which affects recombination. With a data-driven single-β yield model and a clustering strategy that preserves the causal structure of energy deposits, these effects can be modelled without introducing an explicit, ad hoc double-β yield parameterisation.

The framework developed in this chapter predicts the mean detector response of two-electron events starting from the single-beta response, and has been applied in double-beta-decay studies within the XENON collaboration. The observed shift of the $2\nu\beta\beta$ band relative to the single-beta band in XENON1T and XENONnT data is reproduced, and the response converges toward that of other multi-site ER interactions at higher energies.

Beyond predicting the mean response, this framework also provides sensitivity to differences between multi-electron and gamma-induced ER

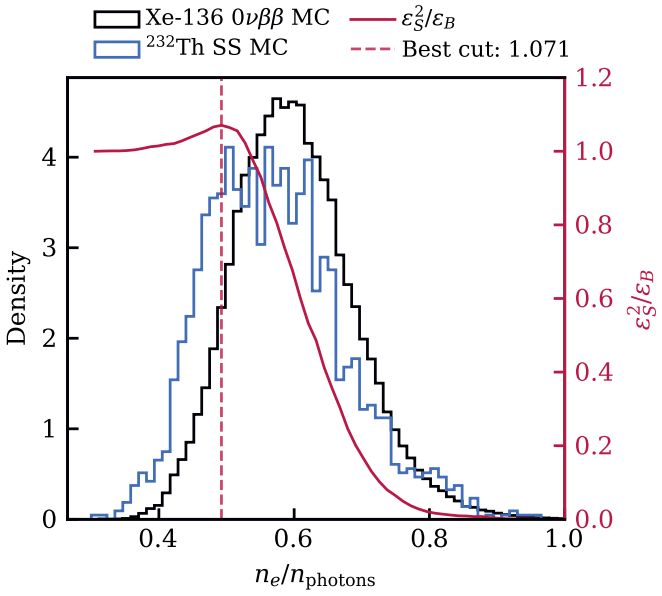


Figure 4.15 *Predicted separation between neutrinoless double-beta decay ($0\nu\beta\beta$) events and gamma-ray backgrounds using the data-driven ER yield model and lineage clustering. The distribution of the electron-to-photon ratio n_e/n_γ is shown for simulated $0\nu\beta\beta$ decays of ^{136}Xe (black) and for single-scatter events from simulated external ^{232}Th calibration data representing the dominant gamma-ray background (blue). The red curve shows the figure of merit ϵ_S^2/ϵ_B, where ϵ_S and ϵ_B denote the signal and background efficiencies for a cut on n_e/n_γ. The vertical dashed red line indicates the optimal cut value, corresponding to a 7.1% increase in this sensitivity metric.*

topologies. Figure 4.15 illustrates an example using the ratio of reconstructed electrons to photons, n_e/n_γ. The distributions predicted for simulated ^{136}Xe $0\nu\beta\beta$ events and for single-scatter gamma-ray backgrounds from external ^{232}Th simulations show a small but visible shift.

Using this observable, an optimal selection can be defined by maximising the figure of merit ϵ_S^2/ϵ_B, where ϵ_S and ϵ_B denote the signal and background efficiencies. The resulting discrimination is modest but non-zero, demonstrating that topology-aware detector-response modelling can provide additional information beyond simple energy reconstruction.

The work presented in this chapter therefore establishes the detector-response modelling required for such studies. A full exploration of double-beta-decay phenomenology, including model dependence and sensitivity projections, is beyond the scope of this thesis, but the tools developed here make such investigations possible.

4.7 summary and implications

In this chapter, we developed a framework to accurately model the detector response to high-energy electronic recoils in XENONnT. The main challenge arises from the interplay between energy-dependent light and charge yields and increasingly complex multi-site interaction topologies at high energies, where the previous simulation framework showed substantial disagreement with data.

We addressed this problem by calibrating ER yields directly on XENONnT data using a high-purity sample of single-site ^{214}Bi beta-decays, and by implementing a physics-based lineage clustering algorithm that groups Geant4 energy deposits according to their particle lineage. This combination ensures that energy-dependent yields are applied at a granularity consistent with the underlying interaction topology.

The combination of these two ingredients leads to a substantial improvement in the agreement between data and simulations. Using a single ER yield model derived from beta-decays, together with lineage-based clustering, we accurately reproduce both the mean response and the event-by-event spread of S1 and S2 signals for a wide range of sources and topologies, including mixed β+γ decays and high-energy γ calibrations dominated by multiple Compton scatters. The observed differences between β- and γ-dominated events emerge naturally from topology rather than from the use of separate yield models.

At the same time, this chapter exposes intrinsic limitations of the current semi-empirical modelling. In particular, a tension remains in the description of recombination fluctuations between low- and high-energy regimes, and transitions dominated by internal conversions such as in ^{83m}Kr and the emissions from activated xenon isotopes are not fully reproduced. These effects point to missing microphysical ingredients related to track structure and atomic de-excitation cascades and motivate future developments beyond the scope of this work.

The yield model and clustering strategy developed here are used as the default simulation configuration in the remainder of this thesis. They provide the foundation for precision studies of high-energy ER processes in XENONnT, including the ^{214}Bi energy spectrum measurement presented in the following chapter.

5

MEASUREMENT OF THE ^{214}BI ENERGY SPECTRUM

Published in: Physical Review C (APS)
Publication date: April 2, 2026

Available online at:

- https://doi.org/10.1103/b3r7-6ff4
- https://arxiv.org/abs/2510.04846

Reference [1].

5.1 INTRODUCTION

Rare-event searches require a precise characterization of backgrounds in the signal region of interest. A ubiquitous source arises from radon, which emanates from detector materials containing trace amounts of uranium and thorium. Radon progeny undergoing beta-decay [15] constitute a major background, with ^{214}Pb and ^{214}Bi decays being key contributors in Weakly Interacting Massive Particle (WIMP) dark matter searches [45, 103, 104] and neutrinoless double beta-decay searches [5, 52, 53].

This study focuses on the ^{214}Bi beta-decay spectral shape, whose decay scheme is shown in Fig. 5.1. The isotope ^{214}Bi has a half-life of 19.9 min, a Q-value of 3.27 MeV, and a beta-decay branching ratio of 99.979%, of which 19.2% decay directly into the ground state [105].

The Q-value at $\mathcal{O}$(MeV) promotes the transition into excited states. The subsequent de-excitation leads to complex gamma emission patterns. Theoretically, the multitude of accessible decay modes complicates the evaluation of the beta spectrum, as calculations must account for a broad spectrum of nuclear transitions. Understanding the nuclear structure of isotopes like ^{214}Bi is crucial across various fields. One example is the modeling of the rapid neutron-capture process (r-process) [106], a key mechanism in stellar nucleosynthesis responsible for producing elements heavier than iron.

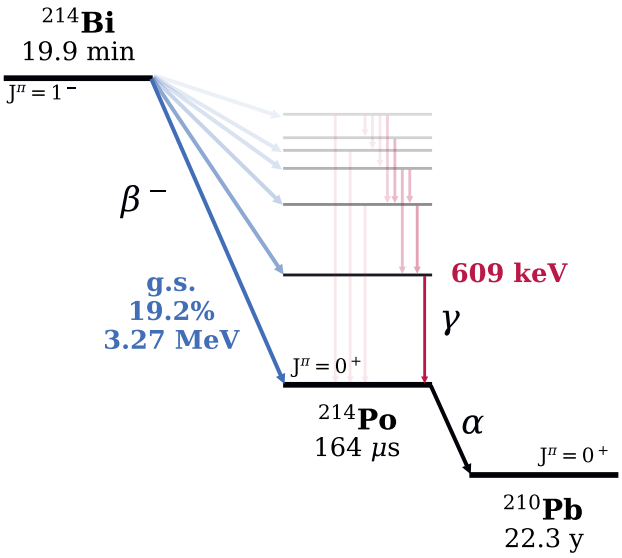


Figure 5.1 *Decay scheme of* ^{214}Bi *to* ^{214}Po *and subsequently to* ^{210}Pb. *The diagram shows beta-decays from* ^{214}Bi *(left) either directly to the ground state (g.s.) or via excited states of* ^{214}Po *(center). Approximately 80 excited nuclear states are accessible during the decay (symbolically represented). All these excited states de-excite via gamma-decays emissions (red arrows), always including at least one transition greater than* 609 keV. *The final step is the alpha-decay of* ^{214}Po *(black arrow) directly to the stable ground state of* ^{210}Pb *(right).*

Experimentally, in a liquid-xenon TPC, beta and gamma interactions typically produce overlapping signals, which prevents separating the two components. However, a well-understood ^{214}Bi beta-spectrum would allow for a continuous detector calibration up to the beta-endpoint at 3.27 MeV, which is not possible with any other calibration sources.

In this study, we report a dedicated analysis in determining the ^{214}Bi ground-state spectrum using data from the XENONnT experiment [43]. Compared to the ^{214}Bi spectrum reconstruction by the NEMO experiment [107, 108], the XENONnT detector has several advantages, including having an ultra-low background and an in-situ ^{222}Rn calibration. As the full ^{222}Rn decay process is contained within the XENONnT detector, the alpha-decays and gamma-decays from the same decay chain are utilized to isolate the ^{214}Bi beta-decays. The selection algorithm developed in this work enables an effectively background-free measurement of the ^{214}Bi spectrum. Particular emphasis is placed on isolating the ground-state ^{214}Bi spectrum, allowing for direct comparisons with theoretical predictions. Furthermore, the XENON collaboration has developed a full-chain simulation framework, incorporating accurate detector response modeling. The simulation was

adapted to develop the selection algorithm, while the ^{214}Bi data serves as a supporting benchmark for the simulation.

5.2 Theoretical framework

Beta-decays are interactions mediated by massive $W^{\pm}$ bosons between hadronic and leptonic currents. Assuming the decaying nucleon does not interact with the other $A-1$ nucleons, the probability density of the beta-decay electron being emitted with total energy E_e can be expressed as [109]:

$$P(E_e)dE_e = \frac{G_{\mathrm{F}}^2}{(\hbar c)^6}\frac{1}{2\pi^3\hbar}C(E_e) \times p_e c E_e (E_0 - E_e)^2 F_0(Z, E_e) dE_e, \tag{5.1}$$

where p_e is the momentum of the electron, E_0 the beta-decay endpoint energy, Z the proton number, G_F the Fermi constant, and $F_0(Z, E_e)$ the Fermi function. The nuclear structure information is encoded by the shape factor $C(E_e)$, which can generally be decomposed into vector, axial-vector, and mixed vector-axial-vector components:

$$C(E_e) = g_V^2 C_V(E_e) + g_A^2 C_A(E_e) + g_V g_A C_{VA}(E_e), \tag{5.2}$$

where g_V and g_A are the corresponding coupling constants for the vector and axial-vector components [15]. The contribution of the shape factors C_V, C_A, and C_{VA} is characterized by the nuclear matrix elements (NMEs).

The theoretical description of the beta-decay has undergone significant advancements in recent years, which have been essential for the background modeling of the XENON experiments [36, 42], addressing the challenges when moving towards heavier nuclei and higher beta-decay Q-values. The complications include incorporating a large number of different transition modes in NMEs [24], modeling strong quenching of the weak coupling constant g_A due to many-nucleon correlations [31] involving numerous approaches [21, 34, 35, 110, 111], and applying the mesonic enhancement correction from the effects of nuclear medium [41].

For the ^{214}Bi beta-decay, the transition to the ^{214}Po ground state proceeds from ^{214}Bi(1^-) to ^{214}Po(0^+), corresponding to a $\Delta J^\pi = 1^-$ change and therefore to a first-forbidden non-unique decay, which requires non-diagonal NMEs for its description [23]. The simplest reference is the zeroth-order spectrum, calculated analytically under the allowed approximation using the BetaShape software module [42], and shown in Fig. 5.2 as *Allowed*.

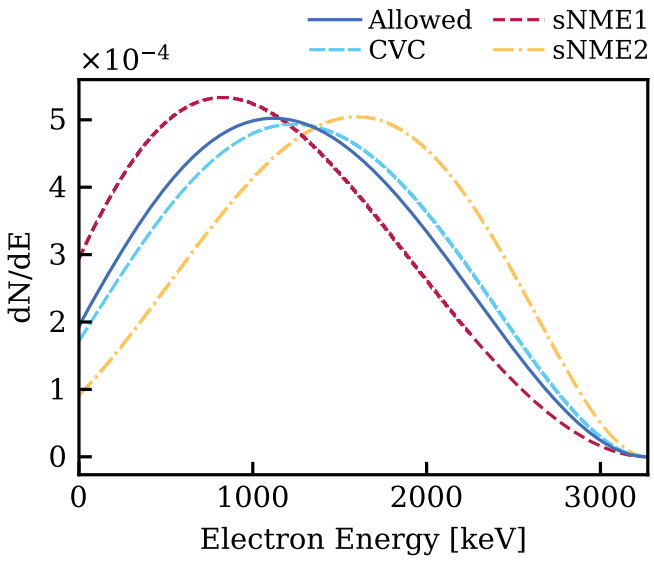


Figure 5.2 *Theoretical beta-decay energy spectra for ground-state to ground-state transitions of* ^{214}Bi *to* ^{214}Po*. The solid blue line shows the Allowed spectrum, the dashed light blue line the CVC-based spectrum, the dashed red line the sNME1 fit, and the dot-dashed yellow line the sNME2 fit.*

A more complete theoretical description is provided in Ref. [24], where the authors perform nuclear shell-model calculations for the atomic mass number $A \approx 214$ region using the KSHELL framework [29] to build the *khpe* Hamiltonian [28] and the NuShellX@MSU library [30] library to extract single-particle energies considering relevant interaction and valence space. This produces the low-lying level schemes of ^{214}Bi and ^{214}Po to within tens of keV, providing a reliable nuclear-structure basis for computing next-to-leading-order beta-decay corrections under the conserved vector current (CVC) hypothesis, which enforces isospin symmetry, fixes the weak vector coupling to $g_V = 1.0$, and constrains the relevant vector form factors [8]. The resulting prediction is shown as the *CVC* spectrum in Fig. 5.2. Although CVC is well supported experimentally [112] and widely used in evaluations of forbidden shape factors [113], small deviations may arise from non-nucleonic effects [22]. The same study also introduced a complementary approach in which the small relativistic vector NMEs (sNMEs) are treated as fit parameters. By varying them simultaneously across all transitions to reproduce the experimental branching ratios, two degenerate solutions are obtained due to a quadratic dependence in the fit. These spectra, labeled

sNME1 and *sNME2*, are also taken from Ref. [24] and shown in Fig. 5.2. Together, the CVC and sNME spectra illustrate how modest variations in the vector NMEs can significantly alter the predicted ground-state beta-decay spectrum, highlighting the need for an experimental determination.

5.3 Experimental Method

5.3.1 *XENONnT Detector and Datasets*

The XENONnT experiment [43] uses xenon as both the target and detection medium inside a cylindrical dual-phase time projection chamber (TPC) that is 1.33 m in diameter and 1.49 m tall. The TPC contains 5.9 t of liquid xenon (LXe), with a thin layer of gaseous xenon (GXe) at the top. An interaction with xenon atoms inside the TPC produces both prompt scintillation light, defined as the S1 signal, and ionization electrons. Meshed electrodes are used to apply an electric drift field in LXe (23 V/cm) and an extraction field in the region of the liquid-gas interface (2.9 kV/cm). This allows the free electrons to drift upward towards the LXe surface, where they are extracted into the GXe and generate secondary scintillation via electroluminescence, defined as the S2 signal. A total of 494 Hamamatsu R11410-21 photomultiplier tubes (PMTs) [60] are arranged at the top and bottom of the TPC to capture both the S1 and S2 signals, measured in units of photoelectrons (PE). A MeV beta-electron in the TPC generates an S1 with a signal area of approximately 10^4 PE and an S2 of order 10^6 PE, while a MeV alpha generates an S1 of several 10^4 PE and an S2 of order 10^4 PE.

An event consists of a pair of S1 and S2 signals. The pairing enables reconstruction of the deposited energy and the three-dimensional interaction position, where the xy-position is derived from the S2 scintillation photon hit pattern on the top PMT array and the z-position from the time difference between S1 and S2. The position reconstruction permits the definition of a fiducial volume, selecting only the data from the interior of the TPC, and thereby avoiding the background events originating from the decay of radioactive contaminants accumulated on the TPC wall. The dual signal also allows the discrimination between interactions with atomic electrons and with nuclei. The former is referred to as electronic recoil (ER) events, and the latter as nuclear recoils (NR) events. In dark matter searches, WIMP signals have the NR signature. In this work, the ^{214}Bi beta-decay signal of interest has the ER signature.

To achieve a minimal background, the XENONnT experiment is operated at the INFN Laboratori Nazionali del Gran Sasso (LNGS) underground laboratory. The cryostat containing the TPC is the innermost of three nested detectors. It is enclosed by a neutron veto [58] that is itself nested within

a muon veto. Both veto detectors are water Cherenkov detectors and are optically separated. All detector construction materials were selected for low radioactivity [64]. The LXe is constantly purified via both gaseous and liquid purification systems [43] that remove electronegative impurities. The noble gas ^{85}Kr beta-decay isotope, which is intrinsically present in xenon, was removed via a dedicated cryogenic distillation system [63] with the remnant measured [114]. For the radon background, a radon removal system also based on cryogenic distillation was installed to continuously remove the radon remnant from the TPC [115]. Finally, degreasing, etching, passivating, and screening were performed to ensure low radon emanation and electronegative impurities for all TPC components [64]. With all these background removal measures, XENONnT has achieved the lowest ever ER background among dark matter detectors [116].

This study includes 95.1 days of data from the first science run (SR0) and 186.5 days from the second (SR1) of the XENONnT experiment. These datasets contain ^{214}Bi decays from natural ^{222}Rn emanation from detector materials at a ^{222}Rn rate of approximately 1.8 µBq/kg of LXe for SR0 and 1.0 µBq/kg for SR1 [2].

To obtain sufficient statistics for this study, the most relevant XENONnT dataset is from a calibration campaign during SR1 using a ^{222}Rn source (SR1 ^{222}Rn). The source was produced by implanting ^{226}Ra into a 2 cm × 2 cm stainless steel plate at the ISOLDE facility at CERN [117], giving an overall emanated ^{222}Rn activity of about 2 Bq. The implantation is sufficiently deep below the surface, such that the ^{226}Ra atoms are mechanically sealed within the sample. Only the noble gas decay daughters ^{222}Rn are ejected from the sample.

During the ^{222}Rn (^{214}Bi) calibration period, the stainless steel sample was installed in-line with the gaseous purification system [43] and was flushed with GXe to carry the ^{222}Rn atoms to the xenon TPC. The time evolution of the source activity is tracked in-situ via ^{222}Rn alpha-decays as shown in Fig. 5.3. Shortly after the source injection, the ^{222}Rn activity increased by two orders of magnitude. Upon closing the source, the ^{222}Rn level decreased by a compound effect of the ^{222}Rn natural decay and the continuous operation of the radon removal system, returning the ^{222}Rn activity to the level observed before the calibration campaign. The expected level of long-term contamination, particularly the beta-decay isotope ^{210}Pb with a half-life of 22.3 yrs that attaches to the TPC surfaces, has been evaluated as negligible.

We analyze the three datasets (SR0, SR1, SR1 ^{222}Rn) with independent treatments to account for their differences in operation and calibration conditions, before combining them into the final spectral data.

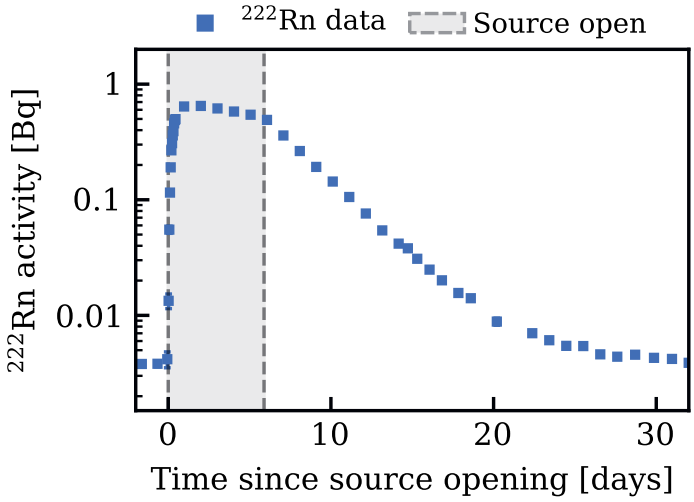


Figure 5.3 *Time evolution of the measured ^{222}Rn activity in the XENONnT TPC fiducial volume during the calibration campaign. The activity is derived from alpha-decays, shown as a function of time since the source opening. The vertical dashed lines indicate the source opening and closing times.*

Finally, to accurately estimate the ER yield model in-situ, we also utilized the XENONnT data from ^{220}Rn (^{212}Pb) calibrations [118].

5.3.2 *Event Topology*

To select ^{214}Bi decays to the ^{214}Po ground state, we exploit the characteristic event topology. The delayed coincidence of a ^{214}Bi beta-decay, followed by the ^{214}Po alpha-decay within the same chain is defined as ^{214}BiPo. Tagging an isolated ^{214}Bi beta-decay is inaccessible because of its broad spectrum and the presence of other ER background components. In contrast, the 7.83 MeV ^{214}Po alpha-decay [105] produces a prominent signal, with S1 signals of order 10^4 PE. The large S1 arises from the high stopping power of alphas, which create dense electron-ion tracks where recombination suppresses the S2 but enhances the S1 [119]. This tagging allows the determination of the alpha rate by isolating and counting the alpha events with negligible background. The derivation of the ^{222}Rn activity in Fig. 5.3, for instance, directly benefits from this tagging scheme. For the ^{214}Po alpha-decays, they are especially well-separated from the nearby alpha peaks by approximately 1 MeV. Moreover, the half-life of ^{214}Po is 164 µs [105], which is relatively short compared to the full TPC drift time of $\sim$2 ms and longer than the typical S2 signal width of $\sim$20 µs. In other words, we can use

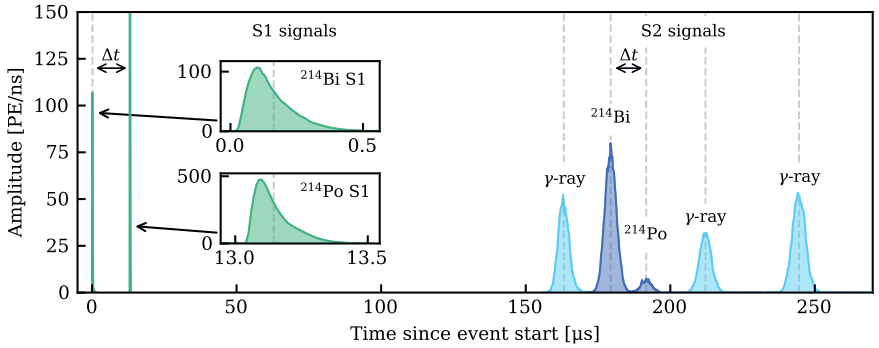


Figure 5.4 *Waveform representation of a typical* ^{214}BiPo *event from measured data. The two green peaks on the left correspond to the S1 signals of* ^{214}Bi *and* ^{214}Po*, separated by a time interval* Δt*. The two dark blue peaks on the right, also separated by* Δt*, are identified as the corresponding S2 signals. The light blue peaks represent S2 signals from gamma emissions. Due to variations in electron drift distance, depending on whether the gamma travels upward or downward, its S2 signal appears earlier or later than the* ^{214}Bi *S2. The alpha-decay (*^{214}Po*) exhibits significant recombination and quenching effects, resulting in a smaller S2 despite its higher energy compared to the preceding beta-decay. Conversely, these effects enhance the alpha's S1 signal. Inset panels show the S1 waveforms on an expanded time scale, highlighting their much shorter duration compared to S2 signals.*

the ^{214}Po alpha to tag the ^{214}Bi decay in a ^{214}BiPo event with effectively no background.

The ^{214}BiPo event topology is illustrated in Fig. 5.4. The event is first triggered by the ^{214}Bi S1 signal and is shortly followed by a significantly larger S1 signal from ^{214}Po after a time difference of Δt. The S2s arrive a drift time later. The ^{214}Po S2 signal is suppressed as expected for an alpha event. The Δt is effectively preserved by the S2s because the alpha signals are localized in position: the ^{214}Po alphas travel on an order of 10 µm in LXe from the preceding ^{214}Bi decay.

A complication arises when considering the ^{214}Bi beta-decay into excited states that subsequently produce gammas from nuclear de-excitation. Given that the de-excitation half-life is in the nanosecond range, the S1s of these gammas are always merged with the S1 of the beta. However, unlike alpha particles, the gamma-rays can travel a few centimeters before interacting with another xenon atom [97]. This shift in positions of the interaction site results in ^{214}BiPo events with more than two S2s, as illustrated in Fig. 5.4. Notably, however, among all the de-excitation gammas of the ^{214}Bi beta-decay as illustrated in Fig. 5.1, there is always one gamma photon with an energy at least 609 keV. We can effectively use these high-energy gamma-decays to tag the excited states.

5.3.3 *Event Selection*

This analysis relies on the ^{214}BiPo *tagging algorithm*, a method that exploits the time coincidence between the ^{214}Bi beta-decay and the subsequent ^{214}Po alpha-decay. Although ^{214}BiPo tagging is commonly employed in low-background experiments [107, 120, 121], our implementation specifically addresses the complexity introduced by gamma emissions accompanying ^{214}Bi decays into excited nuclear states. These emissions produce events with multiple signals, complicating straightforward time-based matching. To reliably pair the correct S1 and S2 signals, our algorithm compares the time differences Δt between all combinatoric pairs formed by the three largest S1 signals (to account for possible spurious signals) and up to ten largest S2 signals. We identify events in which one S1 pair and one S2 pair have matching Δt values within a tolerance of $\pm 3\,\mu s$. Peaks from S2 signals with areas below 1 500 PE, unphysical top-bottom PMT area distributions, or inappropriate timing (such as spurious pulses that occur shortly after the main pulse and can be reconstructed as separate peaks), are excluded from the matching procedure. We use simulations, described in Sec. 5.3.4, to quantify the matching efficiency, which is $> 91\%$ for true ground-state events, providing validation of the matching approach.

Following the matching procedure, we apply a series of *quality selection criteria* to ensure high-quality event reconstruction and minimize contributions from backgrounds or poorly reconstructed signals. Events must pass the fiducial volume constraint, defined by a vertical depth at least 5 cm away from the top and bottom boundaries of the TPC to avoid regions with field inhomogeneities [61]. A minimum separation of 25 µs between the two primary S1 signals is required to prevent potential merging of the corresponding S2 signals, given a typical S2 width at 90% of the area of up to 20 µs. This criterion results in only a modest loss of approximately 3% of events, as shorter S1 intervals typically produce merged, unresolved S2 signals. This requirement also eliminates the background contribution of ^{212}BiPo events originating from the ^{220}Rn decay chain, with a similar topology as the ^{214}BiPo but with a half-life of 300 ns. Additional criteria based on the alpha S1 signal area are employed to remove events originating from other alpha-decays. These selection steps reduce the dataset to approximately 165 000 events, which include both ground-state and excited-state decays.

The final *ground-state selection* step specifically isolates ^{214}Bi decays directly to the ^{214}Po ground state. We require exactly two S2 signals above 1 500 PE, corresponding to the beta-decay and the subsequent alpha-decay. To suppress events with additional gamma interactions, any third S2 is required to be smaller than 10 000 PE. This selection effectively excludes

most excited-state decays, which, due to selection rules, de-excite through high energy gamma-decays leading to clearly identifiable additional S2 signals. Although further criteria, such as a selection on the S2 width to remove partially merged signals, were tested, they provided no significant improvement beyond the initial matching procedure. This selection yields 56 965 events, corresponding to 31% of the events that passed the quality criteria.

Using detailed simulations (described in Sec. 5.3.4), we evaluate the performance and purity of the selection procedure. After applying matching and quality criteria, the requirement of precisely two S2 signals achieves a signal acceptance of around 94% (the fraction of true ground-state events passing the final selection) and a background rejection of approximately 93% (the fraction of excited-state events removed by the selection).

The fraction of excited-state events that pass the ground-state selection criteria (approximately 20% of the total selected events) predominantly belongs to two categories of roughly equal proportions. The first category consists of events in which the emitted gammas completely escape the active detection volume. These events are reconstructed at energies lower than genuine ground-state events, as only the energy deposited by the beta electron is observed. The second category comprises events where gamma-induced signals fully merge with the beta electron signals, resulting in artificially higher reconstructed energies due to the combined gamma and beta contributions. A much smaller additional contribution arises from internal conversion (IC): in these transitions, the gamma-ray is replaced by the emission of an orbital electron, followed by Auger electrons and X-rays from atomic-shell rearrangement. Because this energy is deposited locally, IC events do not provide spatially separated interactions that are characteristic of gamma-rays, and therefore offer limited separation power. The expected IC fraction for the most relevant transitions is at the percent level and is fully accounted for in the GEANT4 simulation. All three contributions are included in the modeling of leakage events used in the comparison between data and simulation.

5.3.4 *Simulations and Signal Modeling*

Simulations for this analysis are produced using the XENONnT framework [68], which incorporates a full-chain simulation from energy deposition to final event reconstruction. Initial interactions, including both the ^{214}Bi and subsequent ^{214}Po decays, are simulated with Geant4 [69, 70]. This step considers all relevant nuclear decay characteristics, including the full decay scheme of ^{214}Bi, associated branching ratios, and emitted gamma energies. To investigate different theoretical models of the ^{214}Bi ground-state

beta-decay spectrum introduced in Fig. 5.2, we adjusted the input beta energy distribution provided to Geant4 accordingly, creating simulation datasets for each considered model.

Subsequent stages of the simulation that include clustering of deposited energies, generation and propagation of photons and electrons using detector-specific efficiency maps, electron drifting, and waveform simulation are performed using the fuse package [68]. These simulated waveforms closely replicate those recorded by the data acquisition system of the XENONnT detector, allowing for the use of an identical processing pipeline as for the real data with the straxen package [74].

For accurate modeling of events with energies up to 3.27 MeV, which is substantially higher than the typical range targeted by the dark matter search analyses (<100 keV) [66], we extended the existing fuse framework with a specialized simulation workflow. Two critical improvements were implemented: a more physically motivated clustering algorithm for energy depositions, and a data-driven yield model specifically tailored for high-energy electron recoils. The latter describes the photon and electron yields per unit energy.

Clustering of energy depositions is necessary because the data-driven yield models, such as the ones provided by NEST [75, 122], require input energies to attribute the correct amount of photons and electrons generated by individual interactions. Due to the incomplete understanding of the underlying microphysics at these scales, energy depositions must be aggregated into physically meaningful clusters before applying the yield model. The default clustering method used in XENONnT [66] groups energy depositions solely based on spatial and temporal proximity. While adequate at lower energies, this method is inadequate in describing complex high-energy event topologies. The new framework employs a lineage clustering method inspired by NEST, which groups interactions based on their physical properties and interaction history. Energy depositions from electron tracks are grouped into single clusters, except for Bremsstrahlung photons, which form separate clusters. Gamma interactions instead initiate new clusters for each distinct interaction site, effectively differentiating gamma from beta events.

Unlike the default NEST model, which employs separate yield models for gamma and beta interactions, our approach uses a single, unified yield model. The apparent difference between the two cases originates from the underlying physics: gamma-rays can interact via Compton scattering and therefore deposit their energy at multiple spatially separated sites, while electrons typically deposit all energy in a contained interaction cluster. Each site contributes its own cluster with energy E_i, and since the charge and light yields depend non-linearly on E_i, the combined yield of these clusters

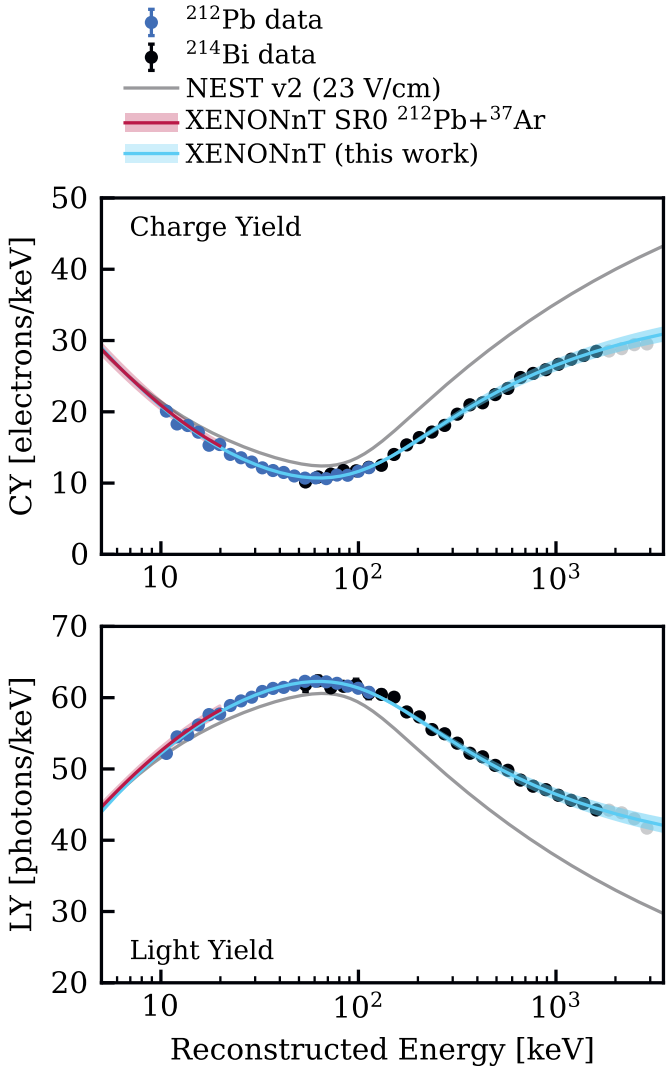


Figure 5.5 *Charge yield (CY, top) and light yield (LY, bottom) as a function of reconstructed energy, extracted from beta-decays of* ^{212}Pb *(blue) and* ^{214}Bi *(black). The ER yield model (light blue) is obtained by fitting the charge yield, with the light yield derived as the complementary fraction of total quanta. The shaded band accounts for the systematic uncertainty from the energy-bias correction. A correction for residual excited-state leakage is applied to the data. Grayed-out* ^{214}Bi *points above* 1.8 MeV *are excluded from the fit due to gamma contamination. For comparison, the default NEST model at* 23 V/cm *(gray) and the XENONnT low-energy yield model (red) are shown. Statistical uncertainties are smaller than the marker size.*

differs from that of a single deposition with the same total energy. In our analysis, this distinction emerges naturally from clustering: the yield difference between beta-decays and gamma-decays is not imposed through separate models, but arises directly from the physical interaction topology when combined with the clustering method.

We extract a new parametrisation of the ER yield model using single beta-decay events, which provide well-defined single-cluster energy depositions. For this purpose, we use ^{212}Pb decays from ^{220}Rn calibrations together with SRo ^{214}Bi decays from this analysis, ensuring broad coverage across the full energy range. The yield parameters are obtained by fitting the ten-parameter NEST function [122] to the measured charge yield, with the light yield calculated as the complementary fraction of quanta.

We account for residual leakage of excited-state events into the ground-state selection by comparing yields in simulation from truth-level ground-state beta-decay events with those obtained when the excited-state events that passed the selection are included. The observed difference is implemented as an effective correction to the charge and light yield in data. Since the leakage events concentrate above 1.8 MeV, this region is excluded from the fit.

Figure 5.5 compares the resulting charge and light yield curves with both the default NEST model at a drift field of 23 V/cm and the official XENONnT yield model used for low-energy searches [66]. The latter, derived from ^{37}Ar and ^{212}Pb calibrations, is in good agreement with our parametrisation below 20 keV. The NEST model, although nominally valid at the MeV scale, is largely based on data below 100 keV. Moreover, only limited experimental input exists at the drift field of XENONnT (23 V/cm), making the extrapolation in this regime uncertain and leading to clear discrepancies in charge and light yields.

To account for the fact that the energy bias correction described in the next section acts on the reconstructed energy but not on S1 and S2 independently, we include in Fig. 5.5 a systematic uncertainty band on our yield curves. The shaded band reflects the propagation of the energy bias into the extracted yields, covering the systematic uncertainty introduced by peak reconstruction and subsequent energy determination.

We validated the resulting yield model using two independent datasets: the clean single-cluster beta selection described above, and the ^{214}BiPo events including excited-state decays that produce combined beta and gamma topologies. These events, with multiple interaction sites and complex clustering, provide a stringent test of the clustering and yield framework. Although the yield parameters were derived only from clean beta-decays below 1.8 MeV, the model reproduces the measured S1 and S2 distributions at the percent level across the full energy range, up to

the ^{214}Bi endpoint. By contrast, simulations based on the default NEST yields and spatial clustering showed discrepancies of up to 50% in S2 for beta–gamma events above a few hundred keV. The agreement obtained here therefore supports the changes in both the clustering method and the yield parametrisation: reproducing multi-cluster topologies at high energy is only possible if both components are accurate.

The clustering procedure and yield parametrisation introduced above are designed specifically to achieve accurate reproduction of the individual S1 and S2 distributions between data and simulations. They do not directly guarantee agreement in the reconstructed energy spectra, which is instead addressed in the following section through dedicated corrections.

5.3.5 *Energy Bias Correction*

To ensure consistent energy reconstruction across simulations and the three experimental datasets (SR0, SR1, and SR1 ^{222}Rn), we derive dataset-specific energy-dependent bias corrections, addressing known effects described in previous XENONnT analyses [66]. These biases affect both the S1 and S2 signals, primarily arising from electronic noise, PMT response effects, afterpulses, photoionization, and threshold effects from the data acquisition system and ancillary software thresholds.

To accurately compare true energies, we first reconstruct the energy for each event as:

$$E_{\text{rec}} = W \cdot \left(\frac{cS1}{g_1} + \frac{cS2}{g_2} \right), \tag{5.3}$$

where $W = 13.7\,\text{eV}$ [123] is the average energy required to produce one quantum (photon or electron) in liquid xenon, cS1 and cS2 are the corrected S1 and S2 signals, and g_1 and g_2 are detector-specific calibration constants determined from the data [66]. In this analysis, we use the following values that were determined in earlier XENONnT analyses. For SR0 we adopt $g_1 = 0.151$ PE/photon and $g_2 = 16.5$ PE/electron [66]; for SR1, we use $g_1 = 0.1367$ PE/photon and $g_2 = 16.9$ PE/electron [45].

Following the procedure described in [66], we then define an empirical energy correction for each dataset by measuring the bias using monoenergetic peaks from background and calibration sources. For each peak, the bias is calculated using $\Delta E = (E_{\text{rec}} - E_{\text{true}})$, and fitted with an arctan function:

$$F(E_{rec}) \equiv \frac{\Delta E}{E_{\text{rec}}} \times 100 = A \cdot \arctan(kE_{\text{rec}}) + c, \tag{5.4}$$

where A, k, and c are free fit parameters.

We use monoenergetic lines from ^{37}Ar (2.8 keV), $^{83\text{m}}$Kr (41.5 keV), $^{131\text{m}}$Xe (163.9 keV), $^{129\text{m}}$Xe (236.1 keV), electron-positron annihilation (511 keV),

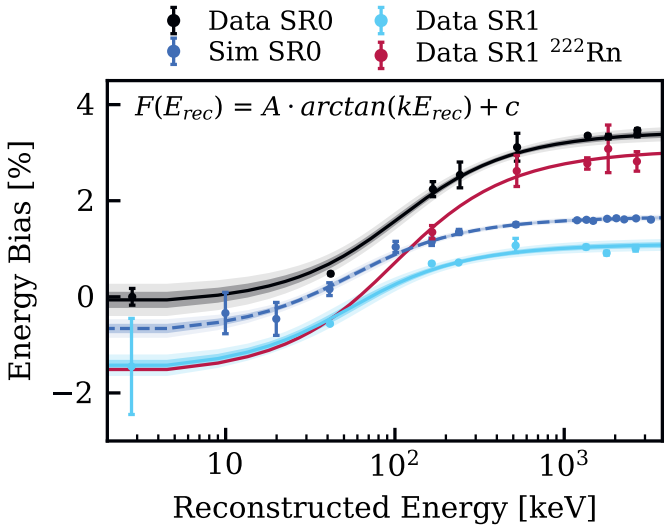


Figure 5.6 *Energy bias as a function of reconstructed energy for SR0 (black), SR1 (cyan), and SR1* ^{222}Rn *calibration data (red), as well as SR0 simulations (blue dashed). The data points correspond to monoenergetic calibration peaks at known energies. The uncertainties on each point correspond to the uncertainty in the mean obtained from the Gaussian fit to the energy distribution of the selected peaks. Bias values are fitted with an arctangent function, and shaded bands indicate the 1σ and 2σ confidence intervals of the fit. Uncertainty bands for the* ^{222}Rn *dataset are omitted for clarity.*

^{60}Co (1 173.2 keV and 1 332.5 keV), gamma-rays from external[1] ^{214}Bi (1 764.5 keV), and ^{208}Tl (2 614.5 keV). For SR1, the ^{37}Ar line originates from an accidental injection of commercial-grade xenon into the detector before the start of SR1. The ^{37}Ar was removed using cryogenic distillation before the ^{222}Rn calibration, so no ^{37}Ar datapoint is available in the SR1 ^{222}Rn dataset. For simulations, monoenergetic gamma-rays and electrons at different energies were uniformly simulated throughout the detector volume.

Figure 5.6 shows the fitted bias curves for each dataset, exhibiting clear differences: SR0 is calibrated such that ^{37}Ar has zero bias by construction, while SR1 shows systematically lower biases, becoming negative at low energies. The SR1 ^{222}Rn calibration data display higher biases at higher energies. We attribute the overall offset of the curves mainly to different choices in the g_1 and g_2 determination procedure, while the increase in

[1] External ^{214}Bi refers to the ^{214}Bi produced in the uranium decay chain within the construction materials. In this case, only the gamma-ray can reach the TPC, while the accompanying electron is absorbed before entering the detector and remains undetected.

amplitude with energy is governed by detector conditions such as the higher event rate during calibrations and photoionization, where absorbed photons release additional delayed electrons. Because no ^{37}Ar anchor point is available in the SR1 ^{222}Rn dataset, the fit was performed under constraints to prevent unphysical divergence. For this reason, we do not show the uncertainty band for this dataset in Fig. 5.6, as the unconstrained errors would be unrealistically large. This choice does not affect the analysis: the lowest energy bin used in the spectrum study is at 65 keV, well above the missing ^{37}Ar line.

The best-fit parameters for each dataset are reported in Tab. 5.2 in the appendix.

5.4 RESULTS

5.4.1 *Models to Data Comparison*

In this analysis, the event selection is rendered effectively free of background contamination as a result of the ^{214}BiPo identification algorithm. The primary remaining background arises from gamma leakage from excited-state decays into the ground-state selection. Since this leakage is accurately modeled by our simulations, which incorporate well-measured gamma emissions from ^{214}Bi, we can compare data directly to full-chain simulations.

We combine data from the three experimental datasets: SR0 (7 708 events), SR1 (5 542 events), and the SR1 ^{222}Rn (37 835 events). The consistency of the ratio between quality-selected and ground-state-selected events across these datasets ($\sim 30\%$) allows combining them into a single dataset. The energies of each dataset have been individually corrected for the energy bias described in Sec. 5.3.5.

Figure 5.7 presents the combined dataset in comparison with four simulation sets corresponding to the different theoretical beta-decay spectra: *Allowed*, *CVC*, *sNME1*, and *sNME2*. The simulated spectra are normalized to match the total number of events observed in data; this normalization involves no free fitting parameters, ensuring a robust, unbiased comparison of spectral shapes. Statistical and systematic uncertainties are included in the simulations, as discussed in Sec. 5.4.2, whereas the data points represent purely statistical uncertainties.

The sNME1 and sNME2 models exhibit significant discrepancies from the measured spectrum and are excluded from further consideration. The allowed model shows a clear energy-dependent deviation visible in the residuals (bottom panel), indicating a systematic mismatch. On the other hand, the CVC spectrum provides an excellent description of the data, with a high p-value larger than 0.05 and no evident residual energy trend.

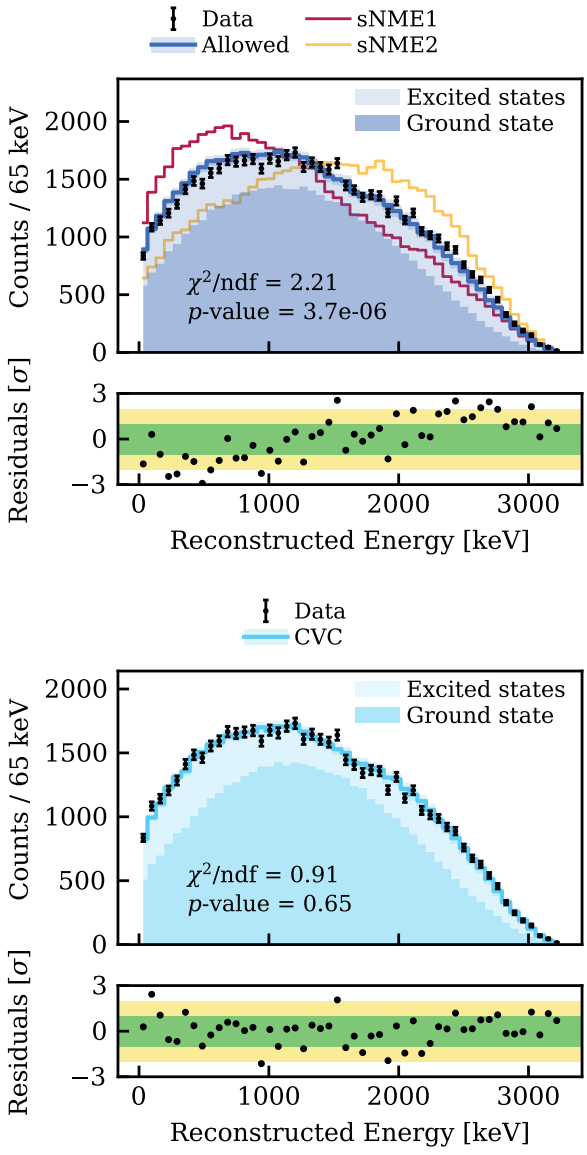


Figure 5.7 *Detailed comparison of the measured energy spectrum of ^{214}Bi ground-state data with simulations. Left panel: Data compared with the allowed model (filled histogram), including the true ground-state and excited-state contributions. The sNME1 and sNME2 theoretical spectra are shown as thin lines for reference. Right panel: Comparison with the CVC spectrum (filled histogram). The normalised residuals with respect to the allowed (left) and CVC (right) spectra are displayed in the bottom panels, along with the corresponding reduced chi-squared (χ^2/ndf) and p-values.*

This comparison strongly favors the CVC spectrum as the most accurate theoretical description of our data.

5.4.2 *Systematic Uncertainties*

We address potential systematic uncertainties arising from mismatches between simulation and data, particularly those with energy-dependent effects. We identify and evaluate three primary sources of uncertainty.

First, we consider uncertainties due to gamma-event leakage into the beta selection. While waveform properties like the S1/S2 amplitudes and signal widths are well matched between simulations and data, slight mismatches could still affect the fraction of excited-state events leaking into the ground-state selection. To account for this, we apply a 1% variation to the total number of leakage events. This choice is motivated by the observed stability (within 1%) of the ratio of selected ground-state events to all quality-selected events across multiple datasets.

Second, we examine uncertainties related to the energy scale of excited-state events. While gamma energies are well known from precise nuclear spectroscopy, the associated beta spectra are more difficult to measure and model. To assess the stability of the result, we apply an effective 5% energy shift to all excited-state events. This corresponds to approximately 150 keV at the beta endpoint, compatible with the spectral shifts observed in the ground-state spectrum studied in this work. The effect on the reconstructed spectrum is limited by the relatively small contribution of excited-state events and the dominant role of the well-measured gamma energies.

Finally, we evaluate the overall energy scale uncertainty. The reconstructed energy is corrected using energy-bias curves derived from calibration peaks, for which the fitted peak positions have uncertainties below 0.25%. To account for possible residual effects, we apply a conservative 0.5% global energy shift. At the beta endpoint, this corresponds to a less than 20 keV shift, which is smaller than the bin widths.

These three sources of uncertainty were combined by evaluating their impact on the simulated energy spectrum in each energy bin. The resulting variations were added in quadrature and incorporated as an additional uncertainty on the simulated prediction, on top of the statistical uncertainties from data. The combined systematic uncertainty remains below 1% across the entire energy range and is subdominant to the statistical error. In the high-energy tail, where event statistics are low, systematic and statistical effects become correlated, but statistical uncertainties dominate.

5.5 conclusion and outlook

We presented a measurement of the energy spectrum of the ^{214}Bi beta-decay to the ground state of ^{214}Po using data from the XENONnT experiment. The event selection employs a dedicated ^{214}BiPo tagging technique, combining background data with a ^{222}Rn calibration campaign. By developing an improved high-energy simulation pipeline, we successfully modeled ER signals up to 3.27 MeV, well beyond the low-energy range defined for WIMP searches. Our analysis strongly favors the theoretical spectrum calculated under the conserved vector current (CVC) hypothesis, significantly rejecting the allowed approximation and alternative nuclear models. These results demonstrate the capability of dual-phase xenon detectors to perform high-precision beta spectroscopy, providing valuable validations for theoretical predictions and reinforcing their potential for probing nuclear processes relevant to fundamental physics and astrophysics.

5.6 appendix

5.6.1 *Yields Parameters*

Table 5.1 *Fit results for the NEST ER yield model parameters.*

Parameter	Description	Fit Value	NEST Default (23 V/cm)
m_1	Stitching-region yield	5.00 ± 0.24	7.10
m_2	Low-energy asymptote	83 ± 3	77.3
m_3	Thomas-Imel shape control	2.00 ± 0.06	0.716
m_4	Low-energy shape control	0.77 ± 0.03	1.83
m_5	High-energy asymptote	29.3 ± 0.7	48.6
m_7	High-energy scaling	125.4 ± 2.2	94.4
m_8	High-energy shape control	2.21 ± 0.08	4.29
m_9	Low-energy asymmetry control	1.072 ± 0.006	0.334
m_{10}	High-energy asymmetry control	0.29 ± 0.09	0.066

We summarize the best-fit parameters for the ER yield model obtained from fitting ^{212}Pb and ^{214}Bi calibration data in XENONnT. The model defines the charge yield as:

Table 5.2 *Energy Bias Fit Results*

Dataset	A	k	c
SR0	2.30 ± 0.12	0.0095 ± 0.0013	-0.2 ± 0.2
SR1	1.70 ± 0.09	0.018 ± 0.003	-1.56 ± 0.13
SR1 ^{222}Rn	3 ± 7	0.01 ± 0.03	-2 ± 12
Simulations	1.56 ± 0.07	0.018 ± 0.002	-0.79 ± 0.12

$$Q_y(E,\mathcal{E}) = m_1(\mathcal{E}) + \frac{m_2 - m_1(\mathcal{E})}{\left[1+\left(\frac{E}{m_3(\mathcal{E})}\right)^{m_4(\mathcal{E})}\right]^{m_9}} + m_5(\mathcal{E}) - \frac{m_5(\mathcal{E})}{\left[1+\left(\frac{E}{m_7(\mathcal{E})}\right)^{m_8}\right]^{m_{10}(\mathcal{E})}}. \quad (5.5)$$

where E is the energy of the interaction and $\mathcal{E}$ is the electric field. The parameters m_1–m_{10} are described in detail in Ref. [75].

Our best-fit values, reported in Tab. 5.1, differ notably from the default NEST values, particularly in the high-energy region (m_5, m_8, m_{10}), highlighting the need for a data-driven recalibration for accurate high-energy ER modeling in XENONnT.

5.6.2 *Energy Bias Fit Results*

In Tab. 5.2 we report the parameters obtained from the fits of the energy bias function described in Sec. 5.3.5. The function and the parameters are introduced in Eq. 5.4. The reported best-fit values can be used to reproduce the bias correction curves shown in Fig. 5.6 for SR0, SR1, SR1 ^{222}Rn, and the SR0 simulation dataset.

6

FERMIONIC DARK MATTER ABSORPTION

The standard approach to dark matter direct detection is built around elastic scattering. A dark matter particle transfers a fraction of its kinetic energy to a target nucleus, producing a recoil at the keV scale. This approach has placed stringent constraints on WIMP–nucleon interactions over a broad mass range [45, 103, 124, 125], but loses reach for dark matter masses below $\sim$1 GeV. At low masses, nuclear recoils fall beneath detector thresholds regardless of the achieved background rate.

To probe lighter dark matter candidates, direct-detection searches have expanded beyond nuclear recoils. Scattering on bound electrons extends sensitivity to sub-GeV masses, while absorption processes provide a qualitatively different detection mechanism. In absorption, the dark matter particle is destroyed and additionally transfers its full rest-mass energy to the target rather than only its kinetic energy. This concept has motivated searches for a variety of dark matter candidates, such as axion-like particles and dark photons [124–126].

These searches typically produce monoenergetic signals with deposited energies directly related to the dark matter mass. Dror et al. showed that light-fermionic dark matter can be detected via absorption if it couples to Standard-Model fields through higher-dimensional effective operators [127]. These interactions allow the dark matter particle to be absorbed by electrons or nuclei, producing detectable final states. Fermionic dark matter (FDM) therefore provides another realization of the absorption paradigm, complementary to bosonic dark matter searches. Depending on the underlying interaction, FDM can be absorbed by electrons [128, 129] or by nuclei [130].

For nuclear targets two channels arise. In the neutral-current (NC) channel, $\chi + (A, Z) \to \nu + (A, Z)$, the absorbed dark matter particle produces an outgoing neutrino while the nucleus recoils. This channel produces a nuclear-recoil signal and requires a separate analysis not pursued here. In the charged-current (CC) channel, $\chi + (A, Z) \to e^- + (A, Z+1)^*$, the absorbed FDM particle induces a nuclear transition accompanied by the emission of an electron [130]. The CC channel is directly analogous to a

Table 6.1 *Natural abundance and ν_e/FDM charged-current capture threshold for each xenon isotope, following [32]. The threshold is the raw ground-state Q-value. For ^{134}Xe and ^{136}Xe the daughter ground state is spin-forbidden, so the accessible threshold is set instead by the lowest allowed excited state (Section 6.1.2).*

Isotope	Abundance (%)	Capture threshold (MeV)
^{124}Xe	0.095	5.93
^{126}Xe	0.089	4.80
^{128}Xe	1.9	3.93
^{129}Xe	26.4	1.20
^{130}Xe	4.1	2.98
^{131}Xe	21.2	0.355
^{132}Xe	26.9	2.12
^{134}Xe	10.4	1.23
^{136}Xe	8.9	0.0903

neutrino-induced β decay, with χ replacing ν_e at the interaction vertex. The interaction is illustrated in Figure 6.1. Because the dark matter particle is absorbed, the outgoing electron carries an energy determined primarily by the dark matter mass, offset by the nuclear Q-value and the excitation energy of the populated daughter state. This produces a monoenergetic signal that can lie in the keV to MeV range accessible to liquid-xenon detectors [53, 66].

Among the xenon isotopes, ^{136}Xe is particularly interesting because the ^{136}Cs daughter produced in the CC interaction has low-lying isomeric states [7], populated in $\approx$ 99% of de-excitations. These states generate a delayed S1 coincidence in the detector, suppressing electron-recoil backgrounds from detector materials and radioactive sources. Table 6.1 lists the neutrino/FDM capture threshold and natural abundance for each xenon isotope [32]. Only ^{129}Xe, ^{131}Xe, ^{134}Xe, and ^{136}Xe have thresholds low enough to be relevant for this search. The delayed-coincidence signature, however, has only been established for ^{136}Xe. The nuclear structure that makes this possible is described in Section 6.1.3, and the comparison with other xenon isotopes in Section 6.1.1.

Solar-neutrino CC interactions on ^{136}Xe produce an identical delayed-coincidence topology and constitute an irreducible background. Both χ and ν_e produce the same final state and cannot be distinguished event by event.

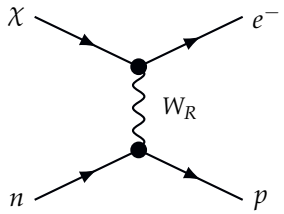


Figure 6.1 *Charged-current FDM absorption at the nucleon level, $\chi + n \to e^- + p$, mediated by the right-handed gauge boson W_R of the UV completion in [130]. In the left-right symmetric extension of the SM, χ sits in the right-handed doublet $(\chi, e)_R$ and the quarks sit in $(u, d)_R$, so W_R couples directly at both vertices. At the nuclear level this produces $\chi + Xe \to e^- + Cs^{(*)}$. The process is directly analogous to solar-neutrino CC absorption, with χ replacing ν_e [127, 130].*

This background is the central challenge of the analysis and is described in Section 6.3.

Natural xenon contains 8.9% ^{136}Xe by mass. The XENONnT active volume of 5.9 tonnes therefore provides a substantial target without isotopic enrichment [43]. The analysis uses science-run data from SR0 and SR1, corresponding to a combined natural xenon exposure of 3.1 tonne·yr. It covers dark matter masses from 0.68 to 6.9 MeV, set by the kinematic threshold and the neutron separation energy of the daughter nucleus respectively.

Section 6.1 introduces the interaction model and the nuclear transition inputs. Sections 6.2 and 6.3 describe the waveform segmentation, signal templates, and background model. Section 6.4 defines the statistical treatment, and Section 6.5 presents the observed limits and a discussion of the unmodelled background contribution at low mass.

This analysis was performed in collaboration with Alejandro Guerrero Ramos. His MSc thesis [131] covers the signal templates and statistical inference in greater detail.

6.1 Theory and Nuclear Inputs

6.1.1 *Effective Interaction and Event Rate*

The possibility of fermionic dark matter absorption was first proposed in the context of neutrino detectors by Kile and Soni [132] and later developed into a systematic framework by Dror et al. [127, 130]. The charged-current absorption process considered in this work is

$$\chi + (A, Z) \longrightarrow e^- + (A, Z+1)^{(*)}, \tag{6.1}$$

where the dark matter particle is absorbed by the nucleus, inducing a nuclear transition and the emission of an electron.

One possible ultraviolet realization is provided by a left-right symmetric extension of the Standard Model, illustrated in Figure 6.1, in which χ and the electron belong to the same right-handed doublet and interact through a heavy charged gauge boson W_R [130]. At energies relevant for direct detection, the heavy mediator can be integrated out, yielding an effective dimension-six interaction between the dark fermion and Standard-Model fields.

At dimension six the dark fermion can couple to the quark currents through four independent Lorentz structures: vector, axial-vector, scalar, and pseudoscalar [130]. The framework is agnostic about the ultraviolet completion, and each structure carries its own nuclear matrix element. Which structure contributes depends on the spin and parity of the initial and final nuclear states, and therefore differs between isotopes. For the dominant $0^+ \to 1^+$ Gamow-Teller transition in ^{136}Xe, only the axial-vector structure contributes at leading order. The vector operator connects $0^+ \to 0^+$ states not accessible near threshold, and scalar and pseudoscalar contributions to this transition vanish by angular momentum selection rules [130].

The event rate in a detector containing N_T target nuclei of isotope (A, Z) is

$$R = \frac{\rho_\chi}{2\, m_\chi}\, N_T\, (A - Z)\, \langle \sigma v \rangle_{\text{abs}}, \tag{6.2}$$

where $\rho_\chi = 0.4\ \text{GeV/cm}^3$ is the local dark matter density [37] and $(A - Z)$ counts the neutrons available for CC absorption. The factor of two averages over the spin states of the incoming Dirac fermion, only one of which couples through the chiral charged-current operator [130]. For an absorptive process, the cross section scales as $\sigma \propto 1/v$ [130]. Slower particles spend more time in the interaction region, increasing the absorption probability. This is the same behaviour seen in thermal neutron capture. As a result, the velocity-averaged product $\langle \sigma v \rangle_{\text{abs}} \propto m_\chi^2/\Lambda^4$ is velocity-independent, where Λ is the scale of the effective operator [130]. The m_χ^2 factor reflects the final-state phase space available to the outgoing electron: well above threshold, $p_e \approx m_\chi$, so the density of final states scales as m_χ^2. The rate depends only on the local dark matter number density ρ_χ/m_χ, not on the velocity distribution of the halo. This distinguishes FDM absorption from elastic-scattering searches, whose predicted rates generally depend on the dark-matter velocity distribution.

Direct-detection searches constrain the combination

$$Y \equiv \frac{m_\chi^2}{4\pi\,\Lambda^4}, \tag{6.3}$$

which has dimensions of a cross section and is expressed in cm^2. For the $0^+ \to 1^+$ Gamow-Teller transition that dominates absorption on ^{136}Xe, the cross section factorises as $\langle\sigma v\rangle_{\mathrm{abs}} \propto Y \cdot B(\mathrm{GT}) \cdot f(m_\chi)$. Here $B(\mathrm{GT})$ is the Gamow-Teller reduced transition probability, the same nuclear matrix element quantity introduced in Chapter 1, and $f(m_\chi)$ is a kinematic factor encoding the electron phase space and the Fermi function [130]. The event rate is therefore separately sensitive to the particle-physics coupling Y and the nuclear structure $B(\mathrm{GT})$, both of which are isotope-dependent. At fixed Λ, $Y \propto m_\chi^2$, producing the characteristic $+2$ slope in log-log sensitivity plots above threshold.

A viable FDM model must be cosmologically stable. The absorptive interaction implies a time-reversed process that would deplete the relic abundance, so stability requires a Dirac fermion or a conserved quantum number that forbids the reverse reaction [127]. Constraints from stellar cooling and supernovae further restrict the viable (m_χ, Λ) space. The same charged-current operator is probed directly at colliders, independent of whether χ constitutes the dark matter: for a heavy, off-shell W_R mediator, the process $u\bar{d} \to e^-\chi$ produces an energy-enhanced lepton plus missing-transverse-energy signal. A CMS search for this final state at $\sqrt{s} = 8$ TeV (LHC Run 1) constrains $g_R^2/M_{W_R}^2 \lesssim (4.5\ \mathrm{TeV})^{-2}$ [130, 133], excluding the grey region shown in Figures 6.2 and 6.15. Nothing in the model restricts χ to a subdominant fraction of the dark matter. Standard thermal or non-thermal production mechanisms, decoupled from the absorption scale Λ, can account for the full relic density $\Omega_\chi h^2 \approx 0.1$ [130]. Thus $\rho_\chi = 0.4\ \mathrm{GeV/cm}^3$ in Eq. (6.2) is taken to be the full local dark matter density.

This framework applies to every xenon isotope in the detector. For each isotope, the ground-state spin and parity determine which Gamow-Teller and Fermi transitions to the daughter are kinematically and structurally allowed, and the summed strength of the accessible transitions, weighted by natural abundance, sets the absorption rate. Each isotope has its own kinematic threshold, set by the nuclear Q-value and the excitation energy of the populated daughter state, and an upper mass limit set by the neutron separation energy of the daughter. Figure 6.2 compares the projected 10-event sensitivity on Y across the xenon isotopes with accessible CC absorption thresholds.

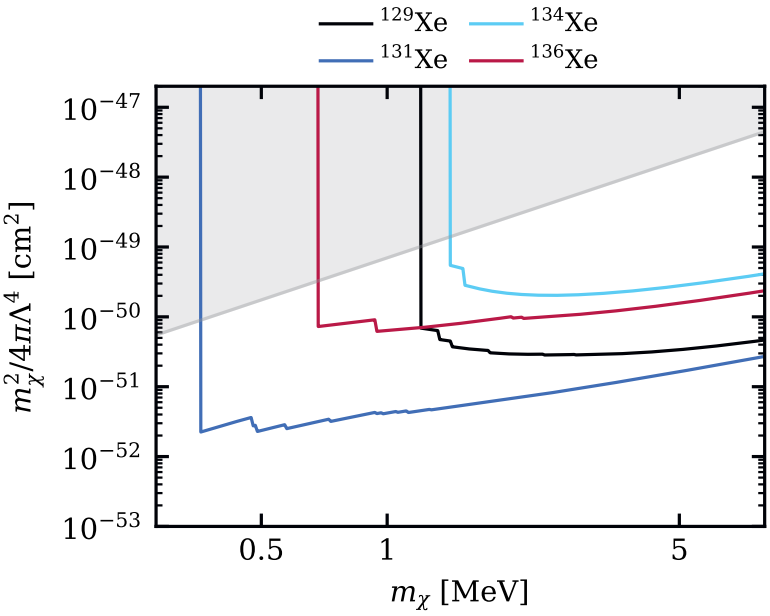


Figure 6.2 *Projected 10-event sensitivity on $Y = m_\chi^2/(4\pi\Lambda^4)$ versus m_χ for four xenon isotopes with accessible CC absorption thresholds (^{129}Xe, ^{131}Xe, ^{134}Xe, ^{136}Xe), each weighted by natural abundance and summed over available GT and Fermi transitions, following [130]. Each curve begins at the lowest kinematic threshold of that isotope. The grey shaded region is excluded by LHC Run 1 searches [133].*

6.1.2 ^{136}Xe Nuclear Structure

A search based on absorption alone would look for the energetic electron, as in inverse beta decay, and any isotope with a suitable threshold would serve. At 21.2% natural abundance, ^{131}Xe would be the obvious choice over the 8.9% of ^{136}Xe. The de-excitation of the ^{136}Cs daughter produced in the CC interaction, however, populates low-lying isomeric states that generate a delayed S1 coincidence in the detector. This coincidence provides strong rejection of electron-recoil backgrounds, a property not available in ^{131}Xe or other xenon isotopes. The cascade is described in Section 6.1.3. The nuclear selection rules and kinematic thresholds specific to ^{136}Xe are established below.

Whether a given operator contributes to absorption on a nucleus is fixed by the spin and parity of the initial and final states. The CC process converts a neutron into a proton, $\chi + n \to e^- + p$, so it connects the parent ground state to states of the daughter through the same selection rules as

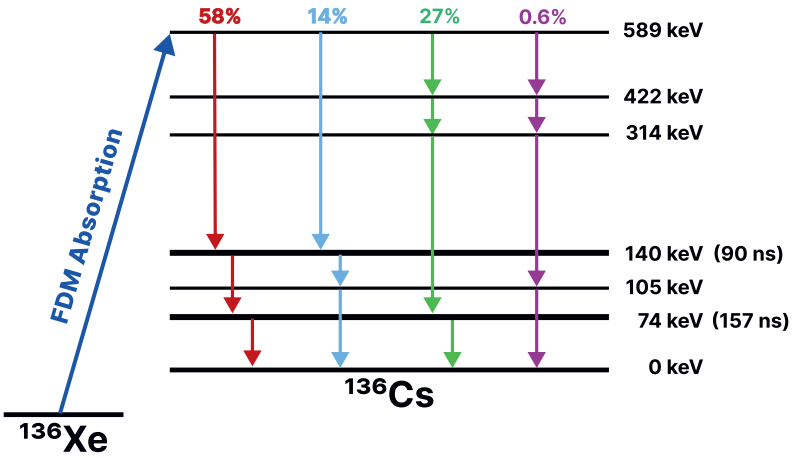


Figure 6.3 *^{136}Cs level scheme below 590 keV following FDM absorption on ^{136}Xe. Figure adapted from [134], with energies and isomeric lifetimes from [7, 135] and mode definitions following [136]. Each coloured branch shows one de-excitation path with its branching fraction. Mode 1 (58%) and Mode 2 (14%) both proceed through the 140 keV state ($\tau = 90$ ns). Of the 72% of decays that reach it, ≈80% continue on to the 74 keV isomeric state ($\tau = 157$ ns) before reaching the ground state (Mode 1: 140 keV → 74 keV → ground, two delayed emissions), and ≈20% decay directly to the ground state via the 105 keV level (Mode 2: one delayed emission). Mode 3 (27%) instead proceeds via 589 keV → 422 keV → 314 keV, reaching the 74 keV isomeric state without passing through the 140 keV state, and produces one delayed emission. A fourth branch (<1%) bypasses both isomeric states and produces no delayed signal. At least one delayed emission occurs in ≈ 99% of all de-excitations.*

ordinary β decay (Chapter 1). The ^{136}Xe ground state has $J^\pi = 0^+$. Allowed transitions therefore reach ^{136}Cs states with $J^\pi = 0^+$ (Fermi, $\Delta J = 0$) or $J^\pi = 1^+$ (Gamow-Teller, $\Delta J = 0, 1$ with no parity change).

No Fermi transitions are kinematically accessible near threshold, so the only open channels are $0^+ \to 1^+$ Gamow-Teller transitions. For ^{136}Xe the absorption rate is therefore set entirely by the Gamow-Teller strength $B(\mathrm{GT})$ of these transitions.

Two 1^+ states of ^{136}Cs are accessible near threshold (Figure 6.3): the 1^+_1 state at 588.8 keV and the 1^+_2 state at 845 keV. The 588.8 keV state carries the dominant Gamow-Teller strength from the ^{136}Xe ground state, approximately four times that of the 845 keV state [7, 135]. Below the higher threshold the 588.8 keV transition is the only open channel. Above it the 588.8 keV transition continues to dominate, with the 845 keV state included as a secondary contribution. The $B(\mathrm{GT})$ values set the absolute

rate normalisation and are taken from [7, 135]. The axial coupling $\lambda = g_A/g_V = 1.2694$ that enters the Gamow-Teller matrix element is the same parameter discussed for β decay in Chapter 1.

Absorption is kinematically allowed when $m_\chi \geq m_{\text{th}} = Q + E^*$, where Q is the ^{136}Xe $\rightarrow$ ^{136}Cs ground-state Q-value and E^* is the excitation energy of the populated daughter state. Based on earlier nuclear mass tables, $Q = 90.3$ keV [7]. A subsequent Penning-trap measurement at JYFLTRAP gives $Q = 79.08 \pm 0.54$ keV [137], reducing m_{th} for the 1_1^+ state from $\approx$680 keV to $\approx$669 keV. The analysis in this chapter uses the earlier value. The 11 keV difference does not affect the mass scan points or the inference. For the 588.8 keV transition this gives $m_{\text{th}} = 680.3$ keV, which sets the lower edge of the mass scan. The 845 keV transition opens at $m_{\text{th}} \approx 940$ keV. The outgoing electron carries kinetic energy $T_e \approx m_\chi - m_{\text{th}}$, placing the visible electron deposit between a few tens of keV and approximately 6.2 MeV across the scanned mass range. The upper edge of the scan is set by the neutron separation energy of ^{136}Cs, $S_n = 6.8$ MeV, giving $m_\chi^{\text{max}} = S_n + Q = 6.9$ MeV [130, 134]. Above this mass the daughter acquires enough excitation energy to emit a neutron, changing the signal topology and requiring a separate treatment.

6.1.3 ^{136}Xe Signal Topology

Haselschwardt et al. established that ^{136}Cs has low-lying isomeric states at 140.3 keV and 73.7 keV, with lifetimes $\tau = 90 \pm 5$ ns and 157 ± 4 ns respectively [7]. When the CC interaction populates the 1^+ state at 588.8 keV, these isomeric states are traversed in approximately 99% of all subsequent de-excitations. In the XENONnT TPC, each traversal produces a photon or conversion electron delayed by $\mathcal{O}(100)$ ns relative to the prompt electron. This creates a second S1 signal after the first. Single-site ER backgrounds produce only one. Requiring the delayed coincidence suppresses backgrounds by orders of magnitude, turning what would be a standard ER rate measurement into a topology-based search. The entire analysis strategy rests on this property.

The de-excitation from the 588.8 keV state proceeds through three modes, as illustrated in Figure 6.3 [7, 136].

In Mode 1 ($\approx$ 58% of events), the 589 keV state decays to the 140 keV isomeric state via a prompt $\approx$ 449 keV gamma. That state then decays to the 74 keV isomeric state, emitting a 66 keV photon delayed by the isomeric lifetime ($\tau = 90$ ns). The 74 keV state subsequently emits a 74 keV photon ($\tau = 157$ ns). Two delayed photons are produced in sequence. This mode yields a three-peak reconstructed topology when both delayed gammas are resolved.

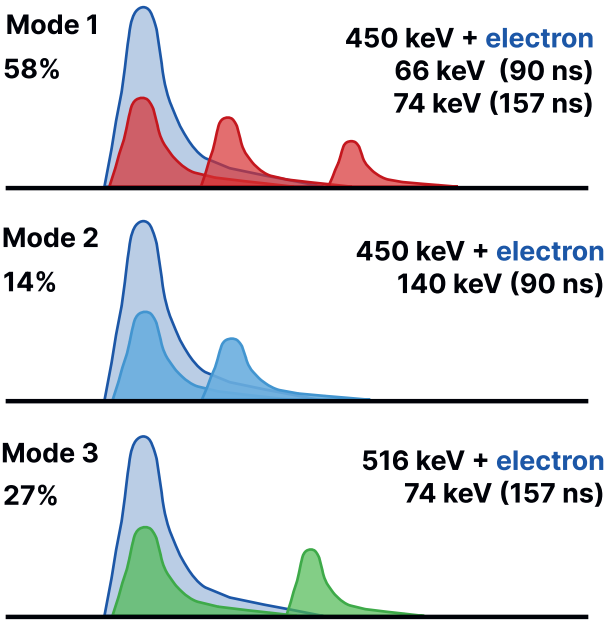


Figure 6.4 *Reconstructed S1 waveform topology for the three de-excitation modes following FDM absorption on ^{136}Xe [136]. In each panel the blue envelope is the total merged S1 and the coloured fill shows the sub-peaks resolved by the segmentation model. The prompt sub-peak ($S1_a$, left) contains the electron energy plus the prompt gammas from the cascade. Mode 1 (58%, top): two delayed sub-peaks from the 140 keV ($\tau = 90$ ns) and 74 keV ($\tau = 157$ ns) isomeric transitions. Mode 2 (14%, middle): one delayed sub-peak at 140 keV ($\tau = 90$ ns), with a larger prompt component (≈ 450 keV) because the cascade reaches the 140 keV state via prompt intermediate levels. Mode 3 (27%, bottom): one delayed sub-peak at 74 keV ($\tau = 157$ ns), with the largest prompt component (≈ 516 keV) because the 589 keV state decays directly to the 74 keV isomeric state.*

In Mode 2 ($\approx$ 14% of events), the 589 keV state decays through the 422 keV and 314 keV levels via prompt transitions, reaching the 140 keV isomeric state. The 140 keV isomeric state decays directly to ground, emitting a 140 keV photon ($\tau = 90$ ns). This mode produces a two-peak topology.

In Mode 3 ($\approx$ 27% of events), the 589 keV state decays directly to the 74 keV isomeric state, bypassing the 140 keV level. A single prompt emission of $\approx$ 516 keV precedes one delayed emission at 74 keV ($\tau = 157$ ns). This mode also produces a two-peak topology.

The three modes produce different prompt energies. The prompt component $S1_a$ contains the electron energy $T_e \approx m_\chi - m_{\text{th}}$ plus the prompt

gammas: ≈ 450 keV for Modes 1 and 2, and ≈ 516 keV for Mode 3. This branching distributes the signal population across a range of $(S1_a, S1_b)$ values that shifts with m_χ, as seen in the signal templates of Section 6.2.4. Figure 6.4 shows the S1 waveform topology for all three modes.

The analysis is performed in the two-dimensional $(S1_a, S1_b)$ plane. Events with two or three reconstructed sub-peaks ($S1^{(2)}$ and $S1^{(3)}$) are both accepted. For three-peak events, $S1_a$ is the prompt sub-peak and $S1_b$ is the first delayed sub-peak. The second delayed sub-peak does not enter the inference. No $S1^{(3)}$ events are observed in the SR0+SR1 dataset.

The cascade therefore produces a prompt energy deposit, the electron together with the prompt gammas, followed by one or more delayed deposits from the isomeric transitions. These appear as two (or, for Mode 1, three) distinct sub-peaks in the S1 waveform. Throughout this chapter, $S1_a$ denotes the area of the prompt sub-peak in photoelectrons (PE), $S1_b$ denotes the area of the first delayed sub-peak in PE, and Δt denotes the difference between the area-weighted centre times of the two sub-peaks. These quantities are reconstructed from merged S1 waveforms by the segmentation model described in Section 6.2.1.

6.2 Analysis Methods

6.2.1 *ML Segmentation Model*

A typical S1 signal in XENONnT has a duration of $\mathcal{O}(300)$ ns, set by the triplet scintillation decay time and photon transport in the active volume (Section 2.4). Two energy deposits separated by a delay shorter than this characteristic S1 width produce a single merged waveform in the standard reconstruction [66]. The prompt and delayed components described in Section 6.1.3 are therefore merged into a single S1 peak for most of the FDM signal population, and the observables $S1_a$, $S1_b$, and Δt cannot be reconstructed from the standard output.

The segmentation model, developed in [138], is a convolutional neural network that takes the top- and bottom-array S1 waveforms as input, sampled at 10 ns resolution, and produces per-sample segmentation masks together with area and timing estimates for each identified sub-peak. Training uses simulated waveforms from `fuse` [68] augmented with Rn calibration data. To train the model to resolve merged S1 pulses, waveforms from separate single-site deposits are superimposed, reproducing the topology in which a prompt and a delayed sub-peak appear within a single merged waveform.

The reconstructed sub-peaks are labelled $S1_a$, $S1_b$, $S1_c$ in time order. An event with n resolved sub-peaks is denoted $S1^{(n)}$. Because the physical roles

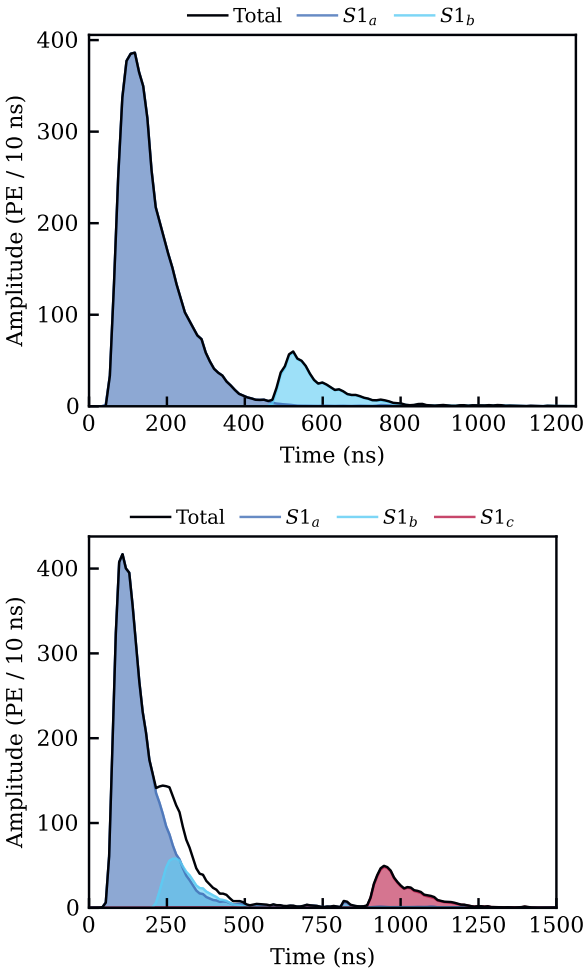


Figure 6.5 *Simulated S1 waveforms with two (top) and three (bottom) reconstructed sub-peaks. In the two-peak case, $S1_a$ carries the prompt electron signal and the de-excitation photon contribution, and $S1_b$ is the delayed sub-peak separated by $\mathcal{O}(400)$ ns. In the three-peak case, the segmentation model resolves an additional delayed component $S1_c$ at $\mathcal{O}(1000)$ ns. Both de-excitation topologies are modelled in the signal templates with their physical branching fractions.*

of the sub-peaks are not necessarily time-ordered, the analysis re-assigns them after reconstruction. $S1_a$ denotes the prompt sub-peak (the larger one, carrying the electron energy) and $S1_b$ denotes the delayed sub-peak (the smaller one, from the nuclear de-excitation), irrespective of their time order. The time separation Δt is the difference of the area-weighted centre

times of the two sub-peaks. The sum $S1_a + S1_b$ approximates the total scintillation yield and provides a cross-check on the energy scale. The S2 signal is used for position reconstruction and quality cuts but is not a primary discriminating observable. The signal search is performed in the $(S1_a, S1_b)$ plane.

Figure 6.5 shows representative simulated waveforms for both the two-peak and three-peak cases. In the two-peak case, the prompt component $S1_a$ and the delayed component $S1_b$ are clearly separated by $\mathcal{O}(400)$ ns, with relative areas reflecting the branching structure described in Section 6.1.3. In the three-peak case the segmentation resolves a further delayed component $S1_c$ at $\mathcal{O}(1000)$ ns.

6.2.2 *Model Performance and Validation*

Solar electron neutrinos from ^{7}Be are monoenergetic at 862 keV and undergo the same charged-current interaction on ^{136}Xe as the FDM signal, producing an identical delayed-coincidence topology. ^{7}Be is the dominant CC background at low m_χ and defines the natural reference point for the analysis. The full background treatment is in Section 6.3.3. The 700–820 keV visible-energy region, spanning the expected ^{7}Be CC peak, was kept blinded throughout SR0+SR1 as part of the protocol for the parallel solar-neutrino CC measurement on ^{136}Xe [138]. Validation and cut development described below use data outside this region.

Performance is evaluated on solar-neutrino CC signal simulations, focusing on sub-peak area and time reconstruction. At a neutrino energy of 862 keV, approximately 47% of simulated signal events are reconstructed as $S1^{(2)}$ and 16% as $S1^{(3)}$ ($\approx$28% of the Mode 1 fraction). Both categories enter the signal template. The remaining $\approx$ 37% are reconstructed as $S1^{(1)}$ and are excluded from the analysis. Reconstruction degrades at short delays where the two sub-peaks are not resolved, and at high energies where the prompt component dominates and the delayed sub-peak is small. The mass-dependent efficiency accounts for this.

A useful validation of the segmentation and reconstruction chain is provided by ^{83m}Kr, a calibration source injected into the xenon. It decays via two transitions, the second delayed by the isomeric lifetime $\tau = 154$ ns: a 32.1 keV transition followed by a 9.4 keV transition, described in Chapter 2. This is a known two-pulse signal at a delay directly comparable to the FDM signal population. The reconstructed $(S1_a, S1_b, \Delta t)$ location of the ^{83m}Kr population is compared against simulation and used as an end-to-end check of the segmentation and reconstruction chain.

Two further populations extend the validation, both arising in the science data rather than from deliberate injection of radioactive tracers. A

small amount of ^{85}Kr entered the detector through an accidental air leak during part of SR1 [45]. It undergoes β^- decay to an isomeric state of ^{85}Rb in 0.43% of decays, with the isomeric state emitting a 514 keV gamma at $\tau = 1015$ ns [139], also producing a delayed coincidence at timescales comparable to the FDM signal population (Figure 6.6). A combined reconstruction acceptance of 40.3% is found from simulation using the chain described in Section 6.2.4 [138]. ^{133}Xe is also present in the dataset, as a residual of neutron activation, produced via $^{132}\mathrm{Xe}(n,\gamma)^{133}\mathrm{Xe}$ during the AmBe neutron calibration campaigns visible in the science-run timeline of Figure 2.9. The ^{133}Xe β^- decay ($T_{1/2} = 5.2$ days) populates the 81 keV state of ^{133}Cs in ≈99% of decays. This state has a half-life of $T_{1/2} = 6.2$ ns [139] and decays to the ground state via an 81 keV transition, observed directly as a gamma ray in 37% of cases and by internal conversion otherwise. The resulting prompt-electron, delayed-gamma pairs populate the analysis sideband in SR1 and verify the reconstruction in the Δt range accessible to the twofold topology.

Figure 6.6 shows all three control populations in the $(S1_a, S1_b)$ plane, with the ^{7}Be solar-neutrino region shown for reference. The $^{83\mathrm{m}}$Kr population clusters at the coordinates corresponding to the 9.4 keV and 32.1 keV transitions, separated by $\Delta t \approx 154$ ns. The ^{85}Kr population appears at larger $S1_b$ and longer delay, consistent with the 1015 ns isomeric lifetime of ^{85}Rb. The ^{133}Xe population occupies the low-$S1_a$ region of the plane. All three calibration populations are broadly consistent with the simulation predictions, confirming that the segmentation and reconstruction introduce no significant bias in area or timing. The ^{133}Xe data points are displaced slightly upward in $S1_b$ relative to the simulation prediction, indicating a small systematic offset in this population.

6.2.3 *Event Selection*

The full event selection is described in [138]. The stages are summarised here.

1. **Valid sub-peak:** rejects events where $S1_a$ or $S1_b$ has zero reconstructed area.
2. **Base quality:** requires $S1_a$ to start within the first 200 ns of the event window and that its risetime and area-fraction-top are consistent with a genuine S1 signal. A pile-up veto rejects events where multiple S2 positions disagree by more than $\Delta z = 16$ cm or $\Delta r = 14$ cm in the transverse plane.
3. **Fiducial volume:** applies a 2D $r(z)$ boundary corresponding to the ≈ 4 t XENONnT fiducial mass.

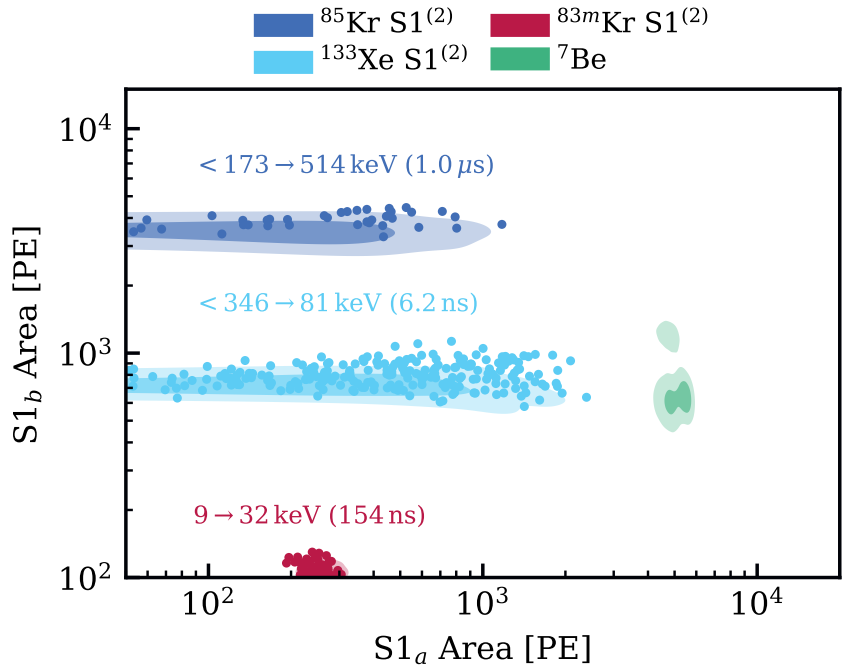


Figure 6.6 *Calibration control populations in the $(S1_a, S1_b)$ plane, used to validate the segmentation and reconstruction chain. Coloured contours show the simulation-predicted distributions for ^{85}Kr, ^{133}Xe, and ^{83m}Kr at the 68% and 95% probability levels. Data points from the same populations are overlaid as coloured scatter points. ^{83m}Kr is labelled with its 9 keV and 32 keV transitions, separated by 154 ns. ^{133}Xe is labelled with its 81 keV transition at 6.2 ns, and ^{85}Kr with its 514 keV transition at 1.0 µs. The ^{7}Be solar-neutrino contour is shown for reference. Data in the corresponding energy region are blinded for the parallel solar-neutrino CC measurement on ^{136}Xe [138]. A small upward displacement in $S1_b$ is visible for the ^{133}Xe population.*

4. **High-energy:** selects the electron-recoil band in the $(cS1, cS2)$ plane and requires an S2 width consistent with longitudinal diffusion at the reconstructed drift time.
5. **Topology:** requires at least two reconstructed S1 sub-peaks. The main failure mode is false-positive splitting, where the segmentation model identifies two sub-peaks in a genuine single-deposit S1. Cuts on $S1_b$ area and the $S1_b$ presence score suppress this. The false-positive rate falls from 8.7×10^{-4} before any selection to below 8×10^{-8} after the full set of topology cuts [138].

Cut development was performed on the sideband outside the ^{7}Be solar-neutrino energy region, which was kept blinded throughout.

6.2.4 *Signal Template Construction*

The inference model (Section 6.4) requires a predicted two-dimensional $(S1_a, S1_b)$ distribution for each dark matter mass hypothesis. These mass-dependent predicted distributions are the signal templates. Each is built from a full simulation chain that reproduces the expected event topology for a given mass [131]. The starting point is `DECAY0` [140], a nuclear decay generator configured to reproduce the three de-excitation modes described in Section 6.1.3: the outgoing electron at energy $T_e = m_\chi - m_{\text{th}}$, together with the de-excitation gammas at their physical branching ratios and energies. The `DECAY0` output is passed to Geant4 [69–71], which propagates the particles through the XENONnT geometry and records the energy deposits in the active volume. The Geant4 output then enters `fuse` [68] (Chapter 3), the detector simulation framework, which applies the full detector response chain. The final step is `straxen` processing, applying the same reconstruction algorithms as are used on the data. This includes the waveform segmentation model (Section 6.2.1) and all topology cuts described in the preceding sections.

This full chain is run at 16 fermionic dark matter mass points: 0.68, 0.75, 0.80, 0.83, 0.85, 0.862, 0.87, 0.90, 0.95, 1.00, 1.20, 1.50, 2.00, 3.00, 5.00, and 6.90 MeV. The spacing is denser below 1.5 MeV, where the signal template changes rapidly near threshold and the ^{7}Be solar-neutrino background overlaps most strongly with the signal region. At each mass point, the chain produces a two-dimensional distribution of reconstructed $(S1_a, S1_b)$ values, spanning $S1_a \in [0, 50\,000]$ PE and $S1_b \in [0, 8\,000]$ PE.

The ^{136}Cs de-excitation proceeds through the three modes described in Section 6.1.3. All three are included in the signal template weighted by their physical branching fractions (58%, 14%, and 27% for Modes 1, 2, and 3 respectively) [136]. Events with two or three reconstructed sub-peaks are both accepted. For triple-peak events, the prompt and first delayed sub-peaks define $S1_a$ and $S1_b$. The branch-weighted distribution is normalised to a PDF and stored at each mass point for use in the inference model.

Figure 6.7 shows the signal contours at three representative masses, incorporating all three de-excitation modes. As m_χ increases, the signal population shifts to higher $S1_a$, tracking the growing electron kinetic energy, while $S1_b$ remains anchored by the fixed ^{136}Cs de-excitation energies.

Applying the event selection cuts (Section 6.2.3) to these templates gives the signal selection efficiency $\varepsilon(m_\chi)$ shown in Figure 6.8, evaluated at each of the 16 mass points and used to normalise the predicted signal rate in the inference. The dominant loss at all masses comes from events where the segmentation fails to resolve any delayed sub-peak, producing a single-peak $S1^{(1)}$ reconstruction that is excluded from the analysis. A second steep fall

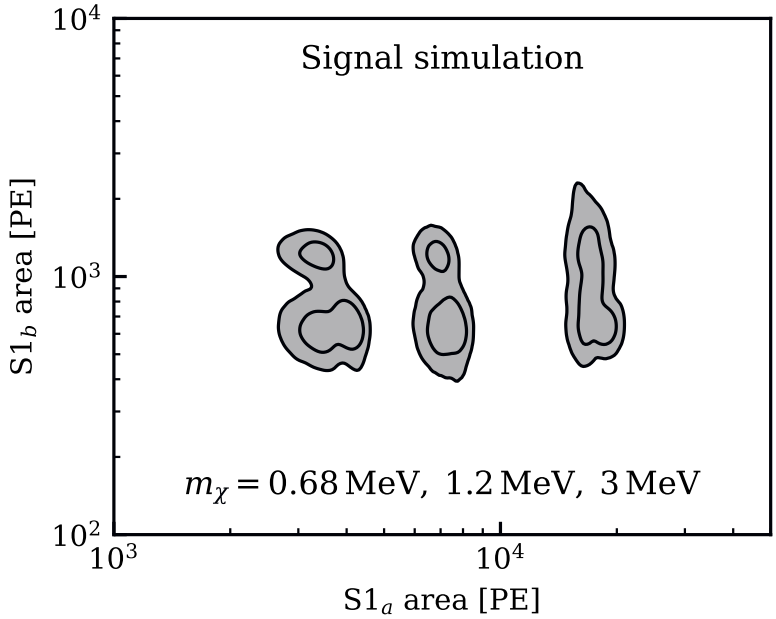


Figure 6.7 *Signal contours in the $(S1_a, S1_b)$ plane at three dark matter masses (0.68, 1.2, and 3 MeV), at the 68% and 95% probability levels, incorporating both two-fold and three-fold de-excitation topologies with their physical branching fractions. As m_χ increases, the signal population shifts to higher $S1_a$, tracking the growing electron kinetic energy. The $S1_b$ range is anchored by the fixed ^{136}Cs de-excitation energies and does not shift.*

appears above $\sim$ 3 MeV, where the growing $S1_a$ increasingly dominates the waveform and the segmentation algorithm fails to resolve the delayed sub-peak in most events.

No selection cut separates the FDM signal from solar-neutrino CC interactions on ^{136}Xe: both produce an identical delayed-coincidence topology. The background is constrained through the unbinned template fit in $(S1_a, S1_b)$ rather than by a signal region definition.

6.3 Background Model

The delayed-coincidence requirement strongly suppresses electron-recoil backgrounds from detector materials and radioactive sources. The residual concern is any process that can produce two S1-like signals within the waveform window and pass the topology cuts.

Four classes of process were considered. Random coincidences between uncorrelated S1 signals are calculable from measured rates and found to be

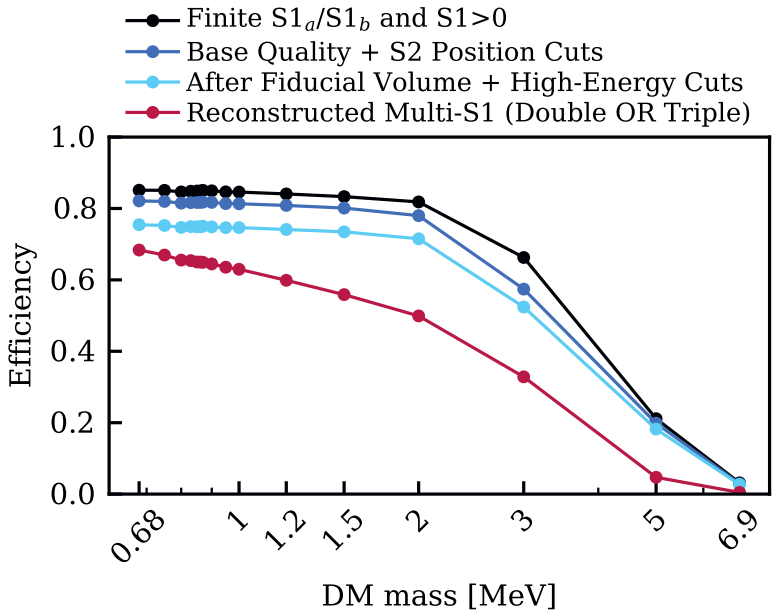


Figure 6.8 *Signal selection efficiency as a function of m_χ, broken down by cut stage (Section 6.2.3), weighted by the physical branching fractions of all three de-excitation modes. Starting from all simulated events, the valid S1$_a$/S1$_b$ cut (black), base quality cuts (dark blue), fiducial volume and high-energy cuts (light blue), and the topology requirement of at least two reconstructed S1 sub-peaks (red) are applied cumulatively. The reconstruction requirement is the dominant source of efficiency loss and drives the steep fall above ∼3 MeV.*

subdominant. Neutron double scatters can produce a two-peak S1 topology, but the rate is suppressed by the neutron veto and fiducial volume cut, and these events populate a different region of $(S1_a, S1_b)$ from the signal. Nuclear de-excitation backgrounds, where a β-emitting isotope produces a delayed gamma or conversion electron from the daughter nucleus, can in principle mimic the signal topology exactly. A systematic scan of the IAEA nuclear database [141] was performed to identify radioactive isotopes resident in the active xenon that could produce the required delayed-coincidence topology. None were found with sufficient rate after the full selection. Spallation products were assessed separately and are described below. The main irreducible background is charged-current solar-neutrino interactions on ^{136}Xe, which produce the same delayed-coincidence topology by the same nuclear mechanism as the FDM signal. This component enters the inference model explicitly.

6.3.1 *Pile-up and Neutron Background*

Accidental coincidences arise when two uncorrelated S1-like signals fall within the delayed-coincidence window by chance. Their rate scales as $R_1 \times R_2 \times \tau$, where R_1 and R_2 are the rates of prompt and delayed S1-like activity and τ is the coincidence window. The expected rate from this source is 0.05 events per year in the $\approx$ 4 t fiducial volume [138], subdominant relative to the solar-neutrino background.

A neutron can scatter twice within the active volume, producing two nuclear recoils separated by tens to hundreds of nanoseconds. If both fall within the S1 waveform window, the resulting two-peak topology mimics the delayed-coincidence signal. The neutron veto [59] and the fiducial volume cut reduce this contribution substantially. The residual expected rate from radiogenic neutrons is below 3 events per year in the $\approx$ 4 t fiducial volume [138], though this is a conservative upper bound limited by simulation statistics. Neutron-induced events tend to populate the region of small $S1_a$ and large $S1_b$, which is inconsistent with the signal topology. Their contribution within the signal ROI is therefore subdominant and neither source is included in the background model.

6.3.2 *Muon-Spallation Products*

When energetic muons pass through detector material, they can fragment target nuclei via inelastic hadronic interactions, producing short-lived radioactive isotopes through a process known as spallation [142]. Muons crossing the detector at LNGS (muon flux $\approx 3 \times 10^{-8}$ cm^{-2} s^{-1} [44]) can therefore produce radioactive spallation isotopes in or near the active xenon volume. If such an isotope undergoes a β decay with a delayed de-excitation of the daughter nucleus, it could produce a signal topology similar to the FDM signal.

A study was carried out to assess this background. Production rates for spallation isotopes at LNGS were obtained from FLUKA simulations [143, 144], provided by Kelly Weerman (XLZD Collaboration).[1] Isotopes with a predicted rate below 0.7×10^{-4} events yr^{-1} kg^{-1} were discarded, corresponding to fewer than 0.5 expected events in the SR0+SR1 exposure before any selection efficiency. Detection efficiency can only be assessed through full-chain simulation and was therefore evaluated only for isotopes passing this filter. In principle, an isotope with a high production rate but low detection efficiency could be missed at this stage. In practice, the absolute rates are small enough that any such contribution is expected to remain subdominant. The surviving isotopes were then compared against

[1] Private communication.

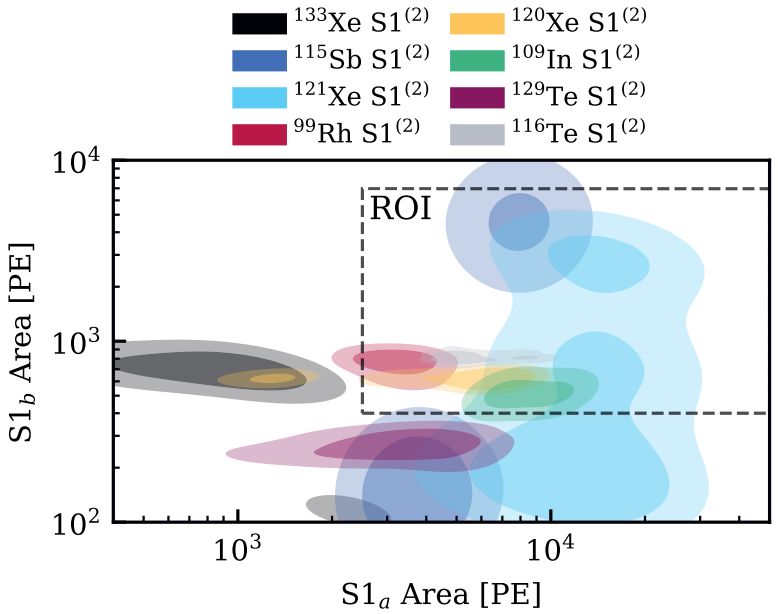


Figure 6.9 *Reconstructed $(S1_a, S1_b)$ distributions of the dominant muon-spallation background candidates after the full selection. Contours show the 30% and 60% highest-density regions of the KDE-smoothed distribution for each isotope. The dashed rectangle marks the region compared quantitatively in Figure 6.12. It is a visual reference only, not an analysis cut (Section 6.4.1). ^{133}Xe has the largest total rate of the fourteen isotopes considered but, unlike ^{120}Xe, ^{99}Rh, and ^{116}Te, its distribution lies almost entirely outside this region.*

the IAEA nuclear database for decay modes that can produce a delayed gamma or conversion electron following a β transition, i.e. topologies capable of generating the two-component S1 signature. Those passing this filter were simulated with `fuse` and the ML segmentation to determine their distribution in the $(S1_a, S1_b)$ analysis space.

Fourteen isotopes passed the initial filter: ^{117}I, ^{99}Rh, ^{100}Pd, ^{95}Tc, ^{115}Te, ^{129}Te, ^{109}In, ^{120}Xe, ^{116}Te, ^{121}Xe, ^{115}Sb, ^{122}Xe, ^{123}I, and ^{133}Xe. After applying the full selection, ^{133}Xe has the highest total rate of the fourteen, with an expected 0.41 events in SR1. Its distribution, however, lies mainly outside the region shown for comparison in Figure 6.9. All other isotopes contribute well below 0.1 events each in total. ^{100}Pd, ^{123}I, and ^{95}Tc also contribute essentially no rate within that region. The isotopes that do contribute there are ^{120}Xe, ^{99}Rh, and ^{116}Te. Summed over all fourteen isotopes, the spallation rate in that region is about 17% of the summed solar-neutrino rate there, and roughly a factor of 11 above the ^{8}B rate alone

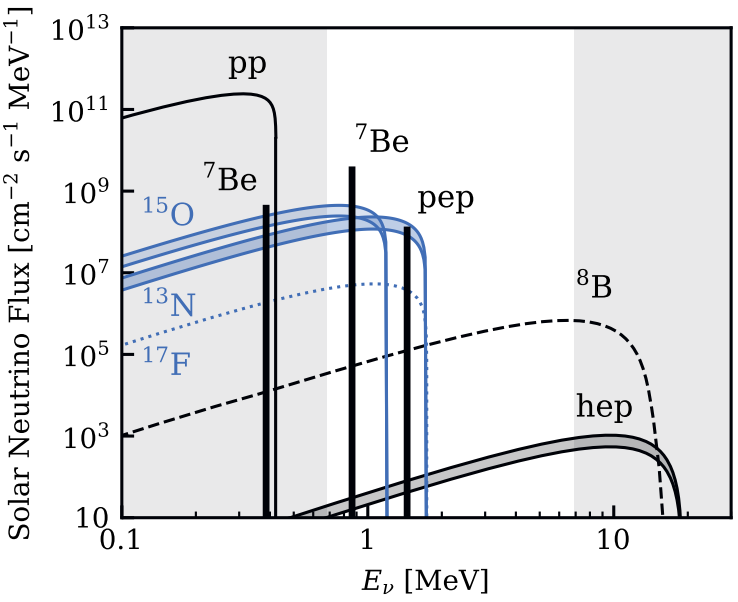


Figure 6.10 *Solar neutrino flux as a function of neutrino energy, for the components relevant to the CC background in this analysis. The ^{7}Be and pep fluxes appear as lines at fixed energies. The ^{8}B and CNO (^{13}N, ^{15}O, ^{17}F) components have continuous spectra. Narrow shaded bands around the CNO and hep curves indicate flux uncertainties from the standard solar model. The wide grey bands mark the energy range outside the $m_\chi = 0.68$–6.9 MeV mass window accessible to this analysis (as discussed in the introduction to this chapter). The four components entering the background model (^{7}Be, pep, ^{8}B, CNO) are those with CC thresholds accessible to ^{136}Xe. Adapted from [145], with CNO spectra from [146].*

(Figure 6.12). This contribution is small compared to the dominant ^{7}Be component but not strictly negligible. These isotopes are not included in the primary full-observable-space inference because the dataset contains a distinct, calibration-induced ^{133}Xe population not described by the steady-state muon-spallation template. Section 6.5.2 reports a spallation-inclusive cross-check, performed in a reduced observable space that avoids this population.

6.3.3 *Solar-Neutrino Interactions*

Charged-current interactions of solar electron neutrinos on ^{136}Xe,

$$\nu_e + {}^{136}\mathrm{Xe} \longrightarrow e^- + {}^{136}\mathrm{Cs}^{(*)}, \tag{6.4}$$

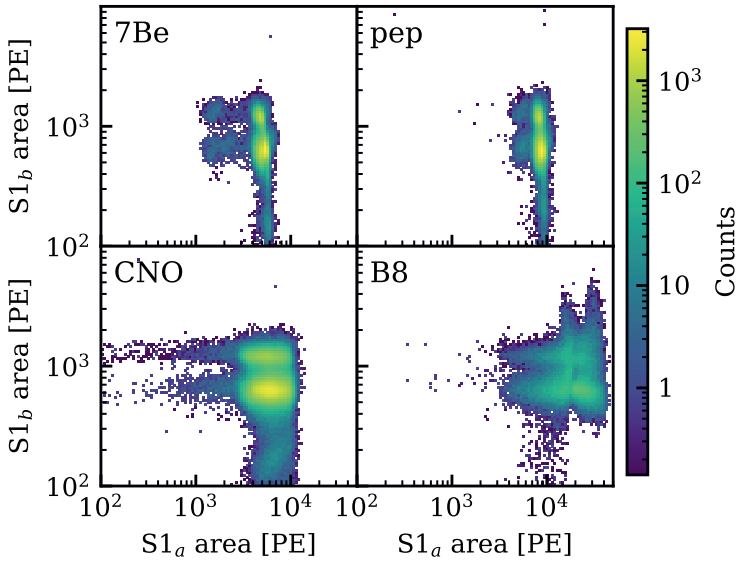


Figure 6.11 *Solar-neutrino CC background templates in the $(S1_a, S1_b)$ plane for the four components included in the model (^{8}B, ^{7}Be, pep, CNO), shown at the 68% and 95% probability levels. Each template is propagated through the full simulation and selection chain. The ^{7}Be, pep, and CNO components cluster at $S1_a \sim 10^4$ PE, corresponding to the electron energy from each neutrino source. The ^{8}B component extends to higher $S1_a$ values and overlaps with the signal region at larger m_χ.*

produce an identical delayed-coincidence topology to the FDM signal and constitute an irreducible physics background. This component cannot be vetoed without direct signal loss. Measuring and constraining it is therefore central to the analysis.

Four solar-neutrino components contribute in the relevant energy range: ^{7}Be, CNO, pep, and ^{8}B. Their flux spectra are shown in Figure 6.10. The corresponding CC interaction rates on ^{136}Xe are taken from [32, 135] and are 5.9, 0.88, 0.78, and 0.53 events per tonne-year of ^{136}Xe for ^{7}Be, CNO, pep, and ^{8}B respectively. Multiplied by the ^{136}Xe natural abundance of 8.9% and the total natural xenon exposure of 3.1 tonne·yr (giving 0.276 tonne·yr of ^{136}Xe), the expected pre-selection event counts are 1.62, 0.24, 0.21, and 0.15 events for ^{7}Be, CNO, pep, and ^{8}B respectively (2.22 total). Each component populates a distinct region of the $(S1_a, S1_b)$ plane because the outgoing electron energy is set by the neutrino energy minus the nuclear transition threshold, which differs between sources. The ^{8}B spectrum, which extends to $\sim$15 MeV, reaches the highest $S1_a$ values and dominates the background at the upper end of the FDM mass range. The ^{7}Be line at 862 keV is close to

the CC threshold on ^{136}Xe and contributes at lower $S1_a$ values, overlapping more strongly with the signal region at low m_χ.

Solar neutrinos are produced as ν_e but undergo matter-enhanced (MSW) flavour conversion inside the sun, arriving at Earth as a mixture of all three flavours. Since the CC interaction on ^{136}Xe couples exclusively to ν_e, only the surviving ν_e fraction contributes. The relevant quantity is the electron neutrino survival probability $P_{ee}(E_\nu)$, which is energy-dependent and has been measured precisely at solar neutrino energies [145, 147]. The CC rates from [32, 135] already incorporate this factor, so no additional correction is applied here. For the fixed-energy sources (^{7}Be and pep), P_{ee} is evaluated at a single energy. For the CNO continuous spectra (^{13}N, ^{15}O, ^{17}F), P_{ee} varies with energy. Within the CNO energy range, however, the variation is small, and applying a single effective value introduces a negligible correction. These rates are therefore used without further modification.

In the inference model, each of the four solar-neutrino components (^{7}Be, CNO, pep, ^{8}B) enters as a separate template source with an independent rate multiplier, treated as a nuisance parameter profiled out in the likelihood fit. The common 5% Gaussian constraint used here is an effective analysis-level constraint on the four solar CC template normalisations. It is anchored by the precision of the dominant ^{7}Be component (2.7% [147]), while the source-specific flux uncertainties of CNO, pep, and ^{8}B ($+30\%/-14\%$ [148], $\approx$ 17% [147], and $\approx 4\%$ [149] respectively) are not separately factorised in the present model. This simplified treatment is adopted for the nominal analysis and is stress-tested below. It should not be interpreted as a complete detector or nuclear-theory systematic model.

To test the sensitivity of the result to this simplified prior treatment, the fit was repeated using flux-oriented, component-specific rate constraints (5% for ^{7}Be, 30% for CNO, 15% for pep, 7% for ^{8}B) in place of the common 5% width. Because ^{7}Be dominates the solar expectation (1.05 of the 1.34 total events), the two treatments give similar effective absolute uncertainties on the total solar rate: approximately 0.054 events for the uniform 5% width versus approximately 0.071 events for the flux-oriented set, a difference of only about 0.017 events over SR0+SR1. The median expected sensitivity is essentially unchanged between the two prior choices, confirming that the expected sensitivity is dominated by counting statistics at the present exposure. The observed limit, separately, is governed by the small observed event sample.

After applying the full event selection and reconstruction acceptances, the expected solar-neutrino CC event counts are 1.05, 0.15, 0.12, and 0.02 events for ^{7}Be, CNO, pep, and ^{8}B respectively, giving a total expected solar-neutrino background of 1.34 events. This is the dominant background in the analysis. Figure 6.11 shows the background template distributions in

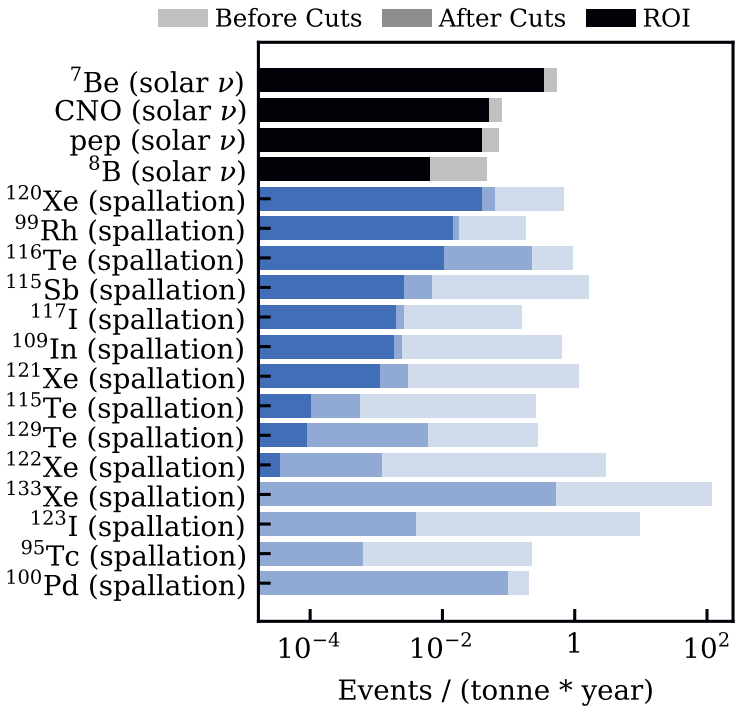


Figure 6.12 *Expected background event rates per tonne-year for all components considered in the analysis, at three successive stages: the total event rate, after the full event selection, and within the region shown for visual reference in Figure 6.9. Solar-neutrino components (black) dominate within that region, with ^{7}Be the leading contribution. The combined spallation rate (blue) is roughly double the combined solar-neutrino rate after the full selection, but drops to about 17% of it once restricted to this region, since most spallation isotopes populate a part of the $(S1_a, S1_b)$ plane away from it.*

the observable space. The relative rates of all background components at successive stages of the event selection are summarised in Figure 6.12.

The ^{85}Kr and ^{133}Xe populations described in Section 6.2.2 are visible in the analysis sideband. Data points consistent with them are pre-assigned to their respective models and excluded from the signal-region inference.

6.4 Statistical Inference

The inference uses an unbinned template profile-likelihood implemented in the `alea` framework [150], which uses `blueice` [151] for the extended

likelihood evaluation. Upper limits are evaluated independently at all 16 mass points.

6.4.1 *Template Profile-Likelihood*

The observable space is $(S1_a, S1_b) \in [0,\, 50\,000]\,\mathrm{PE} \times [0,\, 8\,000]\,\mathrm{PE}$. An unbinned extended log-likelihood is defined over the observed events in this space,

$$\ln \mathcal{L}(\mu_s, \boldsymbol{\theta}) = \sum_{i=1}^{N} \ln \left[\mu_s\, f_s(\mathbf{x}_i;\, m_\chi) + \sum_k \mu_{b,k}\, f_{b,k}(\mathbf{x}_i) \right] - \mu_s - \sum_k \mu_{b,k} - \frac{1}{2} \sum_k \theta_k^2, \tag{6.5}$$

where μ_s is the expected signal count, $\mu_{b,k}$ is the expected count for background component k, and f_s, $f_{b,k}$ are the corresponding normalised PDFs in $(S1_a, S1_b)$. The signal PDF is the mass-dependent template from the simulation chain described in Section 6.2.4. The background PDFs are the four solar-neutrino CC templates (^{7}Be, CNO, pep, ^{8}B) from Section 6.3.3.

Each background component enters with a nuisance parameter θ_k that rescales its nominal rate: $\mu_{b,k} = \mu^0_{b,k}\,(1 + 0.05\,\theta_k)$. The 5% relative Gaussian constraint on θ_k, motivated in Section 6.3.3, is the last term in Eq. (6.5), and the four nuisance parameters are profiled out in the likelihood maximisation. This constraint sets an effective rate uncertainty and does not separately model detector-response or template-shape systematics.

Upper limits at 90% CL are derived using a non-asymptotic Neyman belt construction. Here $\mu \equiv \mu_s$ is the expected signal event count as defined in Eq. (6.5). The test statistic is

$$q_\mu = 2\big(\ln \mathcal{L}(\hat{\mu}_s,\, \hat{\boldsymbol{\theta}}) - \ln \mathcal{L}(\mu,\, \hat{\hat{\boldsymbol{\theta}}}(\mu))\big), \tag{6.6}$$

where $\hat{\mu}_s$ and $\hat{\boldsymbol{\theta}}$ are the unconditional maximum-likelihood estimators and $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ is the profiled estimator under the tested hypothesis μ. For each mass point and each signal hypothesis μ, the 90th percentile of the q_μ distribution is determined from toy MC ensembles, defining a threshold curve $t(\mu)$. The upper limit is the value of μ at which the observed q_μ first exceeds $t(\mu)$.

For masses up to 1.5 MeV the signal hypothesis grid spans $\mu \in [0, 20]$ over 15 non-uniform scan points. For masses of 2 MeV and above it extends to $\mu = 100$, with 3000–5000 pseudo-experiments per grid point. This replaces the asymptotic approximation, which is unreliable in a regime where the total expected background is only $\sim$1.34 events. The fitted signal count is converted to a limit on the absorption cross-section Y using the predicted signal rate at that mass.

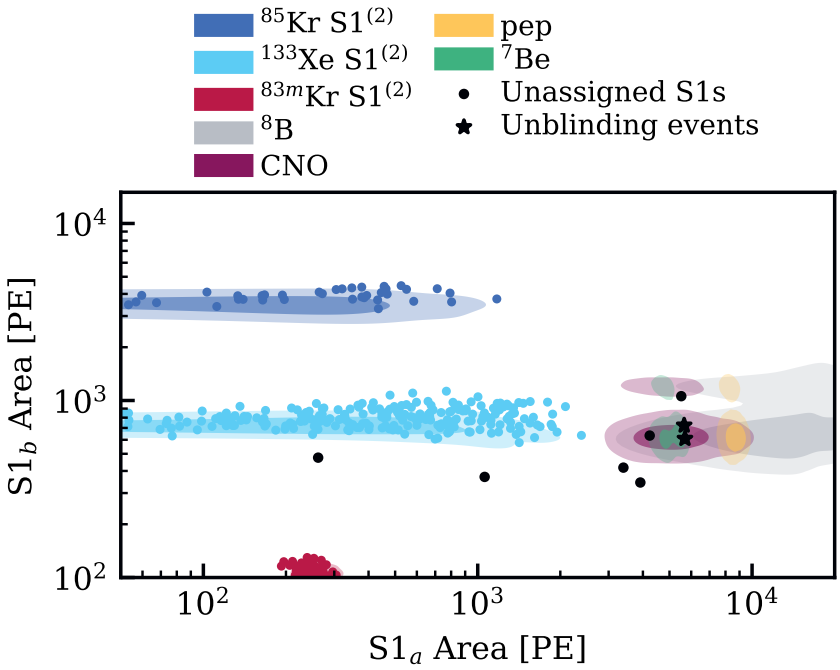


Figure 6.13 *Full unblinded dataset in the $(S1_a, S1_b)$ plane. Coloured contours and scatter points show the three calibration control populations at the 68% and 95% probability levels, as in Figure 6.6. Shaded contours show the four solar-neutrino CC background templates entering the inference (^{7}Be, CNO, pep, ^{8}B) at the 68% and 95% probability levels, each propagated through the full simulation and selection chain. Black circles mark the six events not attributed to any modelled calibration source and outside the previously blinded ^{7}Be solar-neutrino region. All six are observed in SR1, with none in SR0. Star markers indicate the two events in the previously blinded ^{7}Be region, one from SR0 and one from SR1, consistent with solar-neutrino CC interactions [138].*

The expected sensitivity is derived from 2000 background-only pseudo-experiments at each mass point. For each pseudo-experiment, a 90% CL upper limit is computed using the same non-asymptotic threshold curve, and the 16th, 50th, and 84th percentiles (and 2.5th and 97.5th for the outer band) of the resulting distribution define the expected bands on Y.

6.5 results

6.5.1 *Observed Data*

The blinding was partial by design (Section 6.2.3): the sideband shared with the parallel solar-neutrino measurement includes regions where higher-mass FDM signals would appear. The FDM selection and inference were frozen before the blinded region was opened. Unblinding was performed as part of the solar-neutrino CC search [138], and the FDM limits were computed from the same unblinded dataset without further modification. The solar-neutrino CC measurement on ^{136}Xe is the central topic of [138]. The FDM search is a complementary component of the same analysis.

The combined SR0+SR1 dataset consists of SR0 (3.94 t fiducial mass, 82.0 days) and SR1 (4.02 t, 201.2 days), giving a natural xenon exposure of 3.1 tonne·yr and a ^{136}Xe exposure of 0.276 tonne·yr at 8.9% natural abundance.

Figure 6.13 shows the full unblinded observable space, now including all four solar-neutrino CC background templates that enter the inference. Eight events are observed that are not attributed to any modelled calibration source. Six of these were never blinded: they are all observed in SR1, with no counterpart in SR0, and are likely unmodelled background events from sources not included in the inference model. Two of the six have $S1_a < 2000$ PE, placing them well outside the support of any signal or background template. They contribute negligibly to the likelihood and do not affect the inference. The other four lie closer to the solar-neutrino contours in $(S1_a, S1_b)$ and could arise from unmodelled background components, including solar-neutrino CC interactions not captured by the four-component background model. A dark matter interpretation cannot be excluded for any individual event, but the four are not clustered under any single mass hypothesis. Their distributions in Δt, spatial position, reconstructed energy, and time within SR1 were examined and show no statistically significant clustering in any variable [138].

The remaining two events fall inside the previously blinded ^{7}Be region, one from SR0 and one from SR1. Both are consistent with charged-current solar-neutrino interactions on ^{136}Xe and are discussed in detail in [138].

6.5.2 *Observed Limits*

Figure 6.14 shows the 90% CL upper limit on Y as a function of m_χ, with expected sensitivity bands from Section 6.4.1. The best median expected limit is $Y \approx 1.3 \times 10^{-51}$ cm^2 at $m_\chi = 1.2$ MeV. At the lowest mass point the median expected limit is 1.4×10^{-51} cm^2, rising through the 0.68–1.2 MeV

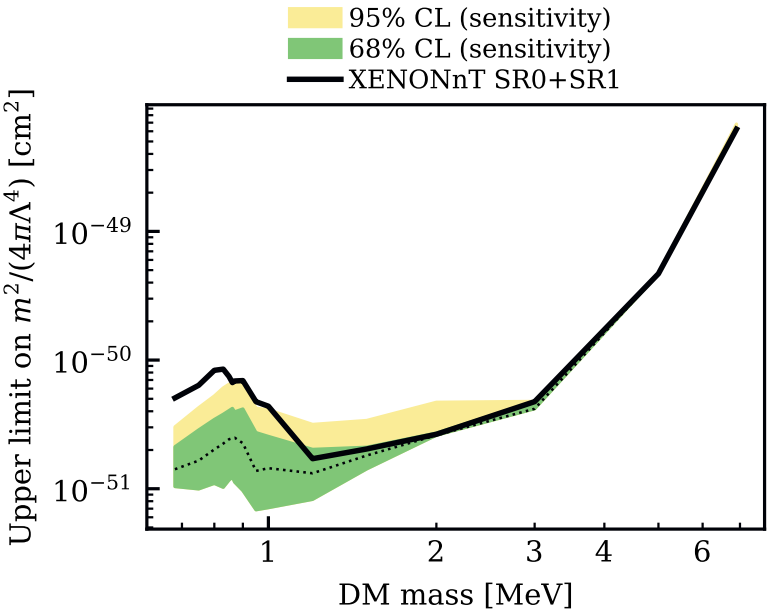


Figure 6.14 *Expected 90% CL sensitivity and observed upper limit on $Y = m_\chi^2/(4\pi\Lambda^4)$ from the XENONnT delayed-coincidence search, as a function of m_χ across the full mass range 0.68–6.9 MeV. Upper limits and sensitivity bands are derived using a non-asymptotic Neyman belt construction: for each signal hypothesis μ, the 90th percentile of the q_μ distribution is determined from toy MC and used as the threshold rather than the asymptotic approximation. The observed limit (solid) is shown with the median expected limit (dotted) and the $\pm1\sigma$ (green) and $\pm2\sigma$ (yellow) coverage bands from 2000 background-only pseudo-experiments. The observed limit is weaker than the median expectation up to $\sim$1.2 MeV, consistent with unmodelled backgrounds in SR1 (Section 6.5.1). Above $\sim$2 MeV the experiment becomes background-free and the median expected limit converges to $\sim$2 events (Section 6.5.2). The steep rise approaching 6.9 MeV reflects the collapse of segmentation efficiency as the prompt $S1_a$ grows to dominate the waveform.*

range before degrading at higher masses as the segmentation efficiency drops. Above $\sim$2 MeV the experiment is background-free and the expected limit converges to $\sim$2 signal events, as shown in Figure 6.16. The median expected limit in this regime is below the classical Poisson value of 2.303 events for zero observed events. The two quantities measure different things: the Poisson 2.303 is the observed limit given zero events, while the median expected sensitivity from the PLR Neyman construction converges to $(Z_{0.9})^2 \approx 1.64$ events in the asymptotic regime [153]. Approaching 6.9 MeV the segmentation efficiency collapses and the curve rises steeply.

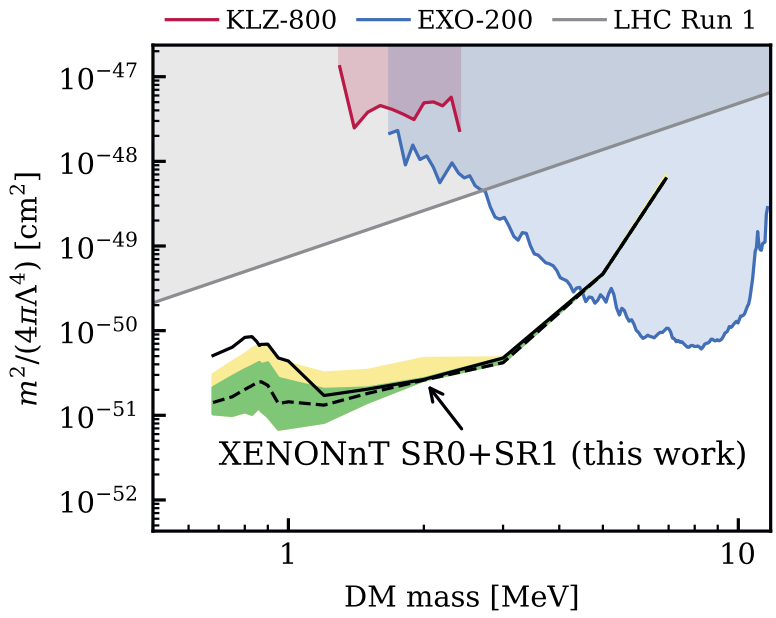


Figure 6.15 *Comparison of 90% CL upper limits on $Y = m_\chi^2/(4\pi\Lambda^4)$ from XENONnT SR0+SR1 (this work, observed limit with $\pm1\sigma$ and $\pm2\sigma$ expected bands), KamLAND-Zen 800 (KLZ-800) [136], and EXO-200 [152]. The grey shaded region indicates parameter space excluded by LHC Run 1 searches [133]. XENONnT achieves the strongest sensitivity in the low-mass region. The weakening of the observed XENONnT limit extends up to ∼1.2 MeV and is consistent with unmodelled backgrounds in SR1.*

At the lowest mass point ($m_\chi = 0.68$ MeV), the observed limit (5.1×10^{-51} cm^2) is weaker than the median expectation (1.4×10^{-51} cm^2) by a factor of approximately 3.6. The weakening extends up to ∼1.0–1.2 MeV. Above 1.2 MeV, the observed and expected limits agree to within the width of the expected bands. The comparison with prior results is shown in Figure 6.15.

The weakening of the observed limit at low mass reflects events not accounted for by the background model. Six unattributed events are observed in SR1 with no counterpart in SR0 (Section 6.5.1). Four of these fall near the solar-neutrino CC contours in $(S1_a, S1_b)$ but are not clustered under any single FDM mass hypothesis. The inference model does not include spallation products, radiogenic neutrons, or accidental coincidences. These were assessed as individually subdominant, but could collectively contribute at the level observed. Above ∼1.2 MeV the limit agrees with the

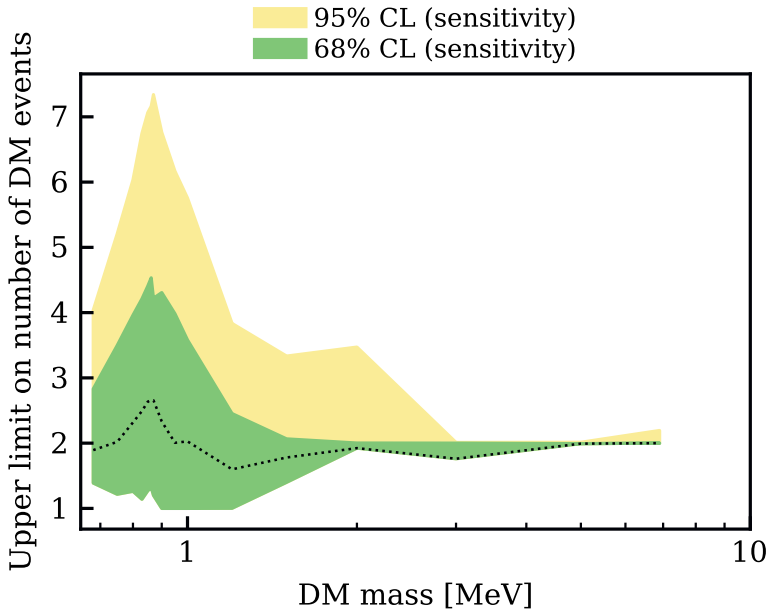


Figure 6.16 *90% CL upper limit on the number of FDM signal events as a function of m_χ, with the same 68% and 95% sensitivity bands as Figure 6.14, derived from 2000 background-only pseudo-experiments using the non-asymptotic Neyman construction. At low masses (0.68–1.5 MeV) the median expected limit (dotted line) ranges from roughly 1.6 to 2.7 events, driven by fluctuations in the expected solar-neutrino background. Above ∼2 MeV the background is subdominant and the median converges to ∼2 events. The width of the bands collapses in this regime because all toys give nearly the same upper limit.*

expected sensitivity and is not affected by these events. The weakening at low mass is attributed to an incomplete background model, not to a signal.

As a robustness cross-check, a separate reduced-observable-space inference was performed with $S1_a \geq 3500$ PE, retaining the full $S1_b$ range. This boundary lies above the labelled calibration ^{133}Xe population. No simulated post-selection ^{133}Xe events survive in the available sample. The same selection was applied consistently to the observed data and to the solar-neutrino, FDM, and individual spallation templates. The remaining spallation products were included explicitly as a merged template with a rate nuisance parameter constrained by an assumed 3% Gaussian width, treated here as an effective cross-check assumption rather than a data-driven precision measurement. The non-asymptotic threshold and sensitivity ensembles for this configuration were generated independently of the primary analysis, following the same procedure described in Section 6.4.1.

In this reduced space, the expected solar-neutrino background is 1.34 events in SR0+SR1, consistent with the primary full-plane expectation, while the total remaining spallation expectation is 0.053 events. The ^{133}Xe contribution is zero after the cut. ^{120}Xe is the largest remaining spallation component. The median expected sensitivity remains within approximately 8% of the primary result across the mass scan, indicating that the expected reach is not materially changed under this reduced-ROI, spallation-inclusive cross-check. The observed low-mass limit becomes stronger in this reduced space, primarily because the cut removes three previously unassigned events below the reduced-ROI boundary. Because the $S1_a \geq 3500$ PE threshold was investigated as a post-unblinding robustness study rather than a pre-defined selection, this result is not used as the primary exclusion limit. This cross-check shows that, after excluding the calibration-^{133}Xe region, explicitly modelling the residual spallation component does not materially change the expected sensitivity. It does not reinterpret the events removed from the primary analysis or replace the primary result.

6.5.3 *Discussion*

Figure 6.15 compares the XENONnT limits to published results from KamLAND-Zen 800 (KLZ-800) [136] and EXO-200 [152]. EXO-200 used a 99.8 kg·yr ^{136}Xe exposure with a one-dimensional energy-based analysis. KLZ-800 performed a delayed-coincidence FDM search covering $m_\chi \in [1.3, 2.4]$ MeV, exploiting the same ^{136}Cs topology in a scintillator-based detector with a different background environment. The present analysis covers $m_\chi \in [0.68, 6.9]$ MeV and achieves the strongest sensitivity across most of this range. In the overlap region with KLZ-800 ($\sim$1.3–2.4 MeV), XENONnT improves on the KLZ-800 limit by approximately four orders of magnitude. Above $\sim$5 MeV EXO-200 sets stronger limits, where the collapse of segmentation efficiency in XENONnT substantially weakens the constraint. Below that mass the deterioration is driven by efficiency loss rather than solar-neutrino leakage.

The larger exposures of future XENONnT science runs and next-generation experiments [46] will extend the reach throughout the mass scan.

6.6 Conclusion

XENONnT has searched for charged-current absorption of fermionic dark matter on ^{136}Xe. The search exploits the delayed-coincidence topology produced when the ^{136}Cs daughter de-excites through low-lying isomeric

states with lifetimes of $\mathcal{O}(100)$ ns. A machine-learning waveform segmentation model decomposes merged S1 pulses into sub-peaks, providing the reconstructed sub-peak areas $(S1_a, S1_b)$ and time separation Δt used in the event selection and inference.

The irreducible background from solar-neutrino CC interactions on ^{136}Xe was modelled with four components (^{7}Be, CNO, pep, ^{8}B), each constrained by a 5% Gaussian prior. All remaining backgrounds (muon spallation, accidental coincidences, and neutron double scatters) are either small in total rate or populate regions of the $(S1_a, S1_b)$ plane away from the solar-neutrino and signal templates, and are excluded from the model. A dedicated reduced-$S1_a$, spallation-inclusive cross-check (Section 6.5.2) indicates that the expected sensitivity is stable under this alternative treatment. It does not replace the primary full-space inference.

After the full event selection, eight events are observed: two in the ^{7}Be region consistent with solar-neutrino CC interactions [138], and six unmodelled events in SR1 discussed in Section 6.5.1.

Upper limits at 90% CL on $Y = m_\chi^2/(4\pi\Lambda^4)$ are set using an unbinned template profile-likelihood in the `alea` framework across 16 mass points from 680.3 keV to 6.9 MeV. Confidence intervals are derived via a non-asymptotic Neyman belt construction: at each signal hypothesis μ, the 90th percentile of the q_μ distribution is determined from toy MC and used as the threshold, rather than the asymptotic approximation. The limits represent the strongest reach to date across most of this mass range.

At $m_\chi = 0.68$ MeV the observed limit is weaker than the median expectation, attributed to unmodelled backgrounds not included in the inference model.

A

XENONNT SOFTWARE: DESIGN AND EXPERIENCE

The XENONnT experiment accumulates around 1 PB of raw data per year of operation, stored across fifteen storage sites on two continents and processed several times with updated calibrations over the course of a Science Run. Managing this at scale, across a collaboration of around 160 physicists in twelve countries and across three Science Runs spanning several years, required building substantial software infrastructure.

This appendix is not a technical manual. It is a case study, written for experimentalists building comparable systems. Section A.1 describes what was built and why. Section A.2 records what operating it taught us.

A.1 THE INFRASTRUCTURE

A.1.1 *The Analysis Framework*

XENONnT's predecessor experiment, XENON1T, used PAX: a monolithic, linearly-ordered processing pipeline with no automatic caching and no intermediate product storage. `strax` [154] was built from scratch to replace it. PAX processed data at $\mathcal{O}(100\,\mathrm{kB/s/core})$; `strax` reaches $\mathcal{O}(10\text{–}100\,\mathrm{MB/s/core})$ [65, 155].

The `strax` framework implements data processing as a directed acyclic graph (DAG) of *plugins*. Each plugin declares the data types it consumes (`depends_on`) and produces (`provides`), along with a `compute()` method. The `Context` object is the central interface for all data access in `strax`. It bundles the registered plugin set that defines the DAG, a list of storage frontends from which products are read and to which they are written, and the full plugin configuration, including the correction version strings for every plugin. When a user requests a data type, the `Context` resolves the dependency graph, checks each frontend in order for a cached product whose lineage hash matches the current configuration, and computes only what is missing. Two contexts that differ in any configuration value produce distinct lineage hashes and therefore fully independent

cached products. All data products are structured numpy arrays processed in time-ordered chunks.

The XENONnT-specific `straxen` [74] implements the plugin chain on top of `strax`. It runs from raw digitiser output (`raw_records`) through waveform reconstruction, peak finding, and position reconstruction to the final `event_info` table used for physics analysis (Figure A.1). `Raw_records` contain the full digitised waveforms from each PMT channel. `Peaklets` are formed by finding and merging hits across channels. They retain the per-peak waveform. `Peak_basics` drop the waveform, storing only integrated quantities: area, timing, and width. `Event_info` pairs S1 and S2 candidates, adds corrected positions and energies, and drops single-peak quantities not relevant at the event level. Waveform-level algorithm development therefore requires `peaklets` or `raw_records`; standard physics analysis works from `event_info`.

Every data product is assigned a lineage hash:

$$h(\text{type}) = H(\text{plugin name},\ \text{plugin version},\ \text{plugin config},\ h(\text{dep}_1),\ h(\text{dep}_2),\ \ldots), \quad \text{(A.1)}$$

where H is a truncated SHA-1 digest encoded as a 10-character alphanumeric string. The hash propagates recursively through all dependencies and is embedded in the filesystem path of every stored data product:

`078621-event_info-`**`kp92mxwt4f`**`/`

Any change anywhere in the processing chain changes the hash of all downstream data types. `strax` will not serve a cached product whose hash does not match the current configuration. Different processing versions of the same run coexist safely on the same storage system. The experiment's reproducibility guarantee is not conventional; it is enforced automatically by the hash. Content-addressed caching of DAG outputs is an established principle in build systems and data versioning tools. `strax` applies the same idea to physics data processing [156].

Science analyses run inside the `xenonnt_offline` context, defined in `cutax` (a private package that adds collaboration-internal selection plugins and analysis cuts). It builds on the base `straxen` plugin chain and Midway3 storage configuration, and applies a pinned `global_vN` correction state via `xedocs` (Section A.1.2).

Simulation and real data share the same analysis chain from `raw_records` onward. The `fuse` framework (Chapter 3) is a native `strax` plugin chain that takes Geant4 energy deposits and produces `raw_records` in the same binary format as the real DAQ output. The

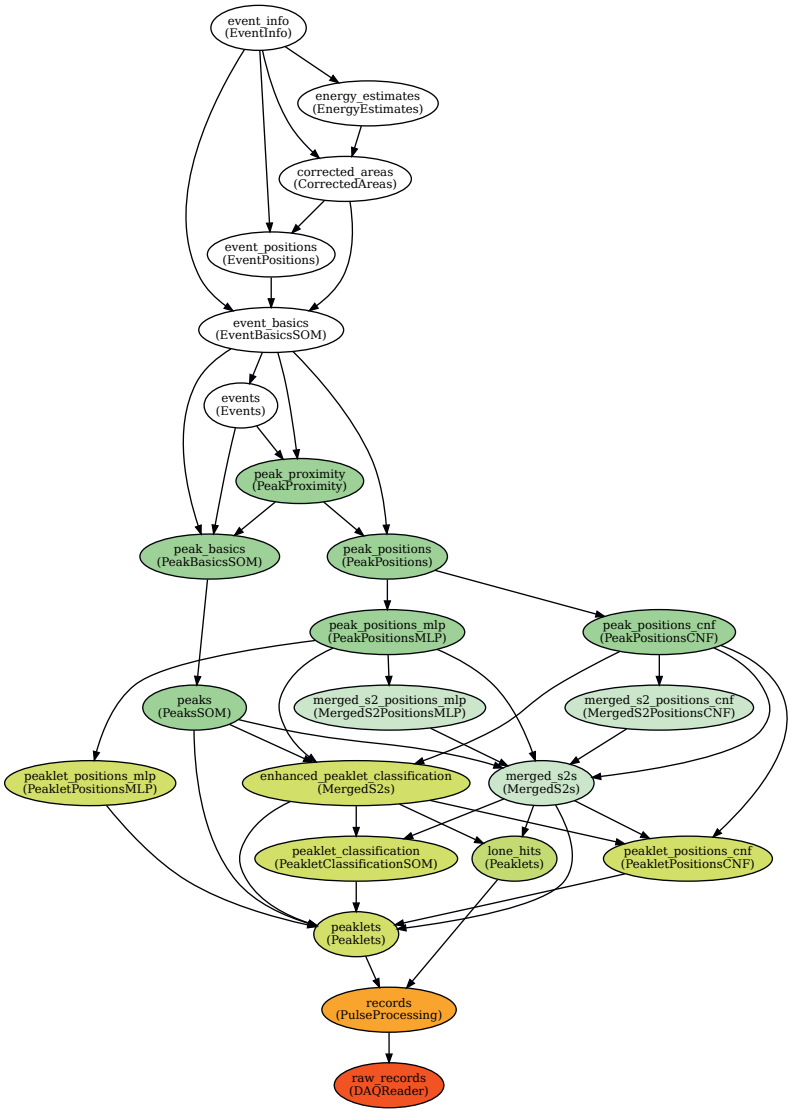


Figure A.1 *The XENONnT data type hierarchy for the main TPC. Data types progress from raw digitiser output (*`raw_records`*) at the bottom to the analysis-level* `event_info` *at the top. Each box shows the data type in lowercase, the name in parentheses is the plugin class that produces it. The arrows indicate the dependency graph: each plugin points to the data types it consumes. Colour indicates data kind: plugins sharing a colour produce data of the same type and can be loaded together as a single array.*

identical `straxen` plugin chain and corrections then apply to both. Sys-

tematic effects from reconstruction and corrections are treated consistently across data and simulation by construction.

A.1.2 *Corrections and Reproducibility*

Several physical effects require time-dependent, three-dimensional corrections: PMT gain variations, electron lifetime, drift field non-uniformity, and position-dependent light collection efficiency. These are stored as versioned JSON files in the `corrections` GitHub repository and synchronised to the `xedocs` [157, 158] MongoDB database by a CI/CD pipeline. MongoDB is a derived artefact; the repository is the source of truth, and direct writes to the database are not permitted. Every correction change is reviewed via pull request and permanently recorded in git history.

Versioned corrections are immutable for past time intervals. A correction found to be wrong is superseded by a new version, not corrected in place. Analyses using the old version are not invalidated; they refer to a distinct processing state, identified on disk by a different lineage hash.

Plugin parameters that depend on calibrations are encoded as URL-like strings that include the correction version explicitly. This version string enters Eq. (A.1) as part of the plugin configuration, so changing a correction version automatically produces a distinct cached data product downstream. No manual bookkeeping is required.

A `global_vN` snapshot is a JSON file that pins one correction version per plugin parameter:

```
[
  {"version": "global_v18",
   "config_name": "elife",
   "value": "xedocs://electron_lifetimes?attr=value
             &run_id=plugin.run_id&version=v11"},
  {"version": "global_v18",
   "config_name": "s1_xyz_map",
   "value": "xedocs://s1_xyz_maps?...&version=v9"},
  ...
]
```

Initialising a context with a global version atomically applies the entire correction state. A Science Run is defined by the conjunction of a run list, a software container pinning the Python environment, and a global corrections version. Given these three inputs, any published XENONnT result can be exactly reproduced.

A.1.3 *Data Acquisition and Storage*

XENON1T used a software trigger: a real-time algorithm decided which waveform windows to retain, and everything else was discarded. The trigger decision was made before the data were fully understood. Any signal type not anticipated by the trigger criteria was lost, and revising event selection later required raw data that no longer existed.

XENONnT adopted the opposite approach. All digitised samples above threshold are saved, and event finding is moved entirely to offline processing. This is practical because the data volume is dominated by a small number of large signals: single-electron S2s account for about 30% of reconstructed signals but only 2% of the data volume [65]. The same approach has been adopted by other liquid noble dark matter experiments [159, 160], reflecting a field-wide convergence on the same trade-off.

The acquisition software, REDAX, writes compressed data in 5-second chunks to a buffer filesystem. An online processing daemon, `bootstrax`, processes new chunks as they arrive using the best available calibration at the time of data taking. Online results are used for detector monitoring only. Science analysis uses a frozen correction state applied during offline reprocessing. Every run and every computed data product are registered in a central MongoDB database; the lineage hash of each product is embedded in the run record.

Long-term storage and distribution is managed by Rucio [161] across fifteen storage elements spanning Europe and the United States. The data volumes involved are substantial (Figure A.2). Across all three Science Runs, `raw_records` total 9.4 PB uncompressed; the waveform data compresses well, to approximately 30% of the uncompressed size, giving roughly 3 PB on disk. `Raw_records` are always stored in two copies, on tape and on disk, at distributed Rucio Storage Elements (RSEs). `Peaklets` total 2 PB uncompressed; the compression ratio for derived data types depends on their content and is not as favourable as for raw waveforms. Analysis-level products (`peak_basics`, `event_info`) total approximately 20 TB across all Science Runs.

High-level data products (`peak_basics`, `event_info`, and related types) are stored on the primary analysis cluster, Midway3, and are immediately accessible. `raw_records` and `records` are held on distributed tape and disk at RSEs across Europe and the United States and require an explicit transfer to access. `Peaklets` occupy a middle position: they are needed for any high-level reprocessing campaign, but the ≈300 TB of project storage on Midway3 was rarely sufficient to hold even a single Science Run's worth, meaning that high-rate calibration data had to be

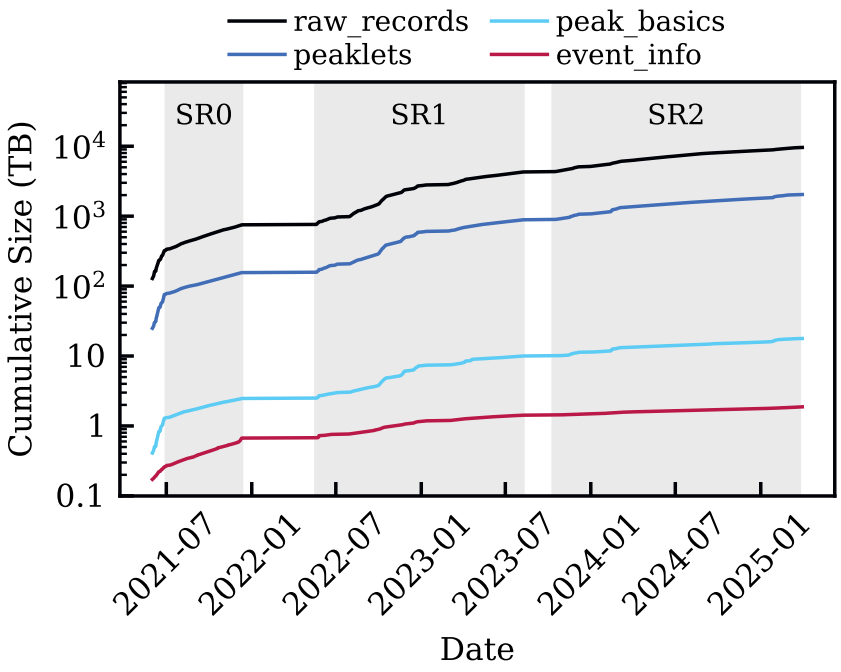


Figure A.2 *Total uncompressed data volume per data type for each Science Run.* `raw_records` *dominate by volume and are always stored in two copies (tape and disk) at distributed storage elements.* `peaklets` *are required for high-level reprocessing and ideally would be kept locally, but at* 1.5 PB *for the largest Science Run they exceeded available local storage. Analysis-level products (*`peak_basics`*,* `event_info`*) are small enough to fit comfortably on the primary analysis cluster. All sizes are uncompressed.* `Raw_records` *compress to approximately 30% of the uncompressed size; compression ratios for other data types vary.*

moved out to make room and that reprocessing campaigns often had to stage peaklets from remote storage elements before proceeding.

The development of a Science Run dataset follows an iterative sequence of correction development and reprocessing campaigns illustrated in Figure A.3. PMT gain corrections are finalised first, enabling a low-level reprocessing from `raw_records` to produce a clean peaks dataset. Subsequent corrections trigger high-level reprocessing campaigns from peaks. After several iterations, a final global version is defined and the Science Run dataset is frozen.

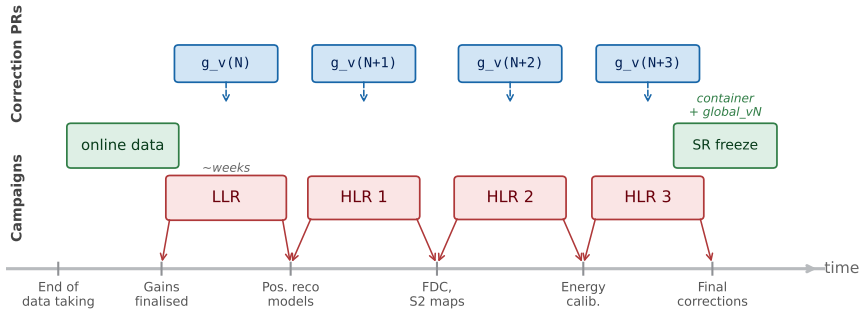


Figure A.3 *The iterative corrections-development and reprocessing cycle for a Science Run. PMT gain corrections are finalised first, enabling an initial low-level reprocessing campaign from* `raw_records`*. Subsequent correction development each produces a new* `global_vN` *snapshot and triggers a high-level reprocessing campaign from peaks. The final* `global_vN` *defines the frozen Science Run dataset.*

A.2 Observations

The following reflect personal observations from operating this system across three Science Runs. They are not a formal evaluation of the collaboration's choices, many of which were made under constraints that are easier to see in retrospect than they were at the time.

A.2.1 *What Worked*

The `strax` plugin model proved its design in practice. Declaring dependencies explicitly through `depends_on` and `provides` made it straightforward to substitute individual components of the analysis chain without touching the rest. New reconstruction algorithms and correction schemes could be tested in isolation by swapping a single plugin or loading a different context. When `fuse` was built as a native `strax` plugin chain, the same flexibility extended to simulation. Analysts working on detector modelling used the same interface as for real data, with the same correction versioning and the same lineage guarantees. Accessible, modular software lowers the cost of entry for new contributors, and the simulation programme drew contributors from across the collaboration as a result.

Versioned corrections behaved as designed throughout all three Science Runs. The `xedocs` [157, 158] package, built on `rframe`, stored corrections with explicit version history, enforced immutability for past time intervals, and made correction changes propagate automatically through the lineage hash. The distinction between online corrections, used by `bootstrax` for

real-time monitoring, and offline versioned corrections, the anchors for Science Run reprocessing, kept the two use cases cleanly separated. Every correction change passed through a pull request review in the `corrections` GitHub repository before being written to MongoDB by a CI/CD pipeline, providing a social check on correction quality that matched the technical guarantees of the system.

MongoDB operated as the experiment's central state: run registry, data location index, corrections database, online monitoring store, and membership management from a single system. Serving multiple purposes from a single database simplified operations and reduced the number of moving parts.

Beyond the core analysis framework, the collaboration built a broad ecosystem of supporting tools over the course of the experiment. Detector monitoring dashboards, run selection interfaces, data quality displays, and shifter support tools accumulated across the Science Runs, with more services added as the experiment matured and the complexity of tracking detector performance and data quality grew. These were maintained and used regularly by large numbers of collaborators. The fact that they worked reliably reflected sustained effort from the people who built and looked after them.

A.2.2 *The Reprocessing Bottleneck*

A full low-level reprocessing campaign, starting from `raw_records` and producing a clean dataset for an entire Science Run, took several months from start to usable output. The bottleneck was not only computing power. It was operating the grid correctly, managing the bookkeeping across hundreds of concurrent jobs, and finding somewhere to put the output.

Low-level reprocessing jobs read `raw_records` from RSEs distributed across Europe and the United States and write `peaklets` back to the storage network. Running this correctly required knowing which grid resources were available, how to route jobs to sites that held the relevant input data, how to diagnose failures and resubmit, and how to track what had and had not completed. This work is largely independent of the physics analysis it enables. Without a complete picture of available resources and data locations, jobs were often submitted to suboptimal sites, adding unnecessary transfer overhead.

Reprocessing campaigns typically fell to a single PhD student, alongside their physics analysis responsibilities. The expertise needed accumulated slowly and was largely specific to the person who held it at any given time.

A.2.3 *The Missing Analysis Layer*

`strax` provides a correct way to load data for a single run. Loading data for a physics analysis requires something beyond that: consistent context configuration across thousands of runs, job submission, and output storage with metadata attached. Tools were built to address parts of this, but none became a general solution adopted across the collaboration.

The canonical path worked in principle. A context initialised with the right global corrections version and the right software container would locate the reprocessed data on Midway3 and return accurate results. In practice, it was fragile. Different computing nodes on the cluster had different access to the storage backends where reprocessed data lived, so the same script succeeded on some nodes and failed on others. The system did not warn when the software container was out of date. The system resolved the request and returned internally consistent results with no indication of whether the correction version matched the current Science Run dataset.

Without a shared tool that enforced these details, analysts worked around them individually. Some built private frameworks: structured scripts for context configuration and multi-run loading, with their own output conventions. These worked reliably for their authors. None became a collaboration standard, and the expertise they encoded did not transfer. Others bypassed `strax` by saving processed output to files and passing those directly between analysts. This reduced the traceability the infrastructure was designed to provide. Two datasets that appeared equivalent could rest on different correction versions, with no warning from the system. The provenance was technically encoded in the filename, but reading it required a database query and knowledge of the producing plugin configuration.

The same dynamic played out for simulation. A standard configuration mechanism existed: a context combined with a simulation file set all relevant parameters, and the lineage hash encoded the full configuration. In principle this was sufficient. In practice, the differences between configurations used by different groups were not always clearly communicated, and the documentation describing what each option controlled was uneven. More concretely, what was absent was a centralised production of standard simulation datasets. Each time detector understanding advanced or a new correction was deployed, each analysis team reran its own background and calibration simulations independently. This was time-consuming. The incentive structure was not purely negative: independent production meant that many analysts engaged directly with `fuse` and developed familiarity with the detector model. The absence of a central production step meant that effort was duplicated across groups, and cross-analysis compar-

isons required first establishing that the underlying configurations were equivalent.

Both paths had the same root cause: no shared, endorsed way to load an analysis dataset at scale existed. The practical effect was that context configuration, run selection, and output tracking were handled differently by different groups, and the guarantees the infrastructure was designed to provide were not consistently realised at the analysis level.

A.2.4 *Shared Storage Constraints*

The ≈300 TB of project storage on Midway3 was a significant constraint on what the collaboration could do scientifically, not just operationally.

Peak waveforms for a complete Science Run could not be stored centrally. Algorithm development requiring waveform-level access meant downloading raw data from tape to local scratch space: a process that was slow, consumed limited quota, and failed for a subset of runs often enough to require manual intervention.

Storage limits interact with reprocessing in a second way. When space ran out, old processing versions had to be deleted to make room for new ones. This forced a choice between keeping multiple reprocessing versions available for comparison and having disk space for ongoing work. The choice was always resolved in favour of ongoing work, which meant the ability to compare results across correction versions was lost at the moments when it would have been most useful.

Storage scarcity also imposed a scientific priority ordering. The WIMP search and the CEνNS analysis ran first, absorbing available storage during the period when the dataset was most current. Secondary analyses arrived later, once the primary effort had concluded. High-energy searches are a typical example. By that point, data had often been partially offloaded to tape to make room for the next Science Run. Secondary analyses frequently required custom processing, sometimes a full low-level campaign from `raw_records`. With no available space on shared storage, analysts ran these campaigns at home institute clusters. The output was then accessible only to that analyst. Every collaborator who needed the result depended on one person for access to data that had no official home in the shared infrastructure.

In practice, available storage shaped which analyses were attempted and when, independently of their physics priority.

A.2.5 *Onboarding Scales with Proximity*

The collaboration held newcomer workshops every year covering the analysis chain, the corrections system, and the computing environment. They were well prepared and, for most members, the primary introduction to the analysis environment. The problem was coverage and frequency: not everyone attended, and one workshop per year rarely reached new members at the moment they needed the material.

The material itself compounded the difficulty. The concepts are too interconnected to absorb from a week of lectures. Understanding what happens when a correction version changes requires simultaneously holding the version encoding, the lineage hash propagation, the global version structure, and the storage backend logic in mind. That mental model has to be built by doing. Analysts with experienced colleagues nearby built it faster. Working visits between institutes transferred practical knowledge that no scheduled workshop replicated: spending a few weeks doing analysis alongside people who already understood the system, with questions driving the work rather than a curriculum.

Documentation lagged behind the system throughout the experiment's lifetime. The people best placed to write it were the most overcommitted. Code-level documentation was at least partially recoverable from the codebase, but design rationale was not. Why corrections are encoded the way they are, what had been tried and abandoned: that knowledge lived with the people who built the system. When they graduated, it could not be recovered from the code. The sessions that came closest to preserving it were informal, experienced members explaining design decisions to incoming ones in real time. That kind of exchange was hard to schedule and harder to record.

A.2.6 *Software Governance*

The XENONnT software stack grew into a large collection of interdependent packages: `strax` [154], `straxen` [74], `cutax`, `xedocs` [157], `rframe` [158], `corrections`, `utilix` [162], `outsource` [163], container infrastructure, and more. Each package was critical to some part of the analysis chain. Each needed continuous maintenance as detector understanding and analysis requirements evolved.

During periods when a designated maintainer existed for a given repository, the development process was clear. Pull requests were reviewed promptly, merges were deliberate, and the implications of changes for downstream packages were considered before they landed. During the long

intervals without a designated maintainer, responsibility was undefined and pull requests accumulated.

The pattern was consistent: defined ownership produced reliable outcomes; undefined ownership did not.

A.3 Conclusions

Across the observations above, the technical infrastructure largely worked as designed. The `strax` plugin model, the versioned corrections system, and the content-addressed caching all behaved correctly and enabled physics that would not otherwise have been possible. The difficulties were mostly in the gap between what the infrastructure could do and what it took to operate it at scale: reprocessing campaigns, analysis tooling, storage planning, and the transfer of knowledge across cohorts.

These are not problems specific to XENONnT. They are a consequence of running a large data-intensive experiment on academic timescales, with rotating student cohorts and constrained resources. Recording them here is not meant as a critique of the collaboration's choices, but as a description of what operating this kind of system actually looked like from the inside.

SAMENVATTING

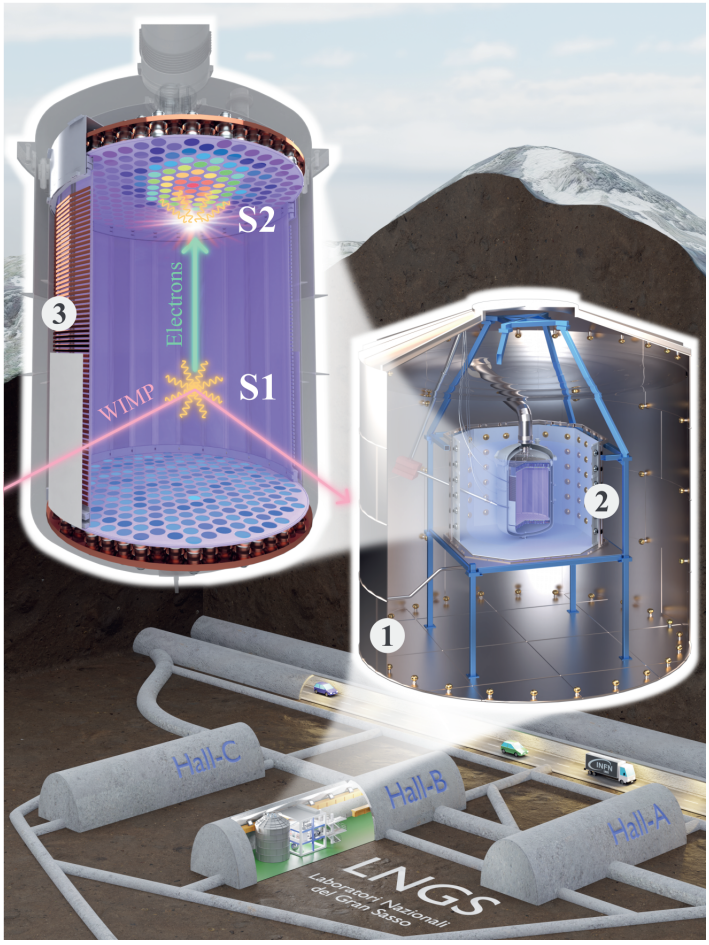
van "A Matter of Coincidences"

"Come sta il barile?" is de vraag die mij de afgelopen jaren op elk familiefeest is gesteld. Het *barile* (het vat) is de bijnaam die mijn familie heeft gegeven aan de detector waarmee ik tijdens mijn promotieonderzoek heb gewerkt: het XENONnT-experiment. Mijn gebruikelijke antwoord is dat het vat het prima maakt en onder een berg ligt te niksen. Helemaal onwaar is dat niet, maar in dit proefschrift probeer ik het wat uitgebreider uit te leggen.



Het vat zelf is een tank vol vloeibaar xenon. Het staat diep onder een berg in de ondergrondse Laboratori Nazionali del Gran Sasso in Midden-Italië, binnen een grotere watertank voor extra afscherming (Figuur 1). Het is gebouwd om een deeltje donkere materie te vangen: de onzichtbare stof die het grootste deel van de materie in het heelal vormt, waarneembaar via haar zwaartekracht maar nooit direct gedetecteerd. Men denkt dat donkere materie ons sterrenstelsel vult en vrijwel ongestoord door gewone materie heen beweegt. Omdat ze nauwelijks ergens mee interactie heeft, gaat ze moeiteloos door een berg rots heen. We hopen dat ze af en toe botst met een van de zware xenonatomen in onze detector.

Tegelijkertijd hebben ook andere deeltjes voortdurend interacties met het xenon, en die veroorzaken ongewenste signalen. Een deel daarvan komt van kosmische straling, al houdt de berg boven het laboratorium het



Figuur 1 *De XENONnT-detector in de Laboratori Nazionali del Gran Sasso (LNGS), Italië. De tank met vloeibaar xenon (3) wordt omringd door lichtsensoren aan de boven- en onderzijde. Een deeltjesinteractie produceert de lichtflits die we S1 noemen, en een tweede, grotere flits, S2, zodra de vrijgemaakte elektronen omhoog driften en de gaslaag bereiken. De detector staat in een watertank (2) voor afscherming, in Hal B van het ondergrondse laboratorium (1). Figuur overgenomen uit [2].*

meeste daarvan tegen voordat het de detector bereikt. Een ander deel komt van neutrino's uit de zon. Elk neutrino heeft maar zelden een interactie, maar de zon zendt er zo veel uit dat sommige toch het xenon raken. Als dat gebeurt, kan het signaal sterk lijken op wat we van donkere materie verwachten.

De belangrijkste bron van deze achtergrond zit echter in de directe omgeving van de detector zelf. Gewone materialen, zoals het gesteente en

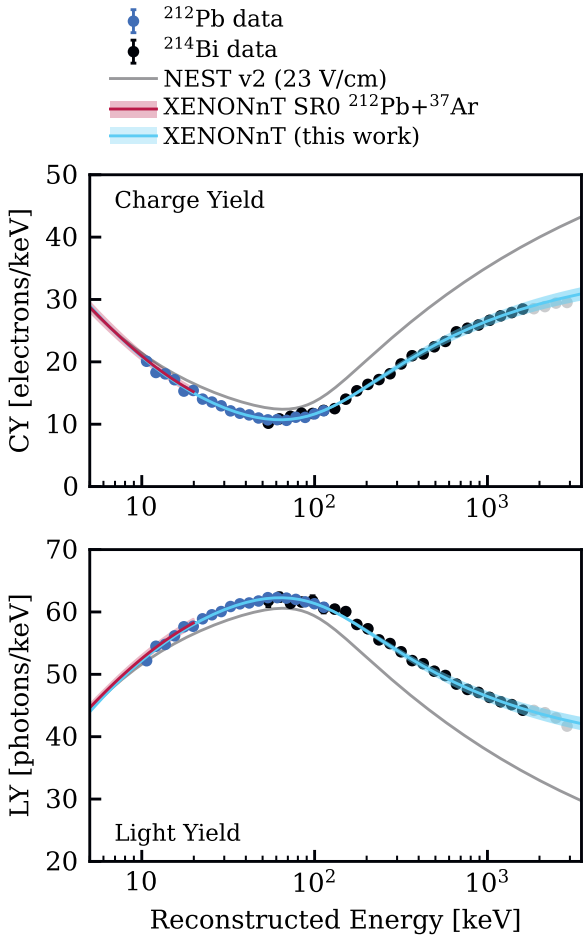
de onderdelen van de detector, zijn licht radioactief, en ook het vloeibare xenon is niet volledig zuiver: er zitten sporen van radon in opgelost, een radioactief gas. Radioactieve atomen zijn instabiel en vervallen één voor één, waarbij ze alfadeeltjes, elektronen of gammastraling uitzenden, en soms ook neutronen. De watertank rond de detector houdt het grootste deel van de neutronen tegen, maar de rest van deze straling bereikt het xenon voortdurend. Om dit onder controle te houden was extreme zorgvuldigheid nodig bij de keuze en reiniging van elk materiaal in de detector, waardoor XENONnT een van de schoonste en stilste plekken op aarde is.

Wanneer een deeltje in het xenon interageert, gebeuren er twee dingen: elektronen worden vrijgemaakt uit xenonatomen, en er ontstaat een lichtflits. Sensoren rond de tank zien deze flits onmiddellijk. We noemen dit eerste signaal S1. De vrijgemaakte elektronen driften vervolgens onder invloed van een elektrisch veld omhoog door de vloeistof. Enkele microseconden later bereiken ze de gaslaag boven de vloeistof en veroorzaken daar een tweede, grotere flits, die we S2 noemen. Samen vertellen S1 en S2 ons waar de interactie plaatsvond en hoeveel energie daarbij werd afgegeven. Dit proces, en de onderdelen van de detector die het mogelijk maken, van de fotomultiplicatorbuizen tot de elektronica die S1 en S2 reconstrueert, zijn het onderwerp van Hoofdstuk 2. Om het nauwkeurig te kunnen interpreteren is ook een betrouwbare simulatie van de detector nodig. Hoofdstuk 3 beschrijft `fuse`, het simulatieraamwerk dat het volledige proces reproduceert, van een deeltjesinteractie tot een gedigitaliseerde golfvorm, en dat gesimuleerde data oplevert die nauwelijks van echte data te onderscheiden zijn.

XENONnT is de nieuwste stap in een zoektocht naar donkere materie die de XENON-collaboratie al bijna 25 jaar voert, met een reeks detectoren (XENON10, XENON100 en XENON1T), elk groter en schoner dan de vorige. De huidige versie bevat 8,6 ton xenon en draait sinds 2020, zonder tot nu toe donkere materie te hebben gevonden.

In afwachting van de ontdekking van donkere materie is dezelfde detector ook voor andere doeleinden nuttig. De zeer lage achtergrond en het vermogen om S1 en S2 met hoge precisie te meten maken hem waardevol voor veel meer dan de zoektocht waarvoor hij is gebouwd. In dit proefschrift gebruiken we data van de detector voor drie doelen: om te modelleren hoe vloeibaar xenon energie omzet in licht en lading, om het energiespectrum van een nucleair verval nauwkeurig genoeg te meten om verschillende theorieën over kernstructuur te kunnen onderscheiden, en om te zoeken naar donkere materie via een alternatieve interactie, absorptie in plaats van verstrooiing.

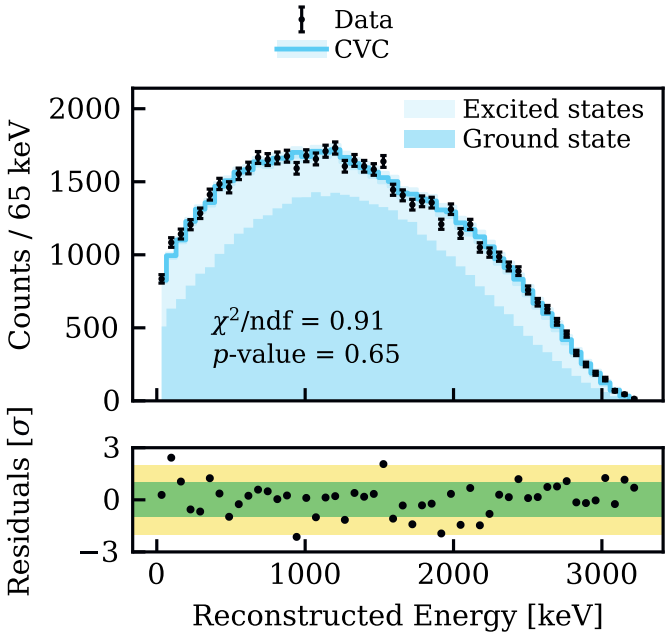
Alle drie de analyses maken gebruik van dezelfde truc om achtergrond te onderdrukken. Telkens produceert het proces waarnaar we zoeken



Figuur 2 *Verdeling van een gegeven hoeveelheid afgezette energie over licht (LY, onder) en lading (CY, boven) in vloeibaar xenon, bepaald uit de ^{214}Bi-vervallen die geselecteerd zijn zoals beschreven in de tekst. Deze meting corrigeert een aanzienlijke discrepantie in eerdere modellen en wordt in dit hele proefschrift gebruikt.*

niet één, maar twee signalen: een eerste signaal gevolgd, na een korte en voorspelbare tijd, door een tweede. Gewone achtergrond doet dit vrijwel nooit, dus door beide signalen te vereisen, in de juiste volgorde en binnen het juiste tijdvenster, wordt bijna al die achtergrond in één keer verwijderd. Deze techniek staat bekend als vertraagde coïncidentie (*delayed coincidence*).

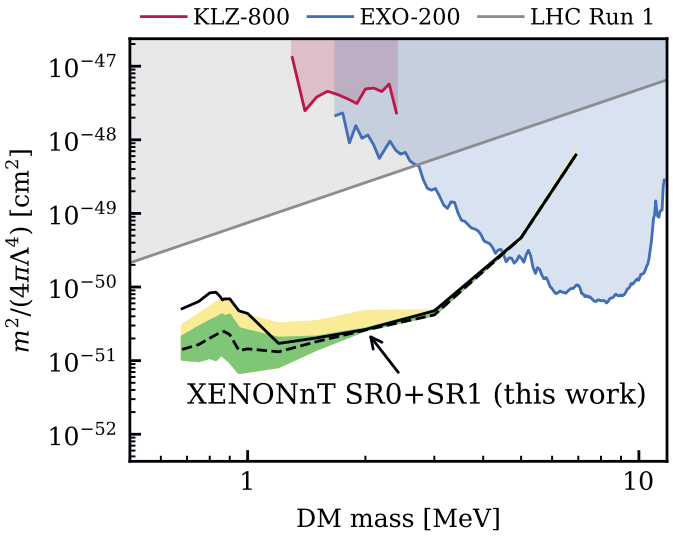
Een van de vervalketens van radon is bijzonder nuttig. Een ^{214}Bi-atoom vervalt door een elektron uit te zenden en verandert daarbij in ^{214}Po, dat op



Figuur 3 *Gemeten ^{214}Bi-betaspectrum, vergeleken met een `fuse`-simulatie van de volledige keten op basis van het algemene symmetrieargument (CVC) uit de tekst. Dit model komt het best overeen met de data. Onderaan staan de residuen. Twee gedetailleerdere kernstructuurberekeningen, hier niet getoond, wijken sterker af van de data.*

zijn beurt een fractie van een milliseconde later vervalt door een alfadeeltje uit te zenden. Dit is de signatuur die we gebruiken: wanneer we dit paar waarnemen, weten we dat we een ^{214}Bi-verval hebben gevangen, duidelijk te onderscheiden van alles wat er verder in de detector gebeurt. Hoofdstuk 4 gebruikt deze geselecteerde dataset om te bepalen hoe een gegeven hoeveelheid afgezette energie zich verdeelt tussen het licht- en ladingssignaal in vloeibaar xenon, een verhouding die eerdere modellen bij deze energieën flink verkeerd voorspelden (Figuur 2). Deze verhouding goed bepalen is belangrijk voor elke andere meting in dit proefschrift, omdat ze de meetlat vormt waaraan al het overige wordt afgemeten.

Met deze gekalibreerde meetlat gaat Hoofdstuk 5 een stap verder en onderzoekt in detail hoe het betaverval er precies uitziet: niet alleen hoeveel energie het gemiddeld vrijgeeft, maar het volledige spectrum van energieën die het elektron kan meenemen. Dat spectrum is geen eenvoudige curve. De exacte vorm hangt af van de interne structuur van de atoomkern op een manier die moeilijk te berekenen is, en verschillende kernmodellen zijn



Figuur 4 *Bovengrenzen (90% betrouwbaarheidsniveau) op de sterkte van het hier onderzochte absorptieproces van donkere materie, als functie van de massa van donkere materie, uit dit werk en uit andere experimenten die gevoelig zijn voor hetzelfde proces. In een deel van het massabereik verbetert deze meting de voorheen beste limiet met maximaal vier ordes van grootte.*

het flink oneens over hoe die vorm eruit zou moeten zien. Hoofdstuk 1 beschrijft de kernfysica hierachter: waarom de vorm moeilijk te voorspellen is, en wat een nauwkeurige meting ervan ons over de kern kan leren. Met dezelfde geselecteerde dataset van schone ^{214}Bi-vervallen hebben we dat spectrum met voldoende precisie gemeten om de concurrerende modellen van elkaar te onderscheiden (Figuur 3).

Hoofdstuk 6 keert terug naar donkere materie, maar vanuit een andere invalshoek. De meeste experimenten zoeken naar een deeltje donkere materie dat tegen een xenonkern botst en die een klein zetje geeft, zoals de ene biljartbal de andere aanstoot. Voor zeer lichte donkere materie is dat zetje te zwak om waar te nemen. Een alternatief is absorptie: het deeltje donkere materie wordt niet afgebogen maar vernietigd, waarbij het zijn volledige massa in één keer aan de kern overdraagt. In xenon kan dit een aangeslagen kern achterlaten die kort daarna tot rust komt door een extra, vertraagde lichtflits uit te zenden, naast het oorspronkelijke signaal. Door die tweede, vertraagde flits te vereisen wordt vrijwel alle gewone achtergrond weggefilterd die de zoektocht anders zou overspoelen. Het is

dezelfde truc die de betavervalmeting mogelijk maakte, nu toegepast op een andere vraag in de natuurkunde.

Er werd geen signaal van donkere materie gevonden. In een deel van het massabereik zijn de limieten die deze zoektocht stelt aan de interactiesterkte echter ongeveer tienduizend keer strenger dan wat voordien bekend was (Figuur 4).

Het proefschrift sluit af met een appendix over de software: de infrastructuur die de data van XENONnT verwerkt, opslaat en verspreidt, en die door veel mensen binnen de collaboratie over meerdere jaren is opgebouwd en draaiende gehouden. Deze appendix geeft een korte beschrijving van de belangrijkste software en infrastructuur, gevolgd door enkele persoonlijke observaties over het gebruik en onderhoud ervan in de praktijk.

Het vat, zo blijkt, had wel wat beters te doen dan daar maar te liggen wachten.

RIASSUNTO

di "A Matter of Coincidences"

"Come sta il barile?" è la domanda che mi viene posta a ogni pranzo di famiglia da qualche anno a questa parte. Il *barile* è il nome scherzoso che la mia famiglia ha dato al rivelatore con cui ho lavorato durante il dottorato: l'esperimento XENONnT. La mia risposta abituale è che il barile sta bene: se ne sta sotto una montagna senza fare niente. La verità non è poi così lontana, ma in questa tesi cercherò di dare una risposta più dettagliata.



Il barile stesso è un serbatoio riempito di xenon liquido, sepolto sotto una montagna nei Laboratori Nazionali del Gran Sasso, nell'Italia centrale, e collocato all'interno di un serbatoio d'acqua più grande per un'ulteriore schermatura (Figura 1). È stato costruito per catturare una particella di materia oscura: la sostanza invisibile che costituisce la maggior parte della materia nell'universo, di cui percepiamo gli effetti gravitazionali ma che non è mai stata osservata direttamente. Si ritiene che la materia oscura riempia la nostra galassia e attraversi la materia ordinaria quasi indisturbata. Proprio perché interagisce così poco, passa senza difficoltà attraverso la roccia della montagna. Speriamo che, ogni tanto, urti uno dei pesanti atomi di xenon nel nostro rivelatore.

Allo stesso tempo, molte altre particelle interagiscono di continuo con lo xenon e producono eventi indesiderati. Una parte di questi eventi è causata dai raggi cosmici, anche se la montagna sopra il laboratorio ne

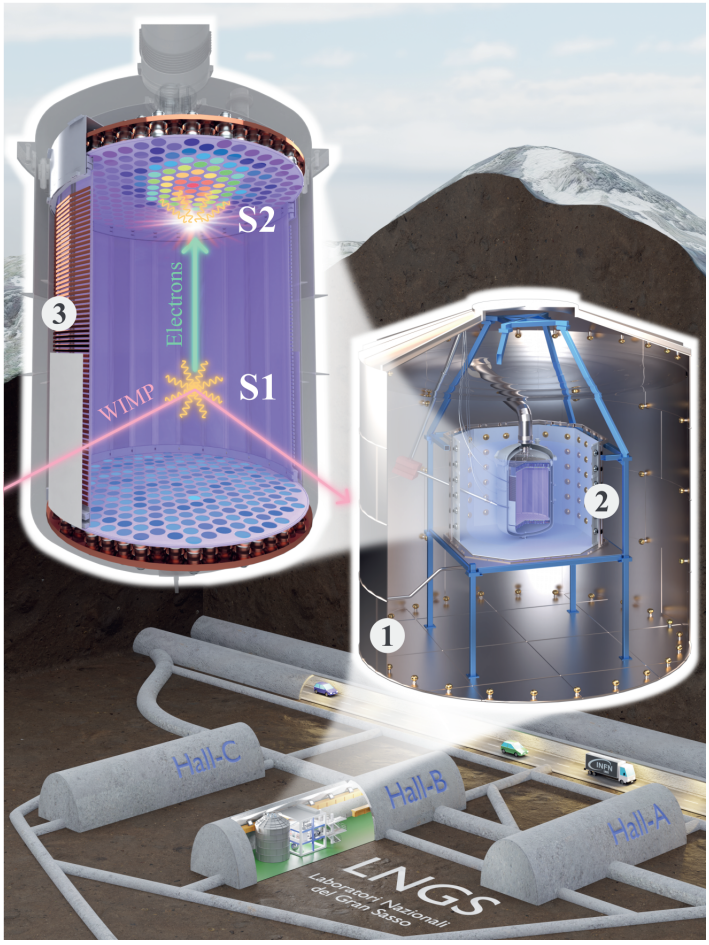


Figura 1 *Il rivelatore XENONnT ai Laboratori Nazionali del Gran Sasso (LNGS), Italia. Il serbatoio di xenon liquido (3) è circondato da sensori di luce sopra e sotto. L'interazione di una particella produce il lampo di luce che chiamiamo S1, e un secondo lampo, più grande, S2, una volta che gli elettroni liberati risalgono fino allo strato di gas. Il rivelatore si trova all'interno di un serbatoio d'acqua (2) per la schermatura, nella Sala B del laboratorio sotterraneo (1). Figura tratta da [2].*

blocca la maggior parte prima che arrivino. Un'altra parte è dovuta ai neutrini provenienti dal Sole. Ciascuno di essi interagisce molto raramente, ma il Sole ne emette così tanti che alcuni colpiscono comunque lo xenon. Quando succede, il segnale può somigliare molto a quello che ci aspettiamo dalla materia oscura.

La sorgente principale di questo background, però, si trova nell'ambiente stesso del rivelatore. I materiali ordinari, come la roccia e i componenti

del rivelatore, sono leggermente radioattivi, e nemmeno lo xenon liquido è perfettamente puro: vi sono disciolte tracce di radon, un gas radioattivo. Gli atomi radioattivi sono instabili e decadono uno alla volta, emettendo particelle alfa, elettroni o raggi gamma, e occasionalmente neutroni. Il serbatoio d'acqua attorno al rivelatore arresta la maggior parte dei neutroni, ma il resto di questa radiazione raggiunge costantemente lo xenon. Tenerla sotto controllo ha richiesto un'estrema cura nella scelta e nella pulizia di ogni materiale del rivelatore, rendendo XENONnT uno dei luoghi più puliti e silenziosi al mondo.

Quando una particella interagisce all'interno dello xenon, accadono due cose: gli elettroni vengono liberati dagli atomi di xenon e si produce un lampo di luce. I sensori attorno al serbatoio rilevano subito questo lampo: è il primo segnale, che chiamiamo S1. Gli elettroni liberati risalgono quindi attraverso il liquido sotto l'effetto di un campo elettrico. Pochi microsecondi dopo raggiungono il gas sopra il liquido e producono un secondo lampo, più grande, che chiamiamo S2. Insieme, S1 e S2 ci indicano dove è avvenuta l'interazione e quanta energia è stata depositata. Questo processo, e i dettagli del rivelatore che lo rendono possibile, dai fotomoltiplicatori all'elettronica che ricostruisce S1 e S2, sono l'argomento del Capitolo 2. Interpretarlo con precisione richiede anche una simulazione affidabile del rivelatore. Il Capitolo 3 descrive `fuse`, il framework di simulazione che riproduce l'intero processo, dall'interazione di una particella fino a una forma d'onda digitalizzata, producendo dati simulati quasi indistinguibili da quelli reali.

XENONnT è la tappa più recente di una ricerca di materia oscura che la collaborazione XENON conduce da quasi 25 anni con una serie di rivelatori (XENON10, XENON100 e XENON1T), ciascuno più grande e più pulito del precedente. La versione attuale contiene 8,6 tonnellate di xenon ed è in funzione dal 2020, senza aver ancora trovato materia oscura.

In attesa di scoprire la materia oscura, lo stesso rivelatore serve anche ad altro. Il suo background molto basso e la capacità di misurare S1 e S2 con alta precisione lo rendono prezioso ben oltre la ricerca per cui è stato costruito. In questa tesi usiamo i dati del rivelatore per tre scopi: modellare come lo xenon liquido converte l'energia in luce e carica, misurare lo spettro energetico di un decadimento nucleare con una precisione sufficiente a distinguere tra diverse teorie della struttura nucleare, e cercare la materia oscura attraverso un'interazione alternativa, l'assorbimento anziché lo scattering.

Tutte e tre le analisi si basano sullo stesso trucco per combattere il background. In ciascun caso, il processo che cerchiamo produce non uno, ma due segnali: un primo segnale seguito, dopo un tempo breve e prevedibile, da un secondo. Il background ordinario non si comporta quasi

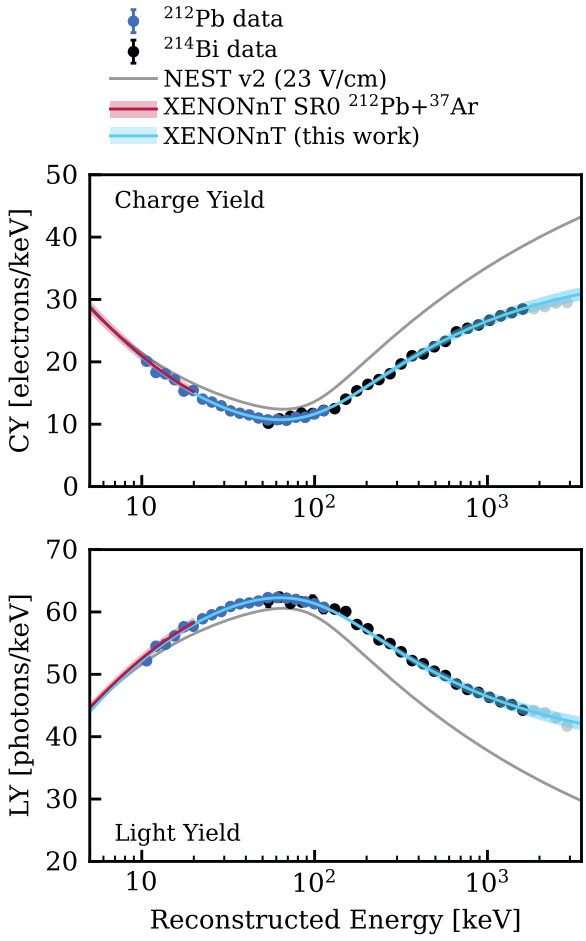


Figura 2 *Ripartizione di una data quantità di energia depositata tra luce (LY, in basso) e carica (CY, in alto) nello xenon liquido, ricavata dai decadimenti di ^{214}Bi identificati come descritto nel testo. Questa misura corregge un disaccordo significativo nei modelli precedenti ed è utilizzata in tutta la tesi.*

mai così, quindi richiedere entrambi i segnali, nell'ordine giusto e nella giusta finestra temporale, elimina quasi tutto il background in un colpo solo. Questa tecnica è nota come coincidenza ritardata (*delayed coincidence*).

Una delle catene di decadimento del radon è particolarmente utile. Un atomo di ^{214}Bi decade emettendo un elettrone, trasformandosi in ^{214}Po, che a sua volta decade una frazione di millisecondo più tardi emettendo una particella alfa. Questa è la firma che usiamo: ogni volta che osser-

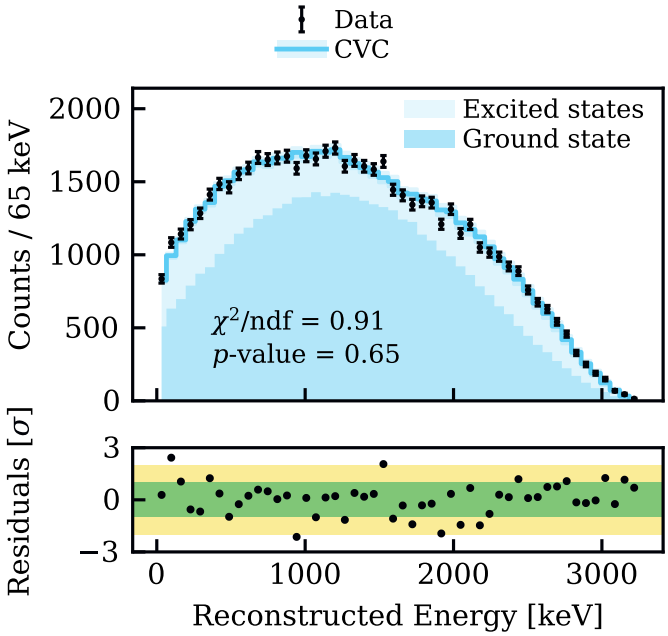


Figura 3 *Spettro beta di ^{214}Bi misurato, confrontato con una simulazione* `fuse` *a catena completa basata sull'argomento generale di simmetria (CVC) descritto nel testo, il modello che meglio riproduce i dati. Nel pannello inferiore sono mostrati i residui. Due calcoli di struttura nucleare più dettagliati, non mostrati, si discostano maggiormente dai dati.*

viamo questa coppia, sappiamo di aver individuato un decadimento di ^{214}Bi, ben distinto da tutto ciò che accade nel rivelatore. Il Capitolo 4 usa questo campione selezionato per determinare come una data quantità di energia depositata si divide tra il segnale di luce e quello di carica nello xenon liquido, una relazione che i modelli precedenti sbagliavano in modo significativo a queste energie (Figura 2). Determinare correttamente questa relazione è importante per ogni altra misura descritta in questa tesi, poiché costituisce il metro di riferimento rispetto a cui tutto il resto viene misurato.

Con questo metro ben calibrato, il Capitolo 5 va un passo oltre e si chiede che aspetto abbia esattamente il decadimento beta: non solo quanta energia rilasci in media, ma l'intero spettro di energie che l'elettrone può portare via. Quello spettro non è una curva semplice. La sua forma esatta dipende dalla struttura interna del nucleo atomico in un modo difficile da calcolare, e diversi modelli nucleari sono in netto disaccordo su come dovrebbe apparire. Il Capitolo 1 illustra la fisica nucleare alla base di tutto ciò: perché la forma è difficile da prevedere, e cosa una sua misura

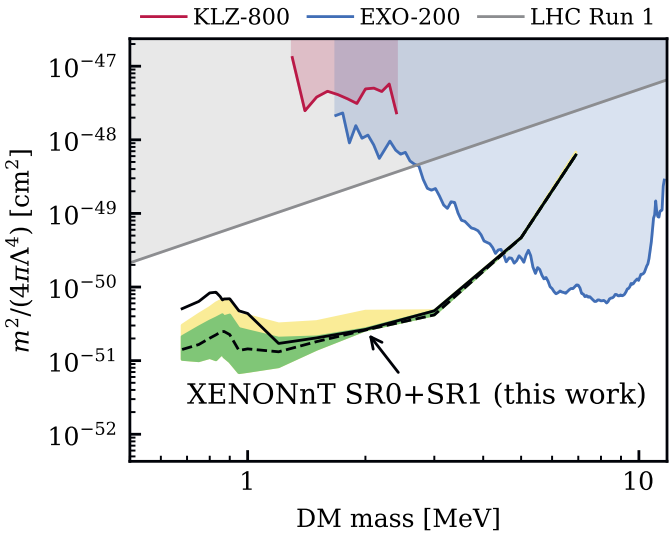


Figura 4 *Limiti superiori con un livello di confidenza del 90% sull'intensità del processo di assorbimento di materia oscura qui cercato, in funzione della massa della materia oscura, da questo lavoro e da altri esperimenti sensibili allo stesso processo. In parte dell'intervallo di massa, questa misura migliora fino a quattro ordini di grandezza il miglior limite precedente.*

precisa può insegnarci sul nucleo. Usando lo stesso campione selezionato di decadimenti puliti di ^{214}Bi, abbiamo misurato quello spettro con una precisione sufficiente a distinguere tra i modelli in competizione (Figura 3).

Il Capitolo 6 torna alla materia oscura, ma da un'angolazione diversa. La maggior parte degli studi cerca una particella di materia oscura che urta un nucleo di xenon e lo smuove leggermente, come una palla da biliardo che ne urta un'altra. Per la materia oscura molto leggera, questa spinta è troppo debole per essere osservata. Un'alternativa è l'assorbimento: la particella di materia oscura non viene deviata ma distrutta, cedendo al nucleo tutta la sua massa in un solo istante. Nello xenon, questo può lasciare un nucleo eccitato che si diseccita poco dopo emettendo un ulteriore lampo di luce, ritardato rispetto al segnale originale. Richiedere quel secondo lampo ritardato elimina quasi tutto il background ordinario che altrimenti sommergerebbe la ricerca. È lo stesso trucco che ha reso possibile la misura del decadimento beta, applicato ora a una domanda di fisica diversa.

Non è stato trovato alcun segnale di materia oscura. In una parte dell'intervallo di massa, però, i limiti posti da questa ricerca sull'intensità dell'interazione sono circa diecimila volte più stringenti di quelli noti in precedenza (Figura 4).

La tesi si chiude con un'appendice sul software: l'infrastruttura che elabora, archivia e distribuisce i dati di XENONnT, costruita e mantenuta in funzione per diversi anni da molte persone all'interno della collaborazione. L'appendice fornisce una breve descrizione del software e dell'infrastruttura principali, seguita da alcune osservazioni personali sugli aspetti pratici del suo utilizzo e mantenimento.

Il barile, a quanto pare, aveva di meglio da fare che starsene lì ad aspettare.

ACKNOWLEDGEMENTS

It feels like yesterday when I was writing the acknowledgements for my master thesis. It starts with "this master project was a blast": so tacky! And now, four years later, I am writing the acknowledgements for my PhD thesis: I have to say it was, again, quite a blast! Somehow it all worked out, the paper I wanted to publish is out (only a few years – years! – later), I survived the responsibility of the DAQ and (A)I wrote this long thesis... and all of this with too little suffering – according to many.

First of all, thank you to all the members of the committee for taking the time to read my thesis. A special thanks to Mara for being my C3 member, it has always been a nice moment to discuss my work and my plans with you. And thanks to Luca for coming from far away for a last opportunity to party together. Jelle, thank you for the inspiring conversations about DAQ and statistics, and Tina, thank you for all the knowledge you brought to the group.

It's time to thank my supervisors: Auke-Pieter and Patrick. Thank you for being such an important point of reference for me. Thank you for always giving me so much trust and freedom, you really taught me how to become an independent researcher and I feel like I learned from the two of you how to think and approach problems. Also, well, "thank you", for never giving me a clear answer to the question "what is the right way to do this", I remember times in which I was begging you for some advice and all I would get was a "what would Carlo want? what do you think is the right way to do this?" – I haaated it but I think this is a very important lesson for me, and I will try to keep it in mind for the rest of my career.

Auke-Pieter, thanking you is a bit difficult. Everybody knows how annoying we are together, always arguing about the most stupid things, asking the same exact questions, making fun of each other, or talking about what crazy thing our AI managed to do. We spent so much time together that at times it was actually scary to see how similar our minds were. But I think we really did learn a lot from each other, and I am very happy to have had you as my supervisor. It's strange how many fun moments come to mind, considering that you are my supervisor, but yeah all the collaboration meetings and conferences have been so much fun with you. And it's so funny you even know all my friends and family in Italy, from the drinks we had in Milan to the incredibly loud dinner we had in Genova, or the lunch at my parents' house – and thank you again for inviting us to Texel for an amazing weekend! I am sure we will keep in touch, just waiting for

the right idea to kickstart our company and make a lot of money with very little work.

Patrick, thank you for always being so supportive. Talking to you is always a pleasure. Thank you for pushing me sometimes, for forcing me to give a talk at conferences even when I didn't feel like it, and for always taking the time to listen to me and to give me advice. Your care for the group and for the people in it is really admirable, and thank you for all the nice BBQs you organised at your house. I loved listening to all your infinite and crazy life stories and adventures.

A special thanks goes of course to Maxime. I've been so lucky to have you around for the entire duration of my PhD. You helped me so much. You were always my point of reference: having you as leader of the high-energy team while I was busy with my analysis, and then as analysis coordinator while I was the team leader of the simulations team, was incredibly helpful. I wish you all the best with your career in academia, you really deserve it!

To the DAQ team: Joran and Peter (I learned so much from you two), Maricke, Andrea, Stefano, Robin, Andrew, Jaron, Alessandro, Gnanesh and Felix. I think we did a great job together, it's been difficult, it's been stressful, it's a lot of responsibility, but we always managed to keep the DAQ running and to keep the collaboration happy. We all put so much effort into this and I think we can be proud of it.

To all my friends from the XENON collaboration, Micha, Henning, Diego, Christian, Daniel, Cecilia, Riccardo, Lanqing, Giovanni, Florian, Luisa, Maxi, Ananthu, Lorenzo, Sebastian and many many others! Thanks for all the nice moments we had together during the collaboration meetings and workshops, it's been a pleasure to work with you!

A special thanks to all the maaany many bachelor and master students who orbited around our group. You were often the ones who gave me the most motivation to believe in this whole science and academia thing. Seeing your excitement and curiosity was always a great reminder of why I wanted to do this in the first place. And thanks for keeping me young, thanks for all the GenZ stuff you taught me – I'll be forever grateful, and thanks for the fun conversations at lunch. I could never fit all your names on this page, but I hope you know how much I appreciated having you around – when you are all gone, what we do feels so empty!

Thanks to Lu and Filippino, my favourite designers, who helped me with the cover of the thesis!

Now the people I shared my days with, at work and long after.

Pranati, let me start with you. For the last three years we've been sitting next to each other every day, talking talking talking. A lot of gossip, but also a lot of work. I know how difficult it can be to work with me, and I am sorry you had to go through that, and sorry if I pushed too much

sometimes, but I feel like we really did learn a lot from each other and that we reached some sort of equilibrium that made our collaboration very productive and our time together very enjoyable and fun. I loved those moments of you talking shit about me with Cecile, and I am sooo happy I could come visit you in India and meet your family – those days were just so nice and special! Thanks for understanding me and for helping me be a little more conscious of all the stupid privileges I have.

Saad.. what can one even say about Saad? I'm so glad I really got to know you in all these different ways. Saad, thank you so much for all the help, for all the physics you explained to me – in a way I have never seen in anyone else. You manage to explain the very basics to people without making anyone feel stupid or judged, and this is really a gift you have! Keep going and keep teaching and helping people the way you do, with all your funny experiments and project and shows and jokes, and your amazing, calming and deep voice.

Saad and Pranati, thanks, really: our office was just amazing, so genuine and so fun. I will miss it a lot, and I will miss you both a lot. I hope we can keep in touch and see each other again soon!

Maricke and Kelly, I miss being in the office with you two. Thanks, Maricke, for the infinite help with the DAQ, and I'm sorry you had to deal with me being a little bit too stressed sometimes. I think we did great together! Kelly, thanks for being so genuine and funny, and for bringing me the best pens from all over the world! Anna, it was really fun to have you around at collaboration meetings, please make a podcast with Saad one day – it's too much fun to listen to you two together!

Marjolein, you are one of those very special people I met during my PhD. I am so happy we got to know each other better during the school in Sicily. I am so proud of you, you grew so much in the last few years and you've been so brave! We're so different but something about our personalities makes us understand each other. Good luck with your new life and job in Heidelberg, I'll miss babysitting Hidde in Amsterdam!

Alejandro, you should have been the first person I thanked! To you I can really say: the final chapter of my thesis would not have been possible without your help. I could not find the motivation to do the work, and when you arrived so eager to learn and wanting to do your best, you managed to get me interested in the work again. Thanks for all your amazing work, for all the endless discussions we had – mainly you teaching me something new every time, and for all the fun we had together. I hope you will have a great time in your new life in the US, and I hope we can keep in touch!

Ali, I'm so glad we both started our PhDs at the same time after our master's. You're my first ever friend in Amsterdam, and you've made me discover so much of the life I have here: new people, music, parties, recipes

and just different ways of living. I think we just know each other very well by now. It feels really nice to have someone who understands what I mean without me having to explain too much, and who is there when things get difficult. Also thanks for joining my bet of finishing our theses on time, and congrats on making it on the last day!

Evri, Evri... I feel like I've known you forever, but it's only been two years since we really started hanging out. I know you tried so hard to get some time for yourself, to recharge your social batteries, to try to avoid seeing me literally every day... but I'm very happy you didn't manage that well! You're the person I think of when I see an interesting event happening, because I know you'll embrace it and enjoy it the same way I do. Thank you also for keeping alive the anarchist corner of my personality ahah. Thank you for all the beers at Fonteyn, for never saying no to a spontaneous plan, for skipping the line with me at the ice cream place in Mallorca, for visiting me in Genova and for all the nice everyday moments we had together. I love your comments that nobody can understand – too sharp for the world.

Andrea – Andrea the Italian – we met at the BND school four years ago, with sooo many people, and a few days later we already had our first date. Such a cute date, the two of us playing ping pong on the street. Since that day you've been a point of reference for me. You are so smart and so kind, always thinking about others, always managing to make everyone feel comfortable, and so respectful even in the most difficult conversations – like the ones about politics and morals with Evri. You've been my daily friend, always down for a ping pong during work, for a beer after work, or for a sketchy scuba diving course in a swimming pool. The way we handled all the problems of the beloved Espace (RIP) during our Christmas road trip was just unbelievable: everything is going to shit but we're enjoying our schnitzel at the restaurant, amazing! Thanks for being there always, for listening to me in the difficult moments, and for never being able to keep a secret from me.

Lluc, I miss you so much in Amsterdam. Thank you for bringing so much energy and fun, thank you for introducing me to INJI (really), for all the tequila shots and the dart games – is it is iiiit? Really you brought so much joy to my life in the middle of a particularly difficult Dutch winter, such a pleasure to have met you! And, of course.. thank you so much for two amaaaazing trips to Formentera and Mallorca, and for all the fun we had there. I'll come visit again soon in Berlin and hope to keep crossing paths in the future! You're my favourite guiri!

Andrea – the Spanish – the first picture of you I have on my phone is you drawing something ugly on my face with a pink pen, and I think this is a very good representation of our friendship. Thanks for always bringing the good mood, and especially: si necesitas el reggaetón, dale;

sigue bailando mami, no pare! Vlad, my favourite Nikhef intruder, thanks for always keeping the good conversations up! It's always nice to be around you, feeling accepted and welcomed in any situation.

Stefano, thank you for all the nice dinners, the dances at the gay street, the iiinfinite goodbyes. I'm happy to be the friend that you and Marta walk to, in pyjamas, crossing the entire city, without your phones and in your neighbours' jackets.. to come get your keys. You're always welcome! Aand just because you happened to be here, how to say no to a morning burraco session?!

To all the rest of the friends and colleagues I met during my PhD at Nikhef, Riccardo il mio amico più tamarro di sempre – he goes around in VanMoof and Tesla but we love him anyway, Luca, Petja, Clara, Thijs, Alex, Bryan, Brian, Zef, Osama, Marion, Vikas, Jose, Pravita, Oliver, Maria, Marta, Mariia, Giacomo, Roberto, Ilaria, Jordan, Leo, Charis, Fabri and so many others: thanks for all the nice moments we had together and for making my PhD experience so memorable – if these years have been so special, it is thanks to all of you!

Thanks also to all of Ceci's family and friends for welcoming me into their lives and for making me feel at home in Amsterdam. A day in Groesbeek with the family, the siblings all together, a birthday where there was a present for me too or a party with the Chubs. This feeling is just something so special and difficult to get anywhere else. I feel so lucky that I got to experience this in my life in the Netherlands.

Finally, I put you all here at the end because I think you are so important that you wouldn't even need to be mentioned in the acknowledgements, but I want to thank Cecile, my family and my friends.

Ceci, you're my best friend. You already know how much I love you and how grateful I am for having you next to me. Thank you for supporting me through these four years of my PhD and for sharing all the good and the less good moments with me. You are so cute, and you even manage to bring out the cute part of me. Thank you for teaching me to see and appreciate the softness and cuteness in people. I think I see the world a little differently because of you. To Papà e Mamma, grazie per tutte le volte che siete venuti a trovarmi, e per accogliermi con così tanto entusiasmo e amore ogni singola volta che torno a casa. Anna e Dede, siete i fratelli migliori del mondo, grazie per capirmi sempre anche quando sono il fratello grande e rompiscatole – con voi mi sento sempre al sicuro. Grazie anche a tutto il resto della mia famiglia, zii e cugini, grazie alla Nonna che ci ha insegnato tutto, e grazie a tutti i miei amici di Genova, grazie a quelli che mi sono venuti a trovare così tante volte e a quelli che saranno qui anche durante la difesa del mio dottorato: siete famiglia.

BIBLIOGRAPHY

[1] E. Aprile et al. (XENON Collaboration), "Spectral measurement of the ^{214}Bi β decay to the ^{214}Po ground state with the XENONnT Experiment", Phys. Rev. C **113**, 044303 (2026).

[2] E. Aprile et al. (XENON Collaboration), "Radon Removal in XENONnT down to the Solar Neutrino Level", Phys. Rev. X **15**, 031079 (2025).

[3] H. Becquerel, "Sur les radiations émises par phosphorescence", Comptes Rendus de l'Académie des Sciences **122**, 420–421 (1896).

[4] E. Rutherford, "Uranium radiation and the electrical conduction produced by it", Phil. Mag., 5th ser. **47**, 109–163 (1899).

[5] S. Abe et al. (KamLAND-Zen Collaboration), "Search for Majorana neutrinos with the complete KamLAND-Zen dataset", Phys. Rev. Lett. **135**, 262501 (2025).

[6] J. Engel and J. Menéndez, "Status and Future of Nuclear Matrix Elements for Neutrinoless Double-Beta Decay: A Review", Rept. Prog. Phys. **80**, 046301 (2017).

[7] S. J. Haselschwardt et al., "Observation of Low-Lying Isomeric States in ^{136}Cs: A New Avenue for Dark Matter and Solar Neutrino Detection in Xenon Detectors", Phys. Rev. Lett. **131**, 052502 (2023).

[8] H. Behrens and W. Bühring, *Electron radial wave functions and nuclear beta-decay*, International Series of Monographs on Physics (Clarendon Press, Oxford, UK, 1982).

[9] L. Hayen et al., "High precision analytical description of the allowed β spectrum shape", Rev. Mod. Phys. **90**, 015008 (2018).

[10] E. Fermi, "Versuch einer Theorie der β-Strahlen. I", Z. Phys. **88**, 161–177 (1934).

[11] G. Gamow and E. Teller, "Selection Rules for the β-Disintegration", Phys. Rev. **49**, 895–899 (1936).

[12] C. S. Wu et al., "Experimental Test of Parity Conservation in Beta Decay", Phys. Rev. **105**, 1413–1415 (1957).

[13] E. C. G. Sudarshan and R. E. Marshak, "The nature of the four-fermion interaction", in Proceedings of the Padua-Venice Conference on Mesons and Newly Discovered Particles, Reprinted in P. K. Kabir (ed.), *The Development of Weak Interaction Theory*, Gordon and Breach, New York, 1963 (1957).

[14] E. C. G. Sudarshan and R. E. Marshak, "Chirality invariance and the universal Fermi interaction", Phys. Rev. **109**, 1860–1862 (1958).

[15] R. P. Feynman and M. Gell-Mann, "Theory of Fermi interaction", Phys. Rev. **109**, edited by L. M. Brown, 193–198 (1958).

[16] W. Pauli, "Dear radioactive ladies and gentlemen", Physics Today **31**, English translation of the open letter to the Physical Society of Tübingen, 4 December 1930, 27 (1978).

[17] M. E. Rose, *Relativistic Electron Theory* (John Wiley & Sons, New York, 1961).

[18] F. N. D. Kurie, J. R. Richardson and H. C. Paxton, "The Radiations Emitted from Artificially Produced Radioactive Substances. I. The Upper Limits and Shapes of the β-Ray Spectra from Several Elements", Phys. Rev. **49**, 368–381 (1936).

[19] K. S. Krane, *Introductory nuclear physics* (Wiley John & Sons, New York, 1987).

[20] K. Grotz and H. V. Klapdor, *The weak interaction in nuclear, particle and astrophysics*, Translated by S. S. Wilson (IOP Publishing, Bristol, 1990).

[21] J. Kostensalo and J. Suhonen, "g_A-driven shapes of electron spectra of forbidden β decays in the nuclear shell model", Phys. Rev. C **96**, 024317 (2017).

[22] M. Haaranen, J. Kotila and J. Suhonen, "Spectrum-shape method and the next-to-leading-order terms of the β-decay shape factor", Phys. Rev. C **95**, 024327 (2017).

[23] H. Ejiri, J. Suhonen and K. Zuber, "Neutrino–nuclear responses for astro-neutrinos, single beta decays and double beta decays", Phys. Rep. **797**, 1–102 (2019).

[24] M. Ramalho and J. Suhonen, "Computed total β-electron spectra for decays of Pb and Bi in the ^{220,222}Rn radioactive chains", Phys. Rev. C **109**, 014326 (2024).

[25] J. C. Hardy and I. S. Towner, "Superallowed $0^+ \to 0^+$ nuclear β decays: 2020 critical survey, with implications for V_{ud} and CKM unitarity", Phys. Rev. C **102**, 045501 (2020).

[26] B. R. Holstein, "Recoil effects in allowed beta decay: The elementary particle approach", Rev. Mod. Phys. **46**, 789–814 (1974).

[27] J. Suhonen, *From nucleons to nucleus: concepts of microscopic nuclear theory*, Theoretical and Mathematical Physics (Springer, Berlin, Heidelberg, 2007), 10.1007/978-3-540-48861-3.

[28] E. K. Warburton and B. A. Brown, "Appraisal of the Kuo-Herling shell-model interaction and application to A=210-212 nuclei", Phys. Rev. C **43**, 602–617 (1991).

[29] N. Shimizu et al., "Thick-Restart Block Lanczos Method for Large-Scale Shell-Model Calculations", Comput. Phys. Commun. **244**, 372–384 (2019).

[30] B. A. Brown and W. D. M. Rae, "The Shell-Model Code NuShellX@MSU", Nucl. Data Sheets **120**, 115–118 (2014).

[31] Q. Zhi et al., "Shell-model half-lives including first-forbidden contributions for *r*-process waiting-point nuclei", Phys. Rev. C **87**, 025803 (2013).

[32] S. J. Haselschwardt et al., "Solar Neutrino Detection in Liquid Xenon Detectors via Charged-Current Scattering to Excited States", Phys. Rev. D **102**, 072009 (2020).

[33] J. Bardeen, L. N. Cooper and J. R. Schrieffer, "Theory of superconductivity", Phys. Rev. **108**, 1175–1204 (1957).

[34] F. F. Deppisch and J. Suhonen, "Statistical analysis of β decays and the effective value of g_A in the proton-neutron quasiparticle random-phase approximation framework", Phys. Rev. C **94**, 055501 (2016).

[35] J. Toivanen and J. Suhonen, "Microscopic quasiparticle-phonon description of odd-mass isotopes and their β decay", Phys. Rev. C **57**, 1237–1245 (1998).

[36] S. J. Haselschwardt et al., "Improved calculations of β decay backgrounds to new physics in liquid xenon detectors", Phys. Rev. C **102**, 065501 (2020).

[37] S. Navas et al. (Particle Data Group), "Review of particle physics", Phys. Rev. D **110**, 030001 (2024).

[38] J. T. Suhonen, "Value of the axial-vector coupling strength in β and $\beta\beta$ decays: A review", Front. Phys. **5**, 55 (2017).

[39] P. Gysbers et al., "Discrepancy between experimental and theoretical β-decay rates resolved from first principles", Nat. Phys. **15**, 428–431 (2019).

[40] S. Zhu and E. A. McCutchan, "Nuclear Data Sheets for A=214", Nucl. Data Sheets **175**, 1–149 (2021).

[41] J. Kostensalo and J. Suhonen, "Mesonic enhancement of the weak axial charge and its effect on the half-lives and spectral shapes of first-forbidden $J^+ \leftrightarrow J^-$ decays", Phys. Lett. B **781**, 480–484 (2018).

[42] X. Mougeot, "BetaShape: A new code for improved analytical calculations of beta spectra", EPJ Web Conf. **146**, 12015 (2016).

[43] E. Aprile et al. (XENON Collaboration), "The XENONnT dark matter experiment", Eur. Phys. J. C **84**, 784 (2024).

[44] G. Bellini et al. (Borexino Collaboration), "Cosmic-muon flux and annual modulation in Borexino at 3800 m water-equivalent depth", JCAP **05**, 015 (2012).

[45] E. Aprile et al. (XENON Collaboration), "WIMP dark matter search using a 3.1 tonne-year exposure of the XENONnT experiment", Phys. Rev. Lett. **135**, 221003 (2025).

[46] J. Aalbers et al. (XLZD Collaboration), "The XLZD Design Book: towards the next-generation liquid xenon observatory for dark matter and neutrino physics", Eur. Phys. J. C **85**, 1192 (2025).

[47] E. Aprile and T. Doke, "Liquid Xenon Detectors for Particle Physics and Astrophysics", Rev. Mod. Phys. **82**, 2053–2097 (2010).

[48] E. Aprile et al., *Noble gas detectors* (John Wiley & Sons, Ltd, 2006), 10.1002/9783527610020.

[49] E. Aprile et al. (XENON Collaboration), "Design and Performance of the XENON10 Dark Matter Experiment", Astropart. Phys. **34**, 679–698 (2011).

[50] E. Aprile et al. (XENON100 Collaboration), "Dark Matter Results from 225 Live Days of XENON100 Data", Phys. Rev. Lett. **109**, 181301 (2012).

[51] E. Aprile et al. (XENON Collaboration), "First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment", Phys. Rev. Lett. **131**, 041003 (2023).

[52] S. A. Kharusi et al. (EXO-200 Collaboration), "Search for two-neutrino double-beta decay of ^{136}Xe to the excited state of ^{136}Ba with the complete EXO-200 dataset", Chin. Phys. C **47**, 103001 (2023).

[53] E. Aprile et al. (XENON Collaboration), "Double-weak decays of ^{124}Xe and ^{136}Xe in the XENON1T and XENONnT experiments", Phys. Rev. C **106**, 024328 (2022).

[54] A. Hitachi et al., "Effect of ionization density on the time dependence of luminescence from liquid argon and xenon", Phys. Rev. B **27**, 5279–5285 (1983).

[55] P. Agnes et al. (DarkSide Collaboration), "Simulation of argon response and light detection in the DarkSide-50 dual phase TPC", JINST **12**, P10015 (2017).

[56] D. S. Akerib et al. (LZ Collaboration), "Pushing the Limits of Pulse Shape Discrimination in a Large Liquid Xenon Detector", (2026), arXiv:2603.26877 [physics.ins-det].

[57] J. Thomas and D. A. Imel, "Recombination of electron-ion pairs in liquid argon and liquid xenon", Phys. Rev. A **36**, 614–616 (1987).

[58] E. Aprile et al. (XENON Collaboration), "The neutron veto of the XENONnT experiment: results with demineralized water", Eur. Phys. J. C **85**, 695 (2025).

[59] E. Aprile et al. (XENON Collaboration), "Conceptual design and simulation of a water Cherenkov muon veto for the XENON1T experiment", JINST **9**, P11006 (2014).

[60] V. C. Antochi et al., "Improved quality tests of R11410-21 photomultiplier tubes for the XENONnT experiment", JINST **16**, P08033 (2021).

[61] E. Aprile et al. (XENON Collaboration), "Design and performance of the field cage for the XENONnT experiment", Eur. Phys. J. C **84**, 138 (2024).

[62] G. Plante et al., "Liquid-phase purification for multi-tonne xenon detectors", Eur. Phys. J. C **82**, 860 (2022).

[63] E. Aprile et al. (XENON Collaboration), "Application and modeling of an online distillation method to reduce krypton and argon in XENON1T", PTEP **2022**, 053H01 (2022).

[64] E. Aprile et al. (XENON Collaboration), "Material radiopurity control in the XENONnT experiment", Eur. Phys. J. C **82**, 599 (2022).

[65] E. Aprile et al. (XENON Collaboration), "The triggerless data acquisition system of the XENONnT experiment", JINST **18**, P07054 (2023).

[66] E. Aprile et al. (XENON Collaboration), "XENONnT analysis: signal reconstruction, calibration, and event selection", Phys. Rev. D **111**, 062006 (2025).

[67] C. E. Dahl, "The physics of background discrimination in liquid xenon, and first results from Xenon10 in the hunt for WIMP dark matter", PhD thesis (Princeton U., 2009).

[68] XENON Collaboration, H. Schulze Eißing et al., *XENONnT/fuse: 1.4.2*, version 1.4.2, 2025, 10.5281/zenodo.14740249.

[69] S. Agostinelli et al. (GEANT4 Collaboration), "GEANT4 - A Simulation Toolkit", Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250–303 (2003).

[70] J. Allison et al., "Geant4 developments and applications", IEEE Trans. Nucl. Sci. **53**, 270–278 (2006).

[71] J. Allison et al. (GEANT4 Collaboration), "Recent Developments in Geant4", Nucl. Instrum. Meth. A **835**, 186–225 (2016).

[72] P. Gaemers et al., *Xenonnt/wfsim: v1.0.2*, version v1.0.2, 2022, 10.5281/zenodo.7216324.

[73] D. Wenz et al., *XENONnT/epix: v0.3.4*, version v0.3.4, 2023, 10.5281/zenodo.7516942.

[74] XENON Collaboration, J. R. Angevaare et al., *XENONnT/straxen: v3.1.4*, version v3.1.4, 2025, 10.5281/zenodo.14996117.

[75] M. Szydagis et al., "A review of NEST models for liquid xenon and an exhaustive comparison with other approaches", Front. Det. Sci. Technol. **2**, 1480975 (2025).

[76] COMSOL AB, *COMSOL Multiphysics®*, https://www.comsol.com, Stockholm, Sweden, 2024.

[77] J. Mock et al., "Modeling Pulse Characteristics in Xenon with NEST", JINST **9**, T04002 (2014).

[78] J. B. Albert et al. (EXO-200 Collaboration), "Measurement of the Drift Velocity and Transverse Diffusion of Electrons in Liquid Xenon with the EXO-200 Detector", Phys. Rev. C **95**, 025502 (2017).

[79] C. H. Faham et al., "Measurements of wavelength-dependent double photoelectron emission from single photons in VUV-sensitive photomultiplier tubes", JINST **10**, P09010 (2015).

[80] R. Veenhof, *Simulation of gaseous detectors*, http://garfield.web.cern.ch/garfield/, 2018.

[81] J. Lindhard et al., "Integral equations governing radiation effects. (notes on atomic collisions, iii)", Kgl. Danske Videnskab., Selskab. Mat. Fys. Medd. **Vol: 33: No. 10** (1963), https://www.osti.gov/biblio/4701226.

[82] L. Baudis, P. Sanchez-Lucas and K. Thieme, "A measurement of the mean electronic excitation energy of liquid xenon", Eur. Phys. J. C **81**, 1060 (2021).

[83] T. Takahashi et al., "Average energy expended per ion pair in liquid xenon", Phys. Rev. A **12**, 1771–1775 (1975).

[84] T. Doke et al., "Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles", Jap. J. Appl. Phys. **41**, 1538–1545 (2002).

[85] A. Mozumder, "Free-ion yield in liquid argon at low-let", Chem. Phys. Lett. **238**, 143–148 (1995).

[86] M. J. Berger et al., *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3)*, Gaithersburg, MD, 2005, `http://physics.nist.gov/Star`.

[87] T. Doke et al., "Let Dependence of Scintillation Yields in Liquid Argon", Nucl. Instrum. Meth. A **269**, 291–296 (1988).

[88] J. B. Birks, "Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations", Proc. Phys. Soc. A **64**, 874–877 (1951).

[89] R. Acciarri et al. (ArgoNeuT Collaboration), "A Study of Electron Recombination Using Highly Ionizing Particles in the ArgoNeuT Liquid Argon TPC", JINST **8**, P08005 (2013).

[90] U. Fano, "Ionization Yield of Radiations. 2. The Fluctuations of the Number of Ions", Phys. Rev. **72**, 26–29 (1947).

[91] T. Doke et al., "Estimation of Fano factors in liquid argon, krypton, xenon and xenon-doped liquid argon", Nucl. Instrum. Meth. **134**, 353–357 (1976).

[92] J. Thomas, D. A. Imel and S. Biller, "Statistics of charge collection in liquid argon and liquid xenon", Phys. Rev. A **38**, 5793 (1988).

[93] D. S. Akerib et al. (LUX Collaboration), "Signal yields, energy resolution, and recombination fluctuations in liquid xenon", Phys. Rev. D **95**, 012008 (2017).

[94] K. Masuda et al., "Relation between scintillation and ionization produced by relativistic heavy ions in liquid argon", Phys. Rev. A **39**, 4732–4734 (1989).

[95] E. Conti et al. (EXO-200 Collaboration), "Correlated fluctuations between luminescence and ionization in liquid xenon", Phys. Rev. B **68**, 054201 (2003).

[96] A. Dobi, "Measurement of the Electron Recoil Band of the LUX Dark Matter Detector With a Tritium Calibration Source", PhD thesis (Maryland U., College Park, 2014), `10.13016/M24P5P`.

[97] M. Berger et al., *XCOM: photon cross section database (version 1.2)*, 1999, `https://physics.nist.gov/xcom`.

[98] E. A. McCutchan, "Nuclear Data Sheets for A = 83", Nucl. Data Sheets **125**, 201–394 (2015).

[99] Z. Elekes and J. Timar, "Nuclear Data Sheets for A = 128", Nucl. Data Sheets **129**, 191–436 (2015).

[100] Y. Khazov, I. Mitropolsky and A. Rodionov, "Nuclear Data Sheets for A = 131", Nucl. Data Sheets **107**, 2715–2930 (2006).

[101] K. Auranen and E. A. McCutchan, "Nuclear Data Sheets for A=212", Nucl. Data Sheets **168**, 117–267 (2020).

[102] H. An et al., "Dark matter implications from the XENONnT and LZ data", (2025), arXiv:2512.05850 [hep-ph].

[103] J. Aalbers et al. (LZ Collaboration), "Dark matter search results from 4.2 tonne-years of exposure of the LUX-ZEPLIN (LZ) experiment", Phys. Rev. Lett. **135**, 011802 (2025).

[104] Z. Bo et al. (PandaX Collaboration), "Dark matter search results from exposure of PandaX-4T", Phys. Rev. Lett. **134**, 011805 (2025).

[105] S.-C. Wu, "Nucl. Data Sheets for A = 214", Nucl. Data Sheets **110**, 681–748 (2009).

[106] M. Arnould, S. Goriely and K. Takahashi, "The r-process of stellar nucleosynthesis: Astrophysics and nuclear physics achievements and mysteries", Phys. Rep. **450**, 97–213 (2007).

[107] R. Arnold et al. (NEMO-3 Collaboration), "Results of the search for neutrinoless double-β decay in ^{100}Mo with the NEMO-3 experiment", Phys. Rev. D **92**, 072011 (2015).

[108] A. S. Barabash et al. (SuperNEMO Collaboration), "The BiPo-3 detector for the measurement of ultra low natural radioactivities of thin materials", JINST **12**, P06002 (2017).

[109] E. D. Commins, *Weak Interactions* (McGraw-Hill, New York, 2007).

[110] E. Caurier, F. Nowacki and A. Poves, "Shell Model description of the $\beta\beta$ decay of ^{136}Xe", Phys. Lett. B **711**, 62–64 (2012).

[111] F. Iachello and P. V. Isacker, *The Interacting Boson-Fermion Model* (Cambridge University Press, New York, 1991).

[112] J. C. Hardy and I. S. Towner, "New Limits on Fundamental Weak-Interaction Parameters from Superallowed β Decay", Phys. Rev. Lett. **94**, 092502 (2005).

[113] A. Kumar, P. C. Srivastava and J. Suhonen, "Second-forbidden nonunique β^- decays of ^{59,60}Fe: possible candidates for g_A sensitive electron spectral-shape measurements", Eur. Phys. J. A **57**, 225 (2021).

[114] S. Lindemann and H. Simgen, "Krypton assay in xenon at the ppq level using a gas chromatographic system and mass spectrometer", Eur. Phys. J. C **74**, 2746 (2014).

[115] E. Aprile et al. (XENON Collaboration), "Online^{222} Rn removal by cryogenic distillation in the XENON100 experiment", Eur. Phys. J. C **77**, 358 (2017).

[116] M. Murra et al., "Design, construction and commissioning of a high-flow radon removal system for XENONnT", Eur. Phys. J. C **82**, 1104 (2022).

[117] F. Joerg, G. Eurin and H. Simgen, "Production and characterization of a 222Rn-emanating stainless steel source", Appl. Radiat. Isot. **194**, 110666 (2023).

[118] E. Aprile et al. (XENON Collaboration), "Results from a Calibration of XENON100 Using a Source of Dissolved Radon-220", Phys. Rev. D **95**, 072008 (2017).

[119] F. Jörg et al., "Characterization of alpha and beta interactions in liquid xenon", Eur. Phys. J. C **82**, 361 (2022).

[120] E. Aprile et al. (XENON Collaboration), "Offline tagging of radon-induced backgrounds in XENON1T and applicability to other liquid xenon time projection chambers", Phys. Rev. D **110**, 012011 (2024).

[121] A. Gando et al. (KamLAND-Zen Collaboration), "Search for Majorana neutrinos near the inverted mass hierarchy region with KamLAND-Zen", Phys. Rev. Lett. **117**, 082503 (2016).

[122] M. Szydagis et al., "NEST: A Comprehensive Model for Scintillation Yield in Liquid Xenon", JINST **6**, P10002 (2011).

[123] E. Aprile et al., "Simultaneous measurement of ionization and scintillation from nuclear recoils in liquid xenon as target for a dark matter experiment", Phys. Rev. Lett. **97**, 081302 (2006).

[124] G. Bertone and T. M. P. Tait, "A new era in the search for dark matter", Nature **562**, 51–56 (2018).

[125] M. Cirelli, A. Strumia and J. Zupan, "Dark Matter", (2024), arXiv:2406.01705 [hep-ph].

[126] E. Aprile et al. (XENON), "Light Dark Matter Search with 7.8 Tonne-Year of Ionization-Only Data in XENONnT", (2026), arXiv:2601.11296 [hep-ex].

[127] J. A. Dror, G. Elor and R. McGehee, "Directly Detecting Signals from Absorption of Fermionic Dark Matter", Phys. Rev. Lett. **124**, 181301 (2020).

[128] J. A. Dror et al., "Absorption of sub-MeV fermionic dark matter by electron targets", Phys. Rev. D **103**, [Erratum: Phys.Rev.D 105, 119903 (2022)], 035001 (2021).

[129] D. Zhang et al. (PandaX Collaboration), "Search for Light Fermionic Dark Matter Absorption on Electrons in PandaX-4T", Phys. Rev. Lett. **129**, 161804 (2022).

[130] J. A. Dror, G. Elor and R. McGehee, "Absorption of Fermionic Dark Matter by Nuclear Targets", JHEP **02**, 134 (2020).

[131] A. Guerrero Ramos, "Fermionic Dark Matter Charged-Current Absorption in XENONnT", MA thesis (University of Amsterdam, 2026).

[132] J. Kile and A. Soni, "Hidden MeV-Scale Dark Matter in Neutrino Detectors", Phys. Rev. D **80**, 115017 (2009).

[133] V. Khachatryan et al. (CMS), "Search for physics beyond the standard model in final states with a lepton and missing transverse energy in proton-proton collisions at $\sqrt{s} = 8$ TeV", Phys. Rev. D **91**, 092005 (2015).

[134] E. Richardson et al. (nEXO Collaboration), "Sensitivity of nEXO to ^{136}Xe Charged-Current Interactions: Background-free Searches for Solar Neutrinos and Fermionic Dark Matter", Phys. Rev. D **112**, 103010 (2025).

[135] B. M. Rebeiro et al., "$^{138}\mathrm{Ba}(d, \alpha)$ study of states in ^{136}Cs: Implications for new physics searches with xenon detectors", Phys. Rev. Lett. **131**, 052501 (2023).

[136] K. Tachibana, "Novel Search for Fermionic Dark Matter via Xe Charged-Current Reactions in KamLAND-Zen", PhD thesis (Tohoku University, 2026).

[137] Z. Ge et al., "β^- decay Q-value measurement of ^{136}Cs and its implications to neutrino studies", Phys. Rev. C **108**, 045502 (2023).

[138] H. Schulze-Eißing, "A Machine Learning Method for Delayed Coincidence Reconstruction in XENONnT: Solar Neutrino Capture on 136Xe and 85Kr Assay", PhD thesis (University of Münster, 2026).

[139] National Nuclear Data Center, *NuDat: Nuclear Structure and Decay Data*, Brookhaven National Laboratory, `https://www.nndc.bnl.gov/nudat3/`, 2024.

[140] O. Ponkratenko, V. Tretyak and Y. Zdesenko, "Event Generator DECAY4 for Simulating Double-Beta Processes and Decays of Radioactive Nuclei", Phys. At. Nucl. **63**, 1282–1287 (2000), arXiv:`nucl-ex/0104018`.

[141] IAEA Nuclear Data Section, *LiveChart of Nuclides*, `https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html`, 2024.

[142] S. Abe et al. (KamLAND), "Production of Radioactive Isotopes through Cosmic Muon Spallation in KamLAND", Phys. Rev. C **81**, 025807 (2010).

[143] T. Böhlen et al., "The FLUKA Code: Developments and Challenges for High Energy and Medical Applications", Nucl. Data Sheets **120**, 211–214 (2014).

[144] A. Ferrari et al., *FLUKA: A Multi-Particle Transport Code*, tech. rep. CERN-2005-010, INFN/TC-05/11, SLAC-R-773 (CERN, 2005).

[145] E. Vitagliano, I. Tamborra and G. Raffelt, "Grand Unified Neutrino Spectrum at Earth: Sources and Spectral Components", Rev. Mod. Phys. **92**, 045006 (2020).

[146] J. N. Bahcall and R. K. Ulrich, "Solar Models, Neutrino Experiments, and Helioseismology", Rev. Mod. Phys. **60**, 297–372 (1988).

[147] M. Agostini et al. (Borexino), "Comprehensive measurement of pp-chain solar neutrinos", Nature **562**, 505–510 (2018).

[148] S. Appel et al. (Borexino), "Improved measurement of solar neutrinos from the Carbon-Nitrogen-Oxygen cycle by Borexino and its implications for the Standard Solar Model", Phys. Rev. Lett. **129**, 252701 (2022).

[149] B. Aharmim et al. (SNO), "Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory", Phys. Rev. C **88**, 025501 (2013).

[150] D. Xu et al., *Xenonnt/alea: vo.2.4*, version v0.2.4, 2024, 10.5281/zenodo.10829030.

[151] J. Aalbers et al., *Jelleaalbers/blueice: v1.2.1*, version v1.2.1, 2024, 10.5281/zenodo.14390807.

[152] S. Al Kharusi et al. (EXO-200 Collaboration), "Search for MeV electron recoils from dark matter in EXO-200", Phys. Rev. D **107**, 012007 (2023).

[153] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C **71**, Erratum: Eur. Phys. J. C 73 (2013) 2501, 1554 (2011).

[154] J. Aalbers et al., *AxFoundation/strax: v2.2.2*, version v2.2.2, 2025, 10.5281/zenodo.18727696.

[155] J. Angevaare, *XENONnT Analysis software*, PyHEP 2021 Workshop presentation, 2021, 10.5281/zenodo.6798231.

[156] R. Kuprieiev et al., *DVC: Data Version Control – Git for Data & Models*, version 3.50.1, 2024, 10.5281/zenodo.11074054.

[157] Y. Mosbacher, L. Sanchez, D. Xu et al., *XENONnT/xedocs: v0.2.44*, version v0.2.44, 2026, 10.5281/zenodo.18347811.

[158] L. Sanchez, *Handling Detector Characterization Data (Metadata) in XENONnT*, CHEP 2023 poster, 2023, 10.5281/zenodo.7945375.

[159] D. S. Akerib et al. (LZ), "The LUX-ZEPLIN (LZ) Experiment", Nucl. Instrum. Meth. A **953**, 163047 (2020).

[160] C. E. Aalseth et al. (DarkSide-20k), "DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS", Eur. Phys. J. Plus **133**, 131 (2018).

[161] M. Barisits et al., "Rucio: Scientific data management", Comput. Softw. Big Sci. **3**, 11 (2019).

[162] E. Shockley et al., *XENONnT/utilix: v0.12.1*, version v0.12.1, 2026, 10.5281/zenodo.18840641.

[163] D. Xu et al., *XENONnT/outsource: v0.4.5*, version v0.4.5, 2026, 10.5281/zenodo.18841671.



Dark matter makes up most of the matter in the Universe and has only ever been seen through its gravity. XENONnT, a liquid xenon time projection chamber operating deep under the Gran Sasso in Italy, is among the most sensitive experiments searching for it, and its low background also makes it a capable instrument for nuclear physics.
This dissertation presents a new measurement of the light and charge yields of liquid xenon, a measurement of the ^{214}Bi beta spectrum precise enough to discriminate between competing nuclear structure models, and a search for fermionic dark matter absorption on ^{136}Xe that improves existing limits by up to four orders of magnitude.

The matter discussed in this thesis is made of many isotopes of different elements: thorium, radon, lead, bismuth, polonium, krypton, argon, xenon. On the cover they are lined up together and shaped into the mountain that surrounds the detector.