



**Commissioning of the world's first water  
Cherenkov neutron veto and first WIMP dark matter  
search results of the XENONnT experiment**

Dissertation zur Erlangung des Grades

Doktor der Naturwissenschaften

vorgelegt dem

Fachbereich 08 - Physik, Mathematik und Informatik  
der Johannes Gutenberg-Universität in Mainz

vorgelegt von

**Daniel Wenz**

geboren am 03. August 1992 in Oberwesel

Mainz, Mai 2023

---

Erstgutacher:  
Zweitgutachterin:  
Datum der Promotion: 04.10.2023







---

## Erklärung

Hiermit versichere ich gemäß § 12 Abs. 3 e.) der Promotionsordnung des Fachbereichs 08, Physik, Mathematik und Informatik der Johannes Gutenberg-Universität Mainz vom 02.12.2013:

- dass ich die vorgelegte Arbeit selbständig verfasst habe. Es wurden ausschließlich die angegebenen Quellen und Hilfsmittel verwendet. Von der Ordnung zur Sicherung guter wissenschaftlicher Praxis in Forschung und Lehre und vom Verfahren zum Umgang mit wissenschaftlichem Fehlverhalten habe ich Kenntnis genommen.
- dass ich die vorliegende Dissertation nicht als Prüfungsarbeit für eine andere Prüfung eingereicht habe.
- dass ich die vorliegende Abhandlung oder Teile davon noch nicht als Dissertation bei einer anderen Fakultät oder einem anderen Fachbereich eingereicht habe.

Mainz, den 01. Mai 2023

---

Daniel Wenz



---

## ABSTRACT

A rich number of astronomical and cosmological observations suggest the existence of a massive, non-luminous, and non-relativistic, matter component in the universe which is five times more abundant than baryonic matter and is commonly referred as to dark matter (DM). Although so far eluding from detection, one class of promising DM candidates are weakly interacting massive particles (WIMPs) which arise naturally from many beyond the Standard Model (BSM) theories.

The XENON Dark Matter Project aims to directly detect WIMPs, and other kinds of rare event signals, by utilizing large-scale liquid xenon (LXe) dual-phase time projection chambers (TPCs). The newest generation of experiment, called XENONnT, utilizes a TPC with a total sensitive LXe mass of 5.9 t, and was designed as a fast upgrade of its predecessor XENON1T. In addition to its larger TPC, XENONnT was augmented with the world's first water Cherenkov neutron veto (NV), which was mounted inside the already existing water Cherenkov muon veto water tank of XENON1T. Neutrons emitted by detector materials can undergo a single back-scatter inside the TPC producing a signal which is indistinguishable from WIMPs. The NV has the task to mitigate this potential threat for the scientific reach of the experiment by tagging these escaping neutrons through their delayed neutron capture on hydrogen.

In the presented work, the results of the first weakly interacting massive particle (WIMP) search science run, called SR0, are discussed. SR0 features a blind analysis between 3.3 keV and 60.5 keV nuclear recoils energies with a total exposure of about 1.1 tonne-year, utilizing the lowest ever achieved electronic recoil background of  $(15.8 \pm 1.3)$  events/(t · y · keV) in a LXe. No significant excess was found in the data, setting the lowest upper limit of  $2.58 \cdot 10^{-47}$  cm<sup>2</sup> for spin-independent (SI) interactions of 28 GeV/c<sup>2</sup> WIMPs at a 90 % confidence level. These results have also been published in [Apr+23b] as part of the presented work.

To obtain these results, this thesis discusses the commissioning of the XENONnT neutron veto (NV), and the calibration of its neutron tagging efficiency. The tagging efficiency was found to be  $(53.1 \pm 2.8)$  % which is the highest efficiency ever measured in a water Cherenkov detector. The efficiency of the NV, as well as the nuclear recoil (NR) response of the time projection chamber (TPC), were calibrated using tagged neutrons from an Americium-Beryllium (AmBe) neutron source. This technique was deployed for the first time in a liquid xenon (LXe) TPC. It enables a calibration of the NR response with high purity and a remaining pollution of less than 0.1 %. Further, the same calibration data was used to determine the thermal neutron capture cross section of hydrogen which was found to be  $336.7 \pm 0.4$  (stat.) $^{+2.0}_{-0.0}$  (sys.) mb. All these analyses are based on the data provided by XENONnT's new processing framework called STRAXEN. As part of the lead developing team, the entire processing chain for the two veto systems of XENONnT was developed, and many additional tools have been implemented. Finally, to enhance the neutron tagging efficiency of the NV even further, the water inside the water tank is going to be doped with Gd-sulfate. As part of the presented work, different Gd-salt samples of the manufacturer Treibacher were analyzed regarding their suitability for the experiment.



---

## ZUSAMMENFASSUNG

Eine große Anzahl an astronomischen und kosmologischen Beobachtungen legt die Existenz einer massiven, dunklen und nicht-relativistischen Materiekomponente im Universum nahe, welche allgemein hin als Dunkle Materie (DM) bezeichnet wird. Bislang hat sich DM jeglichem direkten Nachweis entzogen. Eine vielversprechende Klasse von DM-Teilchen sind weakly interacting massive particles WIMPs, welche aus vielen Theorien jenseits des Standardmodells der Teilchenphysik hervorgehen. Das XENON Dark Matter Project versucht diese und andere seltene Ereignisse mit Hilfe einer Flüssig-Xenon-Zeitprojektionskammer (LXe TPC) direkt nachzuweisen. Das neueste Experiment, genannt XENONnT, umfasst eine TPC mit einer Detektormasse von 5.9 Tonnen und wurde als eine Erweiterung des Vorgängerexperiments XENON1T konzipiert. Zuzüglich zu dieser neuen TPC wurde der Wassertank des XENONnT Experiments um das erste Wasser Cherenkov Neutronenveto (NV) der Welt erweitert. Neutronen, welche von Materialien des Detektors emittiert werden, können WIMP-Signale imitieren, indem sie lediglich ein einziges Mal elastisch in der TPC streuen bevor sie diese verlassen. Das NV markiert diese Signale über den verzögerten Neutroneneinfang von Wasserstoff, welcher zur Emission eines 2.22 MeV Gammas führt.

In der hier präsentierten Arbeit sind die ersten Resultate der Suche nach WIMPs mit XENONnT vorgestellt. Hierfür wurden Daten in einem Energiebereich zwischen 3.3 keV und 60.5 keV mit einer Gesamtlaufzeit von 1.1 Tonnen-Jahren in einer „blinden“ Analyse ausgewertet. Keine signifikanten Anzeichen für die Existenz von WIMPs wurden gefunden. Hierdurch konnte ein neues Ausschlusslimit von  $2.58 \cdot 10^{-47} \text{ cm}^2$  für Spin-Unabhängige Interaktionen für WIMPs mit einer Masse von  $28 \text{ GeV}/c^2$  bei einer Konfidenz von 90 % bestimmt werden. Dieses Ergebnis wurde als Teil der präsentierten Arbeit in [Apr+23b] veröffentlicht.

Die Bestimmung der Neutronennachweiseffizienz des NV in der präsentierten Arbeit ist ein wesentlicher Bestandteil dieses Ergebnisses. Sie wurde zu  $(53.1 \pm 2.8) \%$  gemessen, welches die höchste jemals gemessene Neutronennachweiseffizienz in einem Wasser Cherenkov Detektor darstellt. Die Nachweiseffizienz so wie das Detektorverhalten der TPC gegenüber neutralen Teilchen wie Neutronen oder WIMPs wurde mit Hilfe von „markierten“ Neutronen einer Americium-Beryllium (AmBe) Quelle kalibriert. Diese Technik wurde in der vorliegenden Arbeit das erste Mal in einer LXe TPC verwendet und ermöglicht eine besonders reine Kalibration der Detektorresponse mit einer Restkonatimnation von  $<0.1 \%$ . Die gleiche Technik wurde auch dazu verwendet um mit Hilfe des NV den Neutroneneinfangquerschnitt von Wasserstoff zu messen welcher zu  $336.7 \pm 0.4 \text{ (stat.)}_{-0.0}^{+2.0} \text{ (sys.) mb}$  bestimmt wurde. Alle präsentierten Analysen basieren auf XENONnT's neuer Analysesoftware. Als Teil des Entwicklerteams wurde die gesamte Analysesoftware für das NV entwickelt und getestet. Darüber hinaus ist geplant, das Wasser des NV mit Gadolinium-Sulfat zu versetzen, um die Nachweiseffizienz des NV weiter zu verbessern. Hierzu wurden in der präsentierten Arbeit verschiedene Salzproben des Herstellers Treibacher bezüglich ihrer Eignung hierfür analysiert.

---

## ACRONYMS

<b>ACs</b>	accidental coincidences
<b>ADC</b>	analog-to-digital converter
<b>ALPs</b>	axion like particles
<b>AmBe</b>	Americium-Beryllium
<b>BBF</b>	Bayesian Band Fitting Framework
<b>BBN</b>	big bang nucleosynthesis
<b>BSM</b>	beyond the Standard Model
<b>CE</b>	charge collection efficiency
<b>CE<math>\nu</math>NS</b>	coherent elastic neutrino-nucleus scattering
<b>CMB</b>	cosmic microwave background
<b>CMT</b>	Correction Management Tool
<b>CNC</b>	Computerized Numerical Control
<b>CPU</b>	central processing unit
<b>CSDA</b>	continuous-slowing-down approximation
<b>CUTAX</b>	Cuts for analysing Xenon data
<b>CY</b>	charge yield
<b>DAQ</b>	data acquisition
<b>DM</b>	dark matter
<b>DPE</b>	double photo-electron
<b>EPIX</b>	Electron and Photon Instructions generator for XENON
<b>ER</b>	electronic recoil



## Acronyms

---

**GBDT** gradient boosted decision tree

**GPS** global pointing system

**HEV** high energy veto

**HTML** HyperText Markup Language

**ISM** interstellar medium

**LCE** light collection efficiency

**LED** light emitting diode

**LET** linear energy transfer

**LXe** liquid xenon

**LY** light yield

**MC** Monte Carlo

**MCMC** Markov chain Monte Carlo

**MS** multi-scatter

**MV** muon veto

**ndof** Number of degrees of freedom

**NEST** Noble Element Simulation Technique

**NIM** Nuclear Instrumentation Module

**NR** nuclear recoil

**NV** neutron veto

**PAX** Processor for Analyzing XENON

**PBHs** primordial black holes

**PbS** polycrystalline lead sulfide detectors

**PMT** photomultiplier tube

**PTFE** polytetrafluoroethylene

**QE** quantum efficiency

**RAM** random access memory

**REDAX** "Distributed, parallel high-throughput data acquisition software for XENONnT"

**RICH** Ring Imaging Cherenkov

**ROI** region of interest

**SD** spin-dependent

**SE** single electron

**SI** spin-independent

**SK** Super-Kamiokande

**SNEWS** SuperNova Early Warning System

**SNO** Sudbury Neutrino Observatory

**SPE** single photo-electron

**SRO** first science run

**STRAX** Streaming analysis for xenon experiments

**STRAXEN** Streaming analysis for XENON(nT)

**TPC** time projection chamber

**TTS** transient time spread

**VUV** vacuum ultraviolet

**WFSim** The XENON waveform simulator

**WIMP** weakly interacting massive particle



---

# TABLE OF CONTENTS

	List of Tables	XVII
	List of Figures	XIX
	Introduction	1
<b>1</b>	<b>CHAPTER 1</b> DARK MATTER AND DIRECT DETECTION	3
1.1	<i>Evidence for Dark Matter</i> . . . . .	3
1.2	<i>Dark Matter candidates and properties</i> . . . . .	6
1.3	<i>Astrophysical inputs for the direct dark matter search</i> . . . . .	7
1.4	<i>Direct Dark Matter detection</i> . . . . .	8
<b>2</b>	<b>CHAPTER 2</b> THE XENONNT EXPERIMENT	15
2.1	<i>Liquid xenon as detection medium</i> . . . . .	15
2.1.1	<i>Liquid xenon properties</i> . . . . .	15
2.1.2	<i>Signal generation and reconstruction</i> . . . . .	17
2.2	<i>Dual-phase time projection chambers</i> . . . . .	24
2.2.1	<i>Signal readout</i> . . . . .	24
2.2.2	<i>Particle discrimination and fiducialization</i> . . . . .	26
2.2.3	<i>Signal correction and energy reconstruction</i> . . . . .	28
2.3	<i>Water Cherenkov detectors and neutron detection in water</i> . . . . .	34
2.4	<i>The XENONnT experiment</i> . . . . .	39
2.4.1	<i>XENONnT TPC</i> . . . . .	39
2.4.2	<i>Auxiliary infrastructure: xenon purity</i> . . . . .	42
2.4.3	<i>Water Cherenkov Muon- and Neutron-Veto</i> . . . . .	43
2.4.4	<i>Auxiliary infrastructure: Data acquisition</i> . . . . .	47

---

<b>3</b>	<b>CHAPTER 3</b> XENONNT ANALYSIS SOFTWARE	<b>51</b>
3.1	<i>Strax and Straxen</i> . . . . .	52
3.1.1	<i>Context, Corrections, Lineage and Storage</i> . . . . .	52
3.1.2	<i>Data Structure: Runs, Chunks and Arrays</i> . . . . .	55
3.1.3	<i>Plugins</i> . . . . .	57
3.2	<i>Signal types: Hits, Peaks and Events</i> . . . . .	60
3.3	<i>Plugins for the XENONnT neutron veto</i> . . . . .	64
3.3.1	<i>Raw-data, software trigger and data reduction</i> . . . . .	66
3.3.2	<i>Hitlets</i> . . . . .	67
3.3.3	<i>Veto-Events</i> . . . . .	70
3.3.4	<i>Detector synchronization and linking of unlinked runs</i> . . . . .	71
3.3.5	<i>Veto-Intervals and Deadtime</i> . . . . .	74
3.4	<i>Plugin performance and storage consumption</i> . . . . .	75
3.5	<i>Auxiliary software enhancements</i> . . . . .	76
3.5.1	<i>Interactive event-displays and plots</i> . . . . .	78
3.6	<i>Conclusion and outlook</i> . . . . .	85
<b>4</b>	<b>CHAPTER 4</b> GADOLINIUM SULFATE	<b>87</b>
4.1	<i>Gadolinium-Sulfate properties</i> . . . . .	87
4.2	<i>Transparency measurements</i> . . . . .	88
4.3	<i>Gadolinium Sulfate Procurement</i> . . . . .	91
4.4	<i>Conclusion and Outlook</i> . . . . .	95
<b>5</b>	<b>CHAPTER 5</b> NEUTRON-VETO RESPONSE AND PERFORMANCE	<b>97</b>
5.1	<i>NV detector conditions during SR0</i> . . . . .	97
5.2	<i>NV event center time and background distribution</i> . . . . .	98
5.3	<i>Conclusion and outlook</i> . . . . .	106
<b>6</b>	<b>CHAPTER 6</b> NUCLEAR RECOIL CALIBRATION	<b>107</b>
6.1	<i>Working principle of the AmBe coincidence and SR0 data taking</i> . . . . .	107
6.1.1	<i>Americium-Beryllium source</i> . . . . .	107
6.1.2	<i>SR0 AmBe calibration</i> . . . . .	109
6.2	<i>Data selection, optimization and cut efficiency</i> . . . . .	112
6.2.1	<i>NV gamma selection and detector coincidence</i> . . . . .	112
6.2.2	<i>Data quality cuts, fiducial volume and NR band data</i> . . . . .	120
6.2.3	<i>Cut acceptances, correlation and total cut acceptance</i> . . . . .	129

6.3	<i>NR band fit and response model</i> . . . . .	136
6.3.1	<i>NR band contours</i> . . . . .	137
6.3.2	<i>NR response model</i> . . . . .	138
6.4	<i>Conclusion and Outlook</i> . . . . .	142
<b>7</b>	<b>CHAPTER 7</b> CALIBRATION OF THE NEUTRON VETO	145
7.1	<i>Empirical model of the AmBe energy spectrum:</i> . . . . .	145
7.2	<i>Neutron veto response and AmBe template model using toy MCs</i> . .	150
7.2.1	<i>NV toy MC</i> . . . . .	150
7.2.2	<i>Template modeling and <math>^{12}\text{C}</math> branching ratio</i> . . . . .	153
7.2.3	<i>NV response</i> . . . . .	163
7.3	<i>Tagging efficiency</i> . . . . .	166
7.3.1	<i>neutron capture signals and coincidence time distribution</i> . .	167
7.3.2	<i>Veto window correction</i> . . . . .	169
7.3.3	<i>Geometric correction</i> . . . . .	173
7.3.4	<i>SR0 tagging efficiency and deadtime</i> . . . . .	174
7.3.5	<i>Systematic uncertainties</i> . . . . .	176
7.4	<i>Detection efficiency</i> . . . . .	179
7.5	<i>Neutron veto self-coincidence</i> . . . . .	183
7.5.1	<i>4.44 MeV gamma-ray trigger signal</i> . . . . .	184
7.5.2	<i>Neutron capture on different components:</i> . . . . .	188
7.5.3	<i>Thermal hydrogen neutron capture cross section:</i> . . . . .	190
7.6	<i>Conclusion and Outlook</i> . . . . .	191
<b>8</b>	<b>CHAPTER 8</b> SCIENCE RUN ZERO WIMP RESULTS	193
8.1	<i>SR0 science data summary</i> . . . . .	193
8.2	<i>Neutron veto tagged unblinding and neutron-background</i> . . . . .	195
8.3	<i>WIMP search results</i> . . . . .	201
8.3.1	<i>Remaining background components</i> . . . . .	201
8.3.2	<i>Unblinding and WIMP results</i> . . . . .	202
<b>9</b>	<b>CHAPTER 9</b> CONCLUSION AND OUTLOOK	209
<b>A</b>	<b>CHAPTER A</b> SUPPLEMENTAL MATERIAL	213
A.1	<i>The XYZCamScan inspection setup</i> . . . . .	213
A.1.1	<i>Inspection setup:</i> . . . . .	214
A.1.2	<i>Optical inspection</i> . . . . .	217
A.1.3	<i>Wire sagging measurement</i> . . . . .	223
A.1.4	<i>Conclusion and outlook</i> . . . . .	225

## Table of Contents

---

A.2	<i>Accidental coincidence rate</i> . . . . .	228
A.3	<i>Background rate increase due to Gd-loading</i> . . . . .	231
A.4	<i>NR data quality cuts</i> . . . . .	232
A.5	<i>S2-only analysis using the NR calibration data</i> . . . . .	235
<b>B</b>	<b>CHAPTER B</b> APPENDIX	237
	Bibliography	287
	Curriculum Vitae	298

---

## LIST OF TABLES

<b>2.1.</b>	All natural xenon isotopes, their abundance and half-life if not stable. . . . .	16
<b>4.1.</b>	Summary table of Gd-sulfate samples provided by Treibacher. . . . .	91
<b>5.1.</b>	Best fit values for the upper and lower boundary of the center time cut. . . . .	103
<b>6.1.</b>	Table of the source positions used during the AmBe calibration of SR0. . . . .	111
<b>6.2.</b>	Selection statistics for the NV $^{12}\text{C}$ $\gamma$ -ray selection. . . . .	114
<b>6.3.</b>	Best fit values of the $\gamma$ - $\gamma$ -coincidence peak and timing resolution of the detector coincidence. . . . .	118
<b>6.4.</b>	Comparison of the average number of photons and electrons generated by a 4 keV plus 6 keV multi-scatter NR interaction and a single 10 keV NR interaction. . . . .	123
<b>6.5.</b>	Fit statistics of the sliced NR contour fits. . . . .	137
<b>7.1.</b>	Best fit values of the simulated template fits for the four different AmBe source positions. . . . .	157
<b>7.2.</b>	Best fit results of the time distribution for the NV tagging efficiency. . . . .	170
<b>7.3.</b>	Best fit results of the 4.44 MeV $\gamma$ -ray ellipse in the TPC. . . . .	179
<b>7.4.</b>	Best fit results of the time distribution for the NV tagging efficiency. . . . .	180
<b>8.1.</b>	Number of expected events in the SR0 WIMP analysis. . . . .	201
<b>A.1.</b>	Cut statistics of the NR calibration. . . . .	233
<b>B.1.</b>	Best fit values of the neutron-veto AmBe energy spectrum for the SR0 source positions. . . . .	264





---

## LIST OF FIGURES

<b>1.1.</b>	Rotation curve of the galaxy M33. . . . .	4
<b>1.2.</b>	Anisotropies of the cosmic microwave background (CMB) and the best fit of the Planck 2018 results. . . . .	5
<b>1.3.</b>	Differential WIMP recoil spectra for different WIMP masses. . . . .	10
<b>1.4.</b>	spin-dependent (SD) differential rate for WIMP-nucleus interaction of a 100 GeV/c <sup>2</sup> WIMP. . . . .	11
<b>1.5.</b>	Current best limits for the spin-independent WIMP-nucleon cross section of direct detection experiments as of 07.03.2023, excluding the most recent limits by XENONnT and LZ. . . . .	12
<b>2.1.</b>	Mass attenuation coefficient of $\gamma$ -rays in liquid xenon as a function of photon-energy. . . . .	18
<b>2.2.</b>	Neutron interaction cross section of xenon as a function of energy. . . . .	19
<b>2.3.</b>	Field dependence of the charge and light yield in liquid xenon. . . . .	21
<b>2.4.</b>	Photon and electron yields for nuclear and electronic recoils. . . . .	23
<b>2.5.</b>	Working principle of dual-phase time projection chamber. . . . .	25
<b>2.6.</b>	Electronic and nuclear recoil calibration of the XENON1T experiment. . . . .	27
<b>2.7.</b>	Example fiducial volume of the XENON1T low ER excess analysis. . . . .	28
<b>2.8.</b>	Comparison of the S1 response of XENONnT for the lines of <sup>83m</sup> Kr w/ and w/o S1 corrections applied. . . . .	30
<b>2.9.</b>	Full energy ellipses of the isotopes <sup>129m</sup> Xe and <sup>131m</sup> Xe measured with the XENONnT experiment, and illustration of the "Doke" method. . . . .	33
<b>2.10.</b>	Illustration of the Cherenkov effect when a charged particle traversing a dielectric medium. . . . .	35
<b>2.11.</b>	Mass attenuation of $\gamma$ -rays in water. . . . .	36
<b>2.12.</b>	Neutron interaction cross sections for water. . . . .	37
<b>2.13.</b>	Photograph of the XENONnT experiment. . . . .	40
<b>2.14.</b>	Rendering of the XENONnT TPC. . . . .	41
<b>2.15.</b>	Rendered CAD image of the XENONnT experiment. . . . .	46
<b>2.16.</b>	Schematic of the XENONnT data acquisition (DAQ) system. . . . .	48
<b>3.1.</b>	Overview of the different components provided by the Streaming analysis for xenon experiments (STRAX) package. . . . .	53
<b>3.2.</b>	STRAX data key and lineage hash. . . . .	55
<b>3.3.</b>	STRAX data structure and chunking. . . . .	56

---

<b>3.4.</b>	General structure of a STRAX plugin. . . . .	58
<b>3.5.</b>	Illustration of different Streaming analysis for XENON(nT) (STRAXEN) data types used in the TPC processing chain. . . . .	60
<b>3.6.</b>	Schematic of the most important steps in the TPC event building. . . . .	61
<b>3.7.</b>	Illustration raw_records and records for a single NV photomultiplier tube (PMT) pulse. . . . .	62
<b>3.8.</b>	Dependents chain of the NV processing. . . . .	65
<b>3.9.</b>	Example and schematic illustrating the NV software trigger. . . . .	67
<b>3.10.</b>	Hitlet splitting illustration. . . . .	69
<b>3.11.</b>	Example hitlet illustrating the different shape parameters. . . . .	70
<b>3.12.</b>	Illustration of the synchronization of the NV with the TPC during first science run (SR0). . . . .	72
<b>3.13.</b>	Performance and storage consumption of the major NV data types. . . . .	76
<b>3.14.</b>	Interactive TPC event display. . . . .	79
<b>3.15.</b>	Interactive features of the TPC event display. . . . .	80
<b>3.16.</b>	Interactive zoom feature of the TPC event display. . . . .	81
<b>3.17.</b>	Interactive NV event display. . . . .	82
<b>3.18.</b>	Dynamic features of the NV event display . . . . .	83
<b>3.19.</b>	New TPC event display. . . . .	84
<b>4.1.</b>	Settings screen of the UV/VIS spectrometer Lambda 850 of the manufacturer PerkinElmer. . . . .	89
<b>4.2.</b>	Absorption spectra of the GdSu19-04 sample with respect to ultra pure water and the Japanese Gd-sulfate salt. . . . .	92
<b>4.3.</b>	Absorbance of the filtered Treibacher GdSu20-01A sample with respect to pure water and other salt samples. . . . .	94
<b>4.4.</b>	Transparency of different Gd-sulfate samples with respect to ultra-pure water. . . . .	95
<b>5.1.</b>	NV SR0 background spectrum. . . . .	99
<b>5.2.</b>	Center time of the NV events as a function of event area for SR0 background and AmBe calibration data. . . . .	100
<b>5.3.</b>	Background subtracted AmBe event center time as a function of event area. . . . .	102
<b>5.4.</b>	Spatial distribution of the fifth background component of Figure 5.2 . . . . .	104
<b>5.5.</b>	Area distribution of the NV events found in the fifth population of Figure 5.2. . . . .	105
<b>6.1.</b>	Neutron production cross sections of $^9\text{Be}$ and energy levels of $^{12}\text{C}$ . . . . .	108
<b>6.2.</b>	Neutron yield as function of neutron-energy for different $\alpha$ sources and $^9\text{Be}(\alpha, n)^{12}\text{C}$ reaction channels. . . . .	109
<b>6.3.</b>	Schematic of the TPC NR and NV tagging efficiency calibration. . . . .	110
<b>6.4.</b>	Spatial distribution of the NR calibration events in the TPC for the individual calibration source positions. . . . .	112
<b>6.5.</b>	Selection of NV trigger events for the TPC NR calibration. . . . .	113
<b>6.6.</b>	Time difference distribution for coincident NV and TPC events. . . . .	115

6.7.	Time distribution of the TPC S1 signals in the NR ROI with respect to triggering NV events. . . . .	116
6.8.	Time resolution of the NV-TPC coincidence for the NR response calibration. . . . .	117
6.9.	AmBe nuclear recoil calibration data before and after applying the NV TPC coincidence. . . . .	119
6.10.	Impact of the S2 single-scatter cut on the NR calibration data. . . . .	122
6.11.	Impact of the S2 width cut on the NR calibration data. . . . .	123
6.12.	cS2 versus cS1 space for the AmBe charge-loss events. . . . .	124
6.13.	Spatial distribution of charge-loss events in AmBe calibration. . . . .	125
6.14.	Optimized AmBe fiducial volume and NR calibration data. . . . .	126
6.15.	AmBe charge-loss events after applying the AmBe fiducial volume cut. . . . .	127
6.16.	AmBe NR and electronic recoil (ER) calibration data passing all data quality selections. . . . .	129
6.17.	cS2 versus cS1 space showing all NR single-scatter events within the 90 % contour of the NR band. . . . .	132
6.18.	Example illustrating the Bayesian approach for the cut acceptance uncertainty. . . . .	133
6.19.	Total NR cut acceptance as a function of cS1 and cS2. . . . .	134
6.20.	Correlation matrix of all cuts applied to the NR calibration data. . . . .	135
6.21.	Comparison between the product of the two n-1 cut acceptances compared to the n-2 cut acceptance for the S1 pattern bottom and top cuts. . . . .	136
6.22.	Fit of the upper and lower boundary of 90 % contour for the NR calibration event. . . . .	138
6.23.	P-value evaluation of the best fit of the NR calibration data using the BBF fitting framework. . . . .	140
6.24.	Best fit of the NR calibration data using the BBF fitting framework. . . . .	142
7.1.	Energy spectrum of the AmBe calibration taken at the bottom CW11d2m source position. . . . .	147
7.2.	Best fit of the NV AmBe energy spectrum taken at the bottom CW11d2m source position. . . . .	149
7.3.	Schematic of the effective NV simulation framework. . . . .	151
7.4.	Average PMT quantum efficiency of the NV PMTs. . . . .	152
7.5.	Example probability density distributions of the single photo-electron response for the NV PMTs. . . . .	153
7.6.	Comparison of the NV response for first and second excited state of the $^{12}\text{C}$ with and without source casing. . . . .	154
7.7.	Comparison of the center time distribution between data and simulations for the data taken at the bottom CW11d2m source position. . . . .	155
7.8.	Comparison of the spatial distribution between data and simulations for the data taken at the bottom CW11d2m source position . . . . .	156
7.9.	Best fit of the simulated templates with the AmBe calibration data taken at the bottom CW11d2m source position. . . . .	158

<b>7.10.</b>	Branching ratio and fraction of high energy neutron capture events as a function of radial distance towards the TPC cryostat . . . . .	160
<b>7.11.</b>	Branching ratios of the first excited state based on the data taken from Figure 6.1. . . . .	162
<b>7.12.</b>	Best fits of the simulated mono-energetic pseudo $\gamma$ -lines using the NV simulation framework. . . . .	163
<b>7.13.</b>	Predicted NV energy response and resolution using toy MCs. . . . .	165
<b>7.14.</b>	Schematic showing the working principle of the NV tagging for radiogenic neutrons emitted by the TPC materials (left) and the calibration of the NV tagging efficiency via AmBe (right). . . . .	166
<b>7.15.</b>	Time spectrum showing all NV events found in coincidence with a single scatter NR event for the tagging efficiency study. . . . .	167
<b>7.16.</b>	Hydrogen-capture peak selection for the computation of the NV tagging efficiency. . . . .	168
<b>7.17.</b>	Best fit of the time distribution for the NV tagging efficiency study. . . . .	172
<b>7.18.</b>	Comparison of the simulated and data driven estimates for the tagging efficiency. . . . .	174
<b>7.19.</b>	SR0 NV tagging efficiency. . . . .	175
<b>7.20.</b>	Detector deadtime versus NV tagging efficiency. . . . .	176
<b>7.21.</b>	Systematic bias study for the integration time window correction of the NV tagging efficiency. . . . .	178
<b>7.22.</b>	Best fit of the 4.44 MeV $\gamma$ -ellipse in the cS2 versus cS1 space. . . . .	180
<b>7.23.</b>	Best fit of the time distribution for the NV detection efficiency. . . . .	181
<b>7.24.</b>	Event area distribution for the NV detection efficiency study. . . . .	182
<b>7.25.</b>	Best fit of the NV self-coincidence time distribution using the 4.44 MeV $\gamma$ -ray as the trigger signal for the TopCCW5d0m source position. . . . .	184
<b>7.26.</b>	NV energy spectrum for all events found in a NV self-coincidence during AmBe calibration at the topCCW5d0m source position. . . . .	185
<b>7.27.</b>	Fraction of the high energy tail of the NV energy spectrum found during AmBe calibration as a function of radial distance of the source position towards the outer TPC cryostat. . . . .	186
<b>7.28.</b>	Best fit of the hydrogen neutron capture peak for the topCCW5d0m source position derived from the NV only coincidence data. . . . .	187
<b>7.29.</b>	Neutron capture constant for the four AmBe calibration source positions as a function of radial distance to the TPC cryostat. . . . .	188
<b>7.30.</b>	Decomposed NV energy spectrum of simulated neutrons at the topCW5d9m source position, using custom made Geant4 saveflags. . . . .	189
<b>8.1.</b>	ROI, detection and selection efficiency of the SR0 WIMP analysis. . . . .	195
<b>8.2.</b>	SR0 unblinded veto tagged events. . . . .	197
<b>8.3.</b>	Spatial distribution of the neutron veto tagged neutron candidate events in SR0. . . . .	198
<b>8.4.</b>	Neutron veto events which tagged a SR0 TPC event. . . . .	199
<b>8.5.</b>	Unblinded WIMP region of interest (ROI) for SR0. . . . .	203
<b>8.6.</b>	Spatial distribution of the SR0 unblinded WIMP events. . . . .	204

<b>8.7.</b>	Upper limit on the spin-independent WIMP-nucleon cross section at 90% confidence level . . . . .	206
<b>8.8.</b>	Comparison of the best upper limit obtained in the first science run of XENONnT with other experiments. . . . .	207
<b>8.9.</b>	Upper limits for the spin-dependent WIMP-neutron and WIMP-proton cross section. . . . .	208
<b>A.1.</b>	Photograph of the XYZCamScan inspection setup. . . . .	214
<b>A.2.</b>	Comparison of the image resolution for the different optical systems of the XYZCamScan setup. . . . .	216
<b>A.3.</b>	Exemplary height profiles and images for one of the wire damages found in the XENON1T gate mesh, taken with the confocal microscope. . . . .	217
<b>A.4.</b>	Example measurement showing the normalized contour sharpness estimated by the XYZCamScan autofocus as function of the absolute z-position of the ISEL system. . . . .	219
<b>A.5.</b>	Images belonging to the labeled points in Figure A.4. . . . .	221
<b>A.6.</b>	Illustration of the XENON1T cathode electrode inspection. . . . .	222
<b>A.7.</b>	Distance measurement between the micro-epsilon laser and XENON1T cathode mesh. . . . .	224
<b>A.8.</b>	Measurements of the of the gantry sagging and cathode frame reference plane. . . . .	225
<b>A.9.</b>	Best fit of the catenary curve to wire 47 of the XENON1T cathode. . . . .	226
<b>A.10.</b>	Horizontal wire tension of the XENON1T cathode wires. . . . .	227
<b>A.11.</b>	Image of the extended wire inspection setup. . . . .	228
<b>A.12.</b>	Schematic for the derivation of the analytical accidental rate model explained in the text. . . . .	229
<b>A.13.</b>	Comparison between the analytical and toy MC model for the rate of accidental coincidences for a $n$ -fold coincidence with $m$ counter and a resolving time $T$ . . . . .	230
<b>A.14.</b>	Comparison of the analytical accidental dark rate model with the toy data simulation using two fake detector configurations. . . . .	231
<b>A.15.</b>	Sketch of the neutron-veto aided S2-only NR calibration. . . . .	235
<b>A.16.</b>	Illustration of the S2 width versus "S2-only drift time" relation. . . . .	236
<b>B.1.</b>	S1 correction map of XENONnT's first science run. . . . .	238
<b>B.2.</b>	S2 correction map of XENONnT's first science run. . . . .	239
<b>B.3.</b>	Photograph of the anode electrode showing the transverse wires. . . . .	239
<b>B.4.</b>	Localized region of an intermittent high emission of single electron S2 single in XENONnT. . . . .	240
<b>B.5.</b>	Peak(let) classification boundaries for S1 and S2 in SR0. . . . .	240
<b>B.6.</b>	Maximum gap threshold for the S2 peaklet merging as function of the S2 area. . . . .	241
<b>B.7.</b>	TPC event display with record matrix. . . . .	241
<b>B.8.</b>	Fraction of Gd-neutron capture for a gd-loaded water Cherenkov detector as a function of Gd-mass fraction. . . . .	242
<b>B.9.</b>	Images of the Gd20-01 and Gd20-01A salt samples from Treibacher. . . . .	243

---

<b>B.10.</b> Titration of a Gd-loaded water sample with sulfuric acid. . . . .	244
<b>B.11.</b> Technical drawing of the AmBe source used in XENONnT. . . . .	245
<b>B.12.</b> XENONnT field simulations and charge insensitive mass. . . . .	246
<b>B.13.</b> Example waveform of the NV containing digitizer bit-flips. . . . .	247
<b>B.14.</b> Qualitative comparison of the center time distribution between toy MC data and AmBe calibration data. . . . .	248
<b>B.15.</b> Center time cut boundaries compared against the NV SR0 background data. . . . .	249
<b>B.16.</b> Reconstructed spatial position of NV background events. . . . .	250
<b>B.17.</b> Number of contributing PMTs as a function of event area for SR0 NV background events. . . . .	250
<b>B.18.</b> Hit pattern of a typical population 5 event discussed in Section 5.2. . . . .	251
<b>B.19.</b> Comparison of the spatial distribution of the observed and simulated ER background of XENONnT. . . . .	251
<b>B.20.</b> Comparison of the detector timing resolution of NR data and the high energy $\gamma$ - $\gamma$ -coincidences. . . . .	252
<b>B.21.</b> Fraction of remaining AmBe calibration events in the NR region after applying the NV-TPC coincidence. . . . .	253
<b>B.22.</b> Cut boundaries for the S2 single-scatter cut. . . . .	253
<b>B.23.</b> Signal to charge-loss separation as a function of the maximal radius for the AmBe fiducial volume cut. . . . .	254
<b>B.24.</b> Selection of NR charge-loss events selection. . . . .	255
<b>B.25.</b> Remaining outlier events in the orange region shown in figure 6.15 shown in the alt. S2 versus S2 space. . . . .	255
<b>B.26.</b> Relative difference in light yield and charge yield difference for 5 keV <sub>NR</sub> NR interactions. . . . .	256
<b>B.27.</b> NR charge loss events selection. . . . .	257
<b>B.28.</b> n-1 method for the S1 max PMT cut. . . . .	257
<b>B.29.</b> n-1 method for the S1 naive bayes cut. . . . .	258
<b>B.30.</b> n-1 method for the S1 width cut. . . . .	258
<b>B.31.</b> n-1 method for the S1 width cut. . . . .	259
<b>B.32.</b> n-1 method for the S2 area fraction top cut. . . . .	259
<b>B.33.</b> n-1 method for the S2 pattern cut. . . . .	260
<b>B.34.</b> n-1 method for the S2 pos diff cut. . . . .	260
<b>B.35.</b> Cut acceptance comparison for the correlated cuts cut_s1_naive_bayes and cut_s1_width as well as cut_s2_pattern and cut_s2_recon_pos_diff. . . . .	261
<b>B.36.</b> Single slice fit for the NR band contour estimate. . . . .	262
<b>B.37.</b> Posterior distribution of $\kappa$ for the best fit of the NR response model. . . . .	263
<b>B.38.</b> Best fit of the neutron-veto AmBe energy spectrum for the remaining source positions. . . . .	265
<b>B.39.</b> Best fit of the tagging efficiency time distribution including multi-scatter events. . . . .	266
<b>B.40.</b> Comparison of the tagging efficiency w/ and w/o multi-scatter NR signals. . . . .	267
<b>B.41.</b> Fit range for the 4.44 MeV $\gamma$ -ray ellipse fit. . . . .	268

---

<b>B.42.</b>	Rendering of the XENONnT experimental setup used in Geant4. . . . .	268
<b>B.43.</b>	Impact of different optical properties on the neutron-veto center time. . .	269
<b>B.44.</b>	Best fits of the simulation templates to the AmBe spectra taken at different source positions. . . . .	270
<b>B.45.</b>	Zoom onto two different regions for the best fit of the time distribution of the neutron-veto detection efficiency study. . . . .	271
<b>B.46.</b>	Neutron veto detection efficiency as function of event area threshold. . .	272
<b>B.47.</b>	Example best fit for the simulated 4 MeV full energy line. . . . .	272
<b>B.48.</b>	Zoom in the region $\pm 1 \mu\text{s}$ of the NV self coincidence for the topCCW5d0m source position. . . . .	273
<b>B.49.</b>	Best fits of the neutron capture time distribution recorded in the NV self-coincidence. . . . .	274
<b>B.50.</b>	Best fits of the 2.22 MeV neutron capture gamma line recorded by the NV self-coincidence. . . . .	275
<b>B.51.</b>	Comparison of the time distribution of the H-capture tagged events, simulated with Geant4. . . . .	276
<b>B.52.</b>	Decomposed NV energy spectrum of simulated neutrons at the topCCW5d0m source position, using custom made Geant4 saveflags. . . . .	277
<b>B.53.</b>	Time distribution of simulated neutrons events with (top) and without (bottom) u-tubes. . . . .	278
<b>B.54.</b>	Mono-energetic calibration lines for the XENONnT energy response. . .	279
<b>B.55.</b>	Energy response calibration of SR0. . . . .	280
<b>B.56.</b>	Neutron background templates used in the SR0 WIMP inference. . . . .	281
<b>B.57.</b>	Comparison of the ER and NR band model. . . . .	282
<b>B.58.</b>	ER background model for the SR0 WIMP ROI. . . . .	283
<b>B.59.</b>	SR0 surface background model template in cS1 and cS2. . . . .	283
<b>B.60.</b>	AC background model for the SR0 WIMP search. . . . .	284
<b>B.61.</b>	AC background model validation. . . . .	285





---

## INTRODUCTION

Questions like "Where are we coming from?" or "How did our universe form?" are probably almost as old as humankind itself. In modern physics one important aspect of these questions is the formation of the universe and galaxies, which are undeniably directly related to our existence. In this context, the so-called dark matter (DM) plays an important role. A large number of astronomical and cosmological observations suggest the existence of this invisible matter component which is five times more abundant than baryonic matter. DM seems to influence and shape our universe on all mass and time scales through its gravitational interaction with regular matter. Yet, we still do not know what it is made of [Sch19]. Numerous different theories and models exist which try to explain DM in different ways. These theories reach from beyond the Standard Model (BSM) particles to large macroscopic objects like primordial black holes [BHS05]. However, so far DM eluded any detection beyond gravity.

One class of DM particle which arises from many BSM theories are so-called weakly interacting massive particles (WIMPs). These particles typically have masses in the range of  $\text{GeV}/c^2$  to  $\text{TeV}/c^2$ , and their interaction strength with standard model particles is expected to be on the weak scale or below [Sch19; Coo22]. The quest for the direct detection of WIMPs in a mass range of a few  $\text{GeV}/c^2$  to  $\text{TeV}/c^2$  is led by a number of liquid xenon (LXe) dual-phase time projection chambers (TPCs) [Men+21; Aal+22; Apr+23b]. These highly sensitive, low background experiments can detect potential WIMP signals through energy deposits of recoiling xenon nuclei when scattering elastically with WIMPs. In order to calibrate the nuclear recoil (NR) response, neutrons with an energy of a few MeV are typically used. However, for the very same reason radiogenic neutrons emitted by detector materials, pose a great risk for the scientific reach of these types of experiments. Neutrons can mimic WIMP signals by undergoing a single back-scatter NR inside the sensitive volume of the TPC, before escaping. To mitigate this danger XENONnT is equipped with a new neutron veto (NV), the world's first water Cherenkov NV. It tags these neutron signals by applying a backward veto using the signal of the delayed neutron capture on hydrogen which releases a 2.22 MeV  $\gamma$ -ray.

In this thesis the commissioning and calibration of XENONnT's new NV, and the first WIMP DM search results of the XENONnT experiment are discussed. As part of the presented work, the NR response model of the TPC and the NV neutron tagging efficiency, were calibrated using a novel calibration technique which was never used in a LXe TPC before. The calibration is based on gammas and neutrons which are coincidentally emitted by an Americium-Beryllium (AmBe) neutron source. This enables

a tagging of neutron signals inside the TPC by triggering on the coincident gamma in the NV or vice versa. In addition, as one of the lead developers for XENONnT's processing software, STRAXEN, one major aspect of the presented work was the development of the processing chain for the NV, which was used in the mentioned calibration as well as to study the general signal response of the new NV.

To enhance the neutron tagging efficiency even further, the water of the NV is going to be doped with Gadolinium-sulfate. As part of the presented work, different salt samples of the company Treibacher were analyzed regarding their optical transparency and acidity when in solution with pure water, and compared against a reference salt sample from a Japanese company which is used in EGADS and the Super-Kamiokande Gd-upgrade [Mar+20; Abe+22].

The first WIMP search results of XENONnT have been published as part of the presented work in [Apr+23b] and another publication about the commissioning and performance of the NV is currently being prepared for submission.

The thesis is structured as follows. Chapter 1 discusses the astronomical and cosmological evidence for dark matter from which the basic properties of WIMPs are inferred. The introduction of DM is followed by an overview of the XENONnT experiment in Chapter 2, which opens with a summary of the properties of LXe as detection medium and a discussion about the general working principle of LXe dual-phase TPCs. It further introduces the detection of neutrons with water Cherenkov detectors, before giving an overview of the XENONnT experiment. Chapter 3 continues with a review of the XENONnT processing framework, before the processing chain of the TPC and NV are discussed in more detail. Subsequently, Chapter 4 summarizes the study of different Gadolinium-sulfate samples of the company Treibacher for their suitability in XENONnT. The main aspect of the presented work is the calibration of the NR response of the TPC which is discussed in Chapter 6, as well as the calibration of neutron tagging efficiency of the NV explained in Chapter 7. Both are of vital importance for the WIMP search results of the experiment which are summarized in Chapter 8, before the presented work is concluded in Chapter 9.

Furthermore, the supplemental material documents in Section A.1 the XYZCam-Scan setup, a high-resolution optical inspection setup specifically designed for the inspection of wire surfaces over large areas. The setup was designed for the inspection of the XENONnT TPC wire electrodes and was realized together with the PRISMA detector lab beside the presented main topic of this thesis.

---

## DARK MATTER AND DIRECT DETECTION

The subsequent chapter is divided as follows. Section 1.1 discusses different astronomical and cosmological observations which give evidence for the existence of dark matter. Based on these observations Section 1.2 lists a few potential candidates for DM, among the WIMP which is the main focus of the presented work. Afterwards Section 1.3 discusses important astronomical inputs for the direct search for DM, which is discussed in the closing Section 1.4 of this chapter.

### 1.1 EVIDENCE FOR DARK MATTER

Numerous, gravitational observations provide evidence for the existence of a non-luminous mass component in the universe which is commonly referred to as DM. Evidence can be found on all time and mass scales of the universe, ranging from galaxies to the whole cosmos. On small scales, the radial velocity profile of stars and the interstellar medium (ISM)<sup>1</sup> in galaxies can be used to infer the mass distribution within galaxies, and even beyond the visible disk. The measurements of these so-called rotation curves are based on the Doppler shift, which spectral emission lines experience due to the relative motion of the object under study with respect to the observer. These observations were carried out over many decades starting with Jan Oort, who observed already in the 1930s that the velocities of stars in the local neighborhood of the solar system are too fast compared to expectations based on the mass profile of the visible mass in galaxies and Newtonian gravity [Sch19]. These initial observations were complemented in the 1980s by the systematic measurement of rotation curves of 20 nearby galaxies at optical wavelengths by Vera Rubin (et al.), and a little bit later confirmed by radio-observations of the 21 cm atomic hydrogen line which allowed an observation beyond the visible luminous disk of galaxies. These observations showed that the discrepancy between expected radial velocity and measured velocity is a regular feature of galaxies [Sch19; P19]. The rotation curve of the galaxy M33 is shown as an example in Figure 1.1. To resolve the discrepancy between observation and data, an additional non-luminous mass component can be added to galaxies, embedding galaxies in a DM sphere, called DM halo.

<sup>1</sup>Clouds of dust and ionized atoms and molecules, e.g. hydrogen.

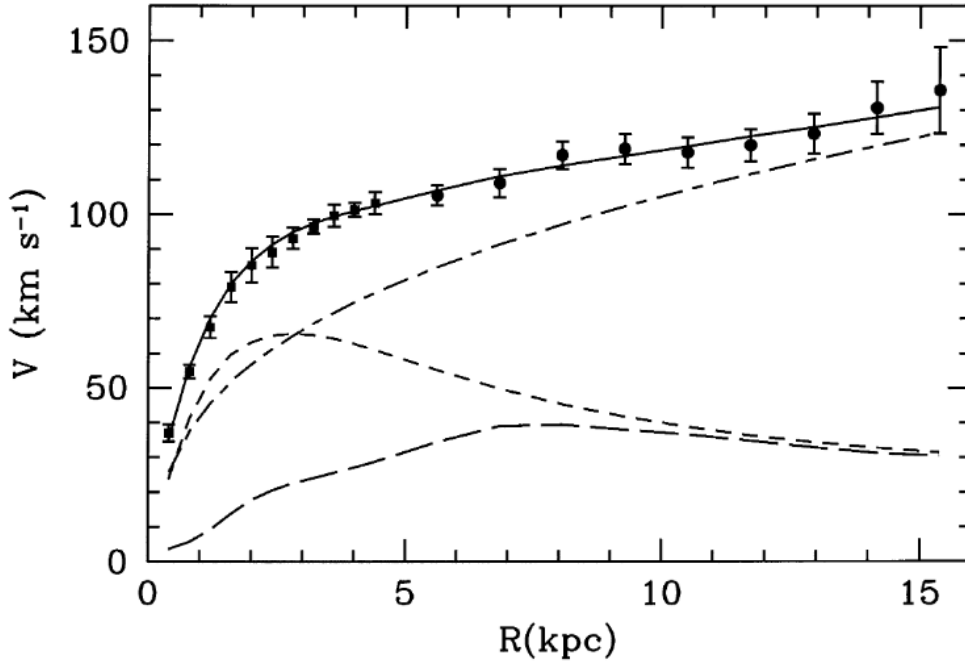


Figure 1.1. – Rotation curve of the galaxy M33. Filled dots show measurements based on the 21 cm line of hydrogen, while filled squares show optical measurements. The best fit is shown as black solid line. Contributions from the halo, stellar disk and gas, are shown as dot-dashed, short-dashed and long-dashed lines respectively [CS00].

At scales of full galaxies and galaxy clusters, gravitational lensing can be used to determine the mass profile of foreground objects. Gravitational lensing occurs when a massive object between the object, under study and the observer, deflects the emitted light due to its gravitational field. Generally, two types of lensing are distinguished: Weak lensing in which the mass profile of a region in space is studied based on the statistical analysis of the distortion of many light-emitting objects, and strong lensing in which a single object is distorted as an arc or ring [P19]. Based on the degree of deflection the mass of the object can be determined and compared against its luminous mass, which is determined based on its spectral energy distribution [P19; UR15]. The masses found by gravitational lensing are usually found to be larger than the determined luminous masses [UR15]. Gravitational lensing paired with the systematic study of galaxy cluster collisions can be used to determine additional properties of DM. In these collisions, the ISM is stripped away from the galaxy clusters and heated due to the ram pressure created during the shock of the collision. Thus, the ISM can be observed in the aftermath of the collision via x-rays, and its spatial distribution be compared against the mass distributions of the luminous matter as well as the total mass distribution observed through gravitational lensing. These studies suggest the existence of a DM component which closely follows the mass distribution of the luminous matter in terms of stars in the galaxies [Har+15; UR15].

On cosmological scales, the cosmic microwave background (CMB) give not only evidence for the existence of DM, but also provides a tool to determine the contribu-

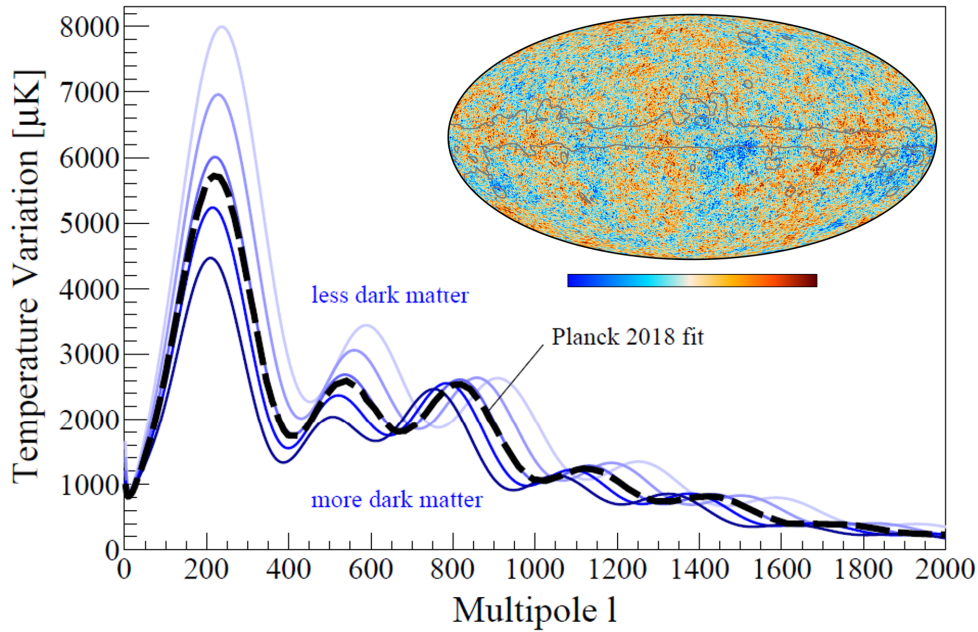


Figure 1.2. – Anisotropies of the CMB and the best fit of the Planck 2018 results. The inset figure shows the temperature anisotropies of the CMB. The color scale encodes variations between  $\pm 300 \mu\text{K}$ . The figure was taken from [col18]. The main plot was taken from [Sch19]. It shows the multipole power spectrum of the CMB anisotropies. The best fit of the  $\Lambda\text{CDM}$  model of 2018 Planck results is represented by the black dashed line. Blue lines show the  $\Lambda\text{CDM}$  with a higher or lower dark matter density. Especially, the third acoustic peak is sensitive to the dark matter density.

tion of DM to the total energy budget of the universe. The CMB can be described as an almost perfect black body spectrum. It was created at the time of electron-proton recombination  $\sim 380000$  years after the big bang when photons decoupled from the electron-proton plasma at a temperature of  $\sim 3000$  K. Today, the CMB has a temperature of  $2.725 \text{ K}^2$  with anisotropies in the order of  $10^{-5}$  which are shown in the inlet of Figure 1.2. The anisotropies can be understood as density perturbations in the baryon-photon fluid, caused by the gravitational pull of DM which decoupled at higher temperatures from the plasma. The temperature anisotropies can be expanded into a multipole power spectrum to describe the temperature variations as a function of angular size. The resulting so-called "acoustic peaks" in Figure 1.2 can be very well described by the six-parameter cosmological standard model called  $\Lambda\text{CDM}$ . The model also includes parameters to describe the entire energy budget of the universe with  $\Omega_B = 0.0486(10)$ ,  $\Omega_{DM} = 0.2589(57)$ , and  $\Omega_\Lambda = 0.6911(62)$ , which represents the energy density of baryonic matter, DM, as well as the energy density for "dark energy" [Sch19; UR15]. Thus, dark matter is about five times more abundant than normal baryonic matter.

In addition, also the big bang nucleosynthesis (BBN) constraints potential DM models on cosmological scales. The BBN describes the formation of light nuclei a few hundred seconds after the Big Bang [UB04]. Their abundance depends on the baryon

<sup>2</sup>Which corresponds to photons with a wavelength in the microwave region and thus the name.

density, which is found to be close to the density found in the CMB [UR15].

Also, the structure of the universe would be different without DM. Comparisons between N-body simulations of non-relativistic DM particles, so-called "cold" DM, and galaxy surveys show a very good agreement over a large range of different scales [KVA12]. These simulations usually simulate the evolution of structure formation on cosmological scales using constraints on density fluctuations observed in the CMB as starting conditions, or smaller substructures based on simulations of individual halos, by simulating several billion DM particles [KVA12]. In addition, these simulation provide valuable information about astrophysical inputs for the direct search of DM [Sch19; UR15] which is discussed in Section 1.3.

### 1.2 DARK MATTER CANDIDATES AND PROPERTIES

Based on the astronomical and cosmological observations discussed in the previous section some key properties of dark matter can be inferred. First of all the potential candidate must be dark, which means DM should be electrically neutral if made of particles, and it must be either long-lived or stable as evidence for DM is found on all-time scales of the universe [RST18]. Based on the observed cluster collisions and structure formation, it can be inferred that the self-interaction cross section must be small. Further, also their interaction strength with standard model particles must be either through the weak force or at least on the weak scale or smaller. In addition, candidates must ensure that they do not violate the observed baryon density inferred from the BBN and CMB, which also means that DM candidates should be non-baryonic.

An exception are **primordial black holes (PBHs)** which would have formed from density fluctuations in the radiation dominate area before the BBN [RST18; Sch19; CK22]. However, based on the number of black hole merger events in the mass range between  $20 - 30 M_{\odot}$ , PBHs in this mass range could only contribute by 1 % to the total DM density. Further, a 100 % DM contribution from PBHs is ruled out for all mass ranges based on other astronomical observations [Sch19]. New instruments like LISA will help to probe PBHs in smaller mass ranges [Bar+19].

One particle candidate for DM which would already fulfill all these requirements are standard model **neutrinos**. Due to their small mass neutrinos travel at relativistic speeds, and thus would represent "hot dark matter" [RST18; Sch19]. However, based on the observed structure formation in the universe, DM needs to be "cold", or "warm", as otherwise the density fluctuations observed in the CMB, which serve as seed for structure formation, would be washed out [RST18]. In addition, their relic abundance is too small to account for the full dark matter density in the universe [BHS05].

A suitable candidate for warm DM would be **sterile neutrinos** which are postulated in order to explain the small neutrino mass. Their masses would be in the range of a few keV to tens of keV, mostly constrained by astronomical x-ray surveys. Warm dark matter could also help to mitigate the missing satellites problem<sup>3</sup> in structure

---

<sup>3</sup>This problem refers to the fact that simulations of cold DM predicts a larger number of subhalos, and therefore a larger number of dwarf satellite galaxies when being observed. However, this problem has been reduced due to the discovery of more ultra-faint dwarf satellites in recent years. [KVA12]



formation [KVA12], however would at the same time also introduce new issues in form of the missing formation of small DM halos at high redshifts where presumably the first stars formed [RST18; UR15].

A model-independent generic class of particle that satisfies all these observations despite the missing satellite problem is the **WIMP**. They are predicted by many well-motivated beyond the Standard Model theories, e.g. as the lightest particle in little Higgs models, of Kalzua-Klein states, or in supersymmetric theories [Sch19; UR15; RST18]. Depending on the exact model, WIMPs have a typical mass in a range between  $1 \text{ GeV}/c^2$  and  $10^5 \text{ GeV}/c^2$ , and interaction cross sections between  $10^{-41} \text{ cm}^2$  and  $10^{-51} \text{ cm}^2$  [Sch19]. The most common production mechanism for WIMPs is thermal freeze-out in the early universe, when the universe expands and the temperature of the plasma falls below the respective WIMP mass [RST18]. The relic DM density is reached when the annihilation rate of the remaining WIMPs falls below the Hubble expansion rate [UR15]. Most of the proposed WIMP candidates introduce a new quantum number which prevents them from decaying into lighter standard model particles and makes them therefore stable [Sch19].

Another type of well-motivated DM candidate are **axions or axion like particles (ALPs)**. Axions were originally postulated as a solution to the "strong CP problem", with masses typically below the keV range due to cosmological and astrophysical constraints. These masses are too small to be observed in dual-phase TPC, however, ALPs can have higher masses on the keV-scale which can be observed in experiments like XENONnT via the axio-electrical effect. The observed signal would be a mono-energetic peak at the rest mass of the ALPs. Such searches were also carried out in XENON1T [Apr+20a], and XENONnT [Apr+22c], but are not further discussed in the presented work.

### 1.3 ASTROPHYSICAL INPUTS FOR THE DIRECT DARK MATTER SEARCH

Besides giving hints for possible dark matter candidates, astrophysical observations give also inputs about the local DM density and velocity distribution, which are important key parameters for the direct search for DM. A variety of different models exist to describe the density profile of DM halos [P19]. To make the results of direct dark matter searches comparable, most of the experiments use the so-called standard halo model as recommended in [Bax+21a]. In this model, the dark matter distribution within a galaxy is described as an isothermal sphere of collisionless particles, with an initial isotropic "Maxwellian" velocity distribution

$$f(\vec{v}_\chi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{|\vec{v}_\chi|^2}{2\sigma^2}\right\}, \quad (1.1)$$

where  $\vec{v}_\chi$  represents the DM velocity and  $\sigma$  the dispersion velocity which is related to the circular velocity through  $\sigma = \sqrt{3/2}v_c$  with  $v_c$  being usually assumed to be  $220 \text{ km/s}$  [UR15; Coo22]. At too high speeds dark matter particles would not be bound to the galaxy anymore and thus the velocity distribution needs to be truncated at the local escape velocity, which is usually assumed to be  $v_{esc} = 544 \text{ km/s}$ . In ad-



dition to the velocity distribution of the DM particles around the solar orbit, also the earth's peculiar motion with respect to the sun's motion through the DM halo must be considered. This leads to an additional modulation of the WIMP speed, and thus a modulation in the measured WIMP rate [Sch19; UR15]. Once a potential DM signal is found, this modulation can help to solidify the claim. However, one must also keep in mind that certain backgrounds also show seasonal variations, like the atmospheric neutrino flux [Sch19].

The density profile of such an isothermal DM halo can be estimated through the equation of state balancing the inward gravitational pull of the DM halo with the pressure of the collisionless gas. It is related to the velocity dispersion and can be written as

$$\rho_\chi(r) = \frac{\sigma}{2\pi Gr^2} \quad (1.2)$$

where  $G$  is the gravitational constant [Sch19]. For the direct search of DM the exact shape of the profile is less important than compared to indirect searches. More important is the local DM density in the vicinity of the solar orbit. Usually, a canonical value of  $\rho_\chi = 0.3 \text{ GeV}/c^2/\text{cm}^3$ , is recommended [Bax+21a], while recent observation prefer a value closer to  $\rho_\chi = 0.46_{-0.09}^{+0.07} \text{ GeV}/c^2/\text{cm}^3$  [Sch19; P19]. In addition, DM N-body simulation show that the local DM density should be rather smooth and that it is fairly safe to assume a constant average local DM density [KVA12]. The precise value of  $\rho_\chi$  also plays only a minor role until a DM signal is observed. It is more important that direct detection experiments agree on a canonical value, in order to make their experimental results comparable.

#### 1.4 DIRECT DARK MATTER DETECTION

Assuming, that DM interacts with standard model particles, it can be observed in one of the three following, complementary ways.

It could be observed **indirectly** through the annihilation of DM particles into standard model particles, and their secondaries like  $e^-e^+$ ,  $p^-p^+$ ,  $\gamma$ -rays or neutrinos. Also, a decay of DM particle could be observed indirectly. Most of the indirect searches focus on objects with large masses which can accumulate DM and therefore enhance the chance of a potential signal. Typical candidates for these objects are stars, dwarf galaxies, or the galactic center. Depending on the type of annihilation signal, DM could be indirectly observed either by  $\gamma$ -ray telescopes like Fermi-LAT, Cherenkov telescopes like HESS or MAGIC, or neutrino observatories like Ice Cube or Super Kamiokande [UR15].

Another viable option is the **production of DM at colliders** through the collision of standard model particles. In this case, DM signals would appear as missing transverse momentum and energies, in detectors like ATLAS or CMS [UR15].

The third option is the **direct detection** of DM through collisions with standard model particles, where the discussion in the following focuses on direct WIMP detection. Given the velocity distribution of WIMPs with a velocity of  $\sim 220 \text{ km/s}$  at the radius of the solar orbit and their masses in the range of a few GeV to TeV, it can be assumed that WIMPs are generally traveling at non-relativistic speeds with typical

kinetic energies of a few keV. Since this is well below the typical nuclear binding energies of heavier nuclei, it can be anticipated that the collision will be elastic, fairly independent of the choice of underlying WIMP candidate [Coo22]. Hence, the recoil energy transferred from an elastic collision between a DM particle with mass  $m_\chi$  to a target with the mass  $m_T$  can be derived in the center of mass frame to be

$$E_R = \frac{\mu^2 v_\chi^2}{m_T} (1 - \cos(\theta)) \quad (1.3)$$

$$\Rightarrow E_R^{\max} = 2 \cdot \frac{\mu v_\chi^2}{m_T} \quad (1.4)$$

with the reduced mass  $\mu = \frac{m_\chi m_T}{m_\chi + m_T}$ , and scattering angle  $\theta$  [Coo22]. As an example, the maximum recoil energy of a WIMP with a mass of  $100 \text{ GeV}/c^2$  and a velocity of  $V_\chi = 0.75 \cdot 10^{-3} c$  leads to a recoil energy on a xenon nucleus of  $\sim 28 \text{ keV}$ .

The expected rate of interactions as unit energy in any direct detection experiment is proportional to the number density of target atoms  $N_T$ , the flux of DM particles through the detector  $\phi_\chi$ , as well as the cross section of the interaction  $\sigma$ , and can be expressed as

$$\frac{dR}{dE_R} = \frac{N_T}{M_T} \cdot \frac{\rho_\chi}{m_\chi} \underbrace{\int_{v_{\min}}^{v_{\text{esc}}} v_\chi f(\vec{v}_\chi)}_{\phi_\chi} \cdot \overbrace{\frac{d\sigma_{\chi,N}}{dE_R}}^{\sigma}, \quad (1.5)$$

where  $M_T$  is the total target mass of the detector,  $m_N$  the mass of a single target nucleus,  $\frac{d\sigma_{\chi,N}}{dE_R}$  the differential DM-nucleus cross section, and  $v_{\min}$  the minimal velocity for which a DM particle leaves a detectable signal in the detector depending on Equation (1.3) [Coo22]. While the first and second terms in Equation (1.5) depend on detector and astrophysical properties of the DM model, the last term depends on inputs from particle physics. Since it is unknown in which way WIMPs may interact with the target nucleus it can be generally assumed that the interaction cross section is the sum of spin-independent (SI) and spin-dependent (SD) interactions. Thus, one can write

$$\frac{d\sigma_{\chi,N}}{dE_R} = \underbrace{\frac{m_N}{2\mu v_\chi^2}}_a \cdot \sigma = \frac{m_N}{2\mu v_\chi^2} [\sigma_0^{SI} F_{SI}^2 + \sigma_0^{SD} F_{SD}^2] \quad (1.6)$$

where the prefactor  $a$  comes from the integration over all possible recoils energies based on Equation (1.3),  $\sigma_0^i$  represent the spin-dependent and independent cross section at zero momentum transfer, and  $F_i$  form factors which describe the dependence on the momentum transfer and nuclear structure [Coo22; UR15].

For the SI case the cross section can be written as

$$\sigma_0^{SI} = \sigma_{\chi,n} \frac{\mu^2}{\mu_n^2} \cdot [Z \cdot f_p + (A - Z) \cdot f_n]^2 \quad (1.7)$$

where  $f_i$  is the contribution from neutron and protons couplings to the total cross

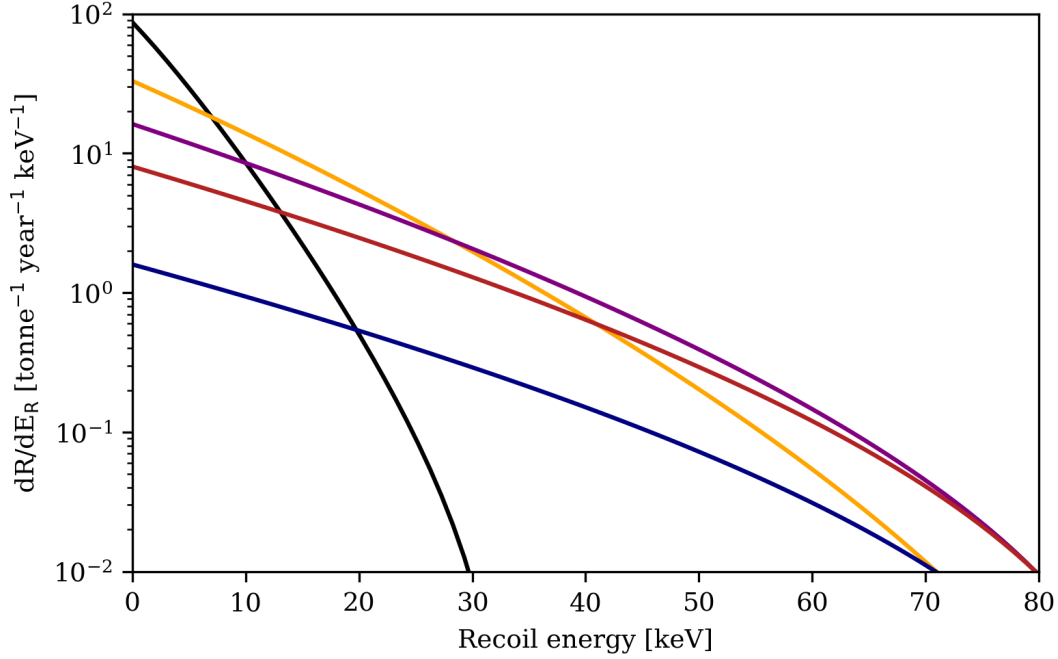


Figure 1.3. – Differential WIMP recoil spectra for different WIMP masses. The plot shows the expected differential WIMP rate as a function of recoil energy in LXe for spin-independent WIMP-nucleon couplings. The rates were calculated using the *wimprates* package [Aal+23] assuming a reference cross section of  $\sigma_{\chi,n} = 10^{-45} \text{ cm}^2$ . The plot shows the rates for the following WIMP masses 20 GeV/c<sup>2</sup>, 50 GeV/c<sup>2</sup>, 100 GeV/c<sup>2</sup>, 200 GeV/c<sup>2</sup> and 1000 GeV/c<sup>2</sup>, in black, orange, purple, red and blue respectively.

section,  $\sigma_{\chi,n}$  the cross section of a single nucleon, and  $\mu_n$  the WIMP-nucleon reduced mass [UR15]. Since the de Broglie wavelength of an elastic scattering WIMP on a xenon nucleus is typically larger than the size of the nucleus<sup>4</sup>, it can be assumed that the interaction cannot resolve the difference between protons and neutrons at the parton level. Thus, it can be assumed that  $f_p = f_n$  which leads to a coherent scattering  $\sigma_0^{SI} \propto A^2$ . This makes especially heavier target materials appealing for the search of SI interactions. At larger momentum transfers the structure of the target nucleus can be resolved and the form factor for the SI interaction  $F_{SI}$  starts to play a role. Usually, the Helm form factor is used [UR15; Coo22] as it is also the case for the XENONnT SR0 WIMP search results presented in this work. The recoil spectra of WIMPs on xenon, with masses between 50 GeV/c<sup>2</sup> to 1000 GeV/c<sup>2</sup>, are shown in Figure 1.3. The spectra are featureless and follow almost an exponential distribution. The highest recoil energies are reached for WIMPs with masses similar to the mass of the xenon nuclei.

In the case of SD interactions, the cross section directly depends on the nuclear angular momentum  $j$  and the nucleon spins  $\langle S_i \rangle$ , where  $i$  represents neutrons or protons. Thus, in contrast to the SI case, the SD case shows much more variation among dif-

<sup>4</sup>Using the same example as before with a WIMP mass of 100 GeV/c<sup>2</sup> and a velocity of  $V_\chi = 0.75 \cdot 10^{-3} c$ , the de Broglie wavelength is  $\lambda \sim 16.5 \text{ pm}$ , while the typical size of a xenon nucleus can be approximated as  $r \approx 1.25 \text{ fm} \cdot A^{1/3} \approx 6.35 \text{ fm}$  [Coo22]

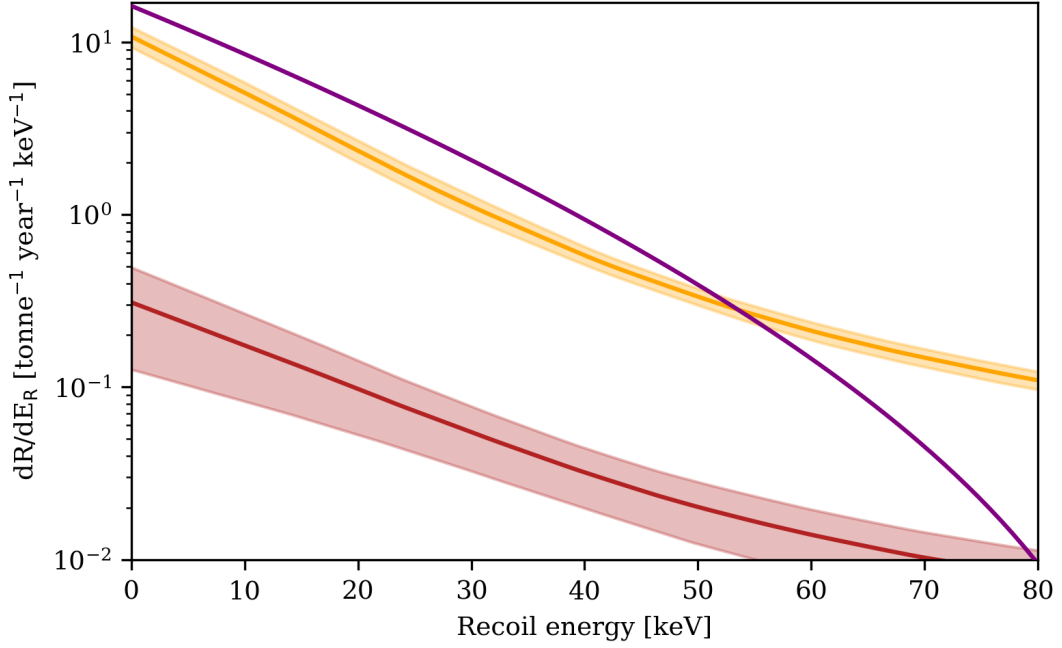


Figure 1.4. – SD differential rate for WIMP-nucleus interaction of a  $100 \text{ GeV}/c^2$  WIMP. The purple curve shows for reference the scatter rate for a SI interaction of a  $100 \text{ GeV}/c^2$  WIMP with a WIMP-nucleon cross section of  $\sigma_{\chi,n} = 10^{-45} \text{ cm}^2$ , as in Figure 1.3. The orange and red curve show the expected rate for SD, WIMP-neutron and WIMP-proton couplings respectively, but with a larger reference cross section of  $\sigma_{\chi,n} = 10^{-40} \text{ cm}^2$ .

ferent isotopes, even if they have an almost identical atomic mass. Generally, the SD cross section for zero momentum transfer can be written as

$$\sigma_0^{SD} = \frac{32G_F^2\mu^2}{\pi} \frac{J+1}{J} [a_p \langle S_p \rangle + a_n \langle S_n \rangle]^2 \quad (1.8)$$

where  $G_F$  is the Fermi coupling constant and  $a_i$  the effective proton and neutron couplings. For xenon only the isotopes  $^{129}\text{Xe}$  and  $^{131}\text{Xe}$  contribute to SD interactions. Their expectation values for the nucleon spins are  $\langle S_p^{129} \rangle = 0.028$ ,  $\langle S_n^{129} \rangle = 0.359$ , and  $\langle S_p^{131} \rangle = -0.009$ ,  $\langle S_n^{131} \rangle = -0.227$  respectively [Coo22; UR15]. Although xenon does not have an unpaired proton it is thus still sensitive to SD proton couplings. However, the expected rates compared to SI interactions are much lower as the  $\propto A^2$  enhancement is missing. This is also illustrated in Figure 1.4 which compares the expected rate of WIMP-nucleus interactions for SI and SD recoils of a  $100 \text{ GeV}/c^2$  WIMP. For SD interaction the form factor is based on the spin-structure function derived from nuclear shell models [UR15; Sch19]. In the traditional analysis for direct WIMP searches it is always assumed that WIMPs either couple only to neutrons or protons. A finer distinction is only needed once a signal is found, and a more precise comparison between complementary measurements with other target materials becomes more interesting.

Depending on the chosen target material and detector technology, signals of DM particles are either detected through photons, charge carriers, phonons (heat), or a

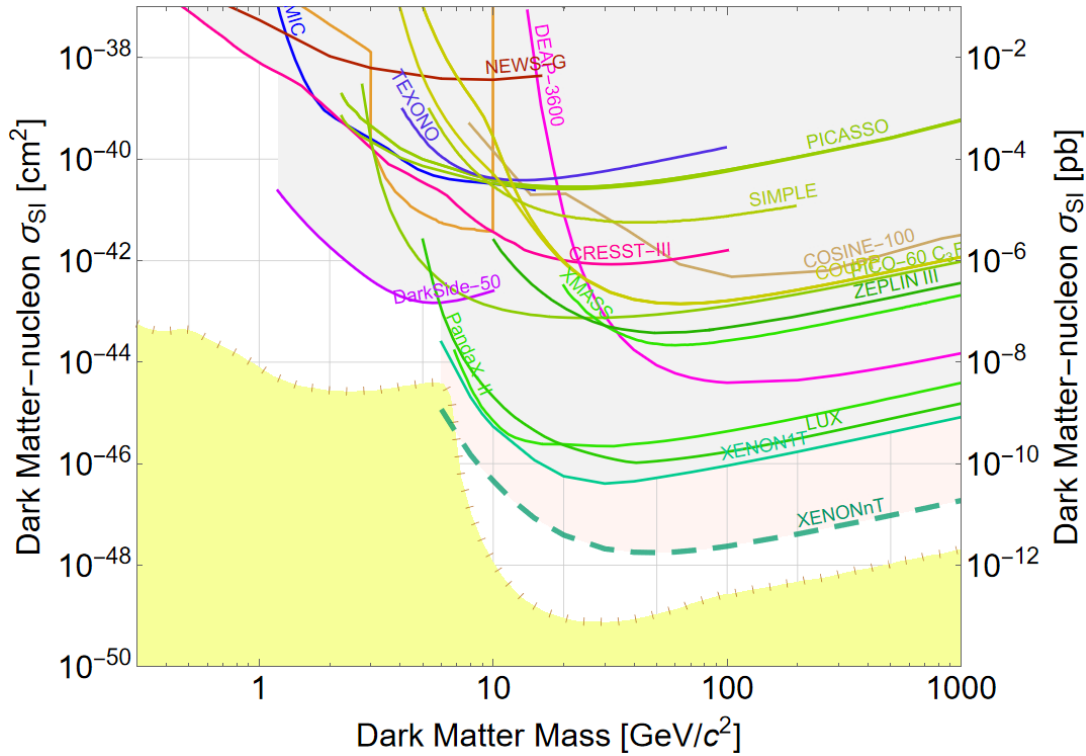


Figure 1.5. – Current best limits for the spin-independent WIMP-nucleon cross section of direct detection experiments as of 07.03.2023, excluding the most recent limits by XENONnT and LZ, which are discussed in the last Chapter of the presented work. Solid lines indicate the exclusion limits for the depicted experiments. The dashed line indicates the projected sensitivity of XENONnT with 20 tonne-year exposure. The yellow shaded region indicates where coherent neutrino-nucleus scattering is expected to be observed. The plot was generated with [Saa+23] v5.18.

mixture of two of these. While liquid noble gas detectors in form of dual-phase TPCs, which detect simultaneously charge and light signals, lead the search for DM in the mass range between  $1 \text{ GeV}/c^2$  and several  $\text{TeV}/c^2$ , cryogenic bolometers are more sensitive to low mass WIMPs due to their ultra-low threshold of a few tens of eV [Abd+19]. However, due to their overall smaller target mass, they cannot compete with liquid noble gas detectors at higher DM masses.

The current status as of 07.03.2023 for the direct search for spin-independent WIMP-nucleon interactions<sup>5</sup> is shown in Figure 1.5. The figure shows that the best exclusion limits at small DM masses below  $\sim 1 \text{ GeV}/c^2$  is given by the CRESST-III experiment which measures DM signals via photons and heat. The mid-range between  $\sim 1 \text{ GeV}/c^2$  and  $\sim 7 \text{ GeV}/c^2$  is dominated by DarkSide-50<sup>6</sup> [Agn+23] a liquid argon dual-phase TPC, and the range starting from  $\sim 7 \text{ GeV}/c^2$  by XENON1T [Apr+18] and LUX [Ake+17b] using LXe dual-phase TPCs, which are discussed more thoroughly

<sup>5</sup>Excluding the recent results from XENONnT and LZ which will be discussed in Chapter 8.

<sup>6</sup>Please, note that these results are based on the theorized Migdal effect which is an inelastic interaction leading to larger energy deposits and thus allows to measure smaller signals which would be below detector threshold otherwise [Bel+22].

in the following Chapter. The yellow shaded region indicates where coherent elastic neutrino-nucleus scattering (CE $\nu$ NS) on xenon nuclei is expected. These signals will have a similar signature as WIMPs, as the neutrinos will scatter coherently on the xenon nucleus. At small DM masses, below  $\sim 10 \text{ GeV}/c^2$ , these signals will be dominated by solar neutrinos from  $^8\text{B}$ , while at higher DM masses signals from atmospheric neutrinos and neutrinos from the diffuse super nova neutrino background dominate [Aki+18]. It is expected that XENONnT [Apr+20b] and LZ will observe these signals for the first time.



---

## THE XENONNT EXPERIMENT

The quest for WIMP dark matter in a mass range of a few  $\text{GeV}/c^2$  to  $\text{TeV}/c^2$  is dominated by a number of different liquid noble time projection chambers [Apr+18; Men+21; Aal+22]. Among these experiments are the XENON1T [Apr+17c] and XENONnT [Apr+22c] experiment as part of the XENON dark matter project. XENONnT was designed as a fast upgrade of its predecessor XENON1T. The experiment features a liquid xenon TPC, primarily designed for the search for WIMPs [Apr+18]. It is also used in variety of other rare event searches like axions, ALPs, solar neutrinos and their properties [Apr+20a], as well as neutrinoless double beta decay and electron capture [AX19]. Further, as part of this upgrade, a new water Cherenkov neutron-veto was added to the experiment which, tags neutrons via the hydrogen neutron capture releasing a single 2.22 MeV  $\gamma$ -ray.

In preparation for the analysis discussed later in this work, the following chapter discusses concepts of liquid xenon TPCs as well as water Cherenkov detectors, before giving details about the XENONnT experiment itself. The chapter is divided as follows: Section 2.1 discusses the general properties of xenon and its behavior as a detection medium for different particles. Based on these properties a discussion about the working principle of TPCs and the most important concepts of particle discrimination, as well as signal and energy reconstruction follow in Section 2.2. Afterwards Section 2.3 introduces the basic concepts of water Cherenkov detectors and particle interactions in water, before introducing the XENONnT experiment in Section 2.4.

### 2.1 LIQUID XENON AS DETECTION MEDIUM

#### 2.1.1 *Liquid xenon properties*

For more than half a century it is already known that noble gases can be used as a detector medium for particle detection [AD10]. In particular argon and xenon have been utilized in many different experiments [Apr+17c; Ake+20; Aal+18] over the past twenty years. As the name suggests the XENON dark matter project uses xenon as target material for its detectors [Apr+12; Apr+17c; Apr+22c].



Xenon has several properties which make it an attractive target in particular for the search for WIMP dark matter. First of all, among all stable noble gases, xenon has the highest atomic mass number which is especially favorable in the search for spin-independent interactions as the cross section directly scales with  $A^2$  (see Equation (1.7)). Secondly, its almost 50-to-50 abundance of even-even and even-odd numbered isotopes allows the effective probing of spin-dependent interactions due to the unpaired neutron, too (see Equation (1.8)). Further, its relatively high mass of  $\sim 122.3 \text{ GeV}/c^2$  maximizes the energy transfer in elastic collisions with WIMPs as their masses are typically predicted to be in a similar range of a few  $\text{GeV}/c^2$  up to a few  $\text{TeV}/c^2$  [Sch19]. However, for the very same reason the search for lighter dark matter particles is more challenging.

All naturally occurring xenon isotopes are listed in Table 2.1. Almost all isotopes are stable except for  $^{134}\text{Xe}$  and  $^{136}\text{Xe}$ . The latter can be considered an additional asset, as both isotopes decay via second-order weak processes and are therefore interesting to study on their own.  $^{134}\text{Xe}$  decays via two-neutrino double electron capture and has the longest half-life ever measured for any isotope. Its direct detection provides important input for nuclear structure models and is a first step for the search of the neutrinoless double electron capture [AX19].  $^{136}\text{Xe}$ , on the other hand, decays via a two-neutrino double beta decay which allows to search for the equivalent neutrinoless process as done in [Alb+14; Apr+22e]. Given the extremely long half-life of both isotopes, which is longer than the age of the universe, highly sensitive low-background detectors are required to study these decays, such as those utilized in the search for dark matter. Both isotopes play only a minor role in the total budget of detector intrinsic backgrounds for the search of WIMP dark matter which is discussed in more detail in Section 2.4.

Another asset of xenon is the fact that it can be liquefied at moderate cryogenic temperatures of about  $-100^\circ\text{C}$  with a vapor pressure of 1.58 bar [Lem+23]. This can be easily achieved through the evaporation of liquid nitrogen which makes liquefaction simpler compared to other noble gases like argon. The density of liquid xenon at these conditions is about  $2.89 \text{ g}/\text{cm}^3$  which is about 200-times higher than in gaseous xenon at the same temperature, but lower pressure [Lem+23]. The high density of liquid

Table 2.1. – All natural xenon isotopes, their abundance and half-life if applicable. The decay times are taken from [Apr+22e; Alb+14], and the abundances from [Cou+15].

Isotope	Abundance	Half-life
$^{124}\text{Xe}$	0.095 %	$1.1 \pm 0.2_{\text{stat.}} \pm 0.1_{\text{sys.}} \cdot 10^{22} \text{ y}$
$^{126}\text{Xe}$	0.089 %	-
$^{128}\text{Xe}$	1.910 %	-
$^{129}\text{Xe}$	26.401 %	-
$^{130}\text{Xe}$	4.071 %	-
$^{131}\text{Xe}$	21.232 %	-
$^{132}\text{Xe}$	26.909 %	-
$^{134}\text{Xe}$	10.436 %	-
$^{136}\text{Xe}$	8.857 %	$2.165 \cdot 10^{21} \text{ y}$

xenon allows for building easily scaleable, compact-sized detectors with a high number of target nuclei, which are directly proportional to the expected signal rate (see Equation (1.5)). In addition, depending on the type of liquid xenon detector, the high density allows to shield off background signals originating from radioactive decays of different detector materials. The downside of xenon compared to other noble gases is its relatively high price tag as it is also used in a variety of other applications such as car industry, medicine and space flight [22]. For a dual-phase TPC this technique is explained more thoroughly in Section 2.4.

### 2.1.2 Signal generation and reconstruction

Particles interacting with liquid xenon can transfer their energy via different mechanisms. For the calibration of the detector and studies presented in this work, the two most important particle types are  $\gamma$ -rays and neutrons. Gammas interact primarily with the electron shell of xenon atoms and transfer their energy either via photo-absorption, Compton scattering, or pair production, depending on the energy of the incident photon. Figure 2.1 shows the mass attenuation of xenon  $\mu_{\text{ph,xe}}$  as a function of energy for all three components. The Compton effect dominates for photon energies of a few MeV which is the most relevant energy range when calibrating TPC and neutron-veto with AmBe as discussed in Section 6.1. At lower energies, photon absorption through the photo effect dominates with the characteristic steps in the mass attenuation due to the electron-shell structure of xenon. The first step around 5 keV is due to the L-shell while the peak around 35 keV corresponds to the K-shell. The mean free path of photons  $\lambda_{\text{ph}}$  in liquid xenon with density  $\rho_{\text{LXe}}$  can be computed via

$$\lambda_{\text{ph}} = \frac{1}{\mu_{\text{ph,xe}} \cdot \rho_{\text{LXe}}} . \quad (2.1)$$

It is about 10 cm for  $\gamma$ -rays with an energy of a few MeV and decreases rapidly for lower photon energies. Electrons produced by  $\gamma$ -rays, or  $\beta$ -particles are dominantly interacting with the shell electrons of xenon and lose their energy quickly by ionization and excitation of nearby xenon atoms. This leads to only short tracks in LXe. For example a 500 keV electron has a range of about 0.1 cm in liquid xenon, estimated with the continuous-slowing-down approximation (CSDA) [Ber+17]. All energy deposits by interaction with xenon shell electrons are usually summarized as electronic recoil (ER), due to the recoiling electron. This also included interactions of particles like neutrinos and hypothetical particles like axions or ALPs [Apr+20a; Apr+22c].

In contrast, neutrons deposit their energy similarly as WIMPs via elastic or inelastic NR, in which the recoiling nucleus transfers its energy to the surrounding xenon. In addition, neutrons can be captured by xenon, most dominantly on the isotopes  $^{129}\text{Xe}$  and  $^{131}\text{Xe}$  [Ama+22]. Figure 2.2 shows the interaction cross sections for neutrons with xenon as a function of energy, averaged in accordance with the abundance of the individual xenon isotopes. The maximum recoil energy  $E_{\text{NR,max}}$  was evaluated using

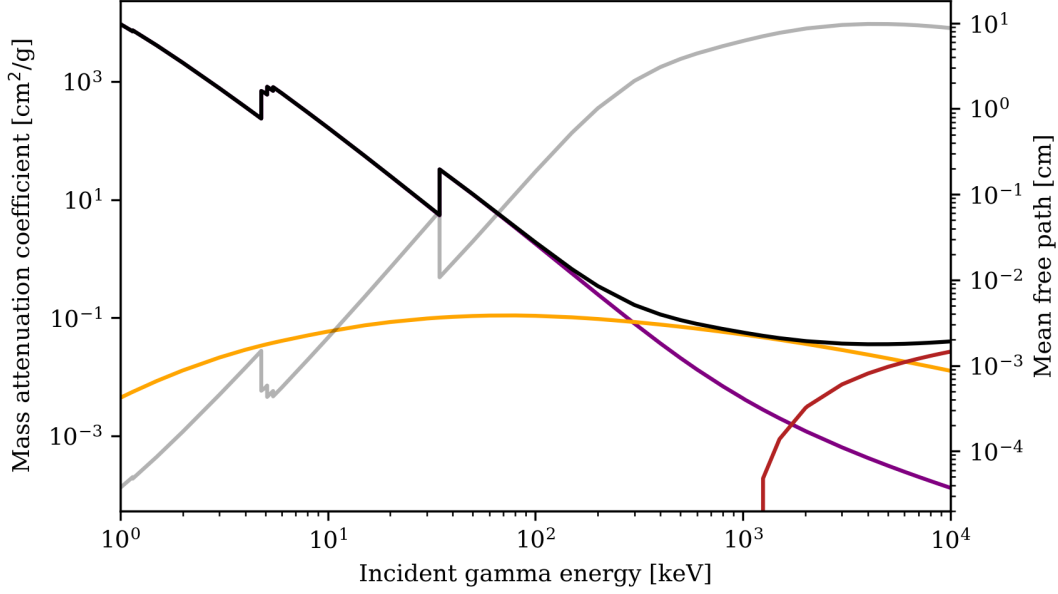


Figure 2.1. – Mass attenuation coefficient of  $\gamma$ -rays in liquid xenon as a function of photon-energy. The black curve shows the total attenuation coefficient, while purple, orange and red are showing the absorption via photo-effect, Compton scattering and pair-production in the field of the nucleus respectively. The gray shaded curve shows the mean free path of a  $\gamma$ -ray of the given energy in liquid xenon. The mass attenuation coefficients are taken from [Ber+10].

Equation (1.3) which simplifies for neutrons to

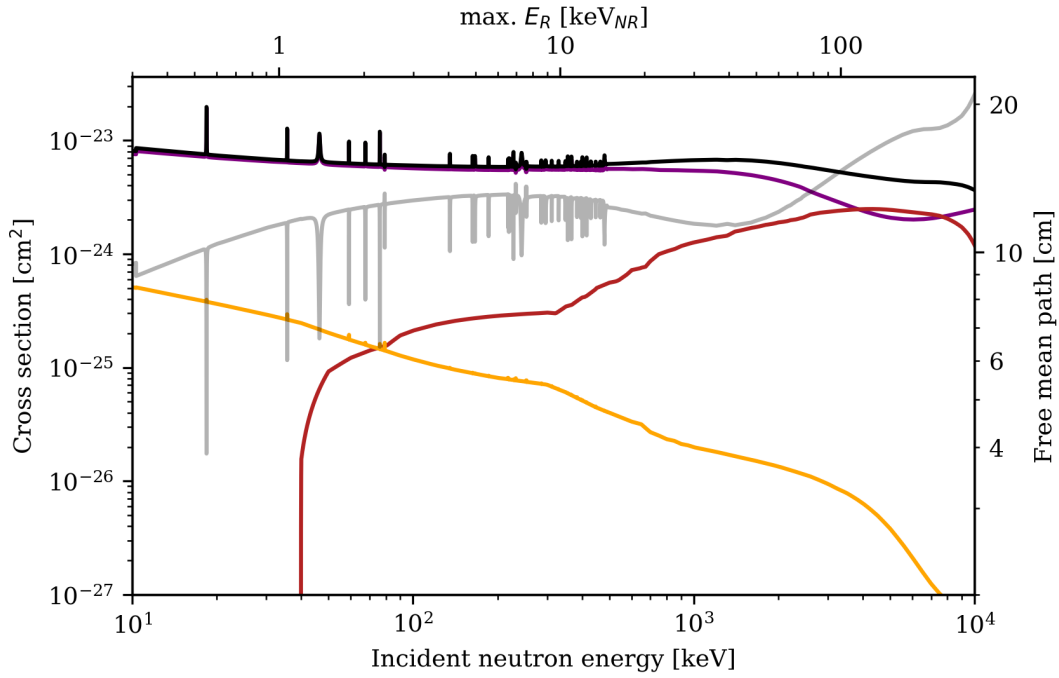
$$E_{\text{NR,max}} \approx 0.03 \cdot E_n \quad (2.2)$$

where  $E_n$  is the kinetic energy of the neutron. Neutrons emitted by AmBe as discussed in Section 6.1 have energies in the MeV range and therefore produce nuclear recoils in the order of a few 10 keV to 100 keV which is of a similar order of magnitude as the recoils of WIMPs with masses of  $\sim 50 \text{ GeV}/c^2$ . Neutrons of these energies have a slightly larger mean free path  $\lambda_n$  in liquid xenon as MeV-scale  $\gamma$ -rays. The mean free path is also shown in Figure 2.2 and can be estimated via

$$\lambda_n = \frac{1}{N_{\text{Xe}} \cdot \sigma_{n,\text{tot}}} \quad (2.3)$$

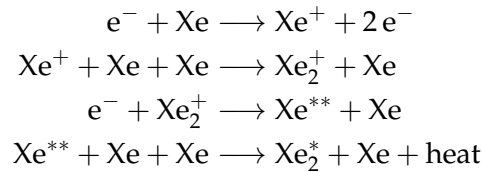
where  $\sigma_{n,\text{tot}}$  is the total neutron interaction cross section and  $N_{\text{Xe}}$  the atom number density of liquid xenon.

Independent of the recoil type (ER or NR), any energy deposit in gaseous or liquid xenon eventually leads to the formation of detectable quanta in the form of scintillation photons and free electrons which are proportional to the amount of deposited energy. In addition, unlike ERs, NRs "lose" a certain fraction of the deposited energy into the non-detectable atomic motion of the recoiling Xe-nucleus or, in other words, into heat. Free electrons are produced via ionization, while scintillation photons can

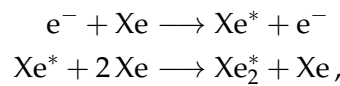


**Figure 2.2.** – Neutron interaction cross section for xenon as a function of energy. The purple, orange and red curve show the contribution from elastic and inelastic nuclear recoils as well as the radiative neutron capture ( $n, \gamma$ ). The black curve shows the total interaction cross section. The gray shaded line indicates the mean free path of neutrons in liquid xenon as a function energy. The x-scale at the top of the figure shows the maximal recoil energy for the corresponding neutron-energy. The data for the cross sections were taken from [Plo+20].

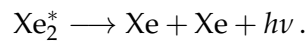
be either produced by ionization ( $\text{Xe}^+$ ), too



or excitation (exitons:  $\text{Xe}^*$ )



which both lead to the formation of so-called diatomic excimers



These excimers eventually dissociate under the emission of vacuum ultraviolet (VUV) photon  $h\nu$  with a wavelength of 178 nm [AD10; CA13]. The formation of such xenon excimers is an important step, as otherwise xenon would not be transparent to its

scintillation light. The initial ratio of xenon excitons-to-ions  $N_{ex}/N_i$  can be described by a field  $F_D$  and energy  $E$  dependent model which is given by

$$N_{ex}/N_i = \alpha F_D^{-\zeta} (1 - e^{-\beta E}) \quad (2.4)$$

where  $\alpha, \beta, \zeta$  are free empirical parameters, which are constraint by measurements [Len+15].

The xenon scintillation light is emitted from the lowest two excited states of the xenon excimer, a singlet and triplet state, to the ground state. Both states are spectroscopically indistinguishable and have a decay time of 4 ns and 20 ns, respectively, which determines the scintillation pulse-shape [CA13]. The ratio between the two states depends on the density distribution of the deposited energy and therefore varies between different particle types. This enables pulse-shape discrimination in other liquid noble gases such as argon as the difference in decay times between singlet and triplet are much larger [CA13]. Delayed recombination between electrons and xenon ions may lead to a third component in the scintillation pulse shape with reported decay times in the order of 40 ns [CA13].

Applying an electric field to the liquid xenon, as done in a dual-phase TPC, quenches the delayed recombination and allows to read out the number of freed electrons as an additional signal channel. Doing so enables an alternative approach to discriminate ER from NR signals as the number of photons and electrons produced in an interaction depends on the linear energy transfer (LET) for a given particle and energy. Interactions with higher LET, such as caused from neutrons and WIMP induced NRs, lead to tracks with higher densities of electron-ion pairs which enhances the probability of recombination and therefore reduces the number of free electrons being extracted from the interaction point. The reduction depends on the strength of the electric field and can be characterized by the Thomas-Imel box model [TI87]. Figure 2.3 shows the behavior of light and charge yield<sup>1</sup> for a mono-energetic calibration source as a function of field strength. Both yields are strongly anti-correlated and in general it was found that the scintillation yield reduces by  $(74 \pm 2)\%$  in the presence of high fields [CA13]. Thus, only about 1/4 of all quanta are generated by xenon excitation. It should be stressed once more that only the number of electron-ion recombination is quenched by the electric field, but the total number of detectable quanta given by the sum of electron-ion pairs plus excitons stays constant<sup>2</sup>, as it is shown in Figure 2.3. The probability of recombination  $r$  is can be described by the Thomas-Imel box model [TI87] which is given by

$$r = 1 - \frac{\ln(1 + N_i \cdot \nu)}{N_i \cdot \nu} \quad (2.5)$$

<sup>1</sup>Light and charge yield refer to the number of detected photons and electrons per unit energy.

<sup>2</sup>It should be mentioned that this is not true in case of interactions with very high LET values, such as for low energetic NR,  $\alpha$ -particles or fission fragments. When the density of excitons becomes too high bi-excitonic quenching through the collision of two excitons, called Penning quenching [Mei+08], starts to play a role ( $Xe^* + Xe^* \rightarrow Xe + Xe^+ + e^-$ ) [CA13]. This leads to a reduction of the light signal.

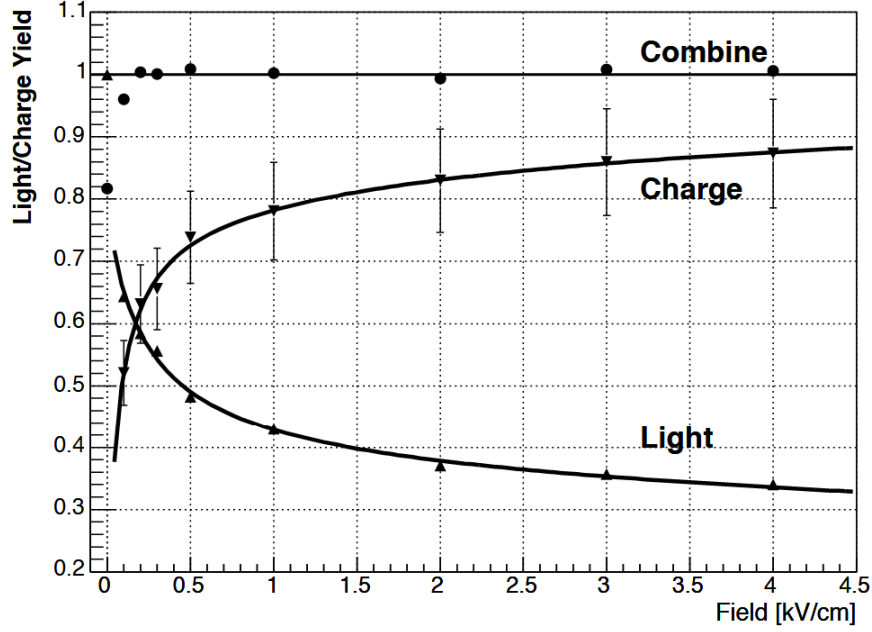


Figure 2.3. – Field dependence of the charge and light yield in liquid xenon. The dependence was measured using 662 keV  $\gamma$ -rays emitted by  $^{137}\text{Cs}$ . The figure was taken from [AD10]

where  $\nu$  can be parameterized with a power law depends on the electric field  $F_d$

$$\nu = \gamma \cdot F_d^{-\delta} \quad (2.6)$$

with the empirical parameters  $\gamma$  and  $\delta$  [Len+15].

Reading out both, scintillation photons and free electrons, simultaneously not only enables particle discrimination but also improves the energy resolution of the experiment. Since both signal channels are anti-correlated, so are their fluctuations. Thus, fluctuations of the combined signal are smaller than those of the two individual signal channels [AD10]. The number of produced excitons  $N_{\text{ex}}$  and the number of electron-ion pairs  $N_i$  for a given energy deposit  $E_0$  is given by

$$(N_{\text{ex}} + N_i) = \frac{E_0}{W} \cdot L \quad (2.7)$$

where  $W$  is the empirically average energy required to produce a detectable quanta. Currently, the majority of the liquid xenon dark matter experiments<sup>3</sup> use as accepted  $W$ -value ( $13.7 \pm 0.2$ ) eV [Dah09], although recent measurements suggest a lower value of about 11.5 eV [Ant+20; BST21].

The factor  $L$  in Equation (2.7) represents an absolute quenching factor to account for the non-detectable energy losses due to the atomic motion of the recoiling nucleus in case of NR. The quenching can be described by Lindhard's theory [LSS63; Szy+13] and plays an important role when comparing the reconstructed energy from NRs and

<sup>3</sup>Including the results presented in this work.

ERs. According to [Len+15] the quenching is parameterized via

$$L(E_0) = \frac{k \cdot g(\epsilon)}{1 + k \cdot g(\epsilon)} \quad (2.8)$$

where  $k$  is a proportionality constant between the velocity of the recoiling nucleus and the electronic stopping power.  $g(\epsilon)$  represents a factor that is proportional to the electronic stopping power to nuclear-stopping power.  $\epsilon$  is a dimensionless quantity that depends on the deposited energy and is parameterized with

$$\epsilon = 11.5 \cdot (E_0/\text{keV})Z^{-7/3} \quad (2.9)$$

where  $Z$  is the atomic number of the target nucleus. Using this parameterization  $g(\epsilon)$  can be expressed as

$$g(\epsilon) = 3\epsilon^{0.15} + 0.7\epsilon^{0.6} + \epsilon \quad (2.10)$$

For ER signals no quenching is observed and  $L = 1$  can be assumed. A precise knowledge and calibration of  $L$  is required to predict the liquid xenon response for WIMP dark matter.

It should be noted that due to the additional quenching the reconstructed energy scales of ER and NR signals cannot be directly compared. Therefore the dark matter community decided to use as a jargon to highlight their difference and use  $\text{keV}_{\text{ee}}$  or  $\text{keV}_{\text{Er}}$  as energy unit for ER signals, and  $\text{keV}_{\text{Nr}}$  for NR signals. The same jargon will also be used in this thesis.

Liquid xenon is a fairly good scintillator with a light yield comparable to other inorganic scintillators like NaI(Tl) [AD10; CA13]. To predict the scintillation yields for liquid xenon the Noble Element Simulation Technique (NEST) [Szy+11; Len+15] can be used. It allows the prediction of light and charge yield for liquid argon and xenon for different particle types, energies, and electric fields. The simulations are based on the semi-empirical models discussed in this section and validated based on measurements of the liquid argon and xenon community. Typical photon and electron yields predicted by NEST for gammas and neutrons are shown in Figure 2.4. The hard truncation of the NR yields at 0.2 keV nuclear recoil energy ( $\text{keV}_{\text{NR}}$ ) is due to missing calibration data needed to validate the semi-empirical NEST model for smaller recoil energies. In addition, it can be observed that NR yields show a weaker field dependence compared to ER which is due to the higher LET and therefore higher electron-ion recombination. This makes the extraction of free electrons harder compared to ER signals.

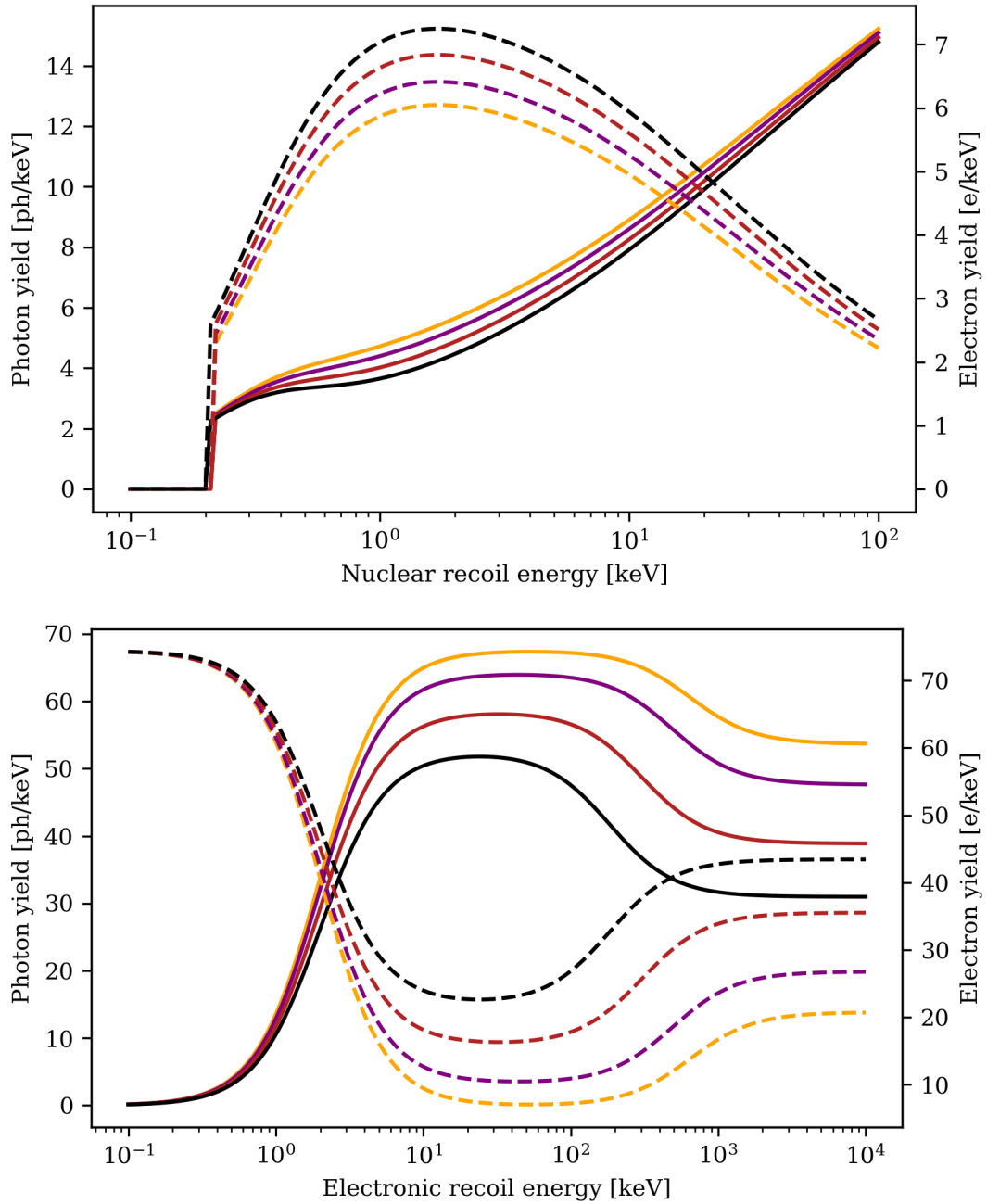


Figure 2.4. – Photon and electron yields for nuclear and electronic recoils. The solid lines show the photon yields and the dashed lines the electron yields. Both plots are showing the yields for an electric field of 10 kV/cm, 30 kV/cm, 100 kV/cm, 300 kV/cm in orange, purple, red and black respectively. The upper plots shows the yield for nuclear recoils and the lower plot for electronic recoils induced by  $\gamma$ -rays. The yield were estimated based on nestpy version 2.0.0. [Far+23].



## 2.2 DUAL-PHASE TIME PROJECTION CHAMBERS

### 2.2.1 Signal readout

One detector type which reads out both scintillation photons and free electrons is a dual-phase TPC. The concept of TPCs was originally invented by David Nygren in the 1970s as a central track detector for the PEP4 experiment [Nyg74]. This very first TPC was not yet operated as a dual-phase detector, but as a gaseous detector to reconstruct spatial tracks and particle energies by drifting electrons in a long cylindrical chamber. The electric field was augmented by a parallel magnetic field to reduce electron diffusion and to measure the momentum of charged particles. Based on this concept Carlo Rubbia suggested in 1977 the first liquid noble gas TPC in form of a large-scale liquid argon detector to study rare phenomena. Over the past decades this concept was developed further and nowadays most of the liquid noble gas dark matter experiments are operated as a dual-phase TPC [AD10].

The working principle of such a dual-phase TPC is sketched in Figure 2.5. The sensitive volume of the detector is given by a liquid xenon phase which is usually confined within a cylindrical container made of highly reflective material for VUV photons. Scintillation photons created by particle interactions are directly recorded by two arrays of photo-sensors which are mounted at the top and bottom of the cylinder. This promptly recorded signal is usually referred to as S1. Free electrons created along with the scintillation photons are recorded as a second delayed scintillation signal called S2<sup>4</sup> which is produced after drifting and extracting the electrons into a thin gas layer on top of the liquid phase.

This is achieved by two electric fields which are created by three semi-transparent electrodes, either made from individual steel wires [Apr+22c], woven meshes [Ake+20] or etched meshes from thin metal sheets [Apr+17c]. Two of these electrodes, gate and cathode, are located at the upper and lower end of the liquid phase producing a drift field to move the liberated electron cloud to the liquid-gas interface. The strength of the drift field varies between experiments and typical values between 20 V/cm and 300 V/cm are reached for large scale detectors [Apr+18; Apr+22c]. The drift speed of the electron cloud is directly proportional to the applied field and defines, together with the height of the cylinder, the maximum drift time of the detector. The third electrode called anode is mounted above the liquid surface. The field between gate and anode has two purposes: First, it helps the drifted electrons to overcome the potential barrier at the liquid-gas interface, to extract them into the xenon gas. The efficiency of the extraction depends on the field strength. It is about 50% for a field of about 3 kV/cm and reaches almost 100% around 5 kV/cm [CA13]. Second, it accelerates the extracted electrons until they reach sufficient kinetic energies to excite additional xenon atoms through collisions. Just like in the liquid phase the xenon excitons form excimers and produce the secondary scintillation signal S2 which is also recorded by

---

<sup>4</sup>S1 and S2 signals are commonly referred to as "peaks", while S1 and S2 which are belonging to the same interaction form an "event".

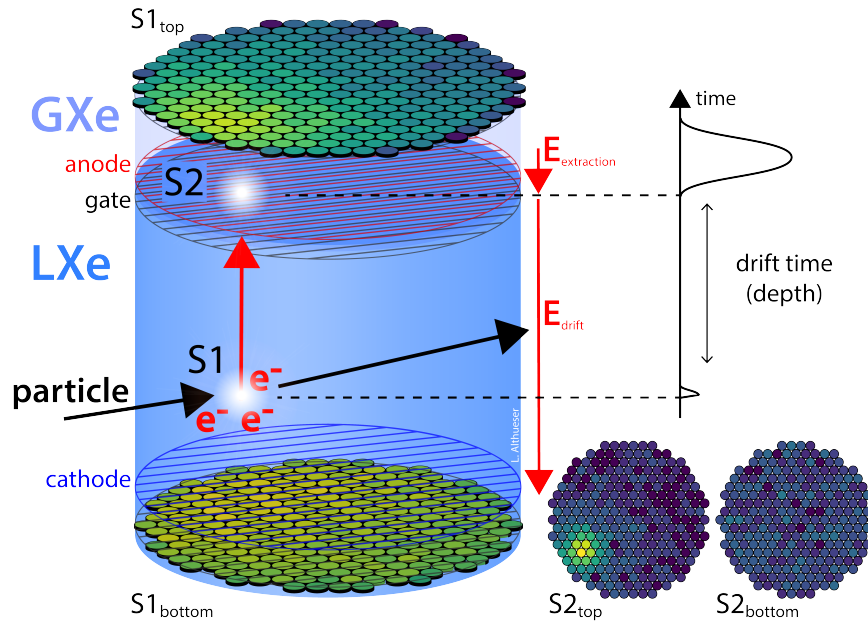


Figure 2.5. – Working principle of dual-phase time projection chamber. The colored circle indicates the number of photons recorded by the individual photo-sensors for an exemplary S2 signal. More details are given in the text. (Credit Lutz Althüser)

the two arrays of photo-sensors. This second scintillation signal is therefore proportional to the number of extracted electrons and thus is called proportional scintillation or electroluminescence [AD10; CA13]. At very high extraction fields even larger amplifications can be achieved through ionization, leading to an electron avalanche as used typically in gas detectors. However, avalanches have not yet been successfully deployed in large-scale dual-phase detectors [AD10]. Secondary scintillation is usually sufficient to detect signals down to a single electron (SE) response with yields of about 100 ph/e/cm<sup>5</sup>. Given the large electroluminescence yields, and the quenching of recombination due to the applied drift field in the liquid phase, the integrated S2 signal area<sup>6</sup> is usually much larger than their S1 counter parts as indicated in Figure 2.5. Additionally, are the shapes of the two signals quite different. While the width of S1 signals is in the order of a few ~10 ns, mainly driven by the scintillation pulse shape of xenon, the photon-propagation, and optical parameters of the detector, the S2 shape is mainly determined by the diffusion of the electron cloud during the electron drift and the shape of the electroluminescence signal of single electron S2s. The time interval in which an S2 signal reaches 50 % of its integrated area  $r_{50}^{\text{model}}$  can be modeled with

$$r_{50}^{\text{model}} = \sqrt{\frac{3.64 \cdot D \cdot t_d}{v_d^2} + w_0^2} \quad (2.11)$$

where  $D$  represents the diffusion constant in liquid xenon,  $t_d$  the drift time and  $v_d$

<sup>5</sup> Estimated via Equation (4.9) in [CA13] for an extraction field of 3 kV/cm and a gas pressure of 1.5 bar.

<sup>6</sup> In the further discussion the jargon "area" will be always used to refer to the integrated signal size of the corresponding signal, e.g. S1, S2 or neutron veto "events"

the drift velocity respectively.  $w_0$  represents the S2 width of single electrons [Sor11; Apr+19c]. The size of S2 signals is typically in the scale of  $\sim 1 \mu\text{s}$  for large-scale detectors like XENON1T and XENONnT.

The S2 width depends on the diffusion is one of the most important characteristics of an S2 signal as it allows to discriminate wrongly paired S1 and S2 signals by the correlation between drift time and S2 width [Apr+19c]. Such a wrong pairing can occur during the signal processing if for example multiple energy deposits are recorded within one drift time. At low energies wrong pairing becomes more likely as "false" S2 signals, of a size of single to a few electrons, are created via photoionization of impurities in LXe and detector materials due to the VUV scintillation light. The delayed extraction of electrons that are trapped beneath the liquid surface for extraction efficiency below  $\sim 100\%$  may play a role, too. Both these effects lead to a constant single to a few electrons S2 background in large-scale detectors. Also "false" S1 signals can lead to wrongly paired signals. "false" S1 signals can either originate from true energy deposits, e.g. if an interaction occurs in an S2 insensitive region of the detector like below the cathode electrode, or originate from fake signals of the used photo-sensors [Apr+19b; Apr+19c].

### 2.2.2 Particle discrimination and fiducialization

Reading out both S1 and S2 signals in a dual-phase TPC has several advantages when searching for rare event phenomena. First of all, it enables particle discrimination as discussed in Section 2.1.2 and shown in Figure 2.6. Secondly, it enables a full position reconstruction of the interaction vertex of the incoming particle. The position in a plane parallel to the footprint of the cylinder can be reconstructed from the hit pattern of the S2 signal measured by the top photo-sensor array. This position is referred to as the  $xy$ -coordinate of the interaction, while the  $z$ -coordinate, along the height of the cylinder, is estimated through the drift time between the S1 and S2 signal, and the drift velocity<sup>7</sup>. The full vertex reconstruction is also sketched in Figure 2.5. The resolution of the position reconstruction usually depends on the number of electrons contained in the S2 signal. Large-scale detectors can reach a resolution in the order of 1 cm [Apr+19c].

The vertex reconstruction paired with the relatively good position resolution allows the discrimination of background from dark matter or other rare-event signals in two different ways. First of all, only the signals in an inner virtual volume, the so-called fiducial volume, are considered in an analysis. Figure 2.7 shows exemplarily the fiducial volume used in the XENON1T low ER excess analysis [Apr+20a]. Choosing such a fiducial volume allows to shield background signals induced by radiation emitted from detector materials. Especially, ER signals induced by  $\beta$ - and  $\gamma$ -particles can be easily shielded in this manner given the high stopping power of liquid xenon. Neutrons, emitted through radiogenic processes like spallation and  $(\alpha, n)$ -reactions, cannot be shielded, as it can be inferred from Figure 2.2. Neutrons with an energy of about

<sup>7</sup>For further discussion  $x = y = z = 0$  will be defined as the center of the TPC at height of the liquid-gas interface.

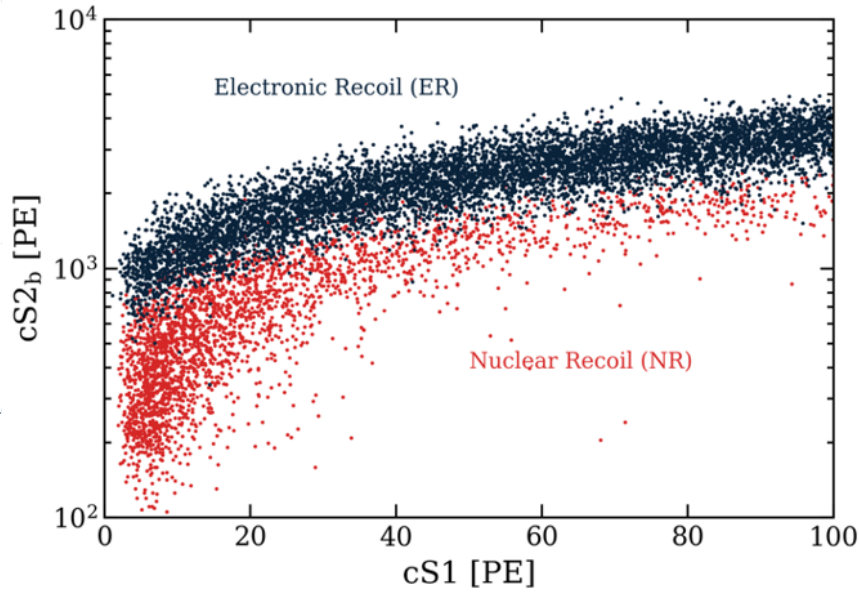


Figure 2.6. – Electronic and nuclear recoil calibration of the XENON1T experiment in the  $(cS1, cS2_b)$  space. The letter “c” indicates that both S1 and S2 signals were corrected for detector dependent effects. The subscript letter “b” for the S2 signal indicates that only photons detected by the bottom photo-sensor array are considered for the S2 signal. The calibration of the ER and NR band were performed using  $^{220}\text{Rn}$  and AmBe respectively. Given the higher ion-electron recombination for NR signals with respect to ERs, NRs appear with a quenched S2 signal. (Credit XENON collaboration)

1 MeV have a free mean path in liquid xenon of about 10 cm and only a small fraction of their energy is lost during the NR. Thus, neutrons can travel long distances and scatter multiple times when traveling through the detector. This multi-scatter signature, on the other hand, can be distinguished from dark matter signals, as e.g. WIMPs are expected to scatter only once in the detector given their small interaction cross section. Multi-scatter signal would appear in the detector as a single summed-up S1 signal from all interaction vertices because the neutron time-of-flight is small compared to the typical pulse shape of an S1 signal, and most likely as multiple S2 signals. The S2 signals can be either distinguished as individual signals if the separation in  $z$  is large enough such that the diffused electron clouds can be reconstructed separately, or via the photo-sensor hit pattern if the separation in  $z$  is too small and the S2 signal is reconstructed as a single monolithic S2. Thus, a good spatial resolution of the detector is necessary to maximize the multi-scatter signal rejection. Identifying multi-scatter signals however is limited by the detector’s intrinsic S2 background. Delayed extraction of electrons which are trapped below the liquid-gas interface lead to a constant single to few electron S2 background in the detector. Therefore, it becomes very difficult to distinguish multi- from single-scatter signals at lower energies. This can also be seen in Figure 2.6. A few of the NR calibration points are found below the main band featuring a larger S1 being paired with an atypical small S2. These signals are most likely multi-scatter signals. More severe for the search of WIMP dark matter are NR induced by neutrons which leave the sensitive volume of the TPC after performing a

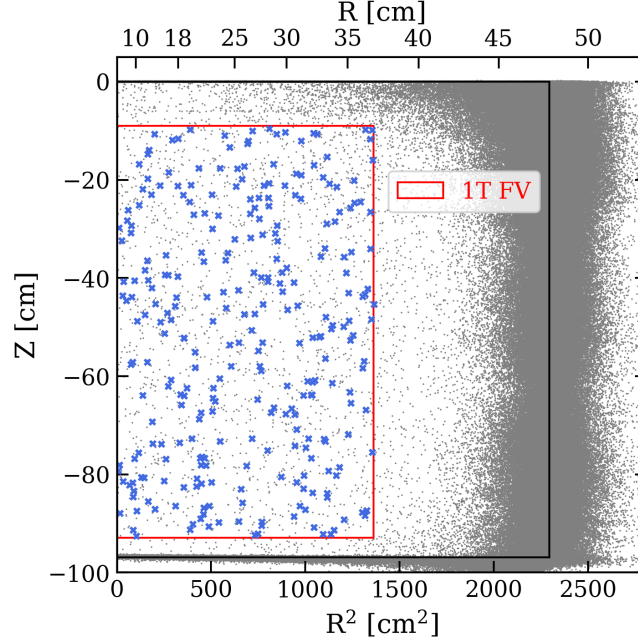


Figure 2.7. – Example fiducial volume of the XENON1T low ER excess analysis. The plot shows the reconstructed interaction vertices in the  $(R^2, z)$ -space. The red contour indicates the 1 t fiducial volume used in the corresponding analysis. Blue dots show all events with an electronic recoil energy between 1 keV and 7 keV, while gray shows all data points below 30 keV. The dimension of the TPC are indicated as black solid lines. Some interaction vertices are wrongly reconstructed outside of the detector. (Credit: XENON collaboration)

single back-scatter. These signals are indistinguishable from WIMPs

### 2.2.3 Signal correction and energy reconstruction

Especially in large-scale detectors, the signal response of S1 and S2 is not uniform throughout the detector. Thus, as already indicated in Figure 2.6, S1 and S2 signals are corrected for their spatial dependence to get a detector-averaged response. In this way uncertainties in the reconstructed recoil energy are reduced for the DM inference. The discussion in the following is based on [Apr+19c] when applicable with additional inputs based on the analysis of the first science run of XENONnT.

S1 signals of a given energy deposit  $E_0$  can be described in terms of the so-called light yield LY as

$$S1(R, \theta, z, E_0, F_d) = E_0 \cdot LY(R, \theta, z, E_0, F_d) \quad (2.12)$$

where  $R, \theta$  and  $z$  indicate the spatial dependence in terms of the radius, azimuth angle and depth of the interaction vertex, and  $F_d$  is the dependence on the drift field at the given vertex. The LY itself can be further broken down into three terms

$$LY(R, \theta, z, E_0, F_d) = \epsilon_{\text{LCE}}(R, \theta, z) \cdot \epsilon_{\text{ph}} \cdot PY(E_0, F_d(R, \theta, z, t)) \quad (2.13)$$



where the product of  $\epsilon_{\text{LCE}}(R, \theta, z)$  and  $\epsilon_{\text{ph}}$  describe the chance of detecting a photon on a given photo-sensor and  $PY(E_0, F_d(R, \theta, z, t))$  describes the photon-yield – the number of photons emitted per unit energy which is shown in Figure 2.4. The chance of detecting a photon is given by the position-dependent light collection efficiency (LCE)  $\epsilon_{\text{LCE}}(R, \theta, z)$ , which describes the chance that a photon impinges on the sensitive area of a photo-sensor given the interaction position, and the photon-detection efficiency  $\epsilon_{\text{ph}}$  of the given photo-sensor type, which describes the chance that an impinging photon is converted into a measurable signal. Since most of the experiments use photo-multiplier tube (PMT) as photo-sensors,  $\epsilon_{\text{ph}}$  can be usually described by the quantum efficiency (QE)<sup>8</sup> and charge collection efficiency (CE) of the PMTs.

The LCE is in general position-dependent because of Rayleigh scattering and photon absorption within the liquid xenon, as well as the total reflection of photons at the liquid-gas interface. To correct for these effects the isomeric  $^{83\text{m}}\text{Kr}$  is diluted into the detector which distributes uniformly inside the liquid and gaseous phase.  $^{83\text{m}}\text{Kr}$  decays via a double transition, emitting first 31.1 keV and then 9.4 keV in form of multiple conversion and Auger electrons, as well as x- and  $\gamma$ -rays [Man+10]. The decay time of the second transition is 154 ns which makes it possible to distinguish a certain fraction of the two corresponding S1 signals. In XNEONnT a relative LCE correction map  $L_c(R, \theta, z)$  is constructed from the combined S1 signals of the total 41.5 keV transition by dividing the detector in multiple small voxels in  $(R, \theta, z)$ . The average integrated S1 signal size of each voxel is normalized by the detector average S1 response  $\langle S1 \rangle$  such that a relative map can be computed via an interpolation between the individual voxel:

$$\begin{aligned} L_c(R, \theta, z) \Big|_{E'_0=41.5\text{keV}} &= \frac{LY(R, \theta, z, E'_0, F_d)}{\langle LY(E'_0, F_d) \rangle} \\ &= \frac{\epsilon_{\text{ph}} \cdot \epsilon_{\text{LCE}}(R, \theta, z) \cdot PY(E'_0, F_d(R, \theta, z, t))}{\langle \epsilon_{\text{ph}} \cdot \epsilon_{\text{LCE}}(R, \theta, z) \rangle \cdot \langle PY(E'_0, F_d(R, \theta, z, t)) \rangle} \end{aligned}$$

This  $L_c(R, \theta, z)$  map is strictly speaking only valid for a uniform drift field. This is because spatial variations in  $F_d$  lead to spatial variations in the photon-yield and therefore to an energy dependence which is indicated via  $E'_0$ . The variation is much smaller compared to  $\epsilon_{\text{LCE}}(R, \theta, z) / \langle \epsilon_{\text{LCE}} \rangle$ , but is still taken into account in the LCE map of XENONnT, using field simulations<sup>9</sup> and a photon-yield model measured in [Jör+22]. Taking these effects into account the correction of S1 into cS1 is then given by

$$\frac{cS1}{E_0} = \langle LY(E_0, F_d) \rangle = \frac{S1}{E_0 \cdot L_c(R, \theta, z)}. \quad (2.14)$$

It should be noticed that cS1 is still not fully spatially independent, as the photon yield

<sup>8</sup>VUV photons have a certain chance to generate two photo-electrons when impinging on the photocathode of the PMT. This effect is referred to in the literature as double photo-electron (DPE) [Apr+19b] and is typically in the order of 20% with a weak dependence on QE. The DPE is usually absorbed in corrections by assuming on average a higher QE of the PMTs.

<sup>9</sup>The simulated field map is shown in the appendix in Figure B.12 on page 246. The map itself is compared against data using the light-yield ratio of the two  $^{83\text{m}}\text{Kr}$  transitions as done in [Ake+17a].

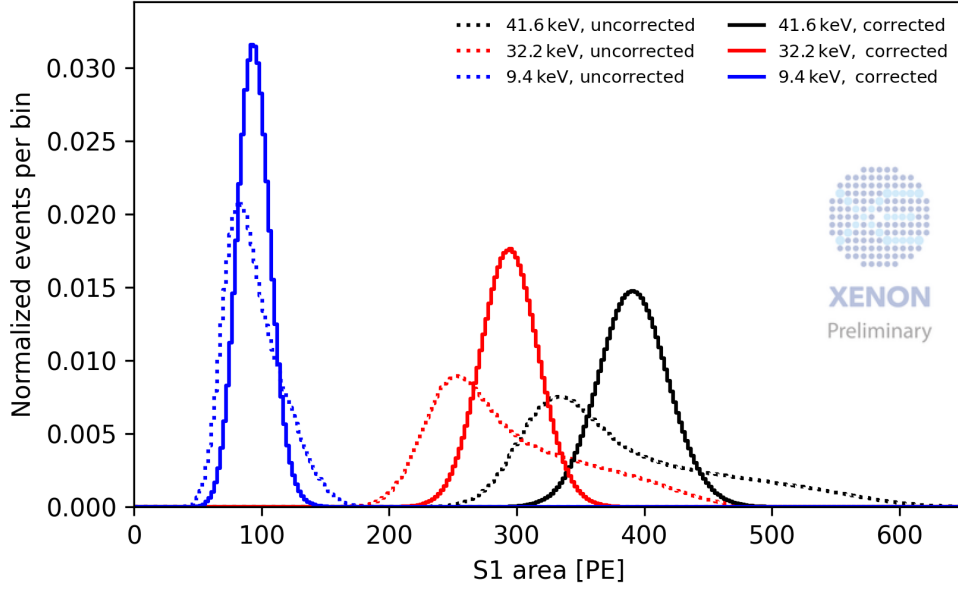


Figure 2.8. – Comparison of the S1 response of XENONnT for the lines of  $^{83m}\text{Kr}$  w/ and w/o S1 corrections applied. (Credit Johannes Jacob, XENON collaboration)

remains dependent on the drift field, particle type and energy. However, this effect cannot be corrected as it would require on an event-by-event basis prior knowledge about the type of interaction. Figure B.1 in the appendix on page 238 shows the S1 correction map used for the first science run of XENONnT. Its impact on the resolution of S1 signals and therefore the energy resolution of the detector is shown in Figure 2.8.

Also S2 signals are corrected for spatial-dependent non-uniformities in the detector response. The corrections can be usually divided into two parts: Electrons liberated in liquid xenon can be lost by attachment to electronegative impurities in the liquid such as oxygen or water. The chance of being lost depends strongly on the concentration of electronegative impurities and the drift time required to move the electron cloud to the liquid-gas interface. The fraction of surviving electrons  $\langle \epsilon_{\text{el}} \rangle$  averaged over  $(R, \theta)$  can be described by simple exponential distribution of the form

$$\langle \epsilon_{\text{el}}(z, \tau_e, F_d) \rangle = \frac{S2(z, G(R, \theta, z, E_0, F_e))}{S2(z=0, G(R, \theta, z=0, E_0, F_d))} = \exp \left\{ - \frac{z}{\tau_e \cdot v_d(F_d)} \right\} \quad (2.15)$$

where  $\tau_e$  is the average electron "lifetime" before being lost,  $v_d(F_d)$  the drift velocity of the electrons for the given drift field, and  $G(R, \theta, z, E_0, F_e)$  the gas gain of an S2 signal for the given position, extraction field  $F_e$  and energy deposit. To correct for the electron lifetime mono-energetic decay lines are used. For example in XENONnT energy lines from  $^{37}\text{Ar}$  and  $^{83m}\text{Kr}$  are used as well as  $\alpha$ -decays from  $^{220}\text{Rn}$ . The second part of the S2 correction is related to the gas gain  $G(R, \theta, z, F_e)$ . The gas gain describes the average signal recorded on the photo-sensor arrays given a single ionization electron which

was extracted into the gaseous xenon. The gas gain can be divided into two terms

$$G(R, \theta, z, F_e) = g_{\text{textSE}}(R', \theta', F_e) \cdot \epsilon_{\text{LCE,S2}}(R', \theta', F_e). \quad (2.16)$$

The first term  $g_{\text{SE}}(R', \theta', F_e)$  describes the SE gain which describes the number of scintillation photons produced for an extracted electron given the extraction field and the height of the gas gap at the extracted position  $(R', \theta')$ . The second term  $\epsilon_{\text{LCE,S2}}(R', \theta', F_e)$  describes again the light collection efficiency of the photo-sensor arrays at the given extraction position. The prime next to the position  $(R', \theta')$  indicates that the extracted position may differ from the interaction position  $(R, \theta, z)$  due to the shape of the drift field. This effect is corrected by the same simulated drift field map which was already mentioned in the discussion of the light yield. The last correction factor is the electron extraction efficiency  $\epsilon_{\text{EE}}(R', \theta', F_e)$  which describes the probability of extracting an electron for the given extraction field at the given extraction position.

To correct for spatial S2 variations the same strategy as for the light yield is used. A mono-energetic gaseous calibration source, e.g.  $^{83\text{m}}\text{Kr}$ , is diluted into the liquid xenon and a relative response map  $Q_c(R, \theta, z)$  is computed. The map is created by dividing again the detector into small voxels in  $(R', \theta')$  and comparing the average S2 response of each voxel with the S2 response of the detector  $\langle S2 \rangle$ . This map usually takes into account the combined effect of  $G(R, \theta, z, F_e)$  and  $\epsilon_{\text{EE}}(R', \theta', F_e)$  and can be described by

$$Q_c(R, \theta, z) \Big|_{E'_0=32.1 \text{ keV}} = \frac{G(R, \theta, z, F_e) \cdot \epsilon_{\text{EE}}(R', \theta', F_e) \cdot QY(R, \theta, z, F_d, E'_0)}{\langle G(R, \theta, z, F_e) \rangle \cdot \langle \epsilon_{\text{EE}}(R', \theta', F_e) \rangle \cdot \langle QY(R, \theta, z, F_d, E'_0) \rangle} \quad (2.17)$$

where  $QY(R, \theta, z, F_d, E'_0)$  represents the electron yield of liquid xenon shown in Figure 2.4. However, in contrast to the light yield map, the charge yield is not correct for a non-uniform electron yield response  $Q(R, \theta, z, F_d, E'_0)$  due to drift field variations in XENONnT. This leads to a systematic uncertainty in the correction which is small for NR as their high LET leads to only a weak field dependence of  $QY(R, \theta, z, F_d, E'_0)$ .

Just as for the S1 signal, S2 signals can be written in terms of the charge yield (CY)

$$\begin{aligned} S2(R, \theta, z, E_0, F_d, \tau_e, F_e) &= CY(R, \theta, z, E_0, F_d, \tau_e, F_e) \cdot E_0 \\ &= \langle \epsilon_{\text{el}} \rangle \cdot G \cdot \epsilon_{\text{EE}} \cdot QY \cdot E_0 \end{aligned} \quad (2.18)$$

and thus the corrected signal is given by

$$\frac{cS2}{E_0} = \langle CY(R, \theta, z, E_0, F_d, \tau_e, F_e) \rangle = \frac{S2}{E_0 \cdot Q_c(R, \theta, z) \cdot \langle \epsilon_{\text{el}} \rangle}. \quad (2.19)$$

The S2 correction map used during the first science run of XENONnT is shown in the appendix in Figure B.2 on page 239.

Based on the correct S1 and S2 signals the energy deposit of a given particle interaction can be estimated via Equation (2.7). To do so the recorded cS1 and cS2 signals must first be converted into the number of photons and electrons of the respective signal. Usually, this conversion is given in terms of the detector-dependent scintillation



gain  $g1$  and the ionization gain  $g2$  which are defined as

$$g1 := \frac{cS1}{N_{\text{ex}}} \quad (2.20)$$

$$g2 := \frac{cS2}{N_i} \quad (2.21)$$

and thus Equation (2.7) can be rewritten as

$$E_0 = \frac{W}{L} \cdot \left( \frac{cS1}{g1} + \frac{cS2}{g2} \right). \quad (2.22)$$

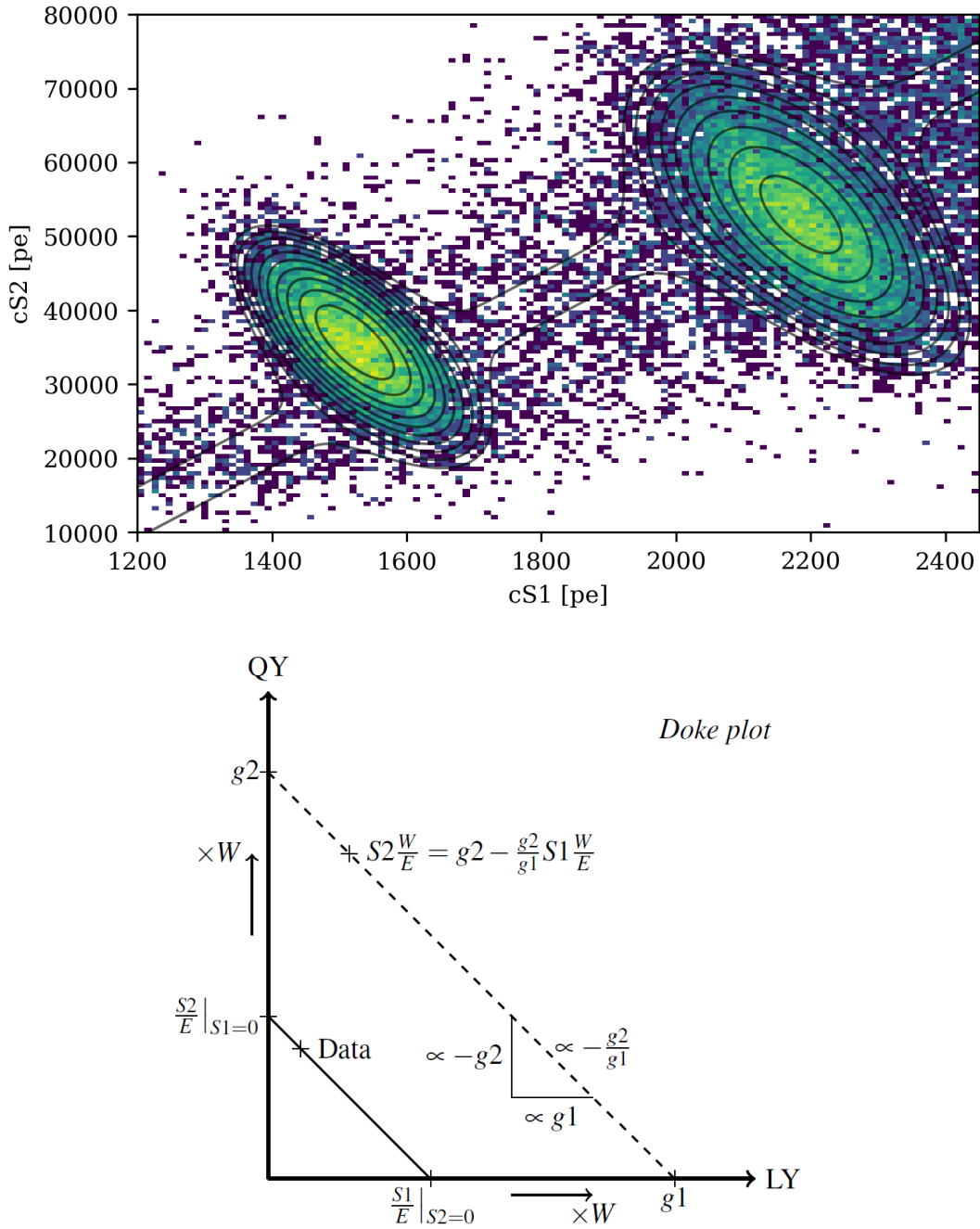
The corresponding energy scale is sometimes referred to in the literature as the "combined energy scale".  $g1$  and  $g2$  represent hereby the detector averaged response for the charge and light yield and can therefore be expressed in the response terms discussed above via Equation (2.14) and Equation (2.19):

$$g1 = \langle \epsilon_{\text{LCE}}(R, \theta, z) \rangle \cdot \langle \epsilon_{\text{ph}} \rangle \quad (2.23)$$

$$g2 = \langle G(R, \theta, z, F_e) \rangle \cdot \langle \epsilon_{\text{EE}}(R', \theta', F_e) \rangle \cdot \langle \epsilon_{\text{el}}(z, \tau_e, F_d) \rangle. \quad (2.24)$$

The calibration of  $g1$  and  $g2$  is typically done via several mono-energetic line sources, e.g.  $^{37}\text{Ar}$  [Apr+22a] or  $^{83\text{m}}\text{Kr}$ , which are assumed to be uniformly diluted in liquid xenon or by the isomeric xenon isotopes  $^{129\text{m}}\text{Xe}$  and  $^{131\text{m}}\text{Xe}$  which are by-products of the NR response calibration via neutrons. Figure 2.8 shows the typical response of the XENONnT TPC for the two isomeric xenon isotopes  $^{129\text{m}}\text{Xe}$  and  $^{131\text{m}}\text{Xe}$ . The full energy peaks of the two lines are represented by two-dimensional normal distributions with a non-zero covariance due to the anti-correlation between charge and light. The correlation angle depends on the deposited energy. After determining the position of the individual peaks,  $g1$  and  $g2$  can be determined from the slope and offset of a simple line in the so-called "Doke plot" which is sketched in Figure 2.9.

Scintillation and ionization gain are detector specific quantities. Values between  $\sim 0.13$  pe/photon and  $\sim 0.15$  pe/photon have been reached for  $g1$  and values between  $\sim 16.5$  pe/electron and  $\sim 56.5$  pe/electron for  $g2$  in large scale experiments [Apr+17b; Apr+22c; Aal+22]. The Doke plot for the first science run of XENONnT is shown in the appendix in Figure B.55 on page 280.



**Figure 2.9.** – Full energy ellipses of the isotopes  $^{129m}\text{Xe}$  and  $^{131m}\text{Xe}$  measured with the XENONnT experiment, and illustration of the “Doke” method. The **upper plot** shows the full energy ellipse of  $^{129m}\text{Xe}$  and  $^{131m}\text{Xe}$  in cS1 and cS2, where the ellipse with smaller cS1 and cS2 belongs to energy deposits of  $^{131m}\text{Xe}$ . The black contour lines indicate two ellipses representing the full energy peaks. The model was fitted via a binned maximum likelihood. The bin color encodes the number of entries per bin. (Credit Henning Schulze, XENON collaboration). The **lower plot** illustrates the computation of  $g1$  and  $g2$  via the “Doke plot” method. The cross labeled data represents the position of a full energy ellipse of a given mono-energetic energy line (e.g.  $^{129m}\text{Xe}$  and  $^{131m}\text{Xe}$ ). The figure was taken from [BST21].

### 2.3 WATER CHERENKOV DETECTORS AND NEUTRON DETECTION IN WATER

The Cherenkov effect, named after Pavel A. Cherenkov who received 1958 the Nobel prizes for its discovery, is used in a wide range of different applications in high energy physics [Alb+02], as well as neutrino [Abe+20; Abb+23] and  $\gamma$ -ray astronomy [Abd+22].

The Cherenkov effect occurs when a charged particle travels faster than the phase velocity of light

$$c_n = \frac{c_0}{n(\omega)} \quad (2.25)$$

in a given dielectric and transparent medium with refractive index  $n(\omega)$  for a given frequency  $\omega$  and the speed of light in vacuum  $c_0$ . The light emission can be understood as coherent due to an asymmetric polarization of the dielectric medium. Atoms in the vicinity of the traversing charged particle are polarized by its electromagnetic field which can be depicted, as shown in Figure 2.10, as a short elliptical deformation of the spherical atoms, before relaxing again as the particle moves further. If the speed of the charged particle is not high enough the medium is spherically, symmetrically polarized such that the relaxation of the individual dipoles does not lead to any measurable radiation. However, if the particle is faster dipoles along the flight direction cannot relax fast enough and an asymmetric polarization field forms. The relaxation of the individual dipoles leads, according to the Huygens principle, to a number of elementary waves which propagated at the phase velocity of the medium. As the charged particle travels faster than the elementary waves, the individual waves can interfere constructively and form a single wavefront which is then seen as Cherenkov radiation under an angle  $\theta_c$  [KW16]. The Cherenkov angle  $\theta_c$  can be determined through the travel speed of the charged particle and the elementary waves after a time  $t$  as

$$\cos(\theta_c) = \frac{\frac{c_0}{n} \cdot t}{v \cdot t} = \frac{1}{\beta \cdot n} \quad (2.26)$$

where  $v$  is the velocity of the charged particle [KW16]. The minimal energy required  $E_{\text{th}}$ , that a given particle produces Cherenkov radiation, can be estimated via Equation (2.26) for the case that  $\theta_c = 0$  which yields  $\beta_{\text{th}} = 1/n$ . Thus, using relativistic kinematics<sup>10</sup> the minimal particle energy required is given by

$$E_{\text{th}} = \gamma_{\text{th}} \cdot m = \frac{m}{\sqrt{1 - \beta_{\text{th}}^2}} = \frac{n}{\sqrt{n^2 - 1}} \cdot m \quad (2.27)$$

where  $m$  is the mass of the given particle.

Most of the modern Cherenkov detectors are Ring Imaging Cherenkov (RICH) detectors. As the name suggests these detectors measure the emitted Cherenkov ring, that is created either when a charged particle slows below the threshold velocity  $\beta_{\text{th}}$  during the Cherenkov light emission, or if only a thin radiator medium is used which does not allow the particle to slow below  $\beta_{\text{th}}$  before exciting the detector. Measuring the

<sup>10</sup>Using the notation  $c = 1$  here.

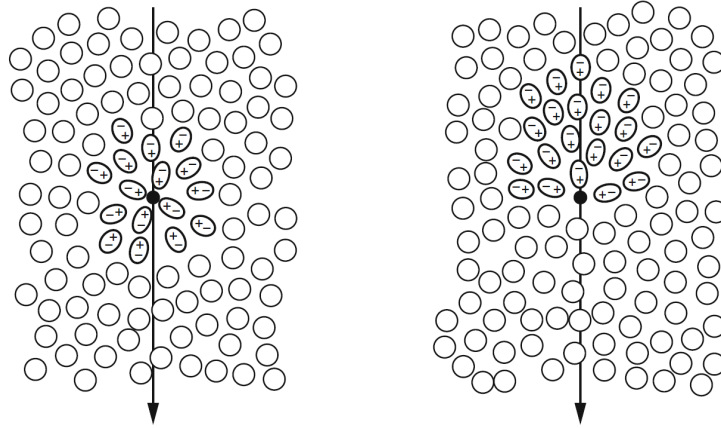
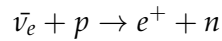


Figure 2.10. – Illustration of the Cherenkov effect when a charged particle traversing a dielectric medium. Left (right) shows the case when the charged particle has a speed lower (higher) than the speed of light for the given medium. The image was taken from [KW16].

ring allows to discriminate between several different charged particles by comparing their respective Cherenkov angles  $\theta_c$ . Among the most famous RICH experiments is the water Cherenkov detector Super-Kamiokande (SK) in Japan [Zha+16]. Although, originally proposed for the measurement of the proton decay, it has significantly contributed to the discovery of neutrino oscillations and serves as a neutrino observatory. One of its signals channels is the inverse beta decay



in which the subsequently emitted neutron is used to tag the reaction for a better signal-to-noise ratio [Zha+16]. The tagging is performed through a delayed coincidence between the prompt  $e^+$  Cherenkov signal and the delayed neutron capture on hydrogen, which emits a 2.22 MeV  $\gamma$ -ray. The  $\gamma$ -ray has a free mean path of  $\sim 30$  cm before it transfers its energy through Compton scattering and photoabsorption to the shell electrons of the surrounding water molecules. Eventually, the freed electrons also emit Cherenkov light which can be measured by the detector. The mass attenuation and free mean path of  $\gamma$ -rays in water are shown in Figure 2.11.

Neutrons on the other hand have a much shorter free mean path of  $\sim 10$  cm. They lose their kinetic energy predominantly through elastic collisions with hydrogen, before they are eventually captured. The neutron capture cross section  $\sigma_{H,n}$  is directly related to the neutron capture time through

$$\tau = \frac{1}{\sigma_{H,n} \cdot n_H \cdot \bar{v}_n} \quad (2.28)$$

where  $n_H$  is the hydrogen density of the given water target and  $\bar{v}_n$  the average neutron velocity [CM77; And+20]. For neutron energies below  $\sim 10$  keV, the capture cross section  $\sigma_{H,n}$  scales with  $1/\bar{v}_n$ , which means that the delay between the prompt  $e^+$

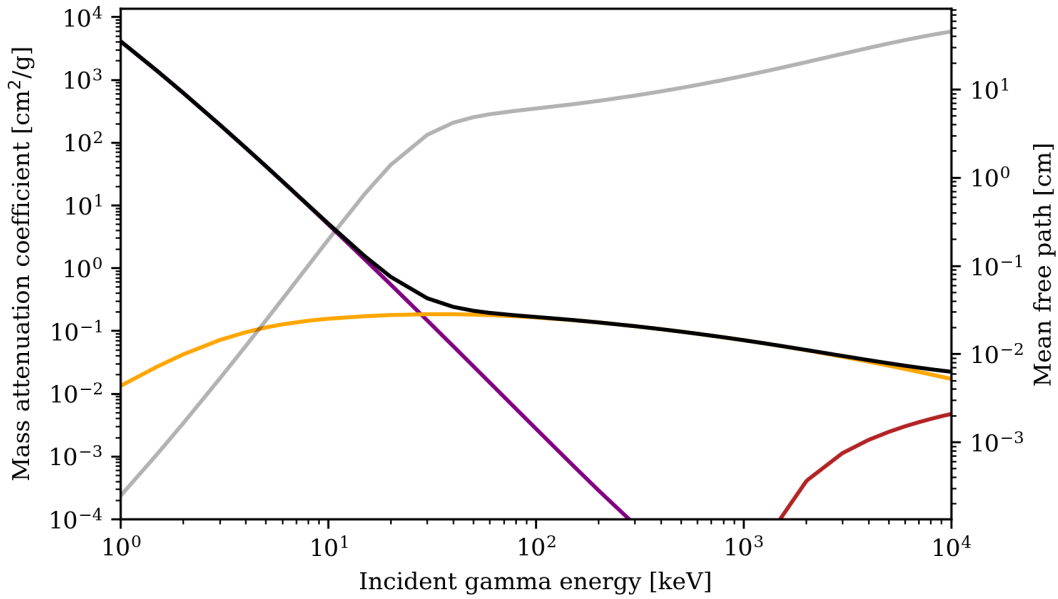


Figure 2.11. – Mass attenuation of  $\gamma$ -rays in water. The black line shows the total mass attenuation, while the purple, orange and red line are showing the contributions from photoabsorption, Compton scattering and pair-production respectively. The gray shaded line indicates the corresponding free mean path. The data was taken from [Ber+10].

Cherenkov signal and the neutron capture signals in SK is mostly independent of neutron moderation and determined through the capture process itself. Figure 2.12 shows the neutron capture and scatter cross sections of water. The neutron capture time is  $\sim 200 \mu\text{s}$  with values of  $(203.7 \pm 2.8) \mu\text{s}$  and  $202.35^{+0.87}_{-0.76} \mu\text{s}$  reported by the water Cherenkov experiments of SK [Zha+16] and Sudbury Neutrino Observatory (SNO) [And+20] respectively.

The very same detection principle of neutrons can also be exploited in a water Cherenkov neutron veto as it is used in XENONnT. Just like in SK, the neutron veto of XENONnT uses a delayed coincidence between the prompt S1 signal in the TPC and the delayed neutron capture in the NV, to tag neutron induced signals. To better understand the requirements of operating a water Cherenkov detector as a neutron veto, which has never been done before, some signal characteristics as the Cherenkov threshold and number of photons produced need to be determined.

Water has a refractive index of  $n(\lambda = 589.29 \text{ nm}) = 1.33$  [IG85] which increases towards smaller wavelengths. This means that the Cherenkov threshold for an electron in water is  $E_{\text{th}} \approx 775 \text{ keV}$  which corresponds to a kinetic energy of  $\sim 264 \text{ keV}$ . While the threshold is well below the energy released by the neutron capture on hydrogen, one needs to determine how many photons are produced for the remaining  $\sim 2 \text{ MeV}$  energy. To estimate the number of produced photons the differential emission formula derived by Frank and Tamm can be used. The formula describes the amount of energy radiated away per unit frequency of the emitted photons and unit distance traveled

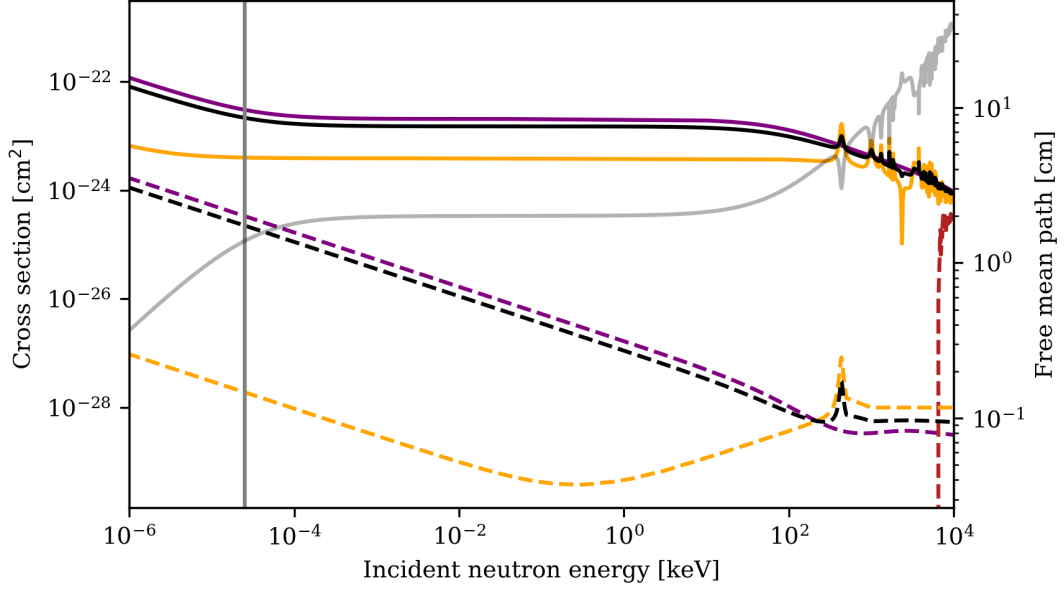


Figure 2.12. – Neutron interaction cross sections for water. The purple and orange lines show the interaction cross sections for hydrogen and oxygen respectively. Solid and dashed lines show the elastic scattering and capture cross sections for neutrons on the respective element. The red dashed line indicates in addition the inelastic neutron scattering cross section for oxygen. Black lines show the corresponding averaged interaction cross sections for water. The gray line indicates the resulting free mean path of neutrons in water. The dark gray vertical line highlights the neutron energy equivalent for the thermal neutrons with a velocity of 2200 m/s. The data was taken from [Plo+20].

by the charged particle. It is given by

$$\frac{d^2E}{d\omega dx} = \frac{z^2 e^2}{4\pi\epsilon_0 c^2} \cdot \omega \cdot \left(1 - \frac{1}{\beta n(\omega)^2}\right) \quad (2.29)$$

where  $z \cdot e$  is the charge of the given particle and  $\epsilon_0$  the vacuum permittivity. Equation (2.29) can also be rewritten as unit energy emitted per unit wavelength using the relation  $\omega = \frac{2\pi c}{\lambda}$  which results in

$$\frac{d^2E}{d\lambda dx} = -\frac{\pi z^2 e^2}{\epsilon_0} \cdot \frac{1}{\lambda^3} \cdot \left(1 - \frac{1}{\beta n(2\pi \cdot c/\lambda)^2}\right). \quad (2.30)$$

Thus, in order to estimate the amount of photons emitted in a given wavelength range one needs to know the distance the charged particle can travel in the given medium. The travel distance of an electron in water is given by the corresponding stopping power due to ionization and radiation losses and has a value of  $1.85 \text{ MeVcm}^2/\text{g}$  at an electron energy of 2 MeV [Ber+17]. This corresponds to a CSDA electron range of  $0.987 \text{ g/cm}^2$ .

Side remark: In comparison to the total stopping power, the energy losses due to Cherenkov radiation is negligible. It can be estimated through Equation (2.30) using

the assumptions:  $n(\omega) = \text{const.} = 1.333$  and  $\beta \approx 0.97 \approx 1$ <sup>11</sup>, resulting in a stopping power of

$$\begin{aligned} -\frac{dE}{dx} &= \frac{\pi e^2}{\epsilon_0} \cdot \underbrace{\left(1 - \frac{1}{n^2}\right)}_{=0.435} \cdot \int_{200\text{ nm}}^{700\text{ nm}} \frac{1}{\lambda^3} d\lambda \\ &= 0.435 \cdot \frac{\pi e^2}{\epsilon_0} \cdot \left[ -\frac{1}{2}\lambda^2 \right]_{200\text{ nm}}^{700\text{ nm}} \\ &\approx 0.002 \text{ MeV/cm} \end{aligned}$$

which is negligible compared to the energy losses through ionization and radiation of 1.85 MeV/cm assuming a water density of  $\rho_{\text{H}_2\text{O}} = 1 \text{ g/cm}^3$ .

In order to estimate the number of emitted photons Equation (2.29) has to be divided by  $E = \hbar\omega$  to get the number of photons emitted per unit frequency and unit distance traveled. Translating this equation again into units of wavelength results in

$$\frac{dN}{d\lambda dx} = \frac{2\pi z^2 \alpha}{\lambda^2} \cdot \left(1 - \frac{1}{\beta^2 n(\lambda)^2}\right) \quad (2.31)$$

where  $\alpha$  is the fine structure constant. Thus, the number of emitted Cherenkov photons in the region [250 nm, 700 nm] for a 2 MeV electron which travels about 1 cm in water with a constant refractive index of  $n = 1.333$  is  $\sim 500$  photons. This is a very optimistic estimate as the slowing down of the electron is completely neglected and it is unlikely that the 2.22 MeV  $\gamma$ -ray transfers its energy only to a single electron<sup>12</sup>. However, already this optimistic scenario shows that the produced signals are very faint. Further, Equation (2.31) has a  $\lambda^{-2}$  dependence which means that most of the Cherenkov radiation is emitted in the ultra-violet and blue regime. Both need to be taken into account in the design of an appropriate water Cherenkov neutron veto which is discussed along with the rest of the experiment in the subsequent section.

<sup>11</sup>Estimate though Equation (2.27) using an kinetic electron energy of 2 MeV.

<sup>12</sup>Simulation with Geant4 [Ago+03] discussed in Chapter 7 produce on average  $\sim 260$  photons within the same optical range.



## 2.4 THE XENONNT EXPERIMENT

The XENONnT experiment is already the 4<sup>th</sup> generation of dual-phase xenon TPC within the XENON Dark Matter project. The experiment is located in the Hall B of the INFN Laboratori Nazionali del Gran Sasso (LNGS) underground laboratory in central Italy. It was designed as a fast upgrade of its predecessor XENON1T [Apr+17c], reusing most of its service infrastructure which was augmented by new systems required to achieve the design goals of XENONnT [Apr+20b]. The experiment itself can be roughly divided into a three-story service building and a 700 t water tank which are both shown in Figure 2.13. The service building accommodates most of the cryogenic infrastructure on the ground and second floor as well as the readout and data processing electronics of the experiment on the first floor. It is connected through a service pipe with the water tank which hosts the three nested detectors of the experiment, two water Cherenkov veto systems, and a dual-phase xenon TPC.

After a short delay due to the worldwide Covid pandemic, the construction of the experiment was completed in late summer 2020, and the filling of the water tank as well as the TPC seven weeks later in fall 2020. In the following the most relevant<sup>13</sup> components of the experiment are explained.

### 2.4.1 XENONnT TPC

The heart of the experiment is a xenon dual-phase time projection chamber which is shown in Figure 2.14. An extended version of the inner vessel of the XENON1T stainless steel cryostat hosts the TPC and is filled with about 8.6 t of xenon at a temperature of  $-96^{\circ}\text{C}$ . The TPC has an almost cylindrical shape with a height of 1.49 m and a diameter of 1.33 m (at room temperature). It uses an improved design compared to XENON1T which uses less material<sup>14</sup>, to reduce background radiation from materials, and to reduce the outgassing to achieve a higher xenon purity. In addition, an extensive material screening campaign was performed to reduce the radioactive contamination from the materials used [Apr+22b]. The main body of the TPC is made out of 48 thin polytetrafluoroethylene (PTFE) panels with a thickness of about 3 mm which are held by 24 PTFE pillars. The active target volume of the detector contains about 5.4 t of liquid xenon which are constantly monitored by two PMT arrays at the top and bottom of the cylinder. The two arrays contain in total 494 Hamamatsu R11410-21 3 inch PMTs which were specifically developed for the XENON1T/nT experiments [Ant+21]. The bottom array contains 241 PMTs while the top array has 253 PMTs. Both arrays are arranged in a hexagonal dense packing which is supported by a single oxygen-free high-conductivity copper disk. The space between neighboring PMTs is filled through a monolithic PTFE disk with holes at the PMT positions. All PTFE surfaces were polished to maximize the reflectivity for VUV photons and thus to increase the light collection efficiency of the detector.

<sup>13</sup>Most relevant for the presented work.

<sup>14</sup>Less material in relative terms. The absolute amount of material is of course large due to the larger size of the detector.





Figure 2.13. – Photograph of the XENONnT experiment. The poster mounted to the water tank shows a CAD rendering of its interior (Credit Henning Schulze).

The electric drift and extraction field are generated by two stacks of electrode grids mounted at the top and bottom of the cylindrical volume. The bottom stack is made out of two individual electrodes, a bottom screening electrode, and a cathode mounted 55 mm above the bottom screening electrode. The latter serves as protection of the bottom PMT array against the high voltage of the cathode which was designed to be operated at  $-30.0$  kV. The gate is mounted in the top electrode grid stack at a distance of 148.6 cm above the cathode and 0.8 cm below the anode electrode. Gate and anode are designed to be operated at a high voltage of  $-1.0$  kV and 6.5 kV respectively. The top screening electrode, which serves a similar purpose as the bottom screening electrode, is mounted 2.8 cm above the anode. Both screening electrodes are designed to be operated at a voltage close two the average voltage of the PMTs which is about  $-1.5$  kV.

All five electrodes are made from parallel annealed stainless steel wires mounted on stainless steel frames. The wires were chosen based on their smooth surface and have a diameter of  $216 \mu\text{m}$ , except for the cathode which uses a larger diameter of  $304 \mu\text{m}$ . The wire pitch was optimized to ensure uniform fields by maximizing their optical transparency as light from both, S1 as well as S2 signals, have to pass at least two of the electrodes before reaching the PMT arrays. The bottom screening electrode and the cathode have a wire pitch of 7.5 mm. Top screening electrode, anode, and gate have a smaller pitch of 5 mm, and the anode wires are offset by half a pitch with respect to the gate wires. The optical transparency of the individual electrodes varies between 95.7% and 97.0%. To prevent wire sagging under electric fields, the gate and

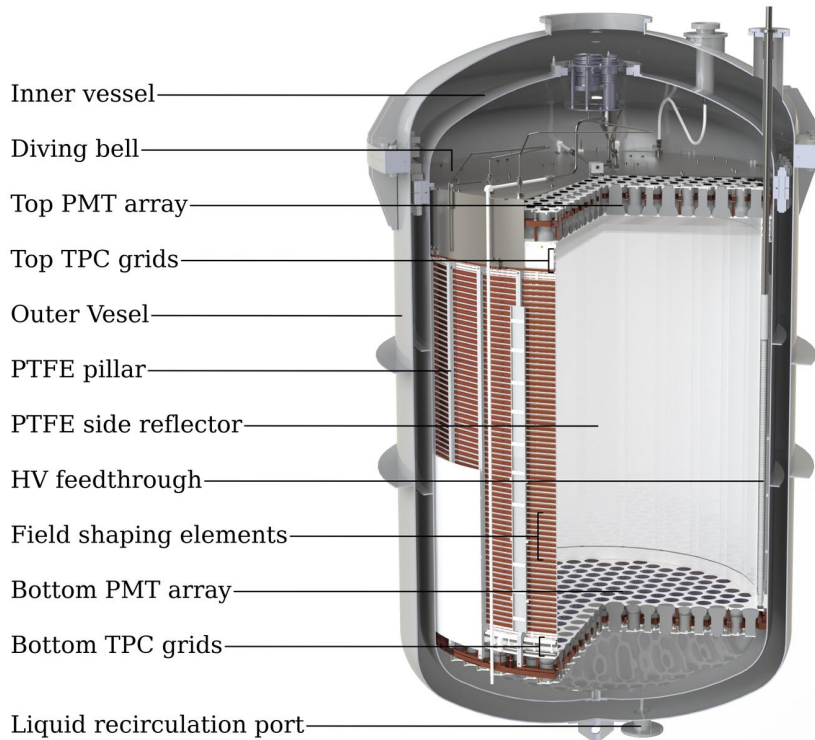


Figure 2.14. – Rendering of the XENONnT TPC. (Credit XENON collaboration)

anode are reinforced by two and four additional  $304\ \mu\text{m}$  wires, respectively, which are mounted in a perpendicular direction using a higher wire tension<sup>15</sup>. Figure B.3 in the appendix on page 239 shows a photograph of the anode wires.

To ensure a uniform drift field, two alternating types of field shaping rings are mounted outside of the PTFE cylinder to the pillar structure. The field shaping rings are made from oxygen-free high conductivity copper and alternate between 2 mm thin copper wire, and 15 mm tall and 5 mm thick rings. Wire and rings are connected to the cathode through two voltage divider networks which define together with the applied voltage the potential gradient along the height of the TPC. In contrast to XENON1T the uppermost field shaping wire and ring are not connected to the gate which allows an independent adjustment of its voltage. The design voltage for the uppermost ring is  $-0.95\ \text{kV}$ .

After a few weeks of commissioning an electrical short between the bottom screening and cathode electrode prevented the operation of the detector at its design fields. Therefore, during the first science run the voltage of the cathode and bottom screening electrode was limited to prevent any additional damage. The cathode was biased with a voltage of  $-2.75\ \text{kV}$  and the bottom screening electrode was set to  $-2.7\ \text{kV}$ . To increase the strength of the drift field, it was decided to operate the gate at a slightly positive voltage of  $0.3\ \text{kV}$ . Thus, also the upper end of the field shaping rings is biased with a positive voltage of  $0.65\ \text{kV}$ . The anode was limited to a voltage of  $4.9\ \text{kV}$ , below

<sup>15</sup>These wires have some implications for the reconstructed S2 signal shape, and will be referred in the further discussion either as "perpendicular wires" or "transverse wires".

its design voltage, to mitigate the localized and burst-like emission of single to few electron S2 signals which can exceed the average S2 background by up to four orders of magnitude. Figure B.4 on page 240 in the appendix shows this behavior resulting in an so-called "hot-spot". In addition, the top screening electrode was biased with a positive voltage of 0.65 kV, because light emission from the electrodes was observed if the potential difference between the anode and top screening electrode exceeds 5.0 kV. The limitation of the fields results in an about 10 times lower drift field of  $23_{-0.3}^{+0.4}$  V/cm than designed, with a maximum drift time of about 2.2 ms over the full drift range of the TPC. The strength of the extraction field varies with xy-position due to the attraction of wires between gate and anode, and has a strength of about 2.9 kV/cm<sup>16</sup>.

The liquid level between the gate and anode is set via the gaseous xenon pressure inside the diving bell. When the pressure inside the bell rises the liquid level decreases until it reaches the opening of a 1/4-inch exhaust pipe. The height of the exhaust can be varied by a linear motion feedthrough and was constantly set to a liquid level of 0.51 cm above the gate electrode<sup>17</sup>. The height of the liquid level is read out through four plate capacitor level meter following the design used in XENON1T [Apr+17c; Gei18]. Each level meter has a dynamic range of 10 mm with a precision of about 0.2 mm.

#### 2.4.2 Auxiliary infrastructure: xenon purity

The relatively low drift field of XENONnT impacts the signal reconstruction of the detector in various ways. A longer drift time leads to larger losses of electrons due to attachments to negative impurities as discussed in Section 2.1.2. To reduce the concentration of electron-negative impurities XENONnT is equipped with a gaseous and liquid purification system. The gaseous purification system, located on the third floor of the service building, is mostly unchanged compared to XENON1T. It uses high-temperature rare gas purifiers which clean the xenon with a flow of 2 slpm [Apr+17c]. The new custom-developed liquid xenon purification system drains liquid directly from the bottom of the TPC cryostat and can reach much higher flows of 58 slpm [Pla+22]. In addition, the system is equipped with a new electron lifetime purity monitor which drifts electrons, produced by a VUV light pulse, in a small 20 cm liquid xenon drift chamber similar as in [Car+90]. The monitor was used to measure the initial electron lifetime of the experiment after filling the TPC which was evaluated to be  $\tau_e \approx 71 \mu\text{s}$ . During the science run the purity monitor was augmented by additional measurements using different calibration sources diluted in LXe<sup>18</sup>. Throughout

<sup>16</sup>The electrode issues encountered during the commissioning phase of XENONnT, motivates once more the necessity to develop new systems for a higher quality control of wire electrodes. Further, it shows that the origin of light emission from the electrodes in high-field environments needs to be studied and better understood. These are both aspects of the new wire inspection set-up which was developed together with the PRISMA detector lab throughout this thesis and is discussed as an extra topic in the supplemental material on 213 and following.

<sup>17</sup>This assumes a flat electrode without attraction of the wires due to electric fields.

<sup>18</sup>This is important as the monitor was designed for a maximum electron lifetime, below the lifetime achieved in first science run (SR0). Thus, the monitor was only used to compare the relative trend of the electron lifetime.

the science run an electron lifetime of about 15 ms was reached, leading to an average loss of free electrons of ~15 % for energy deposits just above the cathode of the TPC. In comparison, the electron lifetime in XENON1T was much smaller ~0.65 ms with a much higher electron loss of ~70 % [Apr+18].

Besides impacting the electron loss due to attachments to impurities, a lower drift field also impacts the photon and electron yields as discussed in Section 2.1.2. In general, a smaller drift field leads to an increased overlap between ER and NR band and therefore to a smaller signal-to-noise ratio for the WIMP dark matter search. While ER background originating from detector materials can be reduced by a careful material screening and fiducialization of the detector volume, ER background induced by decays of radioactive isotopes diluted in LXe has to be mitigated by other means. In XENON1T the two most dominant ER background components in the energy range relevant for the WIMP dark matter search, were given by the  $\beta$ -decays of  $^{85}\text{Kr}$  and  $^{214}\text{Pb}$  which is a daughter of  $^{222}\text{Rn}$ <sup>19</sup>. Both isotopes are rare elements and can therefore be not removed by the same means as electronegative impurities. Already in XENON1T a cryogenic krypton distillation column was used to remove  $^{nat}\text{Kr}$  from xenon [Apr+17c; Apr+17a]. The distillation is based on the fact that  $^{nat}\text{Kr}$  has a higher vapor pressure as xenon at the same given temperature and therefore can be extracted in gaseous form from the liquid. For XENONnT the entire xenon budget was first distilled by the krypton column "offline" before filling the TPC, and then for another three weeks "online" while already operating the detector. In this way it was possible to reduce the  $^{nat}\text{Kr}$  to a factor  $^{nat}\text{Kr}/^{nat}\text{Xe}$  of ~50 ppq(mol/mol) which is about a factor two better than the design goal.

To reduce the content of  $^{222}\text{Rn}$ , the most dominant ER background component, a new cryogenic radon distillation column was added to the purification system in XENONnT [Mur+22]. The cryogenic distillation of  $^{nat}\text{Rn}$  is based on the same idea as for  $^{nat}\text{Kr}$ , but due to its lower vapor pressure gaseous xenon is extracted from a radon-enriched liquid xenon mixture instead. In contrast to krypton distillation, radon distillation must be operated online as  $^{222}\text{Rn}$  is being constantly emanated from the surfaces of the detector materials and cryogenic piping. The column can be operated in two different modes, either extracting just gaseous xenon or gas and liquid xenon from the detector. During the first science run the column was only operated in gaseous mode reducing the initial radon content from  $(4.2^{+0.5}_{-1.0}) \mu\text{Bq}/\text{kg}$  by a factor of about two to  $(1.77 \pm 0.01) \mu\text{Bq}/\text{kg}$ . In comparison XENON1T had a radon concentration of ~13  $\mu\text{Bq}/\text{kg}$ . All three purification systems play an important role in the first WIMP dark matter results as discussed in Chapter 8.

### 2.4.3 Water Cherenkov Muon- and Neutron-Veto

As discussed in Section 2.1, neutrons with energy in the range of a few MeV are an excellent tool to calibrate the NR response of the TPC, which is also the expected signal signature of WIMPs. However, this also means that NR signals induced by neutrons pose a great risk for the sensitivity of the experiment when originating from back-

<sup>19</sup>A spectrum of the ER background components is shown in the appendix in Figure B.58 on page 283.



ground radiation. Especially, after mitigating the total ER background by the novel purification techniques discussed in the previous section, neutrons become the most dominant background which put the science goals of XENONnT at risk [Apr+20b]. There are mainly two sources for neutrons as a background, cosmogenic neutrons, which have been already mostly mitigated in XENON1T, and neutrons emitted from materials through spontaneous fission or ( $\alpha$ , n)-reactions.

Cosmogenic neutrons are either induced by muons through the spallation of nuclei or as a product of an electromagnetic and hadronic shower [Gei18]. To reduce the cosmogenic neutron rate, XENON1T/nT are located at the LNGS underground laboratory at a depth of 3600 m water equivalent [Apr+17c]. This reduces the muon flux by about 6 orders of magnitudes compared to the surface level, to a flux of  $(3.31 \pm 0.03) 10^{-8} \mu/\text{cm}^2\text{s}$  for muons with an energy of  $\sim 260$  GeV [Gei18; Apr+17c]. To reduce the cosmogenic neutron background even further, the TPC is located at the center of a 700 t water tank which has a diameter and height of about 10 m. The water tank serves on the one hand as a passive shield against radiation and neutrons emitted by the surrounding rock of the experimental cavern, but also contains two active veto systems. The outer part of the water tank serves as water Cherenkov muon veto (MV) which was already commissioned during XENON1T [Apr+14; Apr+17c]. It is equipped with 84 Hamamatsu R5912ASSY PMTs which are arranged in 5 rings mounted to the inner walls of the water tank. To improve the light collection efficiency the inner surfaces of the water tank are covered with a reflective foil (3M DF2000MA) which has a reflectivity larger than 99 % at wavelengths between 400 nm and 1000 nm [Gei18; Apr+14]. A CAD rendering of the interior of the water tank is shown in Figure 2.15. The optical response of the MV can be calibrated through optical fibers and four diffuser balls mounted to the support structure at the center of the water tank.

To suppress also the remaining neutron background due to detector materials several additional measures have been taken in XENONnT. As mentioned in Section 2.4.1, all TPC materials were screened for their radio-purity [Apr+22b] and the overall material's mass of the detector parts were reduced. In addition, the central inner part of the water tank was augmented with a new water Cherenkov NV [Apr+20b] which is also shown in Figure 2.15. To ensure a high sensitivity for faint Cherenkov signals all aspects of the neutron veto were designed to maximize its light collection efficiency. The NV is equipped with 120 8 inch Hamamatsu R5912-100-10 PMTs which are distributed over 20 columns surrounding the central TPC cryostat. The PMTs have a high radio-purity, and a high quantum efficiency<sup>20</sup> for Cherenkov light which reaches about 40 % at a wavelength of 350 nm. Each PMT column contains six PMTs equally distributed over the full height of the column which is 3 m. The columns are arranged in an octagonal shape at a distance of about 1 m from the surface of the TPC cryostat and also serve as a support structure for the outer NV walls. The walls of the neutron veto are made out of 1.5 mm thick expanded PTFE (ePTFE) which optical separates neutron and muon veto. ePTFE has a high reflectivity of greater than 99 % above a wavelength of 280 nm. To ensure a high light collection efficiency also other large surfaces inside the neutron veto are covered with ePTFE. The outer reflector has several smaller openings to accommodate service pipes to the TPC as well as openings in the

<sup>20</sup>See also Figure 7.4 on page 7.4.

ceiling of the NV to lower different calibration tools inside the NV. The openings are also shown in Figure 2.15. Among them are two holes that allow to lower a pulsed deuterium-deuterium neutron-generator into the NV. One of the two holes also contains the "I-belt" system, shown in blue in Figure 2.15, which can be used to lower photon-neutron sources like  $^{88}\text{Y}$ -Be mounted in a tungsten shield for calibrations with low energetic neutrons. Two 1 inch stainless steel pipes called "u-tubes", can be used to place calibration sources with a smaller diameter like AmBe next to the TPC cryostat.

A variety of optical calibration tools allow to monitor the PMT response, as well as the optical properties of the neutron veto, like the reflectivity of the ePTFE and water transparency. Each PMT is equipped with an optical fiber to calibrate their gains and to monitor their stability. In addition, four diffuser balls of a smaller diameter than the MV ones<sup>21</sup>, are mounted equally spaced around the TPC cryostat<sup>22</sup>. Each diffuser ball can be operated individually with a pico-second laser with a wavelength of  $(448.3 \pm 2.7)$  nm. Complementary to the duffuser balls, an additional reflectivity monitor is installed to the outer reflector walls of the NV. It is equipped with two pairs of quartz fiber which point to the ceiling and floor of the NV. They can be used to inject a mono-chromatic laser beam with a wavelength of 375 nm to measure the optical stability of the detector.

To ensure high optical transparency of the water inside the water tank, deionized water is used. The water is constantly circulated through a purification plant which was already used in XENON1T, with a throughput of  $2.2 \text{ m}^3/\text{h}$ , leading to a water purity with a residual conductivity of  $0.07 \text{ mS/cm}$  [Apr+17c].

As discussed in Section 2.2.2 and Section 2.3, neutrons can elastically back-scatter on xenon and escape the TPC with a single scatter NR which is indistinguishable from WIMPs. These neutrons are afterwards captured on hydrogen releasing a  $2.2 \text{ MeV}$   $\gamma$ -ray which is eventually recorded as a Cherenkov signal. To further enhance the tagging efficiency of the NV, the water of the water tank is ultimately going to be doped with gadolinium sulfate octahydrate ( $\text{Gd}_2(\text{SO}_4)_3 \cdot 8(\text{H}_2\text{O})$ ). Gd has the highest neutron capture cross section among all stable elements and releases a cascade of multiple  $\gamma$ -rays with a total energy of about  $8 \text{ MeV}$ . The technology needed for Gd doping of water was developed in the EGADS project [Mar+20] which served as a demonstrator for the Super-Kamiokande Gd-upgrade [Abe+22]. One aspect of the technology demonstration of the Gd doping is the purification of the water from residues which may impact the optical transparency, without filtering out any Gd. Thus, also XNEONnT was extended by an additional purification plant which is based on the technology developed in EGADS. While during the first science run of XENONnT, which is discussed in this presented work, the NV was operated with pure water only, the purification plant was already operated and commissioned in a closed circuit without water tank. The procurement of the gadolinium sulfate octahydrate powder is discussed in Chapter 4.

<sup>21</sup>The diffuser balls are made in the same way as the MV ones, but only using the inner quartz sphere. See also [Gei18].

<sup>22</sup>One diffuser ball is placed slightly shifted due to technical constraints.

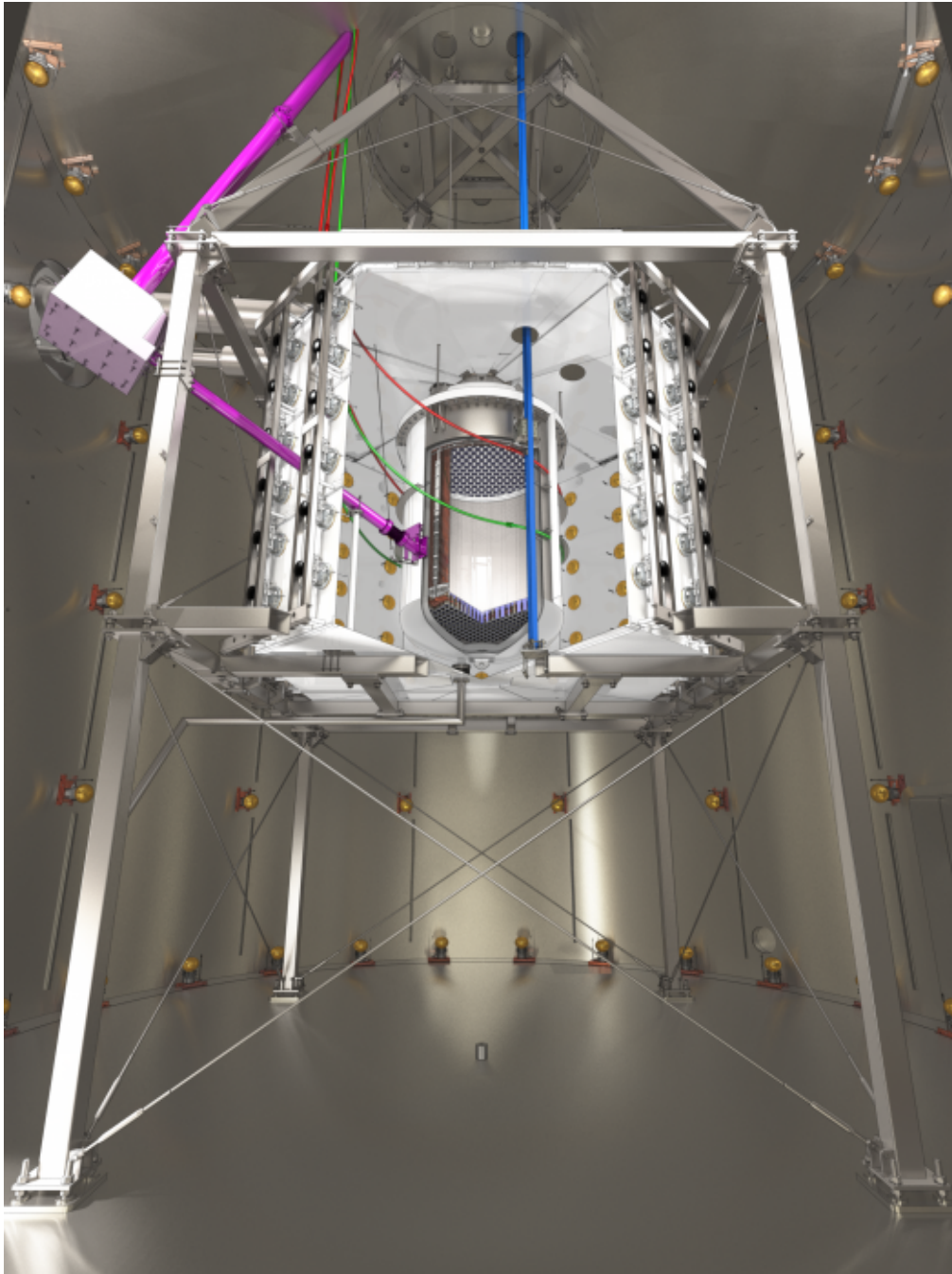


Figure 2.15. – Rendered CAD image of the XENONnT experiment. The outermost gray structure is the inner wall of the water tank which is used as active water Cherenkov muon veto. The white structure within the central metal support structure is the new water Cherenkov neutron veto. In the center of the image a cut view of the TPC cryostat shows the interior of the TPC. Golden spheres in the muon and neutron-veto indicate the individual PMTs. The blue belt is the I-belt system which can be used to lower calibration source next to the TPC cryostat. The purple pipe system corresponds to the neutron back-scatter calibration set-up. It uses a pulsed deuterium-deuterium neutron generator with a  $D_2O$  back-scatter target hosted in the white PTFE box outside of the neutron veto. The green and red pipes are the “U-tube” calibration system. (Credit Xenon Collaboration)

#### 2.4.4 Auxiliary infrastructure: Data acquisition

All three sub-detectors, TPC, MV and NV, are integrated into a single data acquisition (DAQ) system which is shown in Figure 2.16. The following discussion is based on [Apr+22d].

Signals acquired by TPC PMTs are first amplified by dual gain amplifiers with an amplification factor of  $\times 0.5$  and  $\times 10$  respectively. The output of all high gain channels, as well as the low gain channels of the top PMT array, are digitized by CAEN V1724 analog-to-digital converter (ADC) which have a sampling rate of 100 MHz and a 14 bit resolution with a 2.25 V dynamic range. Each channel is recorded in a "triggerless" mode, recording PMT signals independently of any other channel if a channel-specific threshold is crossed. Each recorded signal includes a short pre-trigger region before the actual main pulse as illustrated in Figure 3.7 on page 62. Signals recorded by the high gain channels are used for the regular signal reconstruction of the TPC discussed in Chapter 3, while the low gain channels can be used to improve the detector response for high energetic signals like in the search for neutrinoless double  $\beta$ -decays. The low gain signals of the bottom PMT array serve as an input for the high energy veto (HEV) which vetos large S2 signals and their respective single electron trains<sup>23</sup>. The veto is issued based on the summed waveform of the low-gain channels and varies in length depending on the integrated S2 size.

The neutron-veto PMT signals are directly digitized by eight CAEN V1730 ADCs without any further amplification. The V1730 digitizer has a higher sampling rate of 500 MHz with a dynamic range of 2 V and a 14 bit resolution. The higher sampling rate was chosen to improve the timing resolution of the individual Cherenkov photon signals which is supposed to enhance the discrimination between neutron capture and background signals based on their respective hit-pattern. Just like for the TPC the individual channels are read out in a triggerless mode. However, due to the relatively high dark count rate of the NV PMTs an additional software trigger was developed as part of the presented work. It reduces the overall data rate of the NV and is discussed in Chapter 3.

The MV PMT signals are also recorded by CAEN V1724 digitizers. In contrast to the other two detectors the MV uses a triggered read-out and only PMT signals in a  $5.12 \mu\text{s}$  window around a trigger are stored. A trigger is formed if at least five PMTs measure a signal of 1 pe within a 300 ns window.

Each digitizer of the three detectors is eventually read out by a single custom software package called *redax* [Mas+22] which converts the digitized PMT signals into a data format that can be handled by the *strax* processing framework discussed in Chapter 3. Each signal is stored as an independent time interval which carries information in form of the digital signal waveform as well as information about the interval start and length. All signals are directly processed up to the event-level by one of the three event building machines, located in the service building, before the data is trans-

<sup>23</sup>"Single electron trains" are time intervals of an enhanced background rate of small S2 signals. These trains are caused by photoionization of electronegative impurities or detector parts such as the cathode electrode and usually follow large S2 signals in which a lot of VUV light is produced.



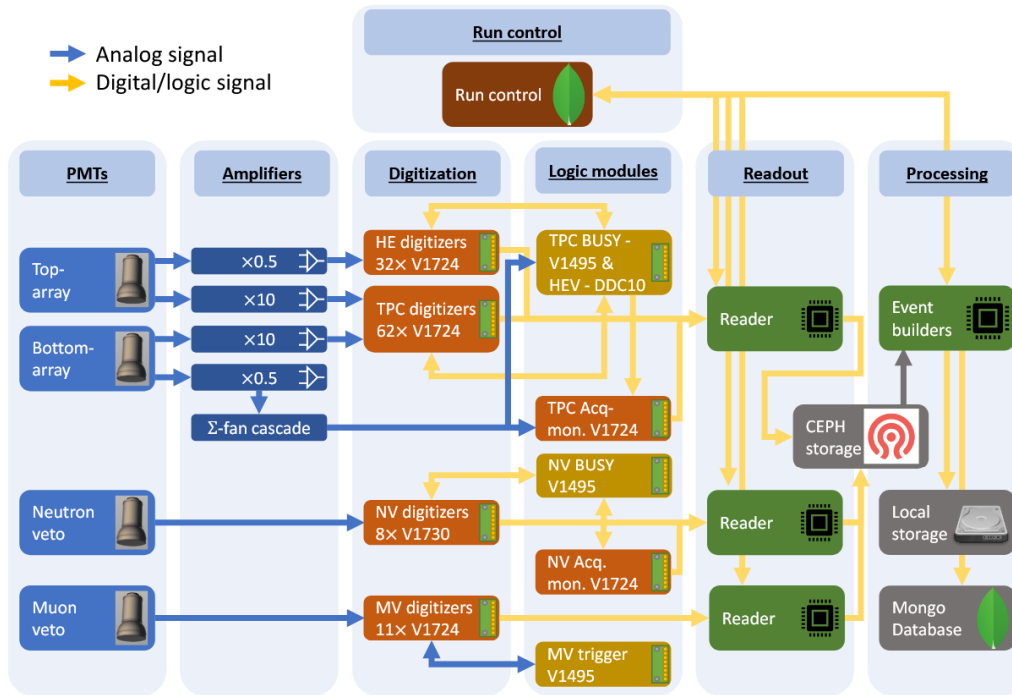


Figure 2.16. – Schematic of the XENONnT DAQ system. The schemataic was taken from [Apr+22d].

ferred to one of XENONnT’s external storage facilities. To synchronize the timestamps between all three detectors a 50 MHz clock generated by a CAEN DT4700 clock generator is distributed to three DAQ subsystems. In addition, each detector is supplied with a 0.1 Hz trigger signal from a custom-made GPS timing module which also provides GPS time stamps along with the trigger. This also allows synchronization of the detector with other experiments.

To record the logic signals issued by the global pointing system (GPS) module or the HEV the DAQ subsystems of the TPC and NV are equipped with a dedicated V1724 digitizer which serves as an acquisition monitor. The MV does not have its own acquisition monitor, but some spare channels of the TPC are used instead.

Although for most of SR0 all three detectors were operated in a synchronized "linked mode", TPC signals, recorded by the regular digitizer, were delayed by a fixed time offset with respect to the MV and NV due to the HEV. As the HEV veto decision is based on S2 signals which can have widths of a few to several  $10\ \mu\text{s}$ , depending on the depth of the interaction, TPC signals must be delayed by a similar amount of time. The delay was measured with the 0.1 Hz GPS trigger signal and found to be 10 153 ns on average between NV and TPC, and 5283 ns for MV and TPC. These two delays must be kept in mind when developing the veto signals for the NV and MV which is discussed in Chapter 3. For data taken after SR0 the corresponding delay is already subtracted during the data readout with `redax`.

The digitizer channels of the three detector systems are organized in channel numbers using different integer ranges. The PMT channels of TPC low gain signals, MV and NV are organized in the ranges  $[0, 493]$ ,  $[1000, 1083]$ , and  $[2000, 2119]$  respectively

and the acquisition monitor channels of the three subsystems in [790, 807], [808, 815] and [1084, 1087] respectively.



---

## XENONNT ANALYSIS SOFTWARE

Not only the experimental apparatus was upgraded and augmented for XENONnT, but also its analysis software. The former processing software called Processor for Analyzing XENON (PAX) was replaced by two new software packages called Streaming analysis for xenon experiments (STRAX) and Streaming analysis for XENON(nT) (STRAXEN). STRAX is a generalized package that provides different base tools and a data processing framework to set up an experiment-specific processing chain for time-sorted peak-like signals. The XENONnT-specific processing structure is hosted by the STRAXEN package. One of STRAX's key features are so-called plugins which are used to divide the data processing logic into small steps, e.g., peak finding and event building. As part of this thesis, the complete plugin structure for the neutron- and muon-veto has been developed. In addition, as one of the current lead maintainers for STRAX and STRAXEN, many technical changes and general tools have been implemented, and the code of other analysts has been tested and reviewed before implementation.

Besides these two packages, XENONnT uses a third, private package called Cuts for analysing Xenon data (CUTAX) to develop and manage data quality cuts. During this thesis also different cuts and other general features were added to CUTAX.

In this chapter, the most relevant terminology and features of the three packages are explained, as well as plugins developed during this thesis, are discussed. The chapter is organized as follows: In Section 3.1 a general overview of STRAX, STRAXEN, and CUTAX is given. In addition, the most important key concept and analysis terms are defined. The subsequent Section 3.2 discusses the computation of the most important TPC signal types called hits, peaks, and events, which refer to the reconstructed form of PMT signals, S1 and S2 signals, and paired S1 and S2 signals, respectively. Section 3.3 follows with a review about the working principle of the neutron- and muon-veto plugins, their underlying processing logic, and data types, before Section 3.4 discusses their overall performance from a computing point of view. The last Section 3.5 deals with a selection of examples for auxiliary tools and modifications which were developed and implemented during this work.

### 3.1 STRAX AND STRAXEN

STRAX and STRAXEN were developed in a joint effort of multiple institutes within the XENON collaboration. The entire software stack is an open source project and available via the version control provider github<sup>1</sup>.

The software itself is written in python with an emphasis on the widely used Scientific python stack including packages like numpy<sup>2</sup>, scipy<sup>3</sup>, pandas<sup>4</sup> and matplotlib<sup>5</sup>. In addition, both projects depend heavily on the novel package numba<sup>6</sup> which enables the "just-in-time" compilation of pure-python and numpy-based code into machine code using the LLVM<sup>7</sup> compiler infrastructure. In this way, performances similar to C/C++ can be achieved.

Both STRAX and STRAXEN have an extensive testing suite with a testing coverage greater than 90%. The testing is based on unittest<sup>8</sup> and randomized tests using the hypothesis<sup>9</sup> package. In addition, both packages provide a short user and developer documentation which is built using the sphinx<sup>10</sup> documentation tool kit.

The main idea of having a software kit that allows to process time sorted peak-like data is based on the fact that the DAQ of XENONnT follows a "triggerless" approach. This means that each single PMT channel only uses its own threshold-trigger to record a signal. Hence, XENONnT requires an analysis software that is capable of analyzing data as a continuous stream of time intervals, rather than pre-defined countable events as would be the case for a DAQ which uses a hardware trigger. The package is not only limited to LXe TPCs but can be applied to any kind of experiment which records time-sorted peak-like data. Figure 3.1 illustrates the main components provided by the STRAX software package. In general, the package can be subdivided into six different blocks. Each block is briefly explained in the following subsections.

#### 3.1.1 Context, Corrections, Lineage and Storage

The central heart of any analysis is the so-called "context"<sup>11</sup>. It is the main entry point for any analysis to access and process data. It can be seen as a container for all settings required to create and manage the experimental data. Depending on the analysis task, different contexts may be used. This makes it easy to share analysis-specific settings

---

<sup>1</sup>The two projects can be found via <https://github.com/AxFoundation/strax> and <https://github.com/XENONnT/straxen>.

<sup>2</sup><https://numpy.org/>

<sup>3</sup><https://scipy.org/>

<sup>4</sup><https://pandas.pydata.org/>

<sup>5</sup><https://matplotlib.org/>

<sup>6</sup><https://numba.pydata.org/>

<sup>7</sup><https://llvm.org/>

<sup>8</sup><https://docs.python.org/3/library/unittest.html>

<sup>9</sup><https://hypothesis.readthedocs.io/en/latest/>

<sup>10</sup><https://www.sphinx-doc.org/en/master/>

<sup>11</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/context.py#L87>

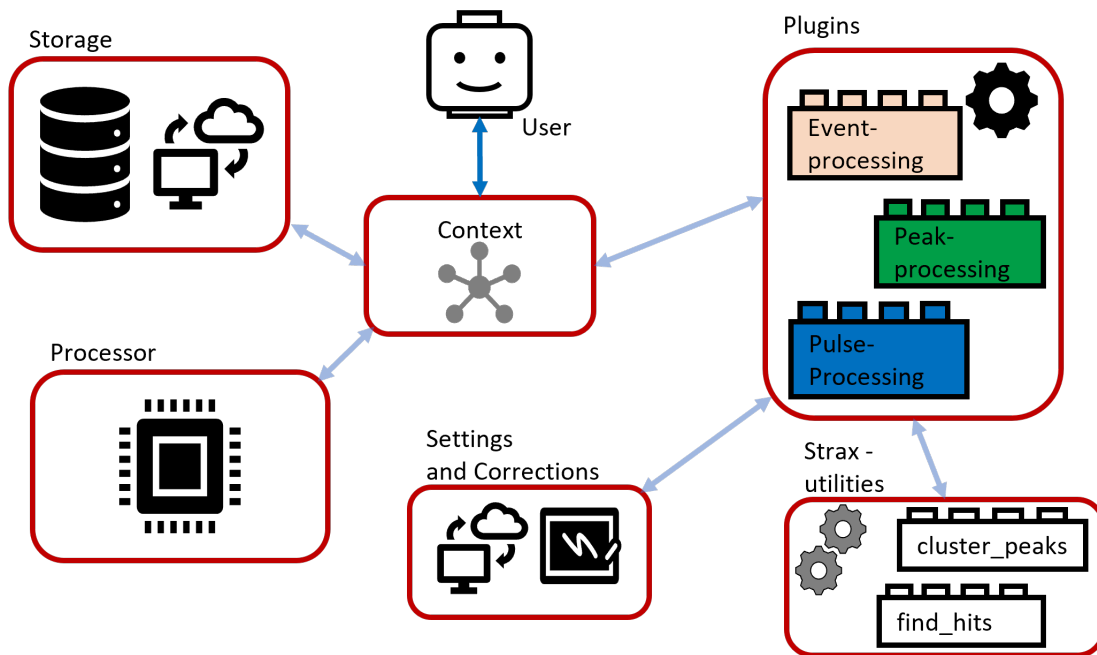


Figure 3.1. – Overview of the different components provided by the STRAX package. The arrows indicate the dependency of the context with the other sub components of STRAX. The user interacts primarily with the context, whereas all other dependencies and interactions are hidden to the user. The colored building blocks represent the stackable plugin structure. The smaller building blocks in Strax-utilities refers to the numba optimized code.

and configurations among different analysts.

The context is designed in a way that it should not be seen as a static object, but as a default set-up of how data can be processed, allowing the user to adopt the data processing during the analysis. These changes can include the reconfiguration of certain settings, e.g., changing the threshold for identifying PMT signals, entail the registration of entirely new plugins to process some new type of data, or exchange an already existing plugin, for example, a new peak classification.

Some of the settings are more dynamic and change many times throughout a science run. This is especially true for settings like the electron lifetime, or light collection efficiency which correct data for detector-dependent effects. To track these changes, STRAX uses Correction Management Tool (CMT), which stores corrections as collections in a Mongo<sup>12</sup> database<sup>13</sup>.

Given all the flexibility to change settings and plugins STRAX requires a way to determine whether certain data have already been produced and can be loaded from a given storage or if the data needs still to be processed. For this reason, STRAX associates each piece of data with a data key<sup>14</sup> which is made of a runid, data type,

<sup>12</sup><https://www.mongodb.com/de-de>

<sup>13</sup>For more information please also refer to the documentation <https://straxen.readthedocs.io/en/latest/cmt.html>

<sup>14</sup><https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/storage/common.py#L22>

and lineage-hash as illustrated in Figure 3.2. The runid is a human-readable way to identify which time interval of data is requested by the user. The data type determines which type of data should be loaded, e.g., peaks or events. The lineage hash encodes the complete processing "lineage" from a human-readable python dictionary, which contains all information about the plugins and settings required to make the requested data type, into a 10-digit hash similarly to commits in version control software.

Based on a data key request by the user, the context iterates over the different storage frontends to check whether the request data has been already processed. For this purpose, the data key is converted by the corresponding frontend<sup>15</sup> in a backend key which tells the underlying backend<sup>16</sup> where the data should be loaded from, or stored to in case new data is being created. The backend key can be as simple as the path to some file that stores the requested data on disk. The default XENONnT STRAXEN context provides several different storage front- and backends to access data stored on multiple local sides on XENONnT's computing cluster, as well as data stored in XENONnT's rundatabase which can be downloaded at any time on users request. In addition, a dedicated online monitor<sup>17</sup> frontend/backend pushes frequently plots to XENONnT's internal slack<sup>18</sup> channel<sup>19</sup>.

In case the requested data has not been stored, the lineage tells STRAX which data types, and therefore which plugins lower in the hierarchy are required to make the requested data and which of those data types are already stored in one of the storages. In this way a change in the high-level processing, e.g., on the event level does not trigger a reprocessing of the data from scratch and high-level data can be processed on the fly when requested by the user.

The XENONnT analysis context can be distinguished into three main categories, an ONLINE context, a few OFFLINE contexts, and a simulation context. The ONLINE context, is used to allow for live processing and monitoring of the currently taken data. It uses an extrapolation or the latest correction settings to correct the online data for different detector-specific effects. The online availability of the data has a delay of a few seconds to a few minutes depending on the selected chunk-size discussed in the subsequent section.

With respect to the ONLINE context, the OFFLINE context usually has a delay in the data availability of 1-2 weeks. The reason is the validity of the corrections used which depends on the frequency of the calibrations performed. For instance the PMT gain is only measured once per week.

The last context category is the simulation context, which is typically used together with the full chain waveform simulation framework. Compared to the other two contexts the simulation context, provides several extra options which are used during data creation.

---

<sup>15</sup><https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/storage/common.py#L105>

<sup>16</sup><https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/storage/common.py#L393>

<sup>17</sup>[https://straxen.readthedocs.io/en/latest/online\\_monitor.html](https://straxen.readthedocs.io/en/latest/online_monitor.html)

<sup>18</sup><https://slack.com/>

<sup>19</sup>For more information please refer to the Ph.D. thesis of Joran Angevaare [Ang23]

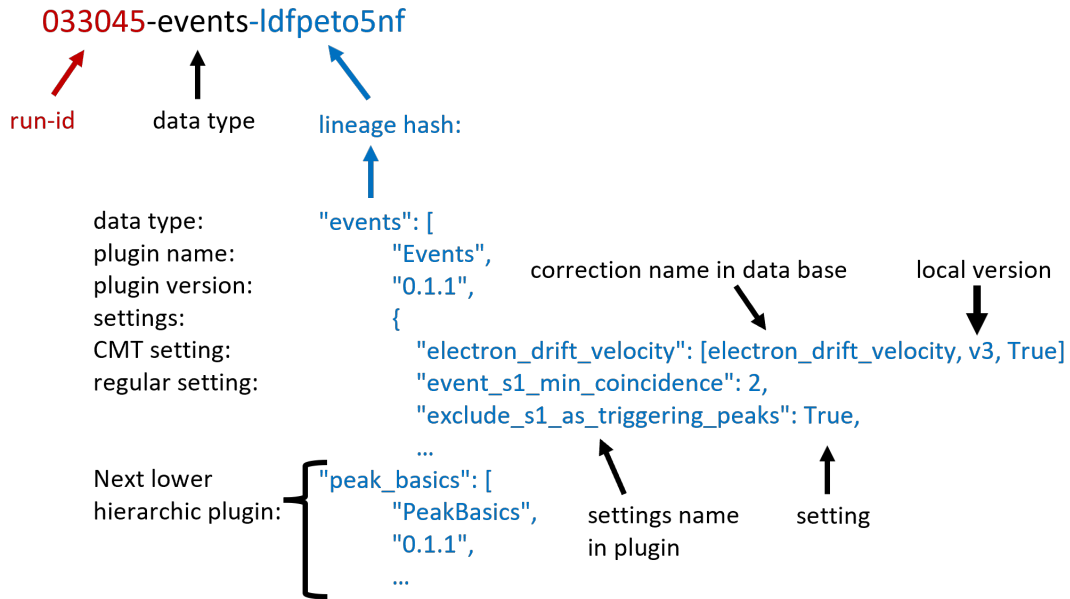


Figure 3.2. – Illustration of a straxen data key as well as the lineage. The lineage (blue) can be seen as a large dictionary which encodes the complete lineage of the corresponding data. For each plugin which is identified by its data type (e.g. "events") the lineage encodes the name of the plugin, its version and the settings currently set in the context.

### 3.1.2 Data Structure: Runs, Chunks and Arrays

Given the fact that XENONnT uses a triggerless approach for its DAQ, its raw data cannot be simply divided into a countable quantity which is usually referred to as "event"<sup>20</sup>. Instead, STRAX processes data streams that can be organized in time intervals. There are different levels of time intervals which are explained in the following.

The highest level is given by a "run". A run marks a fixed period of data taking, typically 1 h for science data taking. Each run is associated with an identifier (runid) which is a zero-padded 6-digit number.

The next smaller unit of a time interval is a "chunk"<sup>21</sup>. The concept of runs and chunks is illustrated in Figure 3.3. Chunks either represent a fixed time interval or a fixed size of data in memory. XENONnT data coming from the DAQ are chopped into fixed time intervals of 20 s each, and are processed as soon as they are readout from the digitizer using the online context. At higher levels of the processing chain, plugins can specify if their results should be rechunked, and the size of the chunk is defined by the byte-size of a chunk. This ensures that files stored on disk are not becoming too small when data gets more refined. For example, the event-level data of a run is usually contained within a single chunk. Section 3.2 discusses the typical data sizes

<sup>20</sup>Please do not confuse with "events" which are used in the following explanations. Here events mean a paired S1 and S2 signal which belong together.

<sup>21</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/chunk.py#L14>



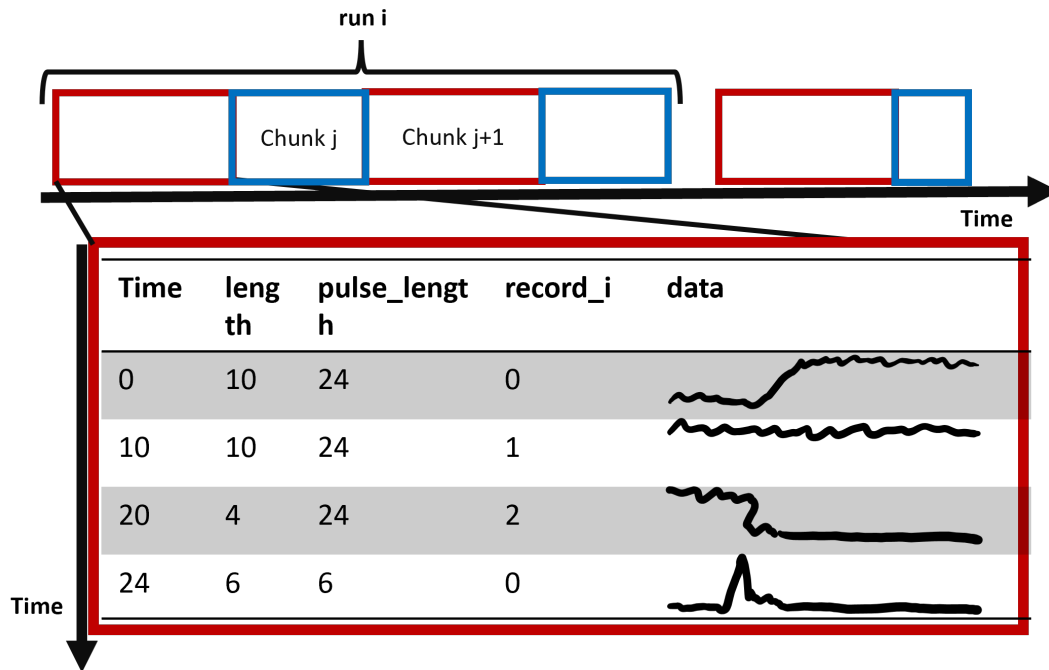


Figure 3.3. – Illustration of the strax chunking. The schematic illustrates how the data of the experiment is divided into different sorts of time intervals. The upper part illustrates runs and chunks while the lower part shows the tabular interval structure for some raw waveforms. The first three rows represent a raw S2 signal seen by a single PMT channel. The signal is too large to fill a single waveform buffer which has a fixed length. Thus it is split across multiple adjacent rows. To track whether a waveform is fragmented across multiple rows, additional information like the total `pulse_length` and a fragment index called `record_i` are stored along with the rest of the raw data. If a waveform is shorter than the buffer zero padding is applied. The last row represents a S1 signal raw waveform.

for the different data types used in XENONnT.

Each chunk carries some metadata information about its start-, end-time, and data size, as well as information about the measured signals within that given chunk in form of numpy structured arrays. The arrays can be seen as a table of time intervals, each containing some information about the different data types of the detector, e.g. a peak or an event. The time intervals are organized by their respective start- and end-times which are encoded as unix-timestamps in nanoseconds<sup>22</sup>. Each interval must be fully contained within the corresponding chunk. In general, the individual time intervals within an array can overlap e.g. PMT signals of different channels, but chunks must not overlap.

Organizing the time-streamed data in chunks has a few advantages from a performance point of view. Chunking facilitates storage of data in manageable file sizes, and it enables multiprocessing, as information must be contained within a single chunk.

<sup>22</sup>In case of low level data, like `raw_records` or `hits`, where multiple time intervals of different channel can overlap, time intervals are first sorted by starttime and then by channel. STRAX always assumes the data to be sorted which analysts must ensure when developing new plugins.

Thus, each chunk can be seen as causally independent and can therefore be processed in parallel. This argument can be used for most of the processing chain, e.g., when computing the properties of individual peaks, and allows for "online" live processing when the data is transferred from the digitizers onto the event builder machines. At some stages, the assumption that the information contained in a chunk is causally independent breaks and data needs to be processed in series, e.g., when building events from peaks, as it is explained in Section 3.1.3.

The usage of numpy structured arrays as a container for the measured detector quantities is also motivated by performance. Numpy structured arrays allows the developer to write custom code which can be compiled into machine code using the numba package. However, numpy arrays have the drawback that tables must always be of a fixed rectangular shape. This means that each row, containing a time interval with information, must be of the same length independent of how the data looks like. This is not a natural choice for a LXe TPC, as illustrated in Figure 3.3. Given the length of an S1 is being on the order of hundreds of nanoseconds and the length of S2 signals on the order of microseconds, the required buffer size to store their waveforms is quite different. Hence, for raw data, S2 signals are usually split across multiple rows of time intervals, where each row represents a fixed-length piece of the waveform.

### 3.1.3 Plugins

Plugins can be pictured as building blocks, defining the processing structure and how to transform, for example, raw waveforms into peaks and events. They contain all the custom code defined by the user. In addition, each plugin can be associated with a set of different options required to process certain data, like signal thresholds or the versions of the to be applied corrections. The general structure of a plugin is illustrated in Figure 3.4. Each plugin can generally be divided into four different sections:

**Options** are specified via the option-class<sup>23</sup> of STRAX. Each option has a name, e.g., "hit\_threshold" which is followed by the options default value and a short "help" description of the option meaning. Static options shared among many plugins might be defined by the context instead, e.g. like the PMT "channel\_map" in Figure 3.4. In addition, the developer can decide if certain option changes should be tracked in the lineage of the plugin which will trigger a reprocessing if the option changes.

The second part is given by the **plugin type and header**. STRAX offers different base classes for plugins from which all custom-made plugins must be inherited. The header contains information which is required for the plugin lineage, like the current version of a plugin, or on which data types the plugin depends and which data types the plugin provides. In addition, the header provides some extra metadata fields to, e.g., specify if the data should be rechunked before saving.

The plugin header is followed by the **plugin initialization** in which two optional functions can be defined, an `infer_dtype` and `setup` method, which are called once during the initialization of the plugin. The `infer_dtype` method defines the numpy

<sup>23</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/config.py#L56>

```

@export
@strax.takes_config(
    strax.Option(
        'hit_threshold', default=15, track=True, infer_type=False,
        help='Signal threshold applied to the inverted and baseline correcte waveform in ADC.'),
    strax.Option('channel_map', track=False, type=ImmutableDict,
        help='ImmutableDict mapping subdetector to (min, max) channel number.'),
    strax.Option('gain_model_nv', default=("to_pe_model_nv", "ONLINE", True), infer_type=False,
        help='PMT gain model. Specify as (model_type, model_config, nT = True)'),
)

class PluginForIllustration(strax.Plugin):
    """Short description of the plugin.
    """
    __version__ = '0.1.1'

    parallel = 'process'
    rechunk_on_save = True

    depends_on = 'records_nv'
    provides = 'hitlets_nv'
    data_kind = 'hitlets_nv'

    def infer_dtype(self):
        dtype = strax.hitlet_dtype()
        return dtype

    def setup(self):
        self.channel_range = self.config['channel_map']['nveto']
        self.n_channel = (self.channel_range[1] - self.channel_range[0]) + 1

        to_pe = straxen.get_correction_from_cmt(self.run_id,
            self.config['gain_model_nv'])

    def compute(self, records_nv, start, end):
        hitlets = find_hits(records_nv, self.to_pe)
        hitlets = compute_some_properties(hitlets)
        return hitlets

def find_hits(records_nv, to_pe):
    ...

def compute_some_properties(hitlets):
    ...

```

Options

Plugin type and header

Plugin initialization

Compute method and custom code

Figure 3.4. – Illustration of a typical `strax` plugin. The plugin can be typically divided into four different parts which are explained in the text. The code itself is only for illustrations and does not show any real code used.

data type of the provided data outputs. It represents the byte length of a single array row together with the associated column names like "time", "area" etc. The setup method facilitates initializing of specific settings, which either require access to CMT or depend on a setting that may be changed by the user.

The last and most important part of a plugin is the `compute`-method which contains the custom code, e.g., to build peaks or to estimate certain signal properties. The input of the compute method is provided based on their "data kinds"<sup>24</sup> instead of the individual "data types". A more thorough distinction between data kind and data type follows in Section 3.2. If a plugin depends on multiple data kinds, e.g., peaks and events, with different sized chunks, STRAX will split the chunks at an appropriate

<sup>24</sup>For the moment the reader can refer to a chunk of the same "data kind" as a chunk of peaks for example.

data gap, such that the same time interval of peaks and events is passed to the compute method. Depending on the provide field of the plugin header, the compute method has to return either a single or multiple numpy structure arrays as output. These arrays are added into new Chunks which contain additional metadata information, before they are passed either to the next higher plugin in the processing chain, or to a front-/backend saver which stores the chunks.

STRAX provides a set of different default plugin types from which all plugins must be inherited. For the development of new plugins the following plugin types are the most important ones:

The regular `strax.Plugin`<sup>25</sup> is used in most cases. This type of plugin assumes that the information stored within a chunk is causally independent of any other chunk.

A `strax.OverlapWindowPlugin`<sup>26</sup> is used in cases in which the causal independence no longer holds, when e.g. building events from peaks. In this case, data is processed in series and each chunk is loaded together with a bit of extra information from the previous and next chunk which is defined by an overlapping window size. During processing, the plugin computes the new data type (e.g., events) for the entire overlapping window region, but only information up to about the first half of the overlapping region is sent out to the receiver<sup>27</sup>. In the iteration of the next chunk the plugin again computes the information for the entire overlapping region, but this time only the remaining half of the result is sent out. In this way, the chunked data can be processed as a continuous stream.

A `strax.CutPlugin`<sup>28</sup> is used when defining data quality selections which will be referred to in the following as "cuts". A typical cut would be the fiducial volume for the analysis. Cut-plugins are a special kind of a plugin in which the user just has to return a boolean array in a `cut_by`-method. A logical True means that the corresponding entity (e.g., an event) is kept.

In addition to these plugins, STRAX offers two more plugin types that are required during data processing. The `strax.MergeOnlyPlugin`<sup>29</sup> loads different data types of the same data kind and returns a single array/chunk of the given data kind. A special case of the `strax.MergeOnlyPlugin` is the `CutList`-plugin that accumulates the boolean of many different cuts in a joint boolean array including a column containing their product. The `strax.ParallelSource`<sup>30</sup> plugin allows to inline multiple low-level plugins for multiprocessing. In this way, the individual compute methods are aligned into a single plugin reducing performance penalties from copying large chunks in and out of memory.

<sup>25</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/plugin.py#L40>

<sup>26</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/plugin.py#L662>

<sup>27</sup>Either the next plugin or a saver.

<sup>28</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/plugin.py#L885>

<sup>29</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/plugin.py#L954>

<sup>30</sup><https://github.com/AxFoundation/strax/blob/ef39e3f94cfabc48a02cd3b7d446df049f2a55a6/strax/plugin.py#L976>

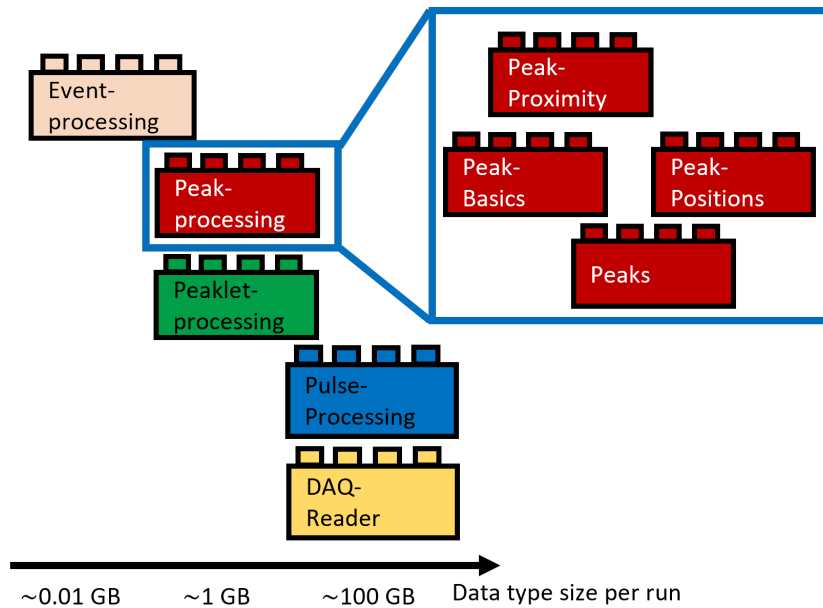


Figure 3.5. – Sketch of the XENONnT TPC data structure. The building blocks on the left illustrate the different stages in the processing chain. Each block corresponds to plugins providing a different data kind. Their position indicates the typical size of their corresponding output for a single 30 min background run. The inbox indicates that a single data kind is made of several plugins which compute different signal properties all reflected by a separate data types.

### 3.2 SIGNAL TYPES: HITS, PEAKS AND EVENTS

Each `strax.Plugin` can either return a single data type or multiple data types. If two data types are describing the same objects, e.g. peaks, but different properties, e.g. positions and shapes, that are computed by different plugins, then these data types are of the same data kind. Structuring data in this way allows to automatically merge information from different data types, but of the same kind when loaded together. This merging is performed by a `strax.MergeOnlyPlugin`.

In XENONnT the TPC data structure can be broken down into five different data kinds which describe most of the TPC data: records, (lone) hits, peaklets, peaks, and events. The entire information of higher level data types (> peaklets) is usually computed through several different plugins. Each plugin can be seen as a building block which can be stacked onto other building blocks to construct the entire processing chain, as illustrated in Figure 3.5. If required by the user individual building blocks can be exchanged, by exchanging registered plugins. In the following the most important TPC data types, exemplary shown in Figure 3.6 are explained.

Raw data recorded by one or more of the three detectors is handled and read out by "Distributed, parallel high-throughput data acquisition software for XENONnT" (REDAX). REDAX<sup>31</sup> reads the data from the digitizer in a block transfer and converts the native digitizer data format into a STRAX-friendly binary format. The data is stored

<sup>31</sup><https://github.com/AxFoundation/redax>

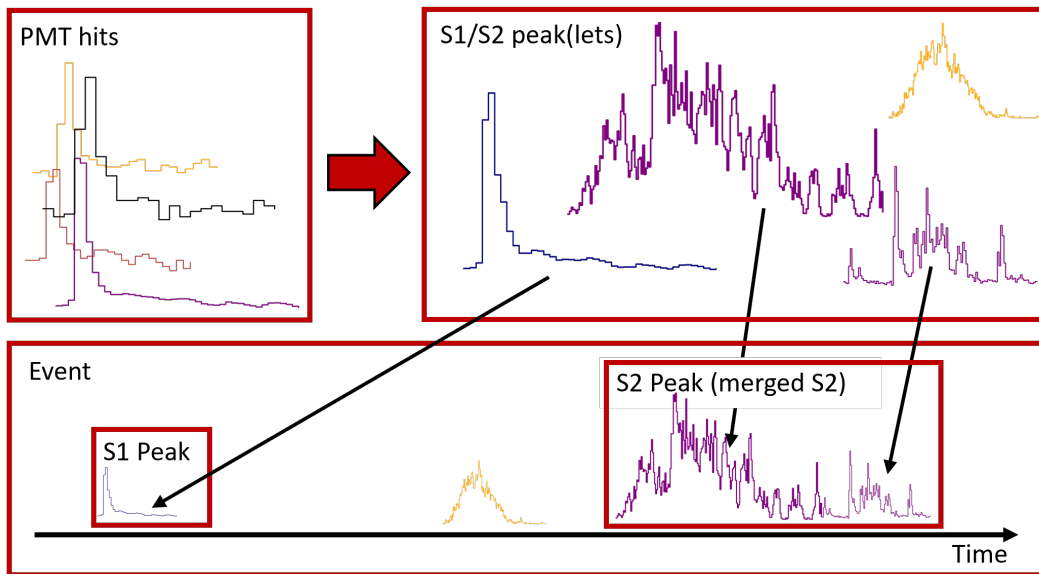


Figure 3.6. – Schematic of the most important steps in the TPC event building which are explained in the text. The depicted hits and peaks are signals taken from a  $^{83m}\text{Kr}$  calibration run.

on disk in chunks typically in a length of 5 s to 30 s depending on the run settings. At this point the data of the individual detectors, TPC, neutron- and muon-veto are stored in a single chunk. Each main-chunk is accompanied by pre- and post-chunk of 0.5 s in which the corresponding 0.5 s of data is saved twice. Therefore, the data of the current post-chunk is identical to the data of the next pre-chunk. In this way, a continuous stream of data with causally independent chunks can be ensured when reading and processing the data in parallel by STRAXEN [Apr+22d].

The entry point into the STRAX framework for the data is given by the DAQReader-plugin<sup>32</sup>. The DAQReader reads in parallel the pre/main/post-chunks written by REDAX and splits the data written in the pre/post-chunk at a suitable gap of at least 1  $\mu\text{s}$ . Similarly, as in a `strax.OverlappingWindowPlugin`, data before that gap is appended to the current chunk, while the remaining data is deleted for the current chunk and will be added to the next chunk. If no suitable gap can be found an artificial dead-time is introduced. In addition to the chunking, the DAQReader divides the data into multiple `raw_records` data kinds depending on the channel the data was recorded in. Each detector has a data kind for the main data of the detector and one for acquisition monitor data like the synchronization signal given by the GPS module.

In the next step `raw_records` are transformed into records. All PMTs in XENONnT produce negative signals, which are baseline subtracted and flipped in the Pulse-Processing-plugin<sup>33</sup>. The baseline is computed in the pre-trigger window over the first 40 samples of a TPC pulse. Afterwards, a hitfinder algorithm is applied. A

<sup>32</sup><https://github.com/XENONnT/straxen/blob/8689fb2d429320b5198e3ce4953cabccaec11415/straxen/plugins/daqreader.py#L59>

<sup>33</sup>[https://github.com/XENONnT/straxen/blob/4ba2209e7d3020222af7653119d04d99d259a59/straxen/plugins/pulse\\_processing.py#L106](https://github.com/XENONnT/straxen/blob/4ba2209e7d3020222af7653119d04d99d259a59/straxen/plugins/pulse_processing.py#L106)

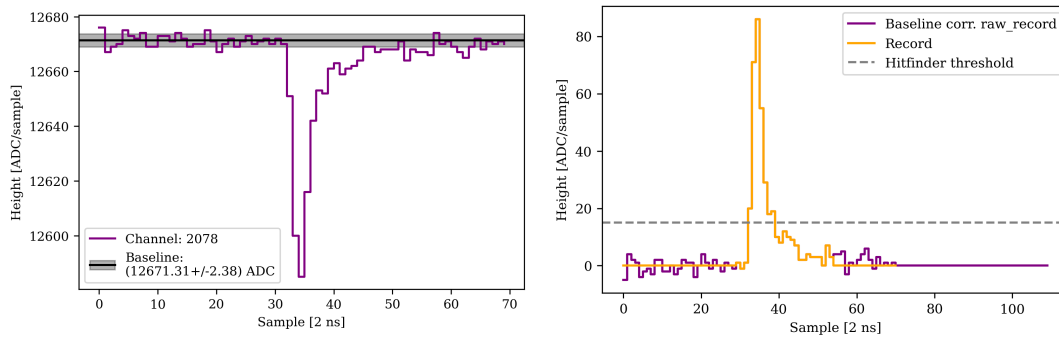


Figure 3.7. – Illustration of a `raw_record` and `records` of a single NV PMT pulse. The left figure shows the raw pulse returned by the `DAQReader`-plugin. The baseline mean and RMS is shown as a black line and gray shaded region. Only the non zero-padded part of the waveform is plotted. The plot on the right hand side shows the baseline corrected and flipped `raw_record` in purple including the zero-padded buffer for samples larger than 70. The orange graph shows the resulting `record`. The gray dashed line indicates the channel dependent hitfinder threshold.

PMT hit forms the smallest unit of a signal in XENONnT which is found if a successive number of samples exceeds a fixed threshold after baseline subtraction. The threshold depends on the PMT channel and is set to the same value as for the corresponding digitizer channel in the DAQ. Hit finding reduces the overall data size of records compared to `raw_records` by applying a zero-padding to the waveform outside of hits. To exclude the PMT signal tails from zero-padding, each hit is extended by a fixed number of samples to the left and right (3 and 20 samples respectively). The hits themselves are not saved along records to save disk space. The `PulseProcessing` step is also used in the processing chain for both vetos, but with slightly different settings. Figure 3.7 illustrates a single `raw_record` and `record` for a NV pulse.

Before categorizing the TPC signals into S1 and S2, hits need to be grouped into peaks. In XENONnT the grouping is done in a two-stage process. At the first stage, hits are grouped into peaklets<sup>34</sup> by splitting groups of neighboring hits if a gap of at least 700 ns between two hits is found. For each peaklet, a summed waveform over all hits is computed. Given the fact that numpy-structured arrays must be of a rectangular shape, the data-field buffer of the waveform must have a fixed length set to 200 samples. This means that a peak(let) that lasts longer than  $2\ \mu\text{s}$  is downsampled to fit into the buffer. This is mostly the case for S2 peak(lets). Single hits which are not grouped into any peaklet are stored as `lone_hits`.

After forming peaklets and computing their summed waveform, peaklets are subdivided based on a modified natural break algorithm<sup>35</sup>. The splitting is necessary as the simple clustering of hits may merge close by signals into a single peak e.g. S1 signals and PMT afterpulses or close by peaks. The algorithm estimates for each sample within the summed waveform a "goodness-of-split" and splits the peak at the given

<sup>34</sup>[https://github.com/XENONnT/straxen/blob/8689fb2d429320b5198e3ce4953cabccaec11415/straxen/plugins/peaklet\\_processing.py#L72](https://github.com/XENONnT/straxen/blob/8689fb2d429320b5198e3ce4953cabccaec11415/straxen/plugins/peaklet_processing.py#L72)

<sup>35</sup>[https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/processing/peak\\_splitting.py#L232](https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/processing/peak_splitting.py#L232)



position if a specified threshold is exceeded. Subsequently, the algorithm recursively tries to divide the split peaks further until either no split point is found, the area of the split peaks would become too small, or a recursion limit is reached. After splitting, a variety of peak(let) properties are estimated to describe the shape, the recorded signal distribution in the top and bottom PMT array, and the hit structure within the peak. These properties are needed for the classification of individual peaks.

Peaklets are classified either as S1, S2, or "unknown" mainly based on their shape properties as well as the hit-pattern distribution of the top and bottom PMT array. To be classified as an S1, at least three PMTs must contribute to the peak(let) in a tight coincidence of  $\pm 50$  ns around the center of the integrated waveform of the peak. This threshold ensures that the number of fake S1 signals due to accidental coincidences between PMT dark counts is reduced. In addition, it is required that each S1 signal has a small rise-time, which is defined as the time to reach from 10 % to 50 % of the total peak area<sup>36</sup>, and a low area fraction top, which is defined as the peak area recorded by the top PMT array divided by the total peak area<sup>37</sup>. Both requirements help to discriminate small S1 signals from single electron S2s, and are shown in the appendix in Figure B.5 on page 240. All remaining peaks which are not classified as an S1 and which have hits contributing from at least four different PMTs are classified as S2s.

All peaklets which have been classified as an S2 are passed to the MergedS2s plugin. As the name suggests this plugin successively merges S2 peaklets, if applicable, until complete S2 peaks are built (illustrated in Figure 3.6 through the purple peaklets/peak waveform). The merging starts with the two peaklets with the smallest gap in between. Neighboring S2 peaklets are merged until the total duration of the new S2 would be larger than  $50 \mu\text{s}$  or no other suitable merging candidates within the vicinity of the S2 can be found. The maximum duration was introduced to prevent the algorithm from accumulating tails of S2 signals<sup>38</sup> into the main peak, which would otherwise lead to signals spanning several hundreds of micro-seconds. The gap size in which other S2 peaklets are considered to be suitable for merging depends on the integrated peak area of the newly merged S2. The area dependence on the merging gap size was chosen due to the low drift field in XENONnT. S2 signals created just above the cathode electrode can span over several  $\sim 10 \mu\text{s}$ , thus achieving a high cluster efficiency without clustering everything into a single peak becomes quite challenging. The gap size threshold is shown in the appendix in Figure B.6 on page 241. S1s are in general excluded from the merging procedure as it is not unlikely that an S1 appears between two single electron background S2s signals. After this step all merged S2 peaklets and the remaining unmerged peaklets, like S1 signals or S2 signals without merging partner, are called peaks.

After classifying the peaks, the x/y position of S2 signals is identified based on their top PMT array hit-pattern. To estimate the position three different algorithms can be used: Convolutional Neural Network (CNN), Graph Convolutional Network (GCN),

<sup>36</sup>The peak area is the integral over the summed waveform of all contributing PMT signals (records).

<sup>37</sup>This requirement is used as S1 signals have typically a smaller area fraction in the top PMT array compared to S2s due to the total reflection of photons at the liquid gas interface.

<sup>38</sup>These tails are caused by a delayed extraction of electrons and photo-ionization caused by the S2 scintillation light.



and a Multilayer Perceptron (MLP) neural network [Lia+22]. By default, the MLP reconstruction is used.

In the last step peaks are grouped into events, as depicted in the lower panel of Figure 3.6. An event is formed if an S2 peak is considered triggering peak requiring an area of at least 100 pe and featuring at most 8 competing neighboring peaks. A peak is considered competing if it has at least half the peak area of the main S2 and is found within a  $\pm 10$  ms window. The event window itself is then defined as an interval of  $[-2450 \mu\text{s}, 250 \mu\text{s}]$  around the triggering S2 peak. If multiple event windows overlap they are merged. After defining the event window the main S1 and main S2 within the event are defined. The main S1 is given by the largest S1 peak which can be found within the event, while the main S2 peak is the largest S2 signal found after the main S1<sup>39</sup>. The second largest S1/S2 in an event are defined as "alternative" S1/S2, illustrated by the orange S2 in Figure 3.6. Also here the alternative S2 must be recorded after the main S1. These alternative S1/S2 peaks are important to identify multi-scatter events. Additional peaks are not considered on the event level. Once main and alternative S1/S2 are identified, important event level properties are determined like for example the drift time between S1 and S2, and the main and the alternative signals are corrected for the detector dependent effects discussed in Chapter 2.

### 3.3 PLUGINS FOR THE XENONNT NEUTRON VETO

One aspect of this work was the development and commissioning of the neutron- and muon-veto plugins for XENONnT. These plugins were developed mostly from scratch, adding new methods and functions to STRAX, STRAXEN, and CUTAX<sup>40</sup>. Figure 3.8 shows the dependency structure of the NV processing chain. The main body of the processing chain is given by six different plugins required to get event-level information for both vetos. Additional plugins are used to compute the synchronization between vetos and TPC as well as the veto intervals which are used to tag TPC events. To separate the data types belonging to the different detectors all data types provided by the vetos have a suffix `mv` or `nv` for the muon- and NV respectively.

If not referred to otherwise in the text, muon-veto plugins are inherited from the corresponding NV plugin. The inheritance is done via so-called "child plugins" which were developed during this thesis. Child plugins ensure that a change of an option in the child is properly propagated to the compute method which is inherited from the parent plugin. It also updates the lineage of the data type accordingly and adds to the lineage the name and version of the child- as well as the parent plugin. In this way code changes of the parent are tracked, too. The exact working principle of the individual plugins, which data types they are providing, and how they are performing are discussed in the following subsections. An actual usage of the data is then shown in the subsequent Chapters 5, 7 and 6. For simplicity, the text will always speak about NV plugins.

<sup>39</sup>"Largest" refers to the largest integrated peak area.

<sup>40</sup>Since cutax is a private repository, references to the code can only be accessed by members of the XENON collaboration. However, for documentation purposes the references are given here anyhow.

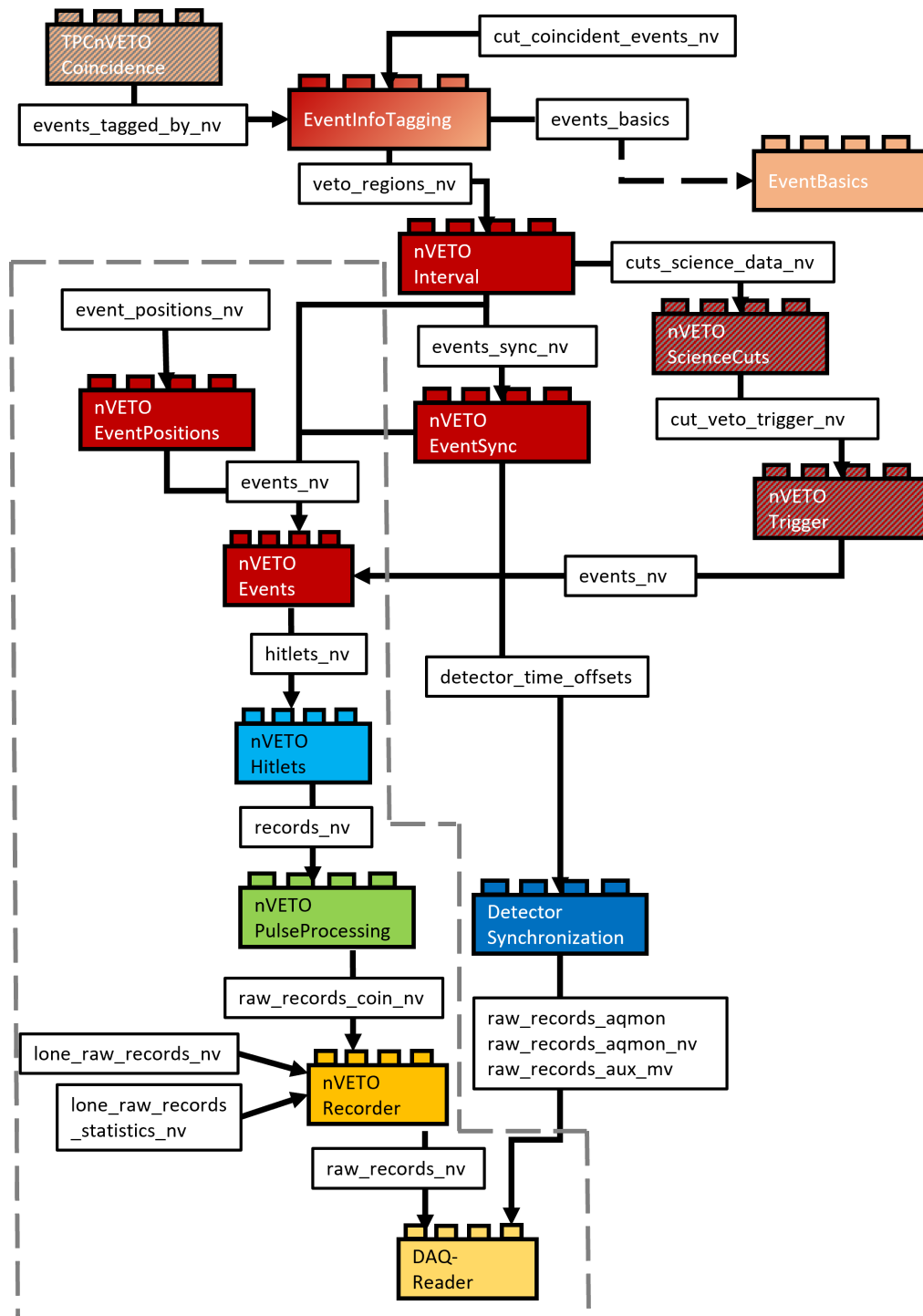


Figure 3.8. – Dependency graph of the NV processing chain. Plugins are displayed as building blocks the arrows between plugins indicate the individual plugin dependencies. Data-types provided by the different plugins are labeled next to the arrows. Arrows not pointing from one to another plugin indicate additional data types provided, but only required for analyses or monitoring purposes. Plugins which share the same color are producing data types of the same data kind. Gray shaded plugins indicate *CutPlugins*. The *EventBasics*-Plugin belongs to the TPC processing chain. The gray dashed region indicates all plugins required to build event-like data for the analysis of the NV and muon-veto data. The other plugins are required to build the neutron- and muon-veto veto-cut for the TPC.

### 3.3.1 Raw-data, software trigger and data reduction

The raw-level processing for the vetos is based on the plugins already provided for the TPC. The `DAQReader` discussed in Section 3.2 is shared among all three detectors providing `raw_records_nv` and `raw_records_mv`. The `raw_records` data type contains a fragmented version of the recorded neutron- and muon-veto pulses. The fragmentation is necessary due to the requirement of squared buffers as discussed in Section 3.1.2. The raw data types of each detector provide a field called `data` which is a buffer of the size of 110 samples. This corresponds for the NV to a time length of 220 ns per fragment and for the muon-veto, as well as the TPC, 1100 ns per fragment. If a single pulse exceeds this buffer size it is fragmented across multiple rows/intervals of the array. Normally, NV PMT pulses are shorter than the buffer. In this case, zero-padding is applied at its end. Information about how many samples are not zero-padded is given by the `length`-field provided in raw data types. Uniquely for the NV, `raw_records_nv` also contains a `baseline`-field which stores the baseline estimate by the digitizer. This differs from TPC and muon-veto as the information cannot be read out from the corresponding digitizer.

Given the larger dark count rate of the NV PMTs a software trigger plugin called `nVETORecorder` is applied between the `DAQReader` and the NV `PulseProcessing` step. Just like in the TPC `PulseProcessing`-plugin, PMT pulses are first baseline corrected and flipped before the same hit finding algorithm is applied. The working principle of the NV trigger is shown in Figure 3.9 based on an example given in the very same figure. The software trigger itself is based on a simple signal coincidence. It is formed by searching for a minimum number of  $n$  hits within a moving resolving time window  $\Delta T$ . For this purpose, the time differences  $\delta t_i$  between the individual hits are computed and a leading zero is appended. In the next step the  $\vec{\delta t}$ -array is convoluted with a simple kernel of the length  $n$  consisting of only ones and a leading zero. The result of the convolution represents the total time difference  $\Delta\tau$  between  $n$ -neighboring hits. If the total time difference is below the required resolving time  $\Delta T$  the first hit of these  $n$  hits is marked as a valid trigger window start. By definition, the last  $n - 1$  hits cannot form a valid trigger window. Each trigger start creates a trigger window of the length  $\Delta T$ . An additional pre-trigger window can be added. If multiple windows overlap they are merged. Only pulses that touch one of the trigger windows are saved. Given that the muon-veto uses a hardware trigger no additional software trigger is required for the muon-veto processing chain.

The settings of the software trigger were optimized to have a very high acceptance for neutron capture Cherenkov signals while keeping the pile-up due to PMT dark counts low. A discussion about this optimization follows in Section 3.3.3 and 3.4. The resolving window was set to  $\Delta T = 600$  ns, and a  $n = 3$ -fold coincidence. A pre-trigger window of 150 ns is used. Compared to the TPC, the NV uses a shorter pre-trigger window of only 32 samples for the recorded PMT pulses. Therefore, the baseline is estimated only on the first 26 samples. The hit finding threshold is set for the most NV channel to 15 ADC counts. A few channels with a higher noise level have a higher threshold up to 20 ADC counts. Figure 3.7 on page 62 illustrates `raw_records_nv` and

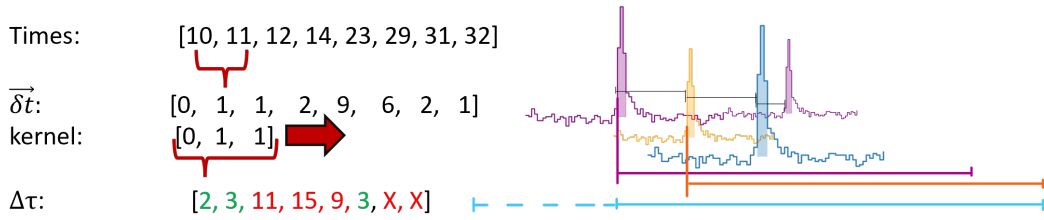


Figure 3.9. – Example and schematic illustrating the NV software trigger. The left hand side illustrates an example for a coincidence trigger requiring  $n = 3$  hits to be found within a resolving time  $\Delta T = 4$  samples. The first row represents the times of the different pseudo hits.  $\vec{\delta t}$  shows the time differences which are convoluted with the kernel moving in the direction indicated by the red arrow. The last row indicates the total time differences computed by the convolution and which time stamp would start a trigger window. In this example 3 trigger starts 10, 11 and 29 are found, resulting in two triggering windows lasting from 10 to 15 and 29 to 33 without any pre-trigger window being specified. The concatenation of trigger windows is also illustrated on the right hand side where the purple and orange window yield the cyan merged window. The pre-trigger region is indicated as a dashed cyan line.

`records_nv` for a single NV PMT pulse.

The data type provided by the `nVETORecorder`-plugin is an exact copy of the `raw_records_nv` which survive the coincidence trigger. In general, the idea is to only temporarily buffer `raw_records_nv` for about 2 days on the NV readout server in case of a problem with the processing chain or a SuperNova Early Warning System (SNEWS)-alert. If not interrupted by such an alarm the regular raw data would be deleted and only `raw_records_coin_nv` kept.

For both vetos the next higher plugin is the `PulseProcessing`-plugin<sup>41</sup>. Like for the TPC, this plugin serves the purpose to further reduce the overall amount of data by extending the zero-padding to the rest of the data-buffer outside of any PMT-hits as it is shown in Figure 3.7 on page 62. However, in SR0 it was decided to store all the raw data information for the neutron veto. This allows to perform extensive testing of the novel processing chain before deleting any data. Hence, to save storage capacities only `raw_records_nv` are stored during SR0 but neither `raw_records_coin_nv` nor `records_nv`. A discussion about the storage consumption for the NV data can be found in Section 3.4.

### 3.3.2 Hitlets

The next higher level data type in the processing chain for both vetos are the so-called hitlets. hitlets are hits but include the left and right extensions discussed earlier. The `nVETOHitlets`-plugin<sup>42</sup> computes hitlets using three steps.

First hits are again searched in records and extended to the left and right. In the second step, overlapping hitlet-intervals are merged into a single hitlet. This is nec-

<sup>41</sup>[https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebbd/straxen/plugins/veto\\_pulse\\_processing.py#L37](https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebbd/straxen/plugins/veto_pulse_processing.py#L37)

<sup>42</sup>[https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebbd/straxen/plugins/veto\\_hitlets.py#L43](https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebbd/straxen/plugins/veto_hitlets.py#L43)

essary as the hitfinder only operates within a fragment of a pulse. If a PMT signal is sitting exactly in between two fragments it might be recognized as two separate hits. In addition, the hitfinder may identify multiple PMT signals only as a subset of overlapping hitlets. After merging the individual hitlet time intervals, a joint waveform is computed per PMT channel, and the unit of the signal height is converted from ADC counts/sample to pe/sample<sup>43</sup>. The conversion factors are based on the weekly PMT light emitting diode (LED) calibrations, and are queried from the corrections database via CMT.

In the third step, a local minimum splitting algorithm<sup>44</sup> iterates over all hitlets checking whether it can be split into multiple separate hitlets. The splitting algorithm used was already implemented in XENON1T and ported into STRAX. It is based on the concept of topographic prominence. A prominence is given by the largest relative height difference between a local minimum in the middle of two peaks and the smaller amplitude of the two peaks. A signal is split at the local prominence if the height of the prominence overcomes a specified threshold. The algorithm supports two different types of thresholds of which both must be surpassed. They are given by a fixed height the prominence must overcome and a minimal ratio multiplied by the signal height at the local prominence. Figure 3.10 illustrates the local-prominence splitting. For SR0 the fixed height threshold was set to 0.063 pe/sample, and the min-ratio was set to 0.75.

After forming hitlets, various properties are computed, e.g. area, amplitude, and time of the amplitude. In addition, several different algorithms were implemented to estimate the shape of the signal as well as a "conditional information entropy"<sup>45</sup>-like quantity which is estimated as

$$H(X|Y) = - \sum_i p(x_i) \cdot \log_{10} \left| \frac{p(x_i)}{p(y_i)} \right| \quad (3.1)$$

where  $p(x_i)$  represents the squared normalized height of a hitlet  $x$  at sample  $i$  and  $p(y_i)$  the squared normalized height of a template  $y$ . This entropy-like quantity can be used to discriminate noise from PMT signals. A noise signal should average to zero leading in general to a smaller entropy than a peak-like signal if a flat template is used.

To estimate the signal shape three different methods are used. The first method is based on the area deciles which are also used to estimate the rise time and width of S1 and S2 peaks for the TPC and was already discussed in Section 3.2 on page 60. For hitlets the central 50% and 80% area width are computed. In addition, two new methods were implemented in strax.

The first methods estimate the full width at a certain percentage of the signal height, e.g. the full width at half maximum. The algorithm searches for the first samples whose heights are above and below the specified threshold on both sides of the signal

<sup>43</sup>PE refers to the unit photo-electron. One photo-electron corresponds to one detected photon measured by a PMT.

<sup>44</sup>[https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/processing/peak\\_splitting.py#L179](https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/processing/peak_splitting.py#L179)

<sup>45</sup><https://github.com/AxFoundation/strax/blob/d6b3575df08f639cd4d054f116bb57f3cd2db92e/strax/processing/hitlets.py#L421>

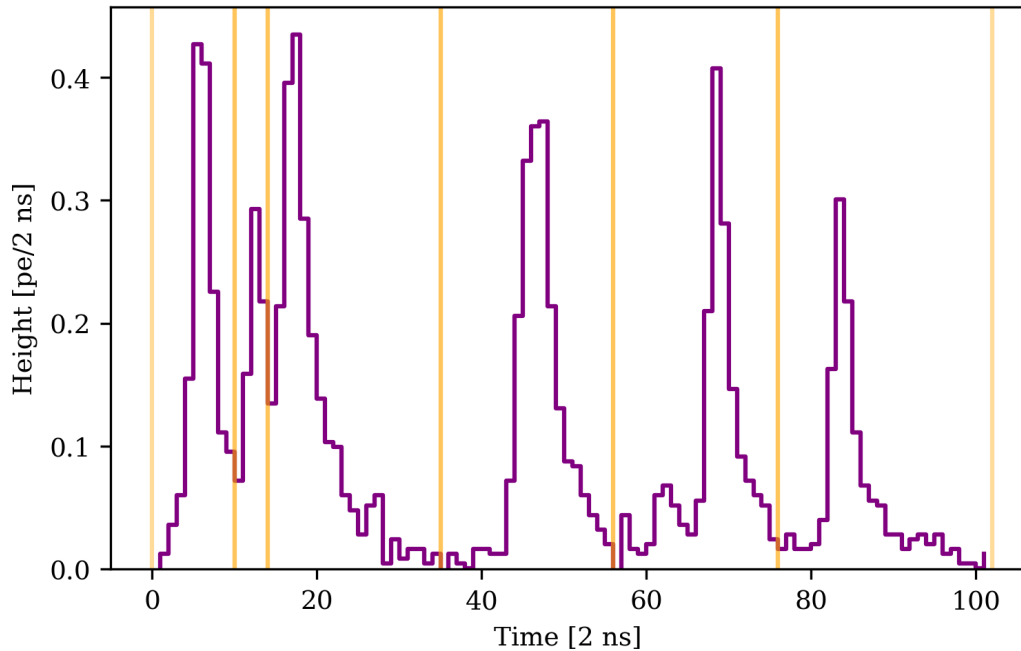


Figure 3.10. – Neutron-veto hitlet for channel 2116 of run 026310 before applying the local prominence splitting. The orange vertical lines are showing the split position when applying the local minimum splitting algorithm for the settings specified in the text.

amplitude. Afterwards, a linear interpolation between the samples is used to approximate the width at the specified height.

The second method estimates the width based on the highest density region of a PMT signal. The highest density region is computed via an "integration" along the hitlet height starting at the amplitude and ending at the baseline. The integration is done by sorting first all samples according to their height and iterating of other the different heights until the desired fraction is reached or surpassed. If the desired fraction is surpassed the width is estimated via a linear interpolation on the last sample according to the remaining fraction desired. Figure 3.11 shows an example of all three width parameters. The hitlet data type stores along each estimated width the left edge at which the corresponding width starts.

The large number of different shape parameters as well as the information entropy was implemented as the NV DAQ suffered from a large noise contribution during the commissioning phase of the experiment. The noise was given on one hand by electronic noise, but also from bit-flips in the digitized waveforms due to a bug in the digitizer firmware. The different parameters were used to discriminate noise from PMT pulses. Eventually, the noise conditions were improved and the bit-flip problem was solved. In total each hitlet contains 78 bytes of information of which 52 bytes are due to the shape parameters and entropy. A down-sampled representation of the signal, like for peaks, is not stored along the data.

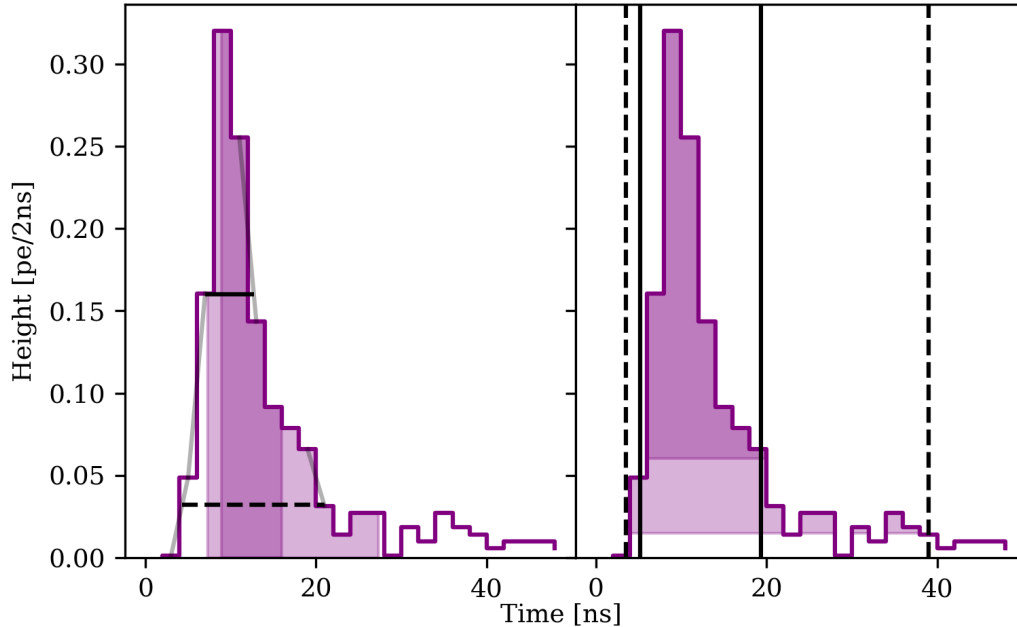


Figure 3.11. – Single photo electron hitlet taken from a SR0 background run. The three different shape properties computed by the *Hitlet*-plugin are illustrated. The plot on the left hand side illustrates the 50 % and 80 % width given by the area deciles as shaded regions. The vertical black lines indicate the full width at half maximum and the full width at the 10 % maximum. The figure on the right hand side illustrates the width estimated via the highest density regions algorithm for an area of 50 % and 80 % as shaded regions. The vertical lines indicate the 50 % and 80 % as solid and dashed line respectively.

### 3.3.3 Veto-Events

Neutron veto events are formed by the *Events*-plugin<sup>46</sup> based on hitlets. For the event-building the same type of moving coincidence is used as for the software trigger explained in Section 3.3.1. After forming the event-intervals hitlets are grouped by events to compute their properties. In some very rare cases a single very long hitlet may be attributed to more than just a single event. In such a case the hitlet will be assigned to the first event recorded in time. Each event start and end is then set according to the start and end of the first and last hitlet within the event. Compared to the software trigger, the event-level coincidence uses a shorter resolving time window of 200 ns and 3-fold coincidence. The time window was chosen based on the fact that the time distribution of a typical Cherenkov event for the neutron veto follows an exponential distribution with a decay constant of  $\sim 60$  ns as it is discussed in Chapter 5. Thus, the event building efficiency for a Cherenkov signal containing only 3 detected photons can be evaluated using a simple toy Monte Carlo (MC) to be  $\sim 90\%$ <sup>47</sup>.

<sup>46</sup>[https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/veto\\_events.py#L14](https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/veto_events.py#L14)

<sup>47</sup>The NV veto threshold for neutron capture event, discussed in Chapter 7, is at 5 detected photons for which the event building efficiency is  $\sim 99.9\%$ .



No additional pre-trigger is applied. The muon-veto uses a slightly larger window of 300 ns.

For each reconstructed event basic properties, as the total area, number of contributing PMTs, number of contributing hitlets and area per PMT-channel are computed and stored. Additionally, the so-called "center-time" and the center-time spread are estimated. The center-time is computed as the area weighted average arrival time of all hitlets within an event. The center-time spread represents the corresponding area weighted standard deviation.

A separate plugin called `nVETOEventsPositions`<sup>48</sup> estimates the event position inside the NV. The position is localized via a simple area weighted mean for all PMT channels which recorded a hitlet within the first 20 ns of an event. Travel time information of the individual photons is not used in the position reconstruction<sup>49</sup>. The reconstructed position is stored in Cartesian coordinates and as an azimuth angle. An example plot for the position reconstruction is shown in Figure 5.4 on page 104. For the muon-veto no position reconstruction is currently implemented. A single veto event contains 95 bytes of information.

A typical NV event is shown for a 4.4 MeV gamma in the NV event display in Figure 3.17 on page 82.

### 3.3.4 Detector synchronization and linking of unlinked runs

As explained in Section 2.4.4 signals recorded by TPC, MV and NV were taken most of SR0 in a "linked" acquisition mode in which the signal timestamps of all three sub-detectors have been synchronized via a 50 MHz clock. Having all three detectors synchronized is of vital importance for the formation of proper veto-intervals and especially for the proposed NR calibration via tagged neutrons, which require synchronization on the time scale of S1 signals which is  $\sim 50$  ns. However, due to the HEV, TPC signals are delayed with respect to the two veto systems, which needs to be corrected. The synchronization of the detector signals was a joint effort together with Andrea Molinario from Turin University who analyzed the time delay between the three detectors using Muon signals, and Ricardo Peres from Zürich University who also worked on the translation of the detector timestamps into GPS timestamps as discussed later in this section.

The synchronization is performed in a single `strax.Plugin` called `DetectorSynchronization`<sup>50</sup> which is then propagated to the event-level data of NV and MV via the `nVTEOEventSync` plugin<sup>51</sup>. The plugin computes for a given run the average time

<sup>48</sup>[https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/veto\\_events.py#L243](https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/veto_events.py#L243)

<sup>49</sup>This is due to the fact that a typical neutron capture event, as it will be shown later in Chapter 7, only contains about 19 detected photons which are exponentially distributed over a time range of a few 100 ns. Thus, any directional information of the Cherenkov light is quickly washed out due to diffuse reflection on the ePTFE surfaces inside the NV.

<sup>50</sup>[https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/acqmon\\_processing.py#L464](https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/acqmon_processing.py#L464)

<sup>51</sup>[https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/veto\\_events.py#L492](https://github.com/XENONnT/straxen/blob/a821086851851d54761b6cafb4ff6fc3aaebabda/straxen/plugins/veto_events.py#L492)



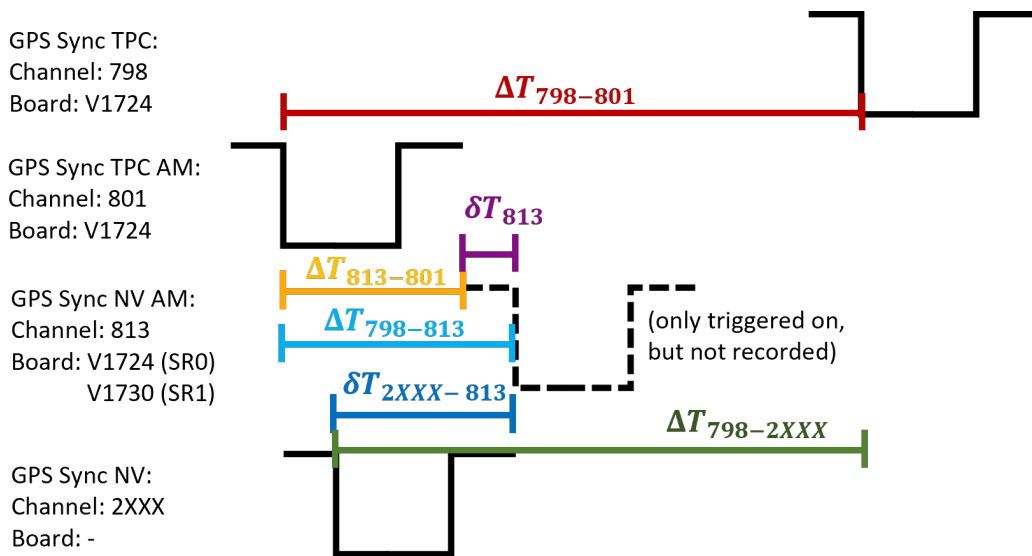


Figure 3.12. – Illustration of the synchronization of the NV with the TPC during SR0. The individual Nuclear Instrumentation Module (NIM) GPS pulses are shown in black and the different involved delays in different colors. The upper two rows indicate the GPS signals recorded by the TPC digitizer and acquisition monitor (AM) channels, while the lower two rows show the corresponding NV channels. For more explanation please refer to the text.

delay between the three detectors using the leading edge of the 0.1 Hz GPS signal recorded by all three detectors. Throughout the commissioning phase and SR0 the distribution of the GPS signal to the individual sub-detector changed which was especially a challenge for the AmBe NR calibration. Figure 3.12 illustrates the corresponding changes and how the delay between TPC and vetos is determined.

Ultimately, to correct for the time delay between TPC and NV, the delay of the GPS signal  $\Delta T_{798-2xxx}$  recorded by the regular TPC and NV channels needs to be determined. However, the GPS signal was never recorded by the regular NV digitizer during SR0, but only by channel 813 of its acquisition monitor which has a small delay  $\delta T_{2xxx-813}$  with respect to the regular NV channel. This delay was found to be 74 ns and is mainly determined by the acquisition start delay of the regular NV digitizer and the acquisition monitor, as well as a 10 ns cable between the two. Further, it was not possible to record the GPS pulse itself by the NV acquisition monitor but only to trigger on it<sup>52</sup>. Thus, the actual GPS pulse is additionally delayed by the corresponding pre-trigger window of  $\delta T_{813} = 160$  ns. Also for the TPC, the GPS signal was only recorded by its acquisition monitor during the AmBe calibration and only added later to a spare channel of the regular TPC digitizer. Thus the delay  $\Delta T_{798-801}$  was determined after the AmBe calibration and was found to be 10 137 ns averaged over SR0 which is 83 ns shorter than the set delay of 10 220 ns. It should also be noticed that this delay was changed during the AmBe calibration to optimize the HEV. Initially, a delay of 5000 ns was used which was also corrected for the 83 ns difference found

<sup>52</sup>The reason was that it was not possible to delay the acquisition window enough through the digitizer to accommodate the GPS pulse itself.

between the set and measured value. Eventually, the NV signals are delayed into the time frame of the TPC by a time delay of

$$\begin{aligned}\Delta T_{798-2xxx} &= \Delta T_{798-801} - (\Delta T_{798-813} - \delta T_{2xxx-813}) \\ \Delta T_{798-2xxx} &= \Delta T_{798-801} - (\Delta T_{813-801} + 87 \text{ ns})\end{aligned}$$

where  $\Delta T_{798-801}$  is either 4917 ns or 10 137 ns depending on the run, and  $\Delta T_{813-801}$  the time delay between the measured GPS pulses of the two acquisition monitors. The accuracy of the synchronization and its precision were found to be  $(-1.21 \pm 0.24)$  ns and  $(14.03 \pm 0.23)$  ns respectively, using a  $\gamma$ -coincidence between TPC and NV discussed in Section 6.2.1.

While delaying NV and MV signals into the time frame of the TPC works for data taken in the synchronized linked mode, it does not work if the three detectors are operated independently, which is referred in the following as unlinked mode. When taking data in unlinked mode the start of the acquisition for a given detector is arbitrary with respect to the other two. Thus, instead of correcting the two vetos relative to the TPC, it is simpler to correct the time stamps of all three detectors into an absolute time scale which is given by the GPS timestamps recorded along with the 0.1 Hz GPS signals. To translate a recorded signal timestamp  $t_{DAQj}$  into a GPS timestamp  $t_{GPSj}$  a linear interpolation between adjacent GPS pulses is used and as such the translation is given by a simple linear equation of the form

$$t_{GPSj} = \frac{T_{GPS,i+1} - T_{GPS,i}}{T_{DAQ,i+1} - T_{DAQ,i}} \cdot (t_{DAQj} - T_{DAQ,i}) + T_{GPS,i} \quad (3.2)$$

where  $T_{GPS,i+1}$ ,  $T_{GPS,i}$  are the GPS timestamps of two adjacent GPS signals recorded at the times  $T_{DAQ,i+1}$ ,  $T_{DAQ,i}$ , and  $j$  runs over all signals found between these two pulses. The GPS correction is performed by a `strax.OverlapWindowPlugin` called `GpsSync`<sup>53</sup> which is shared by all three detector systems. Using an `OverlapWindowPlugin` ensures, that always multiple GPS timestamps are loaded as the frequency of the GPS corresponds to the typical chunk size. This also ensures that the translation can be performed even if an intermediate GPS signal was missed. The recorded  $T_{DAQ,i}$  timestamps of the GPS signals are again inferred from the leading edge of each pulse and the true GPS timestamps  $T_{GPS,i}$  are queried from the run database. Since the GPS signal is only recorded with 0.1 Hz frequency, signals recorded at the very run start or end may not be encompassed by two GPS signals.

In the first science run of XENONnT 97.1 d of science data were taken of which  $\sim 8.7$  d were recorded in unlinked mode at the beginning of SR0. To make still use of the neutron and muon veto during this period,  $\sim 3.9$  d of data were successful, linked "manually" post-fact using GPS corrected timestamps. For the remaining days, the GPS timestamp information was not available.

<sup>53</sup>[https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/plugins/gps\\_syncing.py#L12](https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/plugins/gps_syncing.py#L12)

### 3.3.5 Veto-Intervals and Deadtime

Forming the veto-intervals and tagging TPC events requires several plugins as can be seen from Figure 3.8. In the first step, it must be decided which event recorded by the NV, or MV triggers a veto interval. To make the trigger logic as flexible as possible a CutList-plugin<sup>54</sup> is used. This allows to define multiple different cut plugins for the trigger logic. The final SR0 configuration uses only a single Cut-plugin called n/mVETOTrigger<sup>55</sup>. The trigger plugin requires a veto event to have a minimal charge recorded and a minimal number of contributing PMTs to trigger a veto interval. In principle, by using a CutList, also more complicated trigger conditions can be realized in the future e.g. requiring the NV event and the TPC event to be detected on the same side of the detector.

The veto intervals themselves are computed based on the synchronized timestamps discussed in the previous section, using the center time of the NV and MV events. The center-time is used as it represents a better figure of merit for the event timestamp than e.g. the event start which might be altered due to accidental grouping with a PMT dark count. This is especially important for the tight detector coincidence between NV and TPC which is used in the NR response calibration discussed in Chapter 6. The length of the veto intervals is defined by the needs of the respective task, e.g. the MV uses a veto interval of 1 ms while the NV uses a much shorter veto length which is discussed together with the exact triggering conditions in Chapter 7.

The final tagging of an event is performed by a generalized plugin called Cross-DetectorTaggingPlugin<sup>56</sup> from which the NV and MV specific veto plugins, called EventInfoTagging<sup>57</sup>, are inherited. The generalized plugin allows to define a data type that is used to tag events of another data type, e.g. to tag TPC events based on the veto intervals defined by the NV. Generalizing this plugin also allows to tag events the other way around. E.g. to tag NV events based on a single scatter NR event selection recorded by the TPC, as it is required for the calibration of the NV tagging efficiency which is discussed in Chapter 7. While the tagging logic is defined in the parent plugin the user must define two functions in the inherited plugin that determine which time intervals and which timestamps should be used for the tagging, e.g. NV veto intervals and S1 timestamps. In addition, to a boolean array that tells the user which interval tagged a signal and which signal was tagged, the plugin also provides a column for the tagged objects which contains some time information about the tagging. For all tagged events it contains the time difference to the interval start which tagged the corresponding event, and for all untagged events, it contains the time difference to the closest nearby tagging time interval.

---

<sup>54</sup>[https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/cut\\_lists/nveto.py#L5](https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/cut_lists/nveto.py#L5)

<sup>55</sup>[https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/cuts/nveto/nveto\\_science\\_data\\_cuts.py#L5](https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/cuts/nveto/nveto_science_data_cuts.py#L5)

<sup>56</sup>[https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/plugins/detector\\_coincidence.py#L105](https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/plugins/detector_coincidence.py#L105)

<sup>57</sup>[https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/cuts/nveto/nveto\\_veto\\_intervals\\_science\\_data.py#L7](https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/cuts/nveto/nveto_veto_intervals_science_data.py#L7)

To account for the livetime loss introduced due to the NV and MV a dedicated dead-time plugin<sup>58</sup> was developed together with Robin Glade-Beucke of Freiburg University. The plugin takes into account the livetime losses introduced by the two vetos systems and due to the DAQ busy logic. It accounts for the fact that veto intervals may overlap and prevents double counting of the corresponding time ranges. Further, the plugin also takes into account the artificial deadtime introduced by the DAQReader in case chunks cannot be properly split as well as an artificial deadtime introduced at each run start and end to avoid event building near the run boundaries.

### 3.4 PLUGIN PERFORMANCE AND STORAGE CONSUMPTION

The plugin performance was measured using an interactive batch shell session provided by the slurm workload manager on the midway computing cluster hosted by the University of Chicago, the default computing facility of XENONnT. Multiple sessions with a different number of 1, 2, 4, and 8 central processing unit (CPU)s, and 8 GB of random access memory (RAM)-per-CPU were used for the performance check. The test was performed using the straxer<sup>59</sup> processing script which is a python executable. The following flags were set when executing the script:

- `-context xenonnt_online` using settings like PMT gains as they are used during "online" processing.
- `-workers 1, 2, 4, or 8`, number of threads/processes
- `-notlazy`, prevents from processing in lazy mode which limits processing to a single core.
- `-profile_ram`
- `-multiprocess`, which allows to use multi-processing instead of multi-threading if enabled by the corresponding plugin. Only used by low-level plugins e.g. plugins providing `raw_records_coin_nv` and `records_nv`.

Two different kinds of tests were performed, using three different SR0 runs, 025934, 026060, 026310. Each run lasted for 60 min and was taken in `background_linked` mode. In the first test, all main data types of the NV processing chain were computed step by step using multiprocessing and 4 or 8 CPUs. For each data type, the average computing time and average storage space required were estimated. For the storage space both, the compressed and uncompressed results e.g. when loaded into memory were computed. The results are shown in Figure 3.13.

The majority of the processing time,  $(748 \pm 92)$  s/run on average when using 8 CPUs, is spent for the NV software trigger described in Section 3.3.1. Also, the highest reduction in data is achieved by the software trigger reducing the uncompressed data size from  $(95 \pm 3)$  GB/run to  $(4.5 \pm 0.2)$  GB/run on average. The event-level

<sup>58</sup><https://github.com/XENONnT/cutax/blob/c9edd0a366309d56844c48c6e6a746bbae102f06/cutax/plugins/deadtime.py#L12>

<sup>59</sup><https://github.com/XENONnT/straxen/blob/master/bin/straxer>

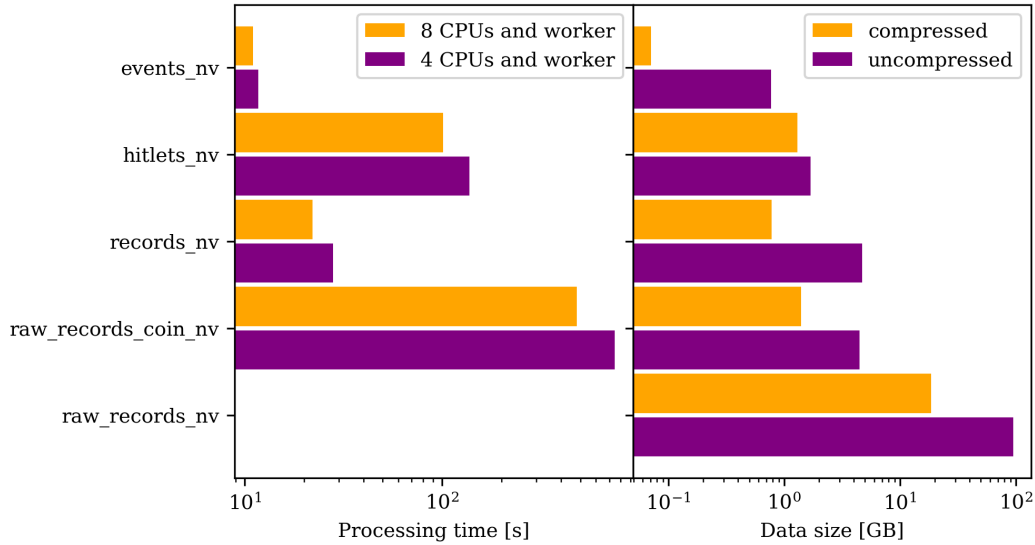


Figure 3.13. – Average performance and storage consumption of the major NV data types during SR0. The plot on the left hand side shows the processing time required to compute a certain data type for two different configurations of CPUs and worker. The performance of the DAQReader cannot be directly determined by the used method as it depends also on the data-rate and the performance of redax. The figure on the right hand side shows the data-size of the different data types when stored on disk (compressed) or being loaded into memory (uncompressed). Uncertainties are too small to be shown.

information is reduced even further to  $(768 \pm 14)$  MB/run uncompressed and  $(71 \pm 1)$  MB/run compressed on average. The large compression factor of the event-level information is given by the fact that for each event the total charge recorded per PMT is stored which will be zero for most PMTs per event. Further, it can be seen that the hitlet data type is quite large compared to events which can be understood from the fact that a single hitlet has a size of 76 B/hitlet which is almost the same as the size of a single NV event. It also has almost the same compressed size as the raw data after applying the software coincidence trigger.

In the second test, all three runs were directly processed from raw\_records\_nv to events\_nv without storing the intermediate data types as it was done during SR0. The processing speed of a single core is  $\sim 20$  MB/s/core/run<sup>60</sup> which is a factor x3 lower than for the TPC processing. However, it is already sufficient to process the data recorded during typical science data runs which produce about 12 MB/s/run. For eight CPU cores a typical processing rate of  $\sim 104$  s/run can be reached which is much larger than any data rate reached during the operation of the NV so far.

### 3.5 AUXILIARY SOFTWARE ENHANCEMENTS

In addition, to the development of the processing chain for the neutron and muon veto, many smaller changes, bug fixes, and other software projects were realized over

<sup>60</sup>MB refers here to the uncompressed data size.

the course of this thesis. These changes include modification to all three analysis packages, STRAX, STRAXEN, and CUTAX. In the following, a few of these smaller changes are briefly described. In the subsequent section, an example that can be visualized more nicely is elaborated in more detail.

- **Superruns:** As discussed in Section 3.1.2, the data-taking in XENONnT is organized in runs that have a typical duration of 30 minutes to one hour during SR0. Due to the very good data reduction when computing events from raw data, typical runs are only a few MB in size. To make the I/O operations of loading these small files faster, and to help analysts to organize the data in a human-readable format, the development of superruns was continued during this work. Superruns<sup>61</sup> allow the user to define new "runs" based on a list of runids. In this way data that is taken over multiple days can be grouped. In addition, superruns allow to rechunk the event-level data across multiple runs which makes I/O operations faster, especially when data quality selections are applied during the loading process.
- **Slow control python client:** To be able to access the parameters of XENONnT's slow control system within python, a small python client was developed which is called SCADAInterface<sup>62</sup>. The client access data stored in a Historian database through an already existing web server application. The different sensors can be accessed through their registered sensor names in the SCADA control software of XENONnT's slow control system. During data requests, the user can decide if raw data, interpolated values, or averaged values should be queried. In addition, the frequency of the data points and the time interval in which data should be queried can be adjusted. The time intervals can be either queried as time intervals in UTC or by runids to compare the detector conditions with data. The latter is especially used in XENONnT's online monitor client<sup>63</sup> which serves regularly plots of the current detector conditions to XENONnT's slack groups.
- **EPIX:** EPIX<sup>64</sup> (Electron and Photon Instructions generator for XENON) is a small side project which was developed together with Diego Ramirez from Freiburg University and Henning Schulze from Münster University. It is a tool that converts energy deposits simulated with Geant4 into a number of quanta in the form of S1 photons and S2, electrons which serve as an instruction for XENONnT's waveform simulator<sup>65</sup>. The quanta are generated after clustering and classifying the initial energy deposits according to their respective recoil types. The quanta generation is either based on Nest[Szy+11; Far+23] or the XENONnT's own LXe response model [Apr+19b], which is adjusted for small drift fields and used in SR0.

<sup>61</sup><https://straxen.readthedocs.io/en/latest/tutorials/SuperrunsExample.html>

<sup>62</sup><https://straxen.readthedocs.io/en/latest/tutorials/ScadaInterfaceExample.html>

<sup>63</sup>[https://straxen.readthedocs.io/en/latest/online\\_monitor.html](https://straxen.readthedocs.io/en/latest/online_monitor.html)

<sup>64</sup><https://github.com/XENONnT/epix>

<sup>65</sup><https://github.com/XENONnT/WFSim>

### 3.5.1 Interactive event-displays and plots

When working with a new detector for the very first time waveform watching is an essential tool to optimize signal reconstruction or to study yet still unknown detector backgrounds. In python, the by far most used plotting package is called `matplotlib`<sup>66</sup> which offers different methods to plot data in different static representations. While `matplotlib` is a very strong tool to display static information, it can be quite cumbersome when applied to waveform watching which is especially true for LXe TPCs. In XENONnT, detector signals span about five orders of magnitudes in time, ranging from S1s signals in the order of 100 ns over S2s which are typically in the order of 10  $\mu$ s to events lasting for 2 ms. Hence, plotting events in a static display would lead to a loss of visual information. For this reason a set of novel python packages called `bokeh`<sup>67</sup>, `panel`<sup>68</sup> and `holoviews`<sup>69</sup> were used to develop dynamic event display for the TPC and NV over the course of this thesis. The packages are designed to be used along web-served applications like `jupyter-notebooks` or `jupyter-lab` which are the most commonly used tools for data analysis in XENONnT.

The dynamic event-display<sup>70</sup> for the TPC is mainly based on the `bokeh` package. An example waveform for a potential NR multi-scatter (MS) event is shown in Figure 3.14. The display itself is composed of 4 to 6 individual plots, depending on the users wishes. At the top of the display basic information about the event like its `runid`, event number per run, and time when it was recorded is shown. The top right corner shows widgets for different interactive tools which are explained in the context of the individual subplots in the following.

---

<sup>66</sup><https://matplotlib.org/>

<sup>67</sup><https://bokeh.org/>

<sup>68</sup><https://panel.holoviz.org/>

<sup>69</sup><https://holoviews.org/>

<sup>70</sup>[https://github.com/XENONnT/straxen/blob/62e1916e9fec656715e9a1c82743f1f1934e1012/straxen/analyses/bokeh\\_waveform\\_plot.py#L18](https://github.com/XENONnT/straxen/blob/62e1916e9fec656715e9a1c82743f1f1934e1012/straxen/analyses/bokeh_waveform_plot.py#L18)



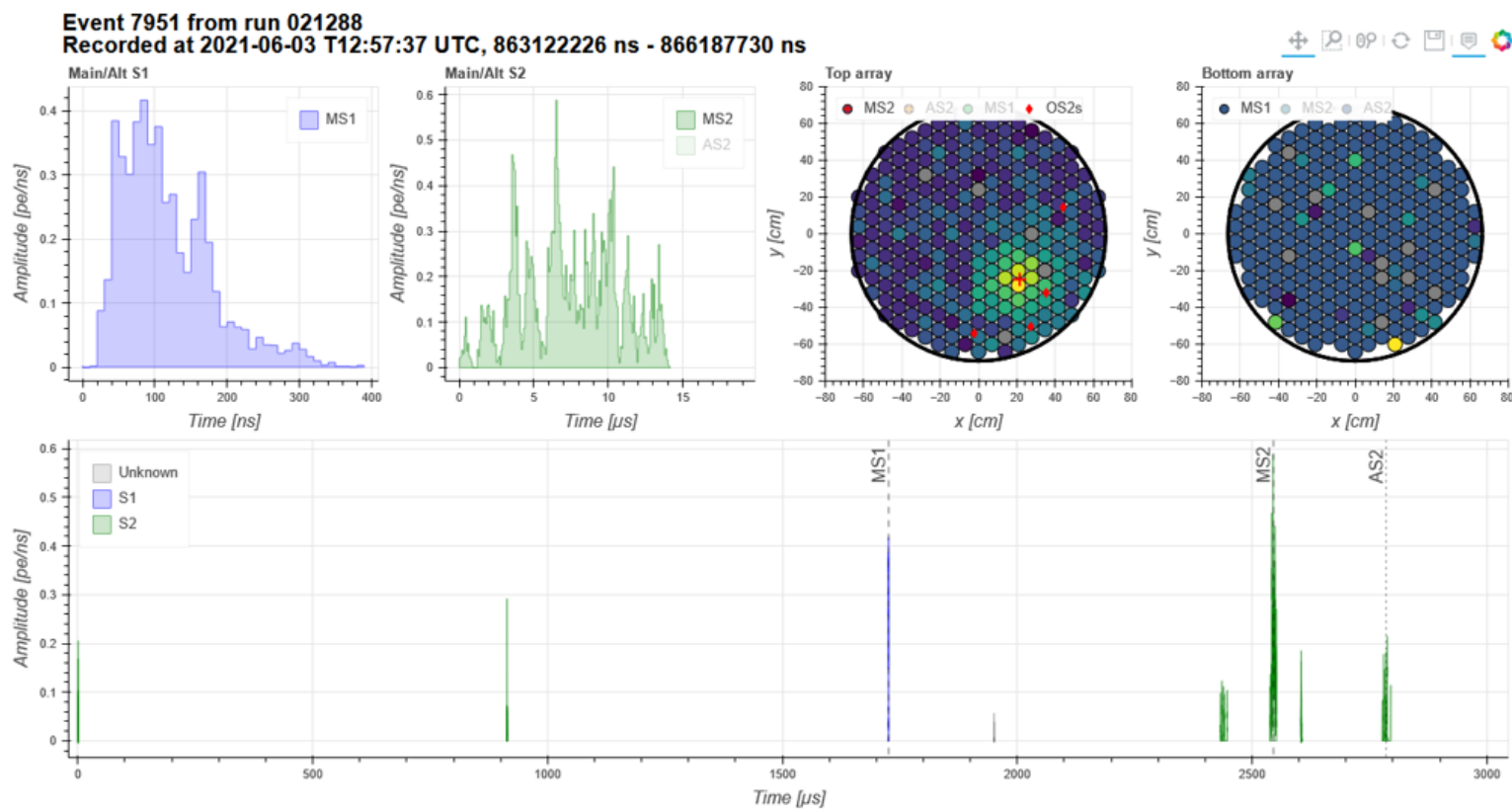


Figure 3.14. – Interactive TPC event display. The shown event corresponds to a NR calibration event during the SR0 NR calibration. The different figures and features are explained in the text.



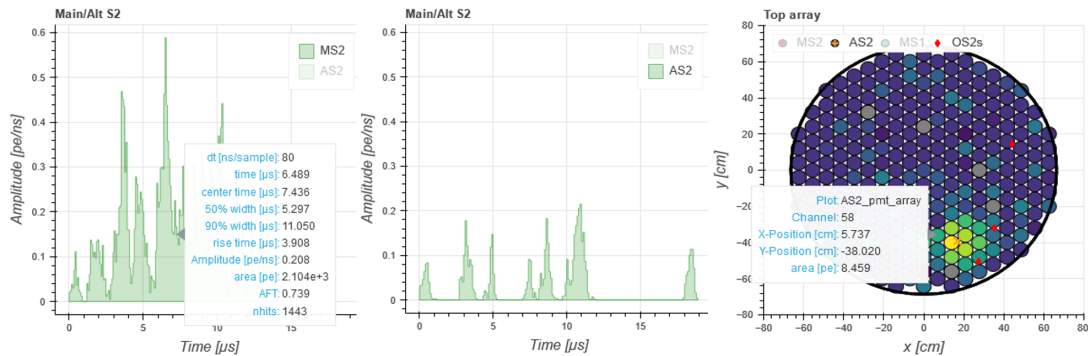


Figure 3.15. – Main and alternative S2 waveform as well as the alternative S2 hit-pattern for the same event as shown in Figure 3.14. The first figure shows the mouse over tool-tip when hovering over the peak waveform. The second figure shows the alternative S2 which is displayed when selecting it via the figure legend. The third plot shows in addition to the alternative S2 hit-pattern the mouse over information for one of the displayed PMTs and the reconstructed position of the alt. S2 as an orange cross.

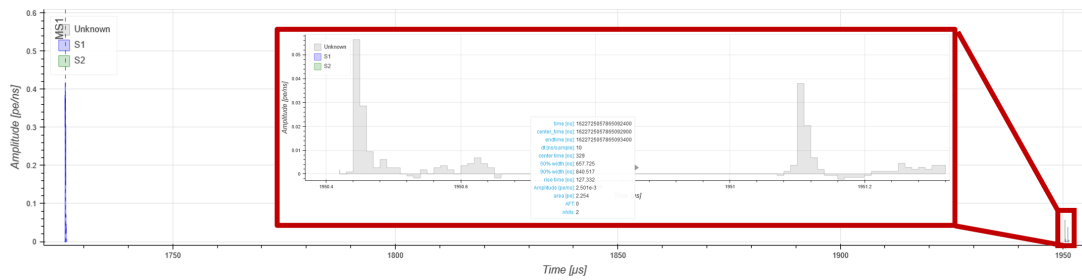
The upper row of plots shows information about the main (M) and alternative (A) S1/S2-signal. The first two plots are showing their waveform respectively. The binning used to display the S1- and S2-signals depends on their down-sampled representation explained in Section 3.2. The third and fourth plot are showing the top and bottom hit-pattern of the individual signals as well as the reconstructed main/alt. S2 position. The color-scale of the individual PMTs represent the integrated charge of the given channel for the displayed S2-signal. The color-scale ranges from dark blue over green and cyan to yellow<sup>71</sup>. PMTs which were not operational during the run or were excluded from analysis are displayed in gray.

The bottom display in Figure 3.14 shows the complete event with all peaks found within its range. The user can decide whether only the main, main and alternative or all peaks of a category should be shown. Horizontal dashed lines indicate the main/alt. S1 or S2 peak within the event. Optionally, the event display offers to plot the so called record matrix beneath the event display. The record matrix depicts all raw\_records recorded by the individual PMT channels as function of time. Figure B.7 shows the full event display including the record matrix and can be found in the appendix on page 241.

Figure 3.15 shows some of the interactive features exemplary for the S2 waveform and hit-pattern plot. All plots of the display feature an interactive legend allowing the user to change between the main/alt. S1 or S2 signal. The same is true for the different peak categories in the main panel of the display. In addition, the user can make use of a mouse over tool to get further information about any of the display signals or markers as illustrated in Figure 3.15.

Especially in the large event panel at the bottom of Figure 3.14 the zoom feature provided by bokeh comes in handy. The XENONnT event display uses two types of

<sup>71</sup>A color-scale was deliberately removed to have more space available. The information of the color-scale can be recovered through an interactive mouse over feature which displays the recorded charge in the corresponding channel as presented in Figure 3.15



**Figure 3.16.** – Illustration of the zoom and panning features of the XENONnT TPC event display. The plot shows a zoom in of the bottom panel of Figure 3.14 between 1700  $\mu\text{s}$  and 2000  $\mu\text{s}$ . The inlet plot shows a zoom onto the very small unknown peak about 200  $\mu\text{s}$  after the main S1. Using the zoom feature in addition to the mouse over tooltip reveals that the unknown peak following the main S1 is composed of two individual single photo-electron PMT signals.

zoom. A simple mouse-wheel zoom which zooms along the time axes to the position of the cursor, or a box zoom allowing to mark a certain region to zoom into. Besides zooming, bokeh also allows to add a panning feature to all plots making it for the analysts possible to move the field of view in different directions or just along specific axes. The working principle of these tools is illustrated in Figure 3.16.

To provide all these interactive features, the bokeh package is supplemented by a JavaScript package called BokehJS. BokehJS handles all the rendering and visualization as well as the user interactions with different tools and widgets via call-back methods written in Java. This has the advantage that the event display can be exported in HyperText Markup Language (HTML)-format keeping all its interactive features. Further, it can be embedded on other websites like for example the XENONnT wiki which is used by analysts for documenting their work. However, to compute and plot new events an active python kernel/interpreter is required, like in a jupyter-notebook session or served via an external python server.

In straxen, the TPC event-display is embedded as a so-called `straxen.mini_analysis`. A mini-analysis can be seen as a wrapper around a regular python class method or function providing the required data information specified by the developer based on user input. The user only needs to specify for which run and time range an event should be plotted and straxen will load the required data and provide the plot. This has the advantage that the syntax is easily recognized by the user as it is equivalent to loading data, but is quite cumbersome and slow when plotting many different events across multiple runs as data is loaded during run-time. On average it takes about 5-8 s to plot a single event depending on the number of peaks within an event. Two seconds are due to the rendering of the display itself.

Based on this lesson learned, the NV event-display was developed using `holoviews` which uses `bokeh` to render plots, but provides additional features to easier streamline larger data sets via python-callbacks. Compared to `bokeh`, `holoviews` does not require a running external server for this purpose, but uses the already running python kernel in a jupyter-notebook session. In addition, `holoviews` allows to write custom python-call back functions which are used to provide some of the features for the NV event display.

Event 94 from run 020400

Recorded at (2021-05-26 T07:40:19) UTC 143508584 ns - 143508846 ns

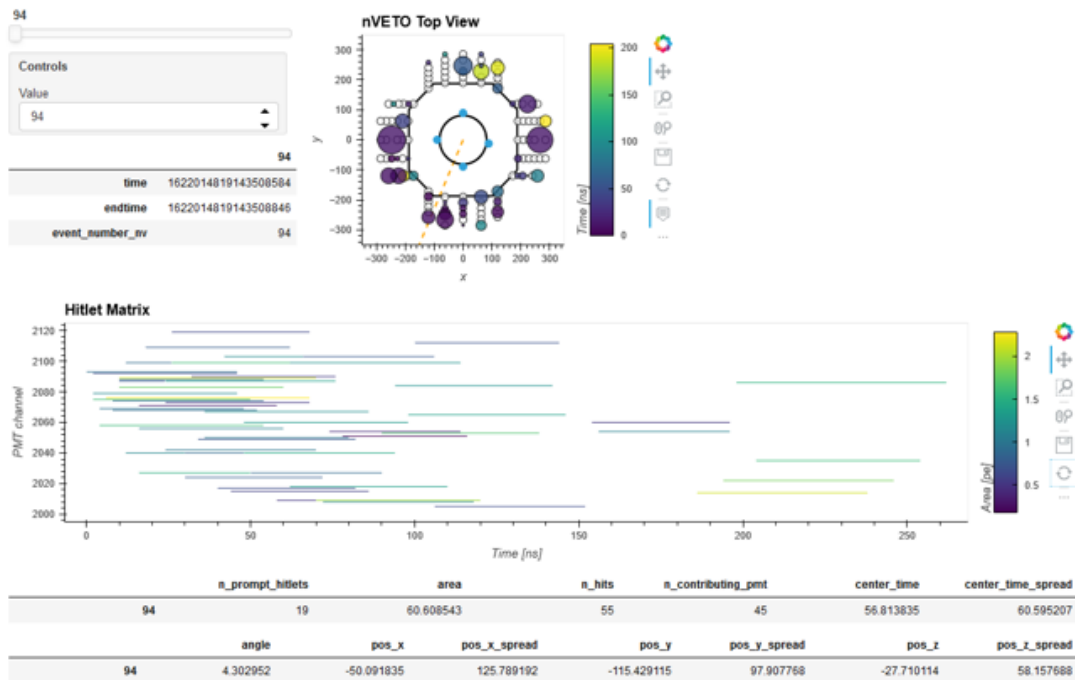


Figure 3.17. – Dynamic NV event display. The shown event was recorded during the SR0 AmBe calibration and represents most-likely an event caused by a 4.4 MeV gamma emitted by the source. An explanation of the different parts of the display is given in the text.

An example of the NV event display is shown in Figure 3.17. It is composed of multiple components which are put together into a single display using the panel package. Just like for the TPC display, the NV event display shows in its headline some basic information about the data like its run- and event-id as well as the time when it was recorded. The display itself is structured in three rows.

The first row shows a slider and control field which allows the user to interactively change the displayed event without rerunning the underlying code by themselves. Below the control panel, a small table shows the start- and end time of the currently selected event as well as its event number. The time units are in nanoseconds unix-time. On the right-hand side of the first row, a spatial representation of the NV is plotted. The black octagonal line indicates the position of the outer NV reflector walls. The inner black circle indicates the outer TPC cryostat. The small light blue circles represent the four diffuser balls. The orange dashed line shows the reconstructed event angle discussed in Section 3.3.3. The 120 PMTs of the NV are displayed as colored or white circles of different sizes. The z-component within each PMT-row is projected onto the xy-plane such that the lowest PMT-row is on top of the black panel reflector line. The size and color of the PMT-circle depends on the integrated charge and time when the first hitlet was observed for the currently shown time range in the hitlet matrix. Larger PMT-circles represent a higher integrated charge. White circles represent

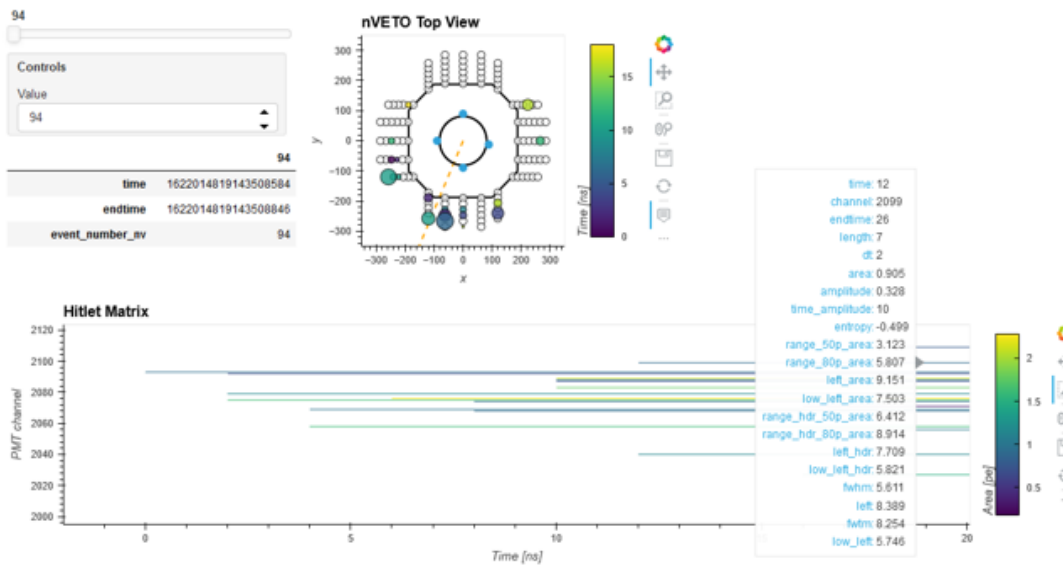


Figure 3.18. – Excerpt of the NV event display as shown in Figure 3.17 after zooming in the hitlet-matrix to the first 25 ns of the event. In addition the mouse over tool tip of a single hitlet is shown. Changing the x-range triggers a recomputing and re-rendering of the PMTs in the spatial distribution plot.

PMTs which have not seen any signal within the selected time range.

The second row is given by the hitlet matrix which shows the individual hitlets seen by each PMT channel as a function of time. Each dash indicates a single hitlet whose color represents its total charge. Below the hitlet matrix, the third row shows a table with base information about the currently selected event.

Just like the TPC event display the NV event display offers various interactive tools. All markers, like PMTs, diffuser balls, and hitlets feature a mouse over tool to display extra information when hovered above. In addition, both plots provide the same zooming and panning features as TPC event display. In contrast to the TPC event display, the NV display features two custom-made python-callback functions which are used together with `holoviews DynamicMap`-method. The first callback changes the data shown in the hitlet matrix when moving the slider or setting a different event number. A change in the hitlet-matrix triggers the second python-callback function which redraws the spatial distribution plot of the NV based on the shown data. This second callback is also triggered when the x-range of the hitlet matrix is changed as it is shown in Figure 3.18. In this way different events are rendered much faster compared to the TPC display. The initial rendering of the display takes about one to two seconds, while changing an event using the slider, or changing the plotting range takes less than 100 ms.

As an outlook, during the writing of this thesis, a new version of the TPC event display is being developed<sup>72</sup>. A beta version of the new display is shown in Figure 3.19 and will make use of some of the features developed for the NV event display. Among those features is the event selection via a slider bar. Also, all other previously

<sup>72</sup><https://github.com/XENONnT/straxen/pull/1077>

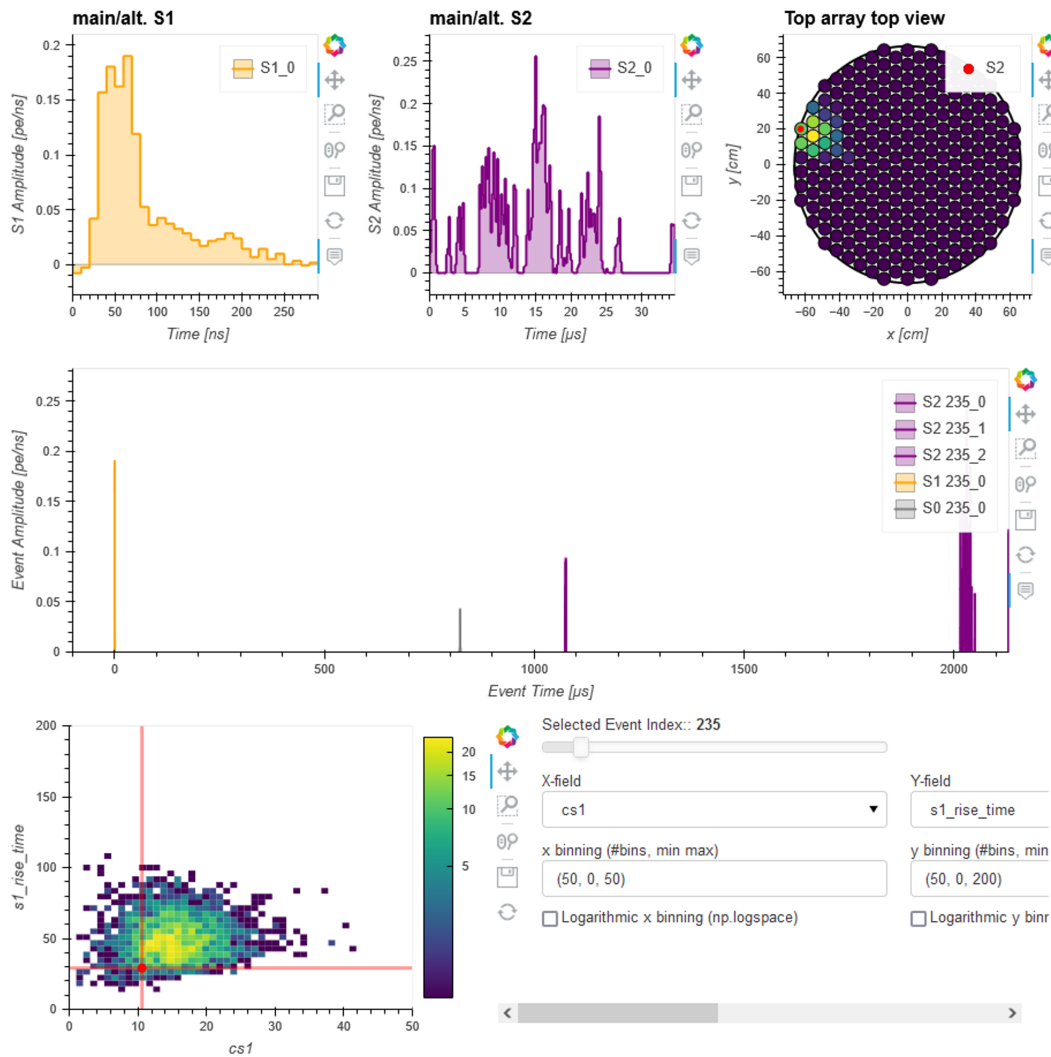


Figure 3.19. – Proof of concept of the new XENONnT TPC event display. Compared to the old TPC event display, the new display has some extra features which are explained in the text. The display shows an event taken during the  $^{37}\text{Ar}$  calibration of SR0.

discussed interactive features are ported to the new display.

In addition, these old features are augmented with a new inspection panel at the bottom of the display which improves the quality of the waveform inspection even further. This additional inspection plot uses a set of different callbacks to display the currently selected event in a two-dimensional histogram showing all events passed by the user. Several widgets next to the inspection plot can be used to dynamically change the information shown in the histogram as well as its binning and plotting range. The selected event itself is displayed as a red dot with a cross-hair to guide the users in case the selected event is outside of the shown plotting range. The color scale indicates the number of entries per bin and can be changed between a  $\log_{10}$  and a linear scale via a widget. As for the NV display, the user provides a selection of events and peaks directly to the display class. However, to decouple the waveform watching

of specific events from events that are shown as a reference in the binned inspection panel, the user does not have to pass the peak-level information for all the entered events. Only events which have additional peak-level information will be accessible via the event selection slider and rendered in the main panel, while all events are presented in the binned histogram of the inspection panel.

Besides these new features also the code is being restructured such that both displays inherit commonly used features from a more general class. Currently, the display is in a preliminary beta status as some of the features are not working as expected from `bokeh` and require input from the `holoviews` developers. The performance of the new display is much better. The initial rendering of the display takes about two to three seconds. Changing between events takes between 100 ms and 500 ms, depending on the complexity of an event.

### 3.6 CONCLUSION AND OUTLOOK

In this chapter, the XENONnT analysis software was discussed. As part of this thesis, the entire processing chain for the neutron and muon veto was developed, implemented, and tested. The processing software provides the fundamental analysis data for both veto systems in XENONnT. In addition, the plugins which are used for the NV-TPC coincidence discussed in Chapter 6 and 7, are explained. The processing software has good performance and allows, in case of the NV, to process about 20 MB/s using a single CPU core, and  $\sim 100$  MB/s using eight cores. Both are well above the average data rates recorded during science data taking and calibration of the NV. On average the NV records about 20 GB/h (compressed) during science data-taking which totals to about 175 TB/year. Furthermore, the Chapter gives an overlook of several other, smaller tools, which were developed and implemented during the presented work. As a highlight, the NV and TPC event displays were shown. The developed tools are used by many analysts in XENONnT and contribute to the overall performance and progress of the experiment.

During the analysis of the NV data for the first science run, which will be discussed later in this thesis, it has become clear that some of the veto data types in their current form are not needed. The `hitlet` data type defines many signal properties for the individual PMT pulses which are mostly irrelevant for the analysis of the NV data. Thus, it is planned to update the processing chain accordingly. First of all, it is advisable to reduce the amount of information stored within each `hitlet` to a minimum. Instead, it is better to extend the event level information for the vetos by adding a summed waveform as well as additional parameters, which describe the shape of the summed signal. This will help to improve the overall signal to background discrimination for the NV. In addition, by reducing the size of individual `hitlets`, the overall storage usage by the NV can be reduced. This can be further augmented by using the NV software trigger as originally intended and starting to delete the `raw_records` data type. This should be fairly safe as the data analysis for SR0 proved the reliability of the software trigger. In this way, the overall data rate would be reduced by a factor of about  $\times 10$ .





---

## GADOLINIUM SULFATE

Gadolinium has two properties that make it an attractive element for the detection of neutrons: It releases a cascade of  $\gamma$ -rays with a total energy of  $\sim 8$  MeV after undergoing a neutron capture, and it has the highest neutron capture cross section among all stable elements. Thus, Gadolinium is foreseen in a variety of different experiments like Super-Kamiokande-Gd [Mar+20; Abe+22], Anni [Fis+20] or LUX-Zepplin (LZ) experiment [Ake+20], to enhance their neutron detection capabilities. Also the new water Cherenkov neutron veto of XENONnT will be eventually doped with Gd in form of gadolinium sulfate octahydrate ( $\text{Gd}_2(\text{SO}_4)_3 \cdot 8(\text{H}_2\text{O})$ ). The procurement of the Gd-sulfate was a shared effort between the university of Tokyo in Japan and the JGU-Mainz. As part of the presented work different Gd-sulfate samples of the Austrian company *Treibacher Industrie AG*, were analyzed for their suitability, and compared against the properties of the Japanese sample which were subjected to a strict selection process for the EGADS demonstrator [Mar+20] and the Super-Kamiokande Gadolinium upgrade [Abe+22].

The chapter is organized as follows: Section 4.1, briefly summarizes the properties of Gadolinium-sulfate, which is followed by a short introduction into transparency measurements using a UV/VIS spectrometer in Section 4.2. The quality assessment of the different salt samples, including the results of transparency measurements, are summarized in Section 4.3 of this chapter. The chapter closes with a short conclusion and outlook for the Gd-loading in XENONnT.

### 4.1 GADOLINIUM-SULFATE PROPERTIES

Gadolinium is a rare-earth metal with the atomic number  $Z = 64$  and has the highest neutron capture cross section among all stable elements of  $\sim 49$  kbarn [Mar+20]. The neutron capture on Gd is dominated by the two isotopes  $^{155}\text{Gd}$  and  $^{157}\text{Gd}$  which have a relative abundance of 14.8% and 15.65%, and a thermal neutron capture cross section of 60.9 kb and 254 kb, respectively [Mar+20]. The neutron capture on both isotopes releases a cascade of  $\gamma$ -rays with an endpoint energy of 8.536 MeV and 7.937 MeV [Tan+20]. In comparison, hydrogen (water) has a neutron capture cross section of only  $\sim 336.3$  mb and emits just a single 2.22 MeV  $\gamma$ -ray. Thus, already small traces of



Gd diluted into water allow to enhance the neutron detection efficiency of a water Cherenkov detector significantly. E.g. about 0.2 % of dissolved Gd in water<sup>1</sup> are sufficient to capture ~90 % of all neutrons on Gd [Mar+20]. A figure showing the capture fraction on Gd as a function of concentration is given in the appendix on page 242.

The doping process of water with Gd is not a simple task. In its natural form as Gadolinium-oxide ( $Gd_2O_3$ ), it is not soluble in water. Different types of Gd-salts on the other hand are known for their solubility in water and are used in different types of industrial applications. Among them are Gadolinium nitrate  $Gd(NO_3)_3$ , Gadolinium chloride  $GdCl_3$  and Gadolinium sulfate  $Gd_2(SO_4)_3$ . While the latter two can be in principle used in water Cherenkov detectors, Gadolinium nitrate is mostly opaque in the UVA region [Mar+20]. Gadolinium chloride has the disadvantage that it is more reactive than sulfate and is thus disfavored due to technical constraints. Although to a much lesser degree, Gd-sulfate in the form of  $(Gd_2(SO_4)_3 \cdot 8(H_2O))$  diluted into water also leads to a reduction of the pH-value which is an issue for various detector materials. The feasibility of this technology was proven by the EGADS experiment [Mar+20] which also provided valuable inputs for the Gd-loading of the XENONnT NV. About 2/3 of the Gadolinium sulfate which will be added to the water of the XENONnT water tank is provided by the same company as the salt used in EGADS, which will be referred to in the following discussion as "Japanese salt".

## 4.2 TRANSPARENCY MEASUREMENTS

Due to the faint nature of Cherenkov signals, one of the most important properties of a water Cherenkov detector is its water transparency. To ensure that the Gd-loaded water, based on the salt procured from Treibacher, has a similar high optical transparency as the Japanese salt, each salt sample was tested using a UV/VIS spectrometer Lambda 850 of the manufacturer PerkinElmer. The working principle of the spectrometer is illustrated in Figure 4.1. The apparatus is equipped with a tungsten-halogen and deuterium lamp. The light of the selected lamp is guided through a mirror system onto a monochromator prism which splits the incoming light into mono-chromatic beams using an adjustable slit. Afterwards, the monochromatic light is guided with a second mirror system through a beam chopper before entering the sample compartment. The compartment can hold two samples one of which serves as a reference standard. After passing the two samples the transmitted light is measured by a PMT or a polycrystalline lead sulfide detectors (PbS) detector in case of infrared light. The absorbance  $A$  of the main sample is computed by comparing the measured relative light intensities which pass the sample with respect to the reference line. Thus, the measured absorbance is given by

$$A(\lambda) = \log_{10} \left| \frac{\phi_R(\lambda)}{\phi_S(\lambda)} \right| \quad (4.1)$$

where  $\phi_i$  represents the light flux through the sample S and the reference R.

The absorbance of liquids can be measured through special glass cuvettes which can be mounted to the sample and reference line of the sample compartment. Their

<sup>1</sup>All percentages will refer to the mass percentage of the given material.

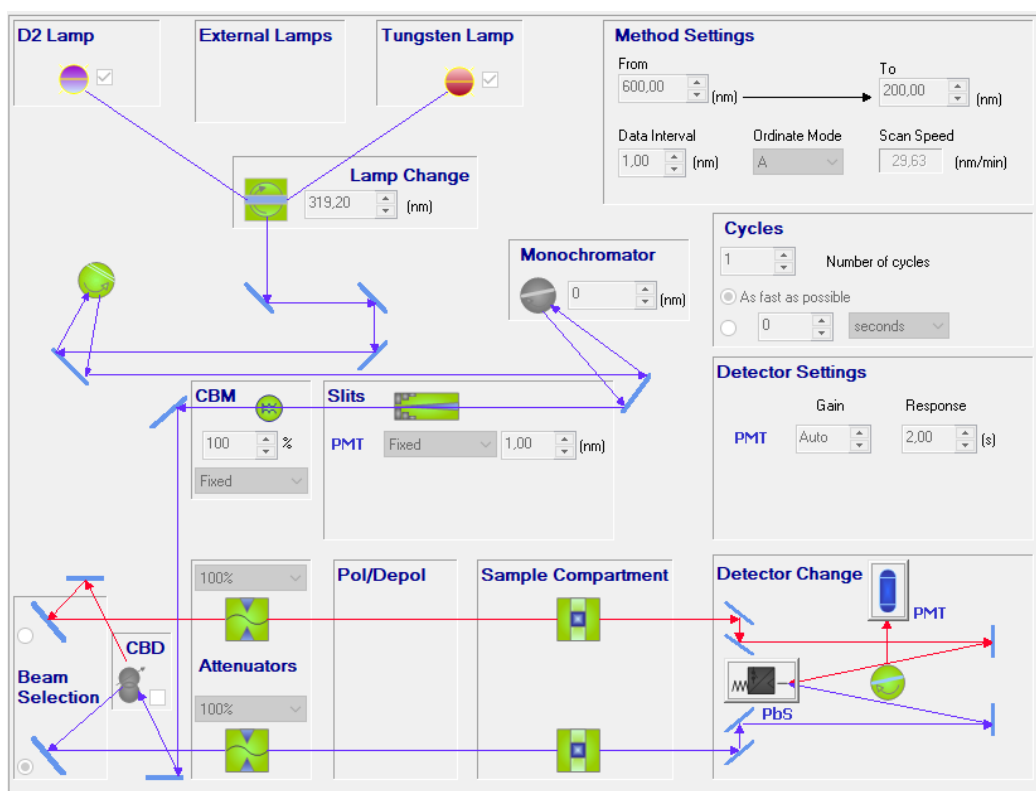


Figure 4.1. – Settings screen of the UV/VIS spectrometer Lambda 850 of the manufacturer PerkinElmer. The different components are explained in the text.

maximum length is determined by the corresponding spectrometer and is limited to 100 mm for the Lambda 850. Cuvettes can be bought in pairs, to guarantee that they are made of the same base material which ensures that they have similar optical properties. This is important as the transmittance  $T$  through the cuvette would otherwise bias the result. For a liquid sample light flux through the reference or sample line of the device can be rewritten as

$$\begin{aligned}\phi_i(\lambda) &= \phi_0(\lambda) \cdot T \\ &= \phi_0(\lambda) \cdot T_C \cdot T_M \cdot (1 - R_{C,M})\end{aligned}\quad (4.2)$$

where  $\phi_0(\lambda)$  is the light flux emitted by the lamp,  $T_C$  and  $T_M$  the transmittance through the glass windows of the cuvette and liquid sample respectively and  $R_{C,M}$  the total reflectance of the incoming light at transitions between different media e.g. the cuvette  $C$  and filled medium  $M$ .

For the measurements conducted during this study only a single 100 mm *Suprasil* quartz<sup>2</sup> cuvette was used, and samples were measured with a blank reference line. Tests with two cuvettes showed a systematic bias in the measurements as the absorbance of the cuvettes is in a similar order of magnitude as the absorbance of the

<sup>2</sup>Quartz glass was used as regular glass is not transparent for UV light.

water sample itself, and thus already small deviations between the optical properties of the two cuvettes lead to a measurable effect. Using Equations (4.1) and (4.2), one can show that, if measured in series, the transmittance of Gd-loaded water relative to ultra-pure water can be estimated via

$$\frac{T_{Gd}}{T_{H2O}} = 10^{-(A_{Gd} - A_{H2O})} \quad (4.3)$$

assuming that the refractive index of Gd-loaded water is the same as for ultra-pure water.

For each measurement, a well-defined sample preparation and accurately defined measurement procedure are important to ensure unbiased results, due to the very high transparency of ultra-pure water. Preceding each measurement, the cuvette was cleaned extensively with isopropanol and flushed two times with ultra-pure water before filling any new sample. The samples themselves were mixed from the corresponding salt and ultra-pure water provided by the in-house chemist of the institute for physics of the JGU-Mainz, Dr. Strübig. It has a conductance of  $0.055 \mu\text{S}/\text{cm}$  which is well below the conductance of the demineralized water inside the water tank of XENON1T/XENONnT<sup>3</sup>. The very same ultra-pure water also served as a reference standard for the relative transmittance measurement. All samples were mixed in sealable glass bottles with a mass concentration of  $(0.20 \pm 0.01) \%$  Gd which corresponds to the design goal of the XENONnT NV [Apr+20b]. The mixing was performed without a stirrer by placing the bottles inside an ultra-sonic bath for half an hour followed by an additional recovery time of at least another hour. Longer stirring times should be avoided as an increase in water temperature leads to a smaller solubility of the Gd-sulfate. The weighing of the water was done using a regular scale with gram precision, while the salt was dosed with a microgram scale. The amount of solution for the different samples varied between 250 ml and 350 ml depending on the analysis task. The filling of each sample into the cleaned cuvettes was done through a one-time use pipette to avoid any contamination and the formation of air bubbles inside the liquid which would increase the absorbance of the sample.

Not only the usage of two different cuvettes can lead to a systematic bias in the absorbance measurement, but also the positioning and orientation of a single cuvette. A slight misalignment of the cuvette in the compartment holder may affect the measurement. To account for this effect, each sample was measured four times and the cuvette has been removed and placed again between measurements of the very same sample. The resulting absorbance is then estimated as the average over these measurements with the standard deviation as a measure for the corresponding uncertainty. To check the time stability of the light flux of the device, a baseline measurement was conducted between two samples by leaving the reference and sample line blank. The settings of each measurement are shown in Figure 4.1. The absorbance was measured between 200 nm and 600 nm with a stepping of 1 nm. The exposure was chosen to be 2 s/nm to minimize statistical uncertainty with respect to the explained systematic effects. The relevant results of the individual salt samples are summarized in the subsequent sec-

<sup>3</sup>See Section 2.4.3

tion.

### 4.3 GADOLINIUM SULFATE PROCUREMENT

In total four different salt samples were produced by the Austrian company *Treibacher Industrie AG*, and tested in Mainz in 2019/2020, before the Covid lockdown in Germany started. Each test sample was about 1 kg of salt which was produced by the laboratory test facility of Treibacher as the main production line can only produce batches of larger sizes. The individual samples were tested for their pH values and transparencies. In addition, radio-purity screenings of the first sample and the final production batch were conducted at LNGS through an ICP-MS<sup>4</sup> measurement and a high purity Germanium  $\gamma$ -ray screening. The different samples are listed in Table 4.1.

The absorbance measurement of the first Treibacher sample, GdSu19-04 is shown together with the absorbance of the ultra-pure water and the Japanese salt in Figure 4.2. Except in the region of small wavelengths below 250 nm, both salts show a similar absorbance. However, the pH-value of the Treibacher salt, listed in Table 4.1, is much lower which may lead to technical problems when detector components are exposed to the Gd-loaded water for a longer time<sup>5</sup>. Thus, Treibacher also produced a variety of additional samples changing the ratio between Gd-oxide and sulfuric acid to create solutions with a larger pH value (Gd19-20, Gd20-01, Gd20-01A). However, all these salts were not fully soluble and partially opaque as can be seen based on the photographs shown in Figure B.9 in the appendix on page 243. Already, by eye, it can be seen that larger ratios of Gd-oxide over sulfuric acid (larger GOSE) lead to a more

<sup>4</sup>Inductively Coupled Plasma - Mass Spectrometry

<sup>5</sup>The pH-value for a solution with a concentration of 2 % is most likely to be even lower.

Table 4.1. – Summary table of Gd-sulfate samples provided by Treibacher. GOSE represents the fraction of Gd-Oxide used during the production, pH the pH-value for the corresponding Gd-loaded water solution, and <sup>238</sup>U, <sup>232</sup>Th and <sup>56</sup>Fe for the concentration of the respective element measured through ICP-MS. The pH-values of the Japanese salt was measured for a Gd concentration of 0.2%, while the pH-values for the Treibacher salts were provided by Treibacher who used a smaller concentration of only 0.16 % of Gd (0.4 % of Gd-sulfate octahydrate).

Sample	GOSE	pH	Th [ppb]	U [ppb]	Fe [ppb]
NYC salt	-	(5.0 ± 0.2)	<0.05	<0.05	(350 ± 100)
GdSu19-04	43.8	(3.1 ± 0.1)	(0.40 ± 0.05)	(0.08 ± 0.02)	(400 ± 100)
GdSu19-20	48.8	(6.5 ± 0.1)	-	-	-
GdSu20-01	48.7	(6.6 ± 0.1)	-	-	-
GdSu20-01A	48.5	(6.3 ± 0.1)	(0.7 ± 0.2)	< 0.05	(250 ± 100)
Gd-batch1	-	-	(30 ± 10)	(4 ± 1)	(500 ± 100)
Gd-batch2	-	-	(44 ± 15)	(2 ± 1)	(500 ± 100)
Gd-oxide	-	-	(50 ± 15)	(175.0 ± 52.5)	-
GdSu20-17	-	-	(31 ± 9)	(33 ± 10)	(350 ± 100)

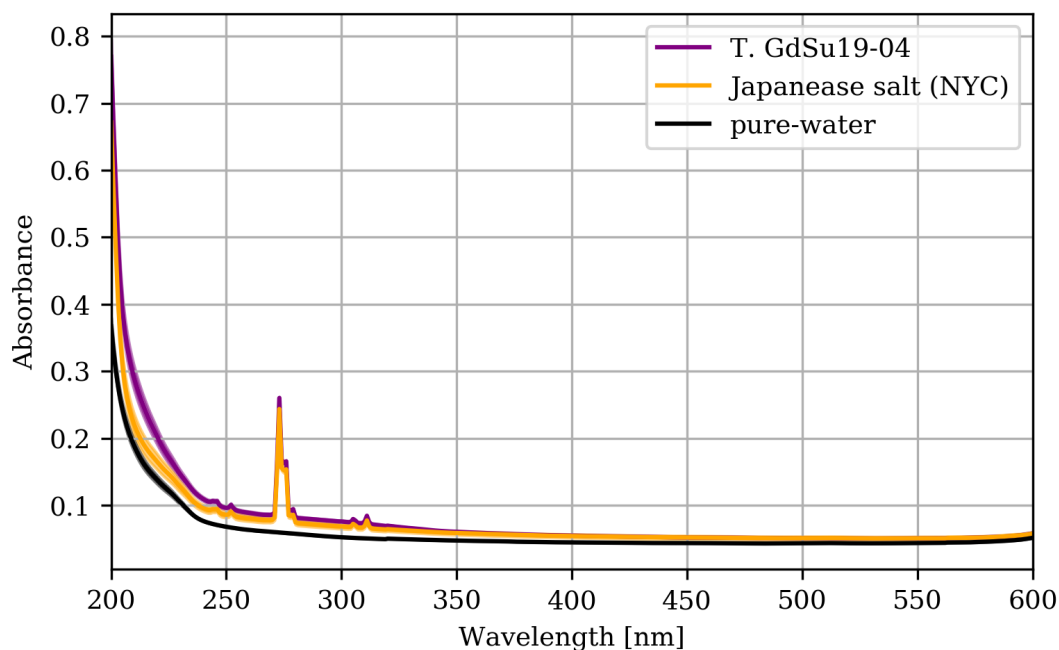
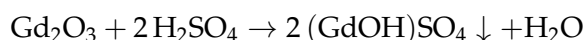


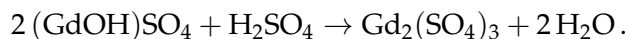
Figure 4.2. – Absorption spectra of the GdSu19-04 sample with respect to ultra pure water and the Japanese Gd-sulfate salt. The solid lines show the mean and the shaded regions the standard deviation of the corresponding measurements as discussed in the previous section. The absorption peaks below 350 nm are characteristic absorption lines of  $Gd^{+3}$  ions [Fis+20].

opaque liquid.

The chemical engineer of the PRISMA detector lab, Dr. Anastasia Mpoukouvalas, pointed out that this observation might be due to the formation of molecules of the hydroxyl group ( $OH^-$  group). These molecules form if the amount of sulfuric acid is not enough to fully transform the  $Gd_2O_3$  into  $Gd_2(SO_4)_3$ . The reaction can be written down as



where the arrow indicates that the  $(GdOH)SO_4$  falls out from the solution. By adding more sulfuric acid the hydroxyl molecules can be transformed into Gd-sulfate through the reaction



To confirm this hypothesis a titration of the second salt sample provided by Treibacher, GdSu19-20, was performed together with Dr. Strübig and Dr. Mpoukouvalas. In this titration droplets of  $\sim 1$  ml of a 0.01 N sulfuric acid were added to a 0.2 % Gd-solution. The solution was constantly stirred and the pH value was measured 5 minutes after adding a droplet. Around a pH-value of  $(5.35 \pm 0.30)$  the solution became completely transparent to the human eye. In addition to this fast titration, 5 sample solutions with a dedicated amount of sulfuric acid were prepared, to measure the pH value after a longer reaction time of one day. The results of the titration and for these 5 sample solutions are shown in Figure B.10 in the appendix on page 244. Four samples showed

after one day an increased pH value of about 5.8 compared to the values measured directly during titration. The pH value of the sample with the highest amount of sulfuric acid dropped to a pH value of 3.6. These qualitative observations are in agreement with the Treibacher salt samples as the first sample with the lowest pH value also had a smaller Gd-oxide to sulfuric acid ratio than the other samples. Filtering the opaque solution without adding any sulfuric acid leads to a pH value of 5.501. The correlation between pH value, opaqueness, and the formation of hydroxyl-molecules was also observed by other groups and is also qualitatively described in [Fis+20].

Based on these observations Treibacher produced the last two samples GdSu20-01 and GdSu20-01A reducing the GOSE fraction for a better Gd-oxide to sulfuric acid ratio. Unfortunately, it was not possible to further optimize the ratio due to production constraints, as the sample GdSu19-04 is made in a different manufacturing approach than the samples GdSu19-20, GdSu20-01, and GdSu20-01A.

Judging by eye based on Figure B.9, GdSu20-01A is almost transparent, except for a few macroscopic remnants. Further, it has a pH value only a little bit smaller than the pH value of ultra-pure water which makes the salt the most promising candidate for the NV. However, measuring the absorption of the sample shows that the remaining macroscopic remnants lead to a significant drop in the transparency of the solution which is indicated by the semi-transparent red curve in Figure 4.3. To see whether a filtration of the solution can improve the transparency, Treibacher provided a filtered solution of the salt using a 0.2  $\mu\text{m}$  pore-sized syringe filter. The solution was produced using the in-house ultra-pure water of Treibacher of which a sample was shipped along with the filtered solution to Mainz. Figure 4.3 compares the absorbance of the filtered solution with the unfiltered sample as well as the previous samples shown in Figure 4.2. The filtered solution shows an improved absorbance which is comparable to the absorbance of the Mainz ultra-pure water sample. Further, it can be observed that the increase in absorbance for the GdSu19-04 sample below 250 nm, is also present in the in-house pure water sample provided by Treibacher. This trend is neither observed for the filtered GdSu20-01A sample nor the Japanese sample which was made from the Mainz ultra-pure water. Thus, the higher absorption observed in the GdSu19-04 sample might originate from contamination of the Treibacher pure water which was used during the manufacturing of the salt. Since the Treibacher water sample shows in general a higher absorbance than the filtered solution, it cannot be used as a reference standard to estimate the transmittance of the solution relative to water. This makes a direct comparison with the Japanese salt quite challenging. Nevertheless, to give an estimate for the transparency the filtered Gd-solution can be referenced to the Mainz ultra-pure water sample. The resulting transmittance is shown in Figure 4.4.

The transmittance of the filtered solution referenced to the Mainzer ultra-pure water shows a clear systematic bias towards larger transparencies as it even exceeds a transparency of 100%. Assuming, based on the other two samples (Japanese salt and Gd19-04), that the transparency loss for wavelengths beyond 500 nm is independent of the precise salt composition one can scale the transparency curve of the filtered solution accordingly. The resulting blue curve in Figure 4.4 shows a transparency that is a little bit better than the sample from Japan, however, this result must be taken with



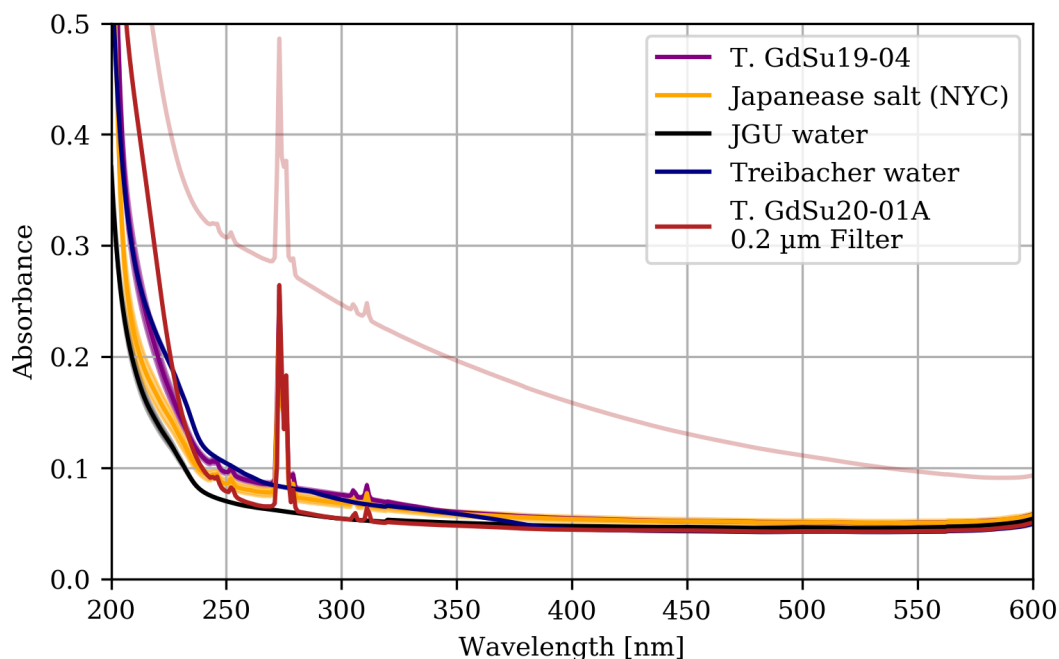


Figure 4.3. – Absorbance of the filtered Treibacher GdSu20-01A sample with respect to pure water and other salt samples. The meaning of the different lines are explained in the legend. The red shaded line shows the absorbance of the unfiltered GdSu20-01A sample. The solid lines show the mean and the shaded regions the standard deviation of the corresponding measurements as discussed in the previous section.

caution as it is not clear if the assumption is valid. The pH-value of the filtered solution is a bit more acidic compared to pure water, but with a value of  $(5.5 \pm 0.1)$  at a Gd concentration of  $(0.21 \pm 0.01)$  % it is less acidic than the Japanese salt. Unfortunately, it was not possible to conduct any additional tests before procuring the salt for the experiment due to the Covid lockdown which started in March 2020.

As mentioned before also the radio-purity of the salt is important as a high background rate would increase the overall lifetime-loss of the experiment. ICP-MS and  $\gamma$ -screening results for the GdSu19-04 and GdSu20-01A salts were conducted at dedicated facilities at LNGS. Both samples were meeting the requirements which were set at 2.5 ppb and 0.4 ppb for  $^{232}\text{Th}$  and  $^{238}\text{U}$  respectively. The results for the ICP-MS measurements are also summarized in Table 4.1. In addition, all samples were screened for  $^{56}\text{Fe}$  which is a typical absorber for UV light. No significant  $^{56}\text{Fe}$  contamination was found either. Thus, it was decided to buy 1 t of the GdSu20-01A salt from Treibacher, in addition to the 2.4 t of Japanese salt which was already procured before.

Unfortunately, additional quality checks of the 500 kg batches produced by Treibacher revealed that the thorium and uranium contamination of the production samples is much higher than the lab samples tested before. Additional screenings of the raw Gd-oxide material and another laboratory sample, GdSu20-17, provided by Treibacher show also higher contamination compared to previous samples. The results are also listed in Table 4.1. The current hypothesis is, that the raw Gd-oxide itself has higher

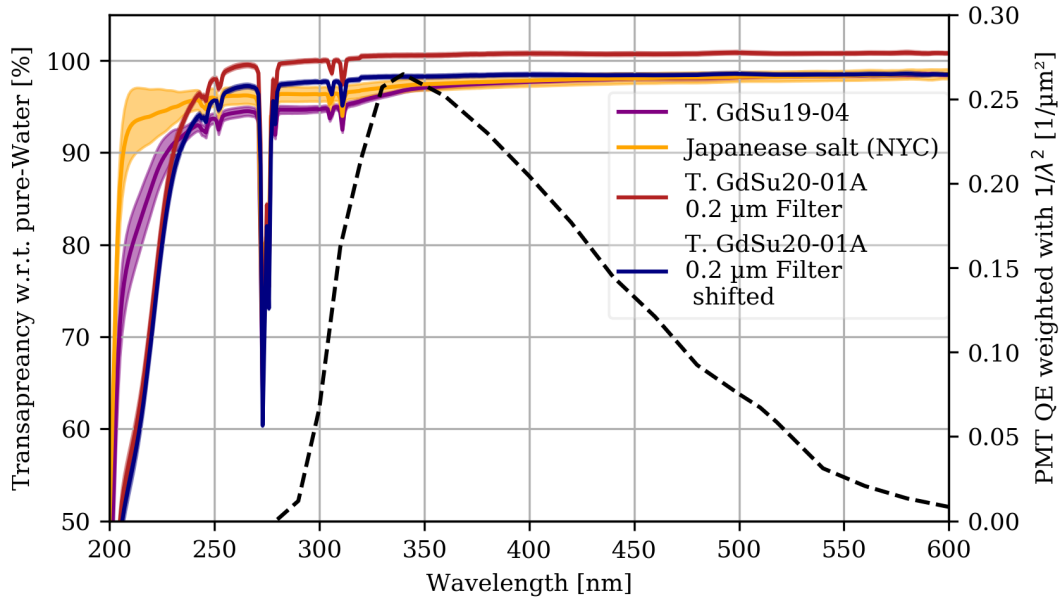


Figure 4.4. – Transparency of different Gd-sulfate samples with respect to ultra-pure water. The transparency was estimated according to Equation (4.3). The black dashed line indicates the region of interest for the NV which is given by the quantum efficiency of the NV PMTs folded with  $1/\lambda^2$  due to the spectrum of the Cherenkov light.

contamination than earlier batches used to produce the samples GdSu19-04 and GdSu-20-01A. Unfortunately, Treibacher does not have any remaining Gd-oxide which was used in the production of the GdSu19-04 and GdSu-20-01A samples. Thus, it remains a bit unclear whether the raw material is truly the origin of this higher contamination, although it seems to be likely.

To assess whether the experiment can withstand this higher background rate Geant4 simulation, conducted by Pietro Di Gangi, can be used to estimate the expected increase in the background compared to the specified requirement. The specified requirements of a maximal concentration of 2.5 ppb and 0.4 ppb for  $^{232}\text{Th}$  and  $^{238}\text{U}$  respectively, lead to an increase in background rate at a 10-fold PMT coincidence of  $\sim 1.5$  Hz, while using the salt with the higher contamination would result in a rate increase of  $\sim 6$  Hz assuming the background contribution from the Japanese salt is negligible. A detailed calculation about the background rate increase is given in the supplemental material on page 231. This renders a significant increase in the overall background rate for the NV considering that the measured background rate at a 10-fold coincidence is  $\sim 30$  Hz as it is discussed in the next Chapter.

#### 4.4 CONCLUSION AND OUTLOOK

In this chapter, the procurement of Gd-sulfate for the neutron veto was discussed. The salt was acquired in a joint effort between the university of Tokyo and the JGU-Mainz. The salt bought by Mainz was manufactured by the Austrian company *Treibacher In-*



*dustrie AG*. Different salt samples of the manufacturer were analyzed for their transparency, pH value, and radio-purity before the Covid lockdown in Germany started. A final salt sample with the reference number Gd20-01A was selected. It fulfilled all requirements on radio-purity and showed a good pH-value of  $(5.5 \pm 0.1)$  and high optical transparency once filtered with a  $0.2 \mu\text{m}$  syringe filter.

Unfortunately, the two 500 kg batches delivered by Treibacher are not fulfilling the same requirements as the measured test sample. A much higher contamination of  $^{238}\text{U}$  and  $^{232}\text{Th}$  was found. This increased contamination would lead to a  $\sim 20\%$  higher background rate for the neutron veto for a 10-fold PMT coincidence. In addition, the salt also shows a much lower transparency closer to the opaqueness that the sample Gd20-01 showed. Thus, while the neutron veto will be eventually doped with Gd, it is not yet decided if the Treibacher salt will be used or if a replacement will be ordered from the company. Investigations of whether at least the transparency can be recovered by adding additional sulfuric acid to the solution are currently explored by Christopher Hils at the JGU-Mainz. As an alternative, if only the Japanese salt is used the fraction of neutron captures on Gd only decreases to  $\sim 85\%$ .

---

## NEUTRON-VETO RESPONSE AND PERFORMANCE

In this Chapter, the performance of the neutron veto during SR0 is discussed, and some key characteristics of the background data are highlighted. The focus will be especially on the NV event-level data, which was mostly used in the analyses discussed in Chapter 6 and 7.

The first section of this chapter reviews the general performance of the NV during SR0 and the energy spectrum of the science data. Section 5.2 continues with a more detailed decomposition of the different background components of the NV background data based on the "center time" parameter. The subsequent section defines a quality cut to discriminate Cherenkov signals from other backgrounds. This cut is also used in the later Chapters 6 and 7.

### 5.1 NV DETECTOR CONDITIONS DURING SR0

The detector conditions of the neutron veto were monitored by frequent PMT calibrations, and by measurements of the reflectivity monitor, to monitor its optical properties. Both found the performance to be stable over the course of SR0. After fixing some initial noise and bit flip problems in the digitizer firmware<sup>1</sup>, the PMT baselines were found to be stable within  $\sigma_{\text{base}} = \sim 2.5$  adc counts for most of the PMTs. This is well below the average PMT threshold which is set to 15 adc counts. From the 120 NV PMTs, only two were excluded from the analysis due to fluctuating noise conditions, and 16 PMTs use a slightly higher threshold of 20 adc counts. The gain of the PMTs were found to be stable with variations of less than 5%, and a constant gain model was used for the entire SR0.

The dark rate of the PMTs varies between 500 Hz and 1100 Hz with an average dark rate of  $\langle R \rangle_{\text{DC}} = (618.8 \pm 5.3)$  Hz. For the event building discussed in Section 3.3, which uses a 3-fold coincidence requirement, this leads to an average event rate of  $\sim 152$  Hz, where  $\sim 8$  Hz are due to accidental coincidence from PMT dark counts. The latter was estimated using an analytical model which was derived during the planning phase of the NV processing chain. The model describes the chance of finding an

---

<sup>1</sup>An example waveform illustrating the bit flip problem is attached in the appendix in Figure B.13 on page 247

accidental coincidence rate  $\langle R_{AC} \rangle$  for a  $n$ -fold coincidence with  $m$  different counter, e.g. PMTs, and an average dark rate of  $\langle R_{DC} \rangle$  per counter as

$$\langle R_{AC} \rangle \approx (m + 1) \binom{m}{n-1} (\langle R_{DC} \rangle \cdot T)^{n-1} (1 - 2 \langle R_{DC} \rangle T)^{m-(n+1)} \cdot \langle R_{DC} \rangle. \quad (5.1)$$

The derivation of the formula can be found in Section A.2 of the supplemental material on page 228. Increasing the threshold to a 5-fold coincidence and using a 5 pe event area threshold, the event rate reduces further to  $\sim 66$  Hz with a  $\sim 40$  mHz contribution from accidental coincidences. This threshold condition is also used as a requirement for the NV neutron tagging discussed in Chapter 7. The 10-fold coincidence rate which was used as a benchmark threshold in the sensitivity study for XENONnT [Apr+20b] reduces the event rate to  $\sim 30$  Hz. This is better than expected as the overall deadtime due to the neutron veto at a 10-fold coincidence would be only 0.72 %<sup>2</sup> while the paper assumed a deadtime of 1.5 %<sup>3</sup>.

The rate at these higher thresholds is mostly dominated by decays of the various isotopes within the TPC and NV detector materials, for example decays of isotopes within the uranium and thorium chains. The typical NV event area distribution of background signals is shown in Figure 5.1. The spectrum is fairly featureless without any peak-like structures. Figure 5.1 also shows the spectrum after applying the so-called center time cut which is discussed in the following section.

## 5.2 NV EVENT CENTER TIME AND BACKGROUND DISTRIBUTION

One of the most important key parameters of the NV events is the so-called "center time". The center time is defined as the area-weighted average start time of the individual hitlets recorded within a single event. The distribution of the event center time as a function of the event area is shown in Figure 5.2 for 2.1 d of background, and 3.6 d of AmBe calibration data which is discussed in more detail in Chapter 7. The distribution of the latter was bin-wise background subtracted to highlight only true water Cherenkov events.

Five different distributions can be identified in Figure 5.2. The **first component** is given by events that distribute almost vertically below  $\sim 20$  pe up to very large center time values of about 500 ns. Further, these events can only be seen in the background data. Based on an analysis of their hit patterns it can be concluded that they originate from accidental coincidences between PMT dark counts.

The **second and fourth component** are present in both plots of Figure 5.2. Thus, it can be concluded that these events are caused by true Cherenkov signals. Events of the fourth component are dominantly distributed in a horizontal band around  $\sim 60$  ns up to  $\sim 60$  pe for background data and up to  $\sim 250$  pe in case of the calibration data. The characteristic center time of  $\sim 60$  ns originates from the optical properties of the NV.

<sup>2</sup>The NV neutron tagging and deadtime estimate are explained in more detail in Chapter 7.

<sup>3</sup>The estimates in the paper are based on Gd-loaded water, so with an reduced neutron capture time compared to the pure-water case of the presented studies. This highlights even more the excellent performance of the NV

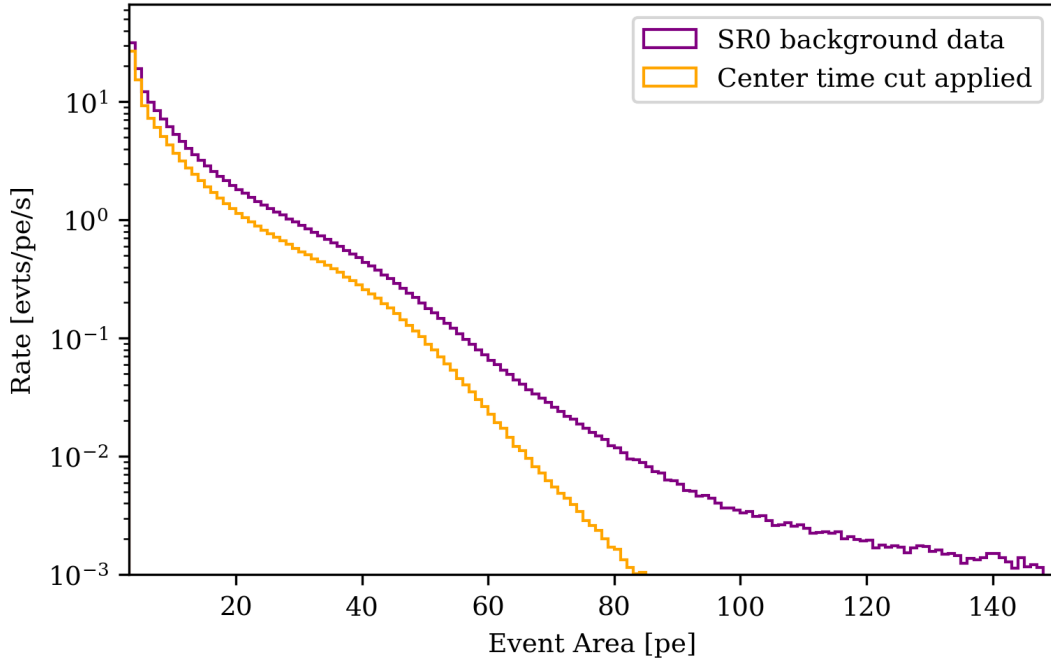


Figure 5.1. – NV SR0 background spectrum. The plot shows the differential rate of the NV events as a function of event area. The purple line shows all background signals which contain more than 3 pe area and which have at least 3 contributing PMT. The orange curve selects in addition only events within the central 95% contour of the center time distribution derived in Section 5.2. Errorbars are too small to be shown.

The recorded time distribution of the Cherenkov photons within an event can be described with an exponential distribution where the decay constant is characterized by the size of the detector, its water transparency, surface reflectivity, and photo-sensor coverage. The center time value, which represents an estimator for the expectation value of the exponential distribution, corresponds to its decay constant. Hence, the center time cannot only be used as a discrimination parameter to select Cherenkov signals but also serves as an important quantity to monitor the stability of the optical properties of the detector. In addition, the fourth component in the calibration data exhibits two peak-like structures at an event area of  $\sim 20$  pe and  $\sim 80$  pe, respectively. These peaks correspond to the 2.22 MeV and 4.44 MeV  $\gamma$ -ray signals induced by AmBe<sup>4</sup>. At smaller event areas the center time distribution widens and becomes slightly asymmetric. In addition to the horizontal band, two column-like structures can be observed at the positions of the full energy peaks which is labeled as second component in Figure 5.2. The columns abruptly stop at  $\sim 250$  ns and form an almost box-like structure. The same behavior can also be observed in background data, although partially overlapping with the first component.

To understand the peculiar shape of center time distribution of these Cherenkov

<sup>4</sup>For more information about these lines and the AmBe calibration the reader is referred to Chapter 6 and 7.

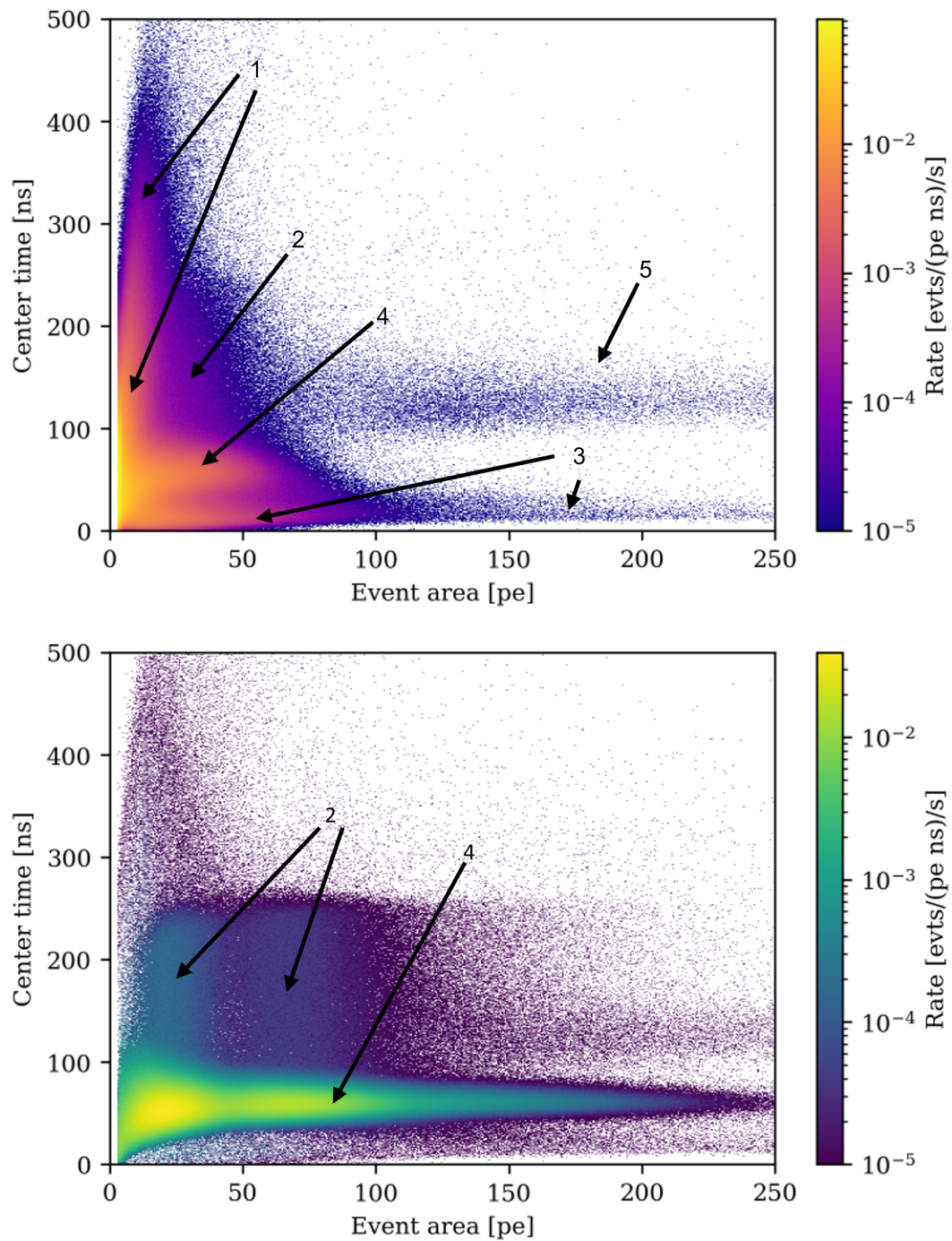


Figure 5.2. – Center time of the NV events as a function of event area for SR0 background and AmBe calibration data. The upper figure shows NV background data, while the lower figure shows the AmBe calibration data. The figure on the bottom was bin-wise background subtracted using the upper plot. The background distribution can be roughly divided into five different components of which two are also found in the lower figure. The different components are explained in the text.

events qualitatively, a simple toy MC was performed based on first principles. The simulation is given by an exponential distribution which uses as decay constant the center time of the 4.44 MeV peak. The center time was estimated by selecting all events between 43.5 pe and 82.0 pe, which corresponds to the central 1 sigma region of the peak. It was found to be  $\tau_{AmBe} = (59.84 \pm 0.06)$  ns. The charge of each pseudo photon was assumed to be 1 pe. The distribution of resulting toy events is shown in the appendix in Figure B.14 on page 248 overlaid with the AmBe calibration data. Data and toy MC show a good agreement within the horizontal band where most of the true Cherenkov signals are allocated. However, the events found within the box-like structure labeled as component 2 cannot be described in this way.

To confirm the hypothesis that these events are due to an accidental pile-up between Cherenkov signals and PMT dark counts, the toy MC sample was extended by a second toy signal component which follows a uniform distribution. To add these pseudo dark counts some approximations for the event building were made: Events in the NV are built via a 3-fold coincidence using a self-extending coincidence window of 200 ns. Hence, to approximate this behavior dark counts are simulated within a window of (-200 ns, 400 ns) around the exponential distribution which starts at 0 ns. No real event building is performed, and the center time is simply computed based on all toy hits within a toy event. The number of dark counts within an event is drawn from a Poisson distribution using a probability value  $p_{DC}$  which is given by

$$p_{DC} = n_{PMTs} \cdot \langle R_{DC} \rangle \cdot T \approx 0.0445 \quad (5.2)$$

with a total window length  $T = 600$  ns, number of NV PMTs  $n_{PMTs} = 120$  and an average PMT dark rate of  $\langle R_{DC} \rangle$ . To compare data and toy model qualitatively the median, as well as the 2.5 % and 97.5 % quantiles of the respective distribution, were estimated as a function of the event area. The quantiles were derived from the accumulative of their respective binned representations, and are shown in Figure 5.3. In addition, the resulting distribution from toy data is also compared against the unbinned AmBe calibration data in Figure B.14 in the appendix. The additional dark counts explain very well the box-like structure up to a center time of 250 ns.

The shift of the center time to higher values can be understood as a shift of the start time of the event due to leading dark counts preceding the actual Cherenkov event. Thus, the center time describes more robustly the true start time of a Cherenkov, than the actual event start time. This must be taken into account when building any coincidence between NV and TPC as done in Chapters 6 and 7. The median of the distribution from toy data and AmBe calibration data follows a similar shape although calibration data tends to have a bit larger center time at higher energies. The 2.5 % and 97.5 % quantiles show a larger deviation between data and simulation. The toy data prefers an overall more narrow distribution. This might be for two reasons. Firstly, the simple toy approach assumes a delta-like pulse shape for the PMT signals, while the real signal shape may lead to an extra broadening of the distribution. Secondly, the recorded charge of each pseudo photon was assumed to be 1 pe, while a normal distributed charge distribution would lead to an extra smearing in the event area. The deviation at the 97.5 % quantile is too large to be explained only by the shape



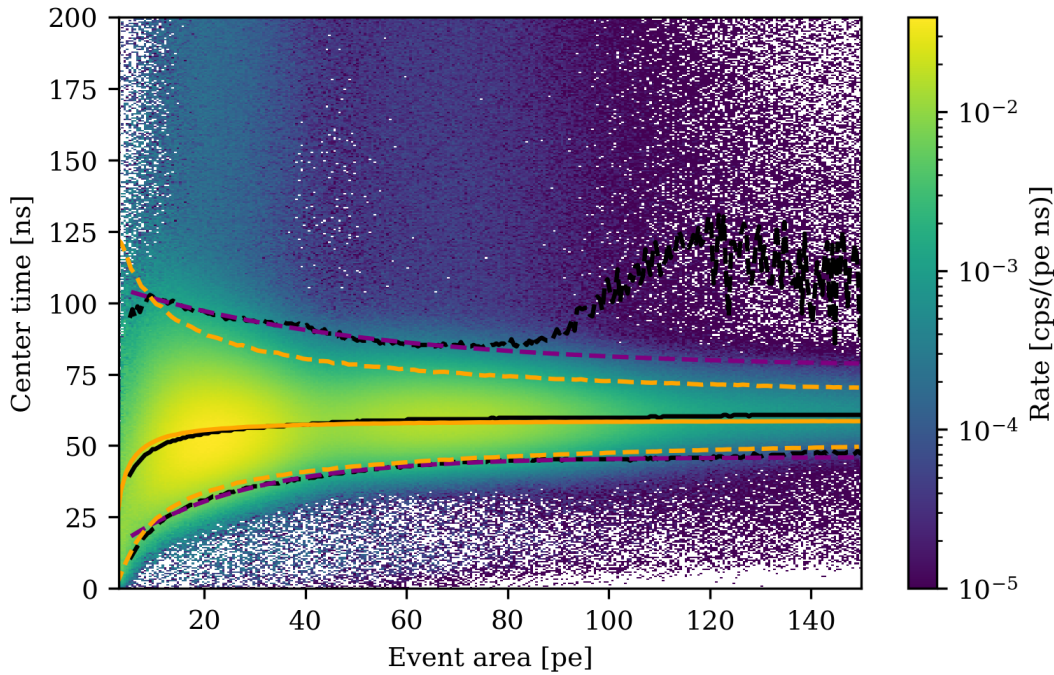


Figure 5.3. – Background subtracted AmBe event center time as a function of event area. The histogram shows the same data as in Figure 5.2. The black dashed lines indicate the lower and upper 2.5% and 97.5% quantile of the distribution. The black solid line indicates the median of the distribution. The orange lines indicate the same, but for the toy data as explained in the text. The purple dashed lines indicate the respective quantile boundaries of the center time cut.

difference. The data appears to prefer a larger fraction of events with dark count pollution when predicted by the toy data. This might come from the fact that the true underlying distribution for PMT dark rates plays a role here, and that using only the averaged PMT dark rate might be an oversimplification. In addition, the toy MC does not include any other background signals which also could lead to event pile-up and a shift in the center time distribution to larger values. However, as the toy MC is only used for a qualitative description of the data, the deviation is less important.

Based on the 2.5% and 97.5% percentiles determined from the background-subtracted calibration data, cut boundaries for the central 95% of the distribution can be set. The upper and lower boundary of the cut is defined empirically by an exponential function of the form

$$f_{\text{bound}}(x) = A \cdot \exp\{-\lambda \cdot x\} + B \quad (5.3)$$

which was fitted to the respective boundaries. The optimization was done using a chi-square method. Since the quantiles were estimated based on the accumulative distribution of binned representation of the data in Figure 5.3, the uncertainties of the percentiles were set as the width of a single bin. Both fits were performed between 5 pe and 70 pe to avoid regions with low statistics. The resulting cut limits are also shown in Figure 5.3 and estimates of the best fits in Table 5.1. Both fits show a good matching with their respective quantiles except for small event areas below 8 pe. In

this region, the cut boundaries are systematically biased towards smaller acceptances. Only for the upper boundary the  $\chi^2$  and the number of degrees of freedom indicate to reject the model. This might be because, in contrast to the lower boundary, the upper boundary is less well defined due to of the pollution of events with PMT dark counts.

At high energies beyond 70 pe the boundaries show at least by eye a good matching with data. However, in this region, the cut should only be applied with caution as the boundaries are just extrapolated. A comparison between background data and the center time cut boundaries is shown in the appendix in Figure B.15 on page 249. After applying the center time cut to the background data, the rate for a 5-fold coincidence and a 5 pe threshold reduces by  $\sim 20\%$ .

Coming back to the discussion about the remaining background components of Figure 5.2, the **third component** shows a similar horizontal distribution as regular Cherenkov signals, but shifted towards smaller center time values of about 20 ns. These events also span over a larger event area range than regular background Cherenkov signals, more similarly to the AmBe calibration data. All these events have in common that a single PMT dominates the event as it can be seen from Figure B.17 in the appendix on page 250. Thus, the shape of the events is dominated by the signal shape of a single PMT and not by the optical properties of the NV. At low energies, a likely origin for these types of events is PMT afterpulses which are paired accidentally with other dark counts. At larger energies these events are more likely dominated by Cherenkov signals due to the  $\beta$ -decays within the quartz windows of the PMTs, or at least very close to it. Possible candidates for such signals are  $\beta$ -decays within the chain of  $^{232}\text{Th}$ ,  $^{238}\text{U}$ , as well as the  $\beta$ -decay of  $^{40}\text{K}$  which were found during the material screening of the NV with a specific activity of 0.42 Bq/PMT, 0.60 Bq/PMT and 0.60 Bq/PMT<sup>5</sup> respectively.

The last remaining component is the **fifth component** which is found at higher energies and higher center time than the AmBe calibration data. This background distribution eludes a physical explanation and is still not understood. The typical hit pattern of these events do not exhibit any strange behavior which points to an artificial origin. An example event is shown in the appendix in Figure B.18 on page 251. The accidental pile-up of two events can be mostly ruled out for two reasons. First of all, the typical event window of about 200 ns and the background rate at high energies of about 30 Hz are too small to explain the observed rate of these events which is  $\sim 10$  mHz. In addition, if this population would be indeed caused by pile-up of events the calibration data should show evidence for a similar pile-up as the overall rate in the detector is

<sup>5</sup>These specific activities were determined from the PMT windows only. The PMT body was screened separately.

Table 5.1. – Best fit values for the upper and lower boundary of the center time cut. The first three columns contain the best fit estimates for the respective boundaries. The last column indicates the  $\chi^2$  and number of degrees of freedom (ndof).

	A [ns]	$\lambda$ [1/pe]	B [ns]	$\chi^2$ /ndof
upper boundary	$31.8 \pm 0.8$	$0.0136 \pm 0.0012$	$72.3 \pm 1.0$	351/127
lower boundary	$-39.3 \pm 0.4$	$0.0504 \pm 0.0011$	$44.65 \pm 0.19$	78/127



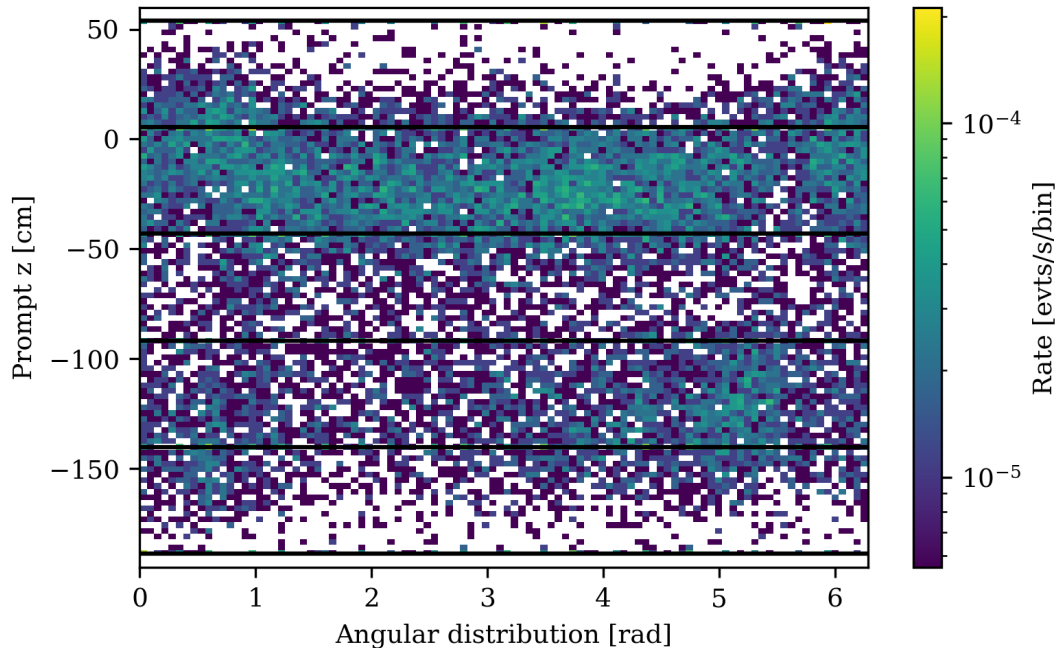


Figure 5.4. – Spatial distribution of the fifth background component of Figure 5.2. The plot shows the reconstructed  $z$  position as a function of the reconstructed angle. The black vertical lines indicate the different height levels of the NV PMTs.

much larger. However, this is not the case. Thus, since there is neither evidence for an artificial origin nor pile-up it must be concluded that these events might have a true physical origin.

To understand the origin of these events a bit better, their spatial distribution can be investigated. The distribution<sup>6</sup> is shown in Figure 5.4, and is very localized compared to regular background events which are uniformly distributed<sup>7</sup>. Events of this distribution seem to be emitted dominantly at two different heights of the detector over its full azimuth angle, with a higher rate being observed between  $-50 \text{ cm} < z < 50 \text{ cm}$  than between  $-150 \text{ cm} < z < -100 \text{ cm}$ . One natural arising suspect for the higher rate might be the stiffening rings which are welded to the inner and outer TPC cryostat as rendered in Figure 2.14 on page 41. These rings are located at two different  $z$  and circuit the full TPC cryostat. Also, the background data of the TPC exhibits a higher background rate close to the positions of the stiffening rings than predicted by simulations<sup>8</sup>. In addition to the two continuous distributions, a small "blob" at the position (5 rad,  $-120 \text{ cm}$ ) can be observed in Figure 5.4. This position corresponds to the bottom of the beam pipe for the neutron back-scatter set-up where it is mounted to the TPC cryostat. The back-scatter set-up is also rendered in Figure 2.15 on page 46.

However, the spatial origin does not explain the peculiar properties of the NV

<sup>6</sup>Including all events which have at least a 5-fold coincidence, contain 80 pe total area, and a center time between 80 ns and 180 ns.

<sup>7</sup>See Figure B.16 on page 250 in the appendix

<sup>8</sup>See Figure B.19 on page 251 in the appendix.

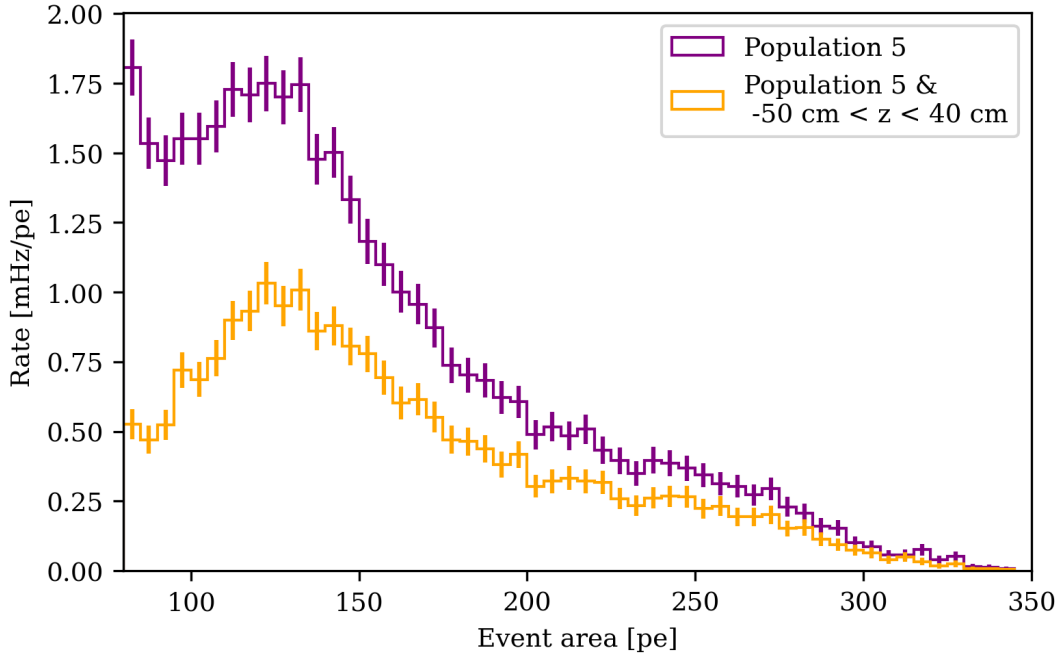


Figure 5.5. – Area distribution of the NV events found in the fifth population of Figure 5.2. The purple figure shows all events with a center time between 80 ns and 180 ns. The orange distribution requires in addition the events to be between  $-50$  cm and 40 cm in  $z$ . The statistical uncertainty for each bin is shown.

events. The shift towards larger center times can be explained if the origin of the event is indeed due to pile-up, but of correlated events, like a chained decay with a short lifetime of the intermittent isotope or state. Based on the event area of these events the total energy released in the two decays would need to be in the order of 7 MeV to 9 MeV based on the energy calibration shown in Figure 7.13 on page 165. Another hint for this hypothesis is the area distribution of the events which is shown in Figure 5.5. At least one peak-like structure can be observed with an average area of about 120 pe, which corresponds to an energy of  $\sim 7$  MeV according to the simulated energy response of the NV discussed in Section 7.2.3.

If these events are indeed due to a timely correlated decay, it would be expected to find its constituents when applying a coincidence between TPC and NV. At least as long as one of the two decays is a  $\gamma$ -decay, which can penetrate the stainless steel cryostat. However, studies of the ER background using a coincidence between TPC and NV, which were conducted for the low ER analysis of XENONnT [Apr+22c], did not reveal any signals which would fit into this overall picture. Thus, either the signals have a different origin or the constituents of the decay might not be able to penetrate the TPC cryostat like  $\beta$ -decays or conversion electrons. However, this would also mean that the higher activity observed in the TPC and the strange background component in the NV are either completely uncorrelated or due to different decays in the same decay chain. At least it seems hard to believe that their similarities in spatial reconstruction are a pure coincidence.

### 5.3 CONCLUSION AND OUTLOOK

In this Chapter, the performance of the neutron veto during SR0 is discussed. This included a summary of its stability and the typical energy distribution of the observed events during science data taking. The total background rate of the NV for a 5-fold coincidence and a 5 pe threshold was evaluated to be  $\sim 66$  Hz, which are the triggering conditions for the NV tagging of TPC signals discussed in Chapter 7. This is much better than expected at the design stage of the detector.

In addition, this chapter discusses the properties and behavior of the "center time" parameter of the NV events which allows to deconstruct the observed background signals into Cherenkov signal and background signal components. All these components were studied and explained as part of the presented work. Only the origin of the "fifth component" is not yet fully understood and needs further investigation. Its contribution to the overall rate of the NV is negligible which makes a precise understanding for the SR0 WIMP search less important.

Further, a data quality selection cut based on the center time was developed. The quality cut is used in the following studies presented in Chapter 6 and Chapter 7. While not used as part of the NV tagging triggering conditions, adding the center time quality cut as a requirement would reduce the overall NV background rate by  $\sim 20\%$  to  $\sim 52$  Hz. However, also the tagging efficiency of the NV, determined in Chapter 7, would be reduced.

As mentioned in the conclusion of Chapter 3, the most important information for the NV analysis is contained in the event level data-type, and especially the shape of the Cherenkov signal can be used to discriminate background from true Cherenkov signals. Thus, to further improve the understanding of the detector this information level should be extended for SR1. The discussion about the center time also showed that an additional signal splitting at the event level might be beneficial for signal reconstruction. The splitting could reduce the influence of leading dark count signals on the estimates for the center time and position reconstruction of an event.

## NUCLEAR RECOIL CALIBRATION

In this chapter, the nuclear recoil calibration of XENONnT is discussed. It is one of the milestones required to understand the detector response for neutral elastic interactions of unknown particles with xenon nuclei. As such it is important for the search of WIMP dark matter or  $CE\nu NS$ . Furthermore, it represents the foundation for the determination of the neutron veto tagging efficiency. The calibration is based on a coincidence technique between NV and TPC, which, to our knowledge, was used here for the first time in a LXe TPC dark matter experiment. The coincidence is built by measuring simultaneously neutrons and gammas which are coincidentally emitted from an Americium-Beryllium source. The very same technique was also applied for the calibration of the NV which is explained in Chapter 7.

This chapter starts with a general discussion about the key characteristics of an AmBe source as well as the working principle of the calibration technique itself. In addition, Section 6.1 gives detail about the NR calibration conducted during the first science run of XENONnT. Section 6.2 discusses the different aspects of the coincidence and data-quality selections, which are used to get a clean NR calibration data sample. Further, the section deals with the acceptance and correlation of the individual data quality cuts. The latter is another key ingredient for the fitting of the NR response model reported in the closing Section 6.3 of this chapter before a conclusion and outlook are given.

### 6.1 WORKING PRINCIPLE OF THE AMBE COINCIDENCE AND SR0 DATA TAKING

#### 6.1.1 Americium-Beryllium source

The compound AmBe emits neutrons mainly via the alpha-capture reaction  ${}^9\text{Be}(\alpha, n){}^{12}\text{C}$ , and, rarely, by the neutron break-up reaction:  ${}^9\text{Be}(\alpha, \alpha'){}^9\text{Be}^* \rightarrow {}^8\text{Be} + n \rightarrow \alpha + \alpha + n$ . The alpha is emitted by  ${}^{241}\text{Am}$  which decays with a half-life of about 432.2 years into an excited state of  ${}^{237}\text{Np}^*$ . The most dominant energy state of  ${}^{237}\text{Np}^*$  is at 60 keV.  ${}^{237}\text{Np}$  itself has a half-life of about 2 million years [GZ75; Sch+15; Mur+14]. Therefore, no considerable background from the  ${}^{241}\text{Am}$  decay itself is to be expected for the

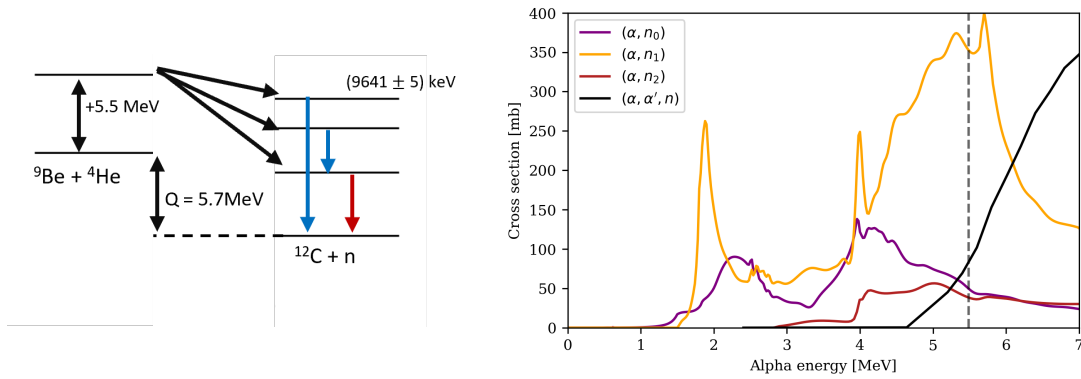


Figure 6.1. – Left: The schematic illustrates the energy levels involved in the  ${}^9\text{Be}(\alpha, n){}^{12}\text{C}$  reaction. The maximal kinetic energy of the  ${}^{241}\text{Am}$   $\alpha$  is indicated via an additional energy level 5.5 MeV above the ground state of  ${}^9\text{Be}$ . Arrows in blue indicate transitions via internal conversion while the red arrow indicates a transition via  $\gamma$  emission. The energy levels, their uncertainties and the transition properties of  ${}^{12}\text{C}$  are taken from the NuDat3.0 database [Bro22]. All energy levels are plotted relatively to  ${}^9\text{Be} + {}^4\text{He}$  and are to scale.

Right: The figure shows the various neutron production cross sections for  ${}^9\text{Be}$  as a function of alpha energy. The index  $n_i$  represents the corresponding final states of  ${}^{12}\text{C}$ .  $n_0$  represents the ground state,  $n_1$  and  $n_2$  the first and second excited states at 4.44 MeV, 7.65 MeV respectively. The reaction  $(\alpha, \alpha', n)$  refers to the neutron break-up reaction explained in the text. The vertical dashed line indicates the average alpha energy for  ${}^{241}\text{Am}$ . The data was taken from JENDL/AN-2005 database [Mur+06].

calibration.

The capture of an  $\alpha$ -particle on  ${}^9\text{Be}$  leads to a compound nucleus of  ${}^{13}\text{C}^*$  which quickly decays via the emission of a neutron into either the ground or an excited state of  ${}^{12}\text{C}$ . In the case of  ${}^{241}\text{Am}$  the two most dominant excited states are the first or second excited state ( $n_1, n_2$ ) with a total energy of 4.44 MeV and 7.65 MeV respectively. Although accessible via energy, the contribution from the next higher excited state of  ${}^{12}\text{C}$  at 9.64 MeV is negligible for  ${}^{241}\text{Am}$   $\alpha$ -decays [GZ75]. A schematic of the involved energy levels is shown in Figure 6.1 on the left-hand side.

In general, the branching ratio of the different excited states depends on the remaining kinetic energy of the alpha particle when impinging on the  ${}^9\text{Be}$  nucleus. Therefore it depends on the exact manufacturing of the source, including the grain size of the  ${}^{241}\text{AmO}_2$  powder as well as the capsular material of the source encasement. In literature, the branching ratio for the first excited state of  ${}^{12}\text{C}$  is usually given to be around 60% [GZ75; Sch+15; Mur+14]. This value is slightly lower than what one would expect based on the individual neutron-production cross sections shown in Figure 6.1 right-hand side. At the typical alpha energy for  ${}^{241}\text{Am}$ , which is about 5.5 MeV, the expected ratio is around 67%.

The energy spectrum of the produced neutrons is a continuous spectrum with a maximal neutron energy of about 11 MeV and an average energy of about 4.5 MeV. It should be noticed that the maximal and mean neutron energy depends on the precise manufacturing of the source, due to the  $\alpha$ -energy dependence of the neutron production cross-section and the final state of  ${}^{12}\text{C}$ . The neutron-spectrum features various peaks given by the production cross-section shown in Figure 6.1. Figure 6.2 shows the

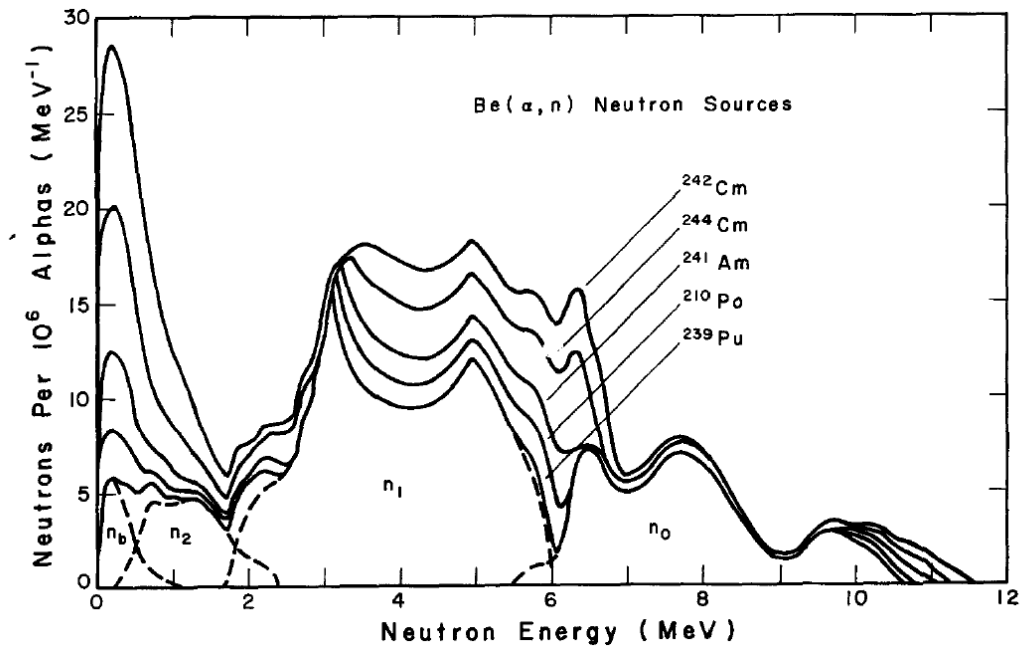


Figure 6.2. – Neutron yield as function of neutron-energy for different  $\alpha$  sources and  ${}^9\text{Be}(\alpha, n){}^{12}\text{C}$  reaction channels.  $n_0$ ,  $n_1$  and  $n_2$  refer to the different excited states of  ${}^{12}\text{C}$  as explained in the text.  $n_b$  refers to the neutron-emission via break-up reactions. The figure was taken from [GZ75].

neutron-energy spectrum and yields for different  $\alpha$  sources and  ${}^9\text{Be}(\alpha, n){}^{12}\text{C}$  reaction channels.

### 6.1.2 SR0 AmBe calibration

The concept of the NR calibration is sketched in Figure 6.3. It is based on a tight coincidence between S1 signals induced by neutrons and 4.44 MeV  $\gamma$ -ray signals from the first excited state of  ${}^{12}\text{C}$  recorded by the NV.

The NR calibration took place at the beginning of SR0 between May 24<sup>th</sup> 2021 and June 7<sup>th</sup> 2021. During this period an AmBe source was placed next to the TPC cryostat using the u-tube calibration system shown in Figure 2.15 on page 46. The AmBe source used for this calibration campaign was purchased and characterized in XENON100. It has a total  ${}^{241}\text{Am}$  activity of about 3.7 MBq with a resulting neutron flux of  $(160 \pm 4)$  n/s<sup>1</sup> as measured by the Physikalisch-Technische Bundesanstalt (PTB), the German National Metrology Institute, in August 2012 [Apr+13]. The source has a diameter of 6.4 mm and a length of 17.6 mm. It is hooked via an M4 thread to the u-tube calibration system. A technical drawing of the source is attached in the appendix in Figure B.11 on page 245.

The source was placed at four different locations next to the TPC cryostat. Table 6.1 shows the different source positions, the distance towards the TPC cryostat as well as

<sup>1</sup>Considering the half life of  ${}^{241}\text{Am}$  which is 433 year [Bro22], a neutron rate of  $(159 \pm 4)$  n/s remains.

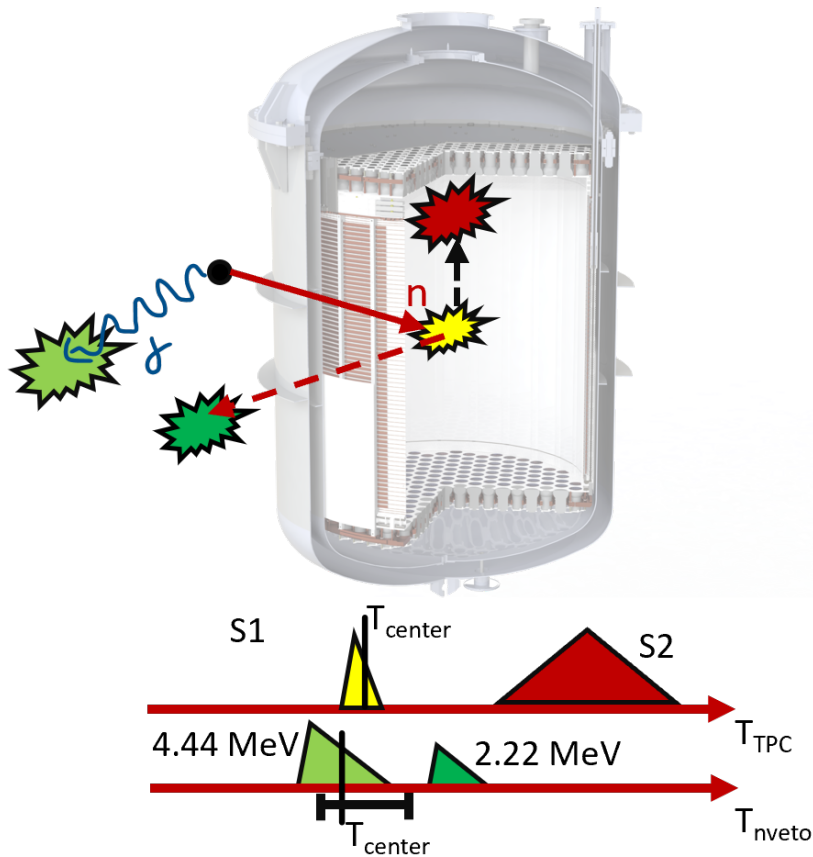


Figure 6.3. – Schematic of the TPC NR calibration and NV tagging efficiency calibration explained in Chapter 7. The black dot illustrates the AmBe source.  $\gamma$ -ray and neutron are coincidentally recorded by the two detectors. The graphic on the bottom illustrates the timing for the various signals involved. The time interval range illustrates the coincidence window which is defined around the center time of the NV event. The two detectors are in coincidence if the center time of an S1 signal occurs within that window.

the time spent at each position.

During calibration, the data acquisition was operated in two different run modes named `ambe_link` and `ambe_linked_hev`. The former is similar to the regular run mode used during science data taking in which TPC and NV are operated synchronously. `ambe_linked_hev` has in addition a hardware high energy veto which vetos the TPC data acquisition for high energy gamma signals in the TPC, e.g. the 4.44 MeV gamma. Vetoing high-energy signals and their tail leads to a reduction of the overall amount of data produced during the calibration. It should be noticed that the high energy veto only affects the data acquisition of the TPC but not the NV.

The data was analyzed using the dedicated STRAXEN base plugins explained in Section 3.3.5. The two detectors were synchronized by delaying the NV signals by 10 157 ns into the time frame of the TPC, prior to building any coincidence. The center time of the NV event and the S1-signal were used as time references for these coincidences. The coincidence window itself was adjusted depending on the analysis task



**Table 6.1.** – Table of the various source positions used during the AmBe calibration of SR0. The position is encoded in the same nomenclature used for the tagging of the runs. **CW** and **CCW** refers to the insertion direction of the source which can be either clockwise or counter clockwise along the u-tubes. The following number is the insertion depth in meter. **d** refers to the decimal point such that CW5d9m reads: clockwise 5.9 m. The insertion depth is also translated into the Cartesian coordinate frame of the TPC given by the columns X,Y,Z.  $\Delta R$  represents the distance between source and outer cryostat surface. The time column contains the time spent at a specific calibration position. The number in brackets refers to the time spent at that position using the high energy veto run mode explained in the text. The total time spent at each position is given by the sum of the two values.

Position:	X [mm]	Y [mm]	Z [mm]	$\Delta R$ [mm]	time [h]
Top CW5d9m:	10.9	-964.0	-102.1	129.8	15.04, (8.69)
Top CCW5d0m:	1084.6	-976.5	404.0	625.2	18.05 (-)
Bottom CW6d6m:	151.5	-1068.9	-955.9	245.4	22.74 (8.32)
Bottom CW11d2m:	987.7	-371.5	-910.1	221.1	12.37 (-)

at hand and is discussed further in Section 6.2.

For the calibration of the NR response model only the positions Top CW5d9m, Bottom CW6d6m, and Bottom CW11d2m are used. This corresponds to a calibration time of 67.16 h in total. The fourth position Top CCW5d0m, is of too large radial distance from the TPC, such that no significant neutron rate above background was observed in the TPC. Thus this source position is only used for the NV only analysis discussed in Chapter 7. A comparison of the spatial distribution of four calibration positions is shown in Figure 6.4. A more detailed discussion about the event selection applied in these plots follows in Section 6.2.2.

At each calibration position, the radial distance between TPC and cryostat surface was chosen to be at least 10 cm, after placing the source first close to the TPC without any water gap. This decision was made for two reasons. First, the data rate very close to the detector was so high that there were concerns regarding storage capacity and DAQ deadtime. Second, due to the short circuit between the cathode and bottom screening mesh and the appearance of a hot-spot discussed in Section 2.4.1, it was unclear if a very high production of electrons might worsen the problem.

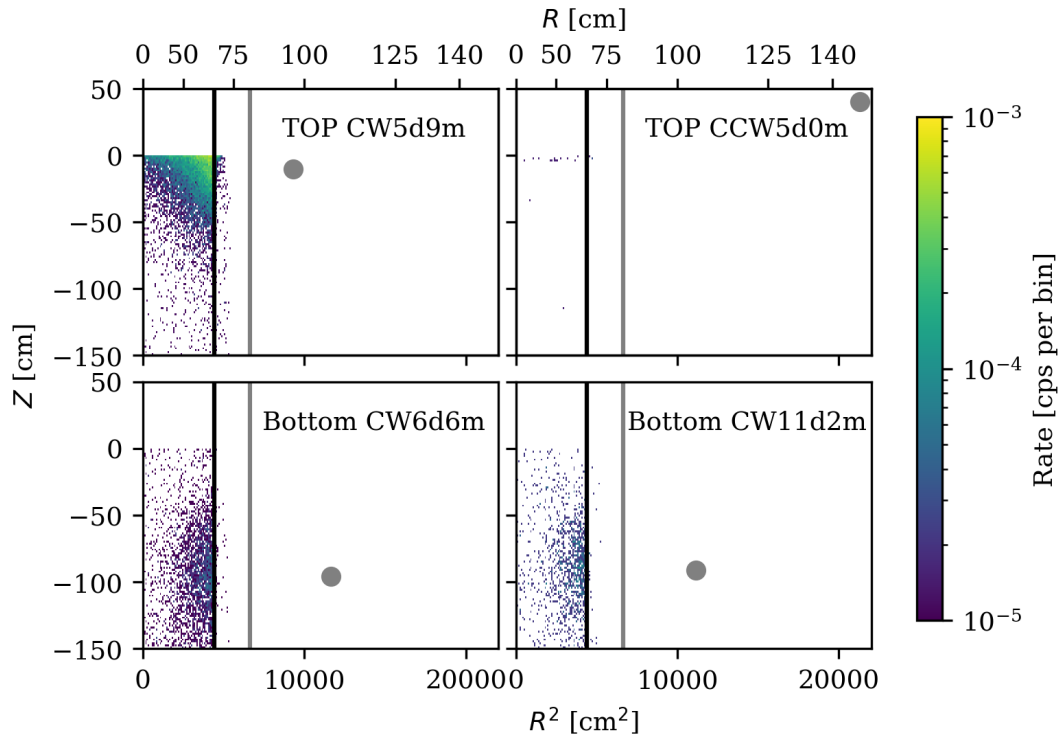


Figure 6.4. – Spatial distribution of the NR calibration events in the TPC for the individual calibration source positions. The plot shows the  $z$  versus  $R^2$  distribution for all events found in coincidence with NV and TPC within the NR ROI which includes all events with a  $cS1$  below 300 pe. The gray dot indicates the source position. The black vertical line indicates the position of the TPC PTFE wall while the gray line indicates the radius of the outer TPC cryostat shell, i.e. the boundary to the NV. The upper axis shows the distance from the center of the TPC. Events that are reconstructed outside of the TPC are an artifact of the position reconstruction explained in Section 3.2.

## 6.2 DATA SELECTION, OPTIMIZATION AND CUT EFFICIENCY

### 6.2.1 NV gamma selection and detector coincidence

To build a tight coincidence between NV and TPC a few analysis steps are required. First, all NV events originating from a 4.44 MeV  $\gamma$ -ray must be selected. Subsequently, a suitable coincidence window must be found and characterized. The characterization should include an estimate of the number of accidental coincidences between the two detectors as well as an analysis of the timing resolution of the detector coincidence. Last but not least, the impact of the coincidence on the NR TPC data must be studied.

The selection of the 4.44 MeV  $\gamma$ -ray in the NV is done in the form of data quality cuts. This is required as the coincidence is built via the corresponding straxen plugin discussed in Section 3.3.5. For the NR-response calibration, triggering NV events are defined via the following selection criteria:

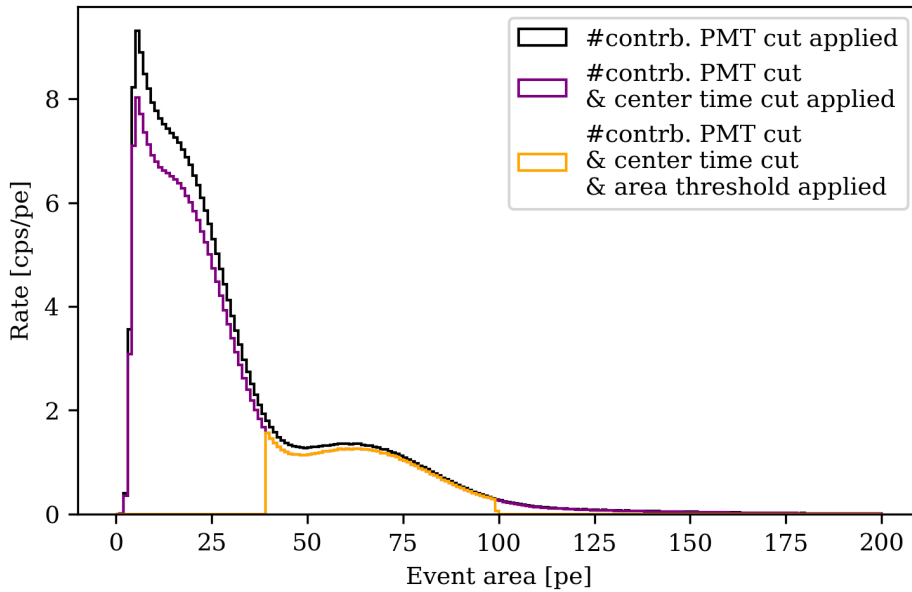


Figure 6.5. – Selection of trigger NV events for the TPC NR calibration. The data shown includes all three relevant calibration positions discussed in the text. The plot shows the event rate per bin as a function of NV event area. The different colors represent different cut selections indicated in the legend. The error bars are too small to be visible.

- The number of contributing NV PMTs must be at least four. This cut reduces the number of false triggers due to PMT afterpulses and due to Cerenkov events caused by  $\beta$ -particles emitted in the glass windows of the NV PMTs. Furthermore, it reduces the number of false triggers due to neutron events formed by accidental coincidences of PMT dark counts.
- Each event must have a center time within the 95 % contour as explained in Section 5.2. This selection further reduces the background due to afterpulse events as well as pile-up events and regular PMT noise.
- An event must have an area between 39.0 pe and 99.23 pe corresponding to the 8.5 % and 97.5 % percentile of the 4.44 MeV full energy peak of  $^{12}\text{C}$ . The lower percentile was chosen deliberately higher to reduce contamination from false triggers due to random coincidence between the 2.22 MeV  $\gamma$ -ray of the hydrogen-neutron capture and the TPC events. The percentiles were evaluated based on the best fit of the NV AmBe calibration spectrum discussed in Chapter 7.

The effect of the different cuts on the event area distribution is shown in Figure 6.5. The spectrum features a fast-rising edge due to the NV threshold requirement followed by a small exponential declining peak below 10 pe. This peak is dominantly caused by accidental coincidences between PMT dark counts forming false events and background signals produced by radioactive decays of the detector materials. At

higher event areas, two Gaussian-like peaks follow with a maximum of around 20 pe and 60 pe respectively. The first peak represents the 2.22 MeV hydrogen-capture line, while the second peak corresponds to the 4.44 MeV  $\gamma$ -ray of the first excited state of  $^{12}\text{C}$ . A more detailed discussion of the event-area distribution follows in Chapter 7.

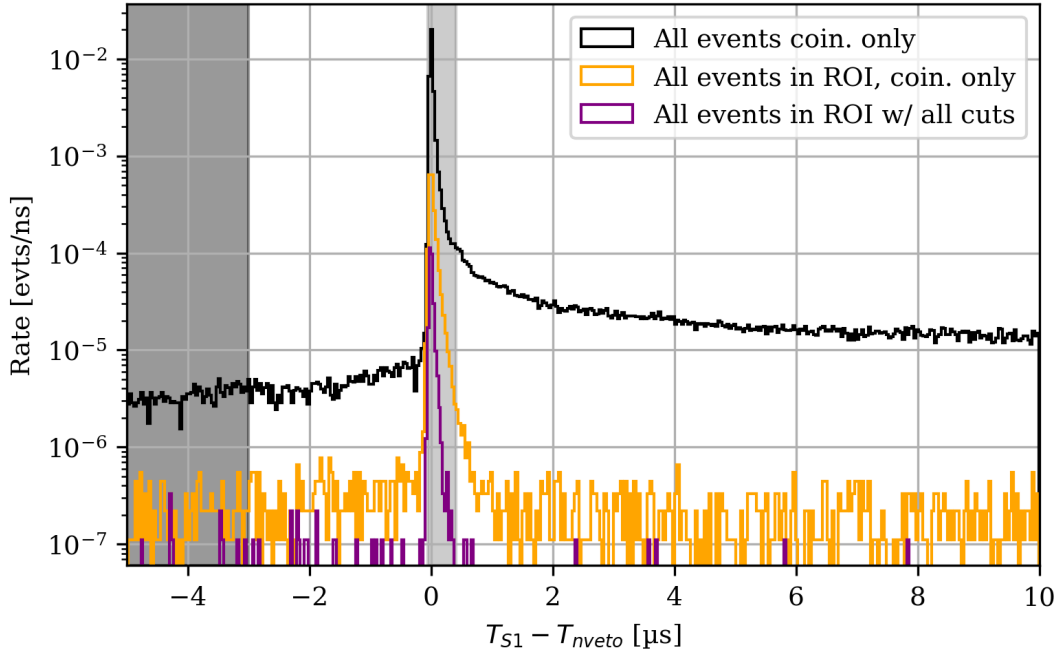
Table 6.2 shows the number of remaining NV events and rate after applying the mentioned cuts. After cuts a rate of  $(62.55 \pm 0.02)$  evts/s remains which is 35 % larger than based on a back-of-the-envelope estimate using the neutron rate of  $(159 \pm 4)$  n/s, a branching ratio of about 50 % between neutrons and gammas, and a solid angle of about 50 % resulting in a rate of  $(40 \pm 2)$   $\gamma$ /s. However, this comparison is truly a back-of-the-envelope estimate as the selection contains remaining pollution of 2.22 MeV  $\gamma$ -ray events, and efficiency losses due to the event selection are not taken into account. However, a precise knowledge of the  $\gamma$ , and therefore neutron rate, is not necessarily required for the calibration of the NR response model. The main focus of this calibration is on the shape of the NR band, i.e. the size of the S1 and S2 for a given energy deposit and the recombination model, converting charge quanta into photons. An absolute measurement of the  $\gamma$  rate, and therefore the neutron rate, would be desirable nonetheless, as this would allow to determine the absolute efficiency of the calibration. To achieve this precise knowledge about the branching ratio of the source is required to match simulation with data. A corresponding study is discussed in Section 7.2.2.

Using 11 h of background data taken before the AmBe calibration campaign and the same selection criteria, the rate of false trigger signals due to background events is estimated to be  $(2.698 \pm 0.008)$  evts/s. This value is a factor 30 smaller than the rate of the selected 4.44 MeV  $\gamma$ -rays.

After selecting triggering NV events, the coincidence window between NV and TPC needs to be optimized. The coincidence window is defined relative to the center time of a triggering NV event. A TPC event is in coincidence with the NV if the center time of its S1 signal is found within this window. For both detectors, the center time has been chosen as it is the most robust definition of a timestamp in sparse signals. To optimize the coincidence window, a larger window was chosen at first. All TPC events with a main S1 within a window of  $-5 \mu\text{s}$  to  $10 \mu\text{s}$  around a triggering NV event are selected. Figure 6.6 shows the distribution of time difference between the center time of S1 signals relative to the center time of the NV events found in coincidence. The black distribution in Figure 6.6 shows the distribution for all events. The distribution

Table 6.2. – Selection statistics for the NV  $^{12}\text{C}$   $\gamma$ -ray selection. The first column indicates the cuts explained in the text. The second column represents the number of remaining events after applying the corresponding cut and all cuts of the previous rows. The third column shows the remaining event rate. The uncertainty was estimated using Poisson statistics for the number of remaining events and considering the uncertainty on the total runtime as negligible.

Cut	Number of events left	Rate [evts/s]
no cuts	150853402	$623.94 \pm 0.05$
cut on $^{12}\text{C}$ area cut	17057454	$70.55 \pm 0.02$
cut on number contributing PMT cut	16411174	$67.88 \pm 0.02$
cut on event center time cut	15122449	$62.55 \pm 0.02$



**Figure 6.6.** – Time difference distribution for coincident NV and TPC events. The plot shows the rate of events found in coincidence per bin as a function of time differences estimated between the S1 center time and NV event center time. The distribution for different sets of TPC data-quality cut are presented. The cut details are explained in the text. The light gray shaded region indicates the coincidence window which is used in the NR response calibration. The dark gray region ranges from  $-5 \mu\text{s}$  to  $-3 \mu\text{s}$  and serves as reference region to estimate the number of accidental coincidence between NV and TPC for a given selection. The uncertainties of the individual bins are not shown.

features a sharp peak centered around zero, which is followed by an exponentially falling edge. The peak sits on top of an almost uniform distribution which shows a factor  $\sim 2$  rise in rate between  $-2 \mu\text{s}$  and  $0 \mu\text{s}$ . In addition, the uniform distribution has a smaller rate before ( $\sim 3 \cdot 10^{-6}$  evts/s) than after the peak ( $\sim 1.3 \cdot 10^{-5}$  evts/s). The orange distribution shows the same data, with the additional requirement that all TPC events are found within the ROI for the NR calibration, i.e., having a cS1 area of less than 300 pe. The purple distribution applies additional TPC data quality cuts, which are explained in more detail in the subsequent Section 6.2.2.

The sharp peak seen in all three distributions corresponds to the TPC NR events which are found in coincidence with a  $4.44 \text{ MeV } \gamma\text{-ray}$ , triggering the NV event. The uniform component of the orange and purple curve is due to accidental coincidence between the two detectors. The fact that the black histogram shows a different height for the uniform component before and after the coincidence peak can be explained by the delayed neutron capture mostly on hydrogen. This leads to an additional background component for the TPC, which increases the chance for random coincidences after the  $4.44 \text{ MeV } \gamma\text{-ray}$  events. The small rise between  $-2 \mu\text{s}$  and  $0 \mu\text{s}$  seems to be an artifact in the reconstruction which disappears when applying additional data quality cuts. This artifact can be either caused by a wrong reconstruction of the S1 center

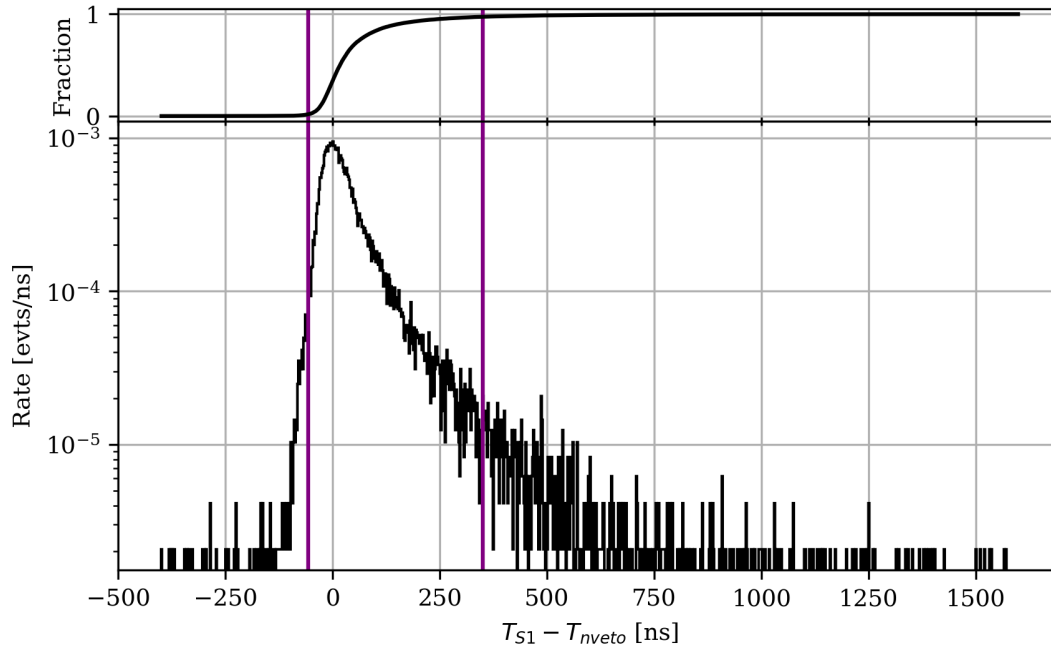


Figure 6.7. – Time distribution of the TPC S1 signals in the NR ROI with respect to triggering NV events. The main panel shows the time distribution and the upper panel its cumulative distribution. Vertical purple lines indicate the selected coincidence window region as explained in the text. The uncertainties of the individual bins are not shown.

time due to additional signal backgrounds, like PMT dark counts or afterpulses being clustered into the S1 signal. However, in this case, one would expect this effect to be more dominant for small signals which does not seem to be the case. Another explanation is that these coincidences are caused by the remaining backgrounds of neutron capture triggering events in the NV, which are paired with TPC signals induced by the 4.44 MeV  $\gamma$ -rays and lead to a mirrored exponential distribution like the one seen in Section 7.5.1 for the NV only case. This explanation seems more likely, as the surplus in coincidences disappears when selecting signals in the NR ROI region.

For the NR response calibration, the sharp peak centered around zero is of importance when defining the coincidence between the two detectors. A zoom into the region between  $-500$  ns and  $1500$  ns for all calibration events with an S1 signal inside the NR ROI is shown in Figure 6.7. The distribution is given by a sharp, normally distributed peak centered around zero on top of an exponential-like tail. The former is given by fast neutrons for which the timing is limited by the timing resolution of the two detectors. E.g. a neutron with an energy of 4.5 MeV has a time of flight of about 7 ns for a distance of 20 cm. The exponential-like tail might be either caused by a reconstruction bias or by slower neutrons which require a longer time of flight. In order to set boundaries for the coincidence window, the cumulative distribution was computed and the boundaries were set at the 2.5% and 97.5% quantile leading to a coincidence window of  $-57$  ns to  $351$  ns.

Many of the 4.44 MeV  $\gamma$ -rays, but also the 2.22 MeV  $\gamma$ -rays from the hydrogen neu-

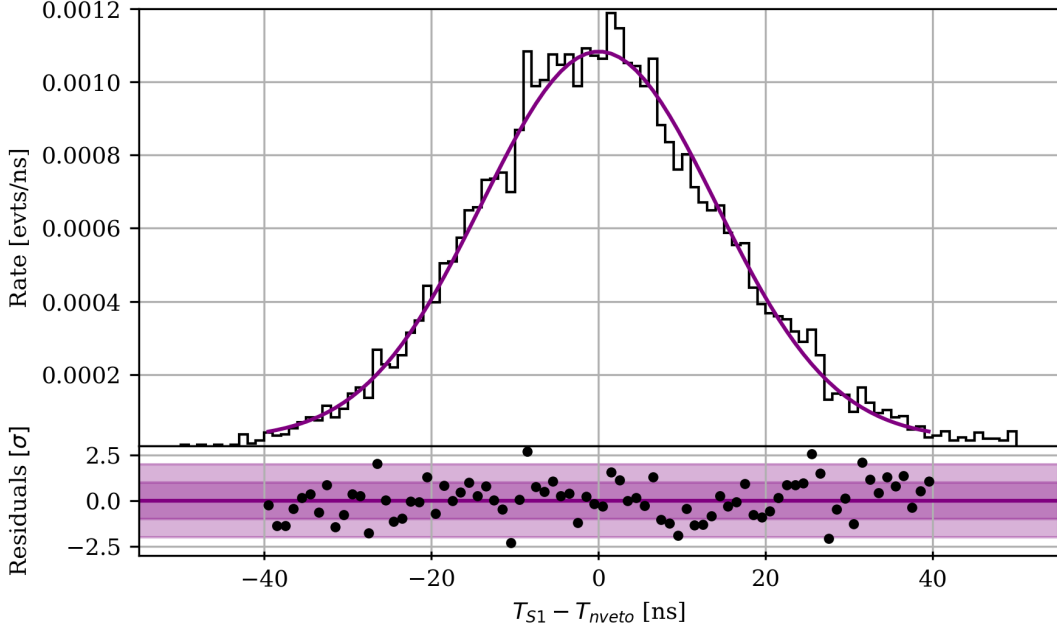


Figure 6.8. – Time resolution of the NV TPC coincidence for the NR response calibration. The black distribution shows the distribution of the S1 center time relative to the NV event center time of coincident events. Only high energy Compton events surviving all TPC data quality cuts are shown. The purple curve shows the best fit to the data as described in the text. The lower panel illustrates the residuals of the fit. The purple line represents the fit function and the shaded regions are the 1 and 2 sigma contour. The black dots represent the relative difference between model and data evaluated at the center of each bin. The uncertainty of the binned data is assumed to follow Poisson statistics, but is not shown.

tron capture, produce signals in both TPC and NV. Thus these signals can be used to determine the overall timing resolution between the two detectors. Compton signals are selected by requiring a cS1 between 5000 pe and 10 000 pe corresponding to an energy range of about 550 keV to 1250 keV with a median of 870 keV in the combined energy scale<sup>2</sup>. The resulting time distribution is shown in Figure 6.8.

Similar to the time distribution of the neutrons, the time distribution shows a peak that consists of a normal distribution centered around zero, but without an exponential-like tail. This supports the hypothesis that the tail seen in Figure 6.7 is due to slow neutrons and not an effect of wrong reconstruction. In order to determine the timing resolution of the two detectors a simple fit model of the form

$$f(x; A, \mu, \sigma, c) = A_1 \cdot \sqrt{2\pi}\sigma N(x; \mu, \sigma) + c \quad (6.1)$$

was fitted to the data using a least-squares method.  $N$  refers to a normal distribution with a mean value  $\mu$  and a standard deviation  $\sigma$ . A constant offset  $c$  is included. The

<sup>2</sup>The energy range was chosen a bit arbitrarily. The lower boundary is motivated by the fact that the S1 signal shape should be well-defined. The upper boundary is motivated by the energy threshold of the NV which is 1 MeV as it is discussed in Section 7.2.3. Finally, the selection should contain sufficient statistics.



amplitude of the normal distribution is modeled via  $A$ . The exact fit results can be found in the appendix, in Table 6.3. The uncertainty of the binned data was estimated assuming Poisson statistics. The uncertainty of the run time is negligible. The fit range was restricted to  $\pm 40$  ns such that each bin contained at least 10 entries. The  $\chi$ -square and the number of degrees of freedom of the fit are 86.98 and 76 respectively which yields a p-value of 0.183 accepting the fit model. The timing resolution is given by the standard deviation of the distribution which is  $(14.03 \pm 0.23)$  ns. This is of the same order of magnitude as the 10 ns time resolution of the TPC digitizers. The accuracy of the synchronization of the two detectors is given by the position of the peak and was found to be  $(-1.21 \pm 0.24)$  ns. For signals with a smaller cS1 the resolution gets worse. A comparison between the best fit of the timing resolution and the timing spectrum for NR events, showing this effect, is added to the appendix in Figure B.20 on page 252.

After evaluating the timing resolution and the size of the coincidence window, the number of accidental coincidences between the two detectors can be evaluated, using the dark gray shaded region between  $-5 \mu\text{s}$  and  $-3 \mu\text{s}$  shown in Figure 6.6 as a reference region. Requiring only that a TPC event must be within the ROI of the NR calibration (orange curve), 106 events are found in the reference region. If all additional TPC data quality cuts are applied, the number reduces to 10 events (purple curve). Scaling these numbers to the actual size of the coincidence window of 408 ns, this leads to  $(18.61 \pm 1.80)$  and  $(1.75 \pm 0.55)$  accidental coincidences respectively, for the entire SR0 NR calibration, making the calibration almost background-free.

The strong background suppression by this detector coincidence is evident when comparing the cS2 versus cS1 data space before and after applying the coincidence as shown in Figure 6.9. A few distinct features are removed by the coincidence cut. Prior to applying the cut, the data shows a vertical distribution of events for cS1 events below 30 pe, labeled "A" in the figure. This distribution is given by so-called accidental coincidences (ACs) which are random pairings between S1 and S2 signals. At low energies, these random pairings are dominantly caused by fake S1 signals due to PMT dark counts or single to few electron S2 signals misinterpreted as S1 signals, originating from photo-ionization in the detector after large light signals, in particular those induced by the 4.44 MeV  $\gamma$ -ray. These fake S1 signals are a dominant driver of the detector energy threshold. At larger cS2 values, the random pairing occurs mostly between fake S1 and S2 signals produced in the gaseous part of the TPC, e.g., between the top screening mesh and anode mesh. ACs can also be produced by random pairs of photo-ionization S2 signals and lone S1 signals which are produced by interactions

Table 6.3. – Best fit values of the  $\gamma$ - $\gamma$ -coincidence peak and timing resolution of the detector coincidence. The fit model is described in equation (6.1)

Parameter	Value
$A$	$(8.9 \pm 0.1) \cdot 10^{-4}$ evts/ns
$\mu$	$(0.05 \pm 0.16)$ ns
$\sigma$	$(14.03 \pm 0.23)$ ns
$c$	$(0.3 \pm 0.1) \cdot 10^{-4}$ evts/ns

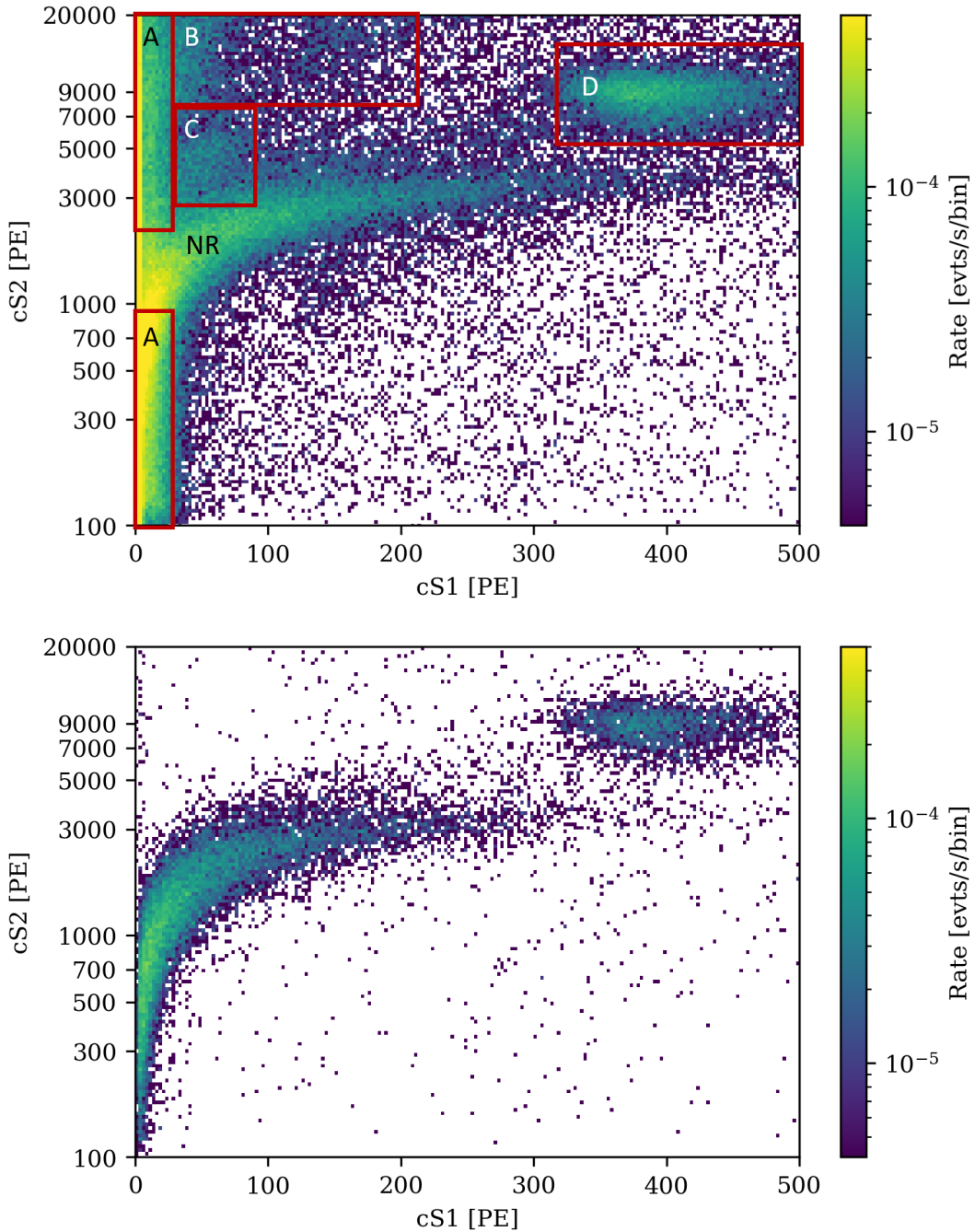


Figure 6.9. –  $cS2$  versus  $cS1$  before and after applying the NV TPC coincidence. The upper plot shows the entire AmBe calibration data before, and the bottom panel after applying the coincidence requirement. The data of all three source positions are shown. The labeled regions of the upper figure are explained in the text.

below the cathode mesh. However, given the source positions used during the AmBe calibration, gas events are much more dominant than below-cathode events, as can be

inferred from the plots shown in Figure 6.4.

Starting from the vertical ACs distribution, a banana-shaped band, labeled as "NR", ranges from cS1 values of 0 pe up to about 400 pe in both plots of Figure 6.9. This band is caused by NR interactions due to the emitted neutrons. The additional peak, labeled "D", is due to inelastic NR on the isotope  $^{129}\text{Xe}$  which emits an additional 39.6 keV de-excitation photon [McC16]. Above the band and on the left-hand side of the 39.6 keV line gas events can be seen in the region labeled as "B", prior to applying the coincidence between the two detectors. These events exhibit a short drift time, and both S1 and S2 show a very large area fraction top<sup>3</sup> which is typical for these types of events. The events in the last region "C" are most likely ER signals originating from the detector background.

After applying the coincidence cut the NR band structure and the 39.6 keV line from inelastic NR interactions are becoming the most dominant features of the data space. The total event rate of the shown events, which are dominated by ACs, reduces from  $(1.699 \pm 0.003)$  Hz to  $(0.088 \pm 0.001)$  Hz after applying the coincidence requirement. If one applies all additional data quality cuts discussed in the subsequent section prior to the coincidence cut, the rate reduces only from  $(24.45 \pm 0.03)$  mHz to  $(10.49 \pm 0.21)$  mHz due to the coincidence cut. This reduction by  $(43 \pm 1)\%$  is similar as expected based on the acceptance losses due to the selection criteria s for the 4.44 MeV  $\gamma$ -ray, discussed above. A back-of-the-envelope calculation yields an identical reduction factor of

$$\epsilon_{tot} = \epsilon_{C12} \cdot \epsilon_{4PMTs} \cdot \epsilon_{area} \cdot \epsilon_{ct} \approx 0.5 \cdot 1 \cdot 0.89 \cdot 0.95 = 0.42$$

where  $\epsilon_{C12}$  is the probability of  $^{12}\text{C}$  being in the first excited state, which is assumed to be 50%<sup>4</sup>,  $\epsilon_{4PMTs}$  is the 4-fold coincidence requirement which is assumed to be about 1,  $\epsilon_{area}$  is the acceptance loss due to the area selection and  $\epsilon_{ct}$  represents the acceptance of the center time cut. Figure B.21 on page 253 in the appendix shows a direct bin-wise comparison of remaining events after applying the NV-TPC coincidence requirement, which yields a similar result.

### 6.2.2 Data quality cuts, fiducial volume and NR band data

Although the coincidence between the two detectors shows already a very strong suppression of background events, additional data quality cuts are required to refine the data further. In general data quality cuts refine data by removing "nonphysical" events, e.g. random pairs of S1 and S2 signals or fake S1s by testing various properties of the S1 and S2 peaks. Each cut is usually optimized by studying the peak behavior for a variety of different additional calibration sources besides AmBe like:  $^{37}\text{Ar}$ ,  $^{85m}\text{Kr}$  or  $^{220}\text{Rn}$  which can be diluted uniformly in the detector. The boundaries of each cut

<sup>3</sup>A large contribution from the top PMT array to the total peak area.

<sup>4</sup>In Chapter 7 a detailed study about the AmBe spectrum is given. The chapter shows that the branching ratio for the first excited state is close to 50% for the used calibration source.

are optimized such that a high acceptance of "good events"<sup>5</sup> is preserved. Most of the data quality cuts were developed by other XENONnT analysts. In this section only the most relevant cuts which show the highest impact on the NR calibration data are discussed. Additional data-quality cuts, which only minimally impact data, are reported in the appendix in Section A.4. Many of these cuts test the quality of S1 and S2 signals based on either their shape or the hit-pattern distribution in the top/bottom PMT array. These tests either apply empirically optimized boundary cuts or pursue likelihood approaches or machine learning techniques.

Besides refining the calibration data, it is important to assess the acceptance of the individual cuts. In XENONnT the cut acceptance for many of the cuts is assessed using the so-called "n-1 method". In this method, the number of surviving events with "n" data quality cuts applied, is compared against the number of remaining events when only "n-1" cuts, excluding the one under study, are applied. A more thorough explanation and discussion of the n-1 method and the overall cut acceptance will follow in the subsequent Section 6.2.3. For the moment the reader can refer in this section to the "n-1" cut acceptance as a "fraction of remaining events" for the individual data quality cuts shown.

Due to the higher interaction cross sections of neutrons compared to WIMPs, neutrons can perform multiple scatters before leaving the sensitive region of the TPC. These interactions are seen as a single summed S1 signal which will either be paired with the largest of the multiple S2 signals if the scatters happen at different z-positions in the TPC or with a summed S2 signal if the separation in z is below the TPC's resolution. This leads to a bias in the NR response calibration for dark matter signals which only scatter once. Multi-scatter events are most dominantly cut by the so-called S2 single-scatter cut and the S2 width cut.

The **S2 single-scatter cut** is defined in the alternative cS2 area as a function of cS2 area space. An alternative S2 is defined as being the second largest S2 within a valid drift time range of the main S1 signal. The cut boundary was optimized using different calibration sources and is shown in the appendix on 253 in Figure B.22. The effect of the cut on the data is shown in Figure 6.10. It primarily removes events on the lower side of the NR band which can be understood from the summing of S1 signals while having same sized S2-signals as for a single-scatter event. In addition, it can be seen that the cut has a very strong impact on the inelastic 39.6 keV line from  $^{129}\text{Xe}$  which by definition is a two-interaction process.

The width of an S2 signal strongly depends on the diffusion of the electron cloud initially produced at the interaction position<sup>6</sup>. Thus it has a strong correlation with drift time or the depth z where the interaction happened. The **S2 width cut** exploits this relation by testing if the width of an S2 signal is as wide as expected depending on its area and drift time. Therefore, it acts as an additional ACs cut against wrongly paired S1 and S2 signals, but also as a cut against MS events for which the two electron clouds produced at the different interactions sites cannot be resolved as individual peaks. The cut is defined in the area-normalized S2 width versus drift time, as

<sup>5</sup>A "good" event refers to events which would also be caused by the unknown particles under study, e.g. WIMPs.

<sup>6</sup>See also Section 2.2

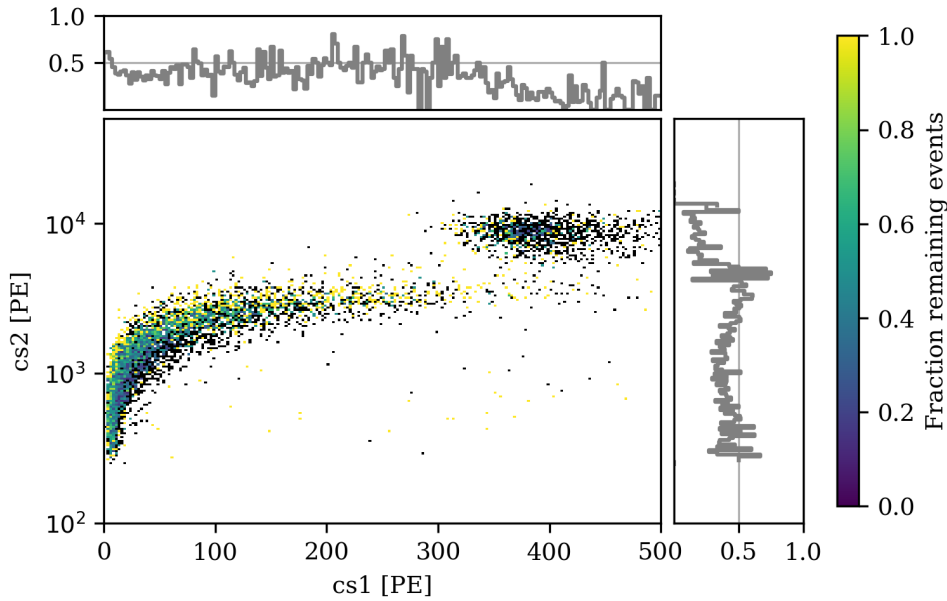


Figure 6.10. – Impact of the S2 single-scatter cut on the NR calibration data. The main graph shows  $cs2$  versus  $cs1$  space. The color scale encodes the fraction of remaining events after applying the cut using the  $n-1$  method. Bins colored in black contained events before the cut, but are empty afterwards. The two plots above and on the right hand side of the main panel show a projection of the fraction of remaining events on the corresponding axis. Note that the  $y$ -axis of the main panel shows a logarithmic scaling.

discussed in [Apr+19c]. It was optimized based on different uniformly distributed calibration sources. In contrast to the S2 single-scatter cut, the S2 width cut primarily cuts at the upper half of the NR band as shown in Figure 6.11. This can be understood by computing a simple example based on the CY and light yield (LY) shown in Figure 2.4 on page 23 as well as the values derived for  $g_1$  and  $g_2$  discussed in Section 8.1. Table 6.4 compares the average number of photons and electrons generated for the given CY and LY as well as the recorded S1 and S2 produced by a single  $E = 10 \text{ keV}_{\text{NR}}$ <sup>7</sup> interaction compared with a  $E = 4 \text{ keV}_{\text{NR}} + 6 \text{ keV}_{\text{NR}}$  multi-scatter interaction. The example shows that a multi-scatter signal tends to have on average less photon and more electrons compared to a regular single-scatter interaction depositing the same amount of energy. Thus, multi-scatter events are shifted diagonally above the single-scatter NR band as it is observed.

One last cut which was optimized based on the AmBe calibration data during this study is the fiducial volume cut. Due to the chosen calibration positions most of the NR interactions are located at the very edge of the detector as it is shown in Figure

<sup>7</sup>The unit  $\text{keV}_{\text{NR}}$  is a LXe TPC typical jargon and refers to the fact that the interaction is caused by a nuclear recoil instead of an electronic recoil. This means that the number of generated quanta is quenched. This quenching leads to different signals for ER and NR for a given energy deposit. See also Chapter 2 for a more detailed explanation.

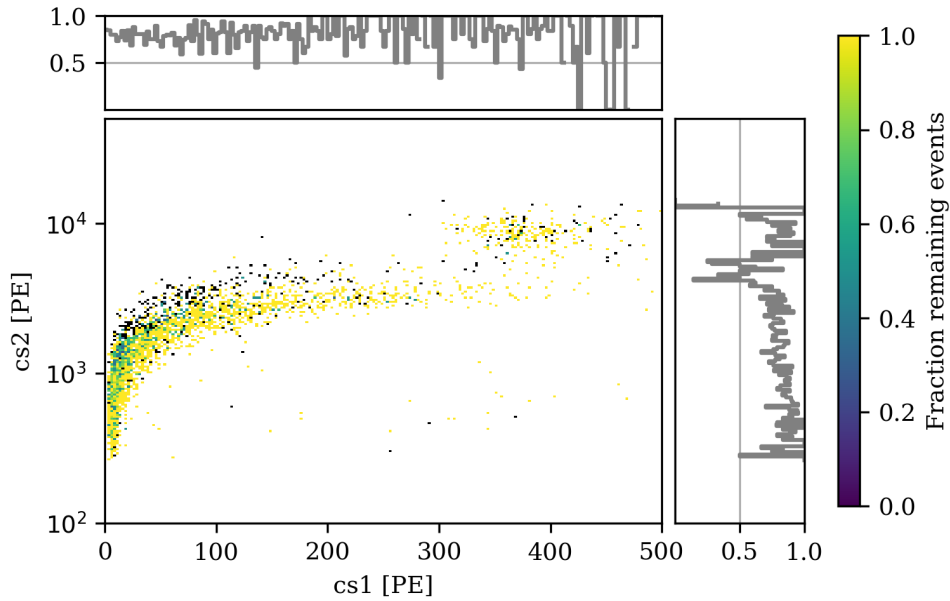


Figure 6.11. – Impact of the S2 width cut on the NR calibration data. The graph shows the same as explained in Figure 6.10, but for the S2 width cut instead.

6.4. Since this region is in general outside of the fiducial volume used in the WIMP analysis a separate fiducial volume must be defined for the calibration. The WIMP fiducial volume is optimized based on simulations to maximize the overall fiducial mass while mitigating backgrounds from neutrons, low energy ER, and other backgrounds as much as possible. Further, it should avoid detector regions in which signal reconstruction is difficult. Given the strong suppression of backgrounds by the detector coincidence requirement, only the latter is important for the NR calibration data. The most problematic region for the reconstruction is the "charge insensitive region" of the TPC in which electrons of an S2 signal are either partially or completely lost during the drift process. This happens especially close to the walls at large depths of the detector, where some of the field lines end up on the PTFE walls of the TPC. This

Table 6.4. – Comparison of the average number of photons and electrons generated by a 4 keV plus 6 keV multi-scatter NR interaction and a 10 keV NR single-scatter interaction. The corresponding S1 and S2 signals were estimated by multiplying the average number of photons and electrons with  $g_1$  and  $g_2$  discussed in Section 8.1. The average drift field used in this estimate is 22.92 V/cm. The number of generated photons and electrons were estimated via NESTpy [Far+23].

Energy [keV <sub>NR</sub> ]	#Photon	#Electrons	S1	S2
4	26.39	24.06	4.00	395.80
6	44.72	34.20	6.78	562.63
4+6	-	-	10.77	958.43
10	86.76	51.79	13.14	852.05



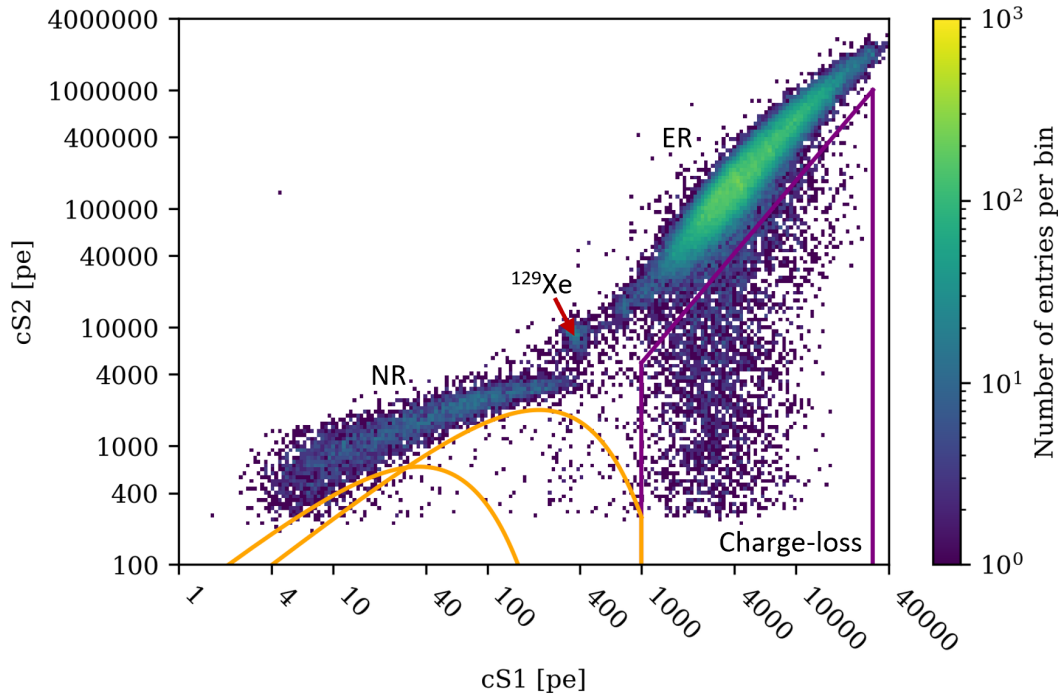


Figure 6.12. –  $cS2$  versus  $cS1$  space all AmBe calibration events surviving the detector coincidence selection and the TPC data quality cuts. Events which are suffering from charge-losses can be found in the orange and purple region.

region depends on the configuration of the drift field, given the voltage configuration of the cathode and gate electrodes as well as the field shaping rings. Charge-up of PTFE walls plays an important role, too, as it is an isolating material. The effect was evaluated by different analysts performing field simulations with COMSOL [COM] and data-driven studies using  $\alpha$ -decays from  $^{222}\text{Rn}$  and  $^{218}\text{Po}$  originating from the PTFE. A plot of the simulated electric field and a simulation-driven map of complete charge-loss can be found in the appendix in Figure B.12 on page 246.

Given that the charge-loss effect purely depends on the field geometry and is independent of whether an interaction is of type NR or ER, another data-driven method was chosen for optimizing the fiducial volume of the AmBe NR response calibration. This method is based on high energy ER interactions originating from the Compton continuum of the 2.22 MeV and 4.44 MeV  $\gamma$ -lines which are much more abundant than NR events. Figure 6.12 shows all AmBe calibration events after applying the detector coincidence and all TPC data quality cuts on a double logarithmic scale in the  $cS2$  versus  $cS1$  space. The plot shows two almost linear bands, the lower one of which belongs to NR interactions and the upper one to ERs. Charge-loss events are expected to be found in the orange<sup>8</sup> and purple colored regions of Figure 6.12. They are characterized by a vertical distribution ranging from the main NR and ER bands down to

<sup>8</sup>The orange region has this particular shape as the selection was done in a different parameter space, which is shown in the appendix in Figure B.24 on page 255.



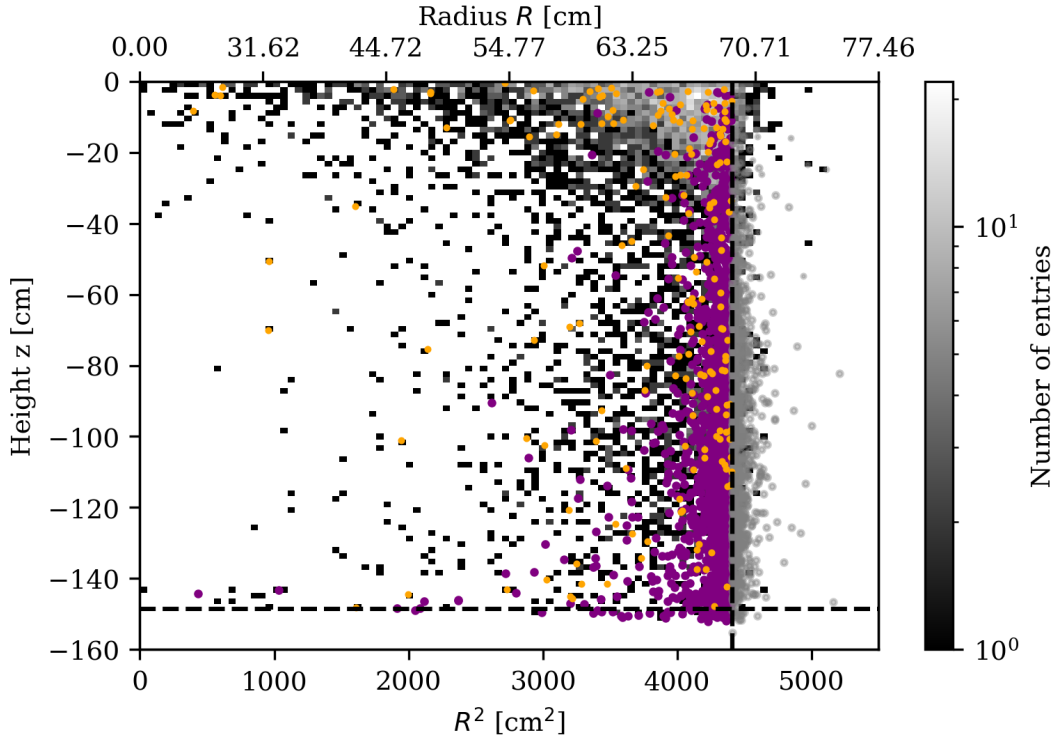


Figure 6.13. – Spatial distribution of charge-loss events and events in the NR ROI. The NR ROI data is shown as a gray-scaled histogram. Events found in the purple and orange region of Figure 6.12 are plotted as points of the same color. Gray points are events found in the purple box, but which were reconstructed outside of the physical region of the TPC. The black vertical and horizontal line indicate the position of the TPC wall and cathode electrode respectively. Note that events reconstructed outside of the detector are an artifact of the reconstruction algorithm used.

200 pe in S2, which is the lower limit as it is the nominal S2 threshold in XENONnT. A spatial comparison of the events found in these two regions compared against all events found in the NR ROI for cS1 signals below 300 pe area is shown in Figure 6.13. Events found in the purple box are more dominantly found near the wall of the TPC and at larger depths, while events selected by the orange contours are following more closely the general distribution of NR events.

The fiducial volume for the NR calibration is given by a simple cylinder. Its radius is optimized by computing the difference between the remaining fraction of charge-loss events within the purple box and the fraction of remaining events in the NR ROI. The highest separation was found at a maximum radius of 64.75 cm. The discrimination parameter as a function of radius is attached in Figure B.23 of the appendix on page 254.

The limits in  $z$  were chosen to be 4 cm and 10 cm away from the gate and cathode mesh respectively. The lower boundary is set to avoid any charge losses due to the cathode mesh itself. The upper boundary discriminates primarily gas events between anode and top screening mesh. The resulting AmBe fiducial volume is shown in Figure 6.14 together with the nominal fiducial volume used in the WIMP analysis. The

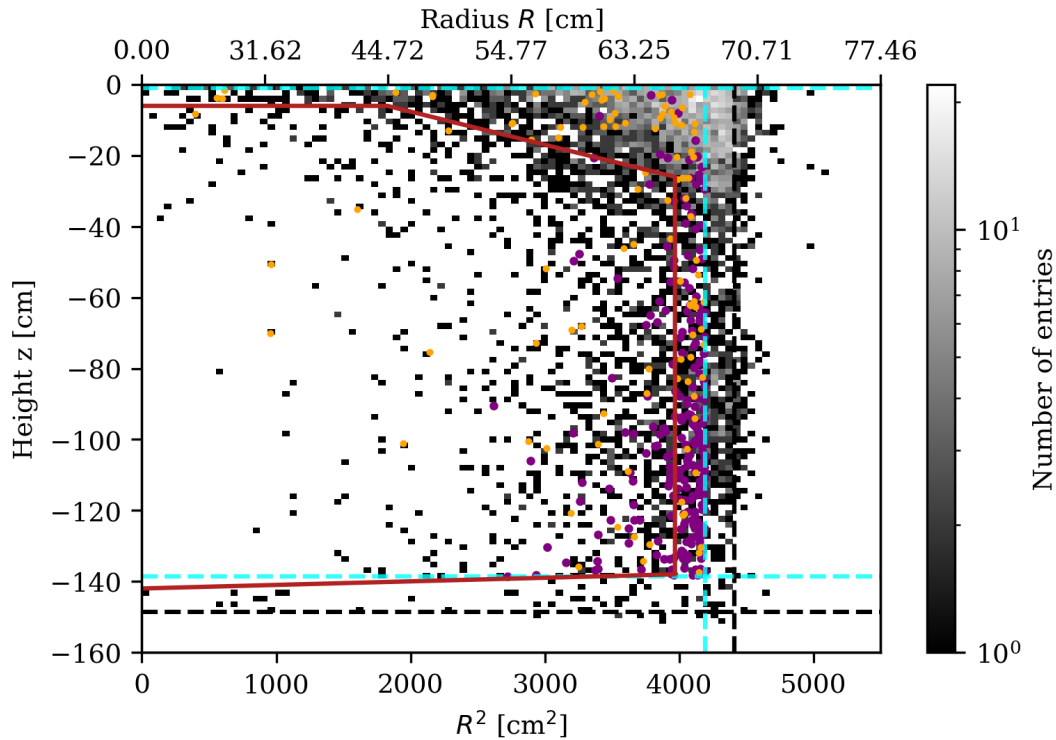


Figure 6.14. – Optimized AmBe fiducial volume and NR calibration data. The figure shows the same data selection as in Figure 6.13, after applying the AmBe fiducial volume cut explained in the text. In addition the plot shows the AmBe fiducial volume in cyan, while the purple line shows the regular WIMP fiducial volume.

fraction of high-energy Compton events which can be found in the charge-loss selection reduces by a factor of about 6 from 6.3 % to 0.9 %, while the fraction of outlier events in the orange box in the NR ROI region reduces only from 5.1 % to 3.3 %. The remaining purple charge-loss events are more dominantly found at larger depths in the detector as expected by the other studies and simulations of charge-loss events. Events in the orange selection are less affected by the optimization of the fiducial volume and more closely follow the spatial distribution of NR events in general. This may point to a different origin for this atypical  $cS2/cS1$ -ratio. Another explanation might be that these events are given by a mixture between charge-loss events and multi-scatter events passing the S2 single-scatter cut. The plot also shows that most of the NR calibration data is found outside of the WIMP fiducial volume. This is not an issue per se as the detector response model is mostly determined by in situ calibration gases like  $^{37}\text{Ar}$ ,  $^{85m}\text{Kr}$  and  $^{220}\text{Rn}$  [Apr+22c] which are directly diluted uniformly into the LXe. The main purpose of the NR calibration on the other hand is to determine the NR response model which is given by the micro-physics of liquid xenon as it is discussed in Section 6.3. Only differences in the electric field need to be considered. The relative change of light and charge yield, as predicted by NEST for the XENONnT drift field, is shown in Figure B.26 in the appendix on page 256. For a 5 keV nuclear

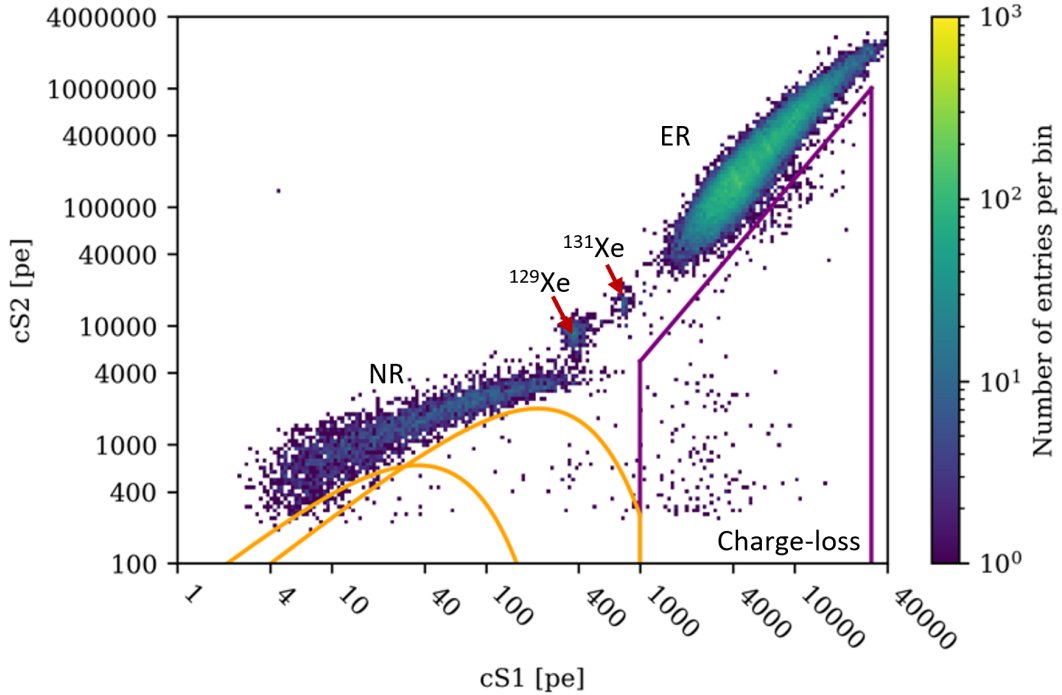


Figure 6.15. – *AmBe* charge-loss events after applying the *AmBe* fiducial volume cut. The plot shows the same as in Figure 6.12, but after applying the *AmBe* fiducial volume cut.

recoil, a variation less than 1 % is predicted by NEST for both yields.

Figure 6.15 shows the remaining events in the double log-scale  $cS2$  versus  $cS1$  space after applying the *AmBe* fiducial volume. Besides showing fewer events in the two highlighted regions two other features can be observed. Compared to Figure 6.12 the lower edge of the ER band moved slightly upwards from about (1000 pe, 10 000 pe) in ( $cS1$ ,  $cS2$ ) to (3000 pe, 40 000 pe) which can be explained by the fact that it becomes harder to deposit low energetic ER signals without performing a second scatter in the active LXe volume and thus being cut by the  $S2$  single-scatter cut. As a consequence of the larger ER "threshold" not only the inelastic NR scatter on  $^{129}\text{Xe}$  at 39.6 keV between NR and ER band, but also the less dominant inelastic scatter line on  $^{131}\text{Xe}$  at an energy of 80.2 keV becomes visible. Both lines are subdominant compared to regular elastic NR-scattering due to their smaller interaction cross-section [McC16]<sup>9</sup>.

The total number of remaining events after applying the individual data selection cuts including: the coincidence between TPC and NV, each data quality cut as well as the fiducial volume cut are summarized in table A.1 in the appendix on page 233. In total 2537 events remain in the NR ROI for  $cS1$  areas below 300 pe area, corresponding to a rate of  $(10.49 \pm 0.21)$  mHz. Restricting the dataset to the Top CW5d9m position, which is closest to the TPC, the rate increases by a factor two to  $(20.42 \pm 0.49)$  mHz. For the NR response model fit discussed in Section 6.3 only events with a  $cS1$  below 100 pe are considered which reduces the number of events to 1986, see also the Table

<sup>9</sup>See also Figure 2.2 on page 19.

**A.1** on page 233. The final selection of events is shown in Figure 6.16 together with the low ER calibration data from  $^{220}\text{Rn}$  and  $^{37}\text{Ar}$ . The NR band is clearly visible below the ER band. Both bands show a non-negligible overlap in the region below  $25\text{ keV}_{\text{NR}}$  which is reflected in the ER leakage discussed in Section 6.3. Few outlier events can still be observed below the NR band. These outliers are probably either remnants of multi-scatter events or charge-loss events. The latter cause seems to be less likely as these events behave differently compared to the charge-loss events at higher energies. Most of these events show alternative S2 signals at the size of a few electrons and thus appear to be remnant pollution of multi-scatter events. A plot illustrating this hypothesis can be found in the appendix in Figure B.24 on page 255. One consequence of this issue is a raise of the S2 threshold from 200 pe in XENON1T to 250 pe in XENONnT for the NR calibration data. In total 46 events remain in the orange region below the NR band after applying all data quality cuts and thresholds.

Raising the S2 threshold for the calibration data to a larger value than ultimately used in the inference<sup>10</sup>, discussed in Chapter 8, certainly impacts the validity of the NR response model for smaller WIMP masses. However, it might be even more harmful to use calibration data which has a higher pollution of multi-scatter NR signals, as WIMPs are expected to perform single-scatters only. In total 67 events are cut by the higher S2 threshold, which corresponds to a loss of 3% of calibration events. Thus, even if these events would be included their impact on the response model might be negligible.

The outliers above the NR band coincide with the ER band. Thus remaining contamination from ER might be a possible explanation. However, this hypothesis is in tension with two observations. First, when evaluating the number of accidental coincidences between NV and TPC, only  $(1.75 \pm 0.55)$  events were found, which is negligible compared to the total number of remaining NR calibration events. Hence, a contamination due to accidental coincidence with ER background signals seems unlikely. A second possibility consists of ER energy deposits from  $\gamma$ - $\gamma$ -coincidence between the two detectors, as used for the estimate of the coincidence timing. However, most of these events create Compton signals which only appear at larger cS1 and cS2 values. In addition, given that the fiducial volume cut increased the threshold for ER interactions way beyond the 80.2 keV inelastic line from  $^{131}\text{Xe}$ ,  $\gamma$ - $\gamma$ -coincidence seem also to be an unlikely explanation. It is more plausible that also these events are caused by multi-scatter events in which the two S2-signals are not resolved, and the combined S2 signal is also not cut by the S2-width cut. One indicator of a remaining pollution of multi-scatter events is the fact that, after applying all data quality cuts, the two inelastic xenon lines are still visible which are by definition two interaction processes. However, for both lines, the two interactions are very hard to resolve as the free mean path of the respective photons is in the millimeter range<sup>11</sup>. This remaining pollution of MS events below and above the NR band must be considered when estimating the NR response model. However, before the response model fit can be performed in Section 6.3, one last key ingredient, besides the calibration data itself, is the missing total cut acceptance, which is discussed in the next section.

<sup>10</sup>For the final SR0 WIMP results, a threshold of 200 pe is used.

<sup>11</sup>See Figure 2.1 on page 18

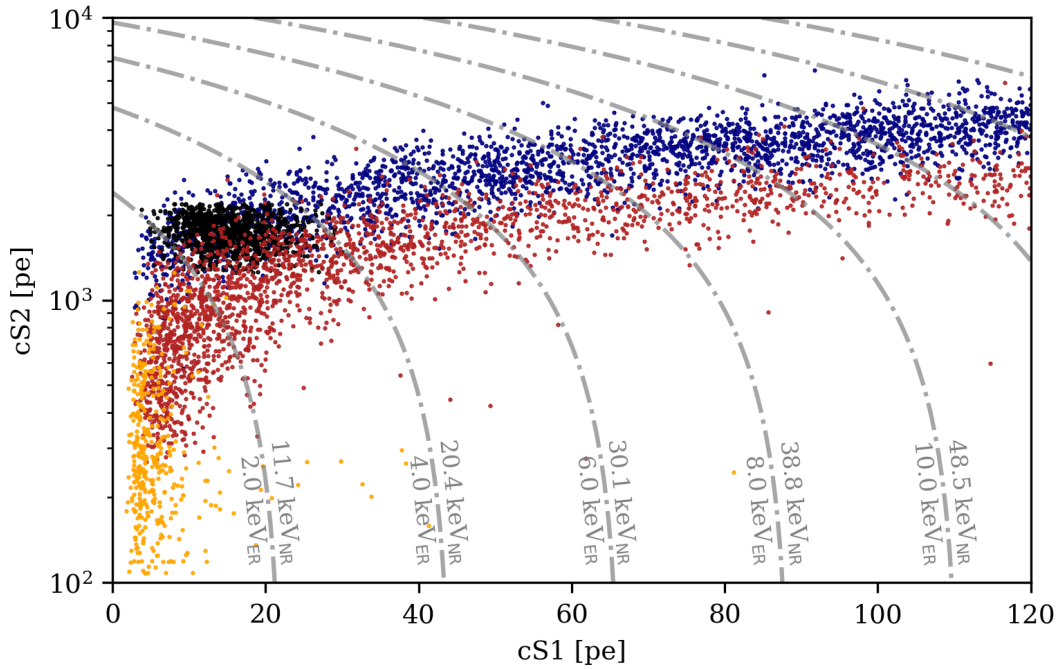


Figure 6.16. – *AmBe NR and ER calibration data passing all data quality selections. The red dots show the remaining NR events including the S1 tight coincidence threshold and the S2 area threshold, while the orange points show the additional NR events when dropping these thresholds. The blue points show  $^{220}\text{Rn}$  ER calibration events, while the black dots indicate the mono-energetic 2.8 keV line of  $^{37}\text{Ar}$ . The gray dashed-dotted lines are showing the energy contours of the combined energy scale computed according to the NR response model described in Section 6.3. A version with linear cS2 axis can be found in the appendix in Figure B.27 on page 257.*

### 6.2.3 Cut acceptances, correlation and total cut acceptance

When accessing the total cut acceptance one needs to distinguish two general types of cut categories:

The first category consists of exposure cuts. The most prominent type of such cuts reduces either the fiducial mass of the experiment or the lifetime, e.g. the NV and muon-veto cuts. Generally speaking, these cuts only affect the total exposure of the experiment and are acting independently of the interaction type or energy deposits<sup>12</sup>. Thus these cuts have a flat cut acceptance for the ROI data space in cS1 and cS2. In the case of the NR calibration, the coincidence between NV and TPC acts as an exposure cut. Uncertainties introduced by these cuts, e.g. in the total exposure of the experiment are later translated into uncertainties in the WIMP rates and are considered as nuisance parameters in the final inference discussed in Section 8.3.

The second category of cuts comprises regular data quality cuts which depend on the S1 and S2 waveform properties and therefore on the type of interaction and en-

<sup>12</sup>It should be noticed that this assumption is only approximately true in the case of the fiducial volume cut as the precision of the x/y position reconstruction depends on the S2 area.

ergy deposits. These cuts show in general a non-uniform cut acceptance in cS1 and cS2. This non-uniformity needs to be understood and modeled in order to accurately describe the NR response.

In XENONnT the overall cut acceptance can be estimated in two different ways. In the first method, a full-chain simulation technique is used which simulates events from the particle propagation up to complete waveforms which can be put into the STRAX/STRAXEN analysis framework discussed in Chapter 3. The simulation is split into three main parts. GEANT4 [Ago+03] is used to simulate the particle propagation and energy deposits in LXe. These energy deposits are afterwards passed to Electron and Photon Instructions generator for XENON (EPIX). EPIX performs a micro-clustering of the corresponding energy deposits and computes the number of generated scintillation photons and electrons. Subsequently, the generated quanta are passed together with the current detector settings such as light propagation maps, electron lifetime, gains, etc. as simulation instructions to The XENON waveform simulator (WFSim) [Gae+22]. WFSim generates waveforms in form of raw data, which are used to determine the cut acceptances by comparing the instructions with the events returned by the XENONnT processing framework after applying the corresponding data quality cut. This, however, only works if the corresponding cut parameter space shows a very good match between simulation and calibration data.

As an alternative, the second method uses the so-called "n-1"-approach to determine cut acceptances. This method assumes that the cut acceptance of a particular cut in a given set of cuts can be estimated by comparing the number of remaining events  $k$  after applying all cuts compared to the number of remaining events  $N$  after applying all other cuts except the cut under study, that is "n-1" cuts. This approach makes two assumptions: First, after applying n-1 cuts mostly good events remain. This means that the estimated acceptance has only a little bias due to the remaining "bad" events, however, it surely also raises the question if the  $n^{\text{th}}$  cut is truly needed. Secondly, the correlation among different cuts is negligible<sup>13</sup>. While one can argue that the first point should be true for ER interactions it is certainly not for NRs. As shown in Section 6.2.2, after applying the NV and TPC detector coincidence, events that are cut most dominantly are multi-scatter interactions. As the interaction cross-section of WIMPs is much smaller compared to neutrons, it is very unlikely that they perform MS signals and therefore these signals cannot be considered as "good events". Thus the n-1 approach cannot be used for cuts that impact the SS/MS discrimination like the S2 single-scatter or S2 width cut.

For the NR response model fit, the total cut acceptance depends on both methods. On the one hand, cuts that are affecting the MS contamination in the NR calibration data must be provided using simulations. On the other hand, not all peak characteristics are yet perfectly matched between simulation and data, thus requiring the n-1 method to evaluate their cut acceptance. Therefore the cuts were divided in the following way between simulations and the n-1 approach which is further subdivided

<sup>13</sup>For example consider the case where two cuts, A and B, cut the very same 10 good events out of a hundred good events. In this case, the n-1 method would yield for both cuts an acceptance of 100% instead of 90%. The correlation among cuts is also further discussed a bit later in this section.



into S1 and S2 area space respectively<sup>14</sup>:

MS: S2 single-scatter cut, S2 width cut, S1 top and bottom pattern cuts

S1: area fraction top cut, max PMT cut, width cut and naive-Bayes cut

S2: S2 reconstructed position difference cut, area fraction top cut as well as the PMT pattern cut.

The study of the MS cut acceptance was performed by Shenyang Shi from Columbia University. Given that the NR model fit does not only get information about the MS cut acceptance from simulations, a discussion is postponed to Section 6.3.

The n-1 cut acceptance was estimated over the relevant energy range for the NR response model fit which only includes NR events in the WIMP inference ROI below a cS1 of 100 pe. Given that the information of the detector threshold is provided to Bayesian Band Fitting Framework (BBF) by other means, the S1 tight coincidence threshold, as well as the S2 threshold, are excluded in the n-1 acceptance estimate. In order to reduce the bias on the cut acceptance due to MS outlier events, and other bad events, it was required that "good events" are found within the 90 % contour of the NR band for the acceptance estimates. The contours were determined via a sliced band fitting approach which is discussed in Section 6.3 too. The resulting input spectrum is shown in Figure 6.17.

Different approaches can be chosen to determine the uncertainty of the n-1 cut acceptance. In general, the cut acceptance  $\epsilon$  follows binomial statistics as we have  $k$  successes in  $N$  trials. In this study, two different approaches were chosen to evaluate the confidence interval / uncertainty for a given  $k$  number of events after  $n$  cuts, and  $N$  number of events after  $n - 1$  cuts. The first method uses the Clopper-Pearson method to estimate the confidence intervals. This approach determines confidence intervals based on the cumulative distribution of a binomial and therefore represents an analytically exact method, which usually gives a conservative estimate of the confidence intervals [CP34]. The second approach by Marc Paterno uses Bayes theorem to convey the binomial probability density function  $P(k|\epsilon, n, I)$ , where  $I$  denotes the prior knowledge that this process follows binomial statistics, into

$$P(\epsilon|k, n, I) = \frac{\Gamma(N + 2)}{\Gamma(k + 1)\Gamma(N - k + 1)} \cdot \epsilon^k(1 - \epsilon)^{(N-k)} \quad (6.2)$$

which represents the probability density function of finding an efficiency  $\epsilon$  given  $k$  and  $N$  [Pat04]. Using this function one can estimate the cut uncertainty by the boundaries of a predefined confidence region. Given that equation 6.2 has no simple analytical expression for its cumulative distribution [Pat04] the coverage and its boundaries are computed numerically via the highest density region method explained in Section 3.3.2. Figure 6.18 illustrates the coverage for a cut acceptance of 90 % using different large sample sizes  $N$ . As expected the method shows an increased (reduced) confidence (uncertainty) when using larger  $N$ . Both methods, the Bayes approach, and the

<sup>14</sup>A short explanation of each cut is given in the supplemental material on page 232



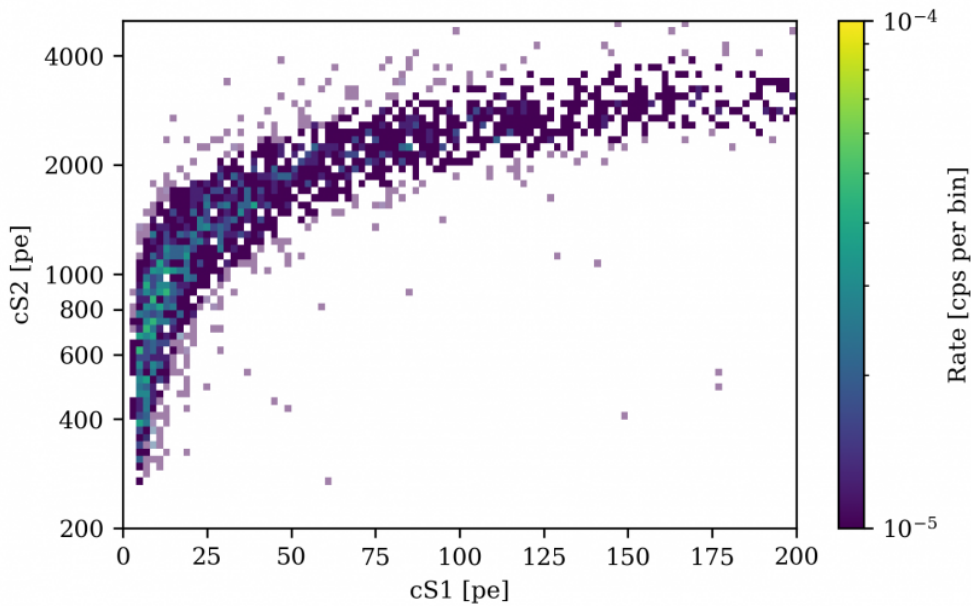


Figure 6.17. –  $cS2$  versus  $cS1$  space showing all NR single-scatter events within the 90% contour of the NR band. Bins which are shaded are not within the contour. The color scale shows the event rate per bin in count per seconds. The same set of events was also used to determine the NV tagging efficiency discussed in the following Chapter 7.

Clopper-Pearson method use a 68.27% coverage as uncertainty for the acceptance. For the presented results the Bayes approach is used.

The acceptances of the individual S1 and S2 cuts listed before are shown in the appendix on pages 257 - 260. Their combined total cut acceptance in  $cS1$  and  $cS2$  is shown in the upper and lower graph of Figure 6.19 respectively. The total acceptance is computed by multiplying the acceptances determined by the n-1 method. Due to the asymmetric uncertainties of the individual cut acceptances, the uncertainty of the total cut acceptance cannot be estimated analytically and Monte Carlo techniques are required. However, as an approximation, it was decided that the uncertainty of the total cut acceptance is estimated via Gaussian error propagation acting on the upper and lower uncertainty separately. Systematic effects introduced by this approximation should be negligible compared to the n-1 approach. The uncertainties are propagated to the NR band fitting in a way, that an additional nuisance term allows to scale of the S1/S2 acceptance up to their respective  $\pm 1\sigma$  values.

The total cut efficiency in  $cS2$  shows a nearly flat response with an average cut acceptance of 99%. In  $cS1$  the cut efficiency is lower with an average value of 91%, and between 5 pe and 25 pe the acceptance shows a small drop in efficiency. This drop is caused by the S1 width cut shown in Figure B.30 on page 258. Two ad hoc explanations for this acceptance drop are:

Single electron S2s, which have areas in the very same range as the efficiency drop, can be misclassified as S1 signals. In this case, the cut would indeed cut bad events.

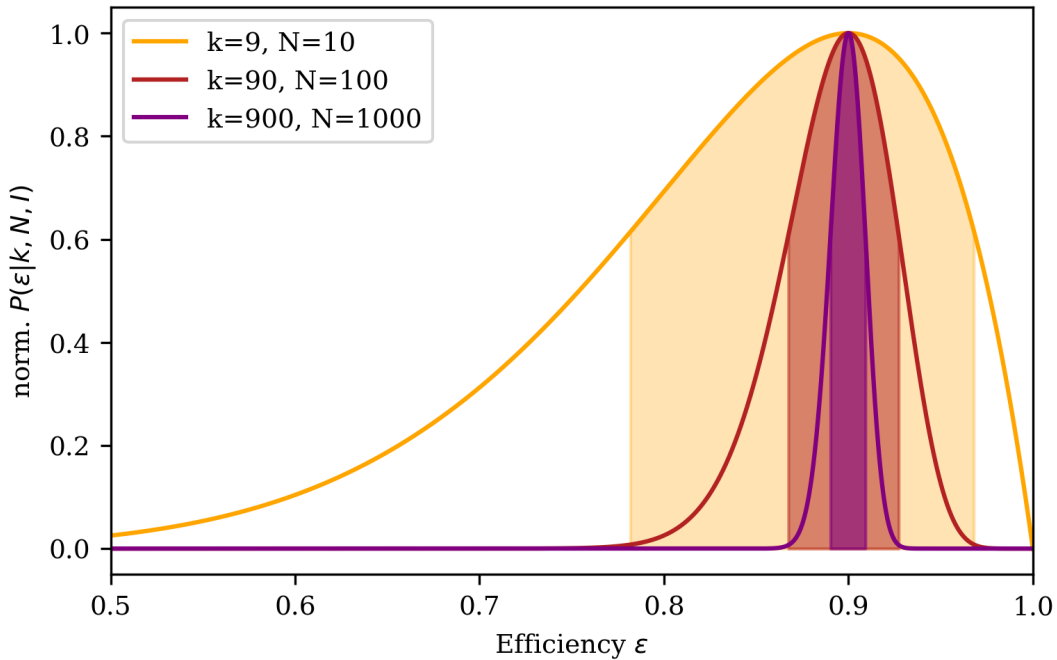


Figure 6.18. – Example illustrating the Bayesian approach for the cut acceptance uncertainty. The plot shows the probability distribution of equation (6.2) as a function of acceptance  $\epsilon$  for three different cases illustrated in the legend. The shaded regions are showing the 68.27% confidence intervals.

Checking the n-1 cut acceptance in the cS2 versus cS1 space shown in the appendix in Figure B.31 on page 259 reveals, that the events cut are mostly found within the NR band. If these signals were falsely classified random S2 signals, one would expect a more random distribution in cS2. In addition, this hypothesis is in tension with the number of random coincidences between NV and TPC assessed in the previous section.

Given that the calibration of the cut boundaries based on  $^{37}\text{Ar}$  and  $^{220}\text{Rn}$  which produce ERs a second explanation might be a slight shape difference in the S1 signals. The dependence of the scintillation behavior on the recoil type is well known [CA13], but was not observed in dedicated studies for XENONnT. Since both explanations are in tension with other observations, further study in SR1 is required.

After deriving the total cut acceptance, one last point which needs to be checked is the correlation among the used cuts to determine the cut acceptance in the n-1 approach. If two cuts are highly correlated, i.e. they target the same events, the cut acceptance will be overestimated. This can be understood from a simple example e.g. take 100 events of which 10 good events are removed by a cut A. If only the very same events are removed by a cut B, the n-1 method would yield a 100% cut acceptance for both cuts individually although in fact it is only 90%. Figure 6.20 shows the standard Pearson correlation coefficient [Pea95] for all data quality cuts applied to the AmBe calibration data. Before computing the correlation coefficient, only events below 500 pe cS1 surviving all exposure cuts were selected. It can be seen that for most of

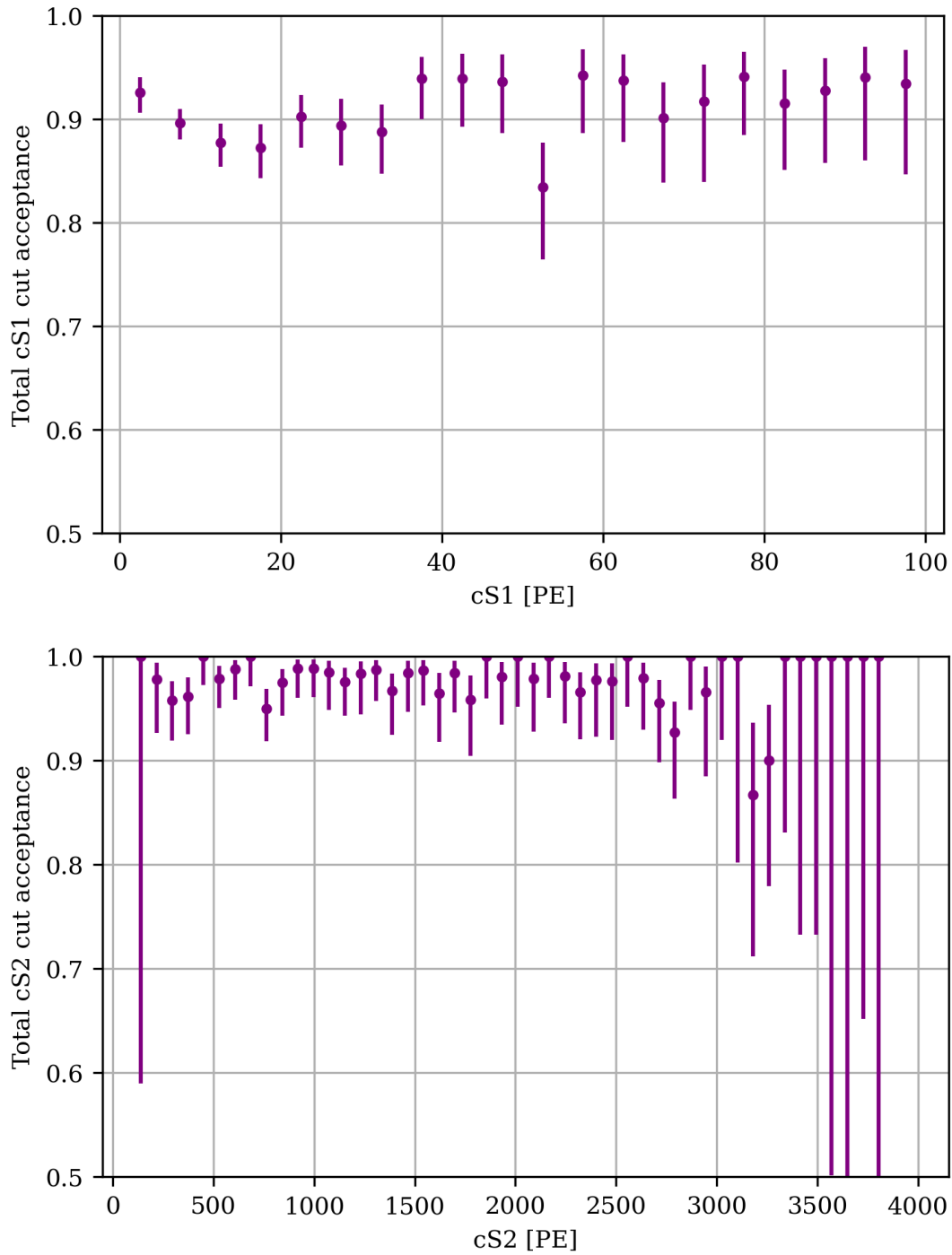


Figure 6.19. – Total cut acceptance as a function of  $cS1$  (top) and  $cS2$  (bottom) as explained in the text.

the cuts the correlation is rather small. For the cuts which are used in the  $n-1$  method, only the following cut combinations show a high correlation: `cut_s2_recon_pos_diff`  $\leftrightarrow$  `cut_s2_pattern`, and `textttcut_s1_naive_bayes`  $\leftrightarrow$  `cut_s1_width`.

## 6.2. DATA SELECTION, OPTIMIZATION AND CUT EFFICIENCY

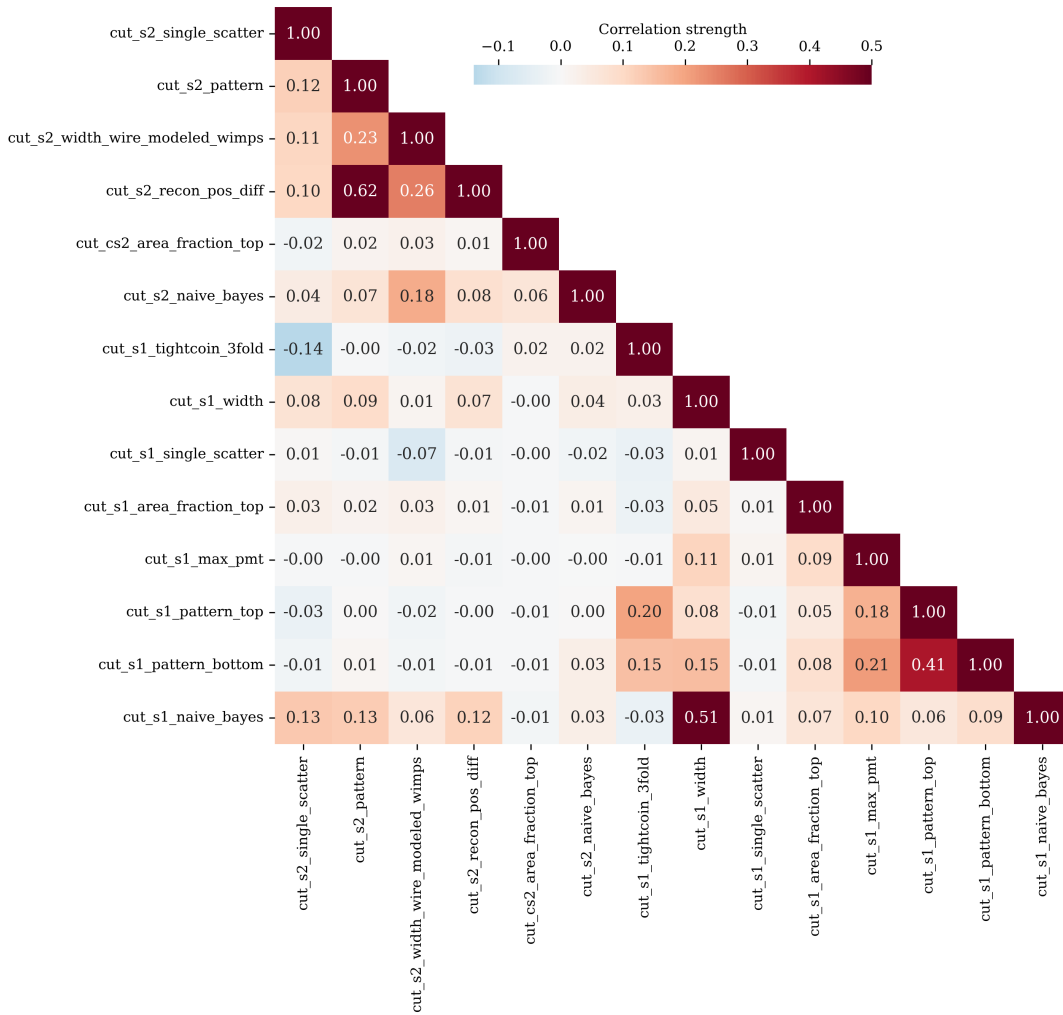


Figure 6.20. – Correlation matrix of all cuts applied to the NR calibration data. The correlation was computed after applying all exposure cuts e.g. the detector coincidence cut, and after selecting events with a  $cs1$  below 500 pe. A positive correlation means that two cuts are cutting to some degree the same events.

To evaluate if the correlations of these cuts, lead to any systematic bias dominating over statistical uncertainties, the "n-2" method can be used instead. n-2 means in this context that one compares the number of surviving events after applying all cuts with the case where all cuts, but the two correlated ones are applied. No significant bias was found. The comparison for those cuts are shown in the appendix in Figure B.35 on page 261.

In addition, a few of the other cuts show a high correlation with the S2-width cut, as well as the two S1-pattern cuts with each other. However, the cut acceptance of these cuts were determined by simulations. Thus their acceptance can be directly determined from the simulated data, and their correlation with other cuts does not bias the total acceptance estimates as it would be the case in the  $n - 1$  approach. Although

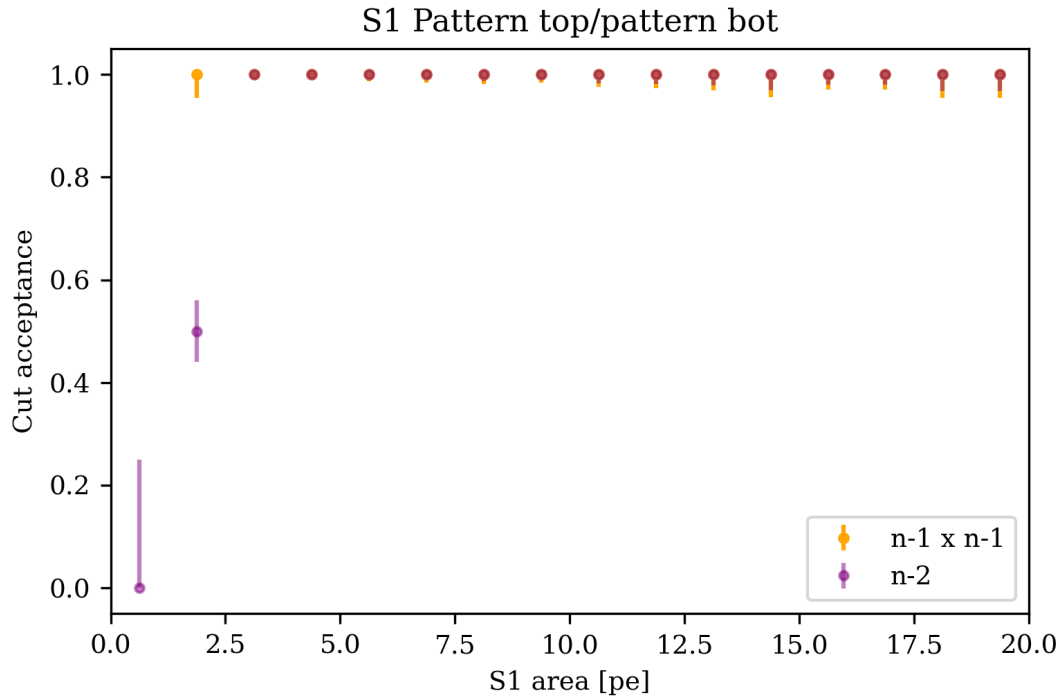


Figure 6.21. – Comparison between the product of the two  $n-1$  cut acceptances compared to the  $n-2$  cut acceptance for the S1 pattern bottom and top cuts. In the  $n-1$  approach the cut acceptance for the lowest bin is not defined as  $N = k = 0$ .

evaluated via simulations, the cut acceptances of the correlated S1-pattern cuts can be used as an illustrative example for the systematic bias which is introduced if their cut acceptance would be evaluated with the  $n-1$  method instead. The comparison of their  $n-1$  acceptance with their  $n-2$  acceptance is shown in Figure 6.21. The cut acceptance evaluated with the  $n-2$  method drops to 50% for small  $cS1$  which is not the case for the product of the two individual  $n-1$  cut acceptances.

Since the  $n-1$  total cut acceptances derived in this section show no significant bias due to correlations, the acceptances can be passed along with the NR calibration to the BBF fitting framework to infer the NR response model. This procedure will be discussed in the following closing section of this chapter.

### 6.3 NR BAND FIT AND RESPONSE MODEL

The last section of this chapter deals with the results derived from the NR calibration data which are twofold: First, the median and contours of the band are determined by fitting the data in slices of  $cS1$ . The contour is used in the previous section to determine the cut acceptance. Based on these contours, a blinded region for the WIMP analysis of SR0 can be defined. The second part of this section discusses the NR response model fit using BBF which is required for the final WIMP analysis discussed in Chapter 8. The fit has the main purpose to determine the micro-physics parameter of liquid xenon, e.g.,

the electron-ion recombination or the Lindhard and Penning quenching as discussed in Section 2.1.

### 6.3.1 NR band contours

To define the NR band contours, the calibration data of Figure 6.16 is sliced in cS1 and fitted for each slice along the  $\log_{10} |cS2/cS1|$  data space. The size of the slices varies between 4 pe and 15 pe. Each slice was fitted using an unbinned maximum likelihood method with a normal distribution summed with a uniform background distribution as model. An example fit for a single slice is shown in the appendix in Figure B.36 on page 262. For each slice, the upper and lower boundary of the 90 % contour were estimated by exploiting the correlation between area and standard deviation of a normal distribution given its normalization constant. The uncertainties were computed using Gaussian error propagation. The derived contour points were fitted by an empirical four-parameter  $p_i$  formula derived for XENON100, which is given by an exponential plus a linear function

$$f(x) = p_0 \cdot \exp(-x/p_1) + p_2 \cdot x + p_3. \quad (6.3)$$

The resulting fits of the upper and lower boundary are shown together with the boundary points derived from the sliced fits in Figure 6.22. The fit of the median, lower and upper boundary was performed using a  $\chi$ -square method. The resulting parameters can be found in Table 6.5. The estimate  $\chi^2/\text{Number of degrees of freedom (ndof)}$  values are 2.9, 3.2, and 1.3 for the lower, median, and upper boundary respectively. The lower boundary and median show a slightly larger reduced  $\chi^2$  pointing to a mismatch between data and the empirical model from XENON100. Judging by eye, the fits of the contour lines show a sufficient agreement with NR distribution. Only at small cS1, deviations between lower boundary and data can be observed which is due to the potential MS leakage explained in Section 6.2.3. A more robust result for the contour lines can be derived from the NR response model fit which is discussed in the following.

Table 6.5. – Fit statistics of the sliced NR contour fits.

Parameter	lower boundary	median	upper boundary
$p_0$	$(0.287 \pm 0.021)$	$(0.500 \pm 0.013)$	$(0.716 \pm 0.022)$
$p_1$ [pe]	$(31.0 \pm 8.0)$	$(23.2 \pm 1.6)$	$(20.2 \pm 1.4)$
$p_2$ [1/pe]	$(-1.79 \pm 0.16) 10^{-3}$	$(-2.32 \pm 0.09) 10^{-3}$	$(-2.02 \pm 0.06) 10^{-3}$
$p_3$	$(1.46 \pm 0.03)$	$(1.602 \pm 0.010)$	$(1.749 \pm 0.014)$
$\chi^2/\text{ndof}$	$(118.31/41)$	$(130.12/41)$	$(55.31/41)$

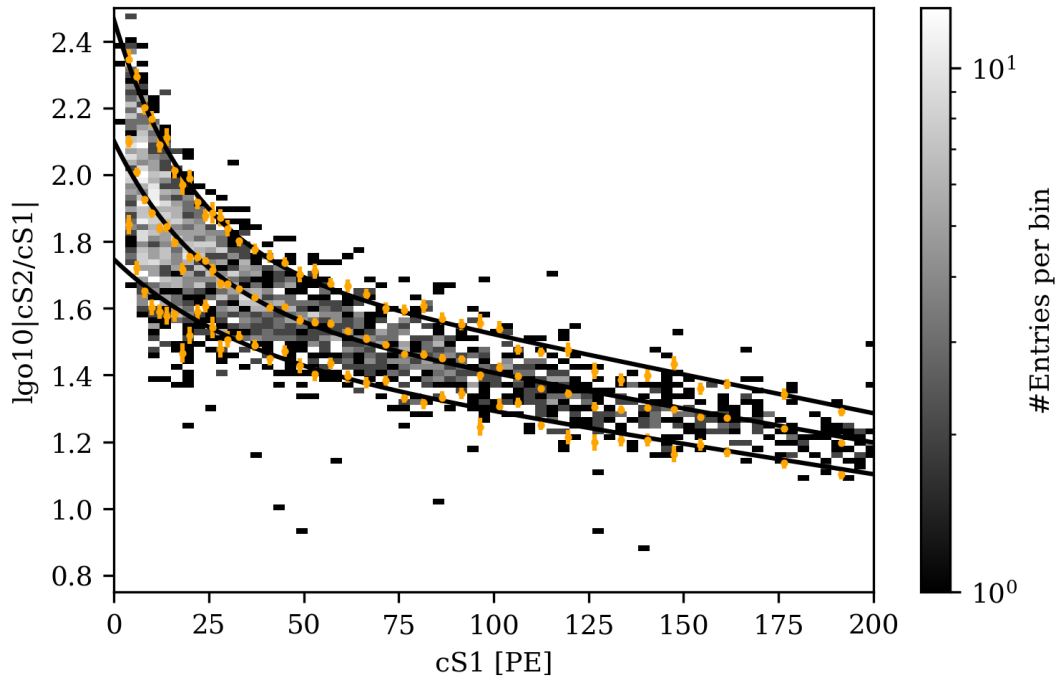


Figure 6.22. – Fit of the upper and lower boundary of 90 % contour for the NR calibration event. The data of the NR band is shown as binned histogram. The median and upper/lower boundary values are shown as orange points including error bars.

### 6.3.2 NR response model

The NR response model fit via BBF was performed by Luisa Höttsch of the Max Planck Institut für Kernphysik in Heidelberg. In the following paragraphs, the fitting procedure and the final NR response model are summarized. For a more detailed discussion, the reader is referred to Luisa Höttsch’s Ph.D. thesis.

BBF was introduced in XENON1T [Apr+19b] as a general tool kit to fit the ER and NR response which is given by a micro-physics model of LXe discussed in Section 2.1. The fitting is based on a Markov chain Monte Carlo (MCMC) method [JJ10] which samples parameters of the micro-physics and detector response model from a multi-dimensional (posterior) probability distribution. The model is described by various parameters and response maps. Among them are single-valued settings like the gain constants  $g_1$ ,  $g_2$ , the drift velocity or the probability for double photo-electron emission of the PMT photocathodes, spatial correction maps for S1 and S2 signals, and a correction map of charge-insensitive regions with partial charge-loss. Most of the parameters are constrained by prior knowledge from calibrations, e.g. from uniformly diluted calibration gases like  $^{83m}\text{Kr}$  and  $^{37}\text{Ar}$ . The response of the analysis software is modeled, including the detector threshold inferred from simulations and a data-driven bootstrapping method using  $^{37}\text{Ar}$  calibration data [Apr+19c], reconstruction bias, and signal broadening introduced by software effects as a function of S1 and S2



size, as well as the cut acceptance derived in the previous section. The LXe micro-physics model is partially constrained by other external studies, for example, NEST.

It should be noted, that given the extremely low electric drift field in XENONnT, a few modifications were made to the exciton-ion ratio in equation (2.4) as well as the electron-ion recombination expressed in (2.5). The field dependencies for  $\zeta$  and  $\delta$  in the NEST model are removed and both parameters are kept constant at their best-fit values described in [Len+15]. The two prefactors  $\alpha$  and  $\gamma$  are treated as free parameters instead. This decision was chosen as it is unclear if the field dependence of  $\zeta$  and  $\delta$  at higher fields can be extrapolated down to 23 V/cm.

The sampling of the model parameters is performed by 200 so-called "walkers". Each walker has its own unique set of values e.g.  $W$ ,  $g_1$ ,  $g_2$ , etc., which are used in a fast toy MC to simulate the full NR band for a given set of interactions. The interactions are bootstrapped from a Geant4 neutron simulation, which provides information about the amount of deposited energy and the position for each simulated interaction with a simulated neutron event. To model the acceptance of the cuts acting against multi-scatter events, and to model the remaining multi-scatter leakage in simulation as found for the NR calibration data in Section 6.2.2, all simulated Geant4 events were processed using the same full-waveform simulation technique as discussed in Section 6.2.2. Only Geant4 events that pass the multi-scatter cuts are provided as input for the fast toy MC. In addition, based on the output of the full-waveform simulation, each energy deposit with an event is backtracked and flagged, whether the S2 signals are identified by the straxen processor as separate S2s, or as a single one. In this way the fast toy MC can simulate merged multi-scatter S2 signals without an expensive waveform simulation.

The toy MC is done in two steps. First, the light and charge quanta are generated based on the LXe response model for the given set of parameters and energy deposits. In the second step, the detector response is simulated yielding cS1 and cS2 values. Subsequently, the simulated NR band is compared to calibration data in a binned Poisson Likelihood in the cS1 cS2 space. If a new set of parameters improves the matching between simulation and calibration the walker accepts the current parameters and updates the posterior distribution. In this way, the walkers "walk" in the direction (converge) of the best fit for the parameters.

An acceptance procedure was defined before performing the final fit since fitting data with an MCMC approach does not provide a method to decide whether a fit has successfully converged. Two different goodness-of-fit measures were chosen. In the first test, the p-value for the entire band was determined by a Poisson  $\chi$ -square test with 6 and 9 equiprobable binned<sup>15</sup> slices in cS1 and  $\log |cS2|$  respectively, where the slices in cS1 have been chosen first. It was required that a p-value of at least 0.1 is reached. The result of this test is shown in Figure 6.23.

In the second test, the same approach was chosen, but only along  $\log |cS2|$  for each equiprobable binned cS1 slice shown in Figure 6.23. All tests were required to yield a p-value larger than 0.1 to accept the final fit of the NR response model.

<sup>15</sup>An equiprobable binning means that a given parameter space is divided for a given number of bins such, that each bin contains the same amount of data points. In this way, p-value "tuning" based on a variation of the binning can be prevented.

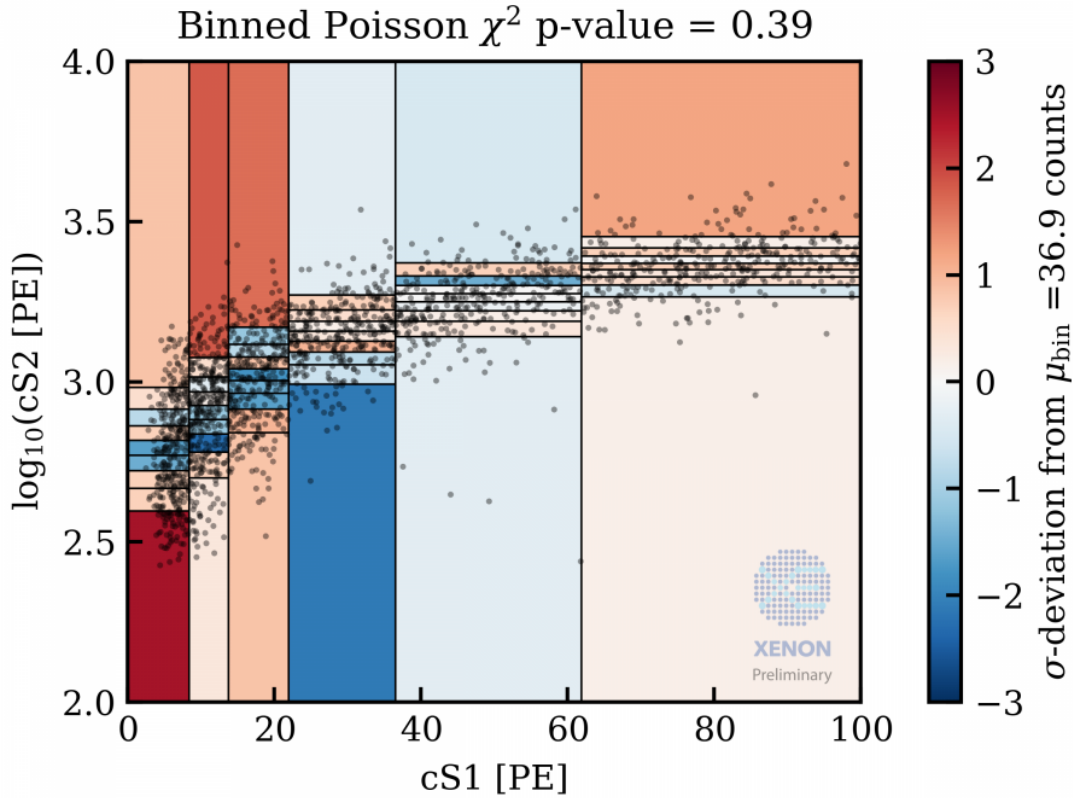


Figure 6.23. – *P*-value evaluation of the best fit of the NR calibration data using the BBF fitting framework. The plot shows the AmBe calibration data as black dots on top of an equiprobable binned histogram which shows the deviation of numbers of entries per bin between data and simulation. The color scale shows the deviation in units of sigma of data from simulation of the best fit. (Credits Luisa Höttsch for the XENON collaboration)

Figure 6.23 shows a p-value of 0.393 for the complete NR band fit. Its color scale can be understood as two-dimensional fit residuals which are showing mostly a random fluctuation around the mean. Only the patch of the lowest bin in cS1 and cS2 shows a larger deviation of about  $2.5\sigma$ . A similar result was found for the sliced fits. Several different studies were conducted as part of this thesis and by other analysts to check the cause of this discrepancy between the model and data. Some of these tests were already discussed in the previous section. They showed that contamination from additional backgrounds like ACs or from badly reconstructed events seems to be unlikely. In addition, the impact of this mismatch on the WIMP inference was tested and turned out to be negligible.

One explanation for this mismatch between the response model and data could be a slightly wrong modeling of the electron-ion recombination in the microphysics of LXe. Usually, these models are studied at much larger drift fields compared to the field reached in XENONnT. Together with the very strong background suppression of the NV TPC coincidence, this effect might not have been observed in previous studies. A second, explanation for this excess is either remnant multi-scatter signals or

mismodeling of multi-scatter events using GEANT4. Multi-scatter events which produce a second alternative S2 below the cutting threshold of the S2 single-scatter cut, will have an anomalously large S1 compared to its main S2. This has two implications. First of all the NR band will be broader at lower energies than expected from pure single-scatter events. Secondly, this effect leads to an overdensity at the detector threshold, as stacked multi-scatter S1 signals can overcome the detector threshold more easily. This hypothesis aligns perfectly with the observation in Section 6.2.3 that the outlier events below the NR band follow the general distribution of good NR events and are neither distributed like ACs nor charge-loss events. However, this effect should be already taken into account by BBF when merging all S1 signals of a single simulated event.

Another possible explanation might be poor modeling of the AmBe source used in the simulation, due to a lack of knowledge of the source parameters at that given time. During SR0 the source was modeled by emitting neutrons according to the neutron-energy spectrum shown in Figure 6.2, including neutron energies even when  $^{12}\text{C}$  ends up in its ground state  $n_0$ . However, given that the NV TPC coincidence only selects events for which  $^{12}\text{C}$  ends up in its first or second excited state, this behavior is clearly different from the calibration. A test in which only simulated events with a primary neutron energy of less than 6 MeV were selected, did not show any significant difference. However, a more complete model of the AmBe source is currently being implemented and a follow-up study will be conducted during SR1.

In principle an additional background component like ACs could explain the difference in data compared to simulations. However, given the strong background suppression due to the detector coincidence, this assumption can be ruled out.

The best fit of the NR calibration data is shown in Figure 6.24. The best fit was reached after about 100k iterations. The posterior value for each parameter is derived as a median from the corresponding distributions which are given by the results of the individual 200 walkers for the last 500 iterations of the fitting. The uncertainties are estimated based on the 15.865 and 84.135 percentile of the corresponding distribution. As an example, the posterior distribution of the proportionality constant between the electronic stopping power and the velocity  $\kappa$  as well as the parameter  $\beta$ , which describes the exponential energy dependents on the exciton-to-ion ratio, is shown in Figure B.37 in the appendix on page 263.

Figure 6.24 shows a good agreement between data and the fitted NR response model. The response model is used to create the templates for neutron and CE $\nu$ NS background models, as well as the WIMP source model, which is created by forward folding the energy spectra of different WIMP masses, shown in Figure 1.3, into the cS1/cS2 space.

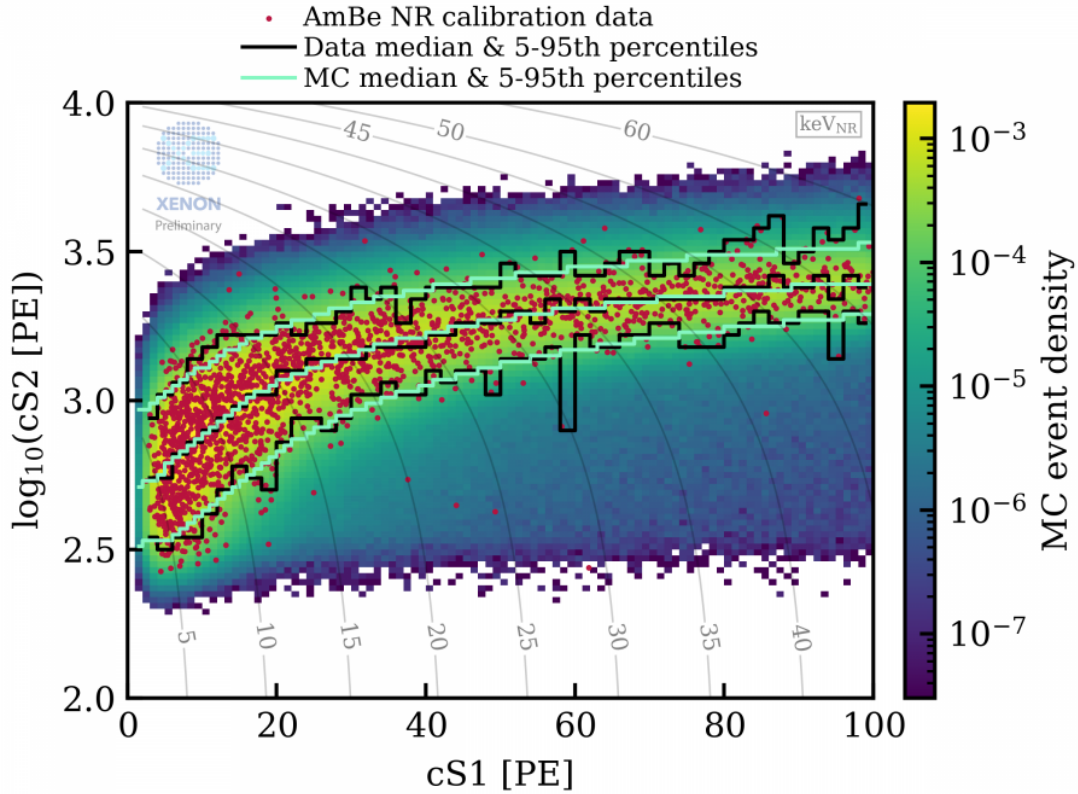


Figure 6.24. – Best fit of the NR calibration data using the BBF fitting framework. The plot shows the NR calibration data as red points on top of the band distribution derived from the simulated data using the posterior values of the best fit. The black and cyan line are showing the 5% and 95% contour as well as the band median for the data and ToyMC data respectively. The gray vertical lines indicate the slicing in  $cS1$  of the equiprobable binning used for the  $p$ -value evaluation. (Credits Luisa Höttsch for the XENON collaboration)

#### 6.4 CONCLUSION AND OUTLOOK

In this chapter, the calibration of the nuclear recoil response model for SR0 of XENONnT is discussed. The NR calibration is based on a novel coincidence technique between TPC and NV which was conducted for the very time during this work. The coincidence uses the 4.44 MeV  $\gamma$ -rays of an AmBe neutron source, which are coincidentally emitted in addition to the neutrons, to tag S1 NR signals in the TPC. The calibration achieved an excellent background suppression with an expected number of accidental coincidences of only  $(1.75 \pm 0.55)$  events, which is negligible compared to the total number of events of 2537 in the NR ROI ( $<0.1\%$ ). In addition, the timing resolution of this detector coincidence was studied and was found to be  $(14.03 \pm 0.23)$  ns which is comparable to the timing resolution of the TPC digitizer of 10 ns. Furthermore, the calibration data was used to determine the total acceptance of XENONnT data selection. Both calibration data and acceptance were provided to determine the XENONnT NR response model which is a key component for the WIMP inference. The resulting

response model shows good agreement with the data, except for the lowest energy bin. A follow-up study showed that this small mismatch has a negligible impact on the final WIMP results [Apr+23b]. In addition, the NR calibration data is used as input for the NV tagging efficiency study.

Although the detector coincidence already shows a very strong suppression of backgrounds with a very small number of accidental coincidences, the selection of the 4.44 MeV  $\gamma$ -ray inside the NV can be further fine-tuned. For example, the same selection of the spatial signal distribution as discussed in Chapter 7 can be added. This would help suppress the contribution from detector backgrounds even further and, in turn, may allow the loosening of some of the other selection requirements, thus increasing the overall acceptance of the  $\gamma$ -ray selection.

Given the uncertainty of whether the broadening of the NR band at low energies is either due to a wrong field dependence of the LXe micro-physics model or due to pollution with multi-scatter events, the multi- to single-scatter discrimination needs to be improved for SR1. One of the reasons for the poor discrimination power is the low drift field, which lowers the z-resolution of the detector, as well as the low extraction field which leads to a higher single electron background. An obvious improvement could be achieved by new electrodes which would allow for better fields. This option is currently investigated within the collaboration, and the XYZCamScan wire inspection set-up discussed in the appendix in Section A.1 plays an important role here. However, such an intervention is very challenging and time-consuming. Hence, it might be more feasible to increase the discrimination power on the analysis side. One action which should be investigated in SR1 is an improvement of the S2 single-scatter cut by using more information from different parameters beyond just the main and alternative S2 size. In addition, an improved GEANT4 AmBe source model, together with a second AmBe calibration at a deeper z-positions and closer to the TPC cryostat, will help investigate this issue further.

Achieving this better understanding will also open up new opportunities for the NR AmBe calibration. Given the strong background and ACs suppression due to the tight detector coincidence, the AmBe data can be used to calibrate the NR response all the way down to a single photon S1 response which was never done before. This should be feasible as the total lone hit rate<sup>16</sup> is  $\sim 4$  kHz, and therefore the chance of finding accidental coincidences in the NV-TPC coincidence using a 400 ns window is less than 0.2 %. The coincidence even allows to calibrate the NR response beyond single photons, by replacing the timestamp of too faint S1 signals with the timestamp of the NV 4.44 MeV  $\gamma$ -ray event. In this way, one of the most important information carriers, the drift time, can be recovered and cS2 can be computed. Further, it provides a tool to study the S2-only response of the detector more deeply, improving the physics reach of all LXe TPC experiments for low WIMP masses. A first feasibility study of this approach was already conducted at the end of this work and is briefly discussed in Section A.5 of the supplemental material on page 235. Last but not least this new cS2-only space, paired with the knowledge that an event must be a NR interaction due to the NV tagging, may also enable a study of the so-called Migdal effect [Bel+22] which is theoretically predicted for LXe but was never directly measured. First stud-

<sup>16</sup>The rate of single PMT signals which are not clustered into any peaks.

ies, testing the feasibility of this type of calibration, are currently conducted including also analysts from other institutes.



---

## CALIBRATION OF THE NEUTRON VETO

As shown in the previous chapter, neutrons are an excellent tool to study the detector response for WIMP DM as both produce elastic nuclear recoils when interacting with xenon. For this very same reason, they also pose one of the biggest challenges for the science goals of XENONnT. Detector materials can emit neutrons via spontaneous fission or  $(\alpha, n)$ -reactions which enter the fiducial volume of the TPC producing WIMP-like signals. A certain fraction of these interactions can be distinguished from WIMPs as neutrons perform multiple scatters. But as seen in the previous chapter, identifying multi-scatter events becomes difficult at small energy deposits. Further, given the mass difference of xenon compared to neutrons, neutrons can undergo a single back-scatter and leave the sensitive detector region. The task of the neutron veto is to tag these neutrons by measuring the 2.22 MeV  $\gamma$ -ray released by the neutron capture on hydrogen and thus mitigate this danger. In order to characterize this "tagging efficiency" the following chapter discusses the calibration of the neutron veto neutron response.

The chapter starts with a general analysis of the AmBe energy spectrum measured by the NV in Section 7.1, which is compared against spectra made with a custom toy MC framework in Section 7.2.2. This analysis is followed by a discussion of the NV tagging efficiency in Section 7.3, which is the most important quantity of the NV. Section 7.4 introduces and characterizes the neutron veto detection efficiency. Compared to the tagging efficiency, the detection efficiency only quantifies the chance of finding a neutron after being emitted. It is an important benchmark to compare the performance of the NV with other water Cerenkov detectors like SNO or SK. The last Section 7.5 of this chapter uses a NV self-coincidence between the 2.22 MeV neutron capture and 4.44 MeV  $\gamma$ -ray, to study the NV response in more detail. In addition, the section attempts to determine the thermal neutron capture cross section on hydrogen. The closing Section 7.6 gives a short conclusion and outlook.

### 7.1 EMPIRICAL MODEL OF THE AMBE ENERGY SPECTRUM:

Before the tagging efficiency of the NV can be studied the key characteristics of the AmBe energy spectrum must be understood. Especially, the 2.22 MeV neutron capture



line must be identified. Already in the previous chapter the AmBe spectrum recorded by the NV was briefly shown in Figure 6.5. The goal of this section is to describe the AmBe spectrum by an analytical fit model and to determine the overall source rate as well as the branching ratio of the final states of its decay product  $^{12}\text{C}$ . Further, the results in this section were used to determine the selection boundaries for the 4.44 MeV  $\gamma$ -ray in Chapter 6 as well as for the last section of this Chapter.

Before a response model for the AmBe data can be developed and studied a clean calibration sample must be selected. The following selection was used in this analysis:

- At least 4 PMTs must contribute to the event (4-fold coincidence)
- Each event must contain at least 4 pe area
- The center time of the event must be within the 95 % contour derived in Chapter 5.
- The reconstructed azimuthal angle and z-position must be within the corresponding 95 % contours derived from data, as it is shown exemplarily in Figure 7.8.

The effect of the different selections on the calibration data is shown in Figure 7.1 for the data taken at the bottom CW11d2m source position. The average signal-to-noise ratio for all four source positions after selections improves by a factor of about 5 from 2.5 to 12.4.

Different models were fitted to the data throughout this work. Given the very large calibration statistics, a  $\chi^2$ -method was chosen as the optimization procedure. The fit model which describes the background subtracted spectrum of Figure 7.1 best is given by

$$f(x) = F(x, \mu_{\text{th}}, \sigma_{\text{th}}, x_0) \left( N(x, \mu_1, \sigma_1) + P_2(\mu_2, \sigma_2, a, b, c) + H(x, \mu_3, \sigma_3, \mu_4, \sigma_4) + c \right) \quad (7.1)$$

where  $N(x, \mu_1, \sigma_1)$  represents a normal distributions describing the 2.22 MeV neutron capture peak,

$$P_2(\mu_2, \sigma_2, a = 0, b = 0, c) = N(x, \mu_2, \sigma_2) + C(x, a, b, c, \mu_2, \sigma_2) \quad (7.2)$$

representing the 4.44 MeV  $\gamma$ -line as normal distribution plus an additional Compton edge term  $C$  which was modeled according to [SDA]. The location of the Compton edge is constrained by the Compton formula and the fitted position  $\mu_2$  for the 4.44 MeV line. The high energy tail of the distribution is modeled by  $H(x, \mu_3, \sigma_3, \mu_4, \sigma_4)$  which consists of two additional normal distributions. These additional peaks are empirically motivated by the spectrum itself. The Compton continuums for these two peaks have not been included as the large degeneracy between the different components of the fit and the missing constraints on the location of the Compton edge made a successful implementation not feasible. Given that the exact origin of the high energy tail is not fully understood an additional constraint is put on the correlation between the mean and variance of the normal distributions modeling the individual peaks. It

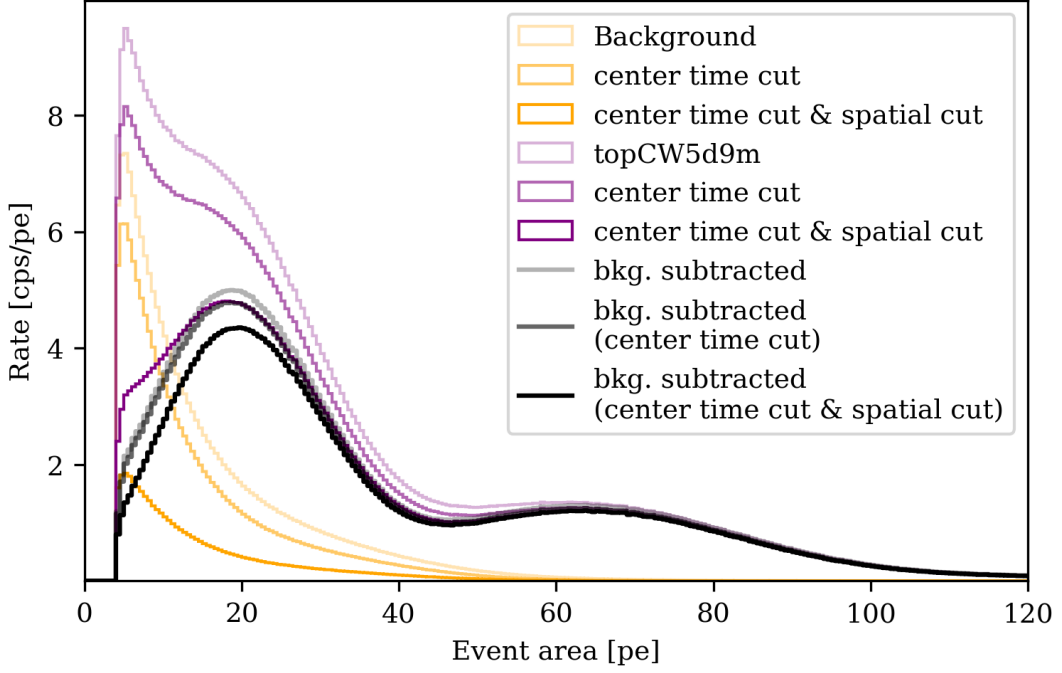


Figure 7.1. – Energy spectrum of the AmBe calibration taken at the bottom CW11d2m source position. The orange and purple distribution show background and calibration data respectively. The different intensities indicate different selections as indicated by the figure legend. The black histogram shows the bin-wise background subtracted representation of the calibration data.

is expected that the width of the peaks is determined by Poisson statistics and that the width can be modeled by

$$\sigma_i^2 = s \cdot \mu_i + s_0 \quad (7.3)$$

where  $s$  and  $s_0$  are two scaling parameters representing the detector response as well as a threshold for the Cerenkov energy light production. An additional constant background term  $c$  is added to the model as an additional degree of freedom.

The selection efficiency of the data quality cuts and thresholds are modeled by a Fermi-Dirac-like function of the form

$$F(x, \mu_{\text{th}}, \sigma_{\text{th}}, x_0) = 1 - \frac{1}{\exp\{(x - \mu_{\text{th}})/\sigma_{\text{th}}\} + 1} \cdot \Theta(x - x_0) \quad (7.4)$$

which is truncated by the Heaviside step function  $\Theta(x - x_0)$  at the position  $x_0 = 4$  pe to take into account the area selection. An additional penalty term in the likelihood ensures that the threshold function evaluated at  $F(x = 0 \text{ pe})$  falls off to zero. The absolute acceptance loss caused by the spatial and center time quality cuts cannot be determined from the analytical fit model itself. To access this cut acceptance an MC study would be required. Thus, the source rate determined by the analytical model is systematically biased towards smaller values which need to be kept in mind.

The best fit for the bottom CW11d2m position is shown in Figure 7.2 in a linear and

logarithmic scaling. The fitting itself was done in multiple steps only freeing a subset of parameters that are fitted over limited fit ranges as many of the parameters are highly correlated. In the last fitting step, all parameters are free and a last combined fit over the entire fitting range is performed. Judging based on the fit residuals, the model shows a reasonable matching with data. The residuals show a slight wobbling structure pointing to a systematic mismodeling or missing component in the fit. The reduced  $\chi^2$  is 6.5 which corresponds to a p-value of zero. This rejects the hypothesis that the current model describes the data completely and also points to a missing component in the fit. However, it should be also noticed that the rather small p-value might also be a problem due to the very high statistics of the data which makes already small deviations between the model and data significant [LLS13].

The fits for the other source positions are attached in the appendix in Figure B.38 on page 265. Their results are similar. The parameter estimates for the best fits are attached in Table B.1. The mismodeling might originate from the fact that the cut acceptance of the data quality cuts is not flat, leading to a bias in the response and the wobbling observed in the residuals. A fit without data quality cuts was performed too, also resulting in a p-value of zero, pointing to an issue with the fit model itself. Finding a good analytical model, however, is very difficult as the very high statistics of the calibration data deem that the fit model must be known up to very high precision. In addition, it is not entirely clear which components and elements contribute to the observed spectrum. The high energy tail is most likely caused by neutron capture on  $^{56}\text{Fe}$  which releases a cascade of  $\gamma$ -rays with a total energy of 7.6 MeV. Such a cascade leads to a broadening of the full energy peak as decay channels with a larger number of gammas emitted must pay a higher toll due to the Cerenkov threshold in water which is about 264 keV. An additional candidate might be the neutron capture on  $^{136}\text{Xe}$  [Alb+16]. Thus a better approach to determine the branching ratio of the  $^{12}\text{C}$  might be given by templates generated from toy MC simulations which should give a more complete picture of the detector response. Further, a template fit via simulations also allows determining the selection efficiency of the quality cuts and therefore the absolute rate of the spectrum. This approach is discussed in the subsequent section.

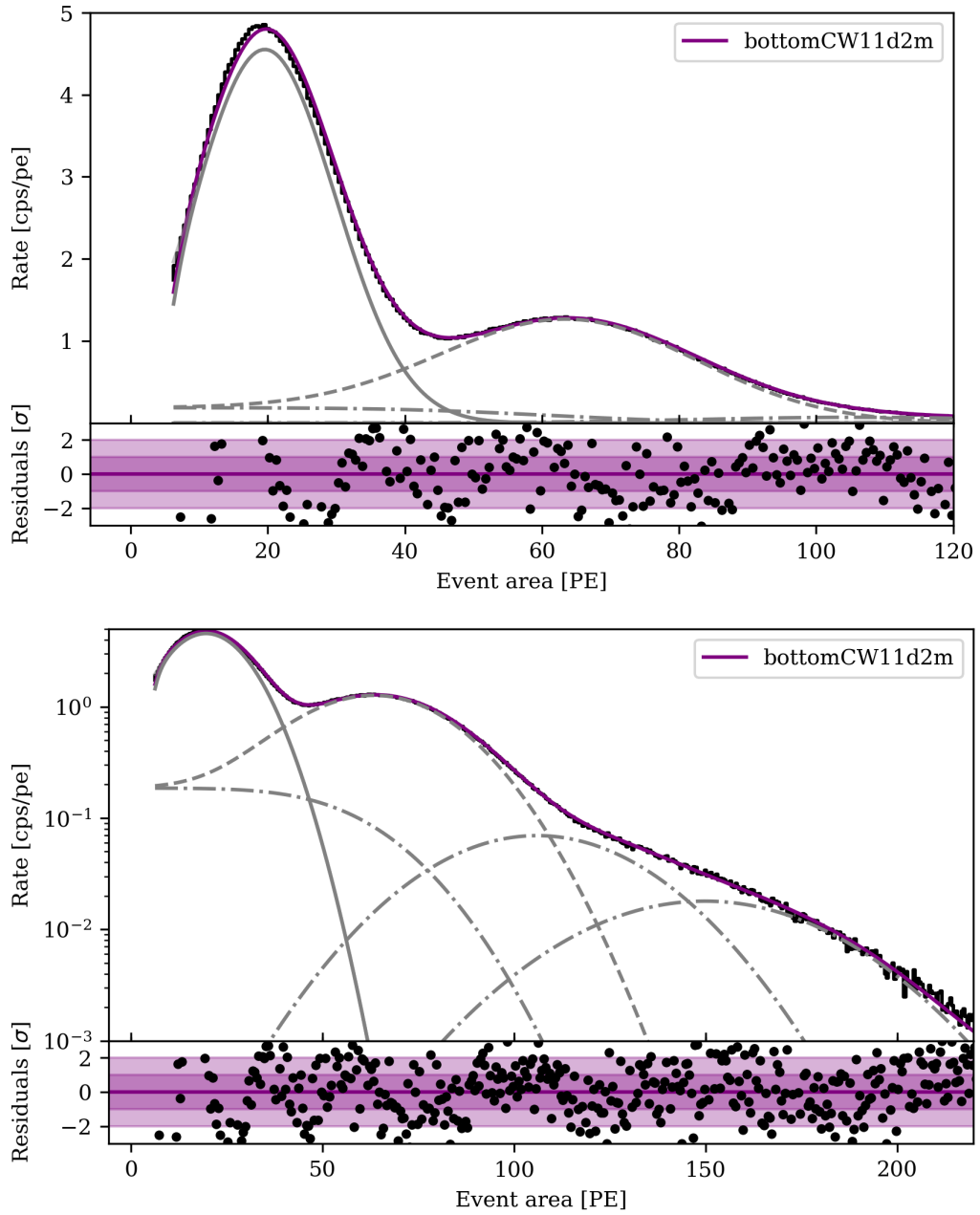


Figure 7.2. – Best fit of the NV AmBe energy spectrum taken at the bottom CW11d2m source position. The upper figure shows the best fit in a linear scale and the lower figure in a logarithmic scaling for a wider range. The bottom panel of both figures shows the fit residuals. The individual fit components are displayed in gray. For the 4.44 MeV line  $P_2$ , as well as the Compton edge term  $C$ , are shown. The threshold function is not displayed. Uncertainties are not shown.

## 7.2 NEUTRON VETO RESPONSE AND AMBE TEMPLATE MODEL USING TOY MCS

### 7.2.1 NV toy MC

As an alternative to the analytical model, toy MCs can be used to fit templates of the individual AmBe source components to data. In this way, the branching ratio of the first excited state of  $^{12}\text{C}$  can be determined. To perform such a template fit an effective toy MC framework was developed during this work. The simulation framework is based on first principles and represents a fast and effective simulator that builds fake NV events based on delta pulse PMT signals. The code of the simulation framework is written in python and depends on a small set of third-party packages like, numpy, numba, uproot, and awkward arrays for numerical and fast operations. The working principle of this fast simulator is sketched in Figure 7.3 and discussed in detail in the following.

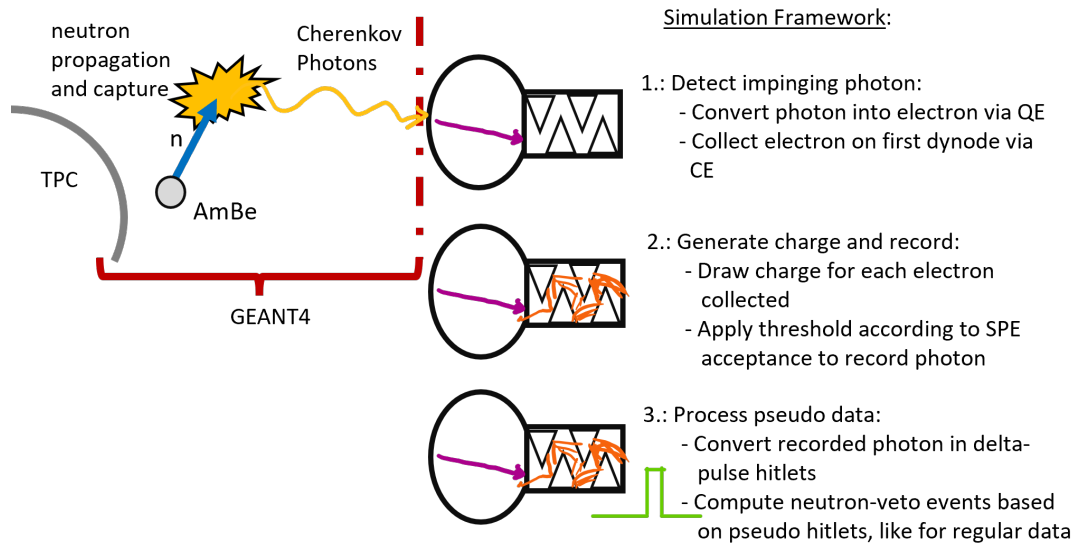
Just like WFSim, the simulator takes as inputs “events” which are generated by Geant4 simulations. Geant4 is used to simulate for a given particle transportation, energy deposition, and energy conversion; e.g. for a neutron the particle propagation inside the water of the NV, the subsequent capture on hydrogen, the release of a 2.22 MeV  $\gamma$ -ray, the emission of Compton electrons via Compton scattering, and the emission of Cerenkov radiation until the impinging of single Cerenkov photons onto the photo-cathodes of the NV PMTs. For each Cerenkov photon, the time of the impact on the photo-cathode, PMT channel, and photon wavelength is stored in a Root TTree.

The effective simulator reads the impinging “PMT hits” from the corresponding TTrees via uproot which provides the information as awkward arrays similarly as it is done in EPIX. To convert the Geant4 photon-hit information into straxen events, these pseudo-hits must undergo the same transformation steps as real photons. As a first step, all impinging photon hits are converted into “detected” photons. A PMT detects photons by converting them into photo-electrons at the photocathode of the sensor. The probability of releasing such a photo-electron is wavelength dependent and is given by the QE of the PMT. Subsequently, the photo-electron needs to be collected on the first dynode of the amplification stage. The chance of collecting photo-electrons depends on the PMT geometry and field and is given by the CE [K07]. The detection probability of a single photon  $p_{det}$  is then given by binomial statistics represented by

$$p_{det}(\lambda) \sim \text{Binom}(n = 1, p = QE_i(\lambda) \cdot CE). \quad (7.5)$$

where  $i$  represents the  $i^{\text{th}}$  PMT channel.

The QE for all NV PMTs was measured by Hamamatsu and the simulation framework uses channel dependent QEs. Figure 7.4 shows a comparison of the photon wavelengths before and after applying the QE. The QE reaches its maximum around 350 nm, however, one should keep in mind that the number of emitted Cerenkov photons per unit wavelength scales with  $1/\lambda^2$  which shift the region of maximal sensitivity slightly below 350 nm as shown by the orange distribution. The simulated photon spectra shown in black and red have a few additional features which are due to the



**Figure 7.3.** – Schematic of the effective NV simulation framework. The simulation is divided into two parts. First particle and photon propagation until the impinging of single photons on the NV PMTs is simulated via Geant4 (left). Afterwards the effective simulator discussed in the text builds NV delta PMT pulses based on the three steps indicated in the schematic.

photon absorption in water.

The CE of the individual PMTs is unknown and is therefore implemented as a free scaling parameter which can be determined via the data MC fitting discussed later in this section. In general, also the CE is PMT dependent, but since it is not feasible to have a free scaling parameter for each PMT only a single averaged value is used.

After converting impinging photons into detected photons, the created charge of each photon must be calculated. The charge creation in a PMT depends on its amplification stage. A model for the charge response of single photo-electrons can be inferred from LED calibration data discussed in Chapter 5. The calibration data is described by a four-component fit model representing the pedestal, under amplified PMT signals, the single photo-electron, and the double photo-electron response. To mimic the charge response of a single photon only the under-amplified and SPE response model are required.

The PMT charge response model used in the fast simulator is based on the best fits for one of the LED calibration campaigns provided by Andrea Mancuso from Bologna University. For each PMT both, under amplified and SPE response model, are normalized into a proper probability density function which returns by definition as average value 1 pe. Charge values are sampled from these distributions and only photon-electrons whose charge exceeds the SPE threshold equivalent charge are transformed into pseudo hits. The SPE threshold of each channel is deduced from the SPE acceptance  $\alpha_{\text{SPE}}$  calibrated during SR0. The threshold is given by the charge for which the cumulative of the charge response model exceeds  $1 - \alpha_{\text{SPE}}$ . Two example SPE distributions for the channels 2000 and 2001 are shown in Figure 7.5. PMTs that were

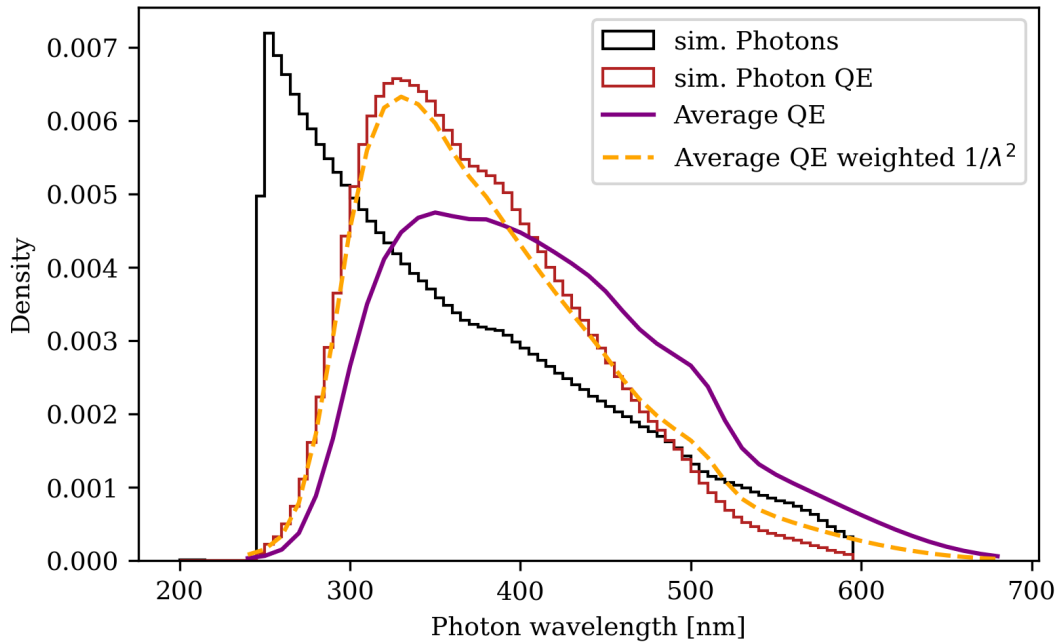


Figure 7.4. – Average PMT quantum efficiency of the NV PMTs. The purple distribution shows the average quantum efficiency as a function of wavelength. The orange distribution shows the QE after weighing it with  $1/\lambda^2$  to account for the number of Cerenkov photons emitted per unit of wavelength. For plotting purposes the amplitude of the orange distribution was scaled. The black distribution shows a simulated Cerenkov photon spectrum produced by 4.44 MeV  $\gamma$ -rays. The red distribution shows the resulting spectrum after applying the PMT dependent quantum efficiency. Both spectra were normalized.

excluded during the science run are also excluded in the toy MC.

In the last step, all recorded hits are converted into straxen’s NV hitlets. Each hitlet is represented as a simple delta pulse with a single sample length, carrying only time, PMT channel, and charge information. Since the timing of the photons is relative to each simulated Geant4 event, each group of hitlets is delayed with an additional time offset which can be either distributed equidistantly to simulate well-separated events or distributed according to a given source rate to mimic pile-up. Subsequently, the generated pseudo hitlets are passed along the regular NV plugins to compute events and their properties. Different other information reflecting the “MC truth” of a simulated event like the deposited energy or interaction type can be propagated, too, depending on the user’s decision.

The performance of the simulator is reasonably fast. The simulation time is dominated by I/O operations when reading the Geant4 input files which take about 50% of the total computing time. The pure simulation time of pseudo hitlets takes about 60 s for 100k 4.44 MeV gammas which correspond to about 30 million Cerenkov photons. This is negligible compared to the simulation time of 100 k Geant4 events which takes about 5 h.



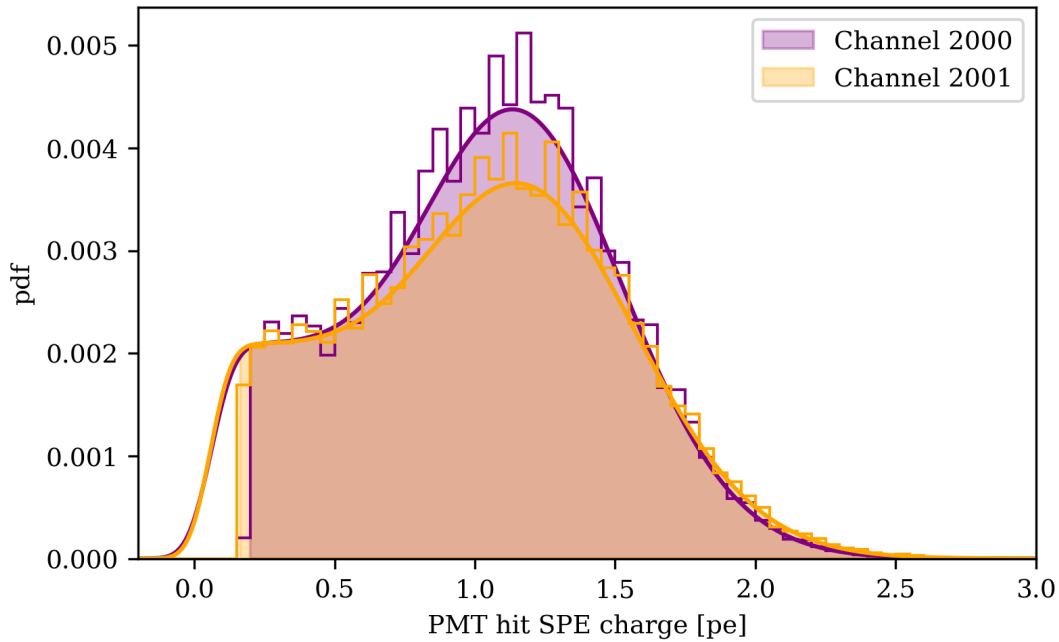


Figure 7.5. – Example probability density distributions of the single photo-electron response for the NV PMTs. The purple curve shows the SPE response of channel 2000 while the orange model shows channel 2001. The shaded region shows the fraction of the distribution which is above threshold. The histograms show the distribution of a few sampled pseudo hitlets for each given distribution.

### 7.2.2 Template modeling and $^{12}\text{C}$ branching ratio

To generate templates for the individual components of the AmBe source 600 k AmBe neutrons, 4.44 MeV  $\gamma$ -rays, and 3.22 MeV electrons were simulated for each of the four calibration source positions used during SR0. The Geant4 model used in the simulation included the full geometry of the experiment, including the u-tubes as well as a simple steel cylinder to approximate the shielding of the source casing for the 3.22 MeV electrons. A rendering of the full simulation configuration is shown in the appendix in Figure B.42 on page 268. As starting value, the CE was assumed to be 80%. The second excited state of AmBe which emits both a 4.44 MeV gamma and a 3.22 MeV electron was mimicked by bootstrapping for each electron event a corresponding 4.44 MeV gamma-ray event. Figure 7.6 shows a comparison of the simulated NV response for the second excited state of  $^{12}\text{C}$  when simulating electrons with and without source casing. The source casing leads to a strong reduction of the deposited energy for the 3.22 MeV electrons. The response of the first and second excited state are almost identical. Only in the low and high area tails the two distributions can be distinguished. The simulation geometry was implemented and provided by Pavel Kavrigin from the Weizman Institute of Science.

Before fitting the templates to the spectra derived from data, a comparison between simulation and data for the most relevant other parameter space was performed. Fig-

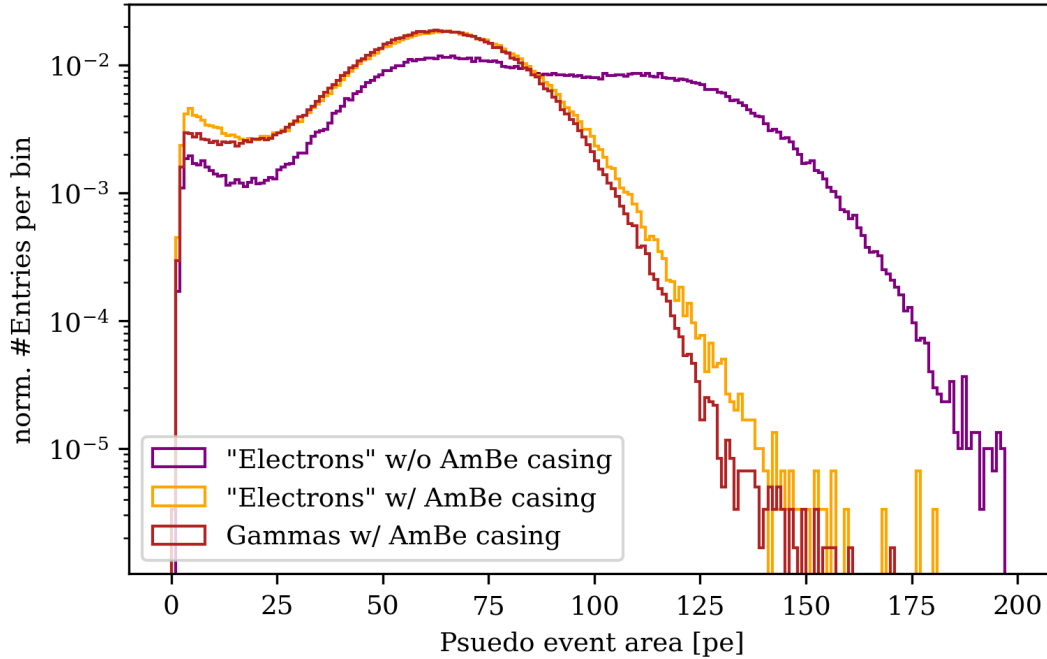


Figure 7.6. – Comparison of the NV response for first and second excited state of the  $^{12}\text{C}$  with and without source casing. “Electrons” refers to the second excited state which emits an additional 3.22 MeV electron with respect to the first excited state, labeled as “Gammas”.

Figure 7.7 shows a comparison between data and simulation for the center time as a function of event area. Both distributions for the simulated neutrons and gammas show a large overlap with the joint distribution from the data. However, simulation also shows a slight systematic bias towards smaller center times. This systematic shift indicates that the current optical model used in the simulation does not match the true optical conditions of the NV. A study, currently conducted by Pavel Kavrigin, compares the center time in simulation and data more systematically. It shows, that especially a change in the reflectivity of the ePTFE walls affects the center time. A few plots of this study are shown in the appendix in Figure B.43 on page 269. A second point that should be noticed is, that the current simulation approach does not include any PMT timing effects like transient time spread (TTS) or digitizer sampling which adds a bit of extra smearing to the photon-timing.

The second important data space is the spatial distribution of the events. Figure 7.8 compares the spatial distribution from data with simulated neutrons and gammas. Data and simulation show a good agreement in this parameter space. The distribution for the neutron capture signals is a bit wider as for the 4.44 MeV  $\gamma$ -rays. Based on the free mean path for neutrons and  $\gamma$ -rays shown in Figure 2.11 and Figure 2.12, the most likely explanation for this observation is, that the  $\gamma$ -rays deposit their energy closer to the source position. Both, MeV neutrons and  $\gamma$ -rays, have a free mean path of  $\sim 10$  cm in water. However, neutrons have to be captured first before the subsequently emitted  $\gamma$ -ray can produce Compton scattering and therefore Cherenkov light, while

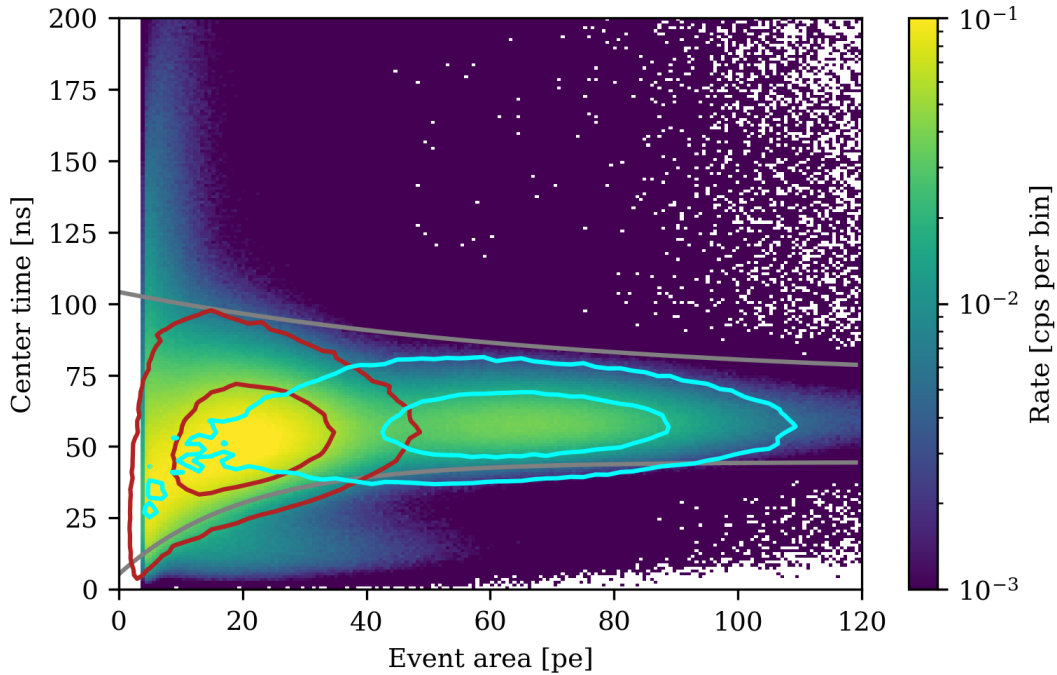


Figure 7.7. – Comparison of the center time distribution between data and simulations for the data taken at the bottom CW11d2m source position. The histogram shows a binned representation of the data. The red and cyan contours are indicating the 50% and 90% percentile of the simulated data distribution for the neutron capture and 4.44 MeV  $\gamma$ -line respectively. The contours were estimated using a highest density region method on a binned representation of the simulated data. The gray lines indicate the same contour as discussed in Section 5.2

the  $\gamma$ -rays emitted by AmBe do not have to undergo this intermediate step.

Given that the center time in simulation slightly deviates from the data, the center time cut is dropped during the template fitting and only the following selections are applied to both data and simulation:

- 4-fold coincidence requirement
- At least 4 pe event area
- The reconstructed azimuthal angle and z-position must be within the 95% contours of the angle and z-position derived from data.

Before fitting, calibration and simulation data are converted into a binned representation of the underlying events. For simulation, the event area is scaled with a free parameter representing an effective collection efficiency  $CE_{\text{eff}}$  before binning. This parameter is needed to account for the fact that the initial choice of the unknown collection efficiency might be wrong and thus needs to be scaled. To avoid a rebinning of the templates during fitting, the fitting itself is performed multiple times scanning over a range of different  $CE_{\text{eff}}$  values between 0.5 and 1.5 in a stepping of 0.01. As in a regular fit, the best fit of this scan is determined by the fit which results in the

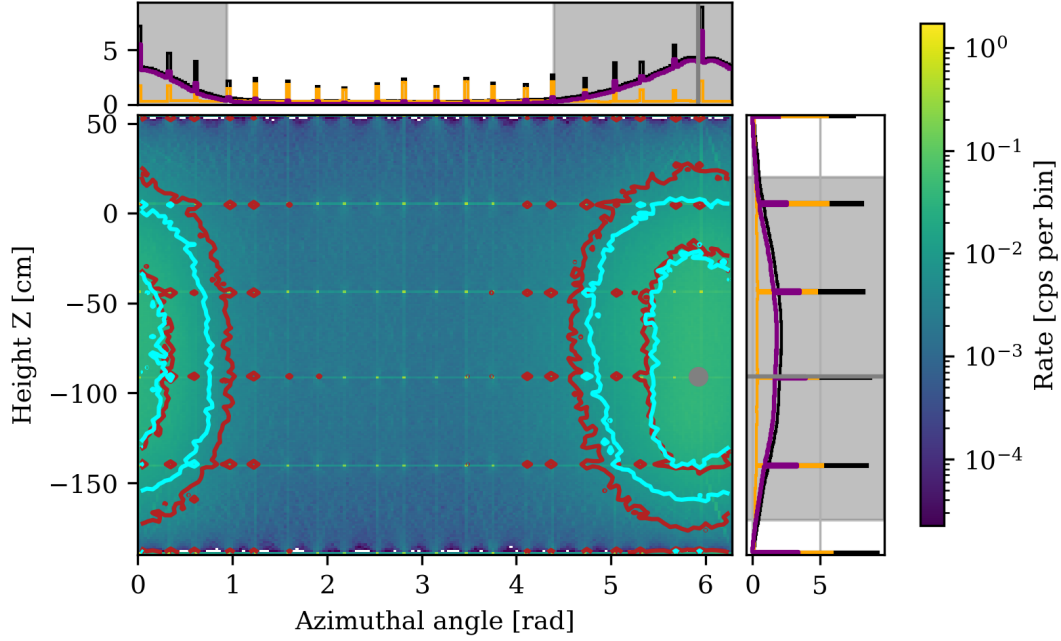


Figure 7.8. – Comparison of the spatial distribution between data and simulations for the data taken at the bottom CW11d2m source position. The main panel shows the calibration and simulation data as described in Figure 7.7. The projections show the angular and  $z$  distribution of the binned data. Black indicates the AmBe data, orange background data and purple the background subtracted AmBe data. The gray shaded regions indicate the central 95% percentile distribution of the given projected axis. The gray dot and lines indicate the position of the calibration source.

lowest  $\chi^2$  between template and data. If the best fit of  $CE_{\text{eff}}$  deviates more than 5% from unity the collection efficiency is updated and the toy data is simulated again.

The template fitting itself is done in the event area space, using a least squares method. Given that the statistical uncertainty of the templates dominates over the statistical uncertainty of the data, both uncertainties were combined.<sup>1</sup> The statistical uncertainty of the calibration data  $\sigma_{\text{data}}$  is combined with the statistical uncertainty of the simulation templates  $\sigma_{\text{sim}}$  such that the total uncertainty  $\sigma_{\text{tot}}$  is given by

$$\sigma_{\text{tot},i} = \sqrt{\sigma_{\text{data},i}^2 + \sigma_{\text{sim},i}^2} \quad (7.6)$$

where  $i$  represents the  $i$ th bin. In both cases, it is assumed that the statistical uncertainties are given by Poisson statistics.

Given that in this approach the magnitude of  $\sigma_{\text{sim}}$  depends on the best-fit parame-

<sup>1</sup>A second valid option would be to increase the statistics of the simulated templates until their uncertainties become negligible. However, this approach has not been chosen as the statistics in data are very high and simulating an adequate number of events would put very high demands on storage capacity.

ters itself, the default least squares likelihood has to be rewritten as

$$L = \prod_i \frac{1}{\sqrt{2\pi}\sigma_{\text{tot},i}(\vec{\theta})} \exp \left\{ -\frac{(y_i - Y_i(\vec{\theta}))^2}{2\sigma_{\text{tot},i}(\vec{\theta})^2} \right\} \quad (7.7)$$

which simplifies to

$$\log L = -\frac{1}{2} \sum_i \frac{(y_i - Y_i(\vec{\theta}))^2}{\sigma_{\text{tot},i}(\vec{\theta})^2} + 2 \log |\sigma_{\text{tot},i}(\vec{\theta})| \quad (7.8)$$

where  $y_i$  and  $Y_i$  represent the rate for the  $i$ th bin in data and simulation respectively<sup>2</sup>. All constant terms in equation (7.8), have been dropped.  $\vec{\theta}$  represent the fitting parameter of the summed template which is given by

$$Y_i(\vec{\theta} = \{R, \epsilon_A, n_1, n_2\}) = R \cdot \epsilon_A (m_i^n + n_1 \cdot m_i^\gamma + n_2 \cdot m_i^e) / N \quad (7.9)$$

in which  $R$  represents the total neutron rate of the source,  $\epsilon_A$  a free scaling parameter modeling the selection efficiency,  $m_i^n, m_i^\gamma, m_i^e$  the number of neutron, gammas and gammas + electron signals found in the  $i$ th bin, and  $N$  the total number of simulated events which is 600 k. Two additional free scaling factors  $n_1$  and  $n_2$  are used to account for the branching ratio of the first and second excited state of  $^{12}\text{C}$ . The rate  $R$  was set to a fixed value of 159 n/s in accordance with the measured source rate and the half-life of  $^{241}\text{Am}$ , as discussed in Chapter 6.

Figure 7.9 shows the best fit of the bottom CW11d2m source position. The majority of the fitted range shows a good agreement between template fit and data. Only in the tail, for areas larger than 70 pe, and between the two peaks the fit shows a deviation from calibration data. The second excited state of  $^{12}\text{C}$  was excluded from the fitting procedure as it has a too large degeneracy with the first excited state.

The estimated parameters of the best fit are shown in Table 7.1, along with the best-fit estimates for the other source positions. The plots of the other source positions can be found in the appendix in Figure B.44 on page 270. Also for the other source positions, templates and calibration data show a decent matching, judging based on the

<sup>2</sup>Also a statistical alternative fit approach has been tested, which is an extension of the Barlow–Beeston likelihood for template fitting [DA22]. This approach uses a likelihood that allows fitting weighted templates to a data distribution. However, the approach deems that all entries of the data distribution can be found in the weighted template distributions. This requirement leads in general to a poorer overall fit, as it is clear that the two templates are not enough to describe the data distribution.

Table 7.1. – Best fit values of the simulated template fits for the four different AmBe source positions. The uncertainties are given by the pure statistical uncertainties deduced from the fit.

source position:	$\epsilon_A$	$n_1$
top CW5d9m	$(0.9556 \pm 0.0016)$	$(0.4825 \pm 0.0010)$
top CCW5d0m	$(0.8964 \pm 0.0014)$	$(0.5247 \pm 0.0015)$
bottom CW6d6m	$(0.9277 \pm 0.0014)$	$(0.5130 \pm 0.0015)$
bottom CW11d2m	$(0.9532 \pm 0.0016)$	$(0.5018 \pm 0.0016)$

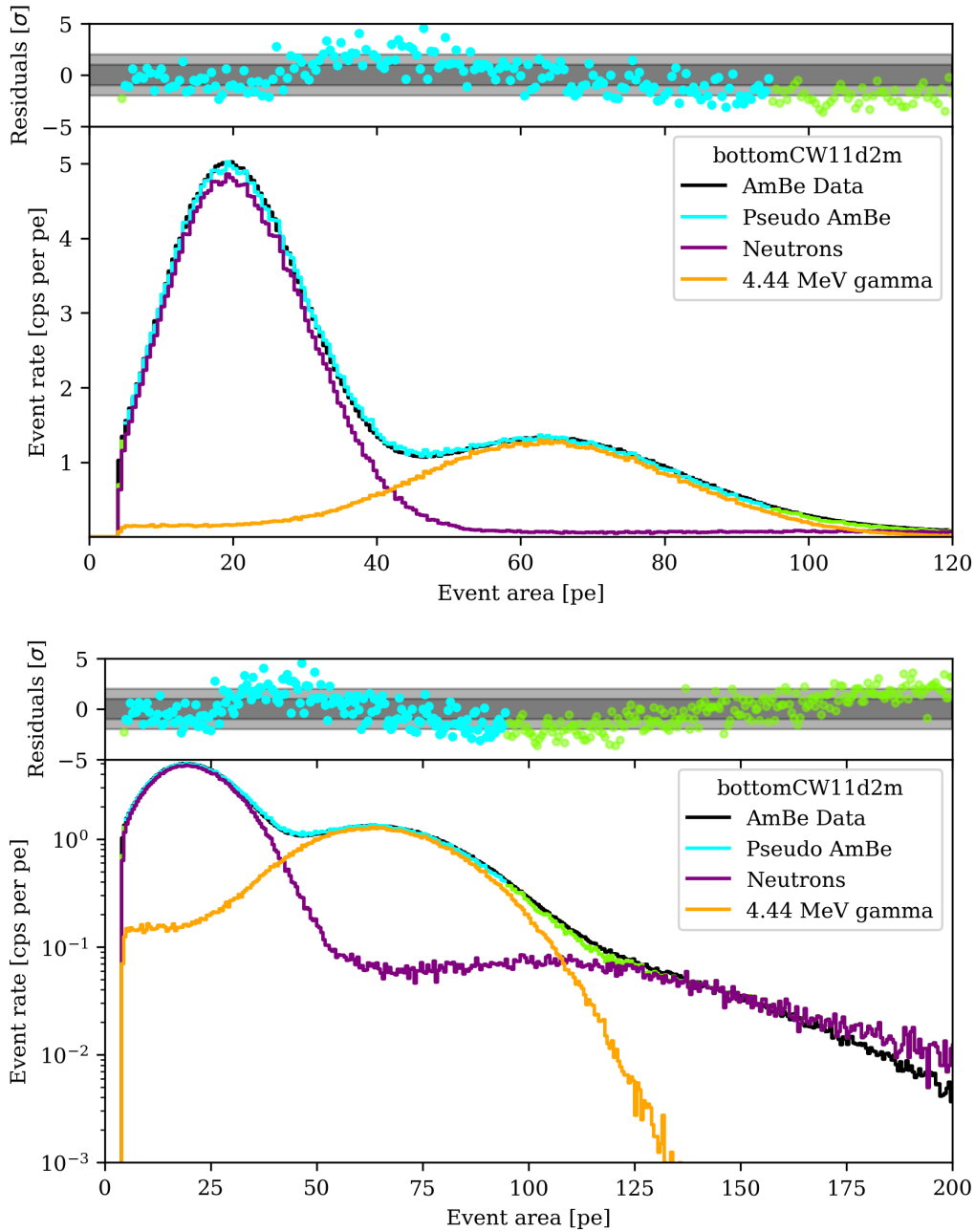


Figure 7.9. – Best fit of the simulated templates with the AmBe calibration data taken at the bottomCW11d2m source position. The upper plot shows the fitted templates for a linear scale while the lower plot shows a logarithmic scaling and a wider range in event area. In both figures the main plot shows the data and best fit in black and cyan respectively. Points shown in green are showing the resulting template beyond the used fit range. The two individual fit components of neutron and 4.44 MeV  $\gamma$ -ray templates are shown in purple and orange. Uncertainties of the individual bins are not shown. The upper panel in both figures shows the distribution of fit residuals normalized by the combined statistical uncertainty  $\sigma_{\text{tot}}$  of toy MC and calibration data.

fitting residuals. All fits show a deviation from data, especially at the tails of the fitted region. The under-fluctuation of the simulated templates at small energies can be understood as a missing background component compared to calibration data. Although the AmBe data spectrum was bin-wise background subtracted, the background conditions during AmBe calibration might be slightly different than during regular science data runs. Given the large amount of additional light produced by the calibration, a higher number of accidental events may exceed the selection threshold. A toy MC approach was used to evaluate the p-value for the individual fits. For each fit, one thousand new toy data distributions were sampled from the best fit and refitted afterwards. The p-values of the best fit for all four source positions were evaluated to be zero, rejecting the hypothesis that the templates are describing the data entirely. This is partially expected as it was not possible to fit all source components to data.

The selection acceptances  $\epsilon_A$  varies between 89 % and 96 % with an average of  $(93.3 \pm 1.4)$  %. It is not entirely clear why the selection acceptance varies so much between the different source positions. Further, it is not clear why some of the fits prefer a selection efficiency larger than 90 %. Already the spatial selection should lead to an acceptance of only  $0.95 \cdot 0.95 \approx 90$  % as the cut boundaries of the angular and  $z$  distribution are defined based on the central 95 % of the corresponding distribution. A weak anti-correlation between radial distance and cut acceptance can be observed. This might point to either a mismodeling of the spatial selection for the individual source positions or an asymmetry in the detector configuration which is not reflected in simulations. The latter would also explain the discrepancy of the  $CE_{\text{eff}}$  for the individual source positions.  $CE_{\text{eff}}$  varies between 97.5 % and 104.0 %.

A stronger discrepancy between data and simulation can be observed by comparing the fitted templates on a logarithmic scale for larger energies. As discussed in the previous section this high energy tail is likely caused by neutron capture events on other isotopes like  $^{56}\text{Fe}$  which can be found in the stainless steel vessel of the TPC. The discrepancy in the high energy tail is caused by an issue of the Geant4 MC code which originates from a wrong treatment of  $\gamma$ -cascades as is discussed in [Men+20]. Given this issue and the large degeneracy between the first and second excited state of  $^{12}\text{C}$ , it is currently not possible to model the second excited state correctly. The model tends to over-predict the contribution from the second excited state with respect to the first one. Thus,  $n_2$  was consequently fixed to be zero in all fits.

The Geant4 issue has additional implications for the estimate of the branching ratio of the first excited of  $^{12}\text{C}$ . The best fit estimates show a clear correlation with the source distance towards the TPC cryostat. This correlation originates from the fact that the fraction of neutron capture events on cryostat materials increases when moving the source closer. The correlation of the branching ratio with radial distance is shown in Figure 7.10. In order to get a true estimate for the branching ratio of the first excited state this correlation must be taken into account as a systematic uncertainty. The source position top CCW5d0m is farthest away from the TPC cryostat and has therefore the smallest potential bias due to the Geant4 bug. In order to get an estimate for the potential systematic bias the template of the neutron-simulation was truncated at 50 pe and the branching ratio reevaluated. The difference of the first excited state with and without high energy contribution from neutrons is about 3 % for the top



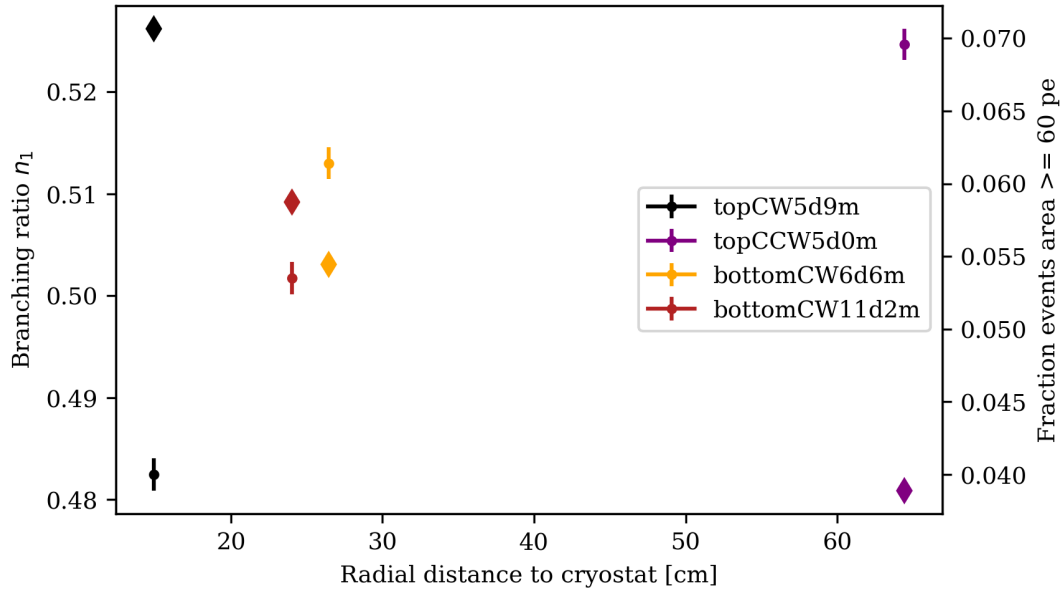


Figure 7.10. – Branching ratio and fraction of high energy neutron capture events as a function of radial distance towards the TPC cryostat. The dots are showing the best fit estimates for the branching ratio of the first excited state of  $^{12}\text{C}$ . The uncertainties are the statistical uncertainties from the best fit. The diamonds are showing the fraction of high energy neutron capture events producing more than 60 pe. Statistical uncertainties are not shown for the fraction estimate as they are too small.

CCW5d0m source position which is considered as a systematic uncertainty of this study.

Another source of systematic uncertainty comes from the fact that the second excited state cannot be fitted to the distribution directly. Depending on the alpha energies and the smearing of the alpha energies the second excited state may also contribute to the 4.44 MeV  $\gamma$ -peak. Figure 7.11 shows the branching ratio of the first and second excited state assuming different alpha energies and smearing together with the best fit of the first excited state branching ratio for the top CCW5d0m source position. Depending on the alpha energy and smearing model the second excited state may contribute up to 5% to the 4.44 MeV  $\gamma$ -peak. However, it should be noted that this is just an estimate and that the precise distribution of alpha energies is unknown as no information about the AmBe powder is available.

Last but not least one has to consider that the spatial distribution is slightly different for the 2.22 MeV and 4.44 MeV line as shown in Figure 7.8. The 4.44 MeV  $\gamma$ -line tends to have a smaller spatial distribution which will lead to a small bias in the branching ratio for the excited states as more neutron capture events than 4.44 MeV events are cut by the spatial selection. To get an estimate for this selection bias the fraction of remaining events for all four source positions of the simulated data was computed. In average over all four source positions  $(1.067 \pm 0.002)$  times more gamma than neutron events pass the spatial selection. Considering this bias as well as the discussed systematic uncertainties the resulting branching ratio of the first excited state is estimated to be:

$$n_1 = 0.491 \pm 0.002(\text{stats.})_{-0.050}^{+0.030}(\text{sys.})$$

## 7.2. NEUTRON VETO RESPONSE AND AMBE TEMPLATE MODEL USING TOY MCS

---

where the value and the statistical uncertainties are coming from the best fit of the top CCW5d0m source position. This value deviates from the usually stated value of 60 % in literature.

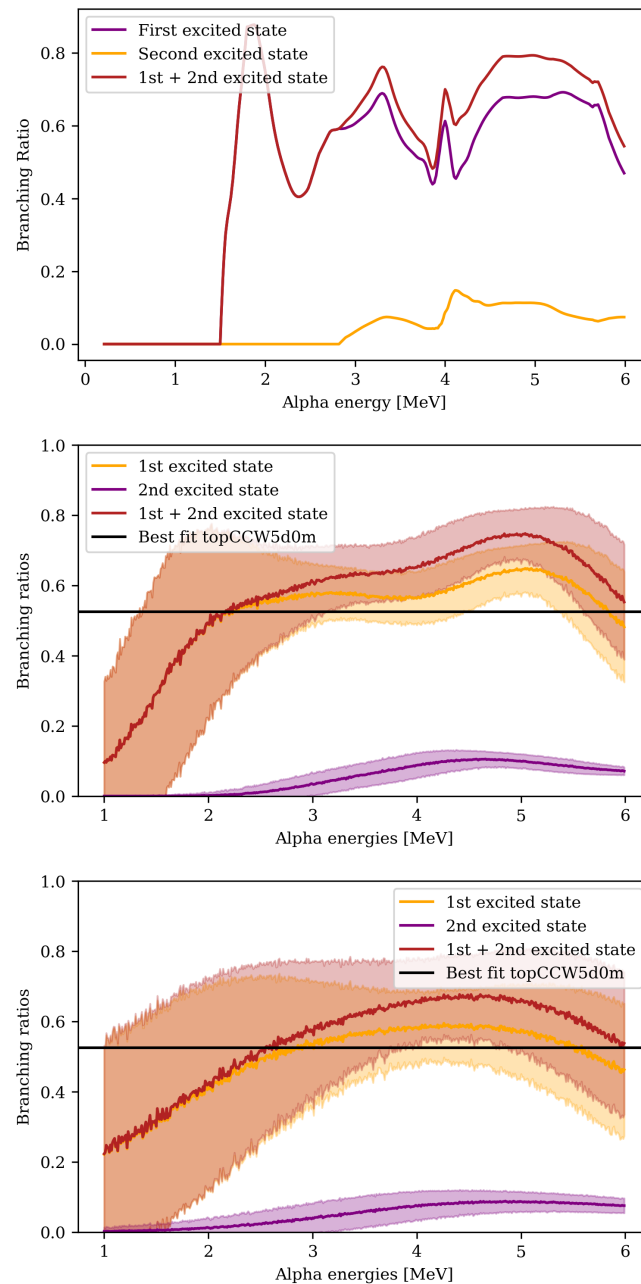


Figure 7.11. – Branching ratios of the first excited state based on the data taken from Figure 6.1. The first plot shows the exact branching ratios for the first and second excited state as well as their sum as a function of alpha energy. The second and third plot shows the resulting branching ratios for a normal distributed alpha-energy with a smearing of 0.5 MeV and 1 MeV respectively. The shaded region indicate a 1 sigma contour. The black horizontal line indicates the branching ratio estimate for the best fit of the top CCW5d0m source position. At the intercept between the black and red line the contribution of the second excited state is about 0% and 5% for the 0.5 MeV and 1 MeV smearing respectively.

### 7.2.3 NV response

Understanding the energy response of the NV just based on the AmBe calibration data is not possible. Additional mono-energetic calibration lines are necessary to model the energy response adequately. To make a prediction and model the energy response nonetheless the simulation framework can be used. For this purpose mono-energetic  $\gamma$ -rays were simulated in an energy range between 1.5 MeV and 9 MeV with a stepping of 0.5 MeV till 6 MeV and a stepping of 1 MeV afterwards. The gammas were simulated as an isotropic source starting at the TopCCW5d0m source position. For each line below 6 MeV 400 000 gammas were simulated and half the statistics for the higher energies. The spectrum of each line is represented using the same model for the Compton continuum and full energy peak as for the 4.44 MeV  $\gamma$ -line in Section 7.1. In addition, the model is truncated with the threshold function expressed in equation 7.4 as done before. The simulated data and best fit for all simulations are shown in Figure 7.12. The optimization was performed with a binned representation of the simulated data using a 0.5 pe binning to capture the shape of the spectra with sufficient resolution. A  $\chi^2$  minimization is applied restricted to bins with at least 40 entries. In addition, the fit range was limited at small event area. The reduced chi-square ranges between 3.1 at the lowest and 1.1 at the highest simulated energy. The corresponding p-values are all below the 10 % threshold rejection that the model describes the data entirely. Based on the fit residuals a slight systematic can be observed for all fits. As an example, the fit residuals of the 4 MeV fit are shown in the appendix in Figure B.47

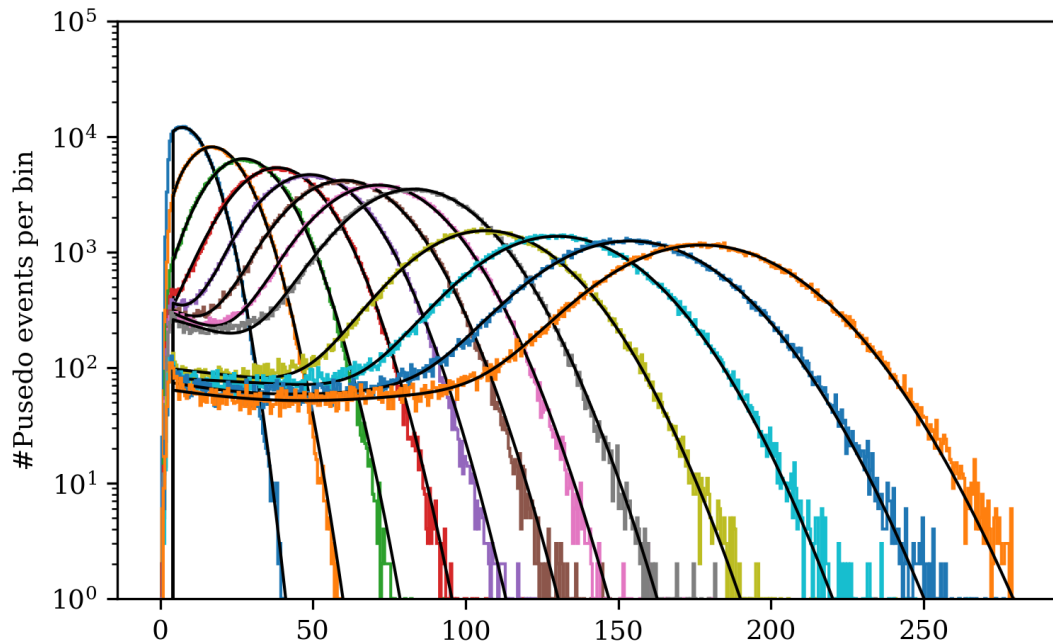
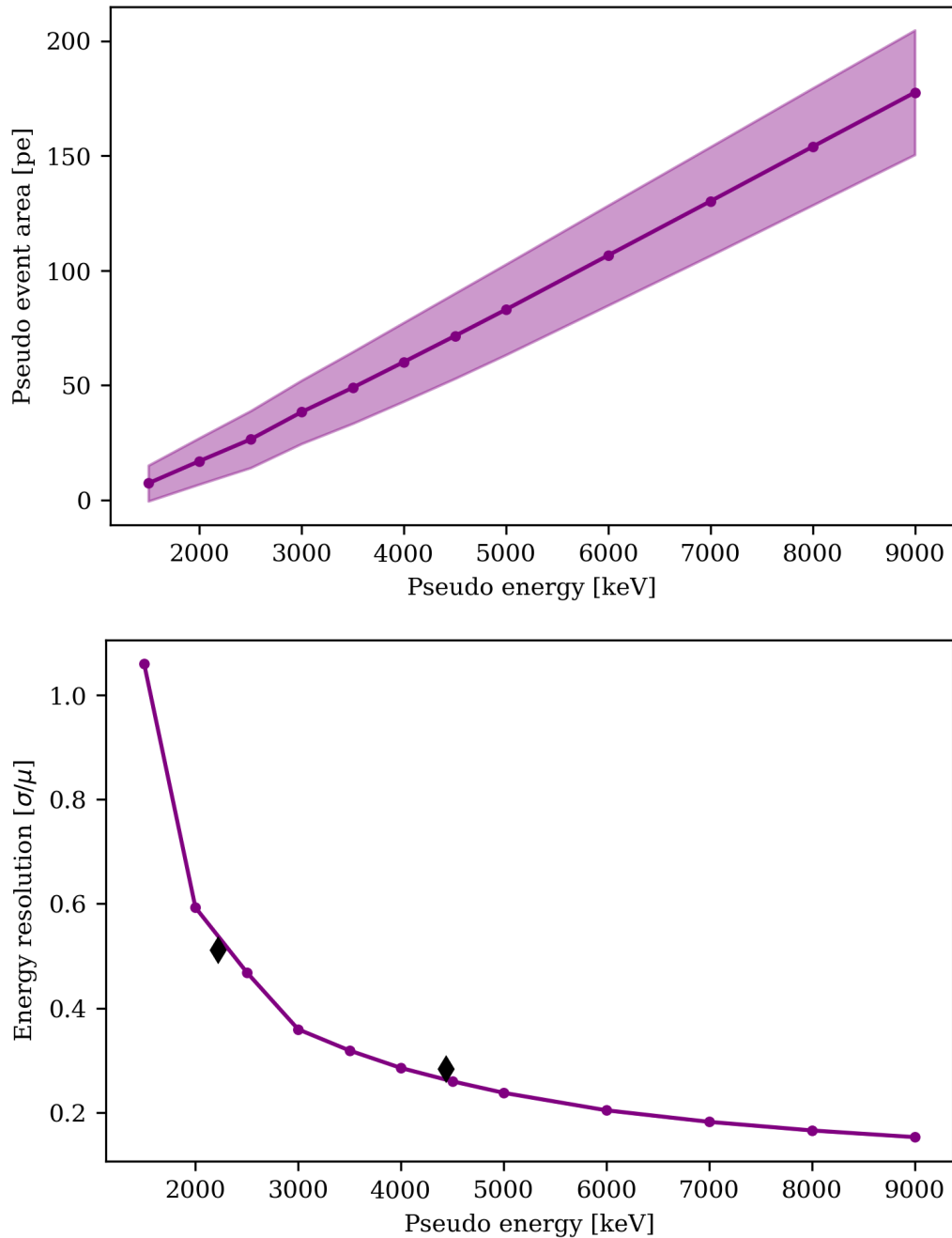


Figure 7.12. – Binned representation of the simulated pseudo  $\gamma$ -lines explained in the text. The black lines indicate the best fit of the individual lines with the explained model.

on page 272. Although the model seems to be incomplete, the individual fits show a good description of the full energy peak. The simulation predicts that for a 1.5 MeV line a fraction of  $(78 \pm 4)\%$  all simulated events can be found above a threshold of 4 pe. Extracting the best-fit values of the corresponding full-energy peaks allows to predict the energy response and resolution of the NV. Both are shown in Figure 7.13. The energy response of the NV is predicted to be almost linear over the entire energy range. Only at small energies the response shows a reduced inclination. Data and simulation show at least by eye a relative good agreement for the derived energy resolution. It is difficult to characterize the agreement in a more robust way as both, the model fit to data in Section 7.1 as well as the simulation shown here, suffer from an incomplete fit model which leads to a systematic uncertainty dominating over their respective small statistical uncertainties. Calibrations with additional mono-energetic lines are required to validate the model further. However, already now it can be used to get a good first order estimate of the NV response.



**Figure 7.13.** – Predicted NV energy response and resolution. The upper plot shows the predicted NV energy response for the given mono-energetic  $\gamma$ -ray simulated. The dots indicate the best fit of the full-energy peak mean shown in Figure 7.12. The shaded region indicate the standard deviation of the fitted full energy peak. The lower plot shows the predicted energy resolution of the NV as a function of the simulated gamma energies. The black diamonds indicate the energy response derived from the average of the best fit values of Section 7.1. Statistical uncertainties of simulation and data are too small to be shown. In both plots, the points derived from simulation are linearly interpolated to guide the readers eye.

7.3 TAGGING EFFICIENCY

The NV tagging efficiency  $\zeta$  is the most important parameter of the NV. It characterizes how many "dangerous" neutron signals in the TPC are tagged by the NV due to a delayed coincidence with the 2.22 MeV hydrogen neutron capture signal. A neutron signal is considered to be dangerous if it produces a single scatter NR which is indistinguishable from WIMP DM. The AmBe NR single scatter calibration data of Chapter 6 can be used to estimate the tagging efficiency. The event topology for calibration and neutron background data is illustrated in Figure 7.14. Both have almost an identical topology except for the positional origin of the neutrons and the additional 4.44 MeV  $\gamma$ -ray signal, detected by the NV, which was used to select the NR single scatter events. This must be taken into account when estimating the true tagging efficiency for background neutrons.

In general, the tagging efficiency depends on the selection thresholds applied to the NV events and the length of the tagging window in which the delayed coincident signals are searched for in the NV. Both must be optimized based on the desired tagging efficiency and affordable lifetime loss of the experiment. The tagging efficiency is then given by the number of neutron capture signals seen by the NV for a given selection divided by the total number of single scatter NR signals seen by the TPC.

To briefly recapitulate, the calibration of the single scatter NR response requires that all events are found in a coincidence between the 4.44 MeV  $\gamma$ -ray and the NR S1 signal

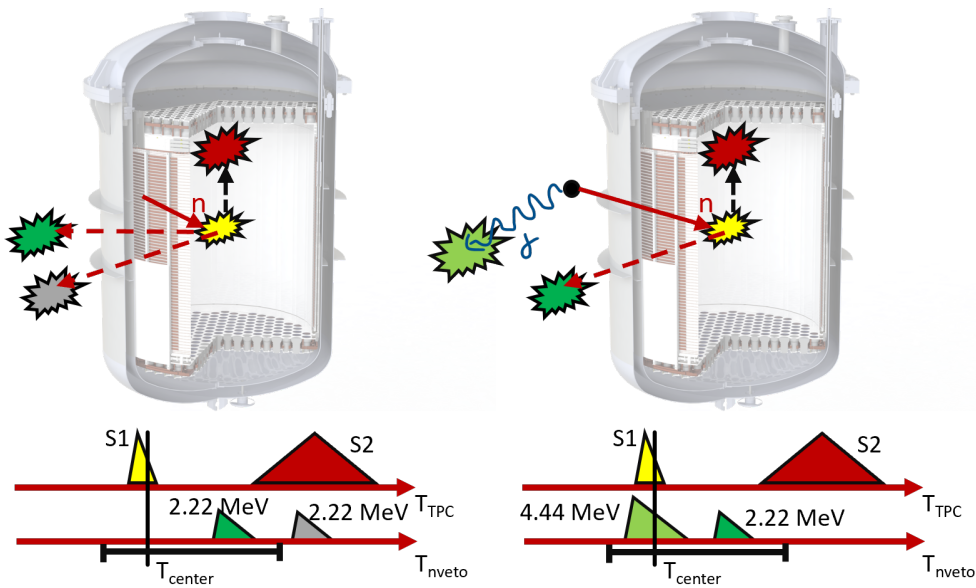


Figure 7.14. – Schematic showing the working principle of the NV tagging for radiogenic neutrons emitted by the TPC materials (left) and the calibration of the NV tagging efficiency via AmBe (right). The bottom panel in both figures illustrates the typical timing of the involved signals. The veto-window used for the neutron-tagging is indicated by the horizontal black bar. The neutron-capture events in dark (and bright) green are tagged, while the gray event would be found to be outside of the tagging window.



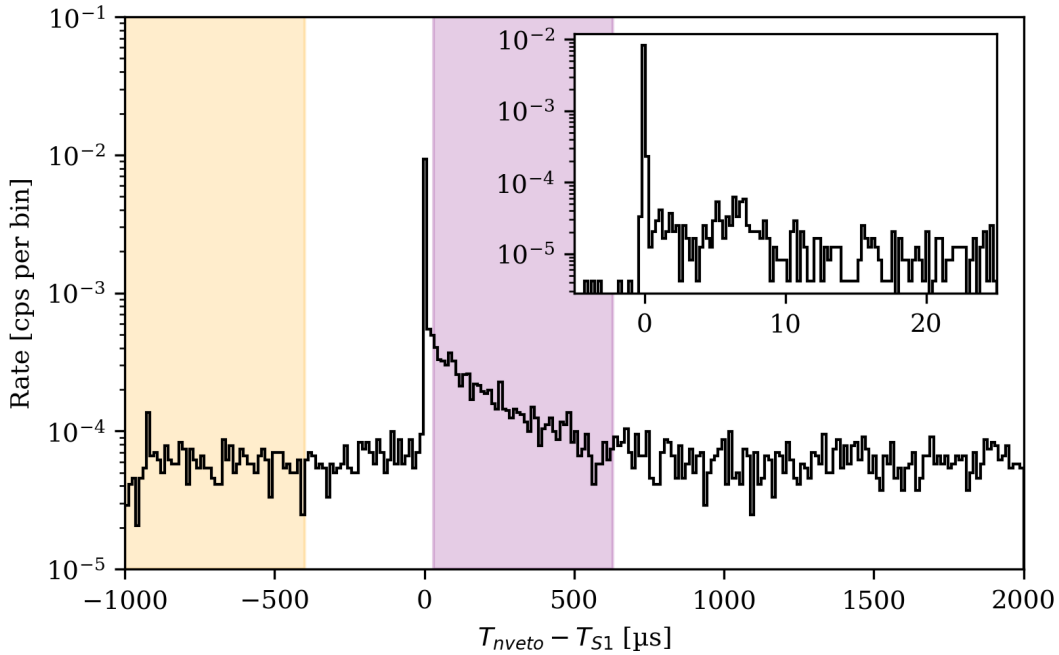


Figure 7.15. – Time spectrum showing all NV events found in coincidence with a single scatter NR event. The inlet figure shows a zoom onto the time range between  $-5 \mu\text{s}$  and  $25 \mu\text{s}$ . The purple and orange shaded regions are explained in the text.

recorded by the NV and TPC respectively. Supplementary data quality cuts are used to select only a pure sample of single scatter NR interactions based on their peak and event properties. To further purify the selection, each event must be found within the 90 % contour of the sliced NR band fit discussed in Section 6.3. Finally, the cS1 of the selected events must not be larger than 200 pe. The selected data is shown in Figure 6.17 on page 132. In total 1989 events are found using this selection.

### 7.3.1 neutron capture signals and coincidence time distribution

To determine the number of tagged NR interactions one has to select all NV signals which can be found in coincidence with one of the NR interactions. The coincidence is computed by a plugin that is inherited from the base coincidence plugins described in Section 3.3.5. Similar to the NR calibration study the coincidence is build between the center time of the S1 signal and the NV event. However, in this study, the S1 signal serves as a triggering signal and the coincidence window was set to  $-1000 \mu\text{s}$  and  $2000 \mu\text{s}$  around its center time. Given the small duration of NV events, which is in the order of 200 ns, multiple events can be found for a single S1 signal. In such a case all events were considered to be in coincidence. No additional selection was required for the coincidence.

The resulting relative time distribution is shown in Figure 7.15. The spectrum is featured by a sharp peak-like structure centered around zero which is followed by an

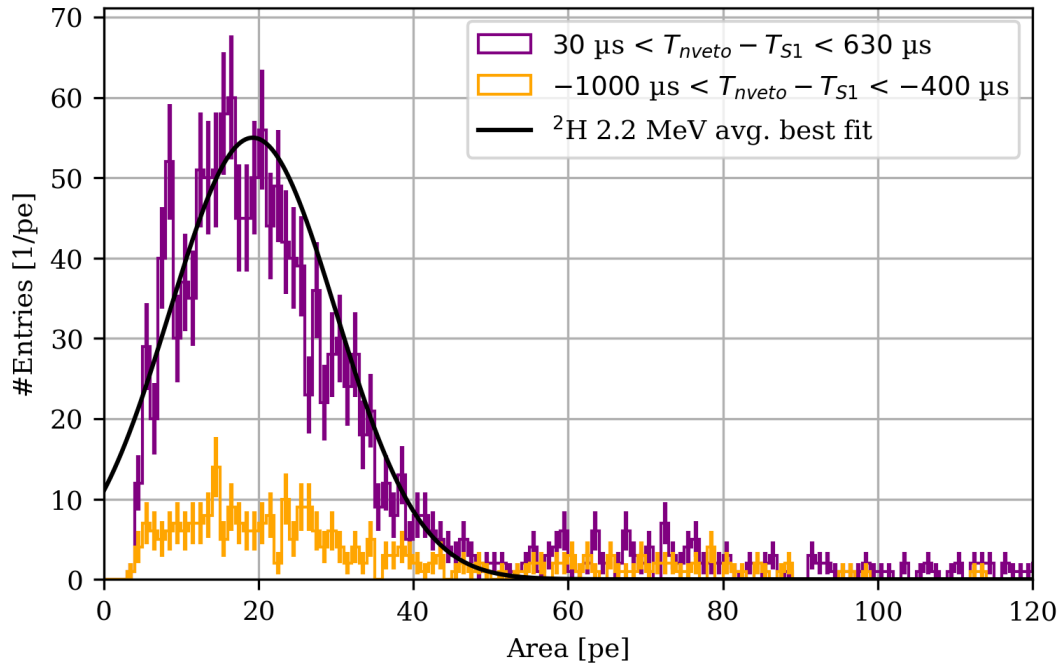


Figure 7.16. – Event area distribution for the signal and background region discussed in the text, shown in purple and orange respectively. The black line indicates the best fit of the 2.22 MeV hydrogen neutron capture line based on averaged best fit parameters of the individual fits discussed in Section 7.5. Its amplitude was scaled to match the shown data. Statistical uncertainties of the individual bins are shown as vertical lines.

exponentially declining distribution. Both are sitting on top of a uniform pedestal. The peak is caused by the prior requirement that all single scatter NR signals must be found in coincidence with a NV signal caused by the 4.44 MeV  $\gamma$ -ray. Thus these  $\gamma$ -events are found back in the coincidence of this study. The zoomed inlet shows two smaller peak-like structures close to the  $\gamma$ -peak around 2  $\mu$ s and 8  $\mu$ s respectively. This increase in the number of accidental coincidences is caused by an increased rate of background events in the NV due to PMT afterpulses occurring after the large 4.44 MeV  $\gamma$ -ray signal. The exponentially falling distribution originates from signals which are caused by radiative neutron capture signals on different materials e.g. hydrogen. Its decay constant is determined by an effective capture time given by all isotopes capturing neutrons and producing signals recorded by the NV. The flat uniform component is given by accidental coincidences between the two detectors.

To estimate the number of neutron capture signals seen by the NV the number of events found in the purple region of interest between 30  $\mu$ s and 630  $\mu$ s is compared against the orange background reference region between  $-1000 \mu$ s and  $-400 \mu$ s. The boundaries of the purple region have been chosen such that the number of neutron capture signals is maximized while avoiding the region between 0  $\mu$ s and 30  $\mu$ s, which is polluted by additional background events due to PMT afterpulses. The area spectra of the selected events are shown in Figure 7.16. Events found in the purple signal region follow a peak-like structure that matches the overall shape of the neutron cap-

ture peak of hydrogen inferred from the best fit discussed in Section 7.5 at the end of this chapter. The binned area distribution shows a statistical upward fluctuation in the bins with an area of 7 pe and 8 pe when compared to the best fit. The cause of this fluctuation is unclear. It can be neither found in other studies computing similar spectra, as discussed in Sections 7.4 and 7.5, nor are these events showing properties different from the remaining events. An explanation that would arise naturally is PMT afterpulses which produce signals of a defined area, but the events found in these bins do neither show a smaller center time compared to regular events nor a large area contribution from a single PMT, which is characteristic for such a signal. Thus the fluctuation might be a statistical fluke.

It should be noticed that Figure 7.16 shows only events which are passing a 5-fold coincidence which is one of the selection criteria used in the final NV tagging, discussed later in this section.

Integrating the purple distribution and subtracting the orange background reference region in Figure 7.16 already provides the number of tagged NR single scatter signals and therefore the tagging efficiency for the given veto window of [30  $\mu$ s, 630  $\mu$ s). However, the window is not very applicable for science data as a large fraction of capture events appear between 0  $\mu$ s and 30  $\mu$ s, and the total window length may also differ from 600  $\mu$ s due to other constraints like the detector deadtime. Thus to compute the correct tagging efficiency one needs to correct the given selection of Figure 7.16 according to the decay constant of the exponential distribution and the chosen veto window used for the WIMP search.

### 7.3.2 Veto window correction

The veto window correction can be computed via the cumulative probability density function of an exponential distribution when the decay constant of the exponential distribution is known. To determine the radiative neutron capture time constant the entire distribution except for the  $\gamma$ -peak and afterpulse region between  $-2.5 \mu$ s and 30  $\mu$ s was fitted using an extended unbinned log-likelihood method. An unbinned method was chosen due to the small statistics of the data which would otherwise require a coarse binning. The fit model is given by a probability density function of the form

$$\Delta T_{pdf}(t, n_{sig}, n_{bkg}, t_0, \tau, \tau_{mod}, a, b) = \frac{n_{sig}}{n_{norm}} \exp_{pdf}(t, t_0, \tau) \cdot \exp_{cdf}(t, t_0, \tau_{mod}) + B \quad (7.10)$$

where  $n_{sig}$  and  $n_{bkg}$  are the number of signal and background events respectively.  $t_0$  represents a time shift with respect to zero. The neutron capture is modeled by an exponential probability density function of the form

$$\exp_{pdf}(t, t_0, \tau) = \frac{1}{\tau} \exp \left\{ \frac{-(t - t_0)}{\tau} \right\} \quad \text{if } t \geq t_0 \text{ and } 0 \text{ else} \quad (7.11)$$

with the neutron capture time constant  $\tau$ . The term  $\exp_{cdf}(t, t_0, \tau_{mod})$  describes the moderation of the neutrons before being captured which is given by the cumulative

of an exponential distribution of the form of Equation 7.11 with the moderation constant  $\tau_{mod}$ . Finally,  $B$  represents a background term describing accidental coincidences which are given by

$$B = n_{bkg} \cdot U_{pdf}(x, a, b) \quad (7.12)$$

in which  $U_{pdf}(x, a, b)$  is a uniform distribution between  $a$  and  $b$ <sup>3</sup>. A normalization constant  $n_{norm}$  given by

$$n_{norm} = 1 + 1/(\tau_{mod} + \tau) \quad (7.13)$$

converts the fit model into a proper probability density function. The best fit of  $\Delta T_{pdf}$  to the data is shown in Figure 7.17 in two different time ranges. The fitting was performed in multiple stages by fitting first the flat background component between  $-1000 \mu\text{s}$  and  $-500 \mu\text{s}$  as well as  $1500 \mu\text{s}$  and  $2000 \mu\text{s}$ , before optimizing the parameters of the exponential distribution. The time offset  $t_0$  was kept fixed at  $30 \mu\text{s}$ . The goodness of fit was evaluated using a binned Poisson  $\chi^2$  test [BC84] with the same binning as shown in the upper graph of Figure 7.17. The best fit has a p-value of 0.31 failing to reject the fit model. The residuals between the binned data and fit model do not show any systematic behavior, thus showing a good fit. The results of the best fit are summarized in Table 7.2. The decay constant was estimated to be  $(180 \pm 8) \mu\text{s}$  and the moderation was found to be  $(0 \pm 15) \mu\text{s}$ . The large uncertainty of the moderation time constant is due to the excluded fit range around zero. The fact that the moderation is preferred by the fit to be zero would be consistent with expectations as the capture cross section of hydrogen scales with  $1/v$  where  $v$  is the velocity of the neutron as discussed in Section 2.3. However, given the large uncertainty, it is not clear if the moderation is indeed zero. The neutron capture time is significantly smaller than expected based on pure-water only<sup>4</sup>. The reason for this is the capture of neutrons on other materials like  $^{56}\text{Fe}$  as shown later in Sections 7.4 and 7.5.

The resulting correction factor for the different veto windows can be estimated via

$$\epsilon_{vw}(t_0, t_1) = \frac{\exp_{cdf}(t_1, 0, \tau) - \exp_{cdf}(t_0, 0, \tau)}{\exp_{cdf}(630 \mu\text{s}, 0, \tau) - \exp_{cdf}(30 \mu\text{s}, 0, \tau)} \quad (7.14)$$

where the denominator describes the fraction of capture events found in the  $[30 \mu\text{s}, 630 \mu\text{s})$  window used in calibration and the numerator the expected fraction of events found in the veto window between  $[t_0, t_1)$  used in the WIMP search. This results in

<sup>3</sup>Due to the excluded  $\gamma$ -peak region in the middle of the distribution the background is subdivided into two regions before and after the excluded range, but share a common amplitude.

<sup>4</sup>See Section 2.3 on page 34.

Table 7.2. – Best fit results of  $\Delta T_{pdf}$  for the NV tagging efficiency time distribution.

Parameter:	Value
$n_{sig}$	$(1200 \pm 40)$
$n_{bkg}$	$(1520 \pm 50)$
$\tau$	$(180 \pm 8) \mu\text{s}$
$\tau_{mod}$	$(0 \pm 15) \mu\text{s}$

a time correction factor of  $(0.92 \pm 0.02)$  and  $(1.18 \pm 0.01)$  for a veto window of  $[0 \mu\text{s}, 250 \mu\text{s})$  and  $[0 \mu\text{s}, 600 \mu\text{s})$  respectively. The uncertainty of the correction factor was estimated using a small toy MC by sampling  $\tau$  from a normal distribution.

One last point which needs to be considered to determine the true tagging efficiency is the geometrical difference between neutrons coming from AmBe and radiogenic neutrons coming from different detector materials which is discussed in the next section.

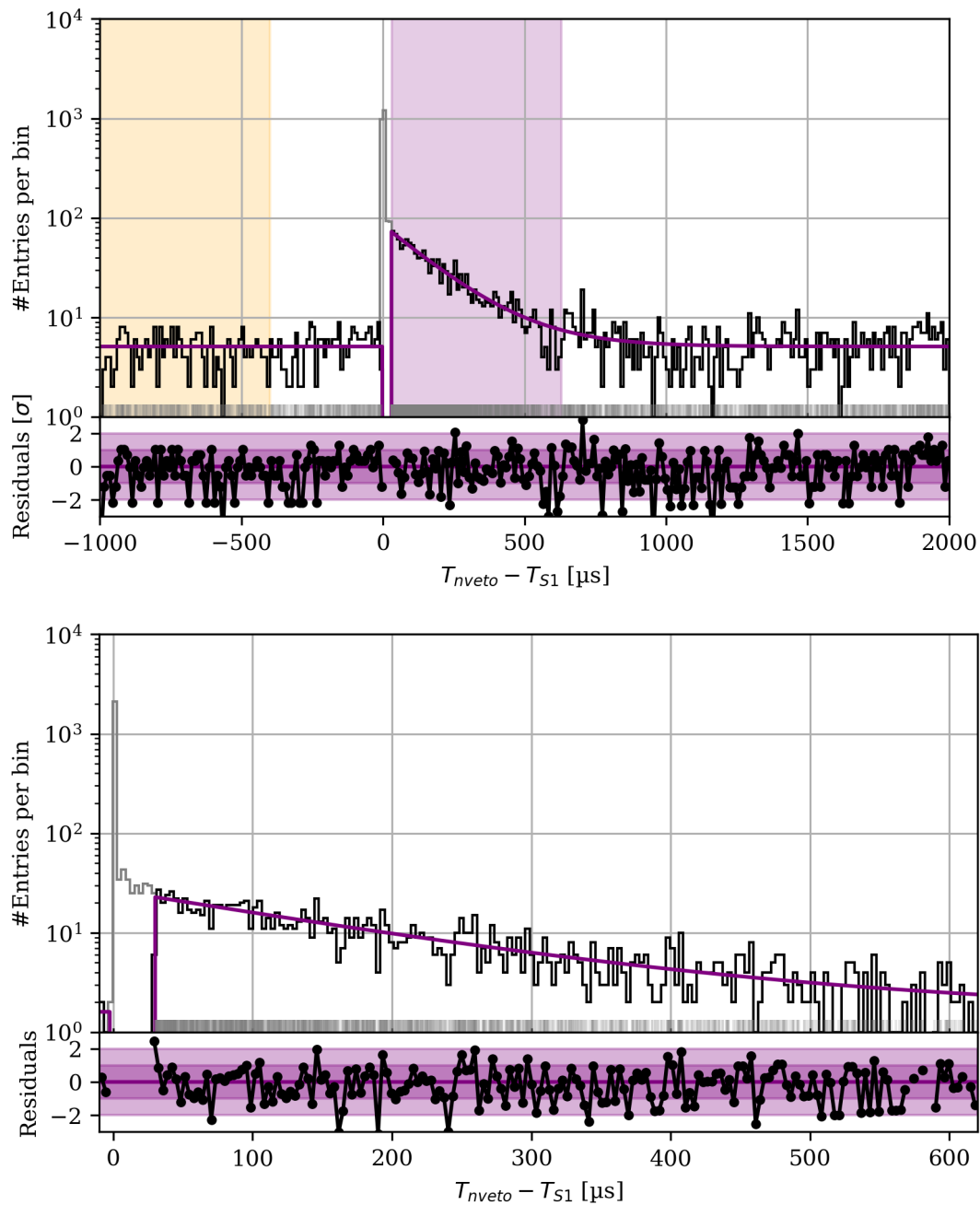


Figure 7.17. – Best fit of the time distribution as explained in the text. The upper plot shows the fit for the entire fit range while the bottom figure shows a zoom onto the most relevant range given by the exponential distribution. Both plots are subdivided into two panels. The main panel shows a binned and unbinned representation of the data as black histogram and gray vertical lines respectively. The gray shaded part of the histogram indicates the data in the area excluded from the fit. The purple line shows the result of the best fit. The bottom panel in both figures shows the residuals of the fit with respect to the binned representation of the data. The purple shaded regions are showing the one and two sigma contours of the best fit.

### 7.3.3 Geometric correction

To estimate the geometrical correction factor  $\epsilon_{geo}$  a series of Geant 4 simulations were conducted by Pavel Kavrigin from the Weizman Institute of Science, simulating different TPC components with the highest neutron yields. Among these components are the ceramic parts of the PMT bases, the PTFE walls of the TPC as well as the inner and outer wall of the stainless steel cryostat. Most of the neutrons emitted by these materials originate from spontaneous fission and  $(\alpha, n)$  reactions due to decays of isotopes in the uranium and thorium chains. In addition, Pavel simulated [Pav22] the AmBe neutrons for the topCW5d9m source position. The resulting tagging efficiencies for the different material backgrounds were averaged in accordance to their respective neutron-yields into a single tagging efficiency  $\zeta_{NRSSsim}^{TPC}$  and compared against the tagging efficiency derived from the AmBe calibration simulation  $\zeta_{NRSSsim}^{AmBe}$ . It should be noticed that in Pavel's simulations no constraint on the capture time window was set which corresponds to an infinitely long tagging window. To mimic this behavior and to compare the data driven tagging efficiency with simulated tagging efficiency the time window correction in equation (7.14) was calculated for a 1200  $\mu$ s long window, resulting in a tagging efficiency of. The result of simulation and data are shown in Figure 7.18. The tagging efficiency for neutrons originating from the different detector materials is slightly higher than the tagging efficiency estimated by AmBe. Thus the data-driven tagging efficiency is corrected by

$$\epsilon_{geo} = \zeta_{NRSSsim}^{TPC} / \zeta_{NRSSsim}^{AmBe} = (1.010 \pm 0.023) \quad (7.15)$$

where the uncertainty was estimated by Gaussian error propagation considering only statistical uncertainties of the simulated results. Data-driven and simulated estimates for the tagging efficiency agree within their statistical uncertainties, where the statistical uncertainties of the calibration dominate. The fact that the AmBe simulation only uses neutrons coming from the topCW5d9m position should not impact the result much as the majority of the calibration data is coming from the very same source position too, as it can be seen in Figure 6.4.

It should be noticed that during the time of Pavel's simulations, the fast effective toy simulator discussed in Section 7.2.1 did not exist yet. As a substitution, a simpler approach was used. The QE was applied in a similar approach as discussed in Section 7.2.1, and the SPE response was modeled by a simple normal distribution with a mean and a standard deviation of 1 pe and 0.4 pe respectively. The threshold was set to a value of 0.5 pe across all channels. Also, the corresponding pseudo-events were computed differently using the DBSCAN<sup>5</sup> algorithm for the event building. Due to this simplified approach, the comparison of the absolute tagging efficiency between simulation and data should be taken with a grain of salt. The impact on the geometrical correction factor on the other hand should be negligible as it is determined purely from simulations and was computed consistently in all simulations.

<sup>5</sup><https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html>



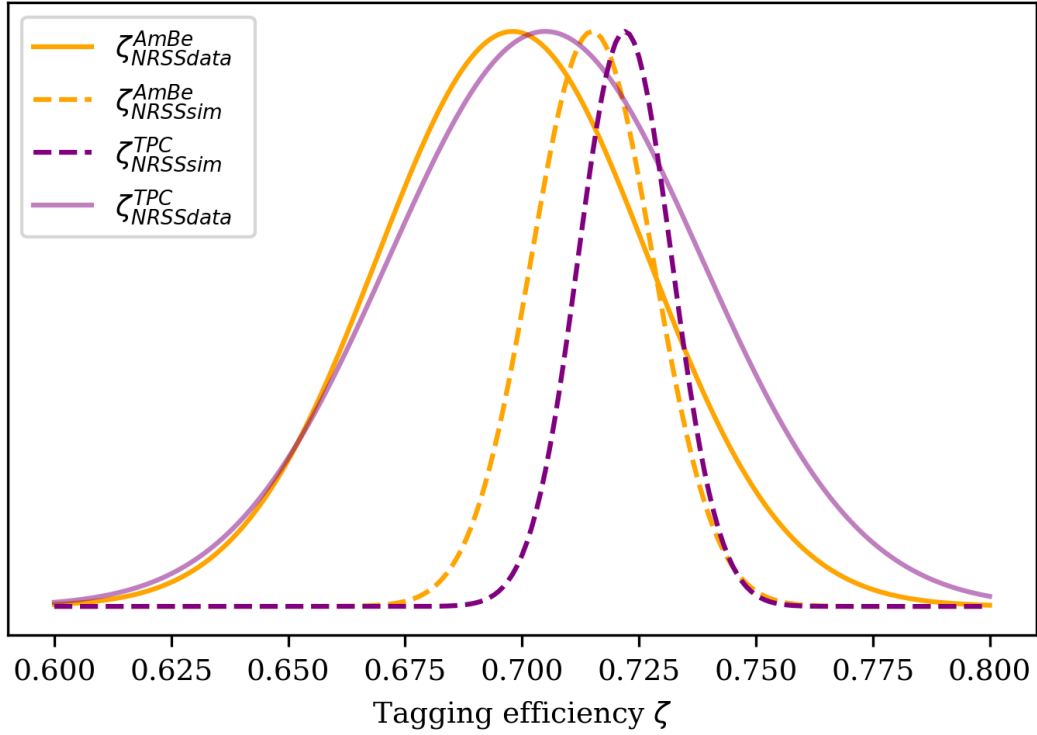


Figure 7.18. – Comparison of the simulated and data driven estimates for the tagging efficiency. The orange and purple dashed lines are showing the tagging efficiency for an "infinitely" long tagging window for simulated AmBe and background data respectively. The orange solid line shows the tagging efficiency derived from data for a 5-fold coincidence, a 5 pe threshold and a 1200  $\mu$ s long tagging window. The purple shaded line shows the data driven tagging efficiency after applying the geometrical correction factor.

### 7.3.4 SR0 tagging efficiency and deadtime

After estimating the number of neutron capture events given by Figure 7.16 and all correction factors the resulting tagging efficiency of the NV can then be estimated via

$$\zeta(t_0, t_1) = \epsilon_{vw}(t_0, t_1) \cdot \epsilon_{geo} \sum_{i=A_{th}}^{\infty} n_{sig,i} - n_{bkg,i} \quad (7.16)$$

where  $n_{sig,i}$ ,  $n_{bkg,i}$  are the number of neutron capture signals and background signals found in the  $i^{th}$  bin of Figure 7.16 for a given n-fold coincidence requirement and an event area threshold  $A_{th}$ ,  $\epsilon_{vw}(t_0, t_1)$  the veto window correction factor according to equation (7.14), and  $\epsilon_{geo}$  the geometrical correction factor. The resulting tagging efficiencies for a 250  $\mu$ s and 600  $\mu$ s long window are shown in Figure 7.19

The first science run of XENONnT use a NV tagging efficiency of

$$\zeta_{SR0} = (0.531 \pm 0.028)$$

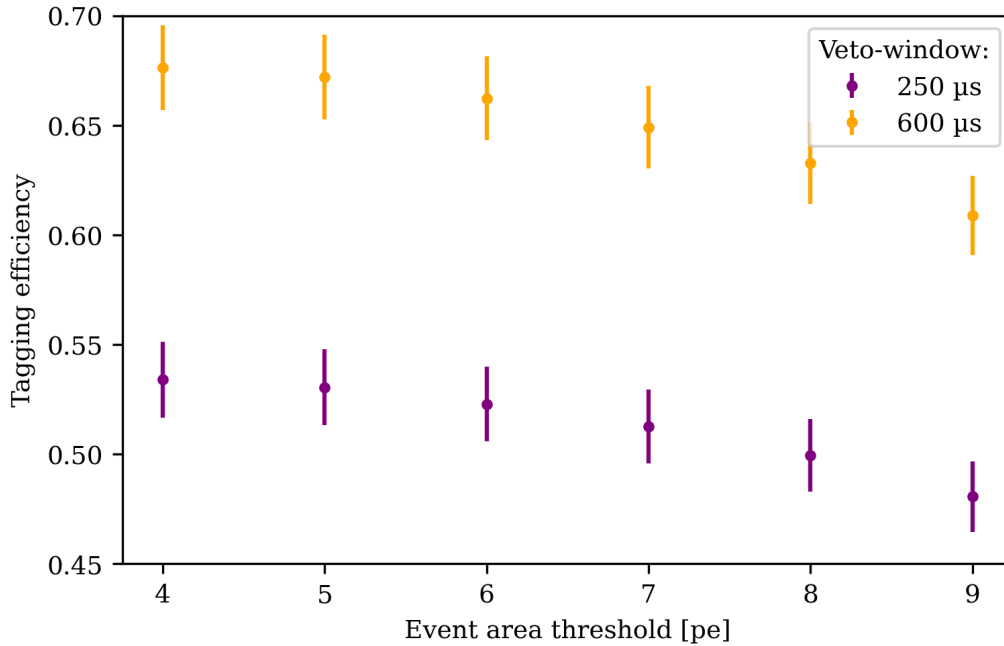


Figure 7.19. – SR0 NV tagging efficiency. The purple and orange points are showing the NV tagging efficiency as a function of event area threshold for a 250  $\mu\text{s}$  and 600  $\mu\text{s}$  respectively.

which is given by a 5-fold PMT coincidence and a 5 pe event area threshold, with a 250  $\mu\text{s}$  tagging window.

The tagging efficiency for a 600  $\mu\text{s}$  long tagging window using the same threshold is  $\zeta_{600\mu\text{s}} = (0.68 \pm 0.03)$ . Uncertainties of the different correction factors as well as the integrated, background-subtracted area spectrum were propagated using Gaussian error propagation. Only statistical uncertainties are included. Potential systematic uncertainties which could affect the result are discussed in the subsequent section.

The smaller tagging window was chosen as the deadtime caused by the NV tagging scales linearly with the window size, while the additional gain in tagging efficiency is smaller. The deadtime introduced by the NV tagging is computed according to the deadtime plugin explained in Chapter 3. Figure 7.20 shows the correlation of tagging efficiency and deadtime for different threshold and window lengths. The NV has the highest contribution to the total deadtime of the experiment. For the chosen tagging window the deadtime introduced by the NV is 1.6 % of the total 2.3 % deadtime of the experiment during SR0.

As a last validation cross-check the single-scatter NR tagging efficiency was compared against the tagging efficiency when allowing multi-scatter NR signals, too. Simulations suggest that there should not be any significant difference between the tagging efficiency of single or multi-scatter signals. To allow for multi-scatter events the single scatter S2 cut explained in Section 6.2.2 is dropped. This doubles the statistics to 4554 events found within the 90 % contour of the NR band. Estimating the number of neutron capture events and using the same correction factors as in the single scatter study leads to a tagging efficiency that is consistent with the pure single scatter

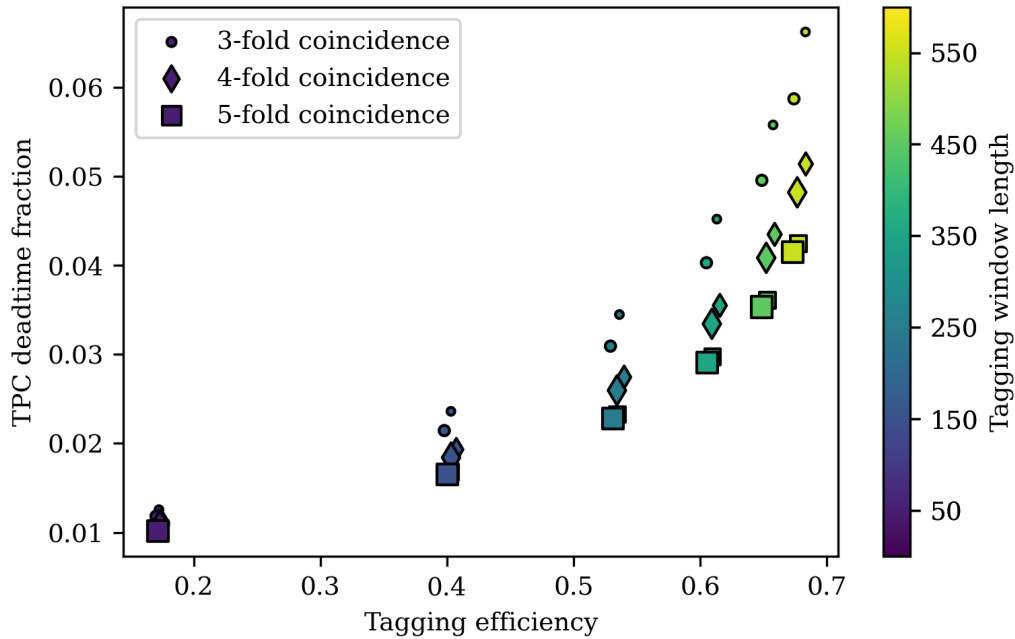


Figure 7.20. – Detector deadtime versus NV tagging efficiency. The different marker symbols show different  $n$ -fold coincidence requirements. The marker size reflects the area threshold. The smaller marker indicates a threshold of 4 pe while the bigger marker indicate a threshold of 5 pe. The length of the tagging window is indicated by the colorbar. Errorbars are not shown.

sample. However, if the veto window correction is evaluated again using the best fit of the corresponding time distribution a slightly smaller capture time of  $(172 \pm 5) \mu\text{s}$  is found, and thus a larger veto window correction must be used. It is slightly larger, by a factor of  $1.02 \pm 0.03$ , but both results are in perfect agreement within their statistical uncertainties. Hence, for SR0 the more conservative single scatter estimate is used. The best fit of the time distribution and the comparison in tagging efficiencies is shown in Figure B.39 and B.40 on page 266 and following.

### 7.3.5 Systematic uncertainties

In this last subsection, different potential sources of systematic uncertainties are highlighted and discussed. These uncertainties are subdominant compared to the statistical uncertainties presented together with the result, but should already be taken into account as additional considerations for future calibrations.

Given that the time window correction factor has a significant impact on the resulting tagging efficiency, a systematic study was conducted to see whether the choice of integration window used in calibration data can lead to a significant bias in the tagging efficiency. Therefore the start time and window length of the  $[30 \mu\text{s}, 630 \mu\text{s})$  calibration window was varied, and the area distribution of the purple signal and orange reference region reevaluated. The length of the reference region was kept at the same length as the signal window, as done before in Figure 7.16. To benchmark the

change in tagging efficiency the tagging efficiency for a  $[0 \mu\text{s}, 600 \mu\text{s})$  was evaluated as explained in Section 7.3.4, applying the time window correction according to equation 7.14 using the varied parameters of the calibration time windows. Figure 7.21 shows a comparison of the tagging efficiency derived for the nominal  $[30 \mu\text{s}, 630 \mu\text{s})$  calibration window with respect to other choices for window start and window length. A small bias can be observed when varying the total window length. Larger calibration windows tend to have slightly larger tagging efficiencies, but not always continuously as sometimes larger windows prefer a smaller efficiency. All values found are still within their statistical uncertainties. However, this effect might play a role in future calibration when higher statistics are achieved in the calibration of the single scatter NR response of the TPC.

The choice of the window start does not show any significant bias as long as start times below  $20 \mu\text{s}$  are avoided, for which the tagging efficiency is estimated to be significantly larger. This bias comes from the afterpulse pollution between  $0 \mu\text{s}$  and  $30 \mu\text{s}$  as explained before. It can conclude that no significant systematic error is introduced by the choice of the calibration window for the SR0 result.

After SR0 one additional effect was found which may introduce a systematic bias of the NV tagging efficiency. Neutrons with sufficient kinetic energy can undergo an inelastic neutron scatter presumably on  $^{16}\text{O}$  which leads to the emission of a prompt  $6.92 \text{ MeV}$   $\gamma$ -ray. This  $\gamma$ -ray is well above the tagging threshold of the NV and thus may increase the tagging efficiency of the NV if the scattered neutron stays undetected afterwards. This effect is discussed in more detail in Section 7.5.2. For simulated AmBe events this effect was found to be  $\sim 3\%$  which is in the same order of magnitude as the statistical uncertainty of the tagging efficiency. However, as this effect depends on the initial energy of the emitted neutron this effect might be different for background neutrons originating from detector materials.

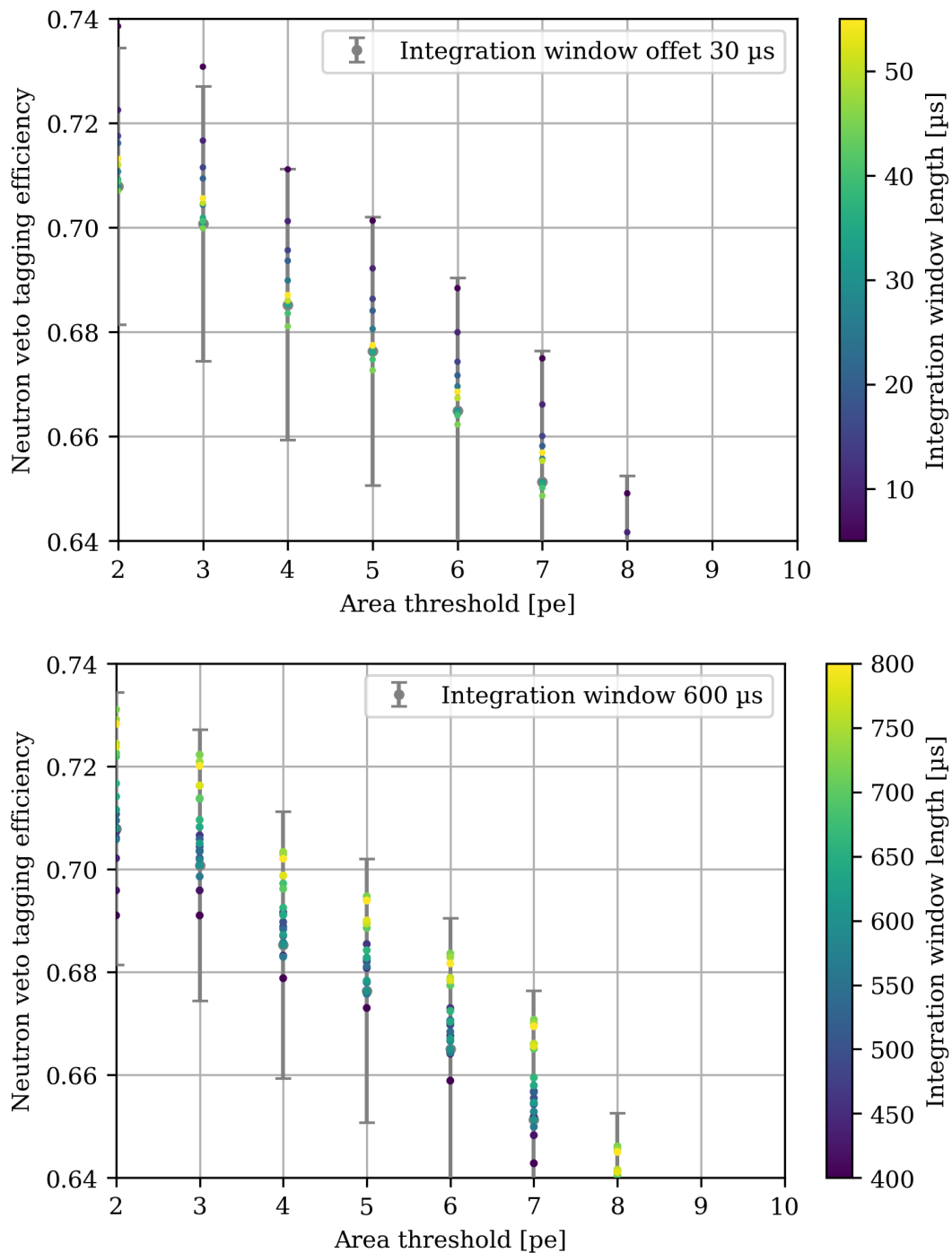


Figure 7.21. – Systematic bias study for the integration time window correction of the NV tagging efficiency. In both plots the gray points indicate the tagging efficiency for the nominal calibration window of  $[30 \mu\text{s}, 630 \mu\text{s})$  and the tagging efficiency window of  $[0 \mu\text{s}, 600 \mu\text{s})$ . The colored dots are representing the difference in tagging efficiency if the start time (top) or the length (bottom) of the calibration window is varied.

## 7.4 DETECTION EFFICIENCY

While the neutron-tagging efficiency is the most important quantity for a "NV" it is rather an abstract and specific parameter for a water Cerenkov detector in general. Most water Cerenkov detectors, like SNO or SK, use the neutron capture on hydrogen as an additional signal to discriminate signal from background. Thus for water Cerenkov detectors the more general "neutron-detection" efficiency is a parameter of greater interest. The detection efficiency  $\eta$  determines the chance of detecting a neutron given the number of emitted neutrons and must be by definition larger than the tagging efficiency. It can be calibrated via the AmBe calibration data, following a similar approach as for the NV tagging efficiency. Although a bit simpler as no simulation and geometric correction are required. To count the number of emitted neutrons, the coincidence between NV and TPC used to calibrate the NR response is inverted. Instead of searching NR events in the TPC by triggering on the 4.44 MeV  $\gamma$ -ray in the NV, the TPC triggers on the 4.44 MeV  $\gamma$ -ray to search for delayed coincident neutron capture events in the NV. The coincidence was computed by inheriting the base plugin described in Chapter 3.

The selection of the 4.44 MeV  $\gamma$ -ray events in the TPC uses the same data quality cuts as for the NR response calibration discussed in Section 6.2.2. To select only 4.44 MeV  $\gamma$ -ray events the full-energy ellipse is fitted in the cS2 versus cS1 space using an unbinned extended negative log-likelihood method. The fit model is given by

$$f(cS1, cS2) = n_{sig} \cdot N(cS1, cS2, \mu_x, \mu_y, \sigma_x, \sigma_y, \rho) + n_{bkg} \cdot U(cS1, cS2, \Delta X, \Delta Y) \quad (7.17)$$

where  $N(cS1, cS2, \mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$  represents a two dimensional normal distribution and  $U(cS1, cS2, x_1, x_2, y_1, y_2)$  a two dimensional uniform background. The data is shown together with the best fit in Figure 7.22. The fit itself was performed in a more limited data range which is shown in the appendix in Figure B.41 on page 268. The results of the best fit are summarized in Table 7.3. In total 9548 events were found within the 3 sigma contour of the ellipse<sup>6</sup>. No background events were found within these contours using 25 h of background data, making this selection background free for the analysis of the detection efficiency.

<sup>6</sup> $n_{sig}$  in Table 7.3 shows less events due to the limited fit range chosen.

Table 7.3. – Best fit results of the 4.44 MeV  $\gamma$ -ray ellipse in the TPC.

Parameter:	Value
$n_{sig}$	$(8950 \pm 90)$
$n_{bkg}$	$(40 \pm 8)$
$\mu_x$	$(31\,075 \pm 15)$ pe
$\mu_y$	$(2\,044\,300 \pm 2000)$ pe
$\sigma_x$	$(1445 \pm 11)$ pe
$\sigma_y$	$(188\,900 \pm 1500)$ pe
$\rho$	$(-0.736 \pm 0.005)$

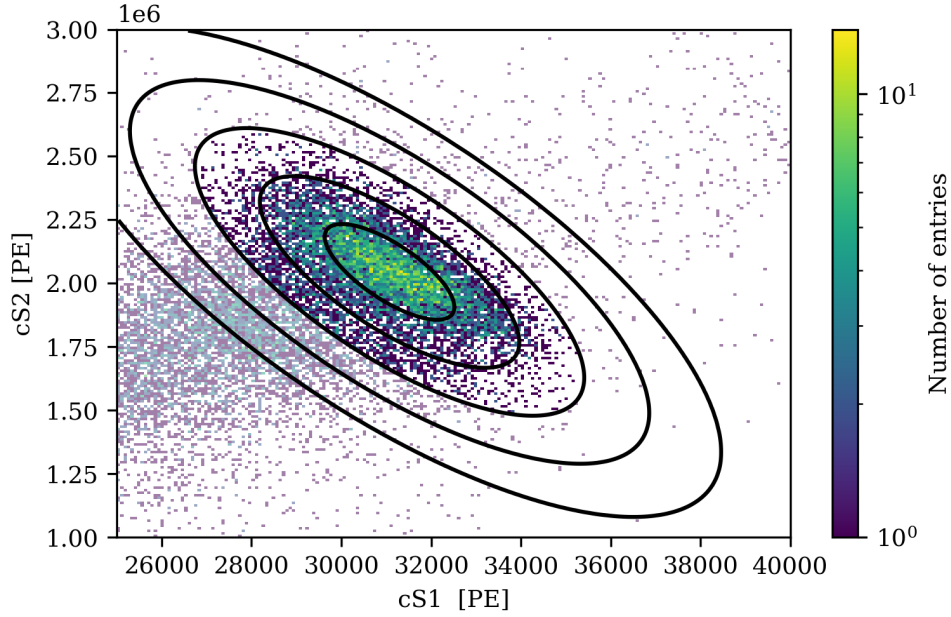


Figure 7.22. – Best fit of the 4.44 MeV  $\gamma$ -ellipse in the  $cS2$  versus  $cS1$  space. The two dimensional histogram shows a binned representation of the data. The ellipses indicates the 1 to 5 sigma contours of the best fit. Events outside of the 3 sigma contour are shaded.

It should be noticed that this selection does not only include events in the full energy peak of the 4.44 MeV-line, but also its Compton edge and continuum. Given that the flat background component cannot model well the Compton continuum the resulting p-value for the goodness of fit is zero. However, since this analysis only requires a clean sample of 4.44 MeV  $\gamma$ -events, but not a precise estimate of the full energy peak position, it does not play a significant role in the estimate of the detection efficiency.

After selecting the 4.44 MeV triggering events in the TPC the same analysis steps as for tagging efficiency are repeated. First, all coincident NV events in a time range between  $-1000 \mu\text{s}$  and  $2000 \mu\text{s}$  are selected. The resulting time spectrum is fitted with the model explained in equation (7.10). In contrast to the tagging efficiency study only the time range between  $-0.1 \mu\text{s}$  and  $0.1 \mu\text{s}$  was excluded from the fit since the 4.44 MeV  $\gamma$ -ray was observed by the TPC and no significant higher afterpulse rate is to be expected. The resulting best fit is shown in Figure 7.23 and the corresponding parameters of the best fit are given in Table 7.4. The time spectrum looks very similar

Table 7.4. – Best fit results of  $\Delta T_{pdf}$  for the NV tagging efficiency time distribution.

Parameter:	Value
$n_{sig}$	$(9570 \pm 140)$
$n_{bkg}$	$(20780 \pm 170)$
$\tau$	$(194 \pm 4) \mu\text{s}$
$\tau_{mod}$	$3e^{-05} \pm 1.88924 \cdot 3e^{-05} \mu\text{s}$



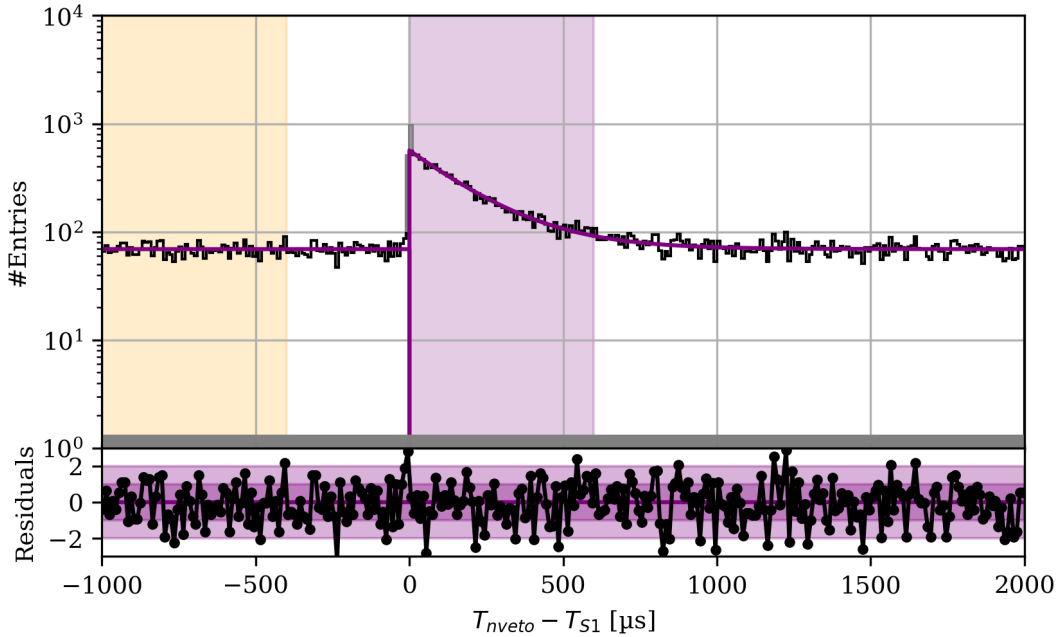


Figure 7.23. – Best fit of the time distribution for the NV detection efficiency. The main panel of the figure shows a binned and unbinned representation of the data as black histogram and vertical gray lines respectively. The region excluded from the fit is shown as gray bins in the histogram. The purple line shows the best fit as described in the text. The bottom panel shows the residuals of the fit. The color shaded regions indicate the one and two sigma contour of the best fit.

to Figure 7.17. Only the peak centered around zero is much smaller compared to the tagging efficiency study. This comes from the fact that no tight coincidence between neutron capture signals in the NV and the 4.44 MeV  $\gamma$ -ray in the TPC is expected. In total 809 events are found in the window  $[-0.1 \mu\text{s}, -0.1 \mu\text{s})$ . These events are partially due to the pile-up of two 4.44 MeV  $\gamma$ -rays which has an expected pile-up of about

$$R_{pile-up} = (R_n \cdot p_{n1})^2 \cdot \Delta T_{tight} \cdot \Omega \approx 147 \text{ evts} \quad (7.18)$$

where  $R_n$  is the neutron rate of the source  $p_{n1}$  the branching ratio for the first excited state of  $^{12}\text{C}$  in AmBe which was estimated to be about 50%,  $\Delta T_{tight}$  the size of the coincidence window which is 200 ns, and  $\Omega$  the solid angle for having a gamma in the NV which was assumed to be about 50%. The number of accidental coincidences between the two detectors for this small window can be inferred from the flat background component of the fit which gives an expectation of  $(1.39 \pm 0.01)$  evts. The remaining prompt coincidence might be caused by the second excited state of  $^{12}\text{C}$  which emits the additional 3.22 MeV electron via internal conversion as discussed in Section 7.1 and, prompt  $\gamma$ -rays emitted by the deexcitation of different isotopes like  $^{16}\text{O}$ . This effect will be further discussed in Section 7.5 of this chapter.

The overall goodness of fit was evaluated using a binned Poisson  $\chi^2$  as an estimator. The p-value was evaluated to be only 0.002 which rejects the hypothesis that the entire distribution is described well by the model. Restricting the range for the goodness of

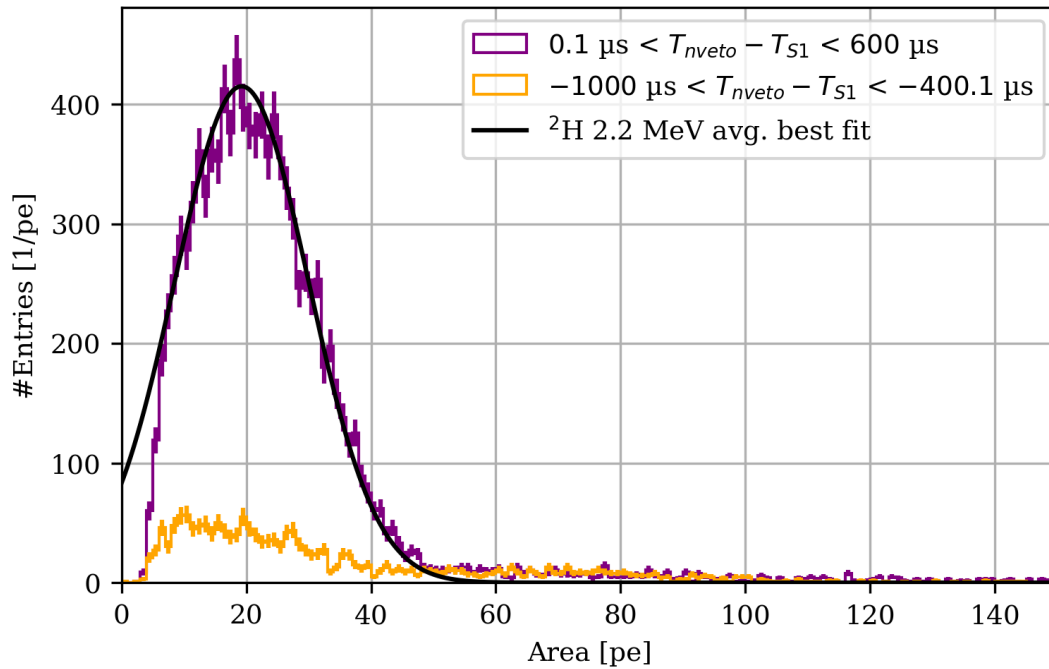


Figure 7.24. – Event area distribution for the signal and background region discussed in the text, shown in purple and orange respectively. The black line indicates the best fit of the 2.22 MeV hydrogen neutron capture line based on averaged best fit parameters of the individual fits discussed in Section 7.5. Its amplitude was scaled to match the shown data. Statistical uncertainties of the individual bins are shown as vertical lines.

fit test to the most relevant regions between  $[0.1 \mu\text{s}, 600 \mu\text{s})$  and  $[0.1 \mu\text{s}, 10 \mu\text{s})$  yields higher p-values of 0.17 and 0.14 respectively. Also, the residuals do not show any mismodeling in the exponential region, such that the results of the fit can be trusted. A zoom onto the two mentioned region between is shown in the appendix in Figure B.45 on page 271.

In contrast to Section 7.3, the time spectrum derived in Figure 7.23 has a much larger neutron capture time constant which is closer to the expected value for the neutron capture on hydrogen. This observation is under the hypothesis that the capture time constant is affected by the neutron capture on other materials. In contrast to the tagging efficiency it is not by definition require that the neutron passes through the TPC cryostat which makes a radiative neutron capture on other isotopes like  $^{56}\text{Fe}$  less likely. A dedicated study of this observation is presented in the subsequent Section 7.5. The moderation time preferred by the best fit is again compatible with zero.

As in Section 7.3 explained for the tagging efficiency, the detection efficiency can be determined by subtracting the event area distribution for events found in the exponential part of the distribution with events found in the orange reference region. For the detection efficiency the signal region was defined between  $0.1 \mu\text{s}$  and  $600 \mu\text{s}$  and the background region between  $-1000 \mu\text{s}$  and  $400.1 \mu\text{s}$ . The two corresponding area spectra are shown in Figure 7.24. The spectrum of the signal region shows a good

agreement with the average best fit of the hydrogen neutron capture derived in Section 7.5. Only at small event area the two distributions deviate from each other due to a 5-fold PMT coincidence threshold applied like in the tagging efficiency study. The excess around 7 pe and 8 pe as in Figure 7.16 cannot be observed.

Using the same strategy as for the tagging efficiency, the detection efficiency is estimated by integrating the background area distribution according to equation (7.16), without geometrical correction, and correcting for the corresponding detection window according to equation (7.14). The latter yields a smaller correction factor of less than a per mille for the 600  $\mu\text{s}$  benchmark window as the time window offset is just 100 ns. For the smaller 250  $\mu\text{s}$  window used in SR0 the correction factor is evaluated to be  $(75.96 \pm 0.01) \%$ .

The resulting detection efficiency for a 5-fold coincidence is shown in the appendix in Figure B.46 on page 272. A 5-fold coincidence threshold and a 5 pe event area threshold with an integration window of 250  $\mu\text{s}$  and 600  $\mu\text{s}$  leads to a neutron detection efficiency of  $(62.1 \pm 1.1) \%$  and  $(81.8 \pm 1.3) \%$ , respectively. Only statistical uncertainties are considered in this result which were propagated using Gaussian error propagation. To our knowledge, this is the highest neutron-detection efficiency ever achieved in a pure water Cerenkov detector.

Systematic uncertainties of the detection efficiency should be negligible. The only potential source of systematics is given by the fact that the number of emitted neutrons is counted via the number of triggering 4.44 MeV  $\gamma$ -rays in the TPC. While the number of false trigger signals by material backgrounds was found to be zero, delayed neutron capture signals on other materials like on  $^{56}\text{Fe}$  can in principle lead to energy deposits of similar magnitudes. However, it seems unlikely these isotopes produce energy deposits similar to those of  $^{12}\text{C}$ . Contamination by prompt signals from the de-excitation due to inelastic scatters from neutrons on other elements is negligible too. Further, would these signals also be valid trigger signals as the neutron can be still captured afterwards.

## 7.5 NEUTRON VETO SELF-COINCIDENCE

In this last section, the AmBe calibration data is discussed when dropping the coincidence requirement between TPC and NV, which has two advantages. Firstly, higher calibration statistics can be achieved as there is neither a passive layer of water nor LXe, particles are required to pass before producing a signal, and the loss of events due to pile-up is smaller since NV events are much shorter compared to  $\sim 2.2$  ms TPC events. Secondly, all four calibration positions summarized in Table 6.1 can be studied individually which is a direct consequence of the higher statistics. This allows to study different parameters as a function of radial distance to the TPC cryostat and thus, some hypotheses made earlier, for example, the dependence of the presence of other materials on the neutron capture time can be studied. Further, using the "TopCCW5d0m" source position allows to estimate the neutron capture cross section for thermal neutrons on hydrogen as the neutron capture on other materials should be reduced, given the distance to the TPC. The foundation of the studies presented in this section is still

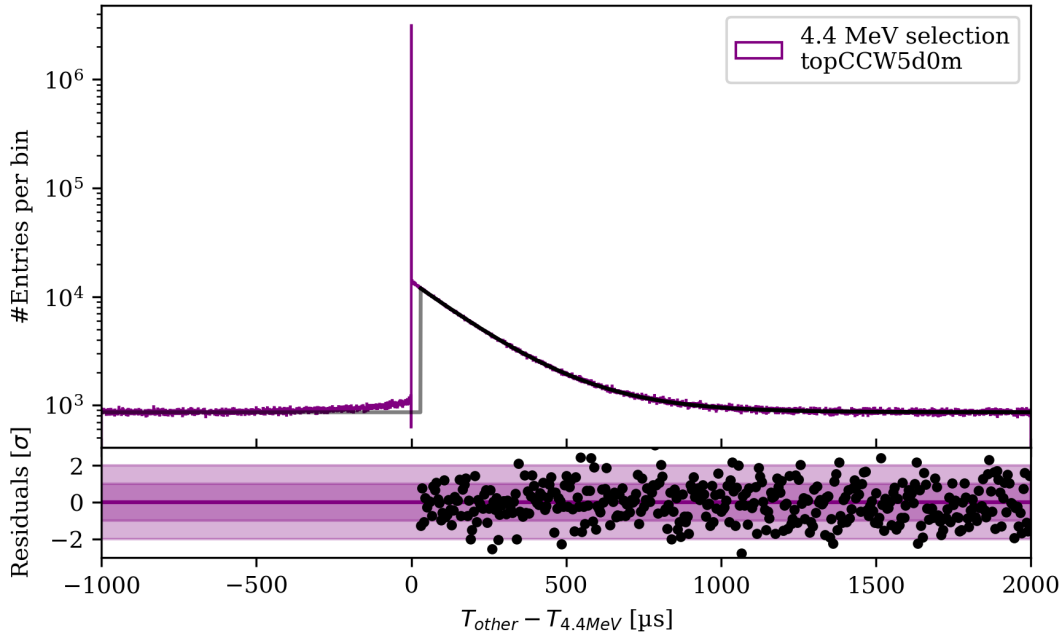


Figure 7.25. – Best fit of the NV self-coincidence time distribution using the 4.44 MeV  $\gamma$ -ray as the trigger signal for the TopCCW5d0m source position. The figure shows the same as explained in Figure 7.17. Only every 5th data point is plotted in the residuals.

the time correlation between neutron capture signals and the prompt 4.44 MeV  $\gamma$ -ray signal. This time correlation can be exploited in a NV self-coincidence such that a clean selection can be made.

### 7.5.1 4.44 MeV *gamma-ray trigger signal*

In this subsection, the 4.44 MeV  $\gamma$ -ray signal is used as a triggering signal for the NV self-coincidence. Triggering signals are selected by requiring their event area to be within  $\pm 1\sigma$  contour of the 4.44 MeV  $\gamma$ -ray signal inferred from the best fits derived in Section 7.1. The selection shows a slight variation depending on the source positron, with upper and lower boundaries around 45 pe and 85 pe respectively. In addition, it was required that at least 5 PMTs contribute to a triggering event. Afterwards, similarly as in previous studies, all coincident NV events in a window of  $[-1000 \mu\text{s}, 2000 \mu\text{s})$  were selected and the resulting time distribution estimated. Figure 7.25 shows exemplarily the resulting time distribution for the TopCCW5d0m position. The time distribution of the other three source positions can be found in the appendix on page 274. The shape of the time distribution is very similar when derived via TPC and NV coincidence. Only around zero and at small negative times the distribution deviates from the flat background observed before. The under fluctuation around zero is caused by the fact that the NV is operated in a self-coincidence and hence a minimal time gap is required to successfully resolve triggering and coincident events. A zoom onto the region around zero is shown in the appendix in Figure B.48 on page 273.

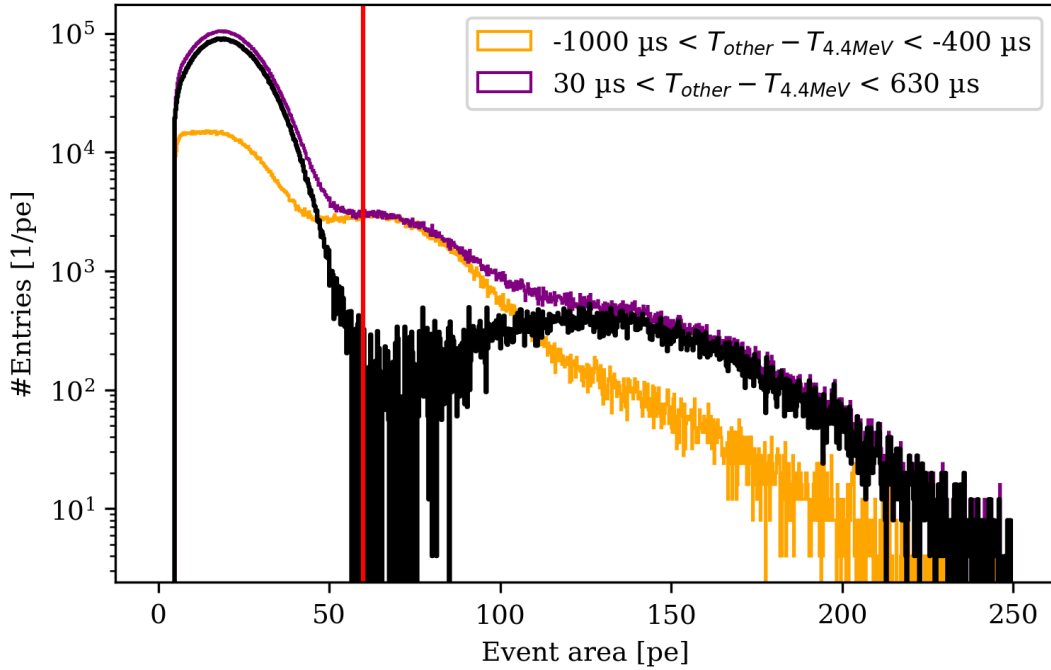


Figure 7.26. – NV energy spectrum for all events found in a NV self-coincidence during AmBe calibration at the topCCW5d0m source position. The plot shows the ROI of capture events in purple and the corresponding background reference region in orange. The bin wise background subtracted spectrum is shown in black. Uncertainties of the individual bins are not shown. The red vertical line is at 60 pe. The line divides the spectrum in neutron capture events and high energy tail events discussed in the text.

At negative time differences, the distribution shows an additional small exponential component which is mirrored around zero compared to the exponential part found at positive time differences. This observation can be explained by the fact, that the selection of triggering signals also contains a small fraction of neutron capture signals, and thus for these events the prompt 4.44 MeV  $\gamma$ -ray is found back with a negative time delay. These coincidences could be either caused by leakage of the 2.22 MeV neutron capture line into the area selection, which is  $\sim 0.5\%$  based on the best fits of Section 7.1, or by neutron capture signals on other isotopes which deposit a larger amount of energy in the NV. If the latter is the case these types of capture signals should also be found back when computing the background subtracted area distribution for events within the exponential distribution found in the positive half of Figure 7.25. The corresponding background subtracted energy spectrum is shown exemplarily for the topCCW5d0m source position in Figure 7.26. The spectrum shows a clear double peak structure which is dominated by a peak below 60 pe which corresponds to the 2.22 MeV neutron capture on hydrogen. The second peak is found at larger event areas beyond the 4.44 MeV  $\gamma$ -ray signal and corresponds to the high energy tail observed in Figure 7.2. In addition, the second peak shows an asymmetric shape with a longer tail to higher energies. Such an asymmetry would fit the picture that neutrons

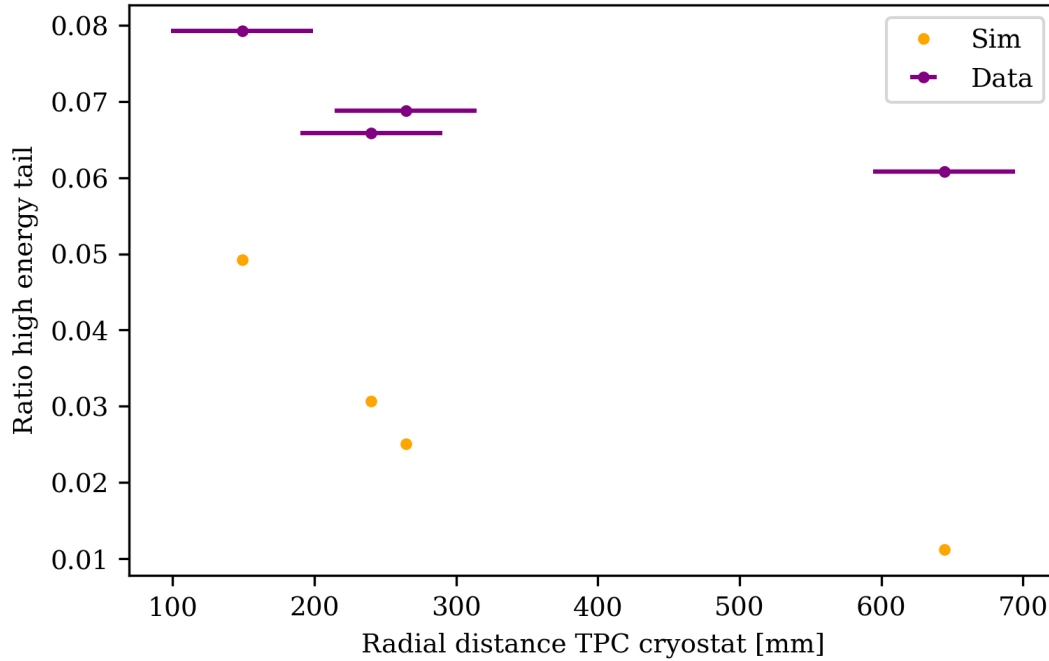


Figure 7.27. – Fraction of the high energy tail of the NV energy spectrum found during AmBe calibration as a function of radial distance of the source position towards the outer TPC cryostat. The purple points show the fraction derived from data, while orange points show the simulated fraction. The uncertainties of the fractions are not shown as the statistical uncertainty is too small. The simulated data does points show the MC truth for the radial distance.

are captured on  $^{56}\text{Fe}$  as the resulting 7.6 MeV excited state has a very branched level structure [Fir+17].

Dividing the spectrum in a low and high-energy region at the red vertical line allows to compute the fraction of high-energy tail events as a function of the various source positions. The fraction ranges from 6% to 8%, and is correlated with the radial distance of the source positions to TPC cryostat as shown in Figure 7.27. Comparing the ratios found in the NV self coincidence with the simulated ratios found in Section 7.2.1 shows a clear mismodeling in simulations. The fraction in data is much higher than in simulations and the deviation increases when moving further away from the cryostat. The latter indicates that either the current model of the u-tubes or the source housing is insufficiently modeled in simulations. The figure also indicates that the estimated source position based on the interpolated calibration points of the u-tube positioning might be wrong as the two bottom positions break the correlation. This effect is not observed in simulations that were modeled according to the interpolated calibration points.

Given that the energy distribution shows a much cleaner selection for the 2.22 MeV hydrogen neutron capture line the spectra for all four source positions were fitted in an event area range between 5 pe and 50 pe. The best fit was achieved using a normal distribution multiplied by the Fermi-Dirac-like threshold function in Equation (7.4). As optimization, a  $\chi^2$  method was used. The best fit for the topCCW5d0m

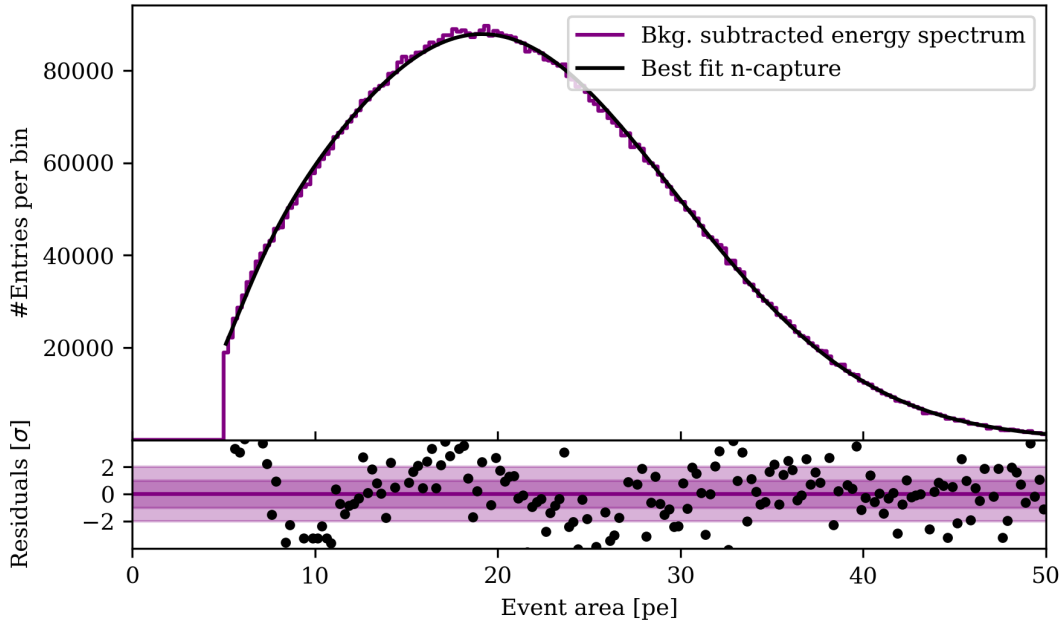


Figure 7.28. – Best fit of the hydrogen neutron capture peak for the topCCW5d0m source position derived from the NV only coincidence data. The top panel shows the data and best fit. The bottom panel the residuals of the fit.

position is shown in Figure 7.28. The remaining fits for the other source positions are shown in the appendix in Figure B.50 on page 275. As before the residuals of all four fits show a slight systematic mismatch between model and data. In addition the p-value computed based on the  $\chi^2$  for all four models is consistent with zero, rejecting that the model describes the data entirely which is again caused by the very high statistics of the data sample [LLS13]. Compared to the full energy spectra fits in Section 7.1 the fits prefer a slightly smaller mean for the hydrogen capture-peak of  $(19.25 \pm 0.18)$  pe on average compared to  $(19.86 \pm 0.17)$  pe found before. The width of the peak is comparable with the other fits. In general, the fits conducted in this Section should yield a better representation of the truth as the selection of the capture line is much cleaner without requiring any additional cuts which could potentially bias the shape and position of the peak.

Coming back to the time distribution of Figure 7.25, due to the high statistics the time distribution of each source position can be fitted, by using a  $\chi^2$  method with a  $1 \mu\text{s}$  binning. The fit range was limited between  $30 \mu\text{s}$  and  $2000 \mu\text{s}$ . A simplified representation of Equation (7.10) was used to fit the data. The model only uses a none-normalized version of an exponential distribution and a constant background model which are described by

$$\Delta T(t, t_0, \tau, A, c) = A \cdot \exp_{\text{pdf}}(t, t_0, \tau) + c \quad (7.19)$$

where  $\exp_{\text{pdf}}(t, t_0, \tau)$  is given by equation (7.11). The fits for the other source positions are shown in the appendix in Figure B.49 on page 274. The p-values of the



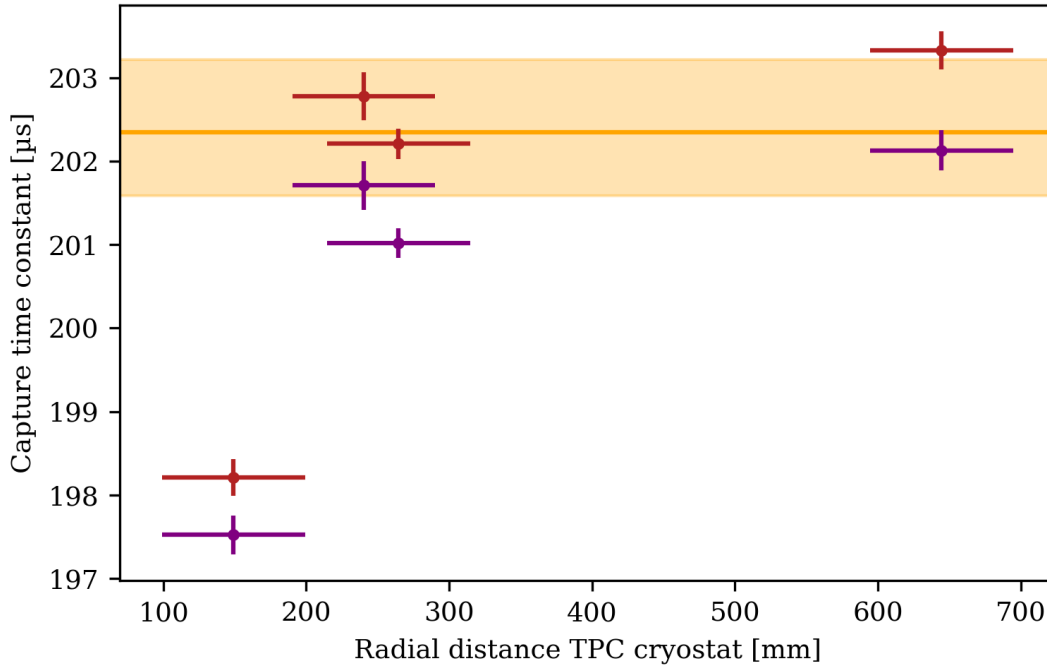


Figure 7.29. – Neutron capture constant for the four AmBe calibration source positions as a function of radial distance to the TPC cryostat. The purple points indicate the neutron capture cross section based on the best fit of the NV self-coincidence time distributions. The red points indicate the best fits when the high energy tail is excluded. The orange band indicates the best estimate of SNO experiment for the neutron capture time constant taken from [And+20].

individual fit range between 0.19 and 0.62 failing to reject the fit model hypothesis for all four distributions. Figure 7.29 shows the results of the individual fits with respect to the estimated neutron capture time constant by the SNO experiment. In addition, the neutron capture time constants when excluding all NV events with an event area larger 60 pe from the NV self-coincidence are shown. In both cases, the neutron capture time constant correlates with the distance to the TPC cryostat, but is always found to be larger than in the previous two studies presented in Section 7.3 and 7.4. The exclusion of the high energy tail leads to an additional increase in the capture time constant as expected, but the radial dependence remains.

### 7.5.2 Neutron capture on different components:

To better understand where the remaining radial dependents originate from, the neutron simulation described in Section 7.2.2 was revisited. The simulation conducted with Geant4 stores a variety of custom defined "saveflags" which allows to study in more detail on which detector component a neutron was captured. The currently implemented flags allow to distinguish if a neutron is either captured on, hydrogen, LXe, a material component in the TPC cryostat, or any other material in the simulation falling in neither of these categories. In addition, a saveflags which indicates inelas-

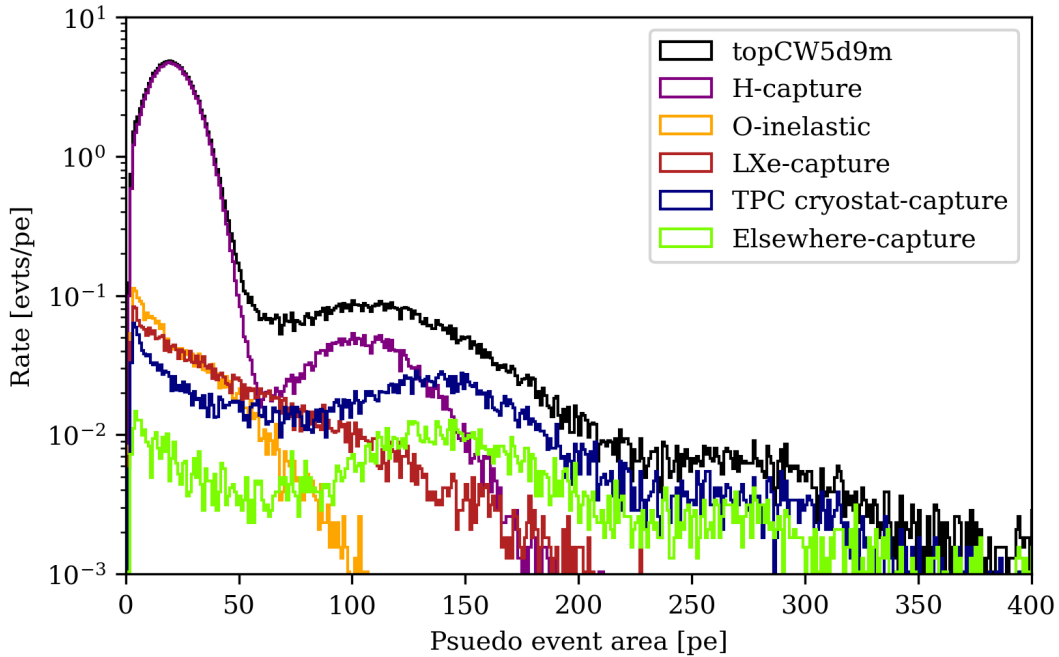


Figure 7.30. – Decomposed NV energy spectrum of simulated neutrons at the topCW5d9m source position, using custom made Geant4 saveflags.

tic neutron-scattering on oxygen is provided. Figure 7.30 shows the simulated energy spectrum, decomposed according to the specified saveflags, for the topCW5d9m source position which is closest to the TPC cryostat. The high energy tail is dominated by neutron capture signals on materials in the TPC and "other" components. At small energies the hydrogen capture peak is convoluted with captures on detector materials, LXe, as well as inelastic scattering on oxygen which leads to a prompt energy deposit in the NV. Completely unexpected is the double peak structure for events that are flagged as hydrogen capture. The second peak reaches its maximum  $\sim 110$  pe. The time distribution of these events, shown in the appendix in Figure B.51 on page 276, indicates that the second peak originates from a prompt  $X(n, n')X^*$  reaction releasing a  $\gamma$ -ray in the order of  $\sim 6.7$  MeV according to the simulated energy response of Figure 7.13. After this process, the scattered neutrons are captured on hydrogen falsely tagging also the prompt interaction as "hydrogen capture". A natural candidate for the inelastic neutron scatter is  $^{16}\text{O}$  which has an excited state at 6.92 MeV [Bro22]. Unfortunately, since this study was conducted only at the very end of this thesis, this prompt energy deposit was not considered during the evaluation of the tagging in Section 7.3. Thus, as this prompt  $\gamma$ -ray has sufficient energy to also lead to a valid NV tag, the tagging efficiency is slightly underestimated. Given this simulation, the fraction of prompt signals which occur within the very first 150 ns after emitting a neutron, is  $\sim 7\%$  of which  $\sim 3\%$  are not followed by a delayed neutron capture signal recorded by the NV. It is currently unclear how this fraction compares with neutrons emitted by detector materials, as only neutrons with sufficient energy can undergo this process.

The topCCW5d0m source position, shown in the appendix in Figure B.52 on page 277, only has capture signals flagged as hydrogen and "other" materials capture, as well as oxygen inelastic scattering. The capture signals produced by other materials originate from the calibration u-tubes as well as the source casing. Thus, the stronger drop in the neutron capture time constant for the topCW5d9m source position compared to the other source positions is explained by the additional material components. In addition, these simulations show that also the topCCW5d0m source position which is farthest away from the TPC cryostat is polluted with neutron capture signals on other materials. This must be taken into account when determining the thermal neutron capture cross section which is discussed in the following subsection.

### 7.5.3 Thermal hydrogen neutron capture cross section:

As discussed in Section 2.3, the neutron capture decay constant of water is directly proportional to the thermal neutron capture cross section in the case of an infinite and homogeneous target. Thus, to get an unbiased estimator for the thermal neutron capture cross section it is advantageous to only use the topCCW5d0m calibration source position as it is the least influenced by other materials compared to the other three calibration positions. In addition, it is required to correct the measured capture cross section from the remaining influences of other materials. For this reason, an additional simulation was conducted in which the calibration u-tubes, as well as the source casing, were removed. The corresponding energy spectrum is shown in the appendix in Figure B.52 on page 277. The time spectra of both simulations, with and without additional material, have been fitted as described in Section 7.5.1. Both fits are shown in the appendix too, in Figure B.53. The p-values of both fits are found to be  $> 0.1$  and thus failed to reject the fit model. For the water-only case, a  $\sim 0.6\%$  larger capture time constant was found with  $(200.61 \pm 0.30) \mu\text{s}$ , while the simulations including materials resulted in a capture time of  $(199.65 \pm 0.21) \mu\text{s}$ . However, it should be noticed that this relative difference of  $0.6\%$  is most likely underestimated with respect to data as the study of the high energy tail of the AmBe energy spectrum, shown in Figure 7.27, indicated a  $\sim 5.5$  times larger materials contribution compared to simulations. Thus, the  $0.6\%$  are only used as a systematic uncertainty rather than a correction, and the thermal neutron capture time of hydrogen for the topCCW5d0m is evaluated to be  $(202.1 \pm 0.2 \text{ stat. } \overset{+1.2}{-0.00} \text{ sys.}) \mu\text{s}$ . Based on Equation (2.28) the neutron capture constant can then be estimated to

$$\sigma_{\text{H},n} = \frac{1}{\tau \cdot n_{\text{H}} \cdot \bar{v}_n} = 336.7 \pm 0.4 \text{ (stat.) } \overset{+2.0}{-0.0} \text{ (sys.) mb}$$

assuming the standard thermal neutron velocity of  $\bar{v}_n = 2200 \text{ m/s}$  and a water density of  $(999.04 \pm 0.16) \text{ kg/m}^3$ . The water density was determined based on the average water temperature measured at the water tank outlet before the water purification plant. It was assumed that the water inside the tank may vary by  $\pm 1 \text{ }^\circ\text{C}$  compared to the water outlet, resulting in a temperature of  $(15.5 \pm 1.0) \text{ }^\circ\text{C}$ . All uncertainties were propagated using Gaussian error propagation. The value found here is in agreement

with the value reported by the SNO collaboration of  $336.3^{+1.2}_{-1.5}$  mb [And+20] and is a bit larger than the value of  $332.6 \pm 0.7$  mb reported in [CM77].

The Geant4 physics list implemented in the simulations for this study uses the neutron capture cross section from the JEFF-3.3 library which is shown in Figure 2.12 on page 37. It has a value of  $(332 \pm 8)$  mb at a thermal neutron velocity of 2200 m/s, closer to the value reported in [CM77].

Additional systematic effects, e.g. due to the neutron capture on oxygen, or the neutron energy of the source which is on average above the  $\sigma_{H,n} \propto 1/v$  requirement of Equation (2.28) have not been considered. Further, the mismatch of the high energy tail of the simulated AmBe energy spectrum with respect to data has not been accounted for, which leads to an underestimate of the systematic uncertainty. A more precise implementation of the calibration u-tubes and source container is required to fully account for this effect.

## 7.6 CONCLUSION AND OUTLOOK

Chapter 7 discusses the calibration of the neutron veto using neutrons and 4.44 MeV  $\gamma$ -rays emitted by AmBe, as done for the NR calibration of the TPC. An empirical model was developed in order to describe the NV energy response to AmBe. It allowed to extract the peak characteristics of the two most dominant full energy peaks of the spectrum given by the 2.22 MeV hydrogen neutron capture line and the 4.44 MeV  $\gamma$ -ray which are used as important inputs for other studies like the NR calibration. The empirical model describes the main components of the AmBe spectrum well, but due to the very high calibration statistics, the data deems that also second order processes, caused by neutron capture on other isotopes as well as radiative inelastic neutron-scattering must be taken into account. Thus the analytical model was augmented with a simulation driven template fit using templates for the most important particles emitted by AmBe. The templates were generated using Geant4 and a custom effective simulation framework for the NV PMT response developed for this work. The template fits were used to extract the branching ratio of the first excited state of  ${}^9\text{Be}(\alpha, n){}^{12}\text{C}^*$  which was determined to be  $0.491 \pm 0.002(\text{stats.})^{+0.030}_{-0.050}(\text{sys.})$ . This value is smaller compared to other values typically found in literature which are  $\sim 60\%$  [Sch+15]. The ratio is a source dependent quantity as it depends on the distribution of  $\alpha$  energies before being capture by the  ${}^9\text{Be}$  target. In addition, the simulation model was used to predict the energy response of the neutron veto which needs to be confirmed by calibrations with additional mono-energetic calibration sources.

Further, the chapter discusses the calibration of the NV tagging and detection efficiency of neutrons. The calibration is based on a similar coincidence technique as for the calibration of the NR response of the TPC. The tagging efficiency, is found to be  $(0.531 \pm 0.028)$  for a 5-fold PMT coincidence and a 5 pe event area threshold with a tagging window of 250  $\mu\text{s}$  and a resulting lifetime loss of 1.6%. In comparison, the SK collaboration reported for their water Cherenkov detector a neutron tagging efficiency of  $17.74 \pm 0.04_{\text{stat.}} \pm 1.05_{\text{sys.}}$  which is about a factor 3 smaller [Zha+16]. However, a study conducted only after SR0 showed that this estimate might be slightly biased

towards smaller values, as prompt signals from a radiative inelastic neutron scatter processes have not been considered. The neutron detection efficiency is estimated to be  $(62.1 \pm 1.1) \%$  ( $(81.8 \pm 1.3) \%$ ) for a  $250 \mu\text{s}$  ( $600 \mu\text{s}$ ) capture time window and the same thresholds as used for the tagging efficiency. This is to our knowledge the highest neutron detection efficiency ever achieved in pure-water Cherenkov detector. The previously highest value was reported by the SNO collaboration with a value of  $\sim 50 \%$  [And+20].

Last but not least a neutron veto self coincidence between the 2.22 MeV and 4.44 MeV was used to determine the neutron capture time constant of thermal neutrons in water. A value of  $(202.1 \pm 0.2 \text{ stat. } ^{+1.2}_{-0.00} \text{ sys.}) \mu\text{s}$  is found which is comparable with values reported by other water Cherenkov experiments like  $202.35^{+0.87}_{-0.76} \mu\text{s}$  SNO [And+20] and  $(202.6 \pm 3.7) \mu\text{s}$  SK [Zha+16], but slightly smaller than the values ( $(203.6 \pm 0.3) \mu\text{s}$  and  $(204.7 \pm 0.4) \mu\text{s}$ ) reported by earlier experiments dedicated for this measurement [CWC65; CM77]. Based on this value an attempt was made to estimate the neutron capture cross section for thermal neutrons in water. The cross section was determined to  $336.7 \pm 0.4 \text{ (stat.) } ^{+2.0}_{-0.0} \text{ (sys.) mb}$  which is in agreement with the values found in literature and  $336.3^{+1.2}_{-1.5} \text{ mb}$  found by SNO [And+20].

The studies conducted in this chapter showed, that a further optimization of the simulation tools is required to improve the understanding of the neutron veto response. Especially, the mismodeling of the high energy tail of the simulated AmBe spectra needs to be investigated. On the one hand it needs to be understood why the fraction of high energy tail events in data exhibits the fraction found in simulations. On the other hand the currently implemented saveflags need to be extended such that it can be better understood which isotopes contribute to this tail.

Also additional calibration of the neutron veto will help to improve the understanding of the neutron veto response, and to better match simulation with data. Dedicated studies need to be conducted to better match the optical response of the neutron veto between simulation and data. The event center time of simulated events is slightly smaller compared to data pointing to a mismatch in water transparency, reflectivity of the ePTFE, or both. Dedicated studies, matching the simulated optical response of the NV with calibrations conducted with either the diffuser balls or the reflectivity monitor, might improve the simulated NV response. Also new calibrations with AmBe, closer to the TPC cryostat and at different positions around it, will not only improve the quality of the NR response calibration described in Chapter 6, but also help to understand better the spatial dependence of the NV tagging efficiency.

Besides all these little improvements, Chapter 7 showed, that AmBe is an excellent calibration source which allows to calibrate and study many aspects of the neutron veto response. Eventually, after SR1 the neutron veto will be doped with gadolinium-sulfate-ocatahydrate which will improve the NV tagging efficiency even further, by an overall much reduced livetime loss due to the smaller neutron capture time as well as the better signal-to-background ratio due to the larger energy deposit. Another way to reduce the livetime loss of the experiment would be an extension of the veto trigger conditions, e.g. by adding the center time cut discussed in Chapter 5 this would reduce the the amount of false triggers due to background signals by  $\sim 20 \%$  assuming the same NV background conditions as in SR0.

---

## SCIENCE RUN ZERO WIMP RESULTS

In this chapter the SR0 WIMP dark matter results are summarized which were also published as part of the presented work in [Apr+23b]. The results are based on the work discussed in previous chapters, and additional studies by other XENONnT analysts. The DM search was performed using a blinded analysis approach. This means that events in the region of interest (ROI) in cS1 and cS2 were not accessible to any analysts before finalizing and freezing the analysis strategy, e.g. fixing the ER/NR response and background models, or the veto efficiencies. Before unblinding the NR region in the WIMP ROI, a veto-tagged unblinding was performed in which all events tagged by one of the two vetos were unblinded.

In the first section of this Chapter, a summary of the SR0 science data is given. The summary is followed in Section 8.2 with a discussion of the veto-tagged unblinding and the validation of the neutron background model. The chapter closes with a discussion about the remaining background components, inference procedure as well as final SR0 WIMP DM results in Section 8.3. The conclusion of this chapter is postponed to the subsequent Chapter 9 where the entire presented work is concluded.

### 8.1 SR0 SCIENCE DATA SUMMARY

The first science run of XENONnT lasted from May 3<sup>rd</sup> 2021 till December 10<sup>th</sup> 2021. In this period 97.1 d of blinded science data was acquired from July 6<sup>th</sup> 2021 to November 10<sup>th</sup> 2021 together with periodically PMT calibrations as well as  $^{83\text{m}}\text{Kr}$  calibrations. The latter was used to correct S1 and S2 signals for a uniform detector response as explained in Chapter 2 as well as to monitor the overall detector stability. Before science data taking, the NR and ER response of the detector was calibrated using AmBe and  $^{220}\text{Rn}$  respectively. The energy response of the detector was calibrated using several different mono-energetic lines which are shown in the appendix in Figure B.54 on page 279. Based on the signals from these lines, corrected for detector response effects as well as reconstruction biases, the photon and electron gains of the TPC were found to be  $g_1 = (0.151 \pm 0.001)$  PE/photon and  $g_2 = (16.5 \pm 0.6)$  PE/electron. The respective calibration plot is also shown in the appendix in Figure B.55 on page 280.

As a short reminder, as discussed in Section 3.3.4, the majority of the blinded sci-



ence data was taken in a so-called detector "linked-mode" in which the data taking of TPC and the two vetos is synchronized. The average deadtime during linked data taking is about 2.3 % of which 0.05 % are due to the TPC DAQ itself, 0.7 % due to the muon-veto tagging and 1.6 % due to the neutron veto tagging. However, at the beginning of SR0, 8.7 d were taken in unlinked mode in which all three detectors operated independently. In this mode, the timestamps of the three detectors are not synchronized by the GPS-module of the DAQ, and thus a tagging of TPC events by one of the two vetos is by default not possible. As part of this work, a manual linking of these runs was performed using the GPS timestamps recorded by the acquisition monitors of the DAQs. In this way 3.9 d were recovered and manually linked. Given that in this period the three detectors were operated asynchronously, a larger deadtime needs to be taken into account in which one of the two veto systems was idle in between two runs. This results in a total deadtime fraction of 4.6 % for the manually linked data. The remaining data was kept as unlinked data in which no veto information is available. This needs to be considered in the tagging efficiency of NV when estimating the neutron background for SR0. The resulting deadtime corrected livetime of SR0 is 95.1 d which results in a total exposure of  $(1.09 \pm 0.03)$  t/y given the  $(4.18 \pm 0.13)$  t fiducial volume used in the WIMP DM search.

The ROI of the WIMP analysis is defined by the recoil spectra presented in Figure 1.3 of Chapter 1. The WIMP analysis is carried out in the cS1/cS2 data space which means that the recoil spectra as a function of recoil energy must be forward folded using the energy and NR response model of the TPC, which was inferred in Chapter 6. To contain most of the spectra for the various WIMP masses, the ROI is defined between [0 pe, 100 pe] in cS1 and  $[10^{2.1}$  pe,  $10^{4.1}$  pe] in cS2. In addition, to the constraints imposed by the ROI, the analysis further takes into account efficiency losses due to detection and selection effects through data quality cuts. The detection efficiency is dominated by the 3-fold coincidence requirement on S1 signals which was also used in the NR response calibration. The efficiency was determined using simulated waveforms as well as fake S1 signals using a data-driven approach<sup>1</sup>.

The selection efficiency loss represents the overall loss of good events through data quality cuts. The selection efficiency was partially determined using a data-driven method, and partially by using simulated waveforms as explained in Section 6.2.3. The latter is especially the case for the S2 shape-based data quality cuts as well as for the data quality cuts acting against the AC background discussed in Section 8.3. S2 signals reconstructed in a 8.9 cm wide range around the transverse wires of the gate and anode mesh suffer due to local variations of the drift and extraction fields, a distortion of their peak shape. This distortion could not be sufficiently modeled in waveform simulations and thus, data quality selections for this region acting on S2 shape are less strict and are optimized based on calibration data instead. However, this has the consequence that the cut efficiency in the near and far regions are slightly different with a 10 % relative higher overall acceptance in the near transverse wire region. This comes at the cost of a higher background rate which is discussed in Section 8.3.2.

<sup>1</sup>In this method waveforms for small S1 signals are artificially created by bootstrapping single PMT hits from larger S1 signals using calibration data like  $^{37}\text{Ar}$  and  $^{83m}\text{Kr}$ . The artificial signals are afterwards fed again into the straxen processing framework.



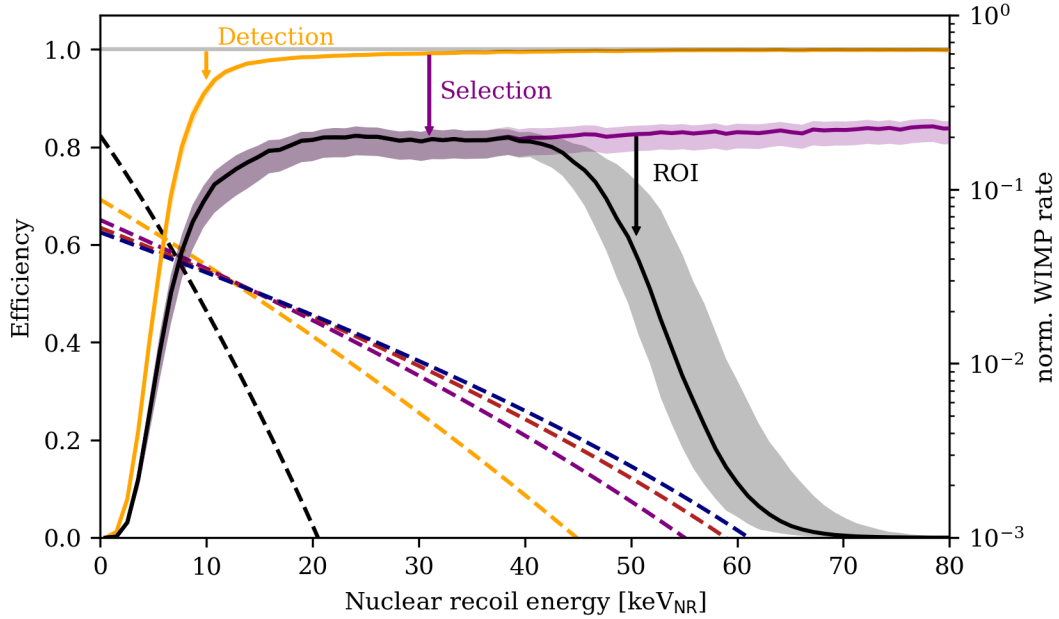


Figure 8.1. – ROI, detection and selection efficiency of the SR0 WIMP analysis. The total efficiency for the SR0 WIMP analysis is shown in black. The shaded region indicate the uncertainties of the corresponding efficiency terms. The colored dashed line show the normalized differential recoil spectra from Figure 1.3. The colors represent the following WIMP masses: 20 GeV/c<sup>2</sup>, 50 GeV/c<sup>2</sup>, 100 GeV/c<sup>2</sup>, 200 GeV/c<sup>2</sup> and 1000 GeV/c<sup>2</sup>, in black, orange, purple, red and blue respectively.

The combined detector efficiency losses averaged over the whole detector volume, are shown together with the WIMP recoil spectra in Figure 8.1. The WIMP analysis has an efficiency of at least 10 % in an energy range between 3.3 keV<sub>NR</sub> and 60.5 keV<sub>NR</sub>.

The blinding region for the science data was defined based on the ROI of the WIMP analysis, and an earlier version of the NR sliced band fitting procedure discussed in Section 6.3 as well as a sliced fit of the ER band. All events between  $\pm 2\sigma$  quantiles in cS2 of the ER and NR band below an energy of 20 keV<sub>ER</sub> were blinded. The ER data was partially unblinded before the NR data unblinding as discussed in [Apr+22c]. During this unblinding also all NR events above 10 keV<sub>ER</sub> were unblinded which corresponds to a nuclear recoil energy of about 50 keV<sub>NR</sub>. Before unblinding the remaining events a veto-tagged unblinding was performed which is discussed in the following section.

## 8.2 NEUTRON VETO TAGGED UNBLINDING AND NEUTRON-BACKGROUND

The main purpose of the veto tagged unblinding is the validation of the absolute neutron rate for the neutron background model in XENONnT which was determined similarly as described in [Apr+19b; Apr+20b]. The neutron background is mainly driven by radiogenic neutrons originating from spontaneous fission and ( $\alpha$ ,n)-reactions in the various detector materials. Hence, the neutron background model can be determined

based on the results from the TPC material screening conducted before the construction of the detector. Especially, isotopes of the uranium and thorium chain contribute to the neutron background. The neutron yields of the different materials and isotopes are evaluated using SOURCES-4A [Wil+99]. Templates of the different neutron backgrounds were simulated using the same full-chain simulation technique as used for the selection efficiency estimate. The estimates of the neutron yields and the simulations of the background templates were performed by various analysts. For SR0 a neutron single scatter background of  $0.179^{+0.061}_{-0.061}(\text{syst.})^{+0.009}_{-0.009}(\text{stats.})$  evts after applying all data-quality cuts as well as the neutron veto tagging, is predicted. The systematic uncertainties mainly originate from the uncertainties of the material screening. The neutron-background templates<sup>2</sup> deduced from these simulations are shown in the appendix in Figure B.56 on page 281. As an additional parameter to validate simulations, the multi-to-single scatter ratio of neutron interactions was determined. It was found to be 2.55 for the given data quality cuts used in SR0, which was confirmed by the AmBe calibration data.

Figure 8.2 shows the SR0 WIMP ROI, after performing the veto tagged unblinding, in which only events which are tagged by one of the vetos were unblinded. The plot shows all events after applying all other data quality cuts<sup>3</sup>. In total 8 (2) events were tagged by the neutron veto (muon veto) of which 5 (2) events are found inside the NR blinded region. Three of the events tagged by the NV inside the NR blinded region are marked as multi-scatter events by the S2 single scatter cut. Below and above the NR blinded region three events are tagged by the neutron veto. The latter is compatible with the expected number of accidentally tagged events which is given by the total number of 157 events found in Figure 8.2 and the exposure loss due to the neutron veto tagging of 1.6%. This leads to a predicted number of accidental coincidences of  $(2.5 \pm 0.2)$  events.

One of the two events tagged by the muon-veto is tagged by the neutron veto, too. This event deposited a total event area of 300.000 pe and 52.000 pe in the muon- and neutron veto respectively, suggesting that this event was indeed caused by a muon crossing the water tank. The other event deposited only an event area of 15 pe in the muon-veto which makes it more likely to be an accidental coincidence between TPC and muon-veto due to some background radiation. This would be in agreement with the expected number of accidental coincidences of  $(1.092 \pm 0.087)$  events. The expected number of muon-induced events inside the TPC was estimated by Andrea Molinaro from Turin University based on [Apr+14]. The expected rate was found to be  $3 \times 10^{-2}$  evts/ty for the larger TPC in XENONnT. Computing the chance of finding a single event for the given  $(1.09 \pm 0.03)$  t/y exposure via Poisson statistics gives a chance of  $P_{0.032}(1) = 0.0310$ . This is a very small chance, but since only a single muon event was found it was decided to not change the estimated muon-induced neutron

<sup>2</sup>In general is the neutron background divided into two components, regular neutron events, and so-called "neutron+x" events. The latter are signals in which a neutron scatters first in a region where no S2 signal can be produced e.g. below the cathode before a second signal is produced in the charge-sensitive region of the detector. As the neutron time of flight is in general too short to resolve the two S1 signals, neutron-x events are shifted towards smaller S1/S2 ratios with respect to regular neutron signals.

<sup>3</sup>Except for the S2 multi-scatter cut in case of veto-tagged multi-scatter events.

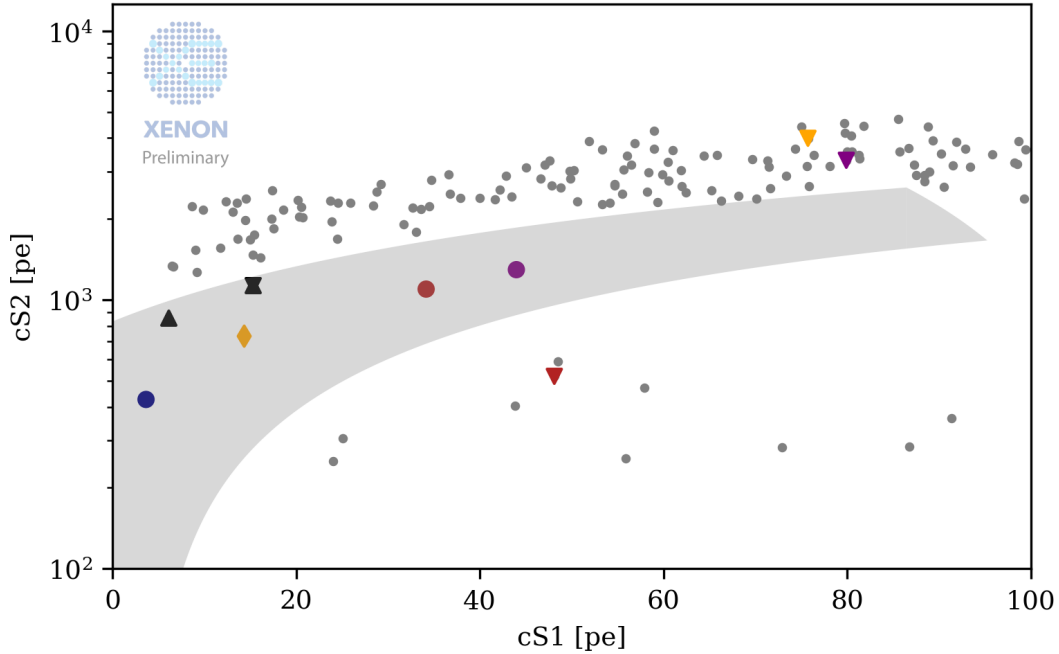


Figure 8.2. – SR0 unblinded veto tagged events. The plot shows all events found in the SR0 WIMP DM ROI passing all data quality cuts outside of the blinded region as gray dots. The gray shaded region indicates the NR blinded region. Events tagged by the vetos are highlighted as colored markers. Colored dots inside the blinded region indicate NV tagged multi-scatter events, while the diamond indicates a NV tagged single-scatter event. The colored inverted triangles show all neutron veto tagged events outside of the blinded region. The black regular triangles indicate two events tagged by the muon-veto. The event reported as a black regular and inverted triangle is tagged by both muon and neutron veto. The color code of the colored events inside and outside of the blinded region encode the event-id of the respective event for the subsequent Figure 8.3 and 8.4.

background which is estimated as an upper limit of 0.01 events for the SR0 exposure after applying the MV tagging. A longer exposure will show if the seen event was just a statistical fluke or whether the muon rate must be re-evaluated.

The multi-to-single scatter ratio observed in the NV tagged unblinded data is in agreement with predictions from simulations. However, the overall neutron background rate is found to be significantly larger than predicted. Thus, to determine the likeliness that the tagged events are indeed neutron capture signals, additional parameter spaces were investigated. Figure 8.3 shows the spatial distribution of the tagged TPC events together with the reconstructed azimuthal angular position for the corresponding NV events. As an approximation for the confidence interval of the reconstructed angular position, the average central 68.27% coverage of the angular distribution for the four AmBe calibration source positions was used. This gives an

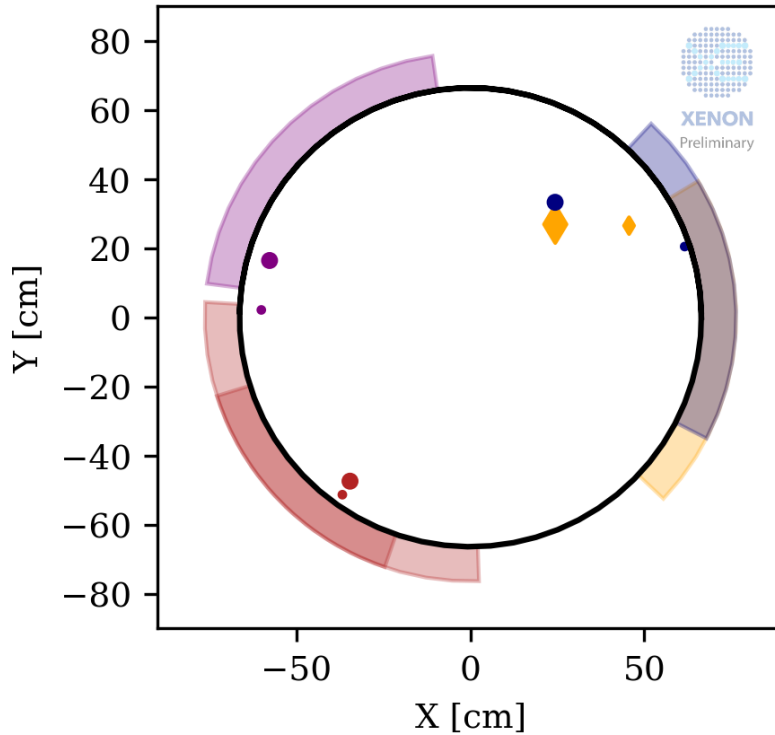


Figure 8.3. – Spatial distribution of the neutron veto tagged NR events in SR0. The markers indicate the reconstructed  $x/y$ -position of the main and alt. S2 signal, where the larger marker indicates the main signal. The color code and marker type represents the same events as in Figure 8.2. The colored shaded ring pieces indicate the reconstructed azimuthal angle position of the corresponding neutron veto events as explained in the text. The same color coding was used as for the corresponding TPC events. The event colored in red has two tagging neutron veto events which are both shown. The manually corrected neutron veto event position is shown in purple.

uncertainty of  $\pm 0.699$  rad ( $\pm 40.0^\circ$ )<sup>4</sup> on the reconstructed angle. The spatial position for one of the shown NV events was reconstructed manually, as a PMT dark count led to a wrong position reconstruction<sup>5</sup>. The alternative S2 signals were manually corrected for field distortion effects just as it is done for the main signals. After correcting for these effects all four tagged events show a plausible position reconstruction in TPC and neutron veto. The plot further shows, that the event passing the S2 single scatter cut might be a multi-scatter signal as well.

In addition to the spatial distribution, the NV event area as a function of the time

<sup>4</sup>This value is only an approximation for the position resolution as this interval determines only the confidence interval in which 68.27% of all AmBe events were reconstructed in the neutron veto. Hence, for a fixed source position, while here it is required to know the reconstructed position resolution for a given reconstructed position in the TPC. Thus, the interval used here might be biased but gives an idea of the confidence in the reconstructed position.

<sup>5</sup>This is one of the weaknesses of the current NV data structure as discussed in Chapter 3 as well as Chapter 5. If a PMT dark count precedes a regular event the position is reconstructed wrongly, as only the first 20 ns are used for the position reconstruction. An additional event level splitting would be required to remove preceding dark counts from true events.

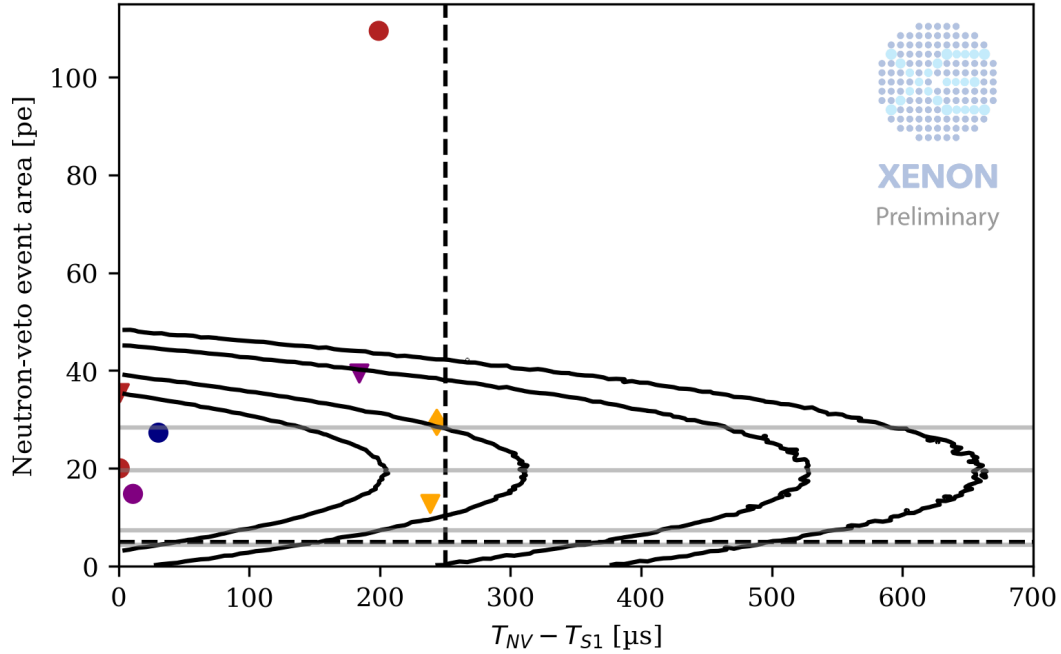


Figure 8.4. – Neutron veto events which tagged a SR0 TPC event. The plot shows the neutron veto event area and time delay between TPC S1 and neutron veto event for all neutron veto events which tagged a TPC event. The color code and marker type encodes the same event-ids as in Figure 8.4. The black dashed lines indicate the area threshold and tagging window length used in SR0. The black contours indicate the central 50 %, 70 %, 90 % and 95 % quantiles in which neutron capture events are expected to be found. The horizontal gray lines show the same, but for background events. The NV event which was also tagged by the MV has an event area beyond the plotted range.

delay between TPC and neutron veto signal can give further information about the likeliness that a tagged event is caused by a neutron. Figure 8.4 shows the distribution of all tagging NV events in the corresponding parameter space together with the highest density contours in one would expect neutron capture signals to be found. The contour regions were estimated numerically using an exponential distribution for the time delay, and a normal distribution for the area, together with the best-fit values for the neutron capture time and peak position as discussed in Chapter 7. All tagged TPC events found within the NR band show a high likeliness of being caused by neutron capture signals. One of the two NV events which tagged the red multi-scatter TPC event has a very large second energy deposit which is beyond 5 MeV according to the simulated energy response of the NV presented in Figure 7.13 on page 165. This points to a neutron emitting process which leaves one of the produced nuclei in an excited state.

The three events outside of the NR blinded region give a less clear picture. The red and purple events are found beyond the 95 % quantile in which one would expect NV background events to be found. The orange event is located between the 70 % and 90 % quantil. The quantiles were estimated based on the accumulative of the background distribution shown in Figure 7.1 on page 147. Especially, the red triangle event shows

a higher likeliness to be also caused by a neutron or a prompt  $\gamma$ -ray seen by both detectors. Given the location of the event in the cS1-cS2 data space the event might be either a neutron-x event or a wall background event discussed in the subsequent section. Also, the purple event shows a larger NV event area. Thus, this event might be an ER event caused by a  $\gamma$ -ray which deposited energy in both detectors.

Since, at least all TPC events within the blinded region, show a high likeliness to be induced by neutrons, it must be concluded that neutron predictions from simulations are wrong. Therefore, it was decided, according to the plan set before performing the veto tagged unblinding, to scale the absolute rate of expected neutrons according to the tagged events within the blinded NR region. Only the shape and the "neutrons-to-neutrons+x"-ratio of the simulated neutron background templates is used as input for the final WIMP inference. To determine the neutron background rate, the number of tagged single and multi-scatter events are combined in a single log-likelihood which is given by

$$\log L = \log |\text{Pos}(3, \mu_{\text{MS}}^{\text{NV}})| + \log |\text{Pos}(1, \mu_{\text{NV}}^{\text{SR0}})| \quad (8.1)$$

where Pos represents probability mass function of a Poisson distribution,  $\mu_{\text{MS}}^{\text{SR0}}$  and  $\mu_{\text{SS}}^{\text{SR0}}$  the expected number of NR multi- and single scatter events tagged by the neutron veto respectively. Both are given by

$$\begin{aligned} \mu_{\text{MS}}^{\text{NV}} &= \mu_{\text{SS}}^{\text{SR0}} \cdot R \cdot \zeta_{\text{eff}} \\ \mu_{\text{SS}}^{\text{NV}} &= \mu_{\text{SS}}^{\text{SR0}} \cdot \zeta_{\text{eff}} \end{aligned}$$

where  $\mu_{\text{SS}}^{\text{SR0}}$  represents the expected number of NR single scatter events before applying the neutron veto tagging,  $R$  the multi-to-single-scatter ratio, and  $\zeta_{\text{eff}}$  the effective neutron veto tagging efficiency considering the fraction of exposure in which a veto tagging was possible. For SR0 the expected number of NR single scatter events before neutron veto tagging is evaluated to be  $\mu_{\text{SS}}^{\text{SR0}} = 2.2_{-0.9}^{+1.3}$  events and  $1.1_{-0.5}^{+0.6}$  events after applying the neutron veto tagging.

It should be noticed, however, that due to a bug in the neutron veto tagging this correction of the neutron background rate was performed only after unblinding of the SR0 WIMP search data. The bug was caused by a wrong setting of the tagging window in the NV-TPC coincidence of the corresponding plugin. The window was first defined as a window ranging from the start of a NV event till  $250 \mu\text{s}$ , while due to the neutron capture delay the respective TPC S1 signal is expected before the NV neutron capture signal. Thus, the window should have been defined the other way around from  $-250 \mu\text{s}$  till the start of the NV event. Hence, with the wrong setting, it was only possible to tag TPC events by accident which led to a total number of zero tagged events in the NR blinded region. This was at first compatible with predictions from simulations, before correcting the mistake and scaling the neutron background as discussed. It should be also noticed that besides fixing the bug and scaling the absolute neutron background rate, nothing else in the analysis was changed because of this finding.

### 8.3 WIMP SEARCH RESULTS

In this last section, the remaining background components are discussed, and WIMP inference is explained before the final results of the SR0 WIMP analysis are presented. The underlying studies were conducted by other analysts and are not part of the presented work.

#### 8.3.1 Remaining background components

In total five different background components are considered in the analysis which are summarized together in Table 8.1. The NR background component can be divided into two parts: The neutron background which was already discussed in the previous section and a background component due to CE $\nu$ NS. The expected rate of CE $\nu$ NS is dominated by  $^8\text{B}$  solar neutrinos whose flux is constrained by measurements of the SNO experiment [Aha+13]. The main uncertainty for the expected number of events from CE $\nu$ NS originates from uncertainties in the detector response model as these signals have not been measured yet.

The most dominant background component in the WIMP ROI is the ER background. The ER background is dominated by beta decays of  $^{214}\text{Pb}$  which is a daughter of the xenon intrinsic  $^{222}\text{Rn}$  as discussed in Chapter 2. In addition, beta decays from  $^{85}\text{Kr}$ , electron-scattering of solar neutrinos, and radioactive decays from detector material contribute to the total ER background. Based on the XENONnT low ER search [Apr+22c] it can be concluded that the ER background is mostly flat in the WIMP ROI as also shown in the appendix in Figure B.58 on page 283. The absolute rate of the background is unconstrained in the final results fit, but the expected number of ER events is given by the number of events found above the blinded WIMP region. To

Table 8.1. – Number of expected events in the SR0 WIMP analysis. The table shows the number of nominal and the number of expected events for the best fit of a 200 GeV/ $c^2$  WIMP with a spin-independent cross section of  $3.22 \cdot 10^{-47} \text{ cm}^2$  in the WIMP ROI and signal-like region explained in the text. The different background components are explained in the text, too. The nominal column shows the number of expected events and their respective constraints if applicable.

	Nominal	Best Fit	
		ROI	Signal-like
ER	134	$135^{+12}_{-11}$	$0.86^{+0.08}_{-0.07}$
Neutrons	$1.1^{+0.6}_{-0.5}$	$1.1 \pm 0.4$	$0.42 \pm 0.17$
CE $\nu$ NS	$0.23 \pm 0.06$	$0.23 \pm 0.06$	$0.022 \pm 0.011$
AC	$4.3 \pm 0.2$	$4.32 \pm 0.15$	$0.366 \pm 0.013$
Surface	$14 \pm 3$	$12^{+0}_{-4}$	$0.35^{+0.01}_{-0.11}$
Total Background	154	$152 \pm 12$	$2.0 \pm 0.2$
WIMP	-	2.6	1.3
Observed	-	152	3



forward fold the flat ER energy spectrum into the cS1/cS2 data space the ER response model of the TPC was calibrated using  $\beta$ -decays of the isotope  $^{220}\text{Rn}$  as well as the mono-energetic 2.8 keV line from  $^{37}\text{Ar}$ . The ER response model is shown in appendix in Figure B.57 on page 282.

In addition, two artificial detector background components must be considered. The first component are so-called "wall" or "surface" events which leak into the fiducial volume and the WIMP ROI. It is believed that these events are caused by  $^{210}\text{Pb}$  plate out at the inner detector surfaces due to  $^{222}\text{Rn}$ . Given that no analytical response model exists, the surface background model is determined in a data-driven approach as in [Apr+19b]. The response model in cS1 and cS2 is derived using a kernel density estimator together with events that are erroneously reconstructed outside of the detector volume. The radial and z-profile of the surface model are determined using  $\alpha$  decays of  $^{210}\text{Po}$  which is a daughter of  $^{210}\text{Pb}$ . The surface background model in cS1 and cS2 is shown in the appendix in Figure B.59 on page 283. The absolute rate of the surface background events inside the blinded region is determined based on the number of events found in the WIMP ROI, but outside of the blinded region<sup>6</sup>. To reduce the contribution of the surface background a fiducial volume with a tighter radial constraint of  $R_{\text{max}} = 61.35$  cm was chosen with respect to the previous low ER search [Apr+22c].

The last detector background component are AC events. AC events are artificial events due to the random pairing of isolated S1 and S2 signals during event building. Isolated S1 signals mostly originate from fake peaks due to PMT dark counts, or misclassified single electron S2 signals, while isolated S2s are created either by photo-ionization or delayed electron extraction after large S2 signals. A data-driven model of AC events can be built by selecting isolated S1 and S2 signals which are more than one full drift length apart. Afterwards, these peaks are randomly paired and assigned to a random drift time within the maximum drifttime boundary. Based on this data-driven sample dedicated data quality cuts were developed to mitigate this background. These cuts are based on the spatial and time proximity to previous large S2 signal, as well as a gradient boosted decision tree (GBDT) cut which cuts events based on their shape, area, and depth relation. Due to the insufficient S2 shape model near the transverse wire region, the GBDT cut is only applied in the far wire region. The absolute rate of AC events is given by the rate of isolated S1 and S2 signal in the detector, and the rate in the near wire region is a factor  $\sim 6$  larger than in the far wire region due to the missing GBDT cut. The AC model is validated based on sideband studies using  $^{37}\text{Ar}$  and  $^{220}\text{Rn}$  calibration data. The AC template as well as its validation are shown in the appendix in Figure B.60 and B.61 on page 284, following

### 8.3.2 Unblinding and WIMP results

The unblinded WIMP search data is shown in Figure 8.5. In total 152 events are found in the ROI, 16 of which are found within the blinded region. At cS1 signals larger 50 pe, but especially between 50 pe and 75 pe, a higher rate of ER leakage is observed, than

<sup>6</sup>In this case outside of the blinded region means outside of the total ER and NR blinded region before the unblinding of the ER data in [Apr+22c].

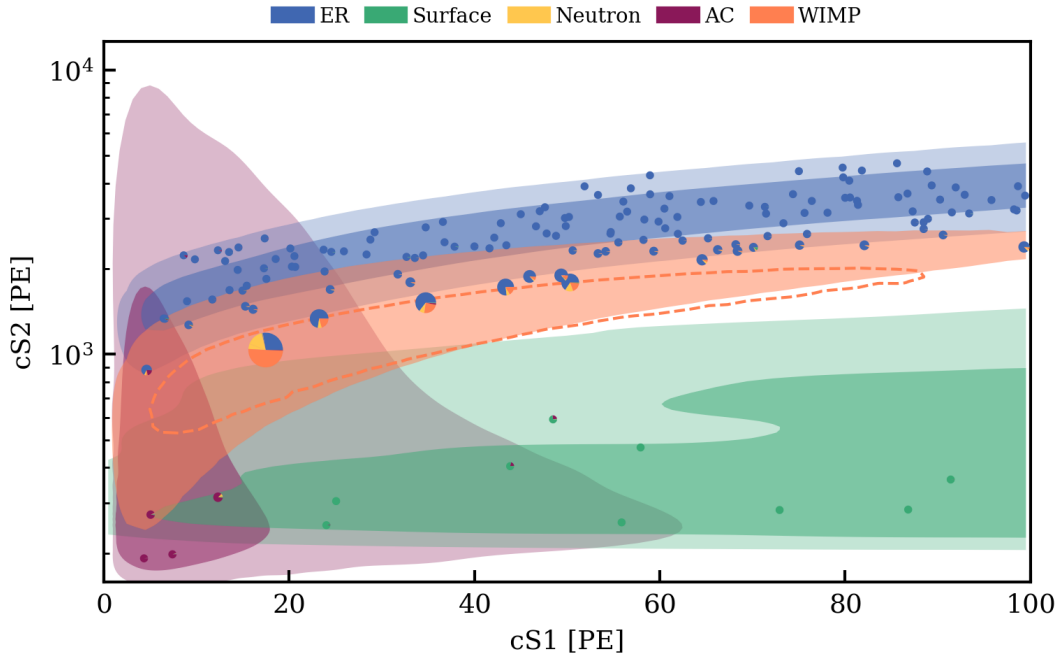


Figure 8.5. – Unblinded WIMP ROI for SR0. The different shaded regions show the one and two sigma contour for the corresponding background components. The yellow background component for neutrons is not shown as it mostly overlaps with the orange shaded region, which shows the 2 sigma signal contour for a  $200 \text{ GeV}/c^2$  WIMP. The orange dashed contour shows the region with the highest signal-to-noise ratio for a region which contains 50% of  $200 \text{ GeV}/c^2$  WIMPs. Each event in the WIMP ROI is depicted as a pie-chart which represents the fraction of the corresponding component evaluated at the given event position including a  $200 \text{ GeV}/c^2$  WIMP. The size of the pie-charts is proportional to the signal fraction at the given position.

expected by the model. Further, the unblinded events exhibit an asymmetry in their spatial xy-distribution as it can be seen in Figure 8.6. The majority of events are found in the "upper half" of the detector. Dedicated checks of calibration data, corrections, and data quality selections did not reveal any systematic mismodeling. The higher ER leakage could be explained by an incomplete ER response model. The model is primarily determined through the calibration with  $^{220}\text{Rn}$  which is a  $\beta$ -emitter. While the ER background is dominated by  $\beta$ -emission from  $^{222}\text{Rn}$  and  $^{85}\text{Kr}$ , it also constitutes other components like solar neutrinos or LL-shell double electron capture from  $^{124}\text{Xe}$ . These signals might exhibit a slightly different response which could explain the higher leakage. However, these signals would not explain the observed asymmetry in the spatial distribution of the unblinded data.

Given the observation in Section 8.2, that the neutron background in XENONnT is higher than expected, one might also wonder whether one of the 16 unblinded events might be due to an additional neutron interaction. Extending the TPC-NV coincidence window to  $600 \mu\text{s}$  and dropping the threshold requirements in the NV, does not show any indication for that. Only two coincidences between NV and TPC were found by this extended window and lower threshold. They are both likely to be accidental

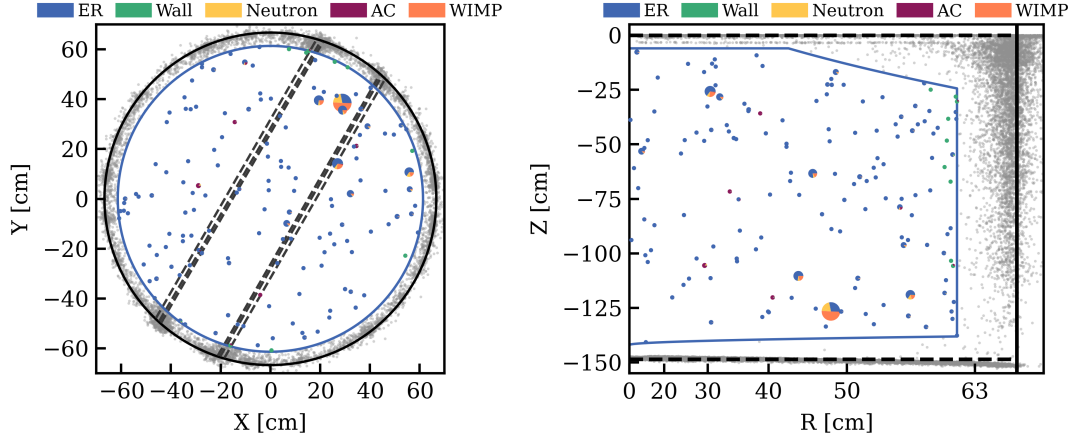


Figure 8.6. – Spatial distribution of the SR0 unblinded WIMP events. The upper plot shows the distribution in  $X$  and  $Y$ , while the lower plot shows the events in  $R$  and  $Z$ . The pie-charts are the same as in Figure 8.5. Events found outside of the fiducial volume are shown as gray dots. The boundaries of the fiducial volume are shown in blue, while the boundaries of the active volume given by the cathode, liquid-gas interface and TPC wall are drawn in black. In addition, the  $x/y$  plot also shows the position of the transverse wires as black dashed lines.

coincidences as the NV event area is below 2 pe for both events.

The corresponding best fit to the unblinded data does not show any significant excess and is consistent with the background-only hypothesis. The expected number of events of each background component expected by the best fit is within the corresponding derived constraints. The inference was performed similarly as in [Apr+19b], using a profile log-likelihood ratio as tests statistics which is given by

$$q(\sigma) = -2 \log \frac{\mathcal{L}(\sigma, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\sigma}, \hat{\theta})} \quad (8.2)$$

where  $\hat{\sigma}$  represents the signal strength,  $\hat{\theta}$  the nuisance parameters, and  $\hat{\hat{\theta}}$  the nuisance parameters which maximize the likelihood under the condition that the signal strength is  $\sigma$ . The science search likelihoods can be decomposed into four individual extended unbinned likelihood terms

$$\mathcal{L}(\sigma, \theta) = \mathcal{L}_{\text{sci, near}}(\sigma, \theta) + \mathcal{L}_{\text{sci, far}}(\sigma, \theta) + \mathcal{L}_{\text{cal}}(\theta) + \mathcal{L}_{\text{aux}}(\theta) \quad (8.3)$$

where  $\mathcal{L}_{\text{sci, near}}$  and  $\mathcal{L}_{\text{sci, far}}$  describe the science search likelihoods in the near and far wire region,  $\mathcal{L}_{\text{cal}}$  a likelihood for the ER calibration to constrain the ER model shape<sup>7</sup>, and  $\mathcal{L}_{\text{aux}}(\theta)$  an auxiliary likelihood term which constraints nuisance parameters like the expected background rates. The two science search likelihoods as well as the cali-

<sup>7</sup>The shape of the ER model is allowed to vary a bit to account for uncertainties in the ER calibration. To constrain these variations the ER calibration data is added in the likelihood.

bration likelihood are described by

$$\mathcal{L}_i = \text{Pos}(N_i | \mu_{i,tot}(\sigma, \theta)) \cdot \prod_j^{N_i} \cdot \sum_c \frac{\mu_c(\sigma, \theta)}{\mu_{i,tot}(\sigma, \theta)} \cdot f_c(\vec{x}_j | \sigma, \theta) \quad (8.4)$$

where  $N_i$  is the total number of events found in the corresponding dataset (near, far, calibration),  $\mu_{i,tot}$  the total expected number of events for the given dataset summed over the individual components  $c$ ,  $\mu_c(\sigma, \theta)$  the contribution from a given component<sup>8</sup>, and  $f_c(\vec{x}_j | \sigma, \theta)$  the probability density function of the different components<sup>9</sup> evaluated at the position  $\vec{x}_j$ . For the far wire region  $\vec{x}_j$  is given by cS1, cS2, and  $R$ , while for the near wire region and the calibration term, the radial dependents is dropped. The expectation value for the WIMP signal  $\mu_{sig}$ , for a given WIMP-nucleon interaction cross section  $\sigma_R$  and a given WIMP mass  $M$ , is expressed as

$$\mu_{sig} = \sigma_R \cdot \epsilon \cdot \mu_{ref}(M) / \sigma_{ref} \quad (8.5)$$

in terms of a reference expectation value  $\mu_{ref}(M)$  with a reference cross section of  $\sigma_{ref} = 10^{-45} \text{ cm}^2$ . In addition, the uncertainty of the signal rate due to uncertainties in the exposure and signal model is encoded via the scaling factor  $\epsilon$  which is constrained through a Gaussian term in the auxiliary likelihood.

The final sensitivity curve and limit are constructed using a Feldman-Cousins approach and toy MC realizations of the experiment. The limit is given at a 90% confidence level, whose threshold  $q_{th}(\sigma)$  is determined through toy MCs as a function of signal strengths. The median upper limit is determined as the median of a distribution of  $\sigma_{UL,i}$  from toy MCs without signal where  $q(\sigma = \sigma_{UL,i}) = q_{th}(\sigma)$ . The same distribution is also used to determine the  $\pm 1\sigma$  and  $\pm 2\sigma$  contours of the sensitivity band.

Figure 8.7 shows the sensitivity band together with the upper limit for the spin-independent WIMP-nucleon cross section derived from the best fit of the SR0 WIMP search data as a function of WIMP mass. To protect the limit against statistical downwards fluctuation the experiment has no sensitivity to, a power constraint [Cow+11] is applied. The constraint is defined based on "rejection power" and uses as a threshold a power of  $\beta_r = 0.5$ . Thus, the upper limit is constrained to the median sensitivity whenever a downward fluctuation in data is observed. This is much higher than suggested in [Bax+21b] which was followed in previous publications of the XENON collaboration. However, [Bax+21b] erroneously defines the power constraint based on "discovery power" instead of rejection power with a minimal discovery power of  $\beta_d = 0.32$ <sup>10</sup>. The very conservative choice of the power constraint for the first WIMP search in XENONnT is motivated by the fact that the authors of [Cow+11] stated that the rejection threshold should be much larger than the chosen significance which is  $\alpha = 0.1$ . Thus, it is not clear if a threshold of  $\beta_r = 0.16$ , which corresponds to the 1 sigma threshold as proposed in [Bax+21b], would satisfy this requirement. Further

<sup>8</sup>E.g. signal, background component or in case of the calibration likelihood just the calibration data

<sup>9</sup>E.g. the colored shaded regions in Figure 8.5.

<sup>10</sup>This corresponds to the one sigma contour of the sensitivity band in the asymptotic case.

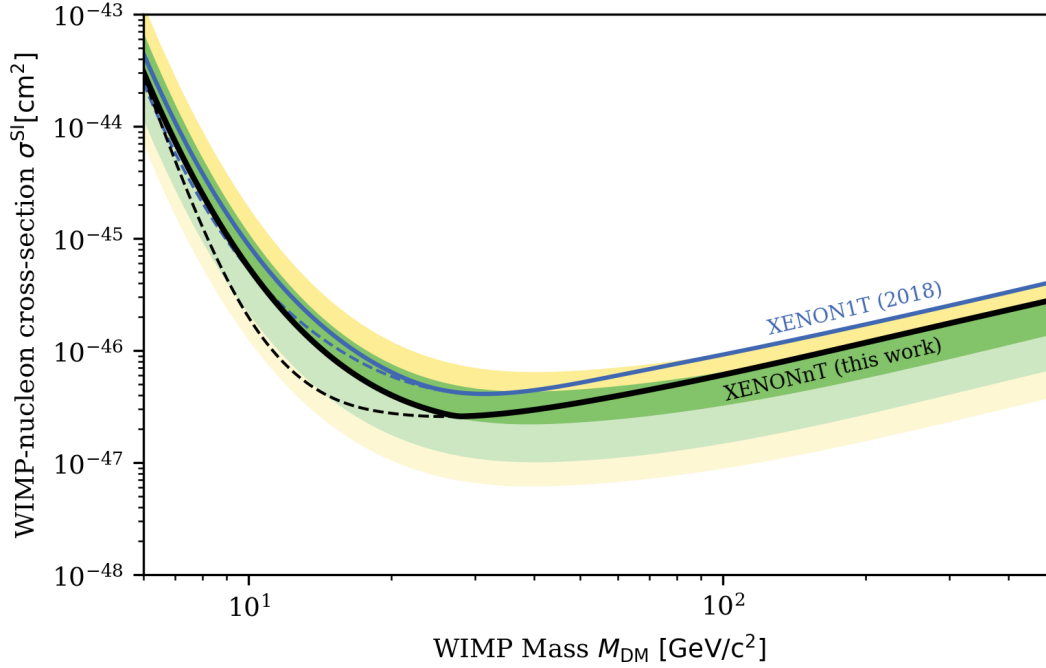


Figure 8.7. – Upper limit on the spin-independent WIMP-nucleon cross section at 90 % confidence level. The plot shows the results of XENONnT’s first WIMP DM search. The black solid line indicates the upper limit derived from the SR0 data. The dashed line indicates the same limit, but without power constraint. The yellow and green shaded region show the  $\pm 1\sigma$  and  $\pm 2\sigma$  contour of the sensitivity band around the median. As reference the previous best result of XENON1T [Apr+18] is shown in blue as a dashed line, and with the same power constraint as solid line.

discussion within the direct detection DM community is needed to agree on a common approach that makes results comparable.

Including the power constraint the lowest upper limit for spin-independent WIMP-nucleon interactions is  $2.58 \cdot 10^{-47} \text{ cm}^2$  at a WIMP mass of  $28 \text{ GeV}/c^2$  and a 90 % confidence level. At higher WIMP masses the limit scales with

$$\sigma_{\text{UL}}(M_{\text{DM}}) = 6.08 \cdot 10^{-47} \text{ cm}^2 \cdot (M_{\text{DM}} / (100 \text{ GeV}/c^2)). \quad (8.6)$$

This represents an improvement of about a factor of  $\times 1.7$  at  $100 \text{ GeV}/c^2$  compared to XENON1T, despite the extremely low drift and extraction fields, as well as the reconstruction issues induced by the transverse wires. The result is also compatible with the results of other LXe direct-detection DM experiments as it can be seen in Figure 8.8. The additional results for the spin-dependent WIMP-proton and WIMP-neutron couplings can be found in the Figure 8.9.

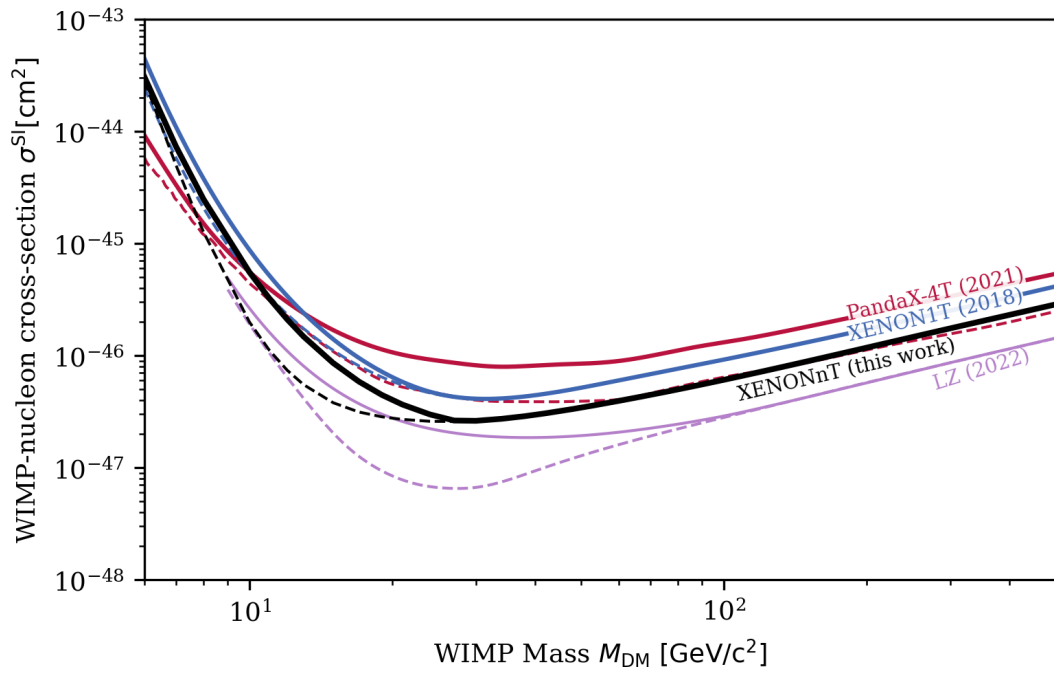


Figure 8.8. – Comparison of the best upper limit obtained in the first science run of XENONnT with the results from XENON1T [Apr+18], LZ [Aal+22] and PandaX-4T [Men+21]. The dashed lines correspond to their published results. The solid lines show their respective limits when the same power constraint as for XENONnT is applied.

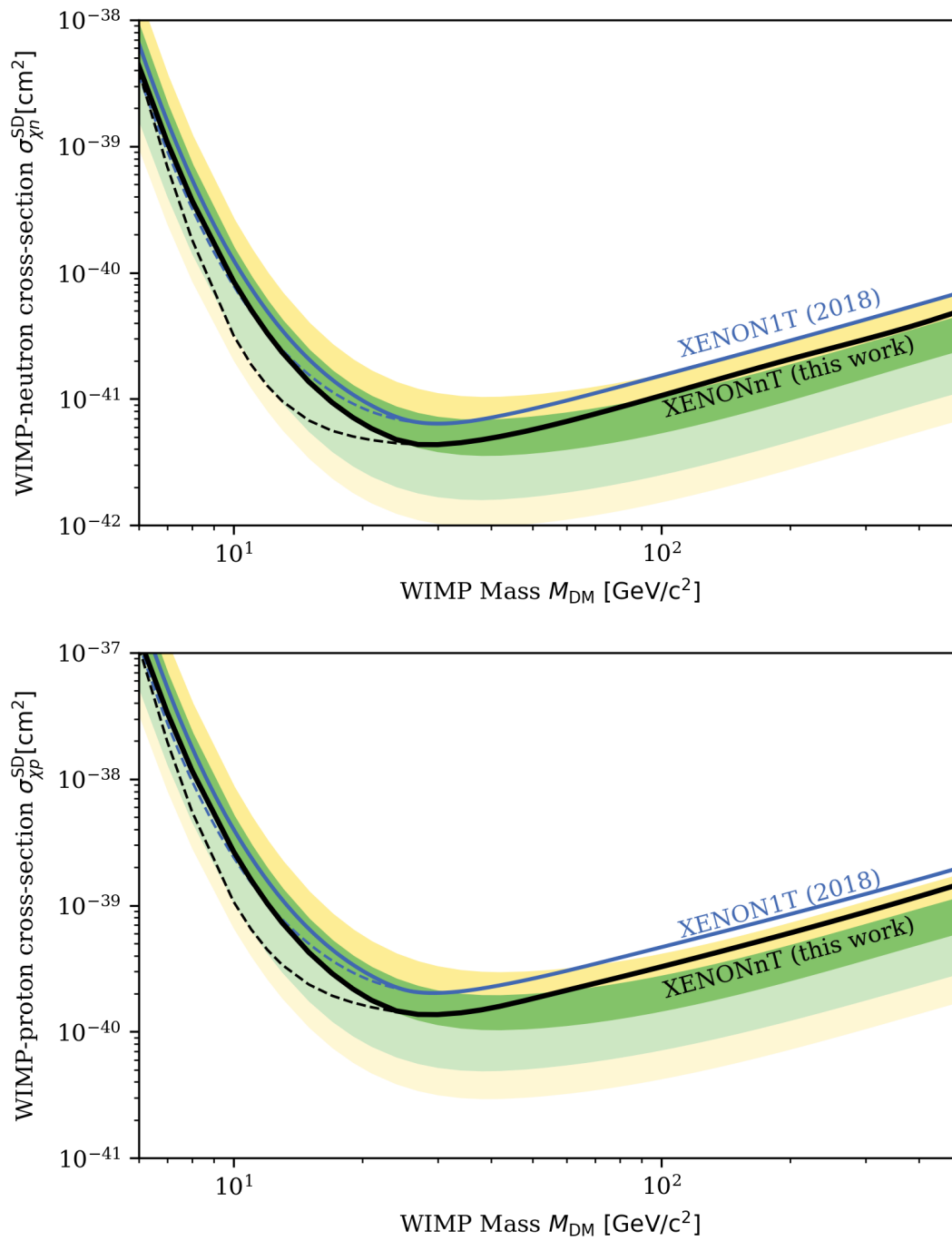


Figure 8.9. – Upper limits for the spin-dependent WIMP-neutron (top) and WIMP-proton cross section (bottom).



---

## CONCLUSION AND OUTLOOK

In this work, the first WIMP search results of the XENONnT experiment are presented. The results have been obtained in a blind analysis of 95.1 d of science data with a total exposure of  $(1.09 \pm 0.30)$  ty. The best fit to the data rejects the signal hypothesis and allows to set new stringent limits on the spin-independent as well as spin-dependent WIMP-nucleon interactions. The best-fit results in a lowest upper limit of  $2.58 \cdot 10^{-47}$  cm<sup>2</sup> for SI interactions of 28 GeV/c<sup>2</sup> WIMPs at a 90 % confidence level. The new results are consistent with results collected by other experiments [Aal+22; Men+21], and are an improvement of a factor of about 1.7 compared to XENON1T [Apr+18] for WIMPs with a mass of 100 GeV/c<sup>2</sup>. The presented results have been also published in [Apr+23b] as part of this thesis.

One key component to obtain the SR0 WIMP search results is the new neutron veto of XENONnT, the world's first water Cherenkov NV. Its efficiency to tag neutrons which produce WIMP-like signals through single-scatter NRs, was evaluated to be  $(53.1 \pm 2.8)$  %, using a backward veto of 250  $\mu$ s with a 5-fold PMT coincidence and 5 pe event area threshold. Both, the NR single scatter response model of the TPC, and the NV tagging efficiency, were determined using a novel calibration technique that was deployed for the first time in a LXe TPC. The calibration is based on neutrons and 4.44 MeV  $\gamma$ -rays which are coincidentally emitted by a  ${}^9\text{Be}(\alpha, n){}^{12}\text{C}^*$  neutron source. A tight coincidence with a window of 408 ns between the 4.44 MeV  $\gamma$ -ray recorded by the NV, and neutron signal recorded by the TPC, was used to obtain a clean calibration of the NR response model with a background leakage of  $<0.1$  %. Inverting the procedure, and tagging the 4.44 MeV  $\gamma$ -ray with the TPC allows to measure the neutron detection efficiency of the NV, which represents the chance of detecting a neutron after being emitted. The detection efficiency was evaluated to be  $(81.8 \pm 1.3)$  % using a 600  $\mu$ s forward coincidence between TPC and NV, which is to our knowledge the highest neutron detection efficiency ever achieved in a water Cherenkov detector. In addition, the signature of the AmBe source was used to measure the thermal neutron capture cross section of water by deploying a NV self-coincidence between  $\gamma$ -ray and neutron capture signal. The water capture cross section was evaluated to be  $336.7 \pm 0.4$  (stat.) $_{-0.0}^{+2.0}$  (sys.) mb which is consistent with the values found by other water Cherenov experiments like SK [Zha+16] and SNO [And+20]. The performance and results of the NV are also currently prepared for publication as part of the presented

thesis.

As part of the presented work, the processing chain of the NV was developed, too. The processing chain is based on a set of custom-made STRAX plugins, which reach a processing speed of about 100 MB/s using eight CPUs. This is a factor of about five larger than any NV data rate measured during SR0. Several other auxiliary software projects were realized over the course of the presented work, which are used by many analysts bringing a general benefit to the experiment.

Currently, XENONnT continues data taking with a reduced  $^{222}\text{Rn}$  background. In addition, the loading of the NV water with Gd-sulfate octahydrate is currently being prepared. Gadolinium will enhance the neutron tagging efficiency of the NV even further. At the moment it is still not decided whether the Gd-salt procured from the company Treibacher AG will be used in this initial loading, as the salt quality and purity are significantly below expectations compared to the laboratory samples, which were analyzed during the presented work. A replacement batch might be added at a later time. It also remains a bit unclear how the NR calibration might perform once the water is loaded. The higher gamma-multiplicity per neutron capture might require a replacement of the current AmBe source, with a source of a lower rate. Further, it might be required to adjust the TPC event reconstruction, such that multiple overlapping events can be found within the max drift time of the TPC. Otherwise small NR signals might be overshadowed by larger ER signals induced by the Gd neutron capture.

Despite this problem, the pioneering work and excellent performance of the NV presented in this thesis demonstrates that the water Cherenkov detector technology is well suited to be used as a NV also for future DM experiments like DARWIN [Aal+16], or other rare-event experiments. Paired with the novel calibration technique presented, it is a strong tool to calibrate the NR response in future science runs or experiments, down to a single photon S1 response and beyond. To realize such a calibration for the current XENONnT experiment a modification of the processing chain is required which is currently realized in cooperation with a PhD student of the WWU Münster. In addition, the coincidence technique will allow to perform for the first time an S2-only response calibration as it allows to substitute the missing timestamp of the S1 signal by the timestamp of NV gamma-tag. Such a calibration will deliver valuable input to validate the low energy S2-only signal model, and therefore for the scientific reach of XENONnT and future DM experiments. The same technique might be also used to measure for the very first time the Migdal effect in a LXe TPC in the near future.





---

## SUPPLEMENTAL MATERIAL

### A.1 THE XYZCAMSCAN INSPECTION SETUP

One of the most important parts of a TPC are the electrodes, which define the potentials for the electric drift and extraction field. In the past many of the large-scale detectors struggled to reach their designed fields [Ake+17b; Apr+18] either due to sparking, or the appearance of "hot spots" which can be described as localized spots of high SE emission or light emission. Also, XENONnT [Apr+22c; Apr+23b] did not reach its design fields due to an electrical short between cathode and bottom screening mesh, as well as as a localized appearing "hot spot". Although for the latter, the precise source of origin is still unclear, it is currently believed, that inhomogeneities of the wire surface, like dents, tips, or notches can lead to points of a higher electric field which act as seeds for the emission of SE and photons. Also, wire sagging plays an important role in the technical design of parallel wire electrodes. Especially, for gate and anode, which are just a few millimeters apart, a precise knowledge of the overall wire sagging<sup>1</sup> is of vital importance for the experiment. To ensure that the wires of the XENONnT gate and anode would not touch, additional transverse wires were added. In order to improve the overall knowledge about harmful wire defects, and to improve the quality of future electrodes, a new optical inspection setup was developed in cooperation with the Mainz PRISMA detector lab.

The first subsection of this small extra section list the different tools and technical details of the new inspection setup, called XYZCamScan. Subsection A.1.2 discusses the optical inspection of the XENON1T cathode and gate electrodes which were conducted before the outbreak of the Covid pandemic. The last Subsection A.1.3 discusses a proof of concept of wire sagging measurements conducted with the XENON1T cathode, before the section concludes with a summary and outlook.

---

<sup>1</sup>In this case sagging, refers to the overall attraction of the wires due to the electric field between gate and anode. The sagging due to gravitational pull only plays a subdominant role.

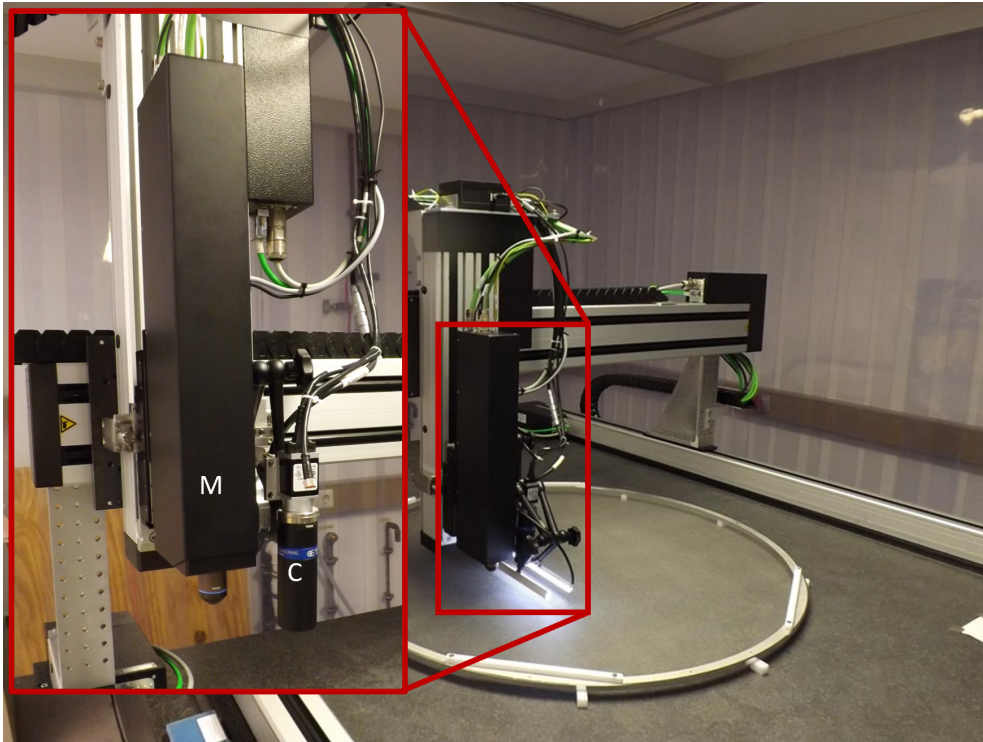


Figure A.1. – Photograph of the XYZCamScan inspection setup. The photo shows the optical inspection of the XENONIT gate mesh. The inset figure shows a zoom of the measurement devices. The black box labeled "M" is the confocal microscope, the object labeled "C" represents the telecentric lens of the high resolution camera which is mounted on top. Behind the camera a laser device for distance measurements is mounted. All devices are mounted to the z-axis of the shown cnc-machine, where the gentry is defined as y-axis, and the two parallel tracks on both sides of the table as x-axis.

### A.1.1 Inspection setup:

The wire inspection setup, XYZCamScan, offers several tools for high-resolution scans over large surface areas. In addition, it provides the means to measure the wire sagging of the individual wires, and therefore wire tension. The setup was designed together with two summer intern students, and a bachelor student, as well as in cooperation with the engineers and experts of the PRISMA detector lab. It was designed for a wide range of different applications, beyond the inspection of TPC electrodes. Therefore, the setup is part of the PRISMA detector lab facility which is a shared facility between many different groups at the JGU-Mainz. The main conceptual challenge of the setup is to achieve the desired high optical resolution  $\sim 5 \mu\text{m}$  for an inspection over large surfaces of a few square-meter.

The current design of the setup is shown in Figure A.1. It was added to an already existing setup which was used for the flatness measurement of detector panels [Ber+16]. The entire setup is mounted in a laminar flow box with an air purity rating of ISO 5-6, depending on the set flow. This is important to ensure the overall cleanliness of the electrodes while performing a scan.

The inspection setup itself is mounted onto a large polish granite table ( $2.2 \times 1.425$ ) m<sup>2</sup> with a surface roughness of about 14  $\mu\text{m}$ . The four main components of the system are given by a CNC machine to place the different devices over the object under study, a high-precision laser system for distance measurements, a high-resolution camera for optical inspections, as well as a confocal microscope for an even higher resolution, and the possibility to take three-dimensional scans of small objects. The granite table, CNC system, and distance laser were already commissioned and extensively operated in [Ber+16], before the new optical systems were added during this work.

All measurement devices are mounted to the z-axis of the CNC machine which allows precise positioning of the tools within the dimensions of the granite table. The CNC system is from the manufacturer ISEL, and its axes have a stepping resolution of 5  $\mu\text{m}$  in x and y directions, and 2.5  $\mu\text{m}$  in z. The overall repeatability of the system is 20  $\mu\text{m}$ .

The optoNCDT 2300-20 laser system from the manufacturer Micro-Epsilon, allows to conduct either high-precision distance measurements along the z-axis or thickness measurements of semi-transparent coatings. It has lateral resolution of 300 nm, a minimum working distance of 40 mm and a dynamic range of 20 mm.

As a high-resolution camera, the model acA4600-7gc from the company Basler was selected. The camera has a sensor size of  $(6.5 \times 4.6)$  mm<sup>2</sup> with  $(4608 \times 3288)$  pixels. The individual pixels are squared with an edge size of 1.4  $\mu\text{m}$ . They are coated according to a BG Bayer pattern to allow for colored image-taking. The camera is equipped with a telecentric lens TC23016 of the manufacturer Opto Engineering. A telecentric lens was chosen due to its property to project non-equidistant objects as same-sized objects onto the image plane. Hence, wires are imaged as flat objects compensating for their strong curvature which would otherwise cause problems in the later image analysis. The selection of the TC23016 was optimized for different important properties such as overall resolution, field of depth, and field of view. It provides a magnification  $M$  of  $\times 0.528$ , a working distance of 43.1 mm and a fixed working f-number  $wf/\#$  of 8. The field of view of system is about  $(12.31 \times 8.71)$  mm<sup>2</sup> which corresponds to a resolution of about  $(5 \times 5)$   $\mu\text{m}^2$ , with a field of depth of  $\sim 2$  mm

To allow for an even higher resolution the system was later augmented with an additional confocal microscope of the company NanoFocus/Mahr<sup>2</sup>. Such a microscope allows to take multiple images over a range of different z-slices to construct a fully three-dimensional image of the object under study. The microscope has its own z-axis which allows it to move the microscope up and down, independently of the ISEL system. The movement is done by two independent systems, a linear motorized positioning unit allowing for a rough z-adjustment over a range of 50 mm, and a second piezo module for fine positioning during the image taking. The piezo module has a dynamic range of 350  $\mu\text{m}$ , and its stepping depends on the magnification of the selected optical lens and chosen measurement sequence. The minimal step size ranges between 4 nm to 40 nm steps. The optical sensor of the microscope has a sensor size of  $1200 \times 1200$  pixel. The field of view and resolution of the system depends on the magnification of the selected lenses. Two lenses were purchased along with the micro-

<sup>2</sup>[https://www.mahr.com/fileadmin/user\\_upload/\\_assets/3\\_Group/Landingpages/Mahr-Showroom/Oberflaeche/MarSurf--3764321--FL--CM\\_Optical\\_3D\\_microscopy--EN--2020-02-11.pdf](https://www.mahr.com/fileadmin/user_upload/_assets/3_Group/Landingpages/Mahr-Showroom/Oberflaeche/MarSurf--3764321--FL--CM_Optical_3D_microscopy--EN--2020-02-11.pdf)



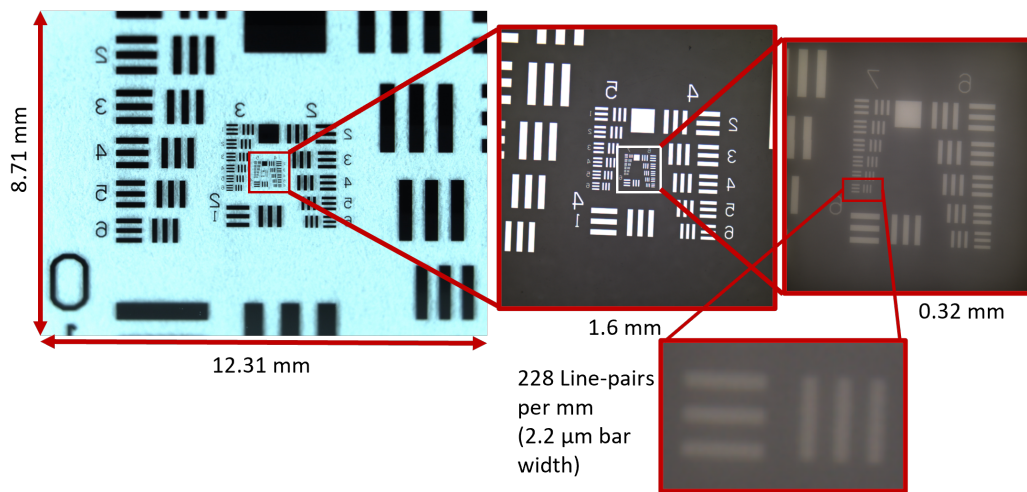


Figure A.2. – Comparison of the image resolution for the different optical systems of the XYZCamScan setup. The images show a USAF 1951 glass target, where the image on the left center and right hand side belong to the camera, and the microscope using a  $\times 10$  and  $\times 50$  magnification respectively.

scope. A  $\times 10$  magnification lens with a working distance of 11 mm, offering a field of view of  $(1600 \times 1600) \mu\text{m}^2$  with a resolution of about  $1 \mu\text{m}^2$ . And, a  $\times 50$  magnification lens with a working distance of just 1 mm<sup>3</sup> which has a field of view of  $(320 \times 320) \mu\text{m}^2$  with a theoretical resolution of about  $0.3 \mu\text{m}^2$ . The typical exposure times of the microscope depend on the task at hand. A single image of just one z-slice takes typically 10 ms, while a full z-profile scan can take several seconds depending on the z-range to be covered.

A comparison between the optical resolution of the camera, and the two magnification levels of the microscope is shown in Figure A.2. The blurriness of the image with the highest magnification is caused by small vibrations of the setup, which presumably originate from various sources in the building. This limits the resolution of the microscope to a larger value, nonetheless the  $2.2 \mu\text{m}$  wide bars of the calibration target can be still very well resolved. This figure also illustrates the general conceptual idea of the inspection. First, the electrodes are scanned in an automatized measurement using the high-resolution camera. This scanning procedure is discussed in more detail in the subsequent section. Afterwards, the microscope would be used to inspect points of interest with a higher resolution in more detail. Figure A.3 illustrates such a point of interest of a broken wire piece found on the XENON1T gate mesh which was taken with both magnification levels of the microscope.

<sup>3</sup>It should be noted that, although the working distance is shorter compared to the  $\times 10$  lens, the absolute z-positioning for the two lenses is almost the same as  $\times 50$  lens is also about 10 mm longer. This must be kept in mind to not crash the microscope into the object under study.

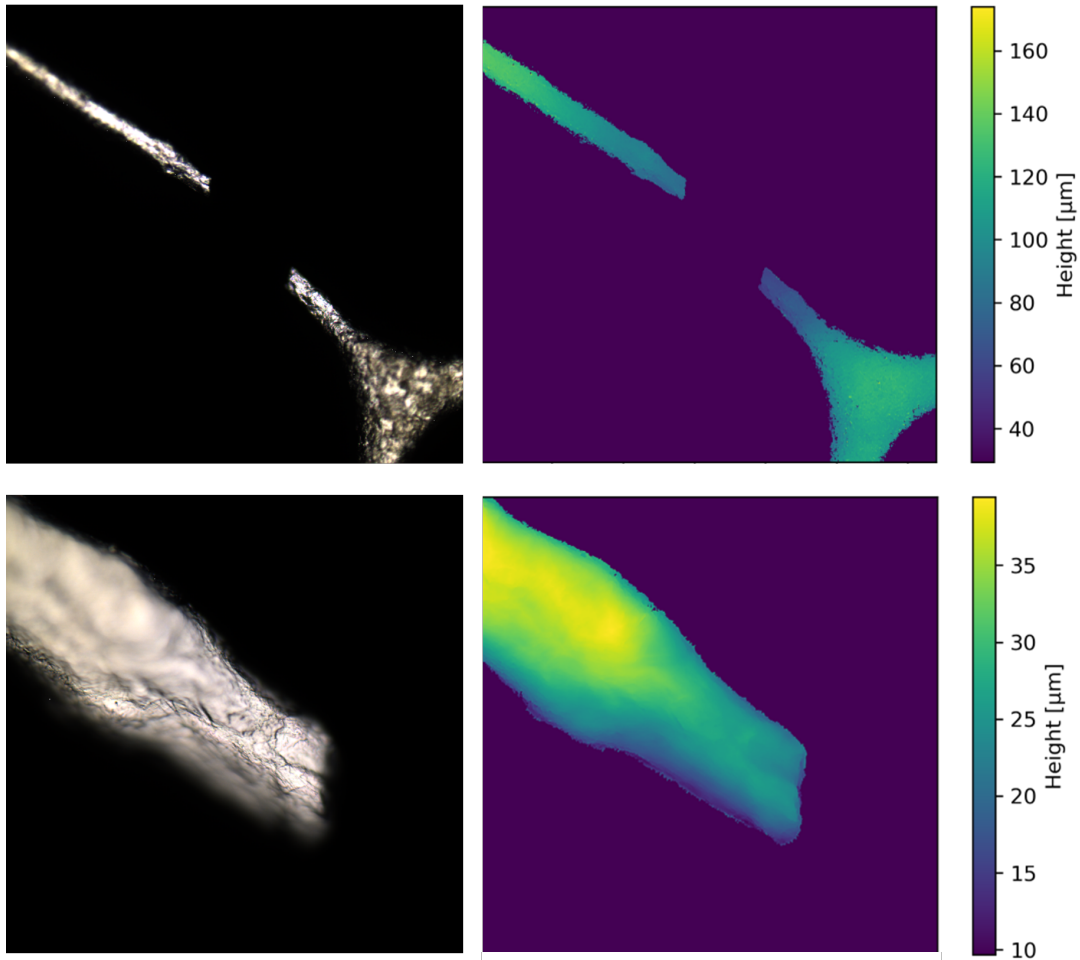


Figure A.3. – Exemplary height profiles and images for one of the wire damages found in the XENON1T gate mesh, taken with the confocal microscope. The upper row shows the full damage using the  $x10$  magnification of the microscope. The lower row shows the upper part of the damage using the  $x50$  magnification. The left column shows images taken at the central height of the height profile which is shown in the right hand-side column.

### A.1.2 Optical inspection

In order to get a better understanding of which wire damages may result in field emissions or sparking, a full optical scan of the XENON1T cathode and gate was performed. Since cathode and gate have been operated in XENON1T, regions in which a higher rate of SE emission were observed, were already known before. Thus, in order to learn which damages are harmful, the idea of this scan was to compare images of these regions with other regions of the electrodes. One important aspect of this task was, that the measurement needed to be automatized as the XENON1T electrodes had an inner diameter of about 95 cm [Apr+17c], which is quite large compared to the field of view of the camera. Thus, the C-libraries of all devices, excluding the confocal

microscope<sup>4</sup>, were ported as python packages which are collected in a single GitLab repository called XYZCamScan<sup>5</sup>. The libraries for the ISEL system and the distance laser were already ported before by Friedemann Neuhaus for the work in [Ber+16], and only added to the shared repository.

To ensure that the images taken during this automated measurement would show the best contour sharpness, a passive autofocus algorithm<sup>6</sup> was developed [Xu+11]. This was necessary as the telecentric lens of the setup has only a fixed focal length with a very small depth of field compared to the sagging of the wires. The working principle of the algorithm is rather simple, images of the object under study, e.g. a piece of wire, are taken at different distances in  $z$  which are adjusted using the ISEL system. These images are afterwards convoluted with a squared Laplacian matrix before the total contrast of the image is estimated by the total sum over the convoluted image. The convolution is normalized such that a chessboard-like pattern between white and black pixels yields a maximum contrast of 1. Figure A.4 and A.5 show the contrast maximization for the shown piece of wire. Three different types of Laplacian masks are compared

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & +4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad \begin{pmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & -4 & -1 \\ -4 & +20 & -4 \\ -1 & -4 & -1 \end{pmatrix} \quad (\text{A.1})$$

which are labeled 1,2 and 3 in Figure A.4 respectively. The peak width of all three masks is the same, and the same number of  $z$ -steps are required to find the point with the highest sharpness. Based on Figure A.5 one can also clearly see by eye, that image 4 has indeed the best contour sharpness as estimated by the algorithm. For the measurements discussed in the following, the first Laplacian matrix was used. The mask can be changed at any given time by the user if required.

The biggest drawback of a passive autofocus is the fact that it is a relative measurement. This means in order to determine if a picture is taken in focus, at least three images are required. The one which is tested, and two images with a smaller and larger distance towards the object as a reference. If one of the reference images shows a better contrast, the autofocus keeps on moving in the respective direction, using a user-defined stepping, until the maximum contrast is reached, or if a safety distance is exceeded. To assist the focus finding the analyst can also decide if only a specific ROI should be used for the focus finding, e.g. by limiting the image range to the wire. However, the  $z$  stepping should not be chosen too small as the whole system needs  $\sim 1$  s to get back to full rest after moving one of the ISEL axes. This makes automated scans of large-scale objects rather slow.

The scan of the XENON1T cathode was performed as illustrated in Figure A.6. To calibrate a coordinate system that has the  $y$ -axis aligned along the wire direction sev-

<sup>4</sup>The libraries of the microscope were not ported for two reasons. Firstly, the microscope is controlled through a separate computer compared to the rest of the system. Secondly, it was considered too dangerous to move the microscope in an automatized scan due to its short working distance.

<sup>5</sup><https://gitlab.rlp.net/prisma-detlab/xyzcamsan>

<sup>6</sup><https://gitlab.rlp.net/prisma-detlab/xyzcamsan/-/blob/master/xyzcamsan/AuxiliaryTools/autofocus.py#L14>

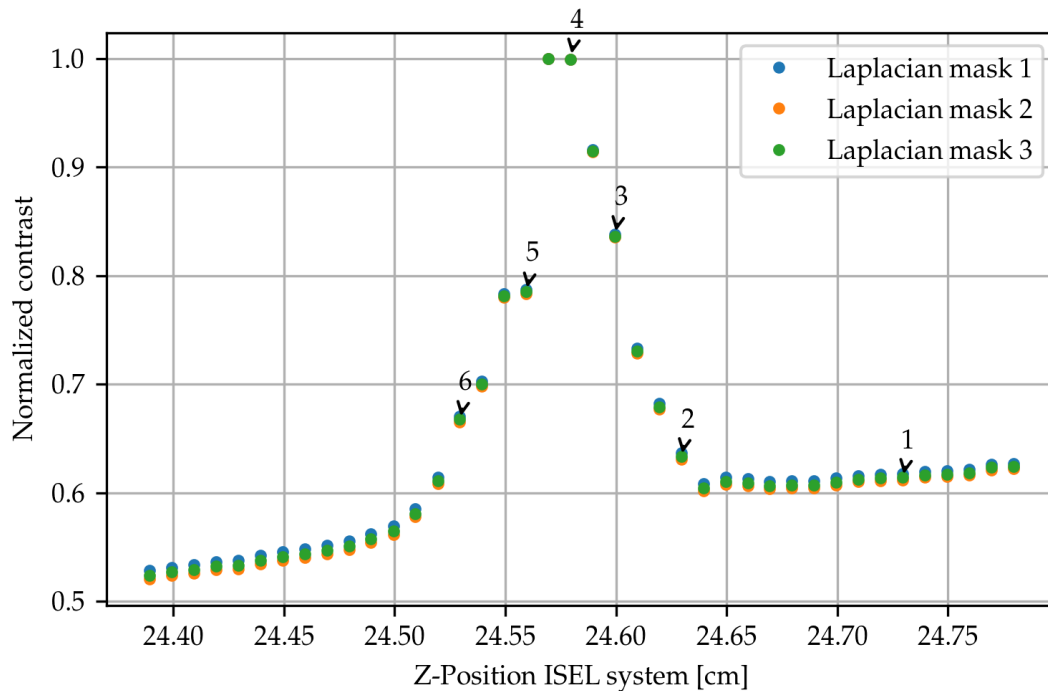


Figure A.4. – Example measurement showing the normalized contour sharpness estimated by the XYZCamScan autofocus as function of the absolute z-position of the ISEL system. Three different Laplacian matrices as explained in the text were used. Their results were normalized to the maximum such that the width of the respective peaks can be compared. The labels next to some of the computed contrasts correspond to the respective labels in Figure A.5.

eral drilling holes in the cathode frame were used. Afterwards, the field of view of the camera was limited to a narrow range to only accommodate the wire plus a  $\sim 100$  pixel on either side as a safety margin. The margin is required as the alignment of the defined coordinate system with the wires is not perfect and a drift of the wires out of the ROI was observed. The drift can be corrected by using simple image edge detection. The scanning of a single wire was always started at its center where the maximum sagging of the wire is to be expected. This was done to exploit the symmetries of the electrode during the automated inspection. The algorithm starts by scanning the first half of the first wire in one direction until a position outside of the frame is reached. Afterwards, the camera is moved back to the center, and the scan of the second half is performed. During the scanning of the wire halves, the algorithm automatically corrects the distance between the camera and the wire by using the passive autofocus discussed before. Once the scan of a wire is completed, the ISEL system moves back to the starting position at the center of the wire, before moving by one wire pitch onto the next wire. A clear tilt of the overall mesh can be observed in Figure A.6, which was also observed in the sagging measurement explained in the next section.

In total, 8.400 images were taken during the scan of the XENON1T cathode which totals to 75 GB of data. The full scan took about 60 h. Cuttings of a few example images are shown in Figure A.6. Many different light conditions were tested, before

the full scan of the XENON1T cathode. The setting which was found to be best is a so-called dark field illumination which is also shown in Figure A.1. In this type of illumination setting the light is emitted under a very shallow angle such that it cannot hit the camera sensor under specular reflection. Thus, a perfectly flat surface appears as dark. However, tips, dents, and notches become visible as they can reflect light onto the sensor. This is also seen in the images presented in Figure A.6. Most of the wires appear black and only a few features are highlighted as bright spots or larger discolorations. Also, salt remnants of the cleaning procedure can be observed (second image), and large-scale destructions (third image). The yellowish color of the cathode wires originates from the fact that the wires are coated with a thin gold layer for a smoother surface.

Two different approaches were developed to analyze the wire images. An autoencoder method that tried to identify pictures into classes of "good" and "bad" wires, where "good" wire pieces would be just mostly black. This method was developed by another Ph.D. student, Jan Lommler. And, a second method, in which features are detected and identified based on their intensity using a DBScan algorithm, after applying some image preprocessing steps first. This approach was developed during the presented work. Identified features would be characterized by their total overall size and brightness. Ultimately, both methods suffered from the same needle in the haystack problem, namely that there are simply too many features found all over the electrodes. Thus, it was not possible to associate the observed emissions of SE with any of the found defects conclusively. To solve this issue the setup is currently extended, which is briefly discussed in the closing conclusion of this extra section.



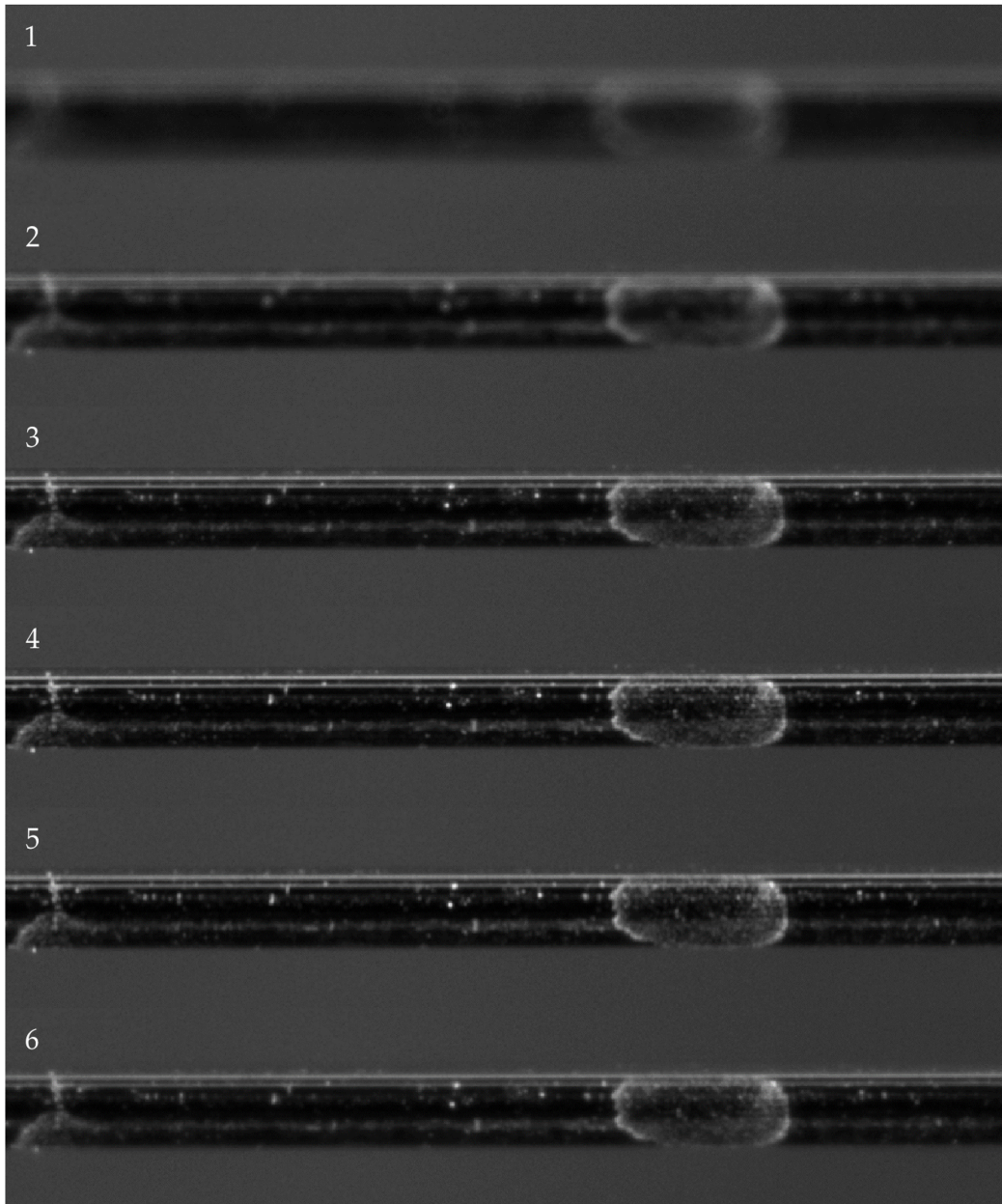


Figure A.5. – Images belonging to the labeled points in Figure A.4.

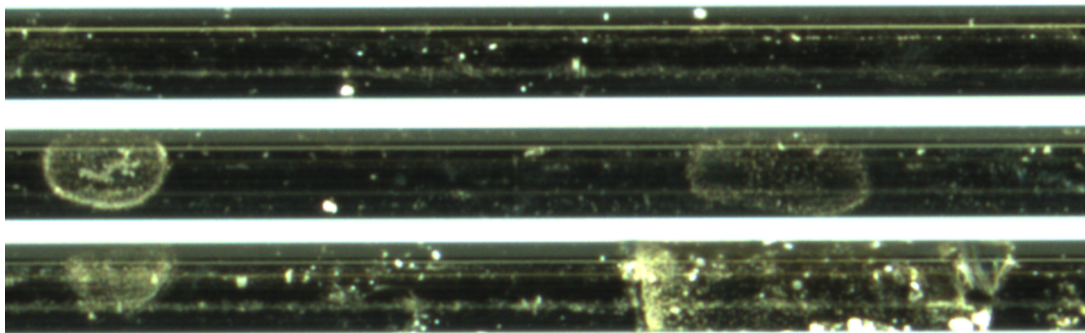
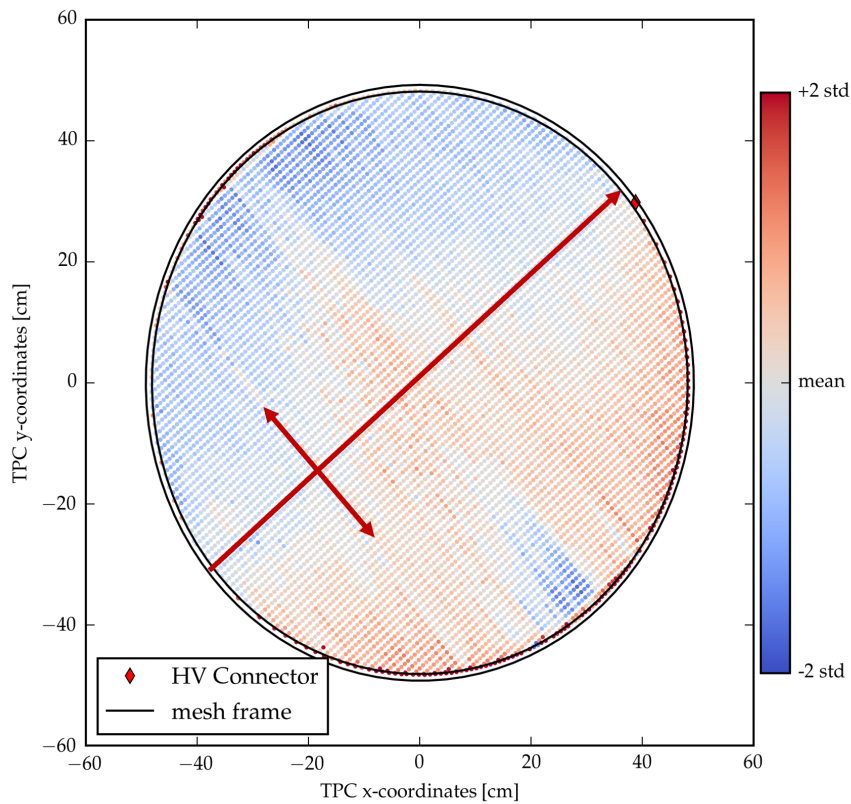


Figure A.6. – Illustration of the XENON1T cathode electrode inspection. The upper plot illustrates the automated scanning procedure of the XENON1T cathode. The scan was started with the outer most wire starting in the third quadrant. Each wire was scanned starting from the center in both directions as depicted by the red double arrow. Afterwards the next wire is scanned by moving along the larger red arrow for one wire pitch. The colorcode indicates the z-position of the ISEL system for the respective wires normalized to the average z position of the given wire. The z position was determined by the autofocus algorithm explained in the text. The images below the plot illustrate different wire defects found on the meshes. The large defect in the bottom picture is one of the defects found near a position where the emission of SE signals was observed.



### A.1.3 Wire sagging measurement

This short section serves as a proof of concept, that the wire sagging of the XENON1T cathode [Apr+17c], and therefore the tension of the wires can be measured by the micro-epsilon laser of the inspection setup. Similarly, as in the automated optical inspection of the XENON1T cathode, the drillings in the cathode frame were used as reference points, to define a coordinate system where one of the axes is aligned with the wire direction. The sagging measurement was performed in a simple approach by moving the ISEL system in a zig-zag pattern over the range of a full square, containing the XENON1T cathode. The pattern is also depicted in Figure A.7 on the left-hand side. The exact position of the wires is found by repeated distance measurements in the vicinity of a wire using a fine transverse stepping of 10  $\mu\text{m}$  to approach the wires. Once a wire was successfully found and the distance measured, the ISEL system was moved by 90 % of the wire pitch to the vicinity of the next wire, where the fine-stepping approach was repeated. Ultimately, the distance between the laser and wire was measured every 50 mm along the individual wires. The distance at each position was averaged over 200 independent measurements. The resulting distances are shown in Figure A.7 on the left-hand side.

To convert the measured distance into a wire sagging, they must be correct for the sagging of the gantry of the ISEL system and referenced to a plane parallel to the wire frame. The former is achieved by measuring the distance between laser and table without cathode electrode over the same range as before but using an equally spaced stepping of 1 cm in both directions. The resulting map is shown in Figure A.8 left-hand side. In order to estimate the reference plane, the distance between cathode frame and laser was measured for 20 points over the full circumference of the cathode. In both measurements the absolute  $z$  position was measured, i.e. the  $z$  coordinate of the ISEL system was added to the measured distances. This was necessary as the two measurements had to be conducted at different heights, to stay within the dynamic range of the sensor. A linear plane was fitted to the measured points of the frame in order to define the reference plane for the wire sagging. As a fitting method, a  $\chi^2$ -method was used. The resulting plane is shown in Figure A.8 right-hand side. The measured heights of the frame vary by about 1.5 mm, and already by eye based on Figure A.8 it becomes clear, that a simple plane is not sufficient to describe the measured data, which is also confirmed by the resulting  $\chi^2$  of the fit. The overall tilt of the frame might be due to PTFE spacer which was put as protection between cathode and table. However, the deviation of the frame heights from a simple flat plane model points to a deformation of cathode frame. For the feasibility study presented in this section, the reference plane is sufficient, however further studies are required to understand whether such a deformation leads to any systematic bias in the measurement.

Based on the gantry sagging model and the frame reference plane the wire sagging is computed. The result is shown in Figure A.7 on the right-hand side. In contrast to the uncorrected measurements on the left, a clear sagging towards the center of the mesh can be observed for the individual wires. Further, some of the wires show a stronger sagging than others. The sagging along one of the central wires of the cath-

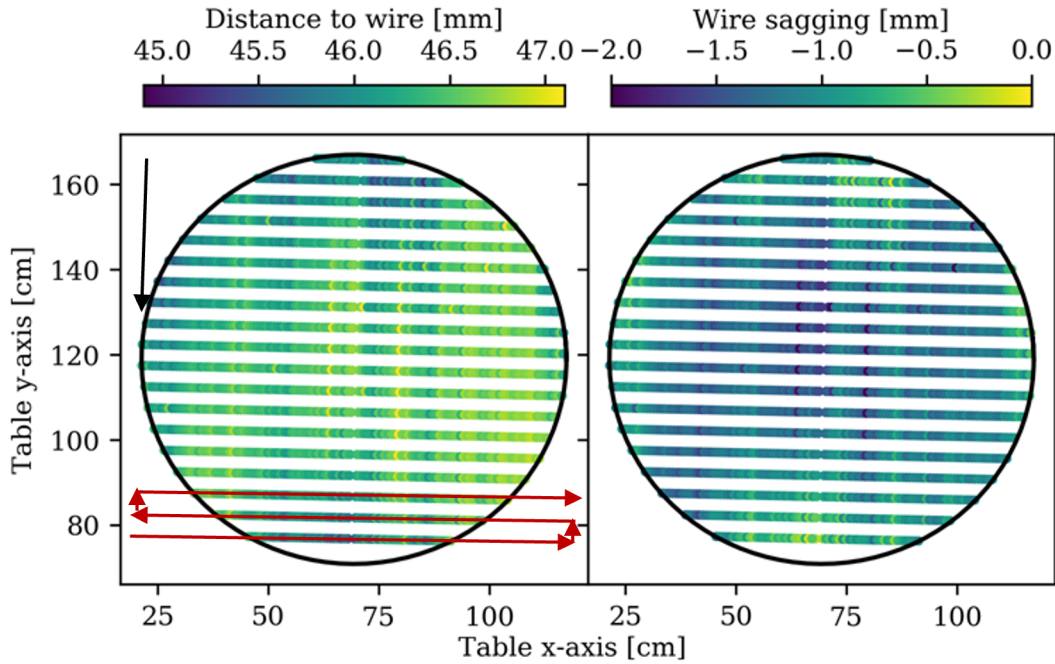


Figure A.7. – Distance measurement between the micro-epsilon laser and XENON1T cathode electrode. The plot on the left hand side shows the raw measured distance between wires and laser. The wires of the electrode are oriented parallel to the black arrow. The red arrows indicate the zig-zag pattern performed during the measurement. The figure on the right hand side shows the sagging of the individual wires after correcting for the gantry sagging, and subtracting the reference plane fitted to the cathode frame as explained in the text.

ode is shown in Figure A.9. The endpoints, A and B, of the wire, show a slight height difference. However, the difference is only about 0.2 mm over the full measured distance which is about 80 cm. In addition, the sagging values are systematically shifted by about one millimeter below the reference plane which also points to a systematic bias in the reference plane as discussed before. This is expected as the simple plane model does not describe the measured points on the frame well. Thus, as consequence the sagging is only determined with respect to the measured points on the wires, ignoring the additional shift in  $z$ . The maximum sagging of wire 47 is evaluated to be  $\sim 0.46$  mm, using the catenary equation

$$y(x) = c \cdot \cosh((x - x_0)/c) + y_0 \quad (\text{A.2})$$

with the catenary constant  $c$ , which was fitted to the data using a  $\chi^2$  method. The fit results in a  $\chi^2/\text{ndof}$  of 5/15, failing to reject the fit model. The catenary constant for the shown wire was estimated to be  $c_{47} = 0.192$  mm. Together with the density of the wire of  $\rho_{\text{wire}} = 7.91644$  g/m<sup>3</sup> and its diameter of  $d = 215.9$   $\mu\text{m}$  [Cha19], the horizontal wire tension can be estimated via

$$T_0 = c \cdot g \cdot \rho_{\text{wire}} \cdot \pi \cdot (d/2)^2 \quad (\text{A.3})$$

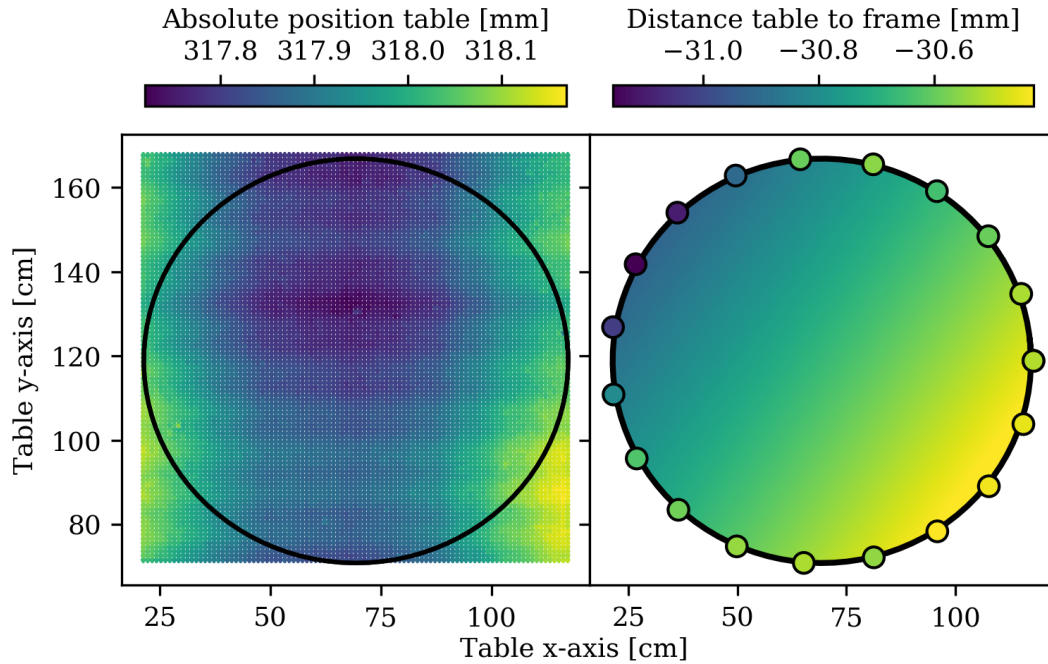


Figure A.8. – Measurements of the of the gantry sagging and cathode frame reference plane. The plot on the **left hand side** shows the absolute position of the table measured by the laser system. The figure on the **right hand side** shows the 20 measured points on the frame, and the fitted reference plane. Both, points and plane, use the same color scale.

where  $g$  is the earth's gravitational field strength. For the shown curve the horizontal tension is measured to  $(0.54 \pm 0.01)$  N, where the uncertainty only represents the statistical uncertainties of the measurement. This is well below the tensions the wires were mounted with during the construction of XENON1T which was between 1 N and 2 N [Cha19]. The reason for this discrepancy is unclear. It might be that the wires lost some of their tension over time, or during the decommissioning of the detector. Due to the issues with the reference plane, also a systematic bias in the analysis cannot be excluded, although the overall tilt of the frame seems to be small. More systematic measurements are required to understand this discrepancy better. The remaining wires of the XENON1T cathode show a similarly reduced tension which rules out a statistical fluke. Their results are shown in Figure A.10.

#### A.1.4 Conclusion and outlook

In this small extra section, the activities with the XYZCamScan setup have been summarized. The setup is an optical inspection setup that allows to scan large surfaces with a high-resolution camera. These scans can be fully automatized due to a passive autofocus algorithm which was developed for the scans of the XENON1T gate and cathode. In addition, as part of the presented work, the system was augmented with a confocal microscope which allows to take three-dimensional images with an even

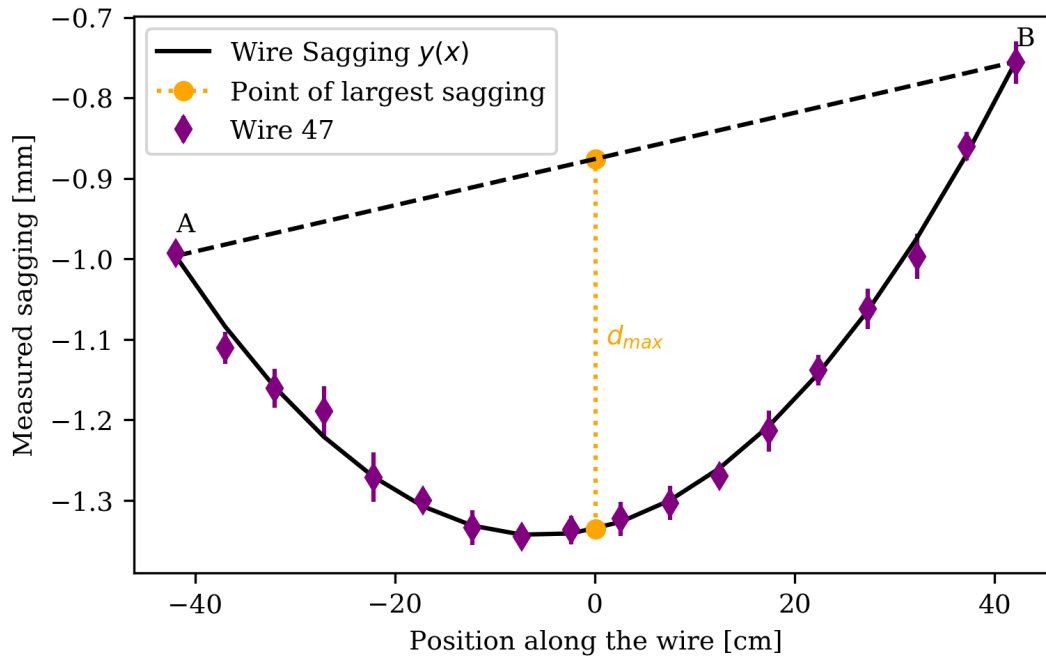


Figure A.9. – Best fit of the catenary curve to wire 47 of the XENON1T cathode. The purple data points show the wire sagging along the wire. The black solid curve shows the best fit of a catenary curve. The black dashed line shows the reference plane which is used to determine the maximum wire sagging depicted in orange. Please mind the different scales on the x and y axis which lead to a very exaggerated sagging of the wires.

higher resolution. A first scan of the old XENON1T electrodes has been performed, and the images were analyzed using two different approaches. Both approaches suffered from the large number of defects found on the wires, such that no conclusive statement about the severity of individual defects could be made.

In addition, a first proof of concept study was performed in which the sagging of the individual wires was determined based on distance measurements. The determined sagging was also used to give an estimate for the wire tension which was found to be significantly lower than expected, based on the technical design of the XENON1T cathode. As for now, it is unclear if this discrepancy is caused by a systematic effect in the measurement, or if the wires lost their tension during the operation of the electrode. To calibrate the distance measurements, and to get an idea about systematic effects additional calibration blocks for CNC machines were purchased which are milled to a micro-meter precision.

At the end of this study, the idea was developed to put the electrodes into an acrylic box such that field emission can be directly observed when applying high voltage in an argon atmosphere. In order to be able to observe the entire electrodes during the slow ramp-up of the high voltage, a second overview camera was added to the system. In this way, the ramping up of the electrodes can be set up in a way that it is stopped as soon as field emissions are observed. This allows for mitigating the danger of sparking, which could lead to irreversible damage to the wires. Afterwards, places

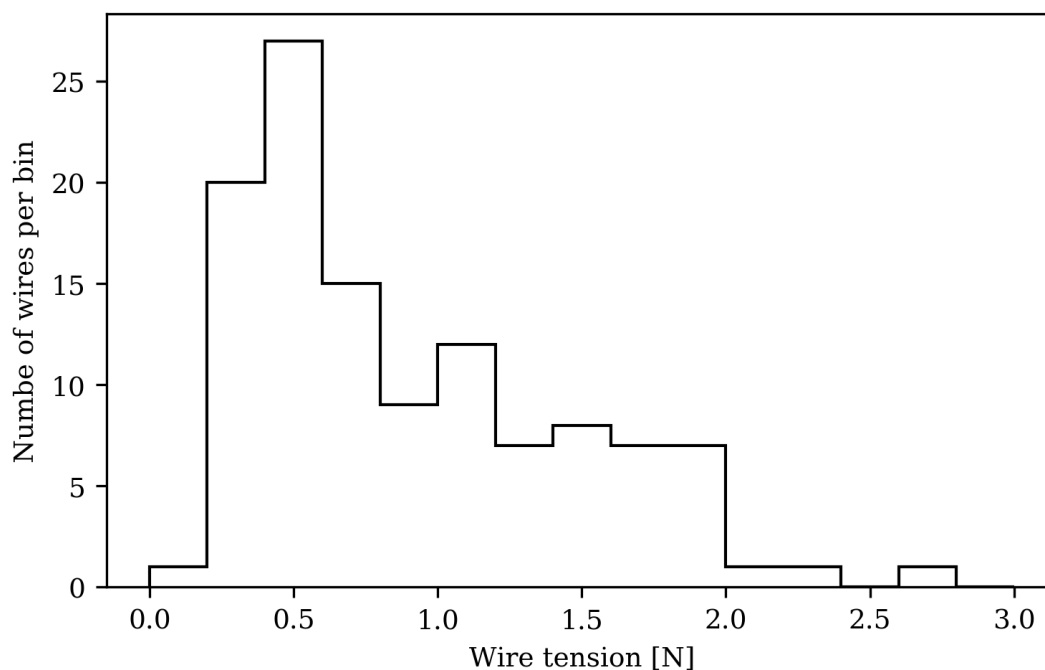


Figure A.10. – Horizontal wire tension of the XENON1T cathode wires estimated based on the catenary equation.

in which field emissions were spotted can be inspected by the already existing system. The work on this extended setup is currently continued by Alexander Deisting and Shumit Mitra. A first image of this new setup is shown in Figure A.11. The first tests have been already successfully performed, which show the creation of corona arcs near places where wire defects can be found. One of these arcs is also shown in Figure A.11. In addition, this new setup allows to perform wire sagging measurements under more realistic conditions including an electric field.

Given, the poor performance of the XENONnT electrodes, a set of backup electrodes is currently prepared, and the new optical inspection at Mainz will play a critical role in their quality assessment. However, there was no final decision yet if these new electrodes will be ever installed into the detector. In either case, the additional studies conducted now will also bring benefit for future larger-scale dual-phase TPCs.



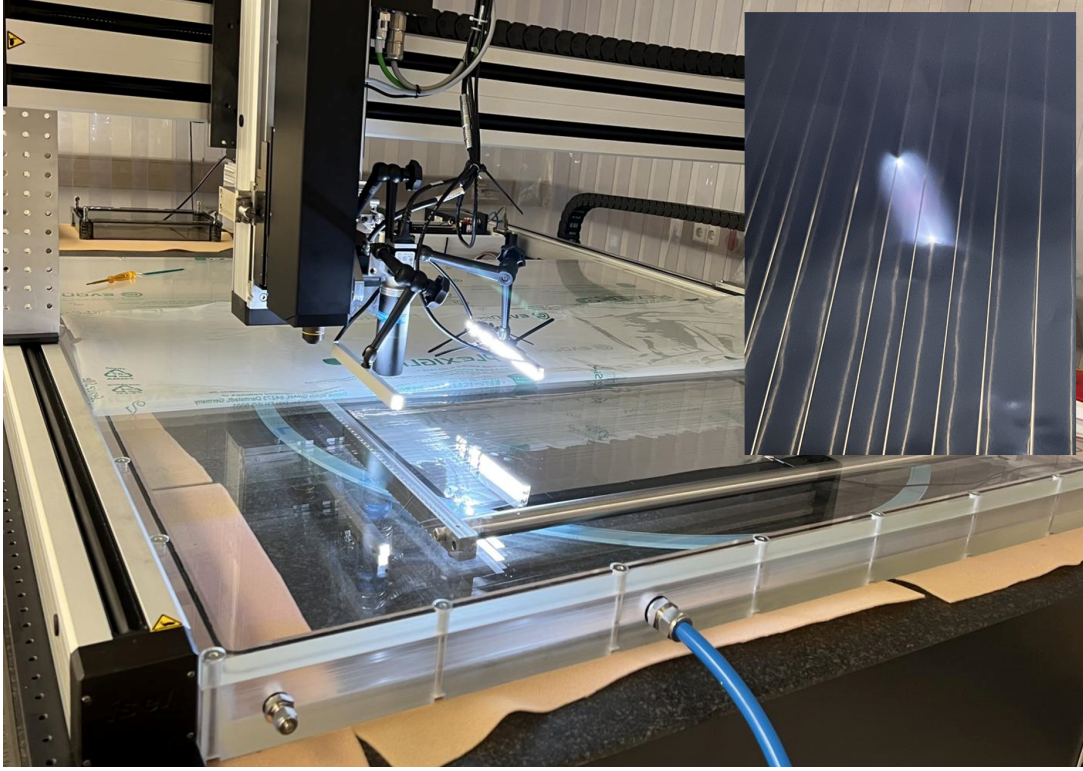


Figure A.11. – Image of the extended wire inspection setup. The main photograph shows a measurement with the extended inspection setup. Inside the acrylic box some old XENONnT test wires are mounted onto a stretching frame. The inset picture shows the field emission in argon atmosphere. The picture was taken with the camera of a mobile phone for a side view.

## A.2 ACCIDENTAL COINCIDENCE RATE

The accidental coincidence rate for a  $n$ -fold coincidence using  $m$  different counter can be estimated based on [JÁN44; Sch44], which discusses a  $n$ -fold coincidence rate for  $n$  counter. Consider the following case in which one counter  $i$  records a signal at the time  $t = 0$ , illustrated in Figure A.12. To satisfy the coincidence requirement  $n - 1$  additional signal must be found within  $\pm T$ , where  $T$  is the revolving time of the coincidence. The chance to find  $n - 1$  signals of  $m$  counter in this window is given by:

$$p(n - 1|m, T) = \binom{m}{n - 1} \cdot p_j(T)^{n-1} \cdot (1 - p_j(T))^{m-(n-1)} \quad (\text{A.4})$$

where  $p_j$  is the chance that the counter  $j$  produces at least one signal in the range  $\pm T$ . The chance for this can be approximated by

$$p_j \approx p \approx 2 \langle R_{DC} \rangle \cdot T \quad (\text{A.5})$$

assuming that all PMTs have the same dark rate. It is further assumed that the difference between  $p = 1 - \text{Pos}(0, \lambda) \approx \lambda$  is small which is true for small values of  $\lambda$ .

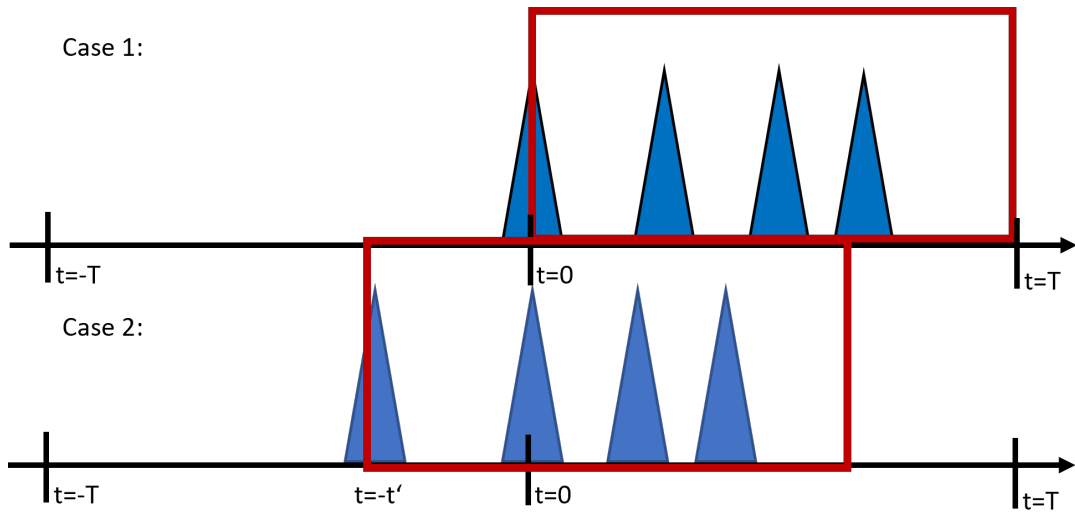


Figure A.12. – Schematic for the derivation of the analytical accidental rate model explained in the text.

The remaining  $n - 1$  signals can be arranged in multiple different ways around the signal at  $t=0$ . In the following, only the two leading order cases are discussed. In the first case, the signal of counter  $i$  is the leading signal and all other  $n - 1$  signals are found in the window  $[0, T]$ . The chance for this is given by

$$Q_1 = \left(\frac{1}{2}\right)^{n-1}. \quad (\text{A.6})$$

In the second case, one of the other  $n - 1$  signals of the  $m$  counter was detected first in the interval  $[-t', -t' + dt']$  and the coincidence is given if  $n - 2$  signals are found in a window  $[-t', -t' + T]$ <sup>7</sup>. Thus this probability is given by

$$Q'_2 = m \cdot \overbrace{\frac{dt'}{2T}}^{\text{A}} \cdot \underbrace{\left(\frac{1}{2}\right)^{n-2}}_{\text{B}} \quad (\text{A.7})$$

where A describes the chance to find a signal in  $[-t', -t' + dt']$ , and B describes the chance of finding  $n-2$  signals in  $[-t', -t' + T]$ . Since this reflects the case that another signal was found before  $t = 0$ ,  $Q'_2$  must be integrated between  $[-T, 0]$  for  $dt'$ . Thus, the accidental dark rate  $\langle R_{AC} \rangle$  for an  $n$ -fold coincidence with  $m$  counter and resolving

<sup>7</sup>Only  $n - 2$  signals are needed as there is already the signal at  $-t'$  and  $t = 0$  which contribute to the coincidence.



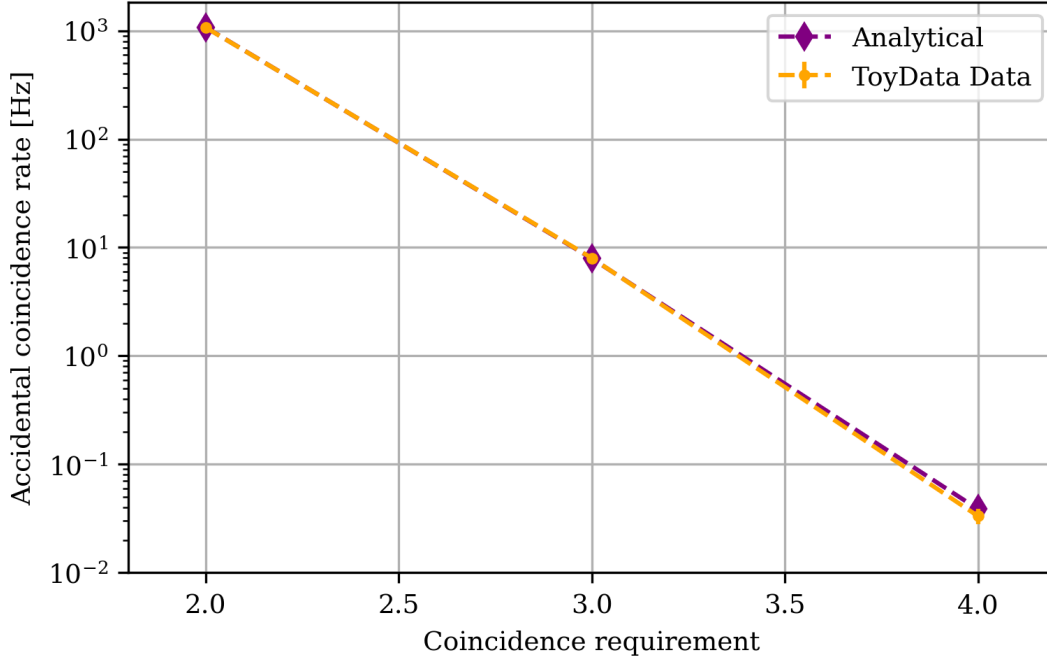


Figure A.13. – Comparison between the analytical and toy MC model for the rate of accidental coincidences for a  $n$ -fold coincidence with  $m$  counter and a resolving time  $T$ . The statistical uncertainty of the toy data are shown, but smaller than the data points itself.

time window  $T$  is given by

$$\begin{aligned} \langle R_{AC} \rangle &= p(n-1|m, T) \cdot (Q_1 + Q_2) \cdot \langle R_{DC} \rangle + O(t^n) \\ \langle R_{AC} \rangle &\approx (m+1) \binom{m}{n-1} (\langle R_{DC} \rangle \cdot T)^{n-1} (1 - 2 \langle R_{DC} \rangle T)^{m-(n+1)} \cdot \langle R_{DC} \rangle \end{aligned} \quad (\text{A.8})$$

where  $Q_2$  represents the integral of  $Q_2'$ . Higher order terms  $O(t^n)$  of having more than one other signal before  $t = 0$  have been neglected.

To validate the derived formula the analytical model was compared against a toy MC in which  $n$  toy signals were uniformly distributed in an interval  $\Delta T$ , where  $n$  was drawn from a Poisson distribution with  $\lambda = 120 \cdot \langle R_{DC} \rangle \Delta T$ . The toy signals were processed by the same coincidence trigger as for regular data, which is discussed in Section 3.3. Simulation and analytical model assume a resolving time of 200 ns as used in the event building for the NV, and a PMT dark rate of  $\langle R_{DC} \rangle = 618$  Hz. The toy experiment was repeated 10 times and the averaged results are used. The resulting rate of both approaches is shown in Figure A.13. Toy data and analytical model show a very good agreement for the 2-fold and 3-fold coincidence case with a deviation of less than one percent. For the 4-fold case the two models deviate for about 10% which means that higher-order terms start to play a role. However, the absolute rate of accidental coincidence becomes at these high thresholds negligible anyhow.

To compare the analytical and the toy model further, Figure A.14 shows the accidental coincidence rate for other detector conditions. One with a high average dark rate

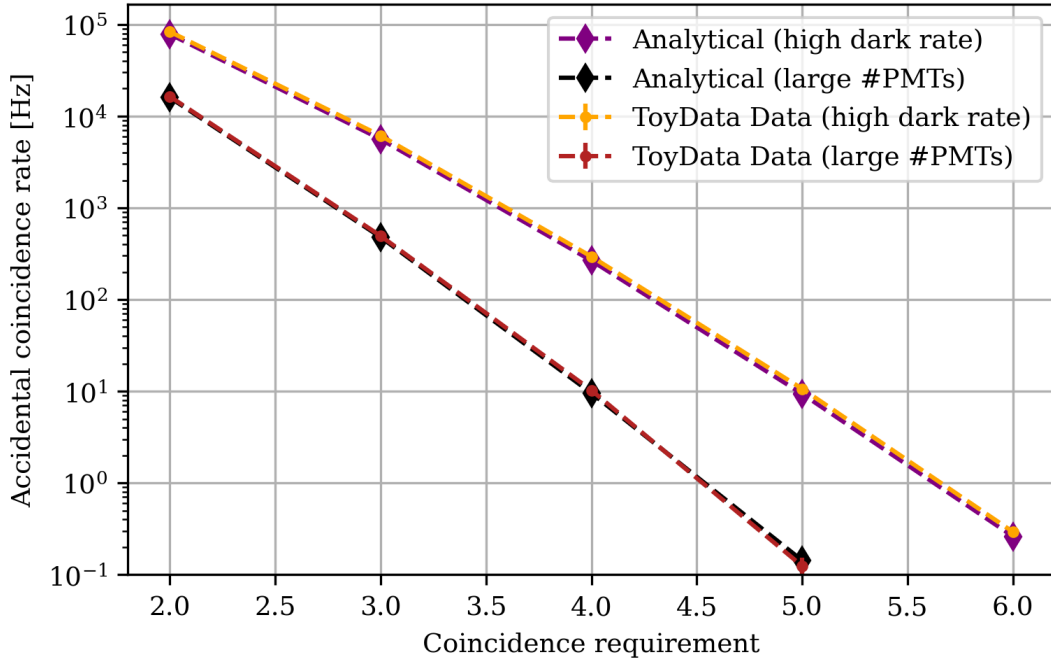


Figure A.14. – Comparison of the analytical accidental dark rate model with the toy data simulation using two fake detector configurations, which are explained in the text.

of 5000 Hz per PMT and 120 PMTs, as well as one with a similar dark rate per PMT as for the NV with 600 Hz, but with 500 PMTs. Also here both models match very well except when the rate of accidental coincidences becomes smaller than 1 Hz.

### A.3 BACKGROUND RATE INCREASE DUE TO GD-LOADING

The simulations used to determine the background rate increase in the NV due to Gd loading, use an older box like geometry for the neutron veto which has an overall larger water volume of ~84 t compared to the final design which only contains ~33 t of water. Assuming that the background rate due to contamination in the Gd sulfate salt is in the first order independent of the shape of the neutron veto the background rate for the final design can be determined by scaling the predicted rates of the box design. The simulations predict a rate increase at a 10-fold PMT coincidence requirement of  $R_U(C_U) = 0.28$  Hz and  $R_{Th}(C_{Th}) = 0.24$  Hz for  $^{238}\text{U}$  and  $^{232}\text{U}$  respectively, for a specific concentration of  $C_U = C_{Th} = 1$  mBq/kg. To convert the concentrations stated in Table 4.1 into the specific activity  $A'_I$  of an isotope  $I$  one can use

$$A'_I = \frac{A_I}{m_s} = \frac{C_I}{M_I} \cdot N_A \cdot \frac{\log |2|}{\tau_{I,1/2}} \quad (\text{A.9})$$

where  $A_I$ ,  $C_I$ ,  $\tau_{I,1/2}$  and  $M_I$  the activity, the concentration, half-life and the molar mass of the specified isotope respectively,  $m_s$  the mass of the sample and  $N_A$  the Avo-

gadrol number. Converting the numbers from Table 4.1 results in a specific activity of 37 mBq/kg and 152 mBq/kg for  $^{238}\text{U}$  and  $^{232}\text{Th}$  respectively. Thus, scaling the rates mentioned above by the relative masses of the two neutron veto designs the resulting background rates due to  $^{238}\text{U}$  and  $^{232}\text{Th}$  are  $R_U \approx 4.1$  Hz and  $R_{Th} \approx 14.6$  Hz. Assuming that only 1/3 of the Gadolinium sulfate is from Treibacher and that the contamination of the Japanese salt is negligible<sup>8</sup> results in a total rate increase of  $\sim 6$  Hz.

#### A.4 NR DATA QUALITY CUTS

In this short section, a brief description of the individual data quality cuts is given, which are used in the TPC NR calibration discussed in Chapter 6. The fraction of remaining events after applying the individual cuts are summarized in Table A.1. All cuts were optimized for an overall high acceptance in the WIMP search ROI, which varies between 95% and 99%.

The first four cuts in Table A.1 are exposure cuts. They remove all events which should not be considered due to different technical considerations. **cut\_interaction\_exists** only ensures that all events have a main S1 and main S2 signal. As events are just built based on triggering S2s peaks, some of them may not contain any S1 signal and are therefore not interesting for the presented SR0 WIMP search. **cut\_main\_is\_valid\_triggering\_peak** goes a level further and ensures that the main S2 has a total peak area of at least 100 pe, as this is the triggering condition for a peak. There can be rare conditions where this is not the case. For example, the triggering S2 peak which created the event is located before the main S1 in that event, and thus can therefore not be the main S2. In this case, a random small S2 might be assigned main S2 which can have an area below the required 100 pe. These edge cases usually appear if the data taking is interrupted, e.g. if the digitizer go busy, or the last event is too close to the edges of a run. The latter two cases are handled by the **cut\_daq\_veto** and **cut\_run\_boundaries** cuts. The **cut\_daq\_veto** removes all events which have a main S1 or main S2 closer than one full drift length to a veto interval issued by the DAQ, e.g. when the digitizer are busy. **cut\_run\_boundaries** removes all events which are found to be in the first second after a run start or the last 100 ms before the run ends.

**cut\_s1\_tightcoin\_3fold** and **cut\_ambe\_s2\_area** represent the thresholds which are applied in the NR calibration analysis. **cut\_s1\_tightcoin\_3fold** only keeps events with an S1 signal which satisfies the 3-fold coincidence, that requires at least three PMT signals within a  $\pm 50$  ns window around the S1 center. Where the peak center is defined as the position where the accumulated peak area reaches 50%. **cut\_ambe\_s2\_area** is the 250 pe S2 threshold which was used in the NR calibration data. The WIMP search uses a smaller threshold of 200 pe.

The next two cuts in Table A.1 are again two exposure cuts and are only applied for the NR calibration. **cut\_nv\_tpc\_coincidence\_ambe** only select TPC events whose S1 is in a coincidence with NV. **cut\_fiducial\_volume\_ambe** removes all events which are reconstructed outside of the fiducial volume which is used in the NR calibration.

<sup>8</sup>See Table 4.1 on page 91.

Table A.1. – Cut statistics of the NR calibration. *cut\_nr\_roi* and *cut\_roi\_bbf* are two simple selections which select all events with a *cS1* below 300 pe and 100 pe respectively. The remaining cuts are explained in the text.

cut	Events (accumulated)	Event rate [Hz] (accumulated)
No cuts:	4565070	18.87989
<i>cut_interaction_exists</i>	4195679	17.35219
<i>cut_run_boundaries</i>	4192862	17.34054
<i>cut_main_is_valid_triggering_peak</i>	4192624	17.33956
<i>cut_daq_veto</i>	3997677	16.53331
<i>cut_s1_tightcoin_3fold</i>	3766654	15.57786
<i>cut_ambe_s2_area</i>	3728634	15.42062
<i>cut_nr_roi</i>	119525	0.49432
<i>cut_fiducial_volume_ambe</i>	46191	0.19103
<i>cut_nv_tpc_coincidence_ambe</i>	10219	0.04226
<i>cut_s2_single_scatter</i>	3618	0.01496
<i>cut_s2_pattern</i>	3505	0.0145
<i>cut_s2_width_wire_modeled_wimps</i>	2839	0.01174
<i>cut_s2_recon_pos_diff</i>	2827	0.01169
<i>cut_cs2_area_fraction_top</i>	2818	0.01165
<i>cut_s2_naive_bayes</i>	2815	0.01164
<i>cut_s1_single_scatter</i>	2692	0.01113
<i>cut_s1_width</i>	2589	0.01071
<i>cut_s1_area_fraction_top</i>	2567	0.01062
<i>cut_s1_max_pmt</i>	2561	0.01059
<i>cut_s1_pattern_top</i>	2560	0.01059
<i>cut_s1_pattern_bottom</i>	2553	0.01056
<i>cut_s1_naive_bayes</i>	2537	0.01049
<i>cut_roi_bbf</i>	1986	0.00821

This volume is less strict than the volume used for the WIMP search. Both cuts are explained thoroughly in Chapter 6.

The *cut\_s2\_single\_scatter*, *cut\_s2\_pattern* and *cut\_s2\_width\_wire\_modeled\_wimp* cuts, most dominantly remove multi-scatter events in the NR calibration data. The *cut\_s2\_single\_scatter* and *cut\_s2\_width\_wire\_modeled\_wimp* cut, are already discussed more extensively in Section 6.2.2. The former cut removes all events which have an alternative S2 larger than a certain threshold, defined by the size of the main S2. The cut threshold was optimized based on  $^{220}\text{Rn}$  calibration data, which compared to AmBe does not contain any multi-scatter events. The S2 width cut removes all events which exhibit an S2 width that is larger than predicted by electron diffusion, see also Section 2.1.2. This cut also acts against accidentally paired S1 and S2 signals which do not have any causal relation. The additional reference "wire modeled" refers to the fact that in the region near the transverse wires, a looser cut definition is used due to a broadening of S2 signals by a local field variation, see also Section 8.3. In

the region far from the transverse wires the cut was optimized based on waveform simulations while in the near wire region, the cut boundaries are defined based on calibration data. The **cut\_s2\_pattern** cut compares the hit-pattern of the main S2 signal recorded by the top PMT array with simulated hit-patterns for the given S2 size and position. The patterns are compared using a  $\chi^2$  test. Thus, like the S2 width cut, it rejects multi-scatter signals with a single merged S2 signal.

The **cut\_s2\_recon\_pos\_diff** cut has usually a very high acceptance and removes only very few events. It removes events where the reconstructed S2 position for the three different position reconstruction algorithms, mentioned in Section 3.2, differ by more than S2 area-dependent threshold.

The **cut\_cs2\_area\_fraction\_top** cut removes events with an abnormally large or small contribution of the top PMT array to the total recorded S2 area. The light fraction recorded by the top array for S2 signals is usually larger than for S1 signals given that their signal is produced above the liquid phase close to the top PMT array. For the same reason, S1 signals also exhibit a smaller,  $z$ -dependent area fraction top. In addition, the area fraction top for S1 signals is systematically shifted towards smaller fractions due to the liquid-gas interface which can lead to a total reflection of photons. These relations are used in **cut\_s1\_area\_fraction\_top** cut, to discriminate for good S1 signals. The area fraction top cuts are optimized based on calibration data like,  $^{37}\text{Ar}$ ,  $^{83m}\text{Kr}$ , and  $^{220}\text{Rn}$  which can be diluted uniformly in the liquid.

The two naive Bayes cuts **cut\_s1\_naive\_bayes** and **cut\_s2\_naive\_bayes**, are based on a Naive Bayes classifier score which describes how S1 or S2 like a given peak is by defining threshold based on that score noncanonical looking S1 and S2 can be rejected. For more information, the reader is referred to [Apr+23a].

Just like the NV, as discussed in Section 5.2, the PMTs of the TPC can produce afterpulse signals which can create fake S1s. Like for the NV, these signals usually have a different signal shape than regular S1 signals, and thus they can be rejected using a **cut\_s1\_width** cut. The cut is optimized based on  $^{37}\text{Ar}$  data, as it plays a role in particular for small signals.

In contrast to the S2 single scatter cut, the **cut\_s1\_single\_scatter** cut does not compare the size of the main and alternative S1. Instead, the cut tests whether the main S2 could be successfully paired with any other preceding S1 peak in a given event. Successful means in this context, that the new alternative S1 and S2 pair must satisfy the S2 width, the S1 area fraction top, and the S1 max PMT cuts. If this is the case the event is rejected

The **cut\_s1\_pattern\_top** and **cut\_s1\_pattern\_bottom** cuts work in principle the same way as the **cut\_s2\_pattern** cut. They compare the observed hit pattern for a given S1 with simulated hit patterns based on the S1 area and position. The only difference is, that the S1 cuts are based on a likelihood-ratio test.

In addition to these data quality cuts, the WIMP search uses extra anti-AC cuts. They are shortly explained in Section 8.3.

## A.5 S2-ONLY ANALYSIS USING THE NR CALIBRATION DATA

In this small section, the basic calibration idea of the cS2-only analysis is sketched. For regular analyses the detector threshold for small WIMP masses is limited by the S1 threshold of the detector. To overcome this issue a so-called S2-only analysis can be performed, dropping the requirement of having an S1 signal as done in [Apr+19a]. However, since no S1 signal is being recorded the drift time is missing as an important information carrier. One can recover this information by exploiting the fact that the S2 width depends on the electron diffusion during drift and therefore on the drift time itself. However, in the past, this effect could only be calibrated at higher energies where S1 signals are present. Using the new calibration approach discussed in Chapter 6 of this work, one can substitute the timestamp of the S1 with the timestamp of the tagged 4.44 MeV  $\gamma$ -ray as a surrogate as sketched in Figure A.15. This approach re-

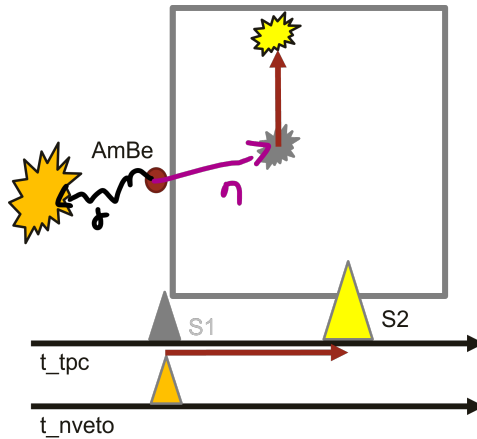


Figure A.15. – Sketch of the neutron-veto aided S2-only NR calibration. Only the signal in the neutron-veto and the S2 in the TPC are recorded.

quires a much larger coincidence window between TPC and neutron veto, resulting in more accidental coincidence between the two detectors. A first proof of concept of this calibration idea is shown in Figure A.16 One can see a population of events following a drift time relation similarly to the one at higher energies. However, given that most of the NR calibration data is found at very shallow depths close to the gate, as shown in Figure 6.4, the majority of the calibration events are below the orange dashed line which is dominated by other background events. Hence, to make use of this method a second AmBe calibration at lower  $z$  with higher statistics is required. In addition, to the S2-only drift time correlation, the cut acceptance of other data quality cuts can be studied for the S2-only analysis.

Further, in this way, it might be possible to measure the Migdal effect [Bel+22] which was never measured before in a LXe TPC. The Migdal effect is an inelastic scattering process between a WIMP and xenon-nucleus in which additional energy is transferred to the shell electrons of xenon, such one of its outer shell electrons gets ionized. During the rearrangement of the electronic shell and filling of the vacancies, additional

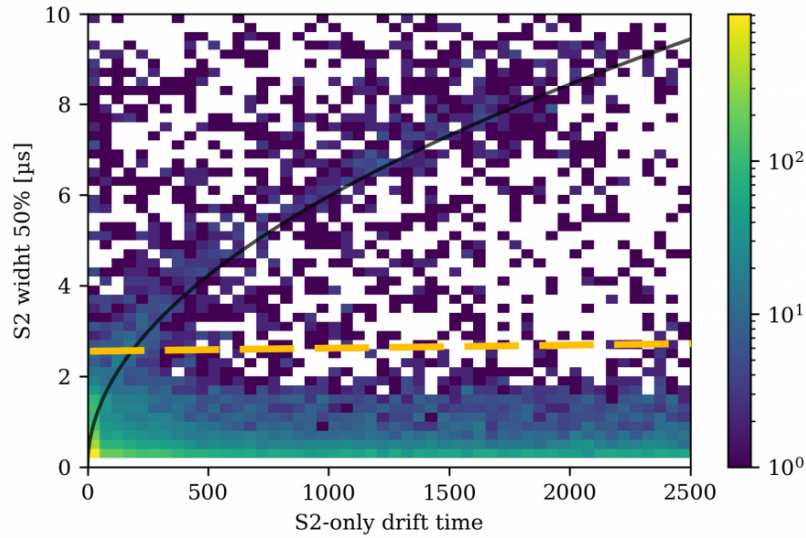


Figure A.16. – Illustration of the S2 width versus “S2-only drift time” relation as explained in the text. The binned histogram shows all events of the AmBe calibration data discussed in Chapter 6 for which no S1 signal was recorded. The black line indicates the S2 width as a function drift time based on other calibrations at higher energies. Events below the orange dashed line shows few electron background signals which are randomly paired with neutron-veto triggering events.

energy is deposited in the detector through a low energetic ER. Hence, this more efficient way of transferring energy from a WIMP to the detector allows to lower the threshold of the experiment and therefore allows to search for smaller WIMP masses. However, the Migdal effect is predicted to be suppressed by several orders of magnitude compared to regular elastic recoils [Bel+22], which makes it difficult to measure. To prove the existence of the Migdal effect for xenon, usually, a calibration with low energetic neutrons with kinetic energy below the detection threshold of the detector is suggested.

However, here I propose a different approach, using tagged S2-only signals with AmBe. In this approach, we exploit the fact that the charge yield of ER is much larger at low energies than for NR signals. This means, that S2-only signals produced by a Migdal event would exhibit much larger S2-only signals than produced by regular NR signals. Thus, the S2-only-spectrum of such a calibration would show a small tail at S2 areas. This excess of tail events must be then compared against the expected S2-only rate from ER background signals which can be done by a side band study of events that do not obey the S2-only width drift time relation of Figure A.16.

One big advantage of using tagged neutrons from, e.g. an AmBe source for such a measurement is the fact that the gamma tag already provides a strong tool to suppress backgrounds. In addition, the tag indicates that the observed signals are most likely induced by neutrons. Studying the effect for S2-only signals has, in addition, the advantage that the pollution of multi-scatter events is less of an issue as each scatter can be treated as a valid event, as long as no summed S1 signal is observed.



---

APPENDIX

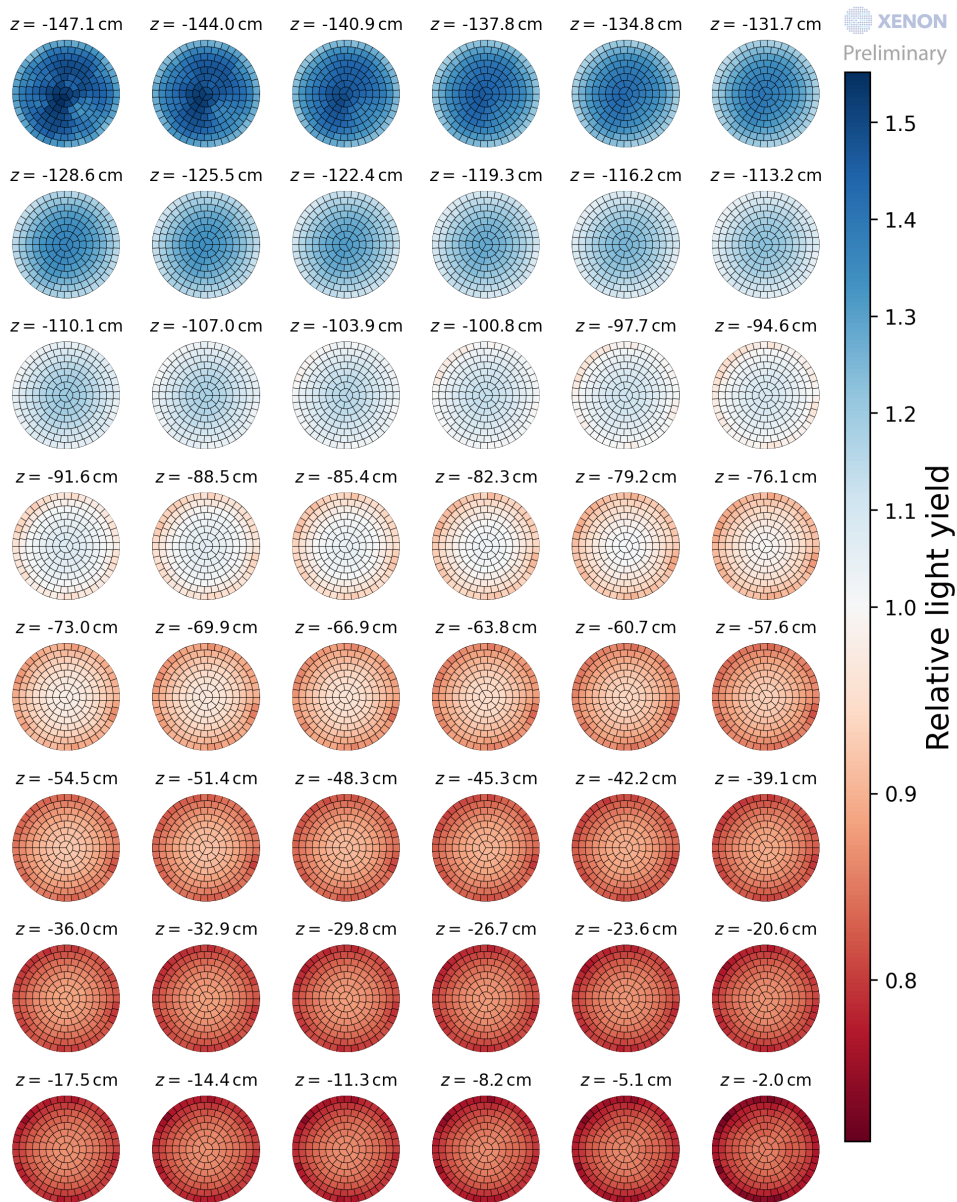


Figure B.1. – S1 correction map of XENONnT's first science run. The different circles show the average S1 response for a different z-slice. The color scale shows the relative S1 area normalized to the detector average. (Credit Johanna Jacobs)

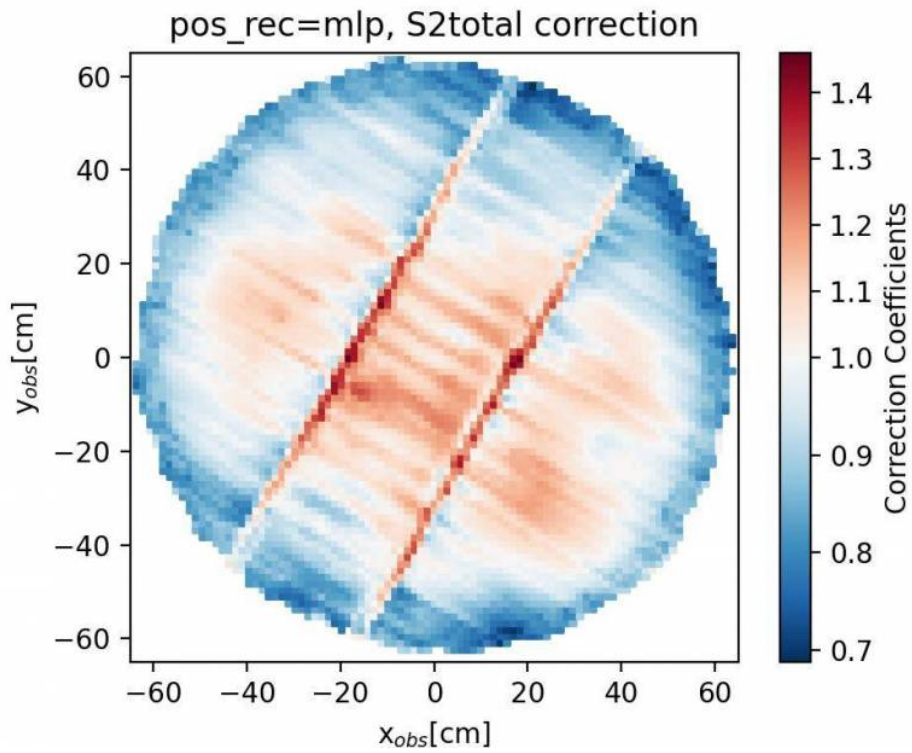


Figure B.2. – S2 correction map of XENONnT’s first science run. The color scale shows the relative S2 area normalized to the detector average. The two diagonal stripe like regions are due to field distortions near the perpendicular wire explained in Section 2.4.1. (Credit Jianyu Long)

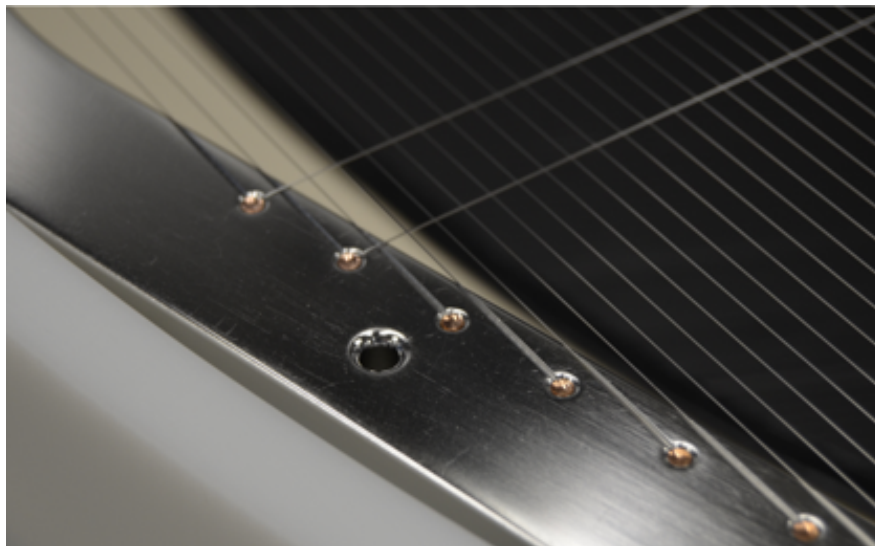


Figure B.3. – Photograph of the anode electrode showing the transverse wires.

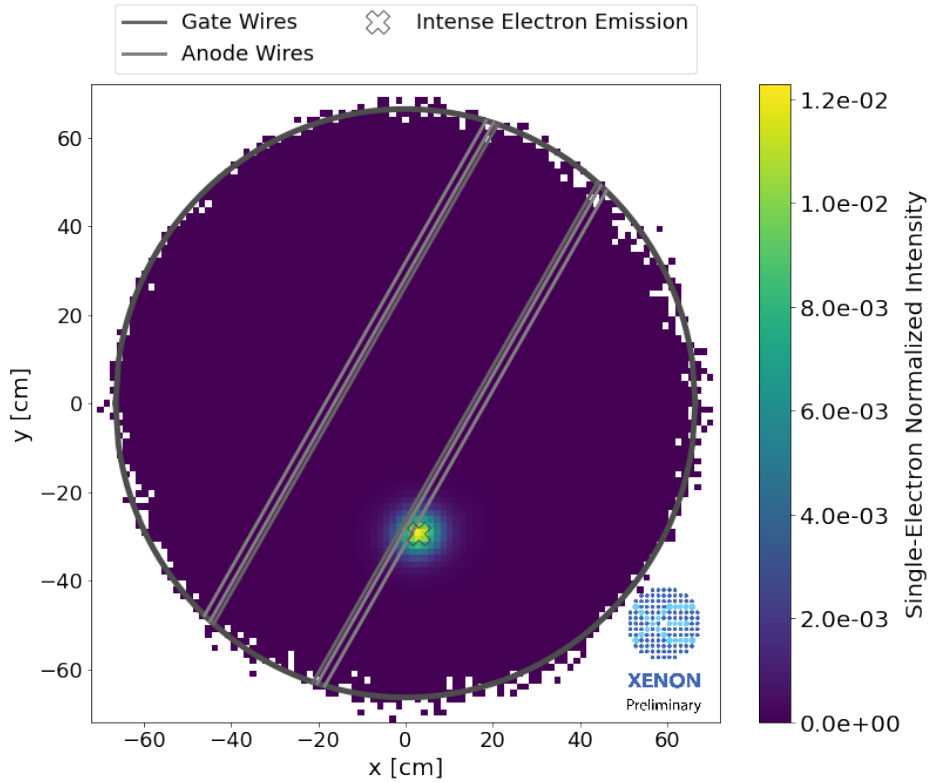


Figure B.4. – Localized region of an intermittent high emission of single electron S2 single in XENONnT.

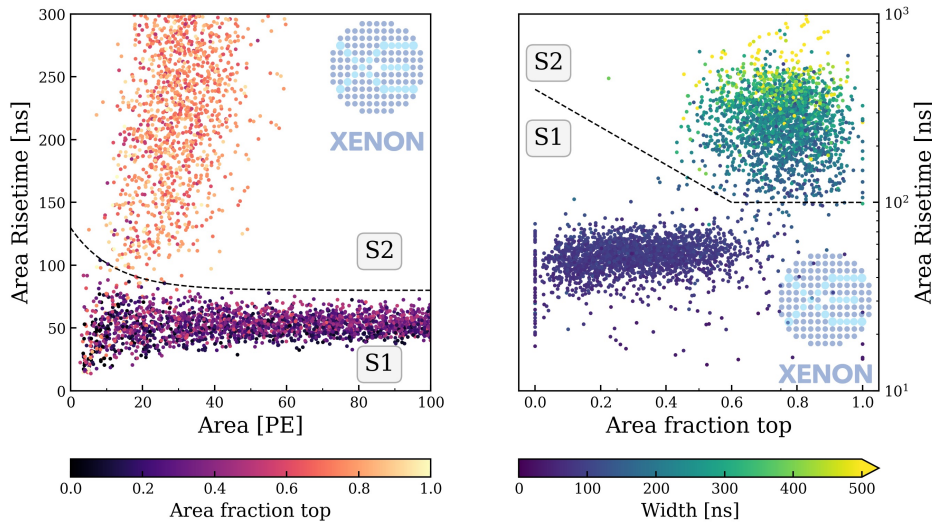


Figure B.5. – Peak(let) classification boundaries for S1 and S2 in SR0. The left plot shows the peak area rise time as a function of peak area. The color scale represents the area fraction top of the corresponding peak. The dashed line indicates the classification boundary. The data shown are S1 signals from the  $^{220}\text{Rn}$  calibration and single electron S2 signals. The plot on the right hand side shows the same data, but as a function of area fraction top. The color scale here indicates the central 50% width of the integrated area of the corresponding peak (Credit XENON collaboration).

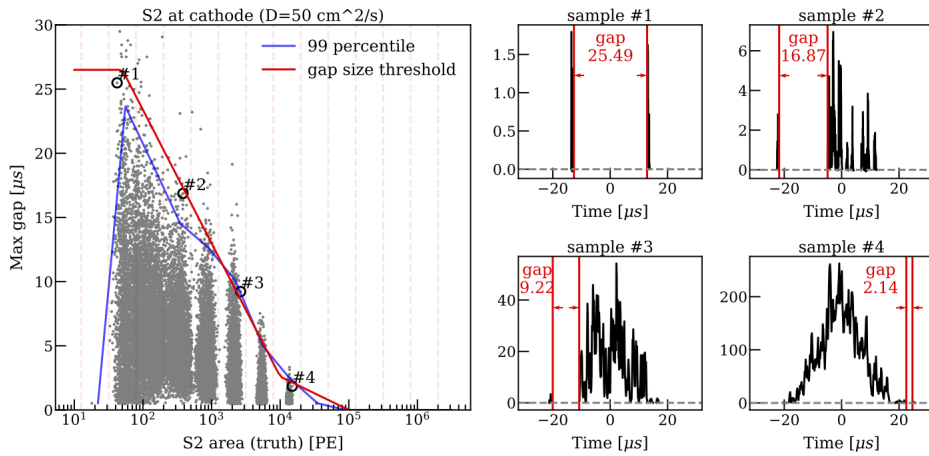


Figure B.6. – Maximum gap threshold for the S2 peaklet merging as function of the S2 area. The plot on the left hand side shows simulated S2 data uniformly distributed in  $xy$  close to the cathode of the detector. The vertical diffusion of the electron cloud is simulated with the diffusion coefficient  $D$  presented in the title of the plot. The merging threshold as a function of S2 area size is shown as red line. Four example waveforms of signal which will be merged are shown on the right hand side. The corresponding events are highlighted by black circles in the plot on the left hand side (Credit XENON collaboration).

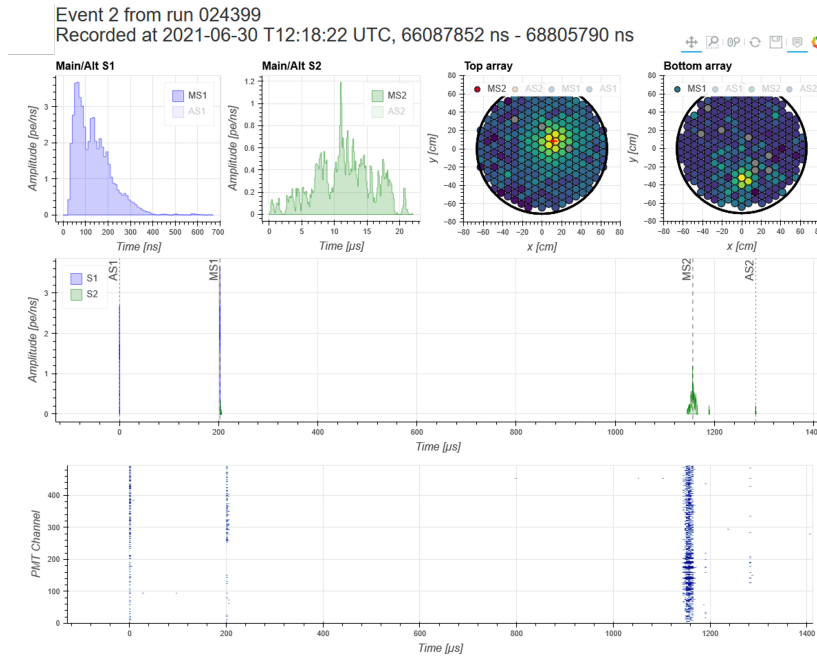


Figure B.7. – Interactive TPC event display as explained in section 3.5.1 showing the optional record matrix in the bottom panel. The record matrix shows the raw data recorded in each PMT channel as a function of time. The matrix is plotted using the datashader function provided by the `holoviews` package. The datashader renders point clouds which are partially overlapping or are too small to be resolved as points of different color intensity. The higher the intensity the darker the color. The shown event is a calibration event taken during one of the SR0  $^{83m}\text{Kr}$  calibration. By zooming in the main panel the raw records point cloud will be rendered until the full length of each individual raw\_record is shown.

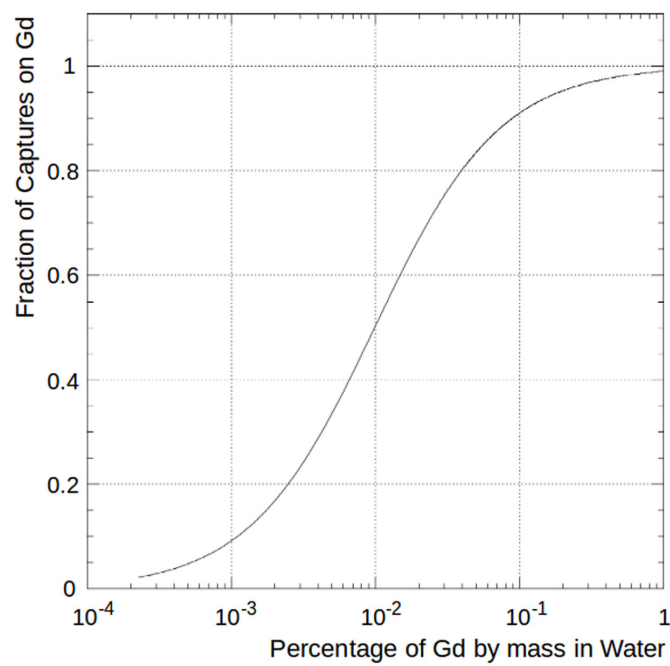


Figure B.8. – Fraction of Gd-neutron capture for a gd-loaded water Cherenkov detector as a function of Gd-mass fraction. The figure was taken from [Mar+20].



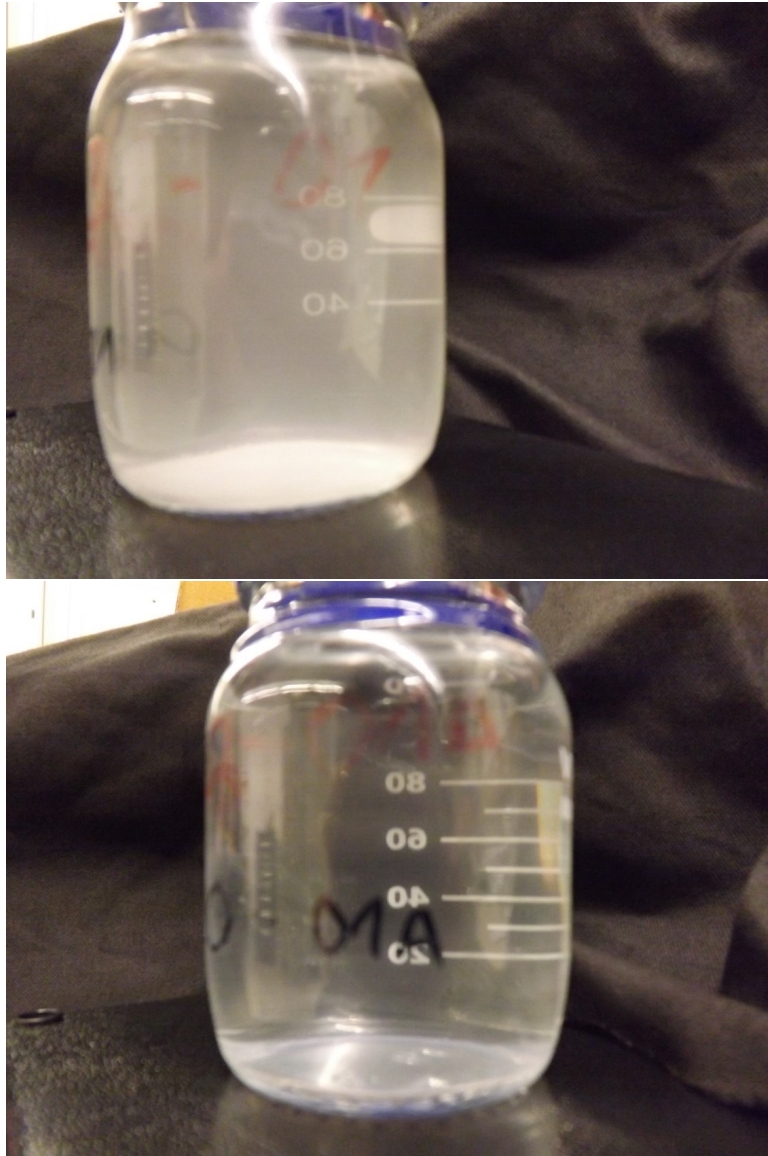


Figure B.9. – Images of the Gd20-01 (top) and Gd20-01A (bottom) salt samples from Treibacher. Both images show a Gd solution with 0.2% concentration.



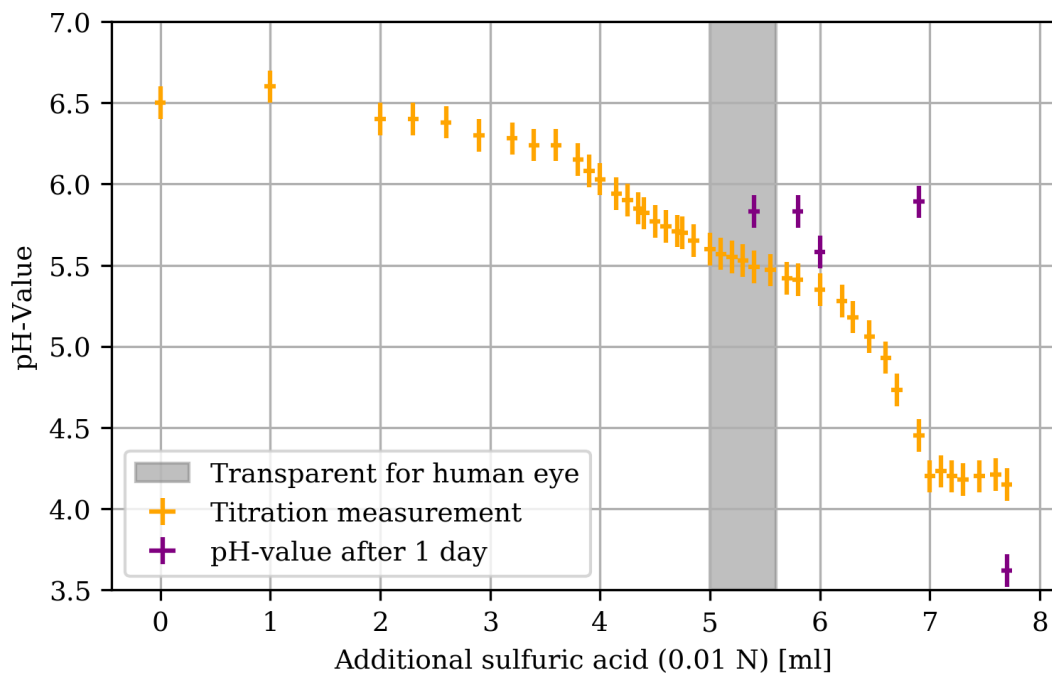


Figure B.10. – Titration of a Gd-loaded water sample with sulfuric acid. The sample has a  $(0.20 \pm 0.01)$  % Gadolinium concentration and a 0.01 N sulfuric acid was used. The orange data points are measured during the titration ~5 min after adding a droplet of sulfuric acid. The purple data points are additional samples which were prepared separately and measured one day after adding the sulfuric acid. The gray shaded regions indicates the region in which the solution became transparent to the human eye.

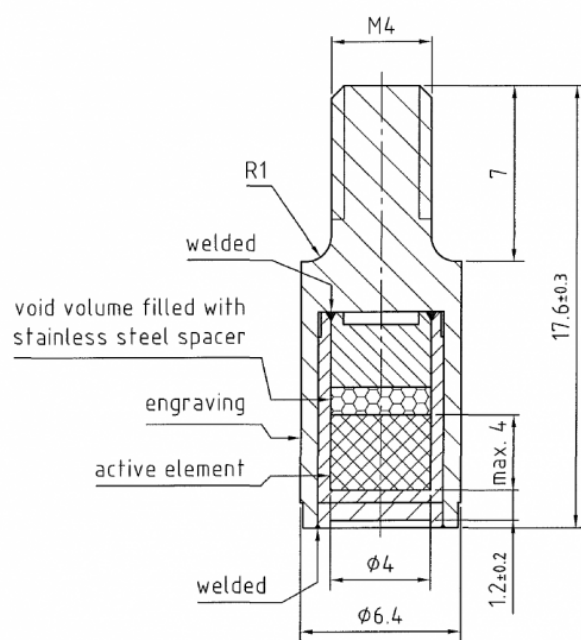


Figure B.11. – Technical drawing of the AmBe source used in XENONnT.

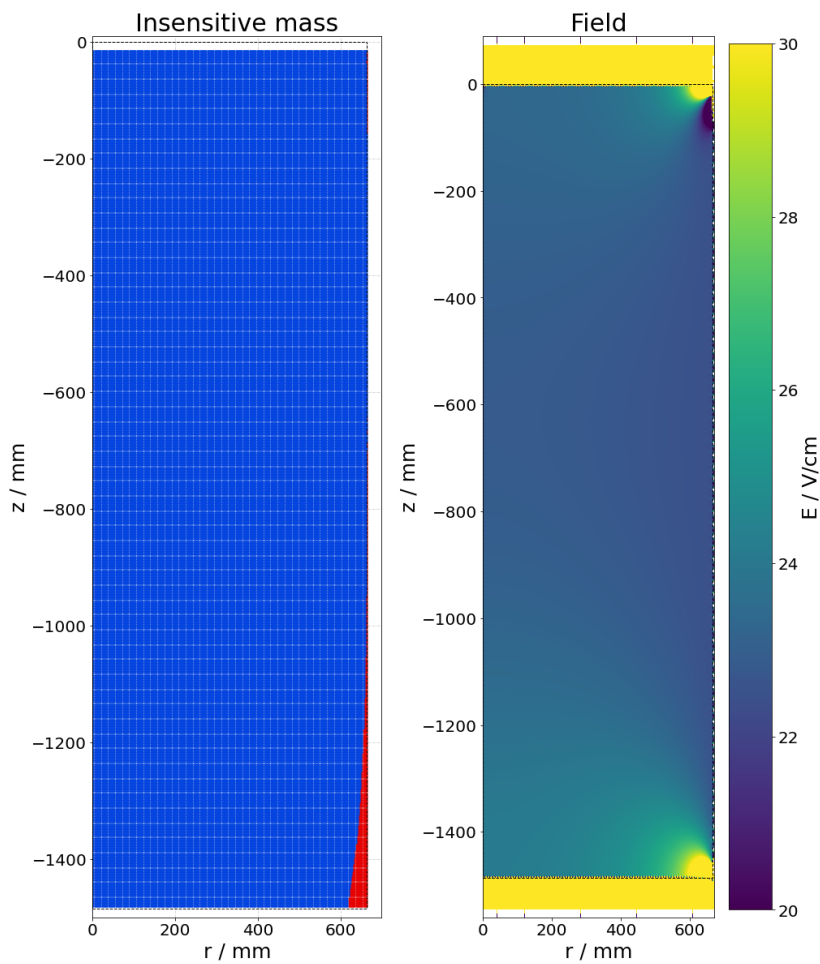


Figure B.12. – XENONnT field simulations and charge insensitive mass. The figure on the left shows the region in which electrons are fully absorbed during the drift process in red. The figure on the right hand side shows the simulated drift field, please note the different z-scale. Both figures are based on simulations with COMSOL. (Credit: Francesco Toschi from Freiburg University for the XENON collaboration).

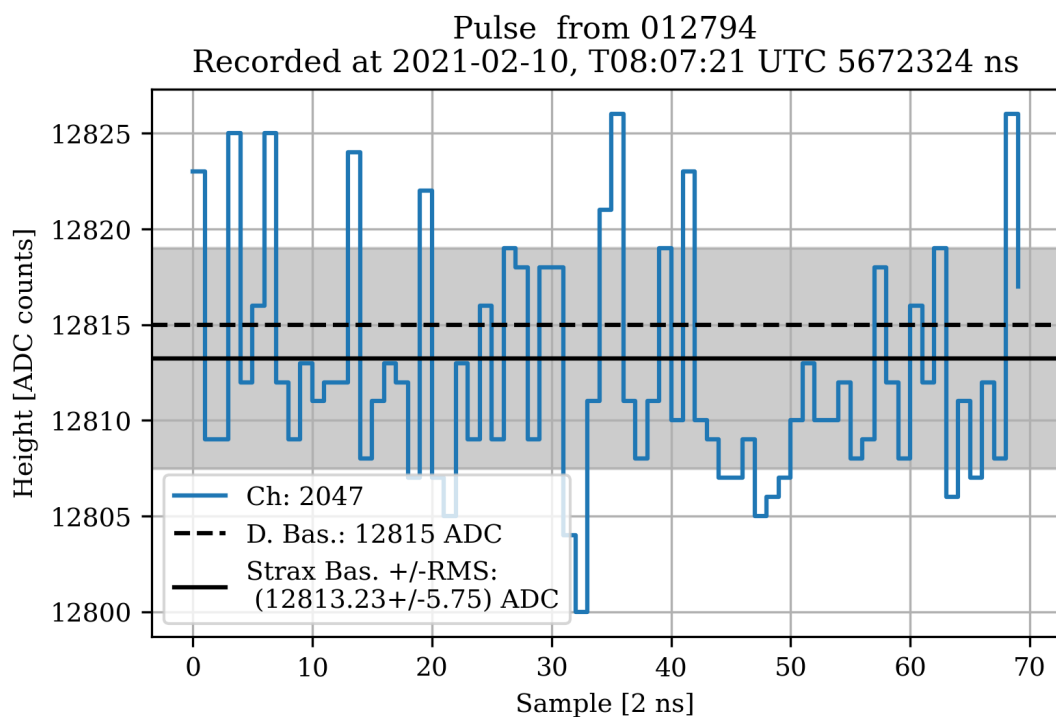


Figure B.13. – Example waveform of the NV containing digitizer bit-flips. The waveform was recorded during the commissioning phase of the detector. The bit-flips appear as single binned large upward fluctuations of the baseline. The gray shaded region indicates the baseline RMS.

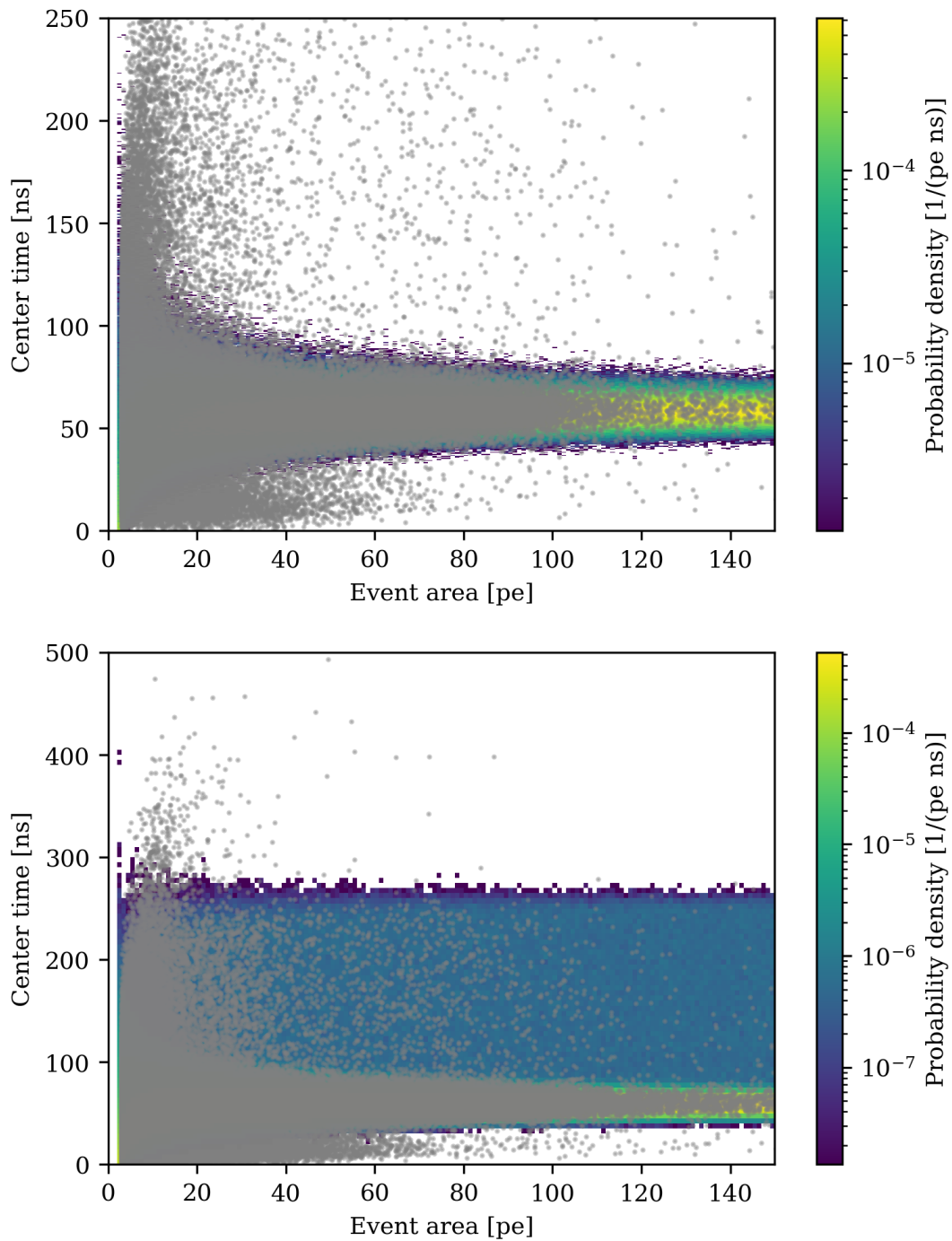


Figure B.14. – Qualitative comparison of the center time distribution between toy MC data and AmBe calibration data. Both plots show a comparison between toy MC data as binned histogram and AmBe calibration data in gray. The upper plot, shows toy MC events which were simulated based on an exponential distribution. The lower figure shows the same, but also adding uniformly distributed PMT dark counts to the toy events.

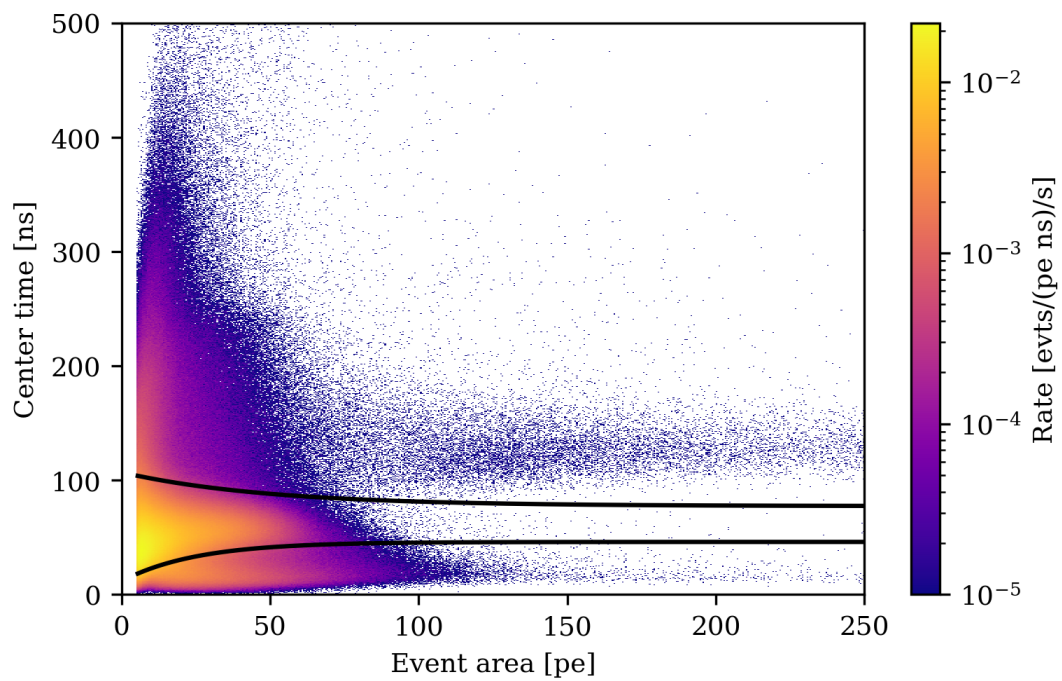


Figure B.15. – Center time cut boundaries compared against the NV SR0 background data. The black line indicate the upper and lower boundary of the center time cut. The histogram shows the distribution of NV background events which satisfy a 5-fold coincidence and a 5 pe event area threshold.

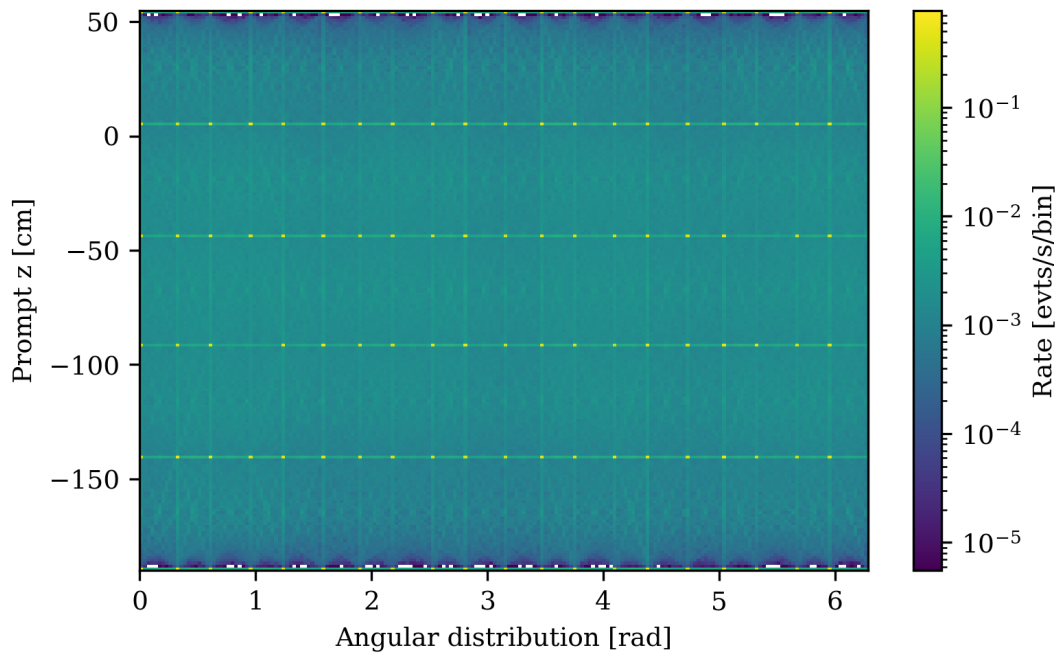


Figure B.16. – Reconstructed spatial position of NV background events. The plot shows the reconstructed  $z$ -position as a function of reconstructed angle. Only background events which satisfy the center time cut discussed in Section 5.2 are shown. The strip pattern is an artifact of the reconstruction method. The bins which show the highest number of events indicate the position of the 118 PMTs which are considered in the analysis.

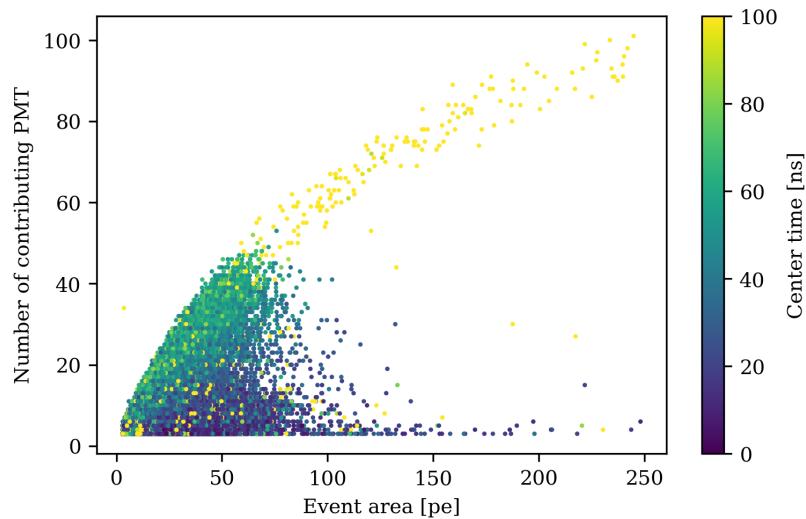


Figure B.17. – Number of contributing PMTs as a function of event area for SR0 NV background events. The color scale shows the center time parameter. At small number of contributing PMTs events of the population three discussed in Section 5.2, are located. The plot shows only a small selection of events.



Event 4 from run 031833 - 031831

Recorded at (2021-11-10 T07:34:05) UTC 558538398 ns - 558539156 ns

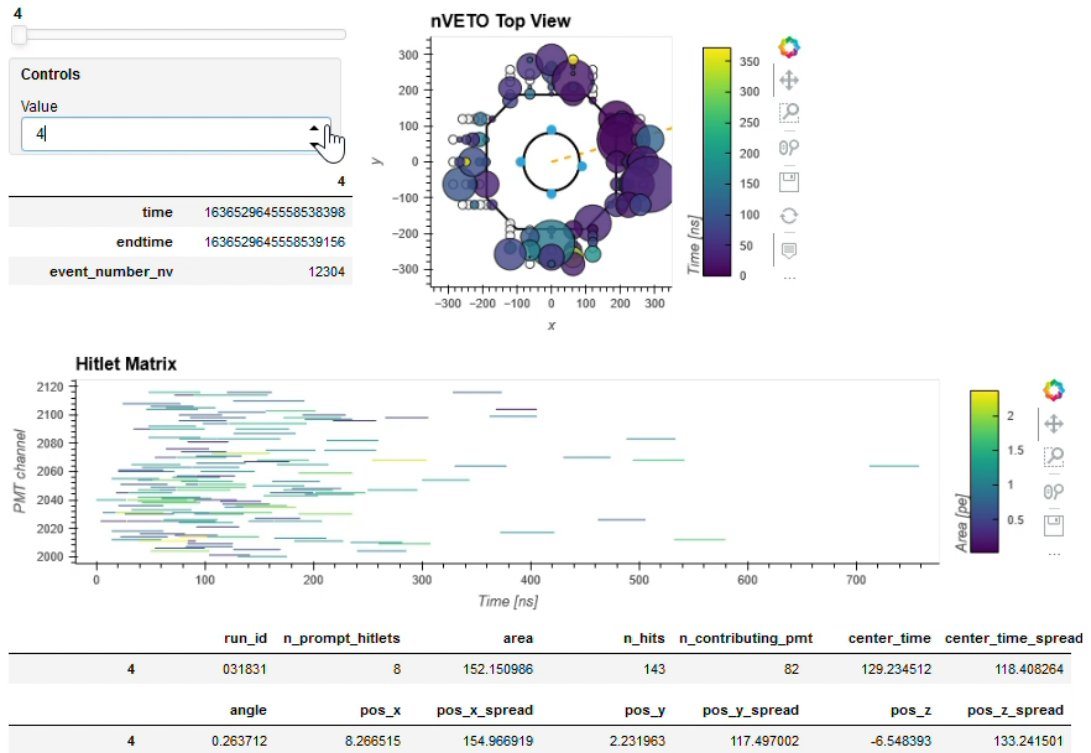


Figure B.18. – Hit pattern of a typical population 5 event discussed in Section 5.2.

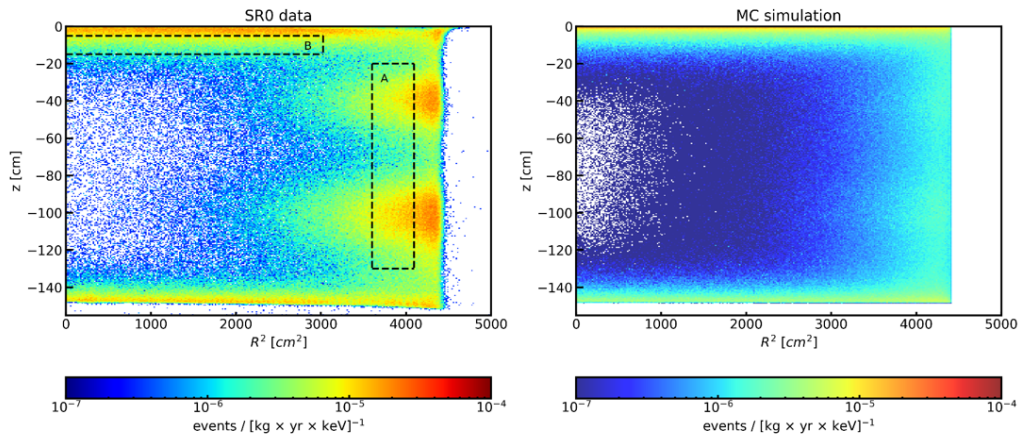


Figure B.19. – Comparison of the spatial distribution of the observed (left) and simulated (right) ER background of XENONnT. The materials background induced by the two stiffening rings can be clearly seen at  $z \approx -40$  cm and  $z \approx -100$  cm (Credit: Micha Weiss for XENON collaboration).

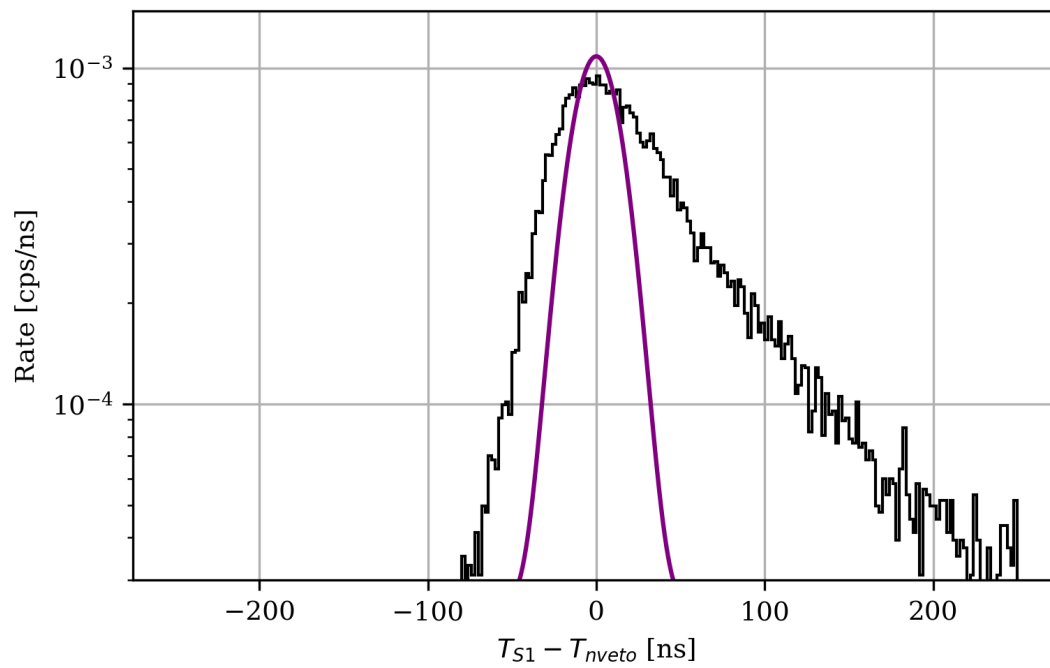


Figure B.20. – Comparison of the detector timing resolution of NR data and the high energy  $\gamma$ - $\gamma$ -coincidences. The black distribution shows the same data as in figure 6.7. The purple curve shows the best fit of figure 6.8.

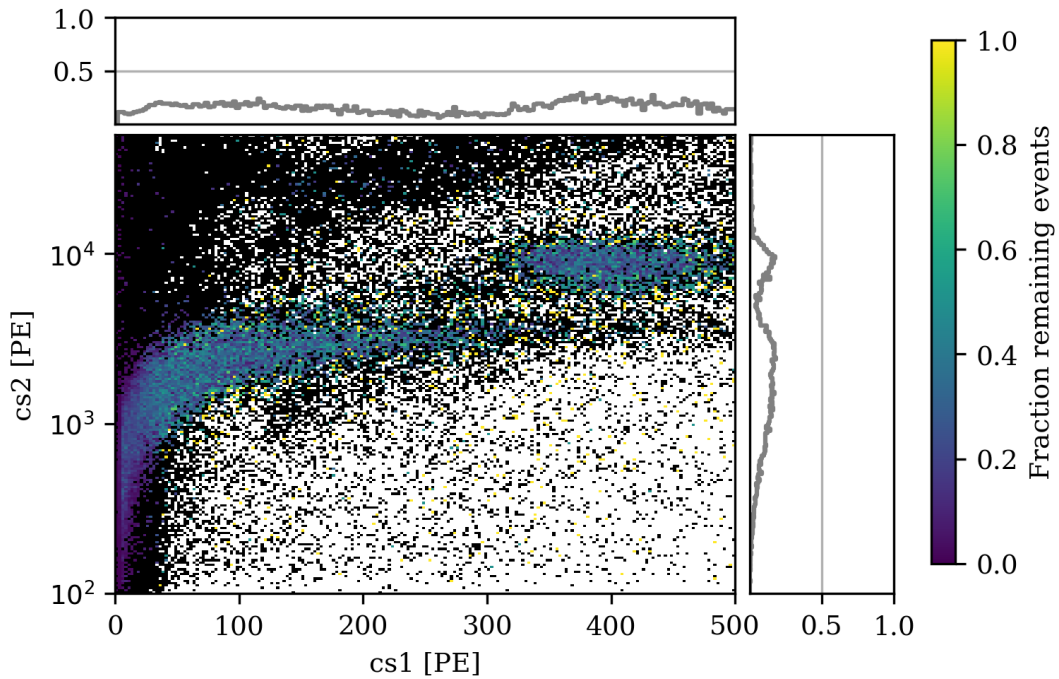


Figure B.21. – Fraction of remaining AmBe calibration events in the NR region after applying the NV-TPC coincidence. The colorcode encodes the fraction of remaining events. Bins displayed in black do not contain any events after applying the coincidence. The  $cS1/cS2$  projection show the remaining fraction of events projected on the respective axis. About 30% to 40% of all events within the NR band are kept.

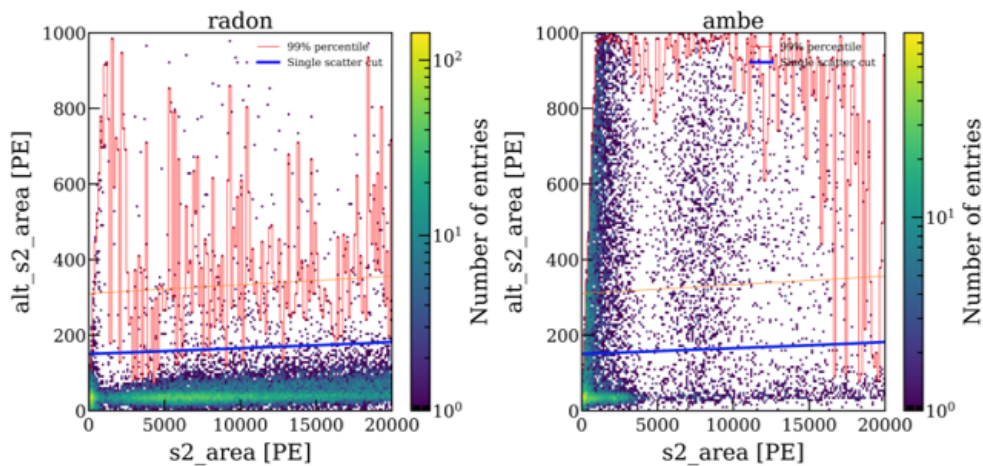


Figure B.22. – Cut boundaries for the S2 single-scatter cut. The cut boundaries were defined based on the  $^{220}\text{Rn}$  calibration data shown in the left plot. The plot on the right hand side shows the cut boundaries together with AmBe calibration data. The vertical distribution at small S2 areas is due to multi-scatter events. (Credit Shenyang Shi for the XENON collaboration)

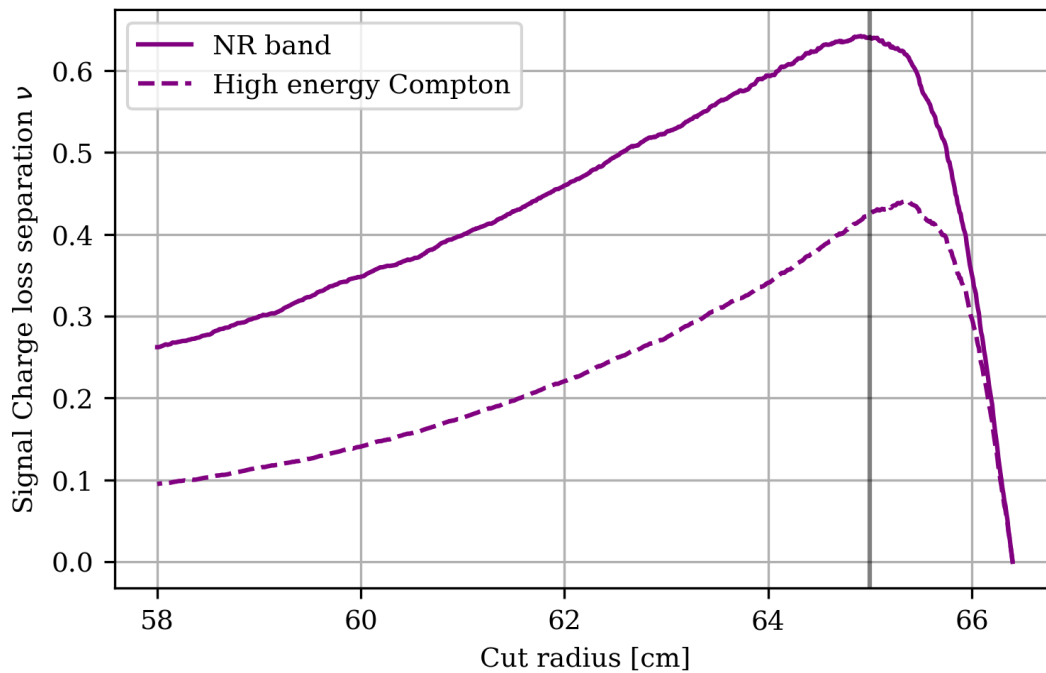


Figure B.23. – Signal to charge-loss separation as a function of the maximal radius for the AmBe fiducial volume cut. The solid purple line shows the difference between fraction of remaining NR and charge-loss events. The dashed line shows the same, but with Compton events of the ER band above the purple box shown in Figure 6.12 instead of events from the NR ROI region. The chosen maximum radius for the cylindrical vertical volume cut is highlighted by the vertical gray line.

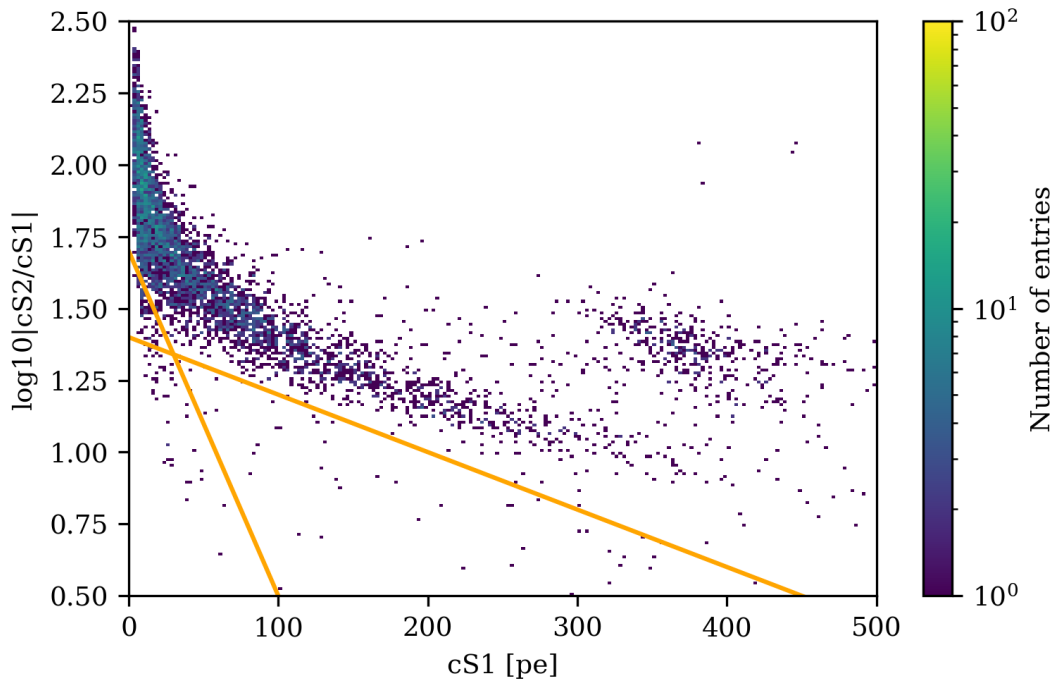


Figure B.24. – Selection of NR charge loss events selection. All events below the two orange lines are selected. The lines were adjusted by eye.

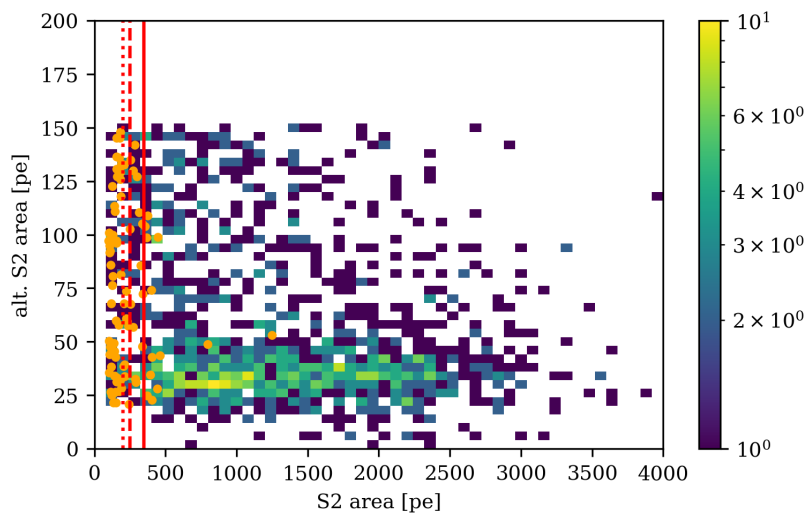


Figure B.25. – Remaining outlier events in the orange region shown in figure 6.15 shown in the  $alt. S2$  versus  $S2$  space. The orange dots are showing the outlier events. The binned data shows all AmBe calibration events in the NR ROI. The vertical red lines are drawn at 200 pe, 250 pe, and 350 pe respectively. Multi-scatter events are expected to follow a distribution where  $alt. S2$  and  $S2$  are of about the same size.

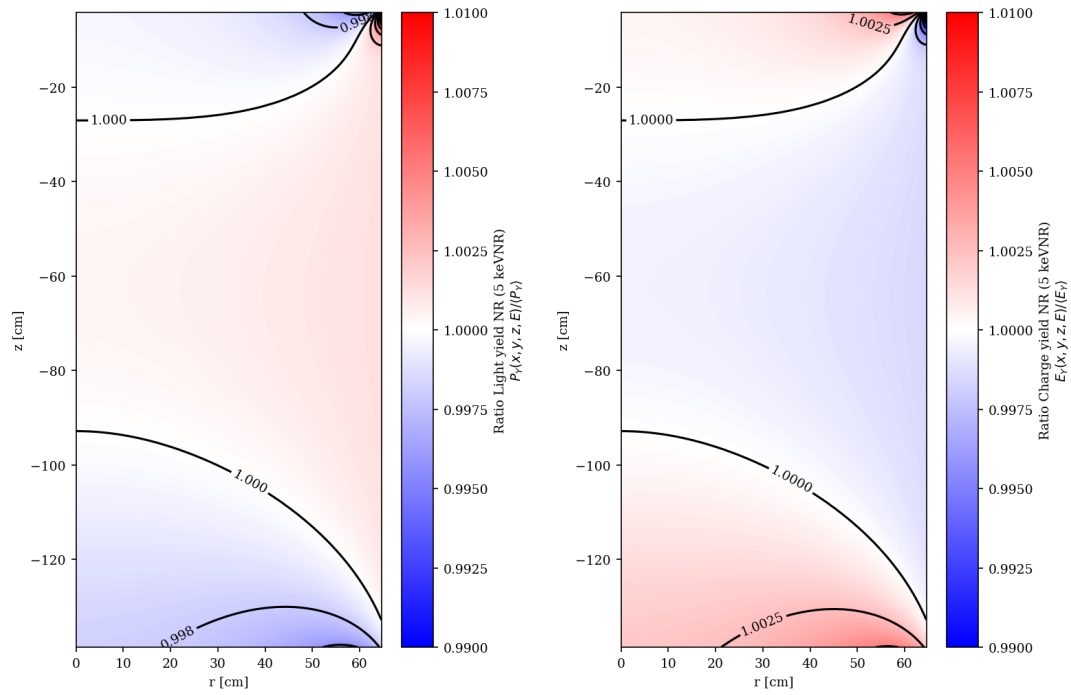


Figure B.26. – Relative difference in light yield and charge yield difference for 5 keV<sub>NR</sub> NR interactions. Both plots show the field dependence of the relative light and charge yield for a 5 keV<sub>NR</sub> NR interactions within the fiducial volume used in AmBe. Deviations from the average yields are less than one percent.

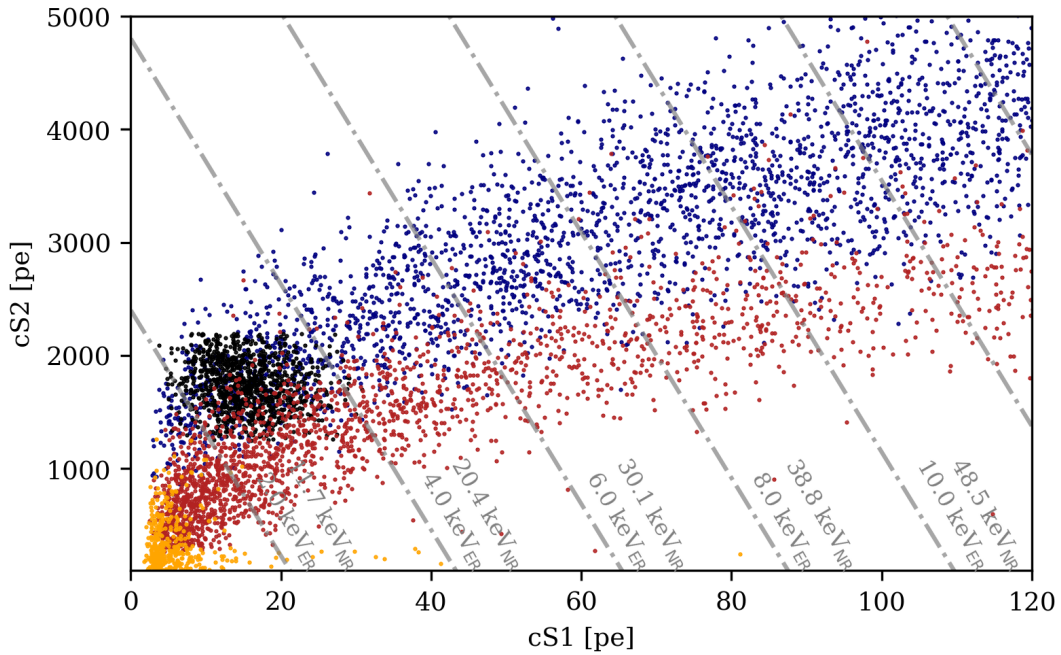


Figure B.27. – ER and NR calibration data passing all data quality cuts. The red dots show the remaining NR events including the S1 tight coincidence threshold and the S2 area threshold, while the orange points show the additional NR events when dropping these thresholds. The blue points show  $^{220}\text{Rn}$  ER calibration events, while the black dots indicate the mono-energetic 2.8 keV line of  $^{37}\text{Ar}$ . The gray dashed-dotted lines are showing the energy contours of the combined energy scale computed according to the NR response model described in Section 6.3.

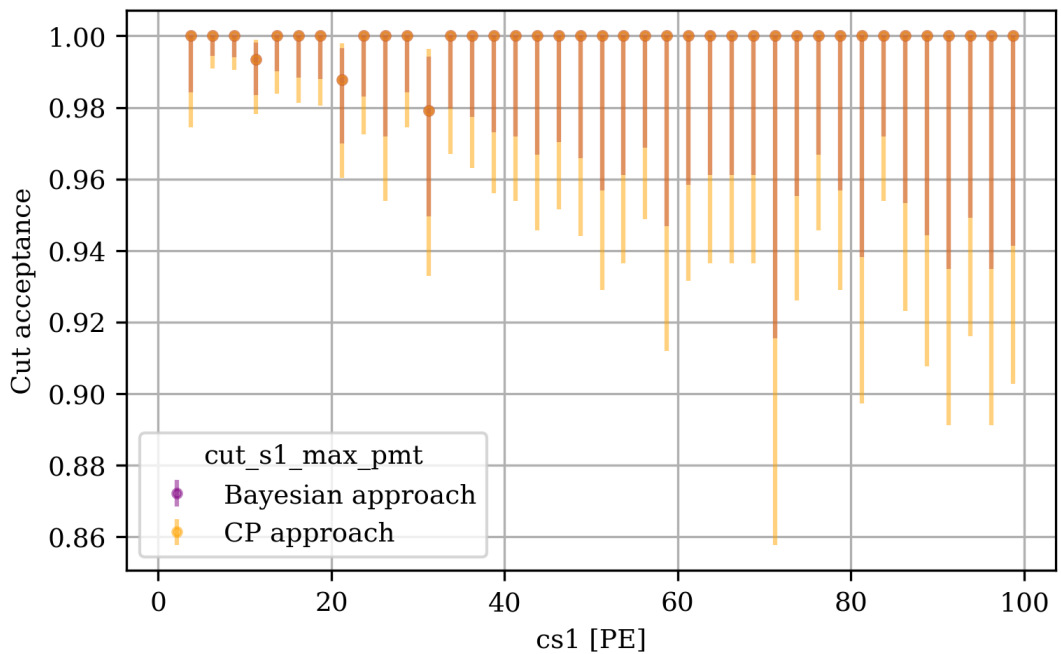


Figure B.28. –  $n-1$  method for the S1 max PMT cut.



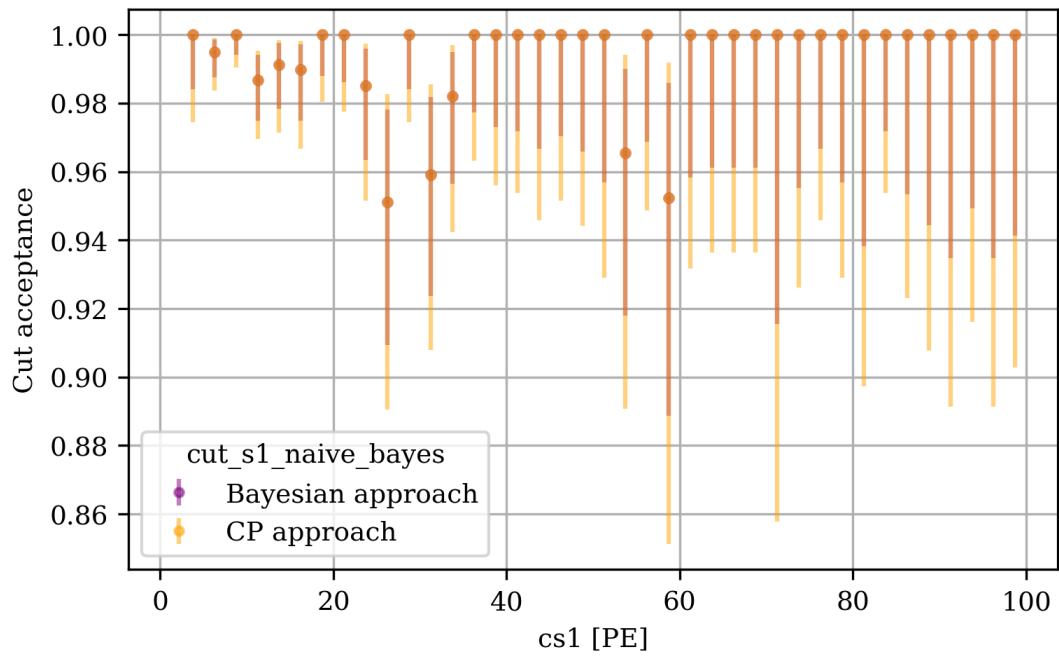


Figure B.29. –  $n-1$  method for the S1 naive bayes cut.

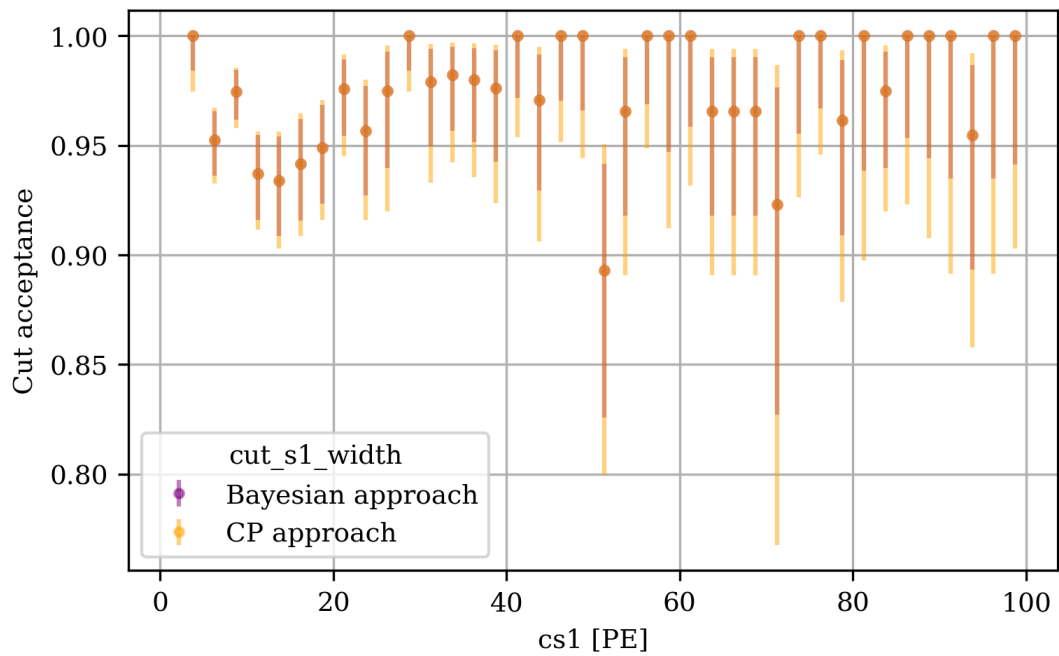


Figure B.30. –  $n-1$  method for the S1 width cut.

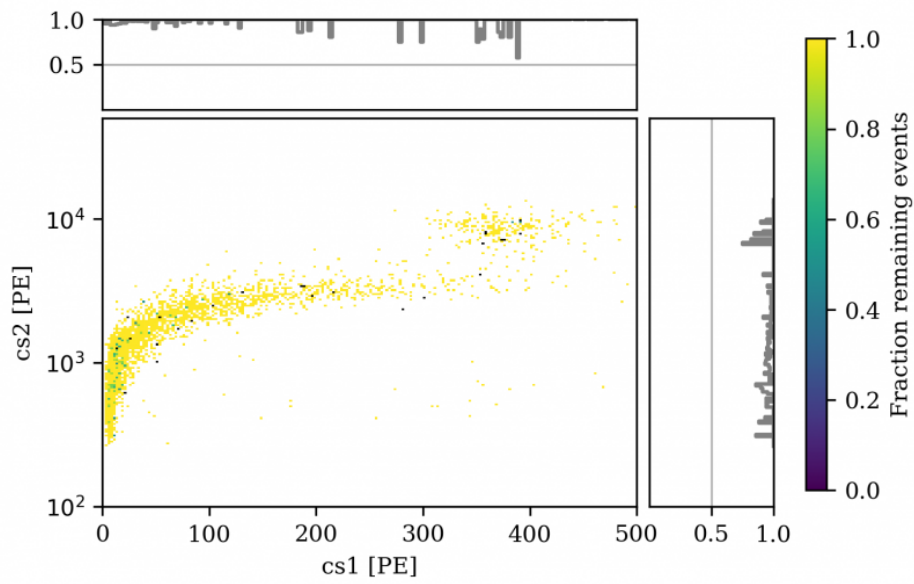


Figure B.31. –  $n-1$  method for the  $S1$  width cut. The plot shows the cut in the  $cs2$  versus  $cs1$  space. The color scale shows the remaining fraction of events according to the  $n-1$  method. The two panels on the side show the overall projection of the acceptance onto the given axis.

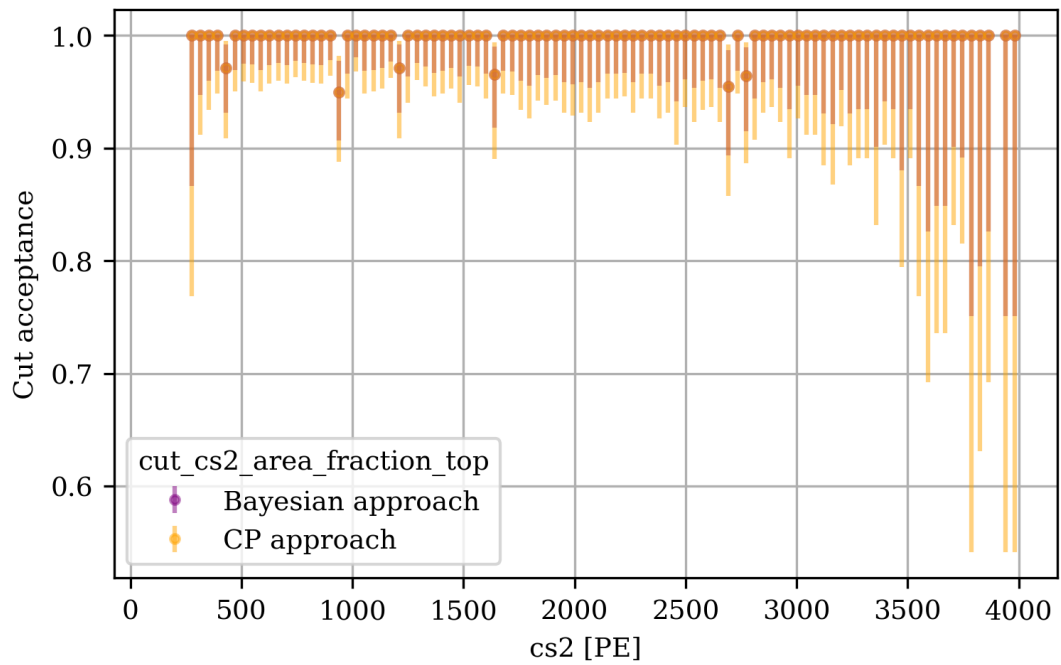


Figure B.32. –  $n-1$  method for the  $S2$  area fraction top cut.

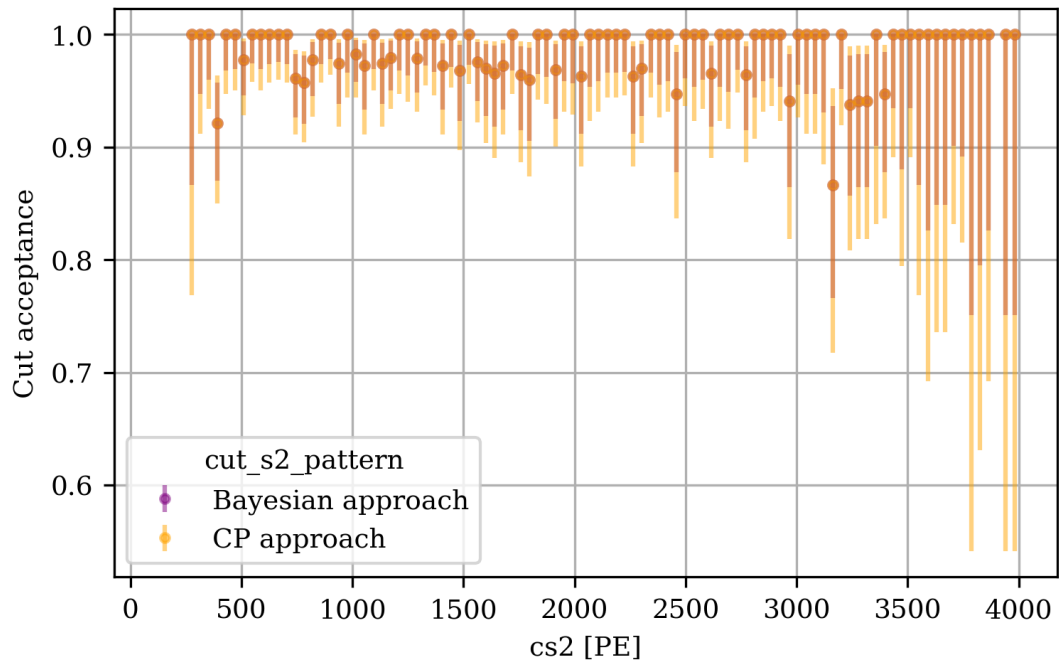


Figure B.33. –  $n-1$  method for the S2 pattern cut.

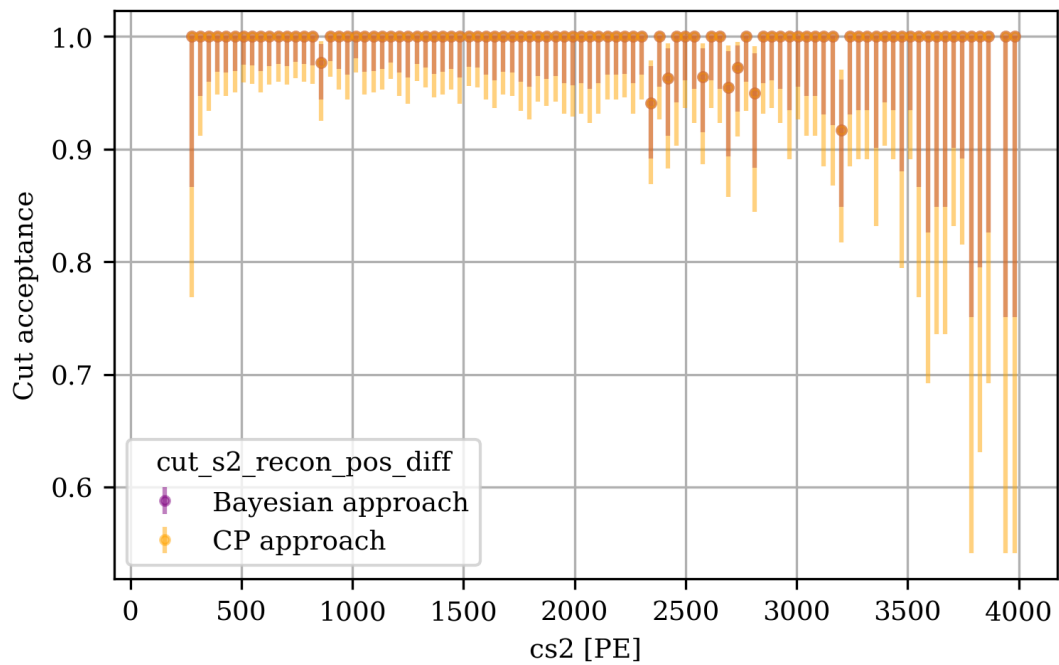


Figure B.34. –  $n-1$  method for the S2 poss diff cut.

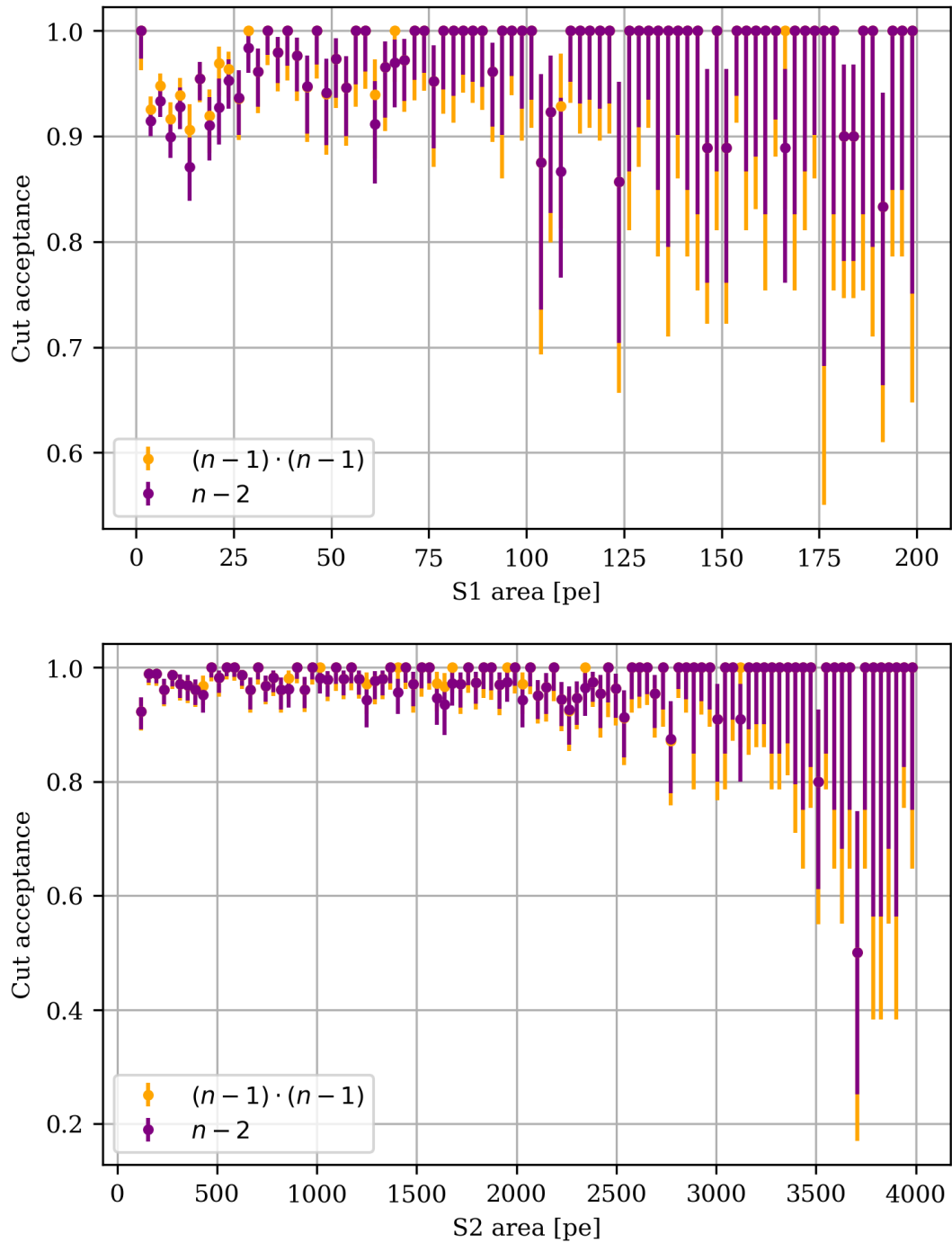


Figure B.35. – Cut acceptance comparison for the correlated cuts *cut\_s1\_naive\_bayes* and *cut\_s1\_width* as well as *cut\_s2\_pattern* and *cut\_s2\_recon\_pos\_diff*. The orange points are showing the product of the cut acceptance of the two individual cuts derived using the *n-1* method. The purple points show cut acceptance based on the *n-2* method explained in section 6.2.3.

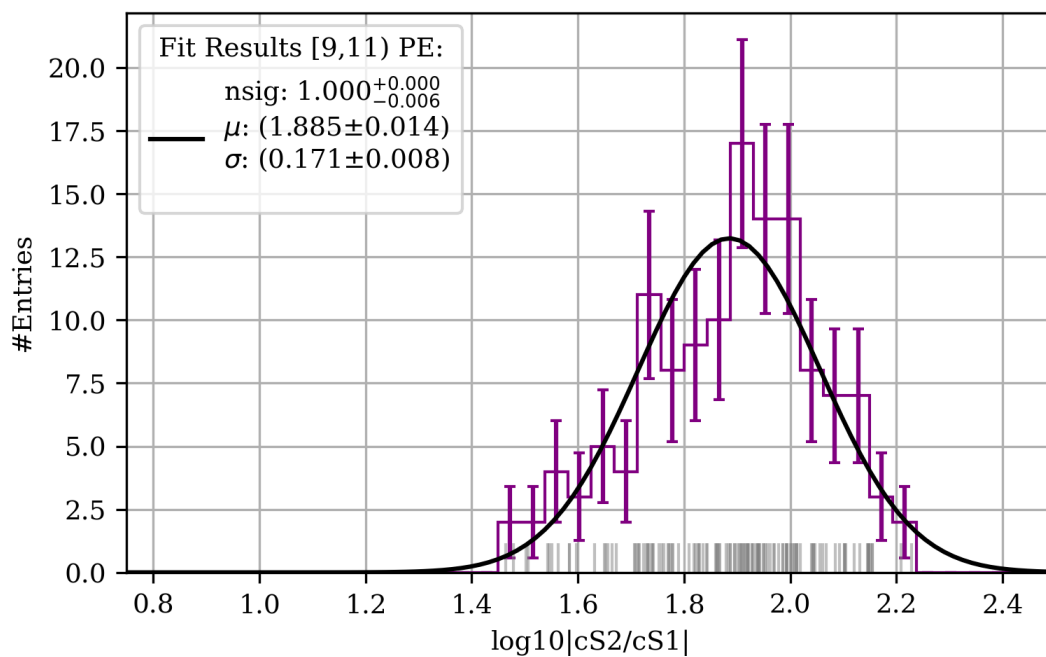


Figure B.36. – Single slice fit for the NR band contour estimate. The result of the best fit parameters are shown in black. Data points found in the corresponding cS1 slice are shown as gray vertical bars. A binned representation of the data is shown in purple. The uncertainties of the bins were evaluated based on Poisson statistics.

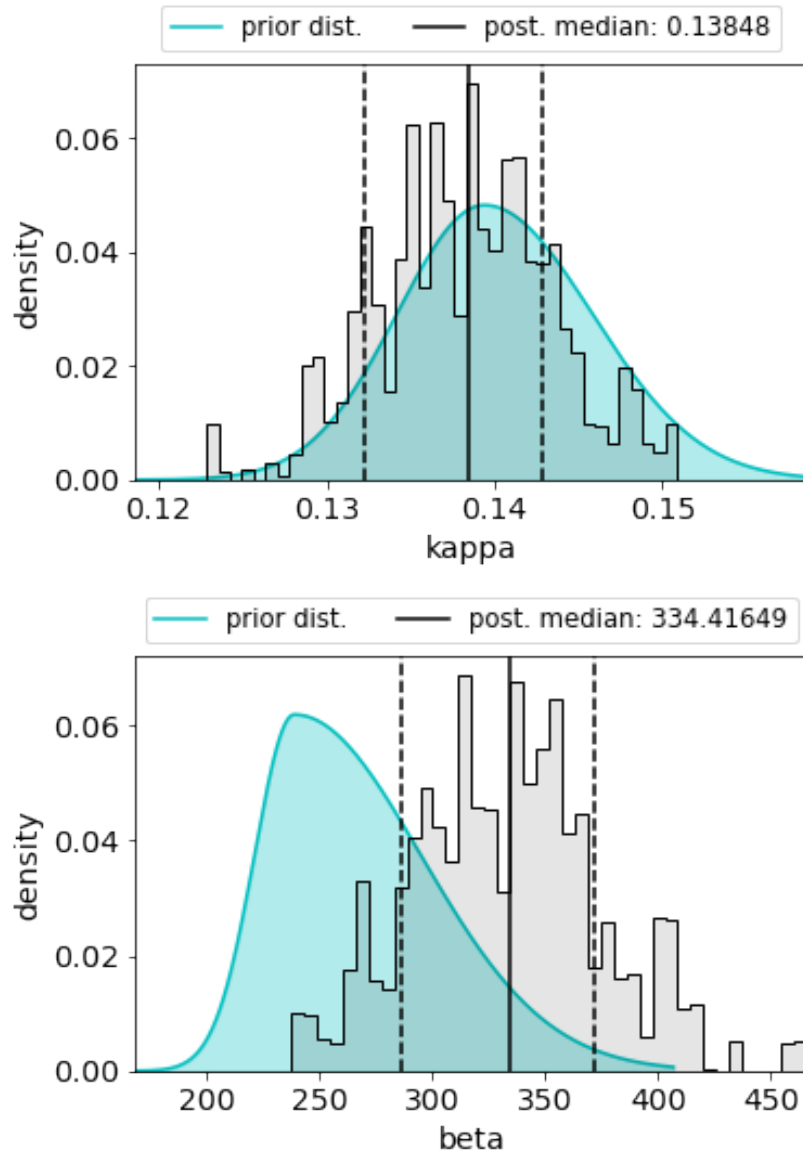


Figure B.37. – Posterior distribution of  $\kappa$  (top) and  $\beta$  bottom for the best fit of the NR response model. The prior distribution of the parameter is shown as cyan shaded region. The posterior as gray shaded histogram. The dashed lines are indicating the 15.865 and 84.135 percentile of the distribution. While, the calibration data has very little impact on the prior of  $\kappa$ , the posterior of  $\beta$  is quite different from its prior distribution. (Credit: Luisa Höttsch for the XENON collaboration,)

Table B.1. – Best fit values of the neutron-veto AmBe energy spectrum for the SR0 source positions.

	TopCW5d9m	TopCCW5d0m	BottomCW6d6m	BottomCW11d2m
$A_1$ [cps/pe]	$(108.2 \pm 0.5)$	$(116.5 \pm 0.5)$	$(121.2 \pm 0.5)$	$(117.3 \pm 0.6)$
$A_2$ [cps/pe]	$(53.9 \pm 0.4)$	$(54.3 \pm 0.4)$	$(59.5 \pm 0.3)$	$(55.4 \pm 0.5)$
$A_3$ [cps/pe]	$(4.03 \pm 0.02)$	$(4.13 \pm 0.03)$	$(4.08 \pm 0.02)$	$(4.17 \pm 0.03)$
$A_4$ [cps/pe]	$(1.35 \pm 0.02)$	$(1.18 \pm 0.02)$	$(1.32 \pm 0.02)$	$(1.28 \pm 0.03)$
$\mu_1$ [pe]	$(19.63 \pm 0.01)$	$(19.81 \pm 0.01)$	$(20.04 \pm 0.01)$	$(19.57 \pm 0.01)$
$\mu_2$ [pe]	$(64.3 \pm 0.1)$	$(64.6 \pm 0.1)$	$(66.2 \pm 0.1)$	$(64.6 \pm 0.1)$
$\mu_3$ [pe]	$(109.6 \pm 0.3)$	$(104.8 \pm 0.3)$	$(109.8 \pm 0.3)$	$(106.0 \pm 0.4)$
$\mu_4$ [pe]	$(151.7 \pm 0.3)$	$(148.7 \pm 0.4)$	$(153.0 \pm 0.3)$	$(149.8 \pm 0.4)$
$s$ [pe]	$(5.42 \pm 0.02)$	$(5.16 \pm 0.01)$	$(5.55 \pm 0.02)$	$(5.40 \pm 0.02)$
$s_0$ [pe <sup>2</sup> ]	$(0.00 \pm 0.01)$	$(0.00 \pm 0.01)$	$(0.00 \pm 0.01)$	$(0.00 \pm 0.02)$
$\mu_{th}$ [pe]	$(4.76 \pm 0.02)$	$(5.23 \pm 0.02)$	$(4.91 \pm 0.02)$	$(5.00 \pm 0.02)$
$\sigma_{th}$ [pe]	$(1.16 \pm 0.01)$	$(1.42 \pm 0.01)$	$(1.33 \pm 0.01)$	$(1.21 \pm 0.01)$
$a$ [cps/pe]	$(0.0 \pm 0.1)$	$(0.0 \pm 0.1)$	$(0.0 \pm 0.1)$	$(0.0 \pm 0.1)$
$b$ [cps/pe]	$(0.0 \pm 0.1)$	$(0.0 \pm 0.1)$	$(0.0 \pm 0.1)$	$(0.0 \pm 0.1)$
$c$ [cps]	$(0.11 \pm 0.01)$	$(0.24 \pm 0.01)$	$(0.06 \pm 0.01)$	$(0.19 \pm 0.02)$



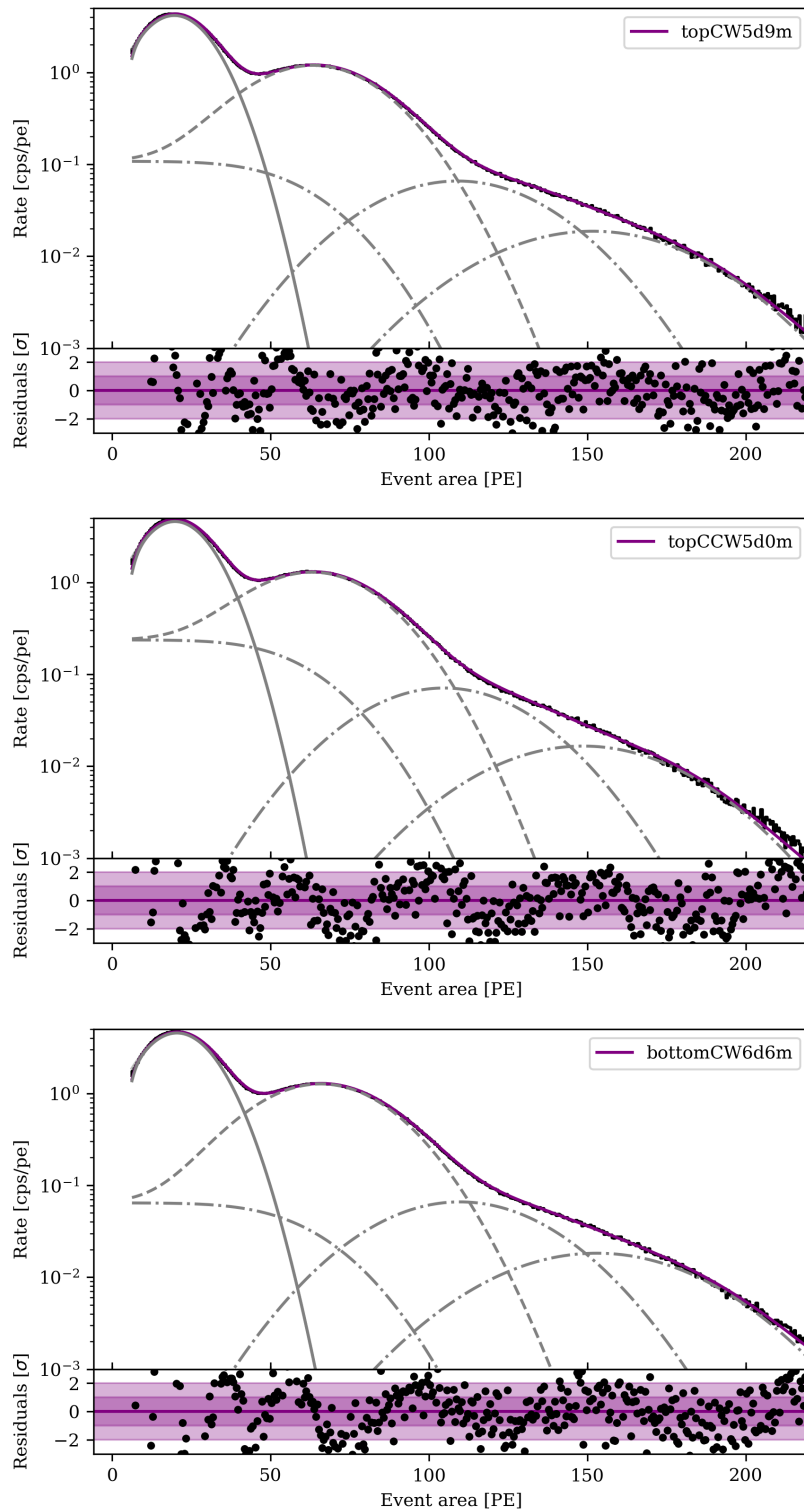


Figure B.38. – Best fit of the neutron-veto AmBe energy spectrum for the remaining source positions. The bottom panels show the fit residuals. The individual fit components are displayed in gray, as in Figure 7.2. Uncertainties are not shown.

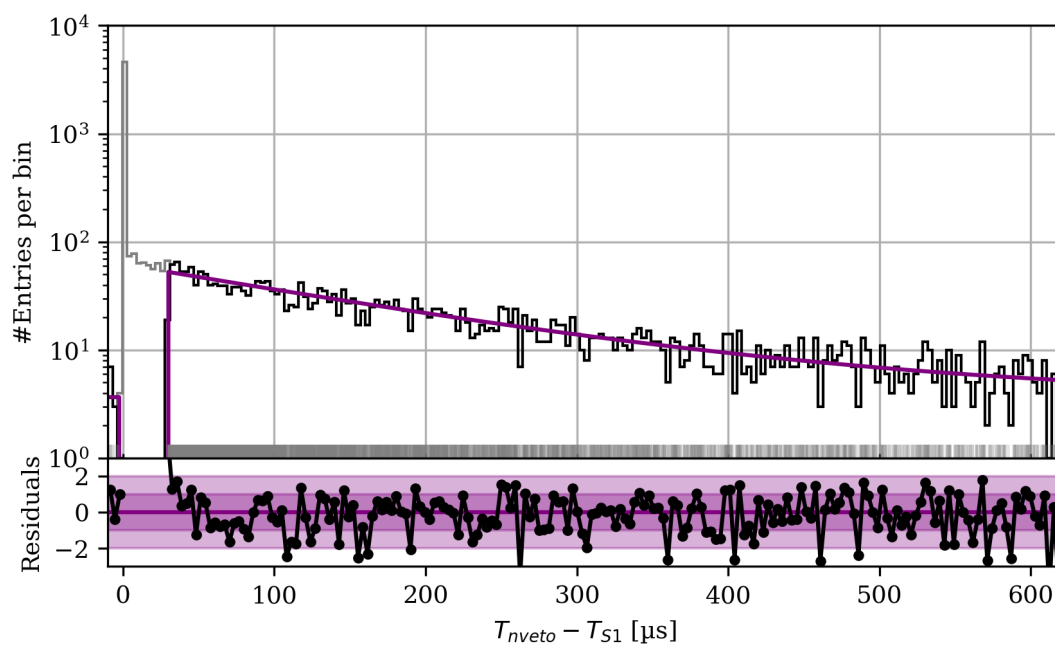


Figure B.39. – Best fit of the tagging efficiency time distribution including multi-scatter events. The figure shows the same as described in Figure 7.17, including multi-scatter NR signals. The decay constant of the exponential distribution was estimated to be  $(172 \pm 5) \mu s$ .

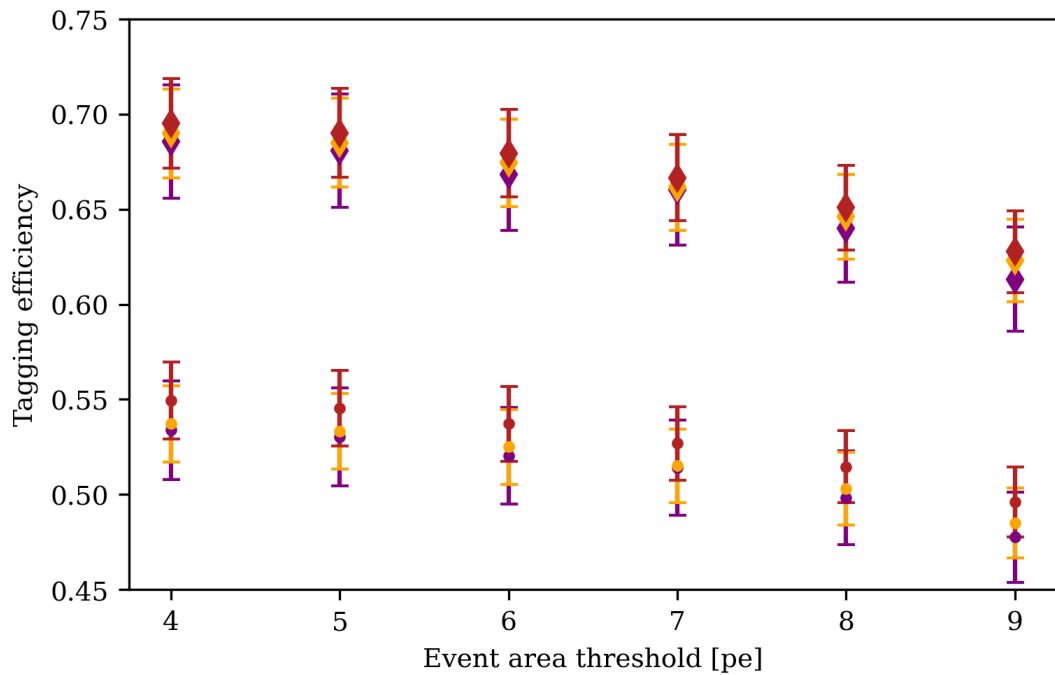


Figure B.40. – Comparison of the tagging efficiency w/ and w/o multi-scatter NR signals. The purple colored points show the tagging efficiency for the single scatter only case. Orange points show the tagging efficiency using the same correction factors as evaluated in the single scatter case, but allowing for multi-scatter events. The red points show the tagging efficiency of the multi-scatter case when evaluating again the time window correction factor according to the best fit of the time distribution shown in Figure B.39. Dots and diamonds show the tagging efficiency for a 250  $\mu\text{s}$  and a 600  $\mu\text{s}$  long tagging window respectively.

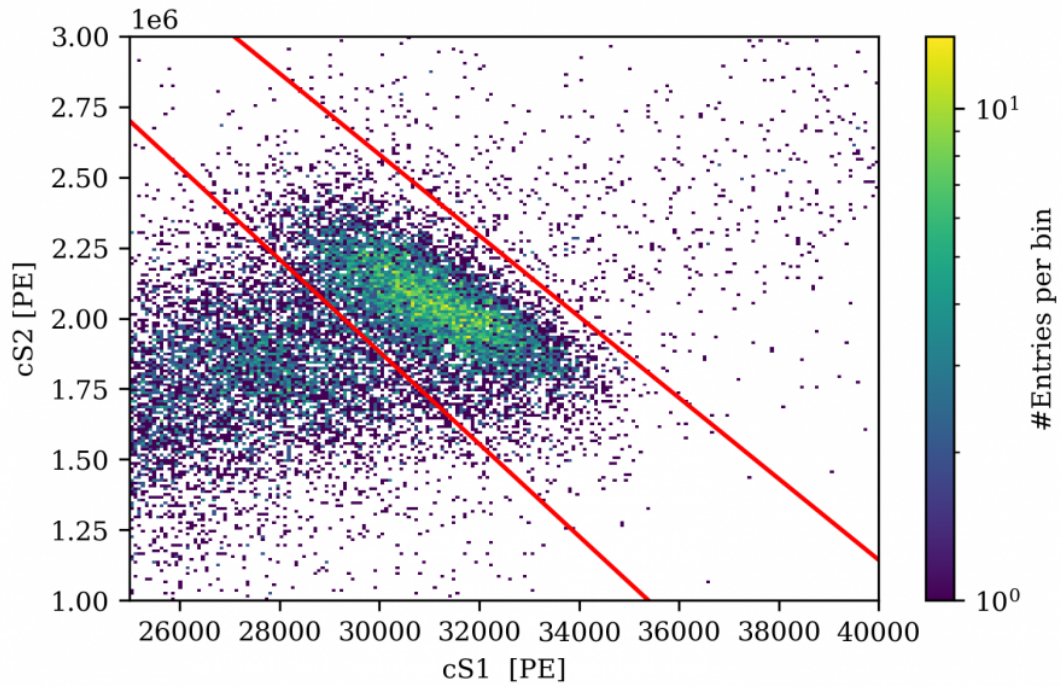


Figure B.41. – Fit range for the 4.44 MeV  $\gamma$ -ray ellipse fit. All data points within the two red lines were considered for the fit.

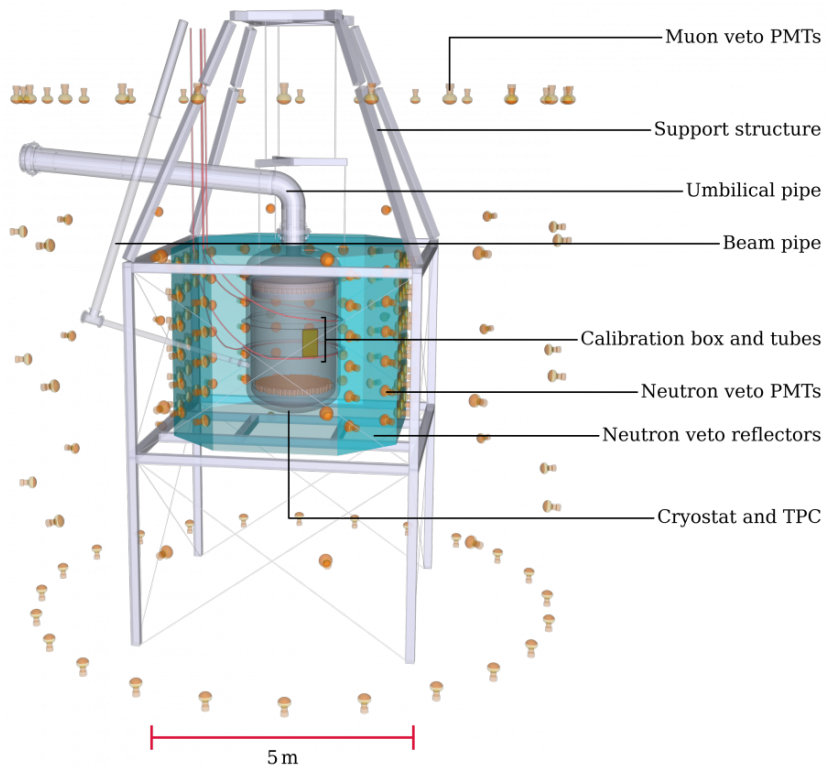
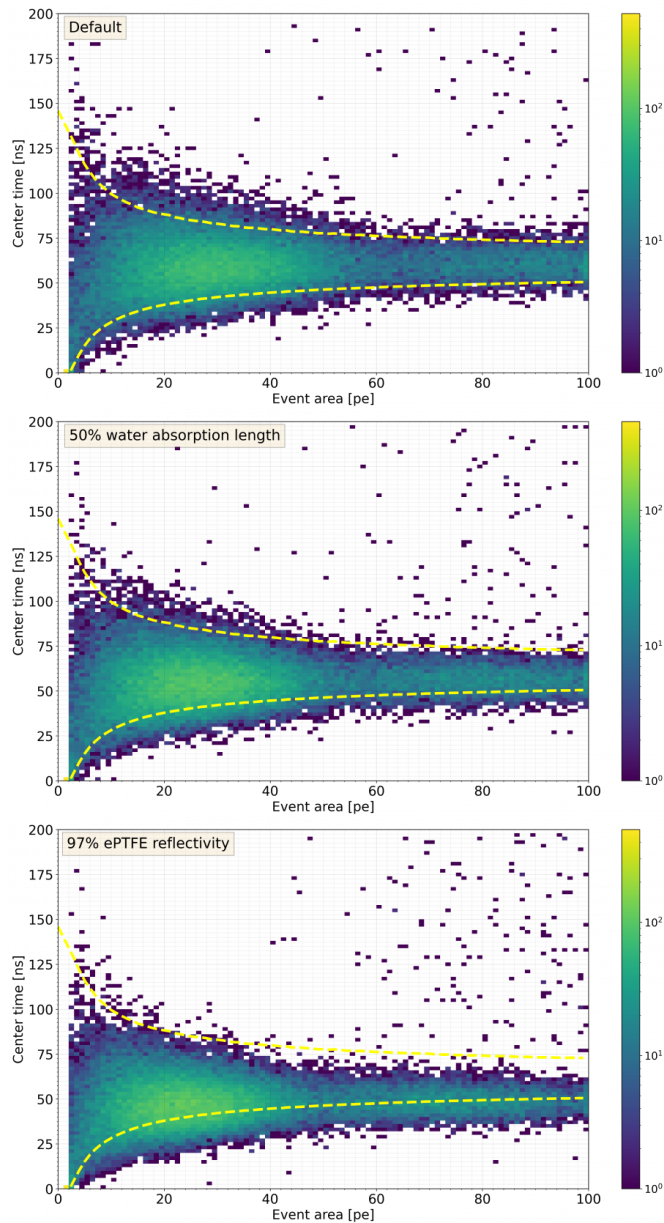


Figure B.42. – Rendering of the XENONnT experimental setup used in Geant4.



**Figure B.43.** – Impact of different optical properties on the neutron-veto center time. The histograms are showing the center time versus event area for some simulated AmBe events. The yellow dashed lines indicate the central 90% percentile derived from data. The upper plot shows the default MC configuration, the second plot shows a reduction of the water transparency by 50% of its default value. The last plot shows a reduction of the neutron-veto reflectivity by a factor 0.97. Credits to Pavel Kavrigin who is conducting this study.

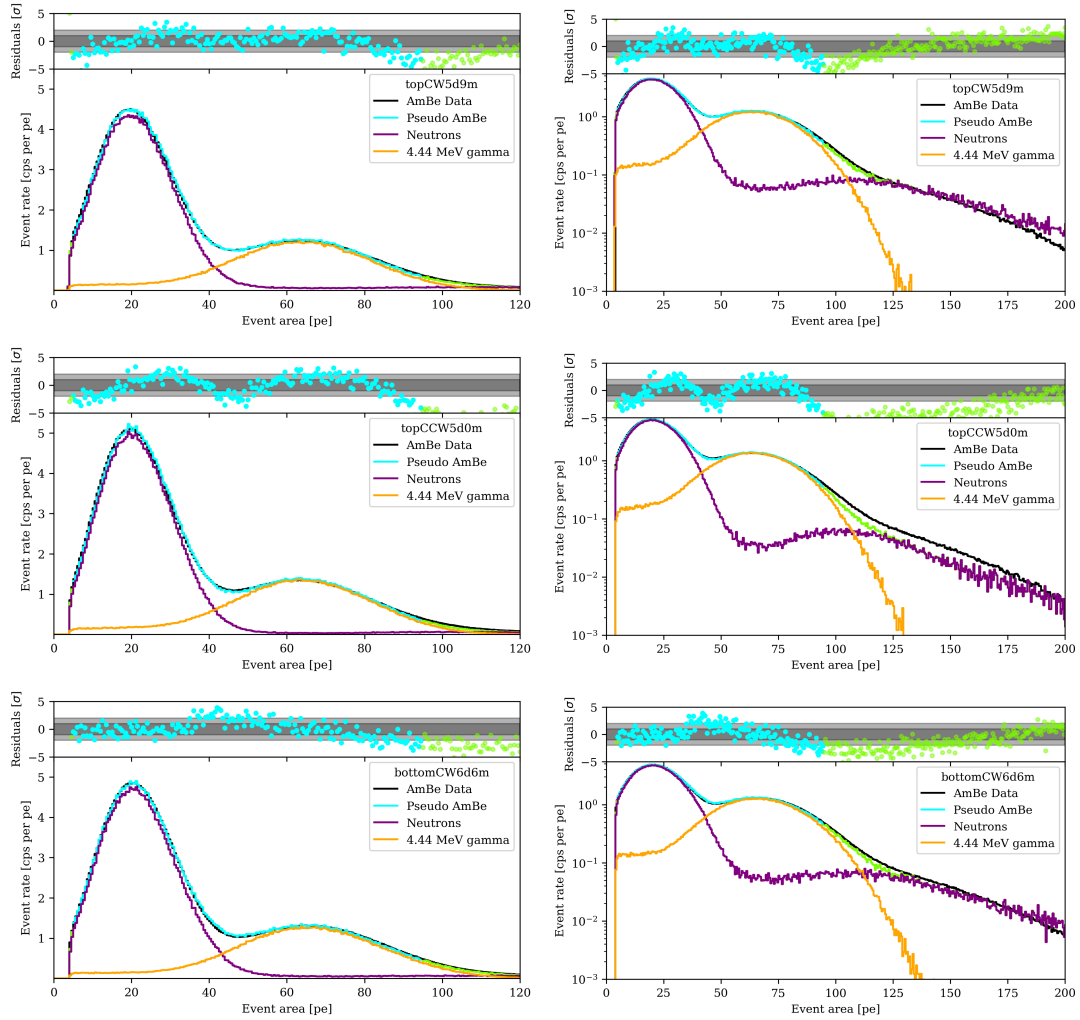


Figure B.44. – Best fits of the simulation templates to the AmBe spectra taken at different source positions. The left column of plots shows the best fit in a linear scaling, the right column in a log scaling with an extended fit range. The shown components are the same as in Figure 7.9. The rows are showing the top CW5d9m, top CCW5d0m and bottom CW6d6m respectively.

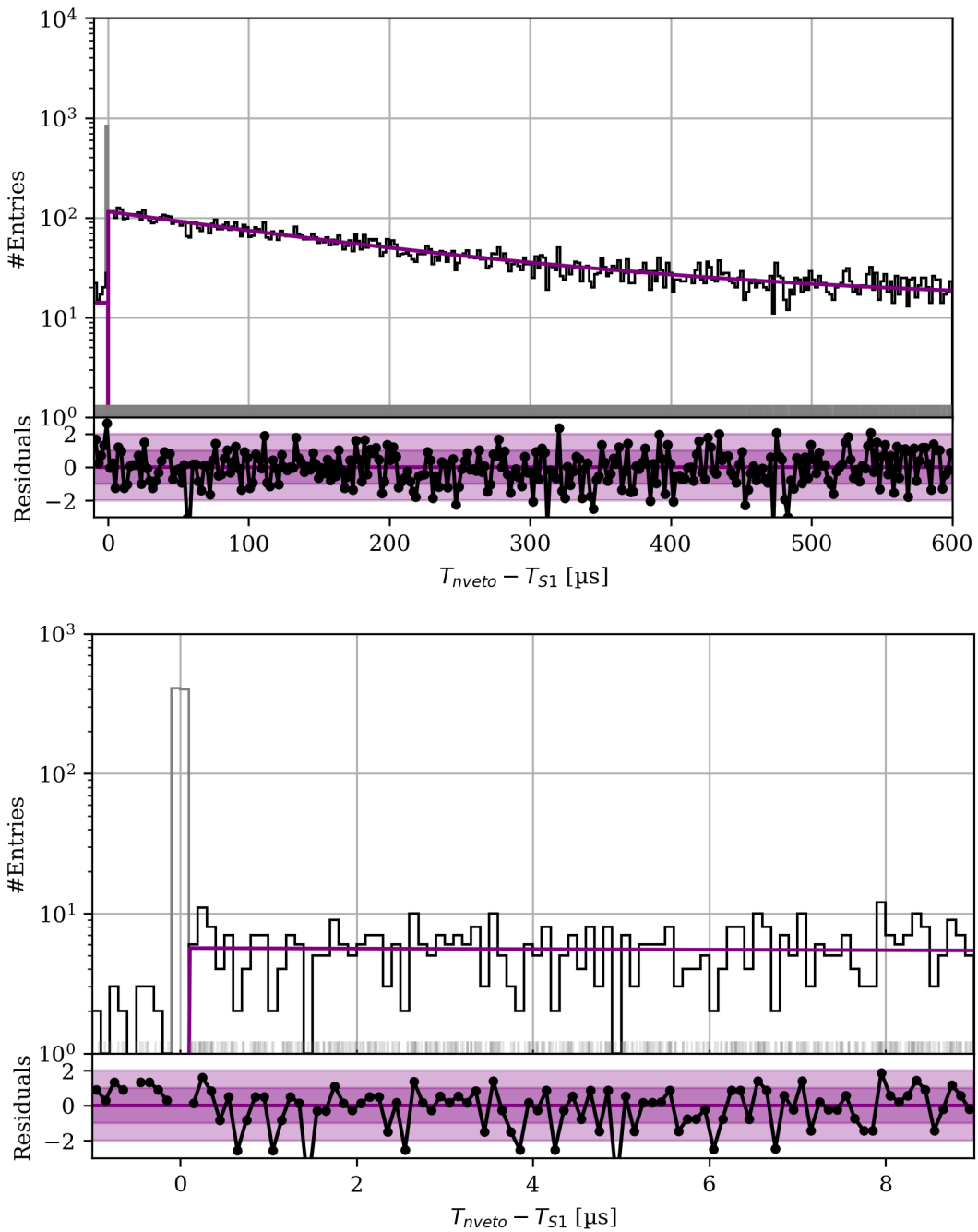


Figure B.45. – Zoom onto two different regions for the best fit of the time distribution of the neutron-veto detection efficiency study. The main panel of the figure shows a binned and unbinned representation of the data as black histogram and vertical gray lines respectively. The region excluded from the fit is shown as gray bins in the histogram. The purple line shows the best fit as described in the text. The bottom panel shows the residuals of the fit. The color shaded regions indicate the one and two sigma contour of the best fit.



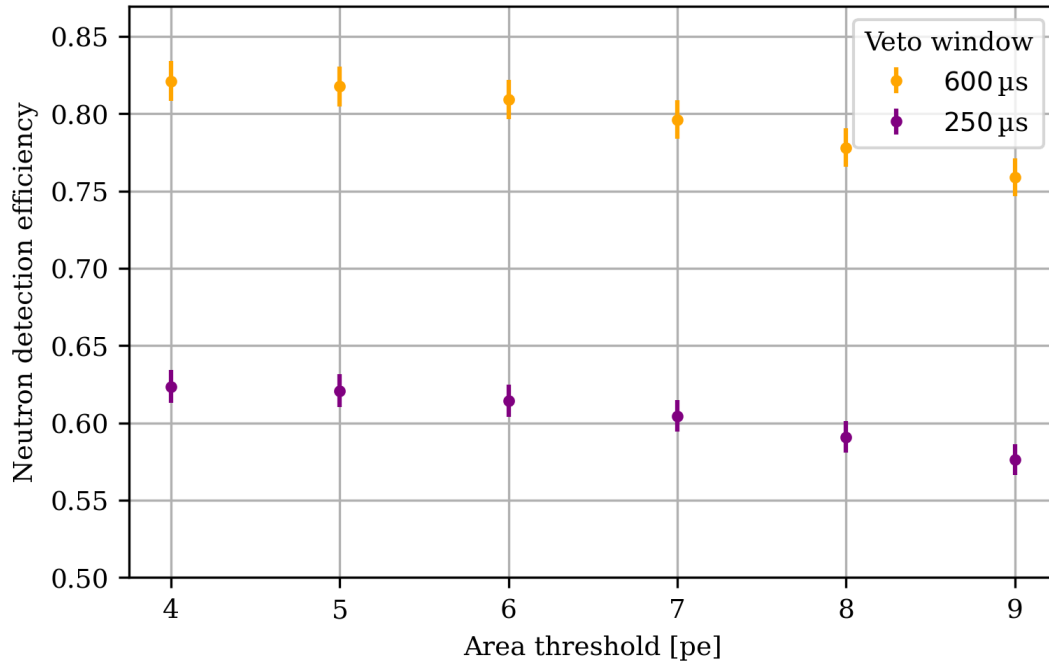


Figure B.46. – Neutron veto detection efficiency as function of event area threshold.

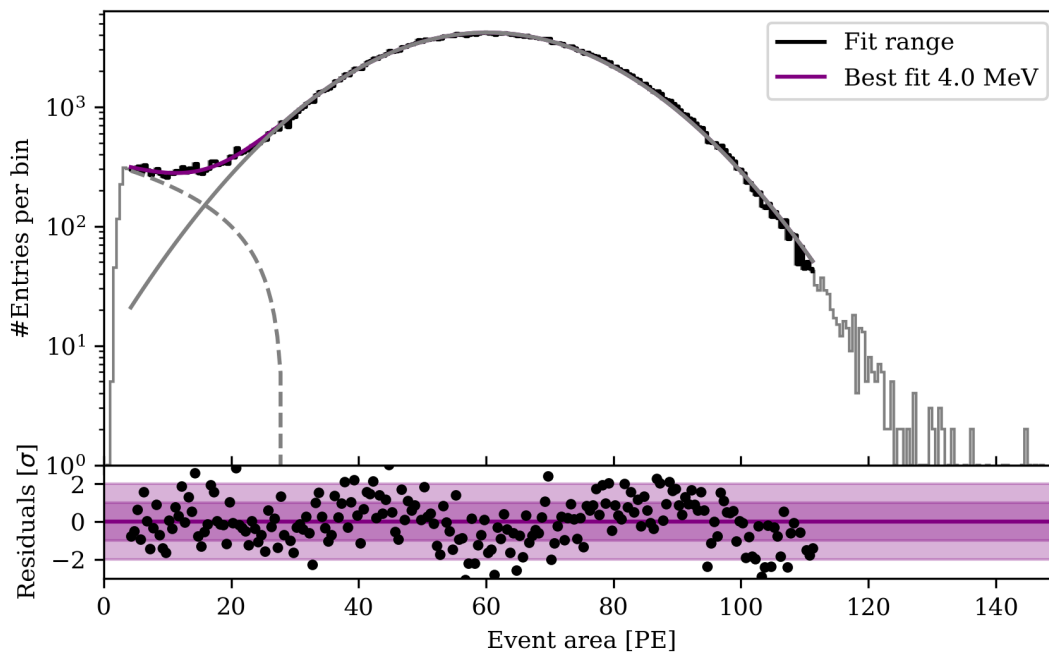
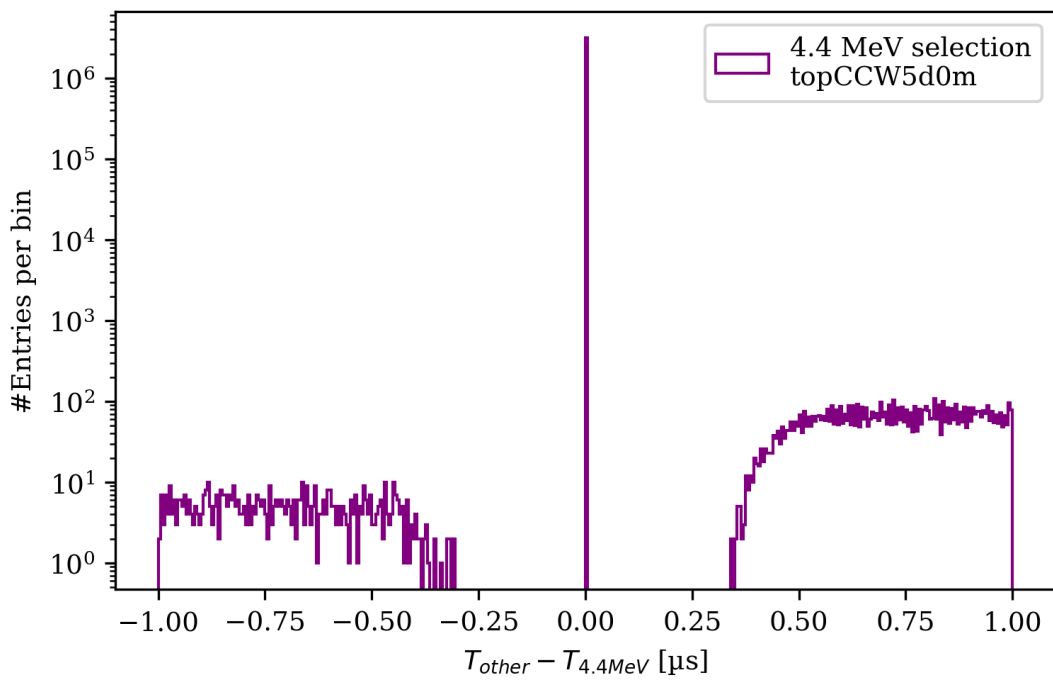


Figure B.47. – Example best fit for the simulated 4 MeV full energy line. The upper plot shows the best fit in black. Individual fit components are shown in gray. The bottom panel shows the corresponding fit residuals.



**Figure B.48.** – Zoom in the region  $\pm 1 \mu\text{s}$  of the NV self coincidence for the *topCCW5d0m* source position. The purple line shows all events found in a NV self coincidence using events with an area corresponding to the 4.44 MeV  $\gamma$ -line as trigger. The gap is caused by the minimal NV event length which is about 300 ns. The peak around zero is caused by the triggering events itself.

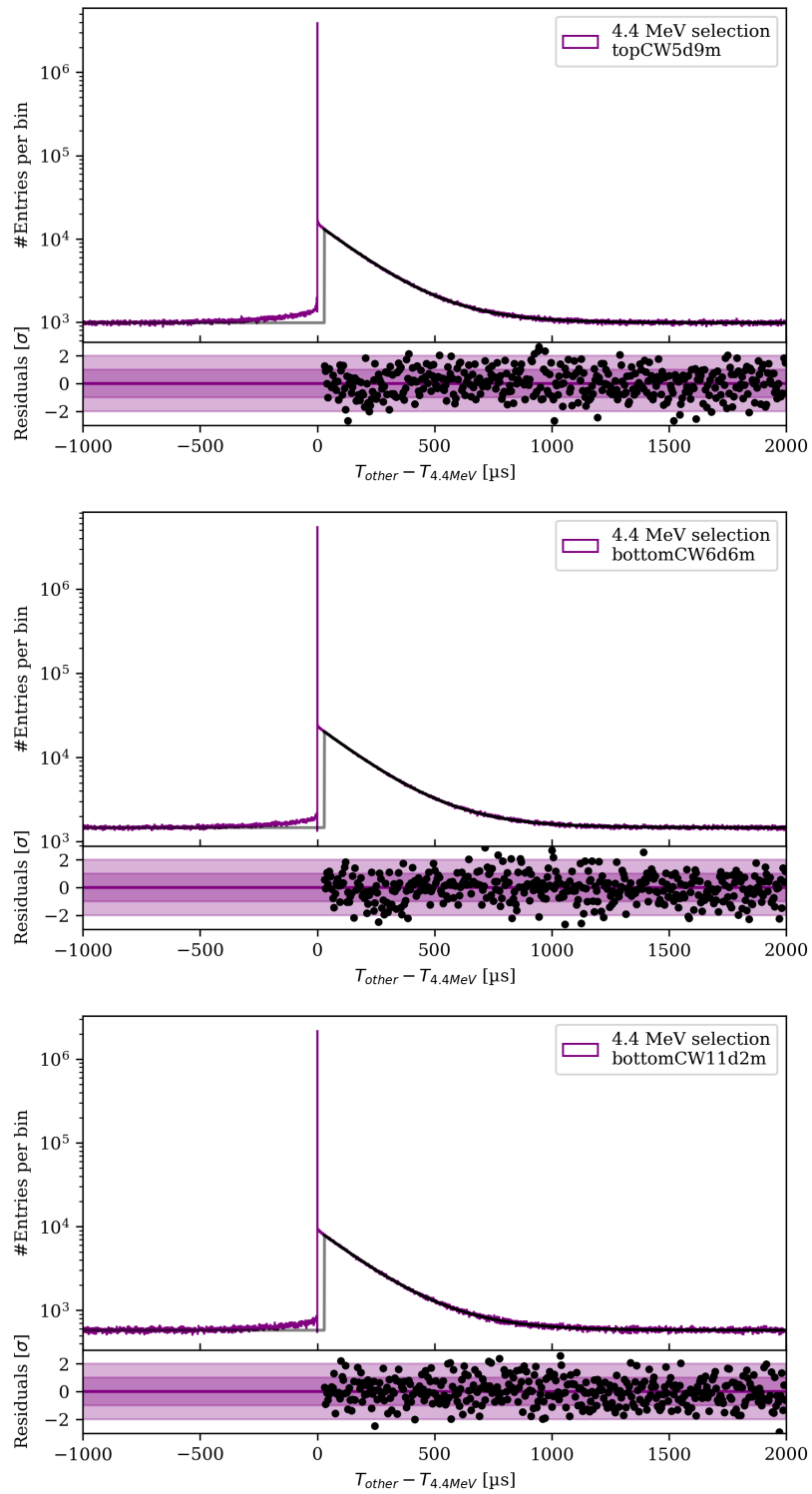


Figure B.49. – Best fits of the neutron capture time distribution recorded in the NV self-coincidence. The plots show the TopCW5d9m, BottomCW6d6m and BottomCW11d2m source positions.

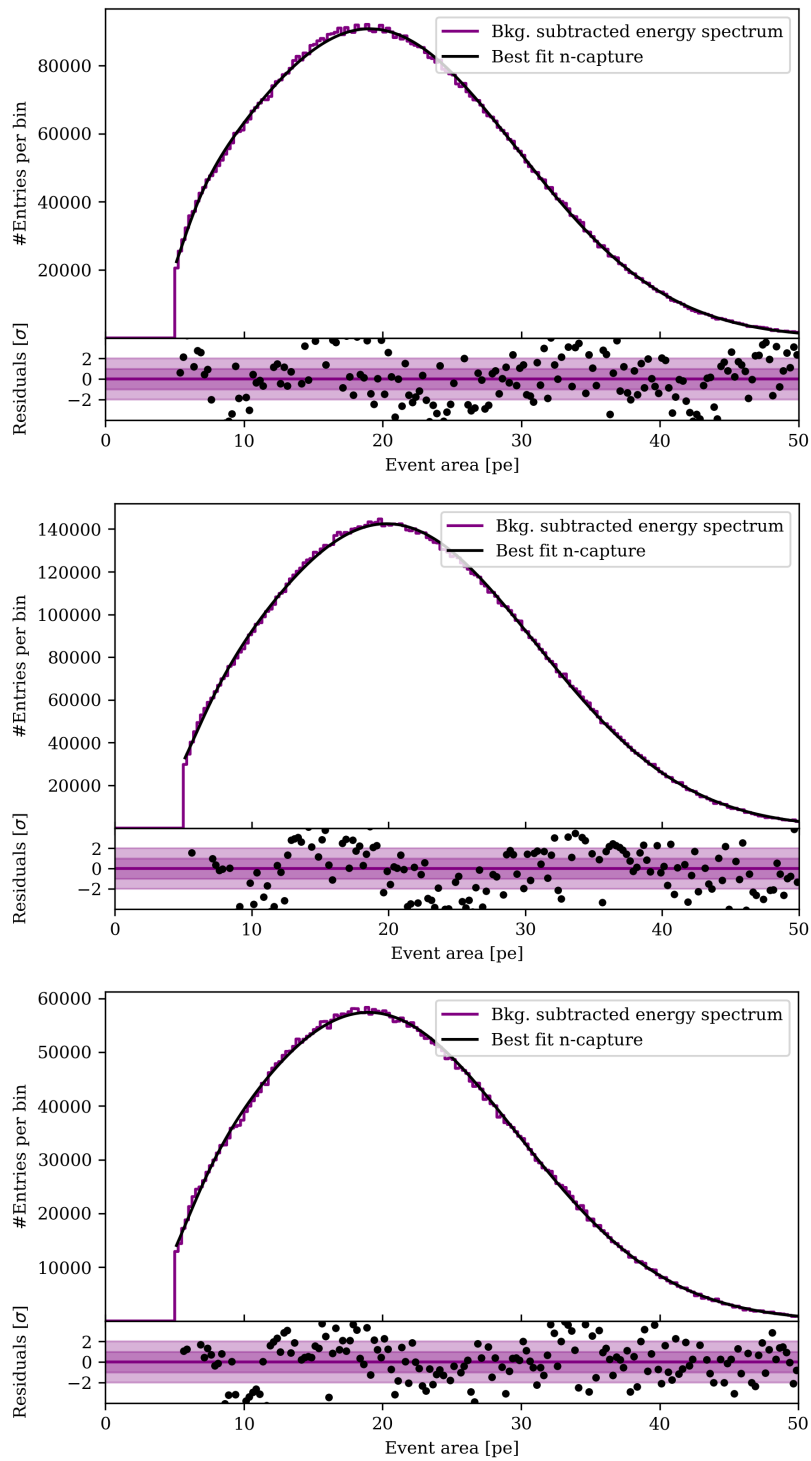


Figure B.50. – Best fits of the 2.22 MeV neutron capture gamma line recorded by the NV self-coincidence. The plots show the TopCW5d9m, BottomCW6d6m and BottomCW11d2m source positions.

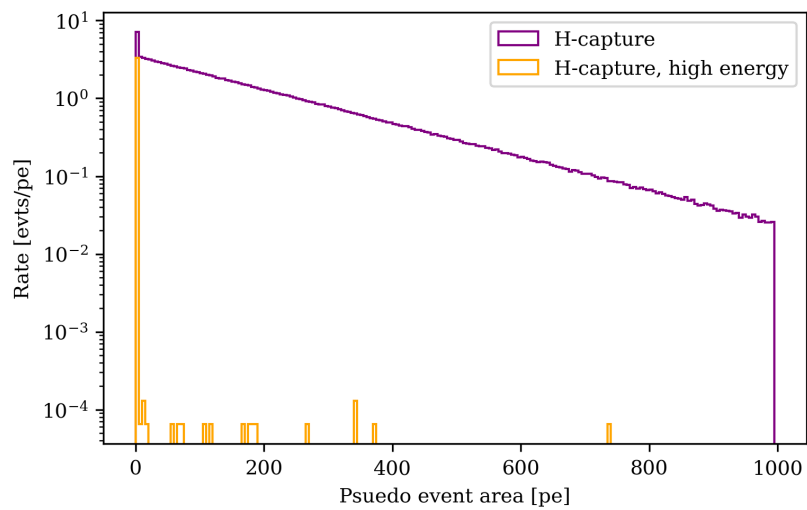


Figure B.51. – Comparison of the time distribution of the H-capture tagged events, simulated with Geant4. The purple distributions shows all events tagged as H-capture, while the orange distribution requires in addition a pseudo event area of at least 70 pe.

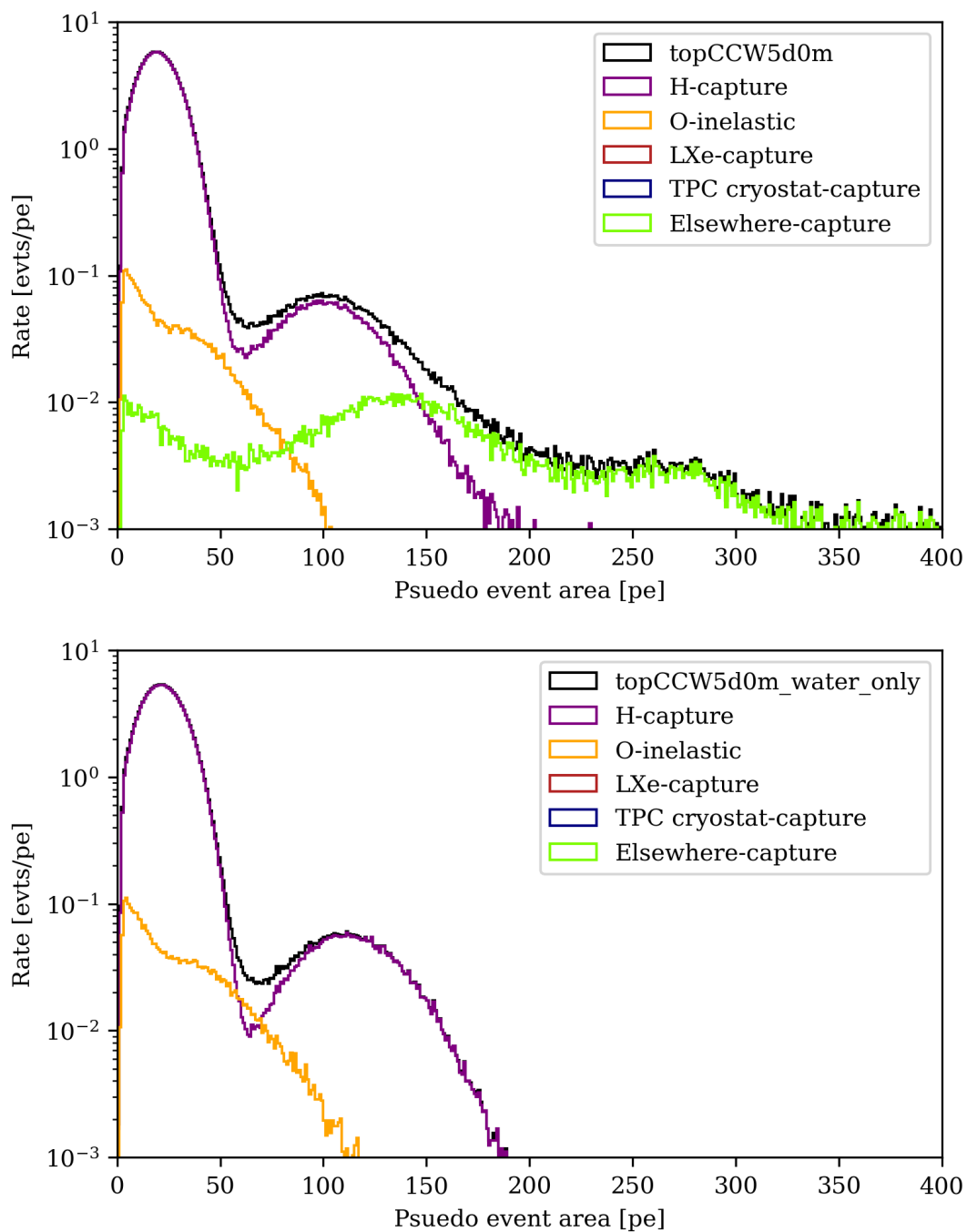


Figure B.52. – Decomposed NV energy spectrum of simulated neutrons at the topCCW5d0m source position, using custom made Geant4 saveflags. The upper plot shows the spectrum with and the lower plot without u-tubes.

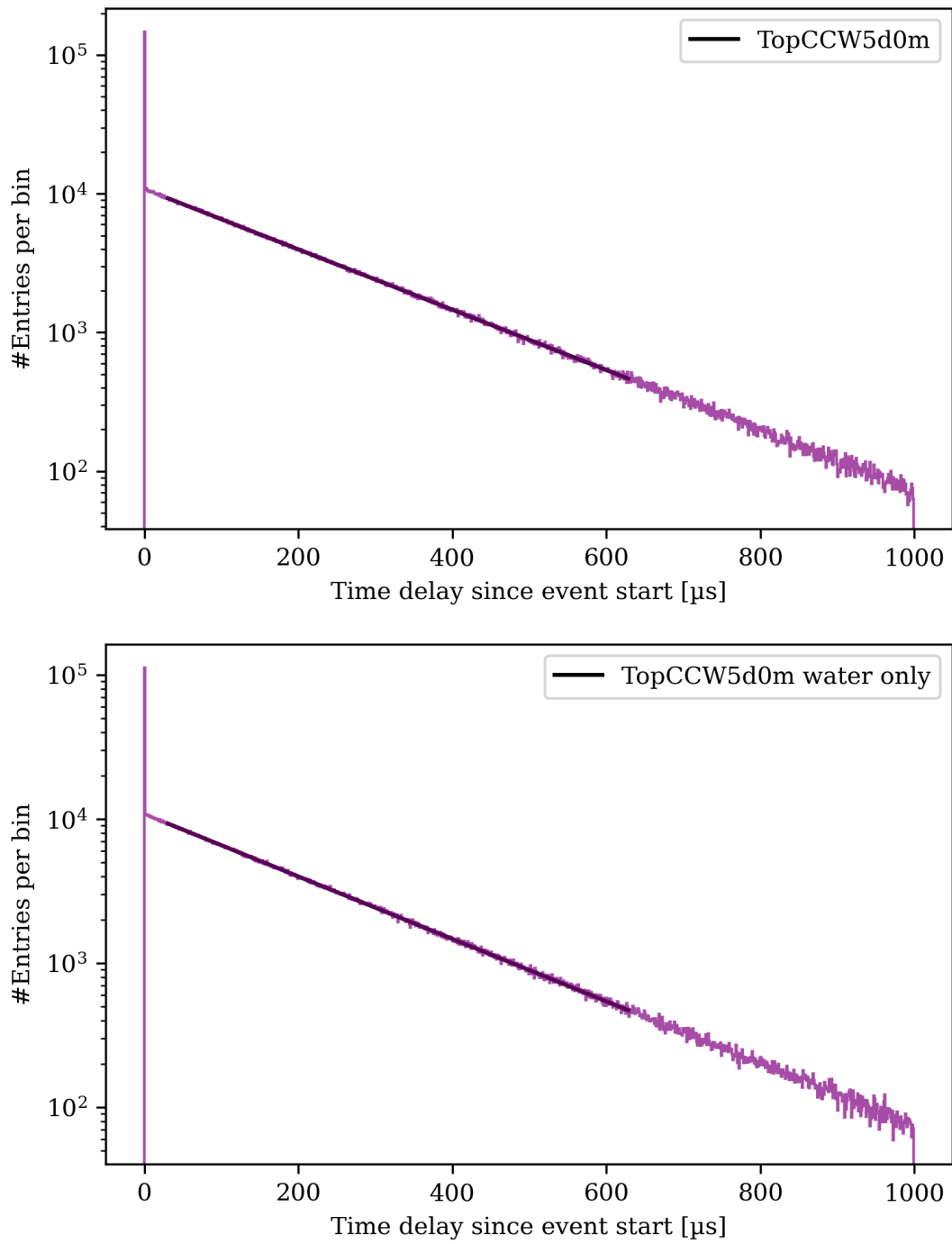


Figure B.53. – Time distribution of simulated neutrons events with (top) and without (bottom) u-tubes. The best fit of both distributions is shown in black.



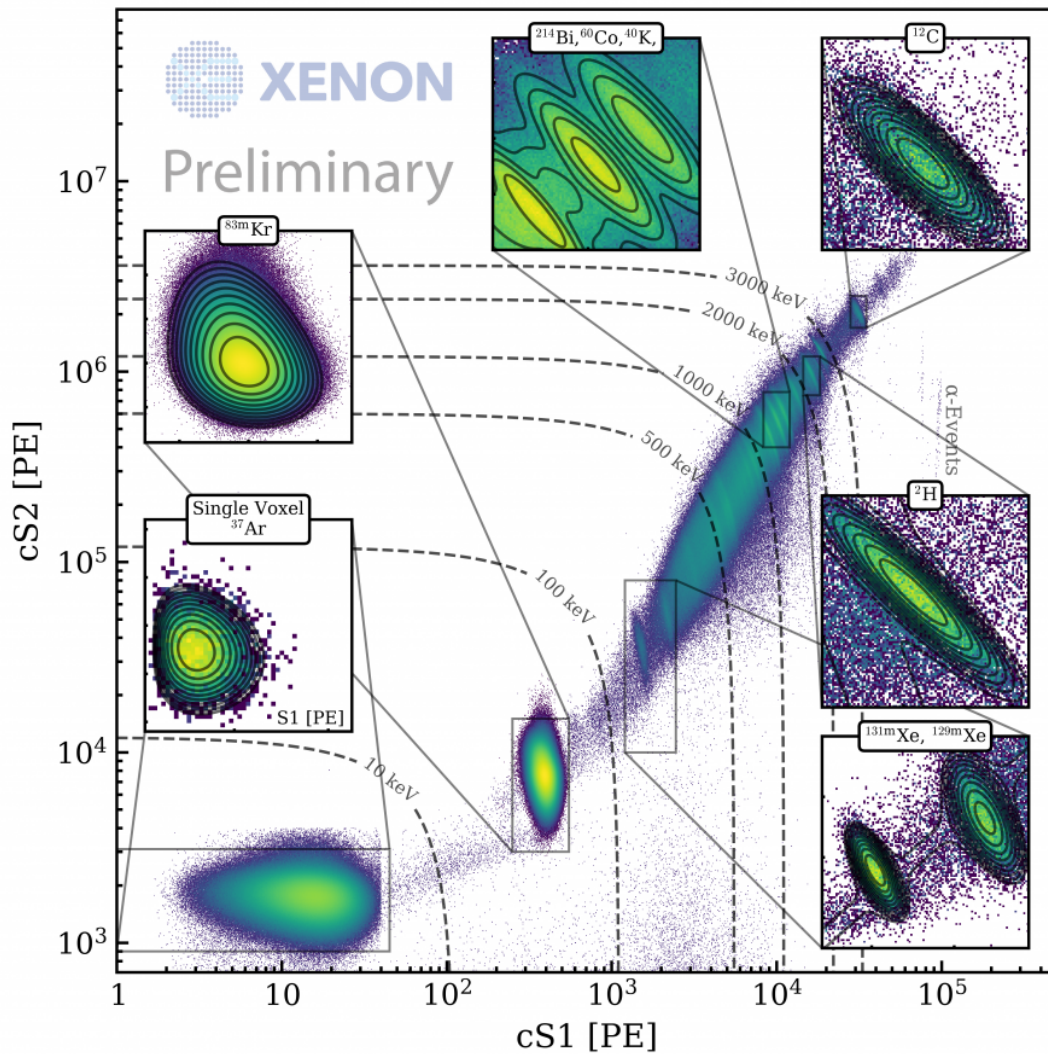


Figure B.54. – Mono-energetic calibration lines for the XENONnT energy response. The plot shows the cS2 versus cS1 data space for several mono-energetic ER lines recorded during SR0. The inset figures show a zoom onto the various different lines. (Credit XENON collaboration)

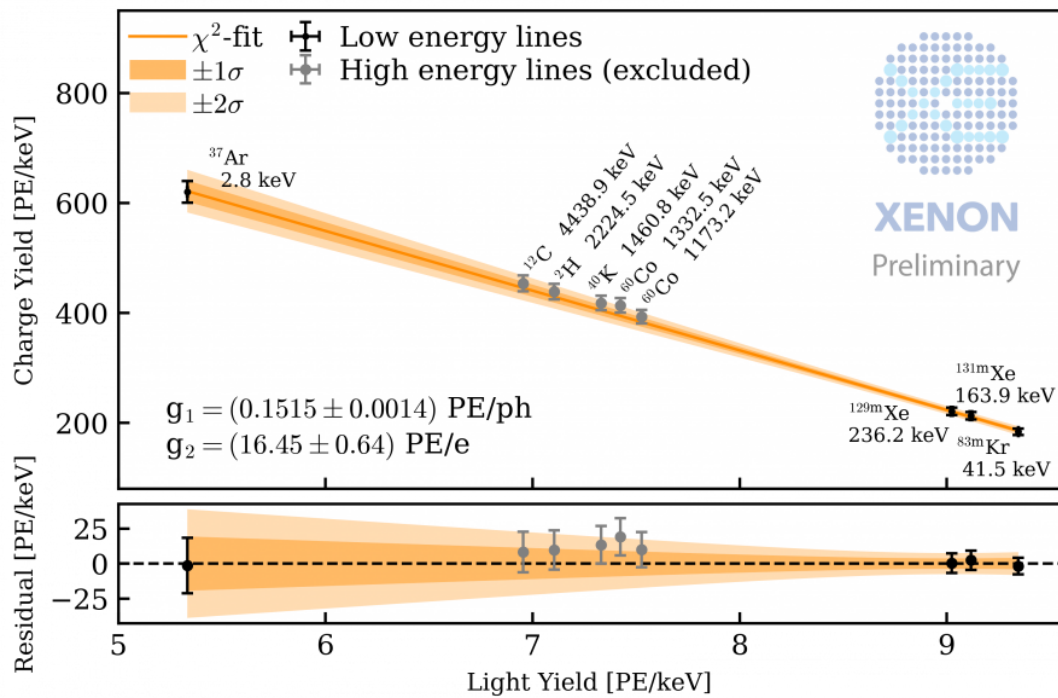


Figure B.55. – Energy response calibration of SR0. The plot shows the charge and light yield of the different mono-energetic ER lines shown in Figure B.54. Energy lines shown in gray are only plotted as reference, but do not participate in the best fit of  $g_1$  and  $g_2$ . The bottom panel shows the residuals of the fit. (Credit XENON collaboration)

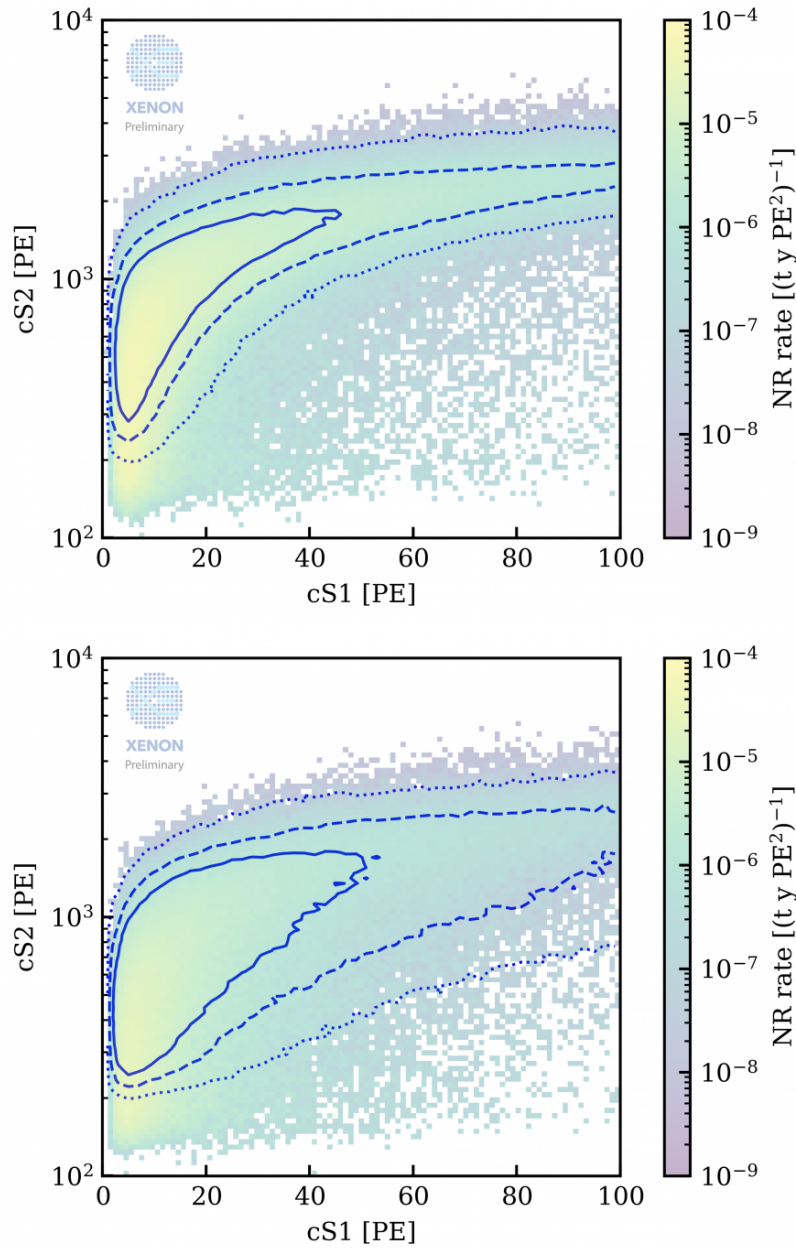


Figure B.56. – Neutron background templates used in the SR0 WIMP inference. The upper figure shows the template for regular neutron background events, while the lower plot shows so-called “neutron-x” events. A neutron-x event is caused by events for which the S2 of a regular NR recoil is paired randomly with an S1 signal. This might happen for events in which the neutron interacts with the LXe below the cathode mesh or if an S1 signal is clustered into an S2 signal, e.g. for very shallow events close to the gate. The contours are showing the 1-3 sigma contours. (Credit XENON collaboration)

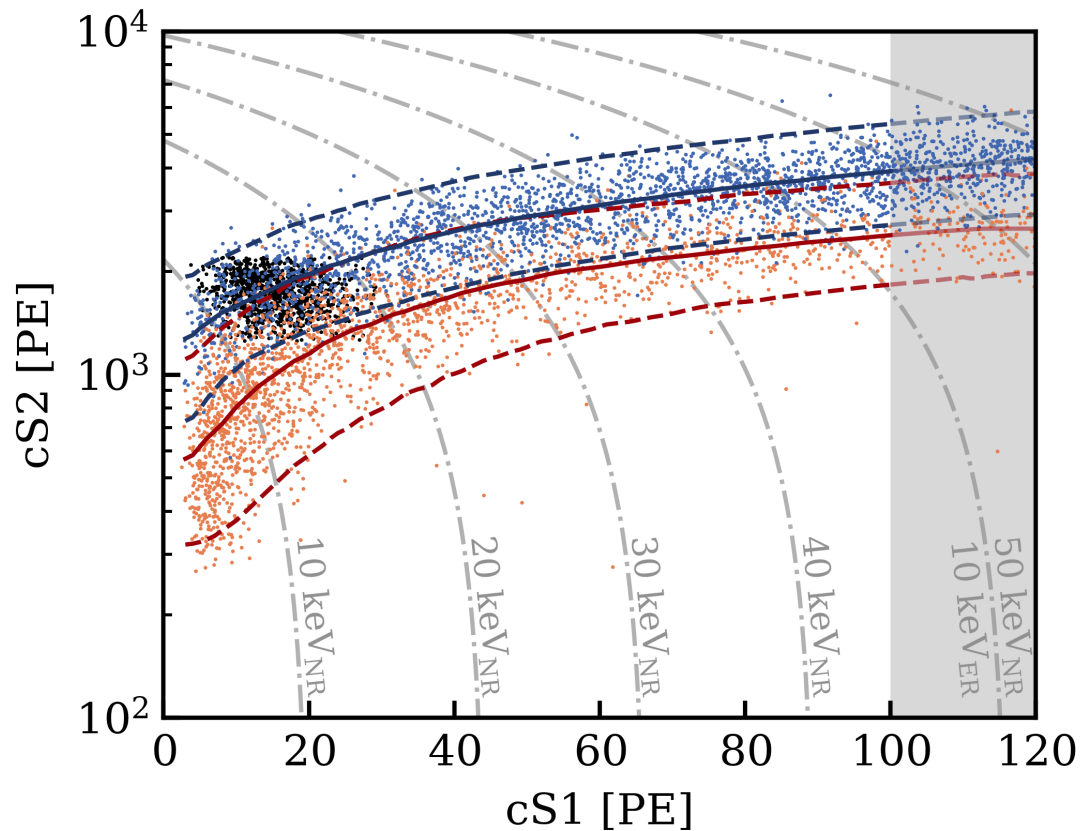


Figure B.57. – Comparison of the ER and NR band model. The orange data points and red lines show the AmBe calibration data as discussed in Chapter 6. The blue and black points show the  $^{37}\text{Ar}$  and  $^{220}\text{Rn}$  ER calibration data respectively. The ER is shown as blue lines. The dashed lines represent the  $\pm 2\sigma$  contours of the respective models. The solid line indicates the median.

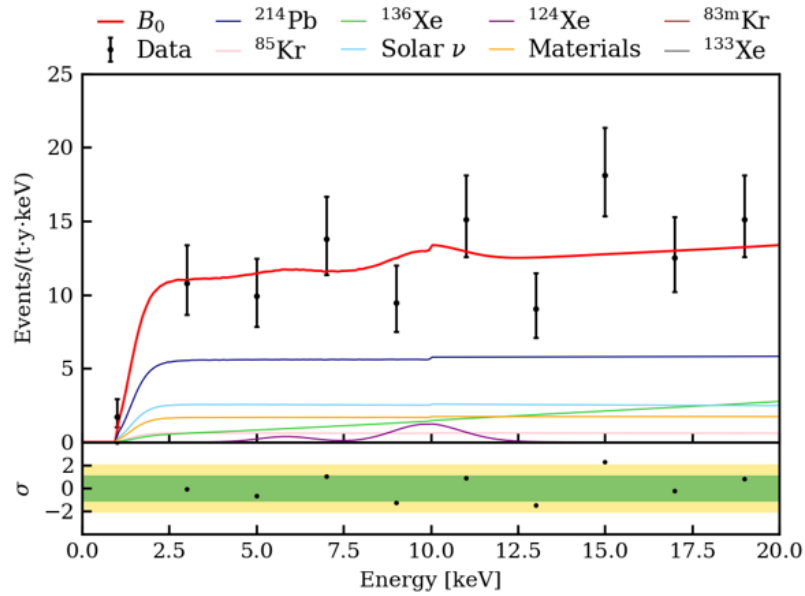


Figure B.58. – ER background model for the SR0 WIMP ROI. Shown are the individual background components derived from the best fit of the low ER search background only model  $B_0$  in [Apr+22c]. The little kink in the red line is due to the still blinded WIMP search data. (Credit XENON collaboration)

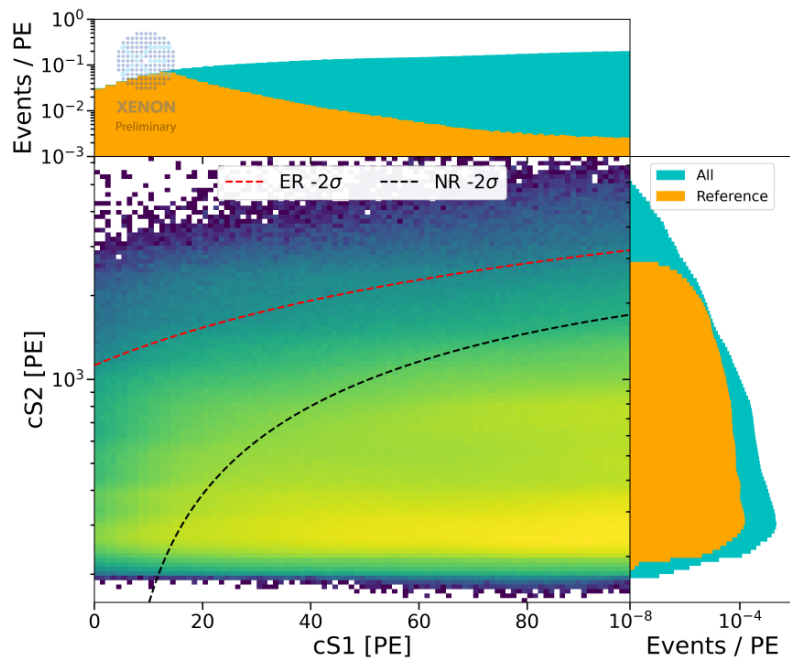


Figure B.59. – SR0 surface background model template in cS1 and cS2. The projection shows the surface background spectrum in cS1 and cS2 for the full ROI (all) and the blinded NR region (reference) (Credit XENON collaboration).

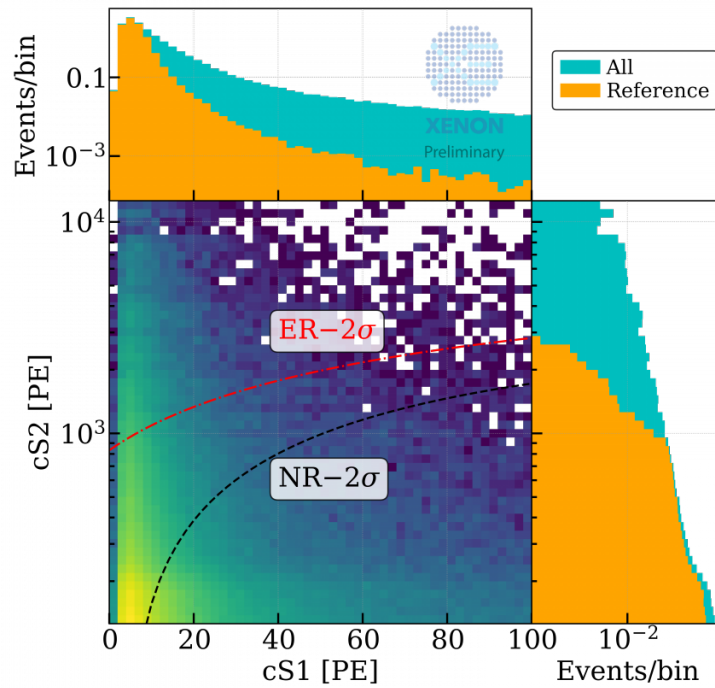


Figure B.60. – AC background model for the SR0 WIMP search. The upper figure shows the AC model in the  $cS2$  versus  $cS1$  space. The blinded region is between the red and black line. The projections show the AC model for the entire WIMP ROI (all) as well as for the blinded region (reference).

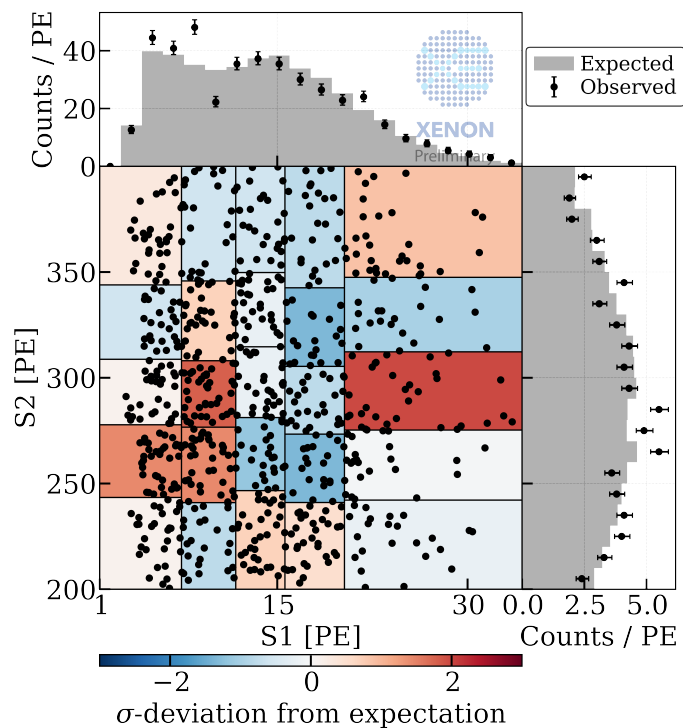


Figure B.61. – AC background model validation using  $^{37}\text{Ar}$  calibration data. The black data points show the events found in data, the shaded regions the deviation of the data from the model for the given bin. The projections show a comparison between data (black marker) and the model (gray) for the corresponding axis (Credit XENON collaboration).





---

## BIBLIOGRAPHY

- [22] *Xenon Market Size, Share & COVID-19 Impact Analysis, By Type (N3, N4.5, and N5), By Application (Imaging & Lighting, Satellite, Electronics & Semiconductors, Medical, and Others), and Regional Forecast, 2022-2029*. <https://www.fortunebusinessinsights.com/xenon-market-101965>. [Online; accessed 06-April-2023]. 2022.
- [Aal+16] J. Aalbers et al. "DARWIN: towards the ultimate dark matter detector". *Journal of Cosmology and Astroparticle Physics* 2016.11 (Nov. 2016), p. 017. DOI: [10.1088/1475-7516/2016/11/017](https://doi.org/10.1088/1475-7516/2016/11/017). URL: <https://dx.doi.org/10.1088/1475-7516/2016/11/017>.
- [Aal+18] C.E. Aalseth et al. "DarkSide-20k: A 20 tonne two-phase LAr TPC for direct dark matter detection at LNGS". *Eur. Phys. J. Plus* 133 (Feb. 2018). DOI: [10.1140/epjp/i2018-11973-4](https://doi.org/10.1140/epjp/i2018-11973-4).
- [Aal+22] J. Aalbers et al. "First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment" (2022). arXiv: [2207.03764](https://arxiv.org/abs/2207.03764) [astro-ex].
- [Aal+23] Jelle Aalbers et al. *JelleAalbers/wimprates: v0.5.0*. Version v0.5.0. Feb. 2023. DOI: [10.5281/zenodo.7636982](https://doi.org/10.5281/zenodo.7636982). URL: <https://doi.org/10.5281/zenodo.7636982>.
- [Abb+23] R. Abbasi et al. *Observation of Seasonal Variations of the Flux of High-Energy Atmospheric Neutrinos with IceCube*. 2023. arXiv: [2303.04682](https://arxiv.org/abs/2303.04682) [astro-ph.HE].
- [Abd+19] A. H. Abdelhameed et al. "First results from the CRESST-III low-mass dark matter program". *Phys. Rev. D* 100 (10 Nov. 2019), p. 102002. DOI: [10.1103/PhysRevD.100.102002](https://doi.org/10.1103/PhysRevD.100.102002).
- [Abd+22] H. Abdalla et al. "Search for Dark Matter Annihilation Signals in the H.E.S.S. Inner Galaxy Survey". *Phys. Rev. Lett.* 129 (11 Sept. 2022), p. 111101. DOI: [10.1103/PhysRevLett.129.111101](https://doi.org/10.1103/PhysRevLett.129.111101). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.129.111101>.
- [Abe+20] K. Abe et al. "Indirect search for dark matter from the Galactic Center and halo with the Super-Kamiokande detector". *Phys. Rev. D* 102 (7 Oct. 2020), p. 072002. DOI: [10.1103/PhysRevD.102.072002](https://doi.org/10.1103/PhysRevD.102.072002). URL: <https://link.aps.org/doi/10.1103/PhysRevD.102.072002>.

- [Abe+22] K. Abe et al. “First gadolinium loading to Super-Kamiokande”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1027 (2022), p. 166248. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2021.166248>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900221010883>.
- [AD10] E. Aprile and T. Doke. “Liquid xenon detectors for particle physics and astrophysics”. *Rev. Mod. Phys.* 82 (3 July 2010), pp. 2053–2097. DOI: [10.1103/RevModPhys.82.2053](https://link.aps.org/doi/10.1103/RevModPhys.82.2053). URL: <https://link.aps.org/doi/10.1103/RevModPhys.82.2053>.
- [Agn+23] P. Agnes et al. “Search for Dark-Matter–Nucleon Interactions via Migdal Effect with DarkSide-50”. *Phys. Rev. Lett.* 130 (10 Mar. 2023), p. 101001. DOI: [10.1103/PhysRevLett.130.101001](https://doi.org/10.1103/PhysRevLett.130.101001).
- [Ago+03] S. Agostinelli et al. “GEANT4—a simulation toolkit”. *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [Aha+13] B. Aharmim et al. “Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory”. *Phys. Rev. C* 88 (2 Aug. 2013), p. 025501. DOI: [10.1103/PhysRevC.88.025501](https://doi.org/10.1103/PhysRevC.88.025501). URL: <https://link.aps.org/doi/10.1103/PhysRevC.88.025501>.
- [Ake+17a] D. S. Akerib et al. “ $^{83\text{m}}\text{Kr}$  calibration of the 2013 LUX dark matter search”. *Phys. Rev. D* 96 (11 Dec. 2017), p. 112009. DOI: [10.1103/PhysRevD.96.112009](https://doi.org/10.1103/PhysRevD.96.112009). URL: <https://link.aps.org/doi/10.1103/PhysRevD.96.112009>.
- [Ake+17b] D. S. Akerib et al. “Results from a Search for Dark Matter in the Complete LUX Exposure”. *Phys. Rev. Lett.* 118 (2 Jan. 2017), p. 021303. DOI: [10.1103/PhysRevLett.118.021303](https://doi.org/10.1103/PhysRevLett.118.021303). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.118.021303>.
- [Ake+20] D.S. Akerib et al. “The LUX-ZEPLIN (LZ) experiment”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 953 (2020), p. 163047. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2019.163047>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900219314032>.
- [Aki+18] D. Akimov et al. “COHERENT 2018 at the Spallation Neutron Source” (Mar. 2018). arXiv: [1803.09183](https://arxiv.org/abs/1803.09183) [physics.ins-det].
- [Alb+02] E Albrecht et al. “COMPASS RICH-1”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 478.1 (2002). Proceedings of the ninth Int.Conf. on Instrumentation, pp. 340–343. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(01\)01823-X](https://doi.org/10.1016/S0168-9002(01)01823-X). URL: <https://www.sciencedirect.com/science/article/pii/S016890020101823X>.

- 
- [Alb+14] J. B. Albert et al. "Improved measurement of the  $2\nu\beta\beta$  half-life of  $^{136}\text{Xe}$  with the EXO-200 detector". *Phys. Rev. C* 89 (1 Jan. 2014), p. 015502. DOI: [10.1103/PhysRevC.89.015502](https://doi.org/10.1103/PhysRevC.89.015502). URL: <https://link.aps.org/doi/10.1103/PhysRevC.89.015502>.
- [Alb+16] J. B. Albert et al. "Measurement of neutron capture on  $^{136}\text{Xe}$ ". *Phys. Rev. C* 94 (3 Sept. 2016), p. 034617. DOI: [10.1103/PhysRevC.94.034617](https://doi.org/10.1103/PhysRevC.94.034617).
- [Ama+22] C. S. Amarasinghe et al. "Feasibility study to use neutron capture for an ultralow energy nuclear-recoil calibration in liquid xenon". *Phys. Rev. D* 106 (3 Aug. 2022), p. 032007. DOI: [10.1103/PhysRevD.106.032007](https://doi.org/10.1103/PhysRevD.106.032007). URL: <https://link.aps.org/doi/10.1103/PhysRevD.106.032007>.
- [And+20] M. R. Anderson et al. "Measurement of neutron-proton capture in the SNO+ water phase". *Phys. Rev. C* 102 (1 July 2020), p. 014002. DOI: [10.1103/PhysRevC.102.014002](https://doi.org/10.1103/PhysRevC.102.014002). URL: <https://link.aps.org/doi/10.1103/PhysRevC.102.014002>.
- [Ang23] Joran Remco Angevaere. "First WIMP results of XENONnT and its signal reconstruction". *to be published* (2023).
- [Ant+20] G. Anton et al. "Measurement of the scintillation and ionization response of liquid xenon at MeV energies in the EXO-200 experiment". *Phys. Rev. C* 101.6 (2020), p. 065501. DOI: [10.1103/PhysRevC.101.065501](https://doi.org/10.1103/PhysRevC.101.065501). arXiv: [1908.04128](https://arxiv.org/abs/1908.04128) [physics.ins-det].
- [Ant+21] V.C. Antochi et al. "Improved quality tests of R11410-21 photomultiplier tubes for the XENONnT experiment". *Journal of Instrumentation* 16.08 (Aug. 2021), P08033. DOI: [10.1088/1748-0221/16/08/P08033](https://doi.org/10.1088/1748-0221/16/08/P08033). URL: <https://dx.doi.org/10.1088/1748-0221/16/08/P08033>.
- [Apr+12] E. Aprile et al. "The XENON100 dark matter experiment". *Astroparticle Physics* 35.9 (2012), pp. 573–590. ISSN: 0927-6505. DOI: <https://doi.org/10.1016/j.astropartphys.2012.01.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0927650512000059>.
- [Apr+13] E. Aprile et al. "Response of the XENON100 Dark Matter Detector to Nuclear Recoils". *Phys. Rev. D* 88 (2013).
- [Apr+14] E. Aprile et al. "Conceptual design and simulation of a water Cherenkov muon veto for the XENON1T experiment". *Journal of Instrumentation* 9.11 (Nov. 2014), P11006. DOI: [10.1088/1748-0221/9/11/P11006](https://doi.org/10.1088/1748-0221/9/11/P11006). URL: <https://dx.doi.org/10.1088/1748-0221/9/11/P11006>.
- [Apr+17a] E. Aprile et al. "Removing krypton from xenon by cryogenic distillation to the ppq level". *Eur. Phys. J. C* 77.5 (May 2017), p. 275. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-017-4757-1](https://doi.org/10.1140/epjc/s10052-017-4757-1). URL: <https://doi.org/10.1140/epjc/s10052-017-4757-1>.
- [Apr+17b] E. Aprile et al. "First Dark Matter Search Results from the XENON1T Experiment". *Phys. Rev. Lett.* 119 (18 Oct. 2017), p. 181301. DOI: [10.1103/PhysRevLett.119.181301](https://doi.org/10.1103/PhysRevLett.119.181301). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.119.181301>.

- [Apr+17c] Aprile, E. et al. “The XENON1T dark matter experiment”. *Eur. Phys. J. C* 77.12 (2017), p. 881. DOI: [10.1140/epjc/s10052-017-5326-3](https://doi.org/10.1140/epjc/s10052-017-5326-3). URL: <https://doi.org/10.1140/epjc/s10052-017-5326-3>.
- [Apr+18] E. Aprile et al. “Dark Matter Search Results from a One Ton-Year Exposure of XENON1T”. *Phys. Rev. Lett.* 121.11 (2018), p. 111302. DOI: [10.1103/PhysRevLett.121.111302](https://doi.org/10.1103/PhysRevLett.121.111302). arXiv: [1805.12562](https://arxiv.org/abs/1805.12562) [astro-ph.CO].
- [Apr+19a] E. Aprile et al. “Light Dark Matter Search with Ionization Signals in XENON1T”. *Phys. Rev. Lett.* 123 (25 Dec. 2019), p. 251801. DOI: [10.1103/PhysRevLett.123.251801](https://doi.org/10.1103/PhysRevLett.123.251801). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.123.251801>.
- [Apr+19b] E. Aprile et al. “XENON1T Dark Matter Data Analysis: Signal and Background Models and Statistical Inference”. *Phys. Rev. D* 99 (11 June 2019), p. 112009. DOI: [10.1103/PhysRevD.99.112009](https://doi.org/10.1103/PhysRevD.99.112009). URL: <https://link.aps.org/doi/10.1103/PhysRevD.99.112009>.
- [Apr+19c] E. Aprile et al. “XENON1T Dark Matter Data Analysis: Signal Reconstruction, Calibration and Event Selection”. *Phys. Rev. D* 100.5 (2019), p. 052014. DOI: [10.1103/PhysRevD.100.052014](https://doi.org/10.1103/PhysRevD.100.052014). arXiv: [1906.04717](https://arxiv.org/abs/1906.04717) [physics.ins-det].
- [Apr+20a] E. Aprile et al. “Excess electronic recoil events in XENON1T”. *Phys. Rev. D* 102.7 (2020), p. 072004. DOI: [10.1103/PhysRevD.102.072004](https://doi.org/10.1103/PhysRevD.102.072004). arXiv: [2006.09721](https://arxiv.org/abs/2006.09721) [hep-ex].
- [Apr+20b] E. Aprile et al. “Projected WIMP sensitivity of the XENONnT dark matter experiment”. *JCAP* 2020.11 (2020), pp. 031–031. DOI: [10.1088/1475-7516/2020/11/031](https://doi.org/10.1088/1475-7516/2020/11/031). arXiv: [2007.08796](https://arxiv.org/abs/2007.08796) [physics.ins-det].
- [Apr+22a] E. Aprile et al. “Low-energy Calibration of XENON1T with an Internal  $^{37}\text{Ar}$  Source”. *submitted to EPJ-C* (Nov. 2022). arXiv: [2211.14191](https://arxiv.org/abs/2211.14191) [physics.ins-det].
- [Apr+22b] E. Aprile et al. “Material radiopurity control in the XENONnT experiment”. *Eur. Phys. J. C* 82 (2022), p. 599. DOI: <https://doi.org/10.1140/epjc/s10052-022-10345-6>.
- [Apr+22c] E. Aprile et al. “Search for New Physics in Electronic Recoil Data from XENONnT”. *Phys. Rev. Lett.* 129 (16 Oct. 2022), p. 161805. DOI: [10.1103/PhysRevLett.129.161805](https://doi.org/10.1103/PhysRevLett.129.161805). arXiv: [2207.11330](https://arxiv.org/abs/2207.11330). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.129.161805>.
- [Apr+22d] E. Aprile et al. “The Triggerless Data Acquisition System of the XENONnT Experiment”. *submitted to JINST* (Dec. 2022). arXiv: [2212.11032](https://arxiv.org/abs/2212.11032) [physics.ins-det].
- [Apr+22e] E. Aprile et al. “Double-weak decays of  $^{124}\text{Xe}$  and  $^{136}\text{Xe}$  in the XENON1T and XENONnT experiments”. *Phys. Rev. C* 106 (2 Aug. 2022), p. 024328. DOI: [10.1103/PhysRevC.106.024328](https://doi.org/10.1103/PhysRevC.106.024328). URL: <https://link.aps.org/doi/10.1103/PhysRevC.106.024328>.
- [Apr+23a] E. Aprile et al. *Detector signal characterization with a Bayesian network in XENONnT*. 2023. arXiv: [2304.05428](https://arxiv.org/abs/2304.05428) [hep-ex].

- 
- [Apr+23b] E. Aprile et al. *First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment*. 2023. arXiv: [2303.14729](https://arxiv.org/abs/2303.14729) [hep-ex].
- [AX19] Aprile, E. and XENON Collaboration. “Observation of two-neutrino double electron capture in  $^{124}\text{Xe}$  with XENON1T”. *Nature* 568 (2019), pp. 532–535. DOI: [10.1038/s41586-019-1124-4](https://doi.org/10.1038/s41586-019-1124-4).
- [Bar+19] N. Bartolo et al. “Testing primordial black holes as dark matter with LISA”. *Phys. Rev. D* 99 (10 May 2019), p. 103521. DOI: [10.1103/PhysRevD.99.103521](https://doi.org/10.1103/PhysRevD.99.103521).
- [Bax+21a] D. Baxter et al. “Recommended conventions for reporting results from direct dark matter searches”. *Eur. Phys. J. C* 81.10 (2021), p. 907. DOI: [10.1140/epjc/s10052-021-09655-y](https://doi.org/10.1140/epjc/s10052-021-09655-y). arXiv: [2105.00599](https://arxiv.org/abs/2105.00599) [hep-ex].
- [Bax+21b] D. Baxter et al. “Recommended conventions for reporting results from direct dark matter searches”. *Eur. Phys. J. C* 81.10 (2021), p. 907. DOI: [10.1140/epjc/s10052-021-09655-y](https://doi.org/10.1140/epjc/s10052-021-09655-y). arXiv: [2105.00599](https://arxiv.org/abs/2105.00599) [hep-ex].
- [BC84] Steve Baker and Robert D. Cousins. “Clarification of the Use of Chi Square and Likelihood Functions in Fits to Histograms”. *Nucl. Instrum. Meth.* 221 (1984), pp. 437–442. DOI: [10.1016/0167-5087\(84\)90016-4](https://doi.org/10.1016/0167-5087(84)90016-4).
- [Bel+22] Nicole F. Bell et al. “Observing the Migdal effect from nuclear recoils of neutral particles with liquid xenon and argon detectors”. *Phys. Rev. D* 105 (9 May 2022), p. 096015. DOI: [10.1103/PhysRevD.105.096015](https://doi.org/10.1103/PhysRevD.105.096015). URL: <https://link.aps.org/doi/10.1103/PhysRevD.105.096015>.
- [Ber+10] M.J. Berger et al. “XCOM: Photon Cross Sections Database”. *NIST Standard Reference Database 8 (XGAM)* (2010). DOI: [10.18434/T48G6X](https://doi.org/10.18434/T48G6X).
- [Ber+16] P. Bernhard et al. “Construction of large-area micro-pattern gaseous detectors”. *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD)*. 2016, pp. 1–4. DOI: [10.1109/NSSMIC.2016.8069756](https://doi.org/10.1109/NSSMIC.2016.8069756).
- [Ber+17] M.J. Berger et al. “Stopping-Power & Range Tables for Electrons, Protons, and Helium Ions”. *NIST Standard Reference Database 124* (2017). DOI: [10.18434/T4NC7P](https://doi.org/10.18434/T4NC7P).
- [BHS05] Gianfranco Bertone, Dan Hooper, and Joseph Silk. “Particle dark matter: evidence, candidates and constraints”. *Physics Reports* 405.5 (2005), pp. 279–390. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2004.08.031>.
- [Bro22] National Nuclear Data Center (NNDC) at Brookhaven National Laboratory. *NuDat 3.0*. <https://www.nndc.bnl.gov/nudat3/>. [Online; accessed 17.11.2022]. 2022.
- [BST21] Laura Baudis, Patricia Sanchez-Lucas, and Kevin Thieme. “A measurement of the mean electronic excitation energy of liquid xenon”. *Eur. Phys. J. C* 81.12 (2021), p. 1060. DOI: [10.1140/epjc/s10052-021-09834-x](https://doi.org/10.1140/epjc/s10052-021-09834-x). arXiv: [2109.07151](https://arxiv.org/abs/2109.07151) [physics.ins-det].

- [CA13] V Chepel and H Araújo. “Liquid noble gas detectors for low energy particle physics”. *Journal of Instrumentation* 8.04 (Apr. 2013), R04001. DOI: [10.1088/1748-0221/8/04/R04001](https://doi.org/10.1088/1748-0221/8/04/R04001). URL: <https://dx.doi.org/10.1088/1748-0221/8/04/R04001>.
- [Car+90] G. Carugno et al. “Electron lifetime detector for liquid argon”. *Nucl. Instrum. Meth. A* 292 (1990), pp. 580–584. DOI: [10.1016/0168-9002\(90\)90176-7](https://doi.org/10.1016/0168-9002(90)90176-7).
- [Cha19] Petr Chaguine. private communication. Aug. 19, 2019.
- [CK22] Bernard Carr and Florian Kuhnel. “Primordial black holes as dark matter candidates”. *SciPost Phys. Lect. Notes* (2022), p. 48. DOI: [10.21468/SciPostPhysLectNotes.48](https://doi.org/10.21468/SciPostPhysLectNotes.48).
- [CM77] D Cokinos and E Melkonian. “Measurement of the 2200 m/sec neutron-proton capture cross section”. *Phys. Rev., C; (United States)* (May 1977). DOI: [10.1103/PhysRevC.15.1636](https://doi.org/10.1103/PhysRevC.15.1636). URL: <https://www.osti.gov/biblio/7311705>.
- [col18] Planck collaboration. *The 2018 Planck map of the temperature anisotropies of the CMB*. June 2018. URL: <https://www.cosmos.esa.int/web/planck/picture-gallery>.
- [COM] COMSOL AB, Stockholm, Sweden. *COMSOL Multiphysics*. Version v. 5.4. URL: [www.comsol.com](http://www.comsol.com).
- [Coo22] Jodi Cooley. “Dark Matter direct detection of classical WIMPs”. *SciPost Phys. Lect. Notes* (2022), p. 55. DOI: [10.21468/SciPostPhysLectNotes.55](https://doi.org/10.21468/SciPostPhysLectNotes.55). URL: <https://scipost.org/10.21468/SciPostPhysLectNotes.55>.
- [Cou+15] J. S. Coursey et al. “Atomic Weights and Isotopic Compositions with Relative Atomic Masses”. *NIST Physical Measurement Laboratory* (2015).
- [Cow+11] Glen Cowan et al. “Power-Constrained Limits” (May 2011). arXiv: [1105.3166](https://arxiv.org/abs/1105.3166) [physics.data-an].
- [CP34] C. J. CLOPPER and E. S. PEARSON. “THE USE OF CONFIDENCE OR FIDUCIAL LIMITS ILLUSTRATED IN THE CASE OF THE BINOMIAL”. *Biometrika* 26.4 (Dec. 1934), pp. 404–413. ISSN: 0006-3444. DOI: [10.1093/biomet/26.4.404](https://doi.org/10.1093/biomet/26.4.404). eprint: <https://academic.oup.com/biomet/article-pdf/26/4/404/823407/26-4-404.pdf>. URL: <https://doi.org/10.1093/biomet/26.4.404>.
- [CS00] Edvige Corbelli and Paolo Salucci. “The extended rotation curve and the dark matter halo of M33”. *Monthly Notices of the Royal Astronomical Society* 311.2 (Jan. 2000), pp. 441–447. ISSN: 0035-8711. DOI: [10.1046/j.1365-8711.2000.03075.x](https://doi.org/10.1046/j.1365-8711.2000.03075.x). eprint: <https://academic.oup.com/mnras/article-pdf/311/2/441/2881340/311-2-441.pdf>. URL: <https://doi.org/10.1046/j.1365-8711.2000.03075.x>.



- 
- [CWC65] A.E. Cox, S.A.R. Wynchank, and C.H. Collie. “The proton-thermal neutron capture cross section”. *Nuclear Physics* 74.3 (1965), pp. 497–507. ISSN: 0029-5582. DOI: [https://doi.org/10.1016/0029-5582\(65\)90197-5](https://doi.org/10.1016/0029-5582(65)90197-5). URL: <https://www.sciencedirect.com/science/article/pii/0029558265901975>.
- [DA22] H. Dembinski and A Abdelmotelieb. “A new maximum-likelihood method for template fits”. *Eur. Phys. J. C* 82 (2022), p. 1043. DOI: <https://doi.org/10.1140/epjc/s10052-022-11019-z>.
- [Dah09] Carl Eric 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.
- [Far+23] Sophia Farrell et al. *NESTCollaboration/nestpy: Sync with NESTv2.3.12beta*. Version v2.0.1. Jan. 2023. DOI: [10.5281/zenodo.7552304](https://doi.org/10.5281/zenodo.7552304). URL: <https://doi.org/10.5281/zenodo.7552304>.
- [Fir+17] R. B. Firestone et al. “Thermal neutron capture cross section for  $^{56}\text{Fe}(n, \gamma)$ ”. *Phys. Rev. C* 95 (1 Jan. 2017), p. 014328. DOI: [10.1103/PhysRevC.95.014328](https://doi.org/10.1103/PhysRevC.95.014328).
- [Fis+20] V. Fischer et al. “Development of an ion exchange resin for gadolinium-loaded water”. *Journal of Instrumentation* 15.07 (July 2020), P07004. DOI: [10.1088/1748-0221/15/07/P07004](https://doi.org/10.1088/1748-0221/15/07/P07004). URL: <https://dx.doi.org/10.1088/1748-0221/15/07/P07004>.
- [Gae+22] Peter Gaemers et al. *XENONnT/WFSim: v1.0.2*. Version v1.0.2. Oct. 2022. DOI: [10.5281/zenodo.7216324](https://doi.org/10.5281/zenodo.7216324). URL: <https://doi.org/10.5281/zenodo.7216324>.
- [Gei18] Christopher W. Geis. “The XENON1T water Cherenkov muon veto system and commissioning of the XENON1T Dark Matter experiment”. PhD thesis. JGU-Mainz, 2018. DOI: <http://doi.org/10.25358/openscience-2175>.
- [GZ75] K. W. Geiger and L. van der Zwan. “Radioactive neutron sources spectra from  $^9\text{Be}(\alpha, n)$  corss section data”. *Nuclear instruments and methods* 131 (1975), pp. 315–321.
- [Har+15] David Harvey et al. “The nongravitational interactions of dark matter in colliding galaxy clusters”. *Science* 347.6229 (2015), pp. 1462–1465. DOI: [10.1126/science.1261381](https://doi.org/10.1126/science.1261381). eprint: <https://www.science.org/doi/pdf/10.1126/science.1261381>. URL: <https://www.science.org/doi/abs/10.1126/science.1261381>.
- [IG85] J. Straub I. Thormählen and U. Grigull. “Refractive Index of Water and Its Dependence on Wavelength, Temperature, and Density”. *JPCRD* 14 (1985). DOI: [10.1063/1.555743](https://doi.org/10.1063/1.555743).
- [JÁN44] L. JÁNOSY. “Rate of n-fold Accidental Coincidences”. *Nature* 153 (165 1944). DOI: [10.1038/153165a0](https://doi.org/10.1038/153165a0).

- [JJ10] Goodman Jonathan and Weare Jonathan. “Ensemble samplers with affine invariance”. *Communications in Applied Mathematics and Computational Science* 5 (2010), pp. 65–80. DOI: [0.2140/camcos.2010.5.65](https://doi.org/10.2140/camcos.2010.5.65).
- [Jör+22] Florian Jörg et al. “Characterization of alpha and beta interactions in liquid xenon”. *Eur. Phys. J. C* 82.4 (2022), p. 361. DOI: [10.1140/epjc/s10052-022-10259-3](https://doi.org/10.1140/epjc/s10052-022-10259-3). arXiv: [2109.13735](https://arxiv.org/abs/2109.13735) [physics.ins-det].
- [K07] HAMAMATSU PHOTONICS K. K. *PHOTOMULTIPLIER TUBES, Basics and Applications*. Japan: HAMAMATSU PHOTONICS K. K, 2007.
- [KVA12] Michael Kuhlen, Mark Vogelsberger, and Raul Angulo. “Numerical simulations of the dark universe: State of the art and the next decade”. *Physics of the Dark Universe* 1.1 (2012). Next Decade in Dark Matter and Dark Energy, pp. 50–93. ISSN: 2212-6864. DOI: <https://doi.org/10.1016/j.dark.2012.10.002>. URL: <https://www.sciencedirect.com/science/article/pii/S2212686412000064>.
- [KW16] Hermann Kolanoski and Norbert Wermes. *Teilchendetektoren Grundlagen und Anwendungen*. Springer Spektrum Berlin, Heidelberg, 2016. DOI: [10.1007/978-3-662-45350-6](https://doi.org/10.1007/978-3-662-45350-6).
- [Lem+23] Eric W. Lemmon et al. “Thermophysical Properties of Fluid Systems in NIST Chemistry WebBook”. *NIST Standard Reference Database Number 69* (Feb. 2023). DOI: [10.18434/T4D303](https://doi.org/10.18434/T4D303).
- [Len+15] Brian Lenardo et al. “A Global Analysis of Light and Charge Yields in Liquid Xenon”. *IEEE Transactions on Nuclear Science* 62.6 (2015), pp. 3387–3396. DOI: [10.1109/TNS.2015.2481322](https://doi.org/10.1109/TNS.2015.2481322).
- [Lia+22] Shixiao Liang et al. “Domain-Informed Neural Networks for Interaction Localization Within Astroparticle Experiments”. *Front. Artif. Intell.* 5 (2022), p. 832909. DOI: [10.3389/frai.2022.832909](https://doi.org/10.3389/frai.2022.832909). arXiv: [2112.07995](https://arxiv.org/abs/2112.07995) [hep-ex].
- [LLS13] Mingfeng Lin, Henry C. Lucas, and Galit Shmueli. “Research Commentary: Too Big to Fail: Large Samples and the p-Value Problem”. *Information Systems Research* 24.4 (2013), pp. 906–917. ISSN: 10477047, 15265536. URL: <http://www.jstor.org/stable/24700283> (visited on 02/24/2023).
- [LSS63] J. Lindhard, M. Scharff, and H. Schiott. “Range concepts and heavy ion ranges”. *Matt. Fys. Medd. Dan. Vid. Selsk* 33.14 (Oct. 1963).
- [Man+10] A. Manalaysay et al. “Spatially uniform calibration of a liquid xenon detector at low energies using  $^{83\text{m}}\text{Kr}$ ”. *Review of Scientific Instruments* 81.073303 (July 2010).
- [Mar+20] Ll. Marti et al. “Evaluation of gadolinium’s action on water Cherenkov detector systems with EGADS”. *Nucl. Instrum. Methods. Phys. Res. A* 959 (2020), p. 163549. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2020.163549>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900220301315>.
- [Mas+22] D. Masson et al. *redax*. Version v2.3.1. June 2022. DOI: [10.5281/zenodo.6606161](https://doi.org/10.5281/zenodo.6606161). URL: <https://doi.org/10.5281/zenodo.6606161>.

- 
- [McC16] Christopher McCabe. “Prospects for dark matter detection with inelastic transitions of xenon”. *Journal of Cosmology and Astroparticle Physics* 2016 (2016). DOI: <https://doi.org/10.1088/1475-7516/2016/05/033>.
- [Mei+08] Dongming Mei et al. “A Model of Nuclear Recoil Scintillation Efficiency in Noble Liquids”. *Astroparticle Physics* 30 (Jan. 2008). DOI: [10.1016/j.astropartphys.2008.06.001](https://doi.org/10.1016/j.astropartphys.2008.06.001).
- [Men+20] Mendoza, E. et al. “NuDEX: A new nuclear cascades generator”. *EPJ Web Conf.* 239 (2020), p. 17006. DOI: [10.1051/epjconf/202023917006](https://doi.org/10.1051/epjconf/202023917006). URL: <https://doi.org/10.1051/epjconf/202023917006>.
- [Men+21] Y. Meng et al. “Dark Matter Search Results from the PandaX-4T Commissioning Run”. *Phys. Rev. Lett* 127 (2021), p. 261802. DOI: [10.1103/PhysRevLett.127.261802](https://doi.org/10.1103/PhysRevLett.127.261802). arXiv: [2107.13438](https://arxiv.org/abs/2107.13438) [hep-ex].
- [Mur+06] T. Murata et al. “Evaluation of the ( $\alpha$ ,xn) reaction data for JENDL/AN-2005”. *JAEA-Research* 052 (2006).
- [Mur+14] I. Murata et al. “Neutron and gamma-ray source-term characterization of AmBe sources in Osaka University”. *Progress in Nuclear Science and Technology* 4 (2014), pp. 345–348. DOI: [10.15669/pnst.4.345](https://doi.org/10.15669/pnst.4.345).
- [Mur+22] M. Murra et al. “Design, construction and commissioning of a high-flow radon removal system for XENONnT”. *Eur. Phys. J. C* 82.12 (Dec. 2022). DOI: [10.1140/epjc/s10052-022-11001-9](https://doi.org/10.1140/epjc/s10052-022-11001-9). arXiv: [2205.11492](https://arxiv.org/abs/2205.11492) [physics.ins-det].
- [Nyg74] David R. Nygren. “Proposal to investigate the feasibility of a novel concept in particle detection” (1974).
- [P19] Salucci P. “The distribution of dark matter in galaxies”. *Astron Astrophys Rev* 27 (2 Feb. 2019). DOI: [10.1007/s00159-018-0113-1](https://doi.org/10.1007/s00159-018-0113-1). URL: <https://doi.org/10.1007/s00159-018-0113-1>.
- [Pat04] Marc Paterno. “Calculating efficiencies and their uncertainties” (Dec. 2004). DOI: [10.2172/15017262](https://doi.org/10.2172/15017262).
- [Pav22] Kavargin Pavel. personal communication. July 28, 2022.
- [Pea95] Karl Pearson. “Note on Regression and Inheritance in the Case of Two Parents”. *Proceedings of the Royal Society of London Series I* (1895).
- [Pla+22] G. Plante et al. “Liquid-phase purification for multi-tonne xenon detectors”. *Eur. Phys. J. C* 82.10 (2022), p. 860. DOI: [10.1140/epjc/s10052-022-10832-w](https://doi.org/10.1140/epjc/s10052-022-10832-w). arXiv: [2205.07336](https://arxiv.org/abs/2205.07336) [physics.ins-det].
- [Plo+20] A. J. M. Plompen et al. “The joint evaluated fission and fusion nuclear data library, JEFF-3.3”. *The European Physical Journal A* 56 (2020). DOI: [10.1140/epja/s10050-020-00141-9](https://doi.org/10.1140/epja/s10050-020-00141-9).
- [RST18] Leszek Roszkowski, Enrico Maria Sessolo, and Sebastian Trojanowski. “WIMP dark matter candidates and searches—current status and future prospects”. *Reports on Progress in Physics* 81.6 (May 2018), p. 066201. DOI: [10.1088/1361-6633/aab913](https://doi.org/10.1088/1361-6633/aab913).

- [Saa+23] T. Saab et al. *Dark matter limit plotter*. Apr. 2023. URL: <https://supercdms.slac.stanford.edu/dark-matter-limit-plotter>.
- [Sch+15] J. Scherzinger et al. "Tagging fast neutrons from an  $^{241}\text{Am}/^9\text{Be}$  source". *Applied Radiation and Isotopes* 98 (2015), pp. 74–79. DOI: <http://dx.doi.org/10.1016/j.apradiso.2015.01.003>.
- [Sch19] Marc Schumann. "Direct detection of WIMP dark matter: concepts and status". 46.10 (Aug. 2019), p. 103003. DOI: [10.1088/1361-6471/ab2ea5](https://doi.org/10.1088/1361-6471/ab2ea5). URL: <https://dx.doi.org/10.1088/1361-6471/ab2ea5>.
- [Sch44] Erwin Schroedinger. "Rate of n-fold Accidental Coincidences". *Nature* 153 (1944), pp. 592–593. DOI: [10.1038/153592b0](https://doi.org/10.1038/153592b0).
- [SDA] M.J. Safari, Abbasi Davani, and H. Afarideh. "Differentiation method for localization of Compton edge in organic scintillation detectors" (). DOI: [arXiv:1610.09185](https://arxiv.org/abs/1610.09185).
- [Sor11] Peter Sorensen. "Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 635.1 (2011), pp. 41–43. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2011.01.089>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900211001756>.
- [Szy+11] M Szydagis et al. "NEST: a comprehensive model for scintillation yield in liquid xenon". 6.10 (Oct. 2011), P10002. DOI: [10.1088/1748-0221/6/10/P10002](https://doi.org/10.1088/1748-0221/6/10/P10002). URL: <https://dx.doi.org/10.1088/1748-0221/6/10/P10002>.
- [Szy+13] M Szydagis et al. "Enhancement of NEST capabilities for simulating low-energy recoils in liquid xenon". *Journal of Instrumentation* 8.10 (Oct. 2013), p. C10003. DOI: [10.1088/1748-0221/8/10/C10003](https://doi.org/10.1088/1748-0221/8/10/C10003). URL: <https://dx.doi.org/10.1088/1748-0221/8/10/C10003>.
- [Tan+20] Tomoyuki Tanaka et al. "Gamma-ray spectra from thermal neutron capture on gadolinium-155 and natural gadolinium". *Progress of Theoretical and Experimental Physics* 2020.4 (Apr. 2020). 043D02. ISSN: 2050-3911. DOI: [10.1093/ptep/ptaa015](https://doi.org/10.1093/ptep/ptaa015). eprint: <https://academic.oup.com/ptep/article-pdf/2020/4/043D02/33040537/ptaa015.pdf>. URL: <https://doi.org/10.1093/ptep/ptaa015>.
- [TI87] J. Thomas and D. A. Imel. "Recombination of electron-ion pairs in liquid argon and liquid xenon". *Phys. Rev. A* 36 (2 July 1987), pp. 614–616. DOI: [10.1103/PhysRevA.36.614](https://doi.org/10.1103/PhysRevA.36.614). URL: <https://link.aps.org/doi/10.1103/PhysRevA.36.614>.
- [UB04] Albrecht Unsöld and Bodo Bachek. *Der neue Kosmos*. 7th ed. Springer-Verlag Berlin Heidelberg, 2004. DOI: [10.1007/978-3-662-06533-4](https://doi.org/10.1007/978-3-662-06533-4).

- 
- [UR15] Teresa Marrodán Undagoitia and Ludwig Rauch. “Dark matter direct-detection experiments”. 43.1 (Dec. 2015), p. 013001. DOI: [10.1088/0954-3899/43/1/013001](https://doi.org/10.1088/0954-3899/43/1/013001). URL: <https://dx.doi.org/10.1088/0954-3899/43/1/013001>.
- [Wil+99] W. Wilson et al. “SOURCES-4A”. *tech. rep. LA-13639-MS, Los Alamos U.S.A* (1999).
- [Xu+11] Xin Xu et al. “Robust Automatic Focus Algorithm for Low Contrast Images Using a New Contrast Measure”. *Sensors* (2011), pp. 8281–8294. DOI: [10.3390/s110908281](https://doi.org/10.3390/s110908281).
- [Zha+16] Y. Zhang et al. “First measurement of radioactive isotope production through cosmic-ray muon spallation in Super-Kamiokande IV”. *Phys. Rev. D* 93 (1 Jan. 2016), p. 012004. DOI: [10.1103/PhysRevD.93.012004](https://doi.org/10.1103/PhysRevD.93.012004). URL: <https://link.aps.org/doi/10.1103/PhysRevD.93.012004>.















