

DISSERTATION

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences
of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

LUISA MARIA HÖTZSCH

born in: Lauingen (Donau), Germany

Oral examination: 25 July, 2024

**Nuclear Recoil modeling for the first
XENONnT Dark Matter Search Results
and
Electronic Recoil Yield Measurements at low
Energies and Fields**

1st Referee: Priv.-Doz. Dr. Teresa Marrodán Undagoitia

2nd Referee: Prof. Dr. James A. (Jim) Hinton

Abstract

Many astrophysical and cosmological observations point to the existence of a massive, non-luminous and non-baryonic component of matter that is called dark matter (DM). The XENON Dark Matter project aims for the direct detection of DM in the form of weakly interacting massive particles (WIMPs), which are one of the most compelling candidates for particle DM. Utilizing the concept of a dual-phase xenon time projection chamber (TPC), the XENONnT experiment is the latest installment in the XENON experiment series, and was designed as a fast upgrade of the predecessor experiment XENON1T. With an increased liquid xenon (LXe) mass of 5.9 tonnes and novel subsystems, the XENONnT detector has reached an unprecedented purity in both electro-negative and radioactive radon contaminations, and has been taking data since 2021. In the first WIMP search results obtained with a blind analysis of the 1.1 tonne-years of exposure taken in XENONnT's first science run, SR0, stringent upper limits on the WIMP-nucleon interaction cross-section were set, with a minimum of $2.58 \times 10^{-47} \text{ cm}^2$ for a WIMP mass of $28 \text{ GeV}/c^2$ at a 90% confidence level.

This work is centered around the XENONnT experiment and its first WIMP search results. During the construction phase of the experiment, the photomultiplier tubes (PMTs) used to detect the scintillation light induced by particle interactions in the TPC were characterized, selected and installed. Their performance in the XENONnT TPC was analyzed and monitored throughout detector commissioning and the first science run, focussing on the identification of vacuum degradation and afterpulse behavior. For the first WIMP search analysis, the detector's response to nuclear recoil interactions expected for a WIMP signal was modeled, using neutron calibration data. Additionally, the small-scale Heidelberg Xenon (HeXe) TPC was used to investigate the response of LXe to γ -ray interactions under the same experimental conditions as in the XENONnT detector.

Kurzzusammenfassung

Viele Beobachtungen in der Astrophysik und der Kosmologie deuten darauf hin, dass es eine massive, nicht-leuchtende und nicht-baryonische Materiekomponente gibt, die als Dunkle Materie (DM) bezeichnet wird. Das XENON Dark Matter Project hat das Ziel, Dunkle Materie in Form von WIMPs (Englisch für "Weakly Interacting Massive Particles", in etwa "schwach wechselwirkende massive Teilchen"), die als eine der überzeugendsten Kandidaten für DM-Teilchen gelten, direkt zu detektieren. Das XENONnT Experiment, eine zwei-Phasen Xenon-Zeitprojektionskammer (TPC, Englisch "Time Projection Chamber"), ist die neueste Installation in der XENON Experimentreihe, und wurde als schnelles Upgrade des Vorgängerexperiments XENON1T entwickelt. Mit einer erhöhten Targetmasse an flüssigem Xenon von 5,9 Tonnen und neuen Untersystemen hat der XENONnT-Detektor eine bisher unerreichte Reinheit in elektronegativen und radioaktiven Radon Kontaminationen erreicht, und nimmt seit 2021 Daten auf. Mit den ersten WIMP-Suchergebnissen, die aus einer Blindanalyse der 1,1 Tonnen-Jahre Exposition in der ersten wissenschaftlichen Datennahme des Experiments gewonnen wurden, wurden strenge obere Grenzen für den WIMP-Nukleon-Wechselwirkungsquerschnitt gesetzt, wobei ein Minimum von $2.58 \cdot 10^{-47} \text{ cm}^2$ für eine WIMP-Masse von $28 \text{ GeV}/c^2$ auf einem Konfidenzniveau von 90 % erreicht wurde.

Diese Arbeit konzentriert sich auf das XENONnT-Experiment und seine ersten Ergebnisse der WIMP-Suche. Während der Bauphase des Experiments wurden die Photovervielfacher (PMTs, vom Englischen "Photomultiplier Tubes"), die zur Detektion des durch Teilcheninteraktionen in Xenon erzeugten Szintillationslichts verwendet werden, charakterisiert, ausgewählt und installiert. Ihr Verhalten in der XENONnT-TPC wurde während der Inbetriebnahme des Detektors und der ersten wissenschaftlichen Datennahme analysiert und überwacht, wobei besonderes Augenmerk auf die Identifizierung von Vakuumdegradation gelegt wurde. Für die Analyse in der ersten WIMP-Suche wurde die Reaktion des Detektors auf Kernrückstoßinteraktionen, die für ein WIMP-Signal erwartet werden, modelliert, unter Verwendung von Neutronenkalibrierungsdaten. Zusätzlich wurde die Heidelberg Xenon (HeXe) TPC verwendet, um die Reaktion von flüssigem Xenon in Interaktionen von Gammastrahlung unter denselben experimentellen Bedingungen wie im XENONnT-Detektor zu untersuchen.

Contents

1. Introduction	11
2. Dual-phase xenon TPCs and the XENONnT experiment	13
2.1. Xenon as a detector medium	13
2.1.1. Properties of xenon	13
2.1.2. Signal generation in xenon	14
2.2. Dual-phase xenon TPCs	14
2.2.1. The dual-phase xenon TPC principle	14
2.2.2. Signal corrections	16
2.2.3. Energy reconstruction	18
2.3. The XENONnT experiment	18
2.3.1. The XENONnT TPC	18
2.3.2. The muon and neutron veto systems	20
2.3.3. Xenon distillation and purification systems	20
3. The Eyes of XENONnT: Photomultiplier tests, selection, and installation	23
3.1. The XENONnT Photomultiplier Tubes	24
3.1.1. Basic PMT properties	26
3.2. Improved PMT qualification tests for XENONnT	27
3.2.1. Overview of facilities and testing procedure	28
3.2.2. Basic characteristics tests	29
3.2.3. Light emission	31
3.2.4. Vacuum degradation tests using afterpulses	33
3.2.5. Summary of qualification test results	36
3.3. The XENONnT PMT arrays	37
3.3.1. PMT selection and array distribution	38
3.3.2. PMT array installation	41
4. PMT afterpulse characterization and monitoring in XENONnT	47
4.1. Afterpulses characteristics and identification	47
4.2. PMT calibration system and afterpulse measurements	50
4.2.1. PMT afterpulse measurements	51
4.2.2. Afterpulse data processing and analysis	52
4.3. Afterpulse monitoring in XENONnT	55
4.3.1. Commissioning phases	56
4.3.2. First science run (SR0)	59
4.4. Summary and Outlook	62

5. Modeling the Nuclear Recoil Response in XENONnT	65
5.1. Nuclear recoil calibration with AmBe neutrons	66
5.1.1. AmBe source and the NV coincidence method	66
5.1.2. NR band data and event selection	67
5.2. Introduction to the basics of the fitting tool BBF	73
5.3. Modeling the NR response to AmBe calibration data	76
5.3.1. NR quanta emission model in LXe	77
5.3.2. Detector reconstruction model	79
5.3.3. Multi-scatter NR handling	83
5.4. Putting it together: Fitting the model	86
5.4.1. Fit results	86
5.4.2. Goodness-of-fit tests and model validation	89
5.4.3. Nuclear Recoil photon and charge yields	93
5.5. Conclusions	95
6. First WIMP search results of XENONnT	97
6.1. SR0 Overview	97
6.2. Signal and Background Models	99
6.3. Statistical Inference and WIMP Results	104
6.4. Conclusions and Outlook	107
7. Measurement of ER yields at low energy and fields in HeXe	109
7.1. The HeXe setup and measurement overview	110
7.1.1. The HeXe TPC	111
7.1.2. Coincidence measurement overview and setup	113
7.1.3. Data taking and processing	115
7.2. Calibrations and detector characterization	118
7.2.1. Energy calibration of the NaI detector	118
7.2.2. TPC response characterization with ^{83m}Kr	120
7.3. Coincidence Measurements	126
7.3.1. Coincidence timing calibration	127
7.3.2. Data selection for low energy γ -ray events	128
7.3.3. Estimation of γ -ray energy and electric field distributions	131
7.3.4. Results for the ER γ -ray yields	134
7.4. Summary, conclusions and outlook	138
8. Summary	141
Appendix	145
A. Author's Publications	147
B. References	149
Acknowledgements	157

Chapter 1.

Introduction

Historically, the first evidence for the existence of DM came from astronomical observations the movement of galaxies in galaxy clusters [1] and of galactic rotation curves [2]. In both cases, the observed velocities were too high to be explained by the visible mass alone, postulating the need for additional mass for the observed objects to remain gravitationally bound. More modern observations of gravitational lensing [3], the cosmic microwave background [4] and large-scale structure formation in the universe [5] have made the impact of DM on the universe even more apparent, concluding that about 27% of the universe's energy density is made up of DM.

The by far most popular explanation for the observed phenomena is the existence of a new dark matter particle, that apart from gravitationally only interacts very weakly with ordinary baryonic matter. DM particles would thus form a halos around galaxies, explaining the observed rotation curves. One of the most popular candidates for DM are Weakly Interacting Massive Particle (WIMPs), which interact with ordinary matter only through the weak nuclear force. This allows for their direct detection in an Earth-based experiment moving through the Milky Way's DM halo, by looking for nuclear recoil signals induced by the interaction.

The most important aspect of direct DM detection is the expected interaction rate between DM particles and a nucleus in the detector target material. The differential event rate of WIMPs is expressed as [6]

$$\frac{dR}{dE_R} = \frac{M_T}{m_N} \cdot \frac{\rho_\chi}{m_\chi} \int_{v_{\min}}^{v_{\text{esc}}} dv v f(\vec{v}) \frac{d\sigma_{\chi,N}}{dE_R}(v), \quad (1.1)$$

where M_T is the total target mass and m_N is the atomic mass of the target nucleus, both determined by the choice of target material. Astrophysical inputs are needed as well, which are the density of the Milky Way's DM halo at Earth's position, ρ_χ , and the DM velocity distribution $f(v)$ in the rest frame of the detector. The velocity distribution of the Milky Way's DM halo is usually assumed to follow an isotropic Maxwell-Boltzmann distribution. The minimum velocity v_{\min} required for a DM particle to induce a nuclear recoil of energy E_R and the escape velocity v_{esc} of the Milky Way form the integration limits. The astrophysical inputs are usually taken from the Standard Halo Model (SHM) [7], which is a set of recommended values for the DM halo parameters, allowing for a comparison of experimental results in direct DM detection.

Apart from the DM particle mass m_χ , the actual particle physics that we're interested in is contained in the differential DM-nucleus cross-section $d\sigma_{\chi,N}/dE_R$. Assuming that the scattering process is isotropic, the differential cross section can then be written as

$$\frac{d\sigma_{\chi,N}}{dE_R}(v) = \frac{m_N}{2\mu_{\chi,N}^2 v^2} \sigma_{\chi,N}. \quad (1.2)$$

Assume spin-independent (SI) interactions, the cross section is then given by

$$\sigma_{\chi,N} = \sigma_0^{\text{SI}} F_{\text{SI}}^2 \quad (1.3)$$

where σ_0^{SI} is the cross-section at zero momentum transfer, and F_{SI} describes the nuclear form factor. The former is given by

$$\sigma_0^{\text{SI}} = \sigma_{\chi,n} \frac{\mu_{\chi,N}^2}{\mu_{\chi,n}^2} (Zf_p + (A-Z)f_n)^2 \approx \sigma_{\chi,n} \frac{\mu_{\chi,N}^2}{\mu_{\chi,n}^2} A^2, \quad (1.4)$$

where $\sigma_{\chi,n}$ is the WIMP-nucleon cross section, $\mu_{\chi,n}$ is the WIMP-nucleon reduced mass, Z is the atomic number of the target nucleus, A is the atomic mass number, and f_p and f_n are the effective couplings of the WIMP to the proton and neutron, respectively. The approximation assumes that the effective couplings are similar for protons and neutrons, for which the SI cross-section scales with the square of the atomic mass number A^2 . This is important for the choice of target material, with heavier nuclei being more sensitive to WIMP interactions.

In order to be able to detect the interaction of potential WIMPs with the target material, it needs to be able to produce observable signals from the energy deposited by the recoiling nucleus in the scattering process. Energy depositions in a detector material are converted into one of three possible signal channels: ionization, scintillation, or atomic motion, corresponding to a detection of electrons, photons, or heat (phonons), respectively. Which of these can actually be read out depends on the detector technology used. In this work, we focus on DM detection using xenon as a target material, the highest- A yet stable noble gas.

Chapter 2.

Dual-phase xenon TPCs and the XENONnT experiment

The concept of noble liquid dual-phase time projection chambers (TPCs) is among the most successful technologies for rare event searches. In particular, liquid xenon (LXe) TPCs have consistently set world-leading limits on the interaction cross-sections of Weakly Interacting Massive Particles (WIMPs) with ordinary matter for almost two decades now. The XENON Collaboration has been at the forefront of this field for many years, developing a series of increasingly larger and more sensitive detectors, culminating in the current iteration, the XENONnT experiment [8].

In this chapter, we will focus on the principles of xenon dual phase TPCs and how they are applied in the XENONnT experiment. General properties of and the processes of signal generation in xenon are introduced in [section 2.1](#). In [section 2.2](#), the working principle of a dual-phase xenon TPC is described, and a short overview of the signal corrections for detector effects and the energy reconstruction process is given. Then, the XENONnT experiment is presented in [section 2.3](#), including the detector design of the TPC and veto detectors, and an overview of the various supporting subsystems needed for the successful operation of the experiment.

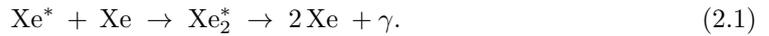
2.1. Xenon as a detector medium

2.1.1. Properties of xenon

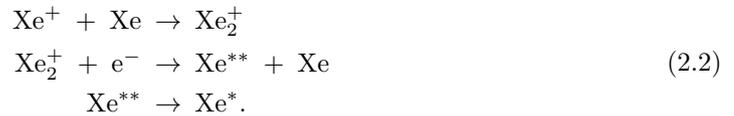
Xenon is a noble gas with atomic number 54 that is naturally present in the Earth's atmosphere at a concentration of about 0.09 ppm, and has many properties that make it an ideal detector medium for rare event searches. Being chemically inert, xenon is relatively easy to handle and purify even in large amounts without strong reactions with other materials. Its high atomic mass number of $A \approx 131$ makes it an excellent target for WIMP-nucleus scattering, greatly enhancing the interaction probability due to the A^2 scaling of the spin-independent cross-section. Liquid xenon has a high self-shielding capability against external backgrounds due to its high density of about 3 g/cm^3 and a resulting high stopping power for γ -ray-rays from detector materials. In its natural isotopic composition, xenon has a low intrinsic background with only two long-lived radioactive isotopes, ^{124}Xe and ^{136}Xe , that do not contribute significantly to the background in the energy region of interest for WIMP searches. The scintillation process in xenon via excimer decay (see [subsection 2.1.2](#)) makes the medium transparent to its own scintillation light, which has a wavelength of 175 nm [9]. Despite being in the vacuum ultraviolet (VUV) range, this wavelength is still detectable at high efficiencies using photosensors such as photomultiplier tubes (PMTs), without the need for wavelength shifters.

2.1.2. Signal generation in xenon

When a particle traverses through a volume of liquid xenon, it interacts with the xenon atoms and thereby deposits energy, leading to the creation of ionization electrons and scintillation photons. Depending on the type of incoming particle, it will either scatter off an electron in the shell of a xenon atom, which recoils and is ejected in a so-called electronic recoil (ER) event, or off a xenon nucleus via elastic scattering in a nuclear recoil (NR) event. In both cases, the recoiling particle – liberated electron or recoiling xenon atom – starts moving through the surrounding medium and thereby ionizes and excites nearby xenon atoms, forming track of electron-ion pairs and excited xenon atoms Xe^* , also called ‘excitons’. The excitons quickly form ‘excimers’ Xe_2^* (excited dimers) with neutral xenon atoms in the vicinity, which decay via the emission of VUV scintillation photons at a wavelength of 175 nm [9],



The ionization electrons may either escape from the electron-ion cloud at the interaction site, or recombine with a xenon ion via the process of electron-ion recombination, where an exciton is formed again:



The last transition from a higher excited state Xe^{**} to the exciton state Xe^* is non-radiative. The recombination excitons themselves undergo the same process as the ones formed by direct excitation in Equation 2.1 above, thereby adding to the number of scintillation photons created in the interaction. If an electric drift field is applied to the xenon volume, ionization electrons can be actively removed from the interaction site and eventually be read out as a second signal in addition to the scintillation photons. The discrimination between ER and NR events is then based on the differing stopping power between electrons and recoiling xenon atoms while moving through the surrounding xenon, and the resulting differences in ionization track density and electron-ion recombination between the two event types. Electrons lose their energy solely via electronic stopping, i.e. the energy loss per unit length in inelastic Coulomb interactions with medium electrons, which produces ionization and excitation, and is the by far dominant process in ER interactions. In contrast, recoiling xenon atoms lose their energy via both electronic and nuclear stopping, where the latter is the energy loss per unit length via elastic collisions with other xenon atoms, a process that translates into atomic motion, i.e. heat or phonons, which can not be detected in LXe. This leads to a reduced number of detectable quanta (scintillation photons and ionization electrons) compared to ER interactions of the same deposited energy, a process called nuclear quenching. Additionally, the overall higher stopping power in NR interactions leads to a higher ionization track density and thus a higher probability of electron-ion recombination, resulting in a smaller charge-to-light ratio in NR events compared to ER events, which forms the basis for particle type discrimination in LXe TPCs.

2.2. Dual-phase xenon TPCs

2.2.1. The dual-phase xenon TPC principle

One of the methods of reading out the scintillation photons and ionization electrons produced by particle interactions in LXe is the principle of a dual-phase time projection chamber (TPC), illustrated in Figure 2.1. The TPC consists of a typically cylindrical LXe volume with a small layer of gaseous xenon (GXe) above the liquid. Two arrays of photosensors are positioned at the top and bottom of the TPC, that directly detect the prompt scintillation light produced in

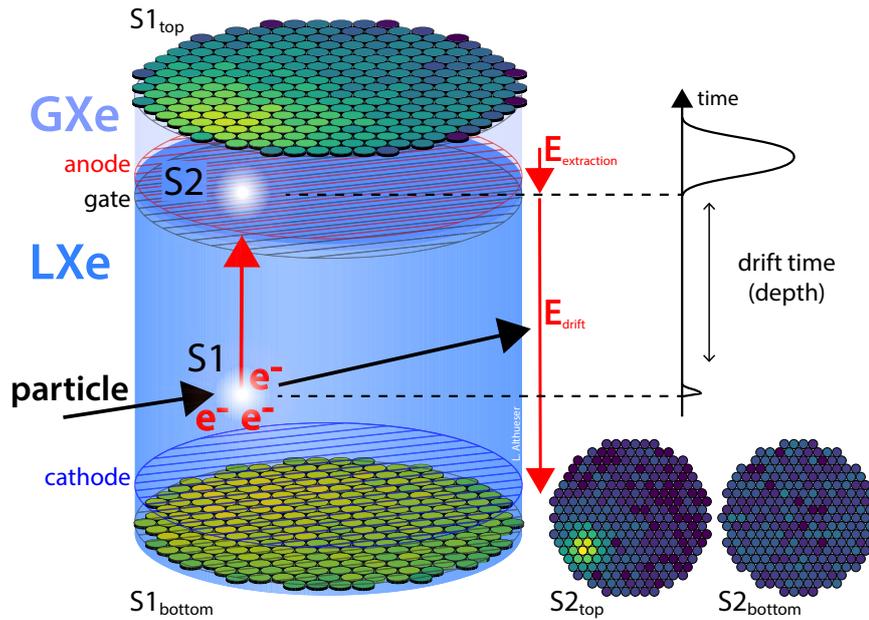


Figure 2.1.: Working principle of a xenon dual-phase time projection chamber (TPC). Figure credit Lutz Althüser, XENON Collaboration.

the liquid, which is called the S1 signal. The ionization electrons produced in the interaction are drifted upwards towards the liquid surface by applying an electric drift field between a cathode electrode at the bottom of the TPC and a gate electrode positioned just below the liquid xenon surface. A second, stronger electric field applied between the gate and an anode electrode positioned above the surface then extracts and accelerates the electrons into the gas phase, producing secondary xenon scintillation light via the electro-luminescence process that is proportional to the number of extracted electrons [10]. These photons are detected by the photosensors as the so-called S2 signal. The depth z of an interaction in TPC can be inferred from the time difference between the observed S1 and the S2 signals, given the electron drift velocity for a given drift field strength, while the position in the horizontal xy -plane is inferred from the photosensor hit pattern in the top array. This results in the full 3D position reconstruction of an interaction in the detector volume which allows for the definition of an inner fiducial volume with a reduced background rate from external sources, aided by the self-shielding properties of LXe. Additionally, in contrast to WIMPs some background sources such as γ -rays and radiogenic neutrons can scatter multiple times inside the LXe volume,

Most commonly, the choice of photosensor in large LXe TPCs are photomultiplier tubes (PMTs), that convert incoming photons into photoelectrons (PEs) via the photoelectric effect, which are then multiplied and read out as a charge signal pulse. After calibration of the PMT responses to single photons, the summed S1 and S2 signal areas from all PMTs are usually given in a detected number of photo-electrons, with unit PE. S1 and S2 signals can then be differentiated from each other based on their signal areas and pulse shape. S1s are prompt signals with a pulse-like shape, with a fast rising edge and an exponentially falling tail from the scintillation decay process in xenon, with signal widths in the order of a few tens of nanoseconds. Due to the amplification in the proportional scintillation process of S2 signals, they are generally much larger than S1s and have a roughly Gaussian pulse shape with widths in the order of microseconds.

After correcting the S1 and S2 signals detected in a particle interaction for spatially dependent signal losses, as explained in the next subsection 2.2.2, plotting the S2 signal versus the S1 signal areas reveals two band-shaped populations, as shown in Figure 2.2. These correspond to ER recoil (upper band, dark blue) and NR recoil (lower band, cyan) events, and are therefore

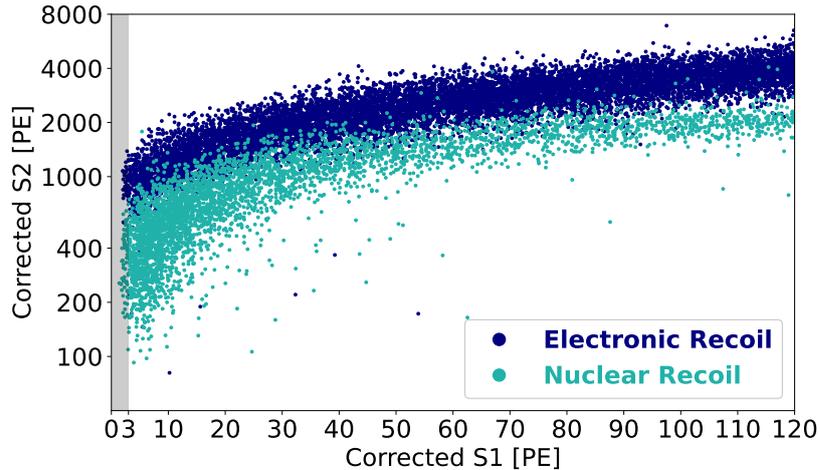


Figure 2.2.: Electronic recoil (ER) and nuclear recoil (NR) band in the S2 versus S1 area space of xenon dual-phase TPC. Data from XENON1T, credit XENON Collaboration.

called, unsurprisingly, the ER and NR band. As explained in [subsection 2.1.2](#), this is a direct consequence of the different stopping power of electrons and recoiling xenon atoms in the medium, and the resulting differences in ionization track density and electron-ion recombination between the two event types. While WIMPs are expected to scatter elastically off xenon nuclei producing NR events, the vast majority of background sources are from natural β - and γ -radiation that produce ER events. The S1-S2 signal space is therefore used to discriminate between the two event types, giving a powerful handle on background rejection in WIMP DM searches.

2.2.2. Signal corrections

Between the process of production of detectable quanta (i.e. the scintillation photons and escaped ionization electrons) in an interaction, and the detected S1 and S2 signals, losses can occur that are dependent on many parameters, such as the detector geometry and the position of the interaction, the electric drift field shape and strength, the purity of the xenon, and the properties of the photosensors. The detected and recorded signal sizes in PE therefore do not directly correspond to the number of produced photons or electrons in an interaction, and are corrected for these effects as described in the following.

Scintillation photons are detected with a certain efficiency that is dependent on the position of the interaction, called the scintillation gain, which is expressed as

$$g'_1(x, y, z) = \varepsilon_{\text{LCE}}(x, y, z) \cdot \varepsilon_{\text{ph}}. \quad (2.3)$$

Here ε_{LCE} denotes the position-dependent light collection efficiency (LCE), defined as the fraction of produced photons that arrive at a photosensor. S1 photons that do not arrive are lost due to Rayleigh scattering and absorption in LXe, as well as reflections on the PFTE surfaces of the TPC wall or total reflection on the liquid-gas interface. Of the photons that do impinge on the sensitive area of a photosensor, only a fraction are actually detected by the sensor, given by the photon detection efficiency ε_{ph} . For photomultiplier tubes, the most commonly used photosensors in xenon TPCs, the photon detection efficiency is the product of the quantum efficiency QE and collection efficiency CE ,

$$\varepsilon_{\text{ph}} = QE \cdot CE. \quad (2.4)$$

The QE is the probability of photoelectron emission from a PMT's photocathode after absorbing a photon, and CE is the efficiency with which an emitted photoelectron (PE) is successfully guided to the first amplification stage within the tubes (details on PMTs will be covered in [chapter 3](#)). The scintillation gain is usually given in units of number of detected photons per produced scintillation photon, PE/n_γ .

The relative changes of the scintillation gain over the detector volume, also called the light yield LY_{rel} , can be measured using mono-energetic ER calibration sources that distribute homogeneously in the sensitive volume of the TPC. Typically, $^{83\text{m}}\text{Kr}$ is used for this purpose [11]. Defining an average scintillation gain over the sensitive volume, $g_1 = \langle g'_1(x, y, z) \rangle$, we can express this as

$$LY_{\text{rel}}(x, y, z) = \frac{g'_1(x, y, z)}{g_1} = \frac{S1(x, y, z)}{\langle S1(x, y, z) \rangle}_{\text{Kr}}. \quad (2.5)$$

A measured S1 signal's size can then be corrected for the spatial dependence of the scintillation gain, independent of the underlying true number of produced scintillation photons, yielding the corrected S1 signal's area called cS1, with

$$\text{cS1} = LY_{\text{rel}}^{-1}(x, y, z) \cdot S1 \quad (2.6)$$

where the inverse of the relative light yield map is used.

For the produced and escaped electrons eventually forming an S2 signal, losses can occur as well. During their drift through the LXe volume towards the gate electrode just below the liquid-gas interface, the electrons can attach to electronegative impurities in the xenon, and are lost. The drift survival probability p_{drift} of an electron is an exponential function of the time the electrons need to drift to the liquid surface, and therefore depends on the interaction depth z and the drift velocity v_d :

$$p_{\text{drift}}(z) = \exp(-z/(\tau_e \cdot v_d)). \quad (2.7)$$

The time constant τ_e is the average time after which an electron is lost, and is called the *electron lifetime*. The resulting z -dependence of an S2 signal is corrected for this effect by determining the electron lifetime (which may vary over time due to changes in impurity concentrations) as well as the drift velocity. S2 signals are then corrected to the size corresponding to $z = 0$, such that the drift-loss corrected S2 signal becomes

$$S2_{\text{elifecorr}} = S2 \cdot p_{\text{drift}}(z)|_{Kr} / p_{\text{drift}}(z = 0)|_{Kr}. \quad (2.8)$$

Similar to the S1 case above, an xy -dependent gain factor is then defined,

$$g'_2(x, y) = G(x, y) \cdot \varepsilon_{\text{extr}}(x, y), \quad (2.9)$$

with the where $\varepsilon_{\text{extr}}(x, y)$ is the extraction efficiency, defined as the fraction of drift-surviving electrons that get successfully extracted and accelerated into the gas phase, and $G(x, y)$ is the gas gain, the amplification factor of electrons into PEs in the final S2 signal. The gas gain itself is the product of the single electron (SE) gain $g_{\text{SE}}(x, y)$, which is the number of scintillation photons produced in the gas phase by one electron being accelerated through it, and, similar to the S1 case, the total efficiency with which these S2 photons reach and are detected by the photosensor arrays $\varepsilon_{\text{LCE}, S2}(x, y)$,

$$G(x, y) = g_{\text{SE}}(x, y) \cdot \varepsilon_{\text{LCE}, S2}(x, y). \quad (2.10)$$

The position-dependence in the xy -plane is a result of both the optical light collection efficiency for S2 photons and local variations in the extraction field strength due to electrode grid geometry, resulting in an inhomogeneous electroluminescence amplification. A relative spatial S2 correction in the xy -plane is derived similar to the S1 case from mono-energetic calibration

data, giving

$$QY_{\text{rel}}(x, y) = \frac{g'_2(x, y)}{g_2} = \frac{S2_{\text{elifecorr}}(x, y)}{\langle S2_{\text{elifecorr}}(x, y) \rangle} \quad (2.11)$$

where $g_2 = \langle g'_2(x, y) \rangle$ is the average *ionization gain*, with units of PE per ionization electron, PE/ n_e . The total corrected S2 signal size cS2 then becomes

$$\text{cS2} = QY_{\text{rel}}^{-1}(x, y) \cdot S2_{\text{elifecorr}} = QY_{\text{rel}}^{-1}(x, y) \cdot S2 \cdot p_{\text{drift}}(z). \quad (2.12)$$

2.2.3. Energy reconstruction

In ER interactions, all the deposited energy is converted into scintillation photons and ionization electrons without any quenching effects. While the exact number of scintillation photons and ionization electrons produced in an interaction varies due to fluctuations in the electron recombination process, the total number of produced quanta N_q is conserved for a given deposited energy, and scales linearly with the energy. With the average energy needed to produce a single quantum W , the total energy E deposited in an interaction can be expressed as reconstructed from the cS1 and cS2 signals as

$$E = W \cdot N_q = W \cdot (n_\gamma + n_{e^-}) = W \cdot \left(\frac{\text{cS1}}{g_1} + \frac{\text{cS2}}{g_2} \right). \quad (2.13)$$

This is also called the *combined energy scale* (CES), and is usually given in units of keV_{ee} (electron-equivalent energy). The combined energy scale can be calibrated using several mono-energetic ER lines, in a method commonly referred to as the ‘Doke plot’ [12]. Here the peak positions of the lines are plotted in cS2/ E versus cS1/ E (with E the known energy of the respective source), which fall on a straight line due to the anti-correlation between the two signals. Rearranging Equation 2.13, we can simply extract the average scintillation and ionization gains g_1 and g_2 from the slope and intercept of the line in the Doke plot, via

$$\frac{\text{cS2}}{E} = \frac{g_2}{W} - \frac{g_2}{g_1} \cdot \frac{\text{cS1}}{E}. \quad (2.14)$$

2.3. The XENONnT experiment

The XENONnT experiment is located at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy, is the latest installment of DM search experiments in the XENON program. Designed as a fast upgrade of the previous XENON1T experiment [13], a large part of the infrastructure and subsystems are reused. The heart of the experiment, the XENONnT TPC, was designed and constructed newly, with an increased active target mass of 5.9 t of LXe, a factor of ~ 3 larger than XENON1T.

The entire experimental setup is described in detail in [8], here we will give a brief overview of the TPC design and active muon and neutron vetos.

2.3.1. The XENONnT TPC

The XENONnT TPC is shown in Figure 2.3. It is housed inside a stainless steel, double-walled cryostat, with the inner vessel holding in total 8.5 t of cryogenic liquid xenon at -98° . The TPC is a cylindrical structure with the wall surface made up of thin PTFE reflector panels. It has a diameter of ~ 1.34 m with a total active xenon mass of 5.9 t, and a height of ~ 1.5 m. The TPC is equipped with 494 Hamamatsu R11410-21 3-inch photomultiplier tubes (PMTs) (details in chapter 3), arranged in a maximum-packing density hexagonal pattern in the two arrays at the top and bottom of the TPC. The array surfaces between the PMT windows facing the active

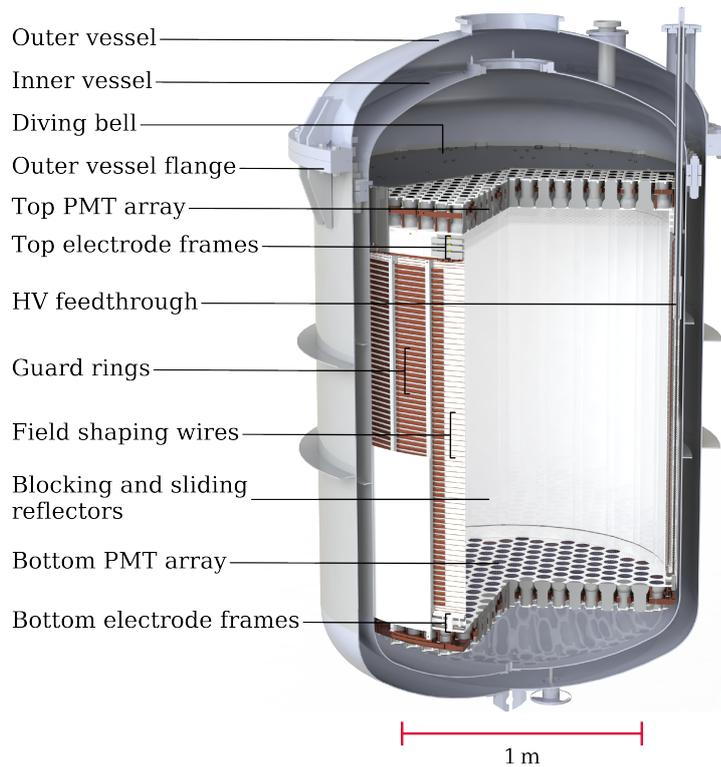


Figure 2.3.: Render of the XENONnT cryostat housing the TPC, with relevant components labeled. Figure from [15].

volume are covered with a PTFE disk. The use of PTFE reflector panels and disks increases the light collection efficiency in the TPC, with diamond-polished surfaces that maximize the reflection of VUV scintillation light.

The electric drift and extraction fields are established by five electrodes, cathode, gate, anode, and two screening electrodes placed close to either of the PMT arrays to reduce the electric field close to the tubes. A field cage formed by concentric field shaping wires and guard rings made from OFHC copper surrounds the wall reflector panels on the outer side. They are connected to the cathode via a resistor chain to ensure a homogeneous electric field in the TPC volume. The top-most ring just below the gate electrode can be biased independently, allowing for a fine-tuning of the electric field strength homogeneity in the top region of the TPC [14]. The cathode, gate and anode electrodes are made from $\mathcal{O}(100\mu\text{m})$ thin, parallel stainless-steel wires with a pitch of 5 mm (7.5 mm) for the anode, gate and top screening (cathode and bottom screening) electrodes. In order to reduce the effect of wire sagging due to the electric field between the gate and anode electrodes, they feature two and four additional support wires, respectively, that run perpendicular to the main wires and are therefore called *perpendicular wires*.

During the commissioning phase of the detector, an electric short occurred between the cathode and the bottom screening electrode, that limited the maximum voltage that could be applied to the cathode such that the designed drift field strength could not be reached. For the first science run SR0, the cathode could only be set to -2.75 kV , which resulted in a drift field strength of 23 V/cm , about an order of magnitude lower than designed. The potential bias between gate and anode had to be lowered as well, due to the observation of the emission of bursts of small S2 signals from single or few electrons, limiting the extraction field to 2.9 kV/cm .

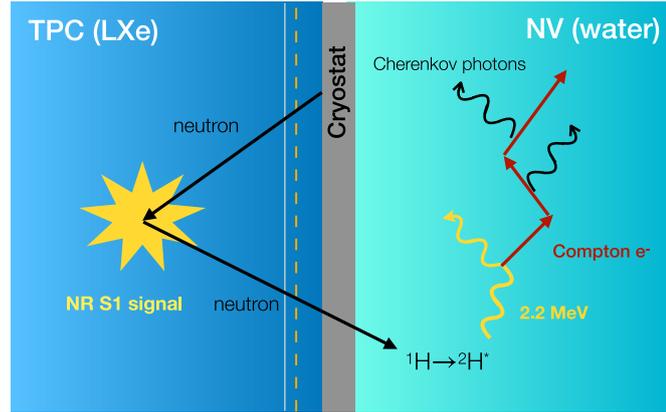


Figure 2.4.: Illustration of the working principle of the XENONnT neutron veto (NV).

2.3.2. The muon and neutron veto systems

The cryostat housing the TPC is nested inside two additional outer veto detector systems that are installed inside a ~ 10 m diameter and height water tank (WT) filled with about 700 tonnes of pure water. Both vetos are designed to tag and reject background events originating from neutrons. Neutron interactions in the TPC produce NR events, and therefore pose a dangerous background source in WIMP searches.

A render of the XENONnT WT is shown in Figure 2.5. Adopted from XENON1T, the outer part of the WT is instrumented with 84 PMTs (Hamamatsu R5912ASSY 8-inch) forming the active muon veto (MV). Muons can interact in the vicinity of the TPC and produce neutrons via spallation reactions, but can be tagged and removed from the data by the detection of the produced Cherenkov light when muons traverse the WT.

In an upgrade from XENON1T, the inner part of the WT is now instrumented with 120 additional PMTs (R5912-100-10 8-inch) forming the water Cherenkov neutron veto (NV), designed to tag and reject radiogenic neutrons produced in the detector materials via (α, n) reactions and spontaneous fission. While the two veto systems share the same water volume, they are optically separated by ePTFE sheets arranged surrounding the TPC cryostat, as can be seen in Figure 2.5. The working principle of the NV is illustrated in Figure 2.4. A radiogenic neutron produced in detector materials, e.g. in the cryostat, can interact in the LXe volume in the TPC producing a nuclear recoil (NR) event. It then may exit the TPC and enter the surrounding water volume of the NV, where it moderates and gets eventually captured on hydrogen, forming deuterium in an excited state ${}^2\text{H}^*$. The deuterium decay produces a 2.2 MeV γ -ray that Compton scatters, producing Compton electrons which in turn produce Cherenkov light that can be detected by the NV PMTs. Thus neutron events in the TPC can be tagged and rejected by the time coincident detection of the S1 signal of the NR in the TPC and the Cherenkov light in the NV.

In order to calibrate both the NV and the TPC, an AmBe neutron source can be inserted into the NV volume via two tubes that are guided from the top of the WT and wrap around the TPC cryostat. Due to their ‘U’-shape, they are called the top and bottom u-tubes, shown in Figure 2.5 in red and green, respectively. This allows for the source to be brought close to the TPC at various positions for NR response calibrations of the TPC and NV detection and tagging efficiency calibrations.

2.3.3. Xenon distillation and purification systems

In addition to the external background sources mentioned so far, intrinsic background sources contribute to the ER background in the TPC as well, originating from the β -decays of ${}^{85}\text{Kr}$

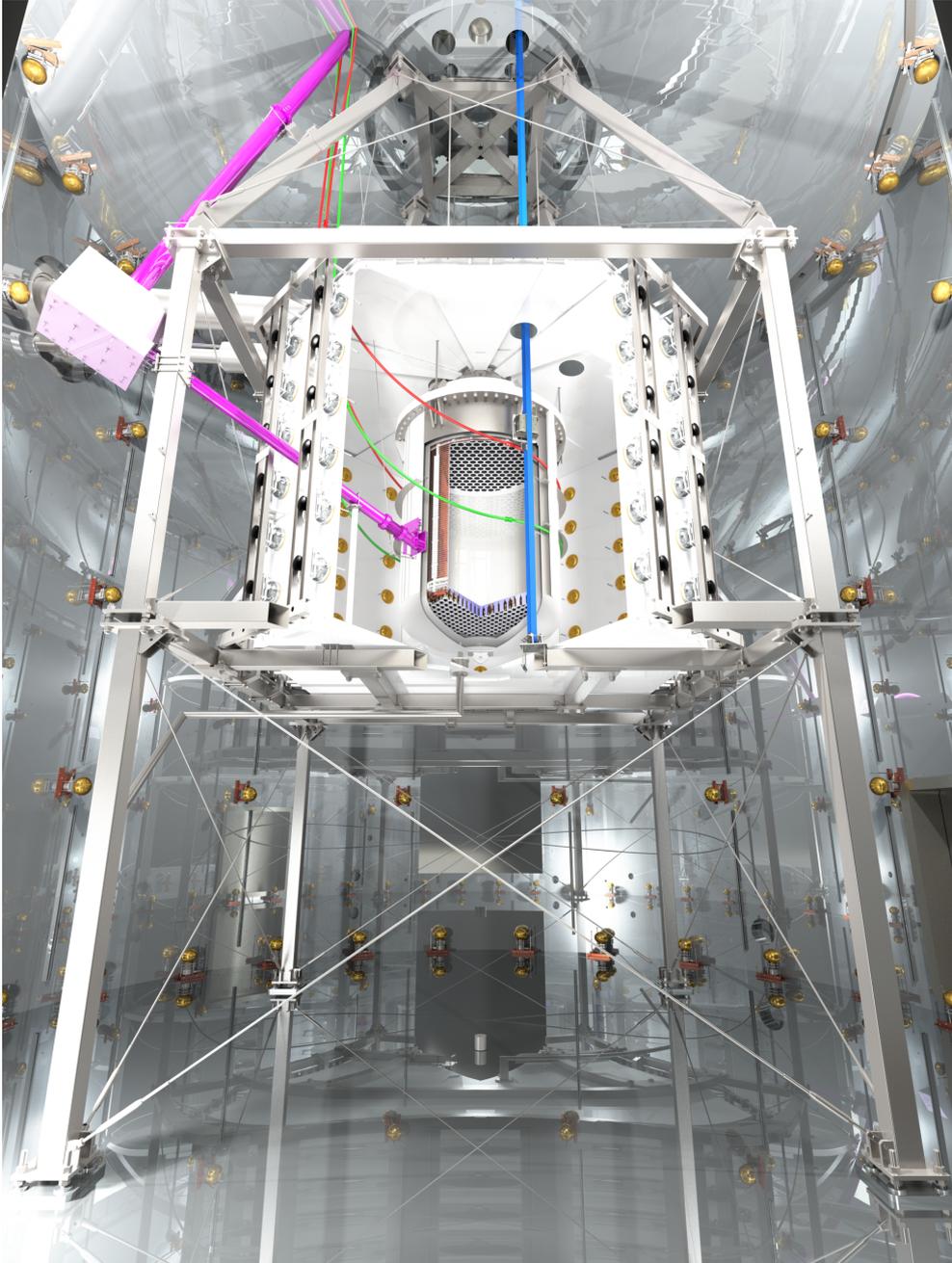


Figure 2.5.: A render of the XENONnT water tank (WT), forming the active muon (MV) and neutron veto (NV) detectors. The TPC inside the cryostat is visible in the center, while NV and MV are nested around. The NV is optically separated by ePFTE sheets arranged around the TPC cryostat. The red and green tubes, called the top and bottom u-tubes due to their rough shape like the letter ‘U’, are used in AmBe neutron calibrations for both TPC and NV. By inserting the source capsule into the tubes, it can be brought close to the TPC at various positions. Figure credit XENON Collaboration.

and ^{214}Pb . While the former is present in xenon due to its production via distillation from air, the latter is a daughter in the decay chain of ^{222}Rn that emanates from the detector materials. Both isotopes can be removed from the xenon volume via cryogenic distillation, based on the differences in vapor pressure between xenon and krypton, and xenon and radon, respectively. Inherited from XENON1T, a Kr distillation column is used for the removal of ^{85}Kr before the start of science data taking [16]. A new addition to XENONnT is the continuous cryogenic radon distillation system [17], which actively removes ^{222}Rn from the xenon volume during active operation of the detector, thus reducing the equilibrium between radon emanation and decay in the xenon volume. Additionally, the ER background caused by the emanation of radon from detector materials is mitigated by the pre-selection of materials with low emanation rates, determined via a dedicated screening campaign [18].

Apart from radioactive contaminants, electronegative trace impurities like water and oxygen can also have a negative impact on LXe TPCs, as drifting electrons attach to them and are lost, reducing the electron drift lifetime and thus the S2 signal size. Impurities are continuously removed from xenon in a gas and a liquid purification system that are based on hot getters. The liquid purification system is another upgrade for XENONnT, using a novel cryogenic purification technique [19] and achieving the high flow rates needed for the effective purification of the increased xenon volume via cryogenic liquid pumps. With the new purification system, an electron lifetime of more than 15 ms is achieved in SR0, sufficiently high for the successful operation of the TPC despite the reduced drift field strength leading to a comparatively large maximum drift time of ~ 2.2 ms.

Chapter 3.

The Eyes of XENONnT: Photomultiplier tests, selection, and installation

The expectedly rather small energy depositions from potential dark matter interactions in a liquid xenon target necessitate the capability of detecting light signals down to the single photon level. This makes photomultiplier tubes (PMTs) a prime candidate for the choice of photosensor. Despite them being a rather old technology – having been around since the 1930ies – PMTs are still the preferred choice for this kind of application in many particle and astro-particle physics experiments. For the XENONnT TPC, the PMT model of choice is the Hamamatsu R11410-21 depicted in [Figure 3.1](#), the same model as in XENON1T. Specifically developed for use in liquid xenon experiments, it and features a high quantum efficiency (QE) in the VUV wavelength range and a low dark count rate. Different subtypes of this model are used in all other major xenon TPC based dark matter searches as well, namely the LZ [\[20\]](#) and PandaX-4T [\[21\]](#) experiments. The ‘-21’ subtype used in the XENON experiments has a specifically low intrinsic radioactivity due to a dedicated development campaign between the XENON Collaboration and the producer [\[22\]](#).

With the new TPC for XENONnT being significantly larger than its predecessor, about twice the number of PMTs is needed to fill the PMT arrays, going from 248 tubes in XENON1T to 494 in XENONnT. While a subset of PMTs from XENON1T are reused, 368 new PMTs of the same type were acquired before the start of construction of the XENONnT detector. However, despite the many advantages of the R11410 PMTs, they are now known to sometimes suffer from issues that can affect the long-term operability of single tubes. A significant fraction of the tubes in XENON1T had to be turned off permanently before the end of the experiment due to these issues. Prompted by this experience, the newly acquired PMTs were submitted to a much more rigorous testing campaign than for XENON1T, including both upgraded and new tests for all PMTs under realistic operational conditions immersed in liquid and gaseous xenon.

In this chapter, we will give a short introduction to the working principle of PMTs in [section 3.1](#), and introduce the characteristics of the R11410 model that are relevant for their use in the XENON experiment. Then we will focus on the results of the XENONnT PMT testing campaign. These results are published in [\[23\]](#) of which the author of this work is a corresponding author, and will be summarized here in [section 3.2](#). Beyond a simple yes-or-no decision for each PMT whether it qualifies for use in the TPC, the results of the testing campaign also form the foundation for a quantitative categorization system of all available tubes. The system was developed in the context of this work, taking into account all PMT characteristics that may lead to unsatisfactory performance in the long-term. Based on this categorization, the best tubes for the XENONnT PMT arrays were selected. Similarly, the decision on the placement of the selected PMTs within the arrays was guided by this system, in an effort to optimize



Figure 3.1.: R11410-21 photomultiplier tubes (PMTs), used in the XENON1T and XENONnT experiments. Photo taken by Christian Föhr.

detector performance even in the case of potential PMT failures. The categorization, selection and distribution procedure for the XENONnT PMT arrays is described in [section 3.3](#), followed by a summary of the subsequent installation of the tubes into the TPC during the assembly of the XENONnT detector.

3.1. The XENONnT Photomultiplier Tubes

Photomultipliers are basically electron tubes that convert incident light into electrons, which are amplified into a detectable electrical current. They usually consist of an evacuated tube either made completely of glass or with a glass window at one end. A thin layer of a photosensitive material is deposited on the inner surface of the window, acting as a photocathode. When an incoming photon hits the photocathode, it can be absorbed and an electron is emitted via the photoelectric effect. This photoelectron is accelerated towards the dynode chain, a metal structure consisting of stages of electrodes, via an applied high voltage. Each dynode is held at a higher potential than the previous one via a voltage divider circuit. The dynodes are coated with a secondary emissive material, such that for each incident electron more than one secondary electrons are emitted, resulting in an electron avalanche that can finally be read out at the last dynode.

Photomultipliers exist in a large variety of types, sizes, and models, that feature different properties and characteristics. A detailed description can be found e.g. in [24]. For the choice of PMT model for the XENON experiments, a very specific set of requirements needs to be fulfilled:

- Spectral response: The photocathode needs to be sensitive to the wavelength of the VUV xenon scintillation photons of 175 nm [9]. Similarly, the window material should be transmissive at that wavelength.
- Detection efficiency: Due to the small expected signal sizes of WIMP-nucleus scatters in xenon, resulting in only a few produced scintillation photons, the detection efficiency of the tubes needs to be as high as possible. This is mainly governed by the quantum efficiency (QE), i.e. the efficiency with which a single incident photon is converted into (at least) one photoelectron in the photocathode.
- Dark count rates: In a detector using many PMTs, the random pile-up of dark pulses

from different tubes, generally defined as pulses not originating from an incoming photon but e.g. from the thermal emission of electrons from the photocathode, can produce fake signals. The dark count rates of the single tubes must therefore be kept as low as possible.

- Radiopurity: Traces of radioactive isotopes in the PMT materials need to be as low as possible to reduce radiogenic backgrounds in the experiment.
- Operation: The stable, long-term operability of the PMTs in the cryogenic environment inside a dual-phase xenon TPC needs to be ensured, with a LXe temperature of about -100°C and pressure of about 2 bar.

Fulfilling the above requirements, the PMT model of choice in the recent generations of XENON experiments is the Hamamatsu R11410-21 model, depicted in [Figure 3.1](#). The R11410 PMT type was specifically developed for liquid xenon experiments, and is rated for operation in the temperature range from -100 to $+50^{\circ}\text{C}$ and pressures of up to 4 bar. It is used in several past and currently running xenon-based detectors, in particular it was already used in XENON1T as well as in other current xenon TPCs searching for WIMP dark matter, using slightly different models of that same type, namely the LUX-ZEPLIN (LZ) [\[20\]](#) experiment and the PandaX-4T [\[21\]](#) experiment.

The design and working principle of the R11410 PMT is illustrated in [Figure 3.2](#). The main structure of the tube consists of a metal body made from a cobalt-free Kovar alloy and a round and flat, 3-inch diameter window made from UV-transparent synthetic fused silica. A thin layer of the photocathode material is applied to the inner surface of the window via vapor deposition. The photocathode material is a so-called special low-temperature bialkali compound, consisting of an alkali antimonide with a composition of (presumably) Sb-K-Cs [\[26\]](#). Due to the small work function energy of the material, the photocathode has a high QE down to the VUV photon range of typically 30 - 40 %. A produced photoelectron is accelerated towards the first dynode by an electric potential between cathode and first dynode. A focussing electrode (also visible in [Figure 3.1](#) as a spiderweb-like structure) held at the same potential as the first dynode provides the correct shape of the electric field, guiding the photoelectron through the geometry of the tube (see also [Figure 4.1](#)). The following amplification chain consists of 12 dynodes arranged in the box and linear-focussed type. Each dynode is connected to one of the pins extruding from the back of the tube, with a ceramic insulation stem acting as the feedthrough. The dynodes are held on different electric potential by attaching a so-called base, which is a voltage divider circuit, to the pins, applying a negative supply voltage bias to the photocathode with the anode on ground. The dynode chain is separated from the main body by an L-shaped piece of insulation material. Finally, a small strip of getter material of undisclosed composition, but that is called “cold getter” by the producer, is fixed to the inside of the body, in order

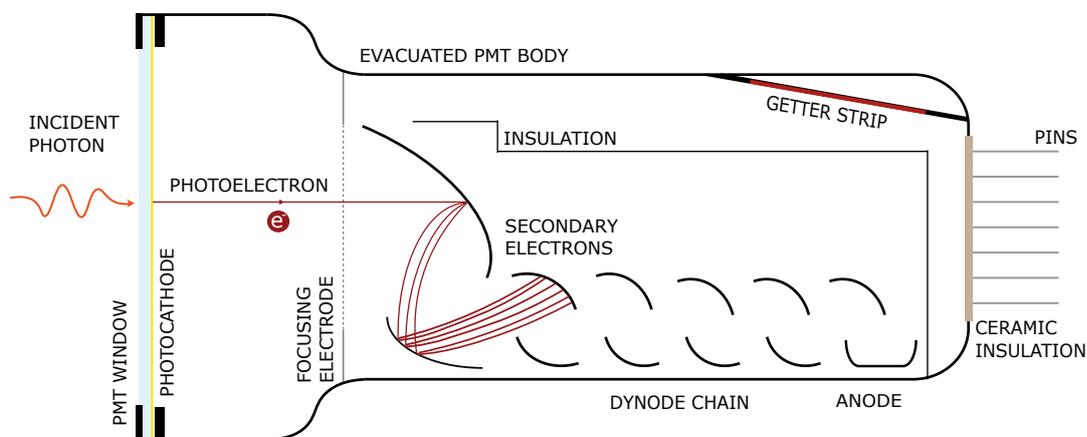


Figure 3.2.: Illustration of the internal structure of the Hamamatsu R11410-21 PMT and its working principle. Figure adapted from [\[25\]](#).

to remove residual gases from the tubes after evacuating and closing them in the production process.

The “-21” version of the R11410 PMT was specifically developed in cooperation between Hamamatsu and the XENON Collaboration towards satisfying the required low radioactivity levels of the experiment, reaching per-PMT activities of less than 13 mBq in ^{238}U and (0.4 ± 0.1) mBq in ^{228}Th [22].

3.1.1. Basic PMT properties

In the following we introduce basic PMT properties and characteristics, with a special focus on the properties that are relevant for the PMTs used in the XENONnT experiment.

Quantum and collection efficiency The quantum efficiency (QE) of a PMT is defined as the ratio of the number of photoelectrons emitted by the photocathode to the number of photons incident on the photocathode, representing the probability of a photon being converted into at least one photoelectron. In the R11410 PMT, double-photoelectron emission (DPE) can occur, an effect where two electrons are emitted from the cathode for a single incident photon. The underlying process behind this behavior is not fully understood, however it is assumed that the short wavelength of xenon VUV photons of 175 nm (corresponding to an energy of ~ 7.1 eV) is high enough in energy that the initially excited electron has sufficient excess energy to excite a second electron during the diffusion process towards the surface of the photocathode material. The collection efficiency (CE), in turn, quantifies the probability of an emitted photoelectron reaching the first dynode, and depends on the geometry and electric field configuration inside the tube. For the R11410 PMT, the QE at 175 nm and the CE are typically about 32.5 % and 90 % respectively, as stated by the producer and confirmed in several studies

Amplification gain The gain of a PMT is the amplification factor of the dynode chain, corresponding to the (average) number of electrons in the final electron avalanche at the last dynode per incoming primary photoelectron. It can be expressed as the product of the individual amplification factors, or secondary emission ratio, δ_i of each dynode stage i ,

$$G = \prod_{i=1}^n \delta_i \quad (3.1)$$

with the number of dynodes n . The secondary emission ratio itself depends on the kinetic energy of the incident electrons and thus on the potential difference U_i between the dynode i and the previous stage:

$$\delta_i = c \cdot U_i^k. \quad (3.2)$$

Here c and k are constants, where k depends on the dynode coating material and the geometry of the dynode chain. With the potential differences U_i being fixed by the choice of resistivities in the voltage divider circuit, the gain of a PMT becomes a power law function of the applied supply voltage. The R11410 PMT has $n = 12$ dynode stages, with k measured to be about 0.67 [23, 27], and a typical gain of about 5×10^6 .

Dark counts Dark counts in a PMT are defined as signals pulses that don't result from photons impinging on the photocathode (note that this definition excludes 'electronic noise' from being dark counts). At room temperature, by far the largest source of dark counts is the thermal emission of electrons from the photocathode. The thermal emission is strongly suppressed at cryogenic temperatures, where other sources of dark counts are dominating, e.g. ionizing particles impacting the cathode or other PMT parts and releasing electrons, such as muons from cosmic rays or external and internal natural radioactivity, as well as electric field emissions of electrons inside the tubes.

Timing performance For the characterization of the timing performance of the PMTs, the transit time spread (TTS), also called jitter, is the parameter used as figure of merit. It is defined as the full-width-at-half-maximum (FWHM) of the transit time distribution, where the transit time is the time between the arrival of the photon on the cathode and the arrival of the electron avalanche at the last dynode. It usually depends on path length differences of the photoelectron when traveling to the first dynode, and therefore depends on the position on the PMT window that the photon arrives at, as well as on the geometry of the tube itself and of the dynode chain. The TTS of the R11410 PMT is typically 9 ns [28].

Spurious pulses Signal pulses that don't arrive at the expected time relative to the photon arrival time are called spurious pulses. They can be either caused by an unusual arrival time of the initial photoelectron pulse, that are called early or late pulses, or be a secondary pulse that is caused by the initial pulse, so-called afterpulses. Early pulses likely result from the incident photon passing through the photocathode and only being absorbed by the first dynode, such that the photoelectron released from there arrives early due to shorter total path it travels. Having missed the first and most important amplification stage at the first dynode, early pulses are also expected to be under-amplified. Late pulses can result from an initial photoelectron elastically backscattering from the first dynode without releasing secondary electrons. Due to the electric field it will circle back and hit the first dynode again, where it may successfully start the amplification process. Due to this extra loop in its path, the resulting pulse will be slightly delayed, and may also be under-amplified due to reduced acceleration before hitting the first dynode. In contrast to the former processes, afterpulses are created by a primary signal pulse via ion feedback. When traversing the region between photocathode and first dynode, an initial photoelectron from the primary pulse can ionize residual gas atoms or molecules within the tubes. Since they are positively charged, the created ions will accelerate back towards the photocathode, where on impact several electrons are released. The pulses resulting from these electrons are thus delayed by the time the ions need to travel back to the photocathode.

3.2. Improved PMT qualification tests for XENONnT

Ensuring the long-term operational stability of the 368 new tubes purchased for XENONnT is paramount to the success of the experiment. Similar to XENON1T, the new tubes were therefore submitted to a PMT testing campaign, aimed at characterizing both the tubes' general properties and catching potential performance issues in the long-term.

The R11410 PMT model has suffered from two main performance issues in the past. The first issue is light emission, where small amounts of light are produced and emitted from the PMT itself. The second one is vacuum degradation, where small amounts of gas enters the PMTs through tiny leaks. The latter in particular can be very problematic, causing excessive afterpulsing induced by the ionization of the gas atoms or molecules in the tube, which in the worst case can render a PMT inoperable.

The problem of light emission was first identified already before the use of this PMT model in the XENON1T detector [29, 30]. Therefore, great effort was made in the initial PMT testing campaign for XENON1T [31] to identify tubes affected by this issue in advance before installing them inside the detector. In contrast, the problem of vacuum degradation was unfortunately only discovered during the operation of the PMTs in the XENON1T TPC. By the end of the experiment, after more than 2 years of continuous operation in xenon, roughly one third of the 248 PMTs in the detector showed afterpulses induced by xenon ions inside the tubes, indicating that they had developed leaks. Of those, 34 (about 14%) had to be turned off before the end of the experiment due to excessive afterpulsing causing frequent bursts of light emission called flashes, accompanied by high-voltage trips, rendering the tubes inoperable.

As a result of this experience in XENON1T, improvements have been made to the XENONnT testing campaign compared to the previous one, specifically targeting the observed issues in

order to significantly reduce the failure rate of the PMTs in the new detector. For the tests in XENON1T, only a small subset of the tubes were tested for leaks at the actual operational conditions in a TPC, immersed in liquid xenon or gaseous xenon at a pressure of ~ 2 bar. In the new XENONnT testing campaign, we therefore ensured that every PMT is tested for at least two weeks in LXe and GXe. Another new effect found during the operation of the tubes in XENON1T is intermittent light emission, a stronger form of light emission that is intermittently turning on and off. It was often observed to be triggered during times of high event rates in the TPC, e.g. during calibrations. A new high-rate stress test targeting this effect was added to the testing campaign in order to identify affected tubes. These improvements to the testing procedure, as well as some minor alterations to the PMT design and production process by the manufacturer, are expected to significantly reduce the failure rate of the PMTs in XENONnT compared to XENON1T.

The XENONnT PMT testing was a joint effort between three institutions within the XENON Collaboration, namely Zurich University, Stockholm University and MPIK in Heidelberg. The results of the tests performed at the MPIK facilities for a subset of the tubes were already reported in [27]. In the following, we will focus on summarizing the overall testing results for all new tubes for the XENONnT experiment, as published in [23].

3.2.1. Overview of facilities and testing procedure

The testing facilities at the three institutions serve different purposes in a complementary way. At the MPIK testing facility, two setups are available. The first consists of a room-sized, light-tight steel box acting as a Faraday cage, where low-noise precision measurements of PMT characteristics can be performed at ambient temperature. The second setup consists of a dewar with a light-tight rubber-sealed flange lid, housing a holding structure that fits 12 tubes at a time. The PMTs face each other window to window, enabling the measurement of light emission from a tube with the opposite tube. The dewar can be filled with nitrogen or argon gas at atmospheric pressure, that is cooled down to cryogenic temperatures via a liquid nitrogen cooling coil. For details on the two setups and the measurement procedures, see [27] and references therein. The setups at Zurich and Stockholm University are used for tests where the tubes are operated in LXe and cold GXe, at realistic operational conditions as in a xenon TPC. While the Stockholm setup was newly constructed for the XENONnT tests, the Zurich setup was already used in the XENON1T PMT testing campaign [31], but was upgraded for the XENONnT tests to be able to test more PMTs simultaneously. Both setups consist of vacuum-insulated cryostats that house holding structures for the PMTs similar as in the MPIK dewar setup, with the tubes facing each other in pairs. The cryostats are filled with xenon gas at ~ 2 bar pressure, that is cooled down and liquified via pulse-tube refrigerators (PTRs). Detailed descriptions of these setups and measurement procedures are reported in [23, 25, 32]. All setups are equipped with optical fibers connected to pulsed LEDs, that allow for the illumination of the tubes as needed for the different measurements.

All PMTs for XENONnT were tested at least twice in either of the two xenon setups – 105 pieces in Zurich, and 260 pieces in Stockholm¹. In one such testing cycle, the tubes are first immersed in LXe for one week, and then in GXe for another week. During this time they are regularly tested for both light emission and vacuum degradation through xenon leakage via afterpulse measurements. A subset of 161 PMTs from the tubes that were tested in Stockholm was tested first at the MPIK facility in addition to the xenon tests. Here, in the standard procedure, first the basic characteristics are measured in the room-temperature setup, namely the amplification gain, single-photon resolution, dark count rates, afterpulse spectra, and transit time characteristics. Then the tubes are transferred to the dewar setup, immersed in argon gas and slowly cooled down to LXe temperature. During the stable cold period, the tests for light

¹The attentive reader may have noticed that this does not sum up to the total number of acquired PMTs of 368 - this is because three tubes that failed already in the first tests at MPIK were consequently not tested in the xenon setups.

emission are performed. The check for vacuum degradation is done after the cooldown cycle by measuring the tubes' afterpulse spectra again in the room-temperature setup. Same as for the xenon tests, each tube undergoes the whole procedure at least twice.

3.2.2. Basic characteristics tests

Similar to the XENON1T PMT testing campaign, the basic characteristics of the new PMTs are determined to ensure their proper functioning within specifications by the producer and needs of the experiment. We measure the PMT gains and related quantities, namely the peak-to-valley ratio and the single photoelectron (SPE) resolution, as well as the transit time, its spread, and the dark count rates. The results of these measurements are summarized in [Table 3.1](#), and compared to the results from the XENON1T PMT testing campaign [31]. While quantum efficiency and transit time spread are very similar between the two sets of PMTs, the amplification gain is considerably higher for the new tubes, along with small improvements in SPE resolution and peak-to-valley ratio. The dark count rates at room temperature (warm) are about an order of magnitude lower for the new tubes compared to the XENON1T tubes, but are comparable at cryogenic temperatures (cold). None of the new tubes were excluded from use in the XENONnT TPC due to their performance in these measurements, since all conformed to the specifications of the producer and their satisfactory working behavior was confirmed.

Characteristic	XENONnT tests	XENON1T tests	Unit
Quantum efficiency (QE)	34.0 ± 2.8	34.5 ± 2.8	%
Gain at 1.5 kV	8.4 ± 2.3	5.4 ± 2.1	10^6
HV for gain of 5×10^6	1.41 ± 0.04	–	kV
SPE resolution (at 5×10^6 gain)	25.1 ± 1.5	27.7 ± 2.1	%
Peak-to-valley ratio (at 5×10^6 gain)	4.3 ± 0.4	3.9 ± 0.4	1
Transit time spread (TTS)	9.2 ± 0.5	9.1 ± 1.3	ns
Dark count rate (warm)	~ 100	~ 1000	Hz
Dark count rate (cold)	~ 40	~ 40	Hz

Table 3.1.: Results of the PMT characterization measurements for the new XENONnT PMTs [23], compared to the results from XENON1T [31]. Values and errors correspond to the mean and standard deviation of the respective distribution, except for SPE resolution and Peak-to-valley ratio where the value is reported as the median. SPE resolution, peak-to-valley ratio, transit time spread and dark count rates were measured for the MPIK subset of PMTs only [27]. None of the new PMTs were excluded due to their performance in any of these measurements.

Quantum efficiency

The QE values of all tubes are provided by the producer, measured at 25°C at a photon wavelength of 175 nm. The distribution of QE values for the new tubes are shown in [Figure 3.3](#) in blue. The darker shaded histogram corresponds to the PMTs that passed all tests and thus qualified for the use in the XENONnT TPC, while the lighter blue histogram stacked on top is for the tubes that failed any of the tests and were thus returned to the producer. The QE distribution has a lower cut-off value at 28%, the minimum value agreed upon with the producer. The mean QE for all new tubes, both including and excluding the failed tubes, is 34.0% with a standard deviation of 2.8%. The distribution of QE values for the 248 XENON1T PMTs is shown as well in red, with a very similar mean and standard deviation of 34.5% and

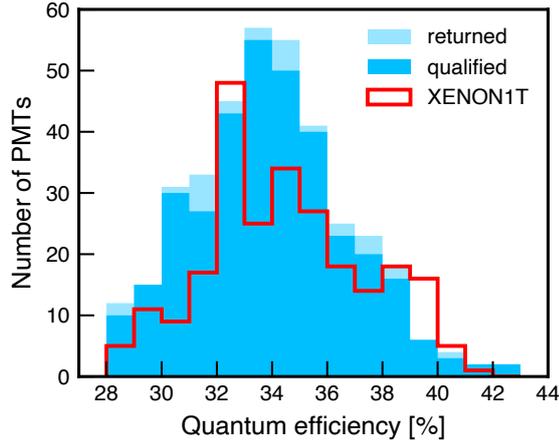


Figure 3.3.: Quantum efficiency distribution of the new R11410-21 PMTs tested in the XENONnT testing campaign (blue). The darker shaded histogram corresponds to the tubes that passed all tests and thus qualified for use in the TPC, while the lighter shaded histogram stacked on top shows the failed and returned tubes. The QE distribution of the PMTs from the XENON1T PTC [31] is shown for comparison (red). Figure published in [23].

2.8%. Note that the QE of this PMT type has been shown to increase at cryogenic temperature by 10 - 15% [33] (relative), caused by a decrease in electrons scattering on lattice phonons while diffusing through the photocathode material towards the surface, making them more likely to be emitted as photoelectrons. The DPE probability of the R11410 PMT has been measured to be around 22% [34, 35], and is included in the QE values measured by the producer. Correcting for both the QE increase at cold temperatures and the DPE effect, the mean QE of the new XENONnT PMTs becomes 30.7%.

Amplification gains

In the full procedure of the XENONnT PMT testing campaign, the gains of all PMTs are measured several times and in different conditions, depending on the facility and setup, for a range of supply voltages between 1.2 and 1.7 kV. The measurement method is the same in all testing facilities: The tubes are illuminated with the pulsed LEDs at single photon intensity, and the charge spectra of the resulting PMT pulses are acquired. These charge spectra feature a noise peak around zero charge, a peak corresponding to the single-photoelectron (SPE) charge, and increasingly smaller contributions at integer multiples of the SPE charge corresponding to multiple initial photoelectrons. Two different methods are used to extract the gain value from the charge spectra, depending on the setup, either by fitting a model to the expected peaks in the spectrum [27, 31] or using a model-independent method exploiting the statistical properties of the PMT response [36]. The gains of all tested tubes, measured in liquid xenon at a bias voltage of 1.5 kV, are shown in Figure 3.4 (left) in blue. The average gain of the set of qualified tubes (darker shaded) is 8.4×10^6 , with a standard deviation of 2.3×10^6 . The gains of the new PMTs are considerably higher than for the XENON1T PMTs, shown in red, which have an average gain of 5.4×10^6 and a standard deviation of 2.1×10^6 . This allows for the operation of the new tubes at lower supply voltages, which is beneficial for the lifetime of the tubes and reduces the amount of micro light emission, which strongly depends on the applied bias voltage (see subsection 3.2.3). Figure 3.4 (right) shows the distribution of supply voltages needed for a gain of 5×10^6 , determined from fitting the power-law dependence of the gain on the supply voltage (see Equation 3.1, Equation 3.2).

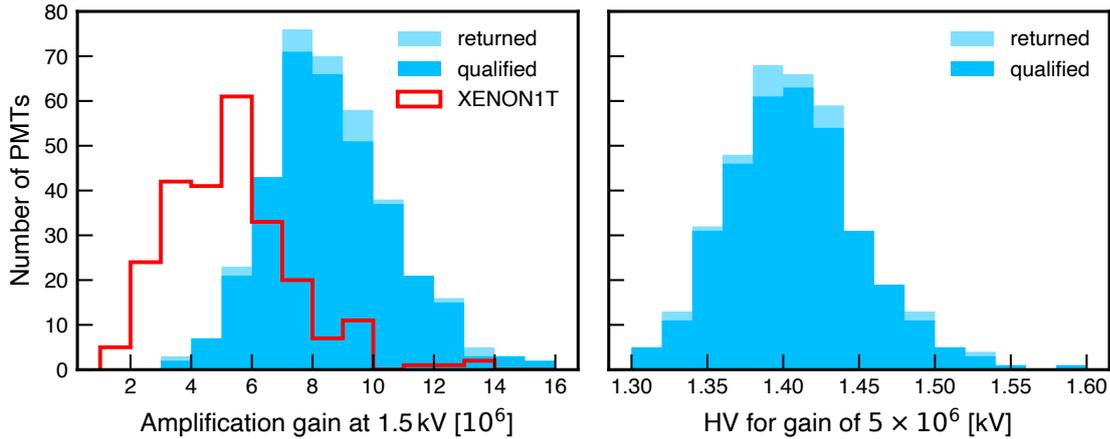


Figure 3.4.: (Left) Amplification gains of all PMTs tested in the XENONnT PMT testing campaign (blue) measured in liquid xenon at about -100°C and a supply voltage of 1.5 kV , averaging at 8.4×10^6 . Shown in red is the gain distribution as measured in the XENON1T testing campaign, showing a considerably lower gain values. (Right) Distribution of HV values needed for reach a gain of 5×10^6 . Figure published in [23].

3.2.3. Light emission

The effect of light emission (LE) of the R11410 PMTs is by now a known phenomenon that has been first observed in [29, 30], as well as in multiple other studies and experiments using these tubes [31, 37–39]. In general this term refers to the emission of photons from the interior of the tubes. Once created, these photons can hit the photocathode from the inside, and are therefore mainly detected by the light-emitting PMT itself. They can also traverse the window and consequently be detected by other PMTs in the vicinity. In a low background and low energy experiment where the expected S1 signal size goes down to the few-photon level, these emitted photons can form background events when piling up with other single-photon signals, forming ‘fake’ S1s. These can in turn coincide with small S2 signals from single electron backgrounds in the TPC, forming ‘accidental coincidence’ events that can mimic real particle interactions. Reducing the amount of light emission from the PMTs is therefore important for the reduction of these background events.

We distinguish between three different categories of light emission that have been observed in the R11410 PMTs. The first category is the so-called micro light emission, a continuous emission of light on the single photon level at low intensities. It is present in all PMTs and varies strongly with the applied bias voltage of the tubes, increasing exponentially at voltages higher than about 1.5 kV . The origin of this effect is not fully understood, but is hypothesized to be caused by field emissions from regions of large electric field inside the PMT, mainly around the dynode chain. According to the producer, micro light emission has been observed to originate specifically from the region around the last dynode stages [40]. In response to the large numbers of PMTs showing strong LE found in the XENON1T PMT testing campaign, the interior design of the tubes was slightly altered for the manufacturing cycle of the new PMTs, blocking the optical path of photons between the dynode chain and the PMT window (mainly by closing the gap between the focussing electrode and the inner wall of the main body, see Figure 3.2). This should prevent the LE photons’ detection by the PMT itself as well as by other tubes, reducing the amount of light coming from the tubes.

The second category of LE is intermittent light emission. In contrast to micro light emission it is not continuous but starts suddenly and may last for minutes, hours or even days, eventually stopping again without any apparent reason. It is present in only few PMTs, and generally has a higher intensity than micro light emission. This behavior was first observed during operation of the PMTs in the XENON1T TPC, often (but not always) triggered during periods of large

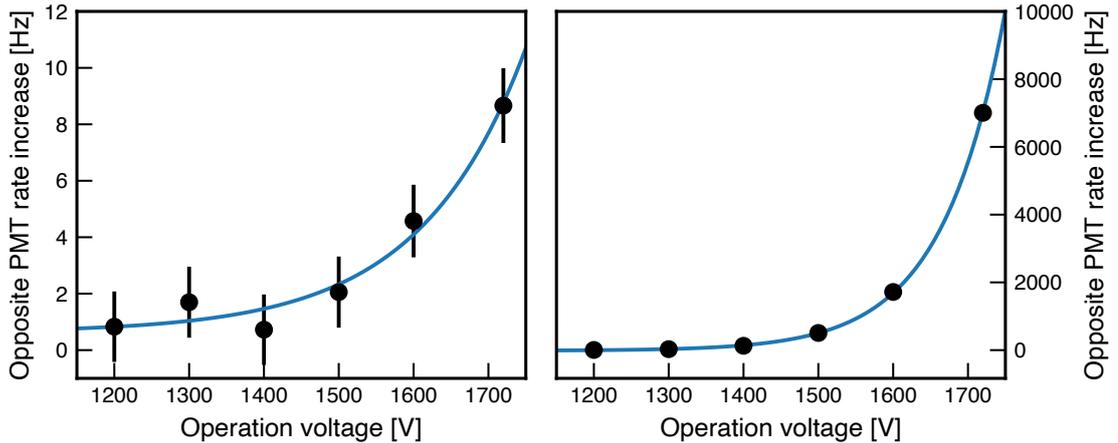


Figure 3.5.: Examples for the test measurements for micro light emission in the XENONnT PMT testing campaign, showing the increase in count rate of an opposite PMT in dependence of the applied HV of the tested PMT. The data points are well described by an exponential function (blue lines). (Left) Micro light emission behavior of a typical, good PMT that passes the test. (Right) Strong micro light emission of a PMT that consequently failed the test. Figure published in [23].

event rates in the detector such as calibrations, or by extremely large events producing large amounts of light in the detector, such as muons.

The third category are so-called flashes, which are sudden but very short bursts of light emission at very high intensity. They are thought to originate from charge build-up on the insulating material around the dynode chain, which may be released in a sudden discharge producing the flash of light. Flashes are observed mainly in PMTs that suffer from vacuum degradation, becoming more frequent and intense as the vacuum in the tubes keeps deteriorating, until the PMT eventually fails completely. Here, too, the producer made changes to the design of the tubes in order to reduce the charge build-up and thus the occurrence of flashes, by exchanging the material of the dynode insulation [40] (labelled L-shaped insulation in Figure 3.2), presumably to one with increased surface conductivity.

In our tests, following the same procedure as for XENON1T PMT tests [31], we measure the severity of micro light emission via the increase in the count rate of a PMT opposite the tested one. The tested PMTs' HV is scanned in a range starting from 1.2 kV, where the tubes show almost no micro light emission, up to the maximum operational voltage of 1.7 kV. Two example measurements from the MPIK cryogenic dewar setup are shown in Figure 3.5, showing the rate increase of an opposite PMT in dependence of the tested PMT's HV. On the left we see the measurement of a tube with the usual micro light emission behavior, where the count rate increase in the opposite PMT becomes apparent for voltages above 1.5 kV, at low intensities with only a few to $\mathcal{O}(10)$ Hz. An example for a tube with very large light emission is shown on the right, where the rate increase is extreme, reaching values of a few hundred Hz already at 1.5 kV, and several kHz at the highest voltages.

In order to detect tubes suffering from intermittent light emission, we added a high illumination stress test to the standard procedure. During the cryogenic testing cycles, the tubes are illuminated with the LEDs at very high intensities and rates, mimicking the conditions during calibrations or large events in a TPC. The illumination is performed for several hours, and the tubes are monitored for light emission during and after the illumination. In cases where a sudden increase in the count rate of one PMT and the opposite PMT is observed, the cause of light emission was confirmed by turning both PMTs off in turn. If the rate increase is caused by light emission, it will disappear when the tested PMT is turned off, and reappear when it is turned on again, while when the opposite PMT is turned off, the rate increase in the tested PMT will remain.

Out of the new 368 new PMTs tested in total, five tubes with micro light emission at the level of several kHz were identified. The high illumination stress test triggered intermittent light emission of a further five tubes. One additional tube showed intermittent light emission during the cryogenic testing cycles, but not during the high illumination stress test. These 11 PMTs were excluded from the use in the XENONnT TPC and returned to the producer for replacement.

3.2.4. Vacuum degradation tests using afterpulses

Vacuum degradation is a serious problem for PMTs, with their basic working principle requiring a high vacuum inside the tubes to function properly, typically between 10^{-6} and 10^{-8} mbar. Too much gas in the tubes will prevent the produced photoelectrons from reaching the first dynode by scattering and deflecting them off their path, making the detection process impossible. But even small amounts of gas inside the tubes can be problematic. These residual gas atoms or molecules can get ionized by the initial photoelectrons on their path from the photocathode to the first dynode, producing positive ions. These are then accelerated by the electric field back towards the photocathode, where on impact secondary electrons are released that can be amplified and detected in the same way as the initial photoelectrons. The produced signals are delayed with respect to the initial photoelectrons by the time it takes the ions to reach the cathode, and are therefore called afterpulses (APs) (see also [subsection 3.1.1](#)). The delay times of ion APs in the R11410 PMTs range from a few hundred nanoseconds to several microseconds. With the typical duration of S2 signals in large xenon TPCs at the few to ten microseconds level, APs often merge into the S2 signals, causing a bias in the reconstructed signal size that can be dependent on the event position due to different rates of APs in different tubes. This bias can also negatively affect the xy -position reconstruction algorithms and event selections that are based on the S2 signals' hit patterns.

Vacuum degradation in PMTs was a serious problem during the runtime of XENON1T, where almost 30% of the PMTs showed vacuum degradation via the penetration of xenon gas by the end of the experiment. This caused issues for the data analysis (e.g. in [\[41\]](#)), causing time-dependent effects because of the continuous leakage of xenon into the tubes and the consequently increasing rate of APs. About half of the leaky PMTs even had to be turned off due to a high frequency of flashes.

In the R11410 PMTs, the AP delay times are characteristic for the mass-to-charge ratio of the ions that created them, as will be described in [chapter 4](#) in more detail. This allows for the identification of the specific gas atoms or molecules that are present within a tube, and thus for the identification of vacuum degradation by measuring the AP delay time spectra and looking for an increase in APs from ions of the gas the tubes are surrounded with. In the XENON1T PMT testing campaign, all PMTs were tested for leaks in a cold nitrogen atmosphere, while only a small subset of 44 tubes could be tested while immersed in liquid xenon in the only cryogenic xenon setup available at the time (a smaller version of the current Zurich setup) [\[31\]](#). To ensure the successful identification of leaks in the new XENONnT PMT testing campaign presented here, all tubes were therefore tested in one of the two now available xenon setups. Since leaks have been observed to open especially during periods of exposure to thermal stress, such as during cool-down or warm-up periods in the setups, every tube undergoes at least two full cycles of the two-week testing procedure, where the first week is spent in LXe and the second in GXe. In addition, for the tubes that were tested at the MPIK facility the testing procedure was slightly altered, switching from testing in nitrogen gas to argon gas in the cryogenic dewar setup. This was prompted by the observation that nitrogen is removed over time by the getter material inside the PMTs, making it difficult to detect leaks in the tubes in a nitrogen atmosphere. The use of nitrogen for leak testing is likely to be also one of the reasons for the low success rate in identifying leaks in the XENON1T PMT testing campaign.

The AP measurements in the XENONnT testing campaign are performed by illuminating the PMTs at pulsed photon intensities of about $\mathcal{O}(10 - 100)$ detected photons per LED pulse, and

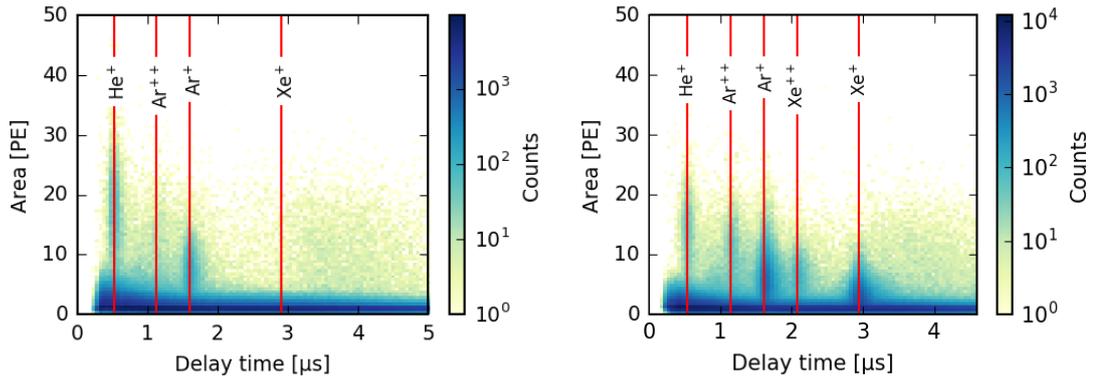


Figure 3.6.: Examples of AP delay time spectrum of two PMTs tested in the XENONnT PMT testing campaign, measured in the Stockholm setup. (Left) A good PMT, showing the characteristic peaks of the most common ions helium (He^+) and argon (Ar^+). (Right) A leaky PMT, showing APs from xenon ions (Xe^+) after being immersed in xenon. APs resulting from double-ionization are visible as well for both argon (Ar^{++}) and xenon (Xe^{++}). Figure from [23].

recording the resulting pulses within a time window of several microseconds following the LED pulse, and extracting the delay time of all recorded pulses as well as their pulse area in units of PE. Two examples of resulting AP delay time spectra are presented in Figure 3.6, showing the AP delay time and the pulse areas on the two axes. The expected delay times of the ions that are typically present are indicated by the red vertical lines. In the left panel, the AP spectrum of a good PMT is shown, with AP populations from helium (He^+) and argon (Ar^+) peaks visible. Helium is always present in the tubes, since it permeates into the tubes over time when they are stored in air due to its small atomic size, while argon is present in a significant fraction of the tubes already from production on. The right panel shows the afterpulse spectrum of a leaky PMT, where a strong xenon AP population is visible as well, indicating xenon penetration into the tube. Populations corresponding to APs from double-ionized xenon (Xe^{++}) and argon (Ar^{++}) are visible as well, however at lower intensities than the single-ionized ones. The severity of ion APs is then quantified by the AP rate, defined as the number of APs from a specific ion above an area threshold of 2.5 PE, normalized by the number of initial photoelectrons from the LED pulses. The Xe^+ afterpulse rate for the measurement shown on the right, for example, is 0.001 APs per PE in the LED pulses, which can be understood as a probability of 0.1 % for an afterpulse to occur after a single PE.

In the XENONnT testing campaign, the afterpulse spectra of all tubes were monitored in the xenon setups at Zürich and Stockholm, and in argon gas at the MPIK dewar setup. Out of the 368 new PMTs tested, we identified in total 15 tubes with vacuum degradation. Of those, five (one) showed an appearance and/or increase of xenon (argon) in the afterpulse spectra during operation in the respective media due to leaks. A further seven tubes showed extremely high overall rates of afterpulses, attributed to a severe nitrogen (N_2^+) contamination of the vacuum. Since in our tests, this was mainly observed either in the very first measurements or after transport between testing facilities, we attributed the nitrogen contamination to air leaks. However, during the commissioning phase of the XENONnT experiment, it was found that the nitrogen contamination is not external, leading to the conclusion that it originates instead from the interior of the PMTs, as will be discussed in detail in chapter 4. A further two tubes were completely inoperable, which we attribute to severe vacuum degradation, likely due to nitrogen contamination as well, bringing the total to 15.

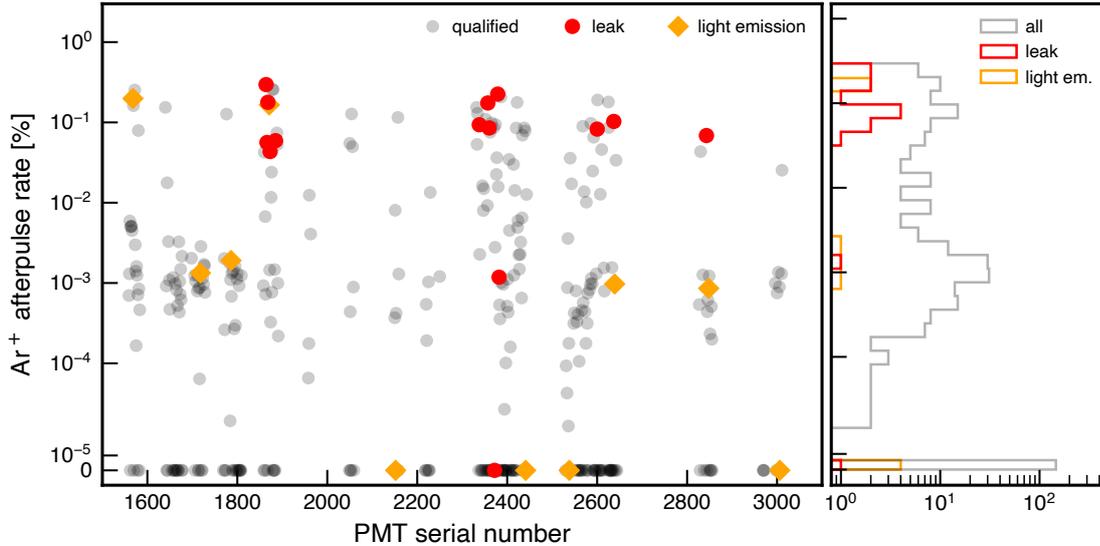


Figure 3.7.: Argon afterpulse rates as a measure for the initial argon amount present in the PMTs from production on, in dependence of the tubes’ serial number. Light gray circles show the PMTs that passed all tests and qualified for use in XENONnT. 60% of the tested PMTs have measurable amounts of argon contamination. Red circles indicate PMTs that later showed vacuum degradation in our tests (either Xe (Ar) leakage or nitrogen contamination), showing a correlation both with the presence and the amount of argon in the tubes. PMTs that showed light emission in our tests are indicated by the golden diamonds, following the general distribution of all tubes and thus not exhibiting a correlation with argon content. The panel on the right displays the respective histograms of the tubes onto the y-axis, showing how many PMTs fall in each bin of argon afterpulse rate. Figure published in [23]

Correlation between argon amount and leakiness

As already observed during the operation of XENON1T [38], we identify a correlation between the initial amount of argon in the tubes from production on, and the development of leaks later on. This is illustrated in Figure 3.7, where the argon afterpulse rate measured in the tubes in the first afterpulse measurement during the XENONnT testing campaign, i.e. before they are exposed to Ar gas in our tests, is shown in dependence of the serial number of the tubes. Gray circles show the tubes that passed our tests, with 60% showing a detectable amount of argon with an afterpulse rate above about 10^{-5} %. The red circles indicate the tubes that failed our tests due to vacuum degradation. Twelve of these show very large argon afterpulse rates above 10^{-2} %, corresponding to a 17% failure rate for tubes with argon afterpulse rates above that threshold. PMTs with argon afterpulse rates below 10^{-2} % show a failure rate of only 1%. The reason for this correlation is not completely understood, but since argon is already present in the tubes from production on, it is likely to be related to the production process itself. We also see that the tubes that failed due to vacuum degradation are clustered in specific ranges of serial numbers, further indicating that the problem may be related to certain production batches. For the final PMT selection and distribution in the arrays, we consider both the argon afterpulse rate and the affiliation of tubes to these ‘suspicious’ production batches as measures for PMT quality (see subsection 3.3.1), indicating a higher risk of failure.

The PMTs that failed due to light emission are indicated in the figure by orange diamonds. Here no correlation is observed, with a failure rate of 3% for tubes with argon afterpulse rates both above and below 10^{-2} %.

3.2.5. Summary of qualification test results

Of the 368 new PMTs tested, 26 tubes in total failed our tests and were excluded from the use in the XENONnT TPC, corresponding to a failure rate of 7%. While 15 tubes failed due to vacuum degradation, 11 tubes failed due to light emission. None of the tubes were excluded based on their performance in the basic characteristics tests, i.e. QE, amplification gain and SPE resolution, or dark count rates. The new tubes show very similar values as the XENON1T tubes regarding QE, SPE resolution, and dark count rates at cryogenic temperatures, but have significantly higher gains (see Table 3.1).

The results of the tests for problematic PMT behavior in the XENONnT PMT testing campaign are summarized in Table 3.2, and are compared to the previous results from the XENON1T PMT testing campaign as well as the PMT performance during operation in the XENON1T TPC. The performance of the new PMTs during the first science run of XENONnT (SR0) is shown in the table already as well, and will be discussed in detail in the following chapter 4.

Comparing the numbers of failures of the tubes due to light emission between the two testing campaigns, we see that the new PMTs show a much lower failure rate of only 3% than the ones tested for XENON1T, where 17% of the tubes showed strong micro light emission. We attribute this to the measures taken by the producer to reduce the amount of light emission from the tubes, that seem to have reduced the problem greatly and thus improved the general PMT quality. Note that the lower failure rate is in spite of the improved testing procedure for the new PMTs, that includes the new high-illumination stress test for intermittent light emission, a problem that was likely largely missed in the XENON1T testing campaign. In addition to the much smaller number of tubes with problems of strong light emission, the new set of PMTs can be operated at lower HV values due to their larger gains, reducing the impact of micro light emission from the good PMTs on the science data in XENONnT further.

Problem type	XENON1T		XENONnT	
	Testing	TPC	Testing	TPC (SR0)
Micro light emission	53 (16.5 %)	8 (3.2 %)	5 (1.4 %)	-
Intermittent light emission	-	-	6 (1.6 %)	1 (0.3 %)
LE total	53 (16.5 %)	8 (3.2 %)	11 (3.0 %)	1 (0.3 %)
Xe (Ar) leak	8 (18.2 %) ¹	71 (28.6 %)	6 (1.6 %)	5 (1.6 %)
N ₂ contamination	-	-	7 (1.9 %)	18 (5.7 %)
Non-operable	-	-	2 (0.5 %)	-
Total vacuum degradation	8 (18.2 %) ¹	71 (28.6 %)	15 (4.1 %)	20 (6.3 %)
Total with problem	73 (22.7 %)	79 (31.9 %)	26 (7.0 %)	21 (6.6 %)
Total excluded	see above	36 (14.5 %)	see above	11 (3.5 %)
Total tested or used	321	248	368	316

¹ Here, only 44 PMTs were tested in LXe [31]

Table 3.2.: Summary and comparison of encountered PMT problems leading to failure or exclusion of the tubes, in XENON1T and -nT for both the testing campaigns as well as during operation in the respective experiment. Note that the numbers for the XENONnT TPC are only for the new tubes that were tested in the XENONnT PMT testing campaign. Numbers for XENON1T from [11, 31, 42].

Regarding vacuum degradation, the new PMTs show great improvement as well, with only 4.1% of the tubes failing due to leaks in our tests, compared to the 18.2% of the small subset of tubes tested for XENON1T. Again, the much lower failure rate here is in spite of a much more rigorous testing procedure for the new PMTs, where all tubes were tested in liquid and gaseous xenon for in total 4 weeks. Also here we suspect the producer made changes in the production process in order to improve leak tightness. The most likely weak points for the development of leaks in the PMTs are assumed to be the glass-to-metal sealing between PMT window and body, and the ceramic stem feedthrough of the pins at the back (see [Figure 3.2](#)). Especially for the former, the producer has possibly taken measures to improve the sealing process or material – this is however speculative and not confirmed.

In XENON1T, already during the tests a high failure rate of about 18% due to leaks was discovered, however only a small subset of 44 tubes were tested in liquid xenon at the time. The other tubes were tested in cold gaseous nitrogen only, which was later discovered to be a suboptimal choice due the getter strip inside the tubes actively removing leaked-in nitrogen from the tubes. This is likely to be the reason why the problem was not discovered earlier. During operation in the XENON1T TPC, the total fraction of tubes showing xenon penetration by the end of the experiment was even higher at 28.6%. Note however that about half of those tubes could still be used in data analysis, since the amount of xenon afterpulses was low enough to not cause any behavioral issues of the tubes themselves, and the impact of increased afterpulse rates on data quality was considered in the analysis procedure [[11](#)].

The expected higher leak tightness of the new PMTs is also supported by the fact that compared to the XENON1T PMTs, much fewer of the new tubes show large amounts of argon in the first place. For the new tubes, 20% show argon afterpulse rates above 10^{-2} %, where the failure rate is highest, while for the old tubes that fraction is 50%.

In summary, due to a much more rigorous and improved testing regime for the new PMTs, together with the changes in PMT production implemented by the producer, we expect a greatly improved PMT performance in XENONnT. This expectation was confirmed in the first science run of XENONnT, where the PMTs showed very good performance, as will be discussed in the next [chapter 4](#).

3.3. The XENONnT PMT arrays

The XENONnT TPC operates with in total 494 PMTs that are distributed over two PMT arrays, one at the top and one on the bottom of the TPC. They hold 253 and 241 tubes, respectively. In each array, the PMTs are positioned in a hexagonal pattern (as can be seen in [Figure 3.9](#)) in order to maximize their packing density. This is beneficial for both position reconstruction resolution in the top array, and for the light collection efficiency in both arrays, but mainly in the bottom array. With this pattern, the fraction of the array surfaces that are covered by the photosensitive area of the PMTs is about 60%.

In order to fill the arrays with PMTs, available tubes come from three different subsets: Old tubes from the XENON1T detector, old tubes that were bought for XENON1T as spare tubes but never installed, and new tubes that were bought for XENONnT and tested as described in the previous section. Here, we describe a categorization system used for the selection of the tubes out of these subsets, as well as for the positioning of the selected individual tubes in the two arrays, in [subsection 3.3.1](#). After the selection and positioning decisions for the PMTs were made, the arrays were constructed and the PMTs installed prior to the full assembly of the XENONnT TPC at LNGS, as we will see in [subsection 3.3.2](#).

3.3.1. PMT selection and array distribution

PMT selection and penalty point system

After completing the XENONnT PMT testing campaign, a total of 539 good PMTs were available for the XENONnT TPC, 45 more than necessary. This number is made up of three subsets: 165 PMTs that were previously installed in the XENON1T TPC, 32 PMTs that were bought for XENON1T as spare tubes but never installed, and 342 new PMTs that passed the tests, as described in the previous sections. The subset of tubes from the XENON1T TPC consists of all tubes that did not show any sign of light emission or vacuum degradation via xenon penetration until the very last week of operation of XENON1T. They can therefore be considered as good tubes. On the other hand, they do belong the old production cycle of generally more failure-prone PMTs, and have been operated for about 2 years in XENON1T already, so it is unclear whether they can be considered generally of better or worse quality than the new PMTs. The subset of spare PMTs from XENON1T times had been tested previously within the XENON1T PMT testing campaign, using the previous testing procedures that were not as rigorous. They are thus considered generally of worse quality.

In order to quantify the quality of the PMTs in a more systematic way, a penalty point

Category	Penalty points	Description
Argon amount	0	no argon visible (leakage fraction (l.f.): 5%)
	1	$2 \times 10^{-5} < \text{AP rate} < 0.0025\%$ (l.f. 5%)
	1.5	$0.0025\% < \text{AP rate} < 0.13\%$ (l.f. 20%)
	2	$0.13\% < \text{AP rate} < 0.265\%$ (l.f. 60%)
	2.5	$0.265\% < \text{AP rate}$ (l.f. 90%)
Suspicious batches	0	not in suspicious batches
	0.5	in suspicious batches
Light emission	0	no LE
	1	possibly small micro-LE or other minor LE-related issues
	2	major indications for micro-LE
Spare PMTs	0.25	all spare PMTs (high failure rate in XENON1T, are older, He afterpulses)
	0.5	spare PMTs that were not retested in LXe within the XENONnT tests
XENON1T issues	0.25	square pulses in count rate
	0.5	low gain $< 2 \times 10^6$ @ 1500V
	1	slow gain decrease

Table 3.3.: Quantitative penalty point system developed for determining the overall qualities of PMTs for the XENONnT detector. The penalty points are assigned based on the results of the PMT testing campaigns and, for the PMTs that are reused, on their performance in the XENON1T detector.

system was developed that assigns each tube penalty points for undesirable characteristics as determined from the results of the testing campaigns. The assignment of penalty points is summarized in [Table 3.3](#).

Vacuum degradation is identified as the most severe issue affecting PMTs during operation in a TPC, causing issues with data quality and stability in milder cases, starting to flash at increasing frequencies in more severe cases, and even rendering the PMTs in-operable in the worst case. As we have seen in [subsection 3.2.4](#), the amount of argon gas present in the tubes from production can be used as a powerful predictor for future leak development of otherwise well-behaved PMTs. Therefore penalty points for the argon amount in each tube are given based on the measured argon afterpulse rates, labelled ‘argon penalty’ in [Table 3.3](#). Here ranges in AP rate are defined for the assignment of penalty points, with the fraction of tubes that showed leakage in that range (using all available PMT data including from the XENON1T TPC) as shown in parentheses. Tubes with highest argon penalty of 2.5 penalty points show a leakage fraction of 90%, and are considered to be at an extremely high risk of failure. In addition to the categorization via the the argon afterpulse rates, the clustering of PMTs showing vacuum degradation around two specific regions in PMT serial number, which correspond to two delivery batches from the producer, are taken into account as well. All PMTs from these batches, a total of 69 tubes, are given 0.5 penalty points in the category labelled ‘suspicious batches’.

Penalty points for light emission are given as well, dependent on whether they showed slightly increased levels of micro-LE at high HV biases. A LE penalty of 1 is given to tubes that showed micro-LE at somewhat more than low intensities of a few tens of hertz, as well as to tubes that showed hints of LE outside of the actual LE tests but not in the LE tests themselves. A penalty of 2 is given to tubes that showed intermediate levels of micro-LE during the tests, with intensities at a few hundred hertz.

The spare PMTs from XENON1T times are given additional penalty points (labelled ‘spare PMTs’ in [Table 3.3](#)). Since they belong to the in general much more failure-prone subset of old PMTs from the previous production cycle, but in contrast the tubes installed in XENON1T have not been proven to operate stably over several years. While a total of 32 spare PMTs were available, 19 of them were retested using the XENONnT testing campaign procedure. PMTs that were retested are thus given 0.25 penalty points, while the others are given 0.5 penalty points.

Finally, minor issues for the PMTs from XENON1T observed during the detectors run time are penalized as well. These issues are: Periodic spikes in the dark count rate of unknown origin that does not affect other PMTs, likely related to electronics issue such as noise (0.25 penalty points, 4 tubes), comparatively very low gain values of less than 2×10^6 (0.5 penalty points, 30 tubes), and a slowly decreasing gain over time (1 penalty point, one tube).

The distribution of the total penalty, i.e. the sum of all penalty points in the individual categories, for all available PMTs is displayed in the bar graph in [Figure 3.8](#). Full-color bars correspond to the subset of new PMTs that were tested in the XENONnT testing campaign, while the lighter shaded color bars show the full set of tubes available for installation. PMTs with zero total penalty make up the largest fraction, with more than 160 tubes in total. A selection criterion was set to exclude the use of PMTs with more than 2.0 total penalty, indicated by the gray shaded exclusion region. Coincidentally, applying this threshold corresponds to exactly the needed 494 pieces, that make up the final selection for the XENONnT PMT arrays.

Distribution of PMTs in the arrays

Having selected the best 494 PMTs, we need to distribute them within the two arrays. For the decision on which tube goes to which spot, we need to consider the different demands on PMT characteristics and quality between the two arrays, as well as for different regions within each array. Because the bottom PMT array is immersed in liquid xenon, the number of scintillation

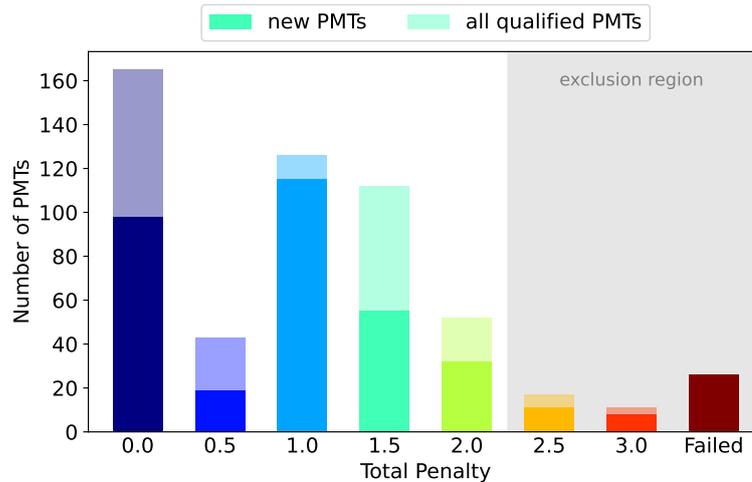


Figure 3.8.: Bar graph of the total penalty points for the new PMTs (full colors) and all PMTs that qualified for use in XENONnT (lighter shaded colors). The selection threshold is set to a total penalty of at most 2.0 points, as indicated by the grey shaded area labelled ‘exclusion region’. Note that in the ‘Failed’ bar on the very right, only the new PMTs that failed in the XENONnT testing campaign are shown, because of the pre-selection of only reusable tubes from the XENON1T detector. For visualization purposes, decimal points in the total penalty of .25 or .75 are rounded down to the nearest multiple of 0.5. The penalty point distribution for the new PMTs is published in [23].

photons from an S1 signal that will arrive at the bottom array is in general larger than for the top array, mainly due to internal reflection of the photons on the liquid-gas interface. Similarly, due to the position-dependent light collection efficiency caused by geometric and optical effects, the amount of incoming photons is generally larger towards the centers of the arrays. In order to increase this effect, the 241 PMTs with the highest QEs are selected for the bottom array, while the remaining ones are assigned to the top array. In a first ordering step, in both arrays the PMTs are then positioned concentrically, with the highest-QE tubes in the centers.

Next, we account for the fact that in the top array, the outermost ‘ring’ of PMT positions is of greater importance for S2 signal reconstruction. In this region the PMT area coverage is already lower due to the proximity to the TPC wall, reducing the resolution of the S2 signal size as well as the position reconstruction in the horizontal xy -plane. Near the walls of the TPC it is however very important to have a reliable position reconstruction, both for defining a reliable fiducial radius and to minimize the effect of surface background events originating from β -decays on the PFTE surface of the TPC wall. PMT failure in this region would therefore disproportionately affect data quality. To reduce the risk of PMT failure in the top array outer ring, PMTs with a high total penalty are swapped with lower total penalty ones from further inward in the array, while maintaining the radial QE gradient as much as possible.

In further resorting steps, in each array the tubes are swapped with tubes of similar QE in a way that avoids clusters of PMTs with high total penalty, with a special focus on tubes with high argon afterpulse rates. This mitigates potential larger blind spots in the arrays in the case of several neighboring PMTs failing. Similarly, PMTs from the two suspicious batches are not placed next to each other, as well as PMTs that have high penalties in any of the categories listed in Table 3.3. Finally, the PMTs’ affiliation to the two large-scale production batches for XENON1T and XENONnT is considered as well, making sure that the old tubes from XENON1T times and the new tubes are evenly distributed. This is done because the old PMTs from the previous production cycle for XENON1T on one hand suffer generally more from failures due to vacuum degradation and light emission, but on the other hand the specific tubes that were chosen to be reused have already survived more than two years of continuous operation in the XENON1T TPC without showing issues. At the same time, the new PMTs are expected to be

of better quality, as has become evident in the new testing campaign, but it is possible that an as-yet-unknown issue that affects PMT quality may emerge during operation in XENONnT. In all these cases, it is beneficial to mix the two sets of PMTs in the arrays as evenly as possible.

The resorting steps described above are repeated for each considered category in an iterative way, until all criteria are fulfilled. The resulting final positioning of the PMTs in the arrays is presented in [Figure 3.9](#). The top panel shows the PMTs' QE distribution over the top (left) and bottom (right) PMT arrays. The radial gradient of QE is visible in both arrays, with average QE values of 31.9% and 36.4% for the top and bottom array, respectively. The center panel shows the argon afterpulse rates of the tubes, with none of the very high rate PMTs located in the outer ring of the top array or close to another one. The total penalty for each PMT is shown in the bottom panel, using the same color scale as for the bar graph in [Figure 3.8](#), with the maximum total penalty of any PMT in the TPC at 2.0 points.

3.3.2. PMT array installation

The XENONnT array assembly and installation of the PMTs took place in late 2019 at LNGS, followed by the assembly of the entire TPC in early 2020. Prior to assembly, all materials used in construction were thoroughly cleaned in order to remove small particulates, grease and lubricants from the production process, and radioactive contaminants that may have attached to the material surfaces during production and transport. The cleaning procedure took place in a cleanroom (CR) with ISO 6 classification at the above-ground facilities of LNGS. To avoid recontamination of the cleaned materials, the assembly took place in the CR as well. A detailed description of the cleanliness procedures in XENONnT can be found in [\[18\]](#).

The top and bottom array were constructed separately in October and December 2019, both following the same procedure. The structures of the arrays are assembled first; a photo of the empty top array structure is shown in [Figure 3.10](#) (Top left). They consist of a diamond-polished PTFE reflector plate with circular cutouts for the PMTs, following the hexagonal pattern already described in [subsection 3.3.1](#). Once placed in the array structure, the PMT windows face the liquid xenon target through these cutouts, while the PTFE plate acts as a reflector for scintillation photons that do not hit a PMT window directly. The PTFE plate is connected to a copper plate via copper rods, providing structural support to the arrays. The copper rods extend through to the other side of the copper plate, where spacers made of PEEK (polyether ether ketone) are screwed on top of them. These serve as support for the PMTs, which are held in place by three-legged, so-called 'tripod clamps' that get screwed onto the top of the rods (the tripods can be seen in [Figure 3.10](#) (Bottom right)). The components of the array structures are cleaned in the CR prior to assembly, described in detail in [\[18\]](#). For the copper parts, a cleaning recipe is used that includes an ultrasound bath with a neutral detergent, a bath in Piranha solution (H_2SO_4 (1%) + H_2O_2 (3%)), and a bath in citric acid (5%) for passivation. The PTFE and other plastic (mainly PEEK) parts are cleaned in a bath with a neutral detergent as well (without the ultrasound for the diamond-shaved parts), as well as a bath in nitric acid solution (HNO_3 (5%)).

The procedure of cleaning and bringing the tubes into the CR was performed in several steps, for batches of about ten tubes at a time. In a first step, before bringing them into the CR, the PMTs were wiped down with cloth wipes dipped in ethanol and blown off with a nitrogen gun, in order to remove grease as well as dust and small particulates. They are then passed into the ante-room of the CR, where they are wiped again with ethanol and finally brought into the CR itself. Once inside the CR, the PMTs underwent a final and most thorough cleaning step, depicted in [Figure 3.10](#) (Top right). In order to also remove dust and other contaminants from areas that are hard to reach with wipes, including a groove between the edge of the windows and the metal body of the tubes or in the vicinity of the pins, the PMTs were submerged in a bath of high purity (> 99.9%), analysis-grade ethanol and carefully moved in a swishing motion. The ethanol in the bath was completely replaced with new ethanol roughly every 50 tubes, avoiding the accumulation of dust and other contaminants in the bath. After bathing, the tubes were

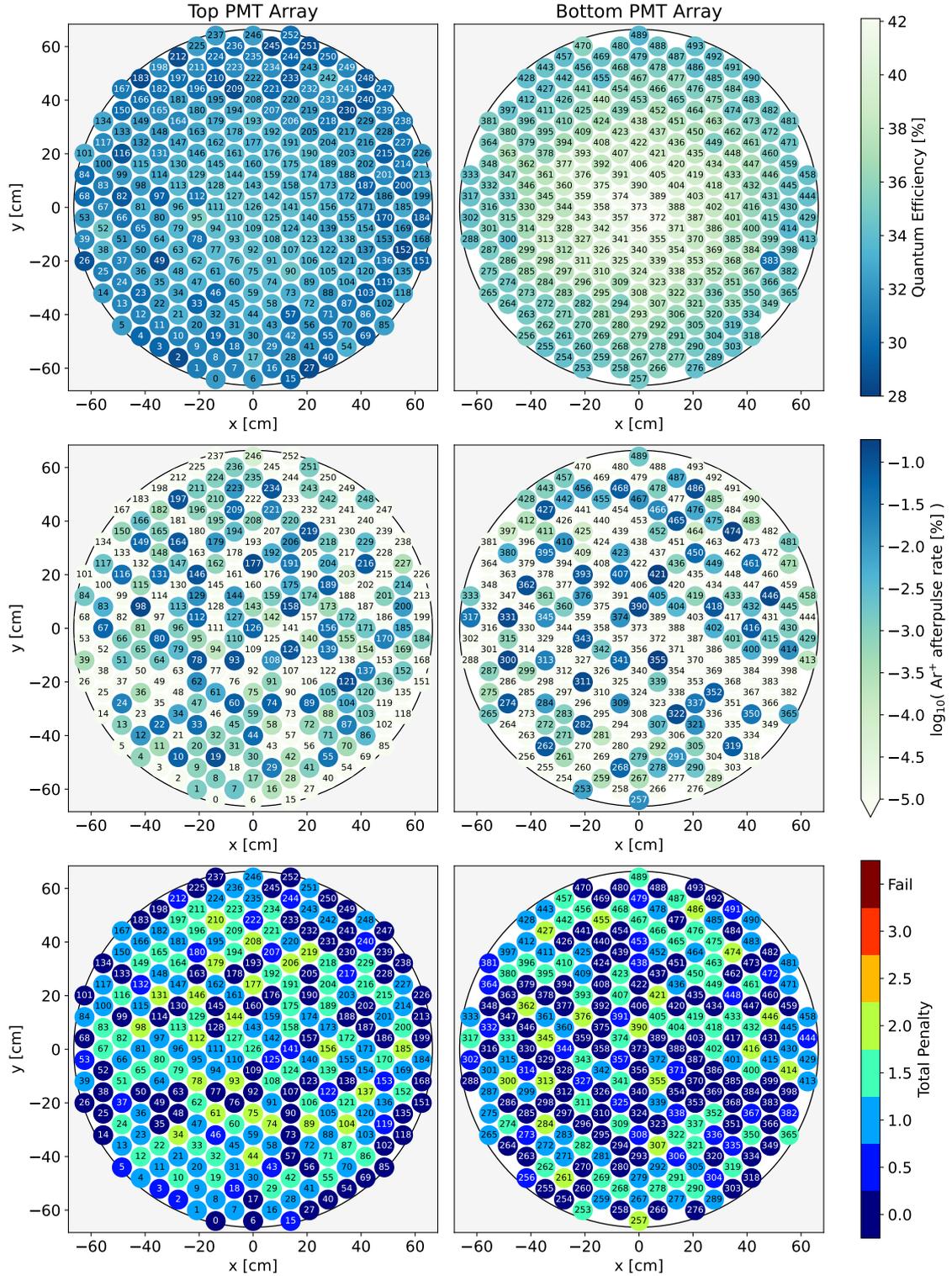


Figure 3.9.: Distributions of the most important characteristics used for the placement decision of the PMTs in the XENONnT PMT arrays (details in text). The coordinate system is identical to the one used in data analysis, corresponding to a bird’s-eye view from the top onto the cylindrical TPC. (Top) Quantum efficiency distribution. Tubes with the highest QE values are placed in the center of the bottom array in order to maximize the light collection efficiency. (Center) Argon afterpulse rates. PMTs with high rates are more likely to develop leaks, and are placed avoiding clusters and critical regions. (Bottom) Total penalty points. The same positioning criteria as for the argon AP rates apply. The color scale is the same as used in Figure 3.8, with the maximum total penalty of any PMT in the TPC at 2.0 points.

wiped again and blown off with nitrogen at a rather low flow for drying. During this step we noticed that some of the PMTs showed stains that were not visible before the ethanol bath, an example can be seen in the photo in [Figure 3.10](#) (Bottom left). These stains were found to be residue of the gloves that got released when holding the tubes while they were still wet with ethanol. They were easily removed by wiping the tubes again with a clean cloth wipe. After this was noticed we took care not to touch the tubes too much with the gloves while they were still wet with ethanol.

The PMTs were then put into their predetermined positions within the array structures in batches of 24. These batches correspond to ‘sectors’ of tubes that have the signal and HV cables from their bases bundled together into so-called ‘connectors’. The installation of one sector of PMTs into the array structure is shown in [Figure 3.10](#) (Bottom right). The bases are connected to the PMT pins, and the tripod clamps are attached to the tubes but not yet screwed into place. Afterwards the signal and HV connectors are assembled and their electric connectivity quickly tested using a multimeter. Once all tubes are set into their designated positions, the bases attached, and the connectors assembled and successfully tested, we fasten the legs of the tripod clamps to the support rods, thus fixing the PMTs to the array support structure.



Figure 3.10.: Photos of the array structure assembly and installation of the PMTs in the CR at LNGS. (Top left) The empty array structure, showing the PTFE reflector plate and the copper holding plate, as well as the copper spacing rods connecting them. (Top right) PMTs being cleaned inside the CR by swishing them in an ethanol bath and wiping afterwards. (Bottom left) Residue found on the PMTs after ethanol bathing, likely originating from the gloves. (Bottom right) PMTs being placed in their designated positions, bases are attached, and the tripod holders set in place.

A top-view picture of the fully assembled top PMT array is shown in [Figure 3.11](#). In a careful visual inspection of the arrays, we found a few small particulates on the structure, likely copper or PEEK shavings from threads in the materials. They were removed with a vacuum cleaner setup consisting of plastic tubing connected to a pump outside the CR.

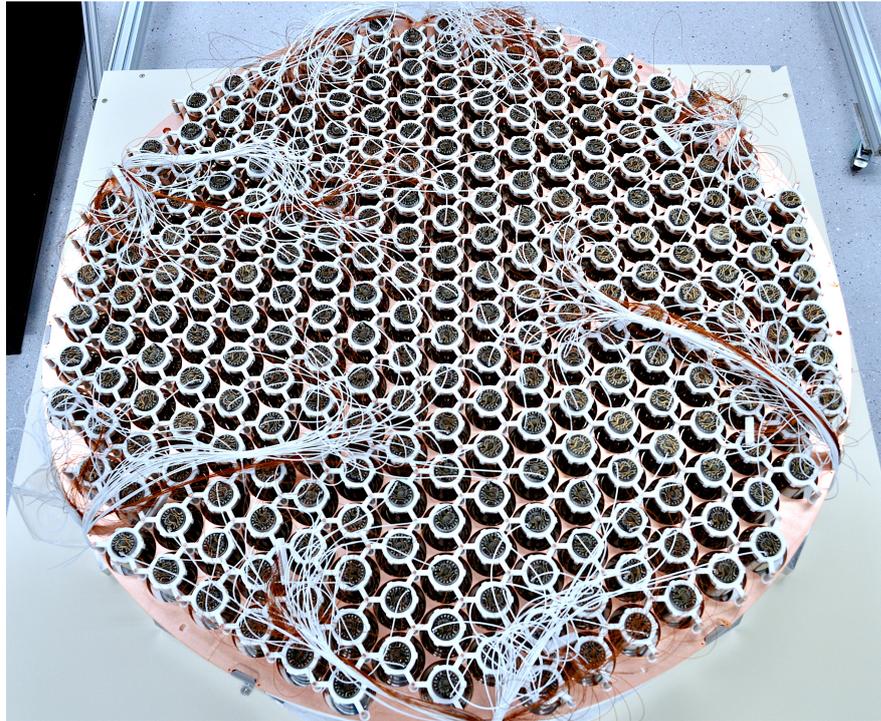


Figure 3.11.: Photo of the top PMT array after installation of all PMTs, showing the ‘backside’ with the attached PMT bases. The PMTs are held in place by the tripod clamps, which are screwed onto the peak spacers at the top of the copper rods. The signal and HV cables (white and copper colored (kapton), respectively) are bundled together into connectors (not visible here, they are underneath the array).

After the PMT installation, the arrays were covered in a dark box in order to verify the proper functioning of all PMTs. The dark box is a light-tight box that mainly consists of a plastic lid that can be put over the arrays while sitting on the table used for assembly (as seen in [Figure 3.11](#)). The lid is sealed to the table using black plastic tape, ensuring light-tightness. Four elbow-shaped tubes on the sides of the box lead the signal and HV connector bundles out of the box, allowing for the connection of the signal and HV cable of each PMT to an oscilloscope and a HV power supply, set up outside of the CR. The PMTs were ramped up one by one to an operating voltage of 1.6 kV and checked for proper dark count signal pulses on the oscilloscope screen. Except for one tube, all 494 PMTs showed proper functioning in the dark box tests. The one exceptional tube could not be ramped up to a proper operating voltage without tripping, likely related to severe vacuum degradation. It was directly replaced with another tube that showed proper functioning (note that in all distributions and numbers shown in this chapter, this tube is already included as if it was the one being used from the beginning). For one other tube, we did initially not see any signals, this was however traced down to a faulty base that could consequently be fixed.

After the successful dark box tests, the arrays were stored in the same box outside of the CR, while constantly being flushed with nitrogen gas in order to avoid any contamination from dust or from radioactive isotopes from the surrounding air attaching to (especially the PTFE) surfaces.

The TPC itself was assembled in February 2020 in the same CR. The assembly was performed from top to bottom, starting with attaching the stainless steel diving bell to an item profile support structure. The top PMT array was then lifted up and inserted into the bell, as seen in [Figure 3.12](#). Afterwards the TPC field cage, consisting of the PTFE wall panels surrounded by the field shaping rings and wires (see [Figure 2.3](#)) as well as the gate, anode and top screening

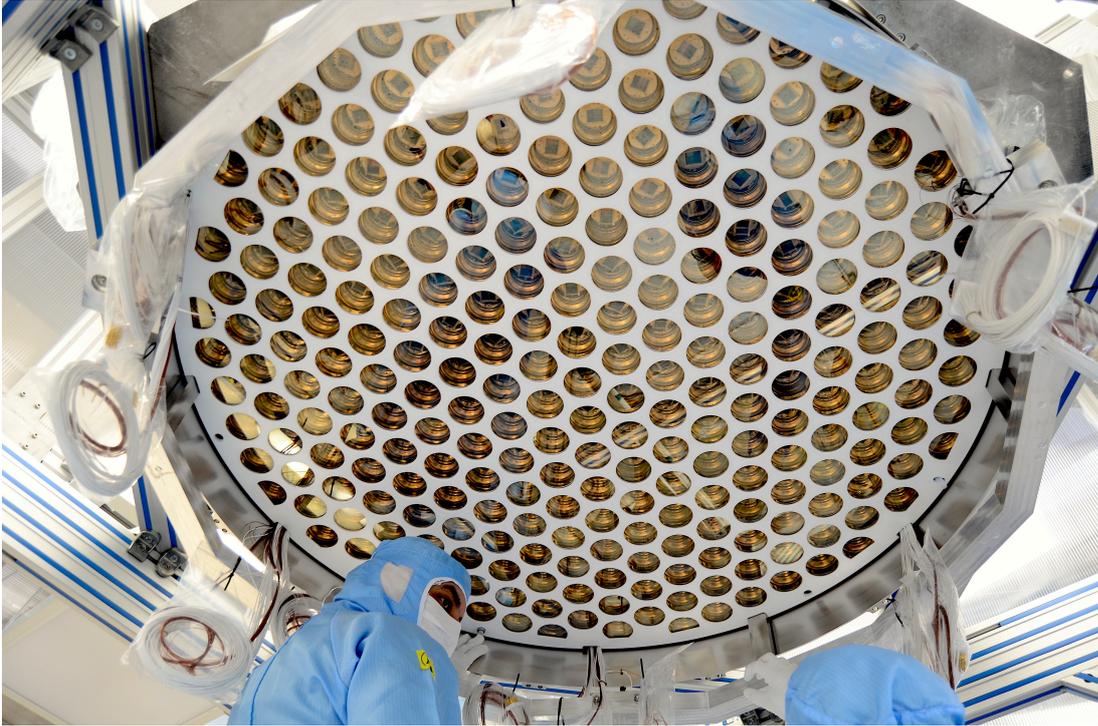


Figure 3.12.: Photo of the top PMT array inside the stainless steel diving bell during the TPC assembly, as seen when standing underneath the assembly support structure.

electrodes attached on top, were assembled and connected to the top PMT array and bell. Having been installed with the PMT windows facing downwards, the bottom PMT array had to be rotated by 180° such that the PMTs face upwards; a photograph of the bottom array during this process is shown in [Figure 3.13](#). After installing the bottom screen and cathode electrodes on top of the bottom PMT array, the TPC was finalized by attaching the bottom array to the bottom of the field cage.

The assembled TPC was then transported to the underground laboratory in early March, where it was installed in a second CR situated inside the empty water tank of the XENONnT experiment, by attaching it to the top dome of the inner cryostat. All PMT signal and HV cables were connected to the corresponding cable connectors coming from the two large 'umbilical' tubes on top of the cryostat dome, which lead the signal and HV cables towards the support building where they exit the gas system via feedthroughs and are connected to the DAQ system. Beforehand, all PMT cable connections were carefully and successfully checked for proper connectivity between the TPC side and the DAQ side. After installing several other components around the outer circumference of the TPC, namely liquid level meters, Pt100 temperature sensors, the HV connections for the electrodes, as well as the tubing for xenon in- and outlets, the inner cryostat vessel was lifted up over the TPC and closed by connecting it to the top dome via a spring-energized metal seal. Immediately afterwards the evacuation of the inner vessel containing the TPC began, officially starting the commissioning phase of the XENONnT experiment on March 25th, 2020.

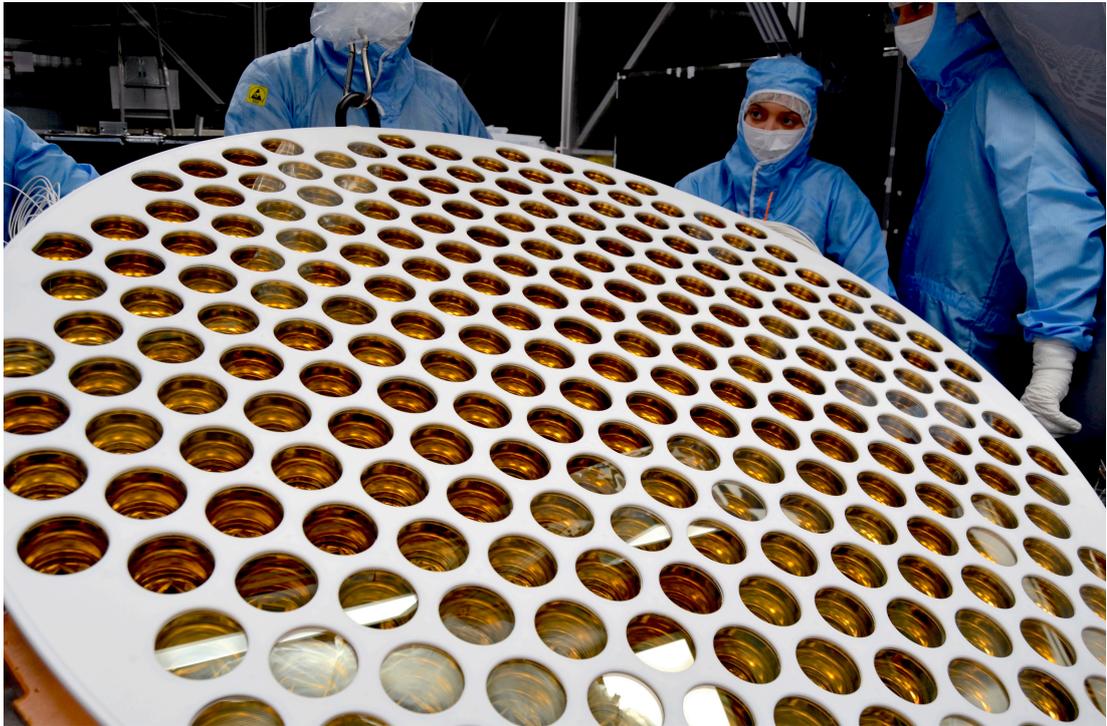


Figure 3.13.: Photo of the bottom PMT array during the full TPC assembly, showing the side of the PTFE plate and PMT windows facing the active xenon volume in the TPC.

Chapter 4.

PMT afterpulse characterization and monitoring in XENONnT

With the start of the commissioning phase of the XENONnT experiment, the monitoring of the photomultiplier tube (PMT) characteristics and performance in the TPC began, a vital part of the data quality monitoring in the experiment. During operation, regular PMT calibration measurements are taken in order to determine the gains, single photoelectron (SPE) acceptances, and the afterpulse rates of the PMTs over time. While the gains and SPE acceptances are used in the very first steps of the data processing chain, the measurements of the afterpulse rates are used for general data quality monitoring purposes, as they may induce a signal size bias by getting merged into signal peaks, an effect that is both signal size and width dependent and therefore mainly affects S2 signals. This can lead to a bias in energy reconstruction, as well as complicating xy -position reconstruction which is based on the relative amount of S2 light detected in each PMT in the top array. More severely, PMTs with growing afterpulse rates due to vacuum degradation additionally cause a time-dependent behavior in the aforementioned effects, that may be difficult to correct for in the data analysis. Measuring and monitoring the afterpulse rates is therefore essential for data quality determination and early detection of PMT issues during operation.

In this chapter, we focus on the characterization, measurement, and monitoring of PMT afterpulses in the commissioning phase and during the first science run of XENONnT. In [section 4.1](#), the characteristics of afterpulses in the XENONnT PMTs and the identification of the ions that produce them are introduced. The PMT calibration system of XENONnT, afterpulse data taking and processing procedures, and the analysis routine developed for monitoring afterpulse rates are described in [section 4.2](#). Results of the afterpulse measurements and monitoring are presented in [section 4.3](#), along with an evaluation of PMT performance in XENONnT with respect to the categorization system described in the previous [chapter 3](#). Finally, a summary and conclusions are given in [section 4.4](#).

4.1. Afterpulses characteristics and identification

The term ‘afterpulses’ generally refers to any kind of PMT pulse that is created inside a PMT and is delayed with respect to a primary pulse that created it. Most commonly, the term is used for afterpulses created via ion feedback, where, as already briefly explained in the previous chapter, residual gases inside the tube vacuum get ionized by a photoelectron of the primary pulse and are accelerated back towards the cathode, causing delayed pulses with delay times that are characteristic for the mass-to-charge ratio of the ion that created it.

The afterpulse delay time of a specific ion corresponds to the time it takes for the ion to travel from its point of creation back to the photocathode, and can be calculated from the mass-to-

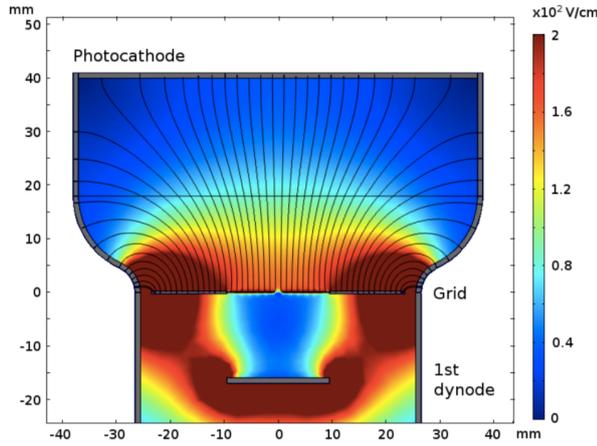


Figure 4.1.: Simulation of the electric field inside the R11410 PMT body between photocathode, focussing grid and first dynode. Figure from [31].

Ion	M/Q	AP delay time [μs]
H_2^+	2	0.39
He^+	4	0.55
N_2^{++}	14	1.04
CH_4^+	16	1.11
Ar^{++}	20	1.24
N_2^+	28	1.47
Ar^+	40	1.75
Xe^{++}	65	2.24
Xe^+	131	3.17

Table 4.1.: Calculated afterpulse delay times following Equation 4.3 for typically observed ions, at a bias voltage of 1300 V.

charge ratio of the ion and the electric field between cathode and first dynode. In the case of the R11410 PMTs, the focussing grid in front of the first dynode (see 3.2) is held at the same potential as the first dynode. The shape of the electric field between cathode and focussing grid, largely determined by the tapered tube geometry of the PMTs, results in ion travel times that only depend on the mass-to-charge ratio of the ion, as well as on the PMT bias voltage determining the strength of the electric field [31, 37, 43]. This can be understood by the general expression of the delay time t of an ion traversing the PMT from its point of creation s_0 to the photocathode at position $s = 0$, given by [44]

$$t = \int_{s_0}^0 \frac{1}{v(s)} ds = \sqrt{\frac{2q}{m}} \int_{s_0}^0 [V(s) - V(s_0)]^{-1/2} ds \quad (4.1)$$

where $v(s)$ is the ion's velocity in dependence of its position as it is accelerated by the electric potential $V(s)$, and is easily obtained following simple kinematics and the electric component of the Lorentz force. q and m are the charge and mass of the ion, respectively. The tube geometry with the body narrowing between window and focussing grid leads to electrical field lines that converge towards the focussing grid, closely resembling that of hemispherical PMTs. For these the delay times have been calculated analytically using a quadratic electric potential distribution of the form $V(s) = V_0 \left(\frac{s}{L}\right)^2$ [43], where V_0 and L are the potential difference and distance between photocathode and focussing grid, respectively. With this potential shape, the delay time becomes

$$t = \frac{\pi}{4} \sqrt{\frac{2m}{qV_0}} L = \left(1.134 \frac{\text{V}^{1/2}}{\text{cm}} \mu\text{s} \right) \sqrt{\frac{L[\text{cm}]^2}{V_0[\text{V}]} \frac{M}{Q}}, \quad (4.2)$$

with the dimensionless mass-number-to-unit-charge ratio M/Q . The delay time is independent of the position of the ion's creation s_0 , as it cancels out when solving the integral in Equation 4.1. With $L = 4.1$ cm in the R11410 PMTs, and $V_0 \approx 0.2162 \cdot U$ set by the resistances in the voltage divider chain on the bases of the PMTs [45], the expected delay time of an ion with mass-to-

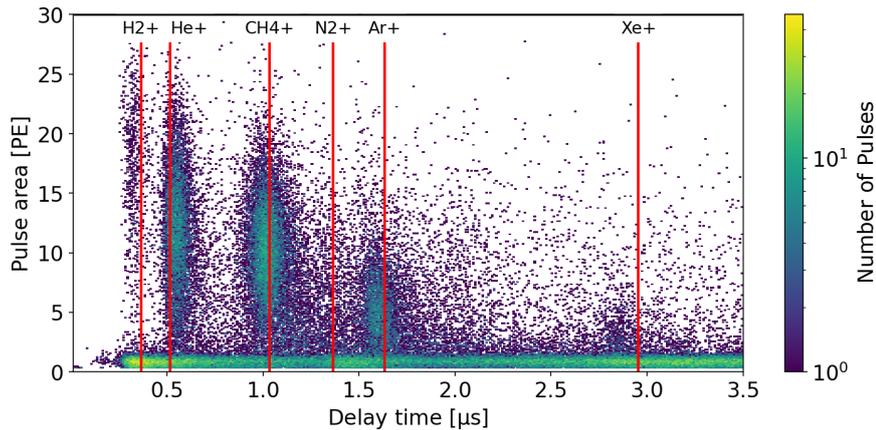


Figure 4.2.: Typical afterpulse spectrum of an R11410 PMT in XENONnT, measured at a bias voltage of 1500 V at room temperature. The expected delay times of ions commonly found in the tubes are indicated by the red vertical lines. A small contamination of xenon ions is visible as well, indicating a leaky PMT.

charge ratio M/Q is fixed for a given operational bias voltage U , and given via

$$t \approx \left(\frac{4.6494}{(0.2162 U/V)^{1/2}} \right) \mu\text{s} \cdot \sqrt{\frac{M}{Q}}. \quad (4.3)$$

The validity of this approach, in particular the assumption of the quadratic potential distribution, has been verified with a COMSOL Multiphysics® [46] simulation of the electric field inside the PMT body in [31], which is shown in Figure 4.1. Using this electric field map, the drift times of several ion species distributed uniformly in the volume were simulated and found to be in good agreement with the calculation following Equation 4.3 as well as with experimental data, deviating on average by about 3% [31]. The calculated delay times for the ions that are typically observed in the AP time spectra of the R11410 PMTs are shown in Figure 4.1, calculated for the typical PMT bias voltage in XENONnT of about 1300 V. Note that the ion identification is however not unique, as the mass-to-charge ratio is degenerate for certain ion species. For example N_2^+ and CO^+ have the same mass-to-charge ratio of 28, as well as Ar^+ and CO_2^+ with a mass-to-charge ratio of 44, and CH_4^+ and O_2^+ with a mass-to-charge ratio of 16. Since we cannot distinguish between these ions in the afterpulse measurements, in this work they are always referred to as N_2^+ , Ar^+ and CH_4^+ , respectively.

An example of a typical afterpulse spectrum is presented in Figure 4.2, showing the delay time on the x-axis and the pulse area on the y-axis. The spectrum was taken at room temperature and a bias voltage of 1500 V. The expected ion delay times are marked by the red vertical lines, calculated from Equation 4.3 with a bias voltage of 1500 V. The elliptical populations of pulses with areas above about 2 PE are identified as the ion-induced afterpulses, as labelled in the figure. At room temperature, all PMTs show afterpulses from hydrogen, helium, and methane, while about 60% of the tubes show argon afterpulses as well, as already discussed in subsection 3.2.4. The presence of xenon afterpulses in the spectrum indicates a leaky PMT.

The pulse areas of the afterpulses show a clear decreasing trend with increasing delay time. This is understood as an effect of the ions' velocity on impact with the photocathode, which is inversely proportional to the square root of the mass-to-charge ratio, such that lighter ions arrive with higher kinetic energy and thus create more secondary electrons upon impact on the cathode, producing larger pulse areas.

In addition to the ion-induced afterpulses, a large population of pulses with areas at the single PE level and at all delay times is visible as well. These SPE-like afterpulses are present in

all PMTs, and show a steeply falling intensity up to delay times of about $2\ \mu\text{s}$, followed by a bump between about 3 and $4\ \mu\text{s}$ (see also [Figure 3.6](#) (left)). At very small delay times below about 100 ns, late pulses from the LED signal itself (see ‘spurious pulses’ in [subsection 3.1.1](#)) are assumed to be the major cause. At larger delay times, however, their origin is unknown, with hypothesized explanations being fluorescence processes of the PMT glass window, or from other materials in the vicinity of the PMTs.

4.2. PMT calibration system and afterpulse measurements

Measuring PMT properties *in-situ*, i.e. during active operation in the experiment, requires a dedicated calibration system which, similarly to the PMT testing setups, is based on pulsed LEDs as a light source. The PMT calibration system of XENONnT was developed and is described in detail in [47], and is sketched in [Figure 4.3](#). A pulse generator used as an LED driver provides the synchronized input signals for four blue LEDs with an emission wavelength of $\sim 460\ \text{nm}$. It is triggered by an external trigger from the DAQ system. In contrast to the triggerless data acquisition mode used in TPC data taking, the data for the PMT calibrations is acquired using the same external trigger signal as the LED pulse generator, triggering data acquisition for all PMT channels simultaneously with fixed-length acquisition windows. The trigger signal for the LED pulse generator can be delayed with respect to the DAQ trigger signal by an adjustable time interval, ensuring the arrival of PMT signals from the LED light within a predefined time range in the recorded waveforms. The LEDs are housed in a dark black box, from which the LED light is guided through an optical adapter into a vacuum breakout chamber located in the cryogenics system of XENONnT, via four acrylic optical fibers covered against ambient light. Using a second set of four fibers made of silica, that are resistant to a large temperature range, the light is then guided through the umbilical pipe connecting the cryogenics system with the cryostat inside the water tank, ending at a set of four connectors on top of the TPC diving bell. Per design, to each of these connectors a bundle of 6 fibers each can be attached, spreading the light over a total of 24 fibers with endpoint positions evenly spaced along the TPC’s circumference, in two rings at heights of $-20\ \text{cm}$ and $-25\ \text{cm}$ in z , respectively. They directly end in the active TPC volume by being inserted through small holes in the TPC wall reflector panels. The fiber positions were designed and optimized for optimal light

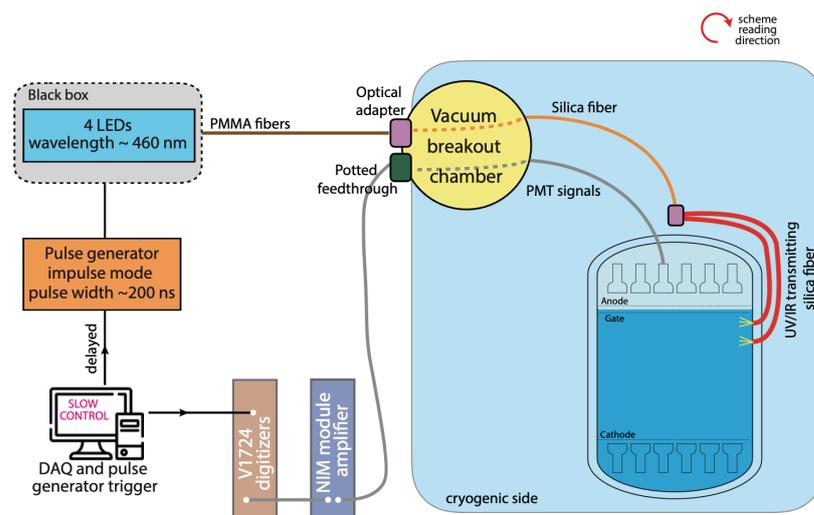


Figure 4.3.: The XENONnT PMT calibration system. For details see main text and ref. [47]. Figure by Ciara Capelli, taken from [25].

collection efficiency in both PMT arrays, as well as for redundancy in case of losses of the very fragile fibers, via optical simulations including several failure cases [47]. During installation of the TPC in the water tank, in fact several fibers were lost. One of the four silica fibers running through the umbilical tube was accidentally severed while arranging the PMT cables running through the same tube, while a second was found damaged likely already during XENON1T disassembly, allowing for the connection of only two of the four fiber bundles. Additionally, two of the fibers from the two connected bundles were accidentally pulled out of their holes in the TPC wall during operations around and at the TPC, leaving 10 of the total of 24 fibers operational. Despite these losses, the light intensity across the PMT arrays is sufficient for all required measurements, owing to the anticipation of fiber losses in the design of the system.

4.2.1. PMT afterpulse measurements

During operation of the PMTs in the XENONnT TPC, LED calibration and measurement runs are taken at least weekly, with the standard procedure consisting of a series of runs with different LED intensities as needed for the corresponding measurements. A full LED data taking set consists of several runs for PMT gain and SPE acceptance measurements, one run for afterpulse measurements, and one run for noise measurements. For the gain and SPE acceptance measurements, due to the inhomogeneous LED light distribution over the arrays several short runs at different LED intensities are necessary in order to have at least one measurement for each PMT with the needed light level for the gain estimation, as well as runs with the LED off (see also [25] for more details on the PMT gain estimation in XENONnT). For the afterpulse measurements, the LEDs are pulsed at a higher intensity, and the runs are taken for a longer duration in order to have a sufficient number of afterpulse events recorded.

In the weekly afterpulse measurement, the LEDs are pulsed at a trigger frequency of 406.5 Hz, with a run duration of 30 minutes. The LED voltage is set to 12 V (already above the maximum rating of the LEDs) in order to maximize the light output. The DAQ window length is set to 800 samples, corresponding to 8 μ s of recorded waveforms per PMT channel, ensuring that the full afterpulse time spectrum with the largest delay times of about 3 μ s for Xe^+ , as well as afterpulses of afterpulses at up to 6 μ s delay time (so-called secondary afterpulses), is covered. The delay between the LED pulser trigger and the DAQ trigger was adjusted such that the LED signals arrive near the start of the recorded waveforms within samples 50 to 100.

Figure 4.4 shows the intensity distribution of the LED light over the top and bottom PMT array for an example afterpulse measurement taken during SR0, extracted from the processed afterpulse data as will be described in the next subsection. Despite the fibers' endpoints being positioned much closer to the top array, the bottom array receives more light due to internal reflection of the photons on the liquid xenon surface, the same effect that is also present for the S1 signal light collection efficiency in TPC data. The intensity distribution over the bottom array is rather uniform with an average LED pulse area of 11.4 PE, with a slight decrease towards the edges. The top array has an average LED pulse area of 5.6 PE, but is less uniform due to the general positioning of the fibers closer to the top array, as well as due to reduced number of working fibers. The regions of highest intensity in the top array, e.g. around PMT 38, around PMT 325, or around PMT 70, closely correspond to the remaining working fibers' positions. The very low intensity at the outer ring of the top array is due to some PMTs being partly covered by the PFTE panels of the TPC walls per design. With the afterpulse run duration of 30 minutes and the 406.5 Hz external trigger frequency, a statistic of about 7.3×10^5 recorded events is achieved, resulting in the average total number of PEs recorded in the LED pulses in one measurement of about 7×10^6 , defining the statistical uncertainty of the afterpulse rate determination to be about 0.05 %.

4.2.2. Afterpulse data processing and analysis

The processing of recorded waveforms in XENONnT is based on the *strax* framework [48] (STReaming Analysis for Xenon TPCs). It was newly developed in the context of XENONnT data analysis, but is publicly available and can be generally used for any xenon TPC. It is a streaming processor purely written in python, and highly modular with a plugin-based structure that allows for easy extension and modification of the data processing chain. The XENONnT-specific processing software is called *straxen* [49], which is built on top of the *strax* framework for the use of data from XENONnT. For this work, a new *straxen* plugin was developed processing of the recorded afterpulse data, extracting all relevant information from the triggered waveforms needed for afterpulse rate monitoring purposes. The plugin is based on the *straxen* hit finder algorithm, which is adapted to the specific requirements of the afterpulse data, working on a per-channel-and-waveform level to extract pulse areas and relative arrival times of pulses within the same waveform. In particular, it allows for the definition of a fixed LED time window around the expected arrival time of the pulses induced by the LED light.

Similar to the TPC data processing chain, first the baseline average and root-mean-square (RMS) is determined from the first 40 samples in each raw waveform. The baseline average is subtracted, and the resulting waveform is flipped (because the raw PMT signals are negative), resulting in waveforms as shown in Figure 4.5. The plugin then searches for ‘hits’ in those waveforms, requiring excursions above a threshold of $4 \times$ baseline RMS for a fixed number of consecutive samples, with the left and right hit boundaries defined by the first sample above and first sample below threshold, respectively. The found hit boundaries for the example waveforms in Figure 4.5 are indicated by the yellow shaded regions. All hits in the same waveform that begin within the LED window, marked by the gray shaded region, are merged into one main LED pulse, in order to avoid artificial splitting. This has been observed to happen especially for PMTs in the top array where the LED illumination is less intense, which, paired with the minimum LED pulser width of 200 ns, can lead to single LED photons being individually identified within the LED window. For all hits, the pulse area is calculated by integrating the waveform within an integration window with extended boundaries 3 and 20 samples to the left and right of the found hit boundaries. Hits starting within the pulse boundaries (= hit boundaries \pm integration extensions) of a preceding hit get merged into the previous pulse in

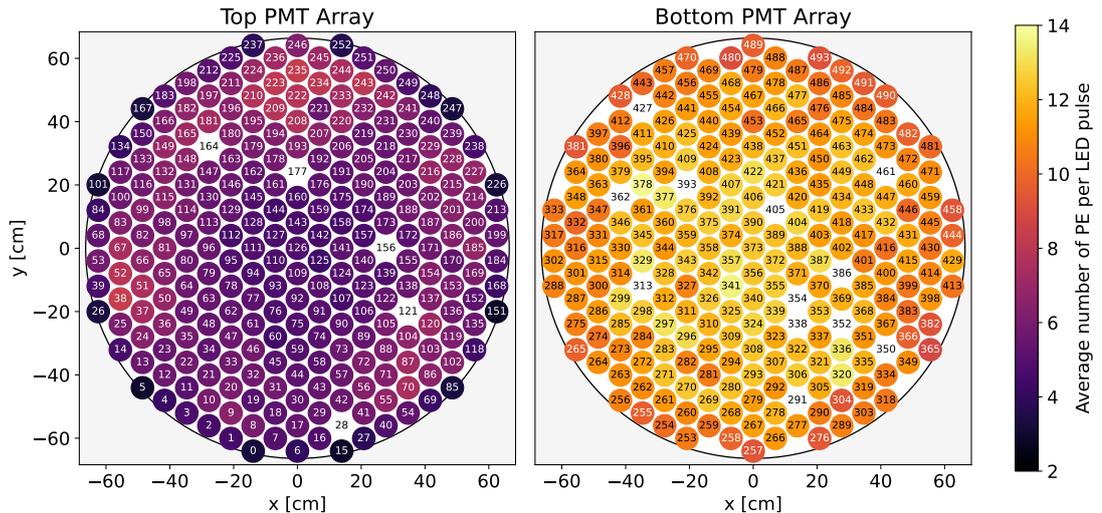


Figure 4.4.: Intensity distribution of LED illumination over the PMT arrays in an afterpulse measurement taken during SR0, with a measurement run duration of 30 minutes and an LED pulser frequency of 406.5 Hz. The color scale indicates the average pulse area for all LED pulses recorded in the run, in units of PE. PMTs without color were turned off during SR0 (see subsection 4.3.2).

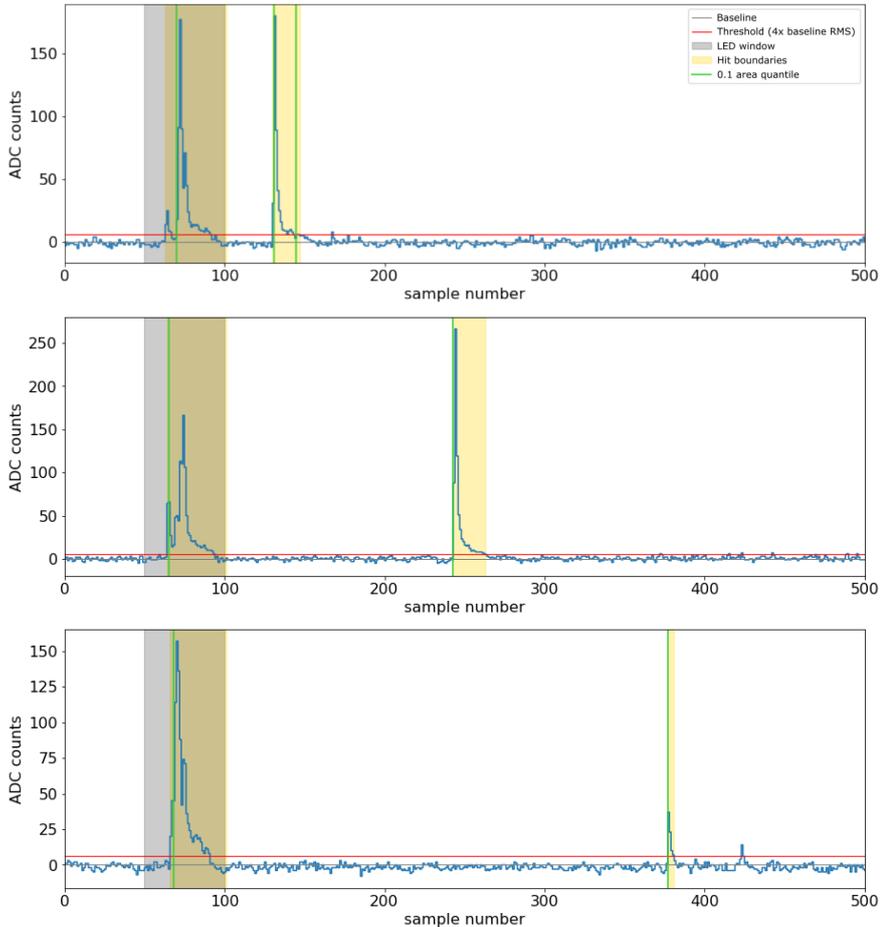


Figure 4.5.: Example waveforms recorded during AP LED measurements, after baseline subtraction and inversion, illustrating the working principle of the *straxen afterpulses* plugin. The set LED window is indicated by the grey shaded region. Found hits are shown by the yellow shaded regions. All found hits inside the LED window get merged, resulting in the right hit boundary of LED pulses extending to at least the right boundary of the LED window. The green vertical lines indicate the 10% quantile of the pulse areas, used as the pulse arrival times. The visible afterpulse hits arriving after the LED window correspond to delay times of He^+ (top), Ar^+ (middle), and Xe^+ (bottom).

order to avoid artificially splitting the falling tails of pulses due to early fluctuations below the excursion-above-baseline threshold, as seen in the top panel in [Figure 4.5](#). The pulse areas are then converted to the number of photoelectrons (PE) using the time-dependent PMT gain model that is determined from the gain calibration runs for use in all TPC data processing (see [\[25\]](#) for details on the gain measurements). The arrival time of the pulses is determined using the 10% pulse area quantile, defined as the time of the first sample with an integrated pulse area above 10% of the total pulse area. The pulse arrival times are indicated in [Figure 4.5](#) by the green vertical lines. The delay time of the pulses is then calculated as the difference between the pulse start time and the LED pulse start time within the same waveform. With this, the delay time of the LED pulses per construction becomes zero, allowing for an easy identification of the LED pulses in the afterpulse data analysis.

An example for the extraction of an afterpulse spectrum for one PMT channel from a measurement run processed with the *afterpulses* plugin is given in [Figure 4.6](#). The left panel shows the histogram of the pulse areas vs. the start time of all pulses found in the recorded waveforms. The pulses found in the LED window (grey shaded region) get merged into the main LED pulse for each waveform. In the right panel, the same measurement is shown, but showing the delay

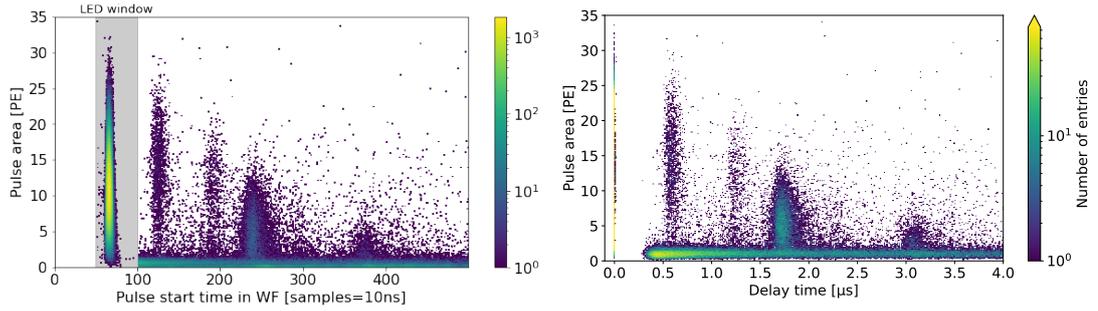


Figure 4.6.: (Left) Histogram of pulse areas vs. start time of all pulses found in all recorded waveforms for one PMT channel during an afterpulse measurement, using the *afterpulses* plugin of *straxen*. The LED window definition is indicated by the grey shaded region. (Right) The resulting afterpulse delay time spectrum, produced by the per-waveform calculation of the delay time of each found pulse with respect to the LED pulse in the same waveform. The color scales in both panels indicate the number of pulses per bin, capped on the right panel at the maximum value outside the zero delay time bins for visibility.

time of each pulse with respect to the LED pulse in the same waveform on the x-axis. The LED pulses are clearly visible at zero delay time.

The afterpulse rates are then determined from the afterpulse delay time spectra, by defining the afterpulse rate for a specific ion i of a PMT as

$$AP_i = \frac{N_i}{N_{\text{PE,LED}}} \cdot 100 \text{ [\%]}, \quad (4.4)$$

where N_i is the number of afterpulses created by ion i , and $N_{\text{PE,LED}}$ is the total number of initial photoelectrons produced by the LED pulses in the measurement. The number of afterpulses N_i is determined for each PMT by fitting the afterpulse peaks in the projected delay time spectra. First, most of the SPE-like afterpulses around pulse areas ~ 1 PE are removed by applying a cut on the pulse height, requiring a minimum pulse height of 0.6 PE/sample. The resulting spectrum is shown in [Figure 4.7](#) in the left panel. Around each expected delay time for the ions, a window of fixed width to the left and right is defined, within which the afterpulse peaks are expected to be located, indicated by the red vertical lines and shaded regions in the figure. To extract the number of afterpulses N_i for each ion species, the delay time spectra are projected onto the delay time axis within these windows; two examples of the resulting delay time spectra are shown in the center and right panels in [Figure 4.7](#). The method developed here uses a fit function that consists of a combination of a Gaussian core and exponential tails on both sides described in [\[50\]](#), defined as

$$f(x; \mu, \sigma, k_l, k_r, A) = A \times \begin{cases} \exp \left[\frac{k_l^2}{2} + \left(\frac{x-\mu}{\sigma} \right) \cdot k_l \right] & \text{for } \left(\frac{x-\mu}{\sigma} \right) \leq -k_l \\ \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right] & \text{for } -k_l < \left(\frac{x-\mu}{\sigma} \right) \leq k_r \\ \exp \left[\frac{k_r^2}{2} - \left(\frac{x-\mu}{\sigma} \right) \cdot k_r \right] & \text{for } k_r < \left(\frac{x-\mu}{\sigma} \right) . \end{cases} \quad (4.5)$$

The parameters μ and σ are the mean and standard deviation of the Gaussian core, A is the amplitude, and k_l and k_r are the decay constants of the exponential tails to the left and right respectively. For the fitting process an additional constant offset is added in order to account for the remaining contribution from the SPE-like afterpulse population. The fit function was chosen motivated by the observation of the tails in the afterpulse peaks not being well described by a Gaussian distribution alone, as can be seen in the example spectra in [Figure 4.7](#) (bottom panels). This is partly a consequence of the per-waveform calculation of the delay times, allowing for a more precise determination of the pulse delay times because their distribution is not convoluted

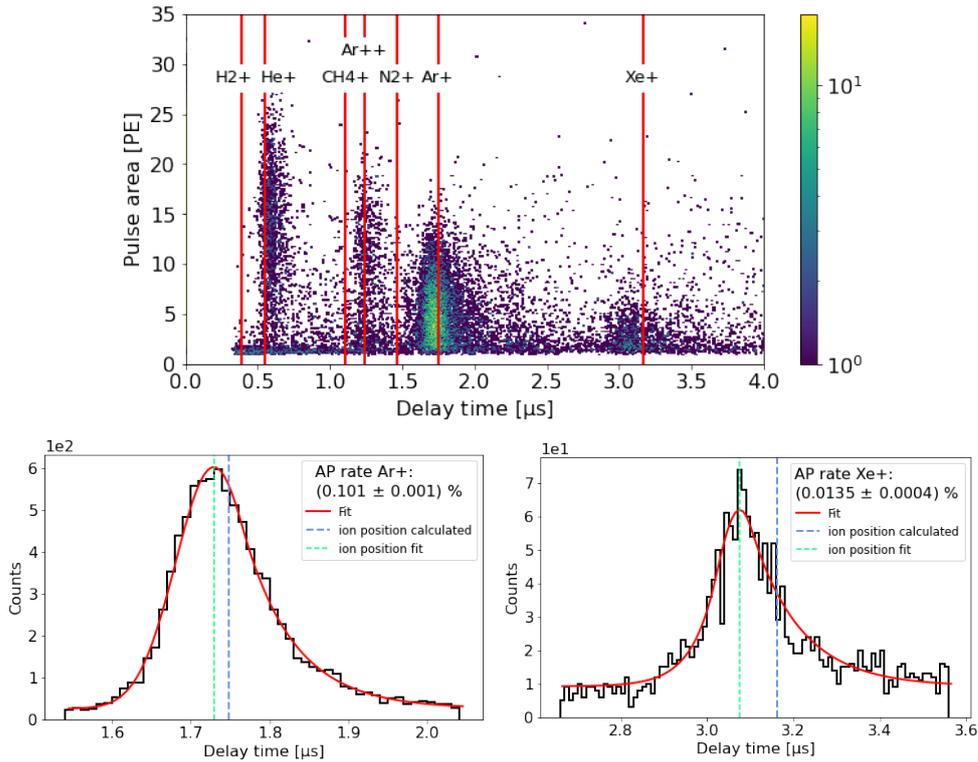


Figure 4.7.: (Top) Afterpulse delay time spectrum for one PMT channel, with SPE contribution removed by the height cut. Calculated afterpulse delay times are indicated by the vertical lines. (Bottom) Fits to the projection onto the delay time axis for Ar^+ (left) and Xe^+ in the respective ranges. The calculated and fitted (μ) delay times for the ions are shown by the blue and green dashed vertical lines, respectively.

with the variation of the LED pulse arrival times in different waveforms¹. The number of afterpulses N_i is then extracted by integrating the fit function with the best-fit parameters in the predefined ion ranges, after subtracting the constant offset. The total number of initial photoelectrons $N_{\text{PE,LED}}$ in a dataset is given by the sum of all LED pulse areas, selected by the condition of delay times equal to zero. Two examples of the fits to afterpulse peaks from Ar^+ and Xe^+ are shown in the bottom in Figure 4.7, resulting in afterpulse rates of $\sim 0.1\%$ and $\sim 0.01\%$, respectively. The stated errors in the afterpulse rates for a single measurement are of statistical nature, estimated by quadratically adding the Poissonian errors of $N_{\text{PE,LED}}$ and N_i . Systematic uncertainties are estimated by the variation of afterpulse rates over all measurements taken during a given time period, and are typically at the level of a few percent (relative).

4.3. Afterpulse monitoring in XENONnT

In this section, the results of the afterpulse monitoring in the XENONnT detector are presented, covering the commissioning phases of the detector and the first science run SR0.

The commissioning phase of the XENONnT detector began on March 25th, 2020, with the start of the evacuation of the inner cryostat housing the freshly installed TPC. The evacuation phase of the TPC lasted for more than three months, with the aim of removing all residual gases from the detector volume as well as from outgassing of the materials, mainly from the

¹Although it is still convoluted with the variation of the arrival times of the single photons within one single LED pulse

PTFE and other plastic components of the detector. During this time, the TPC was filled four times with nitrogen gas for about one week each time, in order to perform radon emanation measurements [18]. Gaseous xenon was then first introduced into the system including the detector volume on July 3rd, 2020, for the commissioning of the xenon circulation, purification, and distillation subsystems. Starting on August 18th, the xenon gas was slowly cooled down to the boiling point at about -95°C at 2 bar pressure until September 8th, when the filling of the TPC with liquid xenon began, which was completed on October 6th. During that time, the neutron veto was constructed around the TPC cryostat, followed by the filling of the water tank with ultra-pure water starting on December 15th. With the water tank (WT) completely filled on January 6th, 2021, the first dual-phase TPC data could be acquired. The period from then on was used for TPC commissioning, with SR0 starting on May 1st, 2021.

During these different periods, the PMT calibration system was commissioned as well, optimizing the LED pulser and DAQ settings for the PMT calibration measurements. The measurements and analysis procedures for the gains and the SPE acceptances were developed and are described in [25]. The HV values of the PMTs to be used in SR0 were optimized during the TPC commissioning period for having a gain of about 2×10^6 for each tube. In order to minimize micro light emission from the PMTs, an upper limit was initially set at 1.4 kV, that was increased to 1.5 kV for 27 PMTs to reach the target gain. The average of the HV values for the PMTs was ~ 1.35 kV, with a standard deviation of 0.1 kV, well below the region where micro-LE is typically observed (see also [subsection 3.2.3](#)). For science data taking, a time-dependent gain model was developed from the results of the weekly gain measurements, either averaging the gains over the whole SR0 period if the overall fluctuations are below a 2.5 % threshold, or using a filtering and smoothing algorithm for PMTs with time-dependent changes in the gains. These may occur due to slight temperature changes in the detector, aging effects due to high illumination periods during TPC calibration, or deliberate changes of supply HV, in which case the resulting discrete gain step is modeled accordingly. This time-dependent gain model is also used for the processing and analysis of the afterpulse data as already described in the previous [section 4.2](#).

4.3.1. Commissioning phases

During the first phase of the commissioning, the PMT calibration system was tested and optimized, and the first LED calibration runs were taken. Two PMTs in the bottom array (numbers 354 and 386) were found to not show any signal despite their HV divider bases showing the expected current draw. The best hypothesis for the cause of this issue is a broken signal cable connection on the TPC side, likely occurring during the transport of the TPC to the underground laboratory at LNGS, specifically when entering the WT through a door on the side of the tank. The height of the door is only slightly larger than the TPC itself (as per TPC design). During transport, the TPC was protected by being wrapped in Mylar foil, which was slightly touching the ground while moving the TPC through the WT door, potentially exerting stress on the signal cables in the bottom array. One further PMT in the bottom array (number 427) could not be ramped up to a working HV without tripping the HV supply in the first PMT tests, possibly due to excessive afterpulsing caused by an air leak into the tube.

Very first afterpulse measurements were taken during the early commissioning phase with the TPC under vacuum. While the settings for the LED pulser and the DAQ were not yet optimized at the time, nevertheless first conclusions could be drawn from the data. Unexpectedly, in six PMTs (numbers 19, 60, 74, 175, 209, 296), Xe^+ afterpulses were identified in their delay time spectra from measurements taken during the vacuum phase, before the introduction of any xenon gas into the XENONnT system. These tubes all belong to the set of reused PMTs from XENON1T, where they did not show any xenon leakage until the end of the experiment. However, the PMTs had not been re-tested for xenon leakage after the decommissioning of XENON1T, during which they went through a period of thermal stress while immersed in GXe during the warm-up phase when the detector was emptied. We conclude that the xenon leakage

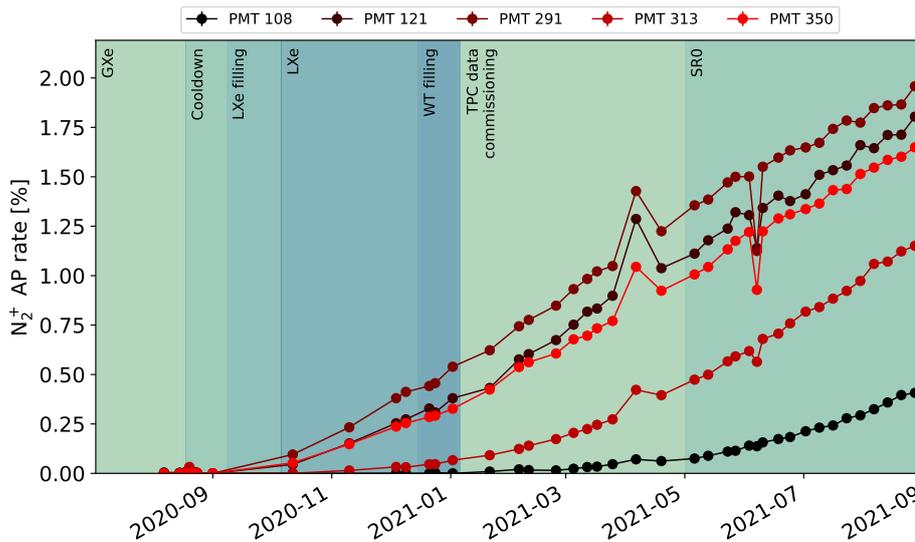


Figure 4.8.: Time evolution of the nitrogen afterpulse rates for several PMTs showing nitrogen contamination after the GXe cooldown and LXe filling during the commissioning phase. All but PMT 108 consequently had to be excluded from the SR0 data analysis or turned off (PMT 350).

likely happened during this period, before the tubes were installed in the XENONnT TPC. While three of these PMTs showed rather large xenon afterpulse rates between about 0.1 % and up to 3.0 %, the rates of the remaining ones were below 0.01 %. The xenon afterpulse rates of these six PMTs remained stable until the end of SR0, with no further leakage observed, consistent with the hypothesis that the leaks opened during the warm-up phase in XENON1T.

In addition, several PMTs started to show nitrogen afterpulses during the early commissioning period. Two PMTs showed high nitrogen afterpulse rates already in first measurements during the vacuum and nitrogen-filling phases, possibly having developed air leaks during transport or while they were immersed in nitrogen gas². Unexpectedly though, after the phases of GXe cooldown and filling the TPC with LXe, several more PMTs started to show nitrogen contamination as well, leading to a total of 18 PMTs with nitrogen contamination during SR0 (more details on that in the next section). The evolution of the N_2^+ afterpulse rates during the commissioning phase of five example PMTs of this set is shown in [Figure 4.8](#), with the different commissioning periods marked by the colored regions. PMTs 121, 291, and 350 showed nitrogen contamination already in the first measurement run taken after the LXe filling period, labelled ‘LXe’, while PMTs 108 and 313 started showing nitrogen contamination only after several months of operation in cryogenic conditions. The nitrogen afterpulse rates, as well as the total afterpulse rates, in these PMTs was observed to increase quickly over time. Because of the appearance and continuous increase of the nitrogen contamination during periods where the tubes were not in a nitrogen atmosphere, the source of the contamination is suspected to be internal, for example originating from outgassing from materials inside the PMTs. A possible explanation would be the presence of small cracks or micro-pockets in some materials, that had nitrogen or air trapped inside during the PMT production process. Similar to external leaks, these internal leaks may open under thermal stress during cooldown and LXe filling and release the previously trapped gas. Alternatively, as already briefly mentioned in the previous [chapter 3](#), in past PMT tests it was observed that the nitrogen afterpulse rates in PMTs with known external nitrogen leaks has a temperature dependence, where the nitrogen contamination acquired during tests in cold nitrogen *decreased* again when the PMTs were warmed up back to room temperature. This may be interpreted as an effect of the getter material strip inside

²It is not possible to distinguish between air or nitrogen from the afterpulse spectra, because other air components than nitrogen, mostly oxygen, are expected to be removed by the getter strip and not be visible.

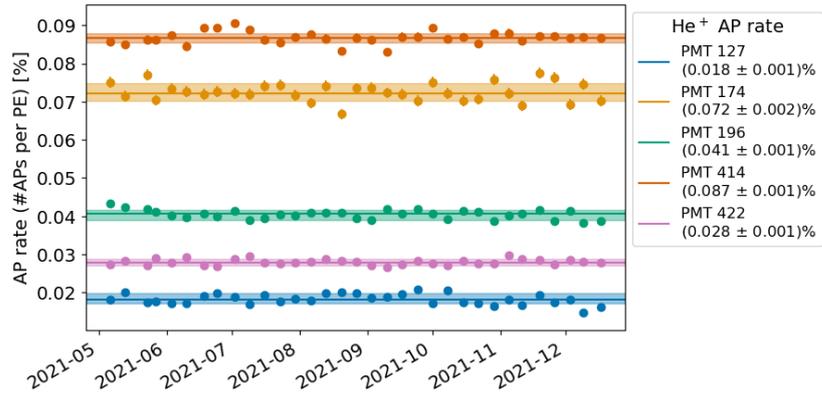


Figure 4.9.: Examples for He^+ afterpulse rate evolution over SR0. Solid lines show the medians, the shaded bands indicate the 68-interpercentile range. The uncertainties stated in the legend are the half-width of the bands.

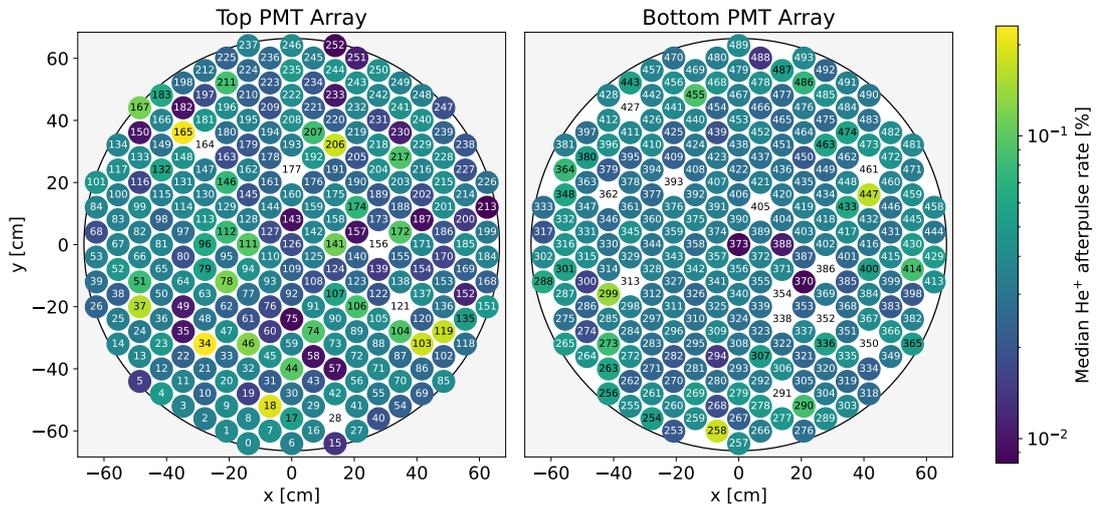


Figure 4.10.: Helium afterpulse rates for all PMTs in during SR0. The color scale indicates the median afterpulse rates over all measurements during the full duration of SR0 for each PMT.

the PMTs being less effective for nitrogen at cryogenic temperatures (despite being labeled a ‘cold getter’ by the producer), and potentially releasing previously bound nitrogen when cold. However, the latter point would contradict the observation that some PMTs started showing nitrogen contamination only after several months of operation in cryogenic conditions. Such behavior more closely resembles what is typically observed for external leaks, where leaks may open and close sometimes without apparent reason. This could therefore be expected to occur for internal leaks as well. Additionally, a general getter inefficiency for nitrogen at cryogenic temperatures is not expected to start suddenly. Most likely, a combination of both effects is the cause, where nitrogen is released from internal sources due to thermal stress and initially removed by the getter material, until it saturates and the nitrogen starts to accumulate inside the PMT volume.

According to feedback from the PMT producer on this issue, confirming after the fact that this is an issue they have observed as well, the internal nitrogen contamination is attributed to outgassing effects from the seal between PMT window and body.

4.3.2. First science run (SR0)

During SR0, the PMT monitoring was performed weekly, as described in the previous section. The afterpulse rates were monitored for all PMTs, and checked for any changes that could indicate problems with the PMTs or lead to a degradation of the data quality. We'll focus first on a quick summary of the results for the afterpulse rates that are expected to be stable over time, namely from helium and argon, and then discuss the afterpulse rates from vacuum degradation via xenon and nitrogen contamination.

Note that the afterpulse rates from hydrogen H_2^+ and methane CH_4^+ are zero for all PMTs from the cooldown period onwards, and during SR0. As previously observed in both PMT testing campaigns [23, 31], hydrogen and methane is seen in the afterpulse spectra at room temperature in all PMTs (see also Figure 4.2), but vanish completely when the tubes are cooled down to cryogenic temperatures. The original afterpulse rates have been observed to be recovered when the PMTs are warmed up again. This is thought to be an effect of adsorption at cold temperatures, either on the getter material strip in the tubes, which is expected to be efficient at cold temperatures (the producer calls it a 'cold getter'), or on other surfaces inside the PMTs.

Argon and helium afterpulse rates

Helium afterpulses are present in all PMTs due to the atoms' size being small enough that they easily diffuse into the tubes when stored in air. Once inside the TPC, the He^+ afterpulse rates are not expected to change over time. This is confirmed by the measured He^+ rates during SR0, which were monitored and found to be constant over time. Exemplary time evolutions for five PMTs of He^+ afterpulse rates during SR0 are shown in Figure 4.9, with the median rates and the central 68 % ranges shown by the solid lines and shaded regions, respectively. On average, the He^+ rates are stable over SR0 within a 3.1 % relative uncertainty. The per-PMT median He^+ afterpulse rates from all measurements taken during SR0 are shown in Figure 4.10, with an average rate over all PMTs of 0.039 %.

Argon afterpulses, on the other hand, are only present in a fraction of the PMTs, as was described already in the previous chapter 3, and are expected to be stable over time as well. About 30 % of all PMTs show a detectable argon afterpulse rate above the sensitivity threshold of the measurement method at $\sim 2 \times 10^{-3}$ %³. The average of the SR0-median rates for the PMTs with argon contamination, excluding PMTs where no argon afterpulse population is found, is 0.058 %, and were stable over time within a relative uncertainty of 4.3 %.

Vacuum degradation: Xenon and nitrogen

In SR0 data, a total of 18 PMTs showed xenon afterpulses, with one additional xenon-leaky PMT (number 177) turned off already during data taking due to showing frequent flashing and HV trips caused by the quickly increasing xenon afterpulse rate (at the time of being switched off it had xenon afterpulse rate of almost 5 %). Except for that one, only four PMTs showed the more severe issue of increasing xenon afterpulse rates, but fortunately with both low overall xenon AP rates and at rather slow growth rates, such that no further action was taken. The time evolution of the PMT with both the highest and the fastest growing Xe^+ afterpulse rate of these four tubes is shown as an example in Figure 4.11 (left), with an increase of the Xe^+ afterpulse rate of about 0.006 % per month and a maximum rate of 0.07 % at the end of SR0. The remaining 14 PMTs with xenon contamination at stable afterpulse rates during SR0 showed small rates below about 0.01 %, except for the three reused XENON1T tubes mentioned in the

³Note that this fraction is for a mixture of reused XENON1T and new PMTs, the latter showed a fraction of 60 % of PMTs showing argon afterpulses, however measured with higher sensitivity to lower afterpulse rates below 10^{-4} %. Adjusting for the threshold of the *in-situ* measurement in XENONnT also yields a fraction of 30 % in the testing data.

previous section that showed large xenon contamination already in the pre-xenon commissioning period. As described already, non-growing contamination is a common observation, with leak rates correlated with periods of thermal stress, which are not present in the stable operation conditions during science data taking. An example of a PMT with a constant afterpulse rates is shown in [Figure 4.11](#) (right), showing a median Xe^+ afterpulse rate of $(0.0060 \pm 0.0007) \%$ over SR0.

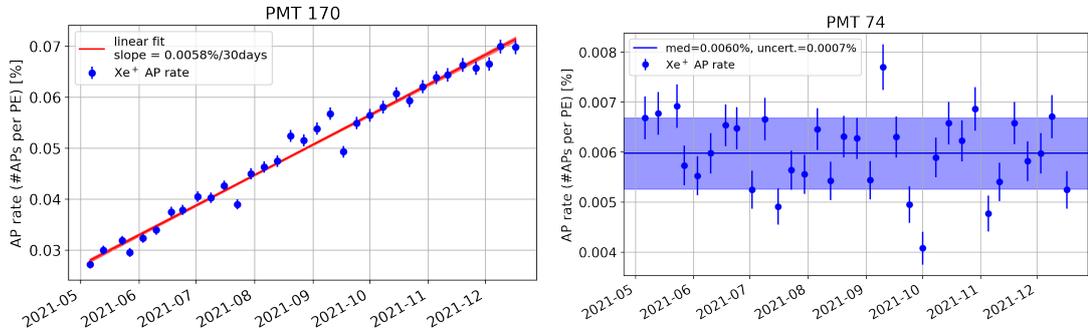


Figure 4.11.: Time evolution of Xe^+ afterpulse rates for two example PMTs during SR0. (Left) A PMT with a relatively small, but linearly growing afterpulse rate, with an absolute growth rate of about 0.006 % per month as determined from a linear fit to the evolution (red line). (Right) A PMT with stable xenon afterpulse rate, representing the large majority of the PMTs that show xenon contamination in SR0.

In contrast to xenon contamination, the afterpulse rates from tubes suffering internal nitrogen contamination were found to be continuously increasing over time for all affected PMTs. During SR0, a total of 12 PMTs showed nitrogen contamination, while six additional tubes were turned OFF already during the commissioning period due to becoming unstable and showing frequent flashing and tripping caused by the quickly degrading vacuum. The afterpulse spectrum of a PMT with large nitrogen contamination is shown in [Figure 4.12](#). Both the N_2^+ and the N_2^{++} peaks are clearly visible at the expected delay times, in addition to the typically observed He^+ and Ar^+ lines. The extracted N_2^+ afterpulse rate for this measurement is 1.7%. Note that due to the proximity of the afterpulse peaks of nitrogen and argon, the fits for the nitrogen afterpulse peaks are performed with a combined fit function, where all peaks between and including the N_2^{++} and Ar^+ peaks are fitted simultaneously. In addition to the N_2^+ peak, so-called secondary and even tertiary afterpulses are visible, indicated by the vertical dashed and dash-dotted lines, respectively. These peaks are caused by afterpulses of the nitrogen afterpulses, leading to an echo-like repetition of the full delay time spectrum that is additionally shifted by the delay time of the primary nitrogen afterpulses. The respective delay times are calculated by adding the delay times of the nitrogen afterpulses to the delay times of the secondary and tertiary afterpulses, in the combinations as stated in the line labels.

All PMTs that showed nitrogen contamination until the end of SR0 are from the set of newly acquired tubes, while the majority of the PMTs (15 out of 19) that showed xenon afterpulses are reused PMTs from XENON1T. Additionally, all PMTs that showed nitrogen contamination also belong to the set of PMTs with the largest argon afterpulse rates above $10^{-2} \%$. We conclude that the previously observed correlation between the argon amount in the PMTs and the later development of an external leak (see [subsection 3.2.4](#) and [Figure 3.7](#)) also holds for internal nitrogen contamination. This points again to a similar origin of the nitrogen contamination as previously for the xenon leaks, potentially being related to the PMT production process. Interestingly, while the failures due to xenon leakage has drastically decreased between the set of old and new PMTs, the internal nitrogen contamination only affects the new PMTs. One could draw the conclusion that the changes in the production process that led to the observed improvement in external leak tightness, may have in turn created the issue of internal nitrogen contamination.

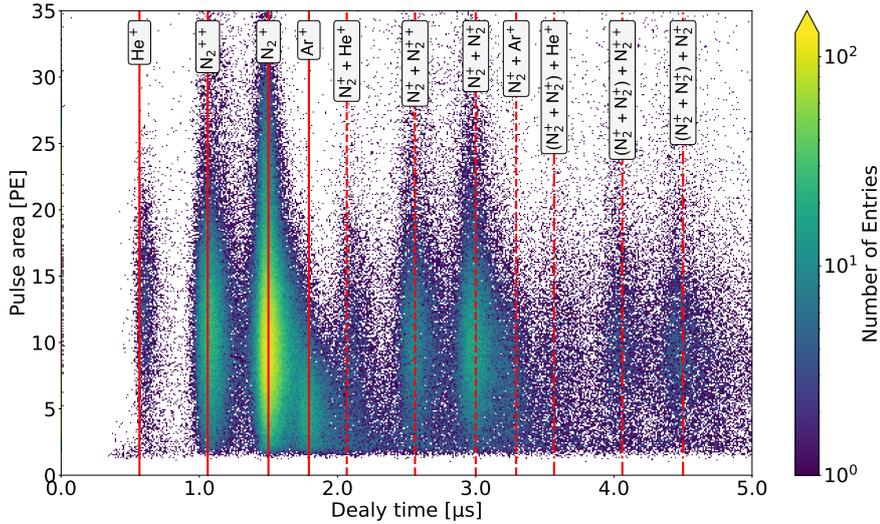


Figure 4.12.: Example of an afterpulse spectrum for a PMT showing nitrogen contamination. Solid vertical lines indicate the calculated delay times for the common afterpulse ions. The largest population at about $1.5 \mu\text{s}$ delay time is the nitrogen afterpulse peak, with the double-ionized N_2^{++} peak visible as well at about $1.1 \mu\text{s}$. Secondary (and tertiary) afterpulses of the nitrogen afterpulses are indicated by the dashed and dash-dotted vertical lines, respectively, calculated by adding the delay time combinations as stated in the label boxes.

Of the 316 new tubes in the arrays, only 5 tubes showed xenon penetration, while the remaining 15 were part of the 178 reused tubes from XENON1T, giving failure fractions of 1.3% and 8.4% in XENONnT, respectively. In total, 34 PMTs showed signs of vacuum degradation in XENONnT SR0, with three tubes showing both xenon and nitrogen contamination. This corresponds to a failure fraction of 6.7% of the total 494 tubes in the XENONnT PMT arrays. Overall, this amounts to a great improvement in PMT performance w.r.t. vacuum degradation compared to the $\sim 30\%$ failure rate in XENON1T (see also Table 3.2).

Selection of PMTs for analysis

For the data analysis in SR0, on top of the already switched off PMTs, seven PMTs were additionally excluded⁴ from the data analysis of SR0, mainly due to increasing afterpulse rates as well as other reasons, as summarized in Table 4.2. Similar to the already turned OFF PMTs, most tubes that were excluded from analysis suffered from severe nitrogen contamination. The decision on which PMTs to exclude was based on the observation of a negative impact on data quality caused by nitrogen afterpulses, in particular on the $^{83\text{m}}\text{Kr}$ data analysis used for the development of signal corrections and stability monitoring. Here the two subsequent S1 signals from the Kr decay via a short-lived metastable state with a half-life of 154 ns are used to select events. The PMTs with the largest nitrogen contamination were found to disproportionately cause an S1 area bias in the second, delayed S1 signal, originating from the merging of N_2^+ and N_2^{++} afterpulses from the first S1 signal into the second S1 signal. This effect was observed for PMTs with a total ion afterpulse rate above 1%, which was consequently set as a threshold for exclusion from the analysis, resulting in the exclusion of four PMTs (including PMTs 121, 291, 313 shown in Figure 4.8). Note that the PMT with the largest, but stable xenon afterpulse rate of about 3% was not excluded, because no negative impact on the Kr data was observed, likely because of the larger delay time of Xe^+ afterpulses of $\sim 3 \mu\text{s}$ making them less likely to merge with the second S1 signal compared to the nitrogen afterpulses.

One PMT showed intermittent light emission and was thus excluded from the analysis as well,

⁴This is done by setting their gains to zero in the processing and analysis software frameworks.

Issue	Xenon	Nitrogen	Light emission	Other
OFF PMTs	177	156, 164, 350, 352, 393, 427	362	354, 386 (no signal)
Excluded PMTs	-	121, 291, 313, 461	338	28 (unstable gain) 405 (noise)

Table 4.2.: List of PMTs that were OFF or excluded from the analysis during the XENONnT SR0. The large majority of PMT failure or exclusion is due to nitrogen contamination.

while two further PMTs were excluded due to showing unstable gain behavior and high noise levels. In total, 17 PMTs were either turned OFF or excluded from the analysis during SR0, corresponding to a working PMT fraction of 96.6 % out of the total of 494 PMTs in the arrays.

Performance of the categorization system and PMT testing campaign

With the results from the PMT monitoring during SR0, we can evaluate the performance of the penalty point system that was introduced in the previous chapter (subsection 3.3.1), used to select and distribute the PMTs in the arrays. The total penalty point distribution over the arrays is shown again in Figure 4.13, with the PMTs that suffered from performance issues in SR0 circled. All except one of the PMTs that suffered from vacuum degradation via xenon or nitrogen contamination (circled in red and magenta respectively), belong to the set of PMTs with the two highest penalty categories in the arrays, of either 1.5 or 2.0 total penalty points. No clusters of PMTs with performance issues are visible in the arrays, with the only two directly neighboring PMTs that are off or excluded being due to broken signal connections and light emission, where the former is not a predictable issue from the testing point of view. Similarly, none of the PMTs in the outermost ring of the top array showed any performance issues during SR0. This is a clear indication that the categorization system was successful in preventing clusters of failing PMTs as well as failures in the most affected regions in the PMT arrays. Additionally, the PMTs with the lowest penalty points were indeed confirmed to be the best-performing ones during SR0. With the penalty points awarded to the PMTs being mostly due to the argon amount in the tubes from the production process, this underlines the importance of a thorough PMT characterization and selection procedure before installation in the detector. Notably, the failure fraction of new PMTs belonging to the suspicious production batches was 13 %, compared to 6 % for the rest of the new PMTs, again confirming the high predictive power of the penalty point system.

4.4. Summary and Outlook

In this chapter, the results of the PMT afterpulse monitoring during commissioning and the first science run of the XENONnT experiment was presented. The LED calibration procedure for afterpulse measurements was optimized, and a *straxen* plugin was developed for processing and analyzing the afterpulse measurement data in the XENONnT analysis framework, along with a dedicated analysis routine for the extraction of the afterpulse rates. This allowed for the continuous monitoring of the afterpulse rates in all PMTs in the arrays during the first science run of the XENONnT detector, and the calculation of the afterpulse rates from all relevant ion species for all PMTs in the arrays. Based on the results of the afterpulse monitoring and an observed, correlated negative impact on data quality from PMTs with high nitrogen contamination, a selection of PMTs for exclusion from the data analysis could be made, based on the total ion afterpulse rate. This method has been successfully applied in predicting PMT failures also post-SR0, allowing for a preemptive exclusion of PMTs from data analysis before long-term data taking periods are started, based on projections for the development of nitrogen

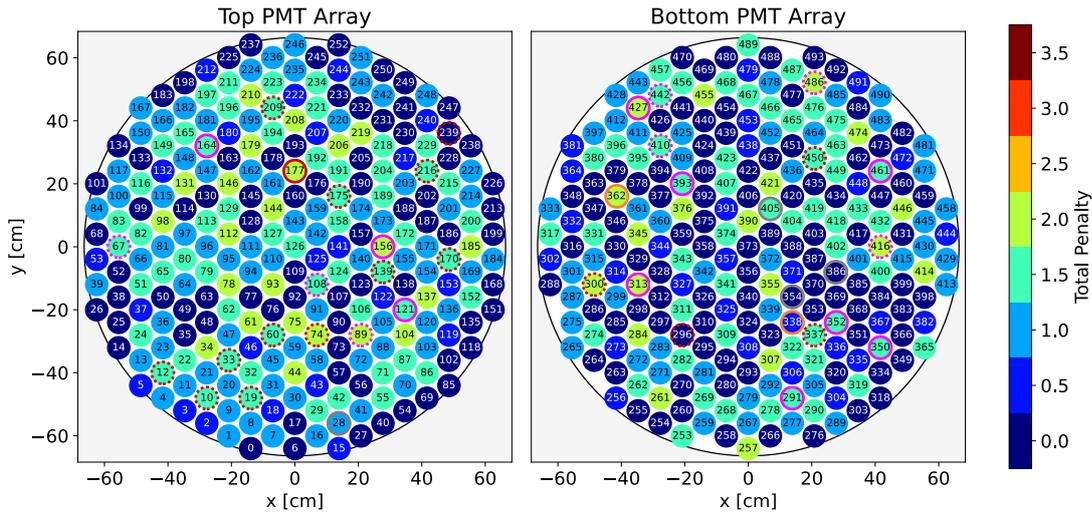


Figure 4.13.: Penalty point array plot with PMT failures due to what circled. Solid and dashed circled positions indicate PMTs that are OFF or excluded from analysis, respectively. The different colors indicate the failure reasons, xenon leaks (red), nitrogen contamination (magenta), light emission (orange), or other reasons (grey). Note that PMTs 337, 486, and 442 show both xenon and nitrogen contamination.

contamination. This avoids the necessity of computing- and time-intensive data reprocessing campaigns in case of PMT failures during data taking.

Overall, PMT performance in SR0 was very good, with only 17 PMTs turned off or excluded from the analysis, corresponding to a working PMT fraction of 96.6 % out of the total of 494 PMTs in the arrays. Additionally, the observed low fraction of new PMTs suffering from xenon penetration confirms that the new set of tubes has greatly improved in terms of vacuum tightness compared to the PMTs used in XENON1T. In turn, a new issue affecting PMT performance of the new tubes was identified, in the form of a continuously increasing internal nitrogen contamination. With 18 PMTs in total affected by this issue, ten of which had to be turned off or excluded, nitrogen contamination emerged as the more severe problem in XENONnT PMT performance compared to xenon leaks. Despite this, the overall PMT performance in XENONnT is greatly improved compared to XENON1T.

Furthermore, the general improvement in PMT performance already predicted from the testing campaign is confirmed by the good PMT performance in SR0. The quantitative categorization system introduced in the previous [chapter 3](#) used for PMT selection and distribution of the PMTs within in the arrays was proven to have a high predictive power for PMT failures, and on top of that was successful in preventing clusters of dead PMTs in the detector or in regions more impacted by PMT failures.

As of the time of writing of this thesis, the XENONnT detector and the PMTs have been operating almost continuously since the end of SR0, for about 3 years in total, with PMT monitoring still ongoing. So far, only three more PMTs were excluded, passing the total afterpulse rate threshold due to increasing nitrogen contamination, bringing the total up to 20. The long-term, overall stable performance of the PMTs over several years of operation is therefore confirmed, and is expected to continue for the remaining future of the experiment.

Chapter 5.

Modeling the Nuclear Recoil Response in XENONnT

A good understanding of the XENONnT detector response to NR and ER interactions is one of the key ingredients in the WIMP DM search, forming the basis of discrimination between signal and background events. As described in [chapter 2](#), NR and ER events form two distinct bands in the main analysis space of the observed S2 vs. S1 signal sizes. The shape of these bands is determined by the microphysics processes of the different interactions in LXe, as well as by the detector's signal reconstruction capabilities. If the shape of the bands can be fully modeled, meaning that we can produce probability density distributions (either analytically or approximately via simulations) for the NR and ER bands, we are able to predict where in the analysis space our WIMP DM events and our background events will lie, allowing us to infer the presence (or absence) of DM events in our data.

In order to characterize the shape of the bands, dedicated calibration datasets are taken with sources that produce NR and ER events with well-known energy spectra in the region of interest (ROI) of the WIMP search ($cS1$ in $[0, 100]$ PE and $\log_{10}(cS2/PE)$ in $[2.1, 4.1]$). The full detector response to these events is then modeled and fitted to the calibration data, taking into account all steps from the energy deposition in LXe to the final reconstructed S1 and S2 signals. For the WIMP search analysis with the first science data taken with XENONnT, SR0, the ER and NR bands were calibrated with a ^{220}Rn source and an AmBe-neutron source, respectively. While the ^{220}Rn source is gaseous and mixed homogeneously into the xenon volume, producing single scatter (SS) ER events with a flat energy spectrum in the ROI from the β -decay of the daughter isotope ^{212}Pb , the AmBe source is deployed externally and produces a broad energy spectrum of fast neutrons. Neutrons interact in the same way as potential WIMPs via elastic scattering with the xenon nuclei, resulting in NR events. In contrast to WIMPs, neutrons have relatively large cross-sections in LXe and therefore often scatter multiple times in the active detector volume. This results in multi-site event topologies with separate energy depositions, each leading to scintillation and ionization. Additionally, the AmBe source being external leads to a spatially non-uniform exposure of the TPC to the neutrons with most of the events being observed close to the TPC walls. The resulting effects on the reconstructed S1 and S2 signals in both multi-scatter (MS) neutron events and events that happen close to the walls must be taken into account in the modeling of the detector response to NR events, in order to correctly infer the NR emission model per energy deposit in LXe for SS, uniformly distributed WIMP NR events.

In this chapter, the characterization of the response of the XENONnT TPC to NR events from the AmBe source is described. First, a summary of the NR calibration data from the AmBe source and the event selection procedure is given in [section 5.1](#), focussing on aspects relevant for the modeling of detector reconstruction effects. The method used to model the NR calibration data is based on a Markov Chain Monte Carlo (MCMC) method, using fast toy simulations of the full response model that are matched to the data. This method is introduced

in [section 5.2](#). The full detector response model is then presented in [section 5.3](#), consisting of both the parameterization for the LXe microphysics processes in NR interactions, and the modeling of detector reconstruction effects with a focus on the effects of MS neutron events. The fit of the full detector response model to the NR calibration data and its results are described in [section 5.4](#), and finally, a discussion of the results and conclusionary remarks are given in [section 5.5](#).

The characterization of the NR response and the calibration fit to the data described in this chapter are included in ref. [\[51\]](#), of which the author of this thesis is a corresponding author.

5.1. Nuclear recoil calibration with AmBe neutrons

The NR events used for the NR band calibration in XENONnT SR0 are induced by elastic scatterers of fast neutrons from an Americium-Beryllium ($^{241}\text{AmBe}$) source. In the source, neutrons are produced by α -capture reactions on ^9Be , where in about 60% of captures a 4.44 MeV γ -ray is emitted from an excited state of ^{12}C . With the source being located inside the NV during calibration, the γ -rays can be detected in the NV, just as the ones from neutron capture on hydrogen (see [subsection 2.3.2](#)). The detection of these γ -rays in time coincidence with the S1 signal of the NR event in the TPC then allows for a clean, background-free NR event selection for the calibration fit.

The analysis of the AmBe neutron data and the NV was performed by Daniel Wenz at JGU Mainz, and is described in detail in [\[52\]](#). In this section, the AmBe source, the γ -ray coincidence selection with the NV, and the NR band data selections and acceptances thereof – which are inputs to the fitting framework – will be summarized. Hereby we focus on aspects observed in the data that are most relevant for the correct modeling of detector reconstruction effects, a necessity for inferring a correct NR emission model from the dataset.

5.1.1. AmBe source and the NV coincidence method

AmBe neutron sources consist of a homogeneous metal powder mixture of Americium dioxide $^{241}\text{AmO}_2$ and Beryllium Be. ^{241}Am is an α -particle emitter with a half-life of 433 years and a typical α energy of about 5.5 MeV (average of five different energies [\[53\]](#)). Neutrons are produced in the source via the α -capture reaction $^9\text{Be}(\alpha, n)^{12}\text{C}$ on ^9Be . In this reaction, a compound nucleus $^{13}\text{C}^*$ is formed which emits a fast neutron when decaying into ^{12}C , either directly into the ground state, labelled n_0 , or an excited state $^{12}\text{C}^*$, labelled n_i . The cross-sections for neutron production from α -capture reactions on ^9Be are shown in [Figure 5.1](#) (left), in dependence of the energy of the captured α -particle. For α energies of ^{241}Am , the accessible excited states with the largest branching ratios are the first and the second one, n_1 and n_2 , with levels of 4.44 MeV and 7.65 MeV respectively [\[53\]](#). While the second excited state of $^{12}\text{C}^*$ decays via internal conversion (IC), the first excited state decays promptly via γ -emission. The energy spectrum of the neutrons from an AmBe source is shown in [Figure 5.1](#) (right), with an average neutron energy of about 4.5 MeV. The shape of the neutron spectrum and the branching ratios for the decays into the different ^{12}C states n_i depend on the kinetic energy of the α -particle from ^{241}Am when it hits the ^9Be , which in turn depends on source parameters such as the ratio between Am and Be, packing density of the powder and the grain size of the $^{241}\text{AmO}_2$ itself.

The AmBe source used in the NR band calibration was already used for XENON100 and XENON1T, and was characterized in [\[55\]](#). It consists of a small capsule containing the compacted powder mixture with an ^{241}Am activity of ~ 3.7 MBq, giving a neutron flux of about 160 n/s. For the calibration, the source is inserted into the u-tube calibration system, with the tubes running around the outer TPC cryostat as seen in [Figure 2.5](#). With the u-tubes going through the NV water volume, the 4.44 MeV γ -ray emitted by the AmBe source from the first excited state n_1 can be detected in the neutron veto via its Compton scatters and resulting Cherenkov light

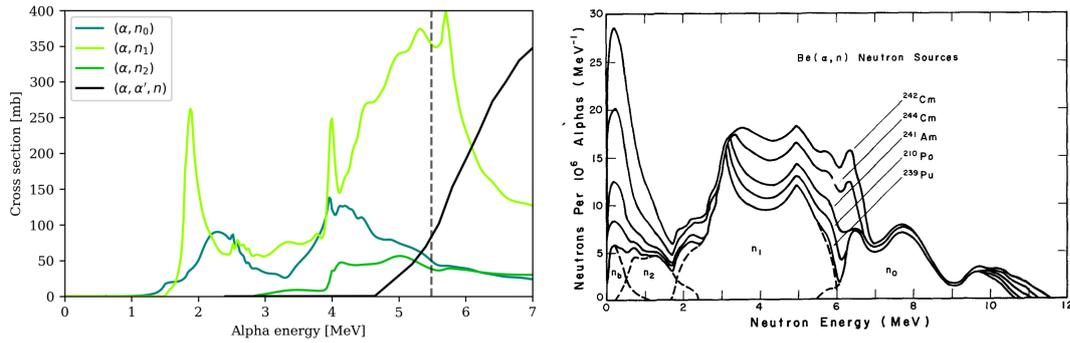


Figure 5.1.: (Left) Neutron production cross sections in the α -capture reaction ${}^9\text{Be}(\alpha, n){}^{12}\text{C}$, for the different states of ${}^{12}\text{C}$ after neutron emission (ground state n_0 , first and second excited states n_1, n_2). The vertical line indicates the maximum kinetic energy of α -particles from ${}^{241}\text{Am}$. Figure from [52]. (Right) Neutron energy spectra of various α -Be neutron sources, including for ${}^{241}\text{AmBe}$ as used for the NR calibration in XENONnT SR0. The contributions of neutrons from the different states of ${}^{12}\text{C}$ are indicated by the dashed lines. Figure from [54].

from the Compton electrons, just as for the 2.22 MeV γ -ray from neutron capture on hydrogen in the NV background-tagging mode. As is illustrated in Figure 5.2, neutron scatters in LXe in the TPC can be tagged via the close time coincidence between the S1 signal of the NR event and the signal from the 4.44 MeV γ -ray in the NV. This allows for an essentially background-free selection of true neutron events for the NR band calibration, making the modeling of the detector response to NRs significantly easier.

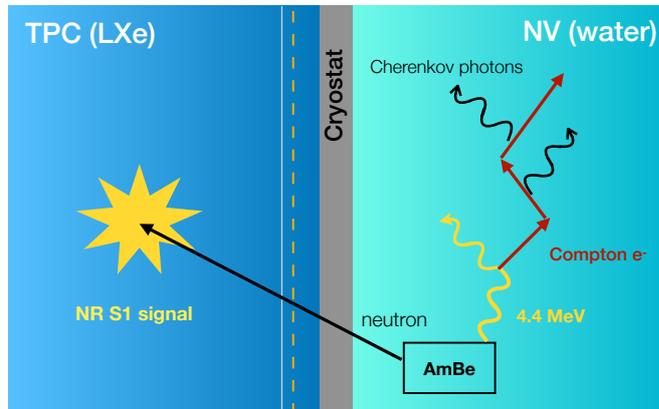


Figure 5.2.: Illustration of the working principle of neutron coincidence selection via the detection of Cherenkov light produced by the Compton scatters of 4.44 MeV γ -rays emitted simultaneously with neutrons from the AmBe neutron source, located inside the NV during NR band calibration of the TPC.

5.1.2. NR band data and event selection

The NR calibration data was taken at the beginning of SR0, from May 24th to June 6th, 2021. The source was positioned in three different locations, one in the top and two in the bottom u-tube (red and green in Figure 2.5, respectively). The position in the TPC coordinate system for the three locations are shown in Table 5.1, along with data acquisition time for each location. While the exposure with the source in location 2 is the largest, the resulting number of detected neutron events in the TPC is lower than for location 1 which has a lower exposure, because of the larger distance of the source to the outer cryostat wall (column Δr) and the resulting loss

Location	u-tube	x [cm]	y [cm]	z [cm]	radius [cm]	Δr [cm]	time [h]
1	top	1.09	-96.40	-10.21	96.41	12.98	23.73
2	bottom	15.15	-106.89	-95.59	107.96	24.54	31.06
3	bottom	98.77	-37.15	-91.01	105.53	22.11	12.37

Table 5.1.: The three positions of the AmBe source during NR calibration data taking, in the TPC coordinate system. Δr is the (radial) distance between the source and the outer cryostat wall, corresponding to the neutron’s path through water. The right-most column shows the time of data taking for each source position, the total time of data taking for the NR calibration was 67.16 hours. The three source locations are illustrated in [Figure 5.6](#).

of neutrons due to capture in the water. The resulting spatial distribution of selected neutron events in the TPC, as detailed in the following paragraphs, are shown in [Figure 5.6](#), along with the angular and z coordinates of the three source locations indicated by orange circles.

NV coincidence selection

For the selection of NR events in the TPC, first the 4.44 MeV γ -ray events in the NV are selected for the tagging, forming the ‘trigger’ events in this analysis. Details on the NV data analysis and selections can be found in [\[52\]](#). The Cherenkov light produced in the water by fast electrons from Compton scatters of high-energy γ -rays is detected by the NV PMTs (see [subsection 2.3.2](#)), with the signals recorded and processed with the same DAQ system and the *strax/straxen* framework as the TPC data, using a dedicated set of plugins for the NV data. The two detectors, TPC and NV, are synchronized in time, and the 4.44 MeV γ -rays are selected via a set of data quality cuts. These remove background events from accidental pile-up of PMT dark counts, PMT afterpulses, and events from β -emitters in the PMT windows that produce relatively small Cherenkov light events, as well as background events from radioactive decays in the detector materials. A minimum number of 4 NV PMTs contributing to an event is required, and the distribution of arrival times of the Cherenkov photons in the event must follow the expected exponential pulse shape given by optical properties of the water Cherenkov detector such as water transparency and surface reflectivity. The 4.44 MeV γ -ray energy signals are then identified as a peak-shaped distribution in the resulting signal area spectrum and selected accordingly, avoiding a peak at lower energies from the 2.22 MeV γ -ray from neutron capture on hydrogen. Using the selected 4.44 MeV γ -ray events in the NV as triggers, S1 signals in the TPC within a time window of -5 to +10 μ s around the NV signal are selected. The resulting time difference spectra between S1 and NV signals are shown in [Figure 5.3](#), for all TPC events in the window (black), for all TPC events in the NR band with $cS1 < 300$ PE (orange), and with additional data quality cuts applied (purple). The grey shaded region around zero time difference indicates the tight coincidence window used for the NR band data selection, defined as the 2.5% to 97.5% quantile range of the orange distribution, giving a range of [-57, 351] ns [\[52\]](#). With this definition, the number of accidental coincidence events between the two detectors in the NR dataset (purple) was determined, from the number of TPC events in the time range preceding the coincidence window. This results in 1.75 ± 0.55 accidental coincidence events [\[52\]](#) in the NR band calibration dataset with $cS1 < 300$ PE. With the NR band fit performed in the region of $cS1 < 100$ PE, the accidental coincidence rate is expected to be even lower, making the selection essentially background-free and removing the need for a background model in the fit.

The impact of the NV coincidence selection on the TPC data is illustrated in [Figure 5.4](#) (left), showing the NR band calibration data in the ROI defined for the band fit, $cS1 < 100$ PE and $2 \leq \log_{10}(cS2 [\text{PE}]) < 4$. The transparent histogram in the background shows all events recorded in the TPC during AmBe data taking without any selections, while the full-color

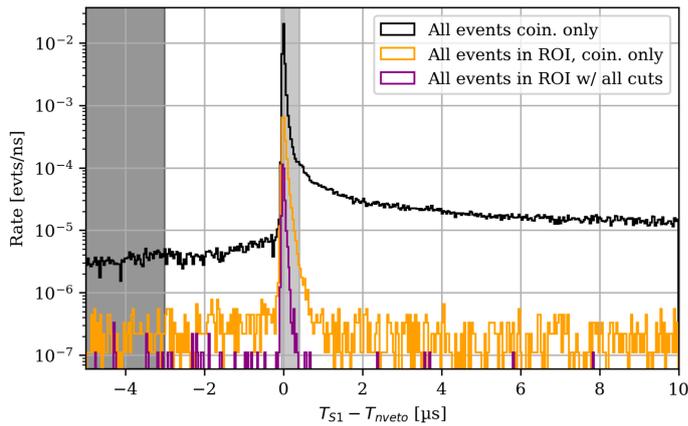


Figure 5.3.: Selection of NR events in the TPC in coincidence with the 4.44 MeV γ -ray from the decay of $^{12}\text{C}^*$ in the first excited state in the AmBe source. The spectra of the time difference between S1 signals recorded in the TPC and the selected 4.44 MeV γ -ray signals in the NV are shown for all events within a 15 μs window round the S1 signal (black), for the subset of events in the NR band ROI (orange), and with TPC data quality cuts applied (purple). The grey shaded region indicates the tight time coincidence window used to select NR events in the TPC in coincidence with the NV events. Figure from [52].

histogram shows the events after applying the NV coincidence selection. The NR band shape becomes immediately apparent in the data after only the NV coincidence selection. The panel on the right shows the same NV coincidence selection of events in the background, while the full-color histogram in the foreground shows the NR band data with additional data quality cuts applied (described in the following paragraph). This is the dataset we will use in the NR band fit.

Data quality cuts

Additional data quality selection requirements (called ‘cuts’) are applied to the events after the NV coincidence selection. Most cuts aim at removing unphysical events produced from a wrong pairing of S1s and S2s, as well as events that are not properly reconstructed due to e.g. pile-up with detector backgrounds. These cuts are based on properties of real S1 and S2 signals such as their expected pulse shapes and the distribution of the detected light on and between the PMT arrays. The cuts were mainly developed on calibration data from the $^{83\text{m}}\text{Kr}$ and ^{220}Rn sources, (both distribute uniformly in the TPC) and optimized towards having a high acceptance of good, physical events for the WIMP search. As briefly mentioned in the introduction of this chapter, neutrons often scatter multiple times in the active xenon volume within one event due to their short mean free path (m.f.p.) in xenon compared to the size of the detector, forming MS events. Here the word ‘scatter’ is used synonymously to ‘interaction with an associated energy deposit’. In the context of the NR band fit, however, we are interested in the LXe quanta emission model, i.e. the number of emitted scintillation photons and escaped ionization electrons per unit of deposited energy in *single* scatter (SS) interactions. Thus, the most important cuts for the NR calibration data are the ones that are able to remove MS events based on their event topology. Since the AmBe-neutrons are very fast (e.g. a neutron with kinetic energy of 5 MeV takes about 5 ns to travel its m.f.p. distance of about 15 cm), the prompt S1 signals from MS events will (almost) fully overlap and be detected as one merged S1 signal. The S2 signals may either be separated or merged, depending on the spatial separation of the SSs in the event in the z -direction. Therefore, the cuts that have a high MS rejection power are related to S2 properties, specifically the so-called S2 single scatter cut, and the S2 width cut, which are described in the following paragraphs. The so-called S1 pattern likelihood

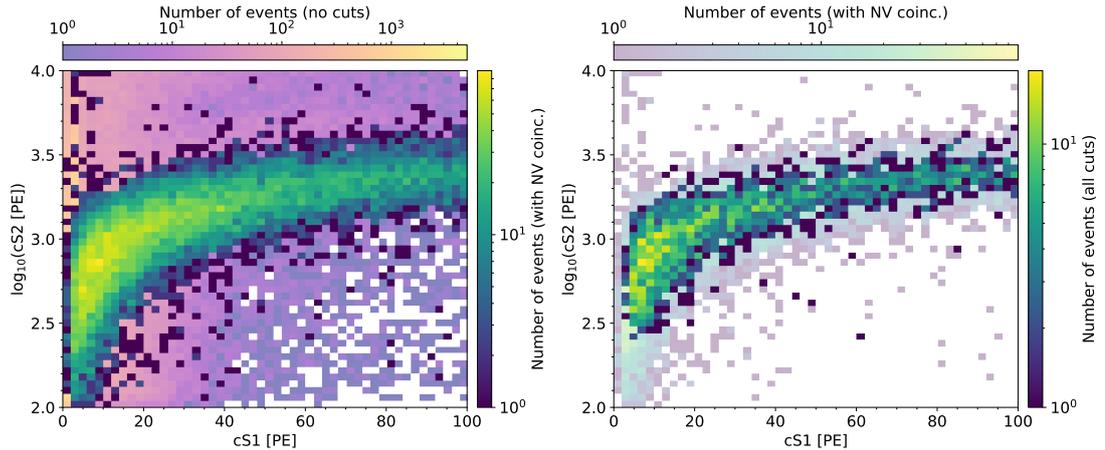


Figure 5.4.: NR band data in $cS1$ – $\log_{10}(cS2)$ space in the NR fit ROI, for various data selections. (Left) Impact on the dataset when applying the NV time coincidence selection. The transparent histogram (colorbar on the top) in the background shows all events without any cuts applied, the full-color histogram (colorbar on the right) shows all events in the tight coincidence window with the NV. The expected band structure of NRs is clearly visible. (Right) The transparent histogram (colorbar on the top) shows the same histogram of events after NV selection as in the left panel, while the full-color histogram (colorbar on the right) shows the events after all additional data quality selections as explained in the text.

cut, which evaluates the probability of a detected S1 PMT hit pattern in the top and bottom arrays to originate from a good event, has some MS rejection power as well, although less than the two S2-based cuts.

The S2 single scatter (S2SS) cut targets resolved MS events where two (or more) distinct S2 signals are found within the valid drift time range, i.e. all S2s must lie within the maximum drift time of the TPC following the (main) S1 of the event. The S2 with the largest area is labelled the mainS2, the one with the second-largest S2 is labelled altS2 (referring to ‘alternative’ S2). The cut is defined to remove events with an altS2 area above a certain threshold, that is dependent on the mainS2 area. The altS2 area threshold as a function of the mainS2 area is determined from ^{220}Rn calibration data where no MS events are present. For the NR band in the $cS1$ – $\log_{10}(cS2)$ space, the S2SS cut mainly removes events on the lower edge of the band. Resolved MS events with separate S2s are expected there, because they have larger merged S1s compared to SS events, while only the (single) scatter with the largest energy deposition in the event will form the mainS2, thus ‘shifting’ these events to the right. This effect is visible in [Figure 5.5](#) (left), showing the fraction of remaining events in the NR band when applying the S2SS cut after first applying all other data quality selections, including the NV coincidence.

The S2 width cut, designed to remove events with unphysically small or large S2 widths, is capable of removing unresolved MS events, where two S2s are overlapping enough to be reconstructed as one S2, but do not fully overlap such that the width of the merged S2 is larger than expected for its size and depth. In the $cS1$ – $\log_{10}(cS2)$ space of the NR band, this cut removes mostly events from the upper edge of the NR band, as seen in [Figure 5.5](#) (right). Several processes in the NR microphysics of quanta production in LXe depend on energy non-linearly, as will be detailed in [section 5.3](#), such that the sum of produced quanta q (either scintillation photon or escape electron) for two separate energy depositions is not equal to the number of produced quanta from an energy deposition with the sum of the energies, i.e. $q(E_1) + q(E_2) \neq q(E_1 + E_2)$. With an increasing (decreasing) photon (charge) yield for increasing deposited energies in the relevant energy range, see e.g. [Figure 5.19](#), this results in a smaller S1 and larger S2 for merged MS events compared to a SS event with an energy deposit corresponding to the sum of the MS energies.

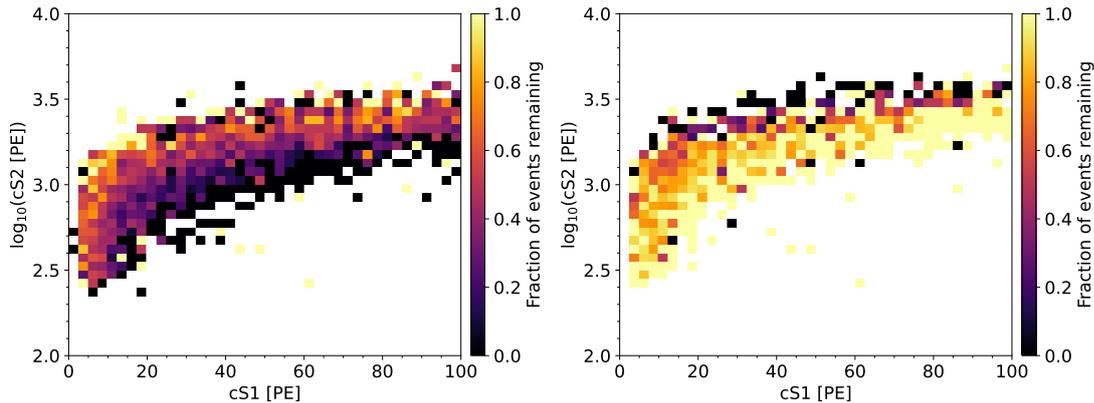


Figure 5.5.: Fraction of remaining events after applying the S2 single scatter cut (left) and the S2 width cut (right) in addition to all other cuts. The two cuts have a high rejection power for resolved and unresolved MS events, respectively, causing the observed impacts on the NR band shape, illustrating the need for a correct modeling of these event topologies in the NR band fit.

The acceptances of the data quality selections that are used for the NR calibration data, i.e. the efficiencies with which good events are retained, generally depend on the S1 and S2 signal sizes. Thus, they may alter the shape of the NR band and therefore need to be accounted for in the NR response fit¹. The acceptances of most of the data quality selections were determined in [52] using the so-called ‘N-1 method’, where the acceptance of each cut is determined by the fraction of events remaining in the NR band when applying the cut, with all other ‘N-1’ cuts applied before. This method assumes that the cuts are not correlated between each other, and that the remaining events after all other cuts are applied form a representative sample of good events. While this is generally a reasonable assumption, for the cuts that remove MS events in a data sample with a significant fraction of MS events that may or may not be resolved, the definition of a good event is not as clear-cut as for a data sample where only SS events are expected. For the NR band fit, the acceptances of the MS rejecting cuts were therefore evaluated by simulation methods, using the *full-chain simulation framework* of XENONnT. Further details on this procedure will be explained in section 5.3.

Fiducial volume and charge-insensitive region

While the definition of data quality selections were kept the same between the NR band data and the WIMP search data, the fiducial volume for the NR calibration data was enlarged compared to the one used in the WIMP analysis. This was necessary because the spatial distribution of the NR events close to the edges of the TPC, due to the short m.f.p. of the AmBe neutrons in LXe. The spatial NR event distribution can be seen in Figure 5.6, both in the xy -plane corresponding to the ‘top view’ of the TPC (left), and in the r^2 - z space (right). Most of the events are located at high radii and at the top of the TPC, outside the boundaries of the WIMP FV (indicated by the red dashed lines). The WIMP FV was optimized by maximizing the signal-to-background ratio, i.e. maximizing the total fiducial mass while avoiding backgrounds from detector materials (radiogenic neutrons and low energy γ -rays from materials, and surface events, see also chapter 6), while excluding regions in the detector where our signal reconstruction is not well understood². The AmBe FV, in contrast, was optimized

¹In contrast, data selections that cut on the exposure, such as the NV coincidence, the FV, but also DAQ busy vetoes and other hardware-based selections, are not expected to alter the NR band shape and are therefore ignored in the band fitting. Note that the NR band fit does not aim at matching the absolute neutron rate with the expected rate from the source.

²In particular the maximum radius was set to 61.35 cm based on the surface background model, which increases with larger radius. It was found that going to larger radii does not increase sensitivity to WIMPs sufficiently compared to the loss of robustness against potential mismodeling in the background model.

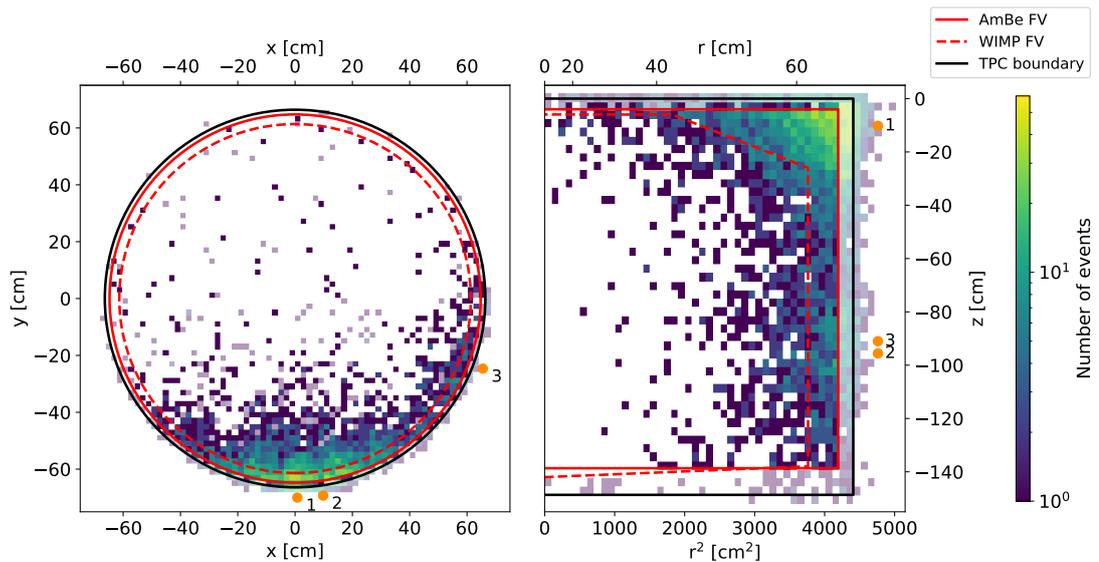


Figure 5.6.: Spatial distribution of the selected NR calibration events. Black lines indicate the physical TPC boundary, the red solid line shows the FV chosen for the AmBe-neutron calibration data (AmBe FV). The dashed red line corresponds the FV used in the WIMP analysis (WIMP FV). The positions of the source locations 1,2, and 3 are indicated by the orange circles, note that the radial distance of the source locations is not to scale (correct are the angles in the left plot, and the z -position on the right). (Left) xy -distribution of NR events. The source position with the most statistics is Position 1. The dashed red line here corresponds to maximum radius of the WIMP FV of 61.35 cm. (Right) z vs. r^2 distribution.

mainly on the latter requirement, by increasing statistics in the calibration data set as much as possible while avoiding going too close to the TPC edges. The boundaries follow a simple cylinder shape and are shown in Figure 5.6 by the red solid lines. The maximum radius of 64.75 cm for the AmBe FV was determined to avoid events suffering from charge loss. Due to the electric field configuration in XENONnT (and exacerbated by the lower-than-designed electrode voltages reached, as well as PTFE charge-up effects), the field lines of the electric drift field defining the drift path of liberated electrons end on the PTFE wall for a small region at high radii, and especially at the bottom of the TPC. Thus, S2 electrons originating in these regions are either partially or fully lost. These regions are dubbed (partially) charge-insensitive volume, pCIV and CIV, respectively. A map of the survival probability for electrons created at a certain position in the TPC was derived from electric field simulations of the XENONnT field cage configuration [14, 56], shown in Figure 5.7. For the definition of the maximum radius, events suffering from charge loss were found by selecting the vertical ‘shadow’ population below the ER band in cS1-cS2 space formed mainly by the Compton continuum of high energy γ -rays in the AmBe calibration data (both the 4.44 MeV from $^{12}\text{C}^*$ and the 2.22 MeV γ -ray from $^2\text{H}^*$ from neutron capture in water). The events in this population have reduced cS2 areas down to the limit set by the S2 threshold in the WIMP analysis of 200 PE, and follow a spatial distribution clustered at the wall and the bottom of the TPC, as expected for events in the pCIV. The maximum radius was then found by maximizing the difference between the fraction of removed charge loss events in the ‘ER shadow’ selection and the fraction of remaining events in the NR band. The lower and upper boundaries in z were set to -138.65 cm (10 cm above the cathode electrode) and -4 cm (below the gate electrode), respectively, avoiding charge loss events caused by the cathode mesh at the bottom and gas events formed between anode and top screening mesh at the top.

After applying the AmBe FV to the NR data, however, several outlier events below the NR band remain (see Figure 5.4). For outlier events with a cS1 above ~ 20 PE, these are expected to

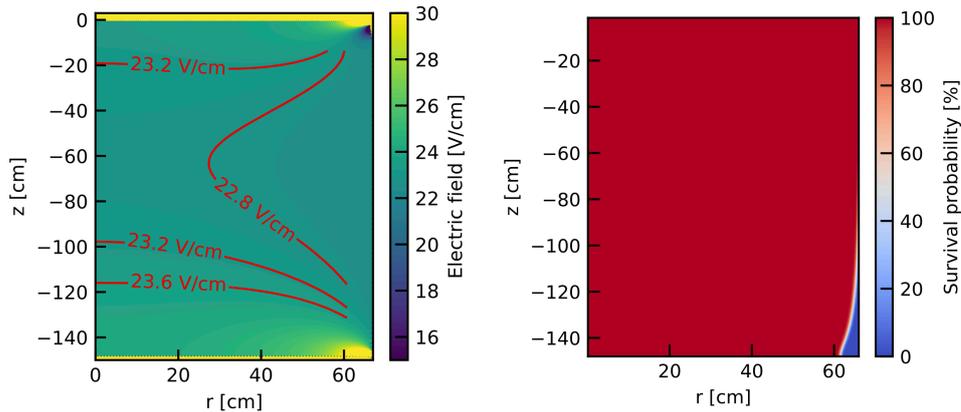


Figure 5.7.: The XENONnT electric field map (left) and the charge survival probability map (right), showing the survival probability for electrons created at each position. The region in the lower right with zero survival probability corresponds to the charge insensitive volume (CIV), while the smaller transition region with $<100\%$ survival probability is the partially charge insensitive volume (pCIV). The charge survival probability map is an input to the fitting framework. Figure from [56].

be caused by charge loss effects in the pCIV, as most of these events are reconstructed close to the pCIV. The outliers at smaller cS1s however do not follow the expected spatial distribution of charge loss events, but closely follow the general distribution of all other NR events. Their origin is hypothesized to be a combination of unresolved MS events and the effect of the (p)CIV, where a neutron could e.g. scatter first in the (p)CIV close to the wall, and then scatter again further inward the TPC. With the S2 from scatters in the (p)CIV significantly reduced in size, but the S1s being merged, this may lead to an untypically small cS2 compared to the cS1. Similarly, an alternative S2 signal may be reduced as well, making the event fall below S2SS cut threshold. Especially for events near the detector threshold, unresolved MS events with merged S2s as well as unrejected MS events can accumulate, due to being more likely to pass the S1 detection threshold. In order to reduce the impact of these events on the NR band fit, the S2 threshold was increased to 250 PE for the NR calibration dataset, removing about half of the outlier events. As we will see in the NR band fit results (section 5.4), the remaining outlier events below the band were modeled by accounting for NR MS effects as well as the charge loss effect in the simulations used in the fitting process.

With all selections applied, i.e. the NV coincidence, data quality selections, the AmBe FV cut, and the increased S2 threshold of 250 PE, 1987 events remain in the AmBe fit ROI for $cS1 < 100$ PE. Of those, 639 events lie within the tighter WIMP FV, corresponding to 32.2% of the final NR calibration dataset.

5.2. Introduction to the basics of the fitting tool BBF

The ER and NR band shapes in the $cS1$ - $\log_{10}(cS2)$ space are crucial for the WIMP search analysis, forming the basis for the background and signal models. In order to characterize the shapes of the bands, we want to fit the full detector response model (i.e. emission models for ER and NR interactions, and the detector reconstruction model) to ER and NR band calibration data. As we will see in section 5.3 for the NR case, these models are rather complex due to the necessity of many parameters and effects that need to be accounted for. The high dimensionality of the parameter space often leads to many local minima that are difficult to handle using standard minimization methods, thus making the fitting process non-trivial. To handle this challenge, a band fitting framework called BBF, an abbreviation for ‘Bayesian Band Fitting’, is used to fit the ER and NR band calibration data in XENONnT SR0. It was originally

developed for and used in the WIMP search analyses in XENON1T [57–59], and is described in detail in [60–62]. It utilizes a Markov Chain Monte Carlo (MCMC) algorithm to sample the posterior probability density function (PDF) of the model parameters by means of matching calibration data to simulations. For this work, it has been adapted for the use in the XENONnT analysis, and extended in particular to be able to handle MS neutron event topologies in the simulation. In this section the concepts behind the BBF tool are introduced, and the main steps of the fitting process are explained.

Bayesian inference using MCMC

As the name of the tool already hints at, BBF uses a Bayesian approach to infer the model parameter distributions that match the calibration data. This is essentially the same as finding the probability density function (PDF) for the model parameters $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$ given the calibration dataset \vec{x}_{cal} , which corresponds to the Bayesian *posterior probability density* given by

$$p(\vec{\theta} | \vec{x}_{\text{cal}}) \propto p(\vec{x}_{\text{cal}} | \vec{\theta}) \cdot p(\vec{\theta}). \quad (5.1)$$

The first term on the right side is usually called the *likelihood* $\mathcal{L}(\vec{x}_{\text{cal}} | \vec{\theta})$ and is the (frequentist) probability of getting the calibration dataset $\vec{x}_{\text{cal}} = (x_1, x_2, \dots, x_N)$ assuming that the set of parameters $\vec{\theta}$ is true. The second term is the *prior probability* distribution of the parameters, encoding previous knowledge on values of the parameters. In our case, each parameter has an independent *prior* term, that may come from e.g. literature measurements of parameters in the micro-physics processes of LXe quanta production and emission, or from dedicated studies of detector effects using other calibration sources such as $^{83\text{m}}\text{Kr}$.

Finding the posterior probability function then relies on varying all parameters $\vec{\theta}$, which becomes computationally very expensive when the dimensionality of $\vec{\theta}$ is high, i.e. when the number of parameters n is large, as is the case here. This problem is solved in BBF by the use of Markov Chain Monte Carlos (MCMC), a method to sample from multidimensional distributions with very high efficiency. In MCMCs, an ensemble of so-called *walkers* iteratively steps through the n -dimensional parameter space, guided by the *target density* which in our case is the posterior distribution $p(\vec{\theta} | \vec{x}_{\text{cal}})$. At each iteration, each walker has a ‘position’ in the parameter space, corresponding to a single-valued vector or set of parameters, and based on a sampling algorithm proposes a new position where to step next. In comparing the posteriors between the proposed new position and the current position, a decision is made to either ‘walk’ to the new position or stay at the current one. The walker positions in every iteration are logged in the MCMC chain, such that after a sufficient number of iterations the positions of the walker ensemble form a representative set of samples from the posterior. The marginalized posterior for each parameter θ_i is then easily obtained by projecting the samples onto the corresponding dimension. BBF uses the affine-invariant sampling algorithm proposed in [63] and implemented in [64, 65], where each walkers’ decision on where to move next is based on the position of all other walkers – hence the reference to the walkers as an ‘ensemble’.

MCMCMC – Likelihood evaluation via MC simulations

For the fitting process via the MCMC algorithm, the likelihood $\mathcal{L}(\vec{x}_{\text{cal}} | \vec{\theta})$ needs to be evaluated at every iteration for every walker parameter set. In BBF, this is done by comparing the calibration data to Monte Carlo (MC) simulations of the full detector response. In every iteration of the MCMC chain, a MC simulation of the full detector response is performed for every walker, using the walker’s new proposed positions in the parameter space. The simulated and data events are then filled into 2D histograms in cS1-cS2 space, and compared via the binned Poisson likelihood,

$$\mathcal{L} = \prod_i \mathcal{L}_i = \prod_i \frac{\hat{n}_i^{n_i} \cdot e^{-\hat{n}_i}}{n_i!}. \quad (5.2)$$

Here n_i and \hat{n}_i denote the number of observed (data) and expected (MC) events in bin i , respectively. As is common, we use the natural logarithm of the likelihood, such that

$$\ln \mathcal{L} = \sum_i n_i \ln(\hat{n}_i) - \hat{n}_i - \ln(n_i!) \quad (5.3)$$

where the last term is a constant since it only depends on the data, and is thus dropped for the fitting process. The likelihood is then multiplied by the prior probability of the parameter set, giving the posterior value to be logged in the MCMC chain for each walker.

In order to fully model the band shapes, all physical processes that occur between particle emission from a calibration source up to the observed and corrected S1 and S2 signals need to be taken into account, including particle propagation through the detector, quanta production via the LXe emission models, scintillation photon and electron drift propagation on a per-quantum level, photosensor detection efficiencies, as well as signal recording, processing, and reconstruction. In XENONnT, these processes can be simulated using the full-chain MC simulation framework [66], which includes the GEANT4 toolkit [67] for source particle propagation as well as scintillation photon propagation, the custom-built EPIX tool [68] for micro-clustering and quanta generation, and the WFSim [69] tool that generates waveforms that can then be processed by the XENONnT data processing framework strax/straxen [48, 49]. Additionally, the electron drift propagation is determined from simulation of the electric drift field in XENONnT [14, 56] using the COMSOL Multiphysics® tool[46]. However, the simulations used in the fitting process need to be very fast in order to be able to perform the fit in a reasonable time – in the order of at most a few seconds per MCMC iteration for all, typically $\mathcal{O}(100)$, walkers simultaneously. In BBF this is solved by using simplified, fast toyMC simulations of the full detector response, which are based on input ‘curves’ (1D) and maps (nD) that encode the above-mentioned effects. These are provided from dedicated detector studies with other datasets, such as for example $^{83\text{m}}\text{Kr}$ calibration data, as well as from studies using the full XENONnT MC simulation framework. In the fast toyMC simulation process, each walker random-samples from these curves or maps. This way, uncertainties can be accounted

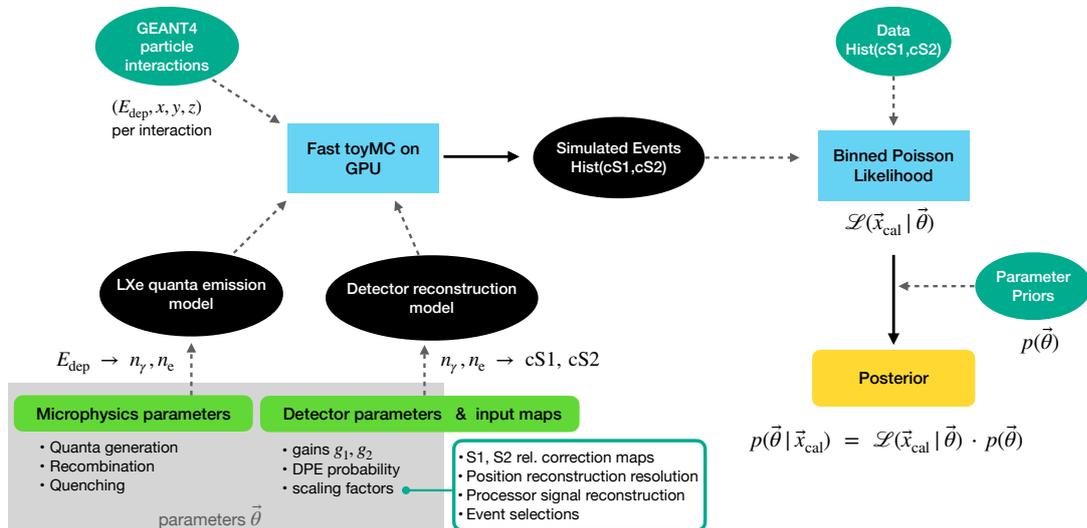


Figure 5.8.: Illustration of the workflow for posterior evaluation of a single walker in one MCMC iteration in the band fitting tool BBF. External inputs are indicated in sea-green, consisting of the histogrammed calibration dataset, a set of simulated calibration source events produced using GEANT4 and micro-clustering, and the prior probability distributions of the parameters. The other external inputs are the maps needed for the effective detector reconstruction model, which are varied using scaling factor parameters.

for by including scaling parameters for the inputs in the fit model. Additionally, the fast toyMC simulations are run on graphical processing units (GPUs), benefitting from increased simulation speed and parallelization, where all walkers' toyMCs per iteration can be performed simultaneously.

The BBF workflow for the posterior evaluation of a single walker in one iteration of the MCMC chain is illustrated in [Figure 5.8](#). Starting from a set of GEANT4-simulated calibration source events, giving deposited energy E_{dep} and position (x, y, z) for each particle interaction (also called scatter) in the events, the fast toyMC simulation is performed according to the full detector response model, consisting of the LXe emission model and the detector reconstruction model. The former describes the production of the observable quanta from an energy deposit in LXe, i.e. n_γ scintillation photons and n_e ionization electrons, and is fully parametric. The latter describes the conversion of the produced quanta into the detected and corrected cS1 and cS2 signals in the TPC, and is simulated by a mixture of parametric and effective modeling via input maps including scaling parameters as explained above. The parameter set of the walker $\vec{\theta}$ consists of the parameters of the full detector response model as indicated by the grey shaded box. The parameterization of the full model will be described in detail in the next [section 5.3](#). After the fast toyMC simulation, the produced events are histogrammed and compared to the data histogram, giving the likelihood of the walker's parameter set, which is then multiplied by each parameter's prior, giving the posterior probability of that particular parameter set.

After having performed the fits to the calibration data, a set of best-fit parameters is extracted from the marginal posterior distributions of the parameters. The same fast toyMCs as in the fitting processes are then used with the set of best-fit parameters to produce finely binned histograms of ER and NR background sources as well as the WIMP signals. These finely-binned histograms represent the probability density functions (PDFs) of each component in the $\text{cS1}-\log_{10}(\text{cS2})$ analysis space for the WIMP search (along with the radial component of the spatial distribution of the components) and are called 'templates'. These form bases of the statistical inference of the WIMP search data. Comparing the sum of all templates to the events in the WIMP search dataset, using an extended unbinned profile likelihood analysis, allows us to assign a 'probability'³ for each events' origin from one of the component templates, giving good sensitivity for a potential WIMP detection or setting upper limits on WIMP interaction cross sections.

5.3. Modeling the NR response to AmBe calibration data

In order to characterize the NR response with AmBe-neutron calibration data, we need to model the full detector response of the TPC taking into account all effects that have an impact on the observed S1 and S2 signals. This includes both the microphysics of the energy deposition in LXe, i.e. the production of scintillation photons and ionization electrons, and the detector physics, such as the propagation of the produced quanta and subsequent detection including efficiencies, as well as signal processing and reconstruction effects. The event topology of the AmBe-neutron calibration data is taken into account as well, with a significant fraction of (unresolved) MS events in the data.

As shown in [Figure 5.8](#), we can separate the full detector response model into two parts: First the LXe emission model for NRs, which describes the production of the observable quanta from an energy deposit in LXe, i.e. scintillation photons and ionization electrons. The second part is the detector reconstruction model, describing the conversion of the produced quanta into the detected S1 and S2 signals in the TPC.

³It should be noted that this is technically not true if a WIMP signal component is included in such a likelihood ratio analysis, since we don't know if a WIMP even exists.

5.3.1. NR quanta emission model in LXe

The processes of excitation and ionization of xenon atoms induced by a nuclear recoil can in principle be calculated if the interaction cross sections are known. These however have not been measured or calculated for xenon, such that these processes have to be modeled semi-empirically. Here, we use the model parameterization described in [70], where a global fit to several datasets from literature measurements of photon and charge yields was performed using the NEST (Noble Element Simulation Technique) framework [71]. In the following we refer to this model as NESTv1. While more recent implementations of the NEST framework are available, that use a different parameterization of the NR emission model (which we call the NESTv2 model in the following), it was decided to stay with the NESTv1 model for XENONnT SR0, as it was already successfully used for the NR modeling in the XENON1T WIMP search analysis [57–59]. Additionally, at the time the NR band fit was performed the NESTv2 model and in particular the parameter values was not published (a review is now available, see [72]), so that no prior information on the parameters was available for the fit.

All parameters in the NR LXe emission model are summarized in Table 5.2, also showing the best-fit values from [70]. Some of these are used as priors in the fit to the XENONnT calibration data, as will be explained in section 5.4.

An illustration of the processes modeled in the NR quanta generation is shown in Figure 5.9. Starting from a given deposited energy E , a total number of quanta n'_q is produced, following a normal distribution

$$n'_q \sim \text{Norm} \left(\frac{E}{W}, \sqrt{f \frac{E}{W}} \right), \quad (5.4)$$

where W is the mean work function, corresponding to the average energy needed to produce one quantum, either an exciton or an ionization electron. For LXe, the commonly adopted value is 13.7 eV [73]. Because the underlying processes of energy deposition in a material are not purely statistical, with e.g. ionization levels of the involved atoms playing a role, the uncertainty of the number of produced quanta is smaller than expected from statistical effects only. This is accounted for by using a Fano factor f [74], decreasing the variance in the number of produced quanta. The fano factor for LXe is determined to be about 0.06 [75, 76].

Because recoiling xenon nuclei have a large nuclear stopping power in LXe, they lose a significant fraction of their energy to atomic motion (heat), which can not be detected in a dual-phase TPC. This loss of detectable quanta is modeled following Lindhard theory [77]. The number of detectable quanta n_q follows a binomial distribution with the ‘success’ probability given by the Lindhard quenching factor L ,

$$n_q \sim \text{Binom} (n'_q, L). \quad (5.5)$$

The quenching factor L , defined as the fraction of detectable quanta, is energy-dependent and parametrized via

$$L = \frac{\kappa g(\epsilon)}{1 + \kappa g(\epsilon)} \quad (5.6)$$

where $g(\epsilon)$ is a function of the dimensionless representation of deposited energy ϵ ,

$$g(\epsilon) = 3 \epsilon^{0.15} + 0.7 \epsilon^{0.6} + \epsilon, \quad (5.7)$$

$$\epsilon = 11.5 \left(\frac{E}{\text{keV}} \right) Z^{-7/3} \quad (5.8)$$

with the atomic number (nuclear charge number) $Z = 54$ for xenon. The parameter κ , sometimes also called the Lindhard-k, is a fitted parameter in the model, and has a value of about 0.15, resulting in a quenching factor of about 0.2 for LXe in the low energy region below about 50 keV – i.e. about 80 % of the energy is lost to heat in NRs [70, 73, 78].

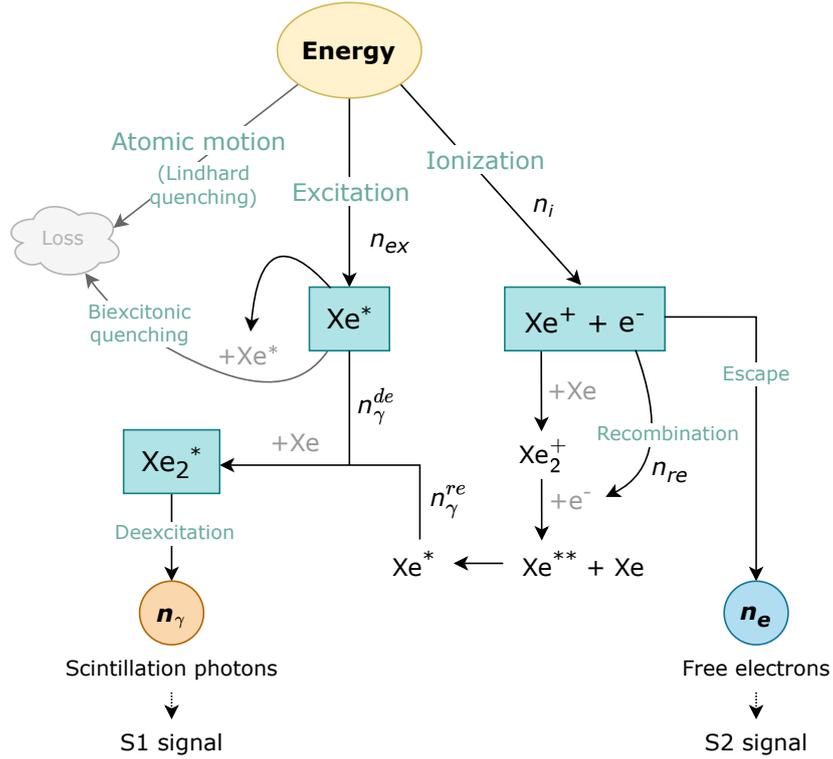


Figure 5.9.: Illustration of the microphysics processes modeled in the NR quanta generation model.

The number of detectable quanta then splits into a number of excitons n_{ex} and a number of ionization electrons n_i . This is modeled by defining the exciton-to-ion ratio $\langle n_{ex}/n_i \rangle$, such that the number of excitons follows a binomial distribution,

$$n_{ex} \sim \text{Binom} \left(n_q, \frac{\langle n_{ex}/n_i \rangle}{\langle n_{ex}/n_i \rangle + 1} \right) \quad (5.9)$$

$$n_i = n_q - n_{ex}. \quad (5.10)$$

Again, the cross sections for excitation and ionization in nuclear recoils are unknown, so the exciton-to-ion ratio is parametrized as a power-law function of both deposited energy and applied electric field F in units of V/cm,

$$\langle n_{ex}/n_i \rangle = \alpha \cdot F^{-\zeta} (1 - \exp(-\beta \epsilon)), \quad (5.11)$$

where α , ζ , and β are fitted model parameters. The energy-dependence in the cross-sections is modeled via the exponential term.

The number of excitons is reduced for higher energies by bi-excitonic quenching, where two excitons collide leading to the emission-less deexcitation of one exciton, thus reducing the number of produced photons. A scintillation quenching factor f_l is defined, such that the remaining number of photons produced from direct excitation n_{γ}^{de} is again a binomial process,

$$n_{\gamma}^{de} \sim \text{Binom}(n_{ex}, f_l). \quad (5.12)$$

Since the exciton density in an NR track is expected to be proportional to the electronic stopping power, the quenching factor is parametrized using the quenching term from Birks' saturation

law [79, 80],

$$f_l = \frac{1}{1 + \eta \epsilon^\lambda}, \quad (5.13)$$

where $\eta \epsilon^\lambda$ is an approximate parameterization of electronic stopping power, with η and λ fitted parameters.

After thermalization of the ionization electrons, a number n_{re} of the n_i electron-ion pairs can recombine, with the recombination fraction $r = n_{re}/n_i$. Because after recombination, the recombined atom is left in the excited state Xe^* (see also Equation 2.2), the number of pairs that recombine adds to the number of produced photons, n_γ^{re} , simulated via

$$n_\gamma^{re} \sim \text{Binom}(n_i, r). \quad (5.14)$$

The recombination fraction r is parametrized following the Thomas-Imel (TI) box model [81],

$$r = 1 - \frac{\ln(1 + n_i \varsigma)}{n_i \varsigma}, \quad (5.15)$$

with the TI box parameter ς , a power law function of the electric field

$$\varsigma = \gamma \cdot F^{-\delta}. \quad (5.16)$$

Finally, we get the total number of produced photons by summing the direct deexcitation photons and the ones from recombination, giving

$$n_\gamma = n_\gamma^{de} + n_\gamma^{re}. \quad (5.17)$$

The total number of produced free electrons after recombination is then given by

$$n_e = n_i - r \cdot n_i = n_i - n_{re} = n_i - n_\gamma^{re}. \quad (5.18)$$

Table 5.2 shows all parameters in the NR LXe emission model, along with their descriptions and related equations for reference and the best-fit values from [70], some of which are used as priors in the fit to the XENONnT calibration data, as will be explained in section 5.4.

5.3.2. Detector reconstruction model

In the detector reconstruction model, we cover the second part of the full response model – the conversion from the produced quanta n_γ and n_e into the detected and corrected S1 and S2 signal sizes. In contrast to the emission model described in the previous section, here we rely largely on input maps produced in dedicated detector studies in order to be able to simulate the full detector response in the BBF fit in a fast and efficient way.

In the fast toyMC simulation in BBF, the following steps are simulated for each scatter in an event separately. For the S1 signal, we first account for the loss of scintillation photons during their propagation and due to the detection efficiency of the PMTs. As described in subsection 2.2.2, this position dependent effect is encoded in the spatially dependent scintillation gain factor $g'_1(x, y, z)$ defined in Equation 2.3. To simulate the number of detected photons N_{phd} (not photo-electrons PE), we define

$$\tilde{g}_1(x, y, z) = g_1 \cdot \text{LY}_{\text{rel}}^{-1}(x, y, z) / (1 + p_{\text{dpe}}), \quad (5.19)$$

where the effect of double photoelectron emission (DPE) from the PMTs' photocathode is removed here as it is per construction included in the average scintillation gain g_1 . Both g_1 and the probability of DPE p_{dpe} are fitted parameters in our model, while the spatially dependent relative light yield map $\text{LY}_{\text{rel}}(x, y, z)$ is a data-derived input to the fit, evaluated in

Parameter name	Description and related equations	Values from NEST [70]	Unit
NR quanta generation			
f	Fano factor, Equation 5.5	0.059 ± 0.001	1
W	Average work function, Equation 5.5	13.7 ± 0.2	eV
κ	Lindhard-k, Equation 5.6	$0.1394^{+0.0032}_{-0.0026}$	1
α	Pre-factor for field dependence in exciton-to-ion ratio, Equation 5.11	$1.24^{+0.079}_{-0.073}$	1
ζ	Exponent in field dependence, Equation 5.11	$0.0472^{+0.0088}_{-0.0073}$	1
β	Exponential pre-factor in energy dependence, Equation 5.11	$239^{+28}_{-8.8}$	1
Scintillation quenching			
η	Pre-factor of energy dependence, Equation 5.13	$3.3^{+5.3}_{-0.7}$	1
λ	Exponent of energy dependence, Equation 5.13	$1.14^{+0.45}_{-0.09}$	1
Recombination			
γ	Pre-factor of field dependence in the TI recombination parameter, Equation 5.16	$0.01385^{+0.00058}_{-0.00073}$	1
δ	Exponent of field dependence in TI recombination parameter, Equation 5.16	$0.0620^{+0.0056}_{-0.0064}$	1

Table 5.2.: Microphysics model parameters for NR interactions in LXe, separated into quanta generation, scintillation quenching, and recombination processes. The reported best-fit values and uncertainties from the NESTv1 [70] are given as well, some of which will be used as priors in the NR band fit, see also [Table 5.3](#).

the simulation at each interaction’s position. The number of photons that hit a PMT and are detected is produced by sampling from a binomial,

$$N_{\text{phd}} \sim \text{Binom}(n_{\gamma}, \tilde{g}_1(x, y, z)). \quad (5.20)$$

Next, the DPE effect is added back to the number of detected photons,

$$N_{\text{dpe}} \sim \text{Binom}(N_{\text{phd}}, p_{\text{dpe}}), \quad (5.21)$$

giving the raw S1 signal size in units of PE via

$$N_{\text{S1,PE}} = N_{\text{phd}} + N_{\text{dpe}}. \quad (5.22)$$

For S2 signals, we need to account for the loss of electrons during their drift due to attachment to electronegative impurities, using the measured electron lifetime τ_e , the drift velocity v_d , as well as the scatter’s z -position, giving the electron survival probability $p_{\text{drift}}(z)$ as defined in [Equation 2.7](#). Additionally, electrons are lost if an interaction happens at the edges of the TPC close to the walls due to the (partially) charge-insensitive volume. This is modeled with the electron survival map shown in [Figure 5.7](#) (right), which encodes the electron survival probability $p_{\text{civ}}(r, z)$, where $r^2 = x^2 + y^2$ is the radial distance from the TPC center. Combining

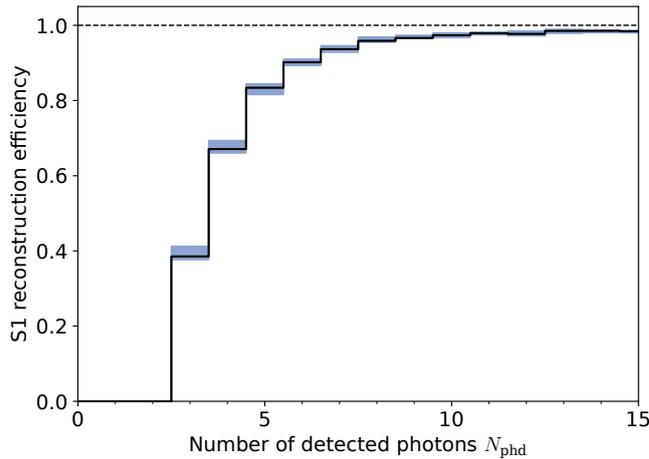


Figure 5.10.: XENONnT S1 reconstruction efficiency for 3-fold PMT coincidence requirement. The curve is derived from simulations and verified on data (see also [82]). The uncertainty is dominated by a small z -dependence, that is taken into account but not explicitly modeled here.

these two effects of electron loss yields the number of electrons that are successfully drifted to the liquid surface via

$$N_{e,\text{drift}} \sim \text{Binom}(n_e, p_{\text{drift}}(z) \cdot p_{\text{civ}}(r, z)). \quad (5.23)$$

The drift velocity is data-derived constant input parameter, while the value for τ_e is sampled from an input map of its evolution during the AmBe calibration data taking period. The electrons that arrive at the liquid surface are extracted with the xy -dependent extraction efficiency that we determine via $\varepsilon_{\text{extr}}(x, y) = (g_2/G) \cdot \text{QY}_{\text{rel}}(x, y)$, where we use the data-derived relative electron yield $\text{QY}_{\text{rel}}(x, y)$ and the average gas gain G as constant input⁴, while g_2 is a fitted parameter. The number of extracted electrons is then simulated via

$$N_{\text{extr}} \sim \text{Binom}(N_{e,\text{drift}}, \varepsilon_{\text{extr}}(x, y)). \quad (5.24)$$

The gas amplification process is assumed to follow Poisson statistics, approximated by a normal distribution due to the large gas gain value of $G \approx 30$ (see also Table 5.3), giving

$$N_{\text{S2,PE}} \sim \text{Norm}(N_{\text{extr}} \cdot G, \sqrt{N_{\text{extr}} \cdot G}). \quad (5.25)$$

This yields the raw S2 signal size in units of PE.

Apart from the physical processes described so far, effects occurring during digitization, processing, and reconstruction of the signals need to be accounted for as well. Small S1 signals are not always successfully reconstructed, due to digitization and PMT hit finding thresholds, and the tight coincidence requirement for at least 3 PMT hits to be found within a 100 ns window centered around the peak maximum for the peak to be considered a signal, leading to a loss of small S1 signals. This is the effect that mainly determines the energy threshold of the detector. The S1 reconstruction efficiency is modeled by a waveform-simulations-driven input map which was verified on data, and is a function of the number of detected photons N_{phd} . It is shown in Figure 5.10. It consists of the median efficiency $\varepsilon_{\text{S1rec}}(N_{\text{phd}})$ as well as lower and upper 1σ percentiles, $\sigma_{\text{S1rec}}(N_{\text{phd}})$. The probability of reconstructing an S1 in the BBF simulation is then sampled from the map, including a fitted scaling parameter f_s for the uncertainties, via

$$p_{\text{S1rec}} \sim \varepsilon_{\text{S1rec}}(N_{\text{phd}}) \cdot (1 + f_{s,\text{S1rec}} \cdot \sigma_{\varepsilon_{\text{S1rec}}}(N_{\text{phd}})). \quad (5.26)$$

A simulated event is then randomly dropped or not based on its sampled S1 reconstruction

⁴its xy -dependence is affectively absorbed into $\text{QY}_{\text{rel}}(x, y)$

probability.

Signals that are successfully reconstructed may not be reconstructed with the correct signal size, resulting in reconstruction biases for both S1 and S2. Reasons for this are e.g. the loss of small, underamplified PMT pulses due to the digitization and hit finding thresholds such that they are not included when building the S1 signal from the single PMT pulses, leading to a negative S1 bias for small S1s. Other sources of bias are noise peaks, which can give both positive and negative biases, PMT afterpulses that get merged into the signals giving positive bias, and the reconstruction software wrongfully splitting away parts of a signal. In the fast simulations we model the relative median biases for S1 and S2, Δ_{S1} and Δ_{S2} , as well as the spread of the biases, $\sigma_{\Delta_{S1}}$ and $\sigma_{\Delta_{S2}}$. Both median and spread are functions of N_{phd} and of $N_{S2,PE}$, for S1 and S2 respectively, and are sampled from input maps which are shown in [Figure 5.11](#). The biases are sampled from a gaussian distributions,

$$\text{S1Bias} \sim \text{Norm}(\Delta_{S1}, \sigma_{\Delta_{S1}}), \quad (5.27)$$

$$\text{S2Bias} \sim \text{Norm}(\Delta_{S2}, \sigma_{\Delta_{S2}}). \quad (5.28)$$

Applying them to the raw signal sizes then yields the reconstructed, uncorrected signals as they would be observed in data,

$$\text{S1} = N_{S1,PE} \cdot (1 + \text{S1Bias}), \quad (5.29)$$

$$\text{S2} = N_{S2,PE} \cdot (1 + \text{S2Bias}). \quad (5.30)$$

Next we need to apply the signal corrections as done in real data. In data, the spatial correction maps are applied based on an event's reconstructed position, which differs from the true interaction position due to the non-zero resolution of the position reconstruction algorithm. Before we correct the signals in the simulation, we therefore apply the position reconstruction resolution as a function of S2 size by smearing the true x, y (the z resolution is much smaller and dropped) to get the reconstructed positions (x_r, y_r) . The corrected S1 signal, cS1, is then given by

$$\text{cS1} = \text{S1} \cdot \text{LY}^{-1}(x_r, y_r, z). \quad (5.31)$$

For the S2 correction, we apply a similar smearing due to the uncertainty of the electron lifetime τ_e , since again, the true value may differ from the reconstructed value used in the correction drift loss correction. The reconstructed electron lifetimes given by $\tau'_e = (1 + \sigma_{\tau_e}) \cdot \tau_e$, where the

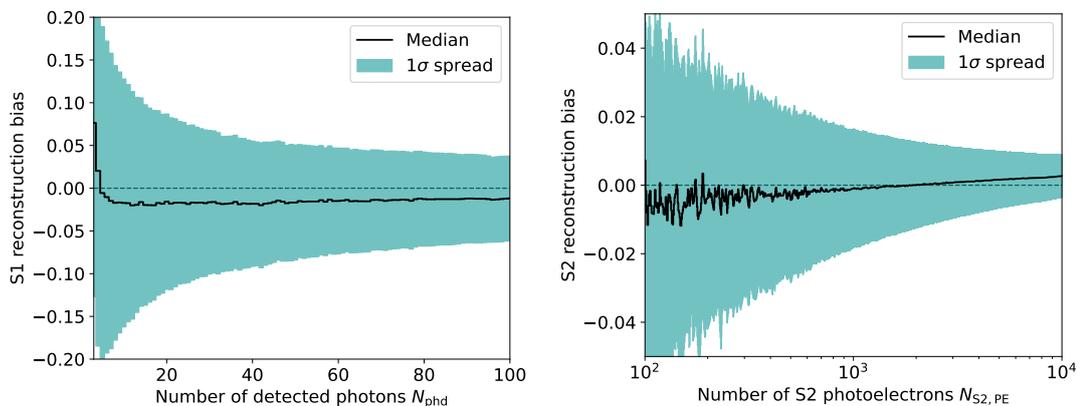


Figure 5.11.: Relative reconstruction biases for S1 (left) and S2 (right) signals derived from simulations in [\[82\]](#). The black line shows the medians, while the shaded regions represent the 1σ spread of the biases.

relative uncertainty σ_{τ_e} is a fitted parameter. The S2 correction is then applied via

$$cS2 = S2 \cdot QY^{-1}(x_r, y_r) \cdot \exp\left(\frac{-z}{\tau_e' \cdot v_d}\right). \quad (5.32)$$

Finally, data selection efficiencies, i.e. the cut acceptances, are applied. Two maps are used for cuts acting on S1 and S2 signals, respectively, which are functions of the corrected signals cS1 and cS2. Similar to how the S1 reconstruction efficiency is modeled, the probability of keeping a given event is sampled from cut acceptance maps with fitted scaling parameters $f_{s,cutcS1}, f_{s,cutcS2}$ for the uncertainties,

$$p_{cutcS1} \sim \epsilon_{cutcS1}(cS1) \cdot (1 + f_{s,cutcS1} \cdot \sigma_{\epsilon_{cutcS1}}(cS1)), \quad (5.33)$$

$$p_{cutcS2} \sim \epsilon_{cutcS2}(cS2) \cdot (1 + f_{s,cutcS2} \cdot \sigma_{\epsilon_{cutcS2}}(cS2)). \quad (5.34)$$

The corrected and accepted signals, cS1 and cS2, are then used as the reconstructed signals for the fit.

5.3.3. Multi-scatter NR handling

The detector reconstruction processes described above are valid on a per-interaction basis, i.e. for the quanta emission model per energy deposition, and are therefore only valid for SS events with one energy deposition. Due to the energy-dependent processes in the NR emission model, we can not simply sum the quanta from the separate scatters to get the total number of quanta produced in the event, and must instead account for the MS event topology in the detector reconstruction model.

The neutrons from the AmBe source are sufficiently fast that the mean time between scatters is on the order of a few ns, considerably smaller than the pulse width of an S1 signal of about 100 ns⁵. Separate S1 signals from different scatters in one event will therefore not be resolved as such, but appear as a single, merged S1 signal. For multiple S2 signals, the time difference with which they appear in an event depends on the z -positions of the respective interaction positions via the drift time. Whether they are resolved or not therefore depends both on the vertical distance between the scatters and their absolute z -position, due to the increasing width of S2 signals from scatters at larger depths due to electron cloud diffusion. Because of the low drift field in XENONnT of 23 V/cm, the maximum drift time for events at the bottom of the TPC is about 2.2 ms, leading to wider S2 pulses due to diffusion of the electron cloud during the drift. S2 signals from the cathode in XENONnT have a central 50% width of about 10 μ s, corresponding to the time of a 0.7 cm drift. This results in a S2 z separation resolution ranging from about 0.5 cm at the top of the TPC to about 1.6 cm at the bottom [42].

While both resolved and unresolved MS events can be rejected to a certain degree by the S2SS (S2 width) cut in the data selection, this only works with a certain efficiency, and we still need to account for the remaining MS events in the fit. However, the reconstruction of MS event topologies in the detector is extremely difficult to model from first principles in the fast simulation used in the BBF framework, as they require a full waveform simulation in order to describe the effects in the software processing. Instead, in the context of this work the detector reconstruction workflow in the fast simulations was modified to be able to take additional input from the full-chain MC simulation framework of XENONnT [66] into account, allowing to effectively account for the effects of (partially) unresolved MS events.

The full-chain MC framework can simulate events in the TPC from incident particle propagation up to the final waveforms that can be put into the same processing framework as the

⁵The S1 pulse width results from the following processes: The xenon scintillation decay time of about 4 ns (20 ns) for the singlet (triplet) excimer state [83], and a broadening due to different photon propagation paths to the PMTs caused by Rayleigh scattering and reflections on surfaces on the scale of several 10 ns, see also [66]

data, *strax*/*straxen* [48, 49]. It consists of three main parts. First, the GEANT4 toolkit [67] (G4 in the following) with an implementation of the full XENONnT geometry including muon and neutron vetos is used for particle propagation and energy deposits, giving position and deposited energy for every interaction in LXe (as well as in the vetos, if needed). Next, the EPIX [68] (Electron and Photon Instructions generator for XENON) package is used to first ‘micro-cluster’ the G4 energy depositions in the TPC on the LXe micro-physics scale (each micro-cluster then corresponding to one scatter), and then generate the resulting scintillation photons and electrons for the given recoil type. Both NEST models or the BBF model with best-fit parameters can be used in that step. Finally, the waveform simulator WFSim [69] generates complete waveforms from these instructions, using a detailed, per-PMT detector reconstruction model⁶. These waveforms then correspond to simulated raw data, and is fed into the *straxen* framework to be processed in the same way as real data. This way, full-chain MC simulations can be used to determine which scatters’ S2s may get merged during processing, and which events will be removed by the MS rejection cuts.

The information from the full-chain simulations is propagated into the BBF framework by using the same G4 events produced in the first step of a full-chain MC simulation of the AmBe neutron data. According to the outcome of the full-chain simulation, the G4 events are back-tagged, indicating whether a scatter’s produced ionization electrons will become part of the main S2 signal, an alternative S2 signal, or neither for each scatter in an event. Events that will be removed by the MS rejection cuts are removed from the back-tagged G4 input set before feeding it to the BBF framework, in order to save computational power and increase the speed of the fitting process. The full-chain MC simulation of the AmBe neutron data was performed by Pavel Kavrigin (Weizmann Institute of Science, Rehovot, Israel) and Shenyang Shi (Columbia University, New York City, USA), accompanied by a detailed study comparing the S2 signal shapes and sizes between simulation and data that achieved a good match in all relevant parameters.

Examples of MS AmBe neutron events from the back-tagged G4 input are shown in [Figure 5.12](#), illustrating different topologies of MS events that can occur. Each scatter’s position (r, z) is indicated by the colored circles, where the color represents the scatter’s deposited energy. The arrows connect the scatters ordered by time in which they occur, starting from the first and ending at the last scatter. Scatters marked by a red or blue star will become part of the main or alternative S2 signal, respectively, corresponding to the tag in the G4 dataset. All examples are for events where the main S2 contains more than one energy deposit, which is the case for about 43% of simulated events. The event shown in the top left panel, for example, is an unresolved MS event with two scatters, where both will become part of the main S2 signal, but will not pass the S2 width cut due to the combination of a shallow depth and relatively large separation in z , while the events shown in the top and bottom right panels will pass both the S2 width and S2 SS cuts. An event that will pass the S2 width cut but fail the S2 SS cut is shown in the bottom left panel. Interestingly, in this event the scatters forming the main S2 both deposit less energy than the scatter forming the alternative S2, underlining the importance of the S2-tagging scheme for the BBF input. In the default BBF simulation workflow, the largest-energy scatter would become the main S2, the second largest the alternative S2, and the S2SS cut would be applied directly in the simulation based on the simulated S2 sizes, ultimately resulting in a mis-modeling of MS events.

The simulation workflow in the detector reconstruction model as presented in the previous subsections is modified for MS events as follows. After all steps have been simulated for each scatter i in an event separately, the merged S1 and S2 signals are formed by summing the

⁶The detector reconstruction model used in WFSim is much more complex than the one used in BBF ([subsection 5.3.2](#)), and includes e.g. simulating photon timing via the scintillation decay times and photon propagation, electron propagation using a COMSOL® simulated electric field map, diffusion, electron extraction and amplification via the GXe electroluminescence process using the Garfield++ tool, per-PMT photon detection efficiencies including incident-photon-angle and radial-photocathode-QE dependence, per-PMT SPE pulse shapes, S1 and S2 PMT light patterns, and background effects such as PMT afterpulses and noise.

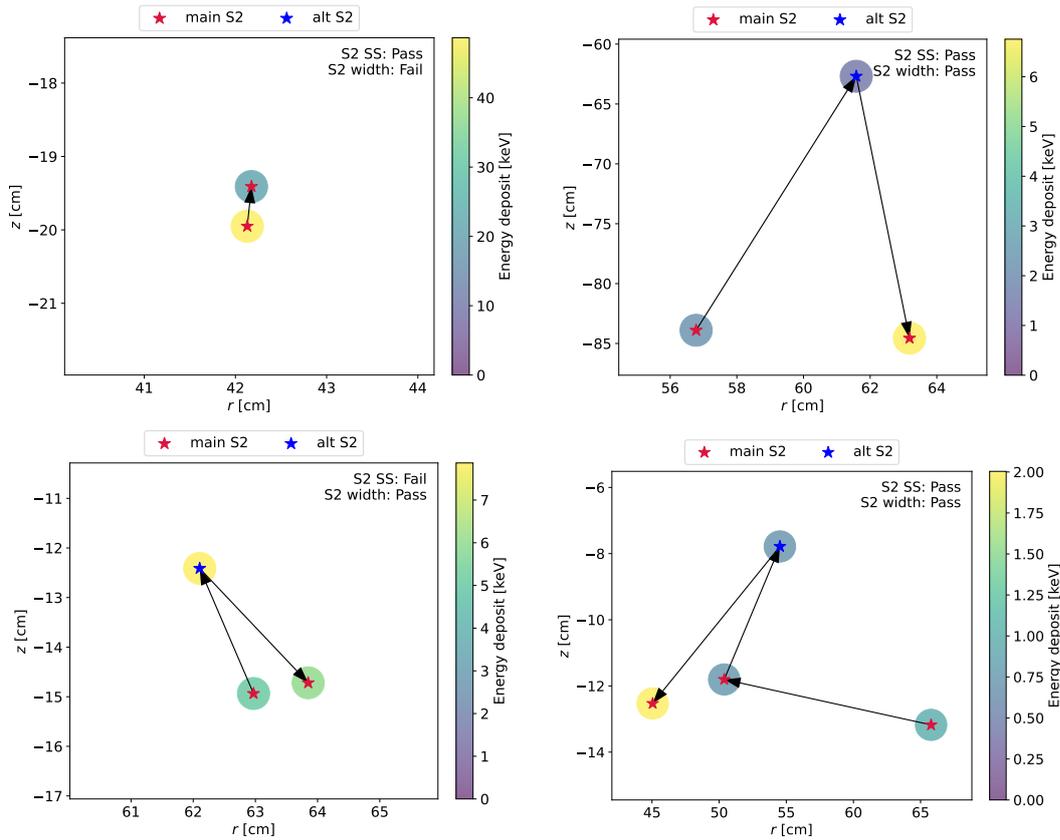


Figure 5.12.: Examples for multi-scatter (MS) event topologies in the AmBe neutron events simulated with the full-chain MC framework. Colored circles represent the position of the scatters in (r, z) in the TPC and the deposited energy, indicated by the colorbars in each panel. Black arrows indicate the order of the scatters in time. Red (blue) stars indicate which scatters will become part of the main (alternative) S2 signals. Each panel also states whether the event passes the S2 single scatter (S2 SS) and S2 width (S2 width) cuts.

signals according to the relevant steps. For S1 signals, the number of detected photons N_{phd}^i from all scatters are summed up to form the total number of detected photons,

$$N_{\text{phd}} = \sum_i N_{\text{phd}}^i, \quad (5.35)$$

and the resulting total raw S1 signal size $N_{\text{S1,PE}}$ as given by Equation 5.21 and Equation 5.22 is determined from this total number of detected photons. Accordingly, the S1 reconstruction efficiency (Equation 5.26) is calculated from the total number of detected photons as well. For S2 signals, we sum at the level of the raw S2 signal sizes $N_{\text{S2,PE}}^i$ given in Equation 5.25, i.e. after the drift and CIV losses have been applied,

$$N_{\text{S2,PE}} = \sum_{(i | \text{tagged mainS2})} N_{\text{S2,PE}}^i, \quad (5.36)$$

where the sum is over all scatters i that are flagged ($=0$) to become part of the main S2 signal. For both S1 and S2 signals, the biases and signal corrections are then applied to the merged signals as it would be for real data. Since the signal corrections have a spatial dependence, the reconstructed positions of events with merged scatters are approximated by the average position of all scatters that are part of the main S2, weighted by the raw S2 size, and these positions are used for the corrections.

5.4. Putting it together: Fitting the model

Now we put the previous three sections together, and fit the full detector response model (section 5.3) to the AmBe calibration data (section 5.1), using the BBF framework (section 5.2).

For the fit, the selected 1987 AmBe calibration data events are binned in $cS1\text{-}\log_{10}(cS2)$ space, from $[0 - 100]$ PE in $cS1$ and from $[2 - 4]$ in $\log_{10}(cS2[\text{PE}])$, with 50 bins in each dimension. This is the same histogram as shown in Figure 5.4 (right, full color histogram).

In Table 5.3, all parameters of the full model are listed, including prior and posterior values where appropriate. We differentiate between fitted and fixed parameters, as indicated in the column ‘Prior type’. For parameters that we fit, the prior type indicates the functional form of the prior distribution, using the values given in the column ‘Prior values’. ‘norm’ refers to a normal distribution with given mean and standard deviation, while ‘asym. norm’ refers to an asymmetric normal distribution with different widths on either side of the mean given by the asymmetric errors of the priors. ‘uniform’ is a uniform prior in the given range. A prior type of ‘free’ means that the parameter is fitted, but no prior is used, the parameters may however still be constrained in the traditional sense of being required to stay within physical ranges. For fixed parameters, the walkers are distributed according to a normal distribution during initialization of the MCMC chain, using the values given in the ‘Prior values’ column, but the walkers are not allowed to move in these dimensions during the fit.

The priors for the NR quanta emission model parameters are based on the values reported in [70], where a global fit to a large set of literature data in a wide range of energies and electric fields was performed using the NEST framework [71] – these are the values given in Table 5.2. However, one difficulty that proved to be challenging for modeling the calibration data is the extremely low drift field in the XENONnT TPC of only ~ 23 V/cm. While the NR microphysics model presented above has been used successfully in the past, in particular in XENON1T [58, 59], its validity at low drift fields is unverified due to the lack of data in literature. Similarly, the fit to global literature data performed in [70] suffers from the same lack of literature data, largely relying on extrapolation from higher drift fields. Thus, the reported best-fit parameters may not be appropriate here. In the fitting process, we therefore fix the two parameters that govern the field dependence in the model to the best-fit values from [70], namely ζ in Equation 5.11 and δ in Equation 5.16, while freeing the corresponding prefactors α and γ (i.e., no priors are applied). All other parameters’ priors are loosened by increasing their widths by a factor of 2 with respect to the published values. Since the electric field dependence in the NR microphysics is expected to be small (see literature measurements shown in Figure 5.19), this is considered a reasonable approach, allowing more freedom in the fit to adjust the parameters to the low-field conditions in XENONnT.

The prior values (or fixed-walker-distribution values) for the parameters in the detector reconstruction model are extracted from dedicated studies of the detector’s performance, and are stated in Table 5.3. The scaling factor priors are set to a normal distribution with zero mean and standard deviation 1, scaling the uncertainties of the input maps in units of sigma.

5.4.1. Fit results

After the fit has converged, the last 500 iteration steps in the MCMC chain for all 200 walkers are used to represent the final posterior distribution. The best-fit parameters are extracted by taking the median of the corresponding marginalized posterior distribution for each parameter, while the negative and positive credible intervals correspond to the 15.9 and 84.1 percentiles, i.e. covering the central 1σ range of the distributions. The best-fit values and uncertainties for all parameters are given in Table 5.3 in the column ‘Posterior values’. The marginalized posterior distributions for all parameters are shown in Figure 5.13, represented by the grey histograms; the 15.9, median and 84.1 percentiles are indicated by the vertical black lines. The prior probability distributions are overlaid in green for comparison. For the two NESTv1 model

Parameter	Description	Prior type	Prior values	Posterior values	Unit
NR quanta emission model parameters					
W	Av. energy per quantum	norm	13.7 ± 0.2	$13.73^{+0.20}_{-0.21}$	eV
f	Fano factor	fixed	0.059 ± 0.001	-	1
κ	Lindhard-k	asym. norm	$0.1394^{+0.0064}_{-0.0052}$	$0.1385^{+0.0050}_{-0.0061}$	1
α	N_{ex}/N_i	free	-	$0.92^{+0.07}_{-0.062}$	1
ζ	N_{ex}/N_i	fixed	0.0472 ± 0.0088	-	1
β	N_{ex}/N_i	asym. norm	239^{+56}_{-18}	334^{+40}_{-43}	1
γ	Recombination	free	-	$0.0164^{+0.0014}_{-0.0014}$	1
δ	Recombination	fixed	0.062 ± 0.006	-	1
η	Scintillation quenching	asym. norm	$3.3^{+10.6}_{-1.4}$	$10.0^{+6.8}_{-5.9}$	1
λ	Scintillation quenching	asym. norm	$1.14^{+0.90}_{-0.18}$	$1.4^{+0.61}_{-0.38}$	1
Detector reconstruction parameters					
G	Gas gain	fixed	31.15 ± 0.97	-	PE/e ⁻
v_d	Drift velocity	fixed	0.0675 ± 0.0006	-	cm/ μ s
F	Electric field	fixed	22.92 ± 0.62	-	V/cm
g_1	Scintillation gain	norm	0.15149 ± 0.00138	$0.1514^{+0.0012}_{-0.0013}$	PE/ n_γ
g_2	Ionization gain	norm	16.45 ± 0.64	$16.42^{+0.66}_{-0.67}$	PE/ n_e
p_{dpe}	DPE prob.	uniform	[0.18, 0.24]	$0.222^{+0.013}_{-0.023}$	1
σ_{τ_e}	Electron lifetime uncertainty	norm	0.0 ± 0.02	$0.0^{+0.021}_{-0.020}$	1
Detector reconstruction maps					
S1 reconstruction efficiency	Scaling	norm	0.0 ± 1.0	$-0.31^{+0.78}_{-0.91}$	1
cS1 cut acceptance	Scaling	norm	0.0 ± 1.0	$0.27^{+0.84}_{-0.58}$	1
cS2 cut acceptance	Scaling	norm	0.0 ± 1.0	$0.63^{+0.66}_{-0.61}$	1
Other					
Number of NR events	Normalization	free	-	1991^{+48}_{-54}	1

Table 5.3.: All parameters used in the NR model for the band fit. The prior type indicates the functional form of the prior distribution. The prior values for the quanta emission model parameters are based on [70] as described in the text. The detector reconstruction parameters are based on dedicated studies of the detector’s performance, see e.g. [14, 84–86]. The ‘Posterior values’ column lists the median and 68% credible region extracted from the marginalized posterior distributions after the fit.

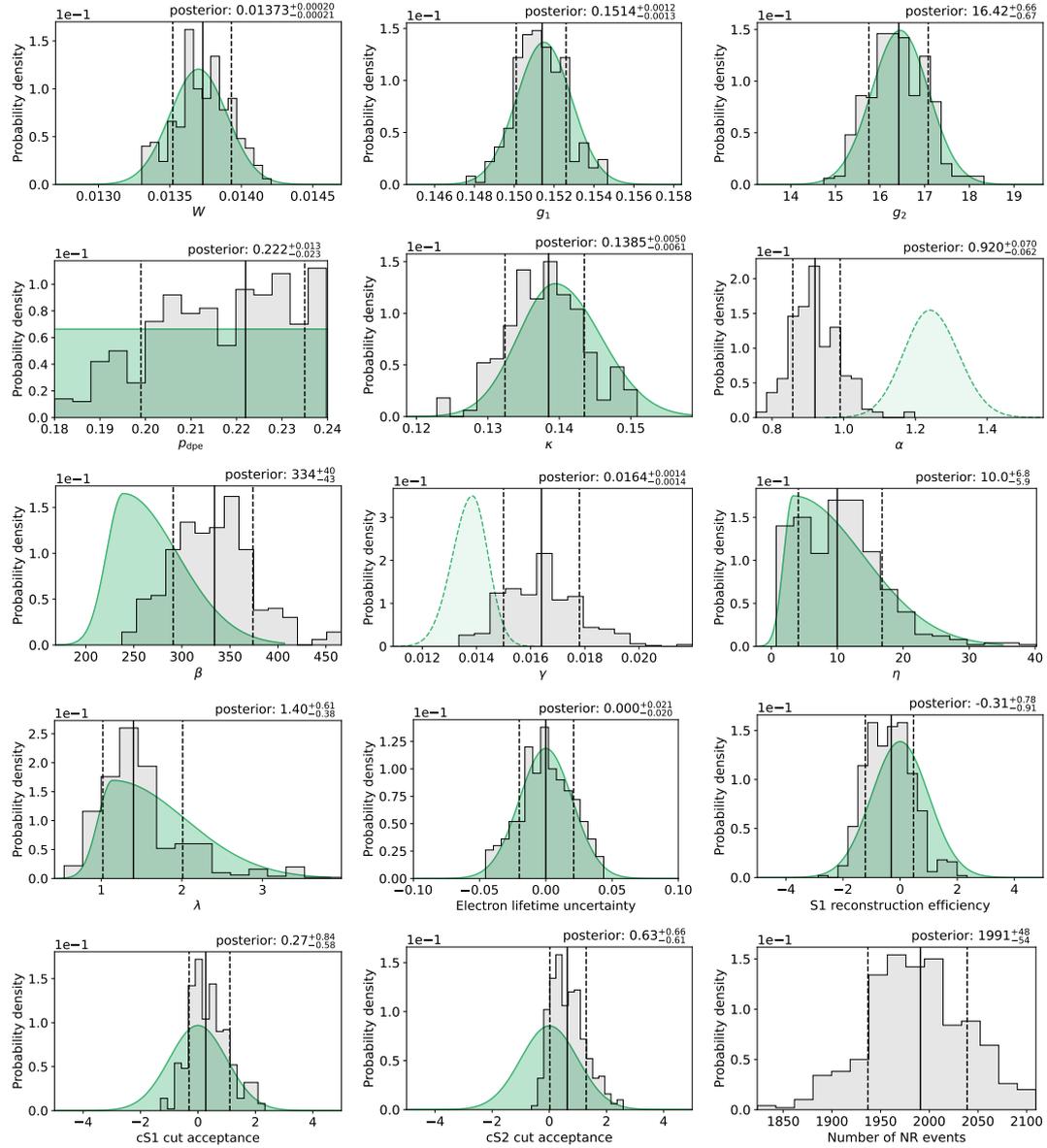


Figure 5.13.: Comparison between the parameter prior and marginal posterior distributions for all fitted parameters in the full detector response model for NR interactions. The analytical prior density PDFs are shown in green, their definitions are described in the main text and in [Table 5.3](#). For the freed parameters α and γ the lighter shaded green distributions are the PDFs according to the NESTv1 model parameter values. The posterior histograms (grey) show the samples from the last 500 iterations of the 200 walkers in the MCMC chain. Median and 1σ credible regions (15.9% and 84.1% percentiles) are indicated by the black vertical lines, solid and dashed respectively.

parameters α and γ that were freed, i.e. where no prior was used in the fitting process, the probability distributions given by the NESTv1 parameter values are shown in a lighter shaded green color, corresponding to the prior distribution that would have been used in the fit had the parameters not been freed.

None of the parameters differ significantly from their prior distributions. For the NESTv1 model parameters, the best-fit values are consistent with the prior distributions. The differences between posterior distributions and priors is visualized in [Figure 5.14](#), showing the so-called pull plot, indicating how far each parameter has been ‘pulled away’ from its prior value. The

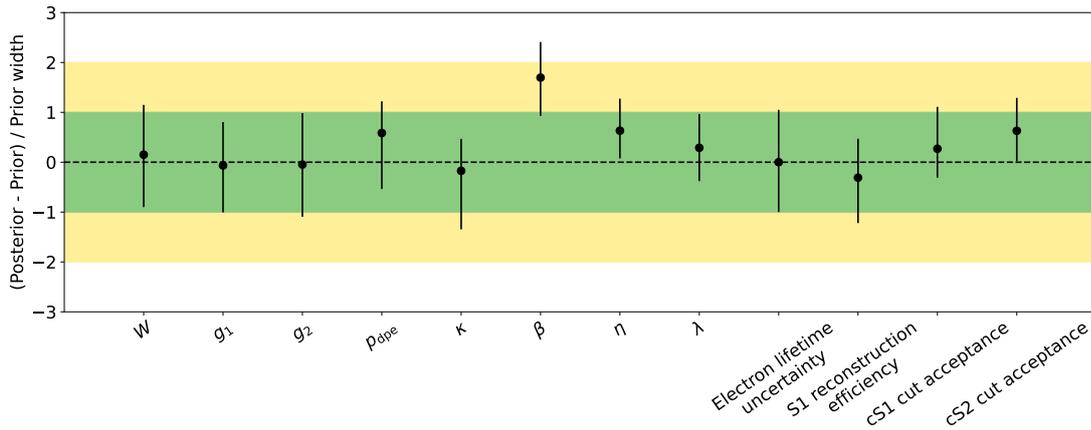


Figure 5.14.: Pull plot of the NR fit, showing the comparison between parameter posteriors and their respective priors. The y-axis shows the difference between prior and posterior in units of σ of the prior; the points and error bars correspond to the median and the 1σ credible regions of the posteriors. Note that for parameters with asymmetric priors, the prior widths are different for values above and below zero. With all pulls below 2σ , none of the parameters significantly deviate from their priors.

pull on the y-axis is defined as the difference between posterior median and prior value as stated in Table 5.3, in units of the standard deviation of the prior. Note that for parameters with asymmetric priors, comparing the posterior *median* with the prior distribution or the value in Table 5.3 is not actually appropriate, since the prior value in this case corresponds to the mode of the distribution. This results in exaggerated values for the pulls shown in Figure 5.14. Even with this, the pulls are all within 2σ of the expected value.

With the set of best-fit parameters, the full detector response model is simulated with the same fast toyMC simulations as used in the fit, yielding the final best-fit MC. The result of this simulation is shown in Figure 5.15, showing our best-fit NR model in $cS1\text{-}\log_{10}(cS2)$ space overlaid with the AmBe calibration data points in red. The median and the 5th and 95th percentile lines in $cS2$ for bins in $cS1$ are drawn for both data and MC, showing good agreement.

5.4.2. Goodness-of-fit tests and model validation

In order to quantify the validity of the fit result, goodness-of-fit (GOF) tests were performed using the *GOFevaluation* package [87]. In the main test, we perform a 2D binned Poisson- χ^2 test in the $cS1\text{-}\log_{10}(cS2)$ space, with an equiprobable binning scheme. Here, an irregular binning is defined that ensures the same number of expected events in each bin, based on the best-fit MC. This avoids the problem of large statistical uncertainties in low-count bins that can negatively affect the validity of χ^2 tests. Before performing the test, we chose a 90% confidence level (C.L.), i.e. we reject the null hypothesis that the data is consistent with the model if the p-value is 0.1 or lower. The number of bins for the GOF test in the equiprobable binning scheme was fixed before the fit, using 6 bins in $cS1$ and 9 bins in $\log_{10}(cS2)$. The result of the 2D GOF test is shown in Figure 5.16, where the equiprobable bins are represented by colored rectangles in the background. The color of the rectangles indicates the σ -deviation of the number of data events (overlaid as black dots) in the respective bin from the best-fit expectation. With the best-fit value for the number of events of 1991 (see Table 5.3) and $6 \times 9 = 54$ bins, each bin has an expectation of $\mu_{\text{bin}} = 36.9$ events. The resulting p-value is 0.39, indicating that the data is consistent with the best-fit model.

For diagnostics purposes, in order to catch potential mismodeling effects that may be missed in the 2D GOF test, additional one-dimensional GOF tests were performed in $\log_{10}(cS2)$, for

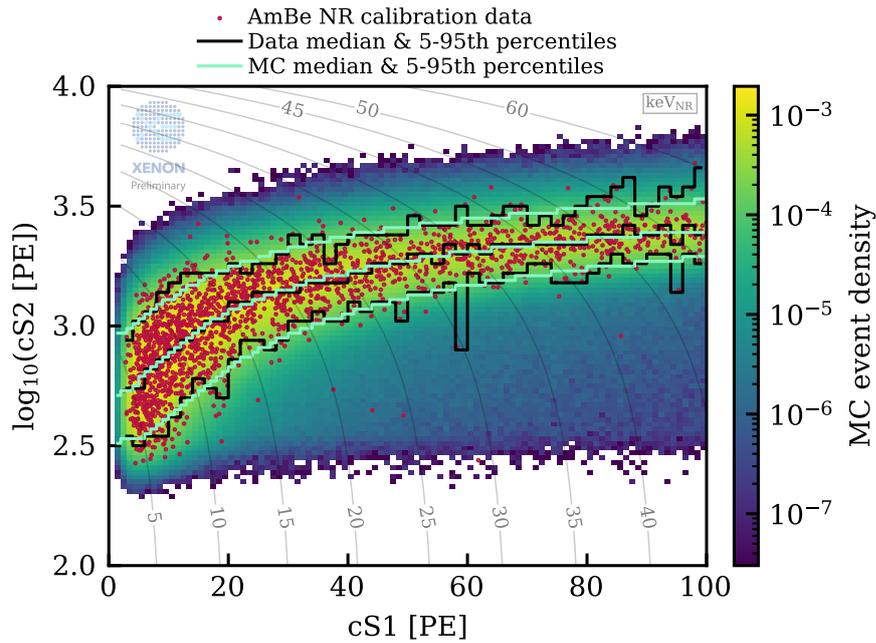


Figure 5.15.: Comparison of the NR band fit result with the calibration data. The histogram is a toy MC histogram of the NR band produced with the best-fit parameters. The AmBe NR calibration dataset is overlaid in red points. Colored lines indicate the median and the 5th and 95th percentile lines in $cS2$ for each bin in $cS1$, for both data (black) and MC (cyan). The grey lines are contours of constant deposited energy assuming NR interaction (i.e. in units of keV_{nr}).

the same equiprobable slices in $cS1$ as in the 2D test. For each slice in $cS1$, both another equiprobable-binned Poisson- χ^2 test and an unbinned Anderson-Darling (AD) test were performed. Especially the Anderson-Darling test is more sensitive to deviations in the tails of

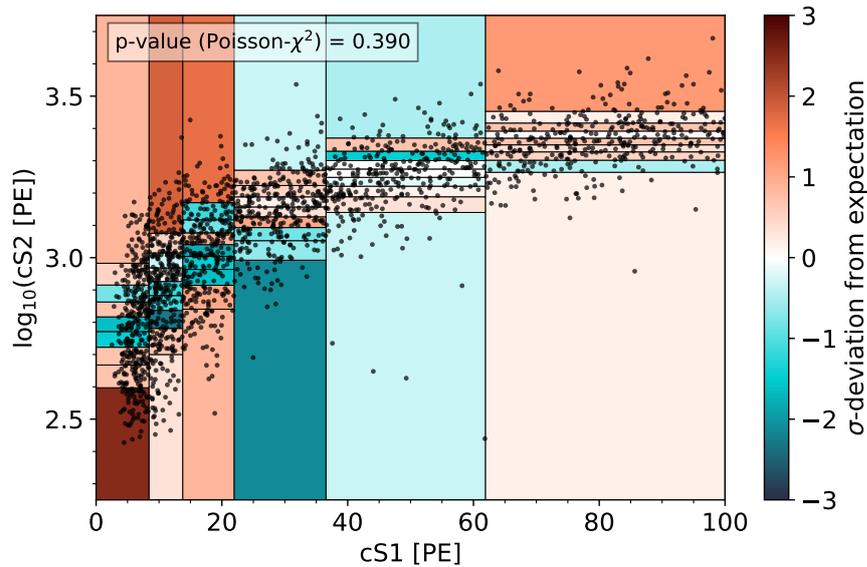


Figure 5.16.: Goodness-of-fit tests for the NR band fit, using an equiprobable binned 2D Poisson- χ^2 test. The black dots show the data events. The colors indicate the σ -deviation of the data from the number of expected events in each bin, $\mu_{\text{bin}} = 36.9$, derived from the best-fit MC. The p-value is 0.39.

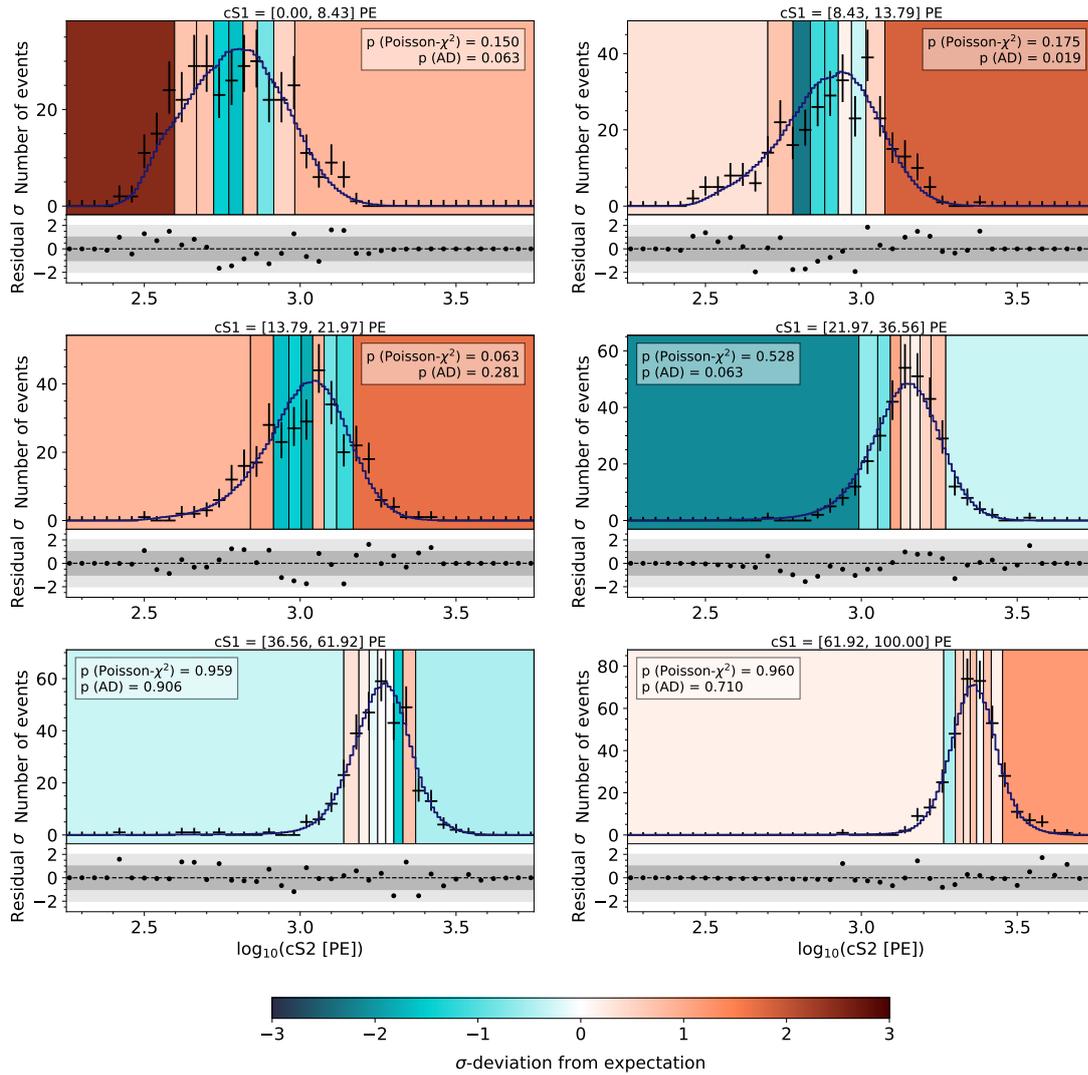


Figure 5.17.: Goodness-of-fit tests in the $\log_{10}(cS2)$ dimension for the same equiprobable slices in $cS1$ as in Figure 5.16. Black error bars show the data in the binning used for the fit, with the best-fit MC overlaid in dark blue. The p-values for the 1D Poisson- χ^2 (equiprobable bins with σ -deviation shown by the colored background) and unbinned Anderson-Darling tests are stated in the corresponding panels.

distributions, serving as a good complement to the Poisson- χ^2 tests. The results of those tests are shown in Figure 5.17. As before, the colored regions in each panel indicate the σ -deviation of the number of data events from the best-fit expectation, corresponding per construction to the same regions as in Figure 5.16. The histogrammed data for each slice is shown by the black error bars, using the same number of bins in $\log_{10}(cS2)$ as in the fit (note that is different from the equiprobable binning), and is compared to the best-fit MC in dark blue. The resulting p-values for both tests are stated in the corresponding panels, where several slices, especially at lower $cS1$, show p-values below the 0.1 threshold. Since here we test the same model six times instead of once, the chance of observing at least one p-value below the threshold of 0.1 is increased compared to the case of a single test, to $1 - 0.9^6 = 0.47$, if the model is correct. Defining a combined 90% C.L. threshold for the six tests gives a minimum p-value threshold for any of the six tests of $1 - 0.9^{1/6} = 0.017$. Both the 1D Poisson- χ^2 and the AD test in the $cS1$ slices show p-values above the combined threshold, indicating also here that the data is consistent with the best-fit model. Nevertheless, it is noticeable from the σ -deviations of the

Poisson- χ^2 tests encoded by the colors of the equi-probable regions (in both the 2D and sliced 1D tests), that the model tends to overpredict the number of events in the central part of the NR band (blue bins), and underpredict the number of events in the tails (red bins). Although not statistically significant, this may suggest that the NR band data has a larger spread in cS2 for low cS1 values than the model predicts. To ensure that this potential issue does not affect the WIMP search, a dedicated study was performed by Tim Wolf (formerly at MPIK Heidelberg) using the WIMP inference framework, where the NR band model was artificially broadened in cS2 for the production of the NR background model, and the resulting WIMP search sensitivity was compared to the nominal case. The result of this study showed that the effect of a potential mismodeling of the NR band shape of this size on the WIMP search sensitivity is negligible. This can be understood as an effect of the very low number of expected NR background events (expectation of 1.1 events in the WIMP ROI in SR0). The difference in the nominal WIMP signal rate due to this potential mismodeling is within the rate uncertainty that is propagated to the final statistical inference (see also [chapter 6](#)).

Two thirds of the events in the calibration dataset used in the fit are outside of the WIMP FV (see [Figure 5.6](#)), representing another possible source of uncertainty in the best-fit model’s reliability for the WIMP search. To ensure that the best-fit model is valid also inside the WIMP FV only, an additional test was performed by constricting both the calibration data and the best-fit MC events to the WIMP FV, resulting in 639 events in the dataset. We use again the 2D Poisson- χ^2 test, with 6×9 equiprobable bins in $cS1-\log_{10}(cS2)$, giving $\mu_{\text{bin}}=13.3$ events, and a p-value of 0.41. Despite this not strictly being a goodness-of-fit test (the proper way would be to repeat the fit in the WIMP FV, but this would likely be less sensitive because of the limited statistics in the data), we do not observe an evident mismodeling of our NR model inside the WIMP FV.

The impact of the merging of non-resolved S2s in the simulation was investigated as well. The average number of scatters for all events per $cS1-\log_{10}(cS2)$ bin is shown in the histogram in [Figure 5.18](#). The median and 5th-95th percentiles in $\log_{10}(cS2)$ per cS1 bin of the event density histograms from [Figure 5.15](#) are overlaid for reference, for the full MC simulation including S2 merging (cyan) and for the calibration data (black). In addition, the percentiles for a subset of the simulated events with exactly one scatter in the main S2 are overlaid as well (orange). While the 5th percentile shows no significant difference between the full MC and the subset of events with exactly one scatter, the median and especially the 95th percentile show a significant difference. The median and the 95th percentile of the full MC simulation on the other hand, as also shown in [Figure 5.15](#), shows a significantly better match to the calibration data. This illustrates that the effect of merging non-resolved S2s has a significant impact on the width in the cS2 dimension of the NR band from AmBe neutrons in the $cS1-\log_{10}(cS2)$ space. Furthermore, it is a good indication that the full model is able to account for the effect of merging non-resolved S2s in the simulation, underlining the importance of modeling this effect in order to be able to infer the NR emission model parameters from the data.

Finally, the few outlier events below the main NR band should be addressed again (see [subsection 5.1.2](#)). The outlier events at larger cS1 values can be modeled successfully by the inclusion of the charge-loss map in the detector reconstruction model, as is visible in [Figure 5.15](#). Here, the observed spread of the MC event histogram below the main NR band is caused by applying the charge-loss map in the simulation, covering the region where the outlier events are observed. Additionally, in [Figure 5.18](#), the impact of MS events on the region below the NR band is visible, showing an overall larger average number of scatters in the main S2 than in the main NR band population. This indicates that in that region, a combination of both charge loss due to the (p)CIV and the MS event topologies may be responsible for the observed events. For the outlier events at small cS1s with a hypothesized cause of non-rejected MS events near the detection threshold, this may be especially relevant. The presence of the pCIV for MS events can lead to a reduced S2 size for each individual scatter, and both a reduced main or alternative S2 leads to events falling below the S2SS cut threshold. Noticeably though, in the 2D GOF test ([Figure 5.16](#)) the bin with the largest σ -deviation is the one at the lowest cS1 and

cS2 values, where most of this kind of events is expected. Some small remaining mis-modeling of MS effects in the full-chain simulations may play a role as well. On the other hand, one outlier-bin should also not be over-interpreted, as its deviation is not statistically significant (the deviation is less than 3σ from the expectation) and the pre-defined GOF test as well as the other tests described in this section do not indicate a major mis-modeling of the data. With more NR band calibration data taken beyond SR0, this effect should and will be further investigated to clarify if the observed overdensity in the lowest cS1 and cS2 bin is a real effect or a statistical fluctuation.

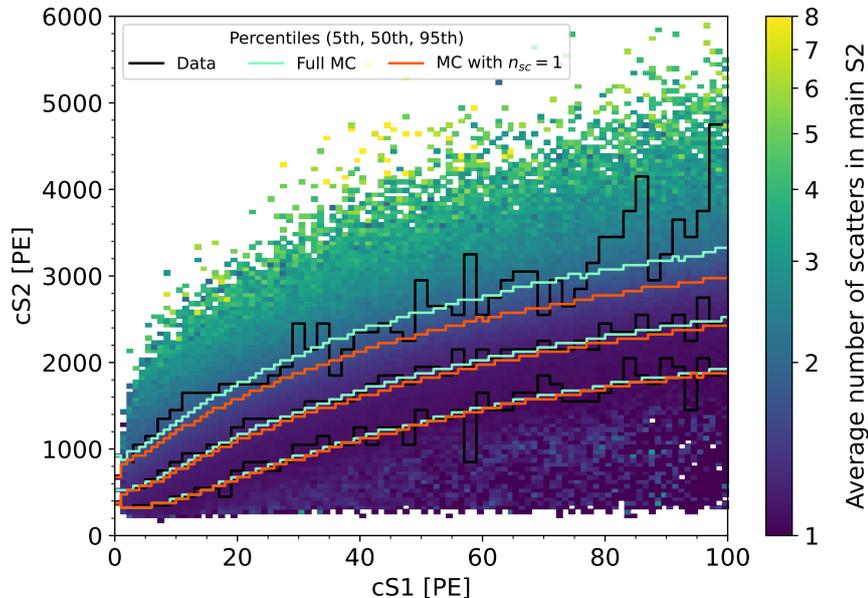


Figure 5.18.: Impact of merging non-resolved S2 signals from MS events in the simulation. The color scale indicates the average number of scatters for all events per $cS1\text{-}\log_{10}(cS2)$ bin. Lines show the median and 5th-95th percentiles in $\log_{10}(cS2)$ per $cS1$ bin, for the full MC simulation including S2 merging (cyan), for the subset of simulated events with exactly one scatter in the S2 (orange), and for the calibration data (black). The MC simulation was performed using the best-fit parameters.

5.4.3. Nuclear Recoil photon and charge yields

The best-fit parameters from the NR band fit are used to calculate the photon and charge yields for NR interactions in LXe as a function of deposited energy, following the chosen parameterization of the LXe quanta emission model from NESTv1 (see subsection 5.3.1). The resulting NR photon and charge yields are shown in figure Figure 5.19 in dark red. The lighter shaded red band shows the 1σ credible region, given by the credible region of the posterior distribution of the quanta emission model parameters. The grey shaded region indicates the WIMP ROI of $[3.3, 60.5]\text{ keV}_{\text{nr}}$, defined by the region where the total efficiency of event detection is above 10% (see also Figure 6.2). The prediction from the NESTv1 model evaluated at the 23 V/cm drift field of XENONnT using the parameter values reported in [70] (see Table 5.2) is shown by the red dashed line, while the more recent NESTv2 [88] model evaluated at 23 V/cm is shown by the black line. The yields resulting from the NR band calibrations in XENON1T [59] and XENON100 [55], are shown by shaded bands as well, where the XENON1T band fit uses the same NESTv1 model parameterization as in this work, but had a larger drift field of 81 V/cm. Note that an error in the plotting function⁷ for the NR charge yield curve in XENON1T (Fig 1, upper right panel in [59]) was found in the context of this work; here the fixed curve is

⁷Only the plotting function producing the curve in the figure is affected, with no impact on the NR modeling in the XENON1T analysis.

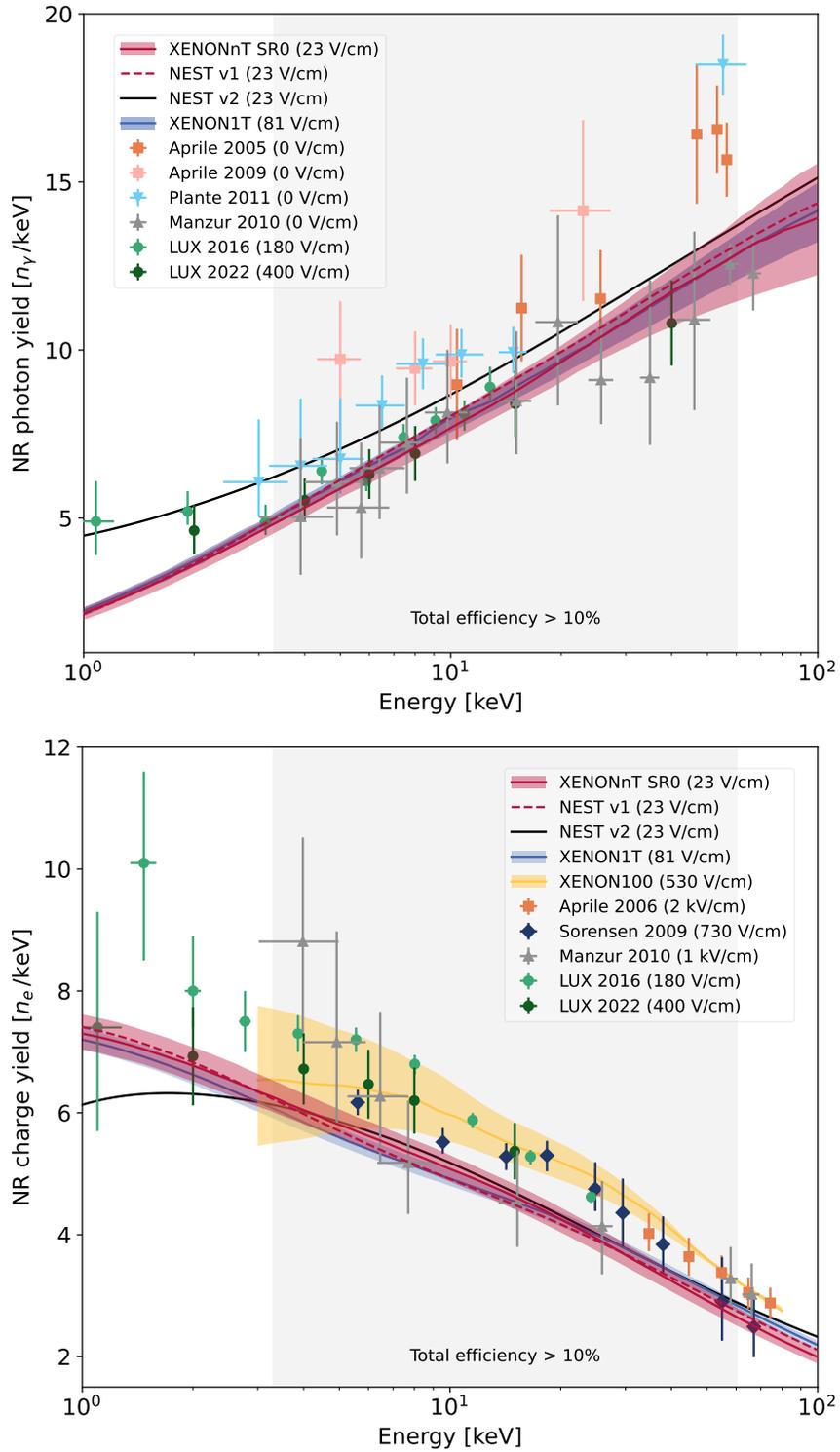


Figure 5.19.: Photon yield (top) and charge yield (bottom) for NR interactions in LXe as a function of deposited energy. The dark red line and shaded band show the LXe quanta emission model resulting from the best-fit parameters and the 1σ credible region in this work, respectively. Blue lines and shaded bands show the photon and charge yields from the NR band fit in XENON1T [59], using the same NESTv1 model parameterization as in this work. Red dashed and black lines show the predicted yields from the NESTv1 [70] and NESTv2 [88] models, respectively, evaluated at 23 V/cm drift field. Literature data from Aprile 2005 [89], Aprile 2006 [90], Aprile 2009 [91], XENON100 [55], Plante 2011 [92], Sorensen 2009 [93], Manzur 2010 [94], LUX 2016 [95], and LUX 2022 [96] is shown for comparison. Figures published in [51].

shown. The literature data points show results from a variety of fixed-angle neutron scattering measurements, as stated in the figure caption. The photon yields in the measurements from Aprile 2005 [89], Aprile 2009 [91], Plante 2011 [92], and Manzur 2010 [94] are reported as relative scintillation yields w.r.t. to the photon yield of 122 keV ERs from ^{57}Co γ -rays, and are scaled here to absolute photon yields using the corresponding absolute ER yield of 69 ph/keV predicted from NEST [88].

Generally, the photon and charge yields from the NR band fit are in very good agreement with the literature data as well as with predictions from the NESTv1 model, with overall little observed field dependence. For the NESTv2 model prediction, particularly the photon yield is slightly higher than from the NR band fit, with an increasing discrepancy for energies below about $3 \text{ keV}_{\text{nr}}$ outside the region that we can probe with the NR band data. In addition to the different parameterization used between NESTv1 and NESTv2, the latter is in the low energy regime mainly constrained by the datasets LUX 2016 [95] and LUX 2022 [96] (shown in green) that were not yet available for NESTv1, both measured at considerably higher drift fields than XENONnT, of 180 V/cm and 400 V/cm respectively.

5.5. Conclusions

In this chapter, we have derived the full detector response model to NR interactions for XENONnT SR0, using the BBF tool to fit the NR band calibration data from the AmBe source. In particular, the implementation of the handling of MS event topologies in the fast toyMC simulations were crucial to be able to account for the effect of non-resolved MS events with merged S2 signals, allowing for a more accurate modeling of the NR band shape. The best-fit model is able to describe the data well, with a p-value of 0.39 in the 2D GOF test, and several checks were performed to ensure the validity of the fit results for the WIMP search.

For future WIMP search runs beyond SR0, additional NR band calibrations have been taken with the AmBe-neutron in different source positions, both increasing the available data statistics significantly as well as a larger coverage of the TPC volume, in particular with more data in the bottom u-tube positions. With the closer proximity to the (p)CIV, this will allow for a better characterization of the outlier events below the NR band. Additionally, in the near future a DD-fusion neutron generator will be available to the collaboration, that produces higher energy neutrons that will penetrate farther into the LXe volume than the AmBe source, increasing the coverage of the WIMP FV with the calibration data. However, in future large-scale LXe TPCs the calibration of the NR response with neutrons will become challenging, as the high self-shielding power of LXe will make it more difficult to calibrate inner detector volumes with external sources. New methods for NR calibrations in large xenon detectors will have to be developed, such as the use of internal sources or other calibration techniques.

Chapter 6.

First WIMP search results of XENONnT

In this chapter, the first results of the WIMP Dark Matter (DM) search [84] with XENONnT's first science run are summarized. We will begin with an overview over SR0 data taking periods in section 6.1 and a short summary of the analysis procedure. In section 6.2, the signal and background modeling is discussed, where signal and NR backgrounds are relying on the NR response model presented in the previous chapter 5. Finally, the statistical inference and the WIMP search results are presented in section 6.3.

6.1. SR0 Overview

The first science run of XENONnT took place from May 1st to December 10th, 2021, during which a total livetime of 97.1 days of science search data was acquired between July 6th and November 10th. Figure 6.1 shows the evolution of acquired livetime for the science search in the white regions. Before science data taking started, ^{220}Rn ER and $^{241}\text{AmBe}$ NR band calibration data (see section 5.1 for the latter) was acquired in the dark blue and yellow marked regions, respectively. A further set of calibration data with an ^{37}Ar source [97] was taken after SR0 (marked in light blue), providing mono-energetic ER events at ~ 2.8 keV. This dataset is used for characterization of the detector response near the detection threshold, and is included in the ER band response model fit as well. During the entire data taking period, $^{83\text{m}}\text{Kr}$ calibration data was collected bi-weekly (regions marked in mint color), used for developing corrections of position and time dependent effects of the detector response as outlined in subsection 2.2.2, as well as for monitoring the stability of the S1 and S2 signal yields. Additionally, weekly LED calibration data was taken for PMT gain and afterpulse monitoring (see chapter 4). In total 17 PMTs were turned off or excluded from analysis in SR0, as previously discussed in subsection 4.3.2. The remaining PMTs' gains were stable within $< 3\%$ over SR0. The end of SR0 science data taking was determined based on the needed exposure for investigating an excess in ER events observed in XENON1T [98], the results of which are published in [85] but are not covered in this thesis. The period labelled 'Getter bypass mode' in Figure 6.1 was used as well for this investigation, with the aim of potentially confirming the presence of tritium as a possible explanation for the observed excess.

For the WIMP search, the resulting SR0 exposure after accounting for deadtime, veto times and the fiducial mass, is $(1.09 \pm 0.03) \text{ t} \times \text{yr}$ (tonne-years), similar to the entire exposure of the XENON1T experiment of $1.0 \text{ t} \times \text{yr}$ [58]. The full data analysis chain was developed and finalized while the WIMP signal region was blinded, avoiding human-induced biases in the search results. The blinding region was defined in the cS1-cS2 space and derived on quantiles from the ER and NR band calibration datasets, with cS2 above the -2σ quantile of the NR band, and below the median of the NR band or the -2σ quantile of the ER band, depending on which is higher. The

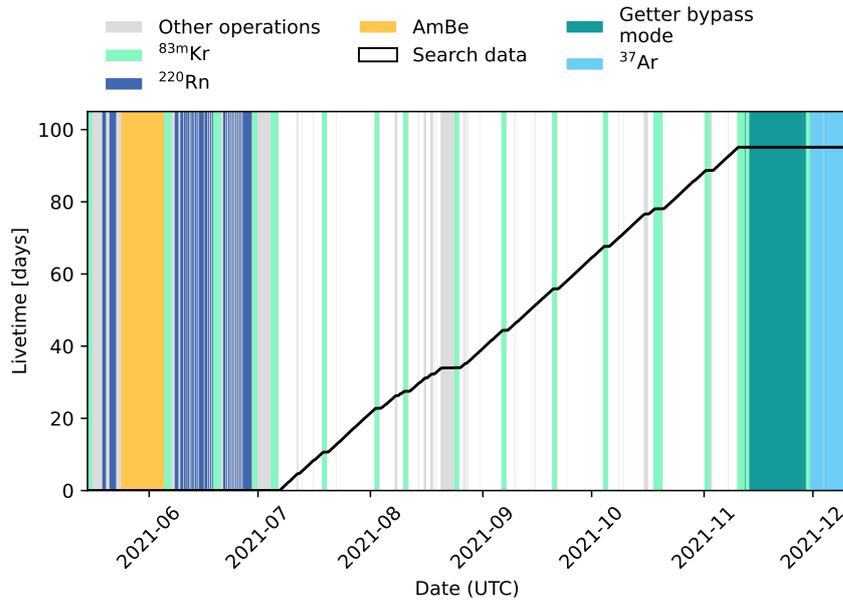


Figure 6.1.: Evolution of SR0 livetime during data taking, shown by the black line. The white regions labelled ‘search data’ correspond to the times of active livetime acquisition, with a total livetime of 97.1 days. Figure courtesy of the XENON Collaboration.

ER band in the WIMP search data was already unblinded for the ER physics searches in [85].

Signal reconstruction, corrections and data selections are the main subject of an upcoming publication [86], and are also described in [11, 84, 85]. Event positions are reconstructed from the PMT hit patterns of S2 signals in the top array, using neural network models trained on optical simulations while accounting for off or excluded PMTs and effects from PMT afterpulses (see chapter 4) using WFSim. The uniformly distributed $^{83\text{m}}\text{Kr}$ calibration data is used to correct reconstructed event positions for a radial component of the electric field, while S1 and S2 signals are corrected for the position-dependent effects described in subsection 5.3.2.

Data selection criteria are applied in order to remove unphysical events, not properly reconstructed events, and events with a high probability of originating from background sources, such as MS events. These data quality cuts are mainly based on the shape of S1 and S2 signals (S1 and S2 width cuts), event topology (S2 single scatter cut), expected PMT hit patterns in the arrays, as well as expected ratios between signal areas detected between the two arrays. The cuts are developed on calibration data as well as the full-chain simulations including WFSim, which is validated to match well with data in all relevant signal area and shape parameters. The data quality cuts are described in [85] as well as in [84]. The background from accidental coincidences (AC) between S1 and S2 signals is reduced by a dedicated cut based on a gradient boosting decision tree (GBDT) classifier, using z -position, S2 area and shape parameters as input features. However, the regions around the perpendicular wires on the gate and anode electrodes (see subsection 2.3.1) show variations in the drift and extraction field that lead to a not well-understood signal shape for S2s. Therefore, the AC cut can not be applied in this region, and a less strict, data-derived S2 width cut is used, leading to a higher AC background rate in this region. Consequently, the TPC is segmented into two regions for the further analysis, near the perpendicular wires (± 4.45 cm around each wire in the xy -plane) and outside, called ‘near-wires’ and ‘far-wires’. The two regions are correspondingly treated separately in the statistical inference. The fiducial volume for the WIMP search (see Figure 5.6) was optimized to reduce backgrounds from radioimpurities in detector materials and so-called surface events, while maximizing exposure, resulting in a LXe fiducial mass of (4.18 ± 0.13) t after accounting

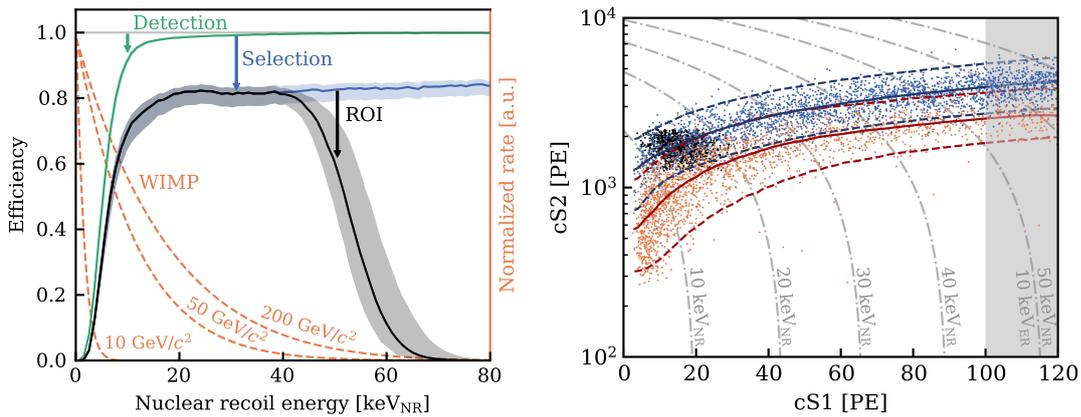


Figure 6.2.: (Left) Efficiency for detection (green) and event selection (blue) in dependence of NR energy. The total efficiency (black) drops off above about $50 \text{ keV}_{\text{nr}}$ due to the upper ROI limit of $\text{cS1} < 100 \text{ PE}$, containing close to 100% of the expected WIMP recoil spectra, shown by the orange dashed lines for several WIMP masses. (Right) ^{220}Rn ER (blue) and AmBe neutron NR (orange) band calibration data. Solid and dashed lines show the median and $\pm 2\sigma$ contours from the respective best-fit models. ^{37}Ar data is shown in black. Grey lines indicate constant reconstructed NR energy (keV_{nr}). Both figures from [84].

for the CIV.

The energy scale calibration for the detector was performed using the Doke-plot method (see subsection 2.2.3), using the mono-energetic lines of $^{83\text{m}}\text{Kr}$ (41.6 keV) and ^{37}Ar (2.8 keV), as well as two lines from the isomeric xenon isotopes $^{131\text{m}}\text{Xe}$ (163.9 keV) and $^{129\text{m}}\text{Xe}$ (236.2 keV) that are produced during AmBe calibrations via neutron activation. The resulting values for the scintillation and ionization gains are $g_1 = (0.151 \pm 0.001) \text{ PE}/n_\gamma$ and $g_2 = (16.5 \pm 0.6) \text{ PE}/n_e$, assuming an average work function of $W = 13.7 \text{ eV}/n_q$ for the LXe.

Figure 6.2 (left) shows the overall signal efficiency curve for the WIMP search inside the FV, overlaid with the expected differential recoil spectra for several WIMP masses. The total efficiency is a product of the detection efficiency, which is dominated by the S1 reconstruction efficiency that drops to zero for low energies due to the 3-fold PMT coincidence requirement, the selection efficiency from the data quality cuts, and the ROI definition. The latter is set to contain almost all expected signal events, as seen in the figure. The plateau of the efficiency curve is at about 80%, with a range where the total efficiency is above 10% of [3.3, 60.5] keV_{nr} .

Figure 6.2 (right) shows the ER and NR band calibration data from the ^{220}Rn and AmBe sources, respectively, with the best-fit models from the BBF tool overlaid. The mono-energetic ER ^{37}Ar data is shown as well in black. For both ER and NR bands, the median and $\pm 2\sigma$ contours are shown by the solid and dashed lines, respectively. The so-called ER leakage fraction, defined as the fraction of ^{220}Rn ER events that fall below the NR band median, is $1.1 (-0.3) (+0.2) \%$.

6.2. Signal and Background Models

The analysis space in the WIMP search is the $\text{cS1-cS2-}r$ space, where the radial coordinate r is included to leverage the position dependence of background components originating from the detector materials or the wall in the inference. In the signal space cS1-cS2 , the WIMP region of interest (ROI) is defined by cS1 in $[0, 100] \text{ PE}$ and $\log_{10}(\text{cS2}/\text{PE})$ in $[2.1, 4.1]$.

To perform the final statistical inference, we require both the shape and rates of the signal and all background components in the analysis space. These correspond to the probability density

distributions (PDFs) of events from each component, which we call templates, which are scaled by the expected event rates from the respective component. As already briefly introduced in [section 5.2](#), the templates for the ER background and for the NR signal and background components are produced from the best-fit ER and NR response models from the BBF tool. This is achieved by forward-folding the expected recoil energy spectra into the fitted detector response models, using the fast-toy MC simulation in BBF with the best-fit parameters from the respective calibration fits. Two further background components originating from detector artifacts are included in the background model for the WIMP search as well, these are the accidental coincidence (AC) background and the surface (or 'wall') background. Since their physical origin is not well understood, they can not be modeled from first principles via BBF simulations, and are instead modeled empirically in a data-driven approach. The templates in the signal space, cS1-cS2, for all background components in the WIMP search are shown in [Figure 6.3](#), together with the expected signal template of a 200 GeV/c² WIMP, and are briefly described in the following. The expected background rates for all components can be found in [Table 6.1](#).

WIMP signal

The distributions of WIMP signal events in the cS1-cS2 space are produced from the expected differential recoil energy spectra for WIMPs of a given mass. The recoil energy spectra are calculated following [Equation 1.1](#), using the standard halo model (SHM) parameters as recommended in [\[7\]](#), resulting in the rate-normalized energy spectra shown in [Figure 6.2](#) (left). Signal templates are produced from these spectra with the BBF fast-simulations, for a range of WIMP masses between 6 and 500 GeV/c². The contours of a 200 GeV/c² WIMP are shown in [Figure 6.3](#) in grey. An uncertainty on the WIMP signal rates is derived from the uncertainties of the NR response model parameters, using the posterior distributions from the BBF fit, resulting in a relative uncertainty of about 10 % on the signal rate for a 50 GeV/c² WIMP, increasing to about 30 % for smaller WIMP masses of 6 GeV/c² [\[59\]](#) due to a stronger impact of model uncertainties near the detection threshold¹. The uncertainty of the shape of the NR band arising from uncertainties in the NR response model is negligible compared to the rate uncertainty, due to the small number of expected NR events from both background and potential signal.

NR backgrounds

The components modeled in the NR background are radiogenic neutrons (RG) from (α , n) reactions or spontaneous fission processes in detector materials, as well as the coherent and elastic scattering of solar neutrinos on xenon nuclei (Coherent Elastic ν Nucleus Scattering, CE ν NS). Similar to the NR band calibration fit, the RG background templates are produced with GEANT4-based inputs that include MS-merging tags from the full-chain simulations as the basis for the fast-simulations, using the best-fit parameters from the BBF NR fit.

The full-chain simulations for neutrons were extensively validated for this purpose, as well as for the estimation of the resulting neutron background rate. Using the radioactive impurity levels determined in the material screening campaigns of both XENON1T and XENONnT [\[18, 99\]](#) for each detector component, the full-chain simulations allow for deriving the expected neutron background rate in the WIMP search region, including the spatial distribution and MS selection and merging effects. In particular, two components of the RG neutron background are modeled separately, the 'normal' RG background and the RG-X component, corresponding to neutron MS events where at least one of the scatters happens in a charge-insensitive region of the detector, mainly below the cathode or in the CIV close to the TPC wall (see [chapter 5](#)). The simulations-derived neutron background rate was validated by comparing simulations with

¹For example, larger photon yields at low energies lead to quickly increased event rates due to a larger fraction of the exponentially increasing WIMP spectra making it above the threshold.

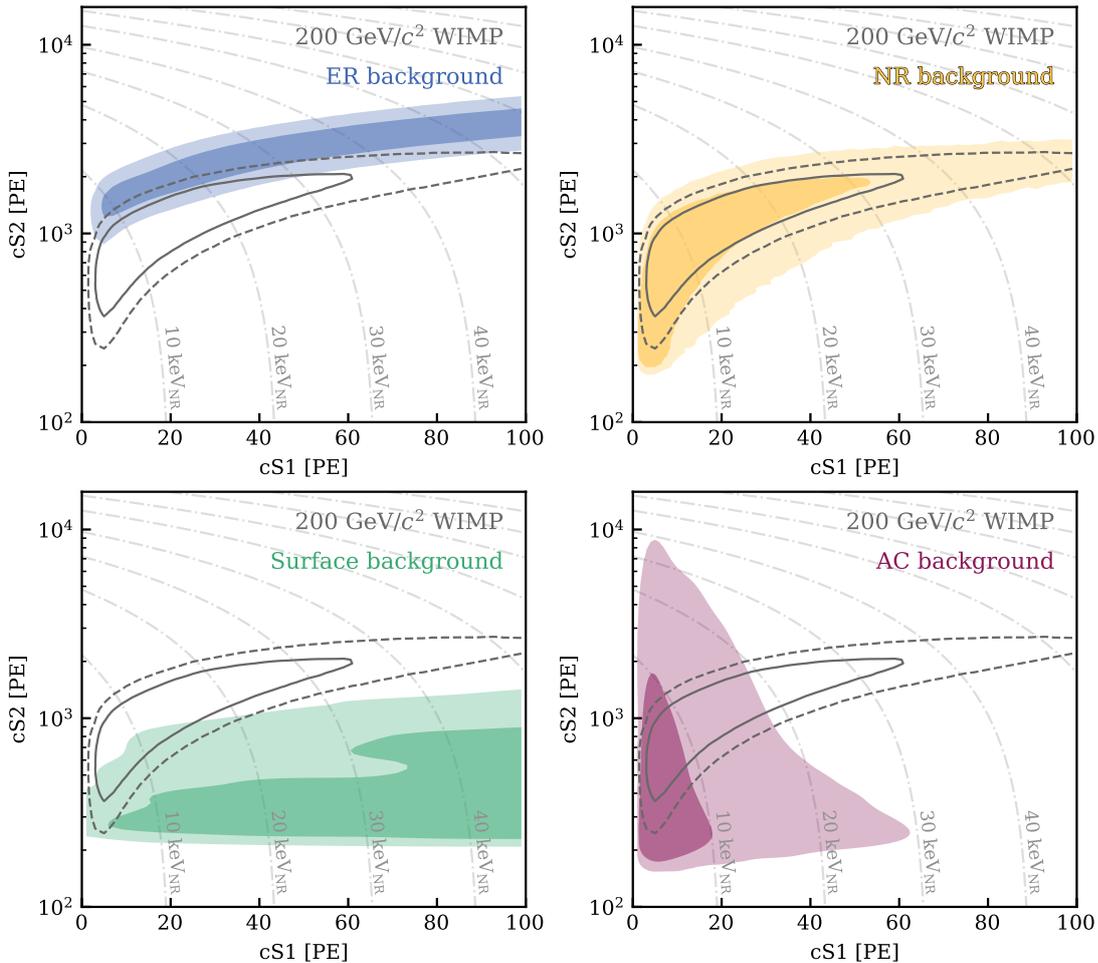


Figure 6.3.: Probability density distributions of the background components used in the WIMP statistical inference, shown here in the $cS1$ - $cS2$ space. The 1σ and 2σ contours of each background are indicated by the darker and lighter colored regions, as labeled in each panel. The contours of the NR signal model of a $200 \text{ GeV}/c^2$ WIMP are shown for comparison by the solid and dashed grey lines. Light grey dash-dotted lines indicate contours of constant deposited energy assuming NR events. Figure from [51].

data-derived estimates, in particular by matching the ratio of MS to (non-rejected) SS events between AmBe data and simulations, giving a ratio of 2.2 [51]. Additionally, the neutron tagging efficiency of the NV, determined from the AmBe TPC-NV cross-calibration data to be $(53 \pm 3)\%$ [52, 84], also matches well between data and simulations. Given the good agreement between data and simulations in these studies, a sideband unblinding was performed for NV-tagged NR events in the WIMP search data, that initially showed good agreement with the expected neutron rate from simulations. After the full unblinding of the WIMP search data, however, a mistake in the NV tagging window definition was discovered, that, after fixing it, made apparent a discrepancy between the simulations-derived neutron rate prediction and the rate observed in the NV-tagged data, which was about a factor 6 higher than predicted. The cause for the discrepancy is under investigation, but is likely due to an underestimated material contribution in the simulations. Thus, the neutron background rate for the SR0 WIMP search uses the data-driven estimate only, resulting in $1.1 (-0.5) (+0.6)$ expected events in the ROI in the SR0 exposure. Note that apart from fixing the mistake in the veto window definition, no other changes were made in the analysis chain as the search data was already unblinded. The muon-induced neutron rate, originating from cosmogenic muon spallation reactions in

surrounding detector materials, is estimated to be about 0.01 events in SR0 after applying the MV veto selection, even without the additional tagging from the NV, and was therefore not included in the background model.

The CE ν NS NR background originates from solar, atmospheric and diffuse supernova (DSN) neutrinos, dominated by ^8B solar neutrinos that produce an NR energy spectrum almost identical to a $\sim 6 \text{ GeV}/c^2$ WIMP [15]. Similar to WIMPS, CE ν NS events are expected to be single scatter events with a spatially homogeneous distribution in the TPC, due to the small cross-section of the CE ν NS process on xenon nuclei. The expected event rate is constrained from neutrino flux measurements [100], resulting in (0.23 ± 0.06) expected events in SR0, with the uncertainty dominated by the uncertainty on the NR response model.

The NR background contours in Figure 6.3 show the sum of all NR components for illustration, with the RG and RG-X neutron contributions the dominant ones, however they are treated separately in the statistical inference. The expected background events for all components separately are stated in [51].

ER model and backgrounds

Similar to the NR response model, the ER band shape is derived by fitting the ER response model simultaneously to the ^{220}Rn ER band calibration and a fraction of ^{37}Ar calibration data², using the BBF tool. This fit was performed by Zihao Xu of Columbia University, and is detailed in ref. [51]. The gaseous ^{220}Rn distributes homogeneously in the LXe volume, with the β -decay of the daughter isotope ^{212}Pb providing SS ER events with approximately flat energy spectrum in the ROI below about $10 \text{ keV}_{\text{ee}}$. While ER and NR model share the detector reconstruction model as described in subsection 5.3.2 for SS events, the LXe emission is different. We use the same ER quanta emission model as in XENON1T, described in both [59] and [51]. An additional background model is added in the fit, accounting for accidental coincidence (AC) events observed in the ^{220}Rn data, caused by a wrong pairing of lone S1 and S2 signals in the high-rate calibration dataset. The ER band fit result was validated by an equiprobable-binned GOF test (as in section 5.4), giving a p-value of 0.42. In contrast the NR band, the uncertainty on the shape of the ER band has a large impact on the WIMP search results, with events at the lower edge of the ER band leaking into the WIMP signal region. The ER model parameterization, same as the NR one, however includes a large number of parameters, too many to propagate their best-fit posterior distributions into the WIMP search. Therefore, a principal component analysis (PCA) [101] was performed on the ER band fit results, selecting the two linear combinations of ER model parameters that result in variations of the ER band shape with the largest impact on the WIMP sensitivity. These two PCA components are used as ER shape nuisance parameters in the inference.

For the ER background model, we take into account the ER background contributions from ^{222}Rn , with the β -decaying daughter isotope ^{214}Pb giving a very similar, flat energy spectrum as ^{212}Pb , as well as the β -decays from the intrinsic ^{85}Kr and ^{136}Xe isotopes. Additionally, solar neutrino-induced ER events and the γ -ray background from the detector materials are included in the background model. The total ER background spectrum being dominated by the ^{222}Rn and ^{85}Kr contributions, it is approximately flat in the ROI [85], and is treated as a single component in the statistical inference. The ER template shape resulting from the ER band fit is shown in Figure 6.3 in blue (top left). The ER background rate is unconstrained in the inference since it was already unblinded previously, with 134 events observed in the ER band in the ROI below about $10 \text{ keV}_{\text{ee}}$ before NR unblinding.

²The downsampling of the ^{37}Ar dataset in the ER fit is motivated by wanting to avoid an overconstraint of the ER model to the small and narrow energy range of ^{37}Ar compared to the full ER band dataset from ^{220}Rn .

	Nominal	Best Fit	
	ROI		Signal-like
ER	134	135_{-11}^{+12}	0.92 ± 0.08
Neutrons	$1.1_{-0.5}^{+0.6}$	1.1 ± 0.4	0.42 ± 0.16
CE ν NS	0.23 ± 0.06	0.23 ± 0.06	0.022 ± 0.006
AC	4.3 ± 0.9	$4.4_{-0.8}^{+0.9}$	0.32 ± 0.06
Surface	14 ± 3	12 ± 2	0.35 ± 0.07
Total Background	154	152 ± 12	$2.03_{-0.15}^{+0.17}$
WIMP	-	2.6	1.3
Observed	-	152	3

Table 6.1.: Expected and observed numbers of events in SR0, for each background component. The column ‘Nominal’ gives the expected event numbers and uncertainties in the WIMP ROI, while the ‘Best fit’ column shows the event numbers resulting from a free fit to the unblinded data with a signal component of a 200 GeV/c² WIMP, both in the ROI and for the ‘signal-like’ region indicated by the orange dashed contour in [Figure 6.4](#). From [84].

Detector Backgrounds

In addition to the ER and NR model based background templates, we model two background components from detector artifacts, the accidental coincidence (AC) background and the surface background. Because the physical origin of these backgrounds is not well-understood from first principles, the templates are produced empirically using data-driven methods.

The AC background originates from wrongly paired isolated S1 and S2 signals and thus corresponds to unphysical events. The isolated signals forming the AC background are mostly small, caused by e.g. pile-up of single photoelectrons and PMT dark counts, and therefore mainly affect the low-mass WIMP limits. The AC model in cS1-cS2 space, shown in [Figure 6.3](#) in purple (bottom right), is produced by sampling isolated S1 and S2 signals from the data and randomly pairing them into events. The method of AC rate prediction was validated on various sideband data, including the ³⁷Ar and ²²⁰Rn calibration datasets, as well as AC sideband data outside of the WIMP ROI defined by inverting the anti-AC cuts, showing good agreement between the predicted and observed rates.

The surface background originates from β -decays of ²¹⁰Pb, another daughter isotope in the ²²²Rn decay chain, that plates out on the PTFE surfaces of the TPC walls. As a result of the proximity to the wall during the decay, a loss of ionization electrons from the S2 signal occurs, leading to a reduced S2 signal for these events with a strong radial dependence towards the wall. The distribution of the surface background in the spatial dimensions r and z is extracted from surface events caused by the α -decay of ²¹⁰Po further down the ²²²Rn decay chain, which are easily identifiable by their S1 signals from the large energy depositions of the α -particles. The cS2-cS2 distribution was derived using a kernel density estimation (KDE) method, utilizing surface events in the WIMP ROI but reconstructed outside the TPC. The model was validated on several sidebands, including the region between TPC wall and WIMP FV, as well as the surface events in the WIMP ROI below the blinding region. The surface background model in the cS1-cS2 space is shown in [Figure 6.3](#) in green (bottom left), with a total expected rate of (14 ± 3) events in SR0.

6.3. Statistical Inference and WIMP Results

The WIMP statistical inference is based on the profile likelihood ratio, with the likelihood functions is constructed from the probability density distributions of the signal and background components in the analysis space of cS1-cS2 space, as well as the radial coordinate r . The test statistic is defined by

$$q(\sigma) = -2 \ln \left(\frac{\mathcal{L}(\sigma, \hat{\boldsymbol{\theta}})}{\mathcal{L}(\hat{\sigma}, \hat{\boldsymbol{\theta}})} \right), \quad (6.1)$$

where σ is the WIMP-nucleon cross section, and $\boldsymbol{\theta}$ are nuisance parameters. A single hat indicates best-fit parameters from the global likelihood maximum, while the double hat on $\boldsymbol{\theta}$ indicates the best-fit parameters for a maximum likelihood given a fixed cross-section σ . The total likelihood function is the product of four likelihood terms,

$$\mathcal{L}(\sigma, \boldsymbol{\theta}) = \mathcal{L}_{\text{farwires}}(\sigma, \boldsymbol{\theta}) \times \mathcal{L}_{\text{nearwires}}(\sigma, \boldsymbol{\theta}) \times \mathcal{L}_{\text{ERcalib}}(\boldsymbol{\theta}) \times \mathcal{L}_{\text{anc}}(\boldsymbol{\theta}). \quad (6.2)$$

The first two terms are extended unbinned likelihood functions of the WIMP search data, which is separated into the two regions far and near the perpendicular wires due to the different AC background levels in those regions. They are given by

$$\mathcal{L}_{\text{wires}}(\sigma, \boldsymbol{\theta}) = \prod_i^N \left(\sum_c \frac{\mu_c(\sigma, \boldsymbol{\theta})}{\mu_{\text{tot}}(\sigma, \boldsymbol{\theta})} \times f_c(\vec{x}_i | \boldsymbol{\theta}) \right) \times \text{Poisson}(N | \mu_{\text{tot}}(\sigma, \boldsymbol{\theta})), \quad (6.3)$$

with the product running over all N events \vec{x}_i in the respective region, each event being a tuple in the analysis space of cS1 and \log_{10} cS2, as well as r for the far wire region. The sum runs over each signal and background component c with template PDFs f_c in the analysis space, as shown in cS1- \log_{10} cS2 in [Figure 6.3](#), with the expected number of events μ_c per component and the total expected number of events μ_{tot} . The nuisance parameters $\boldsymbol{\theta}$ are scaling parameters for the rates of each background component and of the ^{220}Rn calibration dataset, the ER shape parameters, and a WIMP signal efficiency parameter accounting for the uncertainties in the NR rate from The term $\mathcal{L}_{\text{ERcalib}}(\boldsymbol{\theta})$ is a likelihood function of the ^{220}Rn calibration data and the PCA shape parameters from the ER model, allowing for a variation in the ER background model in the WIMP search data that is constrained by the calibration data, by simultaneously fitting the shape parameters to both datasets. The background rate parameters are constrained by the ancillary constraint term in the total likelihood, $\mathcal{L}_{\text{anc}}(\boldsymbol{\theta})$, which is a product of gaussian constraint terms with the nominal rates and uncertainties, derived as summarized in [section 6.2](#).

The nominal rates and uncertainties of the background components in the WIMP search region are summarized in [Table 6.1](#). After unblinding the data, a total of 152 events were observed in the WIMP search ROI, shown in [Figure 6.4](#) in the cS1-cS2 analysis subspace, 16 of which are in the previously blinded region. This is consistent with the expected number of 154 background events. A larger number of ER-leakage events is observed for cS1 above about 50 PE. Additionally, 13 of the 16 unblinded events appear to cluster in the upper right half of the xy -plane of the detector. However, none of the calibration datasets, corrections, or data selection efficiencies show a spatial asymmetry that could explain the observed clustering. Nevertheless, equiprobable-binned goodness-of-fit tests (similar to the test shown in [Figure 5.16](#)) for the best-fit results of both a background-only and background-plus-WIMP-signal fits give p-values above ~ 0.6 , indicating that no significant mismodeling is observed in the data.

We then perform the profile likelihood ratio test for each WIMP mass signal model. The best-fit result for the 200 GeV/ c^2 WIMP signal model is given representatively in [Table 6.1](#) (column 'Best Fit') and are indicated in [Figure 6.4](#), with the fractions of the component's best-fit probability distribution at each event's position shown in pie-charts. No significant excess is observed, with the WIMP discovery p-values above 0.2 for all tested WIMP mass signal hypotheses. The resulting limits on the spin-independent WIMP-nucleon cross sections ($\sigma_{\chi,n}$ in [Equation 1.4](#)) are shown in [Figure 6.5](#). The expected range of sensitivity, shown by the green

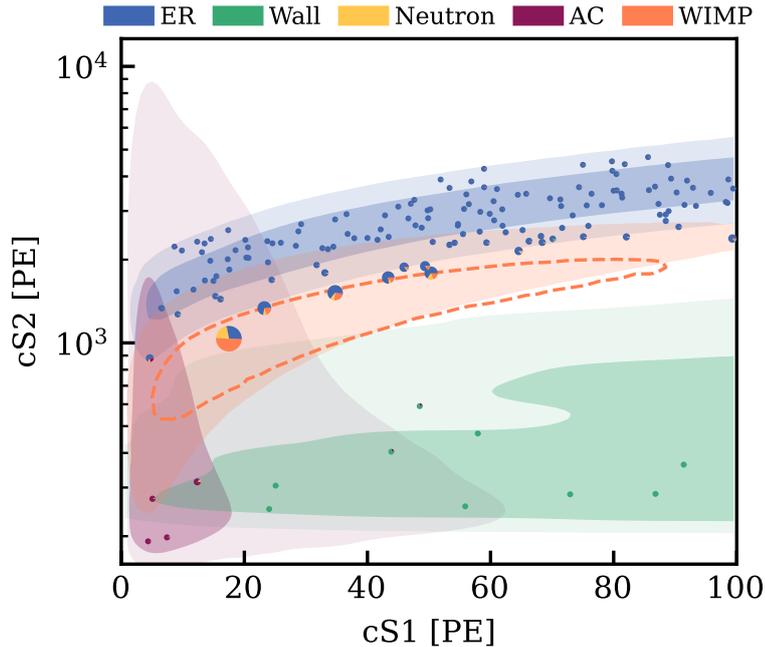


Figure 6.4.: Unblinded data in the DM WIMP search region in the cS1-cS2 space. The darker and lighter shaded colored regions show the 1σ and 2σ contours of the background templates, as well as the 2σ contour of a $200 \text{ GeV}/c^2$ WIMP. Note that the NR background template is not shown as it largely overlaps with the WIMP contour, see also Figure 6.3. All events are represented by pie-charts, indicating the best-fit fractions of each component’s probability distribution at the event’s cS1-cS2- r position, from a fit including a free $200 \text{ GeV}/c^2$ WIMP signal component. The size of each pie-chart is proportional to the fraction of the best-fit WIMP signal component at that position – note that this is NOT equivalent to the ‘probability’ of an event being a WIMP. The orange dashed contour indicates a most ‘signal-like’ region, defined by containing the 50% of the WIMP signal with the largest possible signal-to-background ratio. Figure from [84].

(1σ) and yellow (2σ) shaded regions commonly termed a ‘Brazil-band’, is produced via toy MC experiments of the background-only model, determining the 90% confidence level (C.L.) upper limits on the WIMP-nucleon cross section using the Feldman-Cousins method. Due to the non-asymptotic nature of the profile likelihood ratio test for small expected signal rates, the 90% critical region threshold is pre-determined as a function of number of expected signal events from toy MCs that include an injected signal. The median upper limit sensitivity is indicated by the boundary of the green shaded region in Figure 6.5, while the solid black line shows the observed limit derived from the best-fit results of the unblinded data to each WIMP mass signal model. In order to avoid setting limits in regions far beyond the expected sensitivity of the experiment, which can happen due to spurious underfluctuations in the data (or an overestimated background rate), we apply a ‘power constraint’ on the observed limit, following the PCL (power-constrained limits) method [102]. Here the upper limit is constrained by requiring a minimum rejection power for the hypothesis one wants to exclude, which translates directly to a constraint on the quantile of the expected limit distribution. We conservatively choose a minimum rejection power of 0.5, resulting in constraining the observed limit to median of toyMC upper limits. For reference, the unconstrained limit is shown as a dashed black line in Figure 6.5. Compared to the previous upper limit from the XENON1T experiment [58], the XENONnT SR0 exclusion limit is lower, with the lowest upper limit of $2.58 \times 10^{-47} \text{ cm}^2$ at a $28 \text{ GeV}/c^2$ WIMP mass. For WIMP masses above about $100 \text{ GeV}/c^2$, the upper limit becomes a linear function of the WIMP mass, as the expected number of signal events decreases with increasing WIMP mass due to fix local WIMP DM density ρ_χ , given by $6.08 \times 10^{-47} \text{ cm}^2 \cdot M_{DM}/(100 \text{ GeV}/c^2)$, a factor of 1.7 compared to XENON1T.

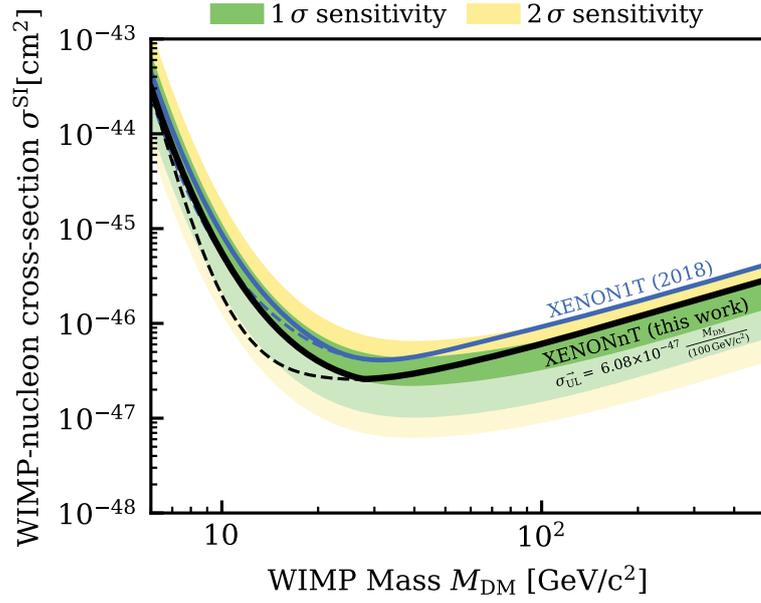


Figure 6.5.: Upper limit on the spin-independent (SI) WIMP-nucleon cross-section in dependence of the WIMP mass from the first science run SR0 of the XENONnT experiment. The green and yellow bands indicate the 1σ and 2σ expected sensitivity bands, respectively. The solid black line shows the 90% C.L. upper limit with a power-constraint applied to the median sensitivity, while the dashed black line shows the unconstrained limit for reference. The solid (dashed) blue line shows the upper limits from XENON1T [58] with the same (without) power-constraint applied. Figure from [84].

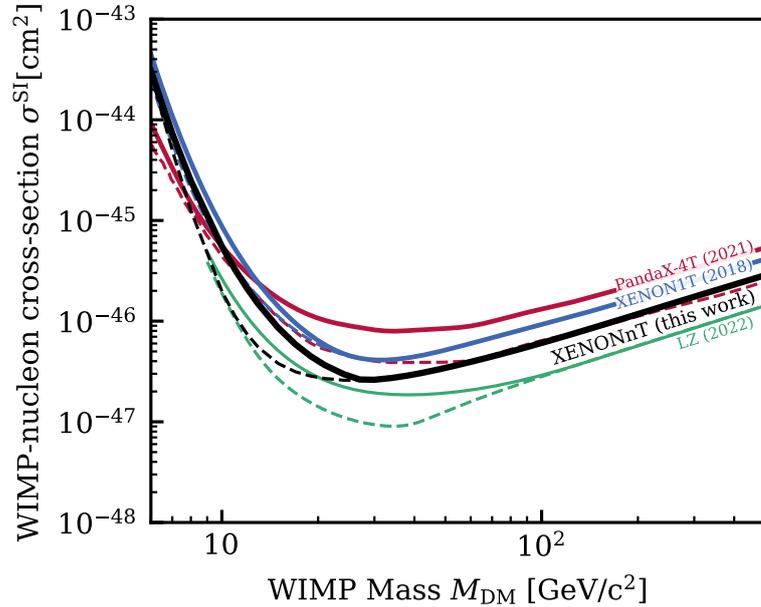


Figure 6.6.: Comparison of the XENONnT 90% C.L. upper limit on the SI WIMP-nucleon cross-section to the previous result from XENON1T [58] (blue), and recent results from other xenon DM experiments, PandaX-4T [103] (red) and LZ [104] (green). The solid lines show the limits with the same power-constraint applied as in the XENONnT analysis, while the dashed lines show the limits as published. Figure from [84] (updated with published results from [104]).

In [Figure 6.6](#), the results from XENONnT are compared to recent exclusion limits from other LXe DM experiments, i.e. PandaX-4T [\[103\]](#) and LZ [\[104\]](#). For better comparison, the limits from the other experiments are shown with the same power-constraint as applied here in the solid lines, while the dashed lines show the limits as published, using a power-constraint at the -1σ level in both cases. The limit from LZ shows a large underfluctuation below the median sensitivity at low WIMP masses below about $40 \text{ GeV}/c^2$, while the limit from PandaX-4T underfluctuates across a large range of moderate to high WIMP masses. Note that in contrast to the XENONnT procedure, these limits were derived from non-blind analyses. Within these considerations, the XENONnT SR0 results can be considered competitive.

6.4. Conclusions and Outlook

In this chapter, the results from the first WIMP search of the XENONnT experiment, SR0, were summarized. The results show no significant excess of events above background expectations in the WIMP search region, leading to competitive exclusion limits on the spin-independent WIMP-nucleon cross-section.

Data taking with XENONnT currently continues, with the second science run SR1 started in May 2022. Due to the operation of the Radon distillation column in an enhanced mode in SR1, a lower radon background activity concentration was achieved, a factor of 2 lower than in SR0. Additionally, the water in the NV is loaded with Gadolinium sulfate, increasing the neutron tagging efficiency and consequently reducing the neutron background rate. Future WIMP search results from XENONnT are therefore projected to have WIMP sensitivity beyond the current best limits, exploring as-yet-untested regions in the WIMP parameter space for the first time.

Chapter 7.

Measurement of ER yields at low energy and fields in HeXe

As we have seen in the previous chapters, understanding the response of a xenon TPC to low-energy electronic recoils (ERs) is crucial for the interpretation of Dark Matter search results, especially in the regime of low drift fields where the discrimination power between ERs and nuclear recoils (NRs) is reduced. However, little literature data is available for the low-field regime for both ER and NR, necessitating the use of extrapolation when modeling and simulating the response of LXe to different particle interactions, using tools such as the NEST [71] framework. With increasing detector sizes, large-scale xenon TPCs have continuously failed to meet their design goals in electric field strength due to the increasing technical challenge of producing stable electrode structures on the scale of meters, requiring ever higher cathode voltages for increasing drift lengths while maintaining high field homogeneity and optical transparency. Therefore, measurements of the light and charge yields in LXe at low drift fields are of great interest for the general understanding of the signal response of xenon dual-phase TPCs, in particular for XENONnT as well as for the next generation of xenon TPCs, informing the development of more accurate signal models and simulations.

In this chapter, a measurement of the absolute photon and charge yields for ERs in a xenon TPC at low drift fields is presented. The measurement was performed with the Heidelberg Xenon (HeXe) setup at the Max-Planck-Institut für Kernphysik (MPIK) in Heidelberg, Germany, featuring a small-scale xenon dual-phase TPC that is particularly well-suited for low-field measurements down to about 10 V/cm while maintaining reasonable field homogeneity throughout the TPC volume. The field homogeneity is confirmed by 3D electric field simulations as well as previous measurements of high-energy α and $^{83\text{m}}\text{Kr}$ interactions spanning large field ranges [105, 106]. In this work, the photon and charge yields of ERs were measured with an external γ -ray source, employing the close time coincidence between γ -ray Compton scatters in the TPC and their consequent detection in a second, external γ -ray detector placed at given scattering angles, allowing for the selection of events with known energy deposition in the TPC based on the kinematics of the scattering process.

This chapter is structured as follows: The HeXe TPC and the Compton coincidence measurement setup are described in [section 7.1](#). In [section 7.2](#), the characterization of the response of the external γ -ray detector, as well as of the HeXe TPC is described, encompassing the development of data selection criteria, signal corrections, and the determination of the ionization and scintillation gains g_1 and g_2 . The details and results of the Compton coincidence measurement are presented in [section 7.3](#), comprising a timing cross-calibration between the two detectors, the selection of low-energy ERs in the TPC, and the resulting photon and charge yields, along with an estimation of the field homogeneity and expected energy distributions in the TPC. The chapter finishes with conclusions and an outlook in [section 7.4](#).

7.1. The HeXe setup and measurement overview

The full HeXe setup, located in a laboratory at the Max-Planck-Institut für Kernphysik (MPIK) in Heidelberg, Germany, is shown in [Figure 7.1](#).

The HeXe dual-phase TPC is housed inside a double-walled cryostat in the center of the setup, indicated by the green box. A gas handling system (GS), indicated by the orange box and schematically drawn in [Figure 7.2](#), is used to fill the TPC with xenon gas from a stainless steel storage bottle. A cold head, consisting of a large copper surface cooled by a helium pulse tube refrigerator (PTR), is mounted on top of the cryostat and used to liquify the xenon gas by condensing it on the cold copper surface. The LXe is then guided via a funnel connected to a tube into the TPC volume. During operation, the xenon inventory is continuously circulated between cryostat and the GS via a recirculation pump (KNF double-diaphragm). LXe is extracted and evaporated from the reservoir at the bottom of the cryostat, guided through the GS, and returned to the gas phase at the top of the cryostat to be liquified again. In the GS, the xenon flows through a hot zirconium getter (SAES MonoTor PS3) in order to maintain chemical purity and thereby avoid signal attenuation in the TPC caused by photon absorption and electron attachment to electronegative impurities. The GS also features two ports for internal calibration sources that can be used to calibrate the TPC's light and charge response. Guiding the xenon flow through an attached gaseous source trap allows for its mixing with the xenon gas, and eventually leads to a homogeneous distribution of the calibration source within the TPC volume.

A Slow Control (SC) system using a custom LabVIEW [107] interface, shown in the blue box, is used to monitor detector parameters relevant for both the measurement and operational safety, such as the cryostat temperature and pressure, liquid xenon level in the cryostat, and the gas circulation mass flow and pressure in the GS. An alarm system is implemented to alert operators via Email and SMS in case of any parameter exceeding predefined thresholds, allowing for the stable, long-term and, most importantly, unsupervised operation of the setup up to several weeks and even months at a time. This is especially important for the measurement presented here, where the coincident event rate is very low at a few mHz (before data quality selections), resulting in the need of long measurement durations in order to reach sufficient statistics. At the end of a measurement run, or in case of emergency such as a cooling failure, the xenon inventory can be quickly recuperated into the storage bottle via cryogenic pumping. For this purpose, the bottle is constantly kept at cryogenic temperature by immersing it in liquid

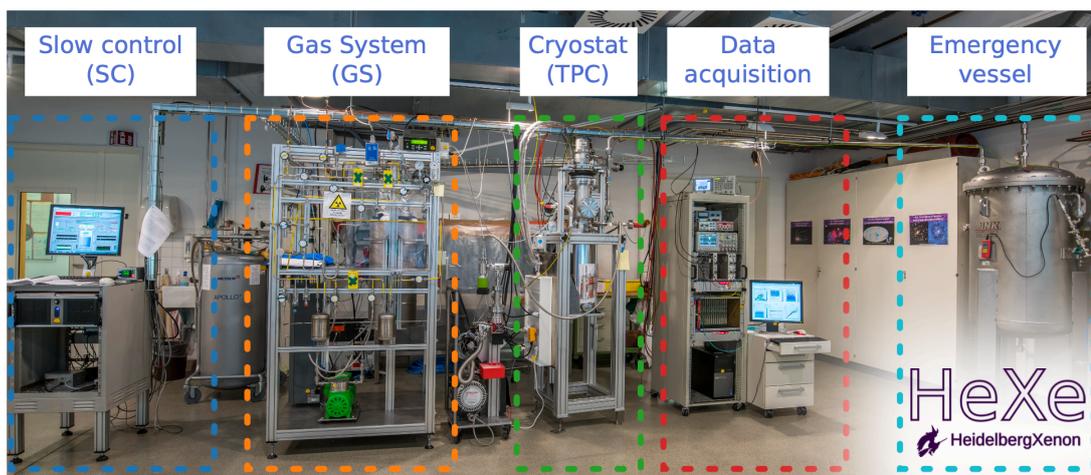


Figure 7.1.: Photograph of the full HeXe setup in the laboratory at the Max-Planck-Institut für Kernphysik (MPIK) in Heidelberg. The different subsystems are indicated by the labeled colored boxes, and are described in the text. Figure from [106].

nitrogen (LN). An automated nitrogen filling device connected to a large LN tank keeps the LN level above a set threshold, keeping the storage bottle constantly cryogenic. Additionally, a large emergency vessel seen in the cyan box on the right, is connected to the cryostat via a rupture disk, allowing to safely capture the entire, very expensive xenon inventory in case of a rapidly developing cryostat overpressure, or when timely xenon recuperation to the storage bottle is not possible. The SC system also features a Python program called HeXeSVM, developed in [106], that allows to monitor and control the high voltage (HV) supply modules, used for both establishing the drift and extraction fields in the TPC and for the potential biasing of the TPC's and the external γ -ray detector's photosensors. Both the LabVIEW and the HeXeSVM program continuously write and store the monitored parameter values to a PostgreSQL [108] database that can be accessed remotely for monitoring and data analysis, using a framework developed in [109].

Finally, the data acquisition system (DAQ) shown in the red box, is used to amplify, digitize and read out the analog signals from the TPCs' photosensors and the external γ -ray detector. Custom-built software is used to read out the available digitizers, store the data to disk, and process the recorded waveforms [109].

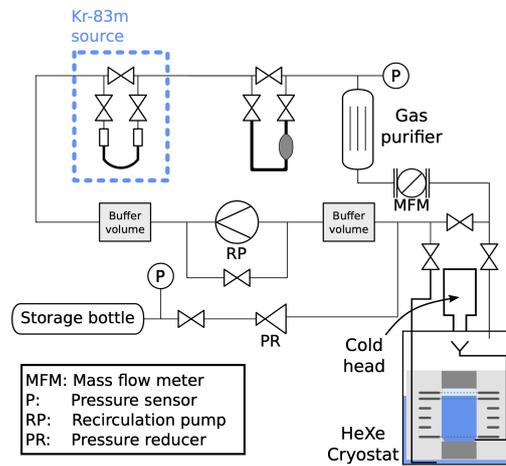


Figure 7.2.: Simplified schematic of the HeXe gas handling system (GS). Figure from [105].

7.1.1. The HeXe TPC

The heart of the HeXe setup is the HeXe dual-phase TPC. The technical drawing in Figure 7.3 shows slice-through view of the TPC and its feature components. The main body consists of three stackable, cylindrical PTFE support structure blocks made to fit the cryostat with an inner diameter of 20.1 cm closely, thereby reducing the necessity for excessive amounts of xenon. An inner cylinder of the central block forms the active TPC volume, measuring 5 cm in height and 5.6 cm in diameter. When filled with LXe, it holds a xenon mass of 345 grams in the active volume.

The TPC employs three electrodes, cathode, gate, and anode, to establish the electric drift and extraction fields. The electrodes are made of stainless-steel, each consisting of a circular holder supporting a hexagonal etched mesh grid with a narrow cell pitch of 1 mm (inner hexagonal cell spacing) and a wire thickness of 120 μm . While the narrow pitch is one of the reasons for the good field homogeneity in the TPC, it leads to a low optical transparency of the grids, estimated to be around 80% [105, 106] per grid. PTFE-insulated cables are attached to each grid, leading outside the cryostat via a custom-built HV feedthrough able to withstand applied voltages up to ± 6 kV on the anode and cathode. In addition to the electrodes themselves, three field shaping rings (FSRs) connected by a resistor chain of 1 G Ω per resistor are placed

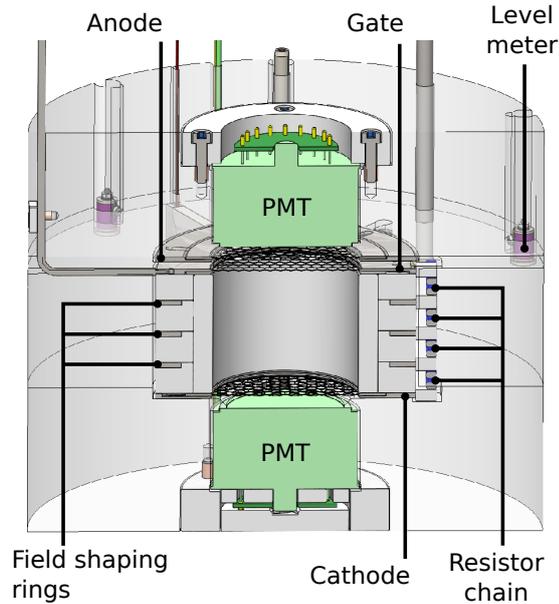


Figure 7.3.: Technical drawing of the HeXe TPC. Light grey indicates PTFE parts, with the main body formed by three stacked support structure block. An inner cylinder of the middle block forms the active volume with a height of 5 cm and a diameter of 5.6 cm, holding a LXe mass of 345 g. Two PMTs shown in green are positioned at the top and bottom of the TPC. Cathode, gate, and anode electrodes are shown in dark grey, with the field shaping rings (FSRs) connecting the cathode and gate via a resistor chain (blue). Figure from [105].

along the drift length, connecting the cathode and gate electrodes thereby increasing the field homogeneity.

The vertical distance between cathode and gate establishing the drift field is 50 mm (at room temperature), corresponding to the TPC height stated above. The distance between gate and anode is 5 mm, defining the extraction field region. During a measurement run, the liquid xenon level is maintained between anode and gate and measured by four capacitive level meters distributed along the exterior of the PTFE support structure blocks. Three of the level meters measure 1 cm in length and are vertically positioned to measure the liquid level in the extraction region between gate and anode, while the fourth level meter covers the entire drift region of the TPC from cathode to gate for monitoring purposes during the LXe filling procedure. Additionally, four Pt100 temperature sensors are positioned on the outside of the structure, one at the bottom of the lowest PTFE block and two situated along the outside of the blocks in order to measure the LXe temperature, with the fourth sensor placed on top of the upper block measuring the GXe temperature as well.

The S1 and S2 signals are detected by two Hamamatsu R6041-406 PMTs, positioned at the top and bottom of the TPC. They have round, 2-inch diameter windows and were specifically selected for high Quantum Efficiency (QE); both PMTs have QEs of about 40% for the xenon VUV scintillation light. In contrast to the PMT model used in XENONnT, they are of the multi-channel dynode type, enabling a good timing response (TTS ~ 0.75 ns), however at the cost of a reduced single PE resolution. Optical fibers positioned close to the inner PTFE cylinder are connected to a customized LED board housing a blue LED, allowing for regular calibration of the PMT gains throughout experimental runs.

7.1.2. Coincidence measurement overview and setup

The measurement of the light and charge yields for low-energy ERs in the HeXe TPC was performed using the coincidence method based on Compton scattering. Here, a high activity mono-energetic γ -ray source is placed in front of the TPC cryostat. The γ -rays emitted by the source may Compton scatter on atomic electrons of the LXe in the TPC, with the Compton electrons depositing their energy in an ER event. Following the well-known Compton scattering kinematics, for an incoming γ -ray of with energy E_γ , the energy of the scattered γ -ray E'_γ only depends on the scattering angle θ between the incoming and outgoing γ -ray,

$$E'_\gamma = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e c^2} (1 - \cos \theta)}, \quad (7.1)$$

where m_e is the electron mass and c the speed of light. Due to energy conservation, the energy of the Compton electron E'_{e^-} at a certain γ -ray scattering angle is fixed, and given by

$$E'_{e^-} = E_\gamma - E'_\gamma = E_\gamma \left(1 - \frac{1}{1 + \frac{E_\gamma}{m_e c^2} (1 - \cos \theta)} \right). \quad (7.2)$$

By placing a second, external γ -ray detector on the other side of the TPC at a desired scattering angle, outgoing γ -rays can be detected in coincidence with the energy deposition in the TPC. Utilizing a tight time-coincidence window between both detectors, combined with the requirement of observing the expected energy deposition E'_γ of the scattered γ -ray in the external γ -ray detector, allows for the selection of events with a known energy deposition from the Compton electrons in the LXe.

For the measurement in this work, the aim is to measure the absolute photon and charge yields of low-energy ERs in the range relevant in DM searches below about 15 keV_{ee}, in particular at low drift fields in a range close to the current XENONnT drift field of 23 V/cm. For this purpose, a γ -ray Compton coincidence measurement setup was developed and implemented. A sketch of the setup is presented in [Figure 7.4](#), showing both the view from a birds-eye perspective (top panel) and a side view (bottom panel) of the setup. The γ -ray source used for the measurement is a ^{137}Cs with a single, mono-energetic γ -ray line at 661.7 keV and an activity of about 1.2 MBq. It is positioned in front of the TPC (shown on the left side in the sketch) at a radial distance of 13.5 cm from the cryostat's outer wall, corresponding to a distance of 27.3 cm to the TPC center, and is surrounded by several layers of lead bricks (indicated by dark green rectangles in the sketch) both for radiation protection shielding and to collimate the γ -ray emission towards the TPC center. The external γ -ray detector is placed on the other side of the TPC, at a horizontal angle θ to the source 'beam'. The used device is a NaI(Tl) scintillation detector unit (SCIONIX type 76B76/3M-E1-X-NEG), consisting of a NaI(Tl) crystal of 3-inch (7.62 cm) diameter and length connected to a 3-inch window PMT (ETL Type 9305) with integrated voltage divider for HV application and signal readout. According to the producer, this particular unit has an energy resolution of 6.3% FWHM at 662 keV, corresponding to a relative gaussian peak width of $\sigma/E = 2.67\%$. It is positioned at a distance of about 47 cm from the cryostat, corresponding to a distance of about 60 cm from the TPC center.

The source and the NaI detector are placed on two high-load-capacity tables, with the source table being adjustable in height in order to be able to position the source, sitting on a 5 cm thick surface layer of lead covering the table, at the desired vertical position corresponding to the center of the TPCs drift length. Both tables are positioned and fixed to the item profile structure that holds the cryostat, thus ensuring the symmetry of the setup and allowing for a simplified alignment of the source and the NaI detector towards the TPC center. The high load capacity of the tables is necessary to support the high total weight of the lead bricks used for shielding and collimation, in the order of up to a few hundred kilograms.

Both the source and the NaI detector's opening angles/field of view are collimated in the

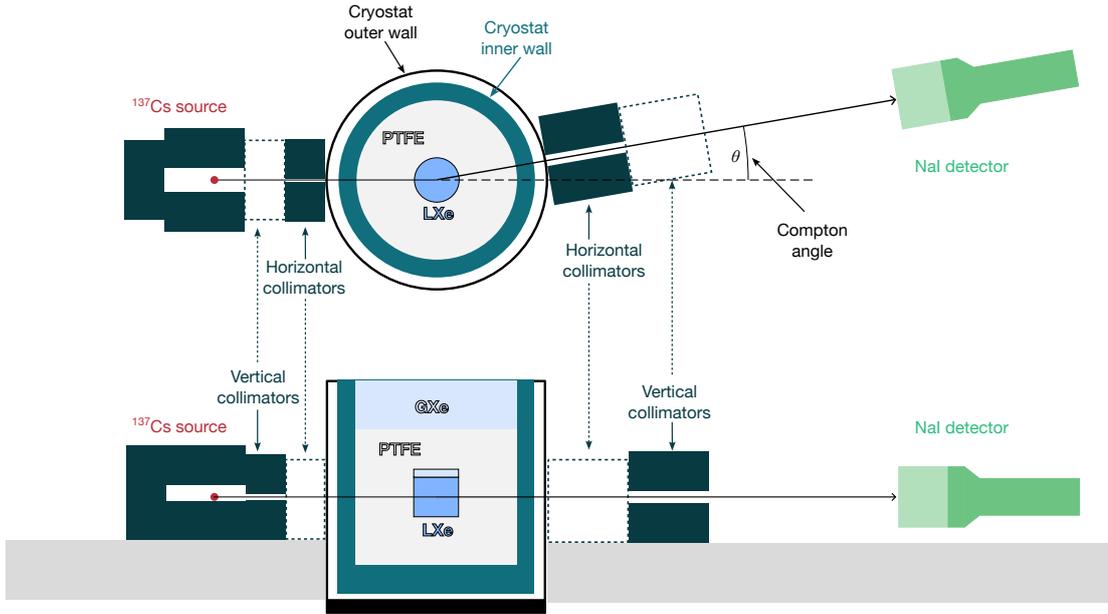


Figure 7.4.: Sketch of the coincidence measurement setup, with a top (top) and side (bottom) view. The ^{137}Cs source is placed in front of the TPC, shown here on the left side, with the external NaI detector placed on the other side at a horizontal angle θ , shown here for $\theta = 10^\circ$. Horizontal and vertical collimation is achieved by placing lead bricks (dark green rectangles) on the source and detector tables as close as possible to the cryostat outer wall. All length measures are to scale.

horizontal and vertical directions by lead bricks placed in parallel on the source and detector tables as close as possible to the cryostat outer wall. The horizontal collimators are positioned directly touching the cryostat, using lead bricks of 5 cm thickness and 10 cm height, but with different lengths, 10 cm for the detector side and 5 cm on the source side (labelled Horizontal collimators in the sketch, on the left and right side of the cryostat respectively). The width of the horizontal collimators, i.e. the horizontal distance between the two bricks forming the collimator walls, are 4 mm and between 6 and 10 mm (varying between different angle setups) for the source and detector sides, respectively. Vertical collimation is achieved similarly, by placing lead bricks in parallel to the surfaces of the source and detector tables, using spacers on either side of the ‘beam’ path, with 8 mm and 17.6 mm height for the source and detector side, respectively. Same as for the horizontal collimators, the vertical source and detector collimators are 5 cm and 10 cm long. The collimation on the source side was chosen to be as tight as possible in order to minimize the angular spread of the γ -rays reaching the TPC, while also keeping the distance between the source and the TPC center short in order to have a decent event rate in the TPC. A geometrical simulation of the resulting angular spread of the setup is presented in [section 7.3](#).

The coincidence measurement was performed at three different central Compton scattering angles θ of 5° , 7° , and 10° , corresponding to expected energy depositions E_{e^-} in the TPC of about 3.2, 6.3, and 12.8 keV respectively, according to [Equation 7.2](#). Three different drift fields, 18, 23, and 28 V/cm were measured, centered around the current field in the XENONnT TPC of 23 V/cm. The total measurement time for all measured angle and electric drift field combinations are shown in [Table 7.1](#).

During the entire run time, $^{83\text{m}}\text{Kr}$ was constantly injected into the TPC by guiding the xenon recirculation flow through the internal calibration source trap in the GS. The source trap contains zeolite beads loaded with ^{83}Rb that decays into the metastable and gaseous $^{83\text{m}}\text{Kr}$ ($t_{1/2} = 1.83$ h), which mixes with the xenon gas and distributes homogeneously in the TPC. It decays via internal conversion (IC) over a short-lived intermediate state ($t_{1/2} = 154$ ns), emitting two IC electrons at 32.1 keV and 9.4 keV in quick succession, giving a clear event

Compton angle θ	Energy in LXe [keV]	Drift field [V/cm]		
		18	23	28
5°	3.24	-	142	116
7°	6.32	61	61	56
10°	12.78	40	50	89

Table 7.1.: Run duration in hours for the combinations of drift field and Compton angle θ that were used in the measurement. The stated drift fields correspond to the targeted median fields in the TPC volume. Energies in LXe deposited by the Compton electrons are calculated from the Compton scattering kinematics using the given central angle.

signature in the TPC. The resulting event signature can be easily and clearly selected in data, allowing for the development of the relative S1 and S2 signal corrections in the TPC, as well as for the monitoring of the stability of S1 and S2 signal yields over the long run time.

Calibrations of the energy response of the NaI detector using an array of several γ -ray sources, and of the gains of the TPC PMTs, were performed throughout the run every few days. Additional cross-calibrations between the TPC and the NaI detector with a ^{22}Na source providing γ -rays from positron annihilation emitted simultaneously back-to-back, were performed as well in order to calibrate the coincidence timing response between the two detectors.

The measurement run was performed between October 20th and December 22nd, 2022, with a total run time of more than two months. During the first week of TPC operation in the LXe dual-phase mode, several commissioning measurements were performed to ensure the correct alignment of the coincidence setup, as well as to check the TPC performance. After this initial commissioning phase, a total power outage in the laboratory led to an emergency recuperation of the xenon inventory in the TPC and the GS, and a following shutdown of the setup. During this time, the $^{83\text{m}}\text{Kr}$ source trap was replenished. The run was then restarted by refilling the TPC with LXe on November 2nd. After an adjustment period for the gas flow through the source trap, as well as a waiting period for the xenon purity to stabilize to a level where the S2 signals were almost flat in drift time, the coincidence measurement could then be started on November 16th. The measurement then ran continuously until December 22nd, with the end determined by external availability constraints¹.

7.1.3. Data taking and processing

The data taking mode for the coincidence measurement is as follows. The top PMT and the bottom PMT of the TPC are operated at a bias voltage of -900 V and -950 V, respectively, below their maximum operational voltage of -1000 V to ensure stable operation and to prevent excessive gain loss due to aging processes potentially caused by the high event rates during the measurement. The external NaI detector is operated at a bias voltage of -800 V, its maximum operational voltage. The analogue signals of the TPC PMTs and of the NaI detector are fed through a custom-made voltage amplifier module, that amplifies the signal voltages by a factor of two for the TPC PMTs, and a factor of ten for the NaI detector. The amplifier produces at least two identical copies of its output signals, one of each is then fed to a digitizer module. The digitizer is a 8-channel CAEN V1724 module that digitizes signal waveforms with a 100 MHz sampling rate and a 14 bit resolution over a 2.25 V dynamic range. The second copy of each amplified signal is fed to a discriminator module with thresholds set to -10 mV, and consequently to a coincidence logic module, which produces an analogue trigger signal for the digitizer. A trigger signal is produced by the coincidence logic module when either both TPC PMTs exceed

¹Commonly known as ‘the Christmas break’.

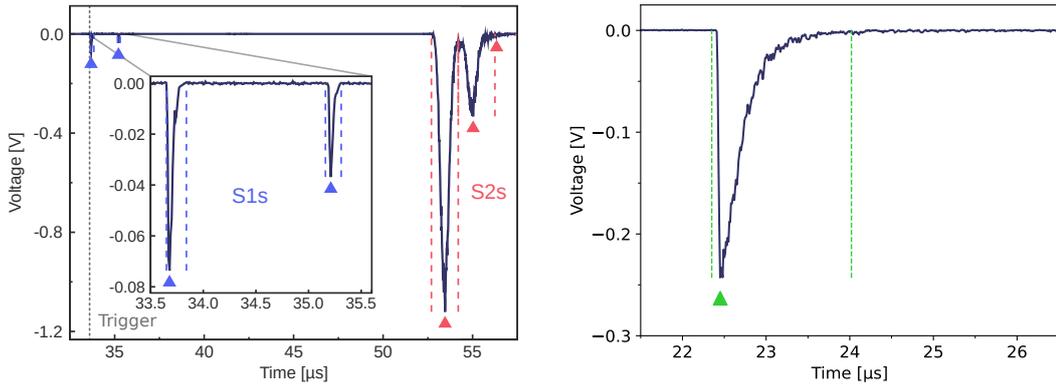


Figure 7.5.: Waveform examples recorded with HeXeR, illustrating the peak finding and classification processes implemented in ‘walgurnisnacht’. (Left) Summed TPC waveform of a $^{83\text{m}}\text{Kr}$ event. Colored triangles and vertical dashed lines indicate the found peak maxima and boundaries, in blue and red for peaks classified as S1s and S2s, respectively. The gray dashed line indicates the trigger position, coinciding with the first S1 signal. Figure from [105]. (Right) Example of a NaI detector waveform, recorded during an energy calibration using ^{137}Cs . Green triangle and dashed lines indicate the found peak position and boundaries, without classification.

the discriminator thresholds at the same time, or when the NaI detector does so. On receiving a trigger signal, the digitizer records the waveforms of the TPC PMTs and the NaI detector in a time window (roughly) centered around the trigger and extending at least one maximum TPC drift time in both directions. The maximum drift times in the TPC for the applied drift fields range from about 65 to 90 μs . The custom-made software HeXeRecorder (HeXeR) [109] is used to read out the digitizer and write the recorded waveforms to disk.

The data processing of the recorded waveforms is performed using another custom-built software developed in [109], called ‘walgurnisnacht’. For each channel, a peak finding algorithm detects pulses in the waveforms that exceed the 5σ baseline fluctuation threshold, and nearby pulses are then clustered together into peaks. The process is repeated for the two TPC PMT signal channels on their summed waveforms as well. For every resulting peak, relevant peak parameters are computed from both the per-channel and the summed waveforms. Relevant peak parameters include area, width, maximum height, start and end times, as well as shape parameters such as the time of the 10th to the 90th percentiles of the signal area in steps of 10, used for peak width determination. The peaks are then classified as S1 or S2 signals based on the peak shape parameters. An example of the summed TPC waveform of a $^{83\text{m}}\text{Kr}$ event is shown in Figure 7.5 (left). S1 signals are characterized by a steep rising edge and an exponentially dropping falling edge given by the short scintillation decay times of LXe, resulting in a narrow width in the order of 100 ns. S2 signals approximately follow a Gaussian shape, due to electron cloud diffusion during the drift as well as the electron multiplication and secondary scintillation processes in the gas phase, with significantly larger peak widths in the order of 1 μs . For peaks to be classified as an S1, their 80% width (defined as the time between the 10th and 90th area quantiles) must be below a threshold of 500 ns, and their rise time (defined as the time between the 10th and 90th height quantiles) must be below a threshold of 100 ns. Peaks above those thresholds are classified as S2 signals. As shown in the inset in Figure 7.5 (left), the two subsequent S1 signals of the $^{83\text{m}}\text{Kr}$ IC electrons at 32.1 keV and 9.4 keV are successfully identified and classified. The S2 signals are successfully classified as well, note though that typically, the two S2 signals from $^{83\text{m}}\text{Kr}$ events largely overlap and are not identified as separate peaks, but are reconstructed as a single S2 signal of the summed energy of 41.5 keV. The example shown in the figure is from an event where the time between the two decays, estimated from the separation of the S1 signals, is about 1.5 μs , corresponding to ten half-lives of the intermediate 9.4 keV state, making it an extremely rare event.

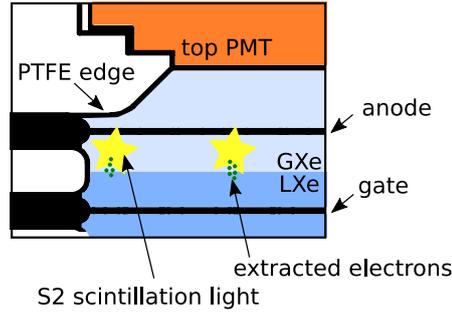


Figure 7.6.: Sketch of the TPC top PMT edge effect. The edge of the PTFE support structure partly blocks the top PMT from seeing S2 photons produced in the gas gap between liquid level and anode, leading to a bias in the S2 signal area seen by the top PMT for events close to the TPC walls. Figure from [110].

An example waveform of a NaI detector peak is shown in the right panel in Figure 7.5. Similar to the LXe scintillation signals, NaI scintillation signals have a steep left flank, but a relatively slowly falling right flank due to the larger scintillation decay time of about $0.23 \mu\text{s}$ [111]. With the NaI peak widths in the order of one microsecond, similar to width of TPC S2 signals, the same pulse finding and clustering algorithm can be used, as indicated by green dashed lines, showing the found peak boundaries. This was further verified by waveform watching of the NaI signals, where the peak boundaries were visually checked. No classification is required for the NaI peaks as only one type of signal is expected.

For the data analysis of the TPC data used in this study, the peak classification based on the summed TPC PMT waveforms is used. All S1 and S2 signal parameters are extracted from the summed TPC PMT waveforms, except for the S2 area where only the signal area seen by the bottom PMT is used. This is necessary because the S2 signal area seen by the top PMT is biased to smaller areas for events close to the TPC walls at large radii, caused by an edge in the PTFE support structure that partly blocks the top PMT from seeing the S2 photons produced in the gas phase. This is illustrated in Figure 7.6. Due to the lack of horizontal (xy) position reconstruction in the HeXe TPC (only one PMT in the top), these events cannot be removed based on their position. While they can in principle be removed by a strict selection on the fraction of the total S2 area seen by the top PMT (area fraction top, AFT), as done in previous studies [106, 109, 110], this would lead to a significant loss of more than half of the already very limited statistics in the data after the coincidence selection is applied, and is therefore not used in this study. The bias in the S2 area seen by the bottom PMT is negligible compared to the intrinsic S2 area resolution of the TPC, such that the S2 area seen by the bottom PMT can be used instead. The peak finding and classification algorithms are still performed on the summed PMT waveforms.

The conversion of the signal areas in the TPC PMTs from the raw ADC counts in the digitizer to photoelectrons (PE) is performed by extracting the gain values from the LED calibration data. For the gain measurements, the PMTs are illuminated with the pulsed LED at a wavelength of 350 nm, with a frequency of about 7 kHz. The gain values are extracted using the model-independent method [36]. Detailed information on the PMT gain calibration procedure and analysis for the PMTs in HeXe can be found in [109]. The two PMTs' gains were calibrated several times throughout the run, the resulting gain values are shown in Figure 7.7. The time evolution of the PMT gains showed a decreasing trend over time, with relative decreases of 6.4% and 8.0% for the top and bottom PMT, respectively, determined from a linear fit to the evolution. This trend is corrected for in the data by scaling the pulse areas seen by the respective PMT accordingly, as a function of the time of the event. After the PMT gain correction, the median S1 and S2 signal areas of the $^{83\text{m}}\text{Kr}$ events, the selection of which will be covered in subsection 7.2.2, showed a stable behavior over time, with relative fluctuations of a standard deviation around the mean of about 1% for both S1s and S2s.

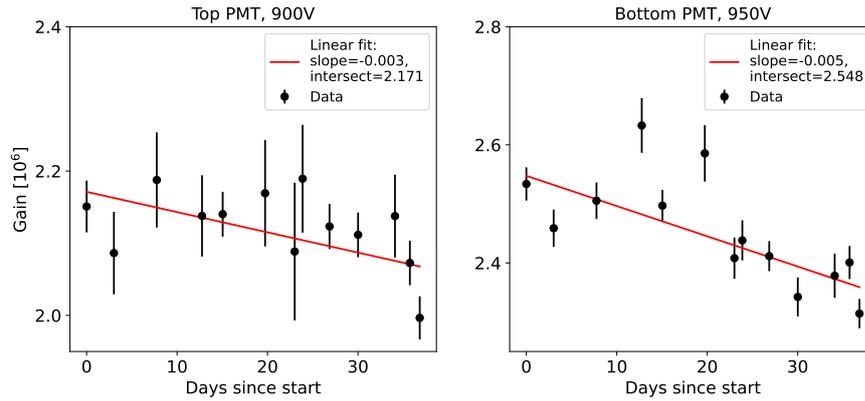


Figure 7.7.: Time evolution of the TPC PMT gains from the LED calibration measurements. The decreasing trend is fitted with a linear function, used to correct the pulse areas in the data analysis.

7.2. Calibrations and detector characterization

In this section, the characterization of the response of both the external NaI detector and the HeXe TPC is detailed. For the NaI detector, the energy scale and resolution are calibrated using several γ -ray sources, the procedure and results is described in [subsection 7.2.1](#). The development of vertical position reconstruction, data selection criteria and signal corrections using $^{83\text{m}}\text{Kr}$ events in the TPC is described in [subsection 7.2.2](#). The selected and corrected $^{83\text{m}}\text{Kr}$ events are used to determine the S1 and S2 signal yields, from which the scintillation and ionization gains g_1 and g_2 are extracted, defining the ER energy scale in the HeXe TPC.

7.2.1. Energy calibration of the NaI detector

The energy response of the NaI detector was characterized with several external γ -ray sources, covering a wide energy range centered around the ^{137}Cs γ -ray line at 662 keV. With the small Compton angles used in the measurement, the expected energies of the scattered γ -rays in the NaI detector are closely approaching the 662 keV line, making the energy range slightly below this peak the most critical for accurate calibration. The used γ -ray sources for the NaI detector calibration are listed in [Table 7.2](#), along with their respective γ -ray energies from literature.

Source	γ -ray energy [keV]
^{241}Am	59.5
^{133}Ba	356.2
^{22}Na	511.0
	1274.5
^{137}Cs	661.7
^{60}Co	1173.2
	1332.5

Table 7.2.: The γ -ray energy of the highest-intensity lines for the sources used in the calibration of the NaI detector. Values from [\[53\]](#).

Energy calibration measurements were repeated seven times throughout the run, performed by

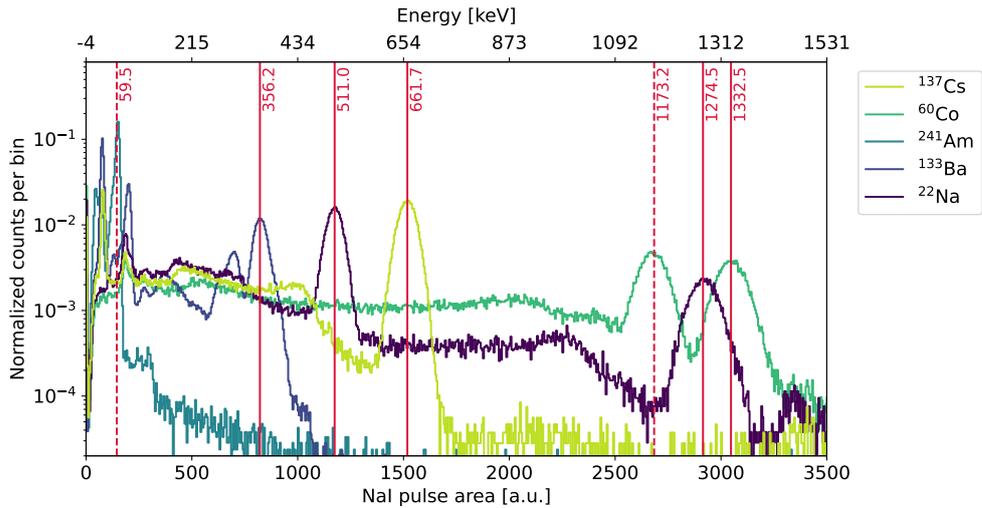


Figure 7.8.: NaI energy calibration measurement from November 30th, 2022. The lower x-axis shows the pulse area of the NaI detector, while the upper x-axis shows the corresponding energy in keV as given by the calibration fit shown in Figure 7.9. Vertical solid lines indicate the γ -ray energies of the peaks used for the calibration, while the dashed lines indicate the expected energies of other prominent peaks in the spectra.

sequentially placing the γ -ray sources in front of the NaI detector and recording and processing the NaI detector waveforms in the same way as for the coincidence measurement. The resulting NaI pulse area spectra for the different γ -ray sources from one of the calibration measurements are shown in Figure 7.8. For each source, the spectrum shows one or more full-absorption photopeaks, corresponding to events where the entire energy of a detected γ -ray is deposited in the NaI crystal via the photoelectric effect as well as (also multiple) Compton interactions inside the crystal, if the scattered γ -ray is eventually absorbed as well. In addition to the full-absorption peaks, the spectra show the Compton continua, corresponding to events where the γ -rays undergo Compton scatters in the crystal and the scattered γ -ray escapes the detector. The Compton edges of the full-absorption peaks, given by the maximum possible energy transferred to the Compton electrons in the back-scattering process at $\theta=180^\circ$, are clearly visible in the spectra as well. To calibrate the NaI detector, the mean pulse areas of the photopeaks indicated by the solid vertical lines are extracted by fitting a Gaussian distribution to the observed peaks in the spectra. The relative energy resolution of the NaI detector at the ^{137}Cs line is determined using the fit result to the full-absorption peak as well, giving a value of $\sigma/\mu = 2.82\%$, which is in good agreement with the value provided by the producer. Note that for the ^{60}Co source, only the highest energy γ -ray line at 1332.5 keV is fitted, because the shape of the lower energy line at 1173.2 keV is distorted by the start of the Compton continuum of the higher energy line, with a Compton edge energy of $E_{e^-, \text{max}} = 1117.8 \text{ keV}$.

After extracting the photopeak positions for each calibration set, a decrease in the NaI pulse area over time is observed, likely due to a decrease in the detector's PMT gain or a slow degradation of the crystal response. The fitted peak positions of the ^{137}Cs line show a relative decrease of 2.0% over the full run duration, a shift that is comparable in size to the 1σ energy resolution of the detector, and therefore not negligible. This downwards trend is approximately linear in time, and is corrected in the data analysis by applying a linear correction to the NaI pulse areas. After correcting for this time dependence, the energy calibration is performed by averaging the corrected peak positions from the seven calibration measurements for each γ -ray line, and fitting a linear function to the average peak positions vs. their literature γ -ray energy. The resulting energy calibration fit is shown in Figure 7.9. The calibration point of the ^{241}Am line at 59.5 keV is not used in the fit, as it is observed to induce a slight bias in the fit for the higher energy lines, likely due to a non-linear energy response of the NaI detector at lower

energies. As can be seen by the residuals in the bottom panel of the figure, the ^{241}Am line is the only calibration point that is not well described by the linear fit, with a deviation of about 2σ . Excluding the ^{241}Am line from the fit therefore leads to a better calibration in the energy range we're most interested in for the γ -ray coincidence measurement, close to the ^{137}Cs energy line at 661.7 keV.

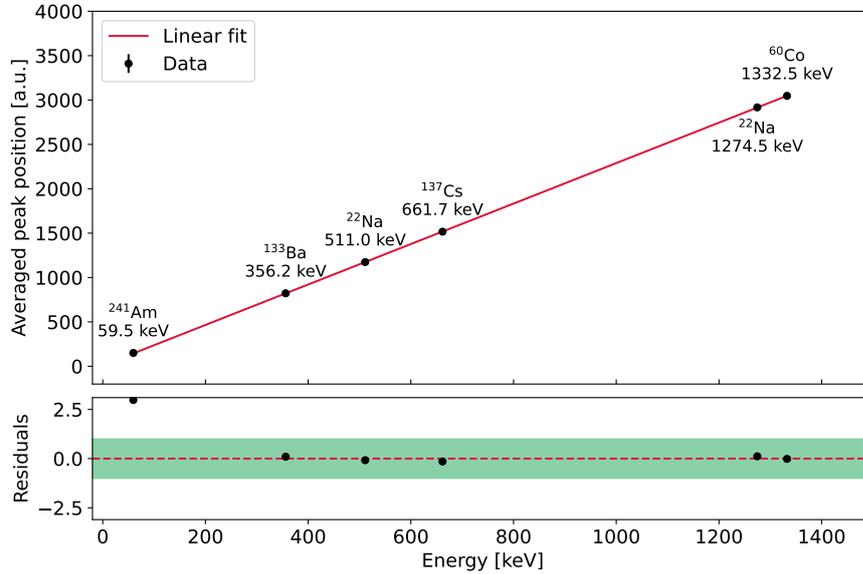


Figure 7.9.: The energy calibration fit of the NaI detector, showing the extracted peak positions vs. the literature line energies for the used calibration sources as labelled in the figure. Note that the lowest energy peak from ^{241}Am at 59.5 keV is not included in the fit. The bottom panel shows the residuals in units of σ -deviation from the best-fit line, with the 1σ region marked in green.

7.2.2. TPC response characterization with $^{83\text{m}}\text{Kr}$

Vertical event position reconstruction

For the event position reconstruction in the vertical coordinate z , the positions of the gate and cathode electrodes in the drift time coordinate need to be determined, with the drift time defined as the time difference between the 10% area quantile of the largest S1 signal and the 50% area quantile of the largest S2 signal in an event. The gate and cathode positions are extracted using the single electron (SE) S2s produced by large S1 signals via the photoelectric effect on the stainless steel electrode grids. If there is no other ‘real’ S2 signal present in the event – an effect that predominantly occurs for interactions in the charge-insensitive region below the cathode – the small SE S2s will become the largest S2 matched with the S1 signal that produced them, and therefore have a drift time corresponding to the time it takes for electrons to drift from the electrodes to the liquid surface. With the continuously deployed ^{137}Cs source in front of the TPC, events with large S1s and SE S2s are readily available in the data, and are selected by requiring large S1s with $S1 > 1000$ PE, and small S2s with $S2 < 100$ PE. This selection is illustrated for the 23 V/cm drift field configuration in [Figure 7.10](#) (left), showing the S2 area seen by the bottom PMT vs. drift time of all events above the S1 threshold. The S2 maximum threshold is indicated by the horizontal dashed line, with the SE populations visible at S2 areas of about 10 PE at drift times of about 1 μs and 75 μs , corresponding to gate and cathode position, respectively. The exact positions are extracted by fitting a Gaussian function with exponential tails on both sides (see [Equation 4.5](#)) to the projection of the SE S2s onto drift time coordinate, as shown in [Figure 7.10](#) (middle) and (right), taking the mean of the Gaussian core as the best-fit estimators for the electrode positions. The conversion of

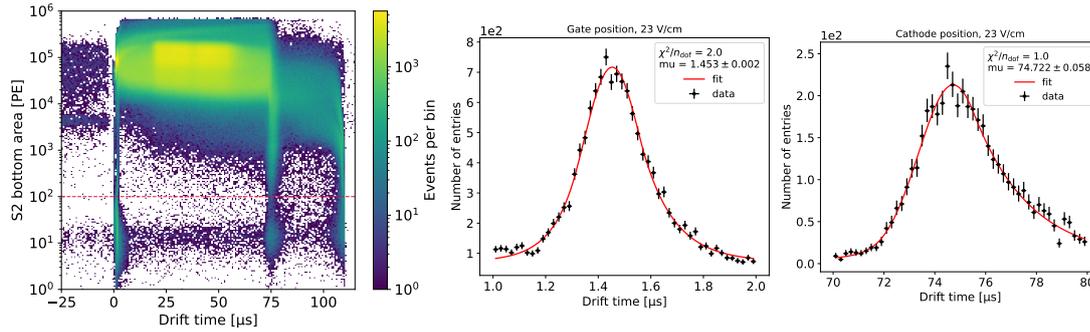


Figure 7.10.: Determination of the gate and cathode positions in the drift time coordinate, here for the 23 V/cm drift field configuration. (Left) Single electron (SE) S2 selection. (Middle) and (right), fits to the projection of the SE S2s in the drift time coordinate for the gate and cathode position, respectively.

drift time to the z -coordinate is then easily performed by linearly interpolating between gate and cathode, setting the gate position to $z = 0$ mm, and the cathode position to $z = 49.3$ mm, where the latter corresponds to the distance between the two electrodes after accounting for the expected shrinking of the PTFE block forming the TPC body due to thermal contraction at LXe temperature.

Data selection criteria

Having determined a conversion from drift time to z coordinate, first a basic fiducial volume (FV) selection is used in order to remove events from below the cathode, in the gas phase, and to avoid events affected by electric field leakage close to the gate electrode from the extraction field region. This basic FV is defined as $5 \text{ mm} < z < 40 \text{ mm}$. Additionally, a general minimum S2 bottom area threshold of 100 PE is applied for all data used in the following analyses.

Events from $^{83\text{m}}\text{Kr}$ decays are then selected by first requiring events to have the largest S1 signal between 0 and 450 PE, and a second-largest S1 between 0 and 200 PE, corresponding to the 32.1 keV and 9.4 keV lines, respectively. Additionally, the two S1s are required to have a time difference between the largest and second largest S1 of at least 300 ns, corresponding to about two times the half-life of the intermediate 9.4 keV state. This is done to avoid an observed dependence of the ratio of the two signal areas on the decay time difference, hypothesized to be caused by the influence of remaining electron-ion pairs that do not recombine from the first decay on the second decay [112, 113]. An upper limit of 900 ns is set on the time difference to avoid events with large time differences that are more likely to be caused by accidental pile-up of other events. The resulting event selection is shown in Figure 7.11. For the development of further data quality criteria and signal corrections, an additional box selection is applied to the events as shown in the figure, requiring minimum S1 areas of 50 PE and 20 PE for the largest and second largest S1, respectively.

The remaining data quality criteria applied to the $^{83\text{m}}\text{Kr}$ events aim at the removal of unphysical or wrongly reconstructed events. The S1 signals are required to have a full range width in a range of (65, 145) ns and (50, 155) ns for the largest and second largest S1, respectively. Additionally, due to the probability of photon absorption depending on the total travelled path between creation and detection of S1 photons, the fraction of S1 area seen by the top PMT (S1 area fraction top, S1AFT) shows a dependence on the z coordinate. S1 signals that follow the expected distribution of S1AFT vs. z are selected by requiring them to lie within the 90% interpercentile range of the entire distribution. For S2 signals, events with more than one S2 signal are removed if the second largest S2 signal is above a threshold that depends on the largest S2 signal, designed to remove both true multiple scatter events and pile-up events, while keeping single scatter events that have SE S2s originating from the electrode grids within

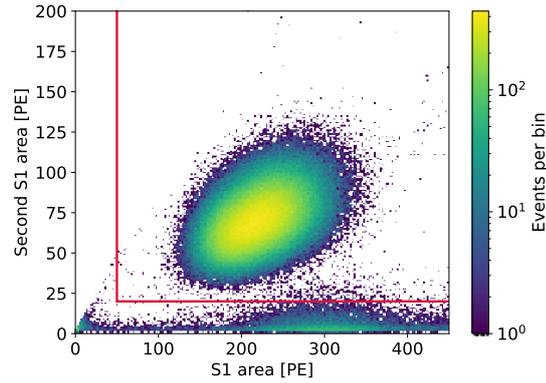


Figure 7.11.: Selection of $^{83\text{m}}\text{Kr}$ events in the 23 V/cm drift field configuration. The second S1 signal area vs. the largest S1 signal area is shown.

their event window. The S2 area fraction seen by the top PMT (S2AFT), as already briefly mentioned, has a strong radial dependence in the HeXe TPC, caused by the shadowing effect of the PTFE edge at the top of the TPC and leading to a dependence of the total S2 area on the S2AFT. This is visible in [Figure 7.12](#), showing the dependence of the S2AFT on the total S2 area (left) and the S2 area seen by the bottom PMT (right) for $^{83\text{m}}\text{Kr}$ events. The S2 area seen by the bottom PMT is almost independent of the S2AFT, and is therefore used in this analysis as the S2 area. While previous studies with data from the HeXe TPC have used a tight selection on S2AFT to remove events closer to the TPC walls, requiring a minimum S2AFT of about 0.7, this is not done in this analysis due to the significant loss of statistics in the γ -ray coincidence dataset. With the dependence of the S2AFT on the radial event position in the TPC, as verified by optical simulations of the TPC [[114](#)] (see also Fig. 7.8 in [[106](#)]), rejecting events with low S2AFT is analogous to a radial FV selection. The FV should however remain the same between the two datasets (Kr and γ -ray coincidence) in order to avoid the influence of a differing geometry on the S1 and S2 light collection efficiencies from optical effects and electric field distributions. Consequently, here only a loose cut is applied to remove unphysical events, requiring the S2AFT to be within 0.5 and 0.76, as indicated by the red lines in [Figure 7.12](#) (right).

Finally, a cut on the S2 signal width is applied to remove events that do not follow the expected distribution of the S2 signal width vs. drift time originating from longitudinal electron cloud diffusion during their drift in the LXe. The cut is defined following [[115](#)], following the model

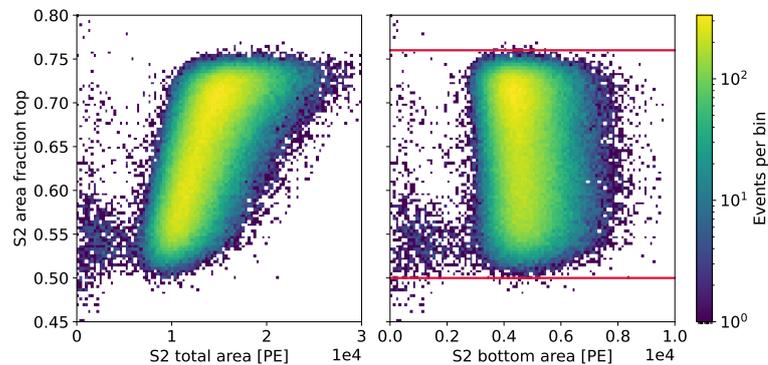


Figure 7.12.: Dependence of the S2 area fraction seen by the top PMT on the total S2 area (left) and the S2 area seen in the bottom PMT (right) for $^{83\text{m}}\text{Kr}$ events. A strong dependence is seen for the total S2 area caused by the shadowing effect of the PTFE edge in the TPC, while the S2 bottom area is almost independent of the S2 area top fraction.

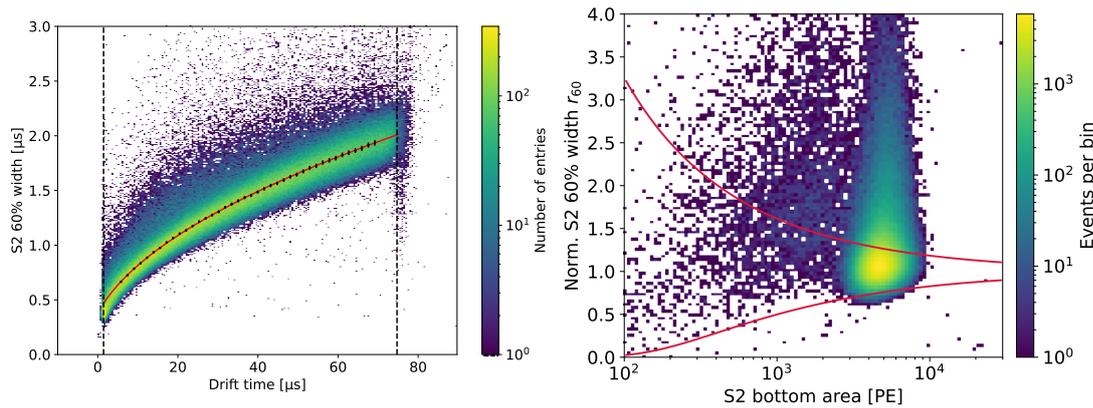


Figure 7.13.: Electron cloud diffusion model fit (red line) to Kr data at 23 V/cm drift field configuration. The inner 60% interpercentile range of the total S2 signal is plotted in dependence of the event drift time, and the model is fitted to the medians of the distribution in bins of drift time (black). Vertical dashed lines indicate the gate and electrode positions.

of a Gaussian-shaped electron cloud diffusion in LXe, described by

$$\sigma(t_d) = \sqrt{\frac{2D \cdot (t_d - t_{gate})}{v_d^2}}, \quad (7.3)$$

where D is the longitudinal diffusion constant, t_d is the drift time, t_{gate} is the gate position in drift time, and v_d is the drift velocity. The longitudinal diffusion constant is field depended, and is obtained for each field by fitting the diffusion model to the medians of the S2 signal width distribution in bins of drift time, as shown in **Figure 7.13**. The S2 signal width parameter used in the fit is the 60% width of the total S2 signal, converted to the 1σ width of the Gaussian model via $\sigma_{60}(t_d) \approx 1.68\sigma(t_d)$. The resulting longitudinal diffusion constant for a 23 V/cm field is $D = (41.6 \pm 0.2) \text{ cm}^2/\text{s}$. The value for v_d is taken as the average drift velocity over the drift region obtained from the known drift length and the gate and cathode positions in drift time for each field, as already shown in **Figure 7.10**, giving $v_d = (0.673 \pm 0.004) \text{ mm}/\mu\text{s}$ for 23 V/cm. The cut is then performed on the ratio of the observed S2 width with the expectation from model, called the ‘normalized S2 width’, defined as

$$r_{60}(t_d) = \frac{(\sigma_{60}(t_d))^2 - (\sigma_{60}^{\text{SE}})^2}{(\sigma_{60}^{\text{model}}(t_d))^2}, \quad (7.4)$$

where σ_{60}^{SE} is the width of the SE S2s. Because the expected width of an S2 generally also depends on the S2 area, the cut boundaries are defined in dependence of the S2 bottom area, allowing for the usage of the same cut definition also for the analysis of the γ -ray coincidence data. The normalized width of larger S2s consisting of multiple electrons, spread out in time according to the Gaussian diffusion model, is expected to follow a χ^2 distribution of the number of electrons in the S2, approximated by dividing the S2 area by the mean area of SE S2s. The cut boundaries are defined by requiring the normalized S2 width to be within the 1×10^{-14} to the $1 - 1 \times 10^{-14}$ percentile range of the expected χ^2 distribution at each S2 area. These boundaries are shown in **Figure 7.13** (right) in red, events between the boundaries are kept. The main population of Kr events at S2 areas of about 5×10^3 PE lies within the cut boundaries, with a tail towards larger widths that is removed by the cut, originating from events where the two merged S2 signals are not as closely merged as the main population. In principle, these events would not need to be removed for the analysis of the Kr data since there is no area bias expected to result from the removal of events with larger S2 widths. For the analysis of the γ -ray coincidence data, however, the removal of these events is necessary for the rejection of

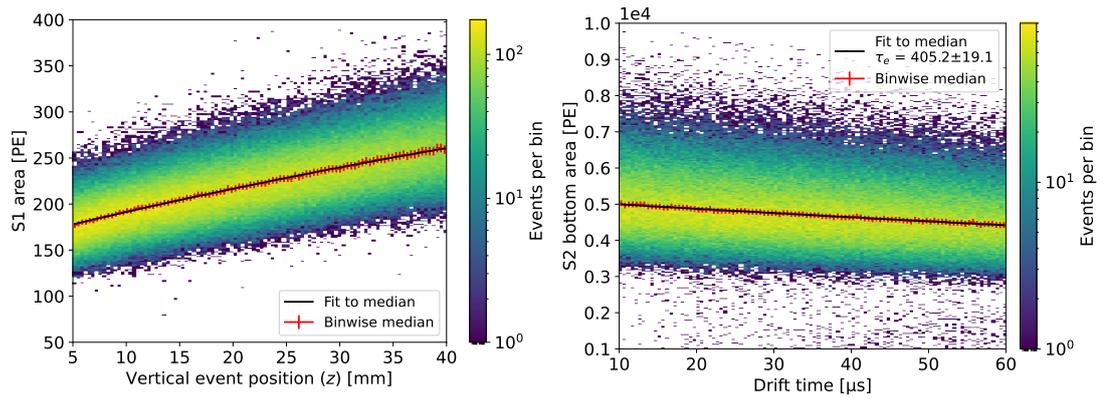


Figure 7.14.: Signal corrections developed from $^{83\text{m}}\text{Kr}$ events in the 23 V/cm drift field configuration. (Left) S1 signal area vs. z for the largest S1 signal. (Right) S2 signal area vs. drift time. In both panels, the red crosses indicate the median of the distribution for each bin in the respective x-axis, which are fitted to extract the relative signal corrections as described in the text.

merged MS events.

Signal corrections

With the data selection criteria as described above applied, we can develop the relative S1 and S2 signal corrections using the selected $^{83\text{m}}\text{Kr}$ events. Because of the use of only one PMT in the top of the TPC, signal corrections in the xy -plane are not possible, such that the signals are only corrected in the z -coordinate. S1 signals are corrected for the z -dependence of the light collection efficiency, shown in Figure 7.14 (left), where the distribution of largest S1 signal area of the selected $^{83\text{m}}\text{Kr}$ events vs. the z coordinate is plotted. The correction is defined by fitting a polynomial of second order to the medians of the S1 signal area distributions in bins of z , shown by the black line. The resulting best-fit function is used to define relative, z -dependent correction factor that scales S1 signals to the one expected at the vertical center of the TPC at 24.65 mm. Corrected S1 signal areas are referred to as cS1. For S2 signals, the exponential-decay-like decrease of the signal size with larger drift times due to electron attachment to impurities in the LXe is corrected, as shown in Figure 7.14 (right). The S2 signal area vs. drift time is plotted for the selected $^{83\text{m}}\text{Kr}$ events, and an exponential function following Equation 2.7 is fitted to the binwise medians of the S2 areas. The resulting electron lifetime is about $\tau_e \sim 400 \mu\text{s}$ for the 23 V/cm drift field configuration, significantly larger than the maximum drift time of about $75 \mu\text{s}$ (see Figure 7.10), with similar values found for the other two drift fields. Therefore, this value is used to correct the S2 signal areas drift time dependence, similarly to the S1 case, by scaling measured S2 signal areas with a relative correction factor according to the event's drift time, to the S2 signal area expected at zero drift time. Corrected S2 signal areas used in the following are referred to as cS2.

Kr yields and g_1, g_2 determination

With the Kr data selection and corrections finalized, the S1 and S2 signal yields, in units of detected PE per deposited energy, are determined from the $^{83\text{m}}\text{Kr}$ dataset at each drift field configuration. For the results for g_1 and g_2 to be consistent with the γ -ray coincidence data analysis, the S1 and S2 signal yields and the g_1 and g_2 values are determined for the tighter ‘coincidence FV’, which is constrained to the central 2 cm in the z -coordinate as will be explained in more detail in the following section.

The corrected S1 and S2 signal area distributions, exemplary for the 23 V/cm drift field con-

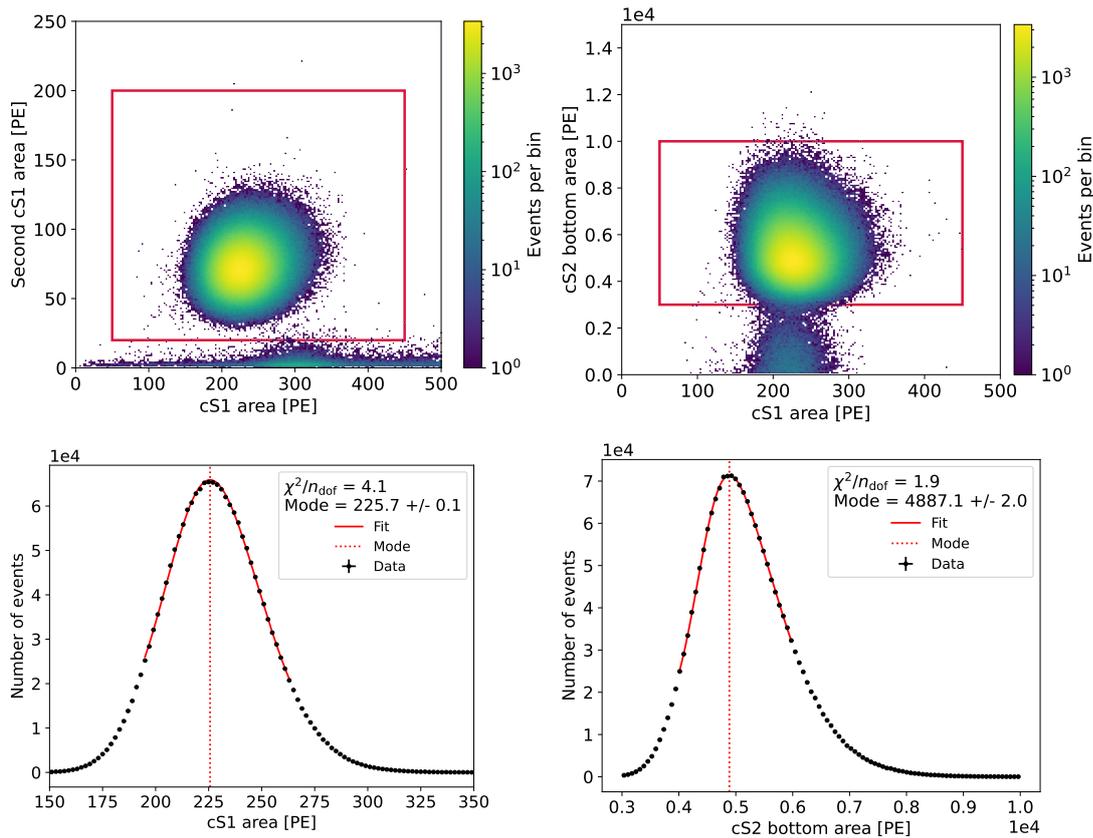


Figure 7.15.: Determination of S1 and S2 yields of the $^{83\text{m}}\text{Kr}$ events in the 23 V/cm drift field configuration. The S1 and S2 signal area distributions are fitted with a skewed normal distribution, and the mode of the fitted distributions is used as the best-fit estimate for the respective fields.

figuration, are shown in [Figure 7.15](#). The two corrected S1 signals from the two energy lines of the $^{83\text{m}}\text{Kr}$ decay are shown in the top left panel, while the largest cS1 and merged cS2 signal distribution is shown in the top right panel. On top of the event selection criteria described above, events are selected according to the red boxes shown in the respective panels. The population of events with cS2 signals below the main population is caused by events that suffer from charge loss, likely due to their proximity to the TPC walls, corresponding to the low S2 area events at S2AFT values below about 0.6, visible in [Figure 7.12](#). The absolute S1 and S2 signal yields (i.e. in units of PE) are extracted from the projected respective area distributions, shown in the bottom panels of [Figure 7.15](#) for the largest cS1 signal (32.1 keV) and the merged cS2 signal (41.5 keV) on the left and right, respectively. For all S1 and S2 signal area distributions, a deviation from the expected Gaussian shape with a tail towards larger signal areas is observed. While the exact reason is not yet fully understood, a likely explanation is the unresolved radial dependence in the geometrical light collection efficiency for the S1 signals. For the S2 signals, inhomogeneities in the extraction and amplification field in the gas phase caused by the shape of the gate and anode grids in the xy -plane, which similarly can not be corrected for lacking xy event position information, are expected to be the cause. A variation in the drift field over the radial coordinate is expected to be a subdominant effect here, as it would lead to anti-correlated signal variations between the S1 and S2 signals, instead of the observed correlated variations. In the absence of a detailed understanding or the possibility of correcting for the proposed causes, the absolute S1 and S2 signal yields are extracted from the projections by fitting skewed normal distributions to the data, and using the mode of the fitted distributions as the best-fit estimate for the yields. For the S1 distribution, a gaussian is fitted first, the best-fit parameters of which are used to determine a fit range of $\mu \pm 1.5\sigma$ for

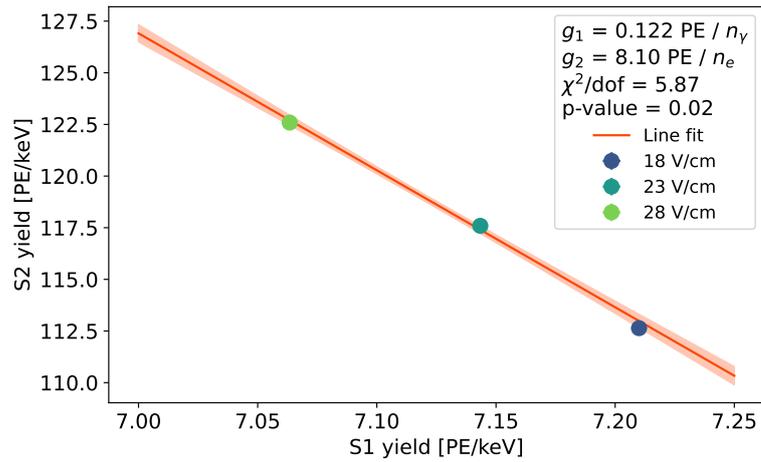


Figure 7.16.: Doke plot with the best-fit S2 signal yields vs. the S1 signal yields for the $^{83\text{m}}\text{Kr}$ events, normalized to summed energy of 41.5 keV. The orange line indicates the best-fit linear function to the data points, used to determine the scintillation and ionization gains, g_1 and g_2 as stated in the legend. The orange shaded region indicates the 2σ uncertainty on the fit.

the skewed normal fit. For the S2 distributions, the fit range is fixed to be between 4×10^3 and 6×10^4 PE for all fields. The errors on the extracted modes are estimated by varying the best-fit parameters according to their uncertainties and correlations in a Monte Carlo approach, and taking the standard deviation of the resulting mode distributions as the error on the mode.

The S1 and S2 signal yields, normalized to the deposited energy, of the $^{83\text{m}}\text{Kr}$ events at the three drift field configurations are shown in Figure 7.16. The S2 yields correspond to the merged S2 signals from both the 32.1 keV and the 9.4 keV, and are thus normalized to the total energy of 41.5 keV. Accordingly, the best-fit absolute S1 yields from the two lines for each field are summed and normalized to the total energy as well. The expected anti-correlation between the S1 and S2 signal yields for different drift fields due to the increased electron extraction from the interaction site and consequently reduced scintillation signal from electron-ion recombination is visible. The average scintillation and ionization gains g_1 and g_2 are determined from these results using Doke-plot method as introduced in subsection 2.2.3, by fitting a linear function to the observed field dependence of the signal yields. The values for g_1 and g_2 are determined from the fit result according to Equation 2.14, giving $g_1 = (0.1221 \pm 0.0006) \text{ PE}/n_\gamma$ and $g_2 = (8.10^{+0.15}_{-0.16}) \text{ PE}/n_e$, with the stated uncertainties resulting from the correlated uncertainties on the best-fit parameters. The uncertainty in the absolute S1 and S2 signal yields caused by the expected but non-resolved radial dependence of the detected S1 and S2 signal areas is estimated by evaluating the variation of the median S1 and S2 signal areas for different slices in S2AFT as a crude proxy for the radial coordinate, weighted by the number of events in each slice effectively accounting for the unknown translation between S2AFT and the radial coordinate assuming a homogeneous distribution of events in the TPC. This results in relative uncertainties of about 1% and 2% on the S1 and S2 signal yields, respectively, which can be directly translated to uncertainties on the g_1 and g_2 values. Adding these uncertainties in quadrature to the statistical uncertainties resulting from the Doke-fit results in the overall values and uncertainties of $g_1 = (0.1221 \pm 0.0013) \text{ PE}/n_\gamma$ and $g_2 = (8.10 \pm 0.22) \text{ PE}/n_e$.

7.3. Coincidence Measurements

For the measurement of the ER yields from coincident γ -ray events, several additional components to the analysis are required. First, the coincidence timing between the TPC and the NaI detector must be calibrated, in order to select events that are time coincident between the two

detectors. This will be described in [subsection 7.3.1](#). Second, the selection of γ -ray events that undergo Compton scattering in the TPC is slightly different from the selection of $^{83\text{m}}\text{Kr}$ events due to the lower energy range of interest, and will be summarized shortly in [subsection 7.3.2](#). Third, the determination of the expected energy spectra in the TPC from the γ -ray events under the given scattering angle setups will be described in [subsection 7.3.3](#), along with the estimation of the spread of the electric field strengths. Finally, preliminary results of the ER yields from the selected γ -ray events will be described and discussed in [subsection 7.3.4](#).

7.3.1. Coincidence timing calibration

For the selection of γ -ray events that undergo Compton scattering in the TPC and are consequently detected in the NaI detector, the timing between the two detectors is determined and validated. With the distance between TPC center and NaI detector of about 0.6 m, and the γ -ray velocity the speed of light (assuming mainly travel through air) of about 0.3 m/ns, the time of flight between the two detectors is expected to be in the order of 2 ns – much smaller than the time resolution of the S1 signals in the TPC and the NaI detector, as well as the sampling width of the digitizer of 10 ns. The events that we want to select can therefore be considered to occur simultaneously in the two detectors, i.e. a TPC event’s S1 signal and the NaI signal from the same Compton scatter event must be within a certain, tight time window of each other, also called coincidence window. In order to define the coincidence window between NaI detector and TPC, a coincidence calibration measurement was performed using 511 keV γ -rays from the ^{22}Na source. ^{22}Na decays via β^+ decay with a branching ratio of $\sim 90\%$ into an excited state ^{22}Ne that subsequently emits a γ -ray with an energy of 1274.5 keV. The positron emitted in the decay annihilates with an electron in the immediate vicinity, producing two 511 keV γ -rays that are emitted simultaneously back-to-back. Placing the ^{22}Na source in between the TPC and the NaI detector then allows for the determination of the coincidence window by selecting the back-to-back γ -ray events arriving simultaneously in the two detectors.

The coincidence calibration measurement was performed every time the NaI energy calibration was performed, with the TPC in each drift field configuration (i.e. 18, 23, and 28 V/cm) for at least eight hours in total. For each recorded event, the time difference between the S1 signal in the TPC and the NaI signal was determined by calculating the difference between the largest S1 signal in the TPC and the largest signal in the NaI detector, using the 10% area quantiles of the signals as the time estimate. A set of basic event quality criteria is applied to the TPC data, requiring the presence of at least one S1 and one S2 signal in the TPC, with the S2 area above 100 PE and within the ‘basic’ fiducial volume in z , and at least one signal in the NaI detector. The determination of the coincidence window is then performed as illustrated in [Figure 7.17](#). In the distribution of the time difference between NaI and S1 signal of all events, a large excess of events close to zero time difference is observed; a zoomed-in view of that region is shown in the histogram in the upper right panel, with the time difference distribution on the x-axis and the energy deposit in the NaI detector on the y-axis. The main coincident population is centered at a time difference of about 70 ns. The offset from zero is likely caused by differences in the signal propagation times in the two detectors, in particular the transit times of photoelectrons through the different PMT models, as well as different lengths of the signal cables. The 511 keV full absorption events in the NaI detector are selected within a range of $\pm 3\sigma$ width of the energy resolution, indicated by the red shaded region. The resulting distribution of cS1 signals in dependence of the NaI-S1 time difference after the NaI energy selection is shown in the top right panel in [Figure 7.17](#). The time difference becomes slightly larger for smaller S1 signals, an effect caused by the determination of the 10% area quantiles of S1 pulses, which becomes dependent on the S1 signal size for small signals presumably due to the interpolation and smoothing algorithms used in the quantile determination. The region of interest (ROI) for the low energy ER yield measurement is below 300 PE in cS1, indicated by the horizontal grey line. The bottom left panel in [Figure 7.17](#) shows the projection of the NaI-S1 time difference for all events after the NaI energy selection in black, and for the events below the upper ROI definition in S1 in grey. The coincidence window is defined to be centered around the peak

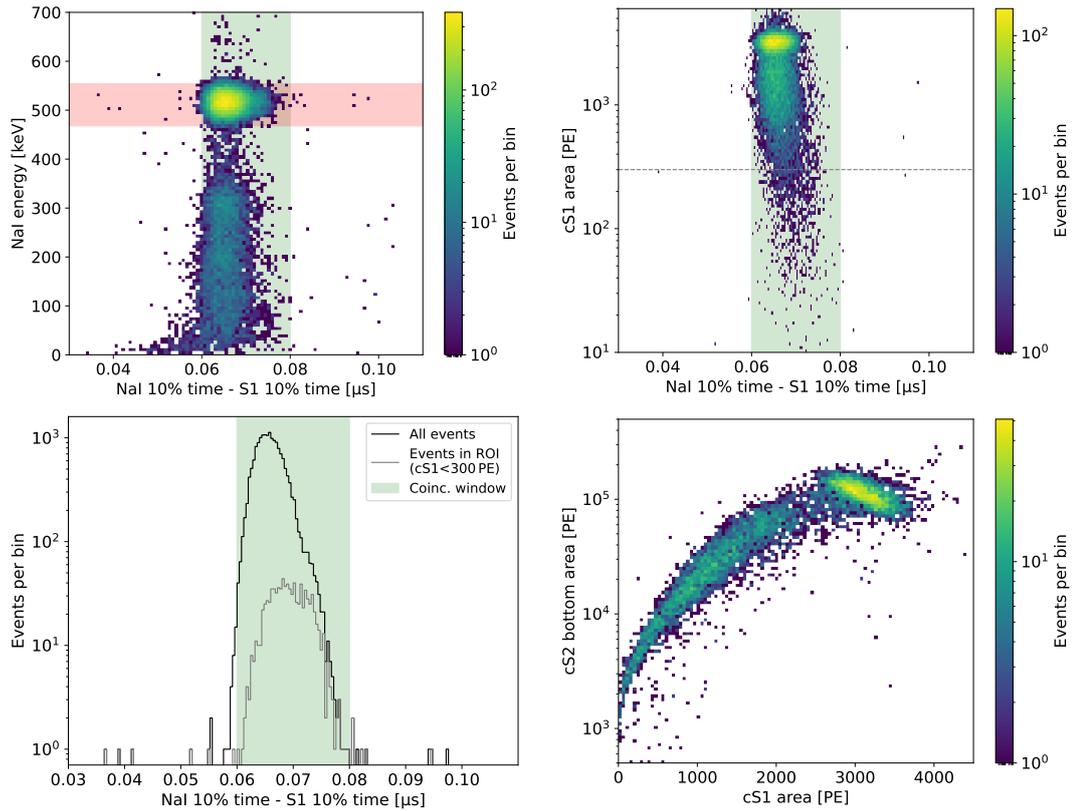


Figure 7.17.: Determination and validation of the coincidence window between the TPC and the NaI detector using 511 keV γ -rays from the ^{22}Na source. (Top left) Energy deposit in the NaI detector vs. time difference between the NaI signal and the largest S1 signal in the TPC. The red shaded region indicates the energy selection of the γ -rays around the 511 keV peak. The green shaded region in all panels indicates the coincidence window definition. (Top right) Largest cS1 signal in the TPC vs. time difference, showing a trend towards larger difference for smaller cS1. Grey dashed line indicates the upper limit of the ROI for the low-energy analysis at $cS1 < 300$ PE. (Bottom left) Projection of the time difference NaI and S1 signal, for all events (black) and for events below the upper ROI definition in S1 (grey). The coincidence window is defined between 60 ns and 80 ns. (Bottom right) cS2 vs. cS1 area distribution of the selected coincident events, showing the expected populations of 511 keV γ -ray events in the TPC as well as the ER band originating from Compton scatters in the LXe.

of the cS1 distribution in the ROI, between NaI-S1 time differences of 60 ns and 80 ns. This region is marked by the green shaded band. The correct selection of coincident events from 511 keV γ -rays in the TPC is validated by the observation of the expected event populations in the cS1-cS2 space, shown in the lower right panel in [Figure 7.17](#). Further event selection criteria to remove MS events, in particular the S2 single scatter and an S2 width cut optimized for high energy events are applied. An elliptical population centered at about 3000 PE in cS1 and 10^5 PE in cS2 is observed, corresponding to the full absorption of 511 keV γ -ray events in the TPC. At lower cS1 and cS2 areas, the ER band formed by Compton scatters of the γ -rays is visible. This confirms that the coincidence window definition allows for an almost background free selection of pure Compton ER events from a few keV up to 511 keV γ -ray events in the HeXe TPC.

7.3.2. Data selection for low energy γ -ray events

The TPC data selection criteria for the low energy Compton scatter events from γ -ray interactions are partially altered compared to the $^{83\text{m}}\text{Kr}$ data selection. While the basic data selection

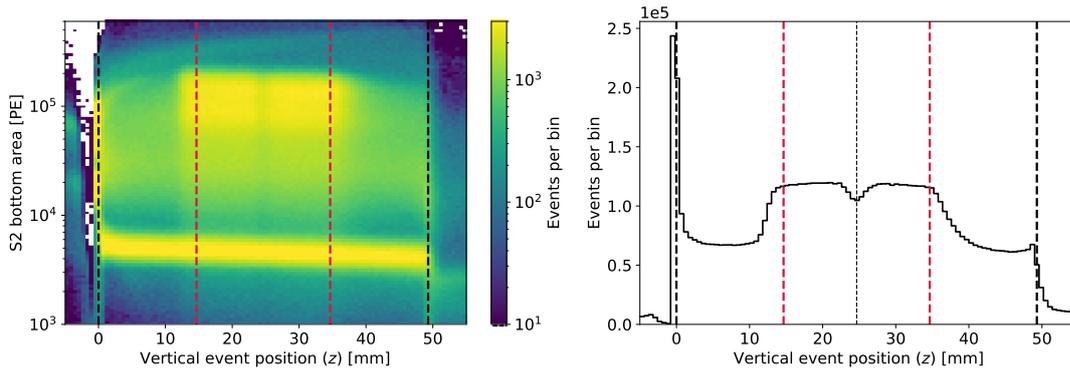


Figure 7.18.: A γ -ray ‘Röntgen’ picture of the HeXe TPC. (Left) The S2 signal area vs. z , showing the effect of the vertical source collimation on the events from the ^{137}Cs source. The observed decrease of S2 area with z originates from electron attachment to impurities during drift. (Right) Projection on the z -axis of all events. In both panels, the red dashed lines indicate the chosen FV used in the coincidence analysis, while the thick black-dashed lines indicate gate and cathode position. The thin black-dashed line in the right panel shows the vertical center of the TPC at $z = 24.65$ mm, with the dip in event density caused by the shadowing effect of the central field shaping ring.

remains the same, i.e. requiring at least one S1 and one S2 signal in the TPC, with an S2 bottom area threshold of 100 PE, several data quality criteria are adjusted.

First, a tighter FV in the z coordinate is defined, designed to select only events from the ^{137}Cs source given the vertical source collimation in front of the TPC. γ -ray events are required to scatter within ± 10 mm around the vertical center of the TPC at 24.65 mm, resulting in the region of $14.65 \text{ mm} < z < 34.65 \text{ mm}$. The definition is illustrated in Figure 7.18, showing the S2 signal area vs. z distribution of all events without any selections applied. In the left panel, the horizontal band at S2 areas of about 5×10^3 PE spanning the full z range of the TPC belong to the $^{83\text{m}}\text{Kr}$ events distributed homogeneously inside the liquid volume. The events seen in the vertical band between about 12 mm and 36 mm in z originate from the ^{137}Cs source, with the visible boundaries caused by the vertical source collimation. The S2 signal areas span the full S2 range of the TPC, originating mainly from γ -ray Compton scatters, with the digitizer channel of the bottom PMT saturating (i.e. clipping at its full dynamic range of 2.25 V signal height) at about 2×10^5 PE (i.e., the full-absorption peak of the 662 keV γ -ray line is not visible here). The small deficiency of events in the vertical center of the TPC is due to the shadowing of the γ -rays by the central field shaping ring (10 mm thick, 2 mm height stainless steel) of the HeXe TPC (see Figure 7.3). The red dashed lines, as also shown in the right panel in Figure 7.18, indicate the chosen FV for the analysis of the coincidence data, containing the main population of the γ -rays within the collimator width. From the set geometry of the source beams opening angle given the vertical collimation height, the expected vertical extent of the γ -ray events in the radial TPC center is about 2.3 cm, nicely in agreement with the observed vertical extent of the events in the TPC.

For the data quality criteria on the TPC data, the S2 single scatter cut, the S2 width cut, and the S2 AFT cut are unchanged from the $^{83\text{m}}\text{Kr}$ selections, as they are already defined in an S2 area dependent way for the first two, and independent of the S2 area for the latter. For S1s, the S1 width cut is adjusted from the constant boundaries defined around the observed widths of the Kr S1s, to a linear function in $\log_{10}(\text{S1})$ as an upper limit for S1s below 100 PE, to account for an observed smaller S1 width for small S1 signals. Additionally, a new cut against multiple S1s is introduced, rejecting events with a second largest S1 signal above 30 PE and therefore mainly rejecting $^{83\text{m}}\text{Kr}$ events from the selection. No selection on S1 area fraction in the top array is applied, since the distribution of coincident events after applying all other corrections and selections were found to be within the expected ranges such that no further selection is necessary.

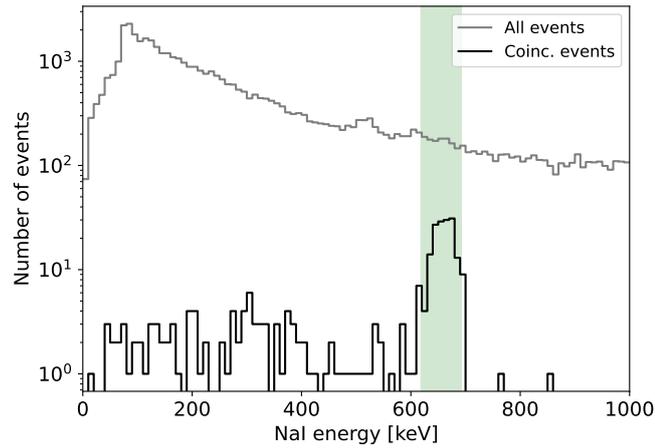


Figure 7.19.: Energy spectra recorded in the NaI detector during the 23 V/cm, 7° configuration. The grey histogram shows all events, while the black one shows only events in the tight coincidence window between TPC and NaI detector. The green shaded area illustrates the NaI energy selection, requiring the energy to be within $\pm 2\sigma$ of the expected γ -ray energy after Compton scattering in the TPC of $E'_\gamma \sim 655$ keV.

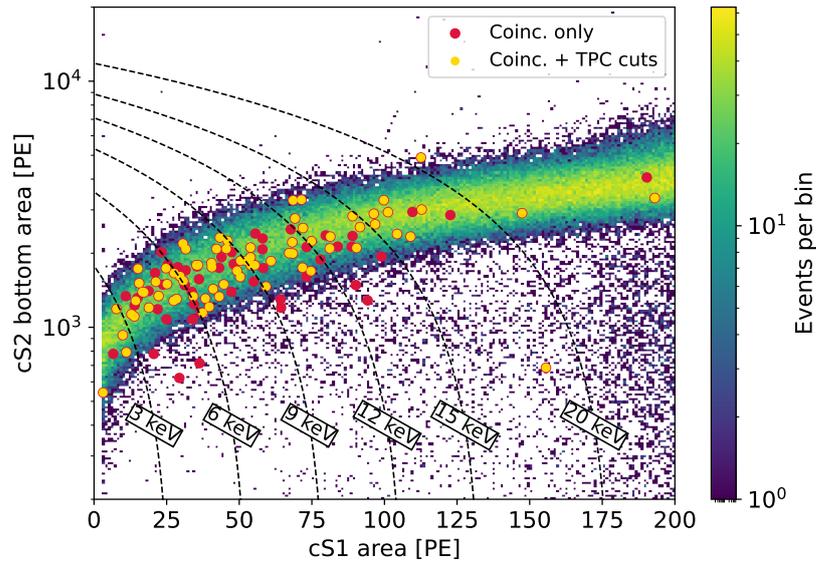


Figure 7.20.: Electronic recoil band in the HeXe TPC, measured at 23 V/cm. The overlaid red datapoints show the events in the tight coincidence window between the TPC and the NaI detector, with the NaI energy selection applied according to the expected energy for the 7° configuration. Golden overlaid points also satisfy the TPC data quality criteria. Black dashed lines indicate equi-energy lines assuming ER events, E_{ER} , in units of keV_{ee} .

In addition to the time coincidence requirement, the γ -ray events seen in both detectors are required to show the expected energy deposition in the NaI detector, as illustrated in Figure 7.19. Here the grey histogram shows the energy spectrum of all events in the NaI detector recorded during the 23 V/cm, 7° configuration, while the black histogram shows the events in the tight coincidence window between the TPC and the NaI detector, with a clear peak at the expected energy of the scattered γ -rays at that angle of about 655 keV. The green shaded area indicates the energy selection applied to the NaI events, requiring the energy to be within $\pm 2\sigma/\mu$ of the expected peak energy.

The final selected dataset of γ -ray events in the TPC is shown in [Figure 7.20](#). The histogram in the background shows the ER band in the HeXe TPC originating from all Compton scatters of the γ -rays at 23 V/cm after applying all data quality selections. All events in the coincidence window and the NaI energy selection for the central scattering angle of 7° are overlaid as red points, while the golden points also satisfy the TPC data quality criteria. The black dashed lines show the equi-energy lines for ER events in the TPC, using the g_1 and g_2 values determined from the $^{83\text{m}}\text{Kr}$ data. These events will be used to determine the low-energy ER yields in the TPC.

7.3.3. Estimation of γ -ray energy and electric field distributions

Before the determination of the low-energy ER yields from the selected γ -ray events, the expected energy spectra in the TPC from the γ -ray events under the given central scattering angles are estimated. The expected spectra of deposited energy in LXe from the Compton electrons of the coincident γ -ray events are determined from an approximate solid-angle MC simulation of the available Compton scattering angles inside the LXe volume, based on the measured geometric setup of ^{137}Cs source, NaI detector and the respective horizontal and vertical collimation widths. [Figure 7.21](#) (top panel) illustrates the geometry in the xy -plane, corresponding to a view from the top onto the setup, used in the calculation for the 7° configuration. The source (red circle) is assumed to be point-like, emitting γ -rays isotropically within a cone (grey shaded region) with an opening angle of 0.85° , given by the 4 mm wide horizontal source collimator positioned directly at the cryostat outer wall (see also [Figure 7.4](#)). The LXe volume is indicated by the blue circle, with the coordinate origin positioned at the center of the TPC. The direct line-of-sight from the NaI detector front (green) through the detector collimator tunnel (blue-green shaded area) is indicated by the solid black lines; in this configuration the width of the collimator tunnel is 9 mm. The dashed black lines, in turn, mark the most extreme possible trajectories through the collimation. For all central scattering angle configurations, all possible trajectories through the tunnel end on the NaI detector front. Additionally, the solid angle cone from the source passing through the LXe volume is completely within the NaI detector's field of view, due to the narrow collimation on the source side. The middle and bottom panels in [Figure 7.21](#) show a zoomed-in view of the central region of the TPC, as indicated by the purple dashed region in the top panel. The scattering angle MC is performed by random sampling interaction points from the solid angle cone within the LXe volume, and calculating the minimum and maximum Compton scattering angles θ_{\min} and θ_{\max} w.r.t. the source position and the point of interaction in the TPC. In the middle panel, the resulting central scattering angles at each interaction point are depicted by the color scale, defined by the central value between θ_{\min} and θ_{\max} . The full range of possible scattering angles, $\theta_{\max} - \theta_{\min}$, is shown in the bottom panel. Noticeably, the most extreme scattering angles are outside of the direct line-of-sight of the NaI detector and therefore show in turn a smaller range of possible scattering angles.

The geometrical calculation is performed for the three central scattering angles θ_o of 5° , 7° , and 10° . The distributions of available θ are produced by stacking the available ranges of each interaction point, weighted by the decreasing γ -ray interaction rate along the path of the γ -ray through the LXe volume following Beer-Lambert, $I(x) \propto \exp(-x/\lambda)$. The mean free path of 662 keV γ -rays in LXe is about $\lambda = 4.5$ cm, estimated using the mass attenuation coefficient for xenon of $\mu/\rho = 7.67 \times 10^{-2} \text{ cm}^2/\text{g}$ [[116](#)] and a LXe density of 2.9 g/cm^3 at 175 K [[117](#)], and results in a relative reduction of the interaction rate of about 70% between nearest and farthest interaction point from the source. [Figure 7.21](#) shows the scattering angle distributions for the three configurations in black. The conversion of the θ distributions to deposited energy distributions of the Compton electrons is done by random sampling from the θ distributions and calculating E_{e^-} according to [Equation 7.2](#) for each sampled θ , and are shown in [Figure 7.22](#) in teal color. While the scattering angle distributions are approximately symmetric around θ_o , the energy distributions exhibit tails towards larger energies due to the non-linear dependence of the Compton electron energy on θ in [Equation 7.2](#), an effect that is more pronounced for smaller θ_o . As a result, the median energy deposit at each θ_o is larger than the one expected from the

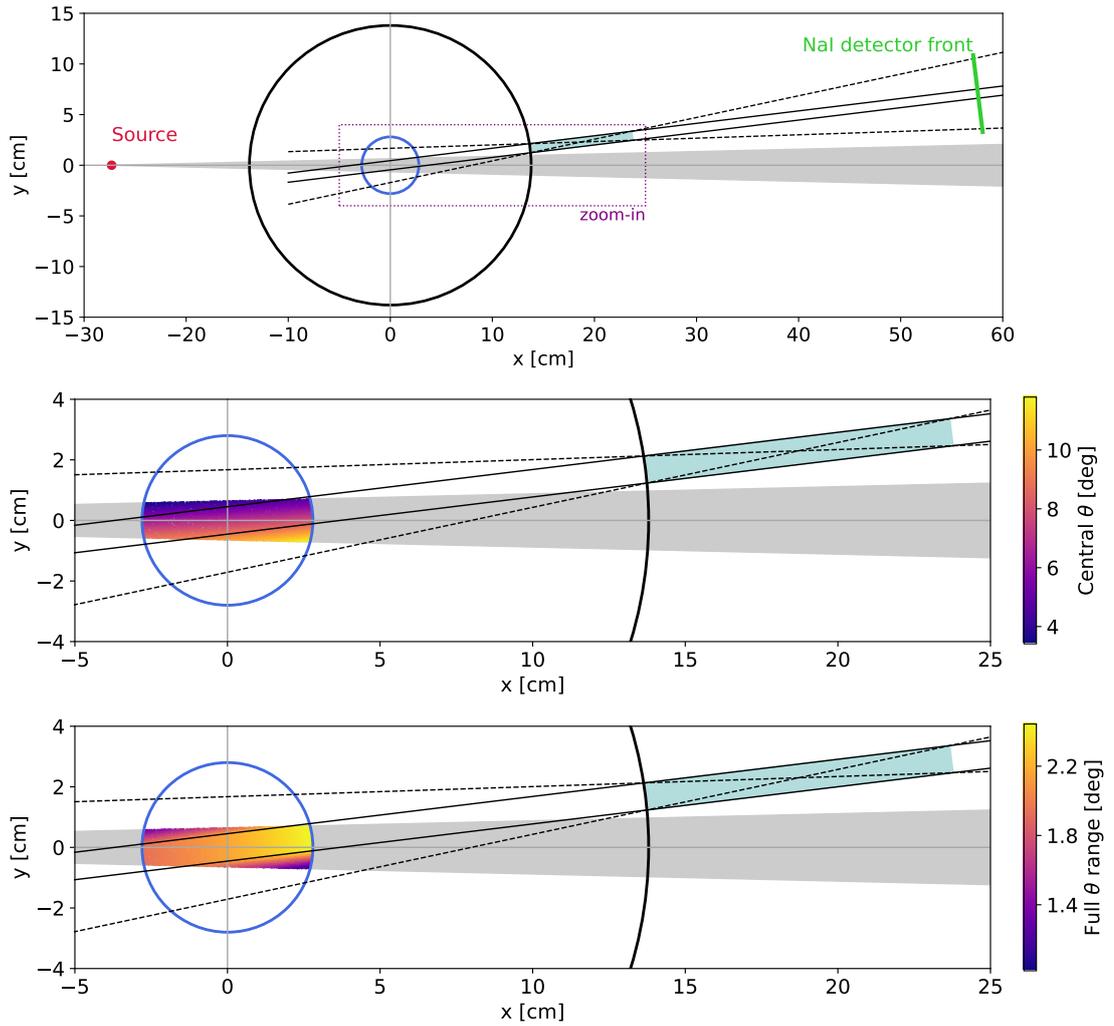


Figure 7.21.: Geometric calculation of the available Compton scattering angles in the xy -plane, shown here for the central scattering angle of 7° . The black and blue circles represent the cryostat outer wall and the LXe volume, with the coordinate origin $x, y = (0, 0)$ in the center. The grey shaded area corresponds to the solid angle cone containing γ -rays assuming a point source and the horizontal source collimation of 4 mm located at the cryostat wall. The middle and bottom panels show a zoomed-in view of the central region of the TPC, with the central scattering angles and the scattering angle spread for each point in the TPC volume indicated by the color scales.

central scattering angle, deviating by 28 %, 11 %, and 3 % for $\theta_o = 5^\circ$, 7° , and 10° configurations, respectively, while the central 68 % spreads relative to the median energy deposits are rather large, with 125 %, 102 %, and 75 %.

It must be stressed that this estimation of the energy spectra is only approximate due to several simplifications in the simulation. First and most importantly, the spread of the scattering angles is calculated only in the xy -plane for simplicity, instead of a full 3D simulation of the scattering angles including the z -coordinate. Second, the Compton scattering cross section's dependence on θ is not taken into account, which decreases the probability of scattering between θ near 0° and $\theta = 15^\circ$ gradually by in total about 10 %, following the Klein-Nishina formula [118]. Third, effects of γ -rays undergoing multiple Compton scatters in the TPC are not taken into account, however the vast majority of these γ -rays are expected to consequently not reach the NaI detector, or fall outside of the NaI energy selection. For γ -rays that do reach the NaI and

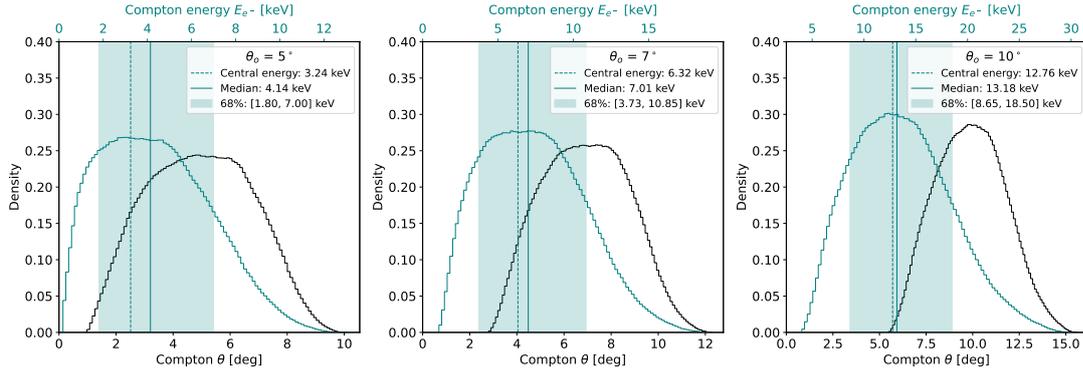


Figure 7.22.: Scattering angle and resulting energy deposit distributions in the TPC.

pass the NaI energy requirement, the resulting MS event topology in the TPC results in a likely removal of the events from the dataset based on the data quality criteria. In particular, the S2 single scatter and the S2 width cut are expected to remove the majority of the MS events, although a quantitative estimate of the MS event rate and the efficiency of the MS event removal remains to be done. Similarly, and more severely, effects of additional interactions of the γ -rays in materials surrounding the LXe volume are not considered as well. Estimating the effects on the resulting energy deposition spectra from multiple γ -ray interactions necessitates a full GEANT4 MC simulation of the HeXe TPC setup and its surroundings, which is beyond the scope of this work, but is currently under development for a future and improved analysis of this dataset.

In addition to the energy distributions, the strength of the electric field and its spread in the TPC are estimated as well, based on 3D COMSOL[46] simulations developed in [106], see also [105]. The simulation employs a superposition method, where the electric field in the TPC is first simulated for each electrical component separately, with the potential set to 1 V for the component in question with all others at ground. The electrical components are the cathode, gate, anode, the field shaping rings, and the PMTs. Summing the fields from each component

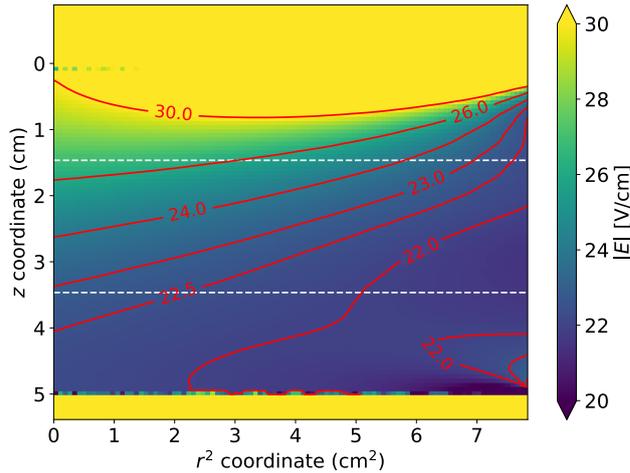


Figure 7.23.: Electric field map in the HeXe TPC produced via a 3D COMSOL simulation [105] for the electric potential configuration used for the set field of 23 V/cm. The color scale shows the magnitude of the electric field in the TPC volume, averaged over the symmetric azimuthal ϕ component. Red lines represent contours of constant electric field strength. The median electric field strength in the FV (between the white horizontal lines) is 22.9 V/cm with a 1σ spread of 5.3%.

Set field configuration	18 V/cm	23 V/cm	28 V/cm
Median	18.2	22.9	27.7
15.9% quantile	17.3	22.1	26.8
84.1% quantile	19.8	24.5	29.2
Minimum	16.6	21.4	26.1
Maximum	22.0	26.7	31.4
Spread $\pm 1\sigma$	6.7%	5.3%	4.3%
Spread full	29.6%	23.3%	19.2%

Table 7.3.: Simulated electric field magnitudes in the HeXe TPC for the three set drift field configurations. All values except the spreads are in units of V/cm.

scaled to the actual applied potentials then results in the overall electric field in the TPC in the given configuration, in a much shorter computational time than the full 3D simulation. For this work, the distribution of electric field strength in HeXe is calculated for the three drift field configurations, targeted at drift fields of 18, 23, and 28 V/cm. The resulting drift field distribution for the 23 V/cm configuration is presented in [Figure 7.23](#), showing the magnitude of the electric field in z vs. r^2 , averaged over the azimuthal ϕ distribution, which is highly symmetric due to the cylindrical symmetry of the TPC. The fiducial volume selection in z is indicated by the white horizontal lines. The spread of the field in the FV is estimated by evaluating the median and the $\pm 1\sigma$ percentile ranges of the field strength in the FV for each drift field configuration, as shown in [Table 7.3](#). For the 23 V/cm drift field configuration, the median field strength in the FV is 22.9 V/cm, with a minimum to maximum spread of 5.3 V/cm, and a relative 1σ spread of 5.3%. The relative spread of the field strength increases (decreases) with decreasing (increasing) set field strength, with a relative spread of 6.7% and 4.3% for the 18 V/cm and 28 V/cm configurations, respectively.

7.3.4. Results for the ER γ -ray yields

For the determination of the ER photon and charge yields from the selected γ -ray events in the TPC, a region of interest (ROI) is defined in the cS1-cS2 space to select low-energy ER events. The cS1 range is set to be between [3, 200] PE for the 5° and 7° central scattering angle configurations (corresponding to the cS1 range already shown in [Figure 7.20](#)), and [3, 300] PE for the 10° one. The lower limit of 3 PE is chosen to avoid the region where the full ER band (see [Figure 7.20](#)) cS1 spectra start showing steeply falling rates towards small PE values, due to the S1 threshold of the TPC caused by PMT QE as well as the light collection efficiency. For all configurations, an upper limit of cS2 of 6×10^3 PE is defined, and no lower limit is set apart from the (uncorrected) S2 threshold of 100 PE already set in the data quality selection.

In the same procedure as for the $^{83\text{m}}\text{Kr}$ data, the γ -ray yields are determined from projecting the selected TPC events onto the cS1 and cS2 axes and extracting the average of the respective S1 and S2 signal yields in PE. [Figure 7.24](#) and [Figure 7.25](#) show the resulting data distributions for all measured combinations of drift field (columns) and central scattering angle (rows). In principle, for mono-energetic or narrow peaks, extracting the S1 and S2 yields from the data would proceed by fitting the shape of the expected energy distributions from [Figure 7.22](#), convoluted with signal reconstruction efficiencies and resolutions, to the measured cS1 and cS2 spectra. However, due to the wide ranges of energy distributions in this measurement, this necessitates the assumption of an underlying model for the energy dependence of the absolute photon and charge yields, i.e. n_γ/keV and n_{e^-}/keV , which is expected to be rather steep in

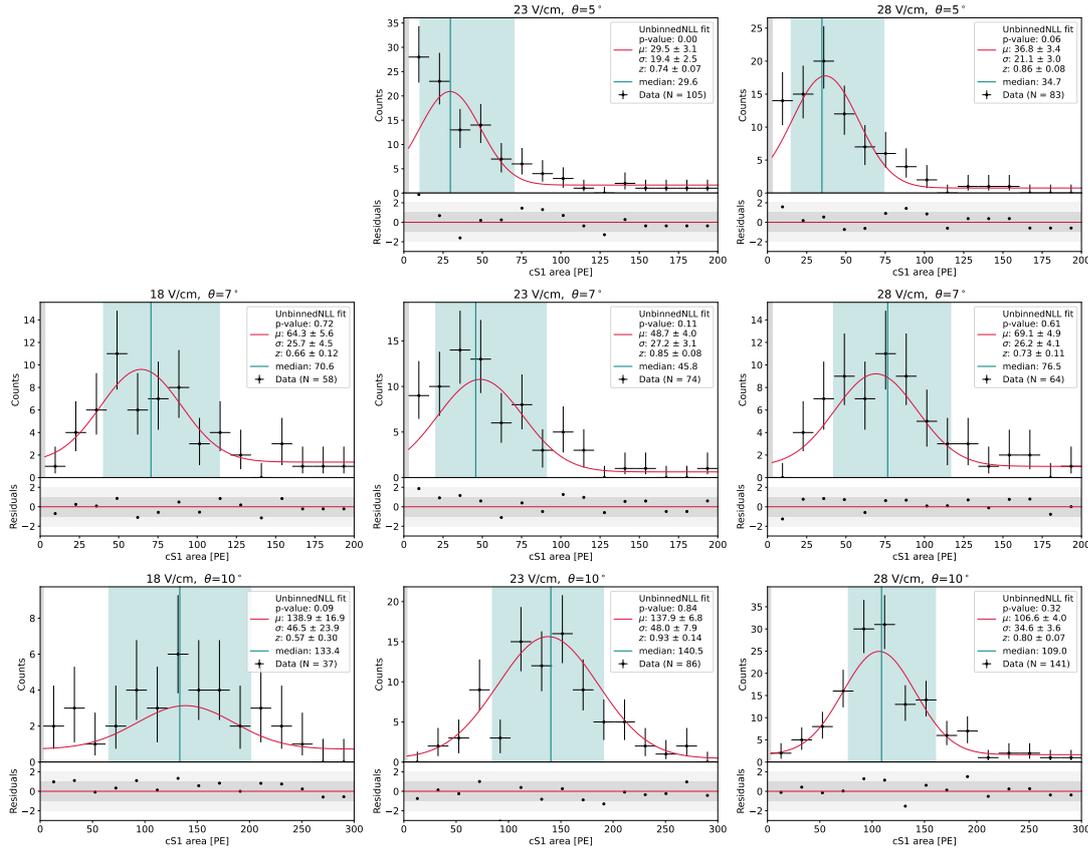


Figure 7.24.: Distributions in cS1 for the selected coincident γ -ray events for all measured drift field (columns) and central scattering angle (rows) combinations. The grey shaded region below 3 PE in cS1 indicates the chosen lower limit of the ROI, above which the S1 detection efficiency is assumed to be close to 100%. Red curves show the results of fitting a Gaussian distribution to each (unbinned) dataset, including a uniform background distribution with the relative fraction between the two PDFs given by the fit parameter z . The teal vertical lines and shaded regions give the median and 68% central interval of the data, respectively. Note the different abscissa in the bottom row.

the energy ranges considered here (see e.g. Fig. 2 in [51]). This approach would be contrary to the aim of a model-independent measurement of the ER yields. In order to keep the results from this measurement model-independent, a Gaussian is chosen instead to fit the S1 and S2 distributions in the data. Given the large p-values for almost all fits and the limited statistics in the data, there is no indication for the need of a more complex fit.

The Gaussian fits to the data are performed using one-dimensional, unbinned negative log-likelihood (NLL) functions of the cS1 and cS2 values in each dataset. The fitted PDFs consist of the sum of a Gaussian and a uniform distribution, where the uniform distribution can account for possible background events with a flat spectrum in the ROI. The total number of events is not included in the fit, instead the relative normalization between the Gaussian and the uniform distribution is fit. The best-fit PDFs for each dataset are shown in red, scaled to the total number of events in the datasets for illustrative purposes. Additionally, in accordance with the estimated energy spectra from Figure 7.22, the median and 68% ranges of the data are indicated in the figures as well, represented by the teal-colored vertical lines and shaded regions, and are found to be in reasonable agreement with the best-fit results. The S1 and S2 yields are extracted from the best-fit results, using μ and the error on μ as the best-fit estimator and its uncertainty, respectively.

Converting the best-fit signal yields to the absolute ER quanta yields is done by dividing the

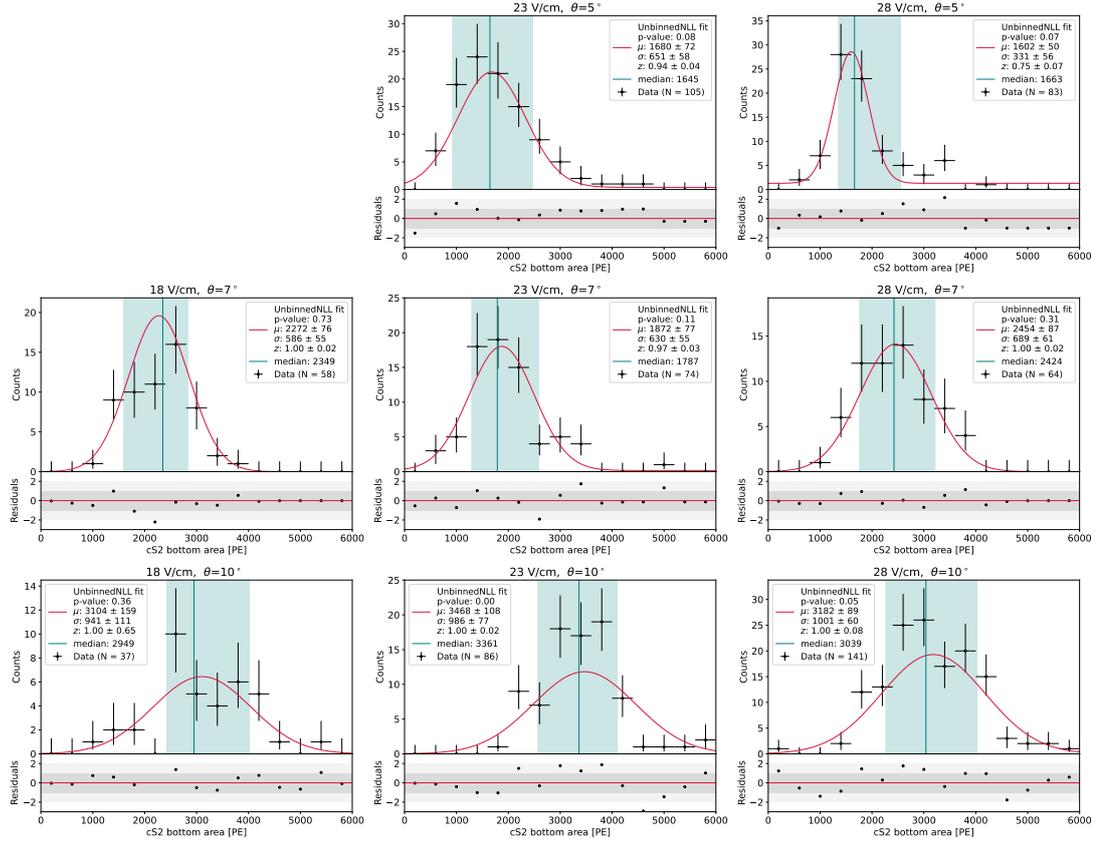


Figure 7.25.: Distributions in cS2 for the selected coincident γ -ray events for all measured drift field (columns) and central scattering angle (rows) combinations. Red curves show the results of fitting a Gaussian distribution to each (unbinned) dataset, including a uniform background distribution with the relative fraction between the two PDFs given by the fit parameter z . The teal vertical lines and shaded regions give the median and 68 % inner range, respectively.

signal yields to the medians of the expected energy distributions, as well as by the scintillation and ionization gains g_1 and g_2 determined from the ^{83m}Kr data. Figure 7.26 shows the absolute γ -ray ER photon and charge yields in dependence of the energy, with the measurements at different drift field strengths encoded in the different colors. The horizontal error bars indicate the 68 % range of the expected energy distributions as shown in Figure 7.22, while the vertical error bars correspond to the uncertainties on the best-fit values for μ . The relative uncertainties on g_1 and g_2 are propagated to the absolute yields as systematic uncertainties, and are plotted as the vertical shaded regions around the best-fit values. The predicted ER photon and charge yields from the NEST framework [71] are shown as dashed lines for comparison, in colors corresponding to the three drift fields.

In general, the expected trend of increasing photon yield and decreasing charge yields with increasing energy is observed, caused by an increasing recombination fraction of electron-ion pairs due to an increasing dE/dx of the Compton electrons in LXe and consequently higher ionization density within this energy range. However, a significant discrepancy between the measured and the NEST-predicted yields is observed, with both the measured photon and charge yields consistently higher than the NEST predictions. A likely explanation for an upwards bias in the extracted yields is the not-accounted-for skewed shapes of the energy distribution in the fits to the cS1 and cS2 spectra. Especially for the two lower energy measurements, this bias is expected to be a non-negligible effect, judging from the observed spectral shapes in Figure 7.24 and Figure 7.25. The datasets for the lowest energy at ~ 4.1 keV ($\theta_o=5^\circ$) are likely additionally affected by not accounting for the energy threshold of the TPC in the estimation of the median

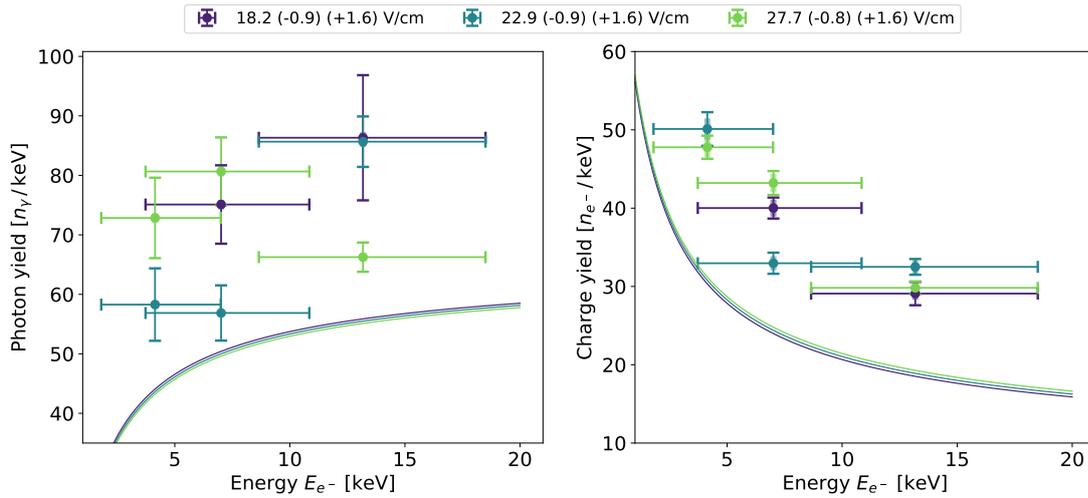


Figure 7.26.: ER γ -ray photon (left) and charge (right) yields in dependence of deposited energy, for the three drift field strengths as labeled in the legend. Solid lines show the NEST predictions for β and Compton ER yields, for the three drift field strengths in the respective colors.

energy. A visual inspection of the CES energy spectrum from all γ -ray events in the ER band at low energies shows a steep roll-off starting for energies below about 2 keV. Comparing this to the expected distribution of energy depositions in the $\theta_o = 5^\circ$ configuration (Figure 7.22) reveals that a not insignificant fraction of the spectrum falls below this threshold, making the observable energy distribution in the TPC even more skewed towards lower energies. This is further supported by the low p-values of the fits in the cS1 spectra for the 5° configuration. In addition to the expected upwards bias in the extracted yields, a bias in the estimation of g_1 and g_2 from the $^{83\text{m}}\text{Kr}$ data may also contribute to the discrepancy between the measured and the NEST-predicted yields, being determined from only three data points of the same energy at very close signal yields. As discussed in subsection 7.2.2, the S1 and S2 distributions of the $^{83\text{m}}\text{Kr}$ data show a significant deviation from the Gaussian shapes expected for mono-energetic signals, an as yet not understood effect. Furthermore, the estimation of the expected energy distributions from the Compton kinematics is based on an approximate simulation, and may not produce the correct energy distributions. As already discussed, a full GEANT4 simulation of the HeXe TPC setup including the ^{137}Cs source and the NaI detector would be necessary to evaluate the energy distributions in the TPC more accurately, including potential effects from multiple Compton scatters.

On the other hand, the results shown here represent the first measurements of γ -ray ER yields in the drift field range around 23 V/cm. As previously stated, only little literature data on photon and charge yields in LXe at small drift fields is available, such that the NEST prediction may not be accurate in this regime. In particular, two measurements of ER photon yields at zero drift field exist in literature [113, 119], as well as one measurement at 60 V/cm [73] for both photon and charge yield. These however also show significantly lower photon and charge yields than the measurements presented here. Figure 7.27 shows the ER photon and charge yields measured in this work in dependence of the drift field in comparison to the NEST predictions. Apart from the overall discrepancy between the measured and the NEST predicted yields already discussed, no strong trend of the yields with the drift field is observed within the scatter of the data points, in accordance with the NEST prediction of a weak dependence in this narrow field range.

Overall, these preliminary results of the first measurements of the ER γ -ray photon and charge yields in LXe at drift fields around 23 V/cm at the current stage of the analysis do not show a good agreement with NEST predictions. A few potential systematic uncertainties are identified and need to be investigated in the future. In combination with the limited statistics of the

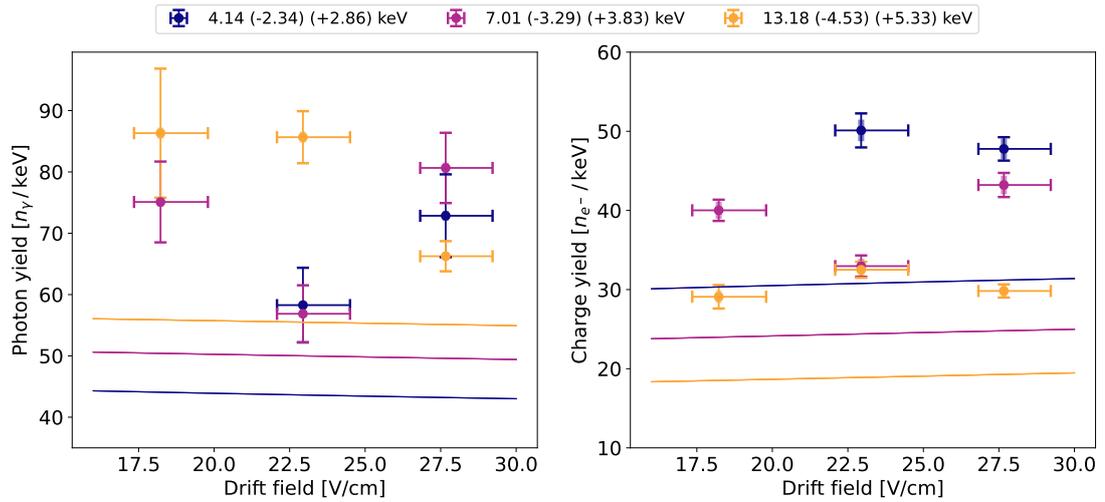


Figure 7.27.: ER γ -ray photon (left) and charge (right) yields in dependence of the drift field strength, measured at the three central scattering angles with corresponding median Compton electron energies as stated in the legend. The NEST prediction of the field dependence for the yields is shown by the solid lines.

data, the current results do not allow for a conclusive statement on the ER yields in LXe at small drift fields at this time.

7.4. Summary, conclusions and outlook

In this chapter, preliminary results of the first measurements of the ER γ -ray photon and charge yields in LXe at drift fields around 23 V/cm in the HeXe TPC were presented. The measurements were performed using a ^{137}Cs source and a NaI detector in coincidence with the TPC, allowing for the selection of γ -ray events at energies given by Compton scattering kinematics for differing central scattering angles. The response of the HeXe TPC was characterized using $^{83\text{m}}\text{Kr}$ calibration data, including the development of data selection criteria and signal corrections, and the scintillation and ionization gains g_1 and g_2 were determined. A cross-calibration of the timing between the NaI and the TPC was performed in order to select coincident γ -ray events in the TPC. The angular distribution and resulting energy spectra of the selected γ -ray events in the TPC were estimated using a simple geometric solid angle calculation, while the electric field strength and spread in the TPC were estimated using 3D COMSOL simulations.

The resulting ER photon and charge yields were determined from the selected γ -ray events in the TPC, and were found to be higher than the NEST predictions for all measured energies and drift field configurations. The discrepancy between the measured and the NEST-predicted yields is likely due to systematic biases in the analysis procedure as well as several effects that are not yet accounted for, most prominently the expected skewed shapes of the underlying energy distributions in the fits to the cS1 and cS2 spectra. While the current method of determining the expected energy distributions from Compton kinematics will be improved as a first step by including the z -coordinate in the estimation, a full GEANT4 simulation of the setup is expected to give more accurate results, and will additionally allow for the study of multiple Compton scatters.

Despite the already ongoing and planned improvements in the analysis procedure, the limited statistics in this dataset will remain a challenge. In order to improve the precision of the ER γ -ray yields in the HeXe TPC, a new measurement run with improved statistics and accounting for the lessons learned from this work should be considered. Improving statistics, firstly, could

be achieved by restricting the measurement to one drift field strength, prioritizing the 23 V/cm configuration, as within the narrow spread of the electric fields utilized in this work, no strong trend of the ER yields with the drift field was observed. Better energy resolution may be achieved by a tighter collimation of the γ -ray beam, which would allow for a more precise selection of the energy range in the TPC. This would however in turn reduce the available statistics, and may not be feasible with the current setup due to the limitations of the manual collimation system. Alternatively, a different external detector with better energy resolution should be considered. In particular, germanium (Ge) detectors represent a good choice for this purpose, with typical energy resolutions below about 1 keV at the ^{137}Cs line, significantly better compared to the ~ 40 keV resolution of the NaI detector used in this work. This choice would additionally allow for scanning over the detected energy range in the external detector even within one central angle configuration, as e.g. done in [120], thereby removing the need for setting up different scattering angle configurations by hand. However, the lower detection efficiency of Ge detectors compared to NaI detectors may again lower the rate of coincident events. Improvements on the HeXe TPC setup can be considered as well, several upgrades are planned for the near future. A change to an array 2×2 1-inch square PMTs in the top array of the TPC would allow for the 3D event position reconstruction necessary to understand the dependence of signal yields in the xy -plane from geometrical effects.

In conclusion, first measurements of ER γ -ray photon and charge yields in LXe at unprecedentedly low drift fields around 23 V/cm were presented here, that, with further improvements, can eventually inform and improve models for low energy ER yields in LXe at small drift fields in XENONnT, as well as other current and future LXe TPCs.

Chapter 8.

Summary

The existence of Dark Matter (DM) is supported by a wealth of astrophysical and cosmological observations, however its nature remains unknown. The XENONnT experiment is aiming for the detection of Weakly Interacting Massive Particles (WIMPs), one of the best motivated particle candidates for DM, using a dual-phase xenon time projection chamber (TPC) with 5.9 tonnes of active target mass. In this detector, particle interactions in liquid xenon (LXe) produce prompt scintillation light as well as delayed scintillation light from ionization electrons, both of which are detected via two arrays of PMTs.

The XENONnT TPC uses 494 Hamamatsu R11410-21 3-inch photomultiplier tubes (PMTs), especially developed for the use in a cryogenic xenon environment for the detection of the VUV scintillation light of liquid xenon at high quantum efficiencies. Their characterization, selection, and distribution in the arrays was the focus of [chapter 3](#). The available tubes for the TPC were selected based on their performance in the predecessor experiment XENON1T as well as in a dedicated PMT testing campaign for new PMTs [23]. The most significant problem for long-term PMT performance in XENON1T was vacuum degradation via the penetration of xenon into the tubes through tiny leaks, that caused a significant fraction of the PMTs to fail over time. Based on the results of the new testing campaign, a quantitative categorization system was developed and used for the selection and distribution of the PMTs in the XENONnT PMT arrays, predicting the probability of vacuum degradation in each tube based on several parameters. The XENONnT PMT arrays were then assembled with PMTs placed according to the scheme derived in this study, followed by the assembly of the TPC and its installation in the water tank.

During both the commissioning phase and the first science run of the XENONnT experiment, SR0, the performance of the PMTs was continuously monitored, which was the subject of [chapter 4](#). For this purpose, weekly LED calibrations were performed, and a dedicated processing and afterpulse analysis procedure was developed within the XENONnT analysis framework. This enabled the continuous monitoring for potential vacuum degradation via the calculation of afterpulse rates from relevant ion species for all PMTs. Consistent with the results from the PMT testing campaign, a much smaller fraction of external xenon leakage is observed in XENONnT compared to XENON1T. A new effect of internal nitrogen outgassing was discovered, that affected a small fraction of about 4% of the tubes. The affected tubes could be identified and excluded from the analysis based on the observed afterpulse rates, avoiding negatively impacting the overall data quality. Additionally, the quantitative categorization system developed for PMT selection and distribution within the arrays was evaluated and proven to be highly effective in predicting PMT failures and preventing clusters of dead PMTs in critical regions of the arrays. Over more than 3 years of operation until now, the PMTs have shown stable performance, with 96% of the PMTs still in operation, confirming the long-term reliability of the PMTs in the XENONnT experiment.

In [chapter 5](#), we turned to the analysis of SR0 in the context of the first WIMP search results of the XENONnT experiment, focusing on the modeling of the Nuclear Recoil (NR) response

of the detector [51]. This task is essential for predicting the expected WIMP signal and NR background distributions in the analysis space, as well as differentiating WIMP signal from Electronic Recoil (ER) background events in the WIMP DM search. A Markov Chain Monte Carlo (MCMC) method was used to match fast simulations of the liquid xenon emission and detector reconstruction models to AmBe-neutron calibration data. In this work, in particular multi-scatter (MS) neutron event topologies with merging signals were addressed, an effect that is especially relevant due to the low drift field in XENONnT. The resulting NR model was extensively validated, including several goodness-of-fit tests that indicated no mismatch between model and data, confirming that the NR band shape in the analysis space accurately describes the calibration data. The extracted scintillation photon and ionization electron yields from the LXe emission model for NR interactions were shown to agree well with available literature.

The NR response model was then used in the analysis of the WIMP DM search data from the first science run of the XENONnT experiment presented in [chapter 6](#). In a blind analysis with a total exposure of 1.09 tonne-years, no significant excess of events above the expected backgrounds was observed. This allowed setting strong upper limits at 90% confidence level on the spin-independent (SI) WIMP-nucleon interaction cross-section, with a minimum of $2.58 \times 10^{-47} \text{ cm}^2$ at a WIMP mass of $28 \text{ GeV}/c^2$. Future WIMP search results from the XENONnT experiment are expected to further improve the sensitivity to WIMP DM and probe new regions in the WIMP parameter space, with an increased exposure, a lower ^{222}Rn ER background rate, and a higher neutron veto tagging efficiency.

A crucial aspect of achieving increasingly sensitive measurements with LXe TPCs is the precise characterization of the target material's response under various detector conditions, which can be effectively conducted using small-scale detectors. In [chapter 7](#), we investigated the absolute photon and charge yields for ERs in LXe at low drift fields, using the local Heidelberg Xenon (HeXe) system that features a gram-scale dual-phase xenon TPC. For this measurement, a dedicated setup was developed to measure the response of LXe to γ -ray Compton interactions in coincidence with an external detector, allowing for the selection of known energy depositions in the LXe. The HeXe TPC and the external detector were calibrated and characterized, the coincidence timing between the two detectors determined, and the expected energy depositions in the TPC under the set scattering angle acceptances were estimated. The coincidence dataset was analyzed to extract the absolute scintillation photon and ionization electron yields in an energy range of about 4–13 keV, at drift fields between 18 and 28 V/cm. Preliminary results indicated that the measured ER photon and electron yields were higher than predictions from dedicated models fit to literature data, that do however not include measurements at low, non-zero drift fields. In this regard, the HeXe measurement is the first in this field range, and provides a valuable contribution to the understanding of ER yields in LXe at low energies and fields. A refined analysis as well as future measurements with an upgraded setup will likely be able to improve ER models that contribute to the understanding of one of the main background sources in DM searches with xenon TPCs.

This work has contributed to the planning and commissioning of the XENONnT experiment and its first WIMP DM search results, focusing on PMT characterization and performance monitoring, the modeling of the NR response of the detector, and the measurement of ER yields in LXe at low energies and drift fields.

Appendix

Appendix A.

Author's Publications

This list contains the publications of the author produced in the context of and used in this work. The author took a significant part in the work described in the publications as well as in the writing process, and is a corresponding author. The publications are referenced in the text by their respective numbers in square brackets. The list is sorted by date of publication.

Publication 1 [23]:

- V.C. Antochi, L. Baudis, J. Bollig, A. Brown, R. Budnik, D. Cichon, J. Conrad, A.D. Ferella, M. Galloway, L. Höttsch, S. Kazama, G. Koltman, H. Landsman, M. Lindner, J. Mahlstedt, T. Marrodán Undagoitia, B. Pelssers, G. Volta, O. Wack, J. Wulf (2021). **Improved quality tests of R11410-21 photomultiplier tubes for the XENONnT experiment.** *arXiv:2104.15051*, published in *Journal of Instrumentation*, 16(08), P08033.
- **Abstract:** Photomultiplier tubes (PMTs) are often used in low-background particle physics experiments, which rely on an excellent response to single-photon signals and stable long-term operation. In particular, the Hamamatsu R11410 model is the light sensor of choice for liquid xenon dark matter experiments, including XENONnT. The same PMT model was also used for the predecessor, XENON1T, where issues affecting its long-term operation were observed. Here, we report on an improved PMT testing procedure which ensures optimal performance in XENONnT. Using both new and upgraded facilities, we tested 368 new PMTs in a cryogenic xenon environment. We developed new tests targeted at the detection of light emission and the degradation of the PMT vacuum through small leaks, which can lead to spurious signals known as afterpulses, both of which were observed in XENON1T. We exclude the use of 26 of the 368 tested PMTs and categorise the remainder according to their performance. Given that we have improved the testing procedure, yet we rejected fewer PMTs, we expect significantly better PMT performance in XENONnT.

Publication 2 [51]:

- XENON Collaboration, E. Aprile *et al.* (2024). **XENONnT WIMP Search: Signal & Background Modeling and Statistical Inference.** Under final internal collaboration review. Prepared for submission to *Physical Review D*, preprint expected to be available on arXiv by July 2024.
- **Abstract (preliminary):** The XENONnT experiment searches for weakly-interacting massive particle (WIMP) dark matter scattering off a xenon nucleus. In particular, XENONnT uses a dual-phase time projection chamber with a 5.9-tonne liquid xenon target, detecting both scintillation and ionization signals to reconstruct the energy, position, and type of recoil. A blind search for nuclear recoil WIMPs with an exposure of 1.1 tonne-years yielded no signal excess over background expectations, from which com-

petitive exclusion limits were derived on WIMP-nucleon elastic scatter cross sections, for WIMP masses ranging from $6 \text{ GeV}/c^2$ up to the TeV/c^2 scale. This work details the modeling and statistical methods employed in this search. By means of calibration data, we model the detector response, which is then used to derive background and signal models. The construction and validation of these models is discussed, alongside additional purely data-driven backgrounds. We also describe the statistical inference framework, including the definition of the likelihood function and the construction of confidence intervals.

Appendix B.

References

- [1] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helv. Phys. Acta*, vol. 6, p. 110, 1933, [Gen. Rel. Grav.41,207(2009)]. DOI: [10.1007/s10714-008-0707-4](https://doi.org/10.1007/s10714-008-0707-4).
- [2] V. C. Rubin *et al.*, “Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa through Sc,” *Astrophys. J.*, vol. 225, p. L107, 1978. DOI: [10.1086/182804](https://doi.org/10.1086/182804).
- [3] M. Bartelmann and P. Schneider, “Weak gravitational lensing,” *Phys. Rept.*, vol. 340, pp. 291–472, 2001. DOI: [10.1016/S0370-1573\(00\)00082-X](https://doi.org/10.1016/S0370-1573(00)00082-X). arXiv: [astro-ph/9912508](https://arxiv.org/abs/astro-ph/9912508).
- [4] P. A. R. Ade *et al.*, “Planck 2015 results. XIII. Cosmological parameters,” *Astron. Astrophys.*, vol. 594, A13, 2016. arXiv: [1502.01589](https://arxiv.org/abs/1502.01589).
- [5] V. Springel *et al.*, “Simulating the joint evolution of quasars, galaxies and their large-scale distribution,” *Nature*, vol. 435, pp. 629–636, 2005. DOI: [10.1038/nature03597](https://doi.org/10.1038/nature03597). arXiv: [astro-ph/0504097](https://arxiv.org/abs/astro-ph/0504097).
- [6] T. Marrodán Undagoitia and L. Rauch, “Dark matter direct-detection experiments,” *J. Phys. G*, vol. 43, no. 1, p. 013001, 2016. DOI: [10.1088/0954-3899/43/1/013001](https://doi.org/10.1088/0954-3899/43/1/013001). arXiv: [1509.08767](https://arxiv.org/abs/1509.08767) [[physics.ins-det](https://arxiv.org/abs/1509.08767)].
- [7] D. Baxter *et al.*, “Recommended conventions for reporting results from direct dark matter searches,” *Eur. Phys. J. C*, vol. 81, no. 10, p. 907, 2021. DOI: [10.1140/epjc/s10052-021-09655-y](https://doi.org/10.1140/epjc/s10052-021-09655-y). arXiv: [2105.00599](https://arxiv.org/abs/2105.00599) [[hep-ex](https://arxiv.org/abs/2105.00599)].
- [8] XENON, E. Aprile *et al.*, “The XENONnT Dark Matter Experiment,” Feb. 2024. arXiv: [2402.10446](https://arxiv.org/abs/2402.10446) [[physics.ins-det](https://arxiv.org/abs/2402.10446)].
- [9] K. Fujii *et al.*, “High-accuracy measurement of the emission spectrum of liquid xenon in the vacuum ultraviolet region,” *Nucl. Instrum. Meth. A*, vol. 795, pp. 293–297, 2015. DOI: [10.1016/j.nima.2015.05.065](https://doi.org/10.1016/j.nima.2015.05.065).
- [10] A. L. et. al., “Development research on a highly luminous condensed xenon scintillator,” *Nuclear Instruments and Methods*, vol. 135, no. 1, pp. 47–52, 1976, ISSN: 0029-554X. DOI: [https://doi.org/10.1016/0029-554X\(76\)90824-7](https://doi.org/10.1016/0029-554X(76)90824-7). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0029554X76908247>.
- [11] XENON Collaboration, E. Aprile *et al.*, “XENON1T Dark Matter Data Analysis: Signal Reconstruction, Calibration and Event Selection,” *Phys. Rev. D*, vol. 100, no. 5, p. 052014, 2019. 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](https://arxiv.org/abs/1906.04717)].
- [12] T. Doke, A. Hitachi, J. Kikuchi, K. Masuda, H. Okada, and E. Shibamura, “Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles,” *Jap. J. Appl. Phys.*, vol. 41, pp. 1538–1545, 2002. DOI: [10.1143/JJAP.41.1538](https://doi.org/10.1143/JJAP.41.1538).
- [13] XENON, E. Aprile *et al.*, “The XENON1T Dark Matter Experiment,” *Eur. Phys. J. C*, vol. 77, no. 12, p. 881, 2017. DOI: [10.1140/epjc/s10052-017-5326-3](https://doi.org/10.1140/epjc/s10052-017-5326-3). arXiv: [1708.07051](https://arxiv.org/abs/1708.07051) [[astro-ph.IM](https://arxiv.org/abs/1708.07051)].

- [14] XENON Collaboration, E. Aprile *et al.*, “Design and performance of the field cage for the XENONnT experiment,” *Eur. Phys. J. C*, vol. 84, no. 2, p. 138, 2024. DOI: [10.1140/epjc/s10052-023-12296-y](https://doi.org/10.1140/epjc/s10052-023-12296-y). arXiv: [2309.11996](https://arxiv.org/abs/2309.11996) [hep-ex].
- [15] XENON, E. Aprile *et al.*, “Projected WIMP sensitivity of the XENONnT dark matter experiment,” *JCAP*, vol. 11, p. 031, 2020. 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].
- [16] XENON, E. Aprile *et al.*, “Removing krypton from xenon by cryogenic distillation to the ppq level,” *Eur. Phys. J. C*, vol. 77, no. 5, p. 275, 2017. DOI: [10.1140/epjc/s10052-017-4757-1](https://doi.org/10.1140/epjc/s10052-017-4757-1). arXiv: [1612.04284](https://arxiv.org/abs/1612.04284) [physics.ins-det].
- [17] XENON100, E. Aprile *et al.*, “Online²²² Rn removal by cryogenic distillation in the XENON100 experiment,” *Eur. Phys. J. C*, vol. 77, no. 6, p. 358, 2017. DOI: [10.1140/epjc/s10052-017-4902-x](https://doi.org/10.1140/epjc/s10052-017-4902-x). arXiv: [1702.06942](https://arxiv.org/abs/1702.06942) [physics.ins-det].
- [18] XENON Collaboration, E. Aprile *et al.*, “Material radiopurity control in the XENONnT experiment,” *Eur. Phys. J. C*, vol. 82, no. 7, p. 599, 2022. DOI: [10.1140/epjc/s10052-022-10345-6](https://doi.org/10.1140/epjc/s10052-022-10345-6). arXiv: [2112.05629](https://arxiv.org/abs/2112.05629) [physics.ins-det].
- [19] G. Plante, E. Aprile, J. Howlett, and Y. Zhang, “Liquid-phase purification for multi-tonne xenon detectors,” *Eur. Phys. J. C*, vol. 82, no. 10, p. 860, 2022. 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].
- [20] LZ, D. S. Akerib *et al.*, “The LUX-ZEPLIN (LZ) Experiment,” *Nucl. Instrum. Meth. A*, vol. 953, p. 163047, 2020. DOI: [10.1016/j.nima.2019.163047](https://doi.org/10.1016/j.nima.2019.163047). arXiv: [1910.09124](https://arxiv.org/abs/1910.09124) [physics.ins-det].
- [21] PandaX, H. Zhang *et al.*, “Dark matter direct search sensitivity of the PandaX-4T experiment,” *Sci. China Phys. Mech. Astron.*, vol. 62, no. 3, p. 31011, 2019. DOI: [10.1007/s11433-018-9259-0](https://doi.org/10.1007/s11433-018-9259-0). arXiv: [1806.02229](https://arxiv.org/abs/1806.02229) [physics.ins-det].
- [22] XENON Collaboration, E. Aprile *et al.*, “Lowering the radioactivity of the photomultiplier tubes for the XENON1T dark matter experiment,” *Eur. Phys. J. C*, vol. 75, no. 11, p. 546, 2015. DOI: [10.1140/epjc/s10052-015-3657-5](https://doi.org/10.1140/epjc/s10052-015-3657-5). arXiv: [1503.07698](https://arxiv.org/abs/1503.07698) [astro-ph.IM].
- [23] V. C. Antochi *et al.*, “Improved quality tests of R11410-21 photomultiplier tubes for the XENONnT experiment,” *JINST*, vol. 16, no. 08, P08033, 2021. DOI: [10.1088/1748-0221/16/08/P08033](https://doi.org/10.1088/1748-0221/16/08/P08033). arXiv: [2104.15051](https://arxiv.org/abs/2104.15051) [physics.ins-det].
- [24] T. Hakamata *et al.*, *Photomultiplier Tubes: Basics and Applications*, Edition 3a, HAMAMATSU Photonics, 2007.
- [25] Volta, G., “Characterization and Monitoring of XENONnT Photosensors and Search for New Physics with the First XENONnT Science Data,” Ph.D. dissertation, Universität Zürich, 2023.
- [26] K. Nakamura, Y. Hamana, Y. Ishigami, and T. Matsui, “Latest bialkali photocathode with ultra high sensitivity,” *Nucl. Instrum. Meth. A*, vol. 623, H. Iwasaki, T. K. Komatsubara, and Y. Sugimoto, Eds., pp. 276–278, 2010. DOI: [10.1016/j.nima.2010.02.220](https://doi.org/10.1016/j.nima.2010.02.220).
- [27] L. Höttsch, “Studies towards an Annual Modulation Search with XENON1T data and Qualification Tests of Photosensors for the XENONnT detector,” Master’s thesis, Ruprecht-Karls-Universität Heidelberg, 2019.
- [28] HAMAMATSU Photonics, *Photomultiplier Tubes and Assemblies for Scintillation Counting and High Energy Physics*, accessed: 2024-02-13. [Online]. Available: [7Bhttps://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99%5C_SALES%5C_LIBRARY/etd/High%5C_energy%5C_PMT%5C_TPMZ0003E.pdf%7D](https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99%5C_SALES%5C_LIBRARY/etd/High%5C_energy%5C_PMT%5C_TPMZ0003E.pdf%7D).
- [29] D. Y. Akimov *et al.*, “Observation of light emission from Hamamatsu R11410-20 photomultiplier tubes,” *Nucl. Instrum. Meth. A*, vol. 794, pp. 1–2, 2015. DOI: [10.1016/j.nima.2015.04.066](https://doi.org/10.1016/j.nima.2015.04.066). arXiv: [1504.07651](https://arxiv.org/abs/1504.07651) [physics.ins-det].

-
- [30] D. Akimov *et al.*, “Peculiarities of the Hamamatsu R11410-20 photomultiplier tubes,” *PoS*, vol. PhotoDet2015, p. 025, 2016. DOI: [10.22323/1.252.0025](https://doi.org/10.22323/1.252.0025).
- [31] P. Barrow *et al.*, “Qualification Tests of the R11410-21 Photomultiplier Tubes for the XENON1T Detector,” *JINST*, vol. 12, no. 01, P01024, 2017. arXiv: [1609.01654](https://arxiv.org/abs/1609.01654).
- [32] A. Brown, “Search for Elastic and Inelastic Dark Matter Interactions in XENON1T and Light Detection for XENONnT,” Ph.D. dissertation, University of Zurich, 2020.
- [33] A. Lyashenko *et al.*, “Measurement of the absolute Quantum Efficiency of Hamamatsu model R11410-10 photomultiplier tubes at low temperatures down to liquid xenon boiling point,” *JINST*, vol. 9, no. 11, P11021, 2014. arXiv: [1410.3890](https://arxiv.org/abs/1410.3890).
- [34] B. López Paredes *et al.*, “Response of photomultiplier tubes to xenon scintillation light,” *Astropart. Phys.*, vol. 102, p. 56, 2018. arXiv: [1801.01597](https://arxiv.org/abs/1801.01597).
- [35] C. H. Faham *et al.*, “Measurements of wavelength-dependent double photoelectron emission from single photons in VUV-sensitive photomultiplier tubes,” *JINST*, vol. 10, no. 09, P09010, 2015. arXiv: [1506.08748](https://arxiv.org/abs/1506.08748).
- [36] R. Saldanha, L. Grandi, Y. Guardincerri, and T. Wester, “Model Independent Approach to the Single Photoelectron Calibration of Photomultiplier Tubes,” *Nucl. Instrum. Meth. A*, vol. 863, G. Xiao, H. Shen, and G. Du, Eds., pp. 35–46, 2017. DOI: [10.1016/j.nima.2017.02.086](https://doi.org/10.1016/j.nima.2017.02.086). arXiv: [1602.03150](https://arxiv.org/abs/1602.03150) [[physics.ins-det](https://arxiv.org/archive/physics)].
- [37] Mayani Paras, D., “Photomultiplier Tubes for the XENON1T Dark Matter Experiment and Studies on the XENON100 Electromagnetic Background,” Ph.D. dissertation, Universität Zürich, 2017.
- [38] Rauch, L., “From Final Dark Matter Results and Background Shape Uncertainties in XENON100 to First Light in XENON1T,” Ph.D. dissertation, Ruprecht-Karls-Universität Heidelberg, 2017.
- [39] Wulf, J., “Direct Dark Matter Search with XENON1T and Developments for Multi-Ton Liquid Xenon Detectors,” Ph.D. dissertation, Universität Zürich, 2018.
- [40] H. Photonics, Private communication.
- [41] XENON Collaboration, E. Aprile *et al.*, “Double-Weak Decays of ^{124}Xe and ^{136}Xe in the XENON1T and XENONnT Experiments,” *Phys. Rev. C*, vol. 106, no. 2, p. 024328, 2022. DOI: [10.1103/PhysRevC.106.024328](https://doi.org/10.1103/PhysRevC.106.024328). arXiv: [2205.04158](https://arxiv.org/abs/2205.04158) [[hep-ex](https://arxiv.org/archive/hep)].
- [42] XENON Collaboration, *Internal Note*.
- [43] K. J. Ma *et al.*, “Time and Amplitude of Afterpulse Measured with a Large Size Photomultiplier Tube,” *Nucl. Instrum. Meth.*, vol. A629, p. 93, 2011. arXiv: [0911.5336](https://arxiv.org/abs/0911.5336).
- [44] Coates, P. B., “The origins of afterpulses in photomultipliers,” *Journal of Physics D: Applied Physics*, vol. 6, no. 10, 1973. DOI: [10.1088/0022-3727/6/10/301](https://doi.org/10.1088/0022-3727/6/10/301).
- [45] A. Behrens, “Light Detectors for the XENON100 and XENON1T Dark Matter Search Experiments,” Ph.D. dissertation, University of Zurich, 2014.
- [46] *COMSOL Multiphysics®*, <https://www.comsol.com/comsol-multiphysics>, online, accessed 4-April-2024.
- [47] Capelli, C., “Search for Dark Matter and Neutrinoless Double Beta Decay in XENON1T and Calibration of the Photosensors in XENONnT,” Ph.D. dissertation, Universität Zürich, 2020.
- [48] J. Aalbers *et al.*, *AxFoundation/strax: Streaming analysis for xenon experiments*, 2023. DOI: [10.5281/zenodo.1340632](https://doi.org/10.5281/zenodo.1340632). [Online]. Available: <https://github.com/AxFoundation/strax>.
- [49] XENON Collaboration, *XENONnT/straxen: Streaming analysis for XENON(nT)*, 2022. DOI: [10.5281/zenodo.5576262](https://doi.org/10.5281/zenodo.5576262). [Online]. Available: <https://github.com/XENONnT/straxen>.

- [50] S. Das, “A simple alternative to the Crystal Ball function,” Mar. 2016. arXiv: [1603.08591 \[hep-ex\]](#).
- [51] XENON Collaboration, E. Aprile *et al.*, “XENONnT WIMP Search: Signal & Background Modeling and Statistical Inference,” 2024, in final preparation.
- [52] Daniel Wenz, “Commissioning of the world’s first water Cherenkov neutron veto and first WIMP dark matter search results of the XENONnT experiment,” Ph.D. dissertation, Johannes Gutenberg-Universität Mainz, 2023. DOI: <http://doi.org/10.25358/openscience-9654>.
- [53] National Nuclear Data Center (NNDC), Brookhaven National Laboratory, *NuDat 3.0 Database*, <https://www.nndc.bnl.gov/nudat3/>, accessed 13-February-2024.
- [54] K. Geiger and L. Van Der Zwan, “Radioactive neutron source spectra from $^9\text{Be}(\alpha, n)$ cross section data,” *Nuclear Instruments and Methods*, vol. 131, no. 2, pp. 315–321, 1975, ISSN: 0029-554X. DOI: [https://doi.org/10.1016/0029-554X\(75\)90336-5](https://doi.org/10.1016/0029-554X(75)90336-5).
- [55] XENON Collaboration, E. Aprile *et al.*, “Response of the XENON100 Dark Matter Detector to Nuclear Recoils,” *Phys. Rev. D*, vol. 88, p. 012006, 2013. DOI: [10.1103/PhysRevD.88.012006](https://doi.org/10.1103/PhysRevD.88.012006). arXiv: [1304.1427 \[astro-ph.IM\]](#).
- [56] F. Toschi, “Design of the field cage and charge response of the xenonnt dark matter experiment,” Ph.D. dissertation, Albert-Ludwigs-Universität Freiburg, 2022. DOI: <https://doi.org/10.6094/UNIFR/234323>.
- [57] XENON Collaboration, E. Aprile *et al.*, “First Dark Matter Search Results from the XENON1T Experiment,” *Phys. Rev. Lett.*, vol. 119, no. 18, p. 181301, 2017. DOI: [10.1103/PhysRevLett.119.181301](https://doi.org/10.1103/PhysRevLett.119.181301). arXiv: [1705.06655 \[astro-ph.CO\]](#).
- [58] XENON Collaboration, E. Aprile *et al.*, “Dark Matter Search Results from a One Ton-Year Exposure of XENON1T,” *Phys. Rev. Lett.*, vol. 121, no. 11, p. 111302, 2018. DOI: [10.1103/PhysRevLett.121.111302](https://doi.org/10.1103/PhysRevLett.121.111302). arXiv: [1805.12562 \[astro-ph.CO\]](#).
- [59] XENON Collaboration, E. Aprile *et al.*, “XENON1T dark matter data analysis: Signal and background models and statistical inference,” *Phys. Rev. D*, vol. 99, no. 11, p. 112009, 2019. DOI: [10.1103/PhysRevD.99.112009](https://doi.org/10.1103/PhysRevD.99.112009). arXiv: [1902.11297 \[physics.ins-det\]](#).
- [60] Matthew Duschl Anthony, “Understanding Low-Energy Nuclear Recoils in Liquid Xenon for Dark Matter Searches and the First Results of XENON1T,” Ph.D. dissertation, Columbia University, 2018. DOI: <https://doi.org/10.7916/D8RN4MBV>.
- [61] Zachary Schuyler Greene, “The XENON1T Spin-Independent WIMP Dark Matter Search Results and a Model to Characterize the Reduction of Electronegative Impurities in Its 3.2 Tonne Liquid Xenon Detector,” Ph.D. dissertation, Columbia University, 2018. DOI: [10.7916/D87M1RTN](https://doi.org/10.7916/D87M1RTN).
- [62] Joseph Howlett, “Liquid-phase purification for multi-ton xenon detectors and a search for dark matter and neutrinos in XENON1T,” Ph.D. dissertation, Columbia University, 2022. DOI: [10.7916/9n8y-7e64](https://doi.org/10.7916/9n8y-7e64).
- [63] Goodman, J., Weare, J., “Ensemble samplers with affine invariance,” *Communications in Applied Mathematics and Computational Science*, vol. 5, no. 1, 65–80, 2010. DOI: <http://dx.doi.org/10.2140/camcos.2010.5.65>.
- [64] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, “emcee: The MCMC Hammer,” *Publications of the Astronomical Society of the Pacific*, vol. 125, no. 925, pp. 306–312, 2013, ISSN: 1538-3873. DOI: [10.1086/670067](https://doi.org/10.1086/670067). arXiv: [1202.3665 \[astro-ph.IM\]](#).
- [65] Foreman-Mackey, Daniel and others, “emcee v3: A Python ensemble sampling toolkit for affine-invariant MCMC,” *The Journal of Open Source Software*, vol. 4, no. 43, p. 1864, 2019. DOI: [10.21105/joss.01864](https://doi.org/10.21105/joss.01864). arXiv: [1911.07688 \[astro-ph.IM\]](#).
- [66] Ramírez García, Diego, “Simulating the XENONnT Dark Matter Experiment: Backgrounds and WIMP Sensitivity,” Ph.D. dissertation, Albert-Ludwigs-Universität Freiburg, 2022. DOI: <https://doi.org/10.6094/UNIFR/228338>.

-
- [67] Geant4, S. Agostinelli *et al.*, “Geant4: A Simulation toolkit,” *Nucl. Instrum. Meth. A*, vol. 506, pp. 250–303, 2003. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [68] D. Wenz *et al.*, *XENONnT/epix: v0.3.4*, version v0.3.4, Jan. 2023. DOI: <https://doi.org/10.5281/zenodo.7516942>.
- [69] Peter Gaemers and others, *XENONnT/WFSim: v1.0.2*, version v1.0.2, Oct. 2022. DOI: [10.5281/zenodo.7216324](https://doi.org/10.5281/zenodo.7216324).
- [70] B. Lenardo *et al.*, “A Global Analysis of Light and Charge Yields in Liquid Xenon,” *IEEE Trans. Nucl. Sci.*, vol. 62, no. 6, pp. 3387–3396, 2015. DOI: [10.1109/TNS.2015.2481322](https://doi.org/10.1109/TNS.2015.2481322). arXiv: [1412.4417](https://arxiv.org/abs/1412.4417) [astro-ph.IM].
- [71] M. Szydagis *et al.*, *Noble element simulation technique*, version v2.0.0, Jul. 2018. DOI: [10.5281/zenodo.1314669](https://doi.org/10.5281/zenodo.1314669). [Online]. Available: <https://doi.org/10.5281/zenodo.1314669>.
- [72] M. Szydagis *et al.*, “A Review of NEST Models, and Their Application to Improvement of Particle Identification in Liquid Xenon Experiments,” Nov. 2022. arXiv: [2211.10726](https://arxiv.org/abs/2211.10726) [hep-ex].
- [73] C. E. Dahl, “The physics of background discrimination in liquid xenon, and first results from Xenon10 in the hunt for WIMP dark matter,” Ph.D. dissertation, Princeton U., 2009.
- [74] U. Fano, “Ionization yield of radiations. ii. the fluctuations of the number of ions,” *Phys. Rev.*, vol. 72, pp. 26–29, 1 Jul. 1947. DOI: [10.1103/PhysRev.72.26](https://doi.org/10.1103/PhysRev.72.26). [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRev.72.26>.
- [75] T. Doke, A. Hitachi, S. Kubota, A. Nakamoto, and T. Takahashi, “Estimation of fano factors in liquid argon, krypton, xenon and xenon-doped liquid argon,” *Nuclear Instruments and Methods*, vol. 134, no. 2, pp. 353–357, 1976, ISSN: 0029-554X. DOI: [https://doi.org/10.1016/0029-554X\(76\)90292-5](https://doi.org/10.1016/0029-554X(76)90292-5). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0029554X76902925>.
- [76] J. Seguinot, J. Tischhauser, and T. Ypsilantis, “Liquid xenon scintillation: Photon yield and fano factor measurements,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 354, no. 2, pp. 280–287, 1995, ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(94\)01236-9](https://doi.org/10.1016/0168-9002(94)01236-9). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0168900294012369>.
- [77] J. Lindhard *et al.*, “Integral equations governing radiation effects. (notes on atomic collisions, iii),” *Kgl. Danske Videnskab., Selskab. Mat. Fys. Medd.*, vol. 33, no. 10, Jan. 1963. [Online]. Available: <https://www.osti.gov/biblio/4701226>.
- [78] P. Sorensen and C. E. Dahl, “Nuclear recoil energy scale in liquid xenon with application to the direct detection of dark matter,” *Phys. Rev. D*, vol. 83, p. 063501, 2011. DOI: [10.1103/PhysRevD.83.063501](https://doi.org/10.1103/PhysRevD.83.063501). arXiv: [1101.6080](https://arxiv.org/abs/1101.6080) [astro-ph.IM].
- [79] J. B. Birks and F. A. Black, “Deterioration of anthracene under α -particle irradiation,” *Proceedings of the Physical Society. Section A*, vol. 64, no. 5, p. 511, May 1951. DOI: [10.1088/0370-1298/64/5/112](https://doi.org/10.1088/0370-1298/64/5/112). [Online]. Available: <https://dx.doi.org/10.1088/0370-1298/64/5/112>.
- [80] J. B. B. et al, “The theory and practice of scintillation counting,” vol. A volume in International Series of Monographs in Electronics and Instrumentation, 1964.
- [81] J. Thomas and D. A. Imel, “Recombination of electron-ion pairs in liquid argon and liquid xenon,” *Phys. Rev. A*, vol. 36, pp. 614–616, 2 Jul. 1987. DOI: [10.1103/PhysRevA.36.614](https://doi.org/10.1103/PhysRevA.36.614). [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevA.36.614>.
- [82] J. R. Angevaere, “First wimp results of xenonnt and its signal reconstruction,” Ph.D. dissertation, Universiteit van Amsterdam, 2023. [Online]. Available: <https://hdl.handle.net/11245.1/140ec608-e872-4715-8417-3b7b2660e529>.

- [83] V. Chepel and H. Araujo, “Liquid noble gas detectors for low energy particle physics,” *JINST*, vol. 8, R04001, 2013. DOI: [10.1088/1748-0221/8/04/R04001](https://doi.org/10.1088/1748-0221/8/04/R04001). arXiv: [1207.2292](https://arxiv.org/abs/1207.2292) [physics.ins-det].
- [84] XENON Collaboration, E. Aprile *et al.*, “First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment,” *Phys. Rev. Lett.*, vol. 131, no. 4, p. 041003, 2023. DOI: [10.1103/PhysRevLett.131.041003](https://doi.org/10.1103/PhysRevLett.131.041003). arXiv: [2303.14729](https://arxiv.org/abs/2303.14729) [hep-ex].
- [85] XENON Collaboration, E. Aprile *et al.*, “Search for new physics in electronic recoil data from XENONnT,” *Phys. Rev. Lett.*, vol. 129, p. 161805, 16 Oct. 2022. DOI: [10.1103/PhysRevLett.129.161805](https://doi.org/10.1103/PhysRevLett.129.161805). arXiv: [2207.11330](https://arxiv.org/abs/2207.11330). [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.129.161805>.
- [86] XENON Collaboration, E. Aprile *et al.*, “XENONnT Analysis: Signal Reconstruction, Calibration and Event Selection,” 2024, in preparation.
- [87] XENON collaboration, *XenonnT/gofevaluation: V0.1.1*, version v0.1.1, Jun. 2022. DOI: [10.5281/zenodo.6619120](https://doi.org/10.5281/zenodo.6619120). [Online]. Available: <https://doi.org/10.5281/zenodo.6619120>.
- [88] M. Szydagis *et al.*, *Noble element simulation technique*, version v2.3.12, Jan. 2023. DOI: [10.5281/zenodo.7577399](https://doi.org/10.5281/zenodo.7577399). [Online]. Available: <https://doi.org/10.5281/zenodo.7577399>.
- [89] E. Aprile *et al.*, “Scintillation response of liquid xenon to low energy nuclear recoils,” *Phys. Rev. D*, vol. 72, p. 072006, 2005. DOI: [10.1103/PhysRevD.72.072006](https://doi.org/10.1103/PhysRevD.72.072006). arXiv: [astro-ph/0503621](https://arxiv.org/abs/astro-ph/0503621).
- [90] E. Aprile *et al.*, “Simultaneous measurement of ionization and scintillation from nuclear recoils in liquid xenon as target for a dark matter experiment,” *Phys. Rev. Lett.*, vol. 97, p. 081302, 2006. DOI: [10.1103/PhysRevLett.97.081302](https://doi.org/10.1103/PhysRevLett.97.081302). arXiv: [astro-ph/0601552](https://arxiv.org/abs/astro-ph/0601552).
- [91] E. Aprile *et al.*, “New Measurement of the Relative Scintillation Efficiency of Xenon Nuclear Recoils Below 10 keV,” *Phys. Rev. C*, vol. 79, p. 045807, 2009. DOI: [10.1103/PhysRevC.79.045807](https://doi.org/10.1103/PhysRevC.79.045807). arXiv: [0810.0274](https://arxiv.org/abs/0810.0274) [astro-ph].
- [92] G. Plante *et al.*, “New Measurement of the Scintillation Efficiency of Low-Energy Nuclear Recoils in Liquid Xenon,” *Phys. Rev. C*, vol. 84, p. 045805, 2011. DOI: [10.1103/PhysRevC.84.045805](https://doi.org/10.1103/PhysRevC.84.045805). arXiv: [1104.2587](https://arxiv.org/abs/1104.2587) [nucl-ex].
- [93] P. Sorensen *et al.*, “The scintillation and ionization yield of liquid xenon for nuclear recoils,” *Nucl. Instrum. Meth. A*, vol. 601, pp. 339–346, 2009. DOI: [10.1016/j.nima.2008.12.197](https://doi.org/10.1016/j.nima.2008.12.197). arXiv: [0807.0459](https://arxiv.org/abs/0807.0459) [astro-ph].
- [94] A. Manzur, A. Curioni, L. Kastens, D. N. McKinsey, K. Ni, and T. Wongjirad, “Scintillation efficiency and ionization yield of liquid xenon for mono-energetic nuclear recoils down to 4 keV,” *Phys. Rev. C*, vol. 81, p. 025808, 2010. DOI: [10.1103/PhysRevC.81.025808](https://doi.org/10.1103/PhysRevC.81.025808). arXiv: [0909.1063](https://arxiv.org/abs/0909.1063) [physics.ins-det].
- [95] LUX, D. S. Akerib *et al.*, “Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics,” Aug. 2016. arXiv: [1608.05381](https://arxiv.org/abs/1608.05381) [physics.ins-det].
- [96] LUX, D. S. Akerib *et al.*, “Improved Dark Matter Search Sensitivity Resulting from LUX Low-Energy Nuclear Recoil Calibration,” Oct. 2022. arXiv: [2210.05859](https://arxiv.org/abs/2210.05859) [physics.ins-det].
- [97] XENON Collaboration, E. Aprile *et al.*, “Low-energy Calibration of XENON1T with an Internal ^{37}Ar Source,” *Eur. Phys. J. C*, vol. 83, p. 542, 2023. DOI: <https://doi.org/10.1140/epjc/s10052-023-11512-z>. arXiv: [2211.14191](https://arxiv.org/abs/2211.14191) [physics.ins-det].
- [98] XENON Collaboration, E. Aprile *et al.*, “Excess electronic recoil events in XENON1T,” *Phys. Rev. D*, vol. 102, no. 7, p. 072004, 2020. 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].
- [99] XENON, E. Aprile *et al.*, “Material radioassay and selection for the XENON1T dark matter experiment,” *Eur. Phys. J. C*, vol. 77, no. 12, p. 890, 2017. DOI: [10.1140/epjc/s10052-017-5329-0](https://doi.org/10.1140/epjc/s10052-017-5329-0). arXiv: [1705.01828](https://arxiv.org/abs/1705.01828) [physics.ins-det].

-
- [100] SNO, B. Aharmim *et al.*, “Combined Analysis of all Three Phases of Solar Neutrino Data from the Sudbury Neutrino Observatory,” *Phys. Rev. C*, vol. 88, p. 025 501, 2013. DOI: [10.1103/PhysRevC.88.025501](https://doi.org/10.1103/PhysRevC.88.025501). arXiv: [1109.0763](https://arxiv.org/abs/1109.0763) [nucl-ex].
- [101] K. Pearson, “LIII. On lines and planes of closest fit to systems of points in space,” *London Edinburgh Dublin Philos. Mag. & J. Sci.*, vol. 2, no. 11, p. 559, 1901. DOI: [10.1080/14786440109462720](https://doi.org/10.1080/14786440109462720).
- [102] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Power-Constrained Limits,” May 2011. arXiv: [1105.3166](https://arxiv.org/abs/1105.3166) [physics.data-an].
- [103] PandaX-4T, Y. Meng *et al.*, “Dark Matter Search Results from the PandaX-4T Commissioning Run,” *Phys. Rev. Lett.*, vol. 127, no. 26, p. 261 802, 2021. 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].
- [104] LZ, J. Aalbers *et al.*, “First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment,” *Phys. Rev. Lett.*, vol. 131, no. 4, p. 041 002, 2023. DOI: [10.1103/PhysRevLett.131.041002](https://doi.org/10.1103/PhysRevLett.131.041002). arXiv: [2207.03764](https://arxiv.org/abs/2207.03764) [hep-ex].
- [105] F. Jörg *et al.*, “Characterization of alpha and beta interactions in liquid xenon,” *Eur. Phys. J. C*, vol. 82, no. 4, p. 361, 2022. 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].
- [106] F. Jörg, “From ^{222}Rn measurements in XENONnT and HeXe to radon mitigation in future liquid xenon experiments,” Ph.D. dissertation, Ruprecht-Karls-Universität, Heidelberg, 2022.
- [107] *National Instruments (NI) LabVIEW*, <https://www.ni.com/en/shop/labview.html>, online, accessed 13-February-2024.
- [108] *PostgreSQL*, <https://www.postgresql.org/>, online, accessed 13-February-2024.
- [109] D. Cichon, “Liquid xenon detector physics with XENON1T and HeXe: Electric noise stability, background discrimination studies and measurements of the scintillation pulse shape,” Ph.D. dissertation, Ruprecht-Karls-Universität, Heidelberg, 2021.
- [110] N. Rupp, “Radon Induced Background in the XENON1T Dark Matter Search Experiment and Studies on Xenon Purity in the HeXe System,” Ph.D. dissertation, Ruprecht-Karls-Universität, Heidelberg, 2021.
- [111] M. Moszyński and others, “Temperature dependences of LaBr₃(Ce), LaCl₃(Ce) and NaI(Tl) scintillators,” *Nucl. Instr. Meth. A*, vol. 568, no. 2, pp. 739–751, 2006, ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2006.06.039>.
- [112] A. G. Singh *et al.*, “Analysis of $^{83\text{m}}\text{Kr}$ prompt scintillation signals in the PIXeY detector,” *JINST*, vol. 15, no. 01, P01023, 2020. DOI: [10.1088/1748-0221/15/01/P01023](https://doi.org/10.1088/1748-0221/15/01/P01023). arXiv: [1911.03999](https://arxiv.org/abs/1911.03999) [physics.ins-det].
- [113] L. Baudis *et al.*, “Response of liquid xenon to Compton electrons down to 1.5 keV,” *Phys. Rev. D*, vol. 87, no. 11, p. 115 015, 2013. DOI: [10.1103/PhysRevD.87.115015](https://doi.org/10.1103/PhysRevD.87.115015). arXiv: [1303.6891](https://arxiv.org/abs/1303.6891) [astro-ph.IM].
- [114] Dominick Cichon, *Personal communication*, 2022.
- [115] P. Sorensen, “Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches,” *Nucl. Instrum. Meth. A*, vol. 635, pp. 41–43, 2011. DOI: [10.1016/j.nima.2011.01.089](https://doi.org/10.1016/j.nima.2011.01.089). arXiv: [1102.2865](https://arxiv.org/abs/1102.2865) [astro-ph.IM].
- [116] National Institute of Standards and Technology (NIST), *XCOM: Photon Cross Sections Database*, <https://www.nist.gov/pml/xcom-photon-cross-sections-database>, online, accessed 2-May-2024. DOI: <https://dx.doi.org/10.18434/T48G6X>.
- [117] National Institute of Standards and Technology (NIST), *NIST Chemistry WebBook*, <https://webbook.nist.gov/chemistry/>, online, accessed 2-May-2024. DOI: <https://doi.org/10.18434/T4D303>.

- [118] O. Klein and Y. Nishina, “Über die streuung von strahlung durch freie elektronen nach der neuen relativistischen quantendynamik von dirac,” *Zeitschrift für Physik*, vol. 52, no. 11, pp. 853–868, 1929. DOI: [10.1007/BF01366453](https://doi.org/10.1007/BF01366453). [Online]. Available: <https://doi.org/10.1007/BF01366453>.
- [119] E. Aprile *et al.*, “Measurement of the Scintillation Yield of Low-Energy Electrons in Liquid Xenon,” *Phys. Rev. D*, vol. 86, p. 112004, 2012. DOI: [10.1103/PhysRevD.86.112004](https://doi.org/10.1103/PhysRevD.86.112004). arXiv: [1209.3658](https://arxiv.org/abs/1209.3658) [[astro-ph.IM](https://arxiv.org/abs/1209.3658)].
- [120] L. W. Goetzke, E. Aprile, M. Anthony, G. Plante, and M. Weber, “Measurement of light and charge yield of low-energy electronic recoils in liquid xenon,” *Phys. Rev. D*, vol. 96, no. 10, p. 103007, 2017. DOI: [10.1103/PhysRevD.96.103007](https://doi.org/10.1103/PhysRevD.96.103007). arXiv: [1611.10322](https://arxiv.org/abs/1611.10322) [[astro-ph.IM](https://arxiv.org/abs/1611.10322)].

Acknowledgements

I've been very fortunate to have the support, guidance and encouragement of many incredible people during my PhD, to whom I owe a great deal of gratitude and appreciation, and to whom I want to dedicate these last few lines.

Above all, I want to express my heartfelt gratitude to my supervisor PD Dr. Teresa Marrodán Undagoitia for her guidance and support in all matters, scientifically, professionally and personally. Thank you for always being available for discussions, for your insightful advice, and for always pushing me to do my best. I also want to thank Prof. Dr. Dr. h.c. Manfred Lindner for giving me the opportunity to work within the XENON group at MPIK for all these years, which allowed me to gain knowledge and experiences that have profoundly influenced my professional and personal development. Many thanks also to Prof. Dr. Jim Hinton for kindly taking on the role of second referee for this thesis.

To all my colleagues at MPIK – Sophie Armbruster, Kai Böse, Matteo Guida, Robert Hammann, Florian Jörg, Ying-Ting Lin, and Giovanni Volta – I am very grateful for the camaraderie and friendship we've had over the years, that have provided a very pleasant working environment and made my PhD time truly enjoyable. Many thanks in particular to my fellow HeXe team members, Dominick Cichon, Robert Hammann, Florian Jörg, and Mona Piottr, without whom my measurements would not have been possible. Thank you for taking on the challenge of taking months-long measurement runs with me, including the occasional pump-explosion-related night shift. I also want to thank the past HeXe team members that largely built and developed the setup: Dominick Cichon, Guillaume Eurin, and Natascha Rupp, as well as Teresa again for leading the HeXe team so expertly. Additionally, I am extremely grateful to Robert Hammann, Florian Jörg, and Giovanni Volta for their very helpful feedback and comments on draft versions of this thesis, as well as for the emotional support during the final stretch of writing. Thank you also to all the past members of the XENON group at MPIK from whom I have learned a great deal over the years – Stefan Brünner, Constanze Hasterok, Sebastian Lindemann, Andrii Terliuk, Joaquim Palacio, Veronica Pizzella, and Tim Wolf. Especially to Tim Wolf, thank you for encouraging and pushing me to take on responsibilities and tasks that I probably would not have taken on otherwise, but that have been critical to my personal growth.

Beyond the local XENON group, I am also very grateful to all my colleagues and friends in the XENON collaboration, with whom I have had the pleasure to work and collaborate over the years. In particular, I want to thank my fellow PMT array assembly crew members Adam Brown, Jörn Mahlstedt, and Giovanni Volta for making the challenging times during the construction phases at LNGS much more enjoyable.

Finally, I want to thank my family. To my parents, Christine and Rainer, thank you for your steadfast support throughout all of my life, for always believing in me, and for giving me the opportunity to pursue my goals in every way possible. To my sister, thank you for your love and encouragement, and for always being there for me with a ready-made plate of pasta.

Without the support and encouragement of everyone mentioned here, as well as many more, this thesis would not have been possible. Thank you all for everything.