Search for the $2\nu\beta\beta$ decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba with XENON1T

master thesis

Henning Schulze Eißing

Westfälische Wilhelms-Universität Münster
Institut für Kernphysik
AG Prof. Dr. C. Weinheimer

First Referee: Prof. Dr. C. Weinheimer
Second Referee: Prof. Dr. A. Andronic

Münster, April 2020

Declaration of Academic Integrity

I hereby confirm that this thesis on **Search for the $2\nu\beta\beta$ decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba with XENON1T** is solely my own work and that I have used no sources or aids other than the ones stated. All passages in my thesis for which other sources, including electronic media, have been used, be it direct quotes or content references, have been acknowledged as such and the sources cited.

Münster, 24. April 2020 ______________________________

I agree to have my thesis checked in order to rule out potential similarities with other works and to have my thesis stored in a database for this purpose.

Münster, 24. April 2020 ______________________________

Contents

1 Introduction

In the past decade, several experiments aimed at the direct detection of Dark Matter. XENON1T is one of these experiments, setting world leading limits on the scattering cross-section of weakly interacting massive particles (WIMPs) with ordinary matter. A low rate of background events is a requirement for these experiments due to the low cross-sections and, thus, low expected event rates. The detectors are specially designed to reduce and mitigate background events as much as possible.

Besides the search for Dark Matter the low background rate allows searches for other rare processes. In 2019 the first observation of the two-neutrino double-electron capture of ^{124}Xe was made with the XENON1T experiment. The measured half-life is the longest half-life directly measured until today. Recent developments in the signal reconstruction make it possible to also search for rare events at higher energies than the keV range associated with the Dark Matter searches. One of the processes of interest at energies of $\mathcal{O}(1\,\mathrm{MeV})$ is the double β-decay of ^{136}Xe. While the $2\nu\beta\beta$-decay of ^{136}Xe to the ^{136}Ba ground state has already been measured directly, no observation of a decay to an excited state has been made yet. This work will focus on the search for the $2\nu\beta\beta$-decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba. An observation would allow a validation of models used to calculate nuclear matrix elements. In turn, this would benefit the search for new physics in terms of the hypothetical lepton number violating neutrinoless double β-decay. The current best limit on the $2\nu\beta\beta$-decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba was set by the EXO-200 collaboration with $T_{1/2} > 6.9 \times 10^{23}\,\mathrm{yr}$ at 90 % C.L. In this work a machine learning discriminator will be developed and used to search for the decay with the XENON1T experiment. This is motivated by the unique signature of the decay which is characterized by the coincidence of two β-electrons and two γ-rays.

Chapter two outlines the theory of double β-decays and their signatures in the XENON1T experiment. General features of double β-decays are described followed by an explanation of the XENON1T experiment and the detection principle. An outline of the analysis is given.

Chapter three is focused on the simulation and data preparation tools as well as on calibration sources. After a description of the XENON1T simulation chain, the data processing tools and event selections are described. An overview of the datasets used is given followed by an explanation of the calibration sources ^{60}Co, $^{129\mathrm{m}}$Xe, $^{131\mathrm{m}}$Xe and ^{212}Pb.

Chapter four describes the development of a machine learning discriminator. First, general machine learning features and techniques are described followed by an overview of the input parameters based on an analysis by the EXO-200 collaboration. These are then defined for XENON1T in order to allow a good discrimination of signal against background events. Two different machine learning discriminators are developed: a multi-layer perceptron and a boosted decision tree. The performance of the discriminators is evaluated and cross-checked on calibration data.

In chapter five, the boosted decision tree discriminator is applied to measured data and a lower limit on the half-life of the decay is set. The last chapter summarizes the work and an outlook is given.

2 Double β-Decay in XENON1T

This first chapter is dedicated to double β-decays in the XENON1T experiment. First a general overview about double β-decays is given followed by a description of the XENON1T experiment and the used detector principle. This is followed by a brief outlook on the analysis carried out in this work.

2.1 Double β-Decay

Double β-decay is the simultaneous occurrence of two β-decays in a nucleus. It was first proposed by Maria Goeppert-Mayer in 1935 [1] only shortly after Enrico Fermi published his theory of β-decay [2]. First indirect measurements of double β-decay were made in the 1960's with geo-chemical methods and the first direct observation was made in 1987 for ^{82}Se [3].

The energetic feasibility of double β-decay is illustrated using the liquid droplet model for the atomic nucleus. The binding energy E_B in dependence on the mass number A and the atomic number Z can be calculated by the semi empirical formula

$$E_B(A,Z) = a_v A - a_s A^{2/3} - a_c \frac{Z^2}{A^{1/3}} - a_a \frac{(N-Z)^2}{4A} + \begin{cases} +\frac{11.2\,\mathrm{MeV}}{\sqrt{A}} & \text{for } Z, N \text{ even} \\ 0 & \text{for } Z \text{ or } N \text{ odd} \\ -\frac{11.2\,\mathrm{MeV}}{\sqrt{A}} & \text{for } Z, N \text{ odd} \end{cases} \tag{2.1}$$

which is also known as Bethe-Weizsäcker formula [4]. Here a_v is called *volume term*, a_s is the *surface term*, a_c is the *coulomb term*, and a_a is the *asymmetry term*. The last term of the equation is called *pairing term*. N is the number of neutrons in the nucleus. In case of a (double) β-decay the mass number A stays constant, so that $E_B(A = \text{const.}, Z) \propto -Z^2$ and thus $M(A = \text{const.}, Z) \propto Z^2$ where $M(A, Z)$ is the mass of the nucleus. Figure 2.1 shows the mass of the atomic nucleus as a function of the atomic number for even/odd, odd/even and even/even, odd/odd combinations of Z and N. The pairing term creates an offset between the binding energies for even/even and odd/odd combinations.

Due to energy conservation, a single β-decay of a nucleus can only occur if a neighboring isobar has a lower mass. In case of even/even nuclei both neighboring isobars can have higher masses due to the paring term in the mass formula and a β-decay is

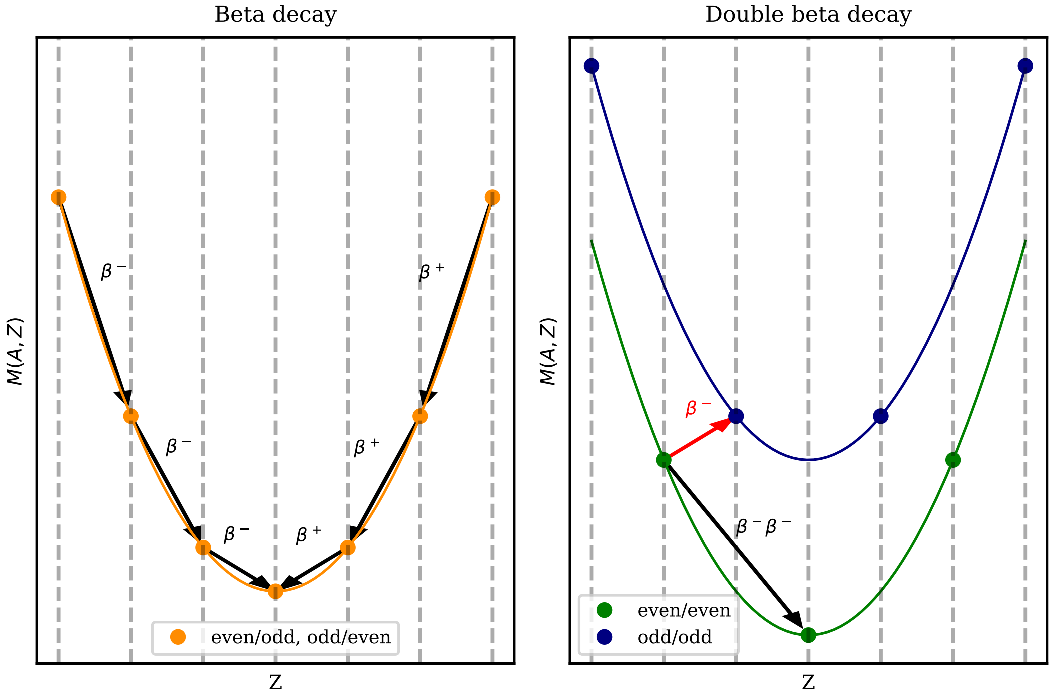


Figure 2.1: Mass of the nucleus $M(A, Z)$ plotted against the atomic number Z for constant A for even/odd, odd/even combinations of Z and N on the left panel drawn in orange and even/even (green), odd/odd (blue) combinations on the right panel of the plot. Allowed decays are indicated as black arrows, a forbidden decay is shown as red arrow. The masses for even/even and odd/odd combinations are separated due to the pairing term in equation (2.1). For some even/even nuclei a single β-decay is forbidden by energy conservation and the only possible decay channel is double β-decay.

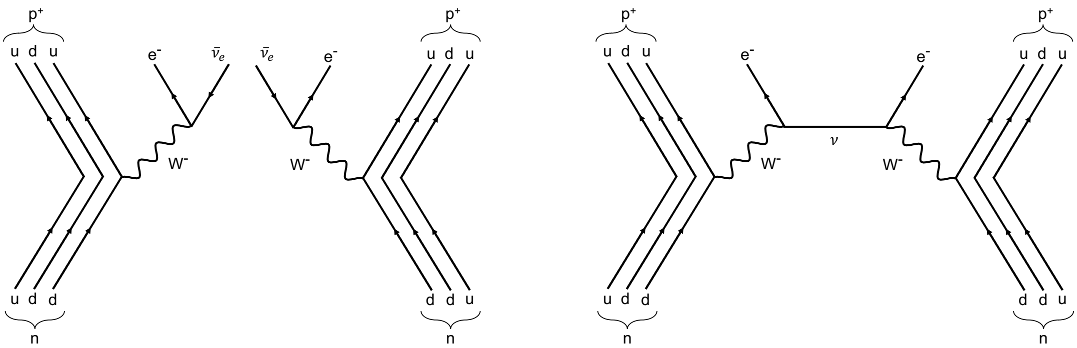
forbidden even if the nucleus is not the one with the maximal binding energy. In this case only a double β-decay can occur [5]. The double β-decay is not detectable when the single β-decay is not forbidden or heavily suppressed due to the large amount of single β-decays obscuring the double β-decay signal [6].

Beside the ordinary double β-decay ($2\nu\beta\beta$) where two β-particles (electrons or positrons) and neutrinos are emitted, a hypothetical lepton number violating neutrinoless double β-decay ($0\nu\beta\beta$) could occur if neutrinos are their own antiparticles, called Majorana particles. Additional a non-zero neutrino mass[1] is required [6].

A double β-decay on the neutron-rich side of the table of nuclides is a double β^--decay either in the process allowed in the Standard Model [5]:

$$2\nu\beta^-\beta^- : (A, Z) \to (A, Z+2) + 2e^- + 2\bar{\nu}_e, \tag{2.2}$$

[1]The fact that neutrinos have a mass was shown by neutrino oscillation experiments [7].



(a) 2νβ⁻β⁻ Decay (b) 0νβ⁻β⁻ Decay

Figure 2.2: Feynman diagrams of the ordinary $2\nu\beta^-\beta^-$-decay (a) and the $0\nu\beta^-\beta^-$-decay (b). In both cases two neutrons turn into protons via emission of a W^- boson. In the ordinary case the bosons decay and two electrons and two electron anti-neutrinos are emitted. If neutrinos are their own anti-particles a decay can happen without emission of neutrinos.

or the hypothetical lepton number violating process [8]:

$$0\nu\beta^-\beta^- : (A, Z) \rightarrow (A, Z+2) + 2e^-. \tag{2.3}$$

On the proton rich side, positrons can be emitted ($\beta^+\beta^+$) and/or electron captures (ECEC, ECβ^+) can occur [6].

The Feynman graphs of the ordinary and the neutrinoless double β^--decay are given in figure 2.2. In both cases two down-quarks convert into up-quarks and virtual W bosons. In case of the $2\nu\beta^-\beta^-$ decay, the W bosons convert to electrons and electron-anti-neutrinos that are emitted in the decay process. In case of the $0\nu\beta^-\beta^-$-decay, the right-handed electron-anti-neutrino originating from one of the W boson decay vertex is absorbed as left-handed electron-neutrino at the other W boson vertex [8]. Thus this process is possible if the neutrino is a Majorana particle and has mass to cause the change in helicity. Only electrons are emitted in the neutrinoless decay with their summed energy corresponding to the Q-value of the decay. While the $2\nu\beta\beta$-decay was observed by several experiments for different isotopes with half-lifes ranging from 1×10^{18} yr to 1×10^{21} yr [7], no observation of the neutrinoless decay has been made yet [8]. Currently several experiments are searching for the neutrinoless decay.

The rate of a $2\nu\beta\beta$-decay $\Gamma^{2\nu}$ can be calculated based on Fermi's golden rule [9][10]. By factorizing the kinematic part containing the phase space of the emitted leptons $G^{2\nu}(Q_{\beta\beta}, Z)$ in dependence on the decay's Q-value $Q_{\beta\beta}$ and the nuclear part giving the transition probability between two nuclear states one obtains [5]

$$\Gamma^{2\nu} = \frac{1}{T^{2\nu}_{1/2}} = G^{2\nu}(Q_{\beta\beta}, Z)|M^{2\nu}|^2. \tag{2.4}$$

Here $M^{2\nu}$ is the nuclear matrix element (NME) of the transition and $T^{2\nu}_{1/2}$ is the half-life of the decay. While phase space factors can be calculated precisely, the calculation of NMEs is challenging [5]. The decay rate of the neutrinoless decay can be calculated using [5]:

$$\Gamma^{0\nu} = \frac{1}{T^{0\nu}_{1/2}} = G^{0\nu}(Q_{\beta\beta}, Z)|M^{0\nu}|^2\langle\eta\rangle^2. \tag{2.5}$$

Here $\langle\eta\rangle^2$ is a lepton number violating parameter representing physics beyond the Standard Model. In order to extract new physics from equation (2.5) in case of an observation, it is important to improve the NME calculations. Even though there is no direct relation between $M^{0\nu}$ and $M^{2\nu}$, the measurement of the $2\nu\beta\beta$-decay allows a validation of the nuclear models. Thus is can benefit the search for new physics [5].

2.2 Excited State Decay of ^{136}Xe

One of the isotopes undergoing a double β^--decay is ^{136}Xe. Its $2\nu\beta\beta$ decay was first measured by the EXO-200 collaboration [11] using a liquid xenon time projection chamber (TPC) with a half-life of $T_{1/2} = (2.165 \pm 0.016(\text{sys}) \pm 0.059(\text{stat}))$ yr [12].

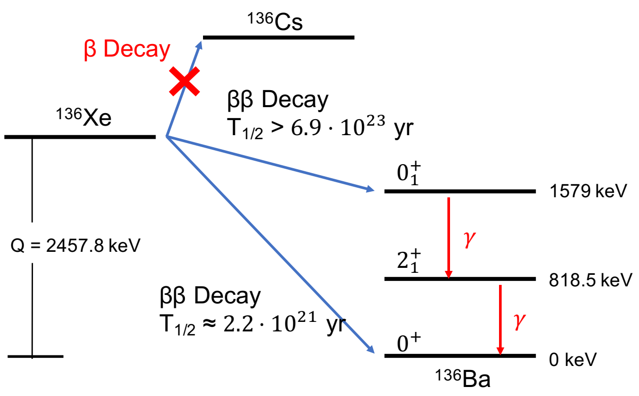


Figure 2.3: Level scheme of the double β-decay of ^{136}Xe to the 0^+ ground and 0_1^+ state of ^{136}Ba. Energies and J^P taken from [13]. Half-life of the ground state decay is taken from [12] and the half-life limit of the excited state decay from [14].

Beside decays to the ground state of ^{136}Ba, the high Q-value of $Q_{\beta\beta} = 2457.83$ keV allows decays to excited states of ^{136}Ba. The decay rate of such a decay to an excited state is substantially suppressed compared to decays to the ground state, due to the smaller transition energies [5]. The corresponding level scheme is given in figure 2.3. This work will focus on decays to the 0_1^+ state of ^{136}Ba.

The state transitions to the ground state via emission of two γ-rays of $E_{\gamma,1} = 760.5$ keV and $E_{\gamma,2} = 818.5$ keV [13]. A graphical illustration of the decay is given in figure 2.4 and the shorthand notation $2\nu\beta\beta^*$ will be used for this excited state decay. Due to the $0_1^+ \to 2_1^+ \to 0^+$ spin structure of the involved nuclear states, the emission angles of the γ-rays are correlated [15]. The angular distribution of the emitted γ-rays is given by

$$W(\theta) = \frac{5}{8} \cdot (1 - 3\cos^2(\theta) + 4\cos^4(\theta)). \tag{2.6}$$

Here θ is the angle between the two γ-rays which are preferredly to be emitted either back to back or in the same direction. The decay is yet unobserved, but searches were performed by the EXO-200 [14] and KamLAND-Zen [16] collaborations. A lower limit on the half-life was set by EXO-200 to $T_{1/2,\mathrm{L}} > 6.9 \times 10^{23}\,\mathrm{yr}$ at 90 % C.L.. This is in agreement with the estimated half-life from theory of $T_{1/2,\mathrm{E}} = 2.5 \times 10^{25}\,\mathrm{yr}$ [14].

The number of expected decays N in a liquid xenon detector can be calculated by

$$N = \frac{t \cdot m \cdot \ln(2)}{m_u \cdot T_{1/2}} \cdot \eta. \tag{2.7}$$

Here t is the observation time and m is the xenon target mass. The atomic mass of xenon is $m_u = 131.293\,\mathrm{u}$ [17] and the abundance of ^{136}Xe in the xenon used by the XENON1T experiment is $\eta = 8.49\,\%$ [18]. In $m = 1\,\mathrm{t}$ xenon monitored over a period of $t = 263.619\,\mathrm{d}$ about $N \approx 8$ νββ*-decays will occur for a half-life of corresponding to the estimated half-life of $T_{1/2,\mathrm{E}}$. For $T_{1/2} = T_{1/2,\mathrm{L}}$, $N \approx 282$ events are expected to occur.

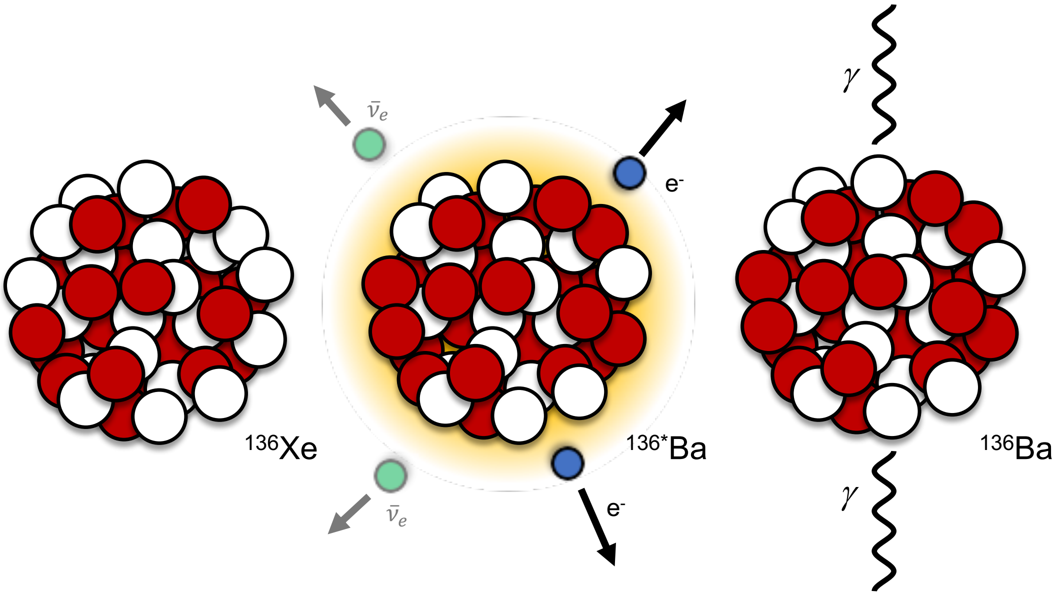


Figure 2.4: Sketch of the 2νββ-decay of ^{136}Xe to the 0_1^+ state of ^{136}Ba. Protons are depicted in red, neutrons in white, neutrinos in green and electrons in blue. The left part of the figure represents the nucleus of ^{136}Xe composed of neutrons and protons. The middle part of the figure shows the νββ*-decay to an excited state of ^{136}Ba. Two electron anti-neutrinos and two electrons are emitted. The excited state of the ^{136}Ba nucleus is indicated in orange. The right part of the figure shows the deexitation to the ground state via emission of two γ-rays, indicated as curved lines. The two neutrinos will not be detectable while the electrons and γ-rays produce detectable energy depositions in a detector.

Even thought the number of expected events in the XENON1T exposure is small, a search can be performed exploiting the unique event signature. As shown in figure 2.4 two electrons are emitted along with two γ-rays. While the electrons will deposit their energy close to the decay position, the two angular correlated γ-rays can travel larger distances is a detector. To measure this decay a detector with a good energy resolution and position reconstruction is required. The energy depositions can than be separately measured and the events identified due to the unique event signature. A detector concept that fulfills these requirements is a xenon dual phase TPC. The detection of a double β-decay to excited states would yield important information for the calculation of NMEs and thus allow a complementary cross-check of the involved models [5].

2.3 Dual-Phase Time Projection Chamber

The XENON1T experiment [19] used a dual phase TPC filled with liquid xenon (LXe) and a gaseous xenon (GXe) layer on top. It searched for interactions of WIMPs with xenon nuclei [20]. The walls of the detector are made of highly reflective polytetrafluorethylen (PTFE) walls and the top and bottom are equipped with photomultiplier tubes (PMTs) [19]. The detector principle is graphically shown in figure 2.5.

When a particle scatters in the detector scintillation photons of 178 nm [22] are produced by two processes [23] given in equation (2.8) and equation (2.9).

$$\begin{aligned} \mathrm{Xe}^* + \mathrm{Xe} + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^* + \mathrm{Xe} \\ \mathrm{Xe}_2^* &\rightarrow 2\mathrm{Xe} + h\nu \end{aligned} \tag{2.8}$$

The first process starts with an excited xenon atom forming an excited xenon dimer. The excited dimer can relax to the ground state by emission of a photon. Since the scintillation light is emitted by a dimer, xenon is transparent to its own scintillation light [23].

$$\begin{aligned} \mathrm{Xe}^+ + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^+ \\ \mathrm{Xe}_2^+ + \mathrm{e}^- &\rightarrow \mathrm{Xe}^{**} + \mathrm{Xe} \\ \mathrm{Xe}^{**} &\rightarrow \mathrm{Xe}^* + \mathrm{heat} \\ \mathrm{Xe}^* + \mathrm{Xe} + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^* + \mathrm{Xe} \\ \mathrm{Xe}_2^* &\rightarrow 2\mathrm{Xe} + h\nu \end{aligned} \tag{2.9}$$

The second process starts with the ionization of a xenon atom and the formation of an ionized xenon dimer. This dimer can than recombine with an electron creating a highly excited xenon atom. Via collisions with surrounding xenon atoms the state relaxes to the single excited state and again the formation of an excited dimer is possible. Scintillation photons are then created via deexitation as shown in equation (2.8). This direct scintillation light is detected by the PMTs and is called S1 [23].

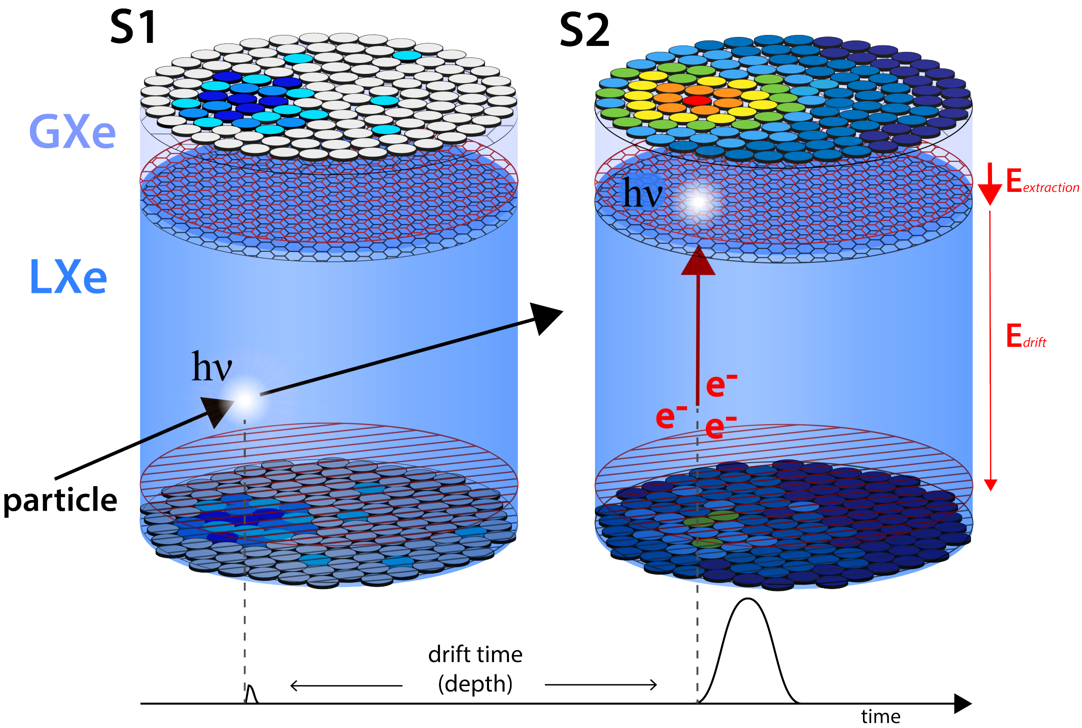


Figure 2.5: Schematic representation of the detector principle of a dual phase xenon TPC. The LXe volume is shown as dark-blue cylinder, the GXe phase in light-blue on top. The left side of the figure illustrates the scattering of a particle in the LXe and the creation of direct scintillation light (S1). The light is detected by PMT arrays on the top and bottom of the detector, here shown as colored circles. The colors used indicate the amount of detected light. The right side of the figure shows the creation and drift of electrons to the liquid gas interface. The electrons are drifted and extracted from the liquid by electric fields. The fields are created by meshes, also shown in the figure. The extracted electrons create secondary scintillation light by electroluminescence (S2) which is also detected by the PMTs. The light distribution on the top PMT array can be used to reconstruct the xy-position of the interaction and the time between the S1 and S2 signal. Image taken from [21].

An electric field E_{drift} created by a cathode on the bottom of the detector and a gate mesh at the liquid-gas interface suppresses the recombination of electrons and ionized xenon dimers and drifts the electrons towards the gas phase. A second electric field $E_{extraction}$ between the gate mesh and anode grid in the gas phase accelerates the electrons into the GXe, creating secondary proportional scintillation light by electroluminescence. The signal is again detected by the PMTs and called S2 [23]. The xy-position of the interaction can be reconstructed from the S2 top PMT hit pattern [24]. The depth of the interaction, the z-position, can be calculated using the time difference between the prompt S1 and the delayed S2 giving this detector a 3D reconstruction of the interaction position [24].The combination of an S1 with at least one S2 signal is called event.

The dual signal allows an independent measurement of charge and light of an interaction to distinguish different recoil types [24]. Electrons and γ-rays interact with the atomic shell of xenon. Such an interaction is called an electronic recoil (ER). Neutrons and WIMPs on the other hand interact with the nucleus, which is called a nuclear recoil (NR). Since light and charge yields, the amount of created photons and electrons per unit energy, differ for ER and NR one can use

$$\left(\frac{\mathrm{S2}}{\mathrm{S1}}\right)_{\mathrm{NR}} < \left(\frac{\mathrm{S2}}{\mathrm{S1}}\right)_{\mathrm{ER}}$$

to reduce the background from ER events in the dark matter search [24].

Xenon is used as detector material due to its high mass number of $\mathrm{A} = 131$ [17]. The spin independent cross-section for interactions of WIMPs with nuclei is proportional to A^2 and a large A will increase the number of WIMP-nucleon scatters. Furthermore, xenon has a high atomic number of $\mathrm{Z} = 54$ [13] which leads to a high stopping power [25] for low energetic γ-rays as possible background source. This self-shielding of xenon can be used for a so-called fiducialization using only the innermost part of the detector volume for the physics search [24]. This fiducial volume has a greatly reduced overall event rate since radiation of most external background sources are stopped in the outer detector region.

2.4 XENON1T

The XENON1T Dark Matter experiment [19] was located at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy shielded by 3000 m water equivalent rock overburden to reduce the flux of cosmic radiation by multiple orders of magnitude. The experiment used in total 3.2 t of xenon from which 2 t are contained in an cylindrical dual phase TPC of 97 cm height between the cathode and gate mesh and 96 cm diameter. The remaining 1.2 t xenon were used as passive shielding around the TPC. The TPC used an drift field of $82\,{}^{\mathrm{V}}/_{\mathrm{cm}}$ and an extraction field of $> 10\,{}^{\mathrm{kV}}/_{\mathrm{cm}}$. To make the field as homogeneous as possible, 74 copper field shaping rings were installed around the detector. In total 248 PMTs with high quantum efficiency and low intrinsic radioactivity were used. The bottom array was equipped with 121 hexagonal closely packed PMTs for a high light collection while the top array was made out of 127 PMTs arranged in a radial pattern [19]. An illustration of the detector is given in section 2.4 along with a picture of the detector assembly in section 2.4.

The whole detector is housed in a double walled cryostat which itself is located in a water tank filled with 700 t of water. The tank is equipped with 84 PMTs to act as passive shield against external radiation and active water Cherenkov muon veto. A picture of the setup can be found in figure 2.7. Next to the water tank a service

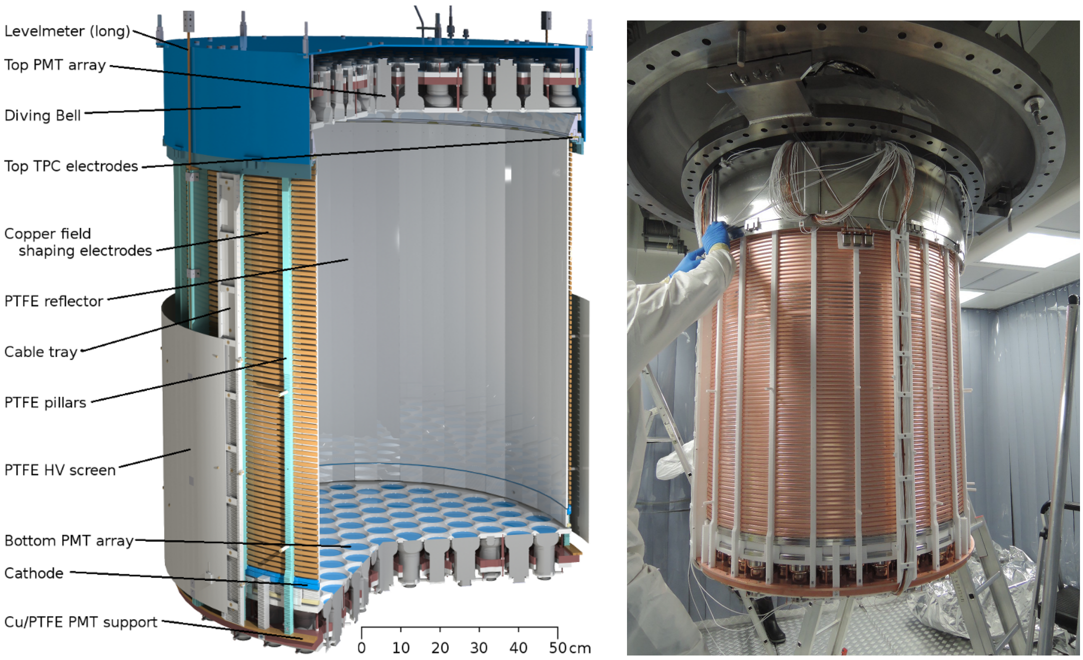


Figure 2.6: Illustration of the XENON1T TPC shown on the left side and a picture of the detector assembly on the right side. Pictures taken from [19] and [26].

building is located and houses several subsystems to run the detector. A description of all subsystems can be found in [19].

Figure 2.7: Picture of the XENON1T experiment. The TPC is located in a cryostat in the water tank on the left side of the image. The service building containing the subsystems is located on the right side of the image. (Image credit: Roberto Corrieri and Patrick De Perio.)

2.5 Analysis Outline

As outlined in section 2.2 only 8 events of the $2\nu\beta\beta$-decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba are expected in the XENON1T data that will be analyzed in this work. Furthermore, the broad energy range of the events make a classical cut-based analysis difficult. In this work a different analysis method will be used to search for the signal events: A machine learning (ML) discriminator trained on the classification of events into signal and background. An graphical overview of the main analysis steps is given in figure 2.8.

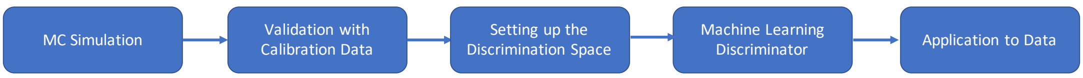


Figure 2.8: Graphical representation of the main analysis steps in this work.

In this work two kinds of data are used: Monte Carlo (MC) simulations and the measured XENON1T data. The working principle of the detector was outlined in the last two sections and the simulations are described in section 3.1. The discriminator is trained on simulated signal and measured background events, thus accurate MC simulations are necessary. Two types of events are simulated: events from calibration sources and signal events of the $2\nu\beta\beta^*$-decay of ^{136}Xe. The measured events from the detector are grouped into calibration data and background data. A range of cuts is applied both to simulated and measured events. An overview of the cuts used is given in section 3.2.2. A graphical overview of the data flow in this work is given in figure 2.9.

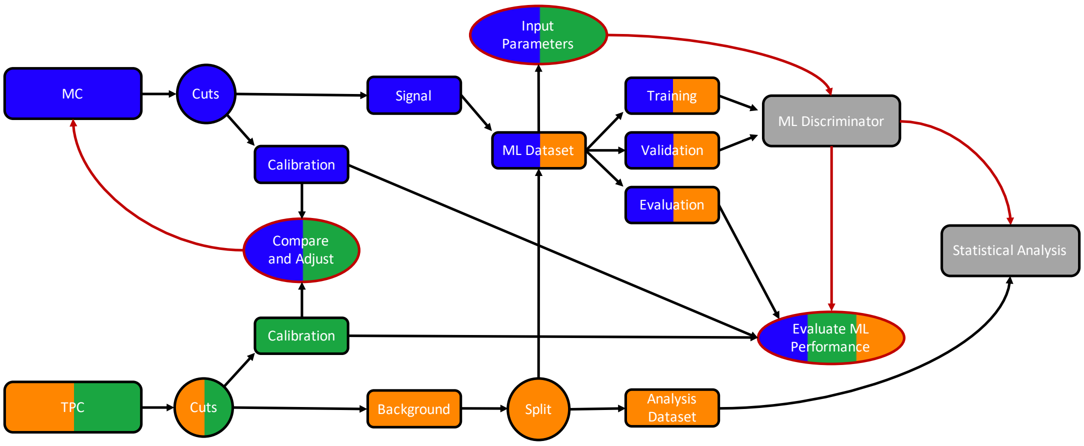


Figure 2.9: Detailed overview of the datasets and performed analyses in this work. Simulated data is shown in blue, measured calibration data in green and measured background data in orange. A dataset containing more than one datatype is indicated as box with more than one color. An analysis is drawn as ellipse and the application of cuts and splitting of datasets as circles. The connection of the datasets is shown as black arrows. The output an an analysis or the application of the discriminator to data is shown as red arrows.

In a first analysis step the simulated and measured calibration data are compared and used to adjust the MC simulations, so that a sufficient agreement between simulated and measured events can be achieved. The calibration sources used are outlined in section 3.3 and the comparison of simulated with measured events is given in section 4.3 for a range of different event parameters. Methods are developed for the charge and light generation in the simulations as outlined in section 3.1.3.

The so-called background data, are events recorded when no calibration source was deployed and the detector was operated under stable conditions. In total a background exposure of 246.7 d is used in this work which corresponds to the exposure[2] of science run (SR) 1 used for the dark matter search of XENON1T in [20]. The background dataset is split in two parts where one third of the exposure is used in combination with simulated signal events for the development of an ML discriminator. The remaining two thirds of the background exposure are used to search for the $\nu\beta\beta^*$-decays using the developed discriminator. Due to the high computational costs, simulations of the XENON1T background spectrum in the energy range of interest were not possible in this work.

[2]The exposure reduction by cuts is already applied.

Based on simulated signal events compared with measured background events, event parameters are chosen as input for the ML discriminator that allow a good separation of signal and background events. These parameters form the discrimination space. The ML dataset consisting of simulated signal and measured background events is further split into three parts for the training, validation and evaluation of the models. An overview of the livetimes of the different datasets is given in table 2.1 and a description on the usage of each dataset is given in section 4.1.2.

Table 2.1: Runtime overview for the datasets used in this work for the development of a ML discriminator and the search for signal events.

	Total	Analysis	Training	Validation	Evaluation
Livetime [d]	246.7	165.29	40.91	20.15	20.35
Fraction	1	0.67	0.17	0.08	0.08

An additional cross check of the simulations and trained discriminator is done with the calibration data. The discriminator is applied to simulated and measured events and the output is compared in section 4.6.2. Finally the discriminator is applied to the analysis data and a limit on the half life of the decay is determined in chapter 5.

3 Simulations and Data Preparation

The training of a machine learning (ML) discriminator requires accurate simulations and a preparation of the measured data that allows to reconstruct multiple spatially separated energy depositions. This chapter will first outline the simulation tools used in this work followed by a description of the data preparation. Finally, calibration sources that are used to validate and tune the MC simulations will be introduced.

3.1 Simulations

This section outlines the XENON1T MC simulation chain used in this work. A graphical overview of the involved steps is given in figure 3.1. The goal of the simulation is to generate events that resemble the measured events in the detector in terms of data format and event parameters. The first step of the simulations chain is an event generator followed by a particle tracking and detector geometry simulation. Based on the output of this simulation the number of charge (S2) and light (S1) quanta are calculated for each resolvable energy deposition. As a final step, the waveform of the event as it would be recorded by the XENON1T data acquisition is simulated. The events are stored in the same format as measured data, so the same data processing tools can be used for simulation and measured data. The individual simulation steps are further described in the following sections.

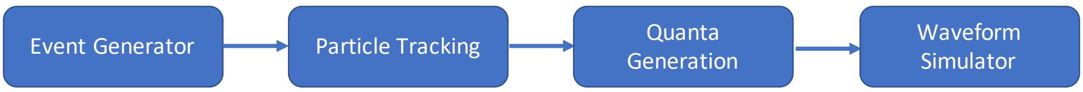


Figure 3.1: Overview of the XENON1T simulation chain. The initial particle kinematics calculated with an event generator are used as the input for the particle tracking and geometry simulations. Based on the provided energy depositions the number of photons and electrons is determined in the quanta generation. In turn, these are used as the input for the waveform simulator.

3.1.1 Event Generator

The first step of the MC chain is the event generator. It calculates the initial momentum vectors and emission times of particles in radioactive decays or other processes of interest. Two different event generators are used in this work, the built-in event generator of the simulation software Geant4 [27] using the *G4RadioactiveDecay* process and Decay0 [28].

The initial kinematics for radioactive decays of background and calibration sources in XENON1T can be calculated with the Geant4 event generator. The kinematics of $2\nu\beta\beta^*$-decay of ^{136}Xe will be generated with Decay0.

The event generator Decay0 is focused on the calculation of initial kinematics of double β decays. Other events like α or single β decays can be generated as well mostly for sources imitating double β-decays. The double β-decays of 40 isotopes can be simulated for 17 decay modes. Besides decays to the ground state, the particle kinematics of decays to excited states can be generated including subsequent deexitations. Internal conversion processes and emission of e^-e^+ pairs are taken into account. Decay0 is used since more than 20 years by multiple groups mostly searching for $0\nu\beta\beta$ decay [28].

For this work the Decay0 Fortran code is modified to include the angular correlation of the two emitted γ-rays during the deexitation of the 0_1^+ excited state of ^{136}Ba. The modified source code is given in appendix A.1.

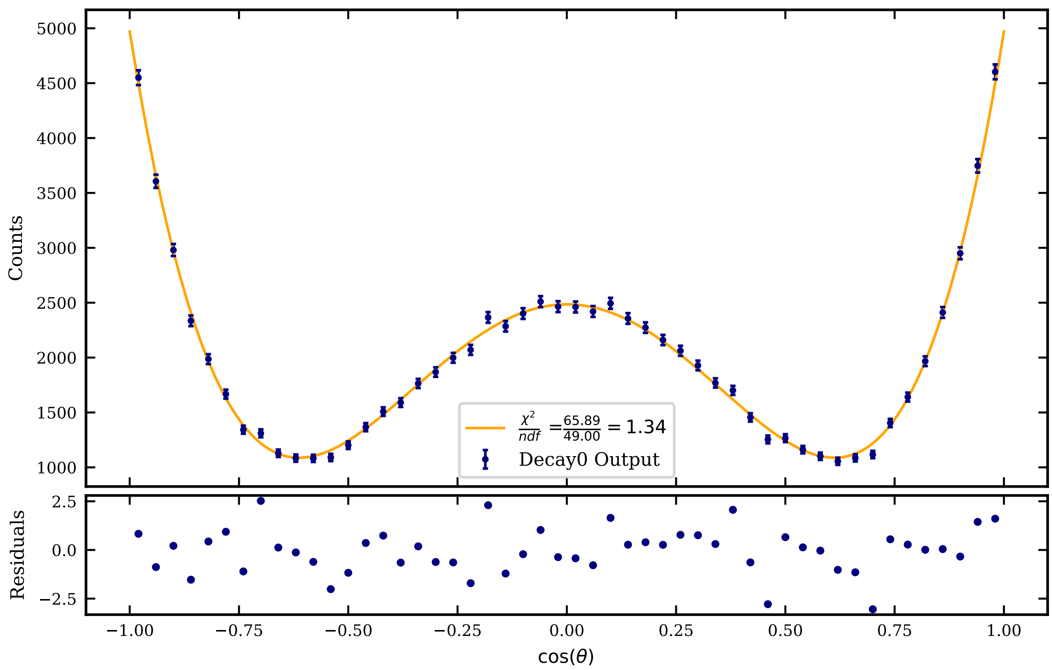


Figure 3.2: Histogram of the angle between the momentum vectors of the two deexitation γ-ray of the $2\nu\beta\beta^*$-decay of ^{136}Xe generated with Decay0 in blue and a fit using equation (2.6) in orange. The residuals are given in the bottom panel.

In order to validate if the implementation of the angular correlation produces the desired output, the angle between the initial momentum vectors of the two γ-rays is calculated for 1×10^5 ^{136}Xe $2\nu\beta\beta^*$ events generated with Decay0. A histogram is given in figure 3.2 together with a fit using the expected angular correlation given in equation (2.6). The fit is carried out as χ^2-minimization with the MIGRAD algorithm [29] via the iminuit [30] Python module. The χ^2 is defined as the sum of the squared deviations between the fit values $\bar{y}_i(x_i, \vec{p})$ and the measured value y_i normalized by the square of the uncertainty σ_i

$$\chi^2 = \sum_i \frac{(y_i - \bar{y}_i(x_i, \vec{p}))^2}{\sigma_i^2}. \tag{3.1}$$

Here, $\vec{p}$ is the parameter vector of the fit function. The residuals of the fit are defined as the deviation between the fit value and the measured value normalized to the uncertainty of each individual data point

$$R_i = \frac{y_i - \bar{y}_i(x_i, \vec{p})}{\sigma_i}. \tag{3.2}$$

One finds a good agreement between the fit and the generated data so that one can conclude that the angular correlation was implemented as intended.

3.1.2 Particle Tracking and Geometry Simulation

The initial particle kinematics generated with the event generators are used as input for the geometry and particle tracking simulations based on Geant4 [27]. Geant4 simulates the propagation of particles through the detector geometry. The detector model used for the XENON1T simulations was built from CAD drawings to resemble the build detector as close as possible. A illustration of the detector geometry is given in section 2.4. The model includes all components of sufficient mass or effects on the optical properties of the detector. Particle tracking is performed in steps, where the step size is determined based on the surrounding detector medium, the type and the energy of the tracked particle. For each step that causes an energy loss in the LXe of the detector, the deposited energy, time, position, particle type and interaction process are stored. Detailed descriptions of the detector model and particle tracking simulations can be found in [21] and [22].

A graphical representation of the output for a simulated ^{136}Xe $2\nu\beta\beta^*$-decay event is given in figure 3.3. The figure shows energy depositions in the LXe in a three dimensional plot. Each deposition is shown as circle. The size of the circles is proportional to the energy deposition and the color gives the time of the deposition measured from the first interaction. One finds energy depositions at 5 different locations, separated by a few centimeters. At each interaction site, multiple energy depositions take place. The timescale involved is well below one nanosecond. As a result, the S1 signal of all energy depositions will be merged while it es likely that multiple S2 signals can be detected.

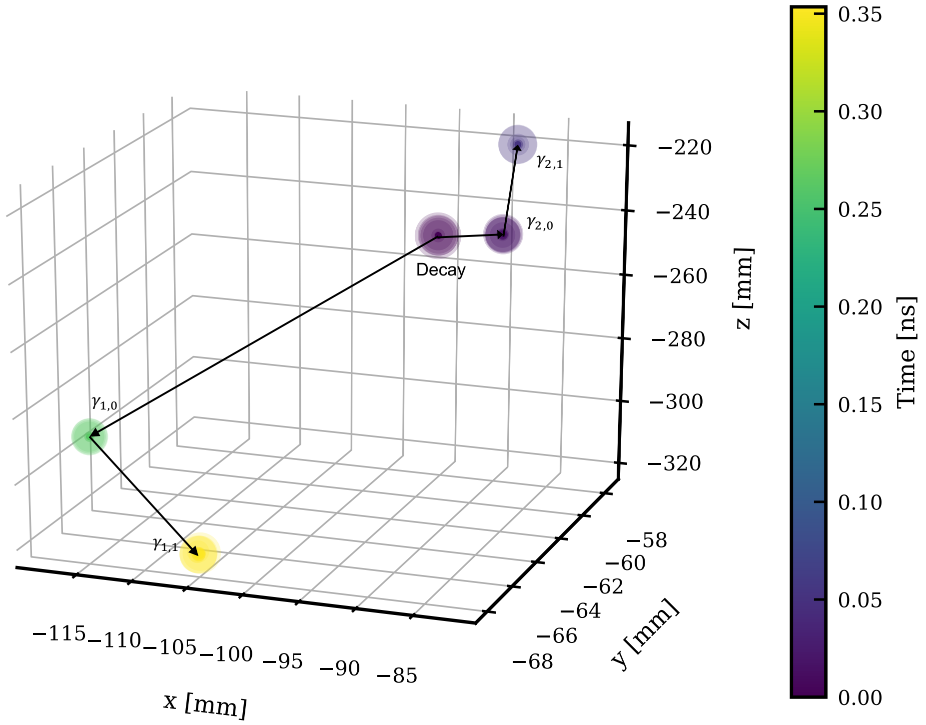


Figure 3.3: Three dimensional representation of energy depositions in the LXe volume of the XENON1T TPC simulated with Geant4. The initial particle kinematics are generated using Decay0, simulating a ^{136}Xe $2\nu\beta\beta^*$-decay. The size of the circles corresponds to the energy deposition and the color indicates the time of the interaction. Arrows are drawn between the interaction positions of the γ-rays.

3.1.3 Quanta Generation

Following the particle tracking simulations the number light and charge quanta are calculated with a software developed in this thesis. The source code is given in appendix A.2. This interface connects the particle tracking simulations with the newly designed waveform simulator WFSim. For each interaction in an event the *position*, *time*, *energy deposition*, *Particle type*, *Parent-type*, *Track ID* and *Parent ID* is used as input. A description of each parameter can be found in table 3.1.

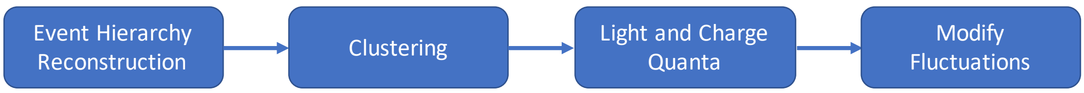


Figure 3.4: Overview of the quanta generation steps. The output of the particle tracking simulations is used to reconstruct the causal particle hierarchy in one event. A clustering is performed to reduce the computational costs and accurately implement the micro-physics processes. The number of light and charge quanta is calculated using NEST and the fluctuations are modified to better represent measured data.

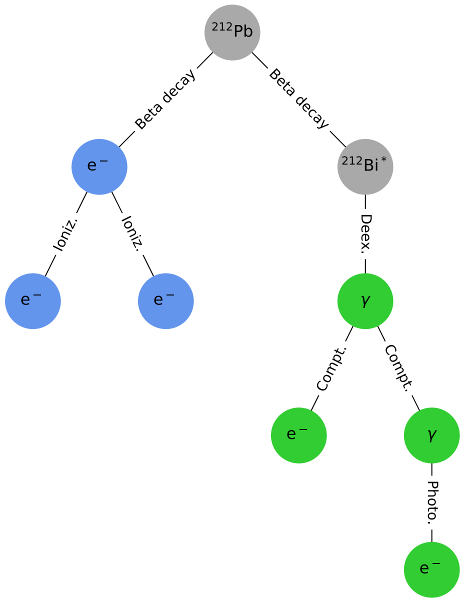
Table 3.1: Description of the input parameter for the quanta calculation. Each parameter is provided by the particle tracking simulation for each interaction in an event.

Parameter	Description
Position	Three dimensional position of an interaction site.
Time	Time of the interaction measured from the generation of the first particle.
Energy deposition	Energy deposition at the interaction site measured in keV.
Particle type	Type of the particle, e.g. electrons or γ-particles, creating the energy deposition.
Parent type	If new particles are created in interactions, the parent type specifies the particle type of the parent particle creating the interaction.
Track ID	A unique number identifying a particle in an event.
Parent ID	The track ID of the parent particle.

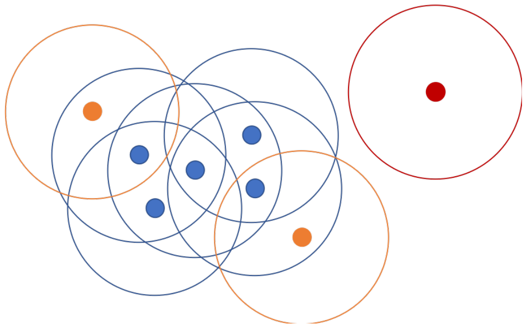
The number of charge and light quanta are determined in a multi step process. A graphical overview is given in figure 3.4. In the first step the causal particle hierarchy of an event is reconstructed using the *Parent* and *Track ID*. This hierarchy lists the causal relations in an event from the primary particles creating the event down to all particles created in interactions and decays. An additional variable is assigned for each interaction indicating the primary particle type. Only electrons, γ-rays, α-particles and neutrons are taken as primary particles. It was found that a better agreement between simulations and calibration data can be achieved when the particle hierarchy is used to determine which model is used in the calculation of charge and light quanta using NEST.

An illustration of an event hierarchy can be found in figure 3.5a. The ^{212}Pb nucleus decays via β^--decay to an excited state of ^{212}Bi, as described in section 3.3.2, leading to the emission of an β^--electron and an deexitation γ-ray. For each interaction causally connected to the β^--electron, the primary particle type is set to "electron" indicated in blue. For each interaction connected to the deexitation γ-ray, the primary particle type is set to "gamma" indicated in green. The particle hierarchy reconstruction is validated only for the β^--decays of ^{212}Pb and ^{60}Co and the deexitation of ^{131m}Xe and ^{129m}Xe. The hierarchy of particles in events generated with Decay0 can also be reconstructed.

Following the hierarchy reconstruction all interactions without energy depositions are removed. The second step of the quanta generation is a temporal and spatial clustering of energy depositions. As shown in figure 3.3, energy depositions can occur at multiple interaction sites. Each site consists of multiple smaller energy depositions. The usage of a clustering will reduce the computational costs of the simulations and account for micro-physics processes by energy depositions in small distances to each other. The clustering is performed using a DBSCAN algorithm [31] included in the scikit-learn



(a) Schematic representation of a simplified ^{212}Pb decay particle hierarchy. Nuclei are marked in gray, particles created by the β^--electron in blue and particles created in interactions caused by the γ-ray in green.



(b) Illustration of a DBSCAN clustering. The neighborhood radius is indicated as circle around the cluster position. *Core points* are drawn in blue, *reachable points* in orange and *noise* in red.

Figure 3.5: Particle relation of a ^{212}Pb decay (a) and graphical illustration of the DBSCAN algorithm.

python library [32]. DBSCAN is a simple but capable algorithm to group one- or multidimensional data into clusters. The number of clusters is determined by the algorithm so it must not be specified at the start unlike other clustering algorithms. The basic principle is illustrated in figure 3.5b. The algorithm is controlled by two parameters, the neighborhood radius ϵ and the minimum number of points to form a cluster $minPTs$. The algorithm classifies points in three categories. So-called *core points* have at least $minPTs - 1$ neighbors within the distance ϵ. *Reachable points* are within the distance ϵ of at least one core point. The third category, called *noise*, are points without connection to core points. All connected core and reachable points form a cluster.

The temporal clustering is performed with $minPTs = 2$ and $\epsilon = 10\,\mathrm{ns}$[1]. Following the temporal clustering a spatial clustering in three dimensions is performed for each of the time clusters. Again DBSCAN is used with $minPTs = 2$ and $\epsilon = 5\,\mu\mathrm{m}$ which is slightly larger as the mean ionization electron-ion thermalization distance [33] in liquid xenon. This length scale was also used in the previous NEST version [34] to select the Thomas-Imel or Doke/Birks formalism for the quanta calculation. The

[1]The same time scale is used in the software *nSort*.

combination of the spatial and temporal clustering prohibits that energy depositions are grouped that happen in close distances from each other but separated in time.

For each cluster the energy of all including interactions is summed and the position and time of the cluster is calculated as weighted average of the individual positions and times with the corresponding energy depositions as weights. The primary particle type of a cluster is determined by the primary particle types of involved interactions. When a cluster has more than one primary particle type, both are used for the calculation of the quanta with NEST and an average is calculated weighted by the summed energy of the involved primary particle types. Following the clustering, all energy depositions happening outside the TPC are cut and the events are artificially separated by 1 s in time while conserving the temporal relations in each event for the usage in WFSim.

In the third step the number of created electrons and photons is calculated with the python bindings of NEST[2] [35] for each cluster. NEST allows the simulation of scintillation and ionization processes of noble gases like xenon. Despite its extensive capabilities, NEST is only used for the quanta calculation which depends on a few detector specific parameters outlined below. The generation of S1 and S2 signals is performed by the waveform simulation, created to accurately simulate the response of the XENON1T experiment. To calculate the number of electrons and photons for an given energy deposition five additional inputs are needed whereas four of them are detector specific and hence constant for all simulated events. The parameters are the mass of xenon atoms $A = 131.293$ [17], the atomic number $Z = 54$ [13], the density of liquid xenon in the TPC $\rho = 2.862\,\mathrm{g/cm^3}$ [36] and the electric drift field strength of $E = 82\,\mathrm{V/cm}$ in the XENON1T TPC. The fifth parameter is the choice of the model used in NEST for the quanta calculation. Different models for nuclear recoils, ions, ^{83m}Kr, electrons and γ-rays can be chosen with the default case being the beta model. Since the signal of interest and used calibration sources in this work consists solely of electrons and γ-rays, only the usage of these two models was evaluated.

A comparison of the light and charge yield for different NEST models used can be found together with light and charge yield values for XENON1T in figure 3.6. If a cluster is caused by an electron as primary particle the beta model will be used for the calculation of the light and charge quanta. If a cluster is created by a γ-ray, an weighted average of the γ and beta model is used for the quanta calculation (cf. [37]). In case of the three calibrations sources $^{131\mathrm{m}}$Xe, $^{129\mathrm{m}}$Xe and ^{212}Pb the measured calibration data could be best matched with a linear transition from the beta to the γ model between 90 keV and 270 keV. A detailed comparison for corrected areas of the S1 and S2 signals for measured and simulated calibration sources is given in section 4.3.2. For energies in the MeV scale larger differences between the NEST predictions and measured data can be found. Fluctuations of the charge and light yield determine the width of the S1 and S2 area distributions. It was found that

[2]Version 1.1.3 is used in this work.

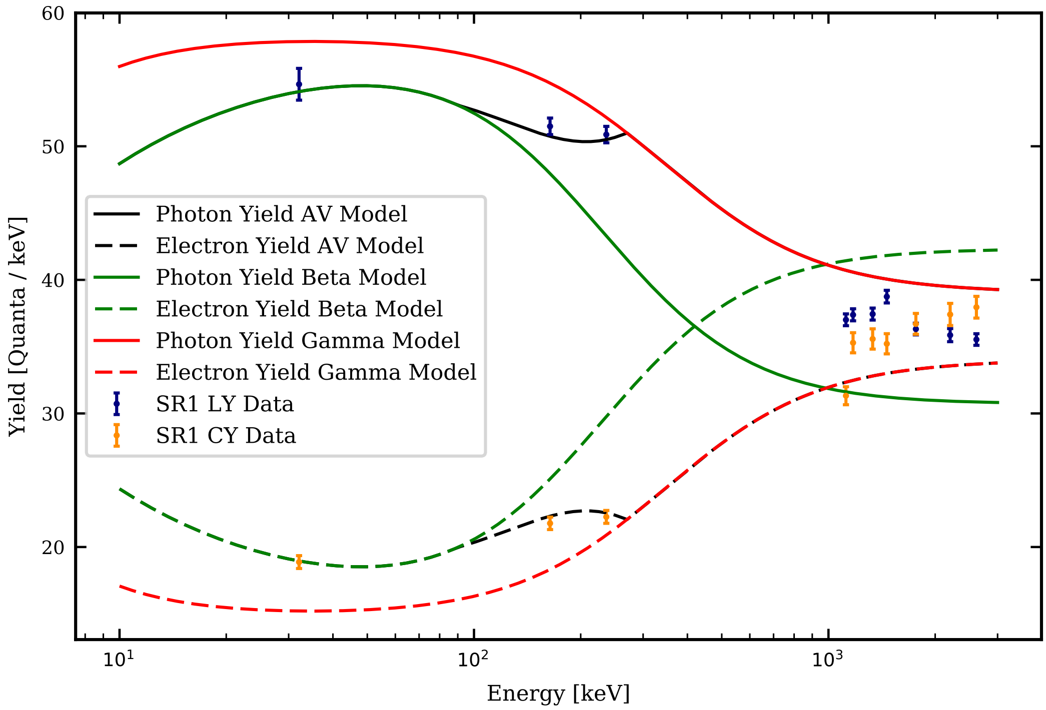


Figure 3.6: Charge and light yield in dependency of the deposited energy. Measured XENON1T charge yield values are given in orange and measured light yield values in blue. The NEST model for γ-rays is given in red, the model for electrons in green and the modified average (AV) model in black. The charge yield predictions are drawn as dashed lines and the light yield predictions as solid lines. The NEST predictions are calculated from single energy depositions whereas the given measured data includes events where multiple energy depositions are merged into one S1 and S2 signal.

the fluctuations simulated with NEST result in to narrow distributions compared to measured calibration data so that a modification was introduced:

$$\tilde{n}_{\mathrm{ph}} = n_{\mathrm{ph}} \cdot N(1, \sigma) \tag{3.3}$$

$$\tilde{n}_{\mathrm{e}} = n_{\mathrm{e}} + (n_{\mathrm{ph}} - \tilde{n}_{\mathrm{ph}}). \tag{3.4}$$

Here n_{ph} is the number of photons, n_{e} the number of electrons and $N(1, \sigma)$ a normal distributed random number with $\mu = 1$ and $\sigma = 0.073$. σ is chosen in such a way, that the width of the Gaussian shaped cS1 and $\mathrm{cS2_b}$ distributions of simulated and measured $^{131\mathrm{m}}$Xe and $^{129\mathrm{m}}$Xe events agrees best.

For future analysis the process quanta generation in the simulation chain needs to be improved in some areas. First the procedure of reconstructing the causal event hierarchy has to be tested on all decay simulations of the future XENONnT experiment. Furthermore it is important improve the agreement between predicted and measured charge and light yield values especially at higher energies.

3.1.4 Waveform Simulator

The last step of the simulation chain is a waveform simulator. It simulates the detector response to photons and electrons including effects of the PMTs and electronics. In this work the waveform simulator *WFSim* will be used which is currently developed for the upcoming XENONnT experiment, based on the Fake Xenon Experiment (FAX) which was developed for XENON1T. While benefiting from the computational performance enhancements implemented in WFSim, it is still possible to simulate XENON1T events. The simulator takes the number of quanta, times and positions as input. The electron drift to the liquid gas interface is simulated considering diffusion effects and reduction of electrons by electronegative impurities represented by an electron lifetime. A scintillation model [22] is used to generate the S2 light from the extracted electrons. Both the S1 and S2 light is distributed to the top and bottom PMTs using light collection efficiency (LCE) maps [21]. Afterwards, measured noise is added and the events are stored as raw data so that they can be processed with the Processor for Analysing XENON (PAX) using the same data analysis framework as measured data. Additional a *truth file* is saved to allow crosschecks on an event by event basis of the simulation and event processing.

The configuration parameters for WFsim are set based on measured data from the TPC and parameters determining the detector geometry, e.g. the number of PMTs. Three important configuration parameters that were investigated in this work are described here. The *s2_mean_area_fraction_top* was set to 0.585. This parameter determines the fraction of S2 light seen on the top PMT array with respect to the overall S2 size.The measured mean area fraction top of 0.627 is reproduced with the chosen configuration value. The *s1_detection_efficiency* was set to 0.126 and controls the number of detected S1 photons, and thus the S1 and cS1 area in an energy independent fashion. The third configuration parameter is the *s2_secondary_sc_gain*[3]. It was set to 28.5 and determines the overall S2 size. Using this set of configuration parameters it was possible to achieve a good agreement for the S1 and S2 sizes between the truth information and output of the full chain simulation after reconstruction with PAX.

3.2 Data Preparation

In the last section, the used simulation tools were described. In this work, both simulated and measured events are used. This section outlines the preparation of events for the analysis. The raw data is processed with the Processor for Analysing XENON (PAX) [38], the data processor developed for the XENON1T experiment. PAX version 6.10.1 is used for this work which also includes several reconstruction

[3] sc: scintillation

improvements for higher energies outlined in [39]. The relevant event information is than extracted from the processed data using a *treemaker* of the Handy Analysis for XENON (HAX) [40]. A detailed description of the event processing and reconstruction can be found in [24].

3.2.1 Data Extraction of Events with multiple S2 Signals

Due to the number of emitted particles and corresponding energies in events of the $\nu\beta\beta^*$-decay of ^{136}Xe, it is likely that an event contains one merged S1 and multiple S2s. Since most of the XENON1T analyses were performed for events with only one S2, which is the more common case for energies below $E \lesssim 400\,\mathrm{keV}$[39], a different method for the data processing had to be used here. A *treemaker* for the analysis of events with more than one S2 was designed for the high energy analysis of XENON1T [39]. The analysis in this work uses a modified version of this *treemaker*.

The *MultiS2 treemaker* analyses all peaks in one event which could in principle be S2 signals, whereas *treemakers* in single site analyses only work on the largest S2 in the event. A peak has to pass five criteria [41] to be classified as S2 in an multi site event for the analysis in this work.

1. The area of a peak must be larger than 150 pe. This reduces the computational cost of each event by omitting small S2s.

2. The reconstructed z-position of the interaction must be within the TPC: $-100 < z < 0$. The position is calculated for each peak by calculating the drift time with the main S1.

3. The goodness of the top pattern fit [24] is used to remove delayed electrons created by S2 signals.

4. The electron diffusion model is used to remove peaks with unphysical width [24]. This criteria further reduces delayed electrons passing the goodness of fit criteria and very small S2s.

5. A second width cut is used to remove miss classified S1 peaks.

Following these criteria, corrections are applied to each S2 individually. This includes a correction of the electron lifetime and a correction of the LCE [21]. Usually, the S1 of an event is corrected on position dependent effects where the position is obtained in combination with the main S2. In the multi side case multiple S1s created at each interaction side are merged into one S1 so that the corrections need to take this into account. Therefore, the corrected merged S1 area is defined as

$$\mathrm{cS1}_m = \mathrm{S1} \cdot \frac{\sum_i \mathrm{LCE}(\vec{x}_i) \cdot \mathrm{cS2}_{\mathrm{b},i}}{\sum_i \mathrm{cS2}_{\mathrm{b},i}}. \tag{3.5}$$

Here, S1 is the uncorrected S1 area, $\mathrm{LCE}(\vec{x}_i)$ the position depended S1 LCE, $\vec{x}_i$ the position and $\mathrm{cS2}_{\mathrm{b},i}$ the corrected S2 area seen on the bottom PMT array of interaction i. $\vec{x}_i$ is obtained using a neural network for the position reconstruction and a three dimensional field distortion correction [42] is applied for the measured TPC data. Since field distortion effects are not included in the MC simulations, the correction is not applied for simulated events. Additionally, the sum of all corrected S2 bottom areas is calculated as

$$\mathrm{cS2}_{\mathrm{b},m} = \sum_i \mathrm{cS2}_{\mathrm{b},i} \tag{3.6}$$

and returned by the treemaker. The mean event position is defined as

$$\vec{x}_m = \frac{\sum_i \vec{x}_i \mathrm{cS2}_{\mathrm{b},i}}{\mathrm{cS2}_{\mathrm{b},m}}. \tag{3.7}$$

Beside the variables introduced so far, the treemaker is modified for this work to return also the individual positions of each interaction $\vec{x}_i$ together with the corrected S2 bottom areas $\mathrm{cS2}_{\mathrm{b},i}$ for the ten largest S2s.

3.2.2 Data Quality Cuts

This subsection will briefly outline the data quality cuts that are used in this analysis. Cuts remove events from the data based on certain criteria. In case of the data quality cuts these criteria are mainly attributed to conditions of the detector. All five presented cuts are be applied to measured TPC data and only two cuts on the simulated events.

DAQ Veto The first data quality cut used is the DAQ-Veto cut [43]. This cut was designed to remove all events which are of technically bad quality caused by problems in the readout system. This includes incomplete events where some channels of the detector were not sensitive and during calibration runs events with large S2s are removed. During normal operation of the detector, the last 21 seconds of each run are removed. Each run has a typical length of 1 h. It was found that due to some effects in the digitizers, in some rare edge cases, events in this time frame are incomplete. To be conservative, these events are removed from the analysis. Additionally, events are removed when one of the digitizers runs out of memory. The cut reduces the overall exposure of the detector but introduces no bias in the events recorded. Since these readout effects are not simulated, this cut is only applied to the measured TPC data.

S2 - Tails It was found during the detector operations that large S2 signals are often followed by many low energetic S2s due to ionization of impurities in the LXe. These low energetic S2s appear in a timescale ranging from a few microseconds for small

S2s to a few seconds for high energetic S2s caused by e.g. muons. For each event, it is checked if it follows a large S2 and if so, this event is cut. This cut [44] does not depend on the event in question so it will not introduce a bias but reduce the total exposure. Since the effects of these tails is not included in the simulations, this cut is only applied to the TPC data.

Flash The third cut [45] only applied to TPC data is the so-called flash cut. This cut was designed to remove events caused by flashes in single PMT. A flash is caused by small xenon leaks in the PMTs. Just like for the previous cuts, the exposure is reduced by application of this cut.

S1 and S2 Greater Zero This general data quality cut requites the S1 and S2 area to be larger than zero and is applied to both the measured and simulated data. A complete event requires at least one S1 and one S2 signal.

Fiducial Volume Cut Like described in section 2.3, the 3D position reconstruction in combination with the high stopping power of LXe allows a background reduction when performing the analysis only in an inner part of the detector, called fiducial volume. In this work two different fiducial volume cuts are used. For the analysis of the xenon γ-lines, only the largest S1 and S2 with their corresponding position is used and the cut is defined as

$$0\,\mathrm{cm} < \quad r \quad < 36.94\,\mathrm{cm}, \tag{3.8}$$

$$-92.9\,\mathrm{cm} < \quad z \quad < -9\,\mathrm{cm}, \tag{3.9}$$

containing 1 t of xenon. Here, r is the radial coordinate of an interaction and z is the depth in the TPC. The TPC dimensions are shown in figure 3.7 as solid red lines and the 1 t fiducial volume as dashed lines in the r^2-z-space. A histogram of simulated $^{129\mathrm{m}}$Xe and $^{131\mathrm{m}}$Xe decays is shown on the left side and measured background data on the right side of the plot. The xenon decays are homogeneously distributed in the detector volume whereas the background events are mostly located towards the TPC walls. The number of events within the fiducial volume is greatly reduced compared to the whole detector volume.

The second fiducial cut in this analysis is a more stringent version of the fiducial volume cut outlined earlier, designed for events with a multiplicity greater than one. These events consist of more than one S2 and the treemaker used returns the 3D positions of each of the interactions in one event. When a decay happens within the fiducial volume cylinder the cut requires all interactions to take place in the fiducial volume as well. Thus, events with at least one energy deposition outside of the fiducial volume cylinder are cut making the outer volume to an active veto. This ensures that

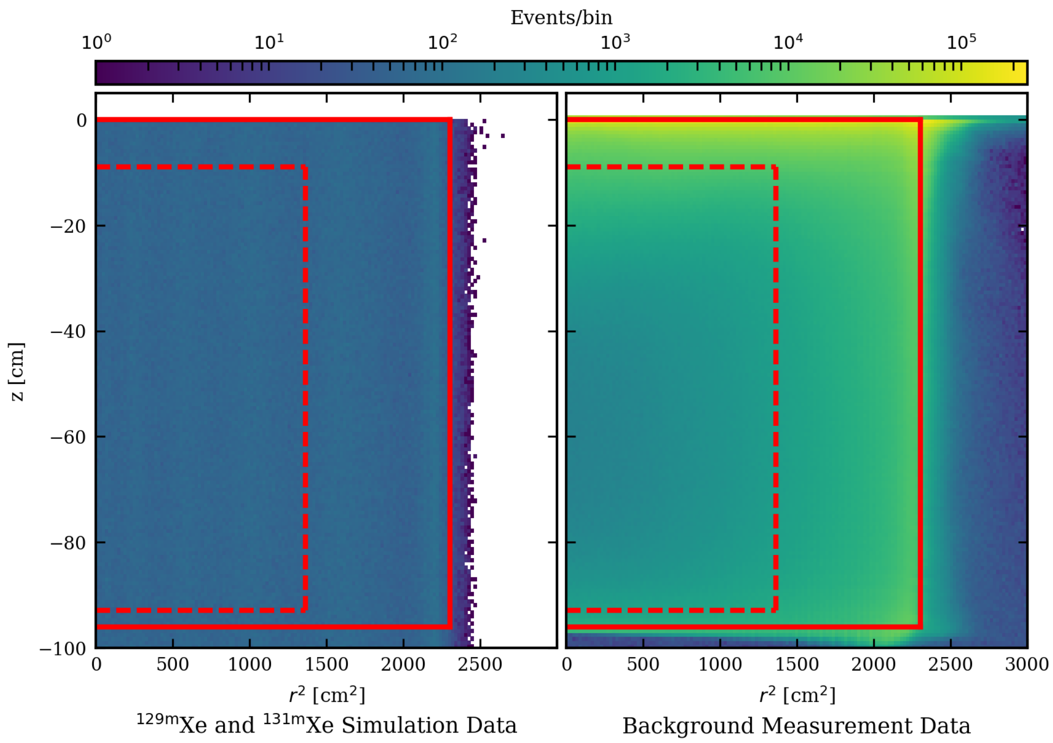


Figure 3.7: Two dimensional histograms of simulated ^{131m}Xe and ^{129m}Xe decays and TPC background data in r^2-z space before data quality cuts. The outer dimensions of the TPC are shown as red lines and the 1 t fiducial volume as dashed lines. The xenon isomers decay homogeneously in the TPC whereas most of the background events are external sources leading to higher count rates towards the outside of the TPC. Measured events can be found outside the detector volume due to position miss-reconstructions and field distortion correction.

events contain the full decay energy since it is rare that a decay happens in the fiducial volume and a γ-ray can escape the detector without interaction in the outer region of the TPC.

3.2.3 Preparation of simulated signal and measured background data

In the last subsection, a set of data quality cuts was outlined. These cuts are now applied to simulated $2\nu\beta\beta^*$-decay events of ^{136}Xe and measured background events.

Only about 8 signal events are expected to occur in the SR 1 data of XENON1T in the one tonne fiducial volume as calculated in section 2.2. In order to ensure that this low number of events is not further reduced by cuts, only a minimal set of data quality cuts is applied. The same cuts are also applied to the calibration data (cf. section 3.3). A brief summary of each cut is given in section 3.2.2. The cut history giving the number of removed and kept events for the background data is given in table 3.2.

Table 3.2: Cut history for the measured background data. Only data quality cuts are applied up to now to keep as much events as possible.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	32820744 (27.71 %)	85625828 (72.29 %)	72.29 %
DAQ Veto	901575 (1.05 %)	84724253 (98.95 %)	71.53 %
S2 Tails	3319739 (3.92 %)	81404514 (96.08 %)	68.73 %
Muon Veto	861819 (1.06 %)	80542695 (98.94 %)	68.00 %
Flash	860 (<0.01 %)	80541835 (>99.99 %)	68.00 %
1T fiducial volume	76450819 (94.92 %)	4091016 (5.08 %)	3.45 %

One finds that 27.71 % of the measured background events are removed by the requirement to have at least one S1 and one S2 signal in an event. The livetime reducing DAQ Veto, S2 Tails, Muon Veto and Flash cuts remove only a small fraction of events. The 1T fiducial volume cut for multi-site events removes 94.92 % of the events passing the previous cuts. This large fraction of removed events demonstrates the self-shielding capabilities of LXe against external sources as outlined in section 2.3. Only 3.45 % of the measured background events pass all applied cuts. The cut history of the simulated signal events is given in table 3.3. A smaller set of data quality cuts is

Table 3.3: Cut history for simulated signal events. Only two data quality cuts are used.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	138366 (3.13 %)	4279565 (96.87 %)	96.87 %
1T fiducial volume	2714988 (63.44 %)	1564577 (36.56 %)	35.41 %

applied to the simulated signal events. Effects that are targeted with the DAQ Veto, S2 Tails, Muon Veto and Flash cut are not included in the simulations and thus the corresponding cuts are not used here.

Only 36.56 % of the simulated signal events pass the fiducial volume cut for multi-site events. Geometrical considerations give an expected ratio of 50 % removed events indicating that more simulated signal events are cut as expected. As described in section 3.2.2 the fiducial volume cut for events with multiplicity greater unity turns the outer detector volume into an active veto. The mean distance a particle can travel in LXe from its origin depends on the particle type and energy. In case of the $2\nu\beta\beta^*$-decay of ^{136}Xe two γ-rays are emitted back-to-back along with two electrons. While the electrons interact close to the position of the decay, the γ-rays can travel distances long enough to create separate S2s. If one of the γ-rays of the $2\nu\beta\beta^*$-decay occurring in the fiducial volume interacts in the outer detector region, the event is cut. It is found that 30.18 % of the signal decays in the fiducial volume are removed by the fiducial volume cut.

3.3 Calibration data

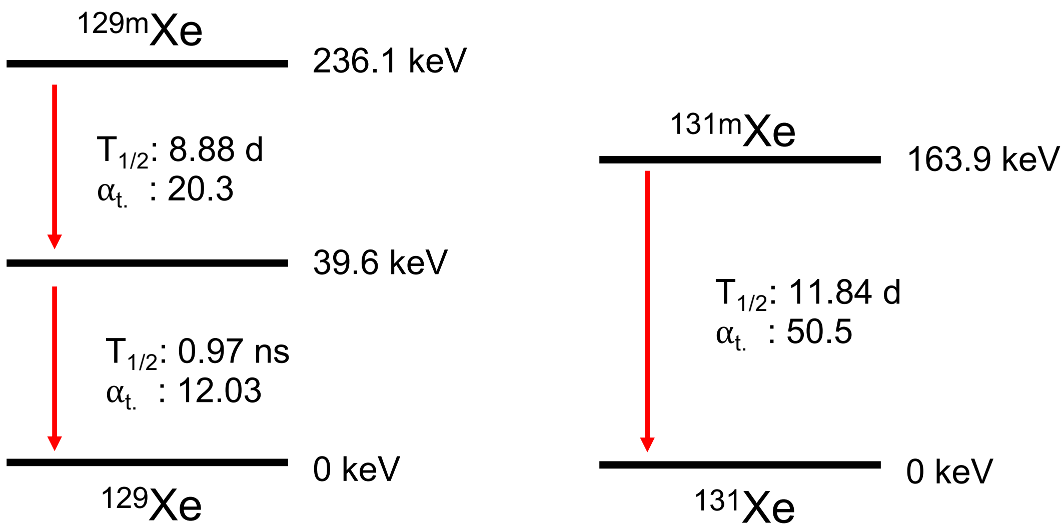
The simulation and data preparation tools outlined in the last two sections have to be validated by comparison to measured data. In this section the four calibration sources ^{129m}Xe, ^{131m}Xe, ^{212}Pb and ^{60}Co used in this work are described including a general source description and the selection of the events.

3.3.1 ^{129m}Xe and ^{131m}Xe

The first two calibration sources are the isomers ^{129m}Xe and ^{131m}Xe. Both are mono energetic γ-ray sources. This makes them important for the validation and tuning of the MC simulations since certain distributions of event parameters like the energy or cS1 and cS2 areas can be compared directly.

Source Description

Both isomers are created by neutron scattering reactions [46] during neutron generator calibration runs. The neutron generator was frequently used during the detector operation in SR 0 and 1. Due to the half-lifes of 8.88 d for ^{129m}Xe and 11.84 d for ^{131m}Xe, the decays of the isomers can be found in the background data of the detector.



(a) Scheme of ^{129m}Xe decay, data taken from [13].

(b) Scheme of ^{131m}Xe decay, data taken from [13].

Figure 3.8: Decay schemes of ^{129m}Xe and ^{131m}Xe. γ-transition are given as red arrows. Both isomers are produced during neutron generator calibration runs and are present in the background data of XENON1T due to the half-lifes in the order of days.

Figure 3.8a shows the decay scheme of ^{129m}Xe. The deexitation [13] is a two staged decay via an intermediate stage of 39.6 keV with short half-life 0.97 ns so that two

γ-rays are emitted. Due to the γ-ray energies of 196.5 keV and 36.6 keV and the short half-life of the intermediate stage, both γ-rays will be merged into a single S2 signal in most cases. ^{131m}Xe decays in a single step to the ground state via emission of a 163.9 keV γ-ray. The corresponding decay scheme is shown in figure 3.8b. Both isomers have a non-negligible conversion factors (cf. figure 3.8a and figure 3.8b) leading to the emission of electrons during the decay.

Event Selection

In order to select only the ^{129m}Xe and ^{131m}Xe decay events, multiple cuts are applied both to the simulated and measured background events. First, the data quality cuts described in section 3.2.2 are applied followed by cuts in cS1, cS2 and reconstructed energy. Since both decays are mono-energetic, they will show up as tilted two-dimensional Gaussians in the cS1-cS2 space and thus as one-dimensional Gaussians in the energy. The overall event selection in cS1 $cS2_b$ space is shown in figure 3.9.

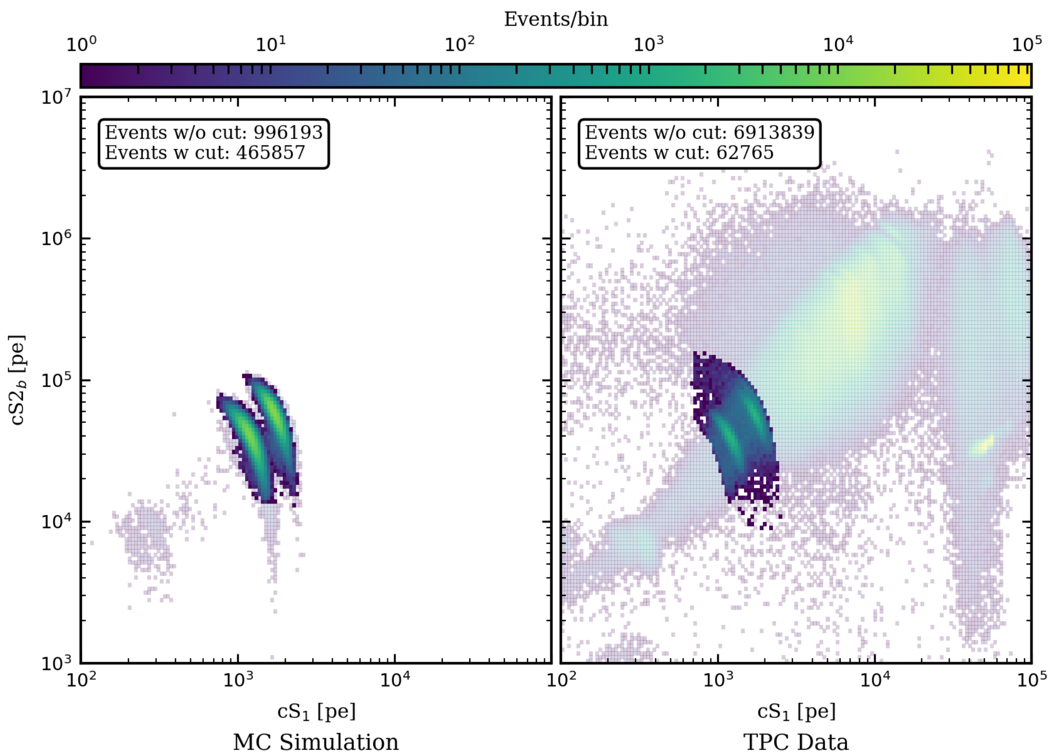


Figure 3.9: Comparison of the event selection of ^{131m}Xe and ^{129m}Xe decays for simulated (left) and measured (right) events in cS1-$cS2_b$ space.The same number of events is simulated for ^{131m}Xe and ^{129m}Xe causing that both peaks have the same area. Bins of events that are removed by cuts are given with a reduced opacity.

The cut history for the simulations of ^{129m}Xe decays is given in table 3.4 and for simulated $^{131\mathrm{m}}$Xe events in table 3.5. The cut history of the measured TPC events can be found in table 3.6.

Table 3.4: Cut history of simulated $^{129\mathrm{m}}$Xe events. Two data quality cuts are applied and the mono energetic source is selected by cuts in cS1, cS2 and energy. A table giving the fraction of events passing a cut is given in table A.1.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	32 (0.01 %)	499462	99.99 %
1T fiducial volume	265770 (53.21 %)	233692	46.79 %
Source Selection			
700 pe < cS1 < 2500 pe	1 (0.01 %)	233691	46.79 %
4×10^4 pe < cS2 < 4×10^5 pe	6 (0.01 %)	233685	46.78 %
190 keV < E < 260 keV	270 (0.12 %)	233415	46.73 %

The cuts in cS1 and cS2 are applied select the region of interest. The cut history shows that the cS1 cut keeps all but one event for simulated $^{129\mathrm{m}}$Xe and $^{131\mathrm{m}}$Xe while removing 97.95 % of the measured background data. The cS2 cut further improves the event selection by removing 18.73 % of the remaining measured background events and keeping almost all simulated events.

Table 3.5: Cut history of simulated $^{131\mathrm{m}}$Xe events. A similar set of cuts is applied compared to the selection of $^{129\mathrm{m}}$Xe events but the energy region is varied to match the differing decay energy. A table giving the fraction of events passing a cut is given in table A.2.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	19 (<0.01 %)	496751	99.99 %
1T fiducial volume	263983 (53.14 %)	232768	46.86 %
Source Selection			
700 pe < cS1 < 2500 pe	0 (0 %)	232768	46.86 %
4×10^4 pe < cS2 < 4×10^5 pe	304 (0.13 %)	232464	46.80 %
130 keV < E < 190 keV	22 (0.01 %)	232442	46.79 %

After this two cuts shared by both $^{129\mathrm{m}}$Xe and $^{131\mathrm{m}}$Xe, a cut in energy is performed to select each source individually. How the energy of an event is reconstructed from cS1 and cS2$_b$ is outlined in section 4.3.3. Just like for the previous cuts, almost all simulated events pass the energy cut while only 22.75 % of the remaining events pass the energy cut for $^{131\mathrm{m}}$Xe and only 22.50 % for $^{129\mathrm{m}}$Xe respectively.

The source selection removed only 277 simulated $^{129\mathrm{m}}$Xe events and 326 simulated $^{131\mathrm{m}}$Xe events. A low rate of removed events is expected since only the two γ-ray sources were simulated here. In contrast, the measured data from the TPC contains also other sources of background. Thus the source selection cuts are required to remove events except of the xenon decays. One can see that only 0.03 % of the measured

Table 3.6: Cut history for the selection of ^{129m}Xe and ^{131m}Xe events in the background data of XENON1T. The data used here is loaded technically different from the simulated events and the S1 > 0 pe, S2 > 0 pe cut is applied implicitly so that no number can be given here. A table giving the fraction of events passing a cut is given in table A.3.

Cut	Events removed	Events passed	Fraction left
Data Quality Cuts			
DAQ Veto	1070018 (1.06 %)	99843936	98.94 %
S2 Tails	4031907 (4.04 %)	95812029	94.18 %
Muon Veto	1011325 (1.06 %)	94800704	93.18 %
Flash	1019 (<0.01 %)	94799685	93.18 %
1T fiducial volume	86461272 (91.20 %)	8338413	8.20 %
General Source Selection			
700 pe < cS1 < 2500 pe	8150754 (97.95 %)	170669	0.17 %
4×10^4 pe < cS2 < 4×10^5 pe	31974 (18.73 %)	138695	0.14 %
^{131m}Xe			
130 keV < E < 190 keV	107143 (77.25 %)	31552	0.03 %
^{129m}Xe			
190 keV < E < 260 keV	107482 (77.50 %)	31213	0.03 %

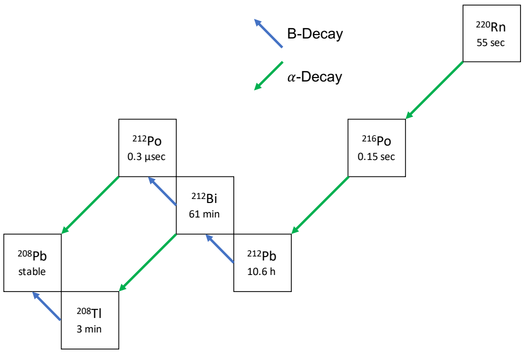
background events pass all cuts for ^{129m}Xe and ^{131m}Xe respectively. Since the source is only present in the background spectrum and not used in dedicated calibration runs with increased count rate, most of the events had to be cut as background events.

3.3.2 ^{212}Pb

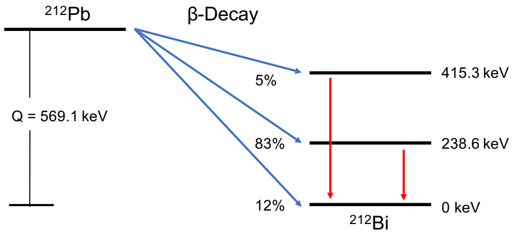
The next calibration source used in this work is ^{212}Pb. ^{212}Pb decays via β^--decay to ^{212}Bi and a sufficient number of events can be measured in calibration runs where ^{220}Rn is injected in the TPC. Due to the β^--decay directly to the ground state ^{212}Pb can create low energetic electronic recoil signals which makes it an important calibration source for XENON1T.

Source Description

^{212}Pb is part of the thorium decay chain from which the relevant part is shown in figure 3.10a. ^{220}Rn is injected in the TPC as calibration source and decays via α-decay to ^{216}Po. ^{212}Pb is obtained after an second α-decay and will undergo a β^--decay to ^{212}Bi. The β^--decay goes in about 12 % of the cases directly to the ground state of ^{212}Bi and subsequent γ-transitions are likely to occur. This event topology is particularly useful for validation of the MC simulations in this work since the excited $2\nu\beta\beta$-decay of ^{136}Xe also involves the emission of electrons and γ-rays. Since only the decay of ^{212}Pb is of interest, only this decay is simulated instead of the full decay chain to save computation time. To allow a proper comparison to the measured calibration data several cuts are applied to remove α-decays and other background events from the datasets.



(a) Decay chain of ^{220}Rn, data taken from [13].



(b) Decay scheme of ^{212}Pb, data taken from [13].

Figure 3.10: Decay chain of ^{220}Rn and decay scheme of ^{212}Pb. ^{212}Pb decay via β^--decay to ^{212}Bi. β-decays are indicated as blue arrows, α-decays as green arrows and γ-transitions as red arrows.

Event Selection

In order to select only ^{212}Pb decays in the simulated and the measured calibration data, a set of basic data quality and more specific event selection cuts were applied to the data. The cut history of the simulated data is given in table 3.7 and the cut history of the measured calibration data is listed in table 3.8. The data quality cuts are outlined in section 3.2.2.

Table 3.7: Cut history of simulated ^{212}Pb decays. Four source specific cuts were used to separate the events of interest. Since only ^{212}Pb decays were simulated, the cuts remove only a small fraction of events. A table giving the fraction of events passing a cut is given in table A.4.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	2268 (0.49 %)	457288	99.51 %
1T fiducial volume	228219 (49.91 %)	229069	49.85 %
Source Selection			
cS1 < 5000 pe	0 (0 %)	229069	49.85 %
cS2 > 10 000 pe	4299 (1.88 %)	224770	48.91 %
$2.3 < \ln\left(\frac{cS2}{cS1}\right) < 5.5$	33 (0.01 %)	224737	48.90 %
E < 570 keV	40 (0.02 %)	224697	48.89 %

The first two source specific cuts are performed in cS1-cS2 space demanding cS1 < 5000 pe and cS2 > 10 000 pe. 98.12 % of the simulated ^{212}Pb events pass this cut while about half of the measured calibration events are removed. The next cut is carried out in $\ln\left(\frac{cS2}{cS1}\right)$ space. The cut removes 33 simulated ^{212}Pb events and 2.27 % of the measured events. The last cut limits the allowed total energy deposition E of an event. The energy reconstruction using cS1 and $cS2_b$ is outlined in section 4.3.3. The Q-value of the ^{212}Pb decay is $Q = 569.1\,\text{keV}$[13] which corresponds to the maximal energy that can be deposited by this decay. Thus, events with $E > 570\,\text{keV}$ are cut. 40 simulated events are removed by this cut and 2.07 % of the calibration data are removed.

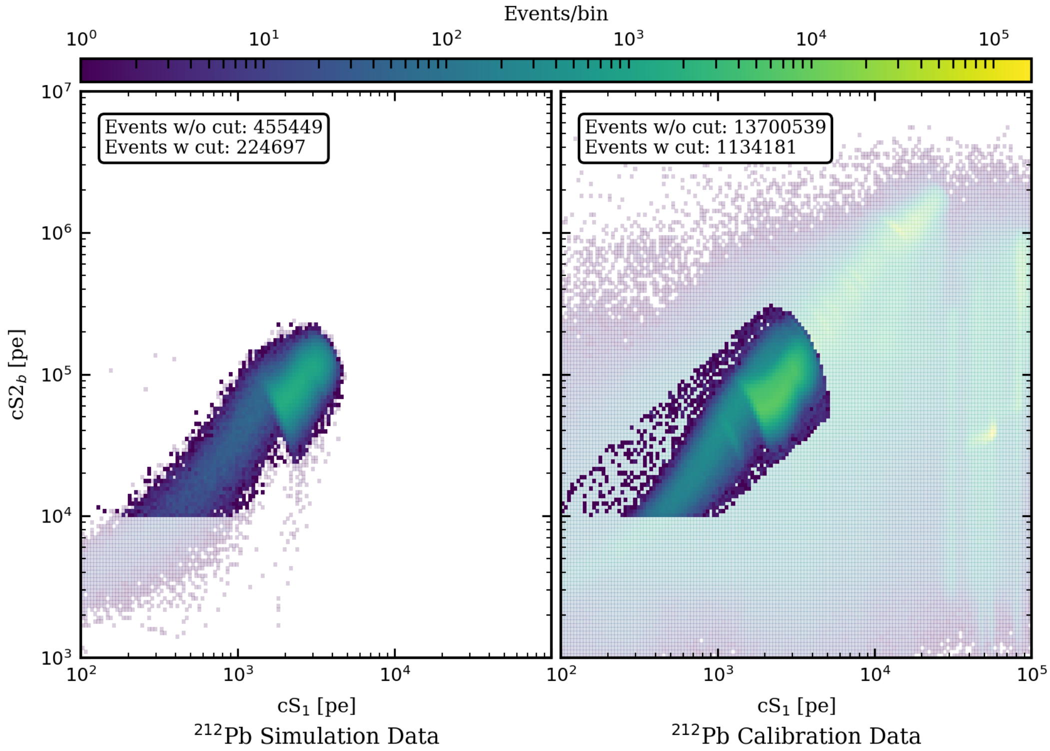


Figure 3.11: Comparison of the event selection of ^{212}Pb decays for simulated events and calibration data from the TPC in cS1-$cS2_b$ space. The decay to the ground state creates a "tail" starting at small cS1 and $cS2_b$. The decays to excited states of ^{212}Bi show up as kink in the data at around cS1$= 1 \times 10^3$ pe. A small population of events above the kink in $cS2_b$ is visible for the calibration data whereas no such population exists in the simulated data. Other events from the decay chain passing the selection could cause this population and wont show up in the simulated data since only the ^{212}Pb decay is simulated.

The resulting event selection is shown in figure 3.11 in the cS1-$cS2_b$ space and in the E-$\ln\left(\frac{\text{cS2}}{\text{cS1}}\right)$ space in figure A.1 in the appendix. Since α-decays usually deposit high energies in the detector, the presented cuts should remove these events efficiently. Possible events passing these selections could come from the β^--decays of ^{212}Bi and ^{208}Tl. Due to the short-half life of ^{212}Po, the β^--decay of ^{212}Bi is in most cases directly followed by an α-decay so that both decays will be merged in one event. This so-called BiPo-coincidence will lead to a removal of these events due to the applied cuts. The β^--decay of ^{208}Tl has a high Q value of 4998.5 keV[13] and goes in all cases [13] to excited states of ^{208}Pb. These events are also cut unless the γ-ray can escape the detector without energy depositions in the outer detector region. Therefore it is also unlikely that these β^--decays can pass the source selection cuts.

Table 3.8: Cut history for ^{220}Rn calibration data. Beside the data quality cuts the same source specific cuts used for the simulated data is applied here to select ^{212}Pb decays in TPC Data. The large fraction of events removed by the fiducial volume cut compared to the 50 % expectation from geometrical effects could be caused by other external sources and by the inlet of the ^{220}Rn in the TPC (cf. figure 4.16). A table giving the fraction of events passing a cut is given in table A.5.

Cut	Events removed	Events passed	Fraction left
Data Quality Cuts			
S1 > 0 pe, S2 > 0 pe	16846910 (54.66 %)	13971868	45.34 %
DAQ Veto	1590225 (11.38 %)	12381643	40.18 %
S2 Tails	582871 (4.71 %)	11798772	38.28 %
Muon Veto	63854 (0.54 %)	11734918	38.08 %
Flash	0 (0 %)	11734918	38.08 %
1T fiducial volume	9547063 (81.36 %)	2187855	7.10 %
Source Selection			
cS1 < 5000 pe	966899 (44.20 %)	1220956	3.96 %
cS2 > 10 000 pe	35825 (2.93 %)	1185131	3.85 %
$2.3 < \ln\left(\frac{cS2}{cS1}\right) < 5.5$	26946 (2.27 %)	1158185	3.76 %
E < 570 keV	23995 (2.07 %)	1134190	3.68 %

3.3.3 ^{60}Co

The next calibration source used in this work is ^{60}Co. Unlike the ^{212}Pb calibration, ^{60}Co decays are present in the background spectrum of XENON1T and not in dedicated calibration measurements. ^{60}Co is used as calibration source due to its emission of 1173.2 keV and 1332.5 keV γ-rays [13] providing a cross-check of the simulations at higher energies.

Source Description

^{60}Co is a radioactive cobalt isotope produced via neutron capture. Since natural cobalt can be found in steel, the usage of the neutron generator can produce ^{60}Co in the cryostat of the detector. In most cases, ^{60}Co decays via β^--decay to the 2505.7 keV excited state of ^{60}Ni [13] which then relaxes to the ground state via emission of two γ-rays. Since the decay happens mostly in the cryostat, only the deexitation γ-rays can reach the sensitive detector volume or even the fiducial volume.

Due to the energies of 1173.2 keV and 1332.5 keV[13], the dominating interaction mechanism of these deexitation γ-rays is Compton scattering. Thus, one can expect events with high multiplicities. Beside the Compton continuum, one should also find mono energetic full absorption peaks for both γ-rays in the data.

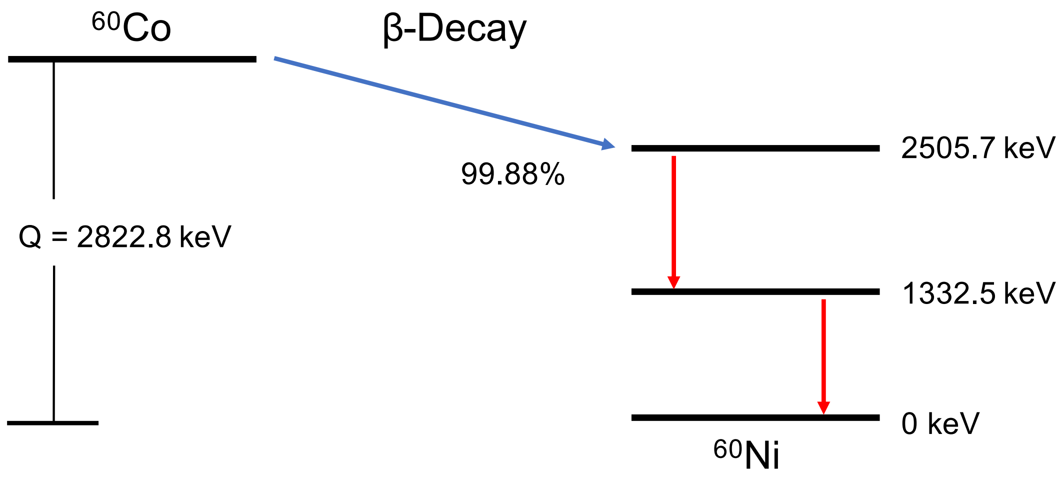


Figure 3.12: Decay scheme of ^{60}Co, data taken from [13]. The β-decay is indicated as blue arrow and the γ-transitions as red arrows.

Event Selection

As described above, ^{60}Co decays are present in the background data of the TPC and it is not a specifically designed calibration source. Thus, the count rate of these events cannot be artificially increased to rates of other calibration sources like ^{220}Rn. Only the mono-energetic parts of the energy spectrum can be separated efficiently from the detector background thereby a comparison of the Compton continuum between simulations and measured data is not possible. Just like for the other calibration data, a set of data quality cuts is applied both to the MC and TPC data. The cut history of the simulated data is given in table 3.9 and for the measured events in table 3.10.

Both for simulated ^{60}Co decays in the cryostat and for measured background data the multi site fiducial volume cut removes about 95 % of the available data. This large fraction of removed events is caused by the external nature of most background sources as-well as the calibration source. After using the fiducial volume cut , a range selection in cS1 and cS2$_b$ is performed to narrow the region of interest. Furthermore a multiplicity cut is applied since the presented analysis in this theses is done on events with more than one S2 in a single event. The signal multiplicity is outlined in section 4.3.1

The full absorption peaks of the two γ-rays are selected by cuts in Energy. Section 4.3.3 outlines how the energy of an signal is reconstructed using the corrected S1 and the corrected S2 signal seen on the bottom PMT array. Different energy windows had to be used for the simulations and measured data. The underlying energy shift is caused by a deviation of the NEST prediction from the measured data. This discrepancy in cS1 and cS2$_b$ is further described in section 4.3.2.

Table 3.9: Cut history for simulated ^{60}Co decays. Only 2.4×10^4 of 1.5×10^6 simulated events passes the cuts so simulations with a sufficient statistics in the fiducial volume are hard to achieve while keeping the computational costs as low as possible. The energy range is varied from the cut applied to the TPC data due to a discrepancy between the predicted light and charge yields from NEST with the measured values in XENON1T. A table giving the fraction of events passing a cut is given in table A.6.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	156412 (10.30 %)	1361967	89.70 %
1T fiducial volume	1301662 (95.57 %)	60305	3.97 %
General Source Selection			
2.5×10^5 pe < $cS2_b$ < 1×10^6 pe	9056 (15.02 %)	51249	3.38 %
4×10^3 pe < cS1 < 1×10^4 pe	4685 (9.14 %)	46564	3.07 %
Multiplicity > 1	13420 (28.82 %)	33144	2.18 %
1173.2 keV ^{60}Co line			
1150 keV < E < 1250 keV	21825 (65.85 %)	11319	0.75 %
1332.5 keV ^{60}Co line			
1300 keV < E < 1500 keV	20157 (60.82 %)	12987	0.86 %

Table 3.10: Cut history for the selection of the 1332.5 keV and 1173.2 keV ^{60}Co lines from the measured background data. A table giving the fraction of events passing a cut is given in table A.7.

Cut	Events removed	Events passed	Fraction left
Data Quality Cuts			
S1 > 0 pe, S2 > 0 pe	32820744 (27.71 %)	85625828	72.29 %
DAQ Veto	901575 (1.05 %)	84724253	71.53 %
S2 Tails	3319739 (3.92 %)	81404514	68.73 %
Muon Veto	861819 (1.06 %)	80542695	68.00 %
Flash	860 (<0.01 %)	80541835	68.00 %
1T fiducial volume	76450819 (94.92 %)	4091016	3.45 %
General Source Selection			
2.5×10^5 pe < $cS2_b$ < 1×10^6 pe	1543775 (37.74 %)	2547241	2.15 %
4×10^3 pe < cS1 < 1×10^4 pe	420188 (16.50 %)	2127053	1.79 %
Multiplicity > 1	541664 (25.47 %)	1585389	1.34 %
1173.2 keV ^{60}Co line			
1130 keV < E < 1210 keV	1256986 (79.29 %)	328403	0.28 %
1332.5 keV ^{60}Co line			
1300 keV < E < 1360 keV	1273819 (80.35 %)	311570	0.26 %

A comparison of the event selections for simulated and measured events in the cS1-$cS2_b$ space is given in figure 3.13. The MC data show a continuum with two mono-energetic Gaussian peaks whereas the spectrum of the measured data also includes peaks from other sources.

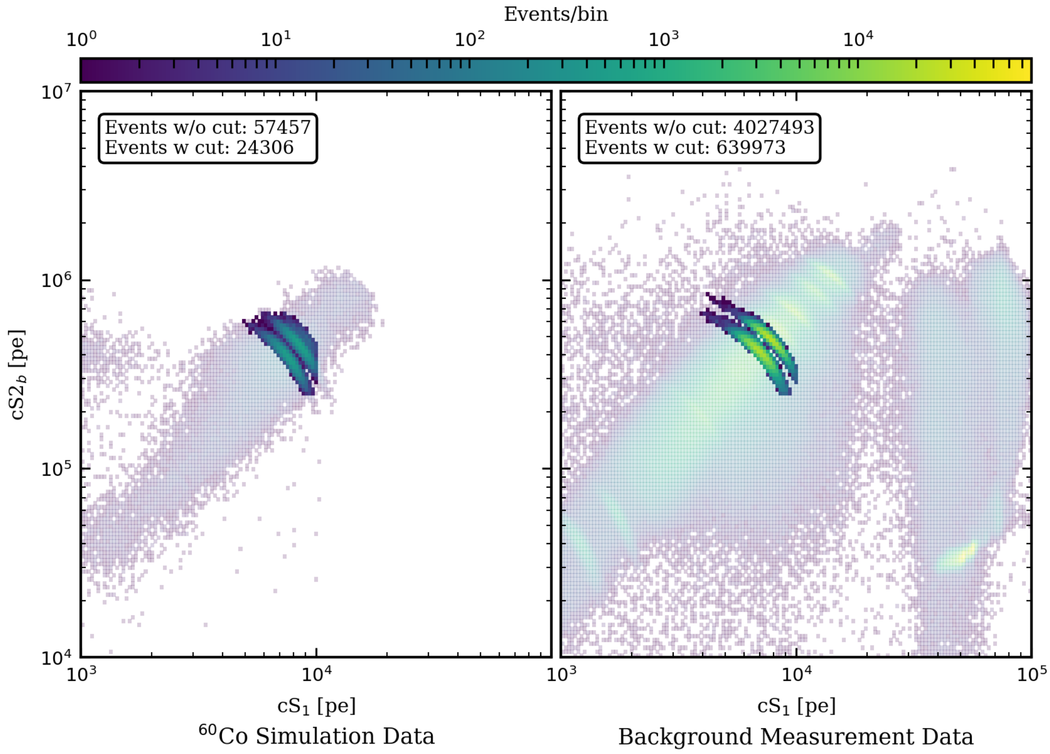


Figure 3.13: Comparison of the event selection in cS1-cS2$_b$ space for simulated ^{60}Co decays (left) and measured background data. (right) The full absorption peaks of the 1173.2 keV and 1332.5 keV γ-rays are visible as tilted two-dimensional Gaussians. The simulations show an additional continuous distribution of events due to energy depositions from Compton scatters. The Compton continuum is not visible in the measured data due to the large number of background events.

4 Development of a Machine Learning Discriminator

In the last chapter, the data preparations and calibration sources used were described. This chapter will focus on the development of a machine learning discriminator to distinguish between signal events of the ^{136}Xe $\nu\beta\beta^*$-decay and background events. First, general machine learning (ML) techniques and features are described. Then the input parameters for a discriminator used in an analysis by the EXO-200 collaboration are outlined. Inspired by the EXO-200 analysis, a multi-dimensional discrimination space consisting of several possible input parameters for this work is set up and investigated. For each parameter, the simulation results are validated using the calibration sources outlined in the last chapter. Two different ML models are trained and the performance is analyzed and compared. Finally the effect of signal events in the background class on the training process is investigated.

4.1 Machine Learning

Machine Learning is a sub-field of the computer science domain of artificial intelligence (AI). Investigations about AI started in the 1950s to use computers to perform intellectual tasks normally done by humans. It includes learning based approaches as well as approaches with large sets of programmed rules usually called *symbolic AI* [47]. Due to the difficulties of manually programming rules for all possible scenarios, learning based approaches are now mainly used in the field of AI.

In ML, data-processing rules are not created by hand but automatically learned by exposure to data. In contrast to classical programming where the output is obtained by human input of explicit rules and data, ML takes the data and corresponding desired outputs and learns the underlying rules which can than be applied to new data [48].

Learning in this field means generalization from seen data so that a task can be accurately performed on new unseen data. Rapid developments in the field of ML were possible in the last decades due to faster hardware and larger available datasets [48].

4.1.1 Categories

Algorithms are usually classified into four broad categories based on their training procedure [48]. These are supervised, self-supervised, unsupervised and reinforcement learning. Supervised and unsupervised learning are applied in this work.

Supervised learning is the most common case in machine learning applications. These algorithms try to map the input of a model to previously known target values. The most frequent applications of supervised learning are classification or regression tasks. In a classification task, the target values are labels that assign a combination of input values to a certain class. A popular example is the classification of handwritten digits. Target values in a regression task are continuous numbers, e.g. the three dimensional position of an interaction in the TPC. In this work, supervised learning is used to separate background from signal events in binary classification tasks. The basic training procedure of a supervised model is shown in figure 4.1. A distance metric, called loss function, is calculated between the output of a model for certain input values and the known target values. The parameter of the model are changed by an optimizer to minimize the output of the loss function, called loss.

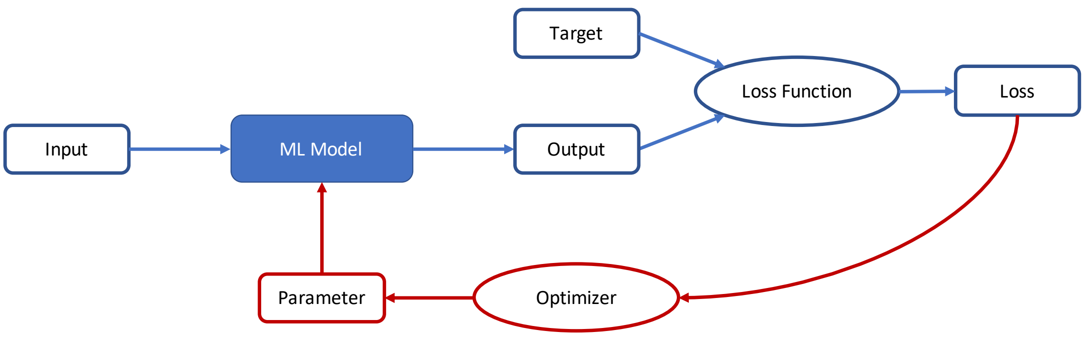


Figure 4.1: Supervised learning training procedure. The output of the machine learning (ML) model is calculated for a set of input values. The loss function is used to calculate a distance between the output and the target values of the specific task. The calculated distance is commonly called loss and is used by the optimizer to change parameters of the ML model. The process is repeated until a sufficient low loss is achieved.

Unsupervised learning is used on data where no targets are available to find useful transformations for e.g. data visualization, compression or de-noising. Another popular example are clustering algorithms where data sample are grouped to clusters based on similarities in certain input parameters. In this work, a DBSCAN [31] clustering is used during the MC simulations outlined in section 3.1.3 to merge energy depositions in the detector based on their three dimensional position and time.

4.1.2 Performance evaluation

The performance of a ML model is measured by its capability to generalize features from seen data, the training data, to unseen data. If a model is continuously trained on a dateset, the performance of the model on this data will increase with the training, but the performance on unseen similar data can decrease. This problem is called overfitting. To prevent overfitting and measure the generalization capability of a model, a common approach is to divide the data available for the development of a ML into three sets[48]: training, validation and evaluation[1] data:

- The machine learning model will be trained on the training dataset.
- The validation data will be used to test the performance of a machine learning model during the training and adjustments of the model's hyper-parameters. These parameters, for example the number of layers in a neural network, are not changed by the optimizer in training itself but by human input based on the performance of the model on the validation data.
- Since the hyper-parameters are adjusted during the development in such a way that the model will perform well on the validation data, a so-called *information leak* can occur. The model could be fine tuned on the validation data to perform artificially well on this data it was never directly trained on. A test with the evaluation data can be performed after the model development and training is finished to independently evaluate the performance of the model on unseen data.

When enough data is available, a simple split into these three datasets can be performed. This method of data division is commonly called *hold out*. When the statistics is limited, one can encounter the problem that a split into these tree sets will lead to a bias due to unequal feature distributions. In this case, a K-fold cross validation [49] can be used.

4.1.3 Neural Networks

Artificial neural networks (NNs) are machine learning algorithms remotely inspired by biological systems like brains. The key components of these networks are neurons and their connections. In general, a neuron receives an input, applies an activation function and returns an output. In most cases, neurons are connected to other neurons where each connection has a certain weight. These weights are changed in the training. During the last decade, many different types and architectures of NNs were developed and a loose classification can be made based on the direction of the information flow [48]. In feed forward networks (FFNs) information is only transmitted in one

[1] This dataset is also frequently called test dataset.

direction, from the input to the output whereas recurrent networks can also include backward connections and loops.

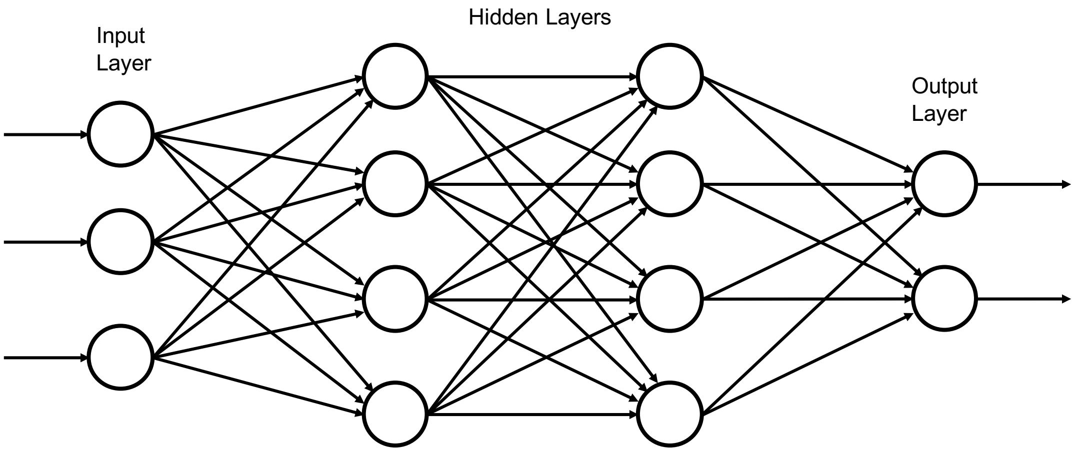


Figure 4.2: Illustration of a simple feed forward network. Neurons drawn as circles are grouped in layers. The output of each neuron is transmitted to neurons in the next layer indicated as arrows. The network in this example receives three input values and calculates two output values. Two hidden layers with four neurons each are shown. This architecture of the network has to be adapted to the specific task and requirements.

One of the simplest FFNs is the multi-layer perceptron (MLP). Like the name suggest, neurons are ordered in densely connected layers. A schematic representation of such a model is given in figure 4.2. This network consist of an input layer, a range of hidden layers and an output layer. The number of neurons in each layer and the amount of hidden layers vary and determine the capacity of the model. Bigger networks with more layers and neurons can deal with increasingly complex problems but tend to overfit.

In case of a MLP with two hidden layers similar to the network illustrated in figure 4.2, the output of the NN y_i is calculated in the following manner [50]. The output $h_i^{(1)}$ of the first hidden layer is calculated as

$$h_i^{(1)} = \varphi^{(2)}\left(\sum_j w_{ij}^{(2)} x_j + b_i^{(2)}\right).$$

Here, x_i is the vector of the input values, $\varphi^{(k)}$ the activation function of layer k and $b_i^{(k)}$ the bias vector. In general, all neurons of a layer use the same activation function. Here, $w_{ij}^{(k)}$ is the weight matrix connecting the neurons of layer k with the neurons of the previous layer. The weight matrices and bias vectors are the trainable parameter of a MLP that can be changed in the training process. The output of the second hidden layer is calculated similar but the output of the first hidden layer is used as input.

$$h_i^{(2)} = \varphi^{(2)}\left(\sum_j w_{ij}^{(2)} h_j^{(1)} + b_i^{(2)}\right).$$

The overall output of the network is then calculated from the output of the second hidden layer.

$$y_i = \varphi^{(3)}\Big(\sum_j w_{ij}^{(3)} h_j^{(2)} + b_i^{(3)}\Big).$$

If non-linear activation functions are used, even a small FFN is a universal function appropriator [51]. Nowadays, a widely used activation function is the rectified linear unit (ReLU) and defined as

$$\varphi(q) = \begin{cases} 0 \text{ , for } q \leq 0, \\ g \text{ , for } q > 0, \end{cases} \tag{4.1}$$

where q is the input of the activation function.

Training of a Neural Network

A NN is trained on data by changing the parameters in the weight matrices $w_{ij}^{(k)}$ and bias vectors $b_i^{(k)}$. Both are usually filled with random numbers at the start of the training and modified to minimize a distance metric, called loss function, between the output of the model and the target. The loss function is specific to the task the network should solve.

The minimization could be solved for only a few parameters analytically but in a typical model thousands or even millions of parameters will make an analytical solution impossible. All operations of the layers are differentiable which allows the calculation of the loss functions gradient with respect to the trainable parameters. The parameters can then be modified by going in the opposite direction of the gradient. This *stochastic gradient descent* procedure in its simplest form is performed in five repeating steps until a sufficiently small loss is achieved:

1. A sample of the training data together with the corresponding targets is selected.
2. The output of the model is calculated for the selected data.
3. The loss function is calculated for the obtained output and the selected target values.
4. The gradient is calculated with respect to all trainable parameters.
5. The parameters are adjusted in the opposite direction of the gradient within a defined step size.

The gradient descent is performed by a so-called *optimizer* that tries to find the global loss minimum. A whole range of *optimizers* was developed over the last decades including sophisticated methods to avoid problems of local loss minima. A frequently used *optimizer* is *Adam* [52]. When dealing with networks of many layers, the chain rule has to be used to calculate the loss gradient. This process, called *backpropagation*, starts at the output layer and works itself backwards through the whole network to adjust each trainable parameter. Modern packages like tensorflow [53] use symbolic differentiation, the calculation of derivatives using computer algebra, so the *backpropagation* algorithm must not be implemented by hand for each model.

4.1.4 Bagging and Boosting

Another common machine learning technique is the ensembling of *weak learners* when dealing with structured data [54]. A *weak learner* is a rather simple ML model like a decision tree which performs only slightly better than random guessing [54]. A decision tree is a ML algorithm partitioning the input data by binary splits into various branches. An overview on decision trees can be found in [55]. By combining multiple *weak learners* one can build a single *strong learner*, i.e. well performing model. The term structured data is used whenever the inputs of the model can be listed in tables, such as numbers or single words. Unstructured data on the other hand are inputs like images or time series. Ensembling of models splits basically into bagging and boosting [54].

Bagging is the combination of multiple independent *weak learners*. Each model is trained independently of each other usually on a sub-sample of the data and the output of all models is then combined by averaging over the outputs. A popular example of this technique is the combination of multiple decision trees to a so-called *random forest* [56].

Boosting on the other hand is the combination of multiple *weak learners* that are not independent of each other. The learners are generated sequentially and added to the model, so that later learners can correct mistakes previous learners did [54]. Again decision trees are commonly used as weak learners in boosting algorithms.

A generalization of the boosting technique is *gradient boosting* which allows the optimization of an arbitrary loss function. A description of the XGBoost algorithm for gradient boosting which is used in this work can be found in [57].

4.2 EXO-200 Analysis

In the last sections general machine learning features and techniques were outlined. In this work a ML discriminator is developed to separate signal events of the $2\nu\beta\beta$-decay of ^{136}Xe to an excited state from background events. Several input parameters, allowing a good distinction between signal and background events have to be found and analyzed.

The EXO-200 collaboration used a boosted decision tree (BDT) for this task based on six input parameters [14]. Histograms of these parameters for signal and background events can be found in figure 4.3. The first parameter used is the deposited energy of an interaction in the detector. This parameter is a commonly used one in physics and is also included in this work. An analysis of the signal and background energy distributions and a comparison of simulated with measured calibration data can be found in section 4.3.3.

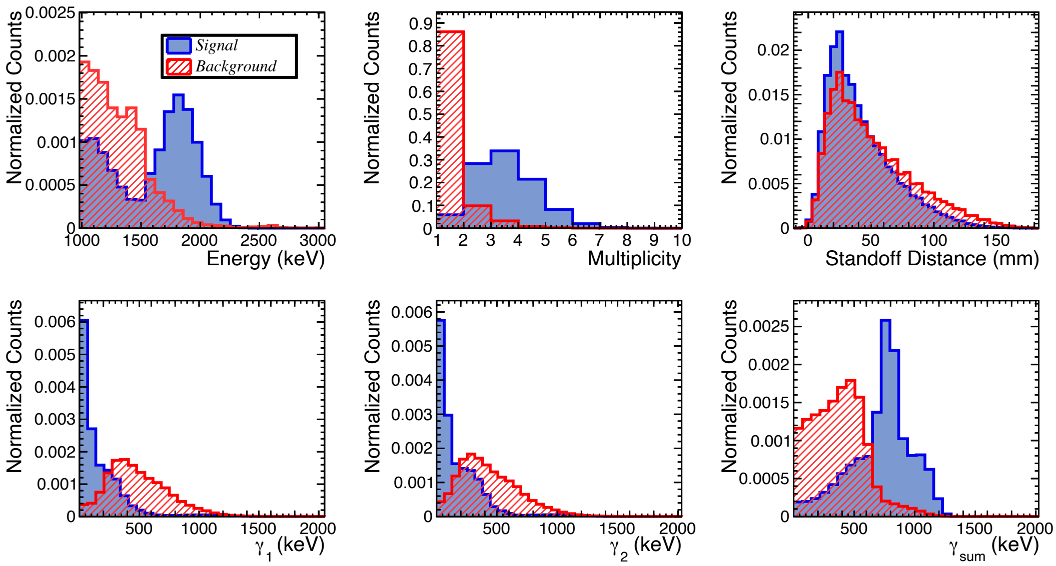


Figure 4.3: Normalized histograms of the discriminator input parameters (from the top left to the bottom right panel) energy, multiplicity, standoff distance, γ_1, γ_2 and γ_{sum} used in the EXO-200 analysis. Distributions obtained from simulated signal events are given in blue and distributions obtained from background simulations in red. The plot is taken from [14].

The second parameter used by EXO-200 is the multiplicity meaning the number of spatial separated energy depositions in the detector. One finds that in case of the EXO detector, background events are mostly single site events with only one energy deposition whereas the signal of the $\nu\beta\beta^*$-decay tends towards higher multiplicities. An analysis of the multiplicity in this work is given in section 4.3.1. The third parameter used in the EXO-200 analysis is the so-called standoff distance. This parameter is the minimal distance of an event to the surrounding walls or grids of the detector. One can expect that background events in most cases are created by external sources while the excited state decays are distributed homogeneously in the detector. An analysis of the standoff distances in this work can be found in section 4.3.4. The next three parameters used are called γ_1, γ_2 and γ_{sum}. These values are the minimal energy difference of one of the energy depositions in the detector to the energies of the two emitted γ-rays and the sum of both. In this work, modified versions of these parameters are investigated in section 4.3.6.

4.3 Discrimination Space

As outlined in the previous section, six input parameters are used for the ML discriminator in the EXO-200 analysis. The combination of the parameters creates a multi-dimensional discrimination space. In order to evaluate which parameters allow a good signal-background discrimination in the XENON1T data, measured background and simulated $\nu\beta\beta^*$-decay events of ^{136}Xe are analyzed in this section. In total seven different parameters will be investigated. Furthermore the simulations are validated by a comparison of simulated and measured calibration events. The first parameter of the discrimination space that is analyzed here is the scatter multiplicity.

4.3.1 Multiplicity

The scatter multiplicity is the first parameter used in this work for the signal-background discrimination. It is defined as the number of S2s in one event. Events with only one S2, now called single-site events, are more frequent at lower energies since in case of γ-radiation the total crosssection is dominated by the photoelectric effect. Events with higher multiplicity dominate the energy spectrum for $E > 400\,\mathrm{keV}$ [39] due to the increasing crosssections of the Compton effect and pair production. These interaction mechanisms lead to multiple spatial separated energy depositions and so to multiple S2s.

These events with more than one S2 can either be created by multiple energy depositions of the same particle like multiple Compton scatters of a single γ-ray, called multi-scatter, or by energy depositions of more than one particle called multi-site events. Multi-site events can either be random coincidences or originate from the same source. While the random coincidence rate is expected to be low due to the low background of XENON1T, multiple particles are emitted by multi staged decays like β^--decays to excited states followed by subsequent deexitation to the ground state. During the $2\nu\beta\beta$-decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba two electrons and two γ-particles with energies of 760.5 keV and 818.5 keV are emitted. Accordingly, one can expect a higher multiplicity for these signal events compared to background events. Depending on the energy of the emitted γ-radiation it is possible that multiple Compton scatters occur in a multi-site event.

In order to validate that the simulations reproduce the multiplicity found in the TPC, this parameter is compared for simulated and measured ^{212}Pb and ^{60}Co decays. A histogram of the multiplicities for ^{212}Pb is shown in figure 4.4. ^{212}Pb can produce multi-scatter events from the Compton scatter of the 238.6 keV and 415.3 keV γ or multi-site events from separated energy depositions of the γ and the electron emitted from the β^--decay. Most of the ^{212}Pb events are single-site events, but about 25 % of the events contain more than one S2 signal. To evaluate where both histograms

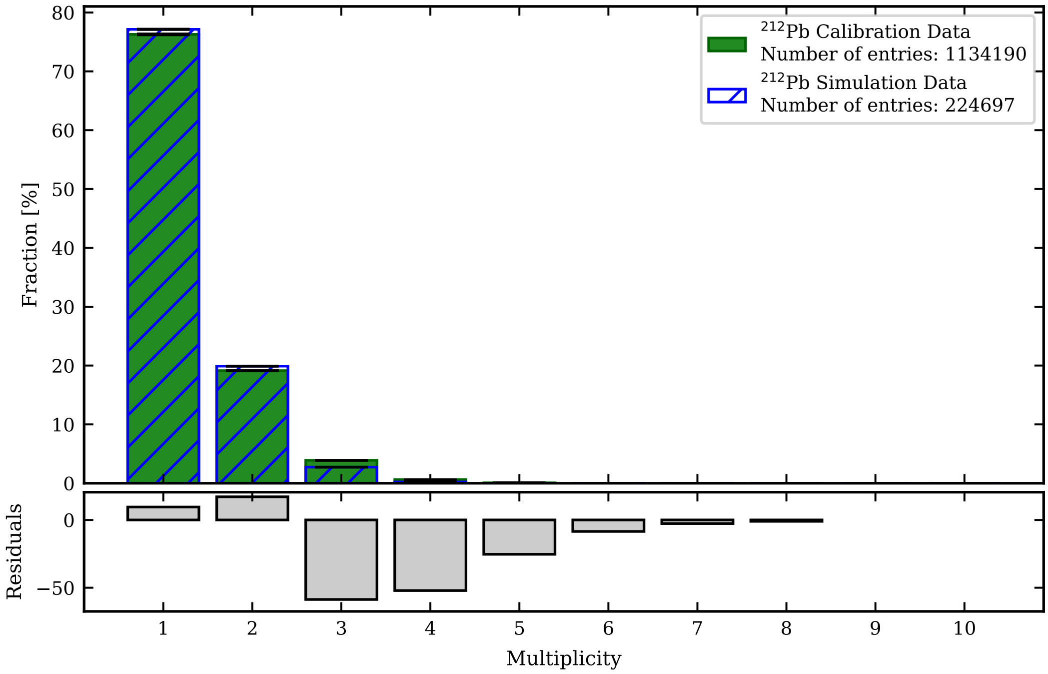


Figure 4.4: Normalized histograms of the multiplicity for simulated ^{212}Pb decays in blue and measured calibration data in green. Residuals are given in the bottom panel of the figure drawn in gray. The multiplicity is evaluated up to a maximum of ten S2s. The differences can be seen owed to high statistics in simulated and measured data. Error-bars are drawn in black but only barely visible due to the high statistics of both datasets.

differ, residuals are given as well. The residuals r are calculated as

$$r = \frac{y_{\mathrm{MC}} - y_{\mathrm{TPC}}}{\sqrt{\Delta y_{\mathrm{MC}}^2 + \Delta y_{\mathrm{TPC}}^2}}. \tag{4.2}$$

Here, y_{MC} are the bin entries of the simulated events and y_{TPC} are the bin entries of the events measured with the TPC. Δy_{MC} and Δy_{TPC} are the corresponding uncertainties. The calculated difference is normalized on the quadratic sum of the symmetric Poisson uncertainties. A negative residual shows that the simulation produces a lower number of events in a particular bin compared to the expectation from the calibration data and vice versa for a positive residual.

The mean multiplicity of the simulated ^{212}Pb events is 1.26 with a standard deviation of 0.51. The measured calibration data have a mean multiplicity of 1.29 with a standard deviation of 0.58 so the MC simulations tends to smaller multiplicities. Especially, the frequency for events with three and four S2s differs between simulation and measured data. $(2.74 \pm 0.04)\,\%$ of the simulated events have a multiplicity of three S2s whereas $(3.90 \pm 0.02)\,\%$ of the measured calibration data show the same multiplicity. This deviation, showing up as significant residuals, could be caused by several effects: One possibility is that the selection of ^{212}Pb events outlined in section 3.3.2 fails to remove all events from the calibration data which are not ^{212}Pb decay events. Since only ^{212}Pb decays are simulated, events from the ^{220}Rn decay chain or events from the background of XENON1T could pass the selections and contribute over-proportionally to the events with high multiplicity leading to deviations between the simulated and measured events. Beside the β^--decay of ^{212}Pb one could find events from α-decays of ^{220}Rn, ^{216}Po, ^{212}Bi or ^{212}Po or β^--events from ^{212}Bi or ^{208}Tl in the calibration data. Due to big differences in the charge and light yields of energy depositions by α-particles compared to energy depositions of β or γ-particles, it is unlikely that α-events can pass the selections. The β^--decay of ^{212}Bi is shortly followed by the α-decay of ^{212}Po so that it is unlikely that these events can pass the selections. ^{208}Tl decays in all cases to an excited state of ^{208}Pb[13] and in addition to the electron, a γ with high energies is emitted. When the γ deposits the energy in the fiducial volume this event would be removed by the energy cut and an energy deposition in the outer detector volume would cause that the event is removed by the fiducial volume cut. In principle ^{208}Tl decay events could pass all selections if the γ escapes the detector with energy depositions in the fiducial volume and without energy deposition in the outer detector region. Beside events originating from the ^{220}Rn decay chain, the measured data also includes events of the XENON1T background without calibration source. A different explanation would be the event generation and reconstruction in the simulations. Several simulations carried out in the development phase of the quanta generation software (cf. section 3.1.3) showed that the performed spatial clustering did not affect the multiplicity so that the differences could show up during the waveform generation or reconstruction with PAX.

A comparison between the multiplicity of background events measured with the TPC and the simulated signal events of the excited state decay of ^{136}Xe is provided in figure 4.5. Like expected, most of the simulated events have a high multiplicity with a mean of 4.4 ± 1.60 separated S2s. The background data tends towards lower multiplicities with a mean of 2.14 ± 1.29 S2s. The standard deviation is given here as uncertainty. Even if there are deviations in percent level for the multiplicity between simulations and experiment, the differences between signal and background events is far larger.Thus this parameter should allow a good differentiation between signal and background events.

To reduce the number of background events while keeping most of the signal, a cut on the multiplicity is applied for the following analysis: single-site events are cut and only events with multiplicities greater one are kept. This cut removes 40.0 % of the background events while keeping 98.4 % of the signal.

In order to evaluate how the multiplicity cut works on calibration data, the multiplicity of simulated and measured ^{60}Co events is shown for the 1173.2 keV line in figure 4.6 and for the higher energetic 1332.5 keV line in figure A.4 in the appendix. ^{60}Co is used

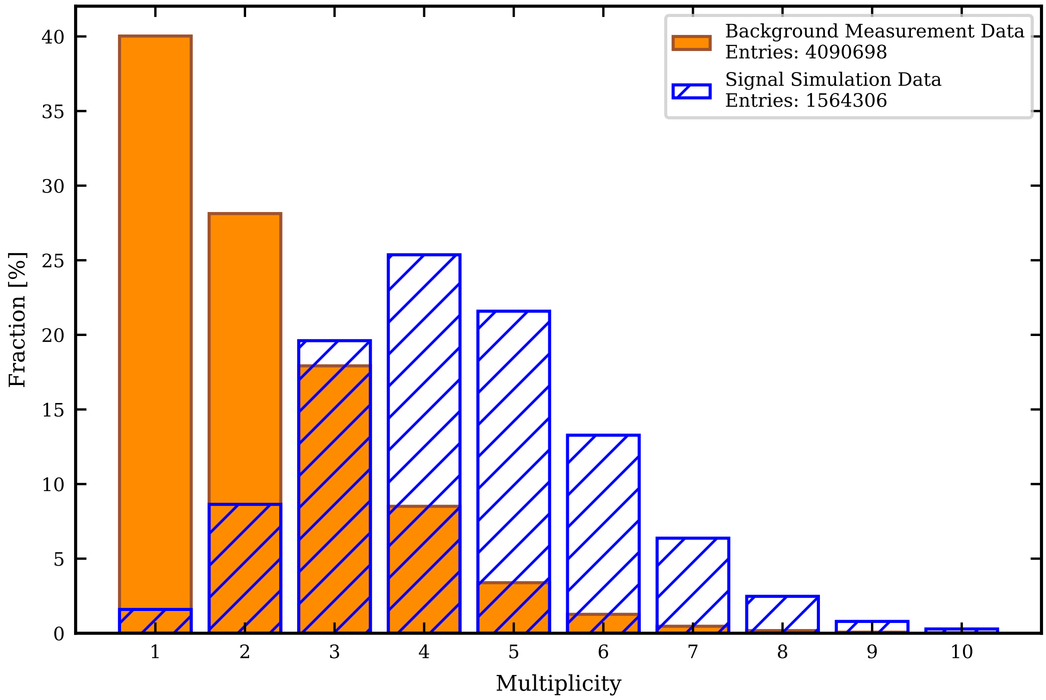


Figure 4.5: Normalized histograms for the multiplicity of simulated ^{136}Xe signal events in blue and measured background events plotted in orange. Background events tend towards a low multiplicity while the signal events are mostly multi-site events.

here as calibration source, since the the γ-rays provide a sufficient amount of events with more than one S2. One can see, that like for ^{212}Pb both distributions look similar, but again the simulation tends towards a slightly lower multiplicity. Simulation of external sources like ^{60}Co is computationally expensive since most of the simulated decay do not create a signal in the fiducial volume. Therefore, the statistics for the ^{60}Co simulations are low compared to the measured data and other simulations used in this work. Only about 1.1×10^4 simulated ^{60}Co events of the 1173.2 keV line are analyzed here whereas the ^{212}Pb simulations contain about 2.2×10^5 events. Additionally, the measured ^{60}Co events are not taken from dedicated calibration runs and a large number of background events can be expected to pass the selections.

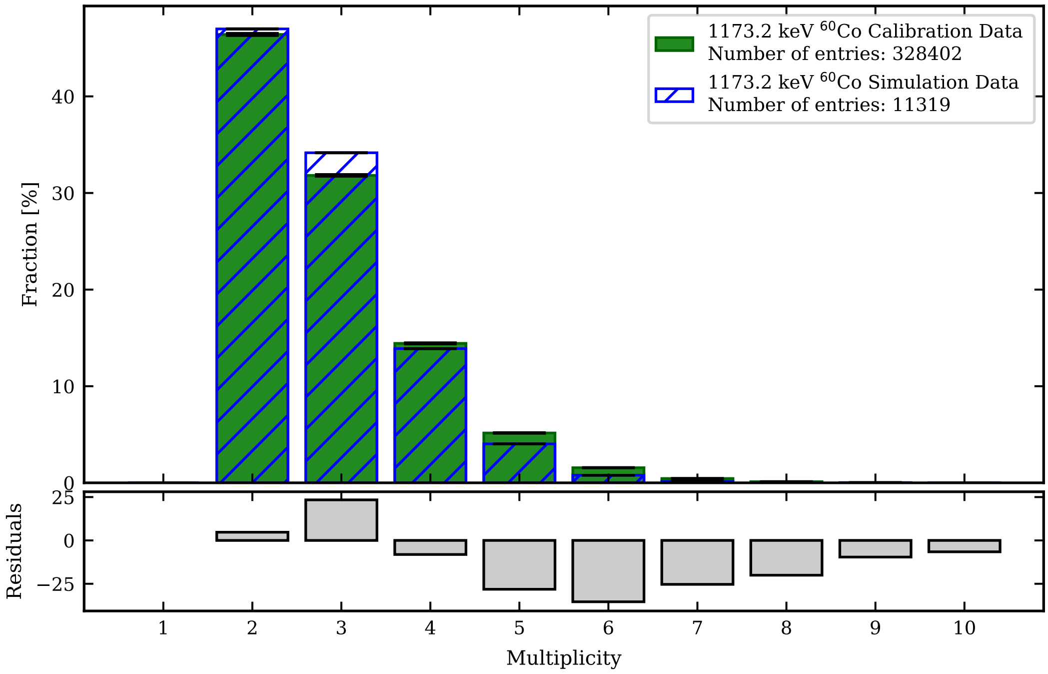


Figure 4.6: Normalized histograms of the multiplicity for simulated 1173.2 keV ^{60}Co γ-rays in the fiducial volume in blue and calibration data in green. A multiplicity cut is applied to remove single-site events. The simulated data tends towards a lower number of S2s. The statistics of the simulation are limited by the computational time needed for external γ-source simulations. Error-bars are drawn in black but barely visible. Residuals are given in the bottom panel of the figure in gray.

4.3.2 Corrected S1 and S2 Areas

The corrected areas of the S1 and the S2 signal seen on the bottom PMT array are used to calculate the energy deposited in the detector and to calculate other discrimination space parameters. Thus, a realistic simulation of cS1 and $cS2_b$ is necessary for the presented analysis.

The MC simulations are compared to measured data for the selected mono energetic xenon γ-lines, the continuous spectrum of ^{212}Pb and two γ-lines from ^{60}Co. The sources and event selections used are outlined in section 3.3. Normalized histograms were calculated for simulated and measured data using the same range and number of bins together with projections on $cS2_b$ and cS1. To directly compare the two-dimensional data in one plot, contour lines are drawn in the same plot for simulated and measured data. To reduce binning effects from varying number of entries in bins close to each other, the data used for the contour lines is smoothed with a Gaussian filter. This procedure keeps the overall shape of the data while reducing complexity of the contour lines allowing an easier comparison. The projections are however calculated from unfiltered histograms. Figure A.2 shows how the filtering simplifies the contours by comparing a filtered and unfiltered histogram of the same data.

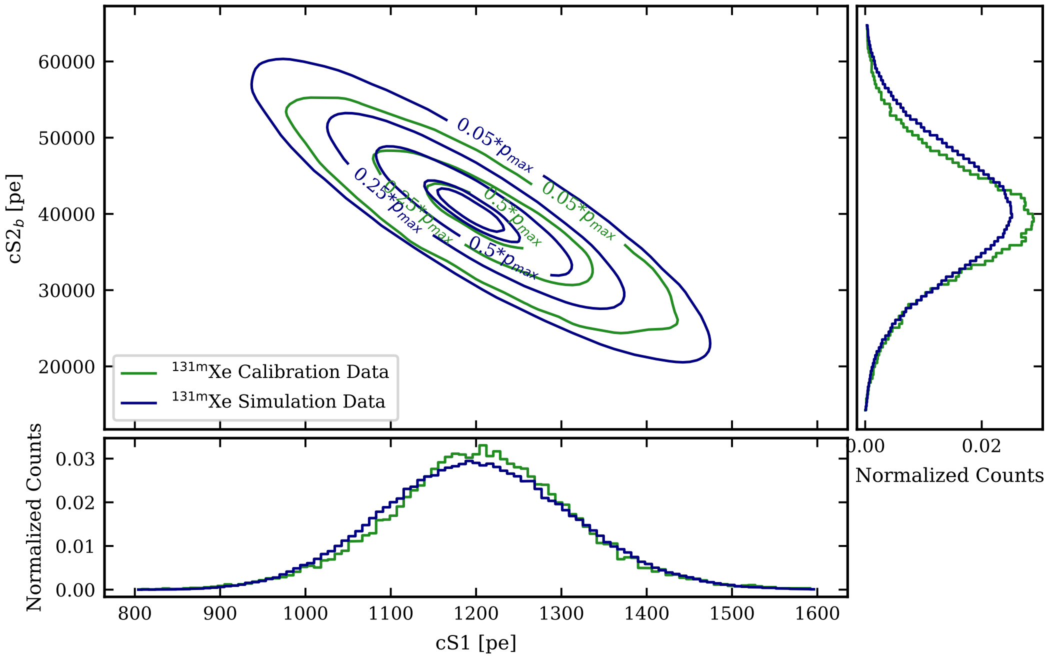


Figure 4.7: Contour plot for cS1 and cS2$_b$ of ^{131m}Xe events with the simulated data in blue and calibration data in green. Projections on cS1 and cS2$_b$ are given as well. The contour lines are calculated from normalized histograms smoothed with a Gaussian filter while the filter is not applied to the projections. A mono-energetic source like ^{131m}Xe shows up as a tilted two-dimensional Gaussian in the cS1 vs. cS2$_b$ space and as one dimensional Gaussian peaks in the projections.Thus, the position and width can be compared easier in contrast to continuous sources like ^{212}Pb.

The lowest mono-energetic source with 163.9 keV in the comparison, ^{131m}Xe, is shown in figure 4.7. The plot for ^{129m}Xe can be found in figure A.3 in the appendix. The positions of the cS1 mean agrees within 0.2 % and the the mean cS2$_b$ position within 2.7 %. The width of the simulated data is about 11 % larger than the measured data. As described in section 3.1.3, the quanta fluctuations were increased manually to best match the data but a compromise had to be found in order to match all benchmark sources.

The comparison for the continuous energy ^{212}Pb decays is shown in figure 4.8. The overall shapes of the projected data look similar, but mean cS1 of the simulated ^{212}Pb events is 2.8 % larger than the measured cS1 mean. The mean of the simulated cS2$_b$ signal is 7.9 % larger than the measured mean. This discrepancy both in the cS1 and cS2$_b$ leads to an simulated energy spectrum shifted to higher values compared to the measured calibration data.

As described in the previous chapter, a multiplicity cut is used in the analysis of the signal and the background events. Since the xenon lines generally produce single-site events due to the low energy of the emitted conversion electrons, a post-cut comparison

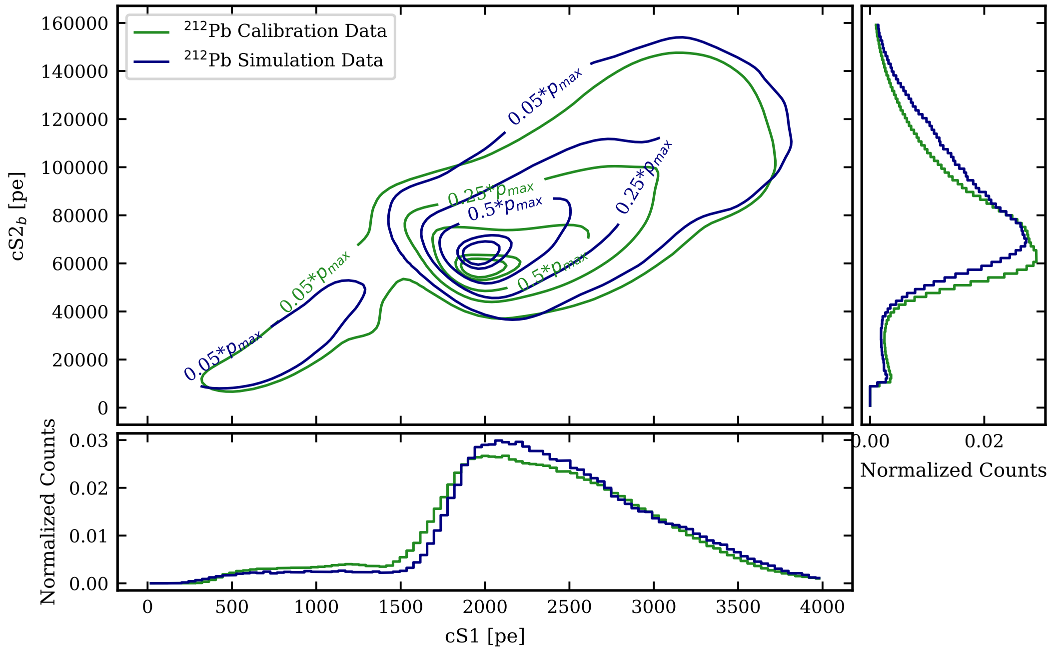


Figure 4.8: Contour plot for cS1 and $cS2_b$ of ^{212}Pb events. The simulated data is given in blue and calibration data in green together with projections on cS1 and $cS2_b$. The contours, obtained from normalized histograms were smoothed with a Gaussian filter. No multiplicity cut is applied. Contour lines are drawn for multiple fractions of the bin with the largest normalized count rate for each histogram.

is only performed for ^{212}Pb and ^{60}Co. The ^{212}Pb events after the multiplicity cut are shown in figure 4.9.

After the cut, the mean of the simulated cS1 signals is 2.0 % larger than the measured values and the mean of the simulated $cS2_b$ signals is now 2.8 % smaller than the measured $cS2_b$ signals in the TPC. With this better agreement compared to the uncut data, one can expect a good agreement of the reconstructed energies for simulated and measured ^{212}Pb events.

The present discrepancy when including the single-site events could possibly be explained by an imperfect event selection like outlined in the previous section and the emission model used for the quanta generation not ready for detector-specific clustering and microphysics at high energy.

The highest energy data in this comparison come from the two γ-lines of ^{60}Co at 1173.2 keV and 1332.5 keV. As outlined before, γ-rays from ^{60}Co produce mostly multi scatter events and so a comparison for events with applied multiplicity cut is shown in figure 4.10. This data contains both γ-lines in the simulated and measured data as well as background events passing the selections for the measured TPC data.

One finds that the simulated ^{60}Co have an approximately 9.6 % larger mean cS1 area and an approximately 5.9 % smaller $cS2_b$ area compared to the measured calibration

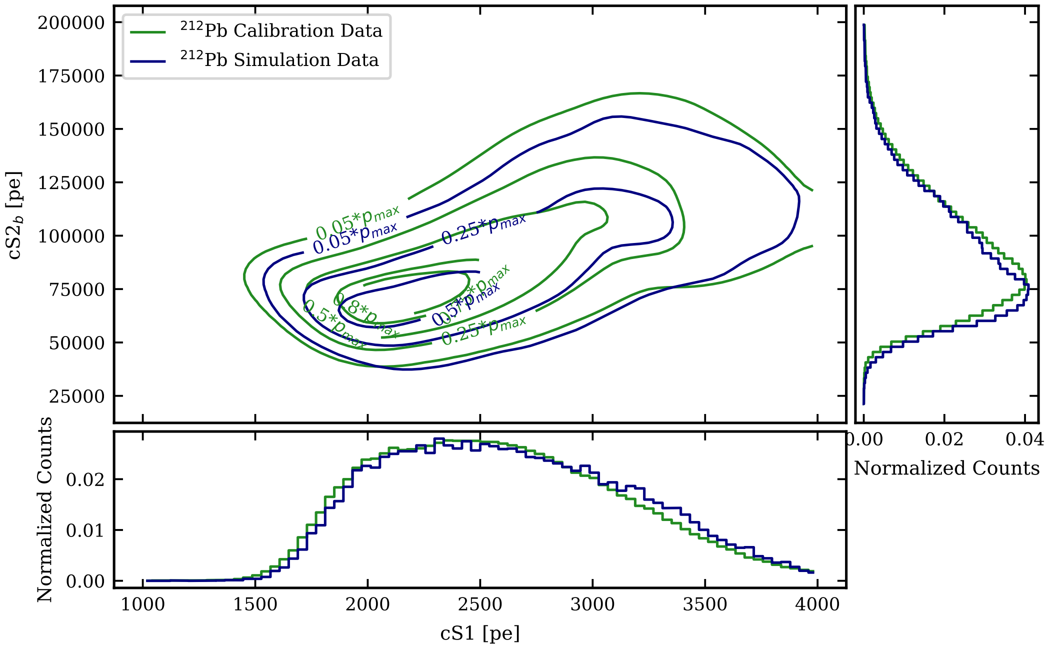


Figure 4.9: Contour plot for cS1 and cS2$_b$ of ^{212}Pb events with applied multiplicity cut for simulated data in blue and calibration data in green together with projections on cS1 and cS2$_b$. The Contours are obtained from normalized histograms smoothed with a Gaussian filter. Contour lines are drawn for multiple fractions of the bin with the largest normalized count rate for each histogram.

data. This result is compatible with the deviations between the NEST model for γ-radiation and the XENON1T data shown in figure 3.6. For future analysis the emission models need further tuning for higher energy to match the observed signals in the TPC with the MC simulations.

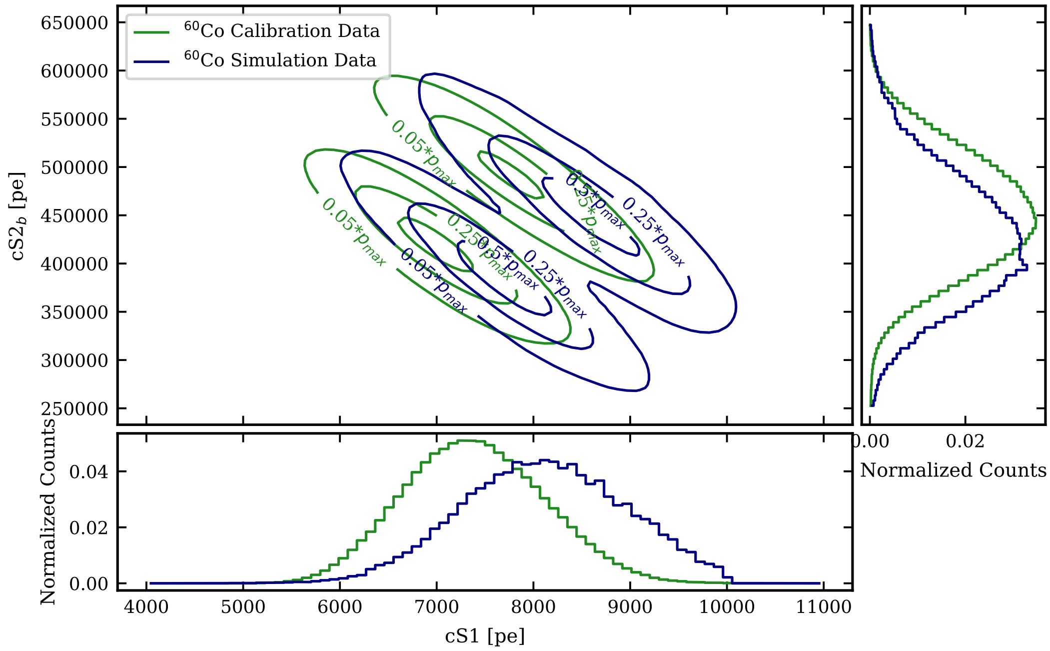


Figure 4.10: Contour plot for cS1 and cS2$_\mathrm{b}$ of ^{60}Co events with applied multiplicity cut for simulated data in blue and calibration data in green. The Contours are obtained from normalized histograms smoothed with a Gaussian Filter. Projections on cS1 and cS2$_\mathrm{b}$ are given as well. The plot show both the 1173.2 keV and 1332.5 keV γ-lines. Contour lines are drawn for multiple fractions of the bin with the largest normalized count rate for each histogram.

4.3.3 Energy

The second parameter that is used for the signal-background discrimination is the total energy deposition of a particle interaction. The energy can be reconstructed from cS1 and cS2$_\mathrm{b}$ using

$$E = w \cdot \left(\frac{\mathrm{cS1}}{g_1(z)} + \frac{\mathrm{cS2_b}}{g_2(z)} \right) \tag{4.3}$$

with a z-dependent photon detection efficiency g_1, charge amplification g_2 and $w = 13.8\,\mathrm{eV}$ [23], the mean energy needed to produce a charge or light quanta. g_1 and g_2 are taken from [39] and are valid for single and multi-site events. The usage of the cS2$_\mathrm{b}$ reduces saturation effects and ensures a position-independent behavior of the PMT array due to a more uniform S2 light-yield. In case of multi-site events with more than one S2, the sum of the corrected S2s is used.

To validate if the energy reconstructed from the simulated events is compatible with the energies found in the measured data, the simulations are compared to the calibration sources introduced in section 3.3. Both xenon lines emit mono-energetic γ-radiation, so the energy distribution has the shape of a Gaussian peak and one can directly compare the mean and width of the MC simulation with the measured calibration data. To achieve the best results, a fit with a Gaussian function is performed

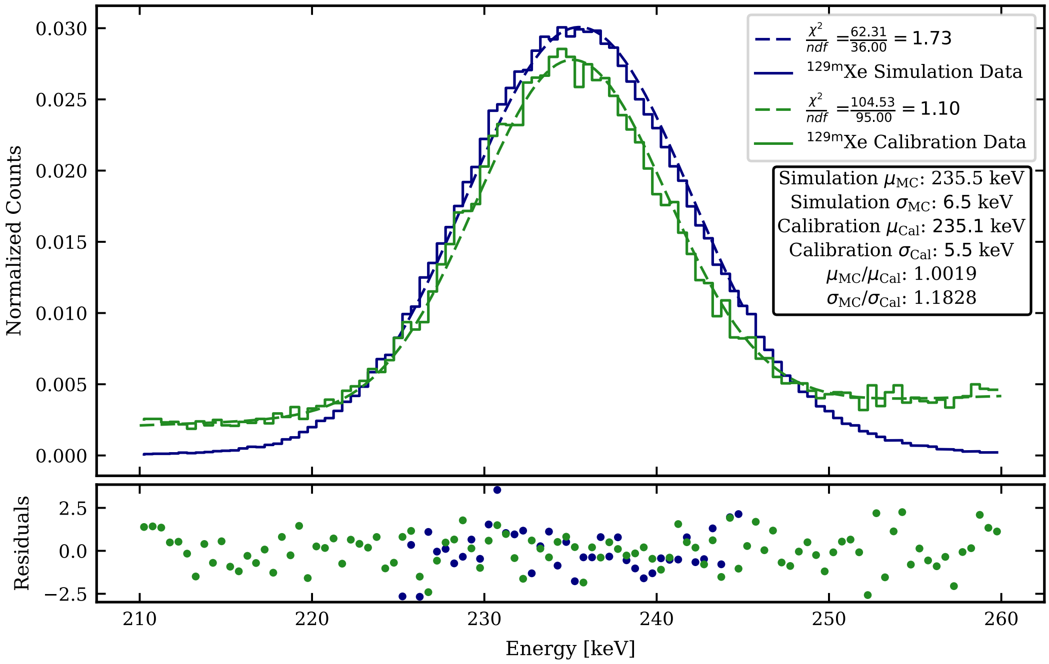


Figure 4.11: Reconstructed energy spectra of simulated (blue) and measured (green) ^{129m}Xe 236.1 keV γ-line in the energy range from 210 keV to 260 keV. Error-bars are not drawn to reduce the complexity of the plot and the residuals of the fit are given in the bottom panel of the figure. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. To determine the mean and the width a Gaussian is fitted to both spectra and a linear term is added to the calibration data fit to account for the background. The fit is performed on the histograms and afterwards normalized to account for different numbers of events for an easier comparison. The fit of the calibration data is performed over the whole energy range given in the plot and the fit of the simulated data is restricted to the central part of the Gaussian to ensure a sufficient number of entries in each bin and allow the usage of symmetric uncertainties.

both for the simulated and measured data for ^{129m}Xe shown in figure 4.11 and for ^{131m}Xe in figure A.5 in the appendix. The fits are carried out as χ^2 minimization (cf. section 3.1.1).To account for background events present in the measured data, a constant term is added to the ^{131m}Xe TPC data fit and a linear term to the ^{129m}Xe TPC data fit. The mean positions of the Gaussian fits agree within 0.6 keV corresponding to deviations below 1 %. The widths agree within 0.9 keV.

The energy distributions of ^{212}Pb cannot be compared in terms of mean and standard deviations of a Gaussian peak since it is not a mono-energetic source, but a β^--decay going to different states of ^{212}Bi. Histograms of the reconstructed energy for the simulation and the selected TPC calibration data are shown in figure 4.12.

Both spectra consist of a lower energy part ($E < 238.6$ keV) which is caused by the β^--decay to the ground state of ^{212}Bi. The decay with the highest branching ratio shows up as β^--spectrum starting at 238.6 keV due to the energy depositions by the emitted

γ-radiation. The third channel includes the emission of an 415.3 keV γ-ray and can be seen as "bump" at the corresponding energy. These features of the spectrum can be found in the measured as well as in the simulated data. One can see that the

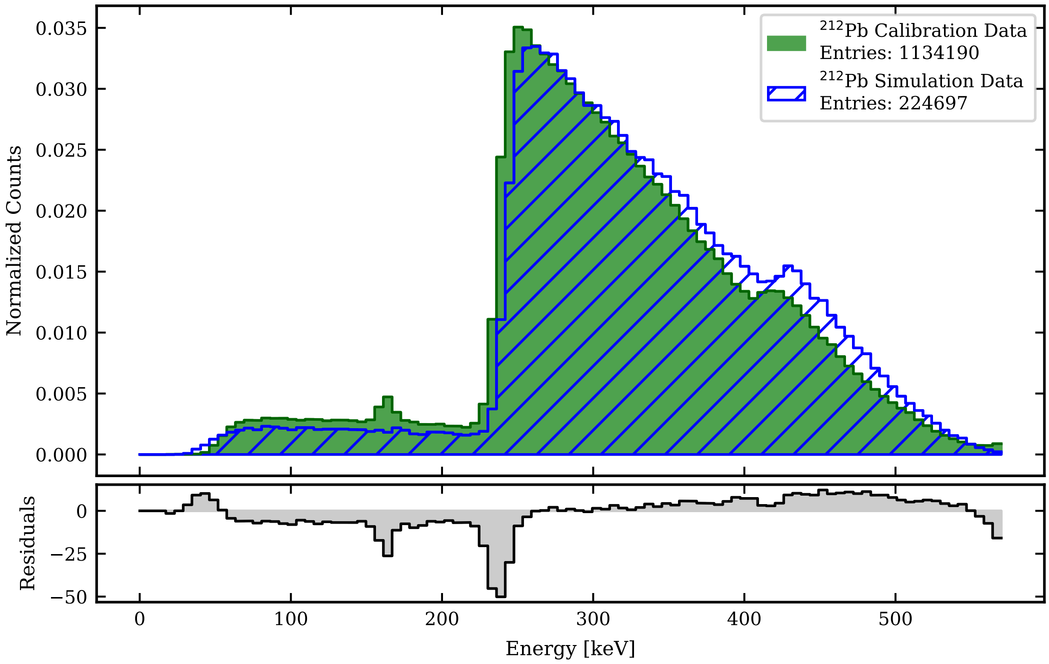


Figure 4.12: Comparison of the reconstructed normalized energy spectra for simulated and measured ^{212}Pb decays. The measured data is shown in green and the simulation in blue. To allow an inspection of the differences, residuals calculated with equation (4.2) are given as well. Error-bars are not drawn to reduce the complexity of the plot. Only small relative uncertainties are present for each bin due to the high statistics of both datasets.

measured ^{212}Pb data show a peak at about 160 keV not present in the simulations. The residuals show a negative peak at the same energy. This peak can be attributed to the ^{131m}Xe γ-line. Since this peak is visible in the measured data it is reasonable to assume that the 236.4 keV peak of ^{129m}Xe is present. The corresponding energy is close to the energy of the γ-ray emitted in decay channel with the highest branching ratio and one cannot see a clearly separated peak. Nevertheless, the spectra show the biggest deviation around this sharp rise in the counts visible as negative peak in the residuals. Together with smaller deviations around 415 keV, the simulated energy spectrum is shifted towards higher energies compared to the measured data. This effect is already expected from the differences in cS1 and $cS2_b$ in section 4.3.2.

The ^{212}Pb energy spectrum after the multiplicity cut is shown in figure 4.13. Most of the events with energies below about 240 keV are removed both for simulated and measured data. The negative residuals in this energy region indicate that background events in the measured data can pass the cut whereas most of the simulated events in this energy region are removed. At around 240 keV, the residuals show a peak-like

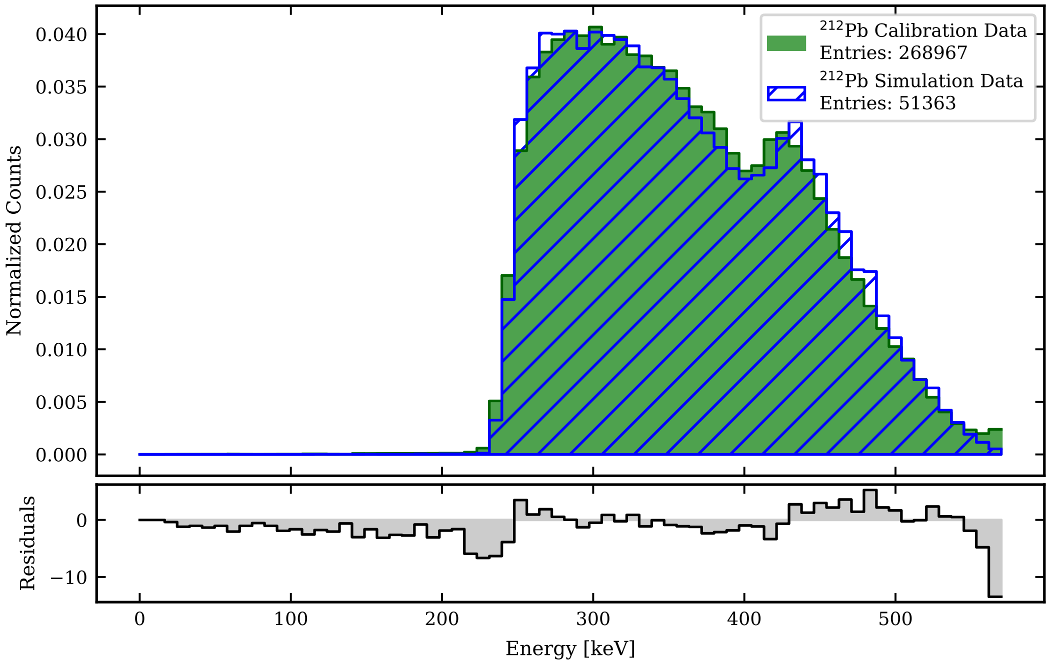


Figure 4.13: Comparison of the reconstructed normalized energy spectra after the multiplicity cut for simulated (blue) and measured (green) ^{212}Pb decays. Most of the events below about 240 keV are removed and only events going to excited states of ^{212}Bi remain. Error-bars are not drawn to reduce the complexity of the plot. Only small relative uncertainties are present for each bin due to the high statistics of both datasets.

structure. This lack of simulated events could be caused by ^{129m}Xe events passing the cuts. For higher energies only ^{212}Pb events remain where a γ is emitted in addition to the β-electron. The simulated ^{212}Pb spectrum decreases smoothly to zero counts at around 570 keV, the Q-value of the decay. The measured data show a not vanishing count rate at this energy which shows up as negative residuals and indicate that a population of background events is passing all data selection cuts.

Figure 4.14 shows the energy spectrum of the XENON1T background data and $\nu\beta\beta^*$-decay simulations of ^{136}Xe in the range from 0 to 3 MeV. The simulations show a double β^--spectrum starting at about 1.58 MeV which corresponds to the sum of the energies of the two emitted γ-rays. This population ranging from about 1.58 MeV to about 2.5 MeV contains signals from all emitted particles (except the neutrinos). A minor component of the spectrum can be found below 1.58 MeV also in the shape of a double β-spectrum. These are events where one of the two γ-rays can escape the TPC without an energy deposition in the outer part of the detector that would lead to a veto of this event by the fiducial volume cut (cf. section 3.2.2). The background spectrum of the TPC consists of a continuum with several mono energetic peaks. The continuous contribution arises from the $2\nu\beta\beta$-decay of ^{136}Xe to the ground state of ^{136}Ba and the Compton spectrum of high energy γ-rays from detector materials. An

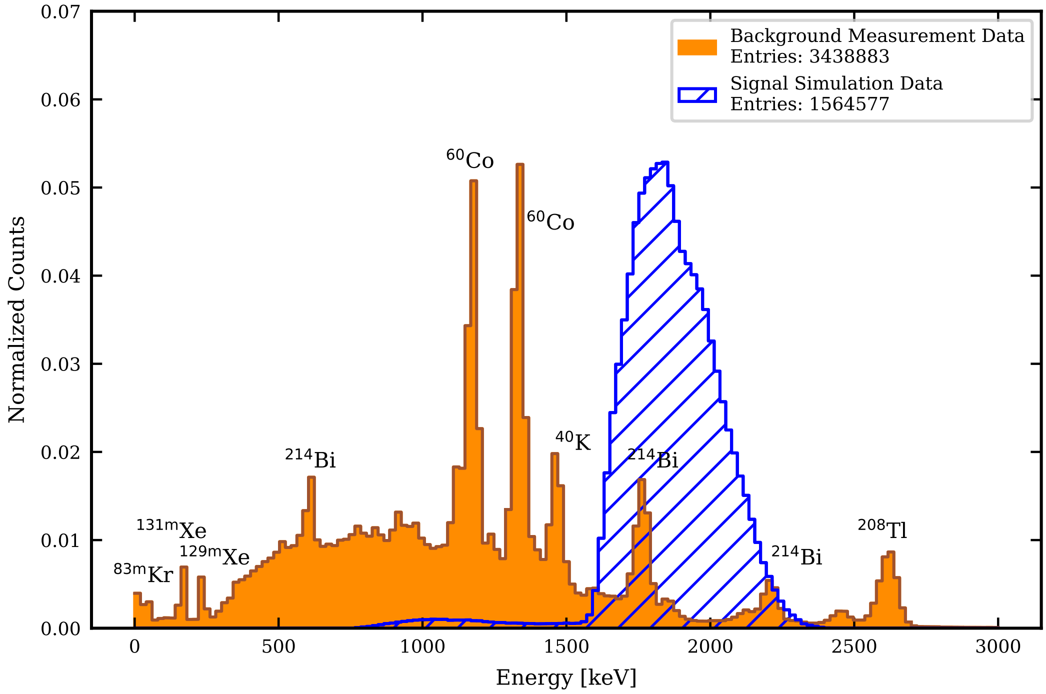


Figure 4.14: Normalized energy spectra of the simulated $\nu\beta\beta^*$-decays of ^{136}Xe shown in blue and the measured background spectrum of SR 1 of XENON1T shown in orange in the energy range from 0 to 3 MeV.

energy cut is applied to the data following the multiplicity cut outlined before in order to increase the signal to background ratio. Only events with $1.5\,\text{MeV} < E < 2.5\,\text{MeV}$ are used for the later analysis and events outside this range are cut. In combination with the multiplicity cut, 89.36 % of the background data is removed while 96.35 % of the signal events are kept. The resulting energy spectrum after cuts is shown in figure 4.15. The background is now dominated by two mono-energetic peaks from ^{214}Bi at 1.764 MeV and 2.204 MeV [13].

In this section, the total reconstructed energy of an event was validated between simulations and calibration data for ^{212}Pb and two xenon γ-lines. Afterwards, the energy spectrum of simulated signal events was compared to the measured background events. The validation of the energy with the sources used has some limitations. The energies of all sources is smaller than the energies deposited by the excited state decay of ^{136}Xe. Additionally, the different number and types of emitted particles has to be taken into account. During the decay of ^{212}Pb, an electron and a γ-ray are emitted whereas two electrons and two γ-rays are emitted in signal decay. Since it is comparably easy to validate simulations of mono-energetic γ-sources, also crosschecks with energy depositions by electrons have to be done. A source of mono-energetic electrons is not available so electrons from a β-decay are used in this work. ^{212}Pb was chosen as calibration source since it is the only available source with high statistics and a low background contamination and an electron γ-coincidence.

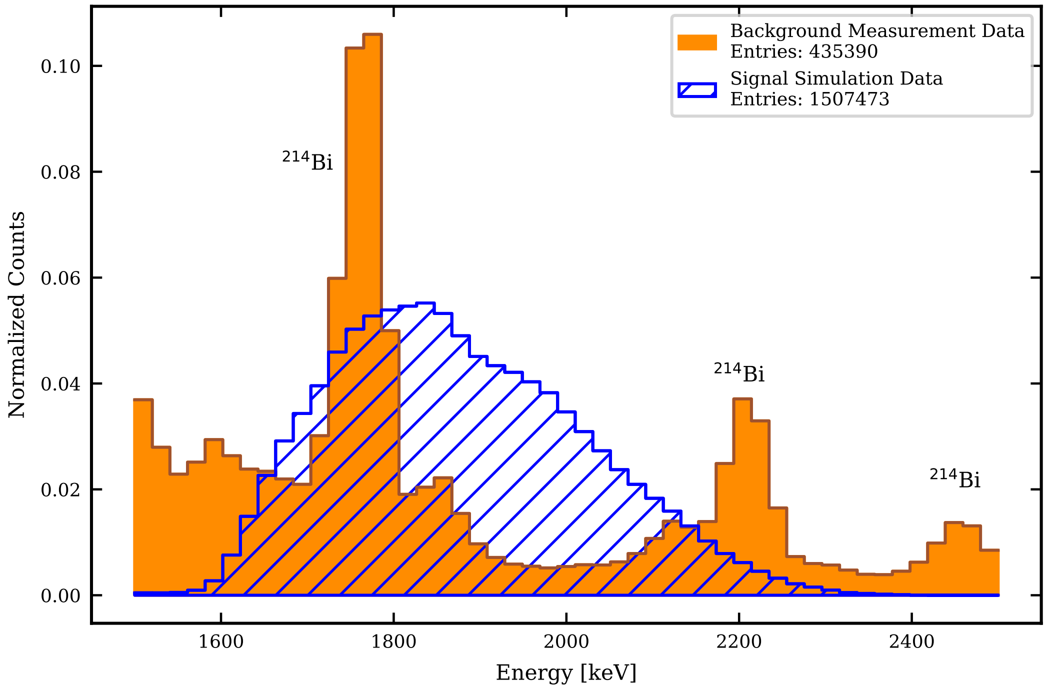


Figure 4.15: Normalized energy spectra of the simulated $\nu\beta\beta^*$-decays of ^{136}Xe shown in blue and the measured SR 1 background spectrum of the TPC shown in orange after multiplicity and energy cut in the range from 1.5 MeV to 3 MeV.

4.3.4 Standoff Distance

The minimal standoff distance is the minimal distance of an event to the TPC boundaries which are defined by the PTFE walls, the gate mesh on the top and the cathode mesh on the bottom of the detector. The minimal standoff distance on an event is calculated using

$$\text{min. standoff distance} = \min(d_i) \tag{4.4}$$

where d_i is the closest distance of energy deposition i to the TPC boundaries

$$d_i = \min(|r_i - r_{\text{TPC}}|, |z_i - z_{\text{Gate}}|, |z_i - z_{\text{Cathode}}|). \tag{4.5}$$

Here r_i is the radial coordinate of the energy deposition, r_{TPC} is the radius of the TPC, z_i is the depth of the interaction, z_{Gate} the position of the gate mesh and z_{Cathode} the position of the cathode.

This parameter should allow a differentiation between homogeneously distributed and external sources. Background events are mostly γ-rays from nuclear decays in the TPC materials that scatter inside the detector ant thus external sources. These events are less frequent in the innermost part of the detector due to xenon's self-shielding. ^{136}Xe is an internal source, so the events are distributed homogeneously in the detector.

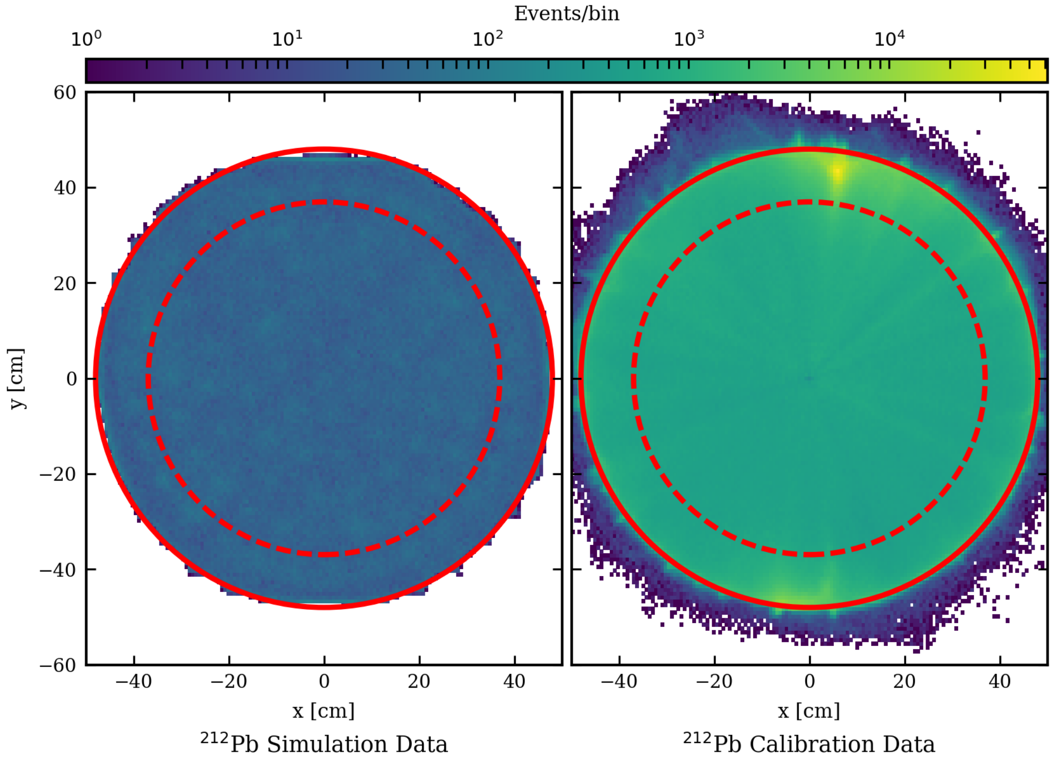


Figure 4.16: The x-y distribution of ^{212}Pb events in the detector for simulated events on the left side and measured data on the right side of the plot. The approximate dimensions of the TPC are shown as the solid red line and the fiducial volume as the dashed red line. It is not possible to measure events with a true position outside of the detector but a small fraction of events is reconstructed here. This is caused by the imperfect position reconstruction algorithm or the applied three dimensional field-distortion correction.

In order to validate the simulations, first simulated and measured ^{212}Pb events will be compared. ^{212}Pb is an internal source and should show a similar standoff distance distribution as the $\nu\beta\beta^*$-decay of ^{136}Xe. Figure 4.16 shows the x-y distribution of simulated and measured ^{212}Pb events. The corresponding z-r^2 distribution is given in figure 4.17. The TPC boundaries are shown as the solid red lines and the fiducial volume is given as the dashed red lines. Under ideal conditions the events should be homogeneously distributed in the whole TPC volume since ^{212}Pb is part of the ^{220}Rn decay chain which was injected as an internal calibration source and should mix homogeneously within the detector.

One can see that the calibration data shows a hot-spot at the top center of the x-y distribution and a slightly increased number of events on the opposite side of the TPC. These features are be caused by the ^{220}Rn inlet to the TPC. It is not possible to measure events with a true position outside of the detector but a small fraction of events is reconstructed here. This is caused by the imperfect position reconstruction algorithm or the applied three dimensional field-distortion correction.

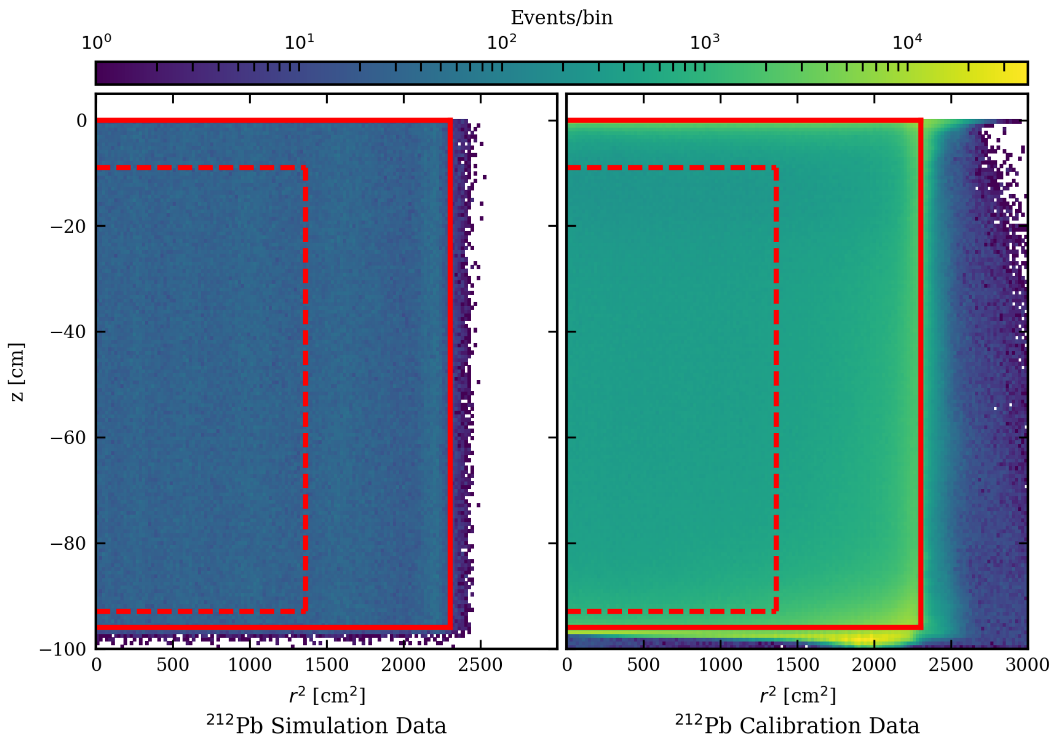


Figure 4.17: z-r^2 distribution of ^{212}Pb events in the detector for simulated events on the left side and measured data on the right side of the plot. The TPC walls are indicated as vertical solid red line and the cathode and gate mesh as solid horizontal lines. The fiducial volume is shown with dashed red lines.

In the z-r^2 distribution, an increased number of events in the measured data can be found at around $r^2 = 2000\,cm^2$ and $z < -90\,cm$. This in-homogeneity leaks, if at all present, only slightly into the fiducial volume where the measured event distribution looks homogeneous like expected from the source characteristics. The MC simulations do not show the hot-spot seen in the TPC, since the decays were simulated homogeneously distributed in the liquid xenon volume. A peak-like structure is present in the x-y projection which shows up as wave pattern in the z-r^2 space. This effect is most likely caused by the process of light distribution in the waveform simulator using PMT pattern maps. Since the effect is small, this should not cause problems in this analysis but has to be checked for the upcoming XENONnT simulations.

A comparison of the resulting standoff distances for ^{212}Pb calibration data and simulations is shown in figure 4.18. The shape of the distributions is defined by the fiducial volume cut used. Events with a low standoff distance are located more towards the outside of the TPC but still located in the 1T fiducial volume. The decrease in counts for distances above about 15 cm is purely caused by the decreasing volume with increasing distance to the TPC walls. Overall, the shapes of both distributions agree, but a oscillation-pattern is visible in the residuals. The largest deviations can

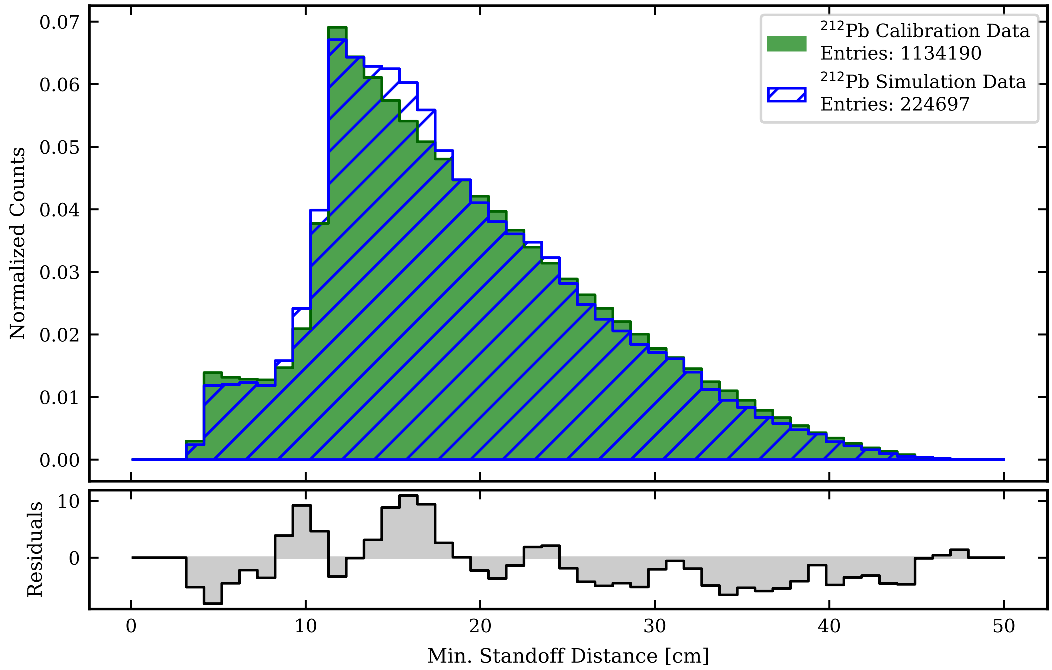


Figure 4.18: Normalized histograms of the minimal standoff distance for simulated ^{212}Pb events shown in blue and measured data in green. Error-bars are not drawn to reduce the complexity of the plot and residuals are given in the bottom panel. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. The shape of the distribution is defined by the fiducial volume cut, the detector geometry and the distribution of the events in the TPC. ^{212}Pb is an internal source and should show a similar standoff distance distribution as the excited state decay of ^{136}Xe.

be found between 10 and 15 cm. The observed differences could be caused by the described peak-like structure in the x-y projection.

The simulated and measured ^{212}Pb data show that the simulations are able to accurately reproduce the minimal standoff distance distribution in the TPC. In order to validate if the distributions of external sources are also reproduces accurately, the minimal standoff distances of simulated and measured ^{60}Co events will be compared. Figure figure 4.19 shows normalized histograms of simulated and measured events. As described earlier ^{60}Co decays mostly in the cryostat of the detector and the events selected are caused by γ-rays reaching the fiducial volume. This external source should show a different standoff distance distribution compared to the internal sources and should be similar to the distribution of the background events. One can see that the overall shape of the distributions is quite similar but differences are present at lower standoff distances. This difference is most likely caused by a slightly different z distribution of the events shown in figure A.6 in the appendix together with the x-y distribution in figure A.7.

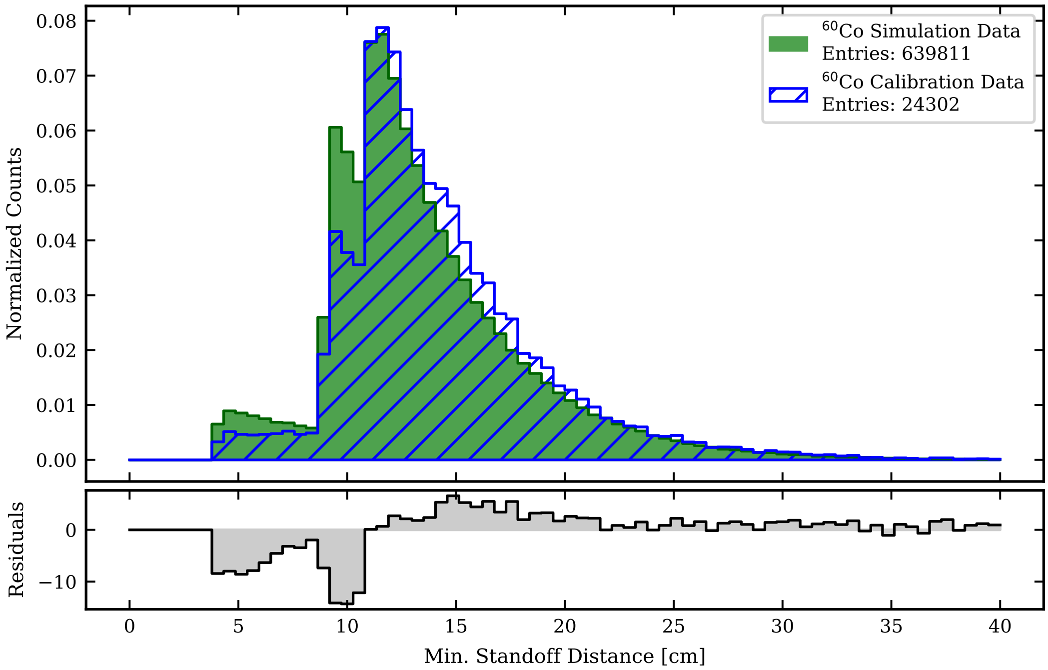


Figure 4.19: Comparison of the standoff distance distribution for simulated decays of ^{60}Co in blue and measured events in green. Error-bars are not drawn to reduce the complexity of the plot. ^{60}Co is an external source of γ-radiation leading to an accumulation of events at smaller standoff distances. The overall shape of the distributions shown here is similar to the shape of the background events shown in figure 4.20 confirming that most background events in the dateset are created by external sources.

The differences in z can be caused by the simulation settings used. For this analysis only ^{60}Co decays in the cryostat of the detector were simulated. Even though, this is the main source of ^{60}Co decays, one can also expect contributions from other detector components like the PMTs which are located at the top and bottom of the detector.

Figure 4.20 contrasts the distribution of the standoff distance for the background events from the TPC with the distances of the signal simulations of the ^{136}Xe decays. One can see that the standoff distance of the simulated signal look similar to the distribution found for ^{212}Pb since both sources are internal. The simulated signal events have a mean minimal standoff distance of 17.72 cm and a standard deviation of 7.00 cm.In contrast to the signal events ,background events tend towards smaller standoff distances with a mean of 14.03 cm and a standard deviation of 5.02 cm like it is expected for sources from the detector materials and shown for the external ^{60}Co source.

The analysis of the minimal standoff distance showed, that obtained distributions for signal events of the excited ^{135}Xe decay and ^{212}Pb calibration data are similar. This similarity is caused by the internal nature of both sources. The standoff distance distribution of background events is similar to the distributions observed for the ^{60}Co

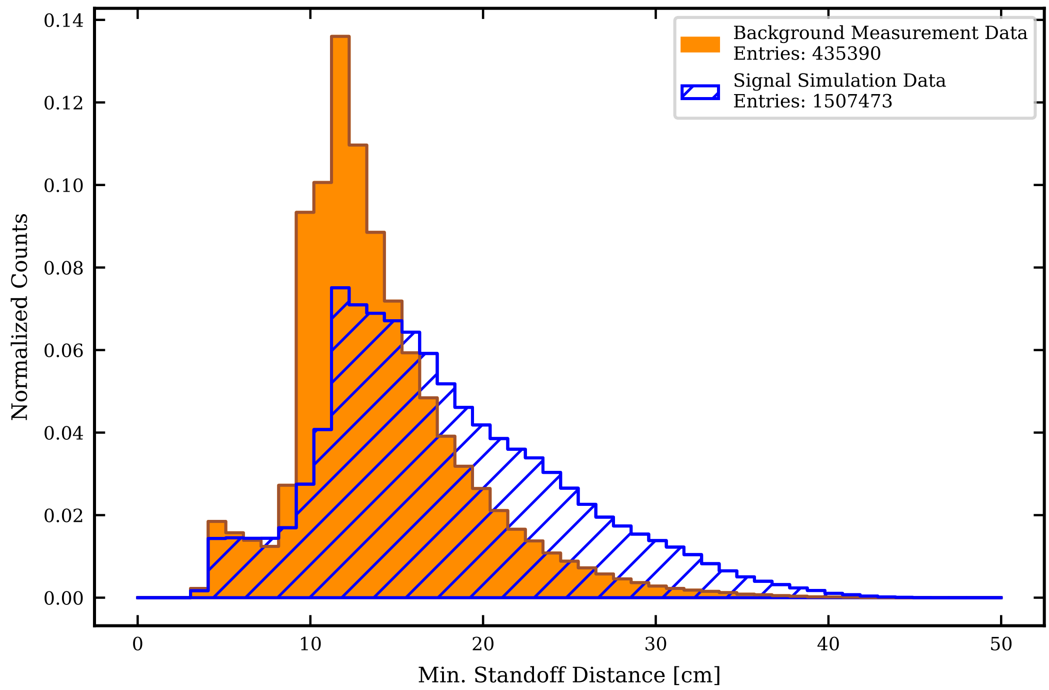


Figure 4.20: Comparison of the standoff distance distribution for simulated signal events in blue and measured background data in orange. The multiplicity and energy cuts outlined in section 4.3.1 and section 4.3.3 are applied to background and signal data. The shape of the signal distribution is similar to the distributions of measured and simulated ^{212}Pb events due to the similar internal origin of the events. The background events tend more towards smaller standoff distances due to the external nature of the most background sources.

source, confirming that most of the background events are created by external sources. The observed differences between signal and background events make the minimal standoff distance a good discrimination space parameter.

4.3.5 Max Distance

The fourth parameter that is used for the discriminator is the maximum distance between the positions of two energy depositions in an event defined as

$$\Delta_{\max} = \max_{i \neq j}(|\vec{x}_i - \vec{x}_j|) \tag{4.6}$$

with $\vec{x}_i$ being the three dimensional position vector of an interaction i in a single event. This parameter was not used in the EXO-200 analysis but investigations during this work suggest that it could improve the discriminator. The $\Delta_{\max}$ distributions for background and signal events differ due to the different source types and energies of the involved particles. Signal events contain two γ-rays of known energy preferred emitted back to back or in the same direction whereas background events lack such unique signatures.

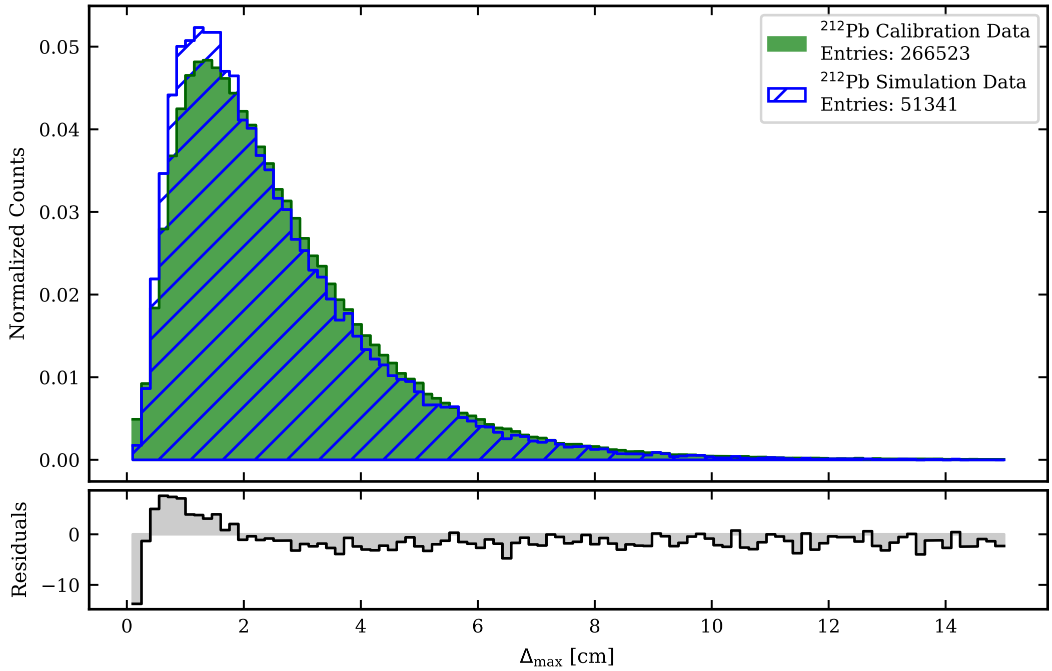


Figure 4.21: Normalized Δ_{max} histograms for simulated (blue) and measured (green) ^{212}Pb events. Error-bars are not drawn to reduce the complexity of the plot and the residuals are given in the bottom panel of the figure. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. Small differences at low Δ_{max} show deviations in the merging of S2 signals between measured and simulated data.

In order to validate the MC simulations, Δ_{max} can be compared for ^{212}Pb and ^{60}Co calibration data. In addition to the source selection and data quality cuts outlined in section 3.3 a multiplicity cut is applied. Thus, only events with at least two S2s are investigated. The Δ_{max} distributions for ^{212}Pb events are given in figure 4.21 and the distributions for ^{60}Co events in figure 4.22. One can see that the normalized histograms for both calibration sources rise quickly with increasing Δ_{max} peaking at around 2 cm. The distributions show a longer tail towards larger Δ_{max}. Except for minor differences in the region $\Delta_{max} < 2$ cm simulated and measured distributions agree.

Since only multi-site and multi-scatter events are used for this analysis, one can assume that the steep increase at low Δ_{max} shows the spatial separation capability for consecutive S2s of the detector and the reconstruction algorithms used in PAX. For small distances between energy depositions in the TPC it is more likely that signals get merged and the events are discarded by the multiplicity cut. The tail towards higher Δ_{max} differs for ^{212}Pb and ^{60}Co whereas the later one drops less steeply. This difference is caused by the different involved particles and energies. ^{60}Co events are

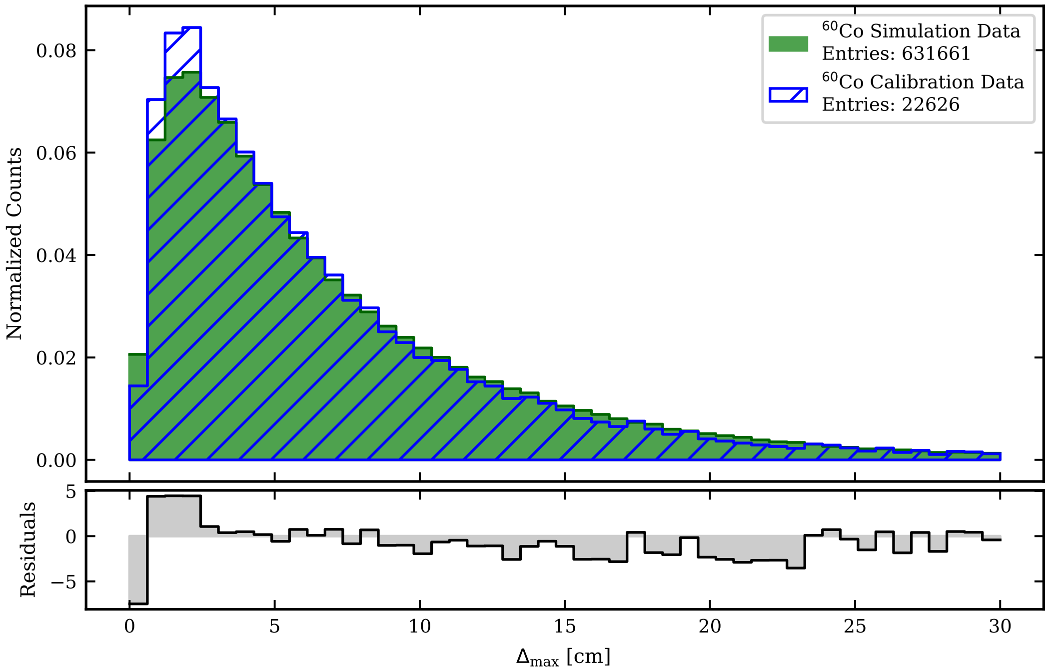


Figure 4.22: Normalized Δ_{max} histograms for simulated (blue) and measured (green) ^{60}Co events. Error-bars are not drawn to reduce the complexity of the plot and the residuals are given in the bottom panel of the plot. The spectrum decreases exponentially after a steep initial increase caused by the spatial S2 separation in the detector.

multi-scatter events of high energy γ-particles (1173.2 keV and 1332.5 keV) whereas ^{212}Pb events are multi-site events containing one lower energetic γ (238.6 keV or 415.3 keV) and a β^--electron. In case of events with multiplicity equal two, Δ_{max} for ^{212}Pb is basically the distance between the γ and β interaction position. For ^{60}Co, Δ_{max} is always the distance between two Compton scatters. In both cases, the distribution of Δ_{max} at higher values is determined by the interaction cross section of photons in liquid xenon. As the γ cross-section increases with decreasing energy[25], one can explain why the higher energetic ^{60}Co events also reach larger Δ_{max}[2].

The ^{136}Xe excited state decay signal consist of two γ-particles with fixed energies and two electrons following a $2\nu\beta\beta$ energy distribution. In this case, Δ_{max} is most likely the distance between two interaction positions of the two involved γ-rays since they are preferably emitted back to back. The distributions of Δ_{max} for signal and background events are given in figure 4.23. One can see that the background spectrum looks similar to the spectra observed for the calibration sources peaking at low Δ_{max} with a tail towards larger Δ_{max}. The Δ_{max} distribution of the signal events on the other hand show an slower increase of counts peaking at around 8 cm followed also by a tail towards larger Δ_{max}. Here, the S2 separation plays only a minor role since

[2]By fitting the falling edge of the distribution one could determine an attenuation averaged over the energy of the γ-rays.

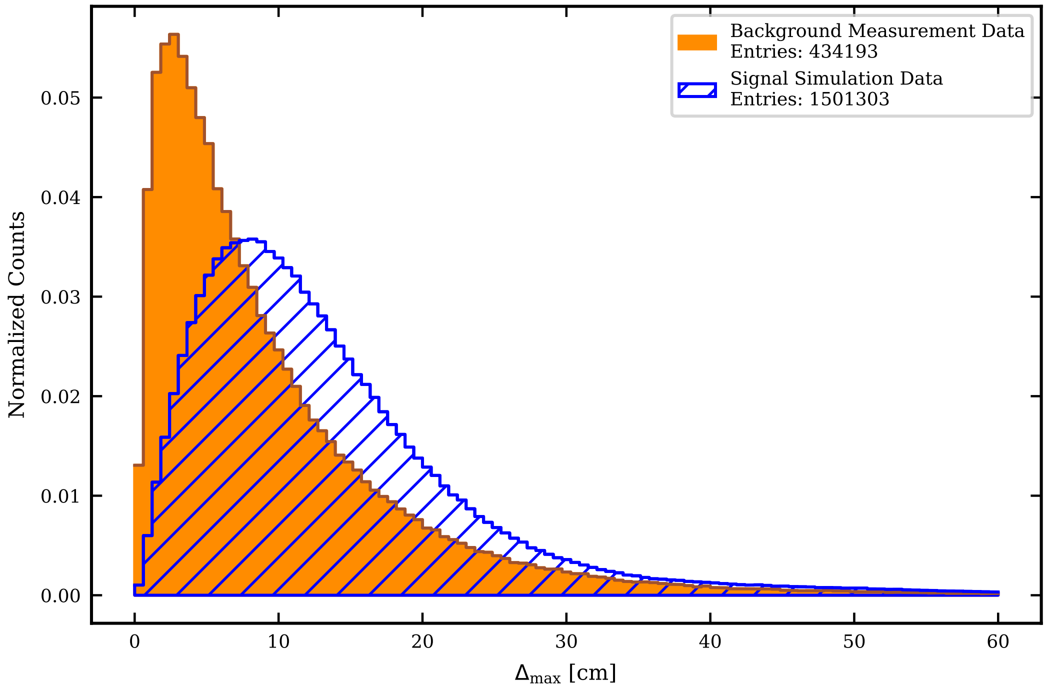


Figure 4.23: Normalized Δ_{max} histograms for simulated signal events in blue and measured background events in orange. The $\nu\beta\beta^*$-decay events show larger Δ_{max} due to the back to back emission of the γ-rays and allow a separation of signal and background.

the probability of two interactions clearly separated in space is quite high due to the back to back emission of the γ-rays.

4.3.6 $\Delta E_{\gamma,1}$, $\Delta E_{\gamma,2}$ and $\Delta E_{\gamma,\mathrm{sum}}$

The next parameters that could be used for the discriminator are modified versions of the three γ-parameters used in the EXO-200 analysis. To clearly distinguish the modified versions from the definitions in the EXO-200 analysis, the parameters are called $\Delta E_{\gamma,1}$, $\Delta E_{\gamma,2}$ and $\Delta E_{\gamma,\mathrm{sum}}$ here and are defined as:

$$\Delta E_{\gamma,i} = \mathrm{sgn}(E_j - \epsilon_i) \cdot \min(|E_j - \epsilon_i|), \text{ for } j \in S, \tag{4.7}$$

where S are the set of S2s in one event, E_j the deposited energy of S2$_j$ and $\epsilon_1 = 760.5\,\mathrm{keV}$, $\epsilon_2 = 818.5\,\mathrm{keV}$ and $\epsilon_{\mathrm{sum}} = \epsilon_1 + \epsilon_2$. $\mathrm{sgn}(x)$ is the sign function defined as

$$\mathrm{sgn}(x) = \begin{cases} -\quad 1, & \text{for } x < 0, \\ \quad\;\; 0, & \text{for } x = 0, \\ +\quad 1, & \text{for } x > 0. \end{cases} \tag{4.8}$$

In contrast to the γ-parameters in the EXO analysis, $\Delta E_{\gamma,i}$ are energy differences which can also go to negative values. If one of the energy depositions in an event is close to the energy of the emitted γ-rays in the excited $2\nu\beta\beta$-decay of ^{136}Xe, the corresponding value should be close to zero. As this is very unlikely for background events this variable should be a strong discriminator.

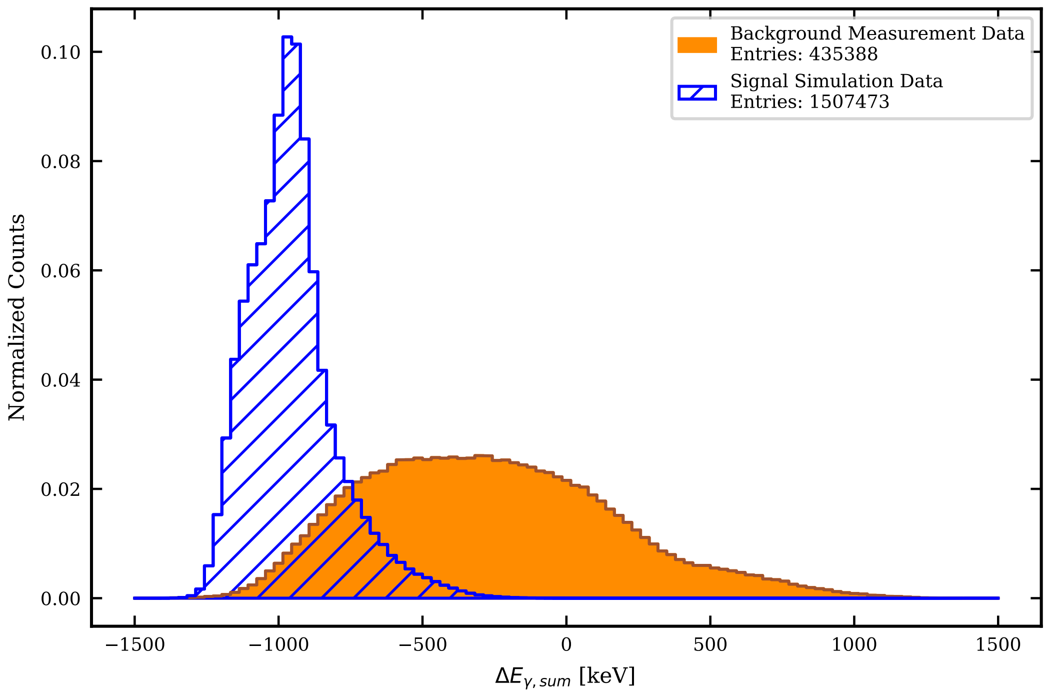


Figure 4.24: Comparison of $\Delta E_{\gamma,\mathrm{sum}}$ for simulated signal events in blue and measured background events in orange. The signal data show a peak-like structure at around $-1000\,\mathrm{keV}$ whereas the background spectrum spreads over a broad energy range.

Figure 4.24 shows $\Delta E_{\gamma,\mathrm{sum}}$ histograms of simulated signal and measured background events. The corresponding plots for $\Delta E_{\gamma,1}$ is given in figure 4.25 and for $\Delta E_{\gamma,2}$ in figure A.8 in the appendix. The signal distribution for $\Delta E_{\gamma,\mathrm{sum}}$ has a peak-like structure at around $\Delta E_{\gamma,\mathrm{sum}} = -1000\,\mathrm{keV}$ whereas the background distributions spreads from about $-1000\,\mathrm{keV}$ to $1000\,\mathrm{keV}$. The signal spectra shows that it is highly unlikely for the two emitted γ-rays of the signal to be merged into one single S2 signal. $\Delta E_{\gamma,1}$ for simulated signal events shows a peak like structure at around $-200\,\mathrm{keV}$ whereas the background distribution spreads in a double-peak structure over a broad energy range. Due to the big differences of the signal and background events in these parameters a good separation should be possible using the $\Delta E_{\gamma,i}$ parameter for the machine learning discriminator.

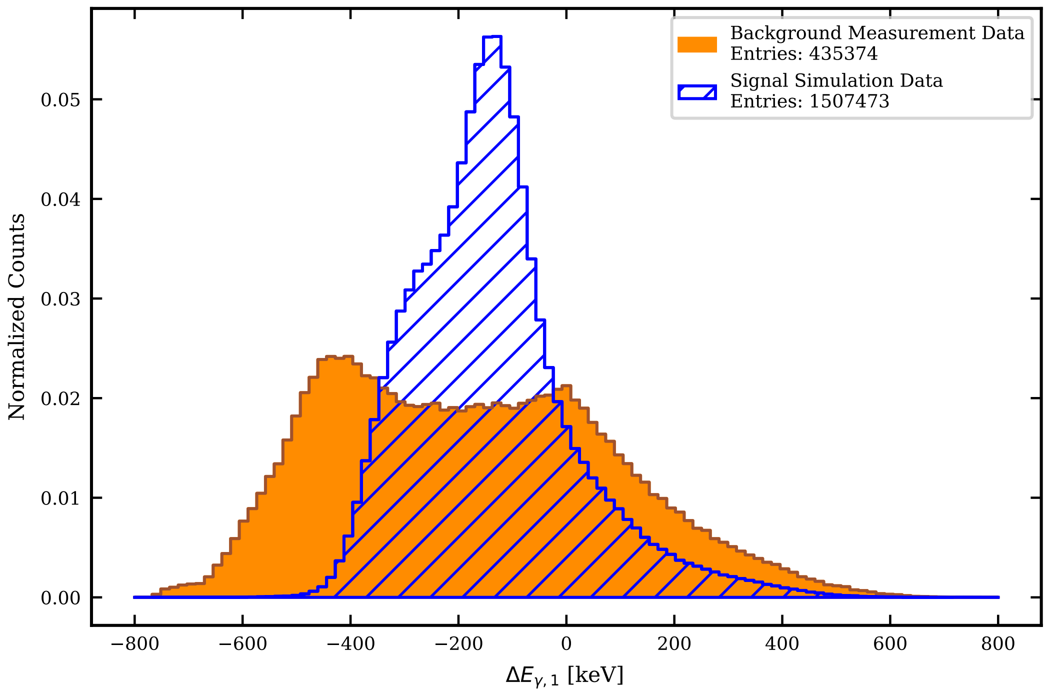


Figure 4.25: Comparison of $\Delta E_{\gamma,1}$ for simulated signal events in blue and measured background events in orange. The signal data show a peak-like structure at around $-200\,\mathrm{keV}$ whereas the background spectrum spreads in a double-peak structure over a broad energy range.

In contrast to the energy reconstruction used for the total event energy, a reconstruction based only on the $cS2_{b,i}$ has to be used here since the S1 signal is merged for all energy depositions. It is developed for this analysis with the modified NEST model for γ-particles used for the MC simulations outlined in section 3.1. The modified model is applied here, since the energy depositions of interest are caused by γ-interactions. The charge yield is calculated for the energy range from $3\,\mathrm{keV}$ to $3\,\mathrm{MeV}$ in 2056 evenly spaced steps in logarithmic space and scaled to the corresponding $cS2_b$ value by multiplication with the charge amplification factor g_2. For simplicity, a mean value for g_2 is chosen instead of the z-dependent charge amplification used in section 4.3.3.

The scaled energy dependent charge yields are now step wise linearly interpolated using the SciPy package [58]. This interpolation allows the approximate energy reconstruction of a single interaction in one event. If one would use a z-dependent g_2, a two dimensional interpolation would be necessary. One has to keep in mind that the energy resolution is worse compared to the usage of the light and charge signal since the anti-correlation between both detection channels is not exploited. The $cS2_b$ based energy reconstruction is plotted in figure 4.26.

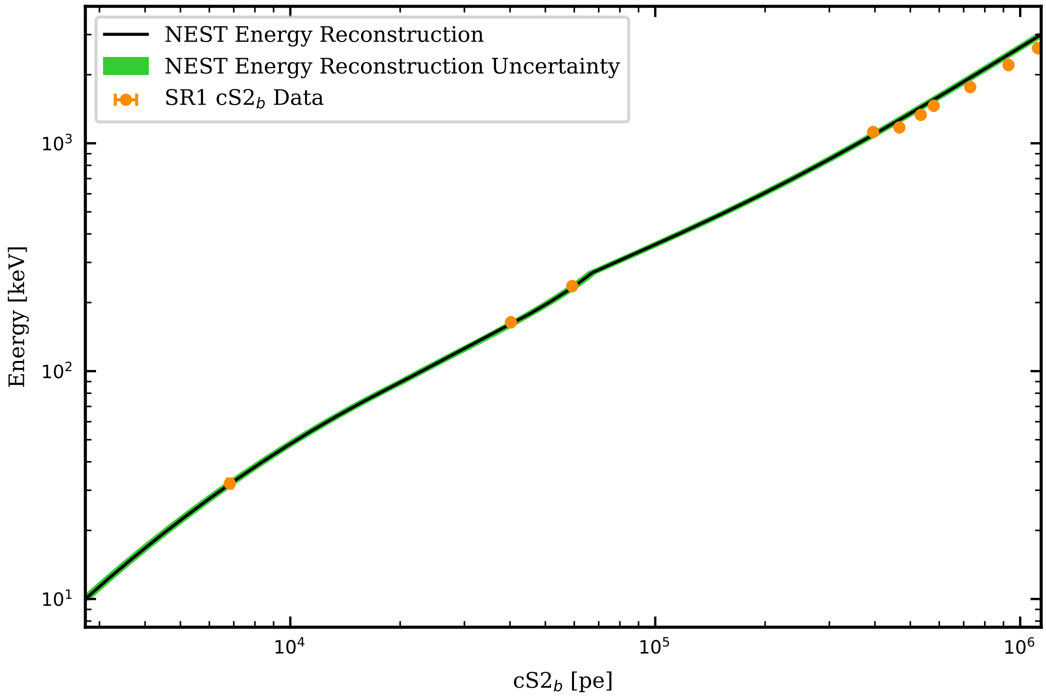


Figure 4.26: Energy plotted against the $cS2_b$ area. The NEST based energy reconstruction interpolation is plotted in black and XENON1T data in orange. The uncertainties of the measured data is drawn but to small to see. The presented interpolation can be used to reconstruct the energy of an interaction based only on the S2 signal. The uncertainty of the energy reconstruction is calculated based on the systematical uncertainty of g_2 and drawn in green.

In order to evaluate how the simulations reproduce the measured $\Delta E_{\gamma,i}$ values, ^{60}Co and ^{212}Pb events with multiplicity larger one are used as benchmark. A histogram of the $\Delta E_{\gamma,1}$ values for ^{212}Pb can be found in figure 4.27. $\Delta E_{\gamma,1}$ ranges from about -650 keV to -300 keV. The overall shape of the distributions looks similar but the calibration data tends towards higher values of $\Delta E_{\gamma,1}$ compared to the simulated data. Since only the $cS2_b$ values are used to calculate these parameters here, it is necessary to accurately reproduce the S2 size with the MC simulations. Other parameters, like the total energy of the events can exploit the S1-S2 anti-correlation to get accurate results even if $cS2_b$ and cS1 differ slightly between simulations and measurements. Thus, the deviations cancel each other out and one obtains the correct energy.

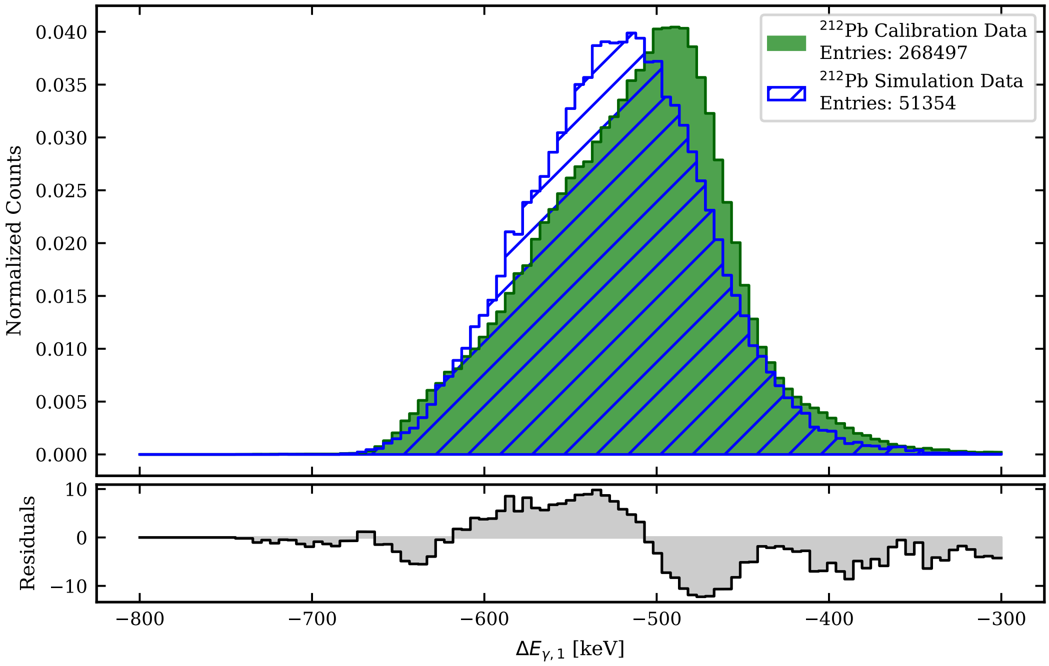


Figure 4.27: Normalized histograms of $\Delta E_{\gamma,1}$ for simulated ^{212}Pb events in blue and calibration data in green. Error-bars are not drawn to reduce the complexity of the plot and the residuals are given in the bottom panel of the figure. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. The simulation tends towards smaller $\Delta E_{\gamma,1}$ values compared to the calibration data.

Since the differences of ϵ_1 and ϵ_2 are small, the spectra of $\Delta E_{\gamma,1}$ and $\Delta E_{\gamma,2}$ are similar but shifted in energy. The histograms of $\Delta E_{\gamma,2}$ and $\Delta E_{\gamma,\mathrm{sum}}$ for ^{212}Pb events are given in the appendix in figure A.9 and figure A.10.

Histograms of the $\Delta E_{\gamma,\mathrm{sum}}$ values for the 1173.2 keV line of ^{60}Co are shown in figure 4.28. The simulated data are Gaussian shaped and centered at around −600 keV, whereas the measured data resembles a double peak with center positions at about −600 keV and −850 keV. The distributions of the 1332.5 keV line look similar to the data presented here and are given in figure A.11 in the appendix.

Deviations in $\Delta E_{\gamma,i}$ are expected for ^{60}Co since it was found in section 4.3.2 that cS1 and cS2$_b$ differ between simulations and TPC data due to differences in charge and light yield predictions between the NEST models and XENON1T data.

Small differences in $\Delta E_{\gamma,i}$ were observed for ^{212}Pb events and even larger deviations for the ^{60}Co data. In order to avoid the training of a ML discriminator on wrongly reproduced parameters, $\Delta E_{\gamma,i}$ are not used in this analysis even though they could provide significant improvements in the classification accuracy when simulated correctly. More work has to be put into the emission models to better match cS2$_\mathrm{b}$ for higher energies with the simulations.

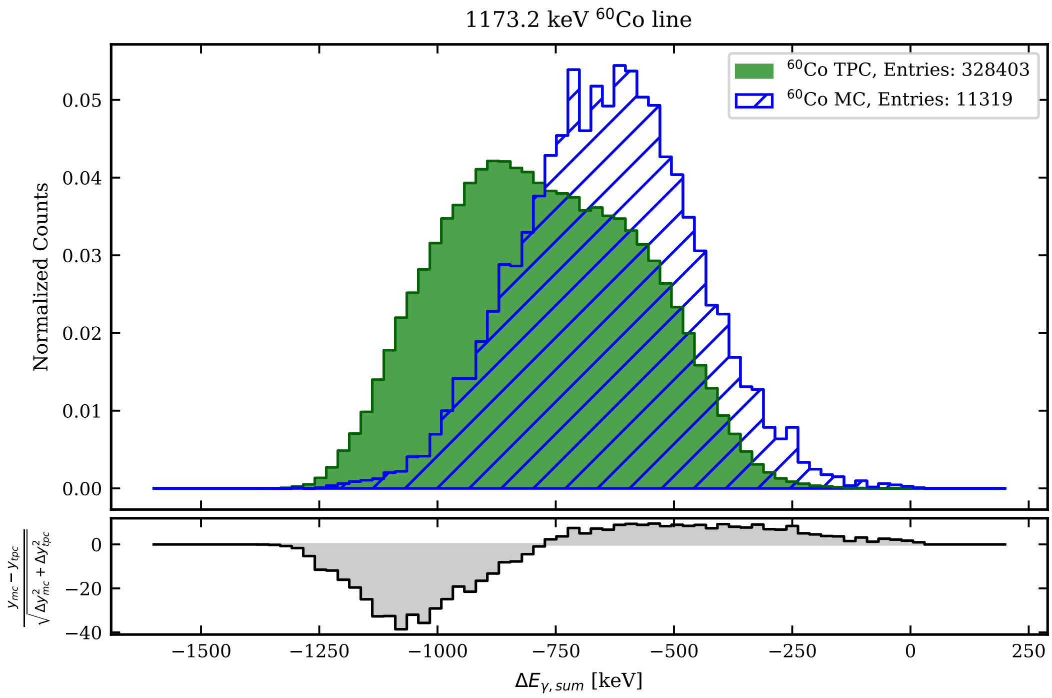


Figure 4.28: Normalized histograms of $\Delta E_{\gamma,1}$ for simulated 1173.2 keV ^{60}Co events in blue and measured events in green. Error-bars are not drawn to reduce the complexity of the plot and the residuals are given in the bottom panel of the figure. The simulated data are Gaussian shaped whereas the measured data show a double peak structure. Deviations in $\Delta E_{\gamma,1}$ are caused by differences between the NEST prediction of the charge yield and the observed yields in XENON1T data.

4.4 Multi Layer Perceptron

In the last section several discrimination space parameters were analyzed with respect to the separation capability between signal and background events and validated using calibration data. In this section a multi-layer perceptron (MLP) is developed for the binary classification of signal and background events based on the four input parameter multiplicity, energy, minimal standoff distance and $\Delta_{\max}$. The model will be trained and validated with the corresponding datasets outlined in section 2.5. The MLP It is implemented with the Keras [59] python library working with TensorFlow [53] as its *backend*[3].

The simulated signal and measured background events have to be preprocessed to increase the performance of the MLP. The mean of each input parameter is subtracted from the values of that parameter and the values are scaled such that the variance is unity. Each simulated signal event gets the target label $y_i = 1$ and each measured background event is assigned $y_i = 0$. The aim of the ML model is then to reproduce

[3]Keras provides high level building blocks and the low level tensor operations are handled by the so-called *backend*.

these labels. The loss-function used for such binary classifications is the mean *binary cross-entropy* or *logistic regression loss* [32]. It is defined as

$$\tilde{L}_{\ln}(y,p) = \frac{1}{N}\sum_{1}^{N} L_{\ln}(y_i, p_i) \tag{4.9}$$

with

$$L_{\ln}(y_i, p_i) = -(y_i \ln(p_i) + (1-y_i)\ln(1-p_i)). \tag{4.10}$$

In case of a *sigmoid* activation function of the output layer such as

$$\varphi(q) = \frac{1}{1+\exp(-q)} \tag{4.11}$$

the output of the model p_i is in the range from zero to one and represents the probability of an event to be a signal event.

Different model architectures, i.e. the number of layers and nodes in each layer, were trained on the training dataset and then tested on the validation data. The layers of the best-performing model are listed in table 4.1. All layers are *dense* layers so each neuron is connected to all neurons in the following layer. Since four input parameters are used, the first hidden layer has 640 trainable parameters in total. These are $4 \cdot 128$ values in the weight matrix and additional 128 values for the bias of each node. The number of parameters for the following hidden layers can be calculated in a similar fashion. Following the first hidden dense layer a *dropout* layer is used. *Dropout* [60] is a regularization technique that reduces overfitting by randomly setting a fraction of a layer's input connection weights to zero. In this work, a *dropout* ratio of 0.3 is used. ReLU activation functions are used for the hidden layers and a *sigmoid* function

Table 4.1: Layer structure of the MLP used in this work. In total 21313 trainable parameters are used.

Layer	Type	Neurons	Activation	Trainable Parameters
Hidden Layer	Dense	128	ReLU	640
Hidden Layer	Dropout	128		0
Hidden Layer	Dense	128	ReLU	16512
Hidden Layer	Dropout	128		0
Hidden Layer	Dense	32	ReLU	4128
Output Layer	Dense	1	Sigmoid	33

for the output layer. The performance of the model on the training and validation data is monitored during the training epochs. In a single epoch the model has seen all available events of the training data. The weight matrices and bias vectors are updated during each epoch after every 128 events, the so-called *batch size*. The mean *logistic regression loss* (cf. equation (4.9)) and accuracy of the prediction are plotted against the epoch number in figure 4.29. The accuracy is defined as the number of correctly classified events divided by the total number of events. In this case an event is classified as signal if $p \geq 0.5$ and as background if $p < 0.5$ for demonstration purposes.

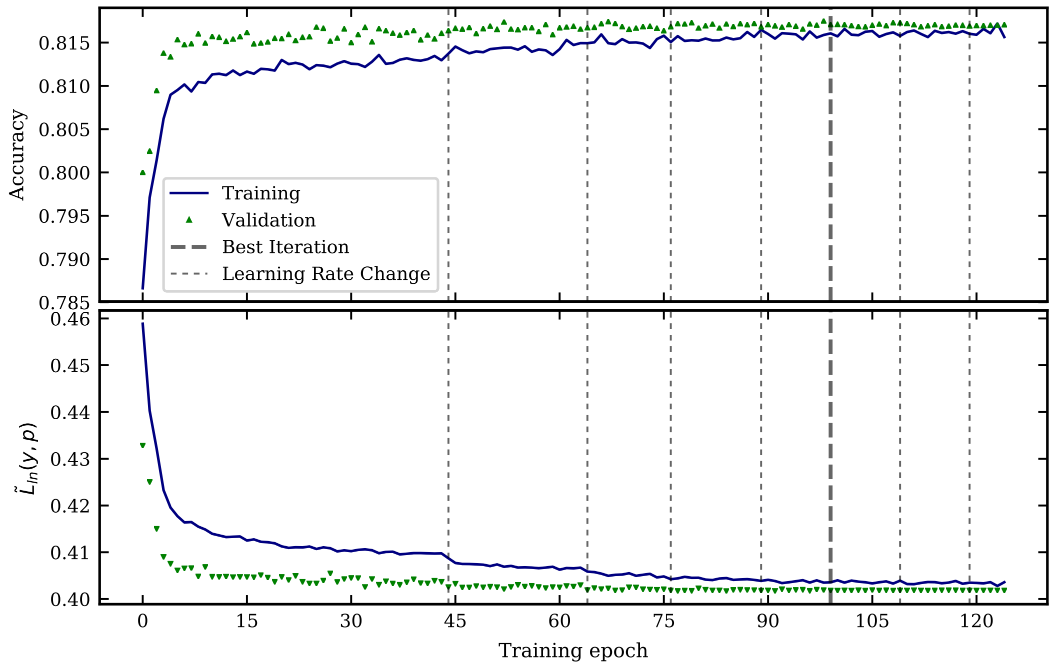


Figure 4.29: Training history of the best performing MLP in this analysis. The training accuracy in dependence of the training epoch is given in the top panel and the epoch dependent loss in the bottom panel. Both performance scores are evaluated on the training and validation data. The *dropout* is only turned on during the training and not when applying the model to other data. Thus, the achieved accuracy and loss are worse on the training data compared to the validation data. The learning rate is decreased multiple times indicated as thin dashed lines. The best performing model is saved after 99 epochs.

One finds that both the training and validation accuracy increase steeply in the first few epochs and that the loss decreases in a similar fashion. At around 7 epochs the training slows down. The *dropout* is only turned on during the training and not when applying the model to other data. Thus, the achieved accuracy and loss are worse on the training data compared to the validation data. Since the performance of the model on the validation data is monitored in each epoch, several actions can be performed during the training. These actions are implemented in so-called *callbacks*. During the presented training three *callbacks* are used.

- The first one saves the model after each epoch if the validation accuracy increased with respect to the best iteration so far.
- The second *callback* decreases the *learning rate* by a factor of 0.5 when the validation accuracy stays constant over 10 epochs. The *learning rate* corresponds to the step size of the optimizer controlling the size of the parameter update. This effect benefits the training process of the model close to the loss minimum.

As shown in figure 4.29 the accuracy could be increased by the first three reductions of the *learning rate*. The last reduction did not increase the accuracy significantly.

- The third and last *callback* stops the training when the validation accuracy stagnates over 40 epochs. This *callback* is triggered after 125 epochs.

The best-performing weight combination during the training with respect to the accuracy is saved after 79 epochs and is used for the following analysis.

4.5 Boosted Decision Tree

In the previous section a MLP was developed to classify signal and background events. In this section a different ML technique will be used for the same task: a boosted decision tree (BDT). Again the training and validation datasets outlined in section 2.5 will be used.

The BDT is implemented with the XGBoost [57] python library. As described in section 4.1.4, weak learners are added iteratively to the model improving the model's performance. A weak learner in the XGBoost library is called *estimator*. It is found that the default hyper-parameter already yield a good signal and background classification accuracy and only minor adjustments had to be made. As a reminder, hyper-parameters define a ML model and are not changed during the training process, but by human input. An overview of the hyper-parameters and their default values can be found in [61].

An upper limit of 2500 *estimators* is used and the training is stopped when the validation loss stays constant over 150 added *estimators* and the best performing model is used for the further analysis. The *learning rate* is set to 0.05 with respect to the default parameter of 0.3. This increases the number of needed *estimators* and computational costs but could also increase the performance of the model. The *maximal dept* of each used decision tree (i.e. the *estimators*) is reduced to 4 from the default value of 6. It limits the number of binary splits a decision tree can perform. A larger *maximal dept* results in a more complex model which can lead to overfitting.

The loss function used here is the *logistic regression loss* that is also used for the MLP. The training history of the BDT is given in figure 4.30. Just like for the MLP, the accuracy increases quickly when adding the first *estimators* and the increase slows down at about 80 *estimators*. At about 240 *estimators* the training accuracy continues to increase while the validation accuracy starts to stagnate. This indicates that the generalization capability of the model is almost reached and overfitting can occur if more and more *estimators* are added. The best performing model with respect to the classification accuracy includes 497 *estimators*.

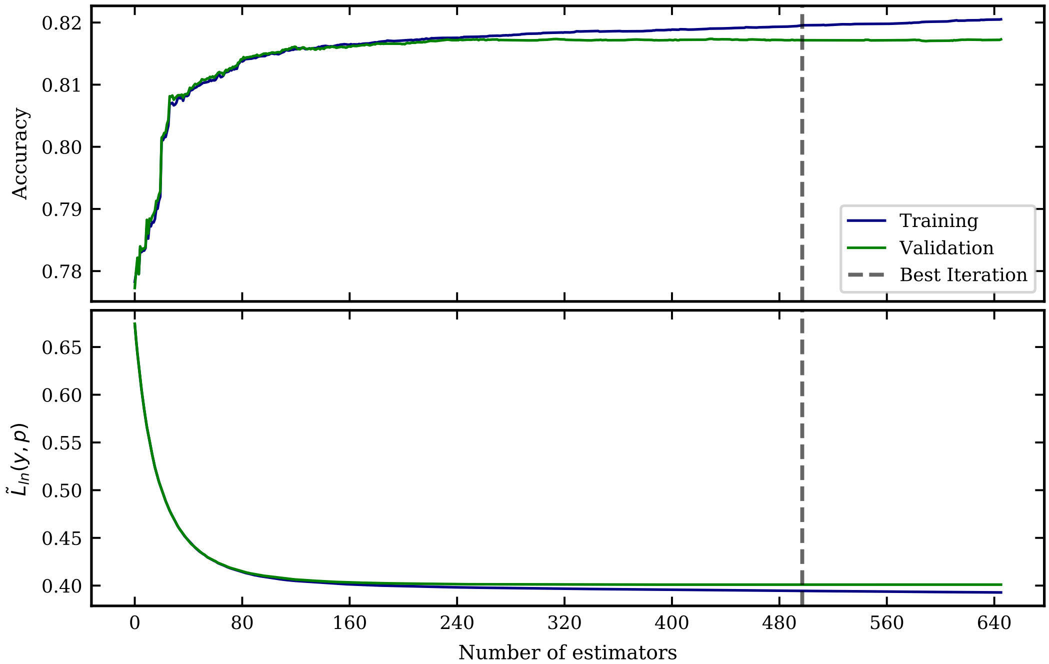


Figure 4.30: Training history of the BDT. The x-axis gives the number of *estimators* added to the model during the training. The accuracy is given in the top panel and the *logistic regression loss* in the bottom panel of the plot. Both performance metrics are evaluated for the training and validation data during. The accuracy increases for both datasets and stagnates for the validation data at about 240 *estimators*. The best iteration including 497 *estimators* is indicated as dashed line.

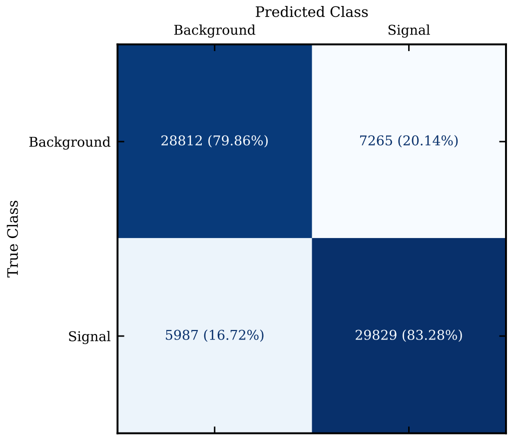
4.6 Performance Comparison

Two different ML discriminators were trained in the last two sections. In this section the performance of both models, the MLP and the BDT will be tested on the evaluation dataset (cf. section 2.5) and on modified calibration data.

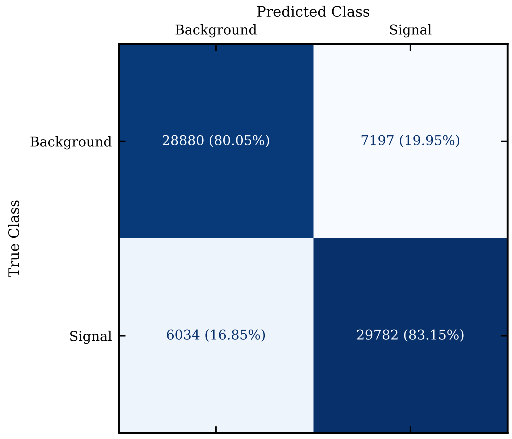
4.6.1 Evaluation Data

The evaluation dataset has a similar signal-background composition as the training and validation data but is still unseen by the two discriminators. A common performance measure for an ML model trained as binary discriminator is the *confusion matrix* given in figure 4.31a for the MLP and in figure 4.31b for the BDT. In this case, all events with an estimated signal probability of p larger or equal the classification threshold of $p_T = 0.5$ get the predicted class *signal* and events with with signal probabilities smaller than p_T get the predicted class *background*.

The *confusion matrix* lists the known true classes against the predicted classes. The diagonal entries are the numbers of correctly classified signal and background events and the off-diagonal values are the numbers of misclassified signal and background



(a) Confusion Matrix for the MLP



(b) Confusion Matrix for the BDT

Figure 4.31: Confusion matrices for the MLP and the BDT. The previously known true class is given in the rows and the predicted classes in the columns. The binary classification is set to a classification threshold of $p_T = 0.5$.

events. Since only a handful of signal events are expected in the whole XENON1T data, it is important to keep the number of wrongly classified signal events low. On the other hand, a low rate of wrongly classified background events is important as well, since the number of background events is of the order 10^5 to 10^6 compared to the expected number of signal events of the order 10^0 to 10^1. Every increase in the rate of wrongly classified background events would lead to many misclassified background events that would outnumber the few expected signal events. Figure 4.31 shows that both models achieve similar performances on the evaluation dateset with a rate of 16.72 % wrongly classified signal events for the MLP and 16.85 % for the BDT. The rate of misclassified background events is also similar for both models with 20.14 % for the MLP and 19.95 % for the BDT. These numbers are too high to just count the number of signal events with $p \geq p_T = 0.5$ and take this number as the number of signal events.

A promising method is to directly use the signal probability output of the models p_i for the determination of the number of signal events in the analysis dataset. Histograms of the probabilities obtained for background and signal events of the evaluation dataset are given in figure 4.32. The distributions obtained with the MLP are similar to the distributions from the BDT. In case of a random guess of the signal probability for each event, one would obtain a flat spectrum. A perfect classifier would contain only histogram entries at zero for the background data and at unity for the signal data.

One can investigate the separation of signal from background events by varying the classification threshold p_T. The fraction of remaining signal and background events is given in figure 4.33 for the BDT and the MLP when only events classified as signal are kept. One finds that for a low classification threshold most of the signal events are kept but almost no background events are removed. By increasing p_T the

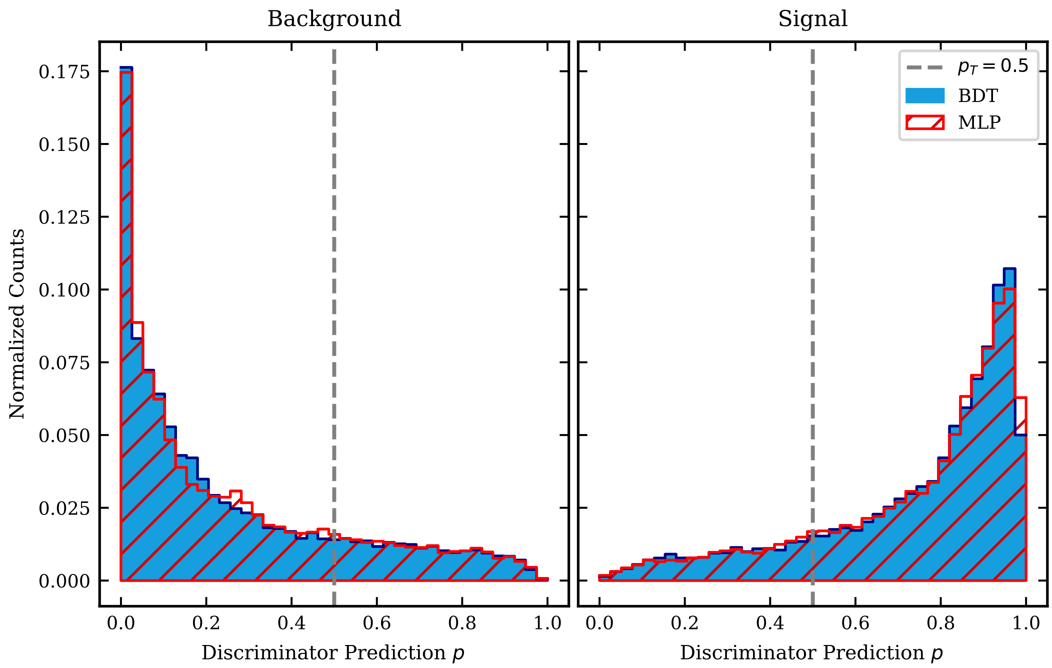


Figure 4.32: Normalized histograms of the discriminator prediction for the events' probabilities to be a signal event obtained from the MLP (red) and BDT (blue) for simulated signal (right) and measured background (left) events of the evaluation dataset. The classification threshold is drawn as dashed line for $p_T = 0.5$. A random guess would give a flat distribution. A perfect classifier would give δ-functions at zero for background and unity for signal events.

fraction of remaining background events drops steeply while the fraction of remaining signal events decreases slowly. At a threshold of $p_T > 0.8$ the fraction of remaining signal events decreases steeply. At a high classification threshold almost all signal and background events are removed.

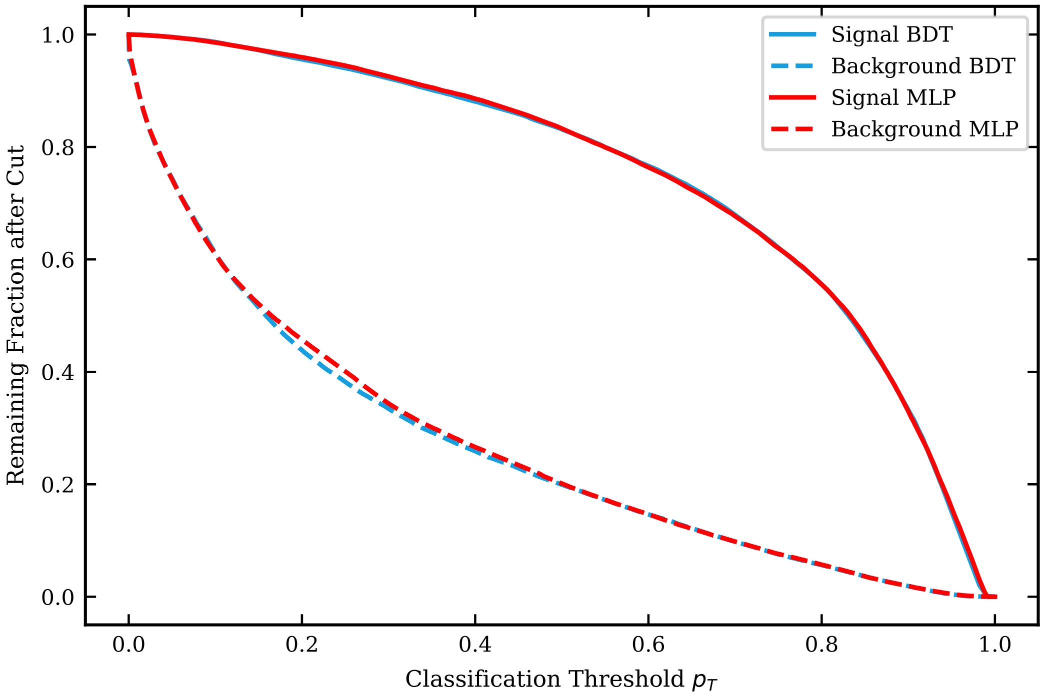


Figure 4.33: The BDT (blue) and MLP (red) are used to binary classify events as signal or background events using a threshold for the estimated signal probability. The plot shows the remaining fraction of signal (solid) and background (dashed) events as function of the classification threshold p_T if all events classified as background are cut. At $p_T = 0$, all events are classified as signal and kept and at $p_T = 1$ all events are classified as background and cut. In both cases no separation of signal and background events can be made.

4.6.2 Calibration Data

Since a mixture of simulated signal events and measured background events is used for the development of the ML discriminators, it is important to evaluate if the models learn to distinguish signal events from background events rather than learning to distinguish measured data from simulations.

One possible method of investigation is to apply the trained discriminators on simulated and measured calibration data and to then compare the output distributions. The output values of the models applied to calibration data can not be interpreted in terms of signal-background discrimination, but the distributions can be compared between simulated and measured input events. In this work the models are tested on ^{212}Pb data since it is the only calibration source used in this work with a continuous energy spectrum and energy depositions from β-electrons and γ-rays similar to the signal events of the ^{136}Xe $\nu\beta\beta^*$-decay. As for the signal and background data, only events with multiplicity greater than one are used here since the models were not trained on single-site events. Furthermore, the energy range of the signal and

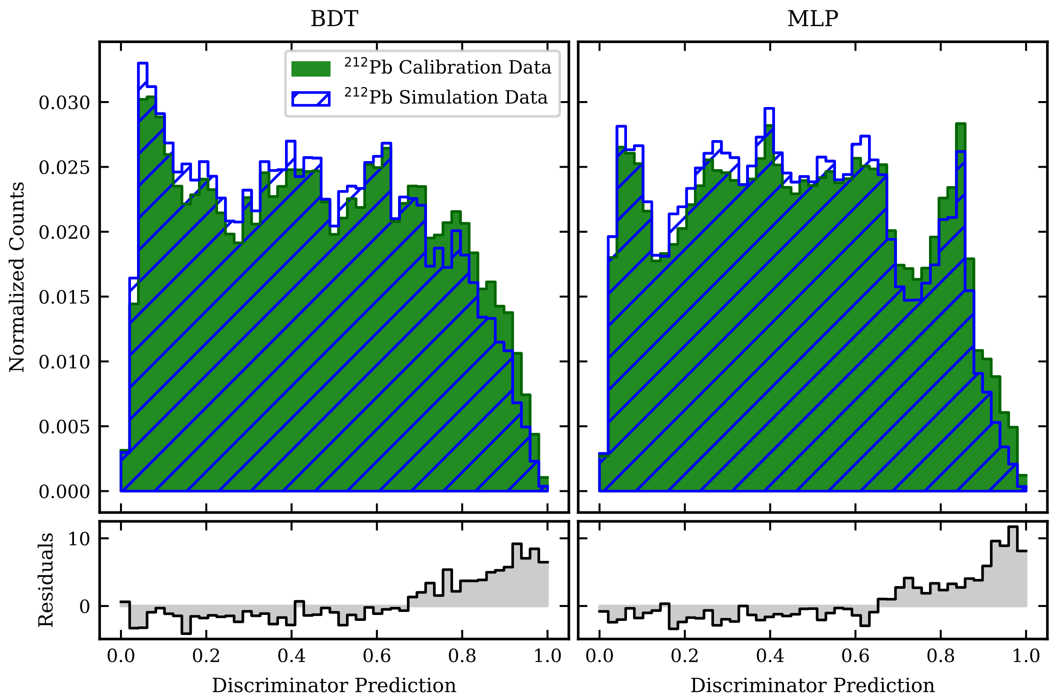


Figure 4.34: Histogram of the discriminator output applied to simulated (blue) and measured (green) multi-site ^{212}Pb data with modified energy. Error-bars are not drawn to reduce the complexity of the plot and residuals are given in the bottom panel of the figure. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. The output of the MLP is given in the right panel and the output of the BDT in the left panel of the plot.

background data is limited from 1500 keV to 2500 keV by an energy cut outlined in section 4.3.3. The energy of the ^{212}Pb decay ranges only up to 570 keV. Since the models were not trained in this energy region 1500 keV are artificially added to the reconstructed energy of the ^{212}Pb decays[4].

Histograms of the discriminator predictions for the BDT and the MLP can be found in figure 4.34. One finds that the predictions both by the BDT and the MLP for simulations and calibration data ranges from zero to one. Unlike the distributions obtained for signal and background events given in figure 4.32 the distributions for the modified ^{212}Pb events are flat with a drop-of towards unity starting at $p \approx 0.8$. Several areas of increased counts are visible next to regions with fewer counts. These substructures of the distributions obtained from the BDT differ from the ones of the MLP data. The residuals show a discrepancy for discriminator predictions with $p > 0.8$ between simulations and measured events and agree quite good for the remaining histogram range.

[4]Simulated and measured multiplicity and $\Delta_{\max}$ values are unchanged but now un-physical due to the correlation with the energy.

The deviations at larger p could be caused by differences in the multiplicity outlined in section 4.3.1. It is found that the ^{212}Pb MC simulations tend towards smaller multiplicity values compared to the measured calibration data. It was discussed that the deviation is probably caused by an unaccounted background in the calibration data that is not included in the simulations. Since section 4.3.1 shows that the expected multiplicity of signal events is larger than the multiplicity found for background events one can expect that a high multiplicity increases the chance of an event to be classified as signal.

The presence and agreement of the substructures both in the simulations and calibration data confirms that the ML models were trained on the discrimination of signal from background events like intended and not on the separation of simulated from measured events.

4.7 Effect of Input Parameters on Discriminator Prediction

In the last section, the performance of the ML models was evaluated on unseen signal and background data and on calibration data. It was found that both models show a similar performance with a slightly better performance observed for the BDT. The application to modified ^{212}Pb calibration data showed that the models are not trained on the separation simulations from measured events but on the classification of signal and background events. In this section The effect of the individual input parameters of the discrimination space on the output of the ML models will be investigated using the evaluation dataset.

Two dimensional histograms are used to evaluate the correlations of input parameters and the outputs of the ML models. The correlations of the input parameters and the (in this case) unknown signal probability of an event would be the starting point of a classical cut-based analysis. In the analysis presented in this work, the signal probability can be obtained from ML classifiers so that only a minimum set of cuts needs to be applied.

The first input parameter of the discrimination space that is investigated is the reconstructed energy of an event. The energy vs. discriminator prediction histograms are shown in figure 4.35. The energy is given on the y-axis and the model's output on the x-axis. The output is calculated both for the simulated signal events and for measured background events. One finds that the output of the MLP is similar to the BDT output. Furthermore, one can find an interesting feature at energies of 1764.49 keV and 2204.06 keV corresponding to the energies of two γ-lines from ^{212}Bi decays in detector components. The decays increase the rate of background events at this energies causing the models to learn that events here are most likely background

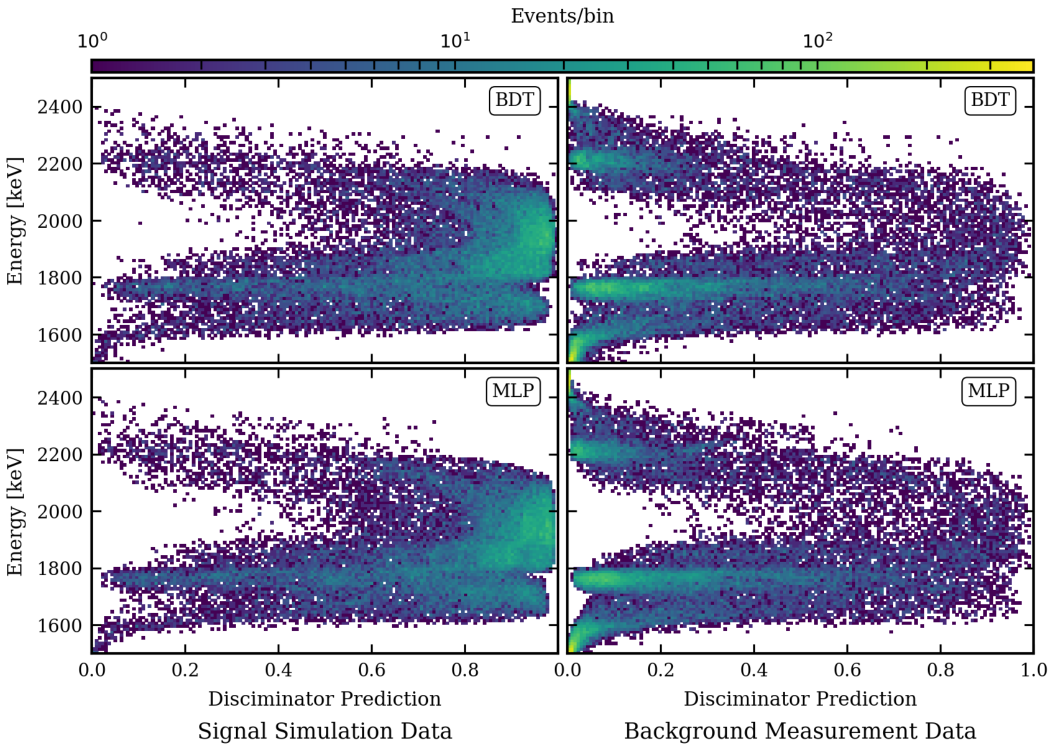


Figure 4.35: Two dimensional histograms of the discriminator prediction obtained with the BDT (top) and MLP (bottom) and energy of signal and background data.

events and an event is more likely classified as background than as signal event as a result..

Figure 4.36 shows the two-dimensional histograms of the multiplicity plotted against the discriminator prediction both for the signal and background events. Again the output of the BDT is very similar to the MLP output. One finds that the chance for an event to be classified as signal increases with the multiplicity. This behavior was already expected in section 4.3.1.

The corresponding plots for the standoff-distance and Δ_{max} discrimination space parameters can be found in the appendix.

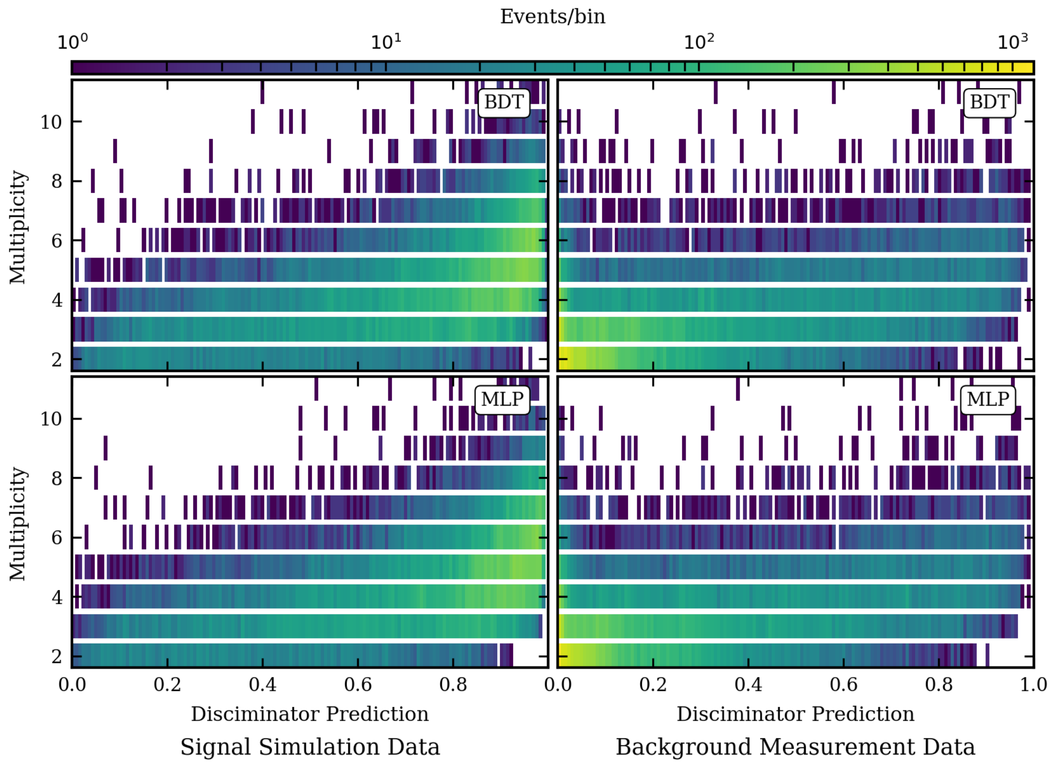


Figure 4.36: Two-dimensional histograms of the scatter multiplicity and discriminator prediction obtained with the BDT (top) and MLP (bottom) for signal (left) and background events (right).

4.8 Signal Contamination of the Background Class

As background data is used as the background class in the development of the ML discriminators, it is possible that signal events from the excited state decay of ^{136}Xe are included in the background class depending on the actual half-life of the decay. Thus, one would have wrongly labeled events in the training, validation and evaluation datasets. An investigation of how this contamination affects the training process and performance of the ML discriminators is necessary.

The investigation is carried out with a modified training dataset. A number n of randomly chosen background events are removed from the training data and replaced by simulated signal events labeled as background corresponding to an effective half-life (equation (2.7)) under the assumption that no other signal events are present in the background class. The added signal events are randomly selected from simulated data not used in the training, validation or evaluation datasets in order to avoid a bias by duplicate events. Since background events are replaced one to one by wrongly labeled signal events the overall balance of the classes in each training dataset stays constant with 50 % signal and 50 % background. This balance is important to ensure accurate calculations of the loss and accuracy during the training process.

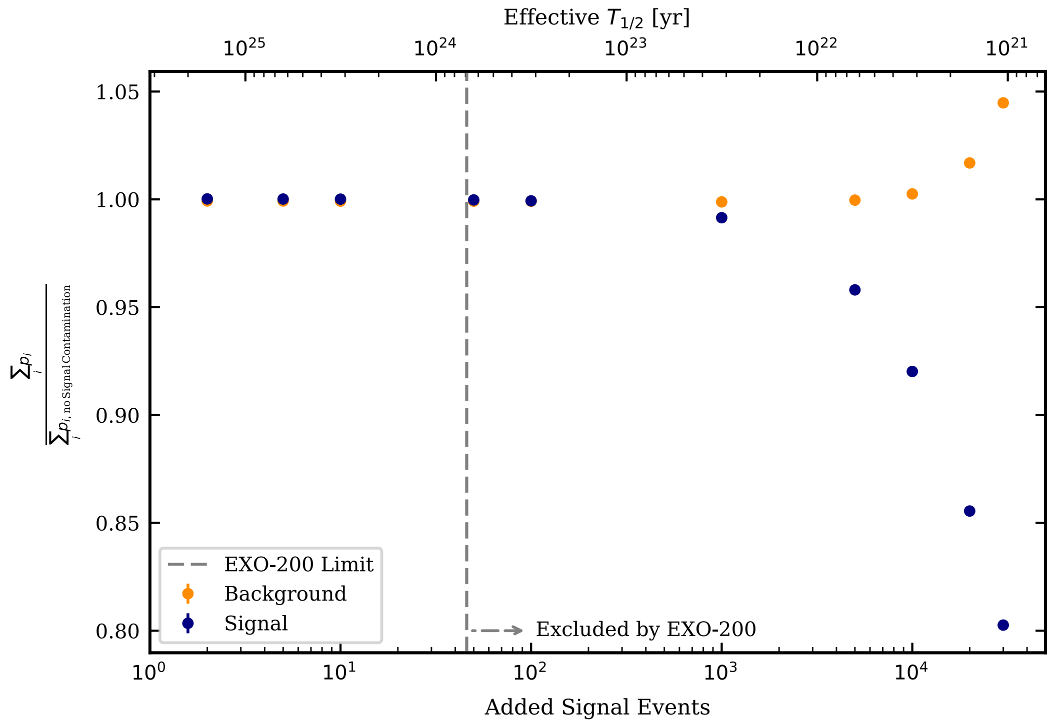


Figure 4.37: Investigation of the signal contamination in the background class. The BDT model is trained on data with signal events artificially labeled as background and applied to signal and background events from the unmodified evaluation data. The number of exchanged events can be used with the exposure of the training data to calculate a decay half-life. The mean and standard deviation is calculated for the summed probabilities and normalized on the mean value without exchanged events and plotted against the number of added signal events. The uncertainties are drawn here but smaller than the marker. One finds deviations from the unmodified case in form of an increased summed probability for background events and decreased summed probability for signal events at an effective half-life that is already excluded by the EXO-200 analysis.

The BDT model is trained on the modified training data and afterwards applied to signal and background events of the unmodified evaluation data to get the probability for each event i to be a signal event. Only the BDT is used here due to shorter training periods reducing the computational costs of this investigation[5]. In order to allow a direct comparison of the model's performance, $\sum_i p_i$ is calculated both for the signal and background events. To reduce statistical uncertainties, this procedure is repeated multiple times for different numbers of exchanged background events. The mean and standard deviation for the summed probabilities is calculated and normalized on the corresponding sum of the probability for the model trained on unmodified data. This ratio is contrasted with the number of added signal events shown in figure 4.37. Using the known exposure of the training dataset of 40.91 d and the number of artificially

[5] The training of the BDT on the training dataset takes about 30 s. The training of the MLP takes about 15 min.

injected signal events n one can use equation 2.7 to calculate a corresponding effective half-life.

One finds that the output of the BDT is not affected for $n \leq 10^2$ exchanged events since both the normalized probability sum for signal and background events is close to unity. Starting at around 10^3 exchanged events the normalized probability sum for signal events decreases which corresponds to an effective half-life of about $T_{1/2} \approx 1 \times 10^{22}$ yr. At $n \approx 10^4$ an increase can be observed for the background events. Since this half-life is excluded by the EXO-200 analysis one can conclude that the presence of a few signal events in the background data does not affect the training of the BDT due to the large number of used signal and background events in the datasets.

In this chapter two machine learning discriminators were developed and tested on simulated and measured data, a MLP and a BDT. It was found that both models show a similar performance on unseen data with a slightly higher signal-background discrimination found for the BDT. It was shown that the training process of the BDT is not affected by a low number of signal events in the background dataset. The BDT will be used in the next chapter to search for events of the ^{136}Xe $\nu\beta\beta^*$-decays in the analysis dataset.

5 Application of the Machine Learning Discriminator to Analysis Data

The BDT discriminator developed in section 4.5 is now used to search for signal events in the analysis dataset. For this data it gives a probability for each event to be a signal originating from the $2\nu\beta\beta$-decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba.

The capability to observe the decay or to set a lower limit on the half-life is limited by two factors: the imperfect discriminator and the lack of a simulated, signal-free background model. If a perfect classifier were available, one could count the number of events classified as signal and take this number as the number of observed decays. The classifiers developed in this work are, however, not perfect as shown in section 4.6, so signal events can be classified as background events and vice versa.

Since no background simulations are available, measured background data from the TPC was used as the background class for the training and evaluation of the ML discriminator. A few signal events could be included in this data depending on the actual half-life of the signal decay. It was shown in section 4.8 that the training of the ML models is not affected by a small number of wrongly labeled signal events.

If a signal free background model were available, one could fit the distribution of the predicted signal probabilities for each event in the analysis dataset with

$$f(n_{\mathrm{sig}}, n_{\mathrm{bg}}) = n_{\mathrm{sig}} \cdot \mathrm{PDF}_{\mathrm{sig}} + n_{\mathrm{bg}} \cdot \mathrm{PDF}_{\mathrm{bg}}. \tag{5.1}$$

This equation includes the number of signal events n_{sig}, the number of background events n_{bg} and the probability density functions (PDFs) of the discriminator output for signal and background events. In this work, the PDFs are obtained from the evaluation data as normalized histograms of the discriminator output of the signal and background classes. Due to the lack of a signal-free background model, n_{sig} would be fitted to zero since the background PDF would already include a sufficient contribution of signal events to describe the signal probability distribution of the analysis data when neglecting statistical fluctuations.

An alternative approach [62] is used here. The number of signal events that can maximally fit into the signal probability distribution of the analysis data is determined by scaling the signal PDF to the histogram of the probabilities of the analysis data. n_{sig} is maximized under the condition that for each bin i with N_i entries

$$N_i \geq n_{\text{sig}} \cdot \text{PDF}_{\text{sig},i} \tag{5.2}$$

is fulfilled. $\text{PDF}_{\text{sig},i}$ is here the signal PDF in bin i. A graphical representation for this procedure is given in figure 5.1. One finds that the maximum number of signal events possible is bound only by the number of entries N_k in the last bin. One now assumes that all entries in the last bin are signal events, so one can determine a 90 % upper limit on the number of signal events in this bin. For $N_k \leq 15$ the upper bound

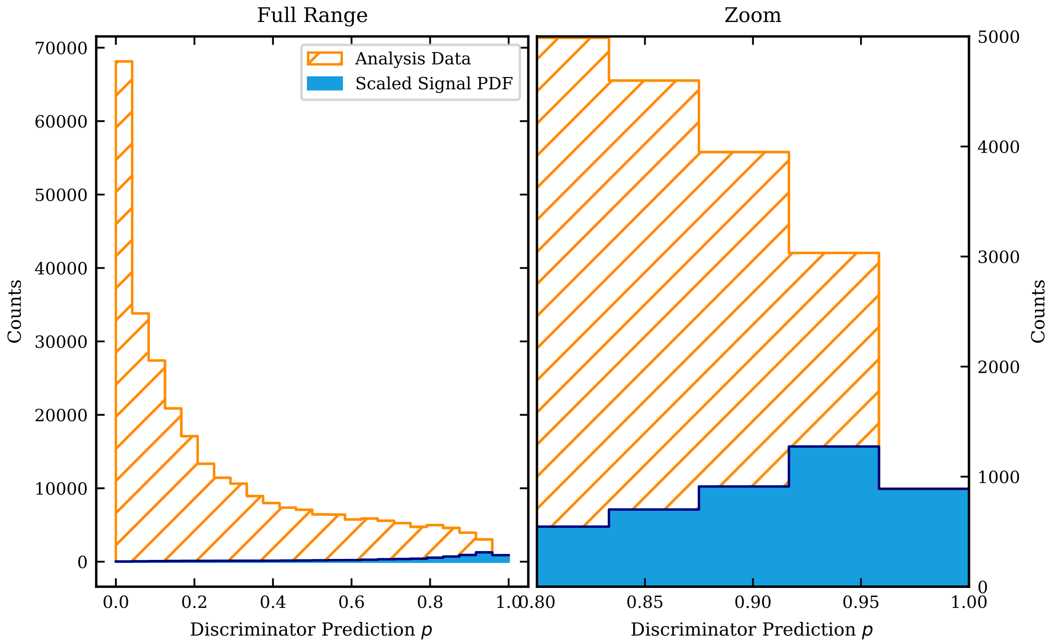


Figure 5.1: Histogram of the signal probabilities of the analysis data in orange and the scaled signal PDF in blue. The full discriminator prediction range is given in the left panel of the plot and a zoom on the last five bins on the right panel of the plot. The signal PDF is scaled to get the same number of entries in the last bin for the analysis data and scaled signal PDF.

$N_{\text{FC},90\,\%}$ of the Feldman Cousins [63] 90 % confidence interval is added to the observed number of entries. One assumes that the signal mean follows a Poisson distribution (cf.[63] table IV and V) and that the mean expected background corresponds to N_k.

$$N_{90\,\%,k} = N_k + N_{\text{FC},90\,\%} \tag{5.3}$$

In case of $N_k > 15$, one can assume that the number of observed events in the last bin varies symmetrically around $\mu = N_k$ with a width of the normal distribution of $\sigma = \sqrt{N_k}$. Then, 90 % of all values are smaller than $\mu + 1.282\sigma$ [64]:

$$N_{90\,\%,k} = N_k + 1.282 \cdot \sqrt{N_k}. \tag{5.4}$$

In order to obtain the 90 % upper limit on the total number of signal events $N_{90\,\%,\mathrm{cut}}$ after cuts, $N_{90\,\%,k}$ is normalized on the signal probability in the last bin k.

$$N_{90\,\%,\mathrm{cut}} = \frac{N_{90\,\%,k}}{\mathrm{PDF}_{\mathrm{sig}}(k)} \tag{5.5}$$

Before using the BDT discriminator, three cuts were used reducing the number of signal events. The energy and multiplicity cuts removed 89.36 % of the background events while keeping $\lambda_{\mathrm{E,M}} = 96.35\,\%$ of the simulated signal events. $\lambda_{\mathrm{Fid}} = 69.82\,\%$ of the signal events occurring in the fiducial volume pass the fiducial volume cut while the other events are removed due to energy depositions in the outer detector region. The maximum number of signal events in the analysis data before usage of cuts is calculated as

$$N_{90\,\%} = \frac{N_{90\,\%,\mathrm{cut}}}{\lambda_{\mathrm{E,M}} \cdot \lambda_{\mathrm{Fid}}}. \tag{5.6}$$

Using the livetime of the analysis dataset of $t = 165.3\,\mathrm{d}$, the xenon mass in the fiducial volume of $m = 1029.4\,\mathrm{kg}$[36], the atomic mass of xenon $m_u = 131.293\,\mathrm{u}$ [17] and the abundance of ^{136}Xe of $\eta = 8.49\,\%$ [18], the lower limit of the half-life at 90 % confidence level can be calculated with

$$T_{1/2} > \frac{t \cdot m \cdot \ln(2)}{m_u \cdot N_{90\,\%}}\eta. \tag{5.7}$$

In order to investigate how the binning of the signal PDF and of the histogram of the analysis data signal probabilities affect the limit, the half-life limit is calculated for different binnings. The result of this investigation is shown in figure 5.2. One finds that for less than 96, bins symmetric uncertainties can be used. The Feldman Cousins method has to be applied when more bins are used. The best limit of

$$T_{1/2} > 2.92 \times 10^{22}\,\mathrm{yr} \text{ at } 90\,\% \text{ C.L.}$$

is achieved for 130 bins, where the last bin of the analysis data has zero entries while still keeping a non-vanishing probability of finding signal events in this bin. The limit deteriorates when the number of bins is further increased due to the shrinking signal probability in each bin. One could combine multiple bins without entries from the analysis data, but the result should be similar to the limit obtained for 130 bins. The half-life limit set by this analysis is a factor of 23.63 smaller than the current best limit set by EXO-200 and thus compatible with this limit.

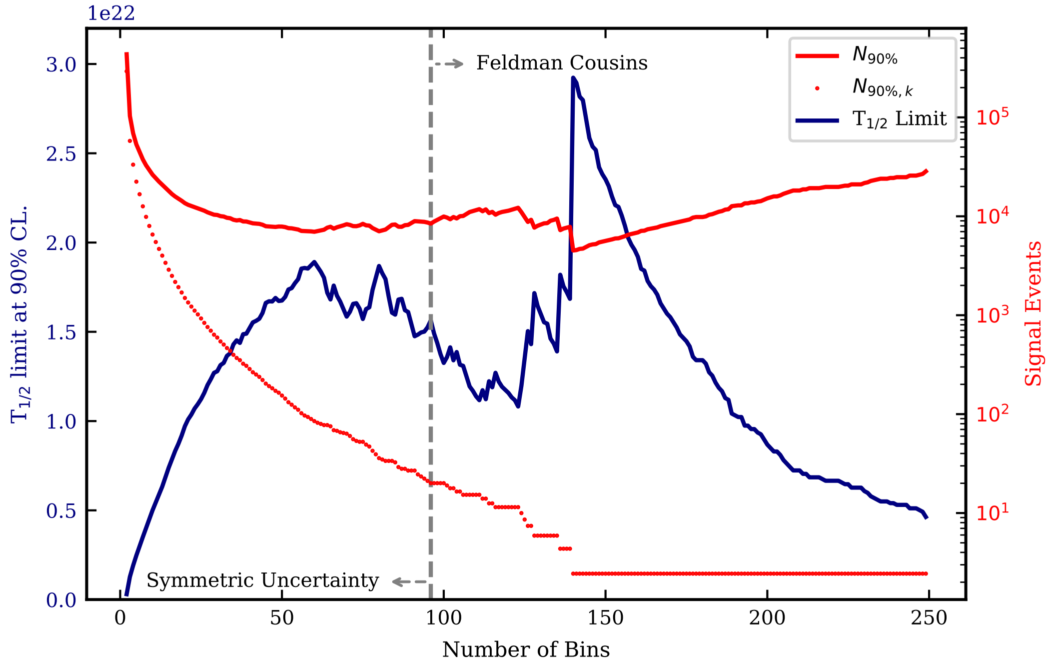


Figure 5.2: Lower limit of the half-life in dependency of the used number of bins plotted in blue. The total number of signal events before cuts is given as solid red line and the upper limit of signal events in the last bin given as red dots. For less than 96 bins, symmetric uncertainties can be used to estimate the 90 % upper limit of signal events in the last bin and for a finer binning the Feldman Cousins approach had to be used. The border of the corresponding regions in the plot are marked with a dashed line.

6 Conclusion and Outlook

A search for the $2\nu\beta\beta$-decay of ^{136}Xe to the 0_1^+ excited state of ^{136}Ba ($2\nu\beta\beta^*$) with the XENON1T experiment was presented in this work. A lower limit on the half-life was set to $T_{1/2} > 2.92 \times 10^{22}$ yr at 90 % C.L.

The $2\nu\beta\beta^*$-decay of ^{136}Xe is a process allowed by the Standard Model of particle physics which has not been observed yet directly due to the long expected half-life of $T_{1/2} = 2.5 \times 10^{25}$ yr from nuclear theory. Chapter two outlined that a direct observation can be used to validate nuclear matrix element calculations. These are needed to constrain new physics in case of an observation of neutrinoless double β-decay. It was outlined that the XENON1T experiment is able to search for the ^{136}Xe $2\nu\beta\beta^*$-decay due to the low background rate, the large exposure and the ability to reconstruct events at MeV energies. The detector principle of the dual phase time projection chamber was outlined along with an overview of the analysis steps in this work. Since only about 8 events are expected in the analyzed XENON1T data and since these exhibit a unique coincidence of two β-electrons and two γ-rays, new techniques have to be used to search for these events. In this work a machine learning discriminator was used, trained on the separation of signal and background events.

Simulated signal and measured background events were used for the development of the discriminator. Thus, simulation tools were needed to simulate events resembling the measured data. An outline of these tools used was given in chapter three. Modifications were made to the event generator Decay0 to include the angular correlation of the γ-rays emitted in the ^{136}Xe $2\nu\beta\beta^*$-decay in order to obtain realistic signal simulations. An interface for the new waveform simulator was developed including a clustering of energy depositions and the calculation of the number of charge and light quanta using the NEST framework. A modified version of the NEST model for γ-rays was used to achieve a better agreement of simulations and measured data. It was found that for energies in the MeV scale differences are still present between the charge and light yield predictions from NEST and the measured data. In order to validate the simulations, calibration data was used. An outline of the sources was given along with the required event selections.

In order to allow a good discrimination between signal and background events, a range of input parameters are needed for the discriminator. These have to be accurately reproduced from the simulation tools and differ between signal and background events.

Four parameters were used in this work: the multiplicity, reconstructed energy, minimal standoff distance and the maximal distance between interactions in an event. The combination of these parameters form the discrimination space which was analyzed in chapter four. A good agreement between simulations and measured calibration data was found for all four parameters as well as reasonable differences between signal and background events. Three other possible input parameters $\Delta E_{\gamma,1}$, $\Delta E_{\gamma,2}$ and $\Delta E_{\gamma,\mathrm{sum}}$ were investigated as well, but differences between simulations and calibration made a usage in the machine learning discriminator not possible. The differences are mainly caused by the emission models calculating the number of charge and light quanta.

Following the analysis of the discrimination space, two machine learning discriminators were trained: a multi-layer perceptron and a boosted decision tree. At a classification threshold of $p_T = 0.5$ the BDT showed a slightly better rate of misclassified background events of 19.95 % compared to the MLP with a rate of 20.14 %. The boosted decision tree was used for the further analysis due to this result in combination with the lower computational costs required. As measured background events were used in the development of the machine learning discriminators a contamination of the background training data with signal events was possible. An investigation showed that the influence on the training process is negligible for $2\nu\beta\beta^*$-decay half-lives larger 10^{23} yr. This is below the current best limit set by the EXO-200 collaboration, so it is unproblematic for the training of the discriminators.

The trained BDT was used to search for $2\nu\beta\beta^*$-decay events in the XENON1T data as outlined in chapter five. It was investigated if a fraction of the measured background data could be used as a background model. A simulated background model was not available due to the high computational costs. In contrast to the training of the BDT, the signal contamination in this data is of the same order as in the analysis data. Thus a fit of the obtained signal-probability distribution with probability density functions for signal and background events was not feasible. As a conservative approach, the maximal contribution of the signal events to the signal-probability distribution of the analysis data was determined. It was assumed that all events in the last bin of the analysis data discriminator output histogram are signal events. This allowed the calculation of an upper limit on the overall signal event number and a lower limit in the half-life was set at $T_{1/2} > 2.92 \times 10^{22}$ yr at 90 % C.L..

Improvements in a future analysis can be made in several fields of the analysis. Significant advancements on the reachable sensitivity can be achieved with improved simulations. Benefits will arise from two effects.

First, not all investigated parameters could be used for the discriminator due to differences between simulation and calibration data. With improved simulations more parameters can be added to the discrimination space and thus improve the separation of signal and background events. Beside the analyzed $\Delta E_{\gamma,1}$, $\Delta E_{\gamma,2}$ and $\Delta E_{\gamma,\mathrm{sum}}$ parameters in this work, other variables like $\ln\left(\frac{\mathrm{cS2_b}}{\mathrm{cS1}}\right)$, $\frac{\max(\mathrm{cS2}_{\mathrm{b},i})}{\mathrm{cS2_b}}$ or the standard deviation of the S2 areas of an event could be used. These values could further improve the signal-background separation by reflecting different energy loss mechanics for events with higher multiplicities. When the simulations are known to reproduce accurate results other machine learning techniques like *deep learning* directly on the waveform-level could be used as well, reducing the need for better discrimination space parameters.

A second benefit from improved simulations can be obtained by simulation of events of the background spectrum. This will allow a fit of the of the discriminator output using probability density functions for signal and background events. Thus, a better limit or even a discovery could be achieved. The latter is not possible without a simulated background model.

Additional improvements will be made with the upcoming XENONnT experiment. A larger exposure of about 20 txyr will result in $\mathcal{O}(100)$ measured $2\nu\beta\beta^*$-decay events. A further reduced background also benefits the sensitivity. In addition to the search for the excited state decay of ^{136}Xe, other rare decays with unique signatures like a $\mathrm{EC}\beta^+$-decay [65] could be analyzed using a similar machine learning based method as presented in this work.

A Appendix

A.1 Decay0 Modification

```
subroutine Ba136low(levelkeV)
!c Subroutine describes the deexcitation process in Ba136 nucleus
!c after 2b-decay of Xe136 to ground and excited 0+ and 2+ levels
!c of Ba136 ("Table of Isotopes", 7th ed., 1978).
!c Call  : call Ba136low(levelkeV)
!c Input : levelkeV - energy of Ba136 level (integer in keV)
    occupied
!c                   initially; following levels can be occupied:
!c                   0+(gs) -     0 keV,
!c                   2+(1)  -   819 keV,
!c                   2+(2)  -  1551 keV,
!c                   0+(1)  -  1579 keV.
!c Output: common/genevent/tevst,npfull,npgeant(100),pmoment(3,100)
    ,ptime(100).
!c VIT, 28.06.1993, 22.10.1995.
common/genevent/tevst,npfull,npgeant(100),pmoment(3,100),
9            ptime(100)
npg1579=0
npg819=0

tclev=0.
if(levelkev.eq.1579) go to  1579
if(levelkev.eq.1551) go to  1551
if(levelkev.eq. 819) go to   819
if(levelkev.eq.   0) go to 10000
                     go to 20000
!c-----------------------------------------------------------------------
1579      thlev=0.

          Egamma = 0.7605
          EbindK = 0.037
          cg = 1.
          cK = 3.2e-3
          cp = 0.
          p = rnd1(d)*(cg+cK+cp)
```

```
      if(p.le.cg) then
         call gamma(Egamma, tclev, thlev, tdlev)
         npg1579=npfull
      else if(p.le.cg+cK) then
         call electron(Egamma-EbindK, tclev, thlev, tdlev)
         call gamma(EbindK, 0., 0., tdlev)
      else
         call   pair(Egamma-1.022, tclev, thlev, tdlev)
      end if
      go to 819
!c-----------------------------------------------------------------------
1551  thlev=0.
      p=100.*rnd1(d)
      if(p.le.50.) go to 15511
                     go to 15512
15511 call nucltransK(1.551,0.037,7.5e-4,0.5e-4,tclev,thlev,tdlev
   )
      return
15512 call nucltransK(0.733,0.037,3.5e-3,0.,tclev,thlev,tdlev)
      go to 819
!c-----------------------------------------------------------------------
819  thlev=1.9e-12

      Egamma = 0.8185
      EbindK = 0.037
      cg = 1.
      cK = 2.6e-3
      cp = 0.

      p=rnd1(d)*(cg+cK+cp)

      if(p.le.cg) then
         call gamma(Egamma,tclev,thlev,tdlev)
         npg819=npfull
      else if(p.le.cg+cK) then
         call electron(Egamma-EbindK,tclev,thlev,tdlev)
         call gamma(EbindK,0.,0.,tdlev)
      else
         call pair(Egamma-1.022,tclev,thlev,tdlev)
      end if

      !c check if a correlation has to be applied and get the
          total momentum of each gamma

      if (npg1579.ne.0.and.npg819.ne.0) then
```

```
            p1579=sqrt(pmoment(1,npg1579)**2+pmoment(2,npg1579)**2+
            9                    pmoment(3,npg1579)**2)
            p819=sqrt(pmoment(1,npg819)**2+pmoment(2,npg819)**2+
            9                    pmoment(3,npg819)**2)

            a2=-15./8.
            a4=20./8.

            twopi=6.2831853

            phi1=twopi*rnd1(d)
            ctet1=1.-2.*rnd1(d)
            stet1=sqrt(1.-ctet1*ctet1)
            phi2=twopi*rnd1(d)
            ctet2=1.-2.*rnd1(d)
            stet2=sqrt(1.-ctet2*ctet2)

!c          see Spherical law of cosines!
            ctet=ctet1*ctet2+stet1*stet2*cos(phi1-phi2)
!c          check if a y coordinate is smaller than the value of the
     corr. function
!c         "box" method of sampling from p distribution
            if(rnd1(d)*(5./8.+abs(a2)+abs(a4))>5./8.+a2*ctet**2+a4*
                ctet**4)
            9    go to 1

            pmoment(1,npg1579)=p1579*stet1*cos(phi1)
            pmoment(2,npg1579)=p1579*stet1*sin(phi1)
            pmoment(3,npg1579)=p1579*ctet1
            pmoment(1,npg819)=p819*stet2*cos(phi2)
            pmoment(2,npg819)=p819*stet2*sin(phi2)
            pmoment(3,npg819)=p819*ctet2

        end if
        return
!c-----------------------------------------------------------------------
10000   return
!c-----------------------------------------------------------------------
20000   print *,'Ba136: wrong level [keV] ',levelkev
!c-----------------------------------------------------------------------
return
end
```

A.2 Quanta Generation Interface

```python
import logging
import time
import pickle
import uproot
import nestpy

import numpy as np
import pandas as pd

from .core import RawData
from sklearn.cluster import DBSCAN

export, __all__ = strax.exporter()
__all__ += ['instruction_dtype', 'truth_extra_dtype']

instruction_dtype = [('event_number', np.int), ('type', np.int), ('t', np.int64),
    ('x', np.float32), ('y', np.float32), ('z', np.float32),
    ('amp', np.int), ('recoil', '<U5'), ('e_dep', np.float32), ('created_by', '<U10')]


#Function adapted from Pietro Di Gangi
#https://github.com/XENON1T/MCAnalysisScripts/blob/master/G4EventDisplay/G4EventDisplay.ipynb
def get_daughters_trackid(event, parentids):
    '''Returns a trackid list of daughters of a set of parent particles.'''
    event = event[event.parentid.isin(parentids)]
    tracks = list(set(event.trackid)) # trackids present in the event
    tracks.sort() # sort trackids
    return tracks

#Function adapted from Pietro Di Gangi
#https://github.com/XENON1T/MCAnalysisScripts/blob/master/G4EventDisplay/G4EventDisplay.ipynb
def get_parentids(event):
    '''Returns a trackid list of all the parent particles in the event.'''
    tracks = list(set(event.parentid))
    tracks.sort() # sort trackids
    return tracks

```

```
#Function adapted from Pietro Di Gangi
#https://github.com/XENON1T/MCAnalysisScripts/blob/master/
    G4EventDisplay/G4EventDisplay.ipynb
def particles_steps(event):

    ids = get_daughters_trackid(event, parentids=[0]) # get
        daughters of primary particle
    particles_by_step=[] # divide particles in hierarchic steps
    particles_by_step.append(ids)
    all_particles = ids # store id of particles
    #Loop over steps as long as a daughter particle is produced
    i=1
    while len(ids)>0:
        ids = get_daughters_trackid(event, parentids = ids)
        if len(ids)>0:
            particles_by_step.append(ids)
            all_particles = all_particles + ids
            i = i+1

    ids = []
    nsteps = i
    multiplicity = 1
    step_max = 0
    multiplicity_max = 1

    #Associate daughter particles to their parent
    for n in (range(0,nsteps)):
        set_ids = particles_by_step[n]
        particles = event[event.trackid.isin(set_ids)]
        parent_ids = get_parentids(particles)
        multiplicity = len(parent_ids) * multiplicity
        list_=[]

        for parent in parent_ids: # collect by parentid
            daughters = get_daughters_trackid(particles,[parent]) #
                set of daughters of this parent
            list_.append([parent,daughters])

            if len(daughters)>=multiplicity_max:
                multiplicity_max=len(daughters)
                step_max = n+2

        ids.append(list_)

    return ids

```

```

def particle_origin(df):

    particle_hierarchy = particles_steps(df)

    modified_particle_hierachy = particle_hierarchy.copy()

    # loop reversed over the hierarchy "layers"
    for i, layer in reversed(list(enumerate(particle_hierarchy))):

        #loop over the individual nodes in the layer
        for node in layer:

            #loop over the nodes in the layer above
            for j in range(len(particle_hierarchy[i-1])):

                #check if the parent particle is in the list of
                    daughter particles in the layer above
                if node[0] in particle_hierarchy[i-1][j][1]:

                    #append the trackids of the daughter particles
                    modified_particle_hierachy[i-1][j][1].extend(
                        node[1])


    #get all the needed informations from the dataframe
    primary_particle = np.unique(df[df.parenttype == "none"].type)
    particle_types = df.type.values
    particle_parent_types = df.parenttype.values
    particle_parent_id = df.parentid.values
    particle_id = df.trackid.values
    particle_created_by = ["0"] * len(particle_id)

    # This part will only work properly for my signal and my
        calibration sources Xe131m, Xe129m and Pb212
    p_list = ["e-", "gamma"]

    if len(np.intersect1d(primary_particle, p_list)) == 0:

        #If the prim particle is an excited state like Xe131m or
            Xe129m
        if "[" in primary_particle[0]:

            last_stage = 2

```

```
        #use this case if the primary particle is a nucleus, tested only for Pb212
        else :
            last_stage = 3

    #use this case if the prim particles are from an eventgenerator like decay0
    else:

        last_stage = 2

    # intercept a super rare case: Pb212 decay to the ground state and the electron with very very low energy
    # (no later particles created by the electron..)
    if len(modified_particle_hierachy) <=2:
        last_stage = 2

    # now loop over the necessesary first stages of the modified particle hierachy
    for i in range(1,last_stage):
        # and loop over the prim. particles of the stage

        for j in range(len(modified_particle_hierachy[i])):

            origin = modified_particle_hierachy[i][j][0]
            target = modified_particle_hierachy[i][j][1]

            #change particle_created_by to the particle type of the origin if the target is caused by the origin particle
            particle_created_by = np.where(np.isin(particle_id, target),
                                           particle_types[particle_id == origin][0],
                                           particle_created_by)

    # Now remove Pb212 and introduce gamma manually...
    # This is the "first generation" of particles comming fromt he Pb212 nucelus. These are the actual primary particles
    list_of_nuclei = ["Rn220","Po216", "Pb212", "Bi214"]
    particle_created_by = np.where(np.isin(particle_created_by, list_of_nuclei), particle_types, particle_created_by)

    #set the particle_created_by for the primary particles (created by geant4 or the event generator) to the
    #particle type itself
```

```python
    particle_created_by = np.where(particle_parent_id == 0,
        particle_types, particle_created_by)

    # A gamma always commes from an excited state in case of an
        radioactive decay.
    # These state are marked by Xx[Energy]. Set the
        particle_created_by to gamma for these cases.
    particle_created_by = np.where(["[" in element for element in
        particle_created_by], "gamma", particle_created_by)

    return particle_created_by

def cluster_function(x):
    d = {}
    #use the average position for each cluster weighted by energy
    d["xp"] = np.average(x["xp"], weights=x["ed"])
    d["yp"] = np.average(x["yp"], weights=x["ed"])
    d["zp"] = np.average(x["zp"], weights=x["ed"])
    d["time"] = np.average(x["time"], weights=x["ed"])
    #Sum the energy
    d["ed"] = np.sum(x["ed"])

    #check which initial particles created this cluster
    types_in_cluster = np.unique(x.created_by)
    nrg_sum = x.ed.sum()
    weights = []

    for p_type in types_in_cluster:

        nrg = x[x.created_by == p_type].ed.sum()
        weights.append(nrg/nrg_sum)

    if len(weights) ==0:

        types_in_cluster = ["gamma"]
        weights =[1]

    d["type"] = types_in_cluster
    d["weights"] = weights

    return pd.Series(d, index = ["xp", "yp", "zp", "time", "ed" ,"
        type", "weights"])


def weights_for_average_model(nrg):

```

```
    only_beta_below = 90
    only_gamma_above = 270

    if nrg < only_beta_below:

        return 0.

    elif (nrg > only_beta_below)&(nrg < only_gamma_above):

        y_1 = 0
        y_2 = 1
        x_1 = only_beta_below
        x_2 = only_gamma_above
        m = (y_2-y_1)/(x_2-x_1)
        y = m*(nrg-x_1)+y_1

        return y

    elif nrg > only_gamma_above:
        return 1.


@np.vectorize
def average_quanta_from_NEST(particle_type, en):

    nc = nestpy.NESTcalc(nestpy.VDetector())
    A = 131.293
    Z = 54.
    density = 2.862  # g/cm^3    #SR1 Value
    drift_field = 82  # V/cm     #SR1 Values


    if particle_type == "gamma":

        gamma_weight_for_average = weights_for_average_model(en)
        beta_weight_for_average = 1-gamma_weight_for_average

        #calculate the gamma and the beta model
        y_beta = nc.GetYields(nestpy.INTERACTION_TYPE(8),
                              en,
                              density,
                              drift_field,
                              A,
                              Z,
                              (1, 1))

```

```
        y_gamma = nc.GetYields(nestpy.INTERACTION_TYPE(7),
                               en,
                               density,
                               drift_field,
                               A,
                               Z,
                               (1, 1))

        y_average = y_gamma
        y_average.ElectronYield = np.average([y_beta.ElectronYield,
            y_gamma.ElectronYield], weights=[
            beta_weight_for_average, gamma_weight_for_average])
        y_average.PhotonYield = np.average([y_beta.PhotonYield,
            y_gamma.PhotonYield], weights=[beta_weight_for_average,
            gamma_weight_for_average])

        event_quanta = nc.GetQuanta(y_average, density)

        photons = event_quanta.photons
        electrons = event_quanta.electrons

    else:

        interaction_dict = {"e-": 8, "gamma": 7, "neutron": 0, "
            Kr83m": 11}

        y = nc.GetYields(nestpy.INTERACTION_TYPE(interaction_dict.
            get(particle_type, 7)),
                               en,
                               density,
                               drift_field,
                               A,
                               Z,
                               (1, 1))

        event_quanta = nc.GetQuanta(y, density)

        photons = event_quanta.photons
        electrons = event_quanta.electrons

    #lets do some modifications to the cs1 cs2 projection width
    photons_mod = photons * np.random.normal(loc=1, scale= 0.073)
    electrons = electrons + (photons-photons_mod)
    photons = photons_mod

    return photons, electrons
```

```
def quanta_generation(particle_type, nrg, cluster_weights):

    quanta = []

    for p_type, E, weights_in_cluster in zip(particle_type, nrg,
        cluster_weights):

        photons, electrons = average_quanta_from_NEST(p_type, E)

        mean_photons_in_cluster = np.average(photons, weights=
            weights_in_cluster)
        mean_electrons_in_cluster = np.average(electrons, weights=
            weights_in_cluster)

        quanta.append(mean_photons_in_cluster)
        quanta.append(mean_electrons_in_cluster)

    return quanta

@export
def read_g4(file, eps=0.005):

    print("Reading the geant4 file...")

    source = uproot.open(file)["generator"]["SourceType"]._fTitle.
        decode("UTF-8")
    data = uproot.open(file)["events/events"]
    df = data.pandas.df(["xp","yp", "zp", "time", "ed", "nsteps", "
        eventid", "type", "trackid", "parenttype", "parentid", "
        creaproc"])

    #Add the interaction type in the correct format
    df["type"] = np.concatenate(data.array("type"))
    df["type"] = df["type"].apply(lambda x: x.decode("UTF-8"))

    df["parenttype"] = np.concatenate(data.array("parenttype"))
    df["parenttype"] = df["parenttype"].apply(lambda x: x.decode("
        UTF-8"))

    df["creaproc"] = np.concatenate(data.array("creaproc"))
    df["creaproc"] = df["creaproc"].apply(lambda x: x.decode("UTF-8
        "))

    df["time"] = df["time"]*1e9 # conversion to ns
```

```

    #lets assign now the prim particles
    print("Find the initial particles")

    #A workaround for Co60 simulations.
    if source == "Co60":

        df["created_by"] = ["gamma"]*len(df)

    else:

        df["created_by"]  = np.concatenate(df.groupby(["entry"]).
            apply(lambda x: particle_origin(x)))

    #Remove all values without energy depositon
    df = df[df.ed != 0 ]
    print("Time Clustering")
    #Time Clustering
    time_scale = 10 #ns
    dbscan_time_clustering = DBSCAN(eps=time_scale, min_samples=2)

    df["time_cluster"] = np.concatenate(df[["time"]].groupby("entry
        ").apply(lambda x: dbscan_time_clustering.fit_predict(x.time
        .values.reshape(-1,1)) ))
    print("Microclustering")
    #Cluster in xyz for each event(entry) and each time_cluster
    dbscan_clustering = DBSCAN(eps=eps, min_samples=2)
    df["cluster"] = np.concatenate(df.groupby(["entry", "
        time_cluster"]).apply(lambda x:  dbscan_clustering.
        fit_predict(np.stack(x[["xp", "yp", "zp"]].values))).values)

    #Apply the clustering for each event(entry), time_cluster and
        cluster
    df = df.groupby(["entry","time_cluster","cluster"]).apply(
        lambda x: cluster_function(x))
    print("Clustering done!")
    df.xp /=10
    df.yp /=10
    df.zp /=10

    #Limit the interactions to the TPC
    tpc_radius_square = 2500
    z_lower = -100
    z_upper = 0
    df = df[(df.xp**2+df.yp**2<=47.9**2)&(df.zp<z_upper)&(df.zp>
        z_lower)]
```

```
    #Sort the df for each event in time
    df["index_dummy"] = df.index.get_level_values(0)
    df = df.sort_values(["index_dummy", "time"])

    #set time of first e deposition in each event to 0
    df["time"] = df.groupby(["entry"]).apply(lambda x: (x["time"]-x["time"].min())).values

    # set the index to start from 0 and run to the final number of events
    # This is important for simulations with external sources where a lot of events never reach the xenon
    idx = df.index.get_level_values(0)
    idx_lims = np.append(np.intersect1d(idx,np.unique(idx),return_indices = True)[1],len(idx))
    n_vals = [idx_lims[i+1] - idx_lims[i] for i in range(0,len(idx_lims)-1)]
    new_entry_idx = pd.Int64Index(np.repeat(np.arange(len(n_vals)), n_vals), dtype = "int64", name = "entry")

    df.index = pd.MultiIndex.from_arrays([new_entry_idx,
                                          df.index.get_level_values(1).values]
                                          )

    #lets reset the time cluster index as well
    new_cluster_idx = pd.Int64Index(np.concatenate(df.groupby("entry").apply(lambda x: np.arange(len(x.index.get_level_values(1).values)))), dtype = "int64", name = "cluster")

    df.index = pd.MultiIndex.from_arrays([df.index.get_level_values(0).values,
                                          new_cluster_idx]
                                          )

    #and separate the events in time by one second
    event_spacing = 1e9
    df.time = np.cumsum(df.time+ (df.index.get_level_values(1).values == 0 )*event_spacing)

    #build the instructions
    n_instructions = len(df)
    ins = np.zeros(2*n_instructions, dtype=instruction_dtype)

```

```
    #shift the time by a constant offset...
    e_dep, ins['x'], ins['y'], ins['z'], ins['t'] = df.ed.values, \
                                                    np.repeat(df.xp.values, 2), \
                                                    np.repeat(df.yp.values, 2), \
                                                    np.repeat(df.zp.values, 2), \
                                                    np.repeat(df.time.values, 2)

    ins["e_dep"] = np.repeat(e_dep, 2)
    ins["event_number"] = np.repeat(df.index.get_level_values(0).values,2)

    ins['type'] = np.tile((1, 2), n_instructions)

    #NEST handling...
    #https://github.com/NESTCollaboration/nest/blob/master/src/NEST.cpp
    #line 406
    #https://arxiv.org/pdf/1106.1613.pdf page 18
    if "Kr83" in source:
        print("Kr83m in source, modify e_dep and types")

        #To use the special 9.4 keV case in NEST the energy has to be exactly 9.4 keV
        e_dep = np.where((e_dep<9.5)&(e_dep>9.3), 9.4 , e_dep)
        #change the particle type for the 9.4 keV line to Kr83m
        df["type"] = [["Kr83m"]]*len(e_dep)
        df["weights"] = [[1]]*len(e_dep)

    recoil_dict = {"e-": "er", "gamma": "er", "neutron": "nr", "alpha": "alpha"}
    ins['recoil'] =np.repeat([recoil_dict.get(particle[np.argmax(weight)], "er") for particle, weight in zip(df["type"].values, df["weights"].values)],2)

    print("Calculate Quanta with NEST")
    quanta = quanta_generation(df["type"].values, e_dep,df["weights"].values )

```

```
    ins['amp'] = quanta
    ins["created_by"] = np.repeat([particle[np.argmax(weight)] for
        particle, weight in zip(df["type"].values, df["weights"].
        values)],2)
    #lets cut interactions without electrons or photons
    ins = ins[ins["amp"] > 0]

    print("Done!")
    return ins
```

A.3 Calibration Source Event Selection

A.3.1 ^{129m}Xe and ^{131m}Xe

Table A.1: Cut history of simulated ^{129m}Xe events. Two data quality cuts are applied and the mono energetic source is selected by cuts in cS1, cS2 and energy.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	32 (0.01 %)	499462 (99.99 %)	99.99 %
1T fiducial volume	265770 (53.21 %)	233692 (46.79 %)	46.79 %
Source Selection			
700 pe < cS1 < 2500 pe	1 (0.01 %)	233691 (99.99 %)	46.79 %
4×10^4 pe < cS2 < 4×10^5 pe	6 (0.01 %)	233685 (99.99 %)	46.78 %
190 keV < E < 260 keV	270 (0.12 %)	233415 (99.88 %)	46.73 %

Table A.2: Cut history of simulated ^{131m}Xe events. A similar set of cuts is applied compared to the selection of ^{129m}Xe events but the energy region is varied to match the differing decay energy.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	19 (<0.01 %)	496751 (>99.99 %)	99.99 %
1T fiducial volume	263983 (53.14 %)	232768 (46.86 %)	46.86 %
Source Selection			
700 pe < cS1 < 2500 pe	0 (0 %)	232768 (100 %)	46.86 %
4×10^4 pe < cS2 < 4×10^5 pe	304 (0.13 %)	232464 (99.87 %)	46.80 %
130 keV < E < 190 keV	22 (0.01 %)	232442 (99.99 %)	46.79 %

Table A.3: Cut history for the selection of ^{129m}Xe and ^{131m}Xe events in the background data of XENON1T. The data used here is loaded technically different from the simulated events and the S1 > 0 pe, S2 > 0 pe cut is applied implicitly so that no number can be given here.

Cut	Events removed	Events passed	Fraction left
Data Quality Cuts			
DAQ Veto	1070018 (1.06 %)	99843936 (98.94 %)	98.94 %
S2 Tails	4031907 (4.04 %)	95812029 (95,96 %)	94.18 %
Muon Veto	1011325 (1.06 %)	94800704 (98.94 %)	93.18 %
Flash	1019 (<0.01 %)	94799685 (>99.99 %)	93.18 %
1T fiducial volume	86461272 (91.20 %)	8338413 (8.80 %)	8.20 %
General Source Selection			
700 pe < cS1 < 2500 pe	8150754 (97.95 %)	170669 (2.05 %)	0.17 %
4×10^4 pe < cS2 < 4×10^5 pe	31974 (18.73 %)	138695 (81.27 %)	0.14 %
^{131m}Xe			
130 keV < E < 190 keV	107143 (77.25 %)	31552 (22.75 %)	0.03 %
^{129m}Xe			
190 keV < E < 260 keV	107482 (77.50 %)	31213 (22.50 %)	0.03 %

A.3.2 ^{212}Pb

Table A.4: Cut history of simulated ^{212}Pb decays. Four source specific cuts were used to separate the events of interest. Since only ^{212}Pb decays were simulated, the cuts remove only a small fraction of events.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	2268 (0.49 %)	457288 (99.51 %)	99.51 %
1T fiducial volume	228219 (49.91 %)	229069 (50.09 %)	49.85 %
Source Selection			
cS1 < 5000 pe	0 (0 %)	229069 (100 %)	49.85 %
cS2 > 10 000 pe	4299 (1.88 %)	224770 (98.12 %)	48.91 %
$2.3 < \ln\left(\frac{cS2}{cS1}\right) < 5.5$	33 (0.01 %)	224737 (99.99 %)	48.90 %
E < 570 keV	40 (0.02 %)	224697 (99.98 %)	48.89 %

Table A.5: Cut history for ^{220}Rn calibration data. Beside the data quality cuts the same source specific cuts used for the simulated data is applied here to select ^{212}Pb decays in TPC Data. The large fraction of events removed by the fiducial volume cut compared to the 50 % expectation from geometrical effects could be caused by other external sources and by the inlet of the ^{220}Rn in the TPC (cf. figure 4.16).

Cut	Events removed	Events passed	Fraction left
Data Quality Cuts			
S1 > 0 pe, S2 > 0 pe	16846910 (54.66 %)	13971868 (45.34 %)	45.34 %
DAQ Veto	1590225 (11.38 %)	12381643 (88.62 %)	40.18 %
S2 Tails	582871 (4.71 %)	11798772 (95.29 %)	38.28 %
Muon Veto	63854 (0.54 %)	11734918 (99.46 %)	38.08 %
Flash	0 (0 %)	11734918 (100 %)	38.08 %
1T fiducial volume	9547063 (81.36 %)	2187855 (18.64 %)	7.10 %
Source Selection			
cS1 < 5000 pe	966899 (44.20 %)	1220956 (55.80 %)	3.96 %
cS2 > 10 000 pe	35825 (2.93 %)	1185131 (97.07 %)	3.85 %
$2.3 < \ln\left(\frac{\mathrm{cS2}}{\mathrm{cS1}}\right) < 5.5$	26946 (2.27 %)	1158185 (97.73 %)	3.76 %
E < 570 keV	23995 (2.07 %)	1134190 (97.93 %)	3.68 %

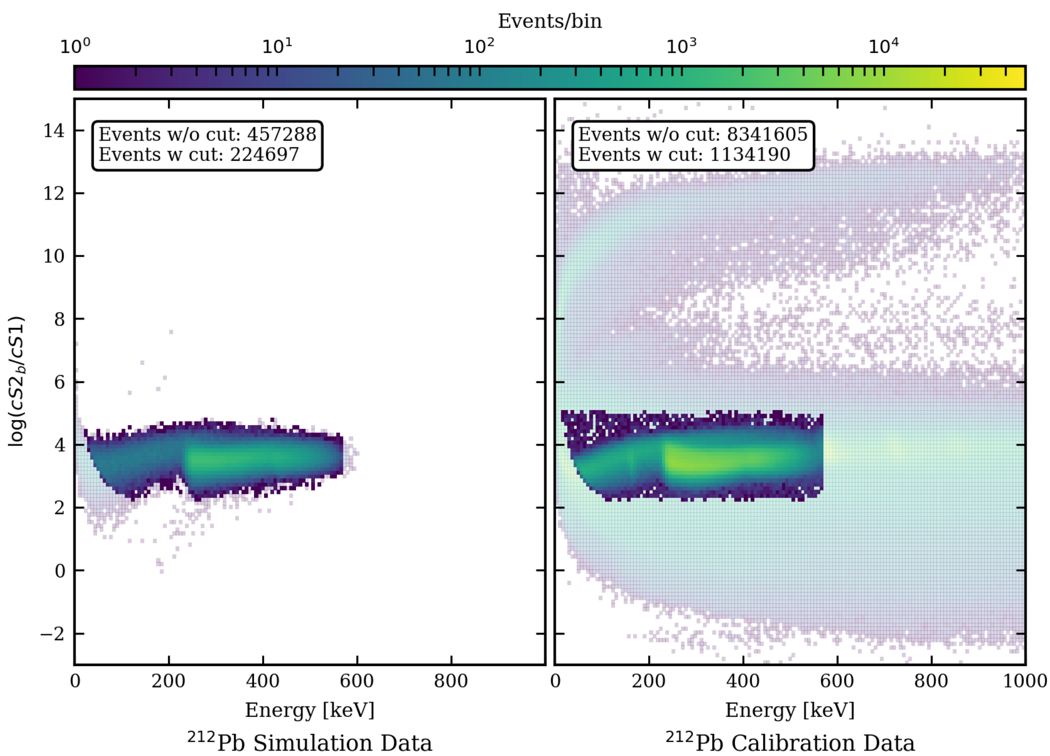


Figure A.1: Two dimensional histograms of simulated (left) and measured (right) ^{212}Pb calibration data in energy-ln($\mathrm{cS2_b/cS1}$) space. Bins of events that are removed by cuts are shown with a decreased opacity.

A.3.3 ^{60}Co

Table A.6: Cut history for simulated ^{60}Co events of the 1332.5 keV and 1173.2 keV γ-lines.

Cut	Events removed	Events passed	Fraction left
Data Quality Cut			
S1 > 0 pe, S2 > 0 pe	156412 (10.30 %)	1361967 (89.70 %)	89.70 %
1T fiducial volume	1301662 (95.57 %)	60305 (4.43 %)	3.97 %
General Source Selection			
2.5×10^5 pe < $cS2_b$ < 1×10^6 pe	9056 (15.02 %)	51249 (84.98 %)	3.38 %
4×10^3 pe < cS1 < 1×10^4 pe	4685 (9.14 %)	46564 (90.86 %)	3.07 %
Multiplicity > 1	13420 (28.82 %)	33144 (71.18 %)	2.18 %
1173.2 keV ^{60}Co line			
1150 keV < E < 1250 keV	21825 (65.85 %)	11319 (34.15 %)	0.75 %
1332.5 keV ^{60}Co line			
1300 keV < E < 1500 keV	20157 (60.82 %)	12987 (39.18 %)	0.86 %

Table A.7: Cut history for the selection of the 1332.5 keV and 1173.2 keV ^{60}Co lines from the measured background data.

Cut	Events removed	Events passed	Fraction left
Data Quality Cuts			
S1 > 0 pe, S2 > 0 pe	32820744 (27.71 %)	85625828 (72.29 %)	72.29 %
DAQ Veto	901575 (1.05 %)	84724253 (98.95 %)	71.53 %
S2 Tails	3319739 (3.92 %)	81404514 (96.08 %)	68.73 %
Muon Veto	861819 (1.06 %)	80542695 (98.94 %)	68.00 %
Flash	860 (<0.01 %)	80541835 (>99.99 %)	68.00 %
1T fiducial volume	76450819 (94.92 %)	4091016 (5.08 %)	3.45 %
General Source Selection			
2.5×10^5 pe < $cS2_b$ < 1×10^6 pe	1543775 (37.74 %)	2547241 (62.26 %)	2.15 %
4×10^3 pe < cS1 < 1×10^4 pe	420188 (16.50 %)	2127053 (83.50 %)	1.79 %
Multiplicity > 1	541664 (25.47 %)	1585389 (74.53 %)	1.34 %
1173.2 keV ^{60}Co line			
1130 keV < E < 1210 keV	1256986 (79.29 %)	328403 (20.71 %)	0.28 %
1332.5 keV ^{60}Co line			
1300 keV < E < 1360 keV	1273819 (80.35 %)	311570 (19.65 %)	0.26 %

A.4 Discrimination Space

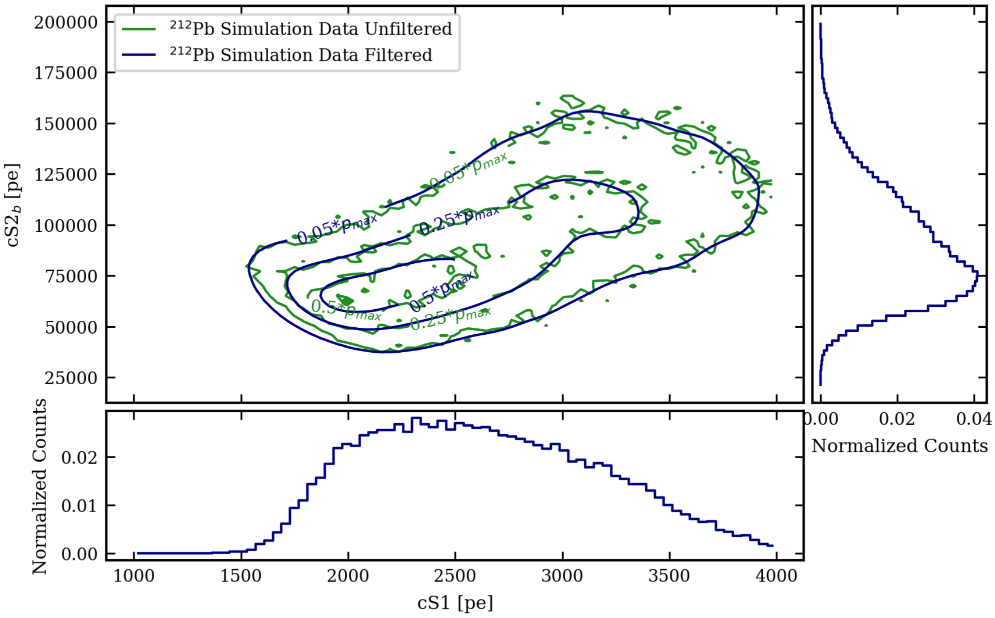


Figure A.2: Contour Plot for filtered and unfiltered simulated ^{212}Pb data with projections to cS1 and cS2$_b$. A Gaussian filter is applied to smooth the contour lines and suppress small scale structures.

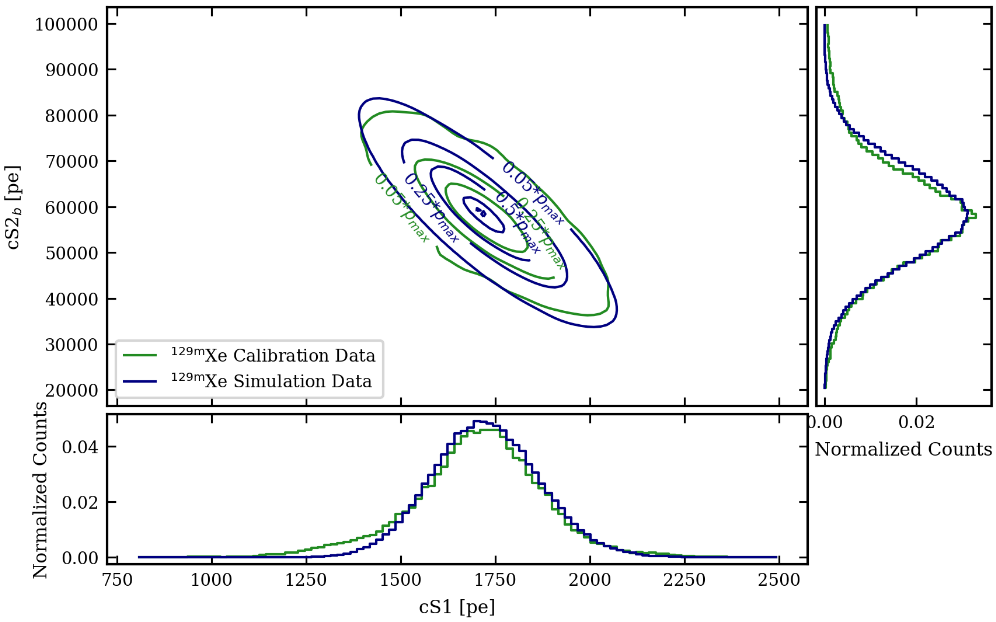


Figure A.3: Contour plot for cS1 and cS2$_b$ of ^{129m}Xe events with the simulated data in blue and calibration data in green.The contour lines are calculated from normalized histograms smoothed with a Gaussian filter while the filter is not applied to the projections.

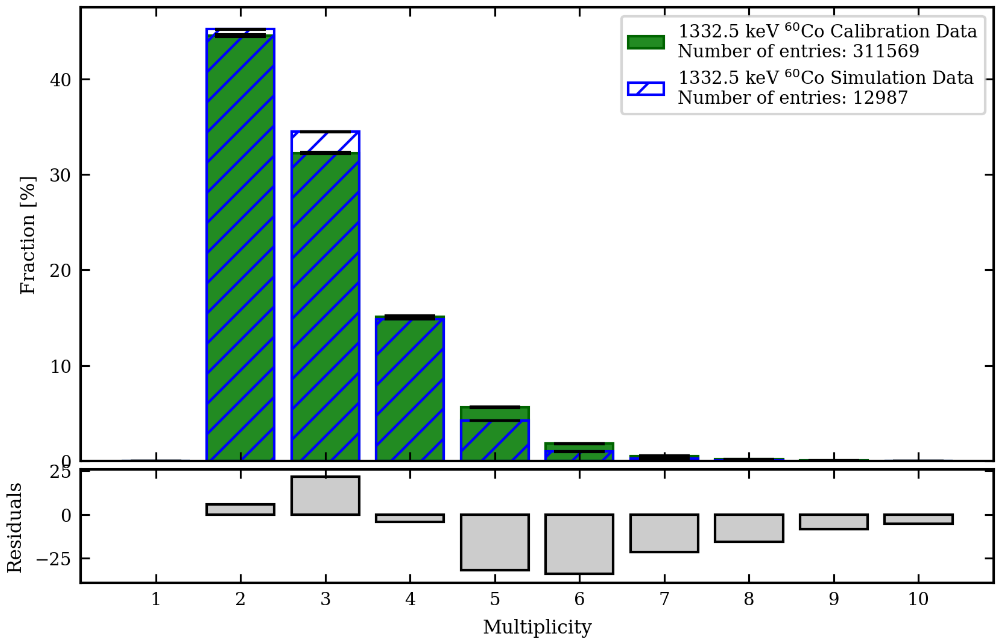


Figure A.4: Histograms of the multiplicity for simulated 1332.5 keV ^{60}Co γ-rays in the fiducial volume in blue and calibration data in green. A multiplicity is was applied to remove single-site events. Error-bars are drawn in black.

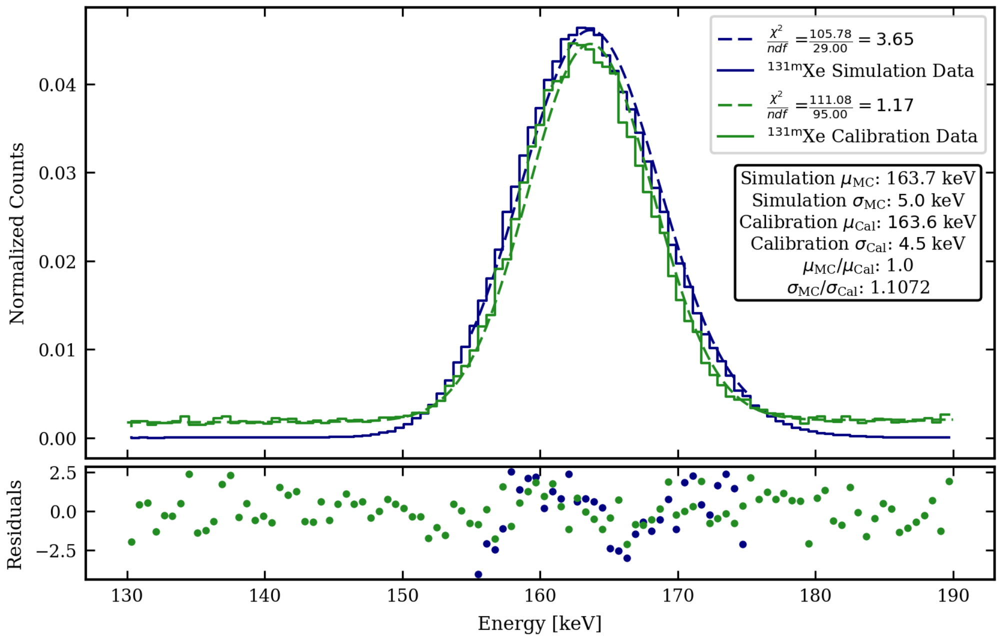


Figure A.5: Energy of simulated (blue) and measured (green) ^{131m}Xe events. Error-bars are not drawn to reduce the complexity of the plot. The mono-energetic 163.9 keV γ line is fitted with a Gaussian function. A constant term was added for the measured data to account for background events. The fit range was limited to ensure that each bin has enough counts.

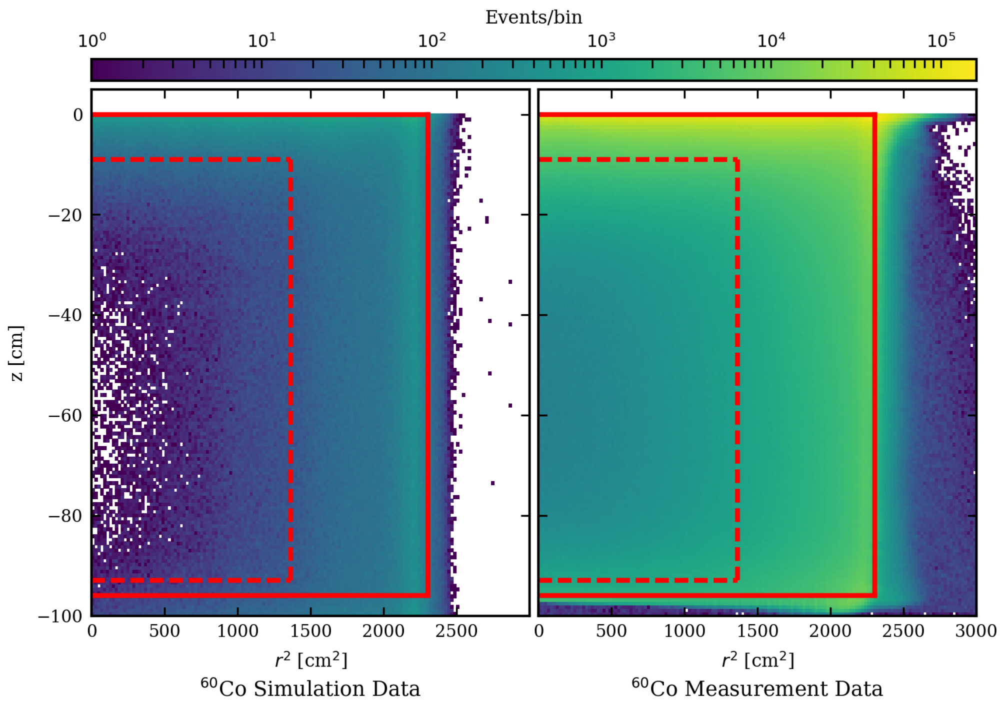


Figure A.6: z-r^2 distribution of ^{60}Co events in the detector for simulated events on the left side and measured data on the right side. The TPC boundaries are given as solid red lines. The fiducial volume is shown with dashed red lines.

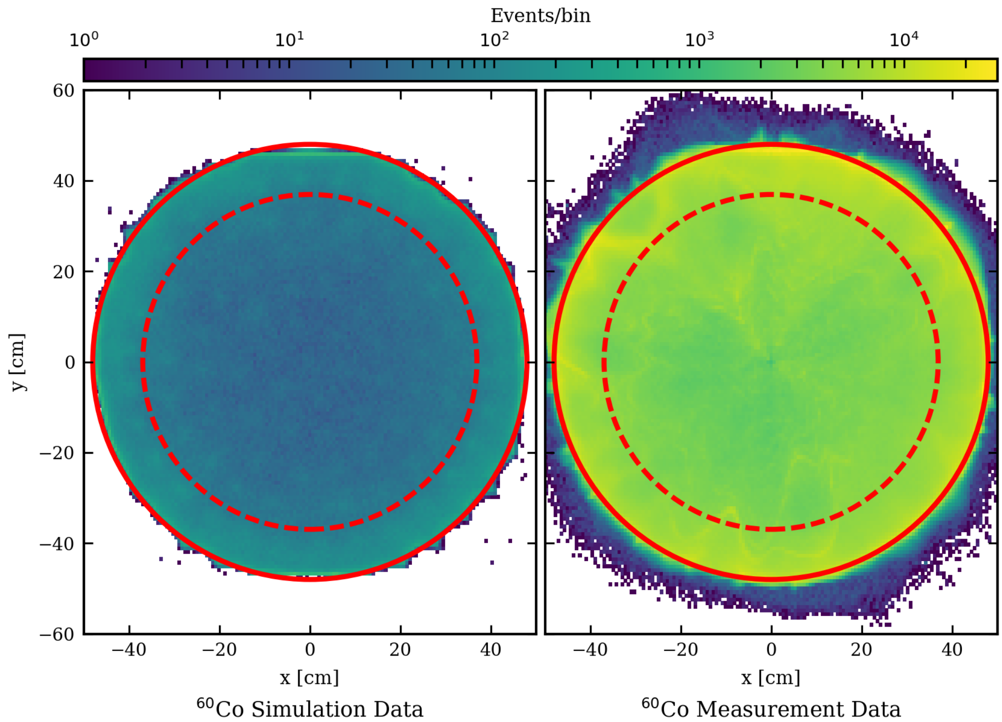


Figure A.7: The x-y distribution of ^{60}Co events in the detector for simulated events on the left side and measured data on the right side of the plot. The approximate dimensions of the TPC are shown as the solid red line and the fiducial volume as the dashed red line.

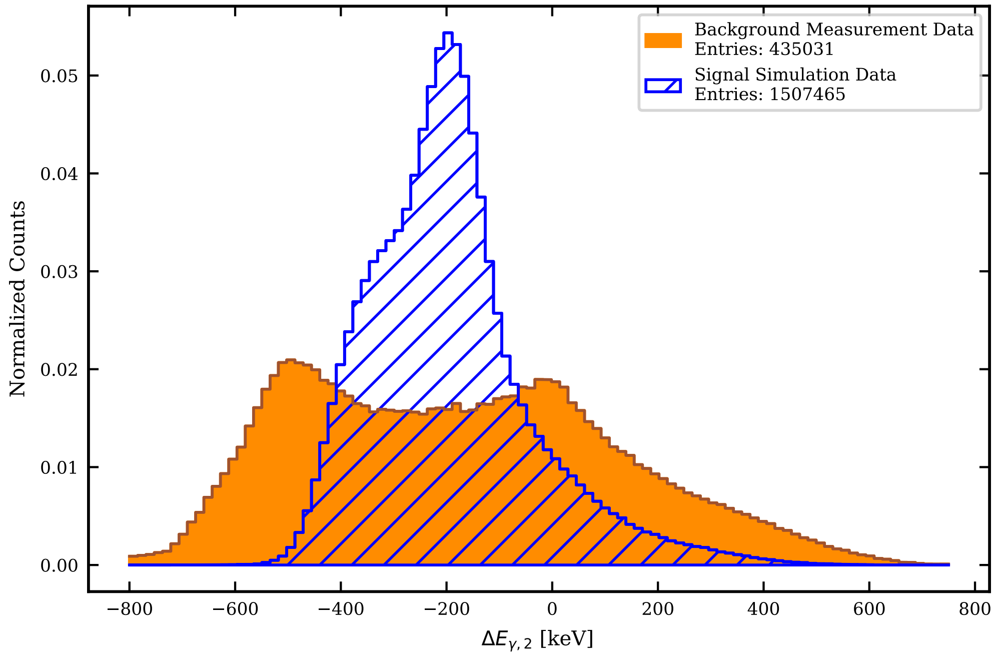


Figure A.8: Comparison of $\Delta E_{\gamma,2}$ for simulated signal events in blue and measured background events in orange. The signal data show a peak-like structure at around -200 keV whereas the background spectrum spreads in a double-peak structure over a broad energy range.

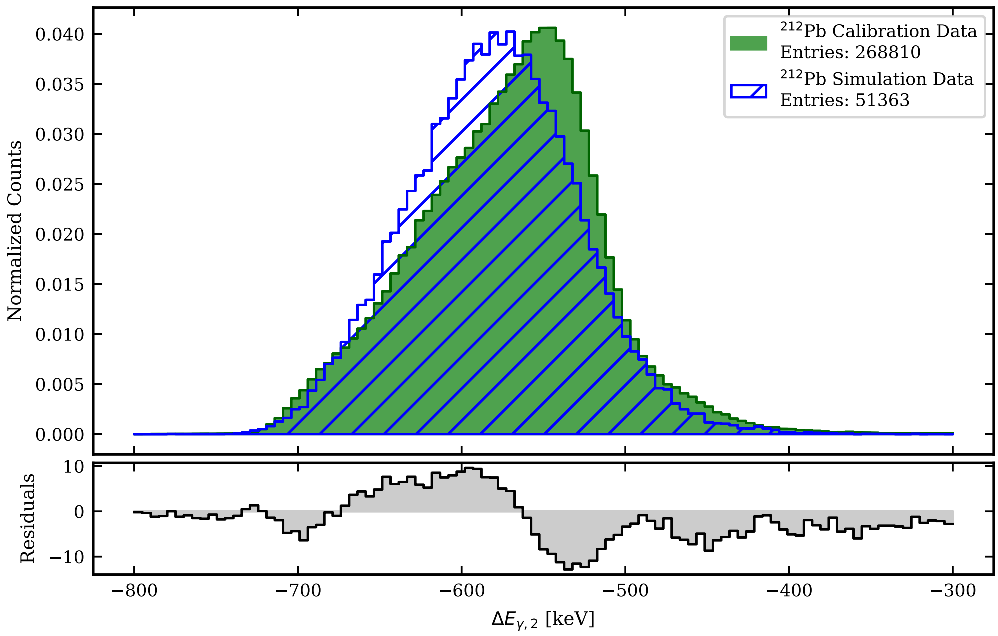


Figure A.9: Normalized histograms of $\Delta E_{\gamma,2}$ for simulated ^{212}Pb events in blue and calibration data in green. Error-bars are not drawn to reduce the complexity of the plot. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. The simulation tends towards smaller $\Delta E_{\gamma,2}$ values compared to the calibration data.

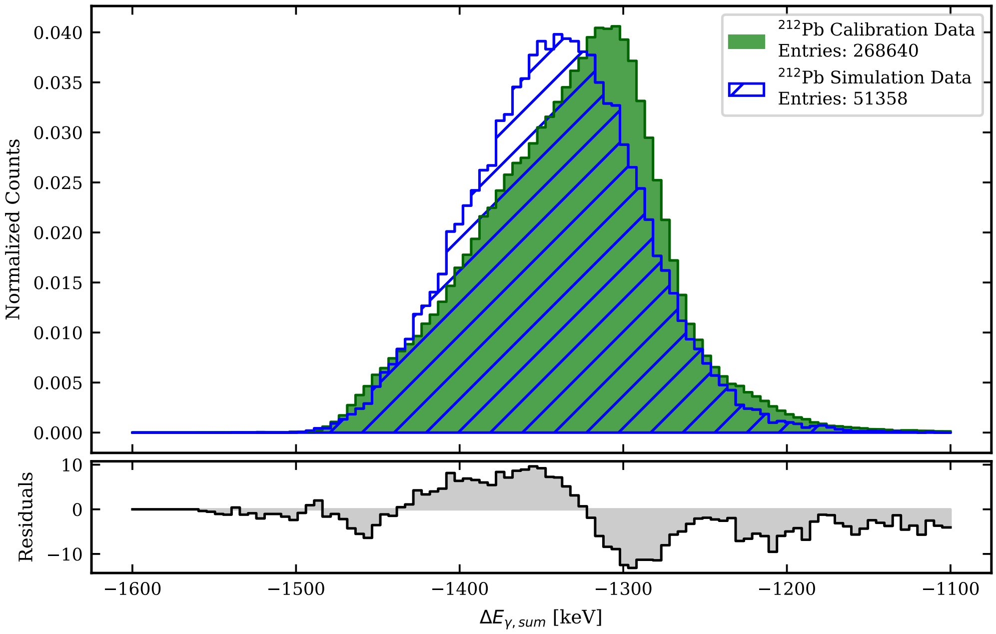


Figure A.10: Normalized histograms of $\Delta E_{\gamma,\mathrm{sum}}$ for simulated ^{212}Pb events in blue and calibration data in green. Error-bars are not drawn to reduce the complexity of the plot. Only small relative uncertainties are present for each bin due to the high statistics of both datasets. The simulation tends towards smaller $\Delta E_{\gamma,\mathrm{sum}}$ values compared to the calibration data.

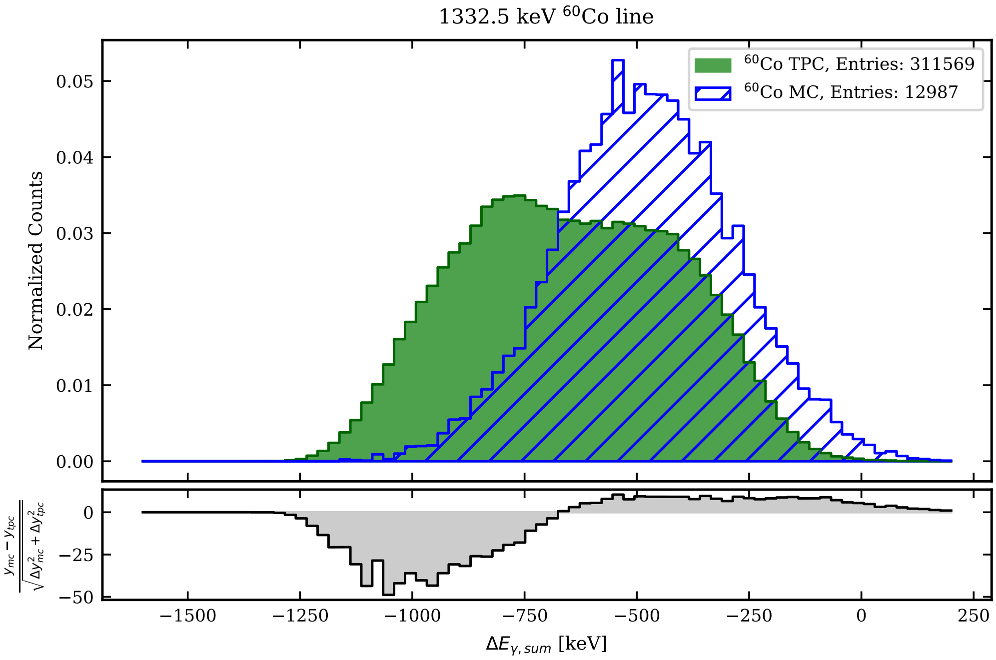


Figure A.11: Normalized histograms of $\Delta E_{\gamma,\mathrm{sum}}$ for simulated 1332.5 keV ^{60}Co events in blue and measured events in green. Error-bars are not drawn to reduce the complexity of the plot. The simulated data are Gaussian shaped whereas the measured data show a double peak structure. Deviations in $\Delta E_{\gamma,\mathrm{sum}}$ are caused by differences between the NEST prediction of the charge yield and the observed yields in XENON1T data.

A.5 Effect of Input Parameter on Discriminator Prediction

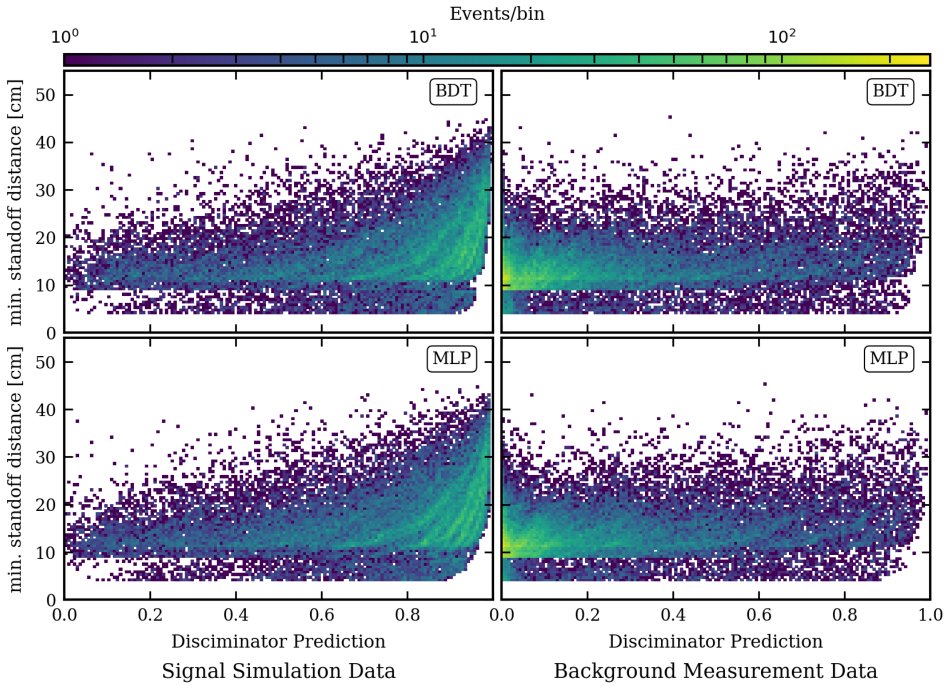


Figure A.12: Two-dimensional histograms of the standoff distance and discriminator prediction obtained with the BDT (top) and MLP (bottom) for signal (left) and background events (right).

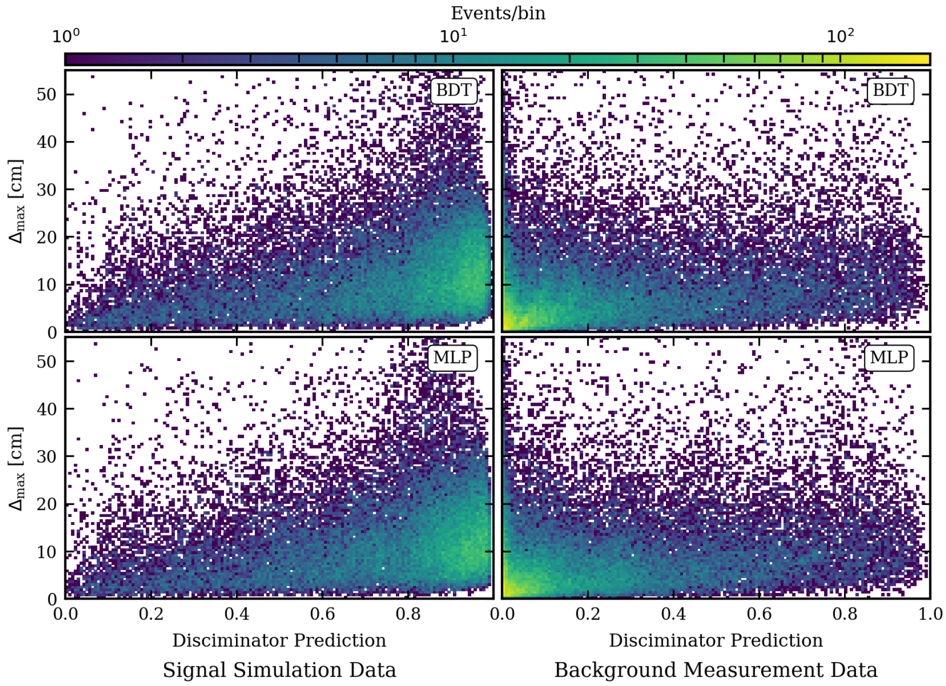


Figure A.13: Two-dimensional histograms of the Δ_{mac} parameter and discriminator prediction obtained with the BDT (top) and MLP (bottom) for signal (left) and background events (right).

List of abbreviations

LXe liquid xenon

GXe gaseous xenon

PMT photomultiplier tube

ER electronic recoil

NR nuclear recoil

WIMP weakly interacting massive particle

LNGS Laboratori Nazionali del Gran Sasso

TPC time projection chamber

PTFE polytetrafluorethylen

NN neural network

FFN feed forward network

MLP multi-layer perceptron

ReLU rectified linear unit

ML machine learning

BDT boosted decision tree

MC Monte Carlo

DBSCAN Density Based Spatial Clustering for Applications with Noise

PAX Processor for Analysing Xenon

NEST Noble Element Simulation Technique

PDF probability density function

AI artificial intelligence

SR science run

LCE light collection efficiency

FAX Fake Xenon Experiment

PAX Processor for Analysing XENON

HAX Handy Analysis for XENON

NME nuclear matrix element

Bibliography

[1] M. Goeppert-Mayer. Double beta-disintegration. *Phys. Rev.*, 48:512–516, Sep 1935. 3

[2] E. Fermi. Tentativo di una teoria dei raggi β. *Il Nuovo Cimento (1924-1942)*, 11(1):1, 1934. 3

[3] SR Elliott, AA Hahn, and MK Moe. Direct evidence for two-neutrino double-beta decay in se 82. *Physical Review Letters*, 59(18):2020, 1987. 3

[4] C. F. v. Weizsäcker. Zur theorie der kernmassen. *Zeitschrift für Physik*, 96(7):431–458, 1935. 3

[5] Ruben Saakyan. Two-neutrino double-beta decay. *Annual Review of Nuclear and Particle Science*, 63(1):503–529, 2013. 4, 5, 6, 8

[6] K Zuber. Double beta decay. *Contemporary Physics*, 45(6):491–502, 2004. 4, 5

[7] Particle Data Group et al. M. tanabashi et al. *Phys. Rev. D*, 98:030001, 2018. 4, 5

[8] Michelle J. Dolinski, Alan W.P. Poon, and Werner Rodejohann. Neutrinoless double-beta decay: Status and prospects. *Annual Review of Nuclear and Particle Science*, 69(1):219–251, 2019. 5

[9] Paul Adrien Maurice Dirac. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 114(767):243–265, 1927. 5

[10] Enrico Fermi. *Nuclear physics: a course given by Enrico Fermi at the University of Chicago*. University of Chicago Press, 1950. 5

[11] N. Ackerman et al. Observation of two-neutrino double-beta decay in xe 136 with the exo-200 detector. *Physical Review Letters*, 107(21):212501, 2011. 6

[12] JB Albert et al. Improved measurement of the two-neutrino double-beta half-life of136xe with the exo-200 detector. *Physical Review C*, 89(1), Jan 2014. 6

[13] IAEA Nuclear Data Section. Live chart of nuclides. `https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html`, 2020. [Online; last accessed 10.04.2020]. 6, 10, 21, 29, 33, 34, 35, 36, 48, 58

[14] JB Albert et al. Search for 2 ν β β decay of xe 136 to the 0 1+ excited state of ba 136 with the exo-200 liquid xenon detector. *Physical Review C*, 93(3):035501, 2016. 6, 7, 45

[15] Michael Duerr, Manfred Lindner, and Kai Zuber. Consistency test of neutrinoless double beta decay with one isotope. *Physical Review D*, 84(9), Nov 2011. 6

[16] K. Asakura et al. Search for double-beta decay of 136xe to excited states of 136ba with the kamland-zen experiment. *Nuclear Physics A*, 946:171–181, 2016. 7

[17] Commission on Isotopic Abundances CIAAW and Atomic Weights. Ciaaw. atomic weights of the elements 2019. `https://www.ciaaw.org/atomic-weights.htm`, 2019. [Online; accessed 22.03.2020]. 7, 10, 21, 89

[18] Alexander Fieguth. *First observation of double electron capture in Xe-124 and detection prospects for underlying nuclear interaction mechanisms in direct dark matter search*. PhD thesis, Westfälische Wilhelms-Universität Münster, 2018. 7, 89

[19] E. Aprile et al. The xenon1t dark matter experiment. *The European Physical Journal C*, 77(12):881, 2017. 8, 10, 11

[20] E. Aprile et al. Dark matter search results from a one ton-year exposure of xenon1t. *Physical Review Letters*, 121(11), Sep 2018. 8, 13

[21] Lutz Althüser. Light collection efficiency simulations of the xenon1t experiment and comparison to data. 9, 17, 23, 24

[22] E. Aprile et al. Physics reach of the xenon1t dark matter experiment. *Journal of Cosmology and Astroparticle Physics*, 2016(04):027, 2016. 8, 17, 23

[23] E. Aprile and T. Doke. Liquid xenon detectors for particle physics and astrophysics. *Reviews of Modern Physics*, 82(3):2053–2097, Jul 2010. 8, 9, 54

[24] E. Aprile et al. Xenon1t dark matter data analysis: Signal reconstruction, calibration, and event selection. *Physical Review D*, 100(5):052014, 2019. 9, 10, 24

[25] M.J. Berger et al. Nist xcom: Photon cross sections database. 10, 66

[26] The XENON collaboration. The xenon1t experiment. `http://xenonexperiment.org`. [Online; last accessed 19.04.2020]. 11

[27] S. Agostinelli et al. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003. 16, 17

[28] Vladimir I. Tretyak. Decay0 event generator for initial kinematics of particles in alpha, beta and double beta decays. 2015. 16

[29] F. James and M. Roos. Minuit – a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, 10:343–367, December 1975. 17

[30] iminuit team. iminuit – a python interface to minuit. `https://github.com/scikit-hep/iminuit`. Accessed: 2018-03-05. 17

[31] M. Ester et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996. 19, 40

[32] F. Pedregosa et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 20, 73

[33] A Mozumder. Free-ion yield and electron-ion recombination rate in liquid xenon. *Chemical physics letters*, 245(4-5):359–363, 1995. 20

[34] M. Szydagis et al. Nest: a comprehensive model for scintillation yield in liquid xenon. *Journal of Instrumentation*, 6(10):P10002, 2011. 20

[35] Chris Tunnell and Nicholas Carrara. NESTCollaboration/nestpy: Improve user friendlyness. `https://doi.org/10.5281/zenodo.3360721`, August 2019. 21

[36] M. Galloway. Estimate of xenon1t target mass. Internal Note. 21, 89

[37] G. Rischbieter. The noble element simulation technique (nest) version 2.0. Identification of Dark Matter conference 2018. 21

[38] E. Aprile et al. The pax data processor v6.8.0. `https://doi.org/10.5281/zenodo.1195785`, March 2018. 23

[39] E. Aprile et al. Energy resolution and linearity in the kev to mev range measured in xenon1t, 2020. 24, 46, 54

[40] E. Aprile et al. Xenon1t/hax: hax referenced in chep. `https://doi.org/10.5281/zenodo.1464018`, October 2018. 24

[41] C. Therreau and T. Zhu. High energy multiple scatters summary i. Internal Note. 24

[42] Miguel Ángel Vargas Jara. *Data Analysis in the XENON1T Dark Matter Experiment*. PhD thesis, Westfälische Wilhelms-Universität Münster, 2019. 25

[43] D. Coderre. Daq veto cut for science run 1. Internal Note. 25

[44] D. Coderre. S2 tail cut summary. Internal Note. 26

[45] O. Wack. Selection of flashes and effect on exposure. Internal Note. 26

[46] K. Ni er al. Preparation of neutron-activated xenon for liquid xenon detector calibration. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 582(2):569 – 574, 2007. 29

[47] M. Mitchell. *Artificial Intelligence: A Guide for Thinking Humans.* Farrar, Straus and Giroux, 2019. 39

[48] Francois Chollet. *Deep Learning with Python.* Manning Publications Co., USA, 1st edition, 2017. 39, 40, 41

[49] Payam Refaeilzadeh, Lei Tang, and Huan Liu. *Cross-Validation*, pages 532–538. Springer US, Boston, MA, 2009. 41

[50] Roger Grosse. Intro to neural networks and machine learning. `https://www.cs.toronto.edu/~rgrosse/courses/csc321_2018/`, 2018. [Online; last accessed 14.04.2020]. 42

[51] M. Leshno et al. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural networks*, 6(6):861–867, 1993. 43

[52] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014. 43

[53] M. Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. 44, 72

[54] Zhi-Hua Zhou. *Ensemble Methods: Foundations and Algorithms.* Chapman and Hall, 1st edition, 2012. 44

[55] Sotiris B Kotsiantis. Decision trees: a recent overview. *Artificial Intelligence Review*, 39(4):261–283, 2013. 44

[56] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001. 44

[57] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016. 44, 75

[58] P. Virtanen et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 2020. 70

[59] François Chollet et al. Keras. `https://keras.io`, 2015. 72

[60] N. Srivastava et al. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014. 73

[61] XGBoost Developers. Xgboost parameters. `https://xgboost.readthedocs.io/en/latest/parameter.html`, 2020. [Online; accessed 06.04.2020]. 75

[62] C. Weinheimer. Private Communication. 87

[63] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Physical Review D*, 57(7):3873–3889, Apr 1998. 88

[64] Wayne W. LaMorte. Computing percentiles. http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Probability/BS704_Probability10.html, 2016. [Online; last accessed 14.04.2020]. 88

[65] Christian Wittweg, Brian Lenardo, Alexander Fieguth, and Christian Weinheimer. Detection prospects for the second-order weak decays of ^{124}xe in multi-tonne xenon time projection chambers, 2020. 93