Studies of Calibration and Electron Recoil Background Modelling for the XENON100 Dark Matter Experiment

DISSERTATION ZUR ERLANGUNG DES GRADES
"DOKTOR DER NATURWISSENSCHAFTEN"
AM FACHBEREICH PHYSIK, MATHEMATIK UND INFOMATIK
DER JOHANNES GUTENBERG-UNIVERSITÄT IN MAINZ

Vorgelegt von: Boris Bauermeister

Johannes-Gutenberg-Universität Mainz/Germany

ACKNOWLEDGEMENTS

My sincerely thanks go to my girlfriend who always supported me and my plans. I thank her in particular for her absence in the first half of 2015 that I had enough time to compose my thesis.

ABSTRACT

According to cosmological studies, there are hints for the existence of *Dark Matter*. Latest results from the Planck satellite mission suggest that our universe consist of 4.9% baryonic matter, 26.8% Dark Matter and 68.3% Dark Energy. Many theories, such as a super symmetric extension (SUSY) of the standard model, provide a natural candidate for particle Dark Matter which has the following properties: It is charge-less and has a mass but its interaction probability with baryonic matter is small. This particle is summarised as WIMP, a weakly interacting massive particle.

In order to detect the WIMP, the XENON collaboration started in 2007 to build an experiment which is designed for direct Dark Matter detection. The experimental setup is located at Laboratori Nazionali del Gran Sasso (Italy). The detector is a two-phase Time Projection Chamber (TPC) filled with 161 kg of liquid xenon that aims at observing Dark Matter by looking at nuclear recoils produces by WIMPs scattering off xenon nucleons. The XENON collaboration archive the best limit in 2012 on the spin-independent WIMP-nucleon cross-section with a minimum of $\sigma_\chi = 2 \times 10^{-45}\,\mathrm{cm}^2$ at a WIMP mass of $m_\chi = 55\,\mathrm{GeV/c^2}$ (90% C.L.). In 2013, using the same Dark Matter data, the XENON collaboration published a more stringent limit for the spin-dependent cross-section minimum of $\sigma_\chi = 3.5 \times 10^{-40}\,\mathrm{cm}^2$ at a WIMP mass of $m_\chi = 45\,\mathrm{GeV/c^2}$ (90% C.L.) for neutrons. Further Dark Matter candidates, such as axions or axion-like particles, are tested with electronic recoil data obtained during the same Dark Matter data taking period. This lead in 2014 to the best upper limit for axion-electron couplings in the $5 - 10\,\mathrm{keV/c^2}$ mass range, assuming that axion-like particles constitute all of the galactic Dark Matter.

Thanks to the underground laboratory and a careful material selection, the XENON100 experiment archived an ultra-low electro-magnetic background of $5.3 \times 10^{-3}\,\mathrm{events/kg/day}$ in the region of interest. In order to identify a WIMP event properly, it is important to understand the background. Therefore a set of calibration sources is used: ^{241}AmBe (Americium-Beryllium) as a neutron source and ^{60}Co (Cobalt) and ^{232}Th (Thorium) as a gamma source (electronic recoil). The neutron source is used to understand how a WIMP signal would look like in the detector and the gamma sources are used to understand the electronic recoil signal, similar to the background one. Weekly photo multiplier-tube calibrations are used, in addition, to determine the gain values and assure a stable read-out condition.

This work is dedicated to a subset of three data analysis topics. An alternative photo-multiplier tube calibration is developed from raw LED data. This method is able to support the actual calibration technique by searching for single photo-electron pulses with the same peak-finder algorithm which is used for Dark Matter data reprocessing. This allows to determine the single photo-electron response directly. An already used method is improved: The signal acceptance for the ionisation energy threshold is developed for the data taking periods 2011/12 and 2013/14. It is also tested with simulated neutrons and further uncertainties on the signal acceptance are discussed. Furthermore there is an alternative phenomenological electronic recoil background model developed from ^{60}Co and ^{232}Th data. This model is compared to the already published model in 2012 and the impact on the WIMP exclusion limit is evaluated.

ZUSAMMENFASSUNG

Kosmologische Beobachtungen bestätigen die Existenz von *Dunkler Materie*. Die neuesten Ergebnisse der Planck Satelliten-Mission zeigen, dass unser Universum aus 4.9% baryonischer Materie, 26.8% Dunkler Materie und 68.3% Dunkler Energie besteht. Theorien, wie zum Beispiel die super-symmetrische Erweiterung des Standard Modells der Kern- und Teilchenphysik, beinhalten ein geeignetes Teilchen. Es besitzt Masse, keine Ladung und zeigt eine schwache Wechselwirkung mit baryonischer Materie.

Die XENON Kollaboration hat im Jahr 2007 mit der Suche nach Dunkler Materie begonnen. Hierfür wird im Untergrundlabor des "Laboratori Nazionali del Gran Sasso" (Italien) eine Zwei-Phasen Zeit-Projektions-Kammer (TPC), mit 161 kg flüssigem Xenon als Detektor verwendet. Das Ziel besteht im Nachweis eines neuen Elementarteilchens: Das WIMP (***W**eakly **I**nteracting **M**assive **P**article*). Die XENON Kollaboration veröffentlichte im Jahre 2012 eine Ausschlussgrenze mit dem minimalen Wechselwirkungsquerschnitt von $\sigma_\chi = 2 \times 10^{-45}\,\mathrm{cm}^2$ bei einer WIMP-Masse von $m_\chi = 55\,\mathrm{GeV/c^2}$ (90% C.L.) für spin-unababhängige WIMP-Nukleon Wechselwirkungen. Im Jahr 2013 folgte mit dem gleichem Datensatz ein minimales Ausschlusslimit von $\sigma_\chi = 3.5 \times 10^{-40}\,\mathrm{cm}^2$ bei eine WIMP-Masse von $m_\chi = 45\,\mathrm{GeV/c^2}$ (90% C.L.) für spin-abhängige WIMP-Nukleon Wechselwirkungen an Neutronen. Darüber hinaus wurden alternative Dunkle Materie Kandidaten, wie z.B. Axionen oder Axion-ähnliche Teilchen, an Elektronrückstößen von Xenonatomen untersucht. Dies führte im Jahr 2014 zu einem Ausschlusslimit für Axion-Elektron Wechselwirkungen im Bereich von $m_A = 5 - 10\,\mathrm{keV/c^2}$ unter der Annahme, dass die gesamte Dunkle Materie aus Axion-ähnlichen Teilchen besteht.

Dank des Untergrundlabors und einer geeigneten Materialauswahl erreicht das XENON100 Experiment einen niedrigen elektro-maganetischen Untegrund von $5.3 \times 10^{-3}\,\mathrm{Ereignissen/kg/Tag}$. Ein detailliertes Verständnis des Untergrunds ist notwendig um nach seltenen WIMP-Ereignissen zu suchen. Hierfür stehen verschiedene Kalibrationsquellen zur Verfügung: ^{241}AmBe (Americium-Beryllium) als Neutronenquelle und ^{60}Co (Cobalt) und ^{232}Th (Thorium) als Gammaquelle. Die Neutronenquelle wird dazu verwendet, WIMP Wechselwirkungen im Detektor zu verstehen. Die Gammaquelle wird zur Beschreibung von Elektronrückstößen verwendet, die sich ähnlich dem elektro-maganetischen Untergrund verhalten. Wöchentliche Kalibrationen der Photoelektronenvervielfacher geben Auskunft über ihren Funktionsstatus und deren zu erwartende Signale.

Die hier vorliegende Doktorarbeit fokusiert auf drei Analysethemen. Es wird eine alternative Methode vorgestellt um einzelne Photo-Elektron Signale in den LED Rohdaten zu identifizieren. Diese Methode kann ebenfalls dazu verwendet werden den Status der PMTs zwischen zwei PMT Kalibration zu überprüfen. Eine bereits vorhandene Methode zur Signalakzeptanzbestimmung der unteren Energieschranke für Ionisationssignale wird verbessert, einschließlich einer vollständigen Disukussion der eingehenden Unsicherheiten. Desweiteren wird ein alternatives Modell zur Bestiummung des zu erwartenden γ induzierten Untergrunds mit ^{232}Th und ^{60}Co Daten entwickelt. Dieses Modell wird im Weiteren mit dem bereits veröffentlichtem Modell aus dem Jahr 2012 verglichen und sein Einfluss auf die WIMP Ausschlusskurve bestimmt.

Contents

Abbreviations

2011/12	Refers to the Dark Matter data taking period of 2011/12 with the necessary calibration data taking before and after.
2013/14	Refers to the Dark Matter data taking period of 2013/14 with the necessary calibration data taking before and after.
ADC	Analog-Digital Converter
CMB	Cosmic Microwave Background
CNB	Cosmic Neutrino Background
DAQ	Data Acquisition
Discrimination space Y	The usual data space which is used in the XENON100 data analysis.
DM	Dark Matter
LED	Light-Emitting Diode
LNGS	Underground lab “Laboratori Nazionali del Gran Sasso” in Italy
Parameter: $\mathscr{L}_{\mathbf{eff}}$	The relative scintillation efficiency.
Parameter: $\mathscr{Q}_{\mathbf{y}}$	The charge yield
PDF	Probability Density Function
PMT	Photo-Multiplier Tube
SPE	Single Photo-Electron
SUSY	Super Symmetry. An discussed extension the standard model of particle physics
Symbol: ♁	The astronomical symbol for the planet Earth.
Symbol: ☉	The astronomical symbol for the Sun.
Symbol: ♃	The astronomical symbol for the planet Jupiter.
WIMP	Weakly Interacting Massive Particle
ZLE	Zero Length Encoding

<value>	If nothing else is mention in the text, the arrow brackets indicate the mean value.

1. Motivation

"If it's so massive, why aren't we being crushed by the gravity?
Well, we would be if it were normal matter, but it's not. Scanners can't make heads or tails of it. Its atomic structure is damn near impenetrable to our science. Amazing. *Dark Matter*."

Crew members of the merchant transport vessel Nestor, The Outer Limits, "Dark Matters"

The search for Dark Matter is an ongoing business for over 80 years. Several earth-bound experiments have searched for new types of particles which interact with baryonic matter weakly. These investigations were triggered by cosmological studies which found out that there are unknown gravitational potentials in the universe which are necessary to describe our nowadays observed structures in, for example, galaxy clusters. The actual observations can be explained by an additional type of matter which is not yet observed directly. Furthermore, this *Dark Matter* is already important in the early universe where the additional gravitational potentials cluster baryonic matter to form our universe as we see it today. This process refers to structure formation. Another *picture* of the early universe is the Cosmic Microwave Background (CMB) which refers to 380,000 years after the Big Bang. Observations with the Planck satellite suggest that our universe consist of 4.9% baryonic matter, 26.8% Dark Matter and 68.3% Dark Energy. Baryonic matter is investigated in accelerator facilities nowadays. From these studies, the standard model of particle physics has been developed which is used to explain physical processes in particle physics. Nevertheless, there are processes which are not fully understood yet and Dark Matter is one of these unsolved mysteries and a hint for physics beyond the standard model.
The favoured explanation is a new particle type: The *WIMP*. According common theories (e.g. the supersymmetric extension of the Standard Model, SUSY), these *W*eakly *I*nteracting *M*assive *P*articles are created during the big bang such as baryonic matter. Duo to the expansion of the universe and the involved cooling process, the interaction probability with baryonic matter has decreased up to nowadays. Nevertheless, such Dark Matter particles interact by gravitation. Ongoing experiments for direct Dark Matter detection wait for a weak interaction of a WIMP with the chosen target material. Up to now, there haven not been any WIMPs detected and hence the WIMP mass m_χ and cross-section σ_χ are still unknown. Depending on the WIMP model, there are mass ranges suggested from some *keV* up to *TeV*. Experiments such as DAMA/LIBRA have claimed to observed Dark Matter in annual modulation. This hypothesis is still discussed by the Dark Matter community. Their claimed WIMP mass and cross-section is already excluded by different Dark Matter experiments.
Another approach in the Dark Matter search are axions and axion-like particles. Their theory is based on Peccei, Quinn, Weinberg and Wilcze who introduced another broken symmetry to explain the observed CP invariance in particle physics. This leads to a new particle with a mass of a few keV which could also explain the missing mass in the universe. Experiments such as CAST use the Primakoff effect to search for that kind of Dark Matter.
In summary, the ongoing experiments could not provide reasonable results for the existence of Dark Matter particle(s). The hunt for Dark Matter is approached in different ways: Looking for Dark Matter annihilation processes (indirect detection), looking for a Dark Matter interaction in earth-bound experiments (direct detection) or by the production of Dark Matter particles at collider facilities where

the missing energy is a hint for the produced DM particle. This thesis will give a broad introduction into Dark Matter from a cosmological and particle physic point of view and will summarise the direct Dark Matter detection and the experimental constrains.
The here presented work will focus on two-phase Time Projection Chamber (TPC) of the *XENON Dark Matter Project* which uses ultra pure xenon as target material for the WIMP search. Besides the great technically achievement to build such a low-radioactive measuring device, past publications exclude a large range of WIMP masses and WIMP-nucleon cross-sections and therefore ongoing studies aim to detect low mass WIMPs in some keV range. The actual experimental setup, the XENON100 TPC with $m_{Xe} = 161\,\text{kg}$ of liquid xenon resulted a WIMP exclusion limit down to a cross-section of $\sigma_\chi = 2 \times 10^{-45}\,\text{cm}^2$ for a WIMP mass of $m_\chi = 55\,\text{GeV}$ (90% C.L.) in 2012. The successor experiment XENON1T which will start to take data at the end of 2015 and will probe the actual results from XENON100 and competitors. Duo to the larger amount of xenon and a lower background expectation it will undergo actual exclusion limits and be 100 time more sensitive to WIMP interactions.
This work is dedicated to a subset of three data analysis topics. A basic calibration is done by the weekly PMT calibration with LEDs. To test the PMT response and determine the PMT gain. This is done by measuring single photons from the LED at the PMTs. At the moment, there are two tools in use: A piece of software analyses the raw data from the calibration sources or the Dark Matter data sets (*xerawdp*) and another one analyses the LED raw data. The goal of this work is to bring both different types of software together and test if the peak-finding ability of *xerawdp* is good enough, in case of analysing LED raw data. This work will give detailed information on the necessary adjustments in the software code of *xerawdp* and show first results for the peak identification with LED raw data. Furthermore the single photo-electron response is tested with a new derived cut set for LED data. This test will search for single photo-electrons directly (signals from scintillation). Once the ability to identify single photo-electrons is proven, a direct gain measurement is carried out which is then compared to the results of the LED calibration with the standard software. The method itself introduce a new application: Single photo-electrons are not only observed in LED raw data, they are also observed in calibration or Dark Matter data. The ability to identify them with *xerawdp* could be useful to support the ongoing PMT calibration and the control the PMT status between two LED calibrations.
A WIMP detection is only possible with a careful calibrated detector. There are several sources which create a (e.g. neutron induced) signal in the TPC, like a WIMP would do. Therefore it is important to characterise the known signal types with measurements and simulations to have a good knowledge of the expected background. Possible signals come from nuclear recoils (NR) by neutrons and electronic recoils (ER) from a γ source. A ^{241}AmBe (Americium-Beryllium) calibration source is used to determine the detector response to neutrons in detail. The core of the analysis is to detect a WIMP signal, based on the signal which is tested against the background with the Profile Likelihood analysis. Therefore the event selection must be optimised by the use of a specific set of cuts, which are designed

to reduce the total amount of events due to energy or position constrains in the TPC. Other typical cuts are looking for single interactions or a proper observed signal which is defined some other criteria. One of the most important cuts is the threshold for signals of the ionisation process. This cut assures that only events are selected which are above a certain energy threshold to grant an 100% trigger efficiency on the ionisation signals. This work describes a proper way to calculate the cut acceptance and furthermore the method is tested with simulated neutrons in the XENON100 TPC geometry. The goal is to extend the studies on the acceptance to develop a uncertainty level for the cut acceptance. This uncertainty will influence the accuracy of the Profile Likelihood analysis.
The electronic recoil is quantified with calibrations from ^{232}Th (Thorium) and ^{60}Co (Cobalt) sources. Both calibration sources have typical energy signatures and therefore the background expectation rate is estimated from this sources. Due to the fact that radioactive decays in the used material are responsible for e.g. further electrons which mimic an additional ionisation signal, the background of the ER background must be modelled in detail. This model is then used during the Dark Matter data analysis to predict the number of expected ER events in a specific energy range. Due to the composition of the ER background in detail, the signal hypothesis must be tested against background events in a certain energy range. Because of the low event rate with real physical motivated signals, the background model must be describe the ER background as good as possible to reduce the uncertainties in the Profile Likelihood analysis. This work present an ER background model which describes the data as good as possible in the usual XENON100 discrimination space. In addition, the goodness of the ER background model is discussed and the impact on the WIMP exclusion limit is tested.
These three topics are based on XENON100 data, but the analysis methods are also interesting for the upcoming XENON1T experiment. Energy thresholds and their acceptances, as well as the search for single photo-electrons, PMT response, calibration and background modelling are also necessary in XENON1T.

2. Introduction to Dark Matter

"I don't want to be human. I want to see gamma rays, I want to hear X-rays, and I want to smell *Dark Matter*. Do you see the absurdity of what I am?"

John Cavil, Cylon Model Number One,
Battlestar Galactica, "No Exit"

2.1. Motivation

Starting with Fritz Zwicky, the search for Dark Matter is going on for 80 years now. Since then we have improved our knowledge about the universe and its content: Beginning with a deeper understanding about fusion processes in our sun up to the structure formation at big scales and finally the precise measurement of the cosmic microwave background (CMB). At the same time the standard model of particle physics (SM) has been developed and tested on many particle accelerator facilities world-wide to understand how our universe works on the smallest scales. Observations like the "ghost-like" neutrino, which was postulated by Pauli in 1930, or the Higgs boson in 2014 are important confirmations of the standard model.
The existence of Dark Matter from a cosmological point of view does not fit into our general understanding of particle physics yet. From here it is necessary to bring both fields, particle physics and cosmology together to search for physics beyond the standard model. The following chapter reviews Dark Matter from cosmological observations and gives a general introduction to particle Dark Matter.

2.2. Cosmological Introduction To Dark Matter

This section gives a detailed view on Dark Matter and its necessity for our nowadays observed universe. A historical review of the past 80 years shows that Dark Matter is observed at the scale of single galaxies, super galaxy clusters up to the CMB which represents our universe $\sim 380\,000$ years after the Big Bang. Simulations about the structure formation support the need of Dark Matter to explain the observed structures in our universe. It turns out that Dark Matter is negligible on scales of single stellar systems and becomes more important the bigger the the observed structures are.

2.2.1. Spiral galaxies

In the middle of the 20^{th} century, several spiral galaxies were investigated with respect to the velocity dispersion of the contained stars and hydrogen clouds. These objects cycle around the core of spiral galaxies and their number, age and luminosity depend mainly on the turning radius. Their radial velocity is determined by measuring the Doppler shift of its emitted light. Vera Rubin and others used spectrographs to measure that light [99].

The Doppler shift predicts that emitted light becomes red-shifted if it moves away from the observer and becomes blue-shifted if it moves towards to the observer. Formula (2.1) shows the radial velocity

calculated from the measured spectral line λ_{obs} and the spectral line λ_0 from a calibrated source in the lab [104].

$$v_r = \frac{\lambda_{obs} - \lambda_0}{\lambda_0} c \tag{2.1}$$

In addition, the tangential velocity is determined from the proper motion of the object. Spiral galaxies provide the observer with regions of different star and cloud types to measure spectral lines there. These regions are compared to each other to determine a rotation curve profile along the radius of the galaxy. A dedicated analysis of stars and hydrogen clouds leads also to a luminosity profile L which is important to determine the mass of such objects by the "mass-luminosity-relation" (see equation (2.2)) [104]. The mass M is important to prove the rotation curves later.

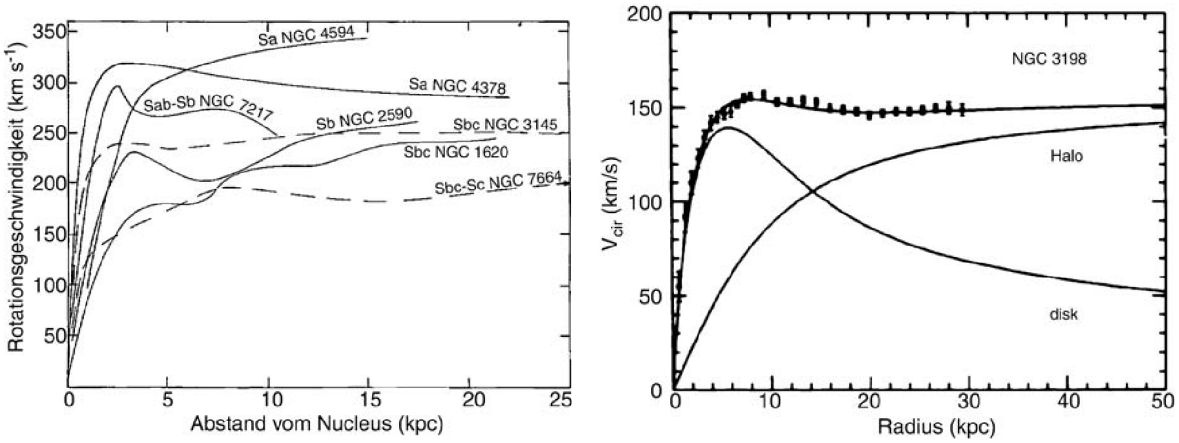
$$\frac{L}{L_\odot} = \left(\frac{M}{M_\odot}\right)^a \tag{2.2}$$

Equation (2.2) uses the sun as reference (luminosity $L_\odot$ and mass $M_\odot$). The exponent a is used to adjust the equation for different types of stars. For example $a = 3.5$ describes common stars of the main-sequence.

Measurements of many spiral galaxies show that the tangential velocities are not decreasing for objects which are far away from the halo of the galaxy. They remain constant at some point. This observation is in conflict with Kepler's law which predicts a decrease of the velocity with a larger radius (see equation (2.3)).

$$v(R) = \sqrt{\frac{GM(R)}{R}} \tag{2.3}$$

Observations of different spiral galaxies show a similar behaviour in figure 2.1a. The larger the distance from the core is, the flatter becomes the rotation velocity curve. This behaviour became a strong evidence for Dark Matter in galaxies. The consequence of such an observation is an additional mass distribution inside the radius of a measured star or cloud with a non-luminous property. This implies a Dark Matter halo with a density $\rho_{DM} \sim R^{-2}$ and a mass distribution which is proportional to its radius: $M \sim R$. Hence the additional Dark Matter rises with the size of the galaxy which makes it difficult to determine the exact size of the galaxy. The predicted Dark Matter halo is shown for a single example of galaxy NGC3198 (see figure 2.1b). The measured rotation curve is plotted with an additional fit to the data. To understand the discrepancy between theory and measurement, the curve labelled "disk", describes the velocity distribution which is expected according Kepler's law.
The velocity of an object inside a given radius R of a galaxy mass distribution rises according $v \sim R$. Outside a radius R, the velocity of a mass sample (m) is proportional to $v \sim 1/\sqrt{R}$ (disk). To explain



(a) (b)

Figure 2.1 – Figure 2.1a: Example of different rotation curves. Each profile shows a flat outfall for a bigger radius. This behaviour is not predicted by Kepler's law and build a strong evidence for Dark Matter. Figure 2.1b: The spiral galaxy NGC3198 in detail – This figure shows the measured object along the distance from the core of the galaxy. The flat rotation velocity spectrum differs from the prediction (disk) by Kepler's Law which assumes that all mass comes from the observed luminosity. In addition a Dark Matter halo (Halo) is shown. Only by adding that Dark Matter component, the measured velocity distribution can be explained. [104]

the discrepancy from data to the expected curve, a Dark Matter halo is added, which presents an additional gravitational component in galaxy NGC3198. This scenario suggests an additional Dark Matter density of ρ_{DM} (over the critical density $\rho_c = 3H^2/8\pi G$) of roughly [76]:

$$\Omega_{DM} = \frac{\rho_{DM}}{\rho_c} = \sim 0.1$$

This value comes from galaxy observations and does imply the type of Dark Matter itself. Section 2.2.5 will introduce a new particle Dark Matter type. Beside an extension to the SM there are other solutions suggested such as MACHOs (see section 2.2.5). Classical solutions are often discussed, but all their constraints make it difficult to create a proper Dark Matter halo model [98].

The search for Dark Matter in the universe is connected to the nowadays observed structures. Dark Matter impacts our own galaxy as well as all galaxies around us and therefore the following discussed methods aim to detect Dark Matter outside our host galaxy. Details about mass and luminosity measurements inside (as well as outside) our galaxy are summarised in the literature, see e.g. [104].

2.2.2. Galaxy Cluster and Super Cluster

Ongoing investigations on single galaxies are necessary to understand their properties in detail. But it turns out that Dark Matter is also important to explain the observed structures in galaxy clusters or super galaxy clusters. In 1933, Fritz Zwicky found first hints of a *missing mass problem* when he studied the Coma cluster [104].

Galaxy clusters consist of three parts: First, a huge number of galaxies which are bound together by gravitation. They consist of different types of galaxies each with their typical shape and luminosity. For example, the Coma cluster consists of roughly 1000 galaxies in total. These galaxies are enveloped by the interstellar medium (gas) which is the second component. Its mass overpowers the mass of all galaxies inside a galaxy cluster. This additional hot gas between the galaxies is almost invisible for optical measurements. New observation technologies such as X-ray astronomy are developed. A third component consists of Dark Matter which makes up almost 90% of the cluster mass [83]. Direct observation of Dark Matter is not possible. Measuring the missing mass with different techniques such as gravitational lensing help to understand its distribution.

A first evidence for additional matter in galaxy clusters was found by applying the Virial theorem to measured data of radial velocities. The cluster mass is determined by its luminosity (see equation (2.2)). The Virial theorem for objects which are bound by gravitation is given in equation (2.4). It connects the kinetic energy T with the potential energy U. Formula (2.5) is developed from equation (2.4) by using the mass and velocity of the individual galaxies i in the cluster. A detailed derivation of the Virial theorem for stellar astrophysics is given in [45].

$$2T = -U \tag{2.4}$$

$$2\sum_i m_i v_i^2 = -\frac{G}{2}\sum_{i!=j}\frac{m_i m_j}{r_{ij}} \tag{2.5}$$

If the time t_{cross} which is needed by a single galaxy to cross a cluster is smaller than the lifetime of the cluster itself, the assumption holds that these objects are bound together by gravitation. Then galaxies relax in a *Virial equilibrium*. A single galaxy velocity and mass is not taken into account if the Virial theorem is applied. The total mass of the cluster is assumed to be:

$$M = \sum_i m_i \tag{2.6}$$

These assumptions lead to a velocity dispersion:

$$\left\langle v^2 \right\rangle := \frac{1}{M} \sum_i m_i v_i^2 \tag{2.7}$$

The potential energy is expressed by the gravitational radius r_G:

$$r_G := 2M^2 \left(\sum_{i!=j} \frac{m_i m_j}{r_{ij}} \right)^{-1} \tag{2.8}$$

Hence, equation (2.5) becomes

$$M \left\langle v^2 \right\rangle = G \frac{M^2}{r_G} \tag{2.9}$$

With this result the cluster mass can calculated by:

$$M = r_G \frac{\left\langle v^2 \right\rangle}{G} \tag{2.10}$$

According to equation (2.10) each galaxy becomes a test particle which is bound to the centre of mass of the galaxy cluster. To calculate the total mass M, the single distances r_i are not useful because their information is not necessarily available. Therefore the distances are expressed as "projected distances". If the distance to a galaxy cluster is known and much larger than its extent and the single galaxies follow an isotropic velocity distribution, then the following expression is valid:

$$\left\langle v^2 \right\rangle = 3\sigma_v^2$$

This expression is based on a Gaussian velocity distribution in three dimensions. Its deviation is set to $\sigma_v = 1000\,{}^{\mathrm{km}}/_{\mathrm{s}}$ for galaxy clusters from observations. Due to the projection, the radius r_G yields:

$$r_G = \frac{\pi}{2} r'_G \quad \text{with } r'_G = 2M^2 \left(\sum_{i!=j} \frac{m_i m_j}{r_{ij}} \right)^{-1}$$

The size of σ_v and r_G are available directly. This effort allows calculating the total mass of galaxy clusters according to formula (2.11)

$$M = \frac{3\,\pi R_G \sigma_v^2}{2G} \tag{2.11}$$

The Virial theorem gives first hints of the missing mass problem because it turns out that the observed total luminosity of galaxy clusters from optical observations L_{opt} is not large enough to explain the estimated total mass M [104, 111].

Hot gas which envelops all the galaxies has its luminosity maximum in the X-ray spectrum. First satellite missions such as UHURU in 1970 observed the interstellar medium in high mass galaxy clusters and successor missions such as the Einstein satellite found evidences for hot gas also in low mass galaxy clusters [104]. In 1990, further missions such as CHANDRA and XMM-Newton measured the X-ray spectrum down to 10 keV in low- and high-mass galaxies with an increased precision of the temperature measurements. From their observations, the density $\rho(r)$ and temperature $T(r)$ are determined for measured radius. This information is used to evaluate the mass of a galaxy cluster.
X-ray investigations yield that hot gas between the galaxies contributes to 15% to the total mass of the galaxy cluster. Compared to $\sim 3\%$ mass which is observed by its luminosity, a huge amount of mass is still missing to describe the total mass of galaxy clusters. This observation shows that the missing mass is not only limited to the number of individual galaxies in a galaxy cluster.

Another important method to determine the total mass of galaxies or galaxy clusters is gravitational lensing. The basic idea is that a huge source of gravity (e.g. a galaxy cluster) is able to bend light around its mass. Einstein formulated that idea in 1915 after he had developed his mass-energy equality. As long as photons carry energy, their track is bend by huge masses in a four-dimensional space-time. Nowadays, the gravitational lensing effect is used in Dark Matter search to determine the mass of galaxies or galaxy clusters without measuring the luminosity of the stars or using X-ray astronomy. Figure 2.2 sketches the basic idea of the method. An object S which is at a distance D_S from the observer O is faded out by e.g. the galaxy cluster itself or the interstellar medium. The distance D_d to the gravitational lens is known from other measurements [41, ff411]. The light is travelling along the path l from the source in direction of the observer:

$$\vec{\hat{\alpha}} = \frac{2}{c^2} \int \nabla_{\perp} \Phi dl \tag{2.12}$$

Here the light trajectory is perpendicular to the gravitation potential under the observation angle $\vec{\hat{\alpha}}$ (angle between "true" source position S and observation position L) at the gravitational lens itself. Position L becomes important for the observer O which is only able to measure angle Θ in the sky after the light is already bent by the gravitational lens. The lens itself is thin compared to the distances D_d, D_s, D_{ds} and the light is bent after a distance D_{ds} instantly at ξ. The parameter of interest is β which describes the true position of the source in the sky. It is calculated by the lens equation ((2.13)) under the assumption of small angles:

$$\beta = \Theta - \alpha(\Theta) \tag{2.13}$$

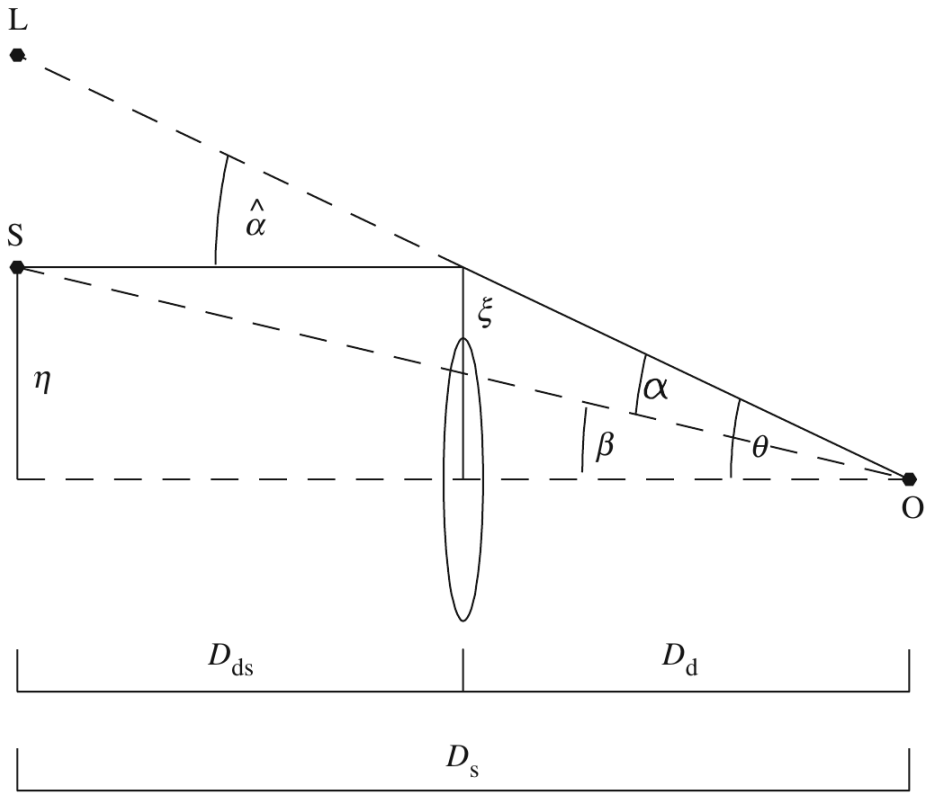


Figure 2.2 – Sketch of the gravitational lensing effect. Light travels along a trajectory in direction to the observer O. The galaxy cluster at D_s bends the light by its gravity to an observer. From the observer's view it looks like that the light comes from position L. (Sketch from [94])

The gravitational lens is described by an arbitrary mass profile $M(\Theta)$ which takes into account that such lens are not perfectly circular. Equation (2.13) becomes:

$$\beta = \Theta - \frac{D_{ds}}{D_d D_s} \frac{4GM(\Theta)}{c^2 \Theta} \tag{2.14}$$

Hence equation (2.14) is simplified if a source is observed at position $\beta = 0$. Hence a ring structure is observed in the sky with a fix "Einstein radius". Figure 2.3 shows an example of an Einstein ring taken by the Hubble space telescope. Here the blue light of a background galaxy looks like a horseshoe around the red light of a foreground galaxy. The Einstein radius for a point mass is given by equation (2.15).

$$\Theta_E = \sqrt{\frac{4GM(\Theta_E)}{Dc^2}} \tag{2.15}$$

This simplified equation contains the parameter of interest for the Dark Matter search. The mass $M(\Theta_E)$ is used to estimate the mass of such a galaxy cluster in between the source and observer.

Figure 2.3 – Example for an Einstein radius. The blue light of a background galaxy looks like a horseshoe around a foreground galaxy which acts as gravitation lens. [87]

Therefore only the light of the source is needed [41]. This method to determine the mass is called ”Strong Gravitational Lensing Effect“. No further information is required about the physical state of the lens itself. Therefore the method becomes independent from X-ray observations or luminosity measurements. Unfortunately, observations show that this method works only for a few galaxy clusters. It requires a good alignment of the foreground mass and the background galaxy, which is rare. Nevertheless the shape of the Einstein rings yield a good prediction of the mass between observer and source.

In reference to the few examples where the Strong Gravitational Lensing Effect is used to determine the foreground mass another effect is exploited: ”Weak gravitational lensing“. The upper part of figure 2.4 shows a sketch of the Euclid satellite mission which measures the Dark Matter distribution along its way of the light trajectory from a 6.5 billion years old source in space [53]. In comparison to the Strong Gravitational Lensing Effect the missing mass is not distributed in one place and a more

complicated structure is observed (see lower part of figure 2.4).

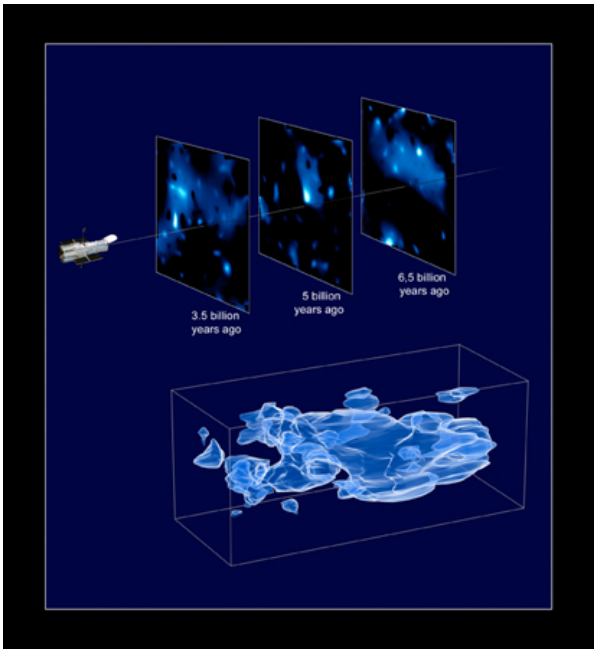


Figure 2.4 – The Euclid Satellite mission creates a 3D map of the Dark Matter distribution. It uses the weak gravitational lens effect which is more accurate in determining the distributions of the missing mass in comparison to the strong gravitational lens effect. [53]

The observed structure results from the fact that light is not bent by one foreground object – It is sheared. Similar to the Einstein rings each observed galaxy shape can be understood as a result of a light distortion by a foreground mass. Hence a single background galaxy changes its observed shape due to the foreground mass during the observation. For example, a circular spiral galaxy becomes elliptical due to the mass in the line of sight [41]. This method requires a huge number of background galaxies. A typical number for them is 50-100 galaxies per square arc minute at faint magnitudes [35]. Large CCD (charge-coupled device) arrays increase the resolution nowadays. Huge sky surveys of background galaxies in galaxy rich clusters are made to determine the mass distribution in the line of sight. A famous example is the Bullet Cluster in figure 2.5. A combination of all previous methods is applied to study the Bullet Cluster. Here two galaxy clusters collided in the past. Its mass is estimated by the luminosity. X-ray astronomy allows estimating the mass of the interstellar medium in between and finally gravitational lensing measures the shear of the background galaxies with high statistics. The result is overlaid in figure 2.5 where the distribution of the hot gas is shown in pink, the galaxies are the ”white dots“ and the two centres of mass are determined by the gravitational lensing in blue. The latter is not at the same position as the hot gas and the galaxies are observed (at the collision

point). A common interpretation is that both centres of gravity are made out of Dark Matter which interacted less during the collision in the past. Hence the Dark Matter moved on through the Baryonic Matter after the collision [40].

Dark Matter searches on cosmological scales use the above explained techniques (Virial theorem, X-ray astronomy and gravitational lensing) to determine the missing mass. It is necessary to combine all three methods to determine the missing mass out of different and independent sources.

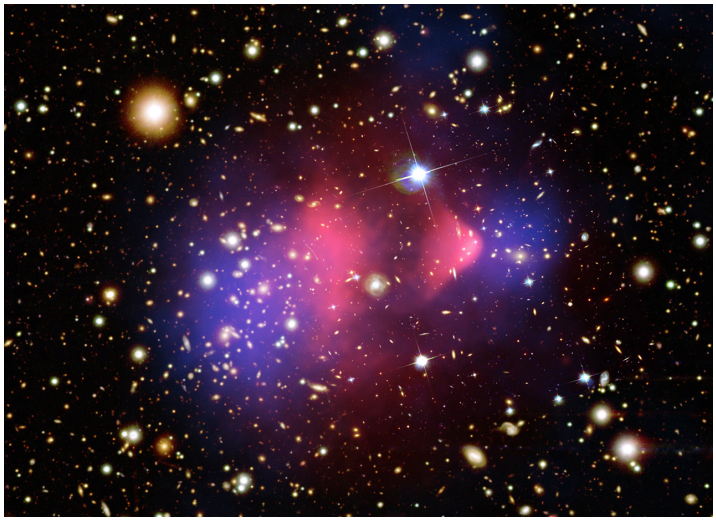


Figure 2.5 – The Bullet Cluster [114] combines different methods to estimate the mass of the colliding galaxy clusters. The result is given by the colour code in the figure. In pink, the X-ray observation from Chandra which determines the hot gas (baryonic matter). In between are single galaxies which are used to determine the mass from optical observation methods. Finally in blue, the distribution of the gravitational source. Both centres of mass do not match to the observed baryonic mass (stars and hot gas). This issue is interpreted as a hint for Dark Matter.

2.2.3. Structure Formation

An important observation at large scales is the structure formation of galaxy clusters and super clusters. This epoch is related to the time when matter has been already the dominating part of the universe and gravitational collapses led to clump the masses. The Sloan Digital Sky Survey (SDSS) maps the universe since 1998 and determines filaments and voids. The outcome is presented in figure 2.6. The previous sections discuss the necessity of Dark Matter in single sub-structures but they do not raise the question if Dark Matter was an important keystone during the structure forming itself. This subsection

discusses three different questions:

1) Is the observed mass at that time in consistence with the Big Bang Nucleosynthesis (BBN)?
2) Which scenario explains the matter cluster process: "top-down" or "bottom-up" structure formation?
3) Is Dark Matter imperative during the structure forming?

The first question addresses the BBN directly. The first lightweight elements formed due to the ongoing expansion of the universe (cooling process). This process happened a few minutes after the Big Bang (BB). During that time the baryons did not contribute much because of their low density, and followed an equilibrium condition (equations (2.16) and (2.17)):

$$p + e^- \leftrightarrow n + \nu \tag{2.16}$$

$$p + \bar{\nu} \leftrightarrow n + e^+ \tag{2.17}$$

$$n \leftrightarrow p + e^- + \bar{\nu} \tag{2.18}$$

In addition, the neutron decays with a decay time of $\tau_n = 887\,s$ (equation (2.18)). As long the rates, and hence their decay times, of equations (2.16) and (2.17) are shorter than 887 s the ratio of neutrons to protons is given by equation (2.19), where the mass difference from neutron and proton is $\Delta m = m_n - m_p = 1.293\,\mathrm{MeV}$.

$$\frac{n_n}{n_p} = \exp\left(-\frac{\Delta m\, c^2}{k_B\, T}\right) \tag{2.19}$$

After the cosmic neutrino background (see section 2.3.2.4) was formed due to the freeze-out of neutrinos, the probability for equation (2.16) and (2.17) decreased and neutrons and protons are left in the state of equilibrium. Free neutrons decayed after τ_n and the reason why neutrons are observed today is that they combined with protons and electrons to atoms. The first elements in the universe with their abundance (values in []) relative to hydrogen are [94]:

$$\mathrm{H} \quad [1], {}^4\mathrm{He} \quad [0.25], \mathrm{D} \quad [3\times10^{-5}], {}^3\mathrm{He} \quad [2\times10^{-5}], {}^7\mathrm{Li} \quad [1\times10^{-10}]$$

The measurement of ^{4}He was performed on different scales in the universe (from stellar atmospheres to gas clouds and clusters) and is one of the most important hints for the Big Bang theory. The primordial abundance at the basis of the BBN for the lightest elements is displayed in figure 2.7. Here the relative abundance to hydrogen is shown for the evolving baryon density in units of $10^{-31}\,\mathrm{g/cm^3}$. From the observation of hydrogen and helium the baryon density ρ_b is calculated in comparison to the critical density ρ_c to:

$$\Omega_b = \frac{\rho_b}{\rho_c} \approx 0.04$$

where the critical density is given in equation (2.20) with the Hubble constant H_0 and the gravitation constant G:

$$\rho_c = \frac{3H_0^2}{8\pi G} \tag{2.20}$$

This result does not agree with observation from galaxies or galaxy clusters from a general matter prediction of $\Omega_{matter} > 0.1$. Therefore another amount of non-baryonic matter is necessary to explain that discrepancy [104].

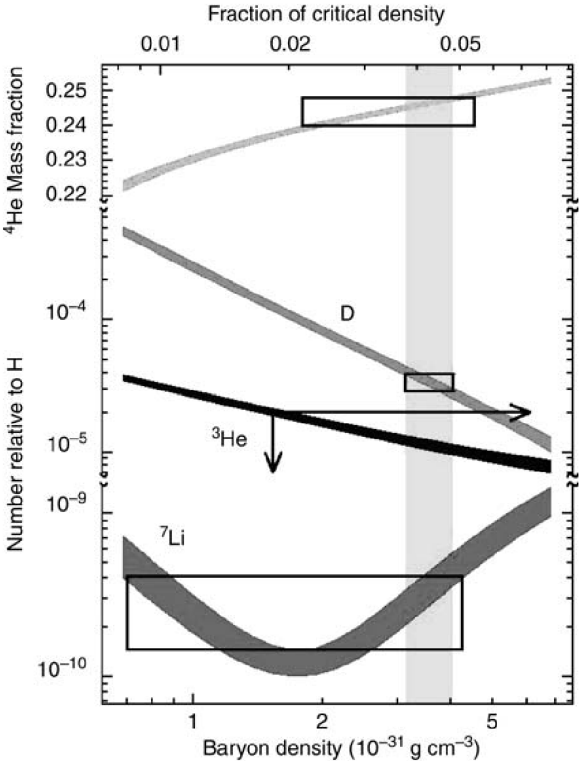


Figure 2.7 – The BBN prediction for the first lightest elements in the universe: H, ^{4}He, D, ^{3}He, ^{7}Li. Their abundances (in relation to H) are shown against the the baryon density. The BBN measurements agree within 10% in comparison to the calculations. The vertical spread is related to the measured distribution of the elements and the horizontal spread shows the overlap with theoretical predictions [104, ff. 166].

At the beginning, the universe was isotropic and homogeneous and became more and more clustered. At some point matter undergoes a gravitational collapse. This idea was first described by Jeans and takes into account that gravitational energy is transformed to heat (kinetic energy). Atoms and molecules collide under their gained energy and due to dissociation or ionisation photons are emitted. The total energy of such a process is conserved, but the single components of the gas clouds are bound together stronger. As soon the critical Jeans mass is reached it acts like a condensation nucleus and further matter clusters around. Emanating from these places of higher density the structure forming

begins [94, ff. 138]. The scenario which is suggested by Jeans is called ”bottom-up scenario“. Matter clusters from small structures up to the biggest structures as they are observed today. In contrast to the bottom-up, the top-down mechanism describes that matter first clusters on largest scales and then smaller substructures develop. To answer that question thoroughly it is important to understand possible Dark Matter (particle) candidates which are discussed in section 2.3. The main difference in the chosen Dark Matter types is their velocity distribution. Particles which move slowly through the universe (”Cold Dark Matter“) are gravitational bound to matter. And as long as the Dark Matter is assumed to be $\sim 26\%$, its mass boosts the clustering process. Another extreme solution are Dark Matter particles with velocities near the speed of light and small mass (”Hot Dark Matter“). These candidates are able to leave every cluster of matter and do not contribute to the structure formation, hence a top-down structure process is expected. Results from SDSS which searches for the largest scales in the universe favour a bottom-up scenario [41].

The last question is difficult to answer from observations only. As long as there is only one universe there is no possibility to compare the growth process to other universes. Therefore simulations are done (e.g. ”Millennium Run”). A simulation of the entire universe is complicated and not executable on nowadays computer. Hence the “Millennium Run” takes a cube of 650 Mpc size and puts in Dark Matter and baryonic matter. The simulation starts after the universe became transparent to photons and takes small density fluctuations into account which, for example, come from Jeans' scenario. The further n-body simulation with the evolution of quasars and galaxies on large scale structures in the universe is carried out until “today”. The simulation outcome is then compared to observations from the SDSS and it turns out that Cold (particle) Dark Matter can explain the observed structures in the universe. First results were published in 2005 [109]. The latest simulation, “Millennium-XXL”, uses a box of 4.1 Gpc and 6720^3 particles in their n-body simulation (30 × more as in 2005). An example for the n-body simulation outcome is given in figure 2.8 [96]. Within the cube, similar structures as in the real universe are observed. The two cubes inside the first cube show the ability to zoom in the later simulation outcome for a detailed analysis. Cold Dark Matter is an imperative to understand the structure of our universe nowadays.

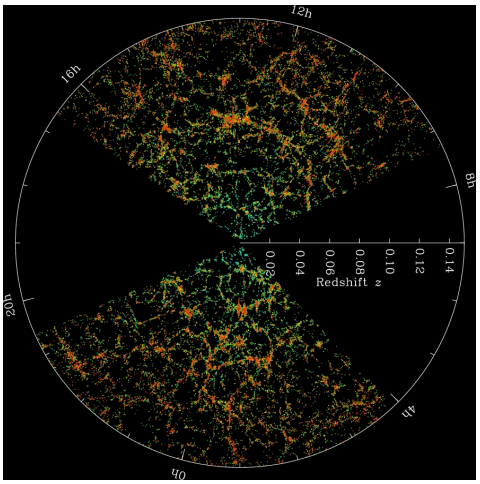


Figure 2.6 – The Sloan Digital Sky Survey (SDSS) with earth in the centre. The points represent galaxies sorted by their red-shift and direction in the sky. [34]

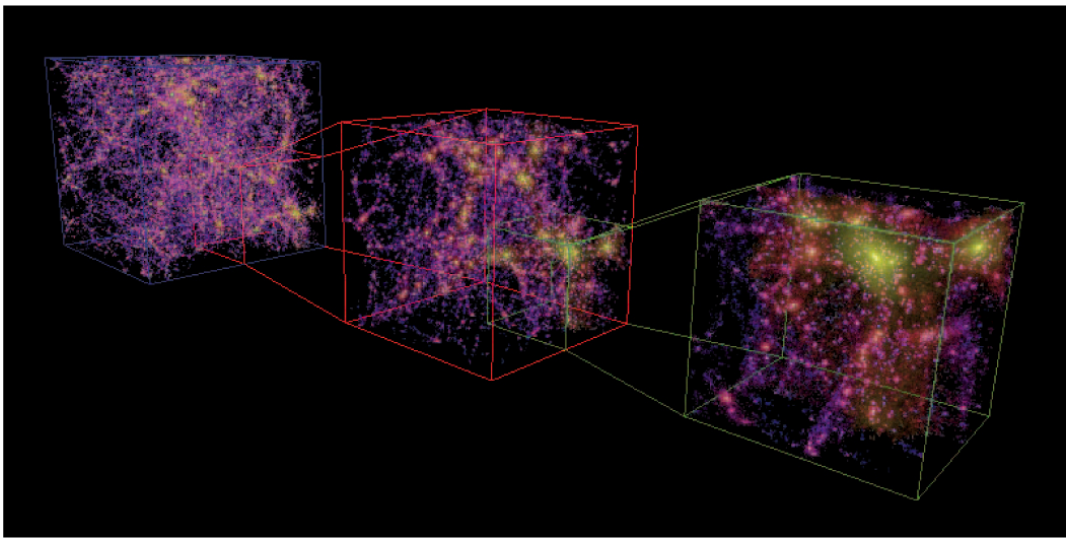
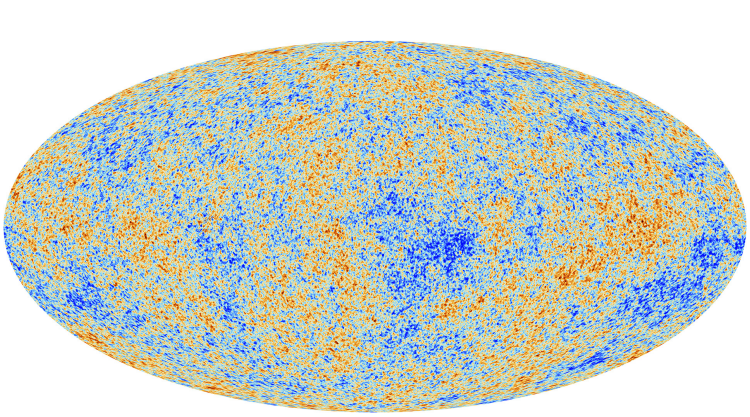


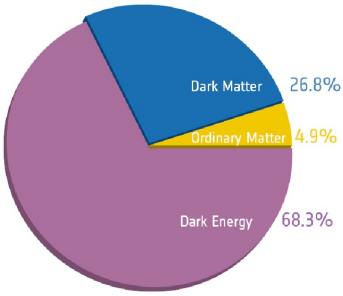
Figure 2.8 – The outcome of the Millennium Run is compared to data (SDSS). Adding Dark Matter to the baryonic matter yields a simulation outcome which agrees very well with sky observations. The simulation suggests a Cold Dark Matter scenario which support the bottom-up scenario. The existence of Dark Matter is an imperative to understand the structure of the universe today. [96]

2.2.4. Dark Matter in the Cosmic Microwave Background

The most important hint for Dark Matter comes from the Planck satellite mission. Previous missions with lower resolution were performed by COBE (launched 1990) and WMAP (launched 2001). The Cosmic Microwave Background (figure 2.9a) present the universe $\sim$ 380,000 years after the Big Bang [54]. This picture shows the last scattering surface of photons with leptons when the universe became transparent. The isotropic CMB was observed at first by Penzias and Wilson in 1964. A black body radiation of $T = 2.72\,\mathrm{K}$ was measured which also has confirmed the theoretical prediction. The newest high precision instruments searching for temperature fluctuations ($\Delta T \sim 2 \times 10^{-5}\,\mathrm{K}$) in the sky that provide information about the mass density in the early universe. It is expected to observe a homogeneous temperature distribution in the whole sky. Small deviations from the homogeneous temperature distribution indicate the existence of further matter contributes, which is in consistence with observations of galaxies or galaxy clusters. Latest measured temperature fluctuations of the CMB are explained by 4.9% baryonic matter, 26.8% Dark Matter and 68.3% Dark Energy (see figure 2.9b) [56]. The CMB tests the Big Bang hypothesis of a cooling universe. Due to the expansion since the Big Bang, the universe becomes cooler from the beginning and matter has been formed which we observe today. This observation fits into the BBN where the content of baryonic matter is predicted by the ratio of photons to baryons [80]. The expansion is slowed down by the gravitational components (baryonic and Dark Matter). Actual results from CMB studies show that our universe is "flat" ($\Omega_{tot} = 1$). The



(a)



(b)

Figure 2.9 – Figure 2.9a: The Cosmic Microwave Background $\sim$ 380000 years after the Big Bang seen by the Planck satellite mission in 2013. Based on a detailed analysis of the small temperature fluctuations (blue to red over the overall temperature $T = 2.72\,\mathrm{K}$) the main components of the universe are determined: baryonic matter, Dark Matter and Dark Energy. Figure 2.9b: The single contributions to the total content of the universe are summarised in the pie chart. The main part of the 4.9% baryonic matter exist in the interstellar medium and gas clouds.

observed accelerated universe does not fit into the picture of a flat universe and hence the additional Dark Energy accelerates the expansion. The according theory is called ΛCDM, where Λ stands for the Cosmological constant which was introduced by Albert Einstein to describe the expansion of the universe.

The total unit-less content (nowadays) in the universe is defined in equation (2.21). Here the parameter Ω_i with $i \in [b, DM, \Lambda]$ describes the particular share (baryonic matter, Dark Matter or Dark Energy) over the critical density: $\Omega_i = \rho_i(t)/\rho_c$.

$$\Omega_{tot} = \Omega_\Lambda + \Omega_b + \Omega_{DM} \tag{2.21}$$

The single contents are determined in a multi-pole analysis of the Planck collaboration. The result is summarised in figure 2.10. The power spectrum starts at largest angle scales and becomes interesting below $\sim 1\,^\circ\mathrm{C}$ where the baryonic Matter part explains the first peak with baryonic acoustic oscillations. The following peaks (second and third) are hints for a Dark Matter part in the temperature fluctuations of the CMB.

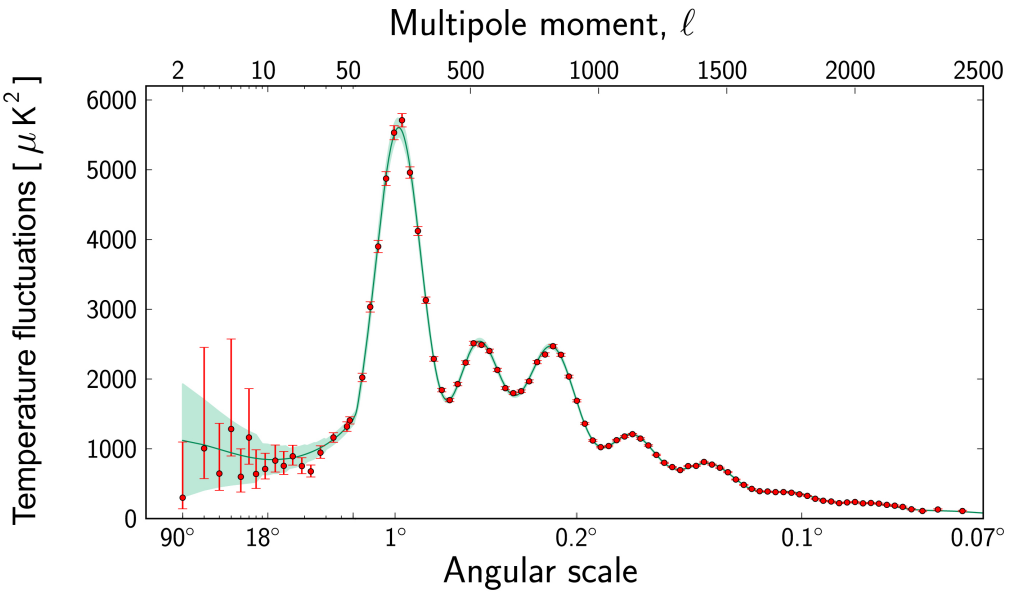


Figure 2.10 – The powers spectrum of the small temperature fluctuations [55]. The spectrum starts at largest angle scales of 90°C. The interesting region begins with angle scales below $\sim 1°$Cwhere the Baryonic Matter share is determined by temperature fluctuations following baryonic acoustic oscillations which are explained by a Dark Matter share in the CMB.

The analysis of the temperature spectrum $T(\theta,\phi)$ is done by a transform into spherical harmonics:

$$T(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm}\, e^{im\phi} P_l^m(\cos\theta) \tag{2.22}$$

In equation (2.22), the $e^{im\phi}$ is the decomposition in azimuths direction. The altitude direction θ is expressed by the according Legendre polynomial. The temperature fluctuations are given by each coefficient a_{lm}. The indices l and m are related to the expansion, which is related to the accuracy of the oscillation size. Therefore the temperature fluctuations should only depend on l and not m and thus the definition of equation (2.23) is valid:

$$C_l \equiv \langle |a_{lm}|^2 \rangle \tag{2.23}$$

The brackets $\langle\rangle$ express that the temperature fluctuations are Gaussian distributed with the expectation value a_{lm}. The resulting power spectrum, the analysis of the C_l, is explained now by the following contributes:

Sachs-Wolfe-effect for the angle scale from 90° to roughly 2°C. The matter distribution was already an-isotropic at small scales 380,000 years ago. Photons which enter such a gravitational potential $\Phi(x)$ gain energy. During the time of travel within $\Phi(x)$ the expansion leads to a lower gravitational potential. Hence, a photon needs less energy to exit $\Phi(x)$. The energy portion ΔE_γ which is left after the travel through the gravitational potential is responsible for the observed red-shift of these photons when they leave the gravitational potential. This process triggers the nowadays isotropic behaviour of the CMB on large scales [101].

The **first acoustic peak** comes from a photon-baryon field before 380,000 years after the BB. Baryons cluster where gravitational potentials $\Phi(x)$ are localised. Therefore the baryonic plasma becomes hotter due to the higher density. As a consequence, energy is released by photons from $\Phi(x)$ (radiation) which prevents further baryons from entering the gravitational potential. Hence the temperature decreases again the density is reduced until the point where the $\Phi(x)$ overtakes again the process and starts to cluster baryons again. The exchange between radiation and gravitation mode results in a longitudinal acoustic oscillation. When the cosmic evolution reaches 380,000 years after the BB, the process stops and only the acoustic oscillations remain in the CMB which are observed today. The best fit to the first peak of the power-spectrum predicts a flat universe $\Omega_{tot} = 1$ which is in consistence with supernova "Ia" results. The term "acoustic" originates from the width of the peak. The related physical length scale suggests the sound horizon which evolves with $v_s = c_s \times t_{rec}$. There c_s is the speed of sound in the baryonic Matter and t_{rec} is associated with the time of the last scattering. As result, the position of the first peak in the spectrum at $l \backsimeq 200$ speaks for a flat universe [41, ff. 383]. The amplitude in the power-spectrum is related to the inert baryon density which is responsible for the acoustic peak itself (compression).

The **following peaks** in the power-spectrum are also related to oscillations. Additional baryons in the potential wells add gravitational force to the system. This results in another compression of the baryons which enter the well. On the other hand, the radiation loss by photons which prevent further baryons from entering the well is still working and the overall oscillation becomes asymmetric and further amplitudes enter the power-spectrum. The additional baryons acts like a damping factor in the oscillation. The second amplitude stands here for the rarefaction of the baryons of the gravitational wells.

Higher order peaks are in general associated with acoustic peaks from the baryon compression (odd peaks) and the rare factions (even peaks). Therefore the power-spectrum presents a temperature fluctuation which consists of two oscillation modes on top of each other. The even peaks always show lower amplitudes than the odd peaks, for example the second peak is slightly lower than the third peak in the power-spectrum of figure 2.10.

At this point Dark Matter enters the stage: Calculations with different baryon densities do not explain why the second peak is just slightly lower than the third peak in the power-spectrum and it is suggested to add another quantity to the universe which explains the increased gravitational potential: $\Phi(x) \rightarrow \Phi(x) + \Phi_{DM}(x)$. Nevertheless, the power spectrum is exponentially damped and higher peaks are strongly suppressed. The damping causes much smaller physical scales in the sky. The reason for that behaviour comes from the random walk of photons through the baryons at that time. The mean free path of photons is determined by Thomson scattering (electron scattering) and therefore physical scales in the universe are suppressed, they were smaller than the Thomson scattering was suppressed at that time.

2.2.5. Cosmological Dark Matter Candidates

A different attempt to explain the observed Dark Matter in galaxies are MACHOs: *M*assive *A*strophysical *C*ompact *H*alo *O*bjects. A first approach is to count the stars and calculate their mass out of their luminosity by the mass-luminosity relation (see equation (2.2)). MACHOs consist of stars which are not seen due to their low luminosity. The Dark Matter mystery remains and possible explanations are given by objects which simply do not shine bright enough and hence are not seen by optical or X-ray telescopes. The range of possible MACHOs reaches from exoplanets and brown dwarfs to neutron stars and black holes. These objects are measured by exploiting their gravitational interaction with neighbour objects and do not emit light by themselves. Other possible objects which emit light are white dwarfs (stellar remnant of stars with $m < 1.4\,\mathrm{M}_{\odot}$) or red dwarfs (cool star of the main sequence). As long as their total number is not determined in single galaxies, their contribution to the total mass is uncertain. This subsection will quickly summarise the possible candidates, give a short introduction into the MACHO search process and summarises the results.

Exoplanets and brown dwarfs bring up the possibility that almost each star is surrounded by a given number of planets similar to our stellar system. These objects do not emit light and therefore they are observed indirectly by a bunch of methods such as: Transit timing/duration variation (TTV/TDV), gravitational micro-lensing and many more. In total 1789 exoplanets are discovered mainly by transit measurements [9]. The heaviest object ever found is "DENIS-P J082303.1-491201" and has a mass of $m = 5.4 \times 10^{28}\,\mathrm{kg} = 28.5\,\mathrm{m}_{\mathrm{♃}}$ (in units of Jupiter masses) [102]. Exoplanet searches are not sensitive to the mass, size and orbital position and therefore the discrepancy from rock planets to huge gas planets is floating. Brown dwarfs are a special type of exoplanet. These objects became not massive enough during the evolution of the stellar system to start a fusion process (hydrogen-burning) like a main sequence star. Hence a huge mass of $m_{bd} < 0.07\,\mathrm{M}_{\odot}$ is clustered there [64, 86]. Nevertheless these objects do not emit visible light and detailed studies are difficult.

White dwarfs, neutron stars and black holes are star remnants from main sequence stars. Their stellar evolution depends mainly on the star itself: Lighter objects with a mass below the Chandrasekhar-mass limit of $m < m_{Ch} \sim 1.45\,M_\odot$ end as white dwarfs. The Chandrasekhar-mass limit is adjusted by the nucleon to electron ratio [46]. Objects above m_{Ch} collapse due to their mass up to the point where the star remnant becomes so dense and hot that atomic nuclei degenerate. Hence neutrons are left and further neutrons are added by the protons which combine with free electrons due to the high pressure. The result is an ultra compact object with a radius of $r < 15\,\text{km}$ and a maximum observed mass of $m_{ns} = 2\,M_\odot$, which is a neutron star. The density of such objects varies with the radius. Due to rotation speed of the star before the collapse the final neutron star rotates with up to $\sim 716\,\text{Hz}$ (from observation). A black hole is the final evolution of high mass objects. They posses the highest known density in the universe. Black holes are able to create accretion disk of baryonic matter around them. Black holes are e.g. the core objects of galaxies and if they interact with their accretion disk they become visible (active galaxy nuclei, AGN). These AGNs are detected in optical light, X-ray or by gravitational measurements. But lighter black holes are hard to find and therefore its difficult to estimate their contribution to the total mass.

The existence of possible MACHO candidates is proved but are there enough to explain the observed Dark Matter distribution? For this it is important to understand more about the nature of MACHOs. Such objects only contribute to the Dark Matter distribution if they are inside the galactic halo. The latest results from Monte Carlo simulations of halo binaries and their comparison with sub-samples lead to a typical MACHO mass of $m = 21 - 68\,M_\odot$ [85]. Earlier micro-lensing studies of the Large Magellanic Cloud (LMC) suggest a total MACHO mass of $m_{tot} = 9 \times 10^{10}\,M_\odot$ within 50 kpc independent from the halo model [11]. In summary the MACHO objects density is not high enough to explain the observed missing mass and thus a particle Dark Matter candidate as it is discussed in section 2.3 is favoured.

2.3. Introduction to Particle Dark Matter

The past section reviewed Dark Matter from a cosmological view, beginning with the missing mass problem in galaxies and galaxy clusters up to the hints from the CMB. Possible baryonic matter candidates such as MACHOs are not sufficient to explain the observed Dark Matter. Therefore a new particle type is a possibility to solve the Dark Matter mystery. This kind of non-baryonic matter requires an extension to the standard model of particle physics. There is a set of possible particle Dark Matter candidates such as neutrinos, axions or WIMPs. Each type provides a specific production mechanism. As long as the true nature of Dark Matter is not revealed, different theories could be true

at the same time and all their single Dark Matter mass predictions contribute to the total mass in the universe.
The following section covers the necessary conditions for such a Dark Matter particle and introduces possible candidates. XENON100 and the successor experiment XENON1T aim to detect (WIMP) Dark Matter directly. Therefore the Dark Matter search methods are summarised finally with focus on *direct* Dark Matter detection.

2.3.1. Characterisation Based on Cosmological Observations

Before the production mechanism and Dark Matter candidates are discussed in detail, a first classification is made. Dark Matter scenarios are distinguished by: *Cold Dark Matter* (CDM), *Warm Dark Matter* (WDM) and *Hot Dark Matter* (HDM). Each scenario is related to a specific production mechanism and particle properties such as mass and cross-section. The difference is less important for the particle Dark Matter type, but the classification is necessary to specify a particle type which explains the cosmological observations as they are discussed in the previous section.

2.3.1.1. Cold Dark Matter

The Cold Dark Matter scenario (CDM) ties up to section 2.2.5. Another source of non-luminous matter is added, a type of particles which moves slowly compared to the speed of light. This scenario allows to describe the observed structures of galaxies and galaxy clusters where Dark Matter acts as a seed where baryonic matter clusters around by gravity. Many theories explain these particles: The most favoured theory is WIMP Dark Matter which fulfils the necessary conditions from a cosmological view (see section 2.3.2.1). Other theories suggest the existence of axions which arise from the QCD. These and further ideas extend the SM. For example, supersymmetric extensions of the standard model have been discussed for quite some time (SUSY).
Unless the true nature of Dark Matter is not revealed it is possible that Dark Matter consists of a bunch of Dark Matter particles which fulfil the necessary conditions. Therefore a selection of experiments look for the different types of WIMP-like CDM. A detailed view on the production mechanism is given in section 2.3.2. Related to direct Dark Matter detection experiments such as XENON100 further observation theory is given in 2.4.

2.3.1.2. Hot Dark Matter

The CDM scenario with a new non-baryonic particle is not necessarily the only explanation for Dark Matter. Three types of neutrinos (ν_e, ν_μ, ν_τ and their anti-particles) contribute to the total mass in the universe. Nowadays these neutrinos are produced in the pp-chain, CNO cycle of stars or supernovae. But there is a relic from the Big Bang, the Cosmological Neutrino Background (CNB). This scenario leaves a neutrino background of $T_\nu = 1.95\,\mathrm{K}$ nowadays, which corresponds to an energy of $E_\nu = 1.676 \times 10^{-4}\,\mathrm{eV}$. These relic neutrinos are too low in energy to detect them because of the energy detection thresholds of modern instruments. When the universe expands, the interaction cross-section for neutrinos decreased. At some point, the neutrino production rate became smaller than the expansion rate and the neutrinos decoupled. The term Hot Dark Matter comes from the fact that neutrinos ($m_\nu < 2\,\mathrm{eV}$) were relativistic at the moment of their decoupling temperature T_{dec} [59]. A detailed view on the CNB and the neutrino production mechanism is given in section 2.3.2.4.

2.3.1.3. Warm Dark Matter

There is an open discussion about the existence of *sterile neutrinos* which also may account for Dark Matter. These additional neutrinos alongside the already known three neutrinos have a non-zero mass. Introduction a forth neutrino implies to extend the mixing matrix with a rang of 3×3 to a rang of 4×4 or higher. The rang depends then on the number of sterile neutrinos which are added as an extension to the standard model. This open up the possibility to introduce a particle which explains the Dark Matter without the need of a non-baryonic particle. The SM of electro-weak interactions is extended to νMSM [27]. The added neutrino(s) are right-handed. Up to now there are only left-handed neutrinos observed [51]. Latest results from the XMM-Newton satellite mission reports a 3.5 keV X-ray emission line in the Andromeda or Perseus galaxy cluster. Both results could be interpreted as a hint for Dark Matter if the line origins from sterile neutrinos. The DM type would have a lifetime of $\tau_\chi = 7.2 \times 10^{29}\,\mathrm{s}$ and a mass of $m_\chi = 7.06\,\mathrm{keV}$ [36].

2.3.2. Dark Matter Candidates

Subsection 2.2 establishes Dark Matter as a necessary occurrence for our universe. Unfortunately, the pure observation of missing mass is not enough to characterise the true nature of Dark Matter. The following section introduces possible Dark Matter candidates from the particle physics point of view. The first subsection summarises the necessary requirements for Dark Matter particles. Different candidates such as WIMPs, axions, neutrinos or supersymmetric extensions are presented. The XENON

Dark Matter Project aims to detect WIMPs directly and therefore the focus in this section points to the WIMP Dark Matter scenario.

2.3.2.1. Requirements on Dark Matter Candidates

Decades of experiments on accelerator facilities have lead to the development of a well understood model of particle physics: The standard model. Figure 2.11 summarises the three quark families with their additional leptons: $(u|d|e^-|\nu_e)$, $(c|s|\mu|\nu_\mu)$, $(t|b|\tau|\nu_\tau)$. These twelve fermions are enough to explain our nowadays observed matter in the universe. Four bosons γ (E/M), $W^\pm/Z$ (weak) and g (strong) intermediate the forces. These forces hold e.g. the quarks together by gluons so that protons or heavier compounds can be formed (mesons, baryons). In general, their most important properties are: mass, charge, spin, parity and lifetime. Finally, the Higgs boson is associated with the scalar Higgs field which is responsible to give mass to each particle. It was discovered at the Large Hadron Collider (LHC) in Switzerland [1]. Detailed studies confirm its mass at $m_H = 129.9\,\mathrm{GeV}/c^2$ with spin 0 and positive parity. The particle properties (fermions with spin ½, bosons with spin 0) in the SM are well measured today. These properties are related to the force exchange particle (bosons) which explain how such particles interact (e.g. conversation principle). No particle Dark Matter candidates have been observed yet, but from cosmological observation, some general properties can be determined:

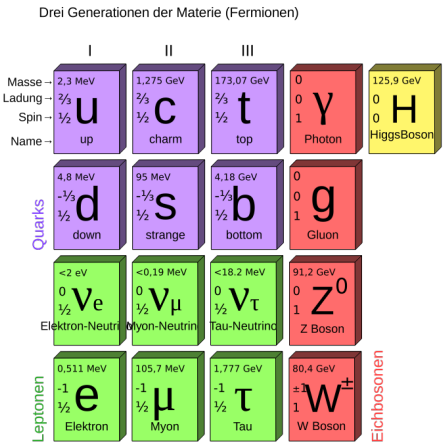


Figure 2.11 – The standard model (SM) of particle physics consists of three families of quarks with their additional leptons which are assembled in columns. Four bosons which intermediate the force (E/M, weak, strong) and the Higgs boson.

1. Dark Matter interacts by gravity and weak interaction.
2. Due to observed missing mass the interaction probability of such a Dark Matter particle with baryonic matter is small. The cross-section is expected in the electro-weak range.
3. It is charges-less and colour-less.
4. Dark Matter is stable, or at least a long-living particle.

These properties were developed from the fact that the missing mass is observed indirectly. The intermediate force is gravitation. The gravitation of force is not yet covered by the SM and a combined

description is not possible. Further direct interaction with photons, $Z/W^{\pm}$ or the strong force are not observed. Related to that, Dark Matter particles are charge-less and colour-less. Due to the fact that we observe Dark Matter nowadays, the Dark Matter particle itself is stable or even long-lived.

Unfortunately there is no proper particle candidate in the SM which explains Dark Matter. Neutrinos are special in terms of being a potential Dark Matter candidate. Due to the observed neutrino oscillations they owe a mass which is limited to $m_{e,\mu,\tau} < 2\,\mathrm{eV}$. Their spin is $1/2$, they are stable and have no charge. But the observed neutrino particle number density is not large enough to explain the huge amount of observed Dark Matter.
The following subsections introduce some particle Dark Matter candidates. This selection is not completely and present only a choice of the most common ideas to describe the missing mass problem with an additional particle.

2.3.2.2. Axion and Axion-like Dark Matter Candidates

The axion is a theoretical particle which was introduced by Roberto Peccei and Helen Quinn in 1977 with a phase SU(2) $\times$ U(1) as single scalar field [73]. In theory, the QCD Lagrangian is extended by another parameter which allows CP (charge, parity) violation. The idea beyond was to explain why the CP symmetry is not broken for strong interactions in QCD. The additional term should be almost zero [93], then CP is violated in QCD. But that additional term leaves a suitable Dark Matter candidate. The postulated axions are in the keV regime and their probability to interact with baryonic matter is very low. Given that, the axion is charge-less and fulfil all required properties to be a Dark Matter particle in a CDM scenario (e.g. non-relativistic particles). The axion is a quasi-Nambu-Goldstone boson with $a(x)$, the axion field. The mass is given by equation (2.24) and depends on the axion decay constant f_a which contains the vacuum expectation value [32].

$$m_a \backsimeq 0.6\,\mathrm{eV}\frac{10^7\,\mathrm{GeV}}{f_a} \tag{2.24}$$

Cosmological constraints demand light axions which appear in two similar scenarios. The simplest axion model refers to Kim-Shifman-Vainstein-Zakharov (KSVZ) where the axion field $a(x)$ couples to heavy quarks [71]. Higgs doubles are assumed to be neutral under $U(1)_{PQ}$ symmetry. This changes for the Dine-Fischer-Srednicki-Zhitnitskii axion model (DFSZ) [49]. There, the axion field couples to the Higgs doubles (H_u/H_d) and there is no need to introduce heavy quark masses [73]. In both cases the axion interacts with photons of the **E** and **B** field by the coupling constant.
Solar axions are produced in the core of stars by four mechanisms: Bremsstrahlung, Compton scattering, axio-recombination and axio-deexcitation [97]. These solar axions leave the core where their flux

is much below the luminosity flux. Another approach are galactic axions, they test the axion decay constant of equation (2.24) which is necessary to determine the axion mass. Therefore it is assumed that galactic cluster in the halo of our galaxy and being studied to find sources of *invisible* axions which are not produced in solar fusion processes.
There are many earthbound experiments which probe the axion hypothesis. The proceeding is similar for all the experiments. An incoming axion couples to photons, electrons or to the nucleus. In every case another coupling parameter becomes relevant and depends on the axion energy E_a, the axion velocity β_a and the target mass. XENON100 searches for electronic recoils (ER). Here an axion interacts with a shell electron of the LXe and produces free charge which is detected in the gaseous phase (ionisation signal). In figure 2.12a the photo-electric absorption is drawn for xenon (A = 54). The photo-electric absorption dominates for energies $E_\gamma < 1\,\mathrm{MeV}$. Hence the axion cross-section is proportional to the photo-electric cross-section and light axion search determines a cross-section of:

$$\sigma_{ae} = \sigma_{pe}(E_a)\,\frac{3\,g_{ae}\,E_a^2}{16\pi\,\beta_a\,\alpha_{em}\,m_e^2}\left(1-\frac{\beta_a^{2/3}}{3}\right)$$

The four production mechanisms are covered by the coupling constant g_{ae} which goes into the cross-section calculation for axions. Latest results from the XENON100 collaboration in 2014 show upper limits for the coupling constant g_{ae} in dependence of the axion mass (figure 2.13). The studies separate axions from the sun (left of 2.13) and from galactic origin (right of 2.13). Both axion models have a different input for the later calculation. In case of solar axions the axion flux is estimated before [97]. The study of galactic axions demand that the whole Dark Matter halo consists of these axions [18]. Therefore the DM density is $\rho_{DM} \sim 0.3\,\mathrm{GeV/cm^3}$ [61].

Besides axion searches via the axion-electric coupling, other experiments such as CAST (CERN Axion Solar Telescope) search for solar axions by the Primakoff effect. Axions leave the core of the sun and interact with photons of a magnetic field in the detector (see Feynman diagram in figure 2.12b). The interaction creates a photon which is detected by photomultiplier cathodes. Actual results from CAST set a maximum solar axion mass to $m_a < 10\,\mathrm{meV/c^2}$ [68]. No experiments report on axion observation up to now.

Due to the expected light axion mass from both solar and galactic axions, the axion particle density is to low to explain cosmological observations only by axions. Therefore axions are expected only to be a minor part of the Dark Matter in the universe if they exist.

2.3.2.3. Weakly Interacting Massive Particles

The nowadays favoured Dark Matter particles are WIMPs: *W*eakly *I*nteracting *M*assive *P*articles. This class of Dark Matter particles fulfil all required conditions of section 2.3.2.1 by definition and fits into the CDM scenario. Keystone of that theory is the production mechanism which is analogous to the relic neutrino background of section 2.3.2.4. Cross-section σ_χ and mass m_χ are free parameters and depend on the WIMP theory itself and the interaction type with baryonic matter (e.g. spin-independent or spin-depend). This subsection will focus on the WIMP production mechanism and the relic density of WIMP DM Ω_χ. The XENON100 experiment aims to detect Dark Matter with a liquid xenon target by direct detection and therefore section 2.4 summarises the necessary theory in detail.
The WIMP production happens in the early thermal plasma of the radiation dominated universe. But that time, the temperature was high enough ($T \gg m_\chi$) to produce WIMPs in a stable equilibrium. Important processes are given in (2.25) [32].

$$\chi\overline{\chi} \leftrightarrow e^- e^+, \mu^- \mu^+, q\overline{q}, W^+W^-, ZZ, HH, ... \tag{2.25}$$

The annihilation rate is $\Gamma_{ann} = \langle\sigma_{ann}v\rangle n_{eq}$ with the annihilation cross-section σ_{ann}, the particle velocity v and the number density n_{eq}. The number density refers to the time when the Dark Matter and baryonic matter were in equilibrium. The brackets denote a thermal dictated distribution with a sub-dominant annihilation process. When the universe expanded, σ_{ann} decreased and also the production rate of WIMPs ($\propto e^{m_\chi/T}$). During the expansion the WIMP particle density n decreased and at the time when the production rate became smaller than the annihilation rate, the WIMP Dark Matter froze out. The freeze-out process is described by two differential equations (2.26) and (2.27) [32].

$$\frac{dn}{dt} = -3Hn - \langle\sigma_{ann}v\rangle(n^2 - n_{eq}^2) \tag{2.26}$$

$$\frac{ds}{dt} = -3Hs \tag{2.27}$$

Equation (2.26) describes the falling number density with time on the left side related to the number density in equilibrium n_{eq} (right side). This situation is compensated by the Hubble expansion which is also used in equation (2.27) to describe the falling entropy density s. Substitute $Y = n/s$ and $x = m/T$ results a dimension-less energy scale and equation (2.26) and (2.27) are combined in equation (2.28).

$$\frac{dY}{dx} = -\left(\frac{45}{\pi m_P^2}\right)^{-1/2} \frac{g^{*1/2}}{x^2}\langle\sigma_{ann}v\rangle(Y^2 - Y_{eq}^2) \tag{2.28}$$

where $g^*(T)$ is derived from thermodynamics and describes the number of degrees of freedom at a given temperature T in the universe and is given by:

$$g^{*1/2} = \frac{h_{\text{eff}}}{g_{\text{eff}}^{1/2}} \left(1 + \frac{1}{3} \frac{T}{h_{\text{eff}}} \frac{dh_{\text{eff}}}{dT} \right) \tag{2.29}$$

Equation (2.28) is solved numerically with $Y = Y_{eq}$ at $x \backsimeq 1$ to obtain the present WIMP abundance Y_0. The WIMP density is then calculated to:

$$\Omega_\chi h^2 = \frac{\rho_\chi^0 h^2}{\rho_c^0} = \frac{m_\chi \, s_0 \, Y_0 \, h^2}{\rho_c^0} = 2.755 \times 10^8 \, \text{Y}_0 \, \text{m}_\chi / \text{GeV} \tag{2.30}$$

Where ρ_c^0 and s_0 are the nowadays critical density and entropy. m_P is the Planck mass. The numerical result is summarised in figure 2.14. The relic WIMP number density decreases due to the decreasing annihilation rate. At a certain point the WIMPs freeze-out happens at a temperature T_{fo}. At lower temperatures the WIMP number density flattens out and settles down at a certain Dark Matter density Ω_χ. This flatten-out point in time depends on the WIMP mass. An important parameter is the remaining WIMP velocity during/after the freeze-out which is given in $\langle \sigma_{ann} v \rangle$. The freeze-out takes place when the WIMP particle density runs out of the equilibrium n_{eq}. This point is proportional to the Hubble constant H and freeze-out temperature T_{fo}. Equation (2.32) can be used to show that, using $x = m/T_{fo}$, the freeze-out is $v = \sqrt{3/x}$. Therefore the WIMP velocity is independent of temperature and mass and only depending of the expansion rate of the universe. A typically freeze-out velocity is $v_{fo} \sim 0.3\,c$.

$$n_{eq} \langle \sigma_{ann} v \rangle \sim H \tag{2.31}$$

$$\frac{1}{2} m v^2 = \frac{3}{2} T \tag{2.32}$$

Based on that discussion, the freeze-out temperature has the largest impact on the WIMP relic density. Therefore other parameters such as m_χ or σ_χ are important to take into account, but overall the Hubble parameter H needs to be measured as well as possible (e.g. by the Planck satellite mission) to calculate the final density. The present relic density is approximately given by [32]:

$$\Omega_\chi h^2 \approx \frac{3 \times 10^{-27}}{\langle \sigma_{ann} v \rangle} \frac{\text{cm}^3}{\text{s}} \tag{2.33}$$

From the previous derived equations, the most important fact is the observed WIMP density Ω_χ which is inverse proportional to the cross-section and velocity from thermal production mechanism. As

long as there is no WIMP detection, the mass and cross-section will remain unknown. Therefore the detection of a particle χ would follow equation (2.34).

$$\Omega_\chi \propto \frac{1}{\langle \sigma v \rangle} \sim \frac{m_\chi^2}{g_\chi^4} \tag{2.34}$$

The WIMP coupling g_χ parameter is taken from the electro-weak scenario where $g_\chi \sim g_{weak} \backsimeq 0.65$. The advantage of equation (2.34) is the WIMP density which is given by independent experiments, such as the Planck satellite mission which studies the CMB. Actual results suggest a $\Omega_\chi = 0.27$ which leads to a WIMP mass of $m_\chi =\sim 100\,\mathrm{GeV} - 1\,\mathrm{TeV}$. This mass range fits the prediction of a new particle from the electro-weak sector. This coincidence is named *WIMP miracle* commonly.

Besides the thermal production mechanisms, non-thermal scenarios are possible which follow other decay processes and relic densities. Furthermore the introduction of WIMPZILLAs is discussed [74]. These scenarios exclude low-mass WIMPs and favour heavier WIMPs. Which theory will ever be proved by particle Dark Matter experiments in the future, its results must fit in the cosmological observations. Whatever WIMP mass or range will be observed, the total expected mass must fit the observations from galaxy clusters and CMB studies. On the other hand it is not enough to predict the mass only. The WIMP production mechanism must fit into the bottom-up scenario and describes the correct amount of expected Dark Matter that galaxies and galaxy clusters could be created. Only this scenario follows observations from cosmology and explains the formation of (super) galaxy clusters [94].

2.3.2.4. Neutrinos

Besides the CDM scenario where *new* particles are introduced to explain the missing mass in the universe, neutrinos also contribute to the total Dark Matter mass. Three types of neutrinos are observed: ν_e, ν_μ and ν_τ (plus their anti-particles). Due to the observed neutrino oscillation from the latest reactor or accelerator experiments such as Daya Bay [12], RENO [8], Double Chooz [4], MINOS [5] and T2K [3], the neutrino mass is essential. The neutrino flavour (e, μ, τ) mixing and hence the mass mixing is measured to $\Delta m_{21}^2 = 7.6 \times 10^{-5}\,\mathrm{eV}^2$ and $\Delta m_{32}^2 = 2.4 \times 10^{-3}\,\mathrm{eV}^2$ [3] in 2014 (T2K Collaboration). In contrast to the measured mass differences the final mass of each neutrino is not fixed yet. The hierarchical structure is not determined up to now. The Troitsk experiment set an upper limit on the electron anti-neutrino mass of $m_e < 2\,\mathrm{eV}$ [28] in 2011.

Besides the earthbound experiments other detectors explore the disappearance from solar neutrinos. Due to the production of an ν_e in the pp-chain or the CNO cycle of the sun the neutrino flux on

earth (assume $P_{\overset{+}{\odot}}(\nu_e \to \nu_e)$, the probability for a non-oscillating neutrino) is given as $\Phi(\nu_e) < 7.7 \times 10^8\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (95% CL) [42]. Hence the non-vanishing neutrino fluxes of ν_μ or ν_τ becomes other strong evidence for neutrino oscillations: $\Phi(\nu_\mu \text{ or } \nu_\tau) = 3.26 \times 10^6\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [62, ff. 247]. This result is in agreement with cosmological studies where upper total neutrino mass limits are set. The latest results from the Planck satellite mission determine the total neutrino mass $\sum m_\nu \leq 0.23\,\mathrm{eV}$ from studies of the CMB [6]. Based on these observations on the mass, the neutrino becomes a possible Dark Matter candidate.
The neutrino masses have a big impact on weather neutrinos are responsible for, or might be a part of the missing mass problem. Neutrinos fulfil the most important properties of Dark Matter: They interact only by gravity and the weak force with baryons and leptons, are charge-less and stable. As long as neutrinos are produced in solar processes or supernovae the neutrino number is proportional to the number of stars and their lifetime. There is another source, the Cosmological Neutrino Background (CNB) which is also explored by WMAP [75] and latest by Planck [6].

Photons of the early universe with an energy of $1\,\mathrm{MeV} \leq T_\gamma \leq 100\,\mathrm{MeV}$ were in equilibrium with the neutrinos of the primordial plasma (see equations (2.35) and (2.36)). The interaction is carried out by the weak force.

$$\nu + \bar{\nu} \rightleftarrows e^+ + e^- \tag{2.35}$$

$$\overset{(-)}{\nu} + e^\pm \rightleftarrows \overset{(-)}{\nu} + e^\pm \tag{2.36}$$

During that time the production rate of such neutrinos was the same as the decay rate. This situation changed when the universe cooled down to photon temperatures of $T_\gamma \sim 1\,\mathrm{MeV}$. The neutrinos decoupled from the equilibrium. After $T_\gamma < T_{dec}$ the temperature profiles decreased different for photons and neutrinos. The neutrino temperature developed with $T_\nu \propto R(t)^{-1}$, where $R(t)$ is the scale factor (expansion factor) of the Robertson-Walker metric. The situation was another one for the photons. Due to e^-/e^+ annihilation their energy was transferred to photons up to their nowadays observed temperature of $T_\gamma^0 = 2.73[K]$ in the CMB (see section 2.2.4). Since the time when neutrinos decoupled, both temperature decreased differently but they are connected by equation (2.37).

$$T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma \tag{2.37}$$

The relic neutrinos count to the HDM scenario. Neutrinos were relativistic before they decoupled. Hence there is still a quantity of such light mass neutrinos below T_ν which are still relativistic. The

number density for relic neutrinos is independent from the neutrino mass m_i ($i = [e, \mu, \tau]$) in the HDM scenario. Their energy density is given by the photon density Ω_γ^0 in equation (2.38).

$$\Omega^0_{\nu-\text{relativistic}} = \left(\frac{4}{11}\right)^{4/3} \Omega^0_\gamma \tag{2.38}$$

From red-shift calculations the neutrino mass is estimated to $m_i = 1.5\,\text{eV}$. Here it is assumed that all three neutrino types possess a mass and in addition, it exists a red-shift where the neutrinos become non-relativistic due the expanding universe. To predict the mass of the light neutrinos it is necessary to calculate the red-shift z_{eq} of matter-radiation equality. Therefore this red-shift must fulfil the condition of:

$$z_{\nu-nr} < z_{eq} \tag{2.39}$$

This equation is true for neutrino masses of $m_i \lesssim 1.5\,\text{eV}$. This result is in good agreement with actual upper limits on neutrino masses. After an upper limit is calculated the number density of light left-handed neutrinos and right-handed anti-neutrinos is used to calculate the total energy density:

$$\Omega^0_\nu = \frac{\sum_{i\in[e,\mu,\tau]} m_i}{94.14\,eV} \tag{2.40}$$

Based on that estimations of the CNB, which decoupled much earlier before the time of matter-radiation density equilibrium, the neutrinos act like a relativistic gas. No large scale structures could be formed during the radiation dominated universe phase. The relic neutrinos move randomly through the universe. This avoids the creation of over-dense structures where baryonic matter clusters. This process continuous on until the neutrinos slowed down to non-relativistic speed at a red-shift of $z_{\nu-nr}$. Therefore only neutrinos with wavelength bigger than the horizon distance at $z_{\nu-nr}$ would be part of the HDM. If the universe would be dominated by HDM during that time the structure was built by a top-down mechanism. There are some statistical observations which favour that scenario. Due to that reason, the scenario of HDM is dis-favoured nowadays and CDM models are more useful to explain the origin of large scale structures by the bottom-up mechanism.
Up to now there are no experiments which probe the CNB. The low neutrino energy needs more precise experiments with much lower energy thresholds. This kind of instrumentation is still under research and development. For example the upcoming PTOLEMY experiment aims to detect relic neutrinos directly with a tritium target on earth [33].

Other candidates for Dark Matter on the neutrino sector are sterile neutrinos. In the beginning of this section, the described neutrinos are left-handed neutrinos and associated with the right-handed anti-neutrinos. Up to now, no left-handed anti-neutrinos or right-handed neutrinos are observed. The extension of minimum one left-handed anti-neutrino or right-handed neutrino (sterile neutrino) would

complete the neutrino model. If sterile neutrinos exist there could be several different sterile neutrinos. Sterile neutrinos become a Dark Matter candidate because they fulfil the necessary requirements (see section 2.3.2.1). The mass of sterile neutrinos is in the range of 1 MeV.
To introduce sterile neutrinos into the already existing picture of normal neutrinos, it is necessary to introduce new mixing angles. As a consequence the Maki-Nakagawa-Sakata-Matrix (MNS-Matrix of rank: 3×3) is extended to allow neutrino oscillations between the already three known types and the possible existing sterile neutrinos. Hence each sterile neutrino increases the rank of the matrix by one (e.g. 4×4) and hence a forth sterile neutrino ν_{st} can oscillate into every other flavour. Neutrinos subjects to the weak interaction and therefore the extended MNS is also parametrised by the Weinberg angle. An important question from cosmology is the half-life of such a sterile neutrino. Without the exact mass it is not possible to answer that question for sure but the Neutrino Minimal Standard Model (νMSM) provides sterile neutrinos with a lifetime of $\Gamma = 10^{17}$ years. Sterile neutrinos are stable at the time scale of our universe.
This is the keystone in Dark Matter sterile neutrinos. Besides the fact that they only interact by gravity, another decay is possible:

$$\nu_{st} \rightarrow \nu + \gamma \tag{2.41}$$

This attractive possibility allows interpreting XMM-Newton satellite data as neutrino Dark Matter. The satellite observes regions in the sky where hints for Dark Matter are already given (e.g. the Andromeda galaxy). Then the data is probed for an unusual γ line in the spectrum. In 2014, an unknown X-ray line was observed in the Andromeda galaxy and the Perseus galaxy cluster. The hypothesis holds a test against the blank sky where neither baryonic nor Dark Matter is expected. The outcome is an unidentified line of $E_\gamma = 3.5\,\text{keV}$ [36]. These excess is not expected in classical galaxy or galaxy cluster spectra and it is possible to interpret the signal as Dark Matter under the assumption of axion Dark Matter models [39]. Besides cosmological hints for sterile neutrinos, earthbound experiments such as Daya Bay also test the hypothesis of observing a four neutrino oscillation. They reject the hypothesis of a forth sterile neutrino in their data from a reactor experiment. Their disappearance spectra of the electron anti-neutrino fits the three flavour neutrino model [13].

2.3.2.5. Supersymmetric Dark Matter

The standard model of particle physics requires six quarks, six leptons and five bosons as shown in figure 2.11. The latest discovery of the Higgs boson and its measured properties result a standard model particle with spin 0 and positive parity. In the SM description quarks and leptons are fermions with spin $1/2$ and the force intermediators are vector bosons (gauge bosons) with spin 0 (γ, $W^{\pm}/Z$, g).

The supersymmetric (SUSY) extension of the SM suggests adding to each fermion its boson equivalent and to each boson its fermion equivalent:

$$Q|\text{fermion}>=|\text{boson}> \quad Q|\text{boson}>=|\text{fermion}>$$

The operator Q provokes the spin change in quantum mechanics. The supersymmetric mirror world is presented in figure 2.15. This step introduces a new symmetry. Hence a fermion (e.g. electron) and its supersymmetric partner (selectron) have the same quantities (e.g. mass or charge, except spin). Lab experiments have not observed any supersymmetric partners yet. If SUSY particles exist, they have a spontaneous broken symmetry in the mass term and therefore they are observed at higher energies. The SUSY extension opens a new parameter space. Possible solutions like the Minimal Supersymmetric Standard Model (MSSM) with 120, up to now, unknown parameters were introduced to solve the occurring hierarchical problems [100]. The heavier SUSY particles are not stable and decay due to their lifetime. Only one expected particle does not decay, the lightest supersymmetric particle (LSP). In the MSSM the R-parity is not conserved for baryon and lepton number. In case of a preserved R-parity for the LSP, its decay is forbidden. This is important to establish a Dark Matter candidate on the SUSY sector.

SUSY candidates for the LSP are neutral winos $\tilde{W}^3$ (supersymmetric partners of the $W^{\pm}$ bosons), binos, neutralinos ($\tilde{B}$) and Higgsinos ($\tilde{H}_1, \tilde{H}_2$) [32, 52] and the resulting linear combination:

$$\chi = \alpha\tilde{B} + \beta\tilde{W}^3 + \gamma\tilde{H}_1 + \delta\tilde{H}_2$$

χ would fit in the CDM scenario. Besides MSSM, further models are discussed such as Constrained MSSN (CMSSN), mSUGRA and many more with a growing number of parameters. Every model yields another CDM density from cosmological observations.

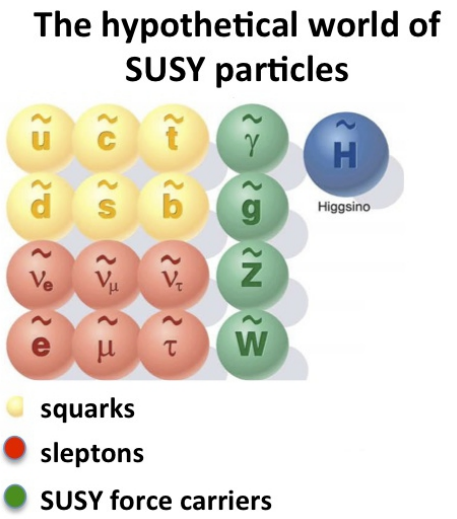


Figure 2.15 – The supersymmetric particle extension to the SM. Each fermion has a supersymmetric partner. [58]

A valid statement which excludes SUSY is the latest observed Higgs boson. The SUSY Higgs boson exists in a complex iso-dublett of Y = 1. Hence there are two, H^+ and H^0, before the electro-weak symmetry is broken. The Q-operator then results two Higgsinos $\tilde{H}^+$ and $\tilde{H}^0$ with spin $s = 1/2$ each. The SM itself provides only one Higgs boson in a minimal config-

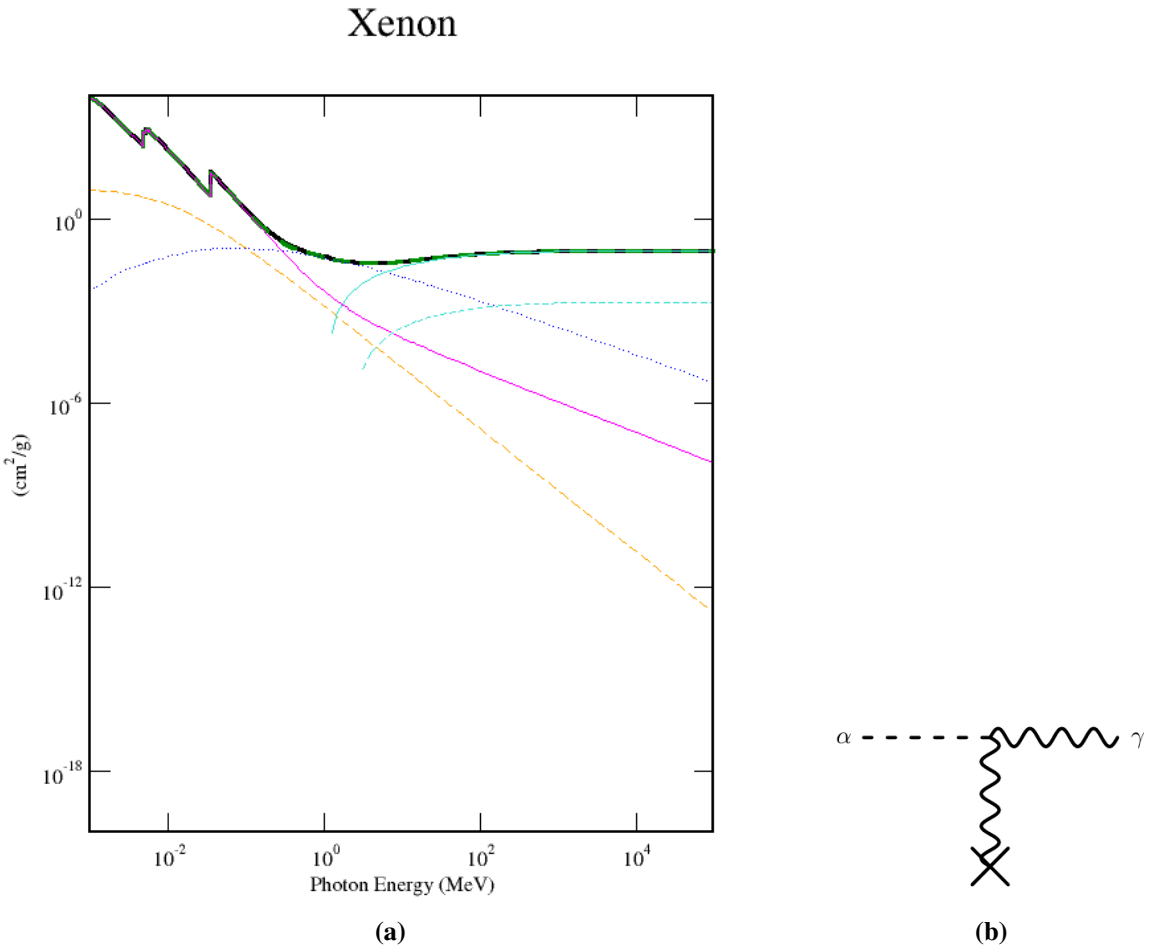
uration because the weak hyper-charge of these processes have to follow a simple rule (see equation (2.42)) to avoid divergences in higher orders.

$$\sum_i Y_i = 0 \tag{2.42}$$

Supersymmetric partners also contribute to the weak hyper-charge and therefore the Higgsinos with $Y = 1$ violates the weak hyper-charge conservation. To solve that problem its necessary to introduce two complex Higgs fields:

$$\Phi_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix}, \ \Phi_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix}$$

This guarantees the conservation of the weak hyper-charge: Φ_2 has a weak hyper-charge of $Y = 1$ and Φ_1 has a weak hyper-charge of $Y = -1$. Latest experimental results suggest that the observed Higgs boson is a SM particle and that there is no second Higgs field [31].



(a) **(b)**

Figure 2.12 – Figure 2.12a: Various cross-sections according the NIST/Xcom data base. The single cross-sections reefer to coherent scattering (green dashed), incoherent scattering (blue points), photo-electric absorption (pink line), pair production in nuclear field (light-blue) and pair production in electron field (light-blue dashed line) [90]. The photo-electric absorption dominates for photon energies $E_\gamma < 1 MeV$. Figure 2.12b: Feynman diagram of an axion which converts by a photon of the electric/magnetic field into a photon.

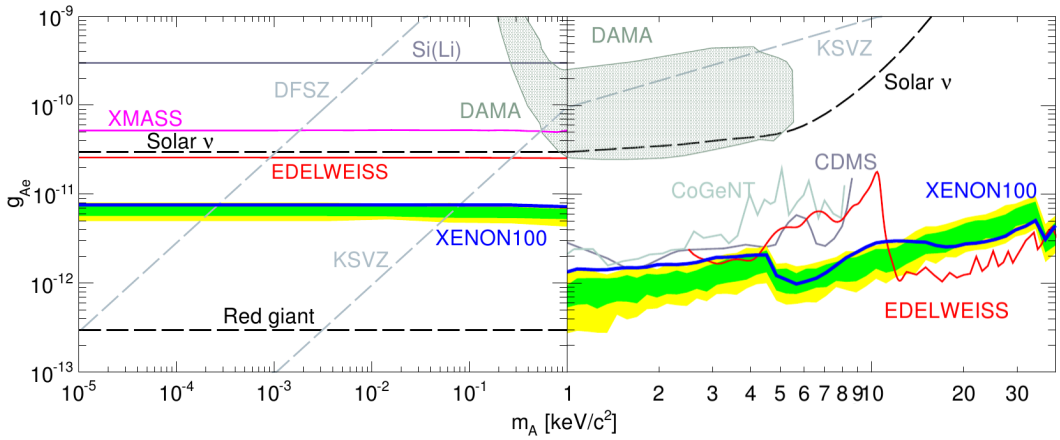


Figure 2.13 – The coupling constant g_{ae} for axions to electrons for (left) solar and (right) galactic axions [18]. For solar axions XENON100 rules out other experiments such as EDELWEISS-II [26], XMASS [106] and Derbin et al. [48]. The DAMA/LIBRA result present the modulation of an axion signal in the detector [10]. In addition the benchmarks of DFSZ and KSVZ models are drawn [49, 71, 81, 117]. On the right site is the limit for galactic axions. The XENON100 undergo the CoGeNT [2], CDMS [7] and EDELWEISS-II [26] results. The green and yellow band gives the 1σ, 2σ sensitivity.

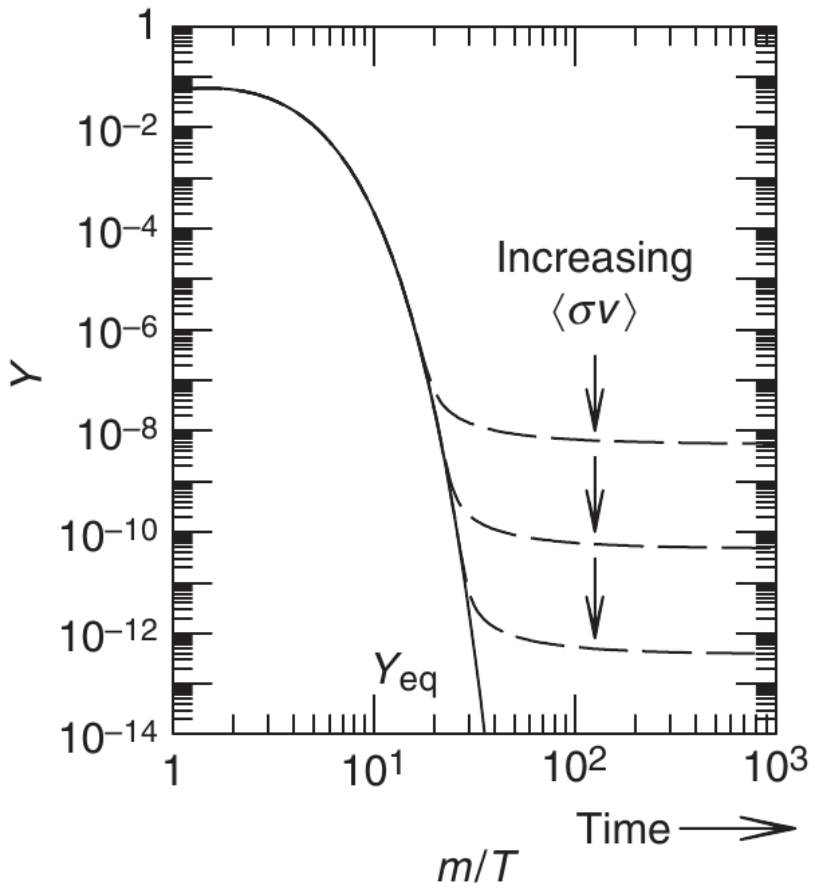


Figure 2.14 – Numerical solution of equation (2.28). The number density Y decreases rapidly. Depending on the increasing annihilation rate there are possible freeze-out values for given temperatures T_{fo}. After the expanding universe cooled down to temperatures $T \ll T_{fo}$ the WIMP number density decrease flattens. These WIMPs are the nowadays assumed relic WIMP background.

2.4. Introduction to (Direct) Dark Matter Search

The previous two subsections explain the necessity of Dark Matter on cosmological scales. Furthermore some favoured particle Dark Matter candidates are discussed. To prove the existence of one or more Dark Matter candidates, it is essential to build sensitive experiments. Three different basic experimental search scenarios are possible:

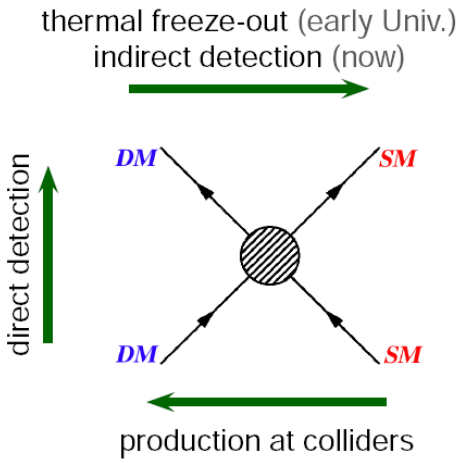


Figure 2.16 – Summary of the basic interaction types for Dark Matter search: *Direct*, *indirect* and *production*. Feynman graphs behind the interactions are much more complicated and depend on the model. [70].

Direct Dark Matter search investigates the interaction of theoretical Dark Matter candidates (e.g. WIMPs) with the nuclei/electron of a target atom. The resulting signal comes from the measured nuclear recoil and depends on the kind of detector. The physical processes are scintillation, ionisation and phonon/heat production. Such experiments need a clean and ultra-low radioactive environment to operate and therefore they are hosted in underground facilities such as the Laboratory Nazionali del Gran Sasso (LNGS) [60] or the Laboratoire Souterrain de Modane [84]. Famous experiments which test the existence of WIMPs are the *XENON Dark Matter Project*, *LUX*, *CRESST* or *DAMA/LIBRA*.

Indirect Dark Matter search takes the possibility into account that two Dark Matter particles annihilate and their decay products are baryonic particles which are detected. The observation probability of such interaction products rises with higher masses, where Dark Matter particles are collected and bound by gravity (e.g. sun). Therefore earthbound detectors look at the sun as a source of the typical decay products. Instruments such as *IceCube* and *ANTARES* (neutrino astronomy), or *HESS* (gamma ray astronomy) are not build to detect Dark Matter only and so indirect Dark Matter search is a *side* product of the instrumentation facility.

Dark Matter Production is another possibility to investigate the nature of Dark Matter. Large collider facilities produce the Dark Matter particles in a reverse process compared to the indirect Dark Matter search. The missing energy in such processes are interpreted as the Dark Matter particles. The cross-section of Dark Matter particles is to low to build suitable targets for an efficient detection. The reconstruction of the missing energy is the only possibility to identify such particles. The biggest disadvantage of that method is the uncertainty about the produced particle after the collision. It is not

possible to measure important parameters such as the lifetime to verify the details. For example, the Large Hadron Collider (LHC) would be able to produce such particles in proton–anti-proton collisions at high energies. The detector (e.g. ATLAS) measures the energies which are deposited and reconstruct the particle type based on theoretical inputs (baryonic matter and/or Dark Matter). Hence the prediction of particle Dark Matter is only as good as the theoretical input.

In summary, even if one experiment claims a discovery, its results must fit to cosmological observations and the outcomes of other experiments. However, these three different methods are complementary to each other and will reveal the mystery of Dark Matter in the future. Figure 2.16 shows the basic interaction methods from the view of the search scenario. Related Feynman graphs depend strongly on the underlying model and are much more complicated if Dark Matter only shows up in higher loop corrections to standard processes. Figure 2.17 shows a non-exhaustive list of direct Dark Matter experiments and the signals they use. Besides different experiments which use other read-out techniques, such as $CaWO_4$ crystals (CREEST), there is also a list of upcoming Dark Matter experiments such as XENON1T (successor experiment of XENON100), LZ (successor experiment of LUX) or the DARWIN experiment, which is in research and development at the moment. The following section discusses the direct Dark Matter search method such it is done at the XENON Dark Matter Project with the XENON100 experiment for WIMP search. The basic requirements for WIMP Dark

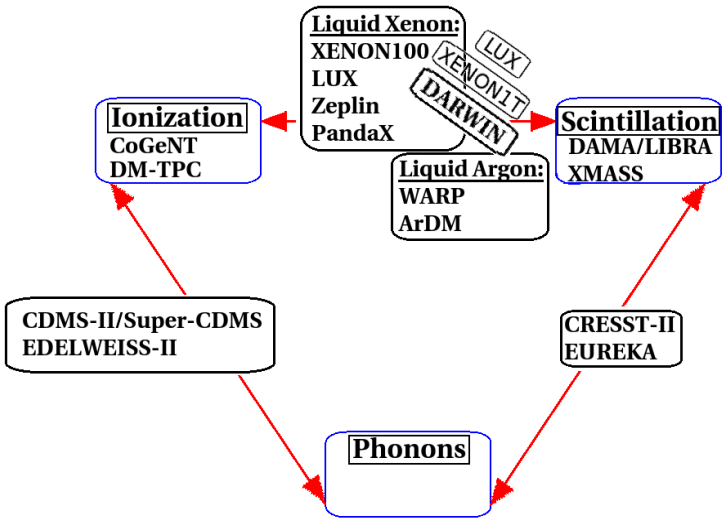


Figure 2.17 – Overview about the experiments which search for particle Dark Matter depending on the signal creation. Besides some exceptions such as CoGeNT where one signal type is read out, most experiments use a combination of ionisation, scintillation and phonon/heat. There is no experiment which exploit three signal creation methods at the same time. Further detection methods are bubble formation chambers or tracking which is independent.

Matter are already discussed in section 2.3.2.1. Cosmological observations suggest a Cold Dark Matter scenario in which WIMPs are much slower than the speed of light. The deposited energy of such a WIMP interaction depends on many parameters, but its most important input comes from assumed Halo model. Here a typical WIMP density distribution is assumed for the Milky Way. Once the cosmological input parameters are fixed, the energy deposition in the instruments and possible particle interaction rates and type can be determined.

2.4.1. Dark Matter Halo Model

Direct Dark Matter detection experiments need information about the Dark Matter density profile in the Milky Way. As already discussed in section 2.2.1 the rotation curves of spiral galaxies do not follow their theoretical predictions. Advanced methods to measure rotation curves in the own galaxy are discussed in further literature [104]. The Milky Way is divided into three main parts: An inner (spherical) bulge with a disk around. The disk contains the five spiral arms in case of the Milky Way. An additional halo of Dark Matter surrounds the bulge and the disk is much bigger in diameter. The sun circles around the centre of the galaxy with a distance of 8.5 kpc. Actual measurements suggest a disk mass of $m_{disk} = 0.95 \times 10^{11}\,\mathrm{M}_{\odot}$ and a bulge mass of $m_{\mathrm{bulge}} = 0.91 \times 10^{10}\,\mathrm{M}_{\odot}$. The Dark Matter halo has a density of $\rho_{\odot}^{DM} = 0.0088\,\mathrm{M}_{\odot}/\mathrm{pc}^{-3}$ ($= 0.35\,\mathrm{GeV}/\mathrm{cm}^{-3}$). Figure 2.18 sketches the three components. Figure 2.19b shows their radial WIMP velocity distribution separated into its components [69]. In case the Dark Matter halo consists of WIMPs, the sun and hence the earth fly through the WIMP particles (see figure 2.19a). Therefore all necessary parameters for the later WIMP detection needs to be calculated for the earth position. As result of the circulation of the sun, an observer on earth observes a WIMP wind which depends on the position of the observer on earth. The upcoming WIMP flux is in the order of $10^5 (GeV/m_{\chi})\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This flux is high enough to observe WIMP interactions at earthbound experiments even if the cross-section is very low [32].

The WIMP velocity distribution becomes important for the later detection. Figure 2.19b shows that the Dark Matter halo density contributes with $\rho(r) \propto r^{-2}$. Therefore a simple approximation for the velocity distribution is[79]:

$$f(\vec{v}) \propto \exp\left(-\frac{|\vec{v}|^2}{v_0^2}\right) \tag{2.43}$$

The picture becomes more complicated by looking at figure 2.19a. The WIMP flux depends on the detector position on earth and because of the annual cycle around the sun the WIMP velocity becomes relative ($v' \rightarrow v + v_{\odot}^{\pm}(t)$). Therefore equation (2.43) is adjusted by $v_{\odot}^{\pm}$. The modification is shown in equation (2.44). The Dark Matter velocity dispersion which is relative to the cycle speed of the sun

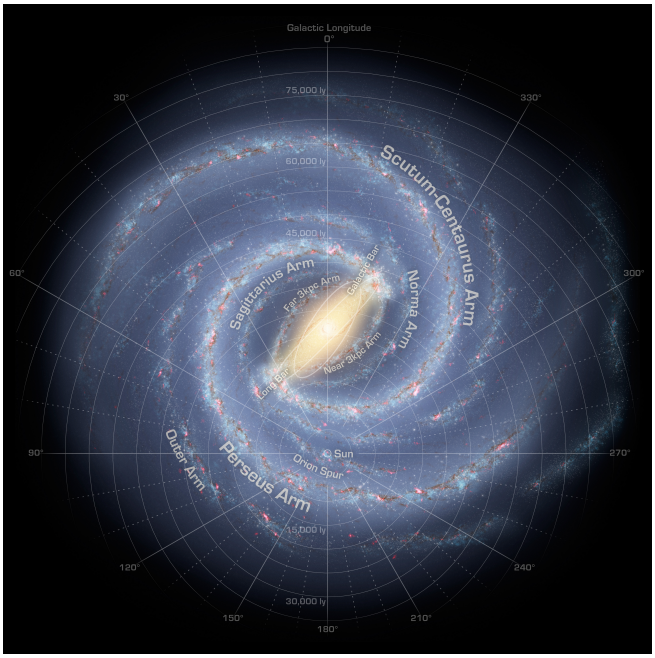
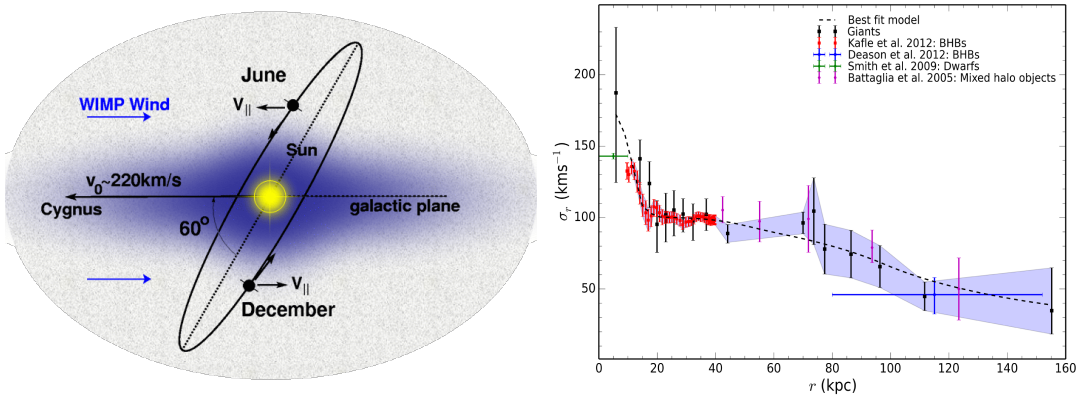


Figure 2.18 – A sketch of the Milky Way consists of a spherical bulge and disk. The disk contains the five spiral arms. The earth is in one of spiral arms at a distance 8.5 kpc from the centre of the galaxy. A much bigger Dark Matter halo surrounds bulge and disk [110] which is also sketched in figure 2.19a.



(a) (b)

Figure 2.19 – Figure 2.19a: Due to the movement of the sun around the centre of the galaxy and the earth movement around the sun, the detected WIMP wind changes with seasons. Figure 2.19b: Actual measurements picture the circular velocity of the galaxy where the distance R is measured from the centre of the galaxy. The curves come from the best fit estimations of the model parameters. See: [69]

around the centre of the galaxy is $v_0 = 220\,\mathrm{km/s}$. The time dependence t is related to the season on earth.

$$f(v, \vec{v}_{\odot}^{\pm}) \propto \exp\left(-\frac{|\vec{v} + \vec{v}_{\odot}^{\pm}(t)|^2}{v_0^2}\right) \tag{2.44}$$

In detail the earth velocity around the sun is given by equation (2.45) where ε_1 and ε_2 are related to the elliptical movement of the earth around the sun and t_1/t_2 are related to the spring equinox (March, 21^{th}) [32].

$$v_{\odot}^{\pm}(t) = v_{\odot}^{\pm}\left(\varepsilon_1 \cos\omega(t - t_1) + \varepsilon_2 \sin\omega(t - t_2)\right) \tag{2.45}$$

2.4.2. Particle Dark Matter Interaction

WIMPs from the Dark Matter halo are supposed to interact with the detector's target material. Every interaction is based on an energy deposit which should yield a detectable signal. The interaction leads to a recoil of the target nucleus (recoil energy E_R). The WIMP mass is an unknown parameter, and different possible WIMP masses cause different recoil energy spectra. The resulting rate is given in events per kg and day ($events\,kg^{-1}\,day^{-1}\,keV^{-1} \equiv dru$). The following section reviews the WIMP-nucleus interaction and formalism of [79]. The differential WIMP rate with a mass m_χ and target nucleus mass m_N is:

$$\frac{dR}{dE_R} = \frac{\rho_0}{m_N\, m_\chi} \int_{v_{min}}^{\infty} v\, f(v)\, \frac{\sigma_{WN}}{dE_R}(v, E_R)\, dv \tag{2.46}$$

With the WIMP density ρ_0 and the differential cross-section for elastic scatter σ_{WN}. The kinetic energy of the WIMPs is taken into account by $f(v, v_{\odot}^{\pm})$ from equation (2.44). The integration boundaries in equation (2.46) are formally driven by the v_{min} which refers to the smallest detectable energy. The equivalent energy E_{min} acts as a detection threshold. The upper boundary is given by ∞ what is right under the assumption that WIMPs are allowed to leave the Dark Matter halo of the Milky Way $v_{max} \gg v_{esc}$. This assumption does not hold under the circumstances that we observe Dark Matter halos nowadays and therefore ∞ acts as an approximation.

Based on equation (2.46) the differential rate dR/dE_R depends on the most probable WIMP energy E_0 for its mass of m_χ:

$$\frac{dR}{dE_R} = \frac{R_0}{E_0} \frac{(m_\chi + m_t)^2}{4(m_\chi\, m_t)} \times e^{-\frac{E_R}{E_0} r} = \frac{R_0}{E_0\, \mu} \times e^{-\frac{E_R}{E_0} r} \tag{2.47}$$

The kinetic factor μ is defined as follows:

$$\mu = 4 \frac{m_\chi\, m_t}{(m_\chi + m_t)^2} \tag{2.48}$$

This results an event rate R per unit mass and total event rate R_0. Equation (2.47) separates the experimental approach to measure a differential rate which depends on the setup (left-side) and total expected Dark Matter signal for different WIMP masses (right-side). The experimental setup of the XENON100 Dark Matter detector is discussed in chapter 3.
Due to the target material and the experimental setup equation (2.47) is adjusted to the observed differential rate in the detector:

$$\left.\frac{dR}{dE_R}\right|_{observed} = R_0\, S(E)\, F^2(E)\, I \tag{2.49}$$

From the left site of equation (2.47), equation (2.49) is developed and contains the total rate R_0, a modified spectral function S, form factor corrections $F(E)$ and an interaction function I. These parameters are described independently from each other and aim to separate physics and experimental setup.

The modified spectral function S describes the detector site:

1. Position relative to the centre of the galaxy
2. Detection efficiency's
3. Detailed target information. (For example the number of targets cells within an experiment varies)
4. Instrumental resolution and thresholds

The form factor correction takes the size of the nucleus into account. The spin interactions (spin-independent – SI and spin-dependent – SD) are described by I. The final outcome is the recoil energy spectrum in pe_{ee} (photo-electron recoil energy equivalent).

The right side of equations (2.47) and (2.49) contains the rate which depends on the WIMP velocity distribution and particle number density n. For the typical DM halo structure, its differential particle density dn takes a Maxwellian velocity distribution into account (see equation (2.44)):

$$dn = \frac{n_0}{k} \exp\left[-\left(\frac{|v + v_{\odot}^{\pm}|}{v_0} \right)^2 \right] d^3v \tag{2.50}$$

With the movement of the sun around the centre of the galaxy $v_0 = 230\,km\,s^{-1}$. To be consistence with the CDM hypothesis the maximum velocity is set to the escape velocity v_{esc}. The pre-factor consist of mean particle Dark Matter number density n_0 which takes the Dark Matter density ρ_χ and mass m_χ into account (see equation (2.51)).

$$n_0 = \frac{\rho_\chi}{m_\chi} \tag{2.51}$$

A normalisation factor k is chosen that the following condition is fulfilled:

$$\int_0^{v_{esc}} dn \equiv n_0 \tag{2.52}$$

As long as WIMPs move in every direction and every velocity the normalisation k is chosen for example according:

$$k = \int_0^{2\pi} d\Phi \int_0^{\pi} \sin(\Theta) d\Theta \int_0^{v_{esc}} f(\vec{v}, \vec{v}_{\odot}^{\pm})\, v^2 dv \tag{2.53}$$

The velocity distribution $f(\vec{v}, \vec{v}_{\odot}^{\pm})$ is taken from equation (2.45) and hence there are two outcomes which depend on the integration boundaries:

$$k_0 = \left(\pi v_0^2\right)^{3/2} \qquad \text{for } v_{esc} = \infty \tag{2.54}$$

$$k_1 = k_0 \left[\text{erf}\left(\frac{v_{esc}}{v_0}\right) - \frac{2}{\sqrt{\pi}} \frac{v_{esc}}{v_0} e^{-\frac{v_{esc}^2}{v_0^2}} \right] \qquad \text{for } |\vec{v} + \vec{v}_{\odot}| = v_{esc} \tag{2.55}$$

Equation (2.54) describes the situation that the expected WIMP velocities can take every possible value even if their velocities become bigger than the escape velocity. Then equation (2.55) limits the WIMP velocity spectrum to the escape velocity and put the relative earth velocity also into account. Both normalisations differ less than 0.5% from each other.

The next step is to connect the differential particle density dn with the differential event rate dR. This takes atomic mass number A and cross-section σ and Avogardos constant N_0 into account:

$$dR = \frac{N_0}{A} \sigma v dn \tag{2.56}$$

Equation (2.56) follows that a differential rate-particle density becomes constant and is determined by a specific velocity and cross-section σ_0 (see equation (2.57)).

$$\frac{dR}{dn} = \frac{N_0}{A} \sigma_0 v \tag{2.57}$$

Present equation (2.56) without a differential leads to:

$$R = \frac{N_0}{A} \sigma_0 \int v dn \equiv \frac{N_0}{A} \sigma_0 n_0 \langle v \rangle = \frac{N_0}{A} \sigma_0 \frac{\rho_\chi}{m_\chi} \langle v \rangle \tag{2.58}$$

By defining R_0 per unit mass as

$$R_0 = \frac{2}{\sqrt{\pi}} \frac{N_0}{A} \frac{\rho_\chi}{m_\chi} \sigma_0 v_0 \tag{2.59}$$

leads to a substitution of N_0 in equation (2.57) and the rate become

$$R = R_0 \frac{\sqrt{\pi}}{2} \frac{\langle v \rangle}{v_0} \tag{2.60}$$

and hence by using the left site of equation (2.58) and use the differential particle density of equation (2.50) leads finally (with R_0 from equation (2.59)) to a velocity dependent rate distribution:

$$R = R_0 \frac{k_0}{k} \frac{1}{2\pi v_0^4} \int v f\left(\vec{v}, \vec{v}_{\odot}\right) d^3 v \tag{2.61}$$

As already discussed are the boundaries of the integral important for the rate determination and hence the possible solutions for three rate calculations is given by:

$$\frac{R(0,v_{esc})}{R_0} = \frac{k_0}{k_1}\left[1-\left(1+\frac{v_{esc}^2}{v_0^2}\right)e^{-\frac{v_{esc}^2}{v_0^2}}\right] \tag{2.62}$$

$$\frac{R(v_{\odot}^{\pm},\infty)}{R_0} = \frac{1}{2}\left[\sqrt{\pi}\left(\frac{v_{\odot}^{\pm}}{v_0}+\frac{1}{2}\frac{v_0}{v_{\odot}^{\pm}}\right)\mathrm{erf}\left(\frac{v_{\odot}^{\pm}}{v_0}\right)+e^{-\frac{v_{\odot}^{\pm 2}}{v_0^2}}\right] \tag{2.63}$$

$$\frac{R(v_{\odot}^{\pm},v_{esc})}{R_0} = \frac{k_0}{k_1}\left[\frac{R(v_{\odot}^{\pm},\infty)}{R_0}-\left(\frac{v_{esc}^2}{v_0^2}+\frac{1}{3}\frac{v_{\odot}^{\pm 2}}{v_0^2}+1\right)e^{-\frac{v_{esc}^2}{v_0^2}}\right] \tag{2.64}$$

Therefore equation (2.63) is the most important ones because it takes the relative WIMP velocity into account relative to earth velocity $\vec{v}_{\odot}^{\pm}$. The according total rate R_0 is calculated out of equation (2.59). The recoil energy of a Dark Matter particle depends on the scatter angle Θ. A WIMP with m_χ lead to a kinetic energy of $E_\chi = 0.5 m_\chi v_\chi^2$ and the therefore the scatter dependent recoil energy is:

$$E_R = \frac{E_\chi \mu}{2}(1-\cos(\Theta)) \text{ with } 0 \leqq E_R \leqq E_\chi \mu \tag{2.65}$$

The recoil energy range of equation (2.65) is limited in reality to a minimum energy $E_{min} = E_R/\mu$. To introduce the energy dependence ($v \to E$) the velocity dependent rate of equation (2.61) is written in differential form:

$$dR = R_0 \frac{k_0}{k}\frac{1}{2\pi v_0^4} v f\left(\vec{v},\vec{v}_{\odot}^{\pm}\right) d^3v \tag{2.66}$$

Hence the differential energy dependent rate is written as (with $E_0 = 0.5 m_\chi v_0^2 = (v_0^2/v_\chi^2)E_\chi$):

$$\frac{dR}{dE_R} = \int_{E_{min}}^{E_{max}} \frac{1}{E\mu} dR(E) = \frac{1}{E_0\mu}\int_{v_{min}}^{v_{max}} \frac{v_0^2}{v^2} dR(v) \tag{2.67}$$

The substitution of E_{min} with the corresponding WIMP velocity is: $v_{min} = \sqrt{2E_{min}/m_\chi} = \sqrt{E_R/(E_0\mu)}v_0$ leads to:

$$\frac{dR}{dE_R} = \frac{R_0}{E_0\mu}\frac{k_0}{k}\frac{1}{2\pi v_0^2}\int_{v_{min}}^{v_{max}} \frac{1}{v} f(\vec{v},\vec{v}_{\odot}^{\pm}) d^3v \tag{2.68}$$

The solution of equation (2.68) is again related to the boundaries of the integration and therefore the already shown interactions of the total rates (see equation (2.62), (2.63) and (2.64)) are given here with their differential rates:

$$\frac{dR(0,v_{esc})}{dE_R} = \frac{k_0}{k_1}\frac{R_0}{E_0\mu}\left(e^{-\frac{E_R}{E_0\mu}}-e^{-\frac{v_{esc}^2}{v_0^2}}\right) = \frac{k_0}{k_1}\left[\frac{dR(0,\infty)}{E_R}-\frac{R_0}{E_0\mu}e^{-\frac{v_{esc}^2}{v_0^2}}\right] \tag{2.69}$$

$$\frac{dR(v_{\odot}^{\pm},\infty)}{dE_R} = \frac{R_0}{E_0\mu}\frac{\sqrt{\pi}}{4}\frac{v_0}{v_{\odot}^{\pm}}\left[\mathrm{erf}\left(\frac{v_{min}+v_{\odot}^{\pm}}{v_0}\right) - \mathrm{erf}\left(\frac{v_{min}-v_{\odot}^{\pm}}{v_0}\right)\right] \tag{2.70}$$

$$\frac{dR(v_{\odot}^{\pm},v_{esc})}{dE_R} = \frac{k_0}{k_1}\left[\frac{R(v_{\odot}^{\pm},\infty)}{dE_R} - \frac{R_0}{E_0\mu}e^{-\frac{v_{esc}^2}{v_0^2}}\right] \approx c_1\frac{R_0}{E_0\mu}e^{-c_2\frac{E_R}{E_0\mu}} \tag{2.71}$$

Equation (2.69) the differential expression is defined by

$$\frac{dR(0,\infty)}{E_R} = \frac{R_0}{E_0\mu}e^{-\frac{E_R}{E_0\mu}} \tag{2.72}$$

and takes an unmodified nuclear recoil spectrum into account for earth velocity at $v_{\odot}^{\pm} = 0$. The approximation in equation (2.71) simplifies the calculation and introduces fitting constants c_1 and c_2 which vary over the year.

2.4.2.1. Nuclear Form Factor

After the differential nuclear recoil spectrum is calculated, further modifications are necessary. Equation (2.49) already shows the required quantities. When the De-Broglie wavelength $\lambda = h/m_{\chi}$ is in the order of the size of the nucleon, the effective cross-section decreases with higher momentum transfer $q = \sqrt{2m_T E_R}$. The cross-section is adjusted to according equation (2.73), where n_r is the effective nuclear radius.

$$\sigma(qr_n) = \sigma_0 F^2(qn_r) \tag{2.73}$$

From textbook the nuclear form factor is calculated by the Fourier transformation of the nucleon density distribution:

$$F(q) = \int_V \rho(r)e^{i\vec{q}\vec{r}}d^3x = \frac{4\pi}{q}\int_0^{\infty} r\sin(qr)\rho(r)dr \tag{2.74}$$

In general the density ρ of the nuclei is described by a Fermi distribution:

$$\rho(r) = \rho_0\left[1+\exp\left(\frac{r-c}{a}\right)\right]^{-1} \tag{2.75}$$

This leads to a numerical solution which needs more computational power. To simplify the problem, Helm [65] suggested the form factor in equation (2.76) which has the advantage of an analytic solution. In addition, a thickness parameter s is introduced which describes a shell around the nucleus.

$$F(qr_n) = 3\frac{j_1(qr_n)}{qr_n}e^{-0.5(qs)^2} \tag{2.76}$$

The Bessel function j_l (here: first order l = 1, see equation (2.77)) describes the decreasing form factor for higher momentum transfer q.

$$j_1(qr_n) = \frac{\sin(qr_n)}{qr_n} - \frac{\cos(qr_n)}{qr_n} \tag{2.77}$$

The parameters r_n and s are connected by [38]:

$$r_n = \sqrt{r_v^2 - 5s^2} \quad \text{with } r_v = 1.2A_T^{1/3} \tag{2.78}$$

The final form factors $F(q)$ for different materials which are used in Dark Matter search are shown in figure 2.20. Here the form factors are given in dependence of the recoil energy E_R.

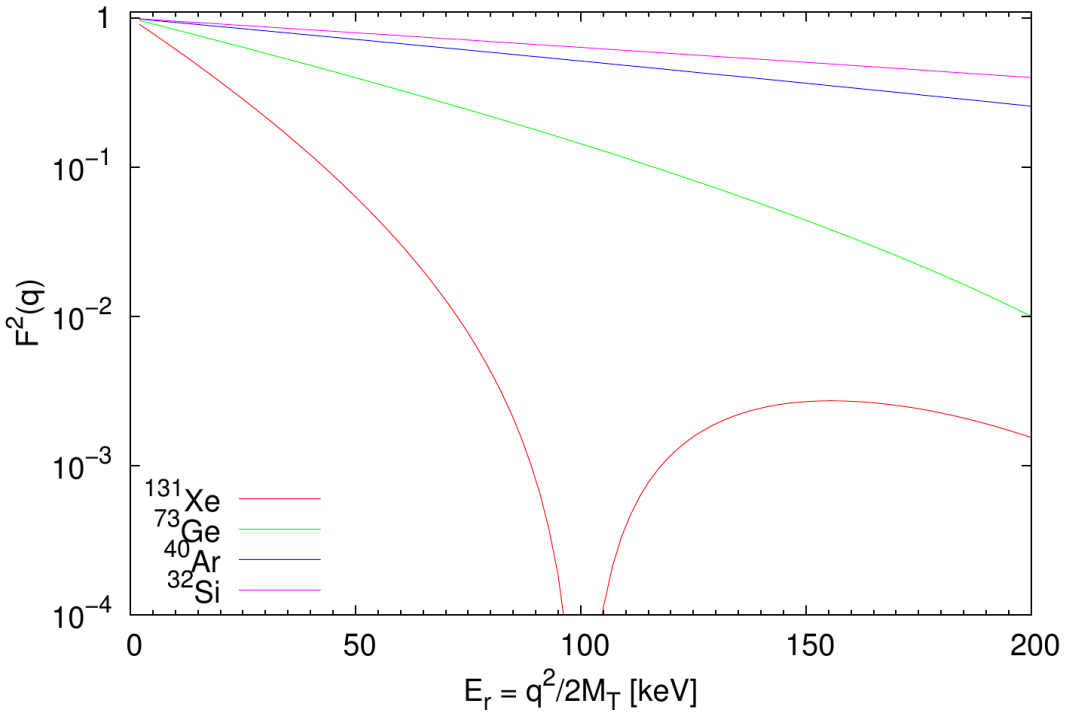


Figure 2.20 – Different nuclear form factors which depend on the recoil energy [82]. The presented nuclei are used in Dark Matter search experiments nowadays. The kind of nucleus/form factor is important for the WIMP scattering process and hence, the observed rate.

2.4.2.2. Different Targets and Spin

The total rate of equation (2.59) is derived without taking different cross-sections σ_0 (for different target materials) into account. At zero momentum transfer the cross-section is proportional to the atomic mass number A_T. This is true for coherent scattering (spin-independent). Figure 2.17 already showed that different experiments use different detection techniques. Therefore the momentum transfer is always measured for a given target. To standardise the procedure the transfer q is normalised to the single proton mass (m_p) by the following factor:

$$\left(\frac{\mu_T}{\mu_p}\right)^2 = \left(\frac{\frac{m_\chi m_T}{m_\chi + m_T}}{\frac{m_\chi m_p}{m_\chi + m_p}}\right)^2 \tag{2.79}$$

and therefore the adjusted and observed recoil energy is:

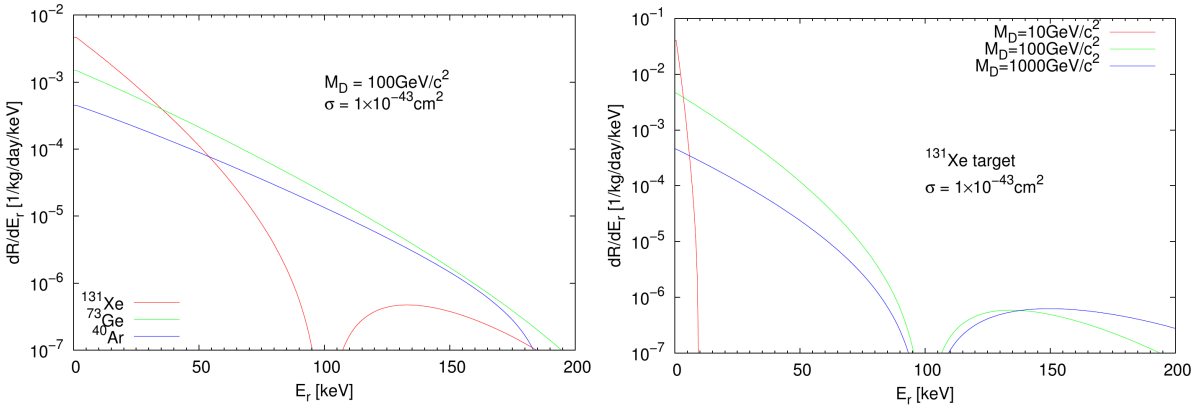
$$\left.\frac{dR}{dE_R}\right|_{observed} = \frac{dR(v_\odot^+, v_{esc})}{dE_R} \sigma_p \left(A_T \frac{\mu_T}{\mu_p}\right)^2 F^2(E_R) \tag{2.80}$$

where σ_p is a normalised single WIMP-proton cross-section. The outcome is shown figure 2.21a where the differential rate is plotted for a WIMP with $m_\chi = 100\,\text{GeV}$ and cross-section $\sigma_\chi = 1 \times 10^{-43}\,\text{cm}^2$ and for different target materials which are used in Dark Matter search experiments: xenon, germanium, argon. On the other hand, figure 2.21b show the result for xenon only and compares three different WIMP masses: $m_\chi = 10\,\text{GeV}, 100\,\text{GeV}, 1000\,\text{GeV}$ at the same cross-section.

Spin-independent (SI) interaction leads to a higher probability for WIMP-nucleus interaction because neutrons and protons can be considered as equal. The situation becomes more complicated when spin-dependent (SD) interaction is assumed: Protons and neutrons, each with their spin s = ½ are not considered as equal any more. A target nucleus (e.g. ^{131}Xe) sums up its single spins pairwise up to $s_{total} = 1/2$. Finally, the single scatter amplitudes on protons and neutrons count zero and the real scattering happens on the unpaired neutrons and protons. This is only possible for odd isotopes. In case of xenon there are two isotopes which allow testing of the spin-dependent scattering. Isotopes ^{129}Xe and ^{131}Xe with abundances of 26.4% and 21.2%, respectively, have a non-zero spin [20].

2.4.2.3. Detection Efficiency's and Limits

With the WIMP rate calculation, the detector set limits on the observed recoil energies. Every instrument has only a finite energy resolution and the observed signal differs from the true signal. The reason for the differences are from the steps of signal creation in the target up to read-out procedure.



(a) (b)

Figure 2.21 – Figure 2.21a: The differential rate for different target materials for WIMPs with $m_\chi = 100\,\text{GeV}$ and $\sigma_\chi = 1 \times 10^{-43}\,\text{cm}^2$. A xenon target is favoured to detect low mass WIMPs which deposit a smaller recoil energy in the detector. Figure 2.21b: A detailed look on a ^{131}Xe target – the differential rate rises for small WIMP masses (under the assumption of a fixed cross-section of $\sigma_\chi = 1 \times 10^{-43}\,\text{cm}^2$) which results in a higher probability for low-mass WIMP detection.

As long as there are several hundreds signals per hour which are counted, it is possible to do a statistical data analysis. The underlying idea is a "counting experiment". Each step between signal creation and read-out underlies its physical properties which are taken into account in the signal interpretation. A result of this analysis are energy dependent detection efficiency's which reduces the expected rate particularly for low recoil energies. To assure a meaningful data interpretation, the efficiency's exceed a given level to avoid that signals are lost in the background. Therefore a very careful background analysis is crucial.

The energy resolution becomes important at the point when measured results must be interpreted. The finite energy resolution ΔE, expressed as the full width of half maximum (FWHM), is not fine enough if the spectrum of a radioactive source is measured and there is no conclusion on the deposited energy in the detector. Fano presented that the absorbed energy, from the ionisation process, in xenon is not Poisson-like distributed [57]. The finite energy resolution in liquid xenon in the optical approximation is given as:

$$\Delta E(E') \propto \sqrt{F \times W \times E'} \tag{2.81}$$

Formula 2.81 includes the Fano factor F which describes how the absorbed energy deviates from a Poisson process. The W describes the average energy which is necessary to create an electron-ion pair during the ionisation process [14].

2.5. Summary

This chapter gives a wide introduction to Dark Matter. Cosmological observations on different scales in the universe are explained in detail and combined with possible Dark Matter candidates (MACHOs) from non-luminous baryonic matter. In addition, some selected Dark Matter search methods are described, which are used on cosmological scales to observe the missing mass. Cosmological observations do not reveal the mystery of the missing mass and therefore a new particle Dark Matter type is suggested. Its properties are derived from cosmological observations: Charge-less with a mass but not interacting by the known forces (except gravity). The list of possible Dark Matter candidates is long and therefore that chapter describes the favoured candidates and their production mechanisms. The WIMP (*W*eakly *I*nteracting *M*assive *P*article) is one of the favoured candidates for Dark Matter. Finally, three Dark Matter search techniques are explained with focus on *direct* search because the XENON100 experiment uses that technique to observe WIMP-nucleus scattering on xenon atoms.

3. The XENON100 Experiment

"I understand how the [*Dark Matter*] engines work now. It came to me in a dream. The engines don't move the ship at all. The ship stays where it is and the engines move the universe around it."

Cubert Farnsworth,
Futurama, "A Clone of My Own"

3.1. Overview

This chapter introduces the XENON100 experiment. The experimental setup, in particular the two-phase time projection chamber (TPC) is described in detail. Then, an overview of the data acquisition system and the data output is given. At the end of this chapter, the data calibration procedures related to the experiment are summarized. This forms the basis of the following data analysis chapters of this thesis.

The XENON100 experiment follows the former XENON10 experiment for direct Dark Matter detection of the *XENON Dark Matter Project*. Liquefied xenon at -91° with a density of $\rho = \sim 2.85\,\mathrm{g/cm^3}$ serves as target material. The density itself varies with temperature and hence pressure. Xenon has the largest density among the noble gases. Seven of its thirteen isotopes are stable (^{126}Xe, ^{128}Xe, ^{129}Xe, ^{130}Xe, ^{131}Xe, ^{132}Xe, ^{134}Xe), two have such a long life time that their contribution can be neglected (^{124}Xe with $\tau > 4.8 \times 10^{16}$ a and ^{136}Xe with $\tau = 2.11 \times 10^{21}$ a) and four isotopes have such short life times that it is possible to establish a radio impure clean environment within a short time. The XENON Dark Matter project uses a two-phase time projection chamber for direct Dark Matter detection. All XENON Dark Matter experiments have been located at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy. Here the experiment is shielded against interactions from cosmic and atmospheric background. The underground laboratory has a depth of 3600 m water equivalent. This reduces the atmospheric muon flux by a factor of 10^6. Figure 3.1 shows a sketch of the underground laboratory.

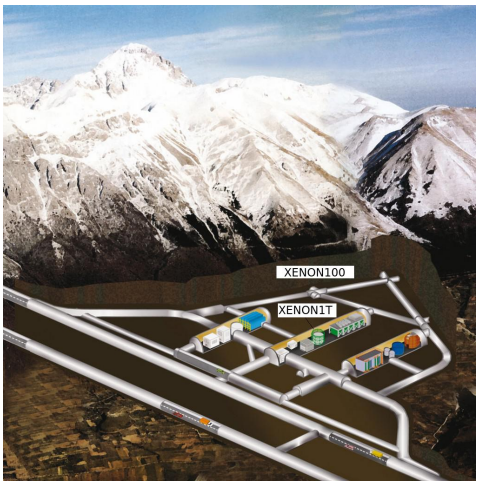


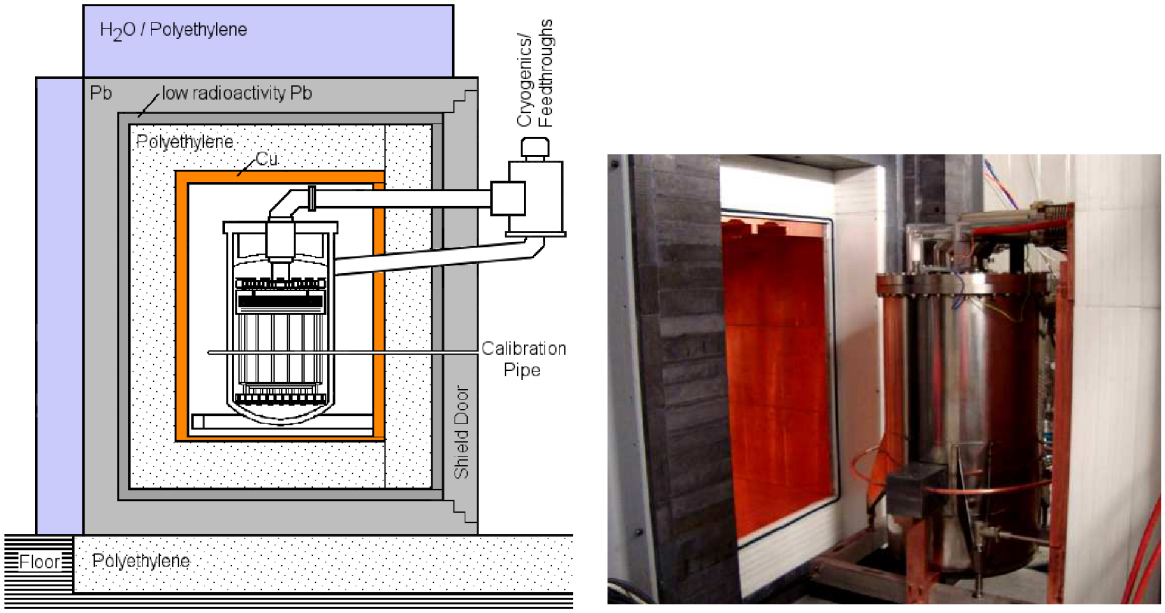
Figure 3.1 – Sketch of the three halls in the underground laboratory with the locations of XENON100 and its successor XENON1T.

3.2. The Experimental Setup

To run the two-phase time projection chamber in the XENON100 experiment and assure a clean and stable environment additional systems are necessary. This section will focus on the experimental setup and go into details about some sub-systems. The centerpiece, the two-phase time projection chamber, is explained in detail in section 3.3.

3.2.1. The XENON100 Shield

The XENON100 TPC is mounted in the inner cryostat, which is thermally insulated by vacuum from the outer cryostat. Outside there is an additional passive shield mounted to reduce the background. The shield (see figure 3.2a) consists of OFHC copper (5 cm), polyethylene (20 cm), lead (20 cm) and water/polyethylene (20 cm). The shield is closed completely and foreseen holes are used to enter calibration sources near the TPC. The shield reduces radioactive background from the cavern [115]. In addition, a picture (see figure 3.2b) of the opened shield shows the copper, polyethylene and lead shield with the mounted cryostat which contains the TPC.



(a) **(b)**

Figure 3.2.: The XENON100 shield around the vessel. Two pipes connect the cooling system with the TPC through the XENON100 shield. An additional calibration line is foreseen to position the calibration sources near to the XENON100 TPC. The different shield components are shown in figure 3.2a and a picture from the opened XENON100 experiments shows the different shielding materials: copper, polyethylene and lead [115]

3.2.2. The XENON100 Cooling System

The cooling system is needed to guarantee a stable temperature of the liquid xenon. Xenon itself is gaseous at room temperature and therefore it is necessary to control pressure and temperature for

a stable measurement environment. Figure 3.3 shows the cryogenic system in XENON 100 which is installed outside the detector shield. The pulse tube refrigerator (PTR) liquefies the xenon gas. The xenon gas is filled into the cooling system during the xenon filling process or gaseous xenon from within the TPC is collected at the cold finger below the PTR. The condensed xenon becomes liquid and is collected in a funnel which is connected to the TPC. The PTR works with a Helium compressor to remove the heat from the liquefied xenon. In case of a cooling failure, the XENON100 cooling system has an liquid nitrogen (LN_2) emergency cooler (dewer). This system assures to cool the XENON100 detector for almost 24 hours. [115]

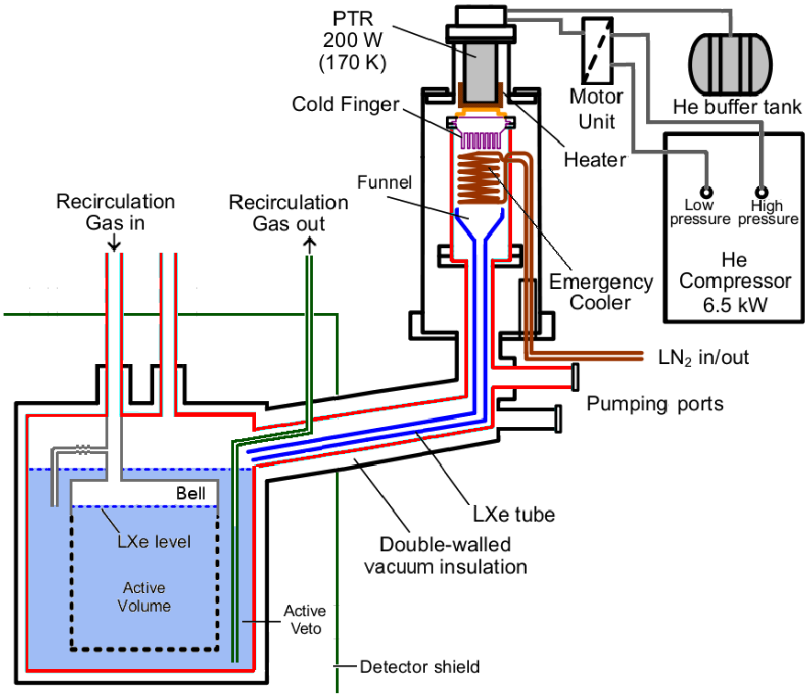


Figure 3.3.: The XENON100 cooling system is mounted outside the detector shield. A pulse tube refrigerator liquefies gaseous xenon and fill it back into the XENON100 TPC. [115]

3.2.3. The XENON100 Gas Recirculation System

Figure 3.3 shows the input/output connections for the gas recirculating system. The gas is extracted from the bottom of the detector and is filled back into the gaseous phase. A schematic draw of the gas system is shown in figure 3.4. Due to O_2 and electro-negative impurities in xenon, a high temperature zirconium getter cleans the liquid xenon continuously. The impurities are bound chemically by the

getter material. An ultra pure xenon environment is necessary in increase the electron lifetime (see section 3.6.1) and enhance the signal readout. [115]

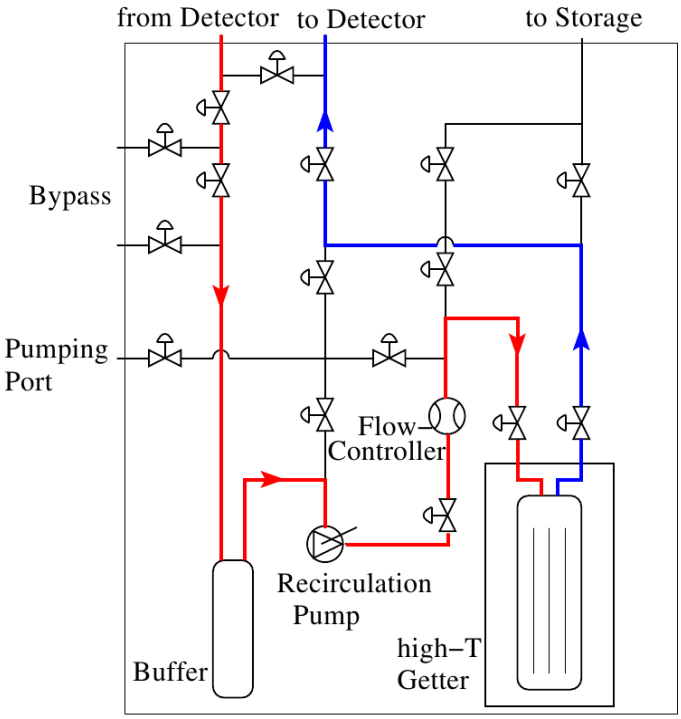


Figure 3.4.: Sketch of the XENON100 gas purification system. The liquid xenon is extracted from the TPC (see figure 3.3) and is cleaned from impurities. This process is necessary to enhance the signal readout process from the ionisation channel. [115]

3.3. The Two-Phase Time Projection Chamber

The centrepiece of the XENON100 experiment is the two-phase time projection chamber which is filled with pure xenon. The following subsections explain the operation principle of a dual phase (liquid/gas) xenon TPC with emphasis on its implementation in XENON100.

3.3.1. Operation Principle

The basic operation principle of the TPC consists of two physical principles: scintillation and ionisation. An overview is given in figure 3.5. An incoming particle (e.g. WIMP or neutron) scatters with

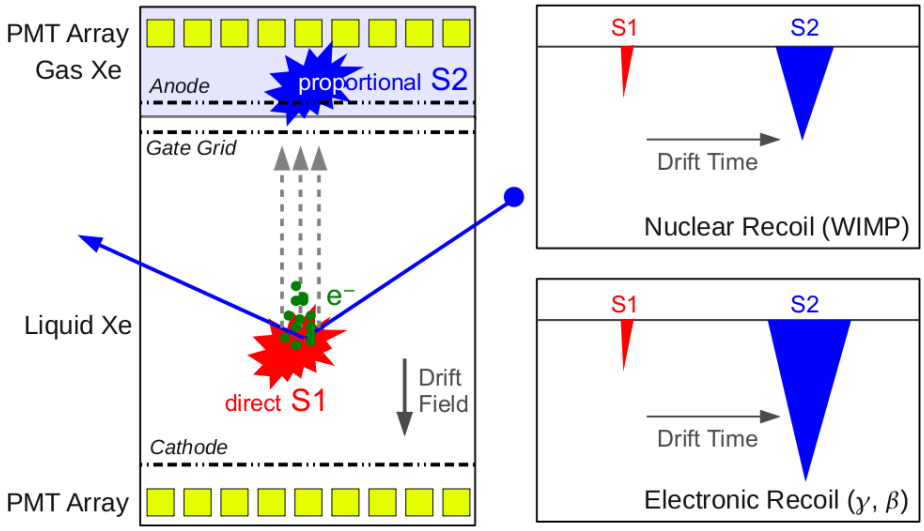


Figure 3.5 – The working principle of a TPC

a xenon nucleus inside the liquid xenon. The scattering process results in scintillation light with a typical wavelength of $\lambda = 178\,\text{nm}$ [67]. These photons can not interact a second time with xenon in the TPC because their energy of 6.97 eV is below the first ionisation energy of xenon. Hence these photons are spread over the TPC and homogeneously distributed. This kind of signal is called the primary scintillation signal S1 in the detector.
At the same time fire electrons out of the ionisation process. An applied voltage between the gate and the cathode mesh creates an electric field. The electrons drift upwards with a velocity of $1.74\,\text{mm}/\mu\text{s}$ (at $0.53\,\text{kV}/\text{cm}$ depending on the electric field strength) to the gate mesh. There, the electrons are extracted from the liquid phase into the gaseous phase by a stronger electric field between the gate and the anode mesh ($\sim 12\,\text{kV}/\text{cm}$) [115]. The electrons enter the upper gaseous part of the TPC and become accelerated by the higher electric field. The electrons gain kinetic energy, are amplified and interact again with further xenon atoms in the gaseous phase and produce scintillation light. This proportional secondary signal from ionisation is called S2. The S1 and S2 signal production is based on the energy deposition from a particle in a xenon atom.

The typical ultraviolet scintillation light in xenon results from atom excitation (Xe^*) and electron-ion (Xe^+) recombination from where a Xe_2^* state results. This state decays according to equation (3.1):

$$Xe_2^* \rightarrow 2Xe + h\nu \tag{3.1}$$

The xenon excitation results from equation (3.2) where a particle P deposit a certain amount of energy in the xenon atom. Surrounding xenon atoms are bound and create the Xe_2^* state which decays according to equation (3.1). The scintillation light from the electron-ion recombination needs an ionised xenon atom Xe^+ (equation (3.4)) which interact with the surrounding xenon to Xe_2^+ (equation (3.5)). An electron triggers the heat production and forms an excited xenon atom Xe^* (equation (3.6)). The Xe^* undergoes the same process such in equation (3.3) and then (3.1) and scintillation light is created. Primary scintillation is produced by recombination and excitation. Excited xenon atoms from excimer eventually yield the scintillation photons of energy below the first ionisation potential of xenon (see equation (3.2), (3.3) and (3.1)).

$$Xe + P \rightarrow Xe^* + P' \tag{3.2}$$

$$Xe^* + Xe \rightarrow Xe_2^* \tag{3.3}$$

$$Xe + E \rightarrow Xe^+ + e^- \tag{3.4}$$

$$Xe^+ + Xe \rightarrow Xe_2^+ \tag{3.5}$$

$$Xe_2^+ + e^- \rightarrow ... \rightarrow Xe^* + \text{heat loss} \tag{3.6}$$

S1 and S2 signals are detected by arrays of photo-multiplier tubes (PMTs, see section 3.3.3). S1 and S2 signals are usually easy to distinguish from each other, that is used to determine the signal type. The idea to distinguish between S1 and S2 signals is shown in figure 3.5 on the right site. The ratio between the S1 and S2 is different whether the signal originates from a electronic recoil (gamma induced) or a nuclear recoil (induced by a WIMP or neutron). An electronic recoil origin from the interaction of gammas with the electron cloud what frees more electrons in comparison to a neutron interaction with the nucleus itself. As a consequence, more electrons result a larger S2 signal. [14]

A key feature of the two-phase time projection chamber is the 3D event reconstruction. The x/y position is determined from the top photo-multiplier array from the S2 signal. From the time difference Δt of the S1 and S2 signal, the interaction depth is determined. Figure 3.6 shows a recorded waveform example, here the S1 and S2 peaks are distinguishable (see figure 3.7 for a detailed view on the S1 and S2 peaks) and figures 3.8 and 3.9 show the according PMT pattern for an S1 and S2 signal. The S1 – S2 time difference (drift time Δt) and the electron drift velocity is used to calculate the interaction depth. Together with the S2 signal information, it is enough to determine the position with corresponding position reconstruction algorithms (see section 3.5.3). The 3D event reconstruction allows to define a fiducial volume in the inner TPC. This greatly reduces background from the edges and surroundings of the TPC from natural radiation and helps to identify double scatter events. Last latter criterion is useful for the WIMP search since it is expected that low WIMP-nucleon interaction lead to single scatter in the TPC.

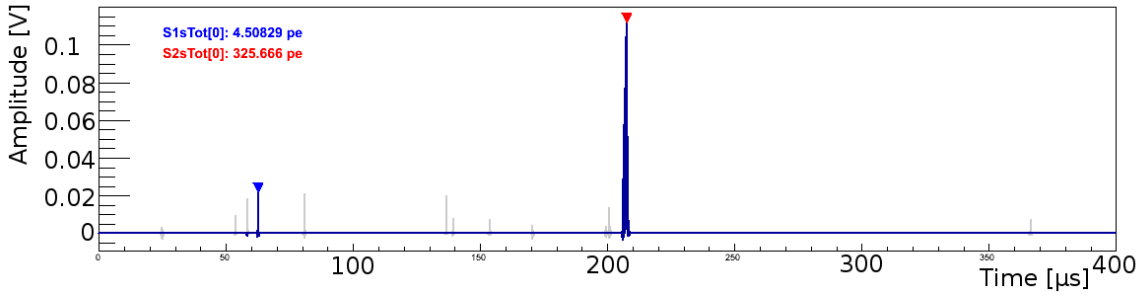


Figure 3.6 – The summed waveform of a typical XENON100 signal. The primary S1 and S2 are separable. With the known electron drift velocity and the time difference between S1 and S2 signal it is possible to determine the interaction depth z.

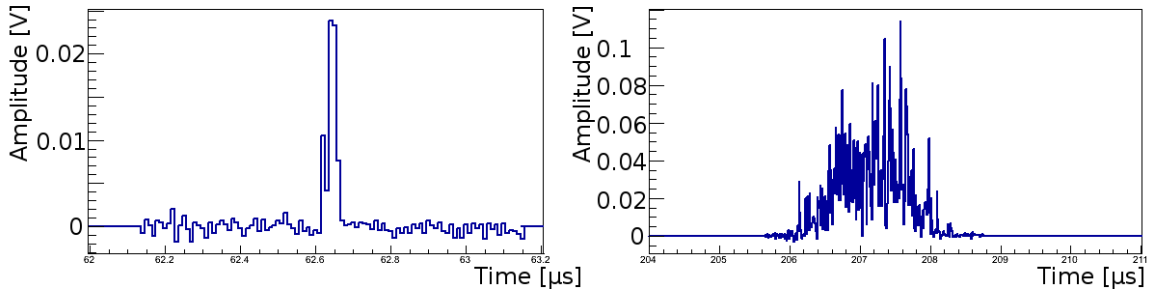


Figure 3.7 – Detailed view on S1 and S2 signal. Both signals show the typical shape on the summed waveform. [115]

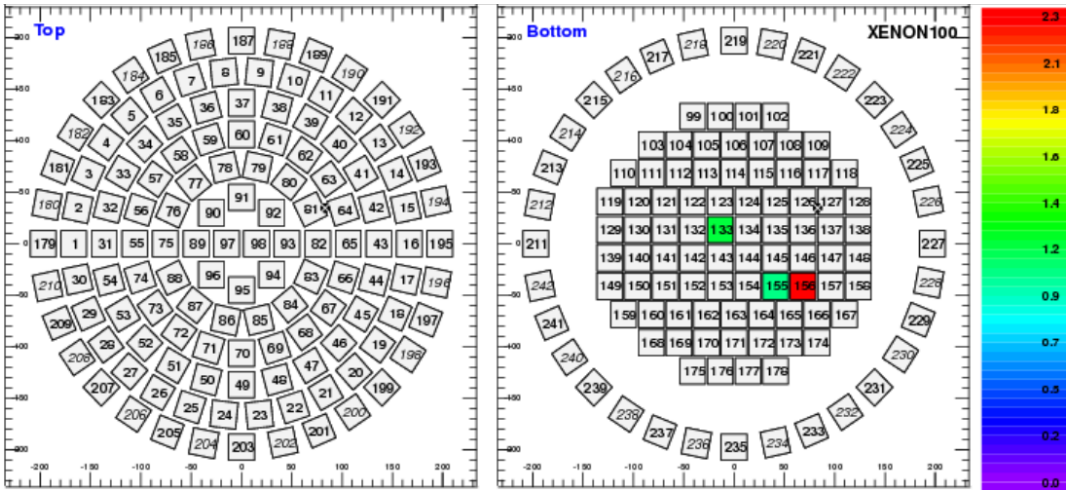


Figure 3.8 – Hit-pattern of the first S1 peak. Its main contribution comes from PMT 156.

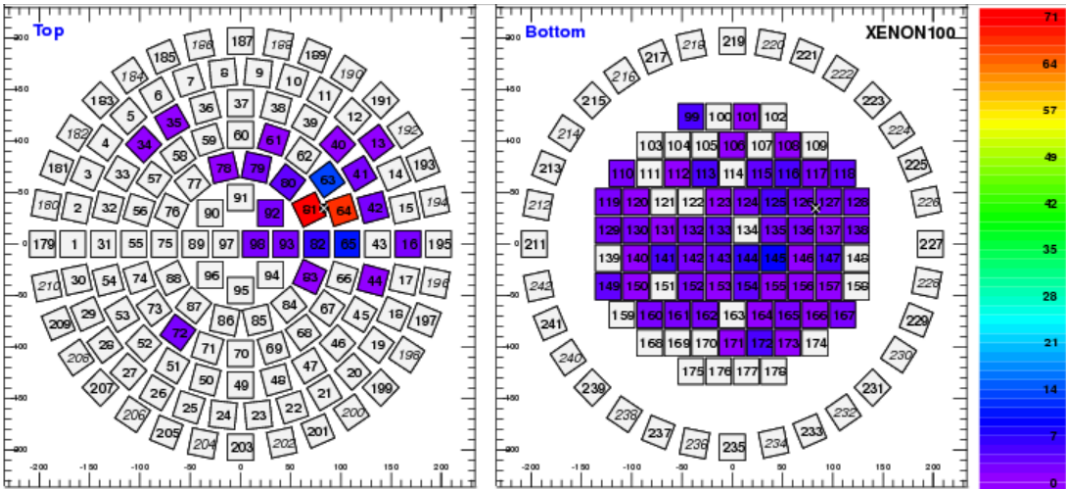


Figure 3.9 – Example hit-pattern of the first S2 peak. The main contribution comes from the top array around PMT 91. The proportional scintillation process creates many photons which are distributed in the whole TPC. Therefore a homogeneous pattern is observed at the bottom PMT array.[115]

3.3.2. The XENON100 TPC

The cylindrical volume of the TPC (30 cm in height and diameter) is delimited by the field cage with Teflon (PTFE) panels to maximise reflections and the cathode, gate and bottom mesh. The TPC is mounted in a low-radioactivity stainless steel vessel. This vessel contains 161 kg of liquid xenon. To separate the outer layer of active veto from TPC inside (inner active volume with 62 kg liquid xenon), there are 24 $^1/_4$ inch PTFE panels installed. Inside the TPC, there is an over-pressure of $\sim$2 bar and above the liquid xenon, a layer of gaseous xenon is in thermal equilibrium with the liquid phase.

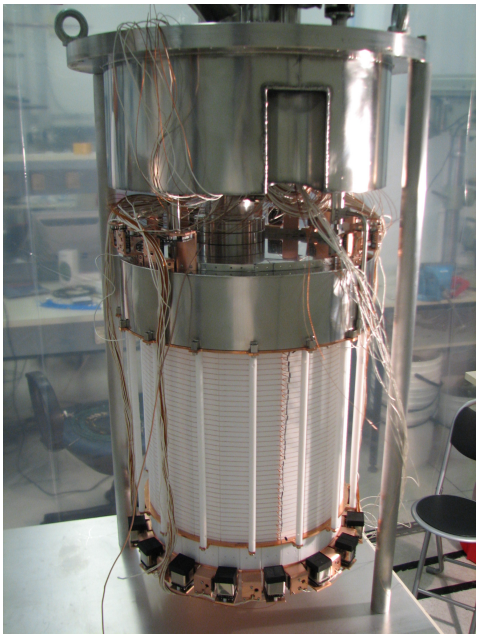


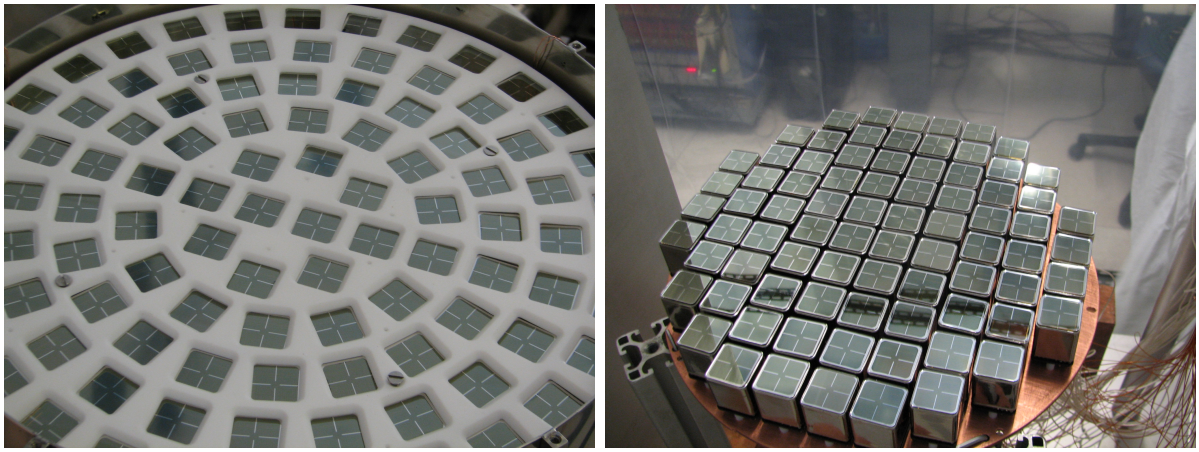
Figure 3.10 – The final XENON100 TPC before mounting into the cryostat. At the bottom a ring of veto PMTs is visible.

Typical operating conditions are at -93° and 2.3 bar absolute pressure. During the cooling process of the TPC the PTFE shrinks. The expected value is 0.0108(5) $^{mm}/_{K}$ and was measured with a similar PTFE piece from the XENON100 supplier before. This is necessary to provide a fully controlled setup in which nothing twists and the exact volume is known. Another advantage is the reflectively of PTFE. Light is distributed homogeneously in the TPC at the first interaction stage. These photons are scattered at the surface of the inner PTFE. This increases their chance to be counted by one of the PMTs. The diffuse reflectively of PTFE dominates over the specular component. The reflectively depends greatly on the surface properties. Values have been reported between 44% to 66% [107], up to exceeding 95% [116]. The PTFE was screened for its radioactive contribution to the background [21].

To operate the TPC in a two-phase mode, the meshes inside the TPC are crucial. There is one above the bottom PMT array and acts as cathode mesh. The resulting electric field is closed at the gate mesh which is below the xenon liquid-gas transition. Another mesh is mounted in front of the top PMT array (anode mesh). The height between the cathode to gate mesh ($\sim$25 cm) is much larger than between gate and anode mesh ($\sim$5 cm). The initial design foresees a negative voltage of -30 kV at the cathode to get an electric field of 1 $^{kV}/_{cm}$ in the liquid phase. The electric field in the gaseous phase has a field strength of $\sim$ 12 $^{kV}/_{cm}$ at a voltage of $+4.5$ kV. Both electric fields are adjusted for stable operation, in particular the cathode voltage is lowered to -16 kV during the data taking periods. Most

important here is the transparency of the gate mesh for electrons. Transport electrons from the liquid to the gaseous phase is crucial for the S2 signal creation. Hence the mesh design is well studied with simulations in XENON100 [82].

At least the photo-multiplier tubes (see section 3.3.3) are assembled in two configurations: bottom and top array. Both arrays are designed to fulfil a given purpose. The top array is used mainly to reconstruct the x/y position of an event (S2). Besides the x/y position the energy of the signal is determined from the produced photons (S1) and hence the PMTs on the bottom array are arranged near to each other to increase the light collection ability. The top and bottom PMT array is shown in figure 3.11a (top) and 3.11b (bottom).



(a) **(b)**

Figure 3.11 – Figure 3.11a shows the top array of XENON100 with its 98 PMTs. They are already placed in a PTFE frame. Figure 3.11b shows the bottom array before the PTFE frame is placed around the PMTs.

3.3.3. The Photo-Multiplier Tubes in XENON100

The signal detection is based on counting photons, produced in the signal creation process, with the photo-multiplier tube (PMT) array. Details about the signal creation process are given in section 3.4. The final signal consists of produced photons with a certain energy. These are observed with arrays of photo-multiplier tubes (PMTs). The XENON100 experiment uses one inch square, metal-channel PMTs of type ”R8520-06-AL“ from Hamamatsu Photonics. These low-radioactively PMTs have a spectral response within $160\,\text{nm} \leq \lambda \leq 650\,\text{nm}$ and are suitable for liquid xenon applications. They show a good response to the typical wavelength of the photons from the signal creation process in xenon ($\lambda = 178\,\text{nm}$).

Figure 3.12 sketches the PMT working principle. Incoming photons hit the photo cathode and excite photo-electrons out of the cathode layer. These electrons are accelerated in the electric field to the first dynode where secondary electron emission multiplies the number of electrons. With each further dynode the electrons are accelerated and further electrons are released. The electrons are collected and present a current pulse. The ratio of the numbers of electrons arriving at the anode to the number of photo-electrons released at the photo cathode, is called *gain*. This current pulse is measured through an impedance of $Z = 50\,\Omega$ as a voltage.

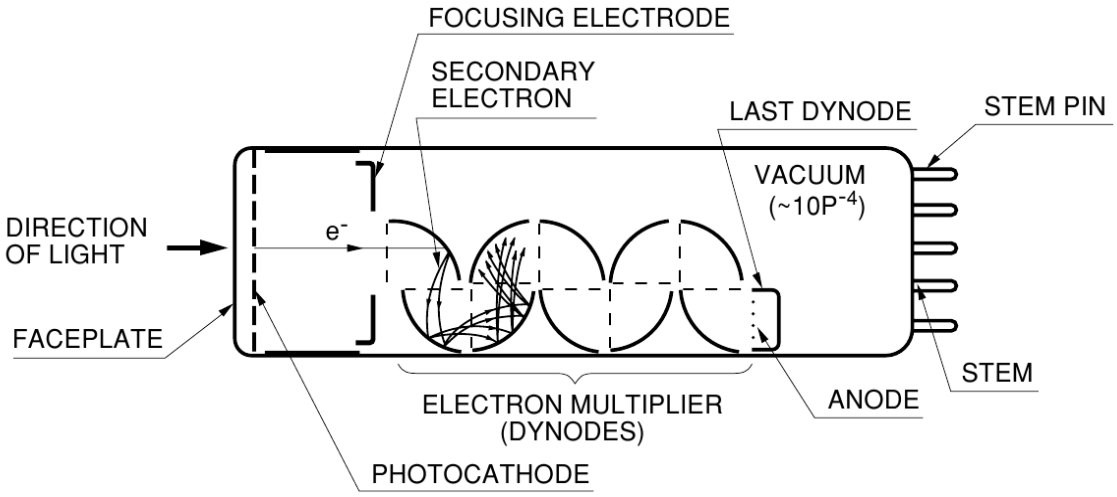


Figure 3.12 – Working principle of a PMT [63]. An incoming photon releases an electron from the PMT photo cathode, which is multiplied by the dynodes. The electrons are collected at the anode.

3.4. Signal Description

The PMT signals are amplified and digitised by Flash Analogue to Digital converters (FADCs). The digitised waveforms of XENON100 are taken with enabled zero length encoding (ZLE) for Dark Matter, calibration and background data. An example for one of these waveforms is given in figure 3.6. This waveform contains the S1 and S2 peak from the scintillation and ionisation process. The ZLE stores only signals above a certain threshold voltage. In addition there is also a time window around the signal stored at the base of each single PMT. Each waveform is 40 µs long and has 40000 samples of 10 ns width. This number is then related to the smallest peak width on the waveform which can be observed for an S1 signal. Both signals, S1 and S2, are determined by the peak-finder of the raw data reprocessing software (*xerawdp*). Since two different physical phenomena are responsible for the S1 and S2 peaks both differ in shape from each other (see 3.4.3 and 3.4.4). In detail there are

at maximum 178 single waveforms (information from top and bottom array) available which build in sum the presented waveform of figure 3.6. Every signal which exceeds a given threshold is used in the summed waveform. This is important for later coincidences between PMTs.

3.4.1. From PMT Pulses to Photo-Electrons

The measured signals are in units of volts. This unit refers to the number of incoming photons at the photo-cathode and the related electron amplification in the PMT. Together with the PMT gain calibration, the measured signal strength is proportional to the incoming particle energy, and hence the signal is measured in units of photo-electrons (pe). The true physical energy of the interaction is determined using the energy calibration (calibration sources). Due to the small amount of deposited energy the usual unit is in electron volt (eV), see section 3.4.5.

3.4.2. Light and Charge Signal Measurement Process

The analytically signal description becomes important for the XENON100 data analysis. Recorded events are analysed due to their origin (light or charge) and therefore a proper formalism is needed to describe the signal expectation in XENON100. This subsection is dedicated to introduce the signal measuring process with its required probability density functions and rate calculations.

The signal chain starts with the creation of S1 and S2 signals which physically means light and charge. The read out ability depends on further parameters such as the electric field strength $\mathscr{E}$ or the interaction type (indicated by $u = nr =$ nuclear recoil or $n = ee =$ electric recoil in the following notation), and hence from the deposited energy. This condition allows defining the combined probability $P(N_\gamma, N_e|E_u, \mathscr{E})$ to get a number of photons (N_γ) and a number of electrons (N_e). Both signal types are treated independent from each other. A signal with a given energy, will show an upward fluctuation in N_γ and at the same time a downward fluctuation in N_e due to energy conversion. This indicates an anti-correlation between charge and light signal. However, there are various other possibilities to lose energy such as heat and therefore the anti-correlation assumption is not guaranteed [14].

$$P(N_\gamma, N_e|E_u, \mathscr{E}) \approx Pois(N_\gamma|n_\gamma) \times Pois(N_e|n_e) \tag{3.7}$$

Equation (3.7) separates the signal creation process in two independent parts with each expectation value n_γ and n_e. Details on the signal creation process are included in the Poisson expectation values, in detail in the W-values which connect the number of expected photons/electrons with the deposited energy. Based on the position of the interaction in the TPC, field corrections becomes necessary.

This is the light reduction factor $S_u(\mathscr{E})$ and the field quenching factor $T_u(\mathscr{E})$ (charge reduction due to recombination). This is written as:

$$n_\gamma(E_u,\mathscr{E}) = \frac{E_u}{W_\gamma(E_u,\mathscr{E})} \approx \frac{E_u}{W_\gamma(E_u,\mathscr{E}=0)}\, S_u(\mathscr{E}) \tag{3.8}$$

$$n_e(E_u,\mathscr{E}) = \frac{E_u}{W_e(E_u,\mathscr{E})} \approx \frac{E_u}{W_e(E_u,\mathscr{E}_{\text{ref}}=\infty)}\, T_u(\mathscr{E}) \tag{3.9}$$

The light reduction factor is $S_u(0) = 1$ and the field quenching $T_u(\mathscr{E} \to \infty = 1)$. With this description, the signal creation process depends on two independent Poisson distribution which are defined by their expectation values for photons and electrons. All necessary corrections are done within the expectation value definition.

3.4.3. S1 Signals and Corrections

Peaks from primary scintillation are only some 10 to 60 ns long on the waveform. The excimer Xe_2^* is suppressed if no additional electric field is applied in the TPC and the recombination of electrons and ionised xenon is dominant. In case of XENON100, the excimer Xe_2^* decay is dominant (see equation (3.3)). Nevertheless, the excited Xe_2^* has two excitation states: A singlet ${}^1\sum_u^+$ which decays with a time constant $\tau_S = 2.2\,\text{ns}$ (short component) and another triplet ${}^3\sum_u^+$ state decay with a long time constant $\tau_T = 27\,\text{ns}$ (long component) [77]. The shape of such a S1 is then formed by a superposition of the fast and slow component as it is described in equation (3.10). Further measurements with α-particles suggest different values of $\tau_S = 4\,\text{ns}$ for the short component and $\tau_S = 22\,\text{ns}$ for the long component [50]. The amplitudes A depend on the kind of the xenon excitation (electronic or nuclear recoil). Both (short and long) components are not distinguishable in XENON100 because of the sampling frequency.

$$S1(t) = A_S\, e^{-\frac{t}{\tau_S}} + A_T\, e^{-\frac{t}{\tau_T}} \tag{3.10}$$

The following notation will differ between the effective measured signal (S1 and S2) and the expected signal (s1 and s2). With this, the measured primary scintillation light from the S1 signal is collected from the *whole TPC*, but due to geometrical boundaries the light collection is not homogeneous. Based on this assumption, each PMT i has an expectation value $s1_i$ from a deposited energy E_u at a given position $\vec{r}$. A position dependent expectation value $s1_i(\vec{r})$ is defined. With this input, the total expected scintillation signal is calculated according to:

$$s1(\vec{r}) = \sum_i^M s1_i \approx n_\gamma(E_{nr},\mathscr{E})\mu(\vec{r}) \tag{3.11}$$

This notation uses $\mu(\vec{r})$ as the light detection efficiency and to be more precise the light detection efficiency depends on the relative scintillation efficiency $\mathscr{L}_{\text{eff}}$ and the measured light yield $\mathscr{L}_{\text{y}}$ which is relative to a reference γ source. Therefore the expected s1 depends on the deposited energy signal becomes:

$$s1(\vec{r}) = E_{\text{nr}}\mathscr{L}_{\text{y}}\left(E_{\text{ee}} = E_{\text{ref}}, \mathscr{E}, \vec{r}\right) \times \mathscr{L}_{\text{eff}}\left(E_{\text{nr}}, \mathscr{E} = 0\right) \frac{S_{\text{nr}}(\mathscr{E})}{S_{\text{ee}}(\mathscr{E})} \tag{3.12}$$

Both parameter $S_{\text{nr}}(\mathscr{E})$ and $S_{\text{ee}}(\mathscr{E})$ result from the field quenching for nuclear and electric recoil and reduce the light yield. The light yield is position dependent and measured at a given reference γ ray source. The relative scintillation yield is also taken with respect to a reference γ ray source at a zero electric field. With this, a S1 signal correction becomes important to balance out the regions where less light is detected (for example at the edges of the top and bottom array). Offline PMTs are also responsible for less observed light. Ideally, the scintillation light is distributed homogeneously over the TPC, including the diffuse scatter at the PTFE walls. Therefore, the TPC is divided into voxels and the S1 signal content is measured from inelastic scattering in xenon (e.g. 40 keV from ^{241}AmBe see subsection 3.6). The 3D S1 corrections allow a detailed study of the light collection. The used coordinates are radius R, azimuths angle θ and depth z. The studies are done for every science run to control the changes. The resulting map, which is shown in figure 3.13 ([115]), is used in the raw data processor to correct the measured S1. This S1 correction map is described analytically as the sum over all PMTs $\mu(\vec{r})$ and with the spacial average $(<\mu>)$ over the TPC the S1 correction map becomes:

$$cS1 = S1 \times \frac{<\mu>}{\mu(\vec{r})} \tag{3.13}$$

The S1 signal correction leads to the possibility to use uncorrected S1 and corrected S1 signals for data analysis. At least, a signal probability density function for S1 is defined with the following two assumptions:

1) The number of generated photons N_γ is Poisson distributed.

2) At each PMT i, the number of photo-electrons ($N_{pe,i}$) is created according to a Binomial distribution.

3) The multiplication of a Poisson and Binomial distribution results a Poisson distribution (bullet points 1 and 2).

and hence the S1 signal PDF becomes:

$$p_{S1,i}\left(S1_i | n_\gamma(E_u, \mathscr{E})\right) dS1_i = \sum_{N_{pe,i}} p_{\text{pmt,i}}\left(S1_i | N_{pe,i}\right) \times Pois\left(N_{pe,i} | n_\gamma, \mu_i(\vec{r})\right) dS1_i \tag{3.14}$$

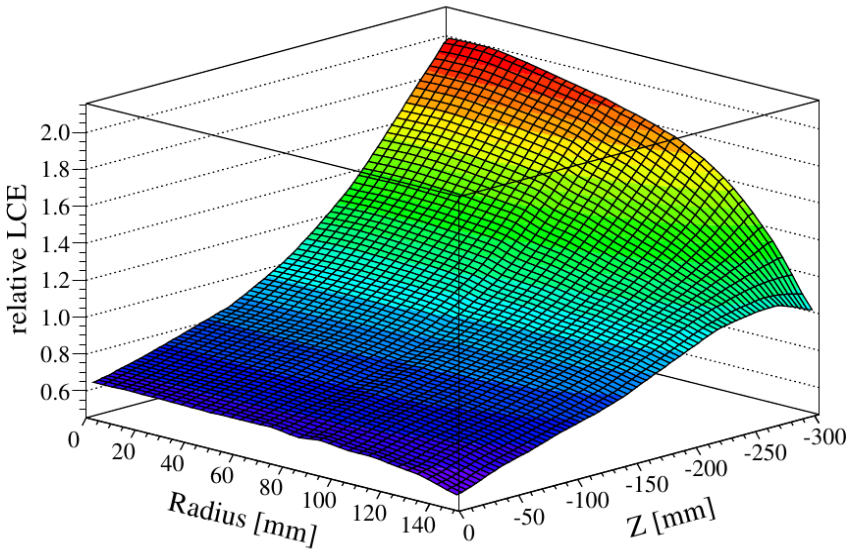


Figure 3.13 – The light collection efficiency from the 40 keV line. Depending on the position in the detector (by radius R and depths Z), the relative light collection efficiency varies within the TPC. A measured S1 signal is divided by a value from this map [115].

The cS1 signal PDF takes the average detector response into account. The number of generated photo-electrons is then summed up by $N_{pe} = \sum_i N_{pe,i}$ over all PMTs and is Poisson distributed with its spatially $cS1 = n_\gamma < \mu >$.

$$p_{cS1,i}\left(S1_i | n_\gamma(E_u, \mathscr{E})\right) dcS1_i = \sum_{N_{pe}} p_{\mathrm{pmt}}\left(cS1 | N_{pe}\right) \times Pois\left(N_{pe} | < \mu > n_\gamma\right) dcS1 \tag{3.15}$$

In both cases the functions p_{PMT} and $p_{\mathrm{pmt,\,i}}$ describe the PMT response. The PMT response is Gaussian distributed with the mean of expected photo-electrons N_{pe} and width of $\sigma_{\mathrm{PMT}}\sqrt{N_{pe}}$ and $\sigma_{\mathrm{PMT}} = 0.5\,\mathrm{pe}$ [115].

3.4.4. S2 Signals and Corrections

Peaks from secondary scintillation are broader and exhibit larger peak voltage on the waveform. The peak width indicates the relation to the interaction depth and is proportional to the relaxation of the produced electron clouds in the ionisation process. Therefore S2 signals are strongly correlated with S1 signals and it is expected that S1 signals come before S2 signals. This knowledge is later applied

by the peak-finder algorithm of *xerawdp*.

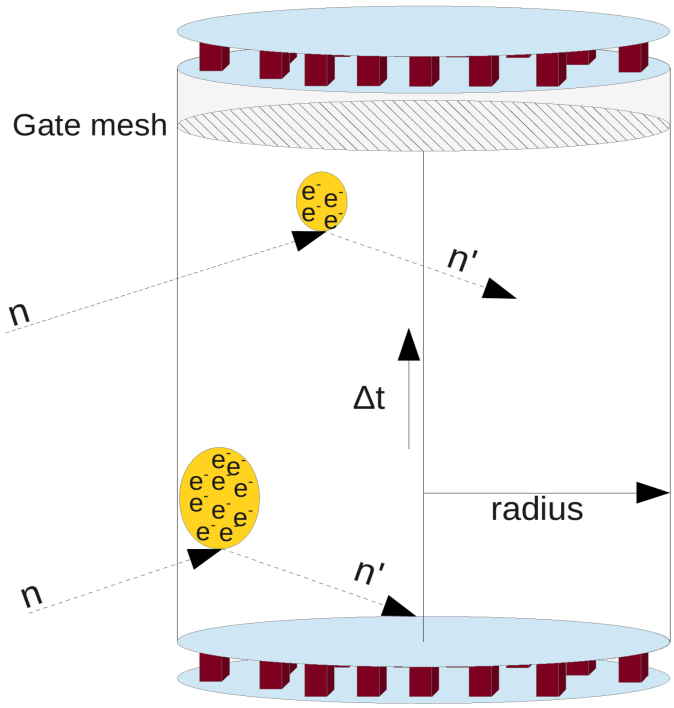


Figure 3.14 – The drift time correction of the S2 signals takes the interaction depth into account. A small energy deposition with a lower interaction depth produces a similarly large S2-signal as can be observed from a large energy deposit at a deeper interaction depth. The drift time correction accounts for the charge losses of he observed S2 signal.

After the ionisation process, the recombination process is suppressed by the electric field in the liquid phase. The process is described by an electron cloud which drifts upwards. The electrons are accelerated in the electric field up to a saturation velocity. On the way to the top of the TPC two processes can occur: Due to impurities in the liquid xenon, the number of electrons decreases. The electrons recombine with the impurities. A higher impurity level reduce the number of electrons extracted into the gaseous phase. Hence the *electron lifetime* must be measured (^{137}Cs, see section 3.6.1). This information is important to estimate if there is a sufficient amount of signal electrons available. The second process comes from the electron cloud itself during its drift. The width of the measured S2 signals depends on the longitudinal diffusion of the electron cloud (L). The upward drifting electron cloud changes shape from a narrow expansion to a broader expansion with increasing drift time. This spacial spread of the electron cloud depends on the electric field and is measured in the broadness of

the S2 signal. This an-isotropic process is described for liquid xenon in [108] for the increasing width in z by:

$$\sigma_z = \sqrt{2\,D_L\,t} \tag{3.16}$$

where D_L presents the longitudinal diffusion coefficient and t is the time after the event occurs.
The operation principle of the XENON100 experiment is explained in section 3.3.1. An important correction which is taken into account during the data analysis is the charge correction. This quantity is relevant to determine the corrected S2 (cS2) signal. The registered S2 signal has to be corrected by the depth of the interaction in the TPC. This correction is referred to drift time correction [16].

$$N(\Delta t) = N_0 \times e^{\frac{\Delta t}{\tau}} \tag{3.17}$$

For example, two S2 signals with the same S2 signal energy are observed on the top array, but their origin depth is a different one (see figure 3.14). Interactions from the lower part of the TPC create a certain amount of electrons and drift upwards. Due to impurities in the LXe, not all of these electrons reach the gate mesh (see equation (3.17)) where they are extracted into the gaseous phase and create a proportional S2 signal. Therefore, interactions which happen at the upper part of the TPC do not need so many electrons in the first place to produce the same amount of S2 signal. This process is described by an exponential decrease, see equation (3.18). The electron loss depends on the typical lifetime τ which is determined by the weekly ^{137}Cs calibration (see section 3.6.1). Based on the electron lifetime τ and the drift time Δt the S2 signal is corrected to cS2 by the exponential electron loss in LXe [115].

$$\mathrm{cS2} = \mathrm{S2}(\vec{r})\, e^{\frac{\Delta t}{\tau}} \tag{3.18}$$

Under this drift time correction, the total expected secondary scintillation light is given by:

$$s2(\vec{r}) = E_{nr}\mathscr{Q}_{\mathrm{y}}(E_{nr}) \exp\left(-\frac{\Delta t}{\tau}\right) \delta(x,y) \tag{3.19}$$

The expected S2 signal uses the $\mathscr{Q}_{\mathrm{y}}$which is the measured charge yield of nuclear recoils events with respect to the applied electric field in $^{e^-}/_{keV_{nr}}$. Equation (3.19) also depends on the deposited energy and $\delta(x,y)$ which summarise the probability that a photon is created at the position x,y in the gas gap (below the top array). By summing up over the PMTs follows $\delta(x,y) = \sum_i \delta_i(x,y)$ and allows to introduce a drift time and spacial corrected S2 signal:

$$cS2 = S2(\vec{r}) \exp\left(\frac{\Delta t}{\tau}\right) \frac{<\delta>}{\delta(x,y)} \tag{3.20}$$

This definition also uses the spatially average $< \delta >$ detection probability. Besides the drift time correction a S2 correction map is applied. As long as the PMTs are not equal to each other in the arrays, the S2 signal depends on the x/y position of the top array. A measurement of the S2 response with fixed energy lines from ^{241}AmBe (e.g. 80 keV or 160 keV) or ^{137}Cs ($E = 660$ keV) determines the light collection at the PMT arrays. These light patterns are determined for each science run and vary only marginally [115]. The observed in-homogeneous result from offline PMTs which bias the S2 response.
Hence, S2 correction means drift time and x/y-plane correction which includes the calibration sources ^{241}AmBe and ^{137}Cs. The energy threshold of XENON100 for S2 signals is at $S2 > 150$ pe, which corresponds roughly to seven measured electrons ($1e^- =\sim 20$ pe, see section 3.4.5) from the signal creation process and is applied to the uncorrected S2 signals to avoid wrong selected corrected S2 signals in case the signal correction fails.
At least, a s2 signal PDF with the same PMT response as in equation (3.14) uses a Poisson distribution for the number of created photo-electrons for the expected number of electrons (n_e) with a given detection probability δ_i:

$$p_{S2,i}\left(S2_i|n_e\left(E_u,\mathscr{E}\right)\right)dS2_i = \sum_{N_{pe,i}} p_{pmt,i}\left(S2_i|N_{pe,i}\right) \times Pois(N_{pe,i}|n_e,\delta_i)dS2_i \tag{3.21}$$

Another important S2 correction is the electric field correction. Dedicated simulations of the electric field in the XENON100 TPC allow estimating edge effects. The radial dependent field line corrections affect the upwards drifting electrons and their true x,y position on the top array [82].

3.4.5. Energy Calibration with XENON100 Signals

With the measured S1 and S2 signals and the applied signal corrections, a translation into physical energy is necessary for a proper signal description (pe $\rightarrow$ eV). Therefore, an energy calibration is essential with the help of the calibration sources (^{60}Co, ^{232}Th, ^{241}AmBe) which are explained in section 3.6. For example, the inelastic scattering lines from neutrons with the xenon isotopes: 40 keV/^{129}Xe, 80 keV/^{131}Xe and further isotopes [115] are used to develop a response energy scale.

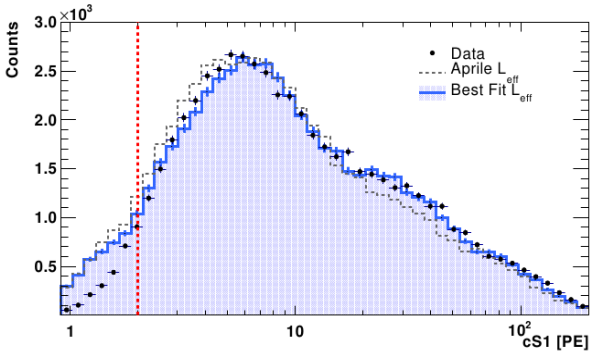
$$E_{\mathrm{S1}} = \frac{cS1}{L_y}\,\frac{1}{\mathscr{L}(E)_{\mathrm{eff}}}\,\frac{S_{\mathrm{ee}}}{S_{\mathrm{nr}}} \tag{3.22}$$

The corrected S1 signal is translated by formula (3.22). Here the electric field suppression factors for electronic and nuclear recoils are given by $S_{\mathrm{ee}} = 0.58$ and $S_{\mathrm{nr}} = 0.95$. Both are assumed to be energy independent. The light yield L_y is determined with electronic recoils in the TCP in respect to

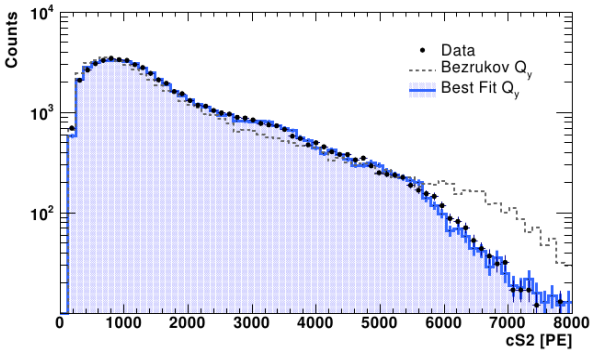
122 $\mathrm{keV_{ee}}$ (electron equivalent, at zero field) for historical reasons. The choice of the γ-ray line of the isotope ^{57}Co with 122 $\mathrm{keV_{ee}}$ is arbitrary but not measured with XENON100 directly. Hence the light yield is measured with calibration sources and the 122 $\mathrm{keV_{ee}}$ measurement is extrapolated. During 2011/12, the light yield is measured as $L_y = 2.28\,\mathrm{pe/keV_{ee}}$. The relative scintillation efficiency $\mathscr{L}_{\mathrm{eff}}$ is unit-less and describes the usual amount of produced light at a given recoil energy. This parameter is determined in scattering experiments, for example [30]. Furthermore the corrected S2 is translated into energy by using equation (3.23).

$$E_{\mathrm{S2}} = \frac{cS2}{Y_a} \frac{1}{\mathscr{Q}(E)_{\mathrm{y}}} \tag{3.23}$$

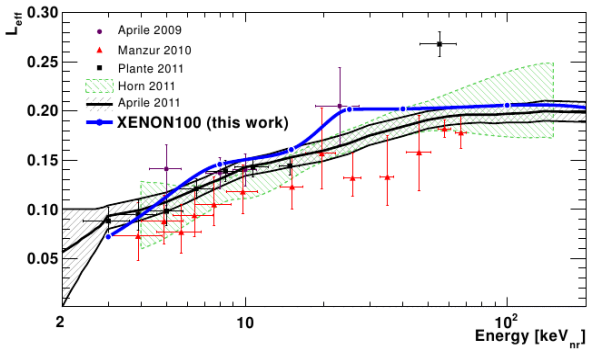
The signal carriers from the interaction point in the liquid phase up to the detection place at the top of the TPC are electrons from the ionisation process. Hence the charge yield $\mathscr{Q}_{\mathrm{y}}$ is important to determine the number of free electrons per deposited energy portion. Due to the electric field, the electrons drift upwards and are extracted from the liquid to the gaseous phase where another stronger amplification leads to the secondary scintillation. This process is summarised by an amplification factor Y_a which describes finally the amplification by photo-electrons per electron.
Studies on signal amplification require a fine data selection to select single upwards drifting electrons. From the latest study about the single photo electrons in XENON100 the signal amplification reaches 19.67 $\mathrm{pe/}e^-$ with 1 σ deviation of 6.98 $\mathrm{pe/}e^-$ [22].
The energy calibration depends on two quantities, the relative scintillation efficiency $\mathscr{L}(E)_{\mathrm{eff}}$ and the ionisation ability $\mathscr{Q}(E)_{\mathrm{y}}$ which is measured in independent experiments with xenon. Besides this, the XENON100 collaboration proofs a well established understanding of the $\mathscr{L}(E)_{\mathrm{eff}}$ and $\mathscr{Q}(E)_{\mathrm{y}}$ by confirming the independent measurements by a Monte Carlo simulation. Details on the simulation and the created S1 and S2 signals are given in section 3.4.6. The simulation reproduces the cS1 and cS2 spectrum and obtain the quantities $\mathscr{L}(E)_{\mathrm{eff}}$ and $\mathscr{Q}(E)_{\mathrm{y}}$ from a fit to the spectra. This is summarised in figures 3.15a and 3.15b where the cS1 and cS2 spectra are shown with their fits. The result shows a good agreement in the spectral shape. Figure 3.15c and 3.15d show $\mathscr{L}(E)_{\mathrm{eff}}$ and $\mathscr{Q}(E)_{\mathrm{y}}$ with external measurements. The simulation shows a good agreement with the data. The quantity $\mathscr{L}_{\mathrm{eff}}$ is not measured for low energies. Therefore, the red line in figure 3.15a indicates the lower boundary for the cS1 spectrum in where fit and simulation/data are in good agreement [23]. Finally the broad spectrum of calibration sources and their activation lines allow to define an energy scale. Due to the anti-correlation of light and charge signal, a combined energy scale from S1 and S2 signals allows a finer energy resolution based on the ^{137}Cs measurement [115].



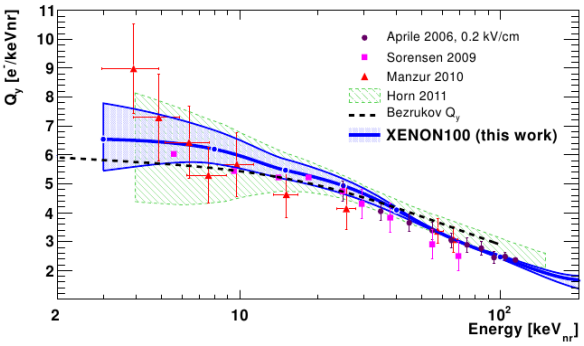
(a) ^{241}AmBe Simulation: cS1 spectrum with fit.



(b) ^{241}AmBe Simulation: cS2 spectrum with fit.



(c) ^{241}AmBe Simulation: Relative scintillation efficiency $\mathscr{L}(E)_{\text{eff}}$



(d) ^{241}AmBe Simulation: Charge yield $\mathscr{Q}(E)_{\text{y}}$

Figure 3.15 – The XENON100 collaboration shows a good understanding of the neutron interaction in TPC. A detailed simulation reproduces the cS1 and cS2 spectra which fit the data quite well. Moreover the $\mathscr{L}(E)_{\text{eff}}$ and $\mathscr{Q}(E)_{\text{y}}$ show a good agreement with external measurements. In case of the relative scintillation efficiency it is simulated down to 2 keV$_{\text{nr}}$. [23]

3.4.6. Monte Carlo Simulation of S1 and S2 Signals

To understand the detector response completely an additional detector simulation is realised. The main goal is a neutron simulation of the XENON100 detector with its most important parts with a full reconstruction of the S1 and S2 signal. The Monte Carlo method is implemented in Geant4 code and uses the geometrical input from the detector design. The following section gives a short introduction to the neutron simulation itself. Furthermore, the neutron simulation is adjusted to simulate the WIMP response in the detector. This is possible because of the similarity of WIMP and neutron interactions with the xenon nucleus.
The simulation is done by the *Nuclear Recoil/WIMP Simulation Code* which is partly based on *XEPHYR*. Both tools are developed and used within the XENON collaboration. The *Nuclear Recoil/WIMP Simulation Code* uses on one hand the outcome of *XEPHYR* to calculate the WIMP recoil spectrum in xenon and on the other hand the input from a calibrated neutron spectrum which is normalised to the ^{241}AmBe calibration in XENON100. Both inputs are plugged in a full detector simulation.

The neutron interaction model takes the XENON100 TPC, the shield configuration and the surrounding environment into account. The Geant4 simulation calculates the neutron cross-sections using the ENDF/B-VI/B-VII database, which is already implemented in G4NDL3.13. The simulated neutron source assumes a standard ^{241}AmBe spectrum (ISO8529-1). For the absolute normalisation, the ^{241}AmBe source of XENON100 was compared to a reference source at the Physikalisch-Technische Bundesanstalt with the Bonner sphere neutron spectrometer. The resulting integral source strength is 160 ± 4 n/s. The subsequent detector simulation of the deposited energy result similar variable such as S1, S2, event position in x/y/z and drift time. The output is stored in a ROOT specific data format analogue to the reprocessed data of the XENON100 experiment. The simulation includes the detector threshold, resolution effects, fiducial volume (X34kgV, see section 3.7) and cut efficiencies. The neutron simulation also takes multiple scatters into account. The XENON collaboration presented a good detector understanding by comparing the simulation with neutron data [23], [113].

As long as no WIMPs are observed, there is evidently no measurement of the WIMP response of the XENON100 detector. Compared to neutrons, WIMPs are assumed to have only single scatters in the detector due to their low cross-section. The WIMP mass itself is unknown and therefore their nuclear recoil spectra are modelled for different WIMP masses. *XEPHYR* calculates the Dark Matter cross-section and rates for XENON100. Furthermore it is possible to choose astrophysical parameters such as the WIMP velocity distribution or the Dark Matter density. The WIMP simulation takes the same detector inputs into account as the neutron simulation.

The signal conversion is in both cases done by the formulas (3.22) and (3.23) in section 3.4.5. The simulation compares a simulated energy distribution with observed S1 and S2 signals. Due to the similarity of WIMPs and neutrons, the WIMP response can be determined for different science run conditions. Keystone in the simulation is the variation of $\mathscr{Q}_y$ and $\mathscr{L}_{eff}$. To begin with, the ionisation channel is used to determine $\mathscr{Q}_y$ with the measured $\mathscr{L}_{eff}$ (according to reference [16]). The simulated cS2 spectrum is adjusted to the cS2 data (see figure 3.15b) by varying the spline interpolation of $\mathscr{Q}_y$. This *best-fit* $\mathscr{Q}_y$(see figure 3.15d) with systematic uncertainties from e.g. the $\mathscr{L}_{eff}$ results a good agreement of simulated cS2 with the cS2 data down to $\sim 3\,\mathrm{keV_{nr}}$. With the information about the best-fit $\mathscr{Q}_y$, the $\mathscr{L}_{eff}$ is determined from the simulation by the same technique. The variation of $\mathscr{L}_{eff}$ (see figure 3.15c) is used to match the simulated cS1 spectrum with the cS1 data (see figure 3.15a) down to 2 pe. The quality of the neutron simulation in the XENON100 TPC is shown in figure 3.16 where the outcome of the simulation is compared to the data. The relative deviation is within 2%. This allows to simulate low neutron energy recoils with a high accuracy and using simulated neutrons instead of ^{241}AmBe calibration data as signal prediction for the actual ongoing data analysis of the 2013/14 data. Even more, the WIMP response is simulated with *Nuclear Recoil/WIMP Simulation Code* for future publication where the the data taking periods 2011/12, 2013/14 and another one before 2011 is combined together.

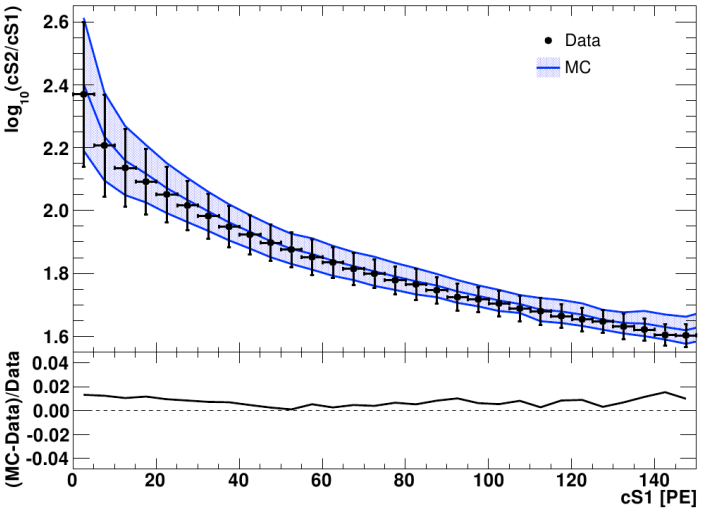


Figure 3.16 – The simulated neutrons are compared to the data in the flattened space. The simulation shows a good agreement with the data within 2% relative deviations.[23]

3.4.7. Event Rate Calculation in XENON100

This subsection is closed by the event rate calculation which is important to determine the expected signal and background rate from whole data taking period [15]. The WIMP-nucleus scattering rate dR/dE_{nr} results from the physically interaction. The resulting S1 and S2 signal is measured as the differential nuclear recoil rate:

$$\frac{d^2R}{dS1\,dS2} = \varepsilon\left(S1,S2\right)\int \frac{dR}{dE_{nr}} p\left(S1,S2|E_{nr}\right) dE_{nr} \tag{3.24}$$

The ε describes the detection efficiencies and it is needed to take a combined probability density function function for S1 and S2 into account. As already shown in subsection 3.4.3 and 3.4.3, both signals are defined their own signal PDF (3.14) and (3.21) and therefore equation (3.24) is separated to:

$$\frac{d^2R}{dS1\,dS2} = \varepsilon_{1,s}\left(S1\right)\varepsilon_{2,s}\left(S2\right) \times \int \frac{dR}{dE_{nr}} p_{S1}\left(S1|E_{nr}\right) p_{S2}\left(S2|E_{nr}\right) dE_{nr} \tag{3.25}$$

This separation allows to apply S1 and S2 detection efficiencies and make use of the S1 and S2 corrections what allows to carry out the data analysis with the corrected or the uncorrected S1/S2 signal. The S1 to E_{nr} signal relation is measured more precise (and to lower values) with $\mathscr{L}_{\text{eff}}$ than the S2 to E_{nr} relation with $\mathscr{Q}_{\text{y}}$. Therefore the final analysis is done with the differential rate of the corrected S1 signal by:

$$\frac{dR}{dcS1} = \varepsilon_1\left(cS1\right) \times \int \frac{dR}{dE_{nr}} \varepsilon_{2,E}\left(E_{nr}\right) p_{cS1}\left(cS1|E_{nr}\right) \tag{3.26}$$

This allows to define the data selection criteria (efficiencies) depend on cS1 signal. The term $\varepsilon_{2,E}\left(E_{nr}\right)$ within the integral make use of the S2 acceptance calculation, which depends on the cS1. The S2 threshold cut influences the registered S1 signal directly and therefore its acceptance is applied before the Poisson fluctuations and uncertainties are taken into account.

3.5. Data Acquisition, Trigger and Raw Data Processing

The XENON100 data acquisition system (DAQ) is responsible for triggering, digitising and storing the information from each of the 242 PMTs. In addition, a slow control system monitors vital parameters like temperature, pressure, gas flow or xenon liquid level in the TPC and stores the data in a SQL database. The following section briefly introduce the DAQ and describes the raw data processing with Xerawdp. Last include different knowledge from the calibration sources such as electron lifetime and necessary S1/S2 corrections. The whole data analysis rely on the peak-finder algorithm which is

implemented in Xerawdp. Later studies in this document on LED calibration data re-processing with Xerawdp include the same peak-finder algorithm and therefore a short introduction is given here.

3.5.1. Trigger Efficiency

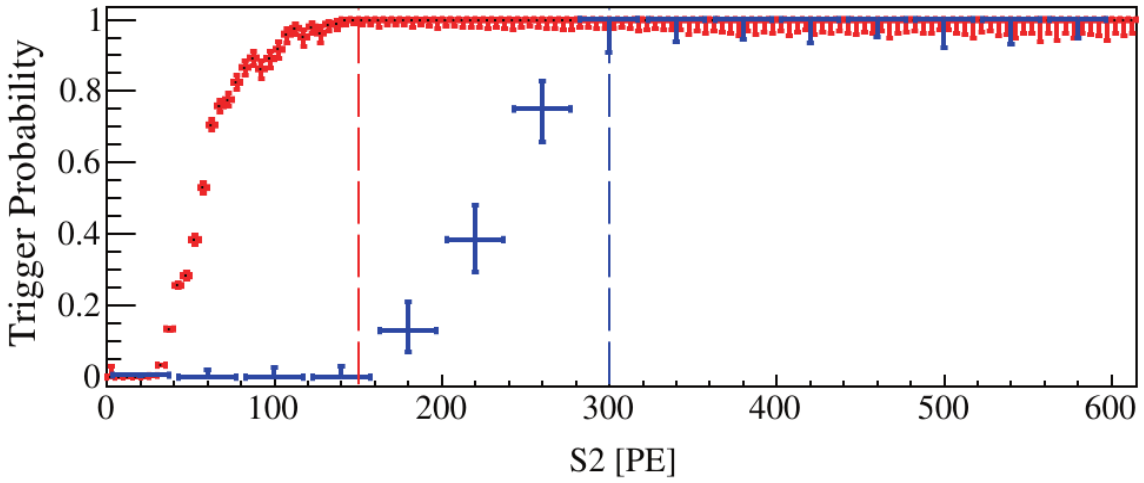


Figure 3.17 – The trigger probability in XENON100. The threshold of 300 pe was lowered to 150 pe from the data taking period in 2011/12. The reasons for the lower threshold are improved run conditions (e.g. higher light yield). The latest trigger probability was evaluated with ^{241}AmBe data from before and after 2011/12. [15]

The threshold for the S2 signal is determined by the S2 trigger efficiency (see figure 3.17), which is $> 99\%$ for $S2_{th} = 150\,\text{pe}$ to detect S2 signals. Based on that assumption the S2 signal acceptance (*Xs2peaks* cut) is defined as

$$\text{Xs2peaks} := S2 > 150\,\text{pe} \tag{3.27}$$

in the XENON100 data analysis.

To measure the S2 trigger efficiency with the XENON100 TPC, 68 inner PMTs of the top array and 16 centre PMTs of the bottom array are used. The analogue summed signal is amplified and passed through a low threshold discriminator [15]. The measurement was done in three different ways [115]: Analysing the (triggered) S2 spectrum of different sources, using a square-wave test pulse of a known size and using real S2 signals of dedicated background and calibration measurements [15]. The S2 trigger efficiency was determined during a previous data taking period (before 2011, blue data samples in figure 3.17) and again for the 2011/12 (red data samples) data taking period with an improved trigger efficiency [15]. The S2 trigger efficiency has not changed since then and the Xs2peaks cut is used in the same way as before. What has changed since then are the slightly different detector conditions which need to be taken into account to re-evaluate the S2 acceptance after a calibration

with Americium-Beryllium (^{241}AmBe).
It is expected that WIMPs and neutrons show a similar interaction process with the xenon nuclei. Differences arise from the fact that neutrons have an energy spectrum up to 10 MeV which is different from the expected WIMP energy spectrum. With this, a different spectral shape for nuclear recoils is expected for WIMPs and elastic scattered neutrons. In addition the mean scattering path of a 1 MeV neutron in liquid xenon is ~ 10 cm (Based on [112] and a scattering cross section of $\sigma_{n-Xe} = 0.7$ barn) This allows double scatters within the TPC from neutrons which is not expected to happen for WIMPs. Moreover the neutrons penetrate the liquid xenon from all directions what leads to a typical spatial event distribution. The neutron introduces events are clustered at the edges of the TPC and are reduced due to the centre of the TPC.

3.5.2. The DAQ System

The data acquisition is done by the amplifier boards (10 $\times$ Phillips Scientific 776 NIM boards, [105]) and the boards (CAEN VME V1724 Flash ADCs) for digitisation, [66]). Each single PMT signal is amplified by a factor of 10 by the Phillips boards. The signal digitisation by the CAEN boards has a resolution of 14 bit and a nominal sampling voltage of 2.25 V at full scale with a 40 MHz bandwidth. A single channel is translated to voltage by formula (3.28).

$$1\,\text{channel} \quad \widehat{=} \quad \frac{2.25\,\text{V}}{2^{14}} = 1.373 \times 10^{-4}\,\text{V} \tag{3.28}$$

Therefore the smallest accessible voltage of one single channel is 0.1373 mV. The digitised waveform has samples of 10 ns width. Figure 3.18 shows an overview about the DAQ system of XENON100.
During calibration and Dark Matter data taking only the relevant information is stored. A baseline is estimated before the data acquisition program is executed. Every time an event triggers the DAQ, the data are recorded with a symmetric window of 20000 samples before and after the event and hence the trigger signal is in the middle of the recorded waveform. The trigger system (see also subsection 3.5.1) is configured to detect even small S2 signals of 150 pe ($\sim$7 electrons) with an efficiency of > 99%. Smaller energies are also recorded but their trigger efficiency decreases rapidly. Data are acquired with a threshold of 30 channels (corresponding to 0.412 $^{\text{mV}}/_{0.3\,\text{pe}}$). Each sample which exceeds this threshold is recorded and in addition a symmetric pre- and post-trigger window of 50 samples each is acquired. The remaining samples on the recorded waveform not transferred from the ring buffer to the DAQ readout computer. This method is called zero length encoding (ZLE) and is already implemented in hardware. Taking into account that a S2 is responsible for the trigger, the corresponding S1 peaks are observed before the S2 peak. This information is used in *xerawdp* to determine proper S1 and S2 peaks from the raw data.

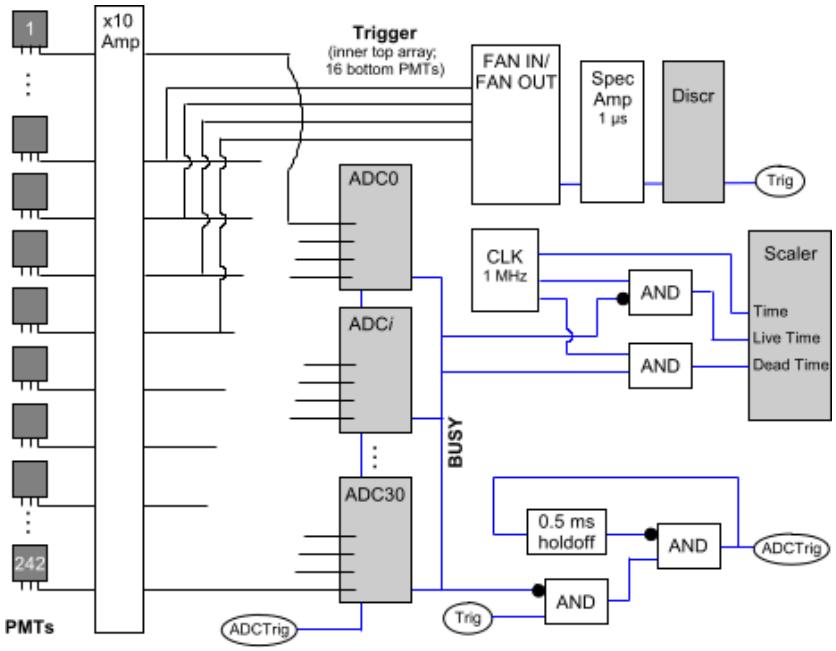


Figure 3.18.: The DAQ system in XENON100 with the Phillips amplifier boards and the CAEN boards [115].

3.5.3. Raw Data Processing in XENON100 with Xerawdp

3.5.3.1. Overview of Raw Data Processing

After a triggered waveform is stored to binary files, the physical information has to be extracted. This is done by another tool, the raw data processor *xerawdp*. It works on 3 levels: First the S1 and S2 peaks are determined by the peak-finder algorithm. Then basic physical information like the S1 and S2 peak areas or their heights are extracted. The final processing level applies signal correction information (e.g. electron lifetime or electric field shape), reconstruct the event position in x/y/z and additional information like the signal entropy is calculated.
The output of *xerawdp* is stored in a separated ROOT files with a database-like structure which represents the three reprocessing levels. These files do not contain the full set of waveforms to save storage on the hard disc and to improve speed during analysis. The raw data are stored in another place to have them as a backup and to reprocess them if necessary, for example if new variables are introduced into the analysis. This is done regularly and *only* cost computation time. This is an advantage to any other experiments (e.g. ATLAS/LHC) where a trigger system decides which data are recorded and which are lost forever. Of course, this is only possible because XENON100 is a low rate experiment.

3.5.3.2. The Peak-finder Algorithm in Xerawdp

The main task of *xerawdp* is to find the recorded peaks in the stored raw data (waveforms). Each waveform has a time stamp to sort them chronologically. The peak-finder algorithm itself is applied to the summed waveform which consists of all waveforms from the top and bottom arrays (178 PMTs) under the condition that some PMTs are switched off because of their high noise contribution. This method is less time consuming in comparison to running the peak-finder algorithm on each single PMT waveform. A low-pass frequency filter (raised cosine) is applied to cut off high frequencies above 3 MHz. A schematic example is given in figure 3.19 where the S2 peak structure from many waveforms is smoothed (magenta).

The procedure starts with identifying S2 peaks first. A possible S2 peak has to exceed a threshold of

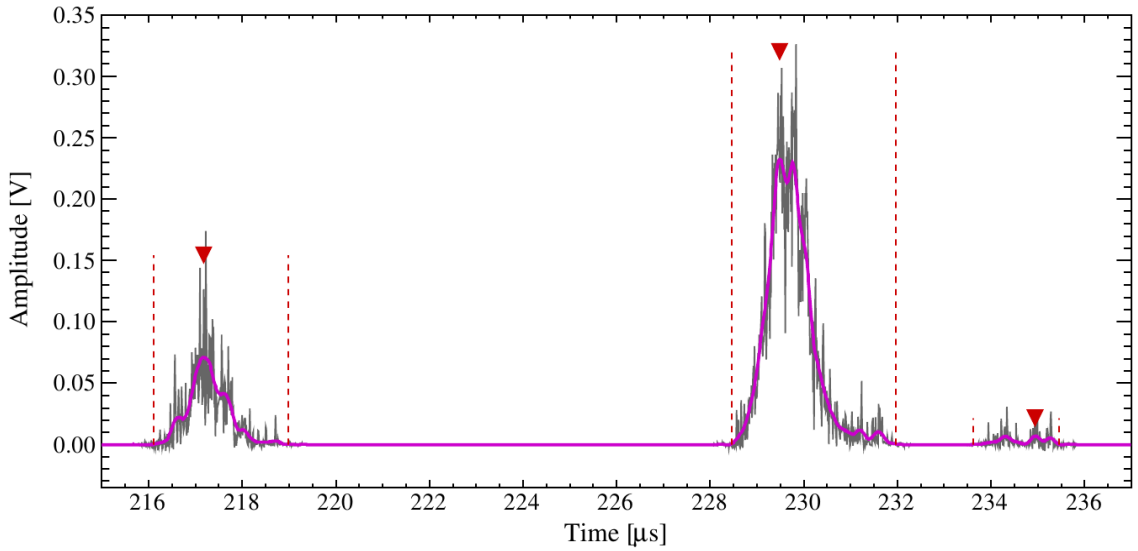


Figure 3.19 – This example shows the raised cosine filter in *xerawdp* which is used to search for large S2 signals on the waveforms [95]. The result of the filter is shown in magenta. The filter smooths the waveform.

10 mV for at least 600 ns. As soon as the average value of 210 samples (2.1 ns) before and after the peak is lower than 25% of the measured amplitude the peak is completed. After the largest S2 peak is identified further smaller S2 peaks are searched (maximum 31 additional ones). This additional S2 peaks differ as multiple S2 peaks on the waveform and small S2-like peaks in between. The actual data-taking procedure takes 32 S2 peaks at maximum into account. The size of the S2 peaks of 150 pe is not a necessary condition for the peak-finder algorithm. Only the trigger efficiency for $S2 < 150$ pe in the DAQ decreases to less than 99% and these identified S2 peaks refer to single electrons when a valid S1 peak is found.

Physically related S1 peaks are distributed only before the first S2 peak which exceeds 150 pe. The S1 identification is more complicated. Initially, the unfiltered waveform is scanned for samples which exceed a threshold of 3 mV ($\sim$0.33 pe). The boundaries are determined where the S1 drops below 0.5% of the peak height for 20 ns. Furthermore the S1 peak has to show the typical decrease behaviour: The average of 0.5 µs before the S1 peak position is less than 1% and 100 ns after the S1 peak position is less than 4% of the S1 peak maximum. To distinguish between S1 and small S2 peaks, the S1 peak FWHM of the filtered waveform then is limited to 0.5 µs. Overall, the S1 peak maximum must be three times larger than the largest negative sample within the boundaries of the identified peak. In total a maximum number of 32 S1 peaks can be identified and stored [115]. After the S1 peak is identified on the summed waveform with its left and right boundaries, the single PMT waveforms are analysed within the same left and right boundaries. Using the gain information of the LED calibration the S1 signal is calculated on the single waveforms and stored in the $S1s[pmt][i]$ variable in units of photo-electrons by *xerawdp*.
In any case an uncertainty on the peak-finder algorithm is left to miss-identify small S2 peaks as S1 peaks and vice versa. Furthermore a necessary condition to apply the peak-finder in *xerawdp* is the activated ZLE encoding to suppress background. Hence, a LED data set can not be processed with the current version of *xerawdp*. Detailed information on the peak-finder algorithm with its detailed settings and parameters is given in [95].

3.6. Detector Calibration

The regular calibration procedure of the XENON100 experiment is essential for the later Dark Matter detection. A set of radioactive sources covers the electronic and nuclear recoil response of the TPC. Furthermore, PMT calibrations with LEDs are necessary to verify the proper functioning of single PMTs and to determine their gain. The electron lifetime is important for the 3D event reconstruction and the S2 signal yield. This quantity is determined by calibration of the photo peak of a Caesium source. The following section explains the sources and provides details about the calibration procedure. Furthermore the PMT calibration is described. During the XENON100 live time multiple calibration data sets are taken. But later data analysis does not use all data sets. Details about the specific data sets in XENON100 analysis work are given in the according subsections of chapter 6.
To assure that enough calibration data sets are available during the science run, a time table foresees a regular calibration with ^{60}Co and ^{232}Th, each for roughly 24 hours a week. The PMT calibration and the ^{137}Cs calibration each take half an hour a week. Hence roughly 120 hours of Dark Matter data are taken per week.

3.6.1. Measurement of the Lifetime with ^{137}Cs

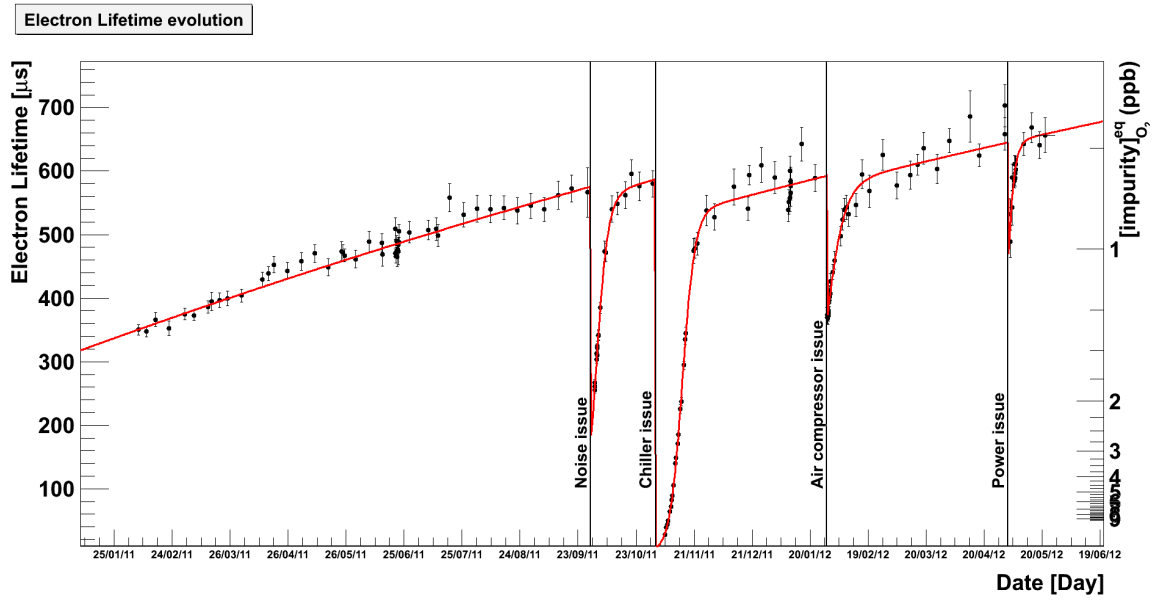
The S2 signal creation process depends on the ionisation yield such it is already discussed in section 3.4.4. The electron lifetime is determined from ^{137}Cs. The decay is shown in figure 3.21, here the ^{137}Cs decay is divided into two branches with a 94.6% chance for an electron with a maximum energy of $E^{\max}_{e^-,1} = 0.512\,\text{MeV}$ and 5.4% for an electron with $E^{\max}_{e^-,2} = 1.174\,\text{MeV}$. The first decay leads to meta-stable barium ($^{137\text{m}}$Ba) and the second to the ground state ^{137}Ba. The meta-stable state decays again with a half-life of $\tau_{1/2} = 2.55\,\text{min}$ to the ground state. Both decays lead to a 100% electron production. A third branch with an end-point energy of 0.892 MeV is strongly suppressed by $\mathcal{O} \sim 10^{-4}$. The fully absorbed gamma at 662 keV produces the calibration peak. The electron lifetime is determined by a the analysis of ^{137}Cs calibration data sets for its measured S2 signal and drift time. This is fitted by an exponential function where the decrease of the electron lifetime is described by τ_e such it is shown in equation (3.29) [115].

$$S2(E) = S2_0(E) \exp\left(\frac{\Delta t}{\tau}\right) \tag{3.29}$$

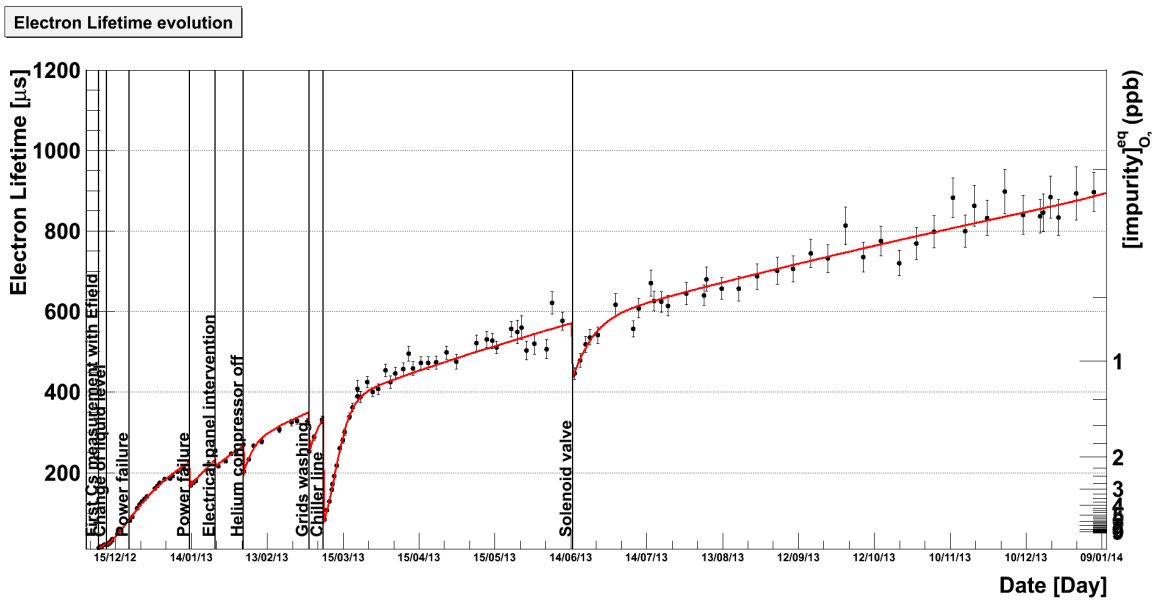
From weekly calibration the outcome of the electron lifetime for 2011/12 and 2013/14 are shown in figure 3.20a and 3.20b by the data points. Both figures show that the ongoing purification process leads to a cleaner xenon and the electron lifetime increases continuously. The composed function which is used in figure 3.20a and 3.20b differs from each other in both data taking periods. The functions consist of different exponential rises to approximate the electron lifetime. Different truncation results from technically issues such as a power failure are also shown. The electron lifetime recovers pretty fast once the experimental setup is restored.

3.6.2. Electronic Recoil Calibration with ^{60}Co and ^{232}Th

The calibration with the isotopes ^{60}Co and ^{232}Th are both performed to measure the electronic recoil response. The isotope ^{60}Co has a half-life of 5.3 years and is produced by activation of ^{59}Co with neutrons. The decay scheme is shown in figure 3.24b. Here ^{60}Co decays by β-decay into an excited state of ^{60}Ni, which promptly decays releasing two photons ($E_1 = 1.172\,\text{MeV}$ and $E_2 = 1.3325\,\text{MeV}$). Both gamma-ray particles with their typical energy may interact with the xenon atoms and generate free electrons mainly by the Compton effect.
The naturally occurring ^{232}Th isotope has a lifetime of 1.4×10^{10} years and is part of the thorium decay series which ends with lead (^{208}Pb). It decays by emitting an α particle with $E_\alpha = 4.083\,\text{MeV}$. Two further gamma particles of 140 keV and 63 keV are released (see decay scheme in figure 3.24a). These gammas again generate electrons mainly by the Compton effect.
The resulting data of ^{60}Co and ^{232}Th are shown in figure 6.1a for 2011/12 and 6.1b for 2013/14



(a) The electron lifetime development in 2011/12.



(b) The electron lifetime development in 2013/14.

Figure 3.20 – Overview about the electron lifetime measurements during 2011/12 and 2013/14with a weekly ^{137}Cs calibration. Single measured electron lifetimes are described by a complicated function (exponential) which describes the different truncation within the data taking periods.

(see chapter 6, page 204). After selecting single scatter events (one S1 and one S2), the data set is represented by a band in discrimination space. This band is above the region of interest for Dark

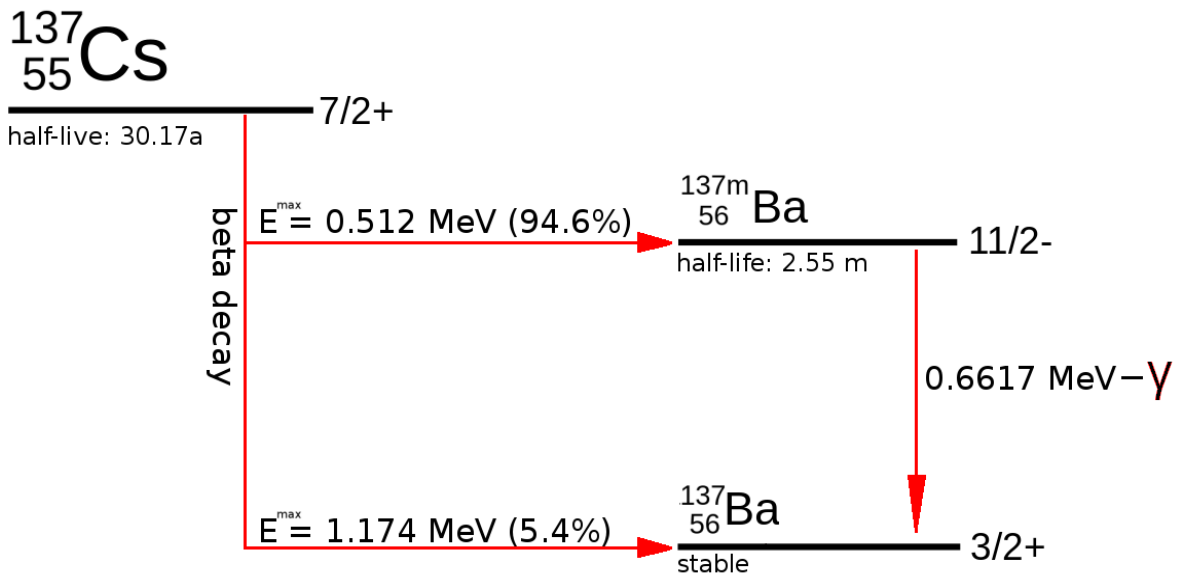


Figure 3.21 – The decay scheme of ^{137}Cs: A first β^- decay with 94.6% branching ratio leads to metastable barium (^{137m}Ba). Remaining β^- decays with 5.4% branching ratio lead to stable barium. Both decays release electrons with maximum energy of $E^{\max}_{e^-,1} = 0.512\,\text{MeV}$ and $E^{\max}_{e^-,2} = 1.174\,\text{MeV}$ energy, respectively the full absorption peak in xenon comes from γ emission of $E_\gamma = 662\,\text{keV}$ which is deposited completely. [88]

Matter search. A dedicated model of the electric recoil background is described in chapter 6. During 2011/12 roughly 25 days of ^{60}Co and 30 days of ^{232}Th data sets were taken. In 2013/14 roughly 21 days of ^{60}Co and 19 days of ^{232}Th data sets were taken (see figure 3.22 and 3.23).

Goal of calibration with gamma sources is to get knowledge about the event distribution of such events in the XENON100 TPC and their typical response in energy. This allows to quantify the background from other electronic recoil, for example ^{129m}Xe, ^{85}Kr or intrinsic radioactivity of all detector components [15]. The typical response to the ER calibration sources is shown in section 3.9 by the blue markers in figure 3.31. The individual energy deposit of the Compton effect is unknown and therefore the whole ER background spectrum is measured. The acquired data from the ER calibration sources are event counts with a certain energy in S1 and S2. Based on this information, it is necessary to determine the ER background from the number of counted events (rate). This is possible for the main share of acquired data. Nevertheless, there are event populations observed which have a lack of statistics. These events are named ”anomalous leakage events“ and are further discussed in section 6. The WIMP search with low-radioactive experiments is only possible with a very good background description. Therefore it would be ideal to model the background based on the physic processes and

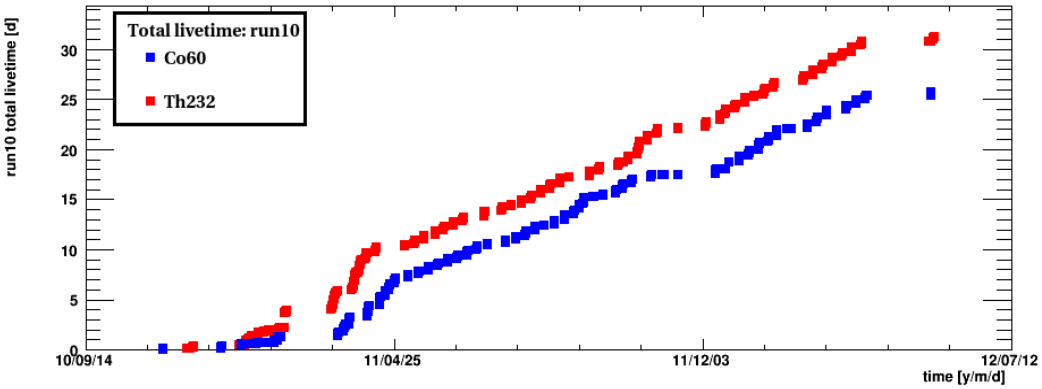


Figure 3.22 – The total time of 2011/12 for ^{60}Co and ^{232}Th data taking.

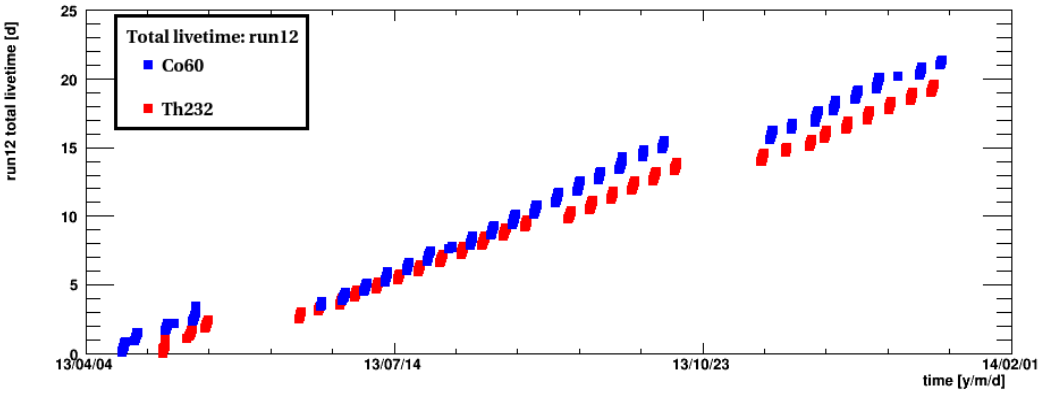
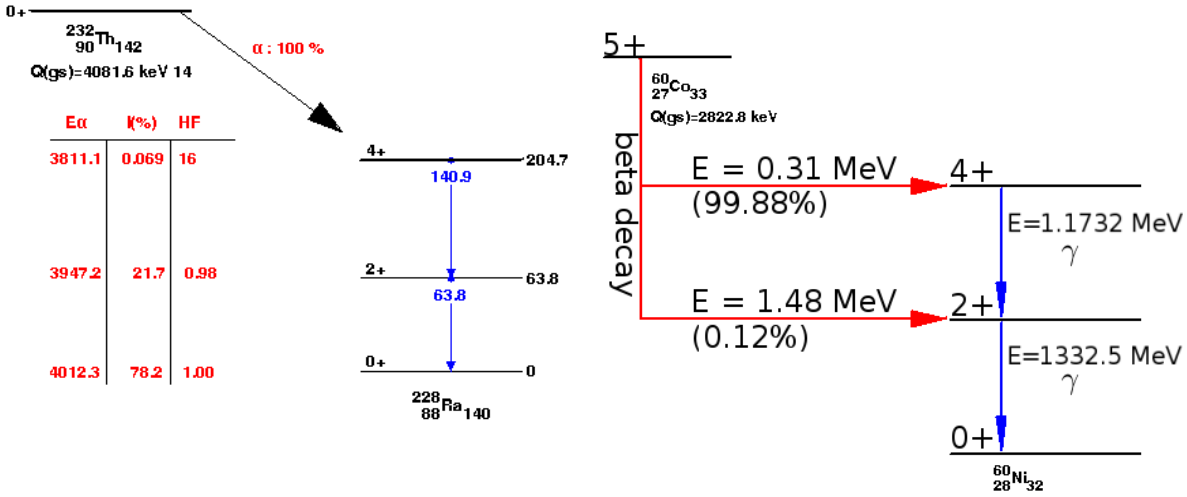


Figure 3.23 – The total time of data taking period 2013/14 for ^{60}Co and ^{232}Th data taking.

the detector response. The goal of section 6 is to present such a model which describes the observed ER background based on the calibration data in the data taking periods of 2011/12 and 2013/14.



(a) **(b)**

Figure 3.24 – The ^{232}Th decay scheme: An α-decay results $^{228}_{88}Ra$. The decay includes two prompt γ particles with typical energies of 140 keV and 63 keV. The ^{60}Co decay scheme: The β^- decay with the largest branching ratio leads to an excited ^{60}Ni by an electron release of $E^{\max} = 0.31$ MeV. The sub argument de-excitation to the ground state yields two gamma particles of 1.17 MeV and 1.33 MeV. Both, the gamma particles from ^{232}Th as well as from ^{60}Co interact with xenon through Compton scattering, which produces a nearly flat electron recoil spectrum at low energies. [88]

3.6.3. Nuclear Recoil Calibration with ^{241}AmBe

The Americium-Beryllium source ^{241}AmBe emits neutrons which are used for the detector calibration. These calibrations are done at the begin and the end of a science data taking period. The neutron source itself is used to determine the acceptance to neutrons and to perform a data/Monte Carlo matching to show a good understanding of the TPC down to 2 pe [23].
In the following, it is focused on the calibration with neutrons from ^{241}AmBe. The neutron is created from two following processes, first in Americium and then in Beryllium. The decay is described in equation (3.30) for the Americium isotope which undergoes a decay and emits an α-particle. This α-particle interacts with Beryllium and results a free neutron (equation (3.31)). The neutrons penetrate the liquid xenon in the TCP and activate the xenon by fast neutron interactions. The resulting nuclear

recoil is measured then by the S1 and S2 signal in the TPC. The created γ particles in equation (3.30) and (3.31) are shielded by lead bricks and only neutrons enter the TPC volume.

$$^{241}_{95}\mathrm{AmBe} \rightarrow ^{237}_{93}\mathrm{Np} + ^{4}_{2}\mathrm{He} + \gamma \tag{3.30}$$

$$^{9}_{4}\mathrm{Be} + ^{4}_{2}\mathrm{He} \rightarrow ^{12}_{6}\mathrm{Co} + ^{1}_{0}\mathrm{n} + \gamma \tag{3.31}$$

The calibration results are summarised in figure 3.27a where the event distribution in the x/y plane of the detector is shown. The excess of events at $x = -150\,\mathrm{mm}$ and $y = 0\,\mathrm{mm}$ indicates the source position outside the TPC. In figure 3.27b three populations can be seen which are generated by the ^{241}AmBe calibration. The "islands" are the result of the fast activated xenon [89]. In particular the two islands at position ($S1 = 120\,\mathrm{pe}|\mathrm{Log}_{10}(S2/S1) = 2$) and ($S1 = 220\,\mathrm{pe}|\mathrm{Log}_{10}(S2/S1) = \sim 2.1$) belong to the activation lines of 40 keV and 80 keV which are used in XENON100 e.g. to develop an S2 correction map (see section 3.4.4). Higher activation lines also leak into the figure too [89]. Figure 3.25 and figure 3.26 summarise the three data taking periods of ^{241}AmBe during the past years, in total roughly 12 days. The created meta-stable isotopes of xenon are used for calibration. Gamma particles are created from fast activated xenon during the ^{241}AmBe calibration. It is possible use these for the light yield determination or the as a spatially uniform gamma source (e.g. 40 keV line). Last is also used to estimate the electron lifetime.
The isotopes with the longest half-life are ^{129m}Xe ($T_{1/2} = \sim 8.9\,\mathrm{days}$) and ^{131m}Xe ($T_{1/2} = \sim 11.84\,\mathrm{days}$). These isotopes introduce at the same time an unwanted background during the Dark Matter data taking. Therefore it is necessary to wait for, at least twelve days before a Dark Matter data taking period starts.

3.6.4. PMT Calibration with LEDs

The weekly PMT calibration is done with two blue LEDs (InGaN with $\lambda = 470\,\mathrm{nm}$). The LED light is fed in the TPC by several fibres which end at the PMT pattern. The positions of the fibre ends are marked by a small dot in figure 3.28. This allows a homogeneous light distribution in the TPC to calibrate the top, bottom and veto PMT section at the same time. The raw data are taken without Zero Length Encoding (ZLE). The resulting waveforms of the raw data are 512 samples long with the usual sample length of 10 ns. The LED is driven by a pulse generator which also triggers the DAQ. An example for such a waveform is given in figure 3.29a. The baseline of the peak is determined by the DAQ from the 50 samples before the registered excess on the waveform. The noise (dark current or electronic noise) around the baseline is not suppressed by the ZLE. Each pulse is then caused by a given quantity of photons which are detected by the PMTs. The triggered excess is roughly in the

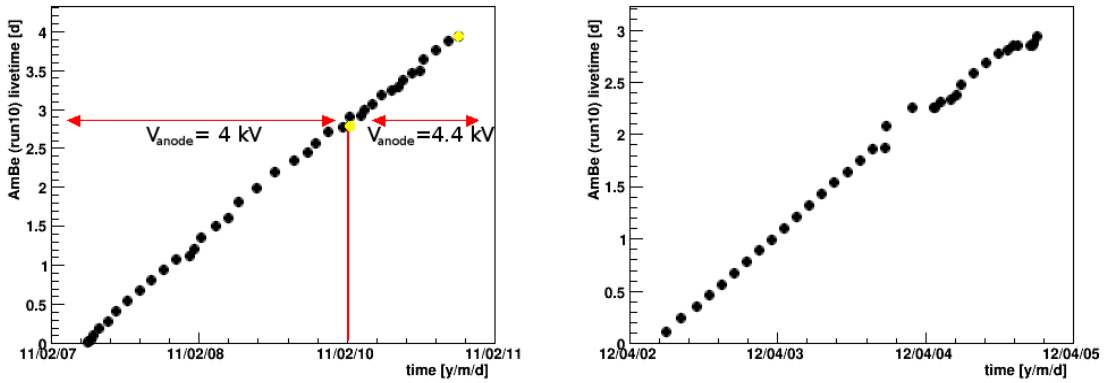


Figure 3.25 – The total time of ^{241}AmBe data taking in 2011/12. **Left:** ^{241}AmBe before Dark Matter data taking period [115]. This sample refers to a low-noise period with different anode voltages of +4.0 kV and +4.4 kV. The two yellow marker indicate a voltage of +4.25 kV. **Right:** The neutron calibration at the end of 2011/12. The applied anode voltage is +4.4 kV. The noise level increases after 2011/12.

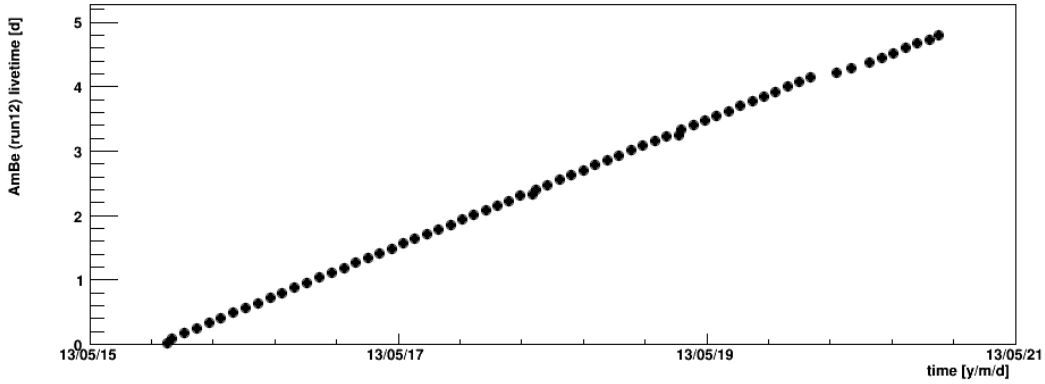
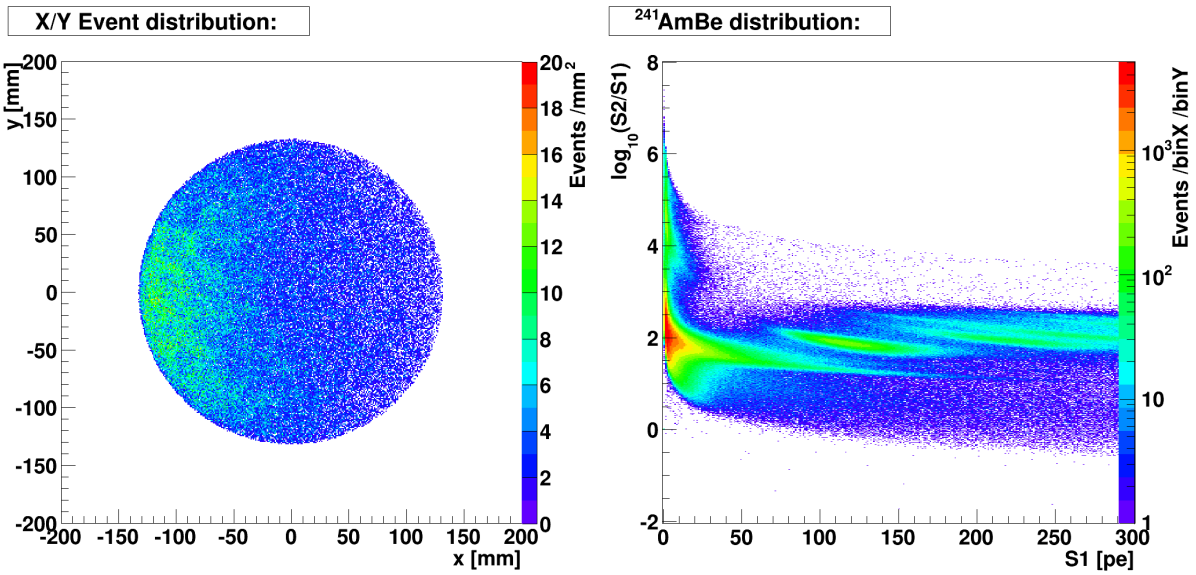


Figure 3.26 – Another neutron calibration before data taking period 2013/14. The data is taken with an anode voltage of +4.4 kV.



(a) **(b)**

Figure 3.27 – Figure 3.27a shows the x/y distribution of ^{241}AmBe in the TPC. The data set refers to 2013/14 data which was taken in May 2013. The event excess indicates the source position during the calibration. The data are presented in log(S2/S1) vs S1 discrimination space in figure 3.27b. Here the specific xenon activation lines from fast neutrons are given [89] by two ellipses at $S1 \approx 120\,\text{pe}$ and $S1 \approx 220\,\text{pe}$.

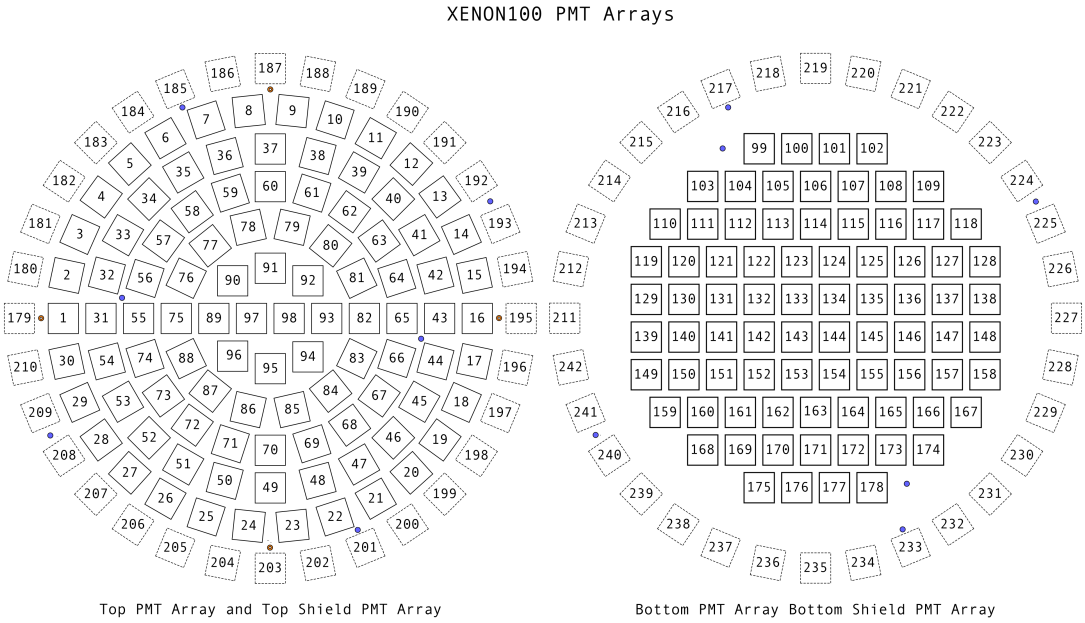


Figure 3.28 – The PMT pattern in XENON100 with its 242 PMTs [115]. The fibre ends where the light of the LED enters the detector is marked by the points with a number beside it. The LED light is split that bottom, top and veto pmt array is covered during the calibration.

middle of the taken waveform around sample 250. The peak marker in figure 3.29a comes from the Xerawdp peak-finder algorithm which is not used in the standard PMT calibration. The peak marker is from a pre-test of the Xerawdp peak-finder algorithm on LED waveforms. A detailed study is given in chapter 4. It is already shown in figure 3.29a that the standard peak-finder algorithm will not work good enough on the summed waveform of a LED calibration data set.

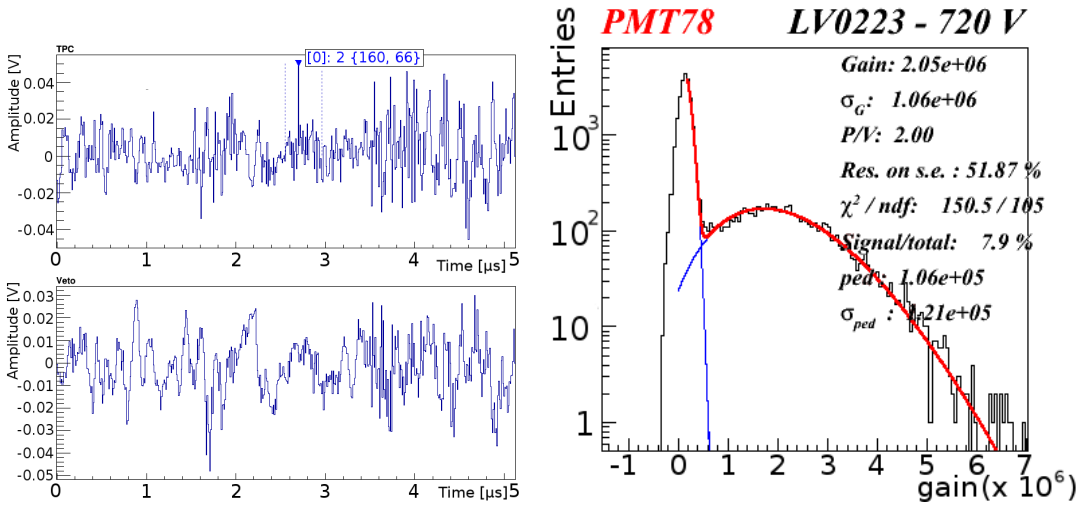
A PMT measures single incoming photons and therefore a careful calibration is necessary. The standard analysis of raw data files is done by another program which looks for single photo-electron (SPE) peaks on the waveform which trigger the sample and determine the gain by a fit to the SPE. The gain is calculated according to equation (3.32). Here the input from the measurement is μ which is the integral of the recorded channels in the digitiser after the amplification of the signal. The digitisation is done by an Analogue-Digital-Converter (ADC) with a given voltage divider. In XENON100, the ADC has a 14 bit resolution, with a given input signal voltage of $V_s = 2.25\,\mathrm{V}$ and hence $r = 2.25/2^{14}\,\mathrm{V/Channel}$. The amplification factor from the Phillips 776 amplifiers $A = 10$ and the sampling frequency in XENON100 is $f = 10^8\,1/\mathrm{s}$. The electron charge is given by $1.62 \times 10^{-19}\,\mathrm{C}$ and the impedance is $Z = 50\,\Omega$ [72].

$$\text{gain} = \frac{\mu\, r}{Z A f e^-} \tag{3.32}$$

Further general information about Hamamatsu PMTs are found in [63]. An example for such a gain calibration is shown in figure 3.29b. Here the gain is measured for a LED calibration at the end of January in 2015 for the PMT 78. The red line indicates the overall fit describing the noise peak on the left side at gain = 0 and also the SPE which peaks around gain $\sim 2.05 \times 10^6$. Gain deviations of $\sim$ 10% are observed [115]. A Gaussian fit function describes the noise part of the SPE spectrum and the SPE peak itself is also fitted by a continuous distribution which is defined in equation (3.33) where $x ==$ gain [115].

$$y(x) = \frac{\mu^x \exp(-\mu)}{\Gamma(x+1)} \tag{3.33}$$

In this work, we develop another method of gain measurement which is implemented in the usual raw data processor (*xerawdp*). It searches for single photo electrons on single PMTs which exploits the peak finding performance of the peak-finder algorithm on LED raw data (see chapter 4 at page 111).



(a) **(b)**

Figure 3.29 – Figure 3.29a shows a LED raw data waveform. The upper part is the sum of the top and bottom PMTs and the lower part consists of the sum of the veto PMTs. The detected peak here comes from the Xerawdp peak-finder software and not from the usually used LED raw data. Figure 3.29b shows an example of the standard PMT calibration for PMT 78. The gain determined to estimate the PMT response to photons.

3.7. Event Selection in XENON100

The most important step during data analysis is a proper pre-selection of the data. XENON100 uses a blinded data analysis, thus the Dark Matter data are recorded, stored and reprocessed with proper information such as electron lifetime, S1 and S2 corrections. A dedicated pre-selection of physically motivated data is prepared by a few non-blinded Dark Matter data sets (outside the region of interest) and calibration data. The following section summarises the cut sets of the past science runs (2011/12 and 2013/14) and explains their meaning. The cuts used later during the data analysis are mentioned at the begin of the according subsection. Detailed information about the cut development are not discussed in this work. The cuts itself are defined by the calibration data and further background predictions are made from the few non-blinded Dark Matter data sets.
During the XENON100 experiment the run conditions changes and therefore the cuts must be adjusted over the time. This leads to different successive versions of the same cut later. The naming convention for the developed cuts is simple: The cut name starts with an X and ends with the version number. In case the cut refers to S1 or S2 only its also added after the X.

The blinded data sets are finally un-blinded and the cut set is applied without further adjustment. The following list explains the most applied cuts in XENON100 during data analysis regarding the later Dark Matter data analysis. The version number is replaced by a V in the list.

- **Xs1coinV:** The first S1 peak is in coincidence with more than one PMT. This avoids single photo-electron signals which are not Dark Matter signal but e.g. dark current from the PMT photo-cathode. This two-fold coincidence is required in every science run for data selection.

- **Xs2singleV:** Select only events where the second (and following) S2 peak(s) are much smaller than the first. This results in a single large S2 peak on the raw data waveform. This cut is used to select clean waveforms without further possible events in the same time window. The cut does not change from 2011/12 to 2013/14 and is used for Dark Matter and calibration data sets.

- **Xs1singleV:** The first S1 peak is in coincidence with more than one PMT (see Xs1coinV) but the second and further S1 peaks are not in coincidence with any other PMTs. This assures a clean selection of S1 signals which are widly spread in the TPC and not just randomly seen by single PMTs (e.g. single photo-electrons or dark current). The cut is designed for high acceptance and take the drift time correction into account in 2011/12. In 2013/14 the measured entropy of the S1 peaks is also taken into account to remove noisy S1 peaks before the coincidence condition is applied.

- **Xs2widthV:** There exists a relation between the typical S2 width and the interaction depth. The cut removes events which are produced in the gaseous phase of the detector with an unusual S2 width. The cut changes from 2011/12 to 2013/14 due to fine tuning.

- **XPL013_97:** The S1 PMT Pattern Likelihood cut is a light pattern estimator to assure a clustered signal. The likelihood is calculated by the 178 PMTs of the top and bottom array. It is designed to reduce so called γ-X events in the TPC. The cut does not alter during 2011/12 and 2013/14. The cut is not altered from 2011/12 to 2013/14.

- **XposrecV:** The position reconstruction works with three independent algorithms. By default a similar result is expected. Events which differ to much in the position reconstruction are sorted out.

- **Xs2chisquareV:** The χ^2 of the Neural-Network (NN) position reconstruction algorithm is normalised by the number of coincidences on the top PMT array. The resulting dimensionless variable is used to suppress double scatters by neutrons from radioactive background in the Dark Matter data. The cut is not altered from 2011/12 to 2013/14.

- **XlownoiseV_m:** The cS2 signal is split into the detected contributions from the top and bottom array. Good events require a proper asymmetry between the top and bottom array. The cut is not altered between 2011/12 and 2013/14 and the index *m* sets an extension on the asymmetry to the previous cut definition.

- **Xn3sigma:** The cut is used to remove events out of the nuclear recoil band from ^{241}AmBe to avoid unexpected events in the later Dark Matter search region.

- **XvetoV:** A S1 signal from the veto PMTs needs to be lower than 0.35 pe to pass events and distinguish so from events from outside which trigger a veto PMT. No change during 2011/12 and 2013/14.

- **Xs2peakposV:** The trigger of the S2 peaks leads to data acquisition of 40 µs in total. A proper S2 trigger should happen then roughly in the middle of the recorded waveform. The cut reject events which have a S2 peak within the first 17.8 µs of the waveform. The cut is very basic and does not change during the science runs.

- **Xs2peaksV:** A very basic cut on the energy threshold in S2 (150 pe). (No change from 2011/12 to 2013/14)

- **Xs2topV:** The cut is only used in 2011/12 to remove an occurring hot-spot (un-reasonable clustering of signals in x/y) at the PMT pattern in the past.

- **XentropyV:** The entropy is calculated for each S1 peak which is identified with the peak-finder of Xerawdp. Some PMTs show a higher noise probability and therefore the entropy of S1 peaks from these PMTs is used to reject noise S1 peaks.

- **Xs1widthV:** A typical S1 signal is narrow about 10-60 ns and a minimum width is demanded. The cut is not altered between 2011/12 and 2013/14.

- **X34kgV:** The spatial dependence of events is very necessary in the later Dark Matter data analysis. As long as liquid xenon is used as detection material and capitalise the self-shielding probabilities it is important to select events only from the inner detector. Therefore a elliptical cut is defined to increase the fiducial volume inside the TPC as much as possible. The cut is robust and does not change during 2011/12 and 2013/14 in the Dark Matter data analysis. Further analysis types require an extension to the standard fiducial volume.

- **XhighlogV:** The cut removes a few events with a low value in S1 but an usual high S2 value.

- **XsignalnoiseV:** The cut compares the measured S1 and S2 signal with the total area of the waveform. Hence a kind of signal-to-noise ratio is defined to reject noisy events. The cut changed from 2011/12 to 2013/14 and is also tuned to for low-noise condition.

3.8. The Nuclear Recoil Background in XENON100

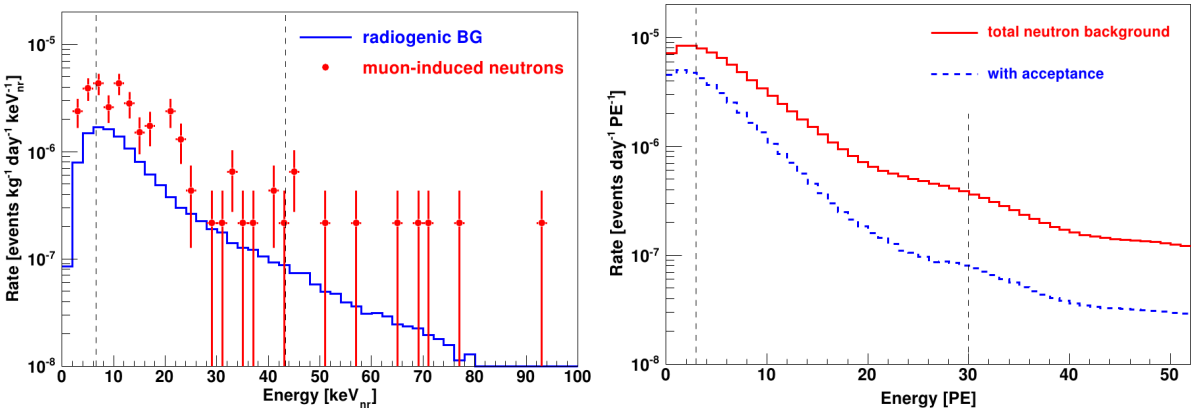
This subsection will introduce the nuclear recoil as additional background in the XENON100 experiment which can not be neglected. Chapter 6 discusses the ER background and introduces a phenomenological model. A complete description of the background needs to take ER and NR background into account.
The NR background is determined by a Monte Carlo simulation[25]. This simulation takes different neutron sources into account:

- Neutron scattering of the xenon nuclei and generate a certain WIMP-like signal.

- Due to intrinsic contamination in the detector materials, additional radiogenic neutrons ($\sim 1\,\mathrm{MeV}$) are produced by (α, n) reactions and spontaneous fission.

- Cosmic myon interactions in the surrounding rock of the cavern introduce a cosmogenic neutron background.

The cosmogenic neutron background depends on the depth of the experiments and is simulated with Geant4 [42] in the experimental setup. The according muon energy and angular distributions are generated with MUSIC and MUSUN [78]. The natural neutron background is simulated with the SOURCES-4A tool [29]. There is a large set of target nucleus and neutron sources defined in the SOURCES-4A code. This includes the chemical composition of the cyrostat, TPC, PMT parts, shield and the surrounding environment.
In both cases, single and multiple scatters are investigated in different detection volumes. The obtained the signal spectra, $\mathscr{L}_{\text{eff}}$ and $\mathscr{Q}_{\text{y}}$ is used. This allows to calculate the total neutron background rate in photo-electrons. Furthermore, the overall detector acceptance corrects the total neutron background. This is shown in figure 3.30 for the neutron background simulation for the data taking period 2011/12. Figure 3.30a shows the radiogenic and cosmogenic components of the neutron rate in units of keV. The converted spectra is shown in figure 3.30b.
The simulation predicts 0.14 ± 0.02 events from radiogenic neutrons and $0.34^{0.34}_{-0.17}$ events from cosmogenic neutrons background, in total a nuclear recoil background of $0.48^{0.34}_{-0.17}$ events in the data taking period 2011/12 for the 34kg fiducial volume (see 3.7) and an energy range of 3 pe to 30 pe. This number is reduced after the benchmark region for the WIMP search is defined to $0.17^{+0.12}_{-0.07}$.



(a) **(b)**

Figure 3.30 – The simulation results a neutron rate which depends on the neutron energy. The result of the Monte Carlo simulation is shown in figure 3.30a for cosmogenic and radiogenic introduces neutrons. The relative scintillation efficiency $\mathscr{L}_{\text{eff}}$ transforms it into photo-electrons. In addition, the overall detector acceptance reduces the rate in figure 3.30b. [25]

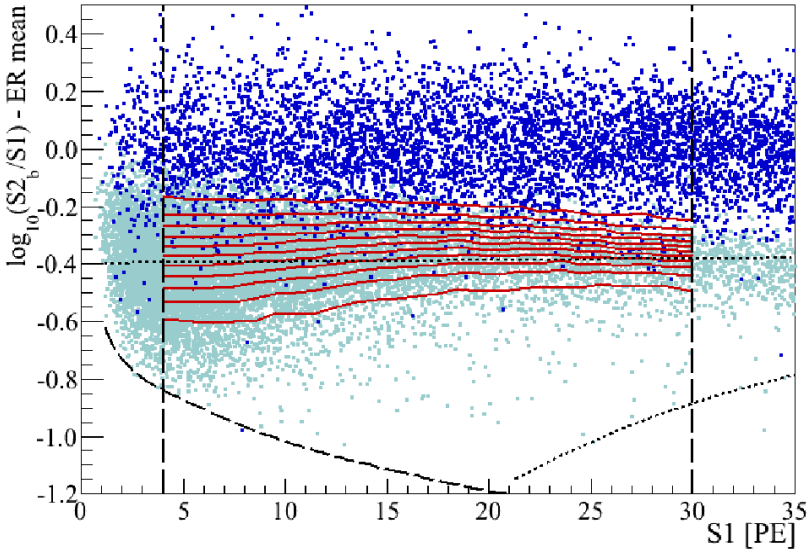


Figure 3.31 – There are two populations shown: In blue the ER background events and in cyan the NR background events. Based on vertical dashed line (3σ standard deviation from ER background), the bottom right dashed line (3σ standard deviation from NR background) and the bottom left dashed line from the S2 threshold cut, a region of interest is defined for the WIMP search. The PL analysis is not limited to the vertical left and right boundaries [15].

3.9. Discrimination Space Y in XENON100 and the Region of Interest

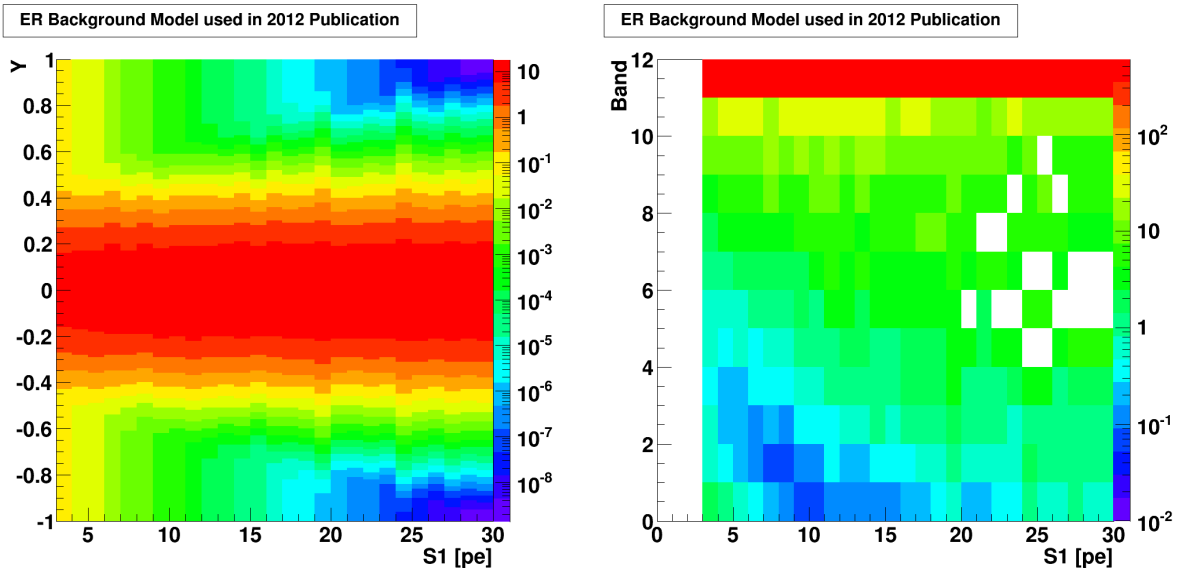
The ER background is characterised by a Gaussian distribution of the abstract variable Y (discrimination parameter) for each slice of $cS1$. This approximation is valid because of the observed shape of the ER background, but there is no physical reason which leads to that assumption. The Gaussian approximation results in an expectation value of $<ER>(cS1)$ which depends on $cS1$ and is understood as constant offset. This leads to the given definition of Y in equation (3.34) which is also named *flattened space* [15]. It includes the corrected S2 signal which is observed on the bottom PMT array ($cS2_b$) and the corrected S1 signal ($cS1$). The advantage of this definition is the flattened ER background where the expectation value varies only a little around $Y \sim 0$. Equation (3.34) allows also a coordinate transform: $cS1/Y \leftrightarrow cS1/cS2$. The reason why Y is defined by $cS2_b$ comes from TPC itself. The S2 signal is created in the gas phase and its scintillation light is spread over the whole TPC down to the bottom PMT array. From there a valid S2 is registered. The discrimination space is given in figure (3.34) where the electronic recoil data are shown in blue and the nuclear recoil data in cyan. Based on a 3σ lower limit for the observed ER background events and for NR events and the S2 signal threshold, a region of interest is defined (e.g. for low mass WIMP searches). The brownish band structure is furthermore discussed in section 3.10 and 3.11, where the band approach is introduced and the Profile

Likelihood (PL) analysis is explained. This statistical method tests the whole XENON100 data space with the Profile Likelihood approach is not limited to the region of interest.

$$Y = \log_{10}\left(\frac{cS2_b}{cS1}\right) - <ER> \tag{3.34}$$

3.10. The Band Approach in XENON100

The method of choice to analyse the Dark Matter data, the Profile Likelihood (PL) analysis, is described in a publication from 2011 [19] (see also section 3.11). The PL needs a probability density function which characterise the background in the XENON100 experiment (f_b). The background consist of the ER and NR background at the same time. This thesis gives a detailed summary on the ER background and introduces a phenomenological model of the ER background. The NR background is summarised in section 3.8.
The PL analysis is based on a binned data space. The usual data space, the flattened space, is introduced in section 3.9. This data space is adjusted by a re-binning into an equal number of bands in the data space. This band approach describes a transform from the parameter space Y and S1 into twelve almost horizontal bands which depend on S1. The bands are defined from NR data (^{241}AmBe) where the data are divided into twelve bands under the condition that each band has the same number of events. A detailed explanation is given in a past XENON100 publication [19] and [15]. Figure 3.31 shows the twelve bands (brown) and the nuclear recoil data in cyan. In addition, the electric recoil band is shown in blue. An example for the band transform is given in figure 3.32. The official ER background model of the data taking period 2011/12 is shown in the flattened space in figure 3.32a. After the transform is carried out, the events are re-distributed into the twelve bands in figure 3.32b. Due to band definition most of the events are transformed into band 12. The leftover eleven bands contain the events from figure 3.32a below a value of $Y < -\sim 0.2$. (This value is a only a rough estimation and the true band definition is more complicated). Hence it is shown that the lower tail of the Gaussian distributed ER background is described by eleven bands. This points also to the region of interest for the WIMP search (e.g. low mass WIMPs) which is located in the the first to fifth bin.



(a) (b)

Figure 3.32 – The ER background model of 2011/12 is given in figure 3.32a. Here the Gaussian approximation is done similar to the method described in section 6.4.2. The leakage events are described by an exponential function which is added after all to the Gaussian ER background. To use the ER background in the actual PL data analysis, a transform from the flattened data to the band space is done.

3.11. Profile Likelihood Analysis

Once the acceptances are calculated for the different cuts and the background rate is determined for the region of interest, the un-blinding process is carried out. Firstly, a set of side-bands is defined which are un-blinded to predict e.g. the expected number of events for the given cuts and acceptances. Secondly, after the full data are un-blinded. The un-blinded events are analysed based on the predefined cuts in discrimination space and the fiducial volume.

This step is used to visually identify potential signal events. This method is only a first step to classify the possible signals. A statistical model is necessary to test the signal hypothesis against the measured background events (N_b). The test is carried out in the pre-defined band space (see section 3.10). The bands are indexed by j. The XENON100 collaboration uses a Profile Likelihood approach to test signals with a certain signal probability ε_s^j and the background with a certain background probability ε_b^j. The signal probability is calculated from ^{241}AmBe calibration signal in 2011/12. This changes in the data taking period 2013/14 where the signal is simulated with a neutron simulation. Due to effectively different energy thresholds in S1 and S2, the distribution of the signal probability density function significantly deviates from the ^{241}AmBe nuclear recoils for WIMP masses below $\sim 20\,\mathrm{GeV}$. Advantages in our understanding of the low energy scintillation and charge yield of nuclear recoils

enable us now to model the signal probability density function from Monte Carlo with smaller systematic uncertainty than the direct use of the ^{241}AmBe data. The background probability results from ER background data. In general the probabilities are calculated in each band. This band approach divides the data before the Profile Likelihood analysis is carried out of equal signal strength. The bands are shown in figure 3.31, where they overlap with the cyan coloured ^{241}AmBe data. The Likelihood function $\mathscr{L}$ is given in equation 3.35.

$$\mathscr{L} = \mathscr{L}_1(\sigma_\chi, N_b, \varepsilon_s, \varepsilon_b, \mathrm{L_{eff}}, v_{esc}, m_\chi) \times \mathscr{L}_2(\varepsilon_s) \times \mathscr{L}_3(\varepsilon_b) \times \mathscr{L}_4(\mathrm{L_{eff}}) \times \mathscr{L}_5(v_{esc}) \tag{3.35}$$

The likelihood function is defined for a given WIMP mass m_χ and cross-section σ_χ. Additional parameters, so-called *nuisance parameters* are *profiled out*. This means to minimise the likelihood for a given set of nuisance parameters, for example the parametrisation of $\mathscr{L}_{\mathrm{eff}}$. This method tests the signal hypothesis against the background only hypothesis for a set of possible WIMP masses. This method also allows to calculate an exclusion limit in case that there is no statistical evidence for a WIMP detection.
The most important part of equation 3.35 is $\mathscr{L}_1$ which is given as:

$$\mathscr{L}_1 = \prod_{j=1}^{K} Pois(n^j | \varepsilon_s^j N_s + \varepsilon_b^j N_b) \times \prod_{i=1}^{n^j} \frac{\varepsilon_s^j N_s f_s(S1) + \varepsilon_b^j N_b f_b(S1)}{\varepsilon_s^j N_s + \varepsilon_b^j N_b} \tag{3.36}$$

The definition tests the Poisson probability for each band with the given total number of signal and background events (N_s and N_b). Furthermore each event, within a band j, is taking the normalised WIMP spectrum f_s into account and is tested against the electron recoil and nuclear recoil background spectrum f_b. The electronic recoil background is shown in figure 3.31 by the blue dots. The band definition shows that almost all of the ER background events are distributed in the highest band $j = 12$ and only few ER events are in the region of interest (bands $j \in [1,5]$). This is indicated by the vertical dashed line which is defined according to a 99.75% two-sided boundary of the ER background under a Gaussian assumption. Terms $\mathscr{L}_2$ and $\mathscr{L}_3$ take calibration measurements as control measurements into account (penalty terms) and $\mathscr{L}_4$ parameterises the uncertainty on $\mathscr{L}_{\mathrm{eff}}$. Uncertainties in the halo velocity distribution are not included. Details are discussed in references [15] and [19].
In the science data set of 2011/12 two events were observed in the region of interest, consistent with the background hypothesis [17]. The PL analysis allows to calculate an exclusion limit for the unknown parameters σ_χ and m_χ. The result is shown in figure 3.33 [17]. The limit shows its minimum at $\sigma_\chi = 2 \times 10^{-45}\,\mathrm{cm}^2$ for WIMP mass of $m_\chi = 55\,\mathrm{GeV/c^2}$ with 90% confidence level for spin-independent WIMP nucleon interaction. In 2012, this result improved the world-wide sensitivity over competing experiments by a factor of 20.

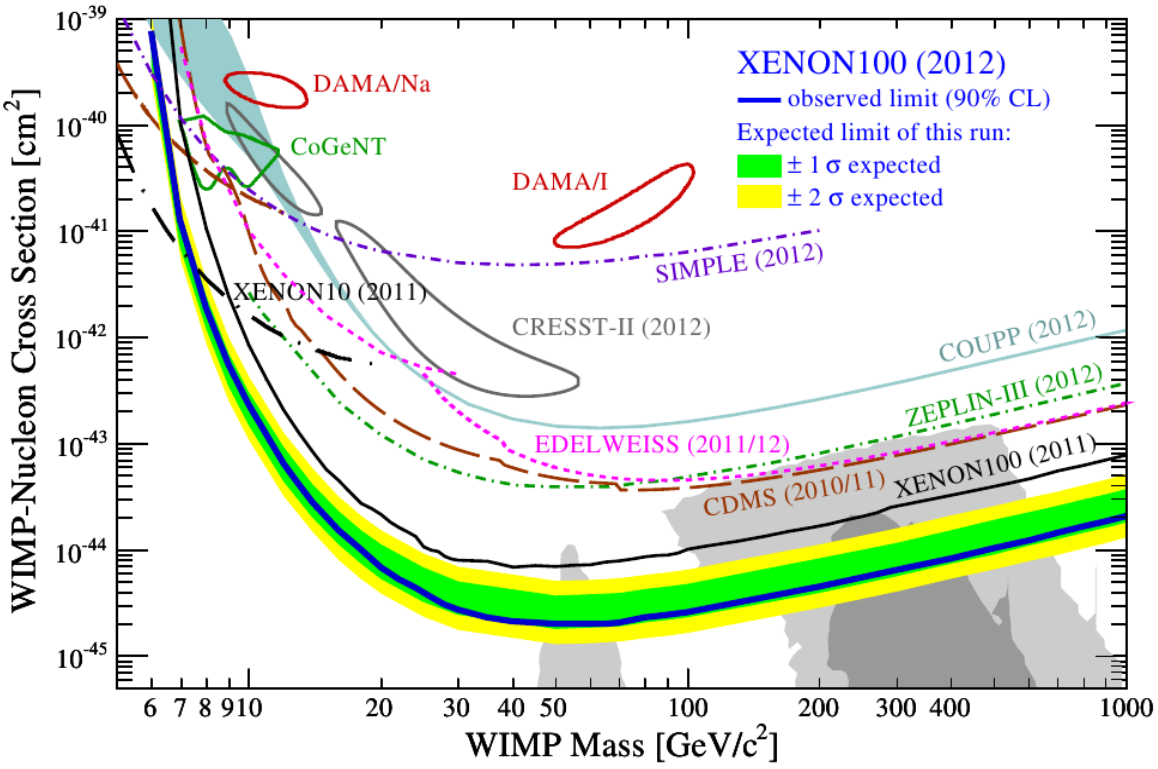


Figure 3.33 – The WIMP-nucleon exclusion limit of the science data taking period in 2011/12. The figure show the lowest exclusion limit in 2012 compared to competitors such as DAMA/Libra [15]. The XENON collaboration showed the lowest level of sensitivity for WIMP search.

4. An Alternative PMT Calibration

"[...] It changes matter into *Dark Matter* and seeks out to host bodies, drawing strength from their life force. Malekith sought to use the Aether's power to return the universe to one of darkness."

Odin, "Thor – The Dark Kingdom"

4.1. Motivation

The calibration with radioactive sources is done to determine the TPC response to certain interaction types and their energies regularly. Another important procedure is the photo-multiplier tube calibration in XENON100. It tests the PMT response to single photons and measures the ability of transfer the information of registered photons on the photo cathode into a signal in units of voltage. The calibration is done weekly with blue LEDs. A detailed review of the PMT calibration is described in section 3.6.4. The obtained LED data is analysed and the gain information is used to determine the single-photon response of each single PMT.
The goal of the standard LED calibration is to measure the PMT gain which is necessary to translate the later measured signal in the detector to its energy (photo-electrons). An example for such a PMT gain is shown in figure 3.29b. Here the single-photon electron peak is centred around gain $= 2.05 \times 10^6$. The left, symmetric part around gain $= 0$ is the PMT noise contribution which comes from dark current.
After the PMT calibration is done, the raw data is analysed by a separate program. This piece of software analyses each single waveform separately and looks only for a certain time frame on the single waveforms, regardless a S1 signal is there or not. The raw data analysis of calibration sources and Dark Matter data is done with another software: *xerawdp*. This program is not able to analyse LED raw data by default. This section presents a method to analyse LED calibration data with *xerawdp*. Therefore it is shown how *xerawdp* is modified, which cuts are necessary to run it on LED data and which variables are used to carry out the LED data analysis. This chapter aims to present an alternative LED calibration with *xerawdp*.
This applications is useful in two terms: First, a second method is available to analyse the LED data sets and check for consistency. Secondly, a further development is the analysis of Dark Matter and calibration data sets in a similar way what allows to check on the PMTs between two LED calibrations.

4.2. Changes in Xerawdp

Xerawdp is not able to read S1 pulses from the LED calibration raw data without modifications. There are two major changes necessary: Firstly, a software zero length encoding (ZLE) must be implemented. This is done according to the software specifications from the data acquisition program (DAQ). Secondly, the *xerawdp* raw data processor works with input scripts which are composed in *xml* language. These scripts depend on the type of acquired data: Dark Matter or calibration data. A new *xml*-file is created where the S1 peak-finding guidelines are tuned to increase the S1 peak identification

rate.

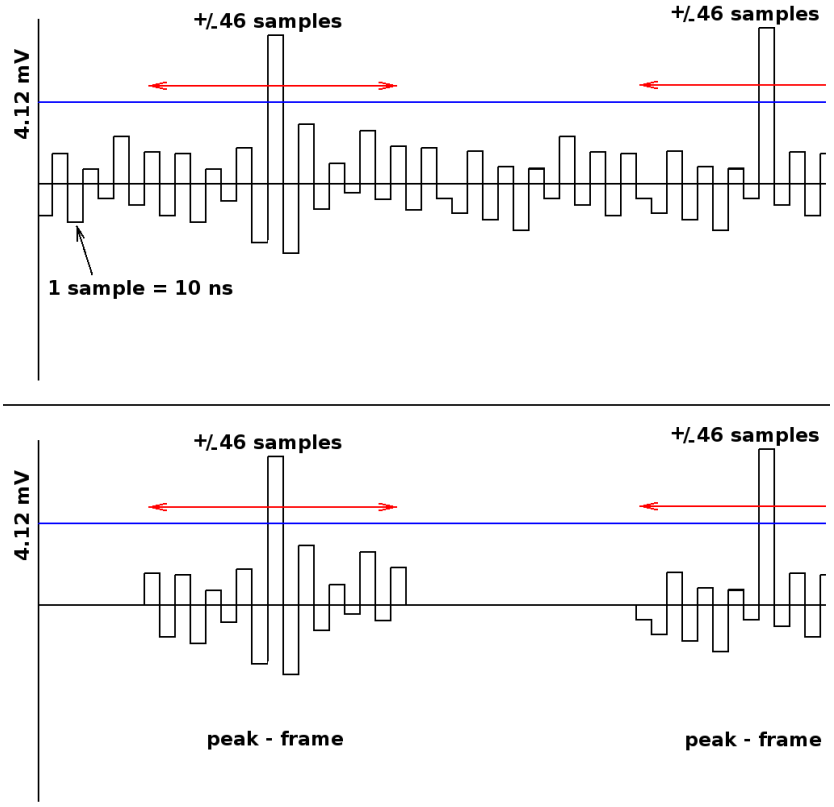


Figure 4.1 – The upper part illustrates the taken data before the ZLE is applied on a single waveform. The DAQ searches for excesses above $U_{ex} = 4.12 \times 10^{-3}$ V and stores a time window of ± 46 samples around this excess. The result is shown in the lower part, here only the excess information is stored. The according time information on the waveform which defines the positions and lengths on the PMT waveform is enough to restore a complete waveform which contains only the important information. These waveforms are analysed by *xerawdp* in case of calibration and Dark Matter data.

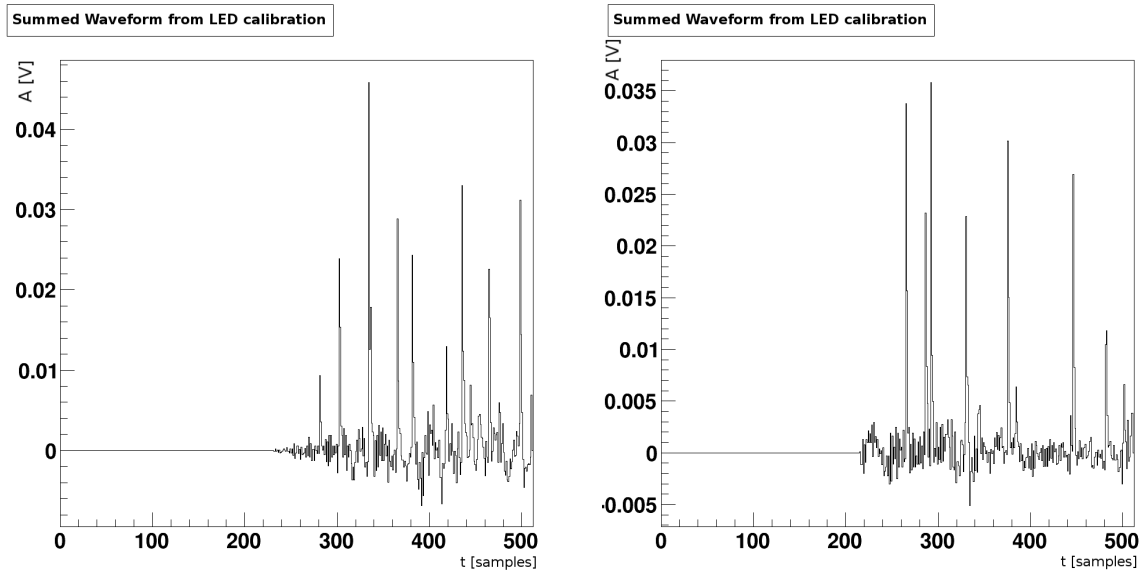
The XENON100 data acquisition program (DAQ) is taking LED data without zero length encoding (see section 3.6.4). But *xerawdp* is not able to find any S1 peaks in the summed waveform given the current implementation of the peak-finding algorithm. This is the reason why another program analyses each single waveform separately. The advantage of taking LED data without ZLE is that there is no information loss.

The ZLE is implemented in the DAQ and is enabled if e.g. Dark Matter data is taken. Zero length encoding is a signal pre-selection which searches for peaks on single waveforms. In case, the ADC registers an excess on the waveform which exceeds 30 channels, which is equal to a voltage of

$U_{ex} = 4.12 \times 10^{-3}$ V, a time window of ± 46 samples around the excess is stored by the DAQ (specification of the DAQ program for data taking) with its time frame information. As a consequence, the samples between two of such excesses are set to zero and only the distance between both is stored. The LED calibration aims to calibrate each PMT with single photons and hence the expected S1 peaks are small. To exploit the calibration data as good as possible, the LED data is stored without ZLE. The upper part of figure 4.1 illustrates the ZLE for a single PMT waveform of the taken data. Due to the time resolution, a single sample is 10 ns broad and varies around a pre- selected baseline. Together with the time frame information about the possible peaks, a waveform is defined with only the measured excesses. These waveforms are analysed by *xerawdp* for S1 and S2 signals by the peak-finding algorithm. Therefore not every excess, which is stored by the DAQ, is part of a peak. This depends strongly on the peak-finding algorithm and the input parameters which defines a S1 and S2 peak (input from *xml*-file).
In case of the LED data taking, there is no ZLE used and the pulse generator of the LED calibration setup is used to trigger the DAQ. Each trigger stores the 242 single waveforms to a raw data file. The stored LED waveforms in the raw data are 512 samples ($t_{wf} = 5.12\,\mu$s) long (1 sample = 10 ns). This is much shorter, compared to the length of the waveforms which are stored in case of Dark Matter data taking. A shorter LED waveform is enough because of the expected transit time of the photons in the TPC after the LED trigger. After the pulse generator triggers the LEDs and the data acquisition system, the single photo-electron peak is expected between 2000 ns and 3000 ns. Due to the reflection properties of the PTFE, the transit time increases and some late photons are recorded too.

The peak-finder algorithm works on the summed waveform: Here the 178 PMT waveforms from the top and bottom array are summed up and then the peaks are identified. An example of such a waveform is given in figure 3.29a. Due to the missing ZLE the summed waveform looks noisy for LED data. The chosen example contains an identified peak which is found by *xerawdp* without any modifications. By analysing the whole waveform and more LED calibration waveforms, there is no reason why this should be an S1 peak. Therefore the software zero length encoding must be implemented into *xerawdp* in case of LED raw data analysis. The *xerawdp* software code is altered in the "AnalysisMananger". Here each set of 242 waveforms is altered by the software zero length encoding before the modified peak-finder algorithm is applied.
To archive that changes, a new *Waveform*-object is created within the AnalysisMananger of *xerawdp*. The *Waveform* object stores the necessary data for the later peak-finding algorithm in case the ZLE encoding is activated: This means it contains the sampling frequency, the waveform length, the number of PMTs and a list of possible peaks (excesses during the data taking period). Each possible peak is assigned to the PMT on which the excess is observed. This *Waveform*-object will hold the LED waveform after the software zero length encoding is applied. This is done in a simple way at first:

An additional member function in the AnalysisMananger tests each bin on the single waveforms if its content exceeds the threshold. In case that condition is true a time window of 46 samples around the excess is written to the *Waveform*-object. This is done bin by bin and waveform by waveform. The more complicated part of this procedure comes from baseline adjustment which is necessary for LED raw data. Therefore each possible sample which shows an excess is analysed due to baseline which is defined right now before *xerawdp* is started. A peak based baseline is determined by the average of 46 samples before the excess. In case the excess happens within the first 46 samples of the LED waveform, the baseline is calculated from the the average of the following 46 samples. This method allows to determine the baseline for each excess separately. The saved information is then ready to be analysed by the adjusted peak-finder algorithm in *xerawdp*.



(a) (b)

Figure 4.2 – Figure 4.2a and 4.2b show two waveform examples of LED calibration data. After the *software zero length encoding* is applied, the sum of the top and bottom array PMT waveforms show the typical S1-like structures of SPE peaks. These peaks are supposed to emerge from the LED calibration and present single photo-electrons. Due to the PMT calibration setup the peak spectrum begins between 2000 ns and 3000 ns and potential successive S1 peaks.

The peak-finder algorithm is explained in section 3.5.3.2. The usual way of identifying peaks in *xerawdp* is to search for S2 peaks first and then localise the much narrower and smaller S1 peaks. During the LED calibration there are no S2 peaks expected and then the peak-finder algorithm is adjusted first of all by disabling the S2 peak search method. In this case *xerawdp* tries to find S2 peaks on the LED waveform, it addresses the necessary computer memory before, even if no S2 peaks are found. In the following, the S1 peak-finder algorithm runs in the same condition as if no S2

had been found. This procedure allows to modify the *C++* code of *xerawdp* very quickly without spending to much time on the S2 peak-finder code. The identified S1 peaks subsequently sorted then by coincidence. The S1 peak-finder algorithm is not altered furthermore and runs according to the input from the *xml*-file for the LED reprocessing. This *xml*-file holds the necessary information for the S1 peak-finder, for example a condition about the minimum distance between two identified S1 peaks.

4.3. *XML*-File for LED Data Reprocessing

After the *C++* code is altered and compiled it is necessary to create a new *xml*-file which is designed to work with LED raw data input only. The *xml*-file contains the necessary information for the (S1) peak-finder algorithm to identify S1 and S2 peaks. For example, the numbers which are cited in subsection 3.5.3.2 are stored in such a *xml*-file. A new LED-*xml*-file adjusts the S1 peak-finder conditions and increases the S1 trigger efficiency on the shorter LED waveforms. According to the used pulse generator settings, the S1 peaks are expected between 2000 ns and 3000 ns but there is no reason to assume that a S1 peak could not accidentally be located in the range before $2\,\mu$s or after $3\,\mu$s. The following LED data analysis takes all potential found S1 peaks into account. An important change is the time difference of the successive S1 peaks on the LED waveform. Due to the much larger waveform of e.g. Dark Matter data, there is also a larger time difference chosen between two possible S1 peaks. This is useful because the peak-finder algorithm does not allow that the time window of a successive peaks overlap with another peak before. The maximal length of such a peak time window is reduced for the LED data to find more S1 peaks (50 ns). In case the maximal length would be 600 ns such as in a Dark Matter *xml*-file, there could be not more than 8 possible S1 peaks identified. Due to the shortened peak window, the maximal length in where *xerawdp* is searching for a peak is decreased to 50 ns (*peak_window*). In addition, the pre- and post-peak average window (*pre_peak_avg_window*/*post_peak_avg_window*) is also shortened to the narrow S1 peaks.
Table 4.1 compares a standard *xml*-file for Dark Matter data taking with the modified version for LED data. The names for the variables which are given here are taken from the *xml*-file. A copy of the full *xml*-file is given in section A.4. It shows the S1 signal related variables which are set externally by the *xml*-file. The *xml*-file uses also the measured gains from the PMT calibration to quantify the single photo-electron (SPE) response in the here presented analysis result in subsection 4.8. From this information, the S1 signal is converted from voltage to photo-electrons. The analysis of the single detected S1 peaks will relay on this information. But there is another opportunity discussed which does not need that information (see section 4.9). In addition, the peak-finder algorithm does not take signals from PMTs which are known as noisy (e.g. PMTs 1 and 2) into account. This does not meant

that these PMTs are switched off, they are just excluded from the peak-finder algorithm.
With the here given constrains on the S1 peak-finder algorithm, there are at maximum 32 S1 peaks possible to identify on the summed waveform. This number is still taken from the Dark Matter *xml*-file to allow a maximum of identified S1 peaks. Therefore the nomenclature for S1 peaks is extended to single S1 peaks on the waveform which get an additional index according to the criterion of ordering. Therefore they are named S1s[pmt][i] peaks in case a specific single S1 peak is meant or just S1s-peaks in case meaning all single S1 peaks from a summed waveform.

Action	LED *xml*-file	Dark Matter *xml*-file
Gain values:	2011/12 or 2013/14	2011/12 or 2013/14
Exclude PMTs (from peak-finder):	No	yes: 1-2, 145, 148, 157, 171, 177
Signal threshold	0.003 V	0.003 V
peak_window	50 ns	600 ns
max_length	50 ns	600 ns
pre_peak_avg_window	20 ns	500 ns
post_peak_avg_window	20 ns	100 ns

Table 4.1 – The table summarises the necessary changes which are taken into account to create an *xml*-file for reprocessing LED raw data with *xerawdp*. The name of the variables which are used in the table are taken from the *xml*-file directly. Chosen gain values refer to the data taking period.

4.4. Results from the Xerawdp Peak Finder

To compare the result from the peak-finder algorithm figures 4.3a and 4.3b show two example waveforms which are analysed by a standard Dark Matter *xml*-file. The red lines present the S1 peak position and the two green lines around show the left and right S1 peak boundary. Compared to the number of S1-like peaks which are observed in the LED waveforms, the number of found peaks is low. This is due to the large peak window (*peak_window*). To test the peak-finder ability, a standard Dark Matter *xml*-file is used with 100000 LED waveforms. Then the number of reconstructed S1 peaks per waveform is counted. The result is given in figure 4.3c. The histogram shows the number of observed S1 peaks per LED waveform. With that information, a standard Dark Matter *xml*-file is only able to find 1.2 peaks per LED waveform (mean value). The new LED *xml*-file increases the S1 trigger efficiency to a mean of 3.6 identified peaks per waveform (see figure 4.3d). In some few cases, such as shown in figure 4.3a, the right reconstructed peak is identified slightly wrong (independently from the used *xml*-file). It can be seen that the S1 peak position (red) is not exactly where the S1 peak is located. These events are also counted in figures 4.3c and 4.3d. To remove such events in the later

analysis there is an entropy cut defined which reduces the number of miss-reconstructed peaks based on their entropy which is determined for each single peak (see section 4.5).

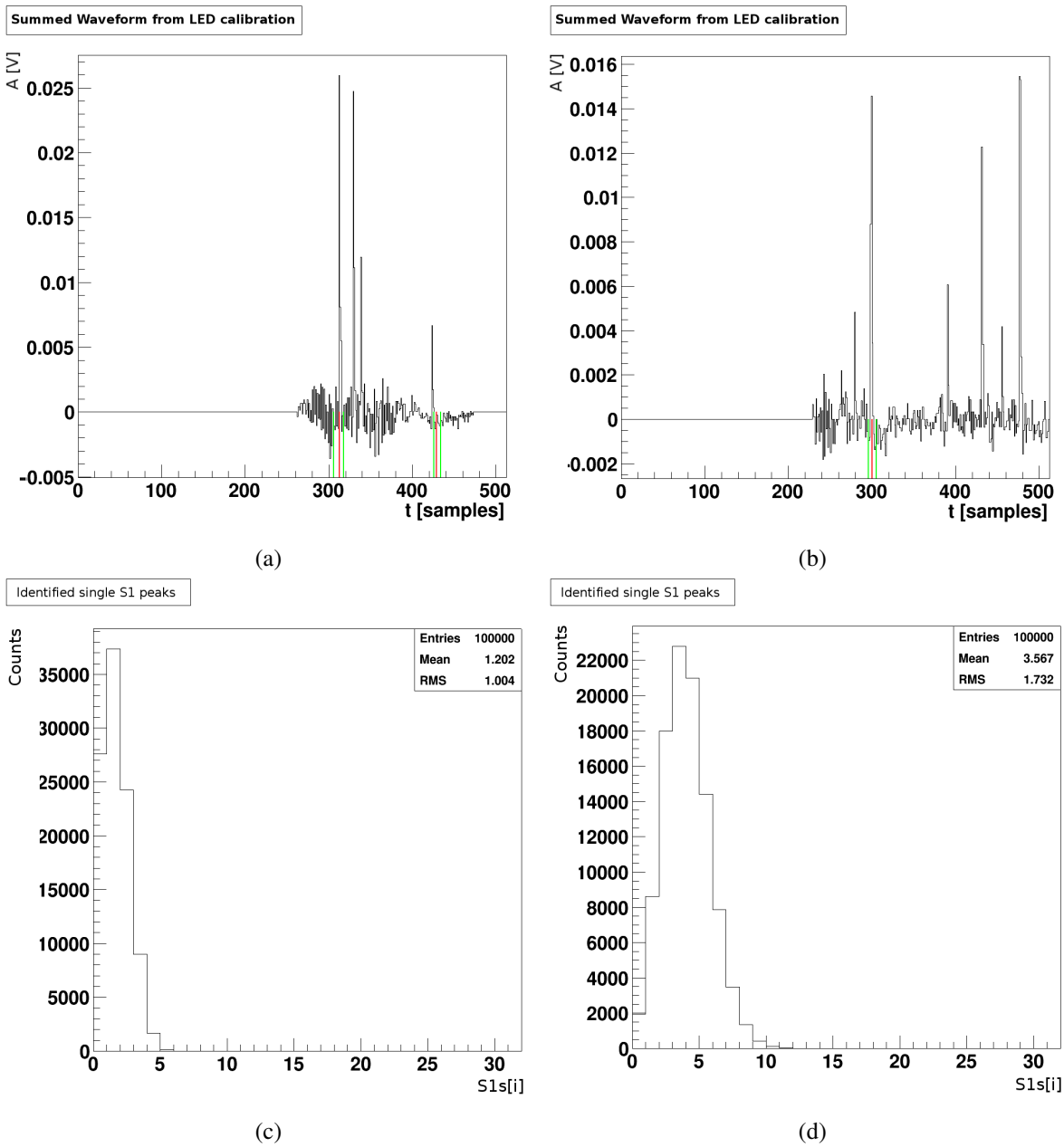


Figure 4.3 – Figure 4.3a and 4.3b show two examples of the peak-finder algorithm of *xerawdp* when a standard Dark Matter *xml*-file is used. The S1 trigger efficiency from 100000 LED waveforms yield in a mean of 1.2 identified S1s peaks per waveform (4.3c). This results increases by replacing the *xml*-file by a new LED *xml*-file with changes according table 4.1. The S1 trigger efficiency increases to a mean of 3.6 identified S1s peaks per waveform (4.3d).

4.5. Cut Definition for LED Data Analysis

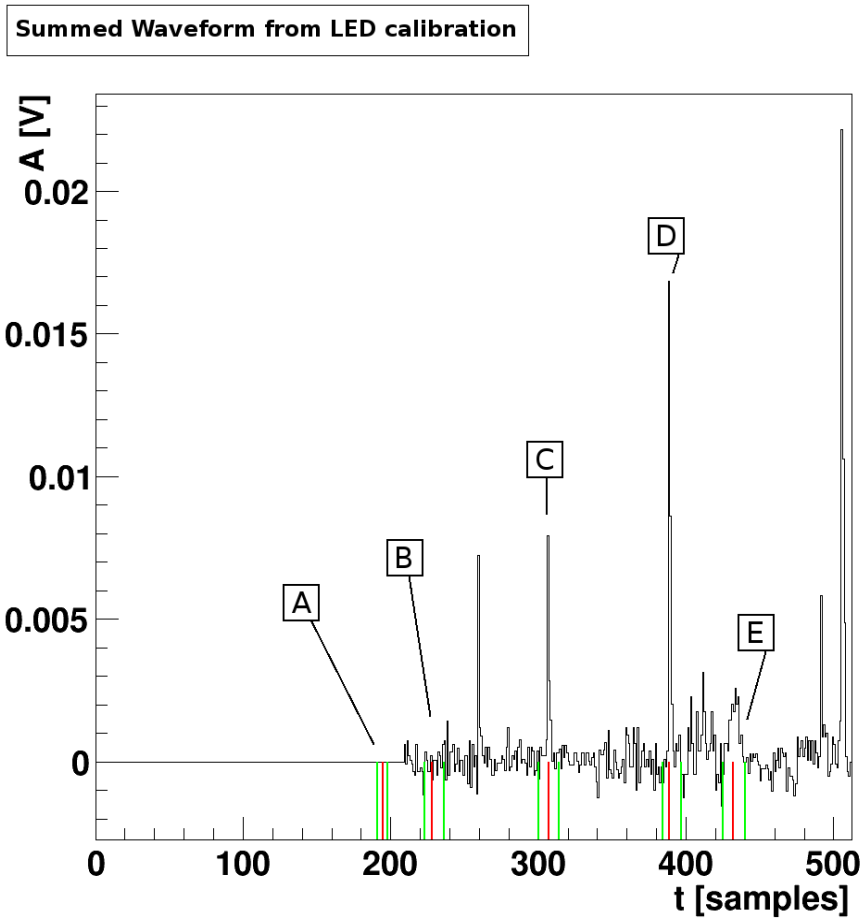


Figure 4.4 – Example waveform with identified peaks by *xerawdp*. It is shown that not every identified peak is a real signal and therefore an additional criterion is needed to increase the number real S1s peaks, the entropy (ε) criterion.

The preliminary results from subsection 4.4 show that a new *xml*-file increases the mean trigger probability, but on the other hand the quality of the identified S1s peaks is not yet clear. Therefore an extended study on the peak properties is done to verify the usability of the S1s peaks, mainly by waveform watching. This is important due to the point that the single peaks are dedicated to measure quantities such as the single photo-electron response (1 pe from a SPE or higher contributions of a ~ 2 pe signal from two photons in coincidence at the same PMT) and it is necessary to distinguish real peaks from noise contributions which are identified wrongly. Figure 4.4 shows that there are peaks miss-identified which makes it necessary to introduce a cut set. The waveform in figure 4.4 is created by using the new *xml*-file.

In case *A*, the peak-finder algorithm shows a weird behaviour in the position reconstruction. The integral of such S1s peaks is zero and therefore a minimum condition for a correct identified S1s peak is $S1sTot[i] > 0$, where i stands for each potential identified a S1 peak in a single LED waveform. Something similar would be necessary to distinguish *B* from a well-defined S1 peak such as *C* and *D*. In the worst case, an identified peak such as *B* yields a negative area below the peak. Therefore the entropy is introduced as additional S1 peak characteristic. In case of *E*, the identified S1 shows a broader behaviour compared to the expected S1 shape. In any case, the S1 characteristic is done with the $S1sHeight[i]$ and the measured quantity $S1s[pmt][i]$. The parameter $pmt \in [1,242]$ names the according PMT LED waveform in the studies. For these characteristic as well as for the entropy ($S1sEntropy[pmt][i]$), each PMT is treated separately.
The first characterisation via the $S1sHeight[i]$ is shown in figure 4.5a, where the amplitudes of the observed peaks (on the summed waveform) are presented. The spectrum does not end at $S1sHeight[i] = 0.5\,\mathrm{V}$ and there are peaks observed up to 1 V. The interesting structure is formed at the beginning. The bin width is chosen to $\Delta S1sHeight[i] = 1\,\mathrm{mV}$ and hence the very first bin present the share of the S1 signals which not exceed a peak height of 4 mV. These signals are noise and can be neglected. A first peak in the $S1sHeight[i]$ spectrum is observed at $S1sHeight[i] = 0.01\,\mathrm{V}$ (marker 1) and a second peak at $S1sHeight[i] = 0.03\,\mathrm{V}$ (marker 2). These two peaks correlate with the single and double photo-electron PMT hit during the PMT calibration.

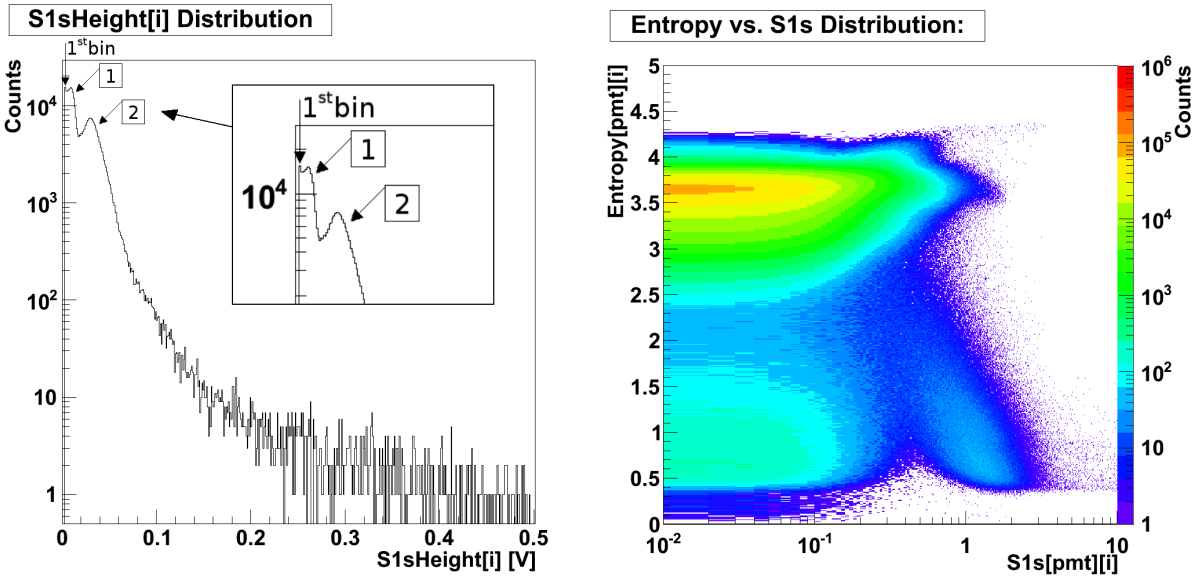
On the other hand, the $S1s[pmt][i]$ is correlated with the entropy information $S1sEntropy[pmt][i]$ in figure 4.5b. The peak entropy is calculated according formula (4.1) between the left and the right boundary of the found S1 peak by *xerawdp*.

$$\varepsilon := \sum_{i=l}^{i<r} p_i \log_2(p_i) \qquad \text{with } l := \text{left peak boundary and } r := \text{right peak boundary} \tag{4.1}$$

$$p_i := \frac{\text{sample content}_i}{\text{area of the potential peak}_i} \tag{4.2}$$

This definition is implied by the *information theory* and also known as *Shannon Entropy* [47]. The p_i in formula (4.1) describes by formula (4.2) with the same boundaries l and r from the potential S1s peak. The entropy ε tends to infinity for badly reconstructed S1s peaks, in case the entropy information is not normalised.
As shown in figure 4.5b there are three populations observed: A dominant share of miss-reconstructed S1s peaks for $S1sEntropy[pmt][i] > 2.5$ and another two populations below $S1sEntropy[pmt][i] < 2.5$. Here the left population comes from the noise contribution to the single PMTs. The right population contains the SPE signals with an obliterate transition to higher hit coincidences. The number $S1sEntropy[pmt][i] = 2.5$ which separates the identified S1s peaks is a conclusion from an extensive

waveform study in where different reconstructed S1 peaks have been analysed due to their contribution in the parameter distributions $S1sHeight[i]$, $S1s[pmt][i]$ and $S1sEntropy[pmt][i]$. Both figures

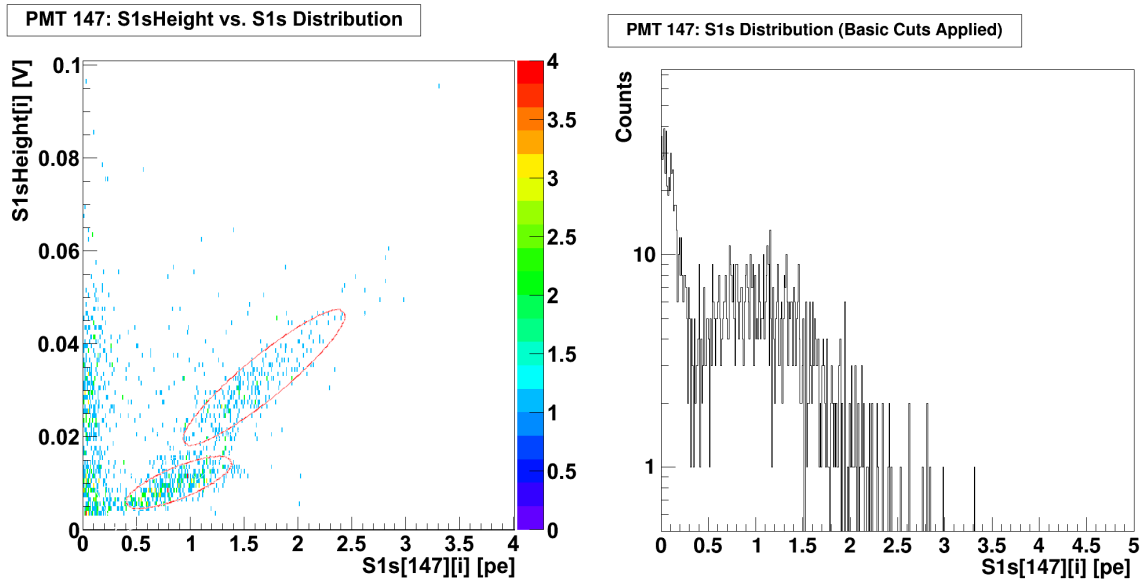


(a) (b)

Figure 4.5 – Both figures 4.5a and 4.5b show the quantities of the detected S1s peaks in the LED calibration data with the according LED *xml*-file. The investigated quantities are $S1sHeight[i]$, $S1s[pmt][i]$ and $S1sEntropy[pmt][i]$. An extensive waveform study shows that the entropy information is able to distinguish very well between good and bad reconstructed S1s peaks in *xerawdp*.

4.5a and 4.5b show their information for all PMTs and S1s peaks together, it is interesting how the result of $S1sHeight$ correlates with S1s for each PMT separately. Therefore the entropy cut of $S1sEntropy[pmt][i] < 2.5$ is applied. An example for one PMT is given in figure 4.6a. Here the PMT 147 is selected and the distribution of $S1sHeight[i]$ and $S1s[147][i]$ is shown. The two peaks of the total $S1sHeight[i]$ spectrum (figure 4.5a) vanish because of a single PMT show less events. Some larger amplitudes with $S1s[147][i] <\sim 0.5\,\text{pe}$ also survive the entropy cut. For $S1s[147][i] >\sim 0.5\,\text{pe}$, two populations are observed which are connected to the single and double photo-electrons from the LED calibration. This information can be deducted from their positions on the $S1s[147][i]$-axis and also their peak height overlaps with the previous observed distribution in figure 4.5a. Similar figures are available for each PMT and a useful single photo-electron calibration can be carried out.

The single photo-electron analysis is done on each PMT separately. The variable of interest is the $S1s[pmt][i]$ as it is already used in figure 4.6a. Due to the calibration with *single* photons, it is expected that each photon which is emitted by the LED hits one single PMT. To assure that every S1s peak comes only from a single PMT a simple coincidence condition is requested (see equation (4.5)).



(a) (b)

Figure 4.6 – The PMT 147 is presented as an example for the parameter distribution of $S1sHeight[i]$ and $S1s[147][i]$ with the applied entropy cut $S1sEntropy[pmt][i] < 2.5$. In figure 4.6a, two populations are related to single and double photo-electron are marked by red ellipses. Figure 4.6b shows the data selection if the basic cut set (see (4.3), (4.4) and (4.5)) is applied to $S1s[pmt][i]$ data. The resulting S1s-spectrum describes the SPE response by the broad distribution around $S1s[147][i] \sim 1\,\text{pe}$. The observed PMT noise (e.g. from dark current) is pictured by the sloppy distribution for $S1s[147][i] < 0.35\,\text{pe}$. Contributions from a second photo-electron in coincidence are suppressed.

If the condition is fulfilled the detected S1s peak is observed at one PMT only and a single photon is detected.

$$S1s[pmt][i] > 0\,\text{pe} \tag{4.3}$$

$$S1sEntropy[pmt][i] < 2.5 \tag{4.4}$$

$$S1sCoin[i] == 1 \tag{4.5}$$

To summarise this subsection, there are three necessary cuts to apply before the reprocessed LED data can be analysed. With this set of cuts, the $S1s[pmt][i]$ distribution is drawn for each PMT as it is shown in figure 4.6b. The chosen example is again PMT 147. The $S1s[pmt][i]$ distribution shows two populations. Values below $S1s[pmt][i] < 0.35\,\text{pe}$ are from noise contributions (e.g. dark current). Values above $S1s[pmt][i] > 0.35\,\text{pe}$ are supposed to be from SPE. Higher order contributions from two photons which hit one single PMT at the same time (e.g. double photo-electrons) are also included in the spectrum. Their part is strongly suppressed in the SPE spectrum (see figure 4.6b).

4.6. LED Raw Data

The raw data analysis of the LED data presented here is carried out with ten LED data files which were taken during the data taking period in 2013/14. This corresponds to five hours of LED raw data. Roughly every tenth LED raw data set is chosen to get a homogeneous data selection over the whole period. A detailed application, e.g. weekly LED calibration would need another data selection.
The LED raw data analysis focuses on the PMTs from the top and bottom array, in total 178 PMTs. Due to known noise issues of some PMTs, some PMTs are switched off in the DAQ during the data taking. These PMTs are not analysed by the here described procedures. The offline PMTs on the top and bottom array are summarised in table 4.2. In addition there are some PMTs observed which show an unusual bias to higher coincidence levels. This mean that coincidence of minimum two or more PMTs is needed to find a valid S1s peak. This comes from noise issues and these PMTs show noisy behaviour during the data taking.

	switched off PMTs	S1 signals out of range
Bottom array	100, 105, 148, 177	145, 157, 170, 171
Top array	9, 39, 58	12, 34

Table 4.2 – Table of PMTs which are switched off by the DAQ and a set of PMTs which yield S1s signals larger than $S1s > 10\,\text{pe}$ by accident. PMTs from the right column are not interesting for a first study of the LED raw data with $S1s[pmt][i]$ such it is presented here. They show an unusual bias to higher coincidence levels (noise issues) what makes them not usable for the required single fold coincidence.

4.7. The Single Photo-Electron Approximation Model

The previous subsections describe the necessary steps to process LED raw data files with *xerawdp*. To assure a higher S1s trigger efficiency, a new LED *xml*-file is also created. The S1 peak quality is then analysed considering their properties such as $S1s[pmt][i]$ and $S1sHeight[i]$. The S1s peak entropy also becomes important to quantify their quality. With these S1s signals, an alternative gain is determined by evaluating the $S1s[pmt][i]$ histograms for each PMT separately. An example for such a $S1s[pmt][i]$ histogram is shown in figure 4.6b.

4.7.1. The Gaussian/Poisson Data Approximation

The following analysis method assumes a Gaussian approximation for the noise section and an adjusted Poisson distribution to describe the single photo-electron response of the PMTs. Each of the

242 $S1s\,[pmt]\,[i]$ histograms is fitted with equation (4.6) in two steps. First the Gaussian and Poisson sections are fitted separately and then the fit information from the single fits is used as input for the combined fit.

$$SPE(S1s) = A_G \times \exp^{-\frac{(S1s-\mu_G)^2}{2\sigma_G^2}} + \quad A_P \times \frac{e^{-\frac{\lambda}{\sigma_P}} \left(\frac{\lambda}{\sigma_P}\right)^{\frac{S1s}{\sigma_P}}}{\Gamma\left(\frac{S1s}{\sigma_P}+1\right)} \tag{4.6}$$

The combined fit consists of six free parameters: The particular amplitudes are described by A_G (Gaussian) and A_P (Poisson). The Gaussian is defined by the expectation value μ_G and standard deviation σ_G. Similar to section 5, the Poisson distribution is adjusted by another parameter σ_P which scales the width. The most important parameter is the Poisson expectation value λ which describes the SPE signal distribution. The most likely value $\hat{\lambda}$, at which the maximum probability is observed, is used to describe the SPE response. The Poisson distribution is defined by the Γ function that real numbers can be described instead of discrete numbers only. This re-definition results a continuous SPE description. In the following analysis, the term Poisson distribution refers to the here defined continuous Poisson-like function. According to the observed *single* photons, the parameter $\hat{\lambda}$ is supposed to be around $\sim 1\,\text{pe}$. With equation (4.6) it is possible to describe the SPE response of each PMT and its noise contribution. The noise share separates from the signal part at $S1s[pmt][i] \sim 0.35\,\text{pe}$. The exact value depends on the PMT itself and defines the lowest boundary for a potential detectable S1s response for a single PMT.
In summary, an analogous description compared to the usual gain calibration (see figure 3.29b) is applied. The following analysis names the Poisson part as *distribution*. This is not mathematically correct anymore because the presented adjustment of Poisson distribution also allows real numbers.

4.7.2. Result of the S1s Histogram Approximation

Fit results with equation (4.6) are shown in figures 4.7a, 4.7b, 4.7c and 4.7d. Four PMTs are selected: from the top array (PMT 13 and 93) and from the bottom array (PMT 110 and 164). The main focus of this work is the PMT calibration of the top and bottom arrays and therefore no PMT examples are selected from the veto PMTs. The fits are carried out by an automatised script where a self-developed minimisation process searches for the *best combined fits* according their likelihood information from the ROOT fit routine. An inferior aim of this work is to present a fully automatised piece of software which analyse all PMTs in a row. With regard to further applications, the fit procedure should run autonomous at each PMT to speed up the analysing process.
The combined fit is shown by the green line. The two pre-fits are indicated by the dotted red (Poisson) and grey (Gaussian) lines and the separated results of the combined fit are sketched by the red (Poisson) and blue (Gaussian) lines. Besides the mentioned six free parameters, the *local minimum*

is determined by the minimum between the Gaussian and Poisson distribution in equation (4.6). This local minimum describes the S1s threshold which is fixed to $S1_{th} = 0.35\,\mathrm{mV}$ in the usual data analysis. The fit by equation (4.6) allows to identify small variations around $S1_{th}$ and the local minimum is used to determine its deviation. This information becomes interesting for the SPE response which is defined by the integrated signal of each PMT. The local minimum is marked in figure 4.7a by the red arrow.

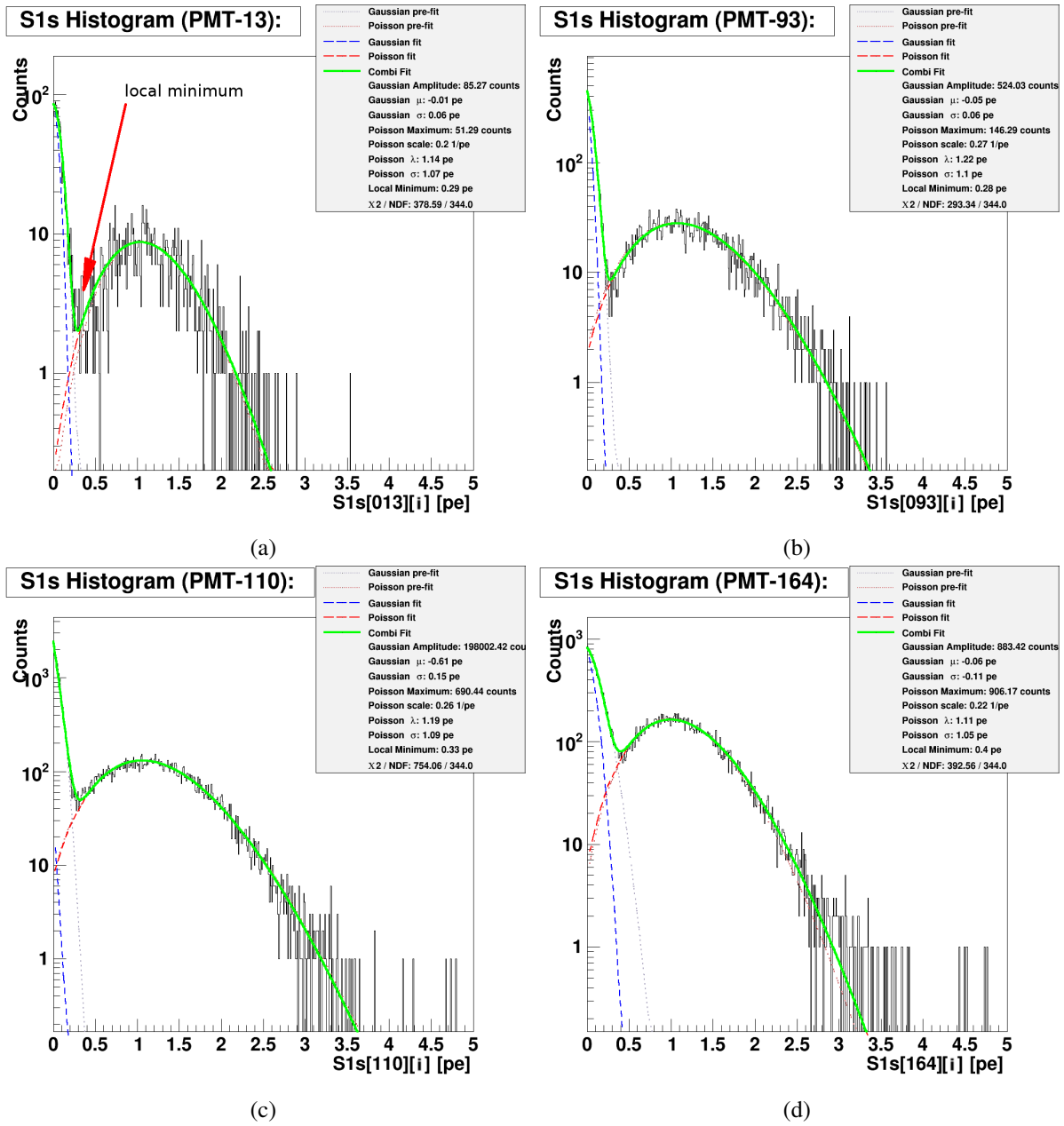


Figure 4.7 – A choice of different $S1s\,[pmt]\,[i]$ distributions: PMT (top array) 13, 93 and (bottom array) 110 and 164. Each figure shows the combined fit in green and the pre-fits (pointed lines in red and gray). The single contribution (Gaussian or Poisson) of the combined fit are drawn in dashed blue and dashed red.

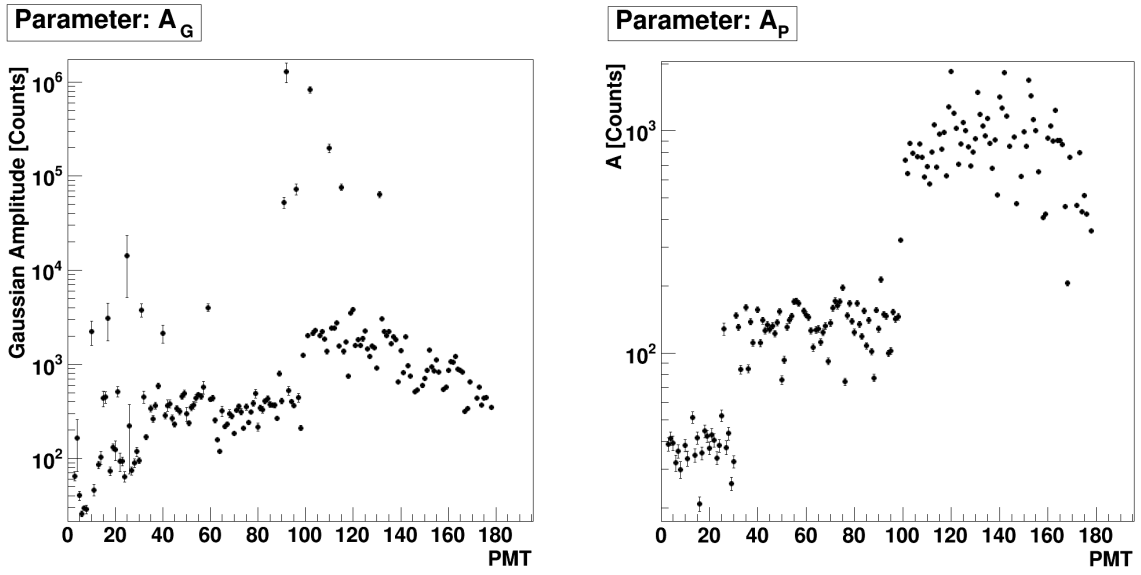
4.8. Single Photo-Electron Analysis with S1s Histograms

The primary analysis is based on the fits of the $S1s\,[pmt]\,[i]$ histograms as shown in section 4.7.2 (or see figures 4.7a, 4.7b, 4.7c and 4.7d). The related fit information is analysed. The most interesting parameter, is the Poisson expectation value λ and the related most likely value $\hat{\lambda}$. Furthermore the local minimum variation is interesting, to determine the exact S1s response of a single PMT and distinguish between signal and noise contribution. Hence, the influence of the varying local minimum is analysed due to the signal part, in comparison to the fixed threshold of $S1_{th} = 0.35\,\text{pe}$.
The goal of this analysis is to quantify the SPE response of the top and bottom array which is then expected to be $\sim 1\,\text{pe}$ from the regular gain calibration.

4.8.1. Discussion: Fit Parameter of the Combined Fit

The following section discusses the outcome of the six fit parameters. The characteristic fit by equation (4.6) is separated into the noise part (Gaussian approximation) and the signal part (Poisson approximation). Due to the two-step fit which starts with separated fits to the noise and signal parts with its characteristic fit functions and continuous with the combined fit, the fit outcome does strongly depend on the start values, especially if the fit parameters of the single fits are used in the combined fit. Hence some of the combined fits show a worse outcome which influences mostly the noise part.

Starting with the less important fit parameters A_G and A_P which normalise the Gaussian and the Poisson distribution to archive the required number of counts in the histogram, both parameters are shown in figure 4.8a and 4.8b. The noise part contains three steps: For PMT 1 to PMT 30, the multiplication factors are $A_G \sim 100\,\text{counts}$, then the count number increases to $A_G \sim 500\,\text{counts}$ for PMTs 31 to 98. The count number increases to $A_G \sim 1000\,\text{counts}$ for the PMTs 99 to 178. The PMT pattern is responsible for the three observed steps. The PMTs 1 to 30 are organised on the outer ring of the top array and therefore they do not collect so much light compared to the PMTs which are located more to the inside of the top array (second step). The third step contains all PMTs from the bottom array which is designed to increase the light detection probability by a compact PMT arrangement and therefore it is expected to detect a higher count number. The decrease in the count number to PMT 178 comes from the PMT box itself where the PMTs are connected. These PMT box shows electronic noise which also responsible for the decrease of the counting. In addition, some PMTs are observed for which the fit parameters differ by a huge factor from the PMTs around. This behaviour is a hint that these fits are not carried out so well by the automatic fit routine. The signal part in figure 4.8b shows the same three steps in the distribution of the counts. The origin is the same such as for the noise part. In contrast to



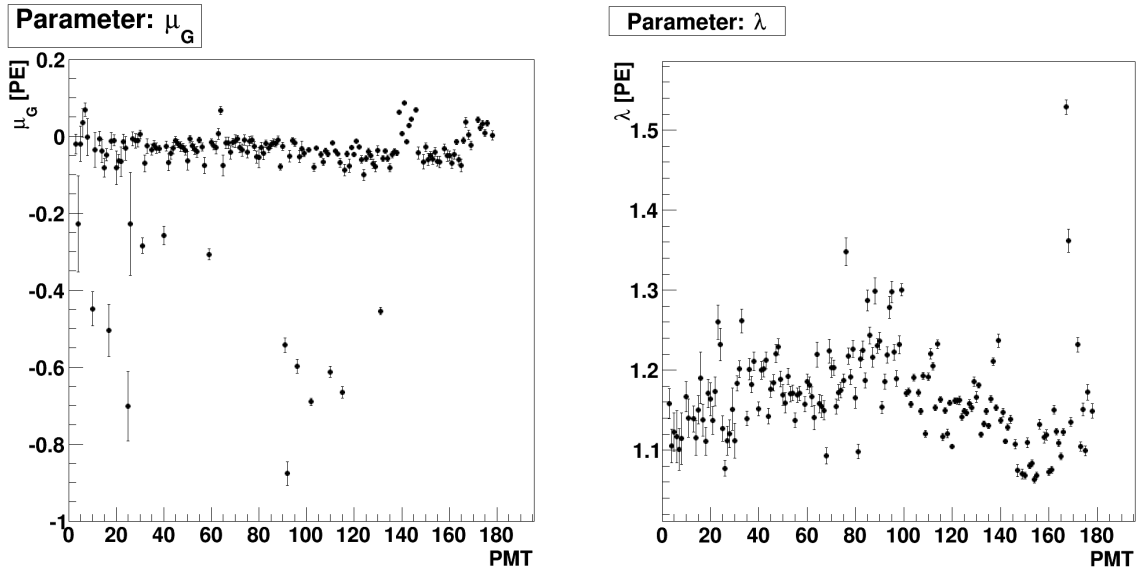
(a) (b)

Figure 4.8 – Comparing the both parameters A_G and A_P which scale up the Poisson and Gaussian share according the event rate in each PMT (log scale). The three populations origin from PMT arrangement on the top and bottom PMT array and hence the a larger scaling factor is somehow proportional to the light collection.

the noise part, no excesses in the error bars are observed which indicate a bad fit.

The next parameters are the particular expectation values. The fit outcome for the noise part is shown in figure 4.9a. Here μ_G varies around $\pm \sim 0.01$ with some PMTs which do not follow this observation. These are the same PMTs which also show exceptional behaviour in figure 4.8a. This is another hint for a worse fit behaviour. The last PMT group up to PMT 178 shows an excess to a positive expectation values. Again, these PMTs are connected to a filter box which shows an increase in noise. This influences the connected PMTs. In detail, most obtained μ_G fit parameters are slightly below $\mu = 0\,\text{pe}$ what is a result of the fit routine which is not limited to positive numbers only. The expectation value λ of the Poisson fits varies between $1.1\,\text{pe} \leq S1s \leq 1.6\,\text{pe}$. The large excesses are come from individual PMTs. These PMTs do not show a bad fit by default, but their distribution itself is biased to this higher expectation value.

The last set of parameters are the widths σ_G and σ_P. Both parameters are defined in a different way and hence a direct comparison is not possible. The width of noise part is defined by the Gaussian deviation which is part the Gaussian definition itself. This fit parameter can be used to describe the width of

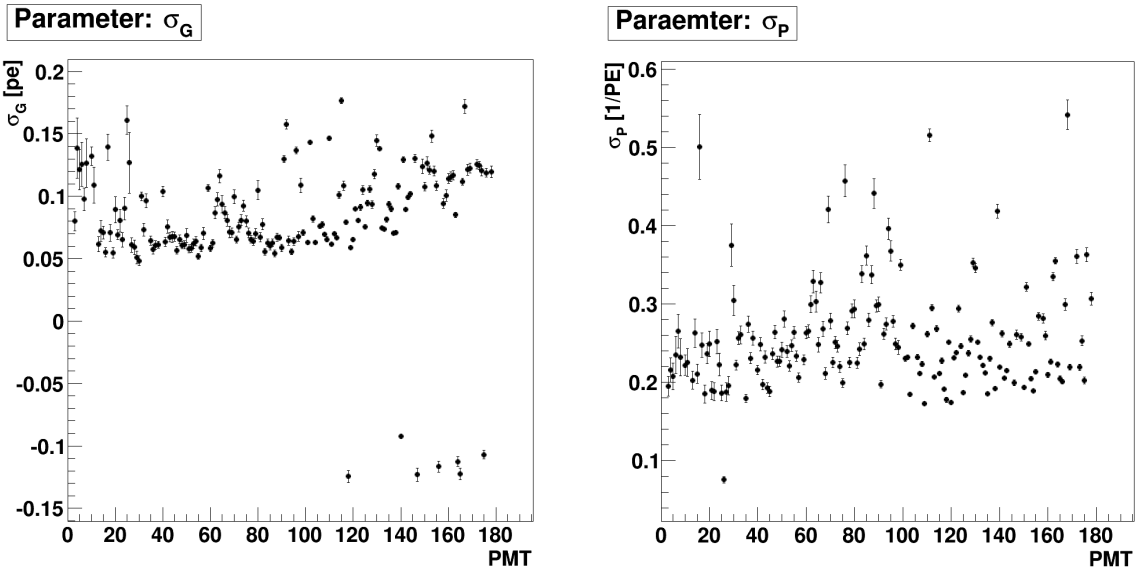


(a) (b)

Figure 4.9 – The expectation value μ_G and λ are plotted for each PMT to describe the noisy and the signal part of the $S1s\,[pmt]\,[i]$ histograms. For most fit outcomes, the μ_G is near zero or slightly negative. This is observed due to the fact that there is no limitation to positive values included in the fits of the $S1s\,[pmt]\,[i]$ histograms and hence the fit parameter μ_G fits pretty well the expected value of $S1s[pmt][i] \sim 0\,\text{pe}$ in the combined fit. The Poisson expectation value λ is related to the SPE by the most likely value of $\hat{\lambda}$. Therefore it is used to determine SPE response.

the noise. The local minimum will vary with the width of the Gaussian distribution what influences the SPE response of the PMTs. The parameter σ_P is used during the fit routine to scale the width of the Poisson distribution when it is scaled up at the same time with parameter A_P. To compensate this scaling in the fit, the additional parameter increases the goodness of fit when the signal part is approximated. Beyond this, that parameter has no further use in this data analysis.

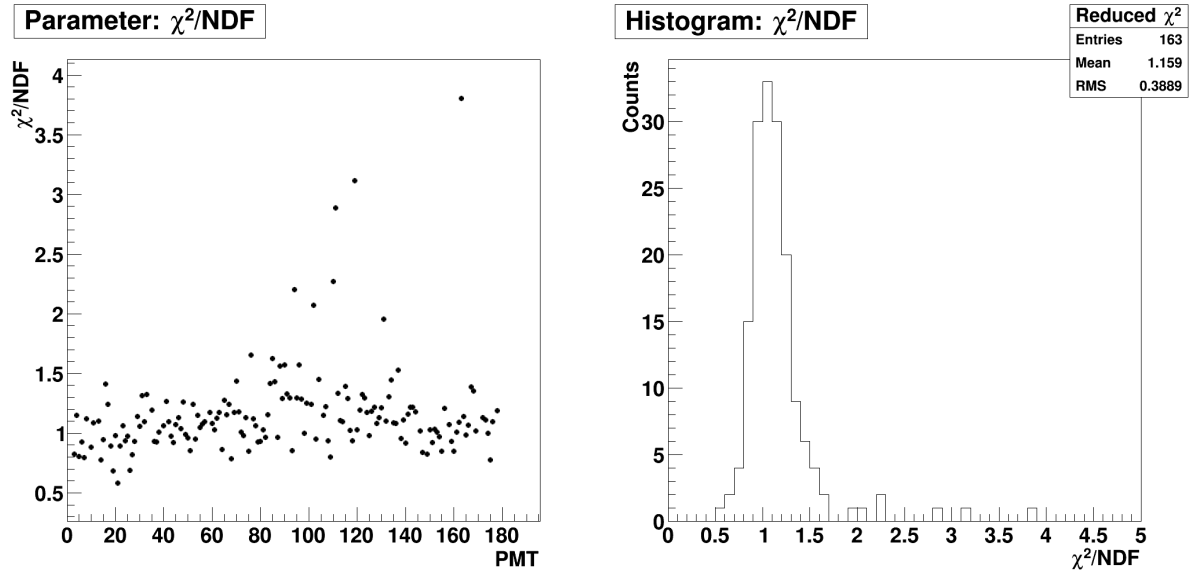
From figures 4.8a and 4.9a there a number of PMTs determined where fit parameters A_G and μ_G deviate at the same time from the general PMT behaviour. An example is the PMT 110 on the bottom array. The fit result is shown in figure 4.7c. Two further examples for PMT 17 and 130 are shown in figure A.21a and A.21b in the appendix. Due to the chosen binning the Gaussian approximation is done on 35 bins ($\Delta S1s = 0.01\,\text{pe}$) and hence the its only possible to fit the right tail of the Gaussian distribution. This reduces the goodness of the fit and this leads to the observed outliers. From the goodness of the fit, determined by the reduced χ^2, it turns out that the combined fit of PMT 110 shows $\chi^2/NDF_{110} = 2.19$ and the other two examples show $\chi^2/NDF_{17} = 1.24$ and $\chi^2/NDF_{17} = 1.21$. PMT 110 is an excess, that is also shown in figure 4.11, where $\chi 2/NDF$ is plotted for each PMT and its distribution. The mean value is $< \chi 2/NDF >= 1.1$. From this outcome for the six parameter fit, the description for



(a) (b)

Figure 4.10 – The fit parameters σ_G and σ_P do not compare to each other. Both describe the width of the according fit distribution (Gaussian and Poisson). In case of the noise part, the σ_G is part of the Gaussian definition and describe how far the noise part leak into the signal model. The additional fit parameter σ_P is added to the Poisson description of the signal part. The fit parameter increases the goodness of fit because the width can scale independent from the expectation value.

the $S1s\,[pmt]\,[i]$ histograms works pretty well. Based on this good description, the physical impact and applications are discussed in the following subsections.



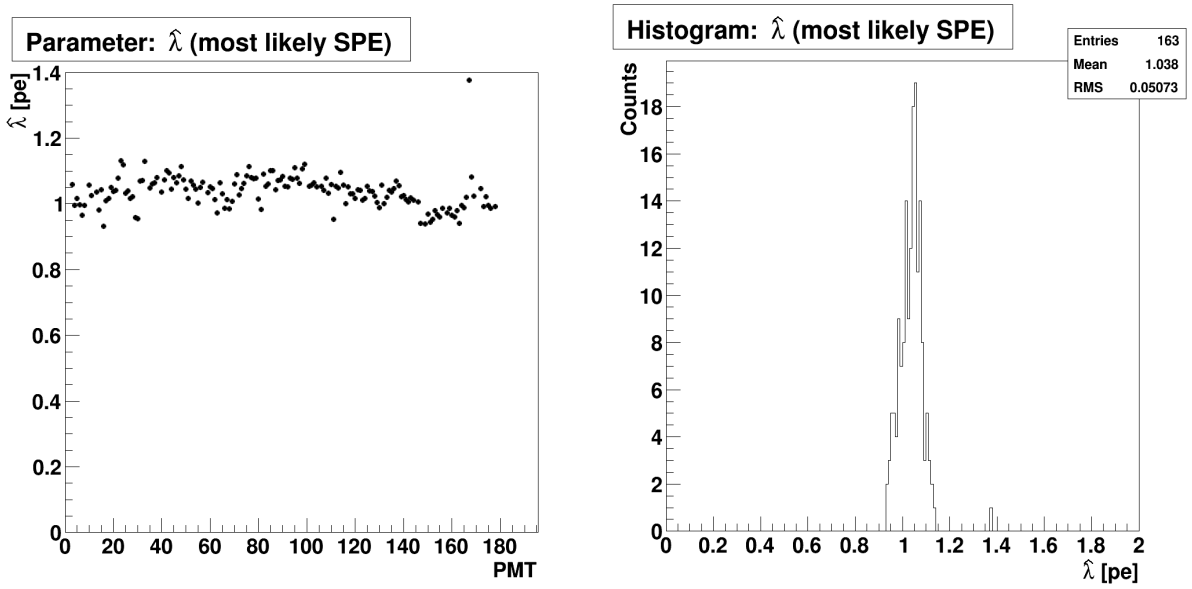
(a) (b)

Figure 4.11 – Both figures show the goodness of the combined fit by their value for $\chi^2/_{NDF}$. Besides some PMTs which show an excess, the goodness of the fit is pretty good. Figure 4.11b shows that their distribution varies around a mean value of $< \chi^2/_{NDF} >= 1.1$.

4.8.2. Single Photo-Electron Response with S1s Histograms

Based on the outcome of the six parameters for the combined fit, the response to single-photo electrons can be measured. The fit function in equation (4.6) describes the single PMT spectra, especially the signal part is described by the Poisson part. Due to the definition of the fit function, mainly the Poisson expectation value λ characterises the spectrum. The expectation values λ are already shown in figure 4.9b. From here, a bias is observed in case the the SPE response is described by λ: The expectation value λ from the fit does not describe the expected SPE response of $\lambda = \sim 1\,\text{pe}$ correctly. For example, PMT 164 (see figure 4.7d) has an expectation value $\lambda = 1.11\,\text{pe} \pm 1.05\,\text{pe}$. The reason for that deviation from the expected response is that the most likely value $\hat{\lambda}$ of a Poisson distribution is not the same as λ. The most likely value is always smaller due to the definition of the Poisson distribution and the impact of that difference becomes stronger for small expectation values λ. Therefore the most likely value $\hat{\lambda}$ should describe the SPE better. This is demonstrated in figure 4.12a, here $\hat{\lambda}$ is shown for each single PMT and figure 4.12b shows the $\hat{\lambda}$ distribution, based on the involved PMTs, the mean value $< \hat{\lambda} >= 1.038\,\text{pe}$ with a RMS of $\sigma_{\hat{\lambda}} = 0.051\,\text{pe}$. There are some PMTs which can be identified in figure 4.12a (Look at the jump around PMT 147). This can be reduced to their filter box connection. PMTs 147 to 162 are connected to the same filter box. They show a similar response. PMTs 163 to 178 are connected to another filter box. But nevertheless, these PMTs do not show large excesses

compared to other PMTs. Hence $\hat{\lambda}$ describes the SPE response pretty well for the involved PMTs. From that result, the alternative gain calibration by *xerawdp* with $S1s\,[pmt]\,[i]$ as variable of interest does work well and shows useful results. It proves the expected SPE response.



(a) (b)

Figure 4.12 – Instead of using λ which characterises the Poisson distribution in the combined fit, the most likely value $\hat{\lambda}$ is used to describe the single photo-electrons for each PMT. Figure 4.12a shows $\hat{\lambda}$ for each PMT. Here the SPE is measured as expected. The histogram in figure 4.12b shows the distribution for the involved PMTs with $<\hat{\lambda}>= 1.038\,\text{pe}$. The outlier in figure 4.12a is also one the PMTs where the fit shows a worse $\chi^2/_{NDF}$ in figure 4.11a.

4.8.3. S1 Yield Analysis

The information which is provided by the six parameter fit allows to test the potential influence of the use of the local minimum to separate signal and noise part from each other. According to the *xml* configuration file, an excess of minimum $U > 3\,\text{mV}$ on the summed waveform is needed to trigger the peak-finder algorithm (see also section 3.5.3.2). The ideal case is that an observed signal on the summed waveform comes from a few PMTs and hence a certain number of PMTs dominate the total S1 signal. In truth, there are also single S1s signals or noise stored from single PMTs which are below the S1 trigger voltage of 3 mV. This is also observed for the LED calibration with $S1s\,[pmt]\,[i]$ histograms and therefore this information can also be used in the analysis of single PMTs. The noise part comes from small contributions of S1 or noise. The width of the noise influences the signal part:

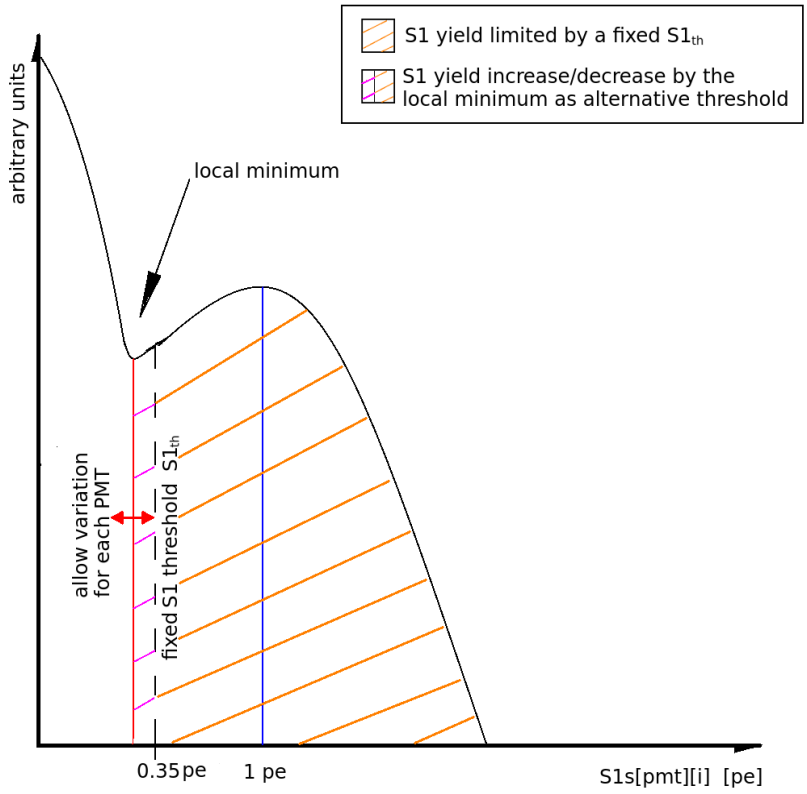
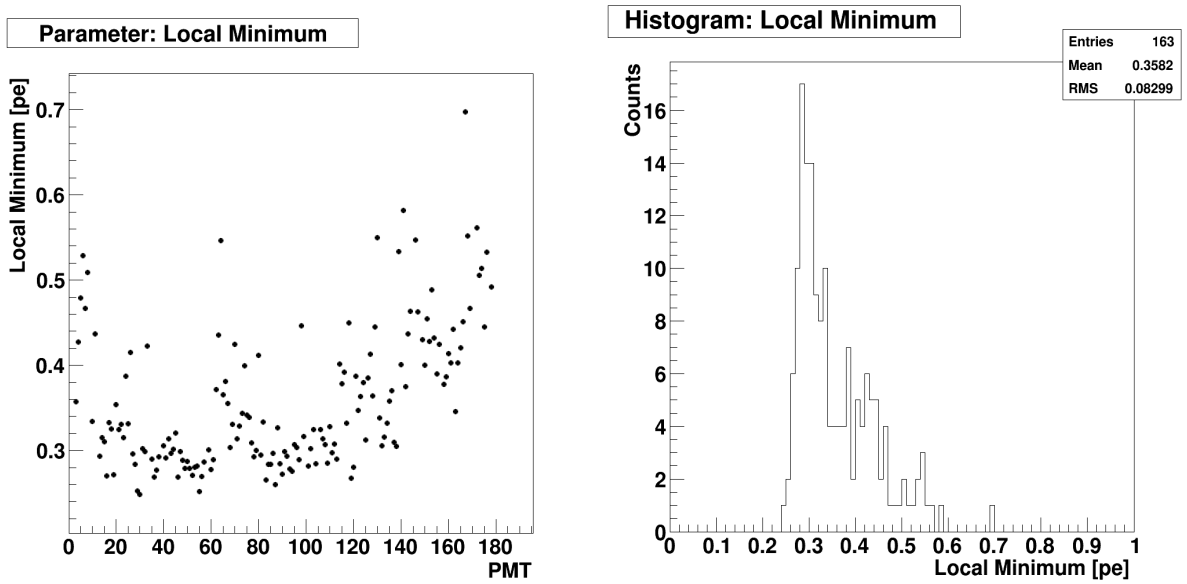


Figure 4.13 – Sketch of the local minimum as the separator between signal and noise in comparison to the fixed S1 threshold in XENON100. It is also shown how the local minimum influence the S1 yield which is understood as the integral from local minimum or $S1_{th}$ to 10 pe.

The larger the noise share, the smaller the signal part. The separation by the S1 threshold divides the signal and the noise part by this fix number ($S1_{th} = 0.35\,\text{pe}$). The $S1s\,[pmt]\,[i]$ histograms and their combined fit allows to determine the noise and the signal part from the data and calculate the local minimum between noise and signal for each PMT. The idea is shown in sketch 4.13 where the local minimum influences the S1 yield (see definition in equation (4.8)). The figures 4.7a, 4.7b, 4.7c and 4.7d already include the value in the fit status overview. Figure 4.14a shows the trend of the local minimum for each PMT, especially for PMTs from the top array the local minimum is slightly below $S1_{th} = 0.35\,\text{pe}$ and the bottom array PMTs show an increase in the local minimum to up to $S1_{th} = \sim 0.5\,\text{pe}$. A summary is also given by the projection in figure 4.14b where the mean value of the observed local minimum is $S1_{th} = 0.35\,\text{pe}$. The structure towards larger values comes from the bottom PMTs mainly. Therefore the fixed value of 0.35 pe can be identified, but single PMTs deviate from that value. From that, the bottom PMTs show a broader noise part which influences the S1 signal part.



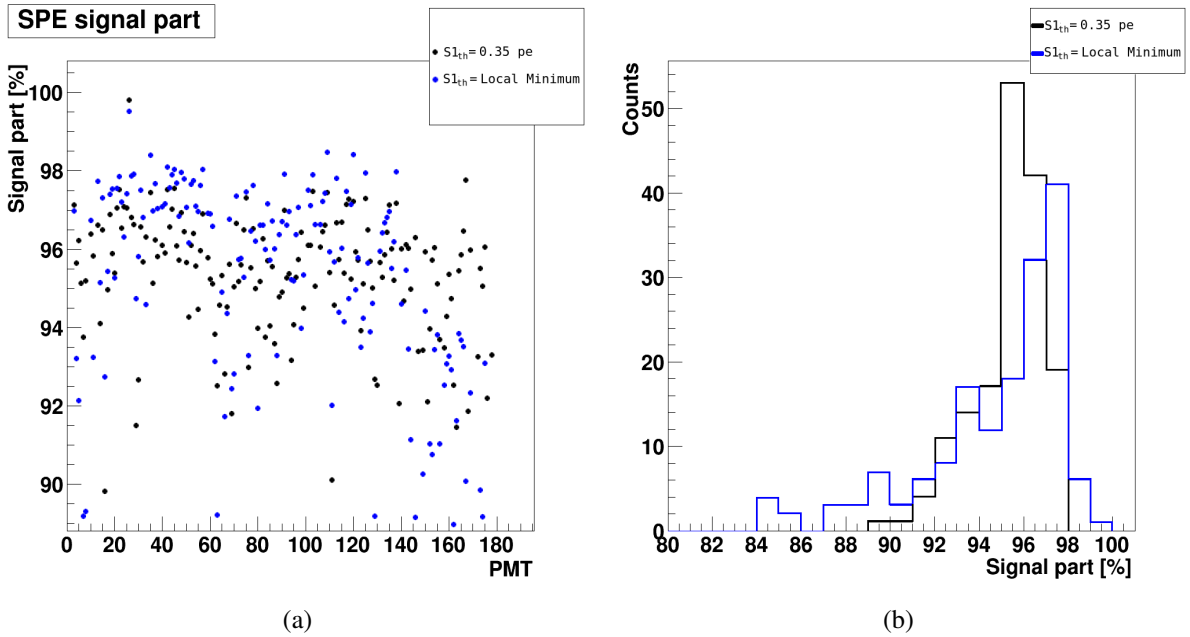
(a) (b)

Figure 4.14 – An overview about the local minimum between the Gaussian noise and Poisson signal part. The local minimum varies around the expected 0.35 pe. The top PMTs show a deviation to slightly smaller values, and the bottom PMT array show an increase in the local minimum. These PMTs from the bottom array show a broader noise share in the signal region.

Based on the S1 yield which comes from the Poisson description, the signal part is reduced by the S1 threshold. Therefore figure 4.15a shows how much S1 signal is left. This is estimated by I_s, the

integrated signal divided by the total signal (see equation (4.7)). The upper boundary is chosen by due to the fact that 10 pe are far enough away from the SPE response. The ”min“ stands for the lower boundary, the local minimum as it is shown in figure 4.14a and marked by the blue dots. This is compared to the outcome with the fixed threshold (black dots). In summary, the signal part has a mean acceptance of 95.3% for the S1 threshold and 94.7% for the local minimum. These values are extracted from figure 4.15b. Hence the local minimum reduces the S1 acceptance signal but not significantly. Considering the histogram in figure 4.15b, taking the local minimum instead of the fixed S1 threshold leads to a broader noise part in some of the PMTs. This results in the larger deviation which can be seen for example in PMT 130 (see figure A.21b in the appendix).

$$I_s = \frac{\int_{min}^{10\,\mathrm{pe}} SPE(S1s) \quad dS1}{\int_{0\,\mathrm{pe}}^{10\,\mathrm{pe}} SPE(S1s) \quad dS1} \tag{4.7}$$



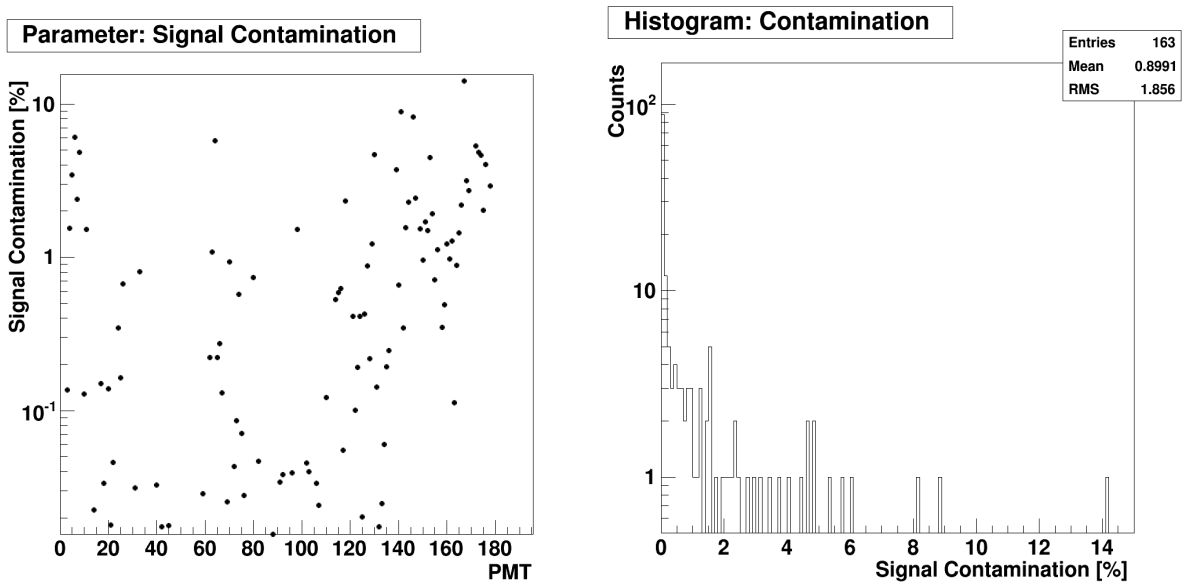
(a) (b)

Figure 4.15 – Comparing the integrated signal of the SPE by using a fixed threshold and the local minimum from the combined fit. The mean integrated S1 signal is the same within the RMS but from some PMTs the local minimum influences the signal range by a broader noise part.

Another important information is the signal contamination which is described as the noise part (Gaussian) which leaks into the signal range. Ideally, the contamination is only small. The contamination is determined by the integrated Gaussian distribution from ”LM“ to 10 pe (see equation (4.8)). It is normalised to the integrated signal of the Poisson part to estimate the contamination in percentage.

The contamination is given in figure 4.16a. The maximum observed contamination is 14% of the SPE signal. This is observed for PMT 167 which is working but shows high fluctuations in the weekly estimated gains from the standard LED calibration. In general, PMTs from 140 to 178 show a higher contamination. Simultaneously, these PMTs also show a higher noisier behaviour during weekly calibrations. For the PMTs below 140, the contamination is less than 1%. From here the contamination is determined for each PMT separately and PMTs which show a contamination less than 1% will less influence the S1 yield.

$$I_c = \frac{\int_{LM}^{10\,\mathrm{pe}} SPE_{Gaussian}(S1s)\,\mathrm{d}S1}{\int_{LM}^{10\,\mathrm{pe}} SPE_{Poisson}(S1s)\,\mathrm{d}S1} \tag{4.8}$$

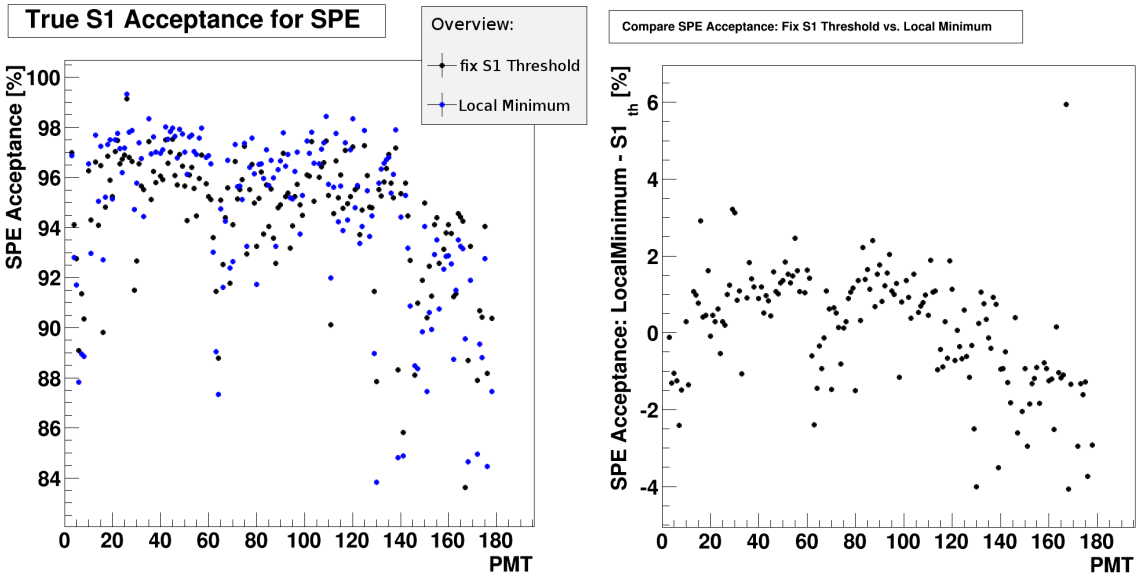


(a) (b)

Figure 4.16 – Based on the outcome for noise approximation with a Gaussian distribution, the contamination is calculated with equation I_c to estimate the part of the noise which leaks into the signal range. Besides the last $\sim$ 50 PMTs from the bottom array the contamination is less than 1%.

$$I_t = \frac{\int_{LM}^{10\,\mathrm{pe}} SPE_{Poisson}(S1s)\,\mathrm{d}S1}{\int_{0\,\mathrm{pe}}^{10\,\mathrm{pe}} SPE_{Poisson}(S1s)\,\mathrm{d}S1} - \frac{\int_{LM}^{10\,\mathrm{pe}} SPE_{Gaussian}(S1s)\,\mathrm{d}S1}{\int_{LM}^{10\,\mathrm{pe}} SPE_{Poisson}(S1s)\,\mathrm{d}S1} \tag{4.9}$$

Based on the information on the contamination as is given in figure 4.16a the *true* S1 yield I_t is calculated. This is done using formula (4.9). Here the signal part is calculated according to the Poisson distribution from the minimum/fixed threshold to 10 pe and the contamination is subtracted from this value. Therefore the true integrated signal will be the same for some PMTs according to the already observed S1 yield and some PMTs will lose in the detected S1 yield due to the contamination in figure 4.16a. This is shown in figure 4.17a where the outcome for equation (4.9) is calculated for



(a) (b)

Figure 4.17 – From the knowledge about the integrated S1 signal and the contamination, the *true* integrated SPE is calculated by equation (4.9) in figure 4.17a. This is done for the fixed S1 threshold and the alternative local minimum. The result is given by the accepted SPE signal range. To answer the question which one is better to estimate the S1 yield the difference for both I_t, shown in figure 4.17b. It turns out that for the PMTs on the top array the local minimum would be an alternative to the regular S1 threshold. On the bottom array the fixed threshold saves the S1 detection ability.

the fixed S1 threshold in black and for the local minimum in blue dots. The decrease in the true S1 yield is already shown by the contamination of figure 4.16a for each PMT. Therefore there is no large impact to expect for most of the PMTs. The interesting comparison comes from the choice of the fixed S1 threshold or the local minimum. Therefore figure 4.17b shows the result of subtracting $S1_{th}$ from the local minimum. Besides one exception for PMT 167, the difference is between $\pm 4\%$ if the S1 threshold is compared to the local minimum for the SPE signal (I_t). Due to the difference, there are two possible cases:

$$\text{local minimum} - S1_{th} > 0 \tag{4.10}$$

$$\text{local minimum} - S1_{th} < 0 \tag{4.11}$$

The first case (equation (4.10)) describes that the $S1_{th}$ is smaller than the local minimum and therefore the difference of the signal parts becomes positive. If the local minimum is used as an alternative S1 threshold, the noise level increases as does the noise contamination in the signal part, and therefore these PMTs show a worse SPE identification ability based on the true integrated signal. An example is given in figure 4.19b for PMT 169. Here the local minimum (LM) is determined to $S1_{LM} = 0.47\,\text{pe}$

what indicates a higher noise level and less S1 yield. The other case is described by figure 4.19a where the $S1_{LM} = 0.27\,\mathrm{pe}$ what is lower than the fixed S1 threshold. In case of equation (4.11), using the local minimum as alternative S1 threshold and hence the S1 yield increases. Divide the outcome for PMTs which fit either equation (4.10) or (4.11), the mean S1 yield detection ability is calculated as $S1_{>0}^{mean} = 1.12\%$ and $S1_{<0}^{mean} = -1.4\%$ (see figure 4.18). The difference becomes interesting because of the arrangement of how the PMTs are connected to the ADCs. By looking at figure 4.17b it seems that the PMTs of the top array (except PMTs 3 to 9, 62 to 80) increase their detection ability by choosing the local minimum instead of the fixed threshold. This changes for the PMTs of the bottom array: The increase in the detection ability goes down and turns into a decrease. This decrease is interesting because the PMTs at which that phenomenon is observed show more often a noisy behaviour during the usual data calibration procedure. The PMTs are named from 99 to 178 in lines from left to right and therefore the observed decrease is not a result of the PMT arrangement itself (compare to figure 3.28 where a map of the top and bottom array is shown). The top array shows two exceptions: PMTs 3 to 9 also show a decrease in the SPE detection ability. Here the PMTs 3 to 8 are connected to the same ADC but not PMT 9, but all are connected to the same filter box. Something similar is observed for PMTs 62 to 80: PMTs 62 to 67 and PMT 68 to 74 are connected to different ADCs but also two different filter boxes. Therefore the filter boxes and ADCs seems to influence their behaviour but are not the only explanations. Other possibilities are higher noise conditions or higher dark current rate. The quantum efficiency (QE) was measured to decide where the PMTs are arranged on the PMT array. Due these measurements, the PMTs with the highest QEs were put onto the bottom array in the middle and lower QE PMTs were arranged around to fill up the bottom array. The PMTs on the top array were homogeneously distributed with respect to their QE. Therefore the QE is not a suitable explanation to describe the decrease on the bottom array. In summary, PMTs from the top array would profit from the choice by using the local minimum instead of a fixed S1 threshold.

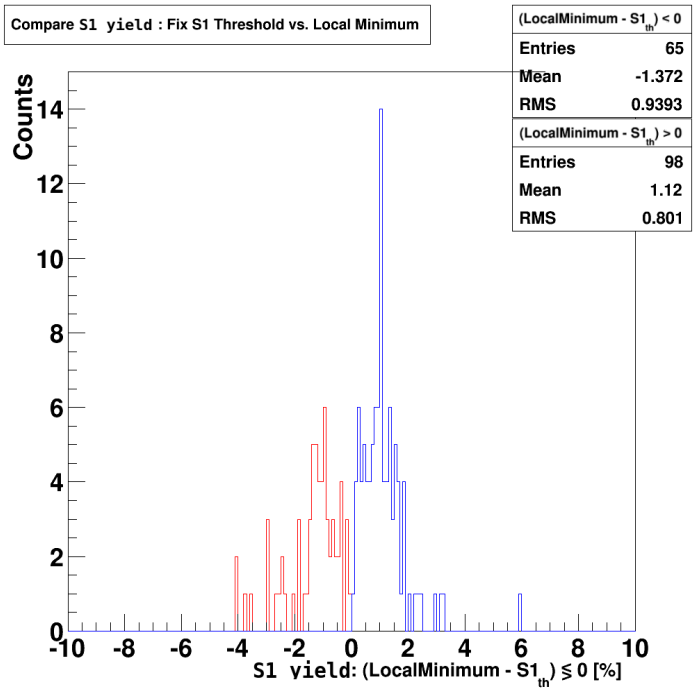
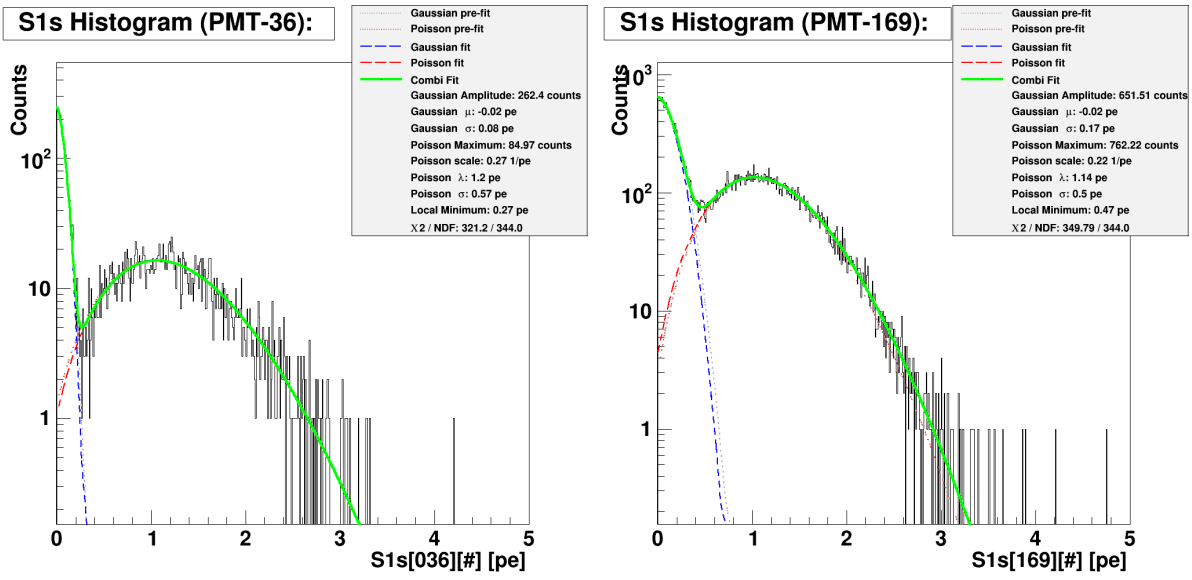


Figure 4.18 – Based on the observation in figure 4.17b, the SPE detection ability is separated in increase (blue) and decrease (red) in figure 4.18. The increase is related to PMTs on the top array and the decrease is more related to PMTs on the bottom array.



(a) **(b)**

Figure 4.19 – Two PMTs examples are given in which the local minimum is larger than the S1 threshold (bottom array, PMT 167 of figure 4.19b) and vice verse for PMT 36 (top array, figure 4.19a).

4.9. Alternative Method for Gain Determination

The previous section 4.7 presented an analysis for the $S1s[pmt][i]$ variable which is the result of the peak identification of *xerawdp*. At least, the measured signal before it is digitised is a voltage peak. The primary goal of the weekly LED calibration is the *PMT gain*. This information is necessary to convert the voltage peak, after it is identified by xerawdp, to units of photo-electrons (according to the SPE pulse). Therefore the previous $S1s[pmt][i]$ analysis helps to identify and support the SPE analysis from LED raw data. It is shown that the SPE identification works as expected for the previous measured gains. The disadvantage here is that the gain information is necessary during the reprocessing with *xerawdp* to yield the SPE with the correct most likely value. The following analysis uses the identified peaks from *xerawdp* but not its $S1s[pmt][i]$ information itself to compare the alternative gain with the standard gain.

4.9.1. Direct Gain Calculation

To calculate the gain directly out of the identified S1 peaks with *xerawdp* needs to take the area of the peak into account. This information is stored in another variable in *xerawdp*: $S1sRaw[pmt][i]$. This area is determined by the integral from the left to the right boundary in units of volts. The gain is calculated according to formula (3.32) in section 3.6.4. The quantity r describes the conversion from volts to bits and is already included in *xerawdp* when the ADC counts are calculated. Therefore the equation is altered to (4.12). Here the μ' describes the total area of the identified S1 peak which is stored in the *xerawdp* variable $S1sRaw[pmt][i]$. The remaining variables are already defined in 3.6.4. The advantage here is that this variable does not take the gain values from the previous PMT calibration into account.

$$\text{gain} = \frac{\mu'}{ZAfe^-} \quad \text{with } \mu' = S1sRaw[pmt][i] \tag{4.12}$$

This formula allows to ”translate“ the already shown histograms such as figure 4.7a, 4.7b, 4.7c and 4.7d to their gain equivalents. Figures 4.20a, 4.20b, 4.20c and 4.20d show the same PMT selection as before but with the according gain values. The result is again a spectrum with an SPE equivalent signal (gain) and a noise part with the difference that these gains are not calculated by the official software the LED calibration.
Based on the traditional PMT calibrations in XENON100 (compare to figure 3.29b), the direct gain calculation shows differences in the observed histograms for the noise part. In the standard gain calibration, the noise part is almost symmetric and can approximately be described by a Gaussian distribution. The noise which results from the identified peaks by *xerawdp* does not behave similar

to the noise from the standard PMT calibration. The peak-finder algorithm needs to exceed a certain threshold to trigger a peak on the summed waveform. The standard PMT software runs on each single PMT waveform and integrates on the waveform which is stored after the LED calibration setup has triggered the event. Therefore the found peaks by *xerawdp* should be "real" S1 peaks ideally, especial after the cut set is applied to the analysis (see section 4.5). Due to some remaining miss-identified peaks a small part of them has a negative gain. Hence the noise becomes more asymmetric and furthermore hard edges are observed in the noise part at the bin zero (see figure 4.20d).

The analysis of the gain calibration is done in the same way as in section 4.7.1 with equation (4.6). The noise part is again approximated by a Gaussian distribution and the signal part is described by the Poisson distribution. There are two pre-fits, and then a combined fit carried out. The combined fit is limited to the gain values with $gain > 0$. This takes the asymmetric noise behaviour into account and neglects that part during the fit. Anyway, negative areas below the identified peak, point to a miss-reconstructed S1 event. The outcome of the fit is also shown in figures 4.20a, 4.20b, 4.20c and 4.20d. Here the pre-fits are drawn in grey-red and grey-blue, as well as the resulting combined fit (green). The Poisson and Gaussian share of the combined fit are also shown by the blue dashed line (noise) and red dashed line (signal). In addition, the fit information of the combined fit is also shown. The most interesting part is again the Poisson expectation value λ and the most likely $\hat{\lambda}$.

4.9.2. Discussion: Fit Parameters of the Gain Histograms

This section discusses the outcome of the combined fit with six parameters: A_G, A_P, μ_G, λ, σ_G and σ_P for the gain histograms. Examples for some of the fit outcomes are shown in figures 4.20b, 4.20c and 4.20d. The previous subsection 4.9.1 describes the asymmetric noise distribution in the calculated gains from *xerawdp*. Therefore this subsection will focus on the outcome of the Poisson fit parameters: A_P, λ, σ_P. These are shown in figures 4.21, 4.22a and 4.23a.

The parameter A_P scales the distribution up to the number of counted SPE from the alternative gain calibration. Figure 4.21 shows this parameter for each single PMT. A similar step pattern is already observed in the analysis of the $S1s[pmt][i]$ histograms (compare with figure 4.8b). The origin of this pattern is again the spatial distribution of the PMTs at the top and bottom array. The number of counted SPE peaks is different for the PMTs 1 to 30 (outer PMT ring on the top array), PMTs 31 to 98 (top array) and the left PMTs from the bottom array.
The Poisson distribution according to equation (4.6) is defined by its expectation value λ. This value is shown for each PMT in figure 4.22a and its distribution is shown in figure 4.22b. The error bars are

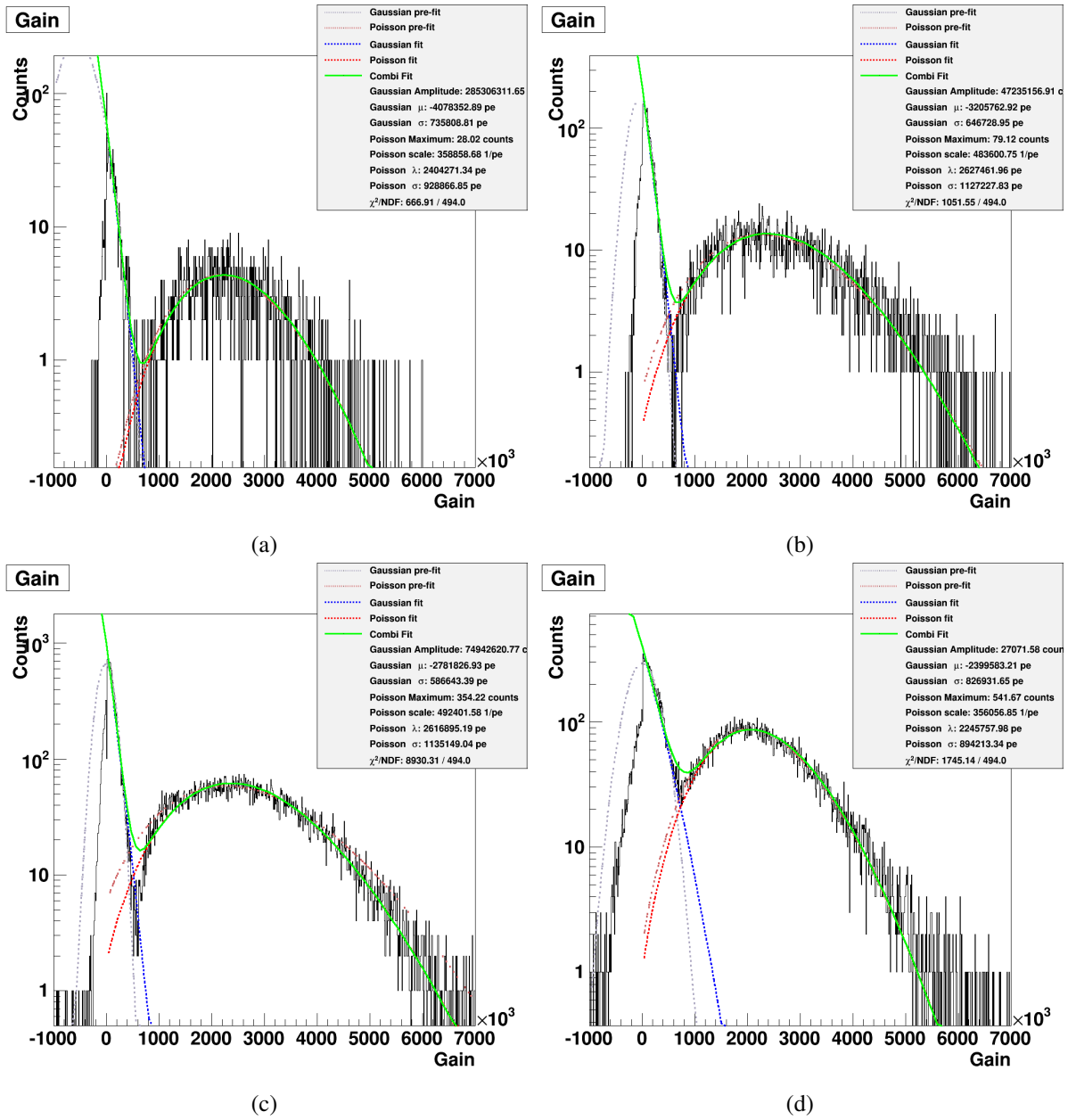


Figure 4.20 – A choice of different gain distributions: PMT (top array) 13, 93 and (bottom array) 110 and 164. Each figure shows the combined fit in green and the pre-fits (dotted lines). The single contribution (Gaussian or Poisson) of the combined fit are drawn in dashed blue and dashed red. The gain itself is calculated according equation (4.12) with the integral of the identified S1 peak.

a result of the fit uncertainty for the λ parameter. A trend to lower λ on the bottom array is observed. This behaviour was already observed in the $S1s\,[pmt]\,[i]$ histogram analysis, especially for PMT 147 and higher. These PMTs sometimes show noisy behaviour during the regular PMT calibrations. The

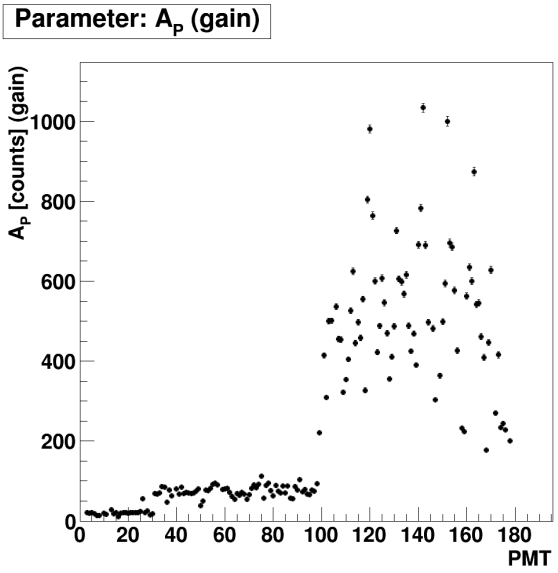
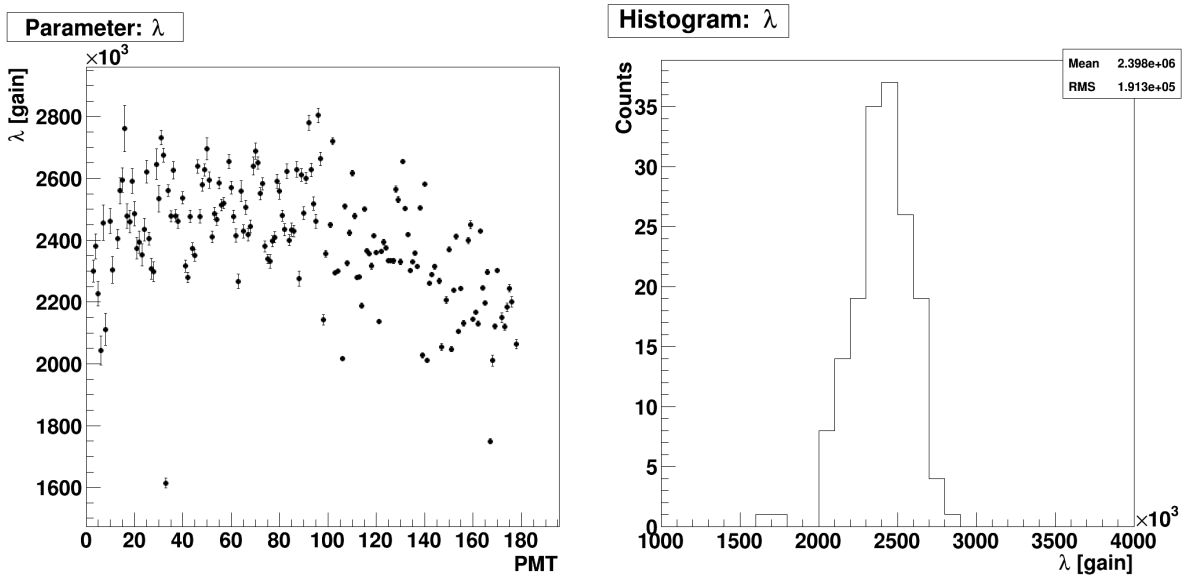


Figure 4.21 – Parameter A_P from the the combined fit with equation (4.6) describes how the Poisson distribution is scaled up due to the higher counting rate which depend on the PMT position: 1) An outer ring of 30 PMTs on the top array. 2) The inner PMTs on the top array. 3) The PMTs on the bottom array.

uncertainty from the fit also decreases for PMTs on the bottom array. This is interpreted as a side effect of the number of counted events. In comparison, also the parameter A_P shows a huge increase. This is related to a higher number of counted photons. The increased number of counted events improves the fit parameter quality. For all PMTs together the mean value is $\lambda^{mean} = 2.398 \times 10^3$ with a $RMS = 0.191 \times 10^3$. Based on the information about the fit parameters for the Poisson distribution, the fits to the gain spectra work quite well.
In addition, the Poisson distribution is scaled with a width parameter σ_P. The fit outcome for σ_P is shown in figure 4.23a and its distribution in figure 4.23b. The alternative calibration shows a good agreement in the width of the Poisson distributions for all PMTs. Its mean value is $\sigma_P^{mean} = 0.432642 \times 10^6$ with a $RMS = 0.089 \times 10^6$. Nevertheless, this parameter is not used in the further analysis.

4.9.3. Compare the Gain Calibration Methods

As already discussed in subsection 4.8.2, the Poisson distribution is characterised by its expectation value λ, but due to the fact that the expectation value is not similar to the maximum of the distribution (= most likely value $\hat{\lambda}$) the gain characterisation is better described by $\hat{\lambda}$. Figure 4.24a shows the values of $\hat{\lambda}$ for each PMT. There, the value of $\hat{\lambda}$ is always smaller than the expectation value λ by a certain offset. This offset depends on the expectation value itself.



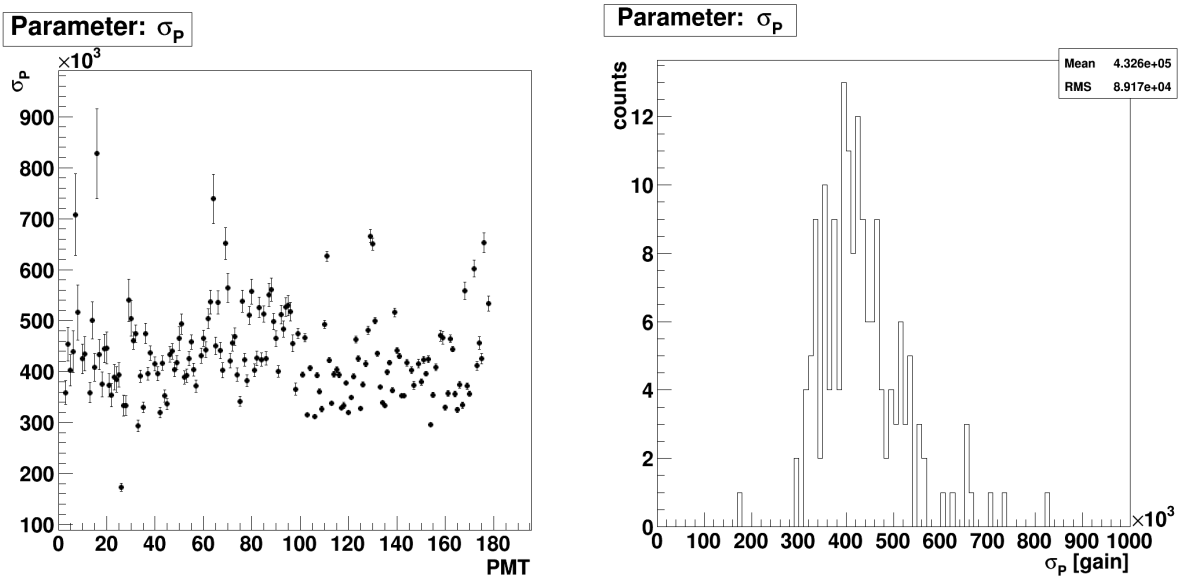
(a) (b)

Figure 4.22 – Both figures describe the expectation value λ from equation (4.6) for the combined fit of the gain PMT spectra. λ characterises the Poisson distribution. The resulting λ shows smaller fit errors for the bottom array compared to the top array. The counting rate is higher on the bottom array which results in a higher statistics in the according gain spectra. For all PMTs together the mean of $\lambda^{mean} = 2.398 \times 10^3$ with an $RMS = 0.191 \times 10^3$ is observed.

Besides the smaller value of $\hat{\lambda}$ for each PMT, compared to λ, the shape and behaviour stays the same as for the fit parameter λ (see figure 4.22 for comparison). The mean value according to 4.24b is $\hat{\lambda}^{mean} = 2177 \times 10^3$ and $RMS = 0.181 \times 10^3$.

The result from the single PMTs becomes interesting to finally verify the method of using the peak-finder algorithm of *xerawdp*. The $\hat{\lambda}$ is comparable to the gain values from the standard PMT calibration. Therefore the same set of LED files is chosen from the weekly calibration which was already reprocessed with the standard LED analysis software. Each PMT has a gain value from the standard LED analysis g_{LED} which can be compared directly to $\hat{\lambda}$. The only constrain is that $\hat{\lambda}$ is developed from all ten chosen LED raw data files together while the standard PMT calibration software provides g_{LED} for each data set separately. Hence a simple mean value is calculated out of g_{LED} which is named g_{LED}^{mean}.

The results are compared directly in figure 4.25a. The black dots mark the result of the alternative gain calibration with $\hat{\lambda}$ and the blue dots indicate the outcome of the standard gain calibration of ten LED raw data files with g_{LED}^{mean}. Both gains show a similar outcome and also the slight decreasing gain values on the bottom array is observed. The mean value $g*_{LED}^{mean}$ is calculated for all involved PMTs from g_{LED}^{mean}. This leads to $g*_{LED}^{mean} = 2049 \times 10^3$ with $RMS* = 0.216 \times 10^3$ which shows a pretty good



(a) (b)

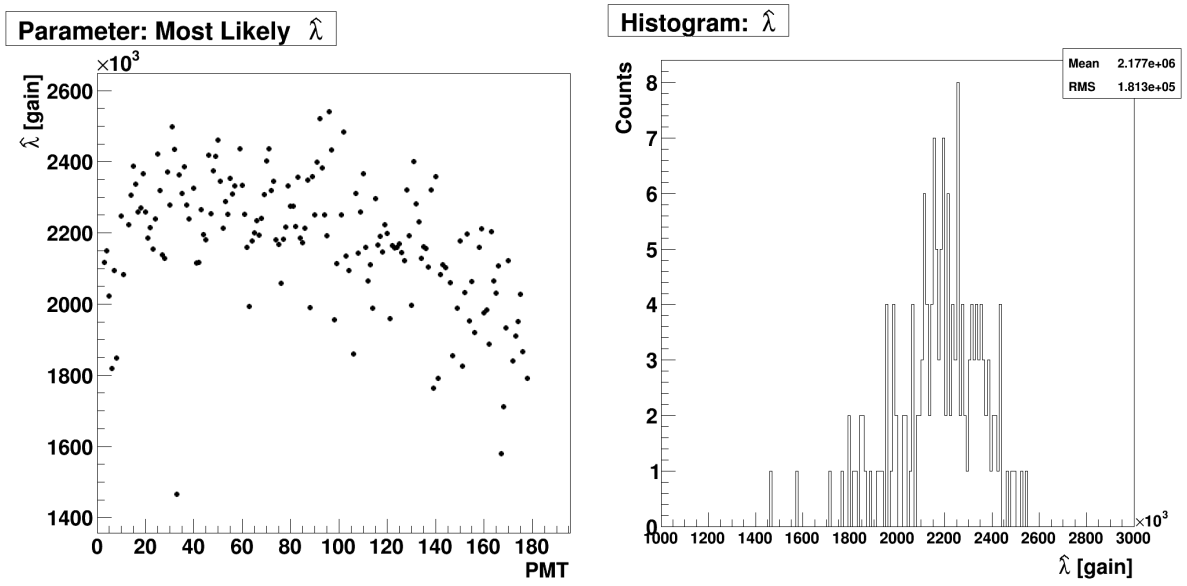
Figure 4.23 – Besides the expectation λ of the Poisson distribution there is another fit parameter which is introduced to increase the quality of the fit: σ_P adjusts the width. The combined fit results in a $\sigma_P^{mean} = 0.433 \times 10^3$ with a $RMS = 0.089 \times 10^3$ for all PMTs. Again the fit errors of the combined fit are smaller for the bottom array.

agreement with $\hat{\lambda}$ from the alternative gain calibration. The deviation in percentage is given in figure 4.25c for each PMT and in figure 4.25d as a distribution. The deviation Γ is calculated simply by equation (4.13). From figure 4.25d, it is observed that the alternative gain $\hat{\lambda}$ is larger than g_{LED}^{mean} in most of the time with a mean of $\sim 6.9\%$. Due to the broad RMS values of figure 4.25d the both gain calibrations overlap.

$$\Gamma = \frac{\hat{\lambda} - g_{LED}^{mean}}{g_{LED}^{mean}} \times 100 \tag{4.13}$$

4.9.4. Uncertainties: Standard vs. Alternative Gain Calibration

Based on the outcome of subsection 4.9.3 the deviations between the standard and the alternative gain calibration fit pretty well to each other. Only the statistical uncertainty on the overlap of both calibration methods is not yet done. From the overlap of both PMT calibration methods, it is expected that the uncertainty of the gain with a 1σ deviation from the Gaussian approximation of the standard PMT calibration is comparable to a similar deviation from $\hat{\lambda}$.
As long as the SPE response is still described by the most likely value of $\hat{\lambda}$, it should be possible to



(a) (b)

Figure 4.24 – Besides the characterisation of the Poisson distribution itself with the parameter λ, the SPE response which is described as the most likely value for the SPE from the gain calibration is $\hat{\lambda}$. This parameter is always smaller compared to λ, but $\hat{\lambda}$ follows the same trend as λ. Its distribution for all involved PMTs shows a mean value of $\hat{\lambda}^{mean} = 2177 \times 10^3$ and $RMS = 0.181 \times 10^3$. It is expected that the most likely value will fit the gain values from the standard calibration.

calculate a 1σ deviation for this case. This is done according sketch 4.26a. Around the most likely gain value from *xerawdp*, a 68.3% (1σ) interval is calculated. Due to the Poisson property, the resulting uncertainty will be asymmetric. The asymmetric deviations together with $\hat{\lambda}$ is compared to g_{LED}^{mean}, the mean value of ten expectation values from ten LED calibration files. Each of the Gaussian fits to the SPE response in the ten LED data files also has a 1σ deviation, the Gaussian σ (compare figure 3.29b, here the Gaussian $\sigma_G = 1.06 \times 10^6$). There is also mean value calculated out of the ten σ_G values which is then used as the 68.3% limit (σ_{LED}^{mean}) for the comparison of both calibration methods. By construction, that fit range will be symmetric around the gain values from the standard PMT calibration.

The outcome is presented in figure 4.26b. As before, g_{LED}^{mean} and $\hat{\lambda}$ overlap and now in addition the σ_{LED}^{mean} deviation is shown in blue. This symmetric deviation for each PMT is included in the 1σ deviation which is calculated from the alternative PMT calibration (black). The Poisson deviation is, as expected, asymmetric around the calculated gain values, but nevertheless it includes the deviation $\sigma > _{LED}^{mean}$ which indicates that both methods overlap within 68.3% and the gain values which are calculated out of both methods are reliable.

Unfortunately, the overlap itself does not tell anything how good both methods match each other in detail. The reason comes from strong asymmetric behaviour which is shown in figure 4.26b. The overlap to the lower values ($\hat{\lambda} - 1\sigma$) overlaps pretty good with the mean standard deviations (σ_{LED}^{mean}) from the standard PMT calibration. The deviations increase asymmetric by a factor two or more onto the other site. In that way, the 1σ deviations increase "to strong" and both deviations are not comparable any more directly. The reason is in the peak-finder algorithm of *xerawdp*. As already shown in section 4.4, the *xml* file influences the peak-finder and therefore more than one peak is found on the summed waveform. Each peak refers to a left and right boundary. This is a basic difference in the conception of the standard LED calibration. There, the trigger system send a signal to the LED and to the DAQ with the result that a waveform is recorded for each PMT, nevertheless if the PMT observed a signal. Finally, the trigger window is integrated and this information is used to determine the gain. The alternative gain calibration allows to observe more signals with different strength on the same summed waveform and hence on each single PMT. Following the design of the peak-finder algorithm, the shown gain histograms will be influenced by counting peaks of different strength from the PMT waveforms. This effect broadens alternative gain.

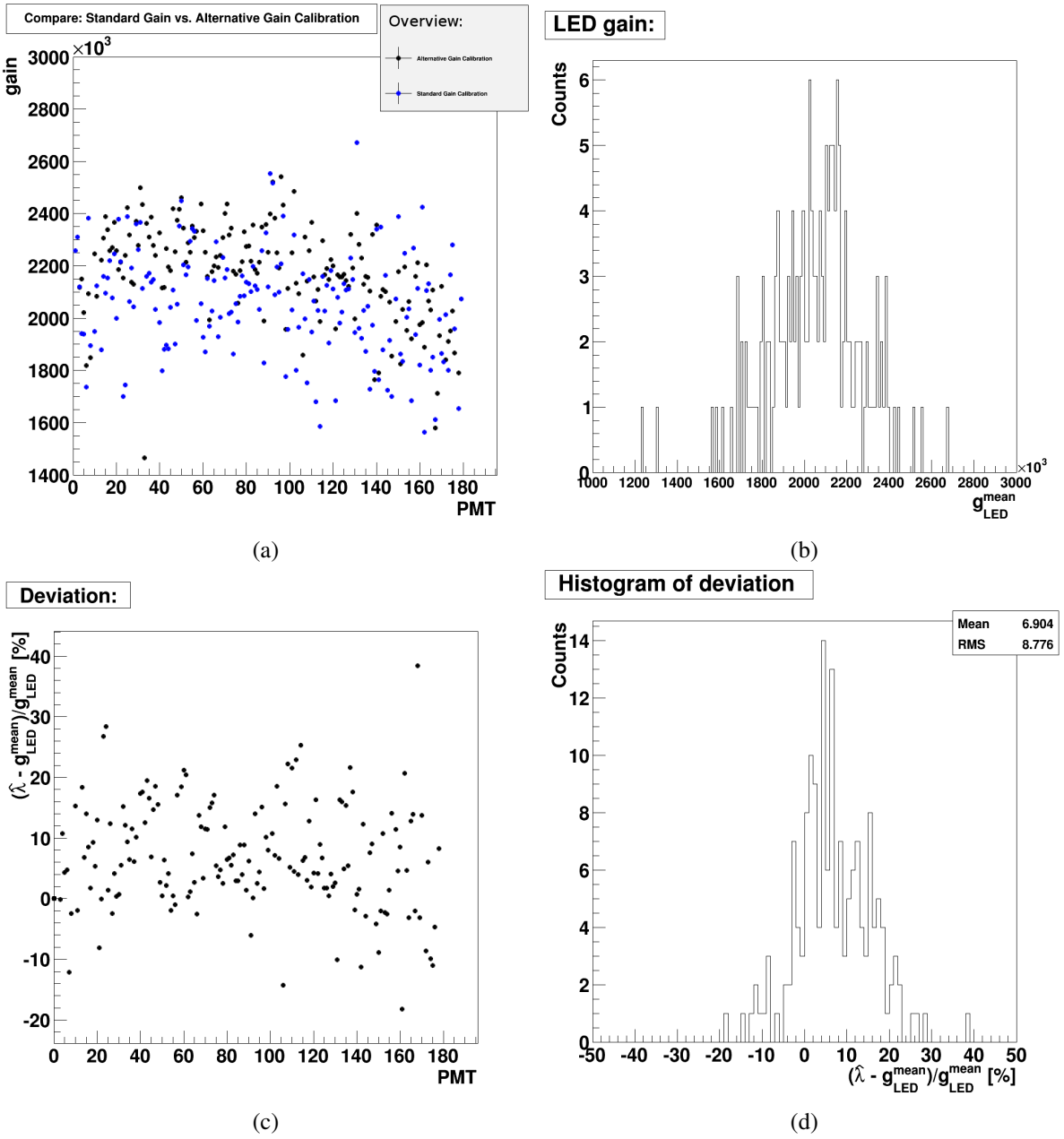
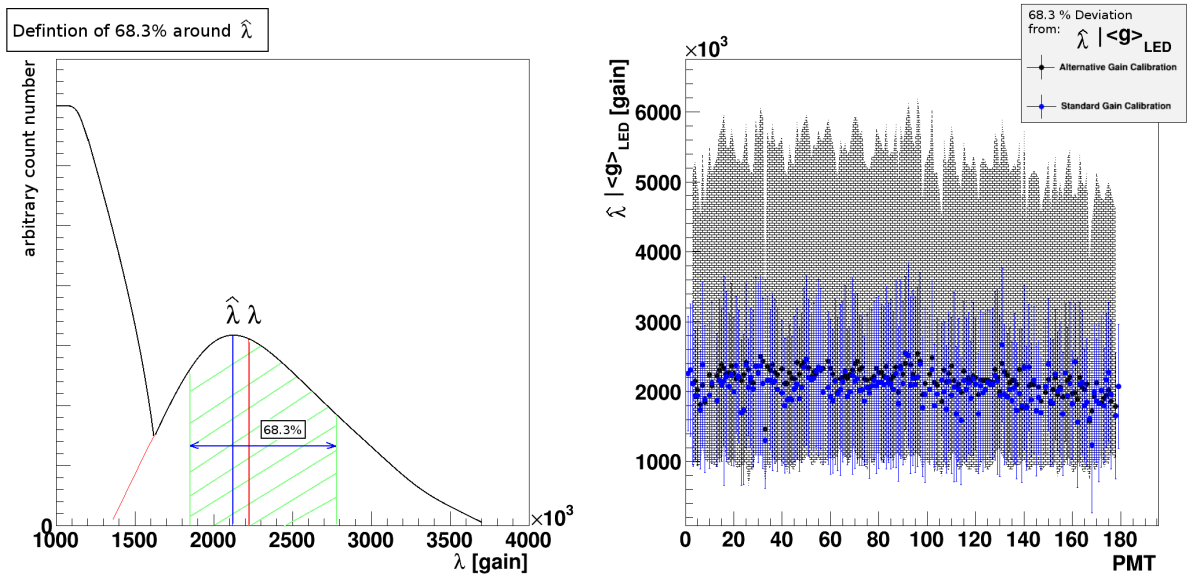


Figure 4.25 – These figures compare the outcome of the standard PMT calibration information (always in blue) with the results from the alternative gain calibration for the SPE response. For this method the most likely value $\hat{\lambda}$ describes the outcome of the SPE response (gain). Figure 4.25a shows the gain values from calibration methods. The distribution of $\hat{\lambda}$ is already presented in figure 4.24b. Therefore figure 4.25b presents a histogram of the gain from the LED calibration (g_{LED}). The tail to the left side comes from some PMTs which show a lower gain compared to the mean value of $\hat{\lambda}^{mean} = 2049 \times 10^3$. These PMTs are widely distributed and not necessarily restricted to one PMT array. In figures 4.25c and 4.25d $\hat{\lambda}$ and g_{LED} are compared directly. There is a trend that the gain determination from the standard PMT calibration leads to $\sim 7\%$ higher values than obtained from the alternative method but due to the wide RMS, the overlap works pretty well.



(a) (b)

Figure 4.26 – To determine the uncertainties from both, the standard and the alternative gain determination method, the 1σ intervals are evaluated. For the alternative gain calibration, the 1σ range is defined by an asymmetric interval around $\hat{\lambda}$. This is chosen to take into account that the SPE response is characterised by that parameter instead of λ. This is compared to the mean value of the σ_G from the ten LED calibration data sets. These symmetric intervals are shown in blue in figure 4.26b. Both 1σ deviations overlap.

4.10. Summary and Outlook

This section presented a detailed description about reprocessing LED raw data with *xerawdp*. Therefore the necessary steps are explained: Adjusting the *xerawdp* software code and creating a new *xml* file for the LED calibration. The basic idea of the successive analysis is to verify that the peak-finder algorithm is able to identify enough S1-like peaks from the summed waveform of 178 PMTs, instead of analysing each waveform separately. After the peak-finder algorithm is carried out the signals are stored in the same data format as it is done with e.g. Dark Matter data.
First the $S1s[pmt][i]$ spectra of each PMT are investigated. To create such a spectrum, the gain information of the standard PMT calibration is necessary during the raw data reprocessing. In case the peak-finder algorithm works well enough and the set of pre-defined cuts is able to select the single-photo electrons correctly, the SPE response can be determined for each PMT. The results presented here show that a fit with an adjusted Poisson distribution describes the SPE response pretty well and the expected SPE response is $\sim 1\,\mathrm{pe}$. A Gaussian fit is applied to describe the noise part of the $S1s[pmt][i]$ spectra. The combined fit yields a local minimum between the noise and signal part which can be interpreted as a variable S1 threshold for each PMT. The analysis on the $S1s[pmt][i]$ histograms presented here shows a signal acceptance of around 95%. On the the other hand, the contamination of the signal by the Gaussian noise is also estimated. It turns out, that for most of the PMTs the noise contamination is less than 1%. Some PMTs show an increase up to 14%. This behaviour is mainly observed on the bottom array with PMTs which are marked as noisy sometimes in the LED calibration.
Besides the direct $S1s[pmt][i]$ analysis, the gain is determined directly from the found S1 peaks. Therefore *xerawdp* identifies S1 peaks and estimates their boundaries. The total area of the peak is calculated and this information is used to calculate the gain. A successive analysis compares the values of the standard gain calibration and the alternatively calculated gains. It turns out that both gains are similar to each other and they overlap within a 68.3% band around their expectation values (compared PMT by PMT).
These results present the status of the LED raw data analysis with *xerawdp*. The two analysis methods and their results, discussed in this section are both possible applications for the XENON100 data processor or the successive data analysis software PAX in XENON1T. The analysis of the $S1sRaw[pmt][i]$ histograms makes it possible to determine the gain directly from the found peaks and furthermore the analysis of the $S1s[pmt][i]$ histograms supports the calibration procedure.
But besides modifying *xerawdp* for LED data analysis, it is also possible to analyse the SPE response with $S1s[pmt][i]$ histograms from Dark Matter or other calibration data such as ^{232}Th. In any case it can be tested, depending on the source rate, if additional SPE response analysis will support the weekly calibration of the PMTs apart from the LED calibration.

5. Determination of the S2 Signal Acceptance

"If there's any *Dark Matter* within 500 kilometres, you'll know it."

Captain Jonathan Archer,
Star Trek: Enterprise, "First Flight"

5.1. Motivation

This chapter discusses the S2 signal acceptance. The XENON100 TPC is triggered by the S2 signal resulting from the upwards drifting electrons, regardless of the interaction depth in the detector. The depth is determined from the time difference between the S1- and S2-signal (drift time). The S2 signal is proportional to the deposited charge at the interaction site and hence a correction, due to the drift time dependent charge loss, is necessary to determine the "corrected S2" (cS2) signal. Details to the S2 correction are given in section 3.4.4. In order to to work with a trigger efficiency of $\sim 100\%$, a cut of

$$S2 > 150\,\mathrm{pe}$$

is introduced named the *Xs2peaks* cut. Detailed information about the most important cuts are given in section 3.7. The origin of the S2 signal is discussed in section 3.4.4. Due to the charge loss along the drift direction, the electron cloud and hence the proportional S2 signal, is reduced exponentially. This process is described by equation (3.18). This step is necessary according to correct the S2 signal to its depth in the TPC.
The following work is motivated by the S2 signal acceptance which is calculated for a S2 threshold energy of 150 pe. This threshold is chosen according to the trigger efficiency which is discussed in section 3.5.1. Following the S2 trigger probability, the final S2 acceptance is important in the data

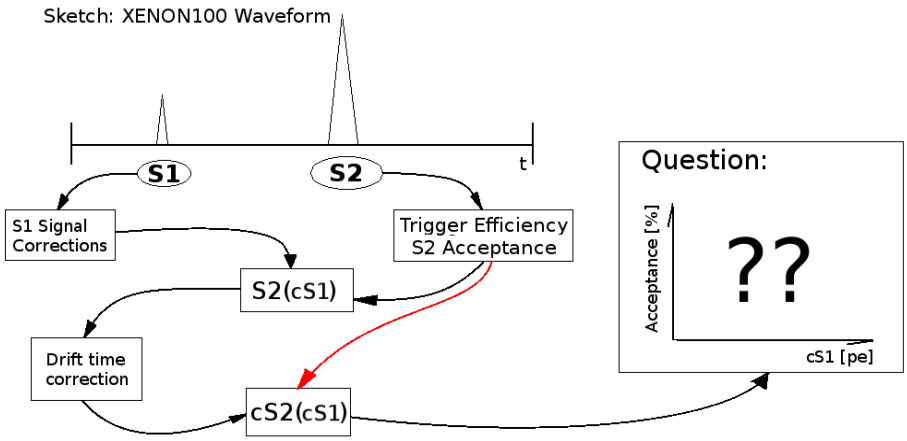


Figure 5.1 – Based on the signal creation process, the S2 signal is selected according to S2 trigger efficiency and following the S2 energy threshold. The final result is a drift time corrected S2 signal which dependence indirectly on the S2 energy threshold. Due to the reason that the cS2 signal is used in the data analysis approach, it is then necessary to determine the S2 acceptance from the uncorrected S2 signal. The black arrows symbolise the applied S1 and S2 corrections to the observed signal. The red arrow symbolise the transition from uncorrected to corrected signals where the S2 acceptance is hidden.

analysis procedure to determine the probability to have an S2 signal above 150 pe for the related S1 signal. The S2 energy threshold is the central issue here. The usual XENON100 data analysis does not provide information below 150 pe. On the other hand, the related S1 signal correlates with the S2 signal and, in general, a larger S1 signal results a larger S2 signal. This introduces a S2 signal which dependent on S1 and allows to define the S2 acceptance according to the S1 signal itself. As a consequence, the corrected S2 signal dependence on the S2 trigger efficiency indirectly. This situation is described in figure 5.1 where a S1 and S2 is detected on a XENON100 waveform. The S2 alone is recorded under the S2 trigger efficiency boundary condition (Xs2peaks cut). The resulting S2 signal is connected to the previous S1 signal in the TPC: S2(cS1). Necessary S1 corrections are shown in the sketch. The S1 correction is discussed in section 3.4.3. The drift time correction allows to calculate the corrected S2 signal which is used in the analysis approach. The red line indicate the indirect dependence of cS2 from the trigger efficiency.
But for all that, it is necessary to model the S2 response below 150 pe to complete the understanding of the S2 acceptance, in particular for low S1 signals. This thesis introduces a phenomenological description of the cS2 signal spectrum by an adjusted Poisson distribution to access the S2 signal acceptance below 150 pe and to calculate the total S2 acceptance with a Monte Carlo technique.
The S2 signal acceptance is determined by neutron calibrations from before and after the Dark Matter data taking in 2011/12 and a third later neutron calibration in 2013/14. The S2 signal acceptance is then part of the overall cut acceptance in the data analysis approach [15].
To validate the applied method, a comparison with simulated neutrons is done in the end. For this, the S1 and S2 signal generator tool (*Nuclear Recoil/WIMP Simulation Code*) from section 3.4.6 is used. Finally, the same tool is used to simulate WIMP interactions in the XENON100 TPC and to determine their simulated S2 signal acceptance according to the WIMP mass.

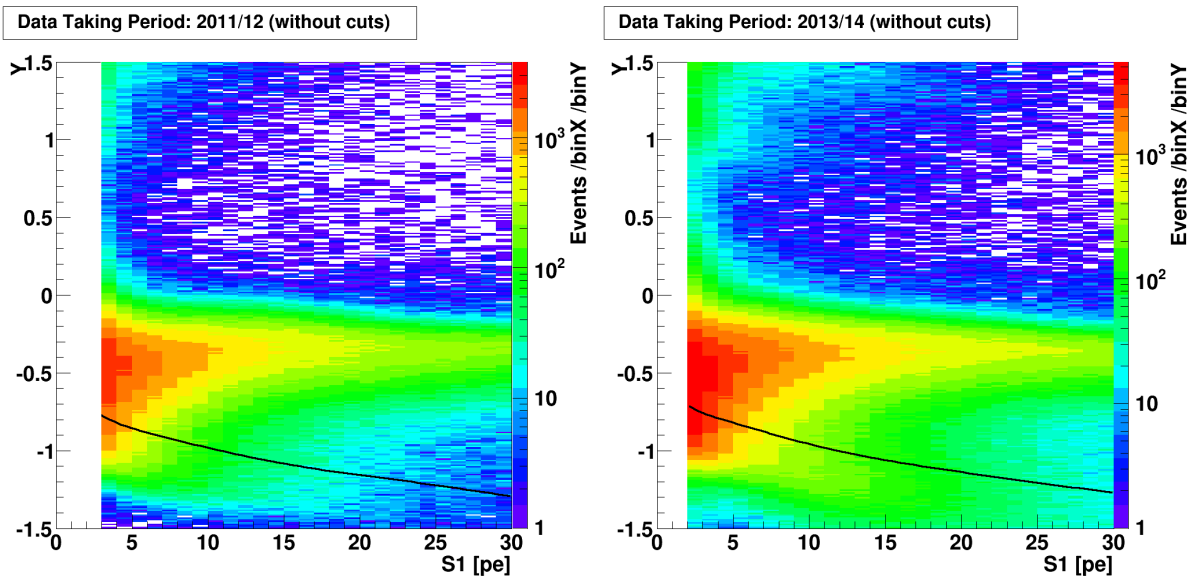
5.2. Data Selection

In the following the S2 acceptance is determined from data. Therefore three data sets are used and will be named as *February11*, *April12* and *May13*. They refer to three data taking periods for ^{241}AmBe: In February 2011 with a period of $\sim$ 2 days, in April 2012 with a period of $\sim$ 4 days and in May 2013 with $\sim$ 6 days. These periods were scheduled before and/or after Dark Matter taking periods to have a continuous operation between 2011/12 and 2013/14. The detector calibration is important to determine the TPC response to neutrons, their energy deposit (nuclear recoil energy) is used to calculate the conversion from measured photo-electrons (pe) to electron-volts (eV). Detailed information is given in section 3.6.3 where the calibration sources are discussed and furthermore in section 3.4.5 where the energy calibration is discussed. All ^{241}AmBe data sets considered here were taken under stable

detector conditions and with the same experimental setup in each data taking period. The data sets February11 and April12 were taken at stable voltage conditions of $V_{\text{Anode}} = +4.4\,\text{kV}$ for the anode voltage and $V_{\text{Cathode}} = -16\,\text{kV}$ for the cathode voltage. For the May13 data sets, the cathode voltage was set to $V_{\text{Cathode}} = -15\,\text{kV}$. The anode voltage did not change. Table 5.1 summarises the data taking periods.

Data set	Number of files	corrected lifetime [h]	corrected number of events	V_{Anode}[kV]	V_{Cathode}[kV]
February11	15	26.3	2389896	+4.4	-16.0
April12	18	44.9	2607102	+4.4	-16.0
May13	63	109.9	9061976	+4.4	-15.0

Table 5.1 – Overview of the selected ^{241}AmBe data sets in data taking periods February11, April12 and May13. All data sets within a data taking period are selected according to stable detector conditions and with the same experimental setup.



(a) (b)

Figure 5.2 – Both figures show the accumulated NR data in 2011/12 and 2013/14 in the discrimination space. Besides a minimum requirement of minimum one S1 and one S2 signal. The black line *symbolise* the Xs2peaks cut in the flattened parameter space. Besides the S1 threshold of 3 pe (2011/12) and 2 pe (2013/14), there are no further cuts used in this presentation. This NR data present the basic input to the S2 acceptance calculation.

The following analysis relies on the applied cut sets to select neutron interactions in the TPC. Table 5.2 summarises the cut sets for the February11, April12 and May13 data taking periods. Both ^{241}AmBe data sets February11 and April12 were taken at the start and end of 2011/12, therefore the applied cut sets are the same. Basically both cut sets (2011/12 and 2013/14) contain the same kind of data selection criteria, but as long as most of the cuts are developed according to the data taking period, some of the cuts are altered. Details about the meanings of the listed cuts are given in section 3.7.
A minimum selection is done by taking only single S1 and S2 peaks into account and by choosing the inner 48 kg volume as analysis volume of the TPC (X48kg0). This volume is furthermore reduced to select only events which originate from the upper third of the TPC by selecting events according to their drift time from $5\,\mu s \leq \Delta t \leq 30\,\mu s$. The drift time selection overlaps with the X48kg0 cut. Selecting events only from the upper third of the TPC assures a low S2 drift time correction and hence the difference between S2 and cS2 becomes small. This is important to minimise a possible bias of the data analysis resulting from the drift time correction. The spatial selection from the inner TPC by radius and X48kg0 cut remove events near the edge of the TPC for a good reasons: The shape of the electric field lines becomes curved with a larger radius and hence the position reconstruction in x/y is biased. To avoid any bias in the analysed cS2 histograms, there is a clean data set selected.

Reason	Cut set (February11/April12)
S1-Signal	Xs1coin2, Xs1single4
S2-Signal	**Xs2peaks**, Xs2single3, Xs2pmtorder0, Xs2asym0, Xs2width8, Xs2top0
Noise reduction	Xsignalnoise3, Xhighlog1, Xlownoise0, Xveto2
Position selection	X48kg0, Radius$<$ 145 mm, $5 \leq \Delta t \leq 30 \mu s$

Reason	Cut set (May13)
S1-Signal	Xs1width0, Xs1coin2, Xs1single5
S2-Signal	**Xs2peaks**, Xs2top0, Xs2single3, Xs2width11, Xs2chisquare1, Xs2peakpos0
Noise reduction	Xsignalnoise5_corrected, Xentropy1, Xhighlog1, Xveto2, XPL013_97, Xposrec1, Xlownoise0_m
Position selection	X48kg0, Radius$<$ 145 mm, $5 \leq \Delta t \leq 30 \mu s$

Table 5.2 – Cut selection for the ^{241}AmBe data files. February11 and April12 belong to 2011/12, hence no different is made in the final cut set for the 2011/12 data taking period. In 2013/14 further cuts were applied to reduce the noise contribution to the signal. In both cases, an inner volume is selected by the X48kg0 cut which is reduced furthermore more to the upper one third by a drift time cut. The additional radius cut overlaps with the previous applied cut X48kg0. The applied Xs2peaks cut is worded in bold.

5.3. Method to Evaluate the S2 Signal Acceptance with ^{241}AmBe Data

In general, the analysis is done with the corrected signals. The same is done here with the corrected S2 signal (named: "cS2sTot[i]" in the data sets). The maximum number of peaks is restricted to 32 by the peak-finder algorithm of *xerawdp*. The uncorrected S2 is called "S2sTot[i]". The S2 peaks are ordered by the peak size. The S1 and S2 signals are stored in different correction levels in the XENON100 data. The S1 signal is corrected to the light collection efficiency (LCE) and its 3D position in the detector. The name of this corrected variable is "cxS1sTot[i]". Again the index i stands for detected peaks on the raw data waveform. They are ordered by coincidence level and then by the peak size. This analysis will focus on the largest peaks in S1 and S2 ($i = 0$) and therefore a simpler nomenclature is introduced in table 5.3. With this nomenclature, the S2 acceptance calculation starts by applying the

cxS1sTot[0]	$\rightarrow$	S1
cS2sTot[0]	$\rightarrow$	cS2
S2sTot[0]	$\rightarrow$	S2

Table 5.3 – Nomenclature for the used variables in the data with their abbreviations in this section.

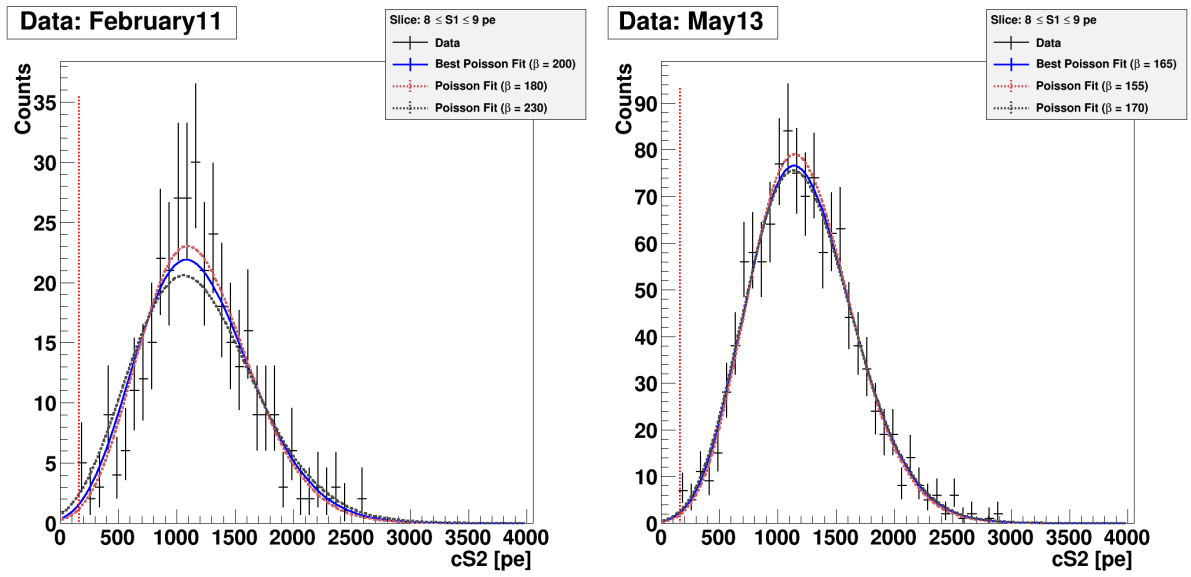
cut set to the ^{241}AmBe data. An overview about the accumulated data is given in figure 5.2a and 5.2b where no further cuts are applied yet. The black line symbolise the S2 acceptance in the discrimination space Y but be aware of the different definition of Y (see section 3.9) and the Xs2peaks cut.
The data is sliced in 15 slices of 1 pe width in S1 to give them a proper Poisson description and in a second step the acceptance is determined by a Monte Carlo method.

5.3.1. Poisson Description of the ^{241}AmBe Data

Before the cut set of table 5.2 is applied to the ^{241}AmBe data of 2011/12 and 2013/14, a huge number of events are in the data. Their distributions are shown in figures 5.2a and 5.2b in the usual flattened data space. The cut selection of table 5.2 reduces these events and remove events which, for example, results from noise or consist of double scatter events.
After the cut set is applied the distribution of the ^{241}AmBe data the S1 space is sliced with $\Delta S1 = 1\,\text{pe}$. Since the Xs2peaks cut affects the acceptance only for low S1, we will restrict the discussion to S1 signals up to 16 pe. The acceptance is already at 100% for S1 signals with $S1 > 16\,\text{pe}$. Each slice of S1 is presented by a histogram of cS2 values. The bin width of each cS2 bin is set to 75 pe, 80 bins span a range of $[0, 6000]$ pe in each S2-histogram. Figure 5.3a shows one of these histograms for the February11 data set and figure 5.3b for the May13 period. Both S2 histograms are chosen to show a S1 slice of $8\,\text{pe} \leq S1 \leq 9\,\text{pe}$. An important information is that these histograms are not filled with

entries down to 0 pe in cS2. According to the S2 correction procedure and the applied *Xs2peaks* cut this energy region of data space in S2 is not accessible. Here we try to estimate the missing number of events below the threshold with a model of the spectral shape in cS2 for each S1 slice which we call "adjusted Poisson" distribution.

By definition the Poisson distribution is only valid for integer numbers n (see equation (5.1)). But the cS2 information is defined in $\mathbb{R}$. For that reason the Poisson distribution is generalised for real numbers by transferring the factorial $n!$ into the Gamma function $\Gamma(n+1)$. The generalised Poisson distribution is then adjusted by further parameters α and β. Parameter α is used to introduce a free multiplication parameter to fit the number of events in each histogram. Parameter β allows to scale the width of the Poisson distribution (x-axis direction). The resulting Poisson distribution is expressed in equation (5.2). The expectation value λ is scaled by a factor 100. In the following, we will use only this adjusted Poisson function, which will be called "Poisson distribution" briefly.



(a) AmBe data: Run10

(b) AmBe data: Run12

Figure 5.3 – Both figures show the S1 slice of $8\,\text{pe} \leq S1 < 9\,\text{pe}$ for two different data taking periods: February11 (5.3a) and May13 (5.3b). Both contain the lower fit boundary (red dashed vertical line) from the drift time corrected S2 energy threshold and the result of the fit with the adjusted Poisson distribution. The boundaries depend on the maximum drift time of $\Delta t = 30\,\mu\text{s}$ and the lowest observed electron lifetime and hence the fit boundary of February11 and April12 is 163 pe and in May13 its 159 pe (see section 5.3.2). The blue line shows the fit (see equation (5.2)). The black and red dashed Poisson distributions show the fit results for the low and high confidence bound in β. The uncertainty on the histogram entries is calculated by the asymmetric Poisson errors.

$$P(n|\lambda) = \frac{e^{-\lambda}\ \lambda^n}{n!} \tag{5.1}$$

$$P(cS2|\alpha,\beta,\lambda) = \alpha\ \frac{e^{-\lambda/100}\ \left(\frac{\lambda}{100}\right)^{\frac{cS2}{\beta}}}{\Gamma\left(\frac{cS2}{\beta}+1\right)} \tag{5.2}$$

$$\lambda' = \frac{\beta}{100}\lambda \tag{5.3}$$

Using equation (5.2) during the fit procedure leaves two free parameters α and β which are determined by the fit. The Poisson expectation value λ which characterises the Poisson distribution for a single slice of S1 is related to β by the definition. The parameter β scales with the width of the Poisson distribution. This relation is defined in equation (5.3) and describe the fact that a larger value of λ needs a smaller value of β and vice versa to fit the same data distribution as good as possible. Hence λ' is not unit-less any more. It is measured in photo-electrons and is interpreted as *true* expectation value of the cS2 signal. The parameter α multiplies the adjusted Poisson distribution by a certain value to measure the number of counted events correctly. Due to the definition of equation (5.2), parameter α becomes only necessary in case the cS2 distribution is normalised, which is not necessary in the following data analysis.

5.3.2. Fit Procedure and Minimisation

Equation (5.2) is the input for the fit procedure of each S1 slice. Besides a given set of start parameters for the Poisson fit (α, β and λ), valid fit boundaries are needed which describe the physics of the TPC correctly (red dashed line of figures 5.3a and 5.3b). It is necessary to set the left boundary to the charge loss corrected cS2 signal. As described in section 3.4.4, drifting electrons are lost within a lifetime τ. The S2 acceptance is determined from the upper part of the detector which is defined by the drift time Δt. This analysis limits the maximum drift time to $\Delta t \in [5;30]\,\mu$s. The minimum value of $\Delta t = 5\,\mu$s is chosen to select only events below the gate mesh to avoid events close to the gaseous phase. Events from the gaseous phase have are related to a large S2 signal in the TPC and hence these events bias the cS2 histograms. The S2 energy threshold of $S2_{th} = 150\,$pe is then adjusted by equation (5.4) to $cS2_{th}$. In these three data taking periods during 2011/12 and 2013/14 the drift time range is not changed. However the electron lifetime τ changed in between due to the improvement in xenon purity. For February11 and April12 the lowest observed electron drift time was $\tau = 354\,\mu$s and for May13 data it was $\tau = 517\,\mu$s. The lowest electron lifetime is determined from the February11 data taking period. As long as both ^{241}AmBe calibrations (February11 and April12) were done during the

same 2011/12, the detector conditions are stable. The resulting corrected thresholds are calculated according formula (5.4).

$$cS2_{\mathrm{th}} = 150\,\mathrm{pe} \times e^{\Delta t/\tau} \tag{5.4}$$

For April12 data the threshold is $cS2_{th} = 163\,\mathrm{pe}$ and due to its higher electron lifetime, the threshold in the May13 data set is $cS2_{th} = 159\,\mathrm{pe}$. The difference between both left boundaries is 4 pe, is much smaller than the chosen bin size ($\Delta cS2 = 75$) of such a histogram of S1 slices in figures 5.3a and 5.3b. This outcome shows the choice of $\Delta t = 5\,\mu\mathrm{s}$ as maximum drift time. The difference between the corrected and uncorrected S2 is less than $\sim 9\%$. Under this condition, the drift correction has less impact on the S2 threshold. A complete set of figures for all S1 slices is shown in appendix A.2. The right boundary during the fit is set to $cS2 = 5000\,\mathrm{pe}$ which is far away from the cS2 distribution of each low S1 slice.

The fit is done by ROOT (TMinute class) [37] within the set boundaries. The outcome of such a fit procedure is also shown in figure 5.3a and 5.3b by the blue line. Further fit outcomes are shown in the appendix (A.9, A.10 and A.11 for February11, in A.12, A.13 and A.14 for April12 and in A.15, A.16 and A.17 for May13). In addition, the histograms contain the statistical error per bin and the red vertical dashed line represents the left fit boundary. Both figures (and also the appendix) take the uncertainty of the parameter β into account which is discussed in section 5.6. This uncertainty is not yet important in case the S2 acceptance is calculated without uncertainties in a first approach. The both dashed Poisson distribution are related to the lower and upper uncertainty in β. In every case, the uncertainties on the parameter β are within the statistical errors.
Subsection 5.3.1 describes the adjusted Poisson distribution for the fit procedure. The scaling parameter β depends on the expectation value λ and vice versa. Hence two free parameters in the fit would yield the best fit result. The later Poisson description will depend only on the expectation value λ and not on the other two parameters to stay conform with the original definition of a Poisson distribution. For this reason the parameter β is kept constant across all single cS2 histograms. Values between $100 \leq \beta \leq 400$ were tested in steps of $\Delta\beta = 5$. To answer the question of which (fixed) parameter β is the best to describe the Poisson distribution, a minimisation process is used with the goal to find the best β by the minimisation of the fit likelihoods for the fixed β and an additional variation in λ to test the robustness of the fit outcome. To achieve this goal, the initial values of the fit parameters do not change during the fitting routine.
The variation in λ is tested by a variation of 200 values around the fit outcome of:

$$\lambda^0 : \lambda^0 - 100 \leq \lambda_j^0 \leq \lambda^0 + 100$$

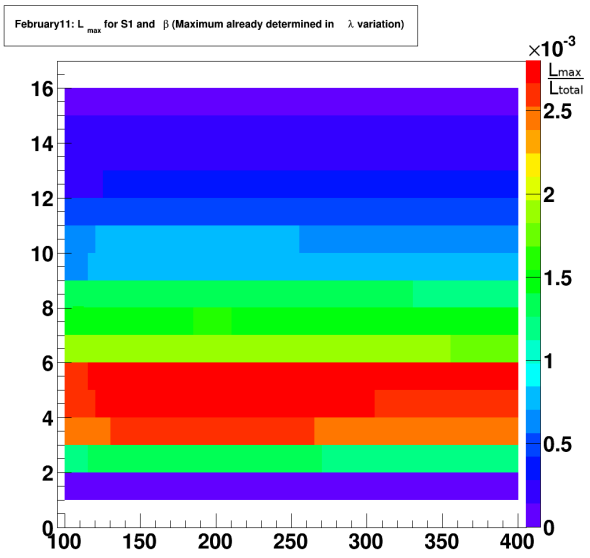
where $\Delta\lambda^0 = 1\,\text{pe}$. With this boundaries, the index j is introduces and refers to λ variation and is within $j \in [-100, 100]$ ($j \in \mathbb{R}$) to describe the iteration with $\Delta\lambda^0$ steps: λ_j^0. This allows a variation in each direction which is broader than one bin with the bin width of 75 pe. Based on the fits to the 15 cS2 histograms for each data period, which contain the primary information, there are 60×200 likelihoods calculated for each cS2 histogram. The likelihoods are calculated for a binned histogram according to formula (5.5) for the expectation value m from the fit to each bin and the observed data d in such a bin i. The index i describes here the number of bins in one of the histograms. The upper limit of $i = 67$ is related to the upper fit boundary of 5000 pe. In detail, 67 bins with a bin width of $\Delta S1 = 75\,\text{pe}$ results an upper boundary of 5025 pe what introduces a difference of 25 pe. But this difference is smaller than the bin width. Equation (5.5) is transformed to $L = \log \mathscr{L}$, the log-likelihood (see equation (5.6)), here the factorial $d!$ becomes $\log(d!)$ and is treated as a constant offset in each fit and hence it is left out in the further calculations.

$$\mathscr{L} = \prod_{i=1}^{i=67} \frac{e^{-m_i} {m_i}^{d_i}}{d_i!} \tag{5.5}$$

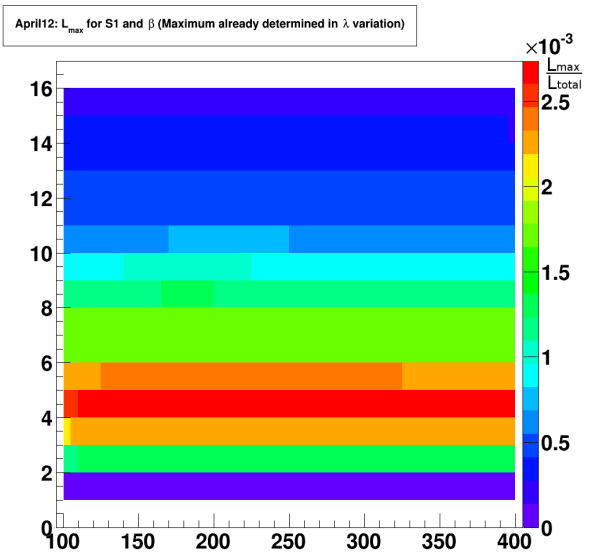
$$L = \log(\mathscr{L}) = \sum_{i=1}^{i=67} -m_i + d_i \log(m_i) \tag{5.6}$$

The log-likelihood is calculated for each of the 60 different test values of β. Once a test value β results in a possible λ^0, the according Poisson distribution is adjusted by the variation of λ. For each λ^0 there are 200 Poisson distributions with an alternative expectation value λ_j^0 with $\lambda^0 - 100 \leq \lambda_j^0 \leq \lambda^0 + 100$. The question of which values for β is the best is then answered by searching for the maximum log-likelihood value, L_{max}. This is shown in figures 5.4a, 5.4b and 5.4c, where L_{max} is determined out of 200 Poisson distributions with λ_j^0 as expectation value and the parameter β is fixed. The result is then sorted for each cS2 histogram. The figures are sorted due to the neutron calibration. From these results a total L_{max} is determined which is described as the maximum of all calculated L_{max}, named L_{max}^{total}. Figures 5.4a, 5.4b and 5.4c uses L_{max}^{total} for normalisation. The L_{max} distributions show that L_{max} is also maximised for a single slice of $\Delta S1$ and hence the maximum L_{max}^{β} is determined by searching the maximum L_{max} for a fixed β out of the cS2 histograms. These values are then normalised by L_{max}^{total}. The result is shown in figure 5.4d for February11, April12 and May13 data sets. The here observed maximum dictates the best choice for the parameter β. Hence the parameter β is fixed to $\beta = 200$ for February11, $\beta = 205$ for April12 and $\beta = 165$ for May13 ^{241}AmBe calibration.

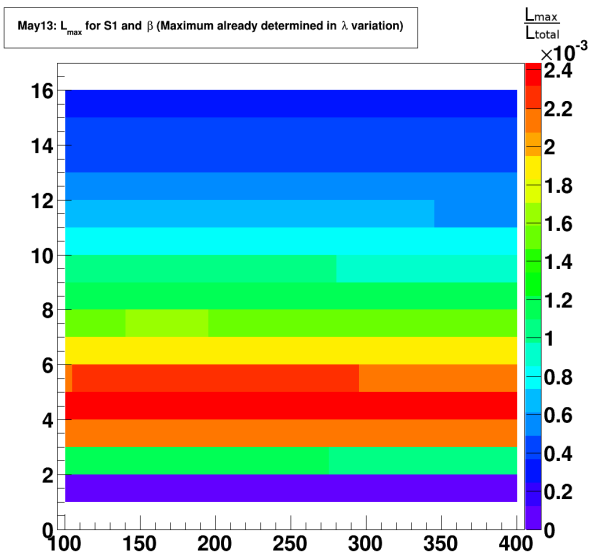
With the outcome of the optimisation in β, the according Poisson fits are analysed using the expectation value λ to characterise each slice of S1. The resulting linear relations $\lambda(S1)$ for the three data sets are given in figure 5.5a. The variation of the parameter λ is necessary for the error calculation. Section 5.6 will introduce a method to estimate the uncertainty on λ and β. The errors for β in table



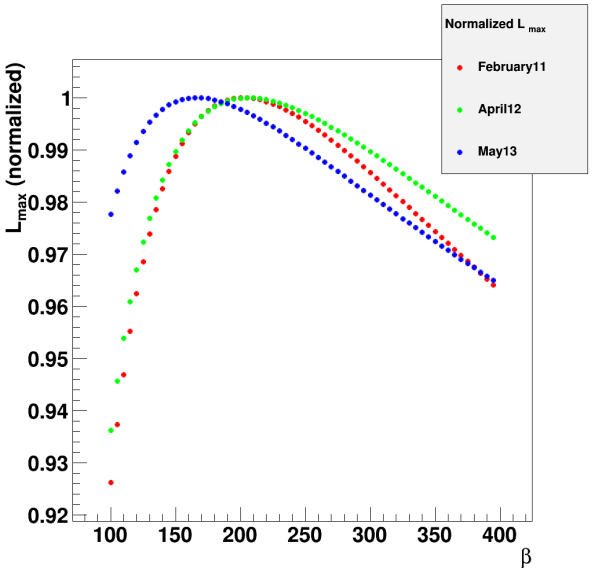
(a) Dataset: ^{241}AmBe - February11



(b) Dataset: ^{241}AmBe - April12



(c) Dataset: ^{241}AmBe - May13



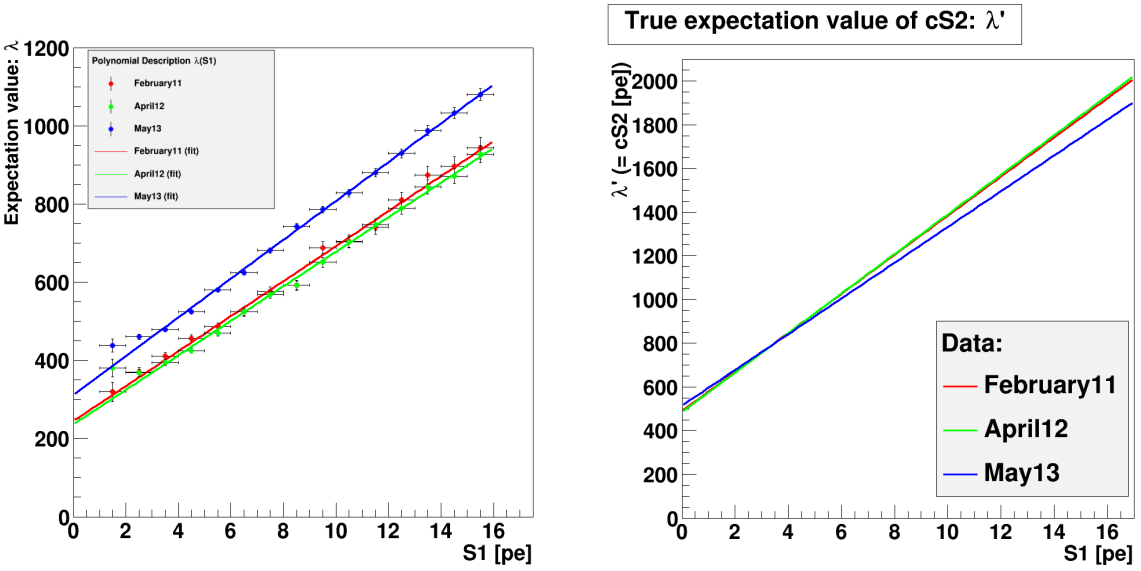
(d) Best choice for β according to the maximum log-likelihood.

Figure 5.4 – The L_{max} values from the variation around the expectation value λ. The result is given for the cS2-histograms and the different choices of β. The here presented data (normalised to L_{max}^{total}) sets describe the data sets of February11, April12 and May13 and based on the maximum likelihood L_{max}^{total}, the data sets are normalised the maximum of their distribution dictates the best choice for β (see figure 5.4d).

5.4 are also taken from this section. A higher value of S1 causes a higher expectation value λ. This is expected and also evident in the histograms. In addition, the fit errors are added: In λ the fit error is taken from maximum likelihood fit (ROOT) and the error bars in S1 direction reflect the size of the S1 slice ($\Delta S1 = 1\,\text{pe}$).
The characterisation of the single slices of S1 by an expectation value $\lambda(S1)$ is parametrised by a linear fit, also included in figure 5.5a. The fit procedure ensures a reliability down to 2 pe in S1. Hence, the extrapolation down to 0 pe depends on the previous values for λ and their error bars. Another aim of the linear approximation is a continuous description $\lambda(S1)$, which is summarised in table 5.4.
The polynomials $\lambda(S1)$ present the connection between the scintillation light via S1 and the expected signal from the ionisation process which results in cS2. Due to the abstract definition of the adjusted Poisson distribution in equation (5.2) and the transform in equation (5.3), it result $\lambda(S1) \propto cS2$. This means, there is less light produced and less electrons at the same time what indicates a low nuclear recoil. The lower the S1 signal the lower the expected cS2 signal (or the other way around). The linear approximation of $\lambda(S1)$ allows to estimate S2 signals below the XENON100 threshold. Such it is observed there, the error bars of the maximum likelihood fit increase for low S1 and a deviation from the linear approximation arise.
The linear fit is not physical motivated and is based on the minimum assumption given the relation between S1 and cS2. The relative scintillation efficiency (L_{eff}) and charge yield (Q_y) which is measured up to know and simulated by the XENON100 collaboration indicates that the linear approximation does not hold necessarily for low S1 energies. The number of electrons per keV nuclear recoil (see figure 3.15d) saturates at $3\,\text{keV}_{\text{NR}}$ and vote against the previous relation that less S1 signal is followed by less S2 signal. There are no further measurements below $E_{NR} = 3\,\text{keV}_{\text{NR}}$ available to determine the trend of Q_y in detail. This *saturation* effect leads to slightly increased observed expectation value λ within $S1 \in [0,2]\,\text{pe}$. The following subsection describes a method to calculate the acceptance from a Monte Carlo simulation which uses $\lambda(S1)$ as input. As already mentioned in section 5.3.1, the Poisson expectation value λ is connected to β by the definition of the adjusted Poisson distribution in equation (5.2) and (5.3). To remove the correlation between λ and β and connect the linear fit with the definition of λ' in equation (5.3), we can extend equation (5.3) by the fit function:

$$\lambda'(S1) = \frac{\beta}{100}\,\lambda(S1) = \frac{\beta}{100}\,(a \times S1 + b) \tag{5.7}$$

The parameters a and b are the fit parameters (slope and constant) according to table 5.4 and the multiplication with $\beta/100$ corrects the abstract expectation value definition to λ'. This means a transform of the polynomials in table 5.4. The result is $\lambda(S1)'$ which is the exact relation between the expectation value of the Poisson distribution and the S1 slice. Due to the transform, abstract definition of the expectation λ becomes easy to understand and furthermore the $\lambda(S1)'$ is measured in photo-electrons.



(a) (b)

Figure 5.5 – Based on the search for the maximum L_{max} for a fixed β ($\beta = 200$ for February11, $\beta = 205$ for April12 and $\beta = 165$ for May13), the best choice for β is made. Figure 5.5a summarises the resulting expectation values λ^0 from the fits to the single cS2 histograms. In addition, the polynomial $\lambda(S1)$ is drawn which is used to get a continuous description down to 0 pe. Figure 5.5b compares the three data taking periods. The *abstract* definition of the expectation value λ, related to the chosen β and is then corrected by formula (5.7) to λ' which presents the expectation value of the cS2 signal.

This presentation allows comparing the different fit polynomials directly. This is seen in figure 5.5b where the three data taking periods are shown together.

^{241}AmBe	Polynomial	β
February11	$\lambda(S1) = 244.8(\pm 17.2) + 44.7(\pm 1.92) \times \mathrm{S1}$	200^{+30}_{-20}
April12	$\lambda(S1) = 236.3(\pm 16.16) + 44.2(\pm 1.76) \times \mathrm{S1}$	205^{+25}_{-25}
May13	$\lambda(S1) = 311.8(\pm 17.0) + 49.6(\pm 1.79) \times \mathrm{S1}$	165^{+15}_{-10}

Table 5.4 – The linear approximation for the data taking periods. The fits are based on the errors which result from the Poisson approximation in λ. The errors in the brackets are the fit errors of the linear approximation by ROOT, not the uncertainty in λ which is determined in 5.6. The errors in β are taken from the uncertainties in section 5.6.

A shift of the linear fit functions is observed in figure 5.5a for February11 (red) to April12 (green). The S2 acceptance calculations of February11 and April12 use the same electron lifetime of 2011/12, $\tau_{DM}^{2011/12} = 374\,\mu\mathrm{s}$. The shift can not be explained by the same electron lifetime. To compare both data taking periods, figure 5.5a with the $\lambda(S1)$ is suitable. It already shows that $\lambda(S1)$ has a lower

^{241}AmBe	Polynomial	β
February11	$\lambda'(S1) = 489.6(\pm 34.4) + 89.4(\pm 3.84) \times \mathrm{S1}$	200^{+30}_{-20}
April12	$\lambda'(S1) = 484.4(\pm 31.1) + 90.6(\pm 3.61) \times \mathrm{S1}$	205^{+25}_{-25}
May13	$\lambda'(S1) = 514.5(\pm 28) + 81.8(\pm 2.95) \times \mathrm{S1}$	165^{+15}_{-10}

Table 5.5 – The linear approximation for the data taking periods after the β correction is applied to the outcome of $\lambda(S1)$. The error in the brackets refers to the β corrected fit error from table 5.4.

offset and slope in April12 compared to February11 and hence a small decrease in the S2 acceptance is expected. A detailed analysis takes the fixed value of β into account. The fixed value of β changes from $\beta = 200$ (February11) to $\beta = 205$ (April12). Both values differ not much from each other and within the uncertainties on the parameter β (see table 5.8 with the uncertainties from section 5.6). The result of the β correction is shown in figure 5.5b. Besides a minimal difference in the slope and constant of February11 and April12, both linear approximations are almost identical (The detailed outcome of the linear approximation after the β correction is applied and is shown in table 5.5). The correlation with the parameter β becomes important if the fit functions are compared to each other. Due to the β corrected fit errors in table 5.5 which are derived from the fit errors in table 5.4, the fit functions overlap within the errors. This observation results a stable detector environment during 2011/12. Moreover the resulting true expectation value $\lambda'(S1)$ from the data taking period 2013/14 overlaps within $S1 \in [0,6]$ pe pretty well with the both result from 2011/12.
This result shows a pretty good overlap of the true expectation value λ' which is in general a proof for the stable S2 signal creation from the drifting electrons in the liquid xenon. Due to impurities and slightly changing noise conditions are small changes expected and observed within the error bars.

5.4. Calculation of the S2 Signal Acceptance with Monte Carlo Method

The acceptance is determined by a Monte Carlo simulation. An infographic (see 5.6) introduces the necessary steps to calculate the S2 acceptance. This technique is chosen instead of integrating over the large parameter space (S1, cS2, fiducial volume, electron lifetime) directly. The input to the simulation comes from the already determined relation of cS2 and S1 by the linear approximation $\lambda(S1)$. Similar to the electron lifetime τ of the ^{241}AmBe measurements, the electron lifetime during the Dark Matter data taking period must be evaluated to estimate the S2 acceptance correct for the related Dark Matter data taking period. Moreover, the S2 acceptance needs to be determined for the fiducial volume of 34 kg.

$$S2_{mc} = cS2_{mc}\, e^{-\frac{\Delta t_{mc}}{\tau_{DM}}} \tag{5.8}$$

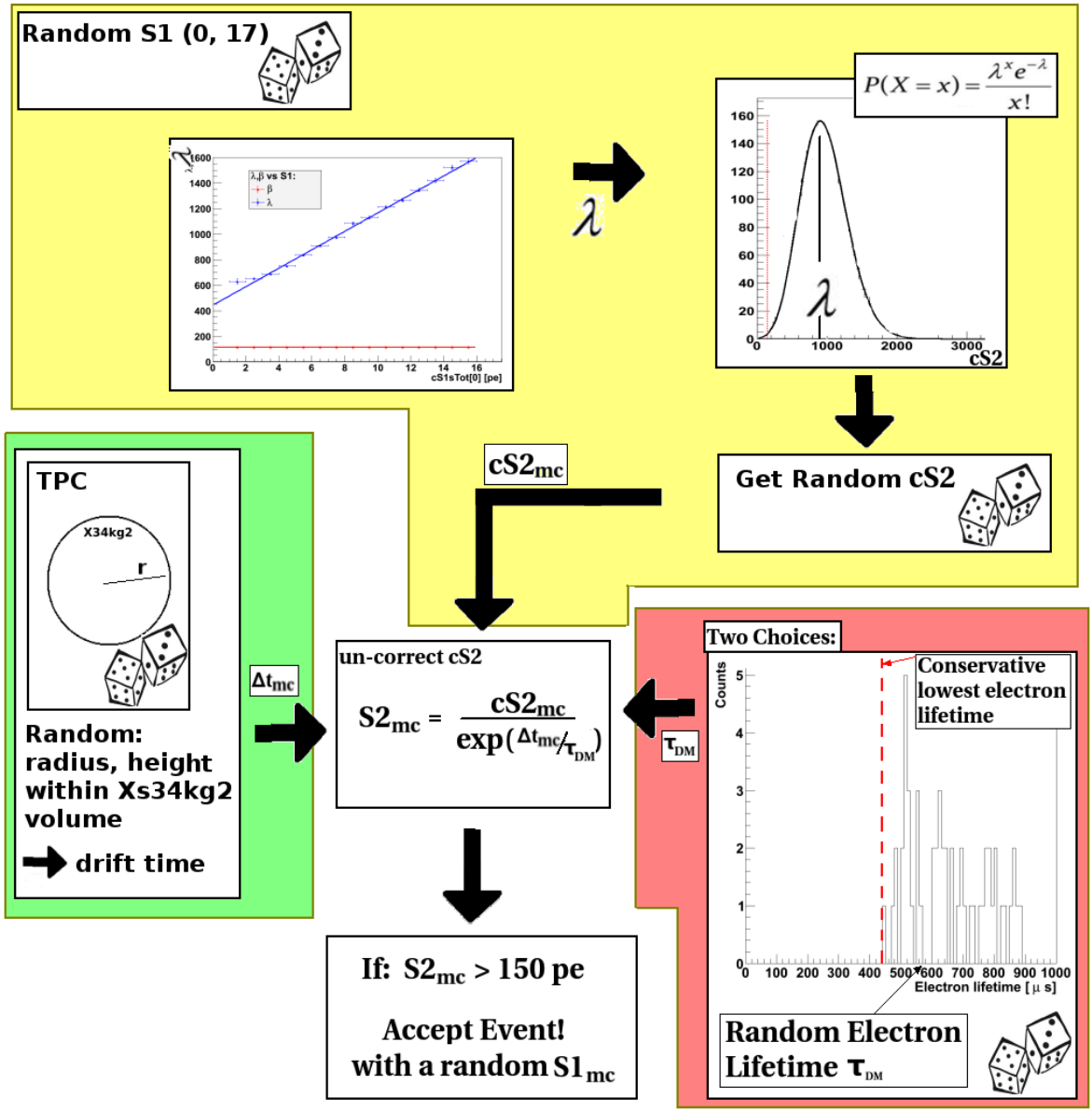


Figure 5.6 – The infographic describes the three different input parameters to equation (5.8). The event position in the TPC and the energy distribution of the S2 are sampled by Monte Carlo methods. Only the electron lifetime τ is chosen conservatively as the lowest observed τ with respect to the data taking periods 2011/12 and 2013/14.

These three inputs are summarised by the infographic around the equation (5.8) (see also the middle of the infographic). To point out the difference to already defined equation (5.4), the variables in equation (5.8) have additional indices *mc* and *DM* to confirm their origin. Its aim is to calculate the uncorrected S2 signal from the Monte Carlo simulation chain $S2_{mc}$ which then tested if its larger or smaller than the given threshold of 150 pe. This test is carried out by equation (5.9):

$$A_i = \frac{\#S2_{mc,i}(> 150pe)}{\#S2_{mc,i}} \quad i \in [0,17]\,\text{pe} \tag{5.9}$$

It is counted how often a generated event exceeds the XENON100 threshold of $S2_{mc} > 150$ an compared to the total number of generated events, yielding the acceptance according to equation (5.9) for each S1 slice i.
To evaluate the three input parameters ($cS2_{mc}$, Δt_{mc} and τ_{DM}) and execute formula (5.8), the fixed value β from the log-likelihood (see section 5.3.2) is a basic ingredient for the choice of the polynomial. Formula (5.9) is evaluated within $S1 \in [0,17]$ pe and uses the the extrapolation down to 0 pe at the left site to calculate the full S2 acceptance. In addition, an extrapolation up to 17 pe is also taken into account. This does not influence the S2 acceptance calculation. At this value of S1, the S2 acceptance becomes $\sim 100\%$.
The three ingredients to equation (5.8) are explained in the following in detail:

- **Parameter $cS2_{mc}$ (yellow):** A random number generator produces this parameter. In a first step, a uniform and random value of $S1_{mc}$ is drawn according to the previous S1 range ($S1 \in [0,17]$ pe). Based on the ^{241}AmBe data set of interest, the related polynomial description from table 5.4 is used to evaluate $\lambda(S1)$ at $S1_{mc}$. This yields a random expectation value λ_{mc} which defines the Poisson distribution with β fixed to table 5.4. With the expectation value λ_{mc} and the fixed β, a random value of $cS2_{mc}$ is drawn from this distribution. The $cS2_{mc}$ generation from the model distribution allows to sample values in cS2 below the left boundary by its Poisson probability.

- **Parameter Δt_{mc} (light red):** takes the possible drift times in the TPC into account. Due to the TPC height of ~ 30 cm, the maximum drift time is approximately $\Delta t = 176\,\mu$s. The Xs2peaks acceptance is supposed to describe the whole TCP for a uniform WIMP distribution and not only the upper third. Therefore a homogeneous drift time distribution of $5\,\mu\text{s} \leq \Delta t \leq 176\,\mu\text{s}$ is assumed. The S2 acceptance is calculated in the Dark Matter search region, an elliptical volume of 34 kg in the inner TPC (X34kg2). Hence the homogeneous drift time distribution is limited to events from inside the X34kg2 fiducial volume. This additional condition rejects simulated events near the edges of the TPC which could also come from background events from the detector surface and exterior.

- **Parameter τ_{DM} (light green):** describes the electron lifetime from the ^{137}Cs calibration (see section 3.6.1). These calibration data sets are taken at least weekly during the Dark Matter data taking period. The 2011/12 Dark Matter data is connected to ^{241}AmBe data sets of February11 and April12 and 2013/14 is connected to May13. The lower right histogram of figure 5.6 shows the electron lifetime distribution of 2013/14 which is developed from ^{137}Cs data. It suggests to sample values of the τ_{DM} histogram to empathise the measured electron lifetimes. That approach biases the S2 acceptance to higher values as it discussed in section 5.7.2. We compute a conservative lowest S2 acceptance from the lowest measured electron lifetime τ_{DM} which is equal to the lowest value in the τ_{DM} histogram of figure 5.6.

According to formula (5.8) the uncorrected $S2_{mc}$ signal from the Monte Carlo method is calculated. The energy threshold of 150 pe is applied to the $S2_{mc}$ to determine the number of accepted events by formula (5.9). Each generated $S2_{mc}$ event belongs to a $S1_{mc}$ which finally allows to present the S2 acceptance $A(S1)$. The result of this method is shown in figure 5.7. The presented outcome is not complete yet. Further details on the uncertainties are introduced in section 5.6 and a full discussion of the S2 acceptance follows in section 5.7

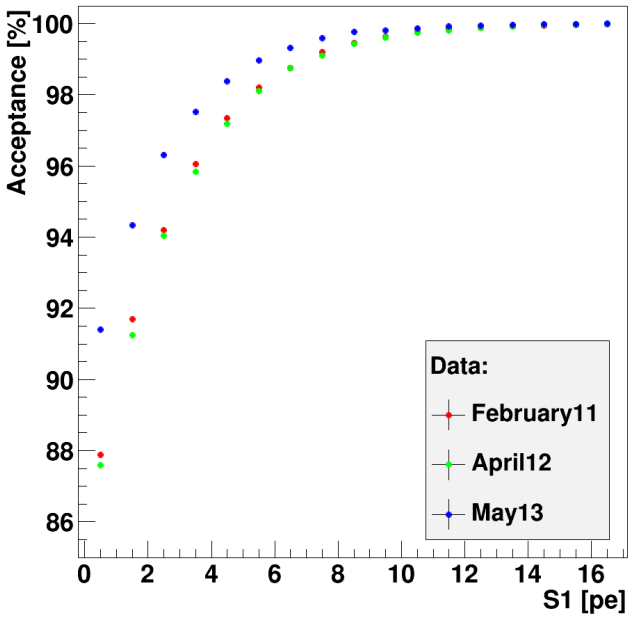


Figure 5.7 – The result of the S2 acceptance calculation based on Monte Carlo method with the input of February11, April12 and May13 ^{241}AmBe calibration data. The here presented result is completed by a full uncertainty discussion in section 5.7.

5.5. Verification of the S2 Signal Acceptance with Simulated Neutrons

The previous subsection 5.4 presents the general method to determine the Xs2peaks acceptance with a Monte Carlo simulation. The used data and cut selections, the Poisson description of the S2 histograms and the realised Monte Carlo method for the final acceptance calculation are derived from ^{241}AmBe data sets. The data based method to determine the S2 acceptance was developed and implemented before a signal response simulation of the XENON100 TPC was available. Meanwhile significant improvements were made in the simulation of the signal response of the TPC. The simulation of neutrons is done by the *Nuclear Recoil/WIMP Simulation Code* which is described in section 3.4.6 and a result of signal response simulation [23]. This tool aims to simulate the detector response for a given energy spectrum of neutrons (^{241}AmBe) in keV. The simulation takes the full TPC as well as the shield configuration into account. Moreover, the *Nuclear Recoil/WIMP Simulation Code* is able to simulate WIMP interactions in the liquid xenon for a given WIMP mass. The output is similar to reprocessed data and simulated S1 and S2 signals are analysed after the simulation.
Goal of the following subsection is to verify the Xs2peaks acceptance calculation with the same method as described in section 3.5.1. The simulation tool allows to generate corrected S1 and S2 signals as well as uncorrected. In general, there would be no need to apply the already described method to simulated neutrons to derive the Xs2peaks acceptance. Due to the applied detector acceptances and thresholds during the neutron simulation itself, there is a spectrum of S2 signals generated down to $S2 = 0\,\text{pe}$. A simple counting analysis of how many single scattered neutrons within the fiducial volume have an energy above $S2 > 150\,\text{pe}$, would be enough to determine the S2 signal acceptance. Figure 5.8 shows the result for an analysis where the events are counted ("simple S2 acceptance"). The binomial error is around ~ 0.001 and is not shown in this presentation. The simple S2 acceptance saturates faster at a 100% acceptance. The larger S1 is, the less events are counted for $S2 > 150\,\text{pe}$ in the cS2 histograms. This bias the S2 acceptance to higher values. A simple counting analysis is carried out in section 5.8 when the WIMP acceptance is calculated for different WIMP masses.
The presented method to calculate the S2 acceptance on data (see section 5.3) and the simple counting analysis can not be compared directly. The simple S2 acceptance does not take the spectral shape of the cS2 distribution into account. The following analysis of simulated neutrons with the previous established Poisson approximation is to verify the usability and correctness of the method by comparing simulated neutrons to ^{241}AmBe data. The comparison provides also feedback on the neutron simulation. Especially, for low values of S1, this is interesting, because the already established S2-histograms are cut at the $S2_{th} = 150\,\text{pe}$. Furthermore, another test proofs that the Poisson distribution approximates simulated neutron data as well as calibration data.

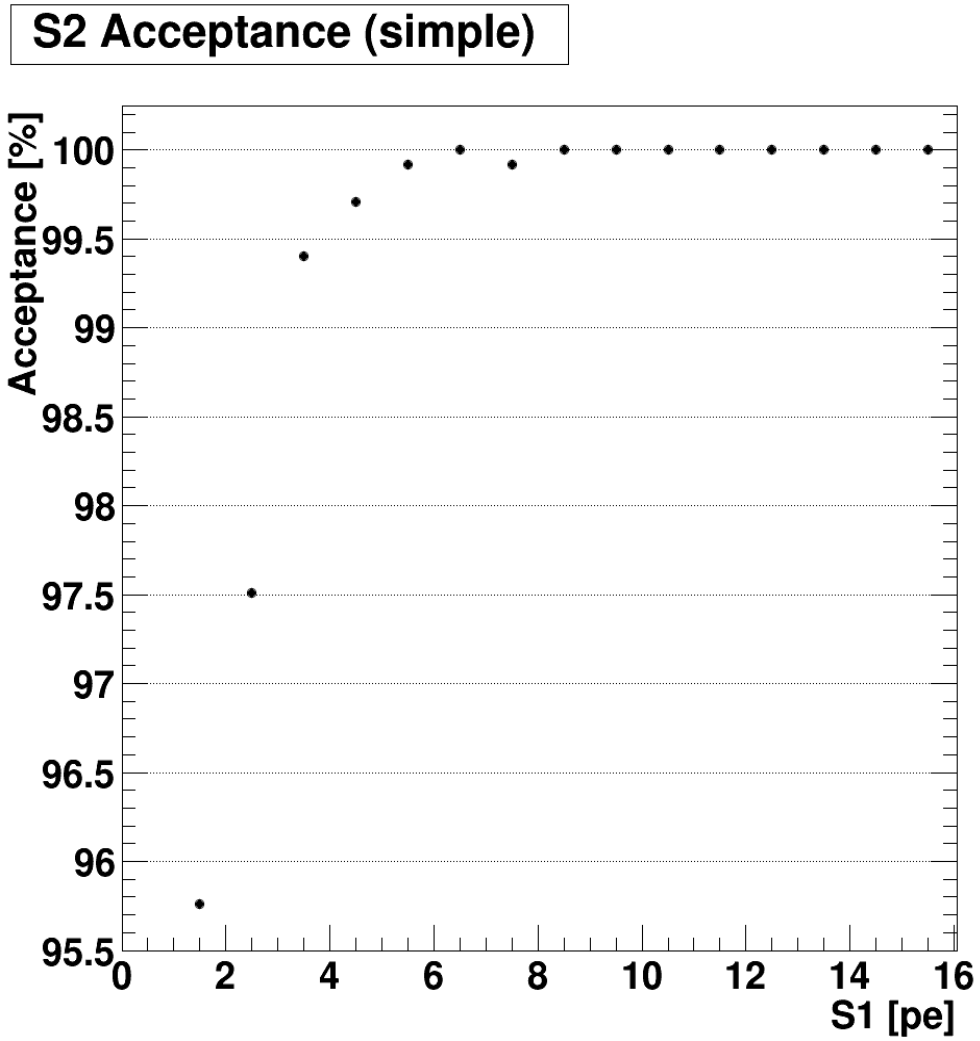


Figure 5.8 – The result of a simple S2 acceptance calculation with simulated neutrons for the data taking period 2011/12. The binomial error is much smaller for each slice of S1 and therefore it not presented in this figure. The simple S2 acceptance saturates at $S1 \in [6,7]$ pe.

5.5.1. Monte Carlo Signal Input and Cut Selection

The *Nuclear Recoil/WIMP Simulation Code* provides a full set of corrected and uncorrected S2 signals from interacting neutrons. This study uses a set of simulated neutrons with a total number of $\sim 1.1 \times 10^6$ events (before applying cuts). Besides the basic simulation input from the Americium-Beryllium energy spectrum, a set of thresholds and detector efficiencies is applied during the simulation. Most important for the signal propagation are the light yield L_y and the relative scintillation efficiency $\mathscr{L}_{\text{eff}}$ for the S1 signal (see formula (3.22)) and the charge yield $\mathscr{Q}_y$ (see equation (3.23)) for

the S2 signal. The XENON100 collaboration proves a good understanding of the signal propagation and the physical processes within the TPC [23].
The input parameters used here are $L_y = 2.28$, $S_{\mathrm{ee}} = 0.58$ and $S_{\mathrm{nr}} = 0.95$ for the S1 signal creation process. The S2 signal creation depends on the number of created electrons and the translation into units of photo-electrons. The conversion of one electron have an (Gaussian) expectation value of $S2 = 19.52\,\mathrm{pe}$ and a standard deviation of $\sigma_{\mathrm{S2}} = 6.71\,\mathrm{pe}$ [22]. The values fit the 2011/12 science run. Furthermore a threshold for $\mathrm{Q}_y > 0.2\,\mathrm{keV}$ is applied, the electron drift velocity is chosen to $v_{e^-} = 1.74\,\mu\mathrm{s/mm}$ and a z-resolution of 3 mm is used in the simulation. The PMT response is similar to the XENON100 experiment with 0.5 pe for a single photo-electron peak. With the given input of the 2011/12 science run a comparison to the Xs2peaks acceptance is realised here.

There are no dedicated cuts available for simulated neutrons and therefore the most important cuts must be reproduced in some simple way. The neutron simulation is done in a background-free environment in the simulated XENON100 TPC and hence a bunch of cuts is not necessary (e.g. cuts which are related to noise) to implement. A dedicated noise simulation is not available in XENON100. The simulation takes the possibility of multiple neutron-xenon scatters into account. The events are selected according to the first neutron scatter in the TPC as in the analysed data. This selection in the data sets is available by applying a single scatter cut (Xs2single3). This cut aims to select events with a large S2 and no other S2 peaks of the same size. The simulation allows to access how often a neutrons scatters in the TPC. A real neutron source produces neutrons with energies high enough to create multiple scatters and hence the Xs2single3 cut must be faked. The simulation allows to select *perfect* single neutron scatters by selecting only events with a single interaction in the TPC. This criterion is much more stringent than the Xs2single3 cut definition (see 5.2). Table 5.6 summarises the applied cuts to the simulated neutrons. The S2 threshold of 150 pe is applied and the fiducial volume of X48kg0 is selected like before in subsection 5.2. The radius of possible events in the TPC is chosen to $r < 145\,\mathrm{mm}$. There is no drift time available which would result the time difference between an S1 and S2 signal in the TPC and hence the depth (z-position) of the neutron scatter is limited to $-100\,\mathrm{mm} \leq z \leq -10\,\mathrm{mm}$ in the TPC. The simulation allows to access the interaction coordinates. The z-position selection is made according to the upper third of the TPC like before. The upper limit of $-10\,\mathrm{mm}$ is chosen according to the smallest accepted drift time of 5 µs which is translated into a z-position by a factor of 1.8 [43].
With *Nuclear Recoil/WIMP Simulation Code*it is possible to simulate corrected and uncorrected signal. To correct the S2 signal a proper electron lifetime is needed which is associated with the data taking period 2011/12. The simulation uses an electron lifetime of $\tau_e = 356\,\mu\mathrm{s}$ which is determined from ^{137}Cs source [23].

Reason	Cut set for simulated neutrons
S2-Signal	**Xs2peaks**, $\mathrm{nr}_{\mathrm{scatter}} \stackrel{!}{=} 1$
Position selection	X48kg0, Radius$<145\,\mathrm{mm}$, $-100\,\mathrm{mm} \leq z \leq -5\,\mathrm{mm}$

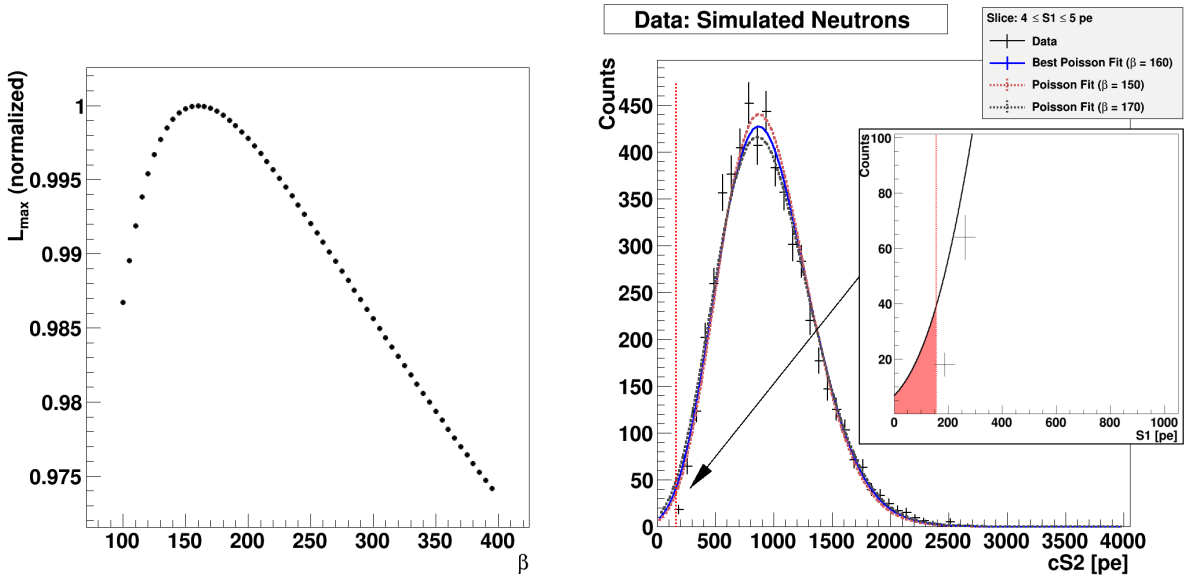
Table 5.6 – A small set of cuts is applied to the simulated neutrons. Noise reduction cuts are not necessary. The fiducial volume is chosen according to the previous data analysis. The selection of single S2 signals is done directly.

5.5.2. Minimisation Procedure with Simulated Neutrons

According to the cut selections of subsection 5.5.1, the cS2-histograms are filled with the corrected S2 signals from the neutron simulation. The cS2-histograms are fitted with a Poisson distribution, as defined in equation (5.2). An example is given in figure 5.9b for the S1 slice of $4\,\mathrm{pe} \leq S1 \leq 5\,\mathrm{pe}$. The red dashed line indicates the low fit boundary of $cS2 = 159\,\mathrm{pe}$ which is chosen according to the corrected S2 threshold (see equation (3.17)) of 2011/12 to compare the simulated neutrons from a 2011/12 configuration with data. The reddish area describes the event distribution below 150 pe which is calculated by the Monte Carlo method. The statistical error of the S2 histograms depends on the number of simulated neutrons.
The Poisson fit in figure 5.9b is performed for a fixed $\beta = 160$. Testing different fixed values of β with the simulated neutrons leads to a $L_{max}1^{total}$ at $\beta = 160$. The $L_{max}^{\beta=160}$ is given in figure 5.9a, which is equivalent to the figure 5.4d. It is maximised following the same procedure as described in 5.3.2.
The simulated neutrons are tested for $100 \leq \beta \leq 400$ in steps of $\Delta\beta = 5$. The appendix (see A.18, A.19 and A.20) includes the full set of S2-histograms for simulated neutrons with their Poisson approximations. The fixed value of $\beta = 160$ results in figure 5.10a where the single expectation values of λ are plotted versus S1. The linear approximation is given by the black line, $\lambda(S1)_{simNR}$. The fit range is set to $2\,\mathrm{pe} \leq S1 \leq 16\,\mathrm{pe}$. To compare the $\lambda(S1)_{\mathrm{simNR}}$ with the outcome of the calibration data, the previous results of February11 and April12 are also shown. The *Nuclear Recoil/WIMP Simulation Code* is able to simulate the nuclear recoil response down to $S1 = 0\,\mathrm{pe}$ (with Poisson smearing) according to an approximation of $\mathscr{L}_{\mathrm{eff}}$ down to 1 pe. Reasonable results are expected only above $S1 \geq 3\,\mathrm{pe}$. The outcome of the linear fit for $\beta = 160$ is given in equation (5.10). The numbers in the brackets are the fit errors of the constant and the slope from the fit (TMinute/ROOT). Section 5.6 describes the uncertainties of the parameters λ and β and hence the uncertainty is $\beta_{\mathrm{simNR}} = 160^{+10}_{-10}$.

$$\lambda(S1)_{\mathrm{simNR}} = 292.2(\pm 22.0) + 66.6(\pm 2.7) \times S1 \tag{5.10}$$

The outcome of the linear approximation for simulated neutrons must be corrected to the chosen best fixed β value. The β correction of equation (5.7) is shown in figure 5.10b and the β corrected poly-



(a) (b)

Figure 5.9 – Figure 5.9a shows the calculated L_{max} values for single fixed values of β in steps of $\Delta\beta = 5$. The maximum of L_{max}, the value L^{β}_{max} is observed at $\beta = 160$. Figure 5.9b is an example of a single S1 slice ($4\,\text{pe} \leq S1 \leq 5\,\text{pe}$). The simulated corrected S2 signals are fitted with the Poisson distribution of equation (5.2) (black). The red dashed line shows the low fit boundary and the filled area shows the loss of cS2 signals due to the S2 threshold which is described ideally by the Poisson fit.

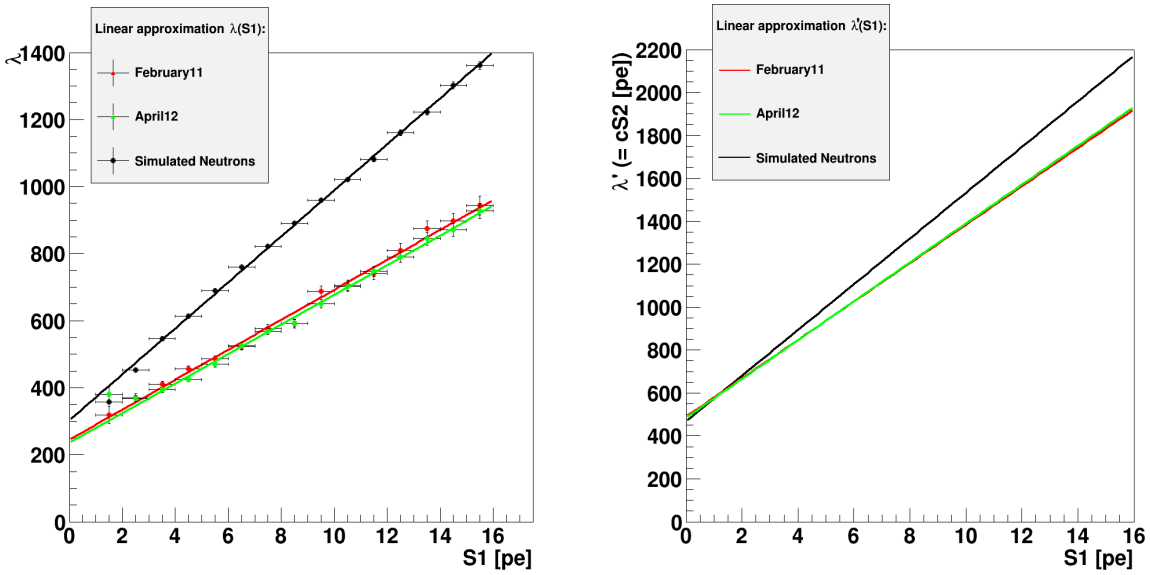
nomials are given in table 5.7.

With the outcome for $\lambda(S1)_{\text{simNR}}$ in equation (5.10), the same method is used to calculate the Xs2peaks acceptance as it is described in section 5.4 and is discussed in section 5.7. The Monte Carlo method uses the same inputs as in 2011/12. The fiducial volume is set to X34kg2 (see section 3.7) and the worst measured electron lifetime in 2011/12 is $\tau_{DM}^{2011/12} = 374\,\mu\text{s}$.

Figure 5.10a and 5.10b show a deviation between the data driven fit polynomials and from the simulated neutrinos. The change in the slope from data to simulated neutrons will impact the S2 acceptance. This is discussed, in particular, in section 5.7.3.

	Polynomial	β
February11	$\lambda'(S1) = 489.6(\pm 34.4) + 89.4(\pm 3.84) \times \text{S1}$	200^{+30}_{-20}
April12	$\lambda'(S1) = 484.4(\pm 31.1) + 90.6(\pm 3.61) \times \text{S1}$	205^{+25}_{-25}
Simulated Neutrons	$\lambda'(S1) = 468.9(\pm 35.2) + 106.3(\pm 4.32) \times \text{S1}$	160^{+10}_{-10}

Table 5.7 – The linear approximation for the data taking periods in February11 and April12 are corrected to their values of $\beta = 200$ and $\beta = 205$. The simulated neutrons are corrected with $\beta = 160$. The β correction yields $\lambda'(S1)$ which is the expectation value of the Poisson approximation measured in photo-electrons.



(a) (b)

Figure 5.10 – Figure 5.10a shows the expectation values λ, currently used, of the Poisson fits to the S1 slices. The linear approximation is fitted in $2\,\text{pe} \leq S1 \leq 16\,\text{pe}$. Due to the XENON100 S1 threshold of 3 pe the fit is relevant for $S1 \geq 3\,\text{pe}$. Furthermore, figure 5.10b shows the β corrected polynomials: $\lambda'(S1)$ for the February11/April12 data taking period and the simulated neutrons.

5.6. Uncertainty of the S2 Signal Acceptance

The following subsection will discuss the uncertainties for the previously discussed methods to evaluate the Xs2peaks acceptance. The error determination will be the same for the ^{241}AmBe data of section 3.5.1 and the simulated neutrons of section 5.5. There are a number of parameters which have a given uncertainty which will influence the final Xs2peaks acceptance.

The uncertainties from the fits of the Poisson model to the data and of the linear fit of $\lambda(S1)$ do not provide a full account of statistical uncertainty in the acceptance. Rather we need to consider the correlations between the of the two parameters β and λ in equation (5.2). As already discussed in section 5.3.1, the common parameter which connects the cS2 histograms is β. The best choice of β is determined by equation (5.6) where L_{max} is calculated from the variation of λ^0 with $\lambda^0 - 100 \leq \lambda^0 \leq \lambda^0 + 100$ for each slice of $\Delta S1$ and $\Delta\beta$. The best choice is then the total maximum L_{max}^{total}. From here the uncertainty for the choice of β is missing. Furthermore, this procedure does not yet take the overall uncertainty of the variation in λ into account.
Therefore the overall uncertainty depends on the deviations of λ and β together. This becomes clear by the fact that λ can be transformed into a *true* expectation value λ' by β. The third fit parameter α

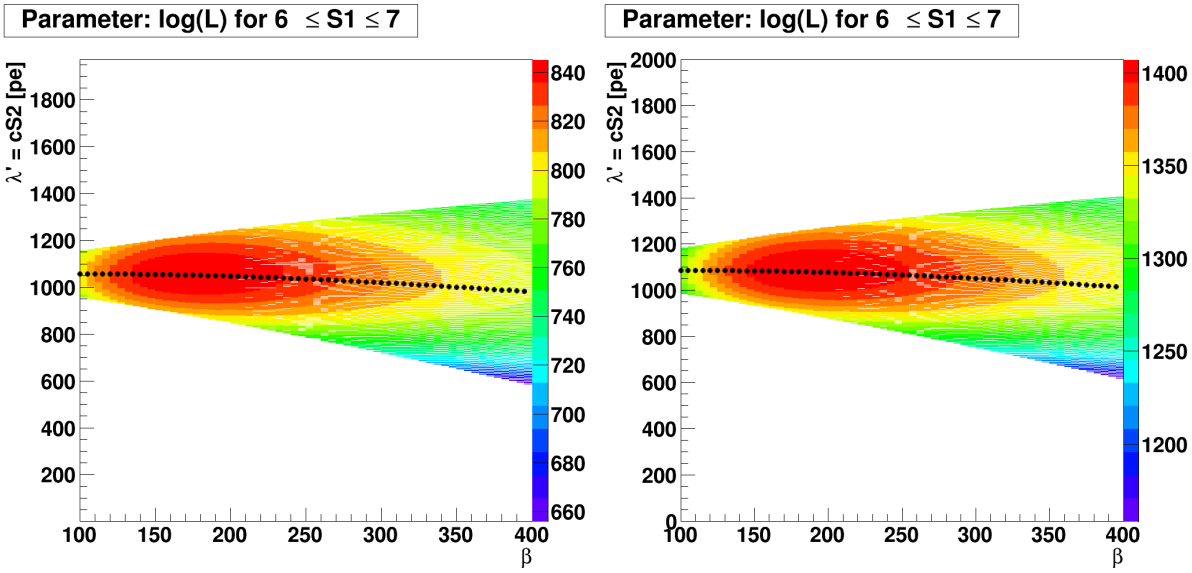
also needs to be taken into account. But as long as the scaling is independent of the choice of β and λ, this becomes only necessary in the later correct choice of a 68.3% uncertainty level determination. Hence, the aim is to calculate a 68.3% confidence region on the already shown S2 acceptances (see figures 5.5b and 5.17a).
The usual one-parameter estimation with the calculated log-likelihoods is given in equation (5.11) for a parameter θ. The log-likelihood L is maximised (L_{max}) and the confidence region is determined by subtracting $N^2/2$ from thus. The level of uncertainty is given by the Gaussian quantile $N = \Theta^{-1}(1-\gamma/2)$ where $1-\gamma$ corresponds to the desired level of confidence region [47].

$$\log(\theta_a^b) = \log(L_{max}^{Total}) - \frac{N^2}{2} \tag{5.11}$$

In case of three parameters, equation (5.11) is extended and a multidimensional confidence region must be estimated. Hence equation (5.12) describes this adjustment for the three parameter fit which is necessary here. The required value of $Q_\gamma = 3.53$ refers to a fit with three parameters and a 1σ confidence region (68.3%).

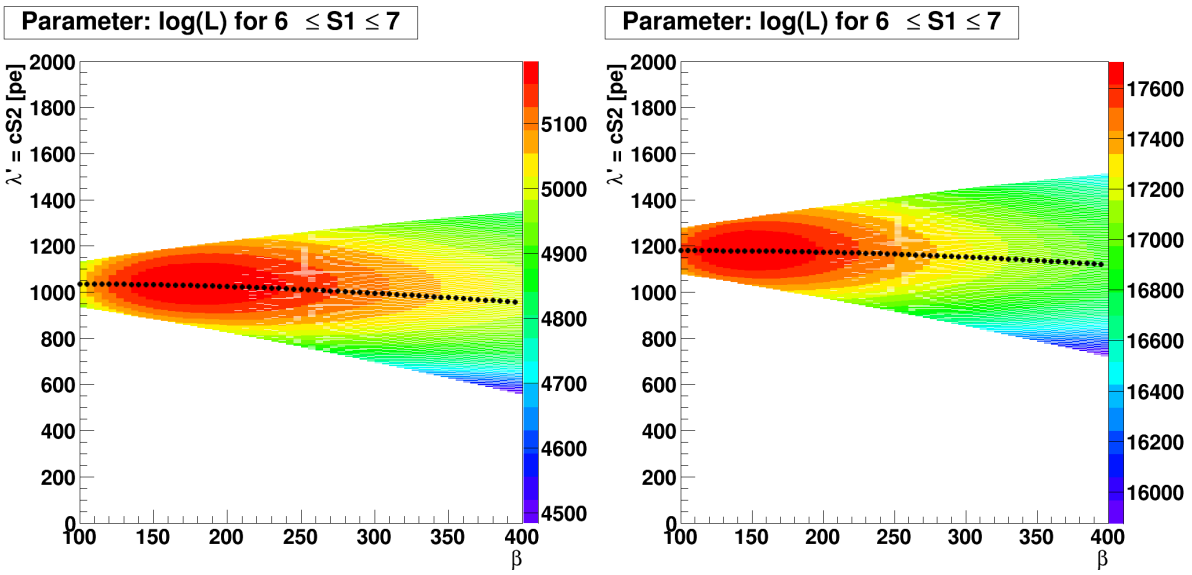
$$\log(\vec{\theta}) = \log(L_{max}^{Total}) - \frac{Q_\gamma}{2} = \log(L_{max}) - 1.765 \tag{5.12}$$

With this confidence level, the uncertainties are estimated for each cS2 histogram to evaluate the overall uncertainty on the choice of λ and β. Examples are given in figures 5.11a, 5.11b, 5.11c, and 5.11d for the February11, April12 and May13 data and the simulated neutrons. In every case a slice of $6\,\text{pe} \leq S1 \leq 7\,\text{pe}$ is shown as an example. The calculations are done for each slice of S1. Therefore the y-axis shows the *true* expectation value λ' and on the x-axis every tested value of β is given. Due to the relation of λ and β, the λ' should be almost the same for each β and therefore the horizontal arrangement is explained. This can be seen by the black dots which present the results of the individual tested values of β for λ', the β corrected expectation values of the cS2 histogram fits. The variation in λ' is shown by the log-likelihoods L in the colour code. Based on equation (5.12) and the variation in λ', the uncertainty is determined by the maximum variation. This is demonstrated in figure 5.12a, where the given 68.3% confidence level is applied to the February11 data set of figure 5.11a. The arrow in vertical direction shows the uncertainty for this slice of S1 in $\Delta\lambda'$. The uncertainties for all slices of S1 and the calibration data are given in figure 5.12b. Also the uncertainty is drawn for simulated neutrons. The uncertainties for λ' drawn here are again transformed back into $\Delta\lambda$ to use these values in the *extended* Monte Carlo technique (explained below) to determine the uncertainty on the Xs2peaks acceptance. The rise in the uncertainties from $S1 = 4\,\text{pe}$ to $S1 = 16\,\text{pe}$ is explained by the Poisson fit assumption to the cS2 histograms. For low values of S1, the cS2 distribution fits pretty



(a) Dataset: February11

(b) Dataset: April12

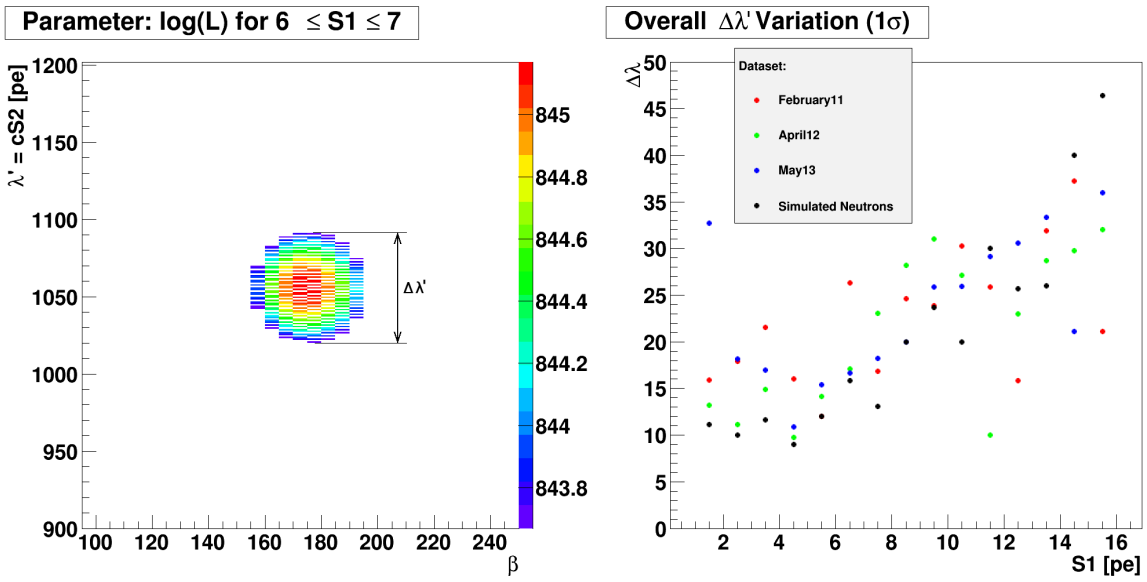


(c) Dataset: May13

(d) Simulated neutrons

Figure 5.11 – The figures shown here present a selection of the S1 energy slice $6\,\text{pe} \leq S1 \leq 7\,\text{pe}$ and compare the outcome of the calculated log-likelihoods to each other for the February11 (5.11a), April12 (5.11b) and May13 (5.11c) data sets as well as the simulated neutrons (5.11d). In every case, the log-likelihood values from the colour code follow the true expectation values λ' which are the direct results of the Poisson fits to the cS2 histograms. The already presented maximum in β varies a little in these figures but agrees with the overall determined values of $\beta_{February11} = 200$, $\beta_{April12} = 205$, $\beta_{May13} = 165$ and $\beta_{simNeutrons} = 160$. The log-likelihood values are the absolute values.

well the Poisson assumption and with increasing S1 energy, the cS2 distribution becomes a more and more symmetric shape. This indicates a better approximation with a Gaussian distribution for higher S1 energies. Therefore the Poisson description is less and less valid which explains the increasing uncertainty. The estimated uncertainty for λ is also in a similar range, compared to the parameter errors of the expectation value λ from the fits.



(a) (b)

Figure 5.12 – Based on the multidimensional confidence level of 68.3% the uncertainty in λ' is determined for each cS2 histogram by investigating the variation in λ' around the expectation value λ^0 which is the outcome of the Poisson fits. The uncertainty in λ' is named $\Delta\lambda'$ in figure 5.12a. The β dependent back transform allows to determine $\Delta\lambda$ (not shown here). Figure 5.12a presents a zoom into the February11 data set of the S1 slice $6\,\text{pe} \leq S1 \leq 7\,\text{pe}$ of figure 5.11a. Summarising all three data sets and the simulated neutrons, figure 5.12b shows the uncertainties for each slice of S1.

After the uncertainties on the parameter λ are determined, the parameter β needs to be evaluated. Therefore the already shown figure 5.4d is used. From the tested values of β, a L_{max}^{total} is estimated and based on equation (5.12) for the three parameter fit, a β_{min} and β_{max} is determined. Therefore not the normalised L_{max} of figure 5.4d and the absolute values of L_{max} are evaluated. An example is given in figure 5.13a where, in addition to the absolute values of L_{max}, a horizontal line indicates the outcome of the multidimensional confidence level of 68.3%. The values of the left and right boundaries for the best choice of β are given in table 5.8 for the February11, April12, May13 data sets and the simulated neutrons.

Data set	β_{min}	β	β_{max}
February11	180	200	230
April12	180	205	230
May13	155	165	180
Simulated Neutrons	150	160	170

Table 5.8 – Overview about the uncertainties on the fit parameter β. The boundaries β_{min} and β_{max} are the result of the multidimensional confidence level from a three parameter fit.

Together with table 5.8 and figure 5.12b the general uncertainty is determined. For this, the already discussed Monte Carlo technique (section 5.4) is used. The Monte Carlo method is based on the linear approximation $\lambda(S1)$ which is determined from the expectation values λ_1 to λ_{15} for the best choice of β. This Monte Carlo technique is extended by the uncertainties of λ and β. This is summarised by the diagram in figure 5.14 and is explained in detail in the following steps:

1. In a first step a random value of β is drawn according to the L_{max} distribution in figure 5.13a (example of data set February11) with the boundary condition of having β_{min} and β_{max} (see table 5.8). The available distribution of the values L_{max} is by definition not a PDF, but as long as the random values of β follow the most likely distribution the deviation of β within its boundaries is taken into account. This is shown once again in step 1 of the diagram. The parameter β^{rand} indicates a random value of β.

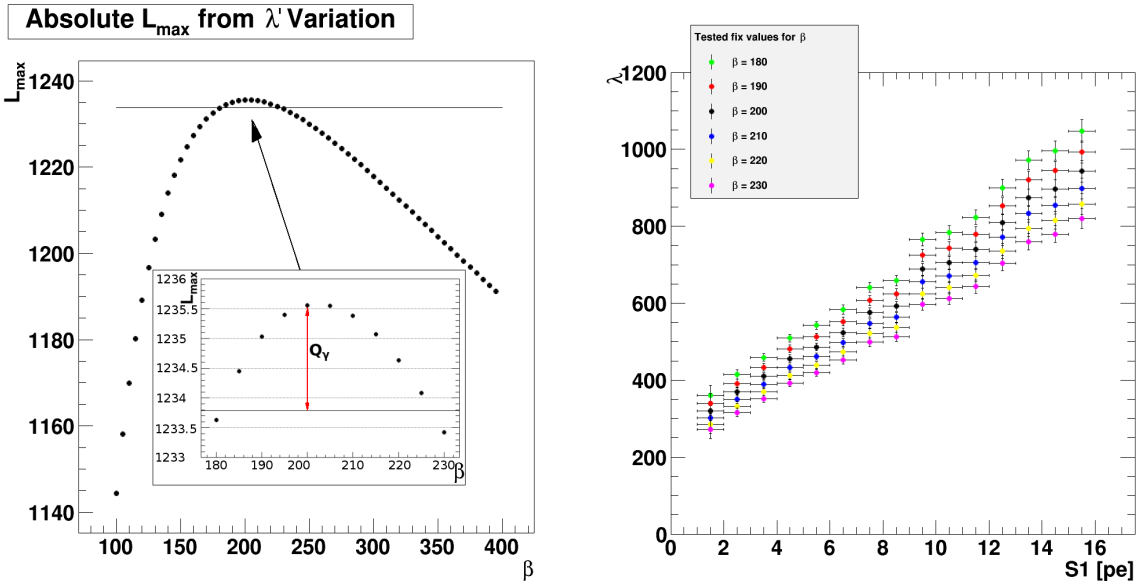
2. The original procedure of finding the best value of β provides also a possible set of expectation values λ_1 to λ_{15} for each β. Based on the uncertainty in β in table 5.8 the according fit outcomes of the Poisson fits are selected. Therefore a β^{rand} is connected to one set of expectation values λ_i with $i \in [1;15]$. A selection of possible expectation values λ_i for different β values is given in the second step of the diagram in figure 5.14.

3. Each of the chosen λ_1 to λ_{15} becomes the expectation value of a Gaussian distribution (see step 3 in the diagram). This corresponds to the a parabolic shape of the log-likelihood around the maximum. The uncertainty on the Gaussian assumption is determined from the variation in λ (see figure 5.12b). From these fifteen Gaussian distributions, fifteen values λ_1^{rand} to λ_{15}^{rand} are drawn randomly. These represent another set of random expectation values which could be understood as outcomes of the Poisson fits to the cS2 histograms.

4. With this set of random expectation values another linear approximation is done. In step four of the diagram 5.14, the outcome of one of these fits is indicated. The linear fit yields another possible slope and constant of $\lambda^{rand}(S1)$.

5. Finally, the new random linear approximation is used as an input to the already discussed Monte Carlo technique in section 5.4 to calculate a new S2 acceptance.

To evaluate the S2 acceptance with proper statistics, the steps 1 to 5 are executed 2000 times for each data set and the simulated neutrons. This assures a high number of draws to estimate the 68.3% confidence level of the Xs2peaks acceptance.



(a) (b)

Figure 5.13 – Figure 5.13a shows the absolute values of the calculated L_{max} for the tested values of β from the variation in λ. The maximum log-likelihood L_{max}^{total} is given here for the February11 data set with $\beta = 200$. According to the multidimensional confidence interval estimation for a three parameter fit, the horizontal line indicates a 68.3% confidence level around the best choice of β. This definition allows a left and right boundary for β: β_{min} and β_{max}. The uncertainty for β allows to select possible outcomes of the Poisson fits (expectation value λ) to the cS2 histograms. Figure 5.13b shows a selection of different values of β which are within β_{min} and β_{max}.

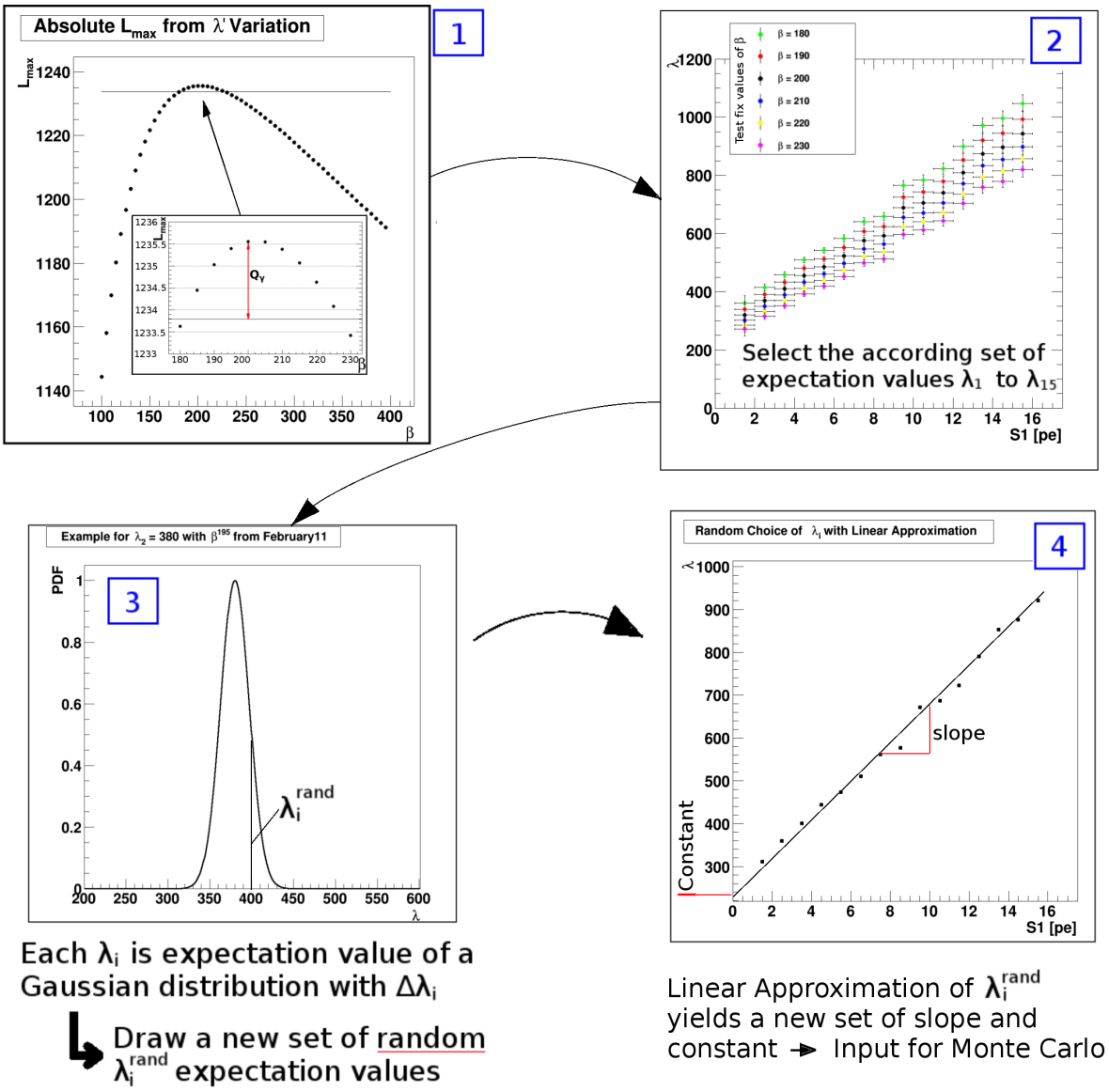


Figure 5.14 – Infographic about the extended Monte Carlo technique which is used to determine the uncertainty of the S2 acceptance from the parameters λ and β.

5.7. Results and Discussion of the S2 Acceptance

The following section shows result of the above discussed method to evaluate the S2 acceptance in XENON100 with its related uncertainty.

5.7.1. Results of the S2 Acceptance and Uncertainty

The previous section 5.4 introduced the Monte Carlo simulation to evaluate the S2 acceptance for the three data taking periods. Section 5.5 introduces a neutron simulation to verify the S2 acceptance calculation. The result of the acceptance calculation is already shown in figure 5.7 as a preliminary result for the three data taking periods which is completed by section 5.6 with the uncertainties in λ and β. The result of the Xs2peaks acceptance cut with its uncertainty is given in figure 5.15a for February11, 5.15b for April12 and 5.15c for the May13 data set and the simulated neutrons in figure 5.15d. The Xs2peaks acceptance is already given in the past section and in green the result of the 68.3% confidence level is drawn. From here the uncertainty behaves as expected, a lower value of S1 will result in an increasing uncertainty and the uncertainty decreases when a saturation of almost 100% is observed in the S2 acceptance calculation.

5.7.2. Discussion: S2 Signal Acceptance with ^{241}AmBe

The acceptances, available for the Xs2peaks cut are calculated with a large Monte Carlo sample ($\sim 16 \times 10^6$ events each) and are presented in figure 5.15. The result indicates a small decrease in the S2 acceptance from the February11 to the April12 ^{241}AmBe calibration data and furthermore an increase from the April12 to the May13 calibration data. That increase/decrease in the S2 acceptance is mainly observed for S1 values below 10 pe. For $S1 > 10$ pe, the difference almost vanishes, and the Xs2peaks acceptance becomes $\sim$100% for all three data taking periods. Compared to the uncertainties on parameter λ and β from where the errors in the S2 acceptances are determined from, the increase/decrease is not significant.
To understand the observed increase and decrease below 10 pe it is necessary to evaluate the parameters which go into the Monte Carlo calculation of section 5.4. The study presented here does not vary the fiducial volume during the Monte Carlo procedures. The changing electron lifetime from April12 to May13 and the upward shift of the fit polynomial (February11/April12 $\rightarrow$ May13, see figure 5.5a) are discussed to explain the observed Xs2peaks acceptance.

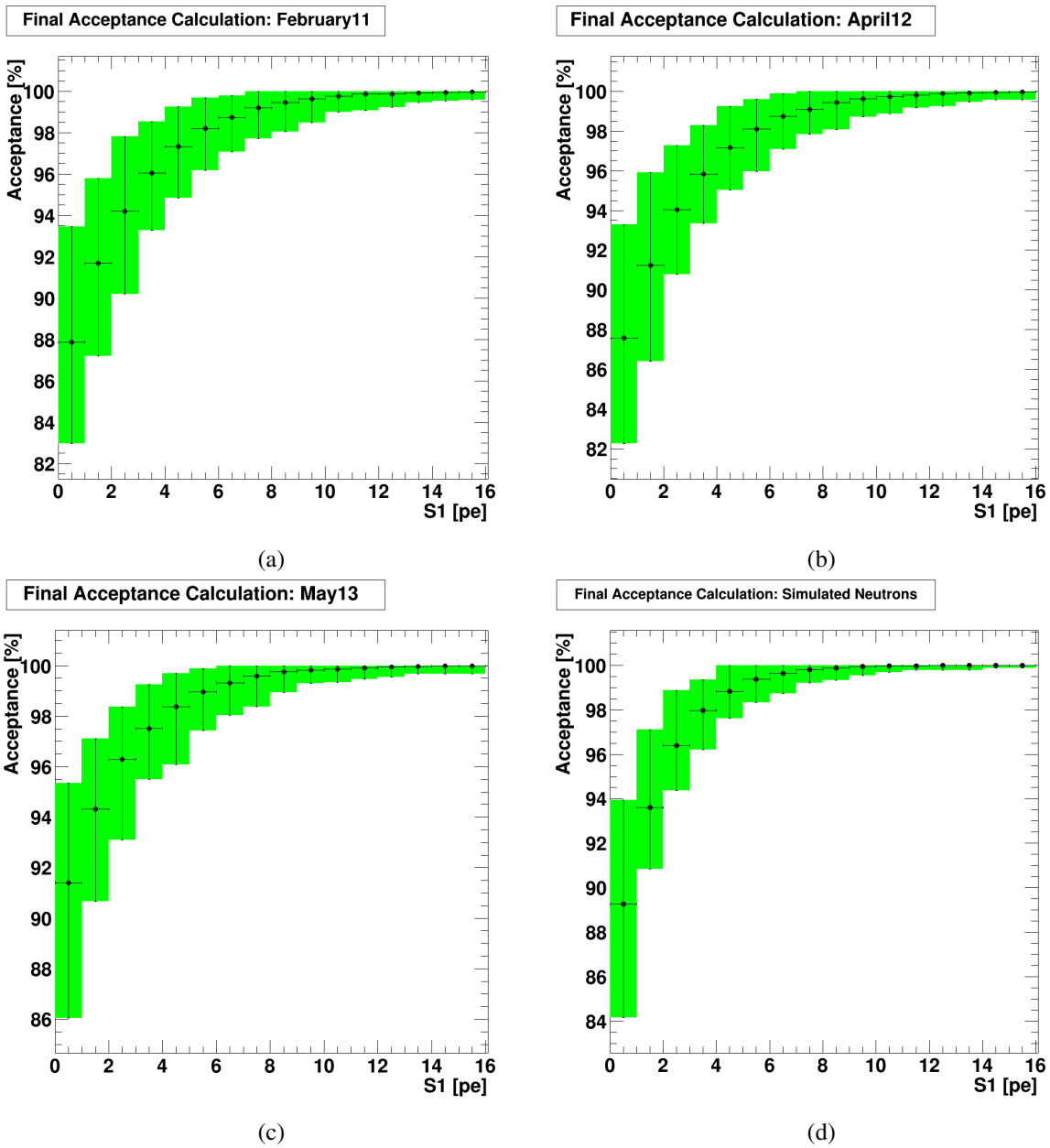
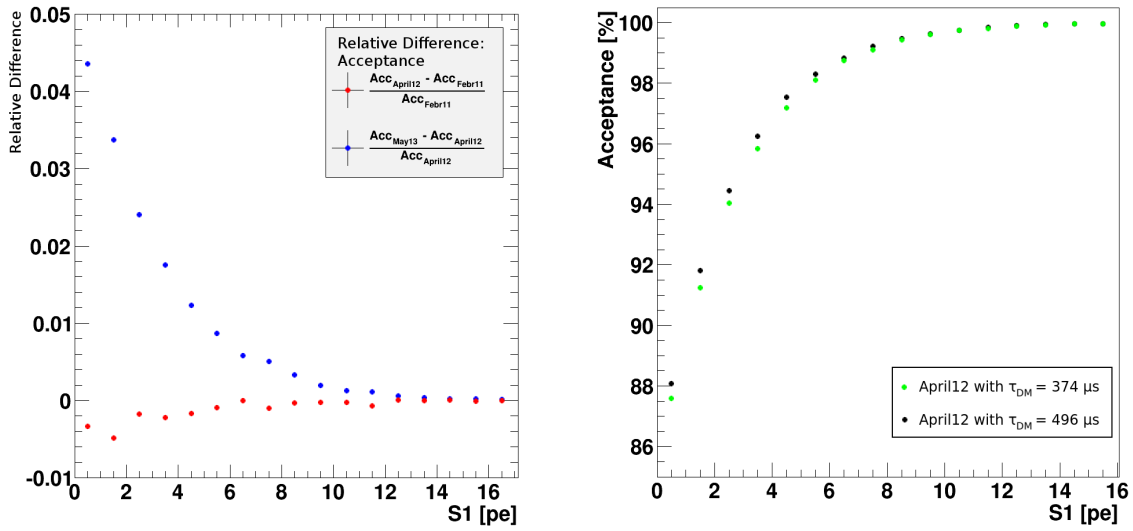


Figure 5.15 – Overview about the Xs2peaks cut acceptance with the according 68.3% confidence level which is determined from the three parameter fit to the Poisson distributions and in depend in detail on the coupled parameters λ and β.



(a) (b)

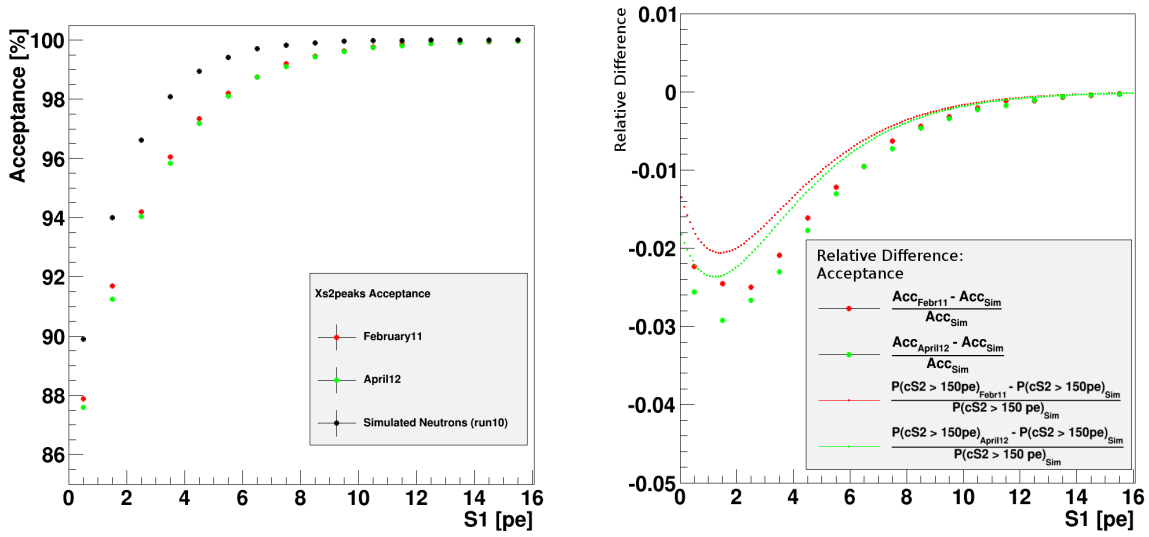
Figure 5.16 – Figure 5.16a present the relative differences of the calculated S2 acceptances. The data taking periods from February11 to April12 (red) and April12 to May13 (blue) are compared. The relative difference increases for the S2 acceptance below $S1 < 6\,\text{pe}$ for the data taking periods April12 to May13 by at most 2.5% above 2 pe in S1. For February11 and April12 calibrations it is nearly zero (within 0.3%). Figure 5.16b shows what happens if the worst electron lifetime $\tau_{DM}^{2013/14}$ is applied to the S2 acceptance calculation with April12 input. The S2 acceptance increases which is expected because the higher electron lifetime.

The result of the Xs2peaks calculation is compared for February11 and April12 in figure 5.16a. The red markers show the residuals of the different calculated S2 acceptances. The deviation is less than 0.4%. Even the strongest deviation lies up between $1\,\text{pe} \leq S1 \leq 2\,\text{pe}$, which is not in the range where the linear fit procedure is defined. From this observation the small increase in Xs2peaks acceptance is expected and can be explained. The electron lifetime is not changing between February11 and April12 and due to the β correction which results an overlap of both β corrected fit functions $\lambda'(S1)$ with in the fit errors (see table 5.5) and the uncertainty in β, there is no impact on the S2 acceptance.
To compare the Xs2peaks acceptance of May13 with February11/April12, the input polynomial $\lambda(S1)$ of figure 5.5a was calculated according to formula (5.7) to $\lambda'(S1)$. The polynomial is given in table 5.5 and the result is shown in figure 5.5b. The Monte Carlo method uses a longer electron lifetime of $\tau_{DM}^{2013/14} = 496\,\mu\text{s}$ compared to 2011/12 with $\tau_{DM}^{2011/12} = 374\,\mu\text{s}$ to calculate the S2 acceptance. The linear approximation in 2013/14 shows a higher constant which means that a higher value of cS2 is expected at low S1. At the same time, the slope of the 2013/14 polynomial decreases. By comparing the polynomials, the point of intersection is at $S1 \approx\sim 4\,\text{pe}$. Below 4 pe the 2013/14 polynomial pro-

vides a larger true expectation value $\lambda'(S1)$ than for 2011/12. The longer electron lifetimes between the two ^{241}AmBe calibrations are the expectation for the larger Xs2peaks acceptance (figure 5.15 or figure 5.7) for $S1 < 6\,\text{pe}$. This is also seen in figure 5.16a where the S2 acceptance difference becomes maximal for $S1 < 4\,\text{pe}$. The observed difference there is up to 3%, which is also out of the previous (linear) fit range of the linear approximation.
Finally, the impact of the change in τ_{DM} from 2011/12 to 2013/14 is evaluated. Therefore the Xs2peaks acceptances of April12 and May13 are compared. As already shown, the February11 is very similar to the April12 data for low values of S1 and therefore we restrict the comparison. Figure 5.16b shows the S2 acceptance computed for the minimum Dark Matter electron lifetime in 2011/12 for the data taking period 2013/14, for example, assuming the earlier data had already been taken at a higher xenon purity. Hence, the Monte Carlo calculation is done for the polynomial of April12 with the worst electron lifetime of 2013/14 ($\tau_{DM}^{2013/14} = 496\,\mu\text{s}$). The result is shown in figure 5.16b by the black markers. As expected, the higher electron lifetime of 2013/14, τ_{DM}, increases the S2 acceptance in the Monte Carlo method. However, the effect is small but explains very well the robustness of the Xs2peaks acceptance calculation for the three data taking ^{241}AmBe data taking periods. The rising electron lifetime during both data taking periods increases the S2 yield according to the corresponding S1 signal, what leads to higher S2 acceptance.

5.7.3. Discussion: Simulated S2 Signal Acceptance

The Xs2peaks acceptance for simulated neutron data is shown in figure 5.15d with the according uncertainty for the parameters λ and β. In addition, figure 5.17a compares the simulated S2 acceptance with the data taking periods February11/April12 without the parameter uncertainties. Figure 5.17b shows the residuals of the simulated neutrons and the February11 and April12 data sets. The simulated neutrons in the 2011/12 condition show a higher S2 acceptance for $0\,\text{pe} \leq S1 \leq 9\,\text{pe}$ compared to the data driven Xs2peaks acceptance of 2011/12. For $S1 \geq 9\,\text{pe}$, the acceptance becomes almost 100%. The relative differences from figure 5.17b show that the simulated S2 acceptance has a maximum deviation of $\sim 2.4\%$ from the February11 Xs2peaks acceptance and a maximum deviation of $\sim 2.8\%$ from the April12 data taking period S2 acceptance. This result is compared to the simple counting analysis where the S2 acceptance for simulated neutrons is shown in figure 5.8. The simple counting analysis results a higher S2 acceptance. To take the spectral shape into account by using the adjusted Poisson distribution bias the S2 acceptance.
The β correction results in the *true* expectation value $\lambda'(S1)$ which is similar to cS2 and is measured in photo-electrons. Even after the correction the slope for simulated neutrons is higher compared to the slopes resulting from ^{241}AmBe data. Therefore the same S1 slice (e.g. $3\,\text{pe} \leq S1 \leq 4\,\text{pe}$) would



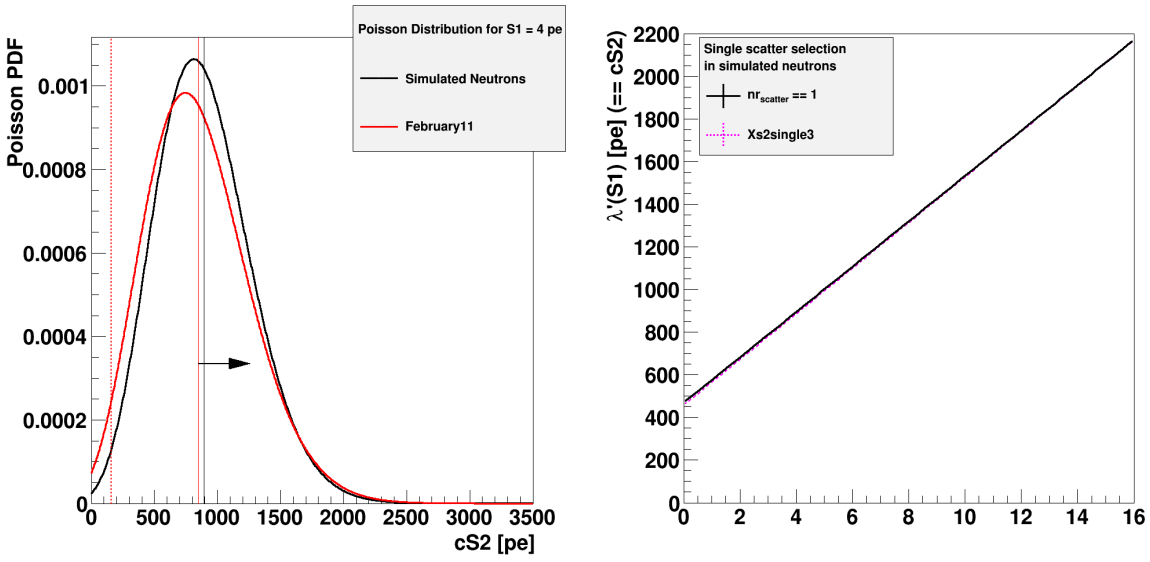
(a) (b)

Figure 5.17 – Figure 5.17a shows the calculated S2 acceptance for simulated neutrons (black) and compares that to the calculation for the February11 and April12 data sets. In figure 5.17b, the residuals are shown between simulated neutrons and February11/April12 data taking periods are shown.

yield a higher true expectation value cS2. During the application of the Monte Carlo method of section 5.4, the linear approximation of $\lambda(S1)$ is used to create a Poisson distribution to describe cS2 for a uniformly distributed value of S1. The chance of a random $cS2 > 150\,\text{pe}$, drawn from this Poisson distribution increases with a larger slope and/or constant of the polynomial. This is observed for the simulated neutrons. That principle is shown in figure 5.18a where such a Poisson probability density function (PDF) is drawn for $S1 = 4\,\text{pe}$ from February11 and the simulated neutrons. The PDF of equation (5.13) is derived from equation (5.2), where the argument of linear approximation $\lambda(S1) = a \times S1 + b$ is put in and α normalises the distribution to 1 to define a valid PDF. The increased slope and/or constant shifts the expectation value λ. This results in a reduced probability to draw events for $cS2 < 159\,\text{pe}$.

$$P(cS2|\alpha,\beta,\lambda,S1,a,b) = \alpha \frac{e^{\frac{-(a\times S1+b)}{100}} \left(\frac{a\times S1+b}{100}\right)^{\frac{cS2}{\beta}}}{\Gamma\left(\frac{cS2}{\beta}+1\right)} \tag{5.13}$$

$$P(cS2 > 150\,\text{pe}|\alpha,\beta,\lambda,S1,a,b) = \int_{150}^{\infty} \frac{\alpha\, e^{\frac{-(a\times S1+b)}{100}} \left(\frac{a\times S1+b}{100}\right)^{\frac{cS2}{\beta}}}{\Gamma\left(\frac{cS2}{\beta}+1\right)} \, d\,cS2 \tag{5.14}$$



(a) (b)

Figure 5.18 – Figure 5.18a: Example of how the simulated neutron involve, compared to the data and what is the impact on the expectation values λ. This behaviour influences the probability to draw random events down to zero photo-electrons for cS2. A comparison between the single scatter cut for simulated neutrons (5.18b) and a single S2 signal cut (Xs2single3), which is developed for data, shows that there are less differences in the polynomial description.

The probability to draw a random $cS2 > 150\,\text{pe}$ from the distribution in equation (5.13) is given in equation (5.14). The probability shifts from February11 $\rightarrow$ Simulated Neutrons and April12 $\rightarrow$ Simulated Neutrons are expressed by a residual which is also drawn in figure 5.17b by the dotted lines. The residuals of the probability shift follow the Xs2peaks acceptance residuals from the Monte Carlo method. Therefore, a higher Xs2peaks acceptance for simulated neutrons follows the increased cS2 for a given S1, compared to February11/April12 data taking periods.

Origin of the Xs2peaks acceptance rise in simulated neutrons is the stringent event selection by the single neutron scatter criterion of $\text{nr}_{\text{scatter}} \stackrel{!}{=} 1$. The analysis of real neutron data uses the Xs2single3 cut which selects events where a single S2 signal is followed by strongly suppressed S2 signals on the same waveform and hence the Xs2single3 cut still allows multiple scatters with a finite probability. The simulation follows each possible scatter of a single event and tracks their deposited energies. The waveforms of real neutron events are analysed by the peak-finder algorithm of *xerawdp* to identify S2 peaks. Besides the largest S2 signal on the waveform which is responsible for triggering the event, further S2 peaks do not necessarily originate from the same physical interaction. Hence, multiple

S2 signals from neutrons are only suppressed by the Xs2single3 cut but not fully eliminated, and the ^{241}AmBe *single scatter* data are not purely single interactions.

$$\text{Xs2single3} := \text{ALT}\,(S2_2;0) < 70 + \frac{S2_1 - 300}{100} \tag{5.15}$$

To compare these different ways of cuts to select single S2 events the Xs2single3 cut is applied to simulated neutrons instead of applying $\text{nr}_{\text{scatter}} \stackrel{!}{=} 1$. All other cuts which are described in section 5.5.1 are the same. The same analysis method with the Poisson fit to the S1 slices is done to calculate the linear approximation of $\lambda'(S1)$ as in figure 5.10b.

The result is given in figure 5.18b where both polynomials do not differ much from each other. The expected outcome would be that the Xs2single3 cut on simulated neutrons yields a polynomial $\lambda'(S1)$ which is similar to the data of February11 and April12 if the simulation describes the data optimally. In this sense the Xs2single3 cut works differently for simulated neutrons and real neutrons. The cut definition is given in equation (5.15). The "ALT" statement in the cut definition (see equation (5.15)) describes the possibility that no second scatter is simulated, then the left site of equation (5.15) is zero and the cut criterion is fulfilled. This is equal to $\text{nr}_{\text{scatter}} \stackrel{!}{=} 1$. A second scatter result in $S2_2$ which has to be smaller than the right side of equation (5.15) to fulfil the cut criterion. That is equal to two detected scatters in the simulation $\text{nr}_{\text{scatter}} \stackrel{!}{=} 2$ with an additional S2 signal limit on the second scatter. Based on this result a possible *Xs2single3* cut for simulated neutrons is defined according to equation (5.16). The second part of the cut introduces a similar cut condition on the second S2 peak such it is done in the original cut definition. This definition uses the parameter α as a suppression factor of the second S2 peak. Due to the reason that the simulation allows multiple scatters of a higher order, the cut definition in equation (5.16) can be extended to higher orders. With

$$\text{Xs2single3}_{\text{Simulation}} := \left[\text{nr}_{\text{scatter}} \stackrel{!}{=} 1\right] \,||\, \left[\text{nr}_{\text{scatter}} \stackrel{!}{=} 2 \;\&\; S2_2 < \alpha \times S2_1\right] \tag{5.16}$$

In summary there is no specific cut set which is developed for simulated neutrons or WIMPs. Therefore, the observed S2 acceptance which is developed from simulated neutrons will always be higher than for real neutrons, the outcome is understood as an upper limit on the S2 acceptance based on simulated neutrons with a perfect single scatter. The suggested $\text{Xs2single3}_{\text{Simulation}}$ cut for single and double scatters is not used in the further analysis of the WIMP acceptance. WIMPs are supposed to interact once in the TPC and therefore it is enough to request a single scatter interaction.

5.8. Determination of a WIMP Acceptance Based on Simulations

The previous two subsections discuss the Xs2peaks acceptance from neutron calibrations and from simulated neutrons. Basic input is the corrected S2 signal (cS2) which is analysed for S1 signals below 16 pe. Due to the S2 threshold of $S2_{th} = 150$ pe and hence the electron drift time corrected threshold in $cS2_{th}$, the S2-histograms contain a blinded region of $0\,\text{pe} \leq cS2 \leq cS2_{th}$. The applied Poisson distribution fit to the S2-histograms describes the cS2 signal below $cS2_{th}$ analytically.
The *Nuclear Recoil/WIMP Simulation Code* allows to simulate the detector response of S1/S2 signals for a neutron spectrum. Furthermore, it is possible to simulate the detector response for WIMP masses m_χ and their cross-sections σ_χ. Given the good fit of the XENON100 neutron simulation with the data and the observed impact of contamination by neutron multiple scatters, we move on to rely in response and acceptance on Monte Carlo simulations of single scatter WIMP-nucleus interaction.

5.8.1. Monte Carlo WIMP Input from Simulations

The basic input to the *Nuclear Recoil/WIMP Simulation Code* are WIMP mass m_χ and cross-section σ_χ which are needed to calculate the correct rates. The cross-section is multiplied to the rate and hence it is enough to investigate a broad spectrum of WIMP masses. The presented analysis fixes the cross-section to $\sigma_\chi = 10^{-40}\,\text{cm}^2$ (default setting in the *Nuclear Recoil/WIMP Simulation Code* code) and scans masses between $5\,\text{GeV} \leq m_\chi \leq 1030\,\text{GeV}$. This introduces a variation in the number of simulated WIMPs for each WIMP mass. The masses below 100 GeV are scanned in steps of 1 GeV, the masses between $100\,\text{GeV} \leq m_\chi \leq 1030\,\text{GeV}$ are scanned in steps of 10 GeV.

The set of simulated WIMPs is collected for single WIMP masses. Taking a constant cross-section into account, the number of simulated WIMPs varies for each WIMP mass. The decision on the number is arbitrary. The minimum amount of simulated WIMPs for a mass of 5 GeV is ~ 65000 events and the maximum for a mass of 54 GeV is $\sim 13.2 \times 10^6$ events. An overview is given in figure 5.19 where the number of simulated WIMPs is shown for each WIMP mass. When the simulation is carried out, the number of WIMPs is fixed and hence the number of interacting WIMPs in the TPC depends on, e.g. the cross-section. This is why the number of interacting WIMPs in figure 5.19 is lower for e.g. $5\,\text{GeV} \leq m_{chi} \leq 10\,\text{GeV}$. Therefore, the simulation itself is carried out more often to collect enough interactions for a proper statistics, what is the reason for the three jumps in figure 5.19. The first segment has 100, the second segment has 20, the third segment has ten and the forth segment has five single simulations with each between $10000 \leq$ interacting WIMPs ≤ 80000 events. This analysis would be also possible with a fixed number of simulated event per WIMP mass and a varying cross-section.

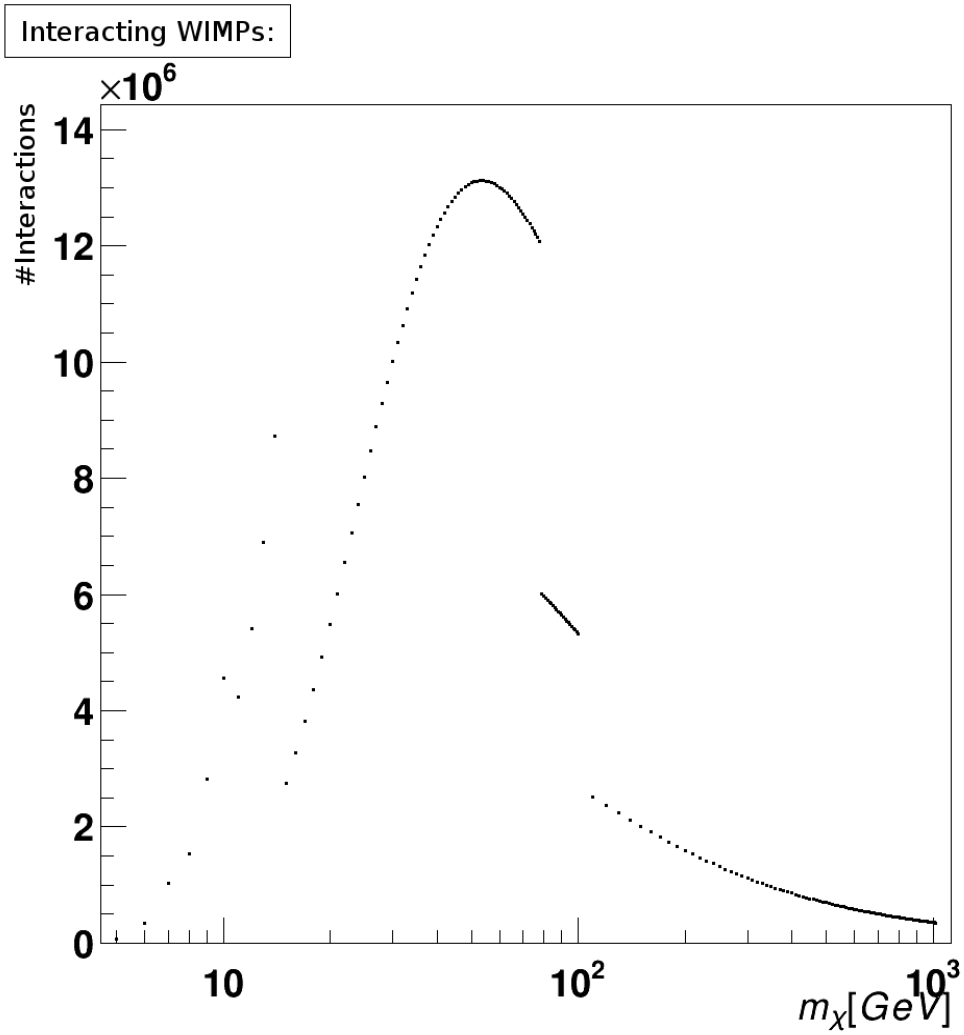


Figure 5.19 – The number of simulated WIMPs in the TPC for a fixed cross-section of $\sigma_\chi = 10^{-40}\,\mathrm{cm}^2$ is divided into four segments. Each one depends on how often the WIMP simulation is carried out. This division allows to collect enough WIMP interactions for each simulated WIMP mass to gain in statistics.

5.8.2. Method to Evaluate the WIMP Acceptance

The WIMP acceptance calculation does not require a complex method as described in section 5.3.1 and 5.3.2. The *Nuclear Recoil/WIMP Simulation Code* allows to simulate WIMPs down to a mass of 5 GeV and describes the TPC response in uncorrected and corrected $(S1, S2)$ pairs. The resulting corrected and uncorrected S2 signals are simulated down to 0 pe. The energy threshold in S2 is the same for WIMPs and neutrons and hence the number of WIMPs is counted above $S2_{th} > 150\,\mathrm{pe}$ for a single WIMP mass and per slice ΔS1. The S1 signal is limited to $0\,\mathrm{pe} \leq S1 \leq 30\,\mathrm{pe}$, where the upper

limit is given by the benchmark search region in the general Profile Likelihood analysis of 2011/12 [15]:

$$\mathrm{S2}^{S1}_{m_\chi} > 150\,\mathrm{pe}, \qquad m_\chi \in [5;1000]\,\mathrm{GeV},\ S1 \in [0;30]\,\mathrm{pe} \tag{5.17}$$

The number of events which pass the criteria (5.17) is n^{S1}_χ which then is normalised by the total number of simulated WIMPs N^{S1}_χ within a given mass m_χ and slice of $\Delta S1$. The normalisation results in a probability density function for a given WIMP mass, p_χ.

$$p^{S1}_\chi = \frac{n^{S1}_\chi}{N_\chi} \tag{5.18}$$

$$A_\chi = \int_{3\,\mathrm{pe}}^{30\,\mathrm{pe}} p^{S1}_\chi dS1 \tag{5.19}$$

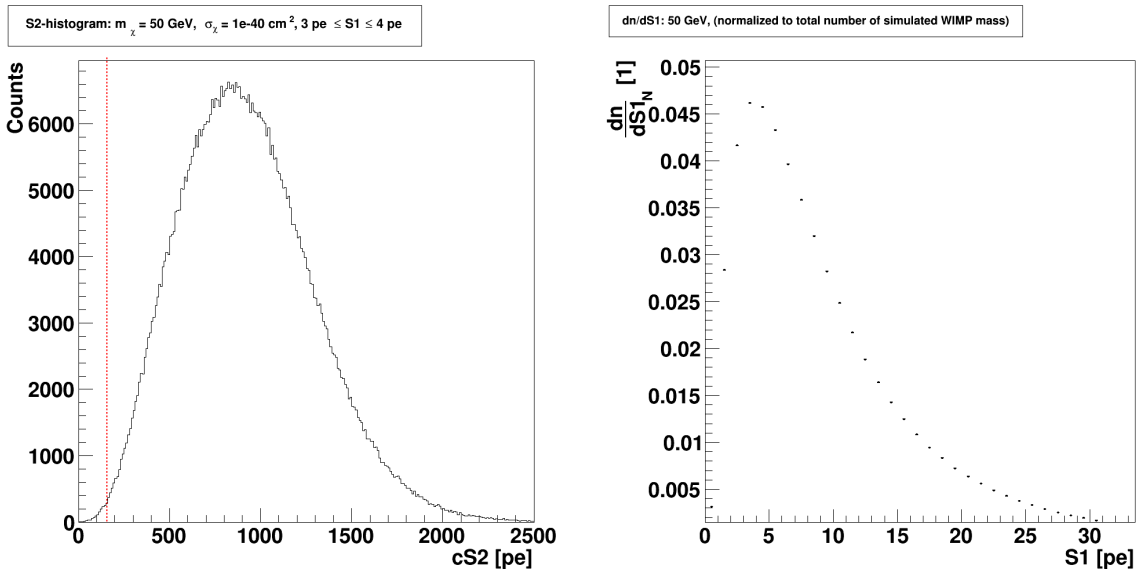
The integrated WIMP acceptance A_χ is calculated according to the detector threshold of $S1 > 3\,\mathrm{pe}$ and integrates equation (5.18) from $3\,\mathrm{pe} \leq S1 \leq 30\,\mathrm{pe}$ as it is shown in equation (5.19). The lower boundary of 3 pe is the analysis threshold in S1 and is determined by coincidence requirements on PMTs to reduce effects from noise and dark counts.

The *Nuclear Recoil/WIMP Simulation Code* allows to also vary the thresholds during the simulation. Therefore this study also includes the two possibilities to lower the S1 and/or S2 threshold to test its impact on the WIMP acceptance. First the S1 threshold is lowered from 3 pe to 0 pe. In reality, this would not be possible as long as $\mathscr{L}_{\mathrm{eff}}$ is measured down to 10 pe ($3\,\mathrm{keV_{nr}}$). The *Nuclear Recoil/WIMP Simulation Code* approximates that energy region by its best fit to $\mathscr{L}_{\mathrm{eff}}$ [23]. The acceptance A_χ is then calculated according to equation (5.19), but with a lower boundary of 0 pe.
In a second step, the impact of the S2 threshold is tested. The S2 threshold depends on the trigger efficiency in XENON100 and therefore the simulation studies the potential of a higher trigger efficiency. We test the WIMP acceptance for a lower threshold of $S2^0_{th} = 100\,\mathrm{pe}$ (see equation (5.17)). In this case, the minimal signal charge size of ~ 7.5 electrons is reduced to 5 electrons.

5.8.3. Results of the WIMP Simulation

In total, 200 WIMP masses are analysed, each with single slices of S1 with a width of $\Delta\mathrm{S1} = 1\,\mathrm{pe}$ and an upper limit of $S1 \leq 30\,\mathrm{pe}$. This yields 6000 single cS2 histograms. An example for such a histogram is given in figure 5.20a, where the red dashed line indicates the S2 threshold in XENON100 of $S2 > 150\,\mathrm{pe}$. The analysis counts events right of the red dashed line.
All S1 slices are summarised in figure 5.20b for a WIMP with $m_\chi = 50\,\mathrm{GeV}$. The S1 spectrum is

normalised to the total number of interacting events for a 50 GeV (example mass) WIMP. All tested WIMP masses are summarised in figure 5.22 as a 2D histogram for the S1 spectra.



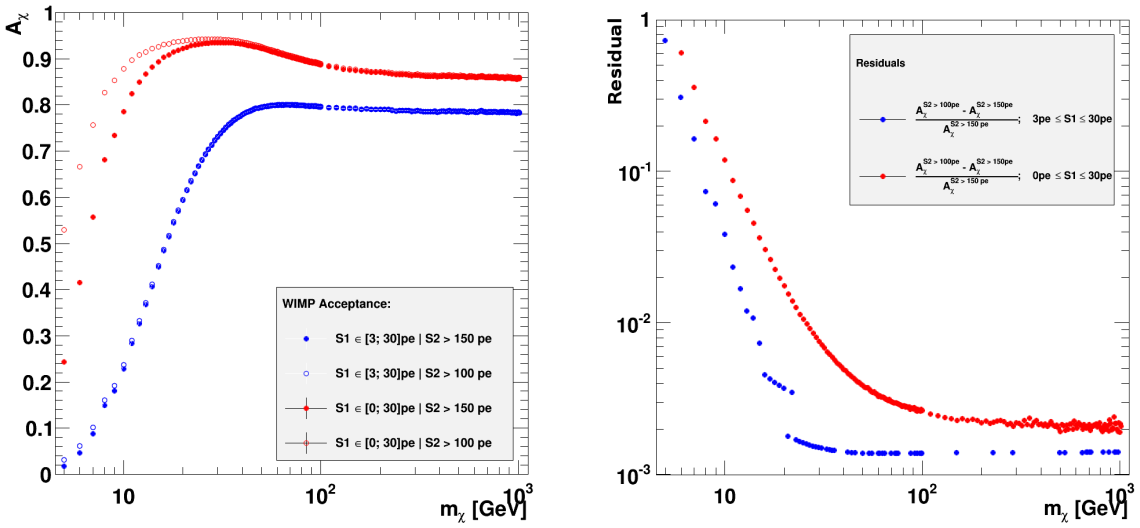
(a) (b)

Figure 5.20 – Figure 5.20a shows an example S2-histogram of a simulated 50 GeV WIMP within a slice of $3\,\text{pe} \leq S1 \leq 4\,\text{pe}$ in S1. In total 30 S1 slices are analysed ($\Delta S1 = 1\,\text{pe}$) for 200 WIMP masses in $5\,\text{GeV} \leq m_\chi \leq 1000\,\text{GeV}$. Figure 5.20b shows the normalised S1 spectrum of a 50 GeV WIMP. The higher the WIMP mass is, the higher is its probability to include higher S1 signals in the recorded S1 and S2 signals of the simulated WIMPs.

The integrated acceptance A_χ as defined in equation (5.19) is presented in figure 5.21a for the detection threshold of $S1 \geq 3\,\text{pe}$ and for an ideal threshold of $S1 \geq 0\,\text{pe}$.

5.8.4. Discussion: Simulated WIMP Acceptance

The integrated WIMP acceptance describes the number of detected WIMP interactions for a certain WIMP mass would be under the 2011/12 science run condition. Figure 5.21a shows the WIMP acceptance for the $3\,\text{pe} \leq S1 \leq 30\,\text{pe}$ range which is similar to the WIMP search benchmark region of the Profile Likelihood analysis of 2011/12 (blue markers). The WIMP acceptance increases up to 80% at a WIMP mass of $m_\chi = 66\,\text{GeV}$ and decreases to 78% for WIMP masses $m_\chi > 200\,\text{GeV}$. The acceptance dross quickly below $m_\chi \lesssim 20\,\text{GeV}$, reaching less than 20% at $m_\chi = 10\,\text{GeV}$ and 2% at $m_\chi = 5\,\text{GeV}$ Under the threshold conditions of $S1 \geq 3\,\text{pe}$ and $S2 > 150\,\text{pe}$, the situation changes if no S1 threshold is considered which requires a measured parameter $\mathscr{L}_{\text{eff}}$ down to 0 pe from external experiments. The S2 threshold is not changed and the outcome is also shown in figure 5.21a (red). There the acceptance



(a) (b)

Figure 5.21 – Figure 5.21a shows the acceptance for four different scenarios: There is no S1 threshold applied and then $S1 = 0\,\text{pe}$ is tested for $S2 > 150\,\text{pe}$ and $S2 > 100\,\text{pe}$ (red dots and red transparent dots). The XENON100 S1 threshold of $S1 > 3\,\text{pe}$ is also shown for $S2 > 150\,\text{pe}$ and $S2 > 100\,\text{pe}$ (blue dots and blue transparent dots). Figure 5.21b shows the residuals of the scenarios if the S2 threshold in XENON100 is lowered to $S2^0_{th} = 100\,\text{pe}$. In both cases, an increase in the acceptance is observed from the rising residual.

is at maximum for a WIMP mass of $m_\chi = 31\,\text{GeV}$ at 93.5%. The acceptance increases to $A_\chi > 90\%$ for WIMP masses of $16\,\text{GeV} \leq m_\chi \leq 75\,\text{GeV}$.
In every case, the WIMP acceptance decreases after its maximum. The reason for that is the analysis benchmark region of up to $S1 < 30\,\text{pe}$. The detector response depends on the WIMP mass and therefore a heavier WIMP reaches to higher energy deposits and hence higher $(S1, S2)$ signals, compared to low mass WIMPs. Extending the analysis range to $S1 > 30\,\text{pe}$ would increase the integrated WIMP acceptance, at the cost of lower signal to background ratio. Background is not considered in this acceptance study. There is barely a decrease in acceptance observed in figure 5.21a and hence the signal gain is little beyond $S1 > 30\,\text{pe}$.

The WIMP simulation shows that the highest acceptance is given for WIMPs in $51\,\text{GeV} \leq m_\chi \leq 100\,\text{GeV}$ for the thresholds of 2011/12. This concludes from the past publication[17], that XENON100 is not sensitive enough in these WIMP masses for spin-independent interactions. The new Dark Matter data set is not yet published. Hence *Nuclear Recoil/WIMP Simulation Code* is not setup yet for the changed run conditions to 2013/14. As soon as the new setup is implemented the Xs2peaks acceptance and the WIMP acceptance will be calculated again.
The *Nuclear Recoil/WIMP Simulation Code* simulates corrected and uncorrected S2 events down to

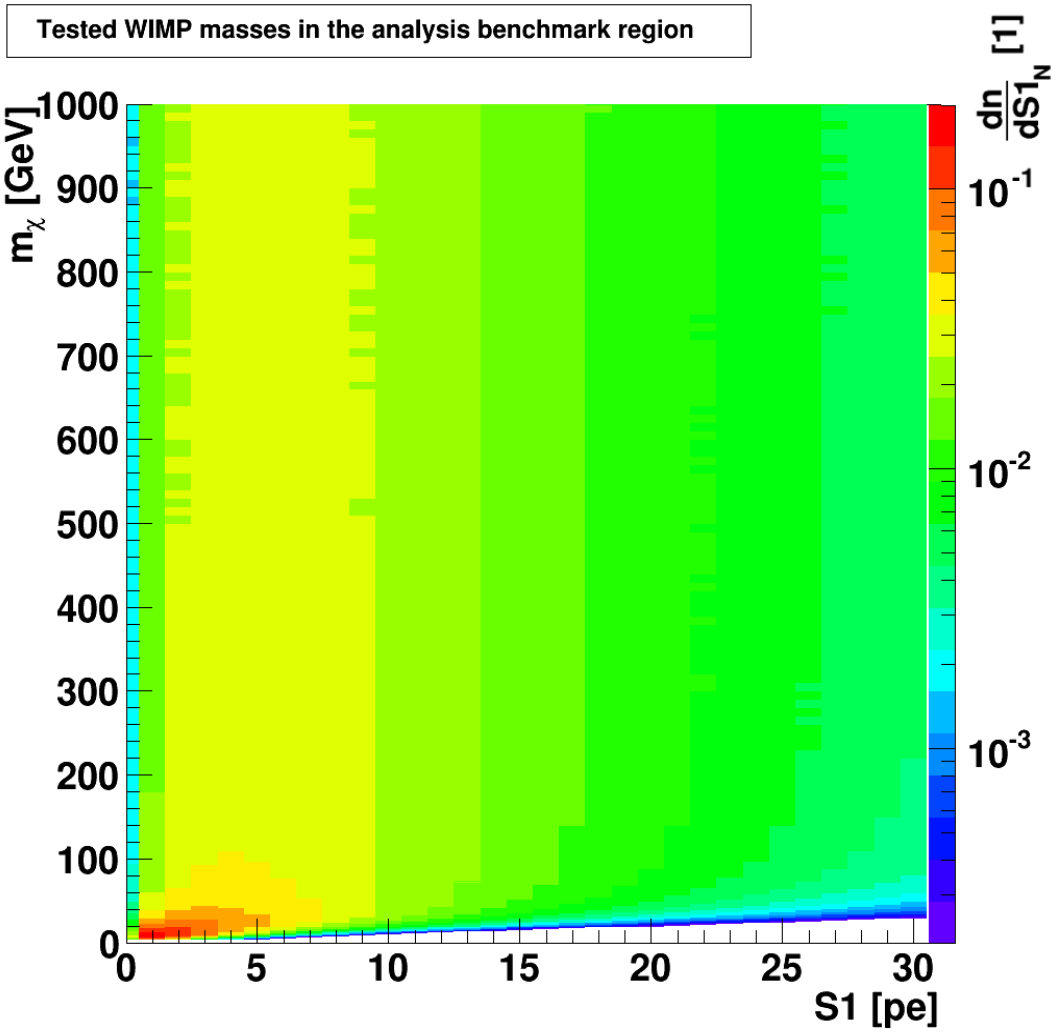


Figure 5.22 – This figure shows a combination of all 200 tested WIMP masses within $5\,\text{GeV} \leq m_\chi \leq 1000\,\text{GeV}$. WIMP masses below 100 GeV are scanned with $\Delta m_\chi = 1\,\text{GeV}$ and masses above 100 GeV are scanned in steps of $\Delta m_\chi = 10\,\text{GeV}$.

0 pe and based on the idea of lowering the S2 threshold by 50 pe to $S2_{th}^{low} = 100\,\text{pe}$. This scenario is shown for both S1 ranges, $0\,\text{pe} \leq S1 \leq 30\,\text{pe}$ (red) and $3\,\text{pe} \leq S1 \leq 30\,\text{pe}$ (blue) by the transparent markers in figure 5.21a. To determine the difference, the figure 5.21b shows the residuals. A strong increase in the WIMP acceptance is not observed for $3\,\text{pe} \leq S1 \leq 30\,\text{pe}$ and hence a further lower of the S2 threshold will not affect the WIMP acceptance in XENON100 with the selected S1 range for most WIMP masses. In case of taking also no threshold in S1 into account, $0\,\text{pe} \leq S1 \leq 30\,\text{pe}$ the residuals are $> 1\%$ for WIMP masses of $5\,\text{GeV} \leq m_\chi \leq 25\,\text{GeV}$ and reach against its maximal residual for a WIMP of $m_\chi = 5\,\text{GeV}$.

In conclusion, an increase in the WIMP acceptance is expected notably if the S1 threshold will be lowered for low mass WIMPs. As shown in figure 5.21b by the blue markers, a higher trigger efficiency to allow a lower S2 threshold will also increase the WIMP acceptance, but this will merely affect WIMPs with $m_\chi \lesssim 7\,\text{GeV}$. Only for WIMP masses of $m_\chi < 8\,\text{GeV}$ would the effect surpass 10%, reaching a 70% increase at a 5 GeV WIMP. Clearly, it is the S1 signal that is limiting the low mass WIMP sensitivity and not the S2 signal. Due to the high number of simulated WIMPs (see figure 5.19), the statistical uncertainties are small and in the order of 10 times lower than the calculated WIMP acceptance.

5.9. Summary

This chapter described the acceptance of the Xs2peaks cut in detail. The cut defines a threshold in uncorrected S2 signals. It is important to set the S2 threshold to its smallest possible value to maximise the acceptance. On the other side, the S2 threshold has to reach a minimum where the observed S2 signals are identified and recorded as *trusted* signal. Therefore the S2 trigger efficiency is used. In case of identifying S2 signals by a probability of almost 100%, the according S2 threshold is selected. This threshold was set to $S2 > 150\,\text{pe}$ in the data taking periods 2011/12 and 2013/14.
The cut results in an overall high acceptance with the largest variation for small S1 signals. The situation is more than one might initially expect complicated, due to the fact that the S2 signals are position dependent and furthermore, no information about the spectral shape of the corrected and uncorrected S2 is available below $S2 < 150\,\text{pe}$. To deal with this situation, the variable of observation becomes cS2 in slices of $\Delta S1$. The observed cS2 distribution is approximated by an extended Poisson distribution which contains further fit parameters. Via minimisation of the likelihood from the likelihood fits results the best choice for the fit parameters. This method results a relation between the expected amount of cS2 signal for each S1 which is approximated by a linear fit. This linear function is used to calculate the S2 acceptance by a Monte Carlo simulation. This technique allows to sample random cS2 values down to $cS2 = 0\,\text{pe}$ for a random S1 value. The simulated cS2 is position corrected via the electron lifetime to S2. The acceptance is calculated by counting the number of simulated S2 signals which exceeds the $S2_{th}$.
This method is presented on ^{241}AmBe calibration data from three calibration data taking periods: February11, April12 and May13. The outcome of the S2 acceptance shows a good agreement between the February11 and April12 data sets. Both calibrations were taken during the same data taking period 2011/12 and therefore the acceptance is not expected to change much. The higher electron lifetime in the following data taking period 2013/14 results a higher Xs2peaks acceptance.
A tool to simulate neutrons is used to verify the general method. To begin with the simulated neutrons

are distributed in slices of $\Delta S1 = 1\,\text{pe}$ and the S2 acceptance is based on counting the number of events which are above 150 pe. This results a high S2 acceptance but neglect the spectral shape, in particular for $S2 < 150\,\text{pe}$. To take the spectral shape with simulated neutrons into account, the same analysis approach such for the data is used. The method itself works quite well on the simulated neutron data, only its outcome is not comparable with 2011/12 data sets directly. The selection criterion of simulated neutrons is different, in particular for the selection of *true* single scatters. This is not possible on data and only a single scatter cut is available which searches for large S2 peaks on the waveform with a highly suppressed second or higher order S2 peak. In contrast, the simulation is background free and allows to select single scatters at a perfect level which is much better than it is possible with data. A dedicated cut set for simulated neutrons is not yet available and therefore the result is understood as an upper limit on the S2 acceptance based on a simulation.
The same tool which simulates neutrons is also used to simulate WIMPs in the XENON100 TPC. Based on that simulation a simple counting experiment is carried out to analyse the S2 threshold and its impact on simulated WIMPs. The question about multiple scatters is not relevant because it is expected that WIMPs interact only once in the TPC. It turns out that the resulting WIMP acceptance follows the expectation from the data analysis for neutrons. The WIMP acceptance has its maximum below a WIMP mass of $m_\chi = 50\,\text{GeV}$ for the given S2 threshold.

6. The Electronic Recoil Background

"Ooh, *Dark Matter*. We better bring a flashlight."

Dr. Rajesh Ramayan Koothrappali,
The Big Bang Theory, "The Pirate Solution"

6.1. Motivation

This section is dedicated to the characterisation of the electronic recoil (ER) background in XENON100. Weekly calibration procedures (see section 3.6.2) are necessary to estimate the event rate, strength and kind of the ER background and to measure the detector response to ER calibration sources. Therefore in each data taking period, two well-known calibration sources are used to get an energy-dependent response of the detector. Prior work has been presented in past publications of the XENON100 collaboration [15], [24].

The characterisation of the ER background is important for a full understanding of the detector. The model, which is discussed in this work, shows that ER event distribution is understood and the model can be used in further statistical analysis. This section introduces a new phenomenological model of the ER background and tries to determine its uncertainties relating to ER background data. Finally, the robustness of the WIMP exclusion limit is tested under the aspect of comparing this ER background model to a previously developed ER background model.

A first characterisation of the ER background allows us to describe the event distribution of ^{60}Co and ^{232}Th in discrimination space in slices of cS1 by a Gaussian approximation. Besides this first approximation of the data, another additional event class is observed. These *anomalous leakage events* appear in contrast to their Gaussian probability in the flattened discrimination space (see 3.9), outside of a 3σ standard deviation. Because of these unusual leakage events, which are not physically explained yet, an adoption of the Gaussian background model is necessary.

In the previous 2011/12 data set, a one-dimensional model of the anomalous leakage events was developed. This model depends only on $cS1$ with an exponential drop and not on the flattened space Y. The following section points to a model of the ER background which includes the full data space ($cS1$ and Y) and contains the observed irregularities. Besides the development of an alternative ER background model which can be used in the Profile Likelihood approach to calculate a limit, this work is extended to compare the previous ER background model with the alternative one.

The advantage of such a 2D model is the joint description of the ER background and the anomalous leakage events by one probability density function (PDF) if the model is complete. Based on the number of expected ER background events in the Dark Matter data, the PDF describes their distribution in the Profile Likelihood analysis [19] (PL) which is used in the latest XENON100 data analysis. The final ER background model is an input to the PL analysis (see section 3.11), but due to its general description it can also be used in alternative analysis methods such as a Bayesian approach [103].

6.2. Data Selection

The following section uses the ER background calibration data (^{232}Th and ^{60}Co). A summary of both data sets is given in table 6.1 for data taking periods 2011/12 and 2013/14. Both are controlled by weekly ER calibrations. In total 2131 hours of calibration data were taken. The data is selected by the "standard" data taking conditions, for example the applied cathode/anode voltage is $-16/+4.5\,\mathrm{keV}$ at the meshes.

Data taking period	Calibration source	Number of data sets	Corrected lifetime [hours]	Corrected events
2011/12	^{232}Th	194	613.15	89303641
2011/12	^{60}Co	186	550	85457088
2013/14	^{232}Th	277	468.89	82024367
2013/14	^{60}Co	238	499.26	69320042
Total	–	895	2131.3	326105138

Table 6.1 – Overview about the selected data sets of ^{232}Th and ^{60}Co during 2011/12 and 2013/14. The term *corrected* refers to a correction which is applied after the data reprocessing with *xerawdp*. This correction filters events which are produced due to high voltage micro dis-charges in the TPC.

Section 3.7 introduces the most important cuts and their definitions. Each analysis is dedicated to a given set of pre-defined cuts which depend e.g. on the detector stability during the data taking period. This set of cuts may change for the individual data taking periods (see section 3.7). The characterisation of the ER background of 2011/12 is done after applying the cuts of table 6.2 and for 2013/14 the cut set of table 6.3 is used. The cut set of 2013/14 is similar to cut set of 2011/12.
Both cut sets contain the typical cuts which are used in the 2011/12 and 2013/14 Dark Matter analysis for the event selection. That Dark Matter cut selection for ER background data is still valid because it matches the most important event selection criterion such as single S1 and single S2 peak selection, shelf-shielding and the Dark Matter search region in cS1. Figure 6.1 shows the remaining events of the ER data for 2011/12 (figure 6.1a) and 2013/14 (figure 6.1b). Both data sets of the calibration sources ^{232}Th and ^{60}Co are presented together. The data is shown in the data space of Y and $cS1$.

Cut type	Data taking period: 2011/12
S1-signal	Xs1coin2, Xs1width0, Xs1single4
S2-signal	Xs2width8, Xs2top0, Xs2single3, Xs2peakpos0
Noise and veto	Xsignalnoise_comb, Xhighlog1, Xlownoise0_m, Xentropy0, Xveto2, XPL013_97
Fiducial volume and position reconstruction	Xposrec1, Xs2chisquare0, X34kg2

Table 6.2 – The cut selection of 2011/12 to select ER events in ^{232}Th and ^{60}Co data sets. The cut set is designed to select single S1- and S2 signals in the fiducial volume of the TPC. Noise and wrong reconstructed events are also reduced 3.7.

Cut type	Data taking period: 2013/14
S1-signal	Xs1coin2, Xs1width0, Xs1single5
S2-signal	Xs2width9, Xs2top0, Xs2single3
Noise and veto	Xsignalnoise5, Xhighlog1, Xlownoise0_m, Xentropy0, Xveto2
Fiducial volume and position reconstruction	Xposrec1, Xs2chisquare0, X34kg2

Table 6.3 – The cut selection of 2013/14 to select ER events in ^{232}Th and ^{60}Co data sets. The cut set is designed to select single S1- and S2 signals in the fiducial volume of the TPC. Noise and wrong reconstructed events are also reduced. Some cuts have changed between two data taking periods to take minor changes in the experimental setup into account 3.7.

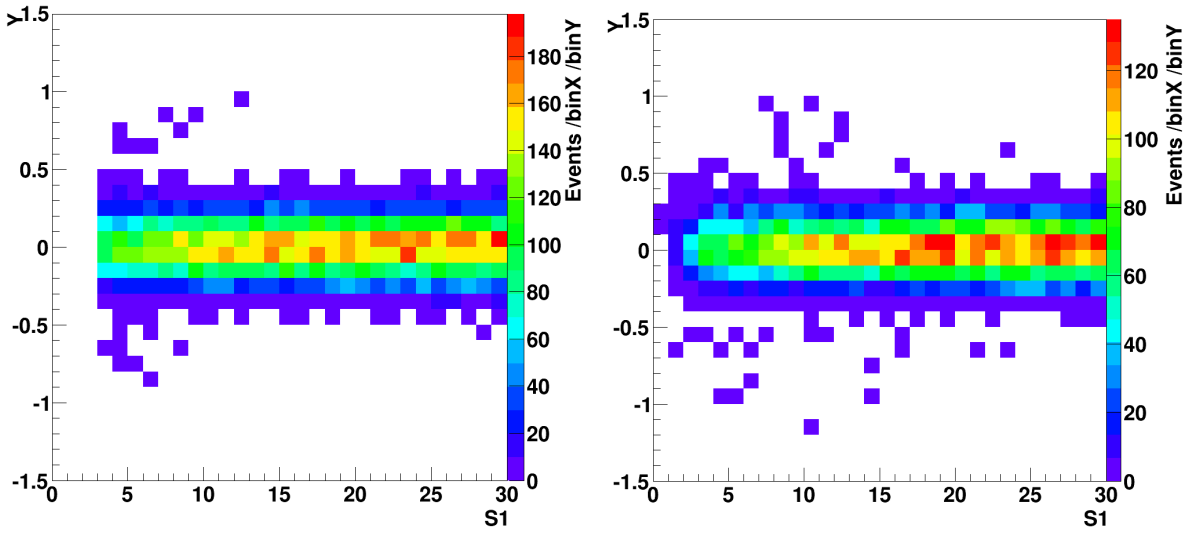
6.3. Nomenclature

The following section uses the nomenclature of table 5.3 (see chapter 5.3). Due to corrections of the S1 and S2 signal (see sections 3.4.3 and 3.4.4), the model is based on the corrected signal. To simplify the following discussions and description the name signals refer to the corrected signals if nothing else is mentioned.

6.4. The Electron Recoil Background Model

The model of the ER background and its anomalous leakage events is built-up in two steps. A first order approximation of the ER background is done by Gaussian distributions. Additional anomalous leakage events and an observed asymmetry are described in the second step. The residual is analysed

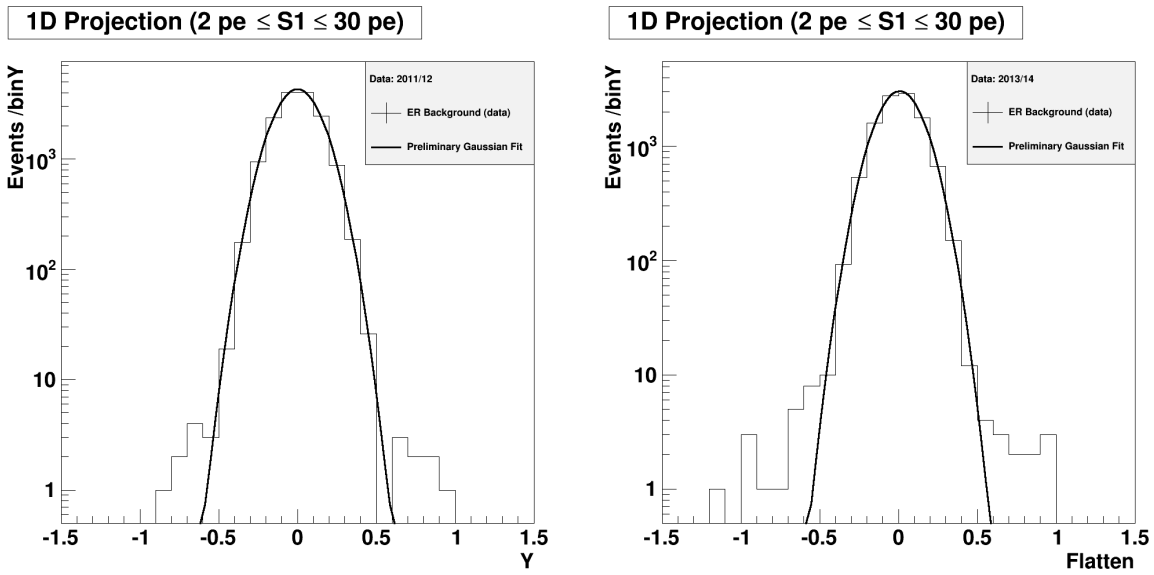
by a 2D Gaussian kernel. That kernel allows to smear the anomalous leakage in the discrimination space. The resulting model of the anomalous leakage events, which is called *leakage model*, also contains a description of the asymmetry within the Gaussian ER background model. By adding up the ER background model and the leakage model, a complete model of the ER background is created.
The ER background is shown in figure 6.1a and 6.1b for 2011/12 and 2013/14 in a binned histogram. The bin size for this analysis is chosen to $\Delta S1 = 1\,\text{pe}$ (x-axis) and $\Delta Y = 0.1$ (y-axis). The chosen binning of Y depends on the instrumental resolution which is related to the width of the Gaussian approximation of the ER background. A detailed overview about the Gaussian approximation is given in subsection 6.4.2. The binning of Y is smaller than the observed Gaussian standard deviation σ_{ER} (see figures 6.2).
The final background model is described by a probability density function which is presented for a data spread of $3\,\text{pe} \leq S1 \leq 30\,\text{pe}$ for 2011/12 and $2\,\text{pe} \leq S1 \leq 30\,\text{pe}$ for 2013/14. The limitation in S1 is given by a signal detection threshold. The data spread in the y-axis of $-1.5 \leq Y \leq 1.5$ results from two observations: The lower boundary overlaps with the S2 threshold of $S2_{th} = 150\,\text{pe}$ and assures that lower leakage events are included in the ER background model. The upper boundary is derived from observations. There are no further leakage events observed for $Y > 1.5$.



(a) ER background: 2011/12

(b) ER background: 2013/14

Figure 6.1 – The 2011/12 and 2013/14 data selections after the cut list is applied to the data. Each data set contains data of the two calibration sources ^{232}Th and ^{60}Co. The leakage events can be seen outside 3σ deviation from the ER background. The binning is chosen to: $1bin = 1\,\text{pe}$ (x-axis) and $1bin = 0.1$ (y-axis).



(a) ER background: 2011/12

(b) ER background: 2013/14

Figure 6.2 – Figure 6.2a shows a projection on the y-axis with a preliminary Gaussian fit to describe the event distribution of 2011/12 within $3\,\text{pe} \leq S1 \leq 30\,\text{pe}$. Based on the preliminary Gaussian fit, the deviation is $\sigma \sim 0.14$. Hence there are more events observed outside a 3σ band in the data than it is expected by the Gaussian model. Figure 6.2b shows the same result ($\sigma = \sim 0.14$) for 2013/14. Based on the numbers in table 6.4, the leakage events do not fit the Gaussian background model.

After the data selection and cuts of section 6.2 are applied, the number of events is determined (see table 6.4). Furthermore a projection on the y-axis allows to observe the Gaussian distribution. The y-projections are shown in figure 6.2a and 6.2b. In addition a preliminary fit is carried out. The fit shows that to many events are observed outside a 3σ standard deviation with a standard deviation of $\sigma \sim 0.14$. Based on the total number of observed events for 2011/12 (15128 events) and 2013/14 (10589 events), it turns out that the number of observed events outside a 3σ standard deviation (48 events for 2011/12 and 45 events for 2013/14) differs from the expected numbers if the ER background is analysed in slices of $\Delta S1 = 1\,\text{pe}$. The total number of expected leakage events is given in table 6.4. That overestimation of the ER background outside a band of $Y < -3\sigma$ and $Y > 3\sigma$ leads to the assumption that another background contributes to the total ER background, the anomalous leakage events.

	S1 threshold	Total events	Anomalous Leakage ($Y < -3\sigma$)	Anomalous Leakage ($Y > 3\sigma$)	Total leakage events	Expected Gaussian Leakage
2011/12	3 pe	15128	23	25	48	41
2013/14	2 pe	10589	29	26	55	28

Table 6.4 – Overview about the observed events from the ^{232}Th and ^{60}Co calibrations in 2011/12 and 2013/14. In addition the leakage event counts are compared to the expected number of events from a pure Gaussian ER background (outside a 3σ standard deviation). This is a hint to an additional background component, the anomalous leakage.

6.4.1. The Anomalous Leakage in the ER Background

The origin of the anomalous leakage events is unknown. A detailed description is not possible due to the few events which are observed yet. Nevertheless, there are different approaches which aim to explain this additional component in the ER background. Due to the few observed events, it is not possible to restrict them to a specific fiducial volume, time dependence or source dependency. Due to the S2 threshold in XENON100 that uncorrected S2 signals below 150 pe are not used for further data analysis, there is a certain probability left that anomalous leakage events origin from this additional event distribution. Figure 6.3 shows the ER background for the 2013/14 data set with the same set of cuts of table 6.3, but without the *Xs2peaks* cut. The ER background data set from 2011/12 shows a similar behaviour. Another event population appears: The *lowS2* events. The anomalous leakage is between both populations and another model of the lowS2 events would allow to determine the probability that the anomalous leakage events emerge from the lowS2 event region.
The characterisation of the lowS2 events is ongoing work and have to take additional issues into

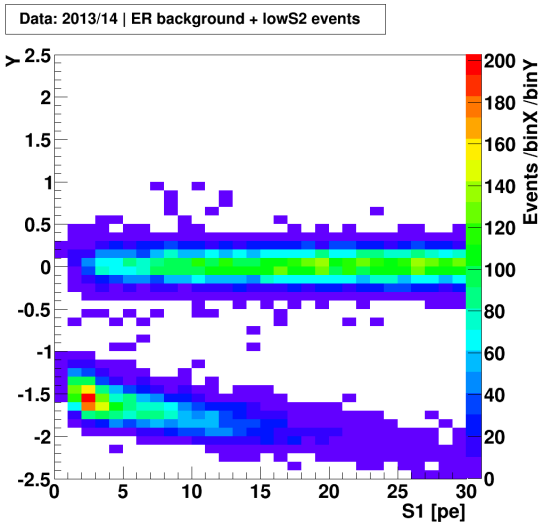


Figure (6.3): Another "lowS2" population arise from removing the S2 energy cut. There is a certain probability left that the anomalous leakage origin from this population and only a few events exceed the S2 energy threshold.

account. The S2 trigger efficiency below 150 pe decreases (see figure 3.17) and further correction are necessary to calculate the true number of S2 signals below 150 pe. To make things worse, not every data selection cut of table 6.2 and 6.3 is defined below 150 pe what requires further studies when lowS2 events are analysed together with the ER background data. The lowS2 event distribution origin from single, double or higher electrons. These electrons leave a typically S2 energy of $S2 = 19.67$ pe with an almost 80% detection efficiency [22].

Another possibility to explain the origin of the anomalous leakage are so called "BiPo" events. This class of events was studied in detail by the XENON collaboration. Due to the Radon background in the experimental setup, the isotopes ^{220}Rn and ^{222}Rn induce along the decay down to lead (^{208}Pb and ^{210}Pb) a specific signal in the TPC which consist of two large peaks from α which are followed by two smaller peaks from β^-'s. In detail, the Bismuth isotopes decay into Polonium along the decay chain and therefore the name "BiPo" is chosen. Extended studied on the BiPo event selection criterion, detection efficiency and the approach to model this event class is summarised in [113]. The resulting BiPo data, which are independent from the ER background studies, shows also a contribution to the region where the lowS2 are located. Therefore this class of events could be a suitable explanation for the anomalous leakage events.

6.4.2. Gaussian ER Background Model

The ^{232}Th and ^{60}Co data sets are analysed together. The XENON100 TPC has a given energy resolution in S1 and therefore the ER background model is made in dependence of $S1$. Therefore the ER data sets (see figure 6.1a and 6.1b) are separated in 30 slices of each $\Delta S1 = 1\,\text{pe}$. Each data slice of $\Delta S1$ is shown in figures A.1, A.2, A.3 and A.4 for 2011/12 data and A.5, A.6, A.7 and A.8 for 2013/14 data (see appendix A). Figure 6.4 presents an example of one of these $\Delta S1$ slices. The selection is made for $4\,\text{pe} \leq S1 \leq 5\,\text{pe}$ from the 2011/12 data set. The overall shape of the data distribution is Gaussian as it is already sketched in figures 6.2a and 6.2b. The Gaussian fit to the data is shown by the black line. The fit range is chosen to the expected $3\sigma_{ER}$ standard deviation from the expectation value μ_{ER}. Figure 6.5c presents the single values of σ_{ER} for each slice of $S1$ for 2011/12 and 2013/14. From these figures, the Gaussian width is about constant with $\sigma_{ER} = \sim 0.14$ for 2011/12 and 2013/14. The exact mean values are $< \sigma >_{ER,2011/12} = 0.138$ and $< \sigma >_{ER,2013/14} = 0.135$. The fit range is then $3\sigma_{ER} = \pm 0.42\,\text{pe}$. The standard deviation σ_{ER} describes the width of the ER background in Y. This information is the most important outcome of the description of the ER background. The region of interest for the WIMP search is below $Y < -3\sigma$ in the ER background in Y direction and therefore it is necessary to understand the ER background itself and to estimate how many events are expected outside a certain deviation from the expectation value μ_{ER}. Figure 6.5a is the amplitude of the Gaussian fit in events per $\Delta S1$ and ΔY. The amplitude describes the scale factor which is used to seize up the Gaussian distribution to the according number of counted events. That parameter depends on the binning. Due to the decreasing acceptance of the detector at low S1 energies, the number of available events in each S1 slice for fitting decreases as shown in figure 6.1 or 6.5a. On the other side, a saturation is already reached for $S1 > 17\,\text{pe}$ and the number of events per S1 slice becomes almost equal. Due to the definition of Y (equation (3.34)), the expectation value is $\mu_{ER} \sim 0$ from the single Gaussian fits (compare with figure 6.4 and see figure 6.5b). Deviations from zero in the fit parameter μ_{ER} are explained by the band flattening method itself. There is no perfect band flattening possible and therefore the band flattening depends on the number of iterations itself.
The reduced χ^2 (figure 6.5d) measures the goodness of the single fits. As long as the reduced χ^2 depends on the degrees of freedom (NDF), the number of bins in each histogram of S1 influences the result later. The data are approximated by a Gaussian distribution and due to the shape it is expected that $\chi^2/_{NDF} \sim 1$, indicating a good fit. Subsection 6.10 discusses the goodness of the model.
The detection S1 threshold is set to 3 pe in the Profile Likelihood analysis in 2011/12 and to 2 pe 2013/14. This means that events with $S1 < 3\,\text{pe}$ are observed and presented in the study of the ER background, but the ER background model will be restricted to the boundaries of the data analysis and their thresholds. Ongoing discussions and studies about lowering the $S1$ threshold from 3 pe to 2 pe are the reason why the ER background model of 2013/14 includes the S1 energy slice $2\,\text{pe} \leq S1 \leq 3\,\text{pe}$,

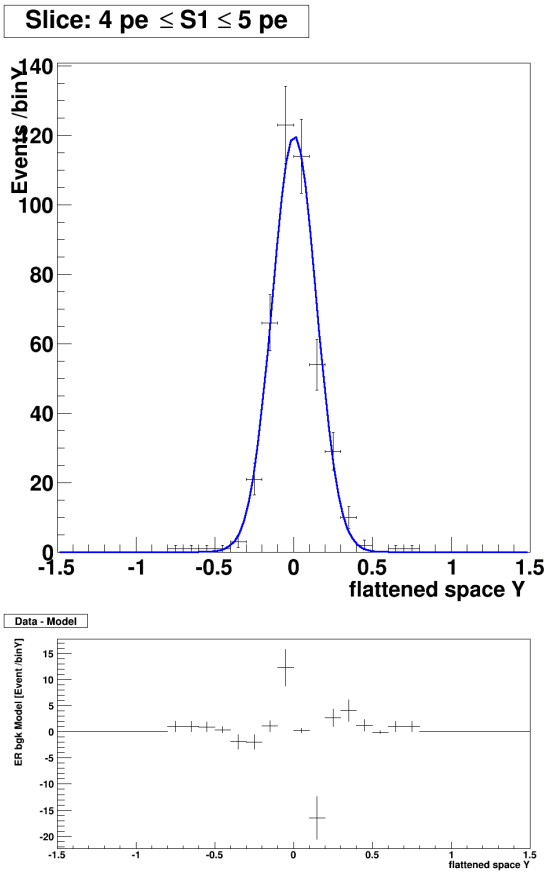


Figure 6.4 – Example for one distribution of the flattened space variable Y for the S1 energy of $4\,\text{pe} \leq S1 \leq 5\,\text{pe}$ (upper figure) for the data taking period of 2011/12. The data is described by a Gaussian model which fits the data very well. The lower figure shows the residual: "Data - Model". Therefore the Gaussian fit is subtracted bin by bin from the data distribution. The residual contains the leakage events. The remaining S1 energy slices are presented in section A.1

.

which was not used in 2011/12. As long as this is an ongoing discussion this extension is a possibility and not mandatory. The final ER background model is influenced by the threshold in $S1$ and the edges of the data space.

The outcome of the single fits in every $S1$ slice is summarised in figure 6.6 for 2011/12 and 2013/14 by a 2D histogram. The single parameters "amplitude", μ_{ER} and σ_{ER} describe the Gaussian background very well and result in a first order approximation of the ER background. Both Gaussian approximations have two features, which are both addressed in the next step:

1. There are no anomalous leakage events in the Gaussian background model.

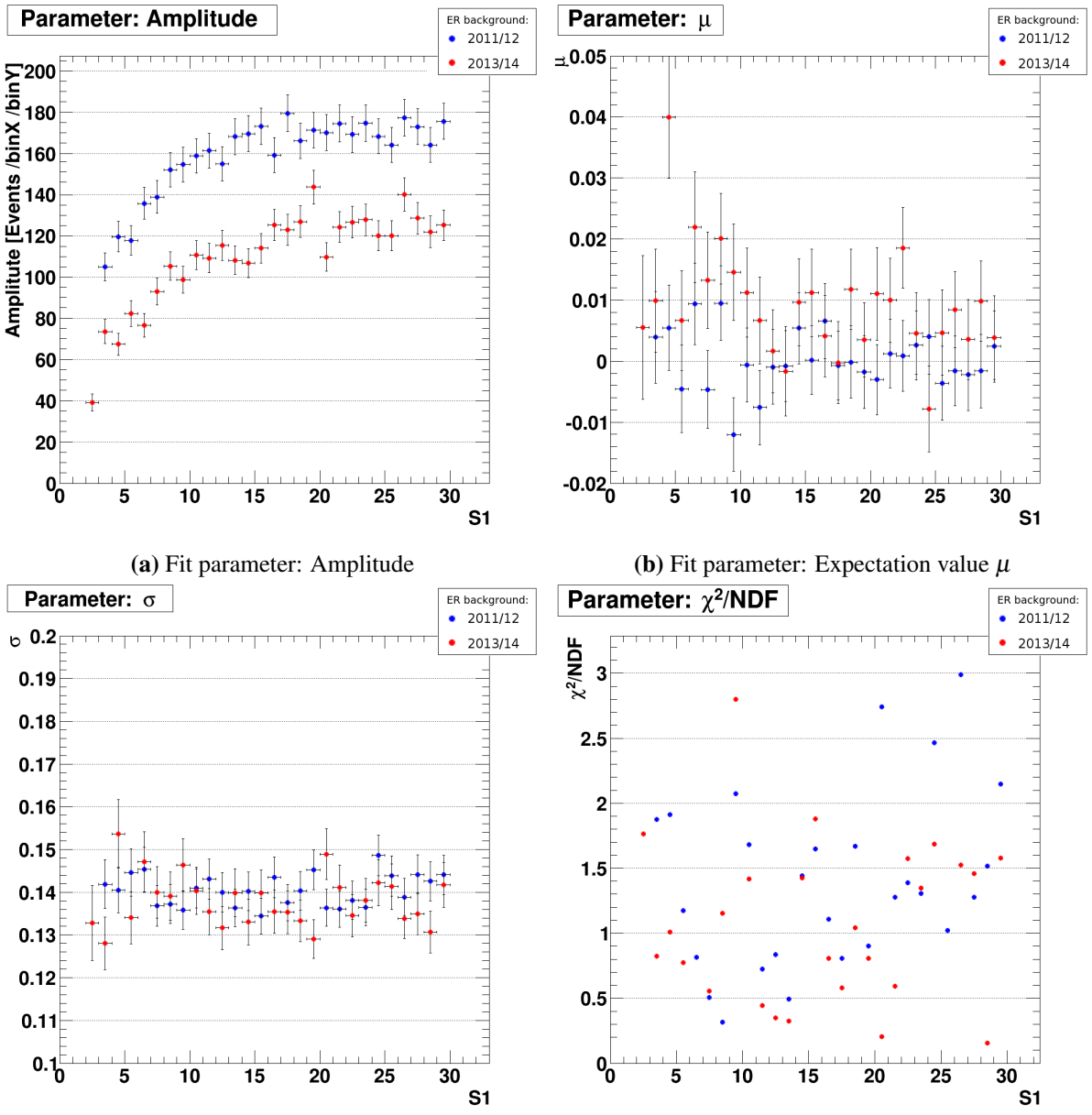
2. The first order approximation is symmetric due to the used Gaussian model.

After establishing the Gaussian ER background model in figure 6.6 for 2011/12 and 2013/14, the residual is calculated. Therefore the fit result is subtracted bin by bin from the histogram of figures 6.6a and 6.6b. Each histogram contains 30 bins (i) in $S1$ and 30 bins (j) in Y, in total 900 bins. The fit yields a continuous description of the ER background model. Therefore the residual for each bin $b_{i,j}$ is calculated by formula (6.1)

$$R_{i,j} = D_{i,j} - M_{i,j}^{\text{Gaussian}} \tag{6.1}$$

where $D_{i,j}$ describes the histogram of the data and $M_{i,j}^{\text{Gaussian}}$ the Gaussian fit. The resulting residual is shown in figure 6.7 for 2011/12 and 2013/14. To control the outcome of the residual, each slice in $S1$ is plotted in figure 6.4 additionally (lower figure). The single $S1$ slices in the appendix contain the residual (see figures A.1, A.2, A.3 and A.4 for 2011/12 data and A.5, A.6, A.7 and A.8 for 2013/14 data). The residual $R_{i,j}$ contains non-Gaussian distributed events after the subtraction. Not every $b_{i,j}$ in $R_{i,j}$ is anomalous leakage necessarily and therefore the non-Gaussian background contain anomalous leakage too.

The residual $R_{i,j}$ contains a full description of the anomalous leakage events which are outside a 3σ boundary around the expectation value μ_{ER}. The anomalous leakage events are not described by the first order approximation of the Gaussian ER background and their expected number of total events is according to table 6.4 is higher. For 10589 (2013/14) observed ER background events with Gaussian distribution, the number of expected leakage events outside 3σ is 55. There are 27 events more in the leakage region of the ER background data than expected (see table 6.4).
Besides a proper explanation of their origin, a useful phenomenological description is needed to model the observed leakage events. That description is called "leakage model".



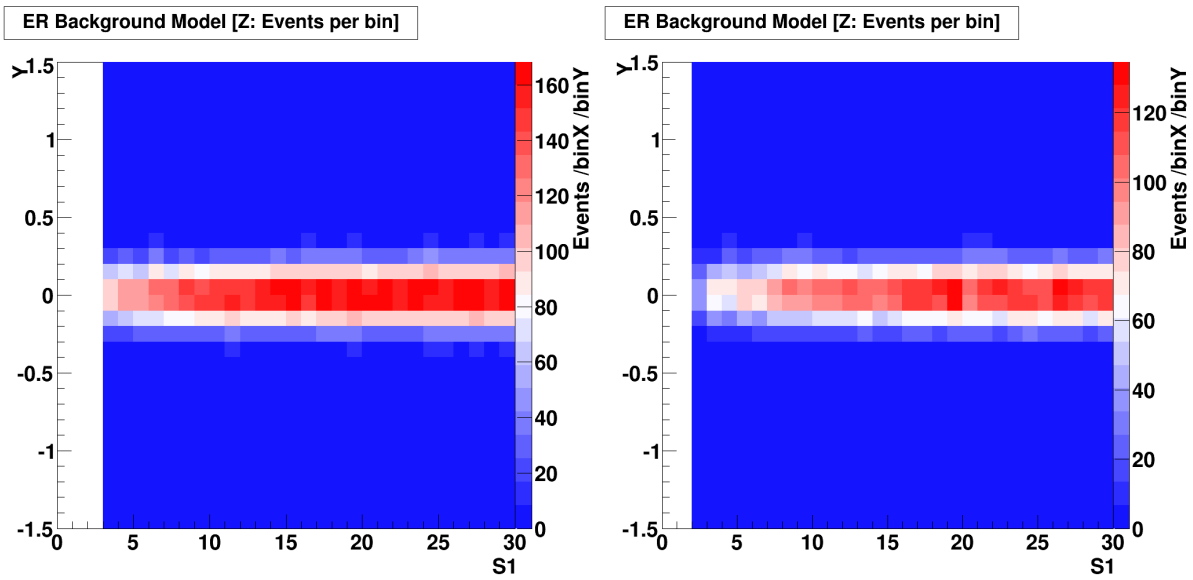
(a) Fit parameter: Amplitude

(b) Fit parameter: Expectation value μ

(c) Fit parameter: Standard deviation σ

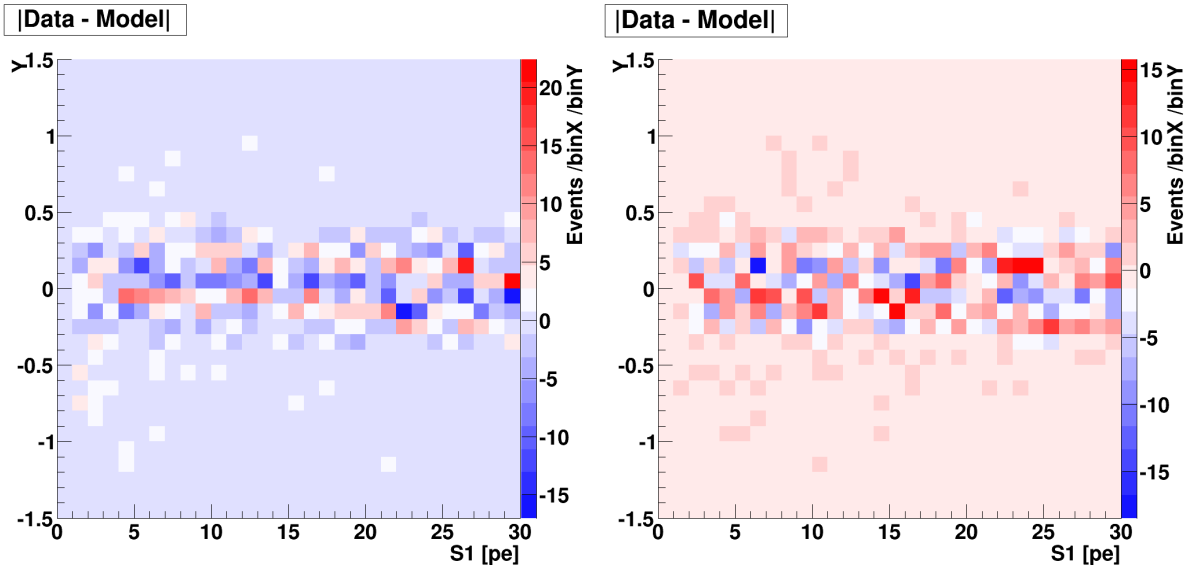
(d) Fit parameter: χ^2/NDF

Figure 6.5 – The fit parameter "amplitude", expectation value μ_{ER} and standard deviation σ_{ER} describe in a first order approximation the ER background. These parameters are fit results with their fit errors (error bars in figures 6.5a, 6.5b and 6.5c) and the reduced χ^2 (figure 6.5d). The amplitude refers to the given numbers of events inside each single slice of S1. The expectation value μ_{ER} varies around $Y = \sim 0$ in the flattened space variable Y and their standard deviation σ_{ER} describes the resolution of the experimental setup in the flattened space. From 2011/12 to 2013/14 the standard deviation does not change within its errors of each bin, its value is $\sigma_{ER} = \sim 0.14$. The reduced χ^2 varies within 40% in 2011/12 and 2013/14 around $\chi^2/\text{NDF} = \sim 1$.



(a) Gaussian ER background model of 2011/12. **(b)** Gaussian ER background model of 2013/14.

Figure 6.6 – These two figures present the ER background model from the Gaussian fits to the data. A more detailed view on single Gaussian distributions is given in section A.1.1 for 2011/12 and A.1.2 for 2013/14. This model describes the pure Gaussian contribution to the real ER background. The distribution follows the flattened band of $Y = \mu_{ER} = 0$. The leakage events are described in a separate model.



(a) Residual data of 2011/12. **(b)** Residual data of 2013/14.

Figure 6.7 – After establishing the ER background model from the single fits, the residual is calculated. The Gaussian background model is subtracted from the data. The residual contains the leakage events and it is the starting point to calculate a leakage model.

6.4.3. Leakage Model of the ER Background

The leakage model in 2011/12 and 2013/14 is developed from the residual of figure 6.7. The basic idea is to describe the distribution of the leftover events with a certain probability. This goal is achieved by a 2D Gaussian distribution (kernel ϕ, see equation (6.2)) which weigh each event in the residual and re-distributes it according to its event density correctly. A similar technique is used in digital photography ("blur effect") to remove hard edges in a picture by smearing the colour value of each pixel with their neighbours.

In a general view, the 2D Gaussian distribution is one possible kernel (ϕ). There is no physical motivation to choose a 2D Gaussian distribution. As long as the origin of the leakage events is not known the model for the anomalous leakage is created from the simplest assumptions. An important condition for ϕ is, that it is closed by definition, what means no singularities. Parameters σ_{S1} and σ_Y describe the width of the distribution in the $(S1, Y)$ data space. This assumption takes into account that each possible anomalous leakage event or non-Gaussian background event is smeared homogeneously in the $(S1, Y)$ data space with a given probability. The most important assumption for the anomalous leakage events is that their width σ_Y is the same such for Gaussian distributed events (σ_{ER}).
In detail, the residual $R_{i,j}$ contains anomalous leakage events and the non-Gaussian background. The latter background describes the leftover events after the Gaussian background model is subtracted from the data $D_{i,j}$ within $-3\sigma \leq Y \leq 3\sigma$. As long as the origin of the leakage events is not explained, it is not possible to distinguish anomalous leakage from non-Gaussian background (within $-3\sigma \leq Y \leq 3\sigma$). By construction, the non-Gaussian background event distribution dominates the residual between $-3\sigma \leq Y \leq 3\sigma$ and only a few leakage events stick out. Equation (6.2) is applied to each bin of the residual plot of figure 6.7 where the expectation values μ_{S1} and μ_Y define the position of each bin in the data space.

$$\phi_{i,j} = \phi(S1, Y) = \frac{A_{i,j}}{I\ 2\ \pi\ \sigma_{S1}\ \sigma_Y}\ \exp\left(-\frac{(S1-\mu_{S1})^2}{2\sigma_{S1}^2}\right)\ \exp\left(-\frac{(Y-\mu_Y)^2}{2\sigma_Y^2}\right) \tag{6.2}$$

$$I = \int_{2\text{ or }3\,\text{pe}}^{30\,\text{pe}} \int_{-1.5}^{1.5} \phi_{i,j}\,\mathrm{d}S1\,\mathrm{d}Y \tag{6.3}$$

$$K(S1, Y) = \sum_{i=0}^{i=30} \sum_{j=0}^{j=30} \phi_{i,j} \tag{6.4}$$

The standard deviation in $S1$ is chosen to $\sigma_{S1} = \sqrt{S1}$. This assumption takes the increasing $S1$ signal uncertainty into account. The deviation in Y is developed from standard deviation σ_{ER} of the Gaussian fits to the data of the ER background (see figure 6.5c). That assumption takes into account that both deviations σ_{ER} and σ_Y for the background and the leakage events are identical.

Applying the kernel ϕ to describe the leakage model corresponds to folding the leftover residual again with the detector resolution. This step reduces the precision in $(S1,Y)$. Nevertheless, the advantage of this method is that localised single leakage events are re-distributed according to a Gaussian probability. As long as the non-Gaussian and leakage events of figure 6.7 are re-arranged, the number of events changes a little due to numerical issues from the calculation of $K(S1,Y)$ and edge effects. Figure 6.8 shows an example for such a kernel with its parameters: $\mu_{S1} = 5\,\text{pe}$, $\sigma_{S1} = \sqrt{5}$, $\mu_Y = 0.5$ and $\sigma_Y = 0.14$ and $A(S1,Y) = 1$ event. The number of points in that distribution is chosen to the number of bins: 30 bins in $S1$ and 30 bins in Y.

These chosen parameters μ_{S1}, σ_{S1}, μ_Y and σ_Y for each bin $b_{i,j}$ result in 900 single 2D Gaussian distributions. To avoid boundary effects during the normalisation by equation (6.3) of the 2D Gaussian distribution, each kernel $\phi_{i,j}$ is normalised by I. This normalisation by the two dimensional integral with the boundaries in $S1 \in [3,30]$ pe (2011/12), $S1 \in [2,30]$ pe (2013/14) and $Y \in [-1.5,1.5]$ is important so that the final ER background model is consistent with the boundaries of the data space. This normalisation results in a truncated Gaussian distribution on each position in the $(S1/Y)$ data space. The truncated Gaussian distribution is multiplied by $A_{i,j}$ to weigh the distribution to the number of events for a given bin $b_{i,j}$. With this definition, each of the 900 Gaussian distributions have $^{\text{Events}}/_{\Delta S1 \Delta Y}$ on the z-axis.

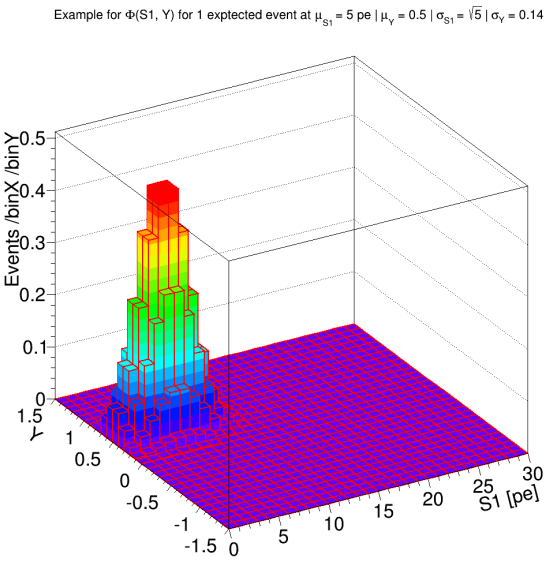


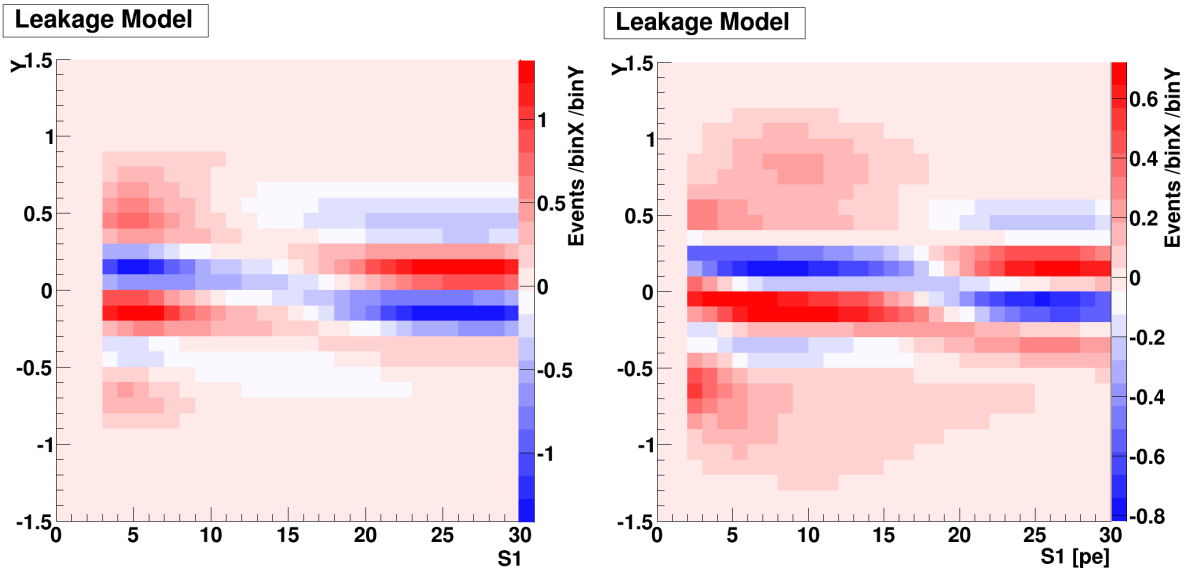
Figure 6.8 – Example of a single 2D Gaussian distribution with parameters: $\mu_{S1} = 5\,\text{pe}$, $\sigma_{S1} = \sqrt{5\,\text{pe}}$, $\mu_Y = 0.5$ and $\sigma_Y = 0.14$ and $A_{S1,Y} = 1$ event. There are 900 of these distributions added up to create a leakage model of the ER background in XENON100.

Figure 6.7 shows that the residual contains bins with positive and negative entries and therefore the amplitude $A_{i,j}$ determines the sign of the kernel $\phi_{i,j}$. The sum of these 900 kernels $\phi_{i,j}$ results in the leakage model ($K(S1,Y)$). Due to the normalisation of each kernel $\phi_{i,j}$ and the multiplication with the according amplitude $A_{i,j}$ the leakage model is given in events per $\Delta S1$ and ΔY. After folding the residual of figure 6.7 with kernel $\phi_{i,j}$, the result is given in figure 6.9. After the smearing there are negative and positive bin contents in the leakage model.

The reason for negative events in the leakage model comes from the fact that the residual is smeared without further constrains. Fluctuations can hence to be negative. Based on the residual plot of figures

6.7, there are negative entries in the histogram. These are expected from the subtraction. Therefore each given histogram entry ($A_{i,j}$) is positive or negative and hence the 2D Gaussian distribution ϕ has an "upward direction" ($+A_{S1,Y}$) or a "downward direction" ($-A_{S1,Y}$). Adding up positive and negative distributions leads to a certain probability of having negative events per bin after all.

A leakage model with negative numbers of events would be useless to describe the whole ER background. But as long as the leakage model is only a part of the ER background description the negative events are valid at this stage and become a feature after all. Reason for the leftover negative events is the assumption of having a first order approximation of the ER background by a Gaussian distribution. The ER background is not perfectly Gaussian shaped (see section 6.9). There is a small asymmetry in the flattened space Y which is responsible for the observed pattern in the leakage model.



(a) Leakage model of 2011/12.

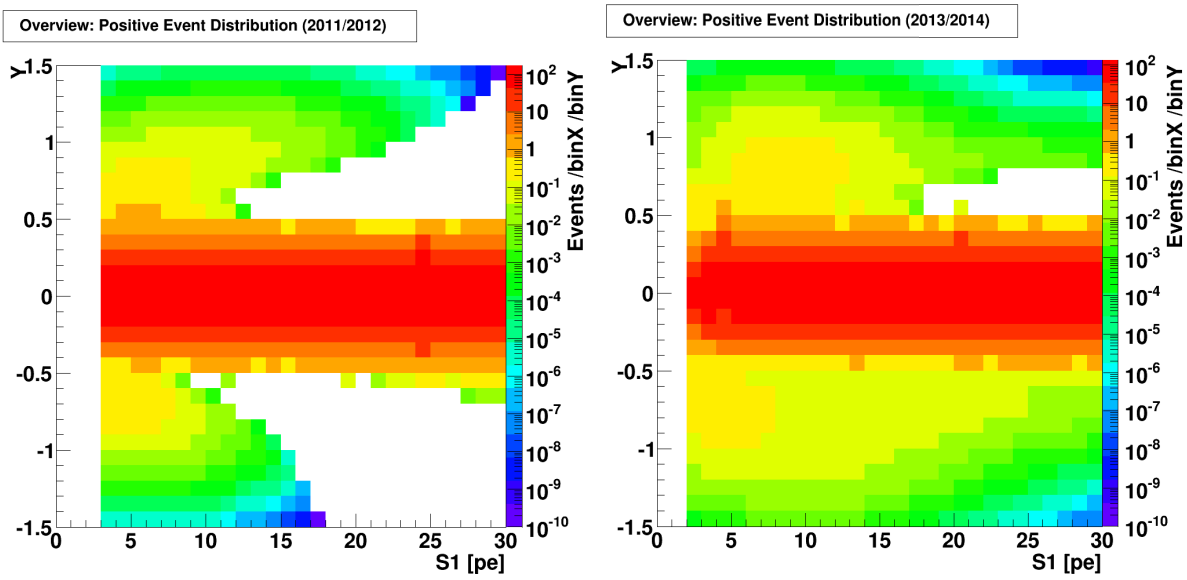
(b) Leakage model of 2013/14.

Figure 6.9 – Using a 2D Gaussian distribution ϕ to convolve the residual from figure 6.7. The result is a model of the leftover leakage events, the "leakage model". This leakage model contains also the small asymmetry in the region of the ER background: $-3\sigma \leq Y \leq 3\sigma$. A hint for the asymmetry is seen in the vertical change from positive to negative event distribution.

6.5. The Basic ER Background Model ("model-0")

The previous subsections describe the Gaussian approximation and the 2D Gaussian smearing. Based on the suitable normalisation during the smearing process, the total number of the Gaussian and the

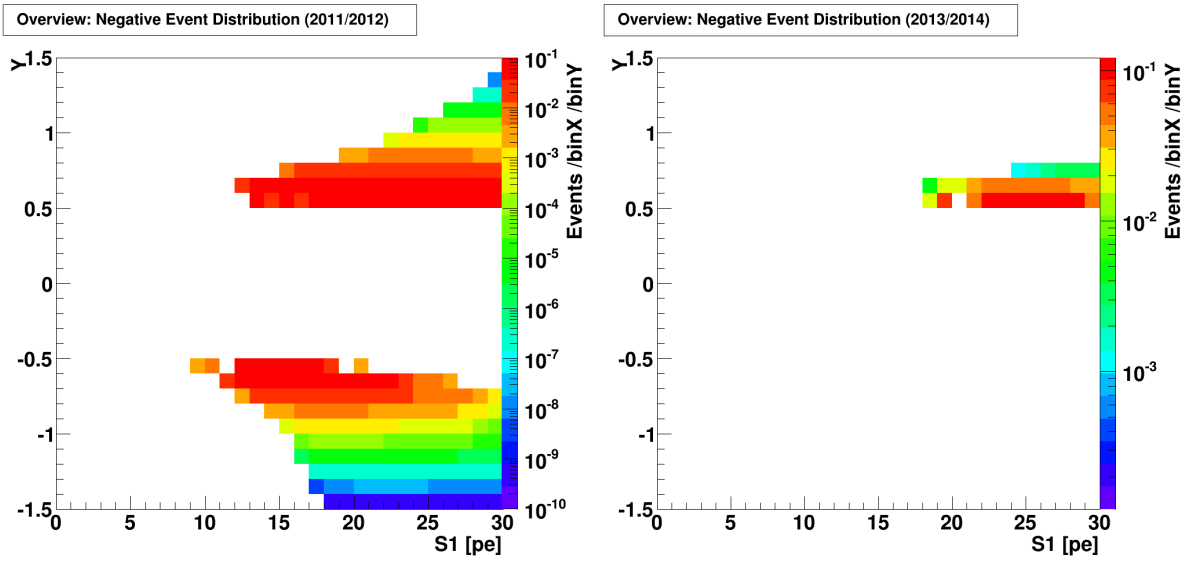
Leakage ER background model events are conserved. The sum of both models (Gaussian and leakage model) is calculated bin by bin. The results are given in figure 6.10 for data taking periods 2011/12 and 2013/14. This model is named in the following as "model-0".
The region of the anomalous leakage is described "by eye" in both data taking periods. Unfortunately, the total event number is not conserved at a satisfying level. The disadvantage of model-0 is the *negative event distribution* which appears in some regions of the data space. As consequence, model-0 predicts a negative number of events which is not in agreement with a physical measurement.
Comparing figures 6.1 and 6.10, it is obvious that the regions where the negative event distributions dominate model-0 are those where no anomalous leakage events are observed. Due to the logarithmic z-axis in figure 6.10 it is not possible to show the negative event distribution in detail. This is done in figure 6.11 where the negative event distribution is multiplied by -1 for presentation. Without further discussion, these areas of the model are inaccessible in a continuative data analysis procedure, for example the Profile Likelihood approach [19]. A proper ER background model have to fulfil that all predicted events are positive or at least zero. A more stronger constrain is to accepted only positive event distributions. This is a fair assumption because after an infinite measurement with XENON100 it is expected to observe events in the whole data space, distributed according to a certain event distribution.
Based on the normalisation during the 2D Gaussian smearing process, the number of total events is conserved. This statement is true for model-0 if all bins are summed up. The total number of events is restored. This is summarised in table 6.5. The table summarises the positive and negative event distribution separately. Finally, it is important to determine the number of leakage events in the model-0 ER background compared to the observed number of events in the data. The data in 2011/12 shows $n_{LE}^{-3\sigma} = 23$ events below -3σ and $n_{LE}^{+3\sigma} = 25$ events above $+3\sigma$. These numbers are compared to model-0 by $n_{LE-model}^{-3\sigma} = 21$ events and $n_{LE-model}^{+3\sigma} = 25$ events. The numbers for 2013/14 are also given in table 6.5. The total number of leakage events are comparable in both data taking periods. By construction the 2D Gaussian smoothing conserves the number of events in the ER background model which indicates the robustness of the used method: The total number of events are re-distributed. The number of anomalous leakage and non-Gaussian events in the residual are also restored.
In summary, the negative event distributions come from a slight overestimation of the Gaussian ER background model and the fact that no or few anomalous leakage events are observed in areas where the final ER background model has a negative event distribution.



(a) Leakage model of 2011/12.

(b) Leakage model of 2013/14.

Figure 6.10 – The model-0 approach for 2011/12 and 2013/14. After summing up the Gaussian ER background model with the leakage model, regions of negative event distribution are left. Due to the presentation with a logarithmic scale at the z-axis, a detailed view on the negative event distribution is given in figure x for completeness.



(a) Leakage model of 2011/12.

(b) Leakage model of 2013/14.

Figure 6.11 – Both figures give an overview about the negative event distribution in 2011/12 and 2013/14 and are the complementary to figure 6.10 (white areas). The negative event distribution is multiplied by -1 to result a positive event distribution for this presentation.

	Final ER background model (total)	Final ER background model (positive)	Final ER background model (negative)	Anomalous Leakage ($Y < -3\sigma$)	Anomalous Leakage ($Y > -3\sigma$)	Total leakage events
2011/12	15128.01	15133.64	-5.64	21	25	46
2013/14	10589.71	10591.10	-1.39	22	27	49

Table 6.5 – Overview about the distribution of the ER background model in 2011/12 and 2013/14. Due to the overestimation of the Gaussian ER background model at some regions in the (S1,Y) data space become the leakage model negative. Negative event counts are non-physical and therefore these bins would be set to zero in the final model. But from studies of systematic uncertainties it is determined if the leakage model contributes to such areas or if the Gaussian ER background is dominant.

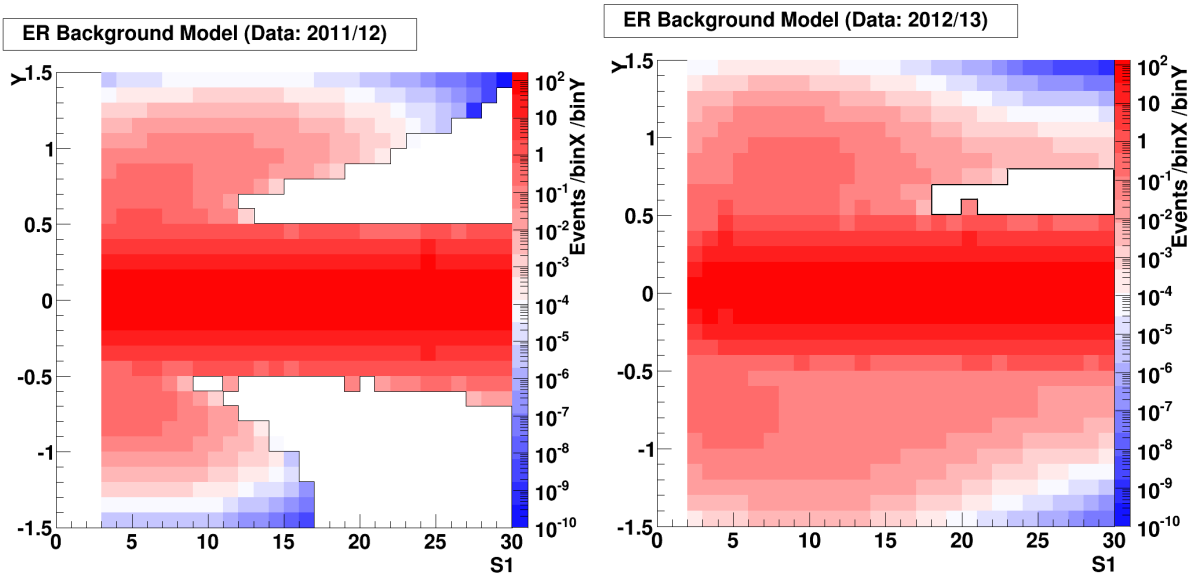
6.6. The Combined ER Background Model (”model-A“)

The combined ER background is the continuation of section 6.5 where model-0 is discussed. Due to the negative event distribution, model-0 does not fulfil the necessary requirements to be used in a PL analysis. The easiest approach is to set the negative event distribution to zero and re-normalise the outcome due to the total number of observed events. This model is named as ”model-A“ in the following. The assumption to set the negative event distribution to zero matches the observation: The negative event distribution appears only out of the $Y > \pm 3\sigma$ area. At these positions are only a few or zero anomalous leakage events observed. From this point of view, model-A fits the data at a certain level. Figure 6.12 shows model-A for 2011/12 and 2013/14.

The negative event distributions ($-A_{S1,Y}$) results from the leakage model. By adding the Gaussian background model, the individual bin i, j does not exceed the zero. This effect becomes dominant for regions of $Y < -3\sigma$ and $Y > 3\sigma$, here the Gaussian tails are already so small that the event distribution is not enough to overcome the negative event distributions from the leakage model. The reason why there is such a difference in the event distribution of figures 6.12a and 6.12b is the event distribution of the anomalous leakage events for the data taking periods in 2011/12 and 2013/14. Compare figures 6.1a and 6.1b, in 2013/14 are more events observed for $S1 > 15\,\mathrm{pe}$. These anomalous leakage events and their remaining probability after applying ϕ to the certain bins lead to the increased number of events in the model-A background model.

Model-A fulfil the minimum requirement to be a valid background model. Furthermore, the areas around the observed anomalous leakage events are restored and described by model-A with a single expectation. During 2013/14, an anomalous leakage event is observed where model-A predicts *exact* zero events. Compare the white area of figure 6.12b with observed data (6.1b). This is a weak point of model-A. This individual bin predicts zero events where obviously one event is observed and hence this bin is not valid described by the model. A proper model should at least predict a positive event

distribution, in particular at an individual bin where data is observed. Nevertheless, model-A predicts events for low S1 energies down to the S1 threshold which is region of interest for e.g. low WIMP mass search.



(a) Final background model of 2011/12.

(b) Final background model of 2013/14.

Figure 6.12 – Model-A is the sum of the Gaussian ER background model and the leakage model. Both are added up and bins which contain negative events are set to zero to avoid inconsistencies (white areas within black lines as boundaries). The resulting model contains a full description of the pure ER, the anomalous leakage and the asymmetry contribution.

6.7. Goodness of the Background Model

This section determine the goodness of the model by a bin by bin Log-Likelihood Ratio (LLR) analysis. The method is explained in detail in section 6.7.1 and is then applied to test the goodness.

6.7.1. Goodness of the Model with Log-Likelihood Ratio

This subsection introduces a possibility to determine the goodness of the model by calculating the log-likelihood ratio. On a regular basis which introduces a way to evaluate the *global goodness* of the model, a first idea on the goodness can be estimated. Furthermore it becomes necessary to determine the goodness of the model on a bin by bin basis to access smaller areas in the data space to quantify

how well a certain model describes the data. The following procedure uses the term M (or $M_{i,j}$) for the model which is tested with this method in general. A specific model is addressed separately. The data are shorten by D (or $D_{i,j}$).
Starting with the basic definition of the log-likelihood ratio in equation (6.5), the single components $Pois(D_{i,j}|M_{i,j})$ and $Pois(D_{i,j}|D_{i,j})$ are defined according to equation (6.6) and (6.7) for the underlying Poisson statistic in a counting experiment. The upper end of the sum in equation (6.5) is specified by N and is related to the largest bin number in $S1$ and Y and is 30 bins each according to the histogram definition of the ER background data.

$$LLR(D|M) = \sum_{i,j=1}^{N} -2 \times \log\left(\frac{Pois(D_{i,j}|M_{i,j}}{Pois(D_{i,j}|D_{i,j})}\right) \tag{6.5}$$

$$Pois(M_{i,j}|D_{i,j}) = \frac{\exp(-M_{i,j}) \times M_{i,j}^{D_{i,j}}}{D_{i,j}!} \tag{6.6}$$

$$Pois(D_{i,j}|D_{i,j}) = \frac{\exp(-D_{i,j}) \times M_{i,j}^{D_{i,j}}}{D_{i,j}!} \tag{6.7}$$

By putting equation (6.6) and (6.7) into (6.5), the LLR is calculated according to:

$$LLR(D|M) = \sum_{i,j=1}^{N} 2 \times (D_{i,j}\ \log(\frac{D_{i,j}}{M_{i,j}}) - D_{i,j} + M_{i,j}) = \sum_{i,j=1}^{N} LLR(D|M)_{i,j} \tag{6.8}$$

The sum in equation (6.8) is shortened by

$$LLR(D|M)_{i,j} = 2 \times (D_{i,j}\ \log(\frac{D_{i,j}}{M_{i,j}}) - D_{i,j} + M_{i,j}) \tag{6.9}$$

This definition takes the number of observed events $D_{i,j}$ in the data and compare it to the expected events $M_{i,j}$ in the model. Therefore the model $M_{i,j}$ is tested against the data $D_{i,j}$. A total agreement results $LLR(D|M)_{i,j} = 0$ (and hence $LLR(D|M) = 0$). The log-likelihood ratio test is positive defined and therefore the large deviation of the model from the data results $LLR(D|M)_{i,j} = +\infty$. The summed log-likelihood ratio $LLR(D|M)$ describes the global goodness (based on the individual goodness $LLR(D|M)_{i,j}$ which is indicated by the addition of the bin index). The $LLR(D|M)_{i,j}$ are understood as a detailed goodness of the model for a certain bin. This assumption is valid because the log-likelihood ratio must not necessarily calculated on the whole data space and therefore the sum in equation (6.8) is reduces to a single addend to calculate $LLR(D|M)_{i,j}$.
To implement equation (6.9) into the data analysis, four different use-cases are possible:

1. $D_{i,j} > 0$ and $M_{i,j} > 0$: This is the usual situation in where the Gaussian ER background model dominates. Hence the values for $LLR(D|M)_{i,j}$ are calculated according to equation (6.9).

2. $D_{i,j} = 0$ and $M_{i,j} = 0$: Due to the requirement that a proper model predicts at every a bin certain number of events, this use-case should not be necessary. But the discussed model-A in section 6.6 needs to take this possibility into account. Equation (6.9) shows that the smallest possible values is $LLR(D|M)_{i,j} = 0$ in case that the predicted number events is equal to the observed number of events. This is also true for zero observed and predicted events. Hence a better match can not be archived.

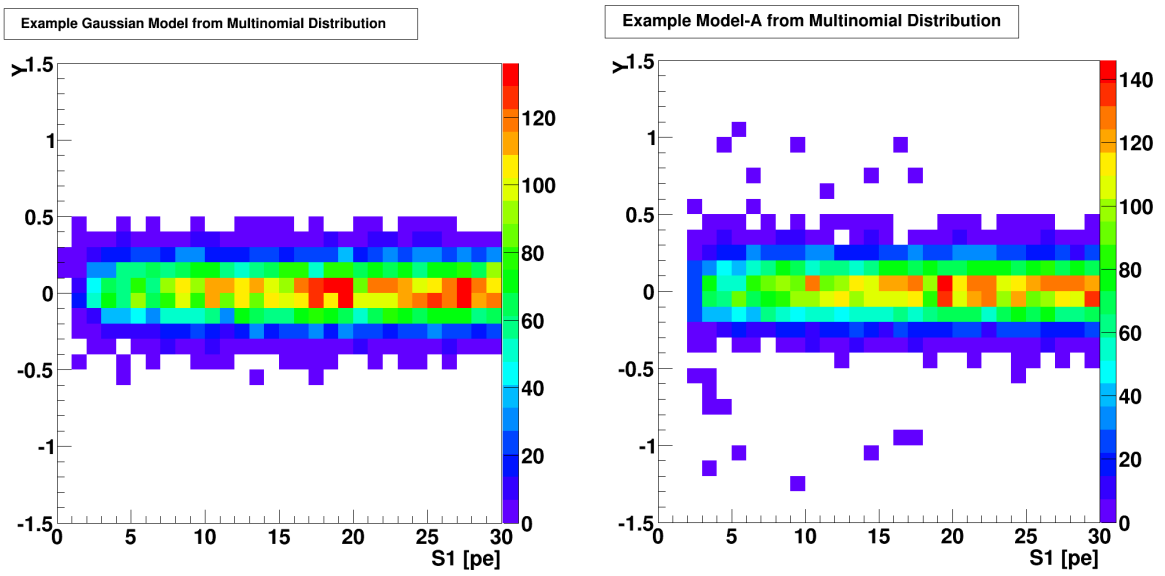
3. $D_{i,j} = 0$ and $M_{i,j} > 0$: This use-case allows to form equation (6.6) and (6.7) to $LLR(D|M)_{i,j} = 2 \times M_{i,j}$. In case that zero events are observed, the log-likelihood ratio only depends on the model, there are no further information (e.g. correlation) available.

4. $D_{i,j} > 0$ and $M_{i,j} = 0$: Calculating this use case with (6.6) and (6.7) results an extreme value of $LLR(D|M)_{i,j} = +\infty$. This situation is observed for the data taking period 2013/14 for a single bin at $S1 = 25\,\text{pe}$ and $Y = \sim 0.7$. In case of testing model-A, this bin information is removed from the summed log-likelihood ratio $LLR(D|M)$ and also not accessible in the later individual bin studies.

Based on this definition it is possible to test two already discussed models, the Gaussian ER background model and model-A and determine the goodness of the model each. Obviously, the model-0 can not be test by a log-likelihood ratio test.

Based on the resulting model-A and the Gaussian ER background model, it is only possible to calculate two global values of the LLR and one set of $LLR(D|M)_{i,j}$ for the individual bins for each model (see section 6.7.2). This is not meaningful and a more complex test is carried out with a Monte Carlo technique. This allows to use the log-likelihood ratio as test variable for the goodness of the model and single bins.
The model is used to create another simulated event distribution (S or $S_{i,j}$), similar to the input data. To fulfil the necessarily boundaries, a multinomial statistic [47] is assumed during the sampling process. This statistic guaranties to conserve the number of simulated events. The model is available as a 2D histogram and therefore each bin is threaten as a probability observe an event at this position in case the simulated event distributions are normalised.
The multinomial sampling is performed by the python package *numpy* [91]. Two possible results are shown in figure 6.13 for the Gaussian ER background model and model-A from the data taking period 2013/14. The simulated event distribution from the Gaussian ER background does not contain any anomalous leakage events whereas the model-A results some anomalous leakage events. The

results for the data taking period 2011/12 are similar. This allows to compute a huge set of simulated



(a) From Gaussian ER background.

(b) From Model-A

Figure 6.13 – Example for two multinomial distributed random draws from the Gaussian ER background (6.13a) and model-A (6.13b). Based on the individual probabilities, a simulated event distribution is created from the models by using the *numpy* package in python.

event distributions and calculate their according log-likelihood ratio values and determine individual distributions of the log-likelihood ratios for each bin. Hence the log-likelihood ratio itself become a test statistic to evaluate the goodness of the model. The following subsections discuss the global and individual log-likelihood ratio distributions and test these with data. As a large sample, there are 10000 simulated event distributions created. To be conform with the previous notation of the log-likelihood ratio calculation rule, the simulated distributions are named S and for example the log-likelihood ratio of a single bin would be named as $LLR(S|M = Gaussian)_{i,j}$ if the simulated event distribution is created from the Gaussian ER background model.

6.7.2. Global Goodness of the Model

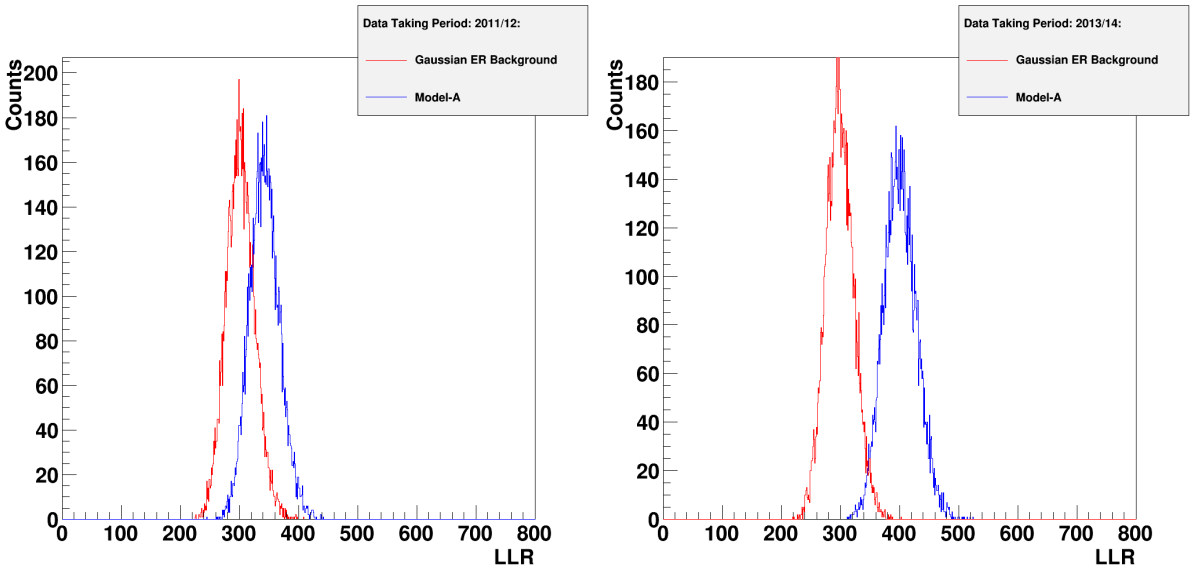
To compare the Gaussian ER background model and model-A with the data the global log-likelihood ratio is calculated such it is defined in equation (6.8). These values are listed in table 6.6 for 2011/12 and 2013/14 (first row). Just by comparing both results without further information it is obvious that model-A matches the data better than the Gaussian ER background model. This is expected because

the anomalous leakage events are not described by the Gaussian ER background. The result is an increase in $LLR(D|M)$ caused by these few bins where anomalous leakage events are observed.

	2011/12		2013/14	
	Gaussian background model	Model-A	Gaussian background model	Model-A
$LLR(D\|M)$	509	286.7	778	329
$<LLR(S\|M)>$	302	342.5	299.2	401
RMS($LLR(S\|M)$)	24.8	25.3	23.6	27.9

Table 6.6 – Overview about both data taking periods and the related calculated log-likelihood ratios where the data are compared to both models. Furthermore the mean and root mean square values are extracted from figure 6.14 to specify the log-likelihood ratio calculations with the 10000 simulated event distributions from both models.

To compare these values with the simulated event distributions, the according quantities $LLR(S|M = Gaussian)$ and $LLR(S|M = Model - A)$ are calculated such it is described in section 6.7.1. The result is given in figure 6.14. It is important to mention that the according log-likelihood ratio test, with e.g. the model-A, is carried out with simulated events from the model-A distribution. The mean and root mean square values are also listed in table 6.6 and extracted from the histograms (figure 6.14). The result of the simulated event distributions is counter-intuitive compared to the results from measured event event distributions: The simulated event distributions from the Gaussian ER background model matches better the Gaussian ER background model ($LLR(S|M = Gaussian)$) better than the event distributions from model-A which is tested with model-A $LLR(S|M = Model - A)$. The reason for this is the dominant Gaussian part of the ER background which is similar in both models. The worse matching of model-A results from the simulated anomalous leakage events. These events are also multinomial distributed and therefore a simulated event distribution from model-A starts to differ in cases of single anomalous leakage events which are somewhere else in the data space by a given probability. From these simulated anomalous leakage events, the $LLR(S|M = Model - A)$ increases by only a few bins i, j in equation (6.8). From the 10000 simulated event distributions for model-A, the anomalous leakage varies and hence the $LLR(S|M = Model - A)$ increases. This compares to the Gaussian part of model-A which also varies but its variation has less influence on the $LLR(S|M = Gaussian)$. This result shows a weak point of the idea to characterise the goodness of the model by a global description. The log-likelihood ratio can be used to compare two models to each other directly as it is done in table 6.6 by the first line but a further approach with multinomial distributed event distributions leads to counter-intuitive result. Every random distributed anomalous leakage event from the model-A distribution increases $LLR(S|M = Model - A)$ when this event does not fit the model-A distribution exactly. Hence the observed result of the simulated model-A distributions is inter-



(a) 2011/12 **(b)** 2013/14

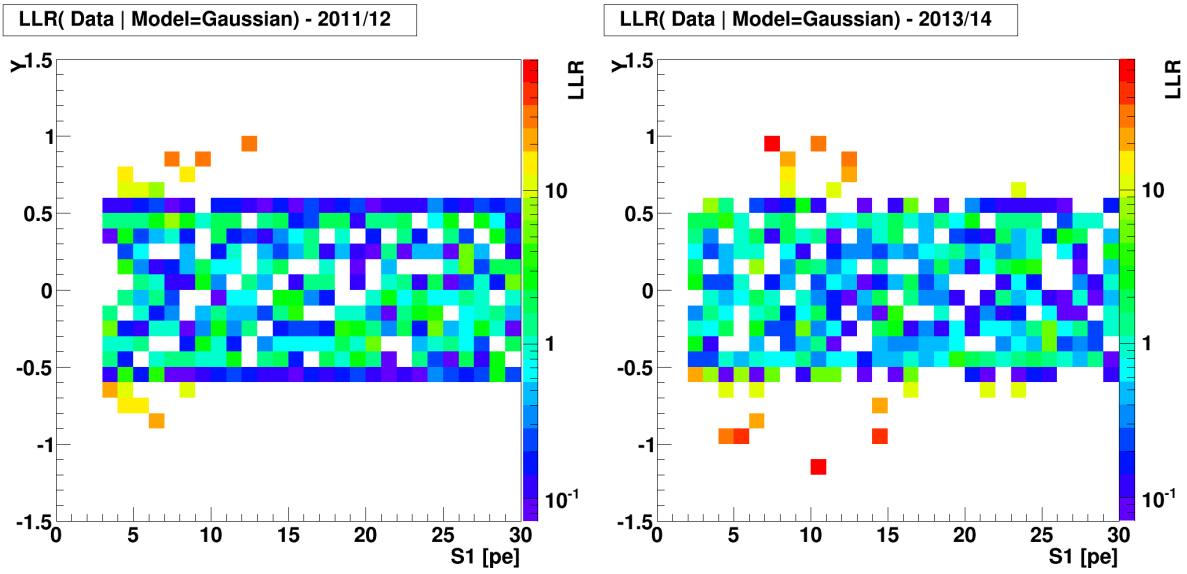
Figure 6.14 – An overview for the data taking periods in 2011/12 and 2013/14 about the log-likelihood ratio tests with simulated event distributions. Therefore 10000 multinomial distributed simulated event distributions are created from the Gaussian ER background and model-A. This information is used to calculate the *global* LLR (see equation (6.8)). It turns out that the simulated Gaussian ER background matches the Gaussian background model better (red histograms) compared to the match of simulated model-A distributions with the model-A.

preted that the smallest possible LLR varies around the mean value of the model-A distribution (blue histograms). Finally, by comparing the outcome of the simulated event distributions with the data ($LLR(S|M) \leftrightarrow LLR(D|M)$), the Gaussian ER background model fits the data worse than model-A. This can be seen by calculating a symmetric p-value around the mean value based on the log- likelihood ratio distributions in figure 6.14 for the red histograms with the according values $LLR(D|M = Gaussian)$ and $LLR(D|M = Model-A)$ from table 6.6. The p-values are in both cases 0%. The situation becomes better when data is compared to model-A. The p-values are $LLR(D|M = Model-A)^{2013/14} = \sim 0.03\%$ and $LLR(D|M = Model-A)^{2011/12} = \sim 1.03\%$. Due to the reason that both $LLR(D|M = Model-A)$ values from table 6.6 are smaller than the mean values of the simulated event distributions leads to the assumption that the model-A over-fits the data, in particular. Comparing the Gaussian ER background ($LLR(S|M = Gaussian)$, red histograms) with $LLR(D|M = Gaussian)$ also leads to the assumption that this model over-fits the data. At this point, the Gaussian approximation over-fits the data already.

6.7.3. Detailed Goodness of the Model

According to the definition of the log-likelihood ratio test in section 6.7.1, the global goodness of the model test is tuned to a more meaningful description on a level of single bins.
The log-likelihood ratio test is carried according to formula (6.9) where a single value of $LLR(D|M)_{i,j}$ is calculated. Based on the chosen number of 900 single bins, the result for 2011/12 and 2013/14 is shown in figure 6.15 for a test of $LLR(D|M = Gaussian)_{i,j}$. Figure 6.16 shows the test with model-A ($LLR(D|M = Model-A)_{i,j}$). The log-likelihood ratio values are presented by the colour code.
Both log-likelihood ratio tests of $LLR(D|M = Gaussian)_{i,j}$ and $LLR(D|M = Model-A)_{i,j}$ show low values of $LLR(D|M)_{i,j}$ in the region of the Gaussian part ($Y < \pm 3\sigma$). This is expected because the first approximation describes the observed data pretty well and the leakage model adjust the Gaussian part only a little. Hence there is almost no different for the choice of the model within the 3σ region. The region of interest is outside the 3σ region where $LLR(D|M)_{i,j}$ is calculated for the anomalous leakage events. Due to the reason, that the Gaussian ER background model does not describe the anomalous leakage event, the large $LLR(D|M = Gaussian)_{i,j}$ values in figure 6.15 are explained. These bins contain anomalous leakage events in the data space. The situation changes for model-A where the anomalous leakage events are described by the additional leakage model. The single bins where anomalous leakage events are observed in the data are distinguishable from their neighbours because of the slightly larger log-likelihood ratio. Nevertheless, the leakage model results a better agreement with the data compared to the Gaussian ER background model. The log-likelihood ratio values are presented by the colour code in figure 6.16 and the lowest value is limited to $LLR_{i,j} \geqq 0.001$. Below this value are log-likelihood ratio values calculated but not shown in this figures. Only exception is

the already discussed *empty* space where the model results zero predicted events (see section 6.6), this regions does not provide any log-likelihood ratio information.

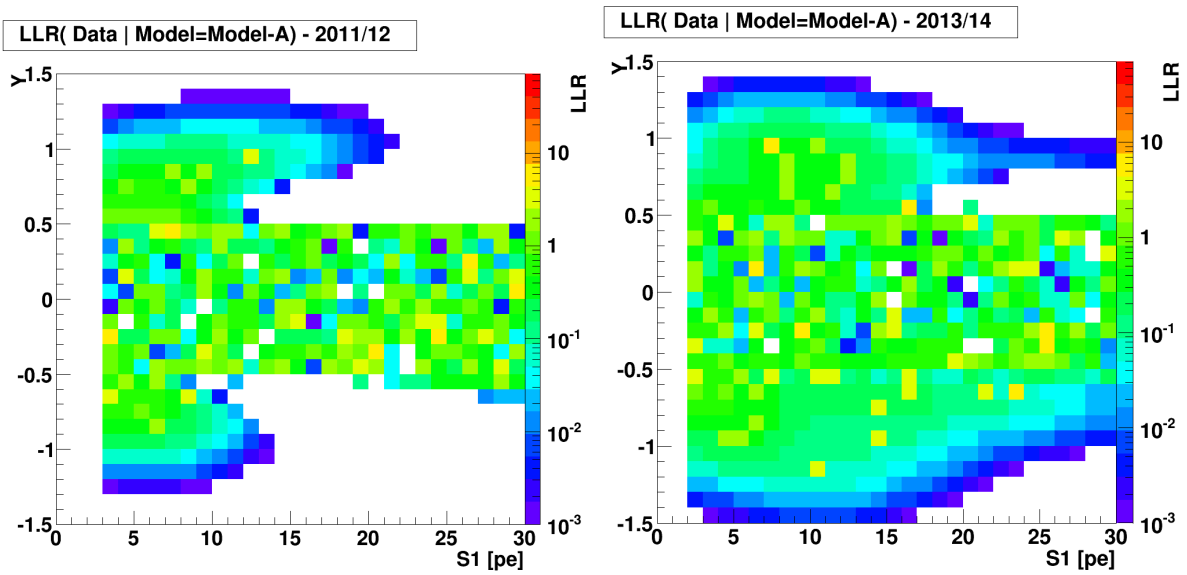


(a) 2011/12 **(b)** 2013/14

Figure 6.15 – The log-likelihood ratio values of single bins for the test: $LLR(D|M = Gaussian)_{i,j}$ for the data taking periods 2011/12 and 2013/14. It shows that the anomalous leakage events are worse described by the Gaussian background model. The Gaussian approximation works pretty well, within the 3σ region are log-likelihood ratio values observed within $\sim 1 \leq LLR_{i,j} \leq \sim 10$.

This result shows that model-A describes the data better in the regions of the anomalous leakage events. Based on a large number of simulated event distributions, a log-likelihood ratio distribution is created for each bin where again the two tests are carried out: $LLR(S|M = Gaussian)_{i,j}$ and $LLR(S|M = Model - A)_{i,j}$. There are two examples shown in figure 6.17 for different bins and tested models.

The $LLR(D|M)_{i,j}$ distributions have a bin size of $\Delta LLR(D|M)_{i,j} = 0.001$. The choice of the bin width is made according to the computing time. Further steps need to calculate different quantiles based on the input of each single log-likelihood distribution. Hence, a finer binning increases the computing time. This is a boundary condition for the log-likelihood interpretation. There are individual bins observed where the log-likelihood ratio test with simulated event distributions from a Gaussian ER background model or model-A results $LLR(D|M)_{i,j} < 0.001$ at any time. These bins predict a good agreement of the model with the data and based on the boundary condition, the model agrees with the data up to an log-likelihood ratio test of $LLR(D|M)_{i,j} < 0.001$. There is a special situation in where this boundary condition of the log-likelihood ratio is translated into predicted events from the model directly: If no data are observed in individual bins and the model predicts a certain number of



(a) 2011/12 **(b)** 2013/14

Figure 6.16 – The log-likelihood ratio values of single bins for the test: $LLR(Data|M = Model - A)_{i,j}$ for the data taking periods 2011/12 and 2013/14. This model describes the observed anomalous leakage events with a better agreement compared to the Gaussian ER background model of figure 6.15. The Gaussian region ($Y < \pm 3\sigma$) results an almost similar description to the Gaussian ER background. The leakage model adjust the Gaussian ER background only a little in this region.

events (see case three of the enumeration in section 6.7.1), the number of predicted events become $M_{i,j} < 0.0005$ at minimum for $LLR(D|M)_{i,j} < 0.001$. If less events are predicted, the model agrees with the data automatically.
The two given examples (figure 6.17) for $LLR(D|M)_{i,j}$ distributions are taken out of 900 distributions in total for the data taking period 2013/14 and for simulated event distributions from model-A ($LLR(S|M = Model - A)_{i,j}$). Both show a wide range of log-likelihood ratios. Other bins such as figure 6.17b have a maximum up to $LLR(S|M)_{i,j} \sim 100$. Log-Likelihood distributions which are derived from the regions where anomalous leakage is observed are looking like figure 6.17a where for almost all log-likelihood ratio a good agreement of model and data is archived. These log-likelihood distributions allow to define a confidence level which is demanded to verify how well a model fits the observed data and at individual bins.

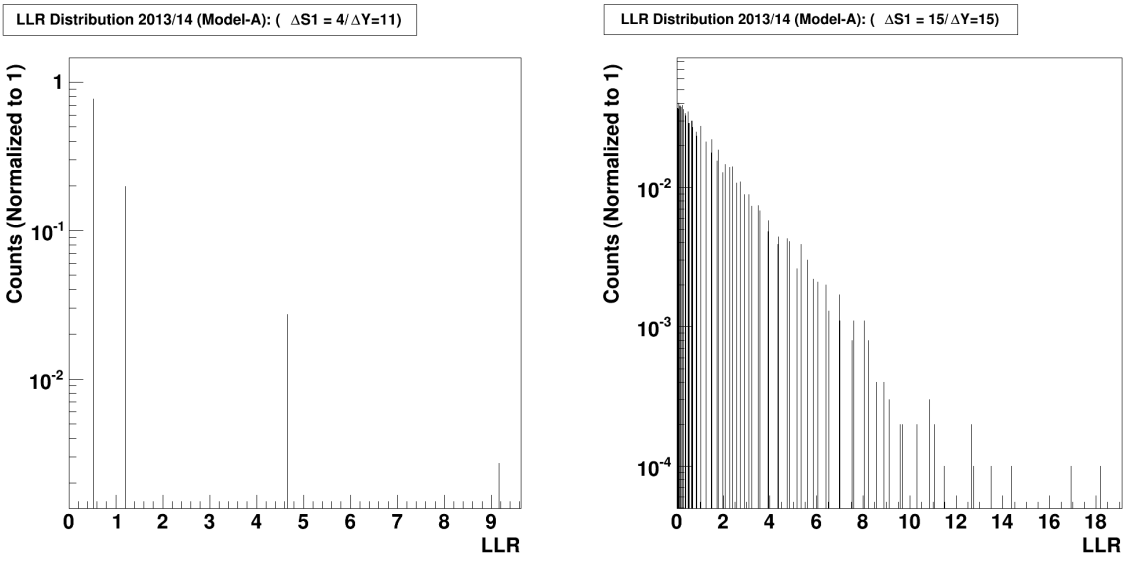
6.7.4. Detailed Goodness of the Model Based on a Confidence Level

With respect to a detailed analysis of the previous section 6.7.3, a criterion is introduced to define the goodness of the model for each single bin. A confidence level of 5% is demanded for the log-likelihood ratio distributions to decide if a model is accepted or rejected. Hence a 95% quantile is calculated for each log-likelihood distribution (two examples given in figure 6.17).
For each quantile a $LLR(S|M = Gaussian)_{i,j}^{95\%}$ value exists to quantify the Gaussian ER background model. These values are shown in figure 6.18 for 2011/12 and 2013/14, where the colour code describes the log-likelihood ratios which fulfil the 5% criterion. For both data taking periods, the result is pretty similar to figure 6.15 (mind the different z-axis).
The same 5% confidence level is also demanded for $LLR(S|M = Model - A)_{i,j}^{95\%}$ to verify the acceptance of single bins for model-A. The result is given in figure 6.19 and is again similar to figure 6.16. The 5% confidence limit is presented by the hard edge in figure 6.19 which does not appear in figure 6.16.

The decision to accept only events which fulfil the 5% C.L. is given in equation (6.10).

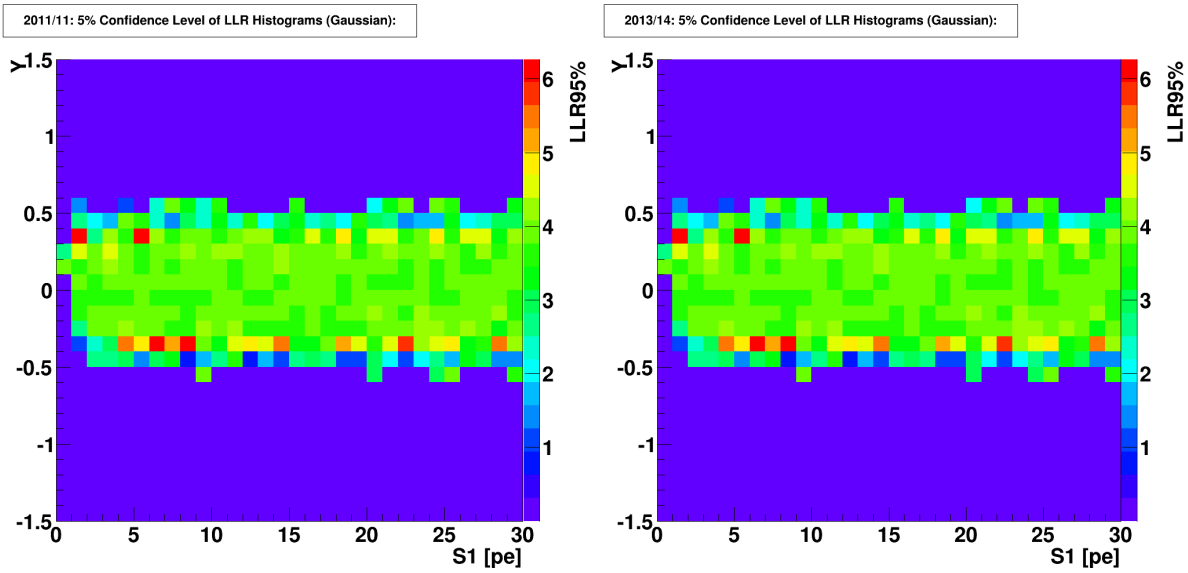
$$LLR(D|M)_{i,j} < LLR(S|M)_{i,j}^{5\%} \tag{6.10}$$

With this confidence level each ER background model is evaluated. The result is given in figure 6.20 for the Gaussian ER background model in data taking periods 2011/12 and 2013/14. It contains the Gaussian background model which is accepted on a 5% C.L. with the corresponding event distributions which is presented by the colour code. The white bins in between indicate single regions where



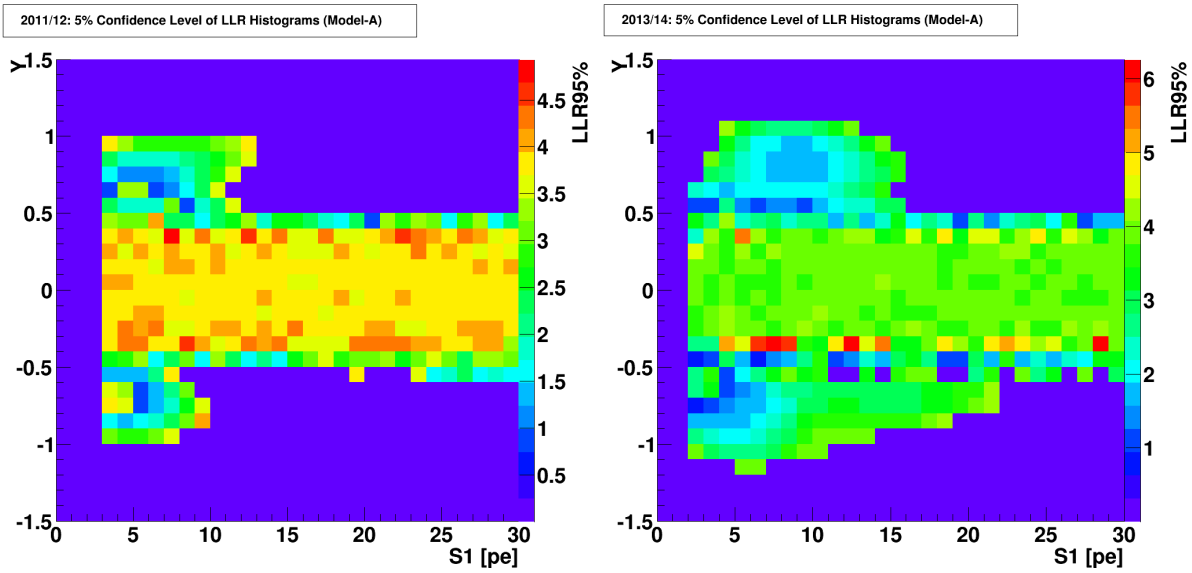
(a) 2011/12 **(b)** 2013/14

Figure 6.17 – These two log-likelihood ratio distributions are created from 10000 simulated event distributions where the $LLR(S|M = Model - A)_{i,j}$ are calculated. The here presented examples show a bin inside (figure 6.17b) and outside (figure 6.17a) the $Y < \pm 3\sigma$ region.



(a) 2011/12 **(b)** 2013/14

Figure 6.18 – Both figures show a 95% quantile (5% C.L.) at which the Gaussian ER background model for both data taking periods results which log-likelihood ratio value based on 10000 simulated event distributions for the Gaussian ER background model.

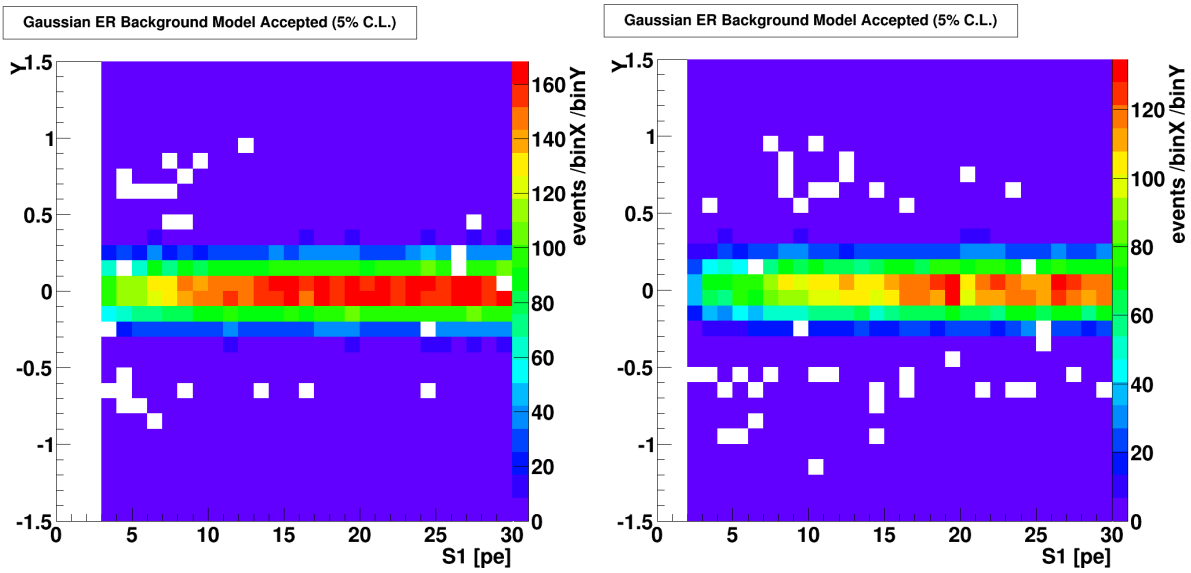


(a) 2011/12 **(b)** 2013/14

Figure 6.19 – Both figures show a 95% quantile (5% C.L.) for model-A and for both data taking periods. The result is based on 10000 simulated log-likelihood ratio test values with simulated event distributions of model-A.

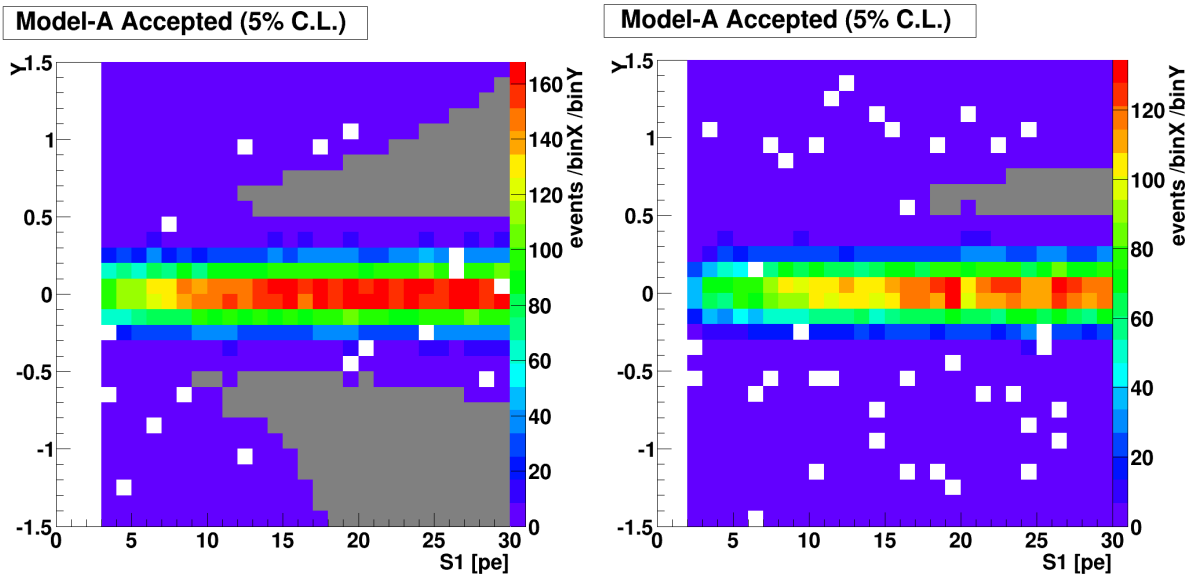
the Gaussian ER background model is rejected. In both data taking periods, the Gaussian distribution approximate the data within $Y < \pm 3\sigma$ pretty well and only three bins are not covered by the 5% C.L. Based on the large number of tested bins, it is statically expected that some bins are not covered by the confidence level. The white bins in the Gaussian tails fail the 5% C.L, this is expected because the Gaussian ER background model does not describe the anomalous leakage events.

The situation changes for Model-A which is given in figure 6.21 for both data taking periods. Again, in the Gaussian tails are a few bins which are not covered by the 5% C.L. This bins are not necessarily the same bins such they are already observed in figure 6.20. This shows that model-A describes the region of the anomalous leakage in a better way. Just by comparing the number of ”white bins“, less bins fail the 5% confidence level criterion in case of looking at model-A. The grey area which appears for the mode-A confidence level calculation shows that there is no 5% C.L. available because in these regions are negative event distributions observed in model-A.



(a) 2011/12 **(b)** 2013/14

Figure 6.20 – A 5% confidence level determines which individual bins of the Gaussian ER background model are accepted and which fail. The Gaussian ER background model is not supposed to describe the anomalous leakage correctly and therefore these bins fail the 5% C.L. (white bins).



(a) 2011/12 **(b)** 2013/14

Figure 6.21 – A 5% confidence level determines which individual bins of model-A are accepted and which fail. The model-A is supposed to describe the anomalous leakage therefore less bins fail the 5% C.L. test in comparison to the Gaussian ER background in figure 6.20. In addition, another set of grey bins arises. Due to the zero event distribution of model-A in the grey region, there is no information available about the acceptance of model-A.

6.8. Alternative Background Model with Log-Likelihood Correction

The question after an alternative ER background model arises from the question if either the Gaussian ER background model or model-A describes the data better. In particular, the definition of model-A with single bins which predict zero events do not fulfil the demands of a well defined background model. The previous subsection 6.7.4 has already determined the acceptance of a certain model based on a 5% C.L. The confidence level introduces a possibility to define an alternative background model which is based on the Gaussian ER background model and model-A. This section will evaluate two possible alternative ER background models which are called in the following "model-B" and "model-C".

6.8.1. ER Background Model ("model-B") with Gaussian Assumption

Model-B uses the Gaussian ER background as main ingredients in the model. As already pointed out, there a few bins which are not covered by the 5% confidence limit. To answer the question for an alternative background model, figure 6.18 is transformed into another presentation: The predicted events are drawn on the x-axis and the $LLR(S|M)_{i,j}^{95\%}$ value on the y-axis. This is shown in figure 6.22 and 6.23 for both data taking periods. This presentation allows to define an envelope (blue line) which acts as upper limit on a 95% quantile (5% C.L.) for the log-likelihood ratio for the Gaussian ER background model. This helps to identify the regions in the followed $LLR(D|M)_{i,j}$ test at which data the log-likelihood ratio test is rejected. This condition is formulated by equation (6.11).

$$LLR(D|M = Gaussian)_{i,j} > LLR(S|M = Gausian)_{i,j}^{95\%} \tag{6.11}$$

The Gaussian ER background is a proper description for the main part of the data. This is quantified by the according log-likelihood ratio test values. In regions where the log-likelihood test fails (see equation (6.11)) the Gaussian ER background description is replaced by Model-A. This definition results a positive ER background model at each single bin. The resulting model-B is shown in figure 6.24 for both data taking periods. Based on equation (6.11) only a single bin fails the upper limit and therefore the Gaussian ER background model dominates model-B. The z-axis is limited a lower boundary of $b_{i,j} > 10^{-10}$ events per bin and therefore the leftover white bins result a positive event distribution. In the 2011/12 data taking period is one bin (grey) which fails the selection criterion on equation (6.11) and at these bins, the leakage model is added to the Gaussian ER background model. This results again to set these bins to zero instead of accepting the negative event distribution. The 2013/14 data taking period shows one single grey bin. This bin is left out during the data analysis with the log-likelihood ratio test. The reason is given in bullet point four of the use-case list in section

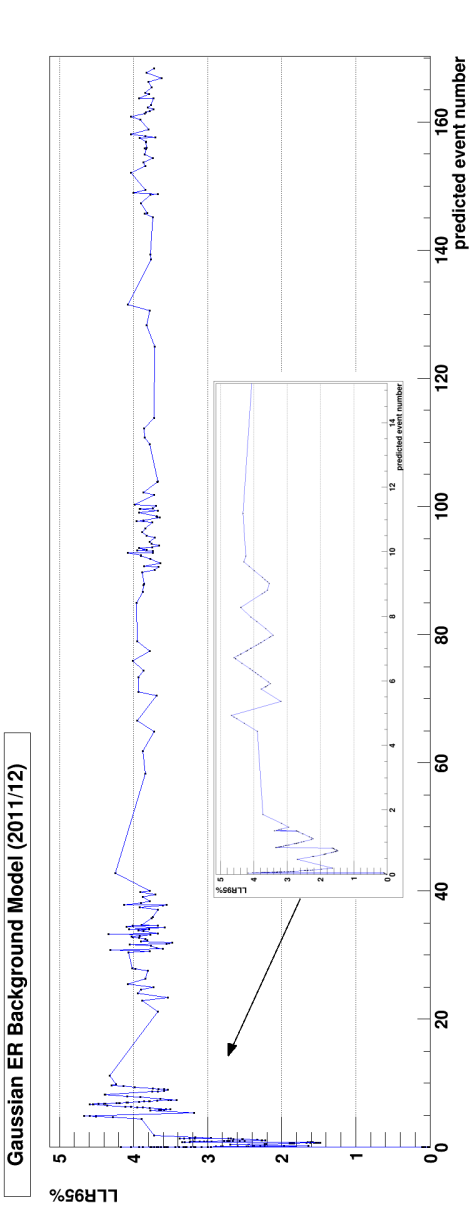


Figure 6.22 – Data taking period 2011/12: The model predicts a certain number of events per bin i, j. These bin information is translated onto the x-axis and connected to the 95% quantile (5% C.L.). This defines a boundary condition at which level the Gaussian ER background model is rejected (see equation (6.11). Below this condition the ER background is accepted.

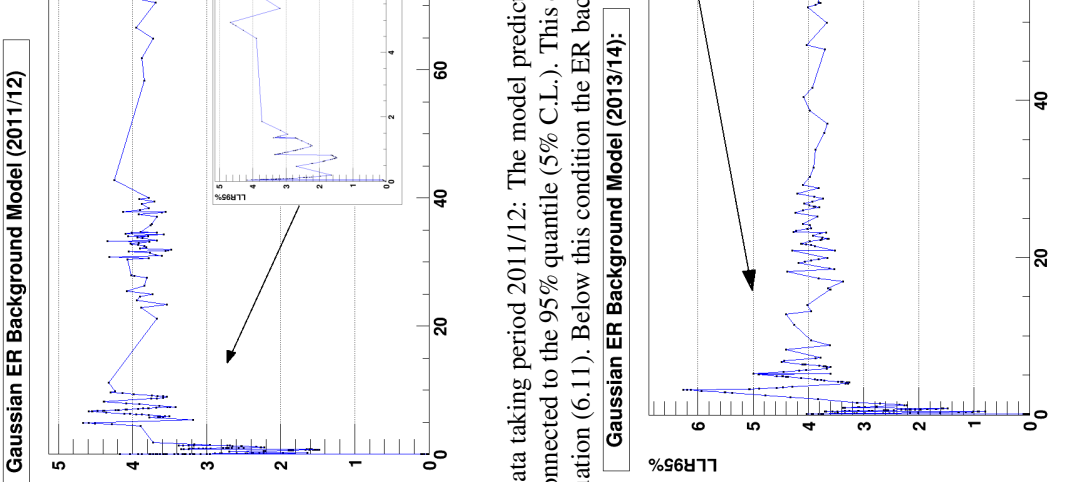
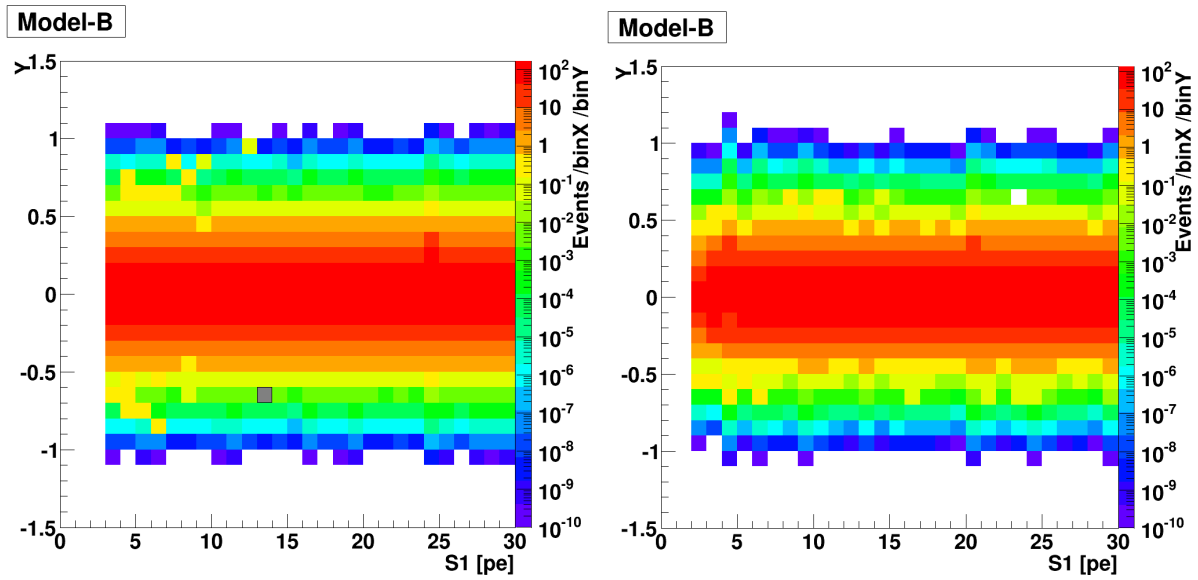


Figure 6.23 – Data taking period 2013/14: The model predicts a certain number of events per bin i, j. These bin information is translated onto the x-axis and connected to the 95% quantile (5% C.L.). This defines a boundary condition at which level the Gaussian ER background model is rejected (see equation (6.11). Below this condition the ER background is accepted.

6.7.1.

Besides this exceptions, the here presented model-A has another disadvantage. The first Gaussian approximation is symmetric and based on the input of Gaussian ER background model, only a few bins are adjusted by equation (6.11). This leads to a dominating Gaussian approximation which does not reflect any asymmetries anymore. A detailed discussion on the asymmetry is in section 6.9



(a) 2011/12 **(b)** 2013/14

Figure 6.24 – A final presentation of model-B which is based on the Gaussian ER background model and minor corrections from Model-A. These corrections are applied to individual bins, in particular bins where anomalous leakage events are observed. The corrections are calculated according to equation (6.11), which is used to reject the Gaussian ER background model and modify it by the leakage model.

6.8.2. ER Background Model (”model-C“) with Model-A Assumption

The previous section introduces a necessary criterion which is used to create an alternative ER background model (model-B). This model makes the valid assumption that the Gaussian ER background model itself is a proper description of the ER background itself and only a few corrections from the leakage model are applied.

There is another approach which solve the question for the correct choice of the model which one works in the opposite way: First it is assumed that model-A is the correct one and there is a choice necessary to decide which of these bins needs to be adjusted by the Gaussian ER background model. This reverse definitions allows to correct model-A by the simplest assumption of the Gaussian ER background model. The tested model model-A is already given in figure 6.6 for both data taking peri-

ods.
Such it is already described in section 6.7.1, the simulated event distributions are used to calculate the log-likelihood ratio test for each bin. Hence each bin is presented by a log-likelihood ratio histogram $LLR(S|M = Model-A)_{i,j}$. The result of these histograms is pretty similar to the already given examples in figure 6.17 for the Gaussian ER background model and therefore no further examples are shown here. The 900 log-likelihood ratio histograms are used to define a similar criterion such it already introduced in section 6.8.1. Again a confidence level of 5% is demanded and hence a 95% quantile is calculated for each log-likelihood ratio histogram. The result of the quantile calculation is shown in figure 6.19. The outcome is similar to the confidence level which is calculated on the basis of the Gaussian ER background model (compare figure 6.18). Only the regions of the anomalous leakage events show a different behaviour. Here the behaviour of the 95% quantile is similar to event distribution in model-A.

$$LLR(D|M = Model-A)_{i,j} > LLR(S|M = Model-A)^{95\%}_{i,j} \tag{6.12}$$

Then the 95% quantile and the number of predicted events of model-A are transferred into a single graph (see figure 6.25 and 6.26). With this, it is possible to define an envelope (blue line) which is used to define a simple criterion at which level the model-A is rejected. This criterion tests the calculated log-likelihood ratio with the data ($LLR(D|M = Model-A)$ in equation (6.12). This allows to keep the model-A at most of the bins and adjust some bins by the most basic knowledge, the Gaussian ER background prediction. The outcome of this method for both data taking periods is given in figure 6.27.
The advantage of model-C is the acceptable definition for the later Profile Likelihood analysis. Each bin of model-C is positive defined and no bin predicts zero events. Figure 6.27 is limited to a lower limit of 10^{-10} events per bin for the presentation but lower event numbers are also calculated. Another advantage of model-C is the restored asymmetry in the Gaussian ER background. Model-C is developed from model-A where the Gaussian ER background is already adjusted by the leakage model and hence the observed asymmetry is re-stamped to the model itself. The asymmetry itself is discussed in section 6.9. The biggest advantage is the Gaussian ER background assumption which is fulfilled at the areas of the negative event distributions.

6.8.3. Compare ER Background Models

In summary, there are four ER background models available which describe the observed ER background data and at least the question remains which of these models is the best one. The previous

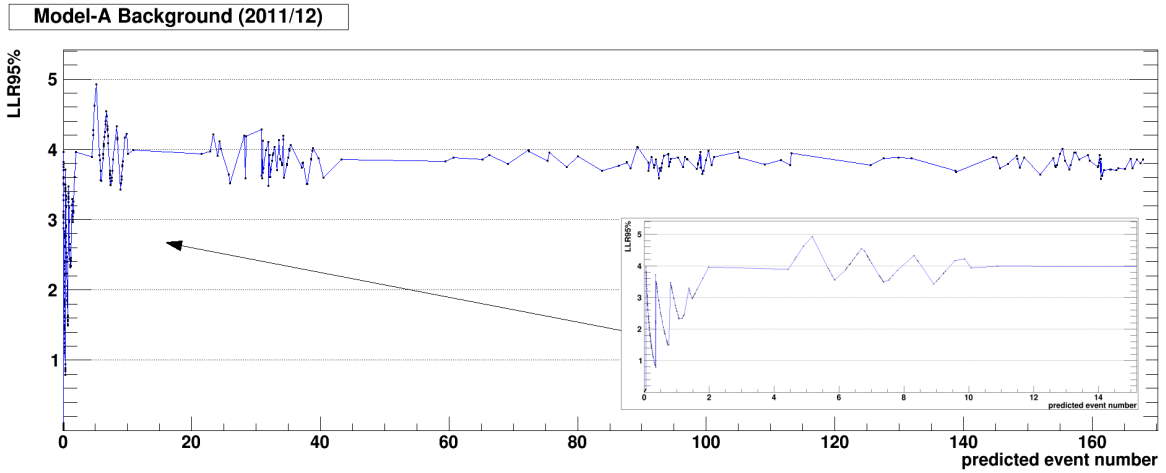


Figure 6.25 – Data taking period 2011/12: The model predicts a certain number of events per bin i, j. These bin information is translated onto the x-axis and connected to the 95% quantile (5% C.L.). This defines a boundary condition at which level the model-A is rejected (see equation (6.12).

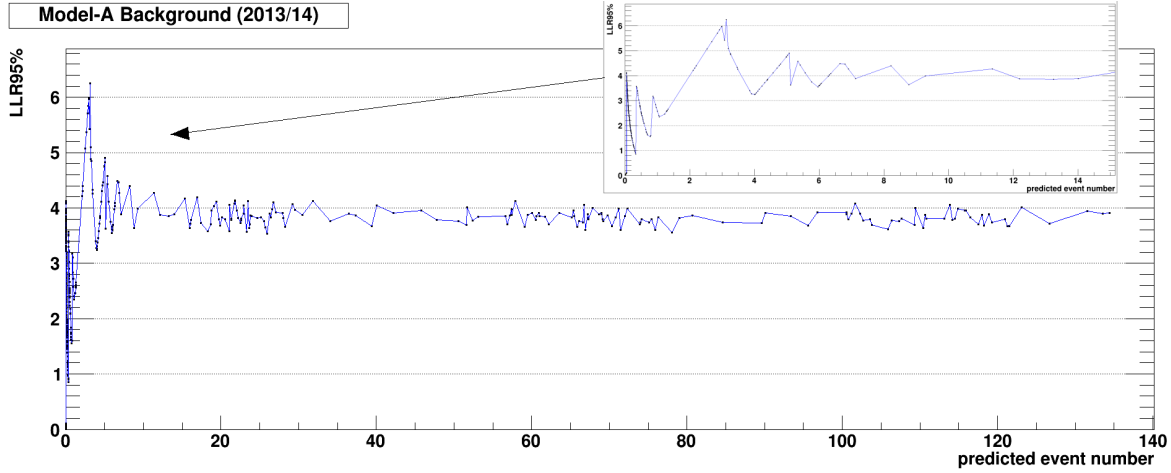
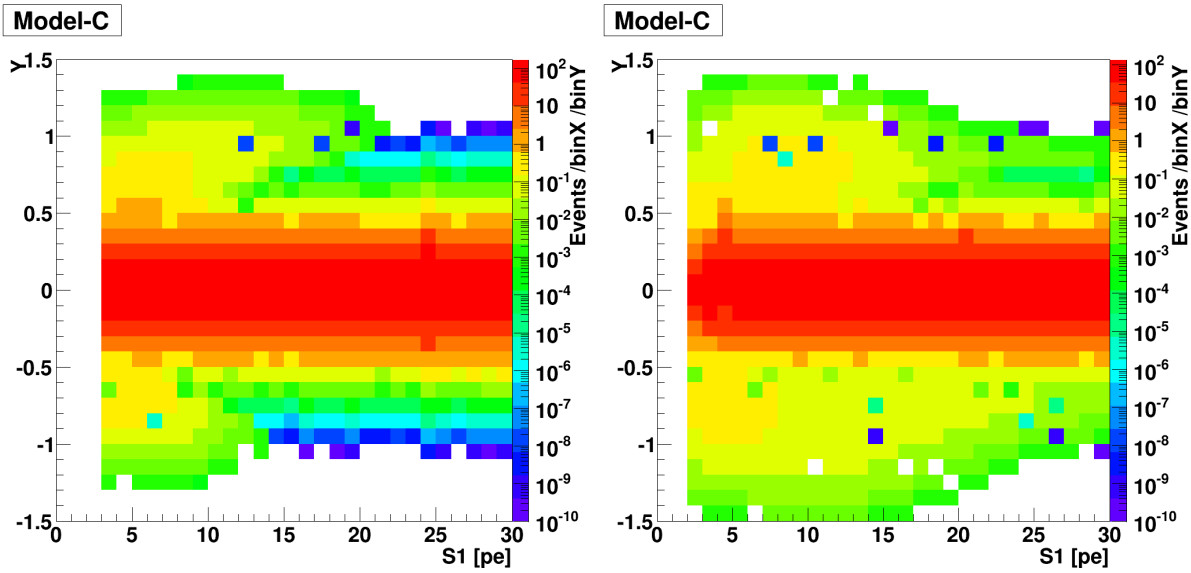


Figure 6.26 – Data taking period 2013/14: The model predicts a certain number of events per bin i, j. These bin information is translated onto the x-axis and connected to the 95% quantile (5% C.L.). This defines a boundary condition at which level the model-A is rejected (see equation (6.12).



(a) 2011/12 **(b)** 2013/14

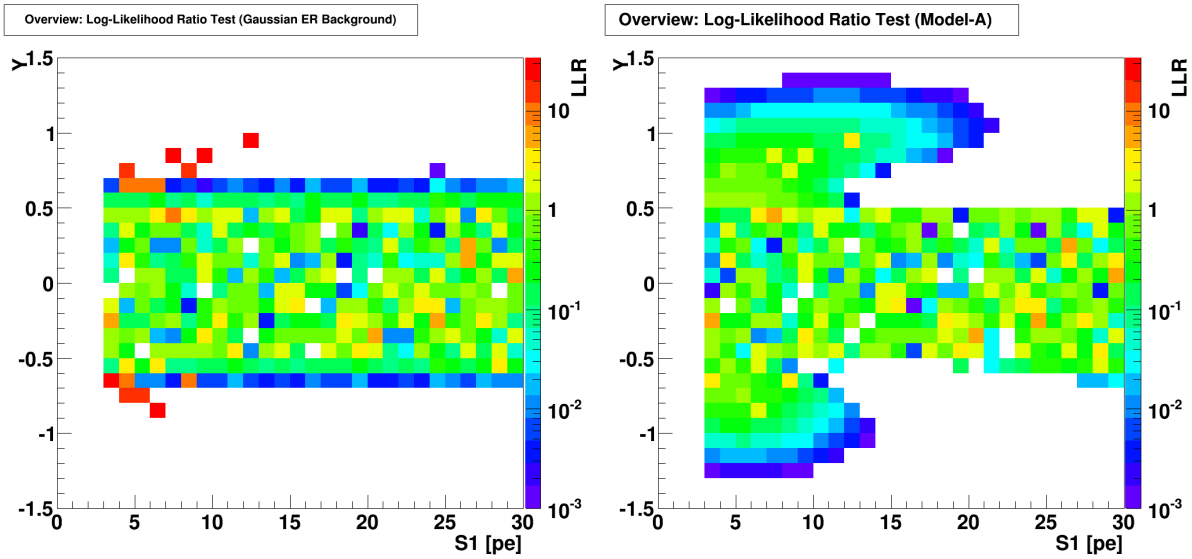
Figure 6.27 – A final presentation of model-C which is developed from model-A. This model assumes that model-A describes the observed data well enough and only a few bins have to be modified. The sufficient assumption of describing the data with the Gaussian ER background model is made for these bins. The result is a well defined ER background which takes the smearing of the anomalous leakage into account and furthermore restore the characteristic asymmetry.

sections evaluated the goodness of the models by a log-likelihood ratio test on single bins. This method is also used to develop to further models model-B and model-C. A final overview is given in figure 6.28 for 2011/12 where the log-likelihood ratio test values for each bin are shown for each model. The logarithmic z-scale is set to the same range in each figure. It is seen that each model has it advantages and disadvantages:
For example, figure 6.28a contains a few bins ($Y \sim 0.9$) where a log-likelihood ratio of almost ~ 34 is reached. This is expected because the Gaussian ER background model does not describe the anomalous leakage events and the result is a larger log-likelihood ratio value in these single bins. Nevertheless, Model-A assures a better coverage of the model with the data (see figure 6.28b). The same bins result a log-likelihood ratio value of $\sim 2-3$. With this knowledge, the model-B (see figure 6.28c) would be preferred because its construction allows to select the model-A event distribution only in these few bins and the other bins are described by the Gaussian ER background model. In contrast, model-C uses model-A as input and replace the model-A bin content by the Gaussian ER bin content in these bins. The outcome is a model where a few bins have a lower event distribution compared to its neighbours. From this point of view, model-B is the best choice because according to its log-likelihood ratio values results the best agreement with the data.
This example shows that a single bin analysis allows to determine the goodness of the model in a very detailed way. This allows to find more bins, such as they are described in the example above, to select another model. After all, it is important to evaluate which model seems to be the best to describe the the whole data space. A first try to quantify the global goodness of the model is already introduced in section 6.7.2. This method is picked up and improved: Instead of calculating the summed log-likelihood ratio on the whole histogram such as it is done (see table 6.6), subtotals are calculated. These subtotal log-likelihood ratio test depend on the data space. The region of interest is below the Gaussian ER background band ($Y \sim -0.42$) and therefore the data space is splitted into:

- $-1.5 \leq Y \leq -0.5$:= Low Y-LLR
- $-0.5 \leq Y \leq 1.5$:= Middle Y-LLR
- $0.5 \leq Y \leq 1.5$:= High Y-LLR

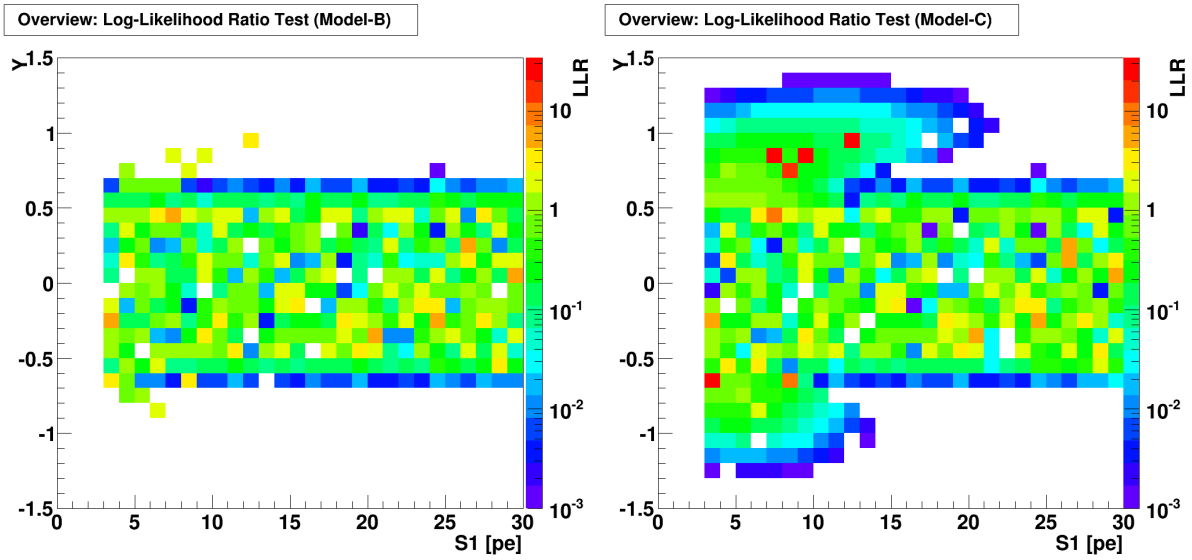
This splitting is chosen according the bin width in Y and allows to determine the partial goodness of the model above and below the Gaussian band. With this boundaries from the splitting, the partial summed log-likelihood ratios are calculated according to equation (6.8) by replacing N by the boundary conditions. The result is given in figure 6.30. In addition the global summed log-likelihood ratio test is also shown in the left column. The right three histograms show the subtotal log-likelihood ratio tests which are normalised to the global log-likelihood ratio test for comparison.
Based on the summed log-likelihood ratio test, model-B matches the data best (blue). The individual

slots show that model-B is the best choice in the regions of the anomalous leakage, followed by model-A and model-C. The Gaussian ER background results the worst agreement. Based on the definitions of the single models, the outcome of the subtotal log-likelihood ratio test agrees with the expectation. Such it is shown, model-A and model-B result a lower subtotal log-likelihood ratio value compared to model-C in the regions of the anomalous leakage. This is also expected because there are more individual bins in the model-C which predict events at position where zero events are observed. This is true in model-C for more bins than in model-B.
By looking at the middle Y-LLR, it is observed that model-C and the Gaussian ER background model matches the data better than model-A and model-B. This is the result of the leakage model which is added to the Gaussian ER background model where difference to the data are stamped on and as a result the subtotal log-likelihood ratio test increases. Due to the definition of model-B, a few bins are adjusted within the Gaussian band, this and statically fluctuations results an increase in the summed partial log-likelihood ratio test.
Nevertheless, Gaussian ER background itself is not as important as the Low Y-LLR region and therefore the choice of the model is done by the lowest summed partial log-likelihood ratio test what results model-B as the favoured model to describe the ER background data. This model does not contain any possible asymmetries in the Gaussian band anymore but as long as the region of interest, for example the low WIMP mass search, is within Low Y-LLR this feature is neglected. From this point, model-B assumes an overall Gaussian ER background which is adjusted in individual bins by the leakage model to describe the anomalous leakage events correctly. The discussed observations are also observed in the data taking period of 2013/14. Hence the according partial summed log-likelihood ratio test is available in figure 6.29. From this result, model-B is also the preferred choice to describe the anomalous leakage events.



(a) 2011/12: Gaussian ER Background

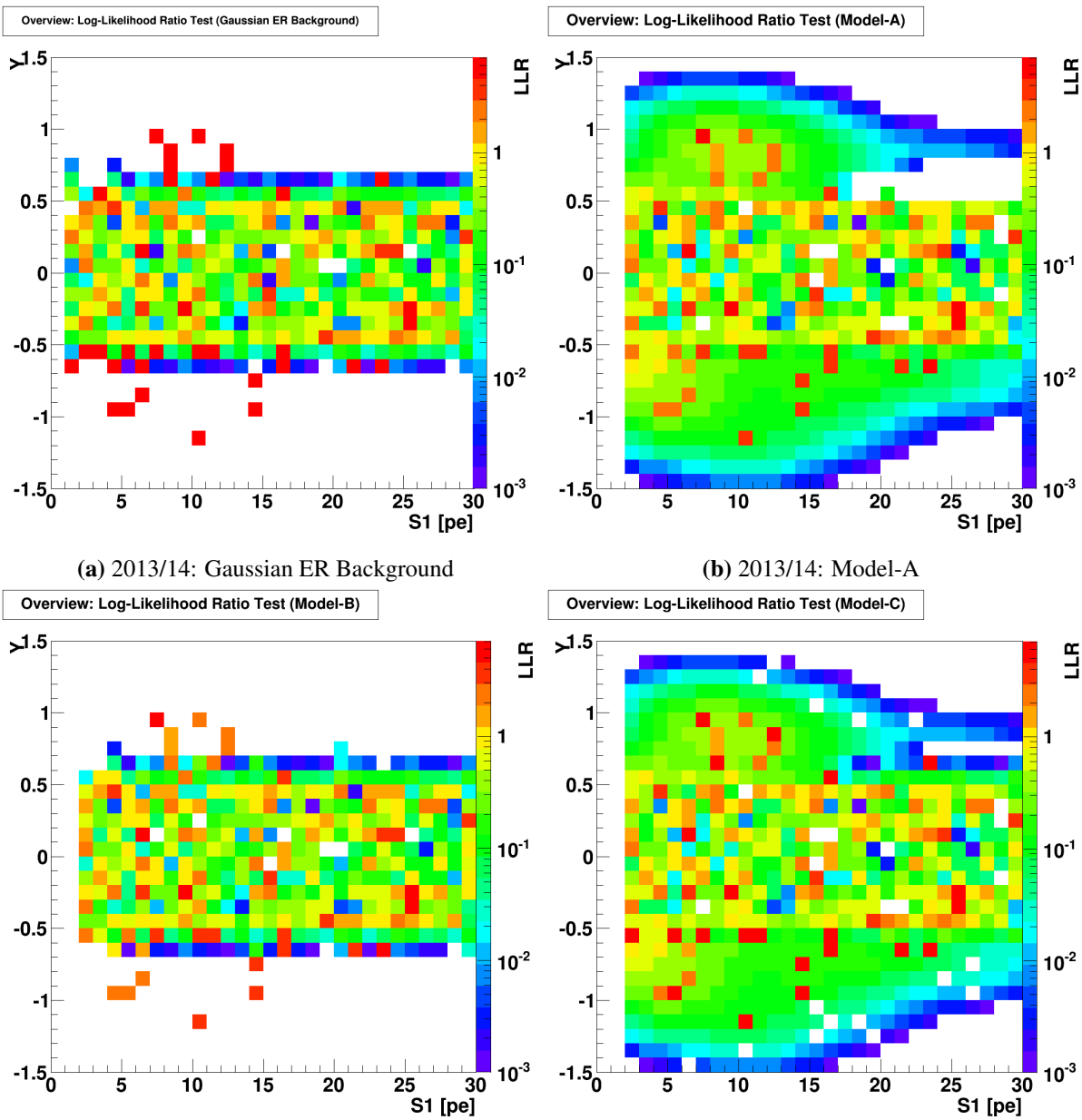
(b) 2011/12: Model-A



(c) 2011/12: Model-B

(d) 2011/12: Model-C

Figure 6.28 – Overview about the log-likelihood ratio test of 2011/12 for Gaussian ER background model, model-A, model-B and model-C. Each figure shows the $LLR(D|M)$ values for each bin in the 2D histograms. Based on this log-likelihood ratio values, a bin dependent choice of the model is possible.



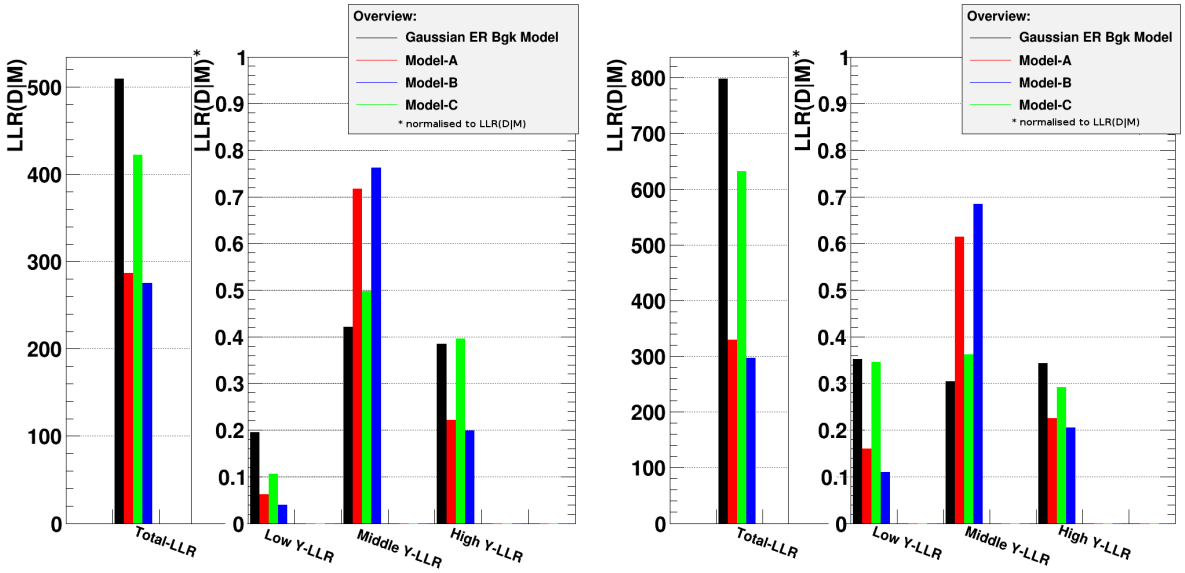
(a) 2013/14: Gaussian ER Background

(b) 2013/14: Model-A

(c) 2013/14: Model-B

(d) 2013/14: Model-C

Figure 6.29.: Overview about the log-likelihood ratio test of 2013/14 for Gaussian ER background model, model-A, model-B and model-C. Each figure show the $LLR(D|M)$ values for each bin in the 2D histograms. Based on this log-likelihood ratio values, a bin dependent choice of the model is possible.



(a) 2011/12 **(b)** 2013/14

Figure 6.30 – Overview about the partial summed log-likelihood ratio test with three pre-defined slices in the data space. The slices are chosen to quantify goodness of the model(s) for the Gaussian part (Middle Y-LLR) and the both regions of the anomalous leakage below and above Middle Y-LLR. These regions are named: High Y-LLR and Low Y-LLR. In partially, the Low Y-LLR is region of interest for low WIMP mass searches. From these results, model-B results the best agreement with the data.

6.9. Asymmetry in the ER Background

An interesting aspect of the ER background model (model-A and model-C) is the asymmetry in Y within $\pm 3\sigma$, indicating small corrections to the Gaussian distributions in individual slices of $S1$. This asymmetry is contained in the residuals of figure 6.7 for 2011/12 and 2013/14. It is not necessarily expected that the ER background data is Gaussian distributed in Y, but if a small asymmetry is a feature of the ER background it has to be taken into account at some point to understand the background model.
Section 6.4.2 show that the first approximation of the ER background data fits almost pretty well. The leftover residual contains the anomalous leakage events and the non-Gaussian background which shows an asymmetry. As long as the Gaussian approximation holds for the region with $-3\sigma \leq Y \leq 3\sigma$ deviation, the residual should contain random fluctuations based on the data distribution. That assumption does not hold during the smearing process with the 2D Gaussian kernel. Within the $\pm 3\sigma$ the smearing produces areas of negative and positive events. The observed pattern in figure 6.9 is not random any more within $\pm 3\sigma$. Reason for that behaviour are small asymmetries in the ER background in the region of the non-Gaussian background. These asymmetries arise from the fact that the approximation with the Gaussian distribution. Another reason is the flattening method: The flattened ER background varies around $Y \sim 0$ with respect to the polynomial which is used for the flattening process. Therefore the flattened polynomial introduces another possible asymmetry.

$$S(X) := \frac{1}{N} \sum_{i=\text{binY}}^{N} \frac{(x_i - < x >)^3}{s^3} \tag{6.13}$$

The following subsection takes a detailed look at the asymmetries observed mainly between $\pm 3\sigma$. The asymmetry itself is measured by the skewness S which is defined according to equation (6.13) [92]. In this formula, the skewness of a random variable x_i is calculated using the mean value $< x >$ and standard deviation s from the event distributions (slices in S1) where N denotes the total event number of the S1 slice. The skewness measures the deviation from symmetric distributed data. To stick with the nomenclature the variable of observation is the discrimination space Y and therefore the index "binY" in equation (6.13) refers to the chosen binning in Y. An excess on the left side of normal distributed data is called "positive skew" and an excess on the right side is called "negative skew" (see figure 6.31). For example, a normal distributed data set results in a skewness of $S = 0$ and can become $\pm\infty$ for huge excesses.

According to the ER background data distribution around $Y = 0$ with a deviation of $\pm 3\sigma$, the Gaussian assumption is tested and measured. Therefore the range of interest to calculate the skewness is between

$\mu_{ER} - 3\sigma_{ER} \leq Y \leq \mu_{ER} + 3\sigma_{ER}$. Events outside these $3\sigma_{ER}$ band are assigned to be anomalous leakage events and not follow a Gaussian distribution any more. The result is shown in figure 6.32 for data period 2011/12 and 2013/14.

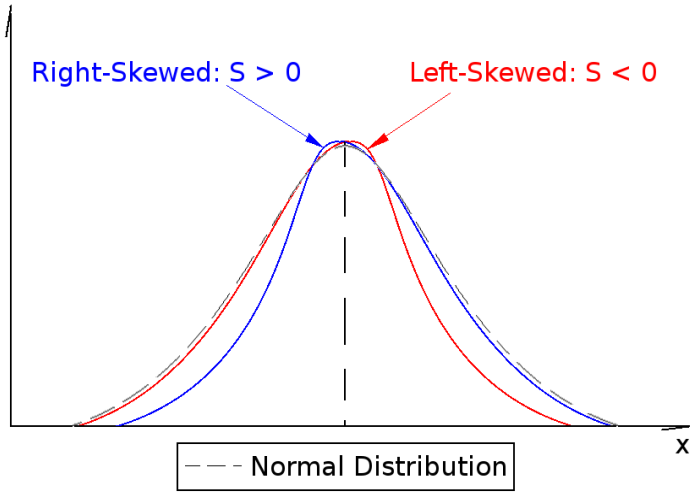
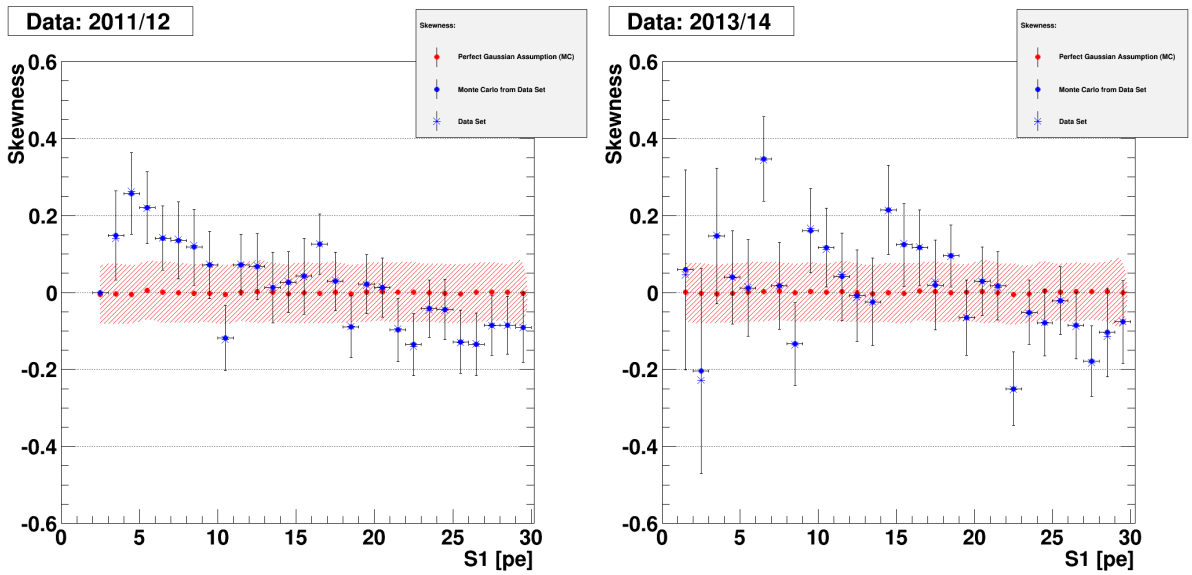


Figure (6.31): The sketch illustrate the behaviour of a left-skewed and right-skewed deviation from a normal distribution. Both possibilities can be separated by the skewness parameter S with $S > 0$ for a right-skewed distribution and $S < 0$ for a left-skewed distribution. An ideal equal distributed data set results a skewness of $S \sim 0$.

The calculated skewness is shown in figure 6.32 by the blue stars. It turns out, that the slices of $S1 <\sim 15\,\text{pe}$ tend to be right-skewed and slices with $S1 >\sim 15\,\text{pe}$ tend to be left-skewed (compare to sketches in figure 6.31). The subtraction of the Gaussian ER background model transfers the skewness into the residuals of figure 6.7. The subtraction underestimates the data in the residual for $Y > 0$ for right-skewed distributions and vice versa for left-skewed distributions. The resulting leakage model contains the observed skewness.

The uncertainties in figure 6.33 are developed with a Monte Carlo technique. For a huge number of Monte Carlo generated ER backgrounds (with underlying Poisson statistics), the skewness is calculated within $-3\sigma \leq Y \leq +3\sigma$. The mean value (blue marker) of the skewness fits pretty well the observed skewness (blue stars). The deviation from the mean value is shown by the error bars in Y. This shows that it is possible to distinguish between the left and right-skewed behaviour. To test the ER background data set, another Monte Carlo study was done. This study with a huge number of simulated Gaussian distributed events results the skewness which is shown by the red markers. The conditions which are applied to Gaussian distributions are $Y = 0$ and $\sigma_{ER} = 0.14$. The red band is the uncertainty based on the deviation from the mean value of the skewness. Such it is expected, Gaussian distributed data result a much lower skewness compared to the ER background data.



(a) **(b)**

Figure 6.32 – The result of the skewness for 2011/12 and 2013/14 is given by the star-shaped markers. An additional Monte Carlos study confirms the observed skewness in the data and allows to determine the uncertainties which are based on the Poisson variation of the input data (blue markers). In the same way, another Monte Carlo study is done with a Gaussian distributed event selection. This confirms that the ideal Gaussian distributed ER data results a much smaller skewness.

The behaviour of the skewness with the underlying data is shown in figure 6.33. This illustrates the event distribution and point to the areas where the Gaussian assumption over and under estimate the data. In any case, the result of the skewness is similar for 2011/12 and 2013/14. Both data sets show a shift from right-skewed to left-skewed behaviour. The polynomial which is used to flatten the ER background data is adjusted from 2011/12 to 2013/14. This indicates that the asymmetry is influenced by the flattening process but that there is also a more global effect that the ER background data are not Gaussian distributed perfectly.
Nevertheless, the asymmetry is small compared to ideal Gaussian assumption because their uncertainties from the Monte Carlo overlap. The general trend of the skewness shift is taken into account by analysing the residual with the 2D Gaussian kernel ϕ.

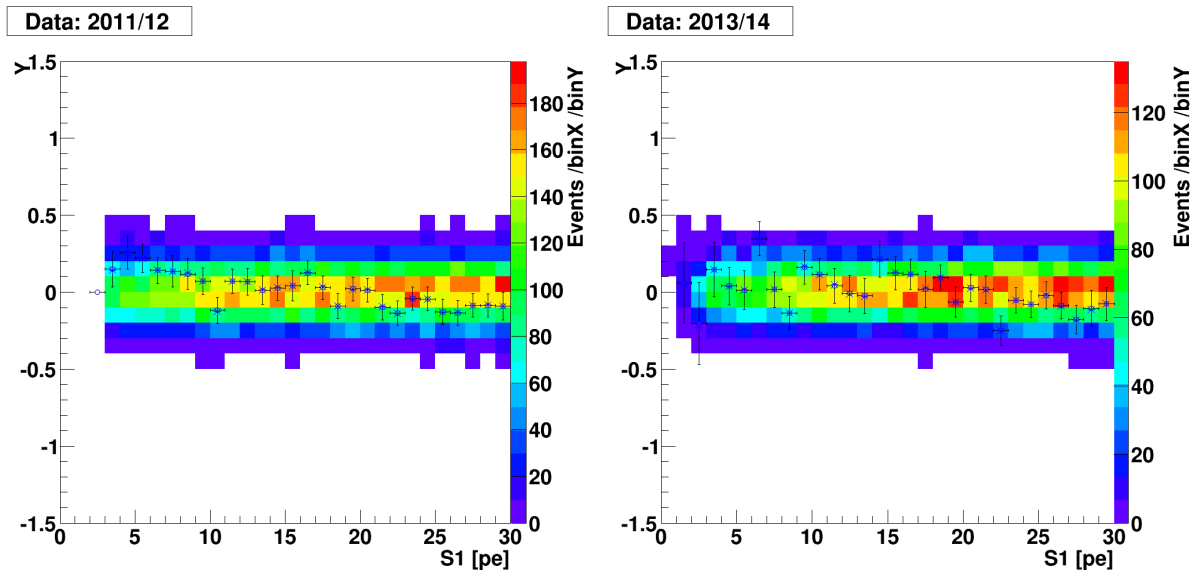


Figure 6.33 – The skewness of 2011/12 and 2013/14 with underlying data distribution. Due to the right-skewed distributions for lower S1 signals and the left-skewed distributions for higher S1 signals, the subtraction of the Gaussian approximation transfers the skewness into the residual of figure 6.7.

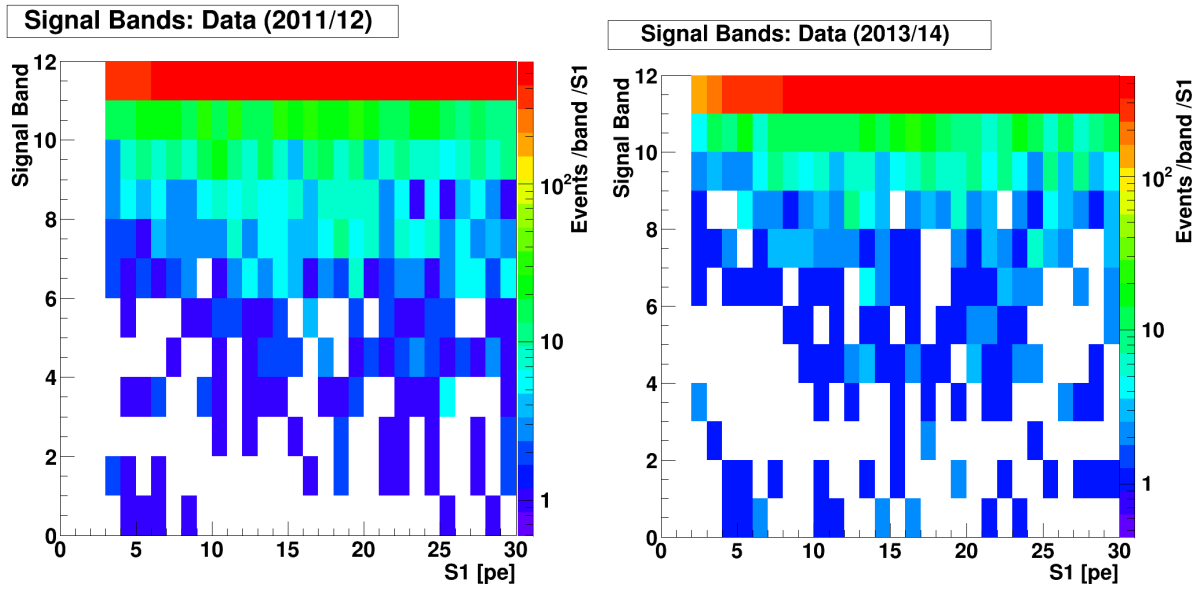
6.10. Goodness of the Model in Signal Bands

The previous sections introduced the log-likelihood ratio test to quantify the goodness of ER background model. The test itself allows to define further alternative models to describe the data. Hence a subtotal log-likelihood ratio is calculated to verify the goodness of the models in a region of interest ($Y < -3\sigma$). This is necessary for future data analysis approaches where the flattened space is used

directly. The actual Profile Likelihood analysis [19] approach foresees to divide the data space into twelve bands. Each of these bands predicts then a certain number of signal, NR background and ER background events. The exact band definition is given in section 3.10 with a well described example. The evaluated ER background models (Gaussian ER, model-A, model-B and model-C) are tested with regard to the band definition. This allows to determine the goodness of a specific model for single bands.

To begin with, the ER background data of figure 6.1 for the data taking periods 2011/12 and 2013/14 is transformed into the band space. The result is shown in figure 6.34. Remarkable is signal band twelve where most of the ER background events are sorted in. Due to the band definition, almost all the Gaussian ER background component and the upper tail is transformed into this band. The remaining eleven bands are related to the *region of interest* which was already used in section 6.8.3 to compare the ER background models. In particular, the lowest bands (b_1 to b_4) describe the region which is far away from the Gaussian ER background band ($> 99.75\%$). Therefore it is important to test the ER background model for lower band numbers. Higher bands ($> b_5$) still overlap with the Gaussian ER background within the 3σ.
Figures 6.35 and 6.36 show the band transform of the discussed models for both data taking periods. The histograms are limited to 10^{-2} events per bin (z-axis). The event distribution fades out to the left lower edge which is similar to the left lower edge in the usual flattened space where the low energy S1 signals are located. The result is pretty similar for both data taking periods. The transformed models show a good comparison to the input data of figure 6.34. A detailed comparison is made band by band. With this band definition and the previous developed ER background models in sections 6.4.2, 6.6, 6.8.1 and 6.8.2 the transformation looks such it is shown in figure 6.35 for the data taking period 2011/12 and in figure 6.36 for 2013/14. In both data taking periods, the similarity between Gaussian ER background model and Model-A is conserved as well as the similarity of model-A and model-C. The lower left edge contains more events in model-A and model-C. This is a result of the dominating leakage model in the construction of the models. It is seen that the Gaussian ER background and model-B predicts less background events for low S1 energies and within bands one to five.

Based on figures 6.35 and 6.36, the total number of entries per band is determined. Figure 6.37 shows the number of observed events in the data by the magenta bars (based on figure 6.34) and compares these numbers to the predicted number of events from the models. In addition the statistically error is presented by the error bars in black. Deviation between different models and the data are observed in the lower signal bands. The higher signal bands show overlap within their error bars. The different



(a) Data taking period: 2011/12

(b) Data taking period: 2013/14

Figure 6.34 – The observed ER background data in 2011/12 and 2013/14 after the transform into the band space. The main part of the Gaussian distributed data is transformed into band twelve and the lower tail is described by eleven bands.

models are compared to the observed data. This is done by the relative deviation in figure 6.38 which is defined by:

$$\text{Relative Deviation} = \frac{M_b - D_b}{D_b} \times 100 \quad [\%] \tag{6.14}$$

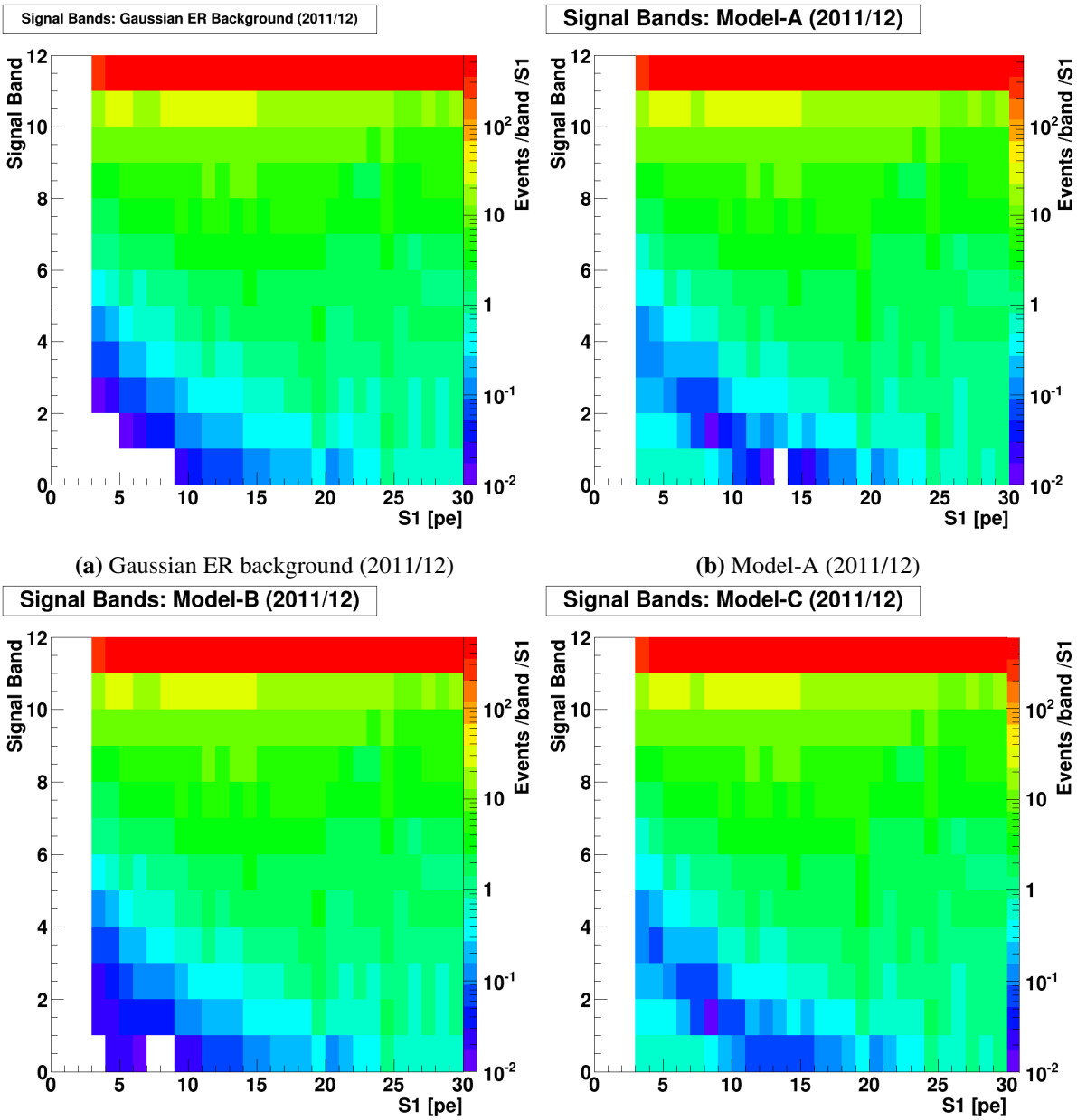
The relative deviation between model and data for band seven is less than 20% and this trend decreases up to a deviation of less than 0.3% in band twelve. The bands one to six show deviations up to $\sim 95\%$, depending on the data taking period and model. These bands show the similarity of the models in detail.

With the differences and similarities a hypothesis test is chosen to compare model and data to predict which model fits the data as best as possible. The test is formulated by the null hypothesis H_0 and H_1.

$$H_0 := \text{The model is true} \tag{6.15}$$

$$H_1 := \text{The model is false} \tag{6.16}$$

Therefore the definition of the p-values is shown in equation (6.17) with the number of predicted events M_b per band. The asymmetry of the Poisson distribution becomes important for less predicted events M_b per band and therefore the statically test is defined by the asymmetric left and right boundary z_0 and z_1 which is ordered by $z_0 < M_b < z_1$. With a proper defined Poisson distribution the following



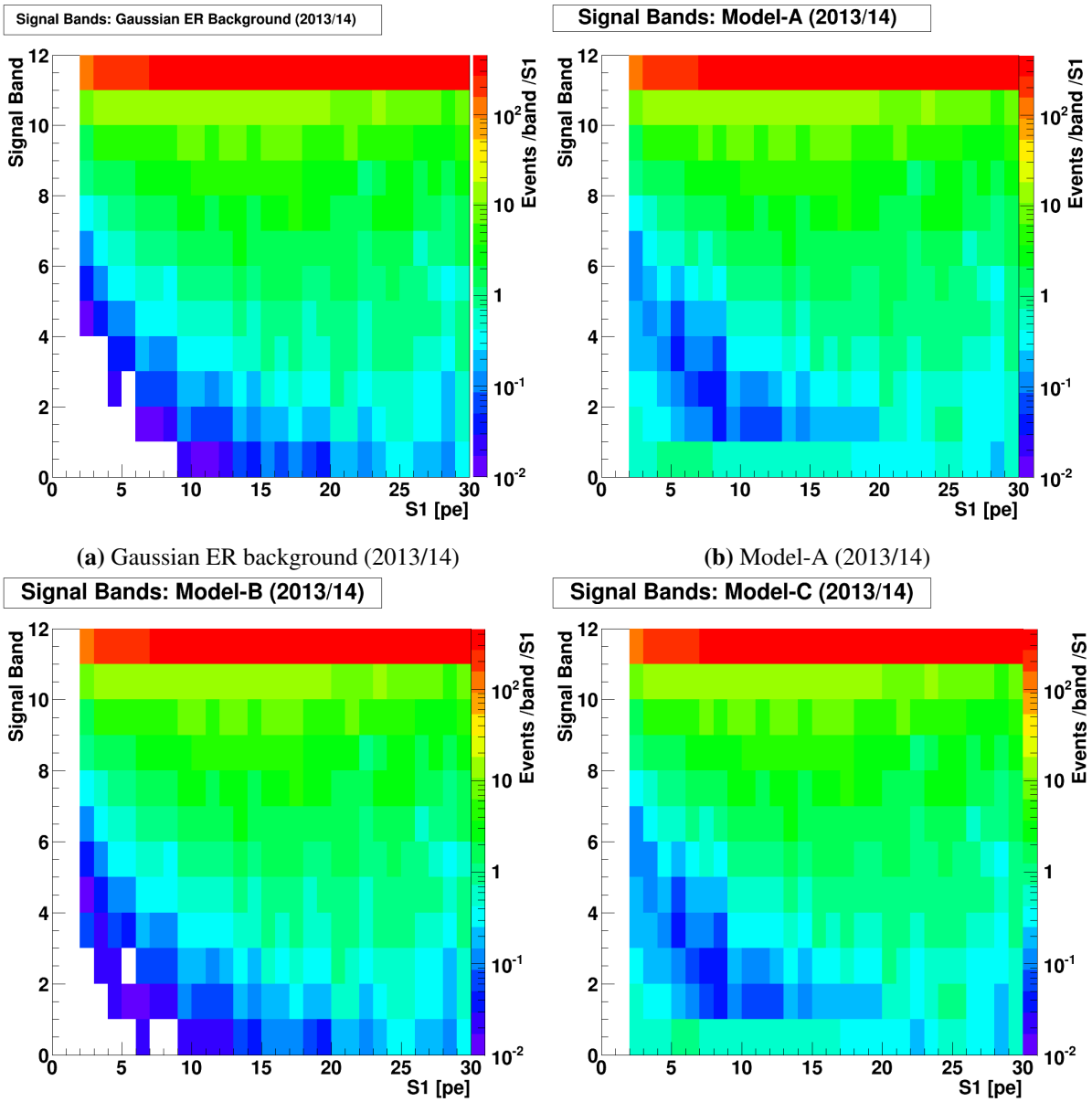
(a) Gaussian ER background (2011/12)

(b) Model-A (2011/12)

(c) Model-B (2011/12)

(d) Model-C (2011/12)

Figure 6.35 – Result of the four discussed ER background model: Gaussian, model-A, model-B and model-C during the data taking period 2011/12. The Gaussian part and the upper tail is transformed mainly into the highest band. The remaining events are distributed into eleven lower bands. Hence it is possible to predict the number of events per band from each model.



(a) Gaussian ER background (2013/14)

(b) Model-A (2013/14)

(c) Model-B (2013/14)

(d) Model-C (2013/14)

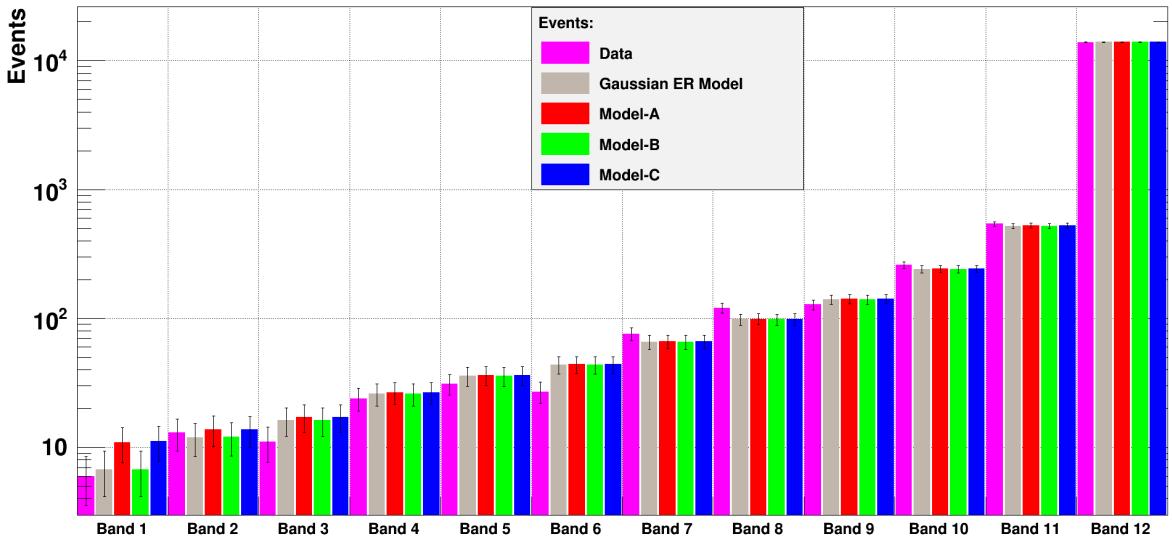
Figure 6.36 – Result of the four discussed ER background model: Gaussian, model-A, model-B and model-C during the data taking period 2013/14. The Gaussian part and the upper tail is transformed mainly into the highest band. The remaining events are distributed into eleven lower bands. Hence it is possible to predict the number of events per band from each model.

criterion must be fulfilled: $Pois(z_0|M_b) = Pois(D_b|M_b) = Pois(z_1|M_b)$ to define an asymmetric region around the Poisson expectation value M_b. This definition allows to quantify the goodness of the model by a p-value. The larger the p-value is as better matches the model the data.

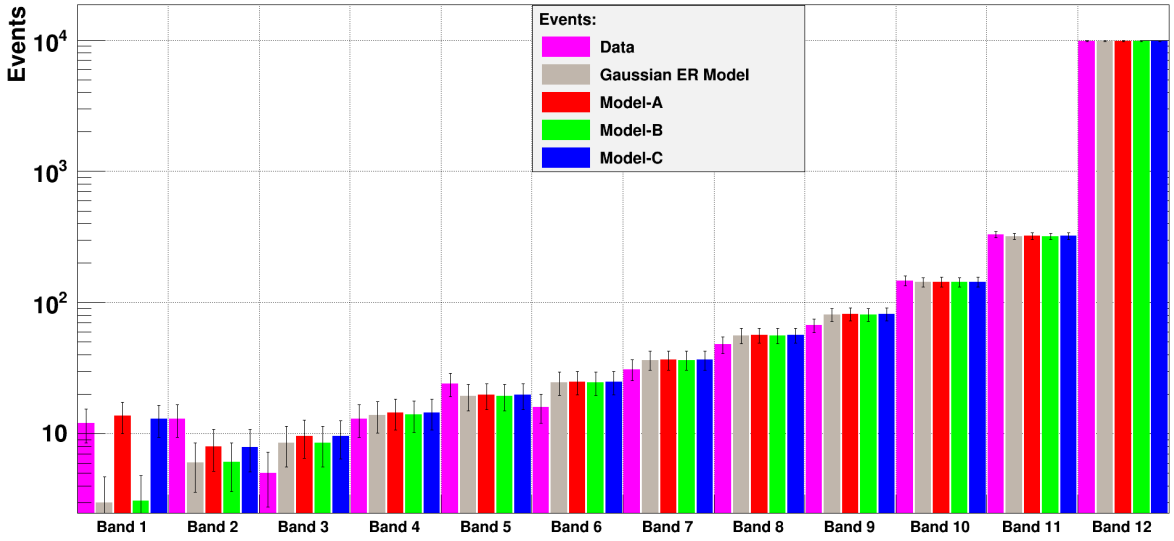
$$\text{p-value} = 1 - \sum_{x=z_0}^{x<z_1} Pois(x|M_b) \tag{6.17}$$

Outcome of the hypothesis test is shown in figure 6.39. Again the similarity between model-A/model-C and Gaussian ER/model-B is observed. By looking and single histogram entries, sometimes the Gaussian dominated ER background fits the data better and vice versa.
From this point, there is no favoured model. In case the four different models compared directly to each other, it is possible to determine a model of choice based on the observed similarities in the p-values. Looking at model-B and the Gaussian ER model both result similar p-values compared to model-A and model-C. Nevertheless, the model of choice is model-B. The reason is given in figure 6.15 where the log-likelihood ratio test has already showed that the anomalous leakage is not well described by the Gaussian ER background. Therefore, the small adjustments in the Gaussian ER background model which result model-B increases the probability to describe the anomalous leakage. A similar argument holds for the choice between model-A and model-C: A proper ER background model predicts a positive event distribution in every bin. However, the definition of model-A set a bunch of bins to zero if no further information is available what is corrected by the approach in model-C. Therefore the leakage model have a larger impact on model-A and model-C and due to the definition of model-C which avoids single bins with zero predicted events, model-C is the model of choice. Nevertheless, by looking only at the data taking period 2011/12 (see figure 6.39a) the Gaussian ER and model-B result a better agreement with the data compared to model-A and model-C in the lower bands 1, 3, 4 and 5. Only in band 2, the additional leakage model results a better agreement with the data but in any case, the calculated p-values exceed 75%. According to the p-values, model-B fits the data as best as possible in 2011/12 under the condition to focus on the lower bands ($b < 5$).
The situation changes in 2013/14, where band 1 and 2 is better described by mode-C. The reason to prefer a model which results a better description of the anomalous leakage events is the anomalous leakage event distribution. The anomalous leakage events are distributed widely in the data space in 2013/14 compared to 2011/12. This affects the leakage model: A single 2D Gaussian distribution is added completely to the leakage model in case a single event has no direct neighbours within a certain distance which contain a negative residual. The widely spread anomalous leakage events in 2013/14 result than a widely spread leakage model which affects in particular band 1 and 2 more than it is possible by the 2011/12 data.
Overall, the band by band comparison with the here calculated p-values shows at last for band 12 a high agreement for data and the four models. The statically error for the predicted and observed events

(see figure 6.37) in band 12 overlaps for all models and data. This points to an over-fitting of the ER background model in band 12 which contains the main part of the Gaussian approximation. Hence the Gaussian approximation already over-fit the data.

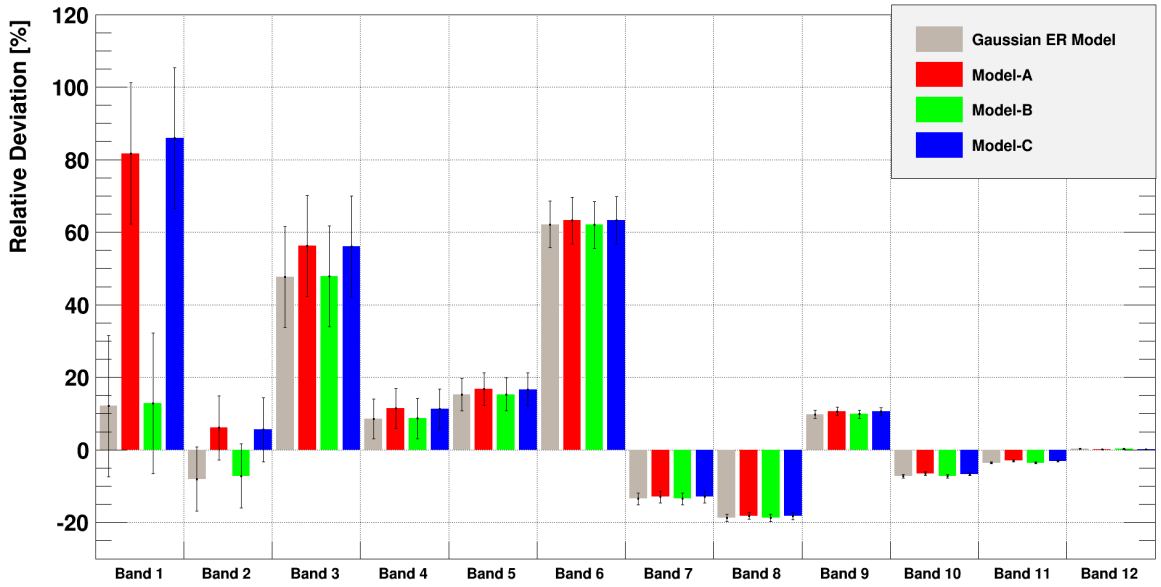


(a) 2011/12

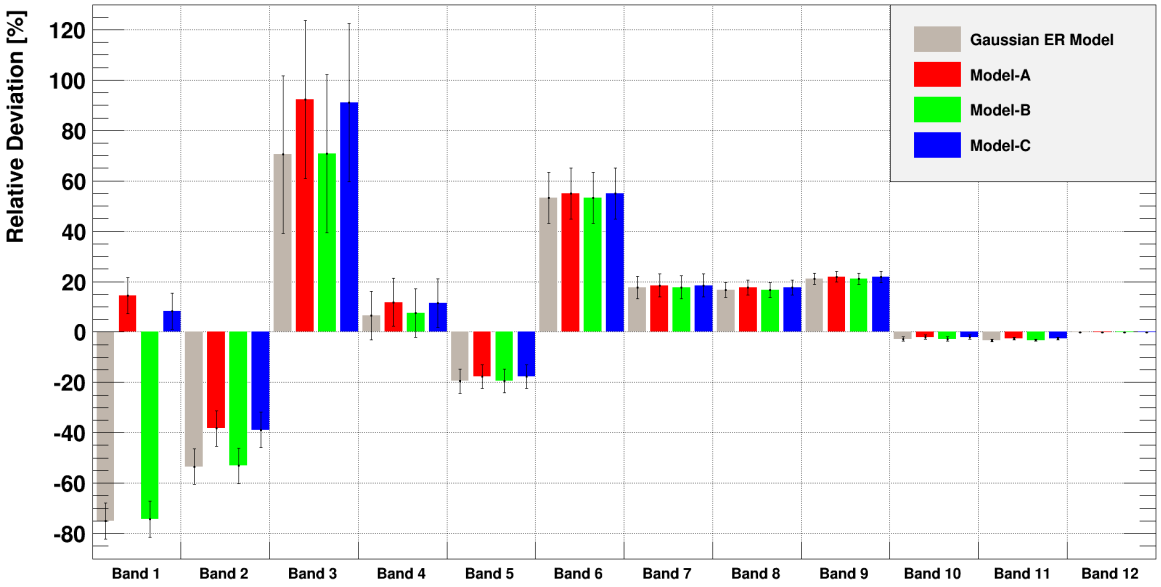


(b) 2013/14

Figure 6.37 – The total number of entries for four different models: Gaussian ER background, model-A, model-B and model-C for each data taking period and band. In addition, the number of observed events per band and data taking period are shown. The black error bars refer to the statistically error of the observed/predicted events.

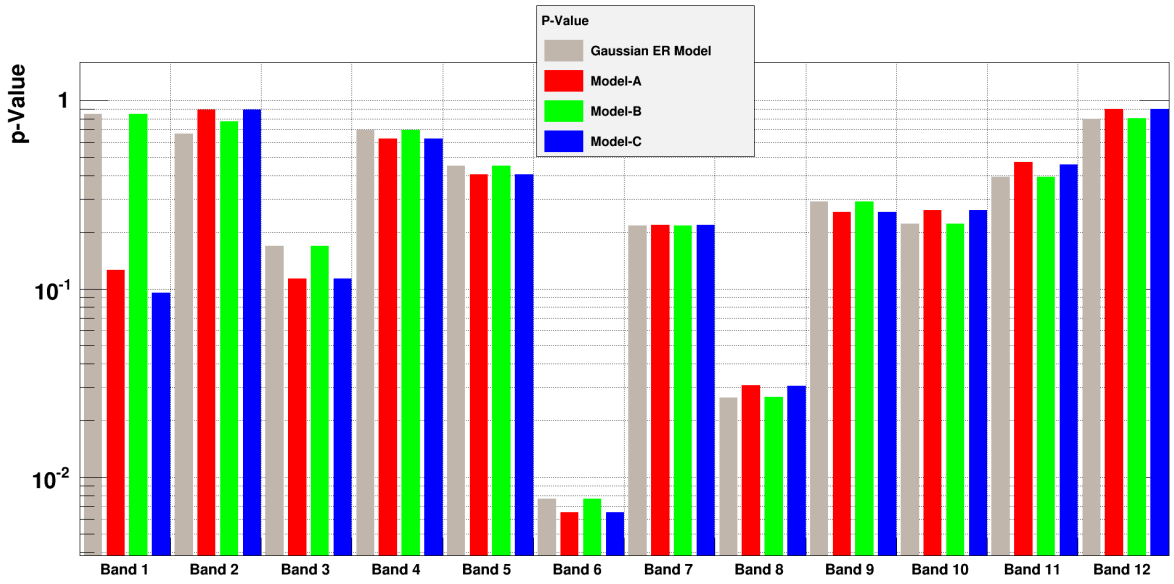


(a) 2011/12

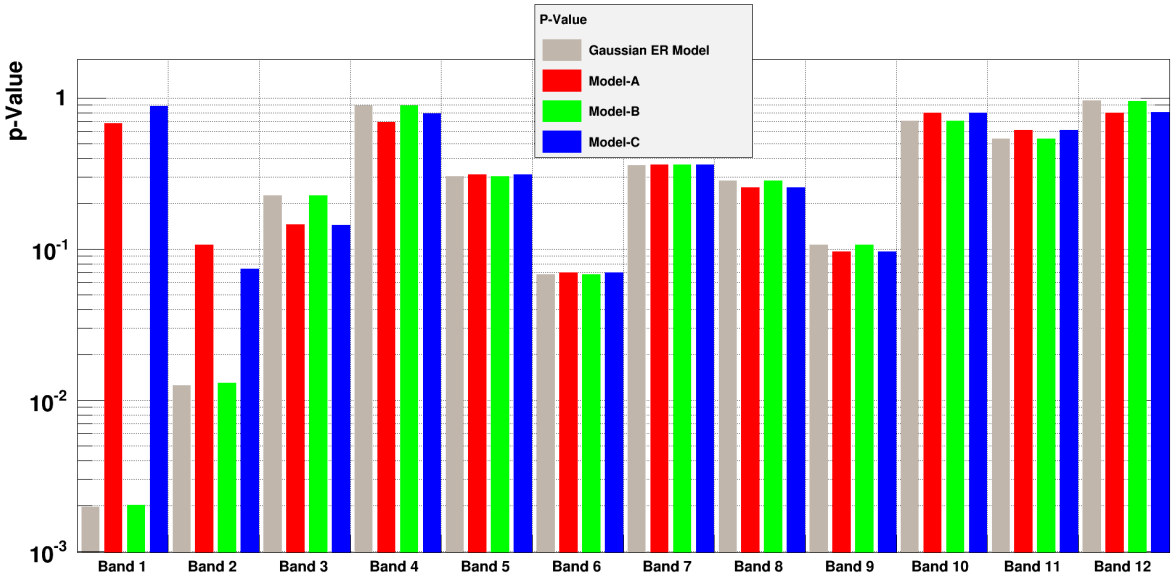


(b) 2013/14

Figure 6.38 – The relative deviation of the models to the observed data for each band. It reflects the similarity of the Gaussian ER/model-B and model-A/model-C. The highest two bands show less than 5% deviation. In general, the deviation becomes important for the lower bands in where the goodness of models needs to be as good as possible to describe the occurring anomalous leakage events. The here presented symmetric error bars are the result of the error propagation with the relative uncertainty on data and model based on the statistical errors.



(a) 2011/12



(b) 2013/14

Figure 6.39 – A hypothesis test is carried out to verify how well the different models fit the data. The p-values describe the goodness of the model for each band separately. As larger the p-value is as better fits the model the data.

6.11. Impact on the XENON100 WIMP Exclusion Limit

Based on the comparison of the four ER background models in section 6.10 there is no final conclusion on the *right* choice of the model. In case of zero WIMP observation a WIMP exclusion limit is calculated. The actual WIMP-nucleus exclusion limit of the data taking period 2011/12 is presented in a past publication [17]. With this result and the four ER background models which are discussed here, it is interesting to determine the impact of the ER background model on the exclusion limit calculation. The published background model in 2011/12 was used in the PL to calculate the limit for spin-independent WIMP [17] interactions. The standard limit takes the ER background and the NR background into account. The PL approach is described in a past XENON100 publication [19]. The exclusion limit calculation is done by *XEPHYR* (see also section 3.11). The program is developed in the XENON collaboration for e.g. limit calculations. The here presented comparison of the exclusion

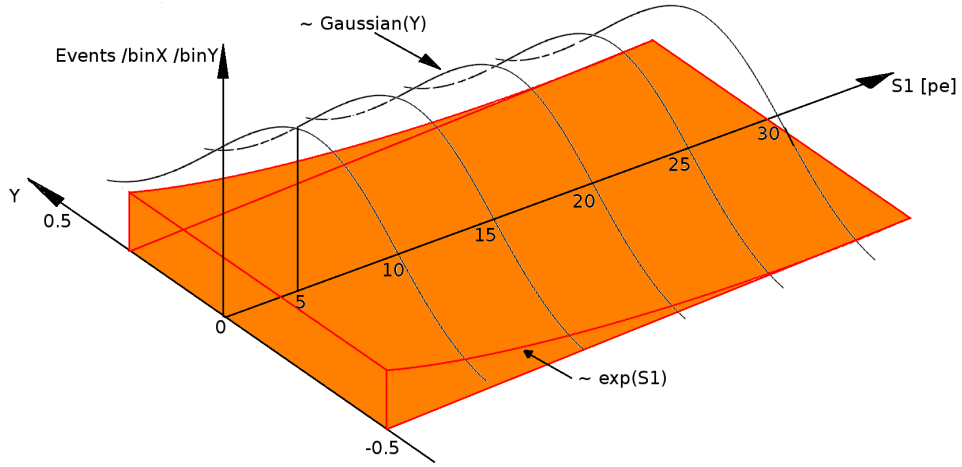


Figure 6.40 – This a sketch of the ER background model in 2011/12. The first approximation with a Gaussian distribution is done similar to section 6.4.2. The remaining anomalous leakage events are analysed by a one dimensional exponential description (red). Adding up both descriptions in a 2D histogram results the figure 3.32a in section 3.10.

limits is done under assumes that there is no NR background. Hence ER background is the only background. With this background contribution, the WIMP exclusion limit is tested for each background model and the comparison is based on the past ER background model from 2011/12. The past ER background is also a phenomenological model which uses the same Gaussian approximation such it is described in section 6.4.2. The remaining anomalous leakage events are described in a one dimensional projection on S1 and are approximated by a exponential function. This is sketched in figure 6.40 where the Gaussian background is drawn in black and the underlying exponential anomalous leakage is drawn in red. The ER background model of 2011/12 is already shown in figure 3.32a (see chapter

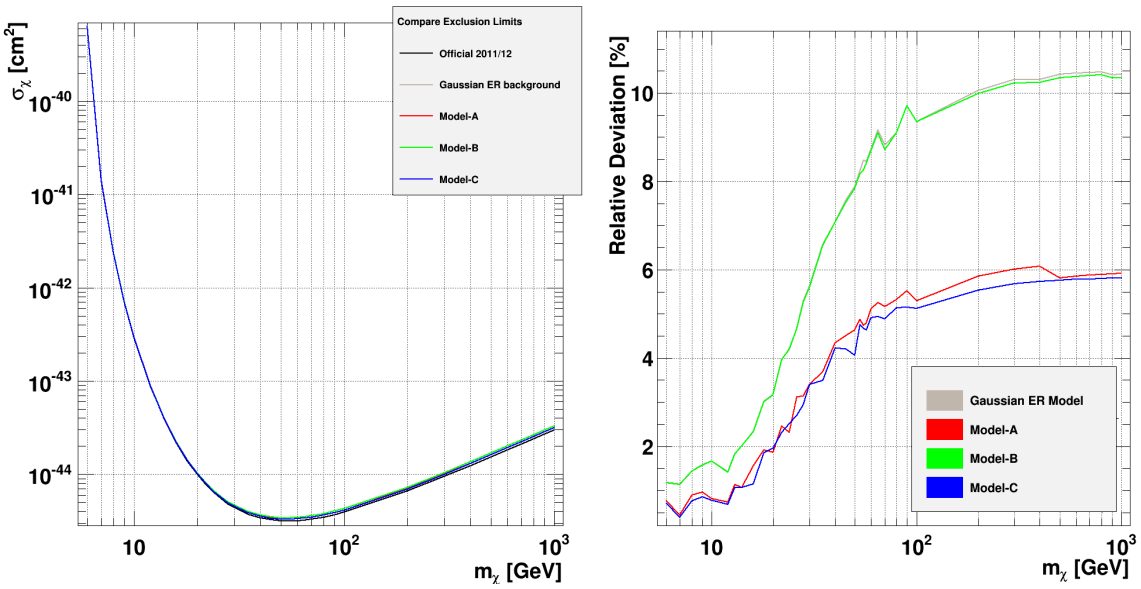
3.10). The comparison is done directly: The four ER background models are plugged into *XEPHYR* and the limits are calculated with the DM data of 2011/12 (after cuts). Based on the limit calculations from *XEPHYR*, it is not possible to conclude which of the four models is better or worse. This is only possible by comparing ER background data to the model. Nevertheless, the choice of the background model has an impact on the exclusion limit which is evaluated in this section.

The outcome of the WIMP exclusion limit is given in figure 6.41a. Due to small changes in the ER background there is no large change expected in the WIMP exclusion limit and therefore all five exclusion curves are near to each other. To access the differences, the relative deviation is calculated in relation to the official ER background model (black line in figure 6.41a). This is shown in figure 6.41b. The relative deviation is defined to ($MD = \{\text{Gaussian ER}, \text{ model-A}, \text{ model-B}, \text{ model-C}\}$):

$$\text{Relative Deviation} = \frac{\sigma_\chi^{MD} - \sigma_\chi^{2011/12}}{\sigma_\chi^{2011/12}} \times 100 \quad [\%] \tag{6.18}$$

First of all, an already observed similar behaviour of the Gaussian ER and model-B, although for model-A and model-C, is observed in the relative deviations of the exclusion limits. The band by band comparison of section 6.10 showed this already. Due to the definition of the relative deviation figure 6.41b shows a higher exclusion limit with the four discussed ER background models compared to the previously used ER background model in 2011/12. The limit rises up to 10% for Gaussian ER background/model-B and up to 6% for model-A and model-C. Based on this observation, the former ER background model results the lowest WIMP exclusion limit. The phenomenological approach to model the anomalous leakage events by a 2D Gaussian distribution describes the observed data with higher details than an exponential approximation and hence differences arise in the limit calculation.
By construction the Gaussian ER and model-B are *Gaussian* dominated models. The Gaussian only approximation of the Gaussian ER background model is adjusted by a few bins which result from the 2D Gaussian approximation. In contrast to the models Gaussian ER and model-B, the models model-A and model-C are strongly adjusted in the regions of *observed* anomalous leakage. This adjustment results a higher detailed model and therefore the relative deviation of the exclusion limits between model-A/model-C is lower than in the Gaussian dominated models.
Any of the four models result a higher exclusion limit what points to the fact, that each of these models predict not enough background events. In particular the inadequate description of the anomalous leakage with the Gaussian ER and model-B background models is responsible for the rise in the exclusion curve up to 10%. Model-A and model-C improve in the event prediction of the anomalous leakage and therefore the additional predicted events increase the relative deviation to maximal 6%.
Pointing to the tested WIMP mass m_χ, the relative deviation is $\sim 0.5\%$ for model-A/model-C and $\sim 1.5\%$ for Gaussian ER/model-B within $m_\chi <\sim 11\,\text{GeV}$. Then a rise is observed in the relative

deviation within $\sim 11\,\text{GeV} \leq m_\chi \leq \sim 100\,\text{GeV}$ and above $\sim 100\,\text{GeV}$ it becomes flat. The relative deviation becomes smaller for lower WIMP masses.



(a) (b)

Figure 6.41 – The four discussed ER background models and the previous used ER background model of 2011/12 result similar WIMP exclusion curves. The lowest exclusion curve is calculated for the previous 2011/12 ER background model. To compare the four discussed models: Gaussian ER, model-A, model-B and model-C, the relative deviation is calculated (figure 6.41b). The Gaussian ER/model-B show larger relative deviations compared to model-A/model-C what comes from the higher detailed description of the anomalous leakage region in model-A/model-C.

This comparison does not vote for a particular background model but shows the impact on the exclusion limit. This favours to reject *Gaussian* dominated ER background models and advertise the need of a proper description of the anomalous leakage, which was the goal of this section. The here presented model-A and model-C show a higher detailed description of the anomalous leakage of the ER background in the discrimination space.

6.12. Summary

The here presented work on the electronic recoil background describes a possible way to develop a phenomenological ER background model from the data. The ^{232}Th and ^{60}Co data are approximated by a Gaussian model in a first step. After the Gaussian ER model is subtracted from the data, the residual

which contains the anomalous leakage events, is smeared by a 2D kernel (second approximation).
By adding up both models to model-0 it turns out that there are *negative* events predicted. To correct this miss-behaviour model-A is created where the negative event distribution is set to zero. Another log-likelihood ratio criterion quantify the goodness of the model and allows to create two other models (model-B and model-C) which combine the Gaussian ER background model with the leakage model in a certain way.
The comparison of these four models (Gaussian ER, model-A, model-B, model-C) in the band space does not result a satisfying criterion which favours a specific model. After all the Gaussian approximation over-fits the data in band 12.
The observed asymmetry is quantified with a Monte Carlo method to verify if the observed asymmetry is observed by chance or results from the data. It turns out that there is a small asymmetry along the S1 axis.
Finally, the impact on the XENON100 WIMP exclusion limit is calculated with *XEPHYR*. It turns out that the choice of the 2D Gaussian models (model-A/model-C) worse the limit by $\sim$6% compared to previous used ER background model of the data taking period 2011/12. The exclusion limit is not a good indicator for the goodness of the model but it turns out that *Gaussian* dominated ER background models worse the limit up to $\sim$ 10% at maximum. From this point, the Gaussian ER background predict the lowest anomalous leakage event rate and every better phenomenological model of the anomalous leakage will lower the exclusion curve.
The here shown techniques to develop the uncertainties on the ER background model are tested with the here described ER background models. There are further methods available to calculate the uncertainties on the ER background and therefore the study is not yet finished. The question of which of the here presented ER background models match the data at best is not yet answered. Besides the phenomenological approach to develop the ER background further physical motivated background models are possible if the anomalous leakage is better understood.

7. Summary and Outlook

"But it has something to do with *Dark Matter*. – It has everything to do with *Dark Matter*."

Dr. Chester Ray Banton,
The X-Files, "Soft Light"

7.1. Summary on the XENON100 Data Analysis

Modern experiments have become much more complex nowadays and the amount of taken data also has risen. The use of computer technology for the data analysis allows to implement statistical methods for an extensive analysis. Experiments such as the XENON100 or its successor experiment XENON1T take a huge amount of data. This starts with the control of the environment such as temperature of the liquid xenon, the pressure inside the TPC or even the room temperature where the experimental setup is mounted. Furthermore the data of interest (Dark Matter or calibration data) is taken and stored as raw data and therefore it is necessary to have a high quality in the data selection and to prove a good understanding of the detector to extract real physical events out of this huge amount of data. This needs to take many methods into account which must be evaluated.
The presented work on some of the specific XENON100 data analysis procedure show this effort. Therefore the goals of the single data analysis procedures were emphasised and the difficulties were discussed in detail. This work uses the input of different calibration sources, which at least are needed to understand the TPC response completely. The here given selection of data analysis procedures are developed from XENON100 calibration data. The method itself, once it is proven that it is working well, can be also applied for future experiments.

The presented work is subtracted into three parts of data analysis topics. Based on the ^{232}Th and ^{60}Co calibration data, a background model is made. Dark Matter search experiments have such a low-radioactive environment, that the few events which describe the background are important to show a good understanding of the background. Therefore a model is made to describe the ER background event distribution which predicts the number expected background events for the Dark Matter data analysis. The model itself relay on a proper description, which is in case of the ER background, a Gaussian distribution in the discrimination space parameter Y vs. S1 (3D and light collection efficiency corrected S1). Unfortunately, the description become more complicate by another background population, the leakage events. This leakage events are populated outside a 99.75% band in Y and hence the number of observed events is not comparable to the number of expected events by the Gaussian assumption.
This work includes an extensive description of the ER background data set (for data taking periods 2011/12 and 2013/14) which separates the data into a first approximation (Gaussian) and is then subtracting the Gaussian model from the data. This procedure yields a residual plot which includes the leakage events and the variations of the Gaussian model inside a 3σ deviation.
The residual is analysed by a kernel $\phi(S1, Y)$ which re-distributes the observed number of events in the residual. This smearing decreases the resolution, but allows to quantify the regions of the leakage events. This is possible because the leakage events stick out of the residual. They are not influence

(and described) by the Gaussian first approximation. This smearing yields to a leakage model and, at least, the leakage model is added up (bin by bin) to the Gaussian model. This yields the final ER background model in where the number of events are conserved.
Finally different tests are done to check the goodness of the model by e.g. different hypothesis tests: A first one calculates a table of p-values for each bin. In that way the data is compared to the final model and to the Gaussian-only approximation. This yields a p-value corrected ER background model. Another approach is tested by a coordinate transform from the discrimination space Y vs. S1 into a pre-defined set of horizontal bands vs. S1. This transformation becomes interesting because only the total number of expected events per band is used in the Profile Likelihood analysis at the moment. In that way, the ER background description lose information but becomes easy comparable to a previous used ER background model in 2011/12. The transformation does not need the p-value corrected ER background model necessarily. Based on another hypothesis test, the p-values shows a better agreement of the here described model, in comparison to the previously used ER background model. Finally, an exclusion limit is calculated with *XEPHYR* to see the impact of the here described model, compared to the previously used: The exclusion limit is higher for the here described ER background model by up to 7% (without p-value correction). This outcome rises the question on the goodness of ER background model itself and leave the unanswered question on a better way to estimate the uncertainties of the background model itself and furthermore what is the best way to compare two different ER background models.
This work on the ER background model presents a method to predict the ER background based on the observed data. In case of not observing any leakage events the smearing with the kernel $\phi(S1, Y)$ does not lead to proper scientific results. This is a limit if a background model is developed from the data directly. In case of modelling the leakage events, there are only a few leakage events from which the model is made from. Different approaches are tested at the moment to describe the ER background as good as possible and hence the presented work here will be a basic working principal for further tests.

Such it is described in section 3.7, it is important to define a set of cuts which select only physic driven events from the data. Therefore a certain set of cuts is defined and applied to data. One important cut is the Xs2peaks and hence its acceptance. This cut introduces a threshold on the S2 signal of 150 pe. This assures a almost 100% trigger efficiency for S2 triggered signals. In the usual case, the S2 signal will be larger than the previous S1 signals and therefore the trigger on the S2 signals is useful. Once a cut is defined it is important to determine its acceptance which means to see how likely it is to observe an event for a given cut and energy. The energy is driven from the defined variables in XENON100 which is a 3D and light detection efficiency corrected S1 signal. Therefore the data (cS2) is sorted into histograms of single slices in S1, where each S1 slice has a width of $\Delta S1 = 1\,\text{pe}$. The shape of

each cS2 histogram is approximated by an adjusted Poisson distribution. This adjustment introduces another fit parameter (β) which allows to scale the width of the Poisson distribution additional. Due to the original definition of the Poisson distribution, it would be enough to have one parameter, the expectation value λ. Therefore a wide range of possible, but fixed, values of β were tested on each of the cS2 histograms at the same time. Based on the log-likelihood values of the different fits, the best choice for the additional parameter β is made at the maximum likelihood. From the best choice of β, the related expectation values λ for each cS2 histogram are selected and fitted by with a simple linear approximation. This results a function $\lambda(S1)$ which predicts the cS2 signal for a certain S1 signal. The advantage is given by the continuous description, even down to 0 pe in S1.
Keystone of the Xs2peaks acceptance calculation is the Monte Carlo technique which is used calculate the acceptance. This method uses $\lambda(S1)$ as an input to draw random values of λ, which present expectation values (together with the best choice β) of a Poisson distribution. Random draws from these cS2 - Poisson distributions allows to create S2 signals below the "corrected S2" threshold. This leaves a certain probability, after the modelled cS2 signal is corrected back into S2 signal, that this signal is acceptance and therefore these signals are counted for each slice of S1 to evaluate the acceptance of S2 signals for a certain S1.
The reliability of the method is proven by simulated neutrons. The simulation takes the XENON100 TPC into account. The simulated neutrons were created under the science run condition of 2011/12 and compared to the both ^{241}AmBe neutron data sets. It turns out that the proper selection of single neutron scatter events in the simulated neutrons is quite easy in comparison to the data, where the number of scatters inside the TPC is not available. This event selection of single scatters for simulated neutrons results a higher S2 acceptance compared to the acceptance which is calculated for the both ^{241}AmBe data sets. Hence this outcome proves the method itself, but results just a kind of upper S2 acceptance limit, based on perfect selected single simulated neutrons scatters in the TPC.
In common for all three ^{241}AmBe data sets and the simulated neutrons, the uncertainty on the S2 acceptance is estimated. This becomes important for analysis at some point when the uncertainties of single cuts were taken into account during the Profile Likelihood analysis. The uncertainties for the S2 acceptance results from the choice of parameter β and λ. Therefore an additional variation around the expectation value λ is tested to estimate its uncertainty. Based on both uncertainties, a multidimensional 1σ level of significance is calculated and the resulting uncertainties are used in an extensive Monte Carlo procedure. This allows to calculate an uncertainty band around the already calculated S2 acceptance. In summary, the significance level yields to larger uncertainties on the S2 acceptance for lower values of S1. This is understood because these S1 energies will refer to lower S2 energy (well almost). The Poisson distribution description becomes more important for lower slices of S1 because its probability change of cS2 events below the "corrected S2" threshold. In this cases the probability increases for these events but also the uncertainties (which again are defined by β and λ).

Finally, another set of simulated WIMPs is used to determine a WIMP acceptance. As long as it is not important to prove the reliability of the chosen method to determine the Xs2peaks cut from data, the WIMP acceptance is evaluated by a simple counting experiment. Different WIMP masses up to 1000 GeV are simulated and the it is tested how much simulated WIMPs exceeds the threshold in S2 of 150 pe. The WIMP acceptance in XENON100, based on simulations yields to highest acceptances for WIMPs between 51 GeV and 100 GeV.

The previous two data analysis procedures use the radioactive calibration sources. Another important procedure is the weekly PMT calibration. Here the gain is determined by blue LEDs which distribute their light homogeneously in the XENON100 TPC. Due to the LED trigger, the DAQ records the LED pulse right after it is sent out. A usual LED calibration is done by taking the full LED waveform (for each PMT) without zero length encoding and analyse that by another software separately. This software is not used when the calibration data or Dark Matter data are reprocessed (*xerawdp*). By default, *xerawdp* is not able to reprocess LED data directly and to determine the PMT gain. This work will introduce a method to use *xerawdp* for the LED data reprocessing. After software adjustments which are necessary in *xerawdp* and a specific cut set which is developed only for LED data, first results of the identified S1 peaks are presented. To verify the outcome and the quality of the identified S1 peaks, the single photo- electron peaks are used to predict the single PMT response to the single LED photons. The results show a good agreement with the expected SPE response at $S1 = \sim 1$ pe.
Nevertheless, to calculate the SPE response it is necessary to use the gain values from the standard PMT calibration when *xerawdp* is reprocessing LED data. After the SPE response is shown, the next approach is to calculate the gain directly out of the identified S1 peaks. Therefore the raw S1 area is used. This variable, which is defined in *xerawdp*, measures the area of a S1 peak between its boundaries. In that way, the gain is calculated from histograms which counts the raw area of the S1 peaks. Hence, the gain is the most likely value of the signal part which is described by a Poisson distribution. By comparing these alternative gains (PMT by PMT) to the gains from the LED calibration, it turns out that both measurements are match within a 68% level.
The here described method have not been used before in XENON100. LED calibration and data reprocessing is done by two different pieces of software. With that new approach it could be possible to use only one piece of software for LED calibration in the future. Due to the end of the XENON100 experiment, further tests with the software code of *xerawdp* are not done any more. Nevertheless, the successor experiment XENON1T will use a piece of software which will be used for all possible data.

7.2. Outlook on XENON1T

The here presented analysis topics are applications which are necessary and useful for the future XENON1T experiment. At that time, the XENON1T experiment is mounted in the *Hall B* of the underground laboratory "Laboratori Nazionali del Gran Sasso" (LNGS). The basic technology is not changes: There is still a TPC used which is filled with liquid XENON. The improvements coming from the physical constrains and hence, a larger target volume would increase the interaction probability. The new TPC has a size of 1 m in height and 1 m in diameter and will there will be up to 3.3 tons of liquid xenon are in use. The target mass will be then around 2 tons which allows to have a fiducial volume of 1 ton. The shelf-shielding properties which were already exploited in XENON100 to reduce γ introduced events will here also help to reduce the background. All materials of the TPC and for the cryostat are screened and chosen for the purpose of low-radioactivity. With this it is expected to have a 100 times lower background than in XENON100.

Figure 7.1 – Picture from the XENON1T experiment in the hall B of the LNGS. The foreground shows the support building where the ReStoX, xenon purification system and further electronic is mounted. In the background, there is the XENON1T water tank shown [44].

Another big advantage, compared to XENON100, is the active muon veto. This is realised by a water tank of 10 m height and 10 m diameter. There are 84 PMTs distributed inside the water tank which looking for Cerenkov light. This helps to select events from atmospheric muons. These events can fake a WIMP signal in the TPC. The active muon veto allows to select events which are in coincidence with events inside the TPC. This type of event is rejected.

The enlarged TPC, together with the active muon veto allows to exploit the WIMP search region. Even in case that there is no WIMP found, it push the actual exclusion curve by a factor of 2 down to lower limit. The water tank and the PMTs of the active muon veto are already mounted in the experimental hall (see background structure in figure 7.1). Inside of the water tank, the cryostat is mounted and first tests are done 7.2. The TPC is not yet mounted. In addition, a support building is right besides the water tank which provides enough space for the xenon recovery and storage facility

(ReStoX). This is able to hold up to 7 tons of liquid or gaseous xenon. Furthermore the ReStoX is used in case of an emergency recuperation. Krypton (^{85}Kr) contamination causes due to its beta decays an addition background and therefore it is distilled out in a Kr-column which is also mounted in the support building. The xenon purification system and further electronic devices (e.g. DAQ) are also mounted there. The support building (foreground of figure 7.1) itself is already mounted together with the xenon purification system, Kr-column and ReStoX. After the TPC is mounted into the cryostat, the XENON1T experiment will start to take data. First data are expected until the end of 2015.

The XENON100 experiment itself is still running and used for further tests with e.g. new calibration sources to test the understanding of XENON100 detector. At least, with the start of the XENON1T data taking the XENON100 experiment will come to an end.

Figure 7.2 – A view inside the active muon veto of XENON1T: The cryostat in where the TPC will be mounted is already installed. The scaffold on the site of the walls will be removed as well as the black protection layers of the veto PMTs [44].

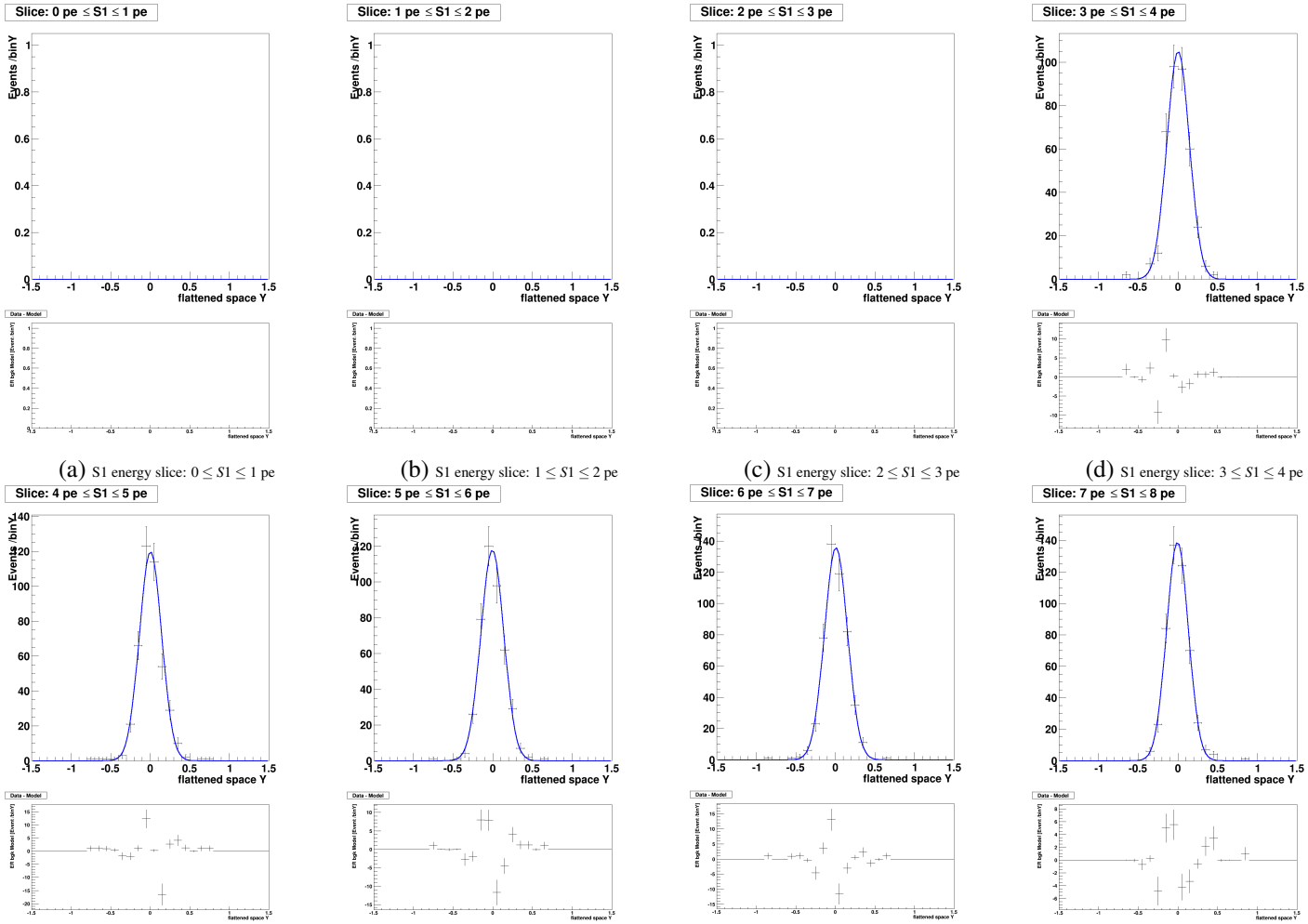
A. Appendix

"*Dark Matter* density is nearly one order of magnitude higher than in similar nebulae. Life-forms here may have developed in ways never before observed."

Lieutenant Commander Data,
Star Trek: The Next Generation, "In Theory"

A.1. ER Background model

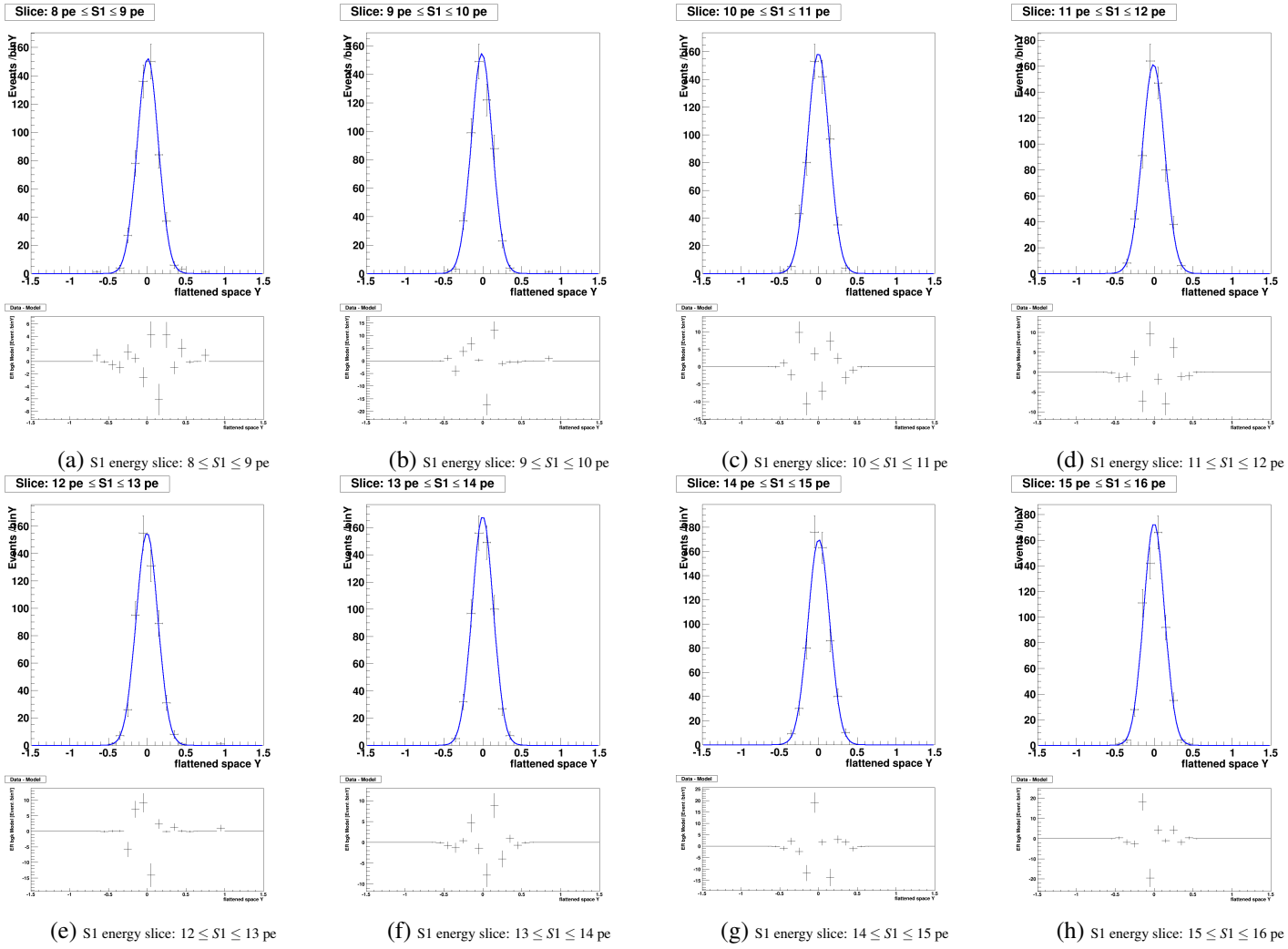
A.1.1. Data Run10



(a) S1 energy slice: $0 \leq S1 \leq 1$ pe (b) S1 energy slice: $1 \leq S1 \leq 2$ pe (c) S1 energy slice: $2 \leq S1 \leq 3$ pe (d) S1 energy slice: $3 \leq S1 \leq 4$ pe

(e) S1 energy slice: $4 \leq S1 \leq 5$ pe (f) S1 energy slice: $5 \leq S1 \leq 6$ pe (g) S1 energy slice: $6 \leq S1 \leq 7$ pe (h) S1 energy slice: $7 \leq S1 \leq 8$ pe

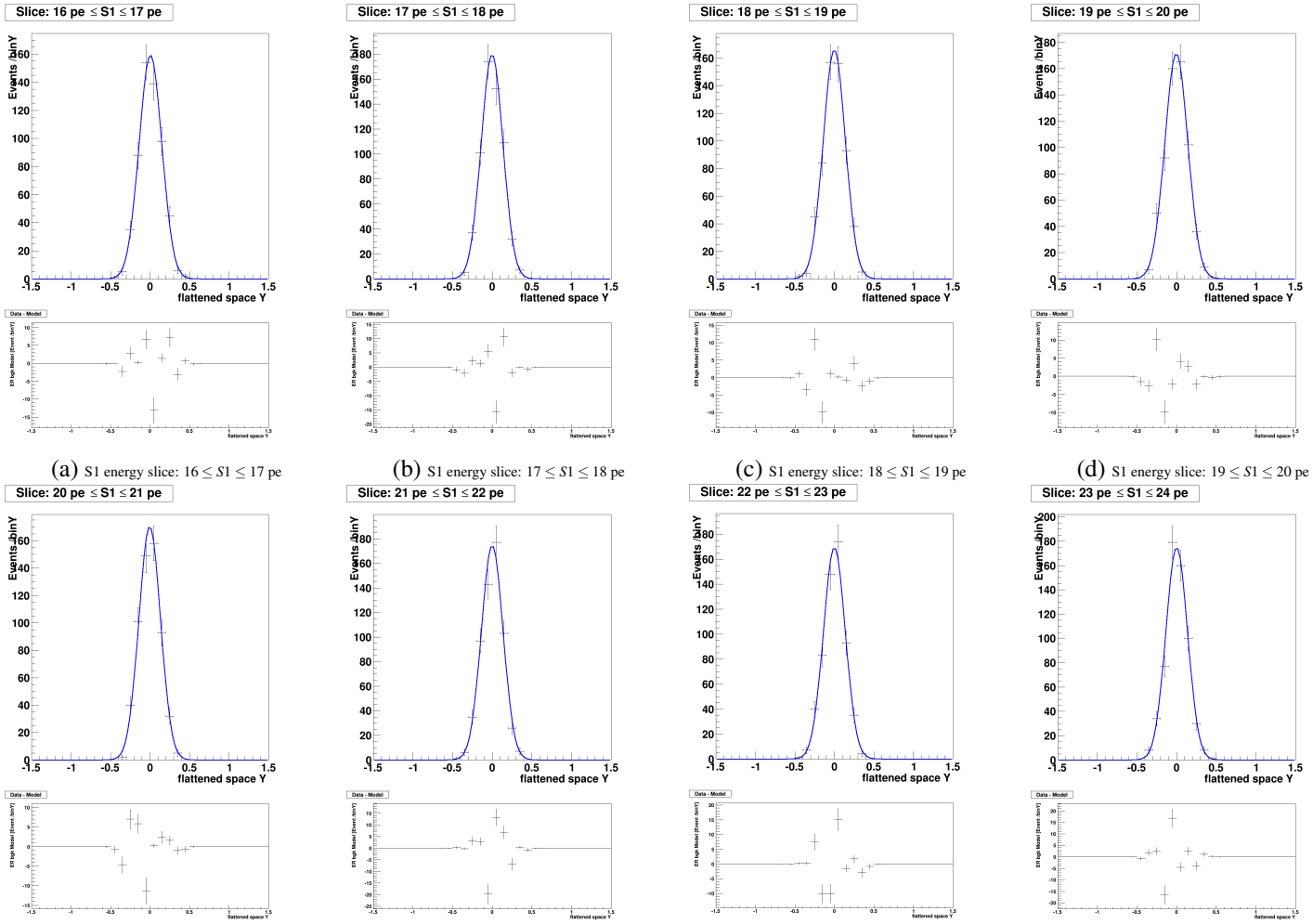
Figure A.1.: Data: 2011/12– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"



(a) S1 energy slice: $8 \leq S1 \leq 9$ pe (b) S1 energy slice: $9 \leq S1 \leq 10$ pe (c) S1 energy slice: $10 \leq S1 \leq 11$ pe (d) S1 energy slice: $11 \leq S1 \leq 12$ pe

(e) S1 energy slice: $12 \leq S1 \leq 13$ pe (f) S1 energy slice: $13 \leq S1 \leq 14$ pe (g) S1 energy slice: $14 \leq S1 \leq 15$ pe (h) S1 energy slice: $15 \leq S1 \leq 16$ pe

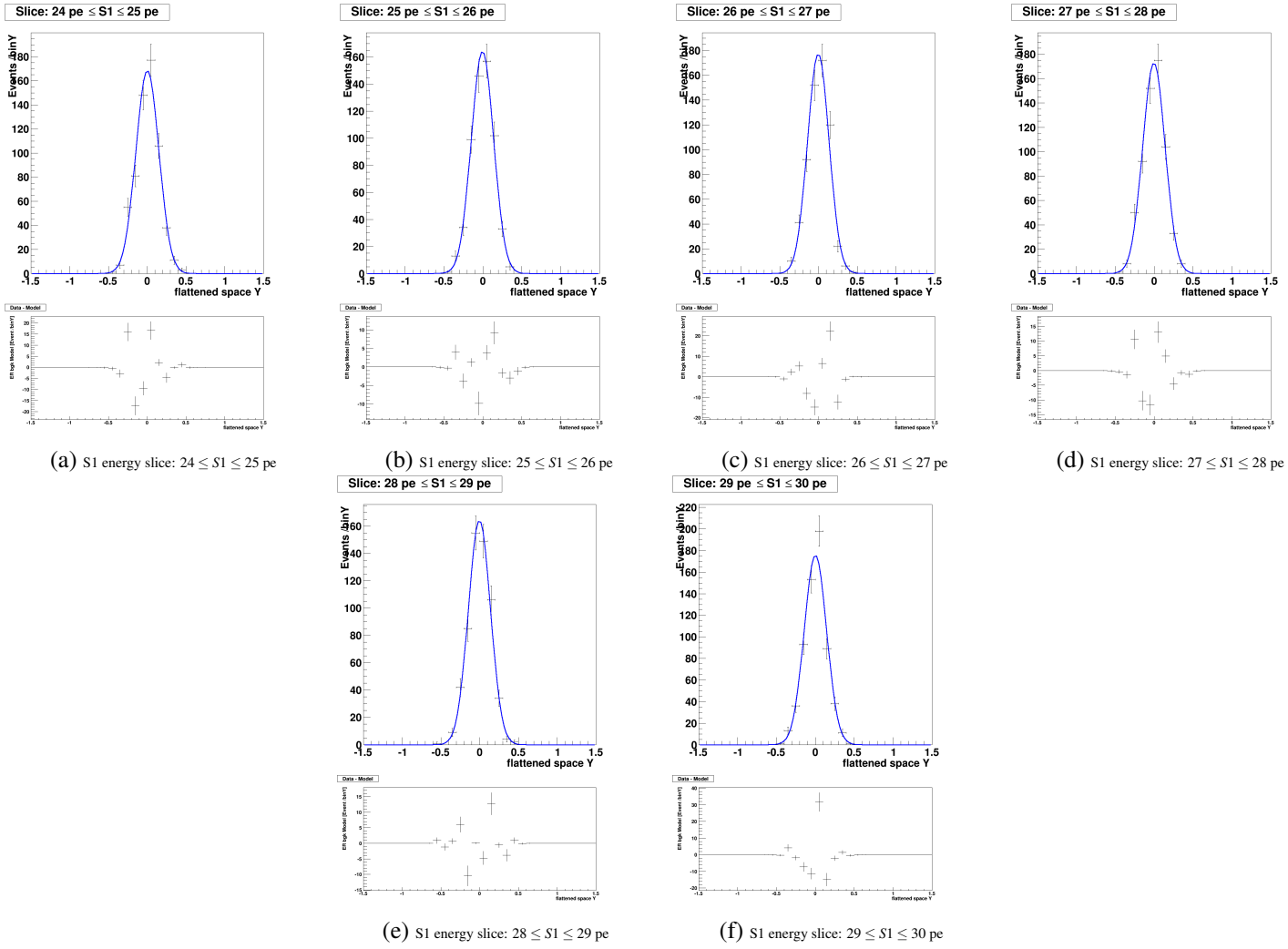
Figure A.2.: Data: 2011/12– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"



(a) S1 energy slice: $16 \leq S1 \leq 17$ pe (b) S1 energy slice: $17 \leq S1 \leq 18$ pe (c) S1 energy slice: $18 \leq S1 \leq 19$ pe (d) S1 energy slice: $19 \leq S1 \leq 20$ pe

(e) S1 energy slice: $20 \leq S1 \leq 21$ pe (f) S1 energy slice: $21 \leq S1 \leq 22$ pe (g) S1 energy slice: $22 \leq S1 \leq 23$ pe (h) S1 energy slice: $23 \leq S1 \leq 24$ pe

Figure A.3.: Data: 2011/12– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: ”Data - Model“

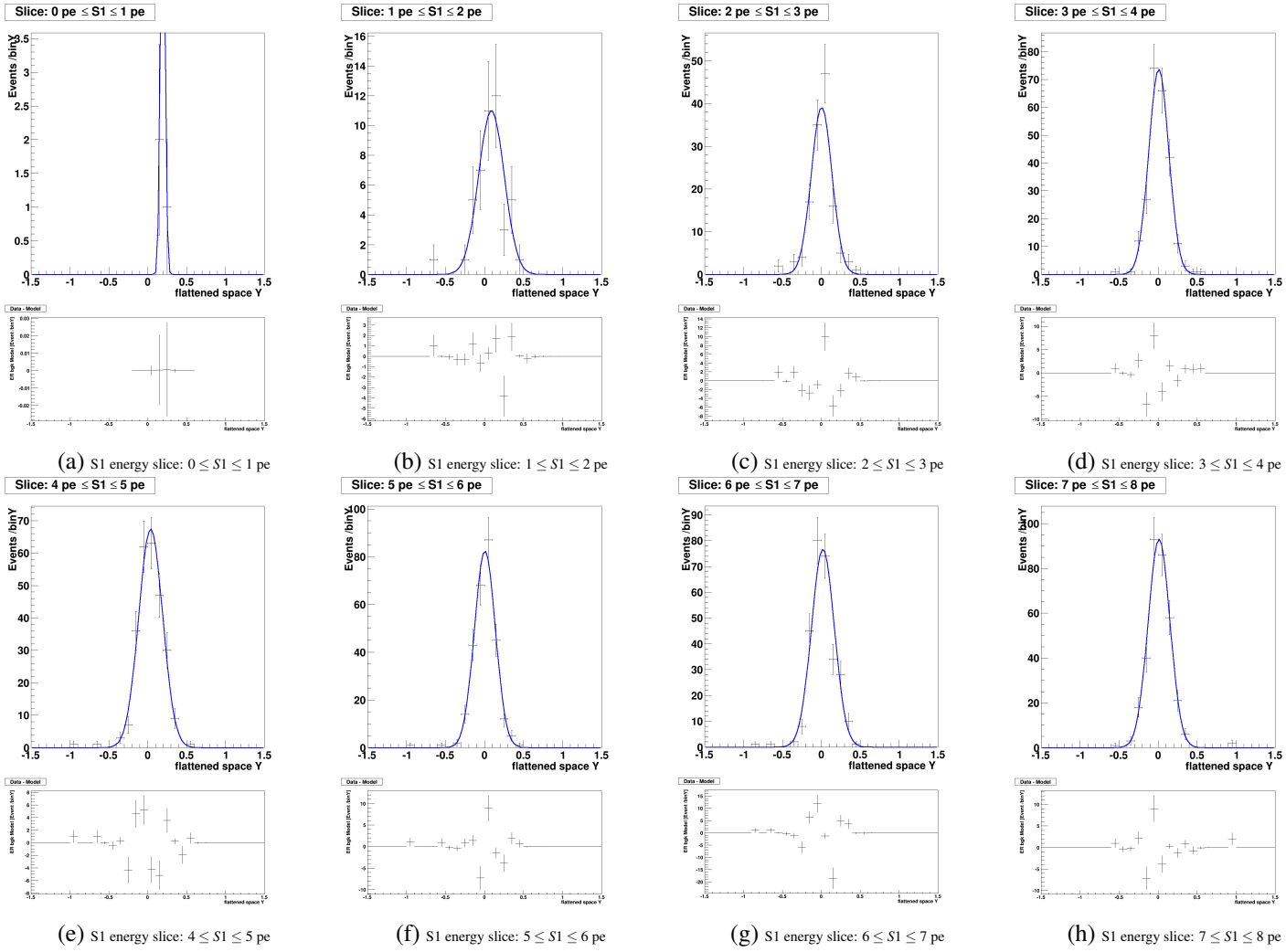


(a) S1 energy slice: $24 \leq S1 \leq 25$ pe (b) S1 energy slice: $25 \leq S1 \leq 26$ pe (c) S1 energy slice: $26 \leq S1 \leq 27$ pe (d) S1 energy slice: $27 \leq S1 \leq 28$ pe

(e) S1 energy slice: $28 \leq S1 \leq 29$ pe (f) S1 energy slice: $29 \leq S1 \leq 30$ pe

Figure A.4.: Data: 2011/12– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"

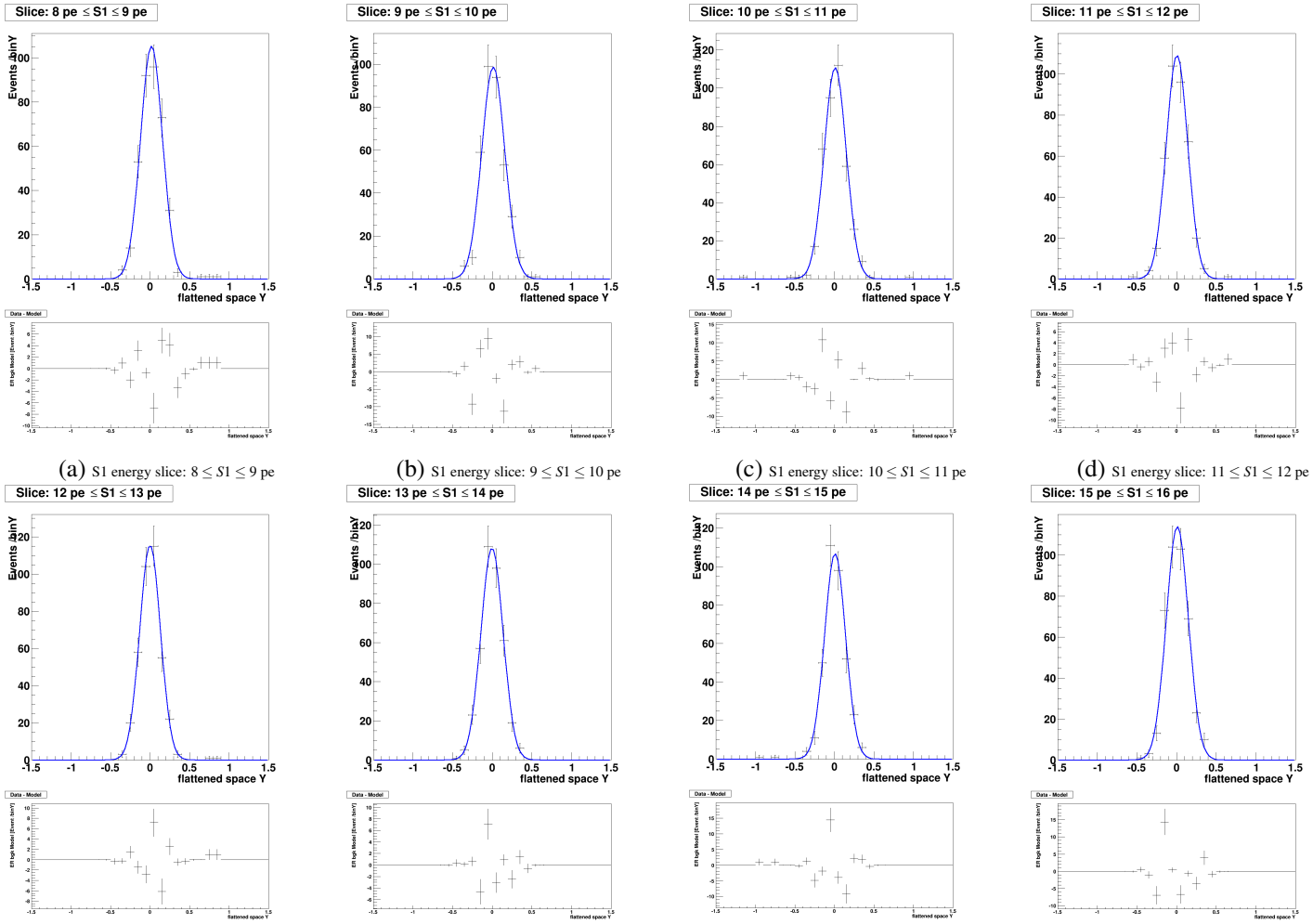
A.1.2. Data Run12



(a) S1 energy slice: $0 \leq S1 \leq 1$ pe (b) S1 energy slice: $1 \leq S1 \leq 2$ pe (c) S1 energy slice: $2 \leq S1 \leq 3$ pe (d) S1 energy slice: $3 \leq S1 \leq 4$ pe

(e) S1 energy slice: $4 \leq S1 \leq 5$ pe (f) S1 energy slice: $5 \leq S1 \leq 6$ pe (g) S1 energy slice: $6 \leq S1 \leq 7$ pe (h) S1 energy slice: $7 \leq S1 \leq 8$ pe

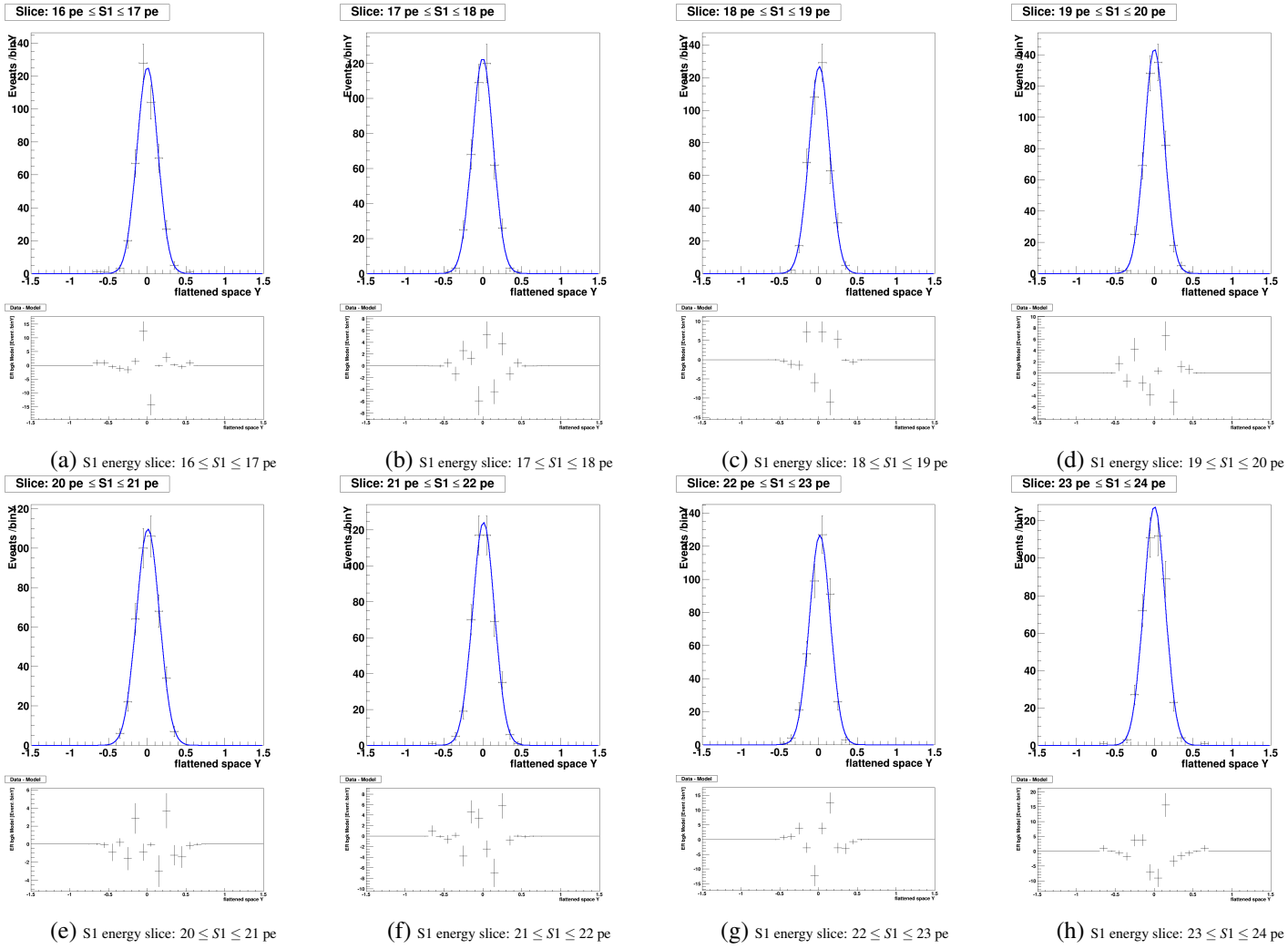
Figure A.5.: Data: 2013/14– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"



(a) S1 energy slice: $8 \leq S1 \leq 9$ pe (b) S1 energy slice: $9 \leq S1 \leq 10$ pe (c) S1 energy slice: $10 \leq S1 \leq 11$ pe (d) S1 energy slice: $11 \leq S1 \leq 12$ pe

(e) S1 energy slice: $12 \leq S1 \leq 13$ pe (f) S1 energy slice: $13 \leq S1 \leq 14$ pe (g) S1 energy slice: $14 \leq S1 \leq 15$ pe (h) S1 energy slice: $15 \leq S1 \leq 16$ pe

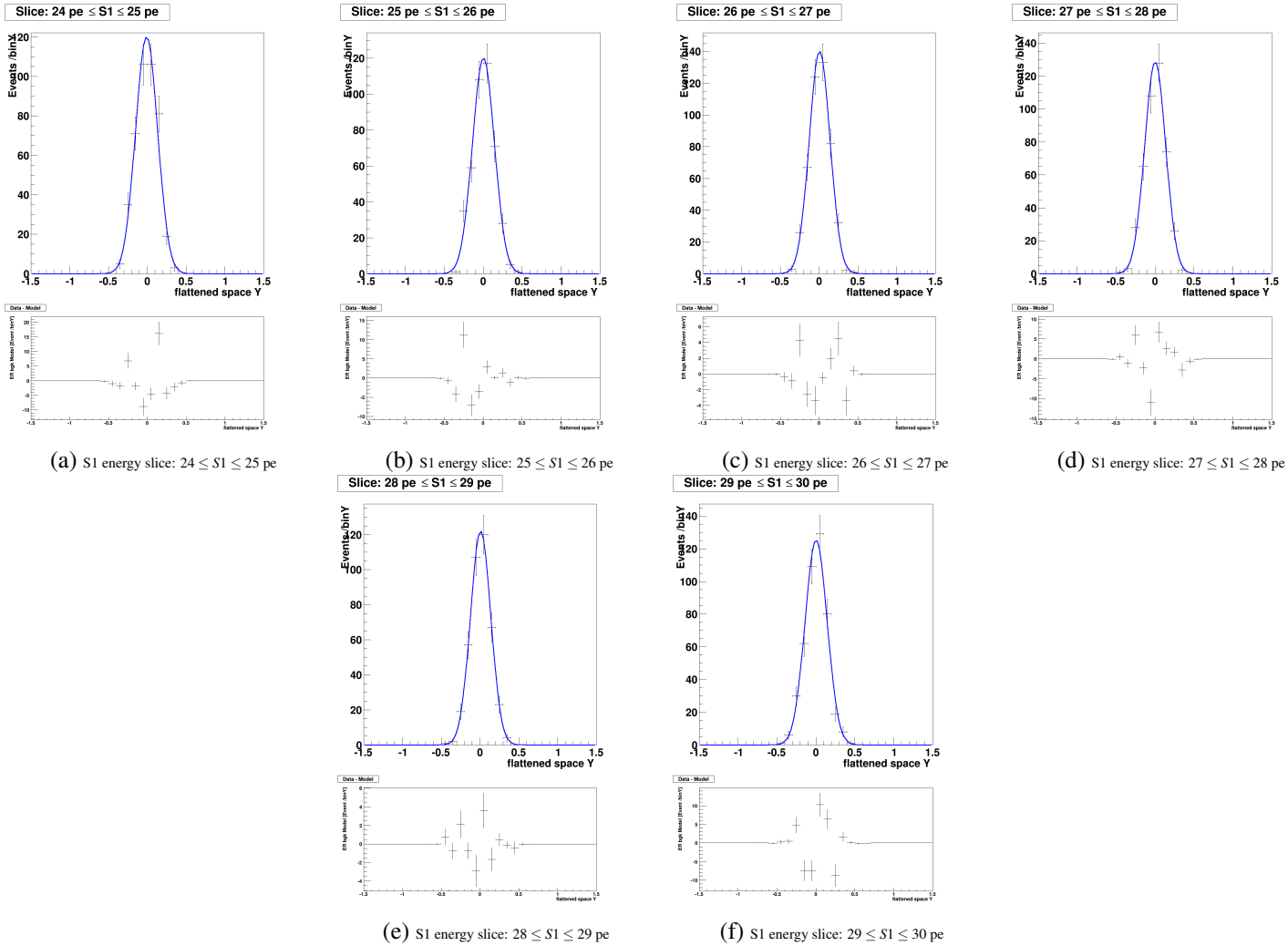
Figure A.6.: Data: 2013/14– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"



(a) S1 energy slice: $16 \leq S1 \leq 17$ pe (b) S1 energy slice: $17 \leq S1 \leq 18$ pe (c) S1 energy slice: $18 \leq S1 \leq 19$ pe (d) S1 energy slice: $19 \leq S1 \leq 20$ pe

(e) S1 energy slice: $20 \leq S1 \leq 21$ pe (f) S1 energy slice: $21 \leq S1 \leq 22$ pe (g) S1 energy slice: $22 \leq S1 \leq 23$ pe (h) S1 energy slice: $23 \leq S1 \leq 24$ pe

Figure A.7.: Data: 2013/14– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"



(a) S1 energy slice: $24 \leq S1 \leq 25$ pe

(b) S1 energy slice: $25 \leq S1 \leq 26$ pe

(c) S1 energy slice: $26 \leq S1 \leq 27$ pe

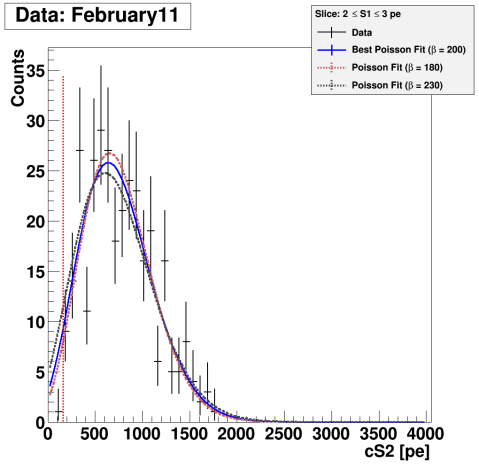
(d) S1 energy slice: $27 \leq S1 \leq 28$ pe

(e) S1 energy slice: $28 \leq S1 \leq 29$ pe

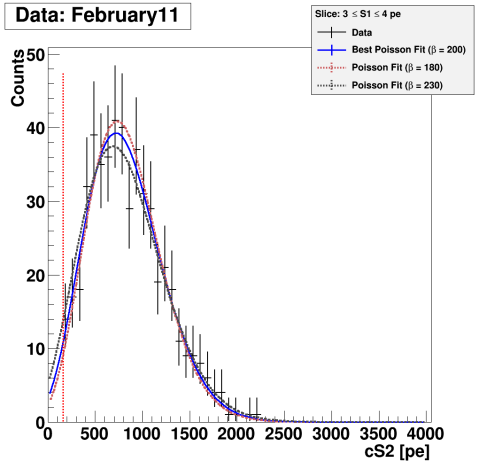
(f) S1 energy slice: $29 \leq S1 \leq 30$ pe

Figure A.8.: Data: 2013/14– Single S1 energy slices of the ER background. The blue line show s the additional Gaussian approximation. The residuum is plotted below each data distribution. Its defined by: "Data - Model"

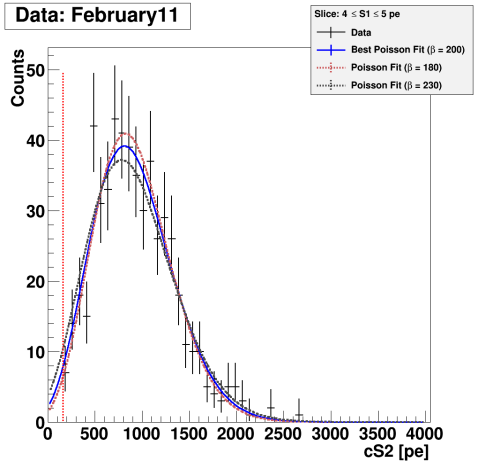
A.2. Evaluate the acceptance of the Xs2peaks cut



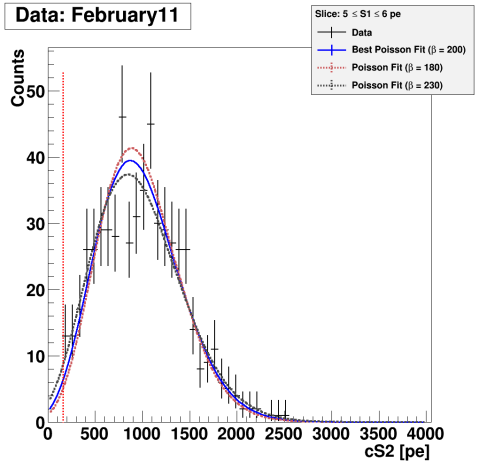
(a) S1 energy slice: $1 \leq S1 \leq 2$ pe



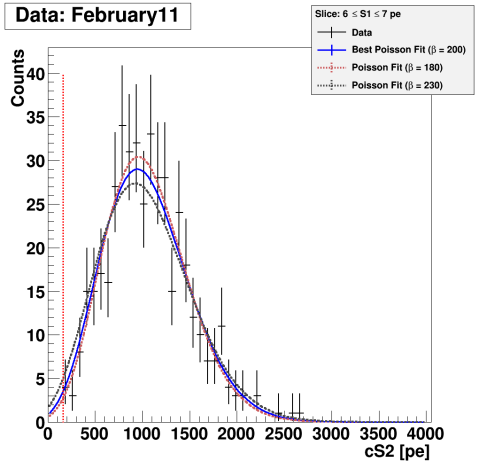
(b) S1 energy slice: $2 \leq S1 \leq 3$ pe



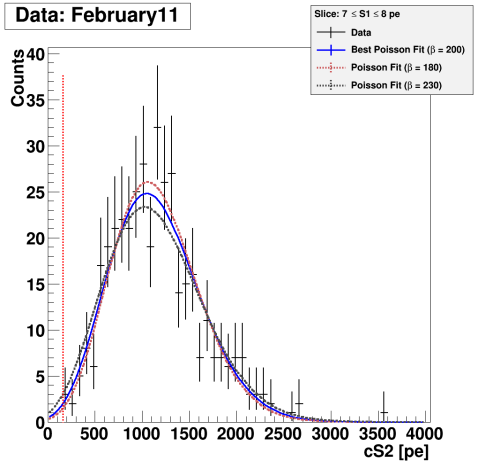
(c) S1 energy slice: $3 \leq S1 \leq 4$ pe



(d) S1 energy slice: $4 \leq S1 \leq 5$ pe

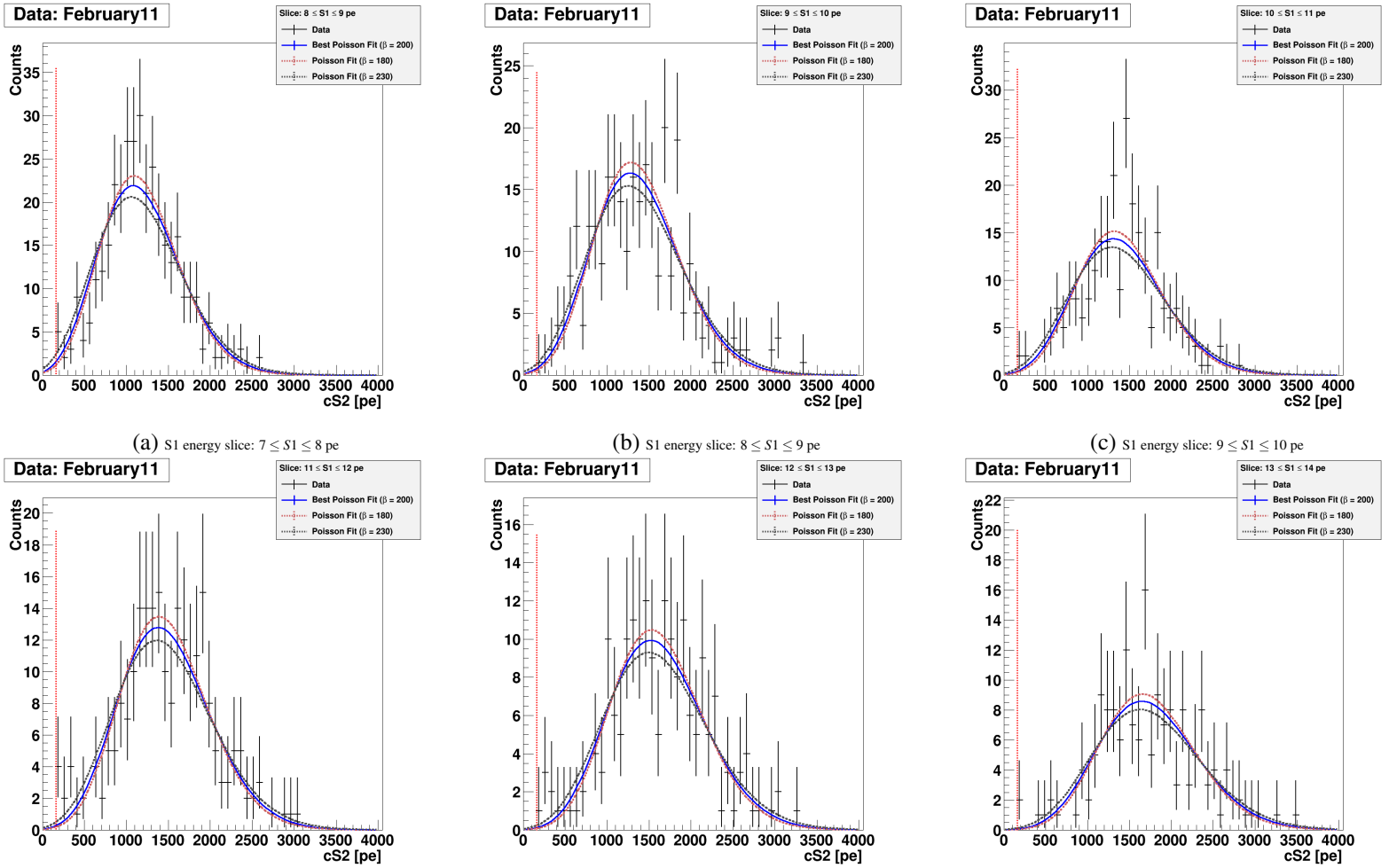


(e) S1 energy slice: $5 \leq S1 \leq 6$ pe



(f) S1 energy slice: $6 \leq S1 \leq 7$ pe

Figure A.9.: ^{241}AmBe data: 2011/12 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $163 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $7 \leq S1 \leq 8$ pe

(b) S1 energy slice: $8 \leq S1 \leq 9$ pe

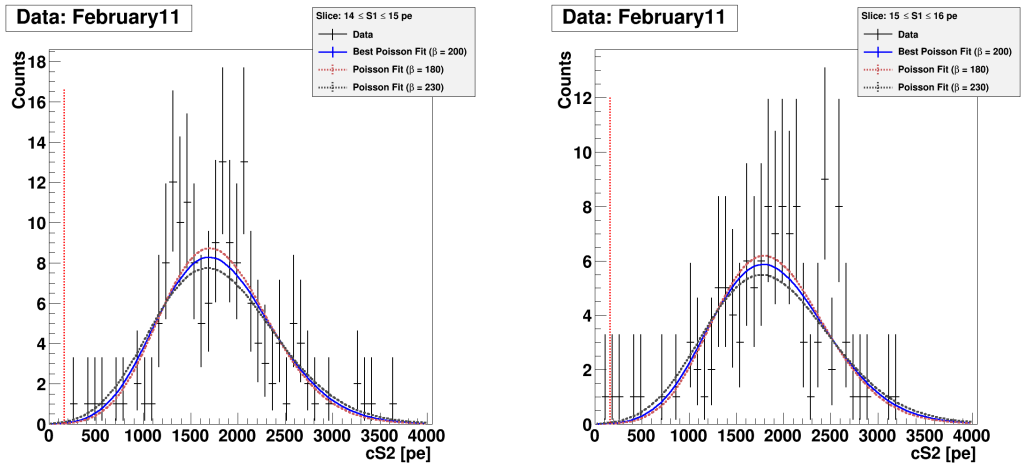
(c) S1 energy slice: $9 \leq S1 \leq 10$ pe

(d) S1 energy slice: $10 \leq S1 \leq 11$ pe

(e) S1 energy slice: $11 \leq S1 \leq 12$ pe

(f) S1 energy slice: $12 \leq S1 \leq 13$ pe

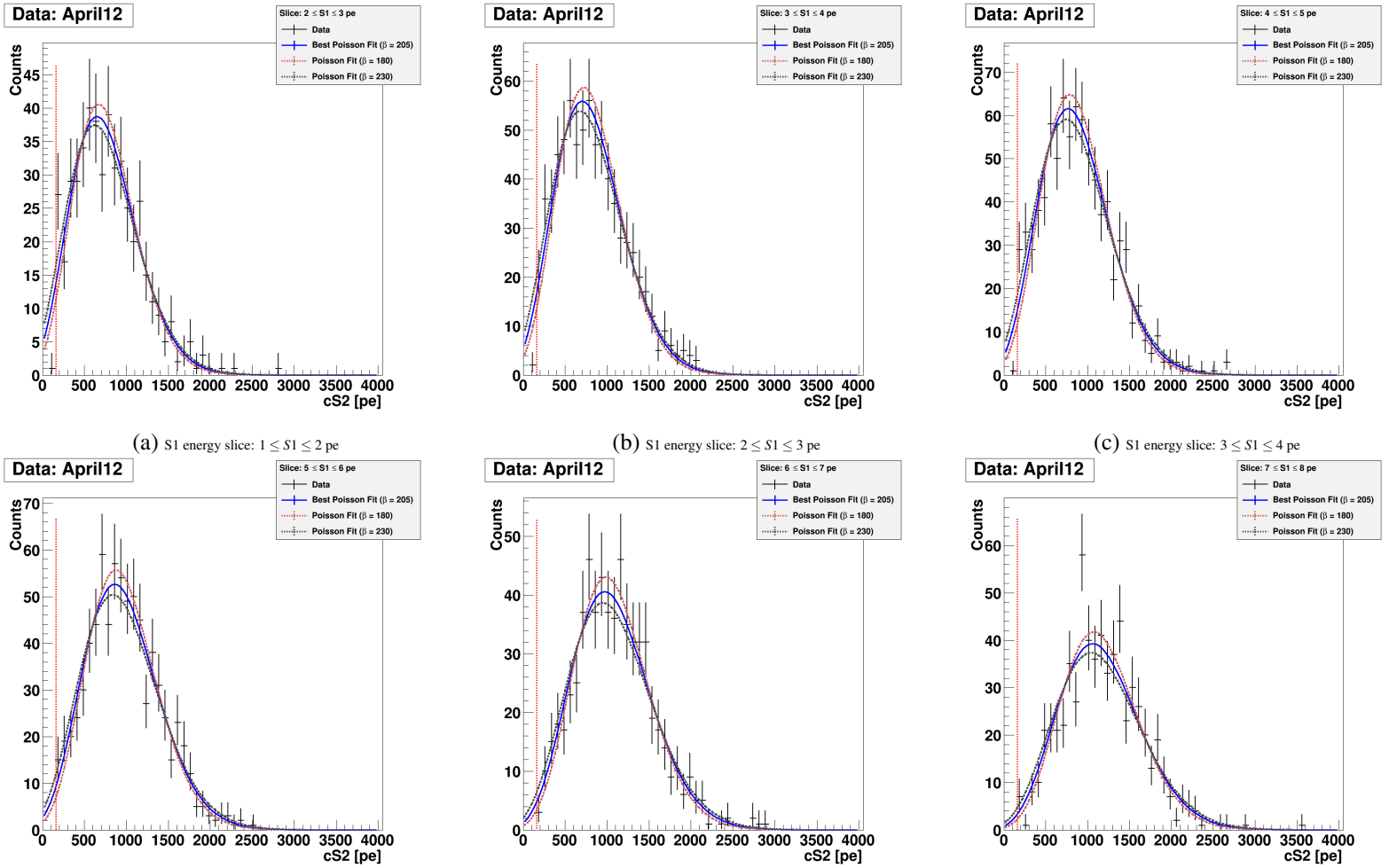
Figure A.10.: ^{241}AmBe data: 2011/12/February11 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $163 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $13 \leq S1 \leq 14$ pe

(b) S1 energy slice: $14 \leq S1 \leq 15$ pe

Figure A.11.: ^{241}AmBe data: 2011/12/February11 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $163 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $1 \leq S1 \leq 2$ pe

(b) S1 energy slice: $2 \leq S1 \leq 3$ pe

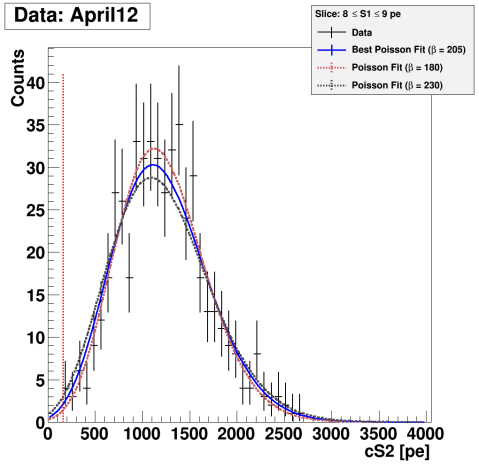
(c) S1 energy slice: $3 \leq S1 \leq 4$ pe

(d) S1 energy slice: $4 \leq S1 \leq 5$ pe

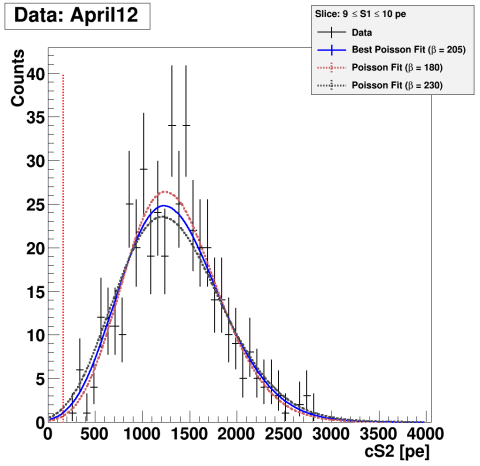
(e) S1 energy slice: $5 \leq S1 \leq 6$ pe

(f) S1 energy slice: $6 \leq S1 \leq 7$ pe

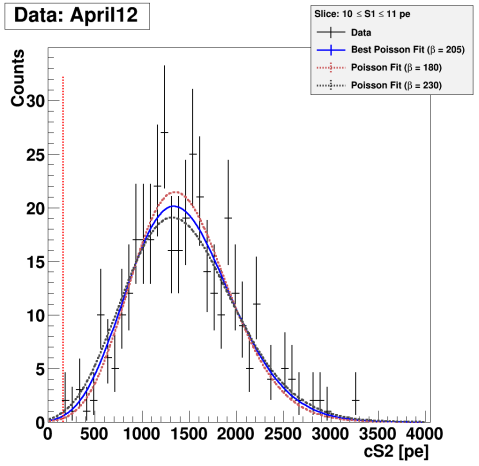
Figure A.12.: ^{241}AmBe data: 2011/12/April12 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $163 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



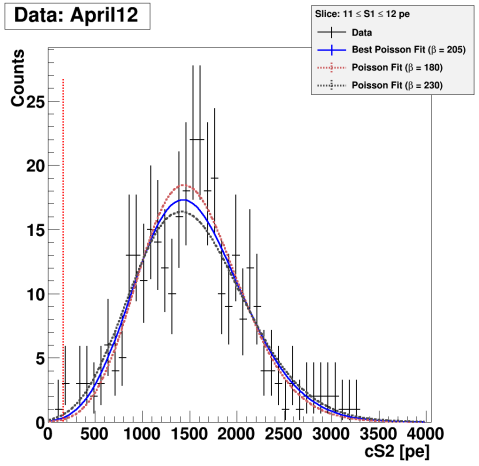
(a) S1 energy slice: $7 \leq S1 \leq 8$ pe



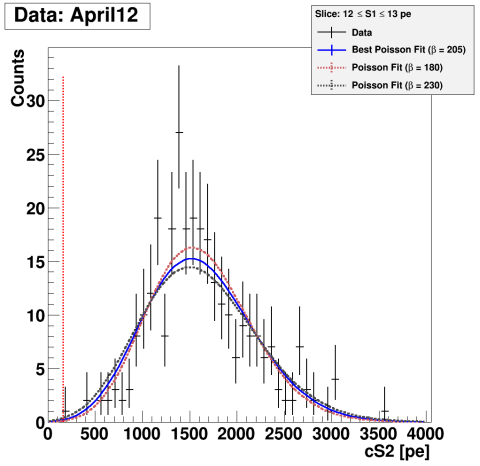
(b) S1 energy slice: $8 \leq S1 \leq 9$ pe



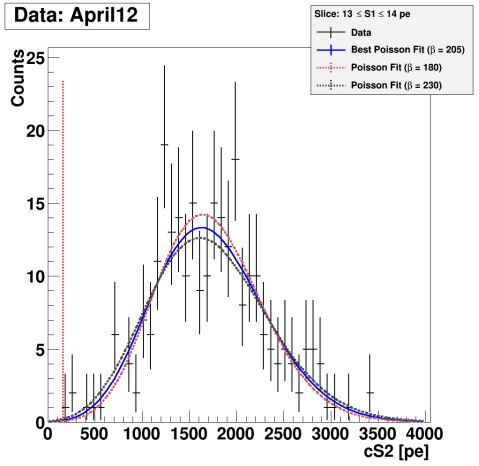
(c) S1 energy slice: $9 \leq S1 \leq 10$ pe



(d) S1 energy slice: $10 \leq S1 \leq 11$ pe

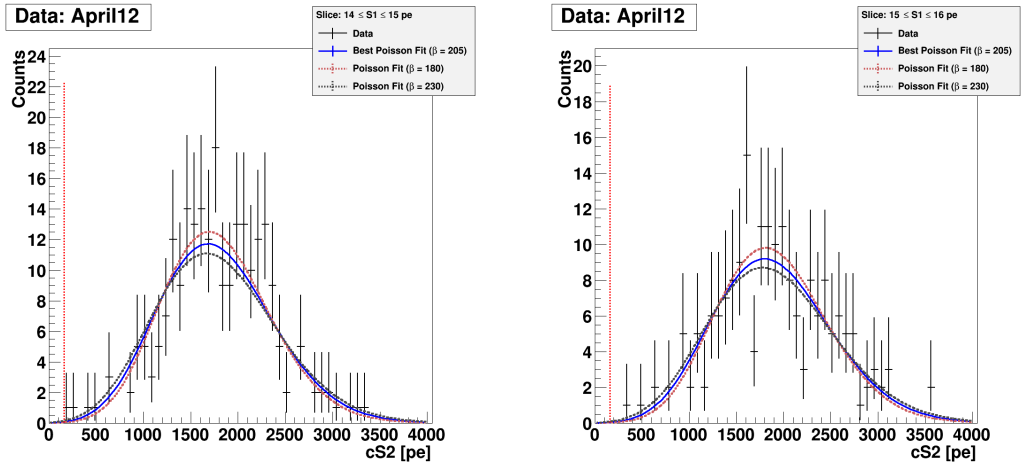


(e) S1 energy slice: $11 \leq S1 \leq 12$ pe



(f) S1 energy slice: $12 \leq S1 \leq 13$ pe

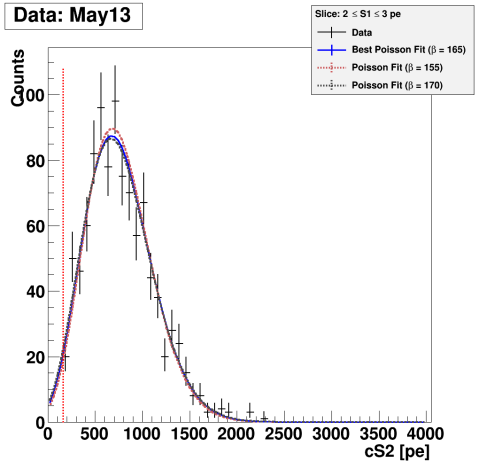
Figure A.13.: ^{241}AmBe data: 2011/12/April12 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $163 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



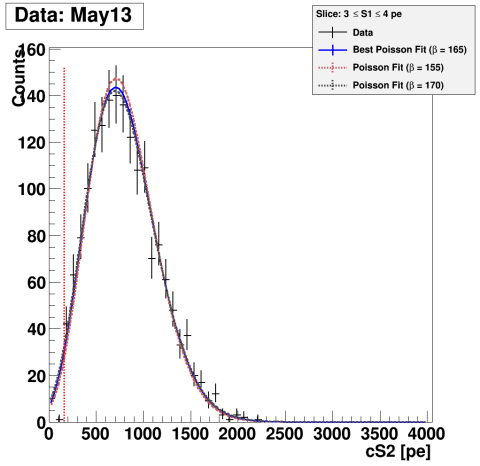
(a) S1 energy slice: $13 \leq S1 \leq 14$ pe

(b) S1 energy slice: $14 \leq S1 \leq 15$ pe

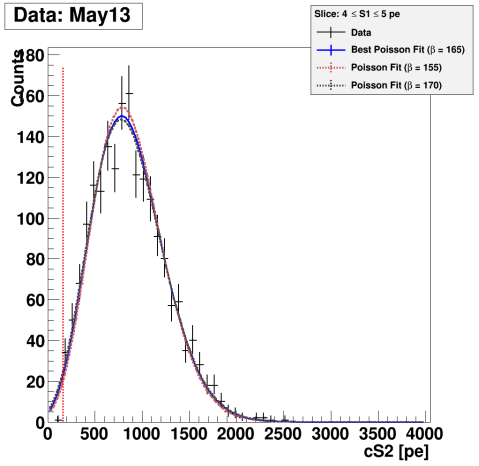
Figure A.14.: ^{241}AmBe data: 2011/12/April12 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $163 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



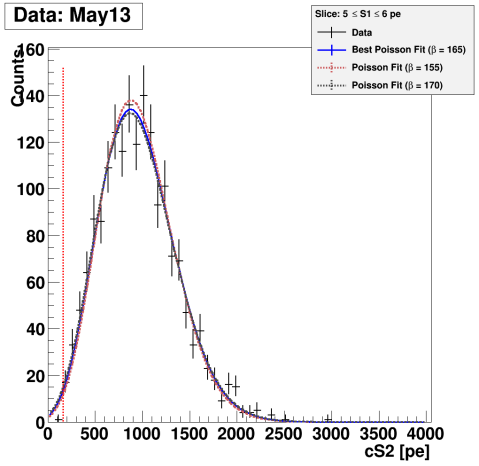
(a) S1 energy slice: $1 \leq S1 \leq 2$ pe



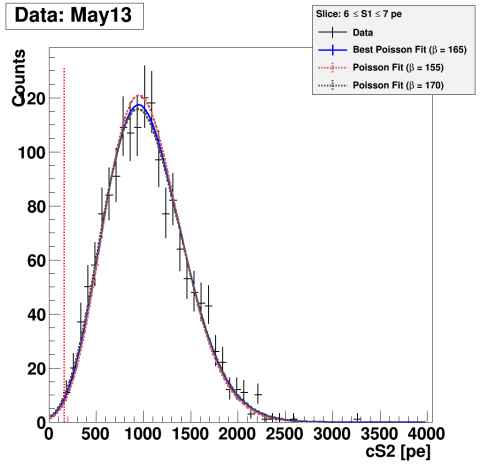
(b) S1 energy slice: $2 \leq S1 \leq 3$ pe



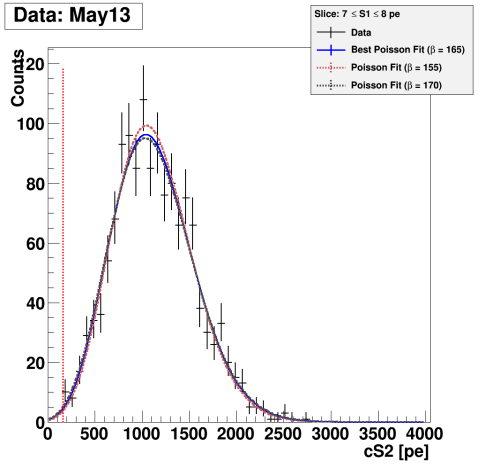
(c) S1 energy slice: $3 \leq S1 \leq 4$ pe



(d) S1 energy slice: $4 \leq S1 \leq 5$ pe

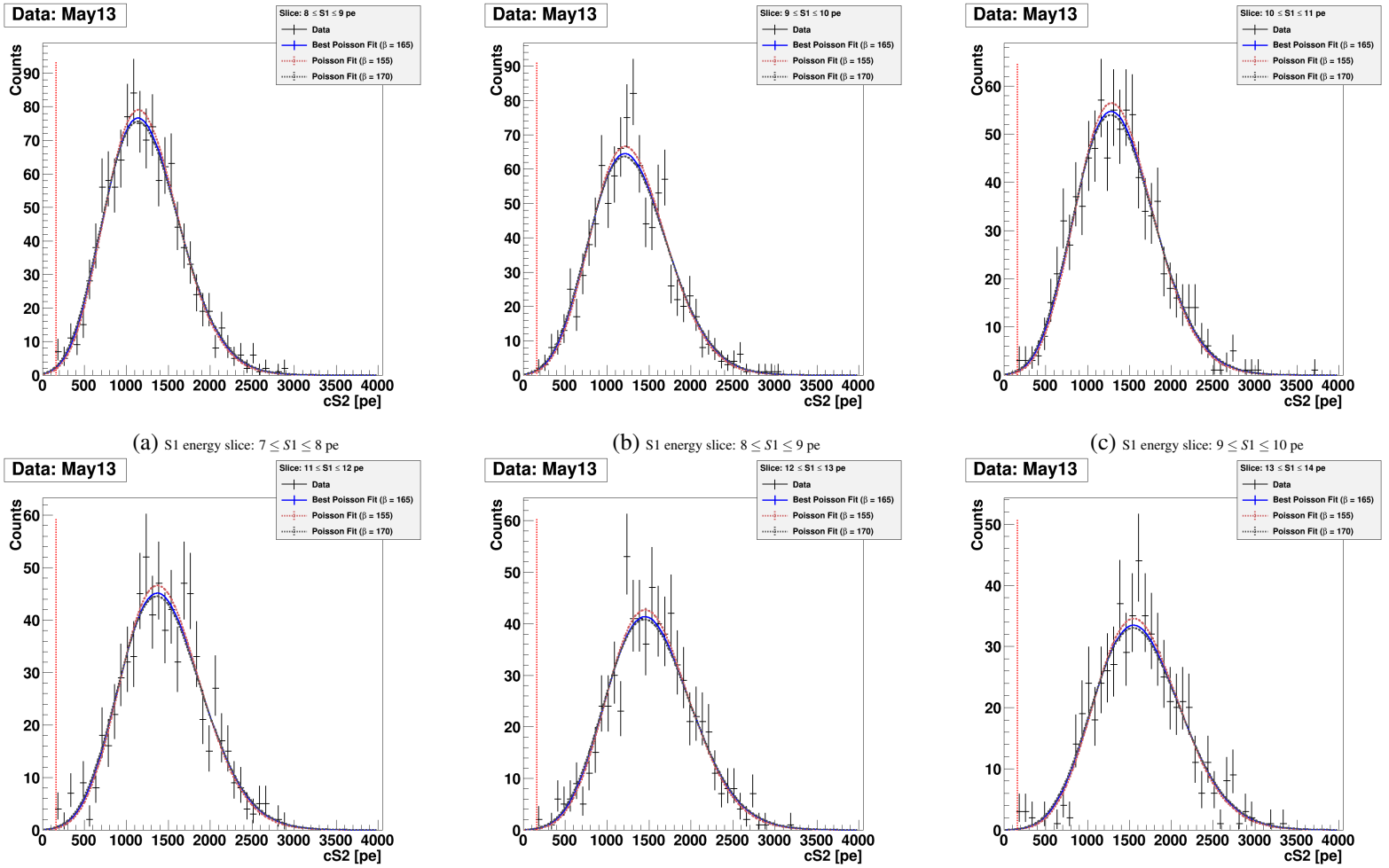


(e) S1 energy slice: $5 \leq S1 \leq 6$ pe



(f) S1 energy slice: $6 \leq S1 \leq 7$ pe

Figure A.15.: ^{241}AmBe data: 2013/14/May13 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $159 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $7 \leq S1 \leq 8$ pe

(b) S1 energy slice: $8 \leq S1 \leq 9$ pe

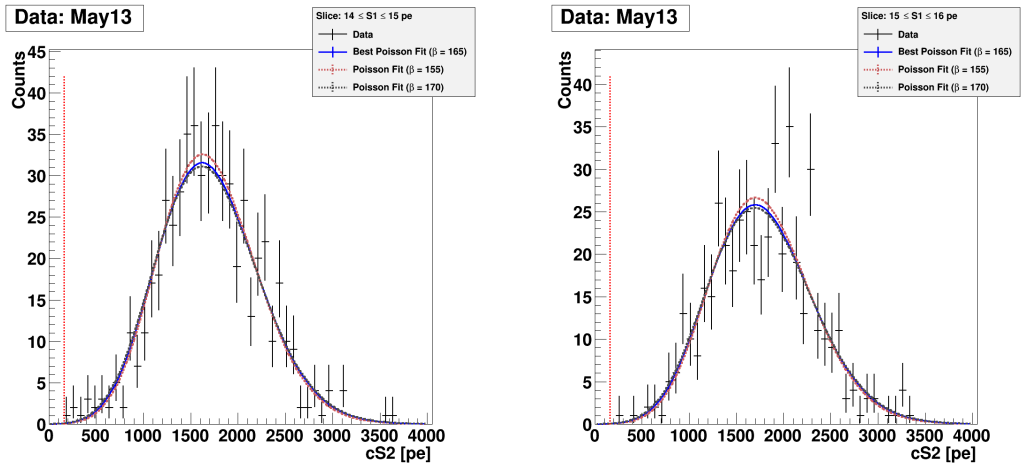
(c) S1 energy slice: $9 \leq S1 \leq 10$ pe

(d) S1 energy slice: $10 \leq S1 \leq 11$ pe

(e) S1 energy slice: $11 \leq S1 \leq 12$ pe

(f) S1 energy slice: $12 \leq S1 \leq 13$ pe

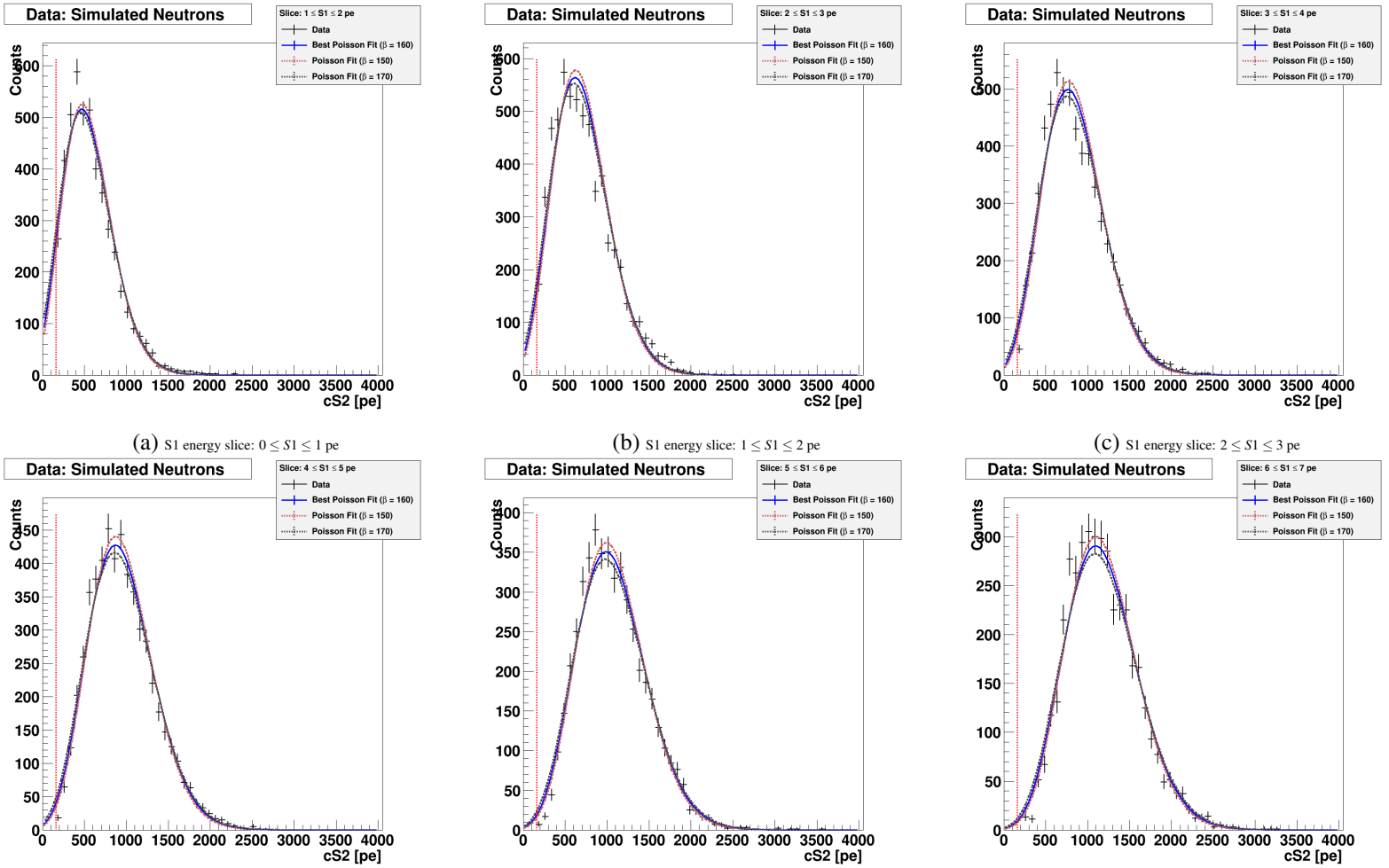
Figure A.16.: ^{241}AmBe data: 2013/14/May13 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $159 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $13 \leq S1 \leq 14$ pe

(b) S1 energy slice: $14 \leq S1 \leq 15$ pe

Figure A.17.: ^{241}AmBe data: 2013/14/May13 – Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $159 \leq S2 \leq 5000$ pe are plotted in blue. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $0 \leq S1 \leq 1$ pe

(b) S1 energy slice: $1 \leq S1 \leq 2$ pe

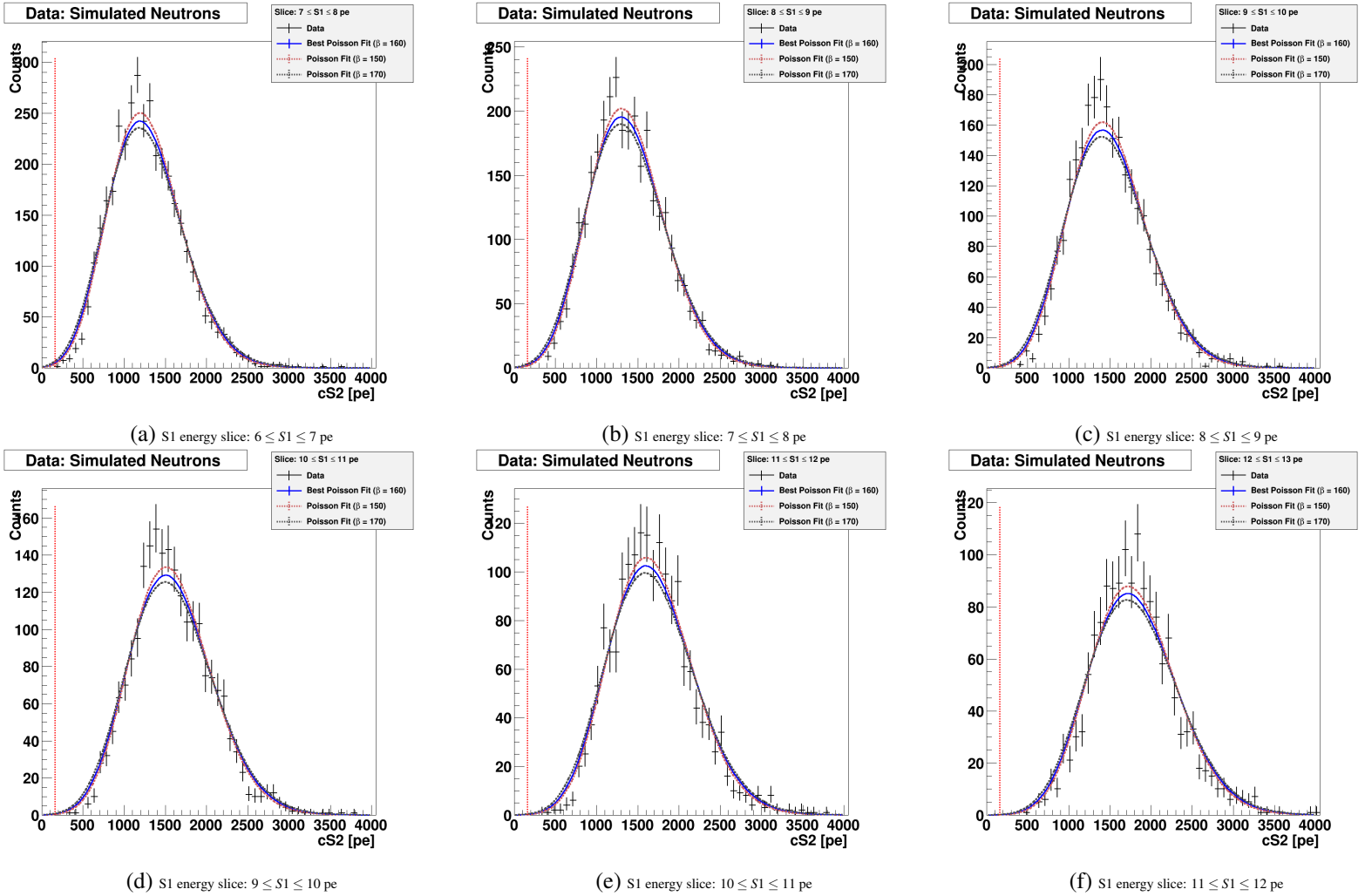
(c) S1 energy slice: $2 \leq S1 \leq 3$ pe

(d) S1 energy slice: $3 \leq S1 \leq 4$ pe

(e) S1 energy slice: $4 \leq S1 \leq 5$ pe

(f) S1 energy slice: $5 \leq S1 \leq 6$ pe

Figure A.18.: Simulated neutron data with *Nuclear Recoil/WIMP Simulation Code* with 2011/12 configuration. Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $159 \leq S2 \leq 5000$ pe are plotted in black. The left fit boundary is calculated out of the electron life time correction for S2 signals.



(a) S1 energy slice: $6 \leq S1 \leq 7$ pe

(b) S1 energy slice: $7 \leq S1 \leq 8$ pe

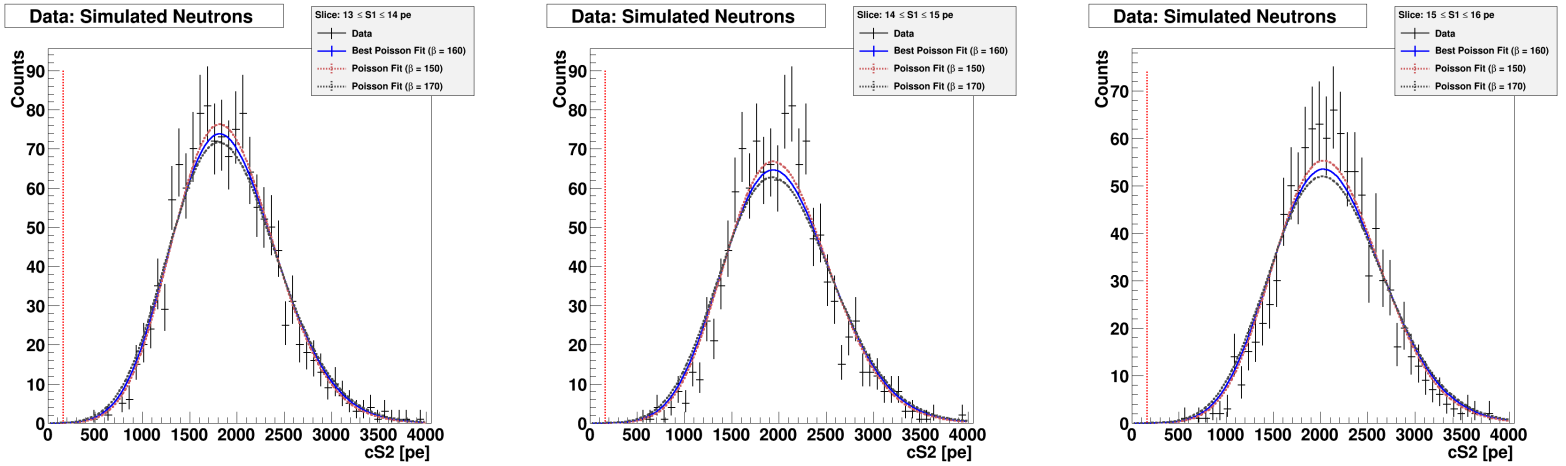
(c) S1 energy slice: $8 \leq S1 \leq 9$ pe

(d) S1 energy slice: $9 \leq S1 \leq 10$ pe

(e) S1 energy slice: $10 \leq S1 \leq 11$ pe

(f) S1 energy slice: $11 \leq S1 \leq 12$ pe

Figure A.19.: Simulated neutron data with *Nuclear Recoil/WIMP Simulation Code* with 2011/12 configuration. Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $159 \leq S2 \leq 5000$ pe are plotted in black. The left fit boundary is calculated out of the electron life time correction for S2 signals.



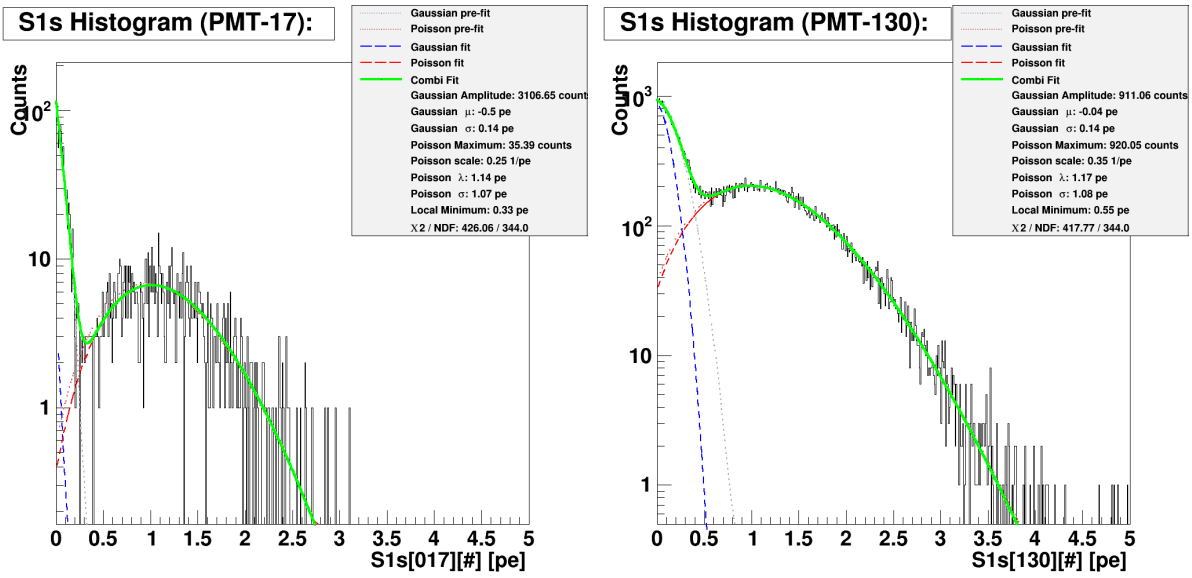
(a) S1 energy slice: $12 \leq S1 \leq 13$ pe

(b) S1 energy slice: $13 \leq S1 \leq 14$ pe

(c) S1 energy slice: $14 \leq S1 \leq 15$ pe

Figure A.20.: Simulated neutron data with *Nuclear Recoil/WIMP Simulation Code* with 2011/12 configuration. Single histograms of the cS2 signal for S1 energy slices. Additional adjusted Poisson fits to the data within $159 \leq S2 \leq 5000$ pe are plotted in black. The left fit boundary is calculated out of the electron life time correction for S2 signals.

A.3. LED Calibration with Xerawdp



(a) (b)

Figure A.21 – Two examples of the combined fit to the noise and signal part (PMT 17 and 130) which show an unusual behavior in the fit parameter outcome (see figure 4.9a and 4.8a). There the fit parameter A_G and μ_G deviates strong from PMTs around.

A.4. LED XML Config File

Listing A.1: Descriptive Caption Text

```
<xerawdp>
  <global>
    <verbosity>0</verbosity> <!-- 0: none, 1: minimal, 2: extensive, 3: debug -->
    <dataset_name_format>xe100_YYMMDD_hhmm_NNNNNN</dataset_name_format>
    <raw_data_dir>/archive/data/xenon100/run_12/</raw_data_dir>
    <processed_data_dir>/home/bauermeister/xenon100_temp/scratch/</processed_data_dir>
    <merged_data_dir>/home/bauermeister/xenon100_temp/run_12_led/</merged_data_dir>
    <minimum_bias_data_dir>/home/bauermeister/xenon100_temp/run_12_ledmb/</minimum_bias_data_dir>
    <processing_level>2</processing_level>
  </global>
  <pmts>
   <top_pmts>1-98</top_pmts>
    <bottom_pmts>99-178</bottom_pmts>
    <top_veto_pmts>179-210</top_veto_pmts>
    <bottom_veto_pmts>211-242</bottom_veto_pmts>
    <calibration_source>
      <here>
      <!-- average gains from run_11 -->
        <header>gain gain_err qe</header>
        <values>
          <!--Table with gain values -->
        </values>
      </here>
    </calibration_source>
  </pmts>
  <raw_data>
```

```
  <nb_baseline_samples>46</nb_baseline_samples>
</raw_data>
<peak_finding>
  <s1>
    <!-- <excluded_pmts>1-2, 145, 148, 157, 171, 177</excluded_pmts>  -->
    <max_nb_peaks>32</max_nb_peaks>
    <filter>
      <raised_cosine>
        <low_pass>1</low_pass>
        <limit>0.03</limit>
        <roll_off>0.2</roll_off>
        <length>31</length>
      </raised_cosine>
    </filter>
    <right_limit_height_threshold>0.05</right_limit_height_threshold> <!-- volts -->
    <signal_threshold>0.003</signal_threshold> <!-- volts -->
    <peak_window>5</peak_window> <!-- samples (standard: 60) -->
    <pre_peak_samples>10</pre_peak_samples> <!-- samples (standard: 10) -->
    <max_length>5</max_length> <!-- samples (standard: 60)-->
    <pre_peak_avg_window>2</pre_peak_avg_window> <!-- samples 50 -->
    <post_peak_avg_window>2</post_peak_avg_window> <!-- samples 10-->
    <pre_peak_avg_threshold>0.01</pre_peak_avg_threshold>
    <post_peak_avg_threshold>0.04</post_peak_avg_threshold> <!-- 0.04 -->
    <filtered_width_threshold>50</filtered_width_threshold> <!-- samples -->
    <negative_excursion_fraction_threshold>3.</negative_excursion_fraction_threshold>
    <height_fraction_threshold>0.005</height_fraction_threshold>
    <samples_below_threshold>2</samples_below_threshold> <!-- samples 2 -->
```

```
    <coincidence_threshold>7.e5</coincidence_threshold> <!-- e- -->
  </s1>
  <s2>
    S2 peak finder conditions are left out here
  </s2>
</peak_finding>
<signals>
  <signal_threshold>0.35</signal_threshold> <!-- pe -->
  <saturation_threshold>2.196</saturation_threshold> <!-- volts -->
 <s1>
   <coincidence_excluded_pmts>1-2, 145, 148, 157, 171, 177</coincidence_excluded_pmts>
   <total_signal_excluded_pmts>1-2, 145, 148, 157, 171, 177</total_signal_excluded_pmts>
   <coincidence_window>2</coincidence_window> <!-- samples -->
 </s1>
 <!--          <s2>
                    <coincidence_excluded_pmts></coincidence_excluded_pmts>
                    <total_signal_excluded_pmts></total_signal_excluded_pmts>
               </s2>   -->
</signals>
<position_reconstruction>
  Details for the position reconstuction are left out.
  Is not important for the S1 peak finder algorithm on LED waveforms
</position_reconstruction>
<position_correction>
  Details for the position correction are left out.
  Is not important for the S1 peak finder algorithm on LED waveforms
</position_correction>
```

```
  <signal_corrections>
    Details for the signal correction are left out.
    Is not important for the S1 peak finder algorithm on LED waveforms
  </signal_corrections>
  <event_selection>
    Details are left out.
    Is not important for the S1 peak finder algorithm on LED waveforms
  </event_selection>
  <extras>
    Details are left out.
    Is not important for the S1 peak finder algorithm on LED waveforms
  </extras>
  <blind_mode>
  <!-- <cut>(log10(cS2sTotBottom[0]/cS1sTot[0])) > ((2.4357) + (-0.0733998)*(cS1sTot[0])
            + (0.00359197)*((cS1sTot[0])^2)+ (-9.39834e-05 )*((cS1sTot[0])^3)
            + (1.34239e-06)*((cS1sTot[0])^4)
            + (-9.79347e-09)*((cS1sTot[0])^5)
            + (2.84814e-11)*((cS1sTot[0])^6)) || (cS1sTot[0] < 2.) || (cS1sTot[0] > 100.)
       </cut> -->
  <cut></cut>
  </blind_mode>
  <minimum_bias>
  <!--  <cut>(NbS2Peaks > 0) && (S1sCoin[0] > 1) && (cS1sTot[0] < 200.)
        </cut> -->
    <prefix>mb_</prefix>
  </minimum_bias>
</xerawdp>
```

List of Figures

List of Tables

Bibliography

[1] Aad, G., Abbott, B., Abdallah et al. "Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector". In: *Phys. Rev. D* 90 (5 09/2014), p. 052004. DOI: 10.1103/PhysRevD.90.052004. URL: http://link.aps.org/doi/10.1103/PhysRevD.90.052004 (cit. on p. 32).

[2] C.E. Aalseth et al. "Experimental constraints on a dark matter origin for the DAMA annual modulation effect". In: *Phys.Rev.Lett.* 101 (2008), p. 251301. DOI: 10.1103/PhysRevLett.101.251301. arXiv: 0807.0879 [astro-ph] (cit. on p. 44).

[3] K. Abe et al. "Neutrino Oscillation Physics Potential of the T2K Experiment". In: (2014). arXiv: 1409.7469 [hep-ex] (cit. on p. 37).

[4] Y. Abe et al. "Improved measurements of the neutrino mixing angle θ_{13} with the Double Chooz detector". In: *JHEP* 1410 (2014), p. 86. DOI: 10.1007/JHEP10(2014)086. arXiv: 1406.7763 [hep-ex] (cit. on p. 37).

[5] P. Adamson et al. "Measurement of the Velocity of the Neutrino with MINOS". In: (2012). arXiv: 1408.6267 [physics.acc-ph] (cit. on p. 37).

[6] P.A.R. Ade et al. "Planck 2013 results. XVI. Cosmological parameters". In: *Astron.Astrophys.* 571 (2014), A16. DOI: 10.1051/0004-6361/201321591. arXiv: 1303.5076 [astro-ph.CO] (cit. on p. 38).

[7] Z. Ahmed et al. "Search for Axions with the CDMS Experiment". In: *Phys.Rev.Lett.* 103 (2009), p. 141802. DOI: 10.1103/PhysRevLett.103.141802. arXiv: 0902.4693 [hep-ex] (cit. on p. 44).

[8] J.K. Ahn et al. "Observation of Reactor Electron Antineutrino Disappearance in the RENO Experiment". In: *Phys.Rev.Lett.* 108 (2012), p. 191802. DOI: 10.1103/PhysRevLett.108.191802. arXiv: 1204.0626 [hep-ex] (cit. on p. 37).

[9] R. L. Akeson et al. "The NASA Exoplanet Archive: Data and Tools for Exoplanet Research". In: 125 (08/2013), pp. 989–999. DOI: 10.1086/672273. arXiv: 1307.2944 [astro-ph.IM] (cit. on p. 28).

[10] R. BERNABEI et al. “INVESTIGATING PSEUDOSCALAR AND SCALAR DARK MATTER”. English. In: *Int. J. Mod. Phys.* 1445.A21 (2006). DOI: 10.1142/S0217751X06030874. URL: http://dx.doi.org/10.1142/S0217751X06030874 (cit. on p. 44).

[11] C. Alcock et al. “The MACHO project: Microlensing results from 5.7 years of LMC observations”. In: *Astrophys.J.* 542 (2000), pp. 281–307. DOI: 10.1086/309512. arXiv: astro-ph/0001272 [astro-ph] (cit. on p. 29).

[12] F.P. An et al. “Independent measurement of the neutrino mixing angle θ_{13} via neutron capture on hydrogen at Daya Bay”. In: *Phys.Rev.* D90.7 (2014), p. 071101. DOI: 10.1103/PhysRevD.90.071101. arXiv: 1406.6468 [hep-ex] (cit. on p. 37).

[13] F.P. An et al. “Search for a Light Sterile Neutrino at Daya Bay”. In: *Phys.Rev.Lett.* 113 (2014), p. 141802. DOI: 10.1103/PhysRevLett.113.141802. arXiv: 1407.7259 [hep-ex] (cit. on p. 40).

[14] E. Aprile and T. Doke. “Liquid xenon detectors for particle physics and astrophysics”. In: *Reviews of Modern Physics* 82 (07/2010), pp. 2053–2097. DOI: 10.1103/RevModPhys.82.2053. arXiv: 0910.4956 [physics.ins-det] (cit. on pp. 58, 68, 74).

[15] E. Aprile et al. “Analysis of the XENON100 Dark Matter Search Data”. In: *Astropart.Phys.* 54 (2014), pp. 11–24. DOI: 10.1016/j.astropartphys.2013.10.002. arXiv: 1207.3458 [astro-ph.IM] (cit. on pp. 85 sq., 93, 105 sq., 108 sq., 156, 192, 201).

[16] E. Aprile et al. “Dark Matter Results from 100 Live Days of XENON100 Data”. In: *Physical Review Letters* 107.13, 131302 (09/2011), p. 131302. DOI: 10.1103/PhysRevLett.107.131302. arXiv: 1104.2549 (cit. on pp. 79, 84).

[17] E. Aprile et al. “Dark Matter Results from 225 Live Days of XENON100 Data”. In: *Phys. Rev. Lett.* 109 (18 11/2012), p. 181301. DOI: 10.1103/PhysRevLett.109.181301. URL: http://link.aps.org/doi/10.1103/PhysRevLett.109.181301 (cit. on pp. 108, 194, 256).

[18] E. Aprile et al. “First Axion Results from the XENON100 Experiment”. In: *Phys.Rev.* D90 (2014), p. 062009. DOI: 10.1103/PhysRevD.90.062009. arXiv: 1404.1455 [astro-ph.CO] (cit. on pp. 34, 44).

[19] E. Aprile et al. “Likelihood Approach to the First Dark Matter Results from XENON100”. In: *Phys.Rev.* D84 (2011), p. 052003. DOI: 10.1103/PhysRevD.84.052003. arXiv: 1103.0303 [hep-ex] (cit. on pp. 106, 108, 201, 216, 247, 256).

[20] E. Aprile et al. “Limits on spin-dependent WIMP-nucleon cross sections from 225 live days of XENON100 data”. In: *Phys.Rev.Lett.* 111.2 (2013), p. 021301. DOI: 10.1103/PhysRevLett.111.021301. arXiv: 1301.6620 [astro-ph.CO] (cit. on p. 57).

[21] E. Aprile et al. "Material screening and selection for XENON100". In: *Astroparticle Physics* 35 (09/2011), pp. 43–49. DOI: 10.1016/j.astropartphys.2011.06.001. arXiv: 1103.5831 [physics.ins-det] (cit. on p. 71).

[22] E. Aprile et al. "Observation and applications of single-electron charge signals in the XENON100 experiment". In: *J.Phys.* G41 (2014), p. 035201. DOI: 10.1088/0954-3899/41/3/035201. arXiv: 1311.1088 [physics.ins-det] (cit. on pp. 81, 173, 207).

[23] E. Aprile et al. "Response of the XENON100 dark matter detector to nuclear recoils". In: *prd* 88.1, 012006 (07/2013), p. 012006. DOI: 10.1103/PhysRevD.88.012006. arXiv: 1304.1427 [astro-ph.IM] (cit. on pp. 81–84, 95, 171, 173, 192).

[24] E. Aprile et al. "Study of the electromagnetic background in the XENON100 experiment". In: *Physical Review D* 83.8, 082001 (04/2011), p. 082001. DOI: 10.1103/PhysRevD.83.082001. arXiv: 1101.3866 [astro-ph.IM] (cit. on p. 201).

[25] E. Aprile et al. "The neutron background of the XENON100 dark matter search experiment". In: *Journal of Physics G Nuclear Physics* 40.11, 115201 (11/2013), p. 115201. DOI: 10.1088/0954-3899/40/11/115201. arXiv: 1306.2303 [astro-ph.IM] (cit. on pp. 103 sq.).

[26] E. Armengaud et al. "Axion searches with the EDELWEISS-II experiment". In: *JCAP* 1311 (2013), p. 067. DOI: 10.1088/1475-7516/2013/11/067. arXiv: 1307.1488 [astro-ph.CO] (cit. on p. 44).

[27] Takehiko Asaka, Mikhail Shaposhnikov, and Alexander Kusenko. "Opening a new window for warm dark matter". In: *Phys.Lett.* B638 (2006), pp. 401–406. DOI: 10.1016/j.physletb.2006.05.067. arXiv: hep-ph/0602150 [hep-ph] (cit. on p. 31).

[28] V.N. Aseev et al. "An upper limit on electron antineutrino mass from Troitsk experiment". In: *Phys.Rev.* D84 (2011), p. 112003. DOI: 10.1103/PhysRevD.84.112003. arXiv: 1108.5034 [hep-ex] (cit. on p. 37).

[29] Wilson W. B. et al. "SOURCES 4A?: a code for calculating ([alpha],n), spontaneous fission, and delayed neutron sources and spectra." In: 09/1999 (cit. on p. 104).

[30] L. Baudis et al. "Response of liquid xenon to Compton electrons down to 1.5 keV". In: *Physical Review D* 87.11, 115015 (06/2013), p. 115015. DOI: 10.1103/PhysRevD.87.115015. arXiv: 1303.6891 [astro-ph.IM] (cit. on p. 81).

[31] Christoph Berger. *Elementarteilchenphysik: Von den Grundlagen zu den modernen Experimenten (Springer-Lehrbuch) (German Edition)*. 2., aktual. u. Überarb. Aufl. 2006. Springer, 04/2006. URL: http://amazon.com/o/ASIN/3540231439/ (cit. on p. 42).

[32] Gianfranco Bertone, ed. *Particle Dark Matter: Observations, Models and Searches*. 1st ed. Cambridge University Press, 11/2013. URL: http://amazon.com/o/ASIN/1107653924/ (cit. on pp. 33, 35 sq., 41, 48, 50).

[33] S. Betts et al. "Development of a Relic Neutrino Detection Experiment at PTOLEMY: Princeton Tritium Observatory for Light, Early-Universe, Massive-Neutrino Yield". In: *ArXiv e-prints* (07/2013). arXiv: 1307.4738 [astro-ph.IM] (cit. on p. 39).

[34] M. Blanton and the Sloan Digital Sky Survey. *SDSS Galaxy Map*. 2013 (cit. on p. 23).

[35] S Bonometto, V. Gorini, and U. Moschella, eds. *Modern Cosmology (Series in High Energy Physics, Cosmology and Gravitation)*. 1st. CRC Press, 12/2001. URL: http://amazon.com/o/ASIN/0750308109/ (cit. on p. 19).

[36] Alexey Boyarsky et al. "An unidentified line in X-ray spectra of the Andromeda galaxy and Perseus galaxy cluster". In: *Phys.Rev.Lett.* 113 (2014), p. 251301. DOI: 10.1103/PhysRevLett.113.251301. arXiv: 1402.4119 [astro-ph.CO] (cit. on pp. 31, 40).

[37] Rene Brun and Fons Rademakers. "ROOT - An Object Oriented Data Analysis Framework". In: *AIHENP'96 Workshop, Lausane*. Vol. 389. 1996, pp. 81–86 (cit. on p. 162).

[38] Spencer Chang et al. "Inelastic Dark Matter in Light of DAMA/LIBRA". In: *Phys.Rev.* D79 (2009), p. 043513. DOI: 10.1103/PhysRevD.79.043513. arXiv: 0807.2250 [hep-ph] (cit. on p. 56).

[39] Michele Cicoli et al. "3.55 keV photon line and its morphology from a 3.55 keV axionlike particle line". In: *Phys.Rev.* D90.2 (2014), p. 023540. DOI: 10.1103/PhysRevD.90.023540. arXiv: 1403.2370 [hep-ph] (cit. on p. 40).

[40] Douglas Clowe et al. "A direct empirical proof of the existence of dark matter". In: *Astrophys.J.* 648 (2006), pp. L109–L113. DOI: 10.1086/508162. arXiv: astro-ph/0608407 [astro-ph] (cit. on p. 20).

[41] P. Coles and F. Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure*. John Wiley, 2002, pp. 1–492 (cit. on pp. 16, 18 sq., 23, 27).

[42] Borexino Collaboration. "Neutrinos from the primary proton-proton fusion process in the Sun". In: *Nature* 512.7515 (08/2014), 383?386. DOI: 10.1038/nature13702;10.1038/nature13702 (cit. on pp. 38, 104).

[43] The XENON Collaboration. private communication (XENON100 wiki entry, Cut page). 02/15/2015 (cit. on p. 173).

[44] The XENON Collaboration. private communication (email). 02/15/2015 (cit. on pp. 267 sq.).

[45] G. W. Collins II. *The virial theorem in stellar astrophysics*. 1978 (cit. on p. 14).

[46] George W. Collins. *Fundamentals of Stellar Astrophysics*. W H Freeman & Co, 03/1989. URL: http://amazon.com/o/ASIN/0716719932/ (cit. on p. 29).

[47] Glen Cowan. *Statistical Data Analysis*. Oxford: Clarendon (Oxford), 1998 (ISBN: 0-19-850156-0 or 0-19-850155-2 in paperback) (cit. on pp. 122, 177, 221).

[48] A.V. Derbin et al. "Constraints on the axion-electron coupling constant for solar axions appearing owing to bremsstrahlung and the compton process". English. In: *JETP Letters* 95.7 (2012), 339?344. DOI: 10.1134/S002136401207003X. URL: http://dx.doi.org/10.1134/S002136401207003X (cit. on p. 44).

[49] Michael Dine, Willy Fischler, and Mark Srednicki. "A simple solution to the strong {CP} problem with a harmless axion". In: *Physics Letters B* 104.3 (1981), pp. 199–202. DOI: 10.1016/0370-2693(81)90590-6. URL: http://www.sciencedirect.com/science/article/pii/0370269381905906 (cit. on pp. 33, 44).

[50] T. Doke and K. Masuda. "Present status of liquid rare gas scintillation detectors and their new application to gamma-ray calorimeters". In: *Nucl.Instrum.Meth.* A420 (1999), pp. 62–80. DOI: 10.1016/S0168-9002(98)00933-4 (cit. on p. 75).

[51] Marco Drewes. "The Phenomenology of Right Handed Neutrinos". In: *Int.J.Mod.Phys.* E22 (2013), p. 1330019. DOI: 10.1142/S0218301313300191. arXiv: 1303.6912 [hep-ph] (cit. on p. 31).

[52] John R. Ellis et al. "Supersymmetric Relics from the Big Bang". In: *Nucl.Phys.* B238 (1984), pp. 453–476. DOI: 10.1016/0550-3213(84)90461-9 (cit. on p. 41).

[53] European Space Agency (ESA). *Weak gravitational lensing*. 03/2010. URL: http://sci.esa.int/euclid/46680-weak-gravitational-lensing/ (cit. on pp. 18 sq.).

[54] Planck Collaboration ESA. *Planck CMB*. 03/2013 (cit. on p. 24).

[55] Planck Collaboration ESA. *Planck CMB*. 03/2013 (cit. on p. 26).

[56] Planck Collaboration ESA. *Planck CMB Pie Chart*. 03/2013 (cit. on p. 24).

[57] U. Fano. "Ionization Yield of Radiations. II. The Fluctuations of the Number of Ions". In: *Phys. Rev.* 72 (1 1947), pp. 26–29. DOI: 10.1103/PhysRev.72.26. URL: http://link.aps.org/doi/10.1103/PhysRev.72.26 (cit. on p. 58).

[58] Pauline Gagnon. *Gluino, Higgsino, bingo!* URL: http://www.quantumdiaries.org/2012/12/10/gluino-higgsino-bingo-2/ (cit. on p. 41).

[59] C. Giunti and C. W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. ISBN 978-0-19-850871-7. Oxford, UK: Oxford University Press, 2007, pp. 1–728 (cit. on p. 31).

[60] Laboratory Nazionali del Gran Sasso (LNGS). *Laboratory Nazionali del Gran Sasso (LNGS)*. Website. Last call 11th January 2015. URL: http://www.lngs.infn.it/ (cit. on p. 46).

[61] A. M. Green. "Astrophysical Uncertainties on Direct Detection Experiments". In: *Modern Physics Letters A* 27, 1230004 (2012), p. 30004. DOI: 10.1142/S0217732312300042. arXiv: 1112.0524 [astro-ph.CO] (cit. on p. 34).

[62] K.A. Olive et al. (Particle Data Group). "Particle Data Book". In: *Chin. Phys. C* 38 (2014), p. 090001 (cit. on p. 38).

[63] Hamamatsu Photonoics K.K. *Photomultiplier tubes – Basics and Applications*. 2007. URL: https://www.hamamatsu.com/resources/pdf/etd/PMT_handbook_v3aE.pdf (cit. on pp. 73, 100).

[64] C. Hayashi and T. Nakano. "Evolution of Stars of Small Masses in the Pre-Main-Sequence Stages". In: *Progress of Theoretical Physics* 30 (10/1963), pp. 460–474. DOI: 10.1143/PTP.30.460 (cit. on p. 28).

[65] Richard H. Helm. "Inelastic and Elastic Scattering of 187-Mev Electrons from Selected Even-Even Nuclei". In: *Phys. Rev.* 104 (5 12/1956), pp. 1466–1475. DOI: 10.1103/PhysRev.104.1466. URL: http://link.aps.org/doi/10.1103/PhysRev.104.1466 (cit. on p. 55).

[66] CAEN Electronic Instruments. *V1724 – 8 Channel 14 bit 100 MS/s Digitizer*. 10/2015. URL: http://www.caen.it/csite/CaenProd.jsp?parent=11&idmod=483# (cit. on p. 87).

[67] J. Jortner et al. "Localized Excitations in Condensed Ne, Ar, Kr, and Xe". In: *jcp* 42 (06/1965), pp. 4250–4253. DOI: 10.1063/1.1695927 (cit. on p. 67).

[68] A. Belov B. Beltran et al. K. Barth. "CAST constraints on the axion-electron coupling". In: *Journal of Cosmology and Astroparticle Physics* 2013.05 (2013), p. 010. URL: http://stacks.iop.org/1475-7516/2013/i=05/a=010 (cit. on p. 34).

[69] Prajwal Raj Kafle et al. "On the Shoulders of Giants: Properties of the Stellar Halo and the Milky Way Mass Distribution". In: *The Astrophysical Journal* 794.1 (2014), p. 59. URL: http://stacks.iop.org/0004-637X/794/i=1/a=59 (cit. on pp. 48 sq.).

[70] Max-Planck-Institut für Kernphysik". *Research: Dark Matter*. Website. Last call 11th January 2015. URL: http://www.mpi-hd.mpg.de/lin/research_DM.en.html (cit. on p. 46).

[71] Jihn E. Kim. "Weak-Interaction Singlet and Strong CP Invariance". In: *Phys. Rev. Lett.* 43 (2 07/1979), pp. 103–107. DOI: 10.1103/PhysRevLett.43.103. URL: http://link.aps.org/doi/10.1103/PhysRevLett.43.103 (cit. on pp. 33, 44).

[72] Alexander Kish. "Dark Matter Search with the XENON100 Experiment". PhD thesis. University of Zrich, 2011 (cit. on p. 99).

[73] Hans Volker Klapdor-kleingrothaus and Geraint F. Lewis, eds. *Dark Matter In Astroparticle and Particle Physics, Dark 2007: Proceedings of the 6th International Heidelberg Conference, University of Sydney, Australia 24 - 28 September 2007*. World Scientific Pub Co Inc, 05/2008. URL: http://amazon.com/o/ASIN/9812814345/ (cit. on p. 33).

[74] Edward W. Kolb, Daniel J. H. Chung, and Antonio Riotto. "WIMPzillas!" In: *Trends in theoretical physics II. Proceedings, 2nd La Plata Meeting, Buenos Aires, Argentina, November 29-December 4, 1998*. [,91(1998)]. 1998, pp. 91–105. arXiv: hep-ph/9810361 [hep-ph]. URL: http://www.adsabs.harvard.edu/abs/hep-ph/9810361 (cit. on p. 37).

[75] E. Komatsu et al. "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation". In: *ApJS* 192, 18 (02/2011), p. 18. DOI: 10.1088/0067-0049/192/2/18. arXiv: 1001.4538 [astro-ph.CO] (cit. on p. 38).

[76] V.I. Korchagin et al. "Local surface density of the Galactic disk from a 3-D stellar velocity sample". In: *Astron.J.* 126 (2003), p. 2896. DOI: 10.1086/379138. arXiv: astro-ph/0308276 [astro-ph] (cit. on p. 13).

[77] S Kubota, M Hishida, and J Raun. "Evidence for a triplet state of the self-trapped exciton states in liquid argon, krypton and xenon". In: *Journal of Physics C: Solid State Physics* 11.12 (1978), p. 2645. URL: http://stacks.iop.org/0022-3719/11/i=12/a=024 (cit. on p. 75).

[78] V.A. Kudryavtsev. "Muon simulation codes {MUSIC} and {MUSUN} for underground physics". In: *Computer Physics Communications* 180.3 (2009), pp. 339–346. DOI: 10.1016/j.cpc.2008.10.013. URL: http://www.sciencedirect.com/science/article/pii/S0010465508003640 (cit. on p. 104).

[79] J.D. Lewin and P.F. Smith. "Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil". In: *Astropart.Phys.* 6 (1996), pp. 87–112. DOI: 10.1016/S0927-6505(96)00047-3 (cit. on pp. 48, 50).

[80] Andrew Liddle. *An Introduction to Modern Cosmology*. 2nd ed. Wiley, 05/2003. URL: http://amazon.com/o/ASIN/0470848359/ (cit. on p. 24).

[81] A. I. Vainshtein M. A. Shifman and V. I. Zakharov. "Can Confinement Ensures Natural CP Invariance of Strong Interactions?" English. In: *Nucl. Phys. B* 166 (1980), p. 493 (cit. on p. 44).

[82] Yuan Mei. "Direct Dark Matter Search with the XENON100 Experiment". PhD Thesis. Rice University, Texas, 2011, 12 et seqq. (Cit. on pp. 56, 72, 80).

[83] D. Merritt. "The distribution of dark matter in the coma cluster". In: *ApJ* 313 (02/1987), pp. 121–135. DOI: 10.1086/164953 (cit. on p. 14).

[84] Laboratoire Souterrain de Modane. *Laboratoire Souterrain de Modane*. Website. Last call 11th January 2015. URL: http://www.lsm.in2p3.fr/ (cit. on p. 46).

[85] M. A. Monroy-Rodríguez and C. Allen. "The End of the MACHO Era, Revisited: New Limits on MACHO Masses from Halo Wide Binaries". In: *APj* 790, 159 (08/2014), p. 159. DOI: 10.1088/0004-637X/790/2/159. arXiv: 1406.5169 (cit. on p. 29).

[86] Takenori Nakano. "Pre-main Sequence Evolution and the Hydrogen-Burning Minimum Mass". English. In: *50 Years of Brown Dwarfs*. Ed. by Viki Joergens. Vol. 401. Astrophysics and Space Science Library. Springer International Publishing, 2014, pp. 5–17. DOI: 10.1007/978-3-319-01162-2_2. URL: http://dx.doi.org/10.1007/978-3-319-01162-2_2 (cit. on p. 28).

[87] ESA/Hubble & NASA. *A Horseshoe Einstein Ring from Hubble*. 12/2011. URL: http://apod.nasa.gov/apod/ap111221.html (cit. on p. 18).

[88] "National Nuclear Data Center (NuDat)". In: (). URL: http://www.nndc.bnl.gov/nudat (cit. on pp. 93, 95).

[89] K. Ni et al. "Preparation of neutron-activated xenon for liquid xenon detector calibration". In: *Nuclear Instruments and Methods in Physics Research A* 582 (11/2007), pp. 569–574. DOI: 10.1016/j.nima.2007.08.180. arXiv: 0708.1976 [physics.ins-det] (cit. on pp. 96, 98).

[90] *NIST/Xcom database*. http://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html. 2015 (accessed January 6th, 2015) (cit. on p. 43).

[91] Travis Oliphant et al. *NumPy is the fundamental package for scientific computing with Python*. 2006–. URL: www.numpy.org (cit. on p. 221).

[92] Athanasios Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill series in electrical engineering. New York: McGraw-Hill, 1991 (cit. on p. 243).

[93] Roberto D Peccei and Helen R Quinn. "Constraints imposed by CP conservation in the presence of pseudoparticles". In: *Phys. Rev. D* 16.6 (1977), pp. 1791–1797 (cit. on p. 33).

[94] D.H. Perkins. *Particle Astrophysics, Second Edition (Oxford Master Series in Physics)*. 2nd ed. Oxford University Press, 02/2009. URL: http://amazon.com/o/ASIN/0199545464/ (cit. on pp. 17, 21, 23, 37).

[95] Guillaume Plante. "The XENON100 Dark Matter Experiment: Design, Construction, Calibration and 2010 Search Results with Improved Measurement of the Scintillation Response of Liquid Xenon to Low-Energy Nuclear Recoils". PhD thesis. Columbia University, New York, 2012 (cit. on pp. 89 sq.).

[96] R. Westermann R. Fraedrich J. Schneider. "Exploring the Millennium Run - Scalable Rendering of Large-Scale Cosmological Datasets". In: *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS* (09/2009), pp. 1251 –1258. DOI: 10.1109/TVCG.2009.142. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5290736&isnumber=5290686 (cit. on pp. 23 sq.).

[97] Javier Redondo. "Solar axion flux from the axion-electron coupling". In: *JCAP* 1312 (2013), p. 008. DOI: 10.1088/1475-7516/2013/12/008. arXiv: 1310.0823 [hep-ph] (cit. on pp. 33 sq.).

[98] Matts Roos. *Introduction to Cosmology*. 3rd ed. Wiley, 12/2003. URL: http://amazon.com/o/ASIN/047084910X/ (cit. on p. 13).

[99] V. C. Rubin and W. K. Ford Jr. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *Astrophysical Journal* 159 (02/1970), p. 379. DOI: 10.1086/150317 (cit. on p. 11).

[100] H. Georgi S. Dimopoulos. "Softly Broken Supersymmetry and SU(5)". In: *Nuclear Physics B.* 193 (1981), pp. 150–162. URL: http://web.stanford.edu/~savas/papers/SoftlyBrokenSupersymmetryAndSU%285%29.pdf (cit. on p. 41).

[101] R. K. Sachs and A. M. Wolfe. "Perturbations of a Cosmological Model and Angular Variations of the Microwave Background". In: *APj* 147 (01/1967), p. 73. DOI: 10.1086/148982 (cit. on p. 27).

[102] J. Sahlmann et al. "Astrometric orbit of a low-mass companion to an ultracool dwarf". In: *AAp* 556, A133 (08/2013), A133. DOI: 10.1051/0004-6361/201321871. arXiv: 1306.3225 [astro-ph.SR] (cit. on p. 28).

[103] Stefan Schindler. "Bayesian approach to the analysis of XENON100 data". In: (2013) (cit. on p. 201).

[104] Peter Schneider. *Einführung in die extragalaktische Astronomie und Kosmologie*. Springer, 2006. URL: http://books.google.de/books?id=EbObvjTDfIcC (cit. on pp. 12–16, 22, 48).

[105] Phillips Scientific, ed. *16 Channel Photomultiplier Preamplifier, NIM Model 776*. Phillips Scientific. URL: http://www.phillipsscientific.com/pdf/776ds.pdf (cit. on p. 87).

[106] "Search for solar axions in XMASS, a large liquid-xenon detector". In: *Physics Letters B* 724.1?3 (2013), 46?50. DOI: 10.1016/j.physletb.2013.05.060. URL: http://www.sciencedirect.com/science/article/pii/S0370269313004425 (cit. on p. 44).

[107] C. Silva et al. "Reflectance of polytetrafluoroethylene for xenon scintillation light". In: *Journal of Applied Physics* 107.6 (03/2010), p. 064902. DOI: 10.1063/1.3318681. arXiv: 0910.1056 [physics.ins-det] (cit. on p. 71).

[108] P. Sorensen. "Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches". In: *Nuclear Instruments and Methods in Physics Research A* 635 (04/2011), pp. 41–43. DOI: 10.1016/j.nima.2011.01.089. arXiv: 1102.2865 [astro-ph.IM] (cit. on p. 79).

[109] Volker Springel et al. "Simulating the joint evolution of quasars, galaxies and their large-scale distribution". In: *Nature* 435 (2005), pp. 629–636. DOI: 10.1038/nature03597. arXiv: astro-ph/0504097 [astro-ph] (cit. on p. 23).

[110] NASA/JPL-Caltech/R.Hurt (SSC-Caltech). *The Structure of the Milky Way Galaxy*. Website. Last call 19th January 2015. URL: http://www.spitzer.caltech.edu/images/1925-ssc2008-10b-A-Roadmap-to-the-Milky-Way-Annotated- (cit. on p. 49).

[111] A. Unsöld and Bodo Baschek. *Der neue Kosmos: Einführung in die Astronomie und Astrophysik (German Edition)*. 7. Aufl. 2002. Korr. Nachdruck 2004. Springer, 09/2004. URL: http://amazon.com/o/ASIN/3540421777/ (cit. on p. 15).

[112] F. J. Vaughn et al. "Total Neutron Cross Sections of Helium, Neon, Argon, Krypton, and Xenon". In: *Physical Review* 118 (1960), p. 683. DOI: 10.1103/PhysRev.118.683. URL: http://dx.doi.org/10.1103/PhysRev.118.683 (cit. on p. 87).

[113] Marc Weber. "Gentle Neutron Signals and Noble Background in the XENON100 Dark Matter Search Experiment". PhD thesis. Ruperto-Carola-University of Heidelberg, Germany, 2013 (cit. on pp. 83, 207).

[114] X-ray: NASA/CXC/M.Markevitch et al., Optical: NASA/STScI; Magellan/U.Arizona/D.Clowe et al., Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al. *Bullet Cluster – Object: 1ES 0657-55.8*. 2006. URL: http://hubblesite.org/ (cit. on p. 20).

[115] Xenon100 Collaboration et al. "The XENON100 dark matter experiment". In: *Astroparticle Physics* 35 (04/2012), pp. 573–590. DOI: 10.1016/j.astropartphys.2012.01.003. arXiv: 1107.2155 [astro-ph.IM] (cit. on pp. 64–67, 69 sq., 76 sq., 79 sqq., 86, 88, 90 sq., 97, 99 sq.).

[116] M. Yamashita et al. "Scintillation response of liquid Xe surrounded by {PTFE} reflector for gamma rays". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 535.3 (2004), pp. 692–698. DOI: 10.1016/j.nima.2004.06.168. URL: http://www.sciencedirect.com/science/article/pii/S0168900204016638 (cit. on p. 71).

[117] A. R. Zhitnitsky. "On possible suppression of the axion hadron interactions". English. In: *Sov. J. Nucl. Phys.* 31 (1980), p. 260 (cit. on p. 44).