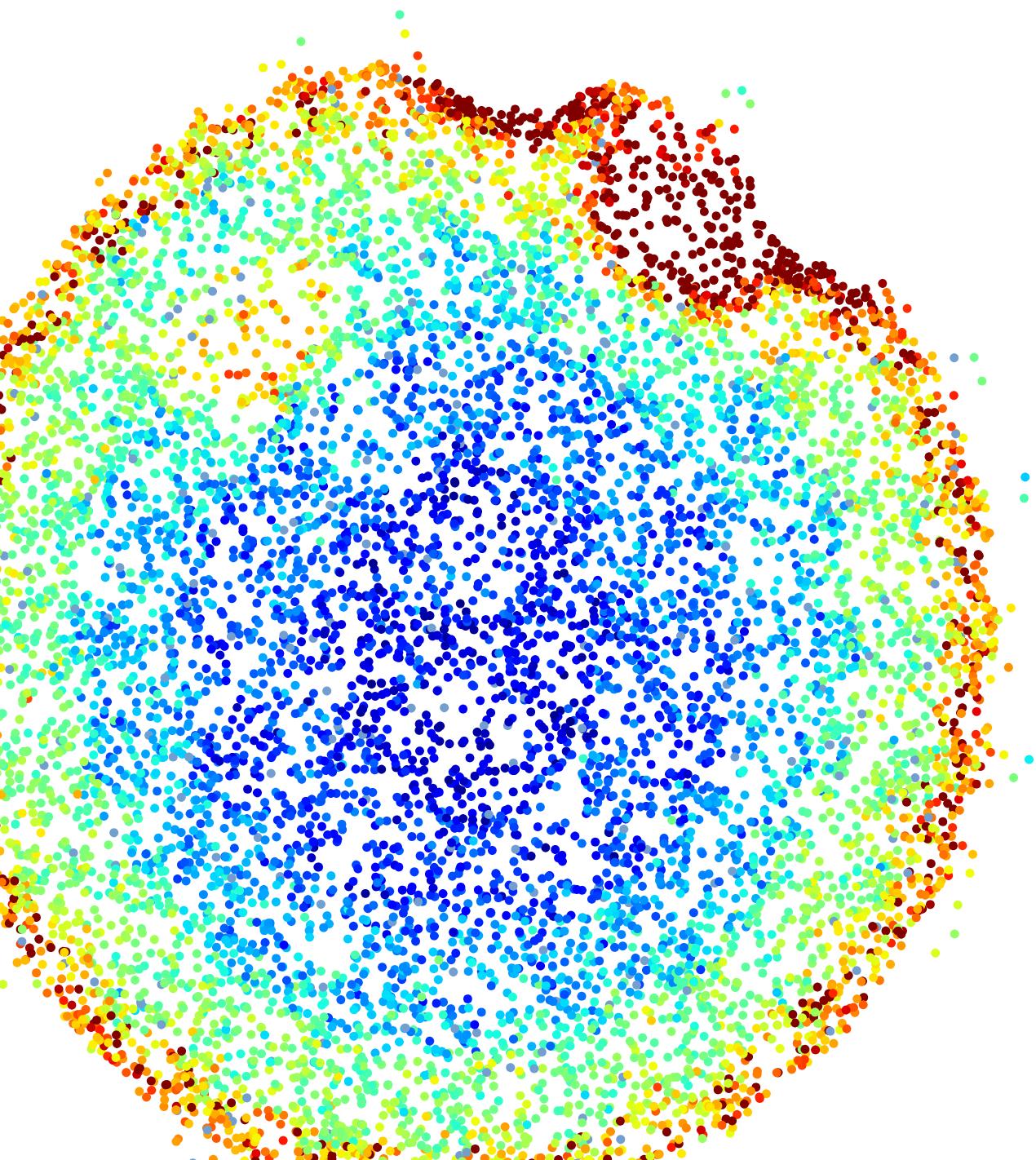


# POSITION RECONSTRUCTION AND DATA QUALITY IN XENON

BART PELSSERS



Cover image: Reconstructed positions using a charge weighted sum.

MASTER THESIS  
Particle Physics

JULY 23, 2015

---

Position Reconstruction and Data Quality in  
XENON

---

**Supervisor Nikhef:**

Prof.Dr. M.P. Decowski

**Supervisor UU:**

Prof.Dr. R.J.M. Snellings

**Daily Supervisor:**

A. Tiseni

**Author:**

B.E.J. Pelssers



**Universiteit Utrecht**

University of Utrecht



Nikhef



## Abstract

The XENON collaboration operates a dark matter detector (XENON100) at Gran Sasso, Italy. A new detector (XENON1T) is under construction. This thesis reports on the working and implementation of a position reconstruction method developed for XENON1T, the definition of several data quality cuts for XENON100 and the first physics analysis performed with a new data processor for XENON1T, called PAX.

The detectors allow the position of an event to be reconstructed in three dimensions. The  $\chi^2_\gamma$  position reconstruction algorithm is physically motivated and provides a per-event error estimation. It is implemented in the data processor for XENON1T and tested using simulations in a XENON100 configuration. The new implementation performs better than its main competitor, a Neural Network.

Two data quality cuts are defined to reject poorly position reconstructed events for the latest calibration run of XENON100. Both cuts are redefined compared to their previous versions to cope with a higher noise level in the XENON100 detector. A third cut on the ratio of light seen in the top and bottom of the detector (asymmetry) is also redefined.

The new data processor PAX is used to study and classify a group of previously unseen events with very low asymmetry. The events have a very unusual combination of properties implying that they may be caused by scintillation light in the liquid xenon, a process very unlikely to happen in XENON100. Further research will be needed to test this hypothesis.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Dark Matter</b>	<b>3</b>
2.1	Rotation curves of galaxies . . . . .	5
2.2	The WIMP hypothesis . . . . .	7
2.3	Searching for particle dark matter . . . . .	8
2.3.1	Current limits . . . . .	10
<b>3</b>	<b>Direct Detection Experiments</b>	<b>13</b>
3.1	Liquid xenon TPCs . . . . .	14
3.1.1	XENON100 . . . . .	16
3.1.2	XENON1T . . . . .	18
<b>4</b>	<b>Position Reconstruction</b>	<b>21</b>
4.1	Reconstruction algorithms . . . . .	22
4.1.1	Light-Collection Efficiency . . . . .	23
4.1.2	Previous position reconstruction algorithms in XENON100 . . . . .	24
4.1.3	Position reconstruction in XENON1T . . . . .	25
4.2	$\chi^2$ Reconstruction . . . . .	25
4.2.1	Description of the $\chi^2$ position reconstruction in PAX . . . . .	29
4.3	Performance . . . . .	29
4.3.1	Local position reconstruction errors . . . . .	32
<b>5</b>	<b>Data Quality Analysis</b>	<b>37</b>
5.1	$\chi^2$ Data quality cut . . . . .	38
5.1.1	Redefining the cut . . . . .	41
5.2	Position discrepancy cut . . . . .	43
5.3	Asymmetry cut . . . . .	46
<b>6</b>	<b>Low Asymmetry S2s</b>	<b>51</b>
6.1	Event selection . . . . .	51
6.2	Clean AmBe data . . . . .	52
6.2.1	Waveforms . . . . .	55
6.3	Comparing other sources . . . . .	58

6.4	Hypotheses . . . . .	60
<b>7</b>	<b>Conclusion and Discussion</b>	<b>63</b>
7.1	Outlook . . . . .	64
<b>A</b>	<b>MC Position Reconstruction plots</b>	<b>67</b>
<b>B</b>	<b>Waveforms</b>	<b>71</b>

# Chapter 1

## Introduction

Every second roughly one billion dark matter particles travel through the human body. Together these particles make up 85% of the mass of the entire Universe, yet no one has ever directly detected a single one.

The above two statements rest on the assumption that dark matter is a particle. An assumption still unconfirmed by experiment making the search for the particle nature of dark matter one of the most intriguing current challenges in physics. Even though there is plenty of evidence for the existence of dark matter on a macroscopic scale its microscopic nature remains unknown.

Searching for dark matter is no easy task. Even though billions of particles are expected to travel through a detector every second, their interaction probability is extremely low. So low in fact that to date only upper limits of this interaction cross section are known. To detect the faint signals deposited by dark matter particles highly specialised detectors have to be designed and build. Often buried deep underground they gather data for years hoping to record an interaction of dark matter.

The XENON collaboration has held the record in setting the lowest dark matter limits for direct detection for several years. These limits are the result of operating the XENON100 detector in an underground laboratory in Gran Sasso, Italy. The construction of a new, bigger and more sensitive detector is ongoing and expected to be completed by the end of 2015. This new detector, XENON1T, will explore a region of parameter space never probed before. Either setting a new limit roughly two orders of magnitude below the current best ones or finding dark matter.

This thesis is the result of a year's work at the Nikhef XENON group, Amsterdam. Chapter 2 gives an introduction in dark matter theory. Detection principles and detectors are discussed in Chapter 3. The main focus of this thesis is on position reconstruction and data quality in the detector. These topics are discussed in Chapter 4 and Chapter 5 respectively. Chapter 6 looks at unusual events seen by using the new data processor on current data.



# Chapter 2

## Dark Matter

The Universe is composed of many different types of matter. Standard model particles composing stars, planets and other luminous objects only account for a small fraction of the entire matter content. The largest part of the Universe is composed of matter that is not visible, called dark matter and dark energy. These rather generic terms refer to an additional massive component and an additional energy component. Both are called dark because they are invisible electromagnetically and their microscopic nature is unknown. Macroscopically the influence of dark matter and energy on the structure of the universe can be measured.

One of the open questions in physics today is the microscopic nature of dark matter. Is dark matter composed of particles and if so which particles? How do these particles interact with Standard Model matter? Cosmology has provided us with models describing the state of the universe. Based on measurements of the Cosmic Microwave Background (CMB) by experiments such as WMAP [1] and Planck [2] we now know the key parameters of the macroscopic properties of the universe such as the energy density of each of its components. One of the results of these experiments is that they showed how little baryonic matter the universe contains: less than 5% of the total energy content. The remaining 95% being composed of dark matter ( $\sim 25\%$ ) and dark energy ( $\sim 70\%$ ). Figure 2.1 shows a breakdown of the different components of the universe showing that dark matter and dark energy are the dominant components by far. The remaining baryonic mass is mostly occupied by free hydrogen and helium with only a tiny fraction belonging to the heavy elements.

The current model for describing the evolution and content of the universe is called  $\Lambda$ CDM or Lambda Cold Dark Matter. Sometimes also referred to as the ‘Standard Model of Cosmology’ and is currently the most accurate model for describing the evolution of the universe. The name  $\Lambda$ CDM refers to the two major components of the model, a cosmological constant  $\Lambda$  to model the dark energy density and a Cold Dark Matter part modelling the dark matter density.

‘Cold’ is this sense refers to the fact that the dark matter particles in this model have low kinetic energy and move at non-relativistic velocities. ‘Dark’ refers to the fact that the particles interact extremely weakly electromagnetically or not electromagnetically

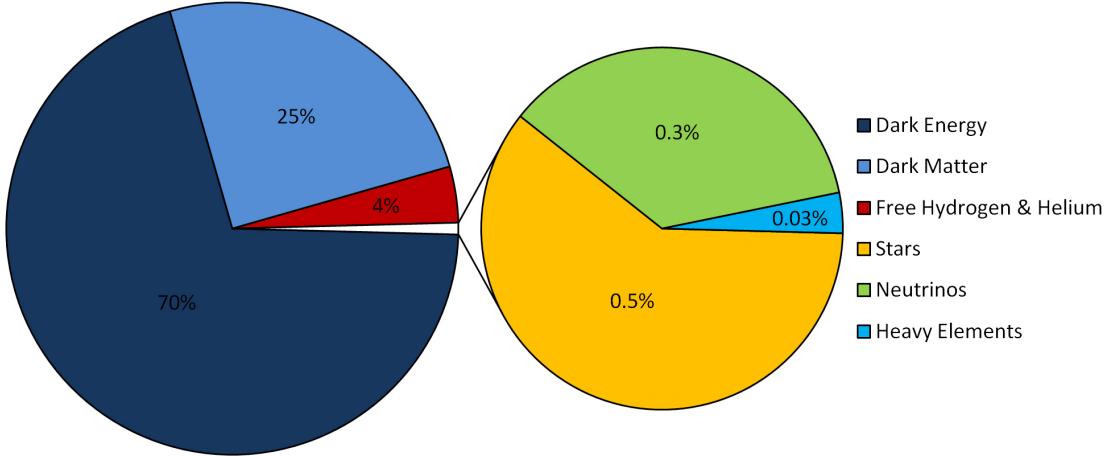


Figure 2.1: The matter-energy components of the universe. Percentages are rounded, measured values in table 2.1

at all.

Exactly how much of these components is present in the Universe is shown in Table 2.1 which shows some of the key parameters reported by the Planck Collaboration in their 2015 paper [2].

Name	Value	Description
$\Omega_\Lambda$	$0.692 \pm 0.012$	Dark energy fraction
$\Omega_m$	$0.308 \pm 0.012$	Matter fraction
$\Omega_b h^2$	$0.02226 \pm 0.00023$	Baryonic mass
$\Omega_c h^2$	$0.1186 \pm 0.0020$	CDM mass
$\Omega_m h^2 = \Omega_b h^2 + \Omega_c h^2$	$0.1415 \pm 0.0019$	Total mass

Table 2.1:  $\Lambda$ CDM parameters by Planck [2],  $\Omega_i \equiv \frac{\rho_i}{\rho_c}$ ,  $h$  is the reduced Hubble constant defined as  $H_0 = 100 h$  (km/s)/Mpc,  $H_0 = 67.74 \pm 0.46$  (km/s)/Mpc

Each  $\Omega$  in Table 2.1 gives an energy density as a fraction of the critical energy density of the universe. The critical energy density is defined to be the energy density at which the universe has a flat geometry. Measurements have shown that the actual energy density is very close to this value yielding a total  $\Omega$  of 1. These relations are shown in Equation 2.1. In  $\Lambda$ CDM the total  $\Omega$  is fixed to one.

$$1 \equiv \Omega_{tot} = \sum_i \Omega_i \equiv \sum_i \frac{\rho_i}{\rho_c} \quad (2.1)$$

Dark matter is non-luminous matter, meaning that it seems not to interact electromagnetically, else it would be visible to telescopes. It must however be massive to explain the missing mass and thus interact via gravitation. One of the more popular hypotheses for the particle nature of dark matter is the WIMP hypothesis, described

in section 2.2. This model proposes that dark matter is composed of Weakly Interacting Massive Particles. These hypothetical, stable, neutral and massive particles have all the properties to be dark matter candidates. The added property that they also interact through the weak nuclear force provides a mechanism for their creation in the early Universe. The weak interaction of these particles can also be exploited to detect them in experiments. These kind of searches require specialised detectors with a very low background, the detector type used by the XENON collaboration is discussed in chapter 3.

## 2.1 Rotation curves of galaxies

Evidence for the existence of dark matter in fields other than cosmology comes from astronomy. One of the earliest and most intuitive indications for the existence of dark matter is the study of rotation curves of spiral galaxies. Their dynamics show that they must be much more massive than can be inferred based on their luminous mass alone. These large astronomical objects slowly rotate around their centre of mass. By measuring Doppler shifts of stars their orbital velocity can be calculated. A rotation curve gives the orbital velocity of stars as a function of their distance from the galaxies centre.

By using classical Newtonian dynamics we can calculate this function for a given mass distribution. By simply balancing the gravitational force with the centripetal force acting on the stars in the galaxy.

$$\vec{F}_G = \vec{F}_C \Rightarrow \frac{GM(r)m}{r^2} = \frac{mv^2}{r} \Rightarrow v(r) = \sqrt{\frac{GM(r)}{r}} \quad (2.2)$$

If we now assume that most of the mass of the galaxy is concentrated in the centre, the rotation curve function is expected to behave as  $v(r) \propto \frac{1}{\sqrt{r}}$ .

Figure 2.2 shows that this is not the case when we look at a measurement of the rotation curve of spiral galaxy M33 by Corbelli [3]. The curve expected from the stellar disk alone is plotted as the short dashed line. It shows the  $\propto \frac{1}{\sqrt{r}}$  dependency for large radii. The data points do not show this dependency at all and the rotation curve (solid line) keeps rising even at large radii. Two different components are added to fit the data points, a contribution for gas present in the galaxy (long dashed line) and one for a halo contribution (dashed-dotted line).

Rotation curves have also been measured for other galaxies such as the 21 spiral galaxies measured by Rubin [4] in 1980. Those galaxies, of varying shapes and sizes, also show a large discrepancy when comparing the measured to the expected curves. All these galaxies have in common that at large radii the orbital velocity of the stars does not go to zero as expected but rather remains roughly constant or keeps rising.

The large discrepancy between the measured and predicted rotation curves is resolved by adding a halo component. This means letting go of our initial assumption that the mass of the galaxy is concentrated in the centre. If the galaxies mass is distributed to far larger radii than the visible parts of the galaxy, the rotation curve can describe the

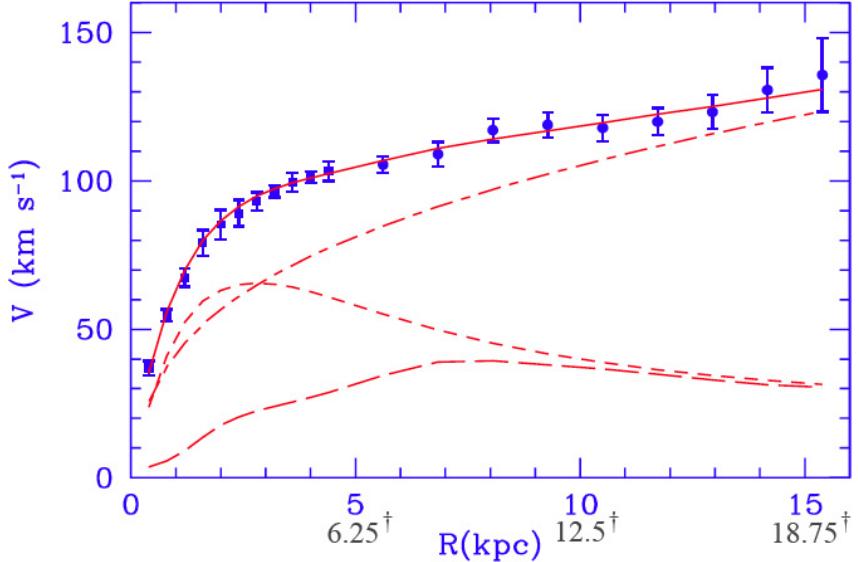


Figure 6. M33 rotation curve (points) compared with the best fit model (continuous line). Also shown the halo contribution (dashed-dotted line), the stellar disk (short dashed line) and the gas contribution (long dashed line)

E. Corbelli & P. Salucci (1999); <http://arxiv.org/abs/astro-ph/9909252v1>

\* “we have assumed the following [arbitrary] conversion:

5 arcmin = 1 kpc corresponding to a distance  $D \approx 0.7$  Mpc.”

\* This assumption assumes  $D = 0.6875$  Mpc, which is a 25.2% decrease in the consensus, measured distance of 0.92 Mpc. † Annotations correct for  $D = 0.92$  Mpc. A substantial increase in ‘missing’ mass is required to produce the rotation curve for these distances.

Figure 2.2: M33 rotation curve by Corbelli [3], the mass from the stellar disk alone can not account for the observations. A halo component is needed to explain the behaviour at large radii.

measured curve. For example if  $M(r)$  does not go quickly to zero but instead behaves as  $M(r) \propto r$  for  $r < r_{Halo}$  with  $r_{Halo}$  being several times the galaxy radius. This would mean that the galaxy is much more massive than the mass that can be inferred from the visible stars, gas and dust alone.

The dark matter halo model states that galaxies and even galaxy clusters are surrounded by a halo of dark matter acting on the galaxy gravitationally. This halo extends far beyond the disk of visible stars and gas.

The existence of dark matter halos is also confirmed by gravitational lensing effects around galaxy clusters. Gravitational lensing is an effect from general relativity that is causing massive objects to act as a lens distorting light from very distant objects. This lensing effect is also influenced by the presence of dark matter, causing a stronger lensing that can be measured and used to infer the distribution of dark matter around clusters. One such example is the Bullet Cluster which also favors a particle dark matter explanation instead of a fluid like mass distribution [5].

Another explanation for rotation curves of galaxies would require modifying Newtonian gravity, turning the ‘missing mass’ problem into a ‘acceleration discrepancy’ problem. Theories such as Modified Newtonian Dynamics (MOND) attempt to explain rotation curves without the need for dark matter. A detailed review of the pros and cons of these two theories can be found in [6].

Dark matter is an integral part of modern cosmology and plays an important role in  $\Lambda$ CDM, its macroscopic influence on matter has been measured and many models for particle dark matter have been proposed.

## 2.2 The WIMP hypothesis

There exist many theoretical models for the nature of particle dark matter. One of the more popular models for particle dark matter has always been the WIMP hypothesis, Weakly Interacting Massive Particles or WIMPs that make up the dark matter component of the universe.

These WIMP particles were originally created in the early universe and in thermal equilibrium with other particles due to the extreme high temperatures. As the universe expands and cools, at some point the temperature ( $T$ ) drops below the mass of the WIMPs ( $M_\chi$ ) causing their number density to drop exponentially following  $\propto \exp(-M_\chi/T)$  since WIMPs now annihilate more often than they are created. If the particles remained in a thermal equilibrium there would be very few WIMPs left today since their number density drops exponentially to zero. However if the number density becomes low enough so that probability of two WIMPs annihilating is negligible then the WIMPs are ‘frozen out’ and remain at a constant number density still present today.

Cosmology tells us that the annihilation cross section for a thermally created particle in the early universe must be on the same order as the electroweak scale. This fortunate coincidence is sometimes referred to as the ‘WIMP miracle’ and is the main reason why this hypothesis is so popular among theorists.

Any stable, neutral and massive particle that does not interact electromagnetically can in principle be a dark matter candidate. Adding the additional properties of the WIMP hypothesis, i.e. a weak scale interaction and a mass in the GeV-TeV range, also gives the correct relic density from cosmology and allows experimentalists to possibly detect this particle directly through its weak interaction with ordinary baryonic matter.

Many theoretical models that extend the Standard Model of particle physics (such as Supersymmetry (SUSY)), naturally have dark matter particles in their model. For example: in SUSY the neutralino (a hypothetical massive neutrino-like particle) is a dark matter candidate. For a review on the status of SUSY dark matter candidates see [7].

### 2.3 Searching for particle dark matter

Suppose that dark matter is composed of WIMPs, how would these particles interact with the ordinary, Standard Model, matter we can observe? If there is such an interaction there are three possible interaction processes, namely the creation, scattering and annihilation of dark matter with standard model matter. Of these processes the creation and annihilation modes are referred to as indirect detection, scattering as direct detection.

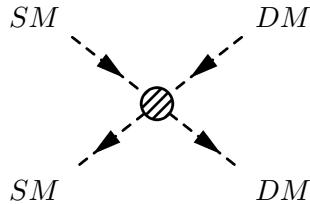


Figure 2.3: Production

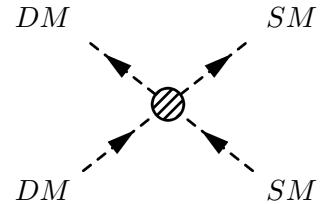


Figure 2.4: Annihilation

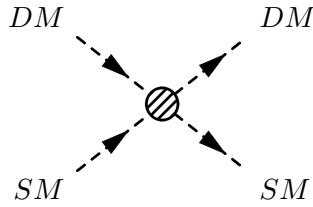


Figure 2.5: Scattering

Figure 2.6: Dark Matter interaction processes,  $DM \Rightarrow$  Dark Matter,  $SM \Rightarrow$  Standard model Matter. Time on horizontal axis.

Figure 2.6 shows the Feynman diagrams of these three processes. As can be seen they are rotations of the same process.

In the production process one could try and produce dark matter particles from ordinary matter, in experiments conducted at CERN's Large Hadron Collider for example. In these type of collisions the production of dark matter would be indicated by events with a missing energy since the dark matter particles created would not be detected and escape the detector unseen. Limits have been published by ATLAS [8].

In the annihilation scenario dark matter annihilates into standard model matter. Experiments such as the Fermi satellite scan the Universe for unknown peaks or emission lines that could originate from dark matter annihilations. One recent example is a 3.55 keV X-ray line and a 1-3 GeV  $\gamma$ -ray flux excess seen by the Fermi-LAT [9] telescope.

Direct detection experiments such as XENON are using the scattering process to look for dark matter. The aim is to have a dark matter particle scatter off standard model matter and by measuring the recoiling atom or nucleus infer the mass and cross section of the dark matter particle. As the name suggests the target material used by

XENON is liquid xenon. This dense cold liquid has several advantageous properties for use in direct detection experiments as will be discussed in chapter 3.

Direct detection experiments are essentially counting experiments, probing a part of the mass-cross section parameter space. Due to the very low interaction probability of dark matter particles these experiments often run for an extended amount of time, taking data for multiple years.

The interaction rate of WIMPs with the target material can be calculated given various astrophysical parameters and a few assumptions (such as the WIMP mass).

Given the fact that galaxies have dark matter halos, as hypothesised by measurements of rotation curves described in section 2.1, we can calculate the flux of dark matter particles at earth. First we assume the WIMP hypothesis to be true, next we assume that the dark matter halo of our galaxy is composed of WIMPs in an isotropic and hydrostatic equilibrium. We use some parameters from the Standard Halo Model for our galaxy [10].

Name	Value	Description
$\rho_0$	$0.35^{+0.08}_{-0.07}$ GeV/cm <sup>3</sup>	Local dark matter density
$v_0$	220 km/s	Orbital velocity at $R_0$
$v_e$	$550.9^{+32.4}_{-22.1}$ km/s	Escape velocity at $R_0$

Table 2.2: Parameters from the Milky Way Halo [10]

The dark matter flux can be expressed as  $\Phi = \frac{\rho_\chi}{m_\chi} v$ , or the local dark matter density over the WIMP mass times the WIMP velocity. The interaction rate with a given target consisting of nuclei with atomic mass  $a$ , Avogadro's number  $N_A$  and WIMP-nucleus interaction cross section  $\sigma A^2$  ( $\sigma$  being the WIMP-nucleon interaction cross section and  $A$  being the atomic number) expressed as given in equation 2.3.

$$R = \Phi(\sigma A^2) \frac{N_A}{a} \Rightarrow R = \frac{\rho_\chi}{m_\chi} v(\sigma A^2) \frac{N_A}{a} \quad (2.3)$$

To see how small a typical rate actually is we calculate a number for it and choose the parameters given in table 2.3. Only the mass of the WIMP is completely unknown and here we assume it to be 100 GeV/c<sup>2</sup> (a sensible assumption given the WIMP hypothesis). The interaction cross section we take to be 10<sup>-38</sup> cm<sup>2</sup>, a typical value for weak scale processes, though as will be shown later this cross section is orders of magnitude too large and the resulting rate too optimistic.

Putting these numbers into equation 2.3 we obtain the following number

$$R = 5,98 \cdot 10^{-8} \text{ g}^{-1}\text{s}^{-1} \Rightarrow R \sim 1.8 \cdot 10^3 \text{ kg}^{-1}\text{year}^{-1} \quad (2.4)$$

As can be seen from equation 2.4 this rate is extremely low, even if a detector with an active volume of 10 kg is left running for a year, only a few thousand interactions of a WIMP with the target would be expected. Moreover experiments have shown that the cross section must be several orders of magnitude lower than the typical weak scale cross section resulting in a rate that is even lower. Current direct detection experiments are

Name	Value	Description
$\rho_\chi$	0.35 GeV/cm <sup>3</sup>	$\rho_0$ from table 2.2
$m_\chi$	100 GeV	Assumed
$v$	220 km/s	$v_0$ from table 2.2
$\sigma$	$10^{-38}$ cm <sup>2</sup>	Typical weak scale cross section
$N_A$	$6.02 \cdot 10^{23}$ mol <sup>-1</sup>	Constant
$a$	131 g/mol	Atomic mass of xenon
$A$	130	Atomic number of xenon

Table 2.3: Parameters to calculate the rate

therefore using large detectors approaching the tonne scale in active detector volume, taking data for multiple years.

### 2.3.1 Current limits

Currently there is only one scientific collaboration that claims to have directly detected dark matter, the DAMA/LIBRA [11] experiment. Many other experiments have however excluded their signal and have set exclusion limits orders of magnitude lower, excluding a large part of the parameter space.

Exclusion limits and discovery claims can be plotted together in plots such as the one in Figure 2.7, this provides a way of comparing results between different detectors and experiments. A WIMP discovery at a certain mass and cross section appears as an elliptical shape in these plots, often a 2 sigma confidence region is defined, outlining the 95% probability for the signal to originate from the encompassed area.

In case no discovery is made a limit can be set excluding a region of parameter space. Above this exclusion limit a signal would have been seen given the conditions of the experiment. Below this limit remains the unexplored parameter space.

Figure 2.7 shows some of the current limits as well as discovery claims by various experiments as presented in [12]. As of 2013 LUX [13] holds the record in excluded parameter space, probing a cross section below  $10^{-46}$  cm<sup>2</sup> for certain WIMP masses. This is about a factor of 2 better than the current XENON100 detector which held the record previously. The plot also shows the projected limit of XENON1T in case no discovery is made. XENON1T is currently under construction and will be commissioned at the end of 2015. It is expected that with only one week of data taking XENON1T will have reached the same sensitivity as the LUX 2013 results and after a two year run will be able to set a limit two orders of magnitude below the current lowest limit.

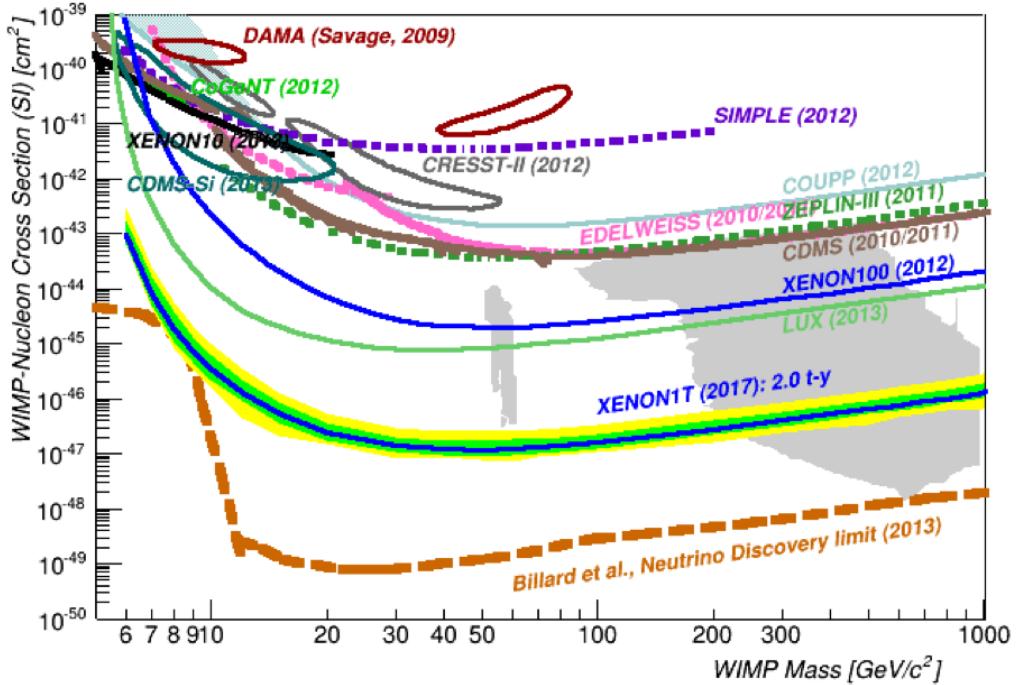


Figure 2.7: Exclusion limits and discovery claims from several dark matter direct detection searches. The current lowest limit is the one set by LUX (2013) (solid green line). The solid blue line at ‘XENON1T (2017)’ is the projected limit for the XENON1T detector. The elliptical shapes are discovery claims, except for the DAMA (2009) claim they have all been withdrawn [12].



## Chapter 3

# Direct Detection Experiments

Direct detection of dark matter involves recording the interaction of dark matter with ordinary matter through scattering. Measuring the recoil of a nucleus hit by dark matter allows the cross section and mass of the incoming particle to be determined. Many detectors have been designed and built to detect these type of interactions. The detection principle they use is based on several physical processes. Generated by the deposition of energy in the detector volume.

Deposition of energy by a particle in the detector can be in the form of a nuclear or electronic recoil. Meaning the incoming particle scatters directly on the nucleus of a target atom or interacts with the electron cloud. Due to the very low interaction probability or cross section of WIMPs we expect only nuclear recoils and no electronic recoils. Having a detector that can discriminate between the nuclear and electronic recoils will turn out to be crucial in the search for WIMPs because it allows analysts to cut away electronic recoil backgrounds caused by various sources.

The process to discriminate between the two interaction types is by measuring different signals generated by the recoil. Three physical processes can occur when energy is deposited in a material [14]:

1. Phonons, lattice vibrations at 10 meV/ph
2. Scintillation, emitted photons at  $\sim 1 \text{ keV}/\gamma$
3. Ionization, released electrons at  $\sim 10 \text{ eV}/e$

Both electronic and nuclear recoils can generate these three signals, but they do so in different proportions. Having a detector that can measure for example the phonon and scintillation signal allows the definition of a discrimination parameter. Also note that the energies involved are very small compared to the energies measured in a typical collider experiment. The typical recoil energies measured in XENON are on the order of a few keV, requiring the use of very sensitive detectors operating in a very low-background (i.e. shielded from any source of radiation) environment.

Figure 3.1 shows various experiments and what signals they use for detection. For the rest of this chapter I will focus on the XENON detectors which use both scintillation and ionization signals to discriminate between electronic and nuclear recoil interactions.

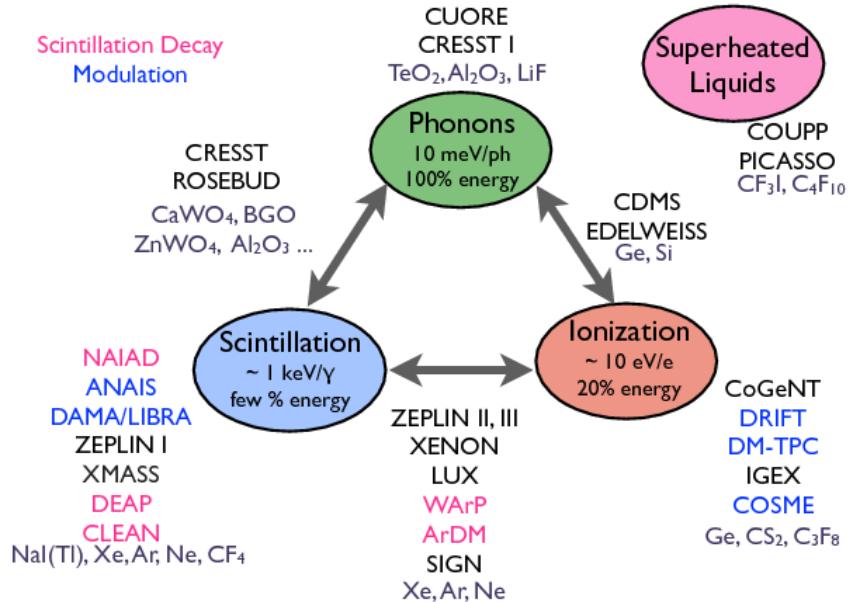


Figure 3.1: Discrimination triangle. Showing many direct detection experiments and what signals they use for detection, from [14].

### 3.1 Liquid xenon TPCs

A Liquid Xenon (LXe) Time Projection Chamber (TPC) is a detector that uses the scintillation and the ionization signal of a recoiling atom or nucleus in the detector. In this way an event can be classified as either a nuclear or electronic recoil. A TPC also allows one to reconstruct the position of the interaction in three dimensions, thus making it possible to use the self shielding properties of the very dense xenon liquid allowing further background subtraction.

A TPC consists of a volume with target material, either a gas, liquid or both. Several metal meshes at different heights in the TPC allow various electric fields to be applied. In the case of XENON the TPC is a dual phase TPC, meaning that part of the TPC is filled with liquid xenon and part with gaseous xenon. Figure 3.2 shows a schematic of a dual phase xenon TPC. The two panels on the right show the two different event signatures expected from an electronic and nuclear recoil event. The largest part of the detector is filled with liquid xenon and contains an electric field to drift the electrons to the top of the

detector containing the gaseous xenon. In this relatively small upper region of the TPC a strong electric extraction field is present to accelerate the electrons that have drifted up from the liquid. This acceleration of electrons through the gaseous xenon causes scintillation light to be emitted. Above and below the TPC are many photomultiplier tubes (PMTs) or light sensors that can detect even a single photon coming out of the TPC. They are arranged in a pattern that allows for the reconstruction of the position in the  $(x, y)$  plane (orthogonal to the electron drift direction). In places where there are no PMTs, such as the sides of the TPC, material with a very high reflectivity is used to reflect as much photons as possible. In this way as much light as possible can be detected in the TPC, allowing for a very low energy threshold. The reflectivity does have a negative effect on the position reconstruction accuracy as we shall see in Chapter 4.

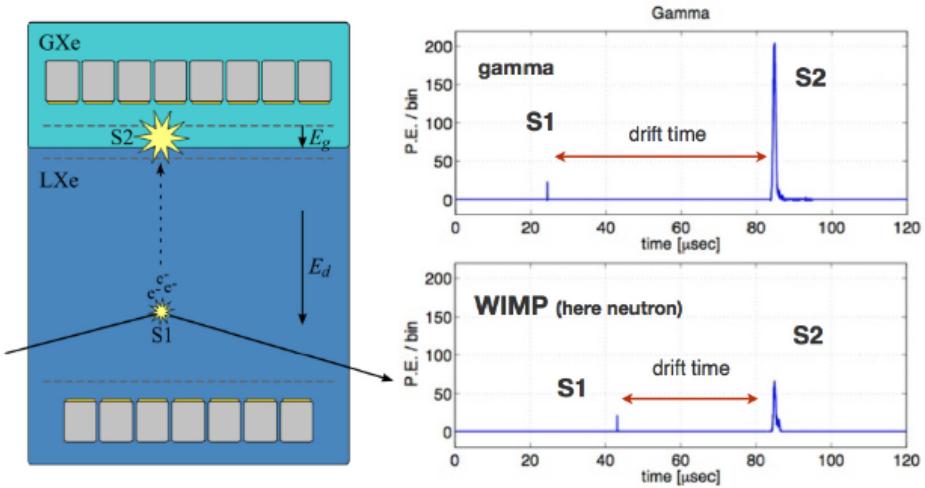


Figure 3.2: LXe TPC principle. The left panel shows a diagram of a LXe TPC. Shown are the liquid and gaseous region, the PMTs, the meshes and electric fields. The S1 and S2 labels show the places where these two signals are generated. The two panels on the right show the difference in event signature for a gamma (electronic recoil) event versus a neutron (nuclear recoil) event. The electronic recoil event has a larger S2/S1 ratio. From [14].

A particle traversing the TPC and interacting with a xenon atom will generate a scintillation and ionization signal. A phonon signal is also generated but is lost as heat (phonon signals are only used by solid detectors). The scintillation photons will travel through the detector and are detected by the PMTs. This is the first signal (light signal or S1). The freed electrons will drift towards the top of the TPC and are further accelerated in the gaseous region, resulting in a second light signal due to proportional scintillation in the gas phase (charge signal or S2). The ratio in energy between the S1 and S2 signals can be used to discriminate between electronic and nuclear recoil. The drift time, the time it takes the electrons to drift from the interaction vertex to the

gas-liquid interface, is a measure of the depth (or  $z$ -coordinate) of the vertex. The drift time separation of S1 and S2 is also illustrated in Figure 3.2. The position of the S2 signal in the plane orthogonal to the electron drift direction can be reconstructed by studying the hit pattern of the S2 in the top PMT array. The array is directly above the S2 signal and by measuring how much light each PMT has seen the  $(x, y)$  position of the S2 can be calculated. The location of the original vertex is assumed to have the same  $(x, y)$  coordinate because the electric drift field is uniform, causing the released electrons to drift vertically upward. Position reconstruction is a major part of this thesis and will be discussed in detail in Chapter 4.

So far the detection principle discussed can be applied to any noble gas dual phase TPC. There are several reasons why XENON chose to use xenon instead of for example the much cheaper argon. Xenon is a noble gas with very good scintillation properties. Most importantly it is transparent to its own scintillation light. The scintillation photons in xenon are not produced by excited xenon atoms decaying to a ground state but rather emitted when xenon dimers ( $\text{Xe}_2$ ) that were formed after the recoil decay. These dimers have different energy levels and thus emit photons that are not absorbed by normal xenon atoms.

Secondly xenon is also a very heavy element and in liquid form has a mass density of almost  $3000 \text{ kg/m}^3$ . This means that particles such as gammas only travel a short distance before they are stopped. This self shielding property of xenon is useful in separating WIMP signals from background. Since a WIMP is expected to only interact once within the detector its distribution in the TPC volume should be uniform. But background signals originating from outside the detector are expected to be stopped close to the edge of the TPC. This means that a fiducial volume can be defined where a certain percentage of background events is excluded. Xenon also has no natural radioactive isotopes causing background signals from the xenon itself.

### 3.1.1 XENON100

XENON100 [15] is the detector currently being operated by the XENON Collaboration and will soon be replaced by the new XENON1T detector. XENON100 has been in operation since 2008 and has set world leading exclusion limits on the WIMP cross section. XENON100 consists of a dual phase liquid/gas xenon TPC contained in a cryostat surrounded by shielding material. The number 100 in the name refers to the order of magnitude of the TPC volume. For the main science runs 161 kg of xenon was used. Similarly XENON1T will use several tonnes of xenon.

Figure 3.3 shows a diagram and photo of the detector and its shielding. The diagram shows the various detector components. The shielding is built up in layers, using both copper, polyethylene, lead and water to stop background signals from entering the detector. All these materials are screened for radio purity and low radioactivity lead is used as the inner layer of the lead shielding.

The detector is housed at the Laboratori Nazionali del Gran Sasso (LNGS) in Italy. This laboratory is located underground in the Gran Sasso mountains and houses many experimental setups requiring a low background, mostly neutrino and dark matter

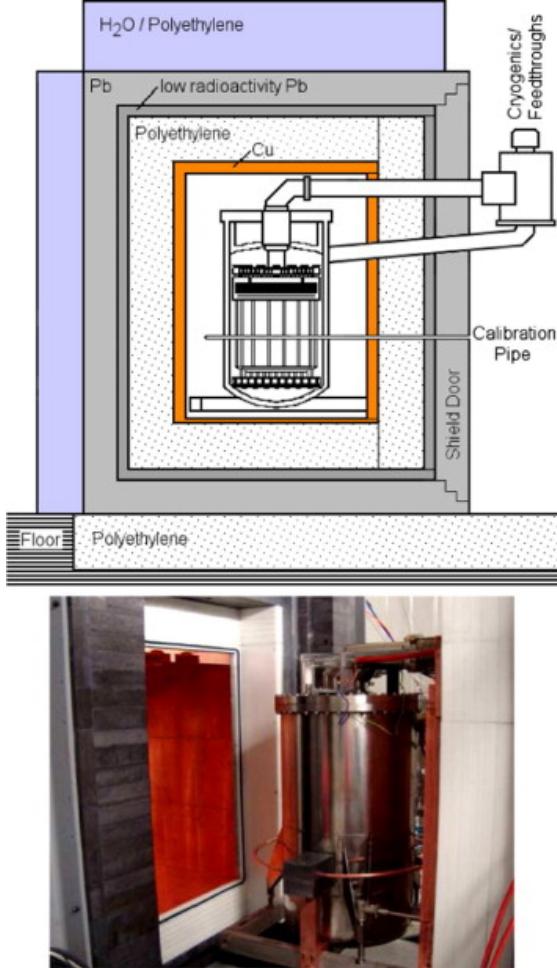


Figure 3.3: A Schematic of the XENON100 detector. In the very centre is the TPC, surrounded by the cryostat and various layers of shielding. A photo shows the cryostat outside the shield attached to the shield door on the right. Figure from [15].

physics experiments. Also XENON1T is being constructed there. The advantage of being underground is the reduced flux from cosmic particles. The muon flux in the underground lab of LNGS is about a factor  $10^6$  lower than the muon flux at sea level. Even though muons can be easily identified in the detector, neutrons induced by a muon interacting near the detector can not. Since neutrons also cause nuclear recoils in the detector that can be mistaken for WIMP signals it is very important to have as little neutron events as possible inside the TPC. Fortunately neutrons will often scatter multiple times in the TPC causing multiple S1 and S2 signals, this provides a way to cut these events from the data sample during analysis.

Compared to detectors commonly seen in particle physics XENON100 is very small. The TPC height and diameter are both 30 cm. Figure 3.4 shows the two PMT arrays

in XENON100, the difference in patterns is due to the fact that the top PMT array is used for position reconstruction while the bottom array is optimized for maximum light collection.

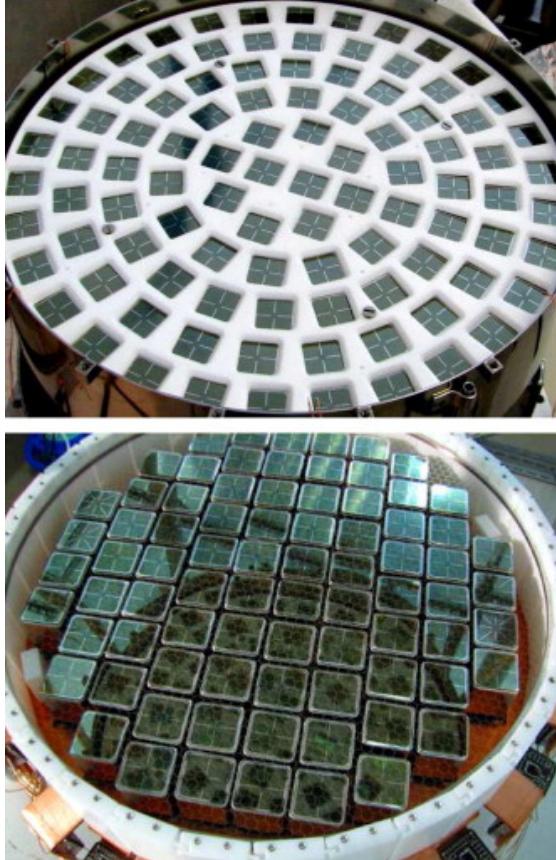


Figure 3.4: Two photos of the PMT arrays of XENON100. The top panel shows the top PMT array, the PMTs in this array have been ordered in a radial pattern to allow for better radial position reconstruction. The bottom photo shows the bottom PMT array, these PMTs are packed in a close grid pattern to maximize light collection efficiency. from [15].

### 3.1.2 XENON1T

Currently under construction XENON1T is the latest addition to the XENON series of detectors, it is also a dual phase TPC containing xenon. Compared to XENON100 it is roughly ten times larger in TPC volume, ten times more sensitive to light and features a water Cerenkov detector to veto muon events. This muon veto replaces the layers of shielding that are used by XENON100. It functions as a sub detector and is in fact a large light tight water tank monitored by multiple PMTs. Inside this tank the cryostat containing the TPC is suspended. A muon traversing the water tank will

generate Cerenkov radiation that will be seen by the PMTs. An event in the TPC can thus be vetoed if at the same time there is also signal in the muon veto.

Figure 3.5 shows a drawing of the water tank containing the cryostat and TPC.



Figure 3.5: A drawing of the water tank of XENON1T containing the support structure holding the cryostat with the TPC. The water tank is over three storeys high.

The TPC of XENON1T is about three times as high as the one in XENON100, giving a maximum drift length for the electrons of 1 meter. XENON1T uses round PMTs instead of square ones. Both PMT arrays are organized in a hexagonal pattern to ensure maximum light collection.



## Chapter 4

# Position Reconstruction

Position reconstruction is a crucial part of the dark matter analysis. It allows cutting away background events at the edges of the detector and making accurate position dependent corrections. The depth or  $z$ -coordinate of the interaction is calculated from the drift time and is independent from the position reconstruction presented here. To determine the depth of an event requires both an S1 and an S2 signal while an  $(x, y)$  position can be calculated for each S2 separately.

The drift time is a very good estimator for the depth and the  $z$ -coordinate can be determined much more accurately than the  $(x, y)$  position. Typically the depth of the interaction is known to  $\sim 0.3$  mm whereas the  $(x, y)$  position is known to  $\sim 3$  mm precision.

The S2 hit pattern is used to reconstruct the  $(x, y)$  position. It is a list of values for each S2 peak and tells how many photoelectrons (pe) were seen by which PMT. The number of pe is directly proportional to the number of photons entering the PMT by multiplying the number of pe by the quantum efficiency (QE) of the PMT. QEs in XENON100 are typically around 30 %.

Figure 4.1 shows a typical XENON100 hit pattern. The top panels show the hit pattern in both the top and bottom PMT array for a single S1 signal. In this case the S1 is only seen in the bottom array, this is mainly due to the reflectivity of the liquid-gas interface causing most of the S1 photons to be reflected down to the bottom array.

The bottom two panels show the hit pattern of an S2 signal. In the top array a signal is seen that is localized. From this pattern the position will be calculated. The S2 signal is typically much larger than the accompanying S1 signal and many of the photons generated also reach the bottom of the TPC. As can be seen in the figure, the bottom hit pattern of the S2 is much more uniform and the information about the position of the S2 is lost. This bottom hit pattern is used to calculate the total energy of the S2, the top pattern is not used for energy calculation since large S2s often saturate one or more PMTs in the top array.

Note that the outer ring of PMTs in the PMT patterns in Figure 4.1 are not inside the TPC but directly outside and thus optically separated. This region is also filled with liquid xenon and is used as an active veto. Events in the TPC are rejected if

a coincident signal is seen by these veto PMTs because such events are very probably background events. For example a neutron scattering once in the veto and once in the TPC.

In XENON100 several PMTs have broken since the detector was commissioned, this includes four PMTs in the top array. This has an effect on position reconstruction as will be shown in Subsection 4.3.

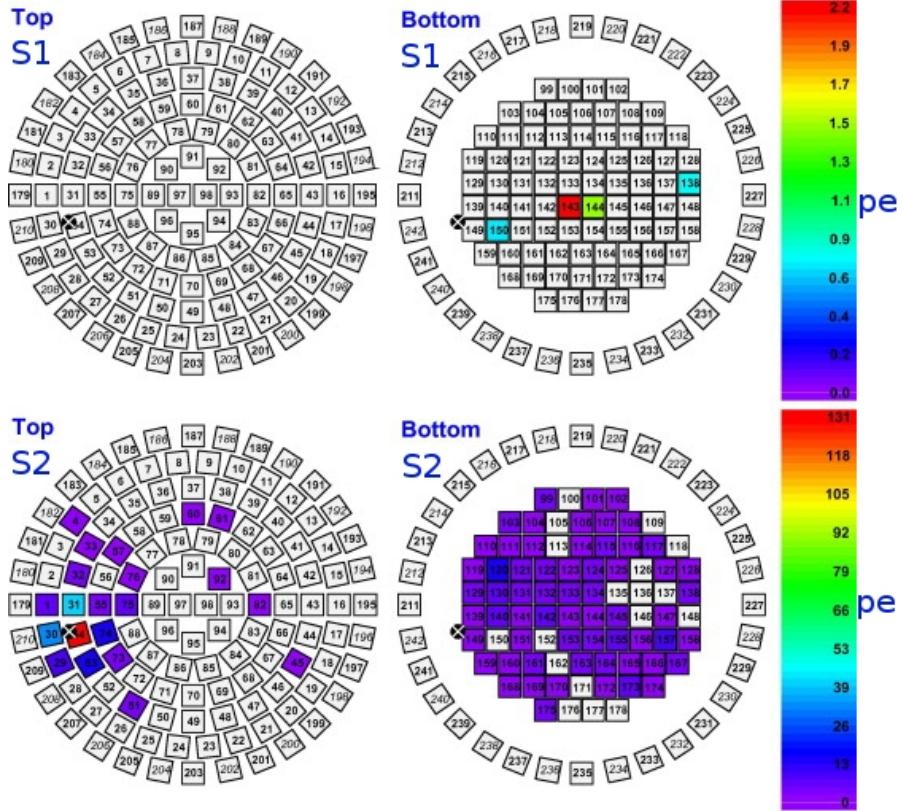


Figure 4.1: Typical hit pattern in XENON100. The left and right panels show the top and bottom PMT array respectively. The top panels show an S1 hit pattern and the bottom panels an S2 hit pattern. Colors indicate the amount of photoelectrons seen in the PMT. The black marker shows the position of the S2 top hitpattern as reconstructed by the Neural Network. From [15].

## 4.1 Reconstruction algorithms

Determining an  $(x, y)$  position for all S2 peaks found in an event can be done in several different ways. The most naive way of obtaining a position from an S2 hit pattern is by looking at the position of the PMT receiving the most signal, the ‘maximum PMT’ method. This method also has the lowest resolution since (in the case of XENON100)

it can only give 98 (minus 4 dead PMTs) unique positions, namely the positions of the live PMTs in the array. One step up is to compute the charge weighted sum of the hit pattern. This method has a resolution of  $\sim 1$  cm, also it is biased towards the centre of the TPC. Especially at large radii (close to the edge of the TPC) events are biased towards the centre since the hit pattern will always lie within the TPC. The main position reconstruction method used in XENON100 is a Neural Network, this has a resolution of  $\sim 3$  mm and no bias is seen in Monte Carlo events.

It would also be very helpful if a position reconstruction algorithm provided a way of quantifying the uncertainty in the reconstructed position. Having an accurate position reconstruction with a quantified uncertainty will (together with a z coordinate from the drift time) allow us to maximally use the fiducialization of the detector volume and position dependent corrections.

#### 4.1.1 Light-Collection Efficiency

Many position reconstruction algorithms that go beyond a simple weighted sum or maximum PMT location have to take into account the detector geometry. Since the TPC walls are made of Teflon to reflect as many photons as possible the geometry influences how photons are distributed.

To find the probability of a photon to reach a certain PMT from a certain origin Light Collection Efficiency (LCE) maps are used. These maps are the result of simulations of the TPC. Given the geometry of the detector a photon can be traced through the TPC. For each PMT in the top array an LCE map is made. For all  $(x, y)$  positions on a regular grid (grid spacing of 2.5 mm) at the gas-liquid interface photons are isotropically emitted and some hit the specific PMT.

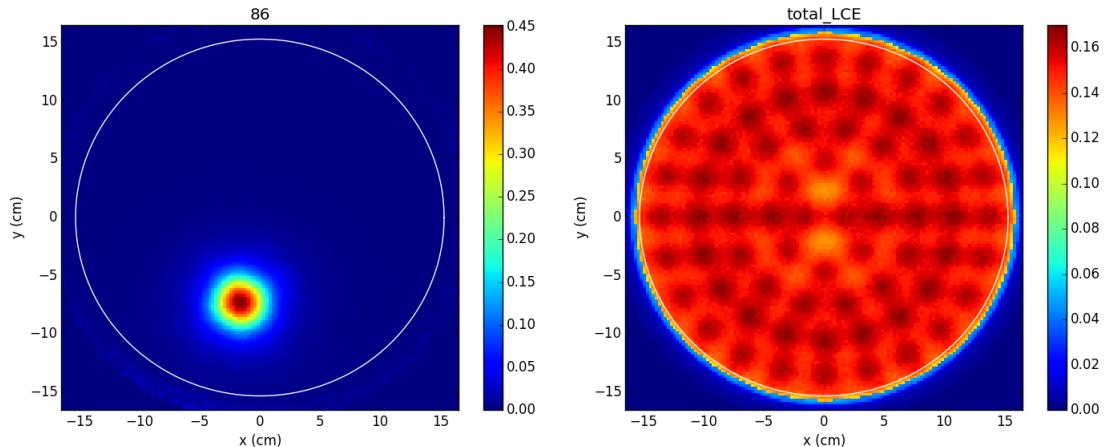


Figure 4.2: Left panel: Interpolated LCE map for PMT 86 showing the relative probability of a photon from position  $(x, y)$  to arrive at the PMT. Right panel: The interpolated total LCE map of all top PMTs combined showing the absolute probability of a photon from position  $(x, y)$  to be seen by any PMT.

Figure 4.2 shows an LCE map for both a specific PMT and the total LCE. The specific map shows the probability for a detected photon originating from position  $(x, y)$  to be seen by PMT 86. The probability is of course maximal when the photon is emitted directly below the PMT since it covers the largest solid angle. Note that due to the reflectivity of the TPC and the detector geometry the LCE map is non-trivial and has to be simulated. It is not possible to calculate the LCE analytically. The total LCE in Figure 4.2 shows that the probability for a photon produced at the gas-liquid interface to be detected by any top PMT is  $\sim 16\%$ . It also shows the positions of the PMTs in the top array where this probability is slightly higher. The individual LCE maps are used for the position reconstruction, since by sampling from the LCE maps an expected or Monte Carlo hit pattern can be constructed for an event originating from a certain position.

#### 4.1.2 Previous position reconstruction algorithms in XENON100

The data processor of XENON100, xerawdp (XENON raw data processor), uses three different algorithms for position reconstruction: a Neural Network (NN),  $\chi^2$  method and a Support Vector Machine (SVM). Each of these algorithms takes an S2 hit pattern as input and gives an  $(x, y)$  reconstructed position as output. The  $\chi^2$  method also needs a simulated Light-Collection Efficiency (LCE) map as input, as output this algorithm also provides a goodness-of-fit parameter called  $\chi^2$  to quantify how well the event was reconstructed. Cutting on high  $\chi^2$  events is one of the ways to exclude poorly reconstructed events from analysis as will be shown in Chapter 5.

Secondly the  $\chi^2$  algorithm can calculate this  $\chi^2$  value not only for its own reconstructed position but also for other positions, such as the position reconstructed by the Neural Network. It is in fact this value, the  $\chi^2$  of the NN, generated by the  $\chi^2$  method that is used as a quality parameter in the analysis since the NN itself only provides a  $(x, y)$  position.

The NN and SVM are both trained machine-learning algorithms. They are trained using the hitpatterns obtained in Monte Carlo simulations that account for the detector geometry and reflections of light inside the TPC (information that is also contained in the LCE maps). At a certain known  $(x, y)$  position isotropically emitted photons are simulated and the hit pattern is stored. These hit patterns can then be reconstructed by the various position reconstruction algorithms and the reconstructed position compared to the true position.

NN is the main algorithm used in analysis since it proved the most accurate when tested on MC data. Figure 4.3 shows a way to quantify the performance of the three algorithms on MC data. For each hit pattern the true position is known. The figure shows the euclidean distance between the true and reconstructed position for  $10^5$  events. Confirming that the Neural Network has a lower position reconstruction error on average. The accuracy of a Neural Network depends on its training. If the events used in training are not representative of the events that are later reconstructed the Neural Network will not work properly. In this sense such a method is a black box. Other methods such as the  $\chi^2$  method do have a physical motivation as explained in section 4.2.

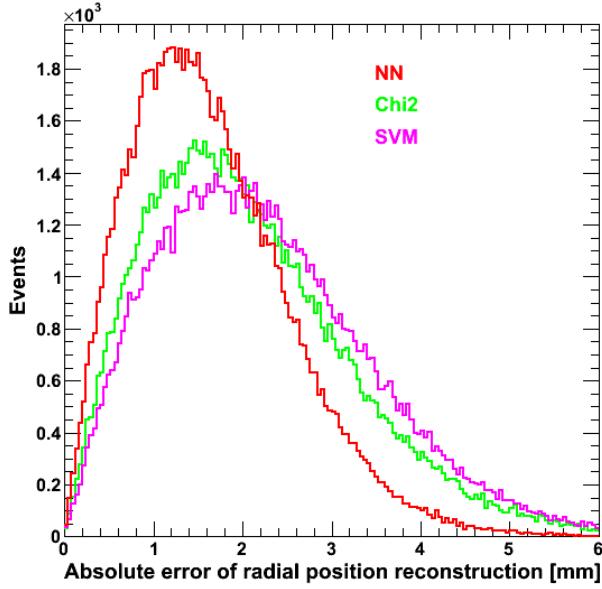


Figure 4.3: Comparison of reconstructed event positions to the MC truth position with three reconstruction algorithms implemented in xerawdp; the Neural Network,  $\chi^2$  method and Support Vector Machine. This result was the main motivation to choose the Neural Network as the main position reconstruction algorithm in XENON100.

#### 4.1.3 Position reconstruction in XENON1T

Position reconstruction in XENON1T is implemented as part of a new data processor called PAX; the Processor for Analyzing XENON. PAX is a software tool developed mainly at Nikhef, its main use is to process the raw data produced by the XENON1T detector. It produces output that is used by analysts in their physics analysis. PAX can also run different detector configurations such as a XENON100 configuration. This allows developers to check and validate the workings of PAX on XENON100 data. In chapter 6 I'll run a PAX analysis on XENON100 data to try and classify unusual events. PAX currently contains four position reconstruction algorithms: the maximum PMT method, weighted sum position, the XENON100 Neural Network and a new  $\chi^2$  algorithm (called  $\chi^2_\gamma$ ). The last two algorithms work only for the XENON100 configuration but the  $\chi^2_\gamma$  method can be run on XENON1T provided that an LCE map for its TPC is made. The goal of the new implementation of the  $\chi^2_\gamma$  method is to provide a physically motivated position reconstruction method for XENON1T data.

## 4.2 $\chi^2_\gamma$ Reconstruction

The  $\chi^2_\gamma$  method is based on the calculation of a test statistic given a hit pattern and simulated LCE map. The minimum function value will yield the reconstructed position

and the function value itself is a goodness-of-fit parameter called  $\chi^2_\gamma$ .

In short a  $\chi^2$  statistic will allow to compare data to model values, a lower  $\chi^2$  meaning a better fit. In our case we compare an S2 hit pattern (number of photons seen by each top PMT) to the expected pattern if the event originated from position  $(x, y)$ , finding the  $(x, y)$  which gives the best agreement yields the reconstructed position. However, an S2 hit pattern has to be modelled with Poisson statistics since there are generally a low number of counts in a PMT. In this case a  $\chi^2$  statistic will not accurately model the data as described by Mighell [16].

Fortunately a modified  $\chi^2$  statistic exists for low counts, this is the modified  $\chi^2$  distribution called  $\chi^2_\gamma$  as proposed by Mighell [16].

As noted before, the method uses an LCE map, this simulated map gives the probability for a PMT in the top array to receive a photon out of the total photons detected in the top array from position  $(x, y)$ ,  $p_i(x, y)$ , where  $i$  is the PMT index and runs from 1 to 98. One complication is that the LCE map is not continuous but simulated on a grid. To provide the reconstruction algorithm with a continuous LCE, interpolation is used.

Converting the number of photoelectrons seen by PMT<sub>*i*</sub> ( $N_{pe,i}$ ) to the number of photons seen by PMT<sub>*i*</sub> ( $n_i$ ) is done by using the Quantum Efficiency ( $QE_i$ ) of the given PMT.

$$n_i = \frac{N_{pe,i}}{QE_i}. \quad (4.1)$$

Summing  $n_i$  over all contributing top PMTs gives the total number of photons seen:

$$N_{ph}^{top} = \sum_{i \in PMTs} n_i. \quad (4.2)$$

Now the test statistic, from the  $\chi^2_\gamma$  distribution, is calculated as follows:

$$T(x, y) = \sum_{i \in PMTs} \frac{(n_i + \min(n_i, 1) - N_{ph}^{top} p_i(x, y))^2}{\sigma_i^2}, \quad (4.3)$$

with  $\sigma_i^2$  given by:

$$\sigma_i^2 = n_i^2 \left[ \left( \frac{\sigma_{QE,i}^2}{QE_i^2} \right)^2 + \left( \frac{\sigma_{gain,i}}{gain_i} \right)^2 \right] + N_{ph}^{top} p_i(x, y) + 1. \quad (4.4)$$

$\sigma_i^2$  includes the uncertainties in the PMT properties related to the gain (with uncertainty  $\sigma_{gain}$ ) and QE (with uncertainty  $\sigma_{QE}$ ).

The adaptation of equations 4.3 and 4.4 from the original distribution proposed by Mighell is made by Yuan Mei [17] who implemented the  $\chi^2$  method in XENON100.

To distinguish between the implementation in Xerawdp and this new implementation in PAX I'll use  $\chi^2$  to denote the Xerawdp (old) implementation and  $\chi^2_\gamma$  to denote the (new) implementation in PAX.

Note that the sum over PMTs in equation 4.3 involves only live PMTs. If a PMT is broken and cannot physically give a signal it will not be used in the computation of  $T(x, y)$  since the signal it would have given is not known. A live PMT that does not see any light however is used since seeing no light also provides information about the position of the event. This summing of the set of live PMTs also means that even if PMTs are added or removed per event the method still works (provided that all PMTs have an LCE map). This feature is not present in the Neural Network which is trained for a specific PMT pattern. If a PMT fails the network needs to be retrained. The  $\chi^2_\gamma$  method stills works even at positions close to the broken PMT.

The LCE map values are normalized on the top PMT array. If a PMT is added or removed the values are renormalized to ensure correct calculation of  $T(x, y)$ :

$$\sum_{i \in PMTs} p_i(x, y) \equiv 1. \quad (4.5)$$

$T(x, y)$  is a continuous function but is only defined inside the TPC. At radii larger than the TPC radius its function values are set to infinity to assure the minimizer only reconstructs inside the TPC ( $T(x, y) \equiv +\infty \mid x^2 + y^2 \geq R_{TPC}^2$ ).

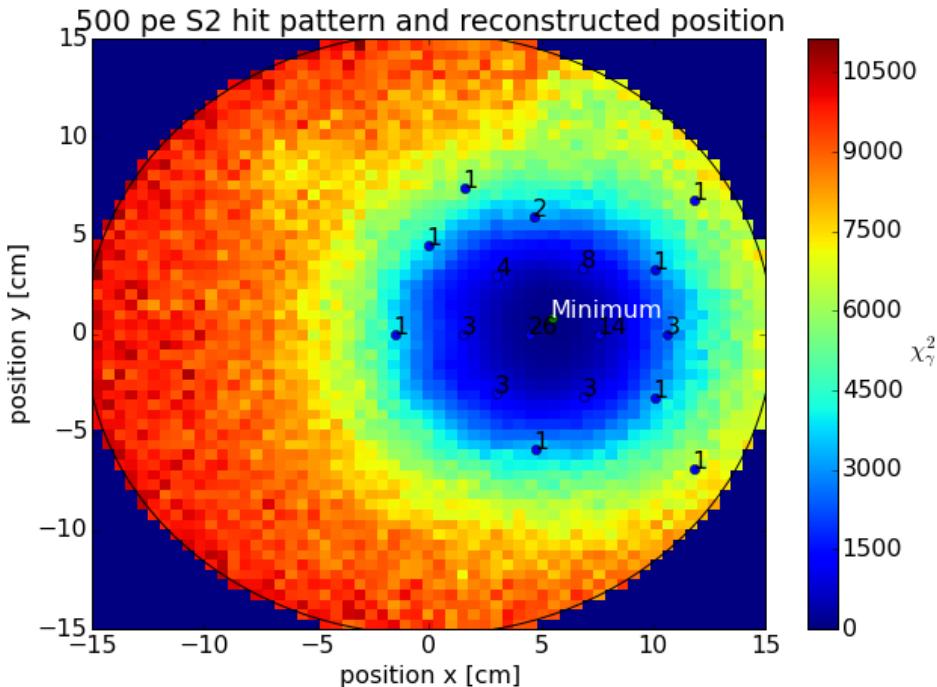


Figure 4.4: Calculation of  $T(x, y)$  for a typical hit pattern. Blue dots denote PMT locations, the numbers the amount of pe seen. The green dot labelled ‘Minimum’ shows the location of the minimum function value. The black circle denotes the TPC radius,  $T(x, y)$  is  $+\infty$  outside it, but for this plot set to 0.

Figure 4.4 shows an example of  $T(x, y)$  being calculated for a typical hit pattern. The blue points are PMT positions where photons are seen, the number denotes how many photoelectrons. The heat map shows the function values of  $T(x, y)$ . The green point at ‘Minimum’ denotes the minimum function value.

Finding the global minimum of  $T(x, y)$  given a certain hit pattern will yield the reconstructed position. The challenge for a good position reconstruction algorithm is to find the global minimum of this function accurately and fast.

Secondly, the value of  $T(x, y)$  at the reconstructed position is the goodness of fit parameter ( $\chi^2_\gamma$ ). As noted before also other reconstruction algorithms can use the  $\chi^2_\gamma$  method to calculate a goodness-of-fit for their reconstructed position. In this case only a single function evaluation has to be made, namely  $T(x, y)$  at the desired position. Suppose the Neural Network reconstructs an event at  $(x_{NN}, y_{NN})$  the corresponding  $\chi^2_\gamma$  value is then directly calculated as  $T(x_{NN}, y_{NN})$ .

The last feature of using a  $\chi^2_\gamma$  method is the ability to compute a confidence region at a desired confidence level. All values of  $T(x, y)$  that are increased by 2.3 describe the  $1\sigma$  confidence contour. For an increase of 6.18 the  $2\sigma$  confidence contour is obtained. Figure 4.5 shows an example of these confidence contours surrounding the reconstructed position. The heat map shows the  $T(x, y)$  function values zoomed in and centred on the reconstructed position. The minimizer start and end point are also plotted. The contours may have various shapes depending on the hit pattern and position in the TPC. The  $1\sigma$  contours are often roughly circular but the  $2\sigma$  contours often have more indistinct shapes.

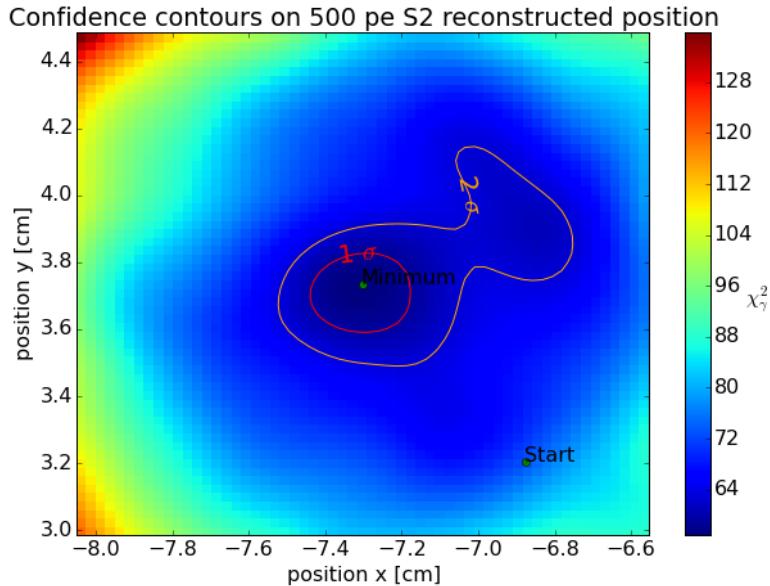


Figure 4.5: Example of the  $1\sigma$  and  $2\sigma$  confidence contours around a reconstructed position. The heat map shows the  $T(x, y)$  function values. The green dot ‘Minimum’ shows the found minimum value. ‘Start’ denotes the starting position of the minimizer.

#### 4.2.1 Description of the $\chi^2_\gamma$ position reconstruction in PAX

The  $\chi^2_\gamma$  implementation exists as a plugin in PAX that calculates the position of each S2 peak with an energy above a certain user-defined threshold. Besides the x and y position also the goodness-of-fit ( $\chi^2_\gamma$ ) and number of degrees of freedom (ndf) are stored.

The plugin can also append a  $\chi^2_\gamma$  value and ndf to already calculated ( $x, y$ ) positions by other reconstruction algorithms. Such as the position of the charge weighted sum of the hit pattern or the Neural Network.

PAX is written in the Python programming language and uses many high level modules (or libraries) for its computation such as SciPy [18]. The biggest advantage being that it allows for very fast code development that lets the developers focus on the physics instead of coding details. The biggest drawback being the sacrifice in speed.

To get continuous LCE map values SciPy interpolation is used (specifically a rectangular bivariate spline approximation). This ensures that  $T(x, y)$  is a continuous function.

To minimize the function the plugin uses a SciPy minimizer using the Powell [19] method. The starting position for the minimizer or seed is the position as reconstructed by the Neural Network. In case there is no Neural Network reconstruction present for the specific peak the charge weighted sum position of the hit pattern is used.

Currently the per-event errors or confidence contours are not calculated since they require too much computation time. A first step towards a solution is to calculate the radial confidence level only since the fiducial volume is rotation symmetric.

The plugin currently uses the XENON100 LCE maps and PMT mapping, these are loaded via a configuration file and can be swapped for XENON1T equivalents in the future. To use the  $\chi^2_\gamma$  method on XENON1T data, only the new LCE maps and PMT properties need to be changed.

### 4.3 Performance

Currently the plugin is optimized for accuracy, meaning that the minimizer makes several function evaluations ( $\sim 50$  per peak), meaning many LCE map lookups and many interpolations. Currently the LCE map lookups (and subsequent interpolation) are the dominant factor in the speed of the plugin, a reconstruction of one S2 peak now takes  $\sim 100$  ms.

The accuracy of the position reconstruction plugins is determined by reconstruction of Monte Carlo events. The position reconstruction error is the Euclidean distance between the known and reconstructed position. Figure 4.3 showed the accuracy for the different position reconstruction algorithms in Xerawdp, the data processor for XENON100. Now the same plot is made for the position reconstruction algorithms in PAX. The new implementation of  $\chi^2_\gamma$ , the XENON100 neural network (which will only work for a XENON100 setup since it is trained on a specific PMT pattern), charge weighted sum position and maximum PMT position are used. The Support Vector Machine method that is part of Xerawdp is not implemented in PAX and not used here. The Monte Carlo events are the same as the ones used in Figure 4.3 and were part of

the validation sample of the Neural Network.

Figure 4.6 shows the absolute position reconstruction error for each of these algorithms using PAX. The Neural Network and  $\chi^2_\gamma$  method both show very similar results to their implementations in the XENON100 data processor shown in Figure 4.3. This was expected from the Neural Network since its implementation is straightforward and its output deterministic. The fact that also  $\chi^2_\gamma$  shows similar behaviour is less trivial since its reconstruction depends on different interpolation and minimization algorithms. The weighted sum and maximum PMT methods show very poor results with errors extending beyond 1 cm.

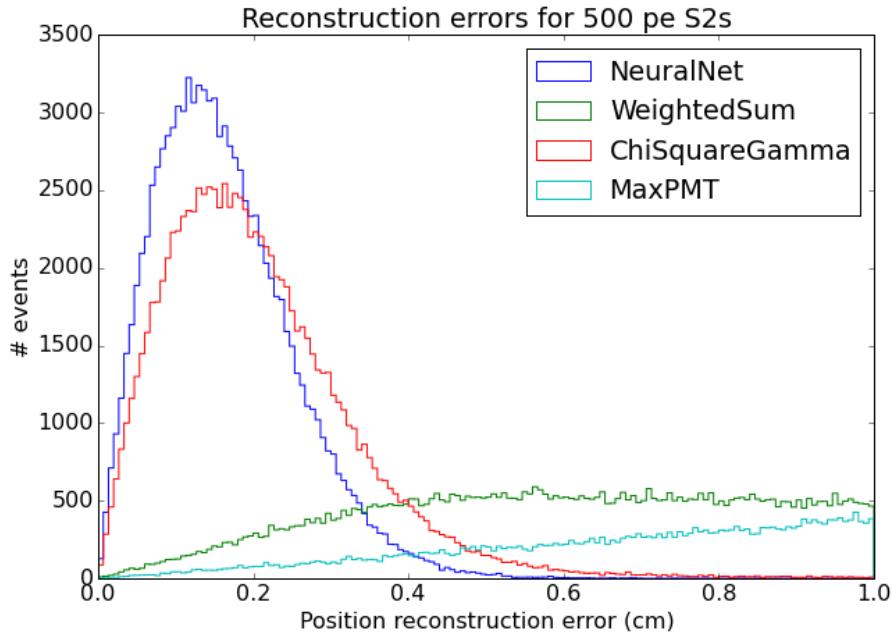


Figure 4.6: Absolute position reconstruction error for the four position reconstruction algorithms in PAX running a XENON100 configuration. This uses  $10^5$  full Monte Carlo events. These Monte Carlo events correspond to S2 signals of 500 pe and were also used to make Figure 4.3.

One question arising from Figure 4.6 is whether or not the  $\chi^2_\gamma$  method can ever be better than the Neural Network. Intuitively the answer is yes since the  $\chi^2_\gamma$  method is the model motivated by the underlying physics. Since the Neural Network is a function approximator it is expected to approximate (and thus never be better) than the  $\chi^2_\gamma$  method.

To test the above assumption a different Monte Carlo is used to test the position reconstruction algorithms in PAX. Figure 4.3 used Monte Carlo events that were the result of simulating photons in the detector. Since an LCE map exists the probabilities of photons arriving at various PMTs are known. This makes it possible to construct a hit

pattern by sampling from the LCE map. In this case the hit pattern would be the result of a perfect Monte Carlo simulation. Any reconstruction errors in this case are caused by intrinsic uncertainties in the reconstruction algorithm. Such as the interpolation of the LCE map, the accuracy of the minimizer or uncertainties in the PMT properties. Another benefit of sampling from the LCE map is that hit patterns for different energy S2s can be made, whereas the full Monte Carlo provides only one energy.

Figure 4.7 shows the result of reconstructing  $10^5$  MC hit patterns sampled from the LCE map, simulating 500 pe S2s similar to the energy of the full Monte Carlo events. Each hit pattern is reconstructed by the four reconstruction algorithms currently implemented in PAX. The new  $\chi^2_\gamma$  method, the XENON100 Neural Network, the charge weighted sum and the maximum PMT. As can be seen the  $\chi^2_\gamma$  method has the lowest error followed by the Neural Network. The weighted sum and maximum PMT methods do not come close to their result. By using the LCE MC hit patterns the  $\chi^2_\gamma$  method is better than the neural net. This result is more in line with the original assumption of the Neural Network being an approximator for the  $\chi^2_\gamma$  method, but is hard to explain after having seen Figure 4.3. One explanation could be that the hit pattern sampled from the LCE map is the result of a perfect Monte Carlo. It results in lower errors since it corresponds to an idealised scenario.

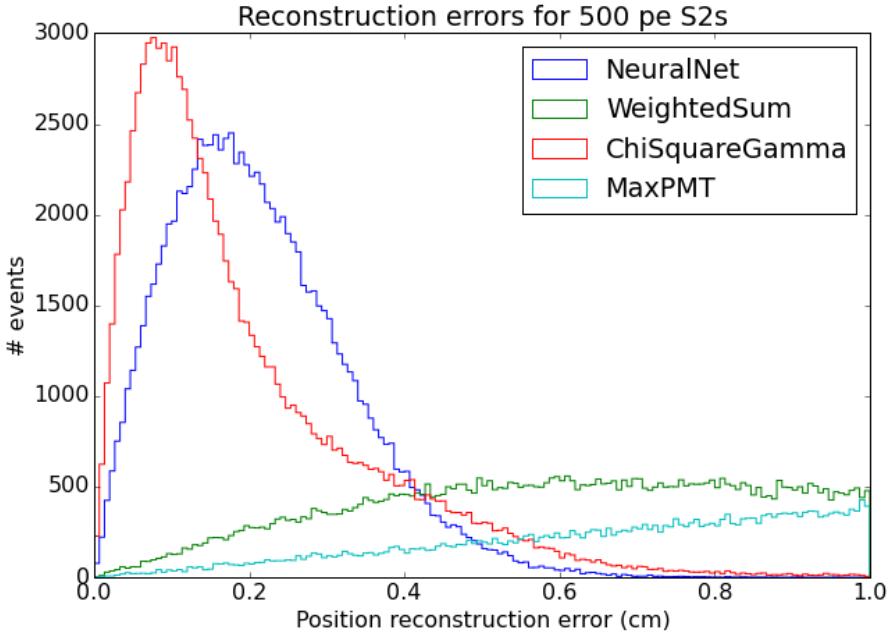


Figure 4.7: Absolute position reconstruction error for all position reconstruction algorithms in PAX running a XENON100 configuration. Using  $10^5$  Monte Carlo hit patterns sampled directly from the LCE maps corresponding to 500 pe S2 signals. The different Monte Carlo results in a much lower position reconstruction error compared to Figure 4.6.

When XENON1T is nearing completion, LCE maps will be made and Monte Carlo studies performed by the collaboration. This allows us to test the various reconstruction algorithms as was done in XENON100. For the remainder of this chapter the LCE map Monte Carlo will be used.

### 4.3.1 Local position reconstruction errors

Figure 4.7 also shows that the tail of the  $\chi^2_\gamma$  reconstruction is higher at very large errors than the tail of the Neural Network reconstruction. This behaviour can be explained by looking at Figure 4.8 and Figure 4.9. These two figures show the same  $10^5$  reconstructed S2 peaks (of 500 pe) as in Figure 4.7. Each bin shows the average position reconstruction error of events in the bin (binned by true position). The arrow in each bin points to the mean reconstructed position of the events in the bin. In this way any bias can be seen. The PMT positions are marked by green squares, dead PMTs are marked in red.

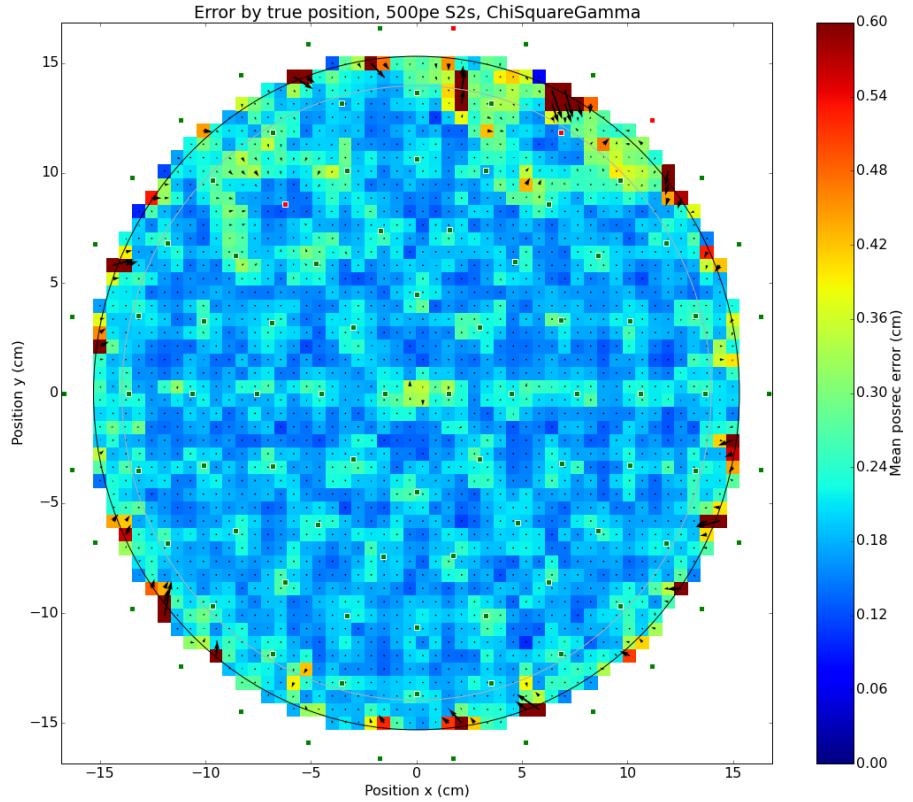


Figure 4.8: Errors on reconstructed positions of the  $\chi^2_\gamma$  method. Hit patterns of  $10^5$  S2 peaks of 500 pe are simulated by sampling from the LCE map and reconstructed. Events are binned in the  $(x, y)$  plane, colors show the mean position reconstruction error in the bin. Arrows are the mean bias of the events in the bin. PMTs are marked by green squares, dead PMTs in red. The black circle denotes the TPC radius, the grey circle denotes the maximum extend of the fiducial volume used by XENON100.

Figure 4.8 shows the results for the  $\chi^2_\gamma$  method. Across most of the  $x, y$ -plane small errors between 1-3 mm are seen. Also at the positions of the four dead PMTs events are properly reconstructed. Near the edge of the TPC larger errors are seen. These explain the tail of the distribution in Figure 4.7. These edge effects are caused by the minimizer and can occur when the minimizer tries to reconstruct outside the TPC radius. Near all live PMTs the reconstruction error is slightly larger. This may be an effect due to the physical size of the PMTs ( $\sim 2.5$  cm).

Figure 4.9 shows the results for the Neural Net. The errors are higher but more uniform. The Neural Network also reconstructs near dead PMTs since it was retrained for this specific PMT pattern. In the majority of the  $(x, y)$  plane however, position reconstruction errors of the  $\chi^2_\gamma$  method are smaller than those reported by the Neural Net in Figure 4.9 favoring the  $\chi^2_\gamma$  method. The Neural Network has slightly larger errors overall and was retrained to cope with the broken PMTs.

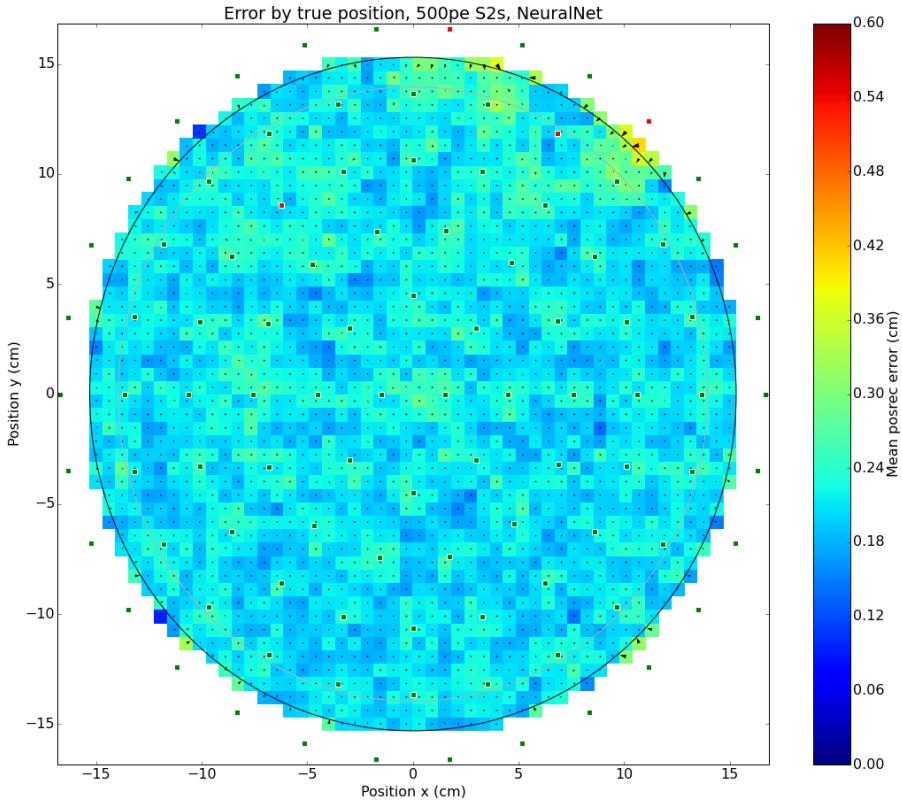


Figure 4.9: Errors on reconstructed positions of the Neural Network. Same labels etc. as in Figure 4.8.

Figures 4.8 and 4.9 show the errors from the true positions. Appendix A shows several more of these plots for different energy S2 peaks and for the other two algorithms used.

At higher energies both the Neural Network and  $\chi^2_\gamma$  method become more accurate. At lower energies they are less accurate. In all cases the  $\chi^2_\gamma$  method is the most accurate

of all four algorithms.

As another performance check the so called leakage events are investigated. A leakage event is defined as an event that originates from one side of the fiducial volume and is reconstructed on the other side. Leak-in meaning an outside event is reconstructed inside. Leak-out is the opposite process. Figure 4.10 shows the leakage for the reconstructed events by the  $\chi^2$  method. In this case  $\sim 7\%$  of the events that originate outside the fiducial volume are reconstructed inside. Table 4.1 shows the percentages for all algorithms used. The  $\chi^2$  and Neural Network showing similar figures. As can be seen the inward bias of the weighted sum reconstruction is so large that almost all events are reconstructed inside the TPC. Even more than leak in with the Max PMT method.

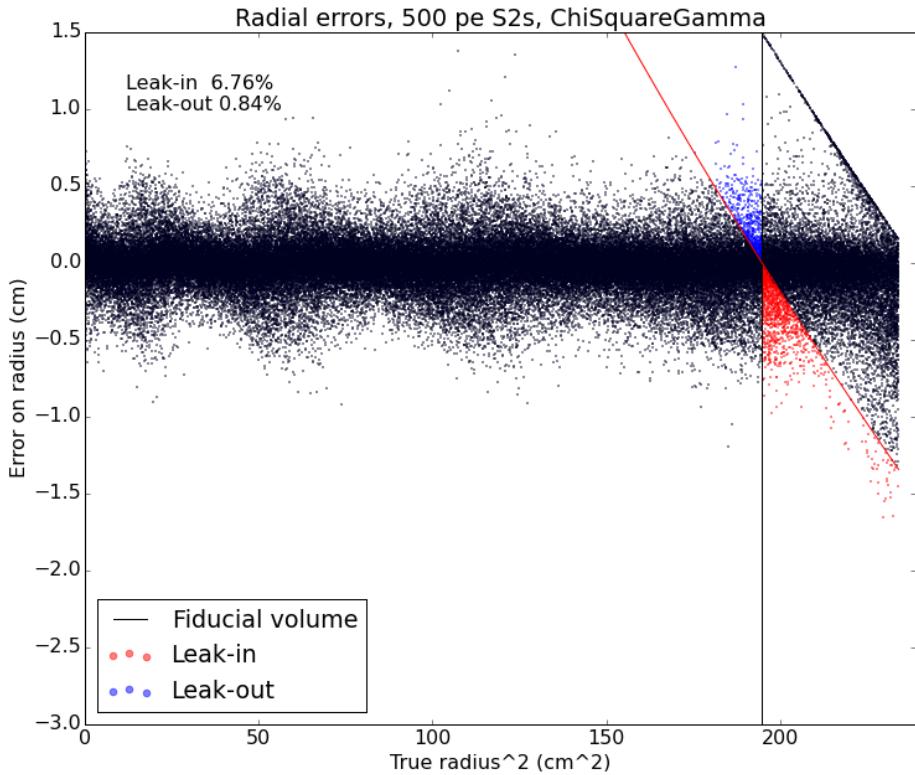


Figure 4.10: Radial errors on  $10^5$  reconstructed S2 peaks of 500 pe by the  $\chi^2$  method. The vertical black line represents the maximum radial extent of the XENON100 fiducial volume. The events marked in red are events that occur outside of the fiducial volume but are reconstructed inside. Events marked in blue occur inside the fiducial volume but are reconstructed outside. The red line being the reconstructed fiducial volume line.

The final feature of the  $\chi^2$  method is the determination of confidence contours per reconstructed position. Figure 4.5 showed that these contours can be roughly circular but also that they can have a more indistinct shape. To provide a measure of how large

Algorithm	Leak-in (%)	Leak-out (%)
$\chi^2_\gamma$	7	1
Neural Network	6	1
Weighted Sum	97	0
Max PMT	92	0

Table 4.1: Percentage of leakage events for the different position reconstruction algorithms.

these confidence regions are the area is calculated for  $10^3$  events. Both  $1\sigma$  and  $2\sigma$  areas are shown in Figure 4.11. The figure shows how the area of the contours is roughly constant, slightly increasing towards larger radii.

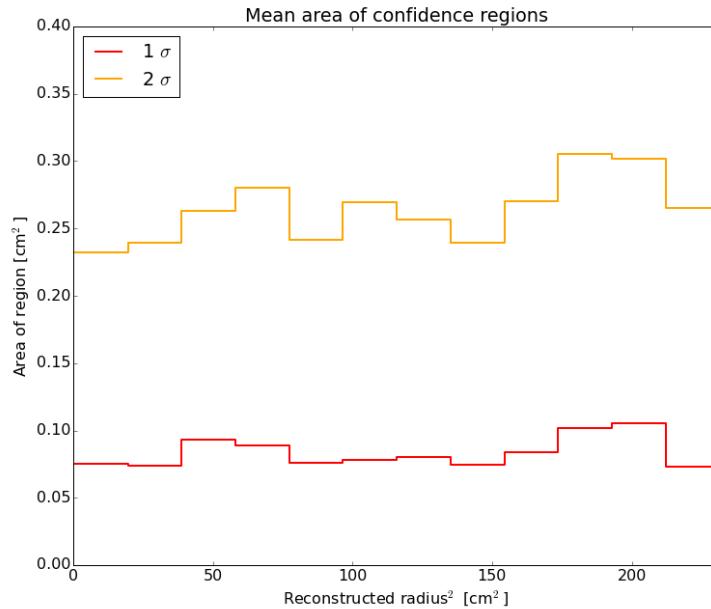


Figure 4.11: The area of the  $1\sigma$  and  $2\sigma$  contours as function of the radius squared. Results of  $10^3$  events (500 pe S2s).



## Chapter 5

# Data Quality Analysis

Chapter 4 explained in detail how position reconstruction works. This chapter shows several ways of using the position of events for data quality. It contains the work on three data quality cuts for the XENON100 run 14 analysis. Run 14 was started on May 2014 and will probably continue until the detector is decommissioned. Run 14 is not a science run to search for dark matter but is a calibration run used for studies on detector response and new sources.

XENON100 is essentially a counting experiment. The detector, shielded against as much radiation as possible, is left to record data for a certain period of time. The events that are found are (mostly) background events from various sources that are still present in or around the detector. By placing radioactive sources of known type and strength close to the detector calibration data can be collected. This calibration data is then used to quantify how the detector responds and which physics events generate what signals.

Background models can be made for the electronic recoil events by using data from gamma sources. Neutron sources provide a model for nuclear recoils modelling WIMP interactions. The most important sources are:

1. Americium-beryllium (AmBe): a neutron source that is used to model the WIMP interactions. The emitted neutrons have a typical energy of a few MeV, extending up to 10 MeV.
2. Cobalt-60 ( $^{60}\text{Co}$ ): a  $\beta/\gamma$  source modelling the electronic background. Emitted photons have energies of  $\sim 1$  MeV.
3. Yttrium-beryllium (YBe): a low energy neutron source first used in run 14. The emitted neutrons have exactly 152 keV energy. This source is especially useful to calibrate the detector at very low energies. The data is used in the low-mass WIMP analysis.

In the following sections AmBe data is often used to define quality cuts that can later be used in analysis. A cut selects events based on a certain requirement. Such as the S2 energy that needs to be in a certain range for the event to be accepted. For more complicated cuts a cut parameter is defined (often as a function of other parameters).

A data quality cut selects good events and rejects bad events. ‘Good’ and ‘bad’ have different meanings depending on the cut. In the case of the first two cuts in this chapter a bad event means an event that has a poor  $(x, y)$  reconstruction. In the third case it means an event of which the main S2 has an unusually high or low asymmetry.

A sample of AmBe data, often on the order of  $10^6$  events, is used to define a quality cut. To these events a certain pre-selection is applied to remove nonphysical events from the sample. The following pre-selection is applied to all data used in this chapter:

1. Radial or fiducial volume cuts: The outermost regions of the TPC are excluded. The radial cut removes events at a radial position larger than 150 mm excluding the outermost 3 mm of the TPC. The fiducial volume cut defines an ellipsoidal inner volume containing 34 kg of liquid xenon. This fiducial cut is also used in the dark matter analysis.
2. S2 Energy range: The main S2 of the event has to have an energy of at least 150 pe and not more than 5000 pe.
3. S1 coincidence: The main S1 of the event has to be seen by at least two PMTs.
4. S1 noise and S1 entropy: Cut on noisy S1 peaks based on per-PMT noise.
5. S2 Top signal: The main S2 of the event has to deposit energy in the top PMT array.
6. Multiple scatter cuts: Remove an event if the second largest S2 has an energy above a certain threshold.

Using the event sample obtained after applying the pre-selection a two dimensional histogram is made of the cut parameter versus the main S2 total energy. This histogram is than binned in energy bins of 100 pe. Each bin is fitted using an appropriate function. A quantile is than calculated given the desired acceptance of the cut. This results in the cut value (value above which the event is cut) for the specific energy bin. After repeating the process for all energy bins a continuous cut value can be constructed by fitting the cut values.

## 5.1 $\chi^2$ Data quality cut

The  $\chi^2$  referred to in this chapter is different from the  $\chi^2$  used in Chapter 4. In this chapter the  $\chi^2$  for a certain S2 peak is defined as the goodness-of-fit of the position found by the Neural Network. This value is then divided by the top array coincidence (the number of PMTs in the top array contributing to the signal) minus one. This is the reduced  $\chi^2$  or  $\chi^2/ndf$  ( $ndf$  = number of degrees of freedom).

$$\chi^2/ndf \equiv \frac{\text{Neural Network position goodness of fit}}{\text{Number of contributing PMTs in top array} - 1} \quad (5.1)$$

The goal of this data quality cut is to exclude poorly reconstructed events from the data sample. This is done by defining a certain maximum reduced  $\chi^2$  value above which an event is cut. Events with a high reduced  $\chi^2$  are often multiple scatter events that can not be properly reconstructed by the position reconstruction algorithms.

The cut value is defined in such a way as to cut a fixed percentage of the event. The cut is defined to have an acceptance of 99.5%, meaning the worst 0.5% of the sample is cut. This data quality is cut also dependent on the energy of the S2. The need for defining an energy dependent cut in this way arises from the fact that position reconstruction is less accurate for lower energy events. To keep a fixed percentage of the events, the acceptance percentage, the cut value has to be higher at the low energy end. In earlier runs several  $\chi^2$  cuts have been used with different cut values, table 5.1 gives an overview. For the standard analysis a constant cut value is used, the S2-only analysis (a low-mass dark matter analysis, using only S2 signals) first used an S2 energy dependent value, mainly because the S2 range used is much smaller and this is the region where the  $\chi^2$  varies most. Above 1000 pe the  $\chi^2$  is mostly flat with only a small linear dependence.

Analysis	Cut Value	Acceptance
Run 08 and 10	$\chi^2/ndf < 7$	99.2 %
Run 12	$\chi^2/ndf < 6$	99.44 %
S2-Only	$\chi^2/ndf < 1.9 - 4.3 \log E_{S2} + 0.73\sqrt{E_{S2}} - 0.0085E_{S2}$	98 %

Table 5.1: Previous  $\chi^2$  cuts,  $E_{S2}$  being the total energy of the main S2 peak in the event.

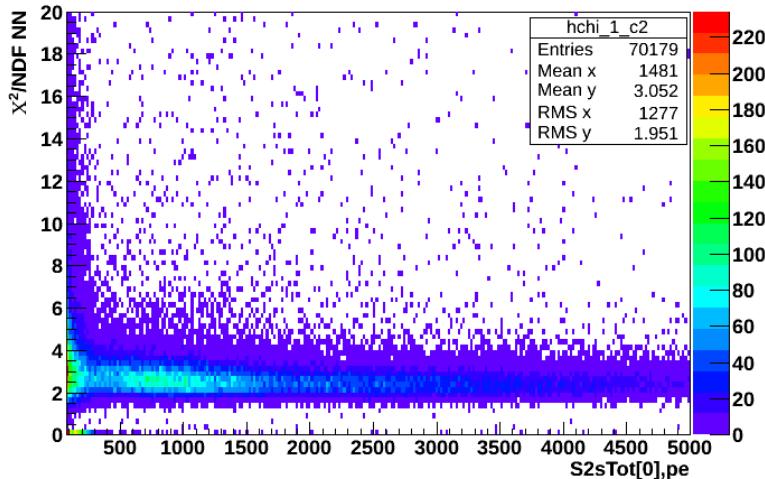


Figure 5.1: Goodness-of-fit of the Neural Network reconstructed position expressed as  $\chi^2/\text{ndf}$  versus the main S2 energy ( $\text{S2sTot}[0]$ ). The color axis shows the number of counts. The main population rises at low energy. At  $\chi^2 = 0$  a second population is visible.

Figure 5.1 shows the expected distribution. The average  $\chi^2/\text{ndf}$  is roughly constant

for large energies and only rises at low energy. The population at very low energy and  $\chi^2/ndf = 0$  is unexpected. Figure 5.2 shows the  $(x, y)$  position of these events. These events practically all originate from the same  $(x, y)$  position indicating a hotspot at  $(-2.5 \text{ mm}, -103.5 \text{ mm})$ . They can be explained as an artefact of the neural net position reconstruction and also affect other cuts as will be shown in section 5.2.

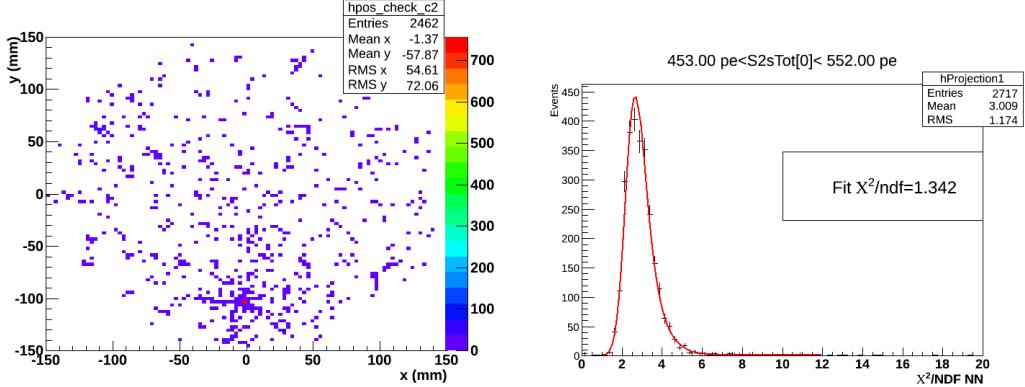
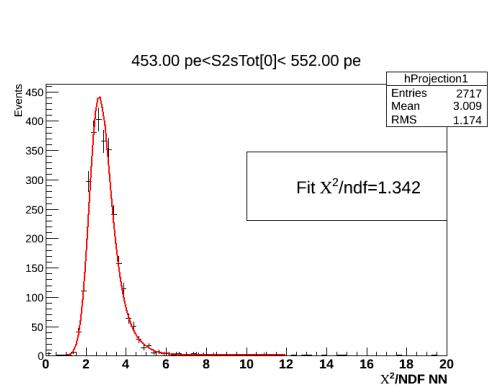


Figure 5.2: (x, y) position of all events in Figure 5.1 with  $\chi^2/ndf < 1$ . The color axis shows the number of counts. Events in the whole TPC are shown (no fiducial volume cut). A hotspot can be seen at position  $(-2.5 \text{ mm}, -103.5 \text{ mm})$ , within the fiducial volume.

Figure 5.3: Histogram of  $\chi^2/ndf$  in energy range  $453 \text{ pe} < \text{S2 Energy} < 552 \text{ pe}$ . Data is fitted with a Gaussian convoluted with an exponential.



Next the histogram in Figure 5.1 is sliced vertically and the distribution in each energy bin is fitted using a Gaussian convoluted with an exponential. The first bin (lowest  $\chi^2$  value) of each slice is not used in the fit to ensure that the result is not influenced by the  $\chi^2 = 0$  population. The fit function describes the data well, with a  $\chi^2/ndf$  (of the fit to data) all below 2. Figure 5.3 shows an example slice from Figure 5.1.

For each slice the 99.5% quantile level from the fit is calculated and this cut value, together with its error, is put into a new histogram. The error is estimated by propagating the errors of the fit. Figure 5.4 shows the resulting cut. At higher energy a more flat cut value of  $\sim 4.5$  is seen, at low energy the cut value rises because the position reconstruction is less accurate here. For S2 energies above 5000 pe the cut value at 5000 pe can be used.

The fit function in Figure 5.4 is the following:

$$f(x) = 32.1 - 5.79\log(x) + 0.488\sqrt{x} - 0.00266x. \quad (5.2)$$

Providing the following data quality cut for an event with main S2 energy ( $E_{S2}$ ):

$$\chi^2/ndf < 32.1 - 5.79\log(E_{S2}) + 0.488\sqrt{E_{S2}} - 0.00266E_{S2}. \quad (5.3)$$

A last check is made to see what events are in the low S2 energy range and have a high  $\chi^2/ndf$  value, since the  $\chi^2$  cut is a cut on poorly reconstructed events, the hit patterns

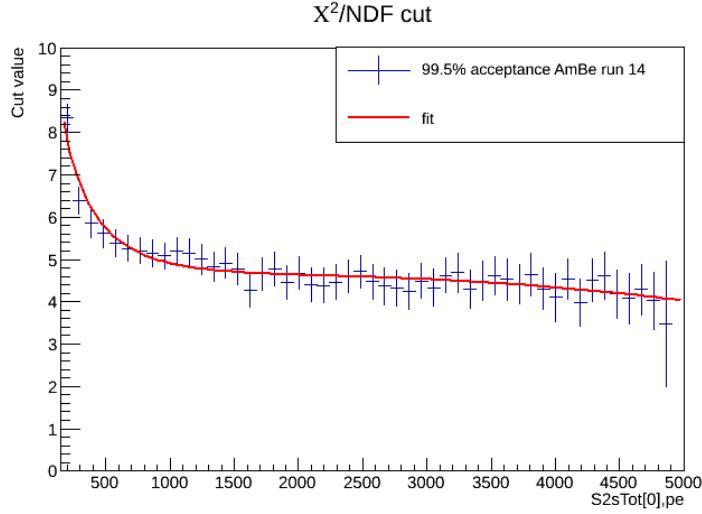


Figure 5.4: Cut values versus S2 total energy ( $S2sTot[0]$ ). The data points show the 99.5% quantile levels for each energy bin. The fit provides a continuous cut value. The fitfunction is an ad hoc function used to describe the data:  $f(x) = 32.1 - 5.79\log(x) + 0.488\sqrt{x} - 0.00266x$

are checked. The event window is  $150 \text{ pe} < S2 \text{ Energy} < 250 \text{ pe}$  and  $6.5 < \chi^2/ndf < 7.5$ , these events have a high  $\chi^2/ndf$  but are still below the cut value and should thus be properly reconstructed for this cut to be useful. Figure 5.5 shows two hit patterns. When looking at the top hit patterns (left panels) the black marker (reconstructed position) is in agreement with the position of the event as seen by eye. The energy of the top hit patterns however is much lower than we would expect from S2 signals in the given energy range. This is explained when we look at the bottom hit patterns. Both events in Figure 5.5 show a very large contribution from PMT 167 and, to a lesser extent, PMT 166 to the event. During run 14 these PMTs were found to be very noisy. In this case they contribute energy to a low energy S2 causing its total energy to be larger and polluting the data sample.

### 5.1.1 Redefining the cut

One way of coping with noisy PMTs in the bottom array is by redefining the cut versus the S2 top energy alone. Repeating the entire analysis in exactly the same way, the following cut is obtained.

$$\chi^2/ndf < 10.07 - 0.7568\log(E_{S2,Top}) - 0.02179\sqrt{E_{S2,Top}} + 0.0005446E_{S2,Top}. \quad (5.4)$$

For energies larger than 2350 pe, the constant cut value of 4.427 is used. Figure 5.6 shows the difference with the previous cut. At low energy the cut value does not rise as steeply as it did when taking the total energy of the S2 which is effectively shifted to

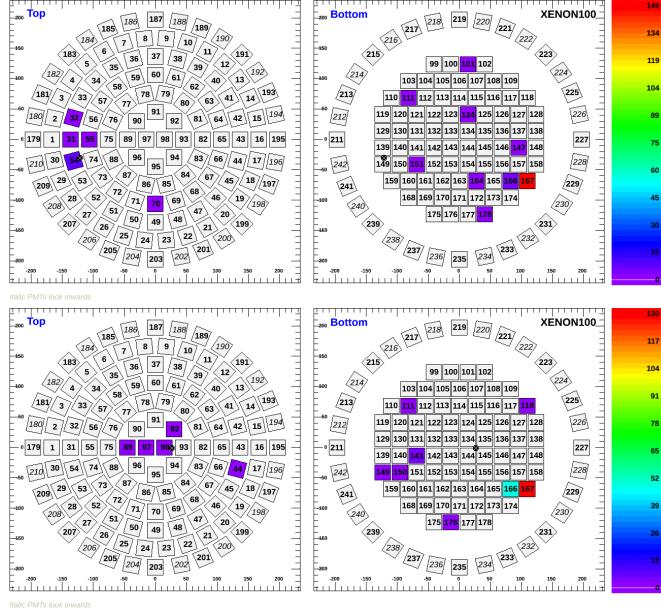


Figure 5.5: Hit patterns of high  $\chi^2/ndf$ , low energy events. Color axis shows the number of pe seen. These events are properly reconstructed S2s with a large energy contribution from a noisy PMT.

the left. Concluding that the pollution caused by events with high energy contributions from noisy PMTs is gone.

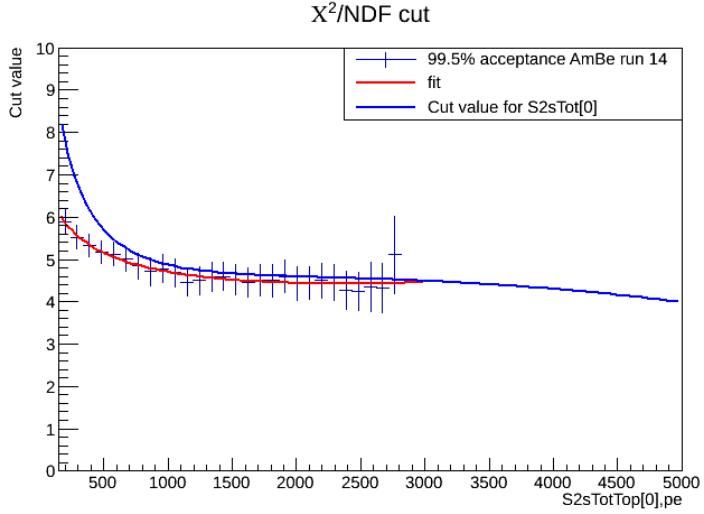


Figure 5.6: Cut values versus S2 total top energy ( $S2sTotTop[0]$ ). The redefined cut is lower because the event sample is cleaner. There are no noise contributions from the bottom PMT array.

## 5.2 Position discrepancy cut

The position discrepancy cut is another way to cut on poorly reconstructed events. This cut removes an event if the various position reconstruction algorithms disagree on the position of the S2 by more than the cut value. The cut parameter, called ‘posparam’, is defined as:

$$\text{posparam} \equiv \sqrt{(x_{NN} - x_{\chi^2})^2 + (x_{NN} - x_{SVM})^2 + (y_{NN} - y_{\chi^2})^2 + (y_{NN} - y_{SVM})^2}. \quad (5.5)$$

With NN, Chi and SVM being the Neural Network,  $\chi^2$  method and Support Vector Machine respectively.

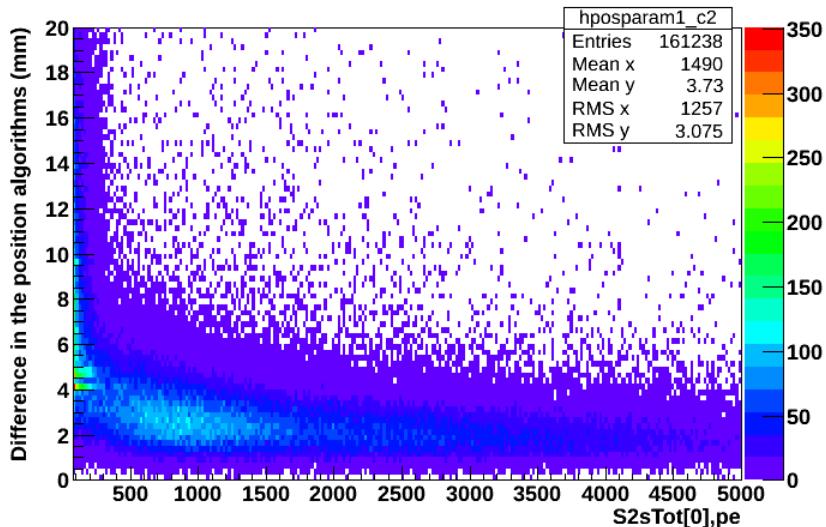


Figure 5.7: Position discrepancy versus main S2 energy ( $S2sTot[0]$ ) in run 14 AmBe data. The color axis shows the number of counts.

The same event sample is used as before. Figure 5.7 shows posparam versus the main S2 energy. At lower energies the position discrepancy rises as lower energy S2s are harder to reconstruct. Some artefacts are seen in the lowest energy bin at a posparam value of  $\sim 4.5$ , these events show up as the bright green and red bins in Figure 5.7. Figure 5.8 shows the distribution of posparam in the first energy bin (158-257 pe). The distribution is fitted with an exponentially modified Gaussian and shows the outliers seen in Figure 5.7. These events originate from the same detector hot spot as seen in the  $\chi^2$  cut analysis. The  $(x, y)$  position of these outlying events are shown in Figure 5.10.

Again the solution lies in the redefinition of this cut with respect to the S2 energy in the top array only. This is a more natural choice since the position reconstruction only relates to the top hit pattern and thus only the energy deposited in the top array. Since all noisy PMTs are in the bottom array we will have a clean sample. There will be no

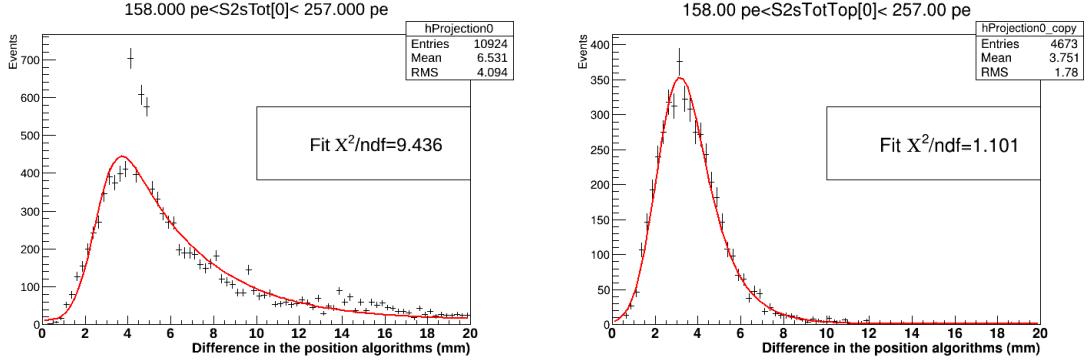


Figure 5.8: The first energy bin of Figure 5.7, showing the posparam distribution fitted by an exponentially modified Gaussian. Three bins are clear outliers, their  $(x, y)$  position is plotted in Figure 5.10.

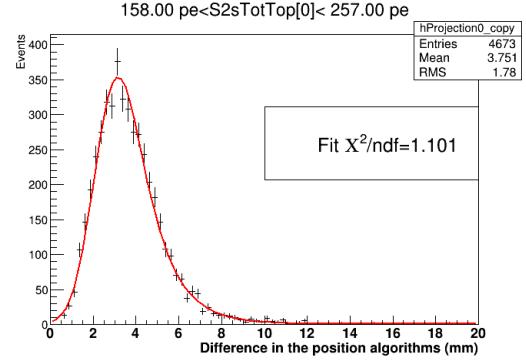


Figure 5.9: The first energy bin of Figure 5.11. A much cleaner sample than in Figure 5.8, no outliers are seen.

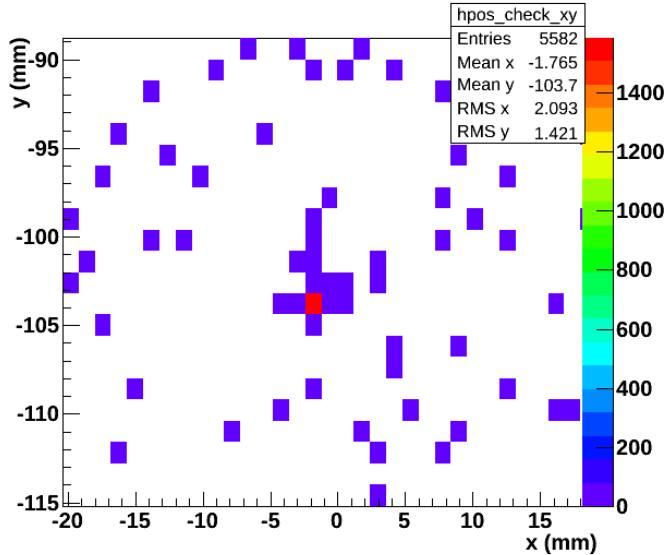


Figure 5.10:  $(x, y)$  position of events with  $3 < \text{posparam} < 5$  and main S2 energy  $< 200$  pe. The color axis shows the number of counts. Zoomed in around the only hot spot.

pollution from lower energy events that end up in the event sample because a bottom PMT is increasing the recorded S2 total energy. Figure 5.11 shows posparam versus the S2 Total Top energy. Figure 5.9 shows a vertical slice in the same energy region as Figure 5.8, the outlier events due to the hot spot are gone and the fit function now fits the data well.

Now an energy dependent cut value with a specific acceptance can be obtained by

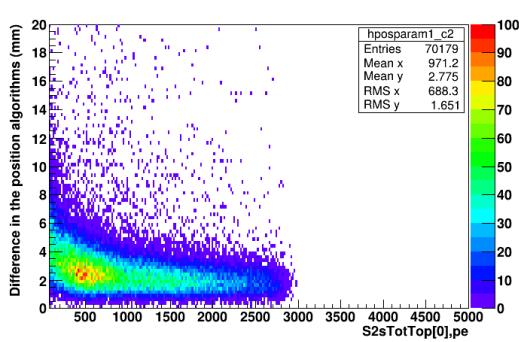


Figure 5.11: Posparam versus the main S2 energy ( $S2sTot[0]$ ) in the bottom array. The color axis shows the number of counts.

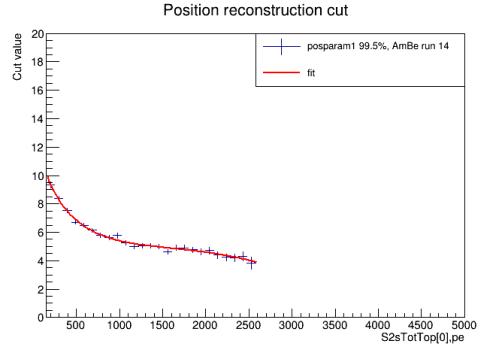


Figure 5.12: Cut values of posparam at the 99.5% acceptance, defined on the S2 Total Top energy. For energies larger than 2600 pe the value at 2600 pe is used.

taking the quantile at the desired acceptance as was also done for the  $\chi^2$  data quality cut. An acceptance of 99.5% is used again. Figure 5.12 shows the resulting cut values. The continuous value is given by the fitted function:

$$posparam < 12.7 + 1.61 \log(E_{S2,Top}) - 1.083 \sqrt{E_{S2,Top}} + 0.0176 E_{S2,Top} - 1.78 \cdot 10^{-6} E_{S2,Top}^2. \quad (5.6)$$

### 5.3 Asymmetry cut

A cut on the asymmetry of the main S2 in an event can remove events that originate from the gaseous region of the detector. These events have a very high asymmetry S2 since they occur above the liquid level. Noise caused by PMTs in the bottom array pollutes the event sample with very low asymmetry events that can also be cut with an asymmetry cut. Chapter 6 looks at a population of low S2 signals that are not due to noise of PMTs. The S2 asymmetry is defined as:

$$\text{S2 Asymmetry} \equiv \frac{\text{S2 Total Top energy} - \text{S2 Total Bottom energy}}{\text{S2 Total energy}}. \quad (5.7)$$

This section defines a cut that only keeps S2s with a ‘correct’ asymmetry. This cut exists in the run 12 analysis and is updated here for run 14. The run 12 cut is simply:

$$-0.2 < \text{S2 Asymmetry} < 0.25. \quad (5.8)$$

The S2 asymmetry is compared for three different data sets, the ‘noise free’ run 12 AmBe data, run 14 AmBe data and reprocessed AmBe data where the five most noisy PMTs are removed from data processing.

Figure 5.13 shows the asymmetry in run 12 (left panel) and run 14 (right panel). Both panels show the expected distribution at asymmetry  $\sim 0.1$ . For a certain energy bin the asymmetry is distributed as a Gaussian. The distribution gets wider at lower energies. In run 14 there is a second population at very low asymmetry and low energy. The low asymmetry S2s are the result of a normal S2 with a very large energy contribution from a noisy PMT in the bottom array causing its asymmetry to drop.

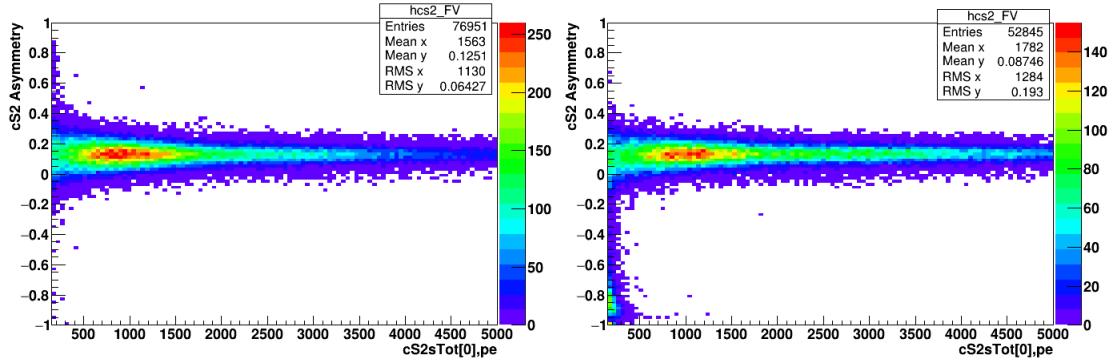


Figure 5.13: S2 Asymmetry versus S2 Energy ( $cS2sTot[0]$ ) in run 12 (left) and run 14 (right). The color axis shows the number of events.

An example of such an event is given in Figure 5.14. This shows the event summary of a run 14 AmBe event with an asymmetry below  $-0.7$  and a main S2 energy below 300 pe. In this particular event PMT 166 and PMT 167 give off a periodic noise through most of the event window. From 150 to 400  $\mu\text{s}$  they are continuously on, distorting the

sum waveform and causing the peak finder to find many false positives. The main S2 will thus have a large energy contribution from these two malfunctioning PMTs which are in the bottom of the TPC; resulting in a very low asymmetry.

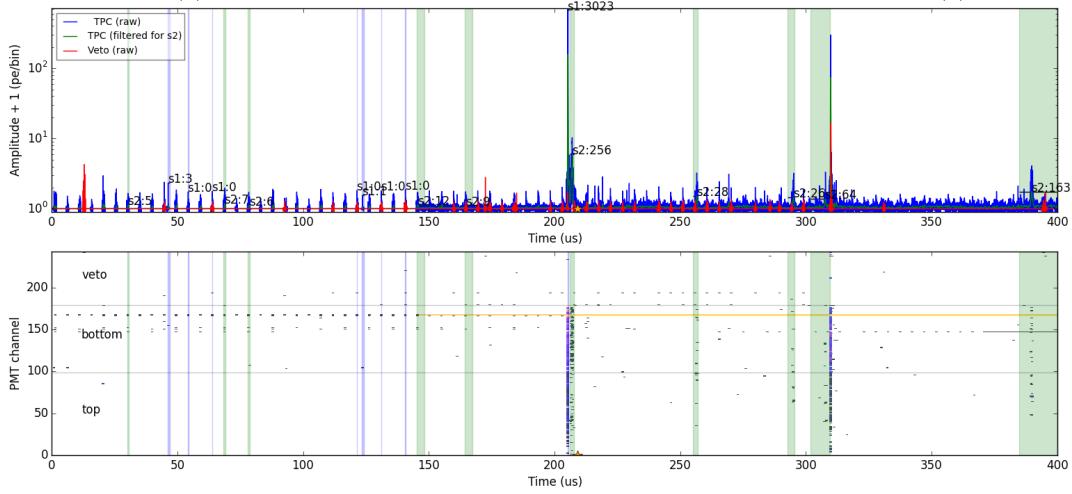


Figure 5.14: Sum waveform (top panel) and per-channel hits (bottom panel) for a typical low energy, low asymmetry event as seen in the low asymmetry population of Figure 5.13. The vertical blue and green bands show the found S1s and S2s respectively. The horizontal yellow line in the bottom panel shows that a PMT is continuously giving off signal.

Figure 5.15 shows the asymmetry of the reprocessed data. Here the low asymmetry population has disappeared leading to the conclusion that the five removed PMTs are the cause of the low asymmetry events.

After binning the energy axis the distributions in each bin are fitted with a Gaussian. For each bin the mean and standard deviation is determined. Figure 5.16 shows the result of all obtained values. As can be seen the run 12 asymmetry mean is constant. In run 14 the mean drops at lower energy due to the presence of the low asymmetry population. When removing the noisy PMTs the mean is constant again but at a higher value. This is caused by the fact that there are now less PMTs in the bottom array and relatively more light is collected at the top causing a shift in the mean asymmetry. Fortunately the sigmas are the same for all data sets. Thus to update the run 12 cut for use in run 14 (with the excluded PMTs) only a shift of the cut value is needed and no energy dependence need be introduced.

Figure 5.17 shows how the shift in asymmetry is calculated. Of both energy dependent mean distributions the mean is determined and the difference between these two values is taken to be the shift. The updated asymmetry cut for run 14 with noisy PMTs excluded thus become:

$$-0.1581 < \text{S2 Asymmetry} < 0.2919. \quad (5.9)$$

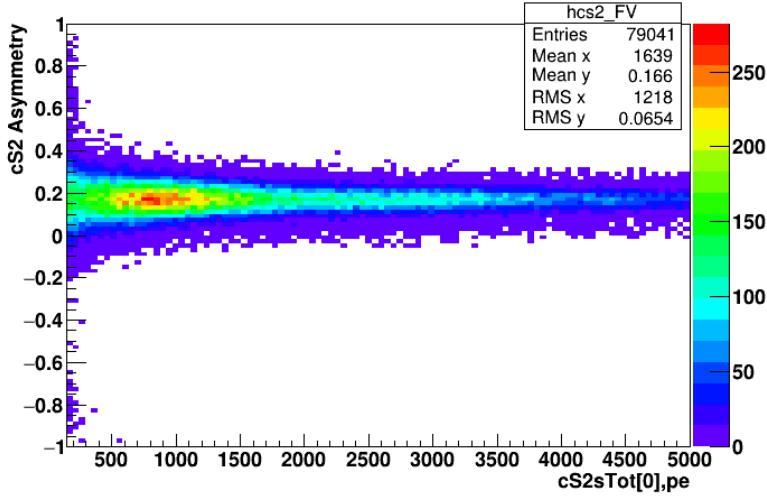


Figure 5.15: S2 Asymmetry versus S2 Energy ( $cS2sTot[0]$ ) in the reprocessed run 14 data with excluded PMTs. Color axis shows number of events.

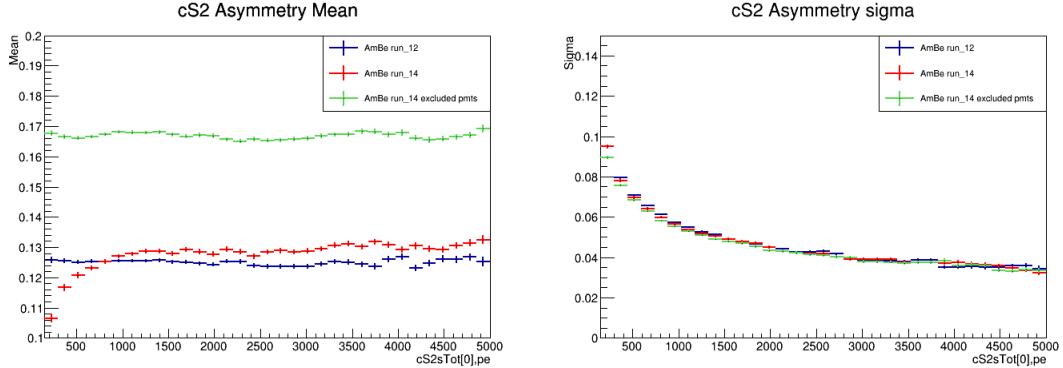


Figure 5.16: Means (left panel) and sigmas (right panel) of each energy bins for the three data sets.

After investigating the S2 asymmetry in run 14 AmBe data a population with very low asymmetry is seen at low energy. After selecting events from the population and looking at wave forms it was concluded that these events have a very high noise contribution from specific PMTs (mostly PMT 147, 149, 152, 166 and 167). Since these PMTs are all in the bottom array, the affected events have a low asymmetry.

In reprocessed run 14 AmBe data the low asymmetry events are gone as shown in Figure 5.15. The mean asymmetry is shifted up because less light is being seen in the bottom array due to the removed PMTs.

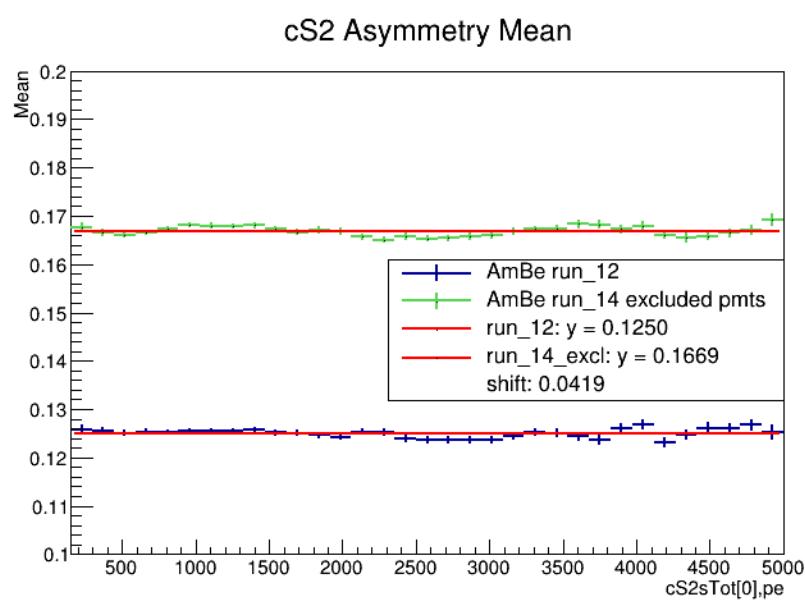


Figure 5.17: Shift of the asymmetry relative to run 12 due to the removes PMTs.



# Chapter 6

## Low Asymmetry S2s

In Chapter 4 and Chapter 5 PAX was mentioned several times, first as the framework that contained the new position reconstruction plugin for XENON1T, secondly as a data processor that is better able to handle data containing noise. In this chapter PAX is used to look at XENON100 data to try and explain a low asymmetry population seen in the S2 asymmetry of run 14 as described in Chapter 5. This analysis is the first analysis done fully with PAX and some surprising events were found.

For each data set analysed by PAX the original XENON100 raw data of that set is reprocessed. This means that instead of merely looking at waveforms of specific events as done in Chapter 5 the reprocessed events are stored in an output file. This file contains the reprocessed events with all their properties to be used in the analysis.

Different data sets of XENON100 data are reprocessed. Starting with a clean sample of AmBe data in run 10. During run 10 (the main science run of XENON100) the detector had little noise and its behaviour during this period is well understood. Therefore it is a good starting point for the analysis as no surprises are expected. After investigating an AmBe run 10 data set, also AmBe run 14 and a dark matter data set from run 10 are looked at.

### 6.1 Event selection

In this analysis the following event selection is used:

1. The event has to have at least one S1 and at least one S2.
2. The largest S1 has to occur before the largest S2 in time.
3. The largest S1 and largest S2 signal have to be in the TPC, not in the veto.
4. The largest S2 energy is between 150 and 5000 pe.
5. The number of S2 peaks is less than 10.

These are minimal cuts to ensure none-physical events are cut. Most of the cuts are self explanatory. The energy range for S2 peaks is typical for AmBe data. The multiple scatter cut (number of S2 peaks  $< 10$ ) is arbitrary. It is low enough to cut events with very many S2 peaks indicating a very high energy deposition. Also it is high enough to allow some single electron S2 peaks (small  $\sim 20$  pe S2 peaks caused by a single electron extracted in the gas) and not drop the statistics too much.

## 6.2 Clean AmBe data

For the clean data sample I selected a run 10 AmBe file, containing 150000 events and with a live time of 7650 seconds. After applying the event selection described above 23576 events remain. Figure 6.1 shows a density plot of these events, S2 Bottom hit pattern spread versus S2 asymmetry. Hit pattern spread is defined as the weighted standard deviation of the distances from the charge weighted sum to each PMT in the hit pattern. It quantifies how local (low spread) or extended (high spread) a hit pattern is. In the case of Figure 6.1 the spread of the largest S2 peak in the event seen by the bottom PMT array is calculated.

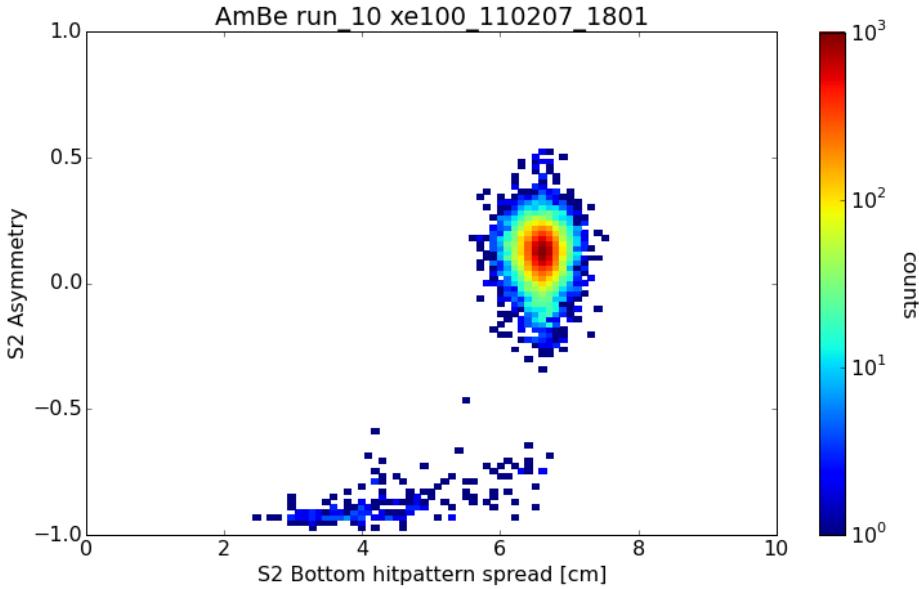


Figure 6.1: Density plot of 23576 S2 peaks in run 10 AmBe data, spread versus asymmetry. Two populations can be identified: a population at the known S2 asymmetry and a population at very low asymmetry. This second population has a low spread meaning a very local distribution of energy.

In Figure 6.1 two populations are visible: a main population with an asymmetry of about 0.1, this is known to be the mean S2 asymmetry of 'proper' events as was shown

in Chapter 5. There is also a small population ( $\sim 1\%$  of the selected events) with very low asymmetry.

These low asymmetry events are usually cut in the standard analysis (see Chapter 5). These events also have a lower bottom hit pattern spread than the main population. The main population has a bottom hit pattern spread of 6 to 7 cm, meaning the S2 energy is very widely distributed on the bottom PMT array as is normal for S2s. The low asymmetry population shows a smaller bottom hit pattern spread, meaning the energy of these S2s is more locally concentrated in the bottom array.

To try and understand Figure 6.1 and classify the low asymmetry events an extra dimension is added. Instead of a density plot Figure 6.2 shows a scatter plot. On the color axis the log of the drift time is plotted.

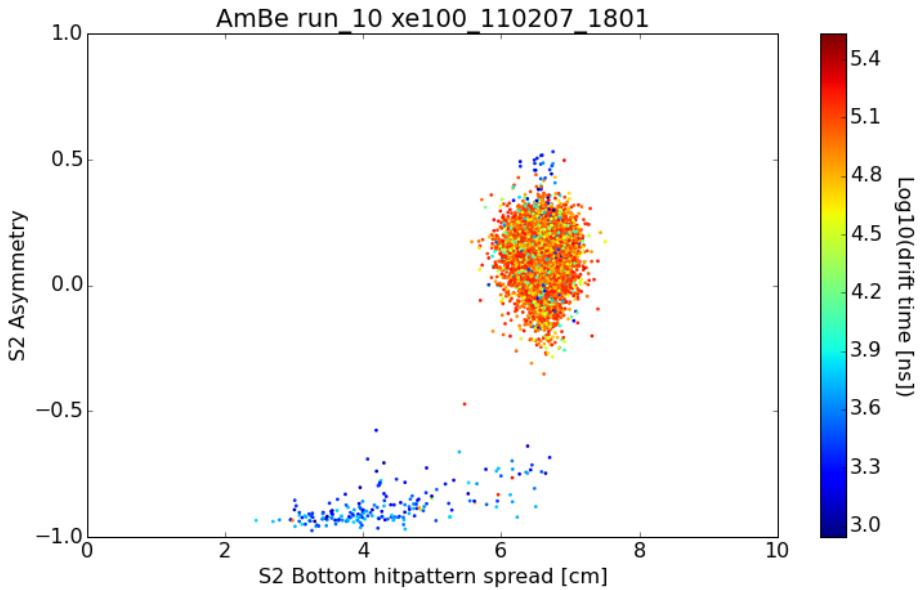


Figure 6.2: Scatter plot of 23576 S2 peaks in run 10 AmBe data, spread vs asymmetry. Drift time on color axis. Events in the low asymmetry population have very short drift times.

Figure 6.2 shows that the low asymmetry population S2's have very short drift times between 1 and 10  $\mu$ s, indicating the event occurs close to the top of the TPC. This seems to contradict the fact that the S2 has a low asymmetry and has a localized bottom hit pattern.

Next the low asymmetry population is isolated by cutting on asymmetry. Events that have an asymmetry smaller than  $-0.4$  and larger than  $-0.98$  are selected. The small region lower than  $-0.98$  is not selected to exclude none-physical events with asymmetry exactly  $-1$ .

After this selection 203 events remain. Small populations away from a known main

population are sometimes the result of detector artefacts (such as the hot spot seen in Chapter 5) or of other none-physical nature. So first we check where these events are located in  $(x, y)$  space. Since the events have low bottom hit pattern spread a charge weighted sum of the bottom hit pattern should give a relatively good position reconstruction (though a charge weighted sum is biased towards the centre as seen in Chapter 4), at least good enough to see whether or not the events are distributed uniformly or are caused by a hotspot or other detector artefact. Also note that since we are looking at the bottom PMT array no Neural Net reconstruction is possible. The  $\chi^2_\gamma$  method can not help us here either since there exist no LCE maps for the bottom PMTs.

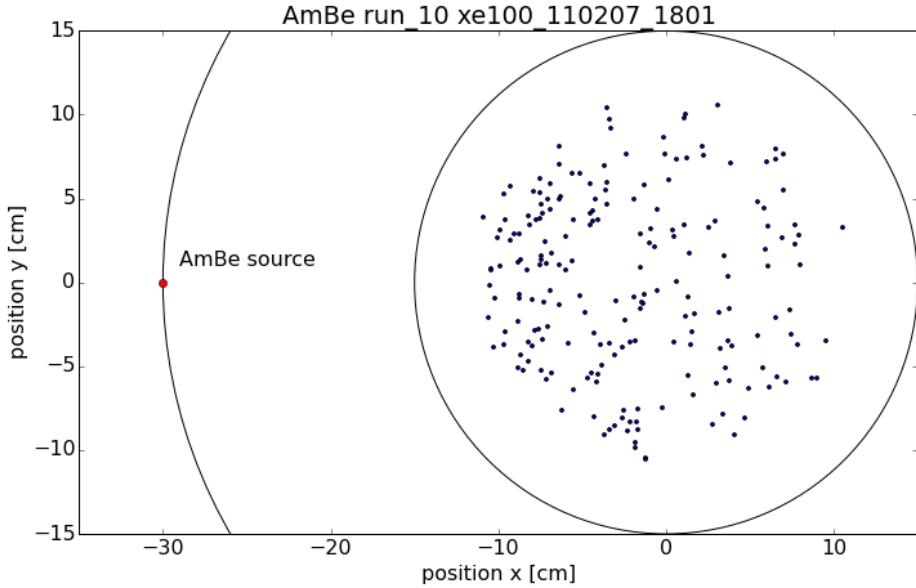


Figure 6.3: Spatial distribution of the low asymmetry population. The red dot marks the AmBe source causing the events. The inner black circle indicates the TPC radius. An explanation for the lack of events close to the TPC edge is because the bottom PMT array does not extend out to the TPC edge and the reconstruction used has a radial bias towards the centre.

Figure 6.3 shows the reconstructed positions of the 203 low asymmetry events using the bottom PMT array. The figure shows that the events are distributed uniformly with slightly more events on the left ( $-x$ ) half consistent with the position of the AmBe source.

As a last check before turning to the waveforms of the low asymmetry events Figure 6.4 shows the drift time, S1 area and S2 area spectra of the events. Of course these three histograms have low statistics and are only meant to give an indication. What we can see from Figure 6.4 is that all events have drift times less than  $8 \mu\text{s}$  (in a normal event this would mean that the event happened  $\sim 1.5 \text{ cm}$  deep). Also we can see that

in these events the S1 has a higher energy than the corresponding S2, inconsistent with regular events.

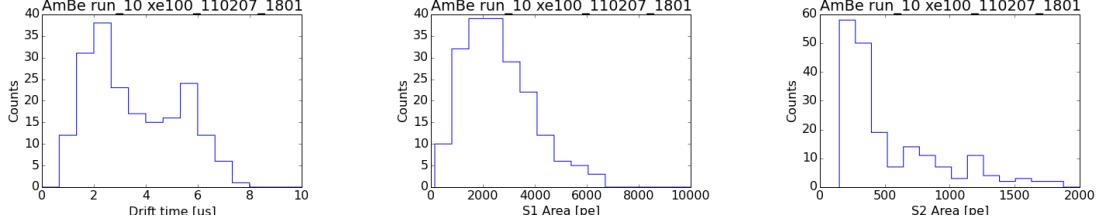


Figure 6.4: Spectra of the drift time, S1 area and S2 area of the low asymmetry events respectively.

From the previous Figures we summarise that the low asymmetry population consists of events where the main S2 has deposited most of its energy locally in the bottom PMT array. The events have a short drift time and in most cases the S1 energy is larger than the S2 energy. The main question being: Where in the TPC (i.e. from what  $(x, y, z)$  position) do these events originate? The drift time (and our current knowledge of the workings of the TPC) suggest close to the top, however the spread and asymmetry seem to suggest an S2 close to the bottom of the TPC.

### 6.2.1 Waveforms

The last step taken is to look at the waveforms and hit patterns of the events themselves. Three events will be discussed here.

For each of the three events two event summaries are shown, one full and one zoomed in on the main S1 and main S2. All these waveforms are also in Appendix B. Each of the event summaries has several panels.

The four panels in the top row show (from left to right):

1. The main S1 peak (linear scale)
2. The hit pattern of the main S1 in the bottom PMT array
3. The hit pattern of the main S2 in the bottom PMT array
4. The main S2 peak (linear scale)

The middle panel shows the sum waveform of all PMTs in a  $400 \mu\text{s}$  time window (with the amplitude on a log scale). The bottom panel shows all hits in the event as red dots in time versus PMT channel. The channels are divided into three groups for the bottom, top and veto PMTs respectively.

Figure 6.5 shows an event with an S1 of 2797 pe and an S2 of 1329 pe about  $2 \mu\text{s}$  apart. The shape of the peaks is consistent with the classification (the S2 is not a misidentified S1 for example). Looking at the channel information (bottom panel of Figure 6.5) we see that the S2 has a low asymmetry since most of the hits are in the

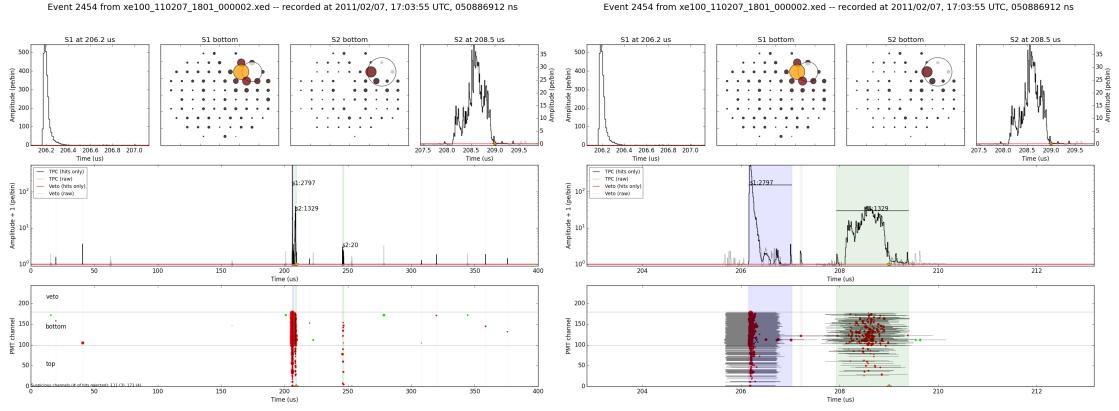


Figure 6.5: Typical low asymmetry event, full (left plot) and zoomed in (right plot). Large version in Appendix B.

bottom PMT channels and few are in the top. By looking at the hit patterns we confirm that the S2 bottom hit pattern is local since most of the energy is in a few PMTs at the top right. The S1 bottom hit pattern shows the same location consistent with a very short drift time between the S1 and S2 signal. Both the localized S2 bottom hit pattern and the asymmetry suggest a signal from deep in the TPC. The drift time suggests the S2 right after the S1 with similar hit pattern in the bottom PMT array. This event also has a small (20 pe) second S2 roughly 50  $\mu$ s after the S1, if this single electron S2 is caused by the S1 then this drift time may be more related to the depth of the S1 than the main S2. However even a drift time of  $\sim 50 \mu$ s is not deep enough ( $50\mu$ s drift time  $\sim 8$  cm depth) to explain the asymmetry and the single electron S2 may well be caused by one of the earlier peaks.

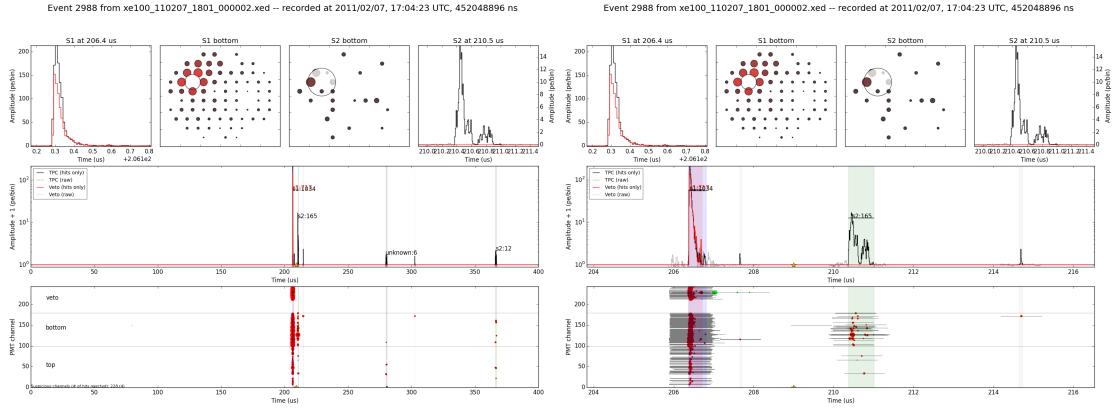


Figure 6.6: Veto low asymmetry event, full (left plot) and zoomed in (right plot). Large version in Appendix B.

The event in Figure 6.6 shows a main S1 of 1034 pe and a main S2 of 165 pe about

$4 \mu\text{s}$  apart. The same type of event as in Figure 6.5. Again we see a low S2 asymmetry, very localized S2 bottom hit pattern correlated with the S1 bottom hit pattern and a short drift time. The event also has a second S2 at  $\sim 170 \mu\text{s}$ , close to the maximum drift time possible in the TPC consistent with a deep S1. The S2 in this event has a double peak structure. Each of the peaks looks more S1 like, but if they are S1s the event has no large S2 which is uncommon since there is a big S1 signal. The main S1 also extends into the veto region of the detector. The fact that the S1 is seen in both the TPC and veto excludes alpha decays as a possible source of these events since an alpha particle cannot penetrate the TPC wall.

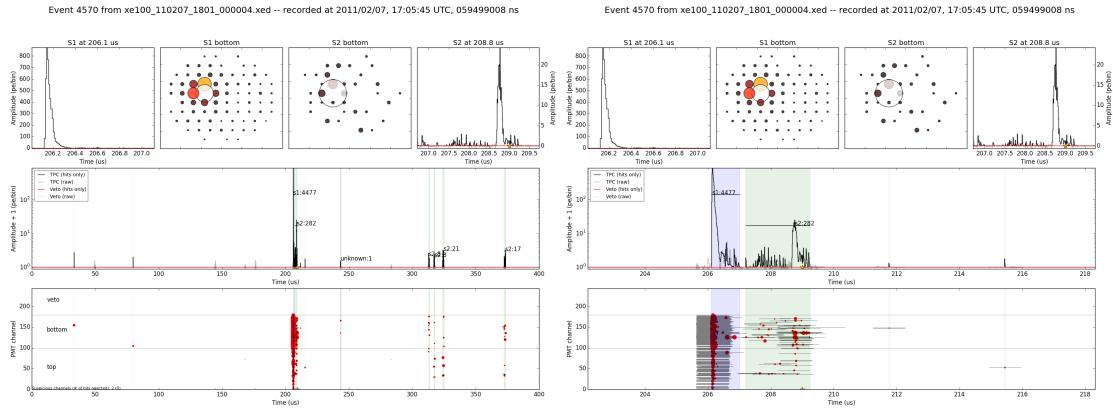


Figure 6.7: Misidentified low asymmetry event, full (left plot) and zoomed in (right plot). Large version in Appendix B.

The event in Figure 6.7 has a main S1 of 4477 pe and a main S2 of 282 pe with a  $2 \mu\text{s}$  drift time. Again a low asymmetry, localized S2 bottom hit pattern correlated with the S1 bottom hit pattern is observed. This event also has several small S2s at very long drift times. Looking at the top right panel of Figure 6.7 (the main S2 on a linear scale) it seems that this particular S2 peak is actually a misidentified S1 peak.

In summary there are several possible hypotheses for the origin of these events.

1. A detector anomaly: If true, the events are only expected to be present in this run or data set.
2. The S2s are misidentified S1s: This could explain some events (such as in Figure 6.7) but not all (the event in Figure 6.5).
3. The S2s do come from deep in the TPC: This implies they are generated in the liquid and are the result of scintillation on or near the bottom screening mesh or cathode (these are two of the metal meshes inside the TPC to which voltages are applied to create the electric fields). This should not be possible in XENON100 due to the electric field strength being too low.
4. ...

In the next section hypothesis 1 is tested by checking different data sets taken a few years later. To exclude the possibility of AmBe being the sole cause of these events a dark matter data set is also used. If hypothesis 3 is correct then the S2s should be created by the electric drift field instead of the extraction field that normally produces the S2 signal. Therefore these events are not expected to occur in a detector configuration where the drift field is much lower or even off, while the extraction field is at its nominal value. Fortunately there exist data sets with lower cathode voltages (weaker drift field) and a nominal extraction field in run 14.

### 6.3 Comparing other sources

In this section the above analysis is repeated on two AmBe run 14 files and on a dark matter (DM, background only) run 10 file. The low asymmetry events described above seem to imply a signal from deep down in the TPC, inconsistent with an S2 created in the gaseous xenon. If the signals seen are from the liquid region we should see a difference when changing the cathode voltage. The following two AmBe run 14 files are taken  $\sim 6$  hours apart, the first file with nominal run 14 conditions (14 kV cathode), the second file with the same configuration except for the cathode voltage which is at 5.4 kV. A dark matter run 10 file was processed to see if the population is still present there and not due to the AmBe source. The noise in the run 14 AmBe data seen in Chapter 5 is practically gone since PAX filters out bad PMT channels. Thus making the low asymmetry events also visible in run 14 data.

1. AmBe run 14, xe100\_140918\_0450 (14 kV cathode voltage)
2. AmBe run 14, xe100\_140918\_1109 (5.4 kV cathode voltage)
3. DM run 10, xe100\_110401\_2058

After event selection the following events remain:

Source	Total events	Live time (s)	Selected	Low	Fraction	Rate
AmBe (14kV)	150000	6435.3	16952	287	0.017	0.0446
AmBe (5.4kV)	150000	6330.7	17351	6	0.0003	0.0009
DM	52302	85055.9	1414	229	0.162	0.0027

Table 6.1: Table showing the analysis results on three data sets. For each data set the source, number of events and live time is given. ‘Selected’ are the events passing the event selection. ‘Low’ gives the number of events with asymmetry  $< -0.4$ . ‘Fraction’ is the fraction ‘Low’/‘Selected’. ‘Rate’ gives the number of low asymmetry events per second.

Figure 6.8 shows the S2 bottom hit pattern spread versus the S2 asymmetry as was also done in Figure 6.1 for the run 10 data. The left plot shows the results for AmBe

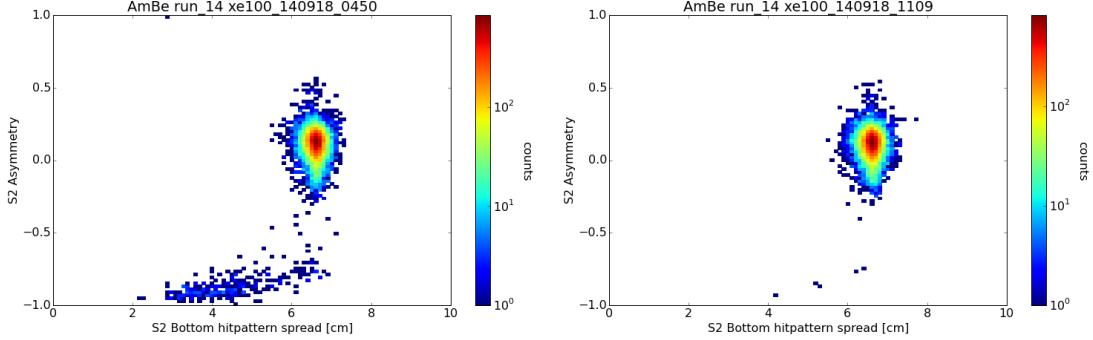


Figure 6.8: S2 asymmetry versus bottom hit pattern spread for AmBe run 14 data. The left plot shows the results with the nominal 14 kV cathode. The right plot shows the results with the low cathode voltage of 5.4 kV.

run 14 under nominal conditions, the right plot for the data set with a lower cathode voltage.

Figure 6.9 shows the result for the dark matter data file. The left panel shows the density plot, the right panel the scatter plot.

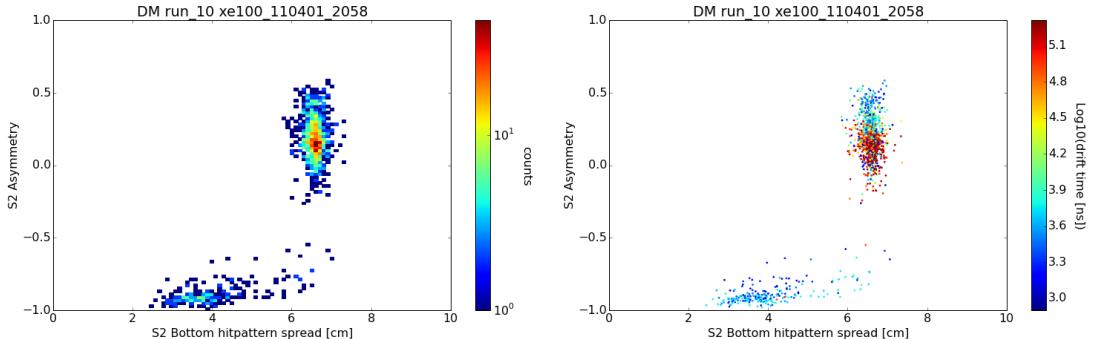


Figure 6.9: S2 asymmetry versus bottom hit pattern spread in DM run 10 data. The left plot shows the density plot. The right plot shows the scatter plot with drift times.

The 14 kV AmBe data shows a similar result as in AmBe run 10 (in which the cathode was at 16 kV) showing that the low asymmetry population is also present in data more than three years later under different detector conditions, making the hypothesis that the events are the result of some detector artefact unlikely. However when we look at the data taken with a cathode voltage of 5.4 kV the population disappears almost completely leaving only six events. The main populations are in both cases very much the same leading to the conclusion that only the low asymmetry events were affected. This result implies that these S2 signals are correlated to the drift field in the TPC. Together with the asymmetry and bottom hit pattern spread this points to the signal being generated in the liquid region of the TPC volume, near the bottom.

The DM data also shows a low asymmetry population as AmBe. In this DM data file the fraction of low asymmetry events is much higher than in AmBe. More than 16 % of the events passing the event selection are made up of these low asymmetry events. As noted before they are cut during the analysis and don't influence the final dark matter analysis result, but the events are nevertheless part of a background that is not understood.

To check if the positions of these events are in line with expectation they are plotted in Figure 6.10. This figure gives the  $(x, y)$  positions of the low asymmetry events in the AmBe (14 kV) and the DM data set. The AmBe events are distributed closer to the AmBe source as expected. The dark matter data set shows the expected uniform distribution when no source is present.

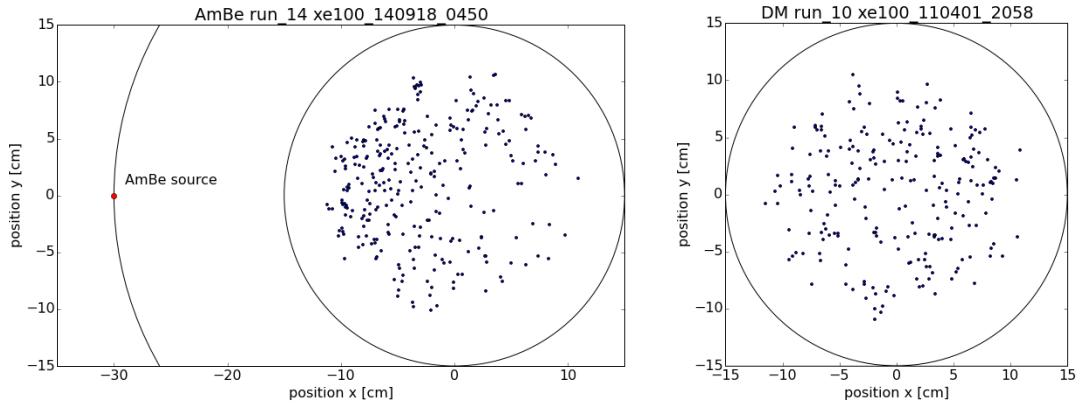


Figure 6.10: Position of the low asymmetry events in AmBe (14kV) (left plot) and DM (right plot).

The waveforms of the low asymmetry events in the AmBe run 14 and the dark matter data set look very similar to the ones seen in the run 10 set. Several more waveforms from the dark matter data are shown in Appendix B. They show the same event types as described in the waveforms from the AmBe run 10 file.

## 6.4 Hypotheses

The main questions raised in this chapter are: Where (i.e. from what  $(x, y, z)$  in the TPC) do these low asymmetry events come from? Also and after examining several waveforms: Are these these are genuine S2 peaks or not. Summarising the unusual combination of properties of the low asymmetry events:

1. The S2s have a very low asymmetry ( $< -0.5$ ).
2. The S2s have a very low bottom hitpattern spread (S2 energy localized in the bottom PMT array).

3. The event has a very short drift time ( $< 10 \mu\text{s}$ ).
4. The S1 has a higher energy than the S2.
5. The events seem to require a high cathode voltage (i.e. disappear at low cathode voltage).
6. The events are present in both AmBe and DM data. (In DM data they occupy 16% of the events that pass event selection defined in this chapter.)

Since the population disappears at low cathode voltage the low asymmetry population seems to come from the liquid part of the TPC. The asymmetry would then suggest the events are from very deep in the TPC, the presence of events with a second S2 after almost the maximum drift time further supports this. The events look like what one would expect from possible cathode effects or scintillation at the bottom screening mesh. However this seems unlikely because the electric field between cathode and bottom screening mesh is not strong enough, the field of 14 kV over 12 mm (11.7 kV/cm) is an order of magnitude lower than the measured value in the paper 'Measurements of proportional scintillation and electron multiplication in liquid xenon using thin wires' [20] which reports a value of  $412^{+10}_{-133}$  kV/cm for proportional scintillation in liquid xenon. At positions very close to the wires of the screening mesh (where the E-field is much stronger) this process could occur, also given the large uncertainty on the required value. Electric field simulations of XENON100 have been performed but not accurately near the bottom screening mesh. For now the origin of the low asymmetry events remains unknown.

Further investigation of these events requires more data as a first step. At the time of writing this thesis, it is not yet possible to analyze very large data sets processed by PAX. Another step would be to try various combinations of electric field configurations and possibly construct a turn-on curve to determine at what field strength these events start to occur.



## Chapter 7

# Conclusion and Discussion

Chapter 4 showed how a new position reconstruction algorithm was implemented in PAX, the data processor for the new XENON1T detector. This algorithm was based on the same principle as an existing algorithm in Xerawdp, the data processor of XENON100. The algorithm is physically motivated and can be used to assign an error on a per-event basis. The new implementation was tested on Monte Carlo data to determine the position reconstruction error. When using Monte Carlo hit patterns sampled from the LCE maps the new position reconstruction algorithm performed excellent, with a lower absolute position reconstruction error than its main competitor, a Neural Network. In PMT arrays that contain dead PMTs the  $\chi^2_\gamma$  method also works without any modification whereas a Neural Network must be retrained.

Various data quality cuts were defined in Chapter 5. These cuts are part of the run 14 analysis of XENON100. Run 14 is one of the last runs of the XENON100 detector and a proving ground for several new techniques and analyses for XENON1T. Run 14 does not contain any dark matter data but only data of various calibration sources. At the beginning of this run and its analysis much was still unknown about the detector behaviour during the run. Noise turned out to be a much larger issue than anticipated. The noise issues resulted in the removal of several PMTs from the data. Partly due to the cuts defined in Chapter 5. The  $\chi^2$  cut and position discrepancy cut were also improved by redefining them as a function of the S2 energy seen in the top array instead of the total S2 energy. The noisy data turned out to be an excellent showcase for PAX’s peak finding routines and gave a glimpse of how data analysis with PAX can be done.

Chapter 6 showed how the new data processor for XENON1T, PAX, was used for its first analysis. When trying to classify a population of very low asymmetry S2 peaks some unexpected events were found. Present in different data sets over several years these events seem to point to S2 signals from the liquid xenon. The strongest evidence being that these events are affected by the electric drift field and not by the extraction field as would be expected from an S2. Though theoretically unlikely no other explanation for these events has yet been found.

## 7.1 Outlook

When XENON1T is commissioned the new position reconstruction algorithm described in Chapter 4 can be used. The generation of the TPC specific LCE map for XENON1T is the only missing piece. To further improve the position reconstructions usefulness a clever way needs to be found to calculate the per-event errors (i.e. confidence contours), this feature is still to be implemented as only a proof of concept was given in this thesis.

Run 14 data is now much better understood than a year ago. The data will be used in studies on the detector response. From this point on XENON100 will exclude 5 more PMTs from analysis due to noise.

An explanation for the low asymmetry events found in Chapter 6 will require more statistics and data. When large scale analysis is possible on data reprocessed by PAX, the analysis could be repeated. Another way of classifying these events would be to take data at various different cathode voltages. Or search for these events in XENON1T which has a different field configuration.

## Acknowledgements

I would like to thank everyone at the Nikhef XENON group for all their help, support and many useful and constructive conversations. I'd like to thank my supervisor Patrick at Nikhef and my supervisor Raimond at the UU. Thanks to Andrea for being a great daily supervisor and office-mate. I'd like to thank Chris for his question ‘Are you sure that  $\chi^2$  is a  $\chi^2$ ?’. Jelle for his Python wizardry. Sander for being the best Diplomacy ally and Froukje for our many coffee-break conversations.



## Appendix A

# MC Position Reconstruction plots

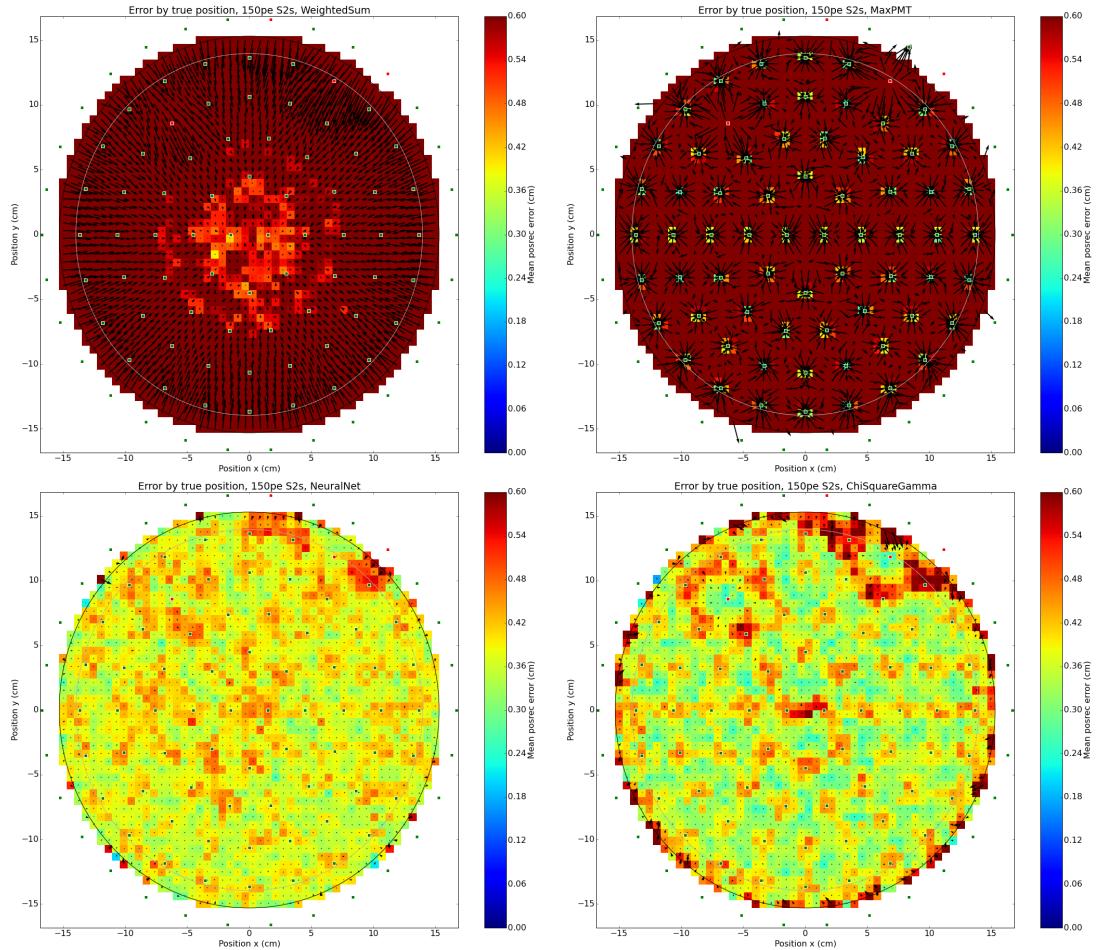


Figure A.1: Binned position reconstruction errors for  $10^5$  150 pe S2 events. Similar to Figure 4.8.

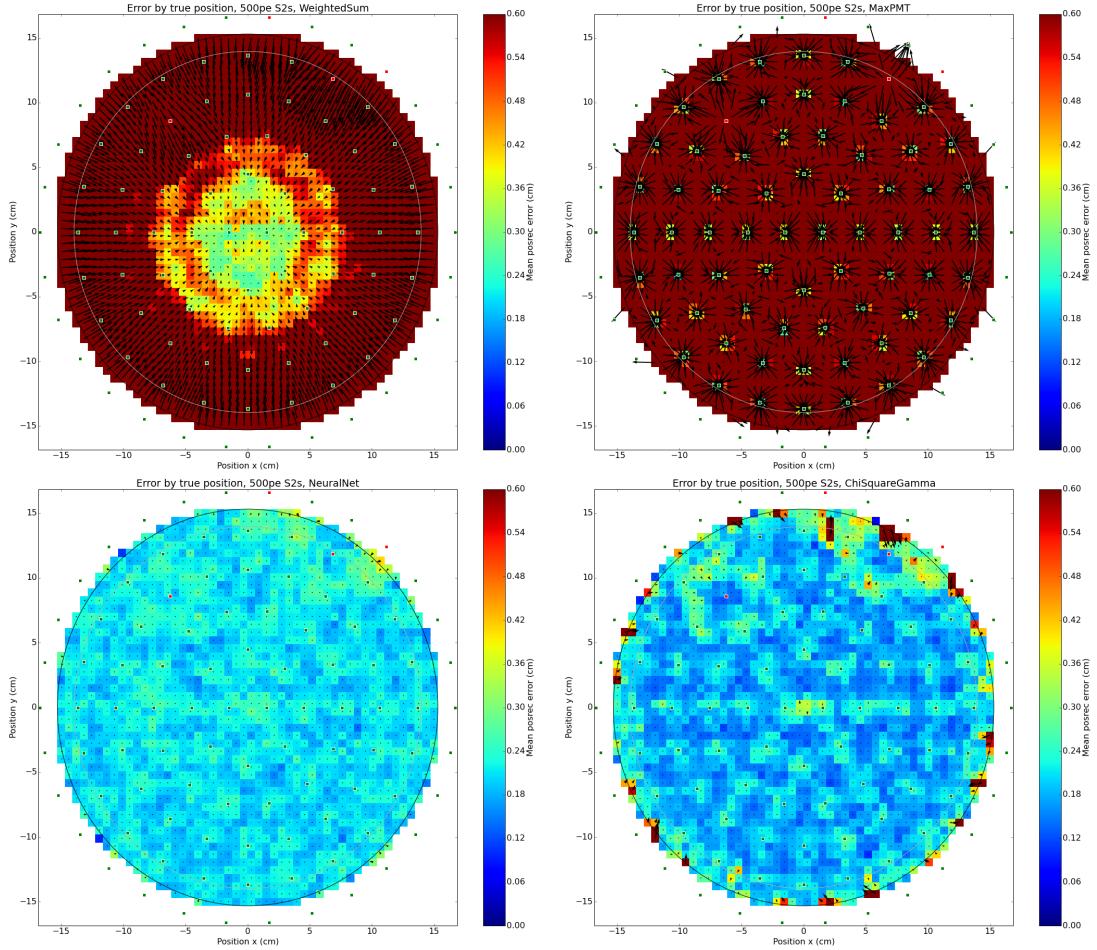


Figure A.2: Binned position reconstruction errors for  $10^5$  500 pe S2 events. Similar to Figure 4.8.

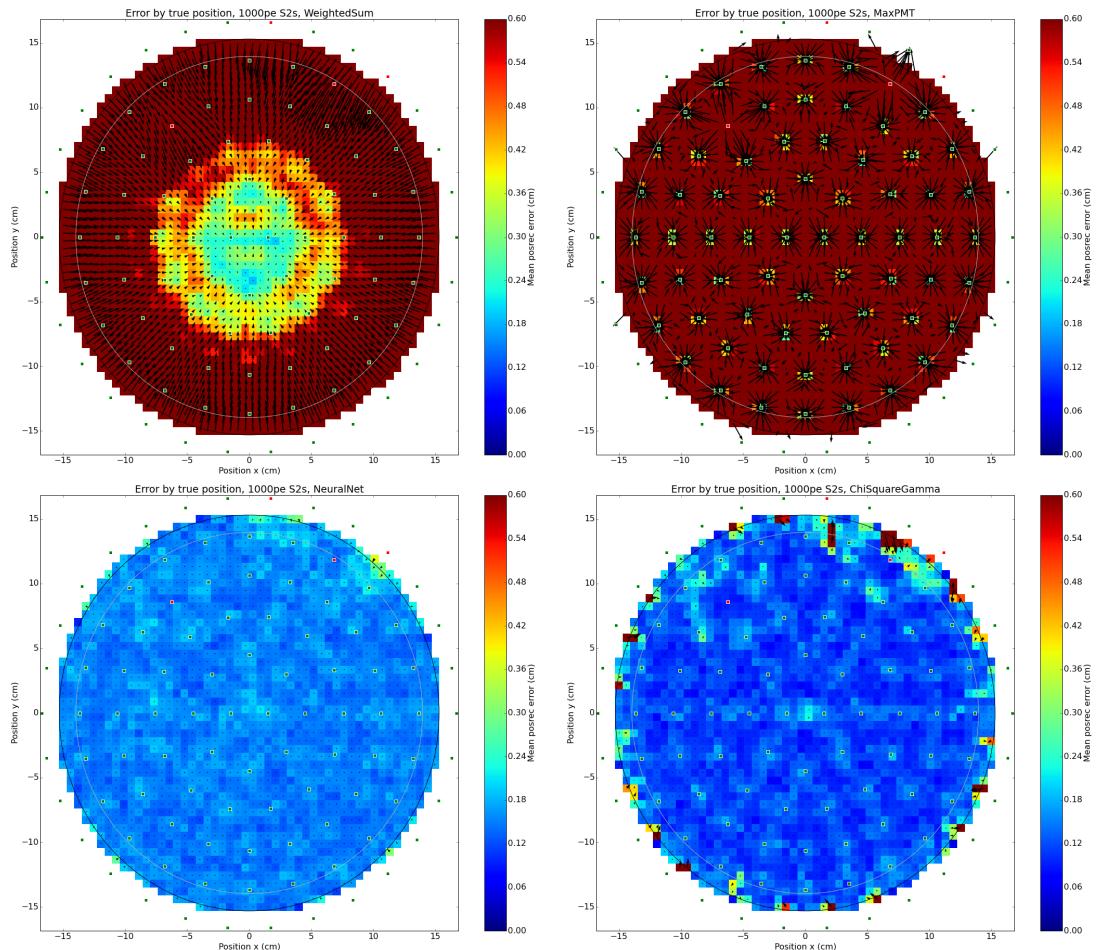


Figure A.3: Binned position reconstruction errors for  $10^5$  1000 pe S2 events. Similar to Figure 4.8.



## **Appendix B**

# **Waveforms**

Several waveforms of events from AmBe run 10 data and Dark Matter run 10 data are shown here. These waveforms all show low asymmetry events as discussed in Chapter 6. For each event the full event summary display is plotted followed by a zoomed in version which shows the main S1 and S2 peak in more detail.

Event 2454 from xe100\_110207\_1801\_000002.xed -- recorded at 2011/02/07, 17:03:55 UTC, 050886912 ns

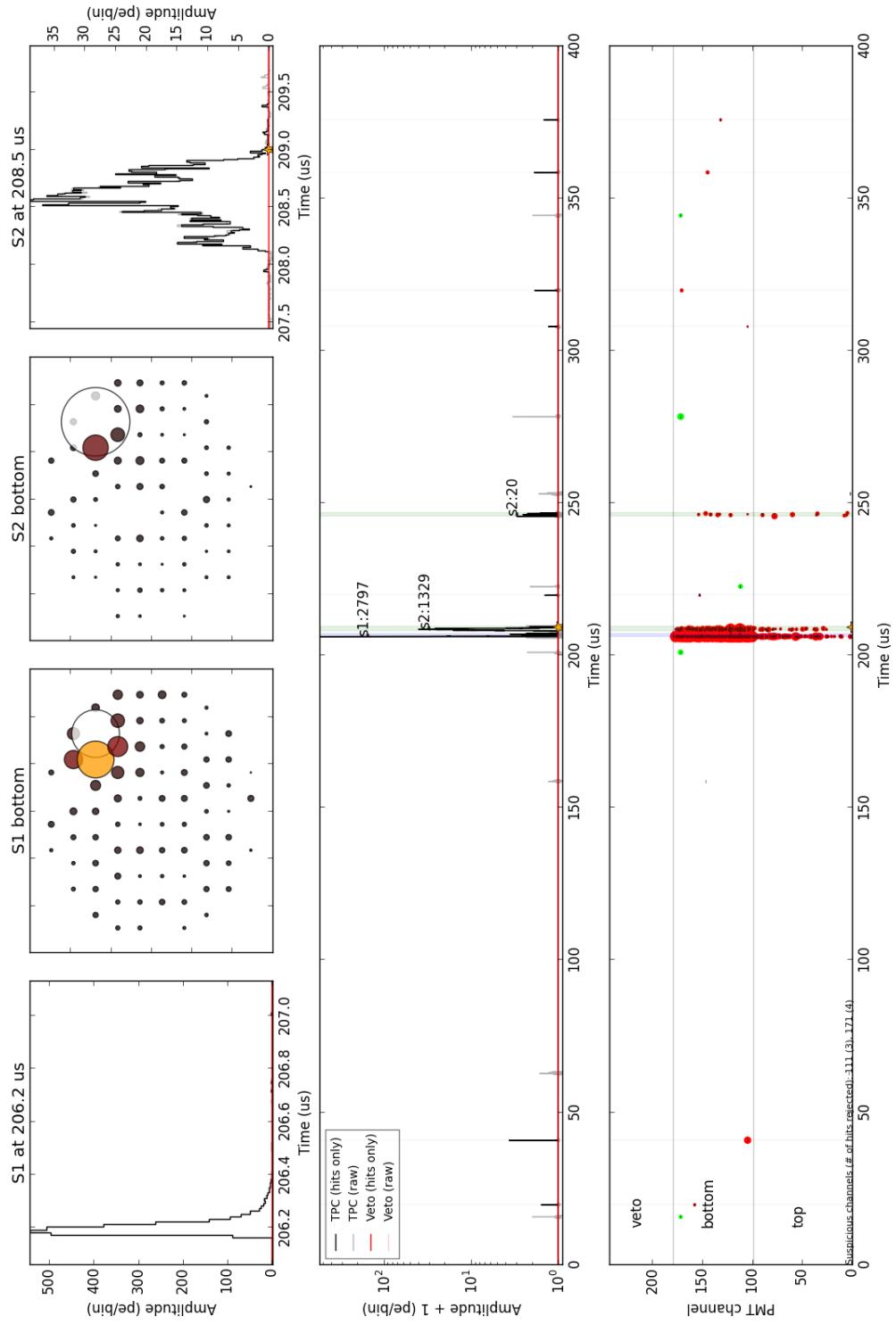


Figure B.1: Low S2 event from AmBe run 10 data.

Event 2454 from xe100\_110207\_1801\_000002.xed -- recorded at 2011/02/07, 17:03:55 UTC, 050886912 ns

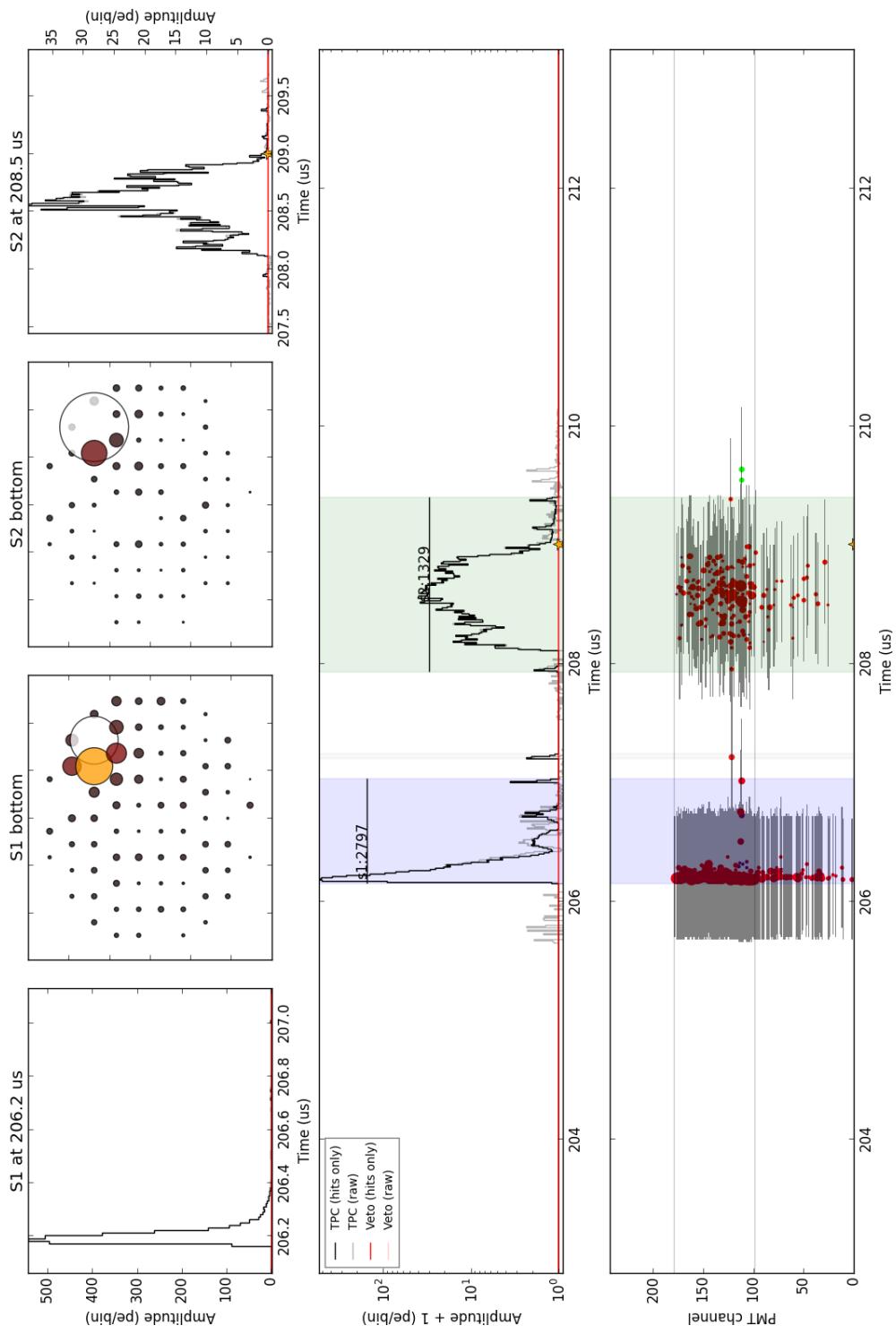


Figure B.2: Low S2 event, zoomed in on Figure B.1.

Event 2988 from xe100\_110207\_1801\_000002.xed -- recorded at 2011/02/07, 17:04:23 UTC, 452048896 ns

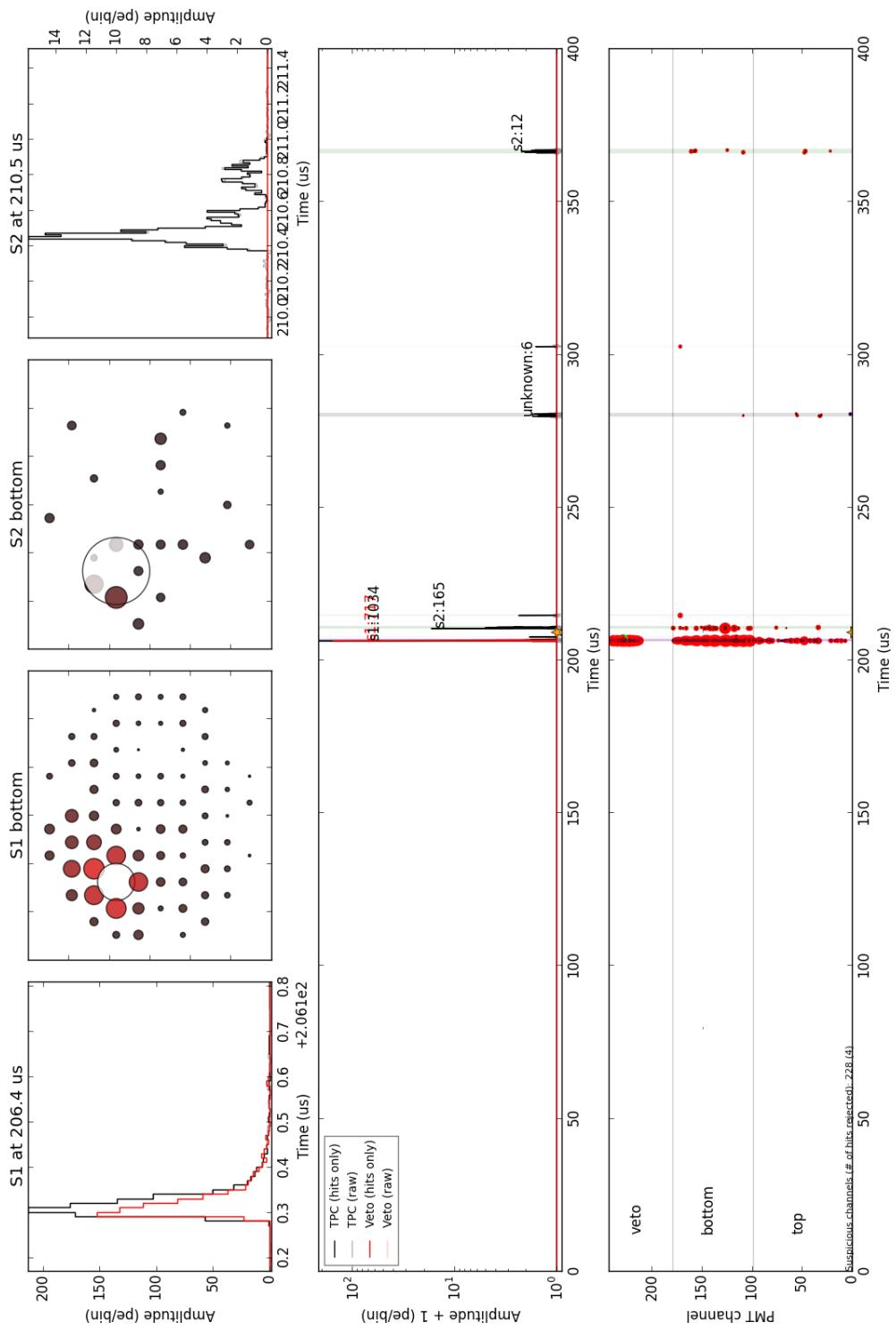


Figure B.3: Low S2 event from AmBe run 10 data.

Event 2988 from xe100\_110207\_1801\_000002.xed -- recorded at 2011/02/07, 17:04:23 UTC, 452048896 ns

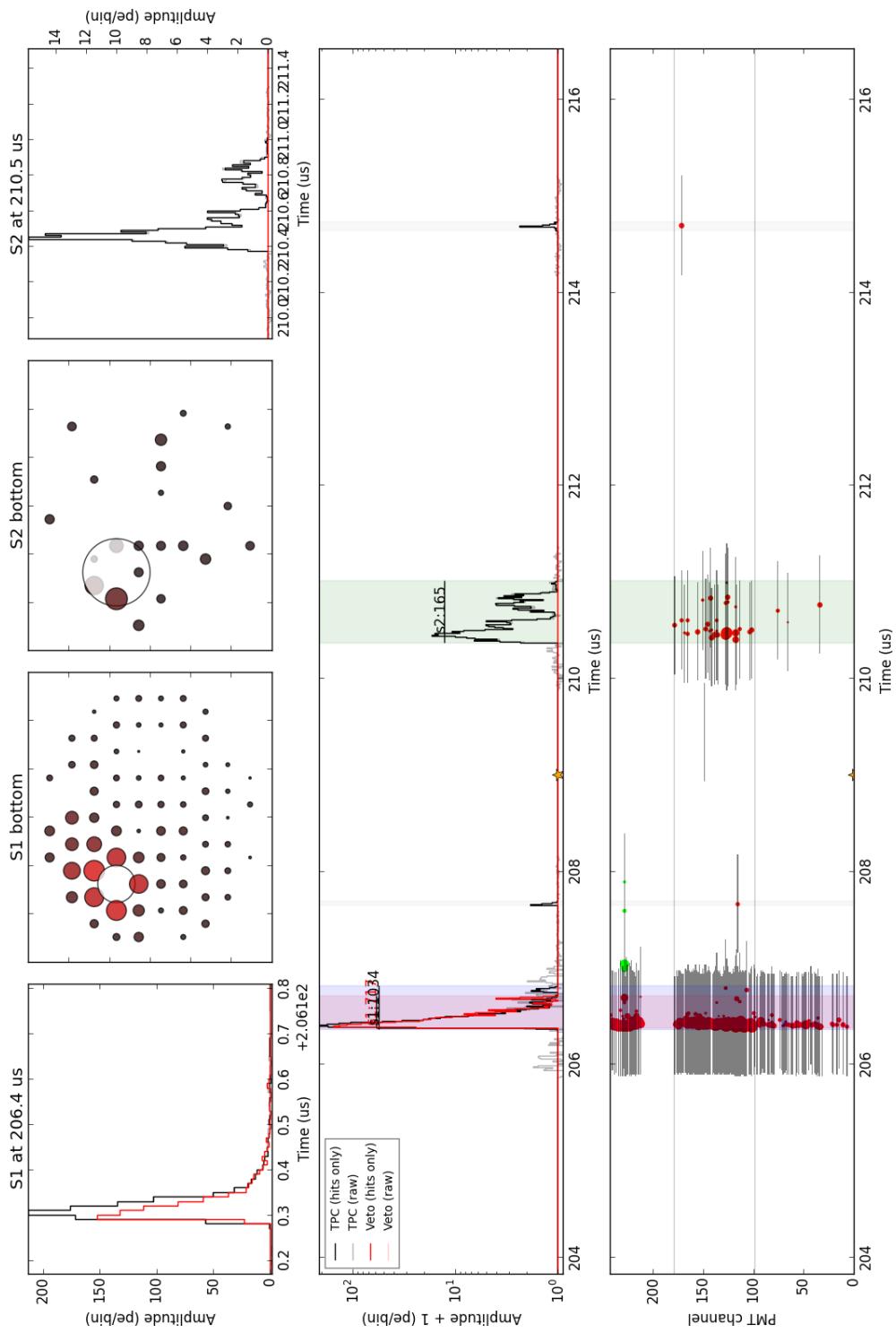


Figure B.4: Low S2 event, zoomed in on Figure B.3.

Event 4570 from xe100\_110207\_1801\_000004.xed -- recorded at 2011/02/07, 17:05:45 UTC, 059499008 ns

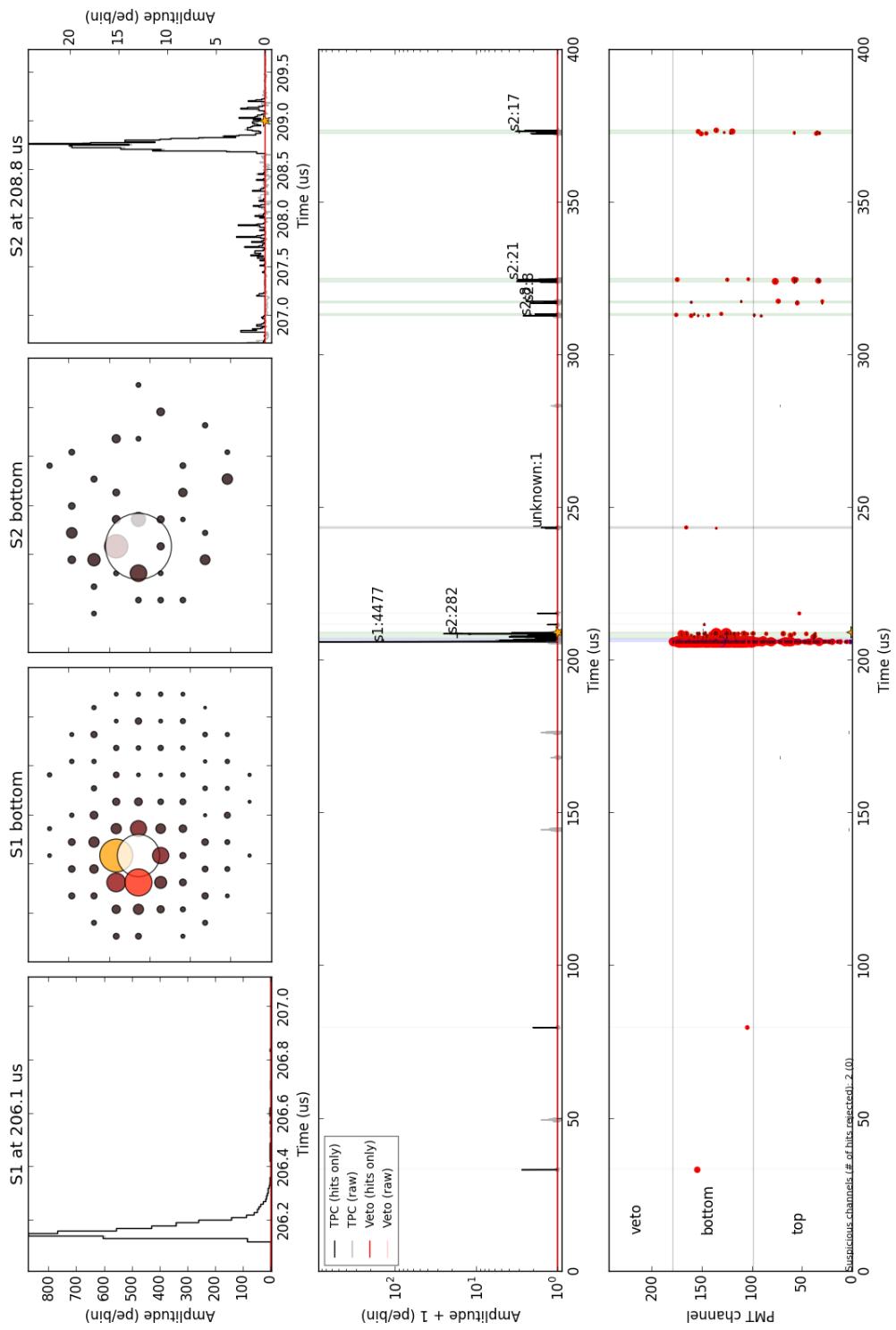


Figure B.5: Low S2 event from AmBe run 10 data.

Event 4570 from xe100\_110207\_1801\_000004.xed -- recorded at 2011/02/07, 17:05:45 UTC, 059499008 ns

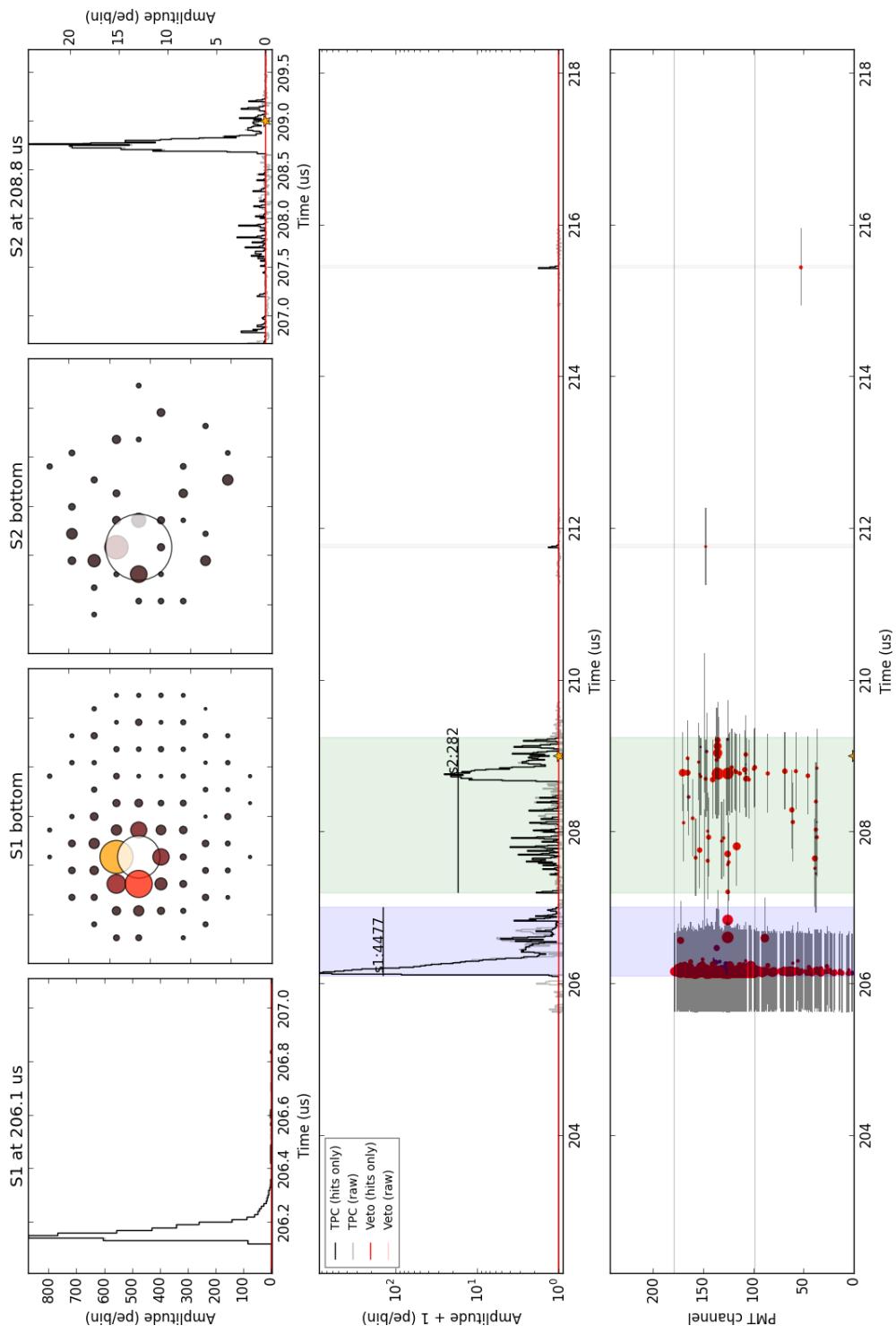


Figure B.6: Low S2 event, zoomed in on Figure B.5

Event 730 from xe100\_110401\_2058\_000000.xed -- recorded at 2011/04/01, 19:08:12 UTC, 577949952 ns

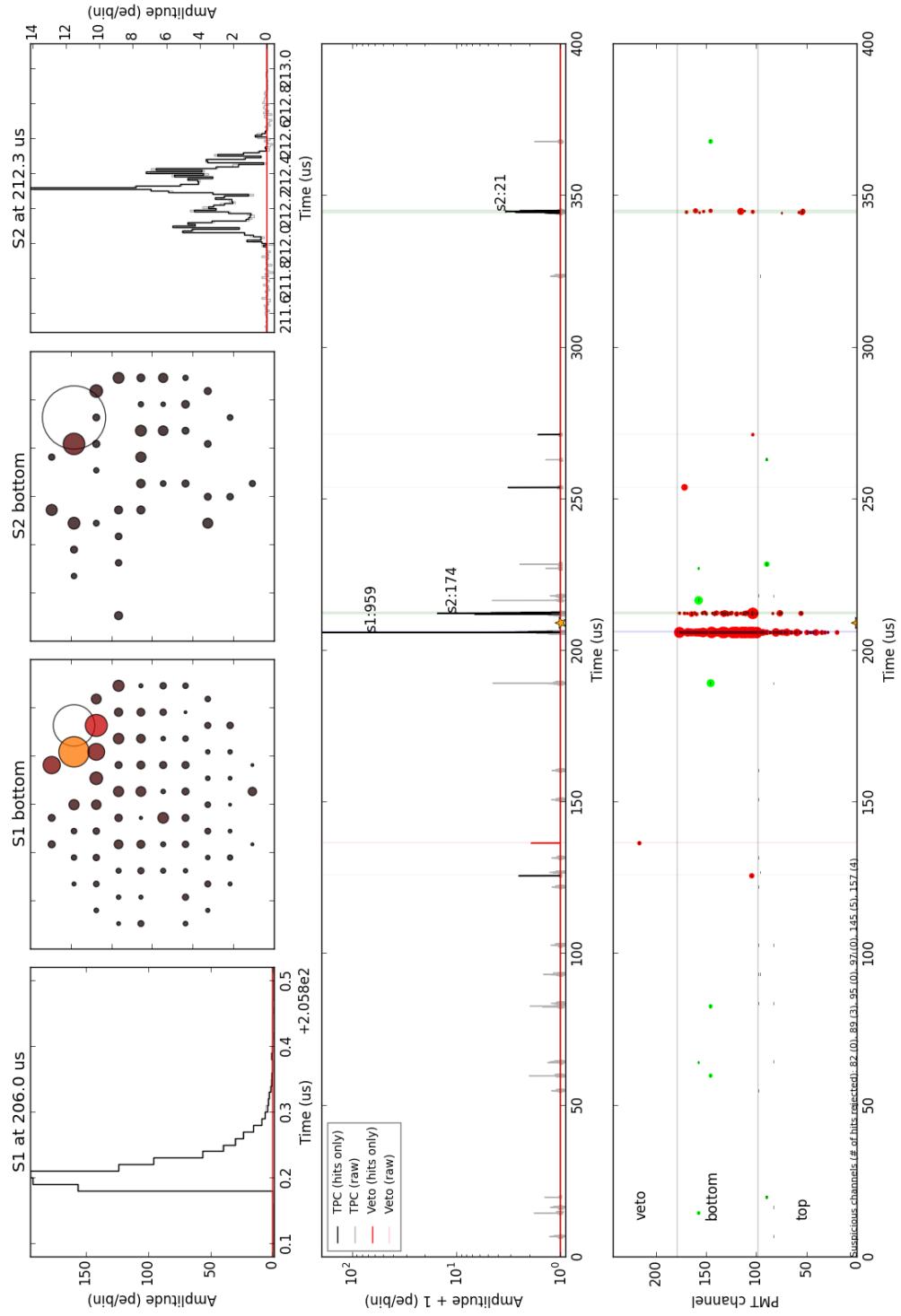


Figure B.7: Low S2 event from DM run 10 data.

Event 730 from xe100\_110401\_2058\_000000.xed -- recorded at 2011/04/01, 19:08:12 UTC, 577949952 ns

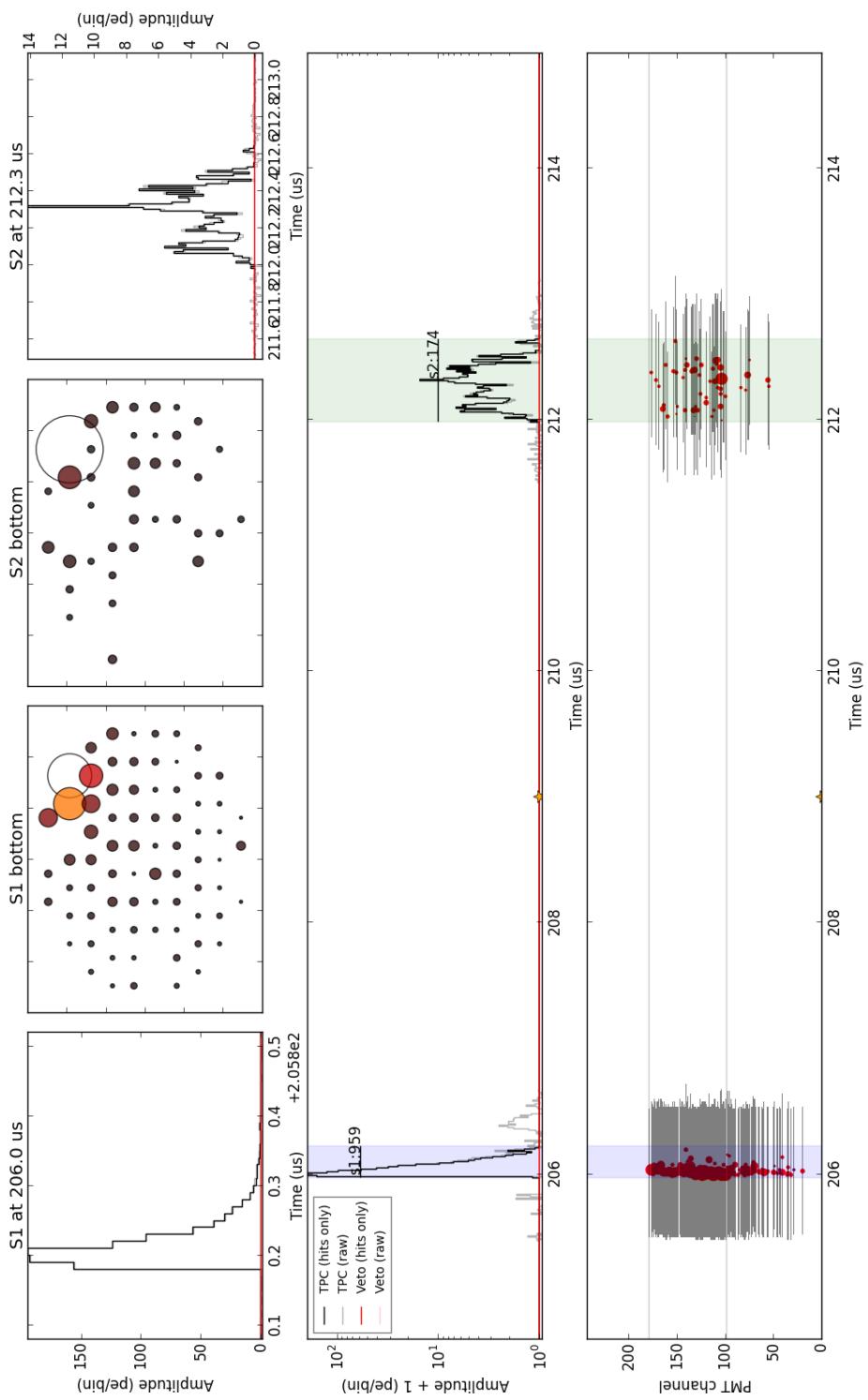


Figure B.8: Low S2 event, zoomed in on Figure B.7

Event 1420 from xe100\_110401\_2058\_000001.xed -- recorded at 2011/04/01, 19:16:57 UTC, 685383936 ns

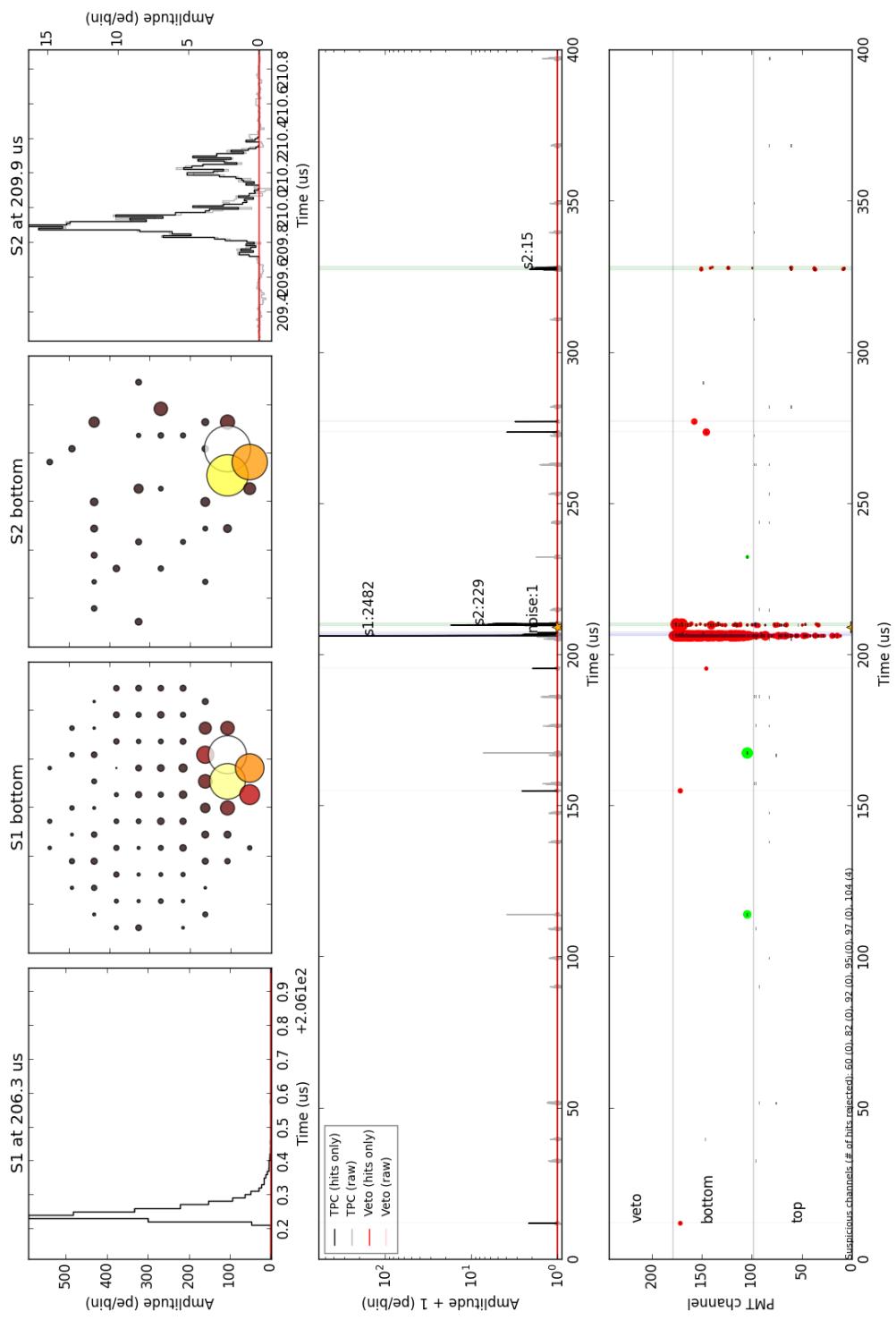


Figure B.9: Low S2 event from DM run 10 data.

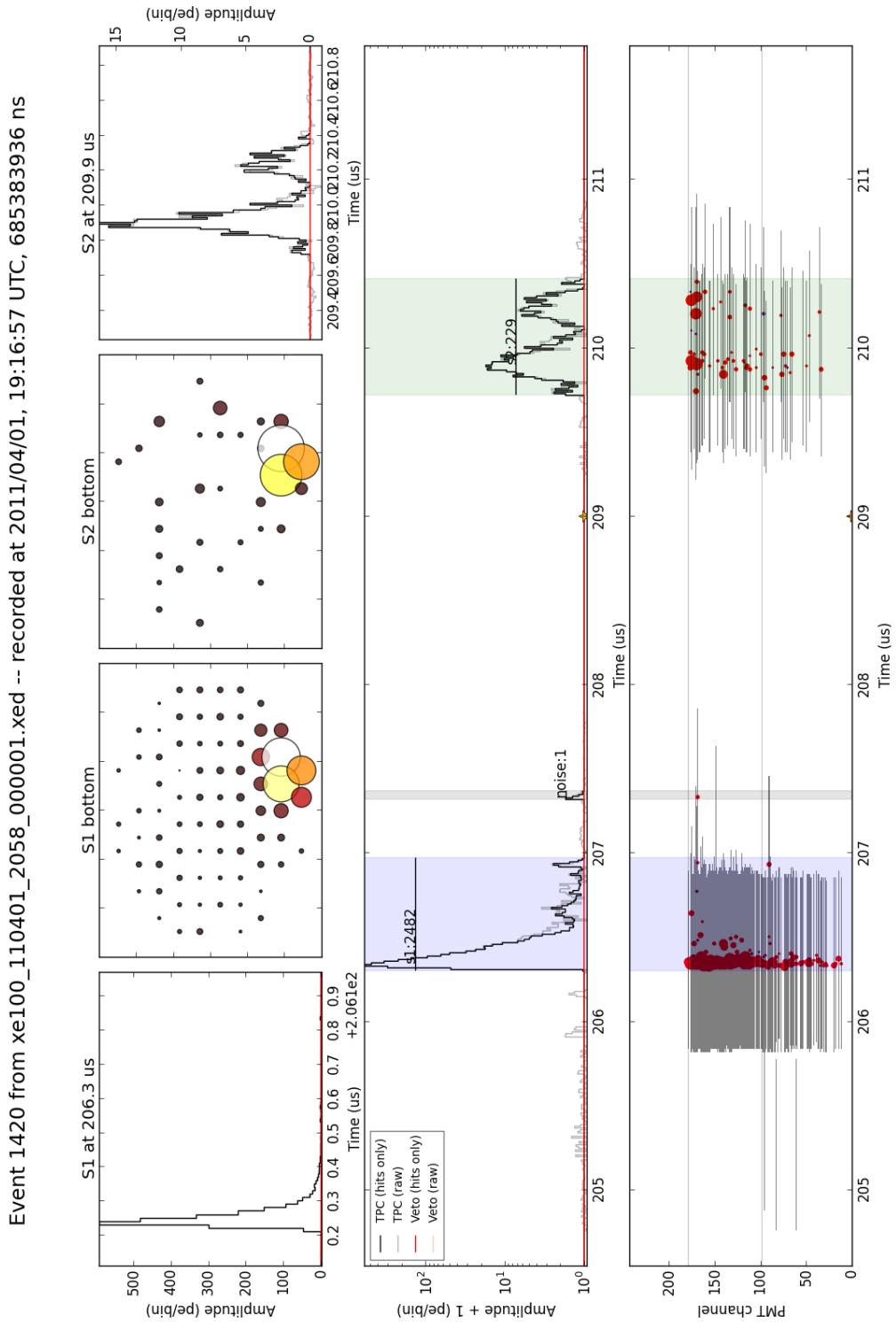


Figure B.10: Low S2 event, zoomed in on Figure B.9

Event 4297 from xe100\_110401\_2058\_000004.xed -- recorded at 2011/04/01, 19:57:59 UTC, 885659904 ns

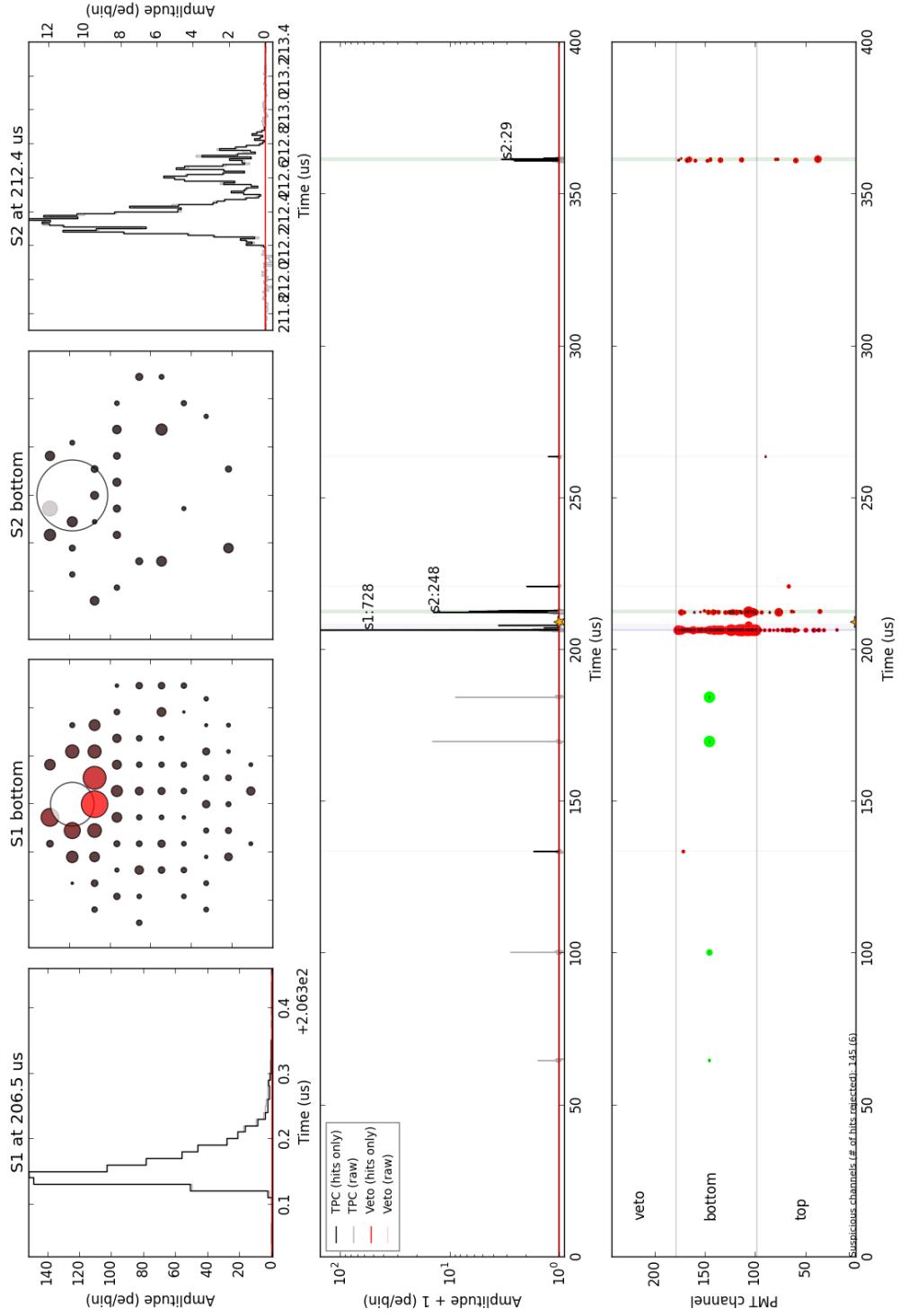


Figure B.11: Low S2 event from DM run 10 data.

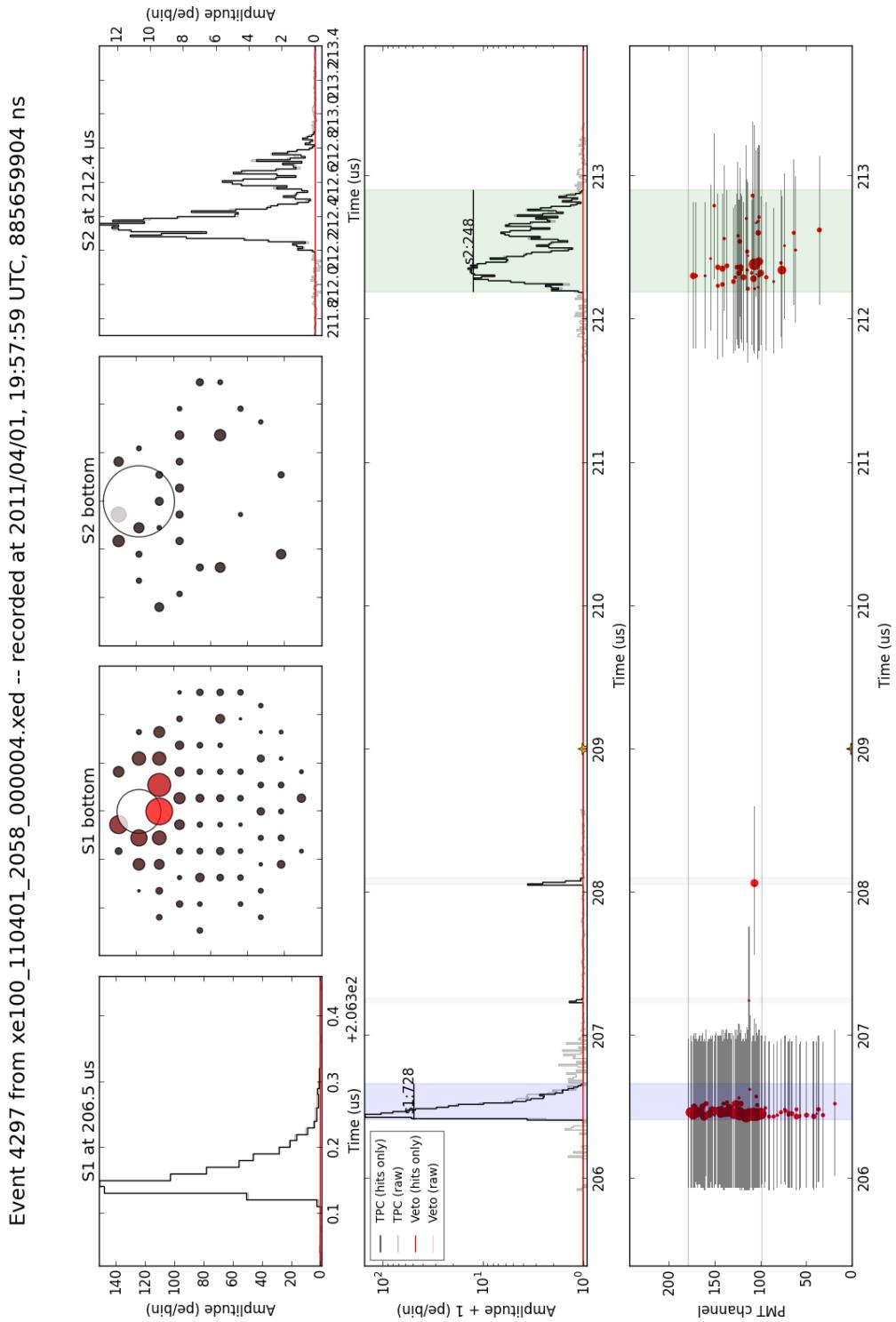


Figure B.12: Low S2 event, zoomed in on Figure B.11



# Bibliography

- [1] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results. *The Astrophysical Journal Supplement Series*, 208(2):19, 2013.
- [2] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*, February 2015.
- [3] Edvige Corbelli and Paolo Salucci. The extended rotation curve and the dark matter halo of m33. *Monthly Notices of the Royal Astronomical Society*, 311(2):441–447, 2000.
- [4] V. C. Rubin, W. K. J. Ford, and N. . Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophysical Journal*, 238:471–487, June 1980.
- [5] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, W. Forman, C. Jones, S. Murray, and W. Tucker. Direct constraints on the dark matter self-interaction cross section from the merging galaxy cluster 1e 0657–56. *The Astrophysical Journal*, 606(2):819, 2004.
- [6] McGaugh Stacy S. A tale of two paradigms: the mutual incommensurability of  $\Lambda$ cdm and mond. *Canadian Journal of Physics*, 93(2):250–259, 2015.
- [7] Riccardo Catena and Laura Covi. Susy dark matter(s). *The European Physical Journal C*, 74(5), 2014.
- [8] Georges Aad et al. Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector. 2015.
- [9] Anirban Biswas, Debasish Majumdar, and Probir Roy. Nonthermal two component dark matter model for fermi-lat -ray excess and 3.55 kev x-ray line. *Journal of High Energy Physics*, 2015(4), 2015.

- [10] Prajwal Raj Kafle, Sanjib Sharma, Geraint F. Lewis, and Joss Bland-Hawthorn. On the shoulders of giants: Properties of the stellar halo and the milky way mass distribution. *The Astrophysical Journal*, 794(1):59, 2014.
- [11] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, S. Castellano, R. Cerulli, C.J. Dai, A. d’Angelo, S. d’Angelo, A. Di Marco, H.L. He, A. Incicchitti, H.H. Kuang, X.H. Ma, F. Montecchia, D. Prosperi, X.D. Sheng, R.G. Wang, and Z.P. Ye. Final model independent results of dama/libra-phase1 and perspectives of phase2. *Physics of Particles and Nuclei*, 46(2):138–146, 2015.
- [12] A. Lyashenko. Xenon dark matter project: past, present and future. Conference on Science at the Sanford Underground Research Facility, 2015.
- [13] D. S. Akerib and et al. First results from the lux dark matter experiment at the sanford underground research facility. *Phys. Rev. Lett.*, 112:091303, Mar 2014.
- [14] Tarek Saab. An Introduction to Dark Matter Direct Detection Searches & Techniques. 2012.
- [15] XENON100 Collaboration, E. Aprile, K. Arisaka, F. Arneodo, A. Askin, L. Baudis, A. Behrens, E. Brown, J. M. R. Cardoso, B. Choi, D. Cline, S. Fattori, A. D. Ferella, K. L. Giboni, A. Kish, C. W. Lam, R. F. Lang, K. E. Lim, J. A. M. Lopes, T. Marrodon Undagoitia, Y. Mei, A. J. Melgarejo Fernandez, K. Ni, U. Oberlack, S. E. A. Orrigo, E. Pantic, G. Plante, A. C. C. Ribeiro, R. Santorelli, J. M. F. dos Santos, M. Schumann, P. Shagin, A. Teymourian, E. Tziaferi, H. Wang, and M. Yamashita. The XENON100 Dark Matter Experiment. 35(9):573–590, 2012. arXiv:1107.2155 [astro-ph].
- [16] Kenneth J. Mighell. Parameter estimation in astronomy with poisson-distributed data. i. the  $\chi^2_\gamma$  statistic. *The Astrophysical Journal*, 518(1):380, 1999.
- [17] Y. Mei. *Direct Dark Matter Search with the XENON100 Experiment*. PhD thesis, Rice University, 2011.
- [18] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.
- [19] M. J. D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, 7(2):155–162, 1964.
- [20] E Aprile, H Contreras, L W Goetzke, A J Melgarejo Fernandez, M Messina, J Naganoma, G Plante, A Rizzo, P Shagin, and R Wall. Measurements of proportional scintillation and electron multiplication in liquid xenon using thin wires. *Journal of Instrumentation*, 9(11):P11012, 2014.