Search for leptonically interacting dark matter with the XENON1T electronic recoil data

Jean-Philippe Zopounidis



SORBONNE UNIVERSITÉ
École doctorale des Sciences de la Terre et de l'environnement et Physique de l'Univers, Paris - ED 560

Laboratoire de Physique Nucléaire et de Hautes Énergies - UMR 7585

Search for leptonically interacting dark matter with the XENON1T electronic recoil data

Presented by

Jean-Philippe ZOPOUNIDIS

Submitted in fulfillment of the requirements for the degree of
DOCTEUR ES SCIENCES DE SORBONNE UNIVERSITÉ

Supervised by Dr. Luca SCOTTO LAVINA

Defended on 30-09-2020 in front of the comittee :

Dr.	Julien	LAVALLE	Rapporteur
Dr.	Jacob	LAMBLIN	Rapporteur
Prof.	Jules	GASCON	Examinateur
Prof.	Laura	BAUDIS	Examinatrice
Dr.	Jaime	DAWSON	Examinateur
Dr.	Piera Luisa	GHIA	Examinatrice
Prof.	Christophe	BALLAND	Président
Dr.	Luca	SCOTTO LAVINA	Directeur de Thèse

Remerciements

À cet endroit, je voudrais remercier un certain nombre de personnes qui m'ont aidé, d'une manière ou d'une autre, pendant les années de cette thèse et envers lesquelles je suis très reconnaissant.

Je voudrais remercier d'abord mon encadrant Luca Scotto Lavina, pour la confiance qu'il m'a démontrée en me choisissant pour réaliser cette thèse et pour sa disponibilité et son aide pendant toute la durée de celle-ci. La transmission de son expérience de XENON100 m'a été très utile pour entreprendre mes travaux de thèse sur XENON1T. Puisque cette thèse est la première du groupe XENON du LPNHE, je souhaite qu'elle soit un premier pas vers l'agrandissement du groupe à l'aube des générations futures de détecteurs aux liquides nobles.

Ensuite je voudrais aussi englober dans mes remerciements mes collègues Sid, Ernesto et Romain, du groupe XENON, avec lesquels j'ai partagé ces années de thèse.

Je remercie beaucoup Dr. Lydia Ross qui a accepté de suivre en observatrice extérieure mes travaux sous la forme de parrainage de l'école doctorale, et m'a apporté son assistance.

Je n'oublie pas les doctorants avec lesquels nous avons partagé le bureau du LPNHE, Jad, Luis, João et Romain, qui ont contribué à créer quotidiennement une agréable ambiance de travail.

Enfin, je tiens également à remercier ma famille pour son encouragement. Surtout ma mère, Danièle, pour son indéfectible soutien. Elle était toujours à mes côtés et sans son appui, je n'aurais certainement pas pu compléter cette thèse. Je tiens aussi à remercier ma grand-mère, mon père et ma sœur pour leur réconfort.

Jean-Philippe Zopounidis
Paris, 20-09-2020

Ego tanquam centrum circuli, cui simili modo se habent circumferentiæ partes : tu autem non sic.
Je suis comme le centre du cercle auquel tous les points de la circonférence se rapportent; il n'en est pas ainsi de toi.

Vita nuova
Dante Alighieri

Contents

Introduction

It is a historical fact that the 20th century has been revolutionary for Physics. The fundamental change that took place in the understanding of the physical world has been slowly prepared in previous centuries. It is then, that one can find the resurgence of the idea about the atom, initially proposed by the pre-Socratic philosophers, leading to the foundation and evolution of atomic theory, at the end of the 18th century. This development, in turn, allowed the expansion of classical mechanics which, until then, concerned an idealized world, continuous and filled with strict geometric shapes, to the application of its fundamental axioms into the existing matter itself. The result was the triumph of the interpretation of thermodynamic phenomena. A set of rules based on a rigorous foundation but treating samples of matter macroscopically, providing only empirical relations regarding its response to actions of the environment, now acquires a profoundly deeper meaning, within the framework of Statistical Mechanics. However, in the evolution of the theoretical understanding of Nature manifested through its experimental investigation, paradoxes always arose. These were, almost always, related to the inability of the concepts emerging from the theoretical understanding to describe the *factum brutum* of observations. Thus, for example, the problematic concept of the action-at-a-distance required a radical revision of the ideas about space and time and the conflict between the classical notions of wave and particle hid a more complex reality about the microcosm, finally described by the Quantum Theory.

However, although the two great theories of the 20th century, General Relativity and the Standard Model of particle physics, have led to an impressive understanding of the Universe, modern observations suggest that the understanding of its evolution, from its primordial phase to its present state, requires invoking a new form of matter, generically called **dark matter**, accounting for the 85% of its matter content. Also, in order to describe 95% of the total mass-energy content of the Universe, in addition to dark matter, appeal must be made to a currently unknown form of energy, generically called **dark energy**.

In the present thesis, we focus on the experimental testing of theories beyond the Standard Model, predicting a specific particle nature for the dark matter and, specifically, in the effort to directly detect these particles with the use of low background, underground detectors. Indeed, from a theoretical point of view, one of the most attractive attempts to

elucidate this profound open question of modern physics emerges through supersymmetry theory, where particle candidates for dark matter are proposed, under the generic name of Weakly Interacting Massive Particles (WIMPs). Nevertheless, all relevant experimental efforts, based on underground detectors, trying to directly detect WIMPs elastically scattering off the nuclei of the target medium, have only excluded part of the parametric space of the theory. It seems, then, that the mystery's answer could be sought in alternative theories, predicting other dark matter candidates. For instance, it is possible that dark matter has tree-level interactions with leptons.

In the present thesis, we will examine the possibility of a direct detection of electronic recoils induced by this kind of dark matter, analyzing the data acquired from the XENON1T experiment, operated from end of 2016 until 2018, with 2t of total mass inside its TPC, reaching the lowest electronic recoil background ever achieved in a dark matter detector.

One method to search for leptonically interacting DM is through the modulation which is expected to induce in the ER background of the detector, due to the relative motion between the Earth and the dark matter halo of our Galaxy. In the absence of observation of WIMP-induced nuclear recoils, an alternative interpretation through a leptophilic dark matter, which has axial-vector or pseudoscalar couplings to electrons, finds its place as a natural hypothesis that should be tested, taking advantage of the long-term stability of the electronic recoil background of the detector, achieved for XENON1T. The present thesis focus on the application of statistical methods in the investigation of possible correlations between the electronic recoil event rates with the detector parameters (such as the liquid level, pressures, and temperatures, electrode's voltages variations etc.) for the experiment's scientific run, in various energy ranges. This procedure is necessary in order to establish the temporal stability of the detector, as small variations of the detector parameters may influence signal generation, potentially affecting the acceptances and the measured event rates. Also a sensitivity analysis for the search for a modulated signal is presented, based on Monte Carlo simulations, in order to evaluate the asymptotic distributions of the relevant test statistics and assess the eventual sensitivity of the data to rate modulations.

Then, we focus on the search of leptonically interacting dark matter particles with mass in the sub-GeV scale, emerging in the framework of so-called,"hidden sector" theories, predicting a leptophilic DM with a mass in $\mathcal{O}$(MeV), that could interact with atomic electrons of the target causing inelastic atomic processes, such as ionization. It happens that the double-phased TPC of XENON1T, is characterized by an exceptional sensitivity for the detection of small charge signals, derived from ionization of xenon atoms, down to the level of one electron. Therefore, the detector is sensitive to the inelastic quantum processes occurring at the atomic level and, *par excellence*, to the ionization process of even a single atom. And yet, it is in this very region of the leptophilic DM expected signal that a huge and, so far, incomprehensible background arises, driven by the so-called

single electrons, small charge signals that appear to derive from a multitude of quantum processes in the atomic level of LXe, for many and yet unknown reasons.

Here, of course, a rich scientific subject emerges in an attempt to explore the nature of this background and eventually to develop a background model. However, part of this background, related to the photoionization of impurities in the liquid Xe, by the scintillation light produced after an main event inside the detector, has already been studied and it has been proposed that it could be used as a natural calibration source to characterize the detector. The second chapter of this thesis focus on this very purpose, using this part of the background of single electrons to infer parameters of the detector, like the secondary scintillation gain and the extraction efficiency (the probability for an electron to be extracted to the gaseous phase), and their dependence on the extraction electric field. A set of parameters that, in addition to their inherent physical importance, are also key elements for Monte Carlo simulations of the detector. Then, in the final chapter, using part of the spectrum of single electrons during the scientific run, we develop the framework with which we could set constrains for sub-GeV dark matter, scattering off electrons with interactions via the exchange of a heavy mediator or a magnetic dipole moment coupling, or exchange of a massless or ultra-light mediator.

What can we infer about leptophilic DM without the possibility of background subtraction due to the incomplete background model? Can we classify the processes that create the single electron background? How can state-of-the-art machine learning techniques, like Neural Networks, mitigate such backgrounds, making use of partial knowledge of them in conjunction with the topological characteristics of the corresponding events? These are some of the questions, about this exciting topic of direct DM detection, that we will try to address in this work.

Chapter 1

The Dark Matter problem and the search for a direct detection

1.1 Indications for a dark matter

A plethora of astronomical observations points towards the existence of a "dark" matter-as opposed to ordinary, visible matter. We can cite, among others: the dispersion of velocities in the cluster of galaxies Coma [1, 2] and the analysis of the rotational curves of galaxies [3, 4, 5]. These observations, combined with data about the Big Bang nucleosynthesis in the early phases of the universe [6], cosmological constraints from large scale structures [7], distant supernovae [8] and the Cosmic Microwave Background (CMB) [9, 10, 11], tend to demonstrate the existence of a non-luminous, non-baryonic matter. We will briefly present these indications, which extend to different scales and refer to different physical processes. Then we will expand on the interpretive possibilities through elementary particle candidates, that are predicted by theories beyond the standard model. The detection of such particles, through their predicted, extremely rare, interactions with ordinary matter is the purpose of direct dark matter detection experiments.

1.1.1 Cosmological scale

On a cosmological scale, the Cosmic Microwave Background, accidentally discovered in 1965 [12], offers information about the universe 380,000 years after the Big Bang [13]. The CMB spectrum is consistent with a black body of temperature $T = 2.7255 \pm 0.0006$ K [14]. The photons that constitute this background are remnants of that cosmic age where the temperature of the universe fell below 3000 K. It is the, so-called, epoch of recombination of the free electrons with protons, since the photons no longer had enough energy to ionize bound atomic states. However, this spectrum presents very small anisotropies, at the level of 10^{-5}. These are considered to originate from quantum fluctuations, remnants of the inflationary epoch, and are a source of valuable information about the early universe, as

they can be used in order to constrain parameters of cosmological models. Experiments like WMAP and Planck essentially offer high resolution full sky maps of these fluctuations. These maps are usually analyzed using the quantity:

$$\theta(\mathbf{n}) = \frac{T(\mathbf{n}) - \langle T \rangle}{\langle T \rangle} \tag{1.1}$$

which describes the deviation of the temperature of radiation T, from the average temperature $\langle T \rangle$, in a direction $\mathbf{n} = (\theta, \phi)$ in the sky. Since $\theta(\mathbf{n})$ is essentially a function of a two-dimensional projection of the temperature fluctuations, it can be developed in spherical harmonic series as:

$$\theta(\mathbf{n}) = \sum_{l=0}^{l=\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\mathbf{n}) \tag{1.2}$$

with the coefficients a_{lm} given from:

$$a_{lm} = \int \theta(\mathbf{n}) Y^*_{lm}(\mathbf{n}) d\Omega \tag{1.3}$$

From this relation we can define the power spectrum of the fluctuations:

$$\langle a_{l',m'} a_{lm} \rangle = \delta_{ll'} \delta m m' C_l \tag{1.4}$$

In (1.4), the averaging is defined, for each multipole, on an ensemble of observations of the independent modes that accompany it, described by the parameter m. As is known from mathematical physics, for a spherical harmonic expansion there are a total of $2l+1$ independent m-modes, for each multipole l, therefore C_l can be written as:

$$C_l = \frac{1}{2l+1} \sum_{m=-l}^{l} \langle |a_{lm}|^2 \rangle \tag{1.5}$$

From this relation we observe that the calculation of C_l contains an error, related to the number of independent m-modes, specifically given by $\Delta C_l = \sqrt{2/(2l+1)}$. This error inevitably increases for small values of l, as can be also seen from the error bars of the data points of the power spectrum of Fig. (1.1). This is a manifestation of the, so called, cosmic variance: Because the temperature fluctuations we observe are outputs of a random process (and in particular of a Gaussian process), as one covers an increasingly larger portion of the observable universe (smaller and smaller l), the number of independent observations, which would allow us to infer a statistical statement with a desired small uncertainty, decreases. In other words, as we are using the one and only observable Universe, as a laboratory, to study its own composition and evolution, it follows that, for increasingly larger scales of observation, we are allowed to make fewer and fewer independent observations. So for example the sample size from which we determine the value of C_l is finite and becomes smaller for lower multipoles, as we cover

ever larger regions of the unique universe, governed by the physical laws under study. In the expansion (1.2), the term $l = 0$ corresponds to the average temperature and is zero. The dipole term $l = 1$, corresponds to a π rad cover of the sky and there can be seen the effect of the peculiar motion of the observer (Earth) with respect to the CMB. Thus, for the part of the sky "approaching" earth the photons of the CMB are blueshifted and redshifted for the other part of the sky. Thus usually the first two terms of the expansion are not taken into account and the power spectrum is plotted as $l(l+1)C_l/2\pi$ for $l \geq 2$, Fig. (1.1). As there is a direct correspondence between multipole moments l and the angular scale, the maximum value for l is determined by the resolution of the observation. In general, multipoles for $l > 2$ encode important information about the early universe and could be used in order to infer constrains on the cosmological parameters, as they are manifestations of density perturbations before the decoupling moment. For example, the angular scale of the first peak of the power spectrum Fig. (1.1) is determined by the curvature of the Universe, the second and third peaks contain information about the baryon and dark matter densities.

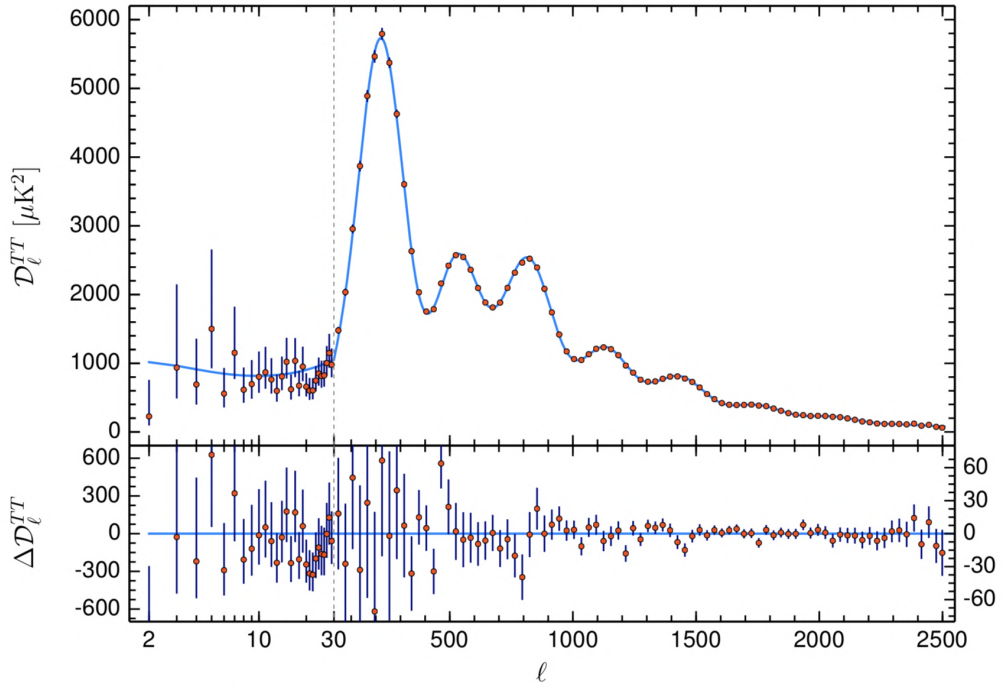


Figure 1.1 – The temperature power spectrum of the CMB from the 2018 results of the Planck Collaboration. Also is shown a theoretical best-fit assuming the ΛCDM cosmology. Also shown are the residuals with respect to the model. Figure from [11].

The so-called ΛCDM (Λ Cold Dark Matter) model [15] gives a theoretical description of the power spectrum of these fluctuations, treating anisotropies of the CMB as small perturbations of the Friedmann–Lemaître–Robertson–Walker metric, with its evolution

being described by the General Relativity. This is a six parameter model containing the baryonic and dark matter density of the universe. Specifically, in this model it is assumed that the energy and matter density of the universe can be written as the sum:

$$\rho = \rho_b + \rho_c + \rho_{rad} + \rho_\Lambda \tag{1.6}$$

In (1.6) ρ_b is the density of baryonic matter, ρ_c the density of the, so called, cold dark matter. This component is hypothesized as consisting of matter other that the ordinary *i.e.*, of non-baryonic origin, with its velocity being much less than the speed of light (cold). It is also assumed that it does not interact with ordinary matter via forces described in the standard model (for example, it does not emit photons), but only through gravitational interactions and possibly other interactions, in the weak scale, described by theories beyond the standard model. ρ_{rad} is the density due to photons and relativistic neutrinos and ρ_Λ, stands for the density of the so-called dark energy, a concept introduced in the model in order to explain the observed accelerated expansion of the universe.

Extracting information about cosmological parameters from the power spectrum of Fig. (1.1), is a complicated process that contains degeneracies between parameters. These can be broken, if constrains of the cosmological parameters are taken into account by other independent measurements, such as data from supernovae or galactic surveys. The Planck collaboration [11], has measured with very high precision the anisotropies of the CMB spectrum from which constrains, with respect to density parameters, can be extracted:

$$\begin{aligned} \Omega_b h^2 &= 0.02207 \pm 0.00033 \\ \Omega_c h^2 &= 0.1196 \pm 0.0031 \end{aligned} \tag{1.7}$$

where h is the reduced Hubble constant, defined as the present-day Hubble parameter divided by 100 km/s/Mpc. The density parameters are related to the observed density ρ[1] as $\Omega = \rho/\rho_{\text{critic}}$. Here ρ_{critic} is the critical density for which the Friedman universe [16], assumed in the ΛCDM model, has a flat geometry. From measurements (1.7) we notice that the density of dark matter is five times that of baryonic matter. The rest of the density of the universe is attributed to dark energy, which formally appears in the Friedman equations as a cosmological constant. This constant compensates the contraction dynamics, due to the gravitational interactions of both the ordinary and the dark matter, resulting in the observed accelerated expansion of the Universe and its geometric flatness.

Another argument, on the cosmological scale, in favour of the existence of non-baryonic matter, is offered by the large-scale structures observed in the universe. Indeed, in order to understand the matter distribution, after the recombination epoch, N-body simulations of dark matter particles have been used [17, 18]. These simulations can be used as evidence that, large-scale dark matter structures, are necessary to explain the present distribution

[1] Defined correspondingly for each component of (1.6)

of matter in galaxies and clusters of galaxies. The same simulations provide indications as to the distribution, later in the history of the universe, of dark matter itself into the smaller-scale structures [19].

In particular, if we take into account that the primordial plasma of baryons and photons can be examined, concerning its dynamic behavior, as a fluid, then the anisotropies observed today in the CMB imply the existence in matter, at the moment of decoupling, of an anisotropy of the same order of magnitude $\delta\rho/\rho \sim 10^{-5}$. If matter had only the baryonic component then this anisotropy, at the present time, it would have grown by $z_{\rm rec} = 1100$ times and in the present age of the universe, the anisotropy of matter should be of the order of 10^{-2}, something that is in stark contrast to existing structures. This observation is consistent with the hypothesis about the presence of a matter component, of a non-baryonic nature, which interacts only in the weak scale with ordinary matter. Therefore, this kind of matter is being in thermodynamic equilibrium with baryons, only in the very early age of the universe. Certainly before the decoupling of ordinary matter from photons, it has already begun to collapse, under its own gravity, forming accretions of matter [20]. After the epoch of recombination, ordinary matter begins to fall into these dark matter concentrations which, in this way, acted as seeds for the later large-scale structures we observe today [21].

1.1.2 Galactic clusters

On smaller scales, of the order of the galactic clusters, it is possible to use an effect predicted from General Relativity, the so-called gravitational lensing [22]. A ray of light is being deflected when it passes through a distribution of mass that locally deforms the geometry of spacetime. This, results in the image of the background being distorted or magnified [23]. Conversely, the position and shape of images of luminous background objects, obtained through this mechanism, can be used to determine the mass distribution of the "lens". It is obvious that this possibility is a first-class tool for the study of the distribution of dark matter, through its gravitational interaction.

We can classify different categories of this effect, as a function of its magnitude, as shown in Fig. (1.2). Its most pronounced form takes place in the so-called strong lensing, where the space-time around a dense concentration of mass is so curved that light rays, initially following a completely different path, can be deflected towards the line of sight of the observer. Thus, the light source can be observed at different positions, corresponding to different initial view angles. Also, the image of the object emitting the light may appear distorted, due to the fact that the magnitude of the deflection, through the object causing the lensing, is a function of the position of emission of the light ray in the surface of the emitting object, if it is spatially extended, such as a galaxy. Through systematic surveys, like the Sloan Lens ACS [24], it was possible to detect galaxies whose image is being distorted by the interposition, in the line of sight, of another galaxy, allowing the

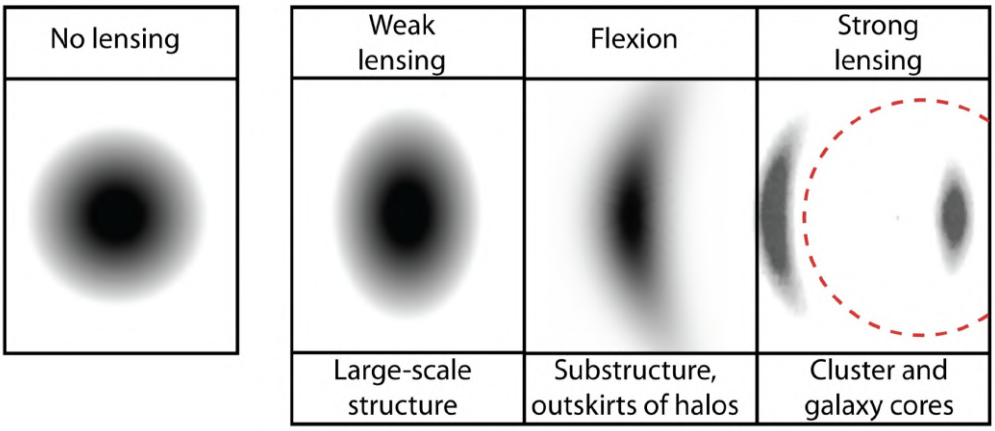


Figure 1.2 – Schematic depiction of the various magnitudes of the gravitational lensing. Is shown the general form of the distortion caused in the background image (no lensing) in the regime of weak and strong lensing, passing from the transitional regime, called flexion. Image taken from [23].

average fraction of dark matter in the galaxy to be measured and also its density profile to be inferred.

In most cases, of course, the effect is not so intense as light rays do not pass through regions dominated by strong gravitational fields so, inevitably, the light deflections are smaller. In this regime, called weak lensing [25], the final image can be considered as a two-dimensional linear transformation of the original, that can be analyzed as a combination of magnification, shear and rotation [26]. Because the effect now is very weak, the analysis cannot be performed on individual objects. However, we can take advantage of the fact that, for instance, the intrinsic shape of elliptical galaxies is uncorrelated. That is, if we look at a large enough ensemble of galaxies, there will be no preferred direction. However the presence of a foreground object, such as a galaxy cluster, can cause shear distortion in the background galaxy shapes that can be perceived as an average shear of an ensemble of galaxies along adjacent lines of sight, that coherently result in an average elliptical shape over the ensemble [23]. Based on this observable field of shear, the mass distribution of the object acting as the gravitational lens can be reconstructed. Observations in more than 300,000 galaxies, have demonstrated the existence of halos of a total mass $1.4 \times 10^{12}\ M_{\odot}$, around galaxies with a total stellar mass of $6 \times 10^{10} M_{\odot}$ [27]. With a combination of weak and strong lensing data, the Hubble Space Telescope SLACS survey, revealed halos of $1.2 \times 10^{13} M_{\odot}$ around elliptical galaxies with mass $2.6 \times 10^{11} M_{\odot}$ [28].

A more dense concentration of mass in the foreground object, begins to cause a curvature in the sheared background image which, in the extreme case of a strong gravitational field, results in the strong lensing regime we described earlier. But this

transitional regime, between weak and strong lensing, is known as flexion [29, 30], This can be considered as the second order effect of the weak lensing, in the sense that, in this case, the mass distribution of the foreground object is characterized by a spatial gradient which makes the magnitude of the shear a function of position onto the image of the background object. Flexing is used to reconstruct the mass distribution, in those regions of gravitational lenses where the deflection of light rays is very small and at the same time, the number of background objects, undergoing the lensing, is too small. In this case, we cannot talk about a strong lensing while, at the same time, a weak lensing analysis, with the suitable averaging procedure described earlier, cannot be applied. For instance, regions surrounding galaxy clusters can be analyzed in this flexing regime.

One of the most striking signs of dark matter concerns the case of the merging galaxy cluster 1E 0657-56, known by the name of Bullet Cluster, which was discovered in 1995 [31]. This object, consists of two colliding clusters of galaxies and we can observe the manifestly different behaviour of its components during the collision. The clusters contain galaxies, gas distributed in the intra-cluster space and possibly dark matter. In these two clusters, that collided relatively recently in cosmic history (just 150 million years ago), the galaxies of the clusters continued their trajectories almost unaffected by the collision as, due to their small size, compared to the clusters as a whole, they have a small collisional cross section. On the contrary, the gas being homogeneously distributed inside the clusters, has a much larger collisional cross section, resulting in a significant deceleration. Hot gases can be detected from X-ray telescopes like Chandra, due to their very intense X-ray emission [32]. The Chandra image also revealed a bow shock, in front of the bullet-like gas cloud, from which it was possible to determine its velocity, relative to the main cluster, thus to determine the moment of the collision, with techniques developed for the use of X-ray data in analyzing merging clusters [33]. On the other hand, weak lensing reconstruction, revealed that the centers of the mass distributions of the two clusters is consistent with the location of the galaxies inside the clusters, with a significance 8σ away of the hot gas peaks. In addition, the total mass of the hot gas is larger, by a factor of two, than the total mass of the clusters, as determined from X-ray and optical images, respectively. But the total reconstructed mass, from weak lensing, is 30-40 times larger than the total visible mass of the clusters, inferred by the galaxies' stars.

This is consistent with the presence of a non-luminous matter that bypassed the gas regions, during the collision, because is only weakly interacting and so is not being decelerated by the collision. From this observation it is also possible to infer an upper bound, with respect to the self-interaction collisional cross section, per unit mass, of this non-luminous matter. It is inferred that an upper limit of $\sigma/m < 1.25$ cm^{2}g^{-1} at a 68% C.L. is needed, in order to explain the observed mass profiles [34].

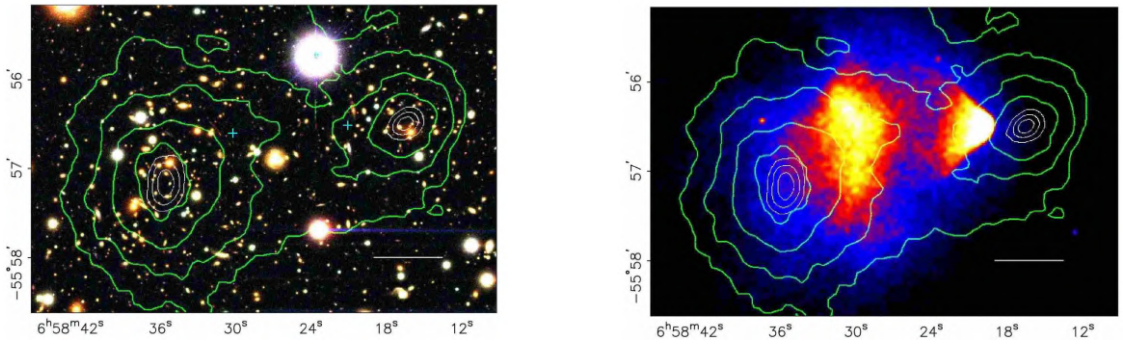


Figure 1.3 – Left: Image of the bullet cluster in visible light, from the Magellan telescope, showing the gravitational contours reconstructed by weak lensing. Their centers are identical to the geometric centers of the two clusters. Right: The gravitational contours of the left image, projected on the X-ray image of the bullet cluster, taken by Chandra. We observe that the intracluster hot gas lag behind the centers of the gravitational contours. Image taken from [35].

1.1.3 Galaxy rotation curves

On smaller scales, a characteristic indication for the existence of dark matter is obtained from the dynamics of clusters and galaxies. In fact, historically one of the first indications about the existence of a non-luminous matter in these objects came from a study by F. Zwicky on the mass of the Coma Cluster, located at a distance of 99Mpc from Earth, in the constellation of Coma Berenices. This object, consists mainly of elliptical galaxies and is almost spherical with a diameter of about 1 Mpc. Considering that galaxy clusters, like that of Coma, containing about 1000 identified galaxies, are ensembles of objects that have been gravitationally bounded for billions of years, it is possible to determine their mass independently of traditional astronomical photometric methods that are based on the ratio of mass to luminosity and thus by definition measuring only the mass emitting light. This is done by application of the virial theorem, that offers the possibility to determine the average kinetic energy of a stable system of components, bounded by potential forces. Specifically, if we have a system with N parts, described by position vectors $(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$ which, throughout the evolution of the system, cover a finite region in space, then if the potential held by the bounded system is a homogenous function $U(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$, it can be shown that the average kinetic energy $\langle K \rangle$ of the system is [36]:

$$\langle K \rangle = \frac{1}{2} \sum_{k=1}^{N} \mathbf{r}_k \frac{\partial U}{\partial \mathbf{r}_k} \tag{1.8}$$

If the potential U is a homogenous function of degree n then, by Euler's theorem on homogeneous functions, (1.8) becomes $2\langle K \rangle = n\mathbf{U}$. For example, in the case of

gravitational attraction $n = -1$. Therefore we have a relation that connects the average potential energy with the average kinetic energy and so, if we consider that the mass distribution in the cluster of galaxies is homogeneous and extends inside the whole spherical region of the cluster with radius R, then the total mass M is:

$$M = \frac{5R\langle u^2 \rangle}{3G} \tag{1.9}$$

where here, the average of the velocity is taken over all components of the cluster and G is the gravitational constant. Zwicky, using Doppler shifts in the galactic spectra, obtained a lower bound for the velocity dispersion and so, from (1.9), he obtained a lower bound on the total mass of the cluster. The value inferred was $M = 4.5 \times 10^{13} M_{\odot}$ which is, orders of magnitudes, inconsistent with the value obtained if one tries to estimate the total mass by photometry, on the 1000 galaxies. In this case, a value of the order of $10^{10} M_{\odot}$ is inferred, which cannot be explained by any systematic error that may have occurred in estimating the mass distribution inside the cluster, as the result would differ by less than an order of magnitude.

In later years it became known that there is about 10% of the total mass of the cluster which is in the form of intragalactic gas and which does not emit light at the visible wavelength, a fact that was not known to Zwicky. But even this, cannot remove the inconsistency between the mass determined by the viral equation and the mass determined by photometry [37, 1, 2].

At the galactic scale, another striking indication about the existence of non-luminous matter arises from the study of the rotational curves of galaxies. Here again, there is a deviation between the observed speed of light-emitting objects, like gas clouds and stars, with what is expected based on their mass, using Newtonian mechanics. In general, for a system of stars and gases orbiting their center of mass, we can assume that the velocity of each object is determined only by the fraction of mass enclosed inside a sphere of radius r, equal to their distance from the galactic center. This is because only this mass determines the intensity of the gravitational field $\mathbf{g}$ on each object. We can verify this using the surface integral of $\mathbf{g}$ inside a sphere S, of radius r:

$$\oint_S \mathbf{g} d\mathbf{S} = 4\pi G M_V \tag{1.10}$$

derived by the Poisson equation for the Newtonian gravitational field. Here M_V is the mass enclosed within a sphere S. From (1.10) it is understood that, the larger this mass, the stronger the gravitational field that determines the velocity of the object in its orbit, as typical of Keplerian orbits (Fig. 1.4), through the relation:

$$u(r) = \sqrt{\frac{GM(r)}{r}} \tag{1.11}$$

where r is the mean distance from the galactic center and $M(r)$ is the total mass enclosed within a sphere of radius r. It is clear from (1.11) that, if the mass of the galaxy follows the light distribution *i.e.*, it is concentrated in a region near the center of the galaxy and, as the distance from the galactic center increases, no significant mass is added, then u(r) should first increase until reaching a maximum value and then decrease, as $u(r) \sim \sqrt{1/r}$, towards the outer regions of the galaxy. On the contrary, the observation imposed the fact that the velocity outside the optical disc of galaxies remained approximately constant. Indeed, the first observation of this discrepancy was made by K. Freeman and concerned the rotational curve of M33 [38] while, in the following years, systematic measurements by Vera Rubin, in dozens of isolated galaxies, resulted to the same conclusion. Because the distances at which the galaxies were located are such that is impossible to use single stars as tracers of the rotational velocity, clouds of hydrogen and helium gas surrounding stars were used, instead. Thus, as material on one side of the galaxy moved away from the observer while, on the other side, approached towards the observer, the Doppler shift of the spectral lines determine the rotational velocities while at the same time giving information about the distance of the examined point from the center of the galaxy.

Given these observations and accepting Newtonian dynamics then, according to (1.10), the observed stability of the rotational velocity, for an increasing radius, by necessity implies the existence of a non-luminous matter which allows the system to remain

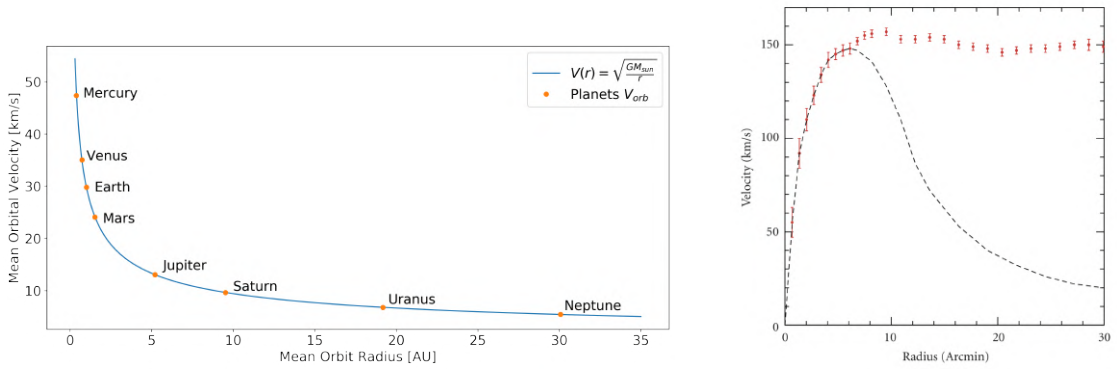


Figure 1.4 – Left: A typical characteristic of a gravitationally bounded system with minor objects rotating around a central object containing nearly all the mass of the system, is that the rotational velocity of the objects is expected to decay as $\sim \sqrt{1/r}$. Here is shown for illustration the mean orbital velocity of the planets of the solar system compared to the velocity expected for a circular keplerian orbit. Right: Comparison of the measurements of velocities of the HI regions of the NGC3198 spiral galaxy situated in the constellation Ursa Major to the Keplerian rotational curve predicted from the mass inferred from the light distribution (black dashed line). Image from [39, 40].

gravitationally bounded. From (1.11), we observe that this stability can be explained if we consider that the non luminous matter, extends beyond the limits of the luminous matter and is characterized by a density profile of the form $\rho \sim 1/r^2$ so that the mass $M(r) \sim r$. We see that the study of the rotational curves of galaxies indicate a model according to which luminous matter is enclosed by a halo of dark matter with a density profile that can be determined by this non-correlation between luminosity and dynamics.

Recently, N-body simulations, concerning the formation of galactic halos, revealed that a universal model, which parameterizes the density of dark matter halos and fit well the observed rotational curves, is provided by the so-called Navarro-Franck-White (NFW) profile [41]:

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2} \tag{1.12}$$

with r_s and ρ_s being parameters determining a scale radius and a scale density. The former is related to the maximum circular velocity and to the concentration of the halo while the latter is related to the halo formation time.

As for Milky Way, our Galaxy presents a rotational curve similar to those observed in many spiral galaxies [42]. In [43] a systematic study of the rotational curve of the Milky Way was performed using all available data from tracers of the galactic potential, using spectral lines from interstellar hydrogen and ionised hydrogen gases (Regions HI and HII), molecular clouds, CO lines, star kinematics and masers mainly concerning the galactocentric distances of 2-20 kpc [44, 45, 46, 47, 48, 49, 50].

These data show that it is not possible for only the baryonic component of galactic mass to interpret the rotational curve and that in the case of the Milky Way it is necessary to consider the existence of a non-luminous and non-baryonic matter which is dispersed throughout the galaxy, even in the solar neighborhood. The important thing about this systematic study is that it is not based on any initial hypothesis about the distribution of dark matter but only on the discrepancy between the observed rotational curve and what is expected from the baryonic component of the galactic matter. In other words, it is an independent model of the dark matter mass in the Milky Way, which is statistically significant at small galactocentric distances and is robust concerning the systematic uncertainties in the determination of the speed.

1.1.4 Nature of dark matter

The evidences listed in the previous section concern observations relating only to the gravitational interactions of dark matter; these gravitational interactions, being universal, do not allow us to identify the microscopic nature of this form of matter. However, the nature and properties of dark matter have, in return, important cosmological consequences.

If the very problem of the interpretation of dark matter, suggested by these astronomical observations, is considered to belong to the domain of particle physics, then

we could reasonably assume that the search for a solution to the well-known problems of the Standard Model becomes, at the same time, itself a search of candidates for the identity of this dark matter. As a result, the dark matter problem becomes a link between astrophysics and cosmology with particle physics. Of course, one of the conditions *sine qua non* for having a link giving an answer to the dark matter problem, is the requirement that the macroscopic behaviour of these candidates, resulting from particle physics, are compatible with astronomical observations, regarding for example, their overall observed density in the Universe today.

Summarizing, from the astronomical observations one can conclude that dark matter unquestionably does not interact electromagnetically. If any interaction with ordinary matter this must be in the weak scale or even weaker. Also, as we saw in the case of the Bullet Cluster, there is a limit to the strength of self-interaction of this dark matter which cannot be larger than the upper bound extracted from the observation. From the combination of CMB data with data from the Big Bang Nucleosynthesis it is possible to determine the baryonic density parameter Ω_b and therefore, from the total matter, to determine the non baryonic DM in the universe today (1.7). A theory of particle physics attempting to describe the microscopic nature of dark matter, must unmistakably interpret this value of the cosmological parameter. Also, since dark matter drives the formation of large scale structures in the later universe, the particles that compose it must be either stable or otherwise have lifetimes of cosmological time scales. In fact, it is possible to extract model independent constrains on this lifetime from the low-l multipoles of the CMB [51].

Regarding the mass of the particles that could constitute this matter, taking into account this small set of indisputable conclusions, which are indicated by the observational data, we can distinguish three categories.

Hot Dark Matter (HDM): In this case the particles are relativistic when disconnected from the thermal bath of the early universe. Their mass must then be of the order of $\mathcal{O}$(eV). An example that could constitute this HDM without having to employ a theory beyond the Standard model, are neutrinos. The problem that emerges from such an interpretation is that the mean free path λ_{mfp} of neutrinos in the early universe is so large that it would not allow the observed clustering scale of the large scale structures of the later universe to be reproduced, if all other cosmological parameters maintain their observed values. This conclusion follows from simulations of the nonlinear growth of structures in an early universe dominated by neutrinos constituting this hot dark matter [52]. Also their speed is not consistent with observations constraining the velocity of DM particles [53].

Warm Dark Matter: In this case the mass of the particles is $\mathcal{O}$(keV). They are still relativistic at the moment of decoupling but become non-relativistic already at the radiation dominated epoch. The smallest λ_{mfp} allows the correction of the HDM behaviour, in terms of its ability to interpret the clustering scale of the later large scale

structures. One particle that could be a candidate for this kind of matter is the sterile neutrino, a hypothetical neutrino that is right-handed and does not interact with any of the Standard Model forces but only gravitationally. Now, all structures on scales smaller than λ_{mfp} is predicted to be suppressed. This presents compatibility problems with the observed structures at scales of 10 Mpc that can be probed from the, so called, "Lyman-α forest" data *i.e.*, Lyman-α absorption lines produced in the spectra of distant galaxies and quasars, due to the Lyman-α electron transition of the galactic neutral hydrogen [54, 55].

Cold Dark Matter (CDM): The particles of this matter can be as heavy as TeV but also new bosons like light axions of axion-like particles with masses below an eV. As the particles are not relativistic, its equation of state is not characterized by radiation pressure which could prevent gravitational collapse. This allows potential wells to be created which will allow ordinary matter after its decoupling from photons to create the observed large scale structures [56]. Indeed, numerical simulations have shown that this dark matter model successfully describes the relevant observations and is also more compatible than the others in terms of the properties of the CMB.

Regarding the candidates for this matter, one idea that was proposed was that it might be a baryon matter in the form of large condensed non-luminous objects such as very faint stars, neutron stars, black holes or even planets, generically called Massive Compact Halo Objects (MACHOs). The EROS-2 collaboration, studied a sample of 7×10^6 bright stars from the Large Magellanic clouds searching for microlensing events that could be created by MACHOs of the Milky Way. These events consist in the apparent change of the brightness of a star due to its alignment with the observer and a relatively small object between them, such as a planet or star that will cause gravitational lensing of much less intensity than the weak lensing [57]. From the visual depth towards the Large Magellanic Cloud that would be caused by the presence of such objects in the Halo of the Milky Way, it was determined that they could not constitute more than 8% of the mass of Halo [58, 59].

Most of the scientific community considers that invoking the existence of non-baryon matter is inevitable for this category of dark matter as well, and in this case there are various particles emerging from theories beyond the standard model. One example is the axion [60, 61] which was proposed in 1977 as a solution to the strong CP problem in quantum chromodynamics and arises in theory as a result of the spontaneous breaking of a new global symmetry. This particle is a boson which can play the role of CDM with a very cold Bose-Einstein condensate filling the universe with an appropriate energy density, indicated by physical cosmology. It is noteworthy that a recent analysis of the low-energy electron recoil events of XENON1T, the same dataset with which in chapter 3 we will search for an annual modulation of the event rate, presented an excess over know backgrounds below 7 keV and in the region of 2-3 keV, where a solar axion model presents a statistical significance of 3.5 σ. Of course this inference is not a discovery and other interpretations for the excess are possible such as, for example, that it could be a

manifestation of β-decays from an extremely small presence of tritium inside the detector, a case that cannot be ruled out nor confirmed [62].

One of the most interesting answers in the question about the microscopic nature of dark matter is given by the so-called weakly interacting massive particles (WIMPs). This is a class of particles that emerges from extensions of the Standard Model such as supersymmetry, theories about possible extra dimensions [63, 64] or "little Higgs" models [65]. These particles are expected to interact gravitationally and also with forces in the weak scale or even weaker, with masses of the order of GeV or even TeV. From a cosmological point of view, they are predicted to have been thermally produced in the early universe and, because of their mass and self-interaction cross section, which is required to be in the weak scale, they produce the DM relic abundance observed today.

Because the WIMP paradigm is one of the most well-established particle candidates for dark matter and because it is the main object of direct searches through the detection of nuclear recoils that are expected to be induced by WIMPs in the target medium of a direct detection detector, in the next subsection, we will go into some theoretical details, as well as in the subsequent production mechanism of the dark matter relic abundance.

WIMPs emerge naturally in the theory of supersymmetry [66] which was proposed as a solution of the Hierarchy problem in particle physics, related to the question about the cause of the huge discrepancy between the weak and the Planck energy scales [67, 68, 69]. This theory is important for particle physics as, in addition to solving the Hierarchy problem, it also predicts the unification of the three gauge interactions of the SM at the energy scale of 2×10^{16} GeV [70, 71]. In the framework of this theory, a symmetry between bosons and fermions is introduced where each particle of the SM with a half-integer (integer) spin, has a supersymmetric partner with an integer (half-integer) spin. This symmetry is able to solve the problem that arises in SM regarding the mass of the Higgs boson.

In particular we will be interested in supersymmetric partners which do not have an electric charge. These, in the supersymmetric models, are the two supersymmetric partners of the two Higgs bosons, which are needed in order to give mass to all the fermions, and the two supersymmetric partners of the electroweak bosons, the weak hypercharge boson B and the weak isospin boson W^0. These partners are the Higgsinos $\tilde{H}_d, \tilde{H}_u$, the bino $\tilde{B}$ and the wino $\tilde{W}^0$. These fermionic states are distinguished only by their electroweak quantum number and, after the spontaneous breaking of the electroweak symmetry, they mix so as to form eigenstates of mass. These eigenstates are called **neutralinos** and are symbolized by a series of increasing mass states:

$$\chi = (\chi_1, \chi_2, \chi_3, \chi_4)$$

If we admit super-partners as a solution to the hierarchy problem then, by introducing a new quantum number, the so called $R-$parity $R_p = (-1)^{3(B-L)+2S}$ – where B, L and S

are respectively the baryonic number, the leptonic number and the spin – the eigenstate of the smallest mass of neutralinos does not decay into SM particles and becomes stable.

This eigenstate, χ_1 constitutes one of the best studied WIMP candidates because its cosmological behavior and its density are compatible with the density observations of dark matter [72]. To show this, we must consider on the one hand the stability of χ_1, and on the other hand the annihilation cross section which determines the evolution of the relic density. In the primordial universe WIMPs are in a thermodynamic equilibrium with the SM particles (creation/annihilation). During the expansion, the temperature of the universe decreases; the numerical density of χ also decreases, according to statistical mechanics ($\propto \exp(-m_\chi/T)$). But because at the same time the universe was expanding, the number density of WIMPs did not tend towards zero, as it would have happened otherwise. Instead, when $n_{eq}\langle\sigma_A u\rangle \sim H\ \sigma_A$, where n_{eq} is the number density of χ in the state of equilibrium, σ_A is the annihilation cross section and H the Hubble constant, then, speaking in terms of statistical mechanics, the degree of freedom of annihilation freezes and the number density will remain constant. That is, because the annihilation rate will be, from this point in time and thereafter, greater than the Hubble constant. Differently stated, one could say that the mean free path (with respect to the annihilation process) has become greater than the size of the universe. Thus the gas of the WIMPs has become so diluted that the annihilation rate is extremely small.

This effect is described using the "competition" equation of the process $\chi\chi \to f\bar{f}$ and of its inverse $f\bar{f} \to \chi\chi$:

$$\frac{dn}{dt} = -3Hn - \langle\sigma_A u\rangle(n^2 - n_{eq}^2) \tag{1.13}$$

Where n is the number density of dark matter particles χ. From (1.13) results the relic density of WIMPs :

$$\Omega_X \sim \frac{x_f T_0^3}{\rho_c M_{Pl}}\langle\sigma_A u\rangle^{-1}$$

Where $x_f = m_\chi/T_f$ with T_f the temperature at the time of freeze-out and T_0 the present temperature of the CMB. The theoretical annihilation cross section for WIMPs is such that, regarding the relation (1.13), if we wanted to explain the (100%) of dark matter with this WIMP paradigm then $m_\chi \sim$ 100 GeV-1 TeV, while to explain the 10% we need $m_\chi \sim$ 30-300 GeV, which corresponds to the masses of the weak scale. Thus we observe the remarkable phenomenon that a theory developed to give a solution to problems intrinsic in particle physics, proposes a stable particle, of mass of the order of the weak scale, which happen to give a compatible relic density with that deduced from astrophysical observations. This concordance explains the attraction of WIMPs as a model of the microscopic nature of dark matter.

1.1.5 Direct detection

Different models have been proposed for the phase space distribution of dark matter in the galaxy. One of the usual assumptions is that of a local energy density of the order of $\rho_{DM} = 0.4$ GeV/(c^2cm^3) and a Maxwellian distribution of its velocity as well as some models about the galactic halo and its (space and velocity) substructures, which also is of importance in the field of direct detection. If we consider as the typical mass of a WIMP $\mathcal{O}$(GeV/c^2) with a speed of the order of magnitude of the speed of our solar system around the galaxy $u \sim 254$ km/s [73], then the WIMPs elastically scattering off the nuclei of the target medium of a detector, give them a recoil energy of the order of $\mathcal{O}$(keV), which allows them, in principle, to be detected by very sensitive detectors, placed in an underground laboratory so as to be protected from background originating from cosmic radiation. The scattering of WIMPs by SM particles can be considered as spin independent for Lagrangian density terms of the form $\bar{X}X\bar{q}q$ or by spin dependent terms under the form of $\bar{X}\gamma^\mu\gamma^5 X\bar{q}\gamma_\mu\gamma^5 q$. Experiments optimized for the detection by interactions of these two types made it possible to set upper limits (in the case of results consistent with a background-only hypothesis) on the corresponding cross sections, as shown in Fig. (1.5), and thereby to set limits on the parametric space of the theory.

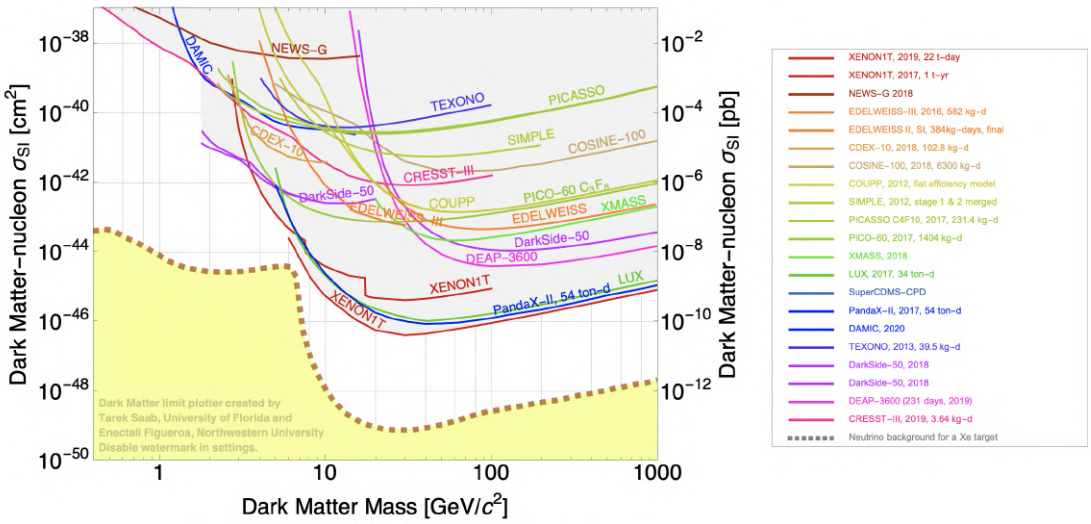


Figure 1.5 – Curent upper limits on the spin-independent nucleon-WIMP interaction cross section as a function of the mass of WIMP, from a plethora of direct detection experiments. This combined plot is generated using the SuperCDMS Dark Matter Limit Plotter https://supercdms.slac.stanford.edu/dark-matter-limit-plotter. The lower limit in yellow correspond to the Solar neutrino floor with a Xenon target. Details about the corresponding experimental limits can be found at [74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94].

We can distinguish two types of direct detection: the non-directional direct direction and the directional one. The first focuses on the measurement of the energy of nuclear recoil induced by the scattering of WIMPs and not on obtaining information on the direction of these recoils. The main purpose of this method is to measure the event rate as a function of the recoil energy which, for WIMP speeds that exceed the minimum speed u_{min} required to produce a recoil with energy E_R by elastic diffusion, can be described theoretically by a relation of the type:

$$\frac{dR}{dE_R} = \frac{\rho_{DM}}{2m_\chi \mu_R^2}\left[\sigma_0^{SD} F_{SD}^2(E_R) + \sigma_0^{SI} F_{SI}^2(E_R)\right] \int_{u_{min}}^{\infty} \frac{f(\mathbf{v})}{u} d^3u$$

Where μ_R is the reduced mass of the WIMP and the nucleus. This relation summarizes the relationship between astrophysics and particle physics that emerges in the study of dark matter. The terms σ_0^{SI} and σ_0^{SD} are the spin-independent and spin-dependent WIMP-nucleon interaction cross sections, which also contain the dependence on the distribution of quarks within nucleons; F_{SI} and F_{SD} are the nuclear form factors which relate to the internal structure of the nucleus and the terms ρ_0 and $f(\mathbf{v})$ are respectively the local halo density and the velocity distribution in the galactic halo, thus providing the link with astrophysics but also the dependence of the expected event rate on the model that will be used for the halo. It is obvious that the velocity distribution of dark matter in the galactic halo is unknown but it can be modeled, for example, with statistical mechanics, by considering the halo as a gas of dark matter particles, essentially without collisions, whose distribution, in phase space, follows the Boltzmann equation. Thus, we can for example propose, based on the rotation curves of galaxies, a particular function for the density of dark matter, which explains the observation of the plateau in the rotational velocity of stars as a function of their distance from galactic center. By solving the Boltzmann equation, using this density, we can prove the following relation [95] :

$$f(\mathbf{v}) = \frac{1}{\sqrt{3\pi\sigma_u^2}} \exp\left(-\frac{\mathbf{v}^2}{2\sigma_u^2}\right)$$

Where σ_u is the dispersion of velocities of WIMPs in the galactic halo[2]. It is this choice which constitutes the Standard Halo Model (SHM) where the gas of WIMPs is supposed not to be in rotation around the center of mass of the galaxy with its density being of the order of $\rho = 0.4$ GeV/(c^2cm^3). In order to measure the event rate and to observe a excess of events with respect to the expected background or to set limits on the cross sections, it is necessary to discriminate between the electron recoil events from nuclear recoils. This is the object of direct dark matter detectors striving for reduced backgrounds, that are due mainly to intrinsic radioactivity from detector components, and therefore increasing their sensitivity in detecting a possible WIMP signal. For example in

[2]The parameters of the velocity distribution of dark matter in the galactic halo can be determined with N-body simulations.

the case of the XENON experiment, in the present status the dominant contribution in nuclear recoils background is due to radiogenic neutrons, produced by the materials of the detector. But even if such background could be strongly suppressed, the coherent neutrino-nucleus scattering (CNNS) from solar and atmospheric neutrinos, and those from supernovae, constitutes an irreducible background and imposes a lower limit on the efficiency of direct non-directional detection. Even if one succeeded in eliminating all other backgrounds, it will be impossible to discriminate between WIMP events and neutrino scattering off nuclei, below a cross section of spin-independent WIMP-nucleon interaction of $\sigma_{SI} = 10^{-49}$ cm^2. In this case the only viable alternative solution is the search for information on the direction of nuclear recoil *i.e.*, direct directional detection or the search for an annual modulation of events that is expected due to the rotation of the earth around the sun.

1.2 The XENON1T dark matter detector

After briefly describing the historical evolution of the XENON dark matter program, we will present the detection principle using a dual-phase TPC filled with xenon as well as specifically the XENON1T detector in order to present its details and to define the concepts that we will use in the analyses that will be presented in the following chapters of this work.

1.2.1 The XENON program

Dual-phase Time Projection Chambers (TPC), using liquid xenon (LXe) as their target medium have demonstrated the ability to achieve exceptional sensitivity in the direct detection of dark matter, through their ability to achieve low electronic and nuclear recoil background conditions.

The XENON dark matter project started in 2005 with its first detector, XENON10, acquiring data between October 6, 2006, and February 14, 2007 using a total mass of 15 kg of LXe, corresponding to a total exposure of 316.4 kg $\times$ days. The analysis resulted in an exclusion limit of 8.8×10^{-44} cm^2 concerning the spin-independent WIMP-nucleon cross section, for a WIMP mass of 100 GeV/c^2 [96]. Also the lower energy threshold from these data allowed a search for a light dark matter making use only of the secondary scintillation signal [97, 98]. In chapter 4 we will perform a similar search for light dark matter with the XENON1T data that will allow us to set limits in the cross section of a possible dark matter-electron interaction and therefore constrain leptonically interacting dark matter models.

In the following years, the XENON100 detector led to an improvement of the results concerning the exclusion limits as the data were still compatible with the background only hypothesis. With 161 kg total mass an exposure of (48 kg $\times$ yr) was acquired between January 2010 and January 2014. In combination with an ultra-low electron recoil background achieved, it became possible to set world leading results on spin independent-nucleon cross section as low as 1.1×10^{-45} cm^2 for a WIMP of 50 GeV/c^2 and and spin dependent WIMP-neutron (proton) limit of 2.0×10^{-40} cm^2 (5.2×10^{-39} cm^2) at 50 GeV/c^2 mass [99]. The same data were analysed in order to get results in various physics channels such as, for example, the investigation of a possible annual modulation of the electron recoil background [100, 101], that is of particular interest and which we will deal with in detail in chapter 3 using the XENON1T data. Using low-energy electron recoils it was also possible to exclude some leptophilic dark matter models as interpretations of the annual modulation signal detected by the DAMA-LIBRA detector [102] (see also chapter 3). Also searches were carried out about a magnetic inelastic scattering [103] or bosonic super-WIMP interactions [104] and exclusion limits on the coupling constants of WIMP-nucleon effective operators in the framework of a non-relativistic effective field theory, taking into account all Galilean-invariant operators up to the second order in

the momentum transfer [105]. It was also possible to study the so-called single electron background and use it as a physical source for the calibration of the detector and the determination of some important physical properties, a technique which we will use extensively in the next chapter for the XENON1T detector [106].

From the end of 2016 until 2018, the XENON1T detector operated with 2t of total mass inside its TPC and one-tonne year exposure, reaching the lowest electronic recoil background ever achieved in a dark matter detector. The analysis of the acquired data resulted once again in the impossibility of excluding the background only hypothesis and, consequently, the current world leading constraint on the spin-independent WIMP-nucleon interaction cross section was set with a minimum of 4.1×10^{-47} cm^2 for a WIMP mass of 30 GeV/c^2 [94]. A series of other physics channels were also probed as we will present in more detail in subsection (1.2.7). The following chapters of this thesis will deal with further analysis of these data for studying the single electron background, the search for an annual modulation in the low-energy electron recoil background and, finally, for setting limits on the dark matter-electron interaction cross-section, using the sensitivity of the detector to small charge signals.

1.2.2 Detection Principle

Production of the Scintillation light

Liquid xenon presents some physical properties that evince its utility as an active medium for the direct detection of WIMPs via the nuclear recoils that are expected to induce in it. The most important of these is its response to radiation through the production of a large number of scintillation photons and ionization electrons for an interaction depositing a certain amount of energy. Its density offers high stopping power and thus through proper fiducialization of the detector's volume it is possible to reduce the rate of external backgrounds in the fiducial volume. Furthermore, the presence of many isotopes, the three most abundant of which are ^{132}Xe (26.9%), ^{129}Xe (26.4%) and ^{131}Xe (21.2%) [107, 108], opens the possibility to probe the spin-dependent interactions of WIMPs with nucleons. From a thermodynamic point of view, its triple point at $T = 161.4$ K, $P = 0.82$ bar [109] is at relatively higher temperatures than other rare gases and thus allows the use of LXe with a reduced degree of difficulty for the process of gas liquefaction.

The process of deposition of energy in liquid xenon and its repartition in various degrees of freedom, is given by the Platzman equation [110] according to which the deposited energy E, can be written as:

$$E = N_i E_i + N_{ex} E_x + N_i \varepsilon \tag{1.14}$$

Where N_i is the number of electron-ion pair created, N_{ex} is the number of exited atoms with E_i and E_x being the average energy needed for the creation of these electron-ion pairs and excited atoms, respectively. The last term concerns free electrons, with average

kinetic energy ε that is lower than the first excited atomic level. Based on (1.14) we can write the average energy W, required to create an electron-ion pair, as:

$$W = E/N_i = E_i + E_x(N_{ex}/N_i) + \varepsilon \tag{1.15}$$

The emission of light, after the deposition of energy in LXe, is called scintillation and is due to the excitation of dimers (excimers) which are created in two ways. Either with the direct excitation of atoms and the consequent creation of excimers or with the recombination of ionization electrons, Fig. (1.6).

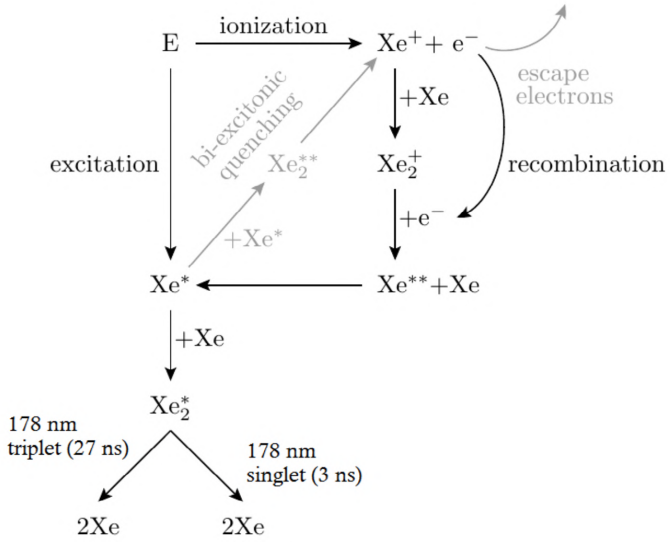


Figure 1.6 – Depiction of the scintillation mechanism in LXe, described in the text. The two main mechanisms, direct excitation and recombination, for the excimer production are shown. The resulting scintillation light has two components related to the singlet and triplet quantum configurations of the excimeric states. Image taken from [111].

The transition from the excimer state to the ground state is accompanied by the emission of vacuum ultraviolet (VUV) light, with wavelength 177.6 nm and a width of 13 nm [112, 113]. The two channels of production of excimers and the subsequent scintillation light, are presented in detail in the following relations [108]:

$$\begin{aligned} Xe^\star + 2Xe &\rightarrow Xe_2^\star + Xe \\ Xe_2^\star &\rightarrow 2Xe + h\nu \end{aligned} \tag{1.16}$$

$$
\begin{aligned}
&\mathrm{Xe}^{+} + \mathrm{Xe} \rightarrow \mathrm{Xe}_2^{+} \\
&\mathrm{Xe}_2^{+} + e^{-} \rightarrow \mathrm{Xe}_2^{+} \\
&\mathrm{Xe}^{**} \rightarrow \mathrm{Xe}^{*} + \mathrm{heat} \\
&\mathrm{Xe}^{*} + 2\mathrm{Xe} \rightarrow \mathrm{Xe}_2^{*} + \mathrm{Xe} \\
&\mathrm{Xe}_2^{*} \rightarrow 2\mathrm{Xe} + h\nu
\end{aligned}
\tag{1.17}
$$

Precisely because of the intermediate step *i.e.*, the excimers formation, the VUV light produced from their decay is not absorbed by the Xe atoms and thus the LXe becomes transparent in the propagation of the scintillation light. If we examine the decay curves of this scintillation we will find that two components are present [114, 112], a short decay followed by a longer one. These can be attributed to the singlet and triplet states of the excimers Xe_2^*. For example, for relativistic electrons in LXe embedded in an electric field of 4 kV/cm, we can observe a characteristic time of about 2 ns from singlet states de-excitation and 27 ns from triplet states. There is a relationship between the way energy is deposited and the relative amplitude of each component. For instance, it has been observed that there is a significant difference in the ratio of intensities of the two components for heavy particles such as alphas of fission fragments compared with β-electrons, due to the dependence of this ratio on the deposited energy density. Thus, the dependence of the scintillation pulse decay shape on the type of particles depositing energy, could theoretically be used to distinguish electron recoils from nuclear recoils, a technique called Pulse Shape Discrimination. However, this is practically applicable only in the case of detectors using liquid argon, as the time difference between the two components is of the order of $\mathcal{O}(\mu s)$ while in the case of the LXe it is of the order of $\mathcal{O}(\mathrm{ns})$ [115, 114].

Dual-phase TPC with LXe

Using a dual-phase TPC such as that depicted in Fig. (1.7), it is possible to take advantage of the partition mechanism of the energy deposited in the LXe to obtain information about the type of particles that interact [116]. Through the process described in Fig. (1.6) a neutron, β-electron, γ or a WIMP, creates a prompt scintillation light and ionization electrons. Scintillation light is observed by photomultipliers as a signal, called S1. The part of the ionization electrons that are not recombined with Xe_2^+ and contributing further to the S1 signal, are drifted by a drift electric field $\mathcal{O}(100$ V/cm) towards the phase boundary where they are extracted to the gaseous phase through a stronger extraction electric field $\mathcal{O}(10$ kV/cm), producing a secondary scintillation light, proportional to the number of extracted electrons [117], called S2. Since, of course, the speed of light is practically infinite compared to the speed of the ionization electrons drifting in the liquid, the time difference between the two signals corresponds exactly to

the time required for the ionization electrons to reach the liquid-gas interface, starting from the point where they were created. Therefore, knowing the drift velocity, which is a function of the drift field, we can obtain the information of the depth of the interaction inside the TPC. The drift field is defined between the negatively biased cathode at the bottom of the TPC and a grounded gate, inserted just below the surface of the LXe (at a 2.5 mm depth for the XENON1T TPC). This region is the active volume where an interaction is in principle possible to produce both signals, S1 and S2. The homogeneity of the field is ensured by a proper placement of field shaping rings. The extraction electric field is defined between the grounded gate electrode and the anode, which is located in the gaseous region of the TPC just above the liquid-gas interface (2.5 mm in XENON1T) resulting in a (5 mm in XENON1T) space surrounding the phase boundary and containing both phases.

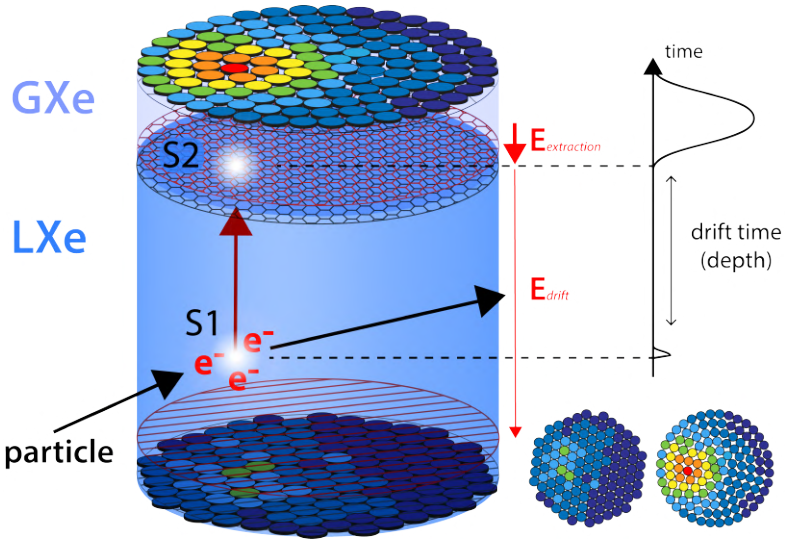


Figure 1.7 – The working principle of a dual-phase TPC.

The light signals are detected by two arrays of PMTs located above the anode in the GXe region and below the cathode, inside the LXe, with a suitable geometric arrangement to maximise the light collection efficiency, as well as, to facilitate the reconstruction of the interaction position. The (x, y) position of an event, at a plane perpendicular to the axis of symmetry of the cylindrical TPC is determined by the hit pattern of the S2 signal on the top PMT array. This, combined with the knowledge of the depth of the interaction, offers a three-dimensional reconstruction of the interaction site. This information, then, allows a fiducial volume to be defined inside the active volume as most events from external background sources are expected in regions close to the TPC boundary. The discrimination between NR and ER can be achieved using the fact that the S1/S2 ratio depends on the type of recoil. This is due to the interplay between the quantity of ionisation and scintillation and the fact that the recombination probability

that (anti)correlates the two quantities, depends on the type of recoil. For long particle tracks[3] the probability of recombination r can be modelized as [118, 119, 120]:

$$r = \frac{A\frac{dE}{dx}}{1+B\frac{dE}{dx}} + C, \quad C = 1 - A/B \tag{1.18}$$

Where dE/dx is the stopping power, proportional to the Linear Energy Transfer (LET) [4], A and B are model parameters and C defines the, so called, Onsager recombination [121] *i.e.*, a constant probability of electrons recombining with the parent xenon ions. Nuclear recoils (NR) expected from WIMPs interactions with the Xe atoms, are characterized by a higher LET than electronic recoils (ER) caused by background sources as betas and gammas. Therefore, for NR a higher probability of recombination is expected and thus the S1 signal is expected to be higher compared to the S1 from ER. Also the S2 signal is expected to be smaller for NR due to the anticorrelation of the two signals. So we can use the S1/S2 ratio as a discrimination parameter, which offers a discrimination between NR and ER greater than 99.5%, at 50% NR acceptance. In Fig. (1.8) is shown the separation of ER and NR events in the parametric space of corrected (S2, S1) for XENON1T (see subsection (1.2.5)). Here the NR are produced from dedicated calibrations using a DD-fusion neutron generator and the ER are produced from ^{214}Pb β-decays originating from calibrations using a ^{220}Rn source (see subsection (1.2.3)).

1.2.3 The XENON1T TPC

In Fig. (1.9), is shown an illustration of the XENON1T TPC. The cylindrical TPC has a diameter of 96 cm and a length of 97 cm, able to contain 2.0 t of LXe. The cylinder is enclosed by 24 interlocking light-tight PTFE (polytetrafluoroethylene) panels used as reflectors of the scintillation light in order to reduce the deterioration of the light collection efficiency resulting from the smaller solid angle coverage of PMTs for large volumes [123, 124]. At the operating temperature of the detector, −96 °C, the length of PTFE panels is reduced by 1.5%. An interlocking design allows the radial dimension to remain constant while the vertical length is reduced.

As described above, it is necessary to use field shaping rings in order to ensure the homogeneity of the drift electric field. For this purpose, the TPC is surrounded by 74 field shaping rings made from low-radioactivity oxygen-free high thermal conductivity (OFHC) copper. The drift electric field is defined between the negatively biased cathode (−12 kV for SR0 and −8 kV for SR1) and the grounded gate, while the extraction field between the gate and the positively biased anode (4 kV). The electrode settings are completed by two screening electrodes situated in front of the top and bottom PMT arrays, in order to minimize the field intensity in front of the photocathodes. The whole electrode settings where optimized by means of electrostatics simulations employing finite elements and

[3] In the sense that they are larger than the mean ionization electron-ion thermalization distance.

[4] LET is the ratio of dE/dx with the density of the material.

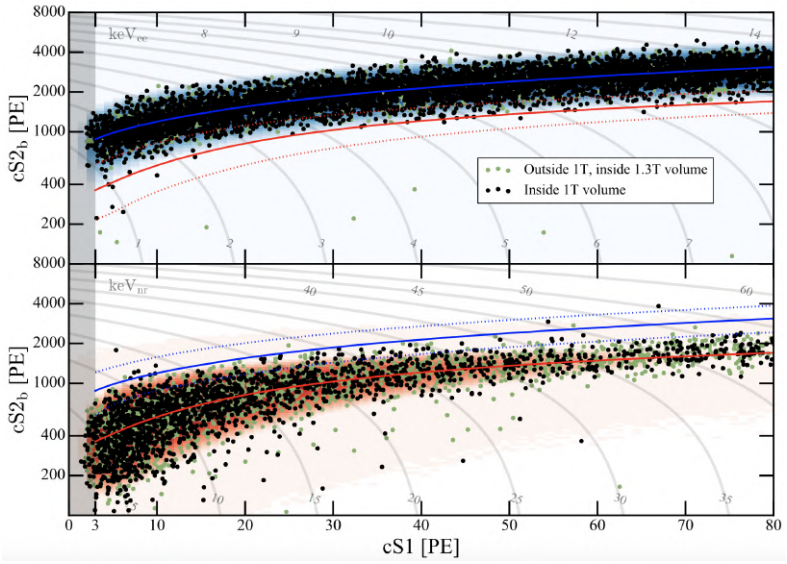


Figure 1.8 – Distribution of events from ^{220}Rn and neutron generator calibrations datasets showing the distribution of ER(top) and NR (bottom) events respectively. The discrimination parametric space here is (cS1, cS2) where the S1 and S2 signals are corrected to take into account detector related effects. Models for ER and NR are extracted and are shown here as $10\% - 50\% - 90\%$ (dotted-solid-dotted) contour lines. Grey contours represent ER and NR energy scales. Figure from [122].

boundary elements methods using software like COMSOL [126] and KEMField [127]. The liquid-gas interface, between the gate and the anode, is formed using a "diving bell", a technique also used in the XENON100 TPC [128]. The height of the liquid level which determines, among others, the gain of the secondary scintillation, is measured using four capacitive level meters installed inside the bell which, with their perimetric placement, can also measure a possible tilt of the TPC. The dynamic range of the four, so-called, short levemeters is 10 mm and have a precision of 30 μm. The LXe level from below the cathode to above the bell is measured by means of two long level meters of 133 mm length, with precision of 2 mm. It is worth noting that the use of the diving bell present the additional advantage that it allows the LXe to rise above the bell, so that there is a layer of LXe, with a thickness of 5cm, above the top PMTs and in the space between the TPC and the cryostat walls and of 3 cm, below the bottom PMTs, acting as a passive shield and offering additional protection from external backgrounds.

The active volume of the TPC is observed by a total of 248 Hamamatsu R11410-21 PMTs of 76.2 mm diameter [129, 130], divided in two arrays. The top array contains 127 PMTs that are radially installed so as to optimize the reconstruction of the position of events in the x-y plane. The bottom array contains 121 PMTs, arranged in a hexagonal close-packed lattice, in order to optimize the light collection efficiency. The PMT are

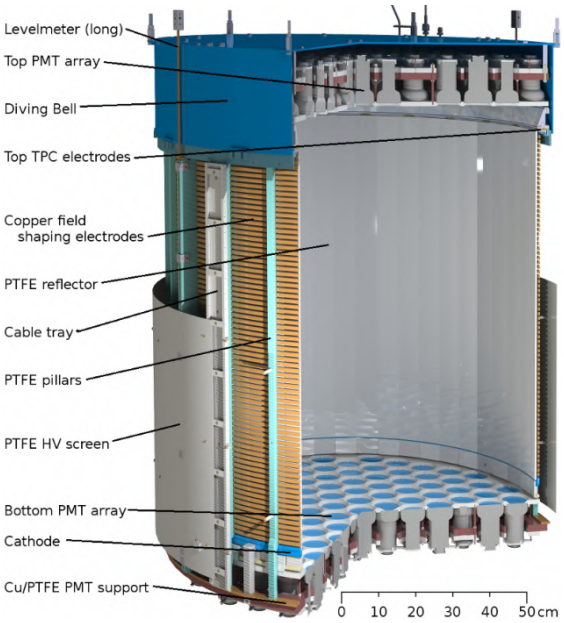


Figure 1.9 – The various parts of the XENON1T TPC. Figure from [125].

designed to operate stably at cryogenic temperature and are characterized by a quantum efficiency of 34.5% at the wavelength of the Xe scintillation light of 178 nm.

All PMTs were screened for their intrinsic radioactivity levels [131] and tested at room temperature and under gaseous nitrogen atmosphere at $-100\ ^{\circ}C$. Then an installation option was made according to their quantum efficiency where, those with the highest efficiency, were placed towards the center of the bottom PMT array in order to maximize the light collection efficiency. Each PMT is held in place by a PTFE plate while PTFE also covers the areas between the PMT windows. The arrays themselves consist of a OFHC copper support plate with cut-outs for each PMT. The 12 dynodes of each PMT are connected to a high-voltage divider with parameters chosen in such a way as to achieve linear operation in the 2.25 V of dynamic range, corresponding to the energy region of interest *i.e.*, below 1.5 MeV. Also, low-pass filters were included on the high-voltage and return line in order to reduce the electronic noise. All cables transmitting the signal are routed through the cryogenic pipe which connects the cryostat with the nearby service building.

The TPC is installed inside a double-walled, cylindrical stainless-steel cryostat made of low-radioactivity material [132]. The inner vessel has a length of 196 cm and a diameter of 110 cm, with its surface being in direct contact with the LXe. The surface was electropolished so as to reduce radon emanation. The outer vessel has a length of 249 cm and a diameter of 162 cm, and is designed large enough in order to integrate the detector of the immediate upgrade stage XENONnT. The cryostat contains the TPC is installed in the center of the water tank and a cryogenic pipe is used to connect the inside of the

TPC with the service building hosting all connections to the cryogenic system concerning cooling, purification, the pressurization of the bell, emergency recovery of the Xe and also the cabling of the PMTs. This cryostat is also attached to a 10m high stainless-steel support frame, (made also by materials chosen for their low intrinsic radioactivity), that is erected inside the water tank. The cryostat is freely suspended from this frame in order that the title of the LXe level with respect to the electrodes can be controlled with a corresponding variation of the length of the three M20 threaded rods that are used to attach it to the frame.

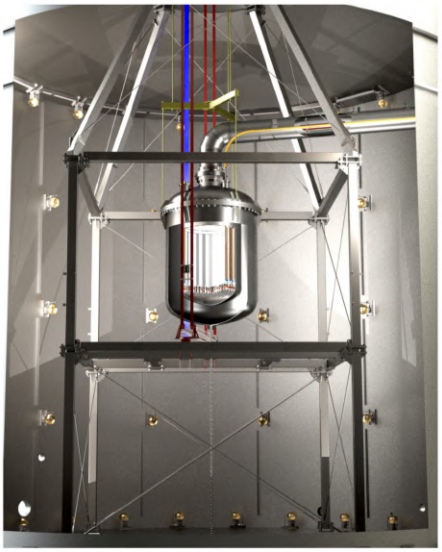


Figure 1.10 – The cryostat hosting the TPC, freely suspended from a support frame inside the water tank. The water tank is equipped with 84 PMTs and is functioning as a Cherenkov detector for cosmic muons. The connection of the interior of the TPC with the service building is done by means of a double-walled vacuum-insulated cryogenic pipe that carries GXe/LXe pipes and all cabling, expect the high voltage cable of the TPC cathode carried by a single-walled pipe (yellow). Blue and red rods represent movable collimators were are installed external sources used for calibration of the detector. Figure from [125].

Every subsystem of the experiment (concerning TPC, Cryogenics, Purification, Recovery and Storage, Muon Veto, Distillation, DAQ, Calibration or Water Tank) is controlled, monitored and recorded by a Slow Control system (SCs). The SCs is based on a hierarchical structure which, given the very large number of parameters is different from the case of the XENON100 monitoring system [133]. This structure is based on a Supervisory Control and Data Acquisition (SCADA) paradigm that is being widely used in industrial instrumentation [134]. The recording of SC data will serve in the present thesis in order to perform a statistical analysis of a possible correlation of their time evolution with the

electronic recoil event rates of the detector, in order to determine the time stability of its operation. The SC system is presented in detail in [135].

Calibration of the Detector

A series of calibration procedures have been designed and performed regularly during the experiment to test the detector response. We will now present the calibration systems used to monitor the gains of the PMTs (LED calibration), to apply corrections for the spatial dependency of the signals (calibration with ^{83m}Kr), monitor the electron lifetime and to infer the response of the detector to ER and NR. To model the former a calibration source of ^{220}Rn is used, while for the latter a neutron generator or a ^{214}AmBe calibration sources is used.

PMT gains calibration The gains are calibrated by stimulating the emission of single photoelectrons from the photocathode by means of low-level light pulses from a blue LED. Four LEDs, installed in the counting room for easy accessibility, are simultaneously controlled by a 4-channel pulse generator. The light is guided into the cryostat via optical fibers. In order to have a uniform illumination of all PMTs, each of the fibers is split into six thin plastic fibers (with a 250 μm core) that transfer the light into the TPC at various angular positions and heights around the field cage. Then, an external signal triggers the pulser and the data acquisition system. The gains were measured on a weakly basis during the experiment with the LEDs producing signals of few photoelectrons (PE).

Usually such calibrations imply a fit of the response of the PMT, using analytical approximations about the single photoelectron (SPE) distribution, for instance, a Gaussian distribution [136] or even models trying to describe the cascade of electrons [137, 138]. However recently a model independent approach has been developed that were used for the XENON1T PMTs calibration, which does not involve any *a priori* assumption about the SPE distribution. This method reduce the bias related to the inability of a particular model to completely describe the SPE distribution [139]. In Fig. (1.11) is shown the gain evolution for three stable PMTs. These are typical cases of the majority of PMTs except specific cases of PMTs that were excluded from the subsequent analyses due to vacuum leaks and light emission causing a decreased performance.

^{83m}Kr calibration Due to the large volume of the TPC and the self-shielding properties of LXe it is impossible to calibrate its central region using external sources like ^{137}Cs. The XENON1T experiment made use of ^{83m}Kr as a calibration source, delivering two monoenergetic low-energy single scatters of ER type, with energies 32.1 keV and 9.4 keV, Fig. (1.12). This isotope is very useful because as a noble gas, it can diffuse uniformly throughout the volume of the TPC. Therefore we have a spatially uniform calibration of the detector in low energy, appropriate for the energy region of interest. On the other hand, its half-life is only 1.8 h *i.e.*, within a few hours from the end of a calibration, its levels have returned to zero and thus does not affect data taking [140]. ^{83m}Kr originates from the electron capture decay of ^{83}Rb. A source of ^{83}Rb is installed

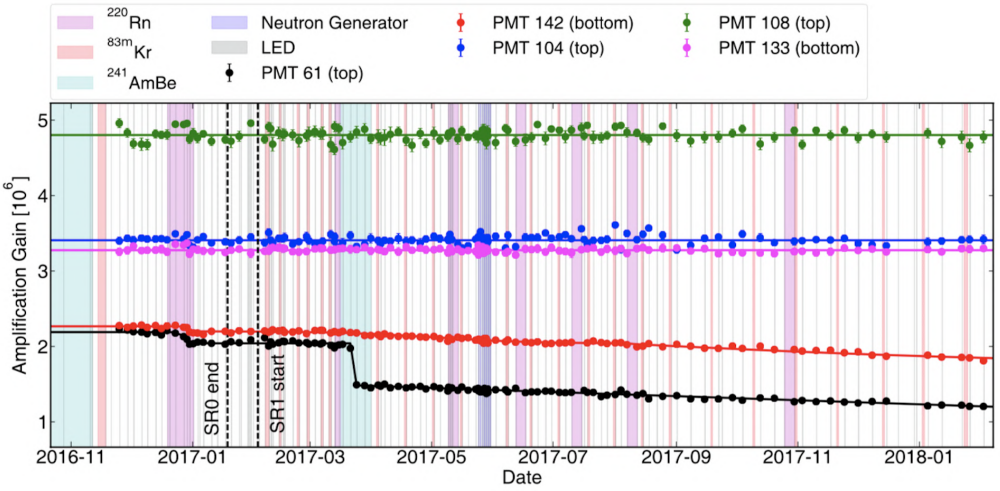


Figure 1.11 – Gain evolution for three PMTs for SR0 and SR1, measured in LED calibrations. Also shown are two PMTs which presented a decreasing performance with time due to vacuum leaks. The coloured bands represent the periods of the various calibrations inserted between the regular background data taking (white regions). Figure from [122].

in the purification system in order for the daughter isotope to be released into the TPC. Normally the long lived ^{83}Rb is not emitted inside the TPC [141]. However, during SR1 it was found that there was a contamination of ^{83m}Kr events in background datasets, possibly originating from an introduction of an amount of ^{83}Rb at some point during SR0, due to an improperly sealed source valve. In (3.2.1) we will present a specific cut that was developed in order to address this contamination at the level of the analysis.

^{220}Rn calibration Another intrinsic calibration source used is again a noble gas, ^{220}Rn with a half-life of 56 s. A source of ^{228}Th is used, which efficiently emanates ^{220}Rn [143]. This isotope is then flushed into the TPC through the stream of GXe. Then, in the ^{220}Rn decay chain Fig. (1.13), α, β and γ radiation is produced that can be used for calibrations in various energy regions, as was also done in the XENON100 detector [144]. For example, the beta decay of ^{212}Pb is used for low-energy ER calibration. The half-life of this process is 10.6 h thus the corresponding activity is completely reduced in 2-3 days after each calibration. Also the α-decays from ^{220}Rn and ^{216}Po, because of their high energy, can be used to derive a correction map for the light collection efficiency across the TPC [145], due to the fact that the response of the PMTs is non linear for such high energy events. This correction is necessary for searches of rare events, as the neutrinoless double-beta decay, in the high energy spectrum. Finally, another possibility provided by this source is that the delayed coincidence between the ^{220}Rn and the ^{216}Po decays can

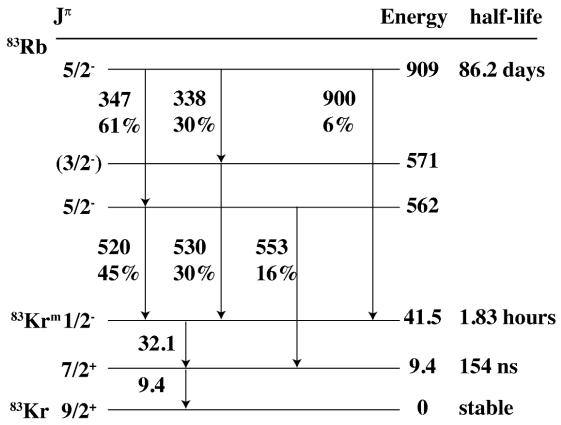


Figure 1.12 – Decay modes of ^{83}Rb to the metastable state ^{83m}Kr and its subsequent decay via emission of two conversion electrons with energies 32.1 keV and 9.4 keV used for the low energy ER calibration of the detector. Figure from [142].

be used to evaluate the atomic velocity as a function of the position, thus obtaining a map of the fluid dynamics of LXe inside the TPC [144].

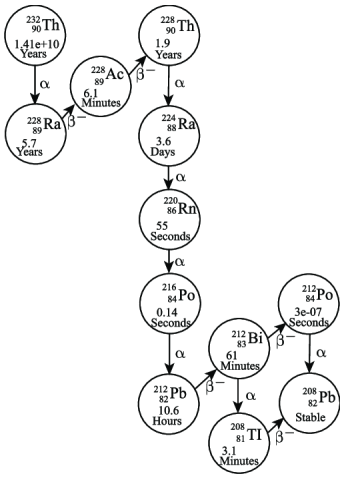


Figure 1.13 – The decay chain of ^{220}Rn. Also shown are the decay modes and their correspondent half-lifes. Figure from [146].

Neutron calibration

The detector response to nuclear recoils is calibrated in two principal ways: First with the use of an ^{241}AmBe source, deployed using the belts shown in Fig. (1.14), allowing also the vertical movement of the source in two angular positions. Also a belt surrounds the cryostat from the bottom. The tungsten collimators, in which the calibration sources are situated, are placed out of the water during normal (background-only) runs. They allow the particles to exit into a cone with 40° opening. This allows the reaching of the

central volume when the collimator is located at half height of the TPC. Second, use is made of a Deuterium-Deuterium (DD) fusion neutron generator, producing neutrons with energies at $\mathcal{O}$(MeV) [147]. This generator consists of a chamber that is filled with ^{2}D. The necessary voltage to start a discharge, between an anode and a cathode inside the chamber, is reached by reduction of the pressure of the deuterium gas. The gas becomes ionized and once in the field cage of the cathode the ions are confined by a strong electric field. When the Coulomb barrier between ions is overcome, the fusion reaction take place:

$$^2\mathrm{D} + ^2\mathrm{D} \rightarrow ^3\mathrm{He}\ (0.82\ \mathrm{MeV}) + n\ (2.45\ \mathrm{MeV}) \tag{1.19}$$

The final energy spectrum of the neutrons is, however, not a single peak centered at 2.45 MeV, but two peaks, at 2.2 MeV and 2.7 MeV. That is because, for the two-body process (1.19), the cross section is dependent on the angle between the momentum of the emitted neutron relative to the momentum of the incident deuteron, resulting in a spectrum, in the lab frame, having two peaks which corresponds to the extreme cases of emission angles π and 0.

Because such a neutron generator can produce a neutron flux of 10^7 n/s it is necessary to be modified in order to reduce this flux to 10 n/s as, for our application of calibration of the response to nuclear recoils, we want to achieve a reduced rate of pile-up events. The generator can also be placed in three locations around the cryostat so as to achieve a uniform illumination of the active volume.

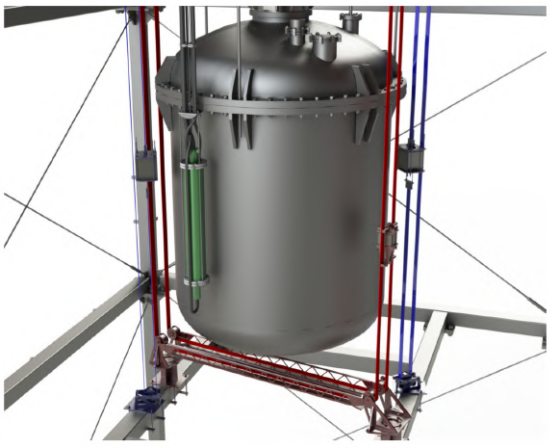


Figure 1.14 – Representation of the system of external calibrations around the cryostat. Blue belts allows a vertical movement of sources placed inside a W-collimator (tungsten alloy (95% W)) for calibrations in two angular positions, while the red belt can also reach the bottom part of the detector. Also shown is the DD-fusion neutron generator positioned vertically near the cryostat.

1.2.4 From LXe microphysics to the observable signals

We will now present how observable signals relate to the microphysics of the intrinsic response of LXe, as well as, the detector reconstruction effects that we must take into account in the analyses that will follow.

The response of LXe to an interacting particle depositing a quantity of energy is described following the model of the Noble Element Simulation Technique (NEST) [148, 149]. Equation (1.14) with the three main ways of energy repartition, ionization, excitation and thermalization, is modified as the latter is undetectable in the XENON1T detector. The deposited energy E can be reconstructed only using the observable quanta, number of excitons and ion-electron pairs, N_{ex} and N_i respectively. We can write the energy deposited in a single interaction as

$$E = W(N_{\text{ex}} + N_i) \tag{1.20}$$

Where W is defined to be a mean work function for the production of one quantum, either exciton or electron-ion pair which, for LXe is found to be (13.7 ± 0.2) eV [148]. Normally we would have to define two different W-values for the two distinct processes (exciton formations and electron-ion pair creation) but these would be difficult to determine experimentally. Instead, an average value describing both can be shown to be sufficient for the interpretation of the experimental results [150, 151]. The energy calculated by (1.20) has advantages over an attempt to calculate it only through scintillation, as in such a case a calibration that would depend on the drift electric field would be required. There would also be non-linearities resulting from the energy dependence of the recombination part of the exciton formation (see Fig. 1.6). Now the difference in the energy scale between electron recoil and nuclear recoil can be described just by the, so-called, Lindhard factor L (see below) [152].

The energy loss due to thermalization of the recoiling particle can be described as a binomial fluctuation of the total number of observable quanta $N_{\text{ex}} + Ni$:

$$N_q \sim \text{Binom}(E/W, L) \tag{1.21}$$

In this relation, L is a factor expressing the fraction of energy loss in thermalization. In the case of electron recoils this loss is negligible because of the very small electron mass compared to the mass of the Xe nucleus, but in the case of nuclear recoils, a non-negligible amount of the kinetic energy in transferred, via elastic scattering, to Xe atoms resulting in an L of 0.1-0.2. This parameter is described by Lindhard's theory [153, 154] as:

$$L = \frac{kg(\varepsilon)}{1 + kg(\varepsilon)}$$

With k a constant and $g(\varepsilon)$ a function of the energy and is proportional to the ratio of the electronic to the nuclear stopping power [155]. Also ε is a dimensionless quantity which, for a nucleus with atomic number Z, is given by [149]:

$$\varepsilon = 11.5(E/\text{keV})Z^{-7/3}$$

As $N_q = N_{ex} + N_i$ and N_q are given by (1.21), the binomial fluctuations of N_i depend on the mean ratio of excitons to ions $\langle N_{\text{ex}}/N_i \rangle$:

$$N_i \sim \text{Binom}\left(N_q, \frac{1}{1+\langle N_{\text{ex}}/N_i \rangle}\right)$$

and also $N_{ex} = N_q - N_I$. In order to calculate the final number of scintillation photons N_γ and ionisation electrons N_e that will eventually generate the signals S1 and S2, we must take into account the recombination process, through the recombination fraction r, which essentially describes the probability of ion-electron recombination and its further contribution to the scintillation light. If we write as $(1-r)$ the probability that an ionisation electron will escape recombination then:

$$\begin{aligned} N_e &\sim \text{Binom}(N_i, 1-r) \\ N_\gamma &= N_i - N_e + N_{\text{ex}} \end{aligned} \tag{1.22}$$

In (1.18) we have described the recombination probability as a function of the linear energy transfer for long tracks. However when the particle tracks are smaller than the thermalization distance of the ionization electrons, then the length of the track is of no relevance and the theory behind equation (1.18) (Onsager Theory) break down. In that regime the so called Thomas-Imel box model is used instead as an alternative and more accurate description of the ionisation process [156]. In that model the mean recombination fraction, which is dependent on the deposited energy and the drift electric field F, is described as:

$$\langle r \rangle = 1 - \frac{\ln(1 + N_i \varsigma/4)}{N_i \varsigma/4} \tag{1.23}$$

And here ς describes the dependence of the recombination on the electric field and is parameterised by a power-law $\varsigma = \gamma F^{-\delta}$ with the parameters γ and δ being extracted from a fit of the detector response model on actual data. One must also take into account the fact that the recombination fraction fluctuates intrinsically [157] and also due to detector effects, such as field non-uniformity (for example near the cathode). If Δr is the recombination fluctuation then one can model r as:

$$r \sim \text{Gauss}(\langle r \rangle, \Delta r)$$

This fluctuation around $\langle r \rangle$ is parameterised with an empirical formula containing two free parameters q_2 and q_3, $\Delta r = q_2(1 - e^{-E/q_3})$. This, takes into account the observation that a constant recombination fluctuation results for energy deposits larger than 2 keV, together with the implicit assumption that $\Delta r \to 0$ as $E \to 0$.

For NR, this modelling is found to be consistent with the data and also that Δr can be set to zero in the detector response model as is much smaller than the fluctuation due to N_e (1.22) [158].

For ER, measurements have shown that, for energies below 3 keV and above 10 keV the process of recombination is still not fully captured by the Thomas-Imel box model. For that reason an empirical modification was introduced to account for these deviations, where the equation (1.23) and ς are modified for ER as:

$$\begin{aligned} \langle r \rangle_{er} &= \frac{1}{1+e^{-(E-q_0)/q_1}} \left(1 - \frac{\ln(1+N_i \varsigma_{er}/4)}{N_i \varsigma_{er}/4}\right) \\ \varsigma_{er} &= \gamma_{er} e^{-E/\omega_{er}} F^{-\delta_{er}} \end{aligned} \tag{1.24}$$

where again the parameters $q_0, q_1, \gamma_{\text{er}}, \omega_{\text{er}}$ and δ_{er} are determined matching the model with actual data obtained from the detector.

Now the mean photon and charge yields can be defined:

$$\begin{aligned} \langle N_\gamma / E \rangle &= \frac{1}{W} \frac{\langle r \rangle + \langle N_{\text{ex}}/N_i \rangle}{1 + \langle N_{\text{ex}}/N_i \rangle} \\ \langle N_e \rangle / E &= \frac{1}{W} \frac{1 - \langle r \rangle}{1 + \langle N_{\text{ex}}/N_i \rangle} \end{aligned} \tag{1.25}$$

using for ER and NR the corresponding quantities from the previous equations. In Fig. (1.15) is shown the mean photon and charge yields for NR and ER for the XENON1T data fit, and comparison from various other measurements.

1.2.5 Reconstruction of the S1 and S2 signals

Let's examine how the physical signals are reconstructed by the detector. Signals at the PMTs exceeding a certain threshold above the baseline are digitized by the data acquisition system [169]. We refer to them as **pulses** of the corresponding PMT channel. An event builder is used online to group these pulses into events, triggering on S1 or S2 candidates and storing a waveform window of 1 ms around each trigger. These grouped pulses are subsequently segmented into smaller intervals called **hits**, by separating individual signals, which may have been grouped into the same pulse waveform. This segmentation is done offline by the data processor PAX [170] developed by the collaboration. Afterwards the hits coming from different channels are grouped into clusters in time, forming structures called **peaks** corresponding to the ionization or scintillation signals. The properties of the peaks (their area, various measures of width[5], their amplitude in the gain-corrected sum waveform), are computed by the data processor. A peak is classified as an S1 if its waveform is rising sufficiently fast and has at least three contributing PMT channels.

[5]That is, the duration containing various percentages of the total area.

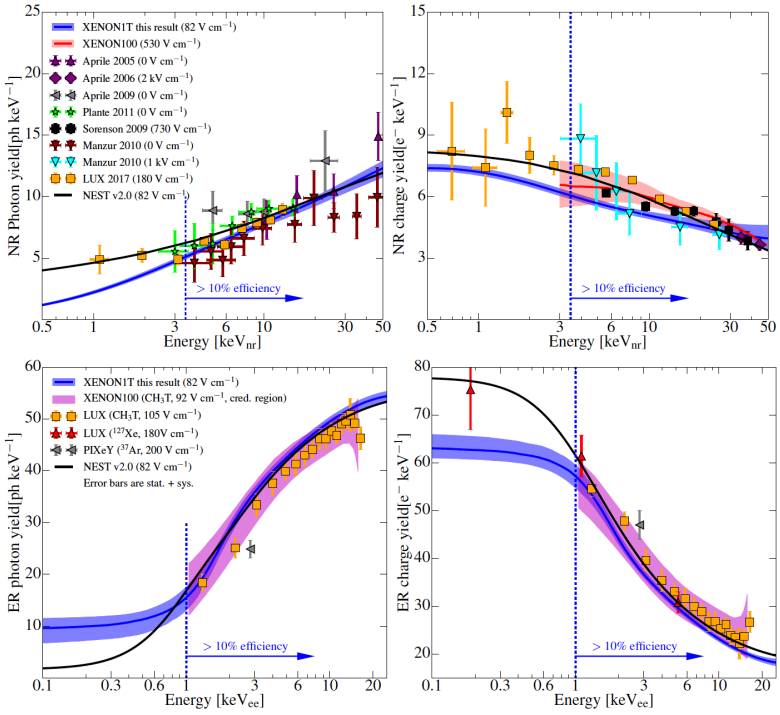


Figure 1.15 – The mean photon and charge yields for NR (upper panel) and ER (lower panel) measured in XENON1T. The blue solid line and region are the median and 15%-85% credible interval. Also the dashed lines and arrows indicate the energy region where the detection efficiency of XENON1T drops to less than 10%. Also shown are the corresponding measurements in XENON100 [158]. The black solid line is the best fit from NEST v2.0 [159] along which are shown various other measurements from [160, 161, 157] for ER and [162, 163, 164, 165, 166, 167, 168] for NR.

An S2 is characterized by a slower waveform rising and is classified as such if it has at least four PMTs contributing Fig. (1.16). For S1s, only hits with maxima within a 100 ns window centered on the maximum of the sum-waveform for all channels are counted for the latter requirement. Then pairings of events, called **interactions**, are formed by searching each event for a valid S1-S2 pairing. The pairing search starts by the largest S1 and S2 of the waveform. For each pair the 3-dimensional position of the event is calculated by the reconstructed lateral (x-y) position of its S2 and the time difference between the maxima of the S1 and the S2 (the drift time). The lateral x-y position of the S2 is computed based on a likelihood maximizer that compares the observed top PMTs hitpatterns with corresponding hitpatterns generated from optical Monte Carlo simulations. The radial resolution achieved was shown to be < 2 cm. Also, the electron drift velocity is measured by the drift time distribution and the known length of the TPC.

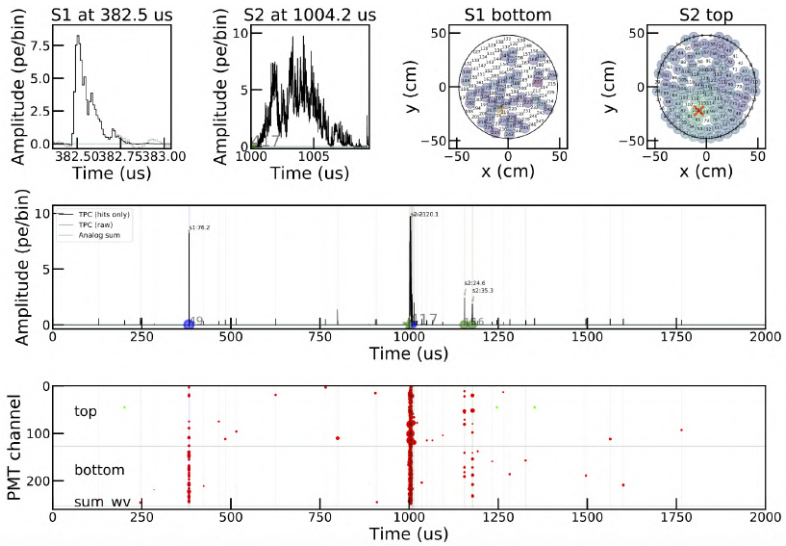


Figure 1.16 – An event, recorded in XENON1T. In the middle panel we see a 2 ms waveform of the event. A peak at 382 μs was identified as an S1 due to its sharp rise as we can see in a zoom presented in the first plot of the top panel. Its area, of 76 PE, was calculated by the data processor. The PMTs contributing in this peak can be seen in the red dots of the lower panel, while their distribution in the bottom PMTs array can be seen in the third plot of the upper panel. Then at a time of 1004 μs a peak was identified as an S2. We can see a zoom of it in the second plot of the upper panel. Its area was evaluated at 2120 PE and the PMT channels contributing are seen in the lower panel. In the last plot of the upper panel is shown the top hitpattern or the S2, used to reconstruct the position of the event.

Further computation of higher level data concerning signal corrections that we will describe below, are done with the custom software package HAX [171].

The performance of the data processor is evaluated using simulated PMT signals from a waveform simulator. This simulator employs data-driven models of XENON1T detector concerning the specific properties like the scintillation light pulse shape, the spatial dependence of the light collection efficiency, the diffusion of electrons during drift, PMT afterpulses, the single electrons generated by photo-ionization of impurities and their time profile (2), and the electronic noise. The simulations are validated by comparison to ^{83m}Kr and neutron calibration data.

1.2.6 Corrections for the reconstruction effects

We have described the generation of the observable signals from the LXe microphysics, after an energy deposition inside the active target of the detector, and the reconstruction of events from the data processor. We must now take into account the reconstruction

effects related to the detector parameters. For example, the position dependence of the S1 and S2 signals due to the spatial dependence of the light collection efficiency, the loss of ionisation electrons due to their attachment to electronegative impurities in LXe, during their drift towards the phase boundary, the effects of single and double photoelectron emission at the photocathode of the PMTs [172, 173].

Between the photons produced by scintillation and their detection as an S1, a series of efficiencies are inserted, associated with the light collection efficiency ε_L, the average quantum efficiency and the average photocathode collection efficiency of the PMTs, ε_{QE} and ε_{CE} respectively. Concerning the S2 signal, this is formed by the amplification of the ionisation electron cloud, drifted towards the phase boundary. This amplification is described by the gas gain G *i.e.*, the number of photoelectrons per electron extracted in the GXe, which is dependent of the x-y position of extraction. Also there is a probability of extraction for each electron reaching the phase boundary, the extraction efficiency $\mathcal{E}$ which also has a spatial dependence. The extraction efficiency will be studied in detail in Section (2.10). Now, if p_{dpe} is the probability for the photocathode of the PMT to emit two photoelectrons when absorbing a photon, then we can define the energy scale parameter $g_1'(x,y,z)$ as the average number of photoelectrons produced per primary scintillation photon and the amplification of the charge signal $g_2'(x,y)$:

$$\begin{aligned} g_1'(x,y,z) &= (1+p_{\text{dpe}})\varepsilon_L\varepsilon_{QE}\varepsilon_{CE} \\ g_2'(x,y) &= \mathcal{E}(x,y)G(x,y) \end{aligned} \tag{1.26}$$

What is used in the various analyses is the average over the active volume of (1.26), denoted g_1 and g_2. These average values will be used also in Section (3.2.1) to define a combined energy scale. If we denote as N_{hit} the number of hits detected and N_{pe} the number of photoelectrons generated from the photocathode of the PMTs, then these could be described by the following binomial processes:

$$\begin{aligned} N_{\text{hit}} &\sim \text{Binom}(N_{\text{gamma}}, \varepsilon_L\varepsilon_{QE}\varepsilon_{CE}) \\ N_{\text{pe}} - N_{\text{hit}} &\sim \text{Binom}(N_{\text{hit}}, p_{\text{dpe}}) \end{aligned} \tag{1.27}$$

However (1.26) does not capture the dependence of the S2 signal on the depth of the events. This dependence arise from the fact that the ionisation electrons, drifting towards the phase boundary can be attached to electronegative elements such as oxide, carbide and nitride impurities existing in ppb concentrations (in the $\mathbf{O}_2$-equivalent level) within the LXe of the detector and derived by outgassing from the detector's materials. This attachment of impurities S with drifting electrons e, consists in a formation of a negative ion:

$$e + S \rightarrow S^-$$

a process which subsequently weaken the ionisation signal. For a certain ionisation signal, the concentration of such impurities $[S]$ and the concentration of the ionisation electrons $[e]$ can be described by a relation of the form:

$$\frac{d[e]}{dt} = -k_s[e][S] \tag{1.28}$$

Where k_s describes the attachment rate. Solving (1.28) for a constant $[S]$ we get the temporal variation of $[e]$:

$$[e(t)] = e(0)e^{-k_s[S]t}$$

from which we can define a characteristic time $\tau_e = (k_s[S])^{-1}$ for the process of the electron attachment. In that context, this will be referred to as the **electron lifetime**. The ratio of this characteristic time to the maximum drift time, determine the maximum charge loss during the drift. To keep the concentration of these impurities in the ppb level, the Xe gas must be continuously purified. This is achieved by a purification loop, where a rare gas purifier (getter) removes the impurities via the formation of chemical bonds with the getter material (zirconium).

Finally the number of electrons survived the drift and the extraction N_{ext} can be described by:

$$N_{\text{ext}} \sim \text{Binom}(N_e, e^{-z/(\tau_e u_d)}\mathcal{E})$$

where u_d is the drift velocity. The secondary scintillation light produced can be approximated as:

$$N_{\text{prop}} \sim \text{Gauss}(N_{\text{ext}}G, \sqrt{N_{\text{ext}}}\Delta G)$$

with ΔG being the spread of the gas gain, determined with the method presented in Chapter (2). There we will also present an alternative model for the secondary scintillation light which best describes the light produced for the case only a few electrons are being extracted.

The S1 and S2 signals are then constructed from N_{pe} and N_{prop} by the digitizer and the clustering and classification software. We account for the biases and the fluctuations inserted in this process, writing S1 and S2 as:

$$\begin{aligned} S1/N_{\text{pe}} - 1 &\sim \text{Gauss}(\delta_{s1}, \Delta\delta_{s1}) \\ S2/N_{\text{prop}} - 1 &\sim \text{Gauss}(\delta_{s2}, \Delta\delta_{s2}) \end{aligned} \tag{1.29}$$

where δ and $\Delta\delta$ for S1 and S2 are parameters accounting for the bias and spread of the reconstruction, respectively. These biases are estimated from the waveform simulator.

Finally the two signals are corrected for their spatial dependence. Assuming that the reconstruction fluctuations in the x and y axes are identical, we can write the reconstructed position $\mathbf{x}_r = (x_r, y_r)$ as:

$$\mathbf{x}_r \sim \text{Gauss}(\mathbf{x}, \sigma_p)$$

with σ_p the position resolution. The corrected S1 and S2, denoted cS1 and cS2, are written as:

$$\begin{aligned} \text{cS1} &= \text{S1}\frac{g_1}{g_1'(x_r, y_r, z_r)} \\ \text{cS2} &= \text{S2}\frac{g_2}{g_2'(x_r, y_r)e^{z/(\langle\tau_e\rangle u_d)}} \end{aligned} \tag{1.30}$$

with $\langle\tau_e\rangle$ the average electron lifetime measured. This is evaluated in intervals of 2-3 weeks using the 41.5 keV line from ^{83m}Kr calibrations, fitting the S2 signal as a function of the drift time. The same quantity can be evaluated using any other monoenergetic line, for example α-decays of ^{222}Rn and ^{218}Po observed in background data. It was observed that these two methods present an offset of 10% of unknown origin, Fig. (1.17). It was hypothesised that the discrepancy originates from small inhomogeneities of the drift field, as the charge yield of α-decays has a larger dependence on the electric field, compared to NR and ER. Due to the fact that the line of ^{83m}Kr source is closer to the energy region of interest for dark matter searches, it was decided to use the τ_e measured from the ^{83m}Kr line for the S2 signal correction.

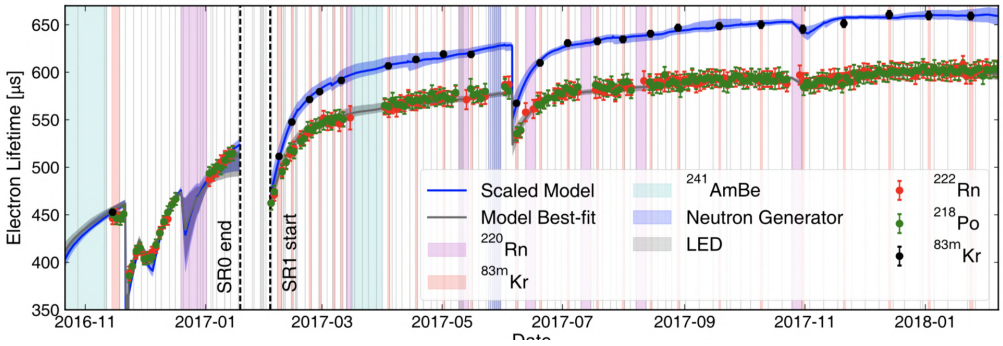


Figure 1.17 – Evolution of the electron lifetime, during SR0 and SR1. Here are shown measurements of the electron lifetime from ^{83m}Kr, ^{222}Rn and ^{218}Po decays. The sudden decreases corresponds to changes in the detector parameters accompanied by releases of impurities which are then removed through the purification process.

1.2.7 Some scientific results of the XENON1T experiment

We will now present some scientific results from the analyses of the data of SR0 and SR1. The analyses presented in Chapter 3 and 4 will be two additional results that will be added to those of this section. In Fig. (1.18) is shown the evolution of the livetime acquired during the two scientific runs accumulating the data that will be used in subsequent analyses.

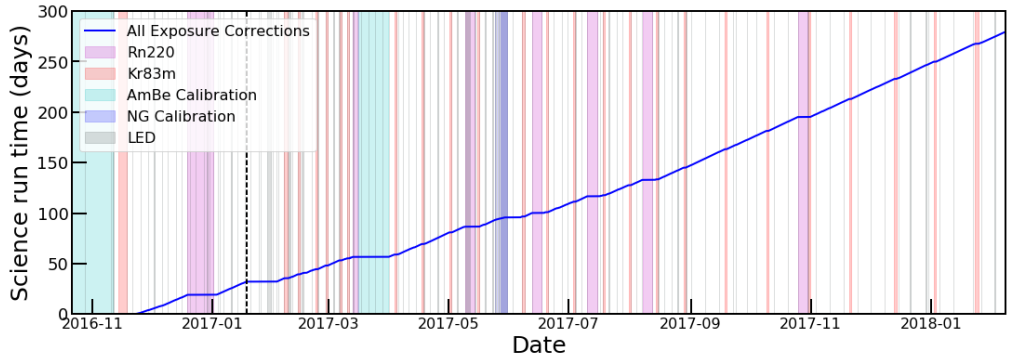


Figure 1.18 – Evolution of the live time accumulated during SR0 and SR1, as a function of the date. The blue line is the accumulated dark matter exposure versus real time, corrected for time periods removed due to DAQ dead-time, muon veto triggers or down-time, photomultiplier malfunctions, or brief periods of increased detector activity, following high energy background events. The final dataset is based on 32.1 (SR0) plus 246.7 (SR1) days of exposure, creating a combined data set of 278.9 days, equivalent to one ton-year using the estimated 1.3 ton fiducial volume. Also shown, with colored bands, are the calibration intervals with the sources presented previously in this Chapter.

WIMP-nucleon spin-independent interaction

The data collected from 278.9 days, corresponding to an exposure of $1.0\,\mathrm{t}\times\mathrm{yr}$, were blinded in the energy region of interest, $[1.4, 10.6]\,\mathrm{keV}_{ee}$ prior to fixing event reconstruction and selection criteria. No significant excess over the background was found and a profile likelihood analysis parameterized in spatial and energy dimensions excluded new parameter space for the WIMP-nucleon spin-independent elastic scattering cross-section for WIMP masses above $6\ \mathrm{GeV}/c^2$, with a minimum of $4.1\times10^{-47}\ \mathrm{cm}^2$ at $30\ \mathrm{GeV}/c^2$ and 90% confidence level.

An imminent detector upgrade, XENONnT, will increase the target mass to 5.9 t. The sensitivity will improve upon this result by more than one order of magnitude.

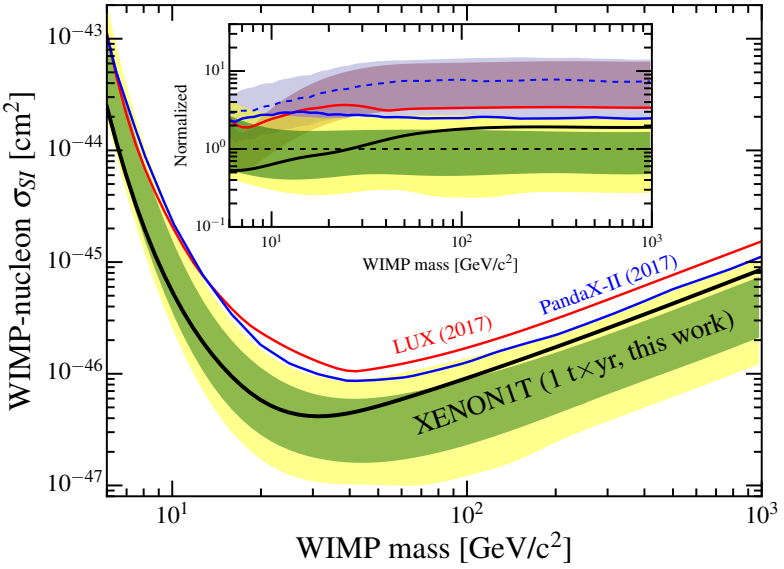


Figure 1.19 – 90% confidence level upper limit of the WIMP-nucleon σ_{SI} with the 1σ (green) and 2σ (yellow) sensitivity bands [94].

WIMP-pion coupling

Other physics channels were explored using the same data, for example scalar couplings of WIMPs to virtual pions exchanged between the nucleons in a nucleus. This interaction is generated when the WIMP couples to a virtual pion exchanged between the nucleons. In contrast to most nonrelativistic operators, these pion-exchange currents can be coherently enhanced by the total number of nucleons and therefore may dominate in scenarios where spin-independent WIMP-nucleon interactions are suppressed. Also this process is the dominant of SD interactions. Again no significant excess was observed over the background and upper limits were set for the wimp-pion coupling.

WIMP-nucleon spin-dependent interaction

Moving down to the hierarchy of the expected nuclear response to WIMPs we have also studied the spin dependent interaction, the axial vector part of which couples with the nuclear spin. For zero momentum transfer the structure factor in the calculation of the differential cross section explicitly depends on the expected values of the nuclear spin operator:

$$S_A(0) = \frac{(2J+1)(J+1)}{4\pi J}|(a_0+a_1')\langle \mathbf{S}_p\rangle + (a_0-a_1')\langle \mathbf{S}_n\rangle|^2 \tag{1.31}$$

where J is the initial ground-state angular momentum of the nucleus, a_0 and a_1 the isoscalar and isovector WIMP-nucleon coupling respectively, $\langle \mathbf{S_p}\rangle$ and $\langle \mathbf{S_n}\rangle$ are the

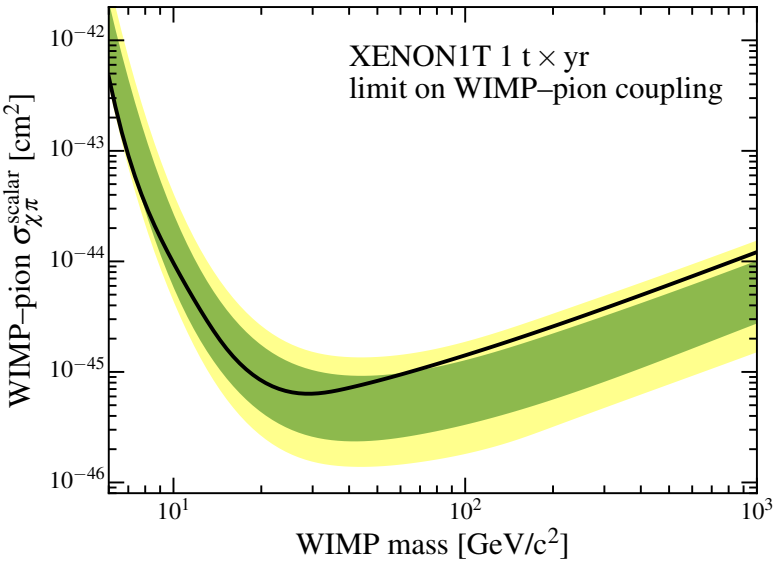


Figure 1.20 – 90% confidence level upper limit of the WIMP-pion coupling as a function of the WIMP mass for the 1 t yr exposure of XENON1T data with the 1σ (green) and 2σ (yellow) sensitivity bands [174].

expectation values of the total proton and neutron spin operators in the nucleus, and a_1' contains a correction to the isovector coupling a_1 due to chiral two-body currents involving the exchange of a pion. Here we study the two special cases $a_0 = a_1 = 1$ ("proton only") and $a_0 = -a_1 = 1$ ("neutron only"). Xenon has two naturally occurring isotopes with nonzero nuclear spin, ^{129}Xe (spin 1/2) and ^{131}Xe (spin 3/2), and as they have an odd number of neutrons we are more sensitive to the neutron only analysis but also for the proton only case a limit can be set, because the total proton spin operator have a non zero expectation value. This limit is, indeed, more than an order of magnitude less restrictive with respect to the neutron only case [175].

Two-neutrino double electron capture

Another physics channel analysed concerns the double electron capture (DEC) of ^{124}Xe, an extremely rare SM process, with an expected -life of the order of $\sim 10^{22}$ yr. In the two-neutrino case (2νECEC), two protons in the nucleus simultaneously convert into neutrons by the absorption of two electrons, mostly from the K atomic shell, with the emission of two electron neutrinos (ν_e). The nuclear binding energy Q released in the process ($\mathcal{O}$(MeV)) is mostly carried away by the two neutrinos but the filling of the vacancies results in a detectable cascade of X-rays and Auger electrons which should produce an observable signal in the detector expected around the energy of $E_0 = 64.3$ keV. The data collected was blinded around this energy region of interest and unblinded after data quality criteria, fiducial volume, and background model had been fixed. In this

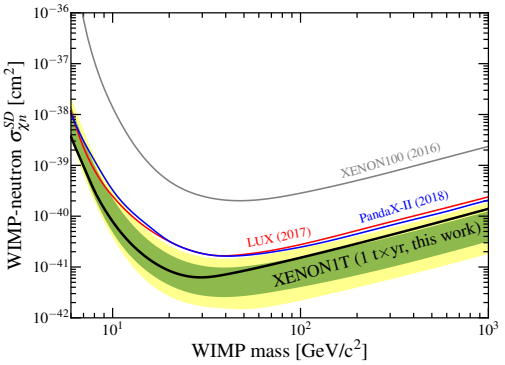


Figure 1.21 – XENON1T 90% C.L. upper limit on the spin-dependent WIMP-neutron cross section from a 1 ton year exposure.

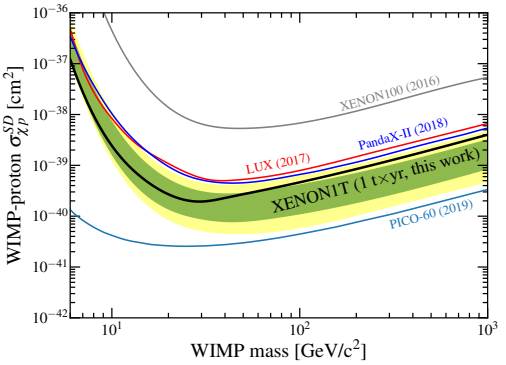


Figure 1.22 – XENON1T 90% C.L. upper limit on the spin-dependent WIMP-proton cross section from a 1 ton year exposure.

energy range the dominant background is ^{125}I coming from neutron activation of ^{124}Xe that occurs after neutron calibrations or interactions with environmental thermal neutrons. The neutron activated Xe decays to ^{125}I via electron capture and ^{125}I decays to an excited state of ^{125}Te that subsequently de-excite producing a mono-energetic peak at 67.3 keV, *i.e.* close to E_0 with the energy resolution of the detector at E_0 being 2.6 keV. The ^{125}I is removed from the detector with a time constant of 9.1 days due to the continuous purification of the detector's xenon inventory by circulation over hot zirconium getters.

After unblinding of the signal region, a clear peak at E_0 was identified. The energy and signal width obtained from the spectral fit to the unblinded data are $E_0 = (64.2 \pm 0.5)$ keV and $\sigma = (2.6 \pm 0.3)$ keV. Converting the fit to a total event count yields $N_{^{125}\mathrm{I}} = (9 \pm 7)$ events from the decay of ^{125}I and $N_{2\nu\mathrm{ECEC}} = (126 \pm 29)$ events from 2νECEC. Compared to the null hypothesis, the $\sqrt{\Delta\chi^2}$ of the best fit is 4.4 σ. The corresponding half-life for the K-shell double electron capture of ^{124}Xe is $T_{2\nu\mathrm{ECEC}} = (1.8 \pm 0.5\ \mathrm{stat} \pm 0.1\ \mathrm{sys}) \times 10^{22}$ y. This is the longest half-life ever measured directly.

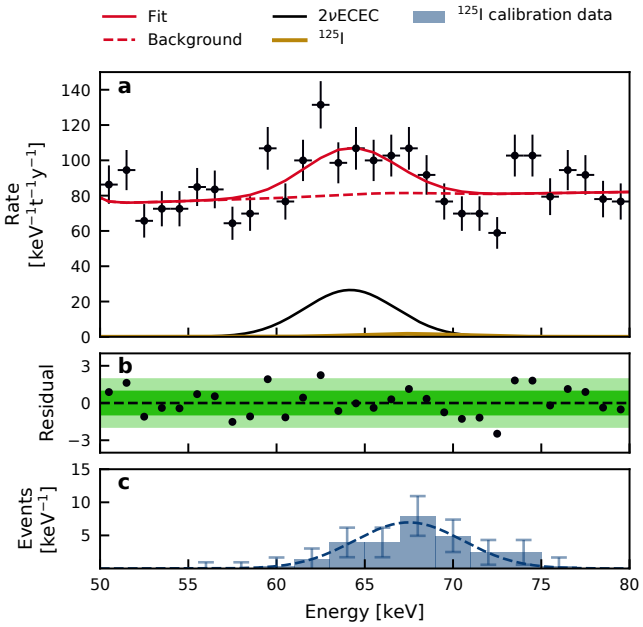


Figure 1.23 – The energy region of interest for the 2νECEC of ^{124}Xe with the best fit contribution from 2νECEC given by the solid black line and the full given by the solid red line. The background-only model without 2νECEC is given by the red dashed line and is clearly incompatible with the data [176].

Chapter 2

The Single Electron background of the XENON1T detector

2.1 Introduction

The events as defined in Chapter 1 are triggered by an S1 or an S2 signal, corresponding to a production of scintillation and ionization. The S2 signal that can trigger the recording of an event corresponds, as we saw, to the proportional scintillation of many electrons that are extracted in the GXe. However, observing the waveform of triggered events we also see that the detector is sensitive to S2 peaks of much lower intensity. This can be seen, for example, from the presence of small peaks that follows (or even precede) the main S2 signal in Fig. (1.16). The total area of these peaks is in the region below 150 PE and therefore cannot be attributed to a deposition of energy by a particle interacting in the active target. The width of these S2s is consistent with just one, or few, electrons being extracted and producing secondary scintillation. These **single electrons** (SE) are of particular interest because their presence dominates the small S2 spectrum of the detector.

Such small charge signals have also been observed in the TPC of ZEPLIN [177, 178] and were the subject of a dedicated study in the XENON100 experiment [179, 106]. There are various mechanisms to which the creation of SE inside the TPC can be attributed. A well known mechanism consists in the photoionization of metal surfaces and impurities from the light of a main S2. Other mechanisms that have been proposed incude effects such as electrons trapped in the liquid-gas interface and their subsequent delayed extraction, a possible existence of long-lived excited states of Xe or of its impurities and the Malter effect *i.e.*, the variation with time of the cathode's local working function resulting in electron emission [180].

In recent years, there has been a growing interest in the study of these mechanisms as well as in finding ways to minimize the presence of SEs, exactly because the sensitivity of the detector, in such small charge signals, can be used for the search of a light and

leptonically interacting dark matter. Indeed in the case of a dark matter interacting with electrons, as suggested by theories beyond the WIMP paradigm predicting a light dark matter particle with mass in the sub-GeV scale, such a particle could scatter off an electron of the Xe atom resulting in its ionization and the eventual production of one or few single electrons that could be extracted and producing proportional scintillation. In Chapter 4 we will calculate the rate of events expected from such a theory and use the small S2 spectrum in order to set limits in the DM-e^- scattering cross section. In such a search, the single electron production mechanisms listed above constitute a background which today is not even fully understood. The consequence is that this search is made without the possibility of a background subtraction and therefore without the possibility to claim an excess over a know background, as can be done with NR of ER. Thus, the study of single electrons is twofold. First it aims in a characterization of as many SE sources as possible and, at the same time, aims to find ways to limit these backgrounds in order to search in the low energy region of the S2 spectrum for a leptonically interacting dark matter.

2.2 Use of Single electrons as a calibration source

Among the various sources of the single electrons background there is one that can be used to study the detector's response, to model its low-energy spectrum, but also to be used as a natural source for calibrating and finding key features related to the physics processes, taking place inside the detector. This category concerns single electrons induced from the photoionization of impurities inside the liquid phase. In particular, the space inside the active volume of the TPC, is dominated by ionization electrons which are drifted towards the liquid/gas interface, following the drift electric field. The presence of oxygen molecules, in very small percentages inside the LXe, results in the capture of some of the drifting electrons by these molecules and the formation of bound anionic states of oxygen, with a small binding energy of about 0.45 eV. These anions can be easily dissociate via interaction with the VUV photons of the S2 scintillation light resulting in the release of the extra electron, Fig. (2.1). This population presents a number of distinct features which we will study below. We can see the presence of this population if we focus on the very low-intensity S2 peaks, found in the waveform of each triggered event of any run, with or without a calibration source. Because, since this background is related to the response of parts of the detector to the two main signals, it is always present. Thus it can be detected in large populations even in dark matter search datasets. In fact, it is exactly this background that constitutes an ultimate limit in the search for dark matter of small masses, as we will see in the last chapter of this work.

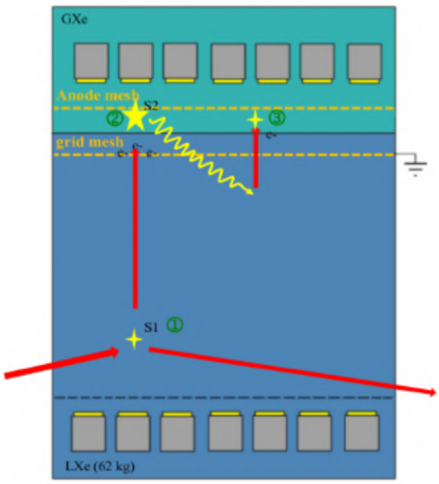


Figure 2.1 – Schematic representation of the main mechanism of production of the single electron background.

2.3 Single electrons from photoionization

In this section we will study the population of small S2s that follows a main S2. Observing the waveform of an event, Fig. (2.2), we see the characteristics of this population by looking at the small peaks recorded immediately after the detection of a main S2, in the time axis but also in their distribution in the PMTs arrays.

We see that, for a certain period of time, a population of peaks is dominant and that these are distributed homogeneously, at first sight, to the PMT arrays. This time period seems to be around 700 μs and corresponds, approximately, to the time required for electrons produced in the lower part of the TPC to reach the liquid/gas interface, to be extracted and to produce the S2 light.

Fig. (2.3), shows the distribution of S2 light in the region $S2 < 150$ PE, which corresponds, as we will see below, to the simultaneous extraction to the gaseous phase of up to 5 electrons. This spectrum has been created by summing up the S2 signals of all PMTs, from a calibration run with a ^{83m}Kr source. For illustrative purposes we have also applied a set of quality cuts before their justification, that will be done in detail in Section (2.5).

We observe a sharp peak around the region of 30 PE that corresponds, approximately, to the average light produced by an electron, as it moves in the gaseous phase of Xe (GXe) under the the strong electric field between the gate and the anode. On a logarithmic scale, we observe a secondary peak around the double value of photoelectrons ≈ 60 PE. As we will see below, this corresponds to the case where two single electrons coincide in time, in the gas gap, and the light produced is detected by the peak finding algorithm as a single peak with an area (corresponding to the intensity of this scintillation light) equal to the sum of

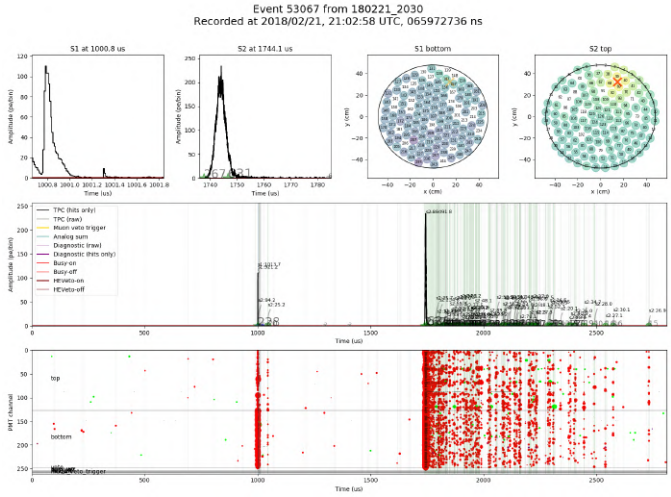


Figure 2.2 – Waveform of a recorded event. A main S1 and main S2 are identified. The main S2 is followed by a population of small peaks, identified as S2 type.

the two signals. The light emitted by electrons, through the proportional scintillation in the gas, is observed both by the top array of PMTs (mainly by the ones closer to the electrons extraction point) and by the bottom array (relatively homogeneously, due to the diffusion of the light into the LXe phase). The S2 signal has been summed over all PMTs channels, the peak finder algorithm is activated to search for a set of simultaneous hits, and the peak has been identified as unique.

2.4 Modelization of the single electrons spectrum

Taking into account the physics of the production of the secondary scintillation light, we can see that the low-energy spectrum can only correspond to light produced by a single electron, in its acceleration in the gaseous phase, while the tail following this S2 distribution, is probably due to accidental coincidences of two or more electrons, according to the previous description. As a first approach, this leads us to formulate a model that could describe the spectrum of Fig. (2.3). The distribution of light produced by a single electron could be described by a Gaussian function centered around a certain mean value, corresponding to the average number of photoelectrons observed per extracted electron, with a specific standard deviation. This is because, through the process of scintillation of the GXe by an electron accelerating into the gas gap, a specific number of photons will be produced, following a Poisson distribution. These photons will be detected by the PMTs, if they can overcome the electronic noise threshold, and then they will be found by the peak-finding algorithm. Let's assume that, on average λ photons are produced

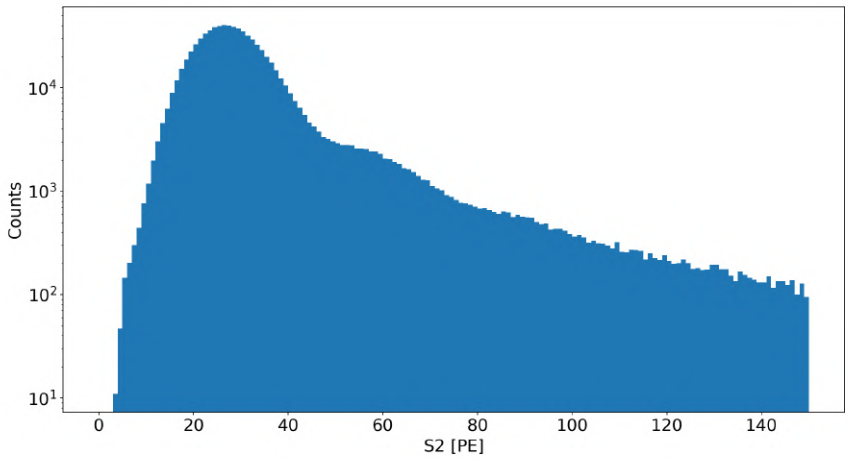


Figure 2.3 – Raw spectrum of the area of small S2 peaks detected during a calibration with a ^{83m}Kr source. A series of data quality cuts have been applied before present them in detail.

per extracted electron. Then, the number n of photons produced per electron could be described by a Poisson distribution:

$$P(n|\lambda) = \frac{\lambda^n}{n!}\exp(-\lambda) \tag{2.1}$$

It will therefore be a distribution with a mean value λ and a standard deviation $\sqrt{\lambda}$, while the pileup of two electrons, simultaneously crossing the gas gap, and the scintillation produced by each of them being described by the relation (2.1), will produce n photons which will be distributed according to the distribution $P(n, 2\lambda)$. With the same reasoning we can generalize the model for all orders of pileups. So, a first thought to describe the spectrum of Fig. (2.3) would be a sum of Poisson's distributions that would describe the S2 distribution of single electrons and their pileups as a sum of Poisson distributions with an average value of integer multiples of λ and standard deviations of $\sqrt{i}$-multiples of $\sqrt{\lambda}$, with $i = 1, 2, \ldots, 5$, to include up to 5 extracted electrons.

Of course, in Fig. (2.3), what we see is not the photons created by the electrons, but, from those photons that were finally observed by the PMTs, we see the signal produced at the anode of the PMTs, normalized to photoelectrons *i.e.*, the average induced voltage on the output of the PMT by the cascade of electrons after an incident photon. This stochastic process, which depends on the gain of the PMTs, is introduced on top on the physical process (2.1) and is essentially what we see in Fig. (2.3). For relatively large values of λ, the conversion of the number of photons into photoelectrons can be perceived as a continuity correction and, thus, we can approach the distribution of photoelectrons

as a normal distribution, with an average value of μ, and standard deviation σ (with σ quite close to $\sqrt{(\mu)}$). As for the pileups of single electrons, they can be described with the sum of multiple normal distributions, with mean value $i\mu$ and standard deviation $\sqrt{i}\sigma$ for $i = 1, 2 \cdots N_p$, to include up to N_p pileups. The model parameters will be determined by a fit on the data.

If we try to fit this model to the spectrum of Fig. (2.3), we will find that it does not describe well the data. This is mainly because, it is clear that, for very small values of S2, the data does not appear to be well described by a normal distribution. It seems that there is a threshold effect, which was already known to the XENON100 detector, due to a combination of effects. One reason may be that, for low-intensity S2s, there is a suppression of the signals due to a drop in the efficiency with which the peak-finder algorithm can detect small S2 signals. Another reason is the fact that there is a sagging and a residual tilt of the TPC, an effect which we will consider in the next section, which implies an dependence of the extraction electric field on the x-y position. This fact introduces in (2.1), a dependence of the rate λ with the horizontal position of the extracted electron. The fact that the spectrum of Fig. (2.3) results from the sum of the S2 light from electrons extracted from many different horizontal positions, not necessarily homogeneously distributed, definitely influences the lower part of the distribution of photoelectrons coming from single electrons. In this section, we will try to take this effect into account by introducing, in the model we developed above, a continuous threshold function, such as a Fermi-Dirac function, which will play the role of a kind of detection efficiency connecting the area of low S2 with the normal distribution for larger S2. In short, our model for single electrons and their N_p pileups, can be written as:

$$f(S_2) = \frac{1}{\exp(-\frac{S_2 - s_{thr}}{\Delta s}) + 1} \sum_{i=1}^{N_p} A_k \exp\Big(-\frac{(S_2 - i\mu)^2}{2i\sigma^2}\Big) \tag{2.2}$$

In this model we have a total of $4 + N_p$ free parameters which will be determined by the fit on the data. Specifically we have N_p amplitudes A_i that represent the frequency of each population, the parameters μ and σ that represent the secondary scintillation gain and the standard deviation in units of PE per extracted electron, as well as two additional parameters s_{thr} and Δ_s, with which we will try to describe the apparent drop in detection efficiency for small S2 signals and the consequent deviation from the normal distribution. In this section we will study the range of energies [0,150] PE, corresponding to up to about 5 extracted electrons, so here we should have $N_p = 5$ and therefore 9 free parameters. However we will make use of two supplementary peaks as the lower tail of Gaussian describing pileups of 6 of 7 electrons can enter the region [0,150] PE, thus their inclusion can contribute to the improvement of the fit near the upper bound of 150 PE. So here we will rather have $N_p = 7$ and 11 free parameters.

2.4.1 The Fit Method

To fit the model (2.2) to the data of the spectrum of Fig. (2.3) we will use a Bayesian Markov Chain Monte Carlo (MCMC) method [181, 182]. Specifically, we will form the likelihood function of the model.

This function refers to the likelihood of finding $(d_1, d_2, \cdots, d_{N_{bins}})$ counts in each, not empty, bin of the histogram, given the parameters of the model $(\mathbf{x}) = (\mu, \sigma, s_{thr}, \Delta_s, h_1, h_2, \ldots, h_7)$.

$$\mathcal{L}(\mathbf{x}|S2) = \prod_{i=1}^{N_{\text{bins}}} \exp\left(-\frac{1}{2}\Big(\frac{d_i - f(S2_i|\mathbf{x})}{\hat{\sigma}_i}\Big)^2\right) = \exp\left(-\frac{1}{2}\sum_{i=1}^{N_{\text{bins}}}\Big(\frac{d_i - f(S2_i|\mathbf{x})}{\hat{\sigma}_i}\Big)^2\right) \tag{2.3}$$

Where d_i is the content of each bin and $\hat{\sigma}_i$ the statistical error associated which here we will consider equal to $\sqrt{d_i}$ for $i = 1, 2, \cdots, N_{bins}$.

Then, given the data and this likelihood, the probability $p(\mathbf{x})$ is:

$$p(\mathbf{x}) = \frac{1}{Z}\mathcal{L}(\mathbf{x}|S2)\text{Prior}(\mathbf{x}) \tag{2.4}$$

where $\text{Prior}(\mathbf{x})$ is a prior estimation on the values of the parameters $\mathbf{x}$ and Z is a normalization constant which is independent of the parameters.

A useful method to identify the optimal parameters and their statistical error as well as the corresponding covariance matrix, is to generate a sample of values $\mathbf{x}_t$ from the posterior distribution $p(\mathbf{x})$. The goal of the MCMC algorithms is to make a sampling of values $\mathbf{x}_t$ from the relation (2.4). This is done with a procedure capable of generating a Markov chain $\mathbf{X}(t) = \mathbf{x}_t$ in the multidimensional space of the parameters $\mathbf{x}$ (lets call it $\mathcal{P}$), that after a certain time of random walk in $\mathcal{P}$ collects a representative set from the distribution. An algorithm that encodes this process is the Metropolis-Hastings Algorithm [183], which is an iterative procedure: Let a point $X(t) \in \mathcal{P}$. A certain transition distribution Q (often, a multivariate Gaussian distribution centered on $X(t)$) randomly selects a new point $Y \in \mathcal{P}$, according to the PDF $Q(Y|X(t))$. This new point is accepted in the sample, (*i.e.*, $X(t+1) = Y$) with a probability:

$$\min\Big(1, \frac{p(Y|S2)}{p(X(t)|S2)}\frac{Q(X(t)|Y)}{Q(Y|X(t))}\Big)$$

Otherwise, $X(t+1) = X(t)$. It can be proved that, as $t \to \infty$, this algorithm converges to a stationary set of samples from the posterior distribution $p(\mathbf{x}|S2)$. Then, the expectation values of the parameters can be approximated with the marginalization of the sampled histogram $\mathbf{X}_t$ *i.e.*, a function of the model parameter x_t has an expectation value:

$$\langle f(x)\rangle = \int p(x|S2)f(x)dx \approx \frac{1}{N}\sum_{t=1}^{N} f(x_t)$$

In this analysis we will select, as the starting point $X(t) \in \mathcal{P}$, the maximum likelihood estimation of the model (2.2). Then making use in (2.4) of a flat uninformative prior, regarding the prior PDF of the parameters, we will start a set of 500 walkers who will start to perform a random walk in the space $\mathcal{P}$. When they reach a state of equilibrium we will select these stationary sets, that correspond to samples of the posterior probability distribution and we will determine the errors around the mean value of each parameter, as well as the covariance matrix, by making all possible 2D projections of the sampled posterior probability distributions.

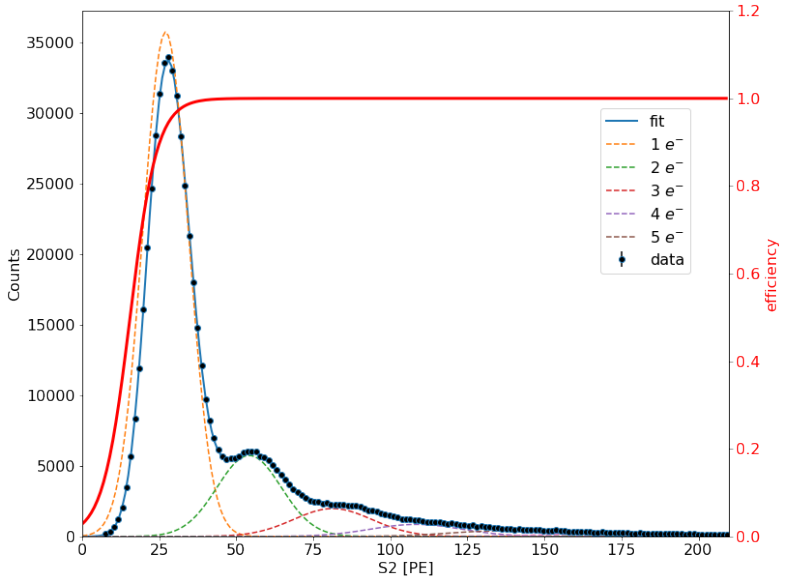


Figure 2.4 – Fit (statistical errors only) of the single electron model to the data obtained from runs where a calibration source has been used. The various components of the model are shown separately.

In Fig. (2.4) we present the result of the fit in blue. Also shown are the contributions of each normal distribution that describes the single electron pileups as well as the detection efficiency as determined by the Fermi Dirac function. Due to the very large number of single electrons, the error bars on this linear plot plot are indistinguishable, while also the error of the parameters is very small. In Fig. (2.5), we can see the stationary sets sampled from the Markov chains, where it seems that, after a very short expansion period, they have stabilized in a specific region of the parametric space, around the maximum likelihood point.

In Fig. (2.6) we can see the covariance matrix of parameters. We observe strong correlations concerning mainly the parameters μ and σ as well as their relation to the parameters of the threshold function.

We see for example an anti-correlation between the parameters μ and σ so that in case of an upward fluctuation of the value of the secondary scintillation gain this is

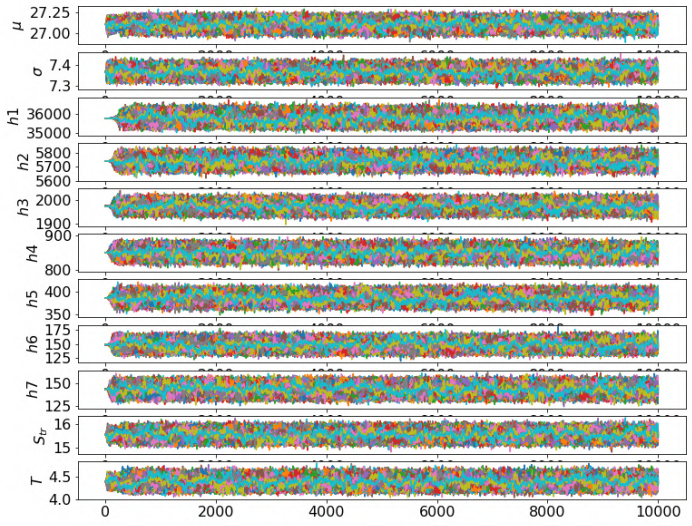


Figure 2.5 – Stationary sets sampled from the Markov chains with 10^4 steps. As we can see, starting from an initial value of the parameters corresponding to the maximum likelihood of the model, very soon, an equilibrium state has been reached. The errors on the parameters can be determined from these sampled distributions.

mitigated on the fit by a decrease of the standard deviation σ. We also observe a strong anti-correlation between the value of the secondary scintillation gain with the parameters of threshold function as, for lower values of μ, it seems that the deviation from the normal distribution is larger for the first peak, therefore it is necessary to increase the parameter s_{thr}, as well as step Δ_s. Ideally if this threshold function was a Heaviside step function θ, which would be 0, for $S2 = 0$, and 1, for $S2 > 0$, or even displaced by an amount of S2, below which the detection efficiency would be 0 and, above which, would be 1, the two parameters would increase or decrease together and inversely to the parameter μ.

Because of these corrections and also because of the data-driven approach, on determining the detection efficiency we will not use the model (2.2). We will rather use an alternative model that we present in detail in the next two sections.

2.4.2 Simulation of the peakfinder efficiency

As we have seen, a major problem with the model (2.2) is the data-driven description of the detection efficiency, via the Fermi-Dirac function. This option was well motivated for the XENON100 detector but the XENON1T clustering and hit-finding algorithm is much more efficient. In fact, using the waveform simulator developed by the Collaboration, we

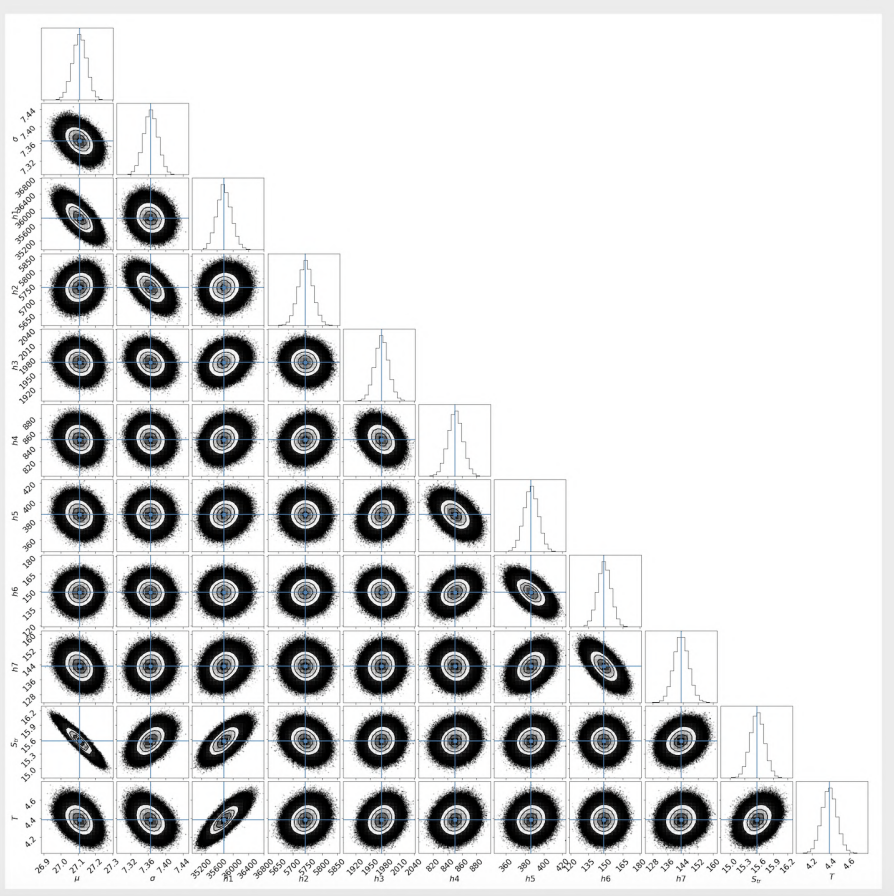


Figure 2.6 – Covariance matrix of the parameters of model (2.2) obtained from two dimensional projections of the posterior probability distribution (2.4). In the "main diagonal" of the matrix are also shown the one dimensional projections of the posterior probability distribution together with the initial values of the fit which correspond to the maximum likelihood estimation of the model (2.2).

did a systematic simulation of the peak-finding process and determined the peak-finding efficiency as a function of the S2 area.

Specifically with the waveform simulator we created 30k events with one single electron each, in random positions inside the TPC. Then the waveforms are processed with PAX and the result is compared with the "truth" file of the simulator. Specifically, each peak detected by PAX in the period between the time of emission of the first and the last scintillation photons, Δt, produced by the single electron, is examined and compared to the "truth" information of the simulation. The following considerations emerge in such a comparison:

- **found**: A unique peak of type S2 was found inside Δt.

- **chopped**: The peak is split into many fragments but exactly one is of type S2.
- **split and misidentified**: The peak is split and at least one fragment is of type S1.
- **split and unclassified**: The peak is split and all fragments are unclassified.
- **missed**: There is no peak found by PAX inside Δt.
- **unclassified**: the peak is found but is not classified.
- **misidentified as S1**: The peak is found but it was identified as S1.

In Fig. (2.7), we can see the fraction of peaks corresponding to these cases, as a function of the S2 area.

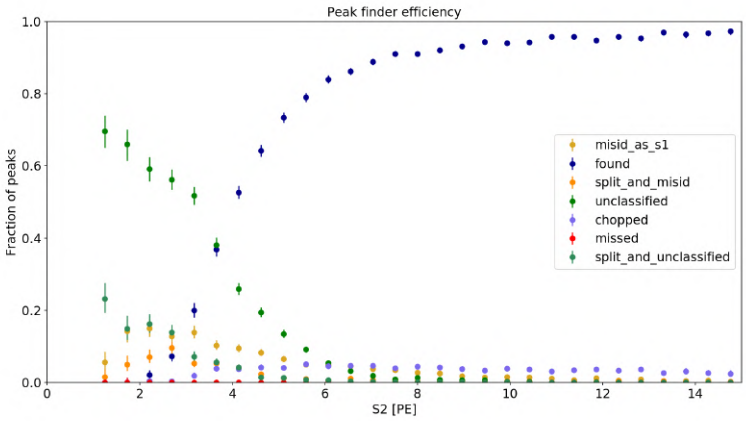


Figure 2.7 – The fraction of peaks generated by the simulator, corresponding to the cases listed in the text. We note that nearly all peaks are detected and the drop of efficiency happens below 10 PE mostly due to wrong classification as S1s.

We note that the main reason for the drop in efficiency in the region below 10 PE is that PAX fails to classify these peaks as of S2-type. On the other hand, the probability of a peak to be lost is negligible. We also notice that, for S2 $>$ 10 PE, more than 96% of the peaks have been found as unique peaks inside the period Δt. But there is still a percentage of peaks, found as the only fragment, of S2-type, coming from a larger peak that it was split by PAX. Because, in this case, the peak of S2-type found was unique, (while all other fragments in the larger peak where unclassified) we can consider that practically the true peak has been found by PAX and add this percentage to the percentage of peaks found as unique. Thus, we can define the detection efficiency of single electrons as the sum of "found" and “chopped” peaks of Fig. (2.7).

In Fig. (2.8) we see that this efficiency is practically 100% for S2s greater than 10 PE, in contrast to the data driven efficiency curve of Fig. (2.4). The two additional

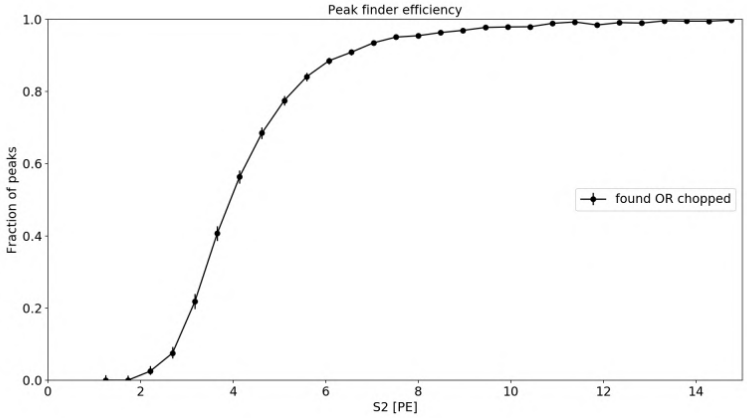


Figure 2.8 – The sum of "found" and "chopped" fractions of the simulated peaks of Fig. (2.7) that constitutes actually the efficiency of detecting an S2.

free parameters S_{tr} and Δ_s of model (2.2) do not seem to really describe the peak-finding efficiency but rather the deviation of the SE area distribution from the Gaussian distribution due to e.g. the dependence of the secondary scintillation gain on the x-y position, or on other factors that may slightly affect it even within a single run, such as small changes in liquid level or in the GXe pressure at the gas gap.

2.4.3 Alternative single electron spectrum model

Taking now into account the knowledge gained by the previous simulation of the peakfinder efficiency by the data processor we see that we have to look elsewhere for the description of the deviation from the Gaussian behaviour in the region of the first peak of the single electron spectrum. A more effective model would be to try a smearing of the underlying Poisson distribution (2.1), that describes the actual physical process of scintillation, with a Gaussian distribution that would describe the smearing around an average value μ, with a standard deviation σ. For example, for n photons this would be:

$$f(n|\mu,\sigma) = \int \text{Poisson}(n|\lambda)\text{Normal}(\lambda|\mu,\sigma)d\lambda \tag{2.5}$$

Here, the parameter μ will be the estimate of the secondary scintillation gain, as it will be centered around the average value of the number of produced photons, while σ will be an estimate of the standard deviation associated with μ as in the model (2.2) but, this time, it will also incorporate the dependency on the position and the other factors that may affect the secondary scintillation gain. If we want to expand the model so that

we can also include the N-order time pileups of single electrons, we can sum up again as in (2.2):

$$f(n|\mu,\sigma) = \sum_{k=1}^{N} \int \text{Poisson}(n|\lambda)\text{Normal}(\lambda|k\mu,\sqrt{k}\sigma)d\lambda \tag{2.6}$$

In addition, we can directly include in this model the area of the S2 peaks, if we consider again that the transition from the number of hits to the area, through the gain of PMTs, is a continuity correction in the Poisson distribution of the relation (2.6). In this case, we can replace in the relation (2.6), the Poisson distribution with the so-called continuous Poisson distribution [184] and write the distribution of the S2 area as:

$$f(x|\mu,\sigma) = \sum_{k=1}^{N} \int \frac{\lambda^x \exp(-x)}{\Gamma(x+1)} \textbf{Normal}(\lambda|k\mu,\sqrt{k}\sigma)d\lambda \tag{2.7}$$

In Fig. (2.9) we present a fit of the SE spectrum based on model (2.2) and on model (2.7). We see that, with two parameters less, the fit is just as good so this model is more robust. More importantly, in Table (2.1), we observe that the best fit value of μ is systematically higher for the model (2.7) than for the model (2.2).

The correlation of the s_{thr} parameter of the model (2.2) with μ led to an underestimation of μ. That is to say, the Fermi Dirac curve, while essentially not describing the peak-finder efficiency, was driven by the data to correct the non-Gaussian form mainly of the first peak, as a result of which it introduced a bias in the determination of μ. For this reason, from now on, we will use this model, both to determine the secondary

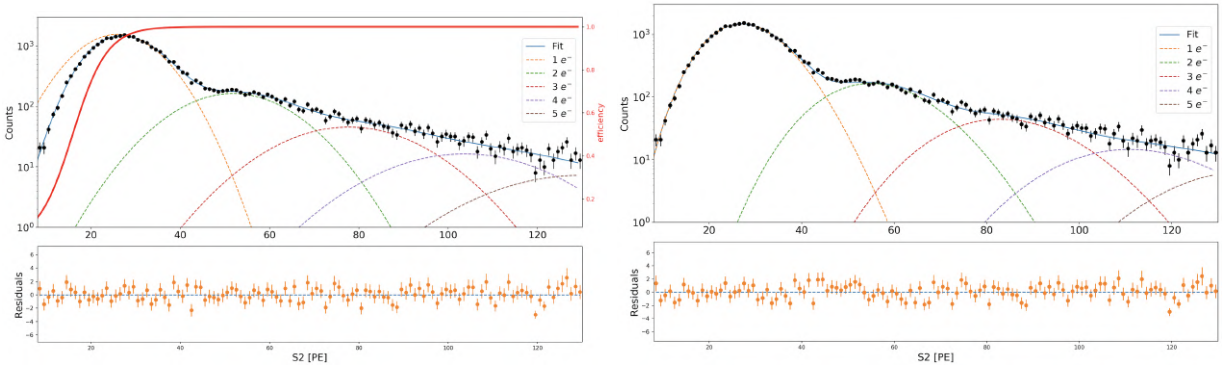


Figure 2.9 – Comparison of the fit to the single electron spectrum, between the multi-Gaussian model (2.2) and the alternative model using a sum of Poisson distributions convoluted by a Gaussian (2.7). We can observe from the distribution of residuals that a model with two free parameters less can describe the data equally well (see also Tab. 2.1), in adittion of being motivated from a physical viewpoint (see text).

Model comparison		
Parameter	multi-Gaussian	multi-Poisson value
μ $[PE/e^-]$	25.9 ± 0.28	27.96 ± 0.06
σ $[PE/e^-]$	7.83 ± 0.1	4.81 ± 0.06
h_1 [Counts]	1559 ± 41	2182 ± 30
h_2 [Counts]	164 ± 4	244 ± 6
h_3 [Counts]	46 ± 2	66 ± 3
h_4 [Counts]	16 ± 2	22 ± 3
h_5 [Counts]	7 ± 2	11 ± 4
s_{thr} $[PE/e^-]$	15.9 ± 0.7	-
Δ_s $[PE/e^-]$	3.82 ± 0.25	-
χ^2/n.d.f.	1.18	1.28

Table 2.1 – Comparison of best fit parameters for the multi-Gaussian and multi-Poisson models. The quality of the fits is also shown, as quantified by the ratio χ^2/n.d.f..

scintillation gain and, later on, for the leptophilic dark matter analysis using the single electron background data.

2.5 Data selections and quality cuts

Here we will present the cuts we applied to the raw spectrum of the small S2s (2.10) we collected from the events of several runs, in order to obtain the spectrum of Fig. (2.3) which allows us to fit the physical model (2.7) and extract with an unbiased way the secondary scintillation gain.

1. From all identified S2 signals of an event we select those, whose intensity is less than 150 PE that correspond, as we have already commented, to about 5 extracted electrons.

2. In addition, in each event, we select those small S2 signals that follow the main S2 of the event, that is, we select exactly those small S2s that are caused via photoionisation of impurities and that constitute a specific and large population that offers abundant statistics and characteristic features which we can easily study as we will see in the next two sections, and which we can use for the next set of cuts that we will develop in this section.

3. From all the events, we choose those that have a main S2 area of at least 10k PE. This is because, as we will see in the next section, there is a linear relation between the light intensity of the main S2 and the rate of single electron that follow

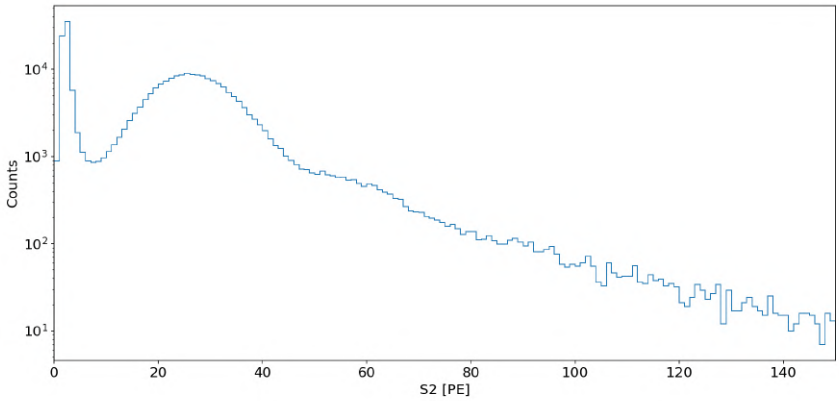


Figure 2.10 – Raw spectrum of raw S2, obtained by the collection of all small S2 peaks of every event in a run

immediately after. Therefore, by selecting a lower threshold for the intensity of the main S2 we increase the rate of single electrons and, a fortiori, the rate of their multiple pileups, which consequently leads to an increase in the amplitude of the secondary peaks of the spectrum (2.3) and therefore to a more robust fit of the model (2.7)

4. In (2.10) we can observe a sharp increase in the number of S2s for very small signals, less than 10 PE. This is mainly due to misidentified S1 or PMTs afterpulses and usually the corresponding light is observed by only one or two PMTs. We can get rid of part of this population by setting a threshold on the minimum number of top PMTs that contributes to the observed S2. In this analysis we required this threshold to be 4 PMTs.

5. Taking advantage of the large number of single electrons that follow a main S2 we can perform a study of the secondary scintillation gain as this is determined by the fit of the spectrum of the Fig. (2.3) as a function of the time distance of single electrons from the main S2. This time interval corresponds to the depth z on which the single electrons are produced. For this purpose we can divide the time after a main S2 into bins of 10 μs so that there is enough single electrons to perform the fit of the model (2.7). In Fig. (2.11) we present the evolution of the parameter μ as a function of the time distance from the main S2. We observe an increase in the parameter μ with the time, which reach a plateau at about 70 μs after the main S2. This is an effect that has been observed also in the XENON100 detector and is related to the fact that after an event with a very intense S2 light, the base current of the PMTs that has observed it, needs some time of the order of μs to return at true zero level, an effect called baseline shift [185]. So, this discharge time

of the PMT circuit cannot be ignored, for the small S2 signals found during this time period, which due to the fact that the baseline is not at its true zero level they will appear biased, resulting in an equally biased estimation of the secondary scintillation gain, as shown in Fig. (2.11). This leads us to the conclusion that, in order to avoid this bias, we must choose only those small S2s that are at least 70 μs away from the main S2.

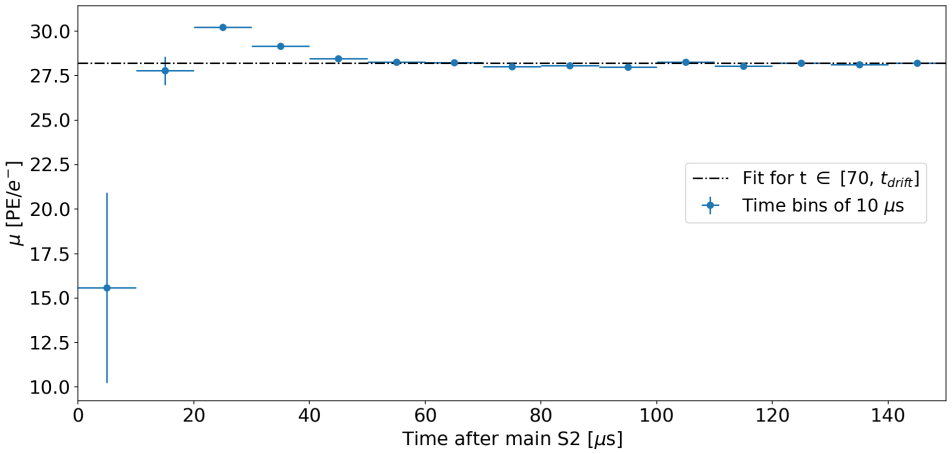


Figure 2.11 – Effect of the proximity to the main S2 peak, in the determination of the parameter μ

6. The requirement of cut 4 is not enough to remove the noisy small S2s at the lower part of the spectrum (2.3). This is because sparks from the electrodes of the cathode can also create such noisy small signals that could be misidentified as S2s. We can get rid of them with a cut on the maximum time that corresponds to one drift of an electron from the bottom of the TPC, ie about 720 μs.

7. It is obvious that, if an event in the waveform has more than one, large S2, then it is necessary, for a period of one maximum drift time, after the main S2, that there is no other S2 greater than 150 PE, otherwise all the problems described in point 4 would appear. We therefore in addition, we require that, for each event, the second largest S2 detected in the waveform by the peak finding algorithm, to be less than 150 PE, in other words, the event must have only one S2 greater than 150 PE.

2.6 Relation with the intensity of the S2 light of the main event

An indication of the fact that single electrons observed immediately after the main S2 of an event are due to the photoionisation of impurities can be obtained from an analysis of the number of small S2s that follow the main S2 of an event, over a period of one drift time, (*i.e.*, throughout the liquid volume of the detector which is exposed to the S2 light), as a function of the intensity of the main S2. In Fig. (2.12) we present the corresponding plot of this analysis.

We make use of calibrations runs with ^{83m}Kr. A vessel containing a ^{83}Rb source is connected to the recirculation system, emanating ^{83m}Kr within the TPC. This isotope is in a metastable state that decays to the ground state by emission of conversion electrons subsequently producing a number of ionization electrons that will create S2 signals ranging in the region around 20000 PE (uncorrected). Using many such calibration runs we partition the S2 region in which we have most of the S2s into 7 bin of 2000 PE, and we measure the number of single electrons that follow the main S2 of each event within a time delay corresponding to one maximum drift time. With a linear fit in the data, we see that the relation between the two quantities is clearly linear, with a proportionality factor that is $(4.51 \pm 0.1) \times 10^{-4}$ small S2s per photoelectron of the main S2. Also noteworthy is the fact that the line intersects the y-axis at a point equal to $b = 1.78 \pm 0.11$ small S2s. Obviously this corresponds to a population of SEs that are not related to the photoionisation of impurities by the main S2 but are rather produced by some other mechanism. For example, they could be single electrons created by the photoionisation of impurities (or by photoionization of the electrode surfaces) by the main S1 light of the event. The single electrons that are created between S1 and S2 are a distinct population that will be studied in a next section and also offers an opportunity to infer important parameters of the detector. Also this population may also be contaminated by SEs that may have been created even before the main S1. These are, for example, SEs created by the S2 light of a previous event that is relatively close to the event under study. There may also be SEs that undergo a delayed extraction as (see in section (2.12)), there is a part of SEs that is trapped on the surface between the liquid and the gaseous phase, and can escape from it stochastically, even up to several ms after the S2 that created them. In this case, these can come from many events before the event under study and, of course, can occur at any time in this event. In fact, these events form an irreducible background of the search for leptonically interacting dark matter, as we will see in the last Chapter.

2.7 Rate of Single electrons and their pileups

The time characteristics of the SEs that are observed immediately after a main S2 give us another indication in favour of the hypothesis that these come indeed from the

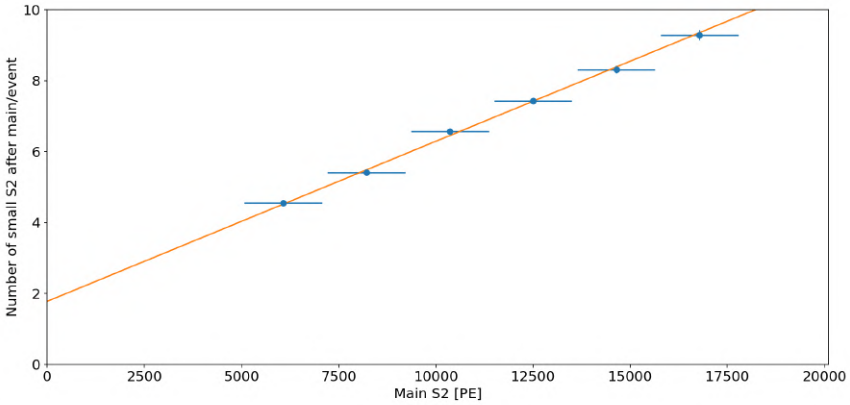


Figure 2.12 – The number of small (below 150 PE) S2s after a main S2 as a function of the intensity of the main S2. A linear relation is revealed. Also the fit shows that the number of small S2s is not zero in the absence of a main S2.

photoionisation of impurities. In this section we will study the rate of SE and their pileups, as a function of their time distance from the main S2 of the event. In Fig. (2.13) we present the rate of the populations of Single electrons as well as the rate of 2 and 3-fold pileups as a function of their time distance from the main S2.

The cuts used are those of section (2.5) and we selected the corresponding populations based solely on the intensity of their S2 signal. We select each population based on those S2s areas that are listed in the figure corresponding to a purity of 95%, in terms of the contamination of each population by the neighbouring ones. We observe that the populations of single electrons and their 2 and 3-fold pileups present a rate that decreases exponentially with time for a period of $720\mu s$ corresponding to the maximum drift time for an electron in the TPC, but with a characteristic time that differs for each population. In fact, from a fit of an exponential function in the three populations (solid line in Fig. (2.13)) $R_i \exp(-t/\tau_i), i = 1,2,3$, we deduce an interesting relationship between the constants τ_i. Specifically we find that:

$$\tau_1/\tau_2 = 2.0 \pm 0.1$$

$$\tau_1/\tau_3 = 3.3 \pm 0.1$$

This result can be explained by the following model. Assuming the rate of single electrons is described by a relation of form $R_1(t) = R_1(0)\exp(-t/\tau_1)$. These single electrons are produced uniformly within the TPC active volume and if they are very close to each other in time, then they will be perceived, as we explained earlier as a single signal. Let's assume that Δt is the time period during which two SEs remain unresolved by the peak-finding algorithm. In this case the number of single electrons

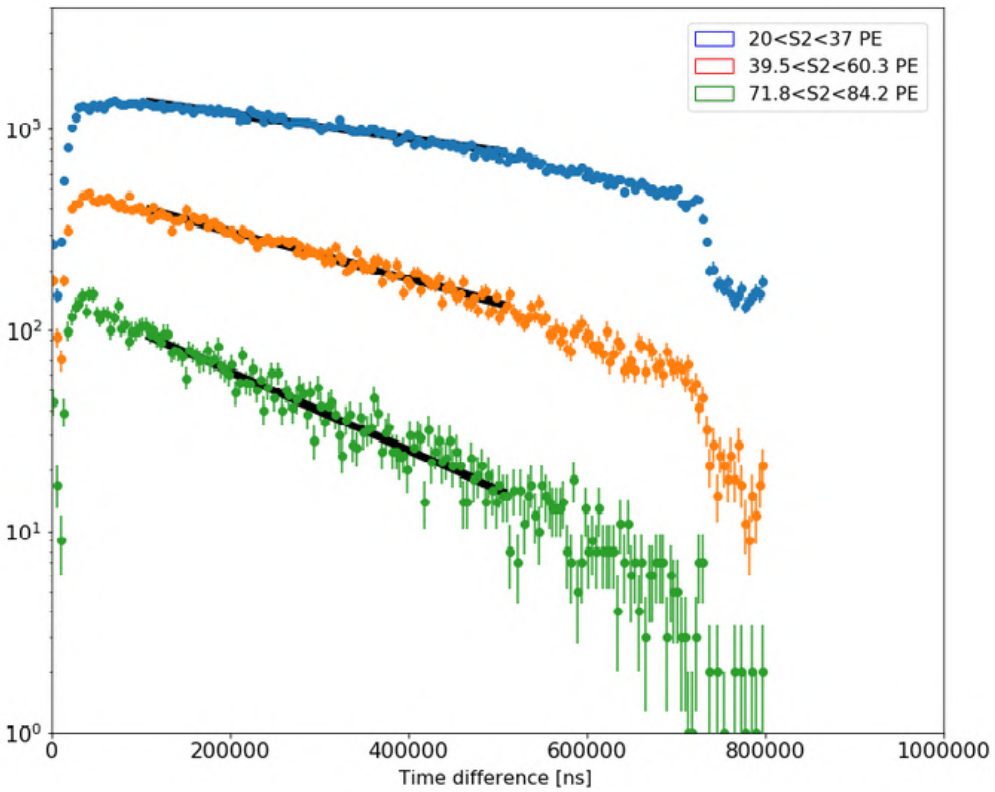


Figure 2.13 – Rate of events from the first peak of the single electron spectrum (blue), from the second peak (orange) and from the third peak (green) as a function of their time difference from the main S2. Also is show an exponential fit ti the rate from which characteristic times τ_1, τ_2, τ_3 are extracted. The relation between them reveals that the events in the second and third peak are indeed temporal pilepus of single electrons.

found during this time is $N(t) = R_1(t)\Delta t$. Thus the rate of the 2-fold pileup will equal $R_2(0) = N(t)R_1(t) = R_1(t)^2\Delta t = R_2(0)\exp(-2t/\tau_1)$. So we see that the 2-fold pileups are characterised by an exponential decrease with characteristic time $\tau_2 = \tau_1/2$. With the same argument, we can easily see that a similar relation should apply to 3-fold pileups.

2.8 Sagging and residual tilt of the anode

Another effect that we can study with the help of single electrons concerns the sagging of the anode's wires. This sagging is due to mechanical and electrostatic effects. On the one hand, even if the mechanical tension is close to the breaking limit, a wire with a length of the order of one meter will be subject to a deformation from the straight line, due to its weight. On the other hand, the extraction electric field between the anode and the gate, will create a force of electrostatic nature on the anode that will tend to deform

it in the same direction as the force due to its weight [186]. In addition, this electrostatic effect is therefore also a function of the applied voltage between the anode and the gate, as we will see in this section. One way to see this deformation of the anode geometry is to study the distribution of the so-called width of the single electron small peaks, that is, the time range in the waveform containing the 50% of the total peak area for each small S2.This width depends directly on the electric field which extracts and accelerates the electrons that produces the corresponding scintillation light. At the same time, the position of the extracted single electrons can be determined from the pattern of the top PMT arrays and thus a study can be made of the width as a function of the position, a relationship that will indirectly reveal the variation of the electric field as a function of the position and therefore the variation of the distance of the anode wires from the phase boundary *i.e.*, the deformation of the anode.

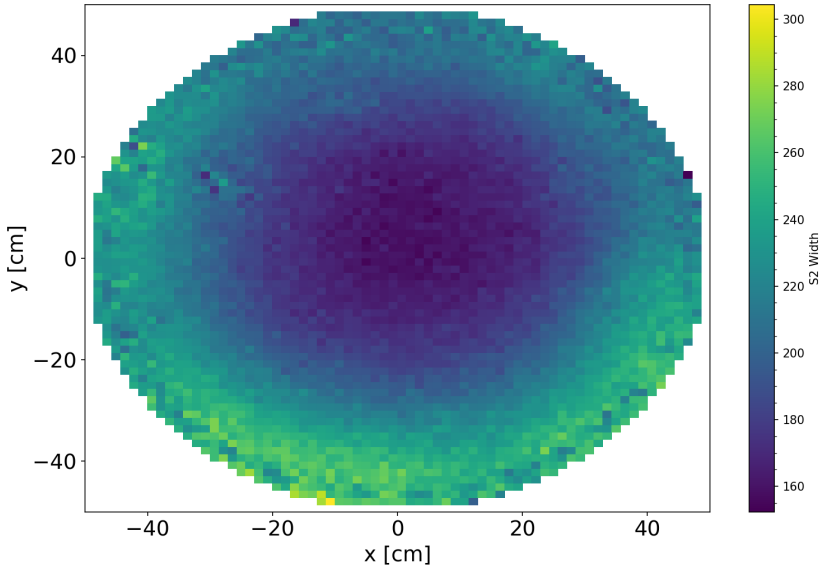


Figure 2.14 – The S2 width of single electrons as a function of their reconstructed position for anode voltage of $V_a = 4$ kV. This 2D histogram is calculated as described in the text.

In Fig. (2.14), we present such a plot for single electrons that are produced in several calibration runs utilising a source of ^{83m}Kr where the potential between the anode and the gate is set to +4 kV. This plot is a projection of a three-dimensional histogram of the position x-y and the width of each peak into a two-dimensional histogram of the x-y position where the colour code corresponds to the average width of the positions contained in each bin, accompanied by the corresponding standard deviation. It should be noted that in this case it was necessary to choose exclusively single electrons without including

their time pileups, as these would introduce a different absolute value of the width (they consist of signals summed over all channels) but also the reconstructed position may not even have a physical significance in the sense that it is possible for two electrons to cross the gas gap simultaneously at two different x-y points of the horizontal projection of the TPC. In this case the position reconstruction algorithm fails because it is based on MC simulations of clusters of electrons produced in the same interaction and so a fortiori derived from the same x-y point. So, in this particular case, the reconstructed position would be a kind of a center of gravity between the clusters of PMTs that would be activated in the neighbourhood of each extracted electron and thus would lack a physical significance for studying the sagging of the anode.

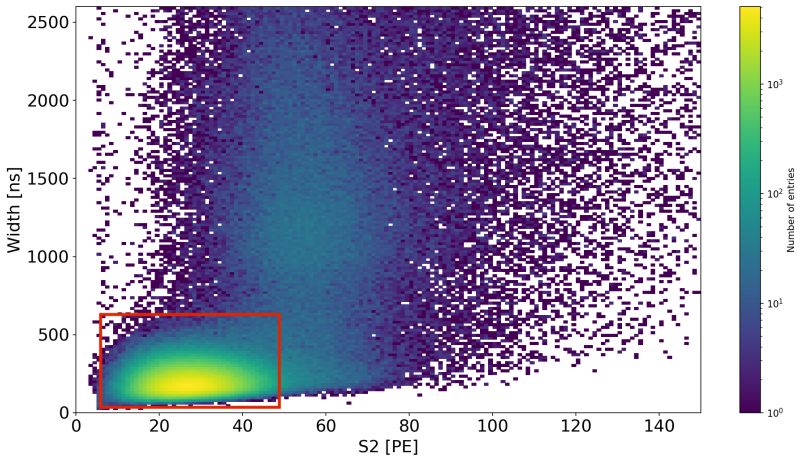


Figure 2.15 – Choice of single electrons without contamination by their pileups, in order to be used for determining the sagging of the anode via the relation between their S2 width with the position of the extraction.

For this, we can select exclusively single electrons using the plot of Fig. (2.15), where we see that single electrons can be distinguished from their pileups due to their smaller width. The plot of Fig. (2.14) was made by selecting the signals inside the square of the Fig. (2.15).

Observing Fig. (2.14) we see that this does not have a radial symmetry, as one would expect from the geometry of the detector. We observe instead that the radial symmetry is broken by what seems to be a residual tilt with respect to the horizontal level of the phase boundary, such that introduces an asymmetry in the S2 width between the 1st and 3rd quadrant of Fig. (2.14), with the former having a larger S2 width (meaning larger gas gap) and the latter having a smaller average width. We will try to fit the x-y distribution of S2 widths using functions that could describe this sagging and tilt. We will make use

of three different possible functional forms for the sagging, shown in Tab. (2.2), which will be added to a function of a plane that will describe the tilt:

$$f(x,y) = p_0 + p_1 x + p_2 y$$

Among the functional forms of Tab. (2.2) we found that the elliptic catenary best describes the data. In Fig. (2.16) we present the result of the fit on the data for this function.

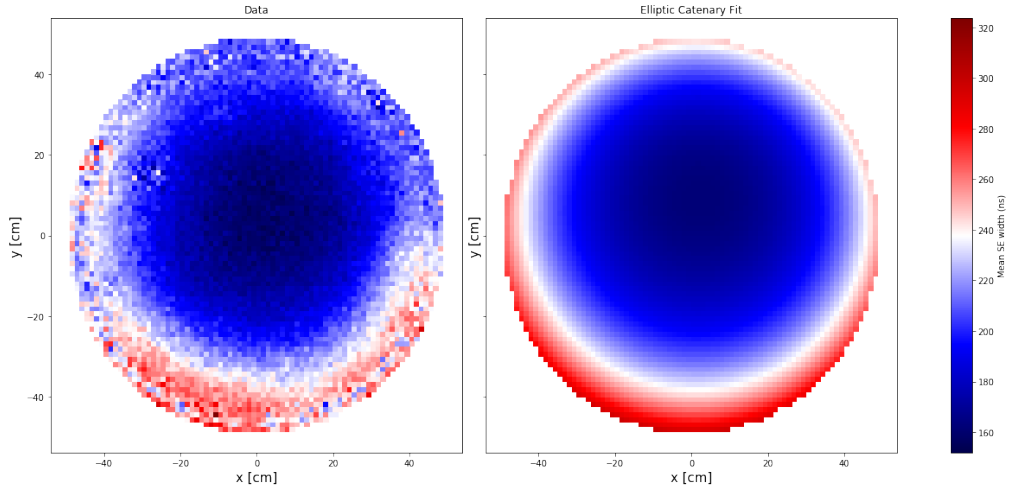


Figure 2.16 – Left: The width of the S2 signal of single electrons as a function of their reconstructed position, for $V_a = 4$ kV. Right: Fit of the elliptic catenary function on the data of the left Figure, which is found to best describe the data compared with the other functional forms of Tab. (2.2).

Functions used to fit the S2 width		
Name	Function	Free Parameters
Parabola	$(x^2+y^2-r_{max}^2)/2p_3$	p_3
Catenary	$p_3\cosh(\sqrt{x^2+y^2}/p_3) - \cosh(r_{max}/p_3)$	p_3
Elliptic Catenary	$p_4\cosh(\sqrt{p_3 x^2+y^2}/p_4) - \cosh(r_{max}/p_4)$	p_3, p_4

Table 2.2 – Three functional forms tested for the description of the sagging of the anode as inferred from the distribution of the S2 width of single electrons.

2.9 Determination of the secondary scintillation gain

The modelling of the spectrum of small S2s derived from single electrons, provide the tools for a systematic study of the secondary scintillation gain of the detector. In the gaseous region of the detector, where the extracted electrons are accelerated, the mechanism of electroluminescence is the result of two competing microscopic processes. The excitation of the Xe atoms takes place through inelastic collisions with electrons, although the kinetic energy of electrons is obtained between two inelastic collisions, where successive elastic collisions with Xe atoms take place. These elastic collisions represent the vast majority of collisions [187]. However, because the mass of the electron is negligible compared to that of Xe atom, the energy loss from these collisions is negligible. But we can imagine that if the gas pressure increases, the number of intermediary elastic collisions will increase significantly, resulting in energy losses that will not be negligible. Then the electrons will reach a kind of limit velocity and will never be able to acquire enough kinetic energy to excite the Xe atoms. The microscopic mechanism of scintillation in Xe gas has been studied experimentally, (see [188]) and is related to radiative de-excitation of the excited dimers Xe_2^*. In terms of the wavelength of the light produced, this is also related to the gas pressure. It has been experimentally observed that, for low pressures, the emission spectrum of the gas is continuous and is characterised by a peak called the first continuum, centered around 147 nm. In quantum mechanical terms, a continuum feature in the emission spectrum of a gas is a typical sign of a transition from an excited state to a so-called repulsive state *i.e.*, an electronic configuration of a molecule where the potential energy lacks a minimum [189]. The quantum state of the system is not a bound state, as the potential energy decreases with increasing interatomic distance, so the atoms repel each other and the vibrational energy levels degenerate into a continuum. Therefore, if a dimer is characterised by a repulsive ground state, the transition to this ground state from an excited molecular state will appear as a continuous feature in the emission spectrum. Indeed, the ground state of the Xe_2 dimer is a repulsive state, while its excited states are Rydberg states with the expectation values of the position of an electron in the molecular orbit being much larger than the interatomic distance Xe-Xe. This first continuum is therefore created by transitions from the higher vibrational levels $^1\Sigma_u^+$ of Xe_2^{**} to the repulsive ground state via the process $Xe_2^{**} \rightarrow 2Xe + h\nu$ with Xe_2^{**} created in the gas through processes like:

$$Xe^* + 2Xe \rightarrow Xe_2^{**} + Xe \tag{2.8}$$

As the gas pressure increases even more, the products of the three-body collision (2.8) begin to collide with each other, resulting in secondary processes like:

$$Xe_2^{**} + Xe \rightarrow Xe_2^* + Xe \tag{2.9}$$

As a result, the lowest vibrational molecular states of Xe_2^*, $^1\Sigma_u^+$ and $^3\Sigma_u^+$, de-excite to the repulsive ground state resulting in a second peak in the continuous emission spectrum centred on wavelengths around 172 nm. For pressures greater than 400 mbar, the process (2.9) dominate over (2.8) resulting in an emission spectrum consisting only of the peak centered at 172 nm, with a FWHM of 10 nm.

In this pressure region, a linear relationship has been observed, per unit pressure, between the electroluminescence yield, defined as the number of secondary scintillation photons produced per extracted electron and per unit length during its drift into the gas gap, with the electric field *i.e.*, a relation of form:

$$\frac{Y}{ph_g} = a\frac{E_g}{p} + b \tag{2.10}$$

Where p is the pressure of the gas inside a the gas gap of total width h_g, in which the electrons accelerate under an electric field E_g. The two parameters, a and b, can be determined by measuring Y, for different values of the electric field. These two parameters are very important for understanding the detector's performance but also for simulating the secondary scintillation process.

2.9.1 Variation with the anode voltage

In this section we will try to determine the parameters a and b of the relation (2.10) through a direct measurement of the electroluminescence yield. For this purpose we will use a series of dedicated calibration runs with a source of ^{83m}Kr taken at 5 different values of the anode potential: [3, 3.2, 3.4, 3.6, 3.8, 4] kV. These runs were taken on the week of June 18-23, 2018 in order to be used both in the present analysis and also in the evaluation of the variation of the extraction efficiency as a function of the electric field of extraction, that we will present in section (2.10). The ^{83m}Kr radioactive source offers a very high rate of triggered events, compared to simple background runs, resulting in many registered events during the run with a consequently high total number of single electrons that will provide statistical abundance in our analysis. It should be noted that what we call the secondary scintillation yield of the detector is a fraction of the electroluminescence yield of the relation (2.10) as, after the physical process of production of the scintillation light, a series of efficiencies related to the observation process are obviously inserted and should be taken into account, as for example the light collection efficiency , the quantum efficiencies of the PMTs and their photocathode collection efficiency. If we call $\bar{\eta}$ the product of the last two and $\bar{\beta}$ the average light collection efficiency for photons emitted inside the gas gap, then the secondary scintillation yield $\tilde{Y}$ is:

$$\tilde{Y} = Y\bar{\beta}\bar{\eta} = (a\frac{E_g}{P} + b)h_g P\bar{\beta}\bar{\eta} \tag{2.11}$$

For the purposes of this analysis, we will calculate the electric field E_g of the gas gap using the parallel plate approximation, although this approach is incorrect because, on

the one hand it leads to a systematic overestimation of the electric field, on the other hand it does not take into account the sagging and tilt of the anode, as we saw in the previous section. So, if d is the distance between the anode and the gate, (about 5 mm) and h_g is the width of the gas gap (about 2.5 mm for SR1) then, because the radius of the cylinder of the TPC is much larger than d, we can approach the electrostatic problem as two parallel plate capacitors connected in series. A capacitor formed between the gate and the isopotential surface of the phase boundary, that is filled with LXe which, from an electrostatic point of view, is a dielectric with a relative permittivity $\varepsilon_l = 1.96$, and a second capacitor, connected in series with the first, formed between the phase boundary and the anode, filled with GXe with $\varepsilon_g \approx 1$. In this context, we can approximate the electric field of the gas gap with the relation:

$$E_g = \frac{\varepsilon_l V_a}{\varepsilon_l h_g + d - h_g} \tag{2.12}$$

Where V_a is the potential of the anode with respect to the gate.

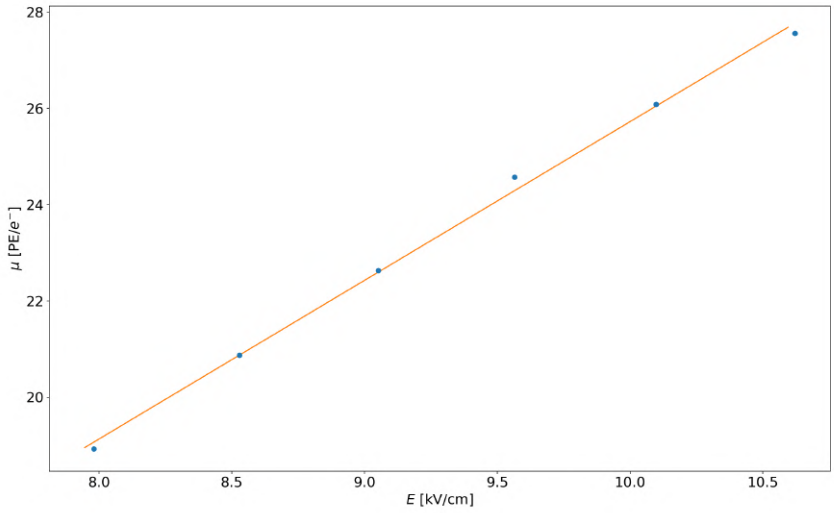


Figure 2.17 – The secondary scintillation gain for the six values of the anode voltage as a function of the electric field of GXe calculated with the parallel plate approximation. Also is show a linear fit which determines the two parameters a and b of (2.11).

Now, the secondary scintillation gain G can be determined by the mean μ of the model (2.7). Using all the available calibration datasets, for the six different values of the anode voltage, we extracted the corresponding spectra of small S2 signals, with an intensity of less than 150 PE appearing after the main S2 of each event, applying also all the data quality cuts described in section (2.5). The average light collection efficiencies,

provided by simulation, is 37.5%, the average collection efficiency from the photocathode to the first dynode is 90% and also the average quantum efficiency of the PMTs is 35% [190]. The values of the parameters a and b, extracted from the fit are:

$$a = 111.69 \pm 1.97$$

$$b = -126.98 \pm 14.35 \tag{2.13}$$

These values are in agreement with other measurements of the corresponding parameters in xenon gas at cryogenic temperatures [191].

2.9.2 Position dependence of the SSG

In Fig. (2.18) we present a two-dimensional histogram of the TPC surface where each bin is represented by a colour code proportional to its average secondary scintillation yield. This histogram can be done in two ways.

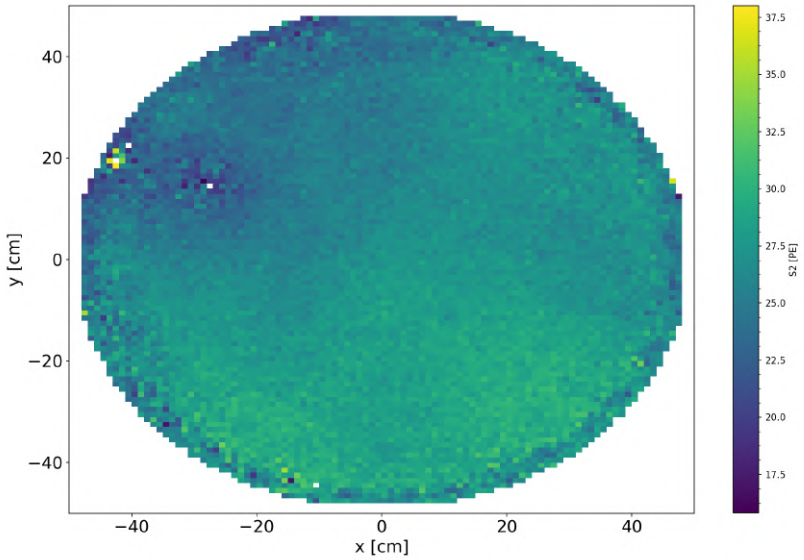


Figure 2.18 – Spatial dependence of the secondary scintillation gain of single electrons calculated as described in the text. The colour of each bin in this 2D histogramm corresponds in the colour scale to the average secondary scintillation gain of single electrons whose x-y position is reconstructed inside the bin.

Either by repeating the study of section (2.4.3) for each bin or, simply, by selecting from the plot of Fig. (2.15) those peaks that correspond only to the single electrons, through

their width and calculating their average area for each bin. In this case the calculated average value is a less accurate estimate of secondary scintillation yield. However, in this section, we used this approach because it is enough to show us the general picture that emerges in relation to its x-y dependence. We observe that it is not homogeneous, nor is it characterised by a radial symmetry, as one might expect at first sight from the analysis of the section (2.8). On the contrary, it seems to follow exactly, not the sagging of the anode but only the residual tilt. This is to be expected if we look carefully at the relation (2.11) using, even the simple approach of the electric field of the gas gap of the relation (2.12). On light of the values of parameters a and b that we inferred in the previous section, and inserting (2.12) in (2.11) we have:

$$Y = ah_g \frac{\varepsilon_l V_a}{\varepsilon_l h_g + d - h_g} + bh_g P$$

where we see that the factor $bh_g P$ is 5 times smaller than the factor $ah_g \varepsilon_l V_a$ for $V_a = 4kV$. This, results in:

$$Y \approx a \frac{\varepsilon_l V_a}{\varepsilon_l + d_l/d_g} \tag{2.14}$$

Where d_l is the liquid level *i.e.*, distance of the gate from the phase boundary. From (2.14) we see that Y (and therefore G) depends more on the ratio d_l/d_g and not so directly on d_g so the variation of G is expected to follow more the residual tilt of the TPC than the sagging of the anode.

2.9.3 Time stability and influence from the calibration source

In order to determine the time stability of the secondary scintillation yield we repeated the study of section (2.4.3) using runs with different calibration source, for a period of time that extends over a period of one year, specifically from 02/2017-02/2018. In this analysis we used runs with calibration sources of ^{220}Rn, Am-Be and ^{83m}Kr as well as background runs. From these runs we extracted again the spectrum of the small S2 that follows the main S2 of an event, where we again used the quality cuts of the corresponding analysis. Although, due to the abundance of single electrons, we could even use a single run and therefore do a monitoring with a high time resolution, here we chose to use time bins of one day to have statistical errors as small as possible. Of course, we do not expect any dependence of the value of the fit parameters and especially the value of μ from the source of calibration, because as we have already explained, the single electrons we use in this analysis are a background that is related to the detector's response to the scintillation light of a main S2 regardless of the source that produces this main S2, and this is related as we have seen exclusively to the physical parameters of the detector such as the anode voltage, the GXe pressure and width of the gas gap. Therefore any changes are not expected to be related to the source of calibration but rather to the time stability of these parameters. Indeed, in Fig. (2.19) we present the result of this analysis where

we actually see that the secondary scintillation yield is the same within the statistical error range, for each calibration source. We also notice that the value of μ is statistically constant in time, which reflects the high stability of the detector parameters on which it directly depends.

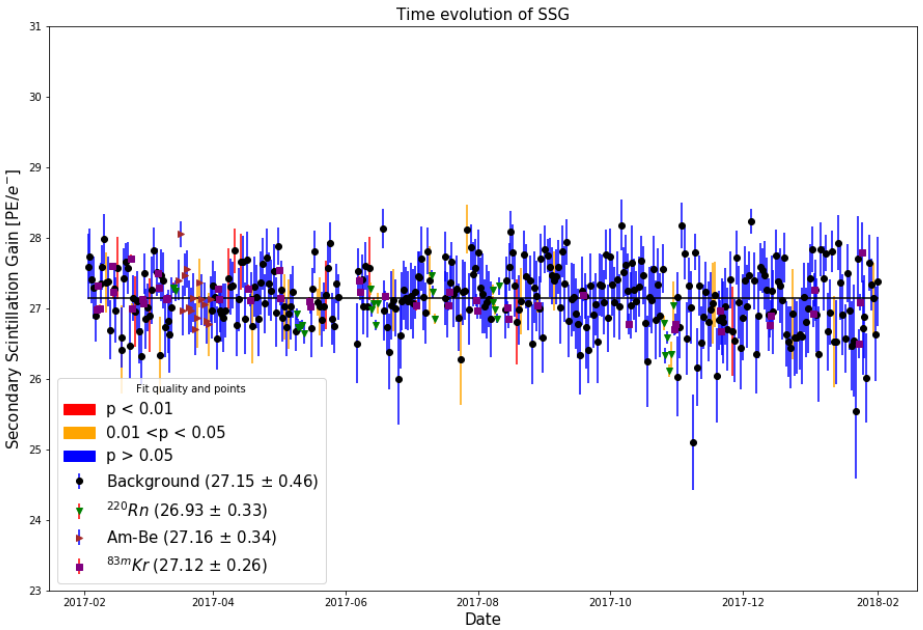


Figure 2.19 – Time evolution of the secondary scintillation gain for various calibration sources and for background runs. The color code correspond to quality of the fit determined from the corresponding p-value.

2.10 Determination of the extraction efficiency

A very important analysis that can be done using the background of single electrons is related to the study of the probability of extractions of an electron from LXe, as a function of the extraction electric field, or otherwise, the extraction efficiency. It is possible to perform an indirect measurement of the extraction efficiency by comparing the number of electrons expected to be generated by a known energy deposition with the number of electrons eventually observed, this last information being essentially derived from the amount of light produced per extracted electron. In this case it is necessary to use a calibration source so that the population of the main S2 of the events created by a known energy deposition is being selected. In this analysis we will use the calibration runs

used also in the previous section as they contain the isomeric ^{83m}Kr that is continuously generated from the decay of the ^{83}Rb decay. The produced ^{83m}Kr is in a level 41.5 keV above its ground state in which decays in two steps, via emission of a 32.1 keV conversion electron, process of a half-life of 1.83 h and subsequently to the ground state via emission of a 9.4 keV conversion electron with a half-life of 154 ns. Due to the time proximity of the two decays in most cases the detector will detect the light and the ionization electrons that will be produced by the two conversion electrons as if they came from the same event, so the two processes will be perceived as a single 41.5 keV energy deposit in the detector. Under given conditions of drift electric field and charge losses during drift due to captures by electronegative impurities, the number of electrons that will reach the separation surface is known and can be compared with the number of electrons that will be measured through the main S2 light corresponding to this radiation decay and the information of the secondary scintillation gain. In this case we must estimate the so-called S2 gain associated to the radiative decay of ^{83m}Kr normalised to the energy of 41.5 keV in such a way as to take also into account the effect of charge loss during the drift of the main electron cloud *i.e.*, an extrapolation of the S2 light for drift times corresponding to small depths.

One point that presents the advantage of using ^{83m}Kr as a source of calibration is that the characteristic energy peak is of relatively low energy. One of the problems that would arise if we used a radioactive source of a higher energy deposit, as was the case in the corresponding analysis of the XENON100 detector, was that we would have to pay attention to the fact that the S2 signal would be so intense that it would lead to saturation the PMTs of the top array which observe the larger part of it, that is a loss of the output linearity with respect to the incident light intensity. We distinguish two types of such saturation: 1) The ADC saturation when the output of a PMT is larger that the dynamic range of 2.25 V of the ADC resulting in a truncated waveform and so to an underestimation of the area of a peak. 2) The Base saturation where here, a very strong signal on the photocathode of the PMT will cause several initial electrons which, being multiplied at the dynodes, will result in a weaknesses of the last dynodes to continue the proportional growth of the signal because there will just not be enough electrons available. This problem is usually mitigated by including suitably a certain number of capacitors to the circuit.

In order to avoid this and therefore any bias in measuring the total light intensity of S2 it would be preferable then to use only that part of the scintillation light observed by the bottom array. This would raise the issue that, in order to be able to compare this part of the S2 gain with the secondary scintillation light we would also have to, *a priori*, repeat analysis of section (2.4.3) using only that part of the light observed from the bottom array. However, as the anode voltage decreases, so does the secondary scintillation gain, with the result that the spectrum of small S2s in Fig. (2.3) is increasingly pushed to the left. This would make the single electrons and their pileups increasingly indistinguishable,

and the corresponding fit of the model (2.7) would become more unstable. The solution used in the XENON100 analysis was to perform an analysis of the light distribution of the small S2s between the two arrays of PMTs, taking advantage of the fact that for these very small light intensities we are far away from the saturation region.

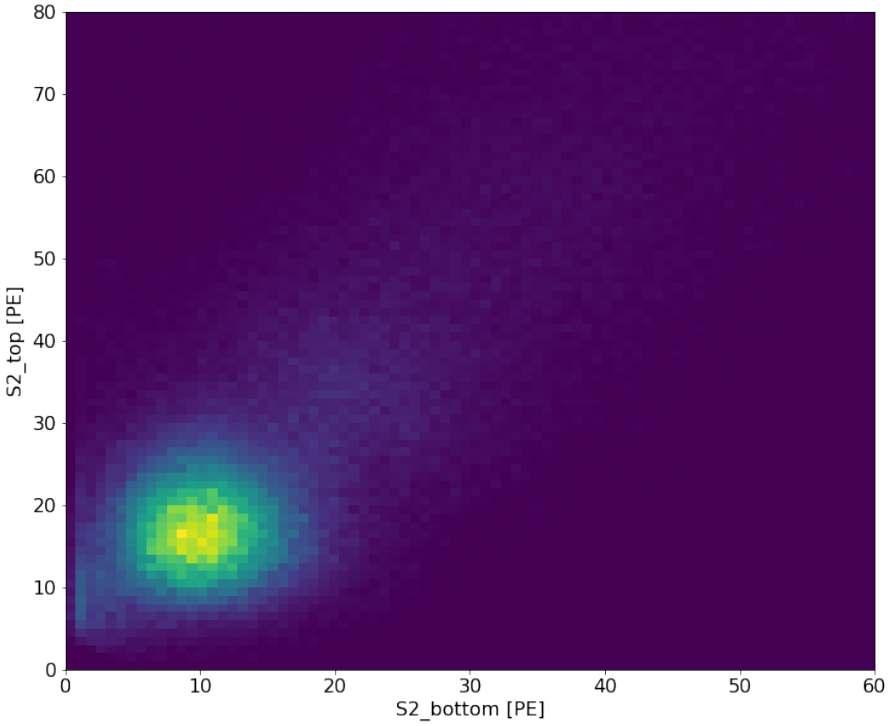


Figure 2.20 – The S2 light repartition between the top and bottom PMT array.

For example, in Fig. (2.20) we present, for single electrons corresponding to a dataset of calibration runs (with ^{132}Cs as source) where the potential of the anode was set to 2.5KV, the light intensity S2 observed by the top array as a function of the corresponding amount on the bottom array. In Fig. (2.21) we present, for the data corresponding to the single electrons of Fig. (2.20), the percentage of light observed by the bottom array. In this distribution we adjust a Gaussian whose average value describes the percentage of total light observed by the bottom array when there is no saturation. Thus, in such an analysis, if we only wanted to use the part of S2 that is observed from the bottom PMT array, we would compare it with the secondary scintillation light multiplied by the average value of the Gaussian of Fig. (2.21).

However, in the XENON1T detector, saturation due to ADC begins to occur for $S2 > 10^5$ PE and base saturation for $S2 > 10^6$ PE, so in our case where the main S2 of ^{83m}Kr is in the range of 2×10^4 PE we are still in the linear response region of the PMTs.

In order to take into account only the reconstructed signal corresponding to the peak of 41.5 keV from the two conversion electrons, we apply a cut to the events that have

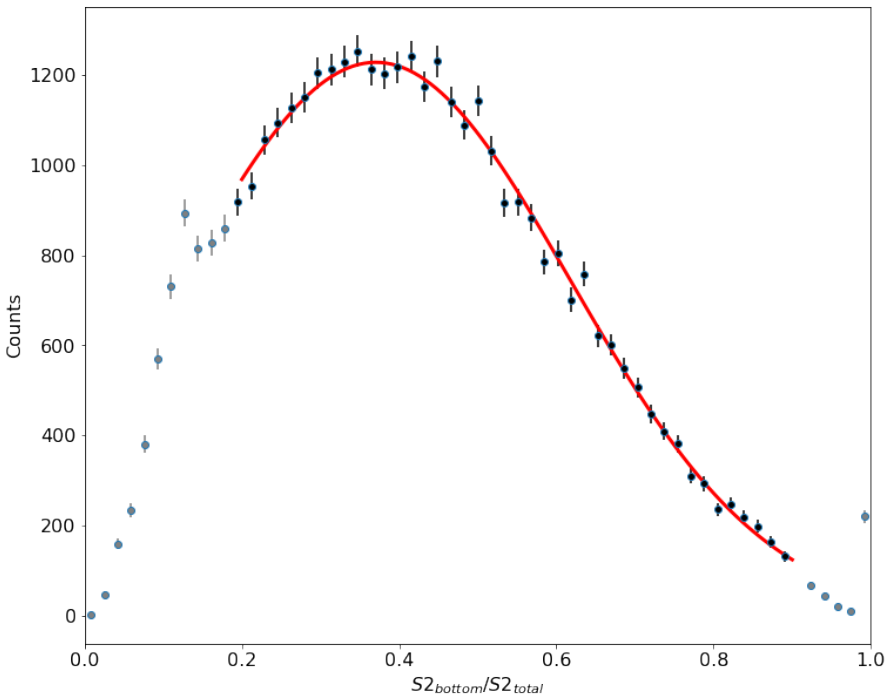


Figure 2.21 – Fraction of the S2 light observed by the bottom PMT array, and fit with a Gaussian function.

second large S1 except from the main S1 *i.e.*, we keep only events where the primary scintillation signal from the two conversion electrons is unresolved.

2.10.1 Position dependence of the extraction efficiency

The extraction efficiency depends on the extraction electric field which becomes a function of the position (x, y) due to the sagging of the anode. We can illustrate this variation if we simply divide the distribution in the x-y plane of the S2 light from the ^{83m}Kr main events corrected for the losses due to drift with the corresponding distribution of secondary scintillation gain. The S2 light corresponding to the ionization electrons of the ^{83m}Kr events depends on the extraction efficiency while the secondary scintillation gain does not as an electron either is extracted to the gaseous phase or not. Therefore the ratio of the two will be directly proportional to the extraction efficiency. In Fig. (2.22) we present the distribution of this ratio at the x-y plane for 4kV anode voltage. This distribution has been scaled to 1 where we can see that the distribution follows the electric field of the gas gap which in turn follows more the sagging and less the residual tilt.

The measurement of the extraction efficiency that we will perform is relative, in the following sense: For the 4kV dataset we will choose the interior of a small circle of radius 5cm around the point $(x, y) = (0, 0)$ where, according to the Fig. (2.14) we expect the

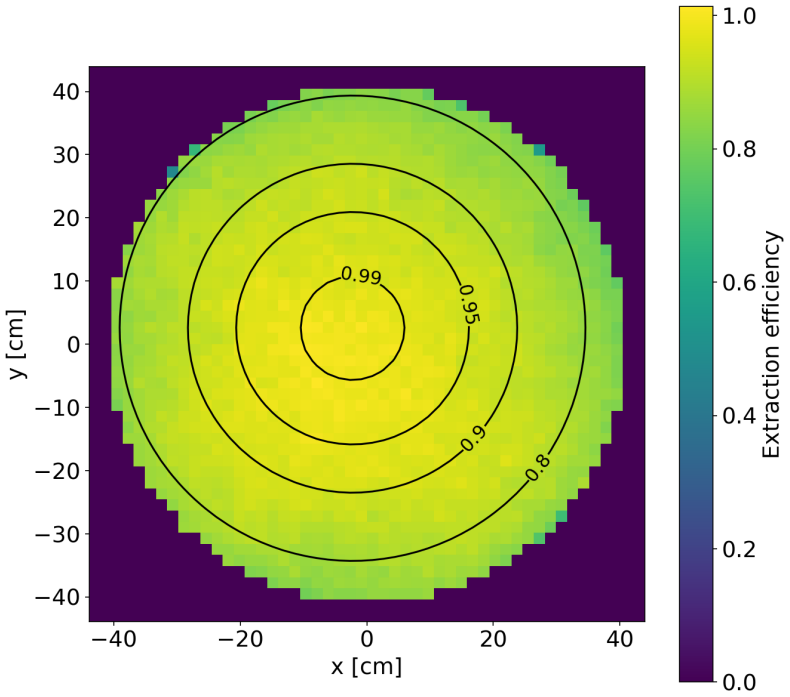


Figure 2.22 – x-y distribution of the ratio of thw S2 gain to the secondary scintillation gain. The result is directly proportional to the extraction efficiency.

smallest width of the gas gap and therefore the higher electric field. There, the extraction electric field just above the phase boundary will be higher than average extraction field at other regions of the surface. We will assume that the number of electrons, per keV, produced by ^{83m}Kr, in a cylinder of the same diameter in the active volume of the detector is extracted with a 100% efficiency through this cycle due to the very high electric field. We will then compare this number with the number of electrons per keV extracted in the $R > 10$ region where we are away from the center of the parabola of the sagged anode and where the gas gap width can be considered compatible with the difference between the 5mm distance between the anode and grounded gate and the width of the liquid level measured by the level-meters. If we consider that 1) $N[e^-/keV]$ is the number of electrons per keV that are extracted for R<5 cm and 2) this number is the 100% of the

electrons produced per keV in a cylinder of radius R<5cm within the active volume, then the extraction efficiency is given by the relation

$$\mathcal{E} = \frac{G_{S2}[PE/keV]}{\mu[PE/e^-]N[e^-/keV]} \tag{2.15}$$

Where $G_{S2}[PE/keV]$ is the S2 gain and $\mu[PE/e^-$ the secondary scintillation gain. $N[e^-/keV]$ is given from:

$$N[e^-/keV] = \frac{G_{S2,R<=5}[PE/keV]}{\mu_{R<=5}[PE/e^-]} \tag{2.16}$$

Where $G_{S2,R<=5}[PE/keV]$ is the S2 gain and $\mu_{R<=5}[PE/e^-]$ is the secondary scintillation gain for the electrons extracted in the region $R < 5cm$.

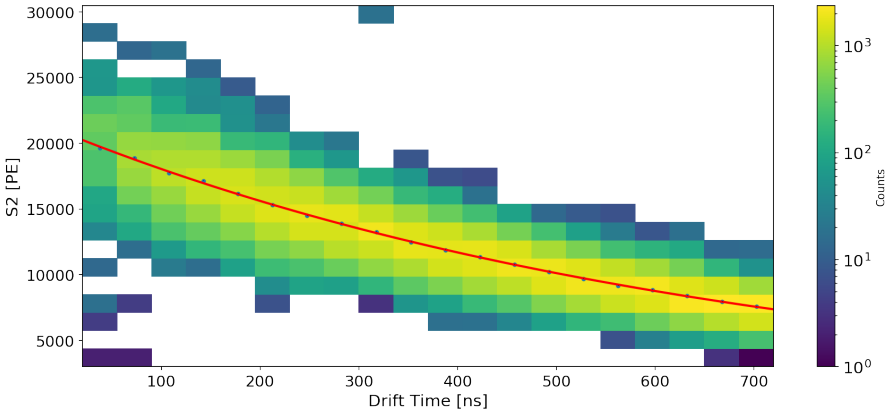


Figure 2.23 – The S2 gain as a function of drift time and the fit with an exponential for electrons extracted in the region $R < 5cm$ were the extraction efficiency is considered to be 1. The fit is used in order to extrapolate the S2 gain to zero drift times in order to remove bias from the charge loss during drifting.

In Fig. (2.23) we present S2 as a function of drift time for the events in the region $R < 5cm$ as well as the fit of the exponential function with which we will determine the S2 gain extrapolating to zero drift times. The values of $G_{S2,R\leq5}[PE/keV]$, $\mu_{R\leq5}[PE/e^-]$ and $N[e^-/keV]$ are given in Table (2.3).

S2 gain and SSG for $R \leq 5cm$	
Parameter	Fit Value
$G_{S2,R\leq5}[PE/keV]$	525 ± 8.1
$\mu_{R\leq5}[PE/e^-]$	28.4 ± 0.14
$N[e^-/keV]$	18.48 ± 0.3

Table 2.3 – Values of the S2-gain, the secondary scintillation gain and of the value $N[e^-/keV]$ for the region $R < 5$ cm, where we assume that the extraction efficiency is 1.

2.10.2 Variation of the extraction efficiency with the anode voltage

We can now calculate the $\mathcal{E}$ from relation (2.15) for the datasets of the 6 different values of the voltage of the anode. In Fig (2.15) we present the graph of $\mathcal{E}$ as a function of the electric field. The same figure shows the corresponding measurements made in the XENON100 experiment [106]. In the present analysis, the electric field was calculated from relation (2.12) *i.e.*, using the parallel plate approximation, and corresponds to an average value of the electric field in the gaseous region, without taking into account the sagging. In other words, it can give an estimate of the electric field, to the extent that the approximation of the two parallel plane capacitors applies and to the extent that the width of the gas gap is determined by the liquid level and the anode-gate distance *i.e.*, if we are relatively far from the center of the catenary of Fig. (2.16).

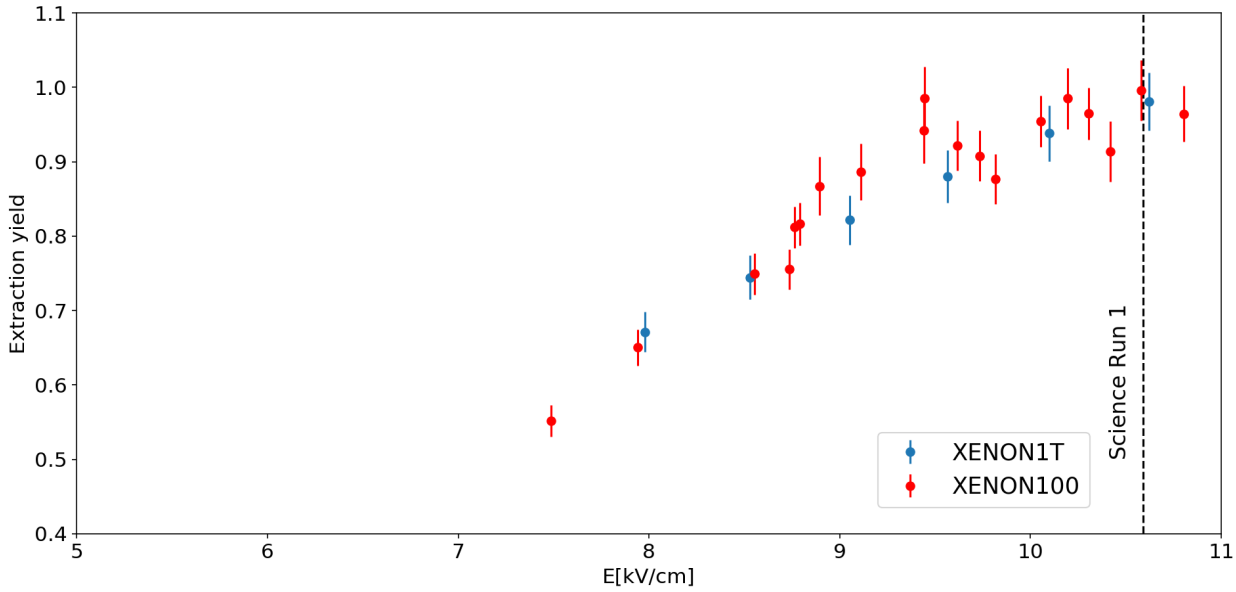


Figure 2.24 – The extraction yield evaluated in this work versus the extraction yield for XENON100 from [106] as a function of the extraction electric field in GXe, calculated with the parallel plate approximation. The dashed line corresponds to the value of the electric field during SR1. With only the variation of the anode voltage we are in the region of electric fields were the extraction efficiency is only increasing.

In [106] the electric field of the relation (2.12) was corrected through a two-dimensional simulation of the electrostatic problem of the extraction region. It was found that the electric field in the gas gap just above the liquid level, as determined by the simulation, is about 11% lower than the one calculated with the parallel plate approximation. However, in the figure, the XENON100 data does not contain this correction, so they are determined by the relation (2.12) as well as the present analysis. A three-dimensional simulation of the electric field in the extraction region, done by the collaboration, showed that, in the

case of the 4kV anode voltage, the extraction field is 24% smaller than the one calculated by the relation (2.12). In the next section we will see that, in fact, the relation (2.12) is an overestimation of the electric field and that even with a analytical two-dimensional approximation of the electrostatic problem we can obtain an order of magnitude of this overestimation.

We observe an agreement with the measurement of [106] and also that, for the 6 values of the anode voltage that we used, we are situated in the region where the extraction efficiency increases in order to reach its maximum value. Unfortunately, we cannot determine the existence of a plateau, due to the lack of data for higher values of the electric field, as in the case of XENON100. Such a measurement would be possible if, in combination with the variation in the value of the anode voltage, we also variated the value of the liquid level. This would result in measuring the extraction efficiency for values of the electric field much smaller than that corresponding to $V_a = 3$ kV and much higher than that corresponding to $V_a = 4$ kV voltage, as shown in Fig. (2.25). However, a change in the liquid level was not possible due to various limitations imposed by the main experimental task during the Science Runs.

2.11 The extraction electric field

In this section, we will limit ourselves to a demonstration of the reason why, the simple parallel plate approximation, overestimates the extraction field and we will also show that it is possible to calculate analytically an order of magnitude of this overestimation. We will make a slightly more realistic representation of the electric field of the gas gap by inserting a second dimension. Specifically, we will assume that the anode is not a plane but consists of parallel wires, which are also parallel to the direction y, in Cartesian coordinates, they are infinite in number and infinite in length and have a specific diameter r_w (Fig. (2.26)). Along the x axis, they are placed at a distance s, from each other (pitch), and every wire is situated at a distance $z = d$ from the plane $z = 0$, which defines the phase boundary. We will continue to consider the grounded gate as a conducting surface located at zero potential, and placed at the plane $z = -d$, under the boundary surface. This electrostatic problem can be solved analytically and is formulated as follows [192]:

$$\varepsilon_g \nabla \mathbf{E} = 0, z > 0 \tag{2.17}$$

$$\varepsilon_l \nabla \mathbf{E} = 0, z < 0 \tag{2.18}$$

And the solution must satisfy the boundary conditions:

$$\lim_{z \to 0^+} (E_x, \varepsilon_g E_z) = \lim_{z \to 0^-} (E_x, \varepsilon_l E_z) \tag{2.19}$$

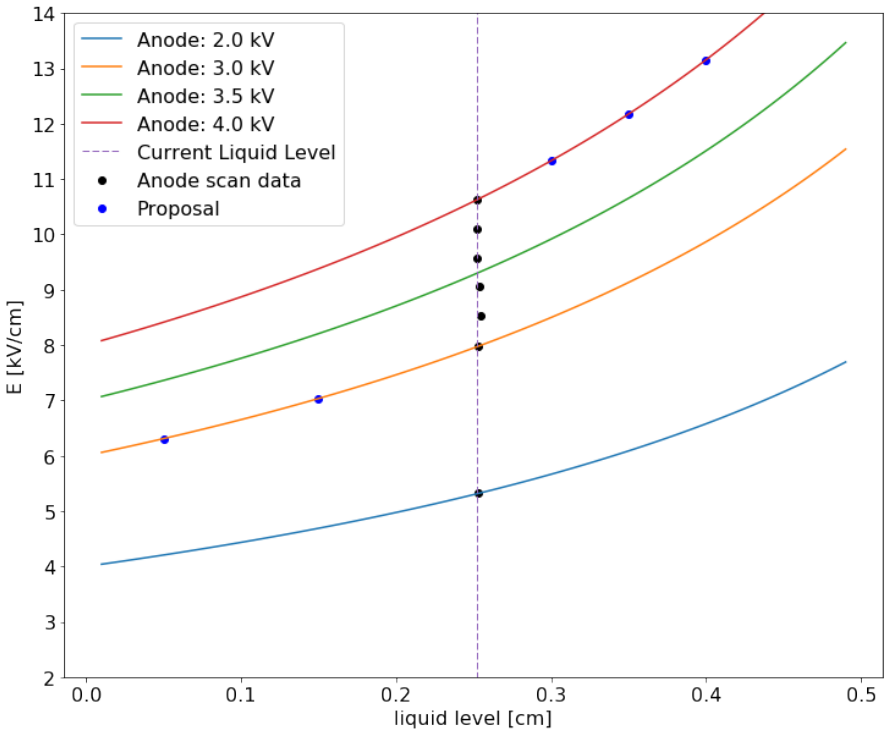


Figure 2.25 – The electric field in GXe as a function of the liquid level for the parallel plate approximation. The XENON1T data points of Fig. (2.24) corresponds to the liquid level indicated by the black points. For this liquid level the variation of the voltage of the anode result in an electric field with intensity in the region were the extraction efficiency is strictly increasing so we cannot observe the saturation region. This could be possible with a modification of the liquid level as proposed (blue points). Eventually this was not possible due to strict conditions concerning the stability of the detector during normal data taking.

Because $\nabla \times \mathbf{E} = 0$ everywhere, this two-dimensional problem can be solved with the help of the complex potential of an infinite charged line with charge per unit length equal to λ. The field of such a line, placed parallel to the y axis and at a point $\rho_0 = x_0 + iz_0$, situated opposite to a grounded surface at $z = -d$, is given at a point $\rho = x + iz$ by the complex potential [186]:

$$\phi(x,z) = -\frac{\lambda}{2\pi\varepsilon_g} \ln \frac{x - x_0 + i(z - d - z')}{x - x_0 + i(z - d + z_0)}$$

We can easily take into account the two different dielectrics, considering that at a point P of the gaseous region (Fig. 2.26) the potential now will be due to the line passing through the point $\rho_0 = x_0 + iz_0$, plus a potential due to the image of this distribution of

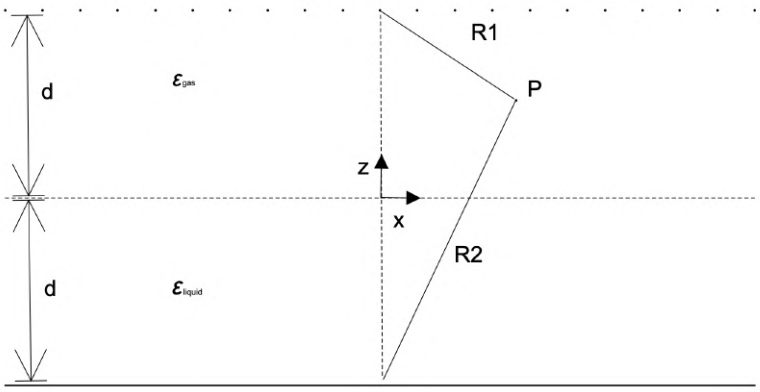


Figure 2.26 – Illustration of the 2D geometry of the electrostatic problem. Here the reference system in centered at the phase boundary surface, the x axis lies on the surface and the z axis is normal to the surface. LXe and GXe are dielectrics with dielectric constants e_l and e_g. The field is calculated at a point P(x,z) inside the space between the gate and the anode, solving the Poisson equation with the appropriate boundary conditions at the surfaces of the two electrodes.

charges, which is also a charged line, with charge per unit length λ_{im}, that passes through the (symmetric with respect to the $z=0$) point $\rho' = x_0 - iz_0$. That is:

$$\phi_{gas}(x,z) = -\frac{\lambda}{2\pi\varepsilon_g}\ln\frac{x-x_0+i(z+d-z_0)}{x-x_0+i(z+d+z_0)} + \frac{\lambda_{im}}{2\pi\varepsilon_g}\ln\frac{x-x_0-i(z-d+z_0)}{x-x_0-i(z-d-z_0)}$$

The real part $\Re(\phi_{gas}(x,z))$ is the solution of the Poisson equation in the region $z>0$ and also satisfies the boundary condition at the grounded gate $\Re(\phi_{gas}(x,-d))=0$. If, now, we have an infinite number of parallel charged lines, extending to the right and left of the previous line, placed with a pitch s *i.e.*, lines passing through the points $\rho'_k = x_0 + ks + iz_0 = \rho_0 + ks$ with $k = \ldots -2,-1,0,1,2,\ldots$ then, the complex potential at a point $P=(x,z)$ of the region $z>0$, is given by the relation:

$$\phi_{gas}(x,z) = -\frac{\lambda}{2\pi\varepsilon_g}\sum_{k=-\infty}^{\infty}\ln\frac{x-x_0-ks+i(z+d-z_0)}{x-x_0-ks+i(z+d+z_0)} + \frac{\lambda_{im}}{2\pi\varepsilon_g}\sum_{k=-\infty}^{\infty}\ln\frac{x-x_0-ks-i(z-d+z_0)}{x-x_0-ks-i(z-d-z_0)} \tag{2.20}$$

The infinite series of (2.20) can be evaluated and gives as a result a function, the real part of witch, is the solution in the region $z>0$:

$$V_{gas}(x,z) \equiv \Re(\phi_{gas}(x,z)) = -\frac{\lambda}{2\pi\varepsilon_g}\ln\Big(\frac{\sin(\frac{\pi}{s}(x-x_0))^2+\sinh(\frac{\pi}{s}(z+d-z_0))^2}{\sin(\frac{\pi}{s}(x-x_0))^2+\sinh(\frac{\pi}{s}(z+d+z_0))^2}\Big)+$$

$$+\frac{\lambda_{im}}{2\pi\varepsilon_g}\ln\Big(\frac{\sin(\frac{\pi}{s}(x-x_0))^2+\sinh(\frac{\pi}{s}(z-d+z_0))^2}{\sin(\frac{\pi}{s}(x-x_0))^2+\sinh(\frac{\pi}{s}(z-d-z_0))^2}\Big) \tag{2.21}$$

Then, the solution in the region $z<0$, is the potential due to the lines in $z>0$, as perceived inside the dielectric with ε_l, that is, the potential of the same geometry of charged lines situated in the place of the real ones but with a different charge per unit length, λ', that will be determined by the necessity that the potential should be continuous at the phase boundary:

$$V_{liquid}(x,z)=-\frac{\lambda'}{2\pi\varepsilon_l}\ln\Big(\frac{\sin(\frac{\pi}{s}(x-x_0))^2+\sinh(\frac{\pi}{s}(z+d-z_0))^2}{\sin(\frac{\pi}{s}(x-x_0))^2+\sinh(\frac{\pi}{s}(z+d+z_0))^2}\Big) \tag{2.22}$$

Now, the constants λ, λ_{im} and λ' will be determined by the two boundary conditions (2.19) together with a third boundary condition on the surface of the anode's wires. There we should have:

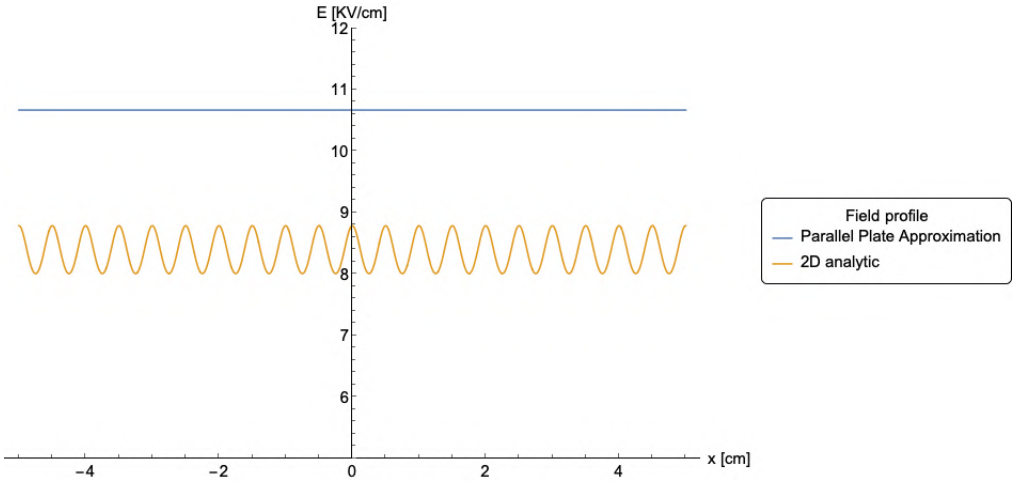
$$\phi_{gas}(x_w,z_w)=V_a \tag{2.23}$$

For every x_w and z_w on the surface of the wires of radius r_w

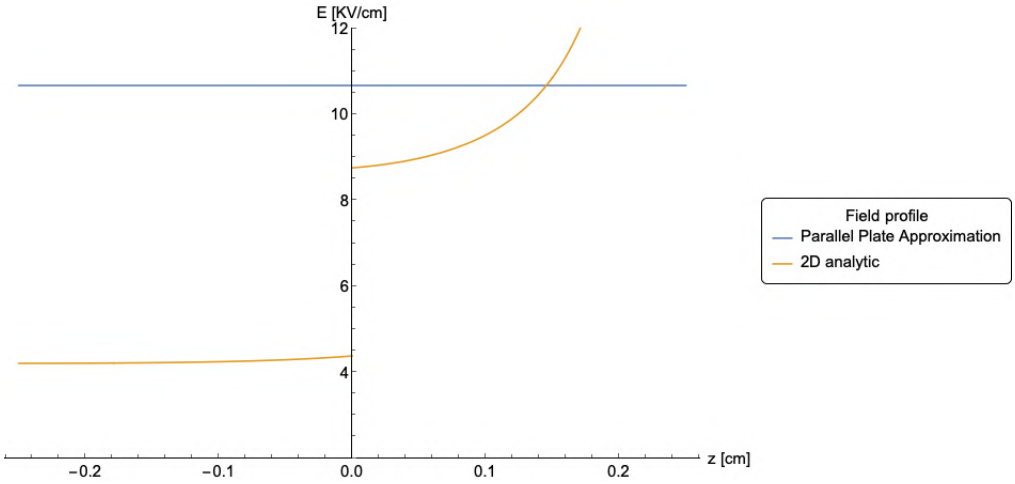
$$(x_w-x_0)^2+(z_w-z_0)^2=r_w^2 \tag{2.24}$$

Fig. (2.27), shows the solution (2.21) for the boundary conditions (2.23), (2.24), as a function of x, for $z=0.1$ mm above the phase boundary, where we used the values $s=3.5$ mm, $r_w=89$ μm, $d=2.5$ mm and $V_a=4$ kV, that is the actual values of the corresponding parameters of the detector [125]. We can first notice that the field, as a function of x, has an undulatory form due to the segmentation of the anode and the fact that the pitch is greater than the distance of the anode from the phase boundary, but that the mean value is, indeed, less than that calculated by the parallel plate approximation and, in particular, a relation of form $E_{analytic}=0.84E_{parallel-plate}$ holds.

For example, for a potential $V_a=4$ kV, the value of the field calculated by the parallel plate approximation is 10.67 kV/cm, while the analytical solution (2.21) gives the value 8.95 kV/cm. It is worth noting that the three-dimensional simulation of the extraction field, made by the collaboration, calculates the value 8.13 kV/cm. It is obvious that, the more realistic description of the anode and the gate as hexagonal etched meshes, as well as, the introduction of the third dimension that further restricts the infinite mesh of Fig (2.26) of this analytic approximation, will result in a further reduction of the estimation of the electric field but it will not be so large, because the field inside the LXe is twice as weak as the field in the GXe, due to the fact that $\varepsilon_l=2\varepsilon_g$, and also because the radius of the anode is much larger than the gas gap, so Fig (2.26) is actually not very far from reality, regarding the average extraction field just above the liquid level. Of course, both this analytic calculation and the 3d simulation, do not take into account the wrapping of the anode.



(a) a



(b) b

Figure 2.27 – 2D analytic solution compared to the parallel plate approximation. Here, the values used for s, r_w, d and V_a, are the actual values of the corresponding parameters of the detector. [a]: Field profile (2.21), as a function of x, for $z = 0.1$ mm above the phase boundary. [b]: Same field profile as a function of z for $x = 0$.

In Fig. (2.27) we present the solution (2.21) as a function of z for $x = 0$. We see the discontinuity of the field across the boundary (corresponding to a surface charge density induced in the boundary) as well as the behavior $\sim 1/(z-d)$ as $z \to d$, that is characteristic of a field created by a wire, because for $z \to d$ we can actually resolve the contribution of the individual wire of the mesh.

It is also worth noting that from (2.21) and (2.22) we can recover the parallel plate approximation. Indeed, if in Fig. (2.26) we consider that the diameter r_w and the pitch s tend to 0 at the same time, then the anode of Fig. (2.26) tends to a plane, in the mathematical sense. Also this plane is at a voltage V_a. Then the solution (2.21) and (2.22) must tend to the field (2.12) of the parallel plate approximation as $(s, r_w) \to 0$. Indeed we can observe this reduction in Fig. (2.28).

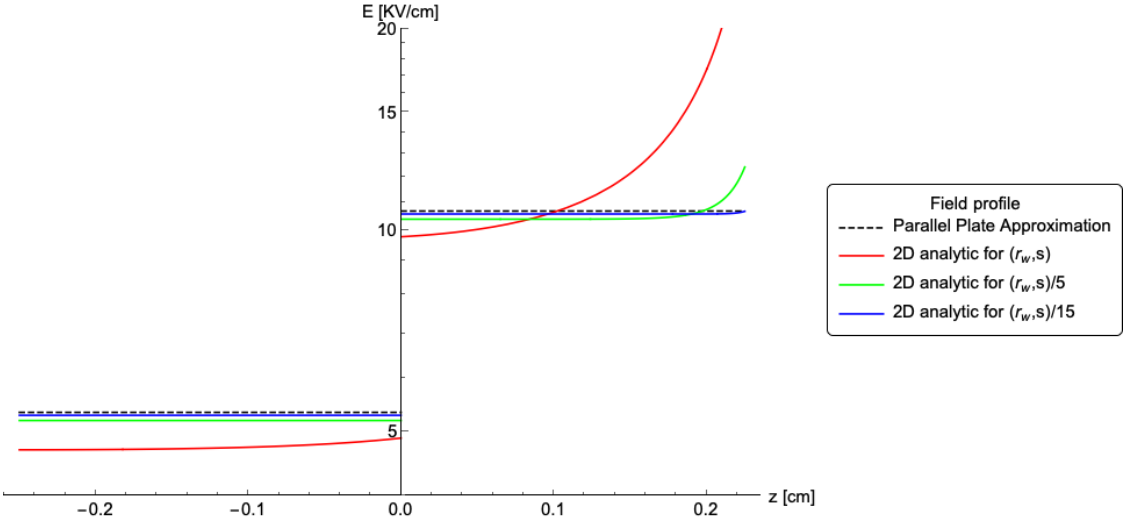


Figure 2.28 – Illustration of the reduction of the 2D analytic solution of this section to the parallel plate approximation as $(s, r_w) \to 0$.

2.12 Comparison with a theoretical model of the extraction efficiency

Of particular interest is an attempt to theoretically interpret the dependence of the electron extraction efficiency on the electric field [193]. At first glance, the microscopic process that governs the extraction is the fact that the external electric field accelerates the electrons but also reduces the potential barrier that is created by the mere presence of this ionization electron which approaches, from the LXe side, the dielectric boundary that separates the liquid from the gaseous phase. Indeed, from an electrostatic point of view, we have a charge situated at a position, say $z_0 < 0$, within a dielectric (LXe), trying to enter into another dielectric (GXe), with a smaller dielectric constant. In this

case, the Poisson equation can be solved satisfying the necessary boundary condition at the dielectric boundary, if we consider that, within the other dielectric (GXe) exists an imaginary "image" charge q_{im}, symmetrically positioned to the real one, with respect to the dielectric boundary *i.e.*, at the point $z'_{im} = -z_0 > 0$. From elementary electrostatics q_{im} is then given from:

$$q_{im} = e\frac{\varepsilon_g - \varepsilon_l}{\varepsilon_g + \varepsilon_l}$$

The potential that the electron perceives because of this image charge is:

$$\phi_{im}(z) = -\frac{e}{8\pi\varepsilon_l z}\frac{\varepsilon_l - \varepsilon_g}{\varepsilon_g + \varepsilon_l} \tag{2.25}$$

And its potential energy is:

$$W_{im} = -eV_{im} = \frac{e^2}{8\pi\varepsilon_l z}\frac{\varepsilon_l - \varepsilon_g}{\varepsilon_g + \varepsilon_l} \tag{2.26}$$

We see from relations (2.25) and (2.26) that if $\varepsilon_l > \varepsilon_g$, this potential is repulsive and the LXe tends to prevent the electron from crossing its surface and entering into the medium with the smaller dielectric constant.

Of course, the concept of the boundary surface between the two mediums, in electrostatics, is purely geometric and this is directly related to the fact one immediately notices, that the potential energy (2.25) tends to infinity as $z \to 0$. In this case we have a breaking down of classical electrostatics as the equation we solve, $\nabla \times \mathbf{E} = 0$, is actually a suitable averaging over microscopically small regions in order to describe the electrostatics of macroscopic media [192]. Physically, the boundary surface is defined in the atomic level by atoms ranged within a certain mean distance between them. If, for example, we consider that the lattice constant of solid xenon is of the order of 10^{-10} m then, from (2.25), the maximum potential energy due to the image potential *i.e.*, the potential barrier is not infinite but rather of the order of $\phi_{im} = 10^{-1}$ eV. Of course, the fact that there is an external electric field E, in addition to accelerating the ionization electrons, also contributes to the reduction of this potential barrier, because the total potential in witch the electron moves is:

$$W_{tot} = W_{im} + W_{field} = \frac{e^2}{8\pi\varepsilon_l z}\frac{\varepsilon_l - \varepsilon_g}{\varepsilon_g + \varepsilon_l} - Ez$$

The potential barrier of $\approx 10^{-1}$ eV, has been reduced by:

$$\Delta\phi = e\Big(\frac{Ee(\varepsilon_l - \varepsilon_g)}{2\pi\varepsilon_l(\varepsilon_l + \varepsilon_g)}\Big)^{1/2}$$

And in fact the potential now has also a minimum, just below the surface (Fig. 2.29). Concerning the energy distribution, $f(\varepsilon)$, of the cluster of ionisation electrons that

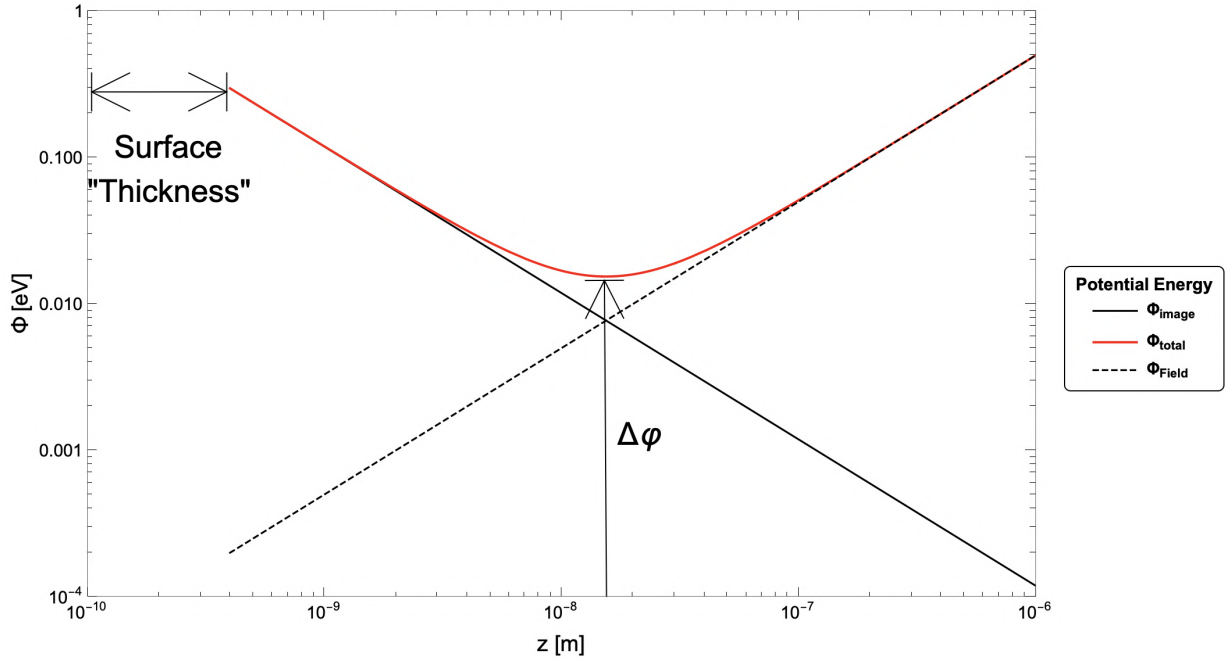


Figure 2.29 – The potential energy due to the electrostatic image of the charge (black solid line) and the potential energy due to the extraction field (black dashed line). Also is show the total energy (red solid line) as a function of the distance from the phase boundary. The extraction field cause a reduction of the potential barrier and the thickness of the surface is effectively defined by the order of magnitude of the interatomic distance.

are accelerated by the extraction field, this can be calculated from the solution of the corresponding Boltzmann equation [194]. Given $f(\varepsilon)$, we can then theoretically define the extraction efficiency as [193]:

$$\mathcal{E} = \int_{\phi_{im}-\Delta\phi}^{\infty} \varepsilon^{1/2} f_0(\varepsilon) d\varepsilon \Big/ \int_0^{\infty} \varepsilon^{1/2} f_0(\varepsilon) d\varepsilon \tag{2.27}$$

That is, the percentage of electrons that have just enough kinetic energy to overcome the reduced potential barrier $\phi_{im} - \Delta\phi$.

A subtle point about the relation 2.27 is that it can be considered that indeed it does describe the probability of extraction, if the width of the liquid/gas interface is less than the order of magnitude of the electron mean free path, in the gas phase. Because, one could imagine that, for high enough pressures of the GXe, the gaseous phase could influence the emission probability. However, if we consider that the phase boundary has a width of the order of the interatomic distance of 10^{-10} m and that the electron mean free path in GXe near the triple point is $\sim 10^{-6}$ cm then we see that such influence of the gaseous phase could occur above a pressure of 30 bar.

In [195] the so-called n-th chance model for describing the extraction efficiency is suggested: Electrons are scattered multiple times during their drift into the LXe. The fact that there is a minimum in the potential leads us to the idea that an electron can try multiple times to cross the surface or, eventually, thermalise at the minimum of the potential. Therefore the equation (2.27) must be seen rather as a single attempt of the electron to cross the surface, which may be repeated many times. So what would best describe the extraction, would be a cumulative process, where the probability κ_n of the electron being finally extracted after n trials, would be given by:

$$\kappa_n = 1 - (1-\kappa)^n \tag{2.28}$$

We will try to use the results we found in the previous section, on the extraction efficiency of XENON1T, along with those of XENON100 (Fig. 2.24), to fit the model (2.28). For $f(\varepsilon)$ we will use the solution of the Boltzmann equation for electrons moving in a steady electric field in gases, liquids and solids, as described in [194].

In that kinetic theory of hot electrons, $f(\varepsilon)$ is given from the relation:

$$f(\varepsilon) = \left(\frac{\varepsilon}{kT} + b\right)^b \exp(-\varepsilon/kT)$$

Where b is given from:

$$b = \frac{1}{3}(eE\Lambda_0)(eE\Lambda_1)/(2m/M)(kT)^2$$

In [194] two distinct mean free paths are defined, Λ_0 and Λ_1, as here, kinetic theory generalises the concept of mean free path in a perfect gas, taking into account that, in liquids and solids, there are phenomena related to some kind of structure and this, in the framework of kinetic theory, appears as a difference between the rates of energy and momentum transfer, which must then be described by two different mean free paths. Specifically Λ_0 is independent of the structure of liquids or solids and is related to the average energy gained by the electrons from the electric field between collisions, and Λ_1 describes the gain of momentum of the electrons due to the drift and this is structure-dependent. For LXe, $\Lambda_0 = 3.6 \times 10^{-9}$ m [196] and Λ_1 is parameterized as:

$$\Lambda_1 = \frac{\sqrt{2} \times 10^{-7}}{1 + (0.91E \times 10^{-4})^{2/3}}$$

We compare the model (2.27) with the data points of Fig (2.24), for the parameters ϕ_b and n inside the range $\phi_b \in [0.25, 0.45]$ eV and $n \in [1, 50]$. The contour plot of the χ^2 is presented in Fig. 2.30, where we can see that there is a global minimum.

The corresponding best fit point and its associated 90% CI is $\phi_b = 0.341 \pm 0.006$ and $n = 16 \pm 1$. In Fig 2.31 we present the best fit over the data together with the 90% CI. This result is consistent with the value 0.38 eV calculated for ϕ_b if we set z at (2.25) equal, for example, to the lattice constant of solid xenon $z = 6.2 \times 10^{-10}$ m [197]. It is

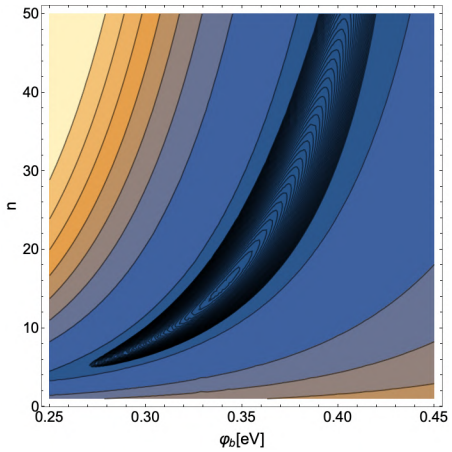


Figure 2.30 – Contour plot of the χ^2 between the data of the extraction yield of Fig. (2.24) and the theoretical model 2.28

also consistent with the result of [195], where this analysis is done using the data of the absolute measurement of the electron extraction yield from [193].

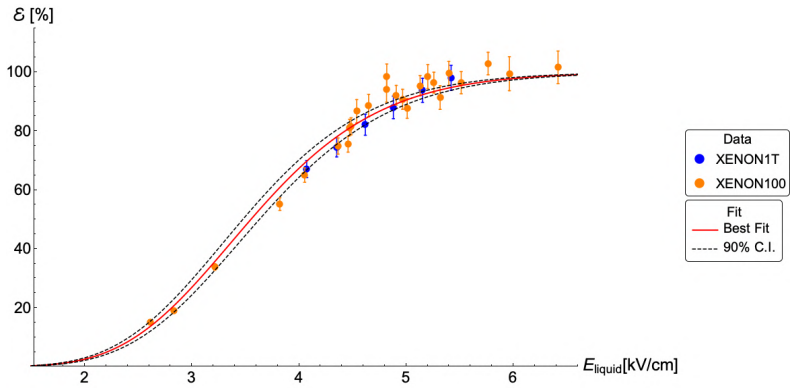


Figure 2.31 – Best fit and 90% C.I. for the "n-th chance" model 2.28. The extraction field here corresponds to the that just below the liquid/gas interface.

Chapter 3

Search for an annual modulation of the electronic recoil background

3.1 Introduction

In addition to the basic experimental approach of direct dark matter detection, which constitutes the main result of the XENON1T experiment [198, 94] *i.e.*, the identification and reduction of any potential background and the statistical inference about the WIMP-nucleon cross section, based on counting observed events in the signal region, there is an alternative direct detection possibility. It is based on the idea of an annual modulation of the expected dark matter event rate, in the energy region of interest. Indeed, such annual modulation is expected due to the movement of the earth around the Sun, as is expected, although much weaker, even a diurnal modulation, due to the rotation of the Earth around its axis.

In Fig. (3.1) we can see a schematic and simplified representation of the above description. Due to the rotation of the galactic disk, the solar system is in relative motion with respect to Milky Way's non-rotating halo of dark matter, a fact which of course, in the first place, opens the possibility to detect dark matter directly. But this same WIMP "wind", due to the additional relative motion of the earth around the Sun, will be perceived by an earth-based detector with an annually modulated strength . The velocity of the dark matter particles, relative to Earth frame, will reach its maximum in June and its minimum in December.

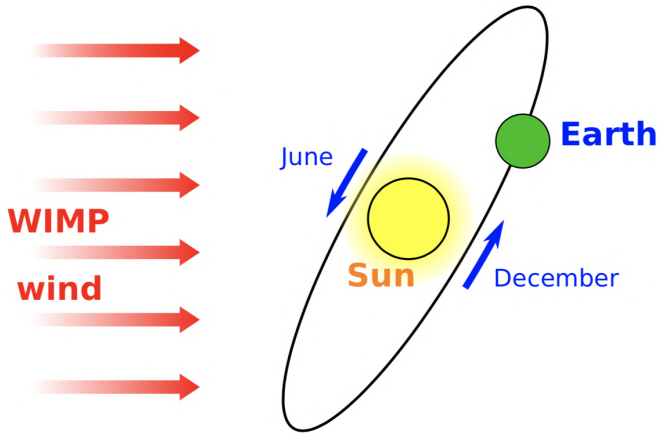


Figure 3.1 – Simplified depiction of the relative motion of the earch in the galactic halo. Figure from [199].

We can calculate the distribution of these relative velocities, as well as their evolution during the solar year, if we use the velocity distribution function of the dark matter particles of the Halo, which is, evidently, always accompanied by specific astrophysical assumptions and corresponding uncertainties. Here we will consider, as is customary in the field of direct dark matter detection experiments, that the dark matter halo, in the region of the solar system, is described by the Standard Halo Model [199, 200], that is, an isothermal sphere. In this case the velocity distribution is Maxwellian and can be described by the following relation:

$$f(\mathbf{v}) = \begin{cases} \frac{1}{N}\left(\frac{3}{2\pi\sigma^2}\right)^{3/2}\left[\exp(-3\mathbf{v}^2/2\sigma^2) - \exp(-3v_{esc}^2/2\sigma^2)\right] & ,\text{for } |\mathbf{v}| < u_{esc} \\ 0, & \text{for } |\mathbf{v}| > u_{esc} \end{cases} \tag{3.1}$$

This Maxwellian is truncated at $\mathbf{v} = u_{\rm esc}$, in order to take into account the fact that the particles with a velocity greater than the escape velocity of the Galaxy gravitational potential wall, will be able to escape the Galactic Halo. This velocity, for the solar neighborhood, is estimated at 498 km/s $< v_{esc} <$ 608 km/s (at 90% C.I.), with a median likelihood of 544 km/s [201]. Also, N is a normalization constant calculated from the requirement $\int_0^{V_{esc}} du^3 f(\mathbf{v}) = 1$.

It should be noted that (3.1) is defined in a reference system in which there is no bulk motion of the dark matter (in the sense that, in this reference system, we have $\int d^3u\mathbf{v}f(\mathbf{v}) = 0$). This reference system, in the case of the Standard Halo Model, is identical to the galactic rest frame. If we want to calculate the velocity distribution of dark matter particles as they are perceived by an Earth-based observer moving, relative to the Sun, with the velocity of the earth $\mathbf{V}_\oplus$, then a Galilean transformation must be

applied to (3.1). Specifically, if $f_{\oplus}(\mathbf{v},t)$ is the dark matter velocity distribution, in the terrestrial reference system at time t, then its relation to (3.1) is:

$$f_{\oplus}(\mathbf{v},t) = f(\mathbf{v}_{\odot} + \mathbf{V}_{\oplus} + \mathbf{v}) \tag{3.2}$$

Here $\mathbf{v}_{\odot}$ is the velocity of the Sun, with respect to the galactic rest frame. This velocity is analyzed as the sum of:

- the velocity $\mathbf{v}_{\text{LSR}}$ of the, so-called, Local Standard of Rest, with respect to the galactic rest frame, (*i.e.* the average velocity of the stars of Milky Way in the local solar neighborhood)[1] and,
- the peculiar velocity of the Sun, $\mathbf{v}_{\odot,\text{pec}}$, in the Local Standard of Rest. This emerges from the fact that the Sun's orbit around the galactic center is not exactly circular, so the Sun moves relative to the Local Standard of Rest.

If we use galactic coordinates[2], then $\mathbf{v}_{\text{LSR}} = (0, 235, 0)$ km/s [203] and $\mathbf{v}_{\odot,\text{pec}} = (11, 12, 7)$ km/s [202, 204].

In (3.2), the velocity $\mathbf{V}_{\oplus}$ is a function of the solar year:

$$\mathbf{V}_{\oplus}(t) = V_{\oplus}[\varepsilon_1 \cos\omega(t-t_1) + \varepsilon_2 \sin\omega(t-t_1)] \tag{3.3}$$

Here the Earth's velocity, with respect to the Sun, was analyzed in the system of galactic coordinates, with $\omega = 2\pi/\text{yr}, V_{\oplus} = 29.8$ km/s. ε_1 and ε_2 is the direction of the velocity at the instants $t = t_1$ and $t = t_1 + \frac{yr}{4}$ respectively [155]. For example, if we choose t_1 as being the Spring equinox then [205]:

$$\varepsilon_1 = (0.9931, 0.1170, -0.01032), \quad \varepsilon_2 = (-0.0670, 0.4927, -0.8676) \tag{3.4}$$

It should be noted that, in (3.3), the time is not measured from the beginning of the solar year but, rather, from the, so called, J2000 *i.e.*, from the noon of Dec 31st 1999 [155]. If time was measured from the beginning of each solar year then this would lead to a systematic annual horizontal shift of the velocity phase by 0.27 days/year, since 2000. Relation (3.3) is a more accurate version of the relation commonly used in astronomy to describe the motion of the earth:

$$\mathbf{v}_e(t) = v_e[-\sin\alpha(t), \cos\alpha(t)\sin\gamma, -\cos\alpha(t)\cos\gamma] \tag{3.5}$$

[1] More precisely the Local Standard of Rest is the rest frame at the local solar neighborhood of a body that would be on a circular orbit in the effective gravitational potential obtained after azimuthally averaging non-axisymmetric features of the actual Galactic potential [202].

[2] The galactic coordinates are Cartesian coordinates where the x-axis point to the direction of the galactic center, the y-axis point to the direction of the rotation of the galactic disc and the z-axis is orthogonal to the galactic plane.

Where $\alpha(t) = 2\pi(t-t_0)/T$, $T = 1$ yr, $t_0 = 153$ days and γ is the angle between the galactic plane and the axis of the ecliptic. Relation (3.5) assumes that the motion of the earth is circular and that the ecliptic lies in the Y-Z plane in galactic coordinates, the motion of the Sun being uniform. This can lead to large errors in calculating the actual velocity of the earth in terms of its X and Z components, although relation (3.3) still ignores the non-uniformity of the motion of the earth but only the non uniformity of the Sun's motion in right ascension.

If we wanted to be even more precise, we had to add to the Galilean transformation (3.2) the velocity that an Earth-based observer has because of the rotation of the earth around its own axis. However, this speed (0.5 km/s) is much lower than the rotational velocity of the earth around the Sun $V_\oplus$ therefore the effect it will have on the final event rate modulation will be subdominant, being even extremely difficult to detect, so we can ignore it[3].

Now, the dark matter event rates show their dependence on velocity (3.3) through the so-called mean-inverse speed:

$$\eta(u_{min},t) = \int_{u>u_{min}} \frac{f(\mathbf{v})}{u} d^3u \tag{3.6}$$

Where u_{min} is the minimum speed for which the dark matter particle has enough kinetic energy to create an observable signal. Specifically, the rate of dark matter events is strictly proportional to (3.6). For example, for the rates of events of nuclear recoils in a dark matter detector:

$$\frac{dR}{dE} = \frac{1}{2m_\chi \mu^2}\sigma(q)\rho_\chi \eta(u_{min}(E,t) \tag{3.7}$$

Substituting (3.2) in (3.6) for the standard halo model (3.1) we can calculate $\eta(u_{min},t)$ for the two extreme values of velocity that appear on June 1st and December 1st, as a function of u_{min}. In Fig. (3.2) we can see these distributions as well as the corresponding distributions $f(\mathbf{v},t)$ for these two dates.

[3] However, it could create a measurable signal in the modulation of the direction of recoils, that could possibly be detected by directional dark matter experiments [206].

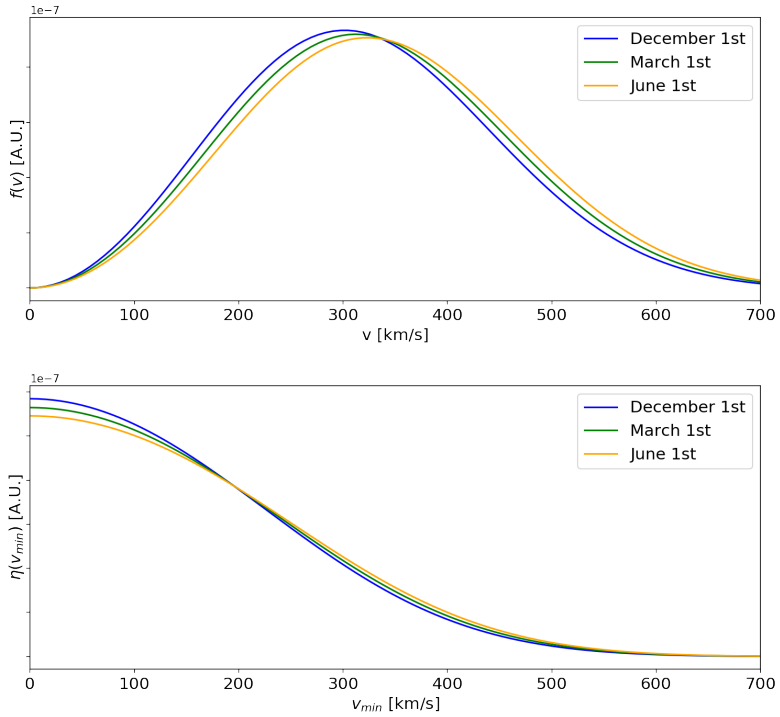


Figure 3.2 – Upper: Velocity distribution of the isothermal sphere representing the galactic halo in the SHM fow three different dates in the solar year. Lower: Same for the mean-inverse speed (3.6) as a function of the minimum velocity u_{min}.

As for (3.7), being a periodic function, with a fixed period of one year, it can be developed in trigonometric Fourier series:

$$\frac{dR}{dE}(u_{min},t) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega t + \sum_{n=1}^{\infty} b_n \sin n\omega t \tag{3.8}$$

We can simplify this relation if we make a time translation $t \to t - t_0$ of this function. If time t_0 is properly chosen so that it coincides with the time of the solar year in which the earth's velocity, with respect to the galactic halo, is maximized, then (3.8) is an even function of time, so if it is developed in Fourier series, the sine factors, b_n, will be zero. Furthermore, if we consider that (3.6) is a smooth and slowly variating function of time, due to the isotropic halo and to the fact that $u_\odot \gg V_\oplus$, then the theory of Fourier analysis shows that its Fourier coefficients a_n should have a rapid decay at infinity. Actually they tend to zero faster than any polynomial at $|1/n|$ [207]. Thus in (3.8) we will have $a_0 \gg a_1 \gg a_{n\geq 2}$ and if we want to capture the time variation of the event rate, for

experimental purposes it suffices to consider only the first cosine term. For the standard halo model and for every practical purposes, (3.8) can then be written as:

$$\frac{dR}{dE}(u_{min}, t) = S_0 + S_m \cos\omega(t - t_0) \tag{3.9}$$

Where the term S_0 corresponds to a time-averaged rate and the term S_m is the modulation amplitude. Here $\omega = 2\pi/1\text{yr}$ and the phase of the modulation is t_0. Thus, a direct detection experiment could attempt to detect a modulation S_m on top of its time constant background rate. Usually, since the rate (3.9) is a function of the energy deposited in the detector, such an experiment, will look for a modulation within an energy region $[E_1, E_2]$, and in this case:

$$\langle S_m \rangle = \frac{1}{E_2 - E_1} \int_{E_1}^{E_2} S_m(E) dE \tag{3.10}$$

In Fig. (3.3) we can see the expected event rates for an elastic scattering of a WIMP of 50 GeV/c^2 off a nucleon of Xe for the two dates, June 1st and December 1st, as a function of the nuclear recoil energy.

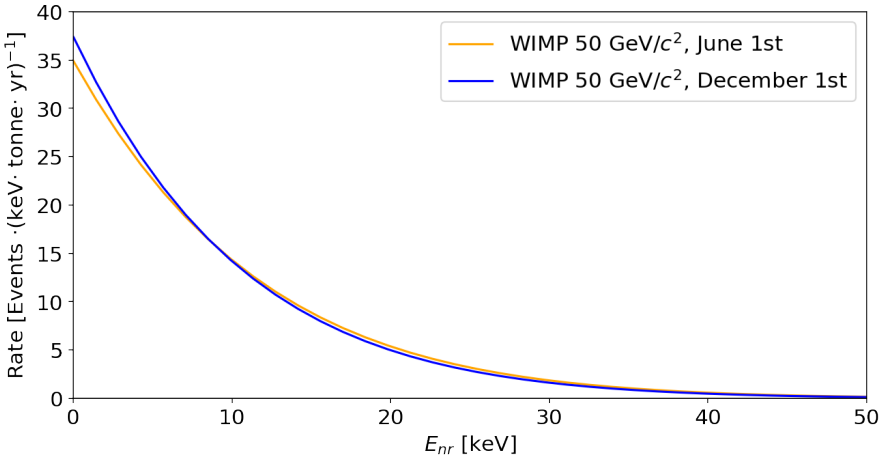


Figure 3.3 – The two dates for which the NR event rate for a WIMP of 50MeV/c^2 reach its extreme spectrums.

We observe that the amplitude of the modulation is relatively small compared to the average expected rate. We will see in a next section that in the case of electronic recoils this modulation fraction can become significantly larger. We also notice a change in the sign of the modulation for small recoil energies, which is consistent with the variation of $\eta(u_{min}, t)$ in Fig. (3.2), where we see that for small values of u_{min} (low recoil energies),

when the velocity of the earth with respect to the galactic halo is minimal, we actually expect a higher rate of events. If we define the amplitude of the modulation as:

$$S_m(E) = \frac{1}{2}\left(\frac{dR}{dE}\Big|_{June} - \frac{dR}{dE}\Big|_{December}\right) \tag{3.11}$$

then we see, from Fig. (3.4), that this amplitude depends on the mass of the WIMP, because this parameter determines the deposited energy and therefore the energy below which there will be a sign change in amplitude. In the direct detection of dark matter such a diagram of the modulation amplitude as a function of the recoil energy can therefore be used to constrain the mass of the WIMP.

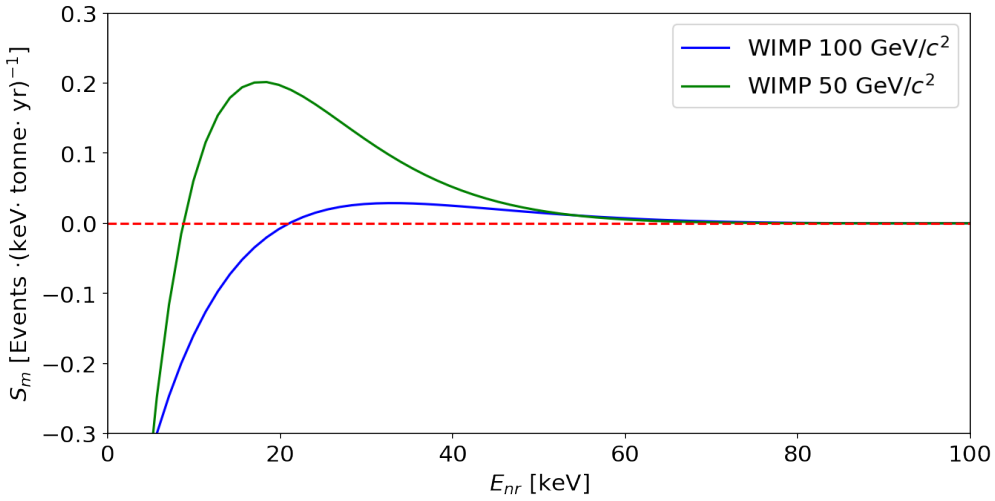


Figure 3.4 – Modulation amplitude S_m as a function of the NR energy E_{nr} for two different masses of WIMP. We can observe here the effect of sign reversal, mentioned in the text, and its dependence on the WIMP mass.

Such an effort becomes more difficult the larger is the background of the detector compared to the amplitude of the modulation, but also to the extent that the detector background is not time constant.

Also, relation (3.9) is based on the simple SHM model, *i.e.* an isothermal sphere with a Maxwelian velocity distribution with a cut off in the high-velocity tail. In a more realistic situation, the dark matter halo also consists of an unvirialized substructure, for example streams and debris flows. Streams consists of material that did not had the time to reach spatial homogeneity, resulting in a distribution with a small velocity dispersion. An example is the Sagittarius stream where it was observed that in this satellite galaxy located on the opposite side of the galactic center relative to the Sun, there is a stream of luminous matter that is pulled by tidal forces and there is evidence that the tail of this stellar material passes, in the form of stream, just a few few kpc away from the solar neighborhood [208]. If the dark matter part of that stream actually passes through the

earth then it can affect the modulation signature, making it very different from the simple form 3.9. Debris flows are structures that are spatially homogeneous, but in momentum space they consist of substructures of velocities created by subhalos that are falling into the Milky Way. In a cosmological simulation of dark matter like Via Lactea [209] it is predicted that for the dark matter of the solar neighborhood, the part of the velocity distribution in the earth's reference system, with $v_{\rm obs} > 450$ km/s consists of such kind of debris flows, with a proportion higher than 50%. Therefore, such flows should be expected for the search for light dark matter where the threshold for a detectable deposited energy may result from the high velocity tail of the velocity distribution. The modulation signal in this region will probably be significantly different from the simple form 3.9. Nonetheless, the SHM can describe well the virialized part of the dark matter halo as well as provide a very useful insight on the modulation signal expected from the earth's motion within it.

3.1.1 The DAMA/NAI and DAMA/LIBRA experiments

For almost a decade, the DAMA experiment measured an annual modulation of the event rate that could potentially be compatible with a WIMP interaction. This experiment was the first to announce the observation of an anomaly in the temporal evolution of the background rate, which was later confirmed by the upgraded phase of the experiment, DAMA / LIBRA. The DAMA detector uses NaI crystals to measure the scintillation light produced by particles that deposit energy in the active volume causing nuclear or electron recoils. Like the XENON1T detector, it is located on the LNGS and initially consisted of 100 kg NaI [210] and was later upgraded to 250 kg [211]. On either side of the crystal, two PMTs detect the scintillation light. The detector is placed inside a low-radioactive copper box in the center of a low-radioactive shield of high-purity copper, lead, and polyethylene and also a 1m of concrete acting as a further neutron moderator.

Data acquisition covers an exposure of 1.04 tons × yr and 1.13 tons × yr for phase I and II respectively and the analysis indicates an annual modulation of the event rate in the energy region (2-6) keV_{ee} at C.L. of 12.9σ. The amplitude of the modulation is $S_m = (0.0103 \pm 0.0008)$ counts/day/kg/keV_{ee} and corresponds to approximately 1% of the total net rate in this energy region. Also the phase of 145 ± 5 days is in the expected region for a dark matter signal according to section (3.1).

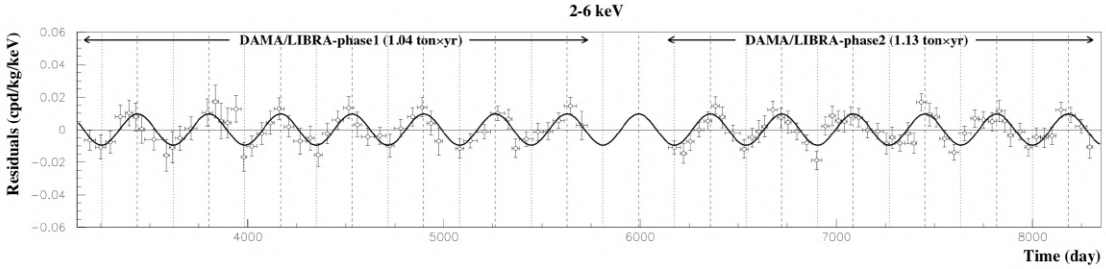


Figure 3.5 – Event rate detected, as a function of time, for the two phases of the DAMA experiment. Solid curve is the fit of a sinusoidal model with parameters constrained to represent WIMP dark matter signal. Image from [212].

The distribution of the modulation amplitude as a function of the deposited energy is shown in Fig. (3.6). There we can clearly see a modulation in the range 2-6 keV$_{ee}$. It should be noted that in this diagram the energy is given in an electron-equivalent energy scale that is related to the energy of the nuclear recoils through a multiplicative factor, the so-called quenching factor.

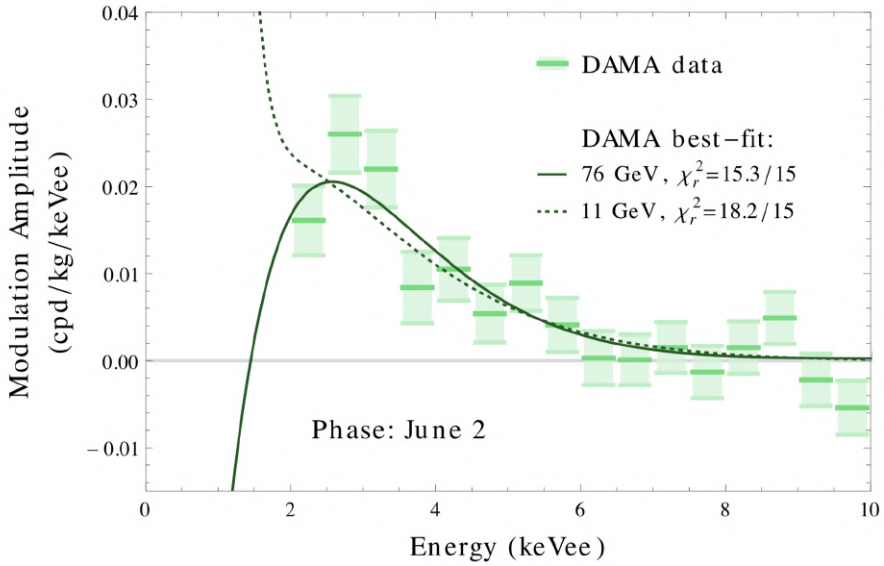


Figure 3.6 – The modulation amplitude observed by the DAMA experiment with two possible masses for a WIMP signal as an interpretation of this apparent modulation. The discrimination between them would be possible with a lower energy threshold were the sign reversal could be verified, see text. Image from [213].

The electron-equivalent of a certain deposited energy is the energy of an electron recoil that would produce the same amount of scintillation or ionization as the amount observed regardless of the actual nature of the event. For example, in the DAMA detector

a recoil event may be due to a Na recoil, an I recoil or an electron recoil, and the detector measuring only the scintillation produced by these recoils cannot distinguish between the three events. And in fact, the two possible nuclear recoils are even characterized by different quenching factors. Therefore the event rates measured by the detector, as well as the modulation amplitude, necessarily contain contributions of all these types of recoils.

Thus in terms of a possible theoretical interpretation of the modulation amplitude of Fig. (3.6), it could be attributed to two different WIMP masses (the signal of which we see in Fig. (3.6) in units keV_{ee}). A WIMP of mass 11 GeV/c^2 in the energy region dominated by sodium-induced nuclear recoils or a WIMP of mass 76 GeV/c^2 that would induce iodine nuclear recoils. Regarding the cross-section for spin-independent WIMP-nucleon scattering, the best fit gives $\sigma_{SI} = 2 \times 10^{-4}$ pb and $\sigma_{SI} = 1.5 \times 10^{-5}$ pb for the two cases respectively.

We see in Fig. (3.6) that these two possible interpretations can be distinguished from the different sign reversal which however occurs in very small energies, in which the detector is not sensitive and in which there was therefore no experimental data. A lower threshold could lead to a distinction between these two interpretations based on the data. Indeed, in phase II of the experiment, the results of which were announced in 2018, the upgraded PMTs of the experiment allowed the reduction of the detection threshold, from 2 to 1 keV_{ee} [212]. This fact demonstrated the incompatibility of both of the above interpretations with the data, however new interpretations are possible as the data now favors models with spin dependent interactions and candidate masses of 10 and 45 GeV [214].

Both the fits of Fig. (3.6) and the new interpretations of phase II of the experiment are in sharp contrast with the null results from other direct detection experiments like XENON100 and XENON1T that rule out a WIMP-nucleon cross section corresponding to the signals of all these interpretations of DAMA by many orders of magnitude.

Another possibility of interpretation is based on the idea that this modulation is due to electron recoils caused by dark matter candidates interacting with the atomic electrons of the target atoms. As explained above, DAMA cannot distinguish the type of recoils in the detector, and in the case of electron recoils the null results from the other direct detection experiments are no longer applicable. With regard to the specific signal detected by DAMA, in the analysis of the XENON100 electron recoil data [102], its interpretations are rejected as due to a leptophilic dark matter with axial-vector coupling to electrons, a kinematically mixed mirror dark dark matter or a luminous dark matter, because the DAMA signal, if interpreted through these models, would produce a signal on the XENON100 detector that could be detected.

However, apart from the results of DAMA, the electron recoil background of detectors like those of the XENON program, although in the context of their main physics search of the WIMP dark matter is perceived as an unwanted background, is also interesting as it could be studied for the search of a possible annual modulation. Precisely because

many well-motivated theoretical models of leptophilic dark matter provide features of electron recoil signatures that, in fact, as we will show in the next section, could lead to larger relative modulation fractions. The rest of this chapter will focus on the study of the electron recoil background of the XENON1T detector and the search for a possible annual modulation.

3.2 Data selection for a modulation search with the XENON1T exposure

In this section we will describe the selection of low energy electron recoil events that will be used in the annual modulation analysis, the various data quality cuts that were used as well as the evaluation of their acceptance as a function of time. The purpose of this analysis, as became apparent from the previous sections, is to collect signals from the energy deposited in the detector in the form of electron recoils. This could be considered as the reverse process of a WIMP search, where the dark matter signal is expected to induce low energy nuclear recoils and thus electron recoils are perceived as an unwanted background. However, in order for this background to be excluded and statistically discriminated, in the parametric space of the experimental observables, from the nuclear recoils that are the candidates of the main nucleophilic search, it is already necessary to well understand and identify the origin of the electron recoils. In the present analysis we therefore use this knowledge of the electron recoil background in order to identify the events that will constitute the population in which we will search for an annual modulation that is not compatible with time stability or even the specific time variation (which, as we will see, exists) of the known ER backgrounds.

In Table (3.1) we present the total number of datasets we have collected and the total lifetime of the events that we will use in our analysis. We will use the background runs of the two scientific runs of XENON1T, (SR0 and SR1) consisting of 7284 datasets each of which has a duration of about one hour with a total lifetime of 300.7 days. Some of the cuts that will be presented in this section affect the data lifetime. After applying all of them, the total lifetime will be reduced at 281.6 days.

Runs and Livetime		
Science Run	**Number of datasets**	**Livetime**
SR0	901	32.705 days
SR1	6383	248.851 days
SR0 and SR1	7284	281.556 days

Table 3.1 – Number of datasets used for each Scientific Run of the XENON1T experiment and the corresponding livetime, after applying all cuts.

In this section, after detailing the data selection cuts, we will present the distribution of the selected events in a series of parametric regions of the experiment for two energy ranges [1-10] keV and [10-20] keV, as well as the calculation of the time evolution of the event rates. This set of data, as well as the rate of events that will be calculated, will be the input data of section (3.4) where a necessary statistical analysis of the stability of the detector will be performed, considering the physical parameters that could affect the signal generation. The hypothesis of the correlation of the time evolution of these electron recoil event rates with the temporal evolution of such a parameters will be investigated in order to exclude the hypothesis that the temporal evolution of the former can be influenced by possible unstable experimental conditions, which could undermine the statistical inference about a time modulation based only on the understanding the physical background.

3.2.1 The combined energy scale and its resolution

The energy deposited in the detector in the form of an electron recoil is almost completely transformed into primary scintillation photons n_{ph} and ionization electrons n_e. From the corresponding signals S1 and S2 that will be generated we can calculate the energy E_d of the interaction, through the relation:

$$E_d = (n_{ph} + n_e)W = \Big(\frac{cS1}{g_1} + \frac{cS2_b}{g_2}\Big)W \tag{3.12}$$

Where cS_1 and $cS2_b$ are the signals S1 and $S2_b$ corrected for light collection efficiency and drift time and other position effects concerning the extraction efficiency such as those we analysed in chapter (2), $W = 13.7 \pm 0.2$ eV is the average energy that is required to be deposited in the LXe in order to produce a scintillation photon or an ionization electron [148]. The parameter g_1 is the average number of photoelectrons produced per primary scintillation photon while the parameter g_2 corresponds to the secondary scintillation gain of chapter (2) but only in terms of the number of photoelectrons detected by the bottom PMT array per extracted electron. Here again, only the part of the secondary scintillation light that is observed by the bottom array is used in order to avoid saturation effects from the PMTs of the top array. The parameters g_1 and g_2 can be determined and therefore a combined energy scale can be defined via relation (3.12) if a number of monoenergetic sources, of known energy, are used in several calibration runs outside the detector or injected into the target volume. So for example in Fig. (3.7) we see such a fit using the 41.5 keV line from ^{83m}Kr, the ^{129m}Xe and ^{131m}Xe metastable isotopes activated during a calibration using a neutron generator outside the TPC, or even γ rays from ^{40}K or ^{60}Co from material radioactivity using the outer region of the active target. The peaks produced by those lines are selected with 3 sigma 2D gaussian fits. The amplitudes are divided by the corresponding energies, obtaining the Light Yield ($LY = S1/E$) and the

Charge Yield ($QY = S2/E$). This results in the points of Fig. (3.7) and then a linear fit determines the parameter values:

$$\begin{aligned} g_1 &= 0.1426 \pm 0.0001 \text{ (stat)} \pm 0.0017 \text{ (syst) PE/ph} \\ g_2 &= 11.55 \pm 0.01 \text{ (stat)} \pm 0.24 \text{ (syst) PE/e} \end{aligned} \tag{3.13}$$

Here the systematic errors are dominated by variations in the volume used for the evaluation inside the TPC.

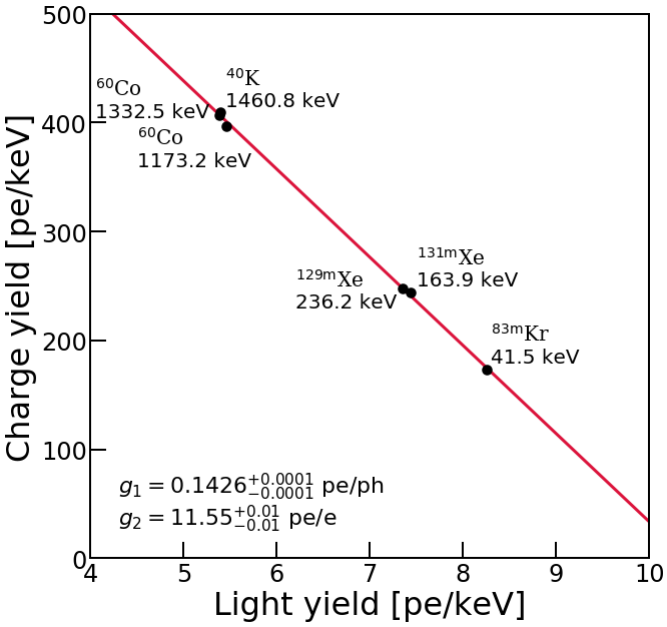


Figure 3.7 – Diagram of the light yield vs the charge yield for various monoenergetic lines used in order to calibrate the detector's responce to different energies. With a linear fit with the function (3.12) a so-called combined energy scale can be defined covering an energy range of three orders of magnitude wide.

The energy resolution is measured by the width σ_E of a monoenergetic peak divided by its mean value E. Its energy dependency can be parameterized as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} + b \tag{3.14}$$

Fig. (3.8) shows the fit of the model (3.14) on the data points corresponding to the single energy lines of ^{83m}Kr (41.5 keV), ^{131m}Xe (136.9 keV), ^{125}Xe (221.6 keV), ^{129m}Xe (236.1 keV), ^{125}Xe (276.6 keV), ^{214}Pb (351.9 keV), ^{208}Tl (510.8keV), ^{214}Bi (609.3 keV), ^{228}Ac (968.0 keV), ^{214}Bi (1120.3 keV), ^{60}Co (1173.2 and 1332.5 keV), ^{40}K (1460.8 keV), ^{214}Bi (1764.5 and 2204.1 keV), ^{208}Tl (2614.5 keV).

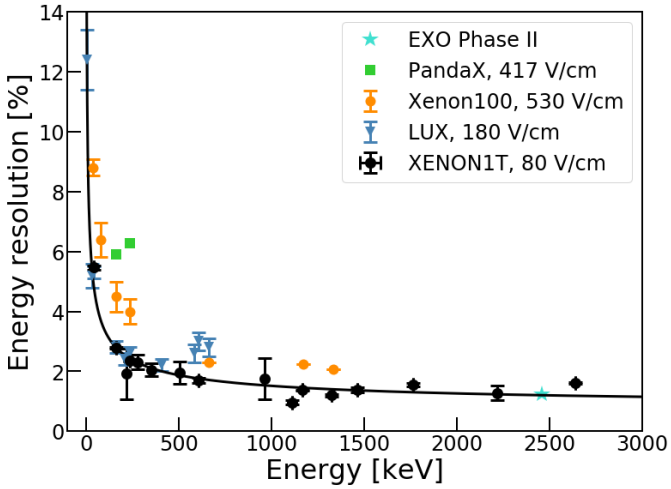


Figure 3.8 – The width of the peaks from energy deposits by various monoenergetic lines can be used in order to derive the energy resolution of the detector as a function of the deposited energy. Here the black points for XENON1T correspond to the isotopes listed in the text, while also shown are the energy resolutions for EXO Phase II, PandaX, XENON100 and LUX with the corresponding drift fields stated.

The following values for the parameters of (3.14) are obtained from this fit:

$$a = 30.98 \pm 0.43\sqrt{\text{keV}}$$
$$b = 0.37 \pm 0.03$$

From Fig. (3.7) we observe a linearity that covers the energy region of keV up to the region of MeV which, based on (3.12) define an energy scale which we can use in order to determine the energy deposition of electron recoils, with the resolution of Fig. (3.8).

The list of the analysis selection cuts

In order to select that set of events that correspond most likely to actual deposits of energy through electron recoils, our understanding of the nature of the events observed in the detector leads us to the following data selection:

1) Fiducial Volume The fiducial volume is a cylinder defined as:

$$\begin{aligned} -92.9 \text{ cm} < z < -9 \text{ cm} \\ \sqrt{x^2 + y^2} < R_{cut} \end{aligned} \tag{3.15}$$

Regarding the depth, the 9 cm cut is chosen in order to exclude events that are expected near the liquid surface due to the intrinsic radioactivity of PMTs and the other

materials. The lower z cut also aims to eliminate events that originate from the anode where, due to the fringe electric field, there may be a partial charge loss or a different recombination rate which would cause an alteration of the distribution of the ER in the (cs1, cs2) space. R_{cut} determines the distance of the surface of the fiducial cylinder (3.15) from the boundary of the TPC and can be selected based on the understanding of the number of events induced from the material radioactivity and in the WIMP search is selected so as to minimize this number as determined by MC simulations in combination with the need to maximize the mass that ensures a greater exposure. Fig. (3.9) shows the MC simulation of the background ER rate in the energy range 1-12 keV, as a function of the fiducial mass (and therefore of the R_{cut} radius) where we see that the component due to material radioactivity is the one that is dominant for large radii and for fiducial volumes approaching the TPC boundary.

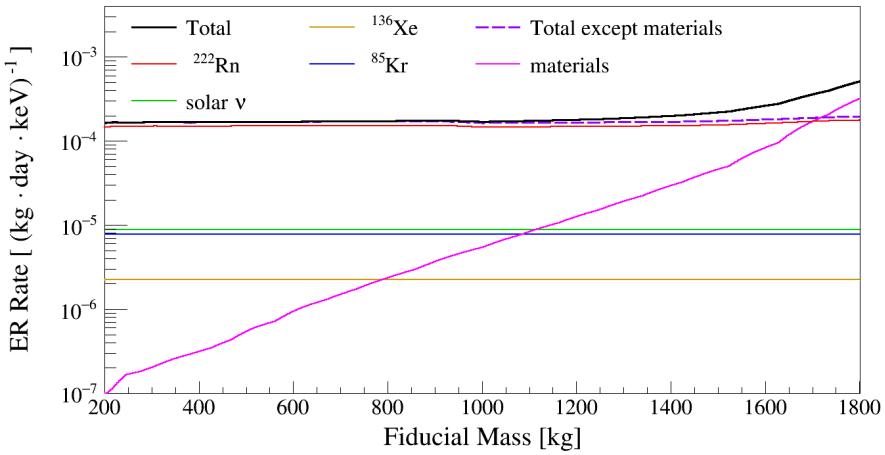


Figure 3.9 – Monte Carlo simulation of the total electronic recoil background as a function of the fiducial mass together with the various contributions in the energy range 1-12 keV. Figure from [190].

Also events generated very close to the TPC wall tend to produce ionization signals which are not fully collected as many ionization electrons are absorbed by the wall charging it. This again results in the deformation of the ER event band in the parametric space of the observables. In SR0 and SR1 two R_{cut} options were made which are well understood in terms of the ER background inside the FV they define: $R_{\text{cut}} = 36.94$ cm and 41.26 cm offering an FV of 1 ton and 1.3 ton respectively. In the present analysis we will choose the 1.3 ton FV which, for the livetime of SR1, corresponds to an exposure of 1 tonne $\times$ yr. The exact calculation of the mass contained within this fiducial volume can be done in two different ways. First by taking into account the density of the LXe and calculating the mass contained in the FV and, second, by using ^{83m}Kr events from corresponding calibration runs, which are expected to be uniformly distributed inside the target mass of

the TPC. By calculating the ratio of the number of events located within the FV on the total number of events, and multiplying with the total mass inside the TPC, we have an alternative estimate of the mass contained within it. The average of these two values, with their difference as a systematic error gives the value of 1.3 ± 0.01 t.

2) S1 and S2 selection. In the waveform of the selected events the S1 and S2 peak of the main event must have the highest intensity of all the S1 and S2 peaks recorded in the waveform.

3) Width of main S2 The width of the S2 signal is related to the depth of the interaction because of the diffusion of the electrons cloud drifting towards the top of the TPC. This relation can be used to remove events whose width is much larger than we can expect from their drift time, such as S2s coming from events in the GXe or accidental coincidences events. Indeed, following [215], the width of a single S2 signal coming from a nuclear or electron recoil (*i.e.* not a single electron like those small S2 signals we dealt with specifically in chapter (2)), can be expressed as:

$$\sigma_{S2} = \sqrt{\sigma_t^2 + \sigma_0^2} \tag{3.16}$$

Where σ_t is the width of the ionisation electron cloud after a drift time t, and σ_0 is the width that will result from the secondary scintillation of each electron of the cloud that will be extracted into the GXe. σ_t, expressed in units of time and corresponding to that time period Δt containing 50% of the area of the S2 peak, is related to the spatial longitudinal width σ_z of the electron cloud through $\sigma_t = \sigma_z / v_d$ where v_d is the drift velocity[4]. If we consider the initial distribution of n_0 electrons in the cloud as point-like, after a time t the distribution can be written as [216]:

$$n = \frac{n_0}{\sqrt{4\pi D_L t}} \exp\left(-\frac{(z - v_d t)^2}{4 D_L t}\right) \tag{3.17}$$

Where D_L is the diffusion constant, 25.26 cm^2/s for SR1. We can then derive σ_z as the second central moment of $n(z,t)$ over z.:

$$\sigma_z = \sqrt{\frac{1}{n_0} \int_{-\infty}^{\infty} n(z) z^2 dz} = \sqrt{2 D_L t} \tag{3.18}$$

Based on this model we can make a fit on the data of S2-width as a function of the drift time, using a relation of the form $\sigma_{S2} = \sqrt{P_0 + P_1 t}$.

[4] v_d is related to the electric field by the relation $v_d = \mu E$ where μ is the electron mobility.

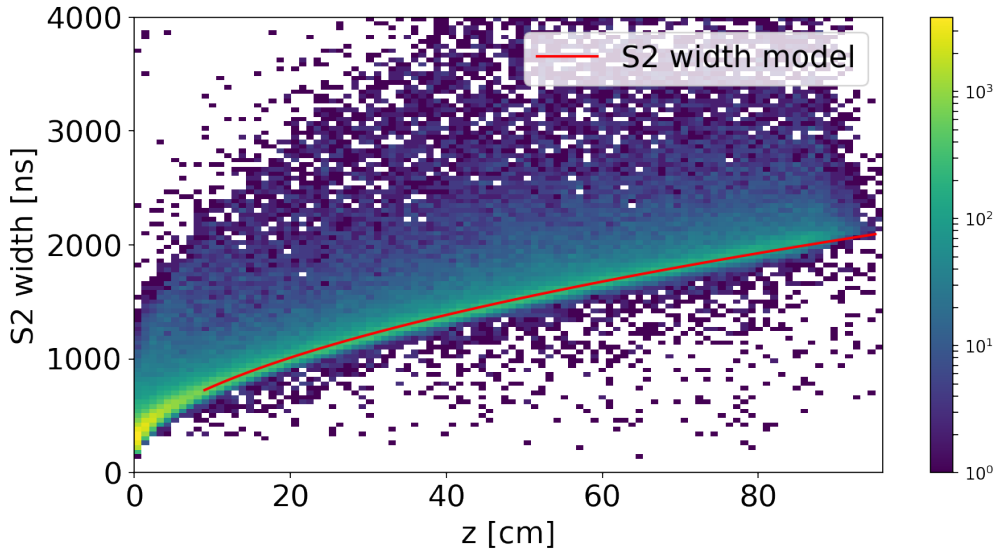


Figure 3.10 – Width of the main S2 signal over drift length (related to drift time by $z = v_d t$) using neutrons induced by AmBe calibrations performed during SR1. The equation used for the fit uses the relation $\sigma_{S2} = \sqrt{P_0 + P_1 t}$ constituting the model presented in the text.

The result of this fit, for SR1, is shown in Fig. (3.10), compared to data from several AmBe calibration runs. Then the cut that will exclude those events whose S2-width is not compatible with the diffusion model, is tuned based on a 2σ band of the ratio of the observed S2-width of calibration datasets over the expected S2-width from the diffusion model. In fig (3.11) we present the cuts boundary on AmBe calibration data.

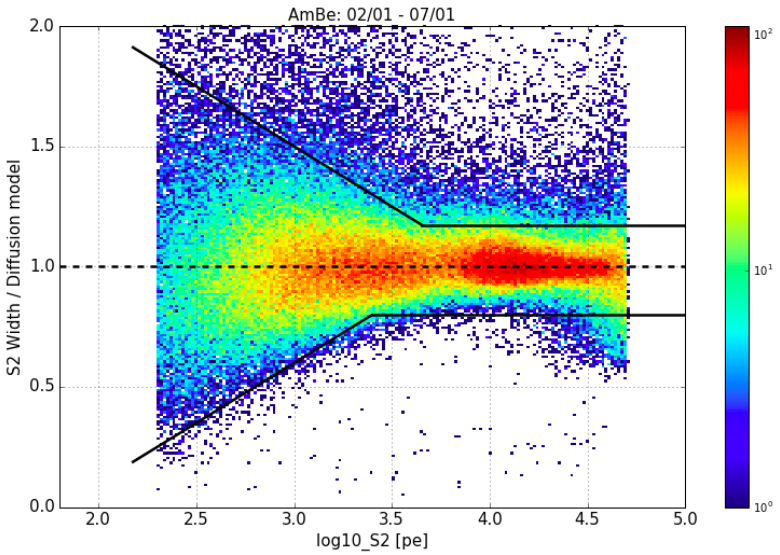


Figure 3.11 – Boundaries of the cut based on the S2-width model. The parameter space here is the ratio of the observed S2-width on the width expected by the drift time of the event versus the S2 area. Data shown are from AmBe calibration.

4) Fraction of light observed by top PMTs Another way to remove gas events is through the percentage of light observed by the top PMT, a parameter called Area Fraction Top (AFT). Gaseous events are characterized by an unusually high percentage of AFT as compared to events that come from a real energy deposit. Using ^{220}Rn calibration datasets we can tune a cut that selects those events that belong to the $[5, 95]\%$ quantiles of AFT distribution as a function of S2. Since no cut can be considered absolutely efficient, cuts 3 and 4 are complementary to the process of removing gaseous events.

5) Double scattering rejection The signal we are looking for consists of a unique true interaction inside the detector, of an electron recoil nature, which will be characterized by a unique pair of S1 and S2, as it is statistically unlikely that a dark matter particle will interact twice during its passage through the active volume of the detector. For this purpose we must make sure that:

- The event was a single scatter *i.e.*, the S2-signal in the waveform is unique. A particle that deposits energy through multiple scatterings inside the detector, such as a neutron, produce events that are characterised by one S1, as the primary scintillation light is observed simultaneously by the PMTs, while in waveform multiple S2s of similar intensity are present, corresponding to different ionization electron clouds produced by several diffusion points of a particle. One way to exclude these events is to require that the second largest S2 signal in the waveform of an event should lie in the region of single electrons *i.e.*, any other S2 detected should be attributed

to the mechanisms described in the previous chapter and not to the deposition of energy of a particle in the active volume.

- For the same reason we must reject those events that might be caused by two particles interacting on the same time window. The way in which the data processor software detects the possible interactions that took place in an event is through pairing the S2s with the S1s that are detected starting with the S1 that has the largest number of coincident channels that are contributing to its sum waveform[5], and the largest S2. Then the search for other possible interactions continues with the pairing of this S2 and the S1 with the next largest number of coincident channel, and so on. Therefore any event that have a second S1, which could be combined with the main S2 and constitute a true interaction should be removed. Such a cut can be made if, in the waveform of each event, we search for the existence of other S1s besides the main one that have been detected by the peak finding algorithm, in the way we described the pairing of S1 and S2 peak above. If such an alternative interaction has been detected by the data processor, then we retrieve the information of the z coordinate of that alternative S1. Then we calculate, as we did in cut 4, the S2-width predicted by the diffusion model based on the depth of this candidate interaction. If this lies within the bounds of the cut of Fig. (3.11), then the event should be excluded, as we cannot be sure about which of the two S1s should be combined with the main S2, therefore we cannot be sure about the uniqueness of the interaction.

6) Expected shape of S1 and S2 peaks Another way to exclude double scatter events as well as poorly reconstructed events, complementary to cut 5, is to evaluate a goodness of fit of the reconstructed position, based on the goodness of fit of the top hitpattern compared to an expected hitpattern from a MC simulation of the top array hitpatterns, using a per-PMT map of the light collection efficiency. The test statistic based on this goodness of fit is a poisson likelihood chi-square and could be used for the test of a null hypothesis of the kind: single scatter with one cluster of hits in the top hit pattern array. Such a cut is defined by the 98% percentile of the 2 dimensional distribution of this test statistics (S2 pattern likelihood) with the S2 area of an event. However, this approach does not take into account the energy dependence of this cut due to PMT saturation effects at higher energies but also, for the low energies involved in this study, it was observed that this cut leads to a gradual drop of acceptance with time through SR1 as is shown by an evaluation of this acceptance on electron recoils from ^{220}Rn calibration runs in Fig. (3.12). This effect can be attributed to an increasing afterpulse rate of some PMTs of the top array during SR1 due to degradation of these photomultipliers over time.

[5] specifically, with the larger number of PMTs contributing within a temporal window of 50 ns centered on its sum waveform maximum

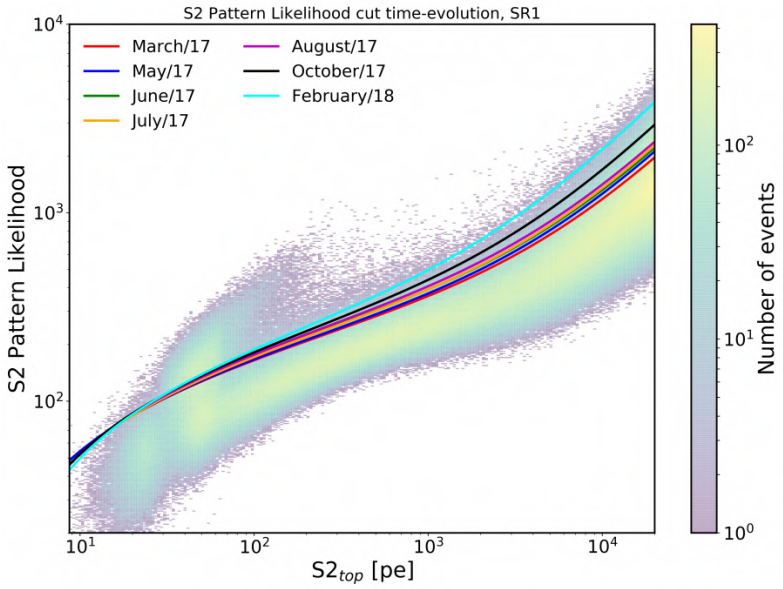


Figure 3.12 – The boundary of the cut removing poorly reconstructed S2s based on the S2-pattern likelihood test statistic (see text) as a function of time during SR1. Is clearly seen a time dependence extended throughout SR1 that is found to be due to the degradation of the performance of some PMTs over time.

One of the causes of these PMT afterpulses is attributed to the residual gas molecules in the vacuum of the PMT. A photoelectron accelerated from the photocathode to the first dynode could ionize one of these molecules, whose ion can then accelerate towards the photocathode and create an afterpulse. From the measured afterpulse time, which depends on the ion's mass to charge ratio, ions such as H^+, He^+, CH_4^+, N_2^+ etc. have been identified as some of the causes of afterpulses for the Hamamatsu R11410-21 photomultipliers used in the XENON1T TPC [217].

It was therefore necessary, for the purposes of the present analysis, to redesign this specific cut so as to remove from the calculation of the top pattern likelihood test statistic those PMTs that show an increasing rate of afterpulses during SR1. Indeed this redefinition of the test statistic allowed the reduction of the acceptance's dependence on time.

A similar cut can be defined also for the S1 signals where a likelihood can be constructed based on the observed hitpattern and the one expected through a MC simulation using a S1 relative light collection efficiency, per-PMT, map. A cut based on this test statistic can exclude a portion of ERs that are not placed inside the ER band in the parameter space of Fig (3.15), for example events due to accidental coincidences of lone S1 and lone S2 signals that are incorrectly perceived by the data processor as a real interaction. Again the optimization and evaluation of the acceptance of such a cut, is done using single scatter events in the ER-band for ^{220}Rn data.

7) Area of the most contributing PMT It has been observed that events may have a specific PMT that is contributing in a very large proportion to the overall S1 signal. During SR0 and SR1 a cut was developed which examines the area of the most contributing channel. This cut was tuned and its acceptance was evaluated on ^{220}Rn calibration data in order to reject this kind of events. This cut works in addition to the cut 6 as such events usually come from accidental coincidences that can be created by light emission, after-pulses or scintillation in the photocathode of a PMT.

8) During SR0 and SR1 a contamination by ^{83m}Kr events was observed, resulting presumably from an amount of ^{83}Rb leaked inside the detector, after a calibration during SR0. We can observe this contamination from the unexpected presence of a strong peak in the energy region of 41.5 keV. Making an elliptical fit of this peak in the parametric space (cS1, cS2) we can measure the number of events that are found inside a 3σ selection. In Fig. (3.13) we present this event count, for a time bin of 3 days, that covers the whole SR1, with a selection of those datasets that are at least one day away from the end of each ^{83m}Kr calibration period, in order to take into account only those events that actually come from the ^{83m}Kr that is the product of the quantity of ^{83}Rb that contaminated the interior of the TPC. If we fit the data with an exponential model we infer a half-life of $86.42^{+2.1}_{-1.9}$ days, value that is consistent with the ^{83}Rb half-life of 86.2 days [218].

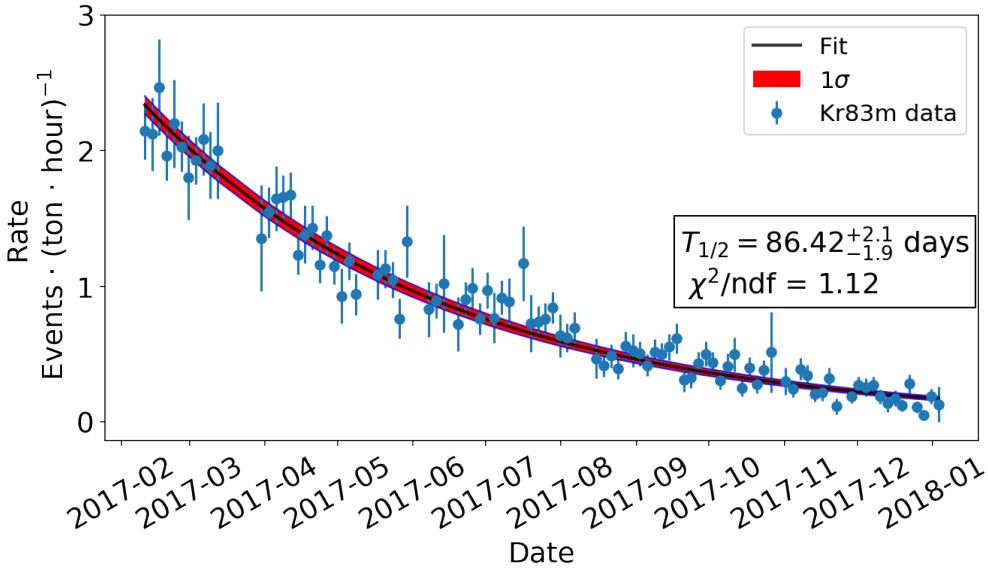


Figure 3.13 – Evolution of the rate of events found in the region of 41.5 keV in background datasets, during SR1. Here this rate is not corrected with the acceptance of the selection cuts used to obtain them, because what is of interest here is the fact that an exponential fit gives a characteristic time consistent with the half-life of ^{83}Rb. This points toward an interpretation of the presence of these events in background datasets as a contamination of the TPC during a certain krypton calibration, probably situated in SR0.

A specific problem arises for our analysis due to this type of contamination. In Fig. (3.14, blue histogram) we present the ^{83m}Kr spectrum in the combined energy scale (CES) using 20 ^{83m}Kr calibration datasets of the SR1, in which we have applied all the cuts we presented in this section up to this point, except those that exclude S1 and S2 multiple scatters. We notice the appearance of four peaks. We commented on the two peaks that are centered at a CES of 32 and 42 keV (250 and 350 PE) in chapter (2). However, of interest here are the two peaks appearing in the energy region of [9, 20] keV. Although, a priori, the energy region of interest for the annual modulation is [1-6] keV, as for a dark matter particle of mass $\mathcal{O}$(10 GeV), this is the region in which we expect the largest modulation amplitude, as we saw in the previous section, however we want to use the energy region of [10-20] keV as a control region for our analysis, so the presence of this additional background, which is also variable in time (due to the decay of the parent ^{83}Rb isotope), can be considered problematic.

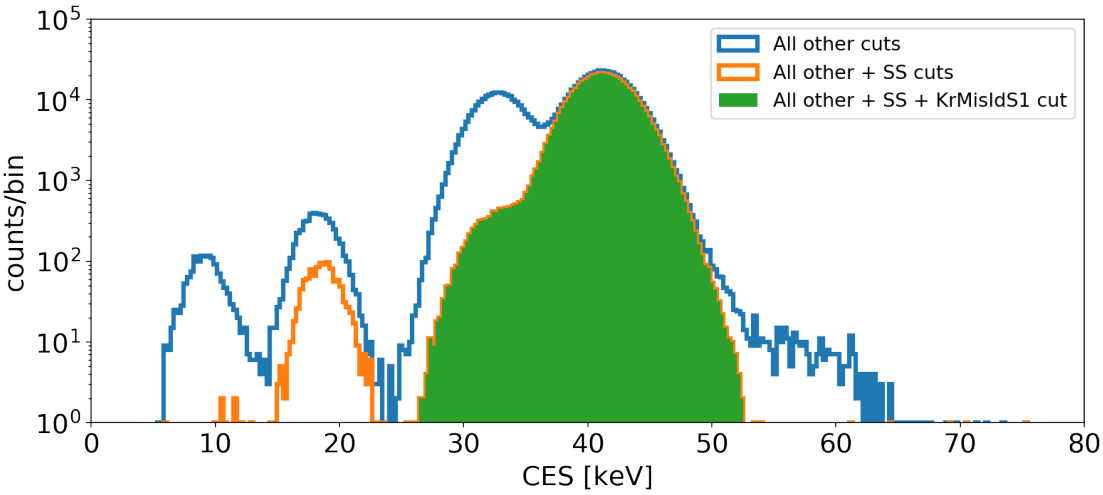


Figure 3.14 – Distribution of events in the CES for 20 ^{83m}Kr calibration datasets of the SR1. Blue: All cuts applied except the single scatter cuts. Orange: All cuts applied including the single scatter cuts. A large portion of events remains in the region of $\sim$ 18 keV. Green: Application of the KrMisIdS1 cut (see text) removes this population.

The peak around 9 keV in Fig (3.14) is composed of events where the two S1 signals, corresponding to the consecutive electron captures of ^{83m}Kr, are merged (*i.e.* perceived by the data processor as one) but they are also misclassified as being an S2. Then, the main S2 of the event is probably paired with an PMT afterpulse, being identified as the main S1 of this spurious interaction.

The peak centered around 18 keV is composed of a population of events where the two S1s, of the ^{83m}Kr decays, are not merged, and the S1 corresponding to the transition

of 32 keV is misclassified as an S2, while the other S1, of the 9 keV transition, is correctly classified as an S1 being paired with the main S2 of the event.

Concerning the population of the first peak, it is obvious that, since the merged S1s are misclassified as an S2, the event will have in its waveform a second S2, beside the main one, of the order of 350 PE, so it could be excluded by the single scatter cut which we mentioned above. Indeed if we apply this cut we see that this population disappears (orange histogram of Fig. (3.14)). However, in the second peak, the largest other S2, beside the main one, is the misclassified S1 of 32 keV transition which corresponds to approximately 250 PE. This area is quite small and the single scatter cut will not be efficient enough to completely eliminate this population. As can be seen in Fig. (3.14) a large fraction of this population will pass the single scatter cut. Therefore, an alternative cut has been developed to deal with this population of misclassified S1s from krypton events which can pass the single scatter cut, taking advantage of the fact that, in the case of this population, the misclassified S1 of the 32 keV transition will appear as an S2 that will precede the main S1 (corresponding to the transition of the 9 keV which is correctly classified). Therefore, the cut that in Fig. (3.14) we call KrMisIdS1 requires the exclusion of events that have a second larger S2, after the main S2, with an area larger than 100 PE and at a time distance of up to $3\mu s$ before the main S1. Indeed, in Fig. (3.14, green histogram) we observe the elimination of the rest of the population of the second peak.

The selection of the data is completed through the rejection of those events that, in the parametric space (cS1, $cS2_b$), are distributed outside the 3σ band of the electron recoils (ER) as in Fig. (3.15), in order to exclude any event that may result from the radioactive contaminants of the materials and which may be characterised by a partial charge loss, as we explained above, resulting in electron recoil events that are systematically distributed lower than the ER band. This band is determined by the detector response to electron recoils that is modelled with a MCMC fit to β-decays of ^{212}Pb selected from ^{220}Rn calibration data [122]. Using the ER model template that result from this fit, we select those events that are distributed within the 99.7 percentile of the band.

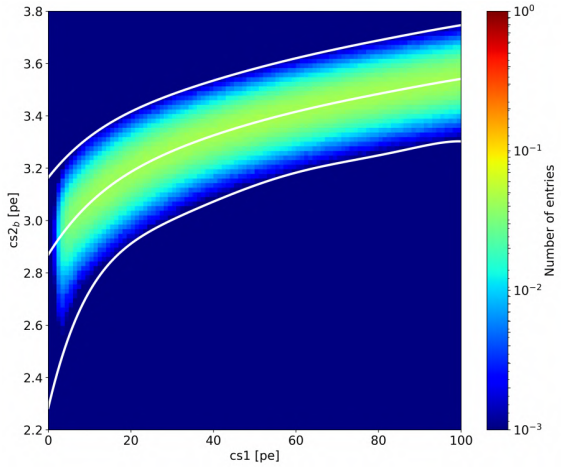


Figure 3.15 – ER 3σ band fit from template models that have been used in SR0 and SR1.

Finally Figures (3.16) and (3.17) In the following figures we present the distribution of the selected ER events, in various parametric spaces, as well as the evaluated electron recoil event rates, in time bins of 7, 15 and 21 days, in the energy range of [1-10] keV of the CES and for a fiducial volume selection of 1.3 tons.

The selected data and evaluated event rates will be used in section (3.4) to study possible correlations with slow control parameters in various energy ranges and choices of fiducial volume.

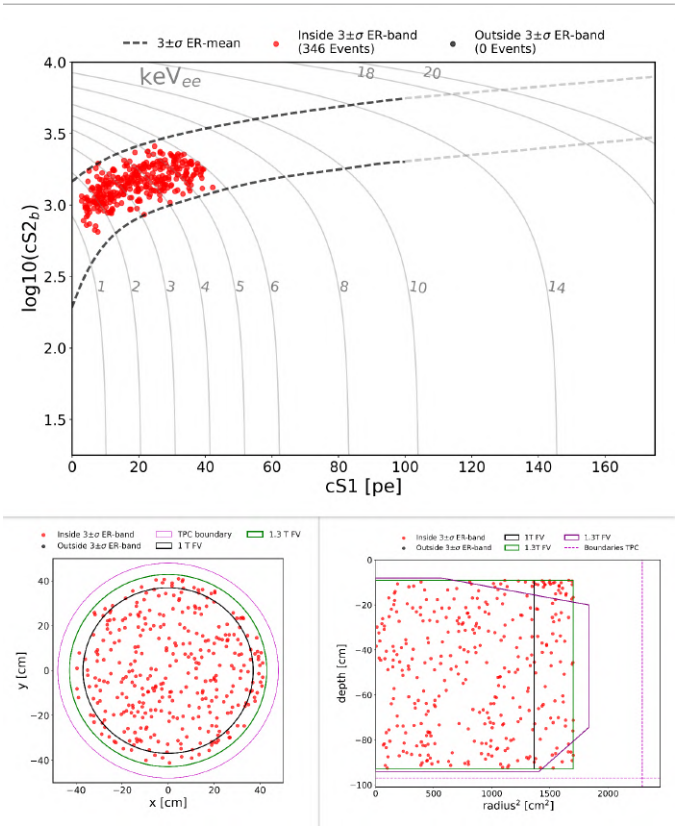


Figure 3.16 – Selected events in the energy region of 1-6 keV, inside the 1.3 tonnes fiducial volume, for SR1. Upper figure: Distribution of the selected events in the parametric space $cS2_b$ vs cS1, inside the 3σ ER band. Lower left: distribution of events in the x-y space where the various fiducial volumes are shown. Lower right, same distribution in the space z vs R^2_{TPC} with the regions of the same fiducial volumes shown.

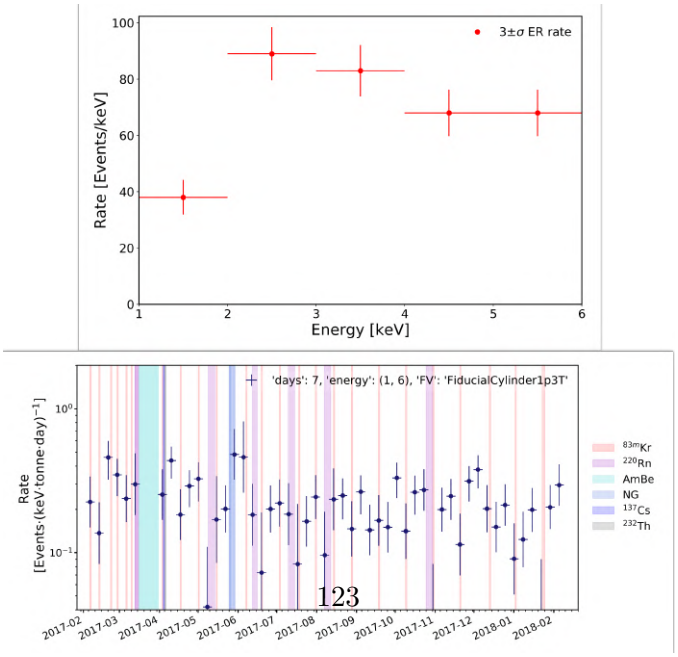


Figure 3.17 – Event rates for the events selected in (3.16). Upper: Event rates as a function of the deposited energy. Lower: Time evolution of the event rates for time bins of 7 days. Also shown is the period of the various calibrations during SR1.

Evaluation of the cut acceptances

In order to have a accurate evaluation of the event rates, as they are evaluated in the binned plots of Fig. (3.17), but mainly for the final statistical inference about a possible modulation of the background rate, we must also estimate the overall acceptance of the cuts we used to select the ER events. The acceptance of these cuts, for the selection of signal like electron recoil events, has been also studied in the context of other analyses, but for our purposes it is necessary to estimate the evolution of the acceptance of each cut used, as a function of time and as a function of the combined energy scale, and then the combined acceptance of all cuts in time and energy. The ideal case would be to make such an estimate using simulation, however it is extremely difficult and certainly accompanied by large uncertainties, an attempt to estimate exactly how the parameters of the simulation would evolve with time. Thus here, as well as in analyses independent of time, we resort to a data driven method. One approach, that is widely used, is the so-called "N-1" method. Specifically, we can consider that the cuts presented above constitutes a series of selections in the parametric space of the observables, which converges on the determination of a "good" sample of events, in the sense that these are signal like (*i.e.* single scatter electronic recoils energy deposition inside the fiducial volume). If we use ^{220}Rn calibration datasets to look at the selection of ER that are induced from the ^{212}Pb β-decay, in the decay chain of ^{220}Rn, then we can make an estimate of the acceptance of a specific cut, from a list of cuts, if we consider that the application of all other cuts except the one under consideration, has already led to a signal like sample of events. If we, then, apply the cut under consideration to this set of events, the further exclusion of some events can be considered as an acceptance loss due to this particular cut. Specifically, the following constitutes the logic of this "N-1" method. Let E^i be the whole set of available events. Then, for a specific data selection cut C, let A and B be the following sets:

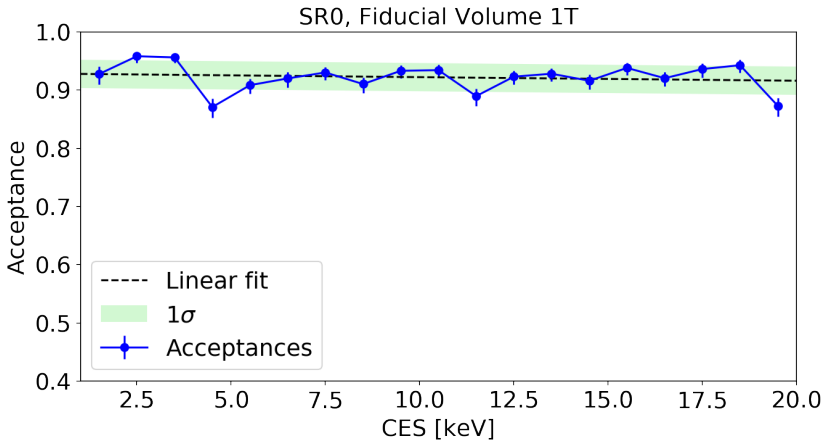
$$A = \{E^i | \text{event i passes all cuts other than C}\}$$
$$B = A \cap \{E^i | \text{event i passes C}\}$$

Then this method defines the acceptance of the cut C as the ratio:

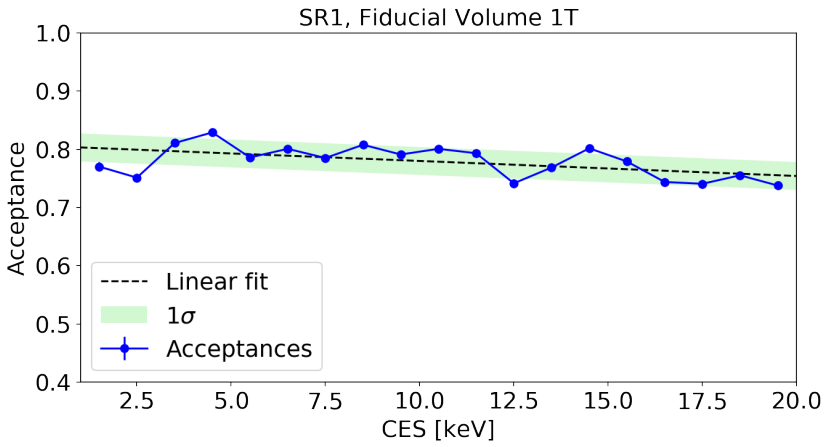
$$\text{acc}_C = \frac{\#\text{ of events in B}}{\#\text{ of events in A}} \tag{3.19}$$

Of course, there are some cuts among the ones we presented above, whose acceptance cannot be calculated with this method, such as the selection of the fiducial volume as this can be considered exactly 1, as expected by all cuts that affect the exposure only. That is because, if an event outside the fiducial volume is genuinely "bad" then by definition it's excluded but if it is a "good" event, its exclusion is accompanied by a corresponding reduction of the exposure.

In the following figures (3.18) we present the cumulative acceptance of the cuts we described in this section as a function of energy in the energy range [1, 20] keV, for 1T fiducial volume, for SR0 and SR1 respectively.



(a) a



(b) b

Figure 3.18 – Cumulative acceptance of the cuts used as a function of the energy for SR0 (up) and SR1 (down). Also is shown a linear fit and the 1σ band.

Fig. (3.19) shows the cumulative acceptance as a function of time, for SR1 calculated in 9 time bins. The corresponding plot for SR0 is not presented, as its lifetime of 32.7 days will be considered as an additional point in Fig (3.19), in the combined final analysis of SR0 and SR1.

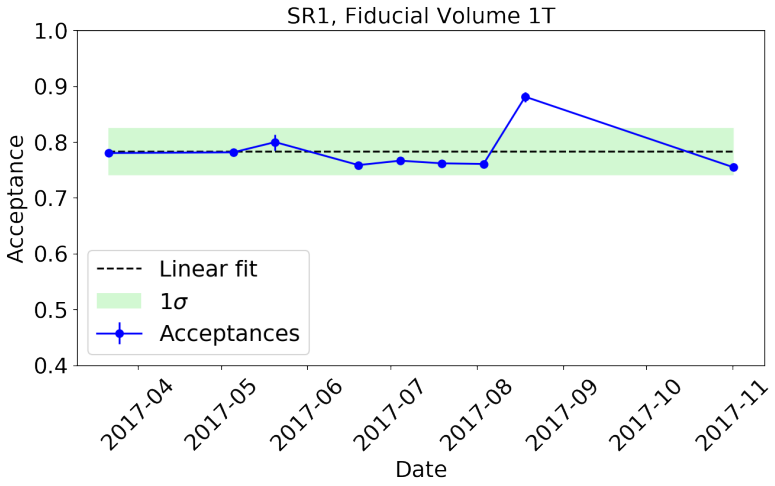


Figure 3.19 – Cumulative acceptance of the cuts used as a function of time for SR1, together with a linear fit and a 1σ band.

The errors of each acceptance point is calculated assuming that the whole process is binomial, since it can be modelled as a process with n draws with a success probability p of passing. However, here we cannot use the normal approximation to calculate confidence intervals as this is invalid for extreme values of the probability p (close to zero or one). Here we can make use of the so-called Wilson score interval [219] that constitutes an improvement of the normal approximation, with good coverage properties even for a small number of samples, or in the case of extreme values of p. This confidence interval is defined as:

$$\frac{n_S + \frac{z^2}{2}}{n + z^2} \pm \frac{z}{n + z^2}\sqrt{\frac{n_S n_F}{n} + \frac{z^2}{4}} \tag{3.20}$$

Where n_S and n_F are respectively the number of successes and fails in n trials and z is the $1 - \frac{\alpha}{2}$ quantile of the standard normal distribution i.e the quantile corresponding to an error rate α. We can parameterize the evolution of the acceptance as a function of energy and time with a linear fit on the data points (dashed line in Figures (3.18, 3.19)). For the cumulative acceptance there is a decreasing trend with the increase of the energy of the electron recoils while, as a function of time, the linear fit indicates a rather constant acceptance within the limits of the statistical error.

3.3 Relevant Backgrounds for a modulation search

For the search of a leptophilic dark matter, our main background would be the electronic recoil events produced inside the detector by various mechanisms. And in fact, unlike the nuclear recoil background of the WIMP search, the electronic recoils are so

frequent that it becomes possible to statistically investigate the case of a time modulation of that background in the search for a possible signal of a lepton interacting dark matter. In Fig. (3.9) we have already shown a MC simulation of the various contributions to this background as a function of the fiducial volume of the detector. In Fig. (3.20) we present these contributions as a function of the energy of the electron recoils.

Regarding the minimisation of this background we can divide the various contributions into two parts. On the one hand we see, from Fig. (3.9), that the contribution from the intrinsic radioactivity of the construction materials of the TPC, can be limited with appropriate fiducialization, as it consists in γ rays with energies of $\mathcal{O}$(MeV) that can cause electronic recoils in the $\mathcal{O}$(keV) energy region, through Compton scattering. But for LXe, the total attenuation of γ radiation of $\mathcal{O}$(MeV) is such that an $1/e$ of its initial intensity is reached after a length of ~ 3 cm.

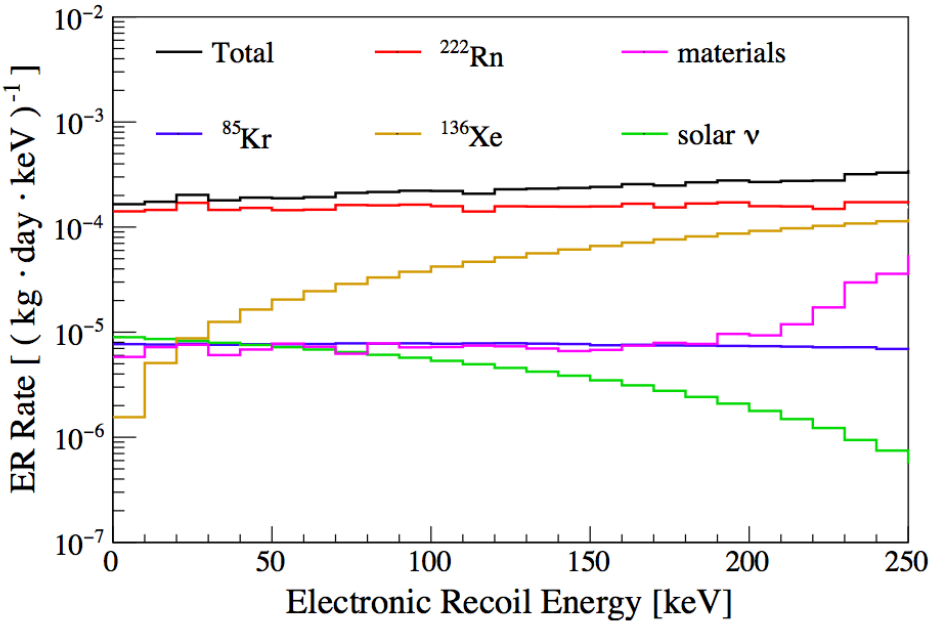


Figure 3.20 – Energy spectrum of the ER background and the various contributions to it in the 1 tonne fiducial volume. Figure from [190].

On the other hand, there are backgrounds, dominant in our region of interest, that are due to beta decays of isotopes whose presence is homogeneous within the detector, such as ^{222}Rn and ^{85}Kr. Here, the choice of fiducial volume is indifferent and a mitigation can be achieved through a dedicated distillation procedure. For the needs of the present analysis it is necessary to examine the time evolution of these contributions to the total background and, in this section, we will briefly present each of them in order of their relative importance, as well as the way we can determine their contribution among the

observed background.

^{214}Pb: As ^{222}Rn is being constantly emanating from the detector materials, where due to its chemical inertness and its relatively long half-life of 3.8 days [218], it becomes homogeneously distributed inside the active volume of LXe. The subsequent decay chain of ^{222}Rn can be seen in Fig. (3.21). In that chain, the daughter that is of relevant importance to our low background experiment concerns the β-decay $^{214}\text{Pb} \xrightarrow{\beta} {}^{214}\text{Bi}$.

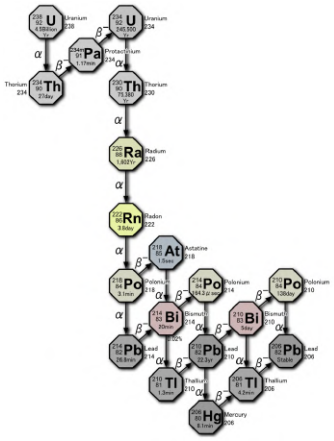


Figure 3.21 – Decay chain of primordial ^{238}U of which ^{222}Rn is transient.

In this case, the low energy tail of this β-spectrum, can produce a single scatter electron recoil in the low energy region of interest. This decay is expected to produce the highest background in the experiment, as can be seen in the simulation (3.20).

In addition to the use of materials selected for low radon emanation, a method that was tested for a short period of time during the experiment in order to further reduce the level of this background was the technique of cryogenic distillation. The working principle here makes use of the fact that in a static equilibrium of LXe with its gaseous phase, at a temperature of $T = 178$ K, radon is a factor of 10 times less volatile than Xe. That is, the ratio of their vapour pressures is $P_{\text{Rn}}/P_{\text{Xe}} \approx 0.1$. Using a distillation column, where a counterflow is created between the evaporated xenon gas and the liquid xenon, due to the difference in partial pressures the radon accumulates in the liquid phase while the purified gaseous xenon returns to the detector. The highly contaminated liquid xenon, then, remains trapped in the reboiler's liquid reservoir until radon disintegration, resulting in a process without loss of xenon during the online operation. Distillation over a period of a few days at the end of SR0 resulted in a 20% reduction in the ^{222}Rn rate. The rate of ^{214}Pb cannot be measured directly but an upper bound can be determined by measuring the rate of ^{222}Rn and ^{218}Po. The measurement of the rate of these isotopes is done via the detection of their alpha decays, Fig. (3.21), that produce large monoenergetic signals that can be measured in the high energy region of the spectrum and thus, indirectly,

offering a monitoring of the presence of the daughter isotope ^{214}Pb. In order to make such an analysis it is necessary to develop specific cuts and corrections in the LCE maps as they are not optimized for those high energy signals. The measurement of the events corresponding to these signals for a specific exposure is then done, with a peak fitting of the corresponding spectrum of the corrected S1 [220].

In Fig. (3.22) are presented the rates of ^{218}Po during SR0 and SR1 which will be used in our analysis as the upper bound of the ^{214}Pb background rate. We can observe that the data indicate a rather stable background over time and therefore, in our analysis, we will exclude the period of radon distillation. From a linear fit in the data results an average rate of $(1.96 \pm 0.08) \cdot 10^{-4}$ events/kg $\cdot$ day $\cdot$ keV.

^{85}Kr: Commercially available xenon is contaminated by natural krypton in the level of part per billion. ^{85}Kr beta decays a 99.57% of times into ^{85}Rb with a half life of ~ 10.8 years and a maximum decay energy of 687 keV. In the natural presence of this isotope in the atmosphere (0.09 PBq) [221] due to its natural production in small quantities by the interaction of cosmic radiation with ^{84}Kr, anthropogenic activity is added mainly by its production in large quantities in nuclear fission, where it appears as one of its main products which is subsequently released into the atmosphere. Its average atmospheric concentration is estimated at 1.3 Bq/m^3 with geographic variations following its major anthropogenic contribution i.e with an increasing concentration near nuclear reprocessing facilities, and a higher mean concentration in the northern hemisphere [222, 223]. As in the case of ^{222}Rn, the ^{85}Kr presence inside the detector can be limited by cryogenic distillation. But here, unlike ^{222}Rn, ^{85}Kr is the high volatile component with a 10.8 times larger vapour pressure than Xe at $T = 177$ K. So the distillation column working inversely than the case of radon distillation[6] will lead krypton at the top and the purified xenon will be collected at the bottom. To achieve this goal, a dedicated distillation column was created by the collaboration [224] and the efficiency of the separation can be monitored using the ^{83m}Kr isotope as a tracer. Using this distillation column, the concentration of ^{85}Kr was reduced by a factor of 10^3 from its initial value, reaching a minimum of 0.36 ± 0.06 ppt. What allows us to include the temporal evolution of the background due to this isotope is the regular monitoring done during SR1 through measurements of the ^{nat}Kr/Xe concentration with a rare gas mass spectrometer (RGMS) [225]. The measurements are presented in the data points of Fig. (3.23). RGMS allows us to monitor the evolution of this concentration during and after krypton distillation, where we see a decreasing by about three orders of magnitude followed by a slight increase during SR1, due to emanation from the detector walls. We can model the temporal evolution of this background in order to use it for the annual modulation analysis through a system of coupled differential equations:

[6] In fact, this is the right operating direction and before, in the case of the ^{222}Rn removal, the column was working in reverse mode.

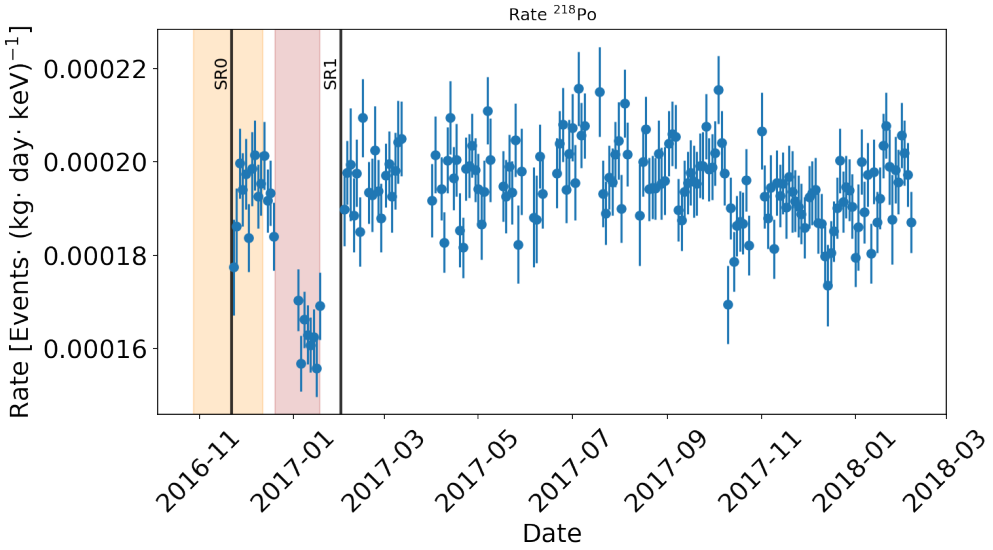


Figure 3.22 – Event rates of ^{218}Po during SR1 for a time bin of 3 days. This rate constitutes an upper bound for the ^{214}Pb background, relevant in our annual modulation of ER, analysis.

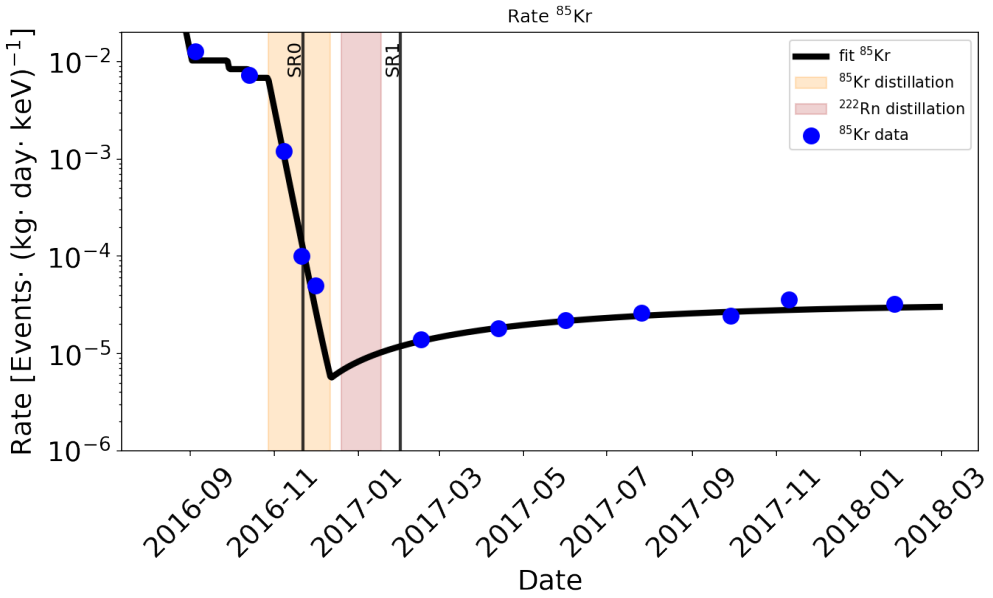


Figure 3.23 – Time evolution of the ^{83m}Kr background during SR0 and SR1. Is also shown a fit on the data taken from RGMS and a fit of the model described by equations (3.21) and (3.23).

If $N_L(t)$ is the concentration of krypton in the liquid phase and $N_W(t)$ is its concentration in the wall of the detector, due to the absorption by its materials, then a state of equilibrium between the two quantities can be described by the system of equations:

$$\frac{dN_W(t)}{dt} = \lambda_a N_L(t) - \lambda_e N_W(t) \tag{3.21}$$

$$\frac{d}{dt}\Big(N_L(t) + N_W(t)\Big) = 0 \tag{3.22}$$

Where the coefficients λ_a, λ_e, control the rate of adsorption and of emanation, respectively. The distillation process intervenes in this closed system by introducing losses in the total krypton concentration, via the second equation of equilibrium with a factor λ_d that determines the distillation rate. In that case the second equation, (3.22), can be written as:

$$\frac{d}{dt}\Big(N_L(t) + N_W(t)\Big) = -\lambda_d N_L(t) \tag{3.23}$$

We can fit the model described by the system of equations (3.21) and (3.23) on the data of Fig. (3.23). Then we obtain the solid curve which will be used as the background model for the time evolution of ^{85}Kr during SR0 and SR1.

Solar neutrinos: Another background that we must consider, although it accounts only for 3% of the total ER background, is the elastic scattering of solar neutrinos off the electrons of the xenon atom [226]. Most of this background (92%) is due to neutrinos produced by the following nuclear fusion reactions:

$$p + p \rightarrow d + e^+ + \nu_e. \tag{3.24}$$

7% is due to electron capture processes of ^7Be (product of fusion of ^{3}He and ^{4}He in the Sun):

$$^7\mathrm{Be} + e^- \rightarrow ^7\mathrm{Li} + \nu_e \tag{3.25}$$

While the most energetic neutrinos are produced by the reaction $p + e^- + p \rightarrow d + \nu_e$ however its flux contribution is suppressed by the fact that this reaction is 400 times less frequent than the pp reaction above, although it has been directly detected in Borexino [227]. Regarding the temporal evolution of this background, it is expected to present an annual modulation due to the eccentricity of the earth's orbit around the Sun. Indeed such a modulation of the ^{7}Be neutrino interaction rate was observed by Borexino with a period, amplitude, and phase that are consistent with a solar origin [228], that is, a 6.7% amplitude modulation with its maximum at the perihelion. In the energy range of [1, 12] keV a rate of about (6.0×10^{-6}) events/kg$\cdot$day$\cdot$keV is expected. For this background we

will use a cosine function with these modulation features, Fig (3.24).

Material radiactivity: Although, as we explained at the beginning of this section, material radioactivity, mainly due to Compton diffusion, can be reduced with appropriate fiducialization, from Fig. (3.9) we see that, within a fiducial volume that includes, for example 1.3 tonnes, the rate of events from this background is expected to be $2 \cdot 10^{-5}$ events/kg·day·keV. It would be useful to analyze its temporal evolution, as the 61% of this background originates from ^{60}Co contamination of the shells and flanges of the cryostat. We can model the temporal evolution through an exponential decay with the ^{60}Co half life of 5.3 yr, which is of the order of magnitude of the time period covered by SR0 and SR1.

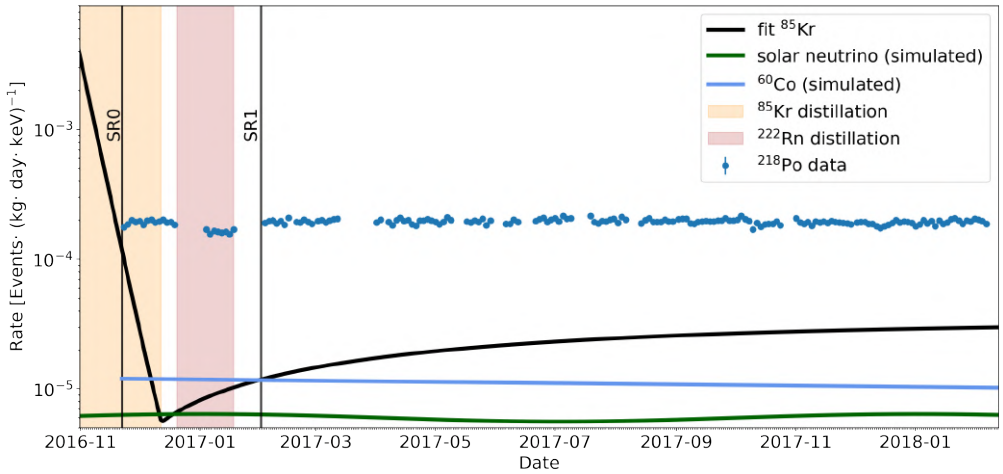


Figure 3.24 – Combined plot of all relevant ER backgrounds, to be used as a background model in the sensitivity analysis and the final statistical inference of the present analysis. Added to the ^{214}Pb (as determined from ^{218}Po a-decay) and ^{85}Kr backgrounds are the contributions from the materials (blue solid line) and the solar neutrinos (green line).

Fig. (3.24) shows these four backround contributions, which are the most important in term of their rate and their time variation, such as ^{214}Pb for which we can have an upper limit but also find that it is rather stable over time, as well as secondary backgrounds which however present time variations and therefore it would be useful to take them into account for the final statistical inference of a modulation of the background that should be consisted with a dark matter interpretation.

Finally, there are a number of additional backgrounds that can be ignored in an annual modulation analysis as they are subdominant and also flat in time. For example the ERs caused by the $2\nu\beta\beta$ decay of ^{136}Xe with a Q-value of ~ 2.5 MeV [229]. The tail of the spectrum can induce events in our low energy region of interest but the rate is

expected to be less than the background rate due to solar neutrinos and also the half-life is 2.17×10^{21} yr *i.e.*, this background is constant in time.

Another background that can be ignored is the one due to accidental coincidences, *i.e.* the pairings, by the data processor, of lone-S1 and lone-S2 events which are incorrectly considered as the main signals of a true interaction. This background can be modelled by randomly matching lone-S1 and lone-S2 detected events. From such background model we find that this is of importance mainly in the region of the parametric space of the observables that corresponds to the NRs while for the ER band, in which our events are collected, this background is insignificant and also constant with time[7].

Lastly, a spurious background that we took into account was the ^{212}Pb in the ^{220}Rn chain. Due to the very short half-life (56 s) of ^{220}Rn it does not have the time to diffuse into the active volume of the TPC. Through the monitoring of ^{212}BiPo delayed coincidences of β followed by an α decay (which follows ^{212}Pb β decay in the ^{220}Rn chain) a lower limit can be inferred about the rate of ^{212}Pb. During the SR1 it was found that this is of the order of 1% compared to the background rate due to ^{214}Pb. However, some datasets were identified in which that ^{212}BiPo amount increased up to 20 times. It was found that these background datasets followed immediately after periods of ^{220}Rn calibrations. They can be attributed to the fact that, in some cases, after a regular ^{220}Rn calibration, the background data taking was not delayed enough, so as to allow the detector to reach again the normal background level. In total, 31 datasets were removed, corresponding to a livetime loss of 1.3 days.

[7] As a consequence of the observed time stability of rates of the lone-S1 and lone-S2 events.

3.4 Detector stability analysis

We will now proceed to an investigation of the possible correlations of ER event rates, as calculated in the previous section, with the physical parameters of the detector such as pressures and temperatures in different parts of the TPC, the xenon cooling power and the recirculation flow rate to the purification system, the measurements of the liquid level, the voltages of the anode and cathode and in general a set of parameters that will hereinafter be called by the generic name slow control (SC) parameters, since they are regularly monitored by a Slow Control SCADA system. Although the final estimate of a possible annual modulation of the background will be based on an unbinned likelihood analysis, a verification is necessary to precede, by checking a possible correlation of the event rate over time with all SC parameters, because implicit in the final inference is the assumption that the time stability of the detector is such that possible changes in the signal generation mechanism, through possible instabilities of the experimental conditions, can not affect a statistical inference based only on the observed events and the background model. Otherwise one should take into account such a correlation with slow control parameters.

This section summarises the results from all studies about the possible correlations between the Electronic Recoil (ER) rate, calculated based on the data selection specified above, and SC parameters associated to important detector performance conditions during SR0 and SR1 data taking. Here, thus, for illustration purposes we will show only the methodology we followed, applied in the case of SR1 and the energy range of 1-6 keV for a FV selection of 1T and a time bin choice of 7 days. This analysis was repeated for the energy ranges of 1-10 keV and 10-20 keV, for fiducial volumes of 1T and 1.3T as well as for time bin choices of 15 and 21 days. Similar analysis was performed for the SR0 data in these three energy regions and for a time bin of 7 days.

We begin by introducing a summary of the different detector parameters that were studied. Then, we present the methodology used to perform the statistical correlation analysis between ER rates and SC parameters. Thus the ER rates for SR0 and SR1 are compared with the mean of all SC parameters considered, and this is repeated for different time-binning conditions and energy ranges. Their statistical correlations are studied by citing point estimates like the Pearson and Spearman's correlation coefficients and the p-values of their corresponding correlation tests. In addition, we also present the attempts to construct confidence intervals for the correlation coefficients and test the hypothesis that their correlation coefficient is different from zero in a different way *i.e.* using general resampling techniques such as bootstrapping and permutation tests.

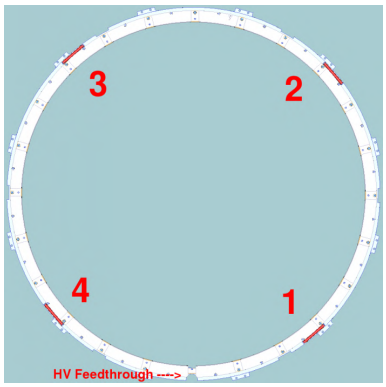
For all the background runs selected for the annual modulation analysis, we collected the values of the SC parameters listed in Tab. (3.2):

The following is a summary of the full analysis focused on a subset of these parameters that are highlighted in Tab. (3.2) and may affect the event rate more directly, such as e.g.

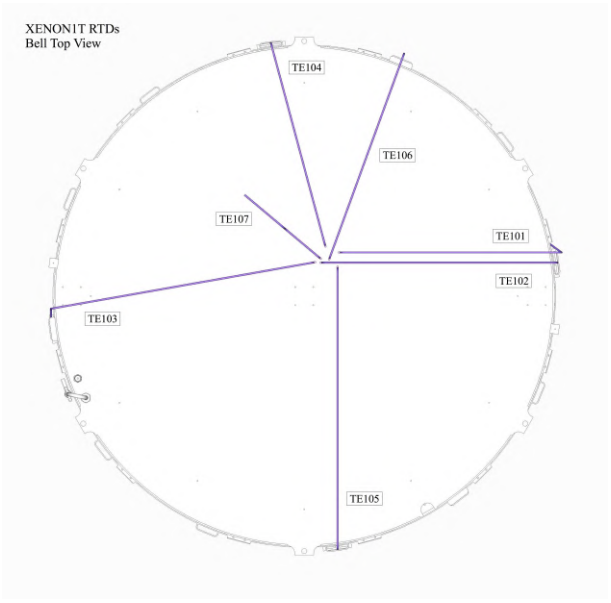
List of Slow Control parameters		
Parameter Name	**Unit**	**Description**
PT101	bar	Detector pressure (cryostat)
PT102	bar	Detector pressure (porcupine)
PT103	bar	Detector pressure (cryogenic system)
PT104	bar	Detector pressure, relative (cryogenic system)
TE101	°C	Cryostat bottom temperature (LXe)
TE102	°C	Cryostat bottom PMT array temperature (LXe)
TE103	°C	Cryostat middle temperature (LXe)
TE104	°C	Cryostat below bell temperature (LXe)
TE105	°C	Reservoir temperature (LXe)
TE106	°C	Inside bell temperature (GXe)
TE107	°C	Cryostat top temperature (GXe)
PI110_PT113	mbar	Vacuum insulation pressure (cryogenic system)
FCV101	SLPM	GXe circulation porcupine mass flow
FCV102	SLPM	GXe circulation cryostat mass flow
FCV103	SLPM	GXe circulation cooling towers mass flow
FCV104	SLPM	GXe bell pressurization mass flow
FC201	SLPM	GXe mass flow through QDrive 201
FC202	SLPM	GXe mass flow through QDrive 202
SLM1	mm	Short level meter 1 reading
SLM2	mm	Short level meter 2 reading
SLM3	mm	Short level meter 3 reading
SLM4	mm	Short level meter 4 reading
LLM2	mm	Long level meter 2 reading
R121	W	PTR-102 tower heater power
Vanode	V	Anode voltage
Vcathode	kV	Cathode voltage (negative)
TILTMETERA	rad	Tilt meter A reading
TILTMETERB	rad	Tilt meter B reading

Table 3.2 – List of slow control parameters examined for a possible correlation with the ER event rates in this section. The parameters that are highlighted will be used as a subset for the presentation of the method and also because they may affect the event rate more directly (see text).

the cryostat top temperature in the gas volume above the bell, as a general increasing should facilitate the outgassing, or the liquid level and pressure that are affecting the secondary scintillation gain. In Fig. (3.25) is shown a view of the positions of the temperature, pressure sensors and the four level-meters.



(a) a



(b) b

Figure 3.25 – a. Positions of the four level-meters measuring the liquid level in the gas gap. b: Top view of the seven temperature sensors of the detector

Fig. (3.26) show the time evolution of the 6 selected SC parameters, for all SR1 datasets used for the annual modulation analysis. The black dashed lines correspond to the average values of the parameters while the two red dashed lines represent the $\pm 1\sigma$ deviation.

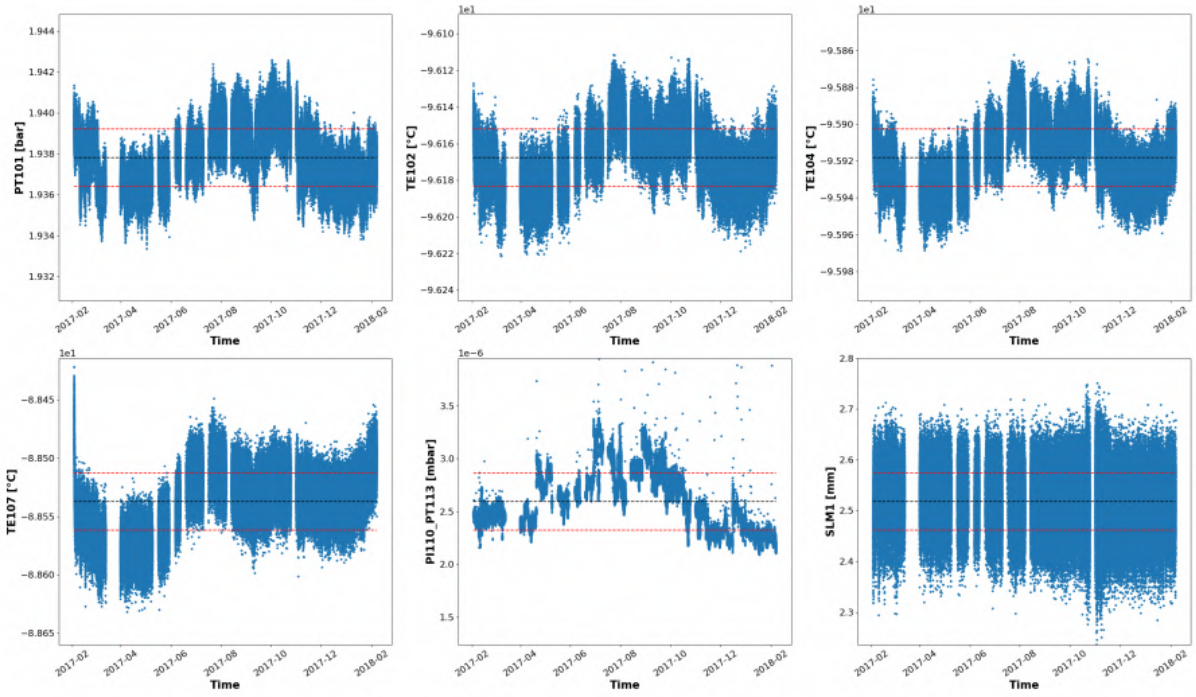


Figure 3.26 – Evolution, during SR1, of the SC parameters used in the illustration of the analysis method in this section. The black lines indicate the mean value and the red lines the 1σ interval

When considering a particular science run, we split the total time of data taken into temporal bins of n days. For these bins, we examined the ER rate within a certain energy range. We examined the correlation of these ER rates with the SC variables for a certain choice of the FV. In Fig. (3.27), we present the correlation plots for the list of SC parameters specified in Tab. (3.2) with the event rates for SR1, for a time binning of 7 days in the energy region of [1, 6] keV. Here, every point corresponds to a particular time bin. The colour code corresponds to the temporal placement of each point during the corresponding SR, such that we can have a visual depiction of the temporal structure of the correlation.

We will proceed by studying some test statistics that quantify the correlations between the pairs variables we are interested in, namely the ER event rates and each of the variables listed in the table (3.2). It is useful to use the so-called Pearson correlation coefficient, which measures the linear correlation between two variables, as well as Spearman's rank correlation coefficient, which expresses whether the two variables can be described by a monotonic function not necessarily linear. For a set of variables $\mathcal{S} = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ the Pearson correlation coefficient for the sample is:

$$r_p = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.26}$$

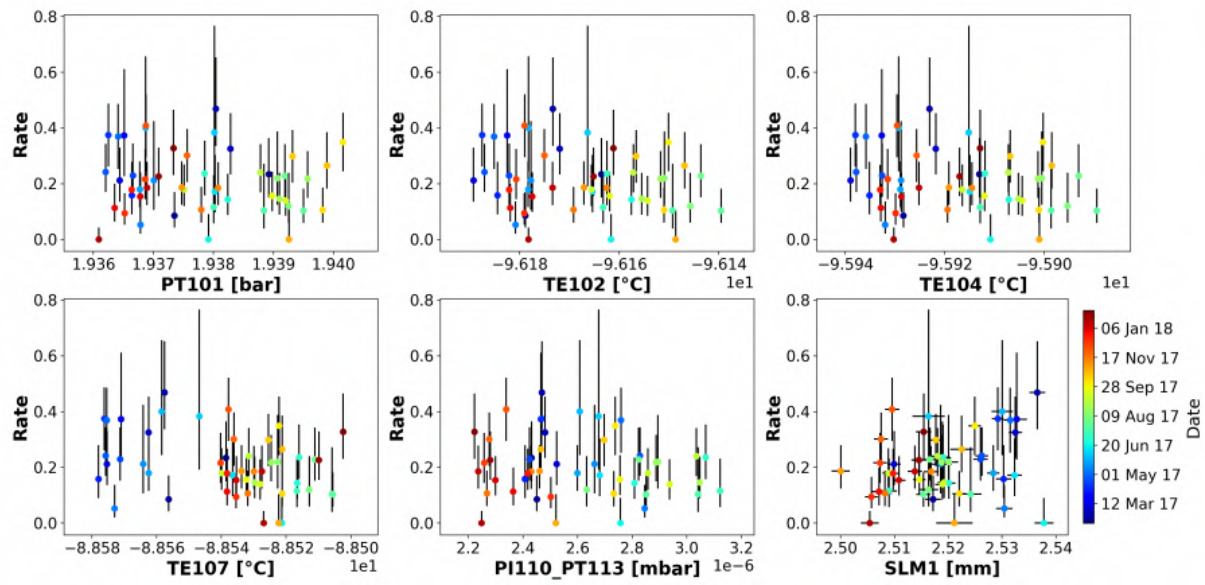


Figure 3.27 – Correlation plots of SC parameters of Fig. (3.26) vs the ER rate in [1, 6] keV range. Each point correspond to a 7 days time-bin.

Where n is the sample size and $(\bar{x}, \bar{y})$ is the sample mean. It is essentially the ratio of the covariance of the two variables with the product of their standard deviations. The Spearman's rank correlation coefficient r_s, on the other hand, expands the range of possibilities for the correlation mode by focusing more generally on a kind of direction of the correlation between the two variables. In particular, this coefficient is functionally the same as the Pearson coefficient except that the numerical values of the variables are replaced by their rank. If rg_X and rg_Y are the ranks of the random variables X and Y, and σ_{rg_X} and σ_{rg_Y} are the standard deviations of the rank variables, then:

$$r_s = \frac{\mathrm{cov}(\mathrm{rg}_X, \mathrm{rg}_Y)}{\sigma_{\mathrm{rg}_X} \sigma_{\mathrm{rg}_Y}} \tag{3.27}$$

What we are interested in, however, is not so much the values of these coefficients for the samples we have observed, but the corresponding values for the statistical population, *i.e.* whether from these samples we can infer an estimate for the correlation of these variables in the population from which we obtain the samples, that is, to infer the properties of the statistical distribution governing them. A common way to attack such problems is to consider a normal distribution approximation, for the population, and try to infer conclusions about the values of the population's parameters based on the statistics calculated from the sample of the observed data. However in our case there is no guarantee that the bivariate distribution of the ER event rates and each one of the SC parameters can be described by a normal distribution. One way to overcome this difficulty is through a non-parametric method involving the use of resampling techniques. Specifically, these

techniques are based on the idea that, in the absence of a model capable of describing the population, we could model the relationship between the sample and the population with the relationship between a randomly selected subsample and the sample itself. In fact, in terms of constructing confidence intervals about parameters of the population, it can be shown that these data-driven methods are asymptotically more accurate than the intervals obtained using sample variance and assumptions of normality [230]. In this analysis we will therefore use two methods from this category. Specifically using bootstrapping [230] we will try to construct a confidence interval regarding the correlation coefficient of event rate with each of the SC parameters and then using permutation tests [231] we will try to test the hypothesis H_0 that the correlation coefficient of the population is zero, that is, we will try to see whether or not the hypothesis of the non-correlation of the two variables can be rejected.

3.4.1 Non-parametric bootstrap

For the observed sample $\mathcal{S}$ of n pairs (x_i, y_i) another sample of n pairs is generated by sampling with replacement from $\mathcal{S}$ and the Spearman correlation coefficient is calculated for the new sample. We repeat this process for a large number of times and we derive an empirical distribution which is an approximation of the sampling distribution of the statistic r, Fig. (3.28). We then construct a 95% confidence interval for the population correlation coefficient ρ by taking the interval spanning from the 2.5th to the 97.5th percentile of the resampled r values [232].

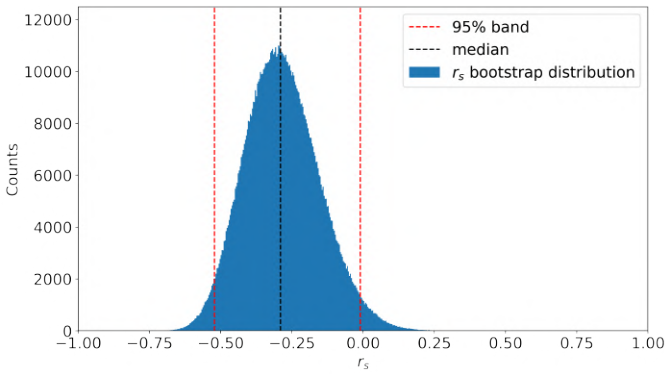


Figure 3.28 – Example of the empirical distribution of r_s constructed with Bootstrapping for the sample of TE107 in SR1 for 7 days time bin.

In Fig.(3.29) we present a Box Whisker Chart plot with the box-and-whisker summary of the bootstrap distribution for the correlation coefficient for each SC parameter with the ER rate data, for the one of the different time-bin options we studied. The fences are indicating the 90% CI and the box is covering the 50% around the median. The color

code is also used indicating the p-value of the permutation test (see text below) applied for each sample correlation.

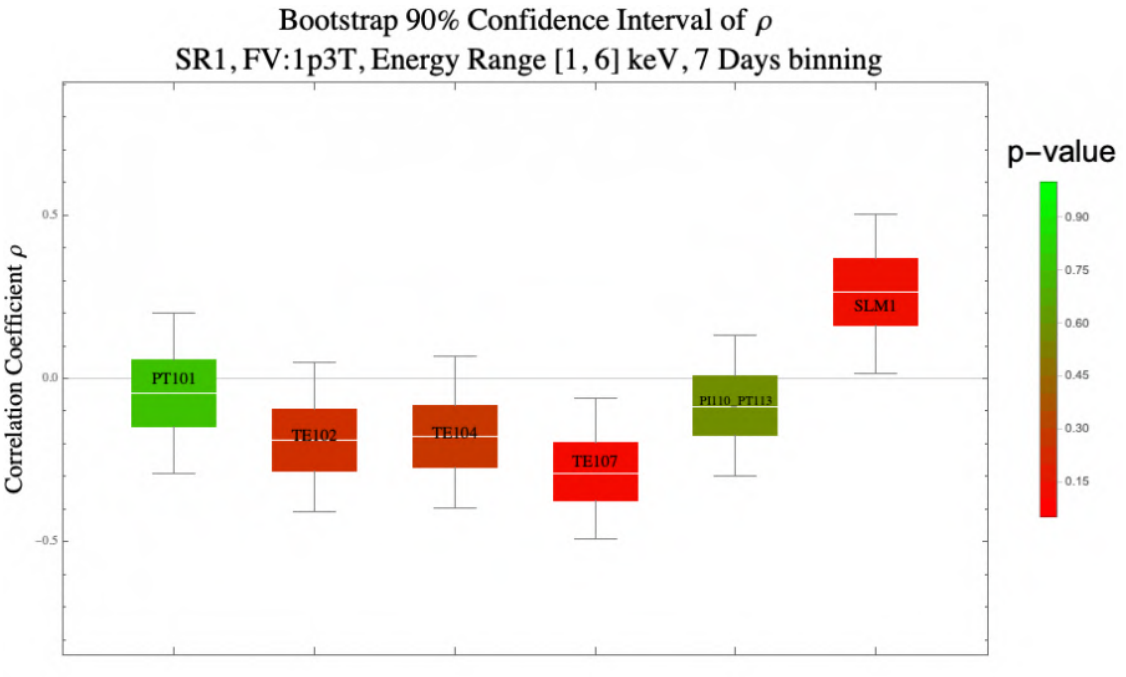


Figure 3.29 – 90% C.I. and p-values for the correlation coefficient between SC parameters and ER rate. SR1, [1,6] keV with 7 days binning.

3.4.2 Permutation tests

In order to test the hypothesis of no correlation in the population ($H_0 : \rho = 0$) we will make use of a resampling technique which we will briefly describe theoretically. For this purpose, let $J(x, y)$ be a continuous PDF describing the joint distribution of a random variable X corresponding to the ER event rates, and of a random variable Y corresponding to one of the parameters of slow control. Also let $E(x) = \int_y J(x, y)dy$ and $S(y) = \int_x J(x, y)dx$ be the marginal distributions. In this case case H_0 can be written as:

$$H_0 : J(x, y) = E(x)S(y), \forall(x, y) \tag{3.28}$$

Then if we have n observed pairs of data $(x_1, y_1), \cdots (x_n, y_n) \equiv (\mathbf{x}, \mathbf{y})$, the likelihood function of H_0, given these data, is:

$$\mathcal{L}_0(H_0|\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} J(x_i, y_i) \tag{3.29}$$

Let's now consider, without loss of generality, an alternative hypothesis, H_A, where (x, y) is distributed under a bivariate normal distribution $f(x,y)$, with correlation coefficient $\rho \neq 0$:

$$f(x,y) = \frac{1}{(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})^n} \exp\left(-\frac{1}{2(1-\rho^2)} \sum_{i=1}^{n} \left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} \right] \right) \tag{3.30}$$

In that case, if we have n observed pairs of data, the likelihood function of H_A, given these data, is:

$$\begin{aligned}
&\mathcal{L}(H_A|\mathbf{x},\mathbf{y}) = \prod_{i=1}^{n} f(x_i,y_i) \\
&= \frac{1}{(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})^n} \exp\left(-\frac{1}{2(1-\rho^2)} \sum_{i=1}^{n} \left[\frac{(x_i-\mu_X)^2}{\sigma_X^2} + \frac{(y_i-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x_i-\mu_X)(y_i-\mu_Y)}{\sigma_X\sigma_Y} \right] \right) \\
&= \frac{1}{(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})^n} \exp\left(-\frac{n}{2(1-\rho^2)} \left[\frac{(\bar{x}-\mu_X)^2}{\sigma_X^2} + \frac{(\bar{y}-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(\bar{x}-\mu_X)(\bar{y}-\mu_Y)}{\sigma_X\sigma_Y} \right.\right. \\
&\left.\left. + \left(\frac{s_x^2}{\sigma_x^2} - \frac{2\rho r s_x s_y}{\sigma_x\sigma_y} + \frac{s_y^2}{\sigma_y^2} \right) \right] \right)
\end{aligned} \tag{3.31}$$

Where $\bar{x}$, $\bar{y}$, s_x, s_y are respectively the observed mean and variance, while r is the observed (or sample) correlation coefficient (3.26). The test of H_0, with H_A as the alternative, can be formulated rigorously as the construction of a critical region Ω in the sample space (x, y) such as to minimize the probability β of an error of the second kind (non rejecting the null hypothesis when it is false), *i.e.* to maximize the so-called power of the test:

$$1-\beta = \int_{\Omega} \mathcal{L}(H_A|\mathbf{x},\mathbf{y}) d\mathbf{x} d\mathbf{y} \tag{3.32}$$

Subject to a condition that provides a desired level of probability α for errors of the first kind (rejecting H_0 when it is true):

$$\alpha = \int_{\Omega} \mathcal{L}(H_0|\mathbf{x},\mathbf{y}) d\mathbf{x} d\mathbf{y} \tag{3.33}$$

According to the Neyman-Pearson lemma [233], such an Ω can be constructed from that part of the sample space for which the ratio $\Lambda = \mathcal{L}(H_A|\mathbf{x},\mathbf{y})/\mathcal{L}(H_0|\mathbf{x},\mathbf{y})$ is maximized, and this happens, if and only if, Ω is exactly that fraction α of the sample space that contains the highest values of the ratio Λ, because (3.32) can be written as:

$$1-\beta = \int_{\Omega} \frac{\mathcal{L}(H_A|\mathbf{x},\mathbf{y})}{\mathcal{L}(H_0|\mathbf{x},\mathbf{y})} \mathcal{L}(H_0|\mathbf{x},\mathbf{y}) d\mathbf{x} d\mathbf{y} \tag{3.34}$$

The problem in our case is that we want to maintain the generality and to not limit ourselves to a specific parametric model of $J(x,y)$ since, for instance, we have no reason to assume that for H_0 the joint distribution of event rates with the SC data is a bivariate normal.

The idea of the permutation test is based on the fact that, if H_0 is true, then each of the $(n!)^2$ pairs, that we can construct with permutations of the (x,y), data are equiprobable under $J(x,y)$. If we are only interested in the relationship of x with y, then we can limit ourselves to examining only the distinct pairings, which are $n!$, and can be obtained fixing x and permuting y.

Thus, if we want to construct a test for H_0 against H_A, we can overcome the problem that we do not have a parametric form for H_0 and, simply assuming that it is a continuous PDF, we can make use of the Neyman-Pearson lemma and choose as a critical region of rejection of H_0, with level of significance α, the fraction α of the sample space containing the permutation pairs that maximize (3.34), because (3.29) is invariant under permutations and so is unchanged in the ratio Λ appearing in the integrand of (3.34), while for (3.31) the permutation leaves unchanged $\bar{x}$, $\bar{y}$, s_x, s_y but affects r. We could then reject H_0 against an H_A with $\rho > 0$, if r is larger than a certain threshold, and against an H_A with $\rho < 0$, if r is smaller than a certain threshold. The critical value of the threshold can be determined from the permutation distribution over the $n!$ distinct and, under H_0, equiprobable pairs of values, subject to our specification about the desired significance level α.

Looking more closely at the permutation distribution, we can try to approximate it by calculating its moments. Without loss of generality for the argument we want to make here, in the case of $(\bar{x},\bar{y}) = (0,0)$, it can be shown [234] that the first 4 moments of this distribution are:

$$\begin{aligned} &E(r) = 0, \quad E(r^2) = \frac{1}{n-1}, \\ &E(r^3) = \frac{n-2}{n(n-1)^2}\text{Skew}(x)\text{Skew}(y), \\ &E(r^4) = \frac{3}{n^2-1}\left(1 + \frac{(n-2)(n-3)}{3n(n-1)^2}\text{Kurt}(x)\text{Kurt}(y)\right) \end{aligned} \tag{3.35}$$

Where Kurt and Skew stand for the kurtosis and the skewness respectively. It appears that the asymptotic behavior of (3.35) in the limit $n \to \infty$ is identical with the moments of the distribution of the sample correlation coefficient from a bivariate normal distribution with $\rho = 0$:

$$f(r) = \frac{1}{B(\frac{n-2}{2}, \frac{1}{2})}(1-r^2)^{\frac{n-4}{2}} \tag{3.36}$$

Where B is the beta function. Therefore even for a small number of data, one can approximate the permutation distribution, with (3.36) and use it in order to construct the critical region of the hypothesis test. Specifically we can use the fact that under (3.36) the test statistic

$$t = r\sqrt{(n-2)/(1-r^2)} \tag{3.37}$$

follows a Student t distribution with $n-2$ degrees of freedom.

Thus the p-value of the hypothesis test in Fig. (3.30) was constructed as follows. From the original sample pairs (x_i, y_i) we randomly redefine pairs $(x_i, y_{i'})$ by drawing $y_{i'}$ randomly without replacement from $[y_1, y_2, \ldots, y_n]$. We then calculate the statistic (3.37). We repeat this random pairing many times and construct the histogram of Fig. (3.30). Then for the observed dataset we can estimate the p-value for the null hypothesis calculating t for the observed dataset (t_{obs}) performing a two-tailed test.

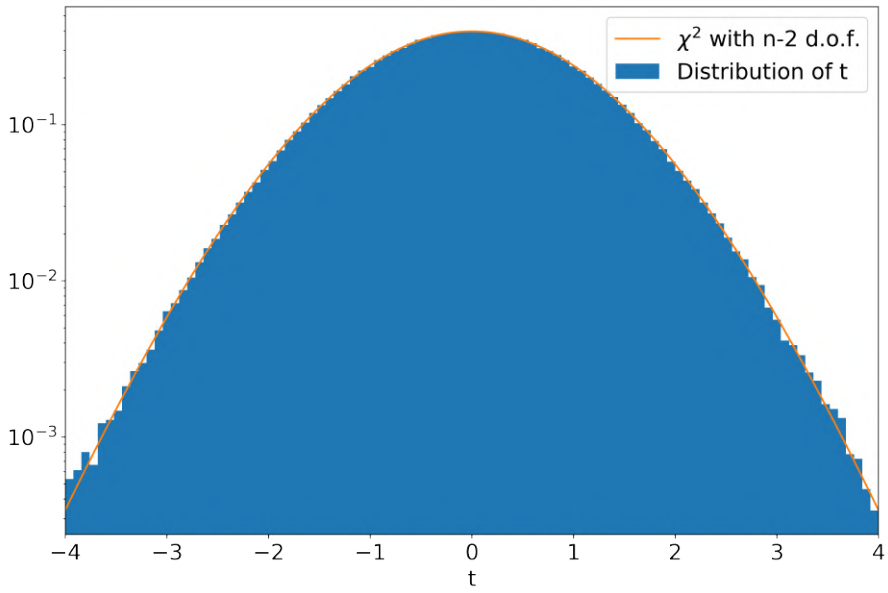


Figure 3.30 – Example of the empirical distribution of t constructed with permutation test for the sample of PT101 in SR1 for a 7 days time bin, 1-6 keV energy range and 1.3t FV.

3.4.3 Trial Factors

In the analysis we performed to calculate the p-values for the hypothesis test of H_0, the same ER data (for different energy regions and fiducial volumes) were used to search for correlation with the SC parameters, in many different bin options and for many different SC parameters. For each one of these degrees of freedom a hypothesis test was conducted and a p-value was evaluated. However, in these cases care should be taken

in order to avoid the, so called, data-dredging [235], a concept closely related to the look-elsewhere effect [236]. When one is conducting a significance test, in a particular subspace of the parameter space, one should also take into account the probability that a significant result (e.g., concerning a correlation) may occur anywhere in the space. For this purpose it is necessary to complete the above analysis by calculating the so-called trial factors *i.e.*, global p-values that will take care of a significance "inflation" occurring when performing many independent hypothesis tests on the same dataset. Specifically, we can search for the maximum value of the parameter t (3.37) that occur in each simultaneous permutation tests of the ER event rates and all SC variables under all bin choices. That is, we evaluate the distribution of the most extreme value of the variable t obtained under the null hypothesis. In order to do so, and as we are only interested in the existence of a correlation and not whether is positive or negative, we can again randomly pair ER rates and SC variables, for each time bin and SC variable, and find the maximum $|t|_{max}$ value obtained. If we repeat this procedure we can get the distribution of $|t|_{max}$, under the null hypothesis, Fig. (3.31). Then, following the same logic we constructed the p-values in the permutation tests of Fig. (3.30), we can define a global p-value from the tail of the $|t|_{max}$ distribution of Fig. (3.31).

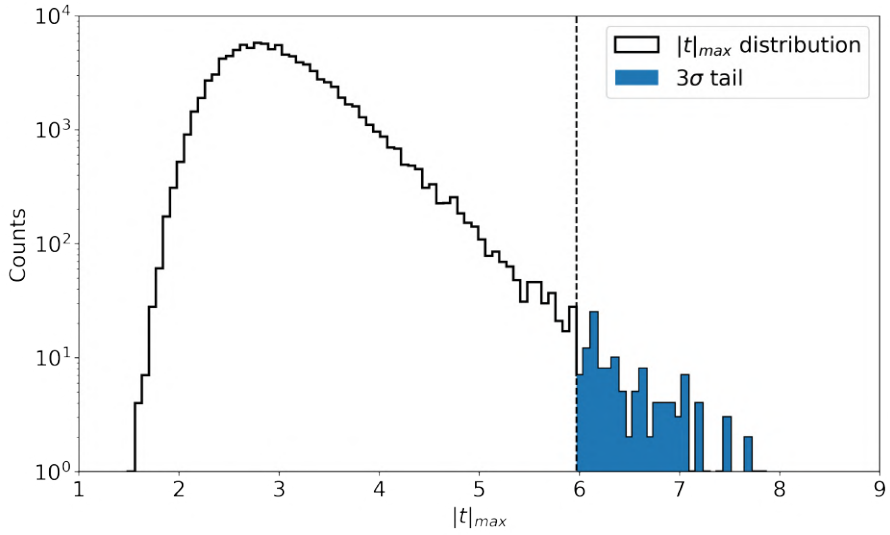


Figure 3.31 – Distribution of the maximum $|t|$ value obtained under the null hypothesis, from permutations, for all time bin options and all SC parameters used in the analysis. Global p-values can be evaluated from the tail of this distribution. In this particular plot the global p-values concerns the datasets in the energy range of 1-6 keV and 1.3t fiducial volume. This p-value is to be compared with the local p-values evaluated previously for each time bin and SC parameter.

We observe that, for some SC parameters, the values calculated for the particular time bin of 7 days and with a permutation test only between the particular SC variable and

the event rates, can reach up to 2σ. However, if we take into account all independent tests we performed, on all different degrees of freedom (time bin and different SC parameters), then, in fact, this value appears inflated with the actual significance threshold being much lower, in terms of p-value. Or, stated otherwise, quite higher in terms of local sigma, as shown by the conversion of that globally evaluated p-value in numbers of σ to the right of Fig (3.32).

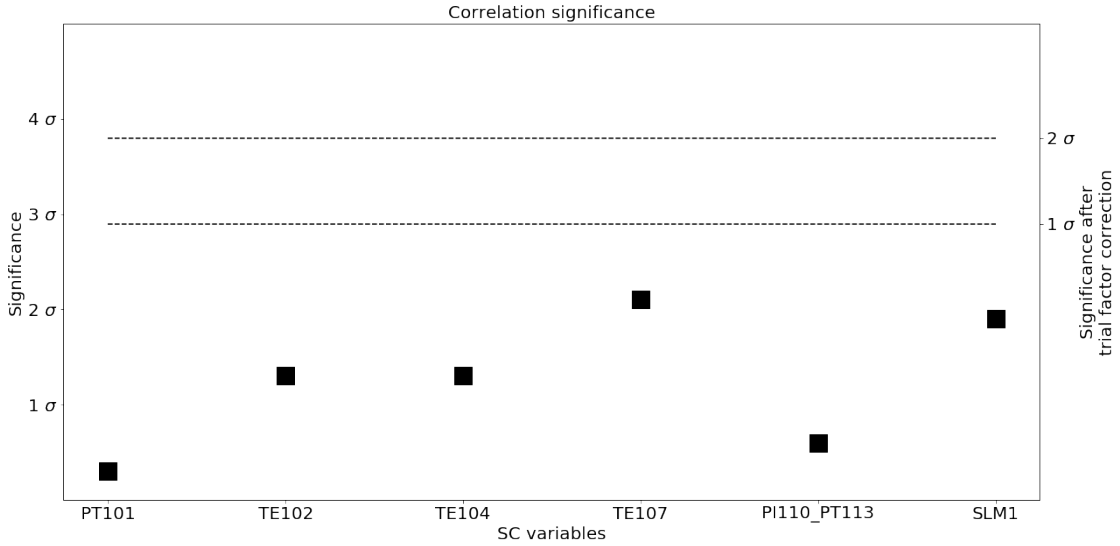
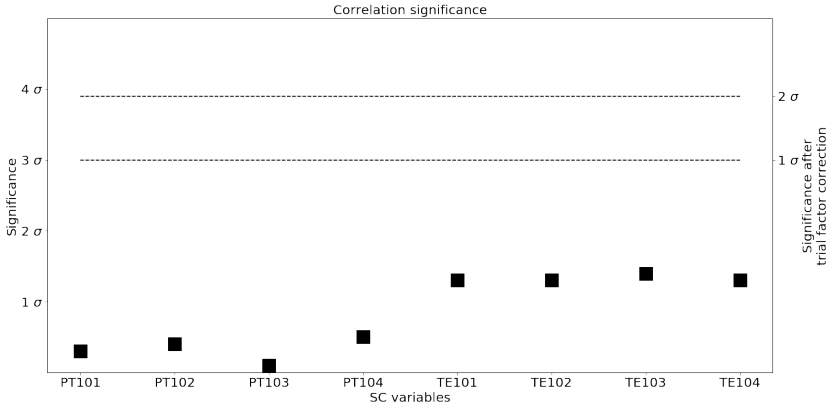


Figure 3.32 – Global significances for 1p3T FV and 1-6 keV energy range and comparison with the local significances concerning the correlations of ER vs SC for the datasets with 7 days time bin (black boxes).

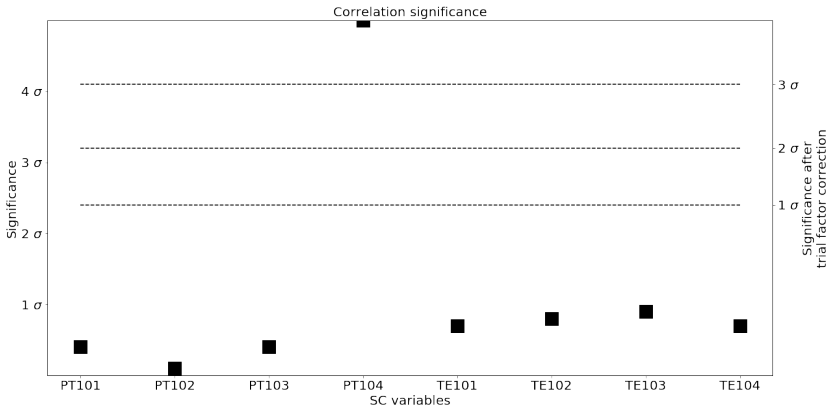
As a cross-check, in order to verify if, nevertheless, our method could detect a significant correlation, we performed a toy MC simulation. For the event rates of Fig. (3.27), corresponding to a specific choice of energy range, fiducial volume and time bin, we obtain the sample covariance matrix of the parameters. For example, if we examine n SC variables, then the vector of parameters is $[X_{er}, C_1, \cdots, C_n]$ where X_{er} is the event rates and C_i is the i-th SC variable, and the full covariance matrix is:

$$\mathbf{K} = \begin{bmatrix} \sigma^2_{X_{er}} & \mathrm{Cov}(X_{er}, C_1) & \cdots & \mathrm{Cov}(X_{er}, C_n) \\ \mathrm{Cov}(C_1, X_{er}) & \sigma^2_{C_1} & \cdots & \mathrm{Cov}(C_1, C_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(C_n, X_{er}) & \mathrm{Cov}(C_n, C_1) & \cdots & \sigma^2_{C_n} \end{bmatrix} \tag{3.38}$$

Then, we inject an artificial correlation between the event rates and one among the slow control variables via a correspondingly modified covariance matrix $\tilde{\mathbf{K}}$. We can create toy MC datasets, sampling from a multivariate normal distribution with $\tilde{\mathbf{K}}$ as the covariance matrix of the population. We can proceed to the calculation of the local and global significances to check whether the injected correlation is retrieved as significant. Let's apply this method by looking at the set of SC parameters shown in Fig (3.33a). For example, the observed correlation coefficient between the event rates and the relative detector pressure of the cryogenic system (parameter PT104) is -0.07. Say that we artificially magnify this correlation by a factor of 10. If we produce a toy MC sampling from a multivariate normal distribution with such a correspondingly modified covariance matrix, we get the output of Fig (3.33b), where we retrieve a globally significant correlation for PT104.



(a) a



(b) b

Figure 3.33

The statistical treatment described above has been applied to the totality of slow control parameters listed in Tab. (3.2). The search for a possible correlation on the

binned data was done over the ranges of [1-6] keV, [1-10] keV and [10-20] keV, for two fiducial volume choices, of 1t and 1.3t and for three time binning options, of 7, 15 and 21 days.

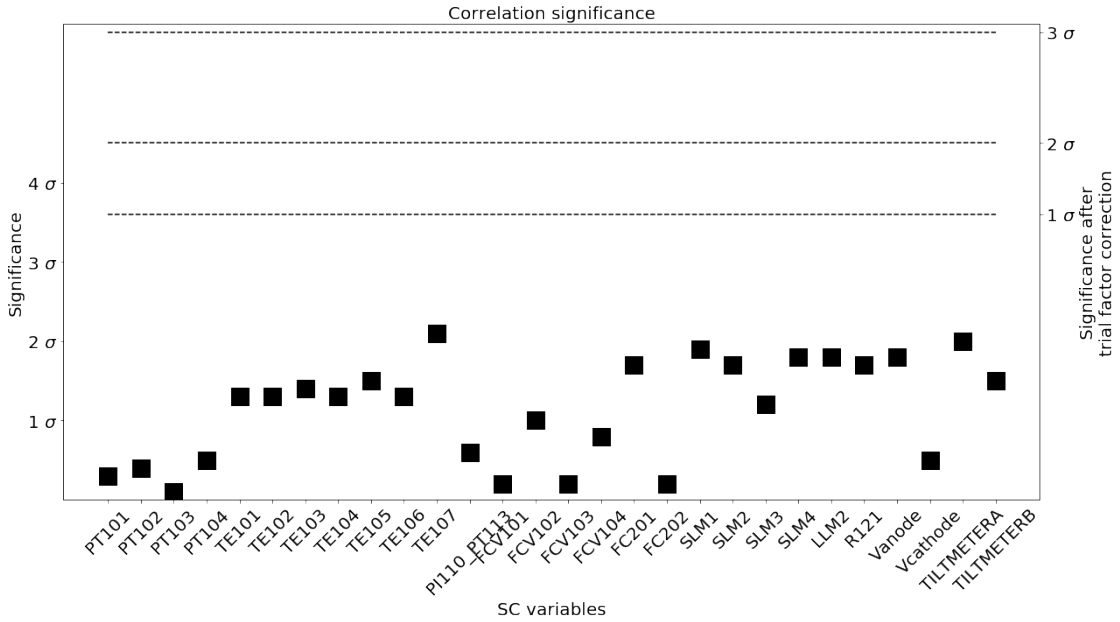


Figure 3.34 – Global significances for 1.3 tons FV and 1-6 keV energy range, for SR1, compared to the local significances concerning the correlations of the ER data with each one of the slow control parameters of Table (**??**), for the datasets with 7 days time bin (black boxes).

Producing diagrams such as Fig.(3.34) and computing the local significance for every degree of freedom, no globally significant correlation was detected over any degree of freedom of our analysis, for both SR0 and SR1. This analysis let us to conclude that, the variation of the detector's SC parameters isn't expected to be significantly correlated with the detected ER event rates, thus the time evolution of the SC parameters of the detector can be considered stable, with respect to the analysis of a possible annual modulation of background events.

3.4.4 Sensitivity analysis

The final statistical inference about a possible modulation of the background will be done with an unbinned likelihood fit on each event selected in Fig. (3.16). Specifically, as we have seen in section (3.1), the time distribution of these events could be described by a PDF with the general form:

$$f(t) = \frac{1}{N}\varepsilon(t)\Big(C + A\cos(\omega(t-\phi))\Big) \tag{3.39}$$

Where $\omega = 2\pi/T$, C is the average background rate and A is the modulation amplitude. $\varepsilon(t)$ is the time variating acceptance, cumulative in the energy range of interest and in the cuts applied, which we identified earlier. Here we can distinguish the two hypotheses that can be tested with null hypothesis H_0 written as $A = 0$ and alternative hypothesis H_A with $A > 0$. Concerning the normalization constant, we must keep in mind that the background data taking does not cover a continuous period of time, from the beginning to the end of the scientific runs, but rather is interrupted several times to allow various calibration procedures to take place. Therefore, time t in (3.39) which is the real time of the earth's rotation around the Sun, differs from the livetime of the experiment. Thus, if we denote as T_s^i and T_e^i the start and end times for the i continuous data taking periods, then the total livetime is $\sum_i (T_e^i - T_s^i)$, and the normalization coefficient N is given by:

$$N = \left(\sum_i C(T_e^i - T_s^i)\varepsilon(T_e^i) + A/\omega \sum_i (\varepsilon(T_e^i)(\sin(\omega(T_e^i - \phi)) - \sin(\omega(T_s^i - \phi))))\right) \tag{3.40}$$

Now, the likelihood function for a set of observed events like those in Fig. (3.16), given the data taking periods, and the parameters of the model (3.39) can be written as:

$$\log \mathcal{L}(A, C, \omega, \phi) = \sum_i \log f(t_i) \tag{3.41}$$

And the unbinned profile likelihood test statistic is defined as:

$$\lambda(\omega) = -2\left(\max_{A,C,\phi}[\log \mathcal{L}(A, C, \omega, \phi)] - \max_{C}[\log \mathcal{L}(A = 0, C, \omega, \phi)]\right) \tag{3.42}$$

The values in the likelihood of the first term are the conditional maximum likelihood estimators for the specific ω, while the value of C, in the second term, is the conditional maximum likelihood estimator under both constrains $A = 0$ and a specific value of ω[8]. As the dimensionality difference between the two terms is 2, we expect that, when the null hypothesis H_0 is true, then $\lambda(\omega)$ should follow a χ^2 distribution with 2 d.o.f [237].

In the above relations, the parameter C has contributions from the time average of the signal but also from the already existing background. If the existing background variates

[8] ϕ stops being an estimator because it disappears with $A = 0$

with time, as is the case for example with the second most important background, due to ^{85}Kr, then this could be taken into account with an additional term $\tilde{C}(t)$ which will be added to (or will replace, depending on the case) the term C in relation (3.39), allowing us to describe exactly our knowledge of the evolution of this background, provided by the analysis of section (3.3).

The sensitivity of the experiment to possible background fluctuations depends on the total livetime, the known background we've studied in section (3.3), as well as, its time evolution. In order to obtain an estimate of this sensitivity we can perform toy-MC pseudo-experiments with samplings from the PDF (3.39) where we can inject a fake modulation signal with a certain modulation fraction and, given the number of events expected from the background, to determine the size of the modulation fraction that would allow a statistically significant inference of the signal.

In order to give a general example of the method, let us consider that the background is flat and specifically the average of the background of Fig (3.22) which is mainly due to the beta decay of ^{214}Pb (as inferred by the monitoring of the α-decay of ^{218}Po). Let us assume also $\varepsilon(t) = 1$. In this case we sample events from (3.39) under the background only ($A = 0$) and the modulation hypothesis. These sampled events of course correspond only to the time periods where the detector records background data, *i.e.* to the periods described above as $[T_e^i, T_s^i]$ for all the acquired datasets.. For the modulation we chose a certain period, phase and modulation fraction $A/(C + A)$. First we verify that the distribution $lambda(\omega)$ follows the expected statistic (3.42) if we repeat many times the pseudo-experiments under the background only hypothesis. As an example, we show in Fig. (3.35) this distribution after considering, for the modulation hypothesis H_A, a period of $T = 365$ days. The distribution clearly follows a χ^2 with 2 d.o.f as expected.

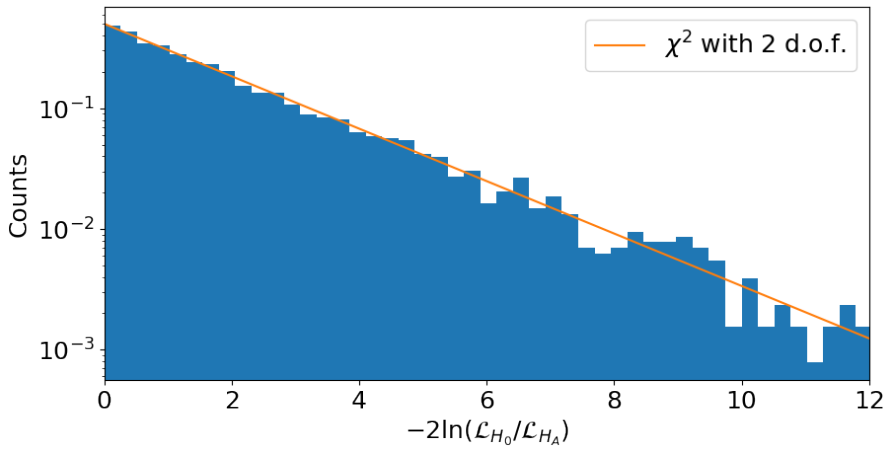


Figure 3.35 – A simulated modulated signal shows that the related unbinned profile likelihood test statistic from formula (3.42) presents (blue histogram), as expected, a χ^2 distribution with 2 d.o.f.

Based on the distribution of (3.42) we can define a confidence interval regarding the asymptotic behaviour of the test statistic q, where the departure from the distribution of Fig. (3.35), for a certain value of the period T, would signify the presence of a signal of this period . We can illustrate this, realising many pseudo-experiments with an injected modulation and performing a scan of $q(\omega)$ over many values of ω. In terms of the period T, such a scan ranges from 0 to a year. For each scanned period we calculate the 1σ and 2σ confidence interval around the median of the distribution of $q(\omega)$.

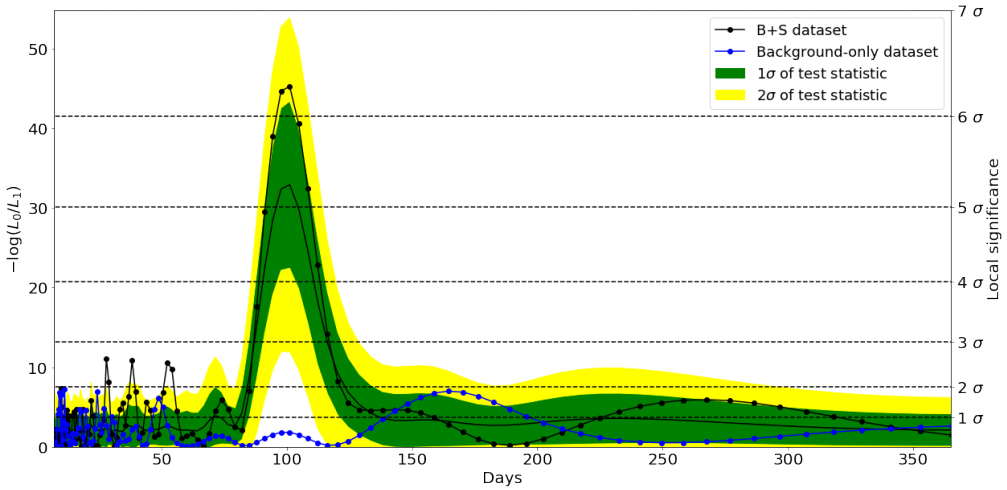


Figure 3.36 – 1σ (green) and 2σ (yellow) confidence intervals of the test statistic, scanning over a possible modulation signal of period T. The pseudo-experiments used for such a test had, on average, a modulation signal with period $T = 100$ days, a phase $\phi = 0$ and a modulation fraction $A/(A+C) = 0.33$. Also shown the observed q values for a particular background-only pseudo experiment (blue dots) and for a particular modulation-injected pseudo-experiment (black dots). This demonstrates how well the statistical method is working.

Fig. (3.36) shows the result of this scan, after having introduced in the pseudo-experiments a modulation signal with a period of 100 days and a phase of 0 days, with modulation fraction 0.33. We observe that the period of the modulation is reconstructed with a peak of the sensitivity band around T = 100 days. Fig. (3.36) also shows the "observed" q values for a particular background-only pseudo experiment (blue dots) and a particular modulation-injected pseudo-experiment (black dots).

In Fig. (3.37) we present the same sensitivity band that would follow the test statistic q, constructed with pseudo-experiments with an injected signal of $T = 365$ days, zero phase and modulation fraction 0.33. We notice that now the resolution of the reconstructed period is less good as we do not have a clear peak around the actual period of the signal but, on the contrary, the sensitivity band is reaching a plateau. This is because the signal period is roughly equal to the time period $\Delta t_{\text{data taking}}$ covered by SR1. And of course

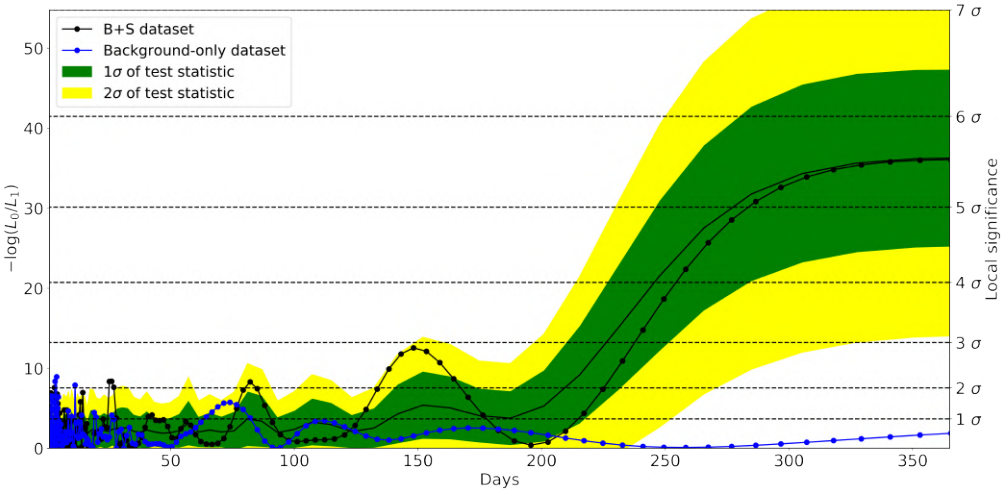


Figure 3.37 – 1σ (green) and 2σ (yellow) confidence intervals of the test statistic, scanning over a possible modulation signal of period T. The pseudo-experiments used for such a test had, on average, a modulation signal with period $T = 365$ days, a phase $\phi = 0$ and a modulation fraction $A/(A+C) = 0.33$. Also shown the observed q values for a particular background-only pseudo experiment (blue dots) and for a particular modulation-injected pseudo-experiment (black dots). This shows the limitation of a data taking whose duration does not exceed enough the period of a year.

modulated signals with periods $T \gg \Delta t_{\text{data taking}}$, will be reconstructed, if at all, with even worse resolution, as a natural consequence of the fact that, within the time period $\Delta t_{\text{data taking}}$, the periodicity of the signal is not observed in its entirety. However, with regard to the signal of a modulation with a characteristic period of one year expected from a dark matter signal, it would at least be possible for a modulation fraction of 0.33 to be distinguished from the unmodulated background, even if the period could not be reconstructed without a large error.

In the analysis we can now include the background as specified in section (3.3). To show the method we will include here only the dominant background of ^{214}Pb as well as the next-to-dominant ^{85}Kr background which presents the largest time variation and which could therefore affect the reconstruction of the period. For this purpose (3.39) must include the two terms describing the background, as determined by Fig. (3.22), via the binned histogram, for ^{214}Pb, and Fig. (3.23) via the fit in the model of time evolution of the ^{85}Kr background. In (3.39) the parameter C includes any background that does not variate with time as well as a possible contribution from the unmodulated part of the signal, *i.e.* in our case, the DM signal due to the movement of the Sun in the galactic halo. Our requirement to describe, manifestly the knowledge of the background but also,

to be conservative, in our assessment of sensitivity to a possible modulation in the data, leads us to write the event rate (3.39) as:

$$f(t) = \frac{1}{N}\varepsilon(t)\Big(C_{\mathrm{Kr}}(t) + C_{\mathrm{Pb}}(t) + A(1+\cos(\omega(t-\phi)))\Big) \tag{3.43}$$

where $C_{\mathrm{Kr}}(t)$ and $C_{\mathrm{Pb}}(t)$ is the time dependent background due to ^{85}Kr and ^{214}Pb and A is the signal rate. That is, here we assume that the background term is fully specified and to this is added a minimum dark matter signal which is modulated by 100% around its constant average rate A *i.e.* there is exactly one day where there are zero expected dark matter events. Also (3.43) is normalised for each run in order to become a PDF.

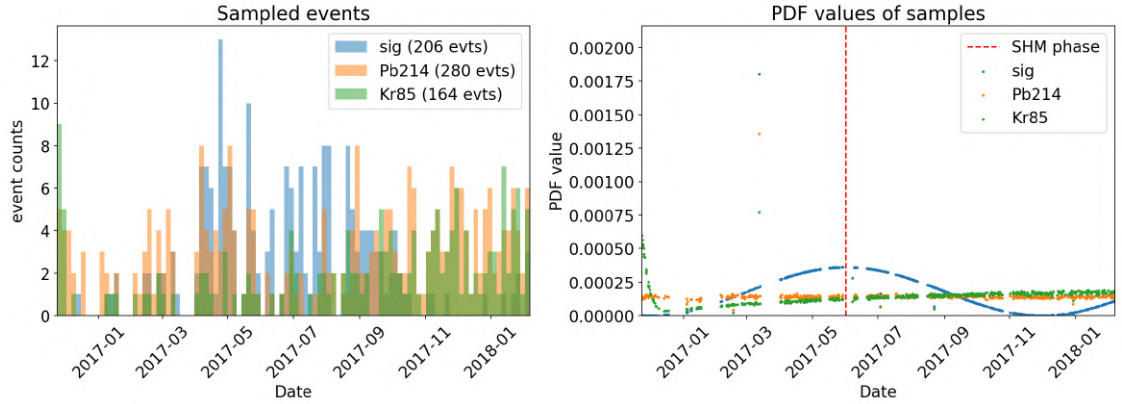


Figure 3.38 – Sampling of toy-MC events (Left) from the background and signal PDFs (Right).

In Fig. (3.38) left we see a sampling from (3.43) where is shown the distribution of events from the signal and the various background contributions. Now in (3.38) we have also introduced the additional 32.7 days of lifetime from SR0. In Fig. ((3.38)) right, is shown the PDFs of the corresponding contribution in (3.43) for the signal and the each one of the background components where the relative weight with which they contribute to the total PDF of the signal is controlled by rate multipliers for each PDF according to the average rate of each background $\langle C_{\mathrm{Kr}}(t)\rangle$ and $\langle C_{\mathrm{Pb}}(t)\rangle$ during SR0 and SR1 and also the signal rate multiplier A, together with the specification of the modulation fraction $A/(A+\langle C_{\mathrm{Kr}}(t)\rangle + \langle C_{\mathrm{Pb}}(t)\rangle)$.

Finally in Fig. (3.39) we repeat the procedure we followed previously where by producing many toy MC datasets from (3.43) with injected signals of different periods we calculate the asymptotic distribution of q (3.42) and thus we estimate the resolution of the reconstruction of the period. We notice again that for the physically interesting case of T = 365 days the fact that we added the livetime of SR0 but also the knowledge of the two largest backgrounds does not change our difficulty on reconstructing with high

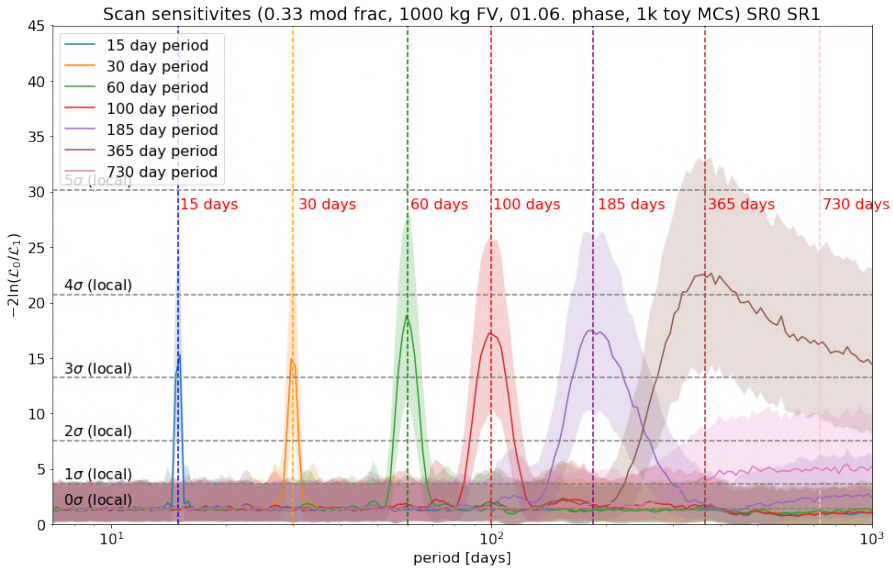


Figure 3.39 – 1σ confidence intervals of the test statistic, scanning over a possible modulation signal of period T. 7 sets of pseudo-experiments have been generated, with modulation signals with different periods and a fixed phase and modulation fraction 0.33. Background model now includes all relevant, time-dependent, components. The data taking range considered is as wide as the sum of SR0 and SR1. Plot done with the computation framework developed by the Annual Modulation Team.

resolution periods greater than or equal to 365 days. However we can observe them, if any, or exclude the null hypothesis of an unmodulated background.

However, as is obvious, the statistical significance of such an inference will be determined chiefly by the actual modulation fraction, as can be seen in Fig. (3.40) where the various sensitivity bands are plotted as a function of the modulation fraction for an injected period of T=100 days in the toy MC datasets sampled from (3.43).

The unbinned likelihood inference (3.42) developed in this section, remains to be applied to the data actually observed along with the smoothed cummulative acceptances estimated in Fig. (3.19), in order to scan over different possible periods and to infer, or to exclude, a possible periodicity followed by the actual data. This analysis is in progress at the time of writing of this work, and the final result will be the subject of a dedicated publication by the collaboration.

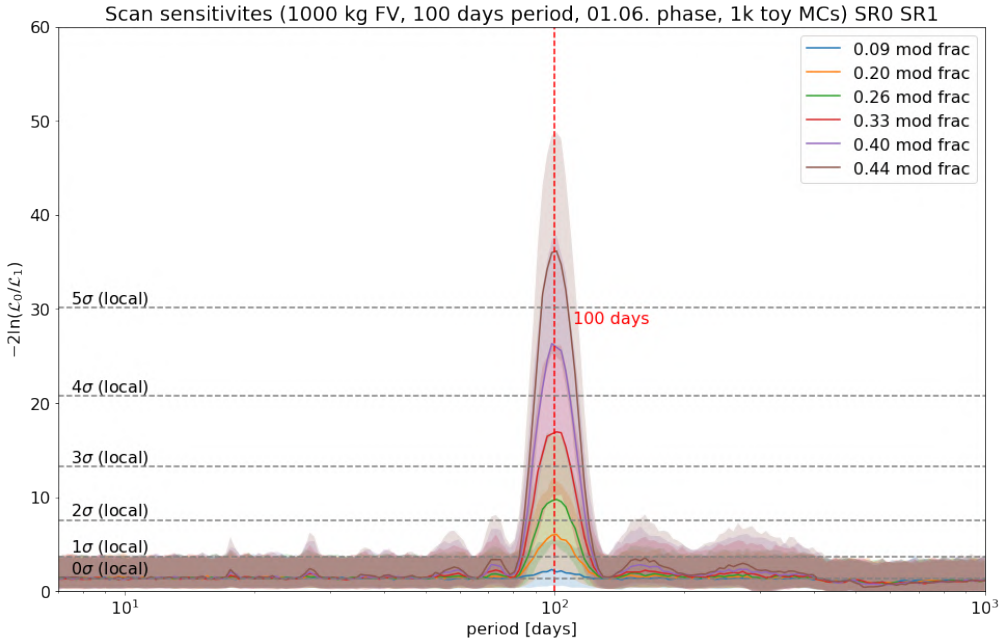


Figure 3.40 – 1σ confidence intervals of the test statistic, scanning over a possible modulation signal of period T. 6 sets of pseudo-experiments have been generated, with modulation signals with different modulation fractions and a fixed phase and period $T = 100$ days. The statistical significance is, as expected, strongly dependent by the intensity of the signal. Plot done with the computation framework developed by the Annual Modulation Team.

Chapter 4

Light Dark Matter Search

4.1 Introduction

Let us now turn our attention back to Chapter 1 and focus on the exclusion limits of Fig. (1.5). We see that the sensitivities of the various experiments, decreases for DM masses m_χ, much higher than 100 GeV/c^2 or less than a few GeV/c^2. In the first case, this is due to the reduction of the rate of expected events as $\sim 1/m_\chi$. In the second case, however, it is due to the fact that WIMPs with smaller m_χ and with relative speeds determined by the DM velocity distribution, will have less kinetic energy to deposit. Below a certain DM mass, the energy of nuclear recoils will be less than the detection threshold (around 5 keV_{nr} for XENON1T) i.e., below the recoil energy necessary to generate the ionization and scintillation signal that could be detected as an S1-S2 pair in the case of our detector.

We can determine the relation between the maximum deposited energy and the mass of a WIMP by studying the kinematic of this elastic scattering. Because the mean velocity of dark matter particles, with respect to an earth-based observer, is of the order of $u \approx 10^{-3}$c we can study this scattering in the non-relativistic regime. Let's consider that the Xe nucleus is initially at rest in the laboratory frame, $\mathbf{p} = 0$, and denote by $\mathbf{k} = m_\chi \mathbf{u}$ the momentum of the incoming WIMP. By the law of conservation of momentum and energy, in the center-of-mass frame the velocities of the two particles remain opposite in direction and with their magnitude unchanged, after the collision. In the center-of-mass frame, the elastic collision is simply rotating the velocities by an angle χ which is determined by the specific interaction.

Elementary kinematics [36] gives us the relation of the momentum of the WIMP $\mathbf{k}$ before the collision, the momenta $\mathbf{k}'$ and $\mathbf{p}'$ of the WIMP and the nucleus after the

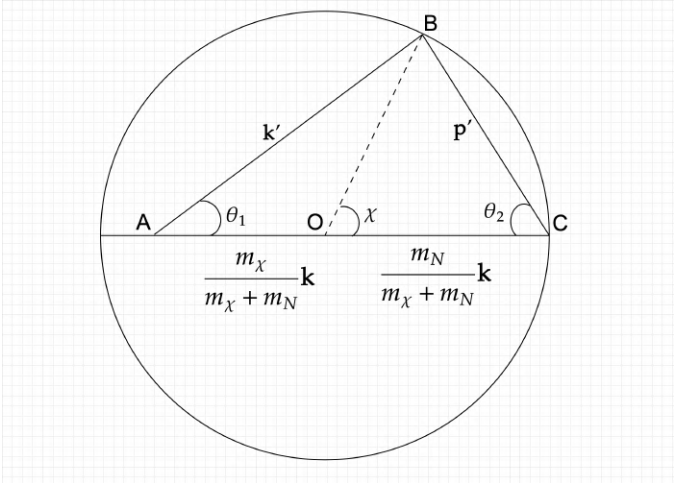


Figure 4.1 – Kinematics of the elastic scattering of a DM particle off a Xe nucleus (initially at rest) for the case where $m_\chi < m_N$. The triangle is formed from momentum conservation between the initial and final momentum of the DM particle, $\mathbf{k}$ and $\mathbf{k}'$ respectively, and the recoil momentum of the nucleus $\mathbf{p}'$, in the laboratory frame. θ_1 and θ_2 are the angles, in the laboratory frame, between the initial ($\overrightarrow{AC}$) and final ($\overrightarrow{AB}$) direction of the DM particle, and between the initial direction of the DM particle and the final ($\overrightarrow{BC}$) direction of the recoiling nucleus. Finally χ, is the angle between the initial and final directions of the motion of the DM particle (or the nucleus), in the center-of-mass frame.

collision and the angle χ Fig.(4.1). From this, it immediately follows that $p' \equiv |\mathbf{p}'|$ is given from:

$$p' = \frac{2m_\chi m_N u}{m_\chi + m_N} \sin(\chi/2) \tag{4.1}$$

where m_N is the mass of the nucleus. Hence, the maximum deposited energy in a nuclear recoil is achieved for $\chi = \pi$ and is:

$$E_{\mathrm{NR,max}} = \frac{p'^2_{\mathrm{max}}}{2m_N} = \frac{2\mu_{\chi,N} u^2}{m_N} \tag{4.2}$$

Where $\mu_{\chi,N}$ is the DM-nucleus reduced mass. From (4.2) we can see that for the Xe nucleus and for a DM mass of $\mathcal{O}$(MeV) we have $E_{\mathrm{NR,max}} \sim \mathcal{O}$(eV) i.e., far below the detection threshold.

However, an interesting observation is that, although for such a small deposited energy of NRs, S1-S2 signals cannot be observed, the total energy of the dark matter particle $\frac{1}{2}m_\chi u^2$, is possibly greater than the binding energy of an electron in the 5p state of the Xe atom or greater than some energy differences between its atomic levels. Therefore, if the dark matter particle had interactions with electrons it would be possible to trigger inelastic atomic processes like an ionization or excitation of the Xe atom.

Let's look at the kinematic aspect of the elastic scattering of such a DM particle off an electron bounded in the Xe atom. Let's again the momentum of the DM particle, before and after the collision, be denoted as $\mathbf{k}$ and $\mathbf{k}'$, respectively. In this case we cannot follow the previous method, as the electron does not even have a specific value of its momentum $\mathbf{p}$, being in an eigenstate of the atomic Hamiltonian. However, the electron have a specific momentum distribution which is determined by its quantum state. In particular, in the language of quantum mechanics, if $|\Psi\rangle$ is its bound state, $|\mathbf{p}\rangle$ and $|\mathbf{x}\rangle$ are the eigenstates of the momentum and the position operators respectively, then the projection of $|\Psi\rangle$ to $|\mathbf{p}\rangle$ is [238]:

$$\langle \mathbf{p}|\Psi\rangle = \int d\mathbf{x}\langle \mathbf{p}|\mathbf{x}\rangle\langle \mathbf{x}|\Psi\rangle = \frac{1}{(2\pi)^{3/2}}\int \exp(-i\mathbf{p}\mathbf{x})\Psi(\mathbf{x})d\mathbf{x}$$

Where we have make use of the bra-ket notation and the identity operator $\mathbf{I} = \int d\mathbf{x}|\mathbf{x}\rangle\langle\mathbf{x}|$. So, the momentum distribution of $|\Psi\rangle$ is

$$|\langle \mathbf{p}|\Psi\rangle|^2 = \frac{1}{(2\pi)^3}\left|\int \exp(-i\mathbf{p}\mathbf{x})\Psi(\mathbf{x})d\mathbf{x}\right|^2$$

That is, the Fourier transform of the coordinate space wavefunction. Thus, the momentum transfer $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ can take various values depending on the momentum-space wavefunction of the electron. However, we can calculate the maximum deposited energy, because the energy transfer to the electron will be:

$$\Delta E_e = \frac{1}{2}m_\chi u^2 - \frac{|\mathbf{k}-\mathbf{q}|^2}{2m_\chi} - \frac{q^2}{2m_N} = \mathbf{q}\mathbf{u} - \frac{q^2}{2\mu_{\chi,N}} \tag{4.3}$$

where here, the term $\frac{q^2}{2\mu_{\chi,N}}$ takes into account the fact that the atom will recoil as a whole. Thus, we can maximise (4.3) with respect to $\mathbf{q}$ in order to find the maximum energy transferred to the electron, $\Delta E_{e,\max}$:

$$\Delta E_{e,\max} = \frac{1}{2}\mu_{\chi,N}u^2 \tag{4.4}$$

From this relation one can immediately find that, for a DM particle of $\mathcal{O}(50\text{ MeV})$, we have a $\Delta E_{e,\max}$ greater than the binding energy of the outer shells of the Xe atomic states. In that case, the ionization electron, if it survives the drift and is extracted to the gaseous phase, it will be detected as a single electron, with exactly the same spectral characteristics of the single electrons we examined in Chapter 2. Further, if the difference of ΔE_e with the binding energy of the electron (i.e., the actual recoiling energy of the ionization electron) is greater than the W-value Fig. (4.2), then the electron will create an additional number of ionization electrons in the LXe. If they survive the drift and are extracted to the GXe, they will produce a signal similar to a temporal pileup of single electrons.

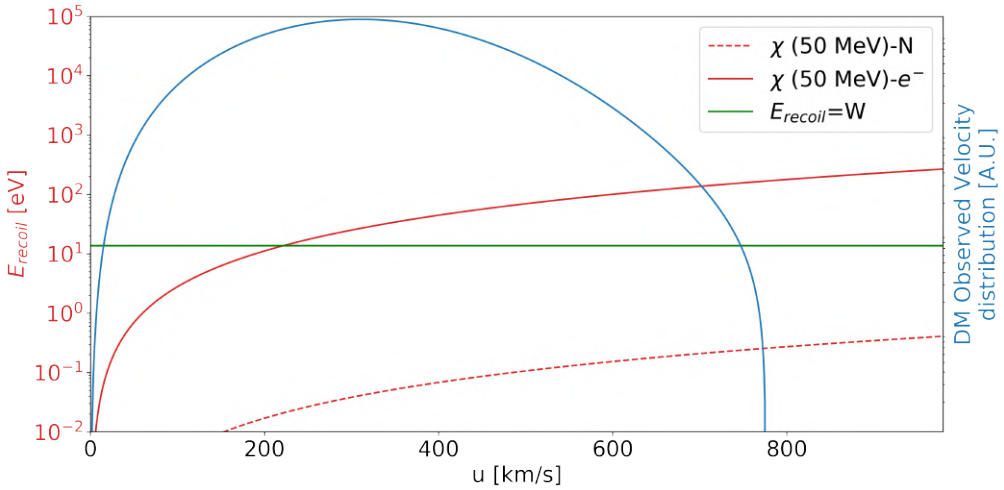


Figure 4.2 – $E_{\mathrm{ER,max}}$ (red solid curve) and $E_{\mathrm{NR,max}}$ (dashed curve) as a function of the velocity. Also is shown the SHM velocity distribution (blue curve). We see that the maximum nuclear recoil energy for a 50 MeV DM cannot even cross the W-value necessary for the production of a quantum in LXe (green line), even for DM particles in the high velocity tail of the halo. But a DM particle with the same mass interacting with electrons can produce an ionization electron with energy hight enough to even produce secondary ionization in LXe, resulting in a small charge signal.

Stated otherwise, the expected signal will have a spectrum similar to that of the detector-related single electron background. This present the unique opportunity to use the detector's extreme sensitivity to such small charge signals in order to search for a signal that will originate from a possible interaction of dark matter with electrons.

4.2 Particle candidates for a light dark matter

A class of candidate particles for light dark matter is derived from a category of theories beyond the standard WIMP paradigm, referring to a light dark matter with mass in the region of MeV. In these theories, DM resides in a, so-called, hidden sector [239, 240, 241, 242] that possibly is provided with a new gauge group, independent from the Standard Model (SM). In such a case, for example, the interaction between DM and electrons could be mediated by a gauge boson from an Abelian group of the hidden sector that couples with the charged SM particles via the mechanism of kinetic mixing with the photon [243].

The concept of a mirror dark matter (a general review can be found in [244]) is a typical example of such a class of theories. In the framework of this theory the content of the hidden sector is a copy of the SM and the mass of the particles is degenerated among the two sectors. The two sectors (SM and hidden sector) have isomorphic Lagrangians

and the particles of the hidden sector interact with each other with exactly the same forces of the SM. That is, the hidden sector gauge group is exactly $SU(3) \times SU(2) \times U(1)$. So these theories are based on generic Lagrangians of the form:

$$\mathcal{L} = \mathcal{L}_{SM} + \mathcal{L}_{DM} + \mathcal{L}_{int}$$

where $\mathcal{L}_{SM}$ and $\mathcal{L}_{DM}$ are both identical to the SM and here, the former describes the actual SM and the latter the hidden sector. The particles of the SM can interact with the particles of the hidden sector, of course through gravity, but also through $\mathcal{L}_{int}$ with all interactions that are allowed from the requirement for gauge invariance and renormalization.

There is an advantage here compared to the usual models of dark matter, for example the WIMP, emerging from supersymmetry or the theories known as "large extra dimension". Because, for example, the "mirror fermions" of the hidden sector which interact with a zero mass gauge boson (like for example the dark photon of the dark QED which resides in $\mathcal{L}_{DM}$) can serve as models where the auto-interaction of dark matter is not negligible. This is in contrast with WIMPs and the Λ-CDM which considers dark matter to be collisionless. This is important because, despite the fact that Λ-CDM is perfect on a large scale, still suffers from the so-called "small scale crisis" (problem of dwarf galaxies, diversity among the real curves of rotation of galaxies in opposition to simulations based on the premises of the Λ-CDM, evidence of a nucleus in the center of dwarf galaxies, to name a few). On the other hand, mirror dark matter asymptotically, on a large scale, reproduces the results of Λ-CDM correcting its defects in the small scale.

From a purely theoretical point of view, hidden sector theories are extremely attractive as they fully restore the Poincaré group as the fundamental symmetry of the theory. That is to say, for the first time, translations in space-time, rotations in space-time[1], time reversals and parity, are simultaneously symmetries of the theory. This is philosophically and aesthetically ideal, because in the physics of the SM there is a violation of P in weak interactions. In the theory of the hidden extension of SM, on the other hand, one can obtain such an extension allowing for left-handed fermions, which are $SU(2)_L$ doublets to be transformed into right-handed hidden sector fermions that belongs to doublets of a new gauge group $SU(2)'_R$. Conversely, the right-handed fermions of the SM, which belongs to the singlets of the $SU(2)_L$, are transformed into left-handed fermions of the hidden sector belonging to the singlets of $SU(2)'_R$.

Thus, eventually the hidden sector gauge group will be:

$$SU(3)_c \times SU(2)'_R \times U(1)$$

[1] Three-dimensional rotations and proper and orthochronous Lorentz transformations.

which is identical, as far as the gauge couplings are concerned, with the SM and simply it has a reverse chirality. The complete group of the theory is symmetric under parity and there is no longer a "right-left" distinction in the Universe[2].

Concerning this latter aspect, an interesting observation emerges also from the history of particle physics. In the publication of Lee and Yang in 1956 [245], in which it is shown that there is a violation of parity symmetry in weak interactions, towards the end of the paper some general remarks are given among which one can read:

> *These experiments test whether the present elementary particles exhibit asymmetrical behavior with respect to the right and the left. If such asymmetry is indeed found, the question could still be raised whether there could not exist corresponding elementary particles exhibiting opposite asymmetry such that in the broader sense there will still be over-all right-left symmetry. If this is the case, it should be pointed out, there must exist two kinds of protons p_R and p_L, the right-handed one and the left-handed one.*

That is to say, the first time that the idea of "mirror particles" was formulated was starting with symmetry problems in particle physics. This places hidden sector theories, with respect to dark matter models, at a similar level of theoretical attractiveness as supersymmetry. Because the latter appears to propose the WIMPs which, on the one hand, have suitable characteristics to describe the thermal history of the universe and the physics of the primordial universe and, on the other, independently solve issues emerging from particle physics, like the hierarchy problem. Thus, this historical detail reveals that even hidden sector theories solve problems in cosmology and independently could be inspired by elementary particle physics.

As the experimental results about WIMPs continued to be compatible with the background-only hypothesis and the exclusion limits became increasingly stricter, in recent years there has been a significant increase of interest in the theory of light dark matter that produces nuclear recoils below the threshold of current detectors but could be detected through its leptonic interactions. As a result, there has been a renewed interest for a variety of theories that have been proposed, among which we find:

WIMPless dark matter: In this theory the correct relic abundance does not result from the weak mass scale but from gauge-mediated supersymmetry breaking [246, 247].

Asymmetric Dark matter: This model emerges in the context of a more general theory of leptogenesis, trying to describe the dark matter and lepton asymmetries. Here, DM resides in the hidden sector and right-handed neutrinos are coupled to both the hidden and the SM sectors. In this model a light dark matter is allowed without violation of other phenomenological constrains [248].

MeV dark matter from SUSY: A supersymmetric model that can accommodate for the lightest supersymmetric particle with a mass in the MeV scale. A small effective

[2]Where the Universe now is meant to be the SM plus the hidden sector.

coupling between DM and the SM fermions will lead to radiative corrections generating a MeV-scale mass for the DM particle. This model was introduced in order to describe the flux of 511 keV photons from the Galactic Bulge [249] as originating from annihilation of this kind of particles [250, 251].

Cold dark matter of Axinos: Axinos, the supersymmetric partners of axions, have masses that could be much smaller than the supersymmetry-breaking scale. If they are created non-thermally via neutrino decays, they could be a candidate for cold dark matter [252, 253, 254].

4.3 Calculation of the direct detection rates

We will now proceed to the presentation of the method used to calculate the differential event rate, as a function of the energy of the electron recoil. This calculation is significantly more difficult than the corresponding calculation of the rates of nuclear recoils, as we must keep in mind that the electrons are bounded in a multi-electron system. Therefore, there is a, non-relativistic, quantum mechanical perspective, driving the complete dynamics of the system. This, makes its appearance in two key points. First, the electron lies in the atomic orbital of the atom i.e., a state of the atomic Hamiltonian. This must take into account the existence of other electrons and the effect they have on the electron in question. Specifically, the full electron configuration of the Xe atom is:

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2 5p^6 \tag{4.5}$$

This is a typical quantum mechanical problem of a many-fermion system, whose solution we must use here in order to have a valid description of the initial wavefunction of the bounded atomic electron.

Second, after the collision with the DM particle, the ionization electron will be in an unbound state of this multi-fermion system. In a first approach, its final state can be considered as a plane wave. However, this will be distorted by an effective nuclear charge i.e, a charge eZ screened by the rest of the atomic electrons. We will examine in detail how we can properly take these effects into account.

Following the model-independent approach of [255], we will parameterise the elastic scattering cross section in the following way:

$$\bar{\sigma}_e = \frac{\mu_{\chi,e}^2 |\mathcal{M}(q = \alpha m_e)|^2}{16\pi m_\chi^2 m_e^2} \tag{4.6}$$

Where $\mathcal{M}(q)$ is the matrix element for an elastic scattering of a **free** electron with a dark matter particle, which is a function of the momentum transfer. In (4.6), $\bar{\sigma}_e$ is defined for a fixed momentum transfer of $q = \alpha m_e$, with $\alpha \approx 1/137$ the fine-structure constant. This fixed quantity is chosen, for the momentum transfer, because its order of

magnitude is relevant for atomic processes. Then, the full momentum dependence of $\mathcal{M}$ can be parameterised through a function called the **dark matter form factor**, $F_{DM}(q)$:

$$|\mathcal{M}(q)|^2 = |\mathcal{M}(q)|^2|_{q=\alpha m_e} \times |F_{DM}(q)|^2$$

The precise functional form of $F_{DM}(q)$ depends on the specific interaction. Three cases are of interest, concerning the phenomenology of the relic abundance of light dark matter, relating with the results of direct detection through ionization of the target atom:

- $F_{DM}(q) = 1$, where the scattering can be approximated as a point-like interaction, for example, resulting from a heavy vector mediator exchange.
- $F_{DM}(q) = \frac{\alpha m_e}{q}$, corresponding to an electric dipole moment coupling.
- $F_{DM}(q) = (\frac{\alpha m_e}{q})^2$, where the interaction is done via the exchange of an ultra-light vector mediator.

It can be shown that the velocity-averaged cross section for the ionization of an atomic electron, situated in the atomic orbital and described by the quantum numbers (n,l), can be written as:

$$\frac{d\langle\sigma^{nl}_{ion}u\rangle}{d\ln E_R} = \frac{\bar{\sigma_e}}{8\mu^2_{\chi,e}} \int q|f^{nl}_{ion}(p',q)|^2|F_{DM}(q)|^2\eta(u_{\min})dq \tag{4.7}$$

where $\eta(u)$ is the mean-inverse speed, defined in (3.6) and u_{min} is the minimum speed required for a DM particle of mass m_χ in order to ionize an electron from its initial bound state, with binding energy $|E^{nl}_{\text{binding}}|$. This $u_{\min}$ is a function of the momentum transfer q and is given by:

$$u_{min} = \frac{|E^{nl}_{\text{binding}}| + E_{er}}{q} + \frac{q}{2m_\chi} \tag{4.8}$$

In (4.7), the fact that the electron is in an atomic bound state is captured by the so-called **ionization form factor** $|f^{nl}_{ion}(p',q)|$. This function essentially constitutes the amplitude of the quantum mechanical process $|\langle\text{unbound}|e^{iqx}|\text{bound}\rangle|$. To calculate it, the wavefunctions of the electron must be determined both in its bound and free states. Also, for the amplitude of this transition, it is assumed that the atomic potential is spherical and with filled shells. Thus, $|\langle\text{unbound}|e^{iqx}|\text{bound}\rangle|^2$ will constitute a coherent sum over all possible bound states.

We will not take into account the observed fact that the LXe presents an electronic band structure, like liquid argon [256]. Indeed, absorption spectroscopy, probing the exciton levels in LXe [257], gives the possibility to also measure the band gap energy between the conduction and the valence bands, just like in a semiconductor [258]. Ignoring this band structure picture and treating the Xe atom as being isolated, leads however

to a conservative estimation of the ionization event rates, as the band would lead to a decrease of the ionization energy i.e., in an enhanced charge yield.

It can be shown that $|f_{ion}^{nl}(p',q)|$, concerning the transition from a bound state i to an unbound state, is given from the relation:

$$|f_{ion}^{nl}(p',q)|^2 = \frac{2p'^3}{(2\pi)^3} \sum_{\text{deg.states}} \sum_{l',m'} \left| \int d^3r \phi^*_{p',l',m'}(\mathbf{r}) \psi_i(\mathbf{r}) \exp(i\mathbf{qr}) \right|^2 \tag{4.9}$$

The first sum is over all degenerate initial states that are occupied by the atomic electrons. Here $\psi_i(\mathbf{x})$ is the wavefunction of the bound state i and $\phi_{p',l',m'}(\mathbf{x})$ is the final state of the ionized electron, which is an unbound eigenstate of the atomic Hamiltonian. The energy of the ionization electron is $E_R = p'^2/2m_e$ and the energy spectrum is a continuum of states $\phi_{p',l',m'}(\mathbf{x})$ described by the quantum numbers l' and m'. Asymptotically, for a large enough radius of separation from the ion, the free electron can be approximated by a plane wave.

4.3.1 Evaluation of the ionization form factor

The aforementioned difficulty of the calculation of the event rates is reflected in the function (4.9). It must be calculated in such a way as to take into account, as fully as possible, the dynamics governing the initial and final states of the electrons.

As for the former, these are the bounded eigenstates of the atomic Hamiltonian that can be considered as non-relativistic[3]. It must take into account the effect of the nucleus and all the other electrons on each electron under consideration. For a nucleus with atomic number Z, it can be written in atomic units as:

$$H = -\frac{1}{2}\sum_{i=1} \nabla_i^2 - \sum_{i=1} \frac{Z}{|r_i|} + \sum_{j>i} \frac{1}{|r_{ij}|} \tag{4.10}$$

where $\mathbf{r}_i$ is the position vector of the electron i and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. In (4.10), the length and energy dimensions are restored through the Bohr radius, a_0, and the Hartree constant, $E_h \approx 27.21$ eV. In [259], approximations of these stationary states are calculated for all atoms with Z in the range [2, 54]. The calculation is based on the so called Roothan-Hartree-Fock method [260]. This consists of expanding the radial atomic orbitals R_{nl} as a superposition of Slater-type orbitals (STO):

$$R_{nl} = \sum_k C_{nlk} \frac{(2Z_{lk})^{n_{lk}+1/2}}{a_0^{3/2}\sqrt{2n_{lk}!}} \left(\frac{r}{a_0}\right)^{n_{lk}-1} \exp\left(-\frac{Z_{lk}r}{a_0}\right) \tag{4.11}$$

with n_{lk} the principal quantum number, Z_{lk} the orbital exponent, for the k-th STO and for an azimuthal quantum number l. The orbital expansion coefficients C_{nlk}, are given

[3] An approach that is accurate in our case as dark matter particles of MeV mass having the velocity distribution of the SHM have enough kinetic energy only for the ionization of the electrons of the outer and next-to-outer shells

by the solution of the RHF pair of equations. In [259] are tabulated all the coefficients C_{nlk} for each one of the STO with quantum numbers $n_{lk}l$, which can also be named with the conventional names 1s, 2p, etc.

Regarding the outgoing electron, we've examined different scenarios for its description. In the simplest case, we can expect that an electron with momentum $\mathbf{p}'$ will be adequately described by a plane-wave wavefunction $e^{i\mathbf{p}'\mathbf{r}}$. However, we must expect that the effect of nuclear attraction should be taken into account. A more accurate description results from assuming that the state of the electron is described by the continuum state solution of the Schrödinger equation for a hydrogen-like atom, with an effective charge Z_{eff}:

$$\begin{aligned}\tilde{R}_{p'l} =&(2\pi)^{3/2}(2p'r)^l \frac{\sqrt{\frac{2}{\pi}}\left|\Gamma(l+1-\frac{iZ_{eff}}{p'a_0})\right|\exp(\pi Z_{eff}/2p'a_0)}{(2l+1)!} \\ &\times e^{-ip'r}{}_1\mathrm{F}_1\left(l+1+\frac{iZ_{\text{eff}}}{p'a_0};2l+2;2ip'r\right)\end{aligned} \tag{4.12}$$

where ${}_1\mathrm{F}_1$ is the confluent hypergeometric function [261]. Initially we tried to calculate (4.9) for the simple case where the wave function of the outgoing electron could be described by the asymptotic behaviour of (4.12) for $r \to \infty$ i.e., exactly the plane-wave $\tilde{R}(\mathbf{r}) = e^{i\mathbf{p}'\mathbf{r}}$. In that case, we can correct for the fact that we ignore the effect of the nucleus in the same way that, in the theory of beta decay, the differential rate is enhanced by the so-called Fermi-factor $F(p, Z_{\text{eff}})$. This factor arises from the ratio of the asymptotic behavior of (4.12) close to the nucleus and its asymptotic behavior at infinity [262]. Here we find:

$$F(p', Z_{\text{eff}}) = \left|\frac{\tilde{R}_{p'l}(r=0)}{\tilde{R}_{p'l}(r=\infty)}\right|^2 = \frac{2\pi Z_{\text{eff}}}{p'a_0}\frac{1}{1-\exp(-\frac{2\pi Z_{\text{eff}}}{p'a_0})} \tag{4.13}$$

Thus, now we could correct for the plane-wave approximation with an enhancement of the final event rates using (4.13). In the plane-wave approximation (4.9) reduces to:

$$|f^{nl}_{ion}(p', q)|^2 = \frac{(2l+1)p'^2}{4\pi^3 q}\int_{|p'-q|}^{p'+q} k|\chi_{nl}(k)|^2 dk \tag{4.14}$$

where $\chi_{nl}(k)$ is the momentum-space radial wavefunction of the bound state. This can be evaluated analytically using (4.11):

$$\begin{aligned}\chi_{nl}(p) =&\sum_k C_{nlk}2^{-l+n_{lk}}\left(\frac{2\pi a_0}{Z_{lk}}\right)^{3/2}\left(\frac{ipa_0}{Z_{lk}}\right)^l\frac{(1+n_{lk}+l)!}{\sqrt{2n_{lk}!}} \\ &\times{}_2\mathrm{F}_1\left[\frac{1}{2}(2+l+n_{lk});\frac{1}{2}(3+l+n_{lk});\frac{3}{2}+l;-\left(\frac{pa_0}{Z_{lk}}\right)^2\right]\end{aligned} \tag{4.15}$$

where ${}_2\mathrm{F}_1$ is the hypergeometric function. From (4.11) and (4.14) we can now evaluate the form factor in this approximation. In Fig. (4.3) we give the plots of $|f^{nl}_{ion}(p', q)|^2$ as a

function of q, for a recoil energy $E_R = p'^2/2m_e = 200$ eV, for the atomic orbitals (4s, 4p, 4d, 5s, 5p) of the Xe atom.

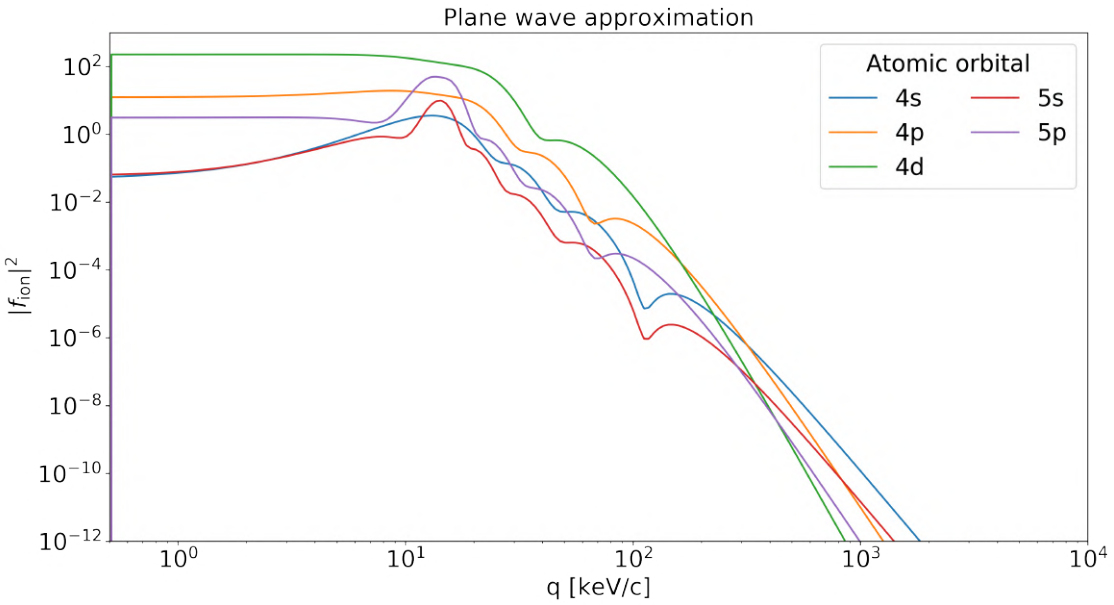


Figure 4.3 – Evolution of the ionization form factor calculated from (4.14) approximating the outgoing electron as a plane wave. For this plot, the recoil energy of the electron was set to 200 eV. The form factor is calculated for all the atomic orbitals of the outer and next-to-outer shells of Xe.

Nevertheless in [263] it is emphasized that such an approach is adequate only for large momentum transfers. This is due to the fact that, in this case, (4.12) is closer to its asymptotic form, but for small q, the correction provided from (4.13) is not adequate for large atoms like Xe as their outer atomic orbitals are far from the asymptotic limit $r \to 0$. Thus the reproduction of the full wavefunction at the origin is not an adequate compensation for the use of the plane wave approximation. It is therefore necessary to use the full hydrogen-like wavefunction of the continuum spectrum (4.12), which is computationally much more demanding.

In ([263]) the following method is proposed: We can treat the bound state R_{nl} as if originating from a Coulomb potential $V_{\text{eff}} = Z^{nl}_{\text{eff}}/r$, with effective charge Z^{nl}_{eff}. Then we can calculate Z^{nl}_{eff} by matching this hydrogen-like energy spectrum, with the RHF spectrum of bound states of [259], thus:

$$-(13.6 \text{ eV}) \times \frac{(Z^{nl}_{\text{eff}})^2}{n^2} = E_{\text{RHF}} \tag{4.16}$$

Finally, this Z^{nl}_{eff} can be used in (4.12) in order to evaluate a more accurate form factor. To do so, we can first develop the factor $e^{i\mathbf{qr}}$ in (4.9) in a linear combination of spherical harmonics, using the plane wave expansion:

$$e^{i\mathbf{qr}} = 4\pi \sum_{L} \sum_{M=-L}^{L} i^L j_L(qr) Y_{LM}(\hat{q})^* Y_{LM}(\hat{r}) \tag{4.17}$$

Where j_L is the spherical Bessel function. Now the integral of (4.9) is written as:

$$\begin{aligned}
&\int d^3x \phi^*_{p'l'm'}(\mathbf{r}) \psi_i(\mathbf{r}) \exp(i\mathbf{qr}) = \\
&4\pi \sum_{L} \sum_{M=-L}^{L} i^L \overbrace{\int dr r^2 j_L(qr) \tilde{R}^*_{p'l'} R_{nl}}^{I_L} \times Y_{LM}(\hat{q})^* \times \int d\Omega Y_{lm}(\hat{r}) Y_{l'm'}(\hat{r})^* Y(\hat{r})_{LM}(\hat{r}) = \\
&4\pi \sum_{L} \sum_{M=-L}^{L} i^L I_L \times Y_{LM}(\hat{q})^* \times (-1)^{m'} \sqrt{\frac{(2l+1)(2l'+1)(2L+1)}{4\pi}} \begin{bmatrix} l & l' & L \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} l & l' & L \\ m & m' & M \end{bmatrix}
\end{aligned} \tag{4.18}$$

Where we denoted the radial integral as I_L and the angular integral of the product of three spherical harmonics is evaluated using the Wigner 3-j symbol $[\cdots]$ [264]. Now we can proceed to perform the sum of (4.9), starting from the double sum over m' and m that can be simplified using the orthogonality relation of the 3-j symbol. Care must be taken with the summation indices in the square of (4.18). We will prime the sum indexes L and M that enters the square:

$$\begin{aligned}
&\sum_{m,m'} \left| \int d^3r \phi^*_{p'l'm'}(\mathbf{r}) \psi_i(\mathbf{r}) \exp(i\mathbf{qr}) \right|^2 \\
&\quad = \sqrt{4\pi}(2l+1)(2l'+1) \sum_{L,M} \sum_{L',M'} I_L I_{L'} Y_{LM}(\hat{q})^* Y_{L'M'}(\hat{q}) \times \\
&\quad \begin{bmatrix} l & l' & L \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} l & l' & L' \\ 0 & 0 & 0 \end{bmatrix} \left(\sum_{mm'} \begin{bmatrix} l & l' & L \\ m & m' & M \end{bmatrix} \begin{bmatrix} l & l' & L' \\ m & m' & M' \end{bmatrix} \right) \\
&\quad = \sum_{L} (2L+1)(2l+1)(2l'+1) \begin{bmatrix} l & l' & L \\ 0 & 0 & 0 \end{bmatrix}^2 \left| \int dr r^2 j_L(qr) \tilde{R}^*_{p'l'} R_{nl} \right|^2
\end{aligned} \tag{4.19}$$

where use has been made of the orthogonality relations:

$$\sum_{mm'} \begin{bmatrix} l & l' & L' \\ m & m' & M' \end{bmatrix} \begin{bmatrix} l & l' & L \\ m & m' & M \end{bmatrix} = \frac{1}{2L+1} \delta_{LL'} \delta_{MM'} \tag{4.20}$$

and the summation identity satisfied by the spherical harmonics:

$$\sum_{M=-L}^{L} Y_{LM}(\hat{q})^* Y_{LM}(\hat{q}) = \frac{2L+1}{4\pi} \tag{4.21}$$

Finally we evaluate (4.9) as a function of the unbound state $\tilde{R}_{p'l'}(r)$ by summing over l':

$$|f_{ion}^{nl}(p',q)|^2 = \frac{4p'^3}{(2\pi)^3} \sum_{l'} \sum_{L=|l'-l|}^{l'+l} (2l+1)(2l'+1)(2L+1) \begin{bmatrix} l & l' & L \\ 0 & 0 & 0 \end{bmatrix}^2 \left| \int dr r^2 \tilde{R}_{p'l'}(r) R_{nl}(r) j_L(qr) \right|^2 \tag{4.22}$$

But here, we will make use of an even more accurate method to determine $R_{p'l}$ and then use (4.22) to evaluate the form factor. This will not even rely on (4.12) but, rather, on a Hartree-Fock method in which the potential $V_{eff} = Z_{eff}^{nl}/r$ is determined as explained above, but in order to then solve numerically the Schrödinger equation for the unbound states. The ionization form factors, calculated by this Hartree-Fock method, was given to our Collaboration by the authors of [265]. In Fig. (4.4) (solid curves) we can see the form factors calculated from (4.22) compared to those based on the plane wave approximation (dashed curves). We notice that the two approaches are very close for high momentum transfers where, however, we have a strong suppression by many orders of magnitude, while for small values of q the Hartree-Fock computation of the unbound state gives a more conservative result.

4.3.2 Calculation of the rate of events

Finally, we can proceed to the computation of the expected event rates for the ionization of a Xe atom via collision with a leptonically interacting dark matter particle. The differential event rate is proportional to the coherent sum of the thermally averaged differential cross section (4.7), for every atomic orbital:

$$\frac{dR}{d\ln E_R} = N_T \frac{\rho_\chi}{m_\chi} \sum_{n,l} \frac{d\langle \sigma_{ion}^{nl} u \rangle}{d\ln E_R} \tag{4.23}$$

where ρ_χ is the local dark matter density and N_T is the number density of Xe atoms in the target. As already mentioned, we will not extend the sum of (4.23) over all Xe orbitals. We will consider only the shells $n=5$ and $n=4$. The relation (4.23) can be made more explicit concerning the order of magnitude of the rate, if we write it for a reference cross section of, say, 10^{-40} cm^2:

$$\frac{dR}{d\ln E_R} = \frac{6.2}{A} \frac{\text{events}}{\text{kg-day}} \left(\frac{\overline{\sigma}_e}{10^{-40}\text{cm}^2} \right) \left(\frac{10\text{MeV}}{m_\chi} \right) \frac{d\langle \sigma_{ion} u \rangle / d\ln E_R}{10^{-3} \overline{\sigma}_e} \tag{4.24}$$

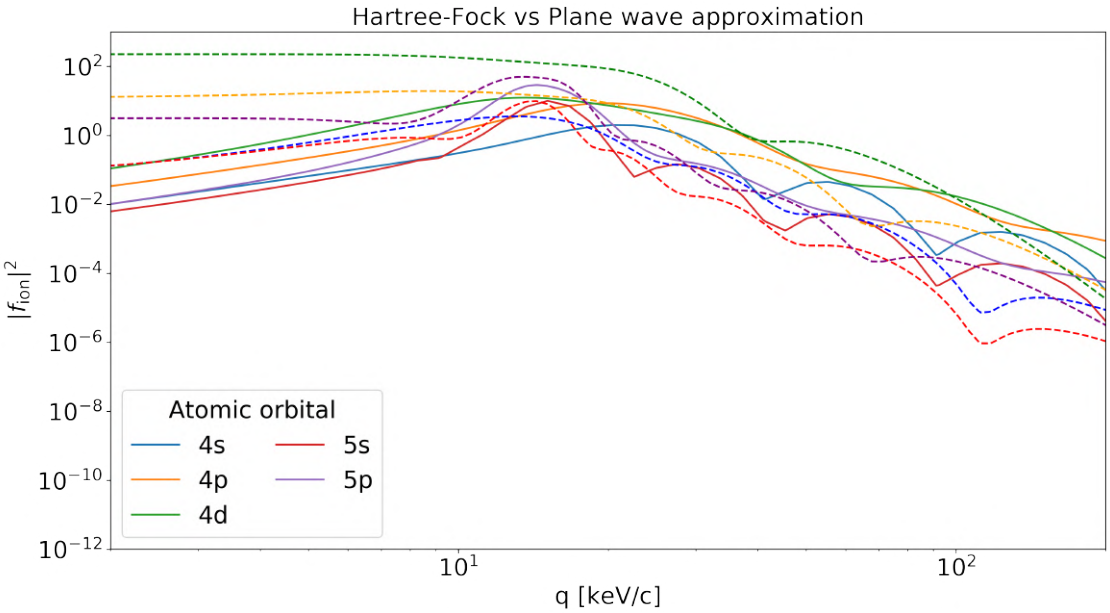


Figure 4.4 – Comparison between the ionization form factors computed with the plane-wave approximation (dashed curves) and the Hartree-Fock method (solid curves). We can see that, for small momentum transfers, the Hartree-Fock method takes into account the effect of the nuclear attraction on the electron, resulting in a reduced value of the form factor. For large momentum transfers the two methods are close but this happens in the region of strong wavefunction suppression.

In Fig. (4.5), we show the event rates as a function of the recoil energy, for two different masses of the dark matter particle and for a cross section of $\overline{\sigma}_e = 10^{-36}$ cm^2. The $F_{DM} = 1$ was used corresponding to an interaction via a heavy vector mediator exchange. Also are show the individual contributions to the rate by each one of the atomic orbital considered, where we can understand the choice we've made for the use of only the higher atomic orbitals.

In Fig. (4.6), the same plots are shown but for an ultra-light mediator of the interaction ($F_{DM} = 1/q^2$). We see a suppressed rate of events for larger energy depositions due to the particular dependence of the dark matter form factor on the momentum transfer.

4.4 Detector responce

We should now take into account the response of LXe to the recoiling electron and fold this response to the signal, in order to have an estimation of the number of electrons that we could observe. As explained earlier, the ionization electron itself, if it "survives" the drift and if extracted to the GXe, it could be detected as a single electron. But if its energy E_R is larger than $W = 13.8$ eV then, following the theoretical understanding of the

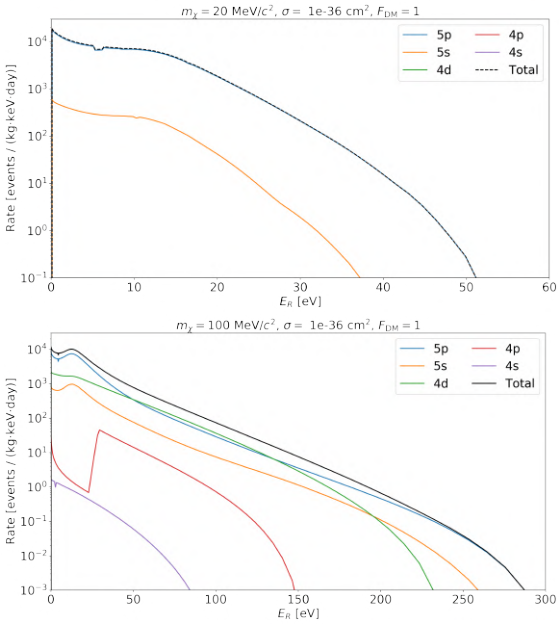


Figure 4.5 – Event rates as a function of the recoil energy, calculated from (4.23), for $F_{DM} = 1$ and $\overline{\sigma}_e = 10^{-36}$ cm^2. We show the rates for two values of m_χ, 20 MeV/c^2 (left) and 100 MeV/c^2 (right). We also show the individual contributions to the total rate of each atomic orbital of the outer and next-to-outer shells. We can see that for $m_\chi = 20$ MeV/c^2 practically the only contribution comes from the 5p orbital. For $m_\chi = 100$ MeV/c^2 also the lower orbital are contributing but they are strongly suppressed. This justifies the use of only the outer and next-to-outer shells for sub-GeV DM masses.

ionization process of an electron recoiling in the LXe, this should create a certain number of secondary electrons. We will use the model proposed in [265], in order to convert the recoil energy of the ionization electron in the number of electrons that will eventually be observed. It should be noted that this model describes the charge yield process as it is understood for electron recoils but the smallest energy for which actually exist a measurement is 186 eV, Fig. (1.15), [157]. However, we can assume that this model could describe the charge yield mechanism in the same way even for smaller deposited energies, based on the plateau of the ER charge yield in the credible interval for XENON1T, Fig. (1.15), namely in the constant value of 73 quanta/keV produced.

Having made this remark, we can assume that the ionization electron with energy E_R, should produce $n^{(1)} = \text{Floor}(E_R/W)$ additional quanta. There is a probability [266]:

$$f_e = (1 - f_R)(1 + N_{ex}/N_i)^{-1} = 0.87 \tag{4.25}$$

that these quanta will be electrons, where f_R is the recombination probability. In the complementary percentage of cases, where they will be photons, these would escape our

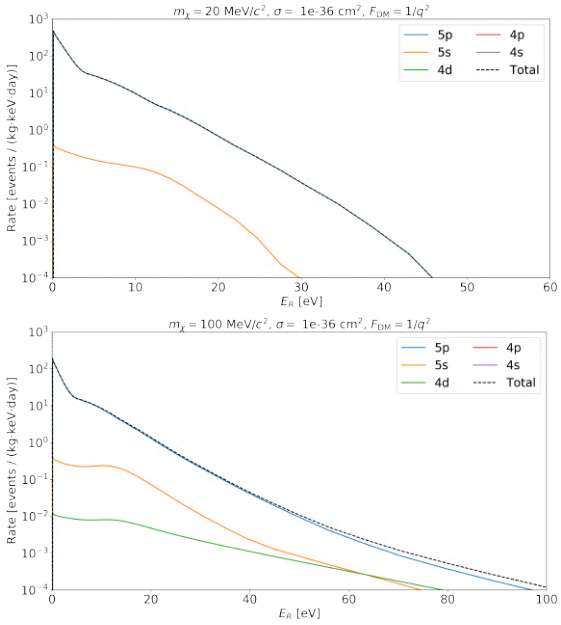


Figure 4.6 – Event rates for $F_{DM} = 1/q^2$ for $m_\chi = 20$ MeV/c^2 (left) and $m_\chi = 100$ MeV/c^2 (right). Also are shown the contributions of each orbital, where in contradistinction with Fig. (4.5), we observe a smaller contribution from lower orbitals due to the dark matter form factor momentum transfer suppression.

S2-only analysis as they will pass unobserved. The recombination probability concerning the original and the additional electrons created, will have a nominal value of $f_R = 0.01$. If the DM particle scatters off an electron of an inner shell, then in addition to the above procedures there will be a supplementary production of secondary quanta, due to the photons that will be created by the electronic transitions that will follow the ionization. The number of these secondary quanta will be $n^{(2)} = \text{Floor}((E_i - E_j)/W)$, where $E_i - E_j = \Delta E$ is the difference of the binding energies between the initial and final states of the transitions. In Tab. (4.1) we reproduce a calculation of the number of these additional quanta made in [265]. This, also takes into account the possibility of multiple available inner shells for the de-excitation of the outer shell electrons after the ionization. To be conservative we will make only use of the minimum number given.

We can now write the probability of production of n electrons due to the primary ionization electron, as the sum of two probabilities, P_1 and P_2 for the following mutually exclusive events:

1. The probability P_1 that the primary electron will recombine AND the probability that the $n^{(1)} + n^{(2)}$ quanta generated will be n electrons:

$$P_1 = f_R \text{Binom}(n | n^{(1)} + n^{(2)}, f_e)$$

Transition energies for Xe shells					
Shell	$5p^6$	$5s^2$	$4d^{10}$	$4p^6$	$4s^2$
Binding energy [eV]	12.4	25.7	75.6	163.5	213.8
Photon energy [eV]	-	13.3	63.2	87.9	201.4
Additional quanta	0	0	4	6-10	3-15

Table 4.1 – Additional quanta that will be produced by a γ from the de-excitation of electrons in the outer shell, after the ionization of an electron in an inner shell. Also shown are the binding energies of the shells and the photon energies. A range of additional quanta is given, concerning the possibility that the outer shell electrons can transition to more than one available inner shells through the whole process of the cascade.

2. The probability P_2 that the primary electron will not recombine AND the probability that the $n^{(1)}+n^{(2)}$ quanta generated will be n-1 electrons:

$$P_2 = (1-f_R)\text{Binom}(n-1|n^{(1)}+n^{(2)}, f_e)$$

Thus, n electrons will be generated with probability[4]:

$$P(n|n^{(1)}, n^{(2)}, f_e, f_R) = \Big(f_R\text{Binom}(n|n^{(1)}+n^{(2)}, f_e) + (1-f_R)\text{Binom}(n-1|N_1+N_2, f_e)\Big) \tag{4.26}$$

Using (4.23), the rate of n produced electrons r_p in the LXe can be written as:

$$r_p = \int \frac{dR}{dE} dE P(n|n^{(1)}, n^{(2)}, f_e, f_R) \tag{4.27}$$

In Fig. (4.7), we can see the rates of events for one, two and three produced electrons as a function of the dark matter mass m_χ for a cross section $\overline{\sigma}_e = 10^{-36}$ cm^2.

But the observed counts should not be compared to these rates, as we didn't include yet the effect of the drift of these electrons i.e., the probability that will be absorbed by electronegative impurities. Also we didn't include the probability of extraction. Concerning the former, the probability $P_{\text{drift}}(\tau_e)$ that the produced electrons will survive the drift to the phase boundary is $e^{-t_{\text{drift}}/\tau_e}$, where t_{drift} is the drift time and τ_e is the electron lifetime. Since these electrons can come from any depth, here we will use an average value for a uniform distribution of electrons in the range $0\ \mu\text{s} \leq t_{\text{drift}} \leq 727\ \mu\text{s}$. Concerning the extraction efficiency, we will use a mean value of $P_{\text{extr}} = 0.95$, inferred from the distribution of Fig. (2.22) over the (x,y) plane for the constant anode voltage of

[4]Note that the probability of recombination for these electrons is taken into account in f_e.

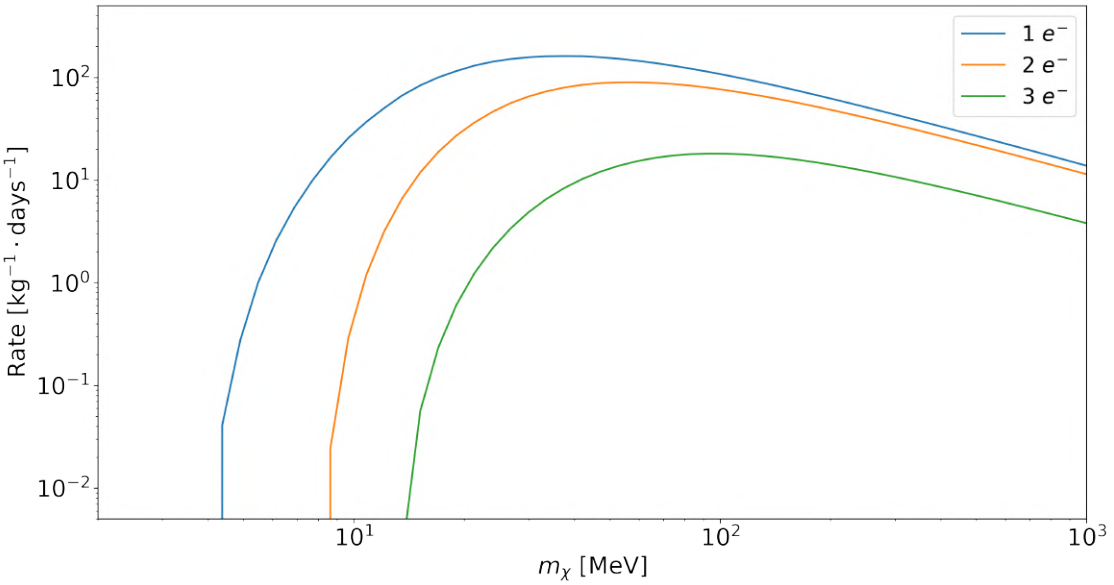


Figure 4.7 – Few electrons event rates from (4.27) as a function of the DM mass. Here we used $F_{DM} = 1$ and $\overline{\sigma}_e = 10^{-36}$ cm^2, thus the underlying DM-e$^-$ scattering event rates correspond to those of Fig. (4.5).

4 kV of SR0 and SR1. We fold these effects in the signal r_p writing the rate of n detected electrons $r_d(n)$ as[5]:

$$r_d(n) = \int \frac{dR}{dE} dE \sum_k P(k|n^{(1)}, n^{(2)}, f_e, f_R) \text{Binom}(n|k, P_{\text{drift}}(\tau_e) \times P_{\text{extr}}) \tag{4.28}$$

In Fig. (4.8) we can compare the event rates of (4.27) and (4.28). We notice that for small values of m_χ the event rate, of course, decreases due to the electron drift. The reduction is greater as more electrons are produced, because the binomial process of equation (4.28) can convert a 3 electron signal into 2 or 1 observed electrons and a 2 electron signal into 1 observed electron. But as m_χ increases and the momentum transfer of the scattering is larger, there is enough energy available to generate events with more than 3 electrons to feed, through the binomial process of (4.28), the 1, 2 and 3 expected electron rates and thus mitigate their drift loses or even increase their rates for $m_\chi > 200$ MeV. Also, in Fig. (4.9), we can compare the expected rates between $F_{DM} = 1$ and $F_{DM} = 1/q^2$.

[5] We can read this formula (4.28) as an expression of the probability that a DM-electron scattering take place AND that the ionization electron produces k electrons in LXe AND that, after drift and extraction, this results in n observed electrons.

Now, the observed events corresponding to one, two, etc. observed electrons of a single electron spectrum, obtained from science run datasets, can be compared to the rates from (4.28) and a limit can be set on $\overline{\sigma}_e$, for various values of m_χ.

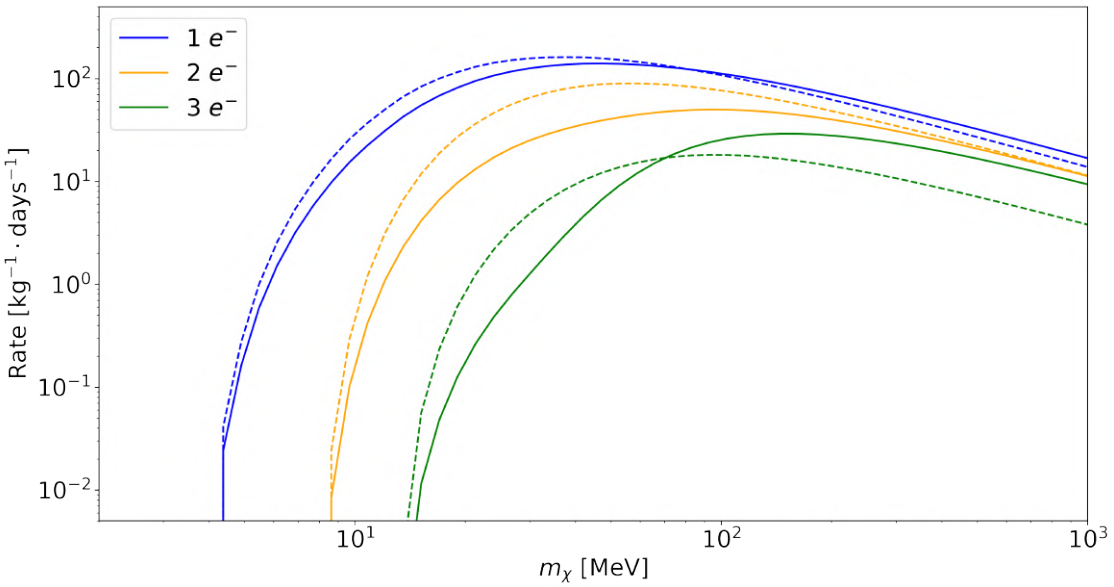


Figure 4.8 – Few electrons event rates, as a function of m_χ, from (4.27) (dashed curves) compared to the rates from (4.28) (solid curves), where the drift and extraction effects are taken into account. Here we used again $F_{DM} = 1$ and $\overline{\sigma}_e = 10^{-36}$ cm^2

4.5 Data selection and analysis cuts

We already explained in (2.1) the problems arising from our ignorance of the single electron (SE) background, that is dominant precisely in the spectral region of the few electrons we expect from a sub-GeV DM interaction. The fact that we cannot eliminate, in a hardware manner, the SE background and that this is extremely high, especially after a main S2, forces us to search for subsets inside the registered waveforms, where we can reasonably expect this background to be smaller. Specifically, our current understanding of the SE background is related to two main sources. The first, concerns the photoionization of impurities, or the photoelectric effect on the metal surfaces, by the light of a main S2 event. This source is well understood, in terms of its temporal characteristics, from studies originally done in XENON100 [106] and repeated here, Fig. (2.13), in order for us to quantify this photoionization region and to proceed in a characterization of the detector. The second is related to, so called, SE trains. These are SEs that can be continued for times scales of $\mathcal{O}$(ms) after a main S2. These kind of trains, that we observed and were also observed by the LUX Collaboration [267], are characterized by a

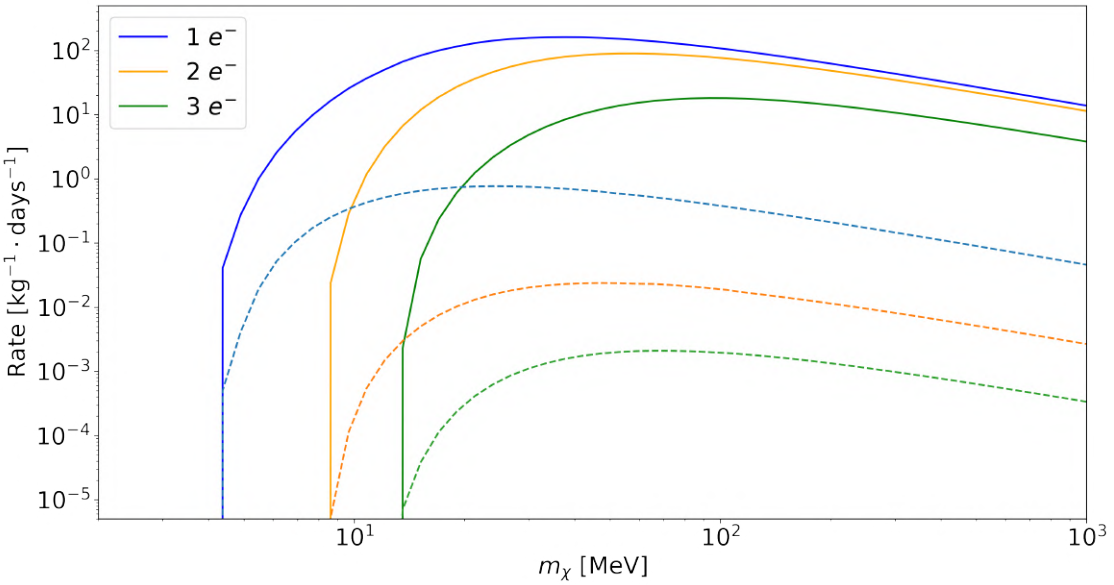


Figure 4.9 – Few electrons event rates, as a function of m_χ, from (4.28) for $F_{DM} = 1$ (solid lines) and $F_{DM} = 1/q^2$ (dashed lines). Here again, $\overline{\sigma}_e = 10^{-36}$ cm^2. We see that the expected rates, for the case of an ultra-light mediator of the interaction, are reduced. We also see that, in this case, the total rate is dominated by the rate of one-electron events.

monotone decrease over such long time scales. This decrease that seems to follow a power law. An important observation is that there is a strong correlation of the positions of these SEs with the position of the preceding S2, as well as, with the purity of LXe, Fig. (4.10). A hypothesis, originally proposed in [268], interpret this observation as being due to delayed emission of electrons, initially captured by impurities. Part of the charge losses due to drift may form negatively charged ions that, due to their low mobility, preserve the initial position of the cloud and, at time-scales of ms, could release the captured electron via certain atomic processes. Also, the theoretical model of section 2.12, reveals that, during the extraction process, there is a percentage of electrons of the main cloud that are not extracted immediately, but they are bouncing in the potential just below the phase separation. They could try multiple times to cross the dielectric boundary and, if they do not succeed, they will eventually thermalize in the minimum of the potential. If we can assume that this model is correct then it could be another mechanism of the delayed emission.

In any case, the empirical observations for the two main sources of the SE background, Fig. (2.13) and Fig. (4.10) guide us about the region of waveforms of the events on which we could focus the search for light DM, aiming for a reduced rate of SEs. First of all, we must definitely avoid the region of one maximum drif time after a main S2, where SEs from photoionization are dominant. Also, we could not even use the rest of the time

region until the end of the event, due to the delayed emission that extends up to ms time scales.

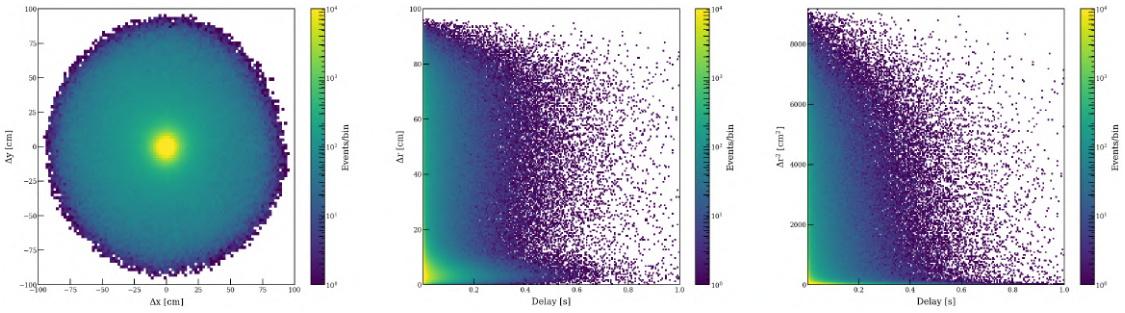


Figure 4.10 – Spatial characteristics of the single electrons after a main S2. Left: Distribution of the differences of (x,y) positions between the SE and the preceding main S2. A clear peak is indicated in the region of Δx, $\Delta y = (0,0)$ revealing the position correlation. Center-Right: Distribution of $\Delta r = \sqrt{(x_{SE} - s_{S2})^2 + (y_{SE} - y_{S2})^2}$ as function of the time separating the SEs from the main S2.

It is also necessary to not even use the, at most ~ 720 μs, region between the main S1 and main S2, as the S1 light can produce single electrons via a photoelectric effect on the metal surfaces (mostly the grounded gate).

We are obliged to limit ourselves to the pre-trigger part of the waveform. The only single electrons that may exist there, are those that come from the mechanism of delayed emission due to the main S2 of the previous event, which may be hundreds of ms away from the event under consideration. Thus, a very strong cut that we will use, concerns the time separation Δt of the single electron peak from the main S2 of the previous event. It goes without saying that this, very powerful cut, will have a dramatic consequence on the exposure of our search. Even without this cut, the waste of exposure, compared with that available for a search using the main S1 and S2 is immense, just due to the fact that small charge signals cannot trigger an event. To give an example, for the average trigger rate of ~ 5 Hz of the detector, the livetime of 1 second of an S1-S2 search, give to us only the pre-trigger window of ~ 5 events i.e., approximately 5×1 ms (because the trigger is usually a main S2 that, by default, is placed by PAX at 1 ms, with respect to the start of the event window). If we also take into account the fact that an S1 could be in an average time of 0.4 ms before the S2 that triggered the event, we see that, for the same target mass, the exposure is approximately the 0.3% of the full science run. Thus, the total exposure expected is around 840 kg $\times$ day . On the other hand, this can be considered very large compared to the exposure of 15 kg $\times$ day of XENON10 which has the best limit, to date, for the DM-e^- cross section in the region $[10-30]$ MeV/c^2. But is also evident that, along with the size of the detector, the rate of SEs has also increased. Nevertheless, a number of cuts have been designed to address as much of the background of the SEs as possible, through its current understanding.

4.5.1 Time separation from previous event

We now present the way we obtain the data from the pre-trigger window. The pre-trigger window is considered here to be the region of the waveform before the main interaction, irrespective of the trigger of the event. Thus we can have: **1)** an S2 trigger that has an S1 before it, **2)** an S2 trigger that is lone and **3)** an S1 trigger. In each case, the livetime is calculated as shown in Fig. (4.11).

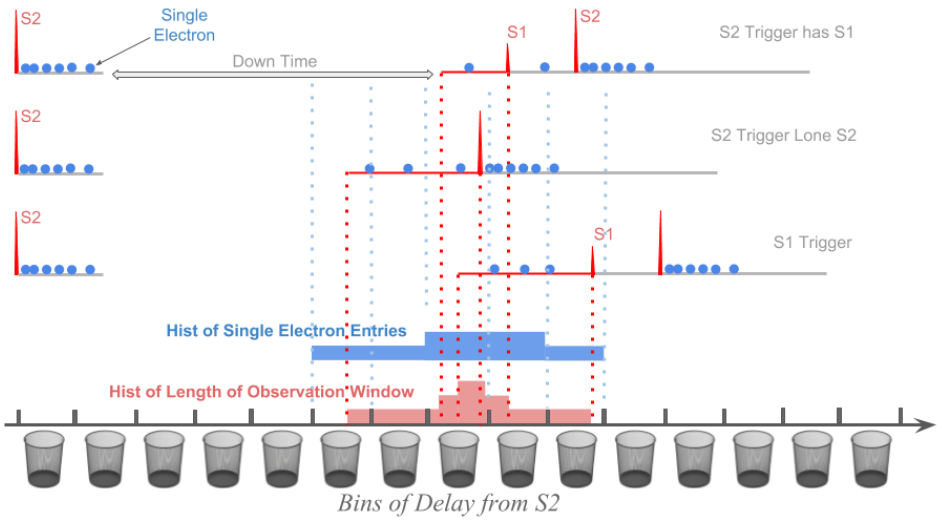


Figure 4.11 – Calculation of the livetime in the pre-trigger window, in various trigger configurations. The calculation is used to correct for the SE rate, when evaluated as a function of the time separation Δt from the main S2 of the previous event, but also to find the actual livetime of the search, when imposing the request of only one peak before any other peak of the main event.

For each SE found, its time separation from the main S2 of the previous event is calculated. The livetime is considered only to be the time between the first peak of the main event and the start of the window. The rate of SEs as a function of Δt can be seen in Fig. (4.12). One can see the background reduction as we move away from the previous S2. Of course, this cut is highly reducing the exposure available thus the analysis aim in a compromise between the exposure left and the suppression of the background of SEs. In the analysis that is currently on-going, we try to optimize this cut in a test dataset of about 8% of the total exposure, before applying the fixed analysis to the full data.

4.5.2 Position Correlation Cut

We are also going to make use of the position correlation with the previous S2, that characterizes the, ms-scale, SE trains. We first try to model the position non-correlated

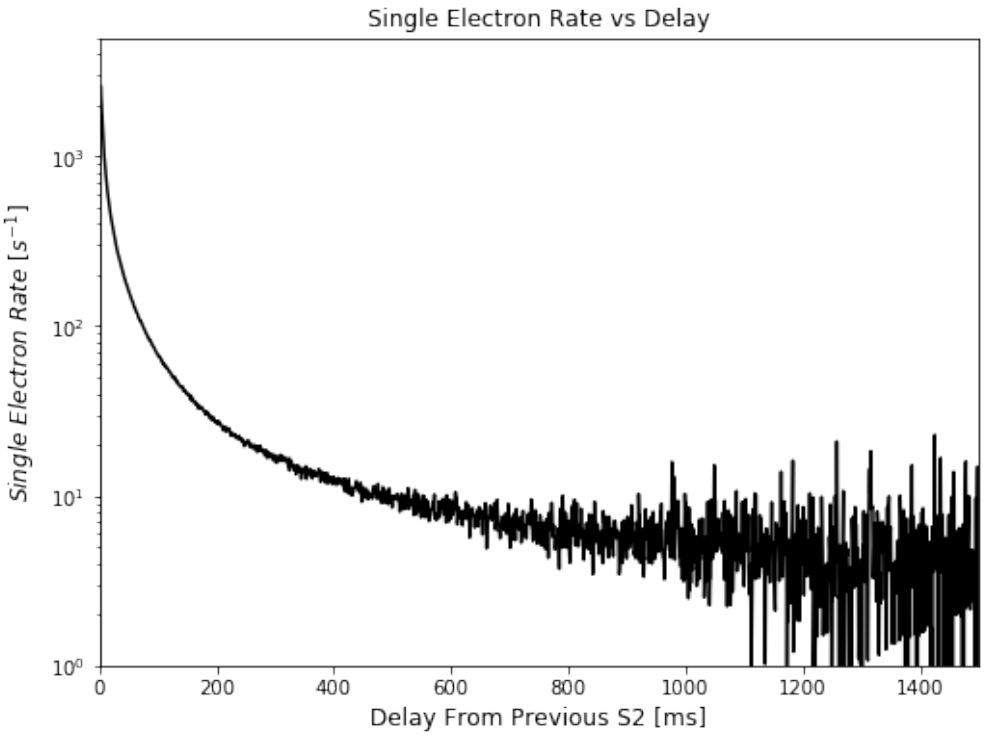


Figure 4.12 – Evolution of the SE rate, as a function of the time separation from the main S2 of the previous event. We see here the effect of those SE trains that extend for hundreds of ms after an S2 and requires the use of a cut in order to be in an region that this background is reduced.

part of the background, that can be seen for large values of Δr in Fig. (4.10). We constructed a toy MC of the distribution of Δx and Δy if there was no correlation, via a random selection of main S2s and SEs, producing a histogram of this (x, y) distances. We then fit this ToyMC distribution in the true distribution of Fig. (4.10), together with a 2D Gaussian that models the central region. The fit on the data is not good but allows us to define a cut in a radius $\Delta r < R_{\text{cut}}$ such that, outside the disk defined by this cut, dominate the non-correlated part of the background. After a systematic study done by the Single electron analysis team, the cut was set at a radius of 15 cm.

We must account for the fact that the exposure now is reduced, in the target mass part. Thus, we must carefully evaluate the fiducial volume that, in the presence of this cut, is limited in the exterior of the cylinder defined by $R_{S2} - R_{\text{cut}} < 15$ cm and -97 cm $< z < 0$ cm, where R_{S2} is the radial position of the main S2 peak of the previous event Fig. (4.13).

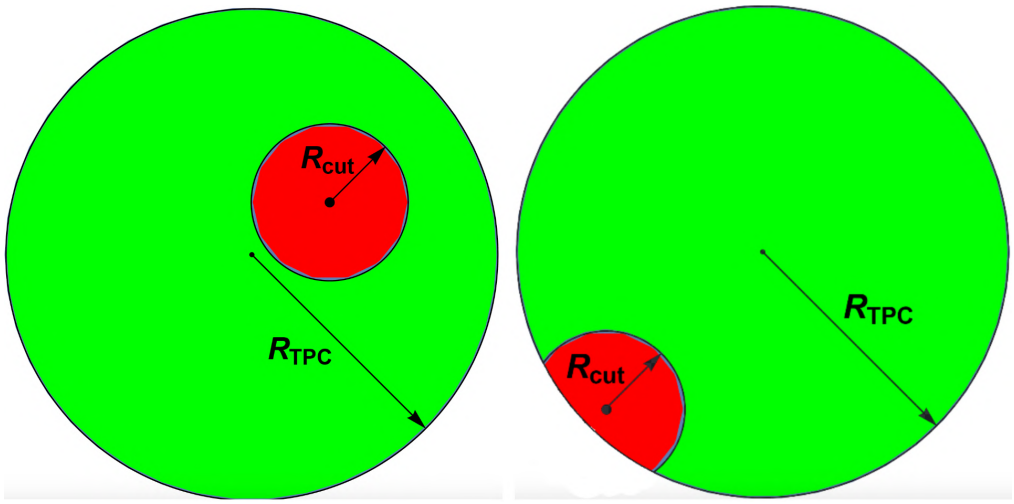


Figure 4.13 – Fiducial volume calculation, taking into account the position correlation cut. For each possible position of the main S2 of the previous event, the fiducial volume is defined by the green area multiplied with the detector length.

The area of the green fiducial surface S_{fid} of Fig. (4.13) is given by:

$$S_{\text{fid}} = \pi R_{\text{TPC}}^2 - A_{\text{lens}}(r_{S2}, R_{\text{cut}}, R_{\text{TPC}}) \tag{4.29}$$

where $A_{\text{lens}}(r_{S2},\ R_{\text{cut}},\ R_{\text{TPC}})$ is the area of the asymmetric "lens", formed by the two intersecting circles, the boundary of the TPC and the circle defined by the boundary of the cut, with its center situated at a distance r_{S2} from (0,0) (Fig. (4.13 right). $A_{\text{lens}}(r_{S2},\ R_{\text{cut}},\ R_{\text{TPC}}$ is given from:

$$\begin{aligned} &A_{\text{lens}}(r_{\text{S2}}, R_{\text{cut}}, R_{\text{TPC}}) \\ &\quad = R_{\text{cut}}^2 \arccos\left(\frac{r_{\text{S2}}^2 + R_{\text{cut}}^2 - R_{\text{TPC}}^2}{2 r_{S2} R_{\text{cut}}}\right) + R_{\text{TPC}}^2 \arccos\left(\frac{r_{\text{S2}}^2 + R_{\text{TPC}}^2 - R_{\text{cut}}^2}{2 r_{\text{S2}} R_{\text{TPC}}}\right) \\ &\quad - \frac{1}{2}\sqrt{(-r_{\text{S2}} + r_{\text{cut}} + R_{\text{TPC}})(r_{\text{S2}} + r_{\text{cut}} - R_{\text{TPC}})(r_{\text{S2}} - r_{\text{cut}} + R_{\text{TPC}})(r_{\text{S2}} + r_{\text{cut}} + R_{\text{TPC}})} \end{aligned} \tag{4.30}$$

Let the livetime of the SE event k be T_k, with the preceding S2 being at $r_{S2,i}$, the LXe density ρ and the TPC length h_{TPC}. Then, the total exposure of the search, taking into account this cut, is calculated as :

$$E = \rho \sum_{k=1}^{N_{\text{events}}} T_k h_{\text{TPC}} A_{\text{lens}}(r_{\text{S2},k}, R_{\text{cut}}, R_{\text{TPC}}) \tag{4.31}$$

4.5.3 Machine Learning against the SE pileup background

In contrast to the previous exposure reduction cuts, we designed and applied a cut whose efficiency can be determined in a robust way, being based on the waveform simulator, which employs all our understanding of the signal reconstruction by the detector. This cut aims in a possible mitigation of the background of $n \geq 2$ electrons, dominated by single electron pileups as we saw in Chapter 2.

The idea here is that, electrons originating from a dark matter-electron scattering process, if they are more than one, they must necessarily be spatially correlated. Despite the fact that the rate of single electrons resulting from this process and for a cross-section $\sim 10^{-37}\text{cm}^2$, inside the whole volume will be about 0.2 every second, however, the probability of having two such electrons simultaneously within the time range $[0.85\mu s, 3.5\mu s]$, corresponding to the width of the peak of a pileup of 2 single electrons is very small. Or even inside the range $[0.4\mu s, 10\mu s]$ which is the corresponding width of the peak of a pileup of 3 or 4 single electrons. Thus, we could assume that every 2 or 3-fold coincidence of single electrons is eligible for being produced from a dark matter-electron scattering, if these electrons are spatially correlated, that is, if they originate from the same point. If not, then the corresponding events cannot be derived from such a process and they should be attributed to single electrons coming from different points (x,y) and simultaneously reaching the surface, being extracted to the gas gap only in time coincidence. These two types of events (two or three electrons coming from the same point (x,y) compared to two or three electrons coming from different points) differ in the patterns they will produce on the top PMT array. It is therefore natural to try the use of neural network architectures, in an attempt to see if it is possible to distinguish between the two cases.

Simulation of single electrons hitpatterns

We have simulated hitpatterns of single electrons which will form the training and testing datasets of the neural network. For up to three observed electrons, we distinguish three cases with their corresponding sub-cases:

1. One single electron
2. 2-fold single electron pileup. In this case the two electrons can be derived
   (a) from the same point (x, y)
   (b) from different points (x_1, y_1) and (x_2, y_2).
3. 3-fold single electron pileup. In this case we have three sub-cases:
   (a) they all come from the same point (x_0, y_0)
   (b) Two of them are coming from the same point (x_1, y_1) but the third originates from a different point (x_2, y_2)

(c) Each of them comes from a different point (x_1, y_1), (x_2, y_2), and (x_3, y_3)

Fig. (4.14) presents the S2 distributions of the simulations of each of the five different possible cases:

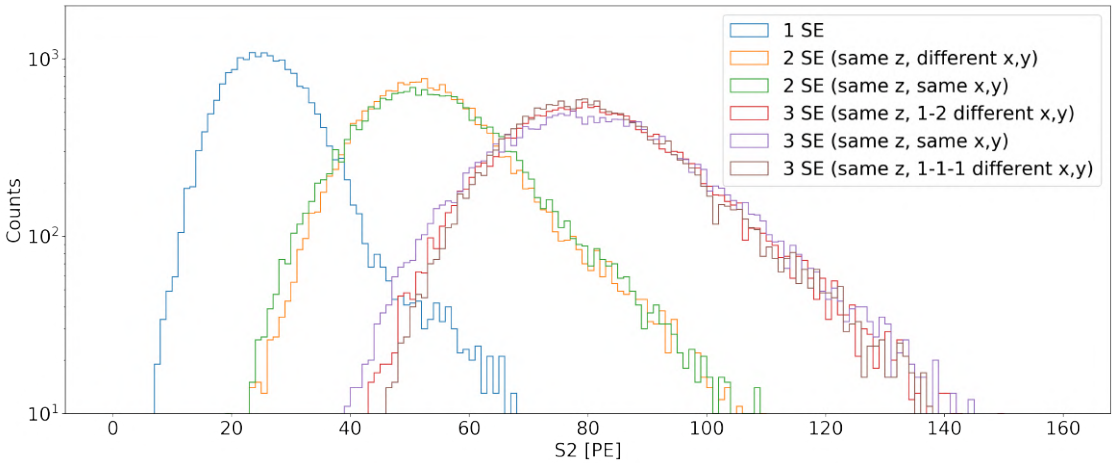


Figure 4.14 – S2 distributions of 10k simulations of each of the 5 different cases studied.

For the simulation I used the full detector configuration of SR1 together with a secondary scintillation yield equal to 28 PE/e^{-1}.

The neural network

We've produced 300k simulations for each of the 5 cases. The idea is to use an architecture that is capable of classifying between the two physically distinct cases to which we will give the labels **0** and **1** respectively:

- **0** all electrons come from the same point.
- **1** at least 1 electron originates from a different point (x, y) than the rest.

It turns out that a very simple neural network, like a fully connected neural network with two hidden layers, can make a fairly accurate classification that can be considered equivalent to the visual perception of an observer. Because the key obstacle to a 100% accurate classification are the similarities between the classes, that will inevitably occur within the training set. For example, the hitpatterns of the cases 2a) and 2b) inevitably will look very much alike, or even become completely indistinguishable, when the distance between the two points $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ becomes very small. And since in the simulation this distance is in the range $0 \leq d < 2R_{TPC}$ it is obvious that misclassification, in cases where d is very small, is inevitable. By the same argument, we can see that, in the same limit of small d, there will also be misclassification between the cases 3a) (label 0) with 3b) and 3c) (label 1).

The neural network architecture we built using TensorFlow [269] is shown in the Fig. (??).

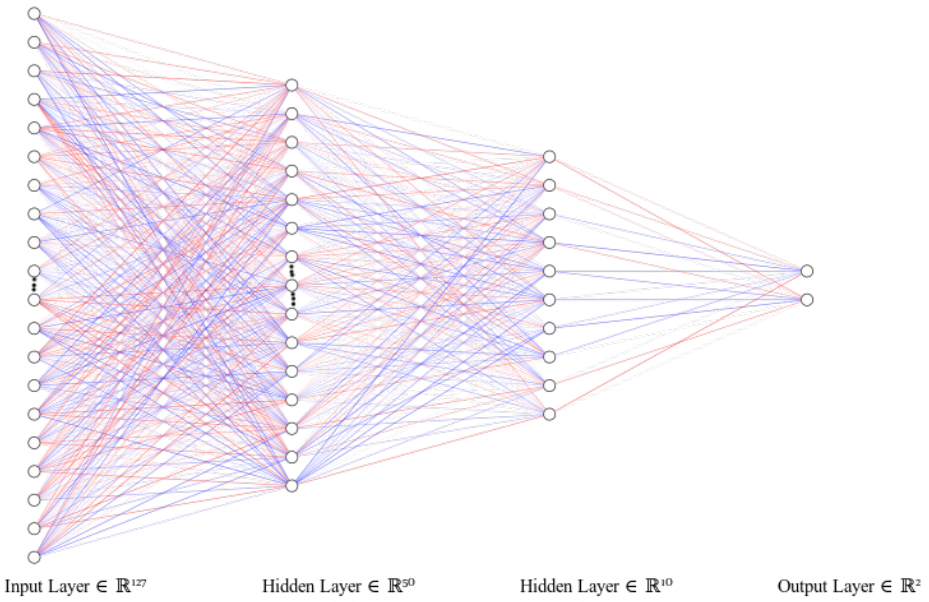


Figure 4.15 – The fully connected neural network architecture used for the classification. The input consists of an array for the S2-area per PMT, that constitutes a one dimensional representation of the top hitpattern. The output is a pair of probabilities characterizing the classification in the two possible classes.

The input layer consists of 127 nodes, corresponding to the equivalent number of the top PMTs. That is, the hitpattern representation is made through a one-dimensional column. Then two hidden layers with 50 and 10 nodes follow and finally an output layer with two nodes corresponding to the two classes. The activation functions for the input and the two hidden layers are sigmoid functions, while the output layer is a softmax as we have here a binary classification problem. The initialization of the weights is done via the so-called Glorot normal initializer [270]. The training set consists of 250k top array hitpatterns for each of the 5 cases (i.e., 1.5M hitpatterns) and the testing set consists of 20k hitpatterns, for each of the 5 cases (i.e., 120k hit patterns).

After 10 training epochs, the network evaluation on the testing set gave a pair of probabilities (p_0, p_1) (such that $p_0 + p_1 = 1$) for each hitpattern of the testing set, concerning the probability that belongs to class 0 or 1. If we set a threshold, $t_0 \in [0,1]$, concerning the classification in class 0 such that, for $p_0 \leq t_0$, the hitpattern is classified as belonging to class 0, otherwise to class 1, and we variate t_0 from 0 to 1, we obtain a

receiver operating characteristic (ROC) curve. This curve characterizes the accuracy of the neural network as well as the variation in the efficiency of the cut.

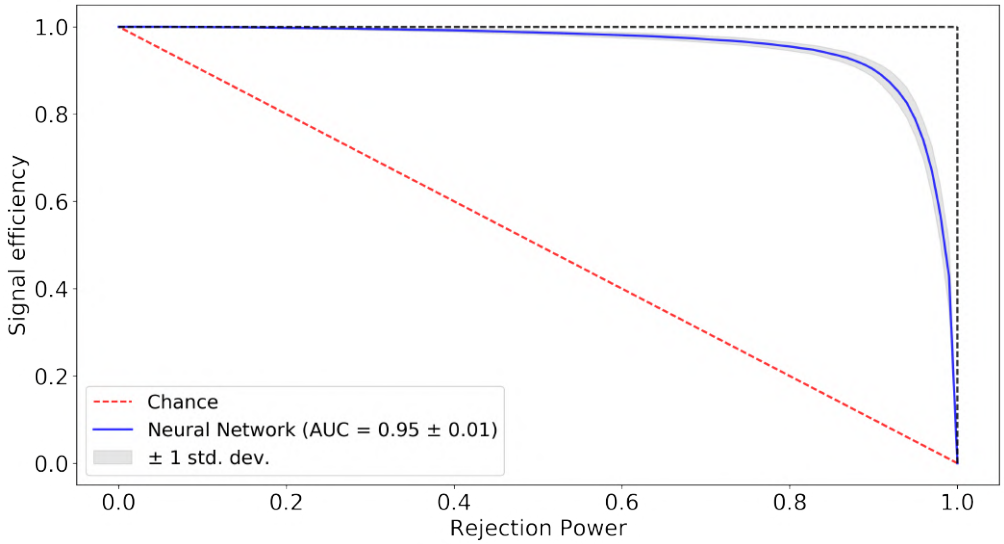


Figure 4.16 – ROC curve corresponding to the neural network cut. The optimal working point on such a curve would be the one offering simultaneously the largest possible rejection power for the largest possible signal efficiency.

In Fig. (4.16), the signal efficiency is equal to the True Positive Rate (TPR) and the rejection power is equal to the True Negative Rate (TNR). If TP, TN, FP, and FN are respectively the cases True Positive (events of class 0 classified as class 0), True Negative (events of class 1 classified as class 1), False Positive (events of class 1 classified as class 0) and False Negative (events of class 0 classified as class 1), then the TPR and the TNR are calculated as follows:

$$TPR = \frac{TP}{TP + FN}$$

$$TNR = \frac{TN}{TN + FP}$$

In Fig. (4.17) are illustrated some examples of true positive cases and true-negative cases of the neural network:

Classification of real data

We can see this neural network classification on real data obtained from 35 datasets of SR1, in various parametric spaces of interest.

In Fig. (4.18) is shown the probability p_0, as a function of the time from the main S2 of the event. We notice that, for a short time separation from the main S2, we have a

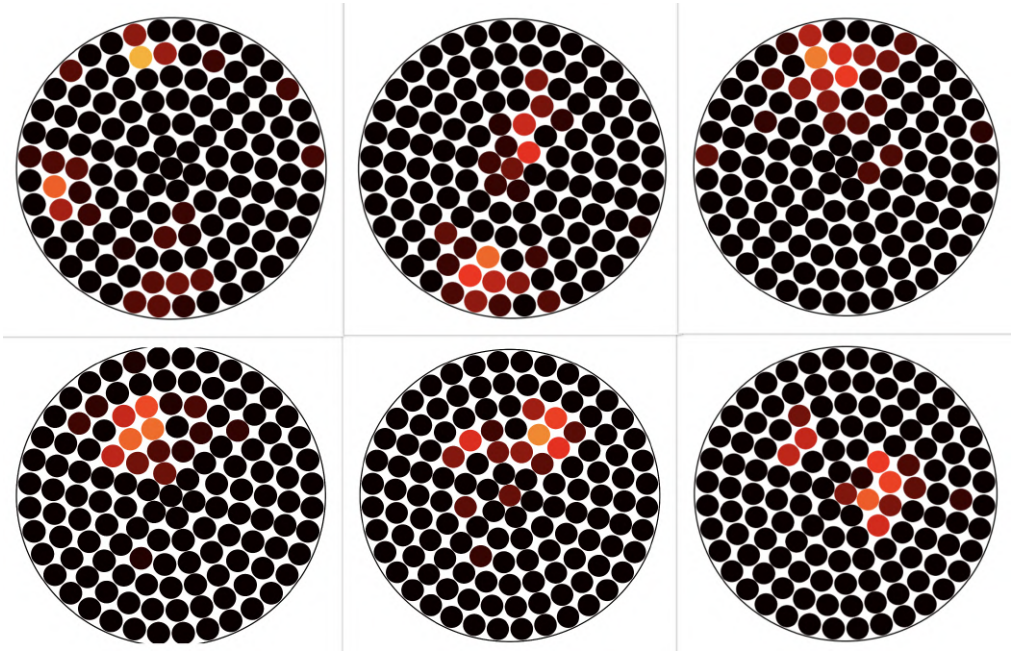


Figure 4.17 – Examples of classified top hitppaterns produced from simulation. Top row: True negative cases. From left to right the degree of difficulty is increasing (probability p_1 goes from above 0.9 to just above 0.5 for the last hitpattern). Bottom row: True positive cases with the same as above increase of classification difficulty from left to right (probability p_0 gets smaller).

large percentage classified as SEs originating from different x-y positions. The majority is classified as originating from the same position, a trend which remains for larger Δt. This is consistent with the observation of the correlation between the positions of the SE trains and the main S2 peak, directly preceding them, as it is expected that pileups from these SEs will produce hitpatterns consistent with a common x-y origin. Nevertheless, a sparse population of events, classified in class 1, persists even for large Δt. These events could be removed by this cut. Also, in Fig. (4.18) left, we see a zoom in the temporal region of one drift time after the main S2 of the event. We clearly see here, nearly a division of events between the two classes, consistent with the observation that single electrons originating from the photoionisation, after a main S2, are homogeneously distributed in the TPC, thus its expected to produce also many spatially non-correlated pileups.

A neural network is also used in order to reconstruct the position x-y of peaks, via the top PMT hit pattern. This neural network is trained on hitpatterns produced by optical simulations. In Fig. (4.19), we can see the probability p_0, as a function of a measure used in the main S1-S2 analysis that we call "goodness_of_fit_tpf". This parameter is actually a Poisson likelihood test statistic, distributed as a chi-square under the null hypothesis [271], that measures the goodness of the fit of the top PMT hitpattern, used

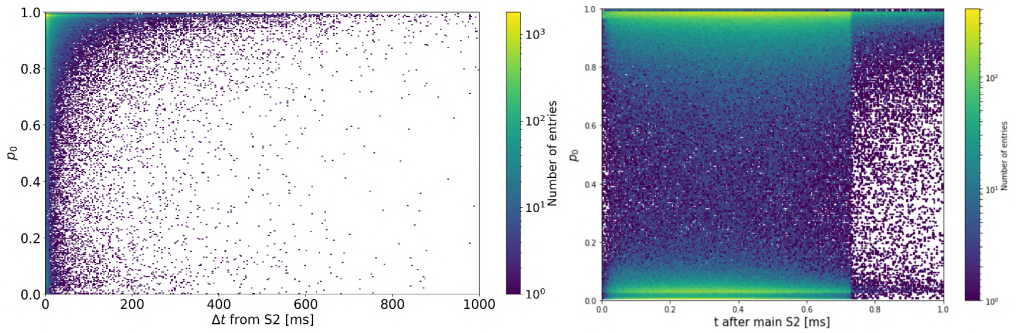


Figure 4.18 – Evolution of the classification probability p_0 as a function of the time, after a main S2. Left: We can see the classification up to 1 s from a main S2, using a dedicated calibration in triggerless mode, were we are not limited by a specific time window around the main S2. We can see a band of events classified as originating from a common origin, persisting up to a second. This is consistent with the SE trains that are in position correlation with the main S2, thus expected to produce position correlated pileups. Right: A zoom in the time window of a typical event from a search in trigger mode. For one drift time after the main S2, we know that the potoionization single electrons are homogeneously distributed inside the TPC . Indeed, this is reflected by the partition of the classification output between extreme values of p_0. After the maximum drift time, for the time left up to the end of the window, the SE trains start to dominate.

to determine the position of the events. A low value is consistent with a more accurate reconstruction, while a higher value signifies the presence of multiple clusters in the top hitpattern (for example, hitpatterns produced from multiple scatters of a neutron) thus the reconstructed position will be a kind of geometric average between these. We observe a general concordance between the two measures i.e., a smaller p_0 for a larger "goodness_of_fit_tpf". But we also notice that, for a narrow band of values of the "goodness_of_fit_tpf", the p_0 can have very different values. This is probably due to the fact that the top pattern fit algorithm is trained on simulated electron clouds from a common origin, in order to be used in position reconstruction. Also these optical simulations doesn't take into account a full simulation of the PMT afterpulses, as is done within the waveform simulator. That is, the present neural network approach is more efficient, concerning the classification of the hitpatterns in the two classes of interest for this study.

Finally, we can see the effect this cut has on the number of events counted, as a function of the S2-area. We can use 3 bins, with widths equal to the SE gain (28 PE/e^-), corresponding to 1, 2 and 3 extracted electrons. We see in Fig. (4.20, left) the number of events with no cut and with the cut applied, for two different working points on the ROC curve, showed in the legend of the plot. In (4.20, right), is illustrated the 90% C.L. limit,

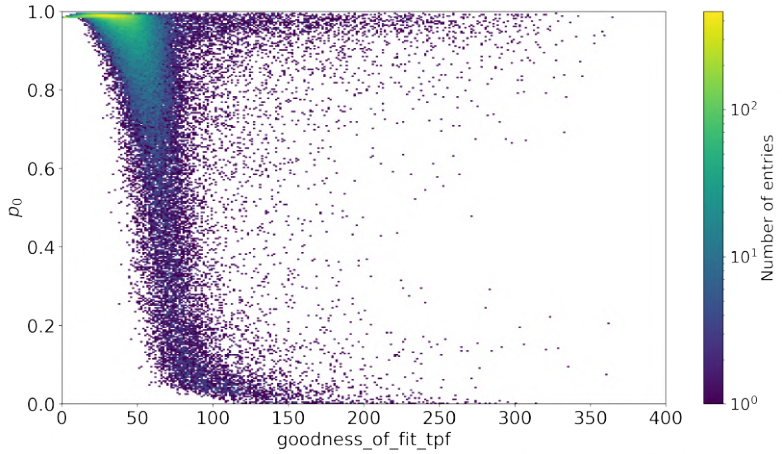


Figure 4.19 – The classification probability p_0 of the neural network as a function of the "goodness_of_fit_tpf" parameter described in the text.

corrected for the cut efficiency. This limit is calculated from $\mu_{up} = \frac{1}{2}F^{-1}_{\chi^2}(1-a_{up};2(n+1))$, with n the number of events and a C.L. of $1-a_{up} = 0.9$.

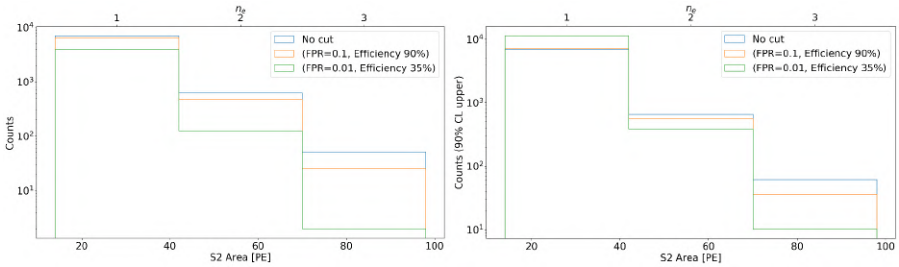


Figure 4.20 – Effect of the cut on the S2 spectrum in bins corresponding to one, two and three observed electrons. Left: Number of events, as a function of S2-area for two values of the cut, corresponding to certain values of false-positive rates (FPR) and efficiency. Right: 90% C.L. upper, for the histogram of left, corrected for the cut efficiency. In a limit setting procedure, this will eventually improve the exclusion, in the region of DM masses were there is a high enough contribution of 2 and three electron signals.

4.6 Exclusion limit

We are going to proceed now to the application of the above analysis, on a small portion (about 8%) of the total exposure. This portion is used, in the ongoing analysis, in order to tune the cuts and fix the analysis before its application in the totality of the data. As is evident, from the discussion in this chapter, the small charge signal that we expect is situated exactly in the S2 region dominated by single electrons. Even worse, as

we saw, the mechanisms behind this detector-related background are various and not yet fully understood. The search for such a signal is therefore ill-starred, in the sense that it takes place without the possibility of a background subtraction, that is, without the ability to claim an excess of events over the background. Nevertheless, what we can do, is to set an exclusion limit on that precise value of the DM-e^- cross section, for which the number of expected signal events will not exceed the number of observed events. Because a larger cross-section than this limit, by necessity would result in a greater number of observed events. The purpose of this work is an **S2-only analysis** (as it is typically called, in the field of dual-phase TPCs) on the smallest possible energy region that produces observable charge signals. This actually constitutes a pushing on the extreme limits of a classical S2-only analysis, that uses the main S2 peaks of triggered events but is simply non requiring the presence of a main S1. As the threshold of an S1 is no longer present this enables the improvement of the sensitivity for low dark matter masses but its lower end is the boundary of the 150 PE, after which the background of single electrons is dominant. Such an analysis has already been published for XENON1T and resulted in the improvement of the best, until then, DM-e^- exclusion limit of XENON10 for masses above 30 MeV/c^2. Our analysis, therefore, intends to expand this improvement for lower DM masses, a region were the XENON10 limit is still world leading, building on the present analysis developed from our current understanding of the SE background.

We are going to apply first all the cuts presented above, except for the neural network cut. This is due to the fact that, until now, the simulated training dataset of the neural network contains the different position combinations of only up to three electrons. Its update, for the inclusion of the simulation for all different possibilities for 4 and 5 detected electrons, is ongoing. The S2 peaks chosen are in the region of $[10-150]$ PE, where the peak finder efficiency (our effective "trigger" efficiency in this analysis) is practically 100%, as we showed in section 2.4.2. We are going to compare the expected rates of $1, 2, \cdots 5$ observed electrons, from 4.28, with the actually observed rates that will be un-folded from the low S2-spectrum of the data by a fit with a specific model. For the moment, and for reasons of comparison with the result of XENON10 [98, 265], we are going to use only the first model of Chapter 2, that is, a multi-Gaussian, without using the Fermi-Dirac threshold function and setting the value of the secondary scintillation gain at 28 PE/e^-.

The fit of the single electron spectrum is done using the formula:

$$f(S2|\sigma, \mathbf{h}) = \sum_{n=1}^{5} h_n \exp\Big(-\frac{1}{2}\big(\frac{S2 - n\mu}{\sqrt{n}\sigma}\big)^2\Big) \tag{4.32}$$

where $\mu = 28$ PE/e^-, is the secondary scintillation yield and $\mathbf{h} = (h_1, h_2, h_3, h_4, h_5)$ are the heights of the five components that will determine the event rates of the, up to five, detected electron signals:

$$r_{\text{detected}} = \frac{h_n \sqrt{2\pi n}\sigma}{\varepsilon \Delta x} \tag{4.33}$$

where ε is the cut efficiency (that here is 1, as we are using only the livetime reducing hard cuts) and Δx is the histogram bin width. This, here, correspond to 0.1 observed electron (i.e., 2.8 PE), as is done in [98].

The fit of the model (4.32) to the data is done as in Chapter 2: We form the likelihood function for the model, concerning finding of $(d_1, d_2, \cdots, d_{N_{\rm bins}})$ counts in each non-empty bin of the histogram, given the data and the parameters of the model $(\sigma, \mathbf{h}) = (\sigma, h_1, h_2, h_3, h_4, h_5)$:

$$\mathcal{L}(S2|\sigma,\mathbf{h}) = \prod_{i=1}^{N_{\rm bins}} \exp\Big(-\frac{1}{2}\big(\frac{d_i - f(S2_i|\sigma,\mathbf{h})}{\hat{\sigma}_i}\big)^2\Big) = \exp\Big(-\frac{1}{2}\sum_{i=1}^{N_{bins}}\big(\frac{d_i - f(S2_i|\sigma,\mathbf{h})}{\hat{\sigma}_i}\big)^2\Big)$$

where d_i is the content of each bin and $\hat{\sigma}_i$ the statistical error associated. Then, given the data and this likelihood, using a MCMC algorithm we are drawing samples for the posterior distribution of the parameters of interest:

$$p(\sigma,\mathbf{h}) = \mathcal{L}(S2|\sigma,\mathbf{h})\text{Prior}(\sigma,\mathbf{h})$$

Using a flat (uninformative) prior over the parameters $(\sigma, \mathbf{h})$.

In this work we will not present the final results of this analysis on the actual data of SR1, as it is expected to be published in a science journal after the publication of this thesis. In this chapter we limit ourselves to outlining the structural features of the analysis in terms of the limit setting process for the case of sub-GeV DM.

Note added a posteriori: The reader can see the final results of the analysis in the publication [272], where the first part of the paper presents the empirical observations regarding the phenomenology of the SE trains, which ultimately lead to the development of the analysis cuts presented earlier. The second part of the paper presents the Light DM search as outlined in this thesis that ends up in the limit setting on the DM-electron scattering for the two cases of heavy and ultralight mediator as well as for other models of light DM.

4.6.1 Effect of the neural network cut on the limit

Due to our very low observed background, in the region of two and three observed electrons, we can obtain a relatively competitive limit. With the help of the neural network cut we could possibly remove a part of the two- and three-electron pileups. Just to give an idea of how this can be achieved, we are going to use the waveform simulator to simulate a small-S2 spectrum. For this purpose, we have generated 10k events, each containing 10 electrons, starting from random positions (x, y) inside the TPC. Thus, the produced spectrum, contains all single electron pileups that may be derived from the same or different points (x,y). Of course, this random simulation cannot exactly reproduce the percentage of single electrons that are spatially correlated nor the correct rate the multiple pileups, since we don't simulate the, unknown, underlying true physical processes, but only the detector response to our homogeneous distribution of single electrons. We are just using it here as a case study, to find out the effect of the cut designed. The spectrum produced by this simulation is illustrated in Fig. (4.22).

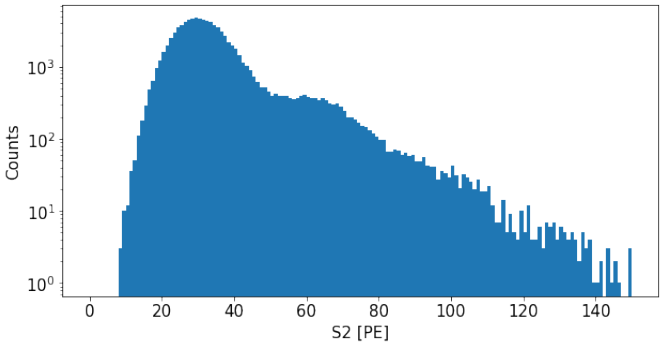


Figure 4.21 – Small-S2 spectrum, obtained from 10k simulations with the waveform simulator of events with 10 electrons each observed in random x-y positions.

In order to simulate a real situation, we randomly select, from the events of Fig. (4.22), 1000 events and say that this corresponds to a certain exposure (for example we set $\sim 30\text{kg}\cdot\text{day}$ exposure corresponding to the order of magnitude of the full exposure expected after all cuts applied). Then, we unfold the few-electron rate, from the fit of the model (4.32) but only using three Gaussians , in order to check the effect of the neural network trained on simulated data with only up to 3 detected electrons, Fig. (4.22).

The two corresponding limits, for the cross-section of dark matter-electron scattering, are shown in Fig. (4.23). It appears that, in the region of the limit, corresponding to DM masses resulting in ionization electrons with kinetic energy high enough to generate two or three observable electrons ($m_\chi \gtrapprox 10$ MeV/c^2), the application of the cut may lead to an improved result.

Of course in real data, the extent of such an improvement will depend on the extent to which, after applying all the exposure reduction cuts, there will still remain several

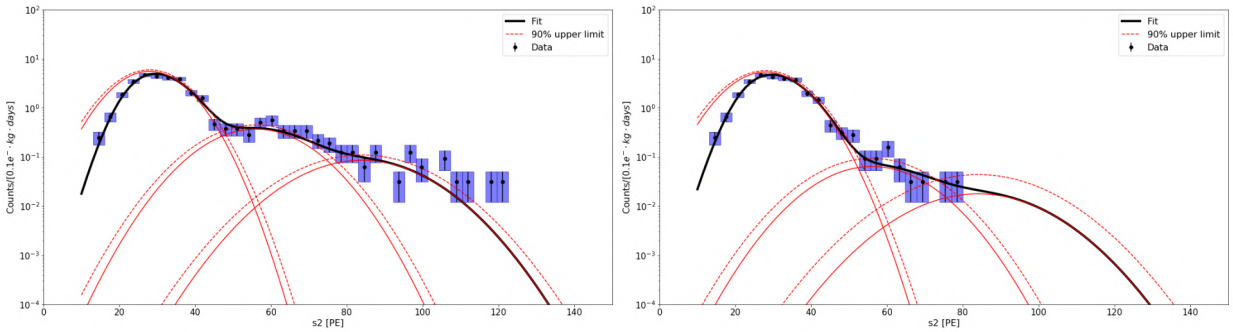


Figure 4.22 – Effect of the neural network cut on the single electron spectrum. The top hitpattern of each small S2 peak, of the simulated spectrum of Fig. (4.22), is used as input of the neural network and, based on its classification, we apply the cut for the optimal point on the ROC curve corresponding to a 90% signal efficiency with 90.2% background rejection power. Right: Left: The same spectrum after the application of the cut removing events classified as originating from different x-y positions. In the two figures, is also shown the fit of the single electron model with only three Gaussians.

two and three single electron pileups that will not be spatially correlated. We see, in Fig. (4.18), that apart the train of SEs which are position correlated, a sparse population of events is classified to originate from different x-y positions, even for large time separations from the previous S2. Thus, to the extent that the other cuts cannot completely eliminate events from every mechanism of production of single-electron backgrounds , inside the TPC, and considering that some of them, such as single-electrons coming from electrons trapped in the phase boundary, are likely to be homogeneously distributed in the x-y plane, such a cut can help to further improve the result of the analysis.

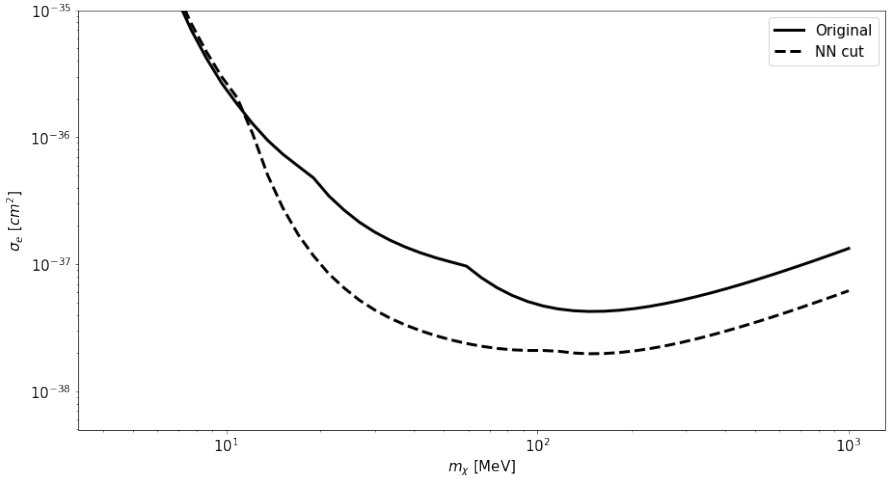


Figure 4.23 – Limit on $\overline{\sigma}_e$ for $F_{DM} = 1$, inferred from the simulated data of Fig. (4.22). We see an improvement on the limit on $\overline{\sigma}_e$, after application of the cut. A small increase in the limit can also be seen, for DM masses below 10 MeV/c^2. This is simply due to the fact that, the cut efficiency of 90% is applied for the whole spectrum of Fig. (4.22). On the other hand, the cut does not remove events on the one observed electron region of S2 (as they are mostly classified as originating from the same position) resulting in a small increase of the limit, in the region of DM masses where the contribution of the one-electron region is dominant.

Conclusion

In this thesis we have studied a number of topics, ultimately related to the leptonically interacting dark matter, through various analyzes of data, acquired during the scientific runs of the XENON1T experiment. Understanding the response of the detector in small charge signals, helped us to proceed in a characterization of the detector with the help of single electrons that offers, in return, important information about the phenomenology of this background. Using this understanding and taking advantage of the detector's sensitivity to few-electron signals, we proceeded to the search for a leptonically interacting dark matter, with mass in the MeV scale. We have also examined an additional possibility of direct detection of DM interacting with electrons, studying the extremely low electronic background of the detector in a search for a possible annual modulation that would be compatible with a dark matter interpretation. We presented a study of the components of this background and a study about the detector stability, that would prevent the attribution of such an annual modulation to spurious effects arising from detector parameters.

Specifically, in Chapter 2 we proceeded to a detailed study of the low-S2 spectrum of the detector, corresponding to the extraction of few electrons and their proportional scintillation in the gaseous phase. We used the abundant single electrons derived from photoionization of impurities in the liquid phase, in order to construct a model that could describe the spectrum of single electrons and their temporal pileups. This model differs from the one used in the XENON100 detector, taking into account mainly the new, more efficient, peak-finding algorithm, used in XENON1T. With this model of the low-S2 spectrum, we have measured the amount of light produced by an electron extracted to the gas gap. It also allowed us to perform a characterization of many detector parameters, such as those that determine the secondary scintillation yield as a function of the pressure of the gaseous Xe, as well as, its dependence on the anode voltage. It was possible to study the sagging of the anode, due to the gravity and the electrostatic force exerted on it. Also, using the photoionization-generated single electrons, from dedicated ^{83m}Kr calibration datasets, we were able to carry out a study about the variation of the extraction efficiency, as a function of the electric field, in the detector's gas gap. This, consisted in a relative evaluation of the electron extraction probability, emerging from a comparison between the secondary scintillation gain with the S2 gain of the monoenergetic calibration source.

The data obtained were compared to a theoretical model about the extraction process referring to the microphysical processes taking place. One of them concerns the trapping of electrons in the potential formed just below the phase boundary and their ability to be extracted at a later time. This is a delayed emission mechanism whose understanding is important to the search of light matter, of which constitutes an irreducible background.

A method to search for dark matter particles, with mass of the order of GeV, interacting with electrons, is given by the annual modulation of the ER background they are expected to induce. The temporal characteristics of this modulation were described in the third Chapter. We proceeded to the detailed presentation of the data analysis cuts used to select true electron recoil events, as well as, the evaluation of their acceptance as a function of time. This was followed by a detailed modeling of each one of the ER background components, including ^{214}Pb, ^{85m}Kr, ^{60}Co and solar neutrino backgrounds. The cut acceptances and the background model, are the necessary elements in order to proceed to a final likelihood inference about a possible modulation. However in order that the confirmation of exclusion of such a modulation could be attributed in the true behavior of the background, it is necessary to establish, in a rigorous statistical manner, the temporal stability of all the detector parameters that could possibly influence the signal generation. Thus, we presented a systematic statistical investigation on the significance of a possible correlation of the electronic recoil event rates with the detector parameters. The corresponding analysis resulted in an exclusion of any significant correlation of this kind. A study followed about the sensitivity of the experiment to background fluctuations, using toy MC datasets that were generated, based on the total livetime and the known background model. An artificial modulation signal, with a certain amplitude and phase, was injected in order to determine the size of the modulation fraction that would allow a statistically significant inference about the signal. We found that, even if the real total time of data taking extends to a little more than one year, the stability of the detector and the temporal stability of the electronic recoil background are such that allows a detection or an exclusion of a modulation, consistent with the one expected from the standard halo model, for certain values of the modulation fraction.

The sensitivity of the detector to the S2-light produced by single electrons finds all its purposefulness in the fourth chapter, where a search is attempted for DM with masses in the MeV scale, scattering off the Xe atomic electrons, a possibility emerging within the framework of hidden sector theories. We presented in detail the theoretical calculation of the event rate, taking into account a many body quantum mechanics treatment of the initial and the final states of the electron. We have folded the theoretical signal with the detector response, based on the theoretical and experimental understanding of the electronic recoil charge yield, and we calculated the expected few-electron $(1, 2, \ldots 5)$ rates. These are compared to the observed rates of single electrons, unfolded from the low-S2 spectrum. In this kind of search, the principal background is due to the plethora of mechanisms producing single electrons inside the detector. As this background scales

up with the size of the detector, the strongest exclusion limits with a noble liquid target to date, for DM masses below 20 MeV, are still given by the XENON10 detector. It is, therefore, necessary to use the knowledge gained on the phenomenology of single electrons in order to find a region of the total exposure that presents, as much as possible, a reduced rate of single electrons. Using the pre-trigger time window of waveforms of acquired events, we developed cuts that helped us remove a significant portion of the background due to single electron trains. Also we developed a neural network that we have trained with simulated top hitpatterns, in order to fight against the background of single electron pileups, something that will further contribute to the improvement of the sensitivity for the search of light dark matter. Finally we found that, despite the very large single electron background, resulting from the size of the detector, it is still possible to set a competitive exclusion limit for DM masses below 30 MeV, exactly because of its better understanding.

It is true that the era of silicon Skipper charge coupled devices (CCD) detectors like DAMIC or SENSEI has risen. They are expected to dominate the search for a DM-electron interaction, due to the very low energy of the band gap together with an extremely low leakage current and the possibility of background identification and rejection, based on an excellent spatial resolution. However, this dominance is restricted in the mass region below 10 MeV, as they rapidly loss sensitivity for higher masses. The present work shows that in the region above 10 MeV, the large dual-phase liquid xenon detectors principally developed for the WIMP dark matter search, still have the ability to probe the parametric space of the theory and haven't reach the limit of their scientific potential. Further, it is necessary to make dedicated efforts with small prototype dual-phase liquid xenon TPCs, in order to carefully investigate the single electron background and to find ways to reduce it in a hardware manner, or developing a complete background model. Such an effort is expected to begin at LPNHE, the laboratory in which this thesis was realized, aiming in the coming years to further strengthen the scientific potential of large xenon detectors in the search of dark matter particles in the sub-GeV scale.

Bibliography

[1] Heinz Andernach and Fritz Zwicky. English and Spanish Translation of Zwicky's (1933) The Redshift of Extragalactic Nebulae. *arXiv e-prints*, page arXiv:1711.01693, November 2017.

[2] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophysical Journal*, 86:217, October 1937.

[3] Vera C. Rubin and Jr. Ford, W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. , 159:379, February 1970.

[4] V. C. Rubin, Jr. Ford, W. K., 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=122kpc). , 238:471–487, June 1980.

[5] Albert Bosma. *The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types*. PhD thesis, 1978. $\mathrm{date}_s ubmitted : 2008 Rights : University of Groningen.$

[6] B. D. Fields and S. Sarkar. Big bang nucleosynthesis. *Physics Letters B*, 667(1):228 – 231, 2008. Review of Particle Physics.

[7] S.W. Allen, A.C. Fabian, R.W. Schmidt, and H. Ebeling. Cosmological constraints from the local x-ray luminosity function of the most x-ray luminous galaxy clusters. *Mon. Not. Roy. Astron. Soc.*, 342:287, 2003.

[8] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, sep 1998.

[9] D. Fixsen et al. Cosmic microwave background dipole spectrum measured by the cobe firas instrument. *The Astrophysical Journal*, 420:445–449, 01 1994.

[10] C. L. Bennett et al. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. , 208(2):20, October 2013.

[11] Planck Collaboration et al. Planck 2018 results. vi. cosmological parameters. 07 2018.

[12] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. , 142:419–421, July 1965.

[13] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. , 142:414–419, July 1965.

[14] D. J. Fixsen. The temperature of the cosmic microwave background. *The Astrophysical Journal*, 707(2):916–920, nov 2009.

[15] S. Dodelson. *Modern Cosmology*. Beijing World Publishing Corporation, 2008.

[16] A. Friedman. On the curvature of space. *General Relativity and Gravitation*, 31(12):1991–2000, 1999.

[17] Joop Schaye et al. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *Monthly Notices of the Royal Astronomical Society*, 446(1):521–554, 11 2014.

[18] Mark Vogelsberger et al. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *Monthly Notices of the Royal Astronomical Society*, 444(2):1518–1547, 08 2014.

[19] Andrew Pontzen and Fabio Governato. Cold dark matter heats up. , 506(7487):171–178, February 2014.

[20] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. , 292:371–394, May 1985.

[21] Edward W. Kolb and Michael S. Turner. *The Early Universe*, volume 69. 1990.

[22] Albert Einstein. The Foundation of the General Theory of Relativity. *Annalen Phys.*, 49(7):769–822, 1916.

[23] Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, jul 2010.

[24] Adam S. Bolton, Scott Burles, Léon V. E. Koopmans, Tommaso Treu, and Leonidas A. Moustakas. The Sloan Lens ACS Survey. I. A Large Spectroscopically Selected Sample of Massive Early-Type Lens Galaxies. , 638(2):703–724, February 2006.

[25] M. Bartelmann and P. Schneider. Weak gravitational lensing. , 340(4-5):291–472, January 2001.

[26] Alan Heavens. Weak lensing: Dark matter, dark energy and dark gravity. *Nuclear Physics B - Proceedings Supplements*, 194:76 – 81, 2009. New Horizons for Modern Cosmology:.

[27] R. Mandelbaum, U. Seljak, G. Kauffmann, C. M. Hirata, and J. Brinkmann. Galaxy halo masses and satellite fractions from galaxy–galaxy lensing in the sloan digital sky survey: stellar mass, luminosity, morphology and environment dependencies. *Monthly Notices of the Royal Astronomical Society*, 368(2):715–731, May 2006.

[28] Raphael Gavazzi, Tommaso Treu, Jason D. Rhodes, Leon V. E. Koopmans, Adam S. Bolton, Scott Burles, Richard J. Massey, and Leonidas A. Moustakas. The sloan lens ACS survey. IV. the mass density profile of early-type galaxies out to 100 effective radii. *The Astrophysical Journal*, 667(1):176–190, sep 2007.

[29] David M. Goldberg and David J. Bacon. Galaxy-Galaxy Flexion: Weak Lensing to Second Order. , 619(2):741–748, February 2005.

[30] Gary M. Bernstein and Reiko Nakajima. Multipole formulae for gravitational lensing shear and flexion. *The Astrophysical Journal*, 693(2):1508–1513, mar 2009.

[31] W. H. Tucker, H. Tananbaum, and R. A. Remillard. A Search for "Failed Clusters" of Galaxies. , 444:532, May 1995.

[32] M. Markevitch, A. H. Gonzalez, L. David, A. Vikhlinin, S. Murray, W. Forman, C. Jones, and W. Tucker. A textbook example of a bow shock in the merging galaxy cluster 1e 0657-56. *The Astrophysical Journal*, 567(1):L27–L31, feb 2002.

[33] Maxim Markevitch, Craig L. Sarazin, and Alexey Vikhlinin. Physics of the merging clusters cygnus a, a3667, and a2065. *The Astrophysical Journal*, 521(2):526–530, aug 1999.

[34] Scott Randall, Maxim Markevitch, Douglas Clowe, Anthony Gonzalez, and Marusa Bradac. Constraints on the self-interaction cross-section of dark matter from numerical simulations of the merging galaxy cluster 1e 0657-5. *The Astrophysical Journal*, 679, 04 2007.

[35] Douglas Clowe et al. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal*, 648, 08 2006.

[36] L.D. Landau and E.M. Lifshitz. *Mechanics: Volume 1.* Number vol. 1. Elsevier Science, 1982.

[37] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, January 1933.

[38] K. C. Freeman. On the Disks of Spiral and S0 Galaxies. , 160:811, June 1970.

[39] K. G. Begeman. HI rotation curves of spiral galaxies. I. NGC 3198. , 223:47–60, October 1989.

[40] David Merritt, Katherine Garrett, and Gintaras Duda. Dark matter: A primer. *Advances in Astronomy*, 2011:968283, 2011.

[41] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Structure of Cold Dark Matter Halos. , 462:563, May 1996.

[42] Yoshiaki Sofue and Vera Rubin. Rotation curves of spiral galaxies. *Annual Review of Astronomy and Astrophysics*, 39(1):137–174, 2001.

[43] Fabio Iocco, Miguel Pato, and Gianfranco Bertone. Evidence for dark matter in the inner milky way. *Nature Physics*, 11(3):245–248, 2015.

[44] N.M. McClure-Griffiths and John M. Dickey. Milky Way Kinematics: Measurements at the Subcentral Point of the Fourth Quadrant. *Astrophys. J.*, 671:427–438, 2007.

[45] Abraham Luna, L. Bronfman, L. Carrasco, and J. May. Molecular gas, kinematics, and ob star formation in the spiral arms of the southern milky way. *Astrophys. J.*, 641:938–948, 2006.

[46] Hou, L. G., Han, J. L., and Shi, W. B. The spiral structure of our milky way galaxy*. *A&A*, 499(2):473–482, 2009.

[47] M. J. Reid, K. M. Menten, A. Brunthaler, X. W. Zheng, T. M. Dame, Y. Xu, Y. Wu, B. Zhang, A. Sanna, M. Sato, K. Hachisuka, Y. K. Choi, K. Immer, L. Moscadelli, K. L. J. Rygl, and A. Bartkiewicz. Trigonometric Parallaxes of High Mass Star Forming Regions: The Structure and Kinematics of the Milky Way. , 783(2):130, March 2014.

[48] Durand, S., Acker, A., and Zijlstra, A. The kinematics of 867 galactic planetary nebulae*. *Astron. Astrophys. Suppl. Ser.*, 132(1):13–20, 1998.

[49] F. Pont, D. Queloz, P. Bratschi, and M. Mayor. Rotation of the outer disc from classical cepheids. , 318:416–428, February 1997.

[50] P. Battinelli, S. Demers, Corinne Rossi, and K. Gigoyan. Extension of the c star rotation curve of the milky way to 24 kpc. *Astrophysics*, 56, 12 2012.

[51] Benjamin Audren, Julien Lesgourgues, Gianpiero Mangano, Pasquale Serpico, and Thomas Tram. Strongest model-independent bound on the lifetime of dark matter. *Journal of Cosmology and Astroparticle Physics*, 2014, 07 2014.

[52] S. D. M. White, C. S. Frenk, and M. Davis. Clustering in a neutrino-dominated universe. , 274:L1–L5, November 1983.

[53] Roland de Putter, Olga Mena, Elena Giusarma, Shirley Ho, Antonio Cuesta, Hee-Jong Seo, Ashley J. Ross, Martin White, Dmitry Bizyaev, Howard Brewington, David Kirkby, Elena Malanushenko, Viktor Malanushenko, Daniel Oravetz, Kaike Pan, Will J. Percival,

Nicholas P. Ross, Donald P. Schneider, Alaina Shelden, Audrey Simmons, and Stephanie Snedden. New neutrino mass bounds from sloan digital sky survey iii data release 8 photometric luminous galaxies. *The Astrophysical Journal*, 761(1):12, nov 2012.

[54] Matteo Viel, George Becker, James Bolton, and Martin Haehnelt. Warm dark matter as a solution to the small scale crisis: new constraints from high redshift lyman-alpha forest data. *Physical Review D*, 88, 06 2013.

[55] Marco Drewes. The phenomenology of right handed neutrinos. *International Journal of Modern Physics E*, 22, 03 2013.

[56] George R. Blumenthal, S.M. Faber, Joel R. Primack, and Martin J. Rees. Formation of Galaxies and Large Scale Structure with Cold Dark Matter. *Nature*, 311:517–525, 1984.

[57] B. Paczynski. Gravitational Microlensing by the Galactic Halo. , 304:1, May 1986.

[58] Tisserand, P., Le Guillou, L., Afonso, C., Albert, J. N., Andersen, J., Ansari, R., Aubourg, É., Bareyre, P., Beaulieu, J. P., Charlot, X., Coutures, C., Ferlet, R., Fouqué, P., Glicenstein, J. F., Goldman, B., Gould, A., Graff, D., Gros, M., Haissinski, J., Hamadache, C., de Kat, J., Lasserre, T., Lesquoy, É., Loup, C., Magneville, C., Marquette, J. B., Maurice, É., Maury, A., Milsztajn, A., Moniez, M., Palanque-Delabrouille, N., Perdereau, O., Rahal, Y. R., Rich, J., Spiro, M., Vidal-Madjar, A., Vigroux, L., and S. Zylberajch (The EROS-2 collaboration). Limits on the macho content of the galactic halo from the eros-2 survey of the magellanic clouds ***. *A&A*, 469(2):387–404, 2007.

[59] T. Lasserre. Not enough stellar mass machos in the galactic halo. *Astron. Astrophys.*, 355:L39–L42, 2000.

[60] R.D. Peccei and Helen Quinn. Cp conservation in the presence of pseudoparticles. *Physical Review Letters - PHYS REV LETT*, 38:1440–1443, 06 1977.

[61] R. D. Peccei and Helen R. Quinn. Constraints imposed by CP conservation in the presence of pseudoparticles. *Phys. Rev. D*, 16:1791–1797, Sep 1977.

[62] E. Aprile et al. Observation of excess electronic recoil events in xenon1t. 06 2020.

[63] Keith R. Dienes, Emilian Dudas, and Tony Gherghetta. Extra spacetime dimensions and unification. *Physics Letters B*, 436(1):55 – 65, 1998.

[64] Ignatios Antoniadis. A Possible new dimension at a few TeV. *Phys. Lett. B*, 246:377–384, 1990.

[65] Nima Arkani-Hamed, Andrew Cohen, and Howard Georgi. Electroweak symmetry breaking from dimensional deconstruction. *Physics Letters B*, 513:232–240, 07 2001.

[66] Stephen P. Martin. *A Supersymmetry Primer*, pages 1–98.

[67] Edward Witten. Constraints on supersymmetry breaking. *Nuclear Physics B*, 202(2):253 – 316, 1982.

[68] Gerard 't Hooft, C. Itzykson, A. Jaffe, H. Lehmann, P.K. Mitter, I.M. Singer, and R. Stora, editors. *Recent Developments in Gauge Theories. Proceedings, Nato Advanced Study Institute, Cargese, France, August 26 - September 8, 1979*, volume 59, 1980.

[69] Nathan Seiberg. Naturalness versus supersymmetric nonrenormalization theorems. *Phys. Lett. B*, 318:469–475, 1993.

[70] S. Dimopoulos, S. Raby, and Frank Wilczek. Supersymmetry and the Scale of Unification. *Phys. Rev. D*, 24:1681–1683, 1981.

[71] A. Popolo. Non-baryonic dark matter in cosmology. *International Journal of Modern Physics D*, 23, 05 2013.

[72] Jonathan L. Feng. Dark Matter Candidates from Particle Physics and Methods of Detection. *Ann. Rev. Astron. Astrophys.*, 48:495–545, 2010.

[73] Pierre Sikivie, I.I. Tkachev, and Yun Wang. The Secondary infall model of galactic halo formation and the spectrum of cold dark matter particles on earth. *Phys. Rev. D*, 56:1863–1878, 1997.

[74] F. Ruppin, J. Billard, E. Figueroa-Feliciano, and L. Strigari. Complementarity of dark matter detectors in light of the neutrino background. *Phys. Rev. D*, 90:083510, Oct 2014.

[75] A. H. Abdelhameed et al. First results from the cresst-iii low-mass dark matter program. *Phys. Rev. D*, 100:102002, Nov 2019.

[76] R. Ajaj et al. Search for dark matter with a 231-day exposure of liquid argon using deap-3600 at snolab. *Phys. Rev. D*, 100:022004, Jul 2019.

[77] P. Agnes et al. Low-mass dark matter search with the darkside-50 experiment. *Phys. Rev. Lett.*, 121:081307, Aug 2018.

[78] P. Agnes et al. Darkside-50 532-day dark matter search with low-radioactivity argon. *Phys. Rev. D*, 98:102006, Nov 2018.

[79] H. B. Li et al. Limits on spin-independent couplings of wimp dark matter with a p-type point-contact germanium detector. *Phys. Rev. Lett.*, 110:261301, Jun 2013.

[80] A. Aguilar-Arevalo et al. Results on low-mass weakly interacting massive particles from a 11 kg-day target exposure of DAMIC at SNOLAB. 7 2020.

[81] Xiangyi Cui et al. Dark matter results from 54-ton-day exposure of pandax-ii experiment. *Phys. Rev. Lett.*, 119:181302, Oct 2017.

[82] D. S. Akerib et al. Results from a search for dark matter in the complete lux exposure. *Phys. Rev. Lett.*, 118:021303, Jan 2017.

[83] K. Abe et al. A direct dark matter search in xmass-i. *Physics Letters B*, 789:45 – 53, 2019.

[84] C. Amole et al. Dark matter search results from the complete exposure of the pico-60 c_3f_8 bubble chamber. *Phys. Rev. D*, 100:022001, Jul 2019.

[85] E. Behnke et al. Final results of the picasso dark matter search experiment. *Astroparticle Physics*, 90:85 – 92, 2017.

[86] M Felizardo et al. Recent results from the SIMPLE dark matter search. *Journal of Physics: Conference Series*, 375(1):012011, jul 2012.

[87] E. Behnke et al. First dark matter search results from a 4-kg $cf_3\mathbf{I}$ bubble chamber operated in a deep underground site. *Phys. Rev. D*, 86:052001, Sep 2012.

[88] Govinda Adhikari et al. An experiment to search for dark-matter interactions using sodium iodide detectors. *Nature*, 564(7734):83–86, 2018.

[89] H. Jiang et al. Limits on Light Weakly Interacting Massive Particles from the First 102.8 kg × day Data of the CDEX-10 Experiment. *Phys. Rev. Lett.*, 120(24):241301, 2018.

[90] E. Armengaud et al. Final results of the edelweiss-ii wimp search using a 4-kg array of cryogenic germanium detectors with interleaved electrodes. *Physics Letters B*, 702(5):329 – 335, 2011.

[91] E. Armengaud et al. Constraints on low-mass WIMPs from the EDELWEISS-III dark matter search. *Journal of Cosmology and Astroparticle Physics*, 2016(05):019–019, may 2016.

[92] Q. Arnaud et al. First results from the news-g direct dark matter search experiment at the lsm. *Astroparticle Physics*, 97:54 – 62, 2018.

[93] E. Aprile et al. Light dark matter search with ionization signals in xenon1t. *Phys. Rev. Lett.*, 123:251801, Dec 2019.

[94] E. Aprile et al. Dark matter search results from a one ton-year exposure of xenon1t. *Phys. Rev. Lett.*, 121:111302, Sep 2018.

[95] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. 2008.

[96] J Angle et al. First results from the xenon10 dark matter experiment at the gran sasso national laboratory. *Physical review letters*, 100:021303, 02 2008.

[97] J. Angle et al. Search for light dark matter in xenon10 data. *Phys. Rev. Lett.*, 107:051301, Jul 2011.

[98] Rouven Essig, Aaron Manalaysay, Jeremy Mardon, Peter Sorensen, and Tomer Volansky. First Direct Detection Limits on sub-GeV Dark Matter from XENON10. *Phys. Rev. Lett.*, 109:021301, 2012.

[99] E. Aprile et al. Xenon100 dark matter results from a combination of 477 live days. *Physical Review D*, 94, 09 2016.

[100] E. Aprile et al. Search for event rate modulation in xenon100 electronic recoil data. *Physical Review Letters*, 115, 07 2015.

[101] E. Aprile et al. Search for electronic recoil event rate modulation with 4 years of xenon100 data. *Phys. Rev. Lett.*, 118:101101, Mar 2017.

[102] E. Aprile et al. Exclusion of leptophilic dark matter models using xenon100 electronic recoil data. *Science*, 349, 07 2015.

[103] E. Aprile et al. Search for magnetic inelastic dark matter with XENON100. *JCAP*, 10:039, 2017.

[104] E. Aprile et al. Search for bosonic super-wimp interactions with the xenon100 experiment. *Phys. Rev. D*, 96:122002, Dec 2017.

[105] E. Aprile et al. Effective field theory search for high-energy nuclear recoils using the xenon100 dark matter detector. *Phys. Rev. D*, 96:042004, Aug 2017.

[106] E. Aprile et al. Observation and applications of single-electron charge signals in the XENON100 experiment. *Journal of Physics G: Nuclear and Particle Physics*, 41(3):035201, feb 2014.

[107] John R. de Laeter, John Karl Böhlke, P. De Bièvre, H. Hidaka, H. S. Peiser, K. J. R. Rosman, and P. D. P. Taylor. Atomic weights of the elements. review 2000 (iupac technical report). *Pure and Applied Chemistry*, 75(6):683 – 800, 2003.

[108] E. Aprile and T. Doke. Liquid xenon detectors for particle physics and astrophysics. *Rev. Mod. Phys.*, 82:2053–2097, Jul 2010.

[109] W.M. Haynes. *CRC Handbook of Chemistry and Physics*. CRC Handbook of Chemistry and Physics. CRC Press, 2011.

[110] R.L. Platzman. Total ionization in gases by high-energy particles: An appraisal of our understanding. *The International Journal of Applied Radiation and Isotopes*, 10(2):116 – 127, 1961.

[111] G. Plante. *The XENON100 Dark Matter Experiment: Design, Construction, Calibration and 2010 Search Results with Improved Measurement of the Scintillation Response of Liquid Xenon to Low-Energy Nuclear Recoils*. PhD thesis, Columbia University, New York, 2012.

[112] Shinzou Kubota, Masahiko Hishida, Masayo Suzuki, and Jian zhi Ruan. Liquid and solid argon, krypton and xenon scintillators. *Nuclear Instruments and Methods in Physics Research*, 196(1):101 – 105, 1982.

[113] Joshua Jortner, Lothar Meyer, Stuart A. Rice, and E. G. Wilson. Localized excitations in condensed ne, ar, kr, and xe. *The Journal of Chemical Physics*, 42(12):4250–4253, 1965.

[114] Akira Hitachi, Tan Takahashi, Nobutaka Funayama, Kimiaki Masuda, Jun Kikuchi, and Tadayoshi Doke. Effect of ionization density on the time dependence of luminescence from liquid argon and xenon. *Phys. Rev. B*, 27:5279–5285, May 1983.

[115] W. H. Lippincott, K. J. Coakley, D. Gastler, A. Hime, E. Kearns, D. N. McKinsey, J. A. Nikkel, and L. C. Stonehill. Scintillation time dependence and pulse shape discrimination in liquid argon. *Phys. Rev. C*, 78:035801, Sep 2008.

[116] B. A. Dolgoshein, V. N. Lebedenko, and B. U. Rodionov. New Method of Registration of Ionizing-particle Tracks in Condensed Matter. *Soviet Journal of Experimental and Theoretical Physics Letters*, 11:351, June 1970.

[117] Alain Lansiart, Alain Seigneur, Jean-Luc Moretti, and Jean-Pierre Morucci. Development research on a highly luminous condensed xenon scintillator. *Nuclear Instruments and Methods*, 135(1):47 – 52, 1976.

[118] K. Ni, E. Aprile, K.L. Giboni, P. Majewski, and M. Yamashita. Gamma Ray Spectroscopy with Scintillation Light in Liquid Xenon. *JINST*, 1:P09004, 2006.

[119] T. Doke, H.J. Crawford, A. Hitachi, J. Kikuchi, P.J. Lindstrom, K. Masuda, E. Shibamura, and T. Takahashi. Let Dependence of Scintillation Yields in Liquid Argon. *Nucl. Instrum. Meth. A*, 269:291–296, 1988.

[120] J.B. Birks, D.W. Fry, L. Costrell, and K. Kandiah. *The Theory and Practice of Scintillation Counting: International Series of Monographs in Electronics and Instrumentation*. International series of monographs on electronics and instrumentation. Elsevier Science, 2013.

[121] L. Onsager. Initial recombination of ions. *Phys. Rev.*, 54:554–557, Oct 1938.

[122] E. Aprile et al. Xenon1t dark matter data analysis: Signal reconstruction, calibration, and event selection. *Physical Review D*, 100, 09 2019.

[123] P. Benetti et al. Detection of energy deposition down to the kev region using liquid xenon scintillation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 327(1):203 – 206, 1993.

[124] M. Yamashita, T. Doke, K. Kawasaki, J. Kikuchi, and S. Suzuki. Scintillation response of liquid xe surrounded by ptfe reflector for gamma rays. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 535(3):692 – 698, 2004.

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

[126] COMSOL. www.comsol.com. https://www.comsol.com/.

[127] Ferenc Glück, Guido Drexlin, Benjamin Leiber, Susanne Mertens, Alexander Osipowicz, Jan Reich, and Nancy Wandkowsky. Electromagnetic design of the large-volume air coil system of the KATRIN experiment. *New Journal of Physics*, 15(8):083025, aug 2013.

[128] E Aprile, Katsushi Arisaka, F Arneodo, A Askin, L Baudis, and A Behrens. The xenon100 dark matter experiment. 01 2012.

[129] K. Lung et al. Characterization of the Hamamatsu R11410-10 3-Inch Photomultiplier Tube for Liquid Xenon Dark Matter Direct Detection Experiments. *Nucl. Instrum. Meth. A*, 696:32–39, 2012.

[130] L Baudis, A Behrens, A Ferella, A Kish, T Marrodán Undagoitia, D Mayani, and M Schumann. Performance of the hamamatsu r11410 photomultiplier tube in cryogenic xenon environments. *Journal of Instrumentation*, 8(04):P04026–P04026, apr 2013.

[131] E. Aprile et al. Lowering the radioactivity of the photomultiplier tubes for the xenon1t dark matter experiment. *The European Physical Journal C*, 75(11):546, 2015.

[132] E. Aprile et al. Material radioassay and selection for the xenon1t dark matter experiment. *The European Physical Journal C*, 77(12):890, 2017.

[133] E. Aprile et al. The distributed slow control system of the xenon100 experiment. *Journal of Instrumentation*, 7, 11 2012.

[134] A. Daneels and W. Salter. Selection and evaluation of commercial SCADA systems for the controls of the CERN LHC experiments. *Conf. Proc. C*, 991004:353–355, 1999.

[135] J. Cardoso, B. Antunes, M. Silva, A. Rizzo, M. Piro, S. Schneider, M. Murra, D. Bar, D. Front, L. Levinson, and A. Manfredini. Experience with a slow control system based on industrial process control hardware and software for the xenon1t dark matter search. In *2016 IEEE-NPSS Real Time Conference (RT)*, pages 1–3, June 2016.

[136] E. H. Bellamy et al. Absolute calibration and monitoring of a spectrometric channel using a photomultiplier. *Nuclear Instruments and Methods in Physics Research A*, 339(3):468–476, February 1994.

[137] Francis J. Lombard and Fred Martin. Statistics of electron multiplication. *Review of Scientific Instruments*, 32(2):200–201, 1961.

[138] J.R. Prescott. A statistical model for photomultiplier single-electron statistics. *Nuclear Instruments and Methods*, 39(1):173 – 179, 1966.

[139] R. Saldanha, L. Grandi, Y. Guardincerri, and T. Wester. Model independent approach to the single photoelectron calibration of photomultiplier tubes. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 02 2016.

[140] A. Manalaysay et al. Spatially uniform calibration of a liquid xenon detector at low energies using 83m-Kr. *Rev. Sci. Instrum.*, 81:073303, 2010.

[141] V Hannen et al. Limits on the release of rb isotopes from a zeolite based83mkr calibration source for the XENON project. *Journal of Instrumentation*, 6(10):P10013–P10013, oct 2011.

[142] L. Kastens, Susana Cahn, Angel Manzur, and D. Mckinsey. Calibration of a liquid xenon detector with kr-83m. *Physical Review C*, 80, 05 2009.

[143] R.F. Lang, A. Brown, E. Brown, M. Cervantes, S. Macmullin, D. Masson, J. Schreiner, and H. Simgen. A220rn source for the calibration of low-background experiments. *Journal of Instrumentation*, 11(04):P04004–P04004, apr 2016.

[144] E. Aprile et al. Results from a Calibration of XENON100 Using a Source of Dissolved Radon-220. *Phys. Rev. D*, 95(7):072008, 2017.

[145] E. Aprile et al. Analysis of the XENON100 Dark Matter Search Data. *Astropart. Phys.*, 54:11–24, 2014.

[146] Hu Wei, Jian Fang, Bo-Xiang Yu, Xuan Zhang, Li Zhou, Xiao Cai, Li-Jun Sun, Wan-Jin Liu, Lan Wang, and Jun-Guang Lü. Laboratory measurement of radioactivity purification for 212 pb in liquid scintillator. *Chinese Physics C*, 40:096202, 09 2016.

[147] R. F. Lang, J. Pienaar, E. Hogenbirk, D. Masson, R. Nolte, A. Zimbal, S. Röttger, M. L. Benabderrahmane, and G. Bruno. Characterization of a deuterium-deuterium plasma fusion neutron generator. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 879:31–38, 2018.

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

[149] B. Lenardo, K. Kazkaz, A. Manalaysay, J. Mock, M. Szydagis, and M. Tripathi. A global analysis of light and charge yields in liquid xenon. *IEEE Transactions on Nuclear Science*, 62(6):3387–3396, Dec 2015.

[150] Tadayoshi Doke, Akira Hitachi, Jun Kikuchi, Kimiaki Masuda, Hiroyuki Okada, and Eido Shibamura. Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles. *Jap. J. Appl. Phys.*, 41:1538–1545, 2002.

[151] Carl Eric Dahl. *The physics of background discrimination in liquid xenon, and first results from Xenon10 in the hunt for WIMP dark matter*. PhD thesis, Princeton U., 2009.

[152] T. Shutt, Carl E. Dahl, J. Kwong, A. Bolozdynya, and P. Brusov. Performance and Fundamental Processes at Low Energy in a Two-Phase Liquid Xenon Dark Matter Detector. *Nucl. Instrum. Meth. A*, 579:451–453, 2007.

[153] J Lindhard, M Scharff, and H E Schioett. Range concepts and heavy ion ranges (notes on atomic collisions, ii). *Kgl. Danske Videnskab. Selskab. Mat. Fys. Medd.*, 33:14, 1 1963.

[154] Peter Sorensen and Carl Eric Dahl. Nuclear recoil energy scale in liquid xenon with application to the direct detection of dark matter. *Phys. Rev. D*, 83:063501, Mar 2011.

[155] J.D. Lewin and P.F. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics*, 6(1):87 – 112, 1996.

[156] J. Thomas and D. A. Imel. Recombination of electron-ion pairs in liquid argon and liquid xenon. *Phys. Rev. A*, 36:614–616, Jul 1987.

[157] D. S. Akerib et al. Tritium calibration of the lux dark matter experiment. *Phys. Rev. D*, 93:072009, Apr 2016.

[158] E. Aprile et al. Response of the xenon100 dark matter detector to nuclear recoils. *Phys. Rev. D*, 88:012006, Jul 2013.

[159] M. Szydagis et al. Noble element simulation technique v2.0, July 2018.

[160] E.M. Boulton et al. Calibration of a two-phase xenon time projection chamber with a37ar source. *Journal of Instrumentation*, 12(08):P08004–P08004, aug 2017.

[161] E. Aprile et al. Signal yields of kev electronic recoils and their discrimination from nuclear recoils in liquid xenon. *Phys. Rev. D*, 97:092007, May 2018.

[162] E. Aprile et al. Scintillation response of liquid xenon to low energy nuclear recoils. *Phys. Rev. D*, 72:072006, Oct 2005.

[163] E. Aprile et al. Simultaneous measurement of ionization and scintillation from nuclear recoils in liquid xenon for a dark matter experiment. *Phys. Rev. Lett.*, 97:081302, Aug 2006.

[164] E. Aprile et al. New measurement of the relative scintillation efficiency of xenon nuclear recoils below 10 kev. *Phys. Rev. C*, 79:045807, Apr 2009.

[165] G. Plante et al. New measurement of the scintillation efficiency of low-energy nuclear recoils in liquid xenon. *Phys. Rev. C*, 84:045805, Oct 2011.

[166] P. Sorensen et al. The scintillation and ionization yield of liquid xenon for nuclear recoils. *Nucl. Instrum. Meth. A*, 601:339–346, 2009.

[167] A. Manzur, A. Curioni, L. Kastens, D. N. McKinsey, K. Ni, and T. Wongjirad. Scintillation efficiency and ionization yield of liquid xenon for monoenergetic nuclear recoils down to 4 kev. *Phys. Rev. C*, 81:025808, Feb 2010.

[168] LUX Collaboration et al. Low-energy (0.7-74 kev) nuclear recoil calibration of the lux dark matter experiment using d-d neutron scattering kinematics. 08 2016.

[169] E. Aprile et al. The XENON1t data acquisition system. *Journal of Instrumentation*, 14(07):P07016–P07016, jul 2019.

[170] XENON Collaboration. The pax data processor v6.8.0, March 2018.

[171] Jelle Aalbers et al. Xenon1t/hax: hax referenced in chep, October 2018.

[172] C.H. Faham, V.M. Gehman, A. Currie, A. Dobi, P. Sorensen, and R.J. Gaitskell. Measurements of wavelength-dependent double photoelectron emission from single photons in VUV-sensitive photomultiplier tubes. *Journal of Instrumentation*, 10(09):P09010–P09010, sep 2015.

[173] B. López Paredes et al. Response of photomultiplier tubes to xenon scintillation light. *Astropart. Phys.*, 102:56–66, 2018.

[174] E. Aprile et al. First results on the scalar wimp-pion coupling, using the xenon1t experiment. *Phys. Rev. Lett.*, 122:071301, Feb 2019.

[175] E. Aprile et al. Constraining the spin-dependent WIMP-nucleon cross sections with XENON1T. *Phys. Rev. Lett.*, 122(14):141301, 2019.

[176] E. Aprile et al. Observation of two-neutrino double electron capture in ^{124}Xe with XENON1T. *Nature*, 568(7753):532–535, 2019.

[177] B. Edwards et al. Measurement of single electron emission in two-phase xenon. *Astropart. Phys.*, 30:54–57, 2008.

[178] E. Santos et al. Single electron emission in two-phase xenon with application to the detection of coherent neutrino-nucleus scattering. *JHEP*, 12:115, 2011.

[179] M. Le Calloch. *Study of the single electron charge signals in the XENON100 direct Dark Matter search experiment.* Theses, Université de Nantes, October 2014.

[180] Louis Malter. Thin film field emission. *Phys. Rev.*, 50:48–58, Jul 1936.

[181] Daniel Foreman-Mackey, David Hogg, Dustin Lang, and Jonathan Goodman. emcee: the mcmc hammer. *Publications of the Astronomical Society of the Pacific*, 125:306–312, 03 2013.

[182] David J. C. MacKay. *Information Theory, Inference amp; Learning Algorithms.* Cambridge University Press, USA, 2002.

[183] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 04 1970.

[184] Andrii Ilienko. Continuous counterparts of poisson and binomial distributions and their properties. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Computatorica*, 39, 03 2013.

[185] Hamamatsu Photonics. *Photomultiplier Tubes: Basics and Applications.* Hamamatsu City, Japan, second edition edition, 1999.

[186] Walter Blum, Werner Riegler, and Luigi Rolandi. *Particle detection with drift chambers; 2nd ed.* Springer, Berlin, 2008.

[187] T. H. V. T. Dias, F. P. Santos, A. D. Stauffer, and C. A. N. Conde. Monte carlo simulation of x-ray absorption and electron drift in gaseous xenon. *Phys. Rev. A*, 48:2887–2902, Oct 1993.

[188] Masayo Suzuki and Shinzou Kubota. Mechanism of proportional scintillation in argon, krypton and xenon. *Nuclear Instruments and Methods*, 164(1):197 – 199, 1979.

[189] Robert S. Mulliken. Potential curves of diatomic rare-gas molecules and their ions, with particular reference to xe2. *The Journal of Chemical Physics*, 52(10):5170–5180, 1970.

[190] E. Aprile et al. Physics reach of the XENON1t dark matter experiment. *Journal of Cosmology and Astroparticle Physics*, 2016(04):027–027, apr 2016.

[191] C M B Monteiro et al. Secondary scintillation yield in pure xenon. *Journal of Instrumentation*, 2(05):P05001–P05001, may 2007.

[192] John David Jackson. *Classical electrodynamics.* Wiley, New York, NY, 3rd ed. edition, 1999.

[193] I.M. Obodovskii E.M. Gushchin, A.A. Kruglov. Emission of "hot" electrons from liquid and solid argon and xenon. *Sov. Phys. JETP*, 55(5):860, 1982.

[194] Morrel H. Cohen and J. Lekner. Theory of hot electrons in gases, liquids, and solids. *Phys. Rev.*, 158:305–309.

[195] Peter Sorensen. Electron train backgrounds in liquid xenon dark matter search detectors are indeed due to thermalization and trapping. *arXiv e-prints*, page arXiv:1702.04805, February 2017.

[196] U. Sowada, G. Bakale, K. Yoshino, and W. F. Schmidt. Electric field effect on electron capture by sf6 in liquid argon and xenon. *Chem. Phys. Lett.*, 34:466–469, 1975.

[197] J. Zemann. Crystal structures, 2nd edition. vol. 1 by r. w. g. wyckoff. *Acta Crystallographica*, 18(1):139–139, 1965.

[198] E. Aprile et al. First dark matter search results from the xenon1t experiment. *Phys. Rev. Lett.*, 119:181301, Oct 2017.

[199] Katherine Freese, Mariangela Lisanti, and Christopher Savage. Annual modulation of dark matter: A review. 09 2012.

[200] Francis Froborg and Alan R. Duffy. Annual modulation in direct dark matter searches. *Journal of Physics G: Nuclear and Particle Physics*, 2020.

[201] Martin C. Smith et al. The RAVE survey: constraining the local Galactic escape speed. *Monthly Notices of the Royal Astronomical Society*, 379(2):755–772, 07 2007.

[202] Ralph Schönrich, James Binney, and Walter Dehnen. Local kinematics and the local standard of rest. *Monthly Notices of the Royal Astronomical Society*, 403(4):1829–1833, 04 2010.

[203] Jo Bovy, David W. Hogg, and Hans-Walter Rix. GALACTIC MASERS AND THE MILKY WAY CIRCULAR VELOCITY. *The Astrophysical Journal*, 704(2):1704–1709, oct 2009.

[204] James Binney. Distribution functions for the Milky Way. *Monthly Notices of the Royal Astronomical Society*, 401(4):2318–2330, 01 2010.

[205] Anne Green. Effect of realistic astrophysical inputs on the phase and shape of the weakly interacting massive particles annual modulation signal. *Phys. Rev. D*, 68, 07 2003.

[206] David N. Spergel. Motion of the earth and the detection of weakly interacting massive particles. *Phys. Rev. D*, 37:1353–1355, Mar 1988.

[207] A. Zygmund and Robert Fefferman. *Trigonometric Series*. Cambridge Mathematical Library. Cambridge University Press, 3 edition, 2003.

[208] V. Belokurov et al. The Field of Streams: Sagittarius and Its Siblings. , 642(2):L137–L140, May 2006.

[209] M Kuhlen, J Diemand, P Madau, and M Zemp. The via lactea INCITE simulation: galactic dark matter substructure at high resolution. *Journal of Physics: Conference Series*, 125:012008, jul 2008.

[210] R. Bernabei et al. Performances of the 100 kg nai(tl) set-up of the dama experiment at gran sasso. *Il Nuovo Cimento A (1971-1996)*, 112(6):545–575, 1999.

[211] R. Bernabei et al. The dama/libra apparatus. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 592:297–315, 04 2008.

[212] Rita Bernabei et al. First model independent results from dama/libra-phase2. *Universe*, 4:116, 11 2018.

[213] R. Bernabei et al. New results from dama/libra. *The European Physical Journal C*, 67(1):39–49, 2010.

[214] Sebastian Baum, Katherine Freese, and Chris Kelso. Dark matter implications of dama/libra-phase2 results. *Physics Letters B*, 789:262 – 269, 2019.

[215] Peter Sorensen. Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 635(1):41 – 43, 2011.

[216] James H. Parker and John J. Lowke. Theory of electron diffusion parallel to electric fields. i. theory. *Phys. Rev.*, 181:290–301, May 1969.

[217] P. Barrow et al. Qualification tests of the r11410-21 photomultiplier tubes for the XENON1t detector. *Journal of Instrumentation*, 12(01):P01024–P01024, jan 2017.

[218] G. Audi, O. Bersillon, J. Blachot, and A.H. Wapstra. The nubase evaluation of nuclear and decay properties. *Nuclear Physics A*, 729(1):3 – 128, 2003. The 2003 NUBASE and Atomic Mass Evaluations.

[219] Edwin B. Wilson. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158):209–212, 1927.

[220] E. Aprile et al. Intrinsic backgrounds from rn and kr in the xenon100 experiment. *The European Physical Journal C*, 78(2):132, 2018.

[221] K. J. P. Schroeder and W. Roether. *The releases of krypton-85 and tritium to the environment and tritium to krypton-85 ratios as source indicators.* IAEA, International Atomic Energy Agency (IAEA), 1975.

[222] Katja Winger, J Feichter, Martin Kalinowski, Hartmut Sartorius, and C Schlosser. A new compilation of the atmospheric 85krypton inventories from 1945 to 2000 and its evaluation in a global transport model. *Journal of environmental radioactivity*, 80:183–215, 02 2005.

[223] Peter Cook and Andrew Herczeg. *Environmental Tracers in Subsurface Hydrology*. 01 2000.

[224] E. Aprile et al. Removing krypton from xenon by cryogenic distillation to the ppq level. *The European Physical Journal C*, 77(5):275, 2017.

[225] Sebastian Lindemann and Hardy Simgen. Krypton assay in xenon at the ppq level using a gas chromatographic system and mass spectrometer. *The European Physical Journal C*, 74(2):2746, 2014.

[226] Aldo M. Serenelli, W. C. Haxton, and Carlos Peña-Garay. Solar Models with Accretion. I. Application to the Solar Abundance Problem. , 743(1):24, December 2011.

[227] G. Bellini et al. First evidence of pep solar neutrinos by direct detection in borexino. *Physical review letters*, 108:051302, Feb 2012.

[228] M. Agostini et al. Seasonal modulation of the 7be solar neutrino rate in borexino. *Astroparticle Physics*, 92:21 – 29, 2017.

[229] J. B. Albert et al. Improved measurement of the $2\nu\beta\beta$ half-life of ^{136}xe with the exo-200 detector. *Phys. Rev. C*, 89:015502, Jan 2014.

[230] Thomas J. DiCiccio and Bradley Efron. Bootstrap confidence intervals. *Statist. Sci.*, 11(3):189–228, 09 1996.

[231] Phillip I. Good. *Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer Series in Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2004.

[232] Bradley Efron. *The jackknife, the bootstrap and other resampling plans*. SIAM, 1982.

[233] Jerzy Neyman and Egon Sharpe Pearson. On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Phil. Trans. Roy. Soc. Lond. A*, 231(694-706):289–337, 1933.

[234] Maurice G. Kendall, Alan Stuart, and J. Keith Ord. *The advanced theory of statistics: in three volumes. 2: Inference and relationship*. Griffin, London, 4. ed edition, 1979. OCLC: 174487424.

[235] George Davey Smith and Shah Ebrahim. Data dredging, bias, or confounding. *BMJ (Clinical research ed.)*, 325(12493654):1437–1438, December 2002.

[236] Eilam Gross and Ofer Vitells. Trial factors for the look elsewhere effect in high energy physics. *Eur. Phys. J. C*, 70:525–530, 2010.

[237] F. James. *Statistical Methods in Experimental Physics*. World Scientific, 2006.

[238] Leslie E. Ballentine. *Quantum mechanics: a modern development*. World Scientific, Singapore, repr edition, 2010. OCLC: 846445677.

[239] Nima Arkani-Hamed, Douglas P. Finkbeiner, Tracy R. Slatyer, and Neal Weiner. A theory of dark matter. *Phys. Rev. D*, 79:015014, Jan 2009.

[240] Nima Arkani-Hamed and Neal Weiner. Lhc signals for a superunified theory of dark matter. *Journal of High Energy Physics*, 2008, 10 2008.

[241] Clifford Cheung, Joshua T. Ruderman, Lian-Tao Wang, and Itay Yavin. Kinetic Mixing as the Origin of Light Dark Scales. *Phys. Rev. D*, 80:035008, 2009.

[242] Rouven Essig, Jared Kaplan, Philip Schuster, and Natalia Toro. On the origin of light dark matter species. 04 2010.

[243] Bob Holdom. Two u(1)'s and charge shifts. *Physics Letters B*, 166(2):196 – 198, 1986.

[244] R. Foot. Mirror dark matter: Cosmology, galaxy structure and direct detection. *International Journal of Modern Physics A*, 29(11n12):1430013, 2014.

[245] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, 104:254–258, Oct 1956.

[246] Jonathan L. Feng and Jason Kumar. The WIMPless Miracle: Dark-Matter Particles without Weak-Scale Masses or Weak Interactions. *Phys. Rev. Lett.*, 101:231301, 2008.

[247] Michael Dine, Ann E. Nelson, and Yuri Shirman. Low energy dynamical supersymmetry breaking simplified. *Phys. Rev. D*, 51:1362–1370, Feb 1995.

[248] Adam Falkowski, Joshua T. Ruderman, and Tomer Volansky. Asymmetric dark matter from leptogenesis. *Journal of High Energy Physics*, 2011(5):106, 2011.

[249] Pierre Jean, J. Knoedlseder, V. Lonjou, Marc Allain, J. Roques, Gerry Skinner, B. Teegarden, G. Vedrenne, P. Ballmoos, B. Cordier, P. Caraveo, Roland Diehl, Ph Durouchoux, P. Mandrou, J. Matteson, N. Gehrels, V. Schönfelder, A. Strong, P. Ubertini, and Christoph Winkler. Early spi/integral measurements of galactic 511 kev line emission from positron annihilation. *Astron. Astrophys.*, 407, 09 2003.

[250] Dan Hooper and Kathryn Zurek. A natural supersymmetric model with mev dark matter. *Physical Review D*, 77, 01 2008.

[251] Celine Boehm, Dan Hooper, Joseph Silk, Michel Casse, and Jacques Paul. Mev dark matter: Has it been detected? *Physical review letters*, 92:101301, 04 2004.

[252] Laura Covi, Jihn E. Kim, and Leszek Roszkowski. Axinos as cold dark matter. *Phys. Rev. Lett.*, 82:4180–4183, May 1999.

[253] Laura Covi. Axinos as Cold Dark Matter. *AIP Conf. Proc.*, 878(1):145–151, 2006.

[254] Laura Covi and Jihn E Kim. Axinos as dark matter particles. *New Journal of Physics*, 11(10):105003, oct 2009.

[255] Rouven Essig, Marivi Fernández-Serra, Jeremy Mardon, Adrián Soto, Tomer Volansky, and Tien-Tien Yu. Direct detection of sub-gev dark matter with semiconductor targets. *Journal of High Energy Physics*, 2016(5):46, 2016.

[256] S. Kubota, A. Nakamoto, T. Takahashi, S. Konno, T. Hamada, M. Miyajima, A. Hitachi, E. Shibamura, and T. Doke. Evidence of the existence of exciton states in liquid argon and exciton-enhanced ionization from xenon doping. *Phys. Rev. B*, 13:1649–1653, Feb 1976.

[257] I. T. Steinberger and U. Asaf. Band-structure parameters of solid and liquid xenon. *Phys. Rev. B*, 8:914–918, Jul 1973.

[258] U. Asaf and I. T. Steinberger. Photoconductivity and electron transport parameters in liquid and solid xenon. *Phys. Rev. B*, 10:4464–4468, Nov 1974.

[259] C.F. Bunge, J.A. Barrientos, and A.V. Bunge. Roothaan-hartree-fock ground-state atomic wave functions: Slater-type orbital expansions and expectation values for z = 2-54. *Atomic Data and Nuclear Data Tables*, 53(1):113 – 162, 1993.

[260] C. C. J. Roothaan. New developments in molecular orbital theory. *Rev. Mod. Phys.*, 23:69–89, Apr 1951.

[261] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications, 1965.

[262] M. Morita. Theory of Beta Decay. *Progress of Theoretical Physics Supplement*, 26:1–63, 03 1963.

[263] P. Agnes et al. Constraints on sub-gev dark-matter–electron scattering from the darkside-50 experiment. *Phys. Rev. Lett.*, 121:111303, Sep 2018.

[264] A. Messiah. *Quantum Mechanics*. Number vol. 2 in Dover books on physics. Dover Publications, 1999.

[265] Rouven Essig, Tomer Volansky, and Tien-Tien Yu. New constraints and prospects for sub-gev dark matter scattering off electrons in xenon. *Phys. Rev. D*, 96:043017, Aug 2017.

[266] E. Aprile et al. Xenon1t dark matter data analysis: Signal and background models and statistical inference. *Phys. Rev. D*, 99:112009, Jun 2019.

[267] D.S. Akerib et al. Investigation of background electron emission in the LUX detector. 4 2020.

[268] P. Sorensen and Kate Kamdin. Two distinct components of the delayed single electron noise in liquid xenon emission detectors. *Journal of Instrumentation*, 13:P02032–P02032, 02 2018.

[269] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[270] Xavier Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. *Journal of Machine Learning Research - Proceedings Track*, 9:249–256, 01 2010.

[271] Steve Baker and Robert D. Cousins. Clarification of the use of CHI-square and likelihood functions in fits to histograms. *Nuclear Instruments and Methods in Physics Research*, 221(2):437–442, April 1984.

[272] E. Aprile et al. Emission of Single and Few Electrons in XENON1T and Limits on Light Dark Matter. *arXiv e-prints*, 12 2021.