Modelling Delayed Electron Emission in XENONnT

by

Conor Ford

Examiner: Jelle Aalbers

Second Examiner: Kristof De Bruyn

Master's Thesis

To fulfil the requirements for the degree of

Master's of Science in Physics

at the University of Groningen

Groningen, November 2025

Abstract

Recent improvements in background mitigation in XENONnT have made it one of the most sensitive experiments worldwide for the direct detection of dark matter. This has extended the search into lower-mass regions, in the order of keV, where new backgrounds become prominent. Among these is single-electron emission, a prominent background of all dual-phase time projection chambers, which can mimic the signal pattern of low-mass dark matter particles. In particular, there exists a delayed emission component which continues for $\mathcal{O}$(s) after large signals in the detector, greatly hindering this search. This work utilises a new method of modelling this background, first developed in Ref. [1], in order to best understand the rate of this delayed electron emission. A new method of data selection is developed in order to maximise the purity of the results. This work develops extensions to the model in the form of a radial function and a numerical count of electrons per signal, and tests these extensions against the original model. This model shows much promise in helping to lower the background limit in the low-mass regime, and would also likely find applicability in other dual-phase time projection chambers.

Contents

1 Introduction

The Standard Model is, at present, the best widespread framework we have with which we can begin to understand the universe and the nature of its composition. Each iteration and addition to it has withstood strict and thorough experimental testing, allowing us each time to further push the boundaries of science; every test either incorporating great discoveries such as the Higgs boson, or else rejecting that which does not hold up to the rigorous scrutiny of the model.

This process means the Standard Model now stands as a testament to the work of countless great physicists, allowing us to explain a huge extent of the physical phenomena observed before us. However, despite these successes, there still remain many open questions with observed phenomena that do not fit into the model; chiefly, that of dark matter, dark energy, and the unification of the four fundamental forces.

When astronomers look out into space and observe the planets and stars; the galaxies and their clusters; they are actually only seeing a small fraction of the total mass-energy content of the universe. About 5% of this total is what we think of as 'normal' matter, called baryonic matter; and in fact, these visible features of stars, planets, and clusters only make up about 6% of that, meaning that the vast majority of this baryonic matter is in the form of interstellar gas: diffuse, ionised clouds circulating around and between galaxies, almost impossible to detect directly [2].

This small fraction that makes up visible baryonic matter, only 0.3% of the universe's mass-energy content, contextualises how much of the universe is invisible or unknown to us. To account for the remaining ∼95% of the universe we must look to the mysterious elements of dark energy and dark matter. Dark energy is, put simply, a hypothesised unknown energy responsible for driving the accelerating expansion of the universe. Aside from mentioning that it makes up ∼69% of the universe, it shall otherwise be put to the side for this thesis, leaving it for the cosmologists and astronomers to work on.

Thus the focus falls now on dark matter, and it is a curious reality faced: approximately five-sixths of all matter, or ∼26.4% of the universe's energy-mass content [3], is so far seemingly invisible to any form of direct observation - revealing itself only in the indirect gravitational effects of celestial bodies, and in underlying structural features which remain unchanged from the early universe. Despite much ongoing effort, there has not yet been any unequivocal evidence for, or detection of, anything that may be considered to be a candidate for this substantial missing chunk.

Dark matter has played a key role in the formation and evolution of the universe, laying gravitational seeds soon after the Big Bang for the large-scale structures of the current universe [4, 5], and while it continues to exert its subtle influence on baryonic matter in the current universe, very little else is

understood about it, or how it might else be observed. Many forms of searches for dark matter are ongoing, both direct and indirect; indirect searches focusing on searching for signals of dark matter annihilation or decay into standard model particles [5, 6]; direct searches searching to discover dark matter particles through their theorised collision and scattering with baryonic matter, which would allow for the measurement of their mass, among other things [5, 7]. Sec. 2 talks more about the theories around dark matter particles and how they might be observed directly.

One way direct searches for dark matter are conducted comes in the form of large dual-phase liquid xenon time projection chambers (LXe TPCs), one of which is the XENONnT detector. The scattering of incident particles with the liquid xenon creates scintillation and ionisation signals to which these types of detectors are highly sensitive. These signals are then analysed to undergo extensive position and energy reconstruction processes in order to categorise each signal and interaction, looking for a potential dark matter signal. Sec. 3 explores exactly how these detectors work in more detail.

XENONnT and its contemporaries are most sensitive in detecting collisions from particles of mass in the GeV/c^2 regime; however, as stronger upper limits are set on the potential cross-sectional area vs. mass by these experiments, the search is increasingly being pushed down into the sub-GeV/c^2 mass range [8]. This poses many problems for the direct search for dark matter, chief among them the matter of an increasing background, resulting from both the detector itself and external factors.

This background is because lighter dark matter particle masses result in lower energy depositions when scattering within the detector, resulting in signals which are harder to pick up amongst the background noise. Below a certain point the dominant background becomes neutrinos, which we are unable to shield against, and their extremely infrequent scattering interactions with the xenon in this regime can mimic low-mass dark matter signals to the point of being nearly indistinguishable [7]. Sec. 3.3 talks more about the low-mass dark matter regime and the problems faced in that search.

One prominent and problematic background present in the lower-mass regime is that of single- or few-electron emission, and in particular that of an increased rate of these signals which are 'delayed' after a scattering event in the detector, between a couple milliseconds to $\mathcal{O}(100)$ms after, presenting as a power law decay model [9–11]. These signals have been detected since some of the first LXe TPCs [12, 13], and even in ones not using xenon [14, 15]. The problem is complex in that we have a few competing hypotheses for it, and furthermore, there is no characteristic timescale to the power law modelling - the emissions do not seem to end some given time after large S2s - meaning we cannot adequately distinguish the backgrounds between any two given interactions, even theoretically those separated by many seconds. Sec. 4 looks more at the background and nature of this problem.

Importantly, these signals mimic those that would be seen from low-mass dark matter particles, as random single- or few-electron signals within the detector, and thus it is vital to the low-mass regime search that this background is accurately modelled and understood, in order for it to be minimised in

analysis. This work presents a method of modelling this delayed electron, first developed in Ref. [1], by which we may model the rate of this delayed electron emission.

Sec. 2 presents the general search for dark matter, and the context of the field within which XENONnT lies; Sec. 3 then describes the experiment itself and the physics at play within the detector. This gives context to the context in which this delayed electron emission takes place, before we outline the problem in Sec. 4, and analyse the evidence for the hypotheses put forward as the possible sources. The main body of work is in Secs. 6, 7, and 8, where we develop a data selection methodology for the model; attempt to improve and extend the model; and then test those extensions to the model against the original of Ref. [1] in order to understand how this model may be best used going forward. Sec. 8 also summarises the successes and challenges faced by this model, and offers improvements which could be made to it in future research.

2 The Search for Dark Matter

2.1 Discovery and Evidence for Dark Matter

In 1922, Jacobus Kapetyn was making observations of the stars in the Milky Way when he realised that their total mass was insufficient to produce the gravitational field required to sustain them in their observed orbits around the galaxy; here we find one of the first references to a 'dark matter' [16] - some invisible mass present throughout the universe, exerting its gravitational force on the observable parts of it. Later understanding of the phenomenon was added by Oort in examining the motion of stars perpendicular to the galactic plane [17] and by Zwicky in his study of the Coma galaxy cluster [18, 19], however their hypotheses still only extended to dark matter just being some 'non-luminous' form of ordinary matter.

It was not until decades later that more evidence came in from studies on the rotation curves of galaxies, and on the Cosmic Microwave Background (CMB) radiation, that the nature of dark matter as this completely invisible, non-interacting substance found in 'halos' around the majority of galaxies, began to be better understood [4, 20, 21]. The CMB is radiation measured from approximately 3.7×10^6 years after the Big Bang [22, 23]. Before this, photons and baryons were coupled in the hot plasma of the early universe, and minute ($\mathcal{O}(10^{-5})$) fluctuations in gravity and pressure caused acoustic oscillations throughout this fluid, resulting in over- and under- energy-dense regions [3]. As the universe cooled the baryons and photons gradually decoupled, leaving the CMB as a slightly-blurred screenshot of these regions over this period - an imprint of the early universe still detectable today [3].

The picture of the CMB we have, and the peaks we observe in its power spectrum, reflect the acoustic oscillations of the photon-baryon plasma and the density variations left after recombination, which can reveal the ratio of baryonic matter to the total amount of matter, and therefore about the abundance of dark matter in the universe [3]. Furthermore, although we do not yet know if dark matter is entirely stable across large cosmological timescales - and indeed its possible decay is the focus of many indirect searches [6] - it is generally calculated to be stable at timespans on the order of the age of the universe [24], and thus the information encoded by the CMB is accurate for calculations of its current-day abundance.

In the early universe, it would still have been far too hot for baryonic matter to effectively start forming larger gravitational objects on its own, and so dark matter, not being affected by coupling to photons in the early universe and thus being able to clump together faster, must have played a role in sowing the seeds for early structure formation [5, 25]. Dark matter was still subject to density perturbations however, and so during and after recombination, over-dense regions of dark matter

grew to a critical density, eventually collapsing into halos - gravitationally bound regions of dark matter in which baryonic matter could begin to be bound into the galaxies and clusters that would eventually be observed today [4, 5]. It is within one of these halos that the Solar System lies, and therefore there should be a constant flux of dark matter passing through the earth, which we may possibly detect via an experiment such as XENONnT.

2.2 Models for Dark Matter

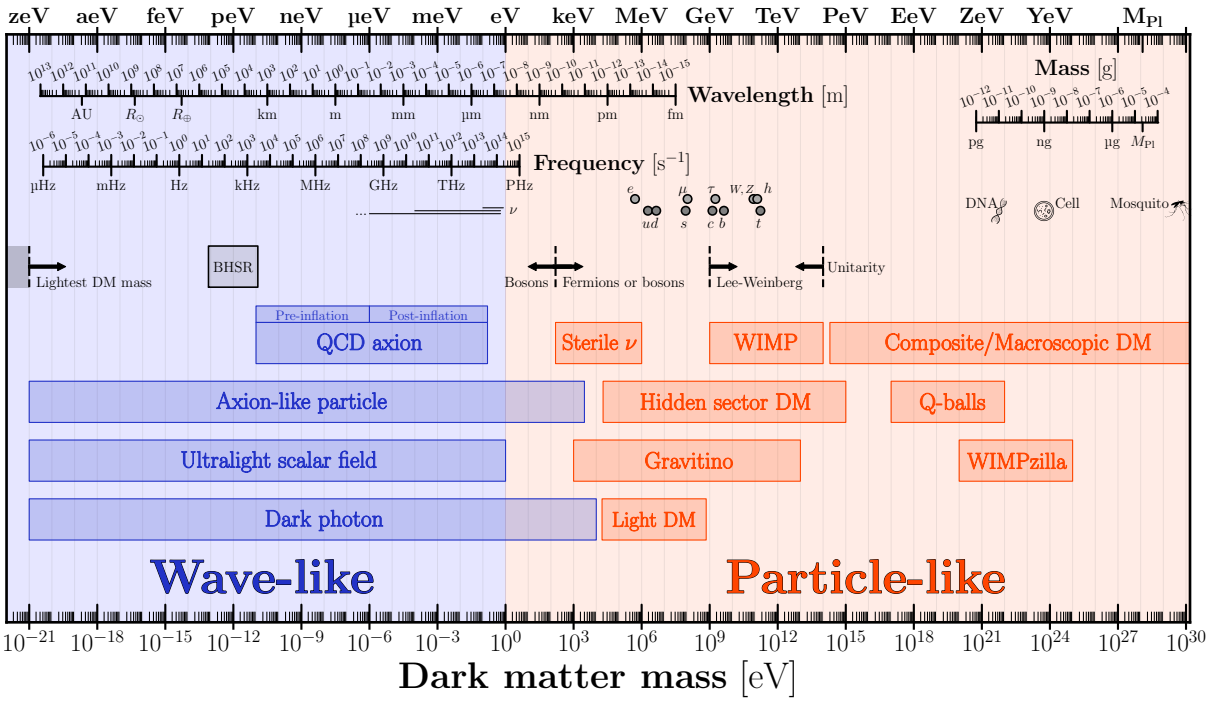


Figure 1: Overview of the mass scales of potential dark matter candidates, from [26]. In XENONnT we are most interested in the candidate of 'WIMPs', in the 1 GeV - 1 TeV mass range.

The theory that best explains this evidence is what is now known as ΛCDM: Λ for a cosmological constant associated with dark energy, coupled with "Cold Dark Matter". Dark matter can be referred to as 'cold' or 'hot' depending on its thermal velocity in the early universe, non-relativistic or relativistic respectively; the most widely-supported theories currently favour a 'cold' model in order to best account for observed structure formation. If all dark matter were 'hot', structures would be spread further and differently than we observe, and the CMB anisotropies would be much larger in size [5].

From this evidence, much can be inferred about the nature and properties of dark matter. We expect it to be neutral [27]; with an extremely small interaction cross-section (as a collision with any baryonic matter has not yet been observed); and while the nature of its colour-charge, if any, is unknown yet, no coupling to a Standard Model particle has yet been observed [6, 28], aside from

its known gravitational effects on baryonic matter. There is currently no Standard Model particle that fits all of these criteria. Neutrinos were one proposed hypothesis for a time, being electrically neutral, stable, and massive, but limits set on their masses show that they would have been too 'hot' in the early universe to allow for structure formation [5], and therefore could not be adequate dark matter candidates. Along with neutrinos, many theories have been proposed and ruled out since the discovery of dark matter, and the remaining face increasing constraints as more and more experiments help narrow down the parameters. Fig. 1 shows the wide range of dark matter candidates and their potential mass-ranges; from the microscopic to the macroscopic. More comprehensive reviews of dark matter candidates can be found at Refs. [3, 4].

One of the candidates for a cold dark matter particle is that of a Weakly Interacting Massive Particle (WIMP), the name indicating theoretical couplings via a force that is as strong as, or weaker than, the weak force. WIMPs are predicted to have been created thermally in the early universe, coupling and annihilating with baryonic matter in that hot thermal plasma, before the universe cooled to a point where these interactions were eventually "frozen out", locking the relative abundance of dark matter [3, 5]. If we assume dark matter particles to have a mass on the electroweak scale ($\sim$1 GeV to 1 TeV, motivated by similarities to other electroweak particles and as well as other early-universe conditions) and a velocity-dependent annihilation cross-section of $\langle \sigma v \rangle \sim 10^{-26} \mathrm{cm}^3/\mathrm{s}$ (again a value motivated by similarities to other electroweak particles), then the values we predict from the thermal freeze-out as a result very closely match the values we get for the relic abundances of dark matter today [4]. This remarkable coincidence is known as the "WIMP miracle", and is what has helped lend a lot of support to the WIMP hypothesis.

It is dark matter in the mass-range of WIMPs which direct-detection experiments such as XENONnT are best set up to detect, yielding high sensitivity in the GeV - TeV range of its predicted mass [29], even possibly probing to the lower-mass MeV regime. The predicted cross-section of WIMPs also make them the most likely to be detected within detectors such as XENONnT, due to the similarity with standard model interaction cross-sections. The next section will explore the details of how the XENONnT experiment works, and how it probes these regions for WIMPs.

3 The XENONnT Experiment

The XENONnT experiment is the latest vessel for the direct detection of dark matter in a multi-staged programme by the XENON collaboration, building on its predecessor, XENON1T. While XENONnT inherited a lot of components from XENON1T in an effort to conserve cost and time in the upgrade, there were a variety of novel methods for background shielding implemented in its construction, such as a neutron veto and radon removal plant (neither present in XENON1T). These, along with a much larger total xenon mass, have made XENONnT more than an order of magnitude more sensitive to the search for WIMPs than XENON1T was, taking it to be one of the most highly sensitive experiments in the field to date [30].

3.1 The Dual-Phase Time Projection Chamber

At the core of the XENONnT experiment is a dual-phase time projection chamber (TPC). 'Dual-phase' refers to the two states of xenon present: the liquid (LXe) and gas (GXe) layers seen in Fig.2. 'Time projection chamber' refers to the ability to reconstruct the 3D position of a given collision within the detector by the hit-patterns of the Photo-Multiplier Tube (PMT) arrays at the top and bottom of the TPC; along with the time discrepancy between the first signal in the LXe (the S1), and the second signal in the GXe (the S2). These signals are explained in Fig. 2. The TPC consists of a cylinder 134 cm across and 148 cm long, with the 5.9 tonne LXe target volume below a GXe layer, and the entire cryostat containing a total 8.5 tonnes of xenon [30]. It is surrounded on the sides by various background shielding methods (see Sec. 3.2), and on the top and bottom by PMT arrays, each covered by screening meshes to reduce the electric fields immediately in front of the PMTs [30]. An anode and cathode placed inside the TPC create a drift field running along the length of the detector [30].

When a particle enters the TPC, it has the possibility of colliding with either a xenon nucleus (called nuclear recoil) or an electron (electronic recoil). This results in a deposition of energy by the particle, creating prompt scintillation light and a cloud of ionised electrons. This scintillation light, the S1, is detected immediately by the photomultiplier tubes (PMTs). The electron cloud then drifts upward to the top of the LXe under the electric field, and is extracted into the GXe, where the electrons collide with the gaseous xenon atoms to produce proportional secondary scintillation light (also called electroluminescence), a much larger signal labelled as the S2. The electron cloud experiences minimal drift in the x-y plane, so the PMT hit-pattern of the signals encode the x-y interaction coordinates, while the drift time between the S1 and S2 gives a z coordinate of depth. In conventional searches, a collision with both an S1 and a clear S2 following it is called an 'event'. An exemplar event within the TPC is shown in Fig. 2. For a more detailed description of these processes see Refs. [30, 31].

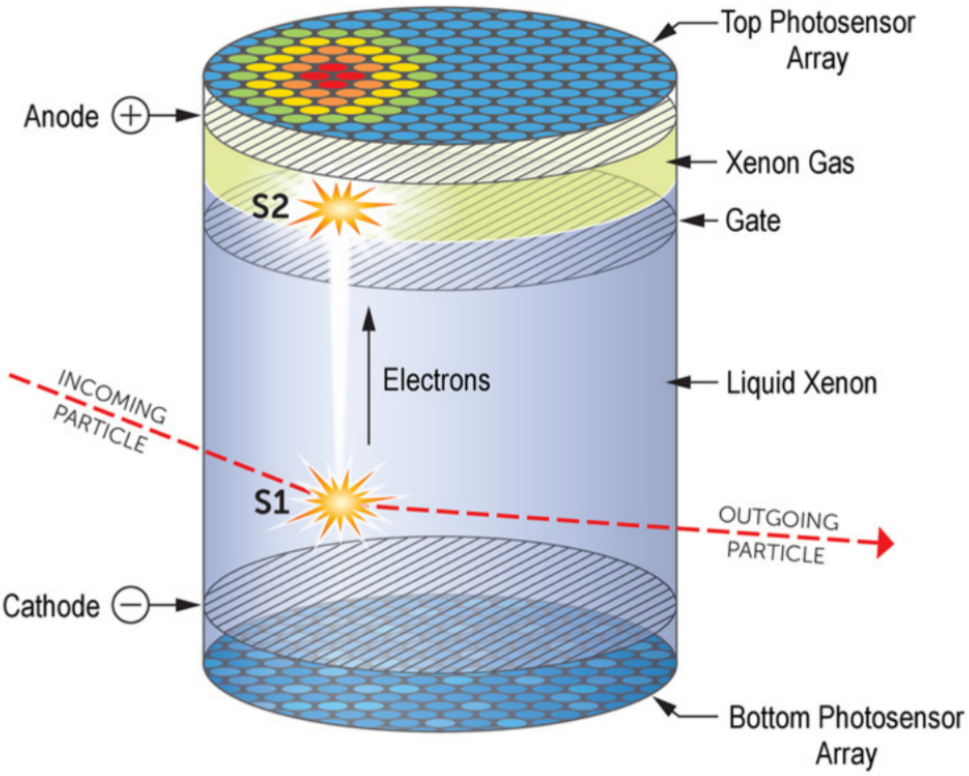


Figure 2: An example of a particle interaction and subsequent signal detections within a dual-phase TPC. An incident particle collides with the LXe, depositing energy and causing prompt scintillation light (an S1). The drift track of this collision ionises electrons in the nearby area, causing a cloud of these ionised electrons to drift upward due to the electric field applied along the length of the TPC, eventually being extracted into the gaseous xenon layer (GXe) to create a second burst of proportionally larger prompt scintillation light (an S2). Features such as the hit patterns on the top and bottom Photo-Multiplier Tube (PMT) arrays, as well as the time difference between the S1 and S2, can allow for precise position-reconstruction of the collision. Figure from [31].

Xenon was chosen as the medium for this series of experiments for several key reasons:

- It is the heaviest stable noble element (functionally, with the only unstable isotopes having half-lives many orders greater than the age of the universe [32]), meaning it is chemically inert, does not produce high background rates of its own, and can be purified with relative efficiency [31, 33].
- At an atomic mass of a $A \approx 131$ ($\simeq 122$ GeV/c^2), it is comparable in size to the theoretical mass of a WIMP, giving a large coherent scattering cross-section for spin-independent interactions, a factor that scales roughly with A^2 [34].
- As a dense, heavy element it self-shields against external radiation and background extremely well, allowing for a fiducialisation of the volume in later research - i.e. excluding some outer volume of the detector from analysis, thus cutting out the highest areas of background rate [31]. We will make use of this fact when extending the model in Sec. 7.3.2.
- An alternative medium like argon would result in less clear signals with WIMPs (due to the lower A^2 factor), and does not self-shield as well, meaning xenon detectors can be smaller by a factor of 5-7 [7].
- Xenon scintillates: when excited by a particle interaction, the subsequent decay of that xenon excimer produces ultraviolet light of 178 nm, to which ground-state xenon is transparent. This allows the light to propagate through the detector and be picked up by the PMTs [33].

Together, these properties make xenon an exceptionally well-suited medium for WIMP searches.

Nuclear recoil (NR) and electronic recoil (ER) events produce distinct signal patterns, with NRs typically being from the neutron background, and ERs typically being from β and γ decay [30]. These produce different ionisation tracks, meaning they can be distinguished between by using the S2/S1 ratios of these different signals. Nuclear recoils result in shorter, more dense ionisation tracks than ERs, with more recombination effects. This results in larger S1s and smaller S2s as compared to ERs. XENON1T was found to have a 99.7% efficiency in rejecting ER interactions, with a 50% NR acceptance [8, 30]. This discrimination allows for the isolation of potential WIMP signals from background events.

The TPC is highly sensitive and versatile for detecting potential dark matter particles. From its construction; to the technology surrounding it; to the choice of medium; all give it extended capabilities to prevent and remove external contamination - the key problem facing the direct search for dark matter. The next section looks at how background signals are mitigated for in XENONnT, both within the TPC and in the wider experiment.

3.2 Background Signals in XENONnT

For direct detection experiments such as XENONnT, the prevention and removal of background signals is essential to the search for dark matter. This suppression is ongoing and occurs at many stages of the experiment, from the location and hardware; to the data acquisition, processing, and analysis. Fig. 3 shows the complexity of the physical XENONnT system, before consideration of the extensive software infrastructure used to process the raw data, and to maintain continuous operation of the experiment.

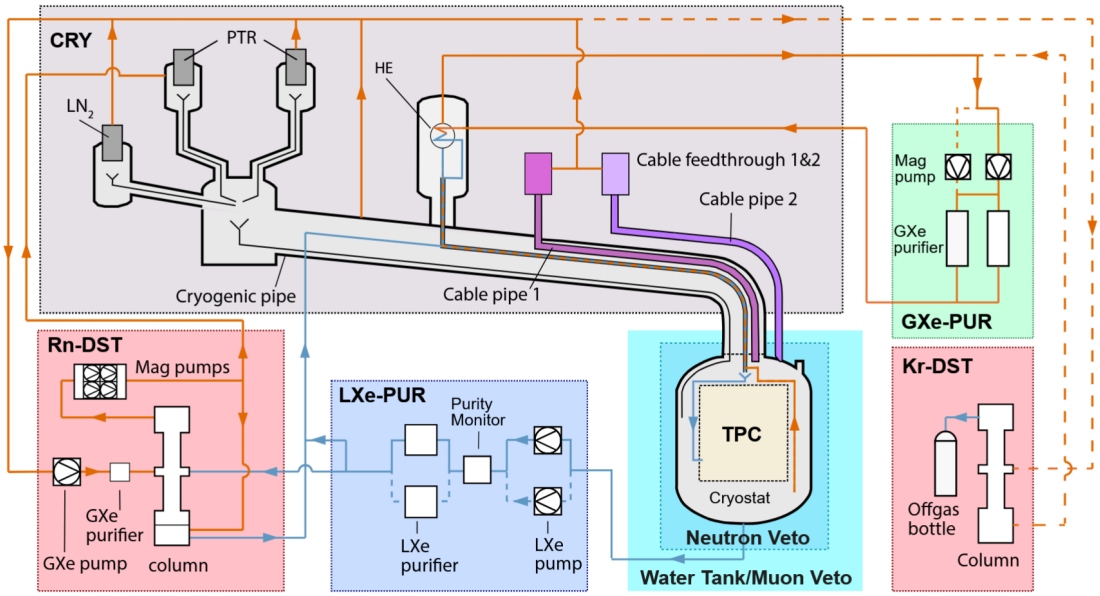


Figure 3: A schematic diagram of the physical XENONnT system. Of particular note is the TPC itself, surrounded by the muon and neutron vetos. Also features are the various subsystems used to purify the xenon. Blue arrows indicate the path of the liquid xenon, while orange ones the gaseous. From [30].

Firstly, extensive hardware measures are taken to minimise background. The location of the detector at the Laboratori Nazionali del Gran Sasso under the Apennine mountains of central Italy provides 1400m of rock coverage (3100 m water equivalent), which leads to a reduction in cosmic ray muon flux of $\mathcal{O}(10^6)$ relative to the surface [35–37]. The choice of xenon as the detector medium provides the unique benefits explained previously; while much care was put into the selection of materials for the detector, such as using polytetrafluoroethylene (PTFE) in the walls of the TPC to reflect UV-photons.

The TPC is enclosed by a water tank with two shielding sections: the muon and neutron vetos. These are separated by reflective foils, although the water itself is shared between. The muon veto, the outer layer, uses PMTs to detect high-energy cosmic-ray muons by the Cherenkov radiation they create as they pass through the detector, tagging them. Inside that is the neutron veto, designed to

flag neutrons that are captured by the hydrogen nuclei of the water, which then release gamma-rays that in turn generate Cherenkov radiation. The water is doped with gadolinium to enhance neutron capture efficiency, as gadolinium possesses highest capture cross-section for thermal neutrons of any element [30, 38].

Alongside these chambers are several purification and distillation systems which remove radioactive and chemical impurities. The LXe and GXe must be purified of electronegative impurities like H_2O and O_2 which dissolve into the LXe and trap drifting electrons, a problem that is discussed more in Sec. 3.3. The LXe purification is another upgrade from XENON1T, where only the GXe purification was performed [39]. Radon is one of the main sources of background in the detector, being emitted from many of the detector materials such as stainless steel, and presenting a long decay chain with many contaminants, the ^{214}Pb part of which can appear to be dark matter interaction. Krypton similarly is present as a result of the initial xenon collection process, and also can present as dark matter interaction, so similarly must be removed [30].

These new methods for background reduction have resulted in XENONnT having a low-energy electronic recoil background of (15.8 $\pm$ 1.3) events/(tonne $\times$ year $\times$ keV$_e e$), a factor 5 reduction compared to XENON1T, making it the lowest electronic recoil background level ever achieved in a dark matter experiment [30, 40].

3.2.1 A Brief Note on Signal Classification

Beyond hardware, there are also extensive software protocols and data analysis techniques involved in removing background and ensuring the data collected is as free from contamination as possible. The Data Acquisition (DAQ) system combines open-source and custom-developed software into one large analysis framework called Straxen, processing 200+ GB of raw data for every hour of a collection run [41]. The muon and neutron vetos are integrated into the DAQ framework, and so the events tagged by these are automatically and removed [41]. Straxen is an extremely versatile framework which allows for reprocessing of the data in various ways, according to the needs of individual researchers; all of the data used in this work was processed by it.

When a signal occurs within the TPC, the initial information arrives in the form of PMT pulses: scintillation photons digitised into waveforms by the DAQ system. This pulse data is then processed and reconstructed into higher-level structures, which allows us to distinguish individual signals and events, from single-electrons to large primary S2s. The processing provides data-types encoding key information about these signals, for example whether it is an S1, an S2, or otherwise unidentifiable; these are labelled as type 1, type 2, or type 0 respectively.

Other data-types of interest are the peak area and width. Area is measured in photo-electrons (PEs)

- a unit calibrated as corresponding to the amount of photons hitting the PMTs. The signal width, often given as the data-type of 'range 50% area', is the time-range (in ns) over which the central 50% of the waveform's integrated area is observed, giving the best idea of how long it took the bulk of the signal to be detected, and so is roughly proportional to the depth at which the interaction occurred in the TPC. These two data-types together can offer a clear distinction between different types of signals, which will also be useful in Sec. 6.

More about how this software is used in analysis will be touched on in Sec. 6, or else can be found at Refs. [42, 43].

3.3 The Search for Dark Matter in the Low-Mass Regime

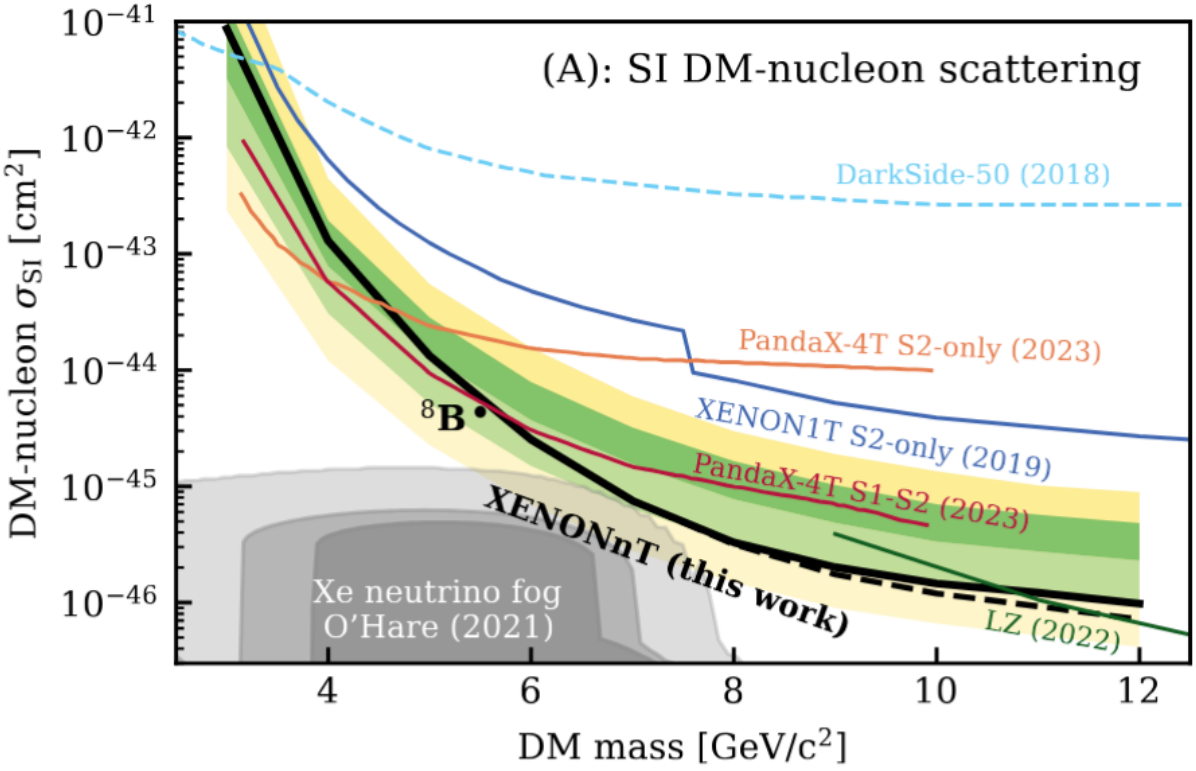


Figure 4: Theoretical WIMP mass vs spin-independent nucleon interaction cross-section, with the limits found by various direct detection experiments, with the 1σ(green) and 2σ sensitivity bands shown. At the bottom of the graph is the neutrino 'fog', in which dark matter interactions would be nearly indistinguishable from those of neutrinos scattering with xenon. From [44].

As the sensitivity of XENONnT improves, the upper limit on WIMP masses continues to lower. Consequently, research has increasingly shifted toward the sub-GeV mass regime - an area still accessible to XENONnT, however one which poses new experimental challenges. Fig. 4 plots some of the latest limits set on the WIMP mass as a function of spin-independent nucleon cross-section,

showing the increasingly stringent limits set on DM in the sub-10 GeV/c^2 mass range, prompting search in even lower mass regions. For dark matter particles in the $\mathcal{O}(100)$GeV/c^2 range, xenon provides a large cross-section of similar mass (in the NR case) with which to scatter and transfer energy, providing clear S1 and S2 signals; this is more difficult for DM particles in the sub-GeV range, as generally they would not be able to transfer enough energy from an NR interaction to create clear and detectable signals.

This does not mean that such particles are completely undetectable. Low-mass DM candidates could still be detected in XENONnT through NR with some difficulty; through ER interactions (though rarer due to a lower interaction cross-section); and through something known as the Migdal effect. This is when a NR scattering, with an exceedingly low probability, causes a small displacement of the nucleus with respect to electrons of the atom, resulting in a small amount of ionisation or excitation of the atom and a subsequent energy deposition [45]. The Migdal effect has not yet been experimentally confirmed, but evidence for it would be extremely promising for the search for low-mass DM, and several investigations into it are ongoing [45–47].

These processes result in much smaller energy depositions compared to standard NR interactions, producing correspondingly smaller S1 and S2 signals. For NR or ER collisions from low-mass DM below $\mathcal{O}$(1 keV), the S1 falls below the energy threshold to be picked up by the PMTs [48], resulting in a signal that we call a 'lone S2', often corresponding to the emission of only a single or few electrons from a xenon atom. This is the signature of would-be low-mass DM candidates, and this area of analysis is known as the 'S2-only' search.

However, the detector is not optimised to detect DM candidates in the low-mass regime, and as such it presents many unique challenges. The lack of an S1 results prevents a calculation of drift time, and thus removes information about interaction depth; similarly lost is the S1/S2 signal ratio patterns used to distinguish between NR and ER interactions in the regular analysis. Despite the work to make XENONnT as shielded from contamination as possible, there remain numerous unpreventable or as-yet unidentified sources of background. Photoionisation of metals in the detector; electron pile-up at the liquid-gas interface; PMT dark-counts; hot-spots along wires in the detector; and many more, all contribute to the background in XENONnT.

Many of the sources listed above contribute to one of the most problematic backgrounds in the experiment, that of single- or few-electron emission. This type of emission is ubiquitous throughout the detector [48, 49]. Single- and few-electron emission has been observed in all XENON-collaboration detectors, as well as other liquid xenon direct detection experiments [9–15, 48–54].

Many attempts to remove, mitigate, and characterise these signals have been attempted; some of these are discussed in Sec. 4. These emissions occur even in long intervals between interactions (on the order of seconds [10]) without clear origin, although a key finding of some of these studies has

been that large S2s seem to be correlated with increased rates in these single- and few-electron signals [51].

This is a phenomenon known as 'delayed electron emission'. Understanding, characterising, and modelling this single-electron background is essential for robust S2-only analyses in the face of lowering WIMP mass-limit. The next section focuses on the ongoing investigation into this single-electron background, and the delayed electron emission component of it in particular; exploring how it has been understood and modelled so far, and how it might better be understood going forward.

4 Delayed Electron Emission

The single- and few-electron background of liquid xenon detectors may be thought of to be made up of, at minimum, two parts: a 'prompt' emission component; and a 'delayed' emission component [53]. Sec. 4.1 describes the phenomenology of these two separate components, before Sec. 4.2 reports on the findings of various studies into specifically the delayed electron emission, and the various theories for its cause. Following that, Sec. 5.1 attempts to evaluate the previous models for the phenomenon; before Sec. 5.2 introduces a new method to serve as a basis for modelling and thinking about this delayed electron emission, sometimes referred to as 'electron trains', into something that might provide a more full background characterisation for them.

4.1 Phenomenology of Delayed Electron Emission

This section will review the relevant physics happening in the detector after a particle collision, partially described in Sec. 3.1; now in the context of understanding the single-electron background, and the potential causes of delayed electron emission in particular; before outlining the hypothesised origins.

Following a scattering event in the TPC, an ionisation track of electrons is produced, which then drifts upward from the electric field to the surface of the LXe. These electrons are extracted almost instantaneously once there, and form the initial S2 signal. For clarity's sake in later discussion, these initial signals are henceforth referred to as the "primary S2s" (pS2s) since an 'S2' may technically be any signal extracted and detected in the GXe. We will thus also here introduce the definition that any S2 which is not a pS2 is thus called a "small-S2".

Within $\sim$1 'full drift time' (fdt) =2.3 ms of an S1 or pS2, another smaller surge of electrons is observed, attributed to the prompt scintillation light of the S1 or pS2 photoionising with exposed metal surfaces and electronegative impurities in the detector [10, 51]. Electrons in an ionised drift cloud can also thermalise at the LXe surface and get trapped at the barrier, some of these may be stimulated and extracted at this time also; this is talked about more in Sec. 4.2.1. This forms the 'prompt' part of the single-electron background observed after a pS2, and hence these signals are referred to as 'prompt electrons'. These electrons may take 3-5 fdt after a pS2 is observed to be fully extracted, as higher-order photoionisation can occur, however the exact number is up to the preference of the analyst. In this work the cut-off for prompt electrons is taken to be 5 fdt. Sec. 6.2.2 develops more on why this boundary was chosen.

However, small-S2 signals are still observed many full drift times after a pS2, even well beyond the maximum electron lifetime in the detector (up to $\approx$15 ms depending on xenon purity [30]). These

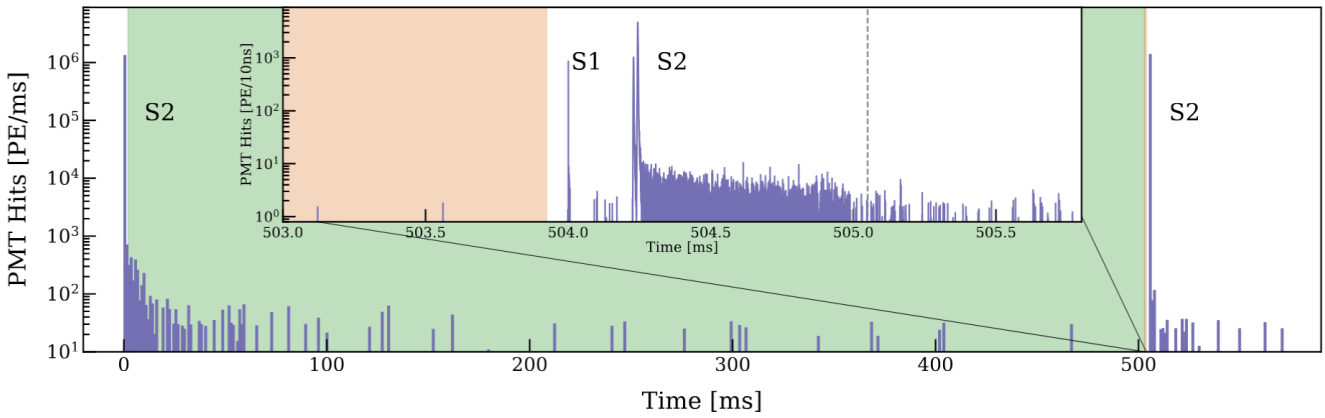


Figure 5: **Main**: An illustrative waveform showing the number and time distribution of small-S2 signals following a pS2 signal. **Inset**: A close-up view of an S1 followed by a pS2. Prompt electrons can be seen immediately following the pS2, from the known background of photoionisation with metal surfaces or electronegative impurities. The dotted line indicates the one full drift time, beyond which we expect the single- or few-electron rate to drop off. Note that this data comes from XENON1T, with a smaller TPC than XENONnT, meaning the maximum drift time is only 750 μs, not 2.3 ms. From [11].

comprise the 'delayed' component, the phenomenon of 'delayed electron trains'. Relative to prompt electrons, this background is far less well-understood. An example waveform of the single-electron background and the two distinct components is shown in Fig. 5.

Also observed in Fig. 5 is the characteristic inverse power law shape of delayed electron trains, i.e. the form $R_e \sim \Delta t^{-n}$, where R_e is the rate of electron emission, and Δt is the time between the pS2 and some arbitrary time afterwards. This shape was again noted in other studies [9–11]. Unlike exponential or Poisson processes which govern other detector phenomena, there are few known processes occurring in the detector which would produce such a long-tailed, temporal distribution, although it has been reported that processes such as electron-ion recombination in liquid xenon [55], and bi- exciton quenching in liquid scintillators [56] can produce power law-like time dependencies. However, the relevance of these processes within XENONnT, and within the context of the delayed emission hypotheses following, is unclear.

This observed power-law model also means that there is no characteristic 'cut-off' time for the emission of delayed electrons after a pS2 has occurred, meaning they can occur on the order of seconds after. This poses a big challenge for the low-mass dark matter search, as even regions far from pS2s in time are not intrinsically 'safe' from the possibility of delayed electron emission. Sec. 5.1 looks more at the nature and modelling of this inverse power law shape, and how it affects the search for light DM.

Two main hypotheses have been proposed for the source of delayed electron emission, though they

are not mutually exclusive, nor do they necessarily rule out other unnamed or unknown theories:

1. **Trapping at liquid-gas interface:** Electrons in the drift cloud following an interaction do not all get extracted into a pS2 signal: some thermalise and get trapped at the interface. Later interactions, as well as fluctuations in detector conditions, can cause them to release at later times, either as part of the prompt or delayed electron emission, and perhaps even uncorrelated to interactions happening in the detector.
2. **Electronegative impurities:** Molecules such as O_2 and H_2O present in the LXe can trap drifting electrons after an interaction, releasing them in an ostensibly stochastic manner up to seconds after the most recent pS2.

While these two processes are known to happen within the detector and have themselves been thoroughly researched, their role in producing delayed electron trains and the factors determining their rates is far less clear. The next section looks more at these two theories; examining the evidence for and against them being the cause of delayed electron trains; and whether they may both be contributing to the phenomenon of delayed electron emission.

4.2 Origin Hypotheses

The presence of a delayed emission background in direct detection detectors has been known about for decades, likely a part of even the earliest liquid xenon and argon detectors, with interest picking up more as the search for light dark matter intensified [12]. One of the earliest detailed studies is from XENON10, where the event trigger was sensitive to a single electron, allowing for a more detailed detection of this background for one of the first times, and where these signals were first referred to as 'electron trains' [50].

Following studies continued to attempt to characterise these emissions and correlate them to detector conditions. A study in XENON100 observed a steep decline in the single- and few-electron populations after the maximum drift time, indicating a correlation between primary S2s and prompt electrons; though also noting that this correlation continues even for signals beyond the maximum drift time [51]. This study also shows that the relative rate of small-S2s is proportional to the size of the most recent prior pS2, indicating this emission is not likely to be from external sources [51].

As a result of this finding, the two main hypotheses on the origin of delayed electron trains revolve around some belief that these trains are roughly correlated, both in time and position, with pS2 signals in the detector. Sec. 5.1 will look at how this aspect of a correlation in time forms a key basis for the initial modelling of delayed electron trains.

However, there has also been investigation into other sources of single-electron emission in noble liquid TPCs [54, 57, 58], suggesting that there are sources of emission, uncorrelated in both time and position to interactions within the TPC. This is an important consideration as light dark matter scattering within the detector would appear as single- and few-electron emission seemingly uncorrelated with interactions happening in the detector, as this itself would be an interaction. Modelling delayed electron emission accurately, as electrons which are causally correlated to the detection of pS2s, is thus a crucial step to isolating this uncorrelated single- and few-electron background. Further to that, there are also methods of helping to remove these known uncorrelated backgrounds mentioned above; Sec. 7.3.2 will return to these studies in justifying why fiducialisation is a necessary step for modelling.

For now, we will work on understanding delayed electron emission, and work under the assumption that the majority of delayed single- and few-electron signals must come from one of the two following sources.

4.2.1 Trapping at Liquid-Gas Interface

The first potential cause of delayed electron emission is electrons that rise to the liquid-gas interface in or around the initial cloud caused by some given interaction, and thermalise just below the LXe surface, becoming trapped until they are eventually extracted many orders of a drift time later. The extraction efficiency of the detector is not unity [11], making this a known process in the detector, for which a more complete treatment can be found at Ref. [59].

One theoretical study by Sorensen [60] looked at the nature of the liquid-gas interface in xenon detectors, in order to find the energy needed for a drifting electron to overcome the barrier and be extracted into the gas. From this he calculated the rate at which electrons can become trapped at this barrier; and thus found the lifetime of these thermalised, trapped electrons, finding it to be "compatible with the original observations of XENON10" [60]. This paper does note however that this is a purely theoretical calculation, missing real-world influences of the detector conditions, such as some surface roughness to the LXe, or non-uniformity to the electric extraction field. Sorensen does offer potential ways to mitigate this background in experiments, for example by increasing the extraction efficiency with stronger extraction fields, or by stimulating electron emission with LED photons [60]. Akerib et al. [10] find physical backing for this process, hypothesising that the liquid surface gets deformed where electrons are trapped, containing them in a small local bubble. This would also have the effect of shielding them from exposure to impurities in the LXe bulk, explaining how they could survive there many orders of magnitude longer than the electron lifetime before being extracted as delayed electrons [10].

Later studies, such as from Kopec et al. [52], also find evidence in support of the trapping hypothesis, by further investigating, among other things, the link between the electron extraction efficiency and the rate of delayed electron emission. This is an in-depth study to the nature of the electron train and its characterisation, the methodology of which will be returned to in Sec. 5.1. This study finds that, modelling the power law rate as $R = A \cdot t^b$, increasing the electron extraction efficiency only increases the amplitude A of delayed electron emission - the amount of signals making up each delayed electron train - while having no effect on the rate b, with Refs. [51, 53] also finding similar linear correlations. This finding is shown on the left of Fig. 6.

Kopec et al. also investigate the effect of the drift field strength, finding that an increase led to a larger measured pS2. This is because there is a finite electron lifetime in the detector, 3 μs for Ref. [52], so an increased drift field results in more electrons being extracted within this lifetime. These drifting electrons also have a higher velocity upon reaching the liquid-gas barrier, and therefore have a higher emission probability [59].

This finding introduces challenges for the hypothesis of electronegative impurities, explored more in the next section. Increasing drift speed should mean that electrons have less chance of attaching to impurities before being extracted [52, 57]. If the impurity hypothesis is the majority contributor to electron trains, we should thus expect to see a decrease in the amount of delayed electrons relative to the total amount of electrons in its related pS2, i.e. $f = \frac{N_{delayed}}{N_{pS2}}$ should decrease: fewer electrons get trapped by impurities in the drift column after an interaction, resulting in larger measured pS2 signals and smaller measured delayed electron trains.

However, Kopec et al. found the reverse to be true: when measuring N_{pS2} as the 'lifetime corrected' pS2 size (an estimated value of the produced charge from an interaction, a necessary calculation when the maximum drift time of an electron in the TPC is longer than the electron lifetime), they found the fraction f increased with drift field, primarily for the single-electron population (right plot on Fig. 6). This suggests the opposite to the impurities hypothesis stated above: delayed electron trains get larger with increasing drift speed.

Conversely, when removing the lifetime correction and measuring N_{pS2} as only the measured pS2 size (i.e. the amount of electrons being detected by the PMTs), the fraction f was found to decrease with increasing drift speed [52, 57]. This indicates that increasing drift field may have no effect on impurities and their chances of capturing electrons after an interaction; however, it does increase the energy of electrons reaching the liquid-gas barrier. This means that proportionally more are extracted as part of the initial pS2 signal, rather than as delayed electrons later - suggesting a surface effect dominating the rate of delayed electrons, rather than one happening in the LXe bulk.

This is not the full empirical story, however, with Sorensen and Kamdin [53] finding that only the earliest parts of an electron train may be accounted for by trapped, thermalised electrons, favouring

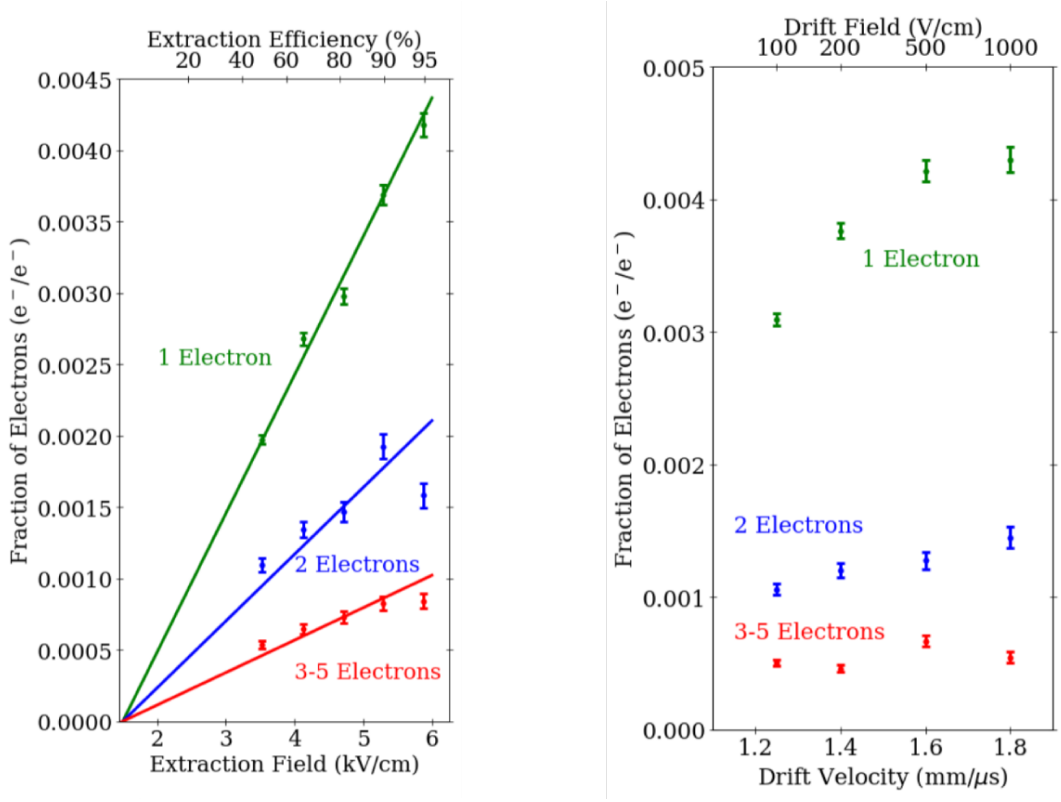


Figure 6: The effect of increasing extraction field strength (**Left**) and increasing drift velocity and field (**Right**) on the fraction of electrons observed in the electron train vs the total produced in the primary interaction (i.e. lifetime-corrected size), for varying electron populations, from single-electron emission ($1e^-$) to multiple-electron emission ($2e^-$, 3-$5e^-$. Each of these levels is identified as a 'multiplicity', a concept which will be important in later sections. From [52].

instead the idea of fluorescence photons photo-ionising with electronegative impurities in the LXe for the more delayed parts of the electron train. This study finds that increasing the strength of the electric field decreases the number of prompt electrons immediately following a pS2 [53]. However, it only linearly increases the rate of delayed electrons, suggesting that the later component of an electron train is related to the size of a pS2, with a reduction only to be seen in purifying the xenon - leading to a conclusion of an origin from electronegative impurities getting ionised [53]. Furthermore, studies from XENON1T [11] and LUX [10] also find more support for the impurities hypothesis, as shall be shown.

4.2.2 Electronegative Impurities in LXe

This second hypothesis comes from the fact that the LXe is not 100% pure, and therefore contains many possible impurities which could capture electrons [30]. As mentioned in Sec. 3.2, there are systems in place to mitigate these impurities and the backgrounds they can cause, however the process is not perfect. Under this hypothesis, after a primary interaction, electrons in the drift cloud have a chance of colliding with an impurity like oxygen. The mechanisms by which these electrons are later released is not fully understood, but photoionisation and direct electron detachment (either through collision or some tunnelling process) are suggested as the two most likely causes [53]. How frequently these processes happen, and whether they would be sufficient to account for the observed rate of delayed electron emission, is less certain.

A XENON1T study from Aprile et al. [11] finds that, as the drift time increases, so too does the fraction of delayed electrons. That is to say, as the drift field and speed increases, the fraction of delayed electrons relative to the size of the lifetime-corrected pS2 decreases, running directly contrary to the findings of Kopec et al. This study observes clear positive trends in the data for all three position-correlated populations [11], shown in Fig. 7. The nature of position-correlation in the context of delayed electron emission is explored more in Sec. 7.3.2; Aprile et al. simply defines a delayed electron as position correlated if it occurs within 15cm of its most recent preceding pS2. Note that Fig. 7 is a measure of drift time, rather than drift velocity as in Fig. 6. This means that it is therefore also a function of interaction depth; however Kopec et al. also investigate this function (Fig. 7 in Ref. [52], finding a decrease for increasing drift time, still contrary to Aprile et al.

This finding from XENON1T supports the hypothesis that impurities in the LXe are trapping electrons after a primary interaction; this is backed up by the conclusion that the electron extraction efficiency has minimal effect on the intensity of delayed electron emission relative to the effect of drift times (though it may still be a secondary effect) [11]. Support for this hypothesis may also be found in Refs. [10, 12, 15, 51]

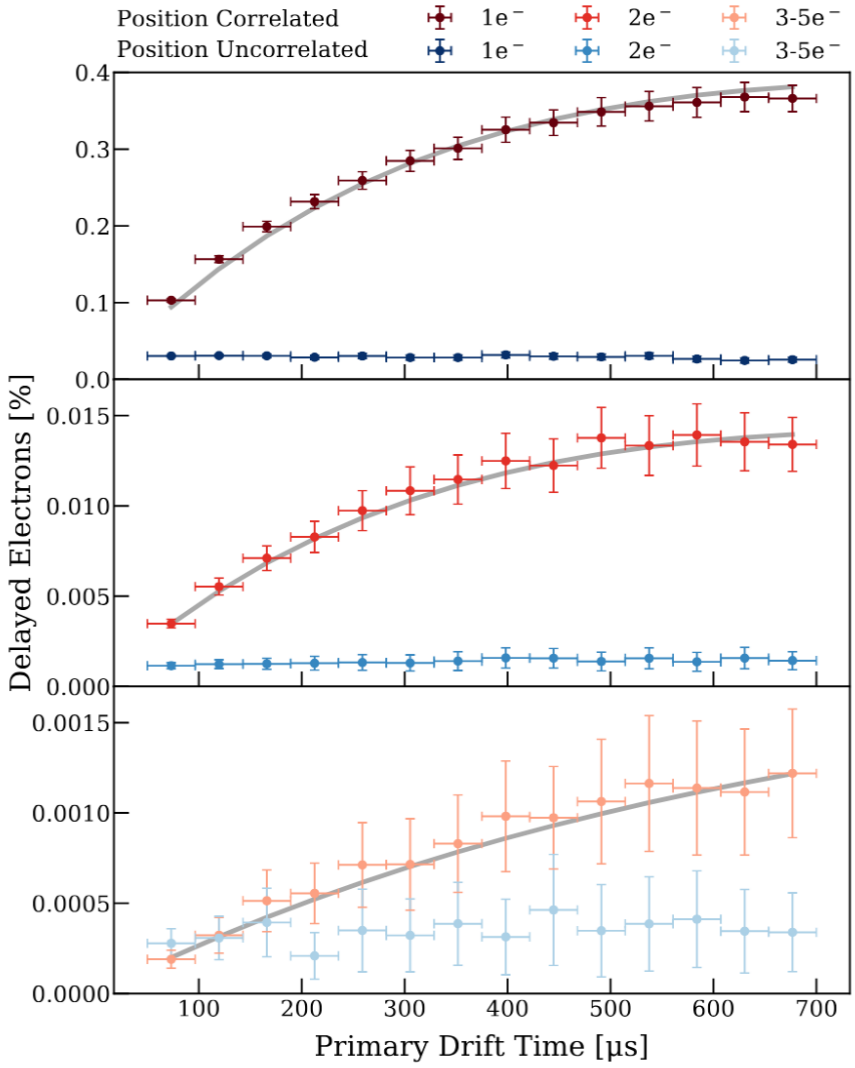


Figure 7: The intensity of delayed electron emissions vs drift time between 2-200ms after a pS2 for $1e^-$ (**Top**), $2e^-$ (**Middle**), and 3-$5e^-$ (**Bottom**) populations. The grey line is the expected number of delayed electrons, assuming the production rate is constant along the drift column of an interaction, and proportional to the number of electrons captured on impurities. Here, position correlation is defined as delayed electrons which occur in a 15 cm radius around their most recent prior pS2, whereas position uncorrelated signals occur more than 20 cm away. From [11].

However, once again not all supports this hypothesis, or at least not as it immediately seems. If electronegative impurities are capturing electrons, they would only be able to capture and release one electron at a time [52]. This may be seen with, for example, O_2, which is a common impurity in the detector. Due to the layout of its electron shells, it is energetically favourable for O_2 to take on to take on an electron and become O_2^-. However, taking on a second electron would require overcoming the repulsive Coulomb force, requiring energy to be put in for it to become O_2^{2-}. This makes the possibility of impurities taking on more than one electron at a time virtually nil.

This is contrary to observation, however, as delayed electron emission is frequently observed to be composed of multiple-electron emission, as well as single-electron emission. Some of this may be 'pile-up' of single electrons, where two or more signals arrive so close together that the software processor clumps them together as one larger signal, however, not to the rate at which this multiple-electron emission is observed [11, 52]. This is difficult to explain if electronegative impurities are the main cause of delayed electron trains.

Kopec et. al also attempted to investigate electronegative impurities by using 1 W of infrared light (1550 nm), a wavelength to which the PMTs are blind, to try and stimulate electron emission [52]. If electronegative impurities such as O_2 were trapping electrons after an interaction, this method should stimulate their release, and an increase in position-correlated electrons would be observed. This is not seen to be the case, instead with only an increase in position-uncorrelated electrons observed.

Furthermore, while Aprile et al. do find evidence for the impurities hypothesis, such as a reduction in electron train intensity as the LXe purity (measured by electron lifetime) improves, they also observe a decrease in intensity over periods in which the purity remains constant [11]. This means that whatever impurity is capturing these electrons, its not something that can be measured by the electron lifetime (like O_2 can), and which can trap electrons for far longer timescales than for typical impurities. In fact, the fraction of electrons observed as delayed electron emission accounts for only 1% of all drifting electrons after an interaction, meaning the majority of captured electrons are never extracted, or at least never measured [11]. The alternative is thus some kind of neutral impurity, that could for example instead capture a scintillation or electroluminescence photon, exciting the molecule and causing to ionise at some unspecified time later.

Electron trains are not greatly understood, with much conflicting evidence in what the true cause of them may be. While this phenomenological work is ongoing, many are also trying to approach this from another angle: modelling the rate of electron trains. If an accurate model of these emissions can be found, its effect is twofold: 1) Understanding of the characteristic rate and timescale of these emissions may allow us to better pin down the processes by which they happen; and 2) we can incorporate it into our background model to better search for light dark matter. This is the main focus of the rest of this work.

5 Modelling

5.1 Conventional Modelling of Delayed Election Emission

As mentioned previously, modelling the electron train phenomenon results in some form of inverse power law in order to best fit the data, with Ref. [9] among the first to note this feature. This section will describe the general methodology by which we arrive at this fit, and look at how we can begin to characterise electron trains. This section draws from the work done by Refs. [9–11, 52], as well as the bachelor's thesis of K. Umesh [1], which will be looked at further in Sec. 5.2. This section will only focus on what is done with the data to get a power law model; some simple selection is explained in the case of this initial modelling, however Sec. 6 examines how this data may actually get selected for the newer model.

To start the analysis, there are some basic cuts and orderings done to the data to get the relevant electron populations:

1. Select all S2-type signals $\leq$ 150 PE. This is an arbitrary cut-off, but since we are interested in single- and few-electron signals, this gives up to the $5e^-$ multiplicity population. These are the 'small-S2s' defined at the start of Sec. 4.1. The definition of S2-type signals is as was given in Sec. 3.2.1. Primary S2s are thus defined as any S2-type signals above 150 PE.

2. Sort the small-S2 signals by time, and assign each to the most recent pS2 prior to it. This is done on the basis of the time-correlation findings mentioned earlier, as well as knowledge of the maximum drift time (2.3 ms) being much shorter than the average time between interactions ($\sim$ 200 ms), making for a reasonable basis that electrons are probably detected close to the pS2s which caused them.

3. Remove from the dataset all signals within one full drift time of the pS2 signal. This is done to cut out prompt electrons, described earlier. We make the cut at 3-5 fdt to fully avoid the background of second-order photoionisation. Since the order of this cut is 20x or more smaller than the average fit time ($\sim$ 250 ms), this can be done without any real loss to the delayed electron emission model. The prominence of these prompt electrons in the dataset and thus the need to remove them, even to 5 fdt (11.5 ms), can be seen in Fig. 8.

With these cuts done, we can next begin to start modelling the power law, by using a maximum likelihood cost function to fit a function of the form $R(t) = A \cdot \Delta t^n$. This thesis uses the extended unbinned version from the CERN ROOT team's iminuit python package throughout. For more information on this package, see Ref. [61].

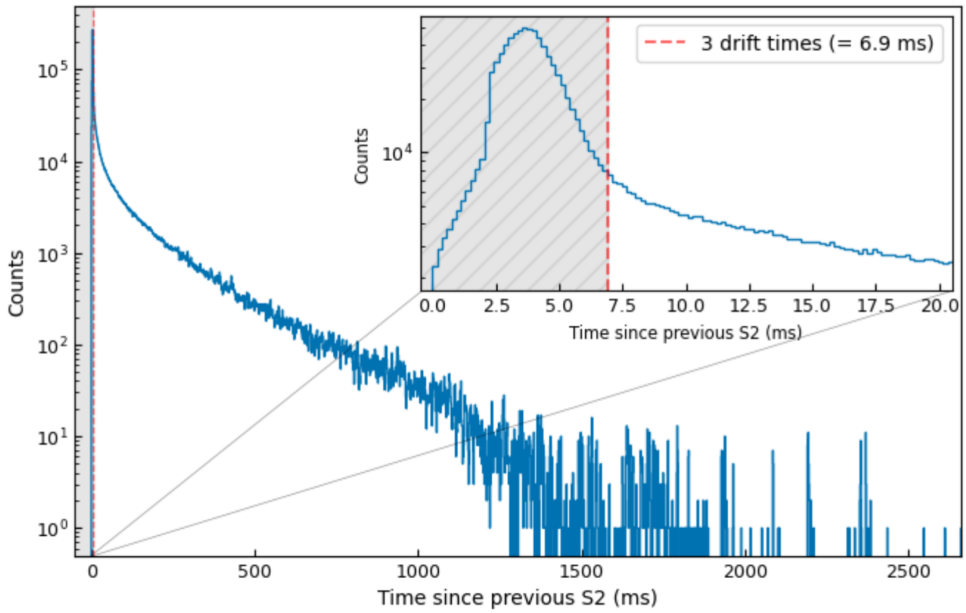


Figure 8: **Main**: A histogram of all small-S2 signals in a 3 hr Science Run 0 (SR0) run as a function of time since the most recent previous pS2. **Inset**: Close-up of the region immediately following a pS2, showing the large electron count caused by prompt electrons. The gray shaded region up to the red dotted line shows the 3 fdt boundary cut applied to the data. Figure from [1].

The simplest fit that can be done is over that of the histogram of delayed electron signals vs time since the previous pS2. On logarithmic axes, this looks like a straight line, as shown in Fig. 9. The reading of this model, and future ones also, is such that taking its rate at a given point in time gives the expected rate of electrons per ms. In this case that rate, $R(t)$ is a function of time since previous pS2, Δt; importantly, this will change later when looking at the model in Sec. 5.2. A power law fit is better than an exponential [9, 52] (what we might expect if delayed electrons were the result of random processes in the detector, unrelated to pS2 signals), and as shown before, there is clear physical basis for fitting the data like this. The fit in Fig. 9 returns a power law exponent of $n = 1.62$, which is slightly high compared to values found from XENON1T (1.1-1.4) [11], LUX (1.0-1.1) [10], ASTERiX (1.2 for $1e^-$ population) [52], and RED-1 (1.4) [9]. However, this is likely a reflection of the decision to cut prompt electrons at 5 fdt rather than the 1-3 fdt of these previous studies.

However, this fit is not perfect, particularly at longer timescales. The reasons for this could be that an uncorrelated electron population begins to dominate at this timescale, or relatedly because delayed electron emission at longer timescales may be caused by different sources. More importantly, the average length between pS2s is only ~200 ms, meaning attempting power law fits at longer timescales get interrupted by the next pS2, thereby causing these longer regions to have fewer data points in general according to the method of assigning each delayed electron to its 'most recent' prior pS2. These poorer statistics allow for more skew in the data at these longer time scales, which can be seen in Fig. 9, where beyond ~500 ms there are fewer than 10 counts for each bin, for over 6 total hours

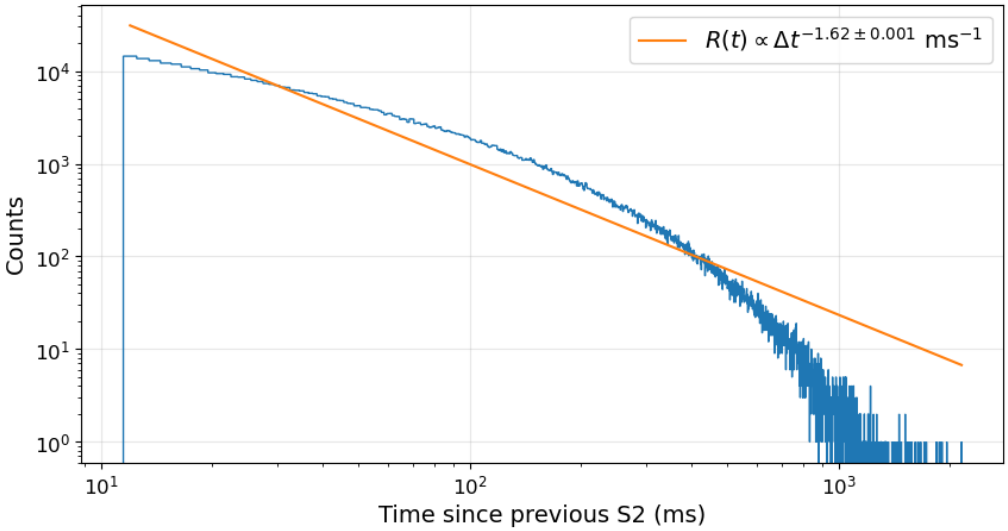


Figure 9: A basic, normalised power law fit to the histogram of single- and few- delayed electron signals as a function of time since the most recent previous pS2. This data comes from 13 30-minute SR0 background runs, used throughout this work. Sec. 6 explains the selection of these runs further. As mentioned before, the prompt electrons are cut from this dataset before fitting. Figure adapted from that in Ref [1].

of data. To counteract this, prior studies sometimes placed limitations on the fitting windows, for example Ref. [11] analyses only 2-200 ms after each pS2, which has the effect of reducing the power law coefficient. Alternatively you can adjust the fit for efficiency, as is done in Ref. [9].

There is however another point to note: within the detector, double-scatters (where a particle interacts again, resulting in a second large S2, usually within 1 fdt) make up 35-40% of all pS2 signals. This means that some delayed electrons caused by the first pS2, but released and detected at some point beyond 3-5 fdt after the second pS2, would become associated with the second pS2, and not the first. We could account for this by combining pS2s within 1 fdt of one another, taking their median time weighted by size; this also corrects for processor errors in peak-splitting. Kopec et al. resolve this by only accepting pS2s of a size "larger than 10% of the previous primary S2's size reduced exponentially according to a time constant of 1 ms" [52]; Akerib et al. at LUX resolve this by only accepting pS2s with no other activity 100 ms before it [10].

However, this limits available statistics and does not offer data guaranteed to be clean of contamination from prior pS2s, it only pushes the problem further along. It was noted previously that delayed electron emission is observed to be on the order of seconds, meaning that any electron train after two given interactions even hundreds of milliseconds apart, has a possibility of being contaminated by the electron train of the first interaction. The nature of a power law model means that soon after a pS2, a delayed electron has a high likelihood of being caused by that pS2; however, further on this likelihood evens out with the tails of prior pS2s. Fig. 19 shows a toy example of what this may look

like in the context of a double scatter. Note that at e.g. 120 ms in the graph, the contribution to the electron rate is near equal from each pS2. This is to say, it is not possible to accurately model delayed electron train emission under the assumption that each delayed electron signal originated from its most recent prior pS2.

Noting these problems, Ref. [1] suggests a slightly different approach to assessing the rate model of electron trains. In this new model, the rate is given as such that each pS2 is expected to generate a certain amount of electrons as part of its delayed electron train. Each train can thus be modelled with a power law in accordance with the real data available. However, each power law is allowed to 'bleed' into one another, such that overlaps are accounted for, rather than cut out, in a cumulative-type model. The nature of the model is set up such that although the most recent prior pS2 will constitute the greatest contribution to the total electron rate, every prior pS2 will also contribute, each to some diminishing extent. The exact form of this model is motivated and explained in the next section.

5.2 A Novel Technique for Modelling Delayed Electron Emission

As has been shown, the prior conventional model has a few problems, chief among them that of electron train overlap, and lack of statistics at longer timescales. Instead, this new model, which is first developed in Ref [1] and then developed throughout this work, allows for the overlap of each electron train throughout the model, with some additional prefactors to normalise and scale. The exact definitions and meanings of these will be covered in due course.

In this model, the contribution of each individual pS2 to the total electron rate at some given time is given by:

$$R_p(t,\, t_p,\, A_p,\, W_p \mid n,\, t_{min},\, \alpha,\, \beta) = \begin{cases} 0 & t \leq t_p + t_{min} \\ A_p^{\alpha} W_p^{\beta} (t - t_p)^{-n}, & \text{otherwise} \end{cases} \tag{1}$$

where t_p is the time the peak occurs (in ms since the start of the run); A_p and W_p are the area (in PE) and width (in ns), with exponents of α and β as scaling parameters. t_{min} is the cut-off time for prompt electrons after a pS2, set in Ref. [1] as 3 fdt, or 6.9 ms for XENONnT. Again, it will be set to 5 fdt in this work. We will use throughout this work the idea of pS2s 'contributing' or 'creating' delayed electrons, but as a reminder, delayed electrons are actually created as a function of electrons being lost in some manner from the ionised cloud created following an interaction in the detector, and then later released. Primary S2s are formed from the majority of this ionised electron cloud

being extracted at once in the first instance; they therefore do not create delayed electron trains. However their properties such as area and width correlate to the properties of the electron cloud, and therefore to the number of delayed electrons which may have been lost from it in one way or another.

Using this model we may assess the total background electron rate at any point in the run by taking the sum of all prior pS2 contributions:

$$R(t \mid k,\, C, ...) = k + C \cdot \sum_{p}^{P} R_p(t,\, t_p,\, A_p,\, W_p \mid n,\, t_{min},\, \alpha,\, \beta) \tag{2}$$

where P is all pS2s prior to t and C is a scaling constant. When put through the likelihood function, k is the value returned as the background rate of electrons we expect without any electron train interference. Lowering this value is how we may lower the limit for low-mass dark matter candidates. From this equation one can see that if we looked at the total electron rate at a point e.g. 10 ms after a large pS2, $R_{p,\,i}$ would give the largest contribution to $R(t)$, with $R_{p,\,i-1}$ a proportionally smaller one, and so on.

Note here how the rate equation, $R(t)$, has changed from being as a function of time since the previous pS2, Δt, to being a function of time since some arbitrary start point in our analysis. This may be, for example, the start of a 'run', the allotments of data collected in some given amount of time by the experiment. More about these runs is talked about in Sec. 6. By redefining the rate function in this way, we shift from looking for light dark matter in the windows between pS2s, to looking for it anywhere. Before, we could only ask whether an excess of single- or few-electrons was expected within a rate model as a function of where that excess occurred relative in time to some pS2. However, light dark matter would be, as mentioned, uncorrelated to the emission and detection of other pS2s in the detector, so this linking of the two is not well motivated physically. With this new model, we can look at the single- and few-electron rate almost anywhere in the time range of our data, and ask whether that rate should be expected according to the amount of other interactions (measured by pS2s and their properties) which have happened in the detector before that point. In this way the model may offer a new angle of insight into the search for light dark matter not previously possible.

If we break apart Eq. 1, we find that the value $C \cdot A_p^{\alpha} W_p^{\beta}$ is a predictive calculation of the total amount of single electrons each pS2 'produces', and therefore contains in its respective electron train. If this number is accurate, it would go a long way to better modelling this background electron rate. Sec. 7 will look more at adding to this prefactor, but for now it is a good starting point. The

exponents of α and β give the model much flexibility also, as if for example the observed electron trains had no dependence on pS2 width, the cost function would naturally put β to 0. This is not likely to be the case though, since pS2 area has already been shown to be correlated to the electron train size, and pS2 width (i.e. how long the signal was observed for) has a rough correlation to drift time and interaction depth, and can be an passable surrogate for an S1 in S2-only searches [48, 62], making it also highly likely to be a correlated property of delayed electron emission.

When plotting to ∼300 s of data, corresponding to approximately 2000 pS2s, Ref. [1] finds that this initial guess does seem to match the pS2 and electron train distribution, at least somewhat. This plot looks very different from previous ones shown: instead of finding a power law fit across a "time since pS2" histogram, we are now looking at a visualisation of the electron rates across a larger span of time, a period of many seconds, with the best fit looking like one where the plotted line best lines up with the histogram counts at every point and for every peak. A snapshot of this plotting is shown in Fig. 10. The model seems to have good agreement with many features of the dataset, different pS2s from large (at ∼900 ms); small (∼400 ms); and clustered (∼1400 ms). This whole region of data produces a power law exponent value of $n = 1.448 \pm 0.005$, which is relatively higher than the previous studies mentioned, however, given this is a method totally apart from previous studies in reasoning and procedure, there is no real reason to believe that this value is any better or worse for being higher than the others.

Finally, Ref. [1] finds a background constant rate of $k = 36.5 \pm 0.6$ electrons/second for this model. This is three times lower than expected from just fitting a uniform distribution across the delayed electron data [1], indicating that, as has been shown already, electron trains have some relation to the pS2s prior to them, and providing a clear physical basis for the model. This is not a comparable feature of the basic power law models, as the nature and statistical limitations of those ones inhibits the addition of some background constant, however it is a value that can be returned to later as this model improves, in order to see how the limit may be lowered further.

However, there are also problems with this model. Fig. 10 shows areas where the rate is not being picked up, such as a large spike at 600 ms. This likely indicates a problem with the data selection, for example where a pS2 has not been selected but its electrons have. Sec. 6 looks at how we can address this issue. Then, a more subtle problem with the model is that although we cut out the prompt electrons after every pS2, the probability density function of the model still predicts the "tails" of each previous electron train to pass through these regions, and thus predicts a non-zero amount of electrons within them, something particularly prominent around double- and multiple-scatter signals. Sec. 7.1 looks at how this can be fixed in the model.

This new model addresses key concerns with how prior models approached the physical reality of electron trains, and offers a new and unique way of interpreting the single-electron emission data,

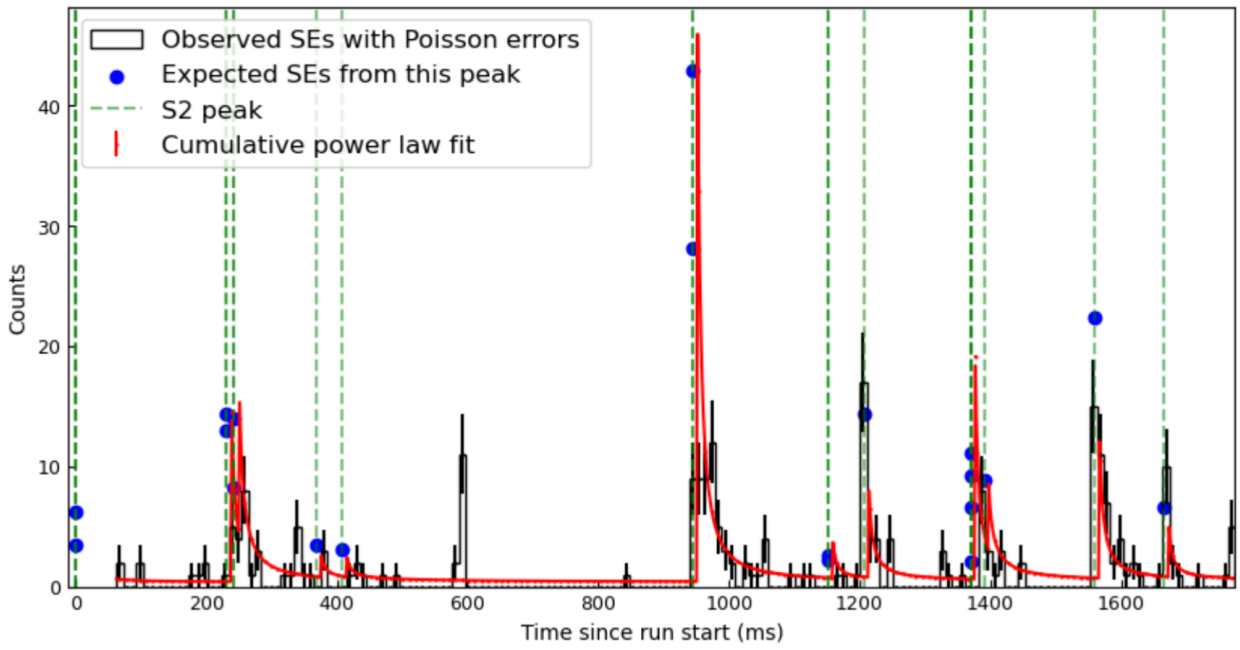


Figure 10: A visualisation of the cumulative power law fit over 10 or so pS2 peaks. The blue points are the value $C \cdot A_p^\alpha W_p^\beta$, the expected number of SEs from the peak, which we can see have relatively good agreement with the heights of both the observed SE values and the cumulative power law fit. For this plot the resolution is 10 ms, and the error bars are the square root of the counts of each bin, with the fit passing through the majority of them. From [1].

worth exploring. The rest of this work examines how the model can be improved and expanded upon in order to best analyse and characterise the phenomenon of delayed electron emission trains.

6 Selection of Data

Before beginning to assess the quality of this model for characterising the phenomenon of delayed electron trains, a thorough and comprehensive methodology for selecting signals must be developed, in order to best capture the signals which we may consider to be pS2s and SEs, and furthermore to distinguish between them.

The desired result for any selection method is to create a cut or filter within the data such that only the relevant signals are retained, while minimising and removing all signals that do not pertain to the modelling of delayed electron trains. A perfect selection method would thereby include all primary S2s, along with both SE signals from the e-trains and low-mass DM signals that mimic these signals, while removing all irrelevant signals, namely that of prompt electrons, mentioned in Sec. 5.1.

Most important is that the data selection is physically justified, both in which signals it chooses to include, and which it chooses to cut, from the dataset. The selection method used by Ref. [1] is explained in Sec. 6.1, setting the baseline by which the final selection methodology is measured. Beyond the work of Umesh, there is also motivation from other prior studies for a more developed selection methodology. As outlined in Sec. 5.1, Refs. [11, 52] define pS2s as those with an area >150 PE, and small-S2s as those below that threshold. Sec. 6.2 describes a methodology developed in this work; Sec. 6.3 shows comparisons between these two methodologies, and the changes made.

Within the S2-only search ongoing within XENON, there are a variety of established cuts which are performed on both primary S2s and the single-electron populations. However, since most of these cuts are performed with the model of Sec. 5.1 in mind, we did not want to take any of these cuts for granted when obtaining the data for this new selection, and instead only implemented cuts if it was certain that they would not affect the integrity of the data later being passed to the model.

We can here begin to utilise the data-types provided by Straxen, mentioned in Sec. 3.2.1, namely that of area and width. Fig. 11 provides a map of the total data landscape available to the selection; every signal of each type from a 30-minute SR0 run is displayed here. Labelled are the core signals of interest: those of pS2s, the large, electron-train-generating signals; the single- and few-electron populations; and a variety of signals we wish to cut, which will be discussed in due course. It is by this map that we may illustrate the data selection going forward.

The data for this work was selected and tested over 13 30-minute-long runs from Science Run 0 (SR0), which ran from May to December 2021. These runs contained no additional sources for testing or calibration, meaning they should assist in making a model for the delayed electron train background which would be applicable in the wider search for light dark matter. A full list of the runs used can be found in Appendix B; for viewing clarity only data from the first of those runs was used in the

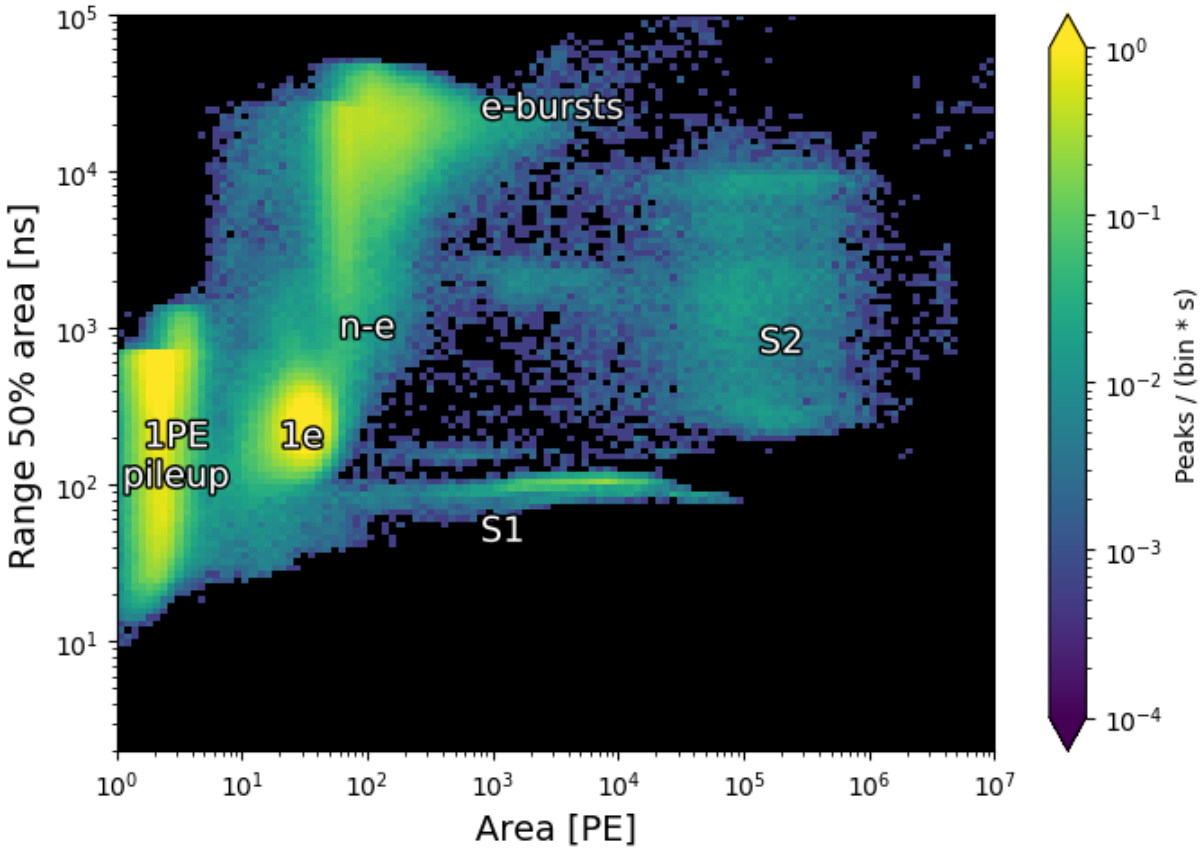


Figure 11: A 2D histogram of area (in PE) vs range 50% area (in ns, analogous to signal width) for all signals in a ∼30 minute SR0 run. This is made using a Straxen mini-analysis, details of which can be found in [42]. The labels act as guidelines for the typical regions in which each signal type is expected to be found. The 1PE pileup region is composed of type-0 signals. These occur from single photon emission in the detector and their pile-up when they hit the PMTs, causing signals too small to be even a single electron. The other labels will be discussed in the course of this section.

plots in this section.

6.1 Previous Data Selection Method

It is from Fig. 11 that a selection like that in Ref. [1] is made. Using the labels as an aid, one can draw boundaries of the regions where we might expect to find the key populations of interest for the model; that of $1e^-$, ne^-, and pS2 signals. We thus select only the S2-type signals in each given region. This was not done in Ref. [1], however it is a simple fix made here for the sake of clarity. Without this cut, the $1e^-$ and ne^- boxes end up with just under 1% of their signals consisting of either S1s or junk (i.e. type 0) - a small but noticeable difference in data quality. The boundary limits were selected based on where these desired populations seem to lie on the graph, arbitrary except to try and maximise the dataset without clear contamination of overlap onto other labels. The visualisation of how these regions were selected is shown in Fig. 12.

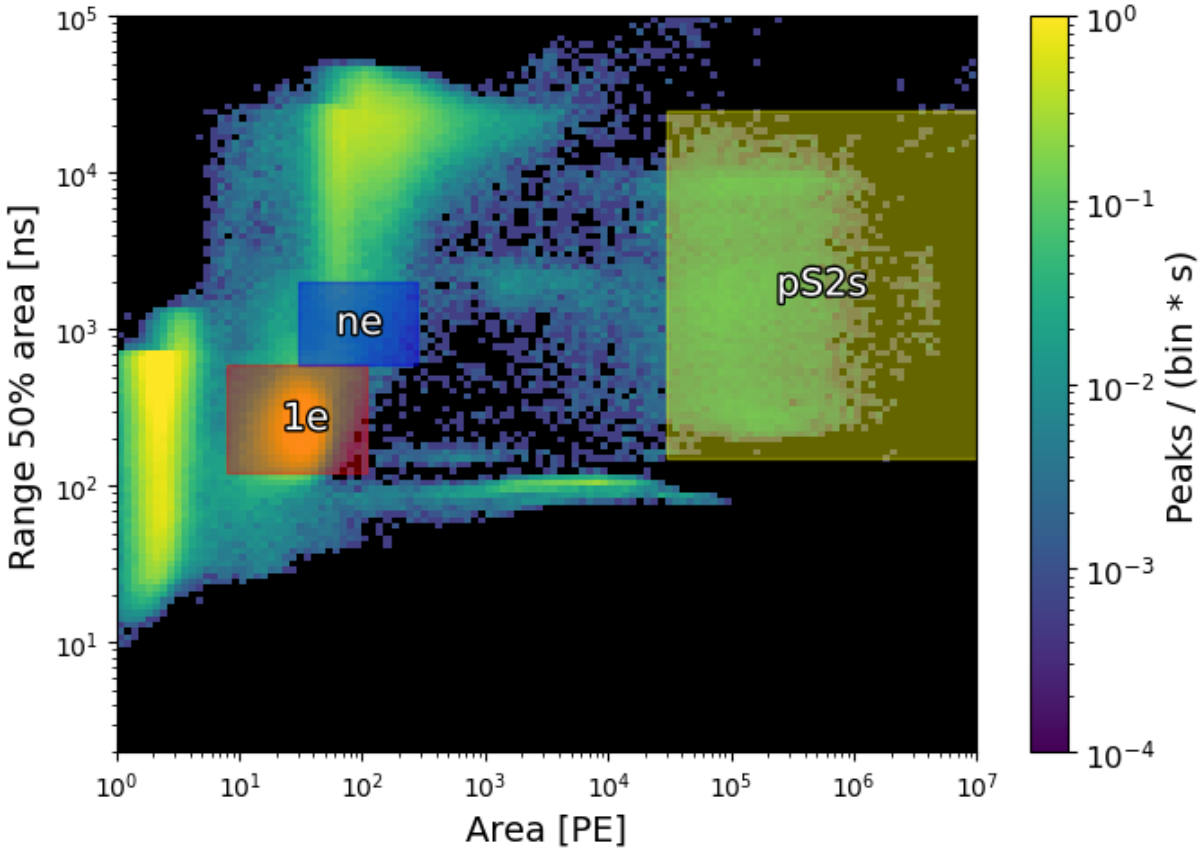
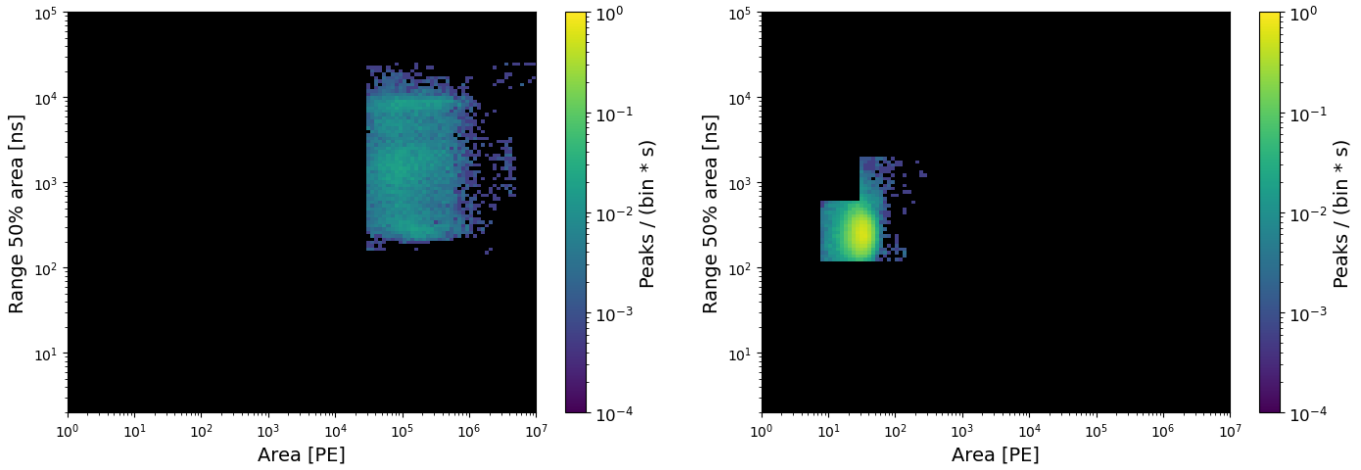


Figure 12: Highlights of the selection regions overlaid on an area vs range 50% area histogram of the same run as Fig. 11, with the $1e^-$ (**red**), ne^- (**blue**), and pS2 (**yellow**) regions shown. This figure is recreated from a similar one in Ref. [1], which uses 'area fraction top' as the colour bar axis instead of density of signals.

As a starting point, this selection is adequate. It encapsulates large parts of each population in a

clear manner. If we further perform a prompt electron cut on the small-s2s selected by the $1e^-$ and ne^- boxes, we return the delayed electron population. The two selections are shown in Fig. 13.



(a) Primary S2s

(b) Delayed Electrons

Figure 13: Final data selection from Ref. [1].

However, there are problems with this method. For example, without knowing more about the underlying nature of the data, there is no reason to think that one could not extend e.g. the pS2 to a lower area value, perhaps even down to the 150 PE limit of prior studies. Similarly for above the ne^- region, there is a vast swathe of signals which may be delayed electrons. Finally, the labels in Fig. 11 give no definitive boundaries; the 1% presence of type 0 and type 1 signals in the small-S2 population indicates as much. This method is not physically justified, which the above points highlight.

In the next section we will create a new selection methodology which will incorporate specific, physically justified cuts and filters to select maximally pure populations and eliminate anomalous signals. Some of these anomalous signals can otherwise appear to be delayed electrons or pS2s by where they lie on Fig. 12, meaning they certainly contaminate the selection used in Ref. [1]. The details of the process for each population are explained more in Secs. 6.2.1 and 6.2.2.

6.2 Improvements to Data Selection

Although the basic Straxen signal-type classifications are a useful starting point, these categories alone are too broad on their own in such busy datasets, when paired with only the area and width cuts shown above. Straxen can also get the classification wrong, for example sometimes mislabelling small S2s as S1s [63]. To fix this, we make use of an internal XENON study into how the peaks can be put into different 'subtypes' based on their areas and relations in time to one another [63].

This program is not complicated: simply scanning in time through the dataset to compare signals

to one another according to area. In doing this it assigns each signal a new 'subtype' - which offers a useful shorthand for better understanding and distinguishing between each signal. It can show which primary S2s have related S1s; which may be solo or a part of multiple-scatter interactions; and finally the status of its related electrons, either as photoionisation ones (the 'prompt' electrons mentioned before) or as ones which are part of a delayed electron train. Most of this is not relevant in our selection of signals: we are most interested in using these subtypes to distinguish between pS2s and small-S2s; and in correcting mis-classifications from Straxen. A more detailed explanation of the process and the related code used to implement this in the selection can be found at Refs. [63–65].

After this subtyping process we perform a variety of quality cuts on the respective pS2 and small-S2 populations to complement the selection, and ensure the removal of certain anomalous populations. These will be described and highlighted in due course.

6.2.1 Selection of Primary S2s

Within this method, primary S2s are distinguished from small-S2s generally as an S2-type signal which is more than 5 single-electrons (SE), approximately 175 PE, in area. The exact treatment of these signals and the subtypes assigned to pS2s in different situations can again be found in Refs. [63, 64]; the only classification relevant here is in how they are being selected relative to other S2-type signals.

All S1s and pS2s generate a series of prompt electron signals from photoionisation detected almost immediately after the initial S1 or pS2 signal; these were discussed earlier. Depending on factors such as initial pS2 size, interaction depth, and proximity to the detector edge, more photoionisation can occur, and in tighter bursts, meaning some of these signals can appear to be pS2s, at least in their area and widths. For this reason the subtyping program also implements a restriction that after a large S2 is detected, in order for another S2-type signal near it (i.e. within 1 fdt) to also be classified as a pS2, it must be more than 0.5x the initial pS2 area. In theory this allows for double- and multiple-scatter interactions, while preventing the selection of large prompt electron signals, as they themselves would need to have been caused by substantially larger pS2s. An example of this is shown in Fig. 14, where the prompt electron signal shown is also followed by multiple signals all >1000 PE in size, within 1 fdt of the initial pS2. The box selection methodology of Sec. 6.1 inadvertently introduces some of these unwanted prompt electron signals, as only the delayed electron dataset undergoes a prompt electron cut.

This is not a perfect method, as this 0.5x factor is arbitrarily chosen with respect to the correlation between pS2 area and quantity of prompt electrons. This has the effect of where, with a large pS2, any genuine co-occurring pS2 < 0.5 the size of the larger pS2 will be instead classified as a cluster of

photoionisation electrons, and thus be cut from the selection for the purposes of this investigation. This is also a problem for multiple-scatters which fall below this threshold. The frequency of these types of instances is unknown, however it is a clear limitation of the selection. It would be best to find the relationship between pS2 and prompt electron sizes in order to choose a more appropriate factor for this cut; however, this is beyond the scope of this work. An example of a double-scatter which would be mis-classified as photoionisation is shown in Fig. 15.

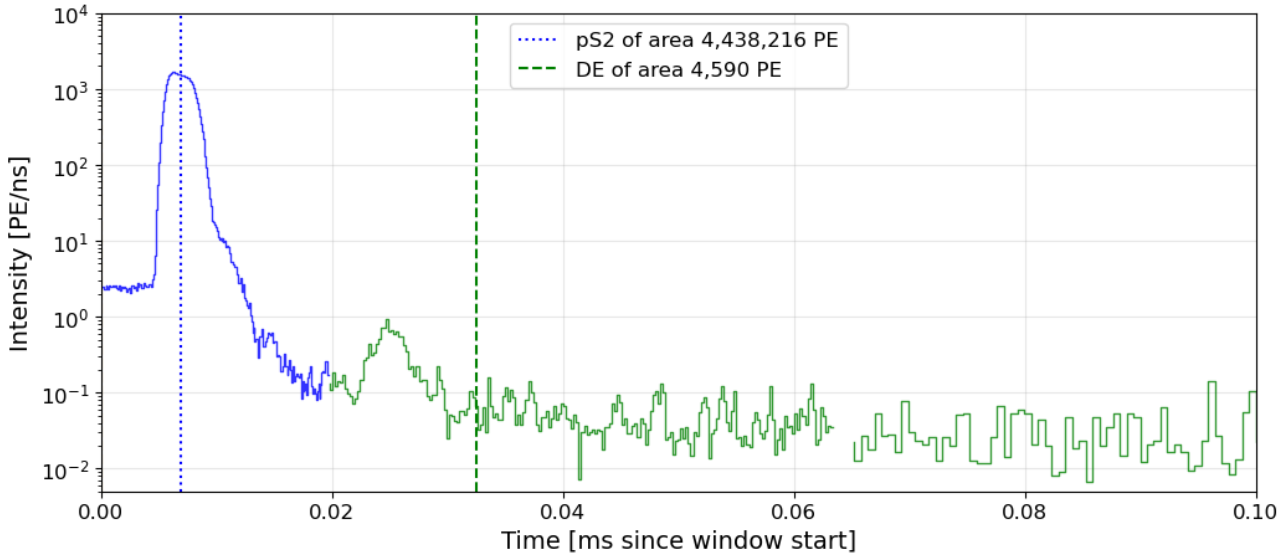


Figure 14: An example of a large primary S2 followed very closely by a series of prompt electron signals, the first of which measures at 4,590 PE in area, large enough to be considered a pS2 in its own right when viewed in isolation. Note the logarithmic y-axis.

We also perform a quality cut to remove any pS2s with an "area fraction top" (AFT) below 0.5. AFT is the ratio of the signal's area detected by the top row of PMTs as compared to the bottom row. Since S2-type signals happen in the GXe layer, we expect a high AFT ($\gtrsim 0.6$). These low-AFT signals usually have a relatively low area for a pS2 ($\lesssim$ 1000 PE), with a typical 'range 50% area', or width, of between 1000-4000ns. Most importantly, they are almost exclusively preceded by abnormally large S1s (>10000 PE, with a normal S1 area being ~1000PE). These factors point to these 'pS2s' actually being PMT 'afterpulses', erroneous signals produced by PMTs themselves after some signals, rather than by any genuine signals. More about afterpulses can be read in Refs. [66, 67]. A full removal of afterpulse signals would be beneficial to the selection, but is beyond the bounds of this work, so the simple AFT cut was made instead. An example waveform of these low-AFT 'pS2s' (in labelling only) is shown in Fig. 16.

Finally, we perform a rough cut on pS2s above 40,000 ns in width, as this region is populated primarily by signals called "e^--bursts", a phenomena noted by a few previous studies [9, 10]. These e^--bursts are large, wide signals sometimes consisting of as many as tens of thousands of electrons,

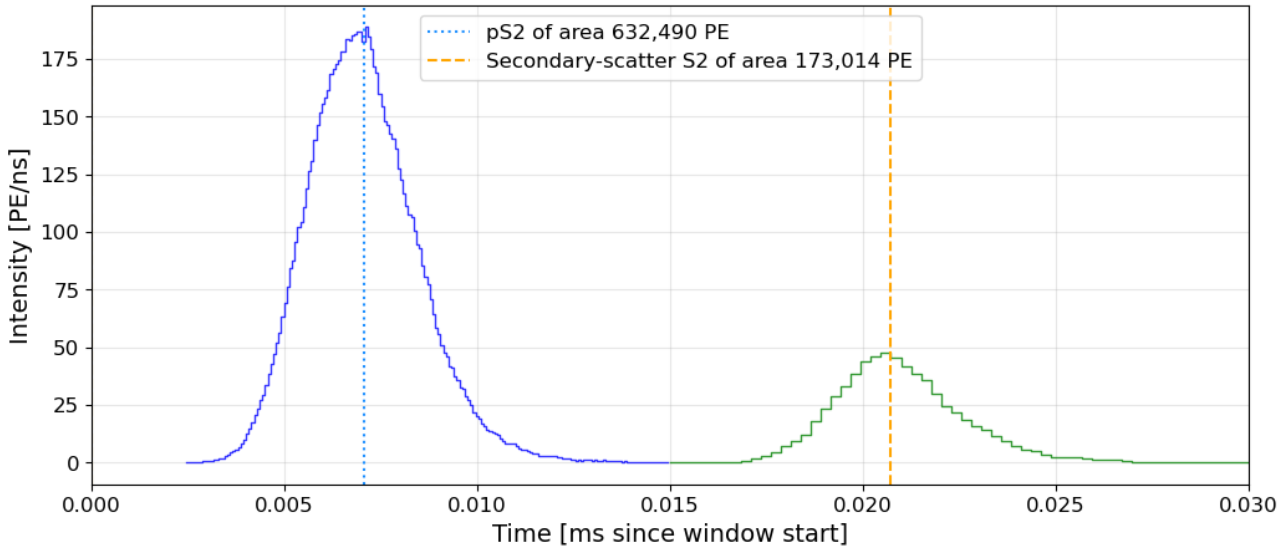


Figure 15: An example of a large primary S2 which is followed by a large secondary scatter S2. The waveform shape of this secondary signal, as well as its proximity in both time and location in the detector ($<$ 2.5 cm separation), make this highly likely to be a double-scatter event. Note now the y-axis is not logarithmic.

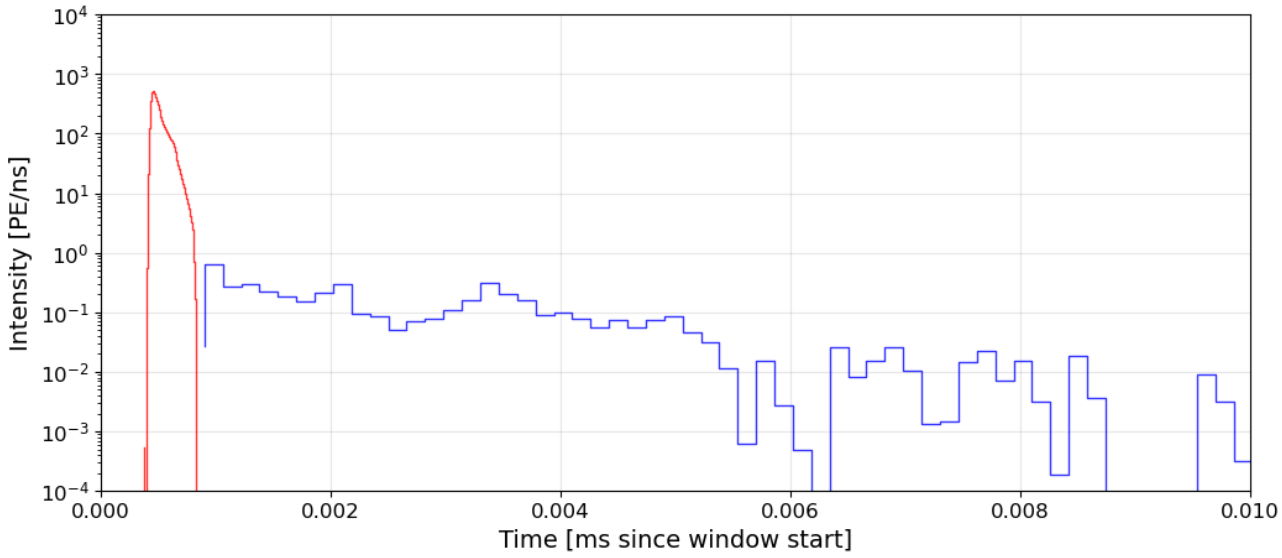
which register as pS2s due to their areas and widths. However, Ref. [10] notes that they contain no S1 or S2 pulses, only clusters of single electrons. Fig. 17 shows a representative waveform of a typical e^--burst signal. The main hypothesis for the physical cause of e^--burst is that electrons trapped at the LXe surface get released all at once, though the exact triggers are not yet well understood, and furthermore whether they themselves are also a cause of delayed electron emission or just a form of it is also in question. They can also form on the surface of metals in the detector, meaning they are commonly found along the edges of the TPC. An internal XENON study and plug-in for selecting them was developed in Refs. [68, 69], though too late to be implemented into this work, and therefore here only a rough cut was performed, which removed $O(100)$ signals per run, 1% of all pS2 signals. Ref. [10] analyses e^--burst as small as $10e^-$ in size, meaning this may also a significant background to the entire pS2 population.

After all of this, the new selection for pS2s appears as in Fig. 18a.

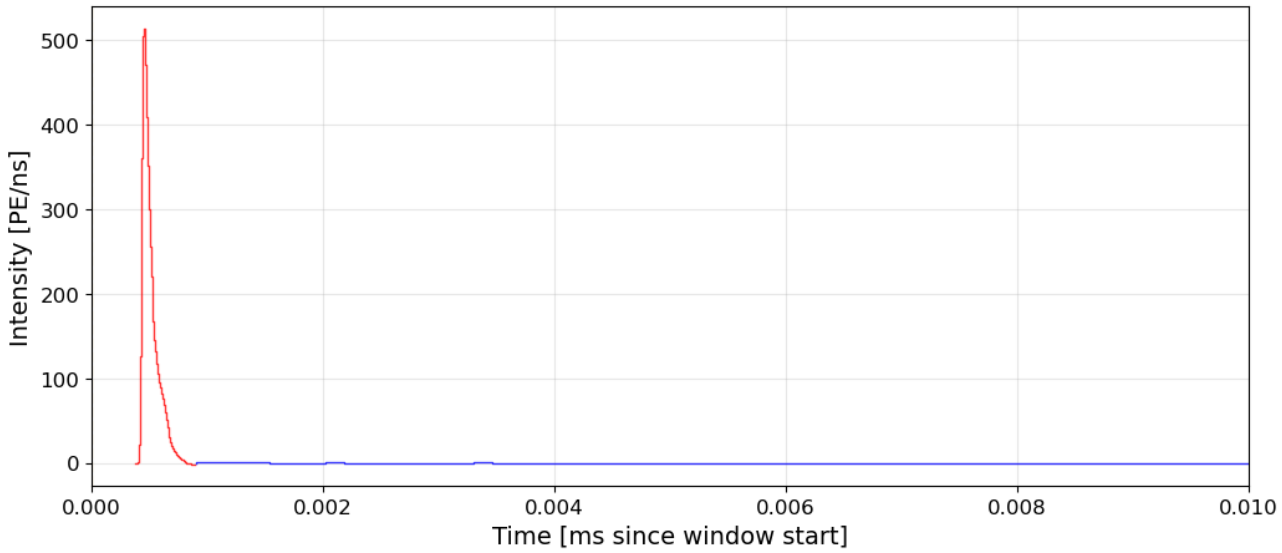
6.2.2 Selection of Delayed Electron Signals

After the selection of pS2s, we are left with a wide dataset of small-S2 signals, in which is our desired delayed electron population. The remaining signals are photoionisation electrons, which make up $\gtrsim$ 80% of this dataset.

To obtain the delayed electron population, a prompt-electron cut of 5 fdt after every pS2, and 2 fdt

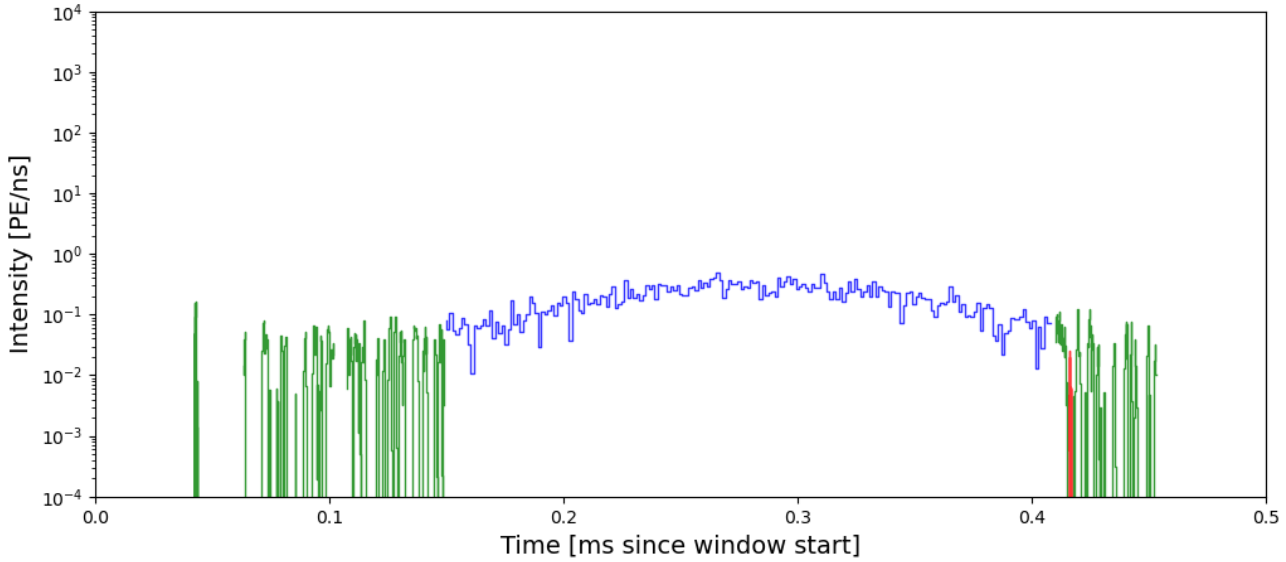


(a) A low-AFT pS2 preceded by a high-area S1, logarithmic y-axis

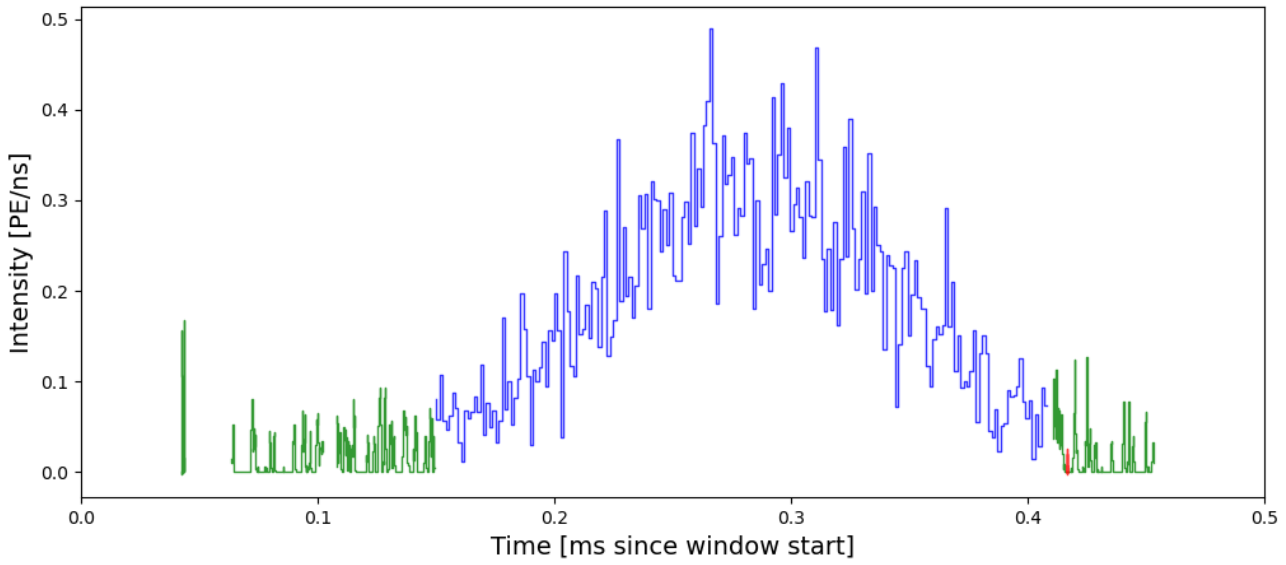


(b) The same low-AFT pS2 preceded by a high-area S1, linear y-axis

Figure 16: Time vs signal intensity waveform plot on a logarithmic y-axis (**Top**), and on a linear y-axis (**Bottom**). They show the waveform of an S1 with an extremely high area, followed by a wide, low-area, low-AFT S2, likely PMT afterpulses, that is mislabelled as a pS2 by the subtyping programme. The data has been coloured to show the Straxen peak classifications of S1s (**Red**) and S2s (**Blue**).



(a) The same example e⁻-burst with width >40,000 ns, logarithmic y-axis.



(b) An example e⁻-burst with width >40,000 ns, linear y-axis.

Figure 17: Time vs signal intensity waveform plot on a logarithmic y-axis (**Top**), and on a linear y-axis (**Bottom**). Although the subtyping has labelled this central signal as a pS2 (**Blue**), e⁻-burst studies show that these signals are single- or few-electrons which have been grouped together into a larger waveform by the Straxen processor [10, 68], and subsequently labelled as a pS2 by the subtyping programme. Delayed electrons are coloured in **Green**, and S1s in **Red**. Also apparent in this waveform is the difficulty Straxen has understanding signals of this type, with an S1 marked just after the pS2 having occurred. In the pairing mechanism of this labelling, this should indicate that this S1 is associated with another pS2 soon after it, which there is not.

after every S1, is performed on the whole small-S2 population, as photoionisation can occur after both signal-types. This can be seen in Fig. 5. This is prompt-electron cut was described before in Sec. 5.1. Figs. 18b and 18c show the separation we get from this cut on the small-S2 signals.

Furthermore, some quality cuts were performed on this delayed electron population where there were known issues. One such example is the removal of any signals which were detected exclusively by the bottom PMT array, as this is indicative of 'dark counts', anomalous signals caused by thermal noise in the PMTs. More information about the removal of these undesired populations can be found in Ref. [70]. The subtyping program also identifies S1s which were misclassified by Straxen as small-S2s, effectively overriding the Straxen type classifications in order to remove these signals from our selection.

Finally, we implemented the 'counting' of electrons for each signal, the details of which can be found in Sec. 7.2, or in Refs. [65, 71]. This was used in the selection to cut the few signals ($\mathcal{O}(100)$ relative to $\mathcal{O}(10{,}000)$ in a 30 minute run) which were >5 electrons in size but still classified as delayed electrons by the subtyping software. While some of these may be true delayed electron signals, many were found to be relatively wide and immediately preceding S1s, indicating some signal reconstruction artefact rather than a genuine delayed electron from some pS2 prior in time to that S1. This also then puts the delayed electron selection in line with prior studies in their selection of delayed electrons as those up to the $5e^-$ multiplicity, such as those of Refs. [11, 52]. Although these studies select the 5 electron limit somewhat arbitrarily, it is a sensible selection within the realms of the potential causes of the power law and the physics of the related processes discussed in Sec. 4.2, and so it was chosen as a restriction for this study also.

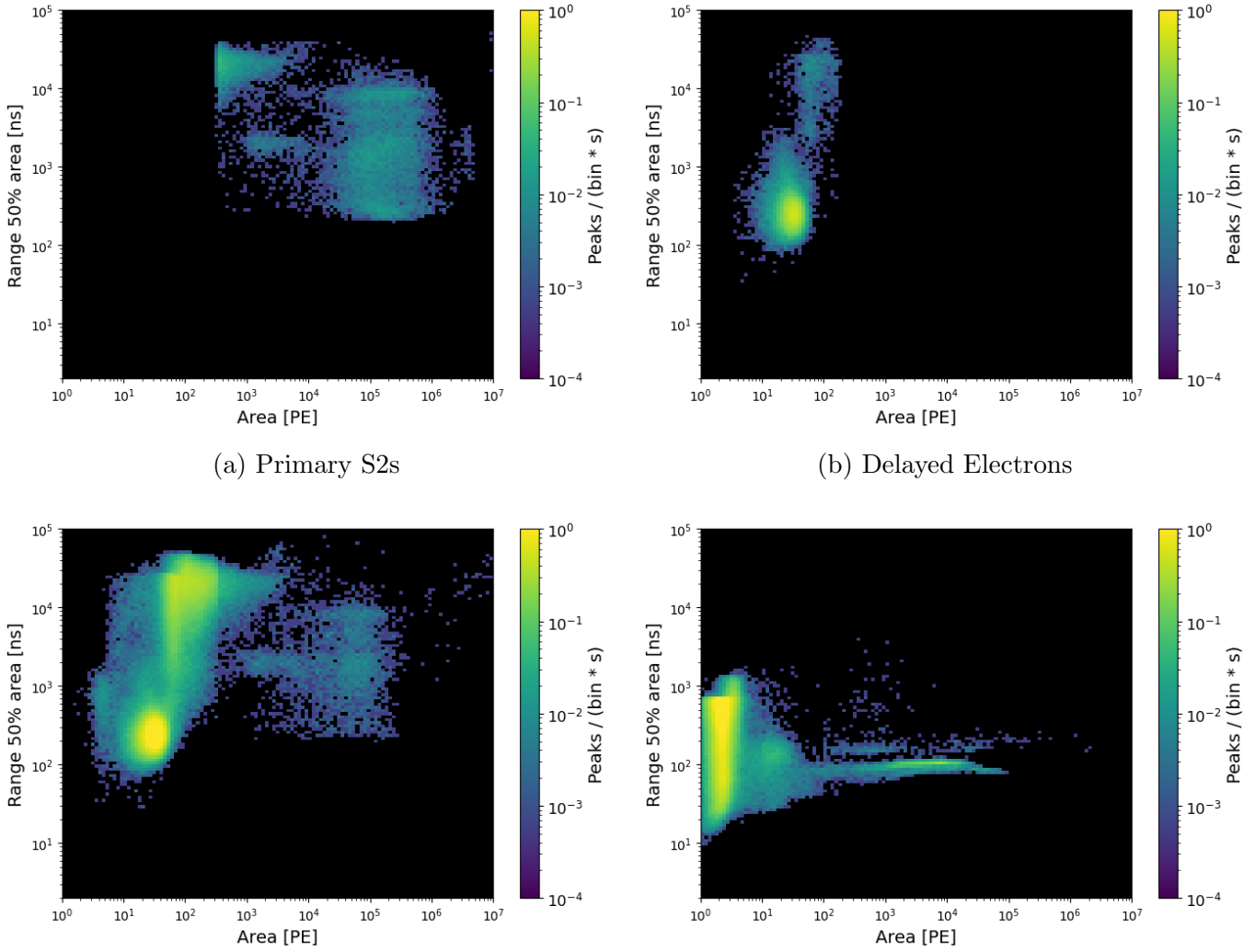
The new selection for delayed electrons is shown in Fig. 18b

6.3 Comparison of Selection Methodologies

With this selection implemented, we now have a more full view of the signals available to analyse. Fig. 18 shows the effect of the described cuts and filters on separating the populations from a single SR0 run. A summary of the process can be viewed as follows:

1. Only S2-type signals are selected initially. This cuts out all S1s and 'junk' signals, which Straxen is unable to classify. The distribution of the signals cut here is shown in Fig. 18d.

2. The S2-type signals are either primary S2s, prompt electrons, delayed electrons, or some form of anomalous signal (e.g. PMT dark counts). We utilise the subtyping program to sort between these and select the pS2s, which then undergo quality cuts. This selection is shown in Fig. 18a.



(a) Primary S2s

(b) Delayed Electrons

(c) Prompt Electrons

(d) Remaining signals (S1s, junk)

Figure 18: Area vs range 50% area density histograms for the same run as Fig. 11 showing the selection for primary S2s (**top left**) and delayed electrons (**top right**); along with the unselected or cut signals in the form of prompt electrons (**bottom left**) and S1s or junk signals(**bottom right**). Note that the prompt electron populations do likely include some double- and multiple-scatter signals, especially when looking at higher-area signals, however it was not possible to separate these signals from high-area prompt electrons in this work.

3. The signals not designated as pS2s are thus some form of small-S2 signal. They then undergo a cut to remove all signals within 5 full drift times (1 fdt = 2.3 ms for XENONnT) after a pS2, or 2 fdt after an S1, to ensure all prompt, photoionisation electrons are removed. These removed signals are shown in Fig. 18c. This then leaves us with the delayed electron dataset.

4. The delayed electron population undergoes its own quality cuts, leaving the selection shown in Fig. 18b.

Now we may compare these final selections to those of Fig. 13. Importantly, there are no signals selected in the previous method which are not also initially selected in this improved methodology; the net is only cast wider. However, more processing is performed afterwards on this wider dataset to attempt to reduce contamination from doing this. In particular, the delayed electron selection differs from that of Refs. [1, 11, 52] in that we now allow for a larger prompt electron cut; removal of anomalous signals; and a more flexible restriction up to and including the $5e^-$ population without a fixed area boundary.

The selection of primary S2s is also expanded. As more genuine pS2s are identified, this allows also for the removal of more prompt electrons in the cut: this is reflected in an extra 10% of all small-S2 signals being cut when performing the prompt electron cut, from $\sim 75\%$ in the dataset of Sec. 6.1, to $\sim 85\%$ in the new selection. This is assuming all else is equal, and so does not account for the increase to 5 fdt from 3 fdt between methods, nor the change to also cut electrons 2 fdt after an S1. This shows that many more electron signals than previously thought are likely caused by photoionisation, just were not being picked up by nature of the limited pS2 selection.

However, there are still open questions when it comes to the selection, in particular that of primary S2s. The existence of smaller e^--bursts in the pS2 dataset is a problem in both selection methodologies, but the expanded search region of the new selection makes it a larger problem there; this will be talked about again in Sec. 8. Other known problems with the selection which were not able to be implemented due to time can be found in Refs. [72, 73]. Again both selection methods will include these problems, however the increased population of pS2s in the new selection means this may have an increased effect; this will be shown later to likely be such the case.

Overall, however, this selection methodology is a promising improvement on prior studies. With it, we can be more confident that the model is analysing a larger and more representative selection of all delayed electrons, and all the signals which would be causally correlated with the emission of these electrons, in the form of pS2s. We will now move on to looking at the way in which the modelling of these primary S2s and their delayed electrons may be expanded, in order that we may begin to better probe the physics at hand in this phenomenon.

7 Changes to the Model

With the new data selection, we can feel that the signals being analysed are a more comprehensive and complete view of the primary S2s and their related delayed electrons. This next section will look at how this delayed electron rate may be modelled better, expanding on Sec. 5.2. With these improvements and extensions, we see how the model can be advanced further to better map to the rate of these delayed electron trains.

7.1 Normalisation of the Model

One of the key issues with the original model was in the normalisation of the model function and its calculation. Although the nature of the model described in Eq. 1 necessarily means that there is no contribution from a given pS2, p, to the electron rate for some time t_{min} afterwards (now 5 fdt = 11.5 ms), the modelled tails of previous pS2s can still continue into this region unrestricted. Consequently, the probability density function (PDF) used in the likelihood fitting predicts a non-zero amount of electrons in each region $[t_p,\, t_p + t_{min}]$ for each of these given tails, where again t_p is the time of a pS2. The 2 fdt cut after each S1 will also cause the same effect. This issue was noted in Ref. [1], with the observation that this quirk may artificially steepen the power law exponent, especially around multiple-scatter clusters. This is certainly a non-negligible effect then, as nearly 40% of all pS2s in a 30 minute SR0 run tested were found to be a part of a double-scatter event, with a total of 60% being multiple-scatter events. Fig. 19 illustrates a toy model of two nearby pS2s, showing how the predicted power law tail from one pS2 can overflow into the prompt electron cut region after another nearby one.

Fixing this problem involved calculating the predicted electron rate as normal, but then applying a global mask on the region $[t_p,\, t_p + t_{min}]$, where t_{min} = 11.5 ms (5 fdt) for every primary S2 peak p (in the chosen analysis time window). We explicitly set those regions to 0, thus matching the real data cut which occurs. The same process was performed on S1 peak times, at t_{min} = 4.6 ms, to match the smaller photoionisation cut performed there. Further implemented was a renormalisation of the power law to remove any "dead-zone" contributions from these prompt electron regions to the total predicted electron rate of an analysis window. Specifically, this was done by subtracting the integrated probability mass (i.e., the cumulative density function of the PDF) between each region $[t_p,\, t_p + t_{min}]$ from the total normalisation term, adjusting t_{min} according to whether it was after an S1 or S2. These adjustments ensure that the physical nature of the prompt electron cut done in the data selection is reflected in both the ongoing rate of the model calculated at each time instance t, as seen in Eq. 2; as well as in its total predicted electron rate across a fit region. This lends the model more mathematical consistency. This is, as mentioned, especially important for multiple-

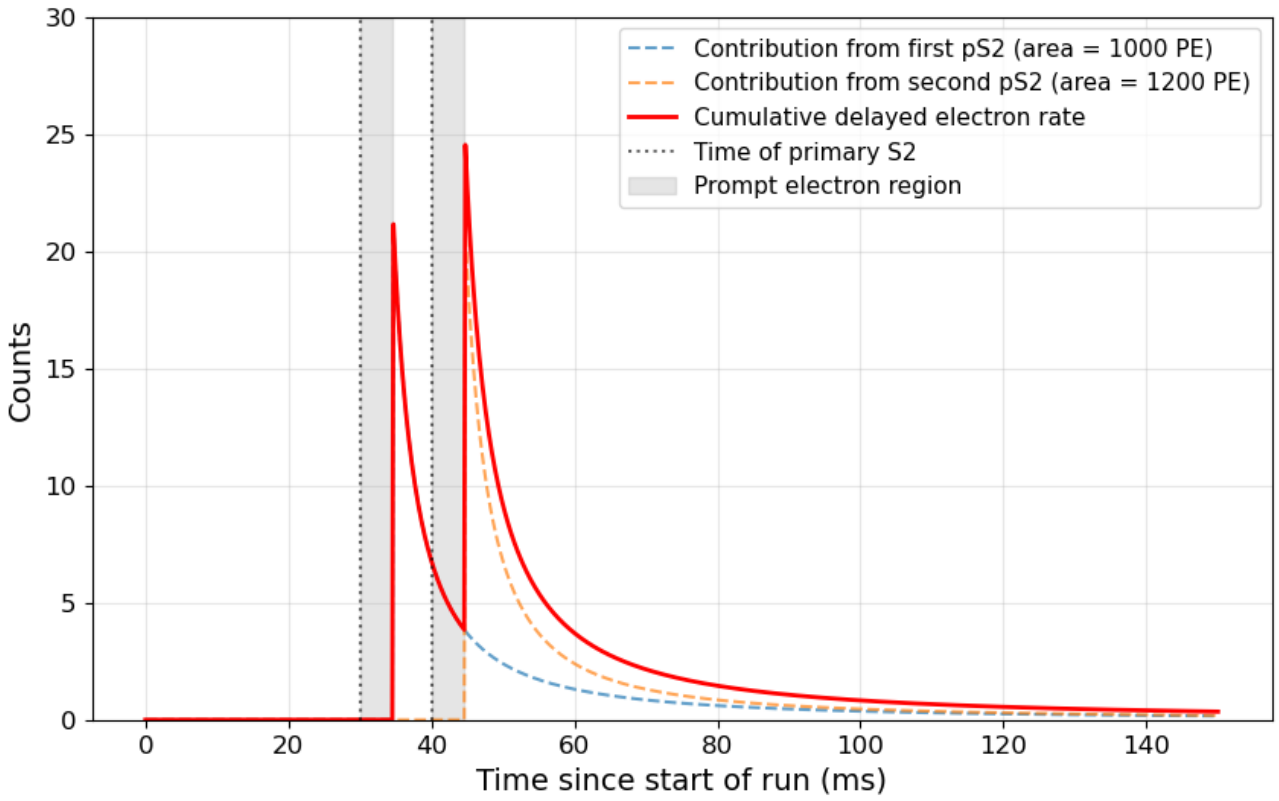


Figure 19: A toy model of a double-scatter event to illustrate how part of one delayed electron train may leak into the prompt-electron region of another subsequent pS2, thus disrupting the model. Due to the nature of the prompt electron cut, all small-S2 signals in this region will be removed from the data-set. However the size of the first pS2 would indicate to the model to expect many electrons in the train, which would then not match with the observed data. Note the prompt electron region has here been set as 2 fdt (4.6 ms) for simpler illustration, and the sizes of the pS2s labelled do not necessarily correspond to the actual number of electrons we expect to observe in their respective trains.

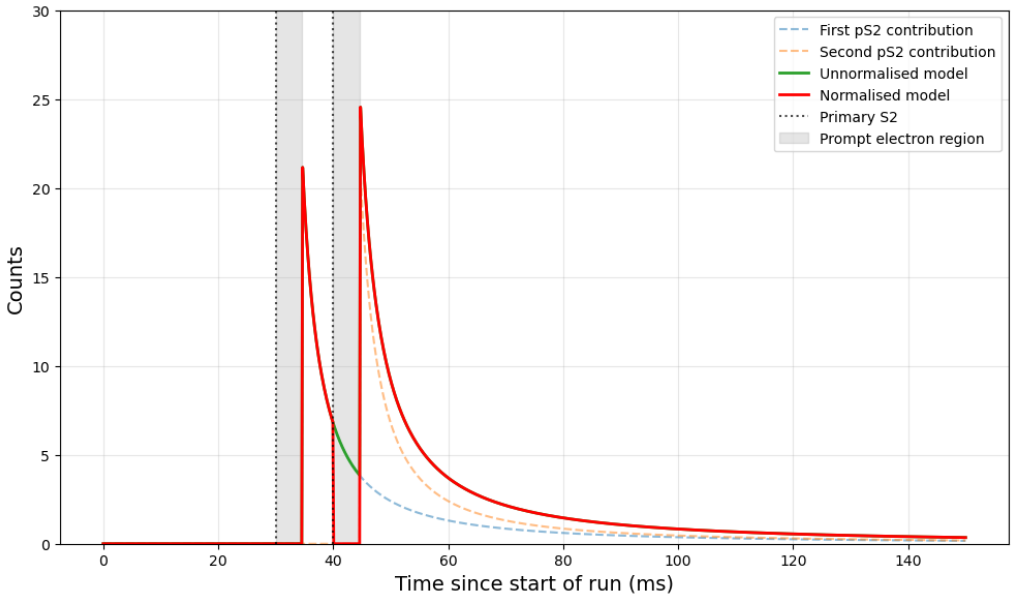


Figure 20: A visualisation of the normalisation fix implemented into the model on all prompt-electron dead-zone regions.

scatter events, which are frequent, and thus certainly allows the model to yield more meaningful results going forward. Continuing the toy model example, an implementation of this fix is seen in Fig. 20.

7.2 Electron Counting

In the model of Ref. [1], the model takes as inputs only the times of the delayed electron signals. However, as mentioned, previous studies have considered up to the $5e^-$ multiplicity as part of the delayed electron population, even calculating different power laws for the various multiplicities [11, 52]. While this new model does not evaluate the delayed electron population in the same way as prior studies, the nature of these populations and how they should be considered by our model is still an interesting question. This section looks at why we might wish to implement electron counting into our model, and why the chosen method was used over a potentially simpler implementation.

There are a couple ways electron counting can be implemented into the model and wider understanding:

1. Aiding in the selection of delayed electrons, as talked about in Sec. 6.2.2.

2. Measurement of the accuracy of the model factors. One of the the aims of the model is to understand the relation between the features of a pS2 and the number of delayed electrons associated with it, with the model prefactor, the value $C \cdot A_p^{\alpha} W_p^{\beta}$, providing a predicted number

of electrons for a given pS2 peak, p. Implementing electron counting may allow a more accurate comparison between this value and the actual number of electrons observed after a given pS2, especially within a given time window as a whole. This comparison is looked at more in Sec. 8.2.2.

3. Implementation into the model itself. If we want to truly understand the physical nature of delayed electron emission, then it may be beneficial to adapt the model to account for the number of electrons being detected, rather than purely their times. Additionally, although Sec. 4.2.2 shows that multiple-electron emission is far more than just a result of single-electron pile-up, this change to the model would also account for the pile-up population in our dataset. Whether this is justified physically is not clear; however, this does give us another avenue of interest to explore with this model.

The ability to 'count' the number of electrons in each signal is not something natively provided by Straxen, and instead needs to be calculated explicitly. One approach is to find the 'single-electron gain' for the run, defined as the average total charge per single electron signal. For SR0 this was found to be approximately 31 PE [74], meaning each single-electron reaching a PMT registered on average 31 PE in area. The amount of electrons per peak is then a trivial matter of dividing each signal's area by this value. This method is simple and effective, and yields distinct populations within the delayed electron dataset.

This method is not without shortcomings however, which make it unsuitable to use on its own. Instead, we here utilise a dedicated plug-in created for Straxen, the details of which can be found in Ref. [71], which employs a waveform quantisation algorithm to reconstruct the electron counts from the raw waveforms. More specifically, the algorithm performs the following process:

1. Each waveform is smoothed over a short time window to reduce baseline noise and fluctuations.
2. The total waveform is then analysed for low-points in intensity and broken up into substructures; each identified as a possible electron pulse.
3. Each substructure is then integrated to find its area, which is then divided by the single-electron gain value.
4. These values are rounded and summed to find a total reconstructed number of electrons for each peak.

This method is far more effective for distinguishing between clustered or overlapping delayed electron signals than a simple area/gain calculation. In the XENONnT data-acquisition system, electrons detected by the PMTs less than a couple microseconds apart will be clustered and labelled as a single

peak. An area/gain calculation would thus see this event as a smooth blob of charge, and divide the total by the single electron gain. While it may sometimes give the correct number of electrons, it is not a calculation with physical motivation or consistency. However, the waveform still encodes the information that these are individual events, meaning the quantisation algorithm can retrieve this and allow for a more physical view of the delayed electron events in the detector.

Figs. 21 and 22 compare calculations of electron multiplicities (up to $5e^-$) for the delayed electron dataset when applying a basic area/gain model as compared to the waveform-based electron counting algorithm. The area/gain method imposes rigid boundaries on the electron populations, while the electron counting algorithm allows for more flexibility in the variation of electron signal areas and widths. This illustrates why it is an important feature to implement, both in the selection of delayed electrons and in improving the physical accuracy of a delayed electron model.

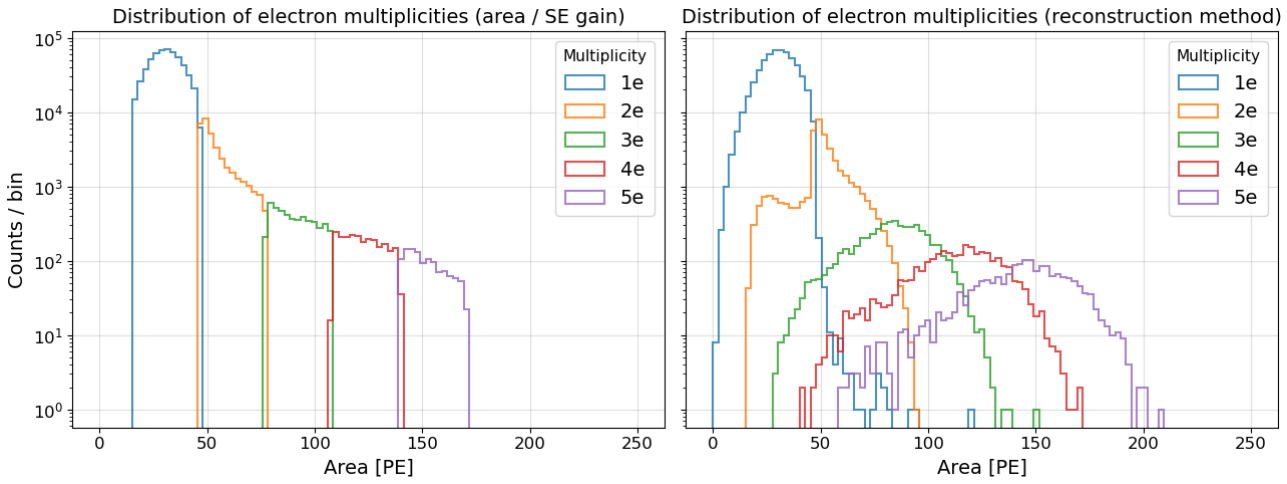


Figure 21: Counts per bin of different electron multiplicities over area for delayed electron signals. **Left**: Distribution using peak area/SE gain (here set at 31 PE); **Right**: Distribution using waveform quantisation and reconstructed electron counting algorithm. The distributions on the right notably primarily follow a Gaussian distribution, with a small exception in the $2e^-$ population, which is the distribution we expect electron multiplicities to follow. This figure is adapted from ones in Ref. [71].

7.3 Fiducialisation & Radial Factor Implementation

One useful feature of xenon as a detector medium is its ability to self-shield from background radiation, a fact mentioned in Sec. 3.1. This means we can restrict our search region for dark matter to only an internal section of the detector, and eliminate a significant fraction of background noise in our 'uncorrelated' delayed electron background. Fig. 23 shows the location densities of signals in the detector for primary S2s and delayed electrons, respectively; showing that the outer regions of the detector are by far the most active, especially for the delayed electron populations.

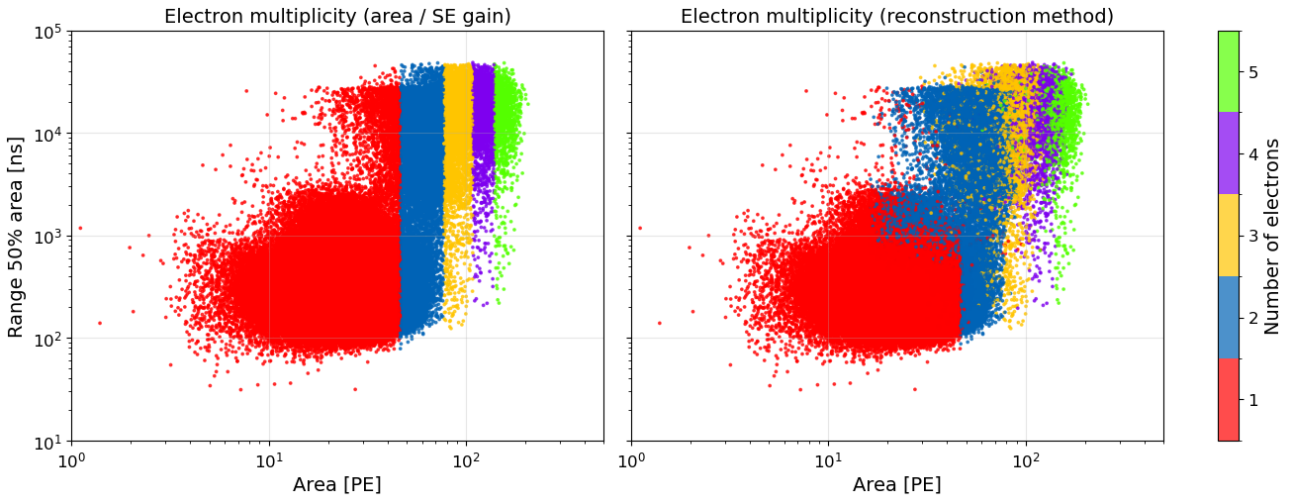


Figure 22: Area vs width for electron multiplicities. We can see there is a large dependence in width on the electron multiplicities, which is not accurately captured in the area/SE gain method (**Left**), but far better accounted for in the electron reconstruction method (**Right**). There is clearly some problem in accounting for electrons of high area but low width, which both methods assign, perhaps falsely, to higher-multiplicity electron populations. This explains the non-Gaussian shape of the $2e^-$ population in Fig. 21.

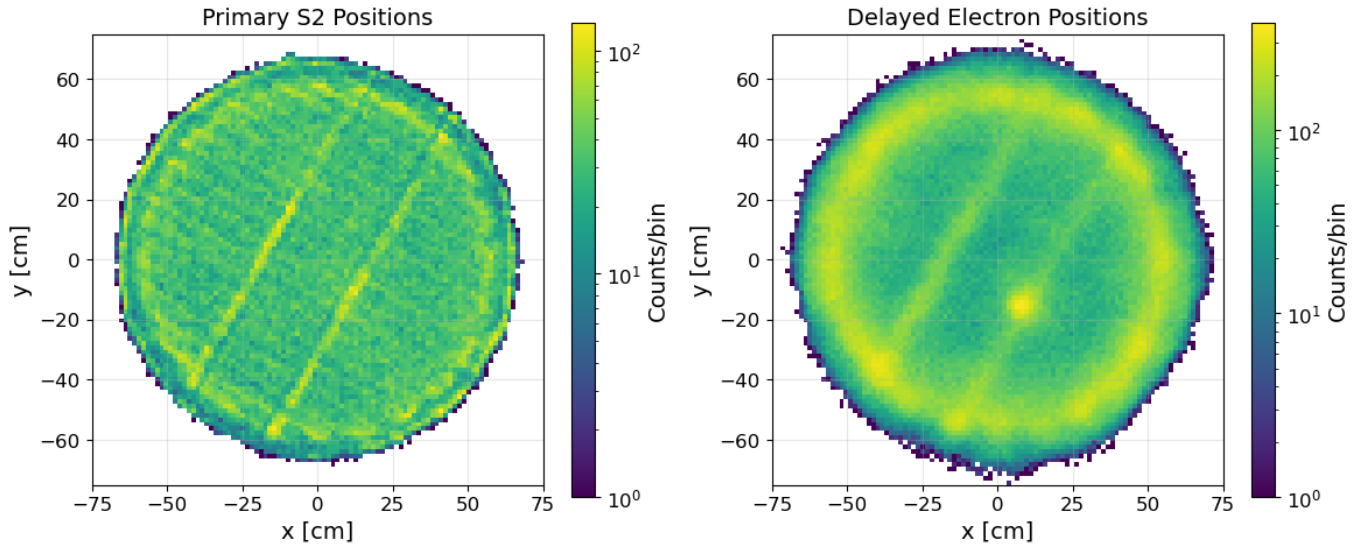
From these graphs, we can also see a disproportionate excess of events along the two transverse wires placed to reduce sagging of the gate and anode electrodes, as well as at a clear "hotspot" region in Fig 23b. Although it is believed that the majority (perhaps as much as $\sim 90\%$) of the delayed electron population is correlated with pS2s, there is still an uncorrelated background population which cannot be ruled out [75]. This population is most likely from photoionisation with the metal walls of the detector, as well as with the wires, contributing somewhat to the excess in these regions. Similarly, the hotspot region is thought to be uncorrelated to electron trains after pS2s [76], contributing perhaps another population of single electrons, though not ones part of delayed electron trains. If we wish to eventually use this model to distinguish between delayed electrons and genuine dark matter signals, then developing a position-dependent aspect to the model to mitigate for these regions would be of great benefit to future analysis. Importantly, although light DM scattering would be its own uncorrelated population, it would not be correlated to the edges of the TPC, and would have the same likelihood of being found anywhere in the detector.

We may further wish to mitigate for these regions due to the uncertain nature of the phenomena of e^--bursts and their contributions to the model. As mentioned in Sec. 6.3, these signals are often mistaken in selection as pS2s, even though phenomenologically they are large clusters of single-electrons. As e^--bursts tend to happen most frequently at the detector edges, as well as along the transverse wires, they would contribute far less to the model under the implementation of a radial scaling factor, the exact mechanics of which will be explored in the following sections. Although a full position-dependent model incorporating the wires and the hotspot regions was deemed to be

beyond the scope of this project, a preliminary radial factor was implemented. However, this still had very promising results, as is shown in Sec. 8, which a complete radial model would likely only improve further.

Implementing a radial factor into the model is a two-step process:

1. Perform a fiducial volume cut, the dimensions of which are determined and justified in Sec. 7.3.1, on the delayed electron population. This ensures some removal of an uncorrelated population at the detector edges.
2. Update the predictive electron rate factor, $C \cdot A_p^{\alpha} W_p^{\beta}$, to accommodate for pS2s throughout the detector having a portion of 'correlated' electrons removed by the fiducial cut. We can't perform a fiducial volume cut on pS2s also, as this then causes problems for their correlated electrons inside the volume. Instead, we scale their contribution with a factor between 0 and 1 depending on the proportion of their electrons remaining inside the fiducial volume. Sec. 7.3.2 explores how this works, is determined for the model, and also provides a more in-depth explanation as to its need.



(a) Primary S2s

(b) Delayed Electrons

Figure 23: Location of each pS2 and DE signal in 13 background SR0 runs (see the run list in Appendix B, after the selection methodology of Sec. 6, meaning prompt electrons were already removed.

7.3.1 Selection of Fiducial Volume Radius

The selection of the radius for the fiducial volume can be calculated using Fig. 24. Fitting a Gaussian over a histogram of electron counts according to their radial distance from the centre of the TPC,

we can see where the majority of events occur, and indeed, most occur within 20 cm of the wall of the TPC (at about 67.5 cm).

To determine the optimal radius to use, we can evaluate the signal sensitivity curve, where sensitivity is defined as exposure/$\sqrt{\text{counts}}$, with exposure scaling with cross-sectional area, πr^2. As r increases, so does the target area for analysis; however, so does the number of background events.

As such there exists an optimal radial value for the maximum sensitivity of analysis, which from Fig. 25 can be seen to be near the 1σ value from the Gaussian fit. In fitting these curves for the list of SR0 background runs, following the approach of Ref. [77] (an SR0 analysis of fiducialisation for the S2-only search), it was found that the value of 45 cm chosen was also an appropriate cut for this work. As can be seen from Fig. 25, it results in only a small loss of sensitivity compared to the 1σ value of 49.3 cm, while greatly increasing the amount of wall-events removed.

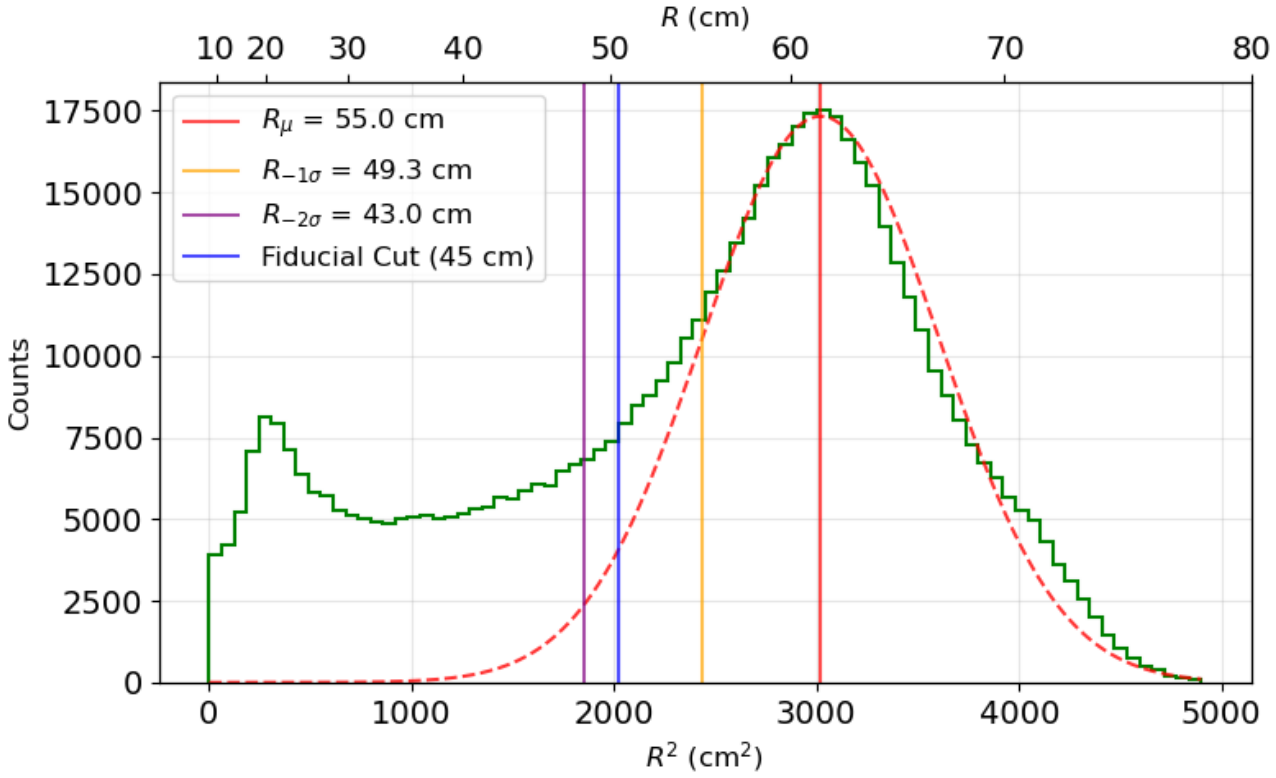


Figure 24: Histogram of (x, y) distribution of delayed electrons as a function of R^2 (Bottom axis). Wall events can be fit with a Gaussian to indicate where a fiducial volume cut might best lie. This and Fig. 25 are adapted to the SR0 runs being used in this work from [74], which settles on a value of 45 cm in order to optimise cutting out the wall volume without significant reduction in signal sensitivity.

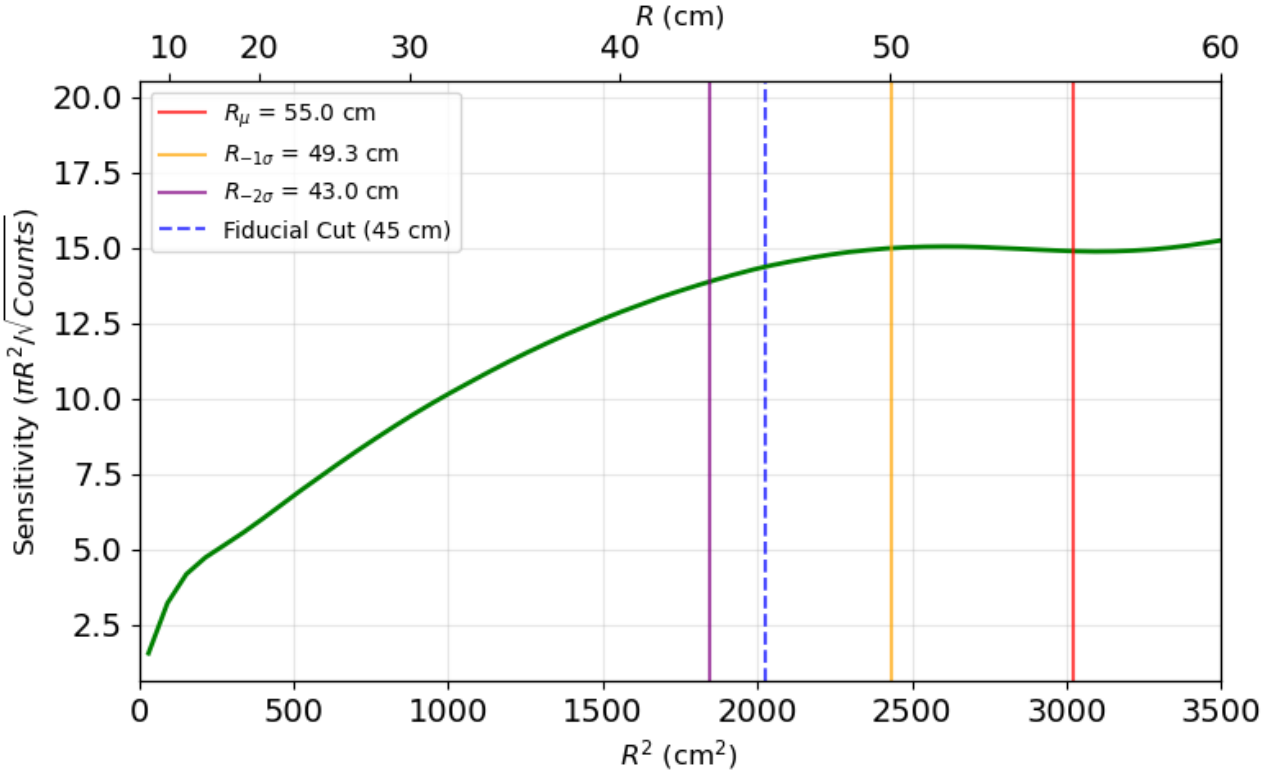


Figure 25: Sensitivity of the delayed electron dataset vs R^2

7.3.2 Adding the Radial Factor

With a suitable boundary for fiducialisation selected, we now require a model which matches the new reality of the data. To understand why this change is necessary, let us review the physics of delayed electron emission.

Within the model, every pS2 is expected to correlate to a certain number of delayed electrons, a function of its size and width. It is this factor which determines the model rate over time, fitting according to the data. It provides a rough statistical approximation as to whether the delayed electrons observed are likely to be 'from' a pS2 or not. A perfect model would be able to identify unexpected excesses in the model from this predictive rate, thereby hopefully identifying low-mass dark matter events. This is why properly normalising the model in the "dead-zone" regions, where we have performed the prompt electron cuts, is important to the model, as well as fiducialisation, which will help remove known background.

However, in introducing this fiducialisation, there then exists a mismatch between the predicted amount of electrons from a pS2 in the model and the real amount in the empirical data. The magnitude and nature of this mismatch is a physical question: asking the extent to which the delayed electron population may be 'position-correlated' with the primary S2 population. As mentioned

in Sec. 3.1, the electron cloud released following an interaction experiences minimal drift in the transverse plane - i.e. the one perpendicular to its upward motion towards the LXe surface. This means that, regardless of which hypothesis in Sec. 4.1 is dominant, we expect a given delayed electron train to be in a similar xy-position to the pS2 that caused it. The exact definition for a delayed electron being position-correlated will be addressed later in this section.

Fig. 26 illustrates a simplified model of the problem occurring from the fiducial volume cut: any pS2 on the inner boundary of the fiducial volume will have some proportion of its related delayed electrons removed by the cut, and thus the model will over-predict the amount of electrons that pS2 contributes to the total electron rate, and under-fit the data on a whole. The inverse scenario, of a pS2 just on the outside of the fiducial volume, illustrates why we cannot perform the fiducial volume cut on the pS2 population also: doing so would leave some portion of its delayed electrons unaccounted for by any pS2 inside the volume, and thus they would act like uncorrelated background signals in the model, effectively making the fiducial volume cut redundant.

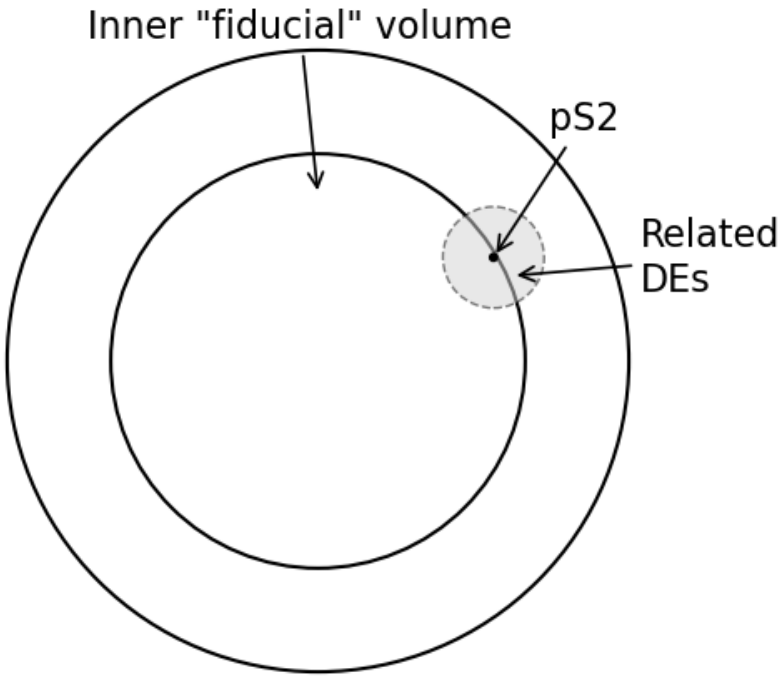


Figure 26: A simplified diagram of the TPC cross-section. This highlights what the fiducial volume would look like schematically, as well as the effect the cut would have on primary S2s and their related delayed electron trains. Primary S2s towards the centre of the volume would be relatively unaffected, while ones closer to the edge would have increasingly more of their related delayed electron population cut.

The solution to this problem is, as mentioned before, to introduce a radial scaling factor into the model, between 0 and 1, such that pS2s towards the centre of the TPC are considered to 'keep' a greater proportion of their related delayed electrons relative to those further from the centre. Those right at the edges may effectively have none of their related delayed electrons remaining inside the

fiducial volume, and so their factor will be close to 0, reflecting the minimal contribution they make to the electron rate inside the fiducial volume. A scaled sigmoid curve, seen in Fig. 27, was decided to be the appropriate function for the model. It takes the form:

$$f(r \mid S) = \frac{1}{1 + exp(S(\frac{r - r_f}{r_c}))} \tag{3}$$

where $r_f = 45$ cm is the fiducial volume radius; r_c is the radius of 'position-correlated' delayed electrons related to a given pS2; and S is a scaling factor on the function which is also allowed to be freely fit by the likelihood function. In Fig. 27 $r_c = 11$ cm, and S is set to 3, for illustrative purposes.

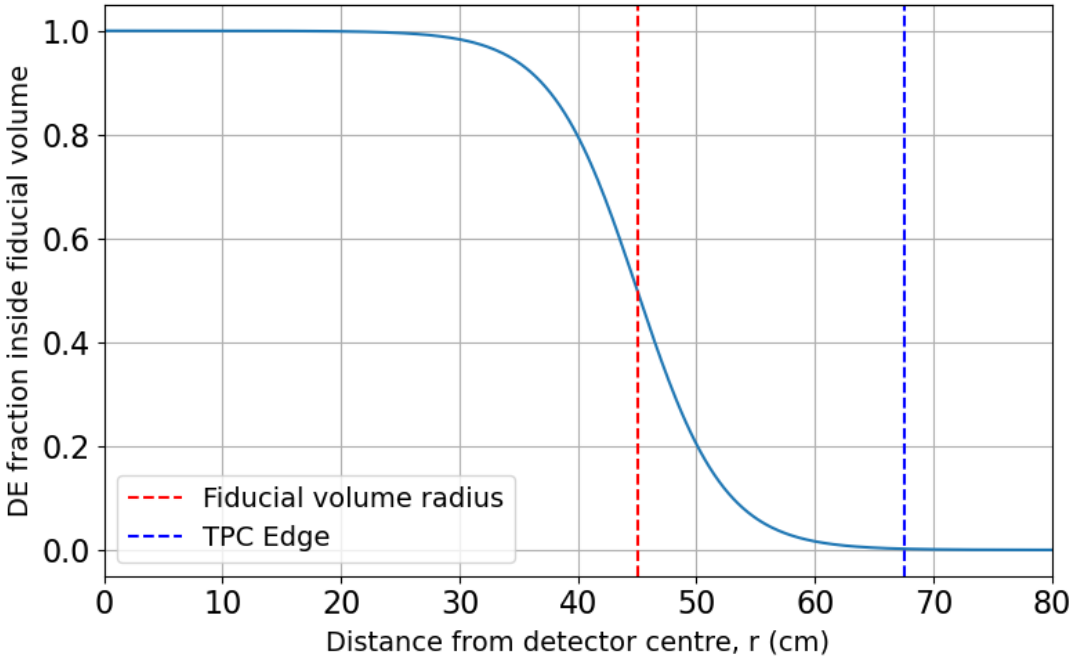


Figure 27: The radial factor, $f(r)$, added to Eq. 1 to accommodate for the fiducial volume cut.

To understand the physical significance of the parameter r_c in the model, note that this is not a position-dependent model in the conventional sense: the model is not matching up delayed electrons with pS2s based on their spatial location in the TPC, nor is the model able to predict the electron rate at a specific location in the detector. The value r_c is simply a measure of how far a delayed electron signal, or population on a whole, can from a pS2 before the rate becomes insignificant. This corresponds physically to the radius of the small, grey circle in Fig. 26, and is therefore key to the tuning of the radial factor as it determines how much a pS2's related signals, empirically-speaking, may have been cut out as the pS2 gets closer to the fiducial volume boundary. Too large and the radial function becomes too shallow, penalising in the model even pS2s at the centre of the detector as potentially having related DEs outside the fiducial volume; too small and the function becomes absurdly restrictive, essentially operating as a fiducial volume cut on pS2s also, thus defeating the

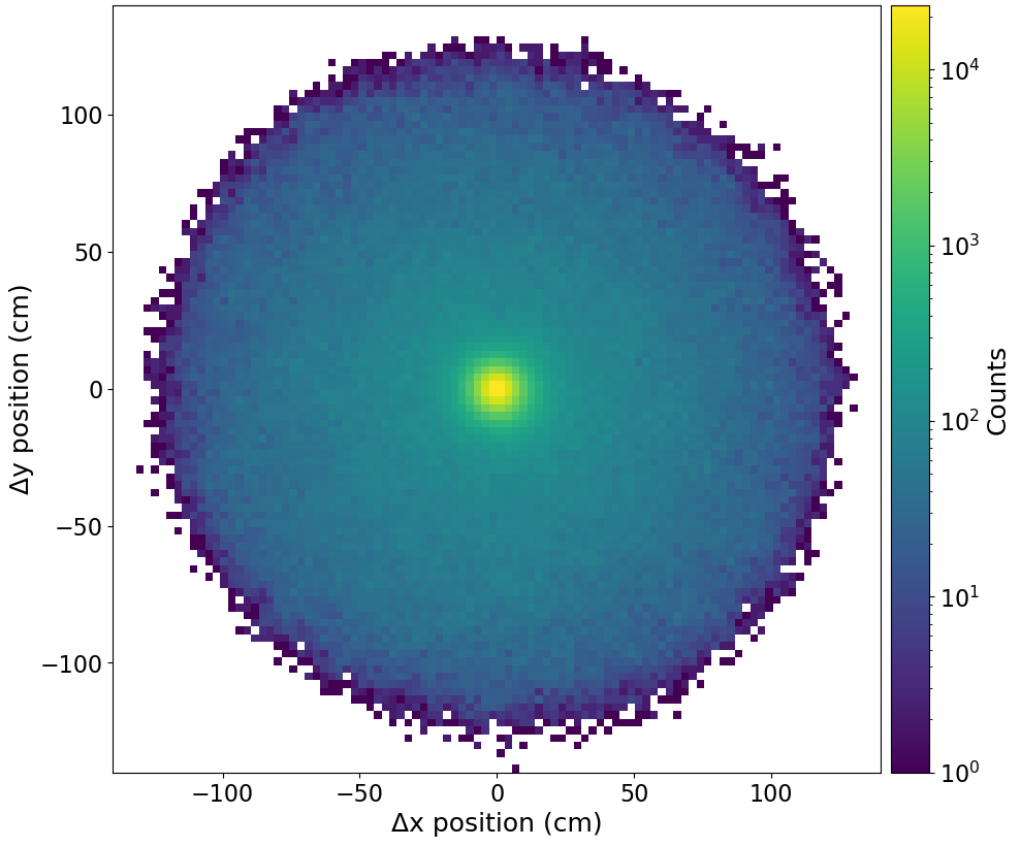


Figure 28: Density histogram of delayed electrons according to their distance to the most recent prior pS2. We can see a large correlated population at the centre, on top of a far smaller uncorrelated background.

purpose.

We can plot a 'position-correlated' population of delayed electrons by assigning each DE to the pS2 most recently preceding it in time, and then plotting the position difference. This is seen in Fig. 28, where there exists a bright hotspot in the centre, justifying the belief that the processes behind delayed electron trains are most likely happening in a similar x-y position to the interaction that caused them. This population can be fit to a Gaussian over a modelled background, from which we find that $r_c \simeq 11$ cm is the value at which 99% of this position-correlated population is contained. This procedure of separating the correlated Gaussian component from a modelled uncorrelated background using a 2D fit can be found in more detail in Refs. [11, 78].

With the values for the fiducial volume radius, r_f, and the position-correlation radius, r_c, now determined, the adjusted model now looks like this:

$$R_p(t,\, t_p,\, A_p,\, W_p,\, r_p \mid n,\, t_{min},\, \alpha,\, \beta,\, S) = \begin{cases} 0 & t \leq t_P + t_{min} \\ A_p^{\alpha}\; W_p^{\beta}\; f(r \mid S)\; (t - t_p)^{-n}, & \text{otherwise} \end{cases} \tag{4}$$

where $f(r_p \mid S)$ is as defined in Eq. 3, and t_P indicates that the rate contribution from a given peak is now being set to 0 for dead-zone regions after all peaks P in the analysis window, not just its own dead-zone region.

This means that the total electron rate at a given time is given as before in Eq. 2:

$$R(t \mid k,\, C, ...) = k + C \cdot \sum_{p}^{P} R_p$$

With this we can now go on to look at fitting the data to this model, and see if these changes do indeed improve the model in meaningful ways.

8 Results

We have now addressed the two key areas of this delayed electron train model: first, that of the data being used by the model, hoping to capture every pS2 signal and every delayed electron signal while cutting out unrelated contaminating signals; second, that of the model itself, fixing the flaws in its logic, and working to expand it to better do its job. This section will now analyse the extent to which these changes have helped, and whether we may further our understanding of delayed electron emission as a result.

We will compare four variations in how they perform: the 'old' data selection outlined in Sec. 6.1; the 'new' data selection outlined in Sec. 6.2; the 'radial' addition to the model defined in Sec. 7.3.2; and the 'electron-counting' model defined in Sec. 7.2. In the electron-counting case, each instance of a delayed electron signal has been cloned in the dataset by the number of reconstructed electrons given for that signal; this will allow us to examine if the power law signal is determined by number of electrons being picked up by the PMTs, or by the number of size-independent electron releases.

The variations on the data selection will both use the same model for fairness, which is the original model of Ref. [1], except with the fix to the rate normalisation described in Sec. 7.1. The new data selection model also includes the 5 fdt cut on the data (and thus also in the power law model) after every pS2, and 2 fdt after every S1; whereas the old data selection only cuts 3 fdt after every pS2 in both the data and the model. The radial and electron-counting variations on the model will each use the new data selection. In this way, we may compare between each model across different timescales and windows of data analysis, in order to see which may best explain the phenomenon of delayed electron emission.

Sec. 8.1 describes how we went about running the model on the data, and general context of the results. Sec. 8.2 looks at the different parameters returned by each of the variations; from this it draws some qualitative interpretations of the model fitting, as well as parsing out physical meaning which may lie in the data as a result. Beyond this, we may also perform goodness-of-fit testing in order to understand quantitatively which model performs best, and which should be used going forward. Sec. 8.3 explains this process and the results in more detail.

8.1 Collection of Results

This section looks at how the results for this model were fit and collected, and how certain graphs may therefore be interpreted. The fundamental modelling loop involved selecting a range of seconds within a run, over which we wish to fit the model to the delayed electron data using the pS2 data. At its core, this model asks how well we can understand the nature of delayed electrons from solely

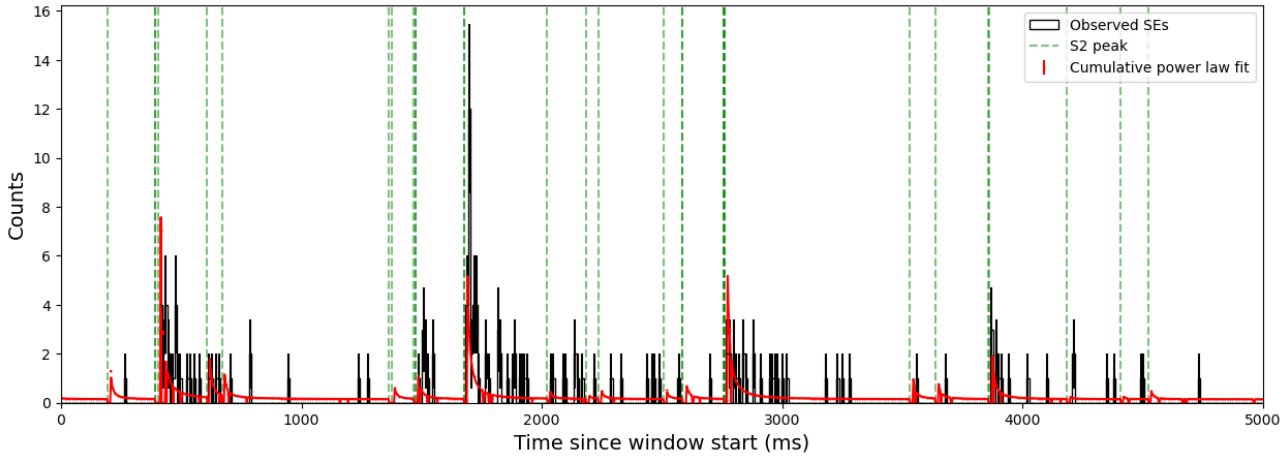


Figure 29: A typical cumulative power law fit (**Red**) for a 5 s window. The fit was performed on a 300 s region of data, and zoomed in to a 5 s window of that region for illustrative purposes. Each dotted green line represents the time of a pS2 in our 'new' selection. Dips in the fit of the model are the S1 and pS2 'dead-zone' regions, where it has been explicitly set to 0.

the presence of primary S2s; the background constant k provided by that fit gives us the extent to which our model fails - either due to data selection, missing features, or genuine light dark matter signals.

Within each fitting 'window', the model parameters were obtained using the CERN ROOT team's iminuit package, specifically the extended unbinned maximum likelihood cost function. It was decided in Ref. [1] that binning the data would lead to the loss of small-scale features in the model, due to the varied and uncertain nature of single-electron emission sources. The maximum likelihood estimation method takes a probability density function and fits it to the data in order to find the best-fit, maximised likelihood function.

To obtain the fitting plots, the data is binned at a resolution of 10 ms, and the error bars of each bin taken as the square root of the counts in each bin. This histogram is overlaid by the model parameter fit. A representative figure for this can be seen in Fig. 29.

To obtain the results discussed in Secs. 8.2.2 and 8.2.3, this modelling loop was ran for a variety of time windows, from 25 s up to 900 s. This upper limit was chosen as it amounted to $\sim \frac{1}{2}$ of a 30-minute run, so gave a representative overview of a large part of the data while also taking an acceptable time to fit. The time ranges were randomly selected across the 13 SR0 background runs listed in Appendix B, and multiple iterations were taken for each window size - with more for the smaller time ranges and fewer for the largest. The parameter values shown below are thus the median values for each window size, to best protect against outliers.

8.2 Discussion of Results

This section examines the results we obtain from the model variations, and what can be learned about both the model and data selection changes as a result. Sec. 8.2.1 shows plots like those of Fig. 29 in order to understand what can be learnt from the behaviour of the model in certain regions of data. Secs. 8.2.2 and 8.2.3 compare model results and parameters across all four variations, in order to understand which performs the best and in which metrics.

8.2.1 Findings from Model Fits

First we will highlight a few regions of fitted data, in order to understand where it is correctly mapping to features of delayed electron emission, and where it is not. For simplicity we will purely be looking at the new data selection fit to the model. There is little difference to the old selection in the basic workings of the model, and so the majority of small-scale features picked up or missed by the model are generally the same in both.

We can look at Fig. 29 as an example of a typical plot. This figure was obtained by running the model for a 300 s region of data to get a balanced fit and then zooming in on a 5 s window to see the smaller-scale features of the data. There is some clear success to this model in that it passes through the majority of error bars, as is the case for the similar plots in Ref. [1]. We can see the dips in the model where we have performed the prompt electron cuts after each S1 and pS2 and subsequently explicitly normalised the model to 0; this was not present in Ref. [1]. Finally, the largest features of the data, in the clusters of delayed electron signals, are being picked up by the model, as we would expect; the clearest example of where this fits well is over the clusters just before 3000 ms into the window.

However, we can also see clear problems with the model in Fig. 29. For example, just before 2000 ms there exists a large cluster of delayed electrons following a pS2, however the red line of the model does not match the count of electrons in this cluster. We also occasionally see instances where, despite a pS2 being indicated, not as many delayed electrons are present in the data as expected, for example with the apparent double-scatter at $\sim$ 1400 ms. This is not necessarily a problem: the nature of delayed electron emission is that the associated delayed electrons from this event could be released much later instead. It is also possible, for example if the interaction happened high in the TPC, that no or very few electrons were lost from the drift cloud of the interaction. And thus we may only see a weak delayed electron train, if any at all.

However, seeing these features too often in these plots could point to other factors with our model. The initial assumption prior to selecting the data was that any genuine pS2 (i.e. a GXe signal with

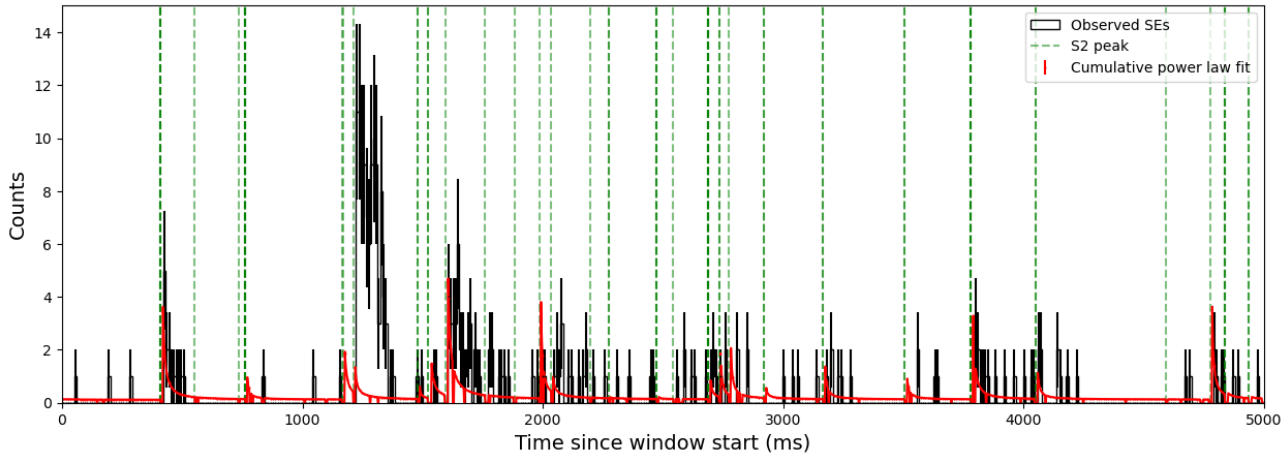


Figure 30: A cumulative power law fit over 300 s, zoomed in to a 5 s region of that data. A large burst of delayed electrons can be seen just after 1000 ms, a clear mis-fit with the model.

a matching S1 - detected or not - which was directly caused by a scatter in the LXe) would have the possibility of being correlated to the emission of delayed electrons. Seeing no delayed electrons after a pS2, and having a low pS2 to delayed electron ratio in the dataset more generally, may indicate that our selection of pS2s was too generous, and in fact not all of them contribute delayed electrons. One anomalous signal mentioned previously was that of e^--bursts, which are known to be as small as $10e^-$ [10]. It is almost certain that a population of these is included in both selections, and moreso in the new pS2 population, and could thus be flagging in the model as pS2s which should be contributing a certain population of delayed electrons, but instead do not.

Another problem can be seen in Fig. 30, where there is a large spike of electrons just after 1000 ms, which is not picked up at all by the model. This indicates a region in the data where the size of the pS2s does not match the amount of delayed electrons actually being seen immediately following it. The proximity of this burst to the pS2 immediately before it indicates that this could be a particularly extended prompt electron region. More likely, this may be an e^--burst, or series of them, which has been broken up and partly tagged as a pS2 by our processor, and the rest assigned as a small-S2. Some of that then gets cut by our prompt-electron cut, but e^--bursts, as mentioned, can be many drift times in width, meaning that the remaining part of it is then counted as delayed electrons in the histogram. This would also explain why the size of the pS2, and therefore the model, is not proportional to the amount of the delayed electrons seen immediately following it.

8.2.2 Total Electron Rates

Another way we can visualise the changes brought about to the model is in the total electron rate calculations made by our model, vs the real amount of electron signals being fed in (without electron counting). As mentioned in Sec. 5.2, the prefactor of $C \cdot A_p^{\alpha} W_p^{\beta}$ from Eq. 2, once fit by the model to

a time range, can provide a prediction as to how many delayed electrons each pS2 in that window likely contributed. The identification of unlikely excesses in this rate is one possible way in which this model may be used to aid in the search for light dark matter; however, this avenue was not explored for the model in this work. Instead, here, we use the total electron rate given by the model fit in a given time window and compare it to the true empirical amount of electrons in that window, as given by our selection methods. It is important to note that the two of these are formed separately: the data selection was not performed with the mechanics of the model in mind, or vice versa. The closeness of these two values, therefore, can tell us about the veracity of our model, and show if the changes made become unjustified.

This total predictive number of electrons from the model is given by:

$$N_{\text{model}} = kt \; + \; \sum_{p}^{P} C \; \cdot \; A_p^{\alpha} W_p^{\beta} \tag{5}$$

where k is the rate of background electrons per ms, multiplied across the live-time t of the window to get an estimated background; C is a constant given by the likelihood fit to scale to the sample size; and P is all primary S2s in the time window given by t.

We compare this value for each iteration in each window to the empirical amount of delayed electrons present in the dataset within that relevant time range, $N_{\text{empirical}}$, and calculate the absolute difference between the two rates: $\mid N_{\text{model}} \; - \; N_{\text{empirical}} \mid$. Examining these values for each window size and model variation gives us Fig. 31, where the absolute difference between the model predicted rate and the actual total rate has been calculated for each iteration of randomly selected time ranges. From this we take the median value for each window size. The error bars are given by the 16th and 84th percentiles around the median error, corresponding to $\mu \; \pm \; 1\sigma$.

In fact, Fig. 31 shows remarkable agreement in all models with the original data. The discrepancy increases with window size, which is expected as longer windows increase the likelihood of anomalous data warping the expected electron rates. However, this discrepancy is below 1 even in the errors for all models, showing that all model variations are estimating the correct total number of electrons with respect to the data for all window lengths.

In reality, this is mainly a result of how flexible the model is in its fitting. However, this figure is a good indicator that none of the changes to the model veer away from the median set by the old selection method. The introduction of truly anomalous or inaccurate data, or in unfounded model extensions, would introduce far larger discrepancies from the model struggling to accommodate. From this we may feel that the new selection, radial factor, and electron-counting would all be valid extensions to the model, at least in principle. Whether this is actually the case is discussed in more

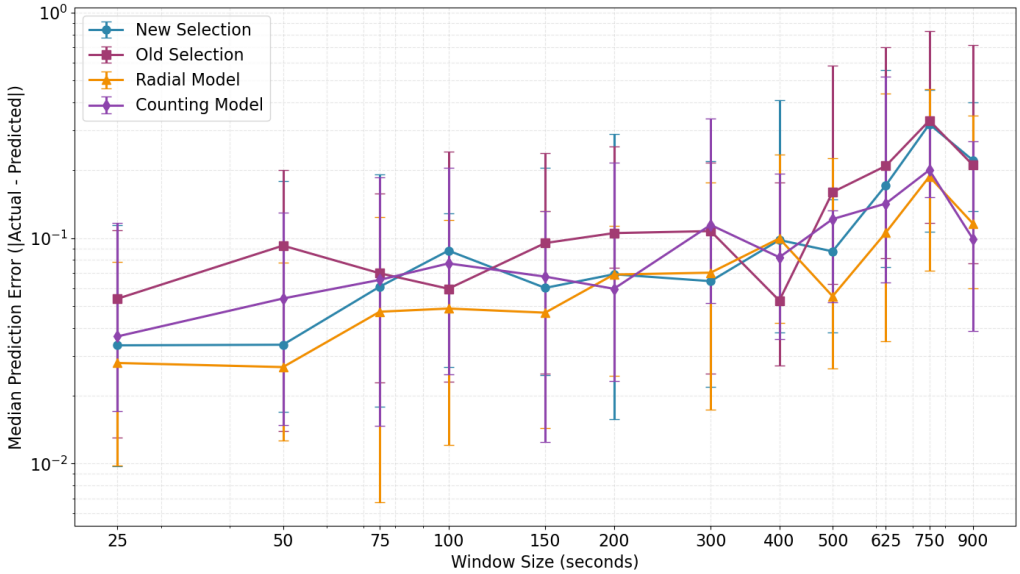


Figure 31: Median absolute difference between the true amount of electron signals in an analysis window vs the predicted total amount of electrons from the model, as fit to longer time ranges. This means that for e.g. 200 s, the model was fit to a time region 200 s long within a run. Note the y-axis is on a logarithmic scale.

detail in Sec. 8.3.

8.2.3 Model Parameters

As we fit our model to different lengths of time across a run, we can get a picture of how each variation changes as the fit length increases. The longest fitting times are of the most interest in terms of the parameters, as then variation in the data becomes more smoothed out, helping to mitigate -albeit not totally remove - the effect of any anomalous data in either selection method. This is why many windows were taken and averaged over for each size of window, however with fewer iterations for longer time spans.

The first parameter of interest is that of the power law exponent itself, n. Fig. 32a shows a large difference in the median n values of the old and new selections. The likely reason for this is that the old selection dataset contains far fewer pS2 signals as compared to the new data selection. When a pS2 is missing from the data (relative to the true physics of the detector), the effect on the delayed electron dataset is twofold:

- Some amount of genuine prompt electrons don't get cut from the dataset as they should, instead appearing as delayed electrons
- These, as well as the genuine delayed electrons caused by missing pS2s, become mis-assigned

to an earlier prior pS2.

Both of these cause a bias in the model towards later-time electrons, causing longer and heavier (i.e. more populated) power law tails. Fewer pS2s increases the average time between pS2s, and the model accommodates for the apparent increased weight of power law trains by adjusting the exponent n to be lower. The skew that can be introduced into delayed electron dataset from all this also leads to some signals being placed at 'unlikely' times relative to the pS2s in the dataset, causing the model to instead view it as a possible source of background. Due to the minimal likelihood fitting of the model, it will often compromise by stretching the parameters in both directions to fit the data: raising the background constant k while lowering the exponent n. The effect on k can be seen in Fig. 32b.

In contrast, the new selection identifies a larger quantity of pS2s, resulting in a larger proportion of delayed electrons being assigned to these, rather than to long tails, or to the background rate. As more delayed electrons are now happening closer to the times of pS2s, this has the effect of steepening the power law exponent. In this way, there may be nothing inherently negative in the new data selection having a higher power law exponent than the old selection, even though on the surface it seems to be further from the result of prior studies (between 1.0 and 1.4 - see Sec. 5.1 for more reference).

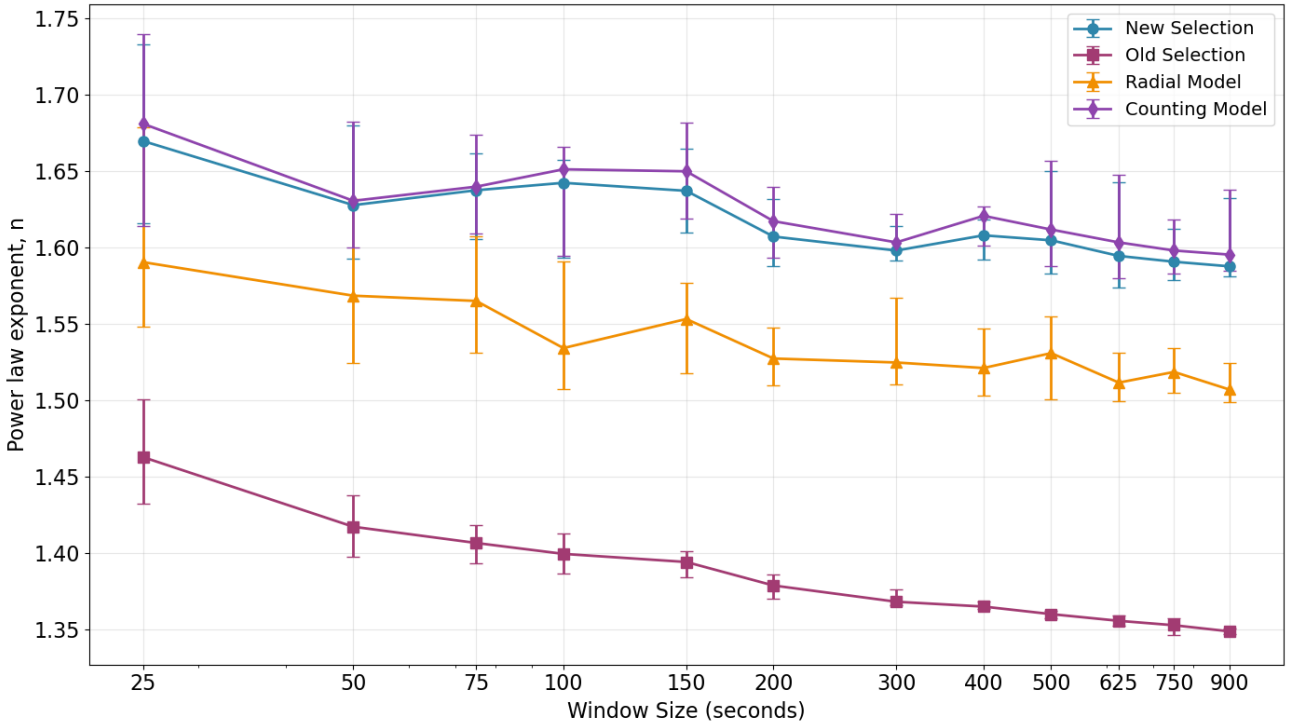
The radial model finds a mid-point in these two exponent values. Although the fiducial volume cut removes a significant portion of the delayed electron population, the associated radial factor also lowers the effective density of pS2s, making it so that pS2s near the edges contribute nearly nothing to the rate. The model flattens the power law exponent slightly to account for this rate, however, importantly, does not raise the background constant to compensate. On the contrary, it lowers by over a factor of 3 compared to the other model versions, as can be seen in Fig. 32b.

The n and k values for the 900 s window are given in Tables 1 and 2, respectively.

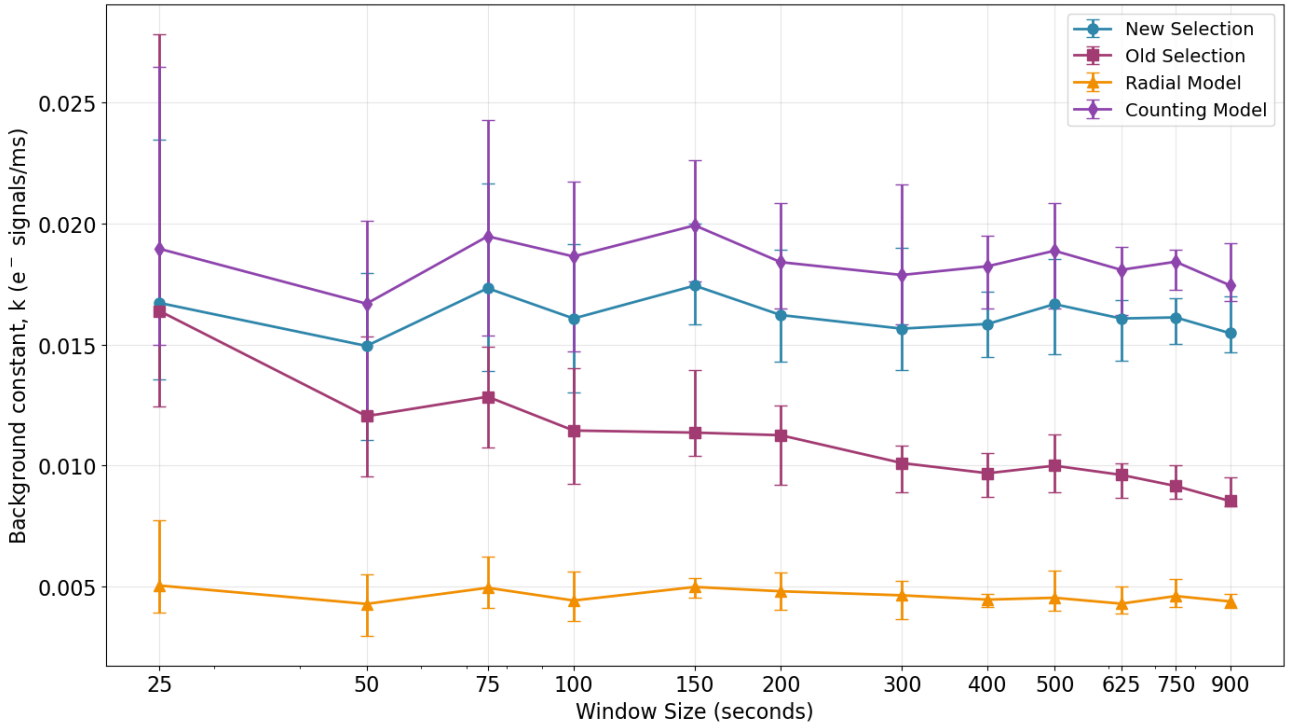
Table 1: Summary of fitted n parameters for the four model variations for 900 s of data.

Model version	$n \pm 1\sigma$
Old selection	1.349 ± 0.002
New selection	1.588 ± 0.025
Radial model	1.507 ± 0.013
Counting model	1.595 ± 0.026

The other two parameters of interest in the model are those of α and β, which tell us the extent to which the model values and depends upon the pS2 areas and width, respectively. A value of 0 for one or the other would indicate that there is no dependency on that value in the cause and correlation to delayed electron emission. It is good, then, to see that both values are not 0. As was the case for

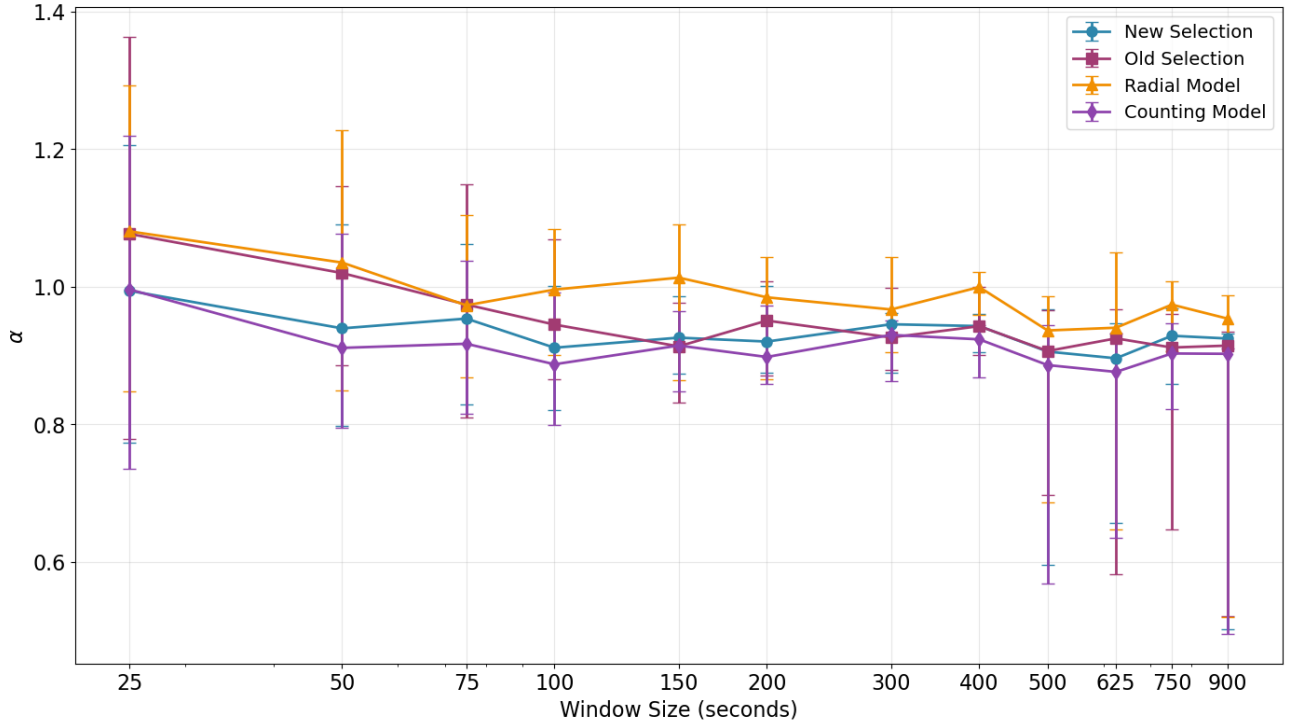


(a) n vs window size

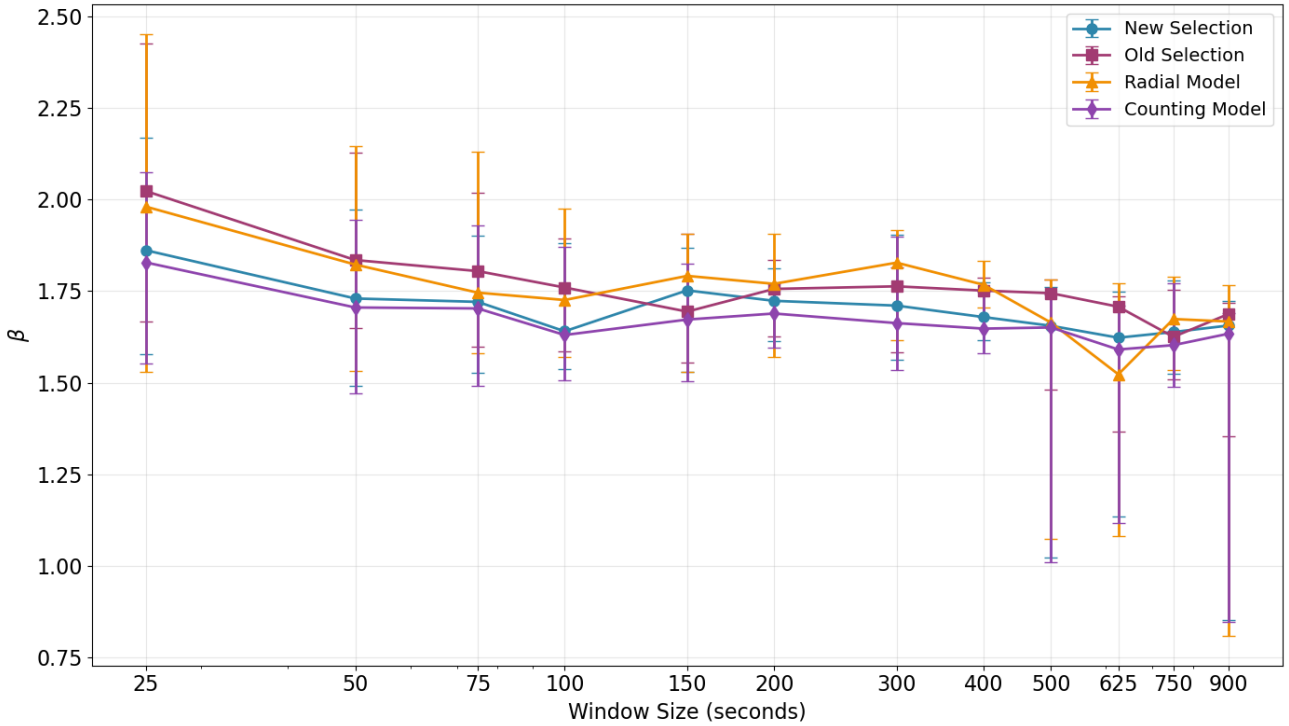


(b) k vs window size

Figure 32: Median power law exponent n (**Top**), and background constant k (**Bottom**), for each model variation, as fit to different time ranges across 13 SR0 runs. The error bars are again given by the 16th and 84th percentiles, $\mu \pm 1\sigma$.



(a) Parameter α vs window size



(b) Parameter β vs window size

Figure 33: Median values for exponents α (**Top**), and β (**Bottom**) from Eq. 1, for each model variation, as fit to different time ranges across 13 SR0 runs. The error bars are again given by the 16th and 84th percentiles, or $\mu \pm 1\sigma$.

Table 2: Summary of fitted k parameters for the four model variations for 900 s of data.

Model version	$k \pm 1\sigma$
Old selection	0.009 ± 0.0012
New selection	0.015 ± 0.0006
Radial model	0.004 ± 0.0003
Counting model	0.017 ± 0.0012

including them initially, there is physical justification for there being dependence on both these values; pS2 area indicates quantity of electrons produced in the interaction, and therefore proportionally the amount of delayed electrons; width is an approximate surrogate for interaction depth, a correlated factor to both the probability of electrons being captured by impurities, or the possibility of some thermalising and getting trapped at the LXe surface. Both of these were mentioned in Sec. 4.1. In this way we can again feel that no model variation has been so unjustified as to completely break from our understanding of the physics at play here.

Looking closer, we can see that all variations conclude on very similar results for both parameters: a heightened emphasis on width via a raised parameter β, and a slightly sub-linear reliance on area via the parameter α. This may tell us slightly about the physics, that the width of a pS2 plays slightly more role in the amount of delayed electrons associated with it than area does. However, it is also important to note that area ranges from $\mathcal{O}(100)$ PE to $\mathcal{O}$ (10,000,000) PE in the new selection, whereas width spans $\mathcal{O}(100)$ ns to only $\mathcal{O}(10{,}000)$ ns. When we adjust to calculate the effective weights of each of these parameters in the model given these differing ranges, we find that the contributions from each parameter to the scaling of the delayed electron rate are approximately equal across any given fitting range, with area just slightly more important. At best, this may indicate that the width of the signal, and therefore the depth of the interaction, may be slightly less important in the delayed electron creation and emission process than the signal area. This does not necessarily point to one hypothesis over the other, however, just that the amount of ionised electrons being created from an interaction plays a slightly larger role in the magnitude of delayed electron trains than the depth of that interaction and of the probability of those electrons encountering an impurity.

8.3 Evaluation of the Models

8.3.1 Implementation of Goodness-of-fit Testing

Although the results above lead to interesting conclusions on the model and the effect of the changes, we also wish to quantitatively evaluate how well each model variation reproduces the empirical

delayed electron distribution. For this we explored goodness-of-fit testing on the model variations. This would hopefully help us to understand whether the changes to the model and data selection brought about genuine improvement in the model, beyond eye-test analysis.

We implemented a non-parametric permutation-based goodness-of-fit test on the same regions of data analysed to collect the model results previously, using the 'Wasserstein Distance' as the metric for our analysis. The Wasserstein metric is also called 'Earth-mover's distance', and can be thought of as the cost of moving from one CDF to another, as if they were piles of earth to be shifted. This was chosen over metrics such as the Kolmogorov–Smirnov test - which measures the single largest distance between two distributions - as it allows for a more holistic view of the region of data being fit, rather than measuring it by its worst parts. A more 'typical' metric of the χ^2 test was also not applicable, as it requires binned data under a Gaussian distribution, whereas these fits are drawn from unbinned data under a physics model which is inherently continuous. More about the computation of the Wasserstein distance and justification for its use can be found at Ref. [79].

The goodness-of-fit test was applied for each model variation (old selection; new selection; radial model extension under new selection; and electron-counting model extension under new selection), for each iteration of randomly selected start times in each window size, from 25 s up to 900 s. For each iteration, the procedure was as follows:

1. Calculate the observed statistic: We compute the model CDF from the differential model rate (the red model line in e.g. Fig. 29). This comes from the best-fit parameters obtained in Sec. 8.1. This is then compared to the empirical CDF from the delayed electron data in the same region. From this we calculate our Wasserstein distance, W_{obs}.

2. Construct the null distribution: We wish to understand if the observed statistic is typical or anomalous, so must sample from our model to understand the typical model distribution. We do this N_{perm} times.

    2.1. Sample a Poisson number of electrons, $N_{sim} \sim \text{Poisson}(C \cdot A_p^{\alpha} W_p^{\beta})$. This is combined with $f(r)$ in the case of the radial model.

    2.2. N_{sim} simulated electron times are generated from the model, accounting for the dead-zone regions after S1s and pS2s.

    2.3. The value $k \cdot t_{livetime}$, where $t_{livetime}$ corresponds to the length of the analysis window, gives us a total amount of background electrons to be added to the simulated sample.

3. We compare each simulated distribution of delayed electron times to our model CDF, to get N_{perm} Wasserstein metrics.

4. The p-value is the fraction of those simulated metrics which are $\geq W_{obs}$

5. The z-score is obtained using the mean and standard deviation of the null distribution, $z = (W_{obs} - \mu_{null})/\sigma_{null}$. This is a measurement of how far the observed statistic is from the simulated distribution.

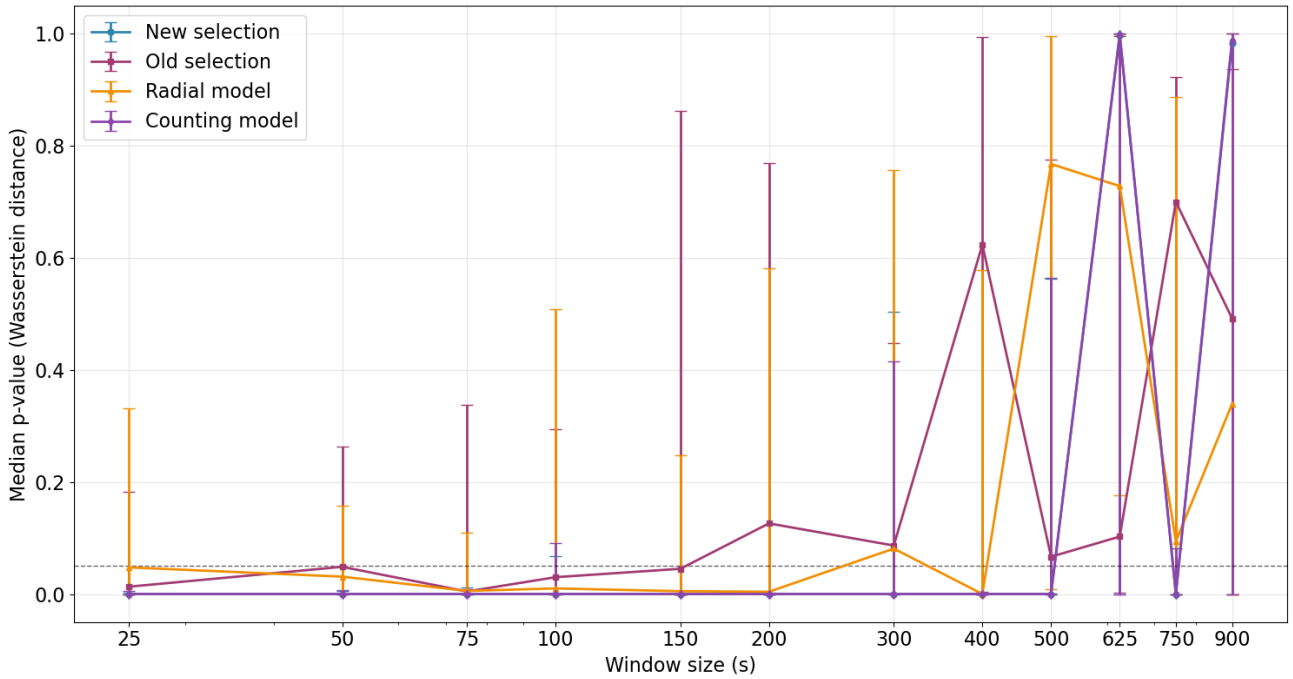
8.3.2 Goodness-of-fit Results

As stated, goodness-of-fit testing was run for every random window iteration and size, for every model variation. This allowed us to obtain the p-value and z-score plots shown in Fig. 34.

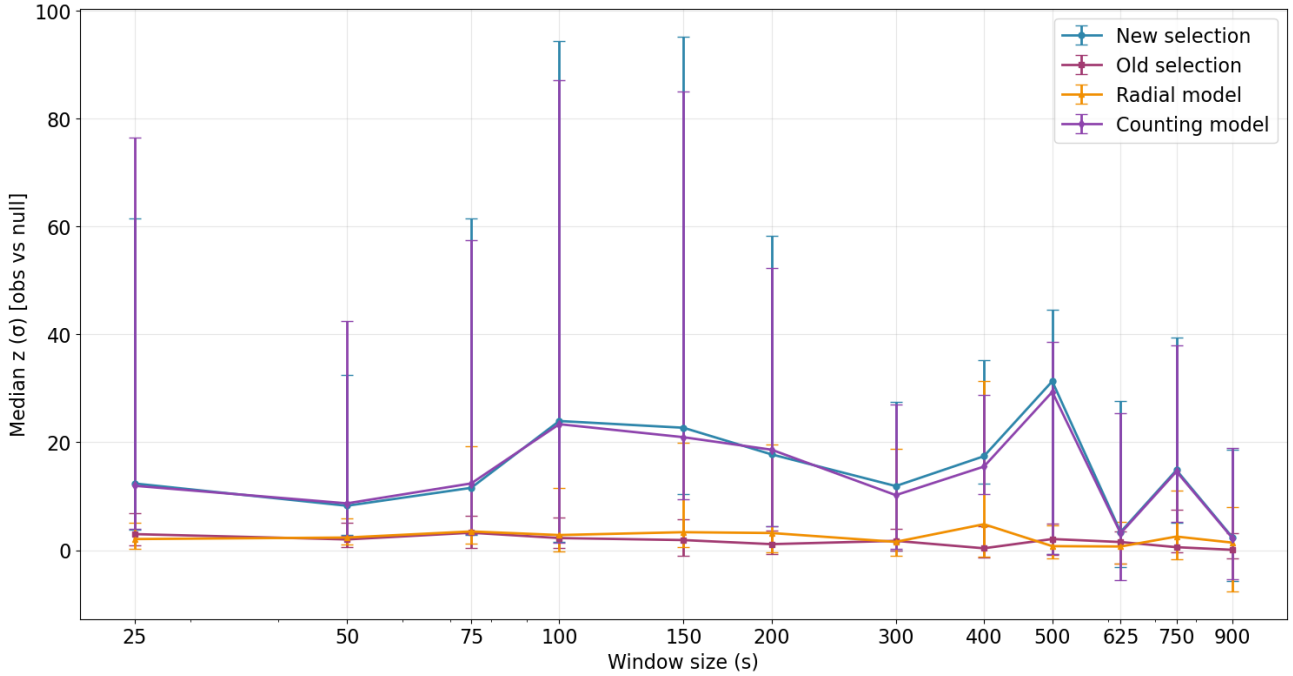
The results from the tests are difficult to draw conclusive results from. The radial model and the old selection both seem to perform better at longer fitting times as compared to at shorter ones, however even at low p-values seem to maintain lower z-scores than the new selection or electron-counting models, showing better similarity in the distributions. Indeed, even in shorter fitting windows, the upper error bars on the p-values for the old selection and radial models indicate that some proportion of the fits do match reasonably well to the data. What is clear from these fits is that the model is highly susceptible to variance in the data, especially when anomalies such as e^--bursts may be involved. The old selection and radial models thus perform better in this goodness-of-fit test as compared to the new selection and electron-counting model as they do more to limit the variance in the data. This is done in the old selection by having more strict boundaries on pS2 and delayed electron populations as compared to the new selection; the radial model limits this variance by reducing the effect of any anomalies which may be happening at the edges of the detector.

The new selection and electron-counting models have extremely similar results - it is not possible to separate the two in order to understand if electron-counting 'improved' the model in any way. It is clear that the effect is extremely minimal: $1e^-$ signals make up approximately 92% of the delayed electron dataset, $2e^-$ signals a further 6%, making the number of signals which are affected by this electron-counting change to the model extremely small. Whether this would be a justified addition to the model requires further analysis, therefore.

The susceptibility to the variance in data is clear when looking at the p-values for longer fitted regions of these two model variations, where they vary between being seemingly extremely accurate, to completely wrong. The large error bars also indicate that some fits within these windows vary in the exact opposite ways. As the fitting window increases, the model CDF accumulates structure which, if it was perfect, would match exactly that of the delayed electron distribution in the detector. If within the fitting region there are sections of data, however long, which consist of particularly sparse amounts of delayed electron signals; over-dense in pS2 signals; or perhaps contain many anomalous signals such as e^--bursts, this will be recorded in the structure of the empirical CDF in that fitting window, but not in the model, which does not have the flexibility or parameters to account for this.



(a) p-values vs fitting length



(b) z-scores vs fitting length

Figure 34: Plots of the median p-values (**Top**) and z-scores (**Bottom**) for each model variation vs length of fitting window in seconds. The error bars are given by the 16th and 84th percentiles, $\mu \pm 1\sigma$

These plots show that in the new selection methodology (and by extension, the electron-counting model), a wide range of data is introduced, not all of which the model can adequately account for. This thus results in some fits closely matching the data (where we may see mostly a typical distribution of genuine pS2s along with a well-proportioned distribution of delayed electrons), and some fits with large discrepancies (where e.g. e^--bursts are counted as pS2s, making the model expect a large number of delayed electrons where empirically there are none). By contrast, the old selection and radial models reduce this variance as mentioned, and thus also experience more stable results over longer fitting windows.

An improvement to the overall model would be to implement stricter quality control on the new selection, especially in that of pS2s, while also adding the radial factor. It is important to note that while the old selection method may appear to be a good fit, it contains a caveat in that it is a good fit only of the data it sees. In a way this is an obvious statement, but should be spelled out. The new selection method sees far more data; this introduces more genuine pS2s and delayed electrons, but also more anomalies. Within the old selection, there is no way of knowing if each pS2 which enters the selection has all of its 'matched' (i.e. causally-related) delayed electrons, and vice versa. This problem is lessened in the new data selection by casting a wider net on the selections of both, however is let down in being unable to properly cut all anomalous data that introduces.

This, as well as the points in Sec. 6, should be a motivation for developing the new selection methodology to be more rigorous, as well as incorporating the radial model; applying the radial model to the old selection methodology would yield improved results also, but would be a greatly incomplete map of all delayed electron emission in the detector. Overall, this makes the radial model seemingly the best version of the model, and the one to continue with going forward should one look to continue improving this model.

9 Conclusion & Future Work

In this work, we have performed a comprehensive review of the literature around delayed electron emission within noble-liquid direct detection experiments, and in XENONnT in particular. Using the information from this, we motivated the use of a new type of model to better map this background. We then analysed the modes of data selection available to develop a more rigorous methodology. Finally, we proposed physically-justified extensions to this model, before testing them against one another and assessing their fitting to the data. This allowed for new insight into the nature of how we may model delayed electron emission more accurately.

This work gets some promising results, such as a lowering of the background single-electron emission rate to 5 Hz for the radial model, down significantly from 36.5 Hz in Ref. [1]; as well as learning more about the nature of anomalies in our data and how they affect the modelling, and therefore, their relation to delayed electron emission on a whole. The power law exponent, n, was found to be between 1.5 - 1.6 for the added models; this is slightly higher than previous studies (between 1.0 - 1.4), though only slightly higher than the original work of Ref. [1] (1.448). This is not a complete model, however, and there is much improvement which could be done in order to better model the background of delayed electron trains.

One such improvement would be to bring this model into a 3D-space by making it position-dependent, such that the electron rate becomes $R(t,\ x,\ y)$. Instead of just understanding the single-electron rate at a given point in time, we could map that to x-y location in the detector also, using the position-correlation findings derived previously. This change would ideally also be able to incorporate the positions of the two transverse wires running across the top of the TPC, as well as the hot-spot which is present, which we were not able to do here in the radial model.

Further, although the selection methodology developed here is a step in the right direction, it clearly introduces anomalies which the model is unable to fully deal with. A few known cuts which could be incorporated were already mentioned, and these would greatly improve the model. The most important thing is that any signals which would causally correlate to delayed electron emission as pS2s do be incorporated, as it is by these signals that our model maps the true data.

Finally, a better understanding of how well this model fits to the data, and more importantly how it fits within the context of other delayed electron emission studies, would be of great benefit to analysis. The fundamental differences in how these two different model-types are fit to the data makes comparisons between parameters difficult, or even impossible. This problem was similarly noted in Ref. [1], but not able to be solved in this work.

In furthering this model more we could contribute a great deal to the understanding of background

noise and single-electron emission in the XENONnT experiment, and aid more and more in the search for dark matter, especially in the low-mass regime.

References

[1] Keerthana Umesh. "Modelling the delayed single and few electron background in XENONnT". Supervisor: J. Aalbers. Deposited on 19 July 2024. Available online at https://fse.studenttheses.ub.rug.nl/33361/. Bachelor's thesis. Groningen, The Netherlands: Rijksuniversiteit Groningen, July 2024.

[2] M. Persic and P. Salucci. "The baryon content of the Universe". In: *Monthly Notices of the Royal Astronomical Society* 258.1 (Sept. 1992), 14P–18P. ISSN: 1365-2966. DOI: 10.1093/mnras/258.1.14p. URL: http://dx.doi.org/10.1093/mnras/258.1.14P.

[3] S. Navas et al. "Review of Particle Physics". In: *Phys. Rev. D* 110 (3 Aug. 2024). Sec. 27; 29.5.2, p. 030001. DOI: 10.1103/PhysRevD.110.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.110.030001.

[4] Gianfranco Bertone and Dan Hooper. "History of dark matter". In: *Reviews of Modern Physics* 90.4 (Oct. 2018). ISSN: 1539-0756. DOI: 10.1103/revmodphys.90.045002. URL: http://dx.doi.org/10.1103/RevModPhys.90.045002.

[5] A. Arbey and F. Mahmoudi. "Dark matter and the early Universe: A review". In: *Progress in Particle and Nuclear Physics* 119 (July 2021), p. 103865. ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2021.103865. URL: http://dx.doi.org/10.1016/j.ppnp.2021.103865.

[6] Carlos Pérez de los Heros. "Status, Challenges and Directions in Indirect Dark Matter Searches". In: *Symmetry* 12.10 (Oct. 2020), p. 1648. ISSN: 2073-8994. DOI: 10.3390/sym12101648. URL: http://dx.doi.org/10.3390/sym12101648.

[7] Marcin Misiaszek and Nicola Rossi. "Direct Detection of Dark Matter: A Critical Review". In: *Symmetry* 16.2 (2024). ISSN: 2073-8994. DOI: 10.3390/sym16020201. URL: https://www.mdpi.com/2073-8994/16/2/201.

[8] E. Aprile et al. "First Dark Matter Search Results from the XENON1T Experiment". In: *Phys. Rev. Lett.* 119 (18 Oct. 2017), p. 181301. DOI: 10.1103/PhysRevLett.119.181301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.181301.

[9] D. Yu. Akimov et al. "Observation of delayed electron emission in a two-phase liquid xenon detector". In: *JINST* 11.03 (2016), p. C03007. DOI: 10.1088/1748-0221/11/03/C03007.

[10] D. S. Akerib et al. "Investigation of background electron emission in the LUX detector". In: *Physical Review D* 102.9 (Nov. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.092004. URL: http://dx.doi.org/10.1103/PhysRevD.102.092004.

[11] E. Aprile et al. "Emission of single and few electrons in XENON1T and limits on light dark matter". In: *Physical Review D* 106.2 (July 2022). ISSN: 2470-0029. DOI: 10.1103/physrevd.106.022001. URL: http://dx.doi.org/10.1103/PhysRevD.106.022001.

[12] B. Edwards et al. "Measurement of single electron emission in two-phase xenon". In: *Astroparticle Physics* 30.2 (Sept. 2008), pp. 54–57. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2008.06.006. URL: http://dx.doi.org/10.1016/j.astropartphys.2008.06.006.

[13] A. A. Burenkov et al. "Detection of a single electron in xenon-based electroluminescent detectors". In: *Physics of Atomic Nuclei* 72.4 (Apr. 2009), pp. 653–661. DOI: 10.1134/S1063778809040103.

[14] P. Agnes et al. "Low-Mass Dark Matter Search with the DarkSide-50 Experiment". In: *Physical Review Letters* 121.8 (Aug. 2018). ISSN: 1079-7114. DOI: 10.1103/physrevlett.121.081307. URL: http://dx.doi.org/10.1103/PhysRevLett.121.081307.

[15] P. Agnes et al. "Characterization of spurious-electron signals in the double-phase argon TPC of the DarkSide-50 experiment". In: (July 2025). arXiv: 2507.23003 [hep-ex].

[16] J. C. Kapteyn. "First Attempt at a Theory of the Arrangement and Motion of the Sidereal System". In: *Astrophys. J.* 55 (1922), pp. 302–328. DOI: 10.1086/142670.

[17] J. H. Oort. "The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems". In: *Bulletin of the Astronomical Institutes of the Netherlands* 6 (Aug. 1932), p. 249.

[18] Fritz Zwicky. "Republication of: The redshift of extragalactic nebulae". In: *General Relativity and Gravitation* 41.1 (2009), pp. 207–224. DOI: 10.1007/s10714-008-0707-4.

[19] F. Zwicky. "On the Masses of Nebulae and of Clusters of Nebulae". In: *Astrophys. J.* 86 (1937), pp. 217–246. DOI: 10.1086/143864.

[20] Vera C. Rubin and W. Kent Ford. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *The Astrophysical Journal* 159 (1970), pp. 379–403. URL: https://api.semanticscholar.org/CorpusID:122756867.

[21] J. G. de Swart, G. Bertone, and J. van Dongen. "How dark matter came to matter". In: *Nature Astronomy* 1.3 (Mar. 2017). ISSN: 2397-3366. DOI: 10.1038/s41550-017-0059. URL: http://dx.doi.org/10.1038/s41550-017-0059.

[22] N. Aghanim et al. "Planck 2018 results: VI. Cosmological parameters". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201833910. URL: http://dx.doi.org/10.1051/0004-6361/201833910.

[23] E. L. Wright. "A Cosmology Calculator for the World Wide Web". In: *Publications of the Astronomical Society of the Pacific* 118.850 (Dec. 2006). Calculator available at: https://www.astro.ucla.edu/~wright/CosmoCalc.html, p. 1711. DOI: 10.1086/510102. URL: https://doi.org/10.1086/510102.

[24] Benjamin Audren et al. "Strongest model-independent bound on the lifetime of Dark Matter". In: *Journal of Cosmology and Astroparticle Physics* 2014.12 (Dec. 2014), pp. 028–028. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2014/12/028. URL: http://dx.doi.org/10.1088/1475-7516/2014/12/028.

[25] Pratika Dayal. "Early galaxy formation and its large-scale effects". In: *Proceedings of the International Astronomical Union* 15.S352 (June 2019), pp. 43–43. ISSN: 1743-9221. DOI: 10.1017/s1743921320001106. URL: http://dx.doi.org/10.1017/S1743921320001106.

[26] cajohare. *HowToMakeAPlot: repository for plotting examples*. https://github.com/cajohare/HowToMakeAPlot. Accessed: 2025-10-06. 2025.

[27] Andrew Gould et al. "Neuton stars: Graveyard of charged dark matter". In: *Physics Letters B* 238.2 (1990), pp. 337–343. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(90)91745-W. URL: https://www.sciencedirect.com/science/article/pii/037026939091745W.

[28] Albert De Roeck. "Dark matter searches at accelerators". In: *Nuclear Physics B* 1003 (2024). Special Issue of Nobel Symposium 182 on Dark Matter, p. 116480. ISSN: 0550-3213. DOI: https://doi.org/10.1016/j.nuclphysb.2024.116480. URL: https://www.sciencedirect.com/science/article/pii/S0550321324000464.

[29] E. Aprile et al. "Projected WIMP sensitivity of the XENONnT dark matter experiment". In: *Journal of Cosmology and Astroparticle Physics* 2020.11 (Nov. 2020), p. 031. DOI: 10.1088/1475-7516/2020/11/031. URL: https://doi.org/10.1088/1475-7516/2020/11/031.

[30] E. Aprile et al. "The XENONnT dark matter experiment". In: *Eur. Phys. J. C* 84.8 (2024), p. 784. DOI: 10.1140/epjc/s10052-024-12982-5. arXiv: 2402.10446 [physics.ins-det].

[31] J Aalbers et al. "A next-generation liquid xenon observatory for dark matter and neutrino physics". In: *Journal of Physics G: Nuclear and Particle Physics* 50.1 (Dec. 2022), p. 013001. DOI: 10.1088/1361-6471/ac841a. URL: https://dx.doi.org/10.1088/1361-6471/ac841a.

[32] E. Aprile et al. "Double-weak decays of ^{124}Xe and ^{136}Xe in the XENON1T and XENONnT experiments". In: *Phys. Rev. C* 106 (2 Aug. 2022), p. 024328. DOI: 10.1103/PhysRevC.106.024328. URL: https://link.aps.org/doi/10.1103/PhysRevC.106.024328.

[33] E. Aprile and T. Doke. "Liquid xenon detectors for particle physics and astrophysics". In: *Rev. Mod. Phys.* 82 (3 July 2010), pp. 2053–2097. DOI: 10.1103/RevModPhys.82.2053. URL: https://link.aps.org/doi/10.1103/RevModPhys.82.2053.

[34] Matthew C. Digman et al. "Not as big as a barn: Upper bounds on dark matter-nucleus cross sections". In: *Phys. Rev. D* 100 (6 Sept. 2019), p. 063013. DOI: 10.1103/PhysRevD.100.063013. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.063013.

[35] J.L. Autran et al. "Characterization of atmospheric muons at sea level using a cosmic ray telescope". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 903 (2018), pp. 77–84. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2018.06.038. URL: https://www.sciencedirect.com/science/article/pii/S0168900218307599.

[36] D.-M. Mei and A. Hime. "Muon-induced background study for underground laboratories". In: *Physical Review D* 73.5 (Mar. 2006). ISSN: 1550-2368. DOI: 10.1103/physrevd.73.053004. URL: http://dx.doi.org/10.1103/PhysRevD.73.053004.

[37] Alessandro Bettini. *The World underground scientific facilities: A Compendium*. Dec. 2007. arXiv: 0712.1051 [hep-ex].

[38] E. Aprile et al. "The neutron veto of the XENONnT experiment: results with demineralized water". In: *Eur. Phys. J. C* 85.6 (2025), p. 695. DOI: 10.1140/epjc/s10052-025-14105-0. arXiv: 2412.05264 [physics.ins-det].

[39] Joran Remco Angevaare. "First WIMP results of XENONnT and its signal reconstruction". PhD thesis. Amsterdam U., 2023.

[40] E. Aprile et al. "Search for New Physics in Electronic Recoil Data from XENONnT". In: *Physical Review Letters* 129.16 (Oct. 2022). ISSN: 1079-7114. DOI: 10.1103/physrevlett.129.161805. URL: http://dx.doi.org/10.1103/PhysRevLett.129.161805.

[41] E. Aprile et al. "The triggerless data acquisition system of the XENONnT experiment". In: *Journal of Instrumentation* 18.07 (July 2023), P07054. ISSN: 1748-0221. DOI: 10.1088/1748-0221/18/07/p07054. URL: http://dx.doi.org/10.1088/1748-0221/18/07/P07054.

[42] XENONnT Collaboration. *straxen: Streaming analysis framework for XENONnT*. Latest version: 3.2.3 (accessed September 17, 2025). URL: https://github.com/XENONnT/straxen.

[43] XENON Collaboration et al. *XENONnT Analysis: Signal Reconstruction, Calibration and Event Selection*. 2024. arXiv: 2409.08778 [hep-ex]. URL: https://arxiv.org/abs/2409.08778.

[44] E. Aprile et al. "First Search for Light Dark Matter in the Neutrino Fog with XENONnT". In: *Physical Review Letters* 134.11 (Mar. 2025). ISSN: 1079-7114. DOI: 10.1103/physrevlett.134.111802. URL: http://dx.doi.org/10.1103/PhysRevLett.134.111802.

[45] E. Aprile et al. "Search for Light Dark Matter Interactions Enhanced by the Migdal Effect or Bremsstrahlung in XENON1T". In: *Physical Review Letters* 123.24 (Dec. 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.123.241803. URL: http://dx.doi.org/10.1103/PhysRevLett.123.241803.

[46] J. Xu et al. "Search for the Migdal effect in liquid xenon with keV-level nuclear recoils". In: *Phys. Rev. D* 109 (5 Mar. 2024), p. L051101. DOI: 10.1103/PhysRevD.109.L051101. URL: https://link.aps.org/doi/10.1103/PhysRevD.109.L051101.

[47] P. Agnes et al. "Search for Dark-Matter–Nucleon Interactions via Migdal Effect with DarkSide-50". In: *Phys. Rev. Lett.* 130 (10 Mar. 2023), p. 101001. DOI: 10.1103/PhysRevLett.130.101001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.130.101001.

[48] E. Aprile et al. "Light Dark Matter Search with Ionization Signals in XENON1T". In: *Physical Review Letters* 123.25 (Dec. 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.123.251801. URL: http://dx.doi.org/10.1103/PhysRevLett.123.251801.

[49] E. Aprile et al. "Search for Light Dark Matter in Low-Energy Ionization Signals from XENONnT". In: *Physical Review Letters* 134.16 (Apr. 2025). ISSN: 1079-7114. DOI: 10.1103/physrevlett.134.161004. URL: http://dx.doi.org/10.1103/PhysRevLett.134.161004.

[50] J. Angle et al. "Search for Light Dark Matter in XENON10 Data". In: *Physical Review Letters* 107.5 (July 2011). ISSN: 1079-7114. DOI: 10.1103/physrevlett.107.051301. URL: http://dx.doi.org/10.1103/PhysRevLett.107.051301.

[51] E Aprile et al. "Observation and applications of single-electron charge signals in the XENON100 experiment". In: *Journal of Physics G: Nuclear and Particle Physics* 41.3 (Feb. 2014), p. 035201. ISSN: 1361-6471. DOI: 10.1088/0954-3899/41/3/035201. URL: http://dx.doi.org/10.1088/0954-3899/41/3/035201.

[52] A. Kopec et al. "Correlated single- and few-electron backgrounds milliseconds after interactions in dual-phase liquid xenon time projection chambers". In: *Journal of Instrumentation* 16.07 (July 2021), P07014. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/07/p07014. URL: http://dx.doi.org/10.1088/1748-0221/16/07/P07014.

[53] P. Sorensen and K. Kamdin. "Two distinct components of the delayed single electron noise in liquid xenon emission detectors". In: *Journal of Instrumentation* 13.02 (Feb. 2018), P02032–P02032. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/02/p02032. URL: http://dx.doi.org/10.1088/1748-0221/13/02/P02032.

[54] A. Tomás et al. "Study and mitigation of spurious electron emission from cathodic wires in noble liquid time projection chambers". In: *Astroparticle Physics* 103 (Dec. 2018), pp. 49–61. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2018.07.001. URL: http://dx.doi.org/10.1016/j.astropartphys.2018.07.001.

[55] Shinzou Kubota et al. "Dynamical behavior of free electrons in the recombination process in liquid argon, krypton, and xenon". In: *Phys. Rev. B* 20 (8 Oct. 1979), pp. 3486–3496. DOI: 10.1103/PhysRevB.20.3486. URL: https://link.aps.org/doi/10.1103/PhysRevB.20.3486.

[56] T. A. King and R. Voltz. "The time dependence of scintillation intensity in aromatic materials". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 289.1418 (1966), pp. 424–439. DOI: 10.1098/rspa.1966.0021. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1966.0021. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1966.0021.

[57] E. Bodnia et al. "The electric field dependence of single electron emission in the PIXeY two-phase xenon detector". In: *Journal of Instrumentation* 16.12 (Dec. 2021), P12015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/16/12/p12015. URL: http://dx.doi.org/10.1088/1748-0221/16/12/P12015.

[58] Adam Bailey. *Dark matter searches and study of electrode design in LUX and LZ*. Aug. 2016. DOI: https://doi.org/10.25560/41878. URL: http://hdl.handle.net/10044/1/41878.

[59] V Chepel and H Araújo. "Liquid noble gas detectors for low energy particle physics". In: *Journal of Instrumentation* 8.04 (Apr. 2013), R04001–R04001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/8/04/r04001. URL: http://dx.doi.org/10.1088/1748-0221/8/04/R04001.

[60] Peter Sorensen. *Electron train backgrounds in liquid xenon dark matter search detectors are indeed due to thermalization and trapping*. Feb. 2017. arXiv: 1702.04805 [physics.ins-det].

[61] Hans Dembinski and Piti Ongmongkolkul et al. "scikit-hep/iminuit". In: (Dec. 2020). Documentation available at https://scikit-hep.org/iminuit/index.html. DOI: 10.5281/zenodo.3949207. URL: https://doi.org/10.5281/zenodo.3949207.

[62] Peter Sorensen. "Anisotropic diffusion of electrons in liquid xenon with application to improving the sensitivity of direct dark matter searches". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 635.1 (Apr. 2011), pp. 41–43. ISSN: 0168-9002. DOI: 10.1016/j.nima.2011.01.089. URL: http://dx.doi.org/10.1016/j.nima.2011.01.089.

[63] Jianyu Long. *Peaks Subtype Study*. Internal note. XENON Collaboration, Aug. 2024. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=jlong:peak_subtyping_study.

[64] Jianyu Long. *Peaks Subtyping Dictionary*. Internal note. XENON Collaboration, Aug. 2024. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=jlong:peak_subtyping_dictionary.

[65] Jianyu Long. "Searching for Dark Matter with Ionization Signals in XENONnT". PhD thesis. Chicago U., 2024. DOI: 10.6082/uchicago.12344.

[66] P. Barrow et al. "Qualification tests of the R11410-21 photomultiplier tubes for the XENON1T detector". In: *Journal of Instrumentation* 12.01 (Jan. 2017), P01024–P01024. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/01/p01024. URL: http://dx.doi.org/10.1088/1748-0221/12/01/P01024.

[67] Luisa Hötzsch. *XENONnT SR0 PMT Afterpulses Summary*. Internal note. XENON Collaboration, May 2022. URL: xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenon1t:hoetzsch:afterpulses:sr0summary.

[68] Alexis Michel. *Overview on e-bursts*. Internal note. XENON Collaboration, Mar. 2024. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:alexis_joshua_paul_michel:overview_on_ebursts.

[69] Alexis Michel. *e-burst strax plugin*. Internal note. XENON Collaboration, June 2025. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:alexis_joshua_paul_michel:e_burst_plugin.

[70] Jianyu Long. *SR0 SE Study*. Internal note. XENON Collaboration, May 2022. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=jlong:sr0_update_se_selection_and_result.

[71] Jianyu Long. *Summary of Ne Counting Algorithm*. Internal note. XENON Collaboration, Apr. 2024. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=jlong:summary_of_ne_counting_algorithm.

[72] Robin Glade-Beucke. *S2-only: Primary S2 selection*. Internal note. XENON Collaboration, Feb. 2025. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:robingb:analysis_s2only:primary_s2_selection.

[73] Tianyu Zhu Yongyu Pan Shenyang Shi. *Summary of cathode background in S2-only analysis*. Internal note. XENON Collaboration, Mar. 2024. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:s2_only:cathode.

[74] Abigail Kopec Jianyu Long. *SR0 Single Electrons*. Internal note. XENON Collaboration, May 2022. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:analysis:sr0_se_conclusive_page.

[75] Noah Hood. *Origin of Position Uncorrelated DE*. Internal note. XENON Collaboration, Oct. 2025. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:noahhood:s2only:uncorrelated_from_mispairing.

[76] Abigail Kopec. *Electron Trains with the Hotspot*. Internal note. XENON Collaboration, Apr. 2021. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:abby:hotspot_etrains.

[77] Abigail Kopec. *XENONnT SR0 S2-only (x,y) Fiducialization*. Internal note. XENON Collaboration, July 2021. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:abby:fiducialization.

[78] Noah Hood. *Position Correlation of Delayed Electrons*. Internal note. XENON Collaboration, Sept. 2023. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:noahhood:s2only:position_correlation.

[79] Aaditya Ramdas, Nicolas Garcia, and Marco Cuturi. *On Wasserstein Two Sample Testing and Related Families of Nonparametric Tests*. 2015. arXiv: 1509.02237 [math.ST]. URL: https://arxiv.org/abs/1509.02237.

Appendices

A Code

Code & Data Availability Code is at: GitHub repository

B Run IDs

name	mode	end	source	start	livetime
031693	background_linked	2021-11-07 08:59:02.709000	none	2021-11-07 08:28:59.022000	00:30:03.687000
031691	background_linked	2021-11-07 07:57:14.305000	none	2021-11-07 07:27:10.918000	00:30:03.387000
031690	background_linked	2021-11-07 07:26:19.429000	none	2021-11-07 06:56:17.550000	00:30:01.879000
031688	background_linked	2021-11-07 06:24:29.650000	none	2021-11-07 05:54:25.366000	00:30:04.284000
031687	background_linked	2021-11-07 05:53:32.980000	none	2021-11-07 05:23:30.029000	00:30:02.951000
031686	background_linked	2021-11-07 05:22:38.083000	none	2021-11-07 04:52:37.468000	00:30:00.615000
031685	background_linked	2021-11-07 04:51:45.035000	none	2021-11-07 04:21:43.435000	00:30:01.600000
031684	background_linked	2021-11-07 04:20:51.818000	none	2021-11-07 03:50:51.205000	00:30:00.613000
031683	background_linked	2021-11-07 03:49:59.758000	none	2021-11-07 03:19:57.821000	00:30:01.937000
031682	background_linked	2021-11-07 03:19:05.850000	none	2021-11-07 02:49:05.523000	00:30:00.327000
031681	background_linked	2021-11-07 02:48:13.539000	none	2021-11-07 02:18:09.654000	00:30:03.885000
031680	background_linked	2021-11-07 02:17:17.299000	none	2021-11-07 01:47:14.472000	00:30:02.827000
031679	background_linked	2021-11-07 01:46:21.954000	none	2021-11-07 01:16:17.893000	00:30:04.061000