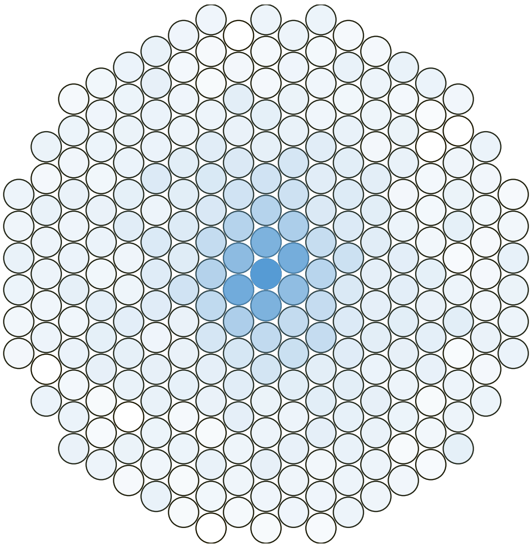
Data Driven Development of Light Collection Efficiency Maps for XENONnT

Felix Semler

Examiner : Jelle Aalbers
Second Examiner : Daan Meerburg



Abstract

The XENONnT experiment utilizes computed S2 light collection efficiencies (LCEs) maps to discriminate between signal and background events. Current Monte Carlo maps fail to fully account for detector physics, causing systematic biases in event reconstruction. This work develops a physics-constrained neural network parametrization of the expected detector response using an extended PMT signal likelihood function, with the full model trained on ^{83m}Kr calibration data. To avoid position-dependent biases, both event positions and model parameters are co-evolved under this likelihood framework. The resulting parametrization achieves a reduced $\chi^2 = 1.223$ compared to Monte Carlo's reduced $\chi^2 = 2.301$, while producing significantly more homogeneous event density distributions throughout the detector volume. This improved, physics-constrained LCE model provides a foundation for enhanced position reconstruction training and systematic studies in XENONnT. The constrained co-evolution methodology may demonstrate applicability to other dual-phase time projection chamber experiments.

Contents

1 Introduction

Despite the success of the Standard Model, it leaves fundamental questions unanswered, notably the nature of dark matter and the origin of the matter-antimatter asymmetry. Direct searches for dark matter aim to observe the minuscule coupling between dark matter and standard model matter particles, while neutrinoless double-beta decay experiments search for this rare nuclear process that could provide insight into the nature of the neutrino and potentially the origin of the matter-antimatter imbalance. Dual-phase xenon time projection chambers (TPCs) like the XENONnT detector are uniquely suited for both searches due to their exceptional sensitivity and background rejection capabilities. These detectors function as large-scale scintillation chambers, where particle interactions produce both scintillation and ionization signals, enabling precise energy reconstruction and powerful event discrimination. XENONnT combines a highly pure multi-ton xenon target with years of exposure time to probe new parameter space for WIMP-like dark matter interactions and rare nuclear decays, including the potential neutrinoless double-beta decay of ^{136}Xe, advancing our understanding of fundamental physics beyond the Standard Model.

The XENON detectors have evolved from the XENON10 prototype, using only 15 kg of xenon, through XENON100, and XENON1T the first ton scale ($\sim 2000\ kg$) TPC, to XENONnT, featuring 5.9 tonnes of active xenon target material, delivering over an order of magnitude greater sensitivity than its predecessor while reusing a significant portion of components. XENONnT aims to push the spin independent WIMP-nucleon cross-section down to $1.4 \times 10^{-48}\ cm^2$ for a 50 GeV/c^2 WIMP at (90% C.L. at 20 $t\ yr$) [1] and to reach a neutrinoless double-beta decay half-life sensitivity of $T_{1/2}^{^{136}Xe} > 2.1 \times 10^{25}\ yr$ (90% C.L. at 275 $kg\ yr$) [2].

In an event observation the ionization signal produces a spatial signal pattern, from which an inferred event position can be drawn. Using a Light Response Function (LRF) to predict an expected pattern for the same location the likelihood of the event can be determined and utilized for event rejection and further analysis. Bias in this process, and unmodeled statistical behavior can result in systematically biased energy reconstruction of events and subsequent analyses. The LRF maps are typically gathered from simulation which themselves are subject to mis-modeling, alternatively these expected patterns can be generated from data driven fitting to calibration sources injected into the detector, with the cleanest example being ^{83m}Kr, a metastable state of Krypton with a very distinctive decay chain.

Within this thesis, a data-driven Light Response Function will be developed that aims at parameterizing the individual contributions to the full response. This development progresses through multiple modeling approaches: starting with generic neural network architectures, advancing to constrained parametric models, and culminating in a comprehensive machine learning framework that integrates relevant detector effects. This progression allows identification of reconstruction biases and light collection biases within the detector, providing deeper understanding of individual components' contributions to the signal distribution. To facilitate this work, a previously developed likelihood function for signal collection is introduced and expanded upon for efficient usage within the XENONnT framework alongside the component functions aimed at parameterizing the LRF.

This thesis is organized as follows. Chapters 2-3 provide background on the physics searches with XENONnT and the XENONnT detector. Chapter 4 presents the first research contribution: studies of optical and electrical effects on light collection using Monte Carlo simulations. Chapter

5 provides theoretical development and implementation of a PMT signal likelihood framework with computational extensions on existing work. Chapter 6 introduces relevant machine learning methods.

The main contributions are in Chapters 7-9, which develop the LRF through progressively sophisticated approaches: a baseline neural network implementation (Chapter 7), a constrained parametric model (Chapter 8), and the primary contribution—a comprehensive machine learning architecture integrating relevant detector effects (Chapter 9). Chapter 10 provides an overview of all results and discusses future directions.

2 Physics with XENON

2.1 Dark Matter

Astronomical observations beginning in the early 1930s revealed discrepancies between observed galactic dynamics and predictions based on visible matter. Oort [3] found stellar velocities along the galactic plane that exceeded expected values based on luminous mass, while later Rubin [4] observed galaxy rotation curves that deviated significantly from Newtonian predictions. These and subsequent observations indicated the presence of substantial amounts of non-luminous matter throughout the universe.

This “dark matter” appears to interact only gravitationally with ordinary matter, is effectively non-luminous, and must be cold (non-relativistic) to account for observed structure formation. These findings launched a wealth of investigations into the origin of this phenomenon, leading to theories of exotic new types of matter.

From astronomical observations, a few key properties of the particles composing dark matter have emerged: they are effectively non-luminous, appear to interact only gravitationally, and due to the observed clumped density distribution must be cold (i.e., non-relativistic and able to form structures rather than streaming freely). Any non-gravitational interactions of these particles that might produce standard model messengers such as photons must be extremely weak to maintain low enough fluxes consistent with observations of galaxy cluster dynamics. Therefore, the coupling between dark matter and standard model matter must be very small, if not null.

Cold Dark Matter (CDM) refers to the class of non-relativistic dark matter that clusters under gravity and lacks significant non-gravitational interactions. The current cosmological paradigm, known as ΛCDM, combines this cold, collisionless dark matter with the cosmological constant Λ (representing dark energy) to describe the evolution and structure of the universe.

Observations of the cosmic microwave background, large-scale structures, and baryon acoustic oscillations strongly support this framework, indicating that 26.4% of the total mass-energy density of the universe corresponds to cold dark matter [5, Sec. 27]. Within this model cold dark matter provides the gravitational seeds for structure growth, allowing the formation of galaxies and clusters from tiny fluctuations observed in the early universe.

A leading candidate among CDM particles is Weakly Interacting Massive Particles (WIMPs), defined by having sufficient mass to become non-relativistic early enough to be classified as CDM, and by having small coupling constants to standard model particles.

Many WIMP candidates are assumed to be produced thermally in the early universe through a process dictated by temperature-dependent creation and annihilation rates. As the universe expanded and cooled, the annihilation rate eventually dropped below the Hubble expansion rate, causing particle numbers to cease decreasing through annihilation. This “freeze-out” process formed a constant co-moving density of dark matter particles. Remarkably, this freeze-out naturally produces the observed dark matter abundance when assuming masses on the electroweak scale ($\sim 10\ GeV$ to $1\ TeV$) with self-annihilation cross-sections on the order of $10^{-26}\ cm^3/s$, produce a relic abundance that matches the observed dark matter density. This remarkable coincidence between theoretically motivated parameters and cosmological observations is known as the “WIMP miracle”.

The ΛCDM paradigm is not without its challenges, particularly in predicting small-scale structure formation on scales < 1 Mpc, where nonlinear effects become dominant. Nevertheless, WIMPs remain a compelling candidate covering a large parameter space of theoretical models while providing concrete targets for both direct and indirect searches through their coupling with standard model matter. This makes the experimental search for WIMPs a particularly worthwhile scientific endeavor with current detector technology.

For comprehensive reviews of dark matter theory and detection strategies, see Bertone and Tait [6].

2.2 Neutrinoless double beta decay

While dark matter detection represents one primary physics goal for liquid xenon experiments, these detectors are also well-suited to search for other rare processes that probe physics beyond the Standard Model.

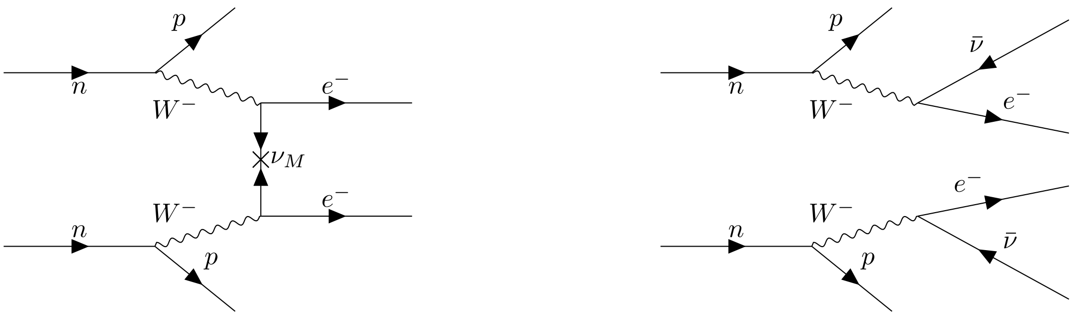
Neutrinoless double beta decay ($0\nu\beta\beta$) is a hypothetical lepton-number violating process in which two neutrons within a nucleus convert to two protons and emit only two electrons. Contrary to standard weak processes, no antineutrinos are emitted as final products (see Figure 1).

In standard lepton number conservation, an antineutrino must be emitted when an electron is produced. However, in this decay, the process can occur through various mechanisms, most commonly modeled as virtual neutrino exchange between the two decay vertices, such that no real neutrinos are produced. This process can only occur if neutrinos are Majorana particles—leptons that are their own antiparticles.

Within the perturbative Standard Model, no lepton number violating processes exist. Observation of this process would provide a very strong probe for physics beyond the Standard Model, with the measured half-life directly related to the effective Majorana neutrino mass through the decay rate formula. Under the assumption that the Majorana mass is the only source of lepton number violation at low energies, bounds on the effective Majorana mass have been set to $m_{ee} < 36 - 180\ meV$ from non-observation in xenon and germanium experiments [5, Sec. 14].

A number of isotopes are viable candidates for ($0\nu\beta\beta$) searches, but xenon and germanium offer the lowest uncertainties from nuclear matrix element calculations. The specialized xenon TPC EXO-200 established a half-life limit of $T_{1/2} > 3.5 \times 10^{25}$ yr for ^{136}Xe [5, Sec. 14]. Although not optimized for ($0\nu\beta\beta$), XENONnT using its natural xenon still expects $T^{0\nu}_{1/2} > 2.1 \times 10^{25}$ yr (275 kgyr of ^{136}Xe) [2], offering a valuable cross-check to dedicated experiments.

For detailed theoretical reviews of neutrinoless double-beta decay, see [7].



(a) $(0\nu\beta^-\beta^-)\,:\,\Delta L = 2$

(b) $(2\nu\beta^-\beta^-)\,:\,\Delta L = 0$

Figure 1: Feynman diagrams for double β^- decay processes: (a) neutrinoless double beta decay $(0\nu\beta\beta)$ showing lepton number violation through virtual Majorana neutrino exchange, and (b) standard two-neutrino double beta decay $(2\nu\beta\beta)$ conserving lepton number.

3 The XENON Experiment

The XENONnT experiment is the latest dual-phase TPC developed by the XENON collaboration. It is optimized for WIMP dark-matter searches with a secondary focus on neutrinoless double beta decay studies. The detector is installed at LNGS beneath the Gran Sasso Mountains, providing $\sim$ 3100 m water equivalent background shielding [8]. It reuses key components from XENON1T while scaling up to 8.6 tonnes of xenon target material.

Concentric water-Cherenkov and gadolinium-doped neutron vetoes further suppress cosmic-ray and radioactive backgrounds by tagging muons and neutrons, respectively. The main detector is a 134 cm diameter, 148 cm tall cylinder instrumented with 494 low-background PMTs for event energy reconstruction. XENONnT recently completed its third science run, SR2 and is currently being opened for maintenance.

3.1 Dual-Phase Time Projection Chambers

A dual-phase Time Projection Chamber (TPC) combines the capabilities of scintillation detectors with ionization tracking to achieve superior background discrimination. Unlike scintillation detectors, TPCs measure both light and charge signals from particle interactions, enabling 3D event reconstruction and full energy reconstruction.

The detector consists of a cylindrical chamber with liquid xenon (LXe) in the lower volume, and gaseous xenon (GXe) above. When a particle interacts with xenon atoms, it produces both prompt scintillation photons and free electrons along the particle's path. The scintillation photons are immediately detected by photomultiplier tube (PMT) arrays positioned above and below the active volume, creating the "S1" signal that marks the interaction time and provides information on the scintillation energy. Figure 2 illustrates the XENONnT detector design and fundamental detection principle.

A uniform electric field drifts these electrons upward through the liquid xenon toward the liquid-gas interface. Upon reaching this interface, the electrons are extracted into the gas phase where they undergo electroluminescence, producing secondary scintillation photons in proportion to the original ionization energy. This delayed signal, detected by both PMT arrays, is called the “S2” signal. The S2 pattern preserves spatial information from the original ionization track. The drift time, corresponding to the time difference between S1 and S2 signals, determines the interaction depth, while the spatial pattern of S2 light on the top PMT array is used for (x, y) position reconstruction of the event.

Particle interactions (events) in dual-phase TPCs can be classified into two fundamental types based on their energy deposition mechanisms. Nuclear Recoil (NR) events occur when incoming particles are incident on the target material nuclei, transferring energy via elastic collisions that produce heat, ionize electrons, and excite electronic states. Electric Recoil (ER) events occur when incoming particles are incident on the atom’s electrons, depositing energy mainly into ionization and excitation processes with little heat loss. While both interaction types can produce the same chain of detection events assuming sufficient energy deposition, they differ significantly in their energy partitioning.

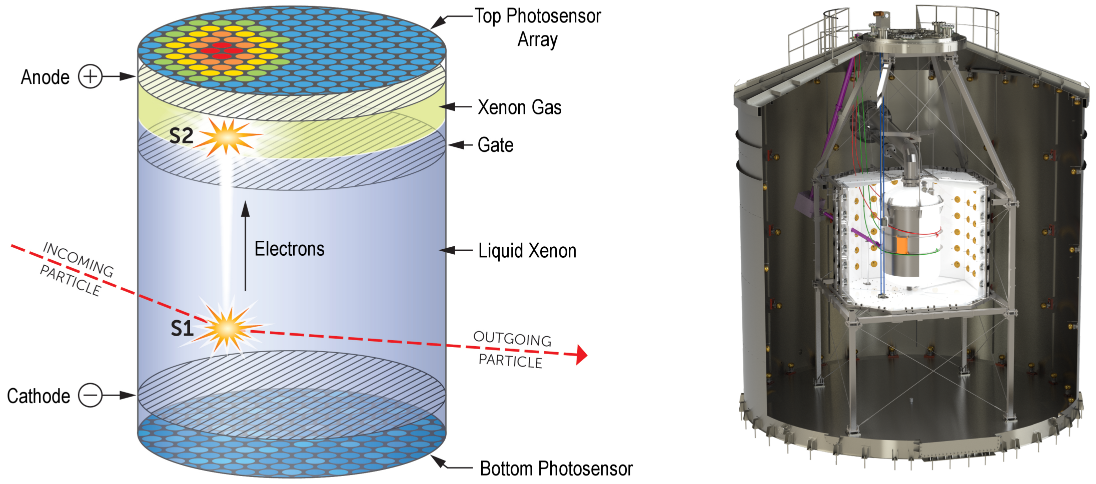
Both interaction types produce signals through the same physical processes but with differing partitioning. Electronic excitation creates excited atoms (Xe^*) that form bound states with ground-state atoms, forming excited dimers (Xe_2^*). These excimers decay within nanoseconds, emitting vacuum ultraviolet (VUV) photons ($\sim$ 178 nm that constitute the source of the detected S1 signal. Liberated primary electrons drift upward under the applied electric field towards the liquid-gas interface, co-moving in what is known as an electron cloud. At the liquid-gas interface stronger extraction fields accelerate these into the gas phase where electrons undergo electroluminescence, the same excimer formation process initiated by energetic electrons rather than the primary collision particle, producing the S2 signal. NR events produce sparser ionization tracks than ER events, making them less affected by electron-ion recombination effects and resulting in characteristically different S2/S1 signal ratios.

The discrimination power of the S2/S1 signal ratio is substantial: XENON1T achieved 99.7% rejection efficiency for electron recoil events while maintaining 50% NR acceptance [9]. ER events, largely resulting from intrinsic beta and gamma decay backgrounds, can thus be effectively separated from the NR signatures expected from WIMP dark matter interactions.

3.2 Xenon

Within the TPC, xenon serves as both the target material and a natural shield against background radiation. Heavy nuclei like xenon provide effective self-shielding through short mean free paths, with xenon’s stopping power capable of reducing 662 keV gamma ray flux by 98% over 20 cm [10]. This self-shielding confines most background events from material contamination, background radioactivity and cosmic rays to the outer detector regions. Such confinement allows for an analysis technique known as fiducialization: by defining a smaller analysis volume within the central regions of the detector and excluding the outer volume, a large portion of background events can be eliminated while retaining a large fraction of signal events from the clean central regions.

Beyond self-shielding, xenon offers several advantageous properties that make it the optimal target material. As the heaviest stable noble gas, the next heavier option, radon, lacks stable isotopes, xenon provides the best combination of mass, inertness, and availability. Its isotopic



(a) Dual-phase xenon TPC detection principle. Particle interactions produce prompt scintillation light (S1) and ionization electrons that drift upward to create delayed proportional scintillation (S2) with a sample hitpattern shown on the top PMT array. Figure from [12].

(b) Complete XENONnT detector assembly showing the time projection chamber (central white cylinder) surrounded by the neutron veto system and muon veto. Figure from [13].

Figure 2: The XENONnT dark matter detector: (a) fundamental detection principle of the dual-phase xenon TPC and (b) complete detector assembly with veto systems.

stability is remarkable: even xenon-124, technically unstable, has a half life of 1.8×10^{22} years, the longest directly measured for any unstable isotope (a measurement achieved by XENON1T itself [11]). Additionally, xenon's excimer decay produces photons to which ground state xenon isotopes are completely transparent, enabling efficient light collection throughout the detector volume with signal reduction only from absorption by detector components and impurities.

3.3 XENONnT Assembly

The XENONnT detector comprises three semi-independent detectors arranged in concentric shells as shown in Figure 2b. The outermost tank contains a muon veto targeting cosmic-ray muons from atmospheric interaction. This veto uses water as the target medium, detecting high-energy muons through Cherenkov radiation via PMTs mounted within the cavity. The intermediate shell contains the neutron veto, designed to flag neutrons primarily from spontaneous fission and (α, n) interactions in surrounding rock and detector materials. The neutron veto operates similarly to the muon veto but employs gadolinium-doped water to facilitate tagging via Cherenkov light.

At its core lies the dual phase TPC itself, featuring two PMT arrays with a total of 494 Hamamatsu R11410-21 3" PMTs that have undergone extensive pre-installation testing [14]. The detector walls consist of numerous PTFE (Teflon) panels arranged in a polygon. PTFE was selected for its high reflectivity at relevant wavelengths alongside its electrical inertness, though these assemblies unfortunately contribute a portion of the dominant (α, n) background through inherent material contamination.

Signal detection for both S1 and S2 events relies on PMT arrays arranged in a hexagonal configuration. The PMT detection process operates through three sequential stages: incident photons first strike the photocathode, ideally producing one photoelectron per incident photon. These photoelectrons are then accelerated toward a dynode stack, where electron cascade multiplication amplifies the signal by factors of $\sim 10^6$, generating a measurable electrical pulse.

However, early XENON experiments revealed that the expected one to one photon-to-electron conversion does not occur as anticipated in the VUV regime. Detector calibration routines indicate that each photon-to-electron conversion has an approximately 20% probability of causing the photocathode to emit a second photoelectron in addition to the primary one [15], thought this varies with measurement methodology. All signal measurements are converted to photoelectrons (PE) forming the primary operating unit within the XENON infrastructure.

The detector assembly requires extensive external facilities housed in a multi-story building adjacent to the main detector. Key components include radon and krypton distillation columns for target material purification, radon being a dominant in-situ background source and krypton serving as a calibration source requiring removal after calibration runs. The facility also contains dedicated liquid xenon and gaseous xenon purification systems. This represents only a subset of required components; comprehensive coverage of all major detector systems can be found in the XENON detector paper Collaboration et al. [9].

3.4 Electric Fields in XENONnT

The electric field system in XENONnT creates two distinctive field domains, a drift field and an extraction field, through a combination of wire grid electrodes and field shaping ring elements. The wire grid electrodes consist of ~ 250 parallel wires each, arranged in five distinct meshes positioned at different heights in the TPC. Starting from the bottom, a shielding mesh protects the bottom PMT array from strong electric fields, followed by the cathode mesh which forms the bottom boundary of the drift field. The concentric field shaping rings maintain field uniformity throughout the drift region until the gate mesh marking the end of the effective drift field. Above the gate mesh, the anode mesh works in conjunction with the gate mesh to create the extraction field, with a final shielding mesh above the anode mesh protecting the top PMT array.

The extraction field region contains the liquid-gas interface, where the anode-gate mesh combination provides a strong electric field that extracts ionization electrons from the liquid phase and accelerated them into the gas region. This acceleration overcomes the energy barrier required for electroluminescence, enabling the production of amplified S2 signals essential for detection of low energy primary events.

The electron energy amplification facilitated by the anode-gate mesh combination exhibits a strong dependence on the local electric field strength, with the light yield from accelerated electrons following a field-dependent relationship.

The electroluminescent light yield exhibits a direct dependence on the local electric field strength according to:

$$\frac{dY}{dx}(x) \propto E(x) \tag{1}$$

where Y represents the number of photons produced per electron and E(x) is the position-dependent electric field. Since the differential light yield changes proportionally to the electric field strength, electrons undergoing electroluminescence closer to the anode wires experience stronger fields and produce higher light yields.

The electric field structure arises from the anode gate mesh combination, where both meshes have a periodic spacing of 0.5 cm with half a half-phase offset between them. This mesh configuration produces electric field line focusing, where field lines converge from the gate mesh region and funnel toward the individual anode wires, this can be seen in Figure 3. This focusing effect concentrates field strength and creates spatial variation in the extraction behavior of individual electrons within a cloud.

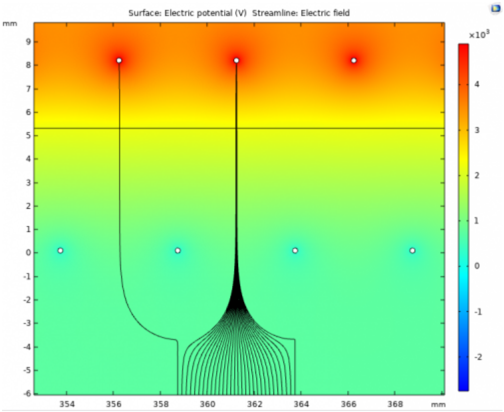


Figure 3: Monte Carlo simulation of electric field lines and potential distribution from gate to anode mesh. The field lines (black curves) demonstrate the funneling effect that concentrates electric field onto the anode mesh, modifying S2 electron arrival locations. Kinks in the field lines are simulation artifacts. Figure from [16].

The detector experiences several problems with the operating electric fields. First, while PTFE is electrically inert, it still accumulates surface static electrical charge, creating field distortion pointing towards the center of the detector. This modifies S2 event locations relative to their primary S1 event sites by shifting them inwards, introducing systematic reconstruction error when using the raw S2 event position to infer primary event position.

During development of XENONnT it was noticed that the anode and gate mesh were subject to gravitational sagging that could hinder proper detector operation. To alleviate this, perpendicular wires were installed; two to support the gate mesh; and four to support the anode mesh. Unlike the primary mesh structure, the support wires are not periodic, creating asymmetric field modulation that drifts electrons away from the perpendicular wire locations and creates high electron densities near the anode support wires.

While the perpendicular wires address significant mesh sagging, the meshes remain free to sag in non-supported areas, albeit at reduced magnitude. Additionally, electrostatic attraction between the gate and anode meshes further deform the parallel structure. As a result, the mesh geometry deviates from the ideal case, introducing spatially varying extraction conditions and light yields.

3.5 Calibration Data and Data Quality

XENONnT uses various calibration sources to characterize detector response and correct systematic effects. This section covers the calibration methods relevant to this analysis, focusing primarily on ^{83m}Kr due to its clean, well-characterized signal properties, with brief mention of LED-based electronic calibration.

^{83m}Kr is a metastable excited state of krypton that decays through a distinctive two-step chain, yielding mono-energetic signals of 32.2 keV and 9.4 keV [9]. The unique timing and energy properties of this decay chain allow for highly effective event isolation and excellent distinction between krypton and background signals.

Krypton is injected into the detector medium and allowed to disperse until homogeneous, providing spatially uniform calibration coverage. Large-scale krypton calibrations are typically performed at the beginning and end of science runs, requiring subsequent removal before data collection. During normal operation, krypton remains present at highly diminished density while maintaining easy identification.

The homogeneous distribution and clean signal properties make ^{83m}Kr ideal for determining key detector parameters such as g_1 and g_2 scaling factors, which convert DAQ readout signals to actual photoelectron counts for S1 and S2 signals respectively. Additionally, the Field Distortion Correction (FDC) accounts for systematic position shifts caused by non-uniform electric fields as explained in Section 3.4.

The work presented in this thesis leverages these properties of ^{83m}Kr calibration data. For this analysis, 88 $^{83m}\mathrm{Kr}$ datasets are selected totaling 1.79 million krypton events, where the only cuts utilized are the krypton cuts outlined in [17]. The complete list of run IDs is provided in Appendix B.

To calibrate the PMT arrays, LEDs are placed within the interior of the XENONnT detector and driven at voltages lower than operating levels such that there is a small likelihood of a single photon produced. The resulting photon detection can then be used to perform gain calibration and electronic noise estimation. Notably, there is a nonzero probability of multiple photons being emitted, and the photocathode will at times emit multiple photoelectrons. To avoid an inductive bias, a model-independent procedure for gain and variance estimation is used [18].

3.6 Data Processing

The XENONnT data processing pipeline transforms raw PMT signals into reconstructed events through a multi-stage procedure. While the specific algorithms and parameters used in this processing chain represents significant technical achievement, the following overview provides context for understanding how calibration data becomes analysis-ready datasets.

The primary signal of the detector corresponds to PMT pulses, digitalized by the DAQ at 100 MHz. Individual PMT readings are classified as a “hit” when they exceed a threshold signal within a defined time window of 30 ns before and 200 ns after the main signal peak. These hits represent the fundamental building blocks of the data processing chain, capturing discrete light detection events from individual PMTs with their associated timing and amplitude information.

Hits are then sequentially grouped with neighboring hits from any PMT into clusters based on temporal proximity. Consecutive hits are assigned to the same cluster when the time gap between them is 700 ns or less. Hits that cannot be grouped into clusters remain as isolated events, corresponding primarily to afterpulses or dark counts, and are processed through a separate pathway.

Clusters are iteratively subdivided into smaller subclusters using a natural break algorithm that analyzes both timing information and the summed waveform characteristics. This splitting procedure is essential for separating S1 signals from PMT afterpulses and distinguishing nearby peaks that may have been incorrectly grouped during initial clustering. When subclusters exhibit saturation effects, they are corrected using pulse models derived from non-saturated channels to recover the true signal amplitude.

The resulting subclusters, now referred to as peaklets, are classified as either S1 or S2 type based on their waveform characteristics. S2 peaklets require additional processing due to their longitudinal drift of electron clouds from the bottom of the detector, which creates broader temporal spreads in the signal. To account for this effect, S2 peaklets undergo remerging using gap-size clustering, where temporally adjacent S2 peaklets are combined until the total duration exceeds 50 µs or no further candidates are found nearby. The remerged peaklets are now classified as peaks.

Following peak processing, events are constructed around single triggering S2 peaks. A triggering S2 must have an area $\leq$ 100 PE and satisfy the n_competing requirement: fewer than eight neighboring peaks within ±10 ms that each exceed 50% of its area. This requirement ensures that only the largest S2 signals in a given time window act as triggering peaks. The event window is defined as extending 2.45 ms before and 0.25 ms after the triggering peak, with overlapping windows merged.

Within an event window multiple S1 and S2 peaks may be present, requiring systematic selection of the primary signals. The primary S1 is defined as the largest S1 peak within the full event window, while the primary S2 is the largest S2 peak occurring after the primary S1. Alternative S1 and S2 signals are identified as the second-largest peaks of each type within the same window. These alternative signals are crucial for identifying and analyzing multiple scatter events, where a particle undergoes more than one interaction within the detector volume.

The n_competing requirement introduces an energy-dependent event building efficiency that varies with the analysis threshold. For electron recoil searches using an S2 area threshold of 500 PE, the event building efficiency reaches 99.3%. WIMP searches employ a lower threshold of 200 PE, resulting in a slightly reduced efficiency of 97.2% due to increased competition from smaller background signals [19].

3.7 Light Response Function

The spatial distribution of S2 light across the top PMT array contains the position information for event reconstruction and can be used for background discrimination. The S2 Light Response Function (LRF) quantifies this spatial dependence by describing the probability for detected photons to be observed by each individual PMT, given a specific event location in the detector. This effectively predicts what fraction of the total detected S2 light signal should be observed by each PMT for a single scatter event.

While position reconstruction algorithms are typically trained using Monte Carlo simulations, see Section 3.8, LRF maps can also serve this training purpose by providing the expected spatial signatures for different event locations. Additionally, deviations from LRF patterns enable identification of anomalous events and multiple scatter interactions.

The LRF is closely related to the Light Collection Efficiency (LCE) map, with the key distinction being the treatment of PMT quantum efficiencies (QE). The mathematical relationship is given by:

$$LRF_i = \frac{LCE_i \times QE_i}{\sum_{j=0}^{n} LCE_j \times QE_j} \tag{2}$$

where both functions are normalized such that $\sum_i LRF_i = \sum_i LCE_i = 1$. The difference between them is that each PMT's LRF value depends on the quantum efficiencies of all other PMTs due to the normalization, creating interdependence across the array. In contrast, the LCE represents the pure geometric light collection probability for each PMT independent of its neighbors, making it a more fundamental descriptor of the detector's optical properties.

While conceptually representing optical light collection, the LRF is not a purely geometric descriptor due to the underlying physics of S2 signal formation. The function describes the behavior of an electron cloud rather than a point light source, making it subject to electric field modulations that affect the S2 emission distribution. These field effects are particularly pronounced for the z-coordinate of S2 emission, where drift field variations can significantly alter the spatial distribution of light. Although x and y coordinate modulations are less critical since Field Distortion Corrections accounts for field-induced position shifts, the complex interplay between electron transport and light emission means the LRF and LCE must inherently capture both optical and electrodynamic effects within the detector.

3.8 Simulation

Comprehensive detector simulation is essential for validating experimental results and understanding detector response. XENONnT employs a sophisticated simulation suite that generates raw data structures identical to real detector output, allowing simulated events to be processed through the same data analysis pipeline as experimental data.

The simulation suite consists of a number of three components that operate sequentially: Geant4, epix and WFSim, each handling distinct aspects of the detector response chain.

The Geant4 toolkit serves as the primary Monte Carlo simulator, managing particle propagation and fundamental physics processes. Each simulation run processes a fixed number of events, tracking primary particles and all secondary interactions through the complete detector geometry. Geant4 handles electromagnetic effects, radioactive decay, and other relevant physics, ultimately producing energy deposits within the detectors sensitive volumes.

Epix takes these energy deposits and converts them into S1 photons and S2 electrons at the liquid-gas interface applying field-dependent and recoil-type-dependent models to determine the light and charge yields for each interaction.

WFSim completes the simulation by propagating these photons and electrons according to precomputed optical maps. It models the extraction field effects, electroluminescence behavior for S2 electrons, and applies PMT-specific responses including quantum efficiencies, timing characteristics, gain variations, and pulse shapes. Finally, WFSim incorporates DAQ electronics effects to produce data files structurally identical to real XENONnT measurements Ramírez García [20].

While the simulation framework provides valuable insights into detector behavior and was instrumental in the design process, certain aspects of S2 signal generation remain problematic. The simulated LRF maps exhibit artificially elevated responses when events occur close to individual PMTs, resulting in poor fit quality compared to experimental data.

A significant contributing factor to these discrepancies is the incomplete modeling of anode mesh electrical field effects. The mesh geometry not only influences electron drift paths but also affects the spatial probability of S2 photon emission. Although preliminary implementations of these field effects exist [21], they have not yet been integrated into the main simulation branch, leaving this important physics component absent from standard simulations.

Given these limitations in the current simulation framework, particularly the absence of anode mesh field effects, alternative approaches for generating reliable LCE functions are necessary. This motivates the development of data-driven methods that can capture the true detector response characteristics observed in experimental measurements.

3.9 Detector Inhomogeneities and LCEs

This analysis utilizes position reconstruction data from the Conditional Normalizing Flow model (CNF) [22] which itself is conditioned on the simulated hitpatterns. This model predicts S2 emission points directly from observed PMT hit patterns and the signal magnitude. Where reconstruction is separate from Field Distortion Correction (FDC) which is applied after the fact to correct for the S1 to S2 position shift induced by inhomogeneous electric fields.

Several spatial inhomogeneities are anticipated in the uncorrected reconstruction data. The CNF model exhibits an inherent bias that shifts reconstructed events away from the detector's walls. Additionally, the detector's support wire geometry introduces density gaps along the supporting wire directions, where the non-uniform electric field structure affects electron drift paths. This reduces the event densities to effectively zero near the gate supporting wires, while increasing event densities around the anode supporting wires, as visible in Figure 4.

Beyond these expected inhomogeneities, the reconstructed event data reveals a systematic bias in the form of periodic density variations aligned with the anode mesh structure. As visible in Figure 4, events exhibit alternating bands of high and low density that run parallel to the anode mesh wires. The periodicity of these variations corresponds approximately to the PMT diameter spacing, with density maxima occurring at the overlap regions between neighboring PMT rows. This striped pattern represents a significant reconstruction bias that was also observed in XENON1T data. This striped pattern represents a significant reconstruction bias that was also observed in XENON1T data. While the exact physical origin remains unclear, this systematic effect likely results from the Monte Carlo simulation not modeling some detector effect correctly.

Analysis of individual PMT light response functions reveals additional spatial inhomogeneity that vary across the detector. As shown in the right panel of Figure 4, most PMTs exhibit reduced light collection efficiency in bands aligned with the anode mesh direction. The pattern is clearly visible for three of the shown PMTs (63, 180, and 72), where the LRF response shows significant suppression along the mesh alignment.

However, this behavior is not uniform across the detector, PMT 216 and neighboring PMTs in the top right portion of the TPC display more radially symmetric response patterns with a qualitatively different mesh-aligned reduction as compared to the other PMTs. The origin of this spatial dependence and the differing behavior between detector regions remains unclear and represents another systematic effect that must be addressed in comprehensive LRF modelling.

These observed spatial inhomogeneities—from wall bias and wire effects to systematic reconstruction patterns—demonstrate that LCE modeling and position reconstruction are inherently coupled problems. Accurate LCE predictions should, in principle, guide events toward their true spatial locations when used appropriately. To quantitatively assess whether different LCE approaches successfully address these inhomogeneities, robust metrics are needed to evaluate the spatial uniformity of position-refined datasets.

To evaluate whether LCE models successfully address these inhomogeneities, homogeneity metrics applied to position-refined datasets provide quantitative assessment. Under the assumption that Krypton is perfectly homogeneously distributed, this should be reflected in the final position refits provided by each LCE model. If homogeneity metrics worsen rather than improve after position refinement, this indicates the LCE model is biased or fails to capture important detector effects, causing artificial event clustering.

However, the parallel wires represent a deviation from this expectation, as they naturally create regions of lower event density due to the drift field. As such, in the computation of these metrics, the areas including all gate and anode perpendicular wires will be masked out, since changes in homogeneity around these points do not provide meaningful information about model performance. To measure the changes in homogeneity, a number of metrics are selected, all chosen for their low computational complexity ($\mathcal{O}(n)$).

The first metric to be utilized corresponds to the coefficient of variation, measuring the ratio of the standard deviation over the mean of binned densities. Where a low C_V then indicates a fairly consistent mean with low variance between bins n_i, and a high C_V indicates strong variation between bins. This metric is mainly useful to determine the variance of the bins.

$$C_V = \frac{\sigma(n_i)}{\mu(n_i)} \tag{3}$$

The next metric corresponds to the Shannon entropy, Shannon entropy quantifies the uncertainty or randomness in the spatial distribution of events. The probability, p_i, of an event falling into a bin is computed for each bin and the entropy is calculated as in equation 4. A higher entropy indicates a more uniform, homogeneous spread of events across the space, while a lower entropy reflects stronger spatial concentration or predictability, this metric is sensitive to both clustering and gaps and will reflect the state of disorder (homogeneity) found within the TPC. This quantity will be normalized to the maximum attainable value to allow easier comparison between run placing the effective range in $H \in [0, 1]$

$$H = -\sum p_i \log p_i \tag{4}$$

The Gini coefficient will also be utilized, this coefficient is a measure of bin inequality, quantifying the significance of deviation shown in equation 5. It essentially functions as a weighted sum, using the sorted bin values c_i with m bin values in total, each bin count is weighted by the term $(2i - m - 1)$ corresponding to its index shifted into the negative regime and stepped twice rather than once, this allows creation of a symmetric weight set about zero. If then all bins are of comparable size the coefficient $G = 0$ as the positive and negative domains share the same weights. If however the bins are not equal the Gini coefficient grows with a maximal value of $G = 1$, this is only satisfied if only a single bin holds all events.

$$G = \frac{\sum_{i=1}^{m}(2i - m - 1)c_i}{m\sum_{i=1}^{m} c_i} \tag{5}$$

And lastly the non-occupancy ratio, simply quantifying the fraction of bins that do not contain any events is used to indicate if at the bin resolution used dead bins are present. All mentioned metrics are bin dependent, with bins only defined within the TPC, $R < 66.4$ cm.

The combination of metrics should provide a general overview of the numerical homogeneity of the data, when taken alongside visual inspection of the density plots this will provide sufficient information to verify the correctness of specific maps.

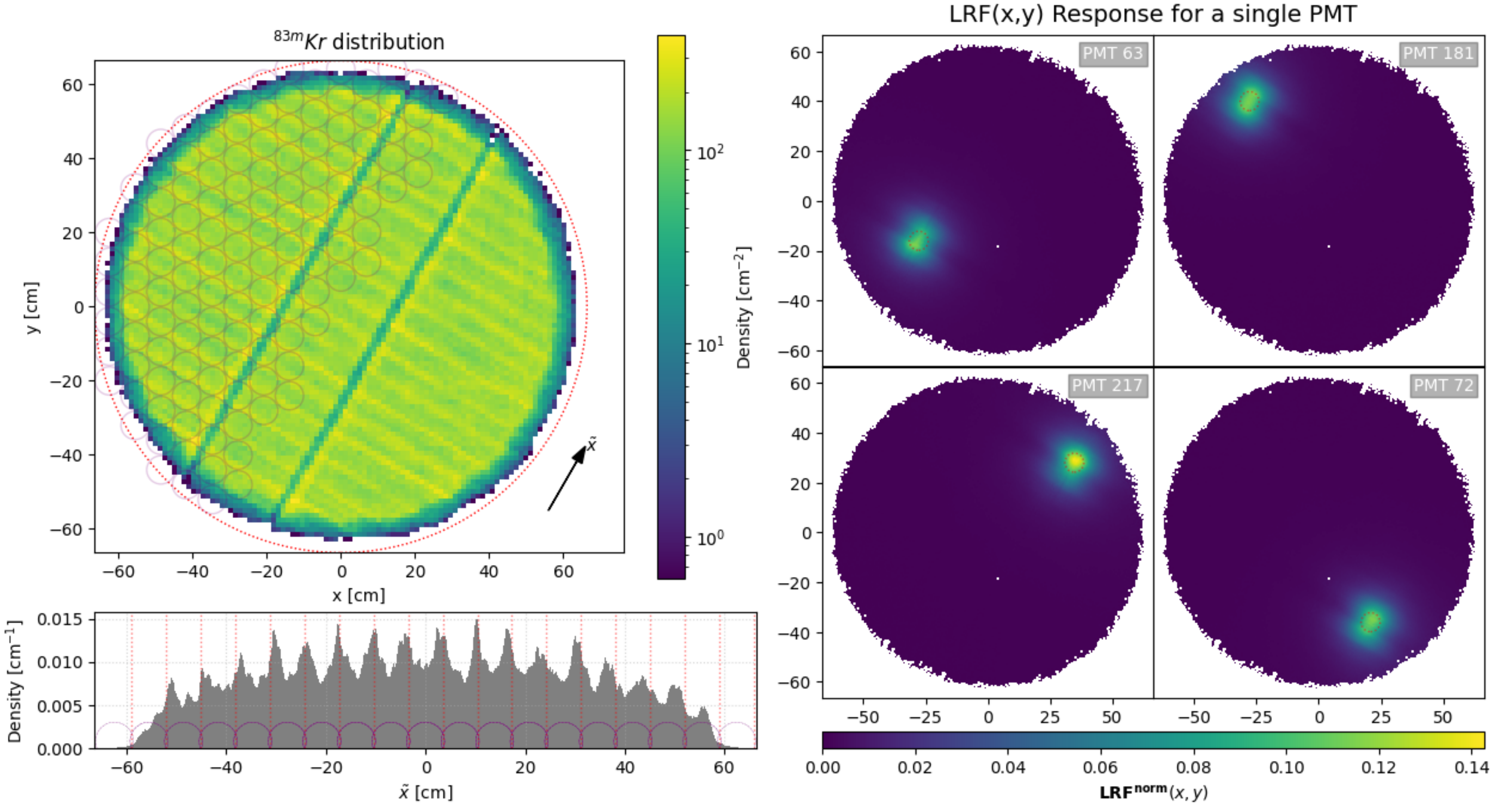


Figure 4: Spatial inhomogeneity within the TPC. **Top left:** ^{83m}Kr event density as a function of position within the TPC, reconstructed using the CNF position reconstruction algorithm without field distortion correction. The plot shows the TPC extent in cylindrical approximation, the rotated $\tilde{x}$ axis aligned with the anode mesh, and PMT locations (gray squares). Visible features include density gaps along perpendicular supporting wires, reduced density at high radius due to field distortion and CNF radial bias, and inhomogeneity parallel to the anode mesh. **Bottom left:** Event density collapsed along the rotated $\tilde{x}$ axis. PMT positions are marked with vertical lines, showing increased event reconstruction at PMT intersection points, particularly near PMT centers. **Right:** Normalized LRF response patterns for individual PMTs as a function of spatial position, derived from real data. Notable variations include significantly reduced light collection along the anode mesh direction for PMTs 63, 180, and 72, while PMT 216 exhibits more radially symmetric behavior with this directional suppression visible only shifted from the center of the PMT.

4 Anode Mesh: Monte Carlo Simulations

To facilitate a description of the S2 response pattern focusing on optical parametrizations for the S2 LCE, a better understanding of the mechanisms by which the anode mesh modulates the signal is required. This Section focuses on the reduction band observed with LCE plots per PMT as a function of space (seen in Figure 4). Several simple Monte Carlo simulations investigate this effect by propagating photons from different S2 event coordinates, all simulations only consider a single PMT, an infinite anode mesh, and various source emission distributions including both electron cloud and point source emission.

The simulation is rudimentary, simply checking if photons intercept with the anode mesh and absorbing it if it does so, and then checking for interception with the PMT surface. Successful PMT hits are counted, while all other photons are discarded. All simulations use an anode mesh pitch of 0.5cm with a wire radius of 0.0152 cm.

The investigation will first focus on fixed z point source emission, moving to the effect of fixed z cloud source emission, followed by an extension including probabilistic z depth S2 event sites and finally an extension of this where the photon yield is scaled alongside the probabilistic z depth emission.

Fixed z point source emission The first trial corresponds to emitting $100,000$ photons from each simulated event site for a fixed z coordinate corresponding to 0.006 cm below the anode mesh, with randomly sampled propagation vectors pointing into the upper hemisphere. This is computed for all positions within a radius of twice the PMT radius ($R = 7.62$ cm), with a resolution equivalent to a quarter of the anode wire diameter, $dx = 0.0076$cm. The resulting positional distribution of photons arriving at the PMT can be seen in Figure 5.

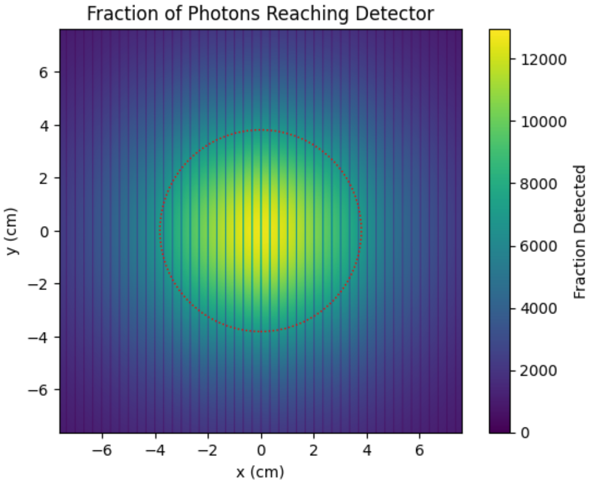


Figure 5: Photon count detected by a single PMT (red dashed outline) as a function of source position for point source photon emission at constant $z = 0.006$ cm below the anode mesh. For each position, $100,000$ photons are emitted, with the simulation only considering mesh interception and PMT detection.

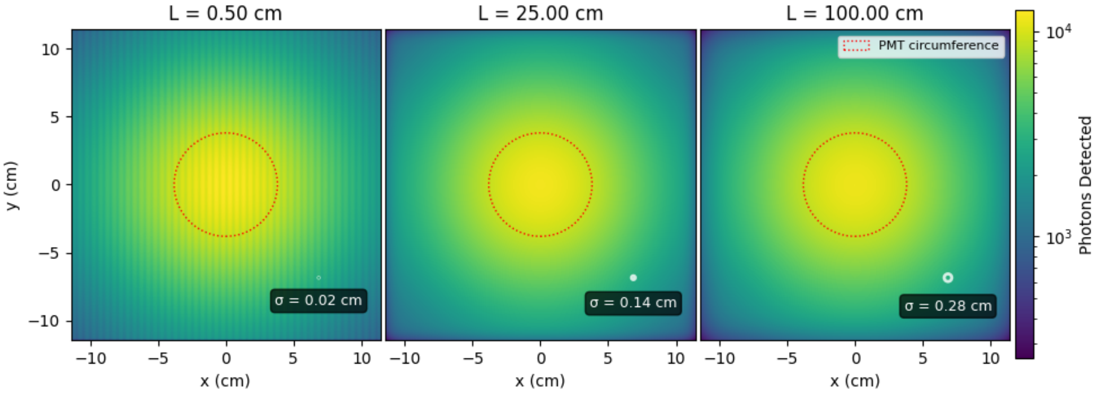


Figure 6: Photon count detected by a single PMT for constant z emission simulated through point source photon emission, subsequently convolved with a Gaussian kernel to simulate electron cloud electron sources. Results show the 99% interval for Gaussian cloud spread computed at various emission depths (L), where the signal at each point represents a weighted average over the cloud distribution. The red dotted circle corresponds to the PMT circumference, while the white circle indicates the 1σ extent with the corresponding numeric value (σ) displayed.

Within this plot, a clear radial distribution of signal count is visible as expected from the direct detection component, a strong reduction of the signal is visible under the anode mesh wires with counts dropping to zero in most cases.

Fixed z cloud source emission A notable difference between this simulation and reality is that this assumes a point source emission rather than emission from an electron cloud. Focusing on a 2D cloud, the expected Gaussian spread of the electron cloud at the S2 emission point can be estimated using standard diffusion equations.

For XENONnT conditions (drift speed $\sim 8.83 \times 10^{-5}\ cm\ ns^{-1}$, diffusion coefficient $D_t = 34.625\ cm^2/s$ [23]), the expected cloud width at mean detector depth ($L = 75\ cm$) is $\sigma = 0.24\ cm$.

This effect is simulated by convolving the point source data with Gaussian kernels corresponding to different drift distances (Figure 6). To quantify potential anisotropy, 2D Gaussian fits are applied to extract x and y spreads. The statistical significance of observed variations is assessed through Poisson bootstrap analysis, where multiple equivalent realizations of the data are generated by applying Poisson random sampling to each pixel, simulating the inherent statistical fluctuations in photon detection.

Results show minimal evidence for mesh-induced anisotropy with x/y spread ratios showing deviation from unity at $\mathcal{O}(10^{-4})$ no statistically significant trends with drift distance, where bootstrap analysis showed variance on the order of $\mathcal{O}(10^{-3})$.

Variable z cloud source emission The second Monte Carlo experiment utilizes a GARFIELD-derived emission distribution that accounts for anode wire electric field effects on emission probability as a function of 3D space [21]. In this framework, electrons follow electric field lines

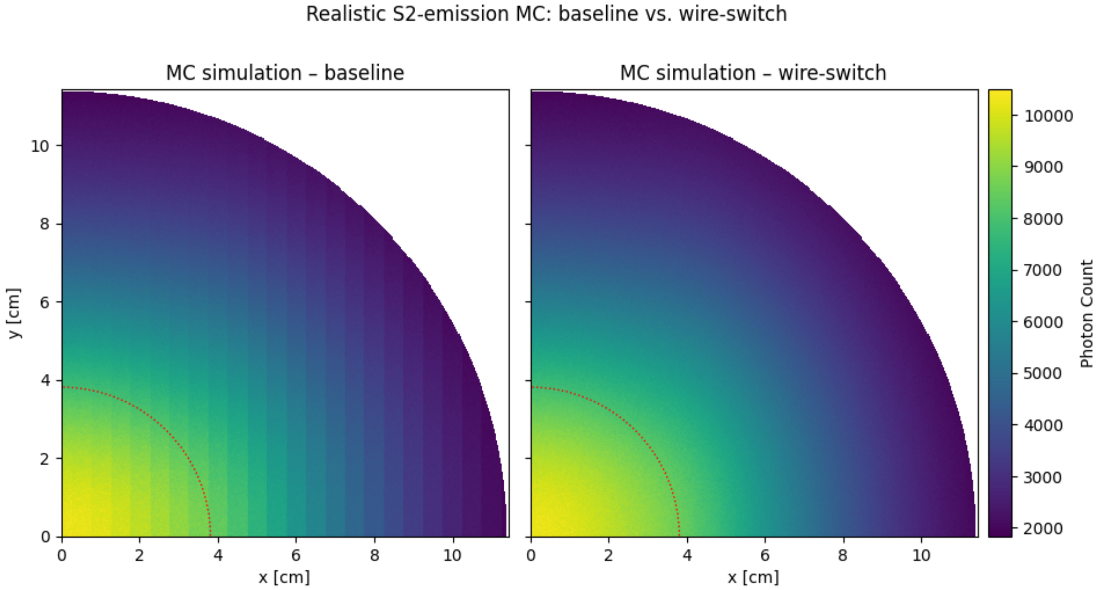


Figure 7: Realistic S2 photon emission optical MC simulation considering anode mesh shadowing, where electrons within a cloud release photons according to a precomputed spatial probability distribution. On the LHS each electron cloud is forced to associate with a single anode wire, while on the RHS electron clouds are free to associate with any neighboring mesh wire depending on the cloud center and a random normal shift to each electron sourced from a Gaussian of mean 0 and standard deviation $\sigma = 0.24$ cm corresponding to a drift distance of $L = 75$ cm.

that terminate on specific anode wire. Electric field lines originating from the region between adjacent gate wires converge upward onto individual anode wires, see Figure 3. This field geometry channels electrons along curved paths that focus onto specific anode wires, constraining the emission region (above the liquid-gas interface) into a narrow band with a maximal spread of $\sigma_x = 0.22$ cm perpendicular to each wire.

Two simulation variants are compared: first, each sampled electron cloud is associated with a singular anode mesh wire, where all of its electrons are attracted to it, with final positional distributions sampled entirely from the GARFIELD derived distribution. Second, individual electrons within a cloud are free to associate with different wires, where wire association is sampled by generating an electron within the Gaussian cloud and associating it with the nearest wire. Due to the narrow x-spread of the S2 emission point around an anode wire, a fine grid resolution of 0.0304 cm is used, evaluating the positive quadrant out to three PMT radii (11.43 cm).

The maximal deviation between the two simulations corresponds to 730 photons, with the major difference being that the wire switching implementation does produce a smoothed version without the discretized steps in x observed in the single-wire case.

For the produced dataset sampling of a confidence ellipse and 2D Gaussian intensity fit as done for the fixed z emission cases produced negligible deviations. To isolate directional effects from the underlying radial dependence, the PMT response must be decomposed into radial and residual components. Since parametric fitting of the radial profile proves numerically challenging,

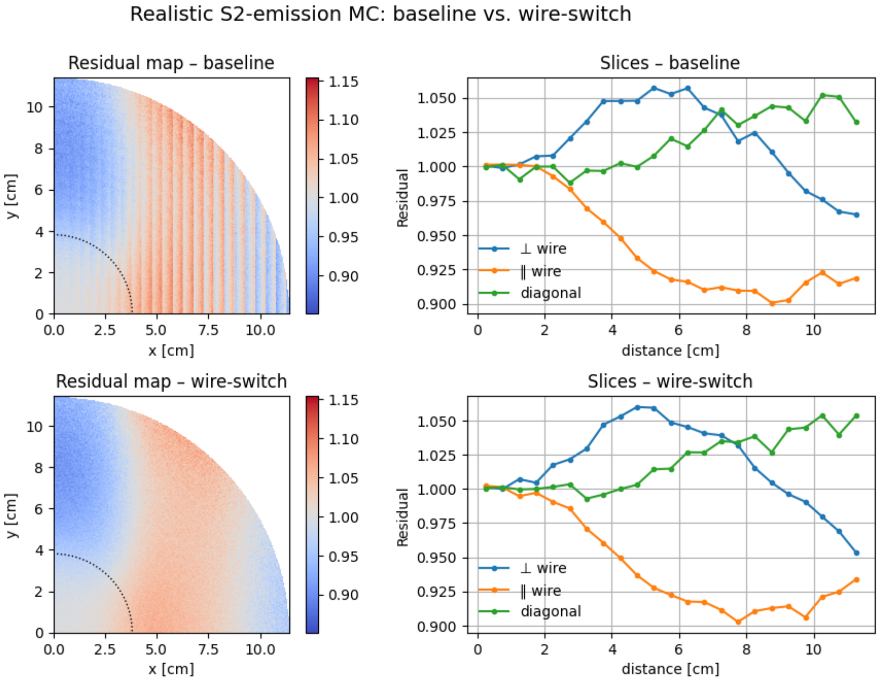


Figure 8: The residual of realistic anode shadowing simulations including electrical emission depth effects, when dividing out the best fit Piecewise Cubic Hermite Interpolating Polynomial (PCHIP). In the top row electrons are forced to associate with the same mesh wire as the cloud center, while in the bottom row electrons are free to associate with any wire based on the Gaussian electron cloud spread, the PMT extent is indicated with a black dashed line. In the second column the residuals are shown as a function of direction starting from the $(0, 0)$ coordinate, with blue perpendicular to the anode mesh, orange parallel to the anode mesh and green along the diagonal.

a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation is employed instead to capture the radial dependence. This interpolation allows extraction of a Direct Detection component modeled as a purely radial contribution with no directional dependence. The result of this procedure alongside the directional dependence can be seen in Figure 8.

Considering the residuals reveals clear directional anisotropy in the S2 optical response. Both simulation variants show consistent behavior: an enlargement of up to 5% from the best fit interpolation perpendicular to the anode wire direction and a reduction of 10% along the wire direction, with the central response largely unmodulated. This directional modulation provides the foundation for the optical parametrization approach developed in subsequent sections.

While this simulation approach approximates the true conditional distribution well, the idealized implementation assumes electrons associate with wires based solely on geometric proximity. A fully accurate model would use conditional probability distributions that account for the electron cloud center position relative to multiple neighboring wires.

Electron Energy Adjustment A last modification needs to be completed, the photon number yield is directly proportional to the electric field strength at the emission point. Since electric field strength decreases with distance from the anode mesh, electrons closer to the mesh produce higher photon counts. To include this effect a simple linear scaling is applied where photon emission furthest away from the anode mesh receive a weight of 0.05 and emission closest to the mesh receives a weight of 1. At the center of the anode mesh, the GARFIELD-derived emission probability is null (no photons spawn within the wire), so no emission events will actually receive the maximum weight. The simulation exploits axial symmetry about the origin: the single-quadrant response is mirrored to generate a full PMT response. This PMT response is then replicated and spatially shifted to create a hexagonal arrangement of 7 PMTs (6 surrounding + 1 central), producing the expected LCE response that includes geometric effects from neighboring PMTs.

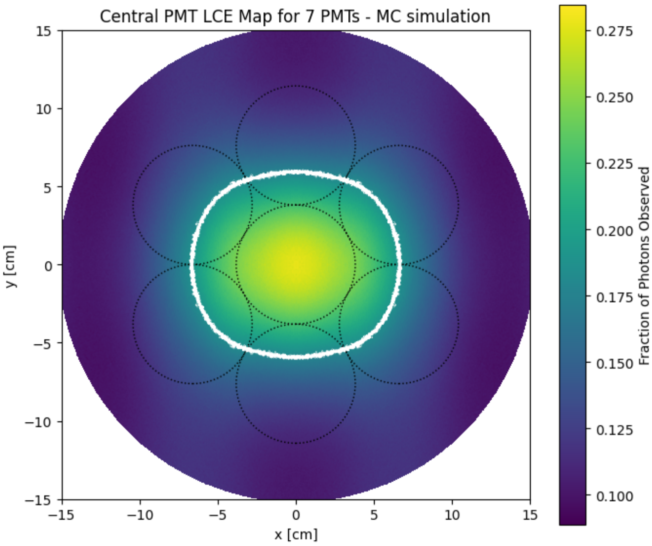


Figure 9: LCE for the central PMT from MC simulation of electron cloud source, where electrons are free to associate with individual anode wires broadcast to hexagonally arranged PMTs. In white the 33% maximum signal contour is shown, for purely radially symmetric responses this contour is expected to be radially symmetric. At the edges strong gradient changes can be observed, this is due to the full plot being generated by replicating the response of a single PMT, these points correspond to areas where the response of a PMT is no longer defined.

While this does not fully explain the observed data, as can be seen by visual comparison to Figure 4, it demonstrates that the radial LCE in conjunction with a fixed z anode parametrization cannot completely describe the direct detection response. The simulation reproduces the directional anisotropy observed in the previous analysis. After PCHIP fitting, the residual response perpendicular to the anode wire is enhanced by 8% while the response parallel to the wires is reduced by 10%, confirming that an emission distribution closer to the mesh, or equivalently with higher yield closer to the mesh, further exacerbates the optical modulation.

This modulation can explain the density variation observed within the raw reconstruction density. Since the XENON MC simulation does not yet account for depth-dependent emission probabilities, position reconstruction is not aware of the magnitude of the differential response.

As a result, if PMTs aligned with the mesh observe a high fractional response, the event is likely reconstructed in the center row by the reconstruction model. However, if the central PMT has a large response but its mesh-aligned neighbors do not, the model perceives this as only possible if the event occurs at the edge of the PMT row, when the event should actually be reconstructed more centrally given the observed response. Producing a high event density at the edges of PMT rows.

Having established the physical mechanisms behind anode mesh effects, the next step requires developing a quantitative framework to fit these effects to experimental data, which necessitates a robust likelihood function for PMT responses.

5 PMT Signal Likelihood's

To fit LRF models to experimental data, we need a metric that quantifies the agreement between model predictions and observed PMT responses. The specific likelihood function used in this analysis was originally developed by Aalbers [24], with minor modifications made here and extended for efficient computation at scale. Likelihood functions provide a natural framework for this optimization, as they quantify the probability of observing specific data given model parameters. Understanding the statistical variance in individual PMT responses is essential for constructing an accurate likelihood function that can serve as both an optimization metric and quality of fit parameter throughout this analysis.

This section develops the theoretical foundation by examining the key physical processes involved in generating a detector response: photon arrival statistics, photoelectron conversion at the photocathode, and electronic amplification. While this represents a simplified model of the complete detection chain, each process contributes well-defined statistical distributions that can be combined to construct a tractable likelihood function for PMT responses. Finally, the practical implementation is addressed, including computational approximations necessary to make the theoretically intractable form evaluable in practice.

5.1 Likelihood Formulation

Photon Statistics In an S2 emission event, photons travel from the event site to the PMT photocathodes. The signal of interest is the number of photons incident on each photocathode surface. This can be described using a Poisson distribution, with N arrivals given an expected μ photons for an emission event. The likelihood, $L_P(N|\mu)$, for the Poisson distribution is written as:

$$L_P(N|\mu) = \frac{\mu^N e^{-\mu}}{N!} \tag{6}$$

Photoelectron Conversion At the photocathode, incident photons undergo photoelectron conversion with an average yield of ~ 1.2 photoelectrons per photon in the VUV regime [15]. This excess arises from double photoelectron emission (DPE), modeled as each photon having probability p of producing an additional photoelectron. This additional photoelectron production is modeled using a binomial distribution, where for N incident photons, the number of extra photoelectrons k follows a Binomial(N, p) distribution. The probability mass function is:

$$L_B(k \mid N, p) = \begin{cases} \binom{N}{k} p^k (1-p)^{N-k} & \text{if } 0 \leq k \leq N, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

The total number of photoelectrons emitted is then $N + k$, where N is the number of incident photons and k is the number of additional photoelectrons from DPE events.

$$L(m|\mu) = \sum_{N=0}^{\infty} L_P(N|\mu) \sum_{k=0}^{N} L_B(k|N,p)\delta(m - N - k) \tag{8}$$

where the delta function ensures $m = N + k$, with m the observed number of photoelectrons.

Amplification and readout After photoelectron emission, each electron is accelerated toward the first dynode, where it triggers the release of multiple secondary electrons. This multiplication process repeats through a cascade of dynodes until a measurable signal is produced. The amplification chain introduces additional variance from dynode to dynode gain variations, vacuum contamination effects, and electronic readout noise in the DAQ.

While the complete statistical description of this cascade is complex, the electronic contribution can be approximated with a Gaussian distribution for computational tractability. This approximation captures the dominant noise behavior, though LED calibration data shows some deviation due to heavy tails in the true single photon response. The computational simplicity of the Gaussian model makes it a reasonable compromise for this analysis.

The final observed signal x (gain-corrected to photoelectron units by the reconstruction pipeline) follows:

$$L_G(x \mid N+k, \sigma\sqrt{N+k}) = \frac{1}{\sqrt{2\pi\,\sigma^2(N+k)}}\exp\left[-\frac{(x-N+k)^2}{2\,\sigma^2(N+k)}\right] \tag{9}$$

where σ represents the per-photoelectron electronic noise contribution, $N+k$ is the mean number of photoelectrons, and the variance scales as $\sigma^2(N+k)$ to account for the noise contribution from each photoelectron in the amplification chain.

The Full Likelihood With the inclusion of the Gaussian noise term, the full likelihood for a single PMT observation can be written for an observed signal x with an expectation of μ photons:

$$L_{full}(x|\mu) = \sum_{N=0}^{\infty} L_P(N|\mu) \sum_{k=0}^{N} L_B(k|N,p) L_G(x|N+k, \sigma\sqrt{N+k}) \tag{10}$$

For model optimization and quality assessment, we use the negative two log likelihood ratio (NLLR):

$$-2\log\frac{L(x|\mu)}{L(x|x_{MLE})} \tag{11}$$

where $L(x|\mu)$ represents the likelihood of the observation given the model prediction, and $L(x|x_{MLE})$ represents the maximum achievable likelihood for that observation, providing normalization to the best possible fit. While this quantity follows a χ^2 distribution under ideal conditions (Wilks' theorem [25]), deviations are expected for low-signal regimes where the response is non-Gaussian and due to imperfections in the likelihood model itself.

We define χ^2_F as this NLLR value for the full likelihood function, with χ^2_{FR} representing the reduced form averaged across PMTs. These metrics serve as both optimization targets and goodness-of-fit measures throughout the analysis.

5.2 Implementation

The theoretical likelihood function developed above requires evaluation of infinite sums for each PMT signal, making direct computation prohibitively expensive for real-time LRF fitting. Each likelihood evaluation involves convolving Poisson, binomial, and Gaussian distributions—a process that scales poorly with the number of PMTs and model evaluations required during optimization. To make this likelihood computationally tractable, we implement a two-regime approach: exact evaluation with lookup table acceleration for low signals ($x < x_{\text{switch}}$), and a Gaussian approximation for high signals ($x \geq x_{\text{switch}}$). The signal likelihood is parameterized as $L(x|\mu;\ \sigma)$, where x is the observed signal in photoelectrons, μ is the expected number of photons, and σ represents the per-PMT electronic noise parameter.

Exact evaluation for $x < x_{\text{switch}}$. For signals below the switching threshold, the likelihood is implemented as a double sum over the photon count N_γ and photoelectron count N_{PE}:

$$L(x|\mu) = \sum_{N_\gamma=0}^{N_\gamma^{\max}} L_P(N_\gamma|\mu) \sum_{N_{PE}}^{N_{PE}^{\max}} L_B(N_{PE} - N_\gamma|N_\gamma, p) L_G(x|N_{PE}, \sigma\sqrt{N_{PE}}) \tag{12}$$

Here $N_{PE} - N_\gamma$ represents the number of double photoelectron emission events (k from the theoretical formulation), and the binomial component is masked to return zero likelihood for negative arguments.

The likelihood evaluation requires matching the model prediction μ to the observed data scale. The LRF produces normalized photon distributions (summing to unity), while the observed signals represent photoelectron counts. For rescaling, only positive-valued observed signals are summed to determine the total signal level, as negative values arise from noise and baseline subtraction rather than physical detections. The LRF prediction is then scaled to match this positive-signal total and applied to all PMT channels during likelihood evaluation.

The implementation uses TensorFlow [26] with XLA compilation for computational efficiency, which requires static array sizes. Therefore, the evaluation ranges $N_\gamma^{\max}$ and $N_{PE}^{\max}$ must be predetermined rather than computed dynamically for each signal.

These ranges are chosen to encompass a 5σ envelope around the expected values, which in the Gaussian limit captures $> 99.9\%$ of the probability mass. To determine the appropriate range sizes, we use the approximation that Poisson distributions have variance equal to their mean in their Gaussian limit. Since we use a conservative 5σ envelope, other variance contributions are neglected for simplicity in range construction.

The evaluation ranges are constructed as $N_{PE}^{\max} = x_{\text{switch}} + 5\sqrt{x_{\text{switch}}}$ and $N_\gamma^{\max} = x_{\text{switch}}/(1 + p) + 5\sqrt{x_{\text{switch}}/(1+p)}$ where the relationship $N_\gamma \approx N_{PE}/(1 + p)$ accounts for the expected photoelectron multiplication from double emission events.

Lookup-table acceleration. The exact evaluation remains computationally expensive, requiring thousands of likelihood evaluations per signal due to the double sum over photon and photoelectron ranges. Each evaluation involves computing Poisson, binomial, and Gaussian components, making real-time optimization prohibitive without acceleration.

To address this, a three-dimensional lookup table (LUT) is implemented on a grid of $\{x, \mu, \sigma\}$ values. The signal x and prediction μ are discretized in steps of 0.1 across their respective ranges, while each PMT receives its own dedicated σ slice for PMT-specific noise characteristics.

The LUT provides an additional crucial benefit: since all likelihood values are precomputed, the maximum likelihood estimate (MLE) for each (x, σ) combination can also be stored directly. This eliminates the need for iterative MLE computation during fitting, which is essential for calculating the negative log-likelihood ratio values used throughout the analysis.

This precomputation strategy dramatically improves evaluation speed at the cost of discretization error and initial overhead. The LUT approach is primarily used for evaluation tasks, while the exact method is retained for fitting operations where continuous evaluation is required.

Gaussian approximation for $x \geq x_{\text{switch}}$. For signals above the switching threshold, the exact sum approaches a normal distribution due to the central limit theorem. The convolution of Poisson photon statistics, binomial DPE events, and Gaussian electronics noise will then be approximated as a single Gaussian distribution characterized by its mean and variance.

In the Gaussian limit, each component has the following variance characteristics:

$$\text{Poisson}\big(N_\gamma\big) \xrightarrow{CLT} \mathcal{N}(N_\gamma,\, N_\gamma)\,, \quad \text{Binomial}\big(N_\gamma, p\big) \xrightarrow{CLT} \mathcal{N}(N_\gamma p,\, N_\gamma p(1-p))\,.$$

For the combined Gaussian approximation, the mean is determined by rescaling the observed signal to photon units: $\mu = x/(1+p)$, representing the maximum likelihood estimate of photons given the observed photoelectrons. The variance combines contributions from all three sources:

$$\begin{aligned} \sigma^2_{\text{switch}} &= N_\gamma(1+p)^2 \\ &+ N_\gamma p(1-p) \\ &+ \sigma^2\sqrt{N_{PE}} \end{aligned} \tag{13}$$

where the terms represent Poisson, binomial, and electronic noise contributions respectively. This Gaussian approximation provides computationally efficient evaluation for high-signal regimes while maintaining statistical accuracy.

Parameter Implementation Attempts were made to compute the Gaussian variance from LED PMT calibration data, however, these variances are not precomputed within the XENONnT infrastructure and computing them proved more complex than expected, falling outside the scope of this thesis. Instead, these variances are fitted directly during likelihood optimization. Incorrect variances values inherently degrade likelihood function—values that are too large or too small produce poorer fits to the observed data, making the variance parameters self-correcting through the optimization process.

The switching signal $x_{\text{switch}} = 40$ was chosen following the recommendations in the XENON1T pattern likelihood development Aalbers [24]. This yields maximum evaluation ranges of $N_{PE}^{\max} = 77$ and $N_\gamma^{\max} = 67$ for the exact evaluation regime.

The probability for emitting a double photoelectron is set to $p = 0.2$, due to the per-PMT and per-measurement method variability hovering around this value [15].

Table 1: Comparison of negative log-likelihood values between exact and LUT evaluation methods.

	$-\log[L(x\|\mu)]$: Exact	$-\log[L(x\|\mu)]$: LUT
Mean	2.539	2.526
Median	2.642	2.629

5.3 Likelihood Behavior & Expected Deviations

Baseline χ^2_{FR} To establish a computational framework for generating χ^2_{FR} baseline values, a Monte Carlo test case is constructed. Using Monte Carlo LCE map[1], $10,000$ patterns are generated at randomly sampled positions across the detector. Each pattern is scaled to a total signal drawn from a normal distribution with $\mu = 5,000$ photons and $\sigma = 950$, corresponding approximately to the mean and variance observed in the Krypton calibration dataset.

To generate signals following the distribution expected by the likelihood function, each event is processed through the complete detection chain: Poisson statistics determine the number of incident photons, binomial conversion represents photoelectron emission at the photocathode, and Gaussian noise ($\sigma = 0.5$ PE accounts for electronic noise sources. The original Monte Carlo LCE prediction serves as μ, while its randomized version represents the observed signal x.

This framework establishes a baseline χ^2_{FR} reference value for theoretical perfect fits. Using the full likelihood method on this validated dataset yields a baseline of $\chi^2_{FR} = 1.181$, which serves as the reference for evaluating fit quality in subsequent analyses.

LUT Accuracy To assess the numerical accuracy of the lookup table implementation, both LUT and exact evaluation methods were applied to the generated dataset. Table 1 compares the mean and median negative log-likelihood values between approaches.

The minimal deviation ($\mathcal{O}(0.01)$) demonstrates that the LUT approach maintains sufficient accuracy for practical applications. However, gradient-based optimization methods perform better with the exact method due to the extra fidelity provided by continuous evaluation. Therefor, the implementation strategy uses exact likelihood evaluation for parameter fitting and LUT evaluation for computational efficiency in grid searches and large-scale evaluations.

Having established the baseline statistical behavior and computational framework of the likelihood function, the next investigation provides a generic overview of machine learning approaches before applying them to the detector response analysis.

---

[1] `XENONnT_s2_xy_patterns_LCE_corrected_qes_MCva43fa9b_wires.pkl`

6 Machine Learning

The focus of this work is to construct a parametrized version of the Light Response Function (LRF), with specific emphasis on modeling the production and propagation components separately. This component-wise approach enables systematic investigation of detector biases and effects that influence the LRF.

As presented in Section 3.9, several problems exist with the currently observed Light Collection Efficiencies (LCEs). While Monte Carlo simulations in Section 4 have partially explained some of these effects, a divide-and-conquer approach is needed to systematically isolate and model individual contributions. By subdividing the LCE and LRF responses into their fundamental physical components, effects can be separated and characterized independently. This component-wise modeling allows verification of Monte Carlo predictions against observed data and enables identification of remaining systematic biases once leading-order contributions are accounted for.

The reconstruction process requires iterative position refinement to correct for systematic biases present in the initial position estimates. Two sequential approaches are employed: first, an encoder-based method where event positions become trainable neural network weights updated through standard machine learning optimization; second, a grid search refinement using predefined step sizes until convergence within specified tolerance. This overall methodology mirrors the approach of Aalbers [27].

As a parallel investigation, Symbolic Regression was explored to determine functional forms for each component. Symbolic Regression uses genetic algorithms to discover equations that satisfy the data's behavior with minimal assumptions about functional form and constrained mathematical operations.

This approach was largely unsuccessful for several reasons. First, genetic algorithms are computationally expensive, as Symbolic Regression constructs equations as binary trees and evaluates function fits through iterative tree node modifications. The parameter space is extremely large, with minor modifications often drastically altering results.

While Symbolic Regression may become viable after all components are well-characterized through other methods, this would require follow-up work. The advantage of such an approach is that the resulting equations would be highly interpretable and transferable to different detector systems.

Therefore, the parametrization is accomplished using machine learning methods, specifically Dense Neural Networks for the component-wise modeling approach. Dense networks are well-suited for this parametrized approach, as they can capture the complex non-linear relationships between event positions and individual physical components of the light response function. Before implementing these models, an overview of the relevant machine learning fundamentals is provided to establish the theoretical foundation for the component modeling approach.

Machine Learning Overview Within the machine learning framework, LRF functions are modeled as differentiable mapping $f_\theta(x) \to \hat{y}$, where x encodes the event positions and $\hat{y}$ represents the per-PMT S2 pattern re-normalized to unit sum over the S2 top array.

For the component-wise modeling approach, neural networks are implemented using Dense (fully-connected) layers that can capture global correlations between input positions and output responses. The basic form of a dense layer corresponds to:

$$y = \sigma(x \cdot w + b) \tag{14}$$

where x is a vector input, w is a tensor weight matrix, b is the bias vector and σ is a non-linear activation function. This activation function enables the network to approximate complex, non-linear functional behaviors. Multiple dense layers in sequence form a Multi-Layer Perceptron, providing the foundation for modeling individual LRF components.

Training proceeds by minimizing a loss function with respect to the network parameters:

$$\mathcal{L}(\theta) = \sum_{i=1}^{N} \ell(f_\theta(x_i), y_i) \tag{15}$$

where i indexes individual training events, and ℓ corresponds to the negative log likelihood described in section 5.

Training proceeds by alternating forward passes (computing f_θ and the loss function), backward passes (computing per-parameter gradients through backpropagation), and parameter optimization via gradient-based optimizers.

Each update step is computed over a batch of predefined size. The batch size determines how much noise from the data propagates to the model, as the loss value computed per batch is used to scale all gradients. Smaller batches introduce more noise but provide more frequent updates, while larger batches give smoother gradients but fewer updates per epoch. As a result, each complete pass through the dataset (epoch) involves multiple weight updates corresponding to the number of batches.

To provide context for how these optimizers operate and evolved, we begin with Stochastic Gradient Descent (SGD), one of the oldest optimizers still in use [28]. SGD updates each weight parameter in the direction of the negative gradient scaled by a learning rate:

$$w_{t+1} = w_t - \eta \nabla_w \mathcal{L}(w_t) \tag{16}$$

A variation known as SGD with Momentum adds a momentum term to each gradient update, where momentum corresponds to an exponentially weighted average of previous gradients accumulated over time. This modification helps smooth out noisy gradient estimates and accelerates convergence when gradients consistently point in the same direction, while dampening oscillations in regions with high curvature.

Adaptive Moment Estimation (ADAM) combines the principles of momentum and adaptive step sizes. It maintains exponential moving averages of both the gradients (first moment) and their squared values (second moment), then scales updates by dividing the first moment by the square root of the second moment plus a small ϵ for numerical stability.

This provides each parameter with adaptive learning rate scaling: parameters with historically large gradients receive smaller steps, while those with consistently small gradients receive larger steps. Effectively, this gives each parameter dynamic self-adjustment based on its gradient history [29]

An optional extension to ADAM is weight decay, which prevents the model weights from becoming too large. On each update, a small fraction λ of each weight is subtracted, gradually shrinking all weights towards zero over time. This prevents the model from overfitting to noise in the training data and improves numerical stability.

In practice, this is implemented by adding $\lambda \cdot w$ to the computed gradients before applying ADAM's adaptive scaling, creating a penalty for large weights during optimization.

ADAM with weight decay can be further extended with decoupled weight decay, known as AdamW, which provides an alternative implementation of weight decay. Instead of adding the weight decay term to the gradient, ADAM is applied normally, then $\lambda \cdot w$ is subtracted directly from the weights themselves. This decoupling of learning rate and weight decay ensures that weight decay remains a straightforward penalty for large weights without distorting ADAM's moment estimates or affecting the adaptive learning rate scaling [30].

Lastly, Layer-wise Adaptive Moments optimizer for Batch training (LAMB) extends the methodology of AdamW by adding layer-wise scaling. After computing the usual AdamW update for each parameter tensor, LAMB rescales that update by the ratio of the weight norm to the update norm. This balances updates across layers of different sizes, preventing small layers from being overwhelmed by larger ones.

Effectively, this adjusts step sizes relative to each layer's scale (its norm), enabling stable learning even with very large batch sizes [31]. Within this thesis, LAMB is used specifically for this layer-wise balancing property. Since the parametrization involves multiple components with differing magnitude contributions, LAMB's per-layer scaling ensures balanced training across all submodels without requiring manually tuned learning rates for each component.

7 Baseline Model

To verify the functionality of the position refitting routine and provide a baseline for comparison with the parametrized architecture, a CNN model following the approach developed in Shi [32] is implemented.

This approach utilizes a generic Convolutional Neural Network (CNN) that directly provides corrections to previously developed Monte Carlo maps. The original implementation employed a scaled form of the Poisson likelihood as the loss function; however, the specific design choices and scaling methodology are detailed in the original work and will not be reproduced here. This model was subsequently reimplemented for SR1 analysis by Shi [33] and is referred to as Shi's model.

To provide a baseline comparison for the refitting routine, Shi's architecture is replicated using the updated likelihood function from Section 5 and trained on the ^{83m}Kr dataset.

Architecture The architecture consists of two main components: a Monte Carlo LCE map predictor and a convolutional correction network. The CNN applies multiplicative per-PMT corrections to the MC map, accounting for effects not captured in the simulation. The network employs a sequence of 1D convolutional layers with progressively reducing channel dimensions $[96, 48, 24, 12, 1]$, followed by an I_0 layer for per-PMT constant scaling and finally pattern renormalization.

The model utilizes multiple coordinate representations as inputs, following the original implementation. These include PMT positions in both radial (r_P, θ_P) and Cartesian (x_P, y_P) coordinates, event positions in radial coordinates (r, θ_e), symmetrized angular coordinates $(\cos(\theta_e), \sin(\theta_e))$, and event coordinates rotated to align with the anode mesh wire orientation. Additional distance measures include the minimum distance to the closest anode mesh wire, projection distance to the detector wall, and distances to the gate's supporting wires. The rationale for these specific input choices is detailed in the original work [32].

Training Procedure Training is performed in three sequential steps to ensure stable convergence: first, the CNN correction network is trained with fixed PMT noise parameters (σ values), next the PMT variances are optimized while keeping the network weights fixed, and finally both components are trained jointly. This staged approach prevents the optimization from blowing up during the initial training phases.

The model uses the LAMB optimizer for 15 epochs with a starting learning rate of 10^{-3}, dropping by one order of magnitude every 5 epochs. Training is performed on the full ^{83m}Kr calibration dataset.

Baseline Results The resulting model achieves a mean $\chi^2_{FR} = 1.238$ on the test dataset. After applying position refinement as outlined in Section 8, the performance improves too $\chi^2_{FR} = 1.225$. This establishes the baseline performance for comparison with the component-based parametrization approach.

To assess the model's performance relative to its theoretical optimum, a baseline χ^2 was established using the Monte Carlo validation method from Section 5.3, but with the PMT noise parameters learned by this model. This yields a theoretical best-fit $\chi^2_{FR} = 1.225$, indicating that the CNN model performs at its theoretical limit.

However, this analysis reveals a significant limitation: one PMT's noise parameter was optimized to the maximum allowed value of 1 PE (variances exceeding 1 photoelectron are physically unrealistic). This suggests the model failed to adequately capture the response behavior for this PMT, and instead compensated by maximizing its uncertainty. This is problematic because the likelihood function inherently penalizes high variances when the underlying response is well-modeled.

This limitation in the variance optimization highlights the need for more sophisticated modeling approaches. The following sections develop a component-based parametrization that addresses these issues by explicitly modeling the physical processes underlying the detector response. The following sections develop a component-based parametrization approach, beginning with the radial light collection efficiency—the dominant component of the detector response.

8 The Radial LCE

As a first step towards component-based modeling, the radial parametrization developed in XENON1T by Aalbers [27] is investigated to capture the leading light-collection dependence on distance. This radial LCE approach represents an empirical parametrization of the direct detection response, utilizing a function that combines inverse-square falloff with linear and constant terms to model how light collection efficiency varies with the distance between an event and each PMT center. While the Monte Carlo analysis of the anode mesh effects in Section 4 already indicates that radial symmetry assumptions will be insufficient for XENONnT's detector geometry, this investigation serves to investigate how well the radial approach performs in practice and to evaluate different optimization techniques for the subsequent full parametrized model development.

For the XENONnT implementation, two portions of Aalbers' approach are tested: first fitting a single global radial LCE function to all PMTs, then attempting individual radial LCE parameters for each PMT. The individual-PMT approach assigns each PMT its own set of radial LCE parameters, with the expectation that natural groupings may emerge during optimization. The position refitting and function refitting routine from the original method is retained for both approaches to enable iterative bias correction.

The radial LCE function adopted from XENON1T is shown in equation 17.

$$\mathrm{LCE}_i^{\text{no-norm}}(R) = I_i \left(\frac{1-b}{\left(1+\left[\frac{R}{d}\right]^2\right)^p} + aR + b \right) \tag{17}$$

Here R corresponds to the event to PMT center distance, I_i is a per PMT scaling parameter that takes into account inherent response differences between PMTs including the quantum efficiency, while a, b, d, and p are four global shape parameters. The function is not self-normalizing and must be sum normalized across all PMTs to ensure proper probability interpretation.

For the XENONnT implementation, two portions of this approach are tested: first fitting a single global radial LCE function to all PMTs, then attempting individual radial LCE parameters for each PMT. The individual-PMT approach assigns each PMT its own set of radial LCE parameters, with the expectation that natural groupings may emerge during optimization. The position refitting and function refitting routine from the original method is retained for both approaches to enable iterative bias correction. However, while the single function approach provides reasonable fits, the individual-PMT radial LCE approach proves unsuccessful during multi-parameter optimization, leading to its abandonment in favor of a full parametrized model architecture outlined in Section 9.

The single radial LCE function achieves reasonable fits using both the Poisson and full likelihood functions, though optimization requires reducing learning rates to 10^{-4} from the default 10^{-3} to accommodate varying parameter sensitivities. The Poisson likelihood exhibits multiple local minima, with only the lowest minimum avoiding introduction of additional position bias during refitting. In contrast, the full likelihood demonstrates more stable behavior by converging to a single unique minimum. However, position refitting using the radial LCE exacerbates the observed density variation along the anode mesh across all likelihood functions, likely due to the mesh's electric field effects and shadowing behavior not behaving radially symmetric, see Section 4.

Using the full likelihood function, the final $\chi^2_{FR} = 1.494$, while after position refit it is $\chi^2_{FR} = 1.478$, producing a normalized Shannon entropy of 0.9698, marginally better than the baseline positions. The investigation of the Poisson-based likelihood and an extensive comparison to the original work can be found in [34], as its relevance to the final results is limited.

Assigning individual radial LCE parameters to each PMT produces mixed results depending on the likelihood function used. Under the Poisson likelihood, this approach yields a lower minimum than the single function case, but with minimal parameter variation and no emergence of sensible PMT groupings. The full likelihood implementation proves entirely unsuccessful, failing to achieve any improvement in loss compared to the single function approach. These results lead to abandoning the individual-PMT radial LCE model in favor of the parametrized implementation.

The encoder methodology for position refinement from Aalbers' original work proves computationally inefficient for XENONnT, with events typically overshooting position minima and requiring many epochs to reconverge. While very low learning rates avoid overshooting, convergence becomes prohibitively slow. A hybrid approach is adopted as the primary optimization method, using grid search to ensure reliable minimum finding within specified resolution, followed by up to 5 single-epoch encoder iterations for post-processing to add randomization and extend the effective search range for events initially far from their true minimum.

The radial LCE investigation demonstrates that while this approach achieves reasonable fits to XENONnT data, it fundamentally cannot eliminate the density variations caused by anode mesh effects, confirming Monte Carlo predictions that radial symmetry assumptions are insufficient for the detector geometry. The single global function provides stable optimization under the full likelihood, though the individual-PMT approach proves unstable and is abandoned due to convergence failures. Additionally, the encoder position refinement methodology proves less computationally efficient than the hybrid grid search approach, likely due to improved reconstruction algorithms since XENON1T. While modeling might theoretically be possible using a superposition of radial LCE functions with separate terms for absolute x and y shifts, this approach is not pursued due to the observed instability and computational complexity. These limitations — particularly the inability to handle mesh-induced detector response variations and multi-parameter optimization instability — necessitate development of the full parametrized model architecture detailed in the following section, which addresses these fundamental constraints while providing the flexibility required for XENONnT's complex detector response.

9 The Full Model

Building on the constrained approach of the radial LCE, a more comprehensive framework is required to address the complex detector response variations observed in XENONnT. The detector's inhomogeneous response arises from multiple physical processes — direct photon detection in conjunction with electric field effects, wall reflections, and wire mesh shadowing effects — each with distinct spatial dependencies. While analytical functions could theoretically describe these individual contributions, constructing such equations without detailed prior knowledge of the behavioral forms would be prohibitively complex. An alternative approach utilizing heavily constrained neural networks is adopted to model these individual components, allowing the data to inform the functional forms.

The essential requirement for this approach is achieving isolation of physical effects by providing each component with a minimal set of input parameters unique to the contributing factors, while ensuring physically valid solutions through monotonicity constraints between input and output variables. This approach leverages the architectural structure of neural networks to create interpretable maps, where the monotonicity constraints enable physically meaningful solutions. However, care must be taken that input features are indeed monotonic with respect to the expected physical output, as the monotonicity requirement fundamentally shapes the network's behavior beyond typical neural network architectures.

The foundation for these constrained subnetworks consists of MonoDense layers, which are specialized Dense layers with enforced monotonicity constraints. In a standard Dense layer, the output can exhibit arbitrary relationships with respect to individual input variables, potentially learning non-physical behaviors such as decreasing light collection with decreasing distance.

MonoDense layers address this by constraining the weight matrix elements to maintain consistent signs relative to specific inputs, ensuring that the layer output is monotonic with respect to those designated input variables. However, simply restraining weight signs with most activation functions prevents convergence during optimization, so the MonoDense implementation utilizes a specialized approach that maintains monotonicity while allowing effective training [35]. Using these constrained layers, the full model architecture can be constructed.

Utilizing MonoDense layers as the backbone for each contributing factor, the complete LRF model is structured around three primary components representing distinct physical processes that modulate the S2 hit pattern. The architecture consists of a direct detection component that captures the fundamental light collection response including anode mesh shadowing effects, a wall reflection component modeling diffuse reflections from the TPC walls, and a perpendicular wire contribution accounting for shadowing from perpendicular wires. A schematic of the complete model architecture is shown in Figure 10. Each component utilizes one to two unique input variables designed to isolate the specific physical effect while preventing the model from reconstructing absolute event positions within the TPC. The outputs from each MonoDense subnetwork are combined through additive (direct detection, wall reflection) or multiplicative (perpendicular wire) operations to produce the final per-PMT LCE prediction. Each component's specific architecture and input variables are detailed below.

Direct Detection Component The direct detection component models the fundamental light collection response, representing the primary light path from S2 events directly to PMTs without intermediate reflections. This includes anode mesh shadowing effects investigated in Section 4.

The input parameterization employs squared coordinate differences (x^2, y^2) relative to the PMT position in a mesh-aligned coordinate system, which are expected to be monotonic with respect to the output response. The use of squared coordinate differences (rather than absolute coordinates) enforces physical symmetry by preventing the network from learning a preferred emission direction, ensuring the response remains symmetric around each PMT position as physically expected.

This component utilizes an extended MonoDense architecture with dimensions $[64, 64, 64, 64, 1]$, sized to provide sufficient model flexibility based on results from radial LCE analysis and MC anode mesh simulations. The component uses ELU activation for hidden layers with softplus final activation for positive additive contributions.

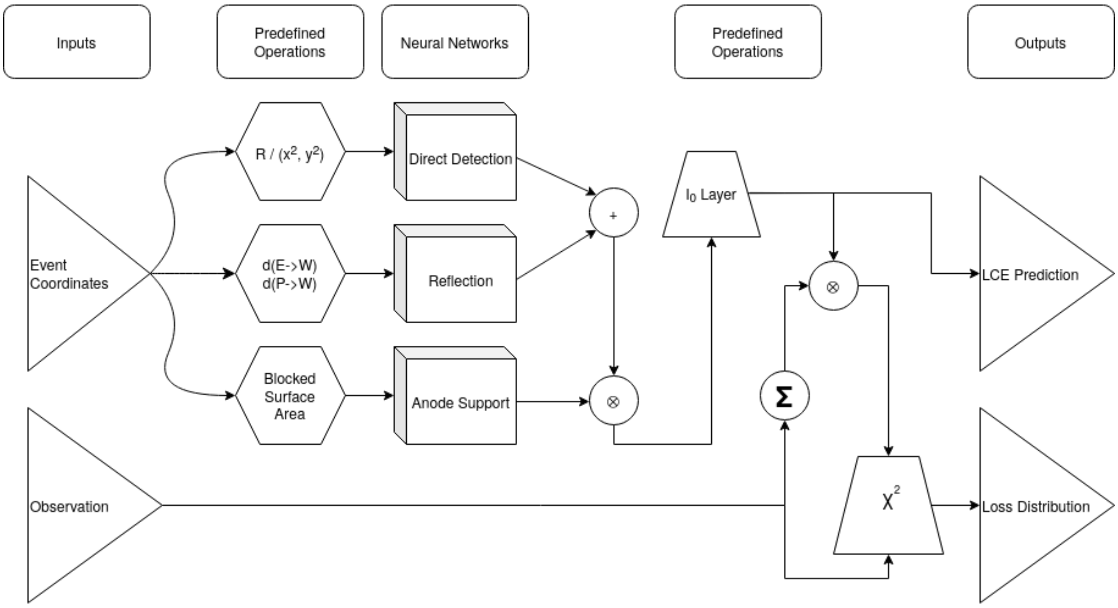


Figure 10: Full parametrized model architecture as a flow chart. Triangles correspond to inputs/outputs, rectangles to neural networks, hexagons to predefined operations, and trapezoids to predefined operations with trainable weights. $\otimes$ represents element-wise multiplication, while Σ represents per-event summation.

Wall Reflection Component The wall reflection component models diffuse light reflections from the TPC cylindrical walls. To parameterize wall reflections without allowing position reconstruction, two input variables are employed: the distance from the event to the wall reflection point, and the distance from the reflection point to the PMT. The reflection point is determined by minimizing the total path length from event to wall to PMT, providing a sensible parametrization for wall reflection effects.

Finding this optimal reflection point requires numerical optimization since no closed-form solution exists. If the event position is $\vec{E}$, the PMT position is $\vec{P}$, and the wall position is parameterized as $\vec{W}(\theta) = R_{TPC} \cdot (\cos(\theta), \sin(\theta))$, the optimization solves for the angle θ that minimizes $L(\theta) = |\vec{E} - \vec{W}(\theta)| + |\vec{W}(\theta) - \vec{P}|$.

A simple minimization routine starts with initial guess $\theta_i = (\theta_E + \theta_P)/2$, then iteratively evaluates $\theta_i \pm \Delta\theta$ and refines the solution. The step size is halved after each iteration until either maximum iterations are reached or the step becomes smaller than the specified tolerance. For computational efficiency during initial training, a tolerance, and step size of 0.1 degrees are used, which are subsequently reduced to 0.01 degrees for final optimization. The geometric setup and optimization process are illustrated in Figure 11.

The wall reflection component employs MonoDense architecture with output sizes $[16, 16, 1]$, using ELU activation for hidden layers and softplus final activation to ensure positive additive contributions. The two input distances are radially symmetric and degenerate by design, meaning multiple event-PMT position combinations can produce identical input values, thereby preventing the model from reconstructing absolute positions within the TPC while still capturing the

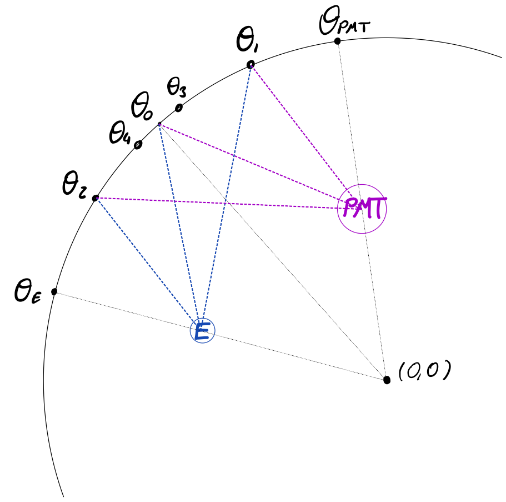


Figure 11: Wall reflection geometry showing event position E, PMT location, and the optimization algorithm's search progression. The angles $\theta_1, \theta_2, \theta_3, \theta_4$ represent successive evaluation points during the iterative minimization process to find the optimal reflection angle that minimizes the total path length.

essential wall reflection behavior.

Perpendicular Wire Component The perpendicular wire component accounts for shadowing effects from perpendicular wires that obstruct the light path between events and PMTs. To avoid providing positional information that could enable position reconstruction, the parametrization uses the surface area fraction blocked by the wires rather than direct distance or coordinate-based measures. The computation projects the light cone from the event position to the height of the perpendicular wires and calculates the extent of wire intersection in one dimension. The blocked surface area fraction is determined by summing the wire intersections within the light path and normalizing by the total projected area. Using a single-dimensional calculation significantly simplifies the computation while allowing the model to learn the appropriate scaling factors to produce physically reasonable results. However, this model assumes constant z emission coordinates, which inherently biases the results. Nonetheless, in the absence of a better parametrization, this approach was chosen as it can still provide partial corrections for most events. The geometric setup is illustrated in Figure 12.

The perpendicular wire component uses the standard MonoDense architecture $[16, 16, 1]$ with ELU activation for hidden layers and shifted tanh activation $(\tanh(x) + 1)$ for the final layer, providing multiplicative modulation factors around unity. The single input parameter ensures the model cannot reconstruct event positions while effectively capturing the wire shadowing effects.

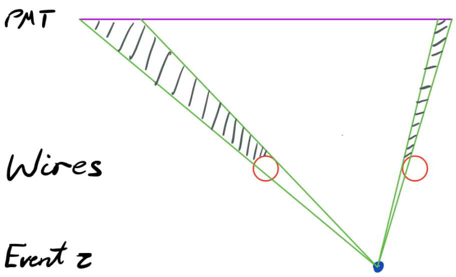


Figure 12: Perpendicular wire shadowing geometry showing the light cone projection from an event to a PMT. The hatched regions indicate wire intersections that block portions of the light path, with red circles highlighting specific wires.

9.1 Training Strategy

The training approach addresses the fundamental challenge that submodels can compensate for each other's failures, potentially creating false local minima where, for example, an overly rapid Direct Detection decay for central events is masked by an increased Wall Reflection response. To mitigate this, two independent iterative approaches were developed to refine both model and event positions until convergence to a well-fitted model is reached.

The sequential approach trains components individually (Direct Detection → Direct Detection + Wall Reflection → all components), with position refits between each stage and likelihood variance training after position refits. However, each refit biases positions to the current partial model, likely requiring multiple additional refit cycles to remove the originally introduced bias. The model training is conducted using the LAMB optimizer with high learning rates to overcome this accumulated positional bias and escape local minima. While sequential component training was initially considered, the co-evolution approach proved more computationally efficient and is the focus of this analysis.

The primary training strategy, called co-evolution, allows for all parameters to be varied in tandem. Starting from a model pretrained on synthetic MC data, the loss distribution of the krypton dataset is assessed and events symmetrically distributed about this median loss are selected for model training. These event positions are then placed as trainable parameters within the model architecture. With every component of the model trainable, it is trained for 200 epochs using LAMB with learning rate decay on a loss plateau. This method allows exploration not only of the functional parameter space that could describe the data, but also the positional parameter space. This approach is applied iteratively with new median events selected after each training and refitting routine.

A useful property of this approach is that if the model finds consistently good solutions even when starting with high learning rates that move away from the original minima, one can be fairly certain that the model is sufficiently constrained to describe the dataset without much redundancy. One potential pitfall is faulty shifting of event positions early in training. If too many events are shifted too far from their original minima, the model may not be capable of restoring the original or correct positions.

9.2 Results

Two training strategies were implemented: sequential component fitting and co-evolution. While early tests showed both approaches converged to similar minima, co-evolution proved computationally faster and more efficient. Due to time constraints, results focus on the co-evolutionary approach, though the consistency between methods provides confidence in the solution quality.

The full model is first pretrained on synthetic data derived from the MC LCE map, processed using the same methodology as the loss function investigation in Section 5.3. This enables rapid pretraining on clean, well-behaved data that approximates the detector response. Training used $10,000$ events with a learning rate scheduler on LAMB decreasing from 10^{-3} to $\mathcal{O}(10^{-6})$ over 100 epochs. The pretrained model achieves $\chi^2_{FR} = 1.23$ on the synthetic training set, compared to $\chi^2_{FR} = 1.18$ when using the MC map directly. When evaluated on real data the pretrained model yields $\chi^2_{FR} = 2.18$ versus $\chi^2_{FR} = 2.30$ for the direct MC map, demonstrating successful transfer from synthetic to real data. All results so far use constant standard deviation $\sigma = 0.5$ PE.

Starting from the MC pretrained model, $2^8 \times 50$ events with losses closest to the median are selected from the real dataset. These event positions are converted to trainable parameters and optimized simultaneously with the network weights and PMT variances, followed by position refinement using both grid search and encoder methods. Training uses the LAMB optimizer with learning rate 10^{-2} decaying to a minimum of $\mathcal{O}(10^{-6})$ over 50 epochs based on loss plateau detection. This coevolution procedure is repeated three additional times, with new median event selection before each iteration based on the current model's performance. The progressive loss evolution across iterations is shown in Table 2.

Table 2: Progressive χ^2_{FR} evolution through coevolution iterations including position refits evaluated over the entire Krypton dataset. Original positions refer to the initial event (CNF) positions, refit positions are the result of the respective position refit.

Model	Original Positions	Refit Positions
MC pretrained	2.301	–
Run 1	1.280	1.284
Run 2	1.237	1.223
Run 3	1.254	1.237
Run 4	1.255	1.237
radial LCE	1.494	1.478
CNN Baseline	1.238	1.225

From the loss progression, the model achieves optimal performance at Run 2, with subsequent iterations converging to suboptimal solutions. While position refinement continues to improve relative to each retrained model, the overall loss degrades in later runs, indicating that the coevolution procedure begins to deteriorate. This suggests the event positions utilized in model training shift too far from their true locations, preventing the model from finding better minima. The identification of Run 2 as optimal is further supported by the spatial homogeneity metrics presented in Table 3.

Table 3: Spatial homogeneity metrics for coevolution iterations compared to original positions and CNN baseline, Section 7. All events are fiducialized (2D radius $<$ 66.4 cm) and evaluated using 300 spatial bins using the homogeneity metrics defined in Section 3.9. Lower CV, Gini coefficients and empty bins indicate better spatial uniformity, while higher entropy indicates a more homogeneous distribution.

Method	CV	Shannon Entropy	Norm. Entropy	Gini Coeff.	Empty Bins (%)	Events
Original Positions	0.7295	10.821	0.9691	0.4102	17.75	1,642,400
Run 1	0.6920	10.850	0.9717	0.3894	14.67	1,647,237
Run 2	0.6793	10.856	0.9722	0.3816	14.72	1,647,680
Run 3	0.6965	10.848	0.9715	0.3933	14.53	1,648,289
Run 4	0.7080	10.843	0.9711	0.4002	14.47	1,648,563
radial LCE	0.6968	10.851	0.9718	0.3921	14.84	1,643,863
CNN Baseline	0.7110	10.829	0.9698	0.4021	17.09	1,642,267

From these results, Run 2 corresponds to the best performing model, with subsequent iterations showing degraded performance across all homogeneity metrics. Since both likelihood and spatial uniformity metrics consistently identify Run 2 as optimal, this model is selected as the final solution. The only metric where Run 2 does not achieve the best value is the total event count within the detector volume, but with variations of only $\sim$ 1000 events and no clear expectation for this metric, this difference is considered insignificant.

The spatial uniformity is further demonstrated in Figure 13, which shows event density as a function of R^2. The refit positions show improved homogeneity compared to the original positions, with reduced oscillation throughout most of the detector volume. The sharp dip at low R^2 corresponds to the exclusion zone created by the perpendicular wires, where event density drops as expected. The distribution remains non-uniform near the detector walls (high R^2), which is also expected due to electric field effects at the boundaries.

Investigation of individual layers reveals that the perpendicular wire parametrization produces constant output for all inputs, showing that a constant z based geometric investigation is insufficient to produce a reasonable output for S2 events produced over variable z. Since this constant output effectively cancels out during the normalization step of the LRF calculation, this component contributes no meaningful modulation to the final detector response and can be removed from the final model.

Parametrization Residuals To investigate the resulting parameterization beyond individual component behavior, individual PMT modulation is first examined by comparing model predictions to real data. The median absolute difference between normalized real and predicted responses is 0.04%, indicating excellent agreement. To further validate performance, a new χ^2_{FR} baseline is established following the methodology of Section 5.3. This produces $\chi^2_{FR} = 1.268$.

Notably, the trained model achieves $\chi^2_{FR} = 1.223$ for the same trained variances, which is lower than its theoretical baseline. With approximately 1.65 million events, this apparent improvement likely reflects limitations in the likelihood function formulation.

The Gaussian approximation ignores the complex PMT response characteristics, particularly the

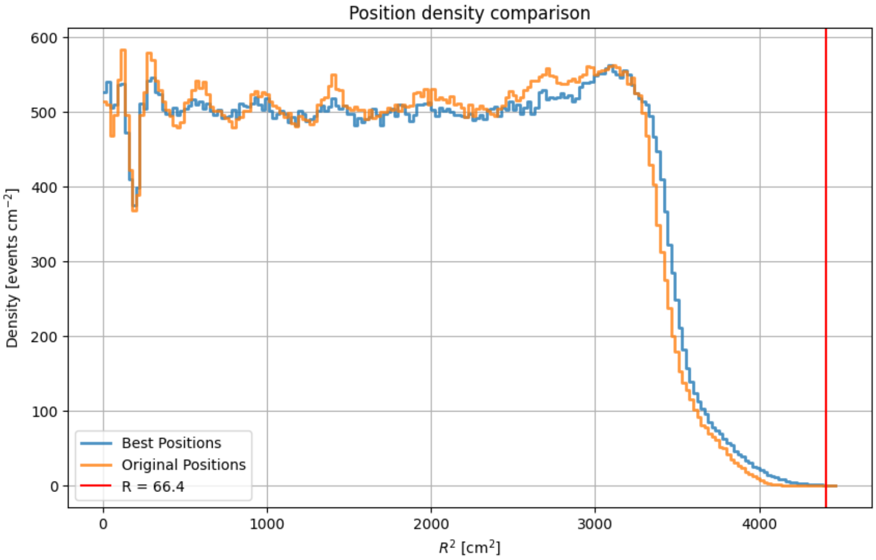


Figure 13: Event density as a function of squared radius for the best fit positions (Run 2 LRF model refit) and the original (CNF) positions of the krypton dataset.

heavy-signal tail observed in PMT calibration studies. This systematic underestimation of the true response variance likely inflates the variance parameters, artificially reducing the χ^2_{FR} value.

While the overall statistical performance appears strong, examining the spatial residuals reveals systematic patterns. Figure 14 shows the spatial modulation of LCE patterns for several selected PMTs after dividing out the predicted parameterized response.

The figure reveals asymmetric response patterns perpendicular to the anode mesh, where one side shows larger response than the other, though the direction varies inconsistently between PMTs. The modulation magnitude is significant, suggesting systematic detector effects not captured by the parameterization.

Several mechanisms could explain the observed asymmetric patters.

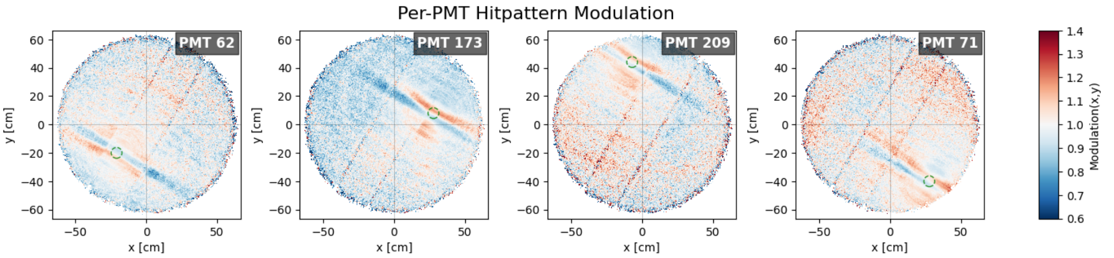


Figure 14: Spatial modulation patterns for selected PMTs after dividing out the predicted parameterized LRF response, showing the residual component not modeled in the LRF parametrization as a function of space per PMT.

The mesh periodicity (0.5 cm) and PMT spacing center (6.94 cm) do not align evenly, creating different mesh-to-PMT arrangements for each PMT row. This geometric mismatch could explain the variable offset between PMT centers and the observed reduction/enlargement bands.

This observed anisotropy may not reflect true detector asymmetry but could arise from the division process itself. When dividing one monotonic function by another (analogous to dividing real response by predicted LCE), the result is not guaranteed to be monotonic, potentially creating artificial patterns, as observed in Section 4. For example, dividing x^3 by x yields x^2, transforming monotonic functions into non-monotonic ones.

However, neither explanation fully accounts for all observed features. Most notably, PMTs 62 and 71 located in the same row but separated only by perpendicular wires show opposite enlargement directions, which challenges both the geometric periodicity and mathematical artifact hypotheses. The current dataset remains insufficient to definitively establish the underlying mechanism.

Comparison: Anode Mesh Simulation To assess the validity of the anode mesh simulation from Section 4 and verify whether the probabilistic S2 emission model agrees with real data, the direct detection component extracted from the parametrized model is compared against MC predictions using the 7-PMT analysis generated in the same way as for the MC data. Figure 15 shows the resulting LCE function for the central PMT.

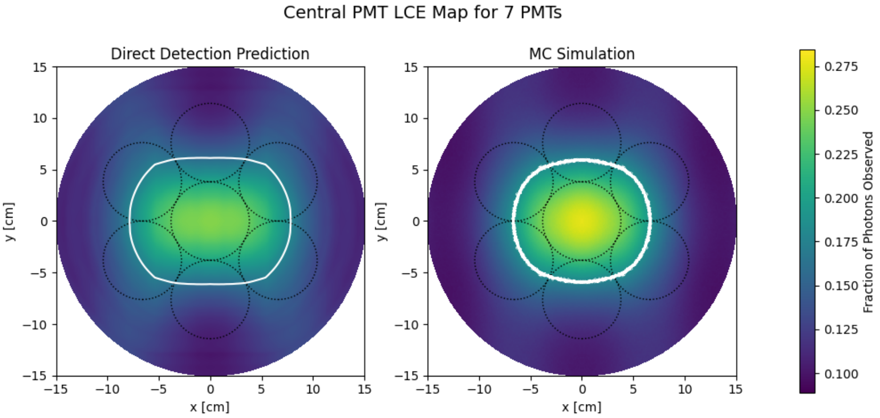


Figure 15: LCE response for the central PMT from direct detection prediction, generated by predicting the upper right quadrant and replicating symmetrically as in Figure 9. The 33% maximum signal contour is shown in white, with PMTs under consideration outlined in black.

Qualitative comparison with the Monte Carlo simulation demonstrates good agreement. The 50% maximum signal contour shows x/y extent ratios of ~ 0.82 (MC) vs ~ 0.74 (Direct Detection prediction). The Root-Mean-Squared error between maps is 0.013, with maximum differences $+0.033/-0.036$. While the MC map produces higher response within the 33% contour and the LCE prediction dominates outside it, this systematic difference is expected given the mock S2 yield weighting used in the MC simulation. Implementation of the probabilistic emission distribution in the full XENON simulation suite should resolve the discrepancy currently observed between real data and Monte Carlo predictions.

To explain this residual effects seen in real data, further analysis is required. Overall, the parametrized architecture has been successful, effectively modeling the expected response, while providing the model freedom during training simultaneously vary its parameters and the event positions used for training. While the parametrization leaves unexplained residual, the framework now enables targeted analysis of the asymmetric patterns observed about PMT centers.

The parametrized architecture successfully demonstrates that the coevolution routine is viable for simultaneous optimization of network parameters and event positions. The key finding is that Run 2 provides optimal performance, with subsequent degradation likely attributable to the learning rate schedule — high learning rates should initially be applied only to model parameters, not event positions, to prevent premature divergence from current true minima prior to establishment of optimizer state.

The coevolution approach produces a high-quality LCE model that outperforms the unconstrained baseline in terms of homogeneity. Position refinements using the parametrized model show improved spatial distributions compared to the unconstrained approach, demonstrating that the heavy architectural constraints successfully generate a more physically correct LCE representation. While also allowing for evolving both the original positions and model architecture simultaneously.

Although unexplained residual patterns remain around PMT centers, the framework enables targeted analysis of detector features and provides a solid foundation for future detector characterization studies.

Model Deployment Attempts Attempts were made to use the final parametrized model's position refit as a foundation for training a new unrestricted network for production use. Several architectures were tested, all based on the coordinate-transformed input from the model outlined in Section 7. Trials included both predefined and trainable variance approaches. Unfortunately, no model achieved better performance than the parametrized model. The best performing unrestricted model achieved only $\chi^2_{FR} = 1.280$ on the refit dataset and $\chi^2_{FR} = 1.298$ on the original dataset, significantly worse than the parametrized architecture and the CNN baseline.

The performance limitation appears to stem from the network's inability to internally model the wall reflection component. To address the computational bottleneck of the wall reflection optimization, separate attempts were made to train neural networks to directly predict the shortest path distances, accepting slightly decreased precision in exchange for faster evaluation. However, these efforts required prohibitively large architectures compared to the target model size. This indicates that the wall reflection optimization represents a computational complexity that cannot be adequately captured by networks of practical size, preventing the unrestricted architecture from accessing this critical information that the parametrized model receives as explicit input.

Given that no tested architecture could provide comparable performance and the wall reflection optimization has no viable neural network replacement, development of a smaller, purpose-built production model is not pursued further in this thesis.

10 Discussion & Future Work

This work has developed and validated a physics-informed machine learning architecture for light collection efficiency modeling in liquid xenon time projection chambers, demonstrating the viability of coevolution training for simultaneous optimization of model parameters and event positions. To provide a comprehensive overview of the performance achievements across different approaches, Table 4 summarizes the χ^2_{FR} metrics alongside the normalized Shannon entropy for the major outcomes of methods and refits investigated throughout this study.

The results demonstrate significant improvements over baseline approaches, with the parametrized coevolution method achieving good fit quality while leading in homogeneity metrics and maintaining physically interpretable constraints.

Table 4: χ^2_{FR} performance comparison across all investigated methods final models evaluated on the full Krypton dataset alongside the respective position refits and their normalized entropy statistics (1 is perfectly homogeneous). The CNN baseline corresponds to a generic neural network approach to fitting the LCE function. The radial LCE corresponds to a functional parametrization of the direct light detection component. And lastly the parametrized best fit corresponds to the constrained machine learning approach developing a separable response for each detector component.

	χ^2_{FR}		Normalized
Method	Original Positions	Refit Positions	Shannon Entropy
CNF Positions	–	–	0.9691
MC LCE Map	2.30	–	–
CNN Baseline	1.238	1.225	0.9698
Radial LCE	1.494	1.478	0.9718
Parametrized Best Fit	1.237	1.223	0.9722

Monte Carlo Validation The Monte Carlo anode mesh simulation successfully identified the physical mechanism behind directional anisotropies in the S2 optical response. The combination of variable emission distribution in z-depth, mock photon yield weighting, and electron convergence onto specific anode wires produces the observed modulation patterns in per-PMT response.

Comparison between Monte Carlo predictions and the direct detection parametrization showed good agreement, with systematic differences primarily attributable to the mock z-weighting scheme. While uncertainty remains regarding full capture of realistic field effects, the demonstrated agreement validates the underlying framework. Integration of this emission model into the full XENON simulation will significantly improve Monte Carlo based S2 light collection efficiency predictions for future analyses.

Coevolution Training: Success & Limitations The overall position refinement routine was successful in producing more accurate event positions, as demonstrated by improved χ^2 and homogeneity metrics. The monotonic dense layers combined with direct detection and wall reflection components created a self-regulating model that enabled effective coevolution training.

The perpendicular wire parametrization suffers from a fundamental geometric oversimplification: it computes shadowing effects assuming fixed z-coordinates, when events actually occur at variable depths throughout the drift region. This means that events experiencing wire shadowing are not properly penalized, as the fixed-z calculation fails to account for their actual three-dimensional position relative to the wire structure. For proper implementation, a parametrization that accounts for variable depth emission must be developed.

Technical Limitations The current likelihood function represents a simplified approximation of the true PMT response, omitting several important physical mechanisms that limit model accuracy. While computationally efficient, the Gaussian approximation fails to capture the heavy-tailed gain behavior observed in PMT calibration studies, particularly the high-signal tail that significantly affects response variance estimates. This foundational limitation affects all subsequent modeling efforts.

The current codebase suffers from accumulated technical debt that significantly impacts training efficiency and convergence stability. All network layers have become unnecessarily complex through iterative development, creating overly intricate computational graphs that likely impede TensorFlow's optimization capabilities and slow backpropagation. A complete architectural rewrite would eliminate this complexity and likely improve training performance.

The likelihood function implementation presents another critical bottleneck. The current pure TensorFlow implementation struggles with XLA compilation, particularly for per-element conditional evaluations that are essential for accurate PMT response modeling. Replacing this with TensorFlow's Python call method to invoke optimized Numba or Cython implementations, see Appendix A, would provide substantial performance improvements while maintaining computational accuracy.

These implementation challenges manifested as significant convergence issues during training. First, attempts to implement a multi-function radial LCE model consistently failed to achieve lower minima compared to the single-function approach. This failure likely stems from the loss landscape complexity and the challenge of optimizing weight matrices containing coefficients with vastly different orders of magnitude within the same parameter group.

Second, the coevolution training exhibited diminishing returns after Run 2, where subsequent iterations could not converge to improved minima despite continued optimization. The high learning rates employed to encourage parameter space exploration appear to have had the opposite effect once event positions shifted beyond the optimal region, creating a scenario where the optimizer became trapped in suboptimal configurations.

The wall reflection parametrization suffers from systematic biases that compromise its accuracy, particularly near detector boundaries where reflection effects are most pronounced. The initial fitting procedure employed a high angular tolerance ($\theta = 0.1$ degrees), which introduces significant errors in individual ray path calculations. This bias is compounded by the coevolution routine's event selection strategy, which samples median-performing events during training, systematically excluding wall events where reflection effects are strongest.

Future Research Directions Several enhancements can be implemented without fundamental architectural changes. The most readily implementable involves developing separate direct detection functions per PMT row to investigate potential differences related to anode mesh alignment. This could be achieved using multiple monotonic dense models or by utilizing a superposition of radial LCE functions in x and y coordinates. However, given the results of the multi-radial LCE fit, convergence may prove problematic.

Alongside this, a more representative perpendicular wire shadowing parametrization would be worthwhile, with a main focus on finding a parametrization that not only allows for spatial reconstruction, but also includes information representative of the emission depth variation in each event.

To improve model performance, the current likelihood function requires extension beyond the Gaussian approximation to capture the heavy high-signal tail observed in PMT calibration data. This enhancement should align the expected mean signal per photoelectron with model-independent gain calibration results, either through improved statistical distributions or by explicitly modeling the individual physical processes for a gain-calibration-independent approach.

The parametrized LCE in its current state does not separate electric and optical effects. To achieve a purely optical description, a complete model of the electric field within the TPC is required. While conventional fitting routines would prove inadequate for this complexity, constrained conditional normalizing flow (CNF) models offer a promising alternative.

This approach would model source electron clouds around event coordinates and propagate them to final S2 emission distributions, enabling realistic application of purely optical detection parametrizations which could be pretrained directly on Monte Carlo simulations, including an anode shadowing parametrization of the same form as the current perpendicular wire shadowing parametrization. Such a framework would properly define S2 event positions prior to anode gate mesh electric field modulation and provide valuable validation for Monte Carlo simulations and S2 field distortion corrections. However, this methodology requires substantial computational resources for three-dimensional electron cloud evaluation and significant development work to validate the relatively novel CNF architecture for constrained detector physics applications.

Conclusion Nonetheless, the approach was a success, producing the best S2 LCE map to date for XENONnT. The combination of parametrized architecture training, coevolution position refinement, and final dense network fitting achieved unprecedented accuracy. Additionally, the introduction of a more complete likelihood function and optimization of its parameters further reduced discrepancies between predicted and observed detector response.

This improved S2 LCE can directly enhance XENONnT's sensitivity to both dark matter and neutrinoless double beta decay searches by providing more accurate event localization and subsequently allowing for better fiducialization of the detector and energy corrections.

The coevolution training methodology represents a novel approach to detector calibration, demonstrating how iterative position refinement in conjunction with constrained dense networks can overcome limitations in traditional neural fitting procedures—a technique applicable to other dual phase TPC experiments and potentially other precision measurement systems.

The successful validation through Monte Carlo simulation confirms that the observed improvements stem from genuine physics understanding rather than overfitting, providing confidence for production deployment. Furthermore, the feasibility of separating detector effects into individual components opens a pathway toward purely optical LCE models, enabling data-driven electric field modeling and improved systematic uncertainty control for future generations of dual phase TPCs.

References

[1] 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`.

[2] E. Aprile et al. “Double-weak decays of Xe 124 and Xe 136 in the XENON1T and XENONnT experiments”. In: *Physical Review C* 106.2 (Aug. 2022), p. 024328. DOI: `10.1103/PhysRevC.106.024328`.

[3] 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, Vol. 6, p. 249* 6 (Aug. 1932), p. 249.

[4] V. C. Rubin. “Dark matter in spiral galaxies”. In: *Scientific American* 248.6 (June 1983), pp. 96–109.

[5] S. Navas et al. “Review of Particle Physics”. In: *Physical Review D* 110 (3 Aug. 2024), p. 030001. DOI: `10.1103/PhysRevD.110.030001`.

[6] G. Bertone and T. Tait. “A new era in the search for dark matter”. In: *Nature* 562.7725 (Oct. 2018), pp. 51–56. DOI: `https://doi.org/10.1038/s41586-018-0542-z`.

[7] M. Dolinski, A. Poon, and W. Rodejohann. “Neutrinoless double-beta decay: status and prospects”. In: *Annual Review of Nuclear and Particle Science* 69.1 (Aug. 2019), pp. 219–251.

[8] M. Haffke et al. “Background measurements in the gran sasso underground laboratory”. In: *Nuclear Instruments and Methods in Physics Research* 643.1 (July 2011), pp. 36–41. DOI: `10.1016/j.nima.2011.04.027`.

[9] XENON Collaboration et al. “The XENONnT dark matter experiment”. In: *European Physical Journal C* 84 (June 2024), p. 784. DOI: `10.1140/epjc/s10052-024-12982-5`.

[10] A. Minamino et al. “Self-shielding effect of a single phase liquid xenon detector for direct dark matter search”. In: *Astroparticle Physics* 35.10 (May 2012), pp. 609–614. DOI: `10.1016/j.astropartphys.2012.01.005`.

[11] E. Aprile et al. “Observation of two-neutrino double electron capture in 124Xe with XENON1T”. In: *Nature* 568.7753 (Apr. 2019), pp. 532–535. ISSN: 1476-4687. DOI: `10.1038/s41586-019-1124-4`.

[12] 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`.

[13] XENON Collaboration. *XENONnT Infrastructure.* `https://xenonexperiment.org/infrastructure/`. Accessed: 2025-05-15.

[14] V. C. Antochi et al. “Improved quality tests of R11410-21 photomultiplier tubes for the XENONnT experiment”. In: *Journal of Instrumentation* 16.08 (Aug. 2021), P08033. DOI: `10.1088/1748-0221/16/08/P08033`.

[15] C. Faham et al. “Measurements of wavelength-dependent double photoelectron emission from single photons in VUV-sensitive photomultiplier tubes”. In: *Journal of Instrumentation* 10.09 (Sept. 2015), P09010. DOI: `10.1088/1748-0221/10/09/P09010`.

[16] J. Qi. *Comparison Between WFSim and Data for Single Electrons.* Internal Note `xenon:xenonnt:jianyang:wfsim_se_comparison`. XENON Collaboration, Dec. 2021. URL:

https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:jianyang:wfsim_se_comparison.

[17] A. Ravindran and F. Pompa. *Krypton Selection for SR2 - v17.* Internal Note `xenon:xenonnt:analysis:kr83m_sr2_v17`. XENON Collaboration, Feb. 2025. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:analysis:kr83m_sr2_v17.

[18] R. Saldanha et al. "Model independent approach to the single photoelectron calibration of photomultiplier tubes". In: *Nuclear Instruments and Methods in Physics Research* 863 (Aug. 2017), pp. 35–46. DOI: 10.1016/j.nima.2017.02.086.

[19] E. Aprile et al. "XENONnT analysis: Signal reconstruction, calibration, and event selection". In: *Physical Review D* 111.6 (Mar. 2025), p. 062006. DOI: 10.1103/PhysRevD.111.062006.

[20] D. Ramírez García. "Simulating the XENONnT dark matter experiment: backgrounds and WIMP sensitivity". PhD thesis. Dissertation, Universität Freiburg, Apr. 2022. DOI: 10.6094/UNIFR/228338.

[21] S. Althüser, B. Andrieu, and O. Dadoun. *Light Collection Efficiency Map for S2 Simulation.* Internal Note `xenon:xenonnt:andrieu:lce_maps_s2_simulation`. XENON Collaboration, Mar. 2023. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:andrieu:lce_maps_s2_simulation.

[22] J. Qin and S. Vetter. *Using Conditional Normalizing Flows for position reconstruction.* Internal Note `xenon:xenonnt:svetter:conditional_norm_flow_posrec_principle`. XENON Collaboration, July 2024. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:svetter:conditional_norm_flow_posrec_principle.

[23] F. Amaro. *Ion - drifting electron recombination in 83mKr calibrations and impact on the electron lifetime estimation.* Internal Note `xenon:amaro:e_lifetime_atempt`. XENON Collaboration, Oct. 2019. URL: https://intranet.xenonnt.org/notes/e_lifetime_atempt.pdf.

[24] J. Aalbers. *Improved pattern likelihood variable.* Internal Note `xenon:xenon1t:aalbers:improved_pattern_likelihood`. XENON Collaboration, Oct. 2019. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenon1t:aalbers:improved_pattern_likelihood.

[25] S. S. Wilks. "The large-sample distribution of the likelihood ratio for testing composite hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (Mar. 1938), pp. 60–62.

[26] M. Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.* Software available from tensorflow.org. Mar. 2016. DOI: 10.48550/arXiv.1603.04467. URL: https://www.tensorflow.org/.

[27] J. Aalbers. *Improving S2 position reconstruction by fitting LCE maps.* Internal Note `xenon:xenon1t:aalbers:lce_fit`. XENON Collaboration, July 2019. URL: https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenon1t:aalbers:lce_fit.

[28] H. Robbins and S. Monro. "A stochastic approximation method". In: *The Annals of Mathematical Statistics* (Sept. 1951), pp. 400–407. DOI: 10.1214/aoms/1177729586.

[29] D. P. Kingma and J. Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (Dec. 2014). DOI: 10.48550/arXiv.1412.6980.

[30] I. Loshchilov and F. Hutter. "Decoupled weight decay regularization". In: *arXiv preprint arXiv:1711.05101* (Nov. 2017).

[31] Y. You et al. “Large batch optimization for deep learning: Training bert in 76 minutes”. In: *arXiv preprint arXiv:1904.00962* (Apr. 2019).

[32] S. Shi. *S2 Pattern: Map and likelihood on Neural Network*. Internal Note `shenyang:s2_pattern_likelihood`. XENON Collaboration, Nov. 2021. URL: `https://xe1t-wiki.lngs.infn.it/doku.php?id=shenyang:s2_pattern_likelihood`.

[33] S. Shi. *XENONnT SR1 S2 Pattern Likelihood Training*. Internal Note `xenon:xenon1t:jacques:nt_sr1_s2pl`. XENON Collaboration, Oct. 2023. URL: `https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenon1t:jacques:nt_sr1_s2pl`.

[34] F. Semler. *Radial LCE Applied to XENONnT*. Internal Note `xenon:xenonnt:analysis:analysts_overview_page:radial_lce_applied_to_nt`. XENON Collaboration, Jan. 2025. URL: `https://xe1t-wiki.lngs.infn.it/doku.php?id=xenon:xenonnt:analysis:analysts_overview_page:radial_lce_applied_to_nt`.

[35] D. Runje and S. M. Shankaranarayana. *Constrained Monotonic Neural Networks*. May 2023. DOI: `10.48550/arXiv.2205.11775`.

Appendices

A Implementation

Code & Data Availability To replicate this work, code is available in a GitHub repository. Data availability alongside replication notes are present in the repository's README files.

Likelihood Function The likelihood function with lookup methods was implemented in Numba, TensorFlow, and JAX, with an additional Cython implementation for exact likelihood evaluation only.

Evaluation times for the exact likelihood with conditional switching to Gaussian approximation were measured across all four implementations. For Numba and Cython, parallelization was implemented using multiple threads at the array level, allowing parallel computation across multiple data rows. JAX and TensorFlow handle parallelization automatically through the XLA framework. Performance was evaluated over 5 iterations using $200,000$ random elements representative of typical pattern evaluation workloads. Results are presented in Table 5.

Table 5: Evaluation times for different implementations of the exact likelihood function.

	Cython	JAX	Numba	TensorFlow
Time (s)	0.007	4.711	0.006	0.022
Standard Deviation (s)	0.001	0.044	0.000	0.002

JAX performs notably poorly, though it is unclear whether this stems from machine-specific issues or implementation limitations. During development, efforts were made to adhere to JAX conventions and ensure the generated XLA representation aligned with TensorFlow's approach. However, neither JAX nor TensorFlow handles conditional dispatch between exact and Gaussian likelihood evaluations efficiently, as XLA compilation is optimized for linear algebra operations rather than per-element conditional evaluation.

TensorFlow ranks as the second-slowest implementation, consistent with XLA's limitations for element-wise conditional logic. Cython, while faster than both JAX and TensorFlow, performs surprisingly slower than Numba, likely due to implementation constraints. The Cython version suffers from limited functionality since closures are not permitted in pure Cython code, requiring a pure Python function for dispatcher generation. While this limitation could be resolved, the complexity of Cython's syntax proved more challenging than anticipated.

Numba emerges as the optimal choice, achieving approximately $4\times$ faster execution than TensorFlow while maintaining implementation simplicity and code clarity.

Implementation of caching for Numba compiled functions further enhances performance. By executing a single array call during initialization, all functions can be stored in precompiled form within the analysis environment, effectively converting Numba's just-in-time compilation into ahead-of-time compilation. This contrasts with TensorFlow, where functions must be JIT compiled at runtime, introducing additional overhead.

These benchmarks demonstrate that Numba provides the optimal balance of performance and implementation simplicity for likelihood evaluation, achieving a 4× speed up over TensorFlow's exact likelihood implementation while maintaining code clarity and numerical stability.

B Tables

RunIDs

RunID	mode	end	start	source	livetime
058652	tpc_kr83m	2024-01-31 11:09:09.840000	2024-01-31 10:39:06.420000	kr-83m	00:30:03.420000
058651	tpc_kr83m	2024-01-31 10:38:15.402000	2024-01-31 10:08:12.331000	kr-83m	00:30:03.071000
058645	tpc_kr83m	2024-01-31 09:05:32.617000	2024-01-31 08:35:29.223000	kr-83m	00:30:03.394000
058640	tpc_kr83m	2024-01-31 08:34:38.839000	2024-01-31 08:04:35.713000	kr-83m	00:30:03.126000
058638	tpc_kr83m	2024-01-31 07:32:50.851000	2024-01-31 07:02:47.648000	kr-83m	00:30:03.203000
058635	tpc_kr83m	2024-01-31 06:00:09.266000	2024-01-31 05:30:06.354000	kr-83m	00:30:02.912000
058632	tpc_kr83m	2024-01-31 05:29:15.824000	2024-01-31 04:59:12.464000	kr-83m	00:30:03.360000
058628	tpc_kr83m	2024-01-31 03:25:36.657000	2024-01-31 02:55:33.611000	kr-83m	00:30:03.046000
058624	tpc_kr83m	2024-01-31 02:23:45.435000	2024-01-31 01:53:42.375000	kr-83m	00:30:03.060000
058621	tpc_kr83m	2024-01-31 00:51:03.584000	2024-01-31 00:21:00.388000	kr-83m	00:30:03.196000
058614	tpc_kr83m	2024-01-30 22:16:26.710000	2024-01-30 21:46:23.822000	kr-83m	00:30:02.888000
058609	tpc_kr83m	2024-01-30 20:43:45.306000	2024-01-30 20:13:42.011000	kr-83m	00:30:03.295000
058608	tpc_kr83m	2024-01-30 20:12:51.150000	2024-01-30 19:42:47.995000	kr-83m	00:30:03.155000
058607	tpc_kr83m	2024-01-30 19:41:57.283000	2024-01-30 19:11:54.421000	kr-83m	00:30:02.862000
058601	tpc_kr83m	2024-01-30 17:38:21.848000	2024-01-30 17:08:18.619000	kr-83m	00:30:03.229000
058600	tpc_kr83m	2024-01-30 17:07:27.911000	2024-01-30 16:37:24.769000	kr-83m	00:30:03.142000
058510	tpc_kr83m	2024-01-25 20:18:56.792000	2024-01-25 19:48:53.059000	kr-83m	00:30:03.733000
058509	tpc_kr83m	2024-01-25 19:47:59.039000	2024-01-25 19:17:55.769000	kr-83m	00:30:03.270000
058502	tpc_kr83m	2024-01-25 17:13:20.395000	2024-01-25 16:43:19.978000	kr-83m	00:30:00.417000
058497	tpc_kr83m	2024-01-25 14:38:38	2024-01-25 14:08:34.887000	kr-83m	00:30:03.113000
058494	tpc_kr83m	2024-01-25 14:07:40.323000	2024-01-25 13:37:36.670000	kr-83m	00:30:03.653000
058490	tpc_kr83m	2024-01-25 12:04:02.241000	2024-01-25 11:33:59.031000	kr-83m	00:30:03.210000
058476	tpc_kr83m	2024-01-25 09:20:28.185000	2024-01-25 08:50:25.543000	kr-83m	00:30:02.642000
058472	tpc_kr83m	2024-01-25 08:49:35.013000	2024-01-25 08:19:31.643000	kr-83m	00:30:03.370000
058465	tpc_kr83m	2024-01-25 06:14:52.468000	2024-01-25 05:44:49.394000	kr-83m	00:30:03.074000
058461	tpc_kr83m	2024-01-25 05:13:02.514000	2024-01-25 04:42:59.274000	kr-83m	00:30:03.240000
058460	tpc_kr83m	2024-01-25 04:42:05.504000	2024-01-25 04:12:02.117000	kr-83m	00:30:03.387000
058459	tpc_kr83m	2024-01-25 04:11:08.236000	2024-01-25 03:41:05.231000	kr-83m	00:30:03.005000
058458	tpc_kr83m	2024-01-25 03:40:10.995000	2024-01-25 03:10:07.841000	kr-83m	00:30:03.154000
058451	tpc_kr83m	2024-01-25 01:05:33.604000	2024-01-25 00:35:30.217000	kr-83m	00:30:03.387000
058444	tpc_kr83m	2024-01-24 22:30:51.658000	2024-01-24 22:00:48.083000	kr-83m	00:30:03.575000
058437	tpc_kr83m	2024-01-24 19:56:12.112000	2024-01-24 19:26:09.104000	kr-83m	00:30:03.008000
058435	tpc_kr83m	2024-01-24 18:54:16.765000	2024-01-24 18:24:13.424000	kr-83m	00:30:03.341000
058433	tpc_kr83m	2024-01-24 17:52:25.382000	2024-01-24 17:22:22.344000	kr-83m	00:30:03.038000
058430	tpc_kr83m	2024-01-24 17:20:17.023000	2024-01-24 16:50:13.953000	kr-83m	00:30:03.070000
058429	tpc_kr83m	2024-01-24 16:49:19.801000	2024-01-24 16:19:16.318000	kr-83m	00:30:03.483000
058427	tpc_kr83m	2024-01-24 15:47:28.840000	2024-01-24 15:17:25.546000	kr-83m	00:30:03.294000
058425	tpc_kr83m	2024-01-24 14:45:33.874000	2024-01-24 14:15:30.435000	kr-83m	00:30:03.439000
058421	tpc_kr83m	2024-01-24 13:43:42.478000	2024-01-24 13:13:39.115000	kr-83m	00:30:03.363000
058418	tpc_kr83m	2024-01-24 12:10:50.749000	2024-01-24 11:40:47.768000	kr-83m	00:30:02.981000
058134	tpc_kr83m	2024-01-09 22:33:19.632000	2024-01-09 22:19:02.027000	kr-83m	00:14:17.605000
058127	tpc_kr83m	2024-01-09 20:14:07.108000	2024-01-09 19:44:06.215000	kr-83m	00:30:00.893000
058121	tpc_kr83m	2024-01-09 18:10:28.533000	2024-01-09 17:40:25.124000	kr-83m	00:30:03.409000
058120	tpc_kr83m	2024-01-09 17:39:35.309000	2024-01-09 17:09:32.264000	kr-83m	00:30:03.045000
058113	tpc_kr83m	2024-01-09 15:04:43.929000	2024-01-09 14:34:40.732000	kr-83m	00:30:03.197000
058112	tpc_kr83m	2024-01-09 14:33:43.473000	2024-01-09 14:03:41.922000	kr-83m	00:30:01.551000
058107	tpc_kr83m	2024-01-09 13:01:00.927000	2024-01-09 12:31:00.271000	kr-83m	00:30:00.656000
058106	tpc_kr83m	2024-01-09 12:30:05.988000	2024-01-09 12:00:02.849000	kr-83m	00:30:03.139000
058102	tpc_kr83m	2024-01-09 10:26:23.762000	2024-01-09 09:56:19.652000	kr-83m	00:30:04.110000
058099	tpc_kr83m	2024-01-09 09:55:24.824000	2024-01-09 09:25:23.975000	kr-83m	00:30:00.849000
058091	tpc_kr83m	2024-01-09 08:22:34.301000	2024-01-09 07:52:31.308000	kr-83m	00:30:02.993000
058088	tpc_kr83m	2024-01-09 07:20:47.934000	2024-01-09 06:50:45.111000	kr-83m	00:30:02.823000
058087	tpc_kr83m	2024-01-09 06:49:50.787000	2024-01-09 06:19:47.846000	kr-83m	00:30:02.941000
058081	tpc_kr83m	2024-01-09 04:46:02.996000	2024-01-09 04:15:59.902000	kr-83m	00:30:03.094000
058080	tpc_kr83m	2024-01-09 04:15:06.118000	2024-01-09 03:45:03.337000	kr-83m	00:30:02.781000
058076	tpc_kr83m	2024-01-09 02:42:26.118000	2024-01-09 02:12:23.168000	kr-83m	00:30:02.950000
058074	tpc_kr83m	2024-01-09 02:11:28.933000	2024-01-09 01:41:25.781000	kr-83m	00:30:03.152000
058072	tpc_kr83m	2024-01-09 01:09:35.028000	2024-01-09 00:39:31.831000	kr-83m	00:30:03.197000
058068	tpc_kr83m	2024-01-08 23:36:43.647000	2024-01-08 23:06:40.441000	kr-83m	00:30:03.206000

RunID	mode	end	start	source	livetime
058066	tpc_kr83m	2024-01-08 23:05:46.642000	2024-01-08 22:35:43.386000	kr-83m	00:30:03.256000
058030	tpc_kr83m	2024-01-08 10:42:01.815000	2024-01-08 10:11:59.160000	kr-83m	00:30:02.655000
057736	tpc_kr83m	2023-12-23 23:48:02.727000	2023-12-23 23:17:59.300000	kr-83m	00:30:03.427000
057733	tpc_kr83m	2023-12-23 22:15:12.794000	2023-12-23 21:45:09.597000	kr-83m	00:30:03.197000
057726	tpc_kr83m	2023-12-23 19:40:25.738000	2023-12-23 19:10:22.718000	kr-83m	00:30:03.020000
057720	tpc_kr83m	2023-12-23 17:36:38.588000	2023-12-23 17:06:35.069000	kr-83m	00:30:03.519000
057719	tpc_kr83m	2023-12-23 17:05:40.565000	2023-12-23 16:35:37.508000	kr-83m	00:30:03.057000
057707	tpc_kr83m	2023-12-23 14:29:43.834000	2023-12-23 13:59:40.765000	kr-83m	00:30:03.069000
057705	tpc_kr83m	2023-12-23 13:27:53.461000	2023-12-23 12:57:50.448000	kr-83m	00:30:03.013000
057704	tpc_kr83m	2023-12-23 12:56:59.930000	2023-12-23 12:26:56.541000	kr-83m	00:30:03.389000
057700	tpc_kr83m	2023-12-23 11:55:08.934000	2023-12-23 11:25:06.101000	kr-83m	00:30:02.833000
057695	tpc_kr83m	2023-12-23 09:20:27.967000	2023-12-23 08:50:24.759000	kr-83m	00:30:03.208000
057691	tpc_kr83m	2023-12-23 08:18:33.565000	2023-12-23 07:48:30.402000	kr-83m	00:30:03.163000
057690	tpc_kr83m	2023-12-23 07:47:36.447000	2023-12-23 07:17:33.243000	kr-83m	00:30:03.204000
057688	tpc_kr83m	2023-12-23 06:45:44.301000	2023-12-23 06:15:41.225000	kr-83m	00:30:03.076000
057681	tpc_kr83m	2023-12-23 04:11:02.466000	2023-12-23 03:40:59.655000	kr-83m	00:30:02.811000
057680	tpc_kr83m	2023-12-23 03:40:05.752000	2023-12-23 03:10:02.849000	kr-83m	00:30:02.903000
057675	tpc_kr83m	2023-12-23 02:07:18.254000	2023-12-23 01:37:15.355000	kr-83m	00:30:02.899000
057674	tpc_kr83m	2023-12-23 01:36:25.145000	2023-12-23 01:06:22.375000	kr-83m	00:30:02.770000
057672	tpc_kr83m	2023-12-23 00:34:35.275000	2023-12-23 00:04:32.503000	kr-83m	00:30:02.772000
057669	tpc_kr83m	2023-12-23 00:03:38.583000	2023-12-22 23:33:35.379000	kr-83m	00:30:03.204000
057667	tpc_kr83m	2023-12-22 23:01:44.454000	2023-12-22 22:31:41.394000	kr-83m	00:30:03.060000
057665	tpc_kr83m	2023-12-22 21:59:49.838000	2023-12-22 21:29:46.447000	kr-83m	00:30:03.391000
057664	tpc_kr83m	2023-12-22 21:28:52.357000	2023-12-22 20:58:49.036000	kr-83m	00:30:03.321000
057661	tpc_kr83m	2023-12-22 20:57:54.638000	2023-12-22 20:27:51.512000	kr-83m	00:30:03.126000
057660	tpc_kr83m	2023-12-22 20:26:57.252000	2023-12-22 19:56:54.044000	kr-83m	00:30:03.208000
057653	tpc_kr83m	2023-12-22 17:52:11.761000	2023-12-22 17:22:08.563000	kr-83m	00:30:03.198000
057652	tpc_kr83m	2023-12-22 17:21:14.615000	2023-12-22 16:51:11.494000	kr-83m	00:30:03.121000
057648	tpc_kr83m	2023-12-22 15:17:30.022000	2023-12-22 14:47:27.172000	kr-83m	00:30:02.850000