

Characterization and Optimization of Calorimeter and Active Veto Systems for Light Dark Matter Searches at DarkMESA

DISSERTATION

zur Erlangung des Grades
„Doktor der Naturwissenschaften“

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität in Mainz

vorgelegt von
MATTEO LAUSS
geb. in Worms

Institut für Kernphysik
Johannes Gutenberg-Universität

Mainz, den 18. Dezember 2025

Matteo Lauß
Institut für Kernphysik
Johann-Joachim-Becher-Weg 45
Johannes Gutenberg-Universität
D-55128 Mainz
mlauss@students.uni-mainz.de

Tag der mündlichen Prüfung: 27. April 2026

Dieses Werk ist lizenziert unter CC BY 4.0.

Abstract

Sub-GeV dark matter searches at surface-level facilities must detect rare low-energy scattering events despite cosmic ray backgrounds exceeding anticipated signals by orders of magnitude. The DarkMESA experiment at the MESA accelerator (150 μ A, 155 MeV) confronts this challenge using PbF₂ and lead glass Cherenkov calorimetry with plastic scintillator veto. The experiment targets dark photon kinetic mixing $\varepsilon \sim 10^{-3}$ – 10^{-5} for masses below 50 MeV.

This thesis answers three questions: Can PbF₂ crystals produce sufficient light for 10 MeV detection? Can photodetector backgrounds be suppressed when scaling from 25 to 2,000 calorimeter crystals? Can cosmic ray rejection be maintained at the targeted 10 MeV threshold energy?

A GEANT4 simulation framework, validated against cosmic ray measurements, provides quantitative answers. PbF₂ crystals yield 1.9 PE/MeV—sufficient for efficient detection. However, single-PMT readout proves incompatible with rare event searches: dark count rates reach 1 Hz at the 10 MeV detection threshold. Silicon photomultiplier arrays with fourfold coincidence among 6–8 SiPMs per crystal suppress spurious triggers to negligible levels while maintaining >90 % detection efficiency at 10 MeV, at one-sixth the cost of PMT alternatives.

Cosmic ray rejection presents the primary experimental challenge. The Phase A veto achieves ~ 98 % efficiency, with gamma rays and secondary particles dominating the residual backgrounds. Characterization of the residual leak-through establishes design requirements for enhanced configurations.

Detection efficiency and spurious backgrounds are solved problems; enhanced veto architecture remains essential for competitive sensitivity. Validated simulations now guide designs for passive absorbers and extended coverage toward Phase B construction.

Contents

1	Introduction: Dark Matter and the Dark Photon Search . . .	1
1.1	The Dark Matter Problem	2
1.2	From WIMPs to Light Dark Matter	3
1.3	Beam Dump Searches for Dark Photons	5
1.4	Technical Challenges and Thesis Structure	7
2	The MESA Facility and DarkMESA Experiment	11
2.1	The Mainz Energy-Recovering Superconducting Accelerator	12
2.2	The P2 Experiment and Beam Dump Configuration.	14
2.3	The DarkMESA Detection Principle	15
2.4	Phase A Detector Design	17
2.5	Phase B Full-Scale Detector	21
2.6	Background Sources and Mitigation	26
2.7	Sensitivity Requirements and Detector Performance Targets	27
3	Calorimeter Characterization	35
3.1	Calorimeter Design Requirements	36
3.2	Simulation Framework Development	37
3.3	Predicted Cosmic Ray Response	43
3.4	Background Suppression Through Shielding	49
3.5	Experimental Validation.	57
3.6	Energy Resolution.	61
3.7	Particle Detection Efficiency	65
3.8	Signal Detection Efficiency at Target Threshold	69
3.9	Summary and Implications	71

4	Photodetector Optimization Studies.	73
4.1	Spurious Events: PMT Dark Counts.	74
4.2	Phase A Solution: Dual-Sided Readout	76
4.3	Alternative Approach: Coupled Crystal Arrays	80
4.4	Silicon Photomultiplier Arrays	85
4.5	Summary and Conclusions	94
5	The DarkMESA Veto System	99
5.1	Cosmic Ray Response.	100
5.2	Energy-Dependent Detection Efficiency.	104
5.3	Veto Performance Characterization	108
5.4	Summary and Conclusions	120
6	Summary and Future Directions	123
6.1	Detection Efficiency: Challenge Resolved	123
6.2	Spurious Background Suppression: Challenge Resolved.	124
6.3	Cosmic Ray Rejection: Challenge Partially Resolved	125
6.4	Implications for DarkMESA Sensitivity	126
6.5	Phase B Development Priorities.	127
6.6	Concluding Perspective	127
	Appendices	129
A	Appendix Statistical Framework for Beam Dump Dark Matter Searches. .	131
A.1	On/Off Measurement Formalism	131
A.2	Exclusion Limit Methodology.	134
A.3	Likelihood Framework	136
A.4	Hypothesis Testing and Upper Limit	140
A.5	Expected and Observed Limits	145
A.6	Toy Sensitivity Study	147
A.7	Unified Sensitivity Model.	151
A.8	Detector Resolution and Response Matrix	154
B	Appendix Dynode Cascade Parameter Determination.	159
B.1	Experimental Setup.	159
B.2	Dynode Gain Distribution	159
B.3	Multiplicative Noise Validation	160

C	Appendix Energy Resolution Model	163
C.1	Model Framework	163
C.2	Parameter Constraints	164
C.3	Fitted Results.	164
D	Appendix SiPM Response Modeling	167
D.1	Photon Detection Efficiency	167
D.2	Charge Resolution Modeling	167
	List of Figures	169
	List of Tables	171
	Bibliography	173

Chapter 1

Introduction: Dark Matter and the Dark Photon Search

“Somewhere, something incredible is waiting to be known.”

— Carl Sagan

Approximately 85 % of the universe’s matter is dark—gravitationally active yet invisible to telescopes. Dark matter shapes cosmic structure from galactic to cosmological scales, but despite decades of searches, its particle identity remains unknown. Physicists must either identify the dark matter particle or establish stringent constraints on candidate models: this challenge defines contemporary particle physics.

The search has entered a pivotal era. For decades, experiments targeting Weakly Interacting Massive Particles (WIMPs) at the 100 GeV scale have yielded null results, steadily excluding the parameter space where theorists once expected discovery. These null results motivate a strategic shift: exploring lighter candidates in the MeV–GeV mass range, where different production and detection mechanisms apply. Beam dump experiments with high-intensity electron beams offer unique sensitivity to such light dark matter, particularly when dark photons mediate the interactions. These hypothetical force carriers link ordinary matter to a hidden sector.

The DarkMESA experiment at Mainz’s MESA accelerator targets this opportunity. Operating parasitically behind the P2 parity-violation experiment, DarkMESA searches for light dark matter produced through mediators such as dark photons. Competitive

sensitivity demands surface-level Cherenkov calorimetry at a 10 MeV threshold. At this threshold, three challenges become critical: detecting enough photons despite limited light yield, suppressing spurious backgrounds to a negligible level, and rejecting cosmic rays with hermetic coverage.

This thesis develops DarkMESA’s technical foundation by addressing these challenges through comprehensive detector characterization. The following sections trace the logical path from dark matter evidence through candidate theories to detection strategies, establishing the requirements that define this work.

1.1 The Dark Matter Problem

Multiple independent observations establish dark matter’s existence, converging on a consistent picture: most matter in the universe neither emits nor absorbs electromagnetic radiation, yet dominates gravitational dynamics across all scales.

Galaxy rotation curves provide the most intuitive evidence. Stars in galactic outer regions orbit at nearly constant speeds even where visible matter becomes sparse [1, 2]. These flat rotation curves contradict Keplerian expectations and require massive dark matter halos extending far beyond visible boundaries. Gravitational lensing independently confirms this mass discrepancy. The Bullet Cluster [3] provides particularly striking evidence: X-ray emission traces hot gas in one location, while lensing reconstruction places the gravitational mass elsewhere. This spatial separation demonstrates that most cluster mass does not interact electromagnetically.

The cosmic microwave background (CMB) yields the most precise dark matter abundance measurement. The Planck satellite’s analysis of temperature anisotropies establishes that dark matter comprises approximately 26 % of the universe’s energy density [4]. The acoustic peak structure tightly constrains dark matter properties: dark matter must have been non-relativistic (“cold”) during structure formation, must not couple significantly to photons or baryons, and must be non-baryonic. Big Bang nucleosynthesis independently confirms this conclusion—observed light element abundances match predictions only for baryon densities far below total matter density [5].

These constraints establish what dark matter cannot be: ordinary atomic matter, compact astrophysical objects, or any particle within the Standard Model [6]. The Lambda Cold Dark Matter (Λ CDM) framework successfully explains observations from CMB anisotropies to galaxy cluster dynamics [4]. Yet the Standard Model provides no

suitable candidate, pointing toward physics beyond established theory. The question is not whether new particles exist, but what properties they possess and how to detect them.

Answering this question requires navigating a vast theoretical landscape of candidates, each with distinct mass scales, interaction strengths, and experimental signatures.

1.2 From WIMPs to Light Dark Matter

What properties might dark matter particles possess, and how might experiments detect them? The theoretical landscape spans wide ranges in mass and coupling strength (Figure 1.1). For decades, Weakly Interacting Massive Particles (WIMPs) dominated experimental efforts. The motivation was the “WIMP miracle”: particles with weak-scale masses (~ 100 GeV) and interactions naturally produce the observed dark matter abundance through thermal freeze-out [7]. Supersymmetric extensions of the Standard Model provided natural WIMP candidates, and the coincidence between cosmological requirements and electroweak physics seemed too compelling to ignore.

However, decades of increasingly sensitive experiments have found no WIMPs. Direct detection experiments like XENON1T [8] and LUX [9] have improved sensitivity by several orders of magnitude, progressively ruling out the parameter space where theorists expected WIMPs to appear. Collider searches at the LHC have likewise found no supersymmetric partners. The WIMP miracle remains theoretically attractive, but nature has not cooperated.

These null results motivate broadening the search. Light dark matter in the MeV–GeV mass range offers a compelling alternative: thermal freeze-out can still produce the correct relic abundance for sub-GeV masses given appropriate annihilation channels [10]. Hidden sector models provide the theoretical framework [11]. These models posit additional particles and forces that interact with ordinary matter only through portal interactions—feeble connections between Standard Model fields and hidden sector fields. The simplest portal arises from kinetic mixing between gauge bosons. If nature contains a second $U(1)$ gauge symmetry beyond electromagnetism, the corresponding gauge boson—the dark photon A' —can mix with the ordinary photon:

$$\mathcal{L}_{\text{mix}} = -\frac{\epsilon}{2} F^{\mu\nu} F'_{\mu\nu} \quad (1.1)$$

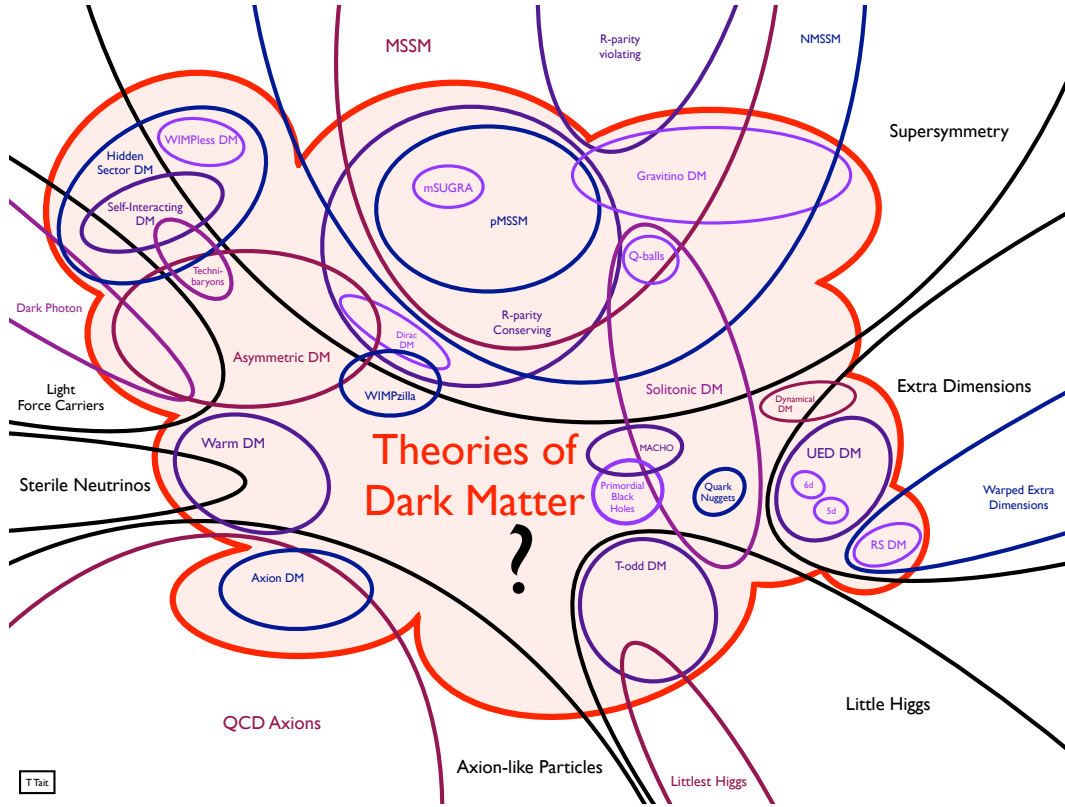


Figure 1.1: Theoretical landscape of dark matter candidates showing interconnections between frameworks. The progressive exclusion of WIMP parameter space by direct detection experiments motivates exploration of alternative candidates, including light dark matter mediated by dark photons (highlighted within hidden sector models). Originally created by T. Tait; reproduced with permission.

where $F^{\mu\nu}$ is the electromagnetic field strength tensor and $F'_{\mu\nu}$ its dark counterpart. After diagonalization, the dark photon couples to electromagnetic current with strength ϵe —suppressed relative to ordinary photon coupling by the small parameter ϵ .

What values of ϵ should experiments target? The mixing parameter is in principle arbitrary, but theoretical considerations suggest characteristic scales. If kinetic mixing arises through quantum corrections involving heavy particles that carry both ordinary and dark charges, one expects $\epsilon \sim 10^{-3}$ [11]. More elaborate scenarios naturally suppress the mixing further, to $\epsilon \sim 10^{-5}$ or below [12, 13]. Rather than pointing to a single natural value, theory motivates searches across many orders of magnitude. The dark photon mass $m_{A'}$ similarly spans a wide range, with symmetry breaking mechanisms naturally placing it in the MeV–GeV range [14].

A dark photon alone cannot constitute dark matter—it decays too quickly to Standard Model particles through the same mixing that enables its production. The scenario

relevant for dark matter searches extends this framework to include a stable particle χ with mass m_χ , coupled to the dark photon with strength α_D . If kinematically allowed ($m_{A'} > 2m_\chi$), the dark photon decays invisibly to $\chi\bar{\chi}$ pairs rather than to electrons or muons.

This framework gains predictive power from cosmology. In the early universe, χ particles maintained thermal equilibrium through creation and annihilation via $\chi\bar{\chi} \leftrightarrow A' \leftrightarrow e^+e^-$. As the universe cooled, annihilation ceased and the χ abundance froze at a value set by the annihilation rate, which depends on ε , α_D , and the masses. This frozen abundance must match the observed dark matter density—a constraint that singles out specific parameter combinations as *thermal relic targets* and transforms an otherwise open parameter space into concrete milestones. For sub-GeV dark matter, these targets require $\varepsilon \lesssim 10^{-4}$.

Other light dark matter candidates exist: axions address the strong CP problem [15], and sterile neutrinos could explain neutrino masses [16]. This thesis focuses on dark photon-mediated dark matter, where beam dump experiments offer particular advantages.

1.3 Beam Dump Searches for Dark Photons

How can experiments detect dark photons and the dark matter they mediate? For dark photon masses in the MeV–100 MeV range, beam dump experiments offer a powerful approach [10, 17]. High-energy electrons striking a dense target produce dark photons through a process analogous to bremsstrahlung. What happens next depends on available decay channels.

When $m_{A'} > 2m_\chi$, the dark photon decays predominantly to dark matter pairs rather than Standard Model particles. The branching ratio depends on relative coupling strengths:

$$\text{BR}(A' \rightarrow \chi\bar{\chi}) = \frac{\alpha_D}{\alpha_D + \varepsilon^2 \alpha_{\text{EM}} R_{\text{SM}}} \quad (1.2)$$

where R_{SM} accounts for available Standard Model final states. For typical parameters where $\alpha_D \gg \varepsilon^2 \alpha_{\text{EM}}$, invisible decays dominate.

The beam dump technique exploits an interaction asymmetry. Dark photons produced in the dump decay promptly to relativistic $\chi\bar{\chi}$ pairs. These dark matter particles traverse meters of shielding essentially unimpeded, while Standard Model backgrounds

are absorbed. A downstream detector then searches for dark matter scattering. Both production and detection scale as ε^2 , yielding an overall rate proportional to ε^4 —a steep suppression demanding high beam intensity and extended running.

When invisible decays are kinematically forbidden ($m_{A'} < 2m_\chi$), the dark photon decays instead to e^+e^- or $\mu^+\mu^-$. The same weak coupling that suppresses these decays also extends the dark photon lifetime: for small enough ε , dark photons travel measurable distances before decaying, producing distinctive displaced signatures. Visible-decay experiments detect these displaced decays with rates scaling as ε^2 —more favorable than the ε^4 of invisible searches. However, they lose sensitivity when dark photons live long enough to escape the detector entirely, which occurs in the small- ε regime where thermal relics reside.

The ε^4 scaling of invisible searches establishes the fundamental experimental challenge. Sensitivity to $\varepsilon \sim 10^{-4}$ requires detecting processes suppressed by 10^{-16} relative to electromagnetic interactions. This demands enormous luminosity, minimal backgrounds, and low detection thresholds—requirements that define the technical challenges addressed in this thesis.

Experimental Landscape

The search for light dark matter has advanced through several complementary experimental strategies, each probing different regions of parameter space. Electron beam dump experiments pioneered the field: the SLAC E137 experiment [18] used a 20 GeV beam and, though originally designed for axion searches, reanalysis of its data now excludes $\varepsilon \gtrsim 10^{-4}$ for dark photon masses below 100 MeV. More recent efforts have refined this approach. At Jefferson Laboratory, the BDX-MINI pilot experiment [19] demonstrated dark matter scattering searches with a 2.176 GeV beam during six months of parasitic operation; the full BDX experiment will extend sensitivity using an 11 GeV beam. The NA64 experiment at CERN [20] takes a complementary approach, using a 100 GeV electron beam to search for visible dark photon decays. Dark photons produced in a compact tungsten dump can escape and decay to e^+e^- pairs in a downstream calorimeter, probing the short-lifetime regime where $10^{-4} \lesssim \varepsilon \lesssim 10^{-3}$.

Proton beam facilities extend sensitivity to higher mediator masses by producing dark photons through meson decays ($\pi^0, \eta \rightarrow \gamma + A'$) and proton bremsstrahlung. The MiniBooNE collaboration [21] performed a dedicated search using an 8 GeV proton beam directed into a steel dump, seeking dark matter–nucleon scattering in a mineral

oil detector 490 m downstream. No excess was observed, providing the most stringent proton beam dump constraints for dark matter masses between 10 MeV and 300 MeV.

Beyond beam dumps, collider and fixed-target experiments offer distinct advantages. BaBar [22] searched for visible decays from $e^+e^- \rightarrow \gamma A'$, setting limits at $\epsilon \sim 10^{-4} - 10^{-3}$ for masses up to 10 GeV. Missing-mass techniques represent a particularly promising direction: by detecting dark photon production rather than subsequent scattering, these experiments achieve ϵ^2 scaling and gain significant sensitivity. At MESA, the MAGIX experiment will employ this approach using a gas jet target [23]. The proposed LDMX experiment [24] would push further still, using a thin tungsten target with tracking and calorimetry to identify events where the recoiling electron carries only a fraction of beam energy. LDMX projects sensitivity one to two orders of magnitude beyond current beam dump limits, though such experiments require dedicated beam time rather than parasitic operation.

Despite this progress, the parameter space where thermal relic dark matter resides remains largely unexplored. For masses between 1 MeV to 50 MeV and $\epsilon \sim 10^{-4}$ to 10^{-5} , visible-decay searches lose sensitivity while invisible searches face severe rate suppression. DarkMESA targets precisely this gap. Operating parasitically behind the P2 parity-violation experiment, DarkMESA uses MESA's 155 MeV, 150 μ A electron beam to produce dark photons in an aluminum beam dump. Dark matter particles from invisible decays scatter in a downstream PbF_2 Cherenkov calorimeter, probing kinetic mixing down to $\epsilon \sim 10^{-5}$ for masses below 50 MeV—reaching into the thermal relic parameter space that motivates light dark matter searches.

1.4 Technical Challenges and Thesis Structure

Achieving competitive dark photon sensitivity at DarkMESA requires detecting electron recoils with energies as low as 10 MeV—the characteristic scale for dark matter scattering in the target parameter space. This threshold requirement, surface-level operation, and the need for very few background events create three interlocking technical challenges. This thesis systematically addresses each challenge, progressing from simulation development through experimental validation to optimization recommendations.

Chapter 2 first establishes the experimental context by describing the MESA accelerator facility and DarkMESA detector configuration, including dual-mode accelerator

operation, detector design, and sensitivity requirements. The chapter also develops the statistical framework relating background levels to sensitivity, quantifying the background suppression required to probe thermal relic dark matter.

Challenge I: Detection Efficiency at Low Threshold

The DarkMESA calorimeter employs PbF_2 crystals operating as pure Cherenkov radiators. Unlike scintillators, PbF_2 produces no scintillation light—optical photons arise exclusively through the Cherenkov mechanism when charged particles exceed the phase velocity of light in the medium.

Detection efficiency depends sensitively on photodetector configuration, coincidence requirements, and threshold settings. Here, detection efficiency is the probability that a genuine 10 MeV deposit triggers the detector. The ϵ^4 scaling of dark photon detection rates means that even modest efficiency losses translate directly into reduced physics reach.

This challenge raises a central question: *Can the simulation framework reliably predict detection efficiency at the 10 MeV threshold?* Chapter 3 addresses this question through simulation of the PbF_2 calorimeter response chain—particle transport, energy deposition, Cherenkov light production, optical propagation, and photodetector conversion—validated against cosmic ray measurements.

Challenge II: Spurious Background Suppression

Establishing adequate detection efficiency addresses only the first obstacle. Even with sufficient light yield to trigger reliably on genuine 10 MeV deposits, a second challenge arises from the photodetectors themselves: spontaneous dark counts that mimic real signals.

Photomultiplier tubes and silicon photomultipliers exhibit dark counts—signals unrelated to incident light—at rates of Hz to several kHz depending on device type and operating conditions. For a single-channel detector, such rates would be negligible. However, DarkMESA's Phase B configuration contains $\sim 2,000$ crystals, each requiring photodetector readout. At 10 MeV threshold with limited light yield, distinguishing genuine signals from dark counts becomes critical.

The question driving this investigation is: *What photodetector configuration suppresses spurious backgrounds to acceptable levels while maintaining adequate signal efficiency?*

Chapter 4 systematically evaluates configurations spanning single PMTs, dual-sided PMT readout, coupled crystal arrays, and silicon photomultiplier arrays.

Challenge III: Cosmic Ray Rejection

Suppressing spurious photodetector backgrounds eliminates internally-generated noise but leaves the dominant external source untreated: cosmic rays. Surface-level operation exposes the detector to cosmic ray fluxes orders of magnitude above anticipated dark matter signal rates.

Muons, electrons, and photons continuously bombard the detector, depositing energy through electromagnetic interactions indistinguishable event-by-event from dark matter scattering. The key distinction: cosmic rays interact with surrounding material as they traverse the detector, while dark matter particles pass through undetected until reaching the active volume.

Hermetic active veto systems exploit this distinction. Plastic scintillator panels surrounding the calorimeter detect charged cosmic rays with near-unity efficiency; the system rejects events coincident with veto signals as background. However, neutral particles—particularly photons from electromagnetic showers initiated above the detector—can deposit energy in the calorimeter without traversing sufficient scintillator to trigger the veto.

The critical question is: *Does the current veto architecture achieve the rejection efficiency required for competitive sensitivity, and which background components limit performance?* Chapter 5 presents the first experimental characterization of the Phase A veto system, comparing it with simulation predictions. The simulation also reveals which particles penetrate the veto undetected.

Synthesis and Outlook

Chapter 6 synthesizes results from these three investigations, establishing both capabilities and limitations of current approaches. This thesis progresses from simulation development through experimental validation to optimization recommendations, demonstrating the viability of Cherenkov calorimetry for beam dump dark matter searches while identifying specific design requirements—SiPM-based readout and enhanced veto systems—that define development priorities for Phase B implementation.

These results provide essential guidance for achieving competitive sensitivity at Dark-MESA, contributing to the broader experimental program probing whether dark matter resides in the light mass range that traditional WIMP searches cannot access.

Chapter 2

The MESA Facility and DarkMESA Experiment

The Mainz Energy-recovering Superconducting Accelerator (MESA), currently under construction at Johannes Gutenberg University Mainz, delivers high-current electron beams uniquely suited for parasitic beam-dump dark matter searches. Operating at 155 MeV with 150 μA beam current, MESA enables the DarkMESA experiment to probe dark photon masses below 100 MeV with kinetic mixing parameters down to $\epsilon \sim 10^{-5}$ —directly targeting the theoretically motivated parameter space identified in Chapter 1.

Parasitic operation behind the P2 parity-violation experiment offers a unique advantage: DarkMESA accumulates beam-on exposure during P2 physics runs while P2 calibration and maintenance periods provide beam-off data essential for background characterization—all without requiring dedicated accelerator time. This operational model, combined with a staged detector program progressing from a 25-crystal Phase A prototype to a 2,000-crystal Phase B expansion, enables systematic technology validation before full-scale commitment.¹

This chapter describes the MESA accelerator facility and DarkMESA detector configuration, culminating in quantitative sensitivity requirements that motivate subsequent characterization studies. Section 2.1 presents accelerator design and dual-mode operation capabilities. Section 2.2 describes the P2 parity-violation experiment and beam dump configuration that enables parasitic dark matter searches. Section 2.3

¹Further extensions in a Phase C are under consideration. However, since these concepts remain preliminary, this thesis focuses on the concrete designs of Phases A and B. For Phase C conceptual studies, see Ref. [25].

introduces the DarkMESA experiment, including detection principle and sensitivity scaling. Sections 2.4 and 2.5 detail the Phase A prototype and planned Phase B expansion. Section 2.6 summarizes background sources, while Section 2.7 translates these challenges into concrete detector performance targets.

2.1 The Mainz Energy-Recovering Superconducting Accelerator

MESA is the first energy-recovering superconducting accelerator dedicated to fundamental physics research, operating parallel to the existing MAMI accelerator at the Institute for Nuclear Physics. Figure 2.1 presents the complete accelerator layout integrated into existing MAMI halls and newly constructed extensions.

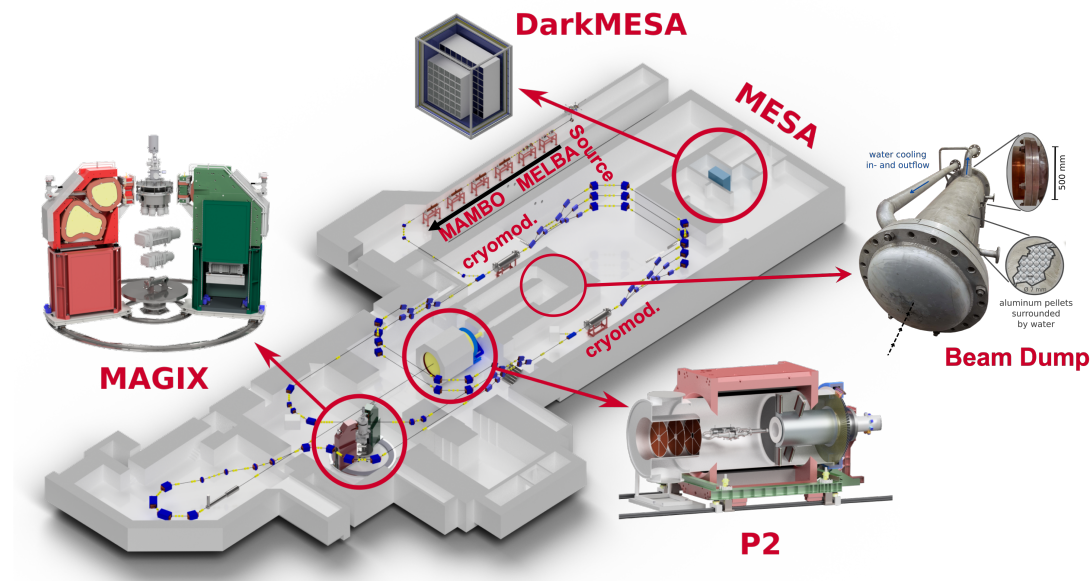


Figure 2.1: Overview of MESA accelerator facility showing beam injection system, recirculation arcs with dipole magnets (blue) and quadrupole magnets (yellow), superconducting cryomodules, and three experimental stations: MAGIX, P2, and DarkMESA. DarkMESA is positioned downstream of P2 beam dump behind substantial concrete shielding.

Acceleration System

The accelerator employs continuous-wave operation at 1.3 GHz radiofrequency, enabling stable, high-duty-factor beam delivery essential for precision measurements.

This frequency matches TESLA/XFEL superconducting cavity technology, leveraging mature RF systems with proven reliability.

Two electron sources serve different experimental needs: a polarized inverted-geometry DC photogun produces spin-polarized electrons, while an unpolarized source enables high-current operation. The MESA Low-energy Beam Apparatus (MELBA) manages beam transport and spin manipulation before injection. The normal-conducting Milli AMpere BOoster (MAMBO) then pre-accelerates the beam to 5 MeV—high enough to minimize space-charge effects during recirculation while enabling efficient coupling into superconducting cavities.

The beam gains energy through recirculation in a racetrack configuration containing two ELBE-type superconducting cryomodules. Each cryomodule houses two 9-cell niobium cavities based on the TESLA/XFEL design [26]. Each cavity provides up to 12.5 MeV energy gain per pass, yielding 50 MeV per complete recirculation. Recirculation arcs employ dipole magnets for beam bending and quadrupole magnets for transverse focusing. The arc geometry must satisfy three requirements: controlled momentum compaction, transverse optics that accommodate varying beam energies, and sufficient physical separation between beam paths.

Dual-Mode Operation

MESA’s operational flexibility derives from two distinct beam modes serving different experimental requirements (Table 2.1).

Table 2.1: Operational parameters for MESA beam modes. Energy-recovering linac (ERL) mode serves internal target experiments with high current; extracted beam (EB) mode provides polarized beam for P2 and DarkMESA.

Parameter	ERL Mode	EB Mode
Recirculations	2	3
Beam energy	105 MeV	155 MeV
Beam current	1 mA	150 μ A
Beam polarization	Unpolarized	85 %
Primary experiment	MAGIX	P2, DarkMESA

Energy-Recovering Linac Mode. In ERL mode, electrons complete two recirculations to 105 MeV before interacting with an internal gas-jet target. The MAGIX

experiment [23, 27] exploits this mode for high-luminosity scattering studies, achieving luminosities of $\mathcal{O}(10^{35} \text{ cm}^{-2} \text{ s}^{-1})$ despite the low areal density of gaseous targets. The physics program includes elastic electron-proton scattering for form factor measurements, few-body nuclear structure studies, and dark photon decay searches—both visible ($A' \rightarrow e^+ e^-$) and invisible.

Most beam electrons pass through the low-density target without scattering and continue into additional recirculation arcs, returning to the acceleration cavities with 180° RF phase shift. This phase shift causes deceleration rather than acceleration, transferring kinetic energy back to cavity electromagnetic fields. Energy recovery enables beam currents of 1 mA (planned upgrade to 10 mA) while eliminating beam dump thermal load. The high current precludes polarized electrons because intense laser illumination limits photocathode lifetime.

Extracted Beam Mode. EB mode prioritizes beam quality and polarization over current. After three recirculations, the beam reaches 155 MeV with 85 % longitudinal polarization at 150 μA . This mode serves the P2 parity-violation experiment and, parasitically, the DarkMESA dark matter search described in Section 2.3.

2.2 The P2 Experiment and Beam Dump Configuration

The P2 experiment aims to measure the weak charge of the proton $Q_W(p)$ with high precision through parity-violating elastic electron-proton scattering at $Q^2 = 4.5 \times 10^{-3} \text{ GeV}^2$ [28]. Because DarkMESA uses P2's beam dump as its dark photon production target, understanding P2's configuration is essential.

In EB mode, 85 % longitudinally polarized electrons at 155 MeV and 150 μA interact with a 60 cm liquid hydrogen target. Over 20,000 h of operation, P2 accumulates approximately 10^{18} elastic scattering events. The parity-violating asymmetry $A_{\text{PV}} = (\sigma_+ - \sigma_-)/(\sigma_+ + \sigma_-)$ demands extraordinary systematic control: P2 must suppress helicity-correlated beam parameter variations below $\Delta A_{\text{app}} < 0.1 \text{ ppb}$. The detection system comprises a toroidal magnetic field, monolithic active pixel sensors for tracking, and quartz Cherenkov detectors that integrate scattered electron flux.

Following the P2 target and detectors, the full-energy beam deposits 23.25 kW in a water-cooled aluminum beam dump originally constructed for the A4 experiment. This beam dump safely absorbs beam power for P2 while serving as DarkMESA’s dark photon production target.

P2’s operational schedule directly determines DarkMESA beam time availability. Parasitic operation provides an additional benefit: P2’s calibration and maintenance periods give DarkMESA beam-off time for background characterization—essential data obtained without dedicated running.

2.3 The DarkMESA Detection Principle

This section describes the detection principle of the DarkMESA experiment; Sections 2.4 and 2.5 detail the detector implementation.

Detection Principle

The beam dump search strategy outlined in Section 1.3 unfolds in three stages at DarkMESA (Figure 2.2).

Stage 1: Dark Photon Production. Electrons at 155 MeV interact in the P2 target and deposit remaining energy in the beam dump through electromagnetic shower development. Kinetic mixing enables dark photon production via bremsstrahlung:

$$e^- Z \rightarrow e^- Z A' \quad (2.1)$$

with production cross section suppressed by ϵ^2 relative to Standard Model photon production.

Electromagnetic showers provide a complementary production channel. Pair-produced positrons can annihilate resonantly with atomic electrons when the center-of-mass energy matches the dark photon mass:

$$e^+ + e^- \rightarrow A' \quad (2.2)$$

This resonant channel enhances sensitivity in mass regions where kinematics favor substantial positron production.

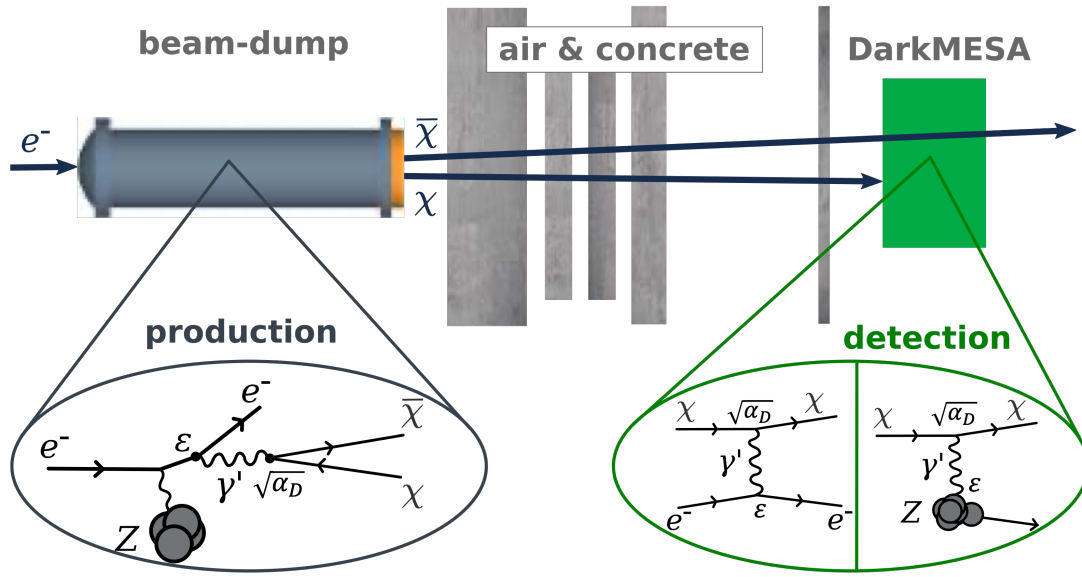


Figure 2.2: Schematic of the DarkMESA detection chain illustrating the three-stage process. Left: electrons impinge on the beam dump, producing dark photons A' (sometimes γ' as in the schematic) via bremsstrahlung that promptly decay to dark matter pairs $\chi\bar{\chi}$. Center: dark matter particles propagate through air and concrete shielding that attenuates Standard Model backgrounds. Right: dark matter scatters from electrons or nuclei in the detector via dark photon exchange. Feynman diagrams show the production vertex (bottom left) and detection vertices (bottom right). The kinetic mixing parameter ϵ appears at each Standard Model–dark sector interface, producing the characteristic ϵ^4 scaling of the total detection rate.

Stage 2: Dark Matter Production and Propagation. If $m_{A'} > 2m_\chi$, the dark photon decays promptly via $A' \rightarrow \chi\bar{\chi}$ within the dump. The relativistic decay products propagate forward with substantial Lorentz boost through approximately 23 m of concrete and air. This shielding attenuates electromagnetic and hadronic backgrounds while dark matter particles pass through essentially unimpeded.

Stage 3: Detection. Dark matter particles scatter elastically from atomic electrons or protons in the calorimeter:

$$\chi + e^- \rightarrow \chi + e^- \quad (2.3)$$

$$\chi + p \rightarrow \chi + p \quad (2.4)$$

with cross sections proportional to $\epsilon^2\alpha_D$, where α_D characterizes dark sector coupling strength. Electron scattering dominates the detectable signal. The planned calorimeter uses PbF_2 , whose Cherenkov threshold ($\beta > 0.56$) requires only 0.11 MeV for electrons but about 200 MeV for protons. Proton recoils from scattering of sub-GeV dark matter

fall far below this threshold, so they produce no Cherenkov light. Electrons that exceed threshold, however, radiate and create the signal.

Sensitivity Scaling

The overall detection probability scales as $\sigma_{\text{total}} \propto \epsilon^4 \alpha_D$, reflecting kinetic mixing suppression in both production (ϵ^2) and detection (ϵ^2). This quartic scaling—the fundamental challenge identified in Section 1.3—demands enormous integrated luminosity. DarkMESA addresses this through years of parasitic operation behind P2, with 150 μA continuous beam accumulating the electron-on-target statistics required to probe ϵ down to $\mathcal{O}(10^{-5})$.

The detector sits approximately 23 m downstream of the beam dump, a distance determined by the available experimental hall geometry. This separation provides substantial shielding for background suppression while maintaining reasonable geometric acceptance for the forward-boosted dark matter flux.

2.4 Phase A Detector Design

The Phase A detector validates detector technology and background rejection strategies while providing initial dark matter search capability during P2 commissioning. This staged approach manages technical risk, enabling optimization of the full-scale Phase B design based on measured performance.

General Configuration

The Phase A detector comprises a 5×5 array of lead fluoride (PbF_2) Cherenkov crystals surrounded by a hermetic active veto system. Figure 2.3 presents an exploded-view CAD rendering illustrating the mechanical integration of calorimeter, veto layers, and lead shielding within an aluminum support framework.

The modular construction enables straightforward calorimeter access for crystal replacement or photodetector modifications without complete system disassembly. Aluminum mountings secure individual veto panels to the MayTec framework, allowing selective panel removal while maintaining structural integrity. This modularity proved essential during optimization studies evaluating multiple readout configurations.

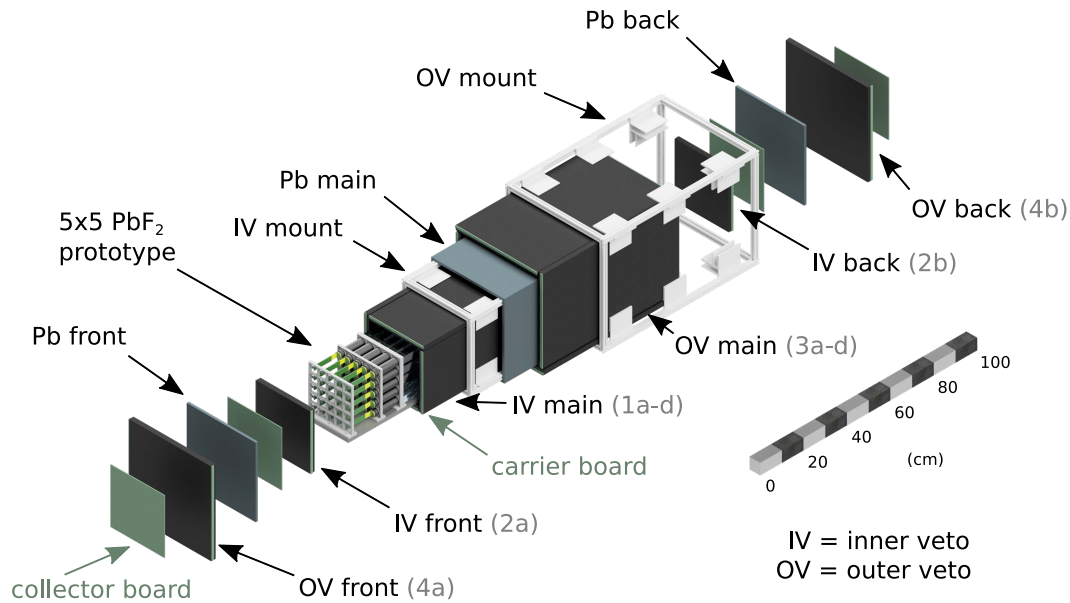


Figure 2.3: Exploded-view CAD rendering of the Phase A prototype detector assembly. The central 5×5 PbF_2 calorimeter is enclosed by the inner veto layer (IV), 10 mm lead absorber (Pb), and outer veto layer (OV). All veto panels consist of 20 mm EJ-200 plastic scintillator. The carrier boards (at panel edges) and collector boards (larger green PCBs) provide SiPM readout electronics for the veto system. The MayTec aluminum profile framework provides structural support and enables modular assembly without welding. Figure adapted from [25]. Note: An additional 10 mm lead layer between calorimeter and inner veto incorporated in the final design is not shown in this rendering.

Calorimeter Design and Photodetector Challenges

Two factors motivated choosing lead fluoride (PbF_2) as the active detection medium: resource availability and superior performance. Approximately 1,000 PbF_2 crystals from the decommissioned A4 parity-violation experiment at MAMI are available, enabling rapid prototyping without new crystal procurement. Beam tests documented in Christmann’s dissertation demonstrated that PbF_2 yields superior light collection compared to alternative Cherenkov radiators of comparable geometry, specifically lead glass types Schott SF5, SF6, and SF57HTU [25].

The calorimeter active volume comprises 25 lead fluoride crystals arranged in a 5×5 matrix, yielding approximately 3.6 L detector volume. Individual crystals taper from 26 mm $\times$ 26 mm at the front face to 30 mm $\times$ 30 mm at the rear, with 185.4 mm length.

Each crystal is wrapped in PTFE (Teflon) reflector tape to enhance light collection through diffuse reflection, with an outer layer of black polyethylene film providing optical isolation between adjacent crystals. The wrapped assemblies mount in custom 3D-printed holders that maintain optical isolation while providing mechanical support.

Figure 2.4 shows representative PbF_2 crystal specimens and the assembled calorimeter matrix.

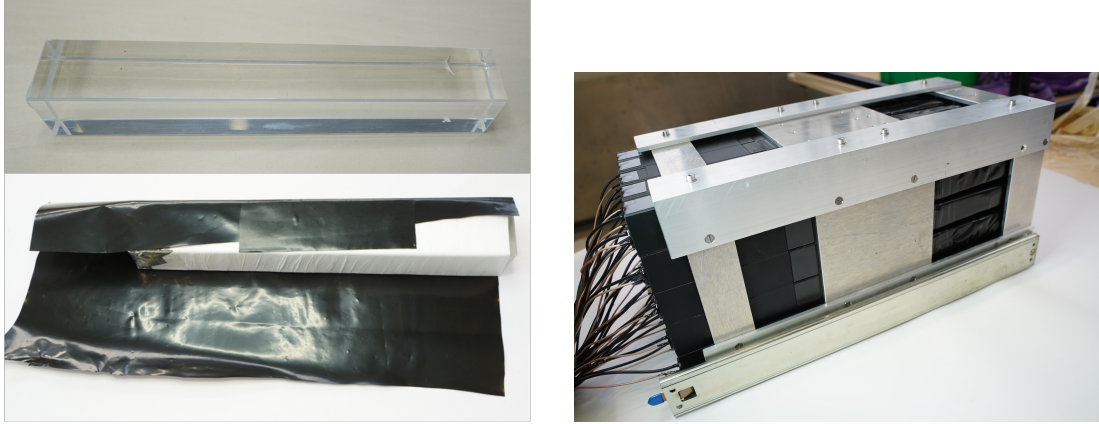


Figure 2.4: Phase A calorimeter components. Left: Individual PbF_2 crystal being wrapped with PTFE reflector and light-tight foil. Right: Assembled 5×5 calorimeter matrix in aluminum framework with individual PMT readout.

The initial Phase A design employed photomultiplier tubes (PMTs) on one crystal face, using Photonis XP2900 tubes from the A4 experiment inventory. However, laboratory characterization revealed that PMT dark count rates produce irreducible backgrounds, rendering this configuration impractical for rare event searches at the target 10 MeV threshold. Chapter 4 presents the comprehensive photodetector optimization study motivating ongoing calorimeter remodeling to implement coincidence-based background rejection.

Active Veto System

Cosmic ray backgrounds dominate at the surface-level DarkMESA location. The hermetic active veto system tags cosmic-induced backgrounds event-by-event, enabling offline rejection through anticoincidence: the analysis rejects events depositing energy simultaneously in both calorimeter and veto system from the dark matter candidate sample. Dark matter particles deposit energy in only a single crystal due to their extremely low interaction probability, producing no correlated veto signal. Chapter 5 presents detailed veto efficiency characterization using cosmic ray measurements and simulations.

The veto system comprises two scintillator layers (inner and outer) separated by a 10 mm lead absorber (Figure 2.3). Both layers employ 20 mm thick EJ-200 plastic

scintillator panels (Eljen Technology), chosen for fast timing response (2.1 ns decay time) and high light yield (10,000 photons per MeV). Each panel is wrapped in PTFE reflector with an outer layer of black polyethylene film for optical isolation.

The inner veto (IV) layer comprises six panels providing coverage around the calorimeter:

- Four lateral panels (1a–d in Figure 2.3): 500 mm × 250 mm × 20 mm, covering top, bottom, and sides
- Two end-cap panels (2a–b): 250 mm × 250 mm × 20 mm, closing front and rear faces

The outer veto (OV) layer follows the same geometry at larger dimensions, positioned outside the lead absorber. The lead absorber attenuates low-energy bremsstrahlung photons and provides additional passive shielding. An additional 10 mm lead layer between the calorimeter and inner veto mitigates backscatter contamination, preventing secondary particles from electromagnetic cascades within the calorimeter from triggering false anticoincidence rejection.

Each scintillator panel employs silicon photomultipliers (SiPMs) on both edges, with SiPM arrays mounted perpendicular to the scintillator length. Figure 2.5 illustrates the readout electronics architecture. Each panel edge mounts a carrier board hosting nine SiPMs (grouped into three triplets) arranged linearly to match the panel width. Carrier boards route signals to collector boards, each accommodating up to six carrier board inputs. The collector board performs two-stage amplification (total gain 100) and generates four independent analog outputs per carrier board: three signals from individual SiPM triplets plus an analog sum of all nine SiPMs. The board also hosts discriminators providing LVDS digital outputs.

The collector board enables remote adjustment of SiPM bias voltages and threshold settings via Ethernet interface. The complete veto system comprises 12 scintillator panels with dual-sided readout, yielding 24 carrier boards serviced by four collector boards.

Commissioning Goals and Expected Performance

Phase A operations during MESA commissioning anticipate approximately 2,200 h of beam time. This dataset enables systematic background characterization through beam-on/beam-off comparison, validates simulation predictions, and demonstrates

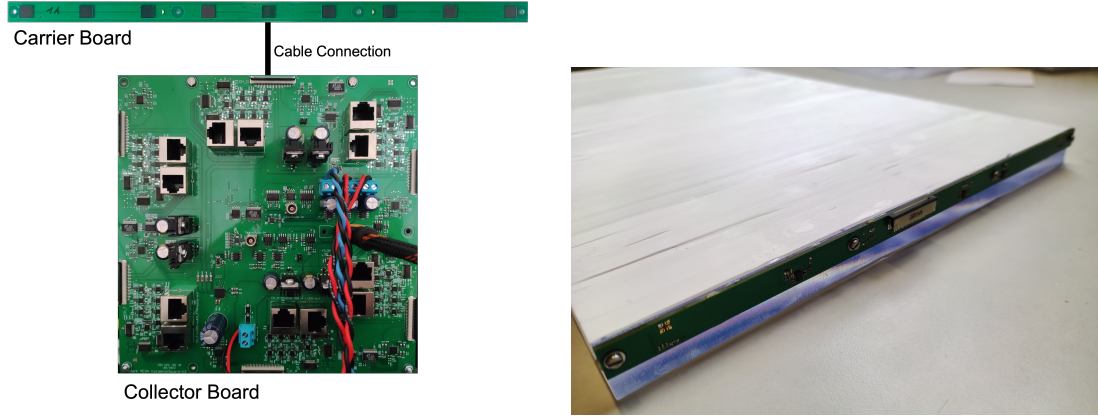


Figure 2.5: Veto system readout electronics. Left: Carrier board (small green PCB with linear SiPM array) and collector board (large PCB). The collector board provides amplification and signal conditioning for up to six carrier boards, with Ethernet connector for remote configuration. Right: EJ-200 scintillator panel wrapped in PTFE reflector, with carrier board mounted at one edge.

long-term detector stability. While the modest active volume precludes competitive dark matter exclusion limits, the Phase A program establishes the foundation for subsequent Phase B scaling.

2.5 Phase B Full-Scale Detector

Phase B scales Phase A technology to a full-scale implementation exploiting the complete A4 PbF₂ crystal inventory supplemented by lead glass modules to maximize detector volume. The transition requires validating background rejection and optimizing photodetector configuration during Phase A commissioning.

Dual-Wall Calorimeter Architecture

Phase B employs a dual-wall architecture with distinct PbF₂ and lead glass detector volumes arranged sequentially along the dark matter propagation direction from the beam dump. Figure 2.6 presents the conceptual detector assembly, while Figure 2.7 shows the SAPHIR lead glass modules under consideration.

PbF₂ Front Wall

The PbF₂ front wall exhausts the usable A4 crystal inventory, with a small number reserved as spares. Each module contains the standard 5 × 5 crystal matrix (25 crystals

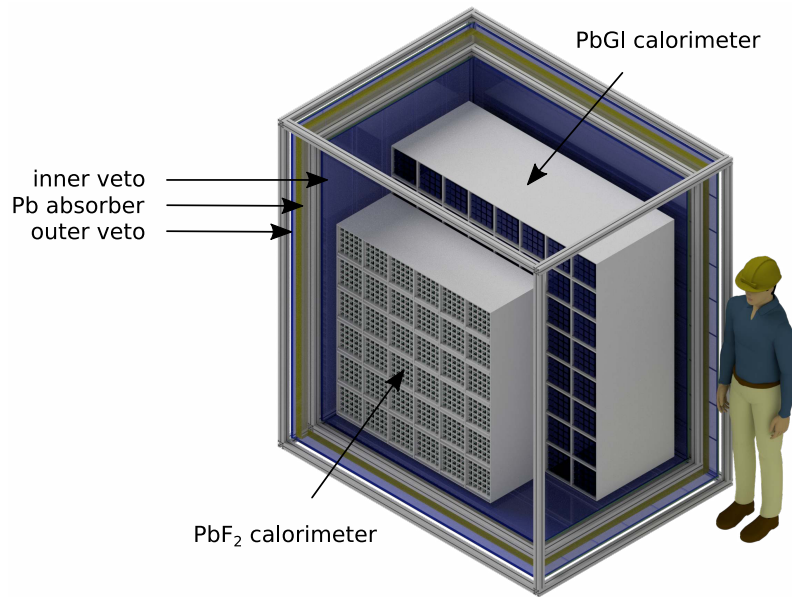


Figure 2.6: CAD rendering of the Phase B detector configuration showing the dual-wall calorimeter architecture. The front wall comprises PbF_2 modules (labeled “ PbF_2 calorimeter”) with total volume ~ 120 L, while the rear wall contains lead glass modules (labeled “ PbGl calorimeter”) totaling ~ 580 L. The complete assembly is enclosed by inner veto, lead shielding, and outer veto. Adapted from [25].

per module), yielding approximately 1,000 PbF_2 detectors total. Individual module volumes vary between 2.9 L and 3.6 L due to crystal length variations (150.0 mm to 185.4 mm), producing aggregate PbF_2 volume of approximately 120 L.

SF5 Lead Glass Rear Wall

Christmann’s comparative beam tests identified Schott SF5 lead glass as the material with the second-highest light yield after PbF_2 . Combined with significantly lower material costs, SF5 offers an economically viable path to expanding active volume.

The decommissioned SAPHIR calorimeter (WA80 heavy-ion experiment) provides 64 supermodules, each comprising a 4×4 SF5 crystal array (1024 individual detectors total). The larger planned SF5 volume (~ 580 L) reflects both greater module count and larger individual crystal dimensions. Combined with the PbF_2 front wall, this configuration provides total active volume approaching 700 L—nearly 200 times the Phase A volume.

However, the SAPHIR modules present significant technical challenges. First, SAPHIR crystals have substantially different geometry from the SF5 samples characterized in Christmann’s beam tests: SAPHIR crystals measure 460 mm in length with 35 mm $\times$

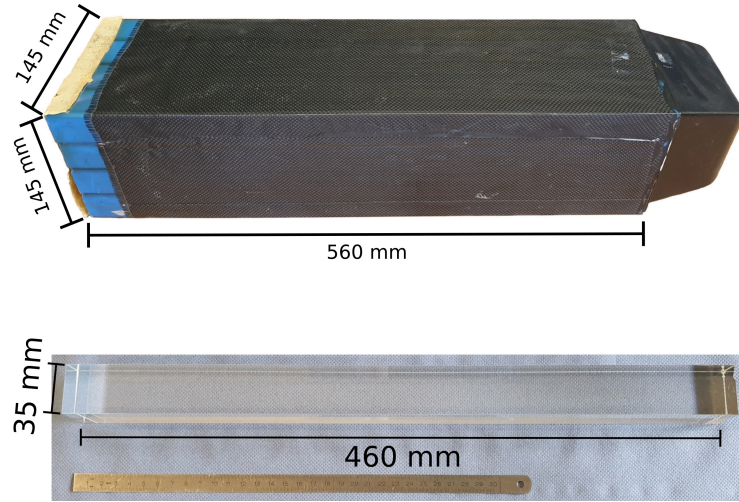


Figure 2.7: SAPHIR lead glass modules under consideration for Phase B rear wall. Top: Complete supermodule containing 4×4 array of SF5 crystals bonded together with carbon fiber reinforcement. Bottom: Individual extracted SF5 lead glass crystal showing substantially larger dimensions (460 mm length, 35 mm $\times$ 35 mm cross-section) compared to PbF_2 crystals (maximum length 185.4 mm).

35 mm cross-section, whereas tested SF5 samples measured approximately 150 mm $\times$ 30 mm $\times$ 30 mm. The threefold increase in propagation length substantially degrades light collection efficiency through accumulated absorption losses, potentially reducing photoelectron yield below levels required for reliable 10 MeV threshold operation.

Second, individual SF5 crystals are permanently bonded within the 4×4 supermodule assemblies using epoxy adhesives with carbon fiber reinforcement (Figure 2.7, top). While disassembly and crystal shortening remain theoretically possible, the brittle nature of lead glass combined with robust bonding risks crystal damage during extraction.

The Phase B lead glass configuration presents four options: utilizing SAPHIR modules intact, attempting systematic disassembly, procuring new SF5 crystals with optimized geometry, or selecting an alternative Cherenkov radiator material entirely. Evaluating these options requires comprehensive optical characterization of SAPHIR module performance at the 10 MeV energy scale and cost-benefit analysis comparing refurbishment, new procurement, and material alternatives.

Veto System Scaling

The Phase B veto system follows Phase A design principles while scaling dimensions for the larger calorimeter volume. Preliminary designs maintain the dual-layer architecture with panel dimensions increased to enclose the approximately $1.3\text{ m} \times 1.3\text{ m} \times 0.7\text{ m}$ calorimeter assembly.

However, larger veto panel dimensions introduce light collection challenges. Christmann's simulations demonstrate that light yield degrades significantly with increasing scintillator panel length: Phase A panels (500 mm) achieve approximately 8.5 % average light collection efficiency, while scaled Phase B panels (2,000 mm) exhibit only 2.7 % efficiency, with systematic degradation of approximately 2 % per 100 mm distance from the SiPM location. Reduced light collection potentially compromises veto efficiency.

Several strategies could mitigate this degradation. First, segmenting panels into smaller tiles would preserve favorable light collection characteristics, albeit with increased SiPM channel count and narrow dead zones at tile boundaries. Second, distributing photodetectors along multiple panel edges could reduce maximum photon propagation distances. Third, integrating wavelength-shifting (WLS) fibers offers potential light yield enhancement, though at increased fabrication complexity.

The definitive Phase B veto architecture depends on experimental validation of large-panel prototypes and quantitative assessment of detection threshold requirements. Segmentation strategy, SiPM configuration, and potential WLS fiber integration all await these results.

Data Acquisition System Scaling

Phase B requires a new data acquisition architecture to accommodate the dramatically increased channel count while maintaining cost-effectiveness and full waveform digitization.

Phase A Data Acquisition Architecture

The Phase A detector employs CAEN V1742 waveform digitizers providing up to 5 GHz sampling rate and 12 bit resolution, enabling detailed pulse shape analysis and offline signal processing optimization. The CAEN system implements self-trigger logic on calorimeter PMT channels: when either PMT signal exceeds threshold, the digitizer captures waveforms from all connected channels simultaneously.

For Phase A with 25 crystals using dual-sided PMT readout, the calorimeter requires 50 PMT channels. Using only the analog sum output from each veto panel's collector board yields 12 additional channels. The complete Phase A detector thus requires 62 total analog channels, implemented using two available CAEN V1742 modules (64 channels total).

Phase B Scaling Challenges and SADC Solution

Extending CAEN-based readout to Phase B reveals significant limitations. The Phase B calorimeter contains approximately 2,000 crystals, requiring roughly 60 digitizer modules. This results in very high acquisition electronics costs.

These costs motivate adopting the Sampling ADC (SADC) system developed by the PANDA collaboration at GSI Darmstadt, specifically optimized for high-rate calorimeter readout. The SADC provides 64 differential analog input channels per board with 14 bit resolution and 80 Msample/s sampling rate. This sampling frequency necessitates analog pulse shaping for reliable amplitude and timing extraction. Dual Xilinx FPGAs enable either full waveform streaming or on-board feature extraction, dramatically reducing data volume. Extensive PANDA beam tests demonstrate reliable energy reconstruction across wide dynamic range.

Critically, the SADC per-channel cost is less than one-quarter the CAEN equivalent, enabling cost-effective scaling to thousand-channel systems. Equipping 2,000 channels requires 32 SADC boards, representing approximately eightfold cost reduction compared to equivalent CAEN implementation while maintaining 14 bit resolution.

The SADC boards interface with the TRB3 timing and trigger distribution system developed at GSI, enabling synchronization through precision clock distribution and trigger propagation. The TRB3 architecture comprises one primary board and multiple secondary boards, each secondary providing eight optical connections for SADC modules. The Phase B configuration with 32 SADCs requires one primary board and four secondary boards.

The DarkMESA project currently has two SADC boards with current PANDA firmware and one TRB3SC controller, providing adequate capacity for Phase A deployment. Full Phase B implementation requires additional SADC and TRB3 secondary boards as construction progresses.

The deployed firmware requires updates before multi-board synchronous operation becomes fully reliable. Integration also requires adapting data acquisition software

for SADC network protocols and data formats. Custom analog signal conditioning electronics must transform single-ended PMT outputs into differential signals with appropriate pulse width extension. Prototype shaper boards have been fabricated and tested with adequate pulse shaping characteristics.

2.6 Background Sources and Mitigation

DarkMESA faces three categories of background, each requiring distinct mitigation strategies.

Beam-Related Backgrounds. Beam-related backgrounds include neutrons and neutrinos from the P2 target and dump. The MESA beam energy (155 MeV) lies only slightly above the threshold for charged pion photoproduction (151.43 MeV for $\gamma + p \rightarrow \pi^+ + n$), suggesting suppressed pion production compared to higher-energy facilities—though the production rate at this energy has not been precisely determined. Pion-induced neutrino backgrounds, which complicate searches at GeV-scale beam dumps, are therefore expected to be reduced at MESA. FLUKA simulations indicate that neutrons from the beam dump do not penetrate the radiation protection walls to reach the DarkMESA room [25], though experimental verification during beam operation is required.

Cosmic Ray Backgrounds. Cosmic ray backgrounds dominate at the surface-level DarkMESA location. The concrete overburden provides partial shielding, but substantial muon and electromagnetic shower fluxes reach the detector. The hermetic veto system provides primary rejection: cosmic ray muons traversing the detector deposit energy in both veto layers, enabling anticoincidence rejection. However, electromagnetic showers initiated above the detector can produce neutral particles—primarily photons—that penetrate veto scintillators without triggering charged particle signals. These photons deposit energy in the calorimeter indistinguishable from dark matter scattering. Chapter 5 quantifies this background and establishes that competitive sensitivity requires veto architectures beyond the current two-layer scintillator design.

Detector-Related Backgrounds. Detector-related backgrounds include photomultiplier dark counts, radioactive contamination, and electronic noise. Dark counts

and noise manifest as uncorrelated single-photodetector events; coincidence requirements effectively suppress these to negligible levels (Chapter 4). Ongoing radiopurity measurements screen detector materials for radioactive contamination.

Of these three categories, cosmic ray backgrounds should dominate for a surface-level experiment. Beam-related backgrounds have not yet been measured at the DarkMESA location; the near-threshold beam energy suggests reduced pion and neutrino production compared to GeV-scale facilities, and simulations predict negligible neutron penetration, but experimental verification is required. The quantitative performance requirements developed in the next section assume cosmic rays constitute the primary background—an assumption that, as the analysis will show, imposes stringent constraints on achievable sensitivity.

2.7 Sensitivity Requirements and Detector Performance Targets

What statistical sensitivity can DarkMESA realistically achieve? This section establishes quantitative performance targets, revealing a stark conclusion: surface-level cosmic ray backgrounds preclude discovery-level claims, restricting the experiment to setting exclusion limits on dark photon parameter space. Background minimization therefore becomes the primary lever for competitive physics reach, motivating the detector optimization studies in subsequent chapters.

The Statistical Reality of Surface-Level Detection

Surface-level operation exposes the detector to substantial cosmic ray flux. No realistic veto system can eliminate backgrounds completely—some fraction of cosmic ray events will inevitably evade rejection. The experiment must therefore operate in a background-dominated regime where residual cosmic ray events substantially outnumber potential signals. How severe are the resulting constraints?

Beam dump experiments distinguish signal from background through beam-on/beam-off comparison. During beam-off periods, only cosmic ray backgrounds contribute. During beam-on periods, potential dark matter signals add to this background. The difference provides the signal estimate, but both measurements carry Poisson fluctuations. For equal measurement periods ($\tau \equiv T_{\text{off}}/T_{\text{on}} = 1$) with expected background

N_{bg} , the combined uncertainty becomes $\sqrt{2N_{\text{bg}}}$ —the factor of two reflecting the statistical cost of background subtraction. Within a simplified Gaussian treatment, adding an effective systematic uncertainty on background prediction yields total uncertainty:

$$\sigma_{\text{total}} = \sqrt{2N_{\text{bg}} + (\sigma_{\text{sys}}N_{\text{bg}})^2} \quad (2.5)$$

where σ_{sys} parameterizes systematic effects as an effective fractional uncertainty on background count. Appendix A derives this result in detail and discusses its limitations.

A discovery claim requires five-sigma significance: the probability of background fluctuating to the observed level must fall below 3×10^{-7} . This demands minimum signal:

$$N_{\text{signal}} \geq 5\sigma_{\text{total}} = 5\sqrt{2N_{\text{bg}} + (\sigma_{\text{sys}}N_{\text{bg}})^2} \quad (2.6)$$

A concrete example illustrates the challenge. Consider $N_{\text{bg}} = 10,000$ residual background events under varying levels of effective systematic uncertainty:

- Negligible systematics ($\sigma_{\text{sys}} \rightarrow 0$): requires $N_{\text{signal}} \geq 707$ events
- Moderate systematics ($\sigma_{\text{sys}} = 2\%$): requires $N_{\text{signal}} \geq 1,000$ events
- Large systematics ($\sigma_{\text{sys}} = 10\%$): requires $N_{\text{signal}} \geq 5,000$ events

These requirements contrast sharply with the extremely low expected dark matter signal rate. Even at much lower backgrounds the gap persists: with $N_{\text{bg}} = 100$ events and negligible systematics, discovery still requires 71 signal events.

The BDX-MINI experiment at Jefferson Laboratory—operating with PbWO_4 crystal calorimetry and plastic scintillator veto as the closest experimental analog to DarkMESA—demonstrates these challenges empirically [19]. Despite careful calibration, that experiment accumulated approximately 3600 events passing all selection criteria over the full running period, illustrating the cosmic ray challenge facing surface-level beam dump experiments. Under similar conditions, DarkMESA’s primary sensitivity lies in setting exclusion limits rather than claiming discovery. These limits provide valuable constraints on dark sector parameter space and complement underground searches.

Exclusion Limits as Operating Mode

Even when signal rates fall far below discovery thresholds, exclusion analyses deliver meaningful constraints on new physics. They address a different question than discovery. Rather than asking “did we observe a signal?”, they ask “are the data inconsistent with a specific theoretical prediction?” If a dark photon model predicts a certain number of signal events and the experiment observes far fewer, that model can be excluded at specified confidence level—typically 90 % in dark matter searches.

This approach demands less stringent statistics than discovery. A simplified Gaussian treatment illustrates the basic principle. Consider an experiment observing $N_{\text{obs}} = 1,050$ events with expected background $N_{\text{bg}} = 1,000$, yielding apparent signal of 50 events. With 10 % systematic uncertainty, total uncertainty reaches approximately 110 events. Discovery significance amounts to only $50/110 \approx 0.45\sigma$ —hopelessly inadequate for a five-sigma claim. Yet exclusion remains powerful: in the Gaussian approximation, any model predicting more than approximately 190 events—the observed estimate plus 1.28 standard deviations—would be excluded at 90 % confidence.

In the idealized limit of zero background and negligible systematics, exclusion limits achieve maximum sensitivity. Observing exactly zero events excludes all models predicting more than 2.3 events at 90 % confidence—a result following directly from Poisson statistics. This idealized limit represents the theoretical benchmark against which realistic performance must be compared.

Appendix A develops the complete statistical framework underlying this analysis. The appendix first derives the Gaussian approximations used throughout this section, which provide closed-form scaling relationships useful for experimental design. It then presents the likelihood-based framework—an exact numerical approach valid at any event count that handles systematic uncertainties more rigorously while recovering the same statistical scaling relationships in the large-sample limit. This framework will form the foundation for data analysis once beam operation begins.

Theoretical Sensitivity and Existing Limits of DarkMESA

Figure 2.8 presents the landscape of dark photon searches in the parameter space relevant to DarkMESA. The exclusion limits use the dimensionless parameter:

$$y \equiv \epsilon^2 \alpha_D \left(\frac{m_\chi}{m_{A'}} \right)^4 \quad (2.7)$$

which governs the dark matter annihilation cross-section $\sigma v_{\chi\bar{\chi} \rightarrow f\bar{f}}$. As discussed in Section 1.2, this cross-section determines the relic abundance through thermal freeze-out. The thermal relic targets in Figure 2.8 mark y values that produce the observed dark matter density. Multiple curves appear because the annihilation cross-section depends on the quantum nature of χ —scalar, Majorana fermion, or pseudo-Dirac fermion—each requiring different coupling strengths to achieve the same relic abundance. Experiments reaching these targets test the thermal dark photon scenario for the corresponding particle type.

The projected exclusion limits of DarkMESA follow the analysis by Christmann [25], using Monte Carlo simulations to link dark matter production, transport, and detection. MADGRAPH simulates dark photon production through dark bremsstrahlung ($e^- + \text{nucleus} \rightarrow e^- + \text{nucleus} + A'$) and positron annihilation ($e^+ + e^- \rightarrow A'$), while GEANT4 models electromagnetic shower development and electron scattering in the PbF₂ calorimeter.

These simulations predict N_χ detected events at a reference coupling $\varepsilon_0 = 10^{-3}$ for each dark matter mass. For an experiment observing zero events, Poisson statistics at 90 % confidence excludes parameters predicting more than 2.3 events. The exclusion limit ε_L for a certain mass then satisfies:

$$N_\chi \left(\frac{\varepsilon_L}{\varepsilon_0} \right)^4 = 2.3 \quad (2.8)$$

where the fourth power arises because both production and detection scale as ε^2 .

The standard choice $m_{A'} = 3m_\chi$ represents a theoretically motivated hierarchy enabling on-shell production and invisible decay $A' \rightarrow \chi\bar{\chi}$. This ratio avoids kinematic suppression while maintaining sufficient phase space for decay, and has become a standard benchmark in beam dump experiments [24]. The choice $\alpha_D = 0.5$ represents moderately strong dark sector coupling, intermediate between electromagnetic strength ($\alpha_{\text{EM}} \approx 1/137$) and strong interaction values. Major collaborations have widely adopted this value [24], though some analyses use $\alpha_D = 0.1$ following more recent recommendations [29].

The projected DarkMESA sensitivity curves assume idealized conditions: zero background events and negligible systematic uncertainties. Under these assumptions, Phase B could probe coupling parameters approaching the thermal relic targets for pseudo-Dirac fermion dark matter. The gap between these projections and existing

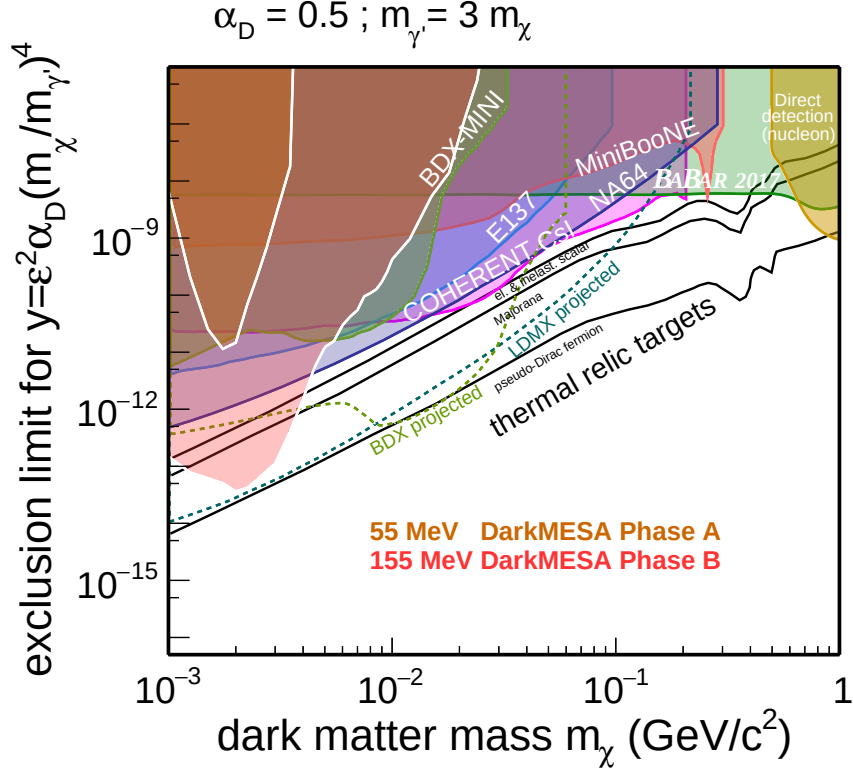


Figure 2.8: Parameter space for dark photons decaying invisibly to $\chi\bar{\chi}$, with dark matter mass m_χ on the horizontal axis and coupling parameter $y = \epsilon^2 \alpha_D (m_\chi/m_{A'})^4$ on the vertical axis, assuming $\alpha_D = 0.5$ and $m_{A'} = 3m_\chi$. Projected DarkMESA exclusion regions at 90 % confidence level are shown for Phase A (2,200 h beam time at 55 MeV) and Phase B (6,600 h at 155 MeV), assuming 95 % detection efficiency at 10 MeV threshold, zero background, and zero systematic uncertainties. Other exclusion limits from BDX-MINI, MiniBooNE, E137, NA64, COHERENT, BABAR, and direct detection experiments are shown for comparison. Thermal relic targets for elastic and inelastic scalar, Majorana, and pseudo-Dirac fermion dark matter are indicated by solid black lines. Figure adapted from [25].

experimental limits defines the target parameter space. Closing this gap requires minimizing backgrounds and controlling systematic uncertainties—subjects addressed in subsequent chapters through veto system performance and detector optimization.

Background Scaling and Detector Requirements

The sensitivity cost of nonzero backgrounds follows from the statistics of Poisson counting. In the on/off measurement framework developed in Appendix A, combining background fluctuations, finite beam-off measurement time, and systematic

uncertainties yields the scaling relationship for event count upper limits:

$$\text{Upper Limit} \propto \sqrt{N_{\text{bg}} \left(1 + \frac{1}{\tau}\right) + (\sigma_{\text{sys}} N_{\text{bg}})^2} \quad (2.9)$$

where N_{bg} is the expected background during beam-on, $\tau = T_{\text{off}}/T_{\text{on}}$ is the ratio of beam-off to beam-on measurement time, and σ_{sys} parameterizes systematic effects as an effective fractional uncertainty on background prediction. Since detection rates scale

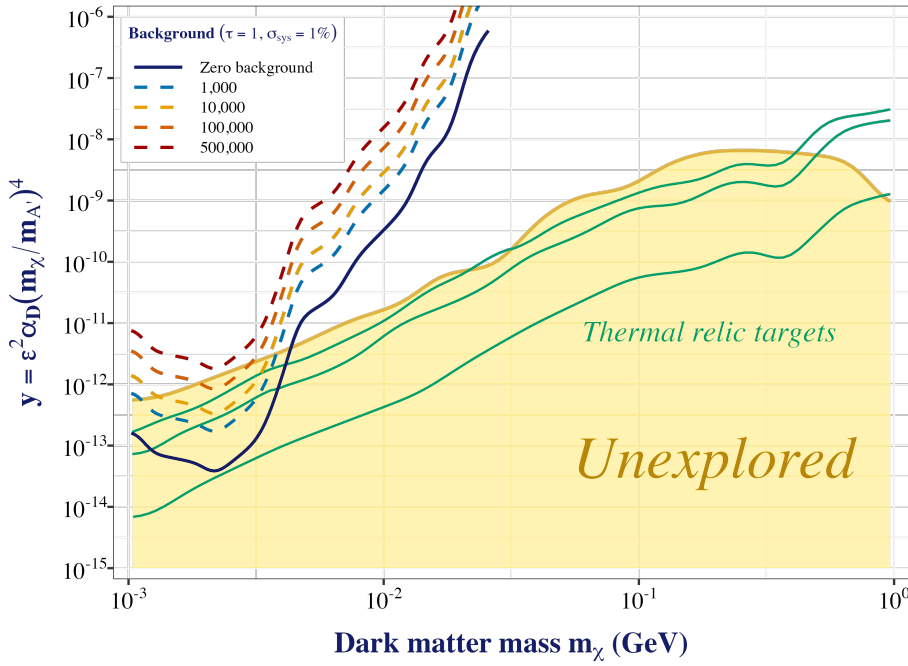


Figure 2.9: Impact of residual background on DarkMESA Phase B exclusion reach in the y - m_χ parameter space, assuming $\tau = 1$ and $\sigma_{\text{sys}} = 1\%$. The solid curve shows the zero-background projection; dashed curves show degraded sensitivity for increasing background levels. The green curves indicate the thermal relic targets; the shaded region remains unexplored by existing experiments.

as ε^4 , the excludable coupling at the exclusion boundary satisfies $\varepsilon^4 \propto \text{Upper Limit}$. With $y \propto \varepsilon^2$, the coupling reach scales as:

$$y \propto \left[N_{\text{bg}} \left(1 + \frac{1}{\tau}\right) + (\sigma_{\text{sys}} N_{\text{bg}})^2 \right]^{1/4} \quad (2.10)$$

Under idealized conditions ($\tau \rightarrow \infty$, $\sigma_{\text{sys}} \rightarrow 0$), this simplifies to $y \propto N_{\text{bg}}^{1/4}$. In the systematics-dominated regime where $\sigma_{\text{sys}} N_{\text{bg}} \gg \sqrt{N_{\text{bg}}}$, the scaling becomes $y \propto (\sigma_{\text{sys}} N_{\text{bg}})^{1/2} \propto N_{\text{bg}}^{1/2}$. The different exponents reveal a crucial insight: coupling reach degrades twice as fast when systematics-limited. The transition occurs when

$\sigma_{\text{sys}} \approx 1/\sqrt{N_{\text{bg}}}$. For $\sigma_{\text{sys}} = 1\%$, this threshold lies at $N_{\text{bg}} \approx 10,000$ events—above which reducing backgrounds without improving systematic control yields diminishing returns.

Table 2.2 quantifies this behavior, comparing idealized conditions against a realistic scenario with $\tau = 1$ and $\sigma_{\text{sys}} = 1\%$. Figure 2.9 visualizes this degradation in the y - m_χ parameter space under these plausible conditions.

Table 2.2: *Coupling reach degradation as a function of residual background events. The idealized case assumes perfect background knowledge ($\tau \rightarrow \infty$) and negligible systematics ($\sigma_{\text{sys}} = 0$); the realistic case assumes equal beam-on/off time ($\tau = 1$) and 1% effective systematic uncertainty. The zero-background Poisson limit of 2.3 events provides the reference.*

N_{bg}	y Degradation Factor	
	Idealized ($\tau \rightarrow \infty, \sigma_{\text{sys}} = 0$)	Realistic ($\tau = 1, \sigma_{\text{sys}} = 1\%$)
0	1×	1×
10,000	6.6×	8.7×
100,000	12×	22×
500,000	18×	47×

The visual impact is striking: with $N_{\text{bg}} = 10,000$ residual events, the exclusion curve shifts upward by nearly an order of magnitude, causing DarkMESA to miss substantial portions of the thermal relic targets. At $N_{\text{bg}} = 100,000$, the y -degradation exceeds a factor of 20, rendering the experiment unable to probe new parameter space where sensitivity is strongest.

This analysis provides only a rough assessment—actual sensitivity depends on details of the likelihood analysis, spectral discrimination between signal and background, and systematic uncertainties characterized during commissioning. Notably, the Gaussian treatment of adding systematic uncertainties in quadrature represents a simplification; the likelihood framework developed in Appendix A shows that systematic effects such as energy calibration enter more favorably when they preserve spectral shape differences between signal and background. Nonetheless, a clear requirement emerges: to probe the thermal relic targets in unexplored parameter space, DarkMESA must keep residual backgrounds at most at the $\mathcal{O}(10^3)$ event level—lower backgrounds yield proportionally better sensitivity. This requirement drives the detector design

choices in subsequent chapters: aggressive cosmic ray veto efficiency, careful shielding optimization, and robust calibration strategies.

Success Criteria for Detector Characterization

The requirements derived above define success criteria for the characterization studies in subsequent chapters. Each technical chapter addresses one component of the detection chain, establishing whether current performance approaches the levels necessary for competitive dark photon sensitivity.

Chapter 3 establishes the foundation for background understanding through comprehensive cosmic ray response studies of the PbF_2 calorimeter crystals. The analysis characterizes energy spectrum composition, quantifies Cherenkov light yield, and systematically evaluates how different particle types interact with the detector material. These particle-specific response characteristics reveal both favorable properties and potential vulnerabilities—findings that inform the veto system evaluation in Chapter 5.

Chapter 4 evaluates photodetector configurations for spurious background suppression. The low Cherenkov light yield in PbF_2 makes low-threshold operation susceptible to dark count contamination. Coincidence requirements between multiple photodetectors can suppress these backgrounds, but at potential cost to detection efficiency. The chapter quantifies this trade-off across configurations ranging from single PMTs to silicon photomultiplier arrays.

Chapter 5 characterizes cosmic ray rejection efficiency—the most critical parameter for sensitivity. In the statistically-dominated regime, upper limits scale as $\sqrt{N_{\text{bg}}}$, translating to $N_{\text{bg}}^{1/4}$ scaling for coupling reach. Each order-of-magnitude reduction in residual background then improves y -reach by a factor of approximately 1.8, as Table 2.2 demonstrates. This favorable scaling holds when systematic uncertainties remain subdominant—a condition the toy study of Appendix A.6 confirms for realistic DarkMESA parameters. Measuring the achieved rejection efficiency and identifying inefficiencies establishes whether Phase A performance suffices or whether Phase B requires enhanced veto architectures.

Together, these characterization results determine the gap between current detector performance and competitive dark photon sensitivity. Quantifying this gap establishes development priorities for Phase B and provides realistic expectations for the physics reach DarkMESA can ultimately achieve.

Chapter 3

Calorimeter Characterization

Detecting sub-GeV dark matter through electron recoils requires calorimeters capable of reliable 10 MeV energy thresholds. Previous work established that PbF_2 crystals provide the best light yield among candidate materials and demonstrated operational detector prototypes. Christmann’s dissertation [25] used 6 MeV to 14 MeV electron beam tests at MAMI to show that PbF_2 outperforms lead glass variants (SF5, SF6, SF57HTU), measuring light yields of approximately 1.86 photoelectrons/MeV to 2.04 photoelectrons/MeV (PE/MeV). Pättschke [30] translated these findings into hardware, constructing the first operational 5×5 array with 3D-printed housing and individual photomultiplier tube (PMT) coupling. Kontogoulas [31] established calibration procedures achieving uniform response across channels despite 50 % PMT gain variations.

This prior work has a common limitation: it characterizes average light yields but cannot predict detection efficiency at a given photoelectron threshold. While 1.9 PE/MeV suggests approximately 19 PE at 10 MeV—nominally adequate for threshold detection—individual events vary substantially around this mean. Poisson fluctuations in photon production, stochastic losses at the photocathode, and multiplicative noise through the dynode cascade cause some 10 MeV deposits to produce 25 PE while others yield only 12 PE. Predicting what fraction of signal events exceed a given threshold requires modeling the complete probability distribution, not just the mean.

This chapter develops a simulation framework extending from Cherenkov production through complete photodetector response, enabling direct prediction of detection efficiency versus threshold. The framework serves two purposes: it answers whether the baseline detector configuration achieves sufficient efficiency at the target threshold, and it provides a validated foundation for evaluating alternative photodetector

configurations—dual-sided PMT arrangements, coupled crystal arrays, and silicon photomultiplier systems—in Chapter 4.

Three developments support these goals. First, the simulation incorporates photocathode conversion, collection efficiency losses, and dynode cascade multiplication with realistic noise, capturing the full statistical distribution rather than mean values alone. Second, the simulation characterizes detection efficiency separately for each cosmic ray particle type—muons, electrons, photons, neutrons, protons—providing quantitative foundations for background discrimination. Third, cosmic ray measurements validate the framework experimentally, demonstrating excellent agreement between predictions and data.

3.1 Calorimeter Design Requirements

The DarkMESA calorimeter must detect energy transferred when light dark matter (LDM) particles scatter elastically off electrons. Christmann [25] showed that the number of recoil electrons is:

$$N_{\text{e-recoil}} = 2N_{\chi}P_{\chi,e} = 2 \left(N_e X_0 n_T \sigma_{\chi\chi} \right) \left(L_d n_e \sigma_{\chi,e} \right) \quad (3.1)$$

The first factor describes dark matter production in the beam dump through dark bremsstrahlung: N_e is the number of beam electrons, X_0 is the radiation length of the dump material, n_T is the target nucleus density, and $\sigma_{\chi\chi}$ is the effective cross-section for dark photon production followed by invisible decay $A' \rightarrow \chi\bar{\chi}$. The second factor characterizes detection through elastic scattering: L_d is the interaction length, n_e is electron density in the detector, and $\sigma_{\chi,e}$ is the elastic scattering cross-section for $\chi + e^- \rightarrow \chi + e^-$. The factor of 2 accounts for both χ and $\bar{\chi}$ particles, each capable of inducing detectable recoils.

Equation 3.1 shows why detector density (n_e) and interaction length (L_d) are crucial: detection probability scales directly with both. The planned Phase B expansion leverages this scaling—adding a lead glass wall increases L_d , directly enhancing sensitivity.

For Phase A, lead fluoride (PbF_2) crystals are the primary detection medium, offering four advantages:

1. **High density:** At 7.77 g/cm^3 , PbF_2 maximizes scattering centers per unit volume (n_e), directly increasing detection probability.

2. **High refractive index:** With $n \approx 1.78$, PbF_2 lowers the Cherenkov threshold velocity ($\beta_{\text{thr}} = 1/n$), enabling light production at lower particle energies. Photon yield per unit length scales as $(1 - 1/n^2)$, enhancing light collection—critical given the low absolute yield discussed below.
3. **Low neutron response:** Experimental studies using 0.5 MeV to 7 MeV neutrons from an AmBe source measured PbF_2 neutron detection efficiency approximately three orders of magnitude below plastic scintillator [32], though experimental limitations likely contributed to this suppression. Chapter 5 presents energy-dependent efficiency comparisons from simulation that reveal a more nuanced picture.
4. **Fast response:** Prompt Cherenkov emission enables excellent timing resolution, facilitating crystal-to-crystal coincidence and possible timing cuts relative to beam structure.

However, PbF_2 presents a significant challenge: low light yield. Christmann [25] measured approximately 1.86 PE/MeV to 2.04 PE/MeV, substantially lower than typical scintillators. This low yield creates challenges for reliable 10 MeV detection, particularly considering statistical fluctuations and potential backgrounds. The following sections evaluate whether this yield suffices and establish detector response characteristics that constrain photodetector selection—addressed comprehensively in Chapter 4.

3.2 Simulation Framework Development

The simulation has two purposes. First, understanding cosmic ray backgrounds is essential for the experiment itself. Cosmic radiation continuously bombards any surface laboratory, producing muons, electrons, photons, neutrons, and protons that constitute the dominant background for dark matter searches. Characterizing how each particle type interacts with the detector is critical for optimizing background discrimination strategies. Second, these same cosmic rays provide the validation dataset. Simulating cosmic ray interactions and comparing predicted photoelectron spectra against measurements validates the complete detector model. Once validated, the simulation becomes a predictive tool for detection efficiencies specific to each particle type—essential for background suppression.

This section describes the extended simulation framework in three parts. The integrated architecture combines cosmic ray generation, particle transport, optical propa-

gation, and photodetector interfaces. The dynode cascade model captures stochastic multiplication processes that dominate energy resolution. Finally, a roadmap shows how subsequent sections build from these foundations toward experimental validation.

Integrated Simulation Architecture

Figure 3.1 illustrates the complete simulation chain from incident cosmic rays to final photoelectron output.

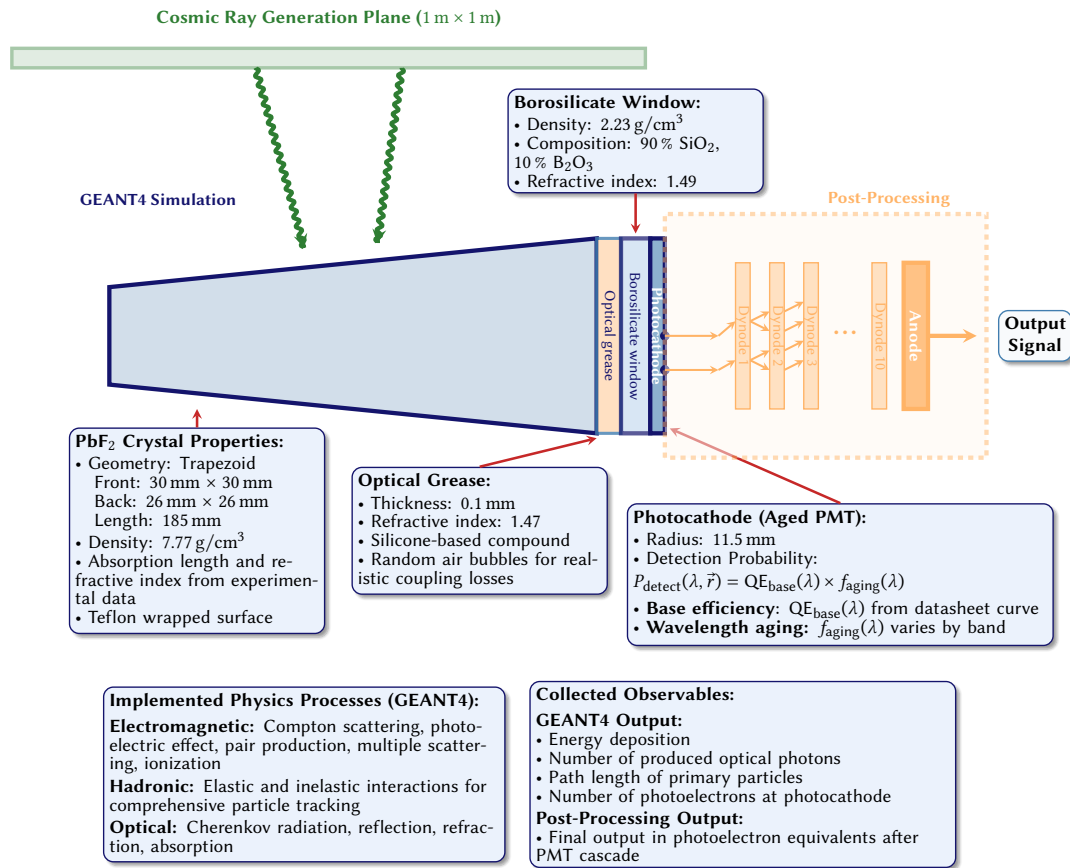


Figure 3.1: Simulation framework architecture for the DarkMESA calorimeter. Cosmic rays generated via the CRY library [33] on a 1 m × 1 m plane undergo GEANT4 transport through PbF₂ crystals with PTFE wrapping. Optical coupling through 0.1 mm silicone grease includes random air bubbles modeling realistic interface losses. The borosilicate PMT window interfaces with an aged photocathode exhibiting wavelength-dependent quantum efficiency (peak at 420 nm, with UV degradation factor 0.7, blue 0.8, green 0.85, red 0.9 reflecting operational history). Post-processing applies dynode cascade simulation (Figure 3.2). The unshielded configuration provides upper-bound background estimates; shielding effects are addressed in Section 3.4.

Cosmic ray generation: The Cosmic-ray Shower Library (CRY) version 1.7 [33] generates correlated particle showers derived from MCNPX simulations. CRY provides muons, neutrons, protons, electrons, photons, and pions over energy ranges from 1 MeV to 100 TeV, matching the spectrum at Earth’s surface.

Particle transport: GEANT4 version 10.4 handles particle propagation through detector materials. The PbF_2 crystals incorporate wavelength-dependent refractive index and absorption length from experimental data [25], with PTFE wrapping included. Optical coupling through silicone grease includes random air bubble distributions modeling realistic interface losses that reduce photon transmission.

Photocathode modeling: The aged photocathode model captures systematic degradation from previous experimental use. UV wavelengths experience 30 % degradation compared to 10 % for red wavelengths. This pattern reflects photocathode degradation physics: high-energy photons cause more severe damage to alkali antimonide surfaces through photo-induced chemical decomposition.

Post-processing: Following GEANT4 particle transport and optical photon tracking, photoelectron signals undergo dynode cascade simulation to predict final PMT output in photoelectron equivalents.

Dynode Cascade Modeling

Predicting detection efficiency requires accurate modeling of signal fluctuations. After photoelectrons are generated at the photocathode, they undergo stochastic multiplication through the PMT dynode chain. Two independent processes introduce fluctuations: collection efficiency losses and per-stage gain variations. Figure 3.2 illustrates the stochastic cascade simulation incorporating these effects.

Collection efficiency represents the probability that a photoelectron reaches the first dynode. Not all photoelectrons follow ideal trajectories due to geometric and electrostatic focusing limitations. The nominal collection efficiency $\varepsilon_0 = 0.70$ lies within the experimentally determined range for similar photodetectors (0.58–0.98) [34], accounting for device age. Per-photoelectron collection probability is modeled as $\varepsilon_i = \varepsilon_0 \times C$, where C represents multiplicative variation sampled from a lognormal distribution $\ln C \sim \mathcal{N}(0, \sigma^2)$ with characteristic spread $\sigma = 0.10$. These parameter values were determined through optimization against experimental cosmic ray measurements (Section 3.5).

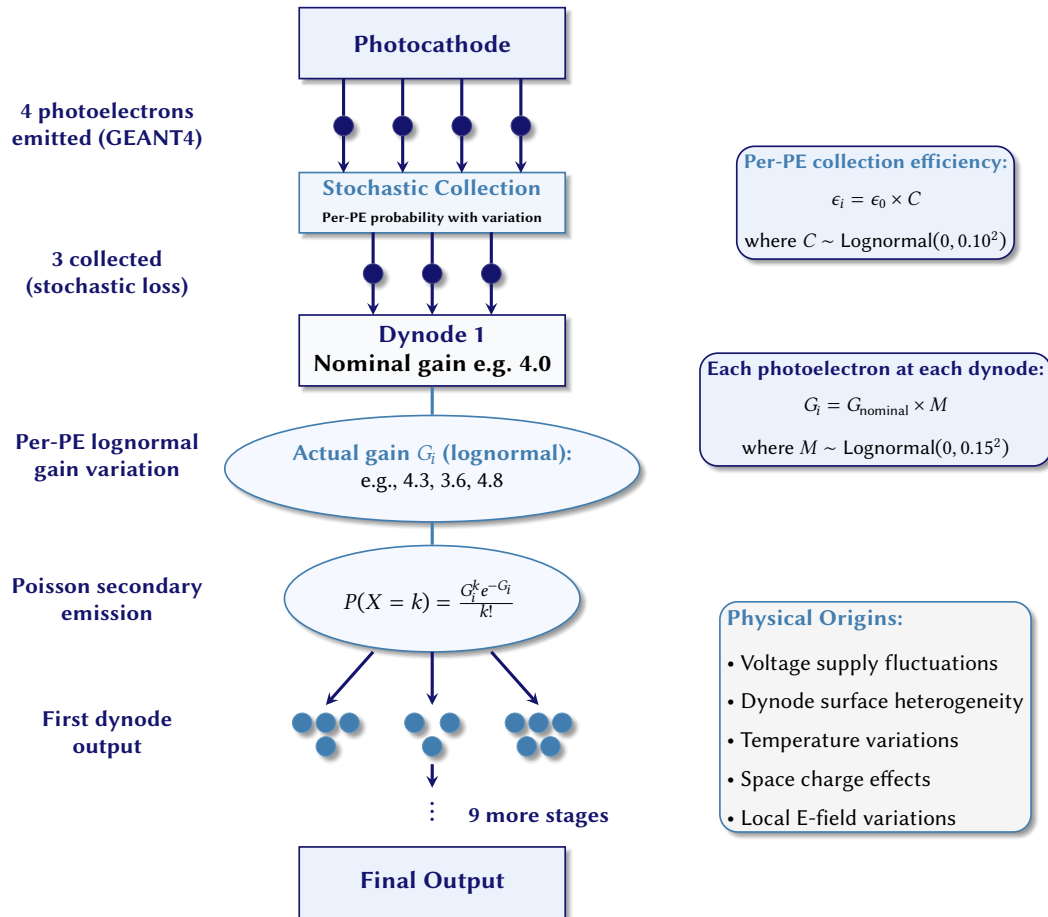


Figure 3.2: Schematic of the PMT dynode cascade simulation. Four photoelectrons emitted from the photocathode each undergo independent stochastic collection with probability $\epsilon_i = \epsilon_0 \times C$, where $C \sim \text{Lognormal}(0, 0.10^2)$. In this example, three photoelectrons reach the first dynode. Each collected photoelectron experiences independent gain variation at each dynode stage: $G_i = G_{\text{nominal}} \times M$, where $M \sim \text{Lognormal}(0, 0.15^2)$. Secondary emission follows Poisson statistics with mean equal to effective gain, producing variable electron numbers through nine additional dynode stages. This stochastic process, repeated independently for each photoelectron, reproduces characteristic statistical fluctuations in photomultiplier-based energy measurements.

Collection efficiency varies because emission angles differ, local electric fields are inhomogeneous, magnetic fields perturb trajectories, and space charge accumulates near the photocathode-dynode interface.

An important caveat: the GEANT4 optical simulation implemented uniform quantum efficiency across a homogeneous photocathode without fluctuations in the photon-to-photoelectron conversion. Consequently, the tuned collection efficiency parameters effectively absorb both true collection losses and unmodeled quantum efficiency variations. The actual geometric collection efficiency may differ from 70 %, with the effective value compensating for missing stochastic processes. This does not affect the complete chain’s predictive accuracy but should be considered when interpreting individual parameter values.

Gain variation affects each photoelectron independently at each dynode stage. While nominal dynode gains are determined by the voltage divider network and dynode materials, each photoelectron experiences fluctuations around these values. The effective gain for photoelectron i at dynode d is $G_{i,d} = G_{\text{nominal},d} \times M$, where M represents multiplicative noise sampled from a lognormal distribution $\ln M \sim \mathcal{N}(0, \sigma^2)$ with characteristic spread $\sigma = 0.15$. Experimental laser-based measurements validated both the lognormal form and fluctuation magnitude (Appendix B).

Gain varies because voltage supplies fluctuate, dynode surfaces are heterogeneous, temperatures shift, space charge accumulates, and local electric fields vary.

The choice of lognormal distributions for both collection efficiency and gain fluctuations follows from elementary statistical physics. When multiple independent multiplicative effects combine, the resulting distribution naturally becomes lognormal—a consequence of the central limit theorem applied in logarithmic space.

From Framework to Predictions

Figure 3.3 illustrates the logical progression from framework development to validation.

Section 3.3 predicts energy deposition and photon production under unshielded cosmic ray exposure. These predictions rely solely on well-established particle transport physics and light production, independent of photodetector modeling. This analysis reveals the characteristic cosmic muon energy deposition peak—a feature serving as the calibration reference for experimental validation. Understanding the unshielded

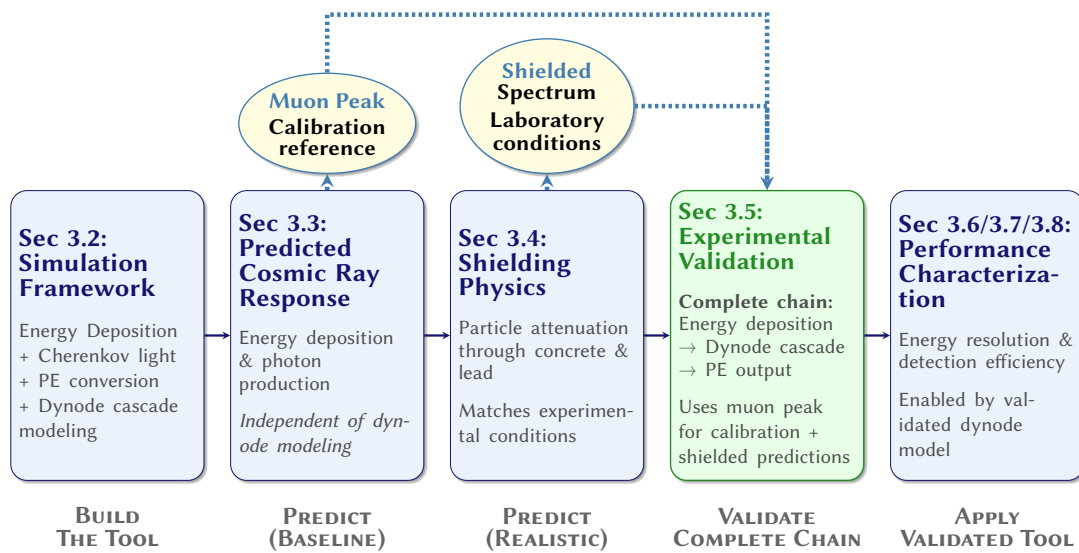


Figure 3.3: Logical flow of Chapter 3.

baseline establishes the reference against which shielding modifications can be quantified.

Section 3.4 incorporates realistic experimental conditions by simulating particle transport through concrete and lead shielding. Validation measurements with the detector inside the laboratory building were performed, making these predictions essential for meaningful comparison. The shielding analysis quantifies attenuation characteristics for each particle type, establishing the expected cosmic ray spectrum incident on the detector under laboratory conditions.

These foundational predictions—the unshielded baseline revealing the calibration peak and the shielded response matching experimental conditions—enable meaningful validation. Section 3.5 validates the complete detector response chain from particle transport and energy deposition through photodetector conversion (including dynode cascade modeling) to final photoelectron output. The validation employs the muon peak identified in Section 3.3 as the energy calibration reference and compares predictions incorporating shielding effects from Section 3.4 against experimental cosmic ray measurements. Excellent agreement between simulated and measured photoelectron spectra establishes confidence in both energy deposition physics and photodetector response modeling.

With the simulation framework validated, the remaining sections apply it to characterize intrinsic detector performance limits. Section 3.6 quantifies energy resolution as a function of deposited energy, establishing the fundamental statistical limits on

energy reconstruction. Section 3.7 evaluates detection efficiency for each cosmic ray particle type—muons, electrons, photons, neutrons, and protons—as functions of threshold settings, revealing which particles contribute to backgrounds and which are effectively rejected. Section 3.8 addresses the central question: signal detection efficiency at the target 10 MeV threshold. These performance predictions rely on the validated photodetector response model and establish the quantitative foundation for photodetector optimization studies in Chapter 4.

3.3 Predicted Cosmic Ray Response

Validating the complete detector response chain requires a calibration reference—a known energy signature that anchors experimental measurements to simulation predictions. Cosmic ray muons provide this reference. Minimum ionizing muons traversing the full crystal height deposit a characteristic energy determined by well-established stopping power physics, creating a sharp peak in the energy spectrum. This section predicts where that peak appears and characterizes the underlying energy deposition and photon production mechanisms. The analysis also reveals the full spectrum shape and particle composition, establishing which cosmic ray species dominate at different energies and how each interacts with PbF_2 .

The predictions presented here rely exclusively on particle transport and Cherenkov light production physics, independent of photodetector modeling. This separation is strategic: establishing baseline energy-deposition characteristics enables subsequent validation of the photodetector response (dynode cascade, collection efficiency, gain variations) by comparing complete-chain predictions against experimental measurements in Section 3.5.

The simulation encompasses 86,400 s (24 h) of detector operation in unshielded configuration, yielding a total interaction rate of 2.2 Hz.

Energy Deposition and Photon Production Characteristics

Figure 3.4 characterizes three fundamental aspects of detector response: the relationship between energy deposition and photon production, energy-dependent trajectory characteristics for each particle type, and the overall energy spectrum revealing the muon calibration peak.

Table: Median Primary Particle Energy [MeV] / Path Length [mm]

Edep [MeV]	5 MeV	12.5 MeV	25 MeV	37.5 MeV	50 MeV	62.5 MeV	75 MeV
Muon	2719/5	2688/13	1790/27	3218/34	3970/38	4518/42	5588/44
Gamma	6/14	17/10	38/9	62/8	89/7	114/7	156/5
Electron	7/5	18/10	37/15	60/19	86/21	126/23	140/22
Neutron	14/16	63/16	97/12	182/15	140/18	207/7	309/12
Proton	417/4	448/9	463/22	455/29	266/28	193/28	75/12
Pion	507/5	385/12	458/31	149/30	—	—	—

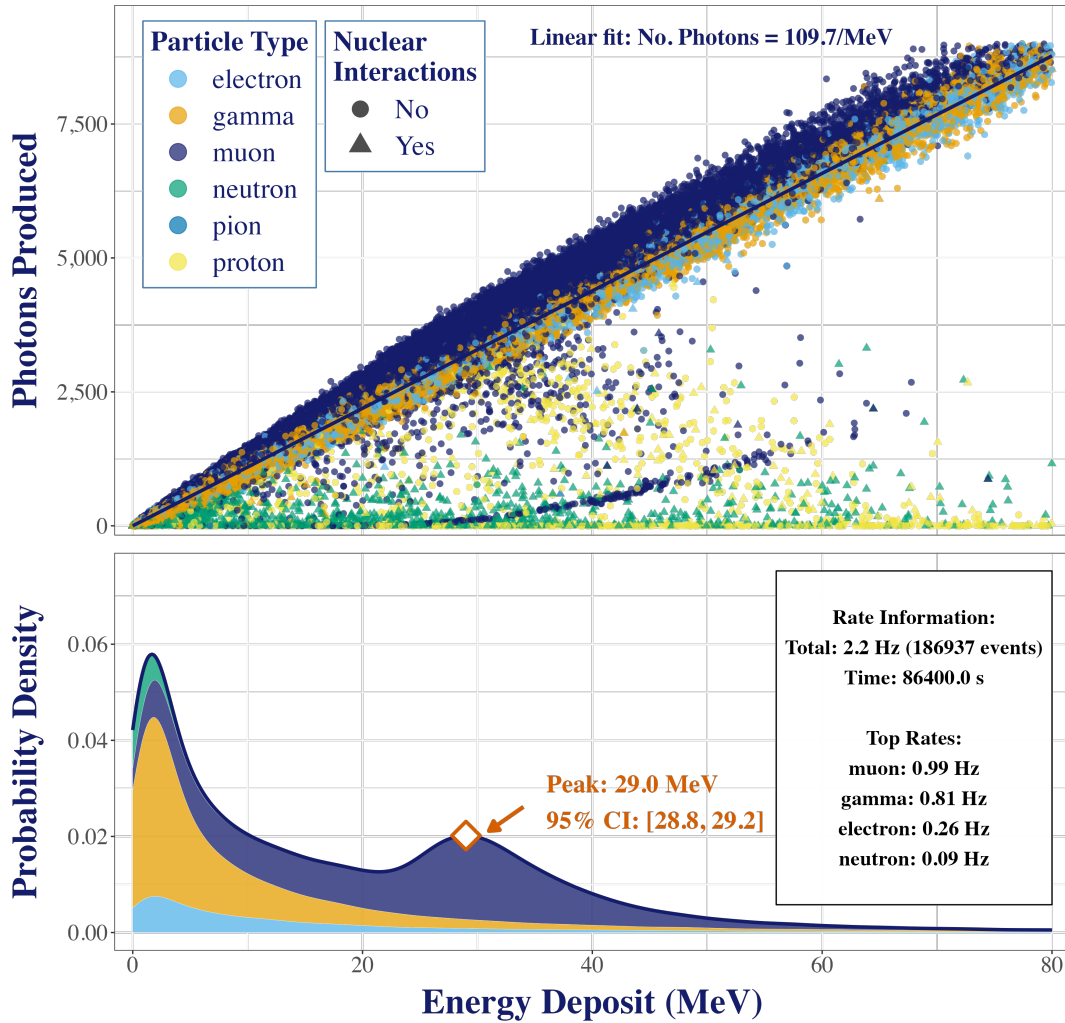


Figure 3.4: Energy deposition and photon production in PbF_2 under cosmic ray exposure. Top: Median particle energies and path lengths versus energy deposit. Middle: Photon yield versus energy deposit, with linear fit yielding 109.7 photons/MeV ($R^2 = 0.96$). Bottom: Energy deposition spectrum showing muon peak at 29.0 MeV (95 % CI: [28.8, 29.2] MeV), with particle-type contributions indicated by stacked colors.

Does Cherenkov Light Production Scale Linearly?

The scatter plot (middle panel) confirms that energy deposition and photon production scale predominantly linearly. The fitted conversion factor of 109.7 photons/MeV with excellent linearity ($R^2 = 0.96$) validates the basic detection mechanism and provides confidence in energy reconstruction capabilities.

Two classes of events deviate from this linear relationship. First, nuclear interactions (triangular markers) appear systematically below the main trend because recoiling nuclei fall below Cherenkov threshold ($\beta < 0.56$), depositing energy without producing detectable light. Second, a subset of muon events also underproduces photons—these are low-energy muons falling below Cherenkov threshold (kinetic energy < 40 MeV) during crystal traversal.

What Do Particle Trajectories Reveal?

The trajectory data (top panel) show how different particle types interact with the detector. Each exhibits characteristic behavior essential for background understanding.

Muons exhibit minimum ionizing behavior (~ 1 MeV/mm) up to 37.5 MeV. Longer path lengths at higher deposits reflect oblique trajectories. Muons striking at grazing angles traverse paths significantly exceeding the nominal 30 mm crystal height while maintaining their characteristic energy loss rate.

Gamma rays show systematically decreasing path lengths with increasing energy: 14 mm at 5 MeV declining to 5 mm at 75 MeV. This reflects the transition from Compton-dominated interactions at lower energies to pair-production-dominated interactions above 8 MeV, consistent with NIST cross-section data [35]. Low-energy gamma rays must undergo multiple Compton scattering events to deposit significant energy, resulting in longer effective path lengths. High-energy gamma rays undergo pair production within the first few millimeters, immediately initiating electromagnetic showers.

Electrons demonstrate a path length plateau at 22 mm to 23 mm above 85 MeV. This plateau reflects radiative energy loss through bremsstrahlung photons escaping the finite detector geometry, combined with geometric truncation. Cosmic electrons exhibit more diffuse angular distribution than muons, causing many to exit the crystal before reaching their theoretical stopping range.

Neutrons exhibit highly variable path lengths (7 mm to 18 mm) with no clear energy correlation, reflecting stochastic nuclear interactions. Energy deposits result from

nuclear recoil and gamma emission following elastic scattering, inelastic scattering, and nuclear reactions—not direct ionization. The wide variation demonstrates the probabilistic nature of neutron-nucleus interactions: individual neutrons may undergo high-energy-transfer collisions immediately upon entering or travel significant distances before interacting. This contrasts with the predictable ranges of charged particles and confirms the inherently low detection efficiency essential for dark matter searches.

Protons show three distinct interaction regimes reflecting complex interplay between angular distribution, energy-dependent stopping power, and detector geometry. Low deposits (5 MeV to 25 MeV) originate from high-energy protons at glancing angles that exit before significant energy deposition. Medium deposits (~ 37.5 MeV) correspond to protons traversing full detector thickness at steep zenith angles. High deposits (50 MeV to 75 MeV) result from lower-energy protons experiencing Bragg peak effects, where stopping power increases as $dE/dx \propto 1/E$ near end of range.

The Muon Calibration Peak

The energy deposition spectrum (bottom panel) reveals the key feature motivating this section: a sharp peak at 29.0 MeV with 95 % confidence interval [28.8, 29.2] MeV. This corresponds predominantly to minimum ionizing muons traversing the full 30 mm detector height, though a small gamma ray contribution slightly affects the peak position. Subsequent analyses throughout this thesis adopt 29 MeV as the calibration reference; while not exact, this approximation introduces negligible error compared to other systematic uncertainties in energy scale determination.

The particle composition varies with energy. Muons (dark blue) dominate the peak region. The low-energy region (0 MeV to 20 MeV) exhibits continuous distribution from multiple sources: partial detector traversal at glancing angles, secondary electrons from electromagnetic showers, partial containment of low-energy electron cascades from gamma interactions, and nuclear recoil events from neutron scattering. Understanding this low-energy behavior proves crucial for achieving the 10 MeV detection threshold required for dark matter searches.

The high-energy tail extending beyond 40 MeV represents complex event topologies: diagonal muon trajectories with extended path lengths, electromagnetic shower development from high-energy electrons and photons, and rare high-energy-transfer

nuclear interactions. This tail follows approximately exponential decay characteristic of stochastic energy deposition in finite detector volumes.

Absolute Rate Validation

The simulated muon rate of 0.99 Hz for the crystal's upper surface area (approximately 55 cm²) agrees well with the established expectation of approximately one muon per minute per square centimeter at sea level. This agreement validates the CRY library implementation and establishes confidence in absolute rate predictions essential for background modeling.

Particle Interaction Characteristics

Beyond energy deposition and photon production, understanding what fraction of cosmic ray particles actually produce detectable signals provides crucial insight into detector response mechanisms. Figure 3.5 categorizes particles by their interaction outcomes.

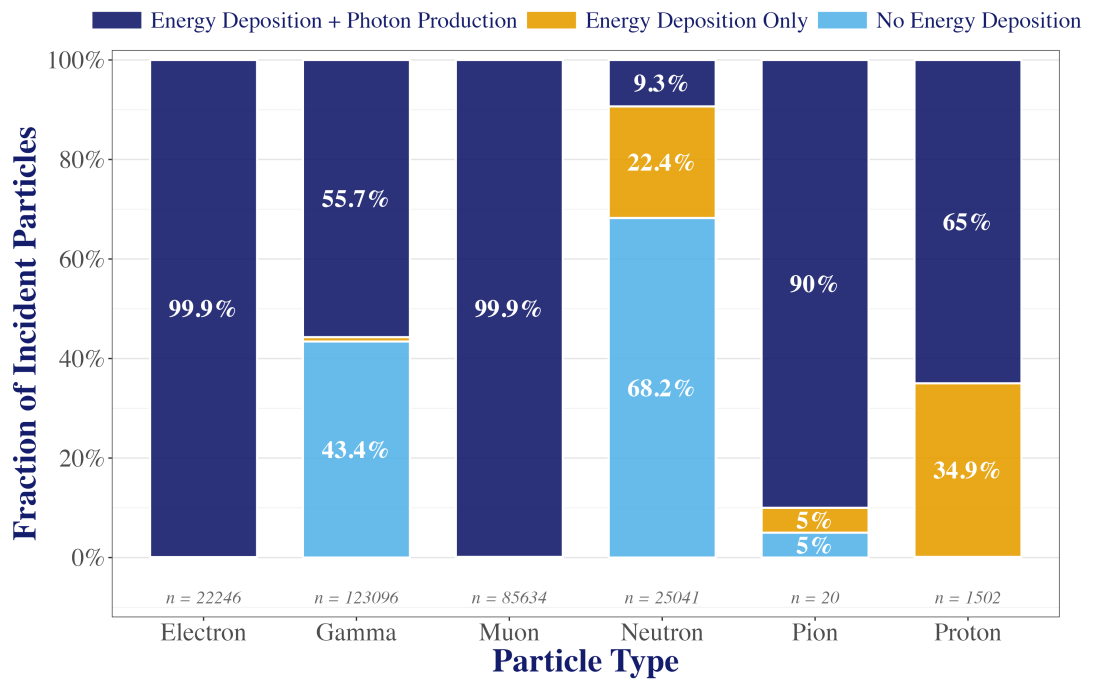


Figure 3.5: Particle interaction outcomes in PbF₂ crystals. Each bar shows the fraction resulting in energy deposition with photon production (dark blue), energy deposition without photon production (orange), or no energy deposition (light blue). Sample sizes indicated below bars. Analysis spans 86,400 s of detector operation.

Electromagnetic Interactions

Electrons and muons demonstrate exceptionally high interaction probabilities: 99.99 % of incidents produce both energy deposition and Cherenkov photons.

Gamma rays exhibit bimodal behavior: 55.7 % produce both energy deposition and photon generation, while 43.4 % traverse without any energy transfer. This 43.4 % non-interaction fraction corresponds to photons undergoing neither photoelectric absorption, Compton scattering, nor pair production within the finite detector volume. The small remaining fraction (1 %) depositing energy without photons represents low-energy secondary electrons below Cherenkov threshold.

Pions also demonstrate predominantly electromagnetic behavior: 90 % of interactions produce both energy and photon signals, reflecting their charged nature. The small fractions of alternative outcomes (5 % each) correspond to nuclear interactions and boundary effects where particles exit before significant energy deposition.

Hadronic Interactions

Neutrons display the most complex signature: 68.2 % produce no detectable signal, 22.4 % deposit energy without photon production, and only 9.3 % generate both energy deposition and Cherenkov light.

Neutrons interact primarily through nuclear processes—elastic scattering, inelastic scattering, and nuclear reactions—which often transfer energy to heavy nuclei. Energy deposited through nuclear recoils produces no optical signal, rendering these events invisible. The 9.3 % producing both energy and photons correspond to interactions creating energetic charged secondaries exceeding Cherenkov threshold. This inherently low neutron detection efficiency represents a crucial advantage for dark matter searches.

Protons exhibit intermediate behavior: 65 % produce both energy deposition and photon generation, 34.9 % deposit energy without Cherenkov light, and a negligible fraction passes without interaction. This distribution primarily reflects the Cherenkov threshold: protons require approximately 200 MeV kinetic energy to exceed $\beta > 0.56$ in PbF_2 . Many cosmic ray protons fall below this threshold, depositing energy through ionization without producing detectable light.

3.4 Background Suppression Through Shielding

The unshielded predictions in Section 3.3 establish baseline detector response and identify the muon calibration peak at 29 MeV. However, the detector operates within laboratory buildings constructed with concrete walls and ceilings and is surrounded by additional lead shielding to reduce electromagnetic backgrounds. The actual background spectrum reaching the detector depends on how these materials attenuate different particle types. This spectrum provides the reference for experimental comparison in Section 3.5.

This section characterizes energy-dependent attenuation for cosmic ray particles in both concrete and lead. The analysis addresses three questions: How effectively does each material suppress electromagnetic backgrounds (electrons, photons)? What attenuation do hadronic particles (neutrons, protons) experience? What background composition remains after realistic shielding configurations? A final consideration concerns the distinction between removing primary particles from a beam and actually absorbing their kinetic energy. Secondary particle production can transmit substantial energy through shielding even when primary particles are efficiently stopped.

Simulation Methodology

Dedicated GEANT4 simulations quantified shielding effectiveness using hollow cubic geometries surrounding the PbF₂ crystal volume. Wall thicknesses ranged from 2 cm to 10 cm (in 2 cm increments) for lead and 10 cm to 40 cm (in 10 cm increments) for concrete.

Material compositions followed standard specifications. Concrete shielding used the standard GEANT4 built-in material G4_CONCRETE¹ ($\rho = 2.3 \text{ g/cm}^3$). High-purity lead was simulated at 11.3 g/cm^3 .

Cosmic ray generation followed the methodology in Section 3.2, with particles generated on a $3 \text{ m} \times 3 \text{ m}$ plane above the shielded detector assembly. Attenuation factors were calculated by comparing event rates with and without shielding for each particle type and energy bin.

¹Composition by mass fraction: H (1.0 %), C (0.1 %), O (52.91 %), Na (1.6 %), Mg (0.2 %), Al (3.39 %), Si (33.70 %), K (1.3 %), Ca (4.4 %), and Fe (1.4 %).

Energy-Dependent Attenuation Analysis

The median attenuation values presented in Figure 3.6 provide useful summary statistics but obscure the strong energy dependence of particle interactions with matter. Extracting energy-resolved attenuation characteristics requires careful consideration of the distinct interaction mechanisms for neutral and charged particles.

Neutral Particle Attenuation

For neutral particles (photons and neutrons), attenuation in matter follows the Beer-Lambert law:

$$I(x) = I_0 e^{-\mu(E)x} \quad (3.2)$$

where $I(x)$ represents transmitted flux after penetrating thickness x , I_0 denotes incident flux, and $\mu(E)$ is the energy-dependent linear attenuation coefficient (cm^{-1}). This exponential form arises because neutral particle interactions are stochastic and discrete: each particle either traverses without interaction or is removed from the beam through absorption or scattering. The probability of surviving to depth x decreases exponentially with the number of potential interaction sites encountered.

Taking the natural logarithm yields:

$$\ln \left(\frac{I(x)}{I_0} \right) = -\mu(E) \cdot x \quad (3.3)$$

This linear relationship enables extraction of $\mu(E)$ from the slope of linear regression between $\ln(I/I_0)$ and thickness x .

Charged Particle Interactions

Charged particles (electrons, muons, protons) interact through different mechanisms that preclude simple exponential attenuation. Charged particles lose energy continuously through electromagnetic interactions with atomic electrons along their trajectory. The Bethe-Bloch formula describes this stopping power $-dE/dx$, which depends on particle velocity and the material's electron density.

For electrons, ionization losses combine with radiative losses (bremsstrahlung) that become dominant above the critical energy (approximately 8 MeV) in lead. Muons, being minimum-ionizing particles at cosmic ray energies, lose energy at a nearly constant rate of approximately 2 MeV per g/cm^2 , penetrating substantial material

thicknesses before stopping. Protons exhibit similar continuous energy loss with a pronounced Bragg peak near end-of-range.

Despite these physical differences, the simulation methodology yields valid empirical transmission fractions for all particle types.

Transmission Fraction Methodology

For each energy bin and shielding thickness, event rates R were calculated by dividing detected events by simulation duration. The transmission fraction:

$$T(E, x) = \frac{R(E, x)}{R(E, x = 0)} \approx \frac{I(x)}{I_0} \quad (3.4)$$

represents the measured fraction of particles at energy E transmitted through thickness x relative to the unshielded configuration. Uncertainties were quantified using bootstrap resampling with 1,000 iterations.

The transmission fraction characterizes the survival of *primary* particles, not the total energy transmitted through shielding. Secondary particles produced in interactions can carry substantial kinetic energy forward even when primary particles are efficiently stopped.

Theoretical Predictions for Neutral Particles

Simulated attenuation coefficients for gamma rays and neutrons were compared against theoretical predictions from nuclear databases to validate the simulation framework. For gamma rays, photon interaction cross-sections from the NIST XCOM database [35] provide mass attenuation coefficients (μ/ρ) as functions of energy. Linear attenuation coefficients follow:

$$\mu_\gamma(E) = \left(\frac{\mu}{\rho} \right) \cdot \rho_{\text{material}} \quad (3.5)$$

Neutron cross-sections for lead and individual elements (H, O, Si, Al, Ca, Fe) were obtained from ENDF-B/VIII.0 nuclear data libraries [36]. For concrete—a composite material—the macroscopic cross-section is:

$$\Sigma_{\text{concrete}}(E) = \sum_i N_i \sigma_i(E) \quad (3.6)$$

where N_i is the number density of element i (cm^{-3}) and $\sigma_i(E)$ is its microscopic cross-section (cm^2 per atom). Number density is given by:

$$N_i = \frac{w_i \rho_{\text{concrete}} N_A}{A_i} \quad (3.7)$$

where w_i is mass fraction, $N_A = 6.022 \times 10^{23} \text{ mol}^{-1}$ is Avogadro's constant, and A_i is atomic mass (g/mol). The complete mixing rule becomes:

$$\mu_{\text{concrete}}(E) = \sum_i \frac{w_i}{\rho_{\text{concrete}}} \rho_i \mu_i(E) \quad (3.8)$$

An important limitation affects this comparison: ENDF libraries contain narrow resonance structures in neutron cross-sections (sometimes only a few keV wide), while simulation energy binning spans tens to hundreds of MeV. Consequently, fine resonance features are averaged in the simulated attenuation spectra.

Figure 3.6 presents comprehensive attenuation analysis for both lead and concrete, showing energy-dependent attenuation coefficients for gamma rays and neutrons with bootstrap-derived confidence intervals, alongside empirical transmission fractions for all particle types.

Gamma Ray Attenuation

Gamma ray attenuation results demonstrate excellent agreement between simulation and theoretical predictions across the energy range from 1 MeV to 1,000 MeV. For both materials, the attenuation coefficient exhibits characteristic energy dependence reflecting transitions between dominant interaction mechanisms.

In lead, the attenuation coefficient decreases from approximately 0.7 cm^{-1} at 1 MeV to a minimum of approximately 0.48 cm^{-1} around 5 MeV, then increases steadily to approximately 1.0 cm^{-1} at 100 MeV. This pattern reflects the transition from photoelectric-dominated interactions at low energies ($\sigma_{\text{PE}} \propto Z^5$), through the Compton minimum, to pair-production dominance at high energies ($\sigma_{\text{pair}} \propto Z^2$). Lead's high atomic number ($Z = 82$) enhances both photoelectric absorption and pair production cross-sections, resulting in gamma attenuation exceeding 97 % for 4 cm thickness and 99.8 % for 10 cm thickness.

Concrete exhibits substantially lower attenuation coefficients due to its lower average atomic number, with values decreasing from approximately 0.12 cm^{-1} at 1 MeV to a

Table: Median Attenuation

Lead Thickn.	2 cm	4 cm	6 cm	8 cm	10 cm
Gamma	86.7%	97%	98.2%	99.4%	99.8%
Neutron	27.7%	42.9%	47%	59.5%	68.3%
Muon	1.5%	2.7%	0.1%	1.6%	2.9%
Electron	96.7%	—	—	—	—
Proton	27.7%	15.1%	19.5%	25.7%	33.3%

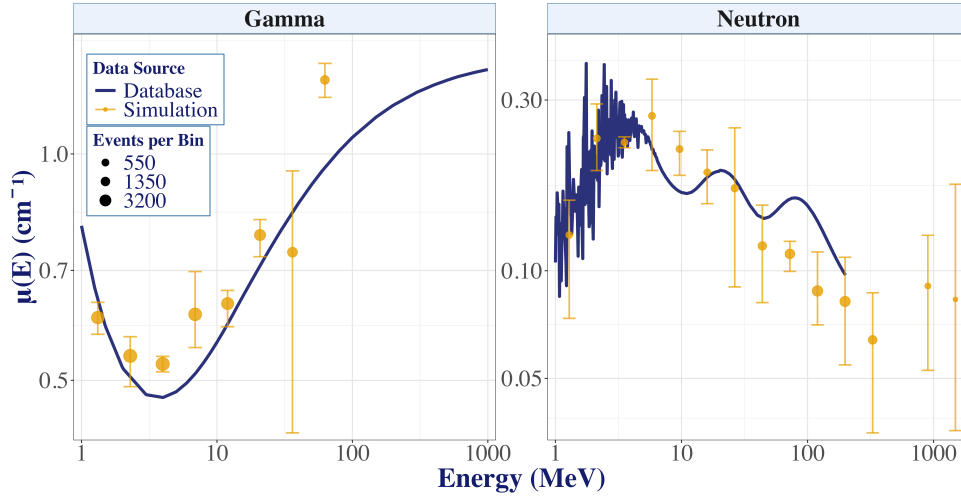


Table: Median Attenuation

Concrete Thickn.	10 cm	20 cm	30 cm	40 cm
Gamma	48.5%	70.6%	87.4%	92.6%
Neutron	47.8%	74.8%	86.8%	85.3%
Muon	1.9%	1.8%	0.6%	3.1%
Electron	45.8%	94.8%	98.3%	99.1%
Proton	35.7%	49.6%	60.7%	80%

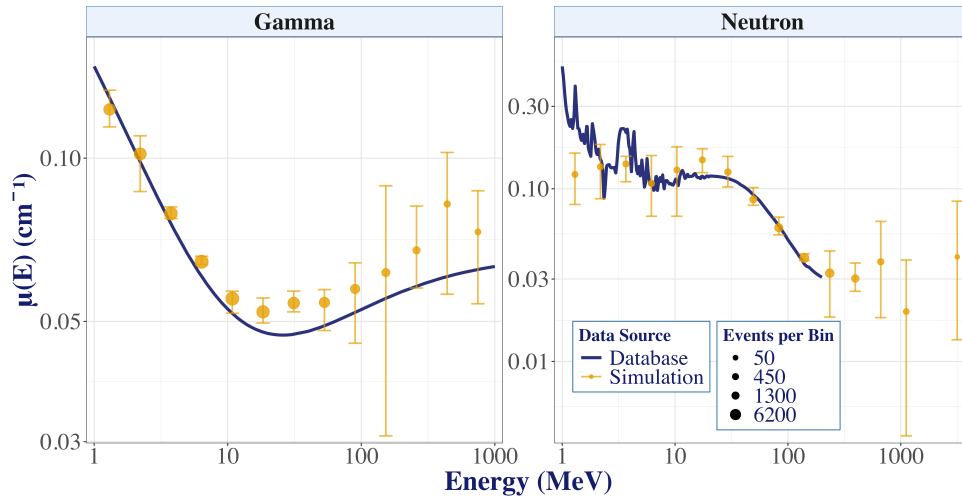


Figure 3.6: Attenuation characteristics in lead (top) and concrete (bottom) shielding. Upper panels: median attenuation probabilities versus shielding thickness for different particle types. Lower panels: energy-dependent linear attenuation coefficients $\mu(E)$ for gamma rays (left) and neutrons (right), where the Beer-Lambert formalism applies. Simulation-derived values (yellow points) with 95 % bootstrap confidence intervals ($n = 1,000$) compared against theoretical predictions (blue curves). These coefficients describe primary particle removal.

minimum of approximately 0.05 cm^{-1} around 20 MeV to 50 MeV, followed by a slight increase at higher energies. The flatter energy dependence reflects dominant Compton scattering contributions from intermediate- Z elements. While less effective per unit thickness than lead, concrete achieves 87.4 % gamma attenuation at 30 cm thickness, providing cost-effective bulk shielding suitable for experimental facility construction.

Hadronic Attenuation

Neutron attenuation characteristics reveal complex energy dependencies reflecting the interplay of elastic scattering, inelastic scattering, and nuclear absorption. The theoretical curves exhibit pronounced resonance structures at low energies.

In lead, the neutron attenuation coefficient peaks at approximately 0.25 cm^{-1} to 0.30 cm^{-1} in the energy range from 2 MeV to 5 MeV, then decreases monotonically to approximately 0.08 cm^{-1} at 1,000 MeV. The theoretical curve displays dense resonance structure below 10 MeV, corresponding to nuclear resonances in ^{208}Pb . Despite lead's heavy nucleus ($A = 208$), which provides less efficient energy transfer per elastic collision than lighter nuclei, the high nuclear density and significant inelastic scattering cross-sections yield measurable neutron attenuation, reaching 68.3 % for 10 cm thickness.

Concrete demonstrates lower neutron attenuation coefficients than lead across most of the energy range. The attenuation coefficient remains relatively constant at approximately 0.10 cm^{-1} to 0.12 cm^{-1} from 1 MeV to 30 MeV, then decreases to approximately 0.03 cm^{-1} at 200 MeV. The theoretical curve shows resonance structures at low energies arising from constituent nuclei, particularly silicon, calcium, and oxygen. Concrete achieves 47.8 % neutron attenuation at 10 cm thickness, increasing to 85.3 % at 40 cm.

Energy Transmission and Secondary Production

The attenuation analysis above quantifies primary particle removal but incompletely characterizes shielding effectiveness. When a gamma ray undergoes pair production, the primary photon disappears from the beam while an electron-positron pair carries most of the original energy forward. When a neutron scatters inelastically, secondary particles may emerge with substantial kinetic energy. Electromagnetic and hadronic showers involve multiple generations of such processes, precluding analytical prediction of total transmitted energy.

To quantify energy transmission, a separate GEANT4 simulation tracked complete energy budgets for monoenergetic particles incident on planar slabs. The geometry consisted of a material slab with front face at $z = 0$ and back face at $z = d$, where d ranged from 2 cm to 10 cm for lead and 10 cm to 40 cm for concrete. Incident energies spanned 10 MeV to 1,000 MeV for both gamma rays and neutrons.

For each event, the simulation recorded three quantities: energy deposited within the slab (E_{dep}), kinetic energy of all particles escaping backward (E_{back}), and kinetic energy of the primary particle if it exited forward ($E_{\text{primary}}^{\text{out}}$). Conservation of energy then gives the total forward-transmitted energy:

$$E_{\text{forward}} = E_0 - E_{\text{dep}} - E_{\text{back}} \quad (3.9)$$

where E_0 is the incident particle energy. This forward energy includes both surviving primaries and all secondary particles propagating downstream. Averaging over many events yields the mean forward fraction $\langle E_{\text{forward}} \rangle / E_0$, which characterizes overall energy transmission regardless of whether primaries or secondaries carry that energy. Figure 3.7 presents these transmission fractions as heatmaps across incident energy and thickness.

The heatmaps reveal behavior that attenuation coefficients alone obscure. For gamma rays in lead, high primary-particle attenuation corresponds to genuine energy absorption: electromagnetic cascades develop rapidly in high- Z material and deposit energy locally by ionizing the dense electron cloud. Forward fractions fall below 25 % for lead thicker than 4 cm across all energies examined. Concrete, despite attenuating gamma rays appreciably (Figure 3.6), transmits substantially more energy. Forward fractions of 50 % to 75 % persist even at 30 cm for higher-energy gammas. The lower atomic number reduces both photoelectric absorption and pair-production cross-sections, allowing more energy to escape as forward-scattered Compton electrons and transmitted secondary photons.

The most striking discrepancy between particle attenuation and energy transmission appears for neutrons in lead. Although 10 cm of lead attenuates 68.3 % of incident neutrons, forward energy fractions remain above 75 % across all thicknesses and energies. Kinetic energy propagates forward through multiple small-angle scatters or emerges in secondary particles from inelastic reactions.

Concrete absorbs neutron energy more effectively, particularly below 50 MeV. Its hydrogen content—only about 1 % by mass but significant by number density—enables

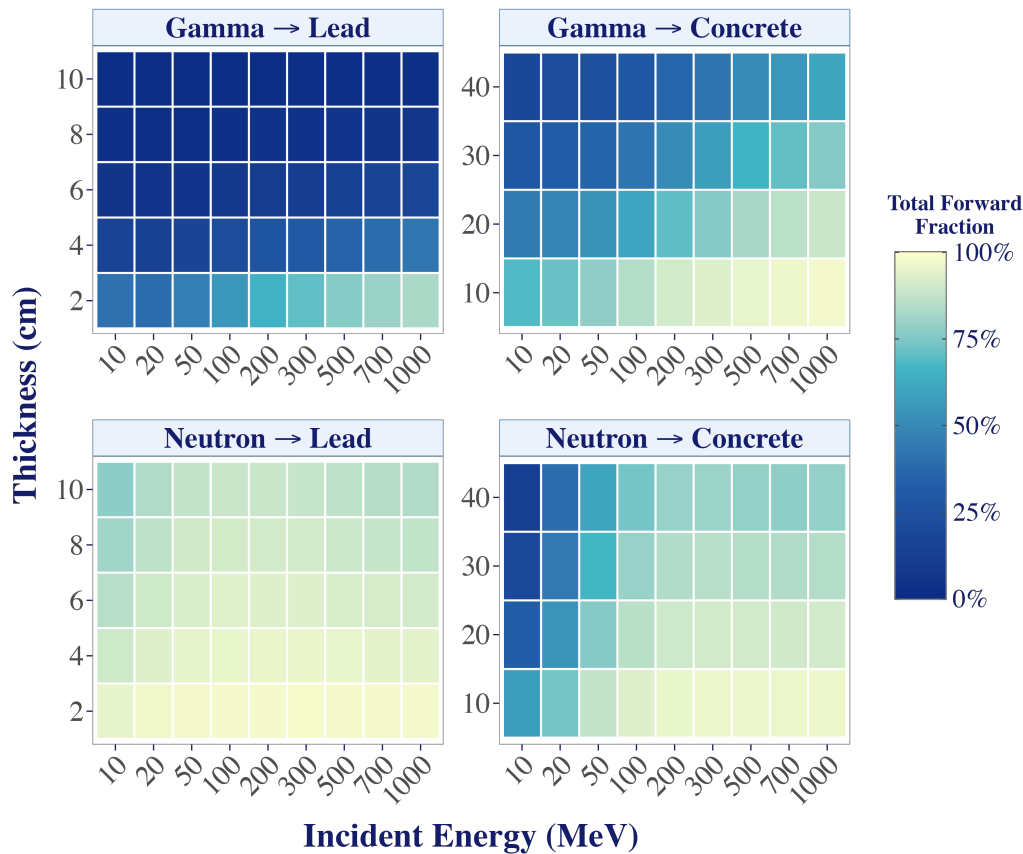


Figure 3.7: Mean forward energy transmission fractions for gamma rays (top) and neutrons (bottom) through lead (left) and concrete (right) shielding. Color scale indicates the fraction of incident kinetic energy emerging downstream, including both surviving primaries and all secondaries.

efficient energy transfer through n-p elastic scattering. Head-on collisions can transfer up to 100 % of neutron kinetic energy to recoil protons, which then deposit that energy locally through ionization. At 40 cm, concrete reduces forward fractions below 25 % for energies under 20 MeV. At higher energies, where elastic cross-sections shrink and inelastic processes dominate, concrete's advantage diminishes.

Expected Background Composition

These attenuation and transmission results together establish quantitative expectations for background reduction. Three key findings emerge:

Electromagnetic suppression proves both substantial and genuine. Two centimeters of lead attenuates primary electrons by more than 95 % and gamma rays by about 87 %. This particle-count reduction reflects actual energy absorption: forward energy fractions fall below 50 % even for thin lead layers, and thicknesses beyond 6 cm

contain nearly all incident energy across the range examined. Combining 20 cm to 40 cm of concrete overburden with 2 cm of lead shielding reduces the initial gamma ray flux of roughly 1 Hz to an estimated 0.01 Hz to 0.04 Hz—small compared to the unshielded rate, but not negligible.

Muons dominate the residual spectrum. Both materials attenuate muons very weakly (less than 5 % reduction through 40 cm of concrete), confirming that muons dominate cosmic ray backgrounds even with substantial shielding. Minimum-ionizing muons lose energy primarily through continuous ionization rather than catastrophic interactions, traversing typical shielding with minimal deflection or energy loss. The muon peak at 29 MeV therefore serves as the primary calibration reference in Section 3.5.

Neutron backgrounds require different strategies. Lead shields poorly against hadronic backgrounds: despite measurable attenuation coefficients, it transmits most incident neutron energy forward through inefficient momentum transfer and secondary production. Effective neutron energy absorption requires either hydrogen-rich moderators (explaining concrete’s superior low-energy performance) or reliance on the intrinsically low detection efficiency of neutrons in PbF_2 (Section 3.3). In practice, both mechanisms contribute. Passive concrete shielding reduces the incident flux and degrades neutron energies through moderation, while neutrons reaching the detector face low probability of generating detectable Cherenkov signals.

These results predict the background environment for experimental validation: a muon-dominated spectrum with residual electromagnetic and hadronic components. This understanding enables meaningful comparison between shielded simulations and laboratory measurements in Section 3.5, and guides interpretation of secondary particle production in veto system studies.

3.5 Experimental Validation

The simulation framework predicts detector response from first principles: particle transport, Cherenkov light production, optical propagation, photocathode conversion, collection efficiency losses, and stochastic dynode multiplication. However, these predictions remain theoretical until confronted with experimental data. Without validation, the framework cannot reliably predict performance for the photodetector optimization in Chapter 4.

This section validates the complete detector response chain through comparison with cosmic ray measurements. The validation exploits the muon peak identified in Section 3.3 as an energy calibration reference ($E_{\text{ref}} = 29 \text{ MeV}$), enabling quantitative comparison between predicted and measured photoelectron spectra. Agreement establishes confidence in both energy deposition physics and photodetector response modeling—particularly the dynode cascade parameterization crucial for efficiency predictions.

Measurement Configuration

Validation measurements were performed using a 4×4 array of PbF_2 crystals, each coupled to individual photomultiplier tubes, as described in Kontogoulas’s calibration work [31]. A single center crystal and its PMT output were selected for detailed analysis. Event selection required simultaneous triggers the selected crystal and at least one of the remaining 15 crystals. This coincidence requirement rejects dark current and electrical noise, ensuring the analyzed dataset contains only genuine particle interactions.

The coincidence requirement also enhances muon selection. Section 3.4 demonstrated that cosmic ray backgrounds are muon-dominated in shielded environments, with electrons, photons, and other charged particles experiencing substantial attenuation through modest concrete thicknesses. Cosmic ray muons possess sufficient penetrating power to trigger multiple crystals in the array; lower-energy background particles are less likely to satisfy the coincidence condition.

The simulation validation therefore focuses exclusively on muon events extracted from the comprehensive cosmic ray simulation described in Section 3.3. This muon-filtered dataset provides the theoretical baseline for comparison against experimental measurements.

Light Yield Validation

Figure 3.8 presents the comparison results. The upper panel illustrates the relationship between energy deposition and photoelectron-equivalent response for the simulation chain. The raw photoelectron output from GEANT4 (green circles) achieves a light yield of 2.8 PE/MeV with exceptional linearity ($R^2 = 0.98$). This value represents theoretical photoelectron generation accounting for Cherenkov photon production, optical transport losses, and aged photocathode quantum efficiency.

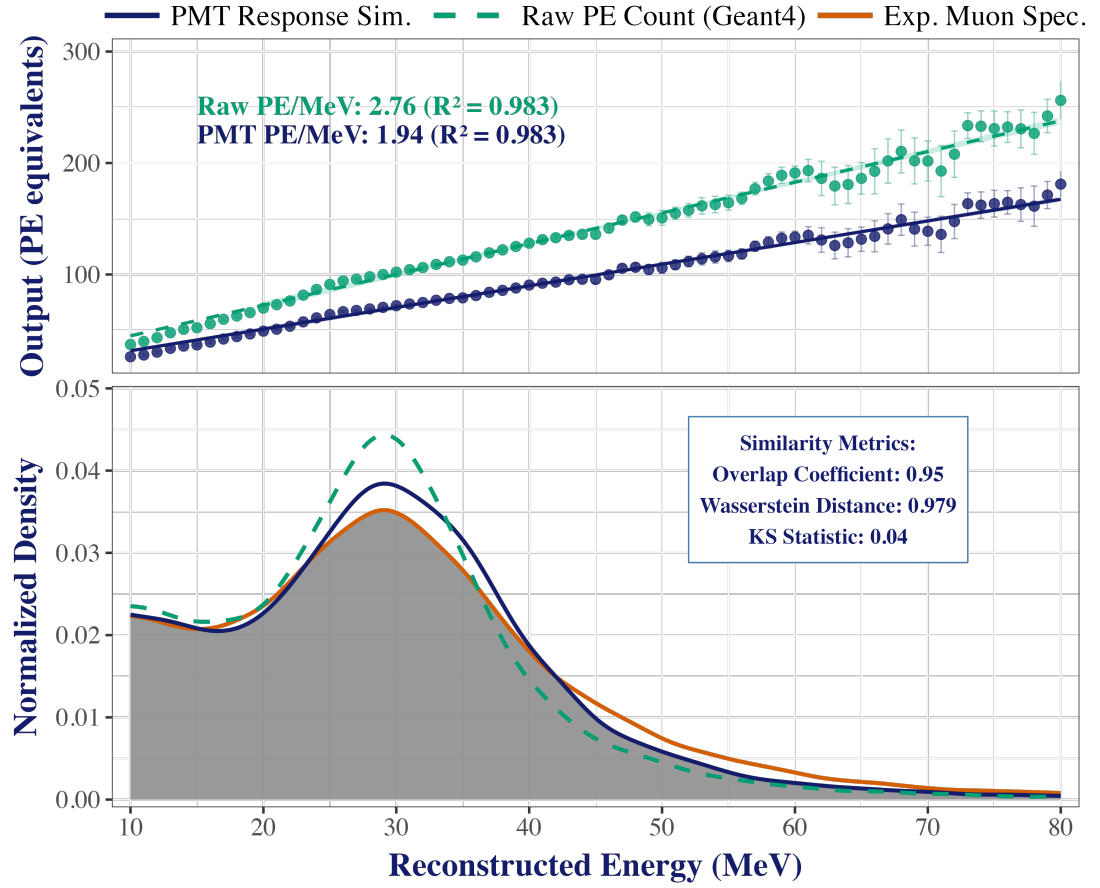


Figure 3.8: Validation of the simulation framework through cosmic ray measurements. Upper panel: photoelectron-equivalent response versus simulated energy deposit, comparing raw photoelectron output from GEANT4 (green circles, 2.8 PE/MeV, $R^2 = 0.98$) with output after dynode cascade simulation (blue circles, 1.9 PE/MeV, $R^2 = 0.98$). Lower panel: normalized energy deposition spectra showing raw photoelectron distribution from GEANT4 (dashed green line), dynode cascade simulation output (solid blue line), and experimental cosmic ray measurements (solid orange line).

After dynode cascade simulation (blue circles), the yield drops to 1.9 PE/MeV ($R^2 = 0.98$). This 30 % reduction arises from collection inefficiencies in the dynode cascade model. This output is consistent with the experimentally determined range of 1.86 PE/MeV to 2.04 PE/MeV reported by Christmann [25] using identical PbF_2 crystals with PMT readout.

The excellent linearity maintained across both stages confirms that dominant response mechanisms—Cherenkov light production and photodetector multiplication—behave linearly across the relevant energy range.

Spectral Shape Validation

The lower panel of Figure 3.8 compares normalized energy deposition spectra, revealing progressive spectral modifications introduced by detector physics modeling. The raw photoelectron spectrum from GEANT4 (dashed green line) exhibits the narrowest distribution, representing energy deposition statistics without instrumental broadening. This distribution reflects only intrinsic stochastic variations in Cherenkov photon production and Poisson photoelectron statistics.

The dynode cascade simulation (solid blue line) introduces additional spectral broadening from statistical fluctuations in photoelectron multiplication, collection efficiency variations, and cascade noise. The broadening from raw spectrum to cascade output quantifies the performance degradation from realistic detector implementation compared to idealized theoretical response.

The collection efficiency modeling parameters described in Section 3.2 ($\epsilon_0 = 0.70$ and $\sigma = 0.10$) were determined through iterative adjustment based on visual comparison with experimental data, rather than formal optimization procedures.

The experimental cosmic ray spectrum (solid orange line) agrees well with the dynode cascade simulation. Quantitative metrics confirm this agreement: overlap coefficient of 0.95 indicates substantial distribution overlap, Wasserstein distance of 0.98 reflects minimal distribution separation, and Kolmogorov-Smirnov statistic of 0.04 indicates no significant difference between distributions at standard significance levels.²

Validation Summary

Simulation and experiment agree through two independent metrics. First, the light yield of 1.9 PE/MeV after dynode cascade simulation matches Christmann’s independent beam-test measurements (1.86 PE/MeV to 2.04 PE/MeV) using identical crystals and PMT readout. Second, the spectral shape agreement confirms that statistical fluctuations governing event-by-event variation are accurately modeled.

This validated framework enables the quantitative predictions developed in the remainder of this chapter. The fluctuation model directly determines energy resolution—how

²The Wasserstein distance quantifies the minimum “cost” of transforming one distribution into another—the integrated probability mass times transport distance—with values near zero indicating identical shapes. The Kolmogorov-Smirnov statistic measures the maximum absolute difference between cumulative distribution functions, $D = \max |F_1(x) - F_2(x)|$, with values near zero indicating statistically indistinguishable distributions.

precisely deposited energy can be reconstructed from photoelectron signals. Combining light yield with the fluctuation model predicts detection efficiency as a function of photoelectron threshold: what fraction of events at a given energy produce signals exceeding a given threshold. This efficiency prediction, applied to both cosmic ray backgrounds and the target 10 MeV signal energy, establishes threshold requirements for reliable dark matter detection—the foundation for photodetector optimization in Chapter 4.

3.6 Energy Resolution

The validated simulation framework predicts resolution from the complete probability distribution of detector response, accounting for statistical fluctuations in Cherenkov photon production, photoelectron conversion, and dynode cascade multiplication. This section characterizes energy resolution from 1 MeV to 100 MeV, then validates predictions against independent measurements spanning three orders of magnitude.

Measurement Approach

Standard calorimeter energy resolution follows the photon counting statistics relationship:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E/\text{MeV}}} \quad (3.10)$$

where σ represents the standard deviation of the energy distribution, E is deposited energy in MeV, and a is the stochastic noise parameter. The general expression includes additional constant (b) and inverse energy (c/E) terms added in quadrature. However, fitting these additional contributions consistently yielded negligible or unphysical negative values, confirming that photon counting statistics dominate and justifying the simplified form.

Determining the stochastic parameter a from simulated energy distributions requires careful statistical treatment. Binned histogram analysis suffers from systematic bias: arbitrary bin width selection introduces variance inflation, making resolution measurements depend on binning choice rather than underlying physics. This problem particularly affects resolution characterization, where measured widths are sensitive to histogram construction.

Maximum likelihood estimation (MLE) circumvents this limitation by fitting probability distributions directly to unbinned data, eliminating binning-dependent artifacts and providing statistically optimal parameter estimates.

Accounting for Asymmetric Response

The measured energy distributions exhibit pronounced asymmetry, particularly at lower energies, due to multiplicative noise in the PMT dynode cascade. Gaussian distributions cannot capture this asymmetry. Therefore skew-normal distributions with energy-dependent parameters model the underlying response.

To represent the decrease in asymmetry with increasing energy, the skewness was modeled using an exponential decay of the form: $\alpha(E) = \alpha_0 \exp(-E/E_{\text{decay}})$. At low energies cascade multiplication produces a significant right tail. At high energies, the large number of photoelectrons and dynode stages causes the central limit theorem to drive distributions toward symmetric Gaussian form. Appendix C provides detailed derivation and statistical methodology.

A pooled estimation methodology addresses simulation statistical limitations. The PMT cascade simulation was executed five times for each event, generating five realizations of photoelectron output for identical input conditions. These five realizations were pooled for maximum likelihood estimation, effectively averaging stochastic fluctuations while preserving underlying resolution characteristics.

Energy Resolution Results

Figure 3.9 presents energy resolution as a function of deposited energy across the 1 MeV to 100 MeV range.

Maximum likelihood estimation yields:

$$\frac{\sigma}{E} = \frac{106.4\%}{\sqrt{E/\text{MeV}}} \quad (3.11)$$

with skewness parameters $\alpha_0 = 7.8$ and $E_{\text{decay}} = 11.3 \text{ MeV}$. The energy-dependent skewness analysis reveals that asymmetric effects become negligible above approximately 60 MeV, where distributions approach the symmetric Gaussian limit.

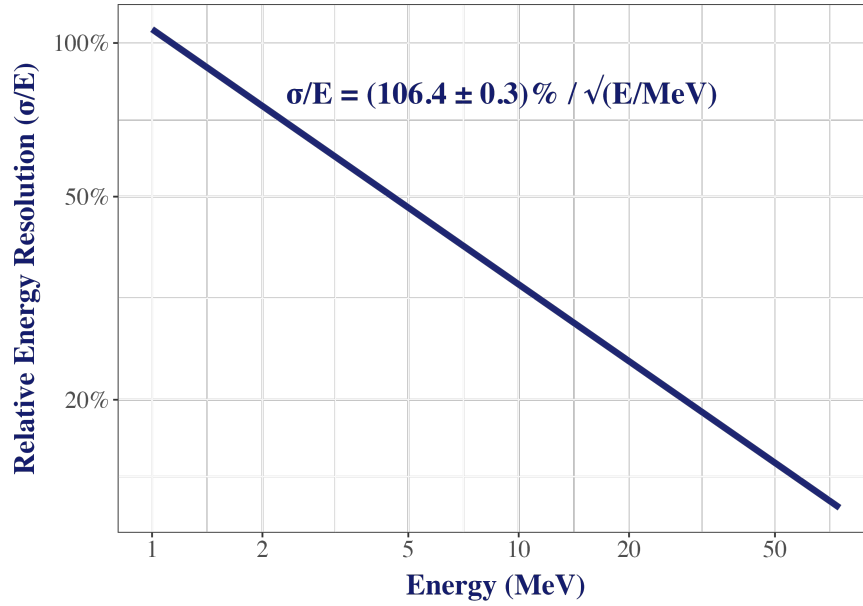


Figure 3.9: Energy resolution σ/E versus deposited energy for the simulated PbF_2 detector with PMT readout. The fitted model $\sigma/E = 106.4\%/\sqrt{E/\text{MeV}}$ (solid line) follows expected $1/\sqrt{E}$ scaling across the full energy range, confirming photon counting statistics dominance. The model demonstrates excellent agreement with empirical measurements from literature across nearly three orders of magnitude, validating both the simulation framework and the resolution characterization methodology.

The $1/\sqrt{E}$ scaling confirms that photon counting statistics dominate detector response. At 10 MeV—the target detection threshold—the model predicts $\sigma/E \approx 34\%$, indicating a standard deviation of approximately 3.4 MeV.

Validation Across Three Orders of Magnitude

The predicted resolution agrees quantitatively with previous experimental studies across vastly different energy regimes, validating both the simulation framework and statistical methodology. Table 3.1 summarizes validation spanning from 12 MeV to 1 GeV.

At low energies, the model predicts $\sigma/E = 30.7\%$ at 12 MeV, agreeing within 4% of Christmann’s beam-test measurements using identical PbF_2 crystals with PMT readout in the 10 MeV to 14 MeV range [25]. This validates the simulation precisely in the energy range most relevant for dark matter detection.

At high energies, the model predicts $\sigma/E = 3.4\%$ at 1 GeV, agreeing within 9% of Fienberg et al.’s measurements of 3.7% for 3×3 arrays of PbF_2 crystals with SiPM readout [37]. This agreement persists despite different readout technologies (PMTs

Table 3.1: *Energy resolution validation comparing simulation predictions with experimental measurements across nearly three orders of magnitude in energy. The predicted stochastic parameter $a = 106.4\%$ reproduces independent measurements within 10 % across the full energy range.*

Energy [MeV]	Predicted σ/E	Measured σ/E	Agreement
12	30.7 %	$\sim 32\%$ ^a	Within 4 %
1,000	3.4 %	$3.7 \pm 0.1\%$ ^b	Within 9 %

^aChristmann [25]: Identical PbF₂ crystals with PMT readout, 10 MeV to 14 MeV range.

^bFienberg et al. [37]: 3×3 PbF₂ arrays with SiPM readout.

vs. SiPMs) and detector configurations (single crystal vs. arrays), demonstrating that energy resolution is dominated by photon counting statistics rather than photodetector-specific effects.

Agreement across three orders of magnitude—from the 10 MeV threshold region to 1 GeV—establishes confidence in resolution predictions throughout the relevant energy range. The validation confirms that the simulation captures fundamental stochastic processes governing energy measurements.

Implications for Dark Matter Searches

Setting rigorous exclusion limits on dark matter production requires comparing predicted signal spectra against measured background distributions. This comparison demands that both spectra share a common energy scale—either true deposited energy or reconstructed energy including resolution effects. Two methods achieve this consistency.

The first method convolves theoretical predictions with detector response. Monte Carlo simulations generate dark matter interaction spectra in true energy, which must then be convolved with detector response to produce the reconstructed spectrum expected in data. This convolution can proceed through full simulation of the detection chain—Cherenkov production, optical transport, and photodetector response including dynode cascade—or more efficiently through a response matrix $R_{ij} = P(E_{\text{rec},i}|E_{\text{true},j})$ encoding the probability that true energy E_j reconstructs as measured energy E_i . The validated resolution model $\sigma/E = 106.4\%/\sqrt{E/\text{MeV}}$ determines this matrix structure.

The second method deconvolves measured spectra to recover true energy distributions. Statistical unfolding techniques can invert the response matrix to transform

reconstructed background measurements into true energy space for comparison with unconvolved signal predictions. This method proves more challenging because noise amplification is inherent in deconvolution, but it avoids requiring full detector simulation for each signal hypothesis.

Resolution matters because it modifies spectral shapes that distinguish signal from background. A dark matter signal producing electron recoils with a characteristic energy distribution becomes broadened after resolution convolution. Sharp spectral features that would clearly distinguish signal from smoothly-falling background become smeared, reducing discrimination power. At the target threshold of 10 MeV, the characterized resolution $\sigma \approx 3.4$ MeV causes events to migrate across energy bins, mixing signal and background in ways that must be accurately modeled to avoid systematic bias in exclusion limits.

Appendix A.8 develops the response matrix formalism in detail, demonstrating how the characterized resolution enters the statistical framework for dark matter exclusion limits. The four-step construction methodology using kernel density estimation produces matrix elements from Monte Carlo simulations, enabling either convolution of signal predictions or providing the basis for background deconvolution.

3.7 Particle Detection Efficiency

With the simulation framework validated, a crucial question can now be addressed: How efficiently does the detector identify particles across energy ranges typical of cosmic ray interactions? This efficiency depends not only on particle energy but also on the paths particles take through the detector volume. Quantifying this relationship enables accurate background predictions for dark matter searches.

A particle is considered detected when its simulated PMT output exceeds a configurable threshold expressed in photoelectrons (PE). Particles producing signals above threshold constitute detected events; those below remain undetected.

Empirical Efficiency Determination

The most direct method divides simulated events into discrete energy bins and calculates the fraction exceeding threshold in each bin. For a bin centered at energy E_i containing N_i events, of which n_i produce signals above threshold, the empirical

efficiency is:

$$\eta_{\text{emp}}(E_i) = \frac{n_i}{N_i} \quad (3.12)$$

with Wilson score confidence intervals accounting for binomial uncertainty. This method validates results without model assumptions but cannot produce continuous efficiency functions across the full energy spectrum. Without a parametric model, comparing detection performance between different detector configurations also becomes cumbersome.

Statistical Modeling with Generalized Additive Models

Generalized additive models (GAMs) overcome these limitations by learning smooth efficiency curves directly from data. For readers familiar with linear regression, a GAM extends that framework by replacing straight-line relationships with flexible curves. Instead of assuming:

$$\text{logit}[\eta] = \beta_0 + \beta_1 E + \beta_2 \ell \quad (3.13)$$

the GAM fits:

$$\text{logit}[\eta(E, \ell)] = s_1(\log_{10} E) + s_2(\ell) + s_{12}(\log_{10} E, \ell) \quad (3.14)$$

where s_1 , s_2 , and s_{12} are smooth functions (typically constructed from spline basis functions) that adapt their shape to the data, and ℓ represents path length through the detector. The logit transformation $\text{logit}(\eta) = \ln[\eta/(1 - \eta)]$ constrains predictions to physically meaningful probabilities between 0 and 1. The energy E is transformed logarithmically because cosmic ray spectra span multiple orders of magnitude. Modern software packages fit these models automatically with minimal tuning, using restricted maximum likelihood to prevent overfitting.

The GAM provides detection probability conditional on both energy *and* path length. However, experimental applications require marginal efficiency—the average detection probability at a given energy across all possible path lengths. This is calculated by integrating over the empirical path length distribution:

$$\eta(E) = \int \eta(E, \ell) p(\ell | E) d\ell \quad (3.15)$$

where $p(\ell | E)$ represents the distribution of path lengths observed at energy E in the simulation. In practice, this integral is evaluated through Monte Carlo sampling: for each target energy E_j , path lengths are drawn from events within an energy window $E_j(1 \pm 0.2)$, and the GAM is evaluated at each (E_j, ℓ_k) pair and averaged.

$$\eta_{\text{GAM}}(E_j) = \frac{1}{K} \sum_{k=1}^K \text{logit}^{-1} [s_1(\log_{10} E_j) + s_2(\ell_k) + s_{12}(\log_{10} E_j, \ell_k)] \quad (3.16)$$

with $K = 1,000$ samples. This procedure yields the expected detection probability for particles of energy E_j when accounting for all possible paths through the detector.

Validation and Results

Figure 3.10 presents detection efficiency results for gamma rays, neutrons, protons, and electrons at the minimum threshold of 1 PE. While realistic operation requires higher thresholds to overcome noise in the signal chain, the 1 PE analysis reveals fundamental detector response characteristics and establishes an upper bound on achievable efficiency.

The GAM predictions (solid lines) agree excellently with empirical validation points (circles), achieving correlation coefficients exceeding 0.99 for all particle types. This agreement validates the marginalization approach and confirms that the GAM accurately captures the underlying detection physics.

The 1 PE analysis reveals distinct detection signatures for each particle type:

Gamma rays exhibit monotonically increasing efficiency from 25 % at 1 MeV to 75 % at 300 MeV, reflecting how electromagnetic cascade development depends on photon energy.

Neutrons demonstrate consistently low efficiency (5 % to 10 %) across all energies, confirming the weak interaction cross-sections and limited neutron detection capability inherent to electromagnetic calorimeters documented in Section 3.3.

Protons show characteristic threshold behavior: near-zero efficiency below 200 MeV and rapid saturation above 300 MeV. This directly reflects the Cherenkov threshold for protons in PbF_2 at approximately 200 MeV. Below threshold, protons deposit energy through ionization but produce no detectable Cherenkov light; above threshold, they become visible.

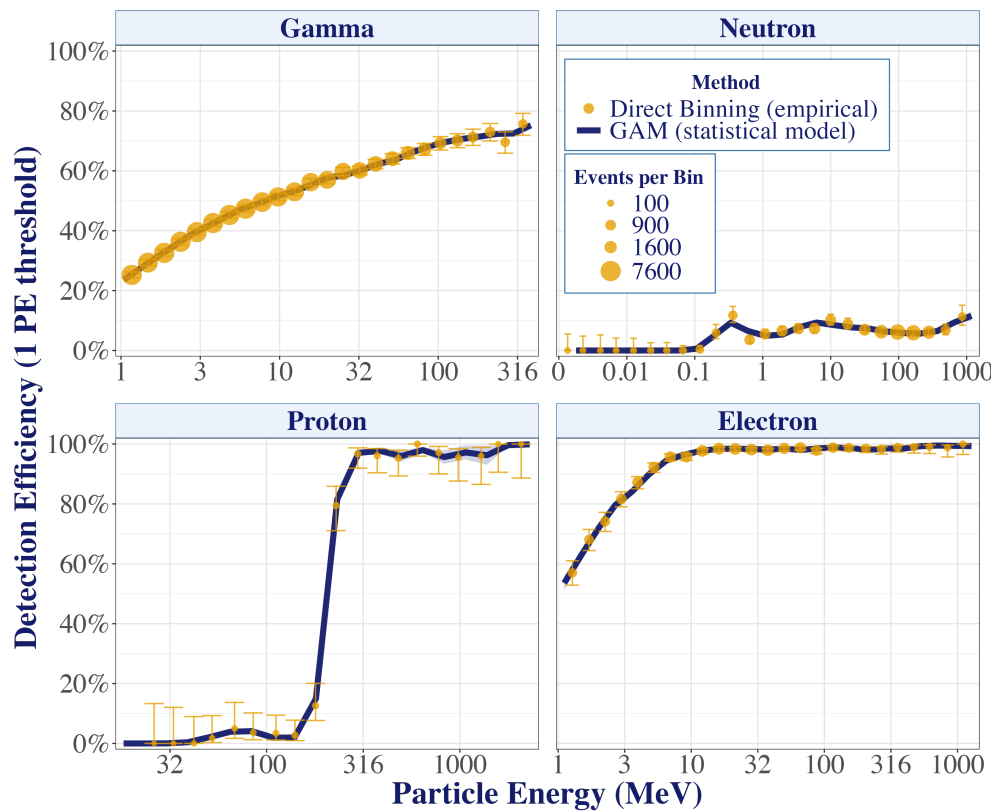


Figure 3.10: Detection efficiency comparison between GAM statistical model (solid lines) and direct empirical binning (circles) for four particle types at 1 PE threshold. The excellent agreement (correlation $r > 0.99$) validates the marginalization approach across the respective cosmic ray energy ranges. Error bars represent Wilson score confidence intervals. The 1 PE threshold provides detection capability insight but poses practical implementation challenges due to noise in the signal chain.

Electrons achieve high efficiency (90 % to 95 %) across typical cosmic ray energy ranges due to direct electromagnetic interactions and reliable Cherenkov light production.

While Figure 3.10 omits muons for clarity, cosmic muons exhibit detection efficiencies approaching 100 % across their naturally occurring energy range of 1 GeV to 100 GeV. This near-perfect detection identifies muons as the most significant background component requiring active suppression through the veto system.

Figure 3.11 extends this analysis across threshold values from 1 PE to 30 PE, providing a comprehensive view of detector response characteristics essential for optimizing background rejection strategies.

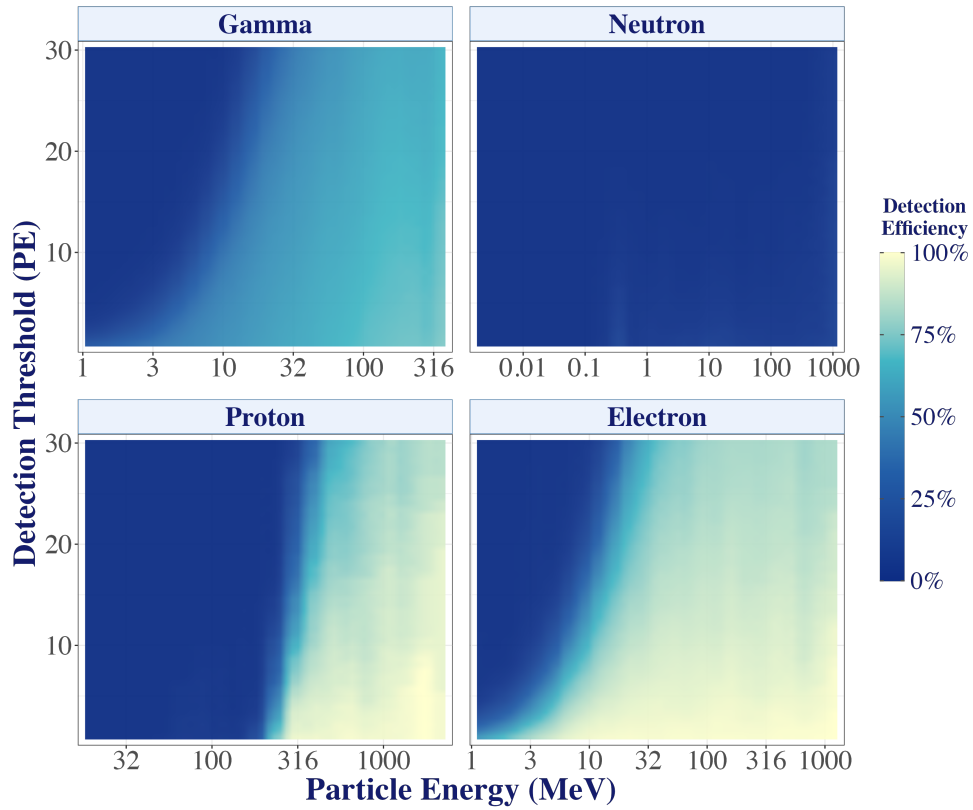


Figure 3.11: Complete detection efficiency response as a function of energy and detection threshold for four particle types. The color scale indicates efficiency from 0 % (light) to 100 % (dark blue). These efficiency maps characterize detector response to cosmic ray backgrounds and provide essential input for background suppression studies and veto system performance comparisons.

3.8 Signal Detection Efficiency at Target Threshold

Dark matter searches require efficient detection of low-energy electron recoils while rejecting backgrounds. The broad efficiency characterization in Section 3.7 established detector response to cosmic ray backgrounds across wide energy ranges, but a critical design question remains: How efficiently can the detector identify physics signals at 10 MeV, the target threshold for DarkMESA?

This efficiency was determined from events within the energy bin 9.5 MeV to 10.5 MeV in the cosmic ray simulation data. Because these events arise from primary particles with diverse entry points and trajectories through the detector volume, the resulting efficiency represents an average across all possible interaction geometries.

Figure 3.12 shows performance characteristics. The detector maintains detection efficiency exceeding 99 % for photoelectron thresholds below 7 PE, demonstrating that

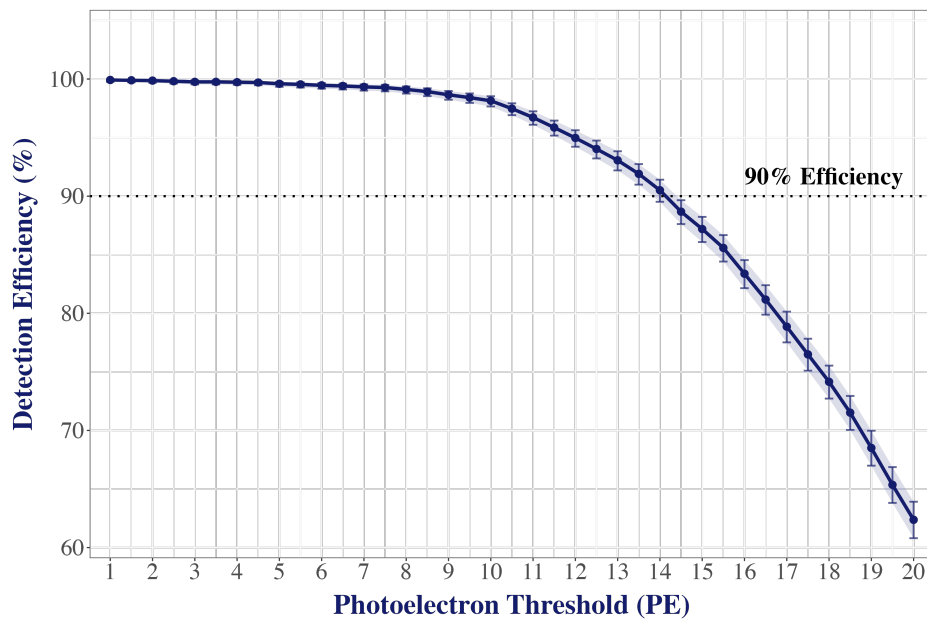


Figure 3.12: Detection efficiency at 10 MeV as a function of photoelectron threshold for single PMT readout. We determined the efficiency curve using cosmic ray data from the validated simulation framework within the energy bin 9.5 MeV to 10.5 MeV ($N=7,419$ events) with Wilson 95 % confidence intervals accounting for binomial statistics.

the validated light yield of 1.94 PE/MeV provides adequate signal detection capability. At 10 MeV, events produce mean photoelectron signals of approximately 19 PE, establishing comfortable margin above thresholds where near-perfect efficiency is maintained.

The 90 % efficiency level—representing acceptable signal loss for rare event searches where background rejection may justify modest efficiency sacrifice—occurs at approximately 14 PE threshold. This substantial margin (14 PE threshold versus 19 PE mean signal) accommodates the statistical and systematic variations inherent to Cherenkov light detection.

The smooth efficiency rolloff—from 99 % at 7 PE through 90 % at 14 PE to 62 % at 20 PE—reflects the convolution of these multiple stochastic processes. Operating at 14 PE threshold provides sufficient safety margin that most 10 MeV events remain above threshold despite photoelectron fluctuations and dynode cascade variations.

The central finding: the PbF_2 Cherenkov calorimeter produces sufficient light for highly efficient signal detection at the target energy threshold. Photoelectron statistics and light collection efficiency do not fundamentally limit 10 MeV threshold achievement.

The validated light yield of 1.94 PE/MeV proves adequate for DarkMESA applications, confirming that the detector concept is viable from a signal detection perspective.

However, adequate signal efficiency addresses only half the implementation challenge. Operating within the identified threshold range (7 PE to 14 PE) requires photodetector systems capable of reliable operation while suppressing spurious triggers that could compromise background discrimination. Chapter 4 evaluates photodetector readout strategies that preserve the signal detection efficiency demonstrated here while addressing these operational challenges.

3.9 Summary and Implications

This chapter establishes the foundation for DarkMESA detector implementation by comprehensively characterizing and rigorously validating the simulation framework. A GEANT4-based simulation framework was developed incorporating detailed dynode cascade modeling to reproduce photomultiplier tube response characteristics. Validation through cosmic ray measurements demonstrates excellent agreement between predicted and measured energy deposition spectra and photoelectron yield distributions, establishing confidence in the model's predictive capabilities. The validated framework predicts energy resolution of $\sigma/E = 106.4\%/\sqrt{E/\text{MeV}}$, matching independent measurements across three orders of magnitude in energy.

The efficiency analysis confirms high sensitivity to electrons and muons (90 % to 100 %) and low neutron response (5 % to 10 %)—essential for background rejection. Comprehensive efficiency maps spanning photoelectron thresholds from 1 PE to 30 PE provide complete response surfaces for each particle type. Shielding studies show that passive attenuation alone cannot eliminate backgrounds to required levels: while concrete and lead substantially reduce electromagnetic and neutron backgrounds, residual rates necessitate active discrimination techniques.

The signal efficiency analysis at the target 10 MeV threshold delivers the chapter's central result: PbF_2 produces sufficient light for highly efficient dark matter detection. The detector maintains 99 % efficiency at photoelectron thresholds up to 7 PE, with 90 % efficiency sustained to 14 PE. The mean signal of 19.4 PE at target energy provides comfortable operational margins against statistical fluctuations. Photoelectron statistics and light collection do not fundamentally limit threshold achievement—the detector concept is viable from a signal detection perspective.

With signal capability established, the remaining implementation challenge concerns practical photodetector operation: can these thresholds be maintained while suppressing spurious triggers that would compromise background discrimination? Chapter 4 addresses this question by characterizing photodetector noise properties and evaluating alternative readout configurations—dual-sided PMT arrangements, coupled crystal arrays, and silicon photomultiplier systems—that preserve signal efficiency while enhancing background rejection. The validated simulation framework enables quantitative assessment of each configuration, guiding optimal implementation strategies for Phase A deployment and Phase B extension.

Chapter 4

Photodetector Optimization Studies

At the 10 MeV detection threshold required for rare event searches, photomultiplier tube dark counts generate spurious triggers that overwhelm the few-events-per-year dark matter signals the DarkMESA calorimeter must detect. Although PbF_2 crystals provide sufficient light yield for low-threshold calorimetry (Section 3.5), dark count rates of 0.5 Hz to 4 Hz at practical detection thresholds remain problematic. Single-PMT readout proves incompatible with rare event detection. This finding, which emerged during characterization measurements, motivates the search for alternative photodetector strategies.

Coincidence-based readout exploits a fundamental distinction between signal and noise: genuine physics events produce correlated signals across multiple photodetectors, while uncorrelated dark counts arise independently in individual devices. Requiring simultaneous detection in multiple photodetectors suppresses accidental coincidences while maintaining sensitivity to real energy depositions. The rate of accidental coincidences scales as [38]:

$$\nu_{\text{accidental}} \approx n \tau^{n-1} \nu_1 \nu_2 \cdots \nu_n \quad (4.1)$$

where n denotes the number of detectors in coincidence, τ represents the coincidence resolving time, and ν_i characterizes the count rate for detector i . Higher coincidence requirements ($n = 3, 4, 5$) provide exponentially stronger background suppression through the τ^{n-1} scaling.

However, higher-order coincidence demands increasingly stringent detection efficiency: each additional requirement reduces the probability that genuine low-energy events produce sufficient photoelectrons across all participating detectors. Viable

configurations must achieve orders-of-magnitude spurious trigger suppression while maintaining the >90 % signal detection efficiency established in Section 3.8. At 10 MeV, where photoelectron yields remain modest, this dual constraint is demanding.

This chapter systematically evaluates photodetector readout strategies under realistic experimental constraints, progressing from the validated single-PMT baseline through increasingly sophisticated multi-detector approaches: dual-sided PMT readout, coupled crystal arrays, and silicon photomultiplier arrays.

4.1 Spurious Events: PMT Dark Counts

Reading out each crystal with a single PMT offers compelling simplicity: straightforward mechanical coupling via optical grease interface, linear scaling of electronics complexity with detector count, well-characterized response that simplifies calibration, and proven long-term stability from previous experiments. The experimental validation (Section 3.5) employed this architecture with individual PbF_2 crystals coupled to Photonis XP2900 photomultiplier tubes, establishing fundamental detection capabilities. However, systematic dark count characterization reveals a problem: spurious trigger rates at practical detection thresholds preclude rare event searches.

Figure 4.1 presents dark count rate measurements as functions of photoelectron threshold for available PMT resources: approximately 1,000 Photonis XP2900 tubes from the A4 experiment inventory and newly procured Hamamatsu R1924A tubes borrowed from MAGIX for comparative evaluation.

Converting charge to photoelectrons using PMT gain values established the photoelectron threshold scale. For the Hamamatsu R1924A device, the manufacturer-specified gain of $G \approx 7.5 \times 10^5$ at 875 V operating voltage was used. For the Photonis XP2900 tubes, gains were determined through laser-based single-photoelectron measurements using the methodology detailed in Appendix B, yielding $G \approx 3.0 \times 10^6$ at 1,700 V—consistent with the gain used in dynode cascade validation (Section 3.5).

The critical problem appears at the 14 PE threshold corresponding to 90 % signal detection efficiency at 10 MeV (Section 3.8). At this threshold, dark count rates range from 0.5 Hz to 1 Hz even for the best-performing tubes. For rare event searches expecting only a few signal events per year, this spurious trigger accumulation represents catastrophic background contamination.

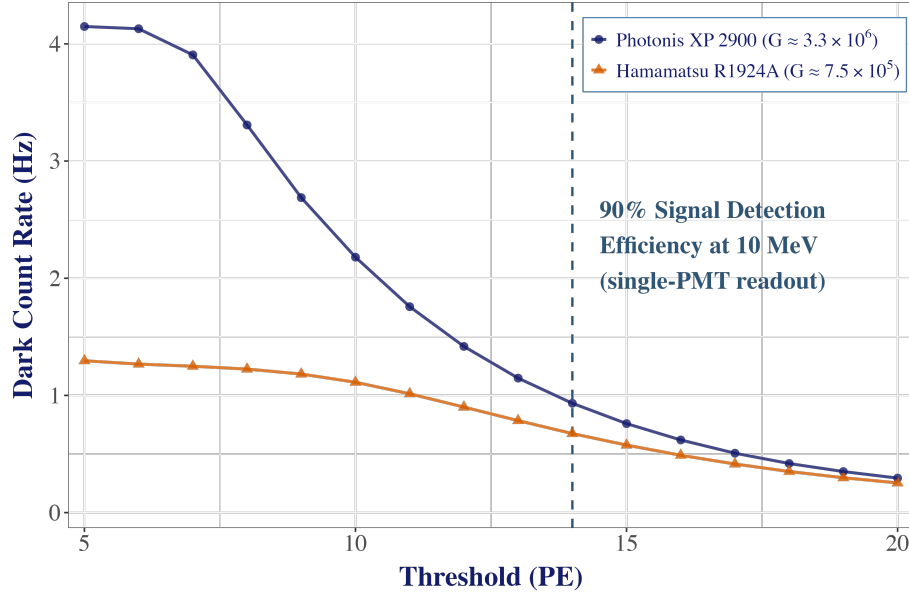


Figure 4.1: Dark count rate as a function of photoelectron threshold for photomultiplier tubes under evaluation. A Photonis XP2900 tube from the A4 experiment inventory (operated at 1,700 V) compared with a Hamamatsu R1924A tube borrowed from MAGIX (operated at 875 V). The photoelectron threshold scale derives from converting measured charge distributions using PMT gain values. The vertical dashed line indicates the 14 PE threshold corresponding to 90 % signal detection efficiency at 10 MeV for single-PMT readout (Section 3.8). Dark count rates of 0.5 Hz to 1 Hz persist at this threshold.

Even aggressive threshold increases provide insufficient suppression. Operating at 20 PE threshold—corresponding to approximately 60 % signal efficiency and representing unacceptable signal loss—still yields dark count rates of 0.2 Hz to 0.4 Hz. Single-PMT readout cannot achieve the background suppression required for rare event detection because no threshold simultaneously provides acceptable signal efficiency and tolerable spurious trigger rates.

These results demonstrate that coincidence-based discrimination through multi-detector architectures represents the only viable path forward. Genuine physics events produce correlated signals across multiple photodetectors, while uncorrelated PMT dark counts arise independently in individual tubes. The following sections systematically evaluate configurations exploiting this distinction: dual-sided PMT readout (Section 4.2), coupled crystal arrays (Section 4.3), and silicon photomultiplier arrays (Section 4.4).

4.2 Phase A Solution: Dual-Sided Readout

Dual-sided crystal readout suppresses PMT dark count backgrounds through twofold coincidence logic. Genuine physics events deposit energy throughout the crystal volume, producing correlated signals at both photodetectors. Uncorrelated PMT dark counts arise independently in individual tubes. Requiring simultaneous detection in both PMTs rejects single-tube noise while maintaining sensitivity to real energy depositions. Figure 4.2 illustrates this readout geometry, where photomultiplier tubes couple to both faces of each PbF_2 crystal via optical grease interfaces.

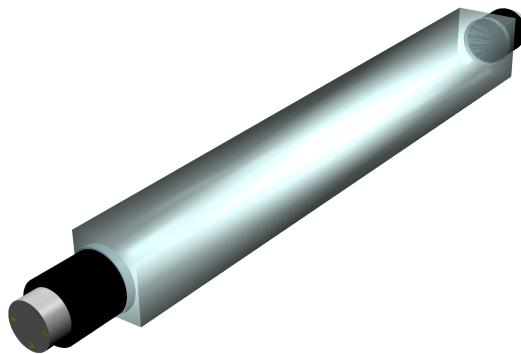


Figure 4.2: Dual-sided readout configuration for Phase A implementation. Each PbF_2 crystal couples to photomultiplier tubes on both faces via optical grease interfaces.

Phase A deployment benefits from straightforward mechanical implementation: standard optical coupling procedures, independent PMT operation with dedicated high-voltage supplies, and well-characterized readout electronics. Each photodetector maintains its own signal processing chain, preserving the operational simplicity that facilitated successful cosmic ray validation studies.

Detection Efficiency Considerations

Positioning photodetectors at opposite crystal faces splits the Cherenkov photon distribution, reducing per-PMT light yield compared to single-sided readout. Single-PMT configurations achieve approximately 1.9 PE/MeV (Section 3.5) through direct optical coupling with minimal losses. Dual-sided architecture yields approximately 1.2 PE/MeV per PMT.

Dual-sided detection efficiency depends on the probability that photoelectron yields in both detectors simultaneously exceed threshold despite this reduced per-PMT

collection. For symmetric threshold operation—where both PMTs employ identical discriminator settings—the coincidence requirement demands adequate light collection at each photodetector.

Figure 4.3 presents detection efficiency at 10 MeV as a function of photoelectron threshold for dual-sided readout (requiring both PMTs above threshold), compared with single-PMT performance for reference.

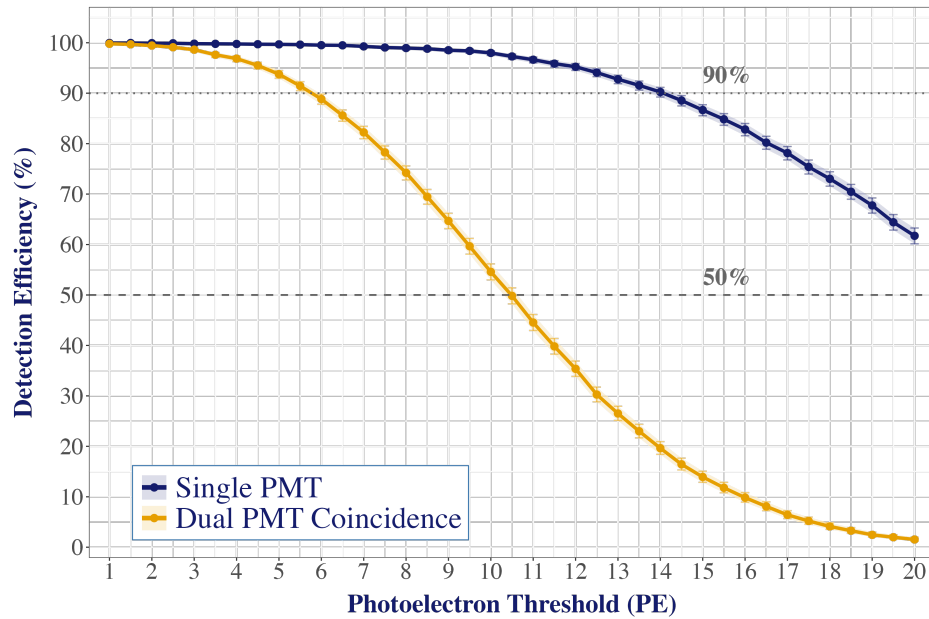


Figure 4.3: Detection efficiency at 10 MeV versus photoelectron threshold comparing single-PMT readout (blue) with dual-sided readout requiring both PMTs above threshold (orange). Error bars represent Wilson 95 % confidence intervals. Horizontal dashed lines indicate 90 % and 50 % efficiency levels.

Dual-sided readout maintains 90 % detection efficiency up to approximately 5 PE to 6 PE threshold per PMT. This threshold range provides adequate operational margin above the noise floor while enabling effective dark count suppression through coincidence logic. Applying identical discriminator settings to both PMTs simplifies calibration and ensures operational uniformity across detector channels.

Accidental Coincidence Background Estimates

Dual-sided readout viability depends on accidental coincidence rates at the operating thresholds that signal efficiency requirements establish.

Coincidence timing windows determine how effectively accidental rates are suppressed. Experimental measurements of time differences between coincident signals in dual-

sided readout yielded $\sigma \approx 1.5$ ns for genuine physics events. A 10 ns coincidence resolving time provides greater than 5σ safety margin, ensuring high acceptance for real coincident signals while minimizing the temporal window for accidental overlaps.

For twofold coincidence between two detectors, Equation 4.1 simplifies to:

$$\nu_{\text{accidental}} = 2\tau\nu_1\nu_2 = 2\tau\nu^2 \quad (4.2)$$

where the factor of 2 accounts for the single possible detector pairing, and the approximation $\nu_1 = \nu_2 = \nu$ assumes identical PMT performance. Table 4.1 presents calculated accidental rates per crystal pair across the realistic dark count range using the 10 ns coincidence window.

Table 4.1: *Accidental coincidence rates per crystal pair (events/year) for dual-sided PMT readout with 10 ns coincidence resolving time across realistic dark count rates at 5 PE to 6 PE threshold. Values rounded to order of magnitude. Multiply by detector count to determine aggregate backgrounds.*

Timing Window	Dark Count Rate per PMT			
	1 Hz	2 Hz	3 Hz	4 Hz
τ				
10 ns	~1	~3	~6	~10

Accidental coincidence rates depend strongly on PMT dark counts. Best-performing PMTs (Hamamatsu R1924A at 1 Hz) produce approximately 1 accidental event per crystal per year—negligible. For A4 PMTs with elevated dark counts, accidental rates per crystal range from approximately 3 to 10 events per year, remaining manageable at the single-crystal level.

However, these per-crystal rates must be scaled by total detector channel count to assess aggregate background contributions. The following subsections evaluate dual-sided readout feasibility for Phase A and Phase B implementations, accounting for substantial differences in detector scale and corresponding accidental background accumulation.

Phase A Implementation Feasibility

For the Phase A calorimeter comprising approximately 25 PbF₂ crystals, dual-sided readout represents a technically and economically viable solution. The implementation

requires 50 PMTs total—manageable both in terms of procurement and operational complexity.

Two photodetector options present distinct trade-offs:

Hamamatsu R1924A tubes offer compact geometry (approximately 80 mm total length including base) that minimizes dead space between crystal array and veto system, low-voltage operation (15 V input) that eliminates high-voltage supply infrastructure, and superior dark count performance (1 Hz at operating threshold) that yields minimal accidental backgrounds. Unit costs typically reach order 1 k€, resulting in total photodetector costs of order 50 k€ for the required 50-tube configuration.

Photonis XP2900 tubes from the available A4 inventory (approximately 1,000 units) eliminate photodetector procurement costs. However, these tubes introduce operational complications: extended physical dimensions (approximately 200 mm total length) increase the overall detector envelope beyond the existing veto system geometry, and elevated dark count rates (1 Hz to 4 Hz) produce increased accidental backgrounds.

For Phase A implementation, dual-sided readout using Hamamatsu R1924A PMTs is recommended based on combined performance, compact geometry, and acceptable cost for a 25-crystal demonstration system. The accumulated accidental background of approximately 25 events per year (assuming 1 Hz dark counts per PMT) remains negligible compared to cosmic ray backgrounds requiring active veto suppression.

Phase B Scaling Limitations

Scaling to Phase B reveals limitations of dual-sided PMT readout. The 2,000-crystal calorimeter (1,000 PbF₂ plus 1,000 lead glass) requires 4,000 PMTs for dual-sided readout, with accidental backgrounds accumulating linearly: even optimal Hamamatsu performance yields approximately 1,000 accidental events per year—three orders of magnitude above anticipated dark matter signal rates.

Substantial photodetector procurement costs (several million euros), the 120 mm additional envelope from PMT dimensions necessitating enlarged veto systems, and infrastructure complexity from mixed A4/Hamamatsu deployment compound these background limitations. These challenges motivate alternative configurations achieving either intrinsically lower noise or higher-order coincidence for stronger background suppression.

The photodetector optimization studies in subsequent sections focus exclusively on the PbF_2 front wall. This restricted scope reflects unresolved technical factors: crystal geometry, readout architecture, and mechanical integration for the lead glass detector configuration remain under evaluation (Section 2.5). While Christmann's beam test program characterized lead glass response using short crystal samples (150 mm length), the light collection efficiency and photoelectron yield at the 10 MeV threshold for the full-length SAPHIR modules proposed for Phase B (460 mm) have not yet been experimentally determined. Substantial optical absorption over extended path lengths may significantly reduce photoelectron yields compared to short-sample measurements, leaving detection efficiency and threshold requirements undetermined for Phase B lead glass configurations.

The PbF_2 analysis uses 185.4 mm crystals as the baseline configuration. This conservative choice establishes worst-case light collection conditions: longer optical paths increase photon absorption, reducing photoelectron yield. Shorter crystals (150.0 mm minimum) exhibit improved light collection through reduced absorption, ensuring that configurations demonstrating adequate performance with 185.4 mm crystals remain valid across the full inventory range.

4.3 Alternative Approach: Coupled Crystal Arrays

The Phase B scaling challenges motivate searching for photodetector configurations that reduce channel count while maintaining detection performance. Optically coupling multiple crystals into unified detector modules promises substantial advantages: light sharing across several crystals reduces photodetector count compared to dual-sided readout of individual crystals, and higher-order coincidence requirements (threefold, fourfold) provide enhanced spurious trigger suppression beyond the twofold coincidence achievable with dual-sided PMT readout.

In a coupled crystal array, four PbF_2 crystals are optically bonded using EJ-500 optical cement, with a single PMT attached to each crystal (Figure 4.4). This architecture cuts photodetector count in half compared to dual-sided readout: four coupled crystals require only four PMTs total versus eight PMTs for individual dual-sided operation. For Phase B implementations with 1,000 crystals, this factor-of-two reduction translates to substantial cost savings and simplified infrastructure. More importantly, higher-order coincidence (threefold $\tau^2 \nu^3$, fourfold $\tau^3 \nu^4$) provides orders-of-magnitude background

rejection beyond the $\tau\nu^2$ scaling of twofold coincidence (Section 4.2), potentially enabling Phase B operation even with PMT dark count rates that prove problematic for dual-sided configurations.

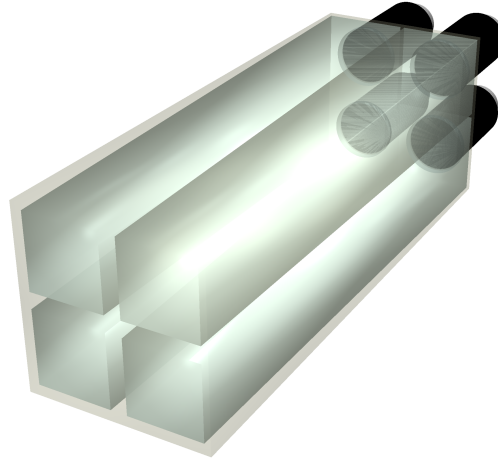


Figure 4.4: Coupled crystal array configuration. Four PbF_2 crystals optically bonded with EJ-500 cement, each with one PMT attached. Optical coupling creates extended light paths compared to individual crystal readout.

However, this promising architecture is not viable for 10 MeV threshold operation. Coupled arrays fail through two compounded effects: severely reduced light yield per photodetector and extreme geometric heterogeneity in light distribution. These failures establish critical requirements for the alternative approaches explored in subsequent sections.

Light Yield Limitations and Distribution Heterogeneity

Coupled crystal architecture dramatically reduces light collection efficiency compared to direct readout geometries. Dual-sided readout of individual crystals achieves approximately 1.2 PE/MeV per PMT (Section 4.2), while coupled arrays yield only approximately 0.5 PE/MeV per PMT after complete dynode cascade simulation. This reduction reflects light sharing across the coupled volume: Cherenkov photons traverse optical cement interfaces with associated losses and distribute among four PMTs rather than concentrating at a single detector. At the 10 MeV target threshold, this reduced light yield produces mean photoelectron signals of only approximately 5 PE per PMT—dramatically lower than the 19 PE achieved with single-PMT readout or 12 PE with dual-sided configurations.

Beyond absolute light yield reduction, the coupled array architecture produces severe heterogeneity in light distribution across photodetectors. Events occurring in different crystals within the coupled volume generate different photoelectron patterns despite depositing identical total energy. This position-dependent light collection is quantified through the coefficient of variation (CV) of photoelectron yields across detectors for events within a fixed energy bin:

$$CV_{\text{event}} = \frac{\sigma_{\text{detectors}}}{\mu_{\text{detectors}}} \quad (4.3)$$

where $\sigma_{\text{detectors}}$ represents the standard deviation across detectors for individual events and $\mu_{\text{detectors}}$ denotes the mean photoelectron yield per detector. The reported CV represents the average of this metric across all events within the 9.5 MeV to 10.5 MeV energy bin.

For photoelectron counting subject to Poisson statistics alone, the CV has an inherent lower bound:

$$CV_{\text{Poisson}} = \frac{1}{\sqrt{N}} \quad (4.4)$$

where N represents mean photoelectron yield. This Poisson limit—approximately 45 % for 5 PE signals characteristic of coupled arrays at 10 MeV—establishes the minimum variation achievable even with perfectly uniform geometric light collection. Deviations above this limit indicate systematic position-dependent effects exceeding statistical fluctuations.

The coupled crystal array exhibits mean $CV \approx 93\%$ —more than double the Poisson statistical limit. This excess variation reflects the geometric origin of light sharing: events in Crystal 1 produce strong signals in the PMT optically coupled to Crystal 1 but weak signals in PMTs viewing Crystals 2–4, and vice versa. Position-dependent light collection introduces systematic heterogeneity exceeding intrinsic Poisson fluctuations.

Coupled crystal arrays thus suffer compounded limitations: reduced absolute light yield (0.5 PE/MeV versus 1.2 PE/MeV for dual-sided readout) and extreme geometric heterogeneity ($CV = 93\%$ versus 45 % Poisson limit). The configuration fails not from insufficient total light production but from inability to distribute that light uniformly across participating photodetectors. Can adequate signal detection efficiency be maintained at practical thresholds while enabling the higher-order coincidence that motivated this approach?

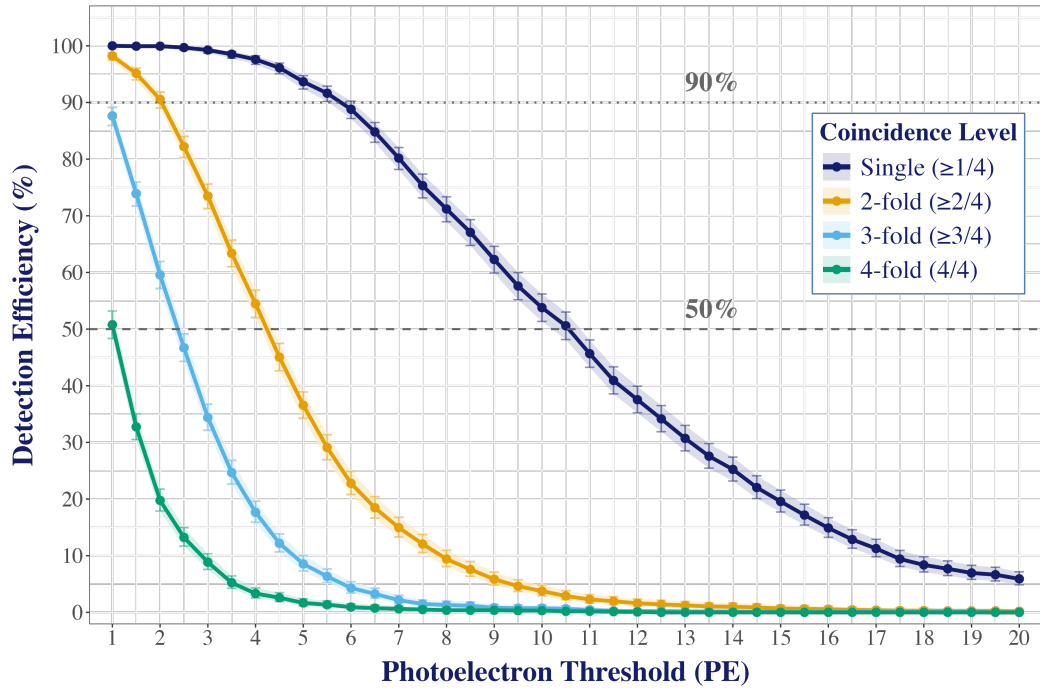


Figure 4.5: Detection efficiency at 10 MeV for coupled crystal array configuration across coincidence requirements. Efficiency curves show single-PMT, twofold, threefold, and fourfold coincidence as functions of uniform photoelectron threshold applied to all participating PMTs. Horizontal lines mark 90 % and 50 % efficiency levels. Error bars represent Wilson 95 % confidence intervals.

Detection Efficiency Analysis

Figure 4.5 presents detection efficiency at 10 MeV versus photoelectron threshold for different coincidence requirements in the coupled array configuration, determined using the validated simulation framework.

The results show that the reduced and heterogeneous light yield is incompatible with practical operational requirements. Single-PMT operation (requiring any one of four PMTs above threshold) maintains 90 % efficiency up to approximately 5 PE threshold, while twofold coincidence (requiring any two of four PMTs) achieves 90 % efficiency only up to approximately 2 PE. These configurations provide no background suppression comparable to dual-sided readout, negating the motivation for coupled array complexity.

Threefold coincidence (requiring three of four PMTs above threshold) maintains 90 % efficiency only for thresholds at or below 1 PE. Operating at 1 PE threshold approaches the noise floor of PMT systems, where single-photoelectron discrimination becomes unreliable and electronic noise contributions become significant.

Fourfold coincidence (requiring all four PMTs above threshold) fails to achieve 90 % efficiency at any practical threshold, reaching maximum efficiency of only approximately 50 % even at 1 PE. This fundamental limitation reflects the geometric heterogeneity: events occurring near any crystal corner produce strong signals in one PMT but weak signals in the three distant PMTs, precluding simultaneous threshold exceedance.

Coupled Array Conclusions

The coupled crystal array configuration fails to meet minimum requirements for Phase B implementation. The combination of reduced light yield (0.5 PE/MeV) and severe geometric heterogeneity ($CV = 93\%$) prevents operation at practical thresholds while maintaining adequate detection efficiency.

Viable solutions must:

- Enable higher-order coincidence through multiple photodetectors per crystal without light-sharing losses
- Reduce channel count through compact photodetector geometries
- Operate reliably at photoelectron thresholds 3 PE to 5 PE above the noise floor

These requirements favor photodetector technologies offering compact form factors that enable deployment of multiple devices per crystal face while maintaining direct optical coupling. Rather than sharing light across an extended coupled volume where position determines distribution, viable configurations must achieve distributed multi-detector coverage of individual crystal exit faces—ensuring all photodetectors view the same Cherenkov emission volume and receive comparable light regardless of event position within the crystal.

4.4 Silicon Photomultiplier Arrays

Individual SiPM devices with 3 mm to 6 mm active areas are small enough to tile complete crystal faces with multiple independent photodetectors. A 30 mm $\times$ 30 mm crystal face accommodates arrays of multiple SiPMs, providing access to higher-order coincidence requirements (fourfold, fivefold, or beyond) that dramatically suppress spurious triggers compared to dual-sided PMT readout. This architecture presents competing trade-offs: compact device dimensions reduce active area per photodetector compared to conventional PMTs, potentially compromising per-device light collection efficiency, while silicon photomultipliers exhibit substantially higher dark count rates than PMTs due to thermal generation in semiconductor junctions. These elevated dark count rates typically necessitate active cooling for viable operation.

Light Yield and Distribution Characteristics

The evaluated configuration arranges five onsemi J-Series silicon photomultipliers (6 mm $\times$ 6 mm active area) on a single 30 mm $\times$ 30 mm PbF₂ crystal face (Figure 4.6). Each SiPM couples directly to the crystal via optical interface, enabling independent photoelectron detection. This single-crystal geometry enables twofold through fivefold coincidence while ensuring all photodetectors view the same Cherenkov emission volume.

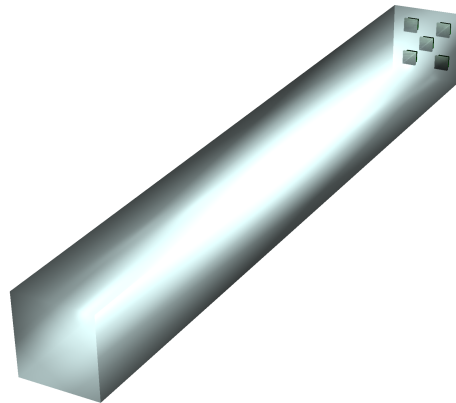


Figure 4.6: SiPM array readout configuration. Five compact silicon photomultipliers (6 mm $\times$ 6 mm) arranged on the 30 mm $\times$ 30 mm crystal face provide distributed light collection across the exit surface.

Silicon photomultiplier physics differs substantially from conventional PMT operation. Detection efficiency depends not only on wavelength-dependent quantum efficiency, but also Geiger discharge triggering probability (overvoltage-dependent), and geometric fill factor. Charge resolution reflects complex semiconductor effects including pixel-to-pixel gain variations, avalanche statistics, and optical crosstalk. The GEANT4 simulation framework incorporated manufacturer-specified photon detection efficiency and experimentally-determined charge resolution parameters extracted from spectroscopic analysis of dark count measurements. This hybrid approach combines GEANT4 optical transport with empirical charge resolution functions, enabling realistic modeling without detailed semiconductor device simulation. Appendix D provides detailed methodology.

Simulated light yield per individual SiPM reaches approximately 0.6 PE/MeV, producing mean signals of approximately 6 PE per device at the target 10 MeV energy. This per-device light collection compares closely to the 0.5 PE/MeV achieved with coupled crystal arrays (Section 4.3)—both configurations exhibit modest photoelectron yields representing substantial reductions from the 1.2 PE/MeV characteristic of direct dual-sided PMT readout.

The critical distinction between SiPM arrays and coupled crystal configurations lies not in total light collection but in distribution uniformity. Both achieve similar aggregate light yields, yet the single-crystal SiPM geometry eliminates the position-dependent systematic variations that plagued coupled arrays. For the five-SiPM array, measured CV reaches approximately 42 %—remarkably close to the 41 % Poisson limit for 6 PE signals. This near-equality demonstrates essentially ideal geometric light collection uniformity. The observed variations arise purely from statistical Poisson fluctuations in photoelectron generation, with negligible contribution from systematic position-dependent effects.

Efficiency Results and Operational Constraints

Figure 4.7 presents detection efficiency at 10 MeV versus photoelectron threshold for the five-SiPM array configuration, evaluated across coincidence requirements from single-SiPM detection through fivefold coincidence. The efficiency curves employ symmetric threshold operation, where all participating SiPMs must exceed identical photoelectron thresholds. This conservative approach simplifies operational uniformity across thousands of detector channels in Phase B implementations.

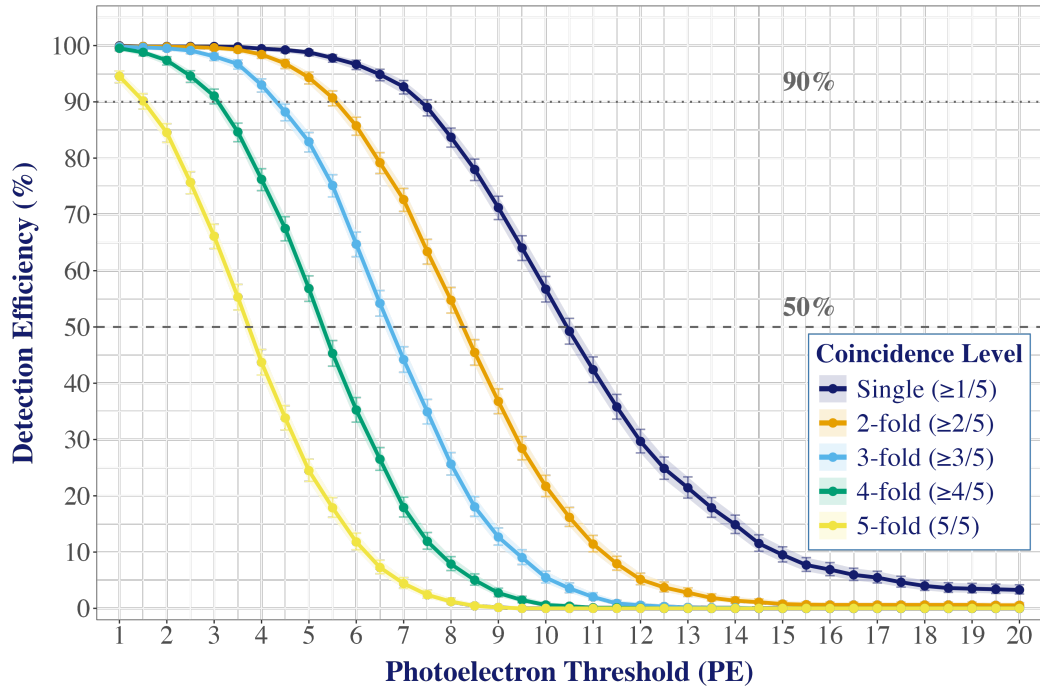


Figure 4.7: Detection efficiency at 10 MeV for five-SiPM array configurations across coincidence requirements. Error bars represent Wilson 95 % confidence intervals. Analysis employed room-temperature SiPM response parameters. With light yield of approximately 0.6 PE/MeV per SiPM, 10 MeV events produce mean signals of approximately 6 PE per device.

Table 4.2 summarizes maximum photoelectron thresholds maintaining 90 % detection efficiency for each coincidence level. Reported thresholds represent the smallest integer photoelectron value at or above the 90 % efficiency crossing point.

Phase B viability depends on two competing factors: threshold tolerance (higher thresholds produce lower dark count rates) and coincidence suppression effectiveness (higher-order coincidence provides stronger spurious trigger rejection). SiPM dark count rates depend strongly on both photoelectron threshold and operating temperature, declining exponentially as threshold increases and temperature decreases.

Dark Count Rate Characteristics

SiPM dark count rates depend strongly on both temperature and threshold, representing the primary limitation for low-energy event detection. [REDACTED] of the DarkMESA collaboration performed systematic measurements of these dependencies using onsemi J-Series devices (6 mm × 6 mm active area). The experimental data are

Table 4.2: Maximum photoelectron threshold maintaining 90 % detection efficiency at 10 MeV for different coincidence requirements in five-SiPM array configuration.

Coincidence Level	Maximum Threshold* for 90 % Efficiency
Single	7 PE
Twofold	5 PE
Threefold	4 PE
Fourfold	3 PE
Fivefold	1 PE

*Integer threshold values at or above 90 % efficiency crossing point.

used with permission; comprehensive measurement methodology will appear in the forthcoming thesis by [REDACTED].

The characterization employed temperature-controlled measurements from -25°C to 25°C in 5°C increments, with devices stabilized to 0.1°C precision. Figure 4.8 demonstrates the differential and integral dark count rate analysis at -10°C , chosen as representative operational temperature. The differential spectrum clearly resolves individual photoelectron peaks separated by distinct valleys, with peak separation becoming more pronounced at lower temperatures.

The differential spectrum was fitted using a generalized Poisson model accounting for optical crosstalk—the phenomenon where photons emitted during avalanche breakdown trigger adjacent cells, producing correlated multi-cell events. The photoelectron multiplicity distribution follows [39]:

$$P(k|\mu, \lambda) = \frac{\mu \cdot (\mu + k \cdot \lambda)^{k-1} \cdot e^{-(\mu+k\cdot\lambda)}}{k!} \quad (4.5)$$

where μ represents the mean number of primary dark events and λ characterizes the optical crosstalk probability. Each photoelectron peak is convolved with Gaussian charge resolution accounting for pixel-to-pixel gain variations:

$$\sigma_k = \sqrt{\sigma_0^2 + k \cdot \sigma_1^2} \quad (4.6)$$

where σ_0 represents electronic noise and σ_1 characterizes single-photoelectron peak width. The fit achieves excellent agreement ($R^2 = 0.991$) with crosstalk parameter $\lambda = 6.6\%$, typical for this SiPM class. Photoelectron thresholds are defined at positions $Q_{n,\text{PE}} = Q_0 + (n + 0.5) \cdot G$ to maximize discrimination between adjacent peaks.

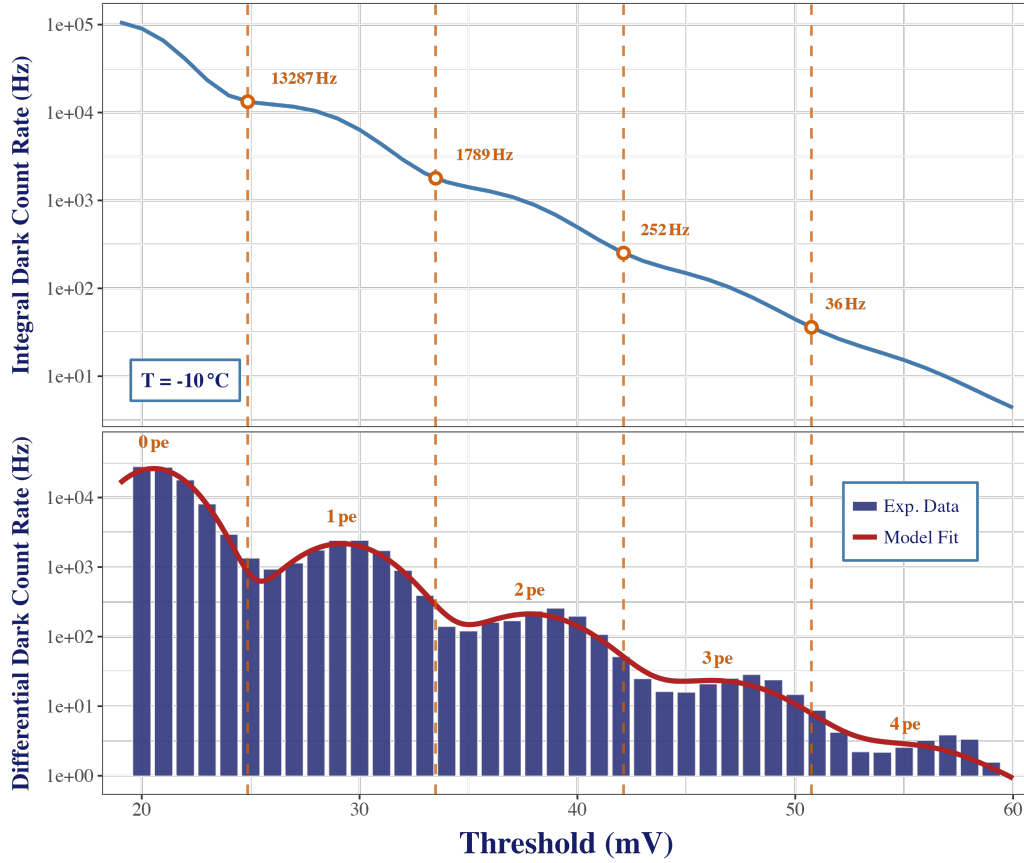


Figure 4.8: SiPM dark count spectroscopic analysis at $-10\text{ }^{\circ}\text{C}$. Upper panel: Integral dark count rates showing specific threshold values corresponding to photoelectron levels defined as $Q_{n,PE} = Q_0 + (n + 0.5) \cdot G$, where Q_0 represents pedestal position and G denotes single-photoelectron gain. Lower panel: Differential dark count spectrum with experimental data (histogram) and generalized Poisson model fit (red curve) incorporating 6.6 % crosstalk parameter ($R^2 = 0.991$). Photoelectron peaks are clearly resolved with threshold positions indicated by dashed vertical lines.

This spectroscopic analysis was performed at all eleven temperature points, with fits achieving R^2 values from 0.958 to 0.996.

Temperature Dependence and Arrhenius Analysis

The temperature dependence of SiPM dark count rates follows Arrhenius behavior characteristic of thermally-activated carrier generation in semiconductor junctions. In silicon photodetectors, dark counts arise from thermal generation of electron-hole pairs that trigger avalanche breakdown in individual microcells. The rate of such thermally activated processes follows:

$$R(T) = A \cdot \exp\left(-\frac{E_a}{k_B T}\right) \quad (4.7)$$

where A represents a pre-exponential factor, E_a denotes activation energy, k_B is Boltzmann's constant, and T is absolute temperature. This exponential dependence explains why modest temperature changes produce dramatic dark count variations.

Silicon photomultipliers exhibit more complex behavior than single-mechanism Arrhenius scaling. At low temperatures, thermal generation occurs primarily through trap-assisted processes where carriers promoted from mid-gap defect states require only half the silicon bandgap energy ($E_a \approx E_g/2 \approx 0.5$ eV). At elevated temperatures, direct band-to-band thermal generation dominates, accessing the full bandgap energy ($E_a \approx E_g \approx 1.1$ eV). This physical transition motivates a piecewise Arrhenius model with distinct activation energies for each regime.

A unified model recognizes that underlying thermal generation mechanisms affect all thresholds similarly. The model incorporates threshold-dependent rate scaling while maintaining common activation energies:

$$R_n(T) = \begin{cases} A_{1,\text{low}} \cdot e^{-\alpha(n-1)} \cdot e^{-E_{a,\text{low}}/(k_B T)} & \text{for } T < T_c \\ A_{1,\text{high}} \cdot e^{-\alpha(n-1)} \cdot e^{-E_{a,\text{high}}/(k_B T)} & \text{for } T \geq T_c \end{cases} \quad (4.8)$$

where $R_n(T)$ denotes dark count rate at photoelectron threshold n and temperature T , $A_{1,\text{low/high}}$ represent pre-exponential factors for the respective thermal regimes, α characterizes exponential threshold scaling (higher thresholds require correlated multi-cell events, systematically reducing rates), $E_{a,\text{low/high}}$ denote activation energies for trap-assisted and band-to-band generation mechanisms, and T_c represents the critical transition temperature.

Figure 4.9 presents temperature-dependent integral dark count rates for photoelectron thresholds from 1 PE to 4 PE with the unified model fit.

The unified model describes all 44 data points with $R^2 = 0.974$, validating the assumption of common activation mechanisms across threshold levels. The fitted parameters are:

- Critical temperature: $T_c = -7.5 \pm 3.4$ °C
- Low-temperature activation energy: $E_{a,\text{low}} = 462 \pm 5$ meV $\approx E_g^{\text{Si}}/2$
- High-temperature activation energy: $E_{a,\text{high}} = 1,042 \pm 92$ meV $\approx E_g^{\text{Si}}$
- Threshold scaling parameter: $\alpha = 1.70 \pm 0.10$

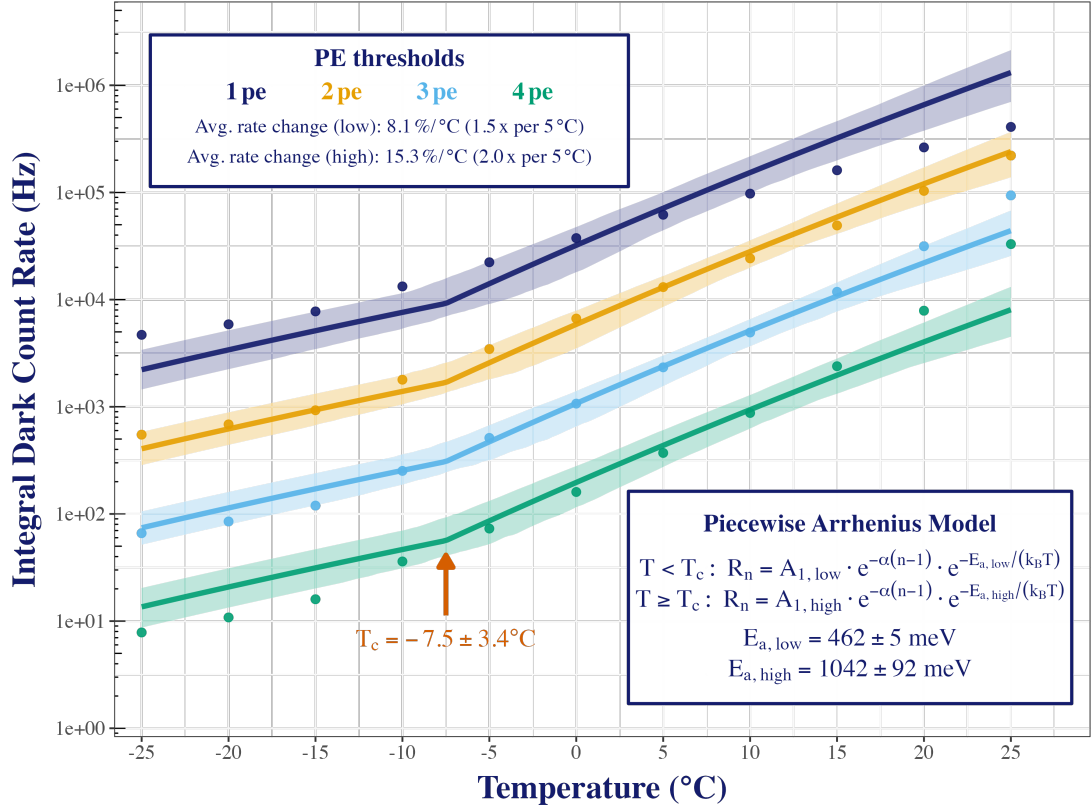


Figure 4.9: Temperature dependence of SiPM integral dark count rates fitted with unified piecewise Arrhenius model. Data points represent integral dark count rates at photoelectron thresholds from 1 PE to 4 PE extracted from spectroscopic analysis at eleven temperatures. Uncertainty bands show 95 % confidence intervals from Monte Carlo propagation of fit parameter covariances.

The low-temperature activation energy corresponds to half the silicon bandgap, consistent with Shockley-Read-Hall generation via mid-gap trap states. The high-temperature value approaches the full bandgap energy, indicating transition to diffusion-limited processes. These values align with established literature [40] and exhibit the well-documented halving rule for SiPM dark counts [41]: rates approximately double for every 5 °C temperature increase.

Accidental Coincidence Predictions for Phase B Implementation

The detection efficiency analysis established maximum photoelectron thresholds maintaining 90 % signal acceptance for each coincidence level, ranging from 5 PE for twofold coincidence down to 1 PE for fivefold. The unified piecewise Arrhenius model enables prediction of accidental coincidence backgrounds at these efficiency-compatible thresholds when scaled to Phase B implementation with 1,000 PbF₂ crystals.

Table 4.3 presents aggregate accidental rates across the detector system, calculated using 10 ns coincidence resolving time with combinatorial factors $\binom{5}{n}$ for all possible n -fold combinations from five SiPMs per crystal face.

Table 4.3: *Accidental coincidence rates (events/year) for Phase B implementation with 1,000 PbF_2 calorimeter crystals using SiPM arrays, calculated using unified piecewise Arrhenius model with 10 ns coincidence resolving time. Rates assume five SiPMs per crystal face with combinatorial factors $\binom{5}{n}$ for n -fold coincidence requirements. Values are rounded. Color coding: green = negligible (<1 event/year), blue = low (1–100 events/year), orange = medium (100–1,000 events/year), red = high ($>1,000$ events/year).*

Coinc.	Thresh.*	Temperature						
		–25 °C	–20 °C	–10 °C	0 °C	10 °C	20 °C	25 °C
2-fold	5 PE	$\sim 2 \times 10^4$	$\sim 5 \times 10^4$	$\sim 2 \times 10^5$	$\sim 4 \times 10^6$	$\sim 9 \times 10^7$	$\sim 2 \times 10^9$	$\sim 7 \times 10^9$
3-fold	4 PE	<1	<1	~ 3	$\sim 2 \times 10^2$	$\sim 3 \times 10^4$	$\sim 2 \times 10^6$	$\sim 2 \times 10^7$
4-fold	3 PE	<1	<1	<1	<1	$\sim 1 \times 10^2$	$\sim 4 \times 10^4$	$\sim 6 \times 10^5$
5-fold	1 PE	<1	<1	<1	~ 10	$\sim 3 \times 10^4$	$\sim 4 \times 10^7$	$\sim 1 \times 10^9$

*Maximum threshold maintaining 90 % detection efficiency.

Fourfold coincidence at 3 PE threshold emerges as the optimal configuration for Phase B implementation. This regime maintains negligible accidental rates (sub-event per year) from –25 °C to 0 °C, with only approximately 1×10^2 events per year at 10 °C—elevated but potentially manageable if aggressive cooling proves impractical.

The other coincidence levels exhibit fundamental limitations. Twofold coincidence proves incompatible with Phase B requirements regardless of cooling: even at –25 °C, the $\tau \nu^2$ suppression yields 2×10^4 accidentals per year. Threefold coincidence requires cooling to –10 °C or below to achieve acceptable rates. Fivefold coincidence, despite offering maximum $\tau^4 \nu^5$ suppression, suffers from the efficiency-imposed 1 PE threshold constraint that produces high per-channel dark rates near the noise floor; at 10 °C, fivefold yields 3×10^4 accidentals per year—comparable to threefold despite the higher coincidence order.

Optimization of SiPM Array Size

The five-SiPM array configuration represents the laboratory test geometry rather than an optimized detector design. The 30 mm $\times$ 30 mm crystal face can accommodate denser packing: with 6 mm $\times$ 6 mm devices, arrays up to 16 SiPMs (4 $\times$ 4 configuration) are geometrically feasible.

Competing effects govern this optimization. For fourfold coincidence from N_{total} SiPMs, possible combinations scale as $\binom{N_{\text{total}}}{4}$: expanding from 5 to 8 SiPMs increases combinations from $\binom{5}{4} = 5$ to $\binom{8}{4} = 70$, a factor of 14 enhancement in potential accidental pathways. However, more SiPMs viewing the same Cherenkov volume increase the probability that sufficient detectors exceed threshold despite Poisson fluctuations, enabling operation at elevated thresholds. Since dark count rates decline exponentially with threshold, even modest threshold increases produce dramatic rate reductions that outweigh the combinatorial growth.

Figure 4.10 presents this optimization, showing maximum threshold maintaining 90 % detection efficiency versus SiPM count for fourfold coincidence, with corresponding accidental rates at five operating temperatures from 5 °C to 25 °C.

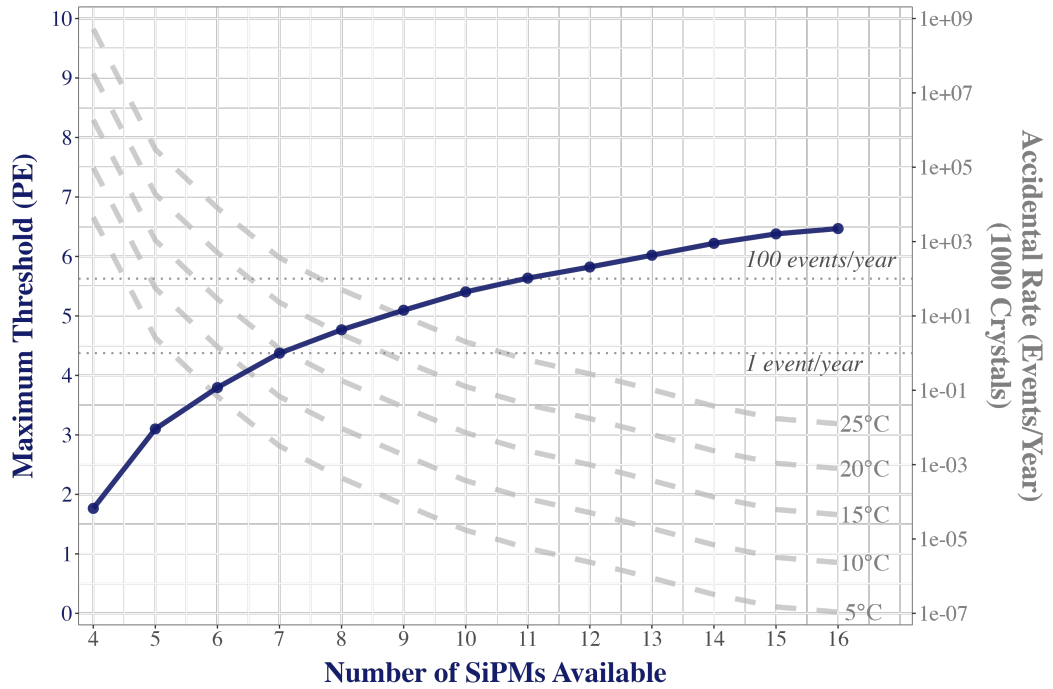


Figure 4.10: Maximum photoelectron threshold maintaining 90 % detection efficiency versus number of available SiPMs for fourfold coincidence. Primary axis shows threshold values from simulated efficiency curves; secondary axis displays corresponding accidental coincidence rates for Phase B implementation (1,000 PbF_2 crystals, 10 ns coincidence window) at operating temperatures from 5 °C to 25 °C. Reference lines indicate 1 and 100 events/year.

Increasing SiPM count can potentially eliminate cooling requirements altogether. For the baseline 5-SiPM configuration at 3 PE threshold, accidental rates exceed 1×10^5 events/year at 25 °C, necessitating cooling to 10 °C or below (Table 4.3). In contrast, 8-SiPM arrays achieve 90 % efficiency at approximately 5 PE threshold—a modest increase

that reduces accidental rates by several orders of magnitude due to the exponential threshold dependence combined with ν^4 scaling. At this configuration, accidental rates remain below 100 events/year at 25 °C and below 1 event/year even at 15 °C, enabling ambient-temperature operation without active cooling infrastructure.

However, actual operating temperatures depend on power dissipation within the readout electronics and thermal coupling to the environment. If SiPMs warm significantly above ambient temperature, dark count rates increase correspondingly, potentially restoring the need for thermal management. The optimization strategy must therefore balance three considerations: SiPM count (higher count enables higher thresholds), cooling infrastructure (active cooling adds complexity and cost), and electronics complexity (more channels require larger readout boards and increased per-crystal cost).

If thermal conditions prove unfavorable and cooling remains undesirable, further increasing SiPM count beyond 8 provides a viable path to suppress accidental rates at elevated temperatures. The additional electronics cost and complexity may prove acceptable compared to implementing and maintaining active cooling systems across 1,000 detector modules. This flexibility allows the final configuration to be optimized based on measured thermal performance during testing.

Fourfold coincidence with 8 SiPMs represents the baseline recommendation, achieving 90 % detection efficiency with negligible accidental backgrounds at ambient temperature and practical threshold while maintaining manageable electronics complexity. Higher SiPM counts remain available as a contingency if operating temperatures exceed expectations.

4.5 Summary and Conclusions

PMT dark count rates at 10 MeV detection thresholds posed the fundamental challenge introduced at the chapter's opening: spurious triggers at rates of 0.5 Hz to 4 Hz overwhelm rare dark matter signals predicted at few events per year. Dual-sided PMT readout serves as a Phase A solution through twofold coincidence suppression, but Phase B scaling to 1,000 PbF₂ crystals presents two significant obstacles: accumulated accidental backgrounds reach approximately 1×10^3 events per year even with optimal Hamamatsu PMTs, and photodetector procurement costs approach 1 M€ for 2,000

PMT channels. Marginal background performance combined with substantial cost motivated the exploration of alternative approaches.

Coupled crystal arrays offered higher-order coincidence and reduced channel count but failed through compounded limitations: reduced light yield (0.5 PE/MeV per PMT) and extreme distribution heterogeneity ($CV \approx 93\%$ versus 45 % Poisson limit) prevented viable operation at practical thresholds.

Silicon photomultiplier arrays overcome these limitations through a fundamentally different architecture. Single-crystal multi-SiPM geometry achieves $CV \approx 42\%$, matching the intrinsic Poisson statistical limit for 6 PE signals. This factor-of-two improvement over coupled crystal arrays eliminates systematic geometric heterogeneity, enabling viable fourfold coincidence operation despite modest per-device light yields of 0.6 PE/MeV. The cost advantage is equally significant: an 8-SiPM array for single-crystal readout costs approximately one-sixth of a Hamamatsu PMT pair, reducing Phase B photodetector procurement to approximately 150 k€—a factor of six savings compared to the dual-sided PMT approach.

The array size optimization reveals a critical finding: modest increases in SiPM count produce reductions in accidental coincidence rates through threshold leverage. The baseline 5-SiPM configuration achieves 90 % detection efficiency at 3 PE threshold, yielding approximately 6×10^5 accidental events per year at 25 °C—requiring cooling to 10 °C or below for viable operation. Increasing to 8 SiPMs raises the maximum threshold to approximately 5 PE, reducing accidental rates by several orders of magnitude due to the exponential threshold dependence of dark counts combined with v^4 scaling of fourfold coincidence. At this configuration, accidental rates remain below 100 events/year even at 25 °C, potentially eliminating the need for active cooling infrastructure.

Further increases in SiPM count push maximum thresholds to 6 PE and above, achieving sub-1 event/year accidental rates even at 25 °C. These configurations provide contingency options if thermal conditions prove unfavorable: power dissipation within readout electronics may warm SiPMs above ambient temperature, increasing dark count rates and potentially restoring cooling requirements. Higher SiPM counts offer a path to maintain negligible accidental backgrounds without cooling, trading increased electronics complexity against thermal management infrastructure.

Fourfold coincidence with 8 SiPMs is the baseline recommendation for Phase B implementation. This configuration achieves 90 % detection efficiency with acceptable

accidental rates at ambient temperature while maintaining manageable electronics complexity and compact board geometries suitable for dense calorimeter integration. Higher SiPM counts remain available as contingency if operating temperatures exceed expectations, allowing final optimization based on measured thermal performance during prototype testing.

Experimental Validation Requirements

The preceding analysis provides operational guidelines and performance estimates for SiPM-based Phase B implementations but relies on simulation-based predictions requiring experimental validation before deployment. Several critical assumptions introduce systematic uncertainties into quantitative predictions.

The accidental coincidence calculations employ simplified expressions neglecting higher-order correlation terms. While adequate for order-of-magnitude estimates, precision predictions at few-events-per-year levels require experimental validation of correlation assumptions and potential corrections for device-to-device coupling effects. The GEANT4 optical simulation predicts approximately 0.6 PE/MeV per SiPM based on manufacturer-specified photon detection efficiency but lacks validation against calibrated beam measurements at 10 MeV. Optical coupling quality, entrance window reflections, and geometric coverage factors could differ between idealized simulation geometry and assembled detectors.

The unified Arrhenius model derives from measurements of a single representative device, assuming consistent behavior across all detectors. However, SiPMs exhibit device-to-device variations arising from fabrication non-uniformities and junction quality differences. While the compound probability nature of fourfold coincidence provides some averaging, individual high-noise outliers could contribute disproportionately to aggregate backgrounds. The model itself represents a phenomenological simplification; quantitative rate predictions should therefore be treated as order-of-magnitude estimates requiring empirical confirmation.

The assumed 10 ns timing resolution represents achievable performance based on manufacturer specifications, but realistic timing jitter from cable length variations, discriminator non-uniformities, and temperature-dependent propagation delays could broaden effective coincidence windows, proportionally increasing accidental rates.

Transitioning to validated detector design requires systematic experimental characterization:

- Dark count characterization across production batches (20–50 devices) to quantify device-to-device variations and establish selection criteria for array assembly
- Long-term stability monitoring (weeks to months) at fixed operational conditions to assess aging effects, gain drift, and reliability for multi-year deployment
- Measurement of actual accidental coincidence rates with existing 5-SiPM arrays across the operational temperature and threshold range to validate Arrhenius scaling and combinatorial factors
- Electron-beam calibration at MAMI to measure absolute photoelectron yields, compare with simulation predictions, and assess position-dependent response across crystal volume
- Design and fabrication of 8-SiPM readout boards optimizing layout, signal routing, and thermal management, with implementation of on-board coincidence logic and analog summing circuits
- Extended operation of complete detector modules under realistic conditions to validate thermal stability and long-term reliability

These validation tasks, currently ongoing under [REDACTED]'s direction, will provide the empirical foundation for Phase B production decisions. The recommendations presented here establish operational guidelines pending experimental confirmation, with the understanding that quantitative thresholds and cooling requirements may require refinement based on measured performance.

Chapter 5

The DarkMESA Veto System

The search for dark matter through rare interaction events demands near-perfect background rejection. Cosmic rays and beam-related particles can mimic dark matter signals in the main detector, overwhelming any genuine dark matter events. The veto system provides the critical first defense: it identifies and tags these background particles while allowing hypothetical dark matter candidates to pass through undetected.

This discrimination is fundamentally challenging. Background particles must be detected with extraordinary efficiency across all energies and interaction angles, while the veto itself must not generate additional backgrounds or obscure potential dark matter signals. Unlike conventional particle physics experiments that tolerate moderate background contamination and rely on statistical subtraction, DarkMESA requires active background suppression approaching theoretical dark matter interaction rates. This threshold is orders of magnitude more stringent than typical experiments demand.

EJ-200 plastic scintillator panels (20 mm thickness) form the active veto layer, detecting particles through scintillation light produced by ionization. Lead absorber layers provide passive shielding, attenuating photons and reducing the energies of penetrating particles. Chapter 2 presents the complete detector geometry in detail, including veto panel arrangement and lead absorber placement.

Prior simulation studies established conceptual feasibility of the scintillator-based veto approach but left critical questions unresolved. How do detection efficiencies for specific particle types compare between the veto and calorimeter materials? How does efficiency vary with energy deposition across the cosmic ray spectrum? Most critically, what physical mechanisms allow particles to penetrate the veto at the target 10 MeV threshold? How can design improvements address them?

This chapter answers these questions through combined simulation and experimental analysis. Section 5.1 establishes baseline scintillator response to cosmic radiation, quantifying energy deposition, photon production, and interaction rates. Section 5.2 compares detection probabilities for each particle type between EJ-200 and PbF₂, revealing substantial differences for neutral particles. Section 5.3 presents experimental measurements of the Phase A configuration, compares them to simulation predictions, and highlights observed performance shortfalls. The section also characterizes leak-through events through simulation particle tracking. The resulting efficiency comparisons and understanding of penetration mechanisms establish the foundation for Phase B optimization strategies.

5.1 Cosmic Ray Response

Energy Deposition and Photon Production

Detailed simulation of EJ-200 plastic scintillator response to cosmic radiation must account for diverse particle types, energy distributions, and interaction mechanisms. The simulation geometry represents a single inner veto tile: a rectangular panel (500 mm × 250 mm × 20 mm, density 1.023 g/cm³). Figure 5.1 presents three complementary views: particle trajectory analysis, photon production efficiency, and energy deposition spectrum. The simulation spans 3,720.0 s (1.03 h) of operation, yielding 23.3 Hz total interaction rate for unshielded 20 mm EJ-200.

The scatter plot (Figure 5.1, middle) demonstrates linear correlation between energy deposition and scintillation photon production: 9,992.7 photons/MeV with exceptional linearity ($R^2 = 1.000$). This conversion factor exceeds the 109.7 photons/MeV observed in PbF₂ Cherenkov calorimeters by two orders of magnitude, reflecting fundamental differences between scintillation and Cherenkov mechanisms. Scintillation occurs through molecular de-excitation following ionization, producing isotropic emission independent of particle velocity. This contrasts with directional, threshold-dependent Cherenkov emission requiring $\beta > 1/n$. The high EJ-200 light yield enables reliable detection of lower energy deposits crucial for veto functionality.

Trajectory Characteristics by Particle Type

The trajectory table (Figure 5.1, top) reveals distinct behavior patterns for each particle type.

Table: Median Primary Particle Energy [MeV] / Path Length [mm]

Edep	1 MeV	2 MeV	3 MeV	4 MeV	5 MeV	7.5 MeV	10 MeV
Muon	2895/6	2518/11	2436/18	2233/22	2502/25	3292/33	4607/42
Electron	1/5	2/10	3/16	29/22	30/25	34/34	26/40
Neutron	2/34	3/32	4/33	6/43	7/24	14/33	42/22
Gamma	2/30	4/27	6/25	8/27	9/18	26/10	44/4
Proton	371/3	594/7	591/10	1534/21	915/21	360/22	225/24
Pion	1598/8	1598/8	16869/21	16869/21	189/25	662/39	—

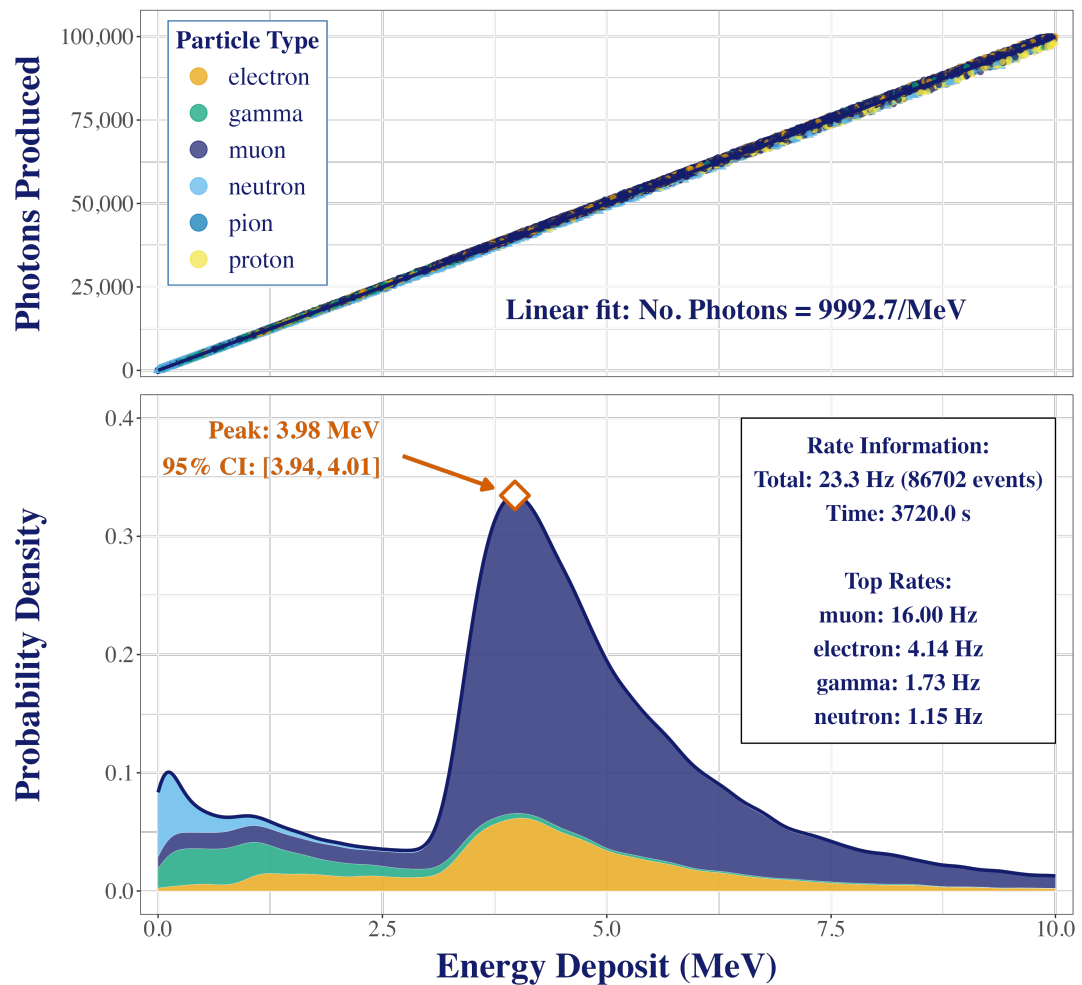


Figure 5.1: Scintillation response of EJ-200 plastic scintillator to unshielded cosmic radiation. Top: Median primary particle energies and path lengths binned by energy deposit. Middle: Energy deposition versus photon production with linear fit yielding 9,992.7 photons/MeV ($R^2 = 1.000$). Bottom: Energy spectrum with prominent 3.98 MeV peak from minimum ionizing muons traversing 20 mm thickness. Note: Pions exhibit identical values in the table across multiple energy bins because their low statistics required adaptive binning, where the search window progressively expands from ± 0.25 MeV up to ± 2.5 MeV until sufficient particles are captured.

Muon path lengths increase monotonically with energy deposit: 6 mm at 1 MeV to 42 mm at 10 MeV. This linear relationship confirms minimum ionizing behavior with characteristic ~ 2 MeV/cm energy loss. At 4 MeV deposit, muons traverse 22 mm—near the 20 mm panel thickness—producing the prominent 3.98 MeV spectral peak. This feature provides natural calibration analogous to the 29 MeV muon peak in PbF_2 , occurring at lower energy due to reduced material density and thickness.

Electron behavior reflects how the 20 mm detector thickness acts as a range filter that depends on energy, naturally separating electrons into two populations. Low-energy electrons (1 MeV to 3 MeV deposits) stop within the scintillator volume, depositing their full kinetic energy along path lengths of 5 mm to 16 mm. A discontinuous jump to 29 MeV median primary energy at higher deposit thresholds reflects a geometric selection effect: intermediate-energy electrons either stop completely or graze through, systematically avoiding the 4 MeV to 5 MeV deposit range. Deposits in this range instead correspond to high-energy electrons traversing the full detector thickness as minimum ionizing particles, depositing only a fraction of their kinetic energy.

Gamma rays show characteristic electromagnetic cascade behavior: path lengths decrease from 30 mm at 1 MeV deposited energy to just 4 mm at 10 MeV. At higher energies, pair production dominates over Compton scattering; gamma rays therefore convert into electron-positron pairs more rapidly. These shorter paths indicate that energetic gamma rays initiate electromagnetic showers closer to their entry point, depositing energy through their secondary particles.

Protons and neutrons show complex energy-path relationships that reflect hadronic interaction physics. Proton path lengths range from 3 mm at 1 MeV deposited to 24 mm at 10 MeV, with median primary energies between 200 MeV to 1,500 MeV. Neutrons can only be detected through secondary charged particles from nuclear interactions—primarily elastic n-p scattering, which generates recoil protons that deposit energy through ionization. Neutron path lengths vary from 22 mm to 43 mm, with primary energies spanning 2 MeV to 40 MeV.

Energy Spectrum Features

The energy spectrum (Figure 5.1, bottom) shows how each particle type contributes across the deposit range. Muons (dark blue) dominate the 3.98 MeV peak, corresponding to minimum ionizing particles traversing the full 20 mm thickness. The peak width reflects path length variations from angular distribution: vertical muons traverse

exactly 20 mm, while inclined trajectories produce longer paths and proportionally higher deposits.

Below the muon peak, electrons (orange) and gamma rays (green) contribute substantially to the low-energy spectrum. The steeply falling distribution reflects the cosmic ray energy spectrum convolved with energy-dependent interaction probabilities. The sub-1 MeV region contains mixed contributions from all particle types, with gamma-induced events becoming increasingly prominent at lowest energies due to Compton scattering producing low-energy electron recoils.

Muons account for 68.7 % of total interactions (16.00 Hz), electrons 17.8 % (4.14 Hz), gammas 7.4 % (1.73 Hz), and neutrons 4.9 % (1.15 Hz). The 23.3 Hz total rate substantially exceeds the calorimeter's 2.2 Hz, reflecting both larger veto effective area and enhanced sensitivity to low-energy deposits from superior scintillation light yield compared to Cherenkov emission.

Particle Interaction Characteristics

Interaction outcome analysis reveals the veto detector's particle discrimination capabilities. Figure 5.2 categorizes cosmic ray particles by interaction signature, revealing critical differences from PbF_2 that directly influence veto performance.

Charged particles achieve near-unity detection probability. Electrons and muons reach 100 % interaction efficiency with photon production—every incident charged particle deposits energy generating scintillation light. Because charged particles traversing material always interact electromagnetically, any charged cosmic ray component necessarily triggers detection. This principle provides the fundamental basis for veto functionality.

Protons achieve 100 % efficiency through ionization energy loss, independent of velocity threshold. This contrasts with PbF_2 Cherenkov detection requiring $\beta > 0.56$ ($E_p > 196$ MeV). Scintillation detection operates without velocity threshold, enabling veto sensitivity across the full cosmic proton spectrum including low-energy particles invisible to Cherenkov-based calorimeters.

Gamma ray interactions reveal a critical limitation: only 7.5 % of incident photons produce detectable signals in 20 mm scintillator, while 92.5 % traverse without interaction. This low efficiency reflects the probabilistic nature of photon interactions in low-Z organic material, where photoelectric absorption probability scales as Z^{4-5} and

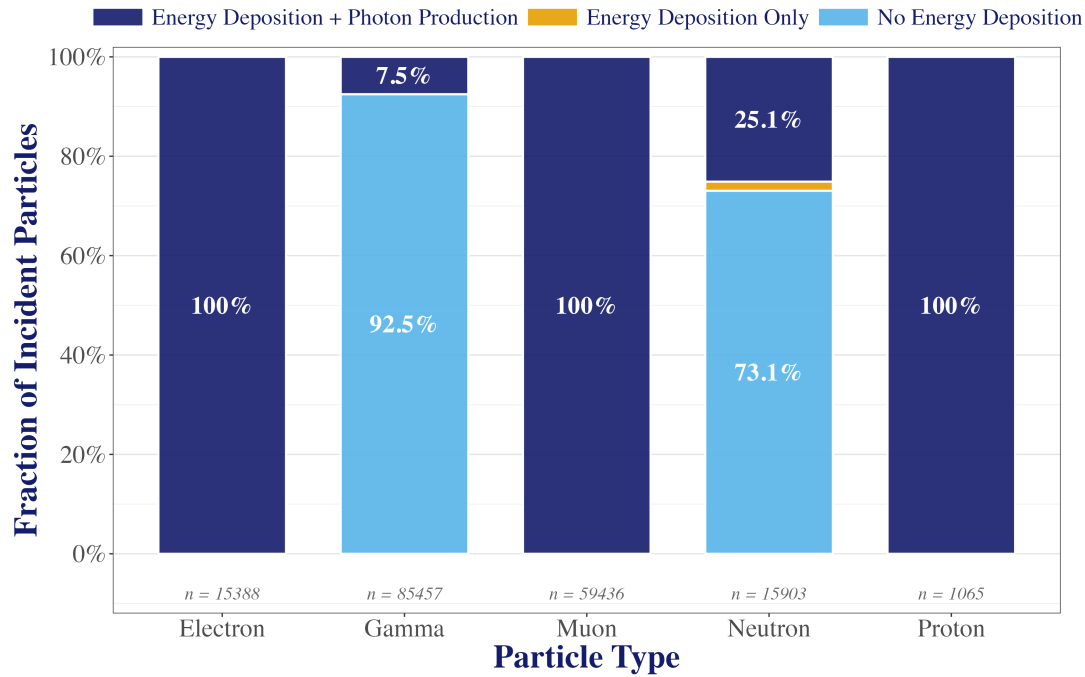


Figure 5.2: Interaction outcome classification for cosmic ray components in EJ-200 scintillator. Each particle type displays fractions producing energy deposition with photon production (dark blue), energy deposition without photons (orange), or no deposition (light blue). Sample sizes indicated below bars. Analysis encompasses 20 mm panels with CRY cosmic ray generation and GEANT4 transport.

Compton scattering dominates at cosmic ray energies. PbF_2 achieves 55.7 % cosmic gamma interaction efficiency (Figure 3.5)—an order-of-magnitude performance gap that limits the scintillator’s effectiveness as a gamma veto.

Neutrons achieve 25.1 % interaction efficiency—substantially exceeding the 9.3 % in PbF_2 . This factor of 2.7 improvement reflects the hydrogen-rich composition of organic scintillator enabling efficient elastic n-p scattering. Recoil protons from these interactions deposit energy through ionization, producing detectable scintillation regardless of neutron energy. This neutron sensitivity advantage validates organic scintillator selection for hadronic background rejection.

5.2 Energy-Dependent Detection Efficiency

Spectrum-averaged interaction probabilities establish baseline performance but conceal substantial energy-dependent variations critical for veto optimization. Plotting efficiency against primary particle energy enables quantitative comparison between

scintillator and PbF_2 , revealing distinct performance regimes where each material offers advantages.

The analysis employs the Generalized Additive Model (GAM) framework established for PbF_2 calorimeter studies (Section 3.7). The detection threshold requires at least one photoelectron detected by any of 18 SiPM readouts, representing the minimum signal discriminable from noise. Figure 5.3 presents energy-dependent efficiency comparisons between detector technologies. Efficiency ratios exceeding unity indicate scintillator advantage; ratios below unity indicate PbF_2 superiority.

Gamma Ray Detection

Gamma rays exhibit the most pronounced efficiency deficit for scintillator. The ratio remains substantially below unity across the entire energy range (1 MeV to 1,000 MeV), with scintillator efficiency reaching only 5 % to 7 % compared to PbF_2 's 25 % to 75 %. The ratio decreases from ~ 0.7 at 1 MeV to ~ 0.05 at high energies. This trend reflects the physics of photon interactions: high- Z materials like PbF_2 ($Z_{\text{Pb}} = 82$, $\rho = 7.77 \text{ g/cm}^3$) provide enhanced photoelectric absorption at low energies and pair production cross-sections scaling as Z^2 at high energies, compared to low- Z organic scintillator ($Z_{\text{C}} = 6$, $Z_{\text{H}} = 1$, $\rho = 1.023 \text{ g/cm}^3$).

This order-of-magnitude efficiency gap demonstrates that thin scintillator layers alone provide minimal gamma ray rejection. Two complementary strategies address this limitation. First, multi-crystal event topology in the PbF_2 array provides discrimination: gamma rays typically traverse multiple crystals before absorption. A gamma ray with 50 % per-crystal detection probability traversing 10 crystals produces signals in at least two crystals with 99 % probability, enabling rejection based on multi-crystal signatures inconsistent with single-site dark matter interactions. Second, lead absorber layers between veto planes attenuate gamma flux. Nevertheless, single-crystal gamma events at detector boundaries represent an irreducible background requiring geometric optimization.

Neutron Detection

Neutron detection exhibits the most striking scintillator advantage, with efficiency ratios spanning three orders of magnitude: $R \approx 1,000$ at 0.1 MeV declining to $R \approx 0.3$ at 1,000 MeV. Scintillator efficiency peaks at approximately 90 % near 0.1 MeV, while PbF_2 remains below 10 % throughout. This performance differential validates organic scintillator selection for hadronic background rejection.

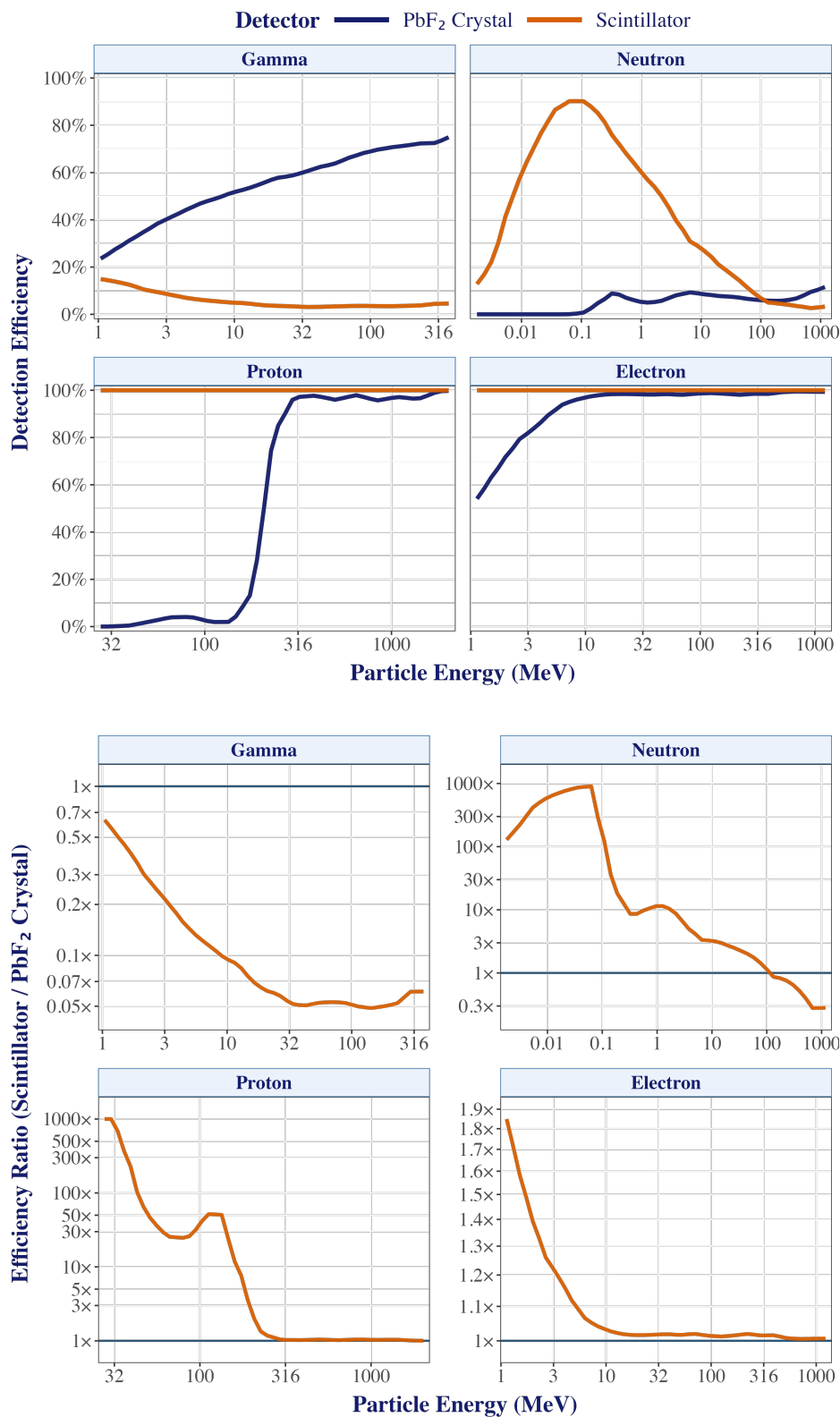


Figure 5.3: Detection efficiency comparison between EJ-200 scintillator and PbF₂ Cherenkov calorimeter. Top: Absolute efficiencies versus primary particle energy for each particle type. Bottom: Efficiency ratios (scintillator/PbF₂) revealing energy regimes where each material dominates.

The underlying mechanism reflects the hydrogen content of plastic scintillator and energy-dependent neutron-proton elastic scattering cross-sections [42, 43]. Elastic n-p scattering transfers substantial kinetic energy to recoil protons—up to the full neutron energy in head-on collisions—which deposit energy through dense ionization tracks producing scintillation light. This mechanism proves particularly effective at low energies (0.01 MeV to 10 MeV), where elastic cross-sections are large and maximum energy transfer to light nuclei is kinematically favored. In contrast, neutron interactions in PbF_2 produce heavy nuclear recoils that cannot generate Cherenkov radiation.

Efficiency convergence at high energies ($E_n > 100$ MeV) reflects increasing importance of inelastic processes: quasi-elastic proton knockout, pion production, and spallation reactions generate charged secondaries capable of exceeding the Cherenkov threshold [44]. Simultaneously, n-p elastic cross-sections decrease at high energies, reducing the scintillator’s hydrogen-based advantage.

The high scintillator efficiency in the sub-MeV regime requires contextual interpretation. DarkMESA targets a 10 MeV energy threshold; neutrons below this energy cannot deposit sufficient energy to mimic dark matter signals. For higher-energy neutrons capable of exceeding this threshold, scintillator efficiency decreases but remains substantially above PbF_2 up to ~ 100 MeV. The dual-layer veto design further enhances cumulative detection probability.

Proton Detection

Proton detection in PbF_2 is governed by the Cherenkov velocity threshold. Protons must achieve $\beta > 1/n \approx 0.56$ (where $n = 1.78$ for PbF_2) to produce Cherenkov radiation, corresponding to kinetic energies exceeding approximately 196 MeV. Below this threshold, protons deposit energy through ionization without generating detectable optical signals.

This threshold behavior reduces background concerns significantly. Cosmic ray protons below 196 MeV cannot produce calorimeter signals and therefore pose no background threat regardless of veto detection. Above threshold, both detectors respond with $>99\%$ efficiency, ensuring reliable veto tagging. In contrast, the scintillator exhibits no velocity threshold: protons of any kinetic energy produce scintillation through ionization, ensuring detection across the full spectrum.

Electron Detection

Electron detection efficiency approaches unity for both detector types above 10 MeV, validating suitability for the primary physics goal: detecting electron recoil signatures from dark matter interactions. Electrons energetic enough to penetrate veto panels and shielding necessarily deposit substantial energy in the scintillator, ensuring efficient tagging. Lower-energy electrons are absorbed within veto and shielding materials before reaching the calorimeter.

5.3 Veto Performance Characterization

The energy-dependent efficiency analysis (Section 5.2) establishes theoretical performance limits based on detection probabilities for single particles, but experimental validation under realistic operating conditions remains essential.

This section characterizes Phase A veto performance using a reduced test configuration: a single calorimeter crystal surrounded by the inner veto layer. This simplified geometry enables systematic efficiency measurements without complications from the full multi-crystal array. A complementary GEANT4 simulation matching the experimental geometry enables detailed event classification, revealing physical mechanisms responsible for veto penetration.

The experimental objectives include three complementary goals:

- **Absolute efficiency measurement:** Quantification of veto detection efficiency as a function of calorimeter energy across the range relevant for dark matter searches
- **Leak-through characterization:** Analysis of the energy spectrum for events depositing energy in the calorimeter without triggering the veto—so-called leak-through events—providing insight into penetration mechanisms and systematic vulnerabilities
- **Simulation validation:** Comparison between measurements and GEANT4 predictions, establishing confidence in simulation-guided optimization

Experimental Configuration

The experimental setup consists of a single tapered PbF_2 calorimeter crystal ($30 \text{ mm} \times 30 \text{ mm} \times 185.4 \text{ mm}$) with dual-sided Photonis XP2900 PMT readout, positioned at the

center of the complete Phase A inner veto assembly. Section 2.4 describes the Phase A detector geometry and readout configuration.

The inner veto consists of six EJ-200 plastic scintillator panels (20 mm thickness) providing hermetic coverage: four lateral panels (500 mm $\times$ 250 mm) and two endcap panels (250 mm $\times$ 250 mm). Each panel employs dual-sided SiPM readout with nine devices per face, providing 18-channel coverage. Carrier boards route signals to collector boards performing two-stage amplification (gain 100) and generating analog sum outputs.

Figure 5.4 presents both experimental hardware and corresponding simulation geometry, enabling direct comparison between measured efficiency and simulation predictions.

Analysis Methodology

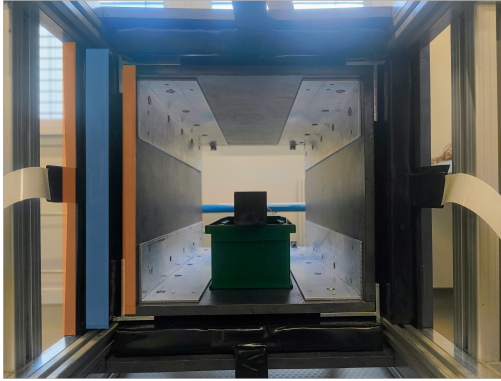
Parallel processing of simulation and experimental data enables direct efficiency comparison while accounting for detector response differences. Figure 5.5 illustrates the complete workflow, highlighting integration of previously established detector characterization results.

Simulation Chain

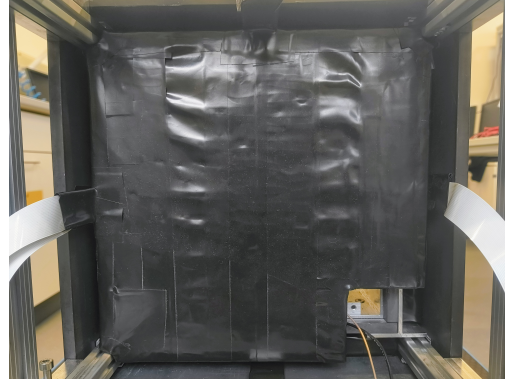
GEANT4 simulations employ the CRY cosmic ray generator for realistic particle flux. For each event, the framework records scintillation photon production in each veto panel and photoelectrons detected at each PMT following optical transport. The PMT cascade model (Section 3.2) transforms photoelectron counts into realistic charge outputs.

Although GEANT4 records true energy deposition in the crystal, the analysis instead reconstructs energy from PMT photoelectron output. The reconstruction employs arithmetic averaging of both PMT signals calibrated to the 29 MeV cosmic muon peak. This approach preserves realistic detector response. The reconstruction procedure ensures that efficiency comparisons between simulation and experiment reflect genuine performance differences rather than methodological artifacts.

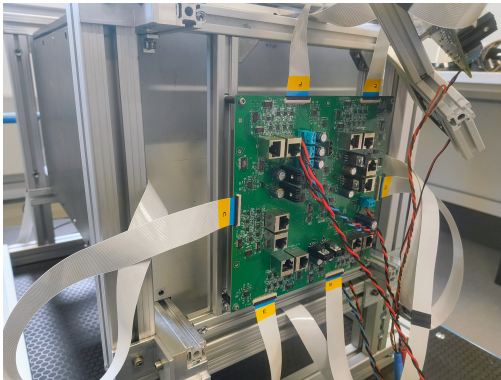
Coincidence requires both PMTs to exceed 6 PE, ensuring 90 % detection efficiency at 10 MeV (Section 4.2).



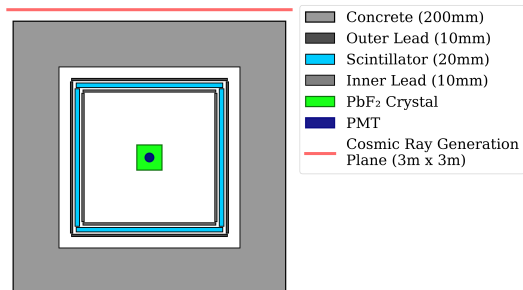
(a) Cross-sectional view of inner veto system showing EJ-200 scintillator panels (blue) between 10 mm lead layers (orange). The PbF_2 crystal with dual-sided PMT readout occupies the central volume. White ribbon cables connect to SiPM readout boards.



(b) Endcap veto panel with cutout for PMT cable routing. In final assembly, the panel is enclosed between 10 mm lead plates, maintaining shielding while accommodating signal and high-voltage feedthroughs.



(c) SiPM collector board implementing analog summation and amplification of six SiPM strips. Each inner veto panel requires two boards for twelve-strip coverage.



(d) GEANT4 simulation geometry (not to scale) implementing identical detector configuration within 200 mm concrete enclosure approximating laboratory environment.

Figure 5.4: Experimental implementation and simulation geometry for veto system performance characterization.

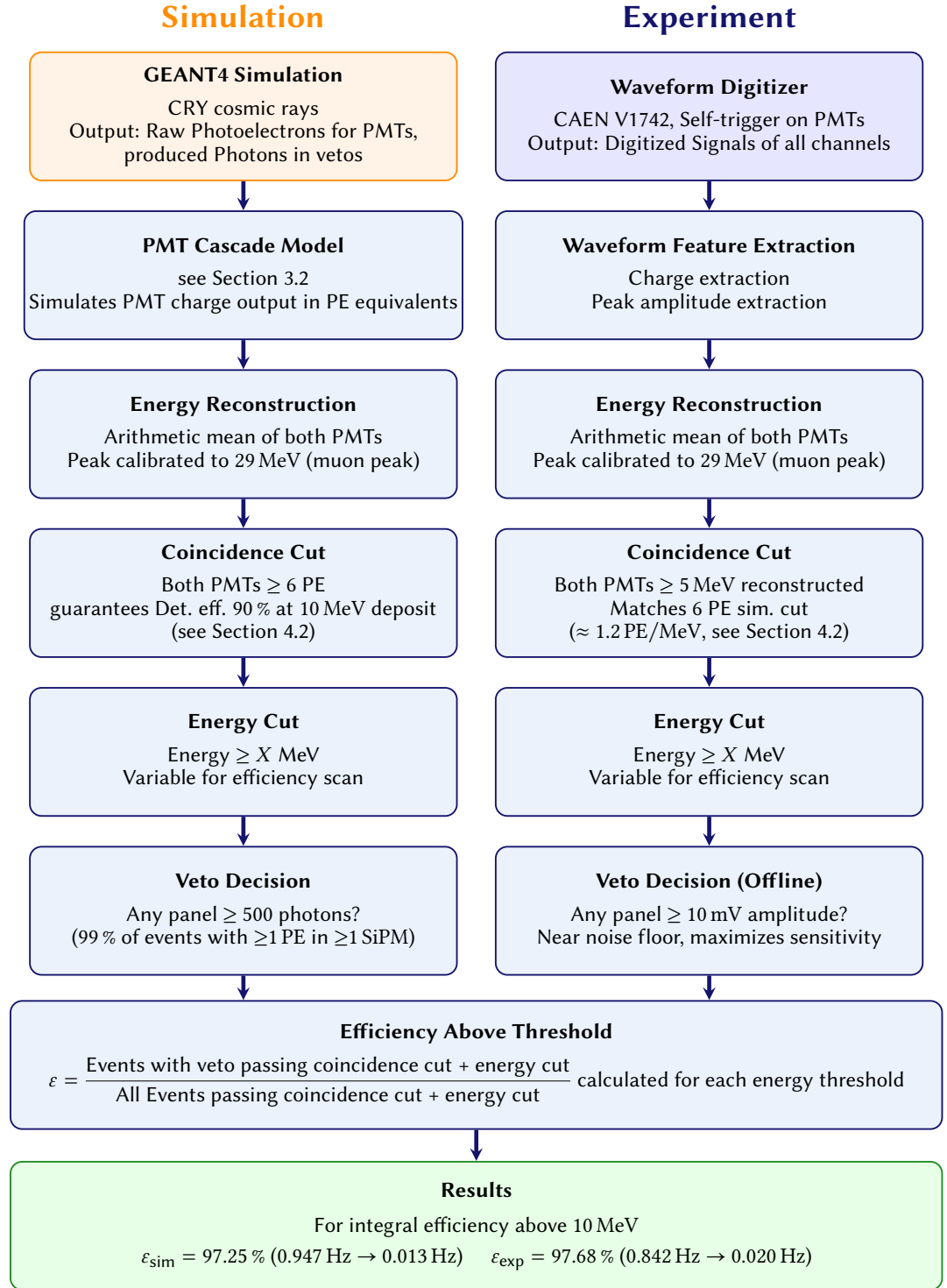


Figure 5.5: Analysis flow diagram showing parallel processing of simulation and experimental data streams. The workflow incorporates validated detector response models, enabling consistent event selection across both branches.

An event is rejected if any panel's scintillation output exceeds 500 photons. Directly simulating complete SiPM photoelectron detection for all scintillation photons would impose prohibitive computational cost given the statistics required. Single-panel simulation with full 18-SiPM readout (Section 5.2) established that 500 photons ensures $\geq 99\%$ probability of at least one SiPM detecting ≥ 1 PE. This threshold represents a conservative baseline.

Experimental Chain

Data acquisition employs a CAEN V1742 waveform digitizer with self-trigger on calorimeter PMT channels. Whenever either PMT exceeds the discriminator threshold, the digitizer captures waveforms from all channels: both calorimeter PMTs and all veto readouts.

Offline processing extracts charge integrals and peak amplitudes. Energy reconstruction follows identical procedures to simulation: arithmetic averaging calibrated to the 29 MeV muon peak. Coincidence requires both PMTs to exceed 5 MeV, corresponding to the 6 PE simulation threshold through the 1.2 PE/MeV conversion established in Section 4.2. An event triggers the veto if any panel amplitude exceeds 10 mV—a threshold positioned near the electronic noise floor to maximize sensitivity.

Efficiency Determination

For both data streams, veto detection efficiency at threshold E_{thresh} follows:

$$\varepsilon(E_{\text{thresh}}) = \frac{N_{\text{vetoed}}(E \geq E_{\text{thresh}})}{N_{\text{total}}(E \geq E_{\text{thresh}})} \quad (5.1)$$

where N_{total} denotes events passing coincidence with energy exceeding threshold, and N_{vetoed} represents events triggering veto detection. Varying the energy threshold from 10 MeV upward enables efficiency characterization across the dark matter search range.

Simulation access to complete event kinematics enables classification of leak-through events by particle type and interaction mechanism—information unavailable from experimental measurements alone.

Experimental Results and Performance Validation

Absolute Veto Detection Efficiency

Figure 5.6 presents veto detection efficiency versus calorimeter energy threshold. The analysis encompasses 24.69 h of simulation and 64.53 h of experimental acquisition, providing statistics to characterize efficiency across the dark matter search energy range.

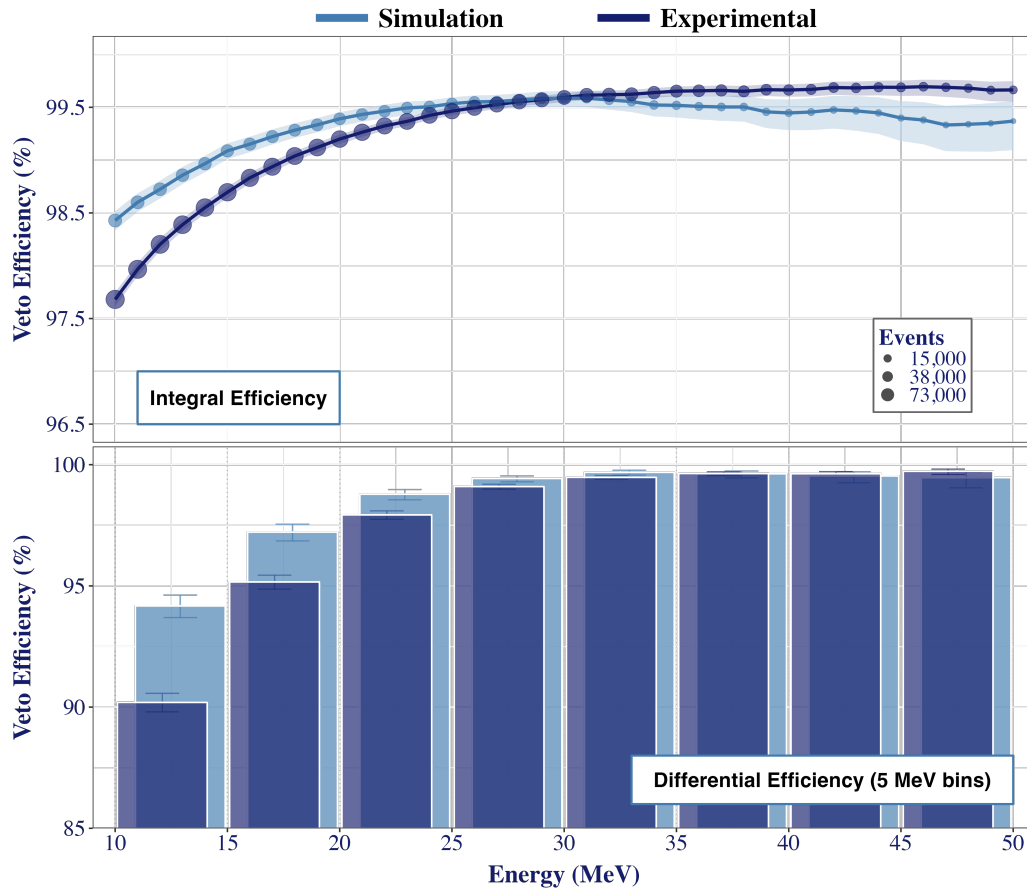


Figure 5.6: Veto detection efficiency versus calorimeter energy threshold. Top: Integral efficiency—probability that events with deposit $\geq E_{\text{thresh}}$ trigger veto. Simulation (light blue) and experiment (dark blue) demonstrate excellent agreement. Marker sizes scale with event counts; shaded bands represent Wilson 95 % confidence intervals. Bottom: Differential efficiency in 5 MeV bins revealing energy-dependent variations.

For events depositing ≥ 10 MeV, the veto achieves 97.25 ± 0.15 % efficiency in simulation and 97.68 ± 0.07 % experimentally. Event rates reduce from 0.947 Hz to 0.013 Hz (simulation) and 0.842 Hz to 0.020 Hz (experiment). This close agreement validates GEANT4 modeling, establishing confidence for larger-scale optimization.

Integral efficiency rises from 97.6 % at 10 MeV to >99.5 % above 30 MeV. However, differential analysis (Figure 5.6, bottom) exposes a critical vulnerability: efficiency degrades precisely in the 10 MeV to 20 MeV range most relevant for dark matter detection. This energy dependence creates a problematic scenario. The veto performs best suppressing high-energy backgrounds potentially distinguishable through energy cuts, while exhibiting weakest performance at the 10 MeV threshold where signal-background discrimination proves most challenging.

The 2 % to 5 % residual leak-through at 10 MeV represents events with signatures indistinguishable from dark matter—the primary systematic limitation for DarkMESA sensitivity. The statistical analysis of Section 2.7 established that probing thermal relic targets requires backgrounds at the $\mathcal{O}(10^3)$ event level or below. At ~ 0.02 Hz, the current leak-through rate translates to $\sim 600,000$ unvetoes annually—nearly three orders of magnitude above this target. Achieving $\mathcal{O}(10^3)$ residual events requires leak-through rates below 3×10^{-5} Hz, corresponding to veto efficiency exceeding 99.99 %. The measured ~ 98 % efficiency falls substantially short. Compounding this limitation, differential efficiency degrades in the 10 MeV to 20 MeV range—precisely where dark matter detection demands peak performance.

Leak-Through Event Characterization

Events penetrating the veto undetected warrant detailed investigation. Figure 5.7 presents energy distributions before and after veto classification, with simulation revealing particle-type composition.

The experimental leak-through spectrum shows two features with opposing implications. The encouraging result: the 29 MeV muon peak has been eliminated, demonstrating successful rejection of >90 % of cosmic flux. However, residual events concentrate in the 10 MeV to 20 MeV range—precisely the dark matter search regime. For searches targeting few signal events per year, even this small residual necessitates improved performance.

Simulation enables classification unavailable from experiment alone. Color-coded stacking (Figure 5.7, right) reveals that primary gamma rays dominate leak-through events across the full energy range. This composition directly reflects the physics challenge: thin organic scintillator provides 7.5 % gamma interaction probability versus 55.7 % in PbF_2 . The complete absence of primary electrons confirms that charged particles reliably trigger veto detection. The leak-through vulnerability targets neu-

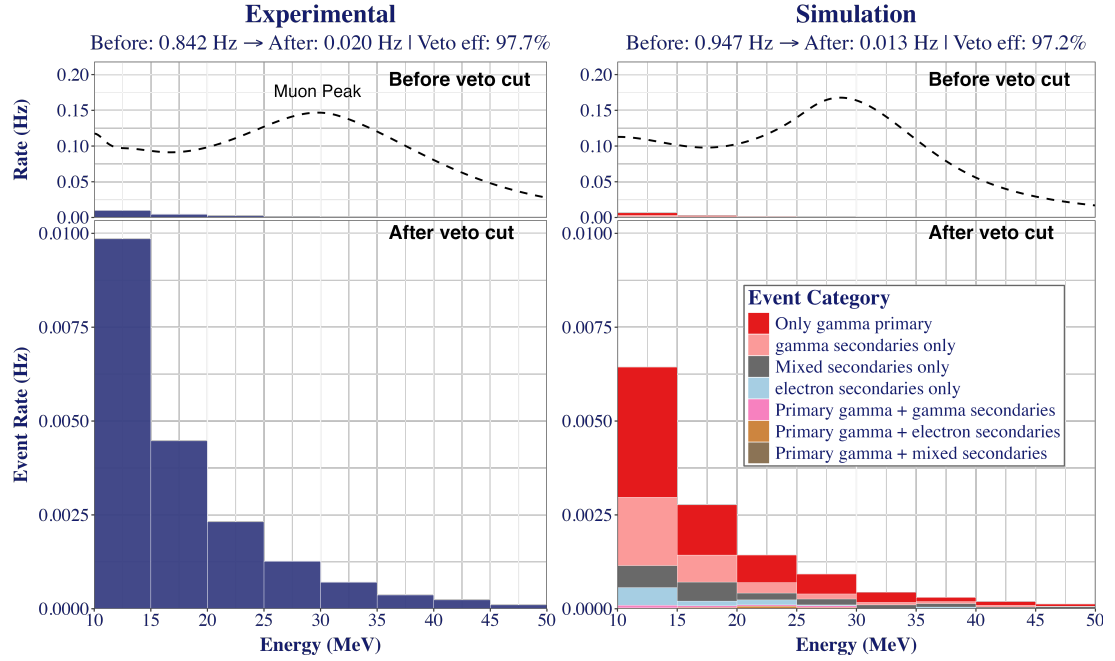


Figure 5.7: Energy spectra before and after veto classification. Experimental (left) and simulation (right). Top: Pre-veto distributions showing 29 MeV muon peak. Bottom: Residual leak-through spectra. Simulation employs particle tracking to classify events by primary type (gamma: red shades) and secondary mechanisms.

tral particles with weak scintillator interaction probability and charged secondaries produced too close to the calorimeter for interception.

The experimental geometry exhibits ~ 5 mm gaps at panel corner junctions from mechanical support structures. Particle tracking confirmed that no unvetoes events traversed exclusively through these gaps; all leak-through events traversed at least one scintillator panel. Geometric inefficiencies from panel gaps contribute negligibly compared to the physics limitation of thin organic scintillator.

Secondary Particle Production

Beyond primary gamma penetration, secondary production in inner lead layers creates geometric blind spots. Figure 5.8 traces creation locations and production processes for secondary particles in leak-through events.

The alluvial diagram reveals a critical vulnerability: secondary particles depositing energy in the calorimeter originate predominantly in passive shielding materials, particularly inner lead layers. Secondary gamma rays arise primarily through bremsstrahlung in lead and building concrete. Bremsstrahlung production concentrates in the inner lead (visible as the dominant purple flow in Figure 5.8). This correlates with a geomet-

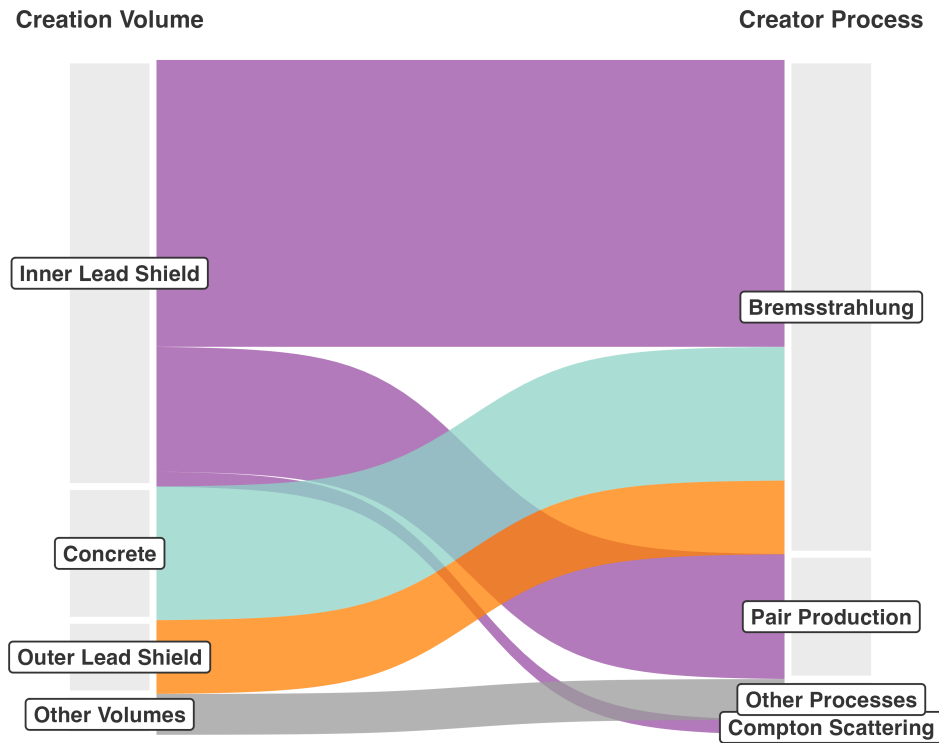


Figure 5.8: Alluvial flow diagram showing creation locations (left) and production processes (right) for secondary particles in leak-through events. Band widths scale with event counts. Secondary gamma production is dominated by bremsstrahlung in inner lead (purple) and concrete (light blue). Pair production and Compton scattering contribute across passive volumes. Problematic secondaries originate predominantly in passive materials adjacent to the calorimeter.

ric vulnerability: particles created closest to the calorimeter face the shortest path to deposit energy while avoiding veto interception.

Figure 5.9 presents energy distributions of particles contributing to leak-through events, distinguishing primary cosmic particles from secondaries produced through cascade development.

Primary gamma rays exhibit peaked distribution centered at 20 MeV to 40 MeV, while secondary gammas demonstrate broader distribution extending to lower energies. The electron distribution reveals exclusively secondary production—no primary cosmic electrons contribute to leak-through. This striking absence confirms that charged particles reliably trigger veto detection. Secondary electrons from pair production or Compton scattering peak at 15 MeV to 30 MeV. The average multiplicity of 1.72 particles per leak-through event demonstrates that many unvetoes events involve complex electromagnetic showers partially contained within detector volumes.

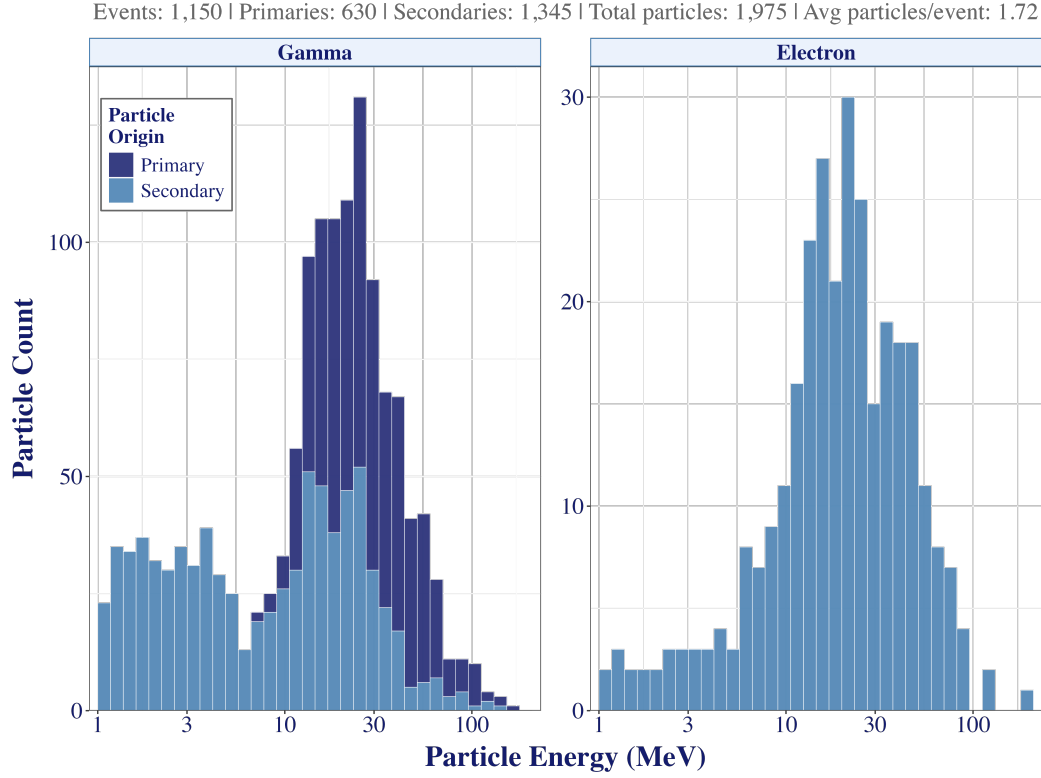


Figure 5.9: Energy distributions of particles in leak-through events, separated by type (gamma: left; electrons: right) and origin (primary: dark blue; secondary: light blue). Header indicates 1,150 leak-through events involving 630 primaries and 1,345 secondaries, averaging 1.72 particles per event. Primary gammas peak at 20 MeV to 40 MeV; secondary gammas extend to lower energies. Electrons appear exclusively as secondaries, peaking at 15 MeV to 30 MeV.

Implications for Detector Optimization

The present measurements characterize a reduced configuration with only the inner veto layer. Adding a second scintillator layer alone cannot substantially improve efficiency against gamma-dominated leak-through—thin organic scintillator exhibits minimal gamma interaction probability regardless of layer count.

Substantial enhancement requires thick passive lead absorbers. The leak-through particle energy distributions (Figure 5.9) reveal that problematic primary gamma rays peak at 20 MeV to 40 MeV. At these energies, lead exhibits attenuation coefficients $\mu \approx 0.5 \text{ cm}^{-1}$ to 0.6 cm^{-1} (Figure 3.6). For 99 % primary particle attenuation:

$$x = \frac{\ln(100)}{\mu} \approx \frac{4.61}{0.55 \text{ cm}^{-1}} \approx 8.4 \text{ cm} \quad (5.2)$$

Crucially, the energy transmission analysis in Section 3.4 demonstrates that this thickness not only removes primary gamma rays but also achieves nearly complete energy containment. Forward energy fractions fall below 10 % for lead thicknesses exceeding 6 cm (Figure 3.7), indicating that secondary particles produced in electromagnetic showers deposit their energy within the absorber rather than propagating to the detector. The required 8.4 cm therefore provides effective suppression of both primary gammas and their shower products.

Optimal placement of this absorber thickness—between veto layers, external to both, or distributed asymmetrically—requires systematic simulation study. Analysis of veto response patterns provides initial guidance for prioritizing shielding placement.

Veto Response Patterns

Characterization of veto response patterns validates the simulation framework while revealing systematic detector characteristics. Figures 5.10 and 5.11 present complementary views of veto system response.

Panel firing frequencies: The asymmetric firing pattern directly reflects cosmic ray angular distribution. Top and bottom panels exhibit highest firing frequencies, while lateral and endcap panels show substantially reduced rates. This confirms that cosmic interactions predominantly involve vertical trajectories.

The 17 % discrepancy in bottom panel frequency between simulation (84.2 %) and experiment (67.1 %) may reflect imperfect environmental modeling or panel-to-panel threshold variations. Additionally, the right panel exhibits lower firing frequency than the left, despite symmetric simulation predictions. This asymmetry reflects detector orientation toward the adjacent A1 experiment hall, which features thick shielded walls attenuating cosmic flux from that direction. The simulation employed simplified 200 mm uniform concrete, neglecting complex directional shielding variations.

Panel multiplicity: The distribution peaks sharply at two panels (simulation: 77.5 %, experiment: 82.2 %), corresponding to vertical penetration through top and bottom coverage. This two-panel dominance confirms hermetic detection for the primary cosmic component. Events triggering three or more panels represent oblique trajectories or shower development. Single-panel events (3 % to 6 %) indicate weak interactions producing signals in only one panel.

Ceiling shielding implications: The pronounced vertical asymmetry suggests targeted shielding strategies. Concentrated flux through the upper surface indicates

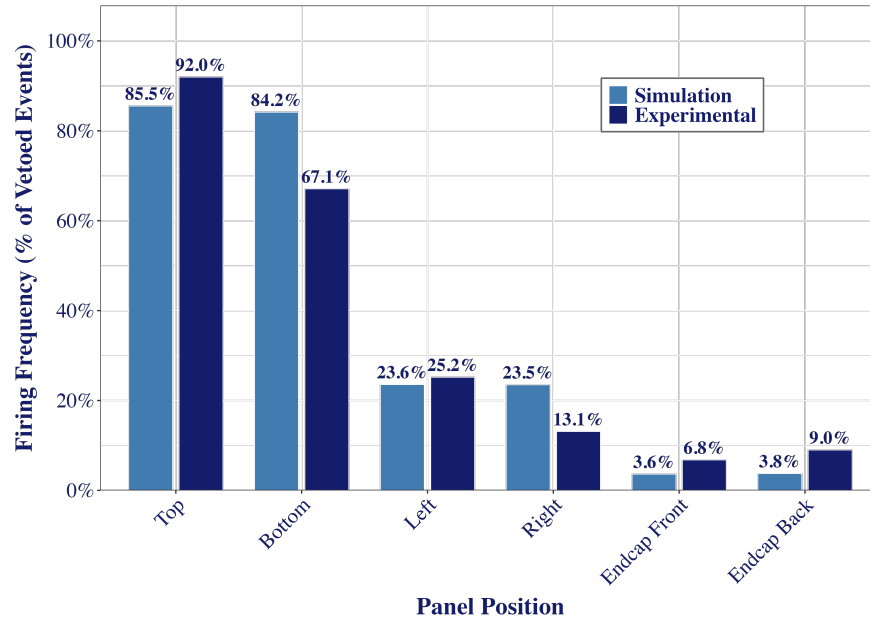


Figure 5.10: Firing frequency by panel position for vetoed events. Predominantly vertical cosmic incidence produces preferential detection in horizontal panels (Top: 85–92 %, Bottom: 67–84 %) compared to lateral (13–25 %) and endcap (<10 %) panels. Excellent simulation-experiment agreement validates geometric modeling.

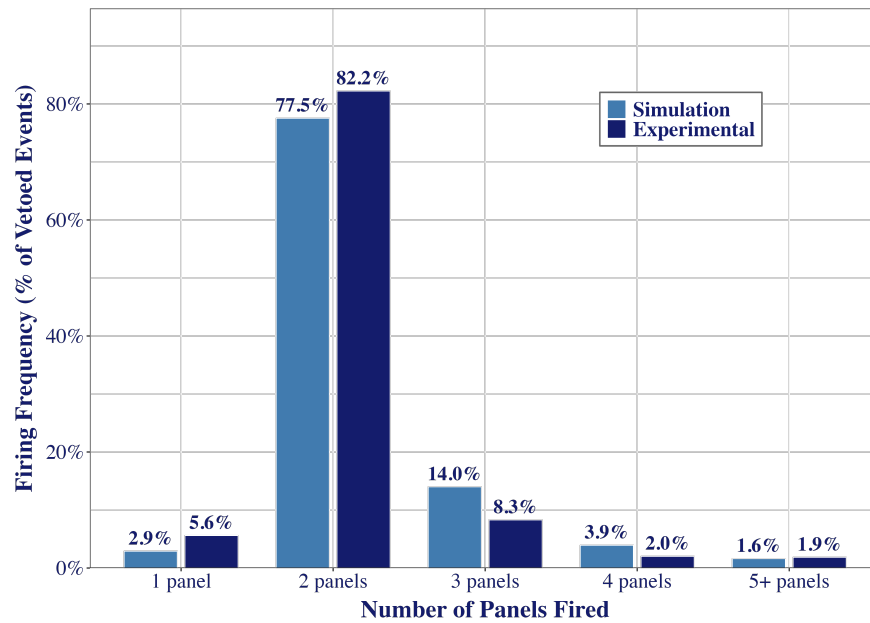


Figure 5.11: Distribution of triggered panel multiplicities per veto event. Most events (77–82 %) trigger exactly two panels, corresponding to vertical trajectories through top and bottom. Higher multiplicities reflect oblique trajectories or shower development. Single-panel events (3–6 %) indicate weak interactions or threshold effects.

ceiling-mounted passive shielding or active veto could provide disproportionate background suppression. However, as the leak-through analysis demonstrates (Figure 5.8), passive absorbers generate secondary particles, potentially replacing primary backgrounds with secondary-induced backgrounds. Detailed simulation incorporating realistic laboratory geometry and secondary production is required to quantify achievable reduction.

5.4 Summary and Conclusions

The statistical requirements established in Section 2.7—backgrounds at the $\mathcal{O}(10^3)$ event level to probe thermal relic targets—demand veto detection efficiency exceeding 99.99 % at 10 MeV. Experimental measurements and GEANT4 simulations characterized Phase A inner veto performance, revealing both achievements and critical limitations of thin-scintillator architectures.

Performance Achievements

Experimental measurements establish 97.68 ± 0.07 % veto efficiency for events depositing ≥ 10 MeV, closely matching simulation predictions (97.25 ± 0.15 %). This agreement validates GEANT4 particle transport modeling across the cosmic ray spectrum. The veto eliminates the dominant 29 MeV muon peak, reducing event rates from ~ 0.85 Hz to ~ 0.02 Hz—demonstrating effective suppression of the primary background component.

Critical Limitations

The ~ 98 % efficiency falls substantially short of the >99.99 % required for Phase B searches. Differential efficiency degrades in the 10 MeV to 20 MeV range—precisely where dark matter detection demands peak performance. The 2 % to 5 % residual leak-through represents the primary systematic limitation. At ~ 0.02 Hz, this yields $\sim 600,000$ unvetoes events annually—orders of magnitude above anticipated signal rates.

Physical Mechanisms

Gamma rays dominate leak-through events. The 20 mm EJ-200 panels detect charged particles with 100 % efficiency but offer only 7.5 % probability for gamma interac-

tions producing detectable photons—versus 55.7 % (Figure 3.5) in PbF_2 . This order-of-magnitude differential in photon-generating gamma interactions explains the gamma dominance in residual backgrounds.

Secondary production in the inner 10 mm lead layers creates geometric blind spots. Electromagnetic cascades in these absorbers generate particles that deposit energy in the calorimeter without traversing sufficient veto material. The average leak-through involves 1.72 particles; complete absence of primary electrons confirms that charged particles reliably trigger detection.

Optimization Pathways

Achieving >99.99 % efficiency requires addressing complementary failure modes: primary gamma penetration and secondary production in blind spots. Adding outer veto layers alone provides modest improvement—substantial enhancement demands thick passive absorbers (≈ 8 cm lead based on attenuation coefficients). Optimal absorber placement remains under investigation. Alternatively, the outer crystals of the active calorimeter could serve as active vetos, leveraging their strong gamma detection capabilities at the cost of reduced detection volume.

These limitations establish clear development priorities for Phase B:

- Enhance passive shielding geometry through simulation optimization
- Develop ceiling-targeted active veto exploiting vertical cosmic ray dominance
- Validate large-panel light collection systematically for scaled implementations

The validated simulation framework enables reliable performance predictions guiding these optimization efforts.

Chapter 6

Summary and Future Directions

The dark matter problem remains unsolved. Despite compelling gravitational evidence across cosmic scales, the particle nature of dark matter continues to elude detection. Traditional WIMP searches have systematically excluded much of the theoretically favored parameter space for heavy dark matter candidates, motivating physicists to explore lighter alternatives. Dark photon mediators coupling to sub-GeV dark matter particles offer a compelling target for beam dump experiments, where high-intensity electron beams can produce dark sector particles through kinetic mixing with ordinary photons.

This thesis established the technical foundation for DarkMESA's contribution to this search. The work addressed three interlocking challenges: achieving efficient detection at 10 MeV threshold, suppressing spurious photodetector backgrounds, and rejecting cosmic ray contamination. The findings demonstrate that Cherenkov calorimetry with PbF_2 crystals provides a viable detection technology, while simultaneously identifying specific limitations that must be overcome for competitive sensitivity.

6.1 Detection Efficiency: Challenge Resolved

The fundamental question underlying Challenge I was whether PbF_2 Cherenkov radiators produce sufficient light for reliable 10 MeV detection. The answer is unambiguously affirmative.

The GEANT4 simulation framework developed in Chapter 3 characterized the complete detection chain: particle transport, energy deposition, Cherenkov photon production, optical propagation through wrapped crystals, and photodetector response. Validation

against cosmic ray measurements confirmed simulation accuracy, with measured and predicted photoelectron spectra agreeing within statistical uncertainties.

At the target 10 MeV threshold, PbF_2 crystals yield approximately 19 PE mean signal with single-sided PMT readout—demonstrating that light yield is not the limiting factor. However, single-sided PMT readout proves operationally unfeasible due to Hz-level dark count rates (Chapter 4). The SiPM array configuration that solves the dark count problem also provides adequate light collection: fourfold coincidence among 6–8 SiPMs per crystal maintains >90 % detection efficiency at 10 MeV and practical threshold (5 PE to 6 PE). The simulation framework further confirms selective sensitivity: electron detection efficiency exceeds 90 %, while neutron response remains below 10 %, preserving the hadronic insensitivity essential for background rejection.

The validated simulation tools enable reliable performance predictions for detector configurations not yet constructed, providing the quantitative foundation for Phase B optimization studies.

6.2 Spurious Background Suppression: Challenge Resolved

Challenge II arose from a practical discovery: PMT dark count rates at operational thresholds produce spurious triggers incompatible with rare event detection. Single-PMT readout generates 0.5 Hz to 4 Hz of background triggers per crystal. This overwhelms the few-events-per-year signal rates by six to seven orders of magnitude.

Chapter 4 systematically evaluated solutions spanning dual-sided PMT readout, coupled crystal arrays, and silicon photomultiplier arrays. The analysis yields a clear recommendation: SiPM arrays with fourfold coincidence provide the optimal balance of background suppression, detection efficiency, and cost.

Dual-sided PMT readout achieves adequate performance for the 25-crystal Phase A detector through twofold coincidence. However, scaling to ~2,000 Phase B crystals accumulates unacceptable accidental backgrounds approaching 1,000 events annually, while photodetector procurement costs reach approximately 1 M€.

Coupled crystal arrays failed through compounded limitations. Light sharing across extended optical volumes introduced severe position-dependent response variations

($CV \approx 93\%$ versus the 45% Poisson limit), preventing viable threshold operation regardless of coincidence configuration.

Silicon photomultiplier arrays overcome both limitations. Single-crystal multi-SiPM geometry achieves $CV \approx 42\%$, matching the intrinsic Poisson limit and eliminating systematic heterogeneity. Fourfold coincidence with 8 SiPMs per crystal maintains 90% detection efficiency at 5 PE threshold while suppressing accidental rates below 100 events annually at 25°C —potentially eliminating active cooling requirements. The cost advantage proves equally decisive: Phase B photodetector procurement reduces to approximately 150 k€, a factor of six savings compared to dual-sided PMT approaches.

6.3 Cosmic Ray Rejection: Challenge Partially Resolved

Challenge III proves the most demanding. Competitive sensitivity requires cosmic ray rejection exceeding 99.99% ; the Phase A inner veto achieves 97.7% . This substantial shortfall in residual background rate represents the primary limitation on currently achievable sensitivity.

Chapter 5 identified the physical mechanisms responsible. Charged particle rejection approaches theoretical limits: muon and electron detection efficiency exceeds 99% across the relevant energy range. The dominant failures arise from neutral particles: gamma rays that penetrate scintillator without ionizing, and secondary photons generated in shielding or structural materials near the calorimeter boundary. These neutral-particle leakage events deposit energy in the calorimeter through Compton scattering or pair production while evading veto detection entirely.

The measured efficiency validates the GEANT4 simulation framework, establishing confidence in simulation-guided optimization. However, the results demonstrate that the current two-layer scintillator architecture cannot achieve the rejection required for competitive dark photon sensitivity, regardless of further incremental improvements.

6.4 Implications for DarkMESA Sensitivity

The three-challenge framework enables a clear-eyed assessment of DarkMESA's prospects. Two challenges admit complete solutions with identified technologies; the third requires architectural changes beyond current Phase A implementation.

Detection efficiency poses no fundamental obstacle. PbF_2 light yield, simulation-validated at 10 MeV threshold, provides comfortable operational margins. Spurious background suppression is achievable through SiPM array readout with fourfold coincidence, at substantially reduced cost compared to PMT-based alternatives. These are solved problems awaiting implementation.

Cosmic ray rejection remains the sensitivity-limiting factor. The $\sim 98\%$ efficiency achieved with current veto architecture leaves residual backgrounds approximately two orders of magnitude higher than required for competitive parameter space coverage. Incremental improvements to the existing design—better optical coupling, reduced thresholds, additional PMT coverage—cannot bridge this gap. The physics of neutral particle penetration demands architectural solutions: enhanced passive absorbers to attenuate gamma rays before they reach the calorimeter, extended veto coverage to capture secondary production, or hybrid approaches combining active and passive elements.

The sensitivity scaling relation clarifies development priorities. Upper limits on signal event counts scale as $\sqrt{N_{\text{bg}}}$, while coupling reach scales as $N_{\text{bg}}^{1/4}$ since signal rates depend on ε^4 . Reducing residual backgrounds by a factor of 100 improves coupling reach by $100^{1/4} \approx 3$ —equivalent to the improvement from tripling beam current or runtime. No other single improvement offers comparable physics impact per development effort. Veto enhancement is therefore the critical path for achieving DarkMESA's physics objectives.

6.5 Phase B Development Priorities

The findings of this thesis establish a prioritized development roadmap for Phase B implementation:

Critical Path: Veto System Enhancement. Achieving competitive sensitivity requires veto architectures beyond current two-layer scintillator designs. Simulation studies should evaluate enhanced passive absorber configurations, extended geometric coverage addressing boundary effects, and hybrid active-passive approaches. Prototype testing of promising configurations should precede full-scale implementation. This development path determines whether DarkMESA achieves discovery-level sensitivity or remains limited to incremental parameter space exclusion.

High Priority: SiPM Array Implementation. The 8-SiPM fourfold coincidence configuration recommended in Chapter 4 requires engineering development before production: readout board design optimizing signal routing and thermal management, on-board coincidence logic implementation, and validation testing under realistic thermal conditions. Beam calibration at MAMI should verify simulation predictions for absolute photoelectron yields and position-dependent response uniformity.

6.6 Concluding Perspective

The search for dark matter particle interactions has spanned decades without definitive detection. Traditional WIMP searches, despite extraordinary sensitivity improvements, have progressively constrained rather than discovered. This sustained null result motivates diversified experimental approaches probing alternative mass ranges and interaction mechanisms.

DarkMESA contributes to this diversification by targeting the sub-GeV mass range through beam dump production and Cherenkov calorimetry. This thesis demonstrates that the core detection technology is viable: PbF_2 crystals achieve the required efficiency, and SiPM arrays provide practical spurious background suppression. The identified limitation—cosmic ray rejection with current veto architectures—is substantial but bounded; enhanced veto designs offer a clear path toward competitive sensitivity.

Whether dark matter particles reside in DarkMESA's accessible parameter space remains unknown. The experiment may achieve detection, establishing a dark photon mass and coupling; alternatively, it may exclude additional parameter space, constraining theoretical models and guiding future searches. Either outcome advances understanding. The value of this thesis lies in realistically assessing what Cherenkov calorimetry with plastic scintillator veto can achieve—and equally important, identifying limitations that require technological advancement. Only through rigorous characterization of both capabilities and constraints can the experimental program chart productive paths forward.

The dark matter problem will likely extend beyond any single experiment's timeline. DarkMESA represents one contribution within a multi-generational scientific endeavor, building incremental understanding through careful measurement and honest assessment of technical limitations. This thesis provides the foundation for that contribution.

Appendices

A. Statistical Framework for Beam Dump Dark Matter Searches

This appendix develops the statistical framework underlying the sensitivity analysis in Section 2.7. The treatment proceeds in two stages.

Section A.1 derives the on/off measurement formalism. The fundamental scaling relationships— $\sqrt{N_{\text{bg}}}$ fluctuations and the $(1 + 1/\tau)$ timing factor—emerge directly from Poisson counting statistics. For large event counts, treating the signal estimator as Gaussian-distributed enables closed-form expressions for discovery thresholds and exclusion limits. Systematic uncertainties enter through quadrature addition, yielding expressions suitable for rapid experimental design calculations. These approximations require $N_{\text{bg}} \gtrsim 30$ and assume systematics behave as uncorrelated Gaussian fluctuations on total counts.

Section A.3 presents the likelihood-based framework required for rigorous data analysis. This exact numerical approach remains valid at any event count, properly handles energy-dependent and correlated systematic uncertainties through nuisance parameters, and prevents spurious exclusions through the CLs method. Binned likelihood construction naturally exploits spectral differences between signal and background, weighting energy regions by their discrimination power rather than treating all events equally. The framework recovers the same scaling relationships in the large-sample limit while providing the statistical machinery for publication-quality results.

Together, these approaches serve complementary purposes: closed-form approximations for design and intuition, likelihood methods for data analysis and final results.

A.1 On/Off Measurement Formalism

Beam dump experiments searching for dark matter employ a fundamental measurement strategy exploiting the distinction between beam-on and beam-off periods. During beam-off periods, the detector observes only background processes—cosmic rays, radioactive decays, detector noise. During beam-on periods, these backgrounds

persist while potential dark matter signals add to the observed rate. Comparing these measurements isolates the beam-correlated signal component.

Statistical Model

Let T_{on} and T_{off} denote beam-on and beam-off measurement periods respectively. Assuming constant background rate λ_{bg} (events per unit time), the expected event counts follow Poisson distributions¹:

$$N_{\text{off}} \sim \text{Poisson}(\lambda_{\text{bg}} T_{\text{off}}) \quad (\text{A.1})$$

$$N_{\text{on}} \sim \text{Poisson}(\lambda_{\text{bg}} T_{\text{on}} + N_{\text{signal}}) \quad (\text{A.2})$$

where N_{signal} represents the expected signal count during beam-on operation. The beam-off measurement contains no signal contribution because dark matter production requires the electron beam.

The timing ratio is defined as $\tau = T_{\text{off}}/T_{\text{on}}$ and the expected background during beam-on as $N_{\text{bg}} = \lambda_{\text{bg}} T_{\text{on}}$. The beam-off expectation becomes τN_{bg} .

Signal Estimation and Variance

The signal estimate derives from subtracting scaled beam-off counts from beam-on counts:

$$\hat{S} = N_{\text{on}} - \frac{N_{\text{off}}}{\tau} \quad (\text{A.3})$$

The factor $1/\tau$ scales the beam-off measurement to the beam-on exposure. This estimator is unbiased: $\mathbb{E}[\hat{S}] = N_{\text{signal}}$.

The variance of this estimator follows from error propagation. Since N_{on} and N_{off} are independent Poisson random variables:

$$\begin{aligned} \text{Var}[\hat{S}] &= \text{Var}[N_{\text{on}}] + \frac{1}{\tau^2} \text{Var}[N_{\text{off}}] \\ &= (N_{\text{bg}} + N_{\text{signal}}) + \frac{1}{\tau^2} (\tau N_{\text{bg}}) \\ &= N_{\text{bg}} \left(1 + \frac{1}{\tau} \right) + N_{\text{signal}} \end{aligned} \quad (\text{A.4})$$

¹The Poisson distribution gives the probability of observing k events given expectation value μ : $P(k | \mu) = \frac{\mu^k e^{-\mu}}{k!}$.

For the background-dominated regime where $N_{\text{signal}} \ll N_{\text{bg}}$, the variance simplifies to:

$$\text{Var}[\hat{S}] \approx N_{\text{bg}} \left(1 + \frac{1}{\tau} \right) \quad (\text{A.5})$$

This result exposes the statistical cost of background subtraction. With equal measurement periods ($\tau = 1$), the variance becomes $2N_{\text{bg}}$ —twice what it would be with perfectly known background. Extended beam-off measurements ($\tau > 1$) reduce this penalty: as $\tau \rightarrow \infty$, the variance approaches N_{bg} , recovering the perfect-knowledge limit.

Incorporating Systematic Uncertainties

Real experiments face systematic uncertainties beyond Poisson counting statistics. Sources include energy calibration stability, detection efficiency variations, background rate temporal fluctuations, and environmental effects. For simplified treatment, these combine into an effective fractional uncertainty σ_{sys} on background prediction.

Modeling systematic uncertainties as additional Gaussian variance yields total uncertainty:

$$\sigma_{\text{total}}^2 = N_{\text{bg}} \left(1 + \frac{1}{\tau} \right) + (\sigma_{\text{sys}} N_{\text{bg}})^2 \quad (\text{A.6})$$

or equivalently:

$$\sigma_{\text{total}} = \sqrt{N_{\text{bg}} \left(1 + \frac{1}{\tau} \right) + (\sigma_{\text{sys}} N_{\text{bg}})^2} \quad (\text{A.7})$$

This treatment assumes that systematic effects behave as uncorrelated Gaussian fluctuations—a simplification, since real systematic uncertainties may exhibit correlations, non-Gaussian distributions, and energy-dependent structure. The likelihood framework (Section A.3) provides more rigorous treatment through nuisance parameters.

Discovery Threshold

The five-sigma discovery requirement demands signal significance exceeding five standard deviations:

$$\frac{N_{\text{signal}}}{\sigma_{\text{total}}} \geq 5 \quad (\text{A.8})$$

Rearranging yields the minimum detectable signal:

$$N_{\text{signal}}^{\text{min}} = 5 \sqrt{N_{\text{bg}} \left(1 + \frac{1}{\tau} \right) + (\sigma_{\text{sys}} N_{\text{bg}})^2} \quad (\text{A.9})$$

For balanced timing ($\tau = 1$) and negligible systematics, this reduces to $N_{\text{signal}}^{\text{min}} = 5\sqrt{2N_{\text{bg}}} \approx 7.1\sqrt{N_{\text{bg}}}$. At 10,000 background events, discovery requires at least 707 signal events—the result quoted in Section 2.7.

A.2 Exclusion Limit Methodology

Simplified Gaussian Treatment

When discovery thresholds exceed theoretical predictions, exclusion limits provide the alternative path to physics results. For large event counts where Poisson distributions approximate Gaussians, exclusion limits follow from straightforward hypothesis testing. The null hypothesis H_0 posits signal strength μ_s atop known background. The experiment excludes μ_s at confidence level $1 - \alpha$ if:

$$\mu_s > \hat{S} + z_\alpha \sigma_{\text{total}} \quad (\text{A.10})$$

where $\hat{S}$ is the observed signal estimate and z_α is the Gaussian quantile ($z_{0.10} = 1.282$ for 90 % CL).

Poisson Treatment for Low Counts

The Gaussian approximation fails for small event counts. In the idealized limit of zero observed events and zero background, pure Poisson statistics apply. The 90 % CL upper limit follows from:

$$P(\text{Poisson}(\mu_s) = 0) = e^{-\mu_s} = 0.10 \quad (\text{A.11})$$

yielding $\mu_s^{90\%} = -\ln(0.10) = 2.30$ events. This represents the theoretical sensitivity limit: observing zero events with zero background excludes all models predicting more than 2.3 events.

Worked Example: Gaussian Exclusion

Consider an experiment observing $N_{\text{on}} = 1,050$ events and $N_{\text{off}} = 1,000$ events with $\tau = 1$. The signal estimate becomes:

$$\hat{S} = 1,050 - \frac{1,000}{1} = 50 \text{ events}$$

With inferred background $N_{\text{bg}} = N_{\text{off}}/\tau = 1,000$ events and 10 % systematic uncertainty:

$$\sigma_{\text{total}} = \sqrt{1,000 \times 2 + (0.1 \times 1,000)^2} = \sqrt{2,000 + 10,000} \approx 110 \text{ events}$$

Discovery significance reaches only $50/110 \approx 0.45\sigma$ —far below five-sigma. However, the 90 % CL exclusion threshold is:

$$\mu_s^{90\%} = 50 + 1.282 \times 110 \approx 191 \text{ events}$$

Any theoretical model predicting more than 191 signal events is excluded at 90 % confidence. A dark photon model predicting 300 events would be decisively excluded, despite the complete absence of discovery sensitivity.

Sensitivity Scaling Relationships

The standard deviation formula (Equation A.7) directly yields the sensitivity scaling relationships employed throughout Section 2.7. In the Gaussian approximation, the 90 % CL upper limit on signal event counts is:

$$\text{Upper Limit} = \hat{S} + 1.282 \times \sigma_{\text{total}} \quad (\text{A.12})$$

Hence the upper limit scales with the total uncertainty:

$$\text{Upper Limit} \propto \sigma_{\text{total}} = \sqrt{N_{\text{bg}} \left(1 + \frac{1}{\tau} \right) + (\sigma_{\text{sys}} N_{\text{bg}})^2} \quad (\text{A.13})$$

In the limit of perfect background knowledge ($\tau \rightarrow \infty$) and negligible systematic uncertainties ($\sigma_{\text{sys}} \rightarrow 0$), this reduces to the characteristic scaling:

$$\text{Upper Limit} \propto \sqrt{N_{\text{bg}}} \quad (\text{A.14})$$

This is the $\sqrt{N_{\text{bg}}}$ scaling referenced in Section 2.7. For finite τ and nonzero systematics, the full expression (Equation A.13) applies.

The upper limit on event counts translates to coupling reach through the signal rate dependence. Since dark photon production and detection both scale with ε^2 , signal rates scale as ε^4 , giving $y \propto \varepsilon^2 \propto \sqrt{\text{UL}}$. The coupling reach therefore degrades as $N_{\text{bg}}^{1/4}$. Table 2.2 presents numerical values for the y -degradation factor under idealized assumptions.

A.3 Likelihood Framework

The Gaussian approximations developed above provide closed-form expressions useful for experimental design, but they exhibit limitations for rigorous data analysis. Most problematically, downward background fluctuations can produce spuriously strong exclusions: an experiment expecting 1,000 events that observes only 980 would set limits on models it has no genuine sensitivity to probe. The Gaussian framework provides no protection against such statistical accidents.

The treatment of systematic uncertainties also differs fundamentally between the two approaches. Adding uncertainties in quadrature assumes they are Gaussian, uncorrelated, and affect only total counts. Real systematics may be correlated across energy bins or modify the signal-to-background ratio in ways that a single effective parameter cannot capture.

The likelihood framework addresses these limitations. The CLs method prevents spurious exclusions by ensuring that limits reflect genuine experimental sensitivity rather than statistical fluctuations. Systematic uncertainties enter as nuisance parameters with explicit physical meaning, properly propagating their effects through the analysis. Binned likelihood construction naturally exploits spectral differences between signal and background, weighting energy regions by their discrimination power rather than treating all events equally.

The following subsections develop this formalism: likelihood construction for the on/off measurement, extension to multiple energy bins, nuisance parameter treatment for systematic uncertainties, and profile likelihood ratios for hypothesis testing.

Likelihood Formulation of On/Off Measurements

The on/off measurement formalism derived above translates directly into a likelihood function. During beam-on periods, the detector observes signal plus background; during beam-off periods, only background contributes. Assuming Poisson-distributed counts, the combined likelihood becomes:

$$\mathcal{L}(\mu_s, \mu_b) = P(n_{\text{on}}|\mu_s + \mu_b) \cdot P(n_{\text{off}}|\mu_b \cdot \tau) \quad (\text{A.15})$$

where n_{on} and n_{off} are observed counts, μ_s is the signal expectation, μ_b is the background rate during beam-on, and $\tau = T_{\text{off}}/T_{\text{on}}$ is the timing ratio. The beam-off term constrains $\mu_b \approx n_{\text{off}}/\tau$ with Poisson uncertainty, enabling signal inference without external background modeling.

Worked Example: Likelihood-Based Background Constraint

Consider $n_{\text{on}} = 15$ beam-on events and $n_{\text{off}} = 10$ beam-off events with $\tau = 1$. The likelihood is:

$$\mathcal{L}(\mu_s, \mu_b) = \text{Poisson}(15; \mu_s + \mu_b) \cdot \text{Poisson}(10; \mu_b)$$

Evaluating for different parameter combinations:

$$\mu_s = 0, \mu_b = 15 : \quad \mathcal{L} \approx 1.5 \times 10^{-6}$$

$$\mu_s = 5, \mu_b = 10 : \quad \mathcal{L} \approx 1.3 \times 10^{-2}$$

$$\mu_s = 15, \mu_b = 1 : \quad \mathcal{L} \approx 1.0 \times 10^{-8}$$

The beam-off measurement strongly favors $\mu_b \approx 10$, making $(\mu_s = 5, \mu_b = 10)$ the most likely scenario. The likelihood automatically penalizes background values inconsistent with beam-off data—no explicit background subtraction required.

This single-bin formulation has limited discriminating power: any excess in beam-on counts is attributed to potential signal, regardless of the energy distribution of those events. A five-event excess from dark matter scattering and a five-event excess from an upward background fluctuation are statistically indistinguishable. Since dark matter signals and cosmic ray backgrounds exhibit distinct energy spectra, partitioning the data into energy bins recovers this lost information.

Binned Likelihood Construction

Partitioning the energy range into N_{bins} bins, the likelihood generalizes to a product over bins:

$$\mathcal{L}(\mu_s, \{\mu_{b,j}\}) = \prod_{j=1}^{N_{\text{bins}}} P(n_{\text{on},j} | \lambda_{\text{on},j}) \cdot P(n_{\text{off},j} | \mu_{b,j} \cdot \tau) \quad (\text{A.16})$$

where $n_{\text{on},j}$ and $n_{\text{off},j}$ denote observed counts in bin j . The expected beam-on counts combine signal and background:

$$\lambda_{\text{on},j} = \mu_s \cdot C \cdot \alpha_j + \mu_{b,j} \quad (\text{A.17})$$

Here μ_s is the total signal strength (the parameter of interest), C is the overall signal acceptance, α_j is the fraction of accepted signal falling in bin j (with $\sum_j \alpha_j = 1$), and $\mu_{b,j}$ is the background rate in bin j .

The signal shape parameters $\{C, \alpha_j\}$ come from Monte Carlo simulation of dark matter production, propagation, and detection. The background shape requires no theoretical input: each per-bin background rate $\mu_{b,j}$ is constrained by its corresponding beam-off measurement through $P(n_{\text{off},j} | \mu_{b,j} \cdot \tau)$. This data-driven approach propagates statistical uncertainties automatically to signal inference.

The formulation above treats C and $\{\alpha_j\}$ as fixed quantities determined by simulation. In reality, systematic uncertainties affect these parameters. The likelihood framework accommodates such effects through *nuisance parameters*: quantities that influence the measurement but are not themselves of primary interest.

Nuisance Parameters and Systematic Uncertainties

The binned likelihood already contains nuisance parameters: the per-bin background rates $\{\mu_{b,j}\}$ are not of primary interest but must be determined to infer signal strength. The beam-off measurements constrain these parameters through the likelihood itself—no external input required.

Systematic uncertainties introduce additional nuisance parameters that are constrained by external measurements. Energy calibration exemplifies this treatment. Let g represent the ratio of reconstructed to true energy: $E_{\text{reco}} = g \cdot E_{\text{true}}$. Miscalibration affects the analysis in two ways:

1. **Bin migration:** Events shift between energy bins. A true 40 MeV event near a bin boundary falls in different bins depending on whether $g = 0.95$ (reconstructed as 38 MeV) or $g = 1.05$ (reconstructed as 42 MeV). The fixed bin fractions $\{\alpha_j\}$ become calibration-dependent functions $\{\alpha_j(g)\}$.
2. **Acceptance variation:** Events migrate across the energy threshold, changing overall acceptance:

$$C(g) = \frac{N_{\text{accepted}}(g)}{N_{\text{generated}}} \quad (\text{A.18})$$

Both $C(g)$ and $\{\alpha_j(g)\}$ are determined through Monte Carlo studies: signal events are generated, detector response is scaled by various g values, and the resulting acceptance and bin fractions are parameterized as smooth functions.

External calibration measurements determine the energy scale to some precision σ_g . This constraint enters the likelihood as a Gaussian penalty term:

$$\begin{aligned} \mathcal{L}(\mu_s, \{\mu_{b,j}\}, g) = & \prod_j \underbrace{P(n_{\text{on},j} | \mu_s \cdot C(g) \cdot \alpha_j(g) + \mu_{b,j})}_{\text{beam-on measurement}} \cdot \underbrace{P(n_{\text{off},j} | \mu_{b,j} \cdot \tau)}_{\text{background constraint}} \\ & \cdot \underbrace{\mathcal{N}(g; 1, \sigma_g)}_{\text{calibration constraint}} \end{aligned} \quad (\text{A.19})$$

The final term $\mathcal{N}(g; 1, \sigma_g)$ penalizes calibration values deviating from the measured value $g = 1$.

This structure reveals a key advantage over the Gaussian quadrature approach: systematic effects propagate through their physical mechanism—bin migration and acceptance change—rather than as an effective fractional uncertainty on total counts. Multiple systematic sources can be included as separate nuisance parameters, with correlations explicitly modeled where they exist. The likelihood optimization simultaneously balances all constraints: beam-on data, beam-off data, and external calibration measurements.

Recovery of Statistical Scaling

Despite the added complexity, the likelihood framework recovers the Gaussian scaling relationships in the appropriate limit. This correspondence emerges from a fundamental result in statistical theory: for large sample sizes, maximum likelihood estimators become approximately Gaussian-distributed, with variance determined by the curvature of the likelihood function at its maximum.

For Poisson-distributed counts with mean λ , this asymptotic variance equals λ itself—exactly matching the Poisson property $\text{Var}[N] = \mathbb{E}[N]$ that underlies the $\sqrt{N_{\text{bg}}}$ scaling. The on/off likelihood structure (Equation A.15) combines beam-on variance $\sim N_{\text{bg}}$ with scaled beam-off variance $\sim N_{\text{bg}}/\tau$, yielding total variance $N_{\text{bg}}(1 + 1/\tau)$ —the same result derived through error propagation in Section A.1.

The Gaussian approximations thus provide correct scaling for experimental design, while the likelihood framework adds three capabilities essential for rigorous analysis: exact Poisson statistics valid at any event count, proper handling of physical boundaries such as $\mu_s \geq 0$, and systematic propagation through explicit nuisance parameters rather than effective fractional uncertainties. Section A.6 validates this correspondence through Monte Carlo studies. Beyond these technical improvements, the likelihood framework provides a principled foundation for hypothesis testing—translating likelihood values into exclusion statements while protecting against statistical pathologies that can afflict simpler approaches.

A.4 Hypothesis Testing and Upper Limit

The likelihood framework provides the foundation; hypothesis testing translates likelihood values into exclusion statements. This section develops the profile likelihood ratio test statistic and the CLs method that prevents spurious exclusions.

Profile Likelihood Ratio

For testing a hypothesized signal strength μ_s , the profile likelihood ratio compares the best achievable likelihood at that signal strength to the global maximum:

$$\lambda(\mu_s) = \frac{\mathcal{L}(\mu_s, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}_s, \hat{\theta})} \quad (\text{A.20})$$

where $\hat{\mu}_s$ and $\hat{\theta}$ maximize the likelihood globally (denominator), while $\hat{\hat{\theta}}$ maximizes the likelihood with μ_s fixed (numerator). The double-hat notation indicates conditional maximization. The ratio quantifies how much worse the best fit becomes when forced to a specific signal strength.

For computational convenience, analyses employ the logarithmic test statistic:

$$q(\mu_s) = \begin{cases} -2 \log \lambda(\mu_s) & \text{if } \mu_s \geq \hat{\mu}_s \\ 0 & \text{if } \mu_s < \hat{\mu}_s \end{cases} \quad (\text{A.21})$$

The one-sided construction reflects that signal hypotheses below the best-fit value provide no exclusion information. Larger $q(\mu_s)$ values indicate greater incompatibility between hypothesis and data.

The CLs Method

Classical frequentist confidence intervals suffer from a fundamental pathology in low-count experiments. Consider an experiment predicting 100 cosmic ray background events that observes only 85 events due to random Poisson fluctuation. A naive frequentist analysis might produce impressively strong exclusion limits based on this observation.

However, the experiment has not actually become more sensitive to signals. The apparent improvement arises entirely from a fortuitous background fluctuation, not from enhanced discriminating power between signal and background hypotheses. If repeated many times under identical conditions, roughly half of such experiments would produce weaker limits due to upward background fluctuations, while only occasionally would downward fluctuations yield such apparently strong results.

This statistical accident—conflating fortuitous background fluctuations with genuine experimental sensitivity—compromises the scientific integrity of hypothesis testing. The CLs method protects against spurious exclusions by accounting for experimental sensitivity. The key idea is illustrated in Figure A.1.

The test statistic $q(\mu_s)$ (Equation A.21) quantifies how well a hypothesized signal strength μ_s fits the observed data. Specifically, it measures how much the likelihood degrades when forced to assume signal strength μ_s compared to the best-fit value:

- Small q : the hypothesis μ_s fits the data well
- Large q : the hypothesis μ_s fits the data poorly

Setting exclusion limits requires knowing what q values to expect under different scenarios. If repeated experiments were performed:

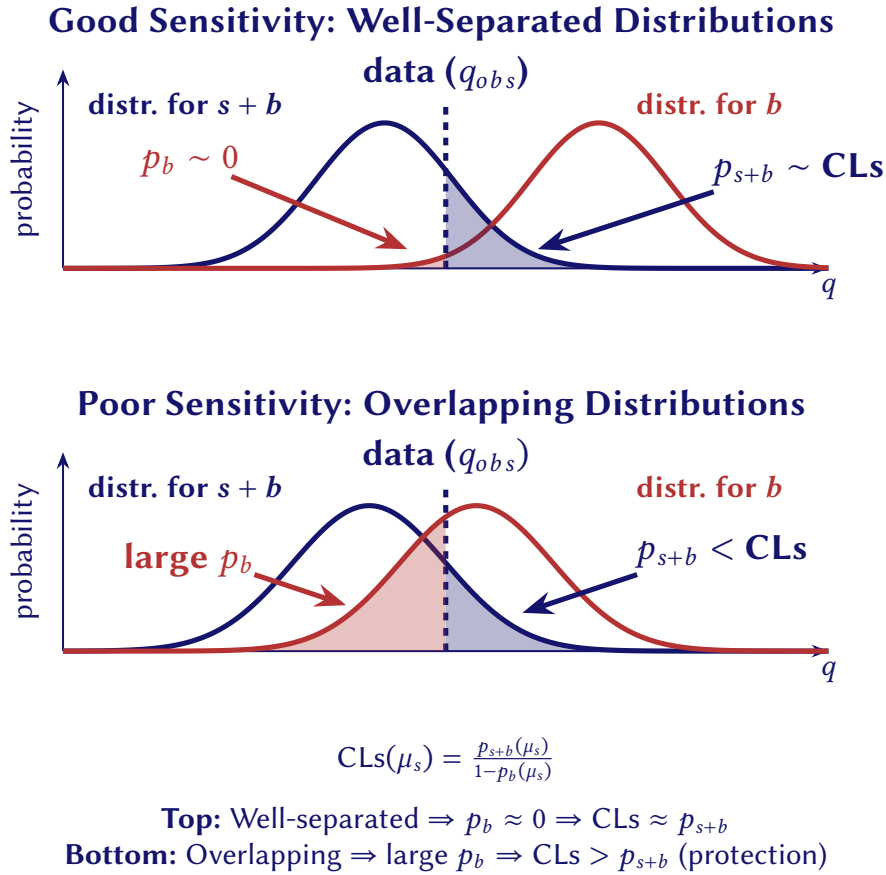


Figure A.1: The CLs method for well-separated distributions (top, good sensitivity) and overlapping distributions (bottom, poor sensitivity). Dark blue curves show expected test statistic distributions under the signal-plus-background hypothesis; red curves under background-only. Small q indicates the signal hypothesis fits well; large q indicates poor fit. When distributions are well-separated, $p_b \approx 0$ and $\text{CLs} \approx p_{s+b}$. When distributions overlap, large p_b inflates CLs, preventing exclusion claims the experiment cannot support.

- Under signal-plus-background: the signal hypothesis should fit well, producing small q values (dark blue distributions on the left in Figure A.1)
- Under background-only: the signal hypothesis should fit poorly, producing large q values (red distributions on the right)

The observed data yield a specific value q_{obs} , shown as the vertical dashed line.

Good sensitivity (top panel): The two distributions are well-separated—the experiment can clearly distinguish between the presence and absence of signal. Two p-values

follow from the shaded areas:

$$p_{s+b} = \text{fraction of } s + b \text{ distribution with } q \geq q_{\text{obs}} \text{ (blue shading)} \quad (\text{A.22})$$

$$p_b = \text{fraction of } b \text{ distribution with } q \leq q_{\text{obs}} \text{ (red shading)} \quad (\text{A.23})$$

When distributions are well-separated, $p_b \approx 0$: under the background-only hypothesis, observing q values as small as q_{obs} would be rare.

Poor sensitivity (bottom panel): The distributions overlap substantially—even without signal, statistical fluctuations often produce small q values that mimic the presence of signal. Now p_b is large: the experiment cannot reliably distinguish the two hypotheses. The CLs statistic accounts for this discriminating power:

$$\text{CLs}(\mu_s) = \frac{p_{s+b}(\mu_s)}{1 - p_b(\mu_s)} \quad (\text{A.24})$$

When sensitivity is good ($p_b \approx 0$), the denominator approaches unity and $\text{CLs} \approx p_{s+b}$ —standard frequentist behavior. When sensitivity is poor (p_b large), the denominator shrinks, inflating CLs and preventing exclusion. The correction factor $(1 - p_b)^{-1}$ penalizes experiments that cannot demonstrate they would have seen a difference had signal been present.

Signal strength μ_s is excluded at 90 % confidence level when $\text{CLs}(\mu_s) < 0.10$.

The CLs method was developed for LEP Higgs searches and has become standard in high-energy physics. Its conservative nature ensures that exclusion claims reflect genuine experimental sensitivity rather than statistical fluctuations.

Implementation

Practical CLs calculation proceeds through Monte Carlo pseudo-experiments, as illustrated in Figure A.2:

1. For hypothesis μ_s , generate ensemble of pseudo-experiments under signal-plus-background assumption
2. Generate separate ensemble under background-only assumption
3. Compute test statistic q for each pseudo-experiment
4. Construct distributions and compute p-values from observed q_{obs}
5. Combine into CLs statistic

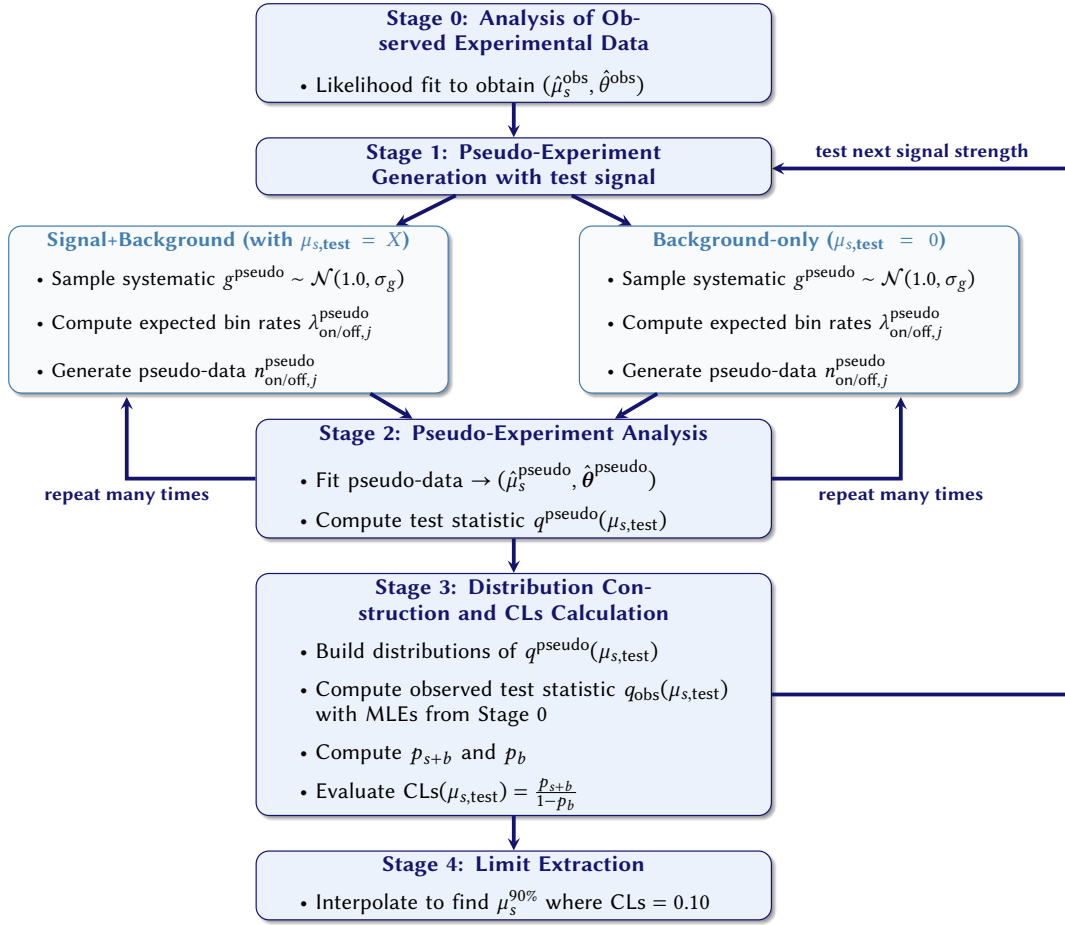


Figure A.2: Monte Carlo implementation workflow for CLs upper limit determination in beam dump dark matter searches. Stage 0 extracts background parameters from data through maximum likelihood fitting. Stage 1 generates pseudo-experiment ensembles under competing hypotheses using fitted parameters. Stage 2 computes test statistics for each pseudo-experiment. Stage 3 constructs p-value distributions from the test statistic ensembles. Stage 4 determines the signal strength satisfying CLs = 0.10 through interpolation.

The RooStats framework provides standardized implementation of this procedure, handling the computational complexity of multi-bin likelihoods with multiple nuisance parameters. For each signal hypothesis μ_s , the framework generates required pseudo-experiment ensembles, performs profile likelihood fits to extract test statistics, and computes resulting p-values and CLs ratios.

A.5 Expected and Observed Limits

The CLs framework provides the machinery for computing exclusion limits; the experimental context determines how that machinery is applied.

Observed Limits

Observed limits derive from actual experimental data. A maximum likelihood fit to the beam-on and beam-off measurements yields the best-fit signal strength $\hat{\mu}_s^{\text{obs}}$, background parameters $\hat{\mu}_{b,j}^{\text{obs}}$, and nuisance parameters $\hat{\theta}^{\text{obs}}$. The observed test statistic $q_{\text{obs}}(\mu_s)$ follows from Equation A.21, comparing the likelihood at hypothesized signal strength μ_s to the likelihood at the best-fit $\hat{\mu}_s^{\text{obs}}$.

Computing CLs requires knowing the distribution of q under both hypotheses (Figure A.1). These distributions are constructed through pseudo-experiments using the fitted background as central values:

$$\lambda_{\text{on},j}^{\text{pseudo}} = \mu_{s,\text{test}} \cdot C(g^{\text{pseudo}}) \cdot \alpha_j(g^{\text{pseudo}}) + \hat{\mu}_{b,j}^{\text{obs}} \quad (\text{A.25})$$

$$\lambda_{\text{off},j}^{\text{pseudo}} = \hat{\mu}_{b,j}^{\text{obs}} \cdot \tau \quad (\text{A.26})$$

where $g^{\text{pseudo}} \sim \mathcal{N}(1.0, \sigma_g)$ samples the calibration parameter from its constraint distribution, and $\mu_{s,\text{test}}$ equals the hypothesized signal strength for the $s + b$ distribution or zero for the b -only distribution. The observed 90 % CL upper limit corresponds to the signal strength where $\text{CLs}(\mu_s^{90\%}) = 0.10$.

Expected Limits

Observed limits answer the question: what signal strengths does this specific dataset exclude? A single number, however, lacks context. Was this a typical result, or did statistical fortune—good or bad—play a significant role? Expected limits provide that context.

A thought experiment illustrates. If the same experiment ran many times under background-only conditions ($\mu_{s,\text{test}} = 0$), Poisson fluctuations would produce different counts each time, and each dataset would yield a different upper limit. Some runs would fluctuate downward, producing stronger (lower) limits; others would fluctuate upward, producing weaker (higher) limits. This distribution defines what “normal” results look like.

The expected limit—the median of this distribution—serves as a benchmark. Comparing the observed limit to this benchmark reveals whether the result matches background-only expectations or reflects something unexpected: a statistical fluctuation or a real signal.

Constructing expected limits requires generating pseudo-experiments under the background-only hypothesis. But what bin rates should these pseudo-experiments use? Before data collection, only theoretical predictions exist. Afterward, fitted background parameters offer an alternative. These two choices lead to distinct approaches, each serving different purposes.

A Priori Expected Limits (Asimov Dataset)

Before data collection, expected limits rely on theoretical background predictions, anticipated detector performance, and estimated systematic uncertainties. The Asimov approach constructs an idealized dataset where observed counts exactly equal background-only expectations:

$$n_{\text{on},j}^{\text{Asimov}} = \mu_{b,j}^{\text{theory}} \quad (\text{A.27})$$

$$n_{\text{off},j}^{\text{Asimov}} = \mu_{b,j}^{\text{theory}} \cdot \tau \quad (\text{A.28})$$

The CLs limit computed from this dataset represents median expected sensitivity. This approach answers: given current understanding of backgrounds and detector capabilities, what sensitivity should the experiment achieve? The result serves as the experimental promise—the sensitivity target before allocating beam time and resources.

A Posteriori Expected Limits

After data collection, expected limits can instead use fitted background parameters $\{\hat{\mu}_{b,j}^{\text{obs}}\}$ extracted from actual measurements. Generating many background-only pseudo-experiments with these parameters and computing the median limit answers a different question: given what was learned about the detector and backgrounds from this specific dataset, what limit would typically be obtained if no signal were present?

Uncertainty Bands

The distribution of expected limits follows from generating many pseudo-experiments, computing an upper limit for each, and extracting quantiles:

- Median (50 %): central expected limit
- $\pm 1\sigma$ band (16 %–84 %): typical fluctuation range (traditionally green in plots)
- $\pm 2\sigma$ band (2.5 %–97.5 %): extended fluctuation range (traditionally yellow in plots)

Comparing the observed limit against these bands provides diagnostic information. An observed limit significantly stronger than the -2σ band could indicate a downward background fluctuation—or, more intriguingly, an actual signal producing the observed excess. An observed limit significantly weaker than the $+2\sigma$ band suggests unexpected backgrounds or underestimated systematic effects. Results falling within the bands confirm consistency with background-only expectations.

A.6 Toy Sensitivity Study

The preceding sections developed the likelihood framework and established that it should recover Gaussian scaling relationships in the large-sample limit. This section validates that prediction through Monte Carlo studies and empirically characterizes how systematic uncertainties affect sensitivity beyond the statistical baseline.

The study employs a simplified experimental model capturing essential features of beam dump searches: a binned likelihood with on/off background constraint, energy-dependent signal and background spectra, and energy scale uncertainty as a representative systematic effect. While realistic experiments face additional complexity—multiple systematic sources, energy-dependent efficiencies, beam-correlated backgrounds—this pedagogical model isolates the fundamental scaling relationships and demonstrates how the likelihood framework handles systematic uncertainties.

Simulation Setup

Figure A.3 presents the signal and background energy distributions. The signal spectrum represents a characteristic dark photon electromagnetic shower profile from Monte Carlo simulation, entering the likelihood through the acceptance $C(g)$ and bin fractions $\alpha_j(g)$ developed in Section A.3. The background follows an exponential distribution typical of cosmic ray energy deposits that evade active veto systems—a functional form consistent with BDX-MINI measurements and DarkMESA veto studies (Chapter 5).

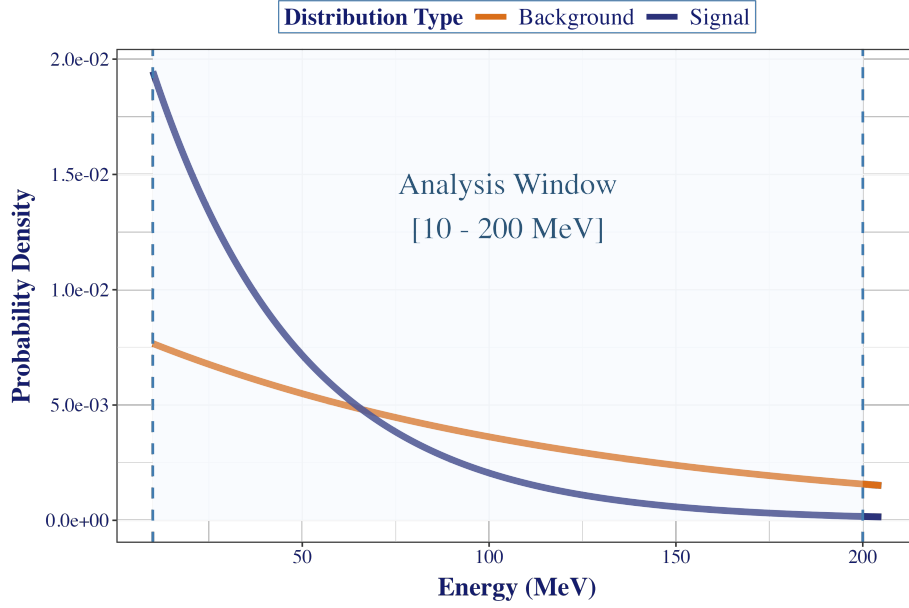


Figure A.3: Signal and background energy distributions for the toy sensitivity study. The signal spectrum (dark blue) represents a characteristic dark photon electromagnetic shower profile with acceptance C and bin fractions $\{\alpha_j\}$ from Monte Carlo simulation. The background spectrum (orange) follows an exponential cosmic ray distribution with spectral fractions $\{f_{b,j}\}$ constrained by beam-off measurements. The analysis window spans 10 MeV to 200 MeV.

The Asimov dataset construction sets observed counts equal to background-only expectations:

$$n_{\text{on},j}^{\text{Asimov}} = N_{\text{bg}} \cdot f_{b,j} \quad (\text{A.29})$$

$$n_{\text{off},j}^{\text{Asimov}} = N_{\text{bg}} \cdot f_{b,j} \cdot \tau \quad (\text{A.30})$$

where $f_{b,j}$ denotes the background spectral fraction in bin j . When testing signal hypotheses, the expected beam-on counts become:

$$\lambda_{\text{on},j}(\mu_s) = \mu_s \cdot C(g) \cdot \alpha_j(g) + N_{\text{bg}} \cdot f_{b,j} \quad (\text{A.31})$$

The signal shape $\{C, \alpha_j\}$ thus determines how additional events from a hypothesized signal would distribute across energy bins—information the likelihood exploits to distinguish signal from background fluctuations.

Asimov datasets eliminate Poisson fluctuations, enabling clean assessment of how experimental parameters affect discrimination power. The CLs limits computed from Asimov data represent median expected sensitivity: if the experiment were repeated

many times under identical conditions, approximately half the resulting limits would fall below this value.

The simulation explores:

- Background events: $N_{\text{bg}} \in \{1,000, 5,000, 10,000, 20,000, 30,000, 40,000\}$
- Timing ratios: $\tau \in \{0.2, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5\}$
- Energy scale uncertainty: $\sigma_g \in \{0, 10, 20, 30, 40, 50\} \%$
- Energy bins: $N_{\text{bins}} = 8$

The large σ_g range—extending to values unrealistic for a single systematic source—serves pedagogical purposes: by employing only energy scale uncertainty, larger values reveal systematic dependencies that would otherwise require combining multiple smaller sources. This yields 288 configurations, each undergoing complete CLs analysis using the RooStats framework.

Statistical Scaling Validation

The Gaussian treatment predicts upper limits scaling as:

$$\text{UL} \propto \sqrt{N_{\text{bg}}} \cdot \sqrt{1 + 1/\tau} \quad (\text{A.32})$$

Taking logarithms, plotting $\log(\text{UL})$ versus $\log[\sqrt{N_{\text{bg}}} \times \sqrt{1 + 1/\tau}]$ should yield unit slope if the scaling holds. Figure A.4 (top panel) confirms this prediction at fixed $\sigma_g = 10 \%$, with fitted slope approximately 1.00. The likelihood framework correctly recovers Gaussian scaling in the statistically-limited regime.

Systematic Effects

The bottom panels of Figure A.4 reveal how energy scale uncertainty degrades sensitivity beyond statistical fluctuations. Two features distinguish the likelihood treatment from Gaussian quadrature addition:

1. The degradation is approximately linear in σ_g , not quadratic as the Gaussian formula $(\sigma_{\text{sys}} N_{\text{bg}})^2$ would suggest.
2. The magnitude is modest: even at $\sigma_g = 50 \%$, additional degradation remains below 10 %.

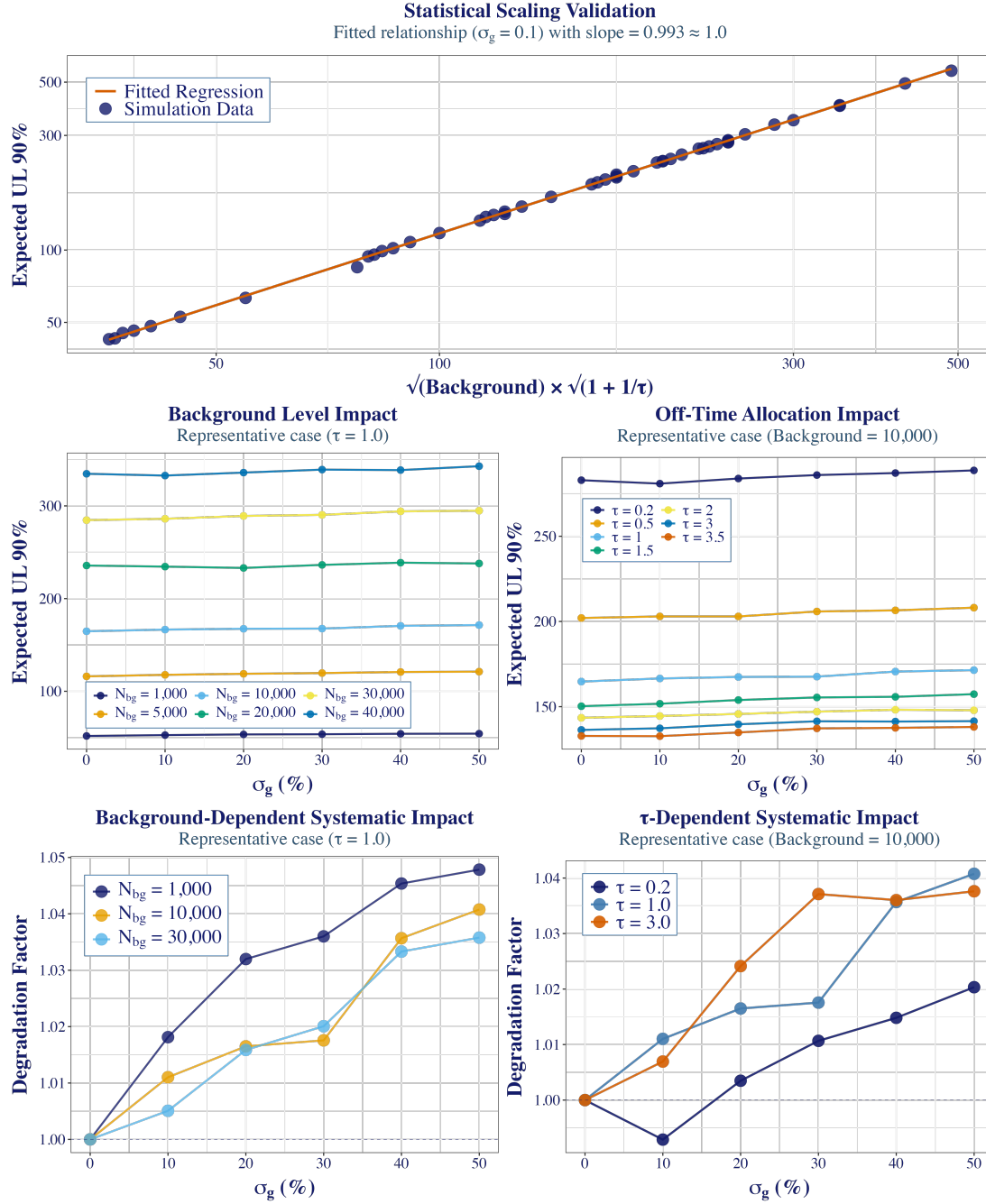


Figure A.4: Validation of theoretical scaling relationships. Top panel: combined scaling confirms $UL \propto \sqrt{N_{bg}} \times \sqrt{1 + 1/\tau}$ with fitted slope ≈ 1.0 at fixed $\sigma_g = 10\%$. Middle panels: upper limit dependence on background level and energy scale uncertainty. Bottom panels: systematic degradation factors relative to the $\sigma_g = 0$ baseline, revealing that degradation rates depend on both N_{bg} and τ —lower backgrounds and higher τ values exhibit greater systematic sensitivity as statistical precision improves.

This modest impact reflects how energy scale uncertainty enters binned likelihood analyses: when both signal and background spectra shift together under calibration changes, spectral shape differences providing discrimination power are largely preserved.

The degradation rates themselves depend on experimental parameters. Lower backgrounds exhibit steeper degradation slopes—with reduced Poisson fluctuations, systematic uncertainties constitute a larger fraction of total uncertainty. Similarly, higher τ values (better background constraint) increase systematic sensitivity as statistical precision improves. This coupling between statistical and systematic effects motivates the unified model developed in the following section.

Limitations

This toy study deliberately simplifies several aspects of realistic beam dump analyses:

- **Single systematic source:** Real experiments face multiple systematics—energy calibration, detection efficiency, geometric acceptance, veto performance—each contributing to total uncertainty. The single-source treatment here isolates scaling behavior but underestimates total systematic impact.
- **No beam-correlated backgrounds:** The on/off methodology perfectly constrains time-independent backgrounds. Beam-correlated backgrounds (neutrons, photons from beam-target interactions) require Monte Carlo prediction with associated modeling uncertainties not captured here.
- **Fixed background shape:** Higher backgrounds in practice often correspond to degraded veto performance, shifting spectra toward lower energies. The fixed-shape assumption isolates normalization effects but misses this correlation.

Despite these simplifications, the study validates the fundamental scaling relationships and demonstrates the likelihood framework’s systematic uncertainty treatment, providing foundation for the unified sensitivity model.

A.7 Unified Sensitivity Model

The Monte Carlo results enable constructing an interpolating model that predicts sensitivity across parameter space without repeated computation.

Model Structure

The unified model combines statistical scaling with empirically-determined systematic effects:

$$\text{UL}(N_{\text{bg}}, \tau, \sigma_g) = \text{UL}_0 \times \sqrt{\frac{N_{\text{bg}}}{N_{\text{ref}}}} \times \sqrt{1 + \frac{1}{\tau}} \times \left(1 + \alpha_{\text{eff}}(N_{\text{bg}}, \tau) \cdot \sigma_g\right) \quad (\text{A.33})$$

The first two factors encode the validated statistical scaling from Equation A.32. The final factor captures systematic degradation through an effective coefficient:

$$\alpha_{\text{eff}}(N_{\text{bg}}, \tau) = \alpha_0 + \alpha_1 \log\left(\frac{N_{\text{bg}}}{N_{\text{ref}}}\right) + \alpha_2 \log\left(1 + \frac{1}{\tau}\right) + \alpha_3 \log\left(\frac{N_{\text{bg}}}{N_{\text{ref}}}\right) \log\left(1 + \frac{1}{\tau}\right) \quad (\text{A.34})$$

The reference normalization $N_{\text{ref}} = 1,000$ centers the logarithmic terms for numerical stability. The interaction term α_3 captures coupling between background level and time allocation observed in the toy study.

Fitting to the 288 simulation points yields excellent model fidelity ($R^2 > 0.999$ on both training and held-out test data), confirming that the multiplicative structure with parameter-dependent systematic coefficient captures the essential physics.²

Translation to Coupling Reach

Since signal rates scale as ε^4 , the excludable coupling scales as $y \propto \varepsilon^2 \propto \text{UL}^{1/2}$. The y -degradation relative to zero background becomes:

$$y\text{-degradation}(N_{\text{bg}}, \tau, \sigma_g) = \sqrt{\frac{\text{UL}(N_{\text{bg}}, \tau, \sigma_g)}{2.3}} \quad (\text{A.35})$$

where 2.3 events is the zero-background Poisson limit at 90 % CL. Substituting the unified model:

$$y\text{-degradation} = C \times N_{\text{bg}}^{1/4} \times \left(1 + \frac{1}{\tau}\right)^{1/4} \times \sqrt{1 + \alpha_{\text{eff}}(N_{\text{bg}}, \tau) \cdot \sigma_g} \quad (\text{A.36})$$

where $C = \sqrt{\text{UL}_0 / (2.3 \cdot \sqrt{N_{\text{ref}}})}$ is a normalization constant. The statistical scaling— $N_{\text{bg}}^{1/4}$ and $(1 + 1/\tau)^{1/4}$ —matches the Gaussian predictions from Section 2.7.

²The fitted parameters are $\text{UL}_0 = 36.8$ events, $\alpha_0 = 1.57 \times 10^{-3}$, $\alpha_1 = -1.71 \times 10^{-4}$, $\alpha_2 = -9.35 \times 10^{-4}$, and $\alpha_3 = 1.84 \times 10^{-4}$, with reference normalization $N_{\text{ref}} = 1,000$.

Impact on DarkMESA Sensitivity

Figure A.5 shows the DarkMESA exclusion reach computed using the unified model for representative background levels, assuming $\tau = 1$ and $\sigma_g = 10\%$ energy scale uncertainty.

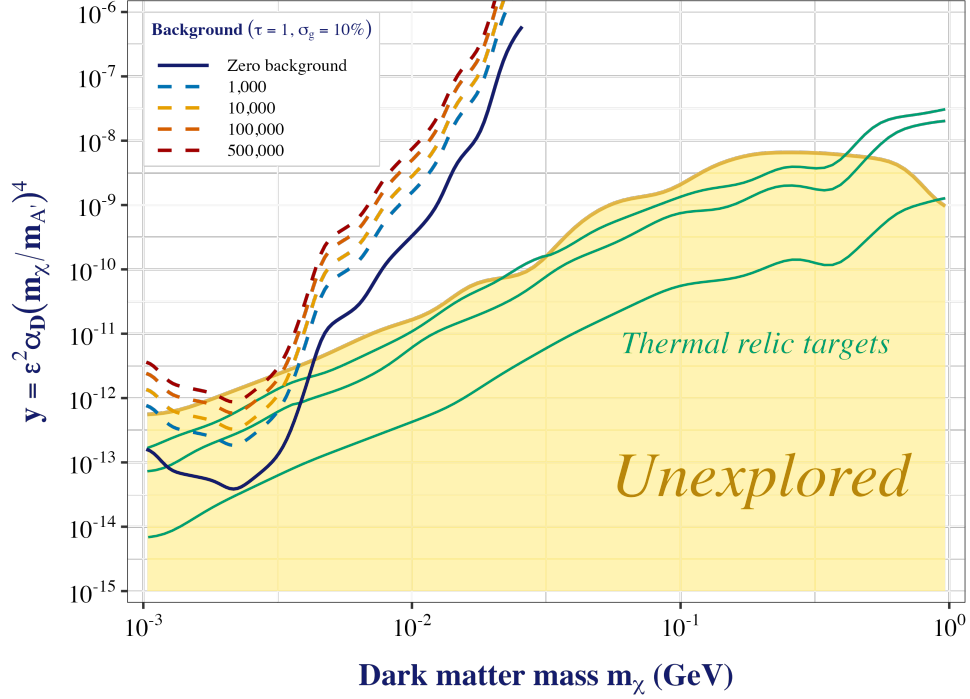


Figure A.5: DarkMESA Phase B exclusion reach in the y - m_χ parameter space computed using the likelihood-based unified model, assuming $\tau = 1$ and $\sigma_g = 10\%$ energy scale uncertainty. The solid curve shows the zero-background projection; dashed curves show degraded sensitivity for increasing background levels. The green curves indicate the thermal relic targets; the shaded region below the yellow curve remains unexplored.

The qualitative behavior matches the Gaussian analysis in Section 2.7: the $N_{\text{bg}}^{1/4}$ statistical scaling dominates, shifting exclusion curves progressively upward as background increases. With $N_{\text{bg}} = 10,000$ events, the exclusion curve misses substantial portions of the thermal relic targets; at $N_{\text{bg}} = 100,000$, sensitivity to unexplored parameter space is largely lost.

The key conclusion from the main text is validated: to probe the thermal relic targets, DarkMESA must keep residual backgrounds at most at the $\mathcal{O}(10^3)$ event level. The likelihood framework confirms this requirement while providing the rigorous statistical machinery—exact Poisson statistics, CLs protection against spurious exclusions, and proper nuisance parameter treatment—for eventual data analysis.

A.8 Detector Resolution and Response Matrix

The likelihood framework developed above implicitly assumed that simulated signal events and measured background data could be compared on a common energy scale. In practice, this comparison requires careful treatment of detector resolution effects, as discussed in Section 3.6.

The Scale Matching Problem

Measured background events from beam-off data are recorded on the reconstructed energy scale—the energy inferred from detected photoelectrons after all detector response effects. Simulated dark matter signal events, however, are initially generated with known true energy deposits. The energy resolution $\sigma/E = 106.4\%/\sqrt{E/\text{MeV}}$ characterized in Chapter 3 causes significant spectral broadening: at the 10 MeV threshold, resolution $\sigma \approx 3.4 \text{ MeV}$ spreads events across multiple energy bins. Meaningful comparison requires transforming signal predictions onto the same reconstructed scale as the measurements.

Response Matrix Definition

The response matrix R_{ij} represents the probability that an event with true energy in bin j produces reconstructed energy in bin i :

$$R_{ij} = P(E_{\text{reco}} \in \text{bin } i | E_{\text{true}} \in \text{bin } j) \quad (\text{A.37})$$

The first index labels reconstructed energy (matrix rows), the second labels true energy (matrix columns). The matrix incorporates all resolution effects characterized in Chapter 3: Cherenkov photon production fluctuations, photoelectron generation statistics, collection efficiency variations, and dynode cascade multiplication noise. Probability conservation requires column normalization: $\sum_i R_{ij} = 1$ for each true energy bin j .

Events with true energies near the analysis threshold may fluctuate above or below the selection cut—an effect called threshold migration that the response matrix captures automatically.

Construction Methodology

Figure A.6 illustrates the four-step construction procedure from Monte Carlo simulations.

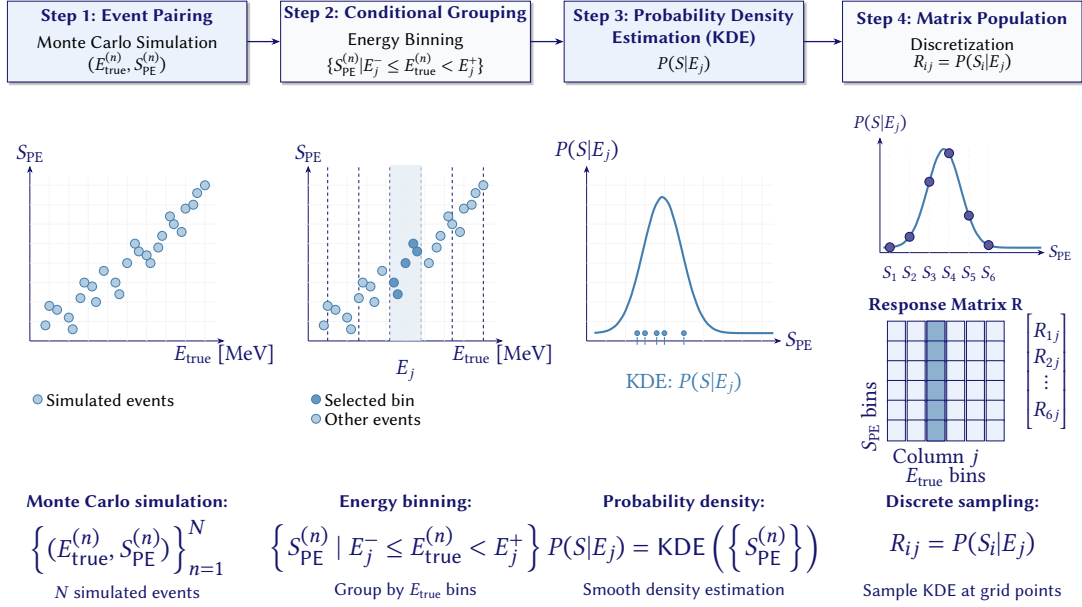


Figure A.6: Response matrix construction from Monte Carlo simulations. Step 1: Generate N signal events with true energies $E_{\text{true}}^{(n)}$ and simulate detector response to obtain measured signals $S_{\text{PE}}^{(n)}$. Step 2: Group events into true energy bins, creating conditional datasets. Step 3: Estimate probability densities $P(S|E_j)$ using kernel density estimation. Step 4: Discretize onto reconstructed energy bins to obtain matrix elements R_{ij} .

Step 1: Event pairing. Monte Carlo simulation generates N signal events, each characterized by a pair $(E_{\text{true}}^{(n)}, S_{\text{PE}}^{(n)})$ of true deposited energy and measured photoelectron signal. The index $n = 1, \dots, N$ labels individual events. The detector response simulation follows the complete chain: Cherenkov photon production, optical transport, photoelectron generation, and dynode cascade multiplication (Section 3.2).

Step 2: Conditional grouping. Events are grouped into true energy bins based on $E_{\text{true}}^{(n)}$. For bin j with boundaries $[E_j^-, E_j^+)$, the conditional dataset contains all measured signals from events in that energy range:

$$D_j = \{S_{\text{PE}}^{(n)} | E_j^- \leq E_{\text{true}}^{(n)} < E_j^+\} \quad (\text{A.38})$$

Step 3: Probability density estimation. For each true energy bin j , kernel density estimation (KDE) provides a smooth, non-parametric estimate of $P(S|E_j)$ from the conditional dataset:

$$P(S|E_j) = \frac{1}{N_j h} \sum_{n \in D_j} K\left(\frac{S - S_{\text{PE}}^{(n)}}{h}\right) \quad (\text{A.39})$$

where $N_j = |D_j|$ is the event count in bin j , K is a Gaussian kernel, and h is the bandwidth. KDE captures the asymmetric response at low energies without assuming a functional form.

Step 4: Discretization. Integrating the probability density over reconstructed energy bins populates the matrix:

$$R_{ij} = \int_{S_i^-}^{S_i^+} P(S|E_j) dS \quad (\text{A.40})$$

Implementation Approaches

Two approaches achieve the required scale consistency between signal predictions and background measurements.

Full Monte Carlo simulation. Signal events are generated with true energies, propagated through the complete detector response chain, and histogrammed in reconstructed energy. This approach implicitly convolves the true spectrum with detector resolution and automatically captures threshold migration effects. The resulting reconstructed spectrum enters the likelihood directly through C and $\{\alpha_i\}$ as developed in Section A.3.

Explicit response matrix convolution. Alternatively, theoretical spectra can be transformed without repeated detector simulation. Starting from the true energy spectrum $\{\alpha_j^{\text{true}}\}$ predicted by dark matter models, the response matrix maps to reconstructed space:

$$\alpha_i^{\text{reco}} = \sum_j R_{ij} \cdot \alpha_j^{\text{true}} \quad (\text{A.41})$$

The expected beam-on counts entering the binned likelihood (Equation A.16) become:

$$\lambda_{\text{on},i} = \mu_s \cdot \sum_j R_{ij}(\theta) \cdot \alpha_j^{\text{true}} + \mu_{b,i} \quad (\text{A.42})$$

where only reconstructed bins i above the analysis threshold enter the likelihood, and the response matrix elements $R_{ij}(\theta)$ may depend on nuisance parameters such as energy calibration.

The acceptance C from Section A.3 no longer appears as a separate factor—it is absorbed into the response matrix formulation. Events with true energy in bin j have probability $\sum_{i>T} R_{ij}$ of reconstructing above threshold T ; the remaining probability represents events lost to downward fluctuations. By including only reconstructed bins above threshold, the likelihood automatically excludes these lost events without requiring an explicit acceptance factor.

This formulation makes explicit how detector resolution propagates into the likelihood: spectral smearing redistributes events across reconstructed bins, while threshold migration removes events that fluctuate below the selection cut. Once constructed from a comprehensive simulation campaign, the response matrix enables rapid evaluation of different signal hypotheses—computationally efficient for parameter space scans where regenerating full detector simulations for each mass point would be prohibitive.

B. Dynode Cascade Parameter Determination

The dynode cascade simulation in Section 3.2 requires two key inputs: the gain distribution across the ten-stage multiplication chain and the multiplicative noise parameter σ characterizing gain fluctuations. This appendix describes how laser measurements determined these parameters.

B.1 Experimental Setup

A monochromatic laser system with variable optical attenuation [30] provided the characterization measurements. A laser source directed through optically dense filter plates onto the PMT in a light-tight enclosure enabled systematic variation of incident photon flux over several orders of magnitude.

Initial gain calibration used extremely dense optical filtering (OD 5.6–6.6) to achieve near-single-photon intensities. Under these conditions, the measured charge spectrum exhibits distinct peaks corresponding to different photoelectron numbers. Fitting these discrete features yields the total PMT gain: approximately 3.0×10^6 .

Measurements across optical densities from OD 3.0 (106.1 PE) to OD 4.0 (12.2 PE) provided experimental charge distributions for simulation validation.

B.2 Dynode Gain Distribution

The gain distribution follows the theoretical framework of [34]: the first dynode provides the highest gain to minimize noise contribution from subsequent stages. For cascaded multiplication, each stage’s relative noise contribution decreases with the gain of preceding stages.

Table B.1 shows the gain distribution for the standard PMT configuration. Percentages represent each dynode’s logarithmic contribution to total gain, ensuring the product equals 3.0×10^6 .

Table B.1: Dynode gain distribution for standard PMT configuration. The first dynode receives highest gain to optimize noise performance.

Stage	Gain (%)	Stage	Gain (%)
1	15.0	6	10.0
2	13.0	7	9.0
3	12.0	8	7.0
4	11.0	9	7.0
5	10.0	10	6.0

B.3 Multiplicative Noise Validation

The cascade simulation incorporates multiplicative noise through independent gain variations at each stage:

$$G_i = G_{\text{nominal}} \times M, \quad \ln M \sim \mathcal{N}(0, \sigma^2) \quad (\text{B.1})$$

Figure B.1 presents experimental charge distributions at four optical density settings, fitted with skew-normal distributions to characterize asymmetric PMT response. Fit quality exceeds $R^2 = 0.997$ for all conditions.

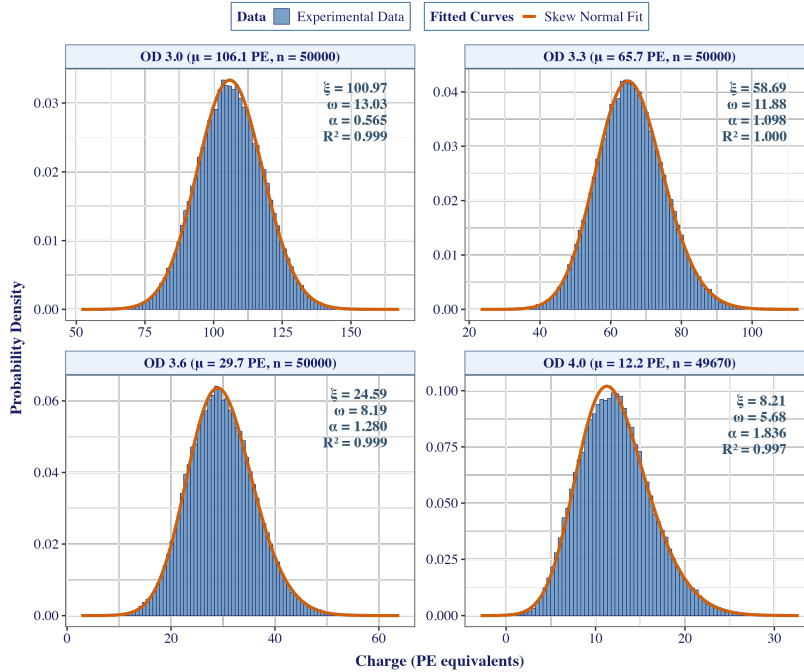


Figure B.1: Experimental charge distributions at four optical density settings with skew-normal fits providing the baseline for cascade simulation validation.

Figure B.2 compares simulated distributions against experimental measurements. Without gain variation (blue curves), simulated distributions appear too narrow and symmetric. Adding multiplicative noise with $\sigma = 0.15$ (orange curves) reproduces the experimental width and asymmetry (green curves) across all intensity ranges. Iterative adjustment of σ until visual agreement was achieved determined this parameter value—a qualitative approach that nonetheless provides confidence in the cascade model’s predictive capability.

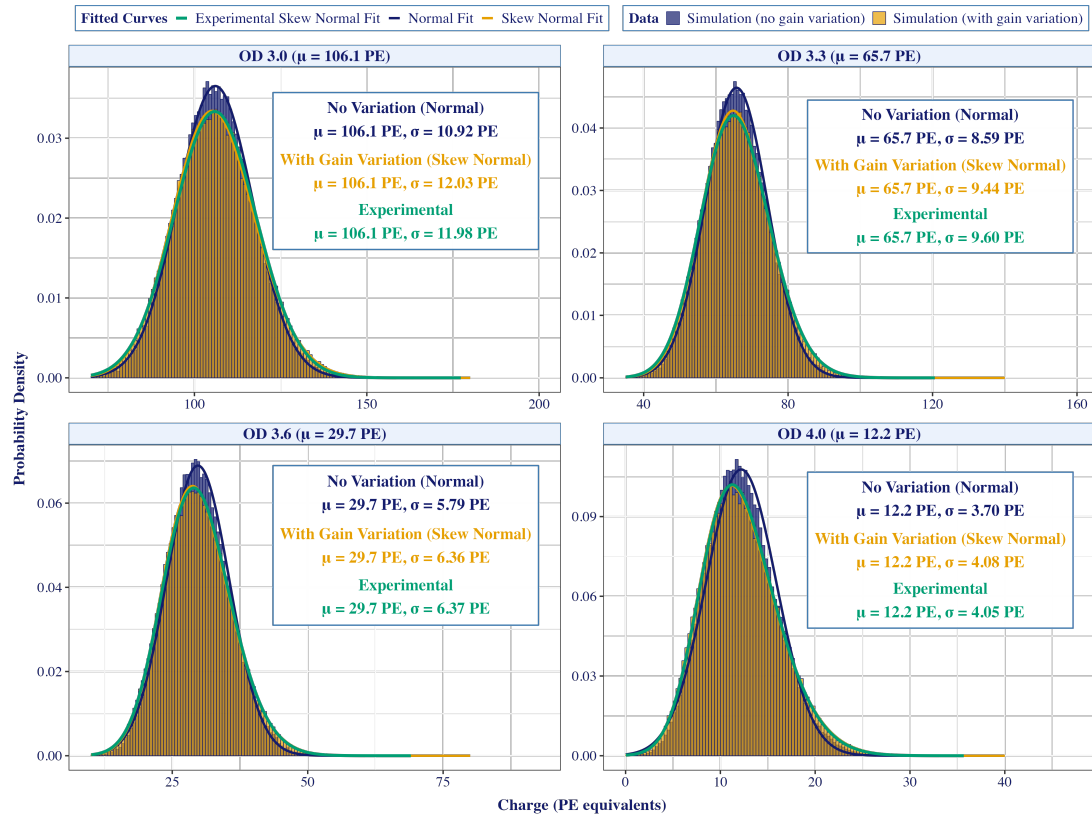


Figure B.2: Validation of dynode cascade simulation. Simulated distributions without gain variation (blue) and with $\sigma = 0.15$ (orange) compared against experimental skew-normal fits (green). Each simulation draws photoelectron counts from Poisson distributions with means matching experimental fitted values.

C. Energy Resolution Model

The energy resolution analysis in Section 3.6 employs skew-normal maximum likelihood estimation with pooled cascade sampling. This appendix provides the mathematical foundation.

C.1 Model Framework

For true energy E_{true} , the measured energy follows a skew-normal distribution:

$$E_{\text{measured}} \sim \text{SkewNormal}(\xi, \omega, \alpha) \quad (\text{C.1})$$

where parameters (ξ, ω, α) depend on true energy. To obtain robust resolution estimates, each simulated event undergoes five independent PMT cascade realizations, creating a pooled dataset that averages cascade fluctuations while preserving physical resolution characteristics.

Two relationships constrain the model. The relative energy resolution follows:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E_{\text{true}}}} \quad (\text{C.2})$$

where a is the fitted noise parameter. The skewness parameter decays exponentially with energy:

$$\alpha(E_{\text{true}}) = \alpha_0 \exp\left(-\frac{E_{\text{true}}}{E_{\text{decay}}}\right) \quad (\text{C.3})$$

reflecting the transition from highly skewed distributions at low energies to nearly Gaussian behavior at high energies.

C.2 Parameter Constraints

Satisfying physical constraints requires careful transformation of the skew-normal distribution parameters. For $X \sim \text{SkewNormal}(\xi, \omega, \alpha)$:

$$\delta = \frac{\alpha}{\sqrt{1 + \alpha^2}} \quad (\text{C.4})$$

$$\text{Mean}[X] = \xi + \omega \delta \sqrt{\frac{2}{\pi}} \quad (\text{C.5})$$

$$\text{Variance}[X] = \omega^2 \left(1 - \frac{2\delta^2}{\pi} \right) \quad (\text{C.6})$$

Requiring the distribution mean to equal true energy (unbiased measurement) yields the location parameter:

$$\xi = E_{\text{true}} - \omega \delta \sqrt{\frac{2}{\pi}} \quad (\text{C.7})$$

The resolution definition determines the scale parameter:

$$\omega = \frac{a \sqrt{E_{\text{true}}}}{\sqrt{1 - \frac{2\delta^2}{\pi}}} \quad (\text{C.8})$$

C.3 Fitted Results

Maximum likelihood estimation on the pooled dataset yields $a = 106.4\%$, $\alpha_0 = 7.8$, and $E_{\text{decay}} = 11.3 \text{ MeV}$. Figure C.1 validates the model through diagnostic plots.

The quantile-quantile plot confirms agreement with skew-normal assumptions across all energies. The energy-dependent skewness decays from strong asymmetry at low energies ($\alpha_0 = 7.8$) to nearly Gaussian distributions beyond 60 MeV, validating the exponential parameterization.

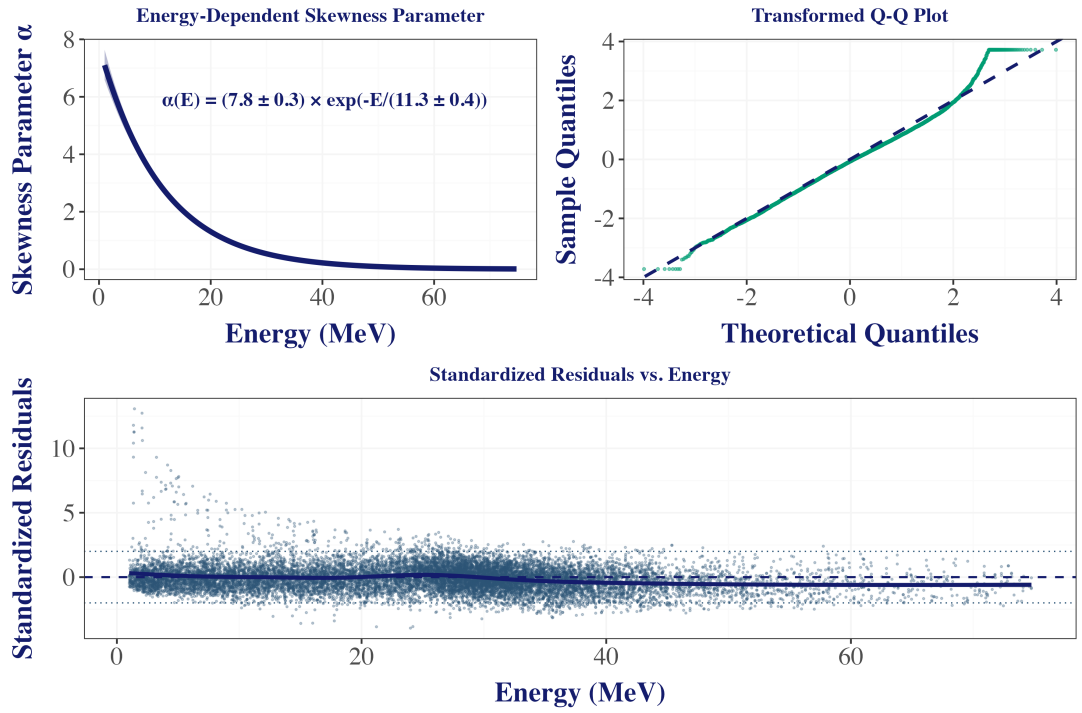


Figure C.1: Diagnostic plots validating the energy resolution model. Top left: Energy-dependent skewness $\alpha(E) = 7.8 \times \exp(-E/11.3)$ showing exponential decay. Top right: Quantile-quantile plot confirming skew-normal assumptions. Bottom: Standardized residuals showing no systematic trends.

D. SiPM Response Modeling

The SiPM detection efficiency analysis in Section 4.4 extends GEANT4 simulations to model silicon photomultiplier response. This appendix documents the photon detection efficiency implementation and charge resolution modeling.

D.1 Photon Detection Efficiency

SiPM detection probability depends on a cascade of processes from photon absorption through avalanche triggering. The simulation implements:

$$\text{PDE}(\lambda, V, T) = \eta(\lambda, T) \cdot \varepsilon(V, T) \cdot F \quad (\text{D.1})$$

The quantum efficiency $\eta(\lambda, T)$ represents the probability of electron-hole pair generation; for silicon, peak efficiency occurs at 400 nm to 550 nm, well-matched to PbF₂ Cherenkov emission. The Geiger discharge efficiency $\varepsilon(V, T)$ represents the avalanche triggering probability, depending critically on overvoltage $V_{\text{over}} = V_{\text{bias}} - V_{\text{breakdown}}(T)$. Because breakdown voltage decreases with temperature at approximately 20 mV/°C to 30 mV/°C, fixed bias produces increasing effective overvoltage at lower temperatures, enhancing ε . The geometric fill factor $F \approx 75\%$ for onsemi J-Series devices represents the active area fraction.

The simulation uses manufacturer PDE specifications at 2.5 V overvoltage and 25 °C, achieving peak PDE of approximately 40 % at 420 nm. Room-temperature values introduce a conservative bias: lower operating temperatures enhance both η and ε , so simulation results represent lower bounds on achievable performance.

D.2 Charge Resolution Modeling

PMT charge resolution is straightforward to simulate through dynode cascade modeling, but SiPM charge resolution involves complex semiconductor physics: pixel-to-pixel gain variations from fabrication non-uniformities, avalanche multiplication statistics from stochastic carrier dynamics, optical crosstalk from photons triggering

adjacent pixels, afterpulsing from trapped carriers, and electronic noise. Modeling these effects from first principles requires device-specific parameters unavailable from manufacturer specifications.

A hybrid approach addresses this limitation by combining GEANT4 optical simulation with empirically-determined charge resolution functions. The simulation generates ideal photoelectron distributions from Cherenkov production, optical transport, and PDE-weighted detection. These distributions—reflecting only Poisson statistics—are convolved with experimental resolution functions:

$$\sigma_k = \sqrt{\sigma_0^2 + k \cdot \sigma_1^2} \quad (\text{D.2})$$

where σ_k is the resolution for k photoelectrons, σ_0 characterizes pedestal width (electronic noise), and σ_1 quantifies single-photoelectron resolution (gain variations, avalanche statistics, crosstalk).

Resolution parameters from dark count spectroscopy at 20 °C (Section 4.4) yield $\sigma_0 = 0.26$ PE and $\sigma_1 = 0.33$ PE—the poorest resolution across the characterized temperature range. Lower temperatures improve resolution through reduced dark carrier generation and enhanced avalanche uniformity.

For each simulated event producing N_{ideal} photoelectrons, the detected signal is sampled from:

$$N_{\text{detected}} = \mathcal{N}(N_{\text{ideal}}, \sigma_{N_{\text{ideal}}}) \quad (\text{D.3})$$

with negative values truncated to zero. Truncating negative values introduces minor asymmetry for low yields ($N \lesssim 3$ PE) but negligibly affects the 5 PE to 10 PE signals characteristic of 10 MeV detection. These conservative parameter choices ensure that detection efficiency results represent worst-case scenarios.

List of Figures

1.1	Dark matter candidate landscape	4
2.1	Overview of the MESA accelerator facility	12
2.2	Schematic of the DarkMESA detection chain.	16
2.3	Exploded-view CAD rendering of the Phase A prototype detector assembly.	18
2.4	Phase A calorimeter components.	19
2.5	Veto system readout electronics	21
2.6	CAD rendering of the Phase B detector configuration showing the dual-wall calorimeter architecture.	22
2.7	SAPHIR lead glass modules under consideration for Phase B rear wall	23
2.8	DarkMESA sensitivity projections and existing limits	31
2.9	Impact of background on coupling reach	32
3.1	Simulation framework architecture for the DarkMESA calorimeter	38
3.2	PMT dynode cascade simulation schematic.	40
3.3	Logical flow of Chapter 3	42
3.4	Energy deposition and photon production in PbF ₂ under cosmic ray exposure	44
3.5	Particle interaction outcomes in PbF ₂ crystals	47
3.6	Attenuation characteristics in lead and concrete shielding	53
3.7	Energy transmission fractions through lead and concrete shielding.	56
3.8	Validation of the DarkMESA calorimeter simulation framework.	59
3.9	Energy resolution of PbF ₂ detector with PMT readout	63
3.10	PbF ₂ particle detection efficiency at 1 PE threshold for four particle types	68
3.11	Complete PbF ₂ detection efficiency response as a function of energy and detection threshold for four particle types	69
3.12	PbF ₂ detection efficiency at 10 MeV as a function of photoelectron threshold for single PMT readout	70
4.1	Dark count rate of Photonis XP 2900 and Hamamatsu R1924A.	75
4.2	Dual-sided readout configuration for Phase A implementation.	76
4.3	Detection efficiency comparison between single-PMT and dual-sided readout	77
4.4	Coupled PbF ₂ crystal array configuration.	81
4.5	Detection efficiency at 10 MeV for coupled crystal array configuration	83
4.6	SiPM array readout configuration	85
4.7	Detection efficiency at 10 MeV for five-SiPM array configurations.	87
4.8	SiPM dark count spectroscopic analysis at -10°C	89
4.9	Temperature dependence of SiPM integral dark count rates.	91
4.10	Maximum photoelectron threshold versus number of available SiPMs for fourfold coincidence	93
5.1	Scintillation photon production and energy deposition in EJ-200	101

5.2	Particle interaction outcomes in EJ-200 veto panels.	104
5.3	Energy-dependent efficiency comparison: EJ-200 versus PbF_2	106
5.4	Experimental and simulation geometries for veto performance characterization	110
5.5	Analysis flow diagram for veto performance characterization	111
5.6	Veto detection efficiency versus calorimeter energy threshold	113
5.7	Energy spectra before and after veto classification	115
5.8	Alluvial diagram of secondary particle creation in leak-through events	116
5.9	Energy distributions of particles in leak-through events	117
5.10	Firing frequency by panel position.	119
5.11	Distribution of triggered panel multiplicities.	119
A.1	Illustration of the CLs method	142
A.2	Monte Carlo implementation workflow for CLs upper limit determination	144
A.3	Signal and background distributions for toy study	148
A.4	Validation of sensitivity scaling.	150
A.5	DarkMESA sensitivity from unified model	153
A.6	Response matrix construction methodology	155
B.1	Experimental charge distributions	160
B.2	Cascade simulation validation.	161
C.1	Energy resolution diagnostics.	165

List of Tables

2.1	MESA operational parameters	13
2.2	Sensitivity scaling versus background level.	33
3.1	Energy resolution validation across three orders of magnitude.	64
4.1	Accidental coincidence rates per crystal pair for dual-sided PMT readout	78
4.2	Maximum PE threshold for coincidence requirements in five-SiPM array	88
4.3	Accidental coincidence rates for Phase B implementation using SiPM arrays	92
B.1	Dynode gain distribution.	160

Bibliography

- [1] V. C. Rubin and W. K. Ford, *Astrophys. J.* **159** (1970), 379, DOI: [10.1086/150317](#).
- [2] V. C. Rubin, W. K. Ford, and N. Thonnard, *Astrophys. J.* **238** (1980), 471, DOI: [10.1086/158003](#).
- [3] D. Clowe *et al.*, *Astrophys. J. Lett.* **648** (2006), L109, DOI: [10.1086/508162](#).
- [4] Planck Collaboration, *Astron. Astrophys.* **641** (2020), A6, DOI: [10.1051/0004-6361/201833910](#).
- [5] B. D. Fields and S. Sarkar, *J. Phys. G* **33** (2006), 1, DOI: [10.1088/0954-3899/33/1/001](#).
- [6] G. Bertone and D. Hooper, *Rev. Mod. Phys.* **90** (2018), 045002, DOI: [10.1103/RevModPhys.90.045002](#).
- [7] G. Jungman, M. Kamionkowski, and K. Griest, *Phys. Rep.* **267** (1996), 195, DOI: [10.1016/0370-1573\(95\)00058-5](#).
- [8] E. Aprile *et al.*, *Phys. Rev. Lett.* **121** (2018), 111302, DOI: [10.1103/PhysRevLett.121.111302](#).
- [9] D. S. Akerib *et al.*, *Phys. Rev. Lett.* **118** (2017), 021303, DOI: [10.1103/PhysRevLett.118.021303](#).
- [10] E. Izaguirre *et al.*, *Phys. Rev. D* **88** (2013), 114015, DOI: [10.1103/PhysRevD.88.114015](#).
- [11] B. Holdom, *Phys. Lett. B* **166** (1986), 196, DOI: [10.1016/0370-2693\(86\)91377-8](#).
- [12] K. R. Dienes, C. F. Kolda, and J. March-Russell, *Nucl. Phys. B* **492** (1997), 104, DOI: [10.1016/S0550-3213\(97\)00173-9](#).
- [13] S. A. Abel *et al.*, *Phys. Lett. B* **666** (2008), 66, DOI: [10.1016/j.physletb.2008.03.076](#).
- [14] M. Pospelov, *Phys. Rev. D* **80** (2009), 095002, DOI: [10.1103/PhysRevD.80.095002](#).
- [15] D. J. E. Marsh, *Phys. Rep.* **643** (2016), 1, DOI: [10.1016/j.physrep.2016.06.005](#).
- [16] R. Adhikari *et al.*, *J. Cosmol. Astropart. Phys.* **2017** (2017), 025, DOI: [10.1088/1475-7516/2017/01/025](#).

- [17] J. D. Bjorken *et al.*, *Phys. Rev. D* **80** (2009), 075018, DOI: [10.1103/PhysRevD.80.075018](https://doi.org/10.1103/PhysRevD.80.075018).
- [18] B. Batell, R. Essig, and P. Schuster, *Phys. Rev. Lett.* **113** (2014), 171802, DOI: [10.1103/PhysRevLett.113.171802](https://doi.org/10.1103/PhysRevLett.113.171802).
- [19] BDX Collaboration, *Phys. Rev. D* **106** (2022), 072011, DOI: [10.1103/PhysRevD.106.072011](https://doi.org/10.1103/PhysRevD.106.072011).
- [20] D. Banerjee *et al.*, *Phys. Rev. Lett.* **120** (2018), 231802, DOI: [10.1103/PhysRevLett.120.231802](https://doi.org/10.1103/PhysRevLett.120.231802).
- [21] A. A. Aguilar-Arevalo *et al.*, *Phys. Rev. Lett.* **118** (2017), 221803, DOI: [10.1103/PhysRevLett.118.221803](https://doi.org/10.1103/PhysRevLett.118.221803).
- [22] J. P. Lees *et al.*, *Phys. Rev. Lett.* **113** (2014), 201801, DOI: [10.1103/PhysRevLett.113.201801](https://doi.org/10.1103/PhysRevLett.113.201801).
- [23] B. S. Schlimme *et al.*, *EPJ Web Conf.* **303** (2024), 06002, DOI: [10.1051/epjconf/202430306002](https://doi.org/10.1051/epjconf/202430306002).
- [24] T. Åkesson *et al.* (2018), arXiv: [1808.05219](https://arxiv.org/abs/1808.05219).
- [25] M. Christmann, “Design Studies for the Beam-dump Experiment DarkMESA”, PhD thesis, Mainz, Germany: Johannes Gutenberg-Universität Mainz, 2022, DOI: [10.25358/openscience-9076](https://doi.org/10.25358/openscience-9076).
- [26] J. Teichert *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **557** (2006), 239, DOI: [10.1016/j.nima.2005.10.077](https://doi.org/10.1016/j.nima.2005.10.077).
- [27] B. S. Schlimme *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **1013** (2021), 165668, DOI: [10.1016/j.nima.2021.165668](https://doi.org/10.1016/j.nima.2021.165668).
- [28] D. Becker *et al.*, *Eur. Phys. J. A* **54** (2018), 208, DOI: [10.1140/epja/i2018-12611-6](https://doi.org/10.1140/epja/i2018-12611-6).
- [29] J. Beacham *et al.*, *J. Phys. G* **47** (2019), 010501, DOI: [10.1088/1361-6471/ab4cd2](https://doi.org/10.1088/1361-6471/ab4cd2).
- [30] J. Pättschke, “Development of a prototype calorimeter for the DarkMESA Experiment”, Bachelor’s thesis, Mainz, Germany: Johannes Gutenberg-Universität Mainz, 2023.
- [31] M. Kontogoulas, “Calibration of the DarkMESA Prototype Calorimeter”, Bachelor’s thesis, Mainz, Germany: Johannes Gutenberg-Universität Mainz, 2024.
- [32] P. Burger, “Antwortverhalten von kalorimetrischen Tscherenkow-Strahlern auf Elektronen und Neutronen im MeV-Energiebereich”, Bachelor’s thesis, Mainz, Germany: Johannes Gutenberg-Universität Mainz, 2019.
- [33] C. Hagmann *et al.*, *Cosmic-ray Shower Library (CRY)*, Technical Report UCRL-TM-229453, Livermore, CA, USA: Lawrence Livermore National Laboratory, 2012.

- [34] A. G. Wright, *The Photomultiplier Handbook*, Oxford: Oxford University Press, 2017, DOI: [10.1093/oso/9780199565092.001.0001](https://doi.org/10.1093/oso/9780199565092.001.0001).
- [35] M. J. Berger *et al.*, *XCOM: Photon Cross Sections Database, NIST Standard Reference Database 8 (XGAM)*, National Institute of Standards and Technology, PML, Radiation Physics Division, 2010, DOI: [10.18434/T48G6X](https://doi.org/10.18434/T48G6X).
- [36] D. Brown *et al.*, *Nuclear Data Sheets* **148** (2018), Special Issue on Nuclear Reaction Data, 1, DOI: <https://doi.org/10.1016/j.nds.2018.02.001>.
- [37] A. T. Fienberg *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **783** (2015), 12, DOI: [10.1016/j.nima.2015.02.028](https://doi.org/10.1016/j.nima.2015.02.028).
- [38] L. I. Schiff, *Phys. Rev.* **50** (1936), 88, DOI: [10.1103/PhysRev.50.88](https://doi.org/10.1103/PhysRev.50.88).
- [39] V. Chmill *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **854** (2017), 70, DOI: [10.1016/j.nima.2017.02.049](https://doi.org/10.1016/j.nima.2017.02.049).
- [40] R. Pagano *et al.*, *IEEE Trans. Electron Devices* **59** (2012), 2410, DOI: [10.1109/TED.2012.2205689](https://doi.org/10.1109/TED.2012.2205689).
- [41] A. N. Otte *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **846** (2017), 106, DOI: [10.1016/j.nima.2016.09.053](https://doi.org/10.1016/j.nima.2016.09.053).
- [42] G. F. Knoll, *Radiation Detection and Measurement*, 4th, Hoboken, NJ: John Wiley & Sons, 2010.
- [43] S. Tavernier, *Experimental Techniques in Nuclear and Particle Physics*, Berlin, Heidelberg: Springer-Verlag, 2010, DOI: [10.1007/978-3-642-00829-0](https://doi.org/10.1007/978-3-642-00829-0).
- [44] C. Grupen and B. Shwartz, *Particle Detectors*, 2nd, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge: Cambridge University Press, 2008, DOI: [10.1017/9781009401531](https://doi.org/10.1017/9781009401531).

